

Université de Montréal

**Contrôle adaptatif et autoréglage :
applications de l'approximation stochastique**

par
Irina Baltcheva

Département d' Informatique et Recherche Opérationnelle
Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention du grade de Maître ès sciences (M.Sc.)
en Informatique et Recherche Opérationnelle

Août, 2004

© Irina Baltcheva, 2004



QA

76

U54

2005

v. 013

Direction des bibliothèques

AVIS

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

NOTICE

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal
Faculté des études supérieures

Ce mémoire intitulé:

**Contrôle adaptatif et autoréglage :
applications de l'approximation stochastique**

présenté par:

Irina Baltcheva

a été évalué par un jury composé des personnes suivantes:

Michel Gendreau,	président-rapporteur
Felisa J. Vázquez-Abad,	directeur de recherche
Pierre L'Écuyer,	membre du jury

Mémoire accepté le: 14 décembre 2004

RÉSUMÉ

Trois applications de l'algorithme d'approximation stochastique sont présentées dans le cadre de procédures adaptatives de contrôle et d'autoréglage. La première traite le problème du contrôle prédictif en ligne d'un réacteur chimique. Le but du contrôle prédictif est de calculer la quantité de produit à fournir au réacteur de sorte qu'un niveau cible de produits résultants soit atteint à la sortie. À chaque étape de calcul, un problème d'optimisation de grande dimension doit être résolu. Utilisant des différences finies pour l'estimation des gradients, les approches traditionnelles d'optimisation non linéaire s'avèrent inefficaces pour ce modèle. Le temps de calcul nécessaire à l'estimation des gradients étant trop grand, le contrôle prédictif du réacteur ne peut être effectué en ligne. Nous proposons d'utiliser un algorithme d'approximation stochastique de type Kiefer-Wolfowitz qui utilise des gradients directionnels. Cette méthode de directions aléatoires est connue pour son efficacité sur des problèmes à grande dimension et saura améliorer la qualité du contrôle prédictif.

La seconde application concerne l'estimation de la probabilité de blocage d'un réseau CDMA mobile. Aucune expression analytique de cette probabilité n'étant connue de nos jours, la simulation est fréquemment employée afin d'approximer cette mesure de performance. Or, le blocage est un événement rare et sa simulation par des méthodes de type Monte-Carlo peut être très lente. La vitesse d'une simulation étant déterminée par la précision désirée, elle peut être augmentée en réduisant la variance des estimateurs. Pour cette raison, nous faisons appel à l'échantillonnage stratégique, technique de réduction de la variance impliquant un changement de mesure. L'approximation stochastique sera employée pour trouver les paramètres optimaux du changement de mesure, sans qu'il y ait besoin de simulations pilotes.

Enfin, nous abordons le problème du contrôle de la congestion dans les réseaux de télécommunication sans fil. Il s'agit de l'auto-ajustement d'un algorithme de renforcement du Protocole de Contrôle de Transmission (TCP). Cet algorithme nommé CLAMP a comme objectif la stabilisation de TCP lorsque ce dernier opère sur des liens sans fils. Pour que CLAMP puisse fonctionner efficacement, la valeur optimale d'un de ses

paramètres doit être trouvée. Nous utilisons un modèle fluide stochastique pour calculer les dérivées stochastiques IPA de la fonction objective qui seront ensuite implantées dans l'approximation stochastique.

MOTS-CLÉS : Approximation stochastique, algorithmes adaptatifs, autoréglage, contrôle en ligne, gradients stochastiques (IPA), échantillonnage stratégique, contrôle prédictif, contrôle de congestion, TCP, modèle fluide stochastique (SFM).

ABSTRACT

Three applications of the stochastic approximation algorithm are presented in the framework of auto-adaptive and self-tuning procedures. The first one addresses the model based predictive control (MBPC) of a chemical reactor. The goal of MBPC is to calculate the values of the input variables in order to achieve a reference level of the output product. At each sampling time, an optimization problem must be solved. The non-linear optimization routines available in the software packages frequently use finite differences to estimate the gradients. These routines proved to be inefficient for the Van der Vusse problem because of the great computational effort required to estimate the gradients. Thus, the control could not be applied on-line. We propose to apply a Kiefer-Wolfowitz-type stochastic approximation method, that uses directional derivatives. This random directions algorithm is known for its efficiency in high dimensional problems and will yield better overall performance of the predictive control.

The second application deals with the estimation of the outage probability of a CDMA mobile network. An analytical expression of this probability not being known today, simulation is frequently used to provide good estimates of this performance measure. However, outage is a rare event and its simulation via Monte-Carlo methods can become very slow. The speed of a simulation being determined by the desired precision, it could be increased by reducing the variance of the estimators. For this reason, we use the Importance Sampling technique, which implies a change of measure, to reduce the variance of outage estimators. The stochastic approximation method will be used to find the optimal values of the parameters of the change of measure, without need of pilot simulations.

Finally, we consider the problem of congestion control in wireless telecommunication networks. We work on the self-tuning of a TCP (Transfer Control Protocol) enhancement algorithm called CLAMP. CLAMP is designed to overcome problems related to wireless links, such as the sawtooth behavior of TCP when dealing with random channel rates. For CLAMP to be efficient, it needs to be tuned, i.e., we need to find the optimal value of one of its parameters. We use a stochastic fluid model to derive IPA estimators

of the objective function which will be optimized using the stochastic approximation algorithm.

KEYWORDS : Stochastic Approximation, Adaptive Algorithms, On-line Control, Self-tuning, Stochastic Gradients (IPA), Importance Sampling, Model-Based Predictive Control, Congestion control, TCP, Stochastic Fluid Model (SFM).

TABLE DES MATIÈRES

RÉSUMÉ	iv
ABSTRACT	vi
TABLE DES MATIÈRES	viii
LISTE DES FIGURES	x
LISTE DES NOTATIONS ET DES SYMBOLES	xii
DÉDICACE	xiii
REMERCIEMENTS	xiv
INTRODUCTION	1
CHAPITRE 1 : APPROXIMATION STOCHASTIQUE - CONVERGENCE	5
1.1 Historique	7
1.2 Convergence avec probabilité 1 - la méthode de l'ODE	8
1.3 Convergence avec probabilité 1 - autres méthodes	14
1.4 Contexte d'application	14
CHAPITRE 2 : APPLICATION AU CONTRÔLE PRÉDICTIF D'UN PROCÉDÉ CHIMIQUE	16
2.1 Description du modèle	18
2.2 Contrôle prédictif non linéaire	19
2.3 Simultaneous Perturbation Stochastic Approximation (SPSA)	23
2.4 Résultats	27

CHAPITRE 3 :	APPLICATION À L'ESTIMATION DE LA PROBABILITÉ	
	DE BLOCAGE DANS UN RÉSEAU CDMA	31
3.1	Modèle et formulation du problème	32
3.2	Simulation Monte-Carlo	35
3.3	Le changement de mesure	37
3.4	Mécanisme d'autoréglage	42
3.4.1	Calcul des gradients	44
3.4.2	Changement de mesure stratifié	49
3.5	Résultats	50
CHAPITRE 4 :	APPLICATION AU CONTRÔLE DE FLOT DANS LES RÉSEAUX	
	SANS FILS	52
4.1	Contrôle de flot : TCP	53
4.2	Amélioration de TCP : CLAMP	55
4.2.1	Description de CLAMP	56
4.2.2	Modèle Fluide Stochastique d'un système à événements discrets	58
4.3	Problème d'optimisation	61
4.4	Calcul de la dérivée IPA	63
4.4.1	Analyse d'une trajectoire	64
4.4.2	Dérivées des temps d'événements	75
4.4.3	Dérivées des mesures de performance	78
4.5	Procédure d'autoréglage et résultats	81
CONCLUSION		83
BIBLIOGRAPHIE		85

LISTE DES FIGURES

1.1	Les interpolations en temps continu $\theta^0(t)$ et $\theta^n(t)$ des itérés de l'approximation stochastique	10
2.1	Réaction Van der Vusse	18
2.2	Les trajectoires de l'algorithme de descente avec l'estimateur du gradient à différences finies (FD) et l'estimateur à perturbations aléatoires (SP) sur le problème déterministe (2.11).	25
2.3	Évolution de la variable contrôlée c_B en fonction du temps. Dimension du problème : $p = 2$. (a) SQP : 52 sec. (b) SPSA : 40 sec.	28
2.4	Évolution de la variable contrôlée c_B en fonction du temps. Dimension : $p = 10$. (a) SQP : 198 sec. (b) SPSA : 59 sec.	28
2.5	Évolution de la variable contrôlée c_B en fonction du temps avec du bruit blanc dans les mesures de c_B . Dimension : $p = 2$. (a) SQP : 1139 sec. (b) SPSA : 50 sec.	28
2.6	Évolution du contrôle en fonction du temps. Dimension : $p = 2$. (a) NAG (b) SPSA : contrôle sous-optimal lorsque près de la borne inférieure	30
3.1	Mobile est stations de base dans un réseau CDMA	33
3.2	Courbe de la variance de l'estimateur de la probabilité de blocage $\hat{p}(\theta, \mu)$ (3.10) en fonction des paramètres du changement de mesure μ et θ , obtenue par estimation fonctionnelle avec nombres aléatoires communs.	43
3.3	Valeurs de θ et μ en fonction du temps, résultat de l'approximation stochastique (3.13). La courbe "non-stratified" correspond au modèle sans stratification (la valeur optimale des paramètres pour ce modèle est indiquée par la ligne pointillée). Les courbes "neighborhood" et "outside neighborhood" sont associées aux usagers à l'intérieur et à l'extérieur du voisinage de la cellule centrale 0 respectivement.	51
4.1	Topologie typique d'un système	54

4.2	Modèle fluide avec rétroaction : file d'attente fermée	58
4.3	Fonctions de performance $E[Q(a)]$ et $E[U(a)]$ et fonction objective $J(a) = E[U(a)] - \lambda (E[Q(a)])^2$ en fonction de a (facteur d'échelle $\lambda = 50$). . .	62
4.4	Trajectoire d'une simulation : $a = 0.1, b = 1.0, \phi = 0.01, d = 0.2, k = 1$. Courbes de $q(t), q^*(t), f(t)$ et $c(t)$ en fonction du temps.	66
4.5	Trajectoire d'une simulation : période vide et ses points de rupture.	68
4.6	Points de rupture d'une période occupée exogène	71
4.7	Points de rupture d'une période occupée endogène	74
4.8	Terminaison d'une période occupée	76
4.9	Enclenchement d'une période occupée endogène	78
4.10	Trajectoire de l'algorithme d'approximation stochastique avec la simulation du modèle fluide de CLAMP. Sont données les valeurs du paramètre a en fonction du nombre d'itérations avec différentes valeurs du gain ϵ et $a_0 = 0.0$	82

LISTE DES NOTATIONS ET DES SYMBOLES

Abréviations

p.s.	presque sûrement
a.p.1	avec probabilité 1
v.a.	variable aléatoire
i.i.d.	indépendants et identiquement distribués
AS	Approximation Stochastique
CDMA	“Code Division Multiple Access”
ODE	“Ordinary Differential Equation”

Notations

$\mathcal{N}(\mu, \sigma^2)$	loi normale de moyenne μ et variance σ^2
$U[a, b]$	loi uniforme sur l'intervalle $[a, b]$
\mathfrak{F}	tribu
\mathcal{F}	filtration

Glossaire des traductions

Accusé de réception	“Acknowledgement”
Inclinaison exponentielle	“Exponential tilting”
Échantillonnage stratégique (ou pondéré)	“Importance Sampling”
Horizon fuyant	“Receding horizon”
Queue de distribution épaisse	“Heavy tail”
Réseau fédérateur	“Core network”
Temps de boucle	“Round trip time” (RTT)

À ma mère, mon père et ma soeur

REMERCIEMENTS

J'aimerais tout d'abord remercier ma directrice Felisa J. Vázquez-Abad de m'avoir dirigé pendant les cinq dernières années et pour tout le savoir qu'elle m'a transmis. Grâce à elle, j'ai eu l'occasion de vivre des expériences de recherche très enrichissantes au Canada, comme à l'étranger. Elle est pour moi un exemple de professionnalisme, de motivation, d'énergie et d'ouverture d'esprit.

Je voudrais aussi remercier les personnes qui ont participé aux projets faisant partie de ce mémoire. Merci au professeur César de Prada Moraga pour m'avoir accueillie à deux reprises dans le Département d'Ingénierie des Systèmes et Automatique (ISA) de l'Université de Valladolid, Espagne. Merci aussi à Smaranda Cristea et à Rami Mukhtar pour m'avoir aidé avec les parties "génie" des problèmes.

Je tiens aussi à dire un grand merci à ma famille et en particulier à ma mère, mon père et ma sœur pour m'appuyer dans ce que je fais et pour être toujours là pour moi. Enfin, merci à Cédric pour les messages d'encouragement.

INTRODUCTION

Les méthodes numériques existent depuis plusieurs décennies. Elles sont utilisées pour approximer la solution de problèmes complexes lorsqu'une solution analytique ne peut être trouvée. Souvent, elles impliquent des paramètres dont la valeur doit être réglée à la main. Aussi, arrive-t-il que les méthodes numériques demandent des temps de calcul considérables pour atteindre une précision raisonnable, ce qui les rend inutilisables dans des applications en temps réel.

Ce mémoire propose des idées originales qui rendent les méthodes numériques plus efficaces, auto-ajustables, donc "intelligentes", de sorte que des essais pilotes ou des ajustements manuels ne soient pas nécessaires. Ainsi, trois sujets seront abordés ici. Il s'agit de projets de recherche indépendants, développés par l'auteur durant les cinq dernières années :

- I. Modification des algorithmes fréquemment utilisés pour une efficacité accrue du contrôle en temps réel d'un réacteur chimique ;
- II. Autoréglage d'une procédure d'estimation de la probabilité de blocage dans un réseau de télécommunications mobiles ;
- III. Mécanisme intelligent pour l'autoréglage d'un algorithme de contrôle de flot dans les réseaux sans fils.

Ces trois projets seront regroupés sous le cadre commun que sont les algorithmes d'*approximation stochastique*. En effet, ces techniques d'optimisation sont à la base des procédures adaptatives que nous proposons pour rendre les méthodes numériques plus efficaces et auto-ajustables.

L'approximation stochastique approxime la valeur optimale des paramètres en ayant seulement des mesures bruitées de la fonction objective. En commençant par une valeur devinée, l'équation récursive de l'approximation stochastique ajuste le paramètre en question dans une direction d'amélioration. Si θ est le paramètre qu'on désire optimi-

ser, les itérations de l'algorithme sont :

$$\theta_{n+1} = \theta_n + \epsilon_n f(\theta_n, X_n), \quad (1)$$

où ϵ_n est une suite de petits gains, X_n est un vecteur représentant l'état (aléatoire) du système à l'itération n et $f(\theta, X)$ est une fonction dont le choix définit précisément les différentes procédures. Par exemple, $f(\cdot)$ peut être un estimé du gradient de la fonction objective ou bien une direction d'amélioration aléatoire. Dans le cas d'un gradient estimé, l'approximation stochastique peut être considérée comme une version stochastique des méthodes déterministes de plus forte pente.

Avec le développement récent des techniques d'estimation des gradients de fonctions aléatoires ^[13, 27, 36, 53], l'approximation stochastique basée sur le gradient est considérée aujourd'hui comme un outil important dans l'optimisation de systèmes dynamiques. Nous allons nous servir de cet outil afin de résoudre les problèmes d'ajustement et de contrôle mentionnés précédemment. Voici un aperçu plus détaillé des trois applications faisant partie de ce mémoire.

La première application concerne le contrôle prédictif d'un réacteur chimique. Il s'agit de la réaction Van der Vusse présente dans divers processus industriels. Les variables de contrôle sont des processus en temps réel dont un niveau cible doit être atteint tout en satisfaisant un certain nombre de contraintes opérationnelles. Le modèle canonique discrétise le temps en utilisant un intervalle d'échantillonnage, ce qui ramène le problème de contrôle à un problème d'optimisation non linéaire avec contraintes de très grande dimension. Comme l'évaluation de la fonction objective est complexe et qu'il n'y a pas d'expression analytique pour les gradients, les procédures disponibles dans les librairies commerciales s'avèrent très inefficaces, le problème principal étant le temps de calcul. En effet, la plupart de ces procédures approximent les gradients par des différences finies, ce qui implique un grand nombre d'évaluations de la fonction objective lorsque la dimension du problème est importante. Par conséquent, l'optimisation numérique ne peut être exécutée en temps réel et les contrôleurs n'atteignent pas leur valeurs optimales. Dans le Chapitre 2, nous implémentons une méthodologie d'optimi-

sation globale, qui ajoute du bruit aux observations des gradients et nécessite seulement deux évaluations de la fonction objective par itération, ce qui apporte une meilleure performance que les méthodes déterministes.

La seconde application vise l'estimation de la performance des réseaux de téléphonie mobile CDMA. Dans ces réseaux, les utilisateurs transmettent à travers une station de base et peuvent créer de l'interférence aux autres stations. Lorsque plusieurs mobiles sont déjà connectés au réseau, un nouvel appel peut faire en sorte que le rapport "signal-bruit" passe en dessous d'un seuil de tolérance. Ce phénomène est appelé *blocage* et c'est un critère de performance utile pour la conception et le contrôle des réseaux. Évaluer la probabilité de blocage de façon analytique s'étant avéré sans succès, il n'existe de nos jours que des approximations de celle-ci. La simulation est dans ce cas un bon moyen d'obtenir ces approximations. Or, une simulation directe d'un réseau mobile CDMA est la plupart du temps inefficace, le blocage étant un événement rare. Parfois, des heures de simulation sont nécessaires afin d'obtenir une précision raisonnable. Ainsi, une approche Monte-Carlo n'est pas très utile lorsqu'il s'agit de comparer la performance de différents modèles. L'échantillonnage stratégique est une méthode permettant de simuler le système en utilisant une distribution de référence sous laquelle l'événement d'intérêt est plus probable. Dans le Chapitre 3, nous appliquons la technique de simulation intelligente de [50, 57], où la distribution de référence appartient à une famille paramétrée. Les paramètres de cette distribution sont optimisés à l'aide de l'approximation stochastique afin de réduire la variance des estimateurs résultants. La procédure est donc capable de s'autorégler, sans qu'il y ait besoin d'expériences pilotes.

Le sujet de la dernière application est le contrôle de flot dans des réseaux contenant des liens sans fil. Il s'agit de l'auto-ajustement de CLAMP¹ [3], une procédure visant l'amélioration du protocole TCP ("Transfer Control Protocol"). Lorsqu'il s'agit d'opérer sur des réseaux mobiles où la qualité du canal de transmission est variable, TCP devient très instable, causant ainsi de grandes fluctuations dans l'évolution des files d'attente responsables de la congestion. CLAMP est un algorithme qui ajuste explicitement la fenêtre

¹CLAMP : "Curtailing Large AWNDs to Maximize Performance"

de réception AWND, trouvant ainsi l'équilibre entre les délais de transmission et l'utilisation du lien, en contrôlant le taux d'envoi des données au niveau du récepteur. Sous des conditions stationnaires (avec un taux de transmission constant de c octets/sec.), il a été démontré que CLAMP stabilise le système en un point où la longueur de la file d'attente est égale à $\bar{q} = p_a^{-1}(\phi/c)$, où $\phi > 0$ est une constante et $p_a(\cdot)$ est une fonction dépendant d'un paramètre $a \geq 0$. Lorsque le taux de transmission est aléatoire, ce qui est le cas des canaux sans fils, la longueur moyenne de la file d'attente peut être contrôlée en trouvant la valeur optimale du paramètre a . Jusqu'à présent, l'ajustement a été fait à la main, ce qui a rendu l'algorithme inutilisable dans des situations réelles. L'objectif du Chapitre 4 est l'implantation d'un algorithme adaptatif (en appliquant l'approximation stochastique) qui pourra trouver en ligne la valeur optimale du paramètre et l'ajustera en fonction des conditions variables du canal de transmission. La méthode d'autoréglage résultante opérera à des vitesses très élevées, sans qu'il y ait besoin d'intervention humaine pour l'ajustement des paramètres.

Ainsi, la contribution de ce mémoire est l'application de la méthode d'approximation stochastique aux algorithmes de contrôle et d'autoréglage des paramètres, donnant lieu à des solutions viables aux problèmes de design.

Le texte est divisé comme suit : dans le premier chapitre sera présenté un aperçu du développement des algorithmes d'approximation stochastique depuis leur invention en 1950. L'accent sera mis sur les différentes preuves de convergence et pour illustrer la méthodologie, le théorème de convergence presque sûre sera démontré par la méthode de l'ODE. Les chapitres subséquents seront consacrés respectivement aux trois problèmes décrits ci-haut.

CHAPITRE 1

APPROXIMATION STOCHASTIQUE - CONVERGENCE

Le but de ce chapitre est de donner un aperçu des méthodes de preuve de la convergence des algorithmes d'approximation stochastique. Il existe différents types de convergence et pour chacun d'entre eux, différentes méthodes de preuves peuvent être employées, selon la structure du modèle considéré. Chacune de ces méthodes requiert des hypothèses plus ou moins fortes.

La convergence presque sûre est le résultat le plus fort de tous les types de convergence. Elle indique que si l'algorithme est exécuté une *infinité* de fois, il convergera vers un optimum local a.p.1. Plusieurs conditions au sujet de la suite des gains et de la structure du bruit doivent être satisfaites pour obtenir ce type de convergence. Tout comme pour les méthodes d'optimisation déterministe^[9], le taux de convergence des algorithmes d'approximation stochastique est grandement influencé par le choix des pas. Les conditions suivantes sont fondamentales pour la convergence presque sûre :

$$\sum_{i=1}^{\infty} \epsilon_i = \infty \quad \epsilon_i \geq 0 \quad \epsilon_i \rightarrow 0.$$

Lorsque le système considéré est *stationnaire*, les gains *décroissants* vont garantir la convergence a.p.1 vers un minimum local. Par contre, si l'approximation stochastique est utilisée pour le contrôle en ligne de systèmes qui changent avec le temps, c'est-à-dire dans un environnement *non stationnaire*, l'algorithme doit pouvoir réagir rapidement aux fluctuations. Dans ce cas, un gain décroissant n'est pas adéquat, car après un certain nombre d'itérations, l'algorithme restera coincé dans un optimum local qui n'est peut-être plus l'optimum recherché, à cause de la non-stationnarité du système. Ainsi, pour le contrôle en ligne, un gain *constant* est plus approprié. Celui-ci induit une plus grande variance dans le comportement limite, mais la convergence est plus rapide et l'algorithme plus robuste.

Néanmoins, lorsque le gain est constant, l'algorithme ne converge pas avec probabilité 1. Pour étudier le comportement asymptotique lorsque le gain est constant et très petit, on utilise les méthodes de convergence faible (convergence en distribution) appliquées aux trajectoires de l'algorithme. Ce type de convergence est aussi utile pour des problèmes où le bruit est dépendant de l'état, corrélé, ou pour des algorithmes décentralisés ou asynchrones ^[34].

La convergence presque sûre peut être démontrée de plusieurs façons. Un argument bien connu est de dire que les itérés de l'algorithme suivent asymptotiquement la trajectoire d'une équation différentielle ordinaire (ODE). Cette ODE "moyenne" a un ensemble de points de stabilité qu'on démontre être la solution recherchée au problème d'optimisation. Cette *méthode de l'ODE* a l'avantage de donner une intuition quant à la vitesse de convergence de l'algorithme. Par contre, comme elle étudie le comportement asymptotique d'une suite de processus interpolés décalés $\{\theta^n(\cdot)\}$, elle nous en dit peu sur la période transitoire (les premières itérations, période d'échauffement) de l'algorithme. Il existe aussi d'autres techniques de preuve de la convergence forte, comme par exemple la *méthode de compacité générale* ou les *méthodes de stabilité*. La section 1.2 discute ces différentes approches.

Les méthodes de preuve de la convergence faible utilisent des hypothèses plus faibles et faciles à vérifier en pratique. Dans le cas de la convergence faible, au lieu de considérer directement le processus des itérés, on caractérise les limites de la suite des *mesures* des processus interpolés $\{\theta^n(\cdot)\}$. Tout comme pour les preuves de convergence presque sûre, on applique l'argument de l'ODE, mais cette fois-ci, on parle de convergence vers un *voisinage* de l'optimum. De plus, dans le cas de la convergence faible avec gain constant, on démontre que la trajectoire au complet s'approche de celle de l'ODE lorsque ϵ est proche de zéro. Il est possible après avoir obtenu la convergence faible de faire une "analyse locale" pour obtenir la convergence a.p.1. Tout comme pour cette dernière, il existe pour la convergence faible différentes approches de preuves, selon la structure du modèle. Par exemple, lorsque le bruit a une structure Markovienne (modèle assez fréquent), l'approche de *mesure invariante* ^[33] est très utilisée. Nous n'aborderons pas en détail les preuves de convergence faible, mais tous les résultats à ce sujet se

trouvent dans ^[35].

Avant de présenter dans la section 1.2 le théorème de convergence presque sûre de l'algorithme avec gain décroissant, la section 1.1 expose brièvement le développement depuis les années 1950 de l'analyse de l'approximation stochastique. La section 1.3 résume les autres méthodes de preuves pour la convergence a.p.1 et la section 1.4 décrit le contexte dans lequel sera utilisée l'approximation stochastique dans les chapitres subséquents.

1.1 Historique

Les premières formes d'approximation stochastique apparaissent au début des années 1950 avec les algorithmes de Robbins-Monro (1951) ^[45] et de Kiefer-Wolfowitz (1952) ^[30]. Le problème qui a motivé les travaux de ces chercheurs est celui de trouver les zéros d'une fonction $M(\theta)$ inconnue, dont on ne peut qu'avoir des estimés pour une valeur de θ donnée. Plus précisément, Robbins et Monro se sont intéressés à des expériences chimiques, où $M(\theta)$ représente la valeur moyenne de l'effet qu'a produit une dose θ d'une substance utilisée pour l'expérience. La fonction $M(\theta)$ est supposée monotone et inconnue et on désire trouver la solution $\theta = \theta^*$ de l'équation $M(\theta) = \alpha$, où α est une constante donnée. L'article ^[45] suggère une méthode faisant des expériences successives à des niveaux de dosage $\theta_1, \theta_2, \dots$ de façon à ce que θ_n tende vers θ^* en probabilité. L'année suivante, Kiefer et Wolfowitz publient leur travail traitant le cas où l'on cherche $\theta^* = \arg \max M(\theta)$ et leur algorithme récursif diffère légèrement de celui de Robbins et Monro.

À la suite de la publication de Robbins et Monro ^[45], l'école russe développe durant les années 1950 l'étude des approximations stochastiques. Une excellente référence à ce sujet est le livre de Tsytkin ^[52]. En 1965, Gladyshev ^[26] publie une note dans laquelle il donne une preuve alternative de l'algorithme de Robbins-Monro, où il fait intervenir pour la première fois des martingales, idée qui sera subséquemment reprise par plusieurs auteurs. En 1974, Derevetskii et Fradkov ^[20] proposent "une approche générale pour l'analyse de la dynamique des algorithmes du type (1) qui consiste en le rempla-

gement de l'équation (1) par un *modèle* qui lui est associé." Le modèle auquel ils font allusion est une équation différentielle. C'est l'origine la "méthode de l'ODE", qui sera formellement introduite par Ljung ^[39] dans un cadre de type Markovien et des gains décroissants.

Il faut noter par la suite, l'importante contribution de Kushner et ses collaborateurs ^[32, 33, 34, 35]. En 1978, Kushner et Clark ^[32] font une première grande étude systématique des algorithmes d'approximation stochastique. On note l'apparition de la méthode de preuve de "compacité générale", ainsi que des arguments impliquant des martingales. On trouve dans leur ouvrage des preuves de convergence presque sûre, ainsi que de convergence faible pour des problèmes avec et sans contraintes. L'article ^[33] de Kushner et Shwartz introduit en 1984 une autre méthode de preuve pour la convergence faible des algorithmes avec bruit dépendant de l'état : la méthode des "mesures invariantes".

Plus récemment, plusieurs autres auteurs ont contribué à l'avancement de la théorie de l'approximation stochastique. De nouvelles conditions nécessaires et suffisantes pour la convergence ont été introduites (voir ^[31] et les références qui s'y trouvent). Les algorithmes pour l'optimisation de fonctions de très grande dimension ont connu des développements récents et des méthodes de directions aléatoires telle SPSA ^[49] ont vu le jour ces dernières années. Avec l'expansion de la recherche opérationnelle et l'importance croissante que détiennent les probabilités dans la modélisation des systèmes, des méthodes d'optimisation telles l'approximation stochastique trouvent des domaines d'application de plus en plus vastes. Pour une vue d'ensemble sur les applications de l'approximation stochastique depuis les 20 dernières années, nous référons le lecteur à ^[1, 6, 51].

1.2 Convergence avec probabilité 1 - la méthode de l'ODE

Afin de donner l'idée principale de la preuve convergence et pour ne pas compliquer les notations, nous allons considérer la forme la plus simple de l'algorithme récursif :

$$\theta_{n+1} = \theta_n + \epsilon_n Y_n, \quad (1.1)$$

où θ est le paramètre que nous voulons optimiser, Y_n est une estimation avec bruit de la direction d'actualisation (possiblement l'estimateur du gradient) et ϵ_n est le gain à l'itération n . D'autres formes récursives existent, par exemple la version avec projection sur un ensemble de contraintes, mais nous allons nous en tenir à la forme donnée par (1.1) pour simplifier l'exposé.

L'idée principale de la méthode de l'ODE est de montrer qu'asymptotiquement, le comportement de l'algorithme est dicté par celui d'une ODE "moyenne". On exprime Y_n de l'équation (1.1) comme la somme d'une espérance conditionnelle et d'un terme de bruit qui est une différence de martingales. On montre que sous des hypothèses raisonnables au sujet du bruit, il y a convergence presque sûre vers un point stationnaire ou une trajectoire limite de l'ODE pour des suites de gains décroissants.

Le cadre mathématique dans lequel nous allons travailler suppose que les variables aléatoires $\theta_0, Y_0, Y_1, \dots, Y_n, \dots$ sont définies sur un espace de probabilité $(\Omega, \mathfrak{F}, \mathbb{P})$ et on note la filtration $\mathcal{F}_n = \sigma(\theta_0; Y_1, \dots, Y_n)$ comme étant la σ -algèbre engendrée par les v.a. $\theta_0, Y_0, Y_1, \dots, Y_n$. La suite $\{\theta_n\}$ sera traitée comme un processus stochastique sur l'espace $(\Omega, \mathbb{F} = \{\mathcal{F}_n; n \geq 1\}, \mathbb{P})$.

L'algorithme d'approximation stochastique est un processus à temps discret. Or, la plupart des preuves de convergence utilisent des *processus interpolés en temps continu* de la suite $\{\theta_n\}$, avec des intervalles d'interpolation ϵ_n . On trouve dans ^[8] (Chapitre 2, p.44) une présentation informelle de ces interpolations, démontrant le lien étroit entre les itérations récursives de l'algorithme de base (1.1) et la trajectoire d'une équation différentielle. L'heuristique qui y est présentée introduit de façon très intuitive le fait que la suite $\{\theta_n\}$ "apparaît comme étant à peu près le schéma de discrétisation standard de l'ODE $\dot{\theta} = g(\theta)$ à l'aide du pas ϵ " (le gain étant supposé petit et constant).

Ainsi, on définit les interpolations reliant le temps continu de l'ODE au temps discret de l'algorithme comme suit. Tout d'abord, on pose

$$t_0 = 0 \quad t_n = \sum_{i=0}^{n-1} \epsilon_i.$$

Soit $m(t)$ l'unique valeur de n telle que $t_n \leq t < t_{n+1}$, pour $t \geq 0$ et $m(t) = 0$ pour

$t < 0$. On définit l'interpolation $\theta^0(\cdot)$ sur $(-\infty, \infty)$ comme étant :

$$\begin{aligned}\theta^0(t, \omega) &= \theta_0 \quad \text{pour } t \leq 0, \text{ et} \\ \theta^0(t, \omega) &= \theta_n \quad \text{pour } t_n \leq t \leq t_{n+1}, t \geq 0.\end{aligned}$$

Le processus décalé $\theta^n(\cdot)$ est (voir figure 1.1) :

$$\theta^n(t, \omega) = \theta^0(t_n + t) \text{ pour } t \in (-\infty, \infty).$$

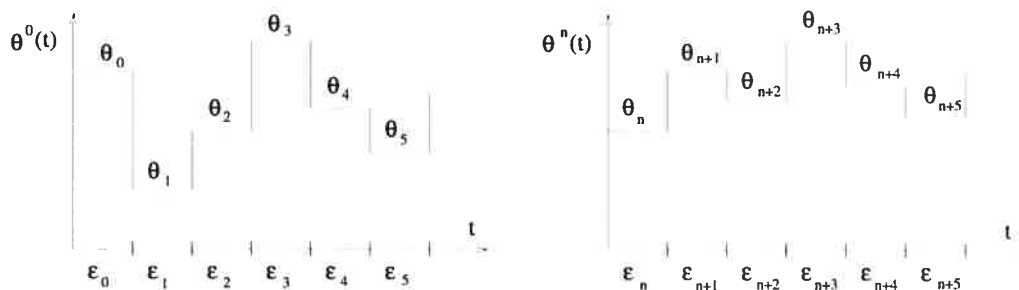


FIG. 1.1 – Les interpolations en temps continu $\theta^0(t)$ et $\theta^n(t)$ des itérés de l'approximation stochastique

Nous allons supposer qu'il existe des fonctions mesurables $g_n(\cdot)$ de θ et des variables aléatoires β_n telles que Y_n peut être décomposé comme suit :

$$Y_n = g_n(\theta_n) + \delta M_n + \beta_n, \quad \text{où} \quad (1.2)$$

$$\delta M_n = Y_n - E[Y_n | \mathcal{F}_{n-1}] \quad \text{et} \quad (1.3)$$

$$E[Y_n | \mathcal{F}_{n-1}] = g_n(\theta_n) + \beta_n. \quad (1.4)$$

Plus loin, nous allons démontrer que la suite $\{\beta_n\}$ est asymptotiquement négligeable. La suite $\{\delta M_n\}$ est une différence de martingales, car le processus défini par $M_n = \sum_{j=0}^n (Y_j - E[Y_j | \mathcal{F}_{j-1}])$ est une martingale.

La proposition suivante au sujet des martingales (tirée du Chapitre 1 de [24]) sera utile pour obtenir des bornes sur les termes impliqués dans l'approximation stochastique.

Proposition 1 Soit $\{M_n\}$ une martingale et $q(\cdot)$ une fonction non décroissante, non négative et convexe. Alors pour tout entier $n < N$ et $\lambda > 0$,

$$P \left\{ \sup_{n \leq m \leq N} |M_m| \geq \lambda \right\} \leq \frac{E[q(M_N)]}{q(\lambda)}. \quad (1.5)$$

Les formes couramment utilisées pour la fonction $q(\cdot)$ sont $q(M) = |M|$, $q(M) = |M|^2$ et $q(M) = \exp(\alpha M)$ pour $\alpha > 0$.

Finalement, on définit les processus suivants : soit $Y^0(t) = 0$ pour $t \leq 0$ et

$$Y^0(t) = \sum_{i=0}^{m(t)-1} \epsilon_i Y_i \quad \text{pour } t \geq 0.$$

Le processus décalé $Y^n(\cdot)$ est défini par

$$Y^n(t) = Y^0(t_n + t) - Y^0(t_n) = \sum_{i=n}^{m(t_n+t)-1} \epsilon_i Y_i \quad \text{pour } t \geq 0.$$

De façon analogue à celle de $Y^n(t)$, on définit les processus $M^n(t)$ et $B^n(t)$ en utilisant δM_i et β_i .

Le théorème suivant, tiré du Chapitre 5 de [35], stipule que pour tout $\omega \in \Omega$ fixe n'appartenant pas à un ensemble de mesure nulle N , les processus interpolés décalés $\theta^n(t, \omega)$ convergent (dans l'espace des fonctions en t continues par morceaux avec la norme sup) vers une fonction continue. De plus, cette fonction limite est indépendante de ω , donc on dit que la suite de fonctions converge presque sûrement.

Théorème 1 (Convergence a.p.1) Soient les hypothèses suivantes :

(A1) Il existe une fonction mesurable $g(\cdot)$ continue et bornée, ainsi qu'une variable aléatoire β_n \mathfrak{F}_{n-1} -mesurable, telles que $E[Y_n | \mathfrak{F}_{n-1}] = g(\theta_n) + \beta_n$.

(A2) $\sup_n E|Y_n|^2 < \infty$

(A3) $\sum_{i=1}^{\infty} \epsilon_i^2 < \infty$

(A4) $\sum_{i=1}^{\infty} \epsilon_i = \infty$, $\epsilon_i \geq 0$, $\epsilon_i \rightarrow 0$

(A5) $\sum_{i=1}^{\infty} \epsilon_i |\beta_i| < \infty$ a.p.1

(A6) L'ODE

$$\frac{d\theta(t)}{dt} = g(\theta(t)) \quad (1.6)$$

a une solution unique pour toute condition initiale. Soit $S \neq \emptyset$, l'ensemble des points limites de l'ODE (points asymptotiquement stables dans le sens de Liapunov).

Si les hypothèses (A1) à (A6) sont vérifiées, alors il existe un ensemble N de mesure nulle tel que pour chaque $\omega \notin N$, $\theta_n(\omega)$ converge vers S . En particulier, si $\bar{\theta}$ est le seul point asymptotiquement stable pour l'ODE, alors $\theta_n \xrightarrow{p.s.} \bar{\theta}$.

Preuve: La présentation de la preuve suit en grandes lignes celle de [35] et [54].

Éq. (1.2) nous permet d'écrire l'algorithme (1.1) comme suit :

$$\theta_{n+1} = \theta_n + \epsilon_n g(\theta_n) + \epsilon_n \delta M_n + \epsilon_n \beta_n \quad (1.7)$$

En utilisant l'interpolation $m(t)$, on écrit

$$\theta^n(t) = \theta_n + \sum_{i=n}^{m(t+t_n)-1} \epsilon_i g(\theta_i) + \sum_{i=n}^{m(t+t_n)-1} \epsilon_i \delta M_i + \sum_{i=n}^{m(t+t_n)-1} \epsilon_i \beta_i. \quad (1.8)$$

Puisque $\theta^n(t)$ est constant par morceaux, la somme des termes $\epsilon_i g(\theta_i)$ est une intégrale et on peut utiliser les processus interpolés pour écrire

$$\theta^n(t) = \theta_n + \int_0^t g(\theta^n(s)) ds + M^n(t) + B^n(t) + \rho^n(t), \quad (1.9)$$

où $\rho^n(t)$ est un terme d'erreur ajouté à cause du fait qu'on a remplacé la première somme par une intégrale. $\rho^n(t) = 0$ aux moments auxquels les processus interpolés ont des sauts et $\rho^n(t) \rightarrow 0$ uniformément en t lorsque $n \rightarrow \infty$.

Soit $M_n = \sum_{i=1}^n \epsilon_i \delta M_i$. M_n étant une martingale, on peut donc utiliser la Propo-

sition 1 avec $q(x) = x^2$ et écrire : pour tout $\mu > 0$,

$$\mathbb{P} \left\{ \sup_{m \geq n} |M_m - M_n| > \mu \right\} \leq \frac{1}{\mu^2} \mathbb{E} \left[\sum_{i=n}^m \epsilon_i \delta M_i \right]^2 = \frac{1}{\mu^2} \sum_{i=n}^m \epsilon_i^2 \mathbb{E} [\delta M_i]^2 \leq \frac{K}{\mu^2} \sum_{i=n}^m \epsilon_i^2, \quad (1.10)$$

où K est une constante quelconque. Dans l'égalité ci-haut, on a utilisé le fait que $\mathbb{E}[\delta M_n \delta M_m] = 0$. La deuxième inégalité de (1.10) tient par (A2) et (A3).

Ainsi, par (1.10) et (A5), il existe un ensemble N de mesure nulle tel que pour tout $\omega \notin N$, $M^n(\omega, t)$ et $B^n(\omega, t)$ convergent uniformément vers zéro sur tout intervalle borné dans $(-\infty, \infty)$ lorsque $n \rightarrow 0$.

Soit $\omega \notin N$. Par la supposition (A1), l'ensemble $\{\theta^n(\cdot)\}$ est équicontinu. Une fois ω choisi, on applique le Théorème d'Arzelà-Ascoli (Chapitre 4 de [35]) pour établir qu'il existe une sous-suite convergente $\{\theta^{n_k}(\omega, \cdot)\}$, dont la limite est Lipschitz-continue. Appelons $\theta(\omega, t)$ cette limite. Éq. (1.9) nous permet d'écrire :

$$\theta(\omega, t) = \theta(\omega, 0) + \int_0^t g(\theta(\omega, s)) ds.$$

On constate alors que chaque sous-suite convergente est solution de l'ODE (1.6)

$$\frac{d\theta(t)}{dt} = g(\theta(t)),$$

qui, par (A6), a une solution unique pour toute condition initiale. Ainsi, pour toute condition initiale $\theta_0 \in S$, tous les points limites de $\{\theta^n(\omega, \cdot)\}$ sont dans S , car les fonctions limites sont toutes les mêmes (par (A6)), peu importe la sous-suite choisie. Donc, a.p.1, $\lim_{t \rightarrow \infty} \{ \lim_{n \rightarrow \infty} \theta^n(\omega, t) \} = \lim_{t \rightarrow \infty} \theta(t)$ appartient à l'ensemble invariant S (par la définition d'ensemble invariant). Par conséquent, $\{\theta^n(\omega, \cdot)\}$ converge a.p.1 vers un ensemble invariant de (1.6) et tous les points limites seront des points asymptotiquement stables.

En particulier, si $\bar{\theta} \in S$ est unique, $\theta_n \xrightarrow{p.s.} \bar{\theta}$.

□

Notons que si $g(\theta)$ est le gradient d'une fonction, le Théorème 1 démontre que l'al-

gorithme d'approximation stochastique va converger a.p.1 vers un point critique de cette fonction. C'est surtout dans ce contexte que sera utilisé l'algorithme dans les chapitres suivants (voir section 1.4).

1.3 Convergence avec probabilité 1 - autres méthodes

Nous venons de voir la méthode de l'ODE, mais il existe aussi d'autres alternatives pour la preuve de la convergence presque sûre, s'appliquant dans des contextes différents et assumant d'autres hypothèses de départ. Voici un bref aperçu de deux d'entre-elles.

Les méthodes de "stabilité" (cf. Section 5.4 de ^[35]) sont utiles lorsque les itérés sont dans un ensemble non borné ou lorsque le problème est sans contraintes. Ces méthodes utilisent des fonctions de Liapunov pour montrer la nature récurrente des itérés, c'est-à-dire, pour montrer que les pas successifs de l'approximation stochastique retournent infiniment souvent dans un ensemble compact a.p.1. Ensuite, la méthode de l'ODE termine la preuve en montrant qu'asymptotiquement, en commençant dans un point de récurrence, les itérations suivent la trajectoire de l'ODE moyenne.

La méthode de "compacité générale" (cf. Section 5.3 de ^[35]) est certainement la plus puissante, car elle utilise des hypothèses plus faibles sur le bruit et sur la suite des gains. Cette approche s'applique aussi bien à des bruits qui sont des différences de martingales qu'à des bruits corrélés. La méthode commence avec des conditions générales sur le taux de variation asymptotique des processus interpolés $\{M^n(\cdot)\}$ et $\{B^n(\cdot)\}$, garantissant l'équicontinuité (et par conséquent le théorème limite) et ensuite trouve des conditions spécifiques qui vérifient les conditions générales. Pour certaines classes de problèmes, les conditions générales sont nécessaires et suffisantes. Pour plus de détails sur ces méthodes de preuves, nous référons le lecteur à ^[35].

1.4 Contexte d'application

Le contexte dans lequel nous allons appliquer l'algorithme d'approximation stochastique dans les chapitres à venir est le suivant : nous chercherons la solution d'un

problème d'optimisation, soit :

$$\min_{\theta} J(\theta),$$

où $J(\cdot)$ est la fonction (aléatoire) représentant la façon dont dépend le système du paramètre θ à optimiser. (Dans le cas du Chapitre 2, l'optimisation sera faite sur le vecteur de contrôle.) Dans ce contexte d'optimisation, se référant à l'équation récurrente (1.1), $Y_n = \hat{g}(\theta_n)$ sera l'estimateur du gradient $g(\cdot)$ de la fonction objectif $J(\cdot)$ au point θ_n . Ce gradient sera estimé à l'aide de perturbations aléatoires dans le cas du Chapitre 2 et à l'aide d'une dérivée stochastique (IPA) dans le cas de Chapitre 3 et Chapitre 4.

Comme mentionné à la fin de la section 1.2, lorsque $g(\theta)$ est un gradient, les itérés de l'approximation stochastique vont suivre approximativement ceux de la méthode de la plus forte pente et les points de stabilité de l'ODE cible seront des points de premier ordre de la fonction objectif.

Pour démontrer la convergence presque sûre de nos algorithmes, nous devons vérifier les hypothèses du Théorème 1. Ainsi, l'estimateur du gradient Y_n devra être borné a.p.1 pour satisfaire l'hypothèse (A2), autrement dit, la fonction objectif devra être Lipschitz-continue. Si tel est le cas, une petite perturbation $\Delta\theta$ du paramètre produira une perturbation de la fonction $J(\cdot)$ bornée par $K\Delta\theta$, où K est une variable aléatoire d'espérance finie. Enfin, la suite des gains $\{\epsilon_n\}$ devra satisfaire (A4), ou bien le pas $\epsilon_n = \epsilon$ sera constant dans le cas du contrôle en ligne et à ce moment-là, l'algorithme convergera en distribution seulement.

CHAPITRE 2

APPLICATION AU CONTRÔLE PRÉDICTIF D'UN PROCÉDÉ CHIMIQUE

Le sujet de ce chapitre est l'application de l'approximation stochastique au contrôle prédictif d'un réacteur chimique nommé Van der Vusse. Ce projet a été réalisé en collaboration avec les professeurs Smaranda Cristea et César de Prada Moraga de l'Université de Valladolid, Espagne. L'article résultant ^[7] a été présenté lors de la Conférence de Simulation Industrielle ISC2003, à Valence, Espagne, en juin 2003. Je tiens à remercier le Département d'Ingénierie des Systèmes et Automatique (ISA) pour m'avoir accueillie en juillet-août 2002 et en mai-juin 2003 afin de poursuivre cette collaboration.

Le contrôle prédictif ^[12, 41, 42] est une technique bien connue dans le milieu industriel, pour sa capacité à traiter un large éventail de problèmes de contrôle à grande dimension, comportant ou non des contraintes. Elle utilise un *modèle dynamique explicite* de l'usine pour prédire l'effet qu'auront les actions postérieures des variables manipulées sur la production. L'évolution future de ces variables est déterminée en résolvant un problème d'optimisation dont l'objectif est la minimisation de l'erreur prédite, sujet à des contraintes opérationnelles.

Cependant, la plupart des contrôleurs industriels étant basés sur des modèles linéaires, leur applicabilité est quelque peu limitée. Pour cette raison, le contrôle prédictif à modèle non linéaire (NMPC) ^[16, 17] a attiré l'attention des chercheurs ces dernières années, du point de vue de ses propriétés et de son implémentation. Quant à ce dernier aspect, l'inconvénient principal de NMPC est son fardeau computationnel. Alors que le contrôle prédictif linéaire avec contraintes peut résoudre le problème d'optimisation associé dans les délais requis en faisant appel à des algorithmes de programmation linéaire ou quadratique, NMPC a besoin de méthodes de programmation non linéaire, telle SQP ("Successive Quadratic Programming"), qui sont connues pour être plus gourmandes en temps CPU. Plusieurs approches ont été proposées pour résoudre ce problème, entre autres, les approches séquentielle et simultanée.

Dans les solutions utilisant SQP avec l'approche séquentielle, la dynamique du

système est simulée à chaque itération de la routine d'optimisation. Seuls les paramètres de contrôle (variables manipulées) demeurent comme degrés de liberté dans le problème de programmation non linéaire. Les routines de simulation et d'optimisation sont exécutées séquentiellement, les unes après les autres. Cette approche peut facilement être jumelée à des outils de simulation avancés.

En revanche, la simulation et l'optimisation de l'approche simultanée incluent les états du modèle et les contrôles comme variables de décision. Les équations du modèle sont ajoutées au problème d'optimisation comme des contraintes d'égalité. Ceci augmente considérablement la dimension du problème d'optimisation et c'est là une différence majeure entre les deux approches. Néanmoins, le temps de calcul est une problématique présente dans les deux cas, surtout lorsque NMPC doit être implémenté en temps réel.

Très souvent, les ingénieurs utilisent les procédures d'optimisation disponibles dans les bibliothèques commerciales, comme par exemple la bibliothèque NAG ¹. La plupart du temps, dans ces bibliothèques, le gradient de la fonction objectif est estimé par des différences finies s'il n'est pas fourni par l'utilisateur. Ceci signifie que pour chaque composante du gradient, deux simulations du système doivent être effectuées afin d'obtenir les différences finies, ce qui rend la procédure d'optimisation très lente, surtout pour des modèles à grande dimension.

Une solution existante du problème de Van der Vusse repose sur la méthode d'optimisation *locale* SQP de la bibliothèque NAG. Cette solution a démontré plusieurs défaillances : le contrôle prédictif ne peut être appliqué en ligne, à cause du temps de calcul de SQP trop élevé ; un contrôle sous-optimal est parfois observé, peut-être causé par des optima locaux dans lesquels SQP est possiblement piégé.

Dans ce chapitre, nous allons remplacer la méthode SQP dans l'optimisation du contrôle prédictif par une méthode d'optimisation *globale*, reconnue en outre pour sa performance sur des problèmes à grande dimension. SPSA ("Simultaneous Perturbation Stochastic Approximation") ^[48, 40, 25] est une variante de la procédure de Kiefer-

¹NAG : Numerical Algorithms Group, <http://www.nag.com/>

Wolfowitz, qui au lieu d'utiliser des différences finies pour estimer les gradients, utilise des perturbations aléatoires et simultanées dans toutes les directions.

Le chapitre est divisé comme suit : la section 2.1 décrit le modèle du réacteur chimique et la section 2.2 présente le problème de contrôle prédictif. La section 2.3 résume les principales hypothèses de convergence de la méthode SPSA et la section 2.4 donne les résultats du contrôle prédictif avec cette méthode.

2.1 Description du modèle

La réaction Van der Vusse est décrite en détail dans ^[17] et les références qui s'y trouvent. En résumé, une substance chimique A entre dans le réacteur et réagit avec d'autres substances pour produire la substance B en sortie, en dégageant de la chaleur comme illustré sur la Figure 2.1.

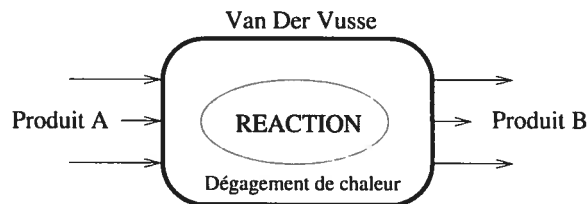


FIG. 2.1 – Réaction Van der Vusse

Définissons les variables réelles suivantes :

c_B la concentration de produit B (variable contrôlée) ;

c_A la concentration de produit A (mesurée) ;

T la température dans le réacteur (mesurée) ;

T_K la température dans le réfrigérant (mesurée) ;

F le flot d'entrée du produit A (variable manipulée) ;

Q_K le dégagement de chaleur (variable manipulée) ;

$c_{A0}, T_0, k_1, k_2, k_3, k_w, \rho, c_p, V_R, \Delta H_1, \Delta H_2, \Delta H_3, m_K$ des constantes.

Les *variables de contrôle* sont F et Q_K . La *variable contrôlée* est c_B et on veut la contrôler en manipulant les variables d'entrée F et Q_K .

Pour décrire la dynamique de ce système, on introduit les équations différentielles non linéaires reliées à la conservation de masse et d'énergie (la notation ' dénote la dérivée par rapport au temps) :

$$\begin{aligned}
 c'_A &= \frac{F}{V_R}(c_{A0} - c_A) - k_1 c_A - k_3 c_A^2 \\
 c'_B &= -\frac{F}{V_R} c_B + k_1 c_A - k_2 c_B \\
 T' &= \frac{F}{V_R}(T_0 - T) - \frac{1}{\rho C_p}(k_1 c_A \Delta H_1 + k_2 c_B \Delta H_2 + \\
 &\quad + k_3 c_A^2 \Delta H_3) + \frac{k_w A_R}{\rho C_p V_R}(T_K - T) \\
 T'_K &= \frac{1}{m_K C_{pK}}(Q_K + k_w A_R(T - T_K))
 \end{aligned} \tag{2.1}$$

La concentration de produit B ne doit pas dépasser certaines bornes inférieures et supérieures, ce qui introduit la contrainte suivante :

$$l \leq c_B \leq L.$$

Les équations différentielles ci-haut sont complètement déterministes. Cette situation est bien sûr très peu courante dans les applications réelles, mais elle est étudiée ici dans le but de résoudre les difficultés reliées à l'utilisation du contrôle prédictif sur un système non linéaire tel le réacteur Van der Vusse. Il s'avère qu'en ajoutant de l'aléa à ce problème déterministe via la méthode SPSA, la performance du contrôle prédictif est accrue. Nous allons voir qu'elle l'est encore plus lorsque le problème lui-même est stochastique (voir section 4.5).

2.2 Contrôle prédictif non linéaire

Considérons le cas déterministe. Soient :

$x(t) = (c_A, c_B, T, T_K)$ le vecteur d'état au temps t ;

$y(t) = (c_B)$ la variable contrôlée au temps t ;

$u(t) = (F, Q_K)$ le contrôle à appliquer au temps t ;

$r(t)$: le niveau de référence à atteindre au temps t .

L'objectif du contrôle prédictif^[12, 41, 42] est de trouver la séquence de contrôle optimale $u^*(t)$ sur un horizon de prédiction fini T , telle que l'erreur quadratique entre la variable contrôlée $y(t)$ et son niveau de référence r soit minimale. De plus, $u^*(t)$ doit être tel que la trajectoire de $y(t)$ soit lisse. Ce sera le cas lorsqu'il n'y a pas de grande variabilité dans les valeurs de $u(t)$. Dans le cas déterministe, la fonction objectif prend donc la forme

$$J(x(0), y(0), \{u(t)\}_{t \geq 0}) = \int_0^T \|y(t) - r(t)\|^2 dt + \beta \int_0^T \|u'(t)\|^2 dt,$$

où $u'(t) = \frac{du(t)}{dt}$ et $\beta \geq 0$ est une constante. Puisque le problème est déterministe, la valeur de la fonction objectif au temps t dépend seulement des valeurs initiales de x et y , ainsi que de la trajectoire du contrôle jusqu'au temps t : $\{u(t)\}_{t \geq 0}$. Ainsi, on minimise sur l'ensemble de contrôles réalisables \mathcal{U} :

$$\min_{u \in \mathcal{U}} J(x(0), y(0), \{u(t)\}_{t \geq 0}).$$

Pour résoudre ce problème, il est nécessaire de discrétiser $u(t)$. On définit le *temps d'échantillonnage* h comme étant l'intervalle de temps pendant lequel les paramètres et le contrôle appliqué au système sont gardés constants. Soit N_p l'horizon de prédiction et N_c l'horizon de contrôle, avec $N_p \geq N_c$. N_p est égal au nombre de périodes d'échantillonnage sur lesquelles sera faite l'optimisation et N_c est le nombre de contrôles futurs qui seront calculés pour *chacune* des variables de contrôle. Si $k \in \mathbb{N}$ est l'indice du temps discrétisé, on a

$$u(k) = \begin{cases} u(t) & \text{si } kh \leq t < (k+1)h \\ u(N_c - 1) & \text{si } k > N_c - 1. \end{cases}$$

Pour que le contrôle soit mieux adapté à l'état actuel du système $x(t)$, un *horizon fuyant*² est utilisé. Cette technique exige que lorsque le contrôle optimal $u^*(x) :=$

²Horizon fuyant : "receding horizon"

$(u_0^*(x), \dots, u_{N_c}^*(x))$ est trouvé, seulement $u_0^*(x)$ est appliqué au système au temps t . Au temps $t + h$, une autre séquence optimale est calculée et de nouveau, seulement la première composante de cette séquence est appliquée au système. À chaque étape, l'horizon de prédiction N_p est déplacé de h unités de temps (horizon fuyant).

Avec la discrétisation, le problème d'optimisation (avec contraintes) peut donc s'écrire ainsi : à chaque temps de calcul $k \in \mathbb{N}$,

$$\min_{u \in \mathcal{U}} J = \int_{t_k}^{t_k + N_p} \gamma [y(t) - r(t)]^2 dt + \sum_{i=0}^{N_c-1} \beta [\Delta u(k + i|t)]^2 \quad (2.2)$$

$$y_{min} \leq y(k) \leq y_{max} \quad (2.3)$$

$$u_{min} \leq u(k) \leq u_{max} \quad (2.4)$$

$$\Delta u_{min} \leq \Delta u(k) \leq \Delta u_{max}, \quad (2.5)$$

où $\Delta u(k) = u(k) - u(k - 1)$ et $\Delta u(k + i|t)$ est la valeur prédite de Δu au temps $k + i$, basée sur l'information disponible au temps t . En pénalisant les contraintes (2.3), le problème devient :

$$\begin{aligned} \min_u J = & \int_{t_k}^{t_k + N_p} \gamma [y(t) - r(t)]^2 dt + \sum_{i=0}^{N_c-1} \beta [\Delta u(k + i|t)]^2 \quad (2.6) \\ & + \sum_{j=0}^{N_c-1} \alpha_1 [y_{max} - y(k + j|t)]^2 + \sum_{l=0}^{N_c-1} \alpha_2 [y_{min} - y(k + l|t)]^2 \end{aligned}$$

où β et γ sont des constantes non négatives, α_1, α_2 représentent une fonction de pénalité et $y(k + j|t)$ est la valeur prédite de y au temps $k + j$, basée sur l'information disponible au temps t .

On remarque que ce problème d'optimisation ne peut être résolu analytiquement, car il est difficile de calculer le gradient de J à cause de la non-linéarité de y en u . C'est pour cette raison que nous allons l'estimer. De plus, il faut noter que la dimension p du problème peut augmenter très rapidement. Celle-ci est déterminée par l'horizon de contrôle N_c qui peut être différent pour les deux variables de contrôle F et Q_K . Par exemple, on peut utiliser un horizon de contrôle $N_{c,1} = 2$ pour la variable F et

$N_{c,2} = 3$ pour la variable Q_K , ce qui donne une dimension de contrôle $p = 5$. Comme on cherche une bonne précision, N_c est désiré assez grand, donc la dimension du problème est importante.

Afin de respecter les contraintes (2.4) et (2.5) du vecteur de contrôle, nous allons faire une projection du Lagrangien de J sur l'ensemble de contrôles réalisables :

$$u(k) = \begin{cases} u_{max} & \text{si } u(k) \geq u_{max} \\ u_{min} & \text{si } u(k) \leq u_{min}, \end{cases} \quad (2.7)$$

et de même pour les contraintes de $\Delta u(k)$.

Pour tester la qualité du contrôle prédictif, une simulation de la réaction Van der Vusse est exécutée à l'aide du logiciel ECOSIM³. Un horizon de T unités de temps est simulé. À chaque temps d'échantillonnage, i.e. lorsque $t = nh$, $n = 1, 2, \dots$, le contrôle optimal est calculé en résolvant le problème d'optimisation (2.6)-(2.7). Les variables sont mises à jour et les nouvelles mesures sont prises. Le temps est ensuite avancé de h unités. Pour illustrer, voici l'algorithme de la simulation du système *déterministe* :

```

WHILE (t < T) DO
  IF (t mod h == 0) DO
    1. Calculer le contrôle optimal4 ( $\min_{u \in \mathcal{U}} J(x(0), y(0), \{u(t)\}_{t \geq 0})$ );
    2. Appliquer le contrôle  $u_0^*$ ;
    3. Mesurer  $c_A, T, T_K$ ;  $x_t \leftarrow (c_A, T, T_K)$ ;
    4. Observer la nouvelle valeur de  $c_B$ ;  $y_{t+1} \leftarrow c_B$ ;
    5.  $t = t + h$ ; aller à 1.
  END IF
END WHILE

```

³ECOSIM : logiciel de simulation ; <http://www.ecosimpro.com/>

⁴Pour calculer le contrôle optimal, la programmation dynamique pourrait être utilisée, mais le but de ce travail est l'amélioration du contrôle prédictif, c'est pour cette raison que nous employons cette technique.

Il est clair que le contrôle optimal doit être trouvé en moins de h secondes (où h est la durée du temps d'échantillonnage), si l'on souhaite que le contrôle soit fait en ligne.

Comme mentionné dans l'introduction de ce chapitre, des approches antérieures pour résoudre le problème d'optimisation apparaissant à l'étape 1 de la procédure ci-haut ont fait appel à l'algorithme de programmation quadratique SQP. Nos collaborateurs de l'Université de Valladolid ont utilisé la méthode SQP implantée dans la librairie NAG, qui utilise des différences finies pour estimer ∇J . Cet estimateur du gradient ralentit la simulation à cause du grand nombre d'évaluations de la fonction objectif (chaque évaluation de J étant une intégration numérique du système d'équations différentielles (2.1)). De plus, la méthode SQP peut être attirée par des minimums locaux. Afin de résoudre ces problèmes, nous allons remplacer la méthode SQP de NAG par la méthode de directions aléatoires SPSA.

2.3 Simultaneous Perturbation Stochastic Approximation (SPSA)

SPSA ^[48, 40, 25] est une variante de l'algorithme d'approximation stochastique, capable de trouver des optimum globaux. Sa principale caractéristique est l'estimation du gradient à l'aide de perturbations aléatoires, qui requiert seulement deux évaluations de la fonction objectif par itération, peu importe les dimensions du problème d'optimisation.

Pour alléger la notation de cette section, au lieu de dénoter la fonction objectif par $J(x(0), y(0), \{u(t)\}_{t \geq 0})$, nous allons indiquer seulement la dépendante en u . Ainsi, comme indiqué dans la section précédente, nous cherchons le contrôle optimal u^* tel que

$$u^* = \arg \min \{ J(u) : u \in \mathcal{U} \},$$

où $J(u)$ est donné par (2.6). Le processus itératif est le suivant :

$$u_{n+1} = u_n - \epsilon_n \hat{g}(u_n), \quad (2.8)$$

où $u_n = ((u_n)_1, (u_n)_2, \dots, (u_n)_p)^T$ représente le $n^{\text{ième}}$ itéré, $\hat{g}(u_n)$ est l'estimateur du

gradient $g(u) = \nabla_u J(u)$ évalué en u_n , et $\{\epsilon_n\}$ est la suite de gains décroissants.

Le gradient $g(u) = \nabla_u J(u)$ peut être estimé de plusieurs façons. Si u_n est un vecteur de dimension p , la $i^{\text{ième}}$ composante de l'estimateur du gradient symétrique à différences finies est

$$\text{FD : } (\hat{g}(u_n))_i = \frac{J(u_n + c_n e_i) - J(u_n - c_n e_i)}{2c_n}, \quad (2.9)$$

$1 \leq i \leq p$, où e_i est le $i^{\text{ième}}$ vecteur unitaire, et c_n est un nombre positif qui décroît avec n . Avec cette méthode, $2p$ évaluations de $J(\cdot)$ pour chaque $g(u_n)$ sont nécessaires. Évidemment, lorsque p est grand, cet estimateur perd de l'efficacité.

L'estimateur à perturbations simultanées utilise un vecteur de perturbations aléatoires à p dimensions $\Delta_n = ((\Delta_n)_1, (\Delta_n)_2, \dots, (\Delta_n)_p)^T$ (voir les détails concernant la distribution plus loin). La $i^{\text{ième}}$ composante de cet estimateur est

$$\text{SP : } (\hat{g}(u_n))_i = \frac{J(u_n + c_n \Delta_n) - J(u_n - c_n \Delta_n)}{2c_n (\Delta_n)_i}, \quad (2.10)$$

où $P[(\Delta_n)_i = 0] = 0, \forall n, i$.

En comparant ces deux estimateurs, on remarque que FD perturbe une seule direction à la fois, alors que SP modifie toutes les directions simultanément. En effet, le numérateur de (2.10) reste le même pour toutes les p composantes. Par conséquent, le nombre d'évaluations de la fonction objectif nécessaires pour l'estimation du gradient avec l'estimateur SP est toujours de deux, indépendamment de la dimension p . Ainsi, la méthode SPSA peut être beaucoup plus efficace lorsque la dimension est grande.

Nous avons comparé la performance des deux estimateurs sur le problème d'optimisation déterministe suivant :

$$\begin{aligned} \min_u J(u_1, u_2) &= (3u_1 + 2u_2 - 5)^2 + 2(u_1 - u_2)^2 \\ \text{s.à} & \quad -2 \leq 3u_1 + 2u_2 \leq 10 \end{aligned} \quad (2.11)$$

Les résultats de ce problème simple à deux dimensions ont démontré que l'algorithme SPSA converge dans le même nombre d'itérations que l'algorithme FDSA (voir fi-

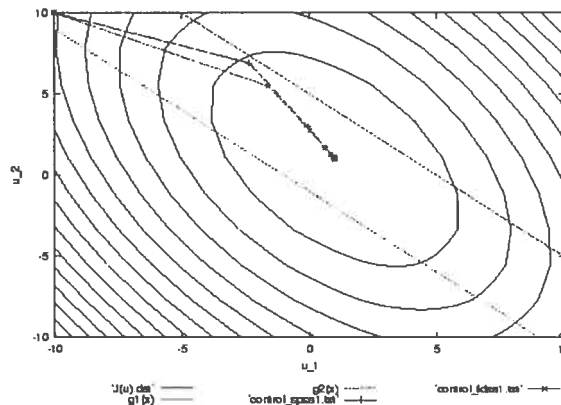


FIG. 2.2 – Les trajectoires de l’algorithme de descente avec l’estimateur du gradient à différences finies (FD) et l’estimateur à perturbations aléatoires (SP) sur le problème déterministe (2.11).

gure 2.2). Ce dernier suit approximativement la trajectoire de l’algorithme du gradient, alors que SPSA, avec ses directions de recherche aléatoires, peut s’en éloigner. Cependant en moyenne, l’estimateur SP suit le vrai gradient, car c’est un estimateur non biaisé, comme démontré dans le Lemme suivant, tiré de [25].

Lemme 1 *Soit*

$$b_n = E[\hat{g}(u_n)] - \nabla_u J(u_n)$$

le biais dans l’estimateur $\hat{g}(\cdot)$. Soit $\{\Delta_n\}$ une suite de vecteurs i.i.d. dont les composantes $\{(\Delta_n)_i, i = 1, \dots, p\}$ sont des variables aléatoires i.i.d., symétriquement distribuées autour de zéro et de moments d’ordre deux bornés. Supposons que $E|(\Delta_n)_i^{-1}|$ est bornée uniformément sur \mathcal{U} . Alors $b_n \rightarrow 0$ a.p. 1.

Preuve: La preuve détaillée se trouve dans [25]. L’idée principale est de conditionner sur Δ_n pour développer la $i^{\text{ème}}$ composante $E[(\hat{g}(u_n))_i]$ et ensuite utiliser une expansion en série de Taylor d’ordre deux de $J(u_n + c_n \Delta_n)$ et $J(u_n - c_n \Delta_n)$. Après quelques manipulations algébriques impliquant l’espérance zéro et l’indépendance des $\{(\Delta_n)_i\}$,

$i = 1, \dots, p$, on obtient

$$E[(\hat{g}(u_n))_i] = (g(u_n))_i + \mathcal{O}(c_n).$$

Le résultat suit de l'hypothèse que $c_n \rightarrow 0$.

□

Le théorème suivant est tiré de ^[40] et démontre la convergence presque sûre de l'algorithme SPSA vers le minimum global de $J(\cdot)$. On retrouve bien évidemment les hypothèses nécessaires pour la convergence a.p.1 de l'approximation stochastique donnée dans le Chapitre 1. Cependant, puisqu'ici nous prouvons la convergence vers un minimum *global*, quelques hypothèses s'ajoutent. ⁵

Théorème 2 (Convergence a.p.1 de SPSA) Soit \mathcal{F}_n la σ -algèbre engendrée par les v.a. u_0, u_1, \dots, u_n . Soient les hypothèses suivantes :

- (J1) Soient Δ_{ni} des v.a. de moyenne zéro, mutuellement indépendantes et symétriquement distribuées autour de zéro, avec $|\Delta_{ni}| \leq \alpha_1 < \infty$ et $E|\Delta_{ni}^{-2}| \leq \alpha_2 < \infty$ p.s., $\forall i, n$.
- (J2) Soient $\xi_n^{(+)}$ et $\xi_n^{(-)}$ le bruit aléatoire dû aux mesures de la fonction objectif. Ce bruit satisfait $E[\xi_n^{(+)} - \xi_n^{(-)} \mid \mathcal{F}_n, \Delta_n] = 0$ p.s. $\forall n$. Les termes de la suite $\{\xi_n^{(\pm)}\}$ ne sont pas nécessairement indépendants.
- (J3) $J(x, y, u)$ est trois fois continuellement différentiable par rapport à u et les éléments individuels de la dérivée troisième sont bornés : $\left| \frac{\partial^3 J(x, y, u)}{\partial u^3} \right| \leq \alpha_3 \leq \infty$. Par ailleurs, $|J(x, y, u)| \rightarrow \infty$ lorsque $u \rightarrow \infty$.
- (J4) $\epsilon_n > 0$, $\epsilon_n \rightarrow 0$ et $\sum_{t=1}^{\infty} \epsilon_n = \infty$ (cf. (A4), Théorème 1). La suite $\{c_n\}$ est de la forme $c_n = c/t^\gamma$, où $c > 0$ et $\gamma \in [\frac{1}{6}, \frac{1}{2})$; de plus, $\sum_{t=1}^{\infty} (\epsilon_n/c_n)^2 < \infty$.
- (J5) Le gradient $g(u)$ est borné (cf. (A2), Théorème 1) et Lipschitz-continu.
- (J6) L'ODE $\dot{u} = g(u)$ a une solution unique pour toute condition initiale (cf. (A6), Théorème 1).
- (J7) Soit $\delta M_n = \hat{g}(u_n) - E[\hat{g}(u_n) \mid \mathcal{F}_n]$. On suppose $E|\sum_{i=1}^n \delta M_i| < \infty$, $\forall i$.

⁵Pour tous les détails, voir ^[40].

Sous les hypothèses J1 à J7, u_n converge a.p.1 vers l'ensemble des minimums globaux de $J(u)$.

Preuve: Pour une preuve détaillée, nous référons le lecteur à l'Annexe de l'article de Maryak et Chin ^[40]. De façon intuitive, le bruit aléatoire Δ_n (non nul) de moyenne zéro qui est injecté à l'estimateur du gradient peut faire sortir l'algorithme d'un minimum local. Comme la suite des gains est décroissante, les pas de l'algorithme deviennent de plus en plus petits et éventuellement, l'algorithme convergera vers un minimum global.

L'idée de la preuve formelle est la suivante : on récrit la récursion de SPSA comme

$$u_{n+1} = u_n - \epsilon_n [g(u_n) + \zeta_n],$$

où $\zeta_n := \hat{g}(u_n) - g(u_n)$ est le bruit causé par la perturbation aléatoire de l'estimateur SP. Des résultats de la théorie des grandes déviations sont utilisés pour compléter la preuve.

□

2.4 Résultats

Les expériences effectuées avec le simulateur ECOSIM montrent qu'effectivement, SPSA offre une meilleure performance que la méthode SQP avec les différences finies comme estimateur du gradient. Les figures de la page suivante comparent, sous différentes conditions, la qualité du contrôle prédictif trouvé à l'aide de (a) la méthode SQP de la librairie NAG, et (b) la méthode SPSA. Dans ces graphes sont tracés la trajectoire de la variable contrôlée $y(t)$ (la concentration de produit B), ses bornes inférieure et supérieure, ainsi que le niveau de référence r qui change avec le temps. Le temps est mesuré en secondes.

Nous avons remarqué que la qualité du contrôle dépend énormément des valeurs initiales des paramètres ϵ_0, c_0 . Des simulations pilotes ont été nécessaires afin de trouver les valeurs justes des paramètres. Pour garantir la convergence de SPSA, i.e., pour satisfaire les hypothèses du Théorème 2, les règles de mise à jour des paramètres ϵ_n, c_n et Δ_n étaient les suivantes :

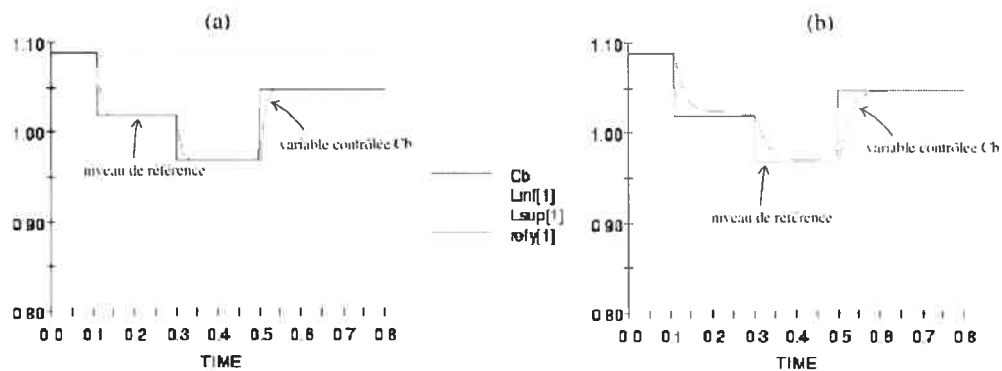


FIG. 2.3 – Évolution de la variable contrôlée c_B en fonction du temps. Dimension du problème : $p = 2$. (a) SQP : 52 sec. (b) SPSA : 40 sec.

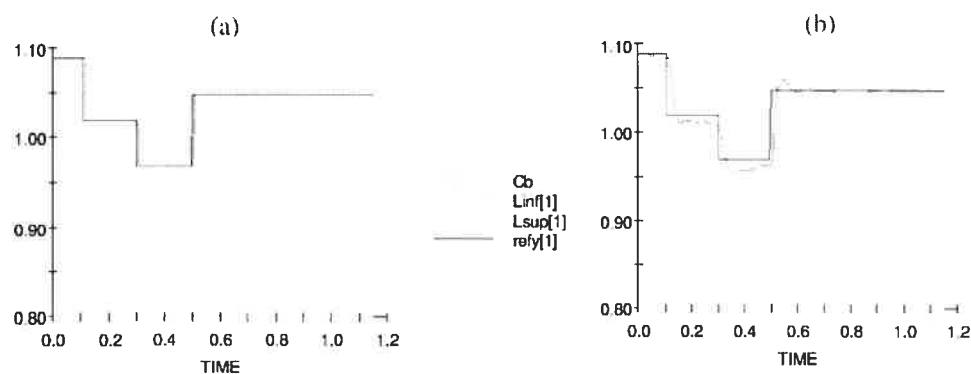


FIG. 2.4 – Évolution de la variable contrôlée c_B en fonction du temps. Dimension : $p = 10$. (a) SQP : 198 sec. (b) SPSA : 59 sec.

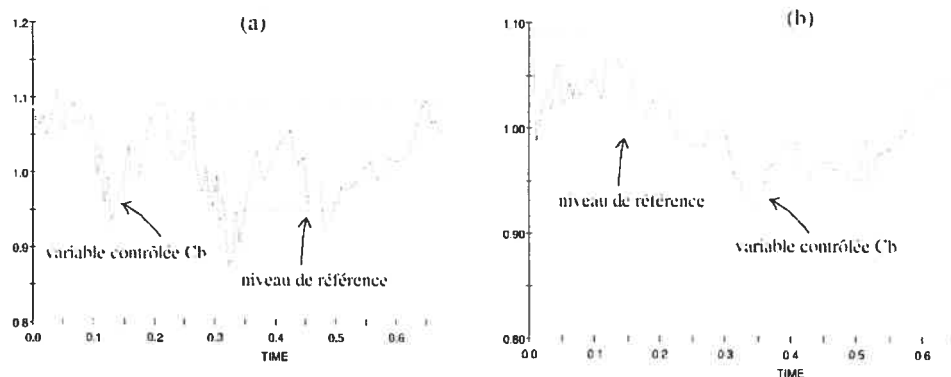


FIG. 2.5 – Évolution de la variable contrôlée c_B en fonction du temps avec du bruit blanc dans les mesures de c_B . Dimension : $p = 2$. (a) SQP : 1139 sec. (b) SPSA : 50 sec.

- $\epsilon_n = \epsilon / (A + n)^\alpha$, $\alpha = 0.6$, $A = 60$ et $\epsilon = 1000$;
- $c_n = c / n^\gamma$, $\gamma = 0.1$ et $c = 10$;
- $\Delta_{ni} \sim \text{Bernoulli}(0.5)$ sur $\{-1, 1\}$.

Pour toutes les figures, l'horizon de prédiction était de $N_p = 10$. Un critère d'arrêt dynamique a été utilisé dans la méthode SPSA de la forme $\|\Delta u\| \leq \beta$, où $\beta = 0.1$.

La dimension du problème de la figure 2.3 est deux, c'est-à-dire, l'horizon de contrôle pour chacune des composantes du vecteur $u(t) = (F, Q_K)$ est égal à un : $N_c = \{1, 1\}$. On voit que dans ce cas, l'efficacité des deux méthodes est à peu près la même. En effet, le temps de simulation qu'a nécessité le contrôle avec la méthode SQP était de 52 sec. et celui avec le méthode SPSA était de 40 sec. Ces résultats ne sont pas surprenants, car pour une dimension $p = 2$, les estimateurs SPSA et FDSA nécessitent le même nombre d'évaluations de la fonction objectif.

Le même schéma est repris pour le graphe de la figure 2.4, sauf que cette fois-ci, $N_c = \{5, 5\}$, i.e., la dimension du problème est $p = 10$. Le temps de simulation des 1.2 sec. du graphe avec la méthode SQP était 198 sec., alors qu'avec SPSA, la simulation a duré 59 sec. SPSA a donc été 3,4 fois plus rapide que SQP. La performance de SPSA n'est pas telle qu'espérée (i.e., 5 fois plus rapide) à cause du nombre d'itérations assez élevé qu'a nécessité la méthode afin de trouver le contrôle optimal de précision 90% à chaque période d'échantillonnage.

Pour l'expérience de la figure 2.5, nous avons simulé le *système aléatoire*, en ajoutant un bruit blanc ($\mathcal{N}(0, 1)$) aux mesures de c_B , afin de tester la qualité de SPSA dans le cas stochastique. La dimension du problème est $p = 2$. C'est dans ce cas que la différence entre les deux méthodes est la plus marquée : SPSA a été 20 fois plus rapide que SQP. On peut donc conclure que dans des cas réalistes où un élément aléatoire est présent dans le système, SPSA est beaucoup plus performante que SQP.

Cependant, un problème a été détecté lors de l'optimisation par la méthode SPSA. Comme on peut voir sur la figure 2.6, lorsque le niveau de référence est proche de la borne inférieure, le contrôle n'atteint pas la valeur désirée et demeure au-dessus de celle-ci. Ce comportement est probablement dû à la façon dont on traite les contraintes, c'est-à-dire causé par la troncation (2.7). Une solution alternative pour traiter ces contraintes

serait d'utiliser la méthode des multiplicateurs ^[9] qui minimise le Lagrangien augmenté. Cette modification fera l'objet de recherches futures.

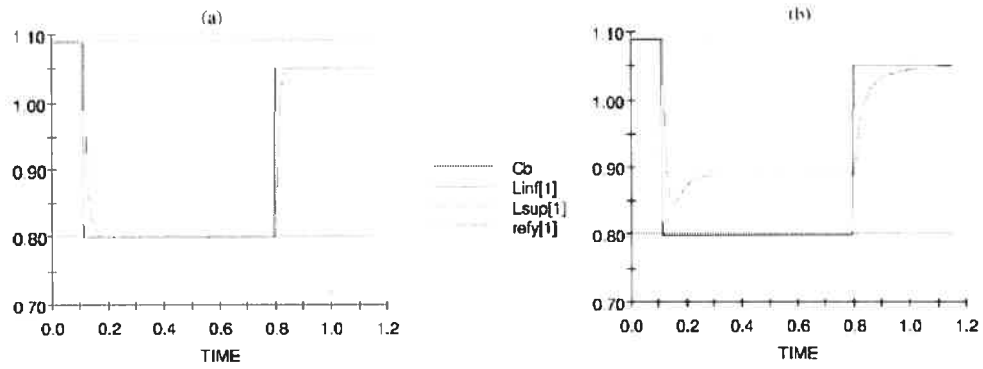


FIG. 2.6 – Évolution du contrôle en fonction du temps. Dimension : $p = 2$. (a) NAG (b) SPSA : contrôle sous-optimal lorsque près de la borne inférieure

CHAPITRE 3

APPLICATION À L'ESTIMATION DE LA PROBABILITÉ DE BLOCAGE DANS UN RÉSEAU CDMA

Ce chapitre présente un algorithme intelligent permettant de simuler efficacement un réseau CDMA sans fil. Ce projet a été réalisé sous la supervision de la professeure Felisa J. Vázquez-Abad dans le cadre de stages d'été entre les années 2000 et 2002 au Laboratoire d'Optimisation et Simulation de l'Université de Montréal. Les résultats de cette recherche ont été présentés dans le 6-ème Atelier International WODES02, à Zaragoza, Espagne, en octobre 2002 ^[55] et le rapport technique ^[56] est publié dans Les Cahiers du GERAD.

Dans un réseau cellulaire CDMA¹, le signal émis par un usager subit une dégradation, causée par des obstacles rencontrés dans l'environnement, tels un édifice ou une montagne. Ce signal est reçu à différentes puissances par les différentes stations réceptrices. La station réceptrice allouée à un mobile est celle qui reçoit la meilleure intensité du signal. L'interférence totale à une station réceptrice est la somme des interférences causées par les mobiles non connectés à cette station. Un blocage est observé à une station réceptrice lorsqu'un mobile se connectant au réseau cause la diminution du ratio signal/interférence en dessous d'un seuil de tolérance.

La probabilité de blocage est une mesure de performance des réseaux cellulaires mobiles CDMA et s'avère importante pour la planification et le design de ces réseaux. Or, des expressions analytiques de la probabilité de blocage ne sont pas connues de nos jours. Malgré que plusieurs études proposent des approximations sous des régimes limites tels un nombre accru d'utilisateurs par cellule, la plupart de ces approximations ne sont pas justes pour des valeurs réalistes des paramètres du modèle.

Dans ce chapitre, nous proposons une méthode de simulation permettant d'évaluer la probabilité de blocage. Puisque celle-ci est un événement rare, des approches de type

¹CDMA : "Code Division Multiple Access"

Monte-Carlo “naïf” ne sont pas appropriées. En effet, un grand nombre d'échantillons doit être simulé avant l'observation d'un blocage et par conséquent, le temps CPU nécessaire à l'obtention d'un estimateur correct peut devenir considérable.

L'échantillonnage stratégique ^[47] est une technique qui a fait ses preuves quant à la simulation d'événements rares. Cette méthode est basée sur le changement de mesure qui permet parfois d'obtenir une erreur relative bornée lorsque les paramètres sont bien choisis. Cependant, il est difficile de trouver analytiquement le changement de mesure optimal. En effet, les bornes de la théorie des grandes déviations capables d'indiquer le changement de mesure optimal ne peuvent être utilisées dans ce problème. La difficulté vient du fait que l'affaiblissement du signal est modélisé par une distribution lognormale pour laquelle la technique habituelle de torsion exponentielle ne s'applique pas. De plus, la distribution des interférences individuelles n'est pas connue.

Pour ces raisons, nous implémentons une méthode de simulation impliquant l'approximation stochastique, qui saura apprendre les paramètres optimaux du changement de mesure, dans le but de réduire la variance de l'estimateur de la probabilité de blocage.

L'échantillonnage stratégique est introduit dans la section 3.3, alors que la méthode d'approximation stochastique, basée sur les travaux de ^[57] et ^[50], est présentée dans la section 3.4. Les résultats se trouvent à la section 3.5. Tout d'abord, voici la description du modèle.

3.1 Modèle et formulation du problème

Afin de bien comprendre le domaine d'application de ce travail, voici quelques définitions au sujet des télécommunications sans fil. Il existe trois types de systèmes à accès multiple, qui se différencient principalement par la façon dont l'onde auditive est utilisée :

FDMA (“Frequency Division Multiple Access”) : chaque usager utilise une fréquence différente.

TDMA (“Time Division Multiple Access”) : une même fréquence est partagée ; les données de chacun y accèdent par petits intervalles du temps.

CDMA (“Code Division Multiple Access”) : les données utilisent la même fréquence, simultanément, mais un code propre à chaque usager permet d’en faire la distinction.

Nous allons étudier le système CDMA qui de nos jours s’avère le plus performant, mais auquel il arrive aussi d’avoir des failles, causant ainsi des désagréments aux utilisateurs.

Un réseau de téléphonie cellulaire fonctionne de la façon suivante : l’espace géographique est divisé en plusieurs cellules, souvent représentées sous forme hexagonale (voir Figure 3.1). Au centre de chaque cellule se trouve une *station de base* (ou station réceptrice). Celle-ci contient des émetteurs et des récepteurs qui connectent les téléphones mobiles au réseau.

L’*interférence* est le phénomène qui se produit lorsqu’une surcharge du réseau provoque un mixage des données causant ainsi une perte d’informations. Si le bruit dans une cellule devient supérieur à un seuil de tolérance, tous les usagers dans cette cellule perdent leur connections et leur appel est coupé. Cet événement est appelé *blocage*. Malgré que le système CDMA minimise l’interférence, il arrive que des appels soient interrompus à cause d’un excès de bruit.

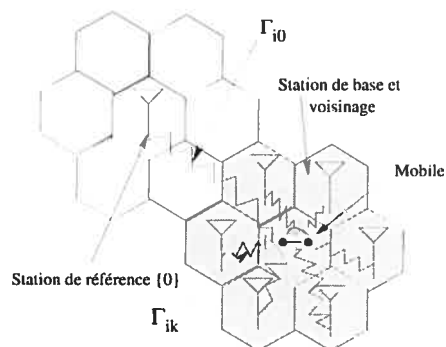


FIG. 3.1 – Mobile et stations de base dans un réseau CDMA

Le modèle que nous allons utiliser est basé sur celui de [15]. Considérons un réseau de K cellules, chacune ayant un *voisinage* de cellules adjacentes dénoté par $\mathcal{V}(k)$, $k = 1, \dots, K$. Dans la cellule k , les utilisateurs mobiles sont placés selon un processus spatial

de Poisson d'intensité $\lambda_k(\cdot)$. Nous considérons un processus de Poisson homogène dans chaque cellule, i.e., λ_k est constant pour une même cellule, mais il peut être différent pour les différentes cellules. Plus loin, dans la section 3.4.2, nous allons étudier un cas particulier où il y a seulement deux intensités différentes : une pour le voisinage de la cellule centrale et une autre pour le reste des cellules. Dans d'autres cas (cf. section 3.2 et section 3.4), pour simplifier l'exposé, nous allons considérer un modèle *homogène*, où $\lambda_k = \lambda, \forall k$.

Soit $N_k \sim \text{Poisson}(\lambda_k)$ le nombre d'utilisateurs dans la cellule k . Pour chaque mobile $i = 1, \dots, N_k$ dans la cellule k , dénotons par $d_{i,k}(l)$ la distance du mobile i jusqu'à la station de base $l \in \mathcal{V}(k)$. Le signal émis par un mobile subit une dégradation causée par des obstacles physiques et une partie de l'information est perdue durant le transfert. La perte d'intensité du signal du mobile i se situant dans la cellule $l \in \mathcal{V}(k)$ est appelé *facteur d'atténuation* à la station l et est dénoté par $\Gamma_{i,k}(l)$. Soit

$$\Gamma_{i,k}(l) = \frac{1}{d_{i,k}(l)^4} 10^{Z_{i,k}(l)}, \quad C_{i,k} = \max_{l \in \mathcal{V}(k)} \{\Gamma_{i,k}(l)\}, \quad (3.1)$$

$$\kappa(i) = \operatorname{argmax} \{\Gamma_{i,k}(l) : l \in \mathcal{V}(k)\} \quad (3.2)$$

où $\{Z_{i,k}(l)\}$ sont des variables i.i.d. $\mathcal{N}(0, \sigma^2)$. Évidemment, $\Gamma_{i,k}(l)$ dépend de la distance entre le mobile et la station de base, mais aussi du terme $10^{Z_{i,k}(l)}$ représentant l'*effet d'ombrage* qui intervient aléatoirement dans l'environnement.

L'utilisateur i sera connecté à la station $\kappa(i)$ ayant la meilleure transmission (le signal le plus fort parmi toutes les stations avoisinantes). Si un mobile émet un signal unitaire, l'effet d'ombrage fera en sorte que celui-ci sera reçu à la station de base avec une puissance $\Gamma_{i,k}(l)^{-1}$. Dans notre modèle, nous supposons qu'il y a contrôle total de la puissance émettrice, ce qui signifie que chaque utilisateur peut choisir la puissance de transmission exacte faisant en sorte qu'il soit reçu avec une puissance unitaire à la station de base [15]. Comme le facteur d'atténuation avec la station choisie est $C_{i,k}$, le mobile ajuste sa puissance de transmission à $\frac{1}{C_{i,k}}$ ce qui garantit que tous les signaux reçus à $\kappa(i)$ seront unitaires.

L'interférence causée à la station de référence 0 par le mobile i appartenant à la cellule k est $\Gamma_{i,k}(0)/C_{i,k}$. On note que si la cellule $\{0\}$ est dans le voisinage $\mathcal{V}(k)$ de la cellule k , alors tous les usagers de la cellule k vont causer une interférence à la station de référence bornée par un. Par contre, si la cellule $\{0\} \notin \mathcal{V}(k)$, il est possible que pour certains usagers i , $\Gamma_{i,0}/C_{i,k} > 1$ et en fait, cette variable aléatoire aura une queue de distribution épaisse.

Dans un réseau, la somme des interférences de tous les usagers est appelée le *rapport signal/bruit*. L'événement *blocage* survient lorsque le rapport signal/bruit mesuré à la station de référence 0 devient inférieur à un seuil α . Puisque les signaux sont toujours reçus avec puissance unitaire, ceci est équivalent à ce que toute l'interférence soit supérieure à $1/\alpha$. Si l'on somme les rapports signal/bruit des usagers de chaque cellule, on obtient l'événement suivant :

$$\mathcal{B}(\{N_k, Z_k\}_{k=1,\dots,K}) = \left\{ \sum_{k=1}^K \sum_{i=1}^{N_k} \frac{\Gamma_{i,k}(0)}{C_{i,k}} > \alpha^{-1} \right\},$$

où l'on souligne la dépendance des v.a. N_k et Z_k , $k = 1, \dots, K$. Soit

$$p = E[\mathbf{1}_{\mathcal{B}(\cdot)}], \quad (3.3)$$

où $\mathbf{1}_{\mathcal{B}(\cdot)}$ est la fonction indicatrice de l'événement $\mathcal{B}(\cdot)$. Le problème principal de ce chapitre consiste à calculer la probabilité de blocage $p = E[\mathbf{1}_{\mathcal{B}(\cdot)}]$ qui s'avère être un événement rare.

3.2 Simulation Monte-Carlo

La simulation directe permettant d'estimer p se résume dans l'algorithme suivant.

0. Initialisation : blocage = 0, $S = 0$;
1. Générer le nombre de mobiles par cellule ;
2. Placer les mobiles dans l'espace ;
3. Connecter les mobiles :

```

FOR chaque mobile  $i$  DO :
  FOR toute cellule  $j \in \mathcal{V}(k), j \neq \{0\}$ , DO
    Calculer  $\Gamma_{i,j}$  et choisir le maximum ;
  END FOR
  Calculer  $\Gamma_{i,0}$  ;
   $S+ = \Gamma_{i,0}/C_i$  ;
END FOR
4. Tester blocage : si  $(S > \alpha)$ , blocage++ ;

```

En implantant cet algorithme, nous avons constaté l'inefficacité de la simulation Monte-Carlo pour l'estimation des événements rares. Le grand nombre de variables aléatoires à générer rend la simulation très lente si elle n'est pas optimisée dans cette direction. On trouve dans ^[56] les détails d'implémentation de notre simulation et en particulier, les méthodes de génération de variables aléatoires que nous avons trouvées les plus efficaces dans notre cas. Dans notre modèle, nous avons besoin de générer des variables Poisson pour le nombre d'utilisateurs par cellule, normales pour les facteurs d'atténuation, ainsi que le placement uniforme des utilisateurs dans les cellules hexagonales. Pour toutes ces variables, nous avons utilisé les méthodes les plus appropriées, celles qui minimisent les temps de calcul. Par exemple, pour les différentes valeurs du paramètre de la distribution Poisson λ , différentes méthodes ont été employées. Malgré cela, le temps de simulation nécessaire à l'obtention d'intervalles de confiance raisonnables peut devenir très grand, comme en témoignent les résultats de la Table 3.1, où nous avons implanté le modèle homogène dans lequel $\lambda_k = \lambda$.

On remarque que la probabilité de blocage varie selon le nombre d'utilisateurs moyen par cellule (λ) et le seuil de tolérance ($1/\alpha$). Par exemple, avec $1/\alpha = 20$ et 8 mobiles/cellule (expérience 3.), $P[\text{blocage}] \approx 0.02$, alors qu'avec $1/\alpha = 20$ et 20 mobiles/cellule (expérience 4.), $P[\text{blocage}] \approx 1$. Ceci est logique, puisque plus le réseau est chargé, plus il y a risque de blocage. On remarque que le temps CPU peut aller jusqu'à 600 minutes par simulation, comme dans la première expérience où le blocage est vraiment un événement rare. Pour les simulations futures, nous allons d'utiliser les

<i>No.exp.</i>	λ	$1/\alpha$	<i>Nb.exéc.</i>	$P[\text{blocage}]$	<i>Temps CPU (min.)</i>
1.	2.0	10.0	4 800 000	0.000320 ± 0.000016	647.783333
2.	8.0	10.0	100 000	0.842360 ± 0.002259	252.516667
3.	8.0	20.0	75 000	0.020187 ± 0.001007	86.933333
4.	20.0	20.0	10 000	0.997300 ± 0.001017	10.683333

TAB. 3.1 – Simulation Monte Carlo pour l’estimation de la probabilité de blocage d’un réseau CDMA (modèle homogène). λ est le nombre moyen de mobiles par cellule, $1/\alpha$ est le seuil de tolérance pour la déclaration de blocage, le nombre d’exécutions indique le nombre de simulations indépendantes qui ont été nécessaires à l’obtention d’intervalles de confiance de niveau 90% .

paramètres de l’expérience 3.

Comme mentionné, l’inefficacité de la simulation directe est due au grand nombre de variables aléatoires à générer, mais aussi au fait que le blocage est un événement rare. En effet, si $X_i, i = 1, \dots, N$ sont des v.a. i.i.d. Bernoulli de paramètre p , alors la moyenne de l’échantillon a une erreur relative proportionnelle à $\sqrt{p(1-p)}/(Np)$, ce qui implique que le nombre de répétitions N doit être considérablement élevé si l’événement en question a une petite probabilité.

L’estimation de la probabilité d’événements rares a été le sujet de nombreuses publications récentes ^[47]. Dans certains cas, la théorie des grandes déviations peut être utilisée et il est possible de montrer qu’un changement de mesure appelé “inclinaison exponentielle” peut donner des estimateurs ayant une erreur relative bornée en p pour des valeurs fixes de N (voir ^[11] et Chapitre 8 de ^[46]), comme nous allons expliquer dans la section suivante.

3.3 Le changement de mesure

Soit X une variable aléatoire de densité $f(x)$. Supposons que nous voulons estimer

$$E_f[h(X)] = \int h(x)f(x)dx$$

où $h(x)$ est une fonction quelconque. Afin de pouvoir calculer $h(X)$, nous devons pouvoir générer des valeurs de la variable aléatoire X . Or, imaginons qu'une simulation directe de X soit inefficace, car il est difficile de générer des variables aléatoires selon la distribution $f(x)$ ou car la variance de $h(X)$ est trop grande. Alors, il serait possible d'estimer $E_f[h(X)]$ en utilisant une autre densité, $f^*(x)$, telle que lorsque $f(x) = 0$, $f^*(x) = 0$ aussi. On obtient ainsi un nouvel estimateur :

$$E_f[h(X)] = \int \frac{h(x)f(x)}{f^*(x)} f^*(x) dx = E_{f^*} \left[\frac{h(X)f(X)}{f^*(X)} \right] = E_{f^*}[h(X)L(X)],$$

où $L(x) = \frac{f(x)}{f^*(x)}$ est le *rapport de vraisemblance*.

La différence entre les deux estimateurs est au niveau de la variance :

$$\begin{aligned} \text{Var}[h(X)] &= E_f[h^2(X)] - E_f^2[h(X)] & (3.4) \\ \text{Var}[h(X)L(X)] &= E_{f^*}[h^2(X)L^2(X)] - E_{f^*}^2[h(X)L(X)] = \\ &= \int h^2(x) \frac{f^2(x)}{f^{*2}(x)} f^*(x) dx - E_f^2[h(X)] = \\ &= \int h^2(x) \frac{f(x)}{f^*(x)} f(x) dx - E_f^2[h(X)] = \\ &= E_f[h^2(X)L(X)] - E_f^2[h(X)] & (3.5) \end{aligned}$$

Les équations (3.4) et (3.5) indiquent que la variance serait réduite si $E_f[h^2(X)L(X)] < E_f[h^2(X)]$.

Dans notre cas, rappelons que nous voulons estimer la probabilité de blocage donnée par (3.3). Supposons que $X = (X_n; n = 1, 2, \dots)$ sont des v.a. i.i.d. de moyenne zéro et soit $h(X) = \sum_{i=1}^n X_i$. Le problème est l'estimation de la probabilité :

$$p_n = P \left(\frac{h(X)}{n} \geq a \right) = E(I_n(X)),$$

où $I_n(X) = \mathbf{1}_{\{h(X) \geq na\}}$, pour $a > 0$. Lorsque $n \rightarrow \infty$, cet événement devient de plus en plus rare et sa probabilité tend vers zéro. D'ailleurs, sous des conditions appropriées, celle-ci tend vers zéro de façon exponentielle : la théorie des grandes déviations ^[11] nous

permet de déterminer le taux auquel p_n décroît. Récrivons l'espérance comme suit :

$$\begin{aligned} E(I_n(X)) &= \int_{\mathbb{R}^n} \mathbf{1}_{\{\sum_{i=1}^n x_i \geq na\}} \prod_{i=1}^n f(x_i) dx \\ &= \int_{\mathbb{R}^n} \mathbf{1}_{\{\sum_{i=1}^n x_i \geq na\}} \left(\prod_{i=1}^n \frac{f(x_i)}{f_\theta(x)} \right) \prod_{i=1}^n f_\theta(x_i) dx \\ &\equiv E(L_n(\theta, \tilde{X}) I_n(\tilde{X})), \end{aligned}$$

où $\tilde{X} = (\tilde{X}_n, n = 1, 2, \dots)$ sont maintenant des v.a. i.i.d. de densité f_θ , telle que lorsque $f(x) = 0$, $f_\theta(x) = 0$. La v.a. $L_n(\theta, X)$ est aussi appelée *dérivée de Radon Nikodym*.

Considérons la variance de ce nouvel estimateur. Puisque $E[L_n(\theta, \tilde{X}) I_n(\tilde{X})] = p_n$, on a :

$$\text{Var}[L_n(\theta, \tilde{X}) I_n(\tilde{X})] = E[L_n^2(\theta, \tilde{X}) I_n(\tilde{X})] - p_n^2 \geq 0,$$

ce qui implique (la variance étant toujours non négative) :

$$\frac{1}{n} \log \left(E[L_n^2(\theta, \tilde{X}) I_n(\tilde{X})] \right) \geq \frac{2}{n} \log p_n.$$

L'idée est d'utiliser le Théorème de Cramer, qui stipule que

$$\liminf_{n \rightarrow \infty} \frac{2}{n} \log p_n = -2 \left(\sup_{\theta} \{a\theta - \log M(\theta)\} \right),$$

où $M(\theta) = E[e^{\theta X}]$ est la fonction génératrice des moments de X_i . Le suprémum du côté droit peut facilement être calculé, puisque la fonction est concave en θ et a donc un seul point critique θ^* :

$$\left. \frac{d}{d\theta} (a\theta - \log M(\theta)) \right|_{\theta=\theta^*} = a - \frac{M'(\theta^*)}{M(\theta^*)} = 0.$$

Par conséquent, on obtient une borne inférieure de l'espérance asymptotique :

$$\text{Var}[L_n(\theta, \tilde{X}) I_n(\tilde{X})] \geq 0 \implies$$

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \left(E[L_n^2(\theta, \tilde{X}) I_n(\tilde{X})] \right) \geq -2(a\theta^* - \log M(\theta^*)). \quad (3.6)$$

Considérons la nouvelle densité $f_\theta(x) = e^{\theta x} f(x)/M(\theta)$, appelée *inclinaison exponentielle* ^[46] de f . La dérivée Radon Nikodym est simplement l'expression :

$$L_n(\theta, X) = \prod_{i=1}^n e^{-\theta X_i} M(\theta) = M^n(\theta) e^{-\theta h(X)}.$$

En utilisant la définition de $I_n(\tilde{X})$, $E[L_n^2(\theta, \tilde{X})I_n(\tilde{X})] \leq M^{2n}(\theta)e^{-2\theta na}$, alors en divisant par n et prenant le logarithme, on obtient :

$$\frac{1}{n} \log \left(E[L_n^2(\theta, \tilde{X})I_n(\tilde{X})] \right) \leq -2a\theta + 2 \log M(\theta), \quad (3.7)$$

ce qui nous donne une borne supérieure du taux de décroissance asymptotique. En comparant les inégalités (3.6) et (3.7), il s'ensuit que la variance asymptotique atteint sa valeur minimale en $\theta = \theta^*$.

Malheureusement, une application directe de cette théorie nécessiterait l'inclinaison d'une variable aléatoire à queue de distribution épaisse, i.e., l'inclinaison de la variable lognormale apparaissant dans le facteur d'atténuation (3.1), pour laquelle la fonction génératrice n'existe pas.

Le changement de mesure que nous implémentons utilise un nombre d'utilisateurs Poisson de moyenne différente de l'originale et une inclinaison exponentielle des facteurs d'atténuation $Z_{i,k}(0)$. Ceci nous évite la transformation des termes problématiques $\Gamma_{i,k}(l)$ ou $\Gamma_{i,k}(l)/C_{i,k}$.

Lemme 2 (Changement de mesure) Soit $N = (N_1, \dots, N_K)$, où les $\{N_k\}$ sont des v.a. i.i.d. Poisson de moyenne λ_k , et soit $Z = (Z_{1,k}(0), \dots, Z_{N_k,k}(0))$, où les $\{Z_{i,k}(0)\}$ sont des v.a. i.i.d. $\mathcal{N}(0, \sigma^2)$, $i = 1, \dots, N_k$, $k = 1, \dots, K$, conditionnellement indépendantes de N . Pour toute valeur des paramètres $(\theta_k, \mu_k) \in \mathbb{R}^{2K}$, $k = 1, \dots, K$ et pour toute fonction réelle $\phi(N; Z)$:

$$E[\phi(N; Z)] = E[L(N^*; Z^*)\phi(N^*; Z^*)], \quad (3.8)$$

où $N_k^* \sim \text{Poisson}(\theta_k)$, $Z_{i,k}^* \sim \mathcal{N}(\mu_k, \sigma^2)$ sont des variables indépendantes, et

$$L(N; Z) = \exp \left\{ \sum_{k=1}^K \left[(\theta_k - \lambda_k) + \frac{N_k \mu_k^2}{2\sigma^2} - \frac{\mu_k}{\sigma^2} B_{N_k, k} \right] \right\} \prod_{k=1}^K \left(\frac{\lambda_k}{\theta_k} \right)^{N_k} \quad (3.9)$$

avec $B_{N_k, k} = \sum_{i=1}^{N_k} Z_{i,k}^*$.

Preuve: Voir Annexe I. □

Le résultat ci-haut suggère la procédure de simulation suivante : K variables aléatoires $N_k^* \sim \text{Poisson}(\theta_k)$ sont générées et pour chaque k , les variables $\xi_{i,k}(l) \sim \mathcal{N}(0, 1)$, $i = 1, \dots, N_k$, $l \in \mathcal{V}(k) \cup \{0\}$, sont générées indépendamment. Ainsi, $Z_{i,k}^*(l) = \sigma \xi_{i,k}(l) + \mu_k$ et les facteurs d'atténuation sont définis comme suit :

$$\Gamma_{i,k}^*(l) = \begin{cases} \Gamma_{i,k}(l) & l \neq 0, \\ 10^{\mu_k} \Gamma_{i,k}(l) & l = 0, \end{cases} \quad C_{i,k}^* = \max_{l \in \mathcal{V}(k)} (\Gamma_{i,k}^*(l)).$$

En utilisant la transformation $Z_{i,k}^*(0) = \mu_k + Z_{i,k}(0)$, le terme $\frac{\mu_k}{\sigma^2} B_{N_k, k}$ dans (3.9) devient sous la mesure étoilée :

$$(\mu_k / \sigma^2) B_{N_k^*, k} = N_k^* (\mu_k^2 / \sigma^2) + (\mu_k / \sigma) \sum_{i=1}^{N_k^*} \xi_{i,k}(0).$$

Finalement, posons $N^* = (N_1^*, \dots, N_K^*)$ et $Z^* = (Z_{1,k}^*(0), \dots, Z_{N_k, k}^*(0))$. Un estimateur non biaisé de la probabilité de blocage à la station $\{0\}$ s'écrit alors :

$$\begin{aligned} \hat{p}(\theta, \mu) &= \exp \left\{ \sum_{k=1}^K \left((\theta_k - \lambda_k) - \frac{N_k^* \mu_k^2}{2\sigma^2} - \frac{\mu_k}{\sigma} \sum_{i=1}^{N_k^*} \xi_{i,k}(0) \right) \right\} \prod_{k=1}^K \left(\frac{\lambda_k}{\theta_k} \right)^{N_k^*} \\ &\quad \times \mathbf{1} \left(\sum_{k=1}^K \sum_{i=1}^{N_k^*} \frac{10^{\mu_k} \Gamma_{i,k}^*(0)}{C_{i,k}^*} > \alpha^{-1} \right) \\ &= L(N^*, Z^*) \mathbf{1}_{B(N^*, Z^*)}. \end{aligned} \quad (3.10)$$

L'efficacité d'une telle estimation est mesurée en termes d'effort de calcul nécessaire pour obtenir un niveau de précision relative. Dans notre cas, améliorer l'efficacité est équivalent à trouver les valeurs de θ_k, μ_k qui *minimisent* la variance $\text{Var}[\hat{p}(\theta, \mu)]$. Puisque $\text{Var}(\hat{p}(\theta, \mu)) = \text{E}(\hat{p}^2(\theta, \mu)) - p^2$ et p est indépendant de $(\theta, \mu) \in \mathbb{R}^{2K}$, il suffit de trouver les valeurs de θ et μ minimisant $\text{E}[\hat{p}(\theta, \mu)^2]$.

3.4 Mécanisme d'autoréglage

Pour simplifier l'exposé, nous allons considérer dans cette section le modèle homogène, où $\lambda_k = \lambda$ pour toutes les cellules.² La notation sera aussi simplifiée en posant

$$M = \sum_{k=1}^K N_k$$

et en numérotant tous les mobiles $i = 1, \dots, M$, peu importe la cellule dans laquelle ils se trouvent. Ainsi, M est une v.a. Poisson de moyenne $K\lambda$. Le changement de mesure sera aussi homogène : $\theta_k = \theta, \mu_k = \mu$ (La section 3.4.2 traitera d'un changement de mesure stratifié). Nous cherchons à minimiser sur (θ, μ) la fonction

$$J(\theta, \mu) := \text{E}[\hat{p}^2(\theta, \mu)] = \text{E}[L^2(M^*, Z^*)\mathbf{1}_{B(M^*, Z^*)}], \quad (3.11)$$

où $M^* \sim \text{Poisson}(K\theta), Z^* \sim \mathcal{N}(\mu, \sigma^2)$ et $\hat{p}(\cdot)$ est donné par l'équation (3.10).

L'idée de la simulation intelligente est d'appliquer l'algorithme d'approximation stochastique comme dans ^[57], pour ajuster les valeurs de θ et μ dans la direction de la variance décroissante. Soit :

$$g(\theta, \mu) := \nabla_{\theta, \mu} \text{E}[L^2(M^*; Z^*)\mathbf{1}_{B(M^*; Z^*)}] = \nabla_{\theta, \mu} J(\theta, \mu), \quad (3.12)$$

et supposons que $(\hat{g}(\theta_n), \hat{g}(\mu_n))$ est un estimateur non biaisé du gradient obtenu avec $S_n > 0$ simulations indépendantes des variables M et Z , les paramètres ayant la valeur

²Le modèle où $\lambda_k = \lambda$ est moins général que celui où chaque cellule a sa propre intensité, mais les idées développées ici sont directement applicables au modèle général.

(θ_n, μ_n) . Considérons la récursion :

$$\begin{aligned}\theta_{n+1} &= \theta_n - \epsilon_n \hat{g}(\theta_n) \\ \mu_{n+1} &= \mu_n - \epsilon_n \hat{g}(\mu_n).\end{aligned}\tag{3.13}$$

Si la fonction $J(\cdot)$ est convexe et si les conditions du Théorème 1 (Chapitre 1) sont vérifiées, alors (3.13) converge vers le seul point de stabilité tel que $g(\theta, \mu) = 0$, i.e., $\theta_n \rightarrow \theta^*$ et $\mu_n \rightarrow \mu^*$ a.p.1.

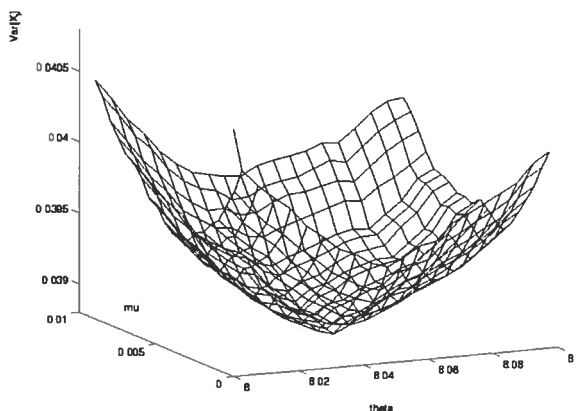


FIG. 3.2 – Courbe de la variance de l'estimateur de la probabilité de blocage $\hat{p}(\theta, \mu)$ (3.10) en fonction des paramètres du changement de mesure μ et θ , obtenue par estimation fonctionnelle avec nombres aléatoires communs.

Pour vérifier l'hypothèse de convexité, nous avons utilisé la méthode d'estimation fonctionnelle ^[37] afin de visualiser la courbe de $J(\theta, \mu)$. Les détails à ce sujet se trouvent dans ^[56]. Le résultat de l'estimation fonctionnelle est illustré à la figure 3.2. Celle-ci laisse croire que $J(\cdot)$ est convexe en ses deux arguments θ et μ , ce qui suggère l'existence d'un minimum global unique. Puisque l'estimation fonctionnelle n'est pas abordée dans cet ouvrage, la section suivante donne une preuve alternative de la convexité de $J(\cdot)$.

3.4.1 Calcul des gradients

Si la fonction $L^2(M^*; Z^*)\mathbf{1}_{\mathcal{B}(M^*, Z^*)}$ est Lipschitz continue p.s. avec constante de Lipschitz intégrable, alors la dérivée et l'espérance dans Éq. (3.12) peuvent être interchangées et le gradient stochastique résultant sera non biaisé pour $\nabla J(\cdot)$ [37]. Cependant, dans notre problème, $L^2(M^*; Z^*)\mathbf{1}_{\mathcal{B}(M^*, Z^*)}$ a des discontinuités, causées par les sauts possibles de l'événement de blocage et du terme M^* , qui peuvent sauter lorsque les valeurs de θ et μ changent. Pour contourner cette difficulté et pour éviter les méthodes d'estimation du gradient qui traitent les discontinuités, Su et Fu [50] proposent de revenir à la mesure originale, afin de calculer le gradient, ce que nous allons faire.

Pour le cas homogène, éq. (3.9) devient

$$L(M^*; Z^*) = \exp \left\{ K(\theta - \lambda) + M^* \frac{\mu^2}{2\sigma^2} - \frac{\mu}{\sigma^2} \sum_{i=1}^{M^*} Z_i^*(0) \right\} \left(\frac{\lambda}{\theta} \right)^{M^*}. \quad (3.14)$$

Le blocage sous la mesure étoilée s'écrit

$$\mathcal{B}(M^*; Z^*) = \mathbf{1} \left(\sum_{i=1}^{M^*} \frac{\Gamma_i^*(0)}{C_i^*} > \alpha^{-1} \right),$$

où $M^* \sim \text{Poisson}(K\theta)$, et $Z_i^*(0), i = 1, \dots, M^*$ sont i.i.d. $\mathcal{N}(\mu, \sigma^2)$. Comme avant, soit $B_{M^*}^* = \sum_{i=1}^{M^*} Z_i^*(0)$.

Il faut remarquer que $L(\cdot)$ et $\mathcal{B}(\cdot)$ sont des fonctions de θ et μ , mais cette dépendance est en quelque sorte implicite, car θ et μ sont les paramètres des lois de distribution des v.a. M^* et Z^* . Cependant, puisque θ et μ sont les paramètres de contrôle par rapport auxquels nous allons dériver, nous allons expliciter leur dépendance dans cette section, afin de rendre la notation plus claire. Notons alors que $L(M^*; Z^*) = L(M^*; Z^*; \theta, \mu)$ et $\mathcal{B}(M^*; Z^*) = \mathcal{B}(M^*; Z^*; \theta, \mu)$.

Lemme 3 *Sous le changement de mesure, les estimateurs*

$$\hat{g}(\theta) = \left(K - \frac{M^*}{\theta} \right) L^2(M^*; Z^*; \theta, \mu) \mathbf{1}_{B(M^*; Z^*; \theta, \mu)} \quad (3.15)$$

$$\hat{g}(\mu) = \left(\frac{M^* \mu - B_{M^*}}{\sigma^2} \right) L^2(M^*; Z^*; \theta, \mu) \mathbf{1}_{B(M^*; Z^*; \theta, \mu)} \quad (3.16)$$

sont des estimateurs non biaisés des dérivées de $J(\theta, \mu)$ par rapport à θ et μ , respectivement.

Preuve: (Le schéma de cette preuve suit celui de Su et Fu ^[50].) En appliquant Lemme 2 avec $\phi(M; Z) = L(M; Z; \theta, \mu) \mathbf{1}_{B(M; Z; \theta, \mu)}$, on peut écrire $J(\cdot)$ et $g(\cdot)$ en termes de la mesure originale :

$$\begin{aligned} J(\theta, \mu) &= E[L^2(M^*; Z^*; \theta, \mu) \mathbf{1}_{B(M^*; Z^*; \theta, \mu)}] = E[L(M; Z; \theta, \mu) \mathbf{1}_{B(M; Z; \theta, \mu)}] \\ g(\theta, \mu) &= \nabla_{\theta, \mu} E[L(M; Z; \theta, \mu) \mathbf{1}_{B(M; Z; \theta, \mu)}]. \end{aligned}$$

Le but de cette transformation est d'éviter la dépendance de θ dans les discontinuités de $L^2(M^*; Z^*; \theta, \mu) \mathbf{1}_{B(M^*; Z^*; \theta, \mu)}$. Ainsi, sous la mesure originale, M et $\mathbf{1}_{B(M; Z; \theta, \mu)}$ sont indépendantes de θ et μ . Par conséquent, on peut écrire :

$$\begin{aligned} \frac{\partial}{\partial \theta} J(\theta, \mu) &= E \left[\frac{\partial}{\partial \theta} L(M; Z; \theta, \mu) \mathbf{1}_{B(M; Z; \theta, \mu)} \right] \text{ et} \\ \frac{\partial}{\partial \mu} J(\theta, \mu) &= E \left[\frac{\partial}{\partial \mu} L(M; Z; \theta, \mu) \mathbf{1}_{B(M; Z; \theta, \mu)} \right], \end{aligned}$$

où $M \sim \text{Poisson}(K\lambda)$ et $Z_i(0), i = 1, \dots, M$ sont i.i.d. $\mathcal{N}(0, \sigma^2)$. Pour tout M et tout Z , la fonction $L(M; Z; \theta, \mu)$ est différentiable en θ et μ , et son gradient a une espérance uniformément bornée sur tout ensemble compact (en θ, μ). On a :

$$\frac{\partial}{\partial \theta} L(M; Z; \theta, \mu) = \left(K - \frac{M}{\theta} \right) L(M; Z; \theta, \mu) \quad (3.17)$$

$$\frac{\partial}{\partial \mu} L(M; Z; \theta, \mu) = \left(\frac{M\mu - B_M}{\sigma^2} \right) L(M; Z; \theta, \mu) \quad (3.18)$$

où $B_M = \sum_{i=1}^M Z_i(0)$.

Pour montrer que ces dérivées ont des espérances bornées sur des ensembles compacts en (θ, μ) , on utilise le fait qu'une v.a. Poisson possède une fonction génératrice des moments finie. Par souci d'espace, seulement le résultat pour (3.17) est démontré ici. Des arguments similaires s'appliquent à (3.18).

Il est clair que $(K - M/\theta)L(\cdot)\mathbf{1}_{B(\cdot)}$ est borné p.s. par $(K - M/\theta)L(\cdot)$, donc il suffit de montrer que cette v.a. a une espérance bornée sur tout ensemble θ -compact. Puisque la mesure originale est utilisée, $M \sim \text{Poisson}(K\lambda)$ et $Y(M) = (\mu/\sigma^2)B_M$, $Y(M) \sim \mathcal{N}(0, M(\mu^2/\sigma^2))$. En conditionnant sur M et en utilisant $E(e^{-Y} | M) = e^{-M\mu^2/\sigma^2}$, il s'ensuit que

$$E \left[\left(K - \frac{M}{\theta} \right) L(M; Z; \theta, \mu) \right] = E \left[\left(K - \frac{M}{\theta} \right) e^{K(\theta-\lambda)} \left(\frac{\lambda}{\theta} \right)^M \right],$$

où la dernière espérance est par rapport à la v.a. Poisson M . Pour de telles variables, la fonction génératrice des moments $\phi(a) = E[e^{aM}]$ existe et

$$E \left[-\frac{M}{\theta} \left(\frac{\lambda}{\theta} \right)^M \right] = \sum_{m \geq 0} -\frac{m}{\theta} \left(\frac{\lambda}{\theta} \right)^m \frac{\lambda^m}{m!} = \frac{\partial}{\partial \theta} \sum_{m \geq 0} \left(\frac{\lambda}{\theta} \right)^m \frac{\lambda^m}{m!} = \frac{\partial}{\partial \theta} \phi[\ln(\lambda/\theta)] < \infty,$$

ce qui prouve que l'estimateur a effectivement une espérance uniformément bornée.

Ce résultat démontre que les estimateurs (3.17) et (3.18) sont non biaisés et peuvent être utilisés dans l'approximation stochastique (3.13). La simulation étant faite avec le changement de mesure, c'est-à-dire, M^* étant généré selon une distribution Poisson($K\theta$), Su et Fu ^[50] proposent de changer la mesure encore une fois, cette fois-ci pour les gradients stochastiques (3.17) et (3.18), ce qui établit le résultat du Lemme :

$$\begin{aligned} g(\theta) &= E \left[\left(K - \frac{M^*}{\theta} \right) L^2(M^*; Z^*; \theta, \mu) \mathbf{1}_{B(M^*; Z^*; \theta, \mu)} \right] \\ g(\mu) &= E \left[\left(\frac{M^* \mu - B_{M^*}}{\sigma^2} \right) L^2(M^*; Z^*; \theta, \mu) \mathbf{1}_{B(M^*; Z^*; \theta, \mu)} \right]. \end{aligned}$$

□

De la même façon que nous venons de trouver les dérivées premières de $J(\cdot)$, il est possible de calculer la matrice hessienne, qui nous indiquera si la fonction est convexe ou pas. Pour la preuve de la convexité, nous aurons besoin de la proposition suivante tirée de [38].

Proposition 2 Une matrice carrée symétrique d'ordre 2, $A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$, est définie positive si, et seulement si les éléments diagonaux a et d sont positifs et $\det(A) > 0$.

Lemme 4 La fonction $J(\theta, \mu)$ est convexe.

Preuve: Pour montrer que la fonction objectif $J(\theta, \mu)$ est convexe, nous allons démontrer que sa matrice hessienne $H(\theta, \mu)$ (la matrice des dérivées secondes) est définie positive. En revenant à la mesure originale comme indiqué dans la preuve du Lemme 3, rappelons que :

$$\begin{aligned} J(\theta, \mu) &= \mathbb{E}[L(M; Z; \theta, \mu)\mathbf{1}_{\mathcal{B}(M; Z; \theta, \mu)}], \\ \nabla J(\theta, \mu) &= \left(\mathbb{E} \left[\frac{\partial}{\partial \theta} L(M; Z; \theta, \mu)\mathbf{1}_{\mathcal{B}(M; Z; \theta, \mu)} \right], \mathbb{E} \left[\frac{\partial}{\partial \mu} L(M; Z; \theta, \mu)\mathbf{1}_{\mathcal{B}(M; Z; \theta, \mu)} \right] \right). \end{aligned}$$

Les quatre composantes de la matrice hessienne sont :

$$\frac{\partial^2}{\partial \theta^2} J(\theta, \mu) = \frac{\partial}{\partial \theta} \mathbb{E} \left[\frac{\partial}{\partial \theta} L(M; Z; \theta, \mu)\mathbf{1}_{\mathcal{B}(M; Z; \theta, \mu)} \right] \quad (3.19)$$

$$\frac{\partial^2}{\partial \mu^2} J(\theta, \mu) = \frac{\partial}{\partial \mu} \mathbb{E} \left[\frac{\partial}{\partial \mu} L(M; Z; \theta, \mu)\mathbf{1}_{\mathcal{B}(M; Z; \theta, \mu)} \right] \quad (3.20)$$

$$\frac{\partial}{\partial \mu} \frac{\partial}{\partial \theta} J(\theta, \mu) = \frac{\partial}{\partial \theta} \frac{\partial}{\partial \mu} J(\theta, \mu) = \frac{\partial}{\partial \theta} \mathbb{E} \left[\frac{\partial}{\partial \mu} L(M; Z; \theta, \mu)\mathbf{1}_{\mathcal{B}(M; Z; \theta, \mu)} \right] \quad (3.21)$$

En utilisant le fait que les dérivées premières ont des espérances bornées sur tout ensemble compact pour échanger de nouveau la dérivée et l'espérance dans (3.19)-

(3.21), on calcule les dérivées secondes de $L(\cdot)$:

$$\begin{aligned}\frac{\partial^2}{\partial \theta^2} L(M; Z; \theta, \mu) &= \left(\left(K - \frac{M}{\theta} \right)^2 + \frac{M}{\theta^2} \right) L(M; Z; \theta, \mu) \\ \frac{\partial^2}{\partial \mu^2} L(M; Z; \theta, \mu) &= \left(\left(\frac{M\mu - B_M}{\sigma^2} \right)^2 + \frac{M}{\sigma^2} \right) L(M; Z; \theta, \mu) \\ \frac{\partial}{\partial \mu} \frac{\partial}{\partial \theta} L(M; Z; \theta, \mu) &= \frac{\partial}{\partial \theta} \frac{\partial}{\partial \mu} L(M; Z; \theta, \mu) = \left(K - \frac{M}{\theta} \right) \left(\frac{M\mu - B_M}{\sigma^2} \right) L(M; Z; \theta, \mu).\end{aligned}$$

À l'aide de ces dérivées, on écrit la matrice hessienne $H(\theta, \mu)$:

$$H(\theta, \mu) = \begin{pmatrix} \mathbb{E} \left[\frac{\partial^2}{\partial \theta^2} L(\cdot) \mathbf{1}_{\mathcal{B}(\cdot)} \right] & \mathbb{E} \left[\frac{\partial}{\partial \theta} \frac{\partial}{\partial \mu} L(\cdot) \mathbf{1}_{\mathcal{B}(\cdot)} \right] \\ \mathbb{E} \left[\frac{\partial}{\partial \mu} \frac{\partial}{\partial \theta} L(\cdot) \mathbf{1}_{\mathcal{B}(\cdot)} \right] & \mathbb{E} \left[\frac{\partial^2}{\partial \mu^2} L(\cdot) \mathbf{1}_{\mathcal{B}(\cdot)} \right] \end{pmatrix}.$$

Réécrivons $H(\theta, \mu)$ sous la forme compacte suivante :

$$H(\theta, \mu) = \mathbb{E} \left[Q(\theta, \mu) L(M; Z; \theta, \mu) \mathbf{1}_{\mathcal{B}(M; Z; \theta, \mu)} \right],$$

$$\text{où } Q(\theta, \mu) = \begin{pmatrix} A^2 + \frac{M}{\theta^2} & AB \\ AB & B^2 + \frac{M}{\sigma^2} \end{pmatrix}, \quad A = K - \frac{M}{\theta} \text{ et } B = \frac{1}{\sigma^2}(M\mu - B_M).$$

Il est facile de voir que $H(\theta, \mu)$ sera définie positive si, et seulement si $Q(\theta, \mu)$ est définie positive (car $L(\cdot) > 0$, $\forall M, Z, \theta, \mu$ et l'espérance d'une v.a. positive est positive).

Ainsi, vérifions les conditions de la Proposition 2, c'est-à-dire que les éléments diagonaux et le déterminant de $Q(\cdot)$ sont positifs. On observe que $A^2 + \frac{M}{\theta^2} > 0$ et $B^2 + \frac{M}{\sigma^2} > 0$, car la v.a. $M \sim \text{Poisson}(K\lambda)$ est non-négative. De même,

$$\begin{aligned}\det \begin{pmatrix} A^2 + \frac{M}{\theta^2} & AB \\ AB & B^2 + \frac{M}{\sigma^2} \end{pmatrix} &= \left(A^2 + \frac{M}{\theta^2} \right) \left(B^2 + \frac{M}{\sigma^2} \right) - (AB)^2 \\ &= \frac{A^2 M}{\sigma^2} + \frac{B^2 M}{\theta^2} + \frac{M^2}{\sigma^2 \theta^2} > 0.\end{aligned}$$

Donc, par la Proposition 2, $Q(\theta, \mu)$ est définie positive. Par conséquent, la matrice hes-

sienne $H(\theta, \mu)$ est aussi définie positive et $J(\cdot)$ est convexe.

□

Nous venons de montrer que la fonction objectif $J(\theta, \mu)$ est convexe et que les gradients stochastiques $\hat{g}(\theta)$ et $\hat{g}(\mu)$ sont non biaisés, ce qui signifie qu'il existe un minimum global unique et l'algorithme d'approximation stochastique va converger presque sûrement vers ce minimum, si la suite des gains satisfait les conditions (A3) et (A4) du Théorème 1, Chapitre 1.

3.4.2 Changement de mesure stratifié

Afin d'obtenir une meilleure performance, nous allons utiliser un changement de mesure différent pour les usagers à l'intérieur et à l'extérieur du voisinage de la cellule centrale. Ainsi, on dénotera par (θ_1, μ_1) et (θ_2, μ_2) les paramètres des cellules en dedans et en dehors du voisinage $\mathcal{V}(0)$, respectivement. Il y a plusieurs motivations pour utiliser une telle stratification : une meilleure performance peut être obtenue à mesure que la dimension des variables de contrôle augmente ; généralement, le temps de simulation augmente avec la dimension, alors plus de stratification ne serait pas utile. Pour ces raisons, nous avons décidé de stratifier le réseau de cette manière, mais notre méthode peut facilement être adaptée à d'autres types de réseaux.

Voici les détails de calcul des gradients avec la stratification. Le rapport de vraisemblance peut être réécrit en termes de voisinage comme suit :

$$L(M^*, Z^*, \theta, \mu) = L_1(M_1^*, Z_1^*, \theta_1, \mu_1) \times L_2(M_2^*, Z_2^*, \theta_2, \mu_2)$$

$$L_1(M_1^*, Z_1^*, \theta_1, \mu_1) = \exp \left\{ 7(\theta_1 - \lambda) + \frac{M_1^* \mu_1^2}{2\sigma^2} - \frac{\mu_1}{\sigma^2} \sum_{i=1}^{M_1^*} Z_i(0) \right\} \left(\frac{\lambda}{\theta_1} \right)^{M_1^*}$$

$$L_2(M_2^*, Z_2^*, \theta_2, \mu_2) = \exp \left\{ (K - 7)(\theta_2 - \lambda) + \frac{M_2^* \mu_2^2}{2\sigma^2} - \frac{\mu_2}{\sigma^2} \sum_{i=1}^{M_2^*} Z_i(0) \right\} \left(\frac{\lambda}{\theta_2} \right)^{M_2^*}$$

où toutes les variables d'indice 1 sont celles associées au voisinage $\mathcal{V}(0)$, alors que celles d'indice 2, sont les autres. Remarquer que puisque les cellules sont hexagonales, il y

a sept cellules faisant partie du voisinage de la cellule centrale et $(K - 7)$ cellules à l'extérieur. Si on dérive chacune des composantes ci-haut par rapport aux deux paramètres d'intérêt, en passant encore une fois par la mesure originale (voir Section 3.4.1), on obtient les estimateurs non biaisés suivants :

$$g(\theta_1) = E \left[\left(7 - \frac{M_1^*}{\theta_1} \right) L_1^2(\cdot) L_2^2(\cdot) \mathbf{1}_{B(\cdot)} \right] \quad (3.22)$$

$$g(\theta_2) = E \left[\left((K - 7) - \frac{M_2^*}{\theta_2} \right) L_1^2(\cdot) L_2^2(\cdot) \mathbf{1}_{B(\cdot)} \right] \quad (3.23)$$

$$g(\mu_1) = E \left[\left(\frac{M_1^* \mu_1 - B_{M_1^*}^*}{\sigma^2} \right) L_1^2(\cdot) L_2^2(\cdot) \mathbf{1}_{B(\cdot)} \right] \quad (3.24)$$

$$g(\mu_2) = E \left[\left(\frac{M_2^* \mu_2 - B_{M_2^*}^*}{\sigma^2} \right) L_1^2(\cdot) L_2^2(\cdot) \mathbf{1}_{B(\cdot)} \right] \quad (3.25)$$

où $B_{M_1^*}^* = \sum_{i=1}^{M_1^*} Z_i^*(0)$ et $B_{M_2^*}^* = \sum_{i=1}^{M_2^*} Z_i^*(0)$.

Ainsi, au lieu d'utiliser la récursion (3.13) pour trouver les paramètres optimaux, nous allons considérer deux récursions, une pour les cellules dans le voisinage de 0 et une pour les autres cellules :

$$\begin{aligned} \theta_{1,n+1} &= \theta_{1,n} - \epsilon_{1,n} \hat{g}(\theta_{1,n}) \\ \mu_{1,n+1} &= \mu_{1,n} - \epsilon_{1,n} \hat{g}(\mu_{1,n}) \\ \theta_{2,n+1} &= \theta_{2,n} - \epsilon_{2,n} \hat{g}(\theta_{2,n}) \\ \mu_{2,n+1} &= \mu_{2,n} - \epsilon_{2,n} \hat{g}(\mu_{2,n}) \end{aligned} \quad (3.26)$$

La section suivante compare la performance du modèle stratifié avec celle du modèle non stratifié.

3.5 Résultats

La simulation intelligente utilise les observations du processus sous l'échantillonnage stratégique pour estimer la probabilité de blocage, en même temps que le gradient de la variance, pour ajuster les paramètres du changement de mesure en ligne. La figure 3.3 montre les trajectoires des paramètres (θ, μ) avec et sans stratification. La ligne

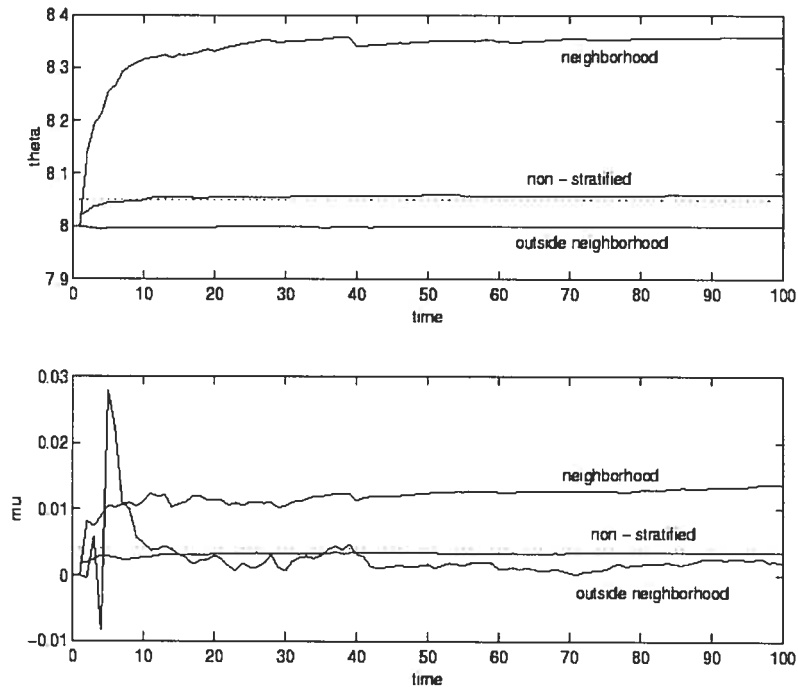


FIG. 3.3 – Valeurs de θ et μ en fonction du temps, résultat de l’approximation stochastique (3.13). La courbe “non-stratified” correspond au modèle sans stratification (la valeur optimale des paramètres pour ce modèle est indiquée par la ligne pointillée). Les courbes “neighborhood” et “outside neighborhood” sont associées aux usagers à l’intérieur et à l’extérieur du voisinage de la cellule centrale 0 respectivement.

pointillée indique la valeur optimale obtenue à l’aide de l’estimation fonctionnelle et correspond au point minimal du graphe de la figure 3.2. Il n’est pas surprenant que la stratification ne change presque pas la mesure des usagers à l’extérieur du voisinage de la cellule centrale, alors que ceux à l’intérieur, voient leur intensité augmentée et leur effet d’ombrage dévié. On remarque aussi que sans stratification, la variance est très peu améliorée (à peine 1% d’amélioration). Par contre, la stratification donne une plus grande sensibilité au modèle et la variance est diminuée de 15%. Quant au temps de simulation, la simulation intelligente a requis un total de 5,1 minutes, alors que l’estimation fonctionnelle qui a produit la figure 3.2 a nécessité 25,75 minutes. L’approximation stochastique s’avère donc une méthode d’ajustement efficace pour ce modèle.

CHAPITRE 4

APPLICATION AU CONTRÔLE DE FLOT DANS LES RÉSEAUX SANS FILS

Les avancées récentes en technologie cellulaire sont en grande partie dues au développement de couches physiques intelligentes. Désormais, des systèmes adaptatifs opérant au niveau de la station de base et au niveau de l'équipement usager peuvent mesurer les caractéristiques du canal radio et basés sur sa qualité, ils peuvent sélectionner un codage et une modulation qui optimisent la capacité de transmission. Ainsi, les méthodes adaptatives permettent d'atteindre des taux de transmission élevés.

Une station de base intelligente utilise des indicateurs pour décider à quel usager transmettre et quel schéma de modulation utiliser. Cependant, un choix optimal peut seulement être fait si les usagers ont des paquets en attente à la station de base. La longueur de la file d'attente est donc d'un enjeu important : si elle est trop petite, les fluctuations du canal de transmission sans fil vont faire en sorte qu'elle se videra souvent, ce qui empêche le fonctionnement optimal des couches inférieures. Par contre, une file trop longue causera moins de vidage, mais des délais plus grands et des pertes de paquets seront observés plus souvent.

L'application décrite dans ce chapitre¹ concerne le contrôle en ligne du nombre de paquets dans la file d'attente à la station de base, contrôle qui permettra la transmission optimale des paquets à travers le canal radio. L'approximation stochastique donne lieu à une procédure intelligente d'autorégulation, capable d'ajuster en ligne les paramètres d'un algorithme conçu pour le contrôle de la file d'attente.

CLAMP^[3] est un algorithme applicable par dessus la procédure de contrôle de flot TCP et vise l'amélioration de celle-ci lorsqu'elle opère sur un réseau sans fils. Son but est d'ajuster explicitement la fenêtre de réception AWND afin de maintenir une longueur ciblée de la file et de différencier la capacité du canal entre usagers et services.

¹Ce projet a été réalisé en collaboration avec Rami Mukhtar (Ph.D.) de l'Université de Melbourne, Australie, en 2003-2004.

Des expériences ont démontré que lorsque le taux de transmission est aléatoire, la performance de CLAMP est très sensible à un paramètre. Ainsi, l'objectif principal de l'étude qui suit est de trouver la valeur optimale de ce paramètre tel que l'utilisation du lien sans fil et le nombre de paquets en attente soient maintenus à des niveaux cibles.

La particularité de ce travail se résume en deux points saillants : l'étude d'un système avec délai qui complique considérablement l'analyse, et l'utilisation d'un modèle fluide pour l'obtention des estimateurs des gradients utilisés dans l'approximation stochastique. Le modèle fluide stochastique d'un système à événements discrets (DES) est une approximation qui au lieu de considérer les événements un à un, considère des "taux instantanés", comme si la dynamique du système était dictée par celle d'un fluide. Ce modèle fluide permet de calculer très facilement des dérivées stochastiques de mesures de performance qui sont ensuite utilisées avec le DES. Des preuves de la validité d'une telle approximation n'existent pas encore, mais le modèle fluide stochastique est de plus en plus courant dans la littérature [14, 58, 59, 60].

Avant de décrire le modèle fluide pour le système de télécommunications (section 4.2.2), nous allons introduire quelques notions de base sur la procédure de contrôle de flot TCP (section 4.1) afin de motiver l'utilisation de CLAMP. Ensuite, CLAMP sera décrit dans la section 4.2.1, avant de poser le problème d'optimisation associé (section 4.3). La section 4.4 donne le calcul des dérivées IPA et la section 4.5 décrit la procédure d'autoréglage utilisant l'approximation stochastique.

4.1 Contrôle de flot : TCP

La procédure de contrôle de flot la plus utilisée de nos jours est le *Protocole de Contrôle de Transmission* ("Transmission Control Protocol"), connu sous l'acronyme TCP [19]. De nos jours, il existe plusieurs versions et implémentations et beaucoup de littérature lui est dédiée. Pour motiver la problématique de ce chapitre, le lecteur doit être quelque peu familier avec TCP. Ainsi, cette section en fait une brève description.

Lors d'une communication entre deux machines, celles-ci établissent une connexion afin d'échanger des données. La machine émettrice (celle qui demande la connexion)

est appelée *client*, tandis que la machine réceptrice est appelée *serveur*. Les machines se trouvant dans un environnement Client-Serveur communiquent en “duplex”, c’est-à-dire, la communication se fait dans les deux sens. Que ce soit un courrier électronique, un fichier de données, un vidéo ou un message vocal, l’information à transmettre est coupée en plusieurs morceaux, appelés *segments*, qui sont transmis du serveur au client séquentiellement.

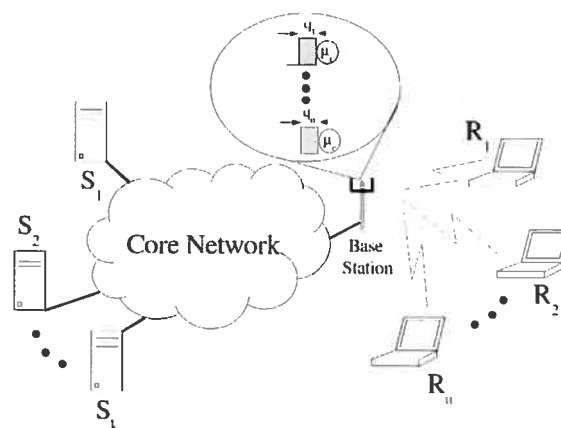


FIG. 4.1 – Topologie typique d’un système

La Figure 4.1 illustre la topologie typique du système qui sera étudié. Les k serveurs et les u clients sont connectés ensemble via le réseau fédérateur². Les paquets d’information arrivent à une station de base où chaque usager possède un tampon-mémoire dans lequel sont stockés temporairement les paquets qui lui sont destinés. C’est à la station de base que se forment des files d’attente responsables parfois des délais de transmission.

TCP est une procédure de transfert de données *fiable*. Ceci signifie qu’il est permis d’envoyer dans le réseau un nombre de segments w et pour chaque segment additionnel, le serveur doit attendre l’arrivée d’un *accusé de réception*³ (ACK) pour un des w paquets

²Réseau fédérateur : “core network”

³Accusé De Réception : “acknowledgement”

en circulation avant d'envoyer le segment suivant. S'il y a perte d'un segment, i.e., si l'accusé de réception n'est pas reçu dans un temps limite, TCP retransmet le segment manquant. Ainsi, il y a garantie que toute l'information sera transmise avec succès. On appelle *temps de boucle*⁴ (RTT) le temps écoulé entre l'envoi d'un paquet et la réception de la confirmation. On définit la *fenêtre* comme étant le nombre maximal d'octets non confirmés qu'il est permis d'envoyer dans le réseau. On distingue deux types de fenêtres :

- CWND : fenêtre de congestion, calculée par l'émetteur en utilisant les algorithmes de *démarrage lent*⁵ et d'*évitement de congestion*⁶, et
- AWND : fenêtre de réception, calculée par le récepteur et indiquant le nombre d'octets que celui-ci est capable de recevoir.

TCP est un algorithme de fenêtrage car les quelques sous-procédures qui le composent ont pour but de contrôler le nombre de paquets dans le réseau en contrôlant la grandeur de la fenêtre. Ainsi, la procédure de démarrage lent initialise CWND à 1 et la fait doubler chaque RTT, jusqu'à ce qu'une perte de paquet soit observée. À ce moment-là, l'algorithme d'évitement de congestion est invoqué. Celui-ci réduit CWND de moitié et l'incrémente ensuite linéairement (un paquet chaque RTT), jusqu'à ce que de nouveau une perte soit observée. La fenêtre TCP w_{TCP} est déterminée comme suit :

$$w_{TCP} = \min(\text{CWND}, \text{AWND}).$$

4.2 Amélioration de TCP : CLAMP

Malgré sa popularité, TCP a démontré à maintes reprises des comportements indésirables et problématiques. Le défaut majeur de TCP est qu'il n'a pas été conçu pour des réseaux comportant des liens sans fils qui se sont développés dernièrement. De tels réseaux possèdent la particularité d'avoir une capacité de transmission aléatoire, ce qui rend impossible le calcul de la taille des tampons-mémoire se trouvant dans les routeurs et les stations de base.

⁴Temps de boucle : "round trip time"

⁵Démarrage lent : "slow start"

⁶Évitement de congestion : "congestion avoidance"

Il est cependant possible de tirer avantage de cette capacité de transmission variable. En effet, il existe des algorithmes d'ordonnancement qui profitent des fluctuations du canal en transmettant plus de données lorsque la qualité du canal le permet et moins lorsque la qualité ne le permet pas. Il est évident que l'ordonnancement ne peut pas être optimal si la file d'attente à la station de base est vide. Idéalement, quelques paquets devraient toujours être maintenus en attente afin qu'ils puissent être transmis lorsque l'opportunité se présente. Par contre, maintenir des files de longueur excessive induira des délais de transmission trop grands et cette solution n'est pas acceptable. De même, maintenir des files trop courtes empêche le bon fonctionnement des couches inférieures. Ce problème a motivé la poursuite des recherches sur des algorithmes améliorant la performance de TCP sur des réseaux sans fils.

Par ailleurs, tel que décrit plus haut, le comportement "sautant" de TCP, causé en majeure partie par son algorithme d'évitement de congestion, interfère avec les mécanismes des couches inférieures, dont le fonctionnement se trouve sous-optimal.

Ainsi, nos collègues de l'Université de Melbourne ont développé un algorithme appelé CLAMP ("Curtailing Large AWNDs to Maximize Performance") qui satisfait aux exigences suivantes :

- Maintenir une longueur ciblée de la file d'attente ;
- Différencier la capacité du canal entre les usagers et les services.

L'idée est d'éviter que le comportement de TCP soit dicté par CWND et donc de faire en sorte que $w_{\text{TCP}} = \min(\text{CWND}, \text{AWND}) = \text{AWND}$. Pour cela, CLAMP va redéfinir la valeur de AWND selon une règle qui est décrite dans la section suivante.

4.2.1 Description de CLAMP

La description détaillée de CLAMP peut être trouvée dans [3, 4, 43]. Comme mentionné précédemment, l'idée principale est de contrôler la file d'attente des usagers à la station de base. En se référant de nouveau à la figure 4.1, supposons que $u = k$, c'est-à-dire, qu'il y a k flots de données qui passent à travers le réseau fédérateur et la station de base. À chaque flot j correspond un serveur S_j et un receveur R_j , $j = 1, \dots, k$. Il y a donc k files d'attente à la station de base. Soit d_j le délai de transmission total, incluant

tous les délais de propagation à travers le réseau fédérateur, mais excluant le délai d'attente à la station de base. CLAMP fait en sorte que la valeur de AWND soit sélectionnée de façon décentralisée, plus précisément, pour chaque flot i et à tout instant t , la valeur d'AWND $w_j(t)$ sera déterminée comme suit.

À des intervalles réguliers, la station de base mesure la longueur de la file d'attente de chaque usager et calcule la valeur d'une fonction de prix $p_a(q_j(\cdot))$ qui est insérée dans l'en-tête des paquets sortants. Soit t_k l'instant où le $k^{\text{ième}}$ paquet du flot j est reçu par le client. Celui-ci ajuste la grandeur de sa fenêtre w_j , $j = 1, \dots, k$, en se servant de la fonction $p_a(q_j)$ selon la règle suivante :

$$w_j(t_k) = \begin{cases} w_j(t_{k-1}) - 1 & \text{si } \Delta w_j(t_k) < -1 \\ w_j(t_{k-1}) + \Delta w_j(t_k) & \text{si } -1 \leq \Delta w_j(t_k) \leq \bar{\Delta} \\ w_j(t_{k-1}) + \bar{\Delta} & \text{si } \Delta w_j(t_k) > \bar{\Delta} \end{cases} \quad (4.1)$$

où

$$\Delta w_j(t_k) = [\phi_j \tau - p_a(q_j(t_k)) \hat{c}(t_k)] (t_k - t_{k-1}) \quad (4.2)$$

et $\phi_j > 0$ est une constante, $\tau > 0$ (paquets/sec) est constant et $\hat{c}(t_k)$ (octets/sec) est un estimé du taux de réception au temps t_k . L'incrément maximale de $w_j(\cdot)$ est limitée à une constante $\bar{\Delta} > 0$. Le taux de réception courant, $\hat{c}(\cdot)$, est estimé en utilisant une fonction de calcul de la moyenne à fenêtre coulissante :

$$\hat{c}(t_k) = \frac{\alpha}{t_k - t_{k-\alpha}}, \quad (4.3)$$

où l'entier α est un facteur de lissage.

La valeur de la fenêtre de réception AWND qui sera envoyée au serveur est :

$$\text{AWND} = \min[w(t_k), \text{AWND}_j],$$

où AWND_j est la valeur de AWND calculée par le client j à partir du système d'opération du receveur.

En décrivant la dynamique du réseau à l'aide d'un modèle fluide stochastique (section 4.2.2), nous allons voir que la performance de CLAMP est très sensible à la valeur du paramètre $a \in \mathbb{R}^+$ de la fonction de prix $p_a(\cdot)$. Le problème d'optimisation sera posé à la section 4.3.

4.2.2 Modèle Fluide Stochastique d'un système à événements discrets

Afin de faire l'analyse de performance de CLAMP et pour trouver le point d'équilibre de l'algorithme, un modèle fluide a été utilisé. Le modèle fluide [14, 58, 59, 60, 61] d'un système à événements discrets est une approximation ou un changement d'échelle. Au lieu de considérer les segments de données un par un et de marquer les arrivées et les départs comme des événements distincts, on considère le flot de données comme un fluide et on observe plutôt des *taux* d'arrivées et de départs. Cette pratique est très courante dans les applications en télécommunications, où il y a un très grand nombre d'événements dans un petit laps de temps. Pour cette raison, on parle aussi de changement d'échelle, puisqu'au lieu d'analyser ce qui se passe à chaque microseconde, on augmente l'échelle et on dit qu'il y a par exemple 1000 entrées par seconde. Dans ce cas, nous parlons donc de taux d'entrée, taux de sortie et taux de service, comme si un fluide circulait dans le système.

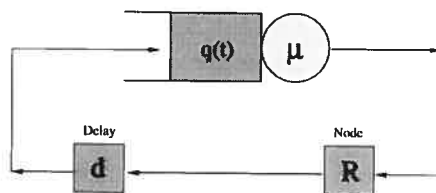


FIG. 4.2 – Modèle fluide avec rétroaction : file d'attente fermée

Avant de donner les équations décrivant la dynamique de la file d'attente, notons qu'il s'agit d'un système fermé avec rétroaction. En effet, avant d'envoyer un paquet, le serveur attend l'arrivée de l'accusé de réception du paquet envoyé antérieurement et par conséquent, l'évolution du système est gouvernée par le mouvement des accusés de

réception. Pour cette raison, il est modélisé à l'aide d'une file fermée, comme montré sur la Figure 4.2. Le contrôleur R modifie la quantité de fluide dans la boucle, ce qui correspond à l'ajustement de la fenêtre donné par CLAMP. Le délai de propagation d correspond au temps requis pour le retour des accusés de réception et est supposé connu.

Pour simplifier l'analyse, nous allons considérer le cas où il n'y a qu'une seule file d'attente, i.e., $k = 1$. L'indice $j = 1, \dots, k$ de la section 4.2.1 sera donc omis jusqu'à la fin de ce chapitre, mais l'analyse qui sera faite pourra facilement être généralisée pour un plus grand nombre de flots.

Soient les processus stochastiques suivants, définis tous sur le même espace de probabilité (Ω, \mathcal{F}, P) :

- $\{q_a(\omega, t)\}$: longueur de la file d'attente ;
- $\{c(\omega, t)\}$: taux de service, i.e., capacité maximale du lien sans fil ;
- $\{f_a(\omega, t)\}$: taux d'entrée des paquets ;
- $\{\mu_a(\omega, t)\}$: taux de sortie, i.e., le taux de transmission.

Remarquer que le paramètre a faisant partie de la fonction de prix $p_a(q)$ est noté en indice pour indiquer la dépendance des processus du paramètre, mais pour ne pas le confondre avec la véritable variable qu'est le temps t . Cependant, ce paramètre sera traité comme une variable jusqu'à la fin de ce chapitre et sa notation sera omise lorsque non nécessaire.

La dynamique du système est décrite par les équations suivantes.

- L'évolution de la file d'attente est donnée par la différence entre le taux d'entrée et le taux de sortie des paquets :

$$\frac{dq_a(t)}{dt} = f_a(t) - \mu_a(t). \quad (4.4)$$

- Comme la file d'attente est fermée avec un délai fixe de d unités de temps, le taux d'entrée $f(t)$ est égal au taux de sortie au temps $(t - d)$ plus (ou moins) le

nombre de paquets ajoutés (ou supprimés) associés à l'ajustement de la fenêtre :

$$f_a(t) = \mu_a(t-d) + \left. \frac{dw}{dt} \right|_{(t-d)} \quad (4.5)$$

$$= \phi + (1 - p_a(q_a(t-d)))\mu_a(t-d). \quad (4.6)$$

Remarque: $f_a(t)$ est une fonction croissante en t et en a et décroissante en $q(\cdot)$.

- Si la capacité maximale du canal au temps t est $c(t)$, le taux de sortie $\mu(t)$ est à son maximum s'il y a des paquets en attente, sinon le minimum de ressources nécessaires est utilisé pour acheminer les paquets :

$$\mu_a(t) = \begin{cases} c(t) & \text{si } q > 0 \\ \min[f_a(t), c(t)] & \text{sinon.} \end{cases} \quad (4.7)$$

- La fonction de "prix" utilisée doit être convexe, monotone et croissante. Dans notre cas, nous allons travailler avec la fonction linéaire

$$p_a(q) = \frac{bq - a}{E[c]}, \quad (4.8)$$

où a et b sont des constantes. Nous allons supposer que la valeur de b est un nombre réel connu et que celle de a appartient à un ensemble fermé et borné $\mathcal{A} \subset \mathbb{R}^+$.

- Le processus stochastique $c(t)$ décrivant les fluctuations du canal prend des valeurs dans un ensemble fini \mathcal{C} , i.e.,

$$c(t) = c_i \geq 0, \quad t \in [\tau_i, \tau_{i+1}), \quad c_i \in \mathcal{C},$$

et les intervalles de temps $[\tau_i, \tau_{i+1})$, $i = 1, 2, \dots$, sont une suite de variables aléatoires i.i.d. Nous allons voir plus loin que les points de rupture τ_i , $i = 1, 2, \dots$, définissent différents modes de fonctionnement du système. $c(t)$ sera donc constant par morceaux et continu à droite. Soit $\xi_i(n)$ la durée de la $i^{\text{ème}}$ visite du processus dans l'état c_i . Nous allons supposer que pour tout i , les v.a.

$\xi_i(n) : \Omega \rightarrow R^+$, $n = 1, 2, \dots$, sont des variables uniformément continues et i.i.d. de densité $f_\xi(x)$.

Pour un taux de transmission déterministe, $c(t) \equiv c$ (octets /sec), il a été démontré que la longueur moyenne de la file d'attente q^* est donnée par

$$q_a^* = p_a^{-1} \left(\frac{\phi}{c} \right) = \frac{1}{b} \left(\frac{\phi}{c} E[c] + a \right). \quad (4.9)$$

Éq. (4.9) indique qu'il est possible de contrôler la longueur de la file d'attente à la station de base en choisissant correctement les paramètres de la fonction $p_a(\cdot)$. Lorsque le taux de service est déterministe, il est possible de faire cela statistiquement. Cependant, dans un système sans fil, nous venons de voir que le taux de transmission $c(t)$ est un processus stochastique et il en est de même pour la longueur de la file d'attente. Celle-ci dépend de $p_a(\cdot)$, du délai de rétroaction et des statistiques de $c(t)$.

Dans les sections suivantes, nous développons l'algorithme qui fera l'ajustement dynamique de la fonction $p_a(\cdot)$ à la station de base, de façon à optimiser une fonction d'utilité impliquant la longueur espérée de la file d'attente et l'utilisation moyenne du lien.

4.3 Problème d'optimisation

Définissons les fonctions de performance suivantes :

$$\begin{aligned} \text{Utilisation du système : } U(\omega, a) &= \frac{1}{T} \int_0^T \mathbf{1}(q_a(\omega, t) > 0) dt \\ \text{Longueur moyenne de la file : } Q(\omega, a) &= \frac{1}{T} \int_0^T q_a(\omega, t) dt \end{aligned}$$

La notation ci-haut fait ressortir le fait que $U(\omega, a)$ et $Q(\omega, a)$ sont des variables aléatoires, fonctions d'une trajectoire de la simulation $\omega \in \Omega$. Par la suite, la dépendance en ω sera omise, car pour la dérivée IPA, ω est fixé et l'analyse faite est *déterministe*.

L'objectif double de CLAMP est de maximiser $E[U(a)]$ et en même temps, minimiser $E[Q(a)]$, ce qui est quelque peu contradictoire. Il s'agit donc de trouver une sorte

d'équilibre économique entre ces deux mesures de performance. Ainsi, nous définissons la fonction d'utilité suivante qui maximise l'utilisation tout en pénalisant les files d'attente trop longues :

$$\max_{a \in \mathcal{A}, a \geq 0} J(a) = E[U(a)] - \lambda (E[Q(a)])^2, \quad (4.10)$$

où $\lambda > 0$ est un facteur d'échelle permettant de comparer les unités d'utilisation avec les unités de file d'attente. Éq. (4.10) peut être vue comme le Lagrangien d'un problème avec une contrainte sur $E[Q(a)]$, mais puisque nous n'avons pas de borne sur l'espérance de $Q(a)$, nous avons préféré travailler avec la fonction d'utilité ci-haut.

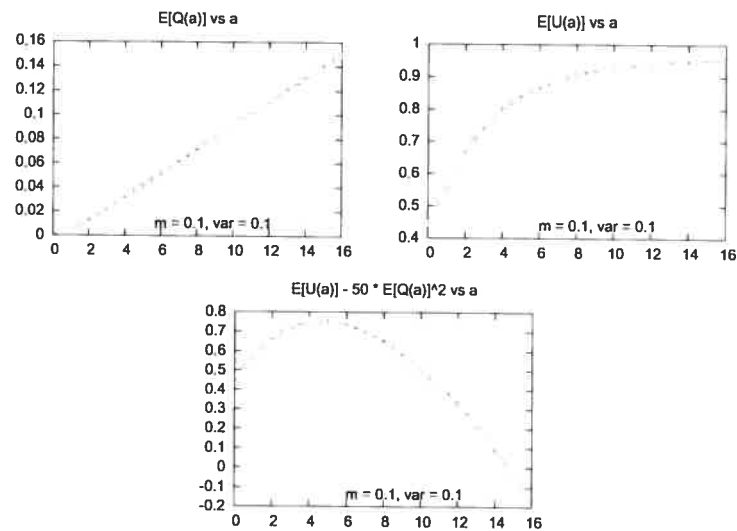


FIG. 4.3 – Fonctions de performance $E[Q(a)]$ et $E[U(a)]$ et fonction objective $J(a) = E[U(a)] - \lambda (E[Q(a)])^2$ en fonction de a (facteur d'échelle $\lambda = 50$).

La figure 4.3 montre les courbes des fonctions de performance ainsi que celle de la fonction d'utilité en fonction du paramètre $a \in \mathbb{R}^+$. On voit que $E[Q(a)]$ et $E[U(a)]$ sont croissantes en a et que la fonction objective (4.10) est concave, ce qui assure l'existence

d'un maximum global unique. Le processus itératif de l'approximation stochastique est dans ce cas :

$$a_{n+1} = a_n + \epsilon_n \nabla J(a),$$

où

$$\nabla J(a) = \frac{d}{da} J(a) = \frac{d}{da} E[U(a)] - 2\lambda E[Q(a)] \frac{d}{da} E[Q(a)].$$

Pour estimer les gradients de la fonction stochastique $J(a)$, nous allons utiliser la technique IPA ("Infinitesimal Perturbation Analysis") [27]. Si la dérivée IPA de $\nabla J(a)$ est non biaisée et si un gain décroissant est utilisé, l'algorithme d'approximation stochastique convergera a.p.1 (cf. Théorème 1, Chapitre 1).

De façon générale, une dérivée IPA $\mathcal{L}'(a)$ est non biaisée si les deux conditions suivantes sont vérifiées :

Condition 1. Pour tout $a \in \mathcal{A}$, la dérivée $\mathcal{L}'(a)$ existe avec probabilité 1.

Condition 2. Avec probabilité 1, la fonction aléatoire $\mathcal{L}'(a)$ est Lipschitz-continue sur \mathcal{A} , et la constante de Lipschitz (généralement aléatoire) possède un moment de premier ordre fini.

Nous devons donc vérifier les deux hypothèses ci-haut au sujet de $E[Q(a)]$ et $E[U(a)]$. Une fois la Lipschitz-continuité de $Q(a)$ et $U(a)$ démontrée, nous pourrions appliquer le Théorème de Convergence Dominée [53] et échanger la dérivée et l'espérance dans $\frac{d}{da} E[U(a)]$ et $\frac{d}{da} E[Q(a)]$. Nous obtiendrons ainsi des estimateurs des gradients non biaisés, soit

$$Q'(a) \equiv \frac{d}{da} Q(a) \quad \text{et} \quad U'(a) = \frac{d}{da} U(a),$$

qui seront évalués en ligne à l'aide du modèle fluide pour être ensuite implantés dans l'approximation stochastique.

4.4 Calcul de la dérivée IPA

Dans cette section sont donnés en détail les calculs des dérivées IPA de $E[U(a)]$ et $E[Q(a)]$. Il existe deux façons de calculer ces dérivées :

- 1) en utilisant la définition d'une dérivée, i.e., en étudiant le comportement du système avec et sans perturbation du paramètre en question puis en prenant la limite lorsque la perturbation tend vers 0 ;
- 2) en évaluant explicitement la dérivée, directement sur les équations du modèle.

La première approche s'avère beaucoup plus ardue, car elle requiert une analyse en profondeur d'une trajectoire avec tous les événements qui peuvent survenir. Par contre, comme elle utilise la définition de la dérivée, elle a l'avantage de fournir les éléments nécessaires pour la démonstration de "non-biais" des estimateurs calculés avec la deuxième méthode. Nous allons utiliser la première approche pour évaluer la dérivée $U'(a)$ et la seconde pour évaluer $Q'(a)$.

4.4.1 Analyse d'une trajectoire

Pour simplifier l'analyse, nous allons sans perte de généralité considérer le cas où le processus stochastique décrivant l'évolution du canal $c(t)$ est donné par un processus "on-off", où à chaque X secondes, avec $X \sim \text{Uniforme}[\epsilon, 10]$, $\epsilon > 0$, deux valeurs déterministes du canal sont alternées : une voie de transmission de bonne qualité avec un taux de transmission de 2 Mbit/sec et une de mauvaise qualité avec un taux de 1 Mbit/sec.

Tout d'abord, nous définissons des *points de rupture* qui marquent un changement de régime dans l'évolution du système. Les points de rupture sont par exemple les moments où le taux d'entrée ou de sortie changent ou bien lorsque la file d'attente devient vide ou commence à se remplir.

Définition 1 Un point τ_i est un "point de rupture" si le signe de $f(\cdot) - \mu(\cdot)$ change en ce point.

Notons qu'il existe deux types de points de rupture :

- (i) un point *exogène*, provoqué par un événement dont l'occurrence est *indépendante* du paramètre a . Dans notre cas, un saut de $c(t)$ est un point de rupture exogène ;
- (ii) un point *endogène*, provoqué par un événement dont l'occurrence est *dépendante* du paramètre a . Dans notre cas, un saut de $f(t)$ ou de $\mu(t)$ est un point de rupture

endogène.

Le lemme suivant stipule que deux événements ne peuvent pas se produire en même temps, ce qui va garantir qu'à chaque moment, seulement un point de rupture peut survenir.

Lemme 5 *Pour tout $a \in \mathcal{A}$, avec probabilité 1, deux événements ne peuvent se produire simultanément.*

Preuve: Le processus stochastique qui dicte le comportement du système est celui décrivant les fluctuations du canal : $\{c(t)\}_{t \geq 0}$. Si deux événements se produisent simultanément au temps τ , c'est que la valeur de $c(t)$ change deux fois au moment τ , c'est-à-dire que le processus reste dans un des états pendant $\epsilon \rightarrow 0$ unités de temps. Or, nous avons supposé que $\xi_i(n)$, la durée de la $i^{\text{ème}}$ visite du processus dans l'état c_i , est une v.a. uniformément continue et positive. Par conséquent, pour $\epsilon \rightarrow 0$, $P[\xi_i \leq \epsilon] = \int_0^\epsilon f_\xi(y) dy = F_\xi(\epsilon) \rightarrow 0$, ce qui prouve le résultat.

□

La trajectoire de $q(t)$ va être décomposée en périodes où des paquets sont en attente dans la file et en périodes où la file d'attente est vide. Des points de rupture marquent le début et la fin de ces périodes.

Définition 2 *Un intervalle de temps de la forme $[\zeta_i, \eta_i)$, $i = 1, 2, \dots$, est appelé " $i^{\text{ème}}$ période occupée", notée PO_i , si $q(t) > 0$ pour $t \in [\zeta_i, \eta_i)$.*

Définition 3 *Un intervalle de temps de la forme $[\eta_i, \zeta_{i+1})$, $i = 1, 2, \dots$, est appelé " $i^{\text{ème}}$ période vide", notée PV_i , si $q(t) = 0$ pour $t \in [\eta_i, \zeta_{i+1})$.*

L'évolution de la file d'attente est constituée par une alternance de périodes vides et de périodes occupées, causées par les fluctuations de la qualité du canal de transmission. Dans ce qui suit, nous allons décrire tous les événements possibles survenant au cours des périodes vides et occupées et analyser le comportement du système entre les points de rupture définis par ces événements.

Afin de mesurer la sensibilité du système au paramètre a , nous allons comparer ce qu'on appelle un processus *nominal*, dont la trajectoire est observée pour une valeur de a

donnée, avec le processus *perturbé* associé, dans lequel on ajoute une petite perturbation au paramètre, soit $\Delta a > 0$. Une analyse similaire pourrait être faite avec une perturbation négative, mais sans perte de généralité, nous allons nous en tenir à la première. Les deux processus sont simulés en parallèle en utilisant des nombres aléatoires communs, afin de mesurer uniquement l'impact du paramètre. À la suite de cette analyse, en faisant la différence entre les deux systèmes et en prenant la limite lorsque Δa tend vers 0, il en résultera des dérivées à droite, trajectoire par trajectoire, de $U(a)$.

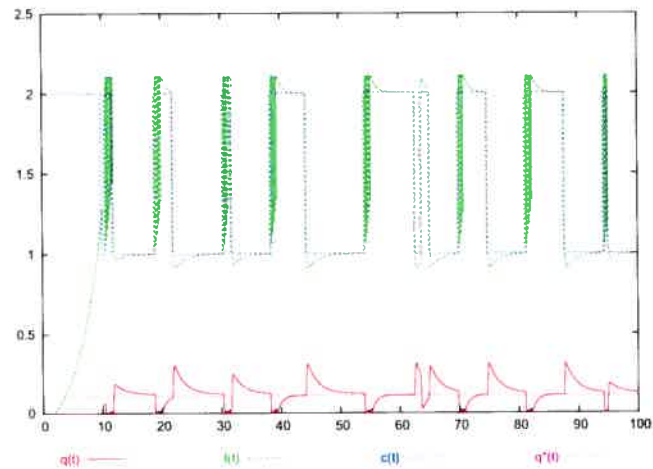


FIG. 4.4 – Trajectoire d'une simulation : $a = 0.1$, $b = 1.0$, $\phi = 0.01$, $d = 0.2$, $k = 1$. Courbes de $q(t)$, $q^*(t)$, $f(t)$ et $c(t)$ en fonction du temps.

La figure 4.4 montre une trajectoire typique du système. On y observe l'évolution de $q(t)$, $c(t)$ et $f(t)$. D'abord, remarquons que la file d'attente réagit aux fluctuations du canal. Lorsque celui-ci est de bonne qualité, les paquets de données sont transmis rapidement et la file d'attente se vide. Lorsque des bruits de transmission apparaissent, moins de données peuvent être acheminées et la file d'attente se remplit. Dans chacun des deux cas, CLAMP ajuste la fenêtre de réception AWND de façon à faire augmenter le flot lorsque le canal est bon et à le diminuer lorsqu'il est mauvais. Donc, malgré les

changements de $c(t)$, CLAMP s'ajuste pour opérer en un point optimal, où la longueur de la file d'attente est égale à $q^* = p^{-1} \left(\frac{\phi}{c} \right)$.

Ensuite, on remarque qu'en même temps que $q(t)$ tend vers q^* , la fonction $f(t)$ tend vers la valeur de $c(t)$. Pour comprendre pourquoi cela se produit, il faut noter d'une part que lorsque le système est en équilibre, le taux d'entrée est égal au taux de sortie :

$$\frac{dq}{dt} = f(t) - \mu(t) = 0 \quad \Rightarrow \quad f(t) = \mu(t). \quad (4.11)$$

D'autre part, si au temps t , $q(t) > 0$, alors Éq. (4.7) implique que $\mu(t) = c(t)$. Par conséquent, comme le système cherche à se stabiliser en $q^* > 0$, $f(t)$ va tendre vers $c(t)$.

Maintenant, analysons de plus près chaque type de point de rupture et le comportement des files nominale et perturbée qui en résulte. Pour chaque régime d'opération, nous allons considérer les différences finies définies ci-dessous.

$$\begin{aligned} \Delta q(\omega, t) &= q_{a+\Delta a}(\omega, t) - q_a(\omega, t) \\ \Delta f(\omega, t) &= f_{a+\Delta a}(\omega, t) - f_a(\omega, t) \\ \Delta \mu(\omega, t) &= \mu_{a+\Delta a}(\omega, t) - \mu_a(\omega, t) \end{aligned}$$

Afin de démontrer la Lipschitz-continuité de $U(a)$ et $Q(a)$, nous devons trouver des bornes pour ces différences. Pour le reste de l'analyse, nous considérons une trajectoire en particulier, donc $\omega \in \Omega$ sera fixé (et on ne notera plus la dépendance). Tout d'abord, voyons comment se comportent les systèmes lorsque les files d'attente sont à l'équilibre.

* * * * *

Système en équilibre

Éq. (4.9) indique que

$$q_{a+\Delta a}^*(t) = q_a^*(t) + \frac{\Delta a}{b}. \quad (4.12)$$

On a donc $\Delta q^*(t) = \Delta a/b > 0$ si $\Delta a > 0$. Il est facile de voir que pour des systèmes qui sont à l'équilibre, $\Delta f(t) = 0$ et $\Delta \mu(t) = 0$.

Période vide

Supposons que $q_a(t) > 0$, $q_{a+\Delta a}(t) > 0$ et $q_{a+\Delta a}(t) > q_a(t)$ ⁷. Une période vide ne peut être engendrée que par un événement exogène, c'est-à-dire, par une augmentation de $c(t)$, la capacité de transmission. Soit τ_i le point de rupture où $c(t)$ passe de c_{i-1} à c_i (Figure 4.5). À ce moment là, $\mu_a(\tau_i) = \mu_{a+\Delta a}(\tau_i) = c(\tau_i) = c_i$. Les deux files d'attente commencent à se vider, car $\left. \frac{dq}{dt} \right|_{\tau_i} = f(\tau_i) - c_i < 0$.

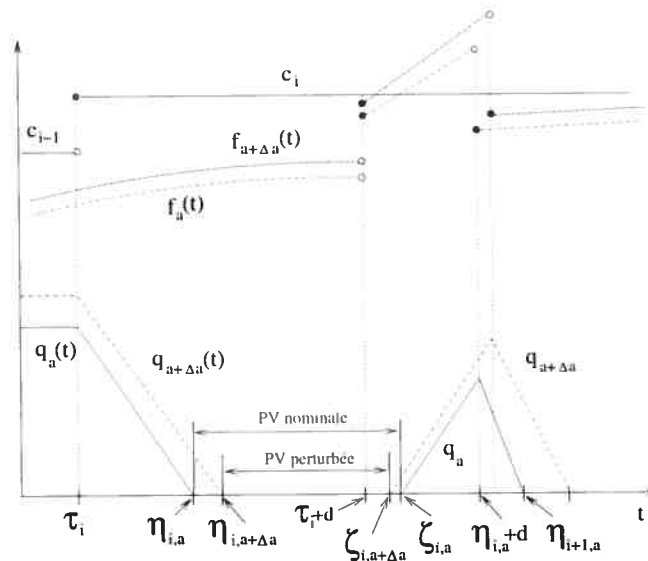


FIG. 4.5 – Trajectoire d'une simulation : période vide et ses points de rupture.

Si avant de commencer à drainer, les deux systèmes étaient en équilibre, i.e., si pour $t < \tau_i$, on avait $f_a(t) = c(t) = f_{a+\Delta a}(t)$, les deux files se videraient au même taux et la file perturbée atteindrait la valeur 0 exactement $\Delta a/b$ unités de temps après la file nominale (Éq. (4.12)).

⁷La domination de $q_{a+\Delta a}(\cdot)$ sur $q_a(\cdot)$ sera démontrée plus loin dans le Corollaire 1.

Soit $\eta_{i,a}$ le moment où la file nominale devient zéro et $\eta_{i,a+\Delta a}$ le moment où la file perturbée devient zéro. Puisque pour $t < \tau_i$, $q_{a+\Delta a}(t) > q_a(t)$, on a $\eta_{i,a+\Delta a} \geq \eta_{i,a}$, c'est-à-dire que la période vide pour le système nominal commence avant celle du système perturbé.

Au temps $\tau_i + d$, $f_a(t)$ et $f_{a+\Delta a}(t)$ ont un saut causé par l'augmentation de $c(t-d)$. Ce saut est suivi d'une incrémentation de $f_a(t)$ et $f_{a+\Delta a}(t)$, causée par la diminution de $q_a(t-d)$ et $q_{a+\Delta a}(t-d)$. Cette incrémentation va éventuellement donner lieu à une période occupée endogène, comme nous allons voir plus loin. Supposons que le délai d est tel que les deux files d'attente ont le temps de se vider avant le début de la période occupée endogène, i.e., $\tau_i + d \geq \eta_{i,a}$ et $\tau_i + d \geq \eta_{i,a+\Delta a}$, comme montré à la Figure 4.5.⁸

Deux types d'événements peuvent terminer la période vide : un événement endogène ou un événement exogène. Soit $\zeta_{i,a}$ le point de rupture marquant la fin de la période vide nominale et $\zeta_{i,a+\Delta a}$ le point de rupture marquant la fin de la période vide perturbée. Si une décrémentation de $c(t)$ (exogène) cause la fin de la période vide, les deux files d'attente commencent à se remplir en même temps, par conséquent, $\zeta_{i,a} = \zeta_{i,a+\Delta a}$. Si $f(t)$ devient supérieur à $c(t)$ (endogène) et provoque l'emplissage des files, cet événement peut survenir à des moments différents pour les deux systèmes.

Voyons comment se comportent les systèmes entre les points de rupture identifiés ci-haut. Tout d'abord, puisque nous travaillons avec un système à rétroaction, nous devons faire la supposition suivante, qui sera démontrée dans le Lemme 6.

Supposition 1 *Supposons que $|\Delta q(t-d)| \leq K' \Delta a$, où K' est une v.a. avec $E[K'] < \infty$.*

Pour $t \in [\eta_{i,a}, \eta_{i,a+\Delta a})$ et pour $t \in [\zeta_{i,a+\Delta a}, \zeta_{i,a})$, $q_a(t) = 0$ et $q_{a+\Delta a} > 0$ (voir figure 4.5). On a

$$0 \leq \Delta q(t) = q_{a+\Delta a}(t) \approx \frac{\partial}{\partial a} q_a(\eta_i) \Delta a + \mathcal{O}(\Delta a^2). \quad (4.13)$$

⁸Si $\eta_{i,a} \leq \tau_i + d < \eta_{i,a+\Delta a}$, il y a une période vide pour le système nominal, mais pas pour le système perturbé. Nous allons supposer qu'il existe Δa assez petit, tel que la probabilité que $\tau_i + d$ tombe entre $\eta_{i,a}$ et $\eta_{i,a+\Delta a}$ est presque nulle.

Pour $t \in [\eta_{i,a+\Delta a}, \zeta_{i,a+\Delta a})$, on a $q_a(t) = q_{a+\Delta a} = 0$ et Supposition 1 nous permet d'écrire :

$$\begin{aligned} \Delta f(t) &= \left(1 - \frac{bq_{a+\Delta a}(t-d) + a + \Delta a}{E[c]}\right) c(t-d) - \left(1 - \frac{bq_a(t-d) + a}{E[c]}\right) c(t-d) \\ &= \frac{c(t-d)}{E[c]} (\Delta a + b(q_a(t-d) - q_{a+\Delta a}(t-d))) \\ &\leq \frac{c(t-d)}{E[c]} (1 + bK') \Delta a \end{aligned}$$

Posons

$$K'' = \frac{c(t-d)}{E[c]} (1 + bK'). \quad (4.14)$$

Notons que puisque K' est une v.a. avec $E[K'] < \infty$, K'' est aussi une v.a. dont l'espérance est finie.

Donc, pour une période vide où $t \in [\eta_{i,a+\Delta a}, \zeta_{i,a+\Delta a})$, on a :

$$\Delta f(t) \leq K'' \Delta a \quad (4.15)$$

$$\Delta \mu(t) = f_{a+\Delta a}(t) - f_a(t) = \Delta f(t) \quad (4.16)$$

$$\Delta q(t) = 0. \quad (4.17)$$

Période occupée

Supposons, sans nous préoccuper des processus nominal et perturbé, que $q(t) = 0$. Ceci signifie que le taux d'entrée est inférieur ou égal au taux de sortie, i.e., $f(t) \leq \mu(t)$.

Deux événements peuvent causer l'incrément de la file d'attente :

- (i) $c(t)$ diminue et devient plus petit que $f(t)$ (exogène) ;
- (ii) $f(t)$ augmente et devient plus grand que $c(t)$ (endogène).

Dans le premier cas, supposons qu'au temps $t = \zeta_i$, $c(t)$ diminue et devient inférieur à $f(t)$. Comme la file d'attente au temps ζ_i est vide, $\mu(\zeta_i) = \min[f(\zeta_i), c(\zeta_i)] = c(\zeta_i)$. Par conséquent, $\left. \frac{dq}{dt} \right|_{\zeta_i} = f(\zeta_i) - \mu(\zeta_i) > 0$ et la file d'attente se remplit. Le point de rupture ζ_i est exogène et la période occupée qu'il engendre est appelée *période exogène*.

Dans le second cas, soit ζ_i le moment où $f(t)$ dépasse $c(t)$. Si au temps $t < \zeta_i$, $q(t) = 0$ et $\mu(t) = f(t)$, alors en ζ_i , $\mu(\zeta_i) = \min[f(\zeta_i), c(\zeta_i)] = c(\zeta_i) < f(\zeta_i)$ et la file d'attente commence à se remplir. Cette fois-ci, ζ_i est *endogène*, tout comme la période occupée engendrée.

Voici l'analyse de ces deux types de périodes occupées.

(i) Période occupée exogène

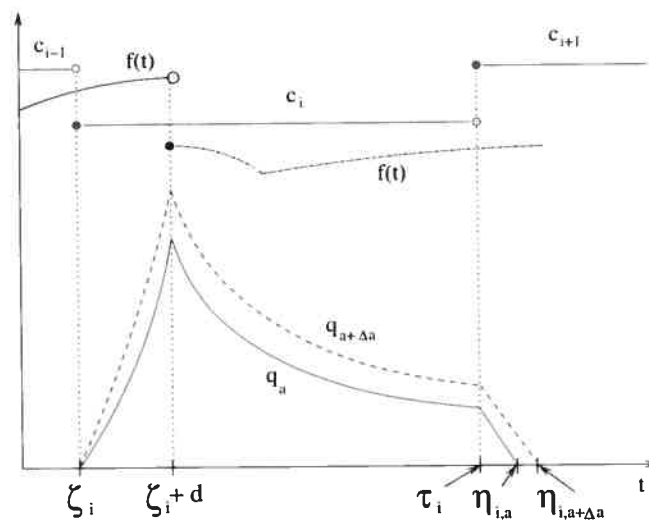


FIG. 4.6 – Points de rupture d'une période occupée exogène

Les points de rupture pour une période occupée exogène sont les suivants (voir figure 4.6) :

1. ζ_i (exogène) : diminution de $c(t)$; début de la période occupée
2. $\zeta_i + d$ (endogène) : saut de $f(t)$ causé par la décrémentation de $c(t - d)$
3. τ_i (exogène) : augmentation de $c(t)$
4. $\eta_{i,a}$ (endogène) : $q_a(t)$ devient nul ; fin de la période occupée nominale
5. $\eta_{i,a+\Delta a}$ (endogène) : $q_{a+\Delta a}(t)$ devient nul ; fin de la période occupée perturbée

Supposons qu'en $t = \zeta_i$, $c(t)$ passe de c_{i-1} à c_i , avec $c_i < c_{i-1}$, $c_i < f_a(\zeta_i)$ et $c_i < f_{a+\Delta a}(\zeta_i)$. Les deux files d'attente commencent à se remplir comme indiqué sur la Figure 4.6. Puisque $f_a(t)$ est une fonction croissante en a , la file perturbée incrémente plus rapidement que la file nominale.

Au temps $\zeta_i + d$, $f_a(t)$ et $f_{a+\Delta a}(t)$ ont un saut causé par la diminution de $\mu(t-d)$. Comme le taux d'entrée décroît, les deux files diminuent pour tendre vers leur point de stabilité q_a^* et $q_{a+\Delta a}^*$, en même temps que $f_a(t)$ et $f_{a+\Delta a}(t)$ tendent vers c_i . Supposons qu'au temps τ_i , $c(t)$ incrémente de nouveau pour passer de c_i à c_{i+1} , $c_{i+1} > c_i$. Les files d'attente commencent à se vider, car $\left. \frac{dq}{dt} \right|_{\tau_i} = f(\tau_i) - c_i < 0$.

Soit $\eta_{i,a}$ le moment où la file nominale atteint zéro et $\eta_{i,a+\Delta a}$ le moment où la file perturbée atteint zéro. Puisque pour $t < \tau_i$, $q_{a+\Delta a}(t) > q_a(t)$, on a $\eta_{i,a+\Delta a} \geq \eta_{i,a}$, c'est-à-dire que la période occupée du système nominal est plus courte que celle du système perturbé.

Durant toute la période occupée nominale, i.e., pour $t \in [\zeta_i, \eta_i)$ on a

$$\Delta\mu(t) = 0. \quad (4.18)$$

Pour $t \in [\zeta_i, \zeta_i + d)$, il y a deux cas⁹ :

(a) si $q_a(t-d) = 0$ et $q_{a+\Delta a}(t-d) = 0$, on a

$$0 \leq \Delta f(t) \leq \left(\left(1 + \frac{a}{\mathbb{E}[c]} \right) K'' + \frac{f_{a+\Delta a}(t-d)}{\mathbb{E}[c]} \right) \Delta a \quad (4.19)$$

$$0 \leq \Delta q(t) \leq \left(d \left(1 + \frac{a}{\mathbb{E}[c]} \right) K'' + \frac{I_f}{\mathbb{E}[c]} \right) \Delta a, \quad (4.20)$$

où K'' est donné par (4.14) et

$$I_f = \int_{\zeta_i}^{\zeta_i+d} f_{a+\Delta a}(t-d) dt < \infty. \quad (4.21)$$

⁹Les autres cas n'arrivent pas, car si τ_{i-1} est l'événement ayant provoqué le vidage des files d'attente durant la période occupée $i-1$ et η_{i-1} le début de la période vide précédent la période occupée exogène, il n'y a qu'une possibilité pour l'événement exogène ζ_i : c'est qu'il arrive avant l'événement endogène qui va faire remplir les files. Donc, $\zeta_i \leq \tau_i + d$. Si la PO précédente est de longueur supérieure à d , $q_a(\zeta_i - d) > 0$ et $q_{a+\Delta a}(\zeta_i - d) > 0$, sinon, $q_a(\zeta_i - d) = 0$ et $q_{a+\Delta a}(\zeta_i - d) = 0$.

(b) si $q_a(t-d) > 0$ et $q_{a+\Delta a}(t-d) > 0$, on a

$$0 \leq \Delta f(t) \leq \frac{c(t-d)}{E[c]}(1+bK')\Delta a \quad (4.22)$$

$$0 \leq \Delta q(t) \leq d \frac{c(t-d)}{E[c]}(1+bK')\Delta a, \quad (4.23)$$

où K' est la constante de la Supposition 1.

Finalement, pour $t \in [\zeta_i + d, \eta_{i,a})$,

$$0 \leq \Delta f(t) \leq \frac{c(t-d)}{E[c]}(1+bK')\Delta a \quad (4.24)$$

$$0 \leq \Delta q(t) \leq (\eta_{i,a} - \zeta_i - d) \frac{c(t-d)}{E[c]}(1+bK')\Delta a. \quad (4.25)$$

(ii) Période occupée endogène

Supposons enfin que l'emplissage de la file est causé par une augmentation de $f(\cdot)$; une période occupée endogène est engendrée. Une suite de périodes occupées endogènes se produit généralement après une période vide, lorsque $f(\cdot)$ a des sauts causés par la rétroaction. Si la valeur de $c(t)$ ne change pas, la suite de période endogènes se termine par la stabilisation du système en son point d'équilibre.

Les points de rupture d'une période endogène sont (voir figure 4.7) :

1. $\zeta_{i,a}$ (endogène) : augmentation de $f(t)$; début de la période occupée PO_i
2. $\tau_{i,a}$: moment où $f(t)$ devient inférieur à $c(t)$ (endogène) ou moment où $c(t)$ devient supérieur à $f(t)$ (exogène)
3. $\eta_{i,a}$ (endogène) : $q_a(t)$ devient nul; fin de la période occupée
4. $\zeta_{i+1,a}$ (endogène) : moment où $f(t)$ devient supérieur à $c(t)$; début de la période occupée PO_{i+1}

Deux cas peuvent se produire : (a) si $\eta_i < \zeta_{i+1}$, la période occupée endogène PO_i termine, (b) si $\eta_i > \zeta_{i+1}$, le système se stabilise en q^* . Voyons ce qui se passe dans le cas (a) (le cas (b) est similaire et ne présente pas d'intérêt).

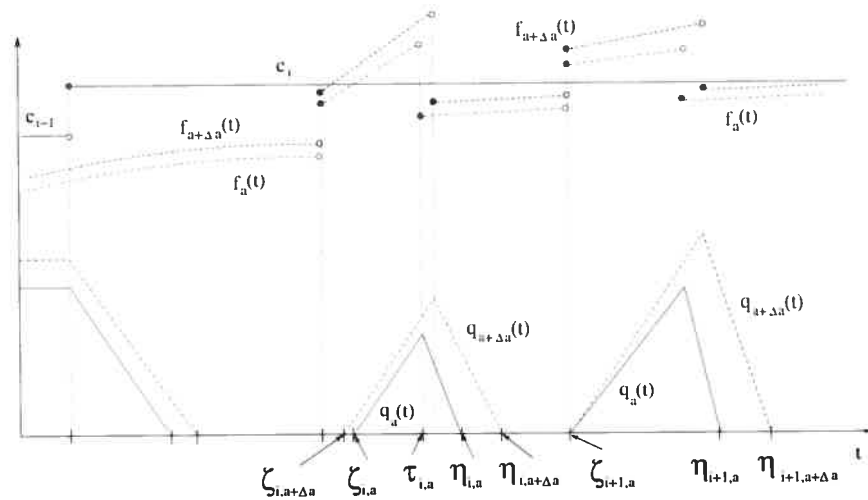


FIG. 4.7 – Points de rupture d'une période occupée endogène

Pour $t \in [\zeta_{i,a+\Delta a}, \zeta_{i,a})$ et pour $t \in [\eta_{i,a}, \eta_{i,a+\Delta a})$, $q_a(t) = 0$, $q_{a+\Delta a}(t) > 0$ et $\Delta q(t)$ satisfait l'inégalité (4.13).

Pour $t \in [\zeta_{i,a}, \tau_{i,a})$, $q_a(t-d) > 0$ et $q_{a+\Delta a}(t-d) > 0$, donc Δf se trouve dans la situation de l'équation (4.22) et

$$0 \leq \Delta q(t) \leq (\tau_{i,a} - \zeta_{i,a}) \frac{c(t-d)}{E[c]} (1 + bK') \Delta a, \quad (4.26)$$

où K' est la constante de Supposition 1.

Pour $t \in [\zeta_{i,a}, \tau_{i,a})$, $q_a(t-d) = 0 = q_{a+\Delta a}(t-d)$ et

$$0 \leq \Delta f(t) \leq \left(1 + \frac{a}{E[c]}\right) K'' \Delta a \quad (4.27)$$

$$0 \leq \Delta q(t) \leq \left[(\eta_{i,a} - \tau_{i,a}) \left(1 + \frac{a}{E[c]}\right) K'' + \frac{I_f}{E[c]} \right] \Delta a, \quad (4.28)$$

où la constante K'' est donnée par équation (4.14) et I_f par équation (4.21).

Après avoir analysé la trajectoire du système, voici un résultat qui va garantir la

Condition 2 de la section 4.3.

Lemme 6 (Lipschitz-continuité) *Pour tout $a \in \mathcal{A}$ et pour tout intervalle de la forme $[\tau_i, \tau_{i+1})$, où τ_i est un point de rupture, $i = 1, 2, \dots$, la fonction $q_a(t)$ est presque sûrement Lipschitz continue, i.e., il existe une variable aléatoire K avec $E[K] < \infty$, telle que*

$$\sup_{\tau_i \leq t < \tau_{i+1}} |q_{a+\Delta a}(t) - q_a(t)| \leq K\Delta a, \quad \text{a.p.1.}$$

Preuve: La preuve est directe si on considère les bornes sur Δq données par les équations (4.12), (4.17), (4.20), (4.23), (4.25), (4.26) et (4.28) avec K choisi respectivement pour chaque intervalle. Pour que la preuve soit complète, il reste à démontrer Supposition 1. Nous devons trouver un instant t pour lequel $|\Delta q(t - d)| \leq K'\Delta a$. En effet, cette inégalité est vérifiée lorsque $q(t - d) = 0$, car à ce moment-là, $\Delta q(t) = 0$ (voir éq. (4.17)). Le système est en quelque sorte régénératif. Comme a.p.1, le système a des périodes vides, le résultat est démontré. □

Corollaire 1 *Pour tout $\Delta a > 0$ et tout $t \geq 0$, le processus perturbé domine le nominal presque sûrement, i.e., $q_{a+\Delta a}(t) \geq q_a(t)$ a.p. 1.*

Preuve: Découle directement des équations (4.12), (4.17), (4.20), (4.23), (4.25), (4.26) et (4.28) et du fait que le taux d'entrée $f_a(\cdot)$ est croissant en a . □

4.4.2 Dérivées des temps d'événements

Les lemmes suivants donnent la dérivée des événements marquant le début et la fin d'une période occupée et seront utiles pour la dérivée de la fonction $U(\cdot)$.

Lemme 7 *Soit $\eta_{i,a}$ le $i^{\text{ème}}$ instant où $q_a(t)$ atteint la valeur 0. On a*

$$\frac{d\eta_{i,a}}{da} = \frac{\frac{\partial}{\partial a} q_a(\eta_{i,a})}{f_a(\eta_{i,a}) - c(\eta_{i,a})}. \quad (4.29)$$

Preuve: Considérons les systèmes nominal et perturbé. Puisque le processus perturbé domine le nominal (cf. Corollaire 1), on sait que $\eta_{i,a+\Delta a} \geq \eta_{i,a}$, c'est-à-dire la file perturbée va atteindre la valeur 0 après la nominale. Nous voulons savoir combien de temps

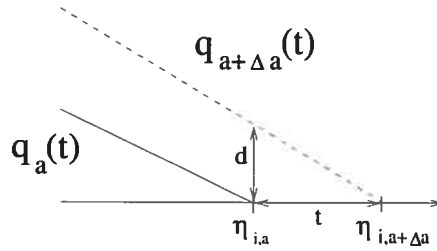


FIG. 4.8 – Terminaison d'une période occupée

après. Se référant à la Figure 4.8, nous cherchons la longueur du vecteur temps qui sépare les deux points de rupture, tel que

$$\eta_{i,a+\Delta a} = \eta_{i,a} + t.$$

On sait que le temps est égal au rapport entre la distance et la vitesse, autrement dit $t = d/v$. Dans notre cas,

$$\begin{aligned} d &= q_{a+\Delta a}(\eta_{i,a}) - q_a(\eta_{i,a}) = q_{a+\Delta a}(\eta_{i,a}) \\ v &= \left. \frac{\partial}{\partial t} q_{a+\Delta a}(t) \right|_{t=\eta_{i,a}} = f_{a+\Delta a}(\eta_{i,a}) - c(\eta_{i,a}). \end{aligned} \quad (4.30)$$

En utilisant un développement en série de Taylor, on exprime $q_{a+\Delta a}$ et $f_{a+\Delta a}$ en fonction de q_a et f_a :

$$\begin{aligned} q_{a+\Delta a}(\eta_{i,a}) &\approx q_a(\eta_{i,a}) + \frac{\partial}{\partial a} q_a(\eta_{i,a}) \Delta a + \mathcal{O}(\Delta a^2) = \frac{\partial}{\partial a} q_a(\eta_{i,a}) \Delta a + \mathcal{O}(\Delta a^2) \\ f_{a+\Delta a}(\eta_{i,a}) &\approx f_a(\eta_{i,a}) + \frac{\partial}{\partial a} f_a(\eta_{i,a}) \Delta a + \mathcal{O}(\Delta a^2). \end{aligned} \quad (4.31)$$

En remplaçant l'équation (4.31) dans (4.30), on a

$$\left. \frac{\partial q_{a+\Delta a}(t)}{\partial t} \right|_{t=\eta_{i,a}} \approx f_a(\eta_{i,a}) - c(\eta_{i,a}) + \frac{\partial}{\partial a} f_a(\eta_{i,a}) \Delta a + \mathcal{O}(\Delta a^2).$$

Ainsi,

$$t = \frac{q_{a+\Delta a}(\eta_{i,a})}{f_{a+\Delta a}(\eta_{i,a}) - c(\eta_{i,a})} \approx \frac{\frac{\partial}{\partial a} q_a(\eta_{i,a}) \Delta a + \mathcal{O}(\Delta a^2)}{f_a(\eta_{i,a}) - c(\eta_{i,a}) + f'_a(\eta_{i,a}) \Delta a + \mathcal{O}(\Delta a^2)}.$$

Finalement, on complète la preuve en appliquant la définition de la dérivée :

$$\begin{aligned} \frac{d\eta_{i,a}}{da} &= \lim_{\Delta a \rightarrow 0} \frac{1}{\Delta a} (\eta_{i,a+\Delta a} - \eta_{i,a}) \\ &= \lim_{\Delta a \rightarrow 0} \left(\frac{\frac{\partial}{\partial a} q_a(\eta_{i,a}) + \mathcal{O}(\Delta a^2)}{f_a(\eta_{i,a}) - c(\eta_{i,a}) + f'_a(\eta_{i,a}) \Delta a + \mathcal{O}(\Delta a^2)} \right) \\ &= \frac{\frac{\partial}{\partial a} q_a(\eta_{i,a})}{f_a(\eta_{i,a}) - c(\eta_{i,a})}. \end{aligned}$$

□

Lemme 8 Soit $\zeta_{i,a}$ le moment où la file nominale commence à se remplir. Alors

$$\frac{d\zeta_{i,a}}{da} = \begin{cases} \frac{\frac{\partial}{\partial a} f_a(\zeta_{i,a})}{\frac{\partial}{\partial t} f_a(\zeta_{i,a})} & \text{si } \zeta_{i,a} \text{ est endogène} \\ 0 & \text{si } \zeta_{i,a} \text{ est exogène.} \end{cases} \quad (4.32)$$

Preuve: Tout d'abord, rappelons qu'un événement exogène ne dépend pas de la valeur du paramètre a , donc si un tel événement provoque l'emplissage de la file d'attente, sa dérivée par rapport à a est nulle.

Que $\zeta_{i,a}$ soit endogène signifie que l'emplissage de la file est dû à l'augmentation du taux d'entrée qui devient supérieur à $c(t) = c_i$, comme illustré sur la figure 4.9. Comme dans la preuve du Lemme précédent, nous cherchons le vecteur temps $t = d/v$

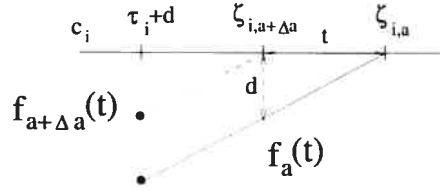


FIG. 4.9 – Enclenchement d'une période occupée endogène

tel que $\zeta_{i,a} = \zeta_{i,a+\Delta a} + t$, mais cette fois-ci, on considère la dynamique de $f_a(\cdot)$. On a

$$\begin{aligned} v &= \left. \frac{\partial}{\partial t} f_a(t) \right|_{t \in \{\zeta_{i,a+\Delta a}, \zeta_{i,a}\}} \\ d &= f_{a+\Delta a}(\zeta_{i,a}) - f_a(\zeta_{i,a}) \\ &\approx f_a(\zeta_{i,a}) + \frac{\partial}{\partial a} f_a(\zeta_{i,a}) \Delta a - f_a(\zeta_{i,a}). \end{aligned}$$

Finalement, pour $\zeta_{i,a}$ endogène,

$$t = \frac{\frac{\partial}{\partial a} f_a(\zeta_{i,a}) \Delta a}{\frac{\partial}{\partial t} f_a(\zeta_{i,a+\Delta a})} \Rightarrow \frac{d\zeta_{i,a}}{da} = \frac{\frac{\partial}{\partial a} f_a(\zeta_{i,a})}{\frac{\partial}{\partial t} f_a(\zeta_{i,a})}.$$

□

4.4.3 Dérivées des mesures de performance

Théorème 3 La dérivée IPA de $U(a)$ est donnée par

$$U'(a) = \frac{1}{T} \sum_{i=1}^{\mathcal{P}} \left(\frac{d\eta_{i,a}}{da} - \frac{d\zeta_{i,a}}{da} \right), \quad (4.33)$$

où \mathcal{P} est le nombre de périodes occupées dans $[0, T]$, $\frac{d\eta_{i,a}}{da}$ et $\frac{d\zeta_{i,a}}{da}$ sont donnés par Lemmes 7 et 8 respectivement.

Preuve: Tout d'abord, écrivons l'intégrale comme suit :

$$U(a) = \frac{1}{T} \int_0^T \mathbf{1}(q_a(t) > 0) dt = \frac{1}{T} \sum_{i=1}^{\mathcal{P}} \int_{\zeta_{i,a}}^{\eta_{i,a}} 1 dt = \frac{1}{T} \sum_{i=1}^{\mathcal{P}} (\eta_{i,a} - \zeta_{i,a}).$$

Avant de dériver, il faut noter que par Lemme 5, \mathcal{P} est localement indépendant de a . En effet, puisque deux événements ne peuvent pas se produire simultanément et comme les événements exogènes responsables de l'écoulement de la file ne dépendent pas du paramètre, il existe un voisinage de a dans lequel avec probabilité 1, le nombre de périodes occupées dans l'intervalle $[0, T]$ est constant. Par conséquent,

$$\frac{d}{da} U(a) = \frac{d}{da} \frac{1}{T} \sum_{i=1}^{\mathcal{P}} (\eta_{i,a} - \zeta_{i,a}) = \frac{1}{T} \sum_{i=1}^{\mathcal{P}} \left(\frac{d\eta_{i,a}}{da} - \frac{d\zeta_{i,a}}{da} \right).$$

□

Théorème 4 *La dérivée IPA de $Q(a)$ est donnée par*

$$Q'(a) = \frac{1}{T} \sum_{i=1}^{\mathcal{P}} \int_{\zeta_{i,a}}^{\eta_{i,a}} \frac{\partial}{\partial a} q_a(t) dt, \quad (4.34)$$

où \mathcal{P} est le nombre de périodes occupées.

Preuve: Pour calculer $Q'(a)$, nous allons nous servir du fait que les fonctions $q_a(t)$ et $\frac{\partial}{\partial a} q_a(t)$ sont intégrables en t pour tout a , et que la fonction $a \mapsto \frac{\partial}{\partial a} q_a(t)$ est absolument continue pour tout $t \in [0, T]$. Alors on peut échanger l'intégrale et la dérivée et écrire

$$Q'(a) = \frac{d}{da} \frac{1}{T} \int_0^T q_a(t) dt = \frac{d}{da} \frac{1}{T} \sum_{i=1}^{\mathcal{P}} \int_{\zeta_{i,a}}^{\eta_{i,a}} q_a(t) dt = \frac{1}{T} \sum_{i=1}^{\mathcal{P}} \int_{\zeta_{i,a}}^{\eta_{i,a}} \frac{\partial}{\partial a} q_a(t) dt, \quad (4.35)$$

où l'on s'est servi du fait que $q_a(\eta_{i,a}) = q_a(\zeta_{i,a}) = 0$.

Comme on ne connaît pas l'expression exacte de $q_a(t)$, nous allons dériver son équation différentielle par rapport à a , puis intégrer numériquement le processus dérivé

par rapport à t afin d'approximer $\frac{\partial}{\partial a} q_a(t)$. Les équations du processus dérivé sont :

$$\frac{\partial}{\partial a} \frac{\partial}{\partial t} q_a(t) = \begin{cases} \frac{\partial}{\partial a} f_a(t) - \frac{\partial}{\partial a} \mu_a(t) & \text{si } q_a(t) > 0 \\ 0 & \text{si } q_a(t) = 0, \end{cases} \quad (4.36)$$

où

$$\frac{\partial}{\partial a} \mu_a(t) = \begin{cases} 0 & \text{si } \mu_a(t) = c(t) \\ \frac{\partial}{\partial a} f_a(t) & \text{si } \mu_a(t) = f_a(t) \end{cases} \quad (4.37)$$

et

$$\frac{\partial}{\partial a} f_a(t) = (1 - p(q_a(t-d))) \frac{\partial}{\partial a} \mu_a(t-d) - \frac{\mu_a(t-d)}{E[c]} \left(b \frac{\partial}{\partial a} q_a(t-d) - 1 \right). \quad (4.38)$$

□

Ainsi, nous pouvons vérifier les Conditions 1 et 2 de la section 4.3 au sujet de $E[U(a)]$ et de $E[Q(a)]$ pour montrer que les estimateurs IPA de leurs dérivées sont non biaisés.

Théorème 5 *L'estimateur (4.33) est un estimateur non biaisé de la dérivée $\frac{d}{da} E[U(a)]$.*

Preuve: Puisque $q_a(t)$ est Lipschitz-continue a.p.1 (Lemme 6), l'intégrale de la fonction indicatrice l'est aussi et par conséquent, $U(a)$ est presque sûrement Lipschitz continue. De plus, sa dérivée par rapport à a existe a.p.1. Donc, $U'(a)$ est un estimateur non biaisé de $\frac{d}{da} E[U(a)]$.

□

Théorème 6 *L'estimateur (4.34), avec l'approximation de $\frac{\partial}{\partial a} q_a(t)$ donnée par les équations (4.36), (4.37) et (4.38), est un estimateur non biaisé de la dérivée $\frac{d}{da} E[Q(a)]$.*

Preuve: Puisque $q_a(t)$ est Lipschitz continue (Lemme 6), le résultat suit directement.

□

4.5 Procédure d'autoréglage et résultats

Les dérivées IPA trouvées dans les sections précédentes nous permettent d'écrire l'algorithme d'optimisation en ligne suivant, dont fait partie l'approximation stochastique :

1. Initialisation : $a = a_0, \epsilon = \epsilon_0$.
2. Simuler CLAMP S fois T secondes afin d'estimer les gradients pour la valeur de a actuelle (par exemple, $S = 10$ et $T = 100$).
3. Mettre à jour les paramètres : $a_{n+1} = a_n + \epsilon \nabla_a J(a_n)$.
(Si ϵ est décroissant, le mettre à jour aussi : $\epsilon_{n+1} = \epsilon_n/n$.)
4. $n = n + 1$; retourner à 2.

figure 4.10 montre la trajectoire de l'algorithme d'approximation stochastique appliquée à la simulation fluide de la file d'attente (intégration numérique de l'équation (4.4)) avec différentes suites pour les gains et la même valeur initiale de $a_0 = 0$.

Pour les expériences 1 à 3, le pas est constant, $\epsilon = 50$, $\epsilon = 110$ et $\epsilon = 150$ respectivement, alors que pour l'expérience 4, $\epsilon_n = \epsilon_0/n$, $\epsilon_0 = 400$. Lorsque le pas est constant, la vitesse de convergence peut être contrôlée en modifiant la taille de ce pas. On constate qu'un gain plus grand donne une convergence plus rapide, mais la trajectoire de l'algorithme est beaucoup plus variable. Les fluctuations des courbes 2 et 3 sont causées justement par l'utilisation d'un gain constant et assez grand. En effet, puisque les estimateurs des gradients ont une variance non nulle, l'erreur de l'estimation est amplifiée par le gain constant.

Par contre, la trajectoire de l'expérience avec pas décroissant est très lisse. De plus, la vitesse de convergence avec ce gain est remarquablement bonne, aussi rapide que l'algorithme avec $\epsilon = 150$ constant. Dans les deux cas, à peu près 10 itérations ont été nécessaires pour que l'algorithme converge vers la valeur optimale qui est $a \approx 12$. On peut donc conclure que pour cette application, la meilleure séquence de pas est $\epsilon_n = \epsilon_0/n$ et c'est probablement celle-ci qui sera implémentée dans les expériences futures avec

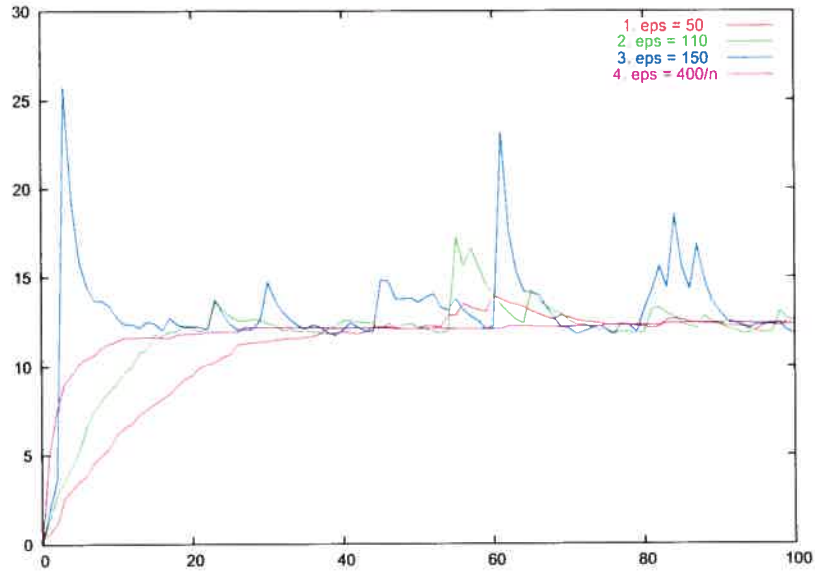


FIG. 4.10 – Trajectoire de l’algorithme d’approximation stochastique avec la simulation du modèle fluide de CLAMP. Sont données les valeurs du paramètre a en fonction du nombre d’itérations avec différentes valeurs du gain ϵ et $a_0 = 0.0$.

CLAMP.

Présentement, nous travaillons en collaboration avec nos collègues de l’Université de Melbourne sur l’implémentation du réglage dynamique avec l’algorithme de CLAMP lui-même. Nous n’avons pas encore de résultats, mais comme les expériences préliminaires se sont avérées prometteuses, nous croyons que le résultat final du projet sera fructueux.

CONCLUSION

Nous avons présenté trois applications de l'algorithme d'approximation stochastique par le gradient. Dans les trois cas, nous avons implanté des procédures intelligentes, faisant le contrôle en ligne des systèmes en question et l'autoréglage des paramètres.

La première application concerne le contrôle prédictif du réacteur Van der Vusse. Celui-ci est devenu un point de repère dans le domaine de l'automatique à cause de la difficulté du problème causée par les importantes non-linéarités du modèle. Le Chapitre 2 a présenté une méthodologie d'optimisation globale qui ajoute du bruit aux observations des gradients menant à une meilleure performance du modèle. Ceci a permis d'obtenir des estimateurs des gradients non biaisés avec seulement deux évaluations de la fonction objectif, peu importe la dimension du problème. Ainsi, nous avons remplacé la méthode SQP au niveau de la routine d'optimisation faisant partie du contrôle prédictif par la méthode d'optimisation globale SPSA. Le contrôle résultant est plus efficace en termes de temps de calcul, malgré que la vitesse de convergence est un peu moindre. Un problème avec le traitement des contraintes étant détecté, des recherches futures auraient pour but l'amélioration de cet aspect. Par exemple, au lieu de tronquer les variables de contrôle lorsqu'elles dépassent les bornes permises, il serait possible de faire la minimisation du Lagrangien augmenté, comme dans la méthode des multiplicateurs. Cette méthode pourrait être combinée avec la méthode SPSA pour donner un résultat satisfaisant.

Le Chapitre 3 a présenté une méthode de simulation appliquée au problème non encore résolu qu'est l'évaluation de la performance des réseaux sans fil CDMA. Puisqu'il n'existe pas d'autres méthodes pour résoudre ce problème, la simulation est devenue un moyen efficace pour tester des solutions approximatives. Nous avons proposé une méthode de simulation intelligente qui donne des résultats utiles en un temps raisonnable. L'optimisation des paramètres est faite en ligne, donc des simulations pilotes ne sont pas nécessaires. Notre algorithme est robuste : il peut facilement être adapté à des modèles avec contrôle de puissance imparfait, un trafic non homogène ou à intensité variable. Nous n'avons pas démontré que notre changement de mesure a une erreur

relative bornée, à cause de la difficulté que pose la déviation de l'interférence. Or, des avancées récentes ^[5, 28, 29] suggèrent des techniques permettant de dévier des distributions à queues épaisses, ce qui pourrait s'appliquer à notre problème. Nous pouvons aussi utiliser d'autres changements de mesure et dans ce cas, un arbre de simulation et une simulation accélérée pourraient être appliqués pour trouver les paramètres optimaux.

Dans le dernier chapitre, l'approximation stochastique a donné lieu à une procédure intelligente d'autoréglage, capable d'ajuster en ligne les paramètres de CLAMP, un algorithme de contrôle de flot des réseaux de télécommunication mobile. À l'aide d'un modèle fluide stochastique, nous avons trouvé les dérivées IPA de la fonction objectif, qui ont été incluses dans l'algorithme adaptatif. Des expériences à l'aide du modèle fluide ont montré que l'algorithme converge rapidement avec un gain décroissant. Nous travaillons présentement sur l'implémentation de l'algorithme intelligent avec la véritable procédure de CLAMP et les résultats s'annoncent prometteurs.

Comme en témoignent les applications que nous venons d'étudier, l'approximation stochastique donne lieu à des procédures d'autoréglage et d'optimisation en ligne efficaces et robustes. Leur existence est primordiale dans des modèles stochastiques, où l'optimisation déterministe n'est pas appropriée. L'importance de ces algorithmes adaptatifs est d'autant plus grande que le besoin de mécanismes intelligents se fait grandissant avec le développement des nouvelles technologies.

BIBLIOGRAPHIE

- [1] S. Andradóttir. "A Review of Simulation Optimization Techniques". In *Proceedings of the 1998 Winter Simulation Conference*, pages 151–158, 1998.
- [2] L. L. Andrew, S. V. Hanly, and R. G. Mukhtar. "CLAMP : A system to enhance the performance of wireless access networks". In *IEEE Globecom 2003*, pages 4142 – 4147, San Francisco, 2003.
- [3] L. L. Andrew, S. V. Hanly, and R. G. Mukhtar. "CLAMP : Differentiated Capacity Allocation in Access Network". In *IEEE International Performance Computing and Communications Conference*, Pheonix, USA, 2003.
- [4] L. L. Andrew, S. V. Hanly, and R. G. Mukhtar. "CLAMP : Maximizing the Performance of TCP over Low Bandwidth Variable Rate Access Links". Technical Report 2004-02-01, CUBIN, University of Melbourne, 2004.
- [5] S. Asmussen, D. P. Kroese, and R. Y. Rubinstein. "Heavy tails, importance sampling and cross-entropy". Presented in the 2nd International Workshop in Applied Probability, Piraeus, Greece, 2004.
- [6] F. Azadivar. "Simulation Optimization Methodologies". In *Proceedings of the 1999 Winter Simulation Conference*, pages 93–100, 1999.
- [7] I. Baltcheva, S. Cristea, F.J. Vázquez-Abad, and C. De Prada. "Simultaneous Perturbation Stochastic Approximation for Real-time Optimization of Model Predictive Control". In *Proceedings of the 1st Industrial Simulation Conference*, pages pp. 533–537, Valencia, Spain, June 2003.
- [8] A. Benveniste, M. Metivier, and P. Priouret. *Algorithmes adaptatifs et approximations stochastiques*. Masson, Paris, 1987.
- [9] D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, Mass. USA, 1999.
- [10] D. P. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, Mass. USA, 2000.

- [11] J. A. Bucklew. *Large Deviations Techniques in Decision Simulation and Estimation*. Wiley, New York, 1990.
- [12] E. F. Camacho and C. Bordons. *Model Predictive Control*. Springer, Berlin, New York, 1999.
- [13] C. G. Cassandras and S. Lafortune. *Introduction to Discrete Event Systems*. Kluwer Academic Publishers, Boston, 1999.
- [14] C. G. Cassandras, Y. Wardi, and al. "Perturbation Analysis for Online Control and Optimization of Stochastic Fluid Models". *IEEE Transactions on Automatic Control*, 47(8), August 2002.
- [15] C. Chan and S. Hanly. "Calculating the outage probability in a CDMA network with spatial Poisson traffic". *IEEE Transactions on Vehicular Technology*, 50(1) :183–204, 2001.
- [16] H. Chen and F. Allgower. "A quasi-infinite horizon nonlinear model predictive control scheme with guaranteed stability". In *Automatica*, 1998.
- [17] H. Chen, A. Kremling, and F. Allgower. "Nonlinear Predictive Control of a Benchmark CSTR". In *Proceedings of 3rd European Control Conference*, pages 3247–3252, 1995.
- [18] W. Colmenares, S. Cristea, C. De Prada, and al. "MLD Systems : Modeling and Control Experience with a Pilot Process". In *Proceedings of IMECE 2001*, New York, USA, 2001.
- [19] D. E. Comer. *Internetworking with TCP/IP, Principles Protocols and Architecture (4th edition)*. Prentice Hall, New Jersey, 2000.
- [20] D.P. Derevitskii and A.L. Fradkov. "Two Models for Analyzing the Dynamics of Adaptation Algorithms". *Automation and Remote Control*, 35(1) :59–67, 1974.
- [21] M. Duflo. *Random Iterative Models*. Springer, Berlin, 1997.
- [22] G. Fishman. *Monte Carlo concepts, algorithms and applications*. Springer-Verlag, New York, 1996.
- [23] D Foata and A. Fuchs. *Calcul des probabilités, 2ème édition*. Dunod, Paris, 1998.

- [24] A. Friedman. *Stochastic Differential Equations and Applications*. Academic Press, New York, 1975.
- [25] M.C. Fu and S.D Hill. "Optimization of Discrete Event Systems via Simultaneous Perturbation Stochastic Approximation". *IEEE Transactions*, 29 :233–243, 1997.
- [26] E.G. Gladyshev. "On Stochastic Approximation". *Theory Probab. Appl.*, 10(2) :275–278, 1965.
- [27] P. Glasserman. *Gradient Estimation via Perturbation Analysis*. Kluwer Academic Publishers, Boston, 1991.
- [28] P. Glasserman, P. Heidelberger, and P. Shahabuddin. "Portfolio Value-at-Risk with Heavy-Tailed Risk Factors". *Mathematical Finance*, 12(3) :239–269, 2002.
- [29] S. Juneja and P. Shahabuddin. "Simulating Heavy Tailed Processes using Delayed Hazard Rate Twisting". *ACM TOMACS*, 12(2) :94–118, 2002.
- [30] J. Kiefer and J. Wolfowitz. "Stochastic Estimation of the Maximum of a Regression Function". *Ann. Math. Stat.*, 23 :462–466, 1952.
- [31] S. R. Kulkarni and C. S. Horn. "An Alternative Proof for Convergence of Stochastic Approximation Algorithms". *IEEE Transactions on Automatic Control*, 41(3) :419–424, March 1996.
- [32] H.J. Kushner and D.S. Clark. *Stochastic Approximation for Constrained and Unconstrained Systems*. Springer-Verlag, Berlin and New York, 1978.
- [33] H.J. Kushner and J. Shwartz. "An Invariant Measure Approach to the Convergence of Stochastic Approximation Algorithms with State Dependent Noise". *SIAM J. Control and Opt.*, 22(1) :13–27, 1984.
- [34] H.J. Kushner and F.J. Vázquez-Abad. "Stochastic Approximation Methods for Systems Over an Infinite Horizon". *SIAM J. Control and Opt.*, 34(2) :712–756, 1996.
- [35] H.J. Kushner and G. Yin. *Stochastic Approximation Algorithms and Applications*. Springer-Verlag, New York, 1997.

- [36] P. L'Ecuyer. "A Unified View of the IPA, SF and LR Gradient Estimation Techniques". *Management Science*, 36(11) :1364 – 1382, November 1990.
- [37] P. L'Ecuyer, B. Martin, and F.J. Vázquez-Abad. "Functional Estimation for a Multicomponent Age Replacement Model (Special issue on Modern Digital Methodology)". *American Journal of Mathematical and Management Sciences*, 19(1 and 2) :135–156, December 1999.
- [38] S. Lipschutz. *Algèbre Linéaire (deuxième édition)*. McGraw-Hill, New York, 1991.
- [39] L. Ljung. "Analysis of recursive stochastic algorithms". *IEEE Transactions on Automatic Control*, 22(4) :551–575, 1977.
- [40] J.L. Maryak and D.C. Chin. "Global Random Optimization by Simultaneous Perturbation Stochastic Approximation". In *Proceedings of the American Control Conference*, pages 756–762, 2001.
- [41] E. Mosca. *Optimal, Predictive, and Adaptive Control*. Prentice Hall, 1995.
- [42] E. Mosca. "Model-Based Predictive Control". *EOLSS*, 2001.
- [43] R. G. Mukhtar, S. V. Hanly, and L. L. Andrew. "Efficient Internet Traffic Delivery over Wireless Networks". *IEEE Communications Magazine*, 41(12) :46 – 53, December 2003.
- [44] B. D. Ripley. *Stochastic Simulation*. John Wiley, USA, 1987.
- [45] H. Robbins and S. Monro. "A Stochastic Approximation Method". *Ann. Math. Stat.*, 22 :400–407, 1951.
- [46] S.M. Ross. *Simulation, 2ème édition*. Academic Press, USA, 1997.
- [47] P. Shahabuddin. "Importance sampling for the simulation of highly reliable Markovian systems". *Management Science*, 40(3) :333–352, 1994.
- [48] J. C. Spall. "Multivariate Stochastic Approximation Using a Simultaneous Perturbation Gradient Approximation". *IEEE Trans. on Automatic Control*, 3(37) :332–341, 1992.
- [49] J. C. Spall. "An Overview of the Simultaneous Perturbation Method for Efficient Optimization". *John Hopkins APL Technical Digest*, 19(4) :482–492, 1998.

- [50] Y. Su and M.Fu. "Importance Sampling in Derivative Securities Pricing". In K. Kang J.A. Joines, R.R. Barton and P.A. Fishwick, editors, *Proceedings of the 2000 Winter Simulation Conference*, pages 587–596, 1994.
- [51] J.R. Swisher, P.D. Hyden, S.H. Jacobson, and L.W. Schruben. "A Survey of Simulation Optimization Techniques and Procedures". In *Proceedings of the 2000 Winter Simulation Conference*, pages 119–128, 2000.
- [52] Y.Z. Tsypkin. *Adaptation and Learning in Automatic Systems*. Academic Press, New York, 1971.
- [53] F.J. Vázquez-Abad. A course on sensitivity analysis for gradient estimation of DES performance measures.
- [54] F.J. Vázquez-Abad. Lectures on the convergence of stochastic approximation algorithms, October 1999.
- [55] F.J. Vázquez-Abad and I. Baltcheva. "Intelligent Simulation for the Estimation of the Uplink Outage Probabilities in CDMA Networks". In *Proceedings of the 6-th International Workshop on Discrete Event Systems (WODES'02)*, pages 399–404, 2002.
- [56] F.J. Vázquez-Abad and I. Baltcheva. "Intelligent Simulation for the Estimation of the Uplink Outage Probabilities in CDMA Networks". Technical Report G-2002-69, Les cahiers du GERAD, Décembre 2002.
- [57] F.J. Vázquez-Abad and D. Dufresne. "Accelerated Simulation for Pricing Asian Options". In J.S. Carson D.J. Medeiros, E.F. Watson and M.S. Manivannan, editors, *Proceedings of the 1998 Winter Simulation Conference*, pages 1493–1500, 1998.
- [58] Y. Wardi and B. Melamed. "Loss Volume in Continuous Flow Models : Fast Simulation and Sensitivity Analysis". In *Proceedings of 8-th IEEE Mediterranean Conference on Control and Automation (MED-2000)*, Patras, Greece, 2000.
- [59] Y. Wardi and B. Melamed. "Variational Bounds and Sensitivity Analysis of Traffic Processes in Continuous Flow Models". *J. of Discrete Event Dynamic Systems*, 11(3) :249–282, 2001.

- [60] Y. Wardi, B. Melamed, C. G. Cassandras, and C. G. Panayiotou. “On-line IPA gradient estimators in stochastic continuous fluid models”. *J. Opt. Theor. Appl.*, 115(2) :369–405, 2002.
- [61] H. Yu and C. G. Cassandras. *Performance Evaluation and Planning Methods for the Next Generation Internet (to appear)*, chapter Perturbation Analysis and Multiplicative Feedback Control in Communication Networks. Kluwer Academic Publisher, 2005.

Annexe I

Annexe du Chapitre 3

Preuve: (Lemme 2) Dénotons par $\bar{\phi}(N) = E[\phi(N, Z) | N]$ l'espérance conditionnelle. Alors $E[\phi(N; Z)] = E[\bar{\phi}(N)]$. Si on multiplie et on divise par la mesure étoilée, on obtient :

$$\begin{aligned}
 E[\bar{\phi}(N)] &= \sum_{k=0}^K \sum_{n_k=0}^{\infty} \bar{\phi}(n_1, \dots, n_K) \times \prod_{k=0}^K \left(\frac{P(N = n_k)}{P(N^* = n_k)} P(N^* = n_k) \right) \\
 &= \sum_{k=0}^K \sum_{n_k=0}^{\infty} \bar{\phi}(n_1, \dots, n_K) \times \prod_{k=0}^K \left(\frac{e^{-\lambda_k} \frac{\lambda_k^{n_k}}{n_k!}}{e^{-\theta_k} \frac{\theta_k^{n_k}}{n_k!}} \left[\frac{e^{-\theta_k} \theta_k^{n_k}}{n_k!} \right] \right) \\
 &= \sum_{k=0}^K \sum_{n_k=0}^{\infty} \bar{\phi}(n_1, \dots, n_K) \times \prod_{k=0}^K \left(e^{-(\lambda_k - \theta_k)} \left(\frac{\lambda_k}{\theta_k} \right)^{n_k} \left[\frac{e^{-\theta_k} \theta_k^{n_k}}{n_k!} \right] \right) \\
 &= \sum_{k=0}^K \sum_{n_k=0}^{\infty} \bar{\phi}(n_1, \dots, n_K) \times \prod_{k=0}^K P(N^* = n_k) \times e^{-\sum_k (\lambda_k - \theta_k)} \times \prod_{k=0}^K \left(\frac{\lambda_k}{\theta_k} \right)^{n_k} \\
 &= E \left[\bar{\phi}(N^*) e^{-\sum_k (\lambda_k - \theta_k)} \prod_k \left(\frac{\lambda_k}{\theta_k} \right)^{N_k^*} \right], \tag{I.1}
 \end{aligned}$$

où $N_k^* \sim \text{Poisson}(\theta_k)$. De même, par indépendance des variables aléatoires normales et en utilisant la notation $d = \sum_{k=1}^K N_k$ et $z = (z_{i,k}) \in \mathbb{R}^d$, il s'en suit :

$$\begin{aligned}
\mathbb{E}[\phi(N; Z) | N] &= \int_{\mathbb{R}^d} \phi(N; z) \mathbb{P}(\phi(N; Z) | N) dz \\
&= \int_{\mathbb{R}^d} \phi(N; z) \frac{\mathbb{P}(\phi(N; Z) | N)}{\mathbb{P}(\phi(N; Z^*) | N)} \mathbb{P}(\phi(N; Z^*) | N) dz \\
&= \int_{\mathbb{R}^d} \phi(N; z) \prod_{k=1}^K \prod_{i=1}^{N_k} \left(\frac{e^{z_{i,k}^2/2\sigma^2}}{e^{(z_{i,k}-\mu_k)^2/2\sigma^2}} \times \frac{1}{\sqrt{2\pi\sigma^2}} e^{(z_{i,k}-\mu_k)^2/2\sigma^2} \right) dz \\
&= \int_{\mathbb{R}^d} \phi(N; z) e^{\left\{ \sum_{k=1}^K \sum_{i=1}^{N_k} \left(\frac{\mu_k^2 - 2\mu_k z_{i,k}}{2\sigma^2} \right) \right\}} \prod_{k=1}^K \prod_{i=1}^{N_k} \frac{1}{\sqrt{2\pi\sigma^2}} e^{(z_{i,k}-\mu_k)^2/2\sigma^2} dz \\
&= \mathbb{E} \left[\exp \left\{ \sum_{k=1}^K \sum_{i=1}^{N_k} \left(\frac{\mu_k^2 - 2\mu_k Z_{i,k}^*}{2\sigma^2} \right) \right\} \phi(N; Z^*) \middle| N \right] = \bar{\phi}(N), \quad (I.2)
\end{aligned}$$

où pour tout k , $Z_{i,k}^*$ sont i.i.d. $\mathcal{N}(\mu_k, \sigma^2)$. En remplaçant $\bar{\phi}(N^*)$ dans (I.1) par (I.2), on obtient le résultat :

$$\begin{aligned}
\mathbb{E}[\bar{\phi}(N)] &= \mathbb{E}[\mathbb{E}[\phi(N, Z) | N]] \\
&= \mathbb{E} \left[\bar{\phi}(N^*) e^{-\sum_k (\lambda_k - \theta_k)} \prod_k \left(\frac{\lambda_k}{\theta_k} \right)^{N_k^*} \right] \\
&= \mathbb{E} \left[\mathbb{E} \left[e^{\sum_{k=1}^K \sum_{i=1}^{N_k^*} \left(\frac{\mu_k^2 - 2\mu_k Z_{i,k}^*}{2\sigma^2} \right)} \phi(N^*; Z^*) \middle| N^* \right] e^{-\sum_k (\lambda_k - \theta_k)} \prod_k \left(\frac{\lambda_k}{\theta_k} \right)^{N_k^*} \right] \\
&= e^{\sum_{k=1}^K \sum_{i=1}^{N_k^*} \left(\frac{\mu_k^2 - 2\mu_k Z_{i,k}^*}{2\sigma^2} \right)} e^{-\sum_k (\lambda_k - \theta_k)} \prod_k \left(\frac{\lambda_k}{\theta_k} \right)^{N_k^*} \mathbb{E}[\mathbb{E}[\phi(N^*; Z^*) | N^*]] \\
&= \exp \left\{ \sum_{k=1}^K \left((\theta_k - \lambda_k) + \frac{N_k^* \mu_k^2}{2\sigma^2} - \frac{\mu_k}{\sigma_k^2} \sum_{i=1}^{N_k^*} Z_{i,k}^* \right) \right\} \prod_{k=1}^K \left(\frac{\lambda_k}{\theta_k} \right)^{N_k^*} \mathbb{E}[\phi(N^*; Z^*)] \\
&= L(N^*; Z^*) \mathbb{E}[\phi(N^*; Z^*)] \\
&= \mathbb{E}[L(N^*; Z^*) \phi(N^*; Z^*)].
\end{aligned}$$

□