



HAL
open science

Contrôle optimal stochastique des processus de Markov déterministes par morceaux et application à l'optimisation de maintenance

Alizée Geeraert

► **To cite this version:**

Alizée Geeraert. Contrôle optimal stochastique des processus de Markov déterministes par morceaux et application à l'optimisation de maintenance. Mathématiques générales [math.GM]. Université de Bordeaux, 2017. Français. NNT : 2017BORD0602 . tel-01557969

HAL Id: tel-01557969

<https://theses.hal.science/tel-01557969>

Submitted on 6 Jul 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

présentée pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE BORDEAUX

par

Alizée GEERAERT

ÉCOLE DOCTORALE DE MATHÉMATIQUES ET D'INFORMATIQUE
SPÉCIALITÉ : Mathématiques appliquées et calcul scientifique

**Contrôle optimal stochastique des processus de Markov
déterministes par morceaux et application à l'optimisation de
maintenance**

Thèse soutenue le 6 juin 2017, devant le jury composé de

| | | |
|--------------------|-------------------------------------|------------------------|
| Madalina DEACONU | INRIA Nancy-Grand Est | Rapportrice |
| Antoine GRALL | Université de technologie de Troyes | Rapporteur |
| Olivier GAUDOIN | Institut Polytechnique de Grenoble | Président du jury |
| Michel PRENAT | Thales Optronique | Examineur |
| Huilong ZHANG | Université de Bordeaux | Examineur |
| Camille BAYSSE | Thales Optronique | Encadrante |
| François DUFOUR | Institut Polytechnique de Bordeaux | Directeur de thèse |
| Benoîte DE SAPORTA | Université de Montpellier | Co-directrice de thèse |

Titre Contrôle optimal stochastique des processus de Markov déterministes par morceaux et application à l'optimisation de maintenance

Résumé On s'intéresse au problème de contrôle impulsionnel à horizon infini avec facteur d'oubli pour les processus de Markov déterministes par morceaux (PDMP). Dans un premier temps, on modélise l'évolution d'un système opto-électronique par des PDMP. Afin d'optimiser la maintenance du système, on met en place un problème de contrôle impulsionnel tenant compte à la fois du coût de maintenance et du coût lié à l'indisponibilité du matériel auprès du client. On applique ensuite une méthode d'approximation numérique de la fonction valeur associée au problème, faisant intervenir la quantification de PDMP. On discute alors de l'influence des paramètres sur le résultat obtenu. Dans un second temps, on prolonge l'étude théorique du problème de contrôle impulsionnel en construisant de manière explicite une famille de stratégies ε -optimales. Cette construction se base sur l'itération d'un opérateur dit de simple-saut-ou-intervention associé au PDMP, dont l'idée repose sur le procédé utilisé par U.S. Gugerli pour la construction de temps d'arrêt ε -optimaux. Néanmoins, déterminer la meilleure position après chaque intervention complique significativement la construction de telles stratégies et nécessite l'introduction d'un nouvel opérateur. L'originalité de la construction de stratégies ε -optimales présentée ici est d'être explicite, au sens où elle ne nécessite pas la résolution préalable de problèmes complexes.

Mots-clés Processus de Markov déterministes par morceaux, contrôle impulsionnel, stratégies optimales, programmation dynamique, optimisation, quantification, modélisation

Title Stochastic optimal control for piecewise deterministic Markov processes and application to maintenance optimization

Abstract We are interested in a discounted impulse control problem with infinite horizon for piecewise deterministic Markov processes (PDMPs). In the first part, we model the evolution of an optronic system by PDMPs. To optimize the maintenance of this equipment, we study an impulse control problem where both maintenance costs and the unavailability cost for the client are considered. We next apply a numerical method for the approximation of the value function associated with the impulse control problem, which relies on quantization of PDMPs. The influence of the parameters on the numerical results is discussed. In the second part, we extend the theoretical study of the impulse control problem by explicitly building a family of ε -optimal strategies. This approach is based on the iteration of a single-jump-or-intervention operator associated to the PDMP and relies on the theory for optimal stopping of a piecewise-deterministic Markov process by U.S. Gugerli. In the present situation, the main difficulty consists in approximating the best position after the interventions, which is done by introducing a new operator. The originality of the proposed approach is the construction of ε -optimal strategies that are explicit, since they do not require preliminary resolutions of complex problems.

Keywords Piecewise deterministic Markov process, impulse control, optimal strategies, dynamic programming, optimization, quantization, modeling

Unité de recherche :
Inria Bordeaux Sud-Ouest - 200818243Z
200, avenue de la Vieille Tour,
33405 TALENCE

Table des matières

| | |
|---|-----------|
| Introduction générale | 1 |
| 1 Processus de Markov déterministes par morceaux (PDMP) | 5 |
| 1.1 Introduction | 5 |
| 1.2 Notations | 6 |
| 1.3 Définition d'un processus de Markov déterministe par morceaux | 7 |
| 1.4 Hypothèses de régularité | 9 |
| 1.5 Exemples | 10 |
| 1.6 Problème de contrôle impulsionnel | 12 |
| 1.6.1 Arrêt optimal pour des PDMP | 13 |
| 1.6.2 Contrôle impulsionnel pour des PDMP | 15 |
| 1.7 Quantification | 21 |
| 1.7.1 Quantification d'une variable aléatoire | 21 |
| 1.7.2 Quantification d'une chaîne de Markov | 23 |
| 1.8 Approximation numérique de la fonction valeur | 25 |
| 1.8.1 Définitions | 26 |
| 1.8.2 Algorithmes | 30 |
| 1.9 Bilan | 35 |
| I Optimisation de la maintenance d'un équipement opto-électronique | 37 |
| 2 Système opto-électronique et modélisation | 39 |
| 2.1 Caméra Sophie | 40 |
| 2.1.1 Descriptif de la caméra | 40 |
| 2.1.2 États de fonctionnement et de panne de la caméra | 42 |
| 2.1.3 Maintenances et coûts | 43 |
| 2.1.4 Problématique industrielle | 44 |
| 2.2 Régimes de la caméra Sophie | 47 |
| 2.2.1 Transition entre les états des éléments de la caméra | 47 |
| 2.2.2 Définition des régimes de la caméra | 49 |
| 2.3 Processus de Markov déterministe par morceaux | 52 |
| 2.4 Problème de contrôle impulsionnel | 55 |
| 2.4.1 Ensemble de contrôle | 56 |
| 2.4.2 Coûts associés | 58 |
| 2.4.3 Facteur d'oubli | 62 |
| 2.5 Conclusion | 63 |
| 3 Simulations : trajectoires et calcul de coûts | 65 |
| 3.1 Politiques d'intervention de référence | 65 |
| 3.2 Simulations de trajectoires | 67 |
| 3.2.1 Généralités sur la simulation de trajectoires | 68 |

| | | |
|-----------|---|------------|
| 3.2.2 | Caméra sans intervention (politique d'intervention n°1) | 69 |
| 3.2.3 | Caméra avec politique d'intervention n°2, n°3 et n°4 | 73 |
| 3.3 | Coûts d'une politique d'intervention | 77 |
| 3.4 | Choix des grilles de discrétisation pour \mathbb{U}_{fini} | 80 |
| 3.5 | Conclusion | 81 |
| 4 | Approximation numérique de la fonction valeur | 83 |
| 4.1 | Quantification du PDMP modélisant l'évolution de la caméra | 84 |
| 4.1.1 | Recherche du plus proche voisin de X_i^{s+1} dans la grille Γ_i^s | 84 |
| 4.1.2 | Initialisation des grilles de quantification et des compteurs associés | 85 |
| 4.1.3 | Choix des variables S , $N_{quantif}$, N et de la suite (δ_s) | 86 |
| 4.2 | Mise en place de la méthode d'approximation de la fonction valeur | 88 |
| 4.2.1 | Horizon de contrôle N_0 | 89 |
| 4.2.2 | Fonction g | 89 |
| 4.2.3 | Fonction F | 90 |
| 4.2.4 | Discrétisation du temps | 90 |
| 4.3 | Résultats autour de l'approximation de la fonction valeur | 91 |
| 4.3.1 | Choix de la constante de pénalisation | 91 |
| 4.3.2 | Discrétisation du temps | 92 |
| 4.3.3 | Influence de la quantification | 94 |
| 4.3.4 | Approximation de la fonction valeur pour différents ensembles de contrôle | 97 |
| 4.3.5 | Approximation de la fonction valeur pour différents coûts d'intervention | 98 |
| 4.3.6 | Évolution de l'horizon de contrôle | 100 |
| 4.4 | Nouveau modèle | 102 |
| 4.4.1 | Définition du modèle | 102 |
| 4.4.2 | Politiques d'intervention | 106 |
| 4.4.3 | Approximation de la fonction valeur | 108 |
| 4.5 | Bilan | 110 |
| II | Stratégies optimales de contrôle impulsif pour des processus Markoviens déterministes par morceaux | 113 |
| 5 | Construction de stratégies ε-optimales | 115 |
| 5.1 | Construction d'approximation des fonctions valeurs | 116 |
| 5.1.1 | Propriété de monotonie des opérateurs K et J | 117 |
| 5.1.2 | Approximation des fonctions valeurs | 118 |
| 5.2 | Construction de stratégies d'intervention explicites | 120 |
| 5.3 | Coût des trajectoires contrôlées | 127 |
| 5.3.1 | Majoration du temps éventuel d'intervention | 127 |
| 5.3.2 | Distributions de la première position et du premier temps de saut pour les processus contrôlés et non contrôlés | 128 |
| 5.3.3 | Propriété de Markov des temps d'intervention | 129 |
| 5.3.4 | Coût associé aux stratégies explicites construites | 131 |
| 5.4 | Stratégies ε -optimales | 138 |
| | Conclusion et perspectives | 139 |
| | Annexe | 141 |
| | Bibliographie | 143 |

Introduction

Une entreprise doit savoir s'adapter pour évoluer avec le marché et ne pas se laisser dépasser par ses concurrents. Pour cela, elle doit prendre de nombreuses décisions avec l'information disponible au moment de la prise de décision. Il peut s'agir d'un changement de technologies, d'investissements dans des projets spécifiques ou bien de gestion de stock. À chaque problème rencontré, les possibilités sont nombreuses et le choix à faire est crucial pour la pérennité de l'entreprise : doit-on investir immédiatement dans le développement d'un nouveau produit ? Est-ce plus rentable de stocker moins de pièces qu'actuellement au risque de se trouver en rupture de stock ? Ou doit-on augmenter le stock et engendrer par là même un coût supplémentaire pour l'entreprise ?

Ces exemples font partie de ce que l'on appelle le contrôle impulsionnel, historiquement introduit dans Bensoussan et Lions (1975). Le principe du contrôle impulsionnel est le suivant. Un processus évolue avec le temps, en faisant intervenir une part d'aléatoire. Une personne doit prendre des décisions en se fondant uniquement sur les informations disponibles, afin de maximiser un gain ou minimiser un coût. Il existe de très nombreux problèmes de contrôle impulsionnel. Ils constituent probablement l'une des plus importantes classes de problèmes d'optimisation appliqués.

Plus précisément, le problème de contrôle impulsionnel se définit comme suit. On appelle contrôleur la personne chargée des prises de décisions. En considérant un processus aléatoire $\{X_t\}$, le contrôleur introduit des sauts dans le processus à des moments spécifiques, représentés par des temps d'arrêt, où il déplace le processus instantanément en un nouvel endroit, représenté par une variable aléatoire. La suite de ces temps d'arrêts couplés aux variables aléatoires définissant les nouvelles positions du processus après une intervention est appelée stratégie de contrôle impulsionnel. L'objectif du problème de contrôle impulsionnel est de choisir une stratégie qui minimise la fonction coût correspondant à l'espérance de la somme du coût continu et des coûts d'intervention jusqu'à l'infini. L'infimum de ce coût est appelé la fonction valeur du problème. Dans le cas où une telle stratégie n'existe pas, le but est de trouver une stratégie qui s'approche au mieux de la fonction valeur.

La catégorie la plus simple de problème de contrôle impulsionnel est celle des problèmes d'arrêt optimal (Hill (2009); Berman (1994); Pham (1997); Davis (1993)). Ce type de problème consiste à trouver une seule date d'arrêt. On calcule le gain (ou le coût) obtenu par rapport à la position du processus au moment de l'arrêt. Un exemple de problème d'arrêt optimal est la gestion d'actions en bourse : on cherche à vendre ses actions au meilleur moment (temps d'arrêt optimal) pour maximiser ses profits (fonction gain), sans pour autant savoir a priori comment le cours de la bourse va évoluer dans les prochaines semaines. Les problèmes d'arrêt optimal sont très importants car ils constituent la première étape dans la résolution de problèmes de contrôle impulsionnel plus généraux.

Le contrôle impulsionnel peut être divisé en de nombreuses classes : à horizon fini ou infini, avec un processus complètement observé ou partiellement observé,... Le cadre de notre étude est celui du contrôle impulsionnel avec horizon infini et facteur d'oubli, dans lequel on minimise une fonction coût. L'horizon infini signifie que l'on tient compte du processus $\{X_t\}$ sur toute sa durée de vie. En particulier, la fonction coût est calculée sur toute la trajectoire du processus $\{X_t\}$. Dans ce contexte, en supposant que le coût continu et les coûts d'intervention sont finis,

le facteur d'oubli permet de s'assurer que la fonction coût est finie. En effet, vu que l'on calcule le coût en fonction de toute la trajectoire du processus, le facteur d'oubli permet d'atténuer le coût du processus au fur et à mesure que le temps passe, via une exponentielle décroissante. Ainsi, la fonction coût est nécessairement finie.

Les processus aléatoires choisis dans le cadre de notre étude sont les processus de Markov déterministes par morceaux (PDMP, *Piecewise Deterministic Markov Processes*), introduits pour la première fois par Davis (1984). Ce sont des processus hybrides non-diffusifs. Leur caractère hybride vient du fait qu'ils contiennent deux variables : l'une discrète, appelée mode ou régime, permettant de décrire l'environnement dans lequel évolue le processus, l'autre continue, appelée variable euclidienne, correspondant généralement à des grandeurs physiques.

Le comportement des PDMP peut être décrit de la manière suivante. Partant d'un point de l'espace d'état, le processus suit une trajectoire déterministe jusqu'au premier instant de saut. Ce dernier est déterminé soit selon une loi donnée, soit de manière déterministe quand le processus atteint la frontière de l'espace d'état. La nouvelle position du processus est sélectionnée de manière aléatoire et le PDMP suit une trajectoire déterministe jusqu'au temps de saut suivant, et ainsi de suite. Finalement, une trajectoire de PDMP est composée de parties de trajectoires déterministes, ponctuées par des sauts aléatoires.

Les PDMP présentent deux propriétés qui les rendent particulièrement intéressants. D'une part, pour chaque PDMP, il existe un espace de probabilité sous lequel le processus vérifie la propriété de Markov forte. D'autre part, un PDMP peut-être entièrement décrit par une chaîne de Markov $(Z_n, S_n)_{n \geq 0}$ à temps discret, où la suite (Z_n) représente les positions après-saut et la suite (S_n) correspond au temps inter-sauts du processus. Ces deux propriétés seront utilisées tout au long des démonstrations de cette thèse.

Les processus de Markov déterministes par morceaux permettent de modéliser le comportement de nombreux systèmes Davis (1993). Ils sont aujourd'hui très utilisés dans le domaine médical Bertail *et al.* (2008), les modèles de biologie Pakdaman *et al.* (2010) ou encore le domaine de la fiabilité et de la maintenance des systèmes de Saporta *et al.* (2015). C'est dans ce dernier domaine que se place l'application industrielle de cette thèse, pour lequel on met en place un problème de contrôle impulsif pour des processus de Markov déterministes par morceaux (voir Costa et Davis (1989); Davis (1993); de Saporta *et al.* (2015)).

Cette thèse est le fruit d'une étroite collaboration avec l'industriel Thales, qui a pour objectif d'optimiser la maintenance d'une caméra thermique. Soucieux d'offrir la meilleure disponibilité des caméras auprès de ses clients tout en maintenant des coûts de maintenance raisonnables, Thales développe et met en oeuvre de nouvelles technologies basées sur des HUMS (*Health & Usage Monitoring System*) permettant d'avoir un aperçu de l'état de santé des éléments qui composent la caméra, au travers de l'enregistrement de données de fonctionnement, d'environnement et de conditions d'utilisation.

Ainsi, une grande quantité d'information sur l'état de santé de la caméra est disponible à tout instant. Il est alors nécessaire de traiter ces informations afin d'en extraire des directives sur la maintenance de l'équipement (date de maintenance conseillée) et des renseignements sur l'utilisation future de l'équipement (potentiel de vie). Le traitement et l'analyse de ces données nécessite le développement de modèles mathématiques.

L'objectif de cette thèse est d'apporter des outils mathématiques pour déterminer, à n'importe quel instant de la vie de la caméra, une date de maintenance optimale, mais également déterminer les opérations de maintenance à effectuer sur la caméra lors d'un retour à l'atelier. Pour ce faire, on modélise l'évolution d'une caméra à l'aide d'un processus de Markov déterministe par morceaux et on définit le problème de contrôle impulsif correspondant. Dans un premier temps, la méthode d'approximation numérique de la fonction valeur développée dans de Saporta et Dufour (2012) est mise en place. Dans un second temps, on poursuit le développement de la théorie du contrôle impulsif pour des PDMP en construisant des stratégies

ε -optimales, c'est-à-dire des stratégies dont le coût associé approche le minimum de la fonction coût.

Ce mémoire de thèse est composé d'une partie appliquée et d'une partie théorique, toutes deux traitant du contrôle impulsionnel pour les PDMP. En préambule à ces deux parties, le chapitre 1 expose la définition des processus de Markov déterministes par morceaux, ainsi que celle du contrôle impulsionnel. Les hypothèses nécessaires à notre travail y sont détaillées. On expose ensuite la méthode d'approximation numérique développée dans de Saporta et Dufour (2012), qui sera mise en œuvre dans le cadre de l'optimisation de la maintenance d'un système opto-électronique.

L'approximation numérique de la fonction valeur associée au problème industriel constitue la première partie de cette thèse. Le descriptif de la caméra et de son fonctionnement, ainsi que la problématique industrielle sont présentés au chapitre 2. On y détaille également la modélisation de la caméra thermique étudiée et son fonctionnement. Enfin, le problème de contrôle impulsionnel lié à la problématique industrielle est mis en place. Par la suite, au chapitre 3, après avoir établi des politiques d'intervention de référence, on discute des simulations des trajectoires représentant l'évolution d'une caméra soumise à une politique d'intervention donnée. On définit également le coût des politiques et on le calcule pour les politiques de référence. Ceci nous donnera un critère pour définir un ensemble de contrôle fini.

Le chapitre 4 concerne la mise en œuvre des algorithmes de quantification et ceux de l'approximation numérique de la fonction valeur. On y détaille les précautions nécessaires à prendre lors de la quantification, afin de tenir compte de la spécificité de la présence d'une variable discrète dans le processus de Markov déterministe par morceaux. On discute alors du choix des variables liées à la méthode numérique d'approximation de la fonction valeur, avant d'exposer les différents résultats obtenus.

Dans toute la partie I, les valeurs choisies dans les applications sont des valeurs arbitraires prises à titre d'exemple dans le seul but de présenter et commenter les résultats des méthodes mises en œuvre.

La seconde partie de la thèse, correspondant au chapitre 5, est consacrée au développement de la théorie du contrôle impulsionnel pour des processus de Markov déterministes par morceaux. On y développe une construction explicite de stratégies ε -optimales. On commence par construire une approximation des fonctions valeurs associée au problème de contrôle impulsionnel avec un horizon de contrôle fini, puis on construit une famille de stratégies. On montre alors que la famille des stratégies construites est ε -optimales.

On termine par une conclusion, synthétisant les principaux résultats de cette thèse et présentant quelques perspectives théoriques et appliquées possibles à ce travail.

Chapitre 1

Processus de Markov déterministes par morceaux (PDMP)

Sommaire

| | | |
|------------|--|-----------|
| 1.1 | Introduction | 5 |
| 1.2 | Notations | 6 |
| 1.3 | Définition d'un processus de Markov déterministe par morceaux | 7 |
| 1.4 | Hypothèses de régularité | 9 |
| 1.5 | Exemples | 10 |
| 1.6 | Problème de contrôle impulsif | 12 |
| 1.6.1 | Arrêt optimal pour des PDMP | 13 |
| 1.6.2 | Contrôle impulsif pour des PDMP | 15 |
| 1.7 | Quantification | 21 |
| 1.7.1 | Quantification d'une variable aléatoire | 21 |
| 1.7.2 | Quantification d'une chaîne de Markov | 23 |
| 1.8 | Approximation numérique de la fonction valeur | 25 |
| 1.8.1 | Définitions | 26 |
| 1.8.2 | Algorithmes | 30 |
| 1.9 | Bilan | 35 |

1.1 Introduction

Les processus de Markov déterministes par morceaux (PDMP, pour *piecewise deterministic Markov process*) ont été introduits par M.H.A. Davis dans Davis (1984, 1993) comme une classe générale de modèles stochastiques hybrides. Ces processus ont deux variables. La première variable, appelée mode ou régime, est discrète et caractérise l'environnement et/ou le régime dans lequel le processus évolue. Elle permet de spécifier par exemple si le système fonctionne ou est en panne. La seconde, appelée variable euclidienne, est à valeurs dans un espace continu et représente les grandeurs physiques du modèle. Il peut s'agir d'une vitesse, d'une position ou de temps de fonctionnement. La trajectoire d'un PDMP consiste en des trajectoires déterministes, ponctuées par de sauts aléatoires. Le processus dépend de trois caractéristiques locales : le flot, l'intensité de saut et le noyau de Markov.

L'idée du comportement du PDMP est la suivante. Partant d'un point de l'espace d'état, le processus suit une trajectoire déterministe, donnée par le flot, jusqu'au premier temps de saut. Ce temps est déterminé soit par une loi liée à l'intensité de saut, soit de manière déterministe quand le processus atteint la frontière de l'espace d'état. Après ce saut, la nouvelle position

et le nouveau régime du processus sont déterminés par le noyau de Markov. La trajectoire du processus est ensuite déterministe jusqu'au temps de saut suivant. Une nouvelle position pour le processus est alors déterminée, et ainsi de suite.

Un certain nombre de résultats ont été obtenus sur les PDMP. Un aperçu de ces résultats est donné dans Azaïs *et al.* (2014a). On peut citer des résultats de stabilité et d'ergodicité Dufour et Costa (1999); Costa et Dufour (2008), de convergence Faggionato *et al.* (2010); Chafaï *et al.* (2010); Bardet *et al.* (2013), d'estimation non paramétrique Azaïs (2014); Azaïs *et al.* (2014b), d'arrêt optimal Gugerli (1986) et de contrôle Lenhart (1989); Dempster et Ye (1995). On pourra se référer à la partie 1.6 pour plus de références dans ces deux derniers domaines. Des applications numériques ont également été développées. Elles sont basées sur la discrétisation du noyau de Markov associé au PDMP ou sur la discrétisation de la chaîne de Markov sous-jacente du PDMP au moyen de la quantification, procédé défini au paragraphe 1.7. Ces méthodes permettent de résoudre des problèmes tels que l'arrêt optimal et le contrôle impulsif Gugerli (1986); de Saporta *et al.* (2010); Costa et Davis (1988); de Saporta et Dufour (2012); Brandejsky *et al.* (2012b,a).

Les processus de Markov déterministes par morceaux ont de très nombreuses applications dans des domaines variés. Les premières historiquement, et les plus simples, sont des problèmes classiques de file d'attente Davis (1984), tels que le temps d'attente pour une file à un serveur. Dans ces modèles, le flot qui détermine la trajectoire du processus entre deux sauts est linéaire. Un exemple de ce type de modèle est donné en cinquième partie de ce chapitre (modèle du TCP). Naturellement, les PDMP ont ensuite été appliqués à des problèmes, toujours classiques, où le flot n'est plus nécessairement linéaire, mais peut être régi par des équations différentielles. Parmi ces problèmes, on trouve des problèmes d'inventaire, de maintenance/remplacement et de stockage/sortie, dont quelques exemples sont donnés dans les articles Bäuerle et Rieder (2011); Dassios et Embrechts (1989) et dans le livre Davis (1993). Un exemple typique de ces problèmes est celui du barrage. L'afflux d'eau dans le barrage constitue la variable continue du processus et les lâchers d'eau la variable discrète donnant lieu aux sauts Davis (1984).

Plus récemment, dans le domaine de la biologie, les PDMP ont été appliqués à des modèles neuronaux Pakdaman *et al.* (2010), en biologie moléculaire Crudu *et al.* (2012) et en dynamique des populations Erban et Othmer (2004); Costa (2016) pour ne citer que quelques exemples. La fiabilité des systèmes est également un vaste champ d'applications des PDMP de Saporta *et al.* (2015).

Ce chapitre est décomposé de la manière suivante : après avoir défini quelques notations dans la partie 1.2, on donne la définition d'un processus de Markov déterministe par morceaux dans la partie 1.3. Les hypothèses de régularité nécessaires à notre travail sont données dans la partie 1.4. Quelques exemples classiques des PDMP sont ensuite exposés au paragraphe 1.5. Dans la partie 1.6, on présente les problèmes de contrôle impulsif pour les processus de Markov déterministes par morceaux, ainsi que les principaux résultats existants. On consacre les deux paragraphes suivants à la méthode d'approximation de la fonction valeur développée dans de Saporta et Dufour (2012), en exposant au paragraphe 1.7 le principe de quantification et en détaillant la méthode d'approximation numérique de la fonction valeur dans le paragraphe 1.8. On termine ce chapitre par une conclusion au paragraphe 1.9.

1.2 Notations

Pour $a, b \in \mathbb{R}$, $a \wedge b = \min(a, b)$ désigne le minimum de a et de b . On notera $a \vee b = \max(a, b)$ le maximum de a et de b . Par convention, on pose $\inf \emptyset = \infty$.

Soit \mathcal{X} un espace métrique muni de la distance $d_{\mathcal{X}}$. Pour un sous-ensemble A de \mathcal{X} , ∂A est la frontière de A et \bar{A} sa fermeture. On note $\mathfrak{B}(\mathcal{X})$ la tribu borélienne de \mathcal{X} et $\mathbf{B}(\mathcal{X})$ l'ensemble des fonctions bornées mesurables définies sur \mathcal{X} et à valeurs réelles. Pour toutes fonctions $w \in \mathbf{B}(\mathcal{X})$,

on note C_w sa borne supérieure :

$$C_w = \sup_{x \in \mathcal{X}} |w(x)|.$$

Soit $w \in \mathbf{B}(\mathcal{X})$ une fonction continue et Lipschitzienne. On note L_w sa constante de Lipschitz :

$$L_w = \sup_{x \neq x' \in \mathcal{X}} \frac{|w(x) - w(x')|}{d_{\mathcal{X}}(x, x')}.$$

Pour un noyau de Markov P sur $(\mathcal{X}, \mathfrak{B}(\mathcal{X}))$ et une fonction w dans $\mathbf{B}(\mathcal{X})$, on pose

$$\forall x \in \mathcal{X}, \quad Pw(x) = \int_{\mathcal{X}} w(y)P(x, dy).$$

1.3 Définition d'un processus de Markov déterministe par morceaux

Soit M l'ensemble fini des régimes possibles du système. Pour chaque régime m dans l'ensemble M , E_m est un sous-ensemble ouvert de \mathbb{R}^d muni de la norme usuelle Euclidienne $|\cdot|$. Soit $E = \{(m, \zeta), m \in M, \zeta \in E_m\}$. On définit pour $x = (m, \zeta)$ et $x' = (m', \zeta') \in E$, la distance sur E par :

$$|x - x'| = |\zeta - \zeta'| \mathbf{1}_{\{m=m'\}} + \infty \mathbf{1}_{\{m \neq m'\}}.$$

Un processus de Markov déterministe par morceaux (PDMP) sur l'espace d'état E est déterminé par trois caractéristiques locales :

- Le flot $\Phi(x, t) = (m, \Phi_m(\zeta, t))$ est défini pour tout $x = (m, \zeta) \in E$ et pour tout $t \geq 0$, où $\Phi_m : \mathbb{R}^d \times \mathbb{R}^+ \rightarrow \mathbb{R}^d$ est continue et telle que

$$\forall t, s \in \mathbb{R}^+, \quad \Phi_m(\cdot, t + s) = \Phi_m(\Phi_m(\cdot, t), s).$$

Le flot décrit la trajectoire déterministe du processus entre deux sauts. On appelle $t^*(x)$ le temps que le flot met pour atteindre la frontière de E quand il commence au point $x = (m, \zeta)$:

$$t^*(x) = \inf\{t > 0 : \Phi_m(\zeta, t) \in \partial E_m\}, \quad (1.1)$$

en rappelant la convention $\inf \emptyset = \infty$.

- L'intensité de saut $\lambda : \bar{E} \rightarrow \mathbb{R}^+$ est une fonction mesurable. Elle vérifie la propriété d'intégrabilité suivante :

$$\forall x = (m, \zeta) \in E, \quad \exists \varepsilon > 0, \quad \int_0^\varepsilon \lambda(m, \Phi_m(\zeta, t)) dt < +\infty.$$

Ainsi, le processus ne peut pas sauter immédiatement. Pour tous $x = (m, \zeta) \in E$ et $t \in [0; t^*(x)[$, on pose

$$\Lambda(m, \zeta, t) = \int_0^t \lambda(m, \Phi_m(\zeta, s)) ds. \quad (1.2)$$

- Le noyau de Markov Q sur $(\bar{E}, \mathfrak{B}(\bar{E}))$ représente la mesure de transition du processus et permet de sélectionner le nouveau point de départ du processus après chaque saut. Il satisfait

$$\forall x \in \bar{E}, \quad Q(x, \{x\} \cup \partial E) = 0.$$

Cela signifie que chaque saut se fait dans E et change la position et/ou le régime du processus. Il n'y a donc pas de saut sur place.

À partir de ces caractéristiques, il a été montré dans (Davis, 1993, section 25) qu'il existe un espace filtré de probabilité $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \{\mathbb{P}_x\}_{x \in E})$ pour lequel le processus de Markov déterministe par morceaux $\{X_t\}$ est un processus de Markov fort. Le processus $\{X_t\}$ a deux composantes : $X_t = (m_t, \zeta_t)$. La première composante m_t est appelée le régime ou le mode. La seconde composante ζ_t est appelée la variable euclidienne.

L'évolution du PDMP $\{X_t\}$ peut être définie de manière itérative comme suit. Considérons le processus démarrant au point $X_0 = (m_0, \zeta_0)$ avec $m_0 \in M$ et $\zeta_0 \in E_{m_0}$. Le premier temps de saut T_1 est déterminé par la fonction de survie :

$$\mathbb{P}_{(m_0, \zeta_0)}(\{T_1 > t\}) = e^{-\Lambda(m_0, \zeta_0, t)} \mathbb{1}_{\{t < t^*(m_0, \zeta_0)\}}. \quad (1.3)$$

Sur l'intervalle $[0; T_1[$, le processus $\{X_t\}$ suit la trajectoire déterministe $\zeta_t = \Phi_{m_0}(\zeta_0, t)$ et le régime m_t est constant, égal à m_0 . Au temps aléatoire T_1 , un saut survient. Ce dernier peut produire une discontinuité dans la variable euclidienne ζ_t et/ou changer le régime. Le processus redémarre à partir d'une nouvelle position $X_{T_1} = (m_{T_1}, \zeta_{T_1})$, selon la distribution $Q((m_0, \Phi_{m_0}(\zeta_0, T_1)), \cdot)$. Un temps inter-saut $T_2 - T_1$ est alors sélectionné de la même manière que précédemment, à l'aide de la fonction de survie (1.3). Sur l'intervalle $[T_1; T_2[$, le processus suit la trajectoire $m_t = m_{T_1}$ et $\zeta_t = \Phi_{m_{T_1}}(\zeta_{T_1}, t - T_1)$.

La figure 1.1 montre un exemple de dynamique d'un PDMP sur l'intervalle $[0; T_2[$: la position initiale du processus est $(m, \zeta) \in E$, un point de l'espace E_m . Il suit une trajectoire déterministe selon le flot Φ jusqu'au premier temps de saut T_1 déterminé de manière aléatoire. À cet instant, le processus redémarre à un nouveau point de l'espace d'état E , selon le noyau de Markov Q du processus. Dans l'exemple choisi, la nouvelle position du processus après le premier saut est dans un régime différent de la position initiale du PDMP. Après ce premier saut, le processus suit à nouveau le flot jusqu'à l'instant T_2 qui est un temps déterministe, correspondant à l'instant $T_1 + t^*(X_{T_1})$ où le processus atteint la frontière de l'espace d'état $E_{m'}$. De nouveau, une nouvelle position du processus est sélectionnée grâce au noyau de Markov Q . Dans notre exemple, la position X_{T_2} a le même régime que la position X_{T_1} .

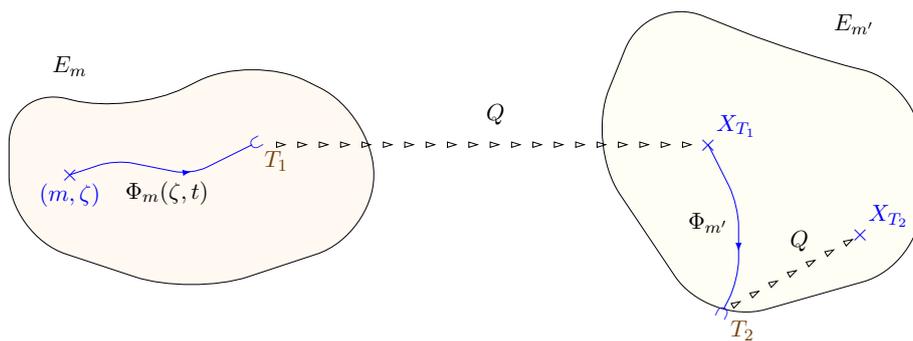


FIGURE 1.1 – Exemple de dynamique d'un processus de Markov déterministe par morceaux.

Pour des convenances de notations, on pose $T_0 = 0$. On définit la suite $(Z_n)_{n \in \mathbb{N}}$, par $Z_n = X_{T_n}$, pour $n \in \mathbb{N}$. Cette suite correspond aux positions du processus $\{X_t\}$ après les sauts. La suite $(S_n)_{n \in \mathbb{N}}$, définie par $S_n = T_n - T_{n-1}$ pour $n \geq 1$ et par $S_0 = 0$, donne les temps de séjours entre deux sauts consécutifs. On peut remarquer que le processus défini par $\{\Theta_n\}$, où $\Theta_n = (Z_n, S_n)$ pour tout $n \in \mathbb{N}$, est une chaîne de Markov à temps discret. Il s'agit de l'unique source aléatoire du processus $\{X_t\}$.

1.4 Hypothèses de régularité

On reprend les hypothèses classiques sur les PDMP, initialement faites dans (Davis, 1993, section 24). Ce travail étant réalisé dans la continuité de (de Saporta *et al.*, 2015, section 1.4), on reprend les hypothèses supplémentaires faites dans ce cas.

Hypothèse 1

Pour tout $x \in E$ et pour tout $t \in \mathbb{R}^+$,

$$\mathbb{E}_x \left[\sum_{n=1}^{\infty} \mathbb{1}_{\{T_n \leq t\}} \right] < \infty.$$

La moyenne du nombre de sauts survenu avant l'instant $t \in \mathbb{R}^+$ est finie, quelle que soit la position initiale du processus.

Cette première hypothèse implique en particulier que l'événement $\{T_k \rightarrow \infty\}$ est presque sûr, quand $k \rightarrow \infty$.

Hypothèse 2

Le temps de sortie t^* est une fonction continue, Lipschitzienne et bornée par C_{t^*} .

Dans la plupart des applications, les propriétés physiques du système assurent que le temps de sortie t^* est borné ou que le problème a un horizon de temps naturel et déterministe T . Dans ce dernier cas, il n'y a pas de perte de généralité à considérer que le temps de sortie t^* est borné par l'horizon de temps déterministe T . Cela permet de remplacer C_{t^*} par T .

Les hypothèses suivantes concernent le caractère Lipschitzien du taux de saut λ et du noyau de Markov Q .

Hypothèse 3

Le taux de saut λ est dans l'ensemble $\mathbf{B}(E)$ des fonctions bornées mesurables, définies sur E et à valeurs réelles. De plus, il existe une constante $[\lambda]_1 \in \mathbb{R}^+$ telle que pour tous $(x, y) \in E^2$ et $u \in [0, t^*(x) \wedge t^*(y)[$,

$$|\lambda(\Phi(x, u)) - \lambda(\Phi(y, u))| \leq [\lambda]_1 |x - y|.$$

On définit $\mathbf{L}_\Phi(E)$ comme l'ensemble des fonctions $w \in \mathbf{B}(E)$ qui sont continues et Lipschitziennes le long du flot, i.e. les fonctions à valeurs réelles, bornées, mesurables, définies sur E et satisfaisant les conditions suivantes :

1. Pour tout $x \in E$, l'application $w(\Phi(x, \cdot)) : [0; t^*(x)[\rightarrow \mathbb{R}$ est continue. De plus, la limite $\lim_{t \rightarrow t^*(x)} w(\Phi(x, t))$ existe et est notée $w(\Phi(x, t^*(x)))$.

2. Il existe $[w]_1 \in \mathbb{R}^+$ tel que pour tous $(x, y) \in E^2$, $t \in [0; t^*(x) \wedge t^*(y)[$, on a

$$|w(\Phi(x, t)) - w(\Phi(y, t))| \leq [w]_1 |x - y|.$$

3. Il existe $[w]_2 \in \mathbb{R}^+$ tel que pour tous $x \in E$, $(t, s) \in [0; t^*(x)]^2$, on a

$$|w(\Phi(x, t)) - w(\Phi(x, s))| \leq [w]_2 |t - s|.$$

4. Il existe $[w]_* \in \mathbb{R}^+$ tel que pour tous $(x, y) \in E^2$, on a

$$|w(\Phi(x, t^*(x))) - w(\Phi(y, t^*(y)))| \leq [w]_* |x - y|.$$

Hypothèse 4

Le noyau de Markov Q est Lipschitzien dans le sens suivant : il existe une constante $[Q] \in \mathbb{R}^+$ telle que pour toute fonction $w \in \mathbf{L}_\Phi(E)$ le noyau de Markov Q satisfait les deux conditions suivantes :

1. pour tous $(x, y) \in E^2$ et $t \in [0, t^*(x) \wedge t^*(y)[$, on a

$$|Qw(\Phi(x, t)) - Qw(\Phi(y, t))| \leq [Q][w]_1|x - y|,$$

2. pour tous $(x, y) \in E^2$, on a

$$|Qw(\Phi(x, t^*(x))) - Qw(\Phi(y, t^*(y)))| \leq [Q][w]_*|x - y|.$$

1.5 Exemples

Dans ce paragraphe, on présente trois exemples de processus de Markov déterministes par morceaux. Le premier est un modèle de transmission de données sur les réseaux. Le deuxième est un modèle où on s'intéresse aux fonds d'une assurance. Enfin, le dernier est un exemple en biologie qui décrit l'évolution de la quantité d'un contaminant alimentaire dans le corps.

Transmission de données : modèle du TCP

Le modèle du protocole de contrôle de transmissions (TCP pour *Transmission Control Protocol*) est un exemple de PDMP où le flot est linéaire. Ses propriétés ont notamment été étudiées dans Chafaï *et al.* (2010); Bardet *et al.* (2013).

Il s'agit d'un protocole de communication réseau qui permet d'envoyer des données via Internet. Il découpe les grandes données en de petits paquets et assure l'intégrité des données une fois les paquets rassemblés pour le destinataire. Ce découpage se fait en fonction des conditions de trafic des données du réseau.

Pour une connexion, le nombre maximal de paquets par envoi est donné par une variable aléatoire W , appelée la taille de la fenêtre de congestion. Si les W paquets ont été transmis avec succès, alors W augmente linéairement, sinon, on dit qu'il y a congestion et le nombre maximal de paquets par envoi est divisé par 2. Les moments de congestion surviennent à des temps aléatoires, avec un taux de saut donné par la valeur du taux de transmission. Entre ces moments, la trajectoire de W est entièrement déterministe.

On peut ainsi définir le processus $\{W_t\}$ comme un PDMP sur l'espace d'état $E = [0; +\infty[$ ayant pour flot $\Phi(x, t) = x + t$ et pour taux de saut λ . Son noyau de Markov satisfait $Q(x, \frac{x}{2}) = 1$, pour tout $x \in E$. La figure 1.2 donne un exemple de trajectoire de ce processus, où on peut observer qu'à chaque instant de congestion, le nombre de paquets par envoi est divisé par 2.

Modèle d'assurance

De nombreux modèles de théorie du risque font appel aux PDMP (Dassios et Embrechts (1989); Schäl (1998)). Le modèle d'assurance développé dans Dassios et Embrechts (1989) s'intéresse aux réserves d'une assurance, notées X_t . Les primes payées par les assurés rapportent c par unité de temps. Les temps d'inter-arrivée des sinistres suivent une loi exponentielle et les montants des sinistres sont indépendants et identiquement distribués suivant une loi donnée à valeurs dans \mathbb{R}^+ . Lorsqu'un sinistre survient, la compagnie paye immédiatement la somme due à l'assuré. Si la compagnie ne dispose pas des fonds nécessaires, elle peut emprunter à un taux β . Elle est nécessairement en faillite lorsque les intérêts de l'emprunt dépassent les primes des

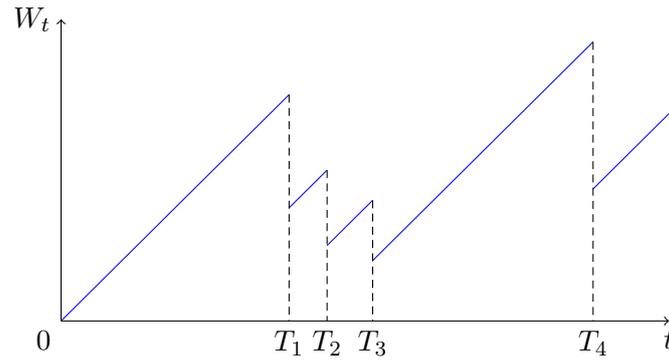


FIGURE 1.2 – Exemple de trajectoire d'un PDMP représentant le nombre de paquets par envoi d'un protocole TCP, en fonction du temps.

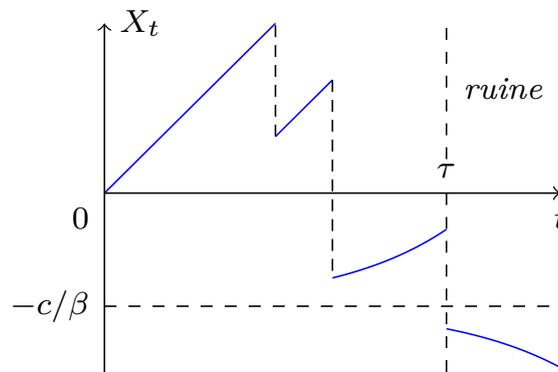


FIGURE 1.3 – Exemple de trajectoire d'un PDMP représentant les réserves d'une assurance jusqu'à sa ruine.

assurés, c'est-à-dire lorsque $X_t < \frac{-c}{\beta}$. Ainsi, l'évolution des fonds de la compagnie suit, entre deux sinistres, l'équation différentielle suivante

$$\frac{d}{dt}X_t = c\mathbb{1}_{\{X_t \geq 0\}} + (c + \beta X_t)\mathbb{1}_{\{X_t < 0\}}.$$

Le processus $\{X_t\}$ peut être modélisé par un PDMP ayant pour espace d'état $E = \mathbb{R}$. Son flot Φ satisfait l'équation différentielle :

$$\frac{d}{dt}\Phi(x, t) = c\mathbb{1}_{\{\Phi(x, t) \geq 0\}} + (c + \beta\Phi(x, t))\mathbb{1}_{\{\Phi(x, t) < 0\}}.$$

Son intensité de saut λ est fixée et son noyau de Markov dépend de la loi des montants des sinistres.

La figure 1.3 donne un exemple de trajectoire du PDMP $\{X_t\}$. Au temps τ , l'assurance ne peut plus couvrir les intérêts de l'emprunt qu'elle a fait. Elle est donc déclarée en ruine.

Contaminant alimentaire

Ce modèle a, entre autre, été étudié dans Bertail *et al.* (2008) et Bouguet (2015). Il étend légèrement les modèles de stockage tels que le stockage de l'eau dans les barrages Davis (1984), étant donné qu'il permet d'avoir des taux de libération/d'élimination aléatoires qui ne suivent pas nécessairement des lois exponentielles.

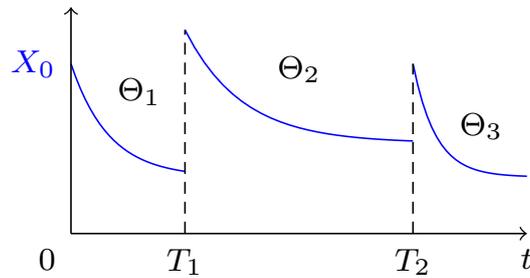


FIGURE 1.4 – Exemple de trajectoire d'un PDMP représentant la quantité d'un contaminant alimentaire dans le corps d'une personne.

Certains aliments peuvent contenir des éléments chimiques, tels que le méthylmercure, présent dans les poissons, et les dioxines, présentes dans la viande. Si ces éléments chimiques s'accumulent dans le corps et dépassent un certain seuil, ils deviennent toxiques pour l'organisme. On peut étudier l'évolution d'un contaminant alimentaire à l'aide des processus de Markov déterministes par morceaux. On note X_t la quantité du contaminant présent dans le corps à l'instant t . Cette quantité augmente à chaque ingestion d'un aliment contenant le contaminant étudié et décroît grâce à l'élimination naturelle du contaminant par le corps.

Les quantités de contaminant prises à chaque ingestion sont indépendantes et identiquement distribuées selon une loi donnée à valeurs dans \mathbb{R}^+ . Le phénomène d'élimination est déterminé par l'équation différentielle suivante :

$$\partial_t X_t = -\Theta X_t$$

où Θ est la vitesse d'élimination du contaminant par le corps. Cette vitesse dépend de facteurs métaboliques et constitue le régime du processus considéré. Elle est supposée constante entre deux ingestions, rendant ainsi les trajectoires du processus $\{X_t\}$ déterministes entre deux sauts.

Le processus $\{X_t\}$ peut être vu comme un processus de Markov déterministe par morceaux. Son espace d'état E correspond à l'ensemble $\mathcal{M} \times \mathbb{R}^+$, où $\mathcal{M} = \{1, \dots, n\}$, avec $n \in \mathbb{N}^*$ est fixé. Son flot est défini sur $E \times \mathbb{R}^+$ par $\Phi_m(\zeta, t) = \zeta e^{-\Theta_m t}$. Son intensité de saut est fixé par le modèle et son noyau de Markov dépend de la loi associée aux quantités de contaminant prises lors des ingestions. La figure 1.4 donne un exemple de trajectoire du processus $\{X_t\}$.

1.6 Problème de contrôle impulsif

Considérons un stock soumis à une demande aléatoire. Le stockage a un coût continu et les commandes pour le réapprovisionnement représentent des coûts fixes. On souhaite optimiser les coûts en déterminant la meilleure stratégie de réapprovisionnement. Une telle stratégie est constituée d'une suite de dates de réapprovisionnement et de quantités commandées à ces instants. En notant $\mathcal{J}(\mathcal{S})$ le coût associé à la stratégie \mathcal{S} , le problème consiste à trouver une stratégie \mathcal{S}^* (si elle existe) qui réalise l'infimum de la fonction \mathcal{J} parmi toutes les stratégies possibles. On appellera fonction valeur du problème cet infimum.

Il s'agit du premier exemple de contrôle impulsif historiquement étudié par A. Bensoussan et J.L. Lions dans Bensoussan et Lions (1975). Les problèmes de contrôle impulsif sont nombreux et constituent une part importante des problèmes appliqués d'optimisation. L'étude initiale du problème de contrôle impulsif faite par A. Bensoussan et J.L. Lions prend place dans le cadre des diffusions. Beaucoup d'auteurs ont ensuite contribué à développer des outils mathématiques pour le contrôle impulsif. Le premier auteur à se dégager du cadre des diffusions est M. Robin, qui dans sa thèse Robin (1978), expose un problème de contrôle impulsif

pour des processus de Markov fellériens, avec un retard déterministe, c'est-à-dire qu'on ne peut prendre une décision tant que l'effet de la dernière décision n'a pas eu lieu.

La théorie du contrôle impulsionnel pour les processus de Markov déterministes par morceaux a été écrite par O.L.V. Costa et M.H.A. Davis dans Costa et Davis (1989), puis a été reprise dans le livre de M.H. Davis Davis (1993). La construction de cette théorie est directe et indépendante des théories déjà existantes, même si les PDMP sont des cas particuliers des processus de Markov droits, utilisés par J.P. Lepeltier et B. Marchal dans Lepeltier et Marchal (1984) pour élaborer leur théorie générale du contrôle impulsionnel Markovien. En effet, repartir d'une théorie déjà existante avait pour conséquence de rajouter beaucoup d'hypothèses restrictives, dont on pouvait se passer pour les PDMP qui sont des processus présentant des spécificités importantes.

La forme la plus simple de contrôle impulsionnel est l'arrêt optimal, pour lequel un seul temps d'arrêt est sélectionné. On s'intéresse ensuite au gain obtenu, qui dépend de la position du processus au moment de son arrêt. Ce problème, ainsi que les principaux résultats le concernant, est présenté au paragraphe 1.6.1, uniquement dans le cadre des PDMP. Il constitue la première étape dans la résolution du problème de contrôle impulsionnel pour les PDMP. Ce dernier est détaillé au paragraphe 1.6.2. On y présente également les hypothèses que l'on choisira pour la suite de notre travail.

1.6.1 Arrêt optimal pour des PDMP

Les exemples de problème d'arrêt optimal sont nombreux et variés : en théorie des jeux Hill (2009), en médecine Berman (1994) pour l'arrêt de traitement, en fiabilité des systèmes de Saporta *et al.* (2015) pour l'arrêt de systèmes avant leur défaillance complète, ou encore en finance Pham (1997) pour le temps d'exercice d'une option.

L'un des plus connus est celui de la princesse Hill (2009) : une princesse doit choisir un mari parmi cent prétendants. Elle a un entretien avec chacun d'entre eux. Ces entretiens se font dans un ordre aléatoire et la princesse n'a aucune information sur les candidats qu'elle n'a pas rencontrés. À l'issue de chaque entretien, la princesse peut choisir d'épouser ou non le prétendant qui vient de se présenter. Si elle décide de ne pas le faire, elle ne peut plus revenir sur son choix et perd à jamais la possibilité de l'épouser. Si la princesse ne choisit aucun prétendant jusqu'au 99^e candidat, alors elle devra épouser le dernier prétendant. L'objectif est d'épouser le meilleur prétendant possible ; la princesse doit donc utiliser un temps d'arrêt optimal dans ce but. Il a été montré que la meilleure stratégie permet de choisir le meilleur prétendant dans plus d'un cas sur trois. Elle consiste à observer les 37 premiers prétendants, afin de se constituer une base de référence, puis sélectionner le premier prétendant qui sera meilleur que tous ceux observés auparavant.

Les problèmes d'arrêt optimaux dans le cadre des PDMP ont fait l'objet de nombreux articles et de chapitres de livres. On peut citer Lenhart et Liao (1985); Gugerli (1986); Costa et Davis (1988); Davis (1993) et plus récemment Costa *et al.* (2000); de Saporta *et al.* (2010, 2015). Cependant, seuls les articles sur lesquels se basent les idées du contrôle impulsionnel étudié dans cette thèse, seront résumés ici, à savoir les articles Gugerli (1986) et de Saporta *et al.* (2010).

Définition du problème d'arrêt optimal

On se place dans le cadre d'un problème d'arrêt optimal avec horizon aléatoire $n \in \mathbb{N}$ fixé pour la suite, que l'on définit comme étant un triplet (\mathcal{M}_n, g, v) où

- $\mathcal{M}_n = \{\tau \in \mathcal{M} \mid \tau \leq T_n \text{ p.s.}\}$, avec \mathcal{M} l'ensemble des \mathcal{F}_t -temps d'arrêts, en notant (\mathcal{F}_t) la filtration naturelle du PDMP,
- la fonction de gain g est mesurable, continue et bornée,

- la fonction valeur v est définie par

$$\forall x \in E, \quad v(x) = \sup_{\tau \in \mathcal{M}_n} \mathbb{E}_x[g(X_\tau)].$$

L'ensemble \mathcal{M}_n est l'ensemble des \mathcal{F}_t -temps d'arrêts sont dominés par le n -ième temps de saut T_n du PDMP. La fonction g est une fonction de gain où $g(x)$ représente le gain obtenu lorsque l'on arrête le processus à la position x . La résolution de ce problème consiste à déterminer un temps d'arrêt optimal $\tau^*(x) \in \mathcal{M}_n$ (s'il existe) qui atteint la borne supérieure du coût, c'est-à-dire qui vérifie :

$$\mathbb{E}_x[g(X_{\tau^*(x)})] = v(x).$$

Cependant, de manière générale, il n'existe pas de temps d'arrêt optimal. On cherche alors un temps d'arrêt appelé ε -optimal, que l'on note $\tau_\varepsilon^*(x)$, qui satisfait la condition

$$v(x) - \varepsilon \leq \mathbb{E}_x[g(X_{\tau_\varepsilon^*(x)})].$$

Construction itérative de la fonction valeur

Dans Gugerli (1986), U.S. Gugerli établit une construction par récurrence de la fonction valeur d'un tel problème d'arrêt optimal. Pour cela, il introduit l'opérateur L_{AO} défini pour toute fonction $w \in \mathbf{B}(E)$ et toute fonction h mesurable bornée, par :

$$L_{AO}(w, h)(x) = \sup_{t \geq 0} J_{AO}(w, h)(x, t) \vee K_{AO}w(x),$$

où les opérateurs K_{AO} et J_{AO} sont définis comme suit

$$\begin{aligned} J_{AO}(w, h)(x) &= \mathbb{E}_x[h(\Phi(x, t \wedge t^*(x))) \mathbf{1}_{\{S_1 \geq t \wedge t^*(x)\}} + w(Z_1) \mathbf{1}_{\{S_1 < t \wedge t^*(x)\}}], \\ K_{AO}w(x) &= \mathbb{E}_x[w(Z_1)], \end{aligned}$$

où S_1 est le temps du premier saut du processus et Z_1 est la position après ce saut, déterminée par le noyau de Markov Q associé au processus.

U.S. Gugerli définit ensuite les fonctions $(v_k)_{k \in \{0, \dots, n\}}$ par :

$$\begin{cases} v_n &= g \\ v_k &= L_{AO}(v_{k+1}, g), \quad \text{pour } k \leq n-1. \end{cases}$$

Il montre alors que la fonction valeur v du problème d'arrêt optimal à horizon aléatoire n est égale à la dernière fonction ainsi construite v_0 .

Construction d'un temps d'arrêt ε -optimal

Toujours dans Gugerli (1986), une suite de temps d'arrêt ε -optimaux est définie à l'aide d'une fonction du temps éventuel d'intervention (fonction r définie p.227), dont le rôle est de déterminer, à un moment donné de la vie du processus, s'il est préférable de s'arrêter ou non. Les détails de la construction de cette fonction et du temps d'arrêt ε -optimal associé sont donnés dans la partie 2 de Gugerli (1986). On en donnera juste l'idée dans ce paragraphe. Sans entrer dans les détails, la fonction du temps éventuel d'intervention est déterminée à partir de la comparaison des valeurs de la borne supérieure de l'opérateur J_{AO} avec la valeur de l'opérateur K_{AO} .

Si K_{AO} gagne la maximisation, alors l'idée est que le gain rapporté par le processus si on ne l'arrête pas avant son prochain saut, est supérieur au gain que l'on aurait si on arrêta le processus avant son prochain saut (quelque soit le moment d'arrêt). On choisit donc pour la fonction du temps éventuel d'intervention le temps d'atteinte à la frontière t^* .

En revanche, si la borne supérieure de J_{AO} gagne la maximisation, cela signifie que le gain rapporté par le processus si l'on s'arrête est supérieur au gain que l'on aurait si on laissait le processus évoluer librement jusqu'au saut suivant. Cependant, la borne supérieure de J_{AO} n'étant pas toujours atteinte, on choisit pour la valeur de la fonction du temps éventuel d'intervention, un temps qui est tel que l'opérateur J_{AO} pris en ce temps est très proche de sa borne supérieure.

L'idée de la construction du temps d'arrêt ε -optimal de U.S. Gugerli est de regarder la position du prochain saut du processus par rapport à la valeur de la fonction du temps éventuel d'intervention pris en compte. Si le saut a lieu avant le temps éventuel d'intervention, alors le processus n'est pas arrêté avant le prochain saut. Au contraire, si le processus n'a pas sauté avant le temps éventuel d'intervention, alors le processus est arrêté à ce dernier temps. Cette construction implique que, dans le cas où l'opérateur K_{AO} gagne la maximisation, on laisse le processus évoluer librement jusqu'à son prochain saut.

Méthode numérique

Une construction numérique des temps d'arrêts ε -optimaux définis par U.S. Gugerli, est faite dans de Saporta *et al.* (2010). Elle repose sur une discrétisation du temps et une discrétisation de l'espace. La discrétisation en temps permet de passer du calcul d'une borne supérieure de l'opérateur J_{AO} , à une recherche de maximum pour des temps compris dans une grille adaptée dans le sens où cette dernière prend en compte la position du processus considéré dans l'opérateur J_{AO} (voir p.1615 de de Saporta *et al.* (2010)).

La discrétisation en espace se fait grâce à la quantification du PDMP. Ce procédé est expliqué en détails au paragraphe 1.7 mais l'idée de la quantification est de construire, pour une variable X donnée, une grille finie telle que la projection \hat{X} de la variable X sur cette grille, minimise au sens de la norme L^p la quantité $X - \hat{X}$. Une des spécificités des PDMP est que toute la partie aléatoire de ces processus est décrite par la chaîne discrète de Markov sous-jacente des temps inter-sauts et des positions après-sauts. Ainsi, lorsque l'on regarde un problème d'arrêt optimal à horizon aléatoire n , il nous faut construire $(n + 1)$ grilles de quantification pour décrire le PDMP considéré. La quantification permet de définir de nouveaux opérateurs \hat{J}_{AO} et \hat{K}_{AO} non plus comme des espérances, mais comme des sommes finies.

Moyennant quelques subtilités, une nouvelle fonction de temps éventuel d'arrêt est alors définie (voir p.1624 de de Saporta *et al.* (2010)) à partir de la fonction du temps éventuel d'intervention de U.S. Gugerli. Il suit une construction d'un temps d'arrêt, pour lequel il est prouvé qu'il s'agit bien d'un temps d'arrêt ε -optimal.

1.6.2 Contrôle impulsif pour des PDMP

L'arrêt optimal, étudié au paragraphe 1.6.1, est la forme la plus simple de contrôle impulsif. En effet, il s'agit d'arrêter le processus à un instant donné pour maximiser une fonction gain. Cependant, si l'on souhaite intervenir sur le processus à un instant donné et lui permettre de repartir après cette intervention, le problème se transforme en un problème de contrôle impulsif.

De manière plus générale, le principe du contrôle impulsif est le suivant. Le processus évolue librement jusqu'à ce qu'un contrôleur décide d'intervenir dessus. Le contrôleur déplace alors instantanément le processus vers un nouveau point de l'espace d'état. Le processus repart alors de ce nouveau point et évolue librement jusqu'à la prochaine intervention, et ainsi de suite. Le but du problème est de trouver une stratégie d'intervention, c'est-à-dire une suite de temps d'arrêt, correspondant aux instants de contrôle, et de variables aléatoires, correspondant aux nouveaux points de départ du processus après intervention, qui minimise une fonction coût.

L'une des difficultés majeures du contrôle impulsif est la sélection du nouveau point de départ du processus après les interventions. Cette complexité supplémentaire par rapport à l'arrêt optimal requiert l'ajout d'un nouvel opérateur dans l'étude du problème.

Les problèmes de contrôle impulsif pour les PDMP ont été étudiés par un certain nombre d'auteurs. On peut citer Lenhart (1989); Gątarek (1992); Dempster et Ye (1995), dans lesquels le problème de contrôle impulsif est considéré au moyen d'inéquations variationnelles. Le travail présenté ici se base sur la théorie et le formalisme développés par O.L.V. Costa et M.H.A. Davis dans Costa et Davis (1989).

Définition du problème de contrôle impulsif

On considère un problème de contrôle impulsif à horizon infini avec facteur d'oubli. Le formalisme probabiliste nécessaire pour définir précisément ce problème est conséquent et ne sera pas utilisé dans la suite de cette thèse. C'est la raison pour laquelle nous ne présentons ici qu'une brève description du problème qui nous intéresse. On pourra se référer à l'article Costa et Davis (1989) ou au livre (Davis, 1993, section 54) pour une définition complète du problème.

Une stratégie d'intervention, notée $\mathcal{S} = (\tau_n, R_n)_{n \geq 1}$, est une suite de temps d'intervention non-anticipatifs $(\tau_n)_{n \geq 1}$ et de variables aléatoires $(R_n)_{n \geq 1}$ non-anticipatives à valeurs dans E sur un espace mesurable $(\bar{\Omega}, \bar{\mathcal{F}})$. À partir de maintenant, on distingue le PDMP initial, noté $\{X_t\}$, appelé encore processus non contrôlé, du processus contrôlé par une stratégie, noté $\{\tilde{X}_t\}$. L'évolution du processus contrôlé est liée à celle du processus initial, de la manière suivante. Entre deux temps d'intervention τ_i et τ_{i+1} , la trajectoire du processus contrôlé $\{\tilde{X}_t\}$ est déterminée par les caractéristiques du PDMP initial $\{X_t\}$, partant du point R_i . On renvoie à l'article (Costa et Davis, 1989, section 2.2) pour la construction détaillée du processus contrôlé $\{\tilde{X}_t\}$.

Lorsqu'une intervention a lieu sur le processus contrôlé $\{\tilde{X}_t\}$, le contrôleur choisit un nouveau point de départ pour le processus, dans un ensemble noté \mathbb{U} , appelé ensemble de contrôle, inclus dans l'espace d'état E .

On dit qu'une stratégie est admissible si elle satisfait les cinq conditions énoncées dans (Costa et Davis, 1989, section 2.3). L'ensemble des stratégies admissibles est noté \mathcal{S} . Soit \mathcal{S} une stratégie d'intervention admissible. Cette dernière induit une famille de mesures de probabilité $\mathbb{P}_x^{\mathcal{S}}$, $x \in E$, sur $(\bar{\Omega}, \bar{\mathcal{F}})$. Le coût avec facteur d'oubli associé à la stratégie d'intervention \mathcal{S} pour un processus partant du point x de l'espace d'état E , est défini par

$$\mathcal{J}^{\mathcal{S}}(x) = \mathbb{E}_x^{\mathcal{S}} \left[\int_0^{\infty} e^{-\alpha s} f(\tilde{X}_s) ds + \sum_{i=1}^{\infty} e^{-\alpha \tau_i} c(\tilde{X}_{\tau_i-}, \tilde{X}_{\tau_i}) \right], \quad (1.4)$$

où

- $\mathbb{E}_x^{\mathcal{S}}$ est l'espérance contre la mesure de probabilité $\mathbb{P}_x^{\mathcal{S}}$,
- α est le facteur positif d'oubli,
- f est la fonction du coût courant,
- c est la fonction liée aux coûts des interventions : $c(x, y)$ correspond au coût de l'intervention qui déplace le processus du point x au point y .

On impose les deux hypothèses suivantes sur les fonctions f et c .

Hypothèse 5

La fonction f du coût courant est une fonction positive appartenant à $\mathbf{L}_{\Phi}(E)$.

Hypothèse 6

La fonction c associée aux coûts des interventions appartient à $\mathbf{B}(\bar{E} \times \mathbb{U})$ et vérifie les assertions suivantes :

1. Il existe une constante $[c]_1 \in \mathbb{R}^+$ telle que pour tous $(x, y) \in E^2$ et $u \in [0, t^*(x) \wedge t^*(y)]$,

on a

$$\max_{z \in \mathbb{U}} |c(\Phi(x, u), z) - c(\Phi(y, u), z)| \leq [c]_1 |x - y|.$$

2. Il existe une constante $[c]_2 \in \mathbb{R}^+$ telle que pour tous $x \in E$ et $(t, s) \in [0, t^*(x)]^2$, on a

$$\max_{z \in \mathbb{U}} |c(\Phi(x, t), z) - c(\Phi(x, s), z)| \leq [c]_2 |t - s|.$$

3. Il existe une constante $[c]_* \in \mathbb{R}^+$ telle que pour tous $(x, y) \in E^2$, on a

$$\max_{z \in \mathbb{U}} |c(\Phi(x, t^*(x)), z) - c(\Phi(y, t^*(y)), z)| \leq [c]_* |x - y|.$$

4. Il existe deux constantes $c_0, C_c \in \mathbb{R}_*^+$ telles que pour tous $(x, y) \in \bar{E} \times \mathbb{U}$,

$$0 < c_0 \leq c(x, y) \leq C_c.$$

5. Pour tous $(x, y, z) \in \bar{E} \times \mathbb{U} \times \mathbb{U}$, on a

$$c(x, y) + c(y, z) \geq c(x, z).$$

L'assertion 5 de l'hypothèse 6 signifie que le coût d'une seule intervention sera moins important (ou égal) au coût de deux interventions simultanées.

Si on considère la stratégie sans intervention, alors son coût avec facteur d'oubli, c'est-à-dire le coût associé au processus non contrôlé $\{X_t\}$, est donné par la fonction h définie par

$$\forall x \in E, \quad h(x) = \mathbb{E}_x \left[\int_0^\infty e^{-\alpha s} f(X_s) ds \right]. \quad (1.5)$$

La fonction valeur du problème de contrôle impulsif à horizon infini avec facteur d'oubli est définie pour tout x dans E par :

$$\mathcal{V}(x) = \inf_{S \in \mathcal{S}} \mathcal{J}^S(x).$$

En règle générale, il n'existe pas de stratégies optimales, dans le sens où elles réalisent l'infimum de la fonction coût $\mathcal{J}(x)$. On cherchera donc une stratégie générale ε -optimale $\mathcal{S}_\varepsilon^*$, c'est-à-dire vérifiant l'inégalité :

$$\mathcal{V}(x) \leq \mathcal{J}^{\mathcal{S}_\varepsilon^*}(x) \leq \mathcal{V}(x) + \varepsilon.$$

Construction itérative de la fonction valeur

La fonction valeur \mathcal{V} associée au problème de contrôle impulsif à horizon infini avec facteur d'oubli peut se construire itérativement, à l'aide d'un opérateur. Ce résultat provient de l'article Costa et Davis (1989). Il est résumé ici et les définitions prises pour les opérateurs sont adaptées comme dans l'article de Saporta et Dufour (2012).

Pour $x \in E$, $t \geq 0$, $(v, w) \in \mathbf{L}_\Phi(E)^2$, on pose

$$\begin{aligned} F(x, t) &= \int_0^{t \wedge t^*(x)} e^{-\alpha s - \Lambda(x, s)} f(\Phi(x, s)) ds \\ &= \mathbb{E}_x \left[\int_0^{T_1 \wedge t} e^{-\alpha s} f(\Phi(x, s)) ds \right], \end{aligned} \quad (1.6)$$

$$\begin{aligned}
J(v, w)(x, t) &= \int_0^{t \wedge t^*(x)} e^{-\alpha s - \Lambda(x, s)} \left[f(\Phi(x, s)) + \lambda Qw(\Phi(x, s)) \right] ds \\
&\quad + e^{-\alpha(t \wedge t^*(x)) - \Lambda(x, t \wedge t^*(x))} v(\Phi(x, t \wedge t^*(x))) \\
&= F(x, t) + \mathbb{E}_x \left[e^{-\alpha(t \wedge t^*(x))} v(\Phi(x, t \wedge t^*(x))) \mathbb{1}_{\{S_1 \geq t \wedge t^*(x)\}} \right. \\
&\quad \left. + e^{-\alpha S_1} w(Z_1) \mathbb{1}_{\{S_1 < t \wedge t^*(x)\}} \right], \tag{1.7}
\end{aligned}$$

$$\begin{aligned}
Kw(x) &= \int_0^{t^*(x)} e^{-\alpha s - \Lambda(x, s)} \left[f(\Phi(x, s)) + \lambda Qw(\Phi(x, s)) \right] ds + e^{-\alpha t^*(x) - \Lambda(x, t^*(x))} Qw(\Phi(x, t^*(x))) \\
&= F(x, t^*(x)) + \mathbb{E}_x \left[e^{-\alpha S_1} w(Z_1) \right]. \tag{1.8}
\end{aligned}$$

Pour toute fonction φ définie sur l'ensemble de contrôle \mathbb{U} et pour tout point $x \in E$, on pose

$$M\varphi(x) = \inf_{y \in \mathbb{U}} \{c(x, y) + \varphi(y)\}. \tag{1.9}$$

Par rapport à l'arrêt optimal, l'opérateur M constitue une nouveauté. Il va permettre de sélectionner la meilleure nouvelle position du processus après une intervention.

Pour toutes fonctions $(v, w) \in \mathbf{L}_\Phi(E)^2$ et pour tout point $x \in E$, on définit les deux opérateurs suivants

$$L(v, w)(x) = \inf_{t \in \mathbb{R}^+} J(v, w)(x, t) \wedge Kw(x) \tag{1.10}$$

et

$$\mathcal{L}w(x) = L(Mw, w)(x). \tag{1.11}$$

La définition des opérateurs J et K est différente de celle donnée dans Costa et Davis (1989). Cependant, l'opérateur L est le même. Il suffit d'utiliser la continuité de la fonction $J(v, w)(x, \cdot)$ sur \mathbb{R}^+ . En effet, en utilisant uniquement nos notations, par construction de J , on a pour tout $t \geq t^*(x)$, $J(v, w)(x, t) = J(v, w)(x, t^*(x))$. Donc on peut écrire :

$$\begin{aligned}
L(v, w)(x) &= \inf_{t \in \mathbb{R}^+} J(v, w)(x, t) \wedge Kw(x) \\
&= \inf_{t \in [0; t^*(x)]} J(v, w)(x, t) \wedge Kw(x).
\end{aligned}$$

Alors, en utilisant la continuité de la fonction $t \mapsto J(v, w)(x, t)$ sur \mathbb{R}^+ , on obtient

$$L(v, w)(x) = \inf_{t \in [0; t^*(x)[} J(v, w)(x, t) \wedge Kw(x).$$

De plus, notre définition de l'opérateur K ainsi que celle de l'opérateur J sur $E \times [0; t^*(z)[$ correspondent à celles de Costa et Davis (1989). En détails, pour l'opérateur J , l'égalité s'obtient par le fait que pour tout $t \in [0; t^*(x)[$, $\mathbb{P}_x(\{S_1 = t\}) = 0$. On a ainsi

$$\begin{aligned}
\forall t \in [0; t^*(x)[, \quad J(v, w)(x, t) &= \mathbb{E}_x \left[\int_0^{T_1 \wedge t} e^{-\alpha s} f(\Phi(x, s)) ds + e^{-\alpha t} v(\Phi(x, t)) \mathbb{1}_{\{S_1 > t\}} \right. \\
&\quad \left. + e^{-\alpha S_1} w(Z_1) \mathbb{1}_{\{S_1 \leq t\}} \right],
\end{aligned}$$

ce qui correspond à la définition de l'opérateur J dans Costa et Davis (1989). Pour l'opérateur K , en utilisant le fait que $T_1 \leq t^*(x)$, on a immédiatement la correspondance de notre définition avec celle de Costa et Davis (1989) :

$$Kw(x) = \mathbb{E}_x \left[\int_0^{T_1} e^{-\alpha s} f(\Phi(x, s)) ds + e^{-\alpha S_1} w(Z_1) \right].$$

On peut donc en conclure que notre opérateur L est le même que celui de l'article Costa et Davis (1989). Par conséquence, les résultats de O.L.V. Costa et M.H.A. Davis sont valables dans notre étude.

L'opérateur \mathcal{L} appliqué à la fonction w peut être interprété comme la fonction valeur d'un problème à un seul saut-ou-intervention, avec une fonction coût w . La fonction valeur \mathcal{V} du problème de contrôle impulsionnel étudié ici peut être calculée en itérant l'opérateur \mathcal{L} . Plus précisément, on rappelle ci-dessous la proposition 4 de Costa et Davis (1989).

Proposition 1.6.1

On pose $\mathcal{V}_0 = h$ le coût avec facteur d'oubli de la stratégie sans intervention et on définit pour tout $n \geq 0$, $\mathcal{V}_{n+1} = \mathcal{L}(\mathcal{V}_n)$. Alors pour tout point x de l'espace d'état E , on a

$$\mathcal{V}(x) = \lim_{n \rightarrow +\infty} \mathcal{V}_n(x).$$

En considérant qu'une intervention est un saut du processus contrôlé $\{\tilde{X}_t\}$, la fonction \mathcal{V}_n correspond à la fonction valeur d'un problème de contrôle impulsionnel avec horizon infini mais où le processus n'est plus contrôlé après le n -ième saut. On a donc au plus n interventions possibles et celles-ci ont forcément lieu avant ou à l'instant \tilde{T}_n du temps du n -ième saut du processus contrôlé.

Stabilité des opérateurs K , J et \mathcal{L}

Quelques propriétés sur les opérateurs K , J et \mathcal{L} sont données à titre de rappel, car utilisées notamment pour justifier l'existence d'objets introduits au cours de notre étude. La première propriété est issue de Costa et Davis (1989). On en donne une démonstration. On donne ensuite une proposition qui porte sur la stabilité des opérateurs K , J et \mathcal{L} , issue de de Saporta et Dufour (2012).

Proposition 1.6.2

Soit $x \in E$. Pour la fonction h définie en (1.5), on a

$$Kh(x) = h(x).$$

Démonstration : Par la définition (1.5) de h , on a

$$\begin{aligned} h(x) &= \mathbb{E}_x \left[\int_0^{+\infty} e^{-\alpha s} f(X_s) ds \right] \\ &= \mathbb{E}_x \left[\int_0^{S_1} e^{-\alpha s} f(X_s) ds + \int_{S_1}^{+\infty} e^{-\alpha s} f(X_s) ds \right] \\ &= \mathbb{E}_x \left[\int_0^{S_1 \wedge t^*(x)} e^{-\alpha s} f(X_s) ds + \mathbb{E}_x \left[\int_{S_1}^{+\infty} e^{-\alpha s} f(X_s) ds \middle| \mathcal{F}_{S_1} \right] \right] \\ &= F(x, t^*(x)) + \mathbb{E}_x \left[\mathbb{E}_x \left[e^{-\alpha S_1} \int_0^{+\infty} e^{-\alpha t} f(X_{t+S_1}) dt \middle| \mathcal{F}_{S_1} \right] \right] \\ &= F(x, t^*(x)) + \mathbb{E}_x \left[e^{-\alpha S_1} \mathbb{E}_x \left[\int_0^{+\infty} e^{-\alpha t} f(X_{t+S_1}) dt \middle| \mathcal{F}_{S_1} \right] \right]. \end{aligned}$$

La propriété de Markov appliquée au processus $\{X_t\}$ nous donne l'égalité suivante

$$\begin{aligned} h(x) &= F(x, t^*(x)) + \mathbb{E}_x \left[e^{-\alpha S_1} \mathbb{E}_{Z_1} \left[\int_0^{+\infty} e^{-\alpha t} f(X_t) dt \right] \right] \\ &= F(x, t^*(x)) + \mathbb{E}_x \left[e^{-\alpha S_1} h(Z_1) \right] \\ &= Kh(x), \end{aligned}$$

en utilisant la définition (1.8) de l'opérateur K . ■

Proposition 1.6.3

Pour toute fonction w de $\mathbf{L}_\Phi(E)$, on a

1. $\mathcal{L}w \in \mathbf{L}_\Phi(E)$,
2. $Kw \in \mathbf{L}_\Phi(E)$,
3. $J(Mw, w) \in \mathbf{L}_\Phi(E)$.

Démonstration : La première assertion provient du lemme A.11. de de Saporta et Dufour (2012). Les deux assertions suivantes sont issues des lemmes A.3, A.6 et A.9 de de Saporta et Dufour (2012). ■

Grâce à cette dernière proposition, on peut justifier que la construction des fonctions valeurs a du sens. En effet, l'ensemble $\mathbf{L}_\Phi(E)$ étant stable par l'opérateur \mathcal{L} , on obtient pour tout $n \in \mathbb{N}$, que la fonction valeur \mathcal{V}_n appartient à l'ensemble $\mathbf{L}_\Phi(E)$.

Stratégies ε -optimales

O.L.V. Costa et M.H.A. Davis proposent dans Costa et Davis (1989) une construction de stratégies ε -optimales pour le problème de contrôle impulsif avec facteur d'oubli et horizon infini, en utilisant pour cela un processus discrétisé : ce dernier ne peut sauter que dans un ensemble fini de points. À partir de la fonction valeur associée au processus discrétisé, O.L.V. Costa et M.H.A. Davis construisent une suite de temps d'arrêt et de variables aléatoires, dont ils prouvent qu'elle correspond à une stratégie ε -optimale pour le problème général de contrôle impulsif. Ils exposent ensuite des techniques numériques pour mettre en oeuvre les stratégies ε -optimales.

Cependant, la construction des stratégies de l'article Costa et Davis (1989) implique de connaître, en tout point de l'espace d'état, la fonction valeur associée au processus discrétisé. De plus, elle nécessite la résolution de plusieurs problèmes d'arrêt optimal. Par conséquent, sa mise en oeuvre peut rapidement devenir très compliquée.

C'est la raison pour laquelle nous développons une autre construction des stratégies ε -optimales au chapitre 5. La construction proposée nécessite la connaissance du coût de la stratégie de non-intervention, ainsi que la connaissance des opérateurs J et K et de leurs itérations. Notre construction est explicite et ne demande pas la résolution de problèmes complexes au préalable. Les idées à l'origine de la construction sont basées sur celles de la construction de temps d'arrêt ε -optimal proposée dans Gugerli (1986) dans le cadre du problème d'arrêt optimal.

Approximation de la fonction valeur

Dans de Saporta et Dufour (2012), les auteurs proposent une approximation numérique de la fonction valeur associée au problème de contrôle impulsif avec facteur d'oubli et horizon infini, avec l'hypothèse supplémentaire suivante :

Hypothèse 7

L'ensemble de contrôle \mathbb{U} est fini et ne dépend pas de la position du processus avant l'intervention.

Le cardinal de l'ensemble \mathbb{U} est noté u :

$$\mathbb{U} = \{y^i : 1 \leq i \leq u\}.$$

Tout comme pour l'arrêt optimal, l'approche de cette méthode numérique repose sur l'utilisation de la quantification pour la chaîne sous-jacente au PDMP, et l'utilisation d'une grille de temps adaptée pour la discrétisation en temps dans le supremum de l'opérateur J . Un taux de convergence est également explicité.

La quantification est présentée dans le paragraphe 1.7 et la méthode numérique de l'approximation de la fonction valeur est détaillée au paragraphe 1.8. L'approximation de la fonction valeur dans le cadre de l'optimisation de maintenance d'un équipement est présentée au chapitre 4.

Elle constitue une étape nécessaire dans la construction d'une approximation numérique de stratégies ε -optimales. La construction d'une telle approximation pour les stratégies ε -optimales développées au chapitre 5 pourra faire l'objet d'un futur travail.

1.7 Quantification

Le terme "quantification" est apparu dans les années cinquante où il était utilisé principalement en théorie du signal et de l'information. Dans ce contexte, la quantification représente la discrétisation d'un signal continu par un nombre fini de codes, appelés quantificateurs. Le but étant d'avoir une transmission efficace du signal, il est nécessaire d'optimiser les codes, afin de minimiser l'erreur commise.

En mathématiques, la quantification consiste à trouver la meilleure approximation d'une loi de probabilité d -dimensionnelle par une loi de probabilité discrète dont le support est un nombre N donné de points. Autrement dit, il s'agit de trouver la meilleure approximation d'un vecteur aléatoire X de dimension d par un vecteur aléatoire Y prenant au plus N valeurs. Le vecteur Y est appelé la version quantifiée du vecteur X .

Ce paragraphe vise à introduire la notion de quantification, ainsi que l'algorithme de quantification d'une chaîne de Markov utilisé dans le chapitre 4 pour la mise en place de l'approximation numérique de la fonction valeur dans le cadre de l'optimisation de la maintenance de la caméra Sophie. Cette introduction est basée principalement sur le cours Pham (2007), les articles Pagès et Printems (2003); Pagès *et al.* (2004) et le livre Graf et Luschgy (2000). Dans un premier temps, on définit la quantification d'une variable aléatoire, puis on s'intéresse à la quantification d'une chaîne de Markov.

1.7.1 Quantification d'une variable aléatoire

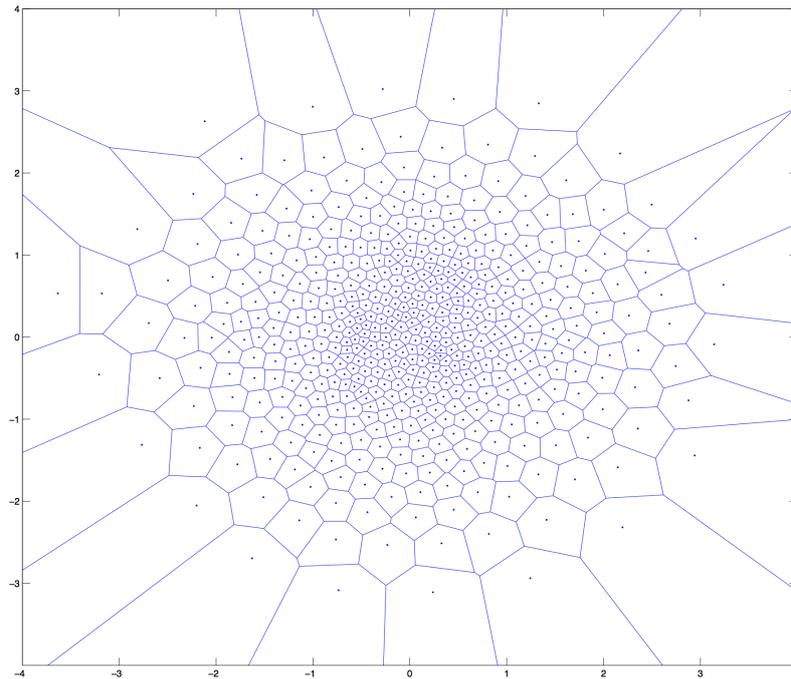
Soit X un vecteur aléatoire à valeurs dans \mathbb{R}^d sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ et tel que $\|X\|_p = \mathbb{E}[|X|^p]^{1/p} < +\infty$. On note \mathbb{P}_X sa loi de probabilité. La L^p -quantification optimale, pour $p \geq 1$, consiste à trouver, au sens de la norme L^p , la meilleure approximation de X par un vecteur aléatoire $\hat{X} := \pi(X)$ où $\pi : \mathbb{R}^d \rightarrow \mathbb{R}^d$ est une fonction borélienne, appelée *quantificateur*, prenant au plus N valeurs. On appelle *grille de quantification* $\Gamma := \pi(\mathbb{R}^d) := \{x_1, x_2, \dots, x_N\}$.

Définition 1.7.1

Soit $\Gamma = \{x_1, \dots, x_N\}$ une grille de N points distincts de \mathbb{R}^d . Une partition Borélienne $(C_i(\Gamma))_{1 \leq i \leq N}$ de \mathbb{R}^d est une mosaïque de Voronoi si et seulement si pour tout $i \in \{1, \dots, N\}$, $C_i(\Gamma)$ satisfait

$$C_i(\Gamma) \subset \{y \in \mathbb{R}^d, |y - x_i| = \min_{j=1, \dots, N} |y - x_j|\}.$$

Un exemple d'une mosaïque de Voronoi à 500 points dans \mathbb{R}^2 est donné par la figure 1.5.

FIGURE 1.5 – Mosaïque de Voronoi à 500 points dans \mathbb{R}^2 .**Définition 1.7.2**

Le quantificateur de Voronoi π_Γ induit par la mosaïque de Voronoi $(C_i(\Gamma))_{1 \leq i \leq N}$ est défini par la projection au plus proche voisin :

Ainsi, on peut définir la quantification de la variable aléatoire X par le quantificateur de Voronoi.

$$\forall y \in \mathbb{R}^d, \pi_\Gamma(y) = \sum_{i=1}^N x_i \mathbb{1}_{C_i(\Gamma)}(y)$$
Définition 1.7.3

La variable aléatoire \hat{X}^Γ définie par

$$\hat{X}^\Gamma = \pi_\Gamma(X) = \sum_{i=1}^N x_i \mathbb{1}_{\{C_i(\Gamma)\}}(X)$$

est appelée la quantification de Voronoi de la variable aléatoire X sur la grille Γ .

Quand il n'y a pas de confusion possible sur la grille Γ choisie, on simplifiera la notation \hat{X}^Γ en \hat{X} . La loi de probabilité discrète $\mathbb{P}_{\hat{X}}$ se caractérise pour tout $i \in \{1, \dots, N\}$, par :

$$p_i = \mathbb{P}_{\hat{X}}(x_i) = \mathbb{P}(\hat{X} = x_i) = \mathbb{P}(X \in C_i(\Gamma)) = \mathbb{P}_X(C_i(\Gamma)).$$

Les p_i sont appelés les masses des *cellules de Voronoi*.

Une grille donnée Γ de taille supérieure à 2 a une infinité de mosaïques de Voronoi. Cependant, toutes les cellules $C_i(\Gamma)$ ont la même fermeture et la même frontière, contenue dans au plus $N - 1$ hyperplans. Par conséquent, si la loi de X ne charge pas les hyperplans, c'est-à-dire si $\mathbb{P}(X \in H) = 0$ pour tout hyperplan H , alors la mosaïque de Voronoi est \mathbb{P} -presque sûrement unique.

Il a été montré que la fonction $\Gamma \mapsto \|X - \hat{X}^\Gamma\|_p^p$ pour une grille Γ de taille au plus N atteint toujours un minimum pour une grille Γ^* qui prend ses valeurs dans l'enveloppe convexe du support de \mathbb{P}_X . Ce résultat se démontre par récurrence sur N dans Bally et Pagès (2003). Par ailleurs, le théorème de Zador permet d'obtenir une vitesse de convergence de $\|X - \hat{X}^\Gamma\|_p^p$ en $O(N^{-\frac{p}{d}})$ (voir théorème II.6.2 de Graf et Luschgy (2000)).

1.7.2 Quantification d'une chaîne de Markov

On s'intéresse dans cette partie à la quantification d'une chaîne de Markov $(X_n)_{n \geq 0}$ à valeurs dans \mathbb{R}^d , définie sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$. On note P_n ses noyaux de transition du temps n au temps $n+1$ et μ la loi initiale de la chaîne. On souhaite approcher la loi du processus (X_n) par la loi d'un processus (\hat{X}_n) à valeurs dans des ensembles finis.

Il existe deux méthodes pour quantifier une chaîne de Markov : la quantification marginale et la quantification Markovienne (voir Pagès *et al.* (2004)). Ces deux méthodes sont basées sur le fait que l'on peut entièrement caractériser une chaîne de Markov par sa distribution initiale et par ses noyaux de transition. Le principe de la quantification marginale repose sur l'approximation de la distribution marginale, à chaque pas de temps n . Cette méthode implique la perte de la propriété de Markov. En revanche, la seconde méthode, dite de quantification Markovienne, préserve cette propriété mais requiert des hypothèses plus fortes. Dans le cadre de notre étude, on s'intéressera uniquement à la quantification marginale de la chaîne de Markov. En effet, notre étude se situe dans la continuité du travail réalisé dans de Saporta *et al.* (2015), qui utilise la quantification marginale. Les lecteurs intéressés par la quantification Markovienne pourront se référer à Pagès *et al.* (2004) pour de plus amples détails.

La quantification marginale se fait en deux étapes. La première consiste à quantifier vectoriellement les marginales de la chaîne de Markov. Cela signifie que si au rang n , on se donne une grille $\Gamma_n = \{x_n^1, \dots, x_n^N\}$ de N points dans \mathbb{R}^d , à qui l'on associe une mosaïque de Voronoi $(C_i(\Gamma_n))_{i=1, \dots, N}$, alors on définit :

$$\hat{X}_n = \pi_{\Gamma_n}(X_n) = \sum_{i=1}^N x_n^i \mathbf{1}_{\{C_i(\Gamma_n)\}}(X_n).$$

Autrement dit, \hat{X}_n est le plus proche voisin de X_n dans la grille Γ_n . Néanmoins, l'application de projection π n'est pas injective. Le processus (\hat{X}_n) ainsi construit ne conserve donc pas la propriété de Markov.

En revanche, on peut construire une chaîne de Markov, à partir du processus projeté (\hat{X}_n) . On pose

$$p_{k-1}^i = \mathbb{P}[\hat{X}_{k-1} = x_{k-1}^i] = \mathbb{P}[X_{k-1} \in C_i(\Gamma_{k-1})], \quad (1.12)$$

$$\beta_k^{i,j} = \mathbb{P}[(\hat{X}_k, \hat{X}_{k-1}) = (x_k^j, x_{k-1}^i)] = \mathbb{P}[(X_k, X_{k-1}) \in C_j(\Gamma_k) \times C_i(\Gamma_{k-1})], \quad (1.13)$$

$$M_k^{i,j} = \mathbb{P}[\hat{X}_k = x_k^j | \hat{X}_{k-1} = x_{k-1}^i] = \frac{\beta_k^{i,j}}{p_{k-1}^i}. \quad (1.14)$$

Il existe une chaîne de Markov (\hat{X}_k^c) de loi initiale p_0 et de matrices de transition $(M_k^{i,j})_{i,j}$. La quantification marginale consiste à approcher la distribution de la chaîne (X_n) par celle de la chaîne (\hat{X}_k^c) ainsi construite.

L'algorithme lié à la quantification marginale est décrit dans Pham *et al.* (2005) et Bally et Pagès (2003). Il s'agit d'une adaptation de l'algorithme CLVQ (voir Pagès et Printems (2003)) qui permet de quantifier une variable aléatoire en norme L^2 . On pose $N_{\text{quantif}} \in \mathbb{N}^*$ l'horizon de quantification, c'est-à-dire que l'on quantifie la chaîne $(X_n)_{0 \leq n \leq N_{\text{quantif}}}$.

Entrées de l'algorithme

- S le nombre de réplifications,
- X^1, \dots, X^S des simulations de la chaîne de Markov que l'on souhaite quantifier, où $X^i = (X_0^i, \dots, X_{N_{quantif}}^i)$ pour tout $i \in \{1, \dots, S\}$, avec $X_0^i = x_0$ un état initial,
- (δ_t) une suite de pas positifs répondant aux critères

$$\sum_{t \geq 0} \delta_t = +\infty \quad \text{et} \quad \sum_{t \geq 0} \delta_t^2 < +\infty. \quad (1.15)$$

Étape 1 : Initialisation

- On initialise la grille $\Gamma_0^0 = x_0$, réduite à un point et les $N_{quantif}$ grilles suivantes par $\Gamma_n = \{\gamma_n^{0,1}, \dots, \gamma_n^{0,N}\}$ pour $n \in \{1, \dots, N_{quantif}\}$.
- On initialise les vecteurs compteurs marginaux :

$$\alpha_0^0 = 1 \quad \text{et} \quad \alpha_n^{0,i} = 1 \quad \text{pour} \quad i \in \{1, \dots, N\} \quad \text{et} \quad n \in \{1, \dots, N_{quantif}\}.$$

- On initialise les matrices compteurs de transition :

$$\beta_n^0 = (\beta_n^{0,(i,j)}) = (0) \quad \text{pour} \quad i, j \in \{1, \dots, N\} \quad \text{et} \quad n \in \{1, \dots, N_{quantif}\}.$$

Étape 2 : Mise à jour des grilles

La phase de mise à jour consiste en une boucle dans laquelle on met à jour les grilles de quantification, la distorsion et les compteurs. On note s la variable correspondant au numéro de la dernière réplification. Les étapes de la boucle sont détaillées ci-dessous.

1. pour $k = 0$,

- mise à jour du compteur marginal : $\alpha_0^{s+1} = \alpha_0^s + 1$,
- stockage de l'indice du plus proche voisin de X_0^{s+1} dans Γ_0^s : $oldwin = 1$.

2. Pour k allant de 1 à $N_{quantif}$, on exécute les lignes suivantes :

(a) *phase de compétition*

On cherche le plus proche voisin de X_k^{s+1} dans la grille Γ_k^s , ce qui revient à sélectionner l'indice win parmi $\{1, \dots, N\}$, tel que

$$X_k^{s+1} \in C_{win}(\Gamma_k^s).$$

(b) *phase d'apprentissage*

Mise à jour de la grille $\Gamma_k^{s+1} = (\gamma_k^{s+1,1}, \dots, \gamma_k^{s+1,n})$:

$$\forall i \in \{1, \dots, N\}, \quad \gamma_k^{s+1,i} = \gamma_k^{s,i} - \delta_{s+1} \mathbb{1}_{\{i=win\}} (\gamma_k^{s,i} - X_k^{s+1}),$$

mise à jour des vecteurs compteurs marginaux :

$$\forall i \in \{1, \dots, N\}, \quad \alpha_k^{s+1,i} = \alpha_k^{s,i} + \mathbb{1}_{\{i=win\}},$$

pour $k \geq 1$, mise à jour des matrices compteurs de transition :

$$\forall i, j \in \{1, \dots, N\}, \quad \beta_{k-1}^{s+1,(i,j)} = \beta_{k-1}^{s,(i,j)} + \mathbb{1}_{\{i=oldwin, j=win\}}.$$

(c) Stockage de l'indice du plus proche voisin de X_k^{s+1} dans la grille Γ_k^s :

$$oldwin = win.$$

Étape 3 : Matrices de transition

Pour $k \in \{1, \dots, N-1\}$, $i \in \{1, \dots, n_k\}$ et $j \in \{1, \dots, n_{k+1}\}$, on calcule les matrices de transitions

$$M_k^{S,(i,j)} = \frac{\beta_k^{S,(i,j)}}{\alpha_k^{S,i}}$$

et pour $k = 0$,

$$M_0^{S,(1,j)} = \frac{\beta_0^{S,(1,j)}}{\alpha_0^{S,1}}.$$

Sorties de l'algorithme

- $(\Gamma_k^S)_{0 \leq k \leq N}$ grilles de quantification,
- $(M_k^{S,(i,j)})_{0 \leq k \leq N-1}$ matrices de transition de la chaîne de Markov approchée.

Utilisation pratique de l'algorithme

Plusieurs choix doivent être faits pour mettre en pratique cet algorithme. Il s'agit du choix de la grille initiale Γ^0 , des masses des cellules de Voronoi associées à cette grille initiale et du choix de la suite (δ_t) . Ces choix sont importants car influencent la convergence de l'algorithme.

Tout d'abord, il existe deux méthodes pour choisir la grille initiale Γ^0 . Si N est relativement petit, on peut utiliser une initialisation aléatoire : on simule N variables aléatoires indépendantes de loi \mathbb{P}_X que l'on choisit pour former la grille Γ_0 . Pour des valeurs de N plus grandes, il existe la méthode appelée *splitting-initializing method*. Cette méthode consiste à ajouter un point à une grille optimale de taille $N-1$ pour obtenir une grille d'initialisation de taille N .

Le choix le plus courant pour l'initialisation des masses p_i^0 des cellules de Voronoi associées à la grille Γ^0 est de prendre toutes les masses égales à $\frac{1}{N}$.

La suite de pas (δ_t) doit vérifier les critères donnés en (1.15). Le choix de cette suite est discuté dans Pagès et Printems (2003) dans le cas de l'algorithme CLVQ. Couramment, on prend une suite telle que

$$\delta_t > \frac{2N^{\frac{3}{d}}}{\pi^2 t}$$

qui correspond au seuil critique pour une loi uniforme pour valider le théorème central limite quand t est assez grand. On pose :

$$\delta_t = \delta_0 \frac{a}{a + \delta_0 b t},$$

où $a = 4N^{\frac{1}{d}}$ et $b = \pi^2 N^{-\frac{2}{d}}$. Pour δ_0 , on peut prendre au choix $\delta_0 = 1$ ou bien

$$\delta_0 = \frac{1}{2} \min_{i \neq j} |x_i - x_j|.$$

1.8 Approximation numérique de la fonction valeur

On présente dans ce paragraphe la méthode d'approximation numérique de la fonction valeur proposée dans de Saporta et Dufour (2012). Cette méthode consiste à implémenter deux algorithmes successifs. Le premier calcule, pour tous les points y appartenant à l'ensemble de contrôle \mathbb{U} , des approximations des fonctions valeurs pour des caméras partant de l'état initial y . Ces approximations servent ensuite pour le second algorithme, qui détermine une approximation de la fonction valeur de notre problème de contrôle impulsif.

1.8.1 Définitions

On suppose que la loi initiale de X_0 est donnée par la distribution de Dirac δ_{x_0} , où x_0 est un point de l'espace d'état E . On choisit une fonction g dans $\mathbf{L}_\Phi(E)$ satisfaisant l'inégalité :

$$\forall x \in E, \quad g(x) \geq h(x) = \mathbb{E}_x \left[\int_0^\infty e^{-\alpha s} f(X_s) ds \right]. \quad (1.16)$$

On rappelle que h correspond au coût de la politique de non-intervention.

L'approximation de la fonction valeur proposée dans de Saporta *et al.* (2015) est basée sur la proposition 1.6.1. On fixe un horizon de contrôle N_0 . Le choix de cet horizon est détaillé dans de Saporta *et al.* (2015) et Costa et Davis (1989). En pratique, pour notre application, ce choix est expliqué aux paragraphes 4.2.1 et 4.3.6. L'horizon de contrôle correspond au moment où il n'est plus possible d'intervenir sur le processus. Plus précisément, les N_0 premiers sauts du processus peuvent être soit des sauts naturels issus du processus non contrôlé, soit des interventions. Après ce N_0 -ième saut, le processus contrôlé suit la même dynamique que le processus non contrôlé ; tous les sauts du processus sont alors des sauts naturels. Cette notion est expliquée plus en détails au paragraphe 5.2, où elle est notamment illustrée par les figures 5.1 et 5.2.

On commence par définir une suite de fonctions valeurs $(v_n)_{0 \leq n \leq N_0}$ par

$$\begin{cases} v_{N_0} = g, \\ \forall n \in \{0, \dots, N_0 - 1\}, \quad v_n = \mathcal{L}v_{n+1}. \end{cases}$$

Par rapport à la proposition 1.6.1, les indices des fonctions valeurs sont inversés. Le but est d'obtenir une approximation de la fonction valeur v_0 au point x_0 .

La procédure d'approximation se fait en trois étapes. Premièrement, on remplace l'infimum sur \mathbb{R}^+ dans l'opérateur J par un minimum sur une grille de temps dépendant d'un point de l'espace z , incluse dans l'intervalle $[0; t^*(z)]$. Deuxièmement, on construit une approximation des fonctions valeurs $(v_n)_{0 \leq n \leq N_0}$ en les points de l'ensemble de contrôle \mathbb{U} . Les approximations ainsi obtenues nous servent pour la dernière étape, qui consiste à construire une approximation de la fonction valeur v_0 prise au point x_0 .

Discrétisation du temps

Définition 1.8.1

Pour $z \in E$, on choisit un pas de temps $\Delta(z)$ dans l'intervalle $]0; t^*(z)[$ et on définit $n(z) = \lfloor \frac{t^*(z)}{\Delta(z)} \rfloor - 1$, où $\lfloor x \rfloor$ désigne la plus grand entier inférieur ou égal à x . On définit la grille de points $G(z)$ associée à l'intervalle de temps $[0; t^*(z)]$ par :

$$G(z) = \{i\Delta(z), i \in \{0, \dots, n(z)\}\}.$$

À l'aide de cette discrétisation de l'intervalle $[0; t^*(z)]$, on peut définir une approximation de l'opérateur L défini initialement par la formule (1.10), en remplaçant l'infimum pour $t \in \mathbb{R}^+$ par un minimum sur la grille de temps adaptée $G(z)$:

$$\forall (v, w) \in \mathbf{L}_\Phi(E)^2, \quad \forall x \in E, \quad L^d(v, w)(x) = \min_{t \in G(x)} J(v, w)(x, t) \wedge Kw(x). \quad (1.17)$$

En effet, par définition de l'opérateur J donné par la formule (1.7), on a l'égalité :

$$\forall (v, w) \in \mathbf{L}_\Phi(E)^2, \quad \forall x \in E, \quad \inf_{t \in \mathbb{R}^+} J(v, w)(x, t) = \inf_{t \in [0; t^*(z)]} J(v, w)(x, t).$$

Approximation des fonctions valeurs sur \mathbb{U}

On rappelle qu'il existe une chaîne de Markov sous-jacente au processus de Markov déterministe par morceaux $\{X_t\}$, définie par $\Theta_n = (Z_n, S_n)$, où $Z_n = X_{T_n}$ sont les positions du processus après chaque saut et S_n sont les temps inter-sauts du processus.

Dans cette partie, on cherche à approcher les fonctions valeurs sur l'ensemble de contrôle :

$$\mathbb{U} = \{y^i, 1 \leq i \leq u\}.$$

Pour cela, on s'intéresse au PDMP $\{X_t\}$ partant d'un point de contrôle y^i appartenant à \mathbb{U} , où $i \in \{1, \dots, u\}$. On note $\Theta_n^i = (Z_n^i, S_n^i)$ la chaîne sous-jacente associée à ce processus. En quantifiant le processus $\{\Theta_n^i\}$ sur un horizon $N_0 - 1$ à l'aide de l'algorithme développé au paragraphe 4.1, on obtient des grilles finies $(\Gamma_n^{i,\Theta})_{0 \leq n \leq N_0-1}$ et leurs matrices de transition $(M_n^i)_{0 \leq n \leq N_0-2}$. On note :

- $\pi_{\Gamma_n^{i,\Theta}}$ la projection au plus proche voisin d'un élément de $E \times \mathbb{R}^+$ sur la grille $\Gamma_n^{i,\Theta}$
- $(\hat{Z}_n^i, \hat{S}_n^i) := \pi_{\Gamma_n^{i,\Theta}}(Z_n^i, S_n^i)$
 $(\hat{Z}_n^i, \hat{S}_n^i)_{0 \leq n \leq N_0-1}$ est l'approximation quantifiée de la suite $(Z_n^i, S_n^i)_{0 \leq n \leq N_0-1}$,
- $\hat{\Theta}_n^i := (\hat{Z}_n^i, \hat{S}_n^i)$,
- $\Gamma_n^{i,Z}$ l'ensemble des points de la grille $\Gamma_n^{i,\Theta}$ projetés sur E ,
- $\Gamma_n^{i,S}$ l'ensemble des points de la grille $\Gamma_n^{i,\Theta}$ projetés sur \mathbb{R}^+ .

On définit les approximations quantifiées des opérateurs K , J et L^d (initialement définis respectivement par les formules (1.8), (1.7) et (1.17)), pour $i \in \{1, \dots, u\}$ et $n \in \{0, \dots, N_0 - 1\}$, pour $v \in \mathbf{L}_\Phi(E)$ et $w \in \mathbf{B}(\Gamma_{n+1}^{i,Z})$, et pour $z \in \Gamma_n^{i,Z}$,

$$\hat{K}_{n+1}^i w(z) = F(z, t^*(z)) + \mathbb{E}_{y^i} \left[e^{-\alpha \hat{S}_{n+1}^i} w(\hat{Z}_{n+1}^i) | \hat{Z}_n^i = z \right], \quad (1.18)$$

$$\begin{aligned} \hat{J}_{n+1}^i(v, w)(z, t) = & F(z, t) + \mathbb{E}_{y^i} \left[e^{-\alpha(t \wedge t^*(z))} v(\Phi(z, t \wedge t^*(z))) \mathbf{1}_{\{\hat{S}_{n+1}^i \geq t \wedge t^*(z)\}} \right. \\ & \left. + e^{-\alpha \hat{S}_{n+1}^i} w(\hat{Z}_{n+1}^i) \mathbf{1}_{\{\hat{S}_{n+1}^i < t \wedge t^*(z)\}} | \hat{Z}_n^i = z \right], \end{aligned} \quad (1.19)$$

$$\hat{L}_{n+1}^{i,d}(v, w)(z) = \min_{t \in G(x)} \hat{J}_{n+1}^i(v, w)(z, t) \wedge \hat{K}_{n+1}^i w(z). \quad (1.20)$$

L'approximation des fonctions valeurs sur l'ensemble de contrôle est faite par retour en arrière dans le temps. Cette procédure est triangulaire.

Définition 1.8.2

On définit la fonction valeur \tilde{v}_{N_0} sur \mathbb{U} par :

$$\forall i \in \{1, \dots, u\}, \quad \tilde{v}_{N_0}(y^i) = g(y^i)$$

et pour $k \in \{1, \dots, N_0 - 1\}$, on définit les fonctions valeurs \tilde{v}_k sur \mathbb{U} par :

$$\forall i \in \{1, \dots, u\}, \quad \tilde{v}_k(y^i) = \hat{v}_k^{i,k}(y^i),$$

où les fonctions $(\hat{v}_n^{i,k})_{i \in \{1, \dots, u\}, k \in \{1, \dots, N_0\}, n \in \{k, \dots, N_0\}}$ sont définies par :

$$\begin{cases} \forall z \in \Gamma_{N_0-k}^{i,Z}, & \hat{v}_{N_0}^{i,k}(z) = g(z), \\ \forall n \in \{1, \dots, N_0 - k\}, \forall z \in \Gamma_{n-1}^{i,Z}, & \hat{v}_{k+n-1}^{i,k}(z) = \hat{L}_n^{i,d}(M \tilde{v}_{k+n}, \hat{v}_{k+n}^{i,k})(z). \end{cases}$$

On peut ainsi approcher les fonctions valeurs $(v_n)_{0 \leq n \leq N_0}$ sur l'ensemble des points de la grille de contrôle \mathbb{U} .

On définit les constantes suivantes :

$$\begin{aligned} E_1 &= C_{t^*}[\lambda]_1 + (C_t + \alpha)L_{t^*} \\ E_2 &= C_\lambda L_{t^*} + [\lambda]_1 \frac{1 + C_\lambda C_{t^*}}{\alpha} \\ E_3 &= \frac{[f]_1}{\alpha} + C_f \left(\frac{C_{t^*}[\lambda]_1}{\alpha} + L_{t^*} \right). \end{aligned}$$

Le taux de convergence de cette approximation est donné par le théorème 10.1 de de Saporta *et al.* (2015), que l'on rappelle ci-dessous.

Théorème 1.8.1

Pour tous $1 \leq k \leq N_0 - 1$, $0 \leq n \leq N_0 - k - 1$ et $1 \leq i \leq u$, on suppose que $\Delta(z)$ pour $z \in \Gamma_n^{i,Z}$ est tel que

$$\left(\frac{2}{C_\lambda} \left(\|S_{n+1}^i - \hat{S}_{n+1}^i\|_p + L_{t^*} \|Z_n^i - \hat{Z}_n^i\|_p \right) \right)^{1/2} < \min_{z \in \Gamma_n^{i,Z}} \Delta(z).$$

Alors on a

$$\begin{aligned} & \|v_{k+n}(Z_n^i) - \hat{v}_{k+n}^{i,k}(\hat{Z}_n^i)\|_p \\ & \leq \|v_{k+n+1}(Z_{n+1}^i) - \hat{v}_{k+n+1}^{i,k}(\hat{Z}_{n+1}^i)\|_p + \max_{y \in \mathbb{U}} |v_{k+n+1}(y) - \tilde{v}_{k+n+1}(y)| \\ & \quad + d_{k,n}^1 \|Z_n^i - \hat{Z}_n^i\|_p + d_{k,n}^2 \|\Delta(\hat{Z}_n^i)\|_p + 3L_{v_{k+n+1}} \|Z_{n+1}^i - \hat{Z}_{n+1}^i\|_p \\ & \quad + 2\alpha C_{v_{k+n+1}} \|S_{n+1}^i - \hat{S}_{n+1}^i\|_p \\ & \quad + 2(2C_{v_{k+n+1}} + C_c) \left(2C_\lambda \left(\|S_{n+1}^i - \hat{S}_{n+1}^i\|_p + L_{t^*} \|Z_n^i - \hat{Z}_n^i\|_p \right) \right)^{1/2}, \end{aligned}$$

avec

$$\begin{aligned} d_{k,n}^1 &= (2[c]_1 + 2[c]_2 L_{t^*} + C_c E_1 + \alpha L_{t^*} (C_{v_{k+n+1}} + C_c)) \vee ([Q][v_{k+n+1}]_* + 2E_3) \\ & \quad + \frac{C_\lambda}{\alpha} [Q][v_{k+n+1}]_1 + C_{v_{k+n+1}} (E_1 + E_2) + L_{v_{k+n}}, \\ d_{k,n}^2 &= 2 \left(C_f + C_{v_{k+n+1}} C_\lambda + [c]_2 + (C_c + C_{v_{k+n+1}})(C_\lambda + \alpha) \right). \end{aligned}$$

Il est à noter que la définition 1.8.2 des fonctions valeurs $(\hat{v}_n^{i,k})_{i \in \{1, \dots, u\}, k \in \{1, \dots, N_0\}, n \in \{k, \dots, N_0\}}$ implique la connaissance de la fonction g en tout point des grilles $\Gamma_n^{i,Z}$, pour tout $n \in \{0, \dots, N_0 - 1\}$ et pour tout $i \in \{1, \dots, u\}$. C'est la raison pour laquelle nous n'initialisons pas les fonctions valeurs avec la fonction h du coût moyen associé à la politique de non-intervention, comme dans la proposition 1.6.1. Il est généralement très difficile de calculer la fonction h sur tous les points des grilles de quantification. Par conséquent, on prend une fonction g supérieure ou égale à h , qui peut être par exemple une constante.

Approximation de la fonction valeur $v_0(x_0)$

Tout comme pour l'approximation des fonctions valeurs sur l'ensemble de contrôle \mathbb{U} , nous avons besoin ici de la version quantifiée de la chaîne de Markov $\Theta_n = (Z_n, S_n)$ sous-jacente au PDMP $\{X_t\}$. En quantifiant le processus Θ_n à l'aide de l'algorithme de la partie 1.7.2 avec un horizon N_0 , on obtient une série de grilles $(\Gamma_n^\Theta)_{0 \leq n \leq N_0}$ et leurs matrices de transitions associées $(M_n)_{0 \leq n \leq N_0-1}$. On note :

- $\pi_{\Gamma_n^\Theta}$ la projection au plus proche voisin d'un élément de $E \times \mathbb{R}^+$ sur la grille Γ_n^Θ ,
- $(\hat{Z}_n, \hat{S}_n) := \pi_{\Gamma_n^\Theta}(Z_n, S_n)$
 $(\hat{Z}_n, \hat{S}_n)_{0 \leq n \leq N_0}$ l'approximation quantifiée de la suite $(Z_n, S_n)_{0 \leq n \leq N_0}$,
- $\hat{\Theta}_n := (\hat{Z}_n, \hat{S}_n)$,
- Γ_n^Z l'ensemble des points de la grille Γ_n^Θ projetés sur E ,
- Γ_n^S l'ensemble des points de la grille Γ_n^Θ projetés sur \mathbb{R}^+ .

On définit de nouvelles versions quantifiées des opérateurs K , J et L de la manière suivante. Pour $n \in \{0, \dots, N_0 - 1\}$, $v \in \mathbf{L}_\Phi(E)$, $w \in \mathbf{B}(\Gamma_{n+1}^Z)$ et pour $z \in \Gamma_n^Z$,

$$\hat{K}_{n+1}w(z) = F(z, t^*(z)) + \mathbb{E}_{x_0} \left[e^{-\alpha \hat{S}_{n+1}} w(\hat{Z}_{n+1}) | \hat{Z}_n = z \right], \quad (1.21)$$

$$\begin{aligned} \hat{J}_{n+1}(v, w)(z, t) = & F(z, t) + \mathbb{E}_{x_0} \left[e^{-\alpha(t \wedge t^*(z))} v(\Phi(z, t \wedge t^*(z))) \mathbf{1}_{\{\hat{S}_{n+1} \geq t \wedge t^*(z)\}} \right. \\ & \left. + e^{-\alpha \hat{S}_{n+1}} w(\hat{Z}_{n+1}) \mathbf{1}_{\{\hat{S}_{n+1} < t \wedge t^*(z)\}} | \hat{Z}_n = z \right], \end{aligned} \quad (1.22)$$

$$\hat{L}_{n+1}^d(v, w)(z) = \min_{t \in G(x)} \hat{J}_{n+1}(v, w)(z, t) \wedge \hat{K}_{n+1}w(z). \quad (1.23)$$

Ainsi, on peut définir une nouvelle suite d'approximations de fonctions valeurs, qui permet d'obtenir une approximation de la fonction valeur $v_0(x_0)$ recherchée, en utilisant les approximations $(\tilde{v}_n)_{1 \leq n \leq N_0}$ des fonctions valeurs sur l'ensemble de contrôle \mathbb{U} .

Définition 1.8.3

On définit la fonction valeur \hat{v}_{N_0} par :

$$\forall z \in \Gamma_{N_0}^Z, \quad \hat{v}_{N_0}(z) = g(z),$$

et pour $k \in \{1, \dots, N_0\}$, les fonctions valeurs \hat{v}_{k-1} par

$$\forall z \in \Gamma_{k-1}^Z, \quad \hat{v}_{k-1}(z) = \hat{L}_k^d(M \tilde{v}_k, \hat{v}_k)(z).$$

L'approximation de la fonction valeur $v_0(x_0)$ correspond à $\hat{v}_0(\hat{Z}_0)$. Le taux de convergence de cette approximation est donné par le théorème 10.2 du livre de Saporta *et al.* (2015), rappelé ci-dessous :

Théorème 1.8.2

Pour tout $0 \leq n \leq N_0 - 1$, on suppose que $\Delta(z)$ pour $z \in \Gamma_n^Z$ est tel que

$$\left(\frac{2}{C_\lambda} \left(\|S_{n+1} - \hat{S}_{n+1}\|_p + L_{t^*} \|Z_n - \hat{Z}_n\|_p \right) \right)^{1/2} < \min_{z \in \Gamma_n^Z} \Delta(z).$$

Alors on a

$$\begin{aligned}
& \|v_n(Z_n) - \hat{v}_n(\hat{Z}_n)\|_p \\
& \leq \|v_{n+1}(Z_{n+1}) - \hat{v}_{n+1}(\hat{Z}_{n+1})\|_p + \max_{y \in \mathbb{U}} |v_{n+1}(y) - \tilde{v}_{n+1}(y)| \\
& + D_n^1 \|Z_n - \hat{Z}_n\|_p + D_n^2 \|\Delta(\hat{Z}_n)\|_p + 3L_{v_{n+1}} \|Z_{n+1} - \hat{Z}_{n+1}\|_p \\
& + 2\alpha C_{v_{n+1}} \|S_{n+1} - \hat{S}_{n+1}\|_p \\
& + 2(2C_{v_{n+1}} + C_c) \left(2C_\lambda (\|S_{n+1} - \hat{S}_{n+1}\|_p + L_{t^*} \|Z_n - \hat{Z}_n\|_p) \right)^{1/2},
\end{aligned}$$

avec

$$\begin{aligned}
D_n^1 &= (2[c]_1 + 2[c]_2 L_{t^*} + C_c E_1 + \alpha L_{t^*} (C_{v_{n+1}} + C_c)) \vee ([Q][v_{n+1}]_* + 2E_3) \\
&+ \frac{C_\lambda}{\alpha} [Q][v_{n+1}]_1 + C_{v_{n+1}} (E_1 + E_2) + L_{v_n}, \\
D_n^2 &= 2(C_f + C_{v_{n+1}} C_\lambda + [c]_2 + (C_c + C_{v_{n+1}})(C_\lambda + \alpha)).
\end{aligned}$$

1.8.2 Algorithmes

L'approximation de la fonction valeur $v_0(x_0)$ est obtenue par la procédure récursive décrite dans la définition 1.8.3. Cette procédure nécessite la connaissance des fonctions $(\tilde{v}_n(y^i))_{1 \leq n \leq N_0, 1 \leq i \leq u}$, issues de la procédure triangulaire récursive présentée dans la définition 1.8.2. Plus précisément, les étapes de l'approximation de la fonction valeur $v_0(x_0)$ sont schématisées par la figure 1.6 et sont récapitulées ici :

- (a) Quantification du processus $\{\Theta_n^i\}_{0 \leq n \leq N_0-1}$:
 $N_{\mathbb{U}}$ nombre de points dans les grilles de quantification $(\Gamma_n^{i,\Theta})_{1 \leq n \leq N_0-1}$,
 $S_{\mathbb{U}}$ nombre de réplifications utilisées pour la création des grilles de quantification
 \Rightarrow Obtention des grilles $(\Gamma_n^{i,\Theta})_{0 \leq n \leq N_0-1}$ et de leurs matrices de transition $(M_n^i)_{0 \leq n \leq N_0-2}$.
- (b) Approximation des fonctions valeurs sur l'ensemble de contrôle \mathbb{U} :
Procédure triangulaire récursive décrite par la définition 1.8.2.
De manière détaillée, on commence par définir $\tilde{v}_{N_0} = g$ et en supposant que l'on a calculé l'ensemble des fonctions \tilde{v}_k , pour $k > j$, où j est un indice fixé, on calcule la fonction \tilde{v}_j en y^1 , puis en y^2 et ainsi de suite jusqu'à y^u . Chaque calcul de valeur $\tilde{v}_j(y^i)$ demande de calculer les fonctions $(\hat{v}_n^{i,j})_{j \in \{k, \dots, N_0\}}$.
- (c) Quantification du processus $\{\Theta_n\}_{0 \leq n \leq N_0}$:
 N_x nombre de points dans les grilles de quantification $(\Gamma_n^\Theta)_{1 \leq n \leq N_0}$,
 S_x nombre de réplifications utilisées pour la création des grilles de quantification
 \Rightarrow Obtention des grilles $(\Gamma_n^\Theta)_{0 \leq n \leq N_0}$ et de leurs matrices de transition $(M_n)_{0 \leq n \leq N_0-1}$.
- (d) Approximation de la fonction valeur $v_0(x_0)$:
Procédure en ligne récursive décrite par la définition 1.8.3.

Les étapes (a) et (c) de quantification peuvent se faire au préalable de l'approximation complète de la fonction valeur. Les étapes (b) et (d) sont résumées dans la figure 1.7 où le triangle de gauche du schéma représente l'étape (b) et la colonne de droite du schéma représente l'étape (d).

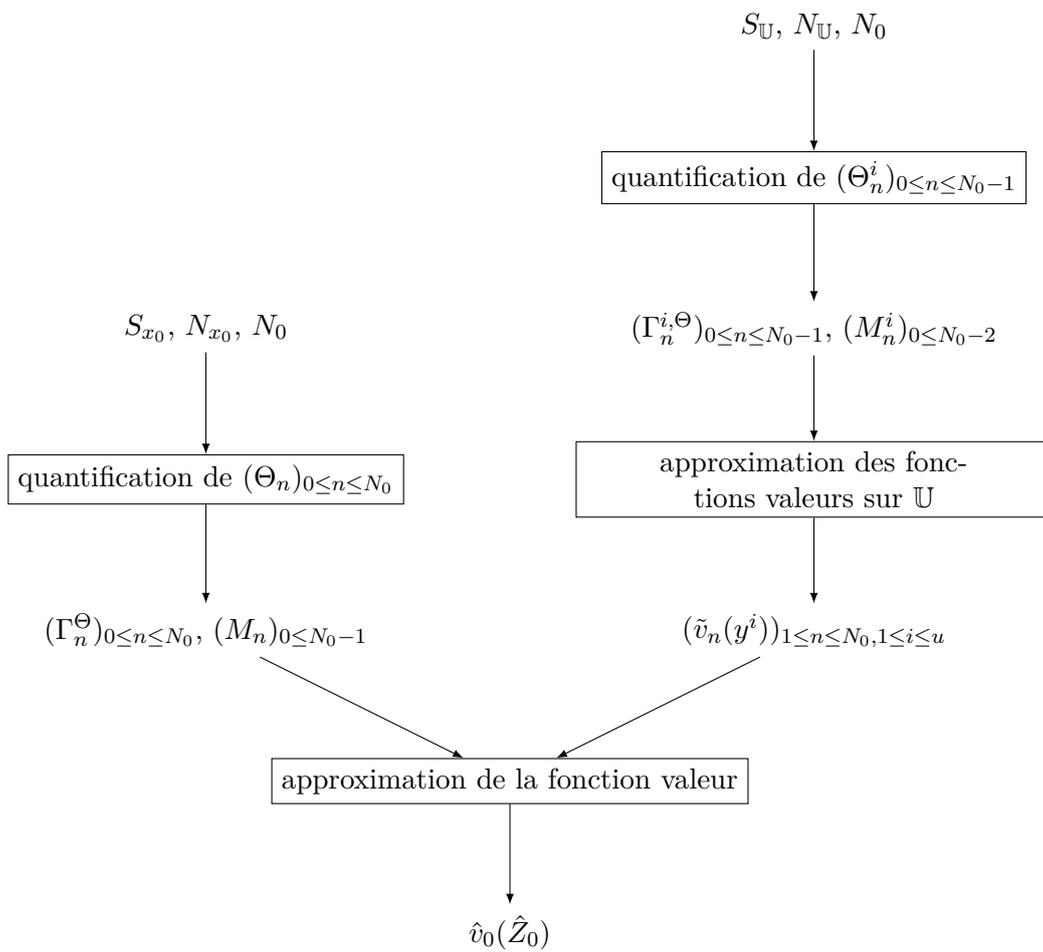


FIGURE 1.6 – Structure complète de l'algorithme de l'approximation de la fonction valeur

APPROXIMATION DES FONCTIONS VALEURS SUR \mathbb{U} (ALGORITHME 2)

APPROXIMATION DE
LA FONCTION VALEUR
(ALGORITHME 3)

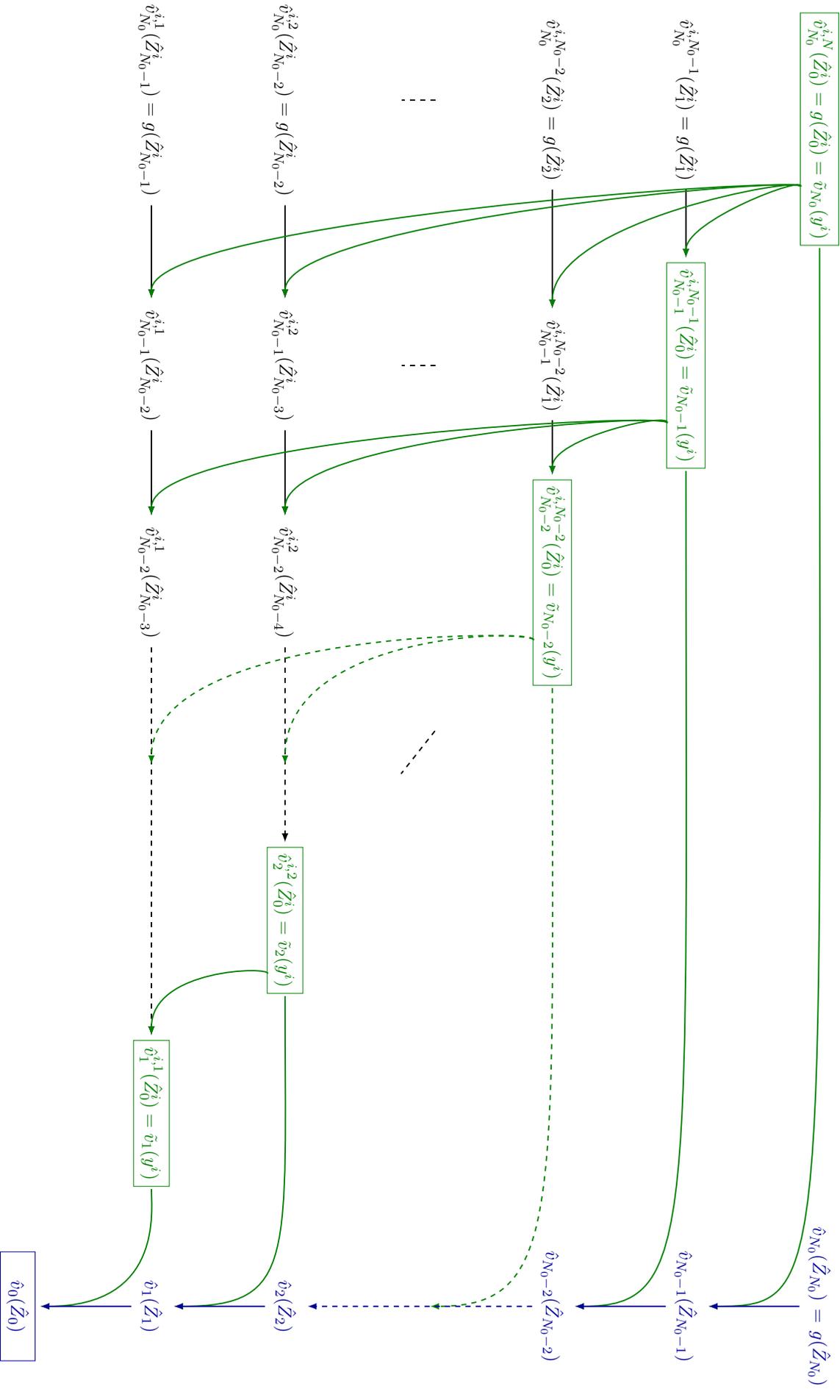


FIGURE 1.7 – Procédure détaillée de l'approximation de la fonction valeur (après avoir réalisé la quantification)

L'approximation des fonctions valeurs sur l'ensemble de contrôle \mathbb{U} (étape (b)) correspond à l'algorithme 1. Dans le pseudo-code, on utilise la notation $\Gamma_n^{i,\Theta}(j)$ pour désigner le j -ième point de la grille de quantification $\Gamma_n^{i,\Theta}$. De même, on utilise la notation $M_k^i(j, l)$ pour désigner la probabilité de transition du j -ième point de la grille $\Gamma_{k-1}^{i,\Theta}$ au l -ième point de la grille Γ_k^Θ .

Algorithme 1 Approximation des fonctions valeurs sur l'ensemble de contrôle \mathbb{U}

Entrées :

$(\Gamma_n^{i,\Theta})_{0 \leq n \leq N_0-1, 1 \leq i \leq u}$ grilles de quantification

$(M_n^i)_{0 \leq n \leq N_0-2, 1 \leq i \leq u}$ matrices de transition associées aux grilles

Sortie :

$(\tilde{V}(n, i))_{1 \leq n \leq N_0, 1 \leq i \leq u}$ matrice dont les éléments sont les fonctions valeurs $(\tilde{v}_n(y^i))_{1 \leq n \leq N_0, 1 \leq i \leq u}$

```

1   $N_n^{i,\Theta}$  le nombre de points de la grille de quantification  $\Gamma_n^{i,\Theta}$ 
2   $\tilde{V}$  matrice de zéros de taille  $N_0 \times u$ 
3  pour  $i=1 : u$  faire
4       $\tilde{V}(N_0, i) = g(y^i)$ 
5  fin pour
6  pour  $n = (N_0 - 1) : -1 : 1$  faire
7      pour  $i = 1 : u$  faire
8           $\hat{V}$  matrice de zéros de taille  $N_0 \times N_{N_0-n}^{i,\Theta}$ 
9          pour  $l = 1 : N_{N_0-n}^{i,\Theta}$  faire
10              $\hat{V}(N_0, l) = g(\Gamma_{N_0-n}^{i,Z}(l))$ 
11          fin pour
12          pour  $k = (N_0 - n - 1) : -1 : 0$  faire
13              pour  $j = 1 : N_k^{i,\Theta}$  faire
14                   $z = \Gamma_k^{i,Z}(j)$ 
15                   $K = F(z, t^*(z)) + \sum_{l=1}^{N_{k+1}^{i,\Theta}} e^{-\alpha \Gamma_{k+1}^{i,S}(l)} \hat{V}(k+n+1, l) M_k^i(j, l)$ 
16                   $m, J_1, J_2$  et  $J$  vecteurs lignes de taille  $n(z) + 1$ , initialisés à 0
17                  pour  $r = 1 : (n(z) + 1)$  faire
18                       $t = (r - 1) \Delta(z)$ 
19                       $m(r) = \min_{1 \leq l \leq u} \{c(\Phi(z, t), y^l) + \tilde{V}(k+n+1, l)\}$ 
20                       $J_1(r) = \sum_{l=1}^{N_{k+1}^{i,\Theta}} e^{-\alpha \Gamma_{k+1}^{i,S}(l)} \hat{V}(k+n+1, l) M_k^i(j, l) \mathbb{1}_{\{\Gamma_{k+1}^{i,S}(l) < t\}}$ 
21                       $J_2(r) = \sum_{l=1}^{N_{k+1}^{i,\Theta}} e^{-\alpha t} m(t) M_k^i(j, l) \mathbb{1}_{\{\Gamma_{k+1}^{i,S}(l) \geq t\}}$ 
22                       $J(r) = F(z, t) + J_1(r) + J_2(r)$ 
23                  fin pour
24                   $\hat{V}(k+n, j) = \min(K, \min(J))$ 
25              fin pour
26          fin pour
27           $\tilde{V}(n, i) = \hat{V}(n, 1)$ 
28      fin pour
29  fin pour

```

L'approximation de la fonction valeur à partir des approximations des fonctions valeurs sur l'ensemble de contrôle \mathbb{U} (étape (d)), est présentée dans l'algorithme 2. On utilise les mêmes notations que pour l'algorithme 1, à savoir $\Gamma_n^\Theta(j)$ désigne le j -ième point de la grille de quantification Γ_n^Θ et $M_k^i(j, l)$ désigne la probabilité de transition du j -ième point de la grille Γ_{k-1}^Θ au l -ième point de la grille Γ_k^Θ .

Algorithme 2 Approximation de la fonction valeur $v_0(x_0)$

Entrées :

\tilde{V} matrice contenant les approximations des fonctions valeurs sur \mathbb{U} : pour $1 \leq n \leq N_0$ et $1 \leq i \leq u$, $\tilde{V}(n, i) = \tilde{v}_n(y^i)$

$(\Gamma_n^\Theta)_{0 \leq n \leq N_0}$ grilles de quantification

$(M_n)_{0 \leq n \leq N_0-1}$ matrices de transition associées aux grilles

Sortie :

$\hat{v}_0(x_0)$ approximation de la fonction valeur

```

1   $N_n^\Theta$  le nombre de points de la grille de quantification  $\Gamma_n^\Theta$ 
2   $\hat{V}$  matrice de zeros de taille  $(N_0 + 1) \times N_{N_0}^\Theta$ 
3  pour  $i = 1 : N_{N_0}^\Theta$  faire
4       $\hat{V}(N_0 + 1, i) = g(\Gamma_{N_0}^Z(i))$ 
5  fin pour
6  pour  $k = (N_0 - 1) : -1 : 0$  faire
7      pour  $j = 1 : N_k^\Theta$  faire
8           $z = \Gamma_k^Z(j)$ 
9           $K = F(z, t^*(z)) + \sum_{l=1}^{N_{k+1}^\Theta} e^{-\alpha \Gamma_{k+1}^S(l)} \hat{V}(k + 2, l) M_k(j, l)$ 
10          $m, J_1, J_2$  et  $J$  vecteurs lignes de taille  $n(z) + 1$ , initialisés à 0
11         pour  $r = 1 : (n(z) + 1)$  faire
12              $t = (r - 1) \Delta(z)$ 
13              $m(r) = \min_{1 \leq l \leq u} \{c(\Phi(z, t), y^l) + \tilde{V}(k + 1, l)\}$ 
14              $J_1(r) = \sum_{l=1}^{N_{k+1}^\Theta} e^{-\alpha \Gamma_{k+1}^S(l)} \hat{V}(k + 2, l) M_k(j, l) \mathbb{1}_{\{\Gamma_{k+1}^S(l) < t\}}$ 
15              $J_2(r) = \sum_{l=1}^{N_{k+1}^\Theta} e^{-\alpha t} m(t) M_k(j, l) \mathbb{1}_{\{\Gamma_{k+1}^S(l) \geq t\}}$ 
16              $J(r) = F(z, t) + J_1(r) + J_2(r)$ 
17         fin pour
18          $\hat{V}(k + 1, j) = \min(K, \min(J))$ 
19     fin pour
20 fin pour
21  $\hat{v}_0(x_0) = \hat{V}(1, 1)$ 

```

L'approximation de la fonction valeur (algorithmes 1 et 2) utilise les fonctions :

- $\Phi, t^*(z)$: fonctions liées à la définition du PDMP $\{X_t\}$,
- F, c : fonctions liées aux coûts d'indisponibilité et aux coûts des interventions,
- g, Δ, n : fonctions paramètres de l'approximation des fonctions valeurs sur \mathbb{U} .

Les trois dernières fonctions influencent directement l'approximation de la fonction valeur. En effet, plus g est choisi proche de la fonction h , meilleur est le résultat. Si h est calculable en tout point de l'espace, alors le mieux est de prendre $g = h$. Sinon, on peut prendre une borne supérieure de la fonction h . De même, le pas de discrétisation Δ (avec sa fonction associée n) a un impact direct sur le taux de convergence, comme le montrent les théorèmes 10.1 et 10.2 de Saporta *et al.* (2015).

1.9 Bilan

Dans ce premier chapitre, nous avons défini les processus de Markov déterministes par morceaux et le problème de contrôle impulsif pour ces processus. Nous avons fixé le cadre de notre étude. Nous avons ensuite détaillé la méthode numérique d'approximation de la fonction valeur de Saporta et Dufour (2012), en exposant notamment le procédé de quantification.

Nous allons maintenant nous intéresser à l'application de cette méthode d'approximation de la fonction valeur, dans le cadre de l'optimisation de la maintenance d'un équipement opto-électronique (partie I). Nous prolongerons ensuite l'étude théorique des problèmes de contrôle impulsif, en construisant une famille de stratégies ε -optimales (partie II).

Première partie

**Optimisation de la maintenance d'un
équipement opto-électronique**

Chapitre 2

Systeme opto-électronique et modélisation

Sommaire

| | | |
|------------|--|-----------|
| 2.1 | Caméra Sophie | 40 |
| 2.1.1 | Descriptif de la caméra | 40 |
| 2.1.2 | États de fonctionnement et de panne de la caméra | 42 |
| 2.1.3 | Maintenances et coûts | 43 |
| 2.1.4 | Problématique industrielle | 44 |
| 2.2 | Régimes de la caméra Sophie | 47 |
| 2.2.1 | Transition entre les états des éléments de la caméra | 47 |
| 2.2.2 | Définition des régimes de la caméra | 49 |
| 2.3 | Processus de Markov déterministe par morceaux | 52 |
| 2.4 | Problème de contrôle impulsionnel | 55 |
| 2.4.1 | Ensemble de contrôle | 56 |
| 2.4.2 | Coûts associés | 58 |
| 2.4.3 | Facteur d'oubli | 62 |
| 2.5 | Conclusion | 63 |

L'opto-électronique est le domaine d'étude et de la conception de systèmes qui allient l'électronique et l'optique. Plus précisément, les dispositifs opto-électroniques sont des systèmes électroniques détectant et contrôlant éventuellement une source de lumière. Cette dernière comprend la lumière visible mais également les formes invisibles de lumière telles que les radiations gamma, les rayons X, les ultraviolets et les infrarouges. Les applications de l'opto-électronique sont multiples mais on peut en citer deux très importantes : les communications par fibre optique et les optocoupleurs qui sont des composants électroniques capables de transmettre des signaux électriques entre deux circuits isolés en utilisant la lumière.

Notre étude porte sur une caméra thermique développée par Thales, utilisant un capteur d'émission infrarouge. La caméra Sophie (figure 2.1) est une paire de jumelle portative permettant de visualiser une scène dans toutes les conditions possibles, y compris de nuit ou dans des conditions climatiques défavorables (brouillard, vent, face au soleil). On s'intéresse dans ce chapitre au fonctionnement et à la gestion de la maintenance de la caméra Sophie, ainsi qu'à la modélisation du problème d'optimisation de la maintenance de la caméra par un problème de contrôle impulsionnel.

Dans toute la partie I, les temps sont exprimés en jours (sauf mention contraire) et les coûts sont exprimés dans une unité arbitraire. Les valeurs choisies dans les applications ne

FIGURE 2.1 – *Caméra Sophie.*

correspondent pas aux vraies valeurs mais sont des valeurs arbitraires prises à titre d'exemple dans le seul but de présenter et commenter les résultats des méthodes mises en œuvre dans ce document.

Dans le paragraphe 2.1, une description de la structure de la caméra Sophie, de son fonctionnement et de ses pannes est exposée. On détaille également le contexte industriel dans lequel se place notre étude. Au paragraphe 2.2, on liste les états des éléments de la caméra et les transitions entre les états d'un même élément, puis on définit les régimes de fonctionnement de la caméra. On modélise ensuite dans la partie 2.3 l'évolution de la caméra par un processus de Markov déterministe par morceaux (PDMP). Le problème de contrôle impulsif associé au problème d'optimisation de la maintenance de la caméra Sophie est présenté au paragraphe 2.4. On y définit notamment les coûts d'indisponibilité et les coûts d'interventions, et on discute de l'ensemble de contrôle associé au problème de contrôle impulsif. Enfin, la partie 2.5 résume ce qui a été fait dans ce chapitre.

2.1 Caméra Sophie

Le fonctionnement de la caméra thermique Sophie consiste à capter l'émission infrarouge de la scène observée à l'aide d'un capteur sous vide, refroidi jusqu'à une température très basse par une machine à froid. Le capteur transmet alors une image visible sur un écran. Un système de mise au point est intégré dans la caméra et est situé avant le capteur.

2.1.1 Descriptif de la caméra

La caméra comporte de multiple composants. Pour notre étude, on considérera que la caméra est constituée de 3 blocs indépendants les uns des autres dans leur fonctionnement. Le premier bloc regroupe toute l'électronique de la caméra. Il est appelé bloc électronique et est noté (E). Le deuxième bloc contient le système de mise au point de la caméra, partie appelée biafocal et notée (BF). Il est composé d'une lentille optique fixée sur un charriot qui se déplace sur une vis sans fin. Le troisième bloc correspond à l'ensemble "détecteur d'infrarouge et machine à froid". Afin de faciliter les appellations, on nommera ce bloc le bloc refroidissement. On distingue deux parties dans celui-ci. La première est un roulement à billes (RB), situé physiquement dans la machine à froid. Le reste du bloc refroidissement, c'est-à-dire le détecteur et la machine à froid sans le roulement à bille, constituera la seconde partie, appelée machine à froid pour simplifier les appellations et notée (MAF). Cette dernière contient notamment de l'hélium pour permettre à la caméra d'abaisser sa température au démarrage. La caméra est ainsi constituée de trois blocs (figure 2.2) mais on distingue quatre éléments qui sont l'électronique, le biafocal, le roulement à billes et la machine à froid (figure 2.3).

La caméra peut contenir également des HUMS (Health and Usage Monitoring System) pour le biafocal et la machine à froid. Le HUMS permet d'enregistrer des données permettant d'évaluer

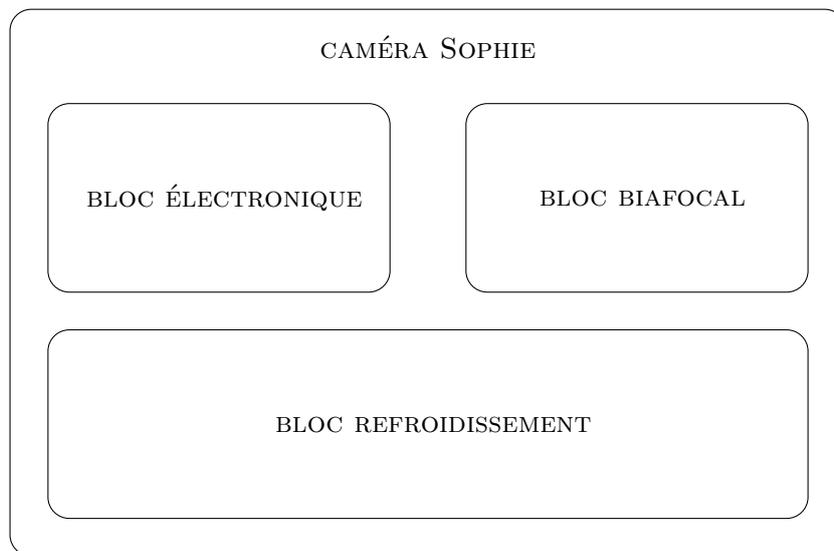


FIGURE 2.2 – Schéma de la caméra Sophie où apparaissent les trois blocs considérés dans notre étude.

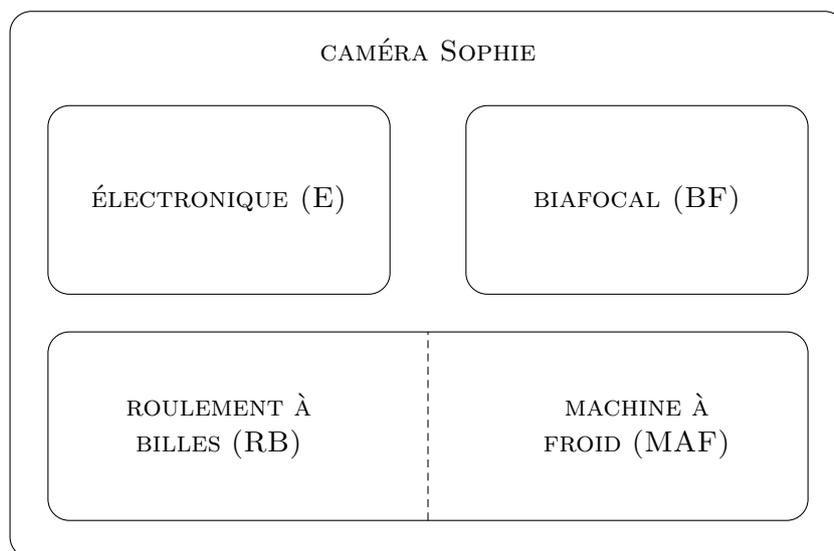


FIGURE 2.3 – Schéma de la caméra Sophie où apparaissent les quatre éléments considérés dans notre étude.

| notation | élément de la caméra |
|----------|--|
| E | électronique |
| BF | biafocal |
| RB | roulement à billes |
| MAF | machine à froid (<i>en réalité : détecteur + machine à froid sans le roulement à billes</i>) |

TAB. 2.1 – Notation des éléments de la caméra

l'état de santé des éléments. Pour la machine à froid, le HUMS enregistre au démarrage de la caméra, son temps de refroidissement jusqu'à la température de fonctionnement du capteur. En fonctionnement normal, ce temps reste stable. En revanche, dès qu'une augmentation du temps de mise à froid est détectée (voir Baysse (2013)), on peut considérer que la machine à froid présente une faiblesse, comme par exemple une fuite d'hélium. Il peut alors être intéressant d'intervenir sur la caméra pour réparer la machine à froid, avant que celle-ci ne tombe en panne.

Pour le biafocal, lors du démarrage de la caméra, la lentille fait un aller-retour complet sur la vis qui la supporte. Le temps de cette action peut être enregistré via un HUMS. Tout comme pour la machine à froid, si une augmentation de ce temps est détectée, le biafocal ne se trouve plus dans des conditions normales de fonctionnement. La panne est donc proche mais une maintenance rapide sur le biafocal peut empêcher la caméra de se retrouver en panne. La maintenance consiste alors à un réajustement du charriot portant la lentille par rapport à la vis sans fin.

2.1.2 États de fonctionnement et de panne de la caméra

On caractérise l'état d'un élément de la caméra de trois manières différentes. On dira qu'un élément est

- dans l'état *stable* s'il fonctionne normalement,
- dans l'état *dégradé* s'il fonctionne mais présente une faiblesse,
- dans l'état *de panne* s'il ne fonctionne pas.

Cependant, caractériser une faiblesse de fonctionnement d'un élément de la caméra n'est pas possible dans tous les cas. Pour la caméra Sophie, seuls la machine à froid et le biafocal peuvent être déclarés dans un état dégradé, grâce notamment aux données récupérées par les HUMS. Ainsi, les états possibles pour les éléments de la caméra sont les suivants.

- L'électronique peut se trouver dans l'état stable ou dans l'état de panne.
- Le biafocal peut être dans l'état stable, dégradé ou de panne.
- Le roulement à billes peut être dans l'état stable ou de panne.
- La machine à froid peut se trouver dans l'état stable, dégradé ou de panne.

Les états de chaque élément de la caméra déterminent l'état général de la caméra, à savoir :

- La caméra est dans l'état *stable* si tous les éléments qui la composent sont dans l'état stable.

- La caméra est dans l'état *dégradé* si l'un au moins de ses éléments est dans l'état dégradé et si les éléments qui ne sont pas dans l'état dégradé sont dans l'état stable. Dans cette situation, la caméra est dans un état de fonctionnement mais présente une faiblesse.
- La caméra est dans l'état *de panne* si l'un au moins de ses éléments est en panne. La caméra ne fonctionne alors pas.

Dans le dernier cas, c'est-à-dire quand la caméra est dans l'état de panne, en réalité seul un élément de la caméra est en panne car deux pannes ne peuvent arriver simultanément.

L'état neuf de la caméra est inclus dans l'état stable. Une caméra partant de l'état stable peut aller soit dans l'état dégradé, si le biafocal ou la machine à froid passe dans l'état dégradé, soit directement dans l'état de panne si l'électronique ou le roulement à billes passe dans l'état de panne.

2.1.3 Maintenances et coûts

Deux types d'opérations de maintenances sont possibles sur la caméra Sophie : la réparation ou le changement. La réparation d'un élément de la caméra ne peut être réalisée que sur la machine à froid et sur le biafocal. Une réparation correspond à une opération réalisée sur l'élément (sans le changer) afin qu'il retrouve un état de fonctionnement normal. Le changement d'un élément est possible pour tous les éléments de la caméra. Il consiste à changer la totalité du bloc qui contient l'élément en question et à le remplacer par un bloc neuf. Au total, cinq opérations de maintenances (individuelles) sont possibles sur la caméra, à savoir :

- la réparation de la machine à froid, qui consiste à regonfler la machine à froid d'hélium,
- la réparation du biafocal, qui consiste à un réajustement du charriot supportant la lentille sur la vis sans fin,
- le changement du bloc électronique, où l'on remplace à neuf toute l'électronique de la caméra,
- le changement du bloc biafocal, pour lequel on remplace à neuf le biafocal,
- le changement du bloc refroidissement, qui consiste à remplacer à neuf la machine à froid et le roulement à billes.

Il est à noter que l'on ne peut pas réparer l'électronique et le roulement à billes. On considère qu'après une réparation, l'élément est remis à neuf. Par ailleurs, on ne peut pas changer le roulement à billes sans changer la machine à froid et vice versa.

La réalisation d'une opération de maintenance sur la caméra ne se fait que quand la caméra est en panne et elle nécessite un retour à l'atelier. Cela implique une immobilisation de la caméra d'une durée fixe, notée d_{atelier} , qui inclut le temps de transport. On fixe arbitrairement la valeur de d_{atelier} à 35 jours. Il peut donc être judicieux d'effectuer plusieurs maintenances sur la caméra lorsque celle-ci est présente à l'atelier. Étant donné qu'on effectue une seule opération de maintenance par élément, c'est-à-dire que l'on s'interdit de réparer et changer un même élément, les combinaisons possibles de maintenances sur la caméra sont les suivantes :

1. changement bloc refroidissement
2. changement bloc électronique
3. changement bloc biafocal
4. changement bloc biafocal + réparation machine à froid
5. changement bloc électronique + réparation machine à froid

| maintenance | coût |
|--|-------------|
| réparation machine à froid | 1000 |
| réparation biafocal | 1000 |
| changement bloc refroidissement (MAF+BF) | 6000 |
| changement biafocal | 6000 |
| changement électronique | 3000 |

TAB. 2.2 – *Coût des maintenances individuelles (ne comprend pas les coûts fixes).*

6. changement bloc refroidissement + réparation biafocal
7. changement bloc électronique + réparation biafocal
8. changement bloc électronique + réparation machine à froid + réparation biafocal
9. changement bloc refroidissement + changement bloc biafocal
10. changement bloc électronique + changement bloc refroidissement
11. changement bloc électronique + changement bloc biafocal
12. changement bloc électronique + changement bloc biafocal + réparation machine à froid
13. changement bloc électronique + changement bloc refroidissement + réparation biafocal
14. changement de tous les blocs (électronique, refroidissement, biafocal)

soit au total 14 combinaisons possibles de maintenances sur la caméra. Par la suite, une telle combinaison est appelée une intervention. À la sortie de l'atelier, on impose que tous les éléments de la caméra soient dans l'état stable. On ne peut notamment pas avoir un élément dans l'état dégradé sur la caméra à la sortie de l'atelier.

Le coût d'une intervention correspond à l'addition de deux coûts. Le premier coût est un coût fixe, d'un montant de 2000, représentant le coût du transport aller-retour de la caméra vers l'atelier et le coût de la logistique liée à l'intervention sur la caméra. Le coût fixe est identique pour toutes les interventions. En revanche, le second coût est variable car correspond aux coûts des pièces et de la main-d'œuvre nécessaires à la réalisation de l'intervention. Il est égal à l'addition des coûts des maintenances individuelles effectuées durant l'intervention. Les coûts des maintenances individuelles sont donnés dans le tableau 2.2. Le coût total de chaque intervention possible sur la caméra est donné dans le tableau 2.3. Il est par exemple de 9000 pour l'intervention numéro 4. En effet, en plus des 2000 de coûts fixes, on additionne le coût d'un changement du bloc biafocal qui est de 6000, au coût d'une réparation de la machine à froid, d'un montant de 1000, soit un total de 9000.

2.1.4 Problématique industrielle

La maintenance des équipements regroupe les actions de dépannage, de réparation, de réglage, de révision, de contrôle et de vérification des équipements. Aujourd'hui, la maintenance des équipements est devenue un véritable enjeu pour les industriels. Elle influe en effet sur les coûts d'entretien des systèmes et sur leur disponibilité auprès des clients. Il est donc nécessaire d'optimiser la maintenance des systèmes. Pour cela, différentes mesures et techniques peuvent

| numéro | intervention | coût |
|--------|--|-------|
| 1 | changement bloc refroidissement | 8000 |
| 2 | changement bloc bifaocal | 8000 |
| 3 | changement bloc électronique | 5000 |
| 4 | changement bloc bifaocal + réparation machine à froid | 9000 |
| 5 | changement bloc électronique + réparation machine à froid | 6000 |
| 6 | changement bloc refroidissement + réparation bifaocal | 9000 |
| 7 | changement bloc électronique + réparation bifaocal | 6000 |
| 8 | changement bloc électronique + réparation machine à froid et bifaocal | 7000 |
| 9 | changement bloc refroidissement et bloc bifaocal | 14000 |
| 10 | changement bloc électronique et bloc refroidissement | 11000 |
| 11 | changement bloc électronique et bloc bifaocal | 11000 |
| 12 | changement bloc électronique et bloc bifaocal + réparation machine à froid | 12000 |
| 13 | changement bloc électronique et bloc refroidissement + réparation bifaocal | 12000 |
| 14 | changement de tous les blocs (électronique, refroidissement, bifaocal) | 17000 |

TAB. 2.3 – Coût total des interventions possibles sur la caméra. Les coûts affichés prennent en compte les coûts fixes.

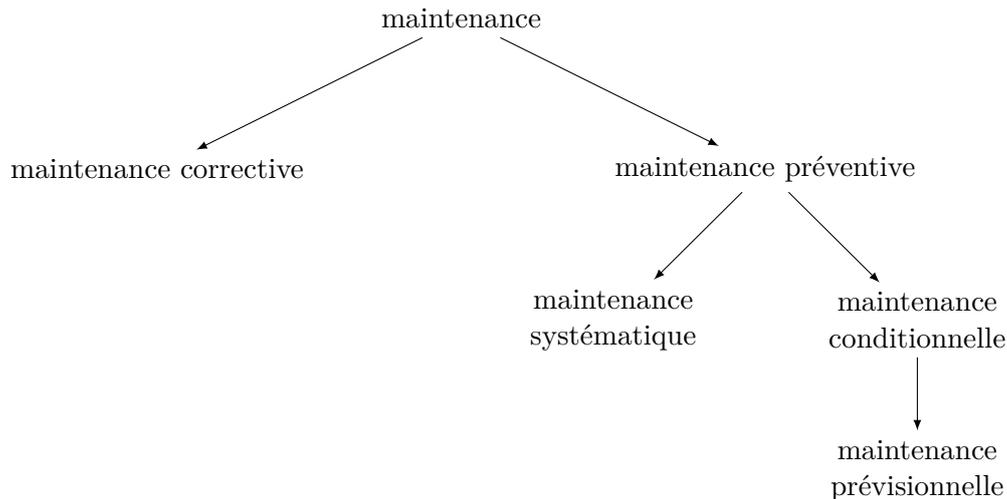


FIGURE 2.4 – Graphique montrant les liens entre les différentes catégories de maintenances.

être mises en place. Celles-ci déterminent le type de maintenance correspondante. On en définit quelques unes ci-dessous, en se basant sur la norme NF EN 13 306 X 60-319 (p.23-24).

On distingue deux grandes catégories de maintenance. La première et la plus simple est la maintenance dite corrective. Il s'agit d'une maintenance exécutée après une panne et destinée à remettre le système en état de fonctionnement.

Cependant, dans certains cas, anticiper les pannes et effectuer une maintenance avant que la panne ne survienne peut être plus intéressant pour deux raisons. D'une part, la réparation d'un élément peut être moins coûteuse et plus facile à mettre en œuvre qu'un changement à neuf de l'élément. D'autre part, la panne d'un élément peut avoir des conséquences sur l'ensemble du système, notamment en terme de sécurité et d'indisponibilité. Un exemple classique est la présence d'une courroie dans un moteur. En cassant, la courroie peut casser d'autres pièces du moteur, ce qui est peut être très coûteux, au regard du coût de changement de la courroie avant la panne. Ces types de maintenances constituent la seconde catégorie de maintenances, dites préventives. Celles-ci regroupent les maintenances exécutées à des intervalles prédéterminés ou selon des critères donnés, qui ont pour but de réduire la probabilité de dégradation ou de défaillance de l'équipement.

Parmi les maintenances préventives, on distingue plusieurs sous-catégories. La figure 2.4 expose les liens entre les catégories de maintenances définies dans cette partie. Font partie des maintenances préventives :

- les maintenances systématiques :
Ce sont les maintenances exécutées "à des intervalles de temps préétablis ou selon un nombre défini d'unités d'usage mais sans contrôle préalable de l'état du bien". Le remplacement de la courroie d'un moteur de voiture en est typiquement un exemple ; celle-ci doit être changée si la durée d'utilisation fixée par le constructeur est atteinte, ou si la voiture a parcouru un nombre de kilomètres fixé également par le constructeur.
- les maintenances conditionnelles :
Ces maintenances sont basées "sur une surveillance du fonctionnement du bien et/ou des paramètres significatifs de ce fonctionnement intégrant les actions qui en découlent". Cette surveillance peut se faire ponctuellement de manière planifiée ou non, ou bien la surveillance peut-être continue. On s'intéressera plus particulièrement à une sous-catégorie de maintenances conditionnelles :
 - les maintenances prévisionnelles (*predictive maintenance* en anglais) :

Ces maintenances sont exécutées "en suivant les prévisions extrapolées de l'analyse et de l'évaluation de paramètres significatifs de la dégradation du bien". Autrement dit, l'objectif est de détecter un état de dégradation de l'équipement (avant sa panne), à partir des données permettant d'évaluer l'état de santé de l'équipement.

Nous nous intéresserons naturellement à la dernière catégorie de maintenance, celle des maintenances prévisionnelles. En effet, la caméra Sophie possède des HUMS, nous permettant de connaître l'état de santé de la machine à froid et éventuellement du biafocal. Il est donc possible de construire des outils de maintenance exploitant les données recueillies par les HUMS.

Un premier travail sur la maintenance prévisionnelle de la caméra Sophie a été réalisé dans Baysse (2013), avec une modélisation de la caméra ne tenant pas compte du biafocal. Une méthode pour détecter l'état dégradé d'un élément de la caméra à partir des données des HUMS y est développée, ainsi qu'une méthode permettant de déterminer le meilleur instant pour la réalisation d'une maintenance en fonction notamment de l'état estimé de la caméra.

Cependant, lorsque la caméra est arrêtée à la date optimale de maintenance ou lorsqu'elle est en panne, il est important de décider des réparations ou changements à effectuer. Cette décision a un impact direct sur le coût de la maintenance de la caméra, mais aussi sur la disponibilité auprès du client. En effet, si un élément est dans l'état dégradé lorsque la caméra est à l'atelier, il est probablement préférable de réparer l'élément au même moment, afin d'éviter un retour prématuré à l'atelier, qui engendrait des coûts et une indisponibilité supplémentaires pour le client.

On souhaite donc déterminer une politique de maintenance (appelée également politique d'intervention) sur la caméra qui minimise les coûts. En pratique, on souhaite, à n'importe quel instant de vie de la caméra, connaître une date optimale de maintenance de la caméra, mais on souhaite également déterminer ce que l'on fait sur la caméra lors d'une maintenance, c'est-à-dire déterminer la ou les pièces à remplacer ou à réparer.

On note *détude* la durée de l'étude de l'optimisation de maintenance réalisée sur la caméra. Cette durée peut être déterminée selon plusieurs critères comme par exemple la durée d'une garantie, le dépassement technologique de la caméra ou encore l'arrêt de la production d'une pièce qui compose la caméra. Dans notre étude, on fixe arbitrairement cette valeur à 10 ans.

2.2 Régimes de la caméra Sophie

Afin de modéliser l'évolution d'une caméra Sophie par un processus aléatoire, il est nécessaire de modéliser les passages d'un état à un autre pour chaque éléments de la caméra, ce qui est fait dans la partie 2.2.1. Il est également nécessaire d'établir les régimes dans lesquels la caméra peut se trouver. Cela est fait dans la partie 2.2.2.

2.2.1 Transition entre les états des éléments de la caméra

Au paragraphe 2.1.2, on a vu que les éléments de la caméra pouvaient se trouver dans l'état stable, dégradé ou panne selon les cas. On résume ci-dessous les états possibles pour chaque élément de la caméra :

- L'électronique (E) peut se trouver dans l'état stable ou l'état de panne.
- Le roulement à billes (RB) peut se trouver dans l'état stable ou l'état de panne.
- La machine à froid (MAF) peut se trouver dans l'état stable, dégradé ou de panne.
On peut détecter l'état dégradé via le temps de mise à froid au démarrage de la caméra.
- Le biafocal (BF) peut se trouver dans l'état stable, dégradé ou de panne.
On peut détecter l'état dégradé du biafocal via le temps mis par la lentille pour effectuer un aller-retour complet sur la vis sans fin au démarrage de la caméra.

Afin de modéliser le passage d'un état à un autre pour les éléments de la caméra, on utilise des lois exponentielles ou des lois de Weibull. Les lois de Weibull modélisent les systèmes qui s'usent au court du temps. Elles caractérisent donc les changements d'état dus à une usure, tel que pour le roulement à billes, où les frottements entraînent une usure sur les billes du roulement et les bagues intérieures et extérieures qui les retiennent. Un roulement à billes trop usé finit par céder et peut entraîner des dégâts importants sur les pièces qui l'entourent. Les lois exponentielles sont des cas particuliers des lois de Weibull. Elles modélisent les systèmes sans usure ou avec une usure très faible et difficilement mesurable, tel que l'électronique. Les caractéristiques principales des lois de Weibull sont données ci-dessous mais on pourra se référer au livre *Coccozza-Thivent (1997)* pour plus de détails.

Proposition 2.2.1

Soient $\alpha, \beta > 0$. Une densité de la loi de Weibull $\mathcal{W}(\alpha, \beta)$ est donnée par la fonction

$$f(t) = \frac{\beta}{\alpha} \left(\frac{t}{\alpha}\right)^{\beta-1} \exp \left\{ - \left(\frac{t}{\alpha}\right)^\beta \right\} \mathbf{1}_{\{t \in [0; +\infty[\}}.$$

Soit T une variable aléatoire de loi $\mathcal{W}(\alpha, \beta)$. Alors

$$\mathbb{P}(T > t) = \exp \left\{ - \left(\frac{t}{\alpha}\right)^\beta \right\}, \quad \mathbb{E}(T) = \alpha \Gamma\left(1 + \frac{1}{\beta}\right) \quad \text{et} \quad \mathbb{V}(T) = \alpha^2 \left[\Gamma\left(1 + \frac{2}{\beta}\right) - \Gamma^2\left(1 + \frac{1}{\beta}\right) \right].$$

Son taux de saut μ est défini, pour tout $t \in \mathbb{R}^+$, par

$$\mu(t) = \frac{\beta}{\alpha} \left(\frac{t}{\alpha}\right)^{\beta-1}.$$

La loi de Weibull a deux paramètres (α, β) qui sont tous les deux strictement positifs. Le paramètre α est le paramètre d'échelle et β est le paramètre de forme. Le paramètre d'échelle α indique quand la panne a le plus de probabilité de se produire. Le paramètre de forme β dépend de l'évolution du taux de panne : si ce dernier augmente avec le temps, comme dans le cas d'une usure du système, $\beta > 1$. En revanche, si le taux de panne est constant, on choisit $\beta = 1$. La loi de Weibull de paramètres $(\alpha, 1)$ correspond alors à une loi exponentielle de paramètre $1/\alpha$. Le dernier cas possible correspondant à un taux de panne qui diminue au cours du temps ($\beta < 1$), n'apparaît pas dans notre problème car il correspond à des défaillances qui surviennent principalement lors de la mise en service de l'élément considéré. Ce type de panne ne se produit pas sur la caméra car celle-ci est testée à la sortie de l'atelier et les éléments défectueux sont détectés et changés avant la mise en service de la caméra.

Remarque 1

La loi exponentielle $\mathcal{E}(\lambda)$, avec $\lambda > 0$, étant un cas particulier des lois de Weibull, il suffit de reprendre la proposition 2.2.1 avec $\alpha = \frac{1}{\lambda}$ et $\beta = 1$ pour en obtenir ses caractéristiques.

Les choix des lois pour les changements d'état des éléments de la caméra, à savoir loi de Weibull ou loi exponentielle, sont basés sur les caractéristiques physiques des éléments (usure mesurable ou non). Les paramètres des lois sont déterminés à partir de relevés statistiques sur les anciennes pannes des éléments. Ils peuvent être donnés par le constructeur de l'élément ou déterminés à l'aide du retour d'expérience (REX) sur la caméra par l'entreprise Thales. Les lois de transitions entre les états des éléments de la caméra sont détaillées ci-dessous.

- Le bloc électronique de la caméra peut se trouver dans l'état de fonctionnement stable ou dans l'état de panne. Comme l'usure de l'électronique est difficilement mesurable à l'heure

actuelle, on modélise le passage de l'état stable à l'état de panne par une loi exponentielle. Le paramètre de cette loi est noté λ_E .

- Le bloc biafocal a trois états possibles : stable, dégradé et panne. L'usure du système est mesurable car on peut quantifier la détérioration sur le pas de vis due au passage du charriot supportant la lentille. On modélise donc le passage de l'état stable du biafocal à son état dégradé par une loi de Weibull, de paramètres $(\alpha_{Bf}, \beta_{Bf})$. On note μ_{Bf} son taux de saut associé.

On considère qu'une fois dans l'état dégradé, l'usure est telle que son évolution n'a plus d'influence sur la probabilité de tomber en panne. Le passage de l'état dégradé à l'état de panne se fait donc au moyen d'une loi exponentielle de paramètre λ_{Bf} .

- Le bloc refroidissement est composé de deux éléments : le roulement à billes et la machine à froid.

- Pour le roulement à billes, la transition entre l'état stable et l'état dégradé se fait par une loi de Weibull de paramètres $(\alpha_{Rb}, \beta_{Rb})$ car le roulement à billes présente une usure au cours du temps. Son taux de panne associé est noté μ_{Rb} .

- La machine à froid a deux états de fonctionnement (stable et dégradé) et un état de panne. Comme le biafocal et le roulement à billes, on peut évaluer l'usure au cours du temps du système. Par conséquent, la loi de passage de l'état de fonctionnement stable à l'état de fonctionnement dégradé est une loi de Weibull, de paramètres $(\alpha_{Maf}, \beta_{Maf})$. Son taux de panne associé est noté μ_{Maf} .

Une fois l'état dégradé atteint, l'usure supplémentaire de la machine à froid n'est plus significative par rapport à l'évolution de la probabilité que la machine à froid tombe en panne. On utilise donc une loi exponentielle, de paramètre λ_{Maf} .

Il est à noter que l'usure de la machine à froid n'influe pas sur le roulement à billes et vice-versa.

Un récapitulatif des états possibles de chaque élément de la caméra et des lois de passage entre les états des éléments est donné par la figure 2.5.

2.2.2 Définition des régimes de la caméra

On appelle régime de la caméra l'ensemble des états des éléments de la caméra. Afin de simplifier les notations, on les numérote de 1 à 27. Le tableau 2.4 regroupe tous les régimes possibles pour la caméra. Parmi eux, on distingue cinq ensembles :

- \mathcal{S} : ensemble du régime stable
Ce premier ensemble de régimes n'est constitué que d'un seul régime numéroté 1. Lorsque l'on se trouve dans ce régime, tous les éléments de la caméra sont dans l'état stable et la caméra fonctionne normalement.
- \mathcal{D} : ensemble des régimes dégradés
Le second ensemble de régimes concerne les états de la caméra pour lesquels la machine à froid et/ou le biafocal sont/est dans l'état dégradé mais aucun élément de la caméra n'est dans l'état de panne. La caméra fonctionne mais au moins un de ces éléments présente une faiblesse. Il y a trois régimes dégradés, numérotés de 2 à 4.
- \mathcal{P} : ensemble des régimes de panne
Le troisième ensemble de régimes correspond aux états de la caméra pour lesquels un des blocs est en panne. Il est à noter que le roulement à billes faisant partie physiquement de la machine à froid, ces deux éléments sont dépendants : si l'un est en panne, le bloc

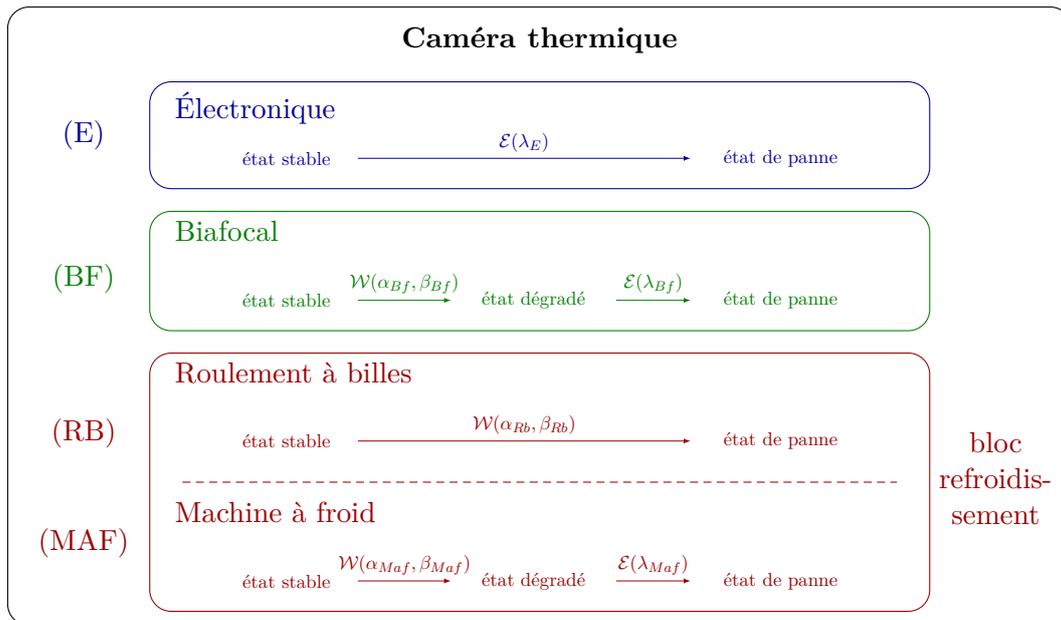


FIGURE 2.5 – Modélisation de la structure de la caméra Sophie et lois des changements d'états des éléments.

refroidissement est en panne.

Par ailleurs, on ne peut pas avoir deux blocs en panne car lorsqu'une panne survient, la caméra ne fonctionne plus. On considère que les éléments qui la composent sont arrêtés et n'évoluent plus. De plus, les lois utilisées étant continues, on ne peut pas mathématiquement avoir deux pannes simultanément.

Il y a au total 8 régimes de panne, numérotés de 5 à 12.

- \mathcal{A} : ensemble des régimes d'atelier

Cet ensemble contient les régimes pour lesquels la caméra est dans l'atelier de maintenance. Lorsque la caméra est à l'atelier, elle ne fonctionne pas et ses éléments n'évoluent pas. En fonction du régime atelier dans lequel la caméra se trouve, on effectue une ou plusieurs opérations de maintenances parmi les suivantes : changement de l'électronique, réparation ou changement du bloc biafocal, réparation de la machine à froid ou changement du bloc refroidissement, c'est-à-dire de la machine à froid et du roulement à billes. À chaque combinaison de ces opérations (listées au paragraphe 2.1.3) correspond un régime d'atelier, soit au total 14 régimes d'atelier, numérotés de 13 à 26.

- \mathcal{C} : ensemble du régime cimetièrre

Le dernier ensemble de régimes de la caméra, contient un seul régime, numéroté 27. Il survient lorsque la caméra a atteint sa durée d'étude $d_{étude}$. Ce régime est appelé le cimetièrre : on considère que la caméra est définitivement arrêtée à cette date quel que soit le régime dans lequel elle se trouve juste avant.

Les passages d'un régime à un autre de la caméra dépendent des contraintes physiques de la caméra et de son fonctionnement. Si on considère une caméra qui part de l'état neuf, c'est-à-dire du régime 1 stable, sur laquelle aucune maintenance n'est effectuée, alors les différentes possibilités d'évolution de la caméra entre les régimes sont données par la figure 2.6. Il est à noter qu'en fonction de l'état initial de la caméra, c'est-à-dire son régime initial, et des choix de maintenances faits, les régimes de la caméra ne sont pas tous accessibles. Dans la figure 2.6, les régimes d'atelier ne peuvent pas être visités car la caméra part du régime stable et aucune maintenance n'est faite durant sa vie. On peut résumer les évolutions possibles d'une caméra

| régime | E | RB | MAF | BF | état de la caméra |
|--------|------------|------------------------|------------|------------|---------------------------------|
| 1 | - | - | - | - | \mathcal{S} : stable |
| 2 | - | - | dégradé | - | \mathcal{D} : dégradé |
| 3 | - | - | - | dégradé | |
| 4 | - | - | dégradé | dégradé | |
| 5 | - | panne du bloc (MAF+RB) | | - | \mathcal{P} : caméra en panne |
| 6 | - | - | - | panne | |
| 7 | panne | - | - | - | |
| 8 | - | - | dégradé | panne | |
| 9 | panne | - | dégradé | - | |
| 10 | - | panne du bloc (MAF+RB) | | dégradé | |
| 11 | panne | - | - | dégradé | |
| 12 | panne | - | dégradé | dégradé | |
| 13 | - | changement | changement | - | \mathcal{A} : atelier |
| 14 | - | - | - | changement | |
| 15 | changement | - | - | - | |
| 16 | - | - | réparation | changement | |
| 17 | changement | - | réparation | - | |
| 18 | - | changement | changement | réparation | |
| 19 | changement | - | - | réparation | |
| 20 | changement | - | réparation | réparation | |
| 21 | - | changement | changement | changement | |
| 22 | changement | changement | changement | - | |
| 23 | changement | - | - | changement | |
| 24 | changement | - | réparation | changement | |
| 25 | changement | changement | changement | réparation | |
| 26 | changement | changement | changement | changement | |
| 27 | arrêté | arrêté | arrêté | arrêté | \mathcal{C} : cimetièrre |

TAB. 2.4 – Numérotation et description des régimes de la caméra. Le symbole – signifie que l'élément est dans un état stable. Les éléments E, RB, MAF et BF correspondent respectivement à l'électronique, au roulement à billes, à la machine à froid et au biafocal.

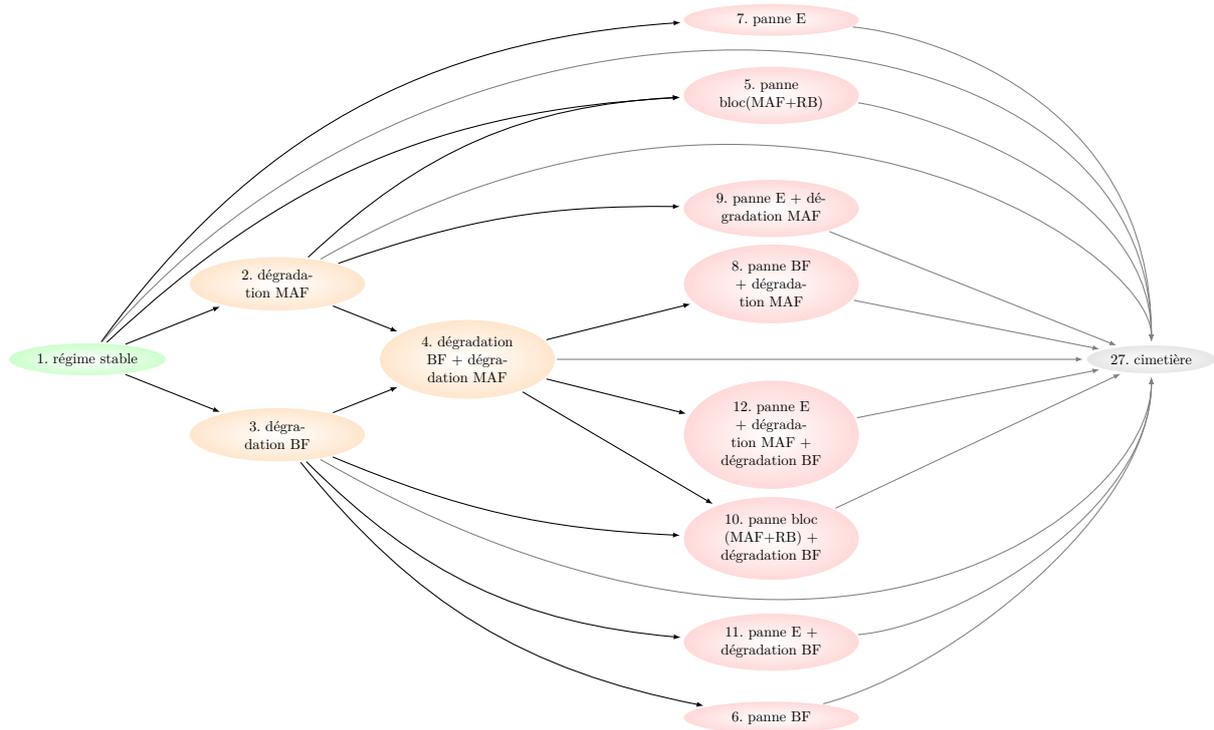


FIGURE 2.6 – Graphique de l'évolution sans maintenance d'une caméra Sophie partant de l'état neuf (stable) jusqu'à son arrêt (cimetière).

pour laquelle aucune maintenance n'est effectuée, par la figure 2.7, où seuls apparaissent les ensembles de régimes de la caméra. On verra par la suite que d'autres évolutions sont possibles si on autorise des maintenances et qu'en fonction de la politique choisie pour les maintenances, le graphique d'évolution change (exemple figure 2.8) et les régimes visités ne sont pas forcément les mêmes.

2.3 Processus de Markov déterministe par morceaux

On modélise l'évolution d'une caméra par un processus de Markov déterministe par morceaux (PDMP), noté $\{X_t\}$. Les PDMP sont particulièrement bien adaptés pour décrire l'évolution des caméras. En effet, dans chaque régime, l'évolution des caméras est connue, ce qui forme la partie déterministe des PDMP. Quant aux changements de régimes de la caméra, comme ceux-ci dépendent de lois de probabilité, ils constituent la partie aléatoire des trajectoires. Ces changements de régimes sont à l'origine de sauts aléatoires dans le processus.

On note \mathcal{M} l'ensemble des régimes de la caméra : $\mathcal{M} = \mathcal{S} \cup \mathcal{D} \cup \mathcal{P} \cup \mathcal{A} \cup \mathcal{C} = \{1, 2, \dots, 27\}$ (voir tableau 2.4). L'état du processus au temps t est noté $X_t = (m_t, \zeta_t)$, où m_t est le régime de la caméra et ζ_t est un vecteur colonne de \mathbb{R}_+^5 qui représente, dans l'ordre,

- le temps de fonctionnement du roulement à billes,
- le temps de fonctionnement de la machine à froid,
- le temps de fonctionnement du biafocal,

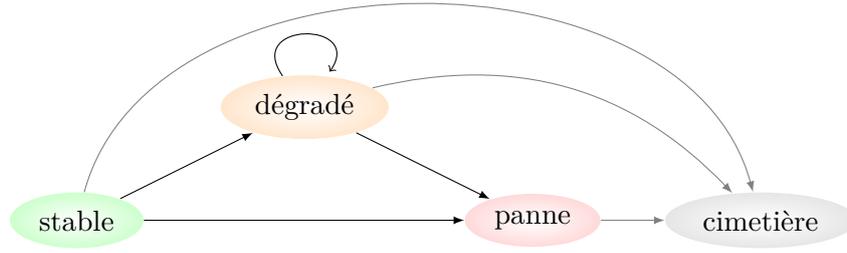


FIGURE 2.7 – Graphique simplifié de l'évolution générale de la caméra Sophie, partant de l'état neuf, quand aucune maintenance n'est effectuée sur celle-ci.

- le temps calendaire de la caméra, c'est-à-dire le temps depuis la mise en service de la caméra,
- le temps écoulé depuis le dernier saut.

Le temps de fonctionnement d'un élément correspond au temps écoulé depuis le dernier changement ou réparation de l'élément en question. On ne considère pas le temps de fonctionnement de l'électronique car ce dernier ne présente pas d'usure. En revanche, le roulement à billes, la machine à froid et le bifaocal s'usent. Il est donc nécessaire de connaître leur temps de fonctionnement pour les lois de Weibull associées, qui sont des lois à mémoire.

La durée d'étude de la caméra est $d_{étude}$. Par conséquent, l'espace d'état du processus $\{X_t\}$ qui regroupe l'ensemble des états possibles de la caméra est le suivant :

$$E = \{(m, \zeta), m \in \mathcal{S} \cup \mathcal{D} \cup \mathcal{P} \cup \mathcal{A}, \zeta \in \mathbb{R}_+^3 \times [0; d_{étude}[\times \mathbb{R}_+]\} \cup \{\Delta\}, \quad (2.1)$$

où $\Delta := (27, \delta) := (27, 0, 0, 0, d_{étude}, 0)$ est l'état cimetière, correspondant à l'arrêt définitif de la caméra. Le choix de δ n'a pas d'importance car la caméra est définitivement arrêtée. Les informations sur les temps (de fonctionnement, calendaire et temps écoulé depuis le dernier saut) ne sont plus utiles, d'où l'introduction d'un unique état cimetière, choisi arbitrairement.

Le flot qui régit l'évolution du processus entre les sauts, dépend du régime dans lequel on se trouve. Soit $(m, \zeta) \in E$ un état de la caméra. On note $\zeta = (\zeta_{rb}, \zeta_{maf}, \zeta_{bf}, \zeta_{cal}, \zeta_{saut})$. Pour $t \in \mathbb{R}^+$, le flot est défini par :

$$\Phi_m(\zeta, t) = \begin{cases} (\zeta_{rb} + t, \zeta_{maf} + t, \zeta_{bf} + t, \zeta_{cal} + t, \zeta_{saut} + t) & \text{si } m \in \mathcal{S} \cup \mathcal{D}, \\ (\zeta_{rb}, \zeta_{maf}, \zeta_{bf}, \zeta_{cal} + t, \zeta_{saut} + t) & \text{si } m \in \mathcal{P} \cup \mathcal{A}, \\ \delta & \text{si } m \in \mathcal{C}. \end{cases} \quad (2.2)$$

Ainsi lorsque la caméra est dans le régime stable ou dans un régime dégradé, comme la caméra fonctionne, tous les temps évoluent. En revanche, lorsque la caméra est dans un régime d'atelier ou un régime de panne, la caméra est considérée arrêtée, y compris durant les phases de maintenance ou de réparation. Les temps de fonctionnement des différents éléments de la caméra sont donc constants. Pour finir, lorsque le cimetière est atteint, l'évolution de la caméra est stoppée et les informations issues des différents temps stockés dans la variable ζ ne sont plus utiles. En conséquence, le flot est constant, égal à δ .

Le temps d'atteinte à la frontière de l'espace d'état E par le flot Φ , noté t^* , correspond au moment où un changement de régime devient obligatoire pour la caméra. Deux cas peuvent se présenter. Le premier, lorsque la caméra atteint la durée $d_{étude}$, elle passe automatiquement dans l'état cimetière. Le deuxième cas possible survient lorsque la caméra est dans l'atelier. On

impose la durée $d_{atelier}$ pour un passage dans l'atelier. Le temps t^* s'exprime donc de la manière suivante :

$$t^*(m, \zeta) = (d_{étude} - \zeta_{cal}) \mathbb{1}_{\{m \in S \cup D \cup P\}} + ((d_{atelier} - \zeta_{saut}) \wedge (d_{étude} - \zeta_{cal})) \mathbb{1}_{\{m \in \mathcal{A}\}} + \infty \mathbb{1}_{\{m \in \mathcal{C}\}}. \quad (2.3)$$

Il est à noter que le temps calendaire et le temps depuis le dernier saut sont nécessaires dans les variables du processus afin de savoir quand la frontière est atteinte.

Le taux de saut qui permet de déterminer les temps de saut, dépend également du régime et des temps de fonctionnements de certains éléments de la caméra. Il est défini comme suit :

$$\lambda(m, \zeta) = \begin{cases} \lambda_E + \mu_{Rb}(\zeta_{rb}) + \mu_{Maf}(\zeta_{maf}) + \mu_{Bf}(\zeta_{bf}) & \text{si } m = 1 \\ \lambda_E + \mu_{Rb}(\zeta_{rb}) + \lambda_{Maf} + \mu_{Bf}(\zeta_{bf}) & \text{si } m = 2, \\ \lambda_E + \mu_{Rb}(\zeta_{rb}) + \mu_{Maf}(\zeta_{maf}) + \lambda_{Bf} & \text{si } m = 3, \\ \lambda_E + \mu_{Rb}(\zeta_{rb}) + \lambda_{Maf} + \lambda_{Bf} & \text{si } m = 4, \\ 0 & \text{si } m \in \mathcal{P} \cup \mathcal{A} \cup \mathcal{C}. \end{cases} \quad (2.4)$$

Si la caméra se trouve dans un régime de panne ou d'atelier, elle ne pourra en sortir que si la frontière de l'espace d'état est atteinte, c'est-à-dire si le temps $d_{étude}$ est écoulé ou si la caméra se trouve à l'atelier depuis la durée $d_{atelier}$.

Soit $(m', \zeta') \in E$, où ζ' s'écrit : $\zeta' = (\zeta'_{rb}, \zeta'_{maf}, \zeta'_{bf}, \zeta'_{cal}, \zeta'_{saut})$. Les positions du processus après un saut sont déterminées par le noyau de transition Q . Si le régime de départ est le régime stable ($m = 1$), alors d'après la figure 2.6, cinq sauts sont possibles à savoir : un saut dans le régime 7 dû à une panne électronique, un saut dans le régime 5 dû à une panne du bloc de refroidissement, un saut dans le régime 2 dû à une dégradation de la machine à froid, un saut dans le régime 3 dû à une dégradation du biafocal, ou un saut dans le régime cimetièrre 27 quand la durée $d_{étude}$ est atteinte. Le noyau de transition s'écrit donc :

$$\begin{aligned} Q((1, \zeta), \{(m', \zeta')\}) &= \left(\lambda_E \mathbb{1}_{\{m'=7\}} + \mu_{Rb}(\zeta_{rb}) \mathbb{1}_{\{m'=5\}} + \mu_{Maf}(\zeta_{maf}) \mathbb{1}_{\{m'=2\}} + \mu_{Bf}(\zeta_{bf}) \mathbb{1}_{\{m'=3\}} \right) \\ &\times \frac{1}{\lambda(1, \zeta)} \mathbb{1}_{\{(\zeta'_{rb}, \zeta'_{maf}, \zeta'_{bf}, \zeta'_{cal}, \zeta'_{saut}) = (\zeta_{rb}, \zeta_{maf}, \zeta_{bf}, \zeta_{cal}, 0)\} \cap \{\zeta_{cal} \neq d_{étude}\}} \\ &+ \mathbb{1}_{\{(m', \zeta') = \Delta\} \cap \{\zeta_{cal} = d_{étude}\}}. \end{aligned} \quad (2.5)$$

Si le régime de départ est le régime dégradé 2, c'est-à-dire si la machine à froid est dans l'état dégradé mais que les autres éléments de la caméra sont dans l'état stable, alors quatre sauts sont possibles (voir figure 2.6) : dans le régime 5 pour lequel le bloc refroidissement est en panne, dans le régime 9 pour lequel l'électronique est en panne, dans le régime 4 qui correspond à une dégradation du biafocal, et enfin dans l'état cimetièrre $m = 27$ si la durée $d_{étude}$ est écoulée. L'expression du noyau de transition est donc la suivante :

$$\begin{aligned} Q((2, \zeta), \{(m', \zeta')\}) &= \left(\lambda_E \mathbb{1}_{\{m'=9\}} + (\mu_{Rb}(\zeta_{rb}) + \lambda_{Maf}) \mathbb{1}_{\{m'=5\}} + \mu_{Bf}(\zeta_{bf}) \mathbb{1}_{\{m'=4\}} \right) \\ &\times \frac{1}{\lambda(2, \zeta)} \mathbb{1}_{\{(\zeta'_{rb}, \zeta'_{maf}, \zeta'_{bf}, \zeta'_{cal}, \zeta'_{saut}) = (\zeta_{rb}, \zeta_{maf}, \zeta_{bf}, \zeta_{cal}, 0)\} \cap \{\zeta_{cal} \neq d_{étude}\}} \\ &+ \mathbb{1}_{\{(m', \zeta') = \Delta\} \cap \{\zeta_{cal} = d_{étude}\}}. \end{aligned} \quad (2.6)$$

Si le régime de départ est le régime dégradé 3, c'est-à-dire si le biafocal est dans l'état dégradé mais que les autres éléments de la caméra sont dans l'état stable, alors, par la figure 2.6, les sauts sont possibles dans le régime 11 si l'électronique est en panne, dans le régime 6 si le biafocal est en panne, dans le régime 4 si la machine à froid est dégradée et dans le régime cimetièrre 27 quand la caméra a atteint le temps $d_{étude}$. Par conséquent, le noyau de transition est le suivant

pour $m = 3$:

$$\begin{aligned} Q((3, \zeta), \{(m', \zeta')\}) &= \left(\lambda_E \mathbb{1}_{\{m'=11\}} + \mu_{Rb}(\zeta_{rb}) \mathbb{1}_{\{m'=10\}} + \lambda_{Bf} \mathbb{1}_{\{m'=6\}} + \mu_{Maf}(\zeta_{maf}) \mathbb{1}_{\{m'=4\}} \right) \\ &\times \frac{1}{\lambda(3, \zeta)} \mathbb{1}_{\{(\zeta'_{rb}, \zeta'_{maf}, \zeta'_{bf}, \zeta'_{scal}, \zeta'_{saut}) = (\zeta_{rb}, \zeta_{maf}, \zeta_{bf}, \zeta_{scal}, 0)\} \cap \{\zeta_{scal} \neq d_{étude}\}} \\ &+ \mathbb{1}_{\{(m', \zeta') = \Delta\} \cap \{\zeta_{scal} = d_{étude}\}}. \end{aligned} \quad (2.7)$$

Dans le régime 4, la machine à froid et le bifaocal sont dans l'état dégradé mais le roulement à billes et l'électronique fonctionnent. Au vu de la figure 2.6, les seuls sauts possibles mènent, soit à un régime de panne (8 si le bifaocal tombe en panne, 10 si c'est le bloc refroidissement et 12 s'il s'agit de l'électronique), soit au régime cimetièrre 27 quand la caméra a atteint la durée $d_{étude}$. On a ainsi :

$$\begin{aligned} Q((4, \zeta), \{(m', \zeta')\}) &= \left(\lambda_E \mathbb{1}_{\{m'=12\}} + (\mu_{Rb}(\zeta_{rb}) + \lambda_{Maf}) \mathbb{1}_{\{m'=10\}} + \lambda_{Bf} \mathbb{1}_{\{m'=8\}} \right) \\ &\times \frac{1}{\lambda(4, \zeta)} \mathbb{1}_{\{(\zeta'_{rb}, \zeta'_{maf}, \zeta'_{bf}, \zeta'_{scal}, \zeta'_{saut}) = (\zeta_{rb}, \zeta_{maf}, \zeta_{bf}, \zeta_{scal}, 0)\} \cap \{\zeta_{scal} \neq d_{étude}\}} \\ &+ \mathbb{1}_{\{(m', \zeta') = \Delta\} \cap \{\zeta_{scal} = d_{étude}\}}. \end{aligned} \quad (2.8)$$

Si la caméra est en panne, le processus est dans un régime de panne : $m \in \mathcal{P}$. La dynamique naturelle du processus (voir figure 2.7) est de rester dans cet état jusqu'à atteindre la frontière, qui correspond au moment où le temps calendaire atteint la fin de l'étude $d_{étude}$. À cet instant, la caméra est envoyée dans l'état cimetièrre. Par conséquent, on définit le noyau uniquement sur la frontière $\partial_{panne} E = \mathcal{P} \times \mathbb{R}_+^3 \times \{d_{étude}\} \times \mathbb{R}_+$ et on le définit par

$$Q((m, \zeta), \{(m', \zeta')\}) = \mathbb{1}_{\{(m', \zeta') = \Delta\}}. \quad (2.9)$$

L'atelier n'est pas accessible naturellement pour le processus : le noyau de transition n'amène jamais aux régimes 13 à 26. Le processus ne peut y accéder que par le biais des maintenances. Cependant, si on part d'un régime atelier ($m \in \mathcal{A}$), la dynamique naturelle de saut va amener à l'état stable au bout de la durée $d_{atelier}$ ou à l'état cimetièrre si le temps calendaire atteint $d_{étude}$. Ainsi, on définit uniquement le noyau de Markov sur la frontière $\partial_{atelier} E = \mathcal{A} \times \mathbb{R}_+^3 \times \{d_{étude}\} \times \mathbb{R}_+ \cup \mathcal{A} \times \mathbb{R}_+^3 \times [0; d_{étude}] \times \{d_{atelier}\}$. Pour $(m, \zeta) \in \partial_{atelier} E$, on pose

$$\begin{aligned} Q((m, \zeta), \{(m', \zeta')\}) &= \mathbb{1}_{\{\zeta_{saut} = d_{atelier}\} \cap \{\zeta_{scal} \neq d_{étude}\} \cap \{m'=1\} \cap \{(\zeta'_{rb}, \zeta'_{maf}, \zeta'_{bf}, \zeta'_{scal}, \zeta'_{saut}) = (\zeta_{rb}, \zeta_{maf}, \zeta_{bf}, \zeta_{scal}, 0)\}} \\ &+ \mathbb{1}_{\{(m', \zeta') = \Delta\} \cap \{\zeta_{scal} = d_{étude}\}}. \end{aligned} \quad (2.10)$$

2.4 Problème de contrôle impulsif

On souhaite optimiser les coûts de maintenance de la caméra thermique Sophie produite par Thales ainsi que sa disponibilité auprès du client. Cela revient à trouver la meilleure politique d'intervention sur la caméra, en matière de coûts. Par politique d'intervention, on entend une manière précise de définir les dates d'intervention ainsi qu'une manière spécifique de choisir les opérations effectuées lors des maintenances. En pratique, on souhaite à n'importe quel instant de vie de la caméra, connaître une date optimale de maintenance de la caméra et déterminer les opérations optimales à effectuer lors des maintenances.

Afin de trouver une politique optimale d'intervention, on se place dans le cadre formel du contrôle impulsif, défini au paragraphe 1.6.2. Ce problème consiste à minimiser une fonction coût qui comprend le coût continu d'indisponibilité de la caméra et le coût des maintenances réalisées dessus. La fonction coût est calculée à partir des trajectoires de la caméra, c'est-à-dire des évolutions de la caméra, qui suivent une politique d'intervention donnée. L'infimum de cette fonction coût est pris parmi toutes les politiques d'intervention possibles. Il est appelé

fonction valeur ou coût optimal. Notre but est de calculer une valeur approchée de ce coût optimal.

La formule du coût d'une stratégie d'intervention \mathcal{S} est donnée par (1.4). Dans toute la suite de l'étude, on considérera uniquement des caméras partant de l'état neuf, c'est-à-dire du point $x = (1, 0, 0, 0, 0, 0)$. Le coût d'une stratégie fait appel à la fonction f de coût courant et la fonction c du coût des interventions, cette dernière étant définie sur $E \times \mathbb{U}$, où \mathbb{U} est l'ensemble de contrôle associé à notre problème. On commence donc par définir l'ensemble de contrôle \mathbb{U} dans la section 2.4.1, puis on définit les fonctions coûts f et c au paragraphe 2.4.2. On finit par une discussion sur le choix du facteur d'oubli dans le paragraphe 2.4.3.

2.4.1 Ensemble de contrôle

L'ensemble de contrôle \mathbb{U} est l'ensemble des points où l'on peut envoyer la caméra lors d'une intervention. Pour l'approximation de la fonction valeur décrite au paragraphe 1.8, on impose que cet ensemble soit fini et soit indépendant de l'endroit où se trouve la caméra avant l'intervention.

Cependant, il est important de noter que ces hypothèses ne seront pas imposées dans l'ensemble du document ; l'étude préliminaire des trajectoires du PDMP modélisant l'évolution d'une caméra, détaillée au paragraphe 3.2, se fait dans le cadre d'un ensemble de contrôle infini, que l'on note \mathbb{U}_∞ et que l'on définit par $\mathbb{U}_\infty = \mathcal{A} \times \mathbb{R}_+^3 \times [0; d_{étude}[\times\{0\}$.

Dans la suite de ce paragraphe, on discutera du choix de l'ensemble de contrôle dans le cadre de l'approximation de la fonction valeur, c'est-à-dire quand l'ensemble de contrôle doit être fini et indépendant de l'endroit où se trouve la caméra avant l'intervention. On note cet ensemble \mathbb{U}_{fini} .

Une intervention correspond à un envoi de la caméra à l'atelier avec une ou plusieurs opérations de maintenance à réaliser durant le passage à l'atelier de la caméra. Par conséquent, tous les points de l'ensemble \mathbb{U}_{fini} doivent avoir pour régime un régime atelier. Une intervention correspondant à un saut du processus, nécessairement le temps depuis le dernier saut doit être égal à 0. L'ensemble \mathbb{U}_{fini} doit donc être un sous-ensemble de $\mathcal{A} \times \mathbb{R}_+^3 \times [0; d_{étude}[\times\{0\}$.

À l'entrée de l'atelier, on impose la remise à 0 des temps de fonctionnements des éléments changés ou réparés durant le passage de la caméra à l'atelier. On distingue six possibilités de remise à 0 des temps de fonctionnements, qui correspondent à 6 sous-ensembles de régimes atelier, notés \mathcal{A}_1 à \mathcal{A}_6 , à savoir :

- \mathcal{A}_1 : temps de fonctionnement du roulement à billes et de la machine à froid remis à 0,
- \mathcal{A}_2 : temps de fonctionnement du bifaocal remis à 0,
- \mathcal{A}_3 : aucun temps de fonctionnement remis à 0,
- \mathcal{A}_4 : temps de fonctionnement de la machine à froid et du bifaocal remis à 0,
- \mathcal{A}_5 : temps de fonctionnement de la machine à froid remis à 0,
- \mathcal{A}_6 : tous les temps de fonctionnement des éléments remis à 0.

Les régimes correspondants à chaque sous-ensemble de \mathcal{A} , ainsi que la liste des temps de fonctionnement remis à 0 sont donnés dans le tableau 2.5.

À chaque sous-ensemble \mathcal{A}_i , $i \in \{1, \dots, 6\}$, on associe l'ensemble \mathbb{U}_i constitué des points ayant un régime appartenant à \mathcal{A}_i , où la caméra peut se retrouver après une intervention. Ainsi l'ensemble de contrôle peut s'écrire comme la réunion de ces six ensembles :

$$\mathbb{U}_{fini} = \bigcup_{i=1}^6 \mathbb{U}_i.$$

| ensemble atelier | régimes | temps de fonctionnement remis à 0 | ensemble de projection associé |
|------------------|----------------|---------------------------------------|---|
| \mathcal{A}_1 | 13, 22 | ζ_{rb}, ζ_{maf} | $A_1 = \{0\} \times \{0\} \times \mathbb{R}^+ \times [0; d_{étude}[\times\{0\}$ |
| \mathcal{A}_2 | 14, 19, 23 | ζ_{bf} | $A_2 = \mathbb{R}^+ \times \mathbb{R}^+ \times \{0\} \times [0; d_{étude}[\times\{0\}$ |
| \mathcal{A}_3 | 15 | - | $A_3 = \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+ \times [0; d_{étude}[\times\{0\}$ |
| \mathcal{A}_4 | 16, 20, 24 | ζ_{maf}, ζ_{bf} | $A_4 = \mathbb{R}^+ \times \{0\} \times \{0\} \times [0; d_{étude}[\times\{0\}$ |
| \mathcal{A}_5 | 17 | ζ_{maf} | $A_5 = \mathbb{R}^+ \times \{0\} \times \mathbb{R}^+ \times [0; d_{étude}[\times\{0\}$ |
| \mathcal{A}_6 | 18, 21, 25, 26 | $\zeta_{rb}, \zeta_{maf}, \zeta_{bf}$ | $A_6 = \{0\} \times \{0\} \times \{0\} \times [0; d_{étude}[\times\{0\}$ |

TAB. 2.5 – Sous-ensembles des régimes de l'ensemble \mathcal{A} des régimes d'atelier, construits en fonction des éléments dont les temps de fonctionnement sont remis à 0 à l'entrée de l'atelier.

On a les inclusions suivantes :

$$\begin{aligned}
\mathbb{U}_1 &\subset \mathcal{A}_1 \times \{0\}^2 \times \mathbb{R}_+ \times [0; d_{étude}[\times\{0\} \\
\mathbb{U}_2 &\subset \mathcal{A}_2 \times \mathbb{R}_+^2 \times \{0\} \times [0; d_{étude}[\times\{0\} \\
\mathbb{U}_3 &\subset \mathcal{A}_3 \times \mathbb{R}_+^3 \times [0; d_{étude}[\times\{0\} \\
\mathbb{U}_4 &\subset \mathcal{A}_4 \times \mathbb{R}_+ \times \{0\}^2 \times [0; d_{étude}[\times\{0\} \\
\mathbb{U}_5 &\subset \mathcal{A}_5 \times \mathbb{R}_+ \times \{0\} \times \mathbb{R}_+ \times [0; d_{étude}[\times\{0\} \\
\mathbb{U}_6 &\subset \mathcal{A}_6 \times \{0\}^3 \times [0; d_{étude}[\times\{0\}.
\end{aligned}$$

Il reste à définir précisément les ensembles $(\mathbb{U}_i)_{1 \leq i \leq 6}$. Comme nous travaillons dans ce paragraphe avec l'ensemble de contrôle \mathbb{U}_{fini} , on ne peut pas choisir $\mathbb{U}_1 = \mathcal{A}_1 \times \{0\}^2 \times \mathbb{R}_+ \times [0; d_{étude}[\times\{0\}$, $\mathbb{U}_2 = \mathcal{A}_2 \times \mathbb{R}_+^2 \times \{0\} \times [0; d_{étude}[\times\{0\}$, ... On introduit donc deux grilles finies G et H et on définit les ensembles suivants :

$$\begin{aligned}
U_1 &:= \{0\} \times \{0\} \times G \times H \times \{0\} \\
U_2 &:= G \times G \times \{0\} \times H \times \{0\} \\
U_3 &:= G \times G \times G \times H \times \{0\} \\
U_4 &:= G \times \{0\} \times \{0\} \times H \times \{0\} \\
U_5 &:= G \times \{0\} \times G \times H \times \{0\} \\
U_6 &:= \{0\} \times \{0\} \times \{0\} \times H \times \{0\}.
\end{aligned}$$

La grille finie G permet de discrétiser les temps de fonctionnements des éléments de la caméra et la grille finie H permet de discrétiser le temps calendaire de la caméra.

Cependant, on ne souhaite pas prendre l'ensemble des points appartenant à $\mathcal{A}_i \times U_i \times \{0\}$ pour définir l'ensemble \mathbb{U}_i . On rajoute une condition que doivent satisfaire les points de U_i : le temps calendaire doit être toujours supérieur à tous les temps de fonctionnement des éléments de la caméra. Cette condition nous permet de décrire précisément les ensembles $(\mathbb{U}_i)_{1 \leq i \leq 6}$ par :

$$\forall i \in \{1, \dots, 6\}, \quad \mathbb{U}_i := \mathcal{A}_i \times \mathbb{U}_i^\zeta, \quad (2.11)$$

où

$$\mathbb{U}_i^\zeta := \{\zeta = (\zeta_{rb}, \zeta_{maf}, \zeta_{bf}, \zeta_{scal}, 0) \in U_i \mid \zeta_{scal} \geq \max(\zeta_{rb}, \zeta_{maf}, \zeta_{bf})\}. \quad (2.12)$$

Ainsi, l'ensemble de contrôle $\mathbb{U}_{fini} = \bigcup_{i=1}^6 \mathbb{U}_i$ est bien fini.

Le choix de la grille finie G de discrétisation des temps de fonctionnement et de la grille finie H des temps calendaires discrétisés fait l'objet du paragraphe 3.4.

2.4.2 Coûts associés

L'objectif de notre étude est d'optimiser la maintenance de la caméra Sophie en terme de coûts de maintenance, mais également en terme de disponibilité de la caméra auprès du client. C'est la raison pour laquelle deux types de coûts sont pris en considération : le coût d'indisponibilité et le coût des interventions.

Coût d'indisponibilité

Le coût d'indisponibilité de la caméra auprès du client peut être vu comme un coût courant. En effet, la caméra ne coûte rien en terme d'indisponibilité quand celle-ci fonctionne. En revanche, quand la caméra ne fonctionne pas, elle a un coût d'indisponibilité journalier. Par conséquent, on fait correspondre le coût d'indisponibilité à la fonction coût courant f dans la formule (1.4) du coût d'une stratégie.

On considère que le coût d'indisponibilité est nul quand la caméra fonctionne ou est à l'atelier. En effet, on ne pénalise pas le temps où la caméra est à l'atelier car on considère que le prix des interventions (au travers du prix fixe) tient compte de l'indisponibilité de la caméra auprès du client. Quand la caméra est en panne, on évalue la perte d'argent pour le client, due à l'indisponibilité de la caméra, au coût d'amortissement de la caméra sur toute la période de panne. On décide donc de poser pour tout état $x = (m, \zeta) \in E$ de la caméra :

$$f(m, \zeta) = \gamma \times \mathbb{1}_{\{m \in \mathcal{P}\}}, \quad (2.13)$$

où γ est le coût d'amortissement journalier de la caméra. La fonction coût courant f ainsi définie respecte bien l'hypothèse 5. On pose arbitrairement $\gamma = 11$ pour toute la suite de l'étude.

Coûts des interventions

Une intervention permet de faire passer la caméra dans un état appartenant à l'ensemble de contrôle \mathbb{U} , de manière instantanée. En particulier, les interventions permettent d'accéder aux régimes d'atelier, ces derniers n'étant pas accessibles par la dynamique naturelle de la caméra. Les interventions ont pour objectif de faire revenir la caméra dans un état de fonctionnement quand celle-ci est en panne. On va donc contraindre notre modèle afin de n'autoriser que les interventions qui répondent aux critères (\mathcal{C}_1) à (\mathcal{C}_3) définis ci-dessous. Au préalable, on introduit une nouvelle notation : pour tout $i \in \{1, \dots, 6\}$, on note Π_{A_i} la projection de l'ensemble \mathbb{R}_+^5 sur l'ensemble A_i défini dans le tableau 2.5.

(\mathcal{C}_1) Lors d'une intervention, la caméra passe d'un régime de panne à un régime atelier.

(\mathcal{C}_2) L'intervention permet à la caméra d'être dans l'état stable à la sortie de l'atelier.

(\mathcal{C}_3) L'intervention qui fait passer la caméra de l'état $x = (m, \zeta)$ à l'état $y = (m', \zeta') \in \mathbb{U}$ doit être telle que, en notant i_0 l'indice qui vérifie $m' \in \mathcal{A}_{i_0}$, on doit avoir :

- si $\mathbb{U} = \mathbb{U}_\infty = \mathcal{A} \times \mathbb{R}_+^3 \times [0; d_{étude}] \times \{0\}$,

$$\zeta' = \Pi_{A_{i_0}}(\zeta),$$

- si $\mathbb{U} = \mathbb{U}_{fini} = \cup_{i=1}^6 \{\mathcal{A}_i \times \mathbb{U}_i^\zeta\}$ où les ensembles \mathbb{U}_i^ζ sont définis par la formule (2.12), ζ' doit être égal au plus proche voisin au sens de la norme euclidienne de $\Pi_{A_{i_0}}(\zeta)$ dans l'ensemble $\mathbb{U}_{i_0}^\zeta$.

Le graphique d'évolution de la caméra contrôlée avec les interventions répondant à ces trois critères est donnée par la figure 2.8.

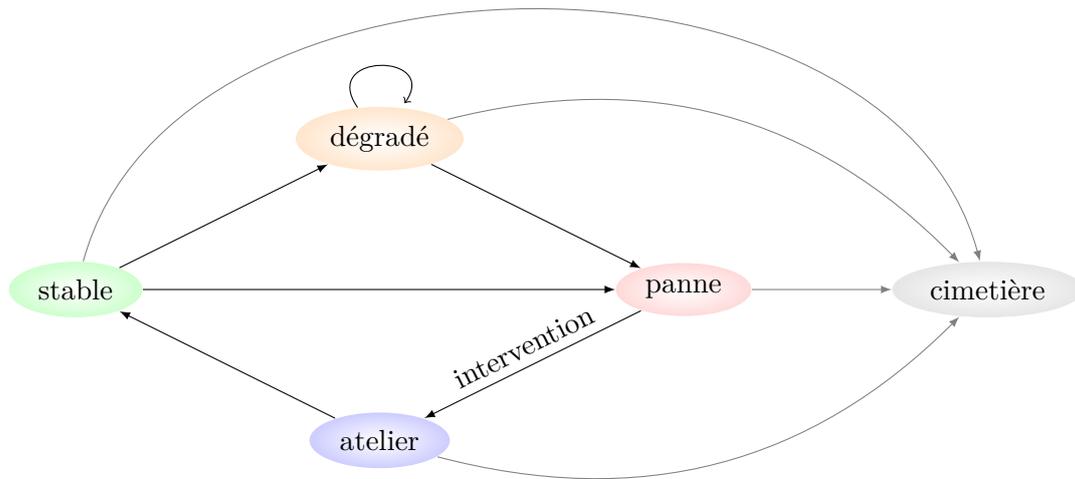


FIGURE 2.8 – Graphique d'évolution générale de la caméra Sophie si celle-ci est contrôlée à l'aide d'interventions satisfaisant les critères (\mathcal{C}_1) à (\mathcal{C}_3) .

Les critères (\mathcal{C}_1) et (\mathcal{C}_2) impliquent que l'intervention choisie doit au minimum changer le bloc en panne sur la caméra et réparer les éléments qui se trouvent dans l'état dégradé. Par exemple, si la caméra est dans le régime de panne $m = 5$ correspondant à la panne du bloc de refroidissement, alors les interventions permises sont les suivantes :

- arriver dans le régime $m = 13$, où le bloc de refroidissement est changé,
- arriver dans le régime $m = 18$, où le bloc de refroidissement est changé et le bifaocal est réparé,
- arriver dans le régime $m = 21$, où le bloc de refroidissement et le bifaocal sont changés,
- arriver dans le régime $m = 22$, où le bloc de refroidissement et l'électronique sont changés,
- arriver dans le régime $m = 25$, où le bloc de refroidissement et l'électronique sont changés et le bifaocal réparé,
- arriver dans le régime $m = 26$, où tous les blocs sont changés.

La fonction c du coût des interventions est définie sur l'ensemble $E \times \mathbb{U}$. Or on ne souhaite pas avoir des interventions qui font passer, par exemple, la caméra de l'état stable à l'atelier. Il nous faut donc empêcher certaines interventions en fonction de l'état dans lequel se trouve la caméra. Néanmoins, l'ensemble de contrôle \mathbb{U} est fixé. On est donc obligé de définir $c(x, y)$, pour tout état x de la caméra et pour tout état y de l'ensemble de contrôle.

Mettre les coûts des interventions que l'on ne souhaite pas voir se réaliser, égaux à l'infini, empêcherait ces interventions de se produire. Cependant, l'hypothèse 6 impose que la fonction c des coûts d'intervention soit bornée. On va donc pénaliser fortement les coûts des interventions ne répondant pas aux critères (\mathcal{C}_1) , (\mathcal{C}_2) et (\mathcal{C}_3) , à l'aide d'une constante de pénalisation p . Le choix de cette constante sera discuté au paragraphe 4.3.1.

Afin de répondre aux critères (\mathcal{C}_1) et (\mathcal{C}_2) , il est nécessaire de construire au préalable la matrice C des coûts de passage d'un régime à un autre, avant de définir le coût d'intervention d'un état à un autre. Cette matrice est de taille 27×27 et contient les éléments $C(m, m')$ qui correspondent au coût de passage du régime m au régime m' si celui-ci répond aux critères (\mathcal{C}_1)

| régime | E | RB | MAF | BF | coût |
|--------|------------|------------|------------|------------|-------|
| 13 | - | changement | changement | - | 8000 |
| 14 | - | - | - | changement | 8000 |
| 15 | changement | - | - | - | 5000 |
| 16 | - | - | réparation | changement | 9000 |
| 17 | changement | - | réparation | - | 6000 |
| 18 | - | changement | changement | réparation | 9000 |
| 19 | changement | - | - | réparation | 6000 |
| 20 | changement | - | réparation | réparation | 7000 |
| 21 | - | changement | changement | changement | 14000 |
| 22 | changement | changement | changement | - | 11000 |
| 23 | changement | - | - | changement | 11000 |
| 24 | changement | - | réparation | changement | 12000 |
| 25 | changement | changement | changement | réparation | 12000 |
| 26 | changement | changement | changement | changement | 17000 |

TAB. 2.6 – Coût total des maintenances pour les régimes d'atelier de la caméra.

et (\mathcal{C}_2) , nul sinon. Par conséquent, si le régime de départ m n'est pas un régime de panne, on ne souhaite pas intervenir :

$$\forall m \notin \mathcal{P}, \forall m' \in \mathcal{M}, \quad C(m, m') = 0.$$

En revanche, si on part d'un régime de panne, certaines interventions sont possibles car répondent bien aux deux critères (\mathcal{C}_1) et (\mathcal{C}_2) . On détaille ci-dessous les cas possibles.

- Si on part du régime 5, pour être sûr d'être dans l'état stable à l'issue de l'atelier, le bloc refroidissement de la caméra doit être changé. On peut donc aller dans le régime 13, qui correspond au changement du bloc refroidissement, ou dans les régimes 18, 21, 22, 25 ou 26 pour lesquels le bloc refroidissement est changé et d'autres opérations sont effectuées sur le reste de l'équipement. D'après le tableau 2.6, aller dans le régime 13 coûte un total de 8000, aller dans le régime 18 coûte 9000, *etc.* L'expression de C est donc la suivante, quand le régime de départ est le 5 :

$$\forall m' \in \mathcal{M}, \quad C(5, m') = 8000 \times \mathbb{1}_{\{m'=13\}} + 9000 \times \mathbb{1}_{\{m'=18\}} + 14000 \times \mathbb{1}_{\{m'=21\}} \quad (2.14) \\ + 11000 \times \mathbb{1}_{\{m'=22\}} + 12000 \times \mathbb{1}_{\{m'=25\}} + 17000 \times \mathbb{1}_{\{m'=26\}}.$$

- Le régime 6 correspond à une caméra dont le bifaocal est en panne. On doit donc au minimum le changer, ce qui donne six régimes possibles, qui sont les régimes 14, 16, 21, 23, 24 et 26 :

$$\forall m' \in \mathcal{M}, \quad C(6, m') = 8000 \times \mathbb{1}_{\{m'=14\}} + 9000 \times \mathbb{1}_{\{m'=16\}} + 14000 \times \mathbb{1}_{\{m'=21\}} \quad (2.15) \\ + 11000 \times \mathbb{1}_{\{m'=23\}} + 12000 \times \mathbb{1}_{\{m'=24\}} + 17000 \times \mathbb{1}_{\{m'=26\}}.$$

- Dans le régime 7, l'électronique est en panne donc on a le choix entre les régimes 15, 17, 19, 20 et 22 à 26 :

$$\forall m' \in \mathcal{M}, \quad C(7, m') = 5000 \times \mathbb{1}_{\{m'=15\}} + 6000 \times \mathbb{1}_{\{m' \in \{17, 19\}\}} + 7000 \times \mathbb{1}_{\{m'=20\}} \quad (2.16) \\ + 11000 \times \mathbb{1}_{\{m' \in \{22, 23\}\}} + 12000 \times \mathbb{1}_{\{m' \in \{24, 25\}\}} + 17000 \times \mathbb{1}_{\{m'=26\}}.$$

- Le régime 8 nous indique que la machine à froid est dégradée et le biafocal en panne. On doit impérativement changer le biafocal dans l'atelier. Pour la machine à froid, deux choix sont possibles. Elle peut être réparée (minimum pour être de retour dans l'état stable après l'atelier) ou le bloc refroidissement, qui comprend la machine à froid, peut être changé. Par conséquent, les régimes d'atelier que l'on souhaite autoriser sont les régimes 16, 21, 24 et 26.

$$\forall m' \in \mathcal{M}, \quad C(8, m') = 9000 \times \mathbb{1}_{\{m'=16\}} + 14000 \times \mathbb{1}_{\{m'=21\}} + 12000 \times \mathbb{1}_{\{m'=24\}} + 17000 \times \mathbb{1}_{\{m'=26\}}. \quad (2.17)$$

- De même, dans le régime 9, la machine à froid est dégradée. On a donc le choix entre la réparer ou changer le bloc refroidissement, en plus de changer l'électronique qui est en panne dans ce régime. Ces contraintes nous obligent à aller dans les régimes 17, 20, 22, 24, 25 ou 26.

$$\forall m' \in \mathcal{M}, \quad C(9, m') = 6000 \times \mathbb{1}_{\{m'=17\}} + 7000 \times \mathbb{1}_{\{m'=20\}} + 11000 \times \mathbb{1}_{\{m'=22\}} + 12000 \times \mathbb{1}_{\{m' \in \{24, 25\}\}} + 17000 \times \mathbb{1}_{\{m'=26\}}. \quad (2.18)$$

- Dans le régime 10, on doit changer le bloc refroidissement et réparer ou changer le bloc biafocal qui est dans l'état dégradé. Les différents régimes possibles sont 18, 21, 25 et 26.

$$\forall m' \in \mathcal{M}, \quad C(10, m') = 9000 \times \mathbb{1}_{\{m'=18\}} + 14000 \times \mathbb{1}_{\{m'=21\}} + 12000 \times \mathbb{1}_{\{m'=25\}} + 17000 \times \mathbb{1}_{\{m'=26\}}. \quad (2.19)$$

- Le régime 11 impose le changement de l'électronique et la réparation ou le changement du bloc biafocal. Ainsi, les régimes que l'on souhaite autoriser depuis 11 sont les régimes 19, 20, 23, 24, 25 et 26.

$$\forall m' \in \mathcal{M}, \quad C(11, m') = 6000 \times \mathbb{1}_{\{m'=19\}} + 7000 \times \mathbb{1}_{\{m'=20\}} + 11000 \times \mathbb{1}_{\{m'=23\}} + 12000 \times \mathbb{1}_{\{m' \in \{24, 25\}\}} + 17000 \times \mathbb{1}_{\{m'=26\}}. \quad (2.20)$$

- Enfin, le régime 12 implique un changement de l'électronique, une réparation ou un changement pour le bloc refroidissement et le biafocal. Les seuls régimes possibles sont donc 20, 24, 25 et 26.

$$\forall m' \in \mathcal{M}, \quad C(12, m') = 7000 \times \mathbb{1}_{\{m'=20\}} + 12000 \times \mathbb{1}_{\{m' \in \{24, 25\}\}} + 17000 \times \mathbb{1}_{\{m'=26\}}. \quad (2.21)$$

La matrice C étant ainsi définie, il ne reste plus qu'à prendre en compte le critère (C_3) pour pouvoir définir la fonction c des coûts d'intervention. Pour cela, on introduit une nouvelle notation : pour tout $z \in \mathbb{R}_+^3 \times [0; d_{étude}[$ et pour tout ensemble $B \subset \mathbb{R}_+^3 \times [0; d_{étude}[$, on note $\text{voisin}_B(z)$ le plus proche voisin de z dans l'ensemble B au sens de la norme euclidienne. Ainsi, la fonction coût d'intervention c , permettant de passer d'un état $x = (m, \zeta) \in E$ à un état $y = (m', \zeta') \in \mathbb{U}$, est définie de la façon suivante :

$$c(x, y) = c_{int}(x, y) \mathbb{1}_{\{c_{int}(x, y) \neq 0\}} + p \mathbb{1}_{\{c_{int}(x, y) = 0\}}, \quad (2.22)$$

où p est la constante de pénalisation et $c_{int}(x, y)$ est un coût intermédiaire, défini par :

- si $\mathbb{U} = \mathbb{U}_\infty = \mathcal{A} \times \mathbb{R}_+^3 \times [0; d_{étude}[\times \{0\}$, alors

$$c_{int}(x, y) = C(m, m') \sum_{i=1}^6 \mathbb{1}_{\{m' \in \mathcal{A}_i\} \cap \{\zeta' = \Pi_{\mathcal{A}_i}(\zeta)\}}. \quad (2.23)$$

| $b \backslash d_{\text{étude}}$ | 5 ans | 10 ans | 15 ans | 20 ans |
|---------------------------------|----------------------|----------------------|----------------------|----------------------|
| 0.90 | 5.8×10^{-5} | 2.9×10^{-5} | 1.9×10^{-5} | 1.4×10^{-5} |
| 0.95 | 2.8×10^{-5} | 1.4×10^{-5} | 9.3×10^{-6} | 7×10^{-6} |
| 0.99 | 5.5×10^{-6} | 2.8×10^{-6} | 1.8×10^{-6} | 1.4×10^{-6} |

TAB. 2.7 – Valeur de $\frac{-\ln(b)}{d_{\text{étude}}}$ pour le choix pour le facteur d'oubli α , en fonction de la durée $d_{\text{étude}}$ et de la borne b choisie.

- si $\mathbb{U} = \mathbb{U}_{\text{fini}}$ (défini dans le paragraphe 2.4.1), alors

$$c_{\text{int}}(x, y) = C(m, m') \sum_{i=1}^6 \mathbb{1}_{\{m' \in \mathcal{A}_i\} \cap \{\zeta' = \text{voisin}_{\mathbb{U}_i^\zeta}(\Pi_{\mathcal{A}_i}(\zeta))\}}. \quad (2.24)$$

Pour rappel, les ensembles \mathbb{U}_i^ζ sont définis par la formule (2.12). Le rôle de la fonction intermédiaire c_{int} est de satisfaire le critère (\mathcal{C}_3) auquel doivent répondre les interventions. Elle impose notamment, à l'entrée de l'atelier, la mise à zéro des temps de fonctionnement des éléments changés ou réparés durant l'atelier. La fonction c des coûts d'intervention ainsi définie respecte bien l'hypothèse 6 du cadre formel du contrôle impulsionnel défini au paragraphe 1.6.2.

2.4.3 Facteur d'oubli

Le problème de contrôle impulsionnel étudié est un problème où intervient un facteur d'oubli, noté α . Comme on suppose que le coût courant et les coûts d'intervention sont bornés, le facteur d'oubli permet d'assurer que l'espérance permettant de définir le coût d'une stratégie (formule (1.4)) est finie. Son rôle n'est pas primordial dans notre étude. En effet, on impose une durée d'étude fixée $d_{\text{étude}}$. Une fois ce temps atteint, on considère que la caméra est arrêtée. Elle ne coûte alors plus rien.

Par conséquent, on souhaite atténuer l'effet du facteur d'oubli pour notre étude. Ce facteur apparaît deux fois dans la formule (1.4) du coût moyen associé à une stratégie. La fonction du coût d'indisponibilité étant nulle dès que la caméra est dans l'état cimetièr, on peut écrire :

$$\mathcal{J}_x^S = \mathbb{E}_x^S \left[\int_0^{d_{\text{étude}}} e^{-\alpha s} f(\tilde{X}_s) ds + \sum_{i=1}^{\infty} e^{-\alpha \tau_i} c(\tilde{X}_{\tau_i-}, \tilde{X}_{\tau_i}) \right].$$

De même, tous les temps d'interventions τ_i doivent avoir lieu avant $d_{\text{étude}}$. L'exponentielle $e^{-\alpha t}$ étant décroissante avec le temps, il suffit par exemple de prendre α tel que

$$e^{-\alpha d_{\text{étude}}} \geq b,$$

où b est un réel inférieur à 1 mais proche de 1. Cela revient à prendre

$$\alpha \leq \frac{-\ln(b)}{d_{\text{étude}}}.$$

Le choix de la borne b vise à limiter l'effet du facteur d'oubli dans le calcul du coût moyen associé à une stratégie : plus b est proche de 1, moins le facteur d'oubli a un impact sur le coût.

Le tableau 2.7 expose les valeurs de $\frac{-\ln(b)}{d_{\text{étude}}}$ pour différentes valeurs de b et de $d_{\text{étude}}$. Il est à noter que même si dans ce tableau $d_{\text{étude}}$ est exprimé en années, le calcul de la borne pour α utilise la valeur $d_{\text{étude}}$ exprimée en jours. La valeur $d_{\text{étude}} = 10$ ans nous servant de référence pour l'application, on s'y intéresse plus particulièrement. Nous pouvons constater que les valeurs de $\frac{-\ln(b)}{d_{\text{étude}}}$ sont du même ordre de grandeur. On décide de poser pour notre étude : $\alpha = 10^{-5}$. Ainsi, pour tout $t \in [0; 3652]$, on a l'inégalité :

$$0.96 \leq e^{-\alpha t} \leq 1.$$

2.5 Conclusion

Dans ce chapitre, nous avons détaillé la structure et le fonctionnement de la caméra Sophie. Nous avons ensuite abordé les opérations de maintenance réalisables sur cette caméra et leur coût. Ceci nous a permis de mettre en place la modélisation de l'évolution d'une caméra par un processus de Markov déterministe par morceaux. L'avantage de cette modélisation est qu'elle prend en compte la partie déterministe de l'évolution d'une caméra (évolution des temps de fonctionnements des éléments de la caméra et évolution du temps calendaire), tout en considérant les événements aléatoires survenant sur la caméra (passage dans un état dégradé ou dans un état de panne d'un élément de la caméra).

La problématique industrielle, présentée au paragraphe 2.1.4, consiste à optimiser la maintenance de la caméra Sophie, au travers des décisions liées à la maintenance (date et opérations effectuées). Cette problématique se traduit par un problème de contrôle impulsif (paragraphe 2.4), où les coûts d'intervention dépendent des coûts des différentes opérations de maintenances effectuées et le coût courant correspond au coût d'indisponibilité de la caméra auprès du client. Ainsi, nous pénalisons les moments où la caméra est dans l'état de panne, ce qui répond à la volonté de Thales d'améliorer le temps de disponibilité du produit auprès de ses clients.

Nous avons choisi d'autoriser les interventions uniquement quand la caméra est en panne (critère (\mathcal{C}_1)). En effet, nous avons souhaité simplifier le plus possible le problème d'optimisation de maintenance de la caméra Sophie. Néanmoins, il est possible d'autoriser les interventions quand la caméra est dans un état stable ou dégradé par exemple. Cela impliquerait l'introduction de nouveaux régimes d'atelier et un changement de définition de la matrice C (voir paragraphe 4.4). Concernant le critère (\mathcal{C}_2) , bien qu'il soit possible de l'enlever, il ne serait pas judicieux de le faire : on aurait en effet des caméras qui se trouveraient alors dans un état dégradé ou de panne à la sortie de l'atelier. La modélisation du contrôle impulsif proposée dans cette thèse est donc susceptible d'être complexifiée, afin de permettre d'intervenir sur la caméra à n'importe quel moment. Cependant, pour une première approche, on imposera les critères (\mathcal{C}_1) et (\mathcal{C}_2) .

Le travail de ce chapitre nous permet maintenant de mettre en place des simulations de trajectoires du PDMP et d'en calculer leur coût. Par conséquent, il nous sera possible de tester différentes politiques d'intervention et de les comparer. Cette étude est réalisée dans le chapitre 3.

Chapitre 3

Simulations : trajectoires et calcul de coûts

Sommaire

| | | |
|------------|--|-----------|
| 3.1 | Politiques d'intervention de référence | 65 |
| 3.2 | Simulations de trajectoires | 67 |
| 3.2.1 | Généralités sur la simulation de trajectoires | 68 |
| 3.2.2 | Caméra sans intervention (politique d'intervention n°1) | 69 |
| 3.2.3 | Caméra avec politique d'intervention n°2, n°3 et n°4 | 73 |
| 3.3 | Coûts d'une politique d'intervention | 77 |
| 3.4 | Choix des grilles de discrétisation pour \mathbb{U}_{fini} | 80 |
| 3.5 | Conclusion | 81 |

On s'intéresse dans ce chapitre à la simulation de trajectoires du processus $\{\tilde{X}_t\}$ contrôlé par une politique d'intervention donnée, ainsi que leurs caractéristiques comme le nombre moyen de pannes par trajectoires. On commence au paragraphe 3.1 par définir quatre politiques d'intervention de référence. La simulation de trajectoires du processus $\{\tilde{X}_t\}$ est ensuite détaillée au paragraphe 3.2. Le choix du nombre de variables considérées pour les simulations des trajectoires est expliqué et on définit ce que l'on appelle une trajectoire simulée. On détaille également les caractéristiques des trajectoires suivant les politiques de référence.

Notre but étant d'optimiser le coût de la maintenance de la caméra Sophie, on calcule ensuite et compare les coûts des différentes politiques de référence, au paragraphe 3.3. L'ensemble de contrôle \mathbb{U}_{fini} , construit au paragraphe 2.4.1, dépend des grilles G et H représentant respectivement la discrétisation des temps de fonctionnement des éléments de la caméra et la discrétisation du temps calendaire. La comparaison du coût d'une politique d'intervention donnée, pour différentes grilles de discrétisation G et H , nous permet d'obtenir un critère de sélection de ces grilles. Cette comparaison est exposée au paragraphe 3.4. Enfin, on conclut ce chapitre au paragraphe 3.5 en résumant le travail effectué.

3.1 Politiques d'intervention de référence

Dans cette section, on introduit quatre politiques de référence pour notre étude, satisfaisant les critères (C_1) à (C_3) . À chacune d'elles est associée une matrice de coûts C_i , où l'on pénalise les interventions qui ne peuvent se produire quand on impose la politique d'intervention choisie.

Politique d'intervention n°1

La politique d'intervention n°1 est la politique dite sans intervention, où aucune intervention n'est permise. Dans ce cas, tous les éléments de la matrice C_1 sont nuls :

$$\forall m, m' \in \mathcal{M}, \quad C_1(m, m') = 0. \quad (3.1)$$

Ainsi pour tout $x \in E$ et $y \in \mathbb{U}$, $c_{int}(x, y) = 0$. La formule du coût (2.22) implique que tous les coûts d'intervention sont pénalisés par la constante p , c'est-à-dire

$$\forall x \in E, \forall y \in \mathbb{U}, \quad c(x, y) = p.$$

Politique d'intervention n°2

La politique d'intervention n°2 autorise les interventions répondant au critère (\mathcal{C}_3) et aux critères suivants :

- (\mathcal{C}'_1) L'intervention fait passer la caméra d'un régime de panne à un régime d'atelier après un temps d_{panne} passé dans le régime de panne.
- (\mathcal{C}'_2) Les seules opérations de maintenance réalisées dans l'atelier sont les suivantes :
 - (a) changer les blocs en panne,
 - (b) changer les éléments en état dégradé.

Les critères (\mathcal{C}'_1) et (\mathcal{C}'_2) sont plus forts que les critères (\mathcal{C}_1) et (\mathcal{C}_2) . La politique d'intervention définie ainsi vérifie bien les critères (\mathcal{C}_1) à (\mathcal{C}_3) . L'idée du critère (\mathcal{C}'_1) vient du fait que la constatation de la panne de la caméra et la prise de décision pour envoyer la caméra à l'atelier prend du temps. On considère dans notre étude que ces actions prennent un temps noté d_{panne} , que l'on fixe arbitrairement égal à un jour.

La matrice des coûts C_2 correspondante à cette politique d'intervention est définie pour $(m, m') \in \mathcal{M}^2$, par :

$$\begin{aligned} C_2(m, m') = & 8000 \times \mathbf{1}_{\{(m, m')=(5,13)\} \cup \{(m, m')=(6,14)\}} + 5000 \times \mathbf{1}_{\{(m, m')=(7,15)\}} \\ & + 14000 \times \mathbf{1}_{\{(m, m')=(8,21)\} \cup \{(m, m')=(10,21)\}} + 11000 \times \mathbf{1}_{\{(m, m')=(9,22)\} \cup \{(m, m')=(11,23)\}} \\ & + 17000 \times \mathbf{1}_{\{(m, m')=(12,26)\}}. \end{aligned} \quad (3.2)$$

Politique d'intervention n°3

Pour la politique d'intervention n°3, les interventions doivent satisfaire le critère (\mathcal{C}_3) et les deux critères :

- (\mathcal{C}'_1) L'intervention fait passer la caméra d'un régime de panne à un régime d'atelier après un temps d_{panne} passé dans le régime de panne.
- (\mathcal{C}''_2) Les seules opérations de maintenance réalisées dans l'atelier sont les suivantes :
 - (a) changer les blocs en panne,
 - (b) réparer les éléments en état dégradé.

La différence entre la politique n°3 et la politique n°2 se trouve dans la deuxième condition. La politique d'intervention n°3 impose de réparer les éléments dans l'état dégradé alors que la politique d'intervention n°2 implique de changer à neuf les éléments dans l'état dégradé.

En conséquence, la matrice C_3 des coûts se définit comme suit. Pour tous les régimes m n'étant pas des régimes de panne, la ligne $C_3(m, \cdot)$ de la matrice est nulle. Pour chaque régime m de panne, un seul régime d'atelier est possible, à savoir $m' = m + 8$. Certains régimes d'atelier (régimes 21 à 26) ne seront donc pas accessibles avec cette politique d'intervention.

$$\forall m, m' \in \mathcal{M}, \quad C_3(m, m') = \mathbb{1}_{\{m'=m+8\}} (5000 \times \mathbb{1}_{\{m=7\}} + 6000 \times \mathbb{1}_{\{m \in \{9,11\}\}} + 7000 \times \mathbb{1}_{\{m=12\}} + 8000 \times \mathbb{1}_{\{m \in \{5,6\}\}} + 9000 \times \mathbb{1}_{\{m \in \{8,10\}\}}). \quad (3.3)$$

Politique d'intervention n°4

Les interventions permises par la politique n°4 répondent au critère (C_3) et aux critères :

(C'_1) L'intervention fait passer la caméra d'un régime de panne à un régime d'atelier après un temps d_{panne} passé dans le régime de panne.

(C''_2) Les seules opérations de maintenance réalisées dans l'atelier sont les suivantes :

- (a) changer le bloc biafocal s'il est en panne,
- (b) changer le bloc refroidissement s'il est en panne,
- (c) réparer les éléments qui sont dans l'état dégradé.

Le critère (C''_2) implique notamment que le bloc électronique n'est pas changé si ce dernier est en panne. Ce cas peut se produire lors d'une rupture de stock sur une pièce, comme une carte mémoire, et que celle-ci ne peut plus être produite.

On ne peut alors plus intervenir dans les régimes 7, 9, 11 et 12. Par conséquent, seules quatre interventions sont possibles et la matrice de coût associée à cette politique est définie par :

$$\forall m, m' \in \mathcal{M}, \quad C_4(m, m') = 8000 \times \mathbb{1}_{\{(m,m')=(5,13)\} \cup \{(m,m')=(6,14)\}} + 9000 \times \mathbb{1}_{\{(m,m')=(8,16)\} \cup \{(m,m')=(10,18)\}}. \quad (3.4)$$

Comparaison des politiques d'intervention n°1 à n°4

Intuitivement, il est possible de classer ces politiques par ordre de performance en terme de coût. La politique d'intervention n°1 ne permettant pas les réparations, une fois que la caméra est en panne, le coût d'indisponibilité de la caméra augmente chaque jour. Au final, le coût d'une telle stratégie devrait être très élevé. En revanche, si on s'autorise des réparations, le coût moyen devrait diminuer. Ce qui est le cas des politiques n°2 à n°4.

Ces trois autres politiques ne diffèrent entre elles que par le second critère. La politique d'intervention n°2 impose de changer les éléments dans l'état dégradé alors que la politique d'intervention n°3 impose de réparer ces éléments. Par conséquent, les coûts des interventions seront plus élevés en moyenne pour la politique n°2, par rapport à la politique d'intervention n°3.

La politique d'intervention n°4 ne permettant pas de changer le bloc électronique si ce dernier est en panne, devrait coûter plus cher que la politique n°3, en raison d'un coût d'indisponibilité plus élevé.

Finalement, la meilleure politique est a priori la politique d'intervention n°3, qui permet les opérations de maintenance les plus adaptées à l'état du système.

3.2 Simulations de trajectoires

Dans cette partie, on s'intéresse à la simulation des trajectoires du processus représentant l'évolution d'une caméra Sophie au cours du temps. On rappelle les valeurs suivantes fixées arbitrairement pour toutes les simulations à :

- durée de l'étude des caméras : $d_{\text{étude}} = 3652$ jours, soit l'équivalent de 10 ans,
- durée du passage dans l'état de panne pour les politiques d'intervention n°2 à 4 : $d_{\text{panne}} = 1$ jour,
- durée du passage de la caméra dans l'atelier : $d_{\text{atelier}} = 35$ jours.

3.2.1 Généralités sur la simulation de trajectoires

Quelle que soit la politique d'intervention choisie, le processus $\{\tilde{X}\}_t$ modélisant l'évolution de la caméra Sophie au cours du temps possède six variables, qui sont :

- m le régime,
- ζ_{rb} le temps de fonctionnement du roulement à billes,
- ζ_{maf} le temps de fonctionnement de la machine à froid,
- ζ_{bf} le temps de fonctionnement du biafocal,
- ζ_{cal} le temps calendaire qui représente le temps écoulé depuis la mise en service de la caméra,
- ζ_{saut} le temps écoulé depuis le dernier saut du processus.

En pratique, on ne considère que les cinq premières variables du processus. On omet ζ_{saut} car cette variable n'est utilisée que pour connaître le temps de sortie de l'atelier. Or le temps de passage à l'atelier est connu et fixé. Cette durée est donnée par d_{atelier} . Dans les algorithmes, nous utiliserons donc uniquement ce paramètre pour imposer à la caméra la sortie de l'atelier.

Ainsi, pour tous les codes de simulation de trajectoires, on aura en entrée le vecteur colonne, noté X , de taille 5, représentant l'état initial de la caméra. L'état neuf de la caméra correspond à $X = (1, 0, 0, 0, 0)'$. Les éléments de X sont rangés dans le même ordre que ceux du processus, à savoir en première ligne le mode, de la deuxième à la quatrième ligne les temps de fonctionnement du roulement à billes, de la machine à froid et du biafocal et en dernière ligne le temps calendaire.

Le processus $\{\tilde{X}_t\}$ peut être entièrement décrit par la chaîne de Markov à temps discret $(\tilde{Z}_n, \tilde{S}_n)_{n \in \mathbb{N}}$, où $(\tilde{Z}_n)_{n \in \mathbb{N}}$ est la suite des positions après-saut du processus et $(\tilde{S}_n)_{n \in \mathbb{N}}$ est la suite des temps inter-sauts du processus (voir paragraphe 1.3). En effet, toute la partie aléatoire du processus $\{\tilde{X}_t\}$ est contenue dans la chaîne de Markov $(\tilde{Z}_n, \tilde{S}_n)_{n \in \mathbb{N}}$. Si l'on considère une trajectoire du processus $\{\tilde{X}_t\}$, alors celle-ci est continue. Cependant, la chaîne de Markov $(\tilde{Z}_n, \tilde{S}_n)_{n \in \mathbb{N}}$ à temps discret la décrit entièrement et il est possible à partir de cette chaîne de Markov de reconstruire la trajectoire continue du processus $\{\tilde{X}_t\}$, en utilisant le flot Φ défini par la formule (2.2).

Par conséquent, on simulera des chaînes de Markov à temps discret $(\tilde{Z}_n, \tilde{S}_n)_{n \in \mathbb{N}}$ correspondant au processus $\{\tilde{X}_t\}$. On appellera une simulation de trajectoire, une simulation de la chaîne de Markov $(\tilde{Z}_n, \tilde{S}_n)_{n \in \mathbb{N}}$ associée au processus $\{\tilde{X}_t\}$. Ainsi, chaque programme de simulations de trajectoires renvoie une matrice Y contenant les positions après-saut et les temps inter-sauts de la trajectoire. La matrice Y contient donc :

- en première ligne, les régimes successifs de la caméra
- en deuxième ligne, les temps de fonctionnement successifs du roulement à billes,
- en troisième ligne, les temps de fonctionnement successifs de la machine à froid,
- en quatrième ligne, les temps de fonctionnement successifs du biafocal,
- en cinquième ligne, les temps calendaires successifs de la caméra,

- en sixième ligne, les temps inter-sauts.

De manière plus détaillée, on stocke les positions après-saut dans la partie supérieure de la matrice Y , c'est-à-dire que $(Y(i, j))_{1 \leq i \leq 5}$ donne la $(j - 1)$ -ième position de la trajectoire, à savoir \tilde{Z}_{j-1} . Les temps inter-sauts sont stockés dans la dernière ligne de la matrice Y : l'élément $Y(6, j)$ donne le $(j - 1)$ -ième temps inter-saut \tilde{S}_{j-1} . Étant donné que le langage de programmation utilisé commence la numérotation à 1 et non à 0, le décalage des indices entre la théorie et le code ne peut être évité.

La matrice Y aura toujours six lignes mais aura un nombre variable de colonnes, dépendant du nombre de sauts de la trajectoire simulée. Il est à noter que le nombre de sauts d'une trajectoire est fini presque sûrement.

Les algorithmes décrits dans toute cette partie se situent dans le cadre de l'ensemble de contrôle $\mathbb{U}_\infty = \mathcal{A} \times \mathbb{R}_+^3 \times [0; d_{étude}] \times \{0\}$. On pourra facilement adapter ces algorithmes avec d'autres choix d'ensemble de contrôle, notamment ceux où ce dernier est fini.

3.2.2 Caméra sans intervention (politique d'intervention n°1)

Les trajectoires du processus $\{\tilde{X}_t\}$ contrôlé par la politique d'intervention n°1 correspondent aux évolutions de la caméra quand aucune opération de maintenance (réparation ou remplacement à neuf) n'est effectuée sur celle-ci. Afin de les simuler, on se base sur la figure 2.6 qui montre les possibilités d'évolution d'une caméra de l'état stable à l'état cimetière.

Algorithme

L'algorithme 3 correspond à la simulation d'une trajectoire de la caméra contrôlée par la politique d'intervention n°1 et partant d'un état initial X donné en entrée du programme. En fonction du régime initial donnée, la simulation se fait de la manière suivante.

- Si le régime initial est un régime cimetière, alors l'algorithme renvoie uniquement en sortie la trajectoire représentée par le vecteur colonne $Y = (27, 0, 0, 0, 3652, 0)'$.
- Si le régime initial est un régime de panne, alors la trajectoire renvoyée est la position initiale du processus, suivie de l'état cimetière. Par exemple, si $X = (5, 1200, 1200, 1200, 1200)'$, alors la trajectoire renvoyée est :

$$Y = \begin{pmatrix} 5 & 27 \\ 1200 & 0 \\ 1200 & 0 \\ 1200 & 0 \\ 1200 & 3652 \\ 0 & 2452 \end{pmatrix}.$$

- Si le régime initial est un régime stable, dégradé ou atelier, alors
 1. On simule un vecteur D des temps de changement de chaque élément de la caméra, en fonction du régime dans lequel est la caméra et des lois associés à chaque état de chaque élément de la caméra. Par exemple, si $X = (1, 500, 500, 500, 500)'$, alors
 - $D(1)$ (associé à l'électronique) correspond à une simulation de la loi $\mathcal{E}(\lambda_E)$,
 - $D(2)$ (associé au roulement à billes) correspond à une simulation de la loi $\mathcal{W}(\alpha_{Rb}, \beta_{Rb})$ ayant dépassé le seuil 500 (voir annexe),
 - $D(3)$ (associé à la machine à froid) correspond à une simulation de la loi $\mathcal{W}(\alpha_{Maf}, \beta_{Maf})$ ayant dépassé le seuil 500,

- $D(4)$ (associé au biafocal) correspond à une simulation de la loi $\mathcal{W}(\alpha_{Bf}, \beta_{Bf})$ ayant dépassé le seuil 500.
2. On introduit une variable t_{saut} donnant le temps du futur saut, dépendant de $d_{atelier}$ si le régime initial est un régime atelier, ou dépendant du vecteur D sinon.
 3. Tant que le temps du futur saut t_{saut} est inférieur strictement à la durée $d_{étude}$ de notre étude de la caméra, on crée un saut dont les variables sont les suivantes :
 - le nouveau régime dépend de l'ancien et éventuellement du gagnant dans la minimisation du vecteur D des temps de changements d'états des éléments,
 - les temps de fonctionnement du nouveau saut sont les temps de fonctionnement de l'ancien saut additionnés à $\min(D)$ si le régime avant le saut n'est pas un régime d'atelier, à la durée $d_{atelier}$ de passage dans l'atelier sinon,
 - le nouveau temps inter-saut est égal au $\min(D)$ si le régime avant le saut n'est pas un régime d'atelier, à la durée $d_{atelier}$ de passage dans l'atelier sinon.
 On met à jour le vecteur D et le temps du futur saut t_{saut} .
 4. Quand le temps t_{saut} du futur saut est supérieur à la durée $d_{étude}$ de notre étude, on rajoute à la trajectoire créée un dernier saut dans l'état cimetièrè.

Exemple de simulation

Pour une caméra partant de l'état neuf, c'est-à-dire avec $X = (1, 0, 0, 0, 0)'$ en entrée de l'algorithme, un exemple de simulation d'une trajectoire suivant la politique d'intervention n°1 est donné par la matrice :

$$Y = \begin{pmatrix} 1 & 2 & 5 & 27 \\ 0 & 1462 & 1530 & 0 \\ 0 & 1462 & 1530 & 0 \\ 0 & 1462 & 1530 & 0 \\ 0 & 1462 & 1530 & 3652 \\ 0 & 1462 & 68 & 2122 \end{pmatrix}.$$

En détails, l'évolution de la caméra est la suivante :

- La caméra part de l'état initial $X = (1, 0, 0, 0, 0)'$ correspondant à l'état neuf de la caméra.
- À $t = 1462$, la caméra passe de l'état stable (régime 1) à un état pour lequel la machine à froid est dégradée (régime 2).
- À $t = 1530$, le bloc refroidissement (MAF+RB) tombe en panne et la caméra passe dans le régime 5.
- À $t = 3652$: la caméra est arrêtée. Elle passe dans l'état cimetièrè (régime 27).

Cet exemple est illustré par la figure 3.1, où sont donnés les régimes de la caméra en fonction du temps calendaire, ainsi que les temps de fonctionnement de la machine à froid, du roulement à billes et du biafocal.

Caractéristiques des trajectoires

Le tableau 3.1 donne le pourcentage de trajectoires en fonction du nombre de sauts, pour 10^7 trajectoires simulées. Il est intéressant de remarquer que les trajectoires ayant 4 sauts, c'est-à-dire les trajectoires présentant deux dégradations successives avant d'arriver dans un régime de panne représentent un nombre très faible de trajectoires simulées. Par conséquent, les régimes

Algorithme 3 simulation d'une trajectoire de la caméra sans intervention (politique n°1)

Entrées : X vecteur colonne de taille 5; état initial de la caméra

$\lambda_E, \lambda_{Bf}, \lambda_{Maf}, \alpha_{Bf}, \beta_{Bf}, \alpha_{Rb}, \beta_{Rb}, \alpha_{Maf}, \beta_{Maf}, d_{étude}, d_{atelier}$ paramètres du modèle

Sortie : Y matrice (positions et temps inter-sauts) de la trajectoire

```

1   $Y = (X', 0)'$ 
2  si  $Y(1, 1)$  n'est pas un régime de panne ou le régime cimetièrè alors
3       $D$  vecteur de zéros de taille 4 : temps de changement d'état des éléments de la caméra
4       $D(1) =$  tirage d'une loi  $\mathcal{E}(\lambda_E)$ 
5       $D(2) =$  tirage d'une loi  $\mathcal{W}(\alpha_{Rb}, \beta_{Rb})$  avec seuil  $Y(5, 1)$ 
6      si  $Y(1, 1) \in \{2, 4\}$  alors
7           $D(3) =$  tirage d'une loi  $\mathcal{E}(\lambda_{Maf})$ 
8      sinon
9           $D(3) =$  tirage d'une loi  $\mathcal{W}(\alpha_{Maf}, \beta_{Maf})$  avec seuil  $Y(5, 1)$ 
10     fin si
11     si  $Y(1, 1) \in \{3, 4\}$  alors
12          $D(4) =$  tirage d'une loi  $\mathcal{E}(\lambda_{Bf})$ 
13     sinon
14          $D(4) =$  tirage d'une loi  $\mathcal{W}(\alpha_{Bf}, \beta_{Bf})$  avec seuil  $Y(5, 1)$ 
15     fin si
16     si  $Y(1, 1)$  est un régime stable ou dégradé alors
17          $t_{saut} = Y(5, 1) + \min(D)$ 
18     sinon
19          $t_{saut} = Y(5, 1) + d_{atelier}$ 
20     fin si
21     tant que  $t_{saut} < d_{étude}$  faire
22          $k$  nombre de colonnes de  $Y$ 
23          $Y = (Y, (0, 0, 0, 0, 0, 0)')$ 
24          $a = \min(D), b = \arg \min(D)$ 
25         si  $Y(1, k)$  n'est pas un mode d'atelier alors
26              $Y(1, k + 1)$  est déterminé en fonction de l'élément qui change d'état (donné
                par  $b$ )
27              $Y(2 : 5, k + 1) = Y(2 : 5, k) + a$ 
28              $Y(6, k + 1) = a$ 
29             si  $Y(1, k + 1)$  est un mode de panne alors
30                  $t_{saut} = d_{étude}$ 
31             sinon
32                 si  $b = 3$ , alors  $D(3) =$  tirage d'une  $\mathcal{E}(\lambda_{Maf})$ 
33                 si  $b = 4$ , alors  $D(4) =$  tirage d'une  $\mathcal{E}(\lambda_{Bf})$ 
34                  $t_{saut} = Y(5, k + 1) + \min(D)$ 
35             fin si
36         sinon
37              $Y(1, k + 1) = 1$ 
38              $Y(2 : 4, k + 1) = Y(2 : 4, k)$ 
39              $Y(5, k + 1) = Y(5, k) + d_{atelier}$ 
40              $Y(6, k + 1) = d_{atelier}$ 
41              $t_{saut} = t_{saut} + \min(D)$ 
42         fin si
43     fin tant que
44 fin si
45 si  $Y(1, 1)$  n'est pas le régime cimetièrè alors
46     Rajout de l'état cimetièrè dans  $Y$ 
47 fin si

```

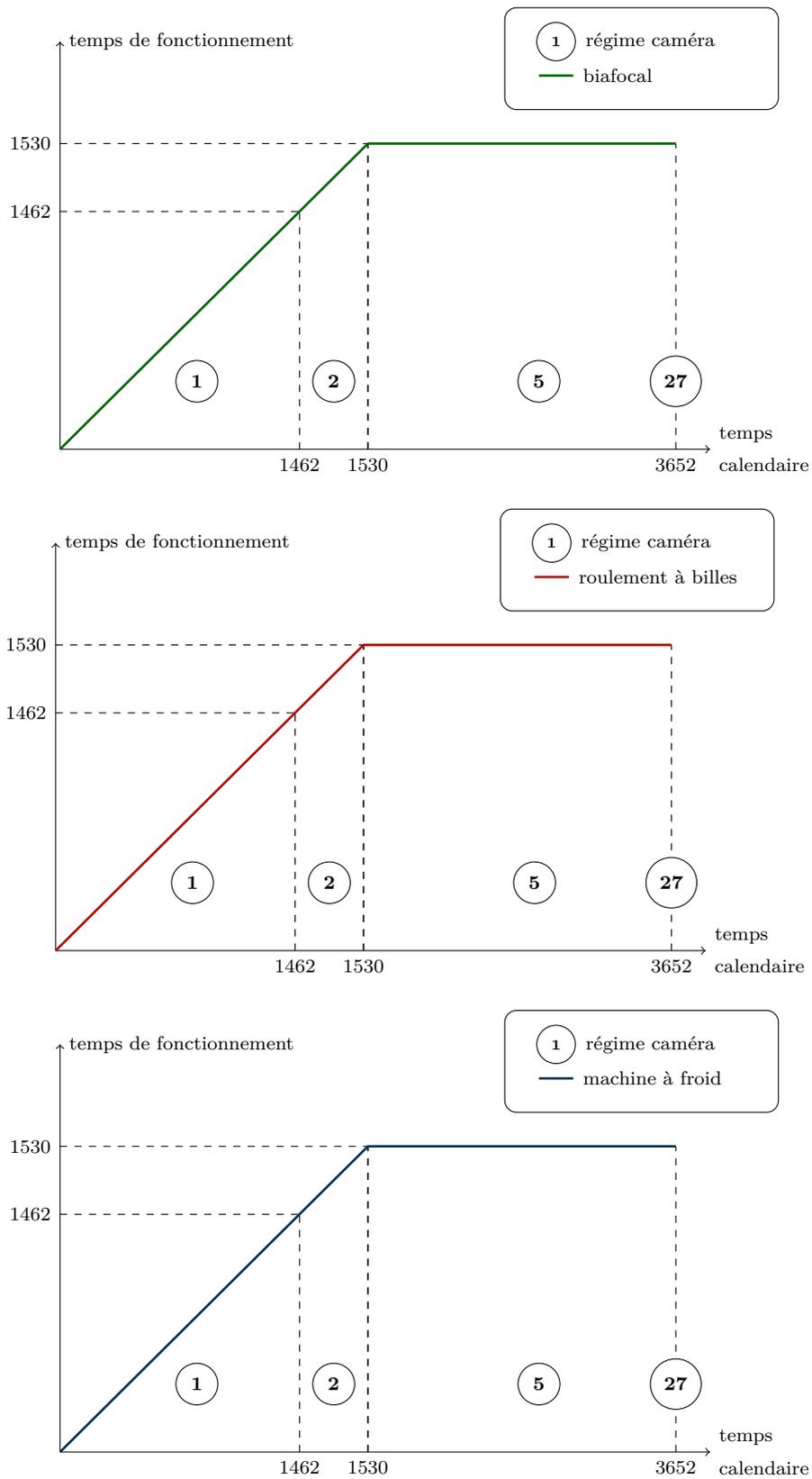


FIGURE 3.1 – Évolution des temps de fonctionnement d'une trajectoire partant de l'état initial $X = (1, 0, 0, 0, 0)'$ contrôlée par la politique d'intervention n° 1.

| nombre de sauts des trajectoires | 1 | 2 | 3 | 4 |
|----------------------------------|--------|--------|--------|-------|
| pourcentage des trajectoires | 13.91% | 50.12% | 35.88% | 0.08% |

TAB. 3.1 – Répartition donnée en pourcentage, de 10^7 trajectoires simulées pour des caméras partant de l'état neuf et avec la politique d'intervention n°1, en fonction de leur nombre de sauts.

| régime de panne visité | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|------------------------------|--------|-------|--------|-------|-------|-------|-------|--------|
| pourcentage des trajectoires | 46.95% | 1.27% | 36.59% | 0.03% | 0.63% | 0.07% | 0.04% | 0.001% |

TAB. 3.2 – Répartition exprimée en pourcentage, de 10^7 trajectoires sans intervention simulées pour des caméras partant de l'état neuf, en fonction du régime de panne dans lequel elles ont été.

8 et 12 étant uniquement accessibles par deux dégradations successives (voir figure 2.6) sont très rares. Ce phénomène demandera des précautions particulières lors de l'initialisation de la quantification des trajectoires, décrite au paragraphe 4.1.2.

Le pourcentage des trajectoires contrôlées par la politique d'intervention n°1, partant de l'état neuf, ne tombant pas en panne est d'environ 14.4%. Les régimes de pannes sont très majoritairement le régime 5 où seul le bloc refroidissement est en panne et le régime 7 où seul le bloc électronique est en panne. Les régimes de panne restants sont très peu visités. La répartition des régimes de pannes pour des trajectoires sans intervention partant d'une caméra neuve est donnée par le tableau 3.2. Il est à noter que les trajectoires visitent au plus un régime de panne car les interventions ne sont pas autorisées.

3.2.3 Caméra avec politique d'intervention n°2, n°3 et n°4

Les politiques d'intervention n°2, n°3 et n°4, définies au paragraphe 3.1 diffèrent par le deuxième critère qu'elles doivent remplir. La politique d'intervention n°2 doit satisfaire le critère

(\mathcal{C}'_2) Les seules opérations de maintenance réalisées dans l'atelier sont les suivantes :

- (a) changer les blocs en panne,
- (b) changer les éléments en état dégradé,

alors que la politique d'intervention n°3 doit satisfaire le critère

(\mathcal{C}''_2) Les seules opérations de maintenance réalisées dans l'atelier sont les suivantes :

- (a) changer les blocs en panne,
- (b) réparer les éléments en état dégradé,

et la politique d'intervention n°4 doit satisfaire le critère

(\mathcal{C}'''_2) Les seules opérations de maintenance réalisées dans l'atelier sont les suivantes :

- (a) changer le bloc bifaocal si en panne,
- (b) changer le bloc refroidissement si en panne,
- (c) réparer les éléments en état dégradé.

Algorithme

La simulation des trajectoires suivant la politique d'intervention n°3 et partant d'une position initiale X , se fait ainsi : si le régime initial est le régime cimetièrre, alors la trajectoire renvoyée est $Y = (27, 0, 0, 0, d_{étude}, 0)'$. Sinon, les étapes sont les suivantes :

1. Simulation d'un vecteur D des temps de changement d'état de chaque élément de la caméra.
2. Détermination du temps du futur saut t_{saut} :
 - si le régime initial est un régime stable ou dégradé, $t_{saut} = Y(5, 1) + \min(D)$,
 - si le régime initial est un régime de panne, $t_{saut} = Y(5, 1) + d_{panne}$,
 - si le régime initial est un régime d'atelier, $t_{saut} = Y(5, 1) + d_{atelier}$.
3. Tant que $t_{saut} < d_{étude}$, on crée un nouveau saut à notre trajectoire, que l'on note $(m', \zeta', t'_{inter-saut})$.
 - (a) Le nouveau régime m' est déterminé en fonction du régime m avant le saut :
 - Si $m \in \mathcal{S} \cup \mathcal{D}$, alors m' dépend du gagnant dans la minimisation du vecteur D (voir figure 2.6).
 - Si $m \in \mathcal{P}$, alors $m' = m + 8$ (critère (\mathcal{C}'_2)).
 - Si $m \in \mathcal{A}$, alors $m' = 1$.
 - (b) Les nouveaux temps de fonctionnement $\zeta' = (\zeta'_{Rb}, \zeta'_{Maf}, \zeta'_{Bf})$ sont déterminés à partir des temps de fonctionnement avant le saut $\zeta = (\zeta_{Rb}, \zeta_{Maf}, \zeta_{Bf})$ et des modes m et m' .
 - Si $m \in \mathcal{S} \cup \mathcal{D}$, alors $\zeta' = \zeta + (\min(D), \min(D), \min(D))$.
 - Si $m \in \mathcal{P}$, alors ζ' dépend du futur mode atelier m' ; les éléments changés ou réparés durant le passage dans le régime m' de l'atelier ont leur temps de fonctionnement remis à 0. Les temps de fonctionnement des autres éléments de la caméra restent égaux aux temps de fonctionnement avant le saut.
Par exemple, si $(m, \zeta) = (5, 1000, 1000, 1000)$, alors $(m', \zeta') = (13, 0, 0, 1000)$ car le bloc refroidissement est changé dans le régime 13 mais le bifaocal n'est ni changé ni réparé.
 - Si $m \in \mathcal{A}$, alors $\zeta' = \zeta$.
 - (c) Le temps inter-saut $t'_{inter-saut}$ vaut :
 - si $m \in \mathcal{S} \cup \mathcal{D}$, $t'_{inter-saut} = \min(D)$,
 - si $m \in \mathcal{P}$, $t'_{inter-saut} = d_{panne}$,
 - si $m \in \mathcal{A}$, $t'_{inter-saut} = d_{atelier}$.
 - (d) Mise à jour du vecteur D : le temps de fonctionnement d'un élément est mis à jour si ce dernier est passé dans un état dégradé ou si la caméra est entrée à l'atelier dans un régime qui imposait de changer ou réparer l'élément en question.
 - (e) Détermination du temps du futur saut :
 - si $m' \in \mathcal{S} \cup \mathcal{D}$, $t_{saut} = t_{saut} + \min(D)$,
 - si $m' \in \mathcal{P}$, $t_{saut} = t_{saut} + d_{panne}$,
 - si $m' \in \mathcal{A}$, $t_{saut} = t_{saut} + d_{atelier}$,
 - si $m' \in \mathcal{C}$, $t_{saut} = d_{étude}$.
4. Quand $t_{saut} \geq d_{étude}$, on rajoute à la trajectoire un saut dans l'état cimetièrre.

Les algorithmes de simulation des trajectoires suivant les politiques d'intervention n°2 et n°4 sont très semblables. On a choisi d'exposer ici l'algorithme associé à la politique d'intervention n°3 car le critère (C_2'') est le plus facile à décrire : si m est un régime de panne, alors le saut suivant va dans le régime d'atelier $m' = m + 8$.

Pour la politique d'intervention n°2, le critère (C_2') peut être traduit par :

si $m \in \{5, 6, 7\}$, alors $m' = m + 8$,

si $m = 10$, alors $m' = 21$,

si $m = 11$, alors $m' = 23$

si $m \in \{8, 9, 12\}$, alors $m' = m + 13$.

Ainsi, il suffit d'adapter l'algorithme décrit ci-dessus dans le cadre de la politique d'intervention n°3 en modifiant la partie 3(a).

En revanche, le critère (C_2''') de la politique d'intervention n°4 demande plusieurs adaptations. En effet, ce critère implique qu'on ne passe pas nécessairement d'un régime de panne à un régime atelier. Plus précisément, les régimes de panne numérotés 7, 9, 11 et 12 ne donnent pas lieu à une intervention. Par conséquent, une fois ces régimes atteints, la caméra reste en panne jusqu'à la fin de la durée de l'étude, où celle-ci est envoyée dans le cimetière. Pour les autres modes de panne, on a les transitions suivantes : si $m \in \{5, 6, 8, 10\}$, alors $m' = m + 8$.

Exemple de trajectoire

Pour une caméra partant de l'état neuf, c'est-à-dire $X = (1, 0, 0, 0, 0)'$, un exemple de simulation d'une trajectoire de caméra suivant la politique n°3, est illustré par la figure 3.2. On y voit l'évolution des temps de fonctionnements du biafocal, du roulement à billes et de la machine à froid au cours du temps. Les régimes de la caméra figurent également sur le graphique.

Dans cet exemple, on peut décomposer la trajectoire en trois parties :

1. Partant de l'état stable à $t = 0$, la machine à froid passe dans l'état dégradé à $t = 1693$, ce qui correspond au passage dans le régime 2.
À $t = 1730$, le bloc refroidissement (MAF+RB) tombe en panne ; la caméra passe au régime 5 et y reste pour la durée $d_{panne} = 1$.
La caméra part ensuite dans le régime 13 d'atelier, pour un changement du bloc refroidissement. Elle y reste pour une durée $d_{atelier} = 35$.
Enfin, la caméra revient dans le régime de stabilité 1, au temps $t = 1766$.
Les cinq premières colonnes de la matrice Y décrivent cette partie de la trajectoire :

$$(Y_{i,j})_{1 \leq i \leq 6, 1 \leq j \leq 5} = \begin{pmatrix} 1 & 2 & 5 & 13 & 1 \\ 0 & 1693 & 1730 & 0 & 0 \\ 0 & 1693 & 1730 & 0 & 0 \\ 0 & 1693 & 1730 & 1730 & 1730 \\ 0 & 1693 & 1730 & 1731 & 1766 \\ 0 & 1693 & 37 & 1 & 35 \end{pmatrix}.$$

2. L'électronique tombe en panne à $t = 2255$. La caméra passe dans le régime 7 et y reste pour une durée $d_{panne} = 1$. Elle passe ensuite à l'atelier pour un changement de l'électronique, ce qui correspond au régime 15. Enfin, elle retourne dans le régime stable 1 au temps $t = 2291$.
3. Jusqu'à la fin de l'étude, la caméra ne présente plus de dégradation ou de panne. Elle rejoint donc le régime cimetière 27 au temps $d_{étude} = 3652$.

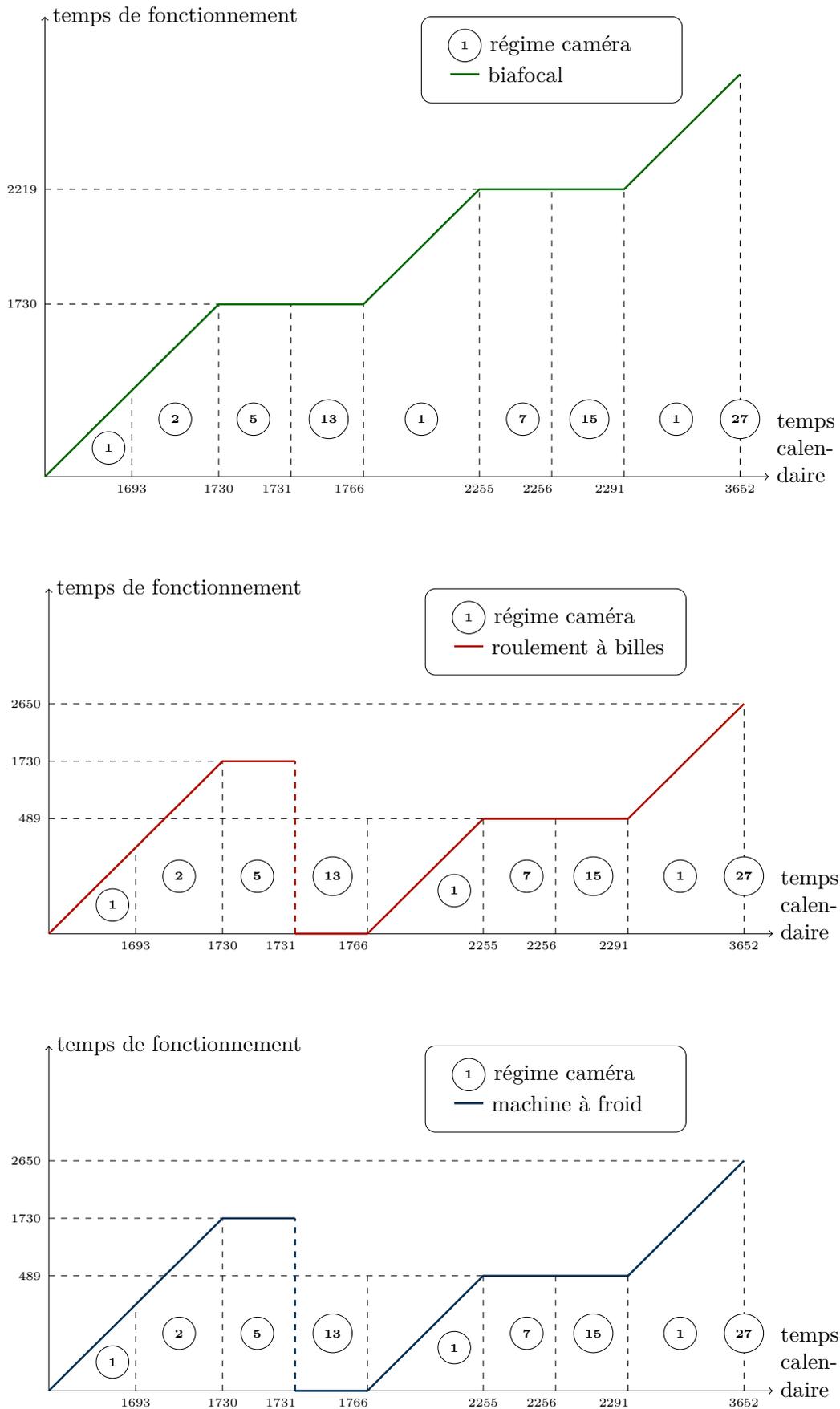


FIGURE 3.2 – Exemple d'évolution des temps de fonctionnement pour une trajectoire partant de l'état initial $X = (1, 0, 0, 0, 0)'$ et contrôlée par la politique d'intervention n° 3.

Caractéristiques des trajectoires

Le nombre de sauts des trajectoires simulées, est compris entre 1 et 31 pour 10^7 simulations de trajectoires partant de l'état neuf et étant contrôlées par la politique d'intervention n°2. L'intervalle dans lequel est compris le nombre de sauts des trajectoires simulées reste le même quand ces dernières suivent la politique d'intervention n°3. Quant aux trajectoires contrôlées par la politique d'intervention n°4, elles ont un nombre de sauts compris entre 1 et 32 mais la grande majorité des trajectoires présente un nombre de sauts plus faible que pour les politiques d'intervention n°2 et n°3. Quelque soit la politique considérée, plus de 95% des trajectoires ont un nombre de sauts inférieur ou égal à 12.

Le tableau 3.3 montre la répartition de ces trajectoires en fonction de leur nombre de sauts pour les politiques d'intervention n°2, n°3 et n°4. On peut constater que les résultats sont très similaires entre les politiques n°2 et n°3. Cela s'explique par le fait que la différence entre ces deux politiques est l'opération réalisée sur les éléments en état dégradé lors d'un passage à l'atelier. Dans le cas de la politique n°2, on le change, tandis qu'on le répare avec la politique n°3. Comme la réparation permet à un élément de repartir dans l'état neuf, il n'y a pas de différence en terme de caractéristiques des trajectoires. Seul leur coût sera différent, comme on le verra au paragraphe 3.3.

Les trajectoires contrôlées par la politique d'intervention n°4 ont généralement un nombre de sauts plus faibles. En effet, lorsque l'électronique est en panne, aucune intervention n'est possible. La trajectoire n'est alors complétée que par un saut dans le cimetière une fois la durée de l'étude atteinte. C'est la raison pour laquelle les trajectoires ayant 2 sauts sont en proportion beaucoup plus importantes dans le cas de la politique n°4 (environ 37%) que dans le cas des politiques n°2 ou n°3 (environ 0.5%).

Les trajectoires sans panne représentent environ 14.4% des trajectoires pour celles suivant les politiques d'intervention n°2, n°3 et n°4 (voir tableau 3.4), ce qui correspond au pourcentage de trajectoires sans panne pour la politique de non-intervention (n°1). Ceci s'explique par le fait que les interventions n'ont aucune conséquence sur les trajectoires ne passant pas par un régime de panne.

Les trajectoires suivant la politique n°2 et n°3 présentent un nombre moyen de pannes de l'ordre de 1.45, soit en moyenne entre 1 et 2 pannes par trajectoires. En revanche, les trajectoires suivant la politique d'intervention n°4 présentent une moyenne de panne plus faible, de l'ordre de 1.12 panne par trajectoire. L'explication est la même que précédemment : l'électronique ne pouvant être changé à neuf, une fois que ce bloc est en panne, la caméra reste en panne jusqu'à la fin de la durée de l'étude. Elle ne peut donc pas être réparée et retombée éventuellement en panne par la suite.

Pour les politiques d'intervention n°2 à n°4, la majorité des pannes correspond à une panne du bloc refroidissement (51.72%), puis vient ensuite la panne du bloc électronique (45.8%). Les autres pannes restent rares (moins de 2.5% des trajectoires au total). Les pourcentages détaillés par régime de panne sont donnés par le tableau 3.5.

3.3 Coûts d'une politique d'intervention

Le coût d'une politique d'intervention, encore appelée stratégie, est défini par la formule (1.4), où $x = (1, 0, 0, 0, 0, 0)$ représentant l'état neuf de la caméra, est le point de départ du processus contrôlé, où f est la fonction du coût courant, définie par la formule (2.13), et où c est la fonction coût des interventions, définie par la formule (2.22).

| nombre de sauts des trajectoires | politique n°2 | politique n°3 | politique n°4 |
|----------------------------------|---------------|---------------|---------------|
| 1 | 13.91% | 13.92% | 13.92% |
| 2 | 0.49% | 0.49% | 37.1% |
| 3 | 0.21% | 0.21% | 0.79% |
| 4 | 25.03% | 25.03% | 8.60% |
| 5 | 16.93% | 16.89% | 17.2% |
| 6 | 0.65% | 0.64% | 8.77% |
| 7 | 10.74% | 10.73% | 0.94% |
| 8 | 13.46% | 13.49% | 3.94% |
| 9 | 5.49% | 5.49% | 5.43% |
| 10 | 2.79% | 2.79% | 1.31% |
| 11 | 3.98% | 3.99% | 0.24% |
| 12 | 2.87% | 2.87% | 0.68% |
| 13 | 1.28% | 1.28% | 0.77% |
| 14 | 0.75% | 0.76% | 0.12% |
| 15 | 0.66% | 0.66% | 0.03% |
| 16 | 0.38% | 0.38% | 0.06% |
| 17 | 0.17% | 0.17% | 0.06% |
| 18 | 0.10% | 0.10% | 0.007% |
| 19 | 0.07% | 0.06% | 0.002% |
| 20 | 0.03% | 0.03% | 0.003% |
| 21 | 0.015% | 0.015% | 0.003% |
| de 22 à 31 | < 0.01% | < 0.01% | < 0.0001% |

TAB. 3.3 – Répartition exprimée en pourcentage, de 10^7 trajectoires simulées pour des caméras partant de l'état neuf suivant les politiques d'intervention n° 2 à n° 4, en fonction de leur nombre de sauts.

| | | | | |
|-----------------------------|---------|---------|---------|---------|
| politique d'intervention n° | 1 | 2 | 3 | 4 |
| nombre moyen pannes | 0.856 | 1.45 | 1.45 | 1.12 |
| % trajectoires sans panne | 14.405% | 14.396% | 14.403% | 14.407% |

TAB. 3.4 – Nombre moyen de pannes par trajectoires et pourcentage de trajectoires sans panne, pour les politiques d'intervention n° 1 à 4.

| régime de panne | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|-----------------|--------|-------|-------|--------|-------|-------|-------|--------|
| politique n°2 | 51.69% | 1.64% | 45.8% | 0.036% | 0.68% | 0.08% | 0.05% | 0.001% |
| politique n°3 | 51.72% | 1.64% | 45.8% | 0.036% | 0.68% | 0.08% | 0.05% | 0.001% |
| politique n°4 | 55.07% | 1.74% | 42.3% | 0.038% | 0.73% | 0.09% | 0.05% | 0.001% |

TAB. 3.5 – Répartition exprimée en pourcentage, des pannes de 10^7 simulations de trajectoire contrôlée à l'aide des politiques d'intervention n° 2 à n° 4, pour une caméra partant de l'état neuf, en fonction du régime de panne.

Calcul du coût d'une trajectoire

Le coût d'une trajectoire \tilde{X}_t contrôlée par une politique d'intervention donnée et partant de l'état neuf de la caméra est donné par :

$$j(\tilde{X}_t) := \int_0^\infty e^{-\alpha s} f(\tilde{X}_s) ds + \sum_{i=1}^\infty e^{-\alpha \tau_i} c(\tilde{X}_{\tau_i-}, \tilde{X}_{\tau_i}). \quad (3.5)$$

Sachant que la fonction f est nulle dès que la caméra est dans l'état cimetièr, on a l'égalité

$$j(\tilde{X}_t) = \int_0^{d_{\text{étude}}} e^{-\alpha s} f(\tilde{X}_s) ds + \sum_{i=1}^\infty e^{-\alpha \tau_i} c(\tilde{X}_{\tau_i-}, \tilde{X}_{\tau_i}). \quad (3.6)$$

On appellera pour la suite :

$$a(\tilde{X}_t) := \int_0^{d_{\text{étude}}} e^{-\alpha s} f(\tilde{X}_s) ds$$

le coût d'indisponibilité de la trajectoire \tilde{X}_t et

$$b(\tilde{X}_t) := \sum_{i=1}^\infty e^{-\alpha \tau_i} c(\tilde{X}_{\tau_i-}, \tilde{X}_{\tau_i})$$

le coût total des interventions de la trajectoire \tilde{X}_t . Cette somme est finie car le nombre d'interventions est borné par la valeur $d_{\text{étude}}/d_{\text{atelier}}$.

Calcul du coût d'une politique d'intervention

Pour une politique d'intervention donnée, on calcule son coût par une méthode Monte-Carlo. On simule un nombre n de trajectoires qui suit la politique donnée, on calcule le coût de chaque trajectoire par la formule (3.6) et on en fait une moyenne.

Pour $\mathbb{U}_\infty = \mathcal{A} \times \mathbb{R}_+^3 \times [0; d_{\text{étude}}[\times\{0\}$, le tableau 3.6 expose les coûts moyens obtenus par cette méthode avec 10^7 trajectoires simulées, pour les quatre politiques de référence. On y

| politique d'intervention n° | 1 | 2 | 3 | 4 |
|------------------------------------|----------------|--------------|--------------|----------------|
| coût moyen d'indisponibilité | 19138 | 15.55 | 15.56 | 10865 |
| coût moyen total des interventions | 0 | 9443 | 9388 | 5016 |
| coût moyen associé | 19138 | 9459 | 9403 | 15881 |
| intervalle de confiance | [19135; 19142] | [9457; 9461] | [9401, 9405] | [15877; 15884] |

TAB. 3.6 – *Coût moyen associé aux politiques de référence, obtenu par une méthode de Monte-Carlo avec 10^7 trajectoires simulées pour chaque politique, et où l'ensemble de contrôle est \mathbb{U}_∞ .*

donne également les intervalles de confiance, le coût moyen d'indisponibilité et le coût moyen des interventions.

Les trajectoires suivant la politique d'intervention n°1 ne présentent aucune intervention. Par conséquent le coût moyen total des interventions est nul. De plus, une fois que la caméra est en panne, on paie chaque jour une somme afin de pénaliser l'indisponibilité de la caméra auprès du client. Raison pour laquelle le coût moyen d'indisponibilité des trajectoires suivant la politique d'intervention n°1 est élevé : 19138. Ce coût correspond au coût moyen associé à la politique de non-intervention.

Pour les politiques d'intervention n°2 et n°3, le coût moyen d'indisponibilité de la caméra est similaire. En effet, en terme de coût, la seule différence entre les politiques n°2 et n°3 vient des changements ou réparations des éléments dégradés dans l'atelier. La réparation d'un élément coûtant moins qu'un changement de ce même élément, le coût moyen associé à la politique n°3 (9404) est inférieur au coût moyen associé à la politique d'intervention n°2 (9459).

Pour la politique d'intervention n°4, le bloc électronique ne peut pas être changé quand celui-ci est en panne. Par conséquent, la caméra peut rester pendant de longue période en panne si une panne électronique survient. Ceci est à l'origine du coût moyen d'indisponibilité élevé de 10865. En effet, la panne du bloc électronique (régime 5) correspond à plus de la moitié des pannes rencontrées. Le coût moyen des interventions est également plus faible que pour les politiques d'intervention n°2 et n°3 car le nombre de pannes par trajectoires est plus faible ; de l'ordre de 1.12 pannes par trajectoire pour la politique n°4 et 1.45 pour les politiques n°2 et n°3.

3.4 Choix des grilles de discrétisation pour \mathbb{U}_{fini}

L'ensemble de contrôle regroupe l'ensemble des points où l'on peut envoyer le système lors d'une intervention. L'ensemble de contrôle \mathbb{U}_{fini} a été défini au paragraphe 2.4.1 à partir d'une grille G permettant de discrétiser les temps de fonctionnement des éléments de la caméra (roulement à billes, machine à froid et bifaocal), et d'une grille H permettant de discrétiser le temps calendaire.

De très nombreux choix sont possibles pour les grilles G et H . Cependant, afin de ne pas avoir des temps de simulations trop longs, on souhaite que la taille de ces grilles reste raisonnable. Notre critère de sélection repose sur la différence entre le coût moyen associé à la politique d'intervention n°3 lorsque l'ensemble de contrôle est \mathbb{U}_∞ (paragraphe 3.3) et le coût associé à cette même politique quand l'ensemble de contrôle est \mathbb{U}_{fini} et tel que défini dans le paragraphe 2.4.1. Le choix de la politique d'intervention n°3 repose sur le fait que le coût moyen associé à cette politique est le plus faible parmi les politiques de référence choisies pour notre étude (tableau 3.6).

Le tableau 3.7 regroupe des tests faits pour différents choix de grilles G et H . Dans ce tableau, la valeur de référence pour calculer l'erreur est le coût moyen associé à la politique n°3

| | | | | |
|---------------------------------|--------------------------|------------------|---------------------|--------------------|
| G | {1000; 2000; 3000} jours | {2, 4, 6, 8} ans | {0, 2, ..., 10} ans | {1, 2, ..., 9} ans |
| H | {1, 2, ..., 9} ans | {2, 4, 6, 8} ans | {0, 2, ..., 10} ans | {1, 2, ..., 9} ans |
| cardinal de \mathbb{U}_{fini} | 246 | 286 | 934 | 3426 |
| coût d'indisponibilité | 14.49 | 15.60 | 15.64 | 15.56 |
| coût interventions | 8734.4 | 9398.5 | 9440.6 | 9382.2 |
| coût moyen | 8748.8 | 9414.1 | 9456.2 | 9397.7 |
| intervalle de confiance | [8745; 8752] | [9410; 9418] | [9452; 9460] | [9394; 9401] |
| erreur/valeur de référence | 6.97% | 0.1% | 0.6% | 0.07% |

TAB. 3.7 – Comparaison du coût moyen associé à la politique d'intervention n° 3 et pour différents choix de grilles de discrétisation G et H .

quand l'ensemble de contrôle est \mathbb{U}_∞ . Ce coût de référence est de 9403 (tableau 3.6).

Pour des grilles G et H différentes (deuxième colonne du tableau), l'erreur est d'environ 7%, ce qui est élevé. Cela s'explique par le fait qu'un point $z = (5, 586, 586, 586, 586)'$ va être envoyé sur le point $y = (13, 0, 0, 1000, 1095)'$ de l'ensemble de contrôle. Ceci a pour conséquence d'avancer de plus d'un an le temps calendaire. On aura en moyenne moins de panne sur la caméra car la durée de notre étude est raccourcie, d'où un coût moyen plus faible par rapport au coût moyen associé à la politique quand l'ensemble de contrôle est \mathbb{U}_∞ .

Les colonnes 3 à 5 du tableau regroupent les tests pour des grilles G et H identiques avec plus ou moins de points. On discrétise par années. Ainsi, dans la troisième colonne du tableau, on prend pour G et H une discrétisation tous les deux ans, en omettant le temps 0 et 10 ans, correspondant à la fin de notre étude dans cet exemple. L'ensemble de contrôle contient alors 286 éléments. On obtient dans ce cas une erreur de 0.1%, ce qui est nettement mieux que précédemment.

Le test suivant (quatrième colonne du tableau), pour $G = H = \{0, 2 \text{ ans}, \dots, 10 \text{ ans}\}$, présente une erreur plus grande que le test avec $G = H = \{2 \text{ ans}, 4 \text{ ans}, 6 \text{ ans}, 8 \text{ ans}\}$: de l'ordre de 0.6% contre 0.1%. Ceci s'explique par le fait que la caméra peut-être arrêtée prématurément, quand le plus proche voisin de la grille H du temps calendaire est 10 ans.

Le dernier test, avec des grilles G et H toujours égales mais plus raffinées, permet de voir une diminution de l'erreur par rapport au test avec $G = H = \{2 \text{ ans}, 4 \text{ ans}, 6 \text{ ans}, 8 \text{ ans}\}$. Cependant, le nombre d'éléments de l'ensemble de contrôle est beaucoup plus grand : 3426 éléments contre 286 dans le cas $G = H = \{2 \text{ ans}, 4 \text{ ans}, 6 \text{ ans}, 8 \text{ ans}\}$.

Par conséquent, on privilégiera le choix $G = H = \{2 \text{ ans}, 4 \text{ ans}, 6 \text{ ans}, 8 \text{ ans}\}$ pour l'approximation de la fonction valeur. En effet, ce choix offre un bon compromis entre l'erreur créée par la projection sur l'ensemble de contrôle \mathbb{U}_{fini} et la taille de cet ensemble de contrôle.

3.5 Conclusion

Dans ce chapitre, nous avons simulé des trajectoires du processus contrôlé et nous avons détaillé leurs caractéristiques, telles le nombre moyen de pannes ou encore le nombre de sauts. Le processus $\{\tilde{X}_t\}$ a été contrôlé à l'aide de quatre politiques de référence. La première politique consiste à ne faire aucune intervention : nous avons pu constater que cette politique présentait le coût le plus élevé. Les trois autres politiques (n°2 à n°4) présentent des coûts plus faibles car elles permettent des réparations et évitent ainsi des coûts d'indisponibilité très élevés. En terme

de coût, la meilleure politique reste la n°3, où les opérations de maintenance réalisées prennent le mieux en compte l'état de la caméra au moment de la panne.

D'autres politiques pourraient être testées, afin de trouver une politique de maintenance plus performante que la politique n°3. On pourrait par exemple considérer des politiques mélangeant de la maintenance prévisionnelle avec de la maintenance systématique. Néanmoins, la description et la mise en oeuvre de ces politiques en terme de simulations deviennent plus complexes. Dans ce travail, nous garderons donc les quatre politiques de référence développées dans ce chapitre.

Par ailleurs, les critères (C_1) à (C_3) que nous avons imposés dans notre modèle, peuvent être modifiés pour disposer d'un panel plus large d'interventions. Par exemple, en modifiant le critère (C_1), on pourrait rendre possible des interventions lorsque la caméra est encore dans un état de fonctionnement (que ce soit l'état stable ou l'état dégradé). Les coûts des réparations étant plus faibles que les coûts de changement à neuf des éléments de la caméra, certaines politiques de maintenance pourraient ainsi avoir un coût moyen plus faible que la politique n°3. Cependant le modèle doit être adapté, avec notamment l'introduction de nouveaux régimes. Un exemple de ce type de modèle est proposé au paragraphe 4.4.

Chapitre 4

Approximation numérique de la fonction valeur

Sommaire

| | | |
|------------|--|------------|
| 4.1 | Quantification du PDMP modélisant l'évolution de la caméra | 84 |
| 4.1.1 | Recherche du plus proche voisin de X_i^{s+1} dans la grille Γ_i^s | 84 |
| 4.1.2 | Initialisation des grilles de quantification et des compteurs associés | 85 |
| 4.1.3 | Choix des variables S , $N_{quantif}$, N et de la suite (δ_s) | 86 |
| 4.2 | Mise en place de la méthode d'approximation de la fonction valeur | 88 |
| 4.2.1 | Horizon de contrôle N_0 | 89 |
| 4.2.2 | Fonction g | 89 |
| 4.2.3 | Fonction F | 90 |
| 4.2.4 | Discretisation du temps | 90 |
| 4.3 | Résultats autour de l'approximation de la fonction valeur | 91 |
| 4.3.1 | Choix de la constante de pénalisation | 91 |
| 4.3.2 | Discretisation du temps | 92 |
| 4.3.3 | Influence de la quantification | 94 |
| 4.3.4 | Approximation de la fonction valeur pour différents ensembles de contrôle | 97 |
| 4.3.5 | Approximation de la fonction valeur pour différents coûts d'intervention | 98 |
| 4.3.6 | Évolution de l'horizon de contrôle | 100 |
| 4.4 | Nouveau modèle | 102 |
| 4.4.1 | Définition du modèle | 102 |
| 4.4.2 | Politiques d'intervention | 106 |
| 4.4.3 | Approximation de la fonction valeur | 108 |
| 4.5 | Bilan | 110 |

On souhaite optimiser les coûts de maintenance de la caméra Sophie, en tenant également compte du coût d'indisponibilité de la caméra auprès du client. De multiples politiques d'interventions sont possibles. Quelques exemples de politiques sont donnés au paragraphe 3.1. Pour déterminer si une politique d'intervention est meilleure ou non qu'une autre, on calcule le coût moyen associé aux trajectoires des caméras contrôlées par la première politique d'intervention. On le compare ensuite au coût moyen associé aux trajectoires des caméras contrôlées par la seconde politique d'intervention. Si le premier coût est inférieur au second, alors la première politique d'intervention est meilleure par rapport à la seconde. L'infimum du coût moyen pris parmi toutes les politiques d'interventions admissibles, correspond à la fonction valeur du

problème de contrôle impulsif (voir paragraphe 1.6.2). Notre but est de trouver une approximation de cette fonction.

On s'intéresse à l'approximation numérique de la fonction valeur développée dans de Saporta *et al.* (2015) et présentée au paragraphe 1.8. Les auteurs proposent une approche basée sur la quantification de la chaîne sous-jacente du processus de Markov déterministe par morceaux et l'utilisation d'une grille de temps adaptée pour la discrétisation de l'opérateur J défini par (1.7).

On commence par exposer au paragraphe 4.1 la quantification du PDMP modélisant l'évolution de la caméra Sophie ; on y détaille notamment les précautions particulières à prendre vis-à-vis de ce type de processus. Dans la partie 4.2, la mise en œuvre de l'algorithme dans le cadre de l'application à l'optimisation de la maintenance de la caméra Sophie est présentée. Au paragraphe 4.3, on expose les résultats obtenus et l'influence du choix des différents paramètres. Une extension du modèle initial (défini au chapitre 2) est explicité au paragraphe 4.4. On y donne les résultats d'approximation numérique de la fonction valeur pour ce modèle. Un bilan de l'utilisation de la méthode numérique d'approximation de la fonction valeur dans le cadre de l'optimisation de la caméra Sophie est donné au paragraphe 4.5.

4.1 Quantification du PDMP modélisant l'évolution de la caméra

Afin de mettre en place la méthode d'approximation de la fonction valeur, il est nécessaire de quantifier les trajectoires du PDMP $\{X_t\}$ modélisant l'évolution d'une caméra, partant d'un état donné, comme l'état neuf, et sur lesquelles aucune intervention n'est réalisée. L'algorithme permettant cette quantification est celui décrit au paragraphe 1.7.2. Cependant, quelques adaptations sont nécessaires pour prendre en compte la spécificité des processus de Markov déterministes par morceaux, en particulier la présence de la variable discrète (régime).

4.1.1 Recherche du plus proche voisin de X_i^{s+1} dans la grille Γ_i^s

Dans l'algorithme de quantification d'une chaîne de Markov, l'étape 2.(a) correspond à la recherche du plus proche voisin de X_i^{s+1} dans la grille Γ_i^s , pour $s \in \{1, \dots, S-1\}$ et $i \in \{1, \dots, N_{quantif}\}$ fixés. Classiquement, pour une chaîne de Markov, cette étape se fait par le biais du calcul de la distance au sens de la norme L^2 de X_i^{s+1} à chaque point de la grille $\Gamma_i^s = \{\gamma_i^{s,1}, \dots, \gamma_i^{s,N}\}$. On cherche ensuite le point de la grille qui réalise le minimum de la distance à X_i^{s+1} .

Cette méthode pose problème pour les processus de Markov déterministes par morceaux. En effet, une des caractéristiques des PDMP est la variable discrète, représentant le régime. Intuitivement, on ne souhaite pas comparer des points de l'espace d'état si ces derniers n'ont pas le même régime. Par conséquent, on procède de la manière suivante : on identifie le régime de X_i^{s+1} , que l'on note m . On sélectionne ensuite tous les points de la grille Γ_i^s ayant pour régime m . Le plus proche voisin de X_i^{s+1} est alors cherché parmi les points précédemment sélectionnés.

Ainsi, on s'assure que le plus proche voisin de X_i^{s+1} possède le même régime que le point X_i^{s+1} . Cependant, afin que cette méthode puisse être mise en place, il est nécessaire qu'il y ait au moins un point dans la grille Γ_i^s ayant pour régime m . Dans le cas contraire, on recherche un voisin dans un ensemble vide, ce qui pose évidemment problème.

Afin de s'assurer la présence de tous les régimes possibles au rang $i \in \{1, \dots, N_{quantif}\}$, il convient de choisir judicieusement les grilles initiales de quantification. C'est l'objet du paragraphe suivant.

4.1.2 Initialisation des grilles de quantification et des compteurs associés

Compte tenu de la variable discrète du PDMP, l'initialisation des grilles de quantification requière un travail conséquent. En effet, chaque régime possible du PDMP doit être présent au moins une fois dans les grilles de quantification afin de garantir un plus proche voisin, comme on a pu le voir au paragraphe 4.1.1. Or certains régimes sont très rares dans notre cas car représentent moins de 0.1% des trajectoires, comme l'indique le tableau 3.2. Afin de garantir la présence de tous les régimes possibles à chaque instant $i \in \{1, \dots, N_{quantif}\}$ dans la grille de quantification Γ_i , on procède de la manière suivante.

Étape 1 : simulations des trajectoires

1. On simule $S + N$ trajectoires du PDMP, sans intervention, partant du point X_0 , où on rappelle que S est le nombre de réplifications données en entrée de l'algorithme de quantification d'une chaîne de Markov (voir paragraphe 1.7.2) et N est le nombre de points dans les grilles $(\Gamma_n^0)_{1 \leq n \leq N_{quantif}}$. On stocke ces simulations dans l'ensemble Ξ .
2. On complète si nécessaire ces trajectoires pour qu'elles contiennent toutes $N_{quantif}$ sauts :
 - Si le nombre de sauts de la trajectoire est $N_{quantif}$, alors on garde la trajectoire telle quelle.
 - Si le nombre de sauts de la trajectoire est inférieur strictement à $N_{quantif}$, on complète la trajectoire en rajoutant son dernier point qui correspond à l'état cimetièrre, le nombre de fois nécessaire pour atteindre une trajectoire ayant $N_{quantif}$ sauts.

En effet, les trajectoires du processus $\{X_t\}$ modélisant l'évolution d'une caméra partant de l'état neuf, ont un nombre de sauts variable, compris entre 1 et 5. Si nous choisissons $N_{quantif} = 3$, alors certaines trajectoires doivent être complétées pour atteindre un nombre de sauts de 3.

Étape 2 : sélection des trajectoires pour les grilles initiales

L'initialisation des grilles de quantification se fait à partir d'une sélection des trajectoires simulées précédemment. On note Υ l'ensemble des trajectoires sélectionnées pour le remplissage des grilles initiales de quantification. Cette sélection est réalisée en deux temps.

1. On parcourt les $S + N$ simulations grâce à une boucle sur $s \in \{1, \dots, S + N\}$ et on sélectionne ou non la trajectoire numérotée s , selon les étapes suivantes :
 - (a) On note $(m_0^s, \dots, m_{N_{quantif}}^s)$ la suite des régimes de la trajectoire $(X_0^s, \dots, X_{N_{quantif}}^s)$, c'est-à-dire
$$(m_0^s, \dots, m_{N_{quantif}}^s) = (\Pi_{\mathcal{M}}(X_0^s), \dots, \Pi_{\mathcal{M}}(X_{N_{quantif}}^s)),$$
où $\Pi_{\mathcal{M}}$ est la projection de E sur l'espace \mathcal{M} des régimes.
 - (b) Si la suite $(m_0^s, \dots, m_{N_{quantif}}^s)$ a déjà été rencontrée dans les $s - 1$ premières simulations parcourues, alors on passe à la trajectoire $(X_0^{s+1}, \dots, X_{N_{quantif}}^{s+1})$.
 - (c) Si la suite $(m_0^s, \dots, m_{N_{quantif}}^s)$ n'a pas encore été rencontrée dans les $s - 1$ premières simulations, alors d'une part, on garde la trajectoire $(X_0^s, \dots, X_{N_{quantif}}^s)$ pour l'initialisation des grilles en la stockant dans l'ensemble Υ , et d'autre part, on la retire de l'ensemble des trajectoires simulées Ξ .

À la fin de cette première étape, on a un nombre n de simulations sélectionnées, contenues dans l'ensemble Υ . Chacune d'elles correspond à une suite de régimes possibles et les n simulations sélectionnées contiennent toutes les suites possibles de régimes présentes dans

Ξ .

L'ensemble des trajectoires simulées restantes Ξ est alors de taille $S + N - n$.

2. On sélectionne $N - n$ trajectoires dans Ξ de manière aléatoire (en s'assurant d'avoir pris $N \geq n$). Chacune des trajectoires sélectionnées est enlevée de l'ensemble Ξ et est stockée dans l'ensemble Υ .

À la fin de la procédure de sélection, l'ensemble Ξ contient exactement S simulations. Elles serviront d'entrée à l'algorithme de quantification. L'ensemble Υ contient quant à lui N simulations. L'ensemble de ces simulations est utilisé pour l'initialisation des grilles de quantifications ; plus précisément, pour tout indice $i \in \{1, \dots, N\}$ tel que la trajectoire $X^i = (X_0^i, \dots, X_{N_{quantif}}^i)$ est dans l'ensemble Υ , on a :

$$\forall j \in \{0, \dots, N_{quantif}\}, \quad \gamma_j^{0,i} = X_j^i.$$

Ainsi faite, l'initialisation des grilles de quantification permet la présence d'au moins un point pour chaque régime possible dans chaque grille. Le remplissage des grilles initiales de quantification étant réalisé à partir de trajectoires simulées complètes, on peut initialiser les compteurs de la manière suivante :

- vecteurs compteurs marginaux :

$$\alpha_0^0 = 1 \quad \text{et} \quad \forall n \in \{0, \dots, N_{quantif}\}, \quad \forall i \in \{1, \dots, N\}, \quad \alpha_n^{0,i} = 1,$$

- matrices compteurs de transition :

$$\forall k \in \{0, \dots, N_{quantif} - 1\}, \quad \forall i, j \in \{1, \dots, N\}, \quad \beta_n^{0,(i,j)} = \mathbb{1}_{\{i=j\}}.$$

Il est important de ne pas initialiser les compteurs α et β à 0. En effet, l'initialisation à 0 présente un inconvénient majeur dans notre cas. En raison de la présence de trajectoires très rares, l'algorithme de quantification peut construire des matrices de transitions ayant des éléments non définis. En effet, malgré un nombre de réplifications s grand, certains régimes peuvent ne pas être visités. Par conséquent, les compteurs α_i^k des points ayant ces régimes dans les grilles correspondantes resteront égaux à 0. Il ne sera alors plus possible de calculer la probabilité de transition car on aura une division par 0.

4.1.3 Choix des variables S , $N_{quantif}$, N et de la suite (δ_s)

Dans le cadre de la quantification des trajectoires du PDMP $\{X_t\}$ n'ayant aucune intervention, on définit la suite de pas $(\delta_s)_{s \geq 1}$ par :

$$\delta_s = \frac{a}{a + bs}, \quad \text{où } a = 4N_{quantif}^{\frac{1}{d}}, \quad b = \pi^2 N_{quantif}^{-\frac{2}{d}},$$

où d est la dimension du processus $\{X_t\}$. Pour les simulations, on prendra $d = 6$ car on ne tient pas compte du temps $t_{atelier}$ (voir paragraphe 3.2.1). Le choix de la suite $(\delta_s)_{s \geq 1}$ repose sur l'article Pagès et Printems (2003).

L'horizon de quantification $N_{quantif}$ sera défini en fonction de l'horizon de contrôle N_0 de l'approximation de la fonction valeur (voir paragraphe 4.2.1). Dans ce paragraphe, nous utilisons l'horizon $N_{quantif} = 4$ car nous quantifions le processus non contrôlé $\{X_t\}$ partant de l'état neuf de la caméra, c'est-à-dire de $X_0 = (1, 0, 0, 0, 0)'$. Or quelle que soit la trajectoire simulée, X_4 se trouve nécessairement dans le régime cimetièrre, alors que X_3 peut se trouver soit dans un régime de panne, soit dans le régime cimetièrre. Ainsi, en quantifiant avec un horizon égal à 4, on quantifie tout le processus non contrôlé partant de l'état neuf.

| | | | | | |
|-------------------------|----------------|----------------|----------------|----------------|----------------|
| N | 10^2 | 10^2 | 10^2 | 10^3 | 10^3 |
| ω | 10^2 | 10^3 | 10^4 | 10^2 | 10^4 |
| $S = N \times \omega$ | 10^4 | 10^5 | 10^6 | 10^5 | 10^7 |
| coût moyen | 19194 | 19100 | 19145 | 19094 | 19125 |
| intervalle de confiance | [19190; 19198] | [19098; 19103] | [19142; 19149] | [19090; 19098] | [19121; 19129] |
| erreur | 0.29% | 0.20% | 0.04% | 0.23% | 0.068% |
| temps quantification | 2s | 20s | 3min30 | 47s | 1h33 |

TAB. 4.1 – Estimation par une méthode de Monte-Carlo, avec $n = 10^7$ simulations, du coût moyen de n trajectoires créées à partir des grilles de quantification $(\Gamma_i)_{0 \leq i \leq 4}$ et de leurs probabilités de transition.

Le nombre N de points dans les grilles de quantification $(\Gamma_i)_{1 \leq i \leq 4}$ doit être supérieur au nombre de suites de régimes possibles pour qu'il n'y ait pas de problème lors de la phase d'initialisation des grilles de quantification. Dans le cas du processus modélisant l'évolution d'une caméra partant de l'état neuf, il y a au maximum 18 suites de régimes possibles. Il est donc nécessaire de prendre $N \geq 18$, ce qui n'est pas contraignant pour notre étude.

Le nombre S de répliques pour la phase de mise à jour de la quantification dépend du nombre de points dans les grilles de quantification. On prend

$$S = N \times \omega,$$

où ω est un facteur multiplicatif.

Le choix de la variable N , représentant le nombre de points dans les grilles de quantification (exceptée la première) a un impact en terme de temps d'exécution des programmes pour l'approximation de la fonction valeur : plus N est grand, plus le temps d'exécution est grand. Cependant, les grilles de quantification étant créées au préalable de l'approximation de la fonction valeur décrite au chapitre 4, le temps d'exécution de la quantification ne sera pas pris en compte dans le choix de la variable N .

Le critère principal que nous avons choisi pour déterminer une valeur intéressante du couple (N, S) est le suivant. Après avoir obtenu des grilles de quantification $(\Gamma_i)_{0 \leq i \leq 4}$ et leurs matrices de transition associées $(M_i)_{0 \leq i \leq 3}$ pour un couple (N, S) fixé, on crée des trajectoires à partir de ces grilles de quantification et de leurs matrices de transition : on part du point $X_0 = (1, 0, 0, 0, 0)$ correspondant à l'état neuf de la caméra, puis on effectue un tirage aléatoire respectant les probabilités de transition données par la matrice M_0 pour déterminer le second point de la trajectoire X_1 . On réitère ce processus en prenant les probabilités de transition issues de la matrice M_i pour déterminer le point X_{i+1} de notre trajectoire, jusqu'à atteindre le point X_4 . Une fois une trajectoire obtenue, nous en calculons le coût. Il est ainsi possible d'estimer le coût moyen de telles trajectoires par une méthode de Monte-Carlo.

Le tableau 4.1 donne les résultats obtenus pour des méthodes de Monte-Carlo effectuées avec 10^7 trajectoires créées à partir des grilles de quantification, pour différentes valeurs du couple (N, S) , où on rappelle que N est le nombre de points dans les grilles de quantification (exceptée la première grille qui ne contient qu'un point) et S est le nombre de simulations utilisées pour créer les grilles de quantification. L'erreur exposée dans le tableau représente le pourcentage de la différence du coût moyen des trajectoires obtenues à partir des grilles et le coût moyen des trajectoires simulées, qui est de 19138 (voir tableau 3.6).

On constate dans le tableau 4.1 que si le nombre N de points dans les grilles de quantification est fixé, le coût moyen pour les trajectoires créées à partir des grilles de quantification tend vers le

| | | |
|------------------------|---------|--------|
| N | 10^2 | 10^3 |
| ω | 10^4 | 10^4 |
| S | 10^6 | 10^7 |
| n | 50 | 50 |
| valeur moyenne du coût | 19127 | 19127 |
| erreur moyenne | 0.055 % | 0.058% |
| écart-type | 87.5 | 11.1 |

TAB. 4.2 – Estimation par une méthode de Monte-Carlo de la valeur moyenne du coût : à partir de n séries de grilles de quantification simulées, on estime le coût moyen de 10^7 trajectoires créées à l'aide de ces grilles de quantification et on effectue une moyenne des n coûts moyens obtenus.

coût moyen 19138, quand le nombre de réplifications S augmente. De plus, à facteur multiplicatif ω fixé, plus le nombre de points des grilles de quantification N est grand, meilleur est le résultat : pour $\omega = 10^2$ par exemple, l'erreur passe de 0.29% pour $N = 10^2$ à 0.23% pour $N = 10^3$.

Cependant, il est important de noter que, à combinaison (N, S) fixée, le calcul du coût dépend de la série de grilles de quantification utilisée. Or l'algorithme permettant de créer la série de grilles de quantification présente une variance. En effet, à paramètres (N, S) fixés, l'algorithme présenté dans la partie 1.7.2 ne donne pas nécessairement la même série de grilles de quantification. On estime la variance due à cet algorithme par une méthode de Monte-Carlo. On répète $n = 50$ fois le processus suivant : on simule une série de grilles de quantification et leurs matrices associées pour un couple (N, S) fixé. Puis on simule 10^7 trajectoires créées à partir de cette série de grille de quantification et on calcule le coût moyen de ces 10^7 trajectoires. On effectue ensuite la moyenne des 50 résultats obtenus, ce qui nous donne la valeur moyenne du coût, exposée dans le tableau 4.2. L'erreur moyenne par rapport au coût moyen des trajectoires simulées normalement (19138) et l'écart-type sont également donnés dans ce tableau. Cette méthode d'estimation a été réalisée pour deux couples : $(N, S) = (10^2, 10^6)$ et $(N, S) = (10^3, 10^7)$. Dans les deux cas, l'erreur entre la valeur moyenne du coût et la valeur de référence du coût (19138) reste inférieure à 0.1% et l'écart-type est raisonnable car représente une erreur de plus ou moins 0.46% pour $(N, S) = (10^2, 10^6)$ et 0.06% pour $(N, S) = (10^3, 10^7)$.

En ayant à l'idée que le temps d'exécution des programmes d'approximation de la fonction valeur décrit au paragraphe 1.8.2 est d'autant plus grand que le nombre N de points dans les grilles de quantification est grand, nous pouvons en conclure que le choix du couple (N, S) égal à $(10^2, 10^6)$ paraît raisonnable. Néanmoins, pour une meilleure stabilité du coût par rapport à la quantification, il peut être intéressant d'augmenter le nombre de points dans les grilles, tout en gardant un facteur de réplifications ω égal à 10^4 , soit le couple $(N, S) = (10^3, 10^7)$.

4.2 Mise en place de la méthode d'approximation de la fonction valeur

On souhaite appliquer la méthode d'approximation numérique de la fonction valeur pour notre problème d'optimisation de maintenance de la caméra Sophie. La méthode nous im-

pose d'avoir une distribution initiale δ_{x_0} , où x_0 est un point de l'espace d'état E , pour le processus $\{X_t\}$ modélisant les évolutions de la caméra (sans maintenance). On choisira donc $x_0 = (1, 0, 0, 0, 0)$. Ainsi, on s'intéressera uniquement à l'optimisation de la maintenance des caméras partant de l'état neuf.

4.2.1 Horizon de contrôle N_0

Le choix de l'horizon de contrôle N_0 peut être fait comme le suggèrent O.L.V. Costa et M.H.A. Davis dans [Costa et Davis \(1989\)](#). Nous avons vu, dans le [tableau 3.3](#), que les trajectoires suivant les politiques de référence présentaient moins de 32 sauts (dont certains étaient des interventions) sur 10^7 simulations de trajectoires réalisées. Il pourrait être naturel de prendre un horizon de contrôle du même ordre, ou légèrement plus faible, par exemple $N_0 = 30$. Cependant, il n'est pas certain que cette valeur soit suffisante : nous nous basons uniquement sur quatre politiques alors qu'il en existe beaucoup d'autres.

Étant donné que ce travail constitue la mise en place de l'approximation numérique de la fonction valeur proposée dans [de Saporta et Dufour \(2012\)](#), nous avons choisi de prendre dans un premier temps un horizon N_0 égal à 5. La principale raison concerne le temps d'exécution des programmes. Afin de pouvoir tester l'influence de différents paramètres, il est plus commode d'avoir des programmes dont les temps d'exécution sont raisonnables. De plus, cet horizon de contrôle permet d'intervenir une à deux fois sur le processus au cours de sa vie. Par conséquent, les opérations de maintenance sont possibles et la fonction valeur devrait être bien inférieure au coût de la politique de contrôle sans maintenance (et sans horizon de contrôle).

Ainsi, la quantification de la chaîne (Θ_n) s'effectuera sur un horizon $N_{quantif} = 5$ et les chaînes sous-jacentes (Θ_n^i) , $1 \leq i \leq u$, associées aux processus $\{X_t\}$ ayant pour loi initiale δ_{y^i} , seront quantifiées sur un horizon $N_{quantif} = N_0 - 1$, c'est-à-dire sur un horizon égal à 4.

Pour l'approximation des fonctions valeurs, la suite (\tilde{v}_n) sera définie pour $1 \leq n \leq N_0 = 5$. Dans le cadre de l'approximation de la fonction valeur $\hat{v}_0(x_0)$ correspondant à l'algorithme [2](#), on construira la matrice \hat{V} de taille $6 \times N_4^\Theta$, qui correspondra à la suite (\hat{v}_n) pour $0 \leq n \leq N_0 = 5$ (définition [1.8.3](#)).

L'horizon de contrôle sera pris égal à 5 dans tout le [paragraphe 4.3](#), excepté dans la partie [4.3.6](#) où on étudiera l'évolution de la fonction valeur quand l'horizon de contrôle croît.

4.2.2 Fonction g

La fonction g sert à l'initialisation des fonctions valeurs. Elle a une influence directe sur l'approximation de la fonction valeur de notre problème. Dans l'idéal, on choisit $g = h$, où h est la fonction du coût moyen d'une trajectoire sans intervention, à savoir

$$\forall z \in E, \quad h(z) = \mathbb{E}_z \left[\int_0^\infty e^{-\alpha s} f(X_s) ds \right].$$

Cependant, dans notre cas, il n'est pas possible de calculer la fonction h pour tout point z de l'espace d'état E . On choisit donc une fonction g supérieure ou égale à h et satisfaisant une certaine régularité ($g \in \mathbf{L}_\Phi(E)$).

On pose $z = (m, \zeta_{rb}, \zeta_{maf}, \zeta_{bf}, \zeta_{cal}, \zeta_{saut})$ un point de l'espace d'état. Étant donné que le coût d'indisponibilité f défini par [\(2.13\)](#) est nul une fois l'état cimetière atteint, on a

$$h(z) = \mathbb{E}_z \left[\int_0^{d_{étude} - \zeta_{cal}} e^{-\alpha s} f(X_s) ds \right].$$

Par ailleurs, la fonction f peut être majorée par une constante :

$$\begin{aligned} h(z) &\leq \gamma \times \mathbb{E}_z \left[\int_0^{d_{\text{étude}} - \zeta_{\text{cal}}} e^{-\alpha s} ds \right] \\ &\leq \gamma \times \mathbb{E}_z \left[\int_0^{d_{\text{étude}} - \zeta_{\text{cal}}} ds \right] \\ &\leq \gamma \times (d_{\text{étude}} - \zeta_{\text{cal}}). \end{aligned}$$

On décide de poser :

$$g(z) = \gamma \times (d_{\text{étude}} - \zeta_{\text{cal}}). \quad (4.1)$$

4.2.3 Fonction F

On peut calculer explicitement la fonction F , définie par la formule (1.6). On la rappelle ici :

$$\forall z \in E, \forall t \geq 0, \quad F(z, t) = \int_0^{t \wedge t^*(z)} e^{-\alpha s - \Lambda(z, s)} f(\Phi(z, s)) ds,$$

où la fonction Λ est définie par :

$$\Lambda(z, t) = \int_0^t \lambda(m, \Phi(z, u)) du.$$

Étant donnée la définition (2.13) de la fonction f du coût d'indisponibilité, on sait que $F(x, t) = 0$ si le régime du point z n'est pas un régime de panne. Il suffit donc de calculer F quand le régime de z est un régime de panne. Dans ce cas, de par sa définition (équation (2.4)), le taux de saut λ est nul. La fonction Λ est donc également nulle. Par conséquent, on a pour tout $z = (m, \zeta) \in \mathcal{P} \times \mathbb{R}_+^3 \times [0; d_{\text{étude}}[\times \mathbb{R}_+$:

$$\begin{aligned} F(z, t) &= \gamma \int_0^{t \wedge t^*(z)} e^{-\alpha s} ds \\ &= \frac{\gamma}{\alpha} (1 - e^{-\alpha(t \wedge t^*(z))}). \end{aligned}$$

En pratique, dans le cadre de l'approximation de la fonction valeur, on utilise la fonction F :

- dans le calcul de l'opérateur K théorique, pour $t = t^*(z)$,
- dans le calcul de l'opérateur J théorique, pour t appartenant à la grille de discrétisation $G(z)$, qui est telle que $\max(G(z)) < t^*(z)$.

Ainsi, pour l'implémentation de la fonction F , on utilisera la forme simplifiée suivante

$$F(z, t) = \begin{cases} \frac{\gamma}{\alpha} (1 - e^{-\alpha t}) & \text{si le régime de } z \text{ appartient à } \mathcal{P} \\ 0 & \text{sinon.} \end{cases}$$

4.2.4 Discrétisation du temps

On rappelle que la discrétisation du temps se fait au moyen de grilles adaptées de temps (définition 1.8.1) :

$$G(z) = \{i\Delta(z), i \in \{0, \dots, n(z)\}\}, \quad z \in E,$$

où $\Delta(z)$ est un pas de temps dans l'intervalle $]0; t^*(z)[$ et $n(z) = \lfloor \frac{t^*(z)}{\Delta(z)} \rfloor - 1$. Le choix de la fonction Δ influe directement sur l'approximation de la fonction valeur. En effet, la grille $G(z)$ détermine les temps potentiels d'intervention. Plus la grille $G(z)$ est raffinée, plus le nombre de possibilités d'interventions est grand. On définit les fonctions Δ_k , pour $k \in \mathbb{N}$, qui nous serviront de référence par la suite. Nous détaillerons dans la partie 4.3.2 leur influence dans les résultats d'approximation de la fonction valeur de notre problème.

Discrétisation 0

L'unité de temps choisie étant la journée, la première discrétisation est celle où le pas est de un jour quelque soit le point z de l'espace d'état E . On définit Δ_0 par

$$\forall z \in E, \quad \Delta_0(z) = 1.$$

Ainsi, les grilles de discrétisation sont les suivantes :

$$\forall z \in E, \quad G_0(z) = \begin{cases} \{0\} & \text{si } t^*(z) < 2, \\ \{0, 1, \dots, \lfloor t^*(z) \rfloor - 1\} & \text{sinon.} \end{cases}$$

Discrétisation 1

Pour cette discrétisation, on différencie le pas en fonction de l'ensemble de régimes auquel appartient le point z de l'espace d'état E . Si le régime de z est un régime de panne, alors on choisit un pas de un jour. Dans tous les autres cas, on choisit un pas de $\frac{t^*(z)}{2}$. L'idée de cette discrétisation vient du fait que dans notre modèle, les interventions ne sont possibles que dans les régimes de panne. Par conséquent, pour les autres régimes, les coûts d'intervention seront très fortement pénalisés, ce qui impliquera des valeurs pour l'opérateur J bien plus grandes que la valeur de l'opérateur K pour les points z ayant un régime différent de celui de panne. La discrétisation de l'intervalle $[0; t^*(z)]$ n'aura donc a priori pas d'influence sur le résultat. En revanche, elle aura une influence sur le temps d'exécution des algorithmes 1 et 2. En choisissant un pas de temps de $\frac{t^*(z)}{2}$, on minimise les temps de calculs. Cette idée sera vérifiée numériquement au paragraphe 4.3.2.

$$\forall z = (m, \zeta) \in E, \quad \Delta_1(z) = \begin{cases} \frac{t^*(z)}{2} & \text{si } m \notin \mathcal{P} \\ 1 & \text{sinon.} \end{cases}$$

Discrétisation k , où $k \in \mathbb{N}$ et $k \geq 2$

Plus généralement, on appellera la discrétisation k , pour $k \in \mathbb{N}$ et $k \geq 2$ la discrétisation avec le pas de temps $\Delta_k(z)$ égal à k jours quand le régime de z est un régime de panne, et égal à $\frac{t^*(z)}{2}$ sinon.

$$\forall z = (m, \zeta) \in E, \quad \Delta_k(z) = \begin{cases} \frac{t^*(z)}{2} & \text{si } m \notin \mathcal{P} \\ k & \text{sinon.} \end{cases}$$

Les grilles de discrétisation associées sont :

$$\forall z \in E, \quad G_k(z) = \begin{cases} \{0\} & \text{si } m \in \mathcal{P} \text{ et si } t^*(z) < 2k, \\ \{0, k, \dots, k * (\lfloor t^*(z) \rfloor - 1)\} & \text{si } m \in \mathcal{P} \text{ et si } t^*(z) \geq 2k, \\ \{0, \frac{t^*(z)}{2}\} & \text{si } m \notin \mathcal{P}. \end{cases}$$

4.3 Résultats autour de l'approximation de la fonction valeur

4.3.1 Choix de la constante de pénalisation

La constante de pénalisation p , introduite au paragraphe 2.4.2, permet d'empêcher certaines interventions de se produire. Elle doit être choisie très largement supérieure au coût d'une trajectoire non contrôlée, afin d'être certain que les interventions non souhaitées ne se produisent

| p | 10^{10} | 10^{30} | NaN |
|------------------------|-----------|-----------|-------|
| $\hat{v}_0(\hat{Z}_0)$ | 11445 | 11445 | 11445 |

TAB. 4.3 – Comparaison de l’approximation de la fonction valeur $\hat{v}_0(\hat{Z}_0)$ pour différentes constantes de pénalisation.

jamais. Une approximation par méthode de Monte-Carlo du coût moyen associé à la politique sans intervention (politique n°1) est de 19138, d’après le tableau 3.6.

Afin de valider le choix de p , on calcule l’approximation de la fonction valeur dans trois cas. Les trois simulations se font avec les paramètres suivants fixés :

- L’ensemble \mathbb{U}_{fini} est défini à partir des grilles $G = H = \{2 \text{ ans}, 4 \text{ ans}, 6 \text{ ans}, 8 \text{ ans}\}$, soit un ensemble \mathbb{U}_{fini} à 286 éléments.
- Les grilles de quantification $(\Gamma_n^{i,\Theta})_{0 \leq n \leq N_0-1}$ et leurs matrices associées $(M_n^i)_{0 \leq n \leq N_0-2}$ sont simulées une seule fois avec $N_{\mathbb{U}} = 100$ points dans chaque grille $\Gamma_n^{i,\Theta}$, $1 \leq n \leq N_0 - 1$ et $S_{\mathbb{U}} = 500N_{\mathbb{U}}$ répliques.
- Les grilles de quantification $(\Gamma_n^{\Theta})_{0 \leq n \leq N_0}$ et leurs matrices associées $(M_n)_{0 \leq n \leq N_0-1}$ sont simulées une seule fois avec $N_{x_0} = 10^3$ points dans chaque grille Γ_n^{Θ} , $1 \leq n \leq N_0$ et $S_{x_0} = 10^3N_{x_0}$ répliques.
- La discrétisation du temps est celle correspondant au pas de temps Δ_1 défini au paragraphe 4.2.4.
- La fonction coût est celle définie par la formule (2.24). On autorise ainsi toutes les politiques d’interventions répondant aux critères (\mathcal{C}_1) à (\mathcal{C}_3) .

Seule la variable p change de valeurs. Le tableau 4.3 regroupe les résultats obtenus. Pour $p = 10^{10}$ et $p = 10^{30}$, on obtient la même fonction valeur, ce qui nous indique qu’a priori la valeur $p = 10^{10}$ est suffisante.

Afin de le confirmer, on exécute les algorithmes d’approximation de la fonction valeur en définissant la variable p comme un NaN (*not a number*). Les NaN sont des éléments que l’on peut voir comme des éléments absorbants : toute addition, multiplication ou fonction appliquée à un NaN renvoie un NaN. On est ainsi certain que les interventions que l’on ne souhaite pas avoir donneront des NaN et ne pourront être choisies par le programme. Le résultat de l’approximation de la fonction valeur obtenu pour $p = \text{NaN}$ étant identique aux deux autres ($p = 10^{10}$ et $p = 10^{30}$), cela confirme le fait que le choix de la constante de pénalisation égale à 10^{10} est suffisant pour notre application. Désormais, p sera fixée à 10^{10} pour toutes les simulations.

4.3.2 Discrétisation du temps

Dans cette partie, on s’intéresse au comportement de l’approximation de la fonction valeur pour différentes discrétisations de l’intervalle $[0; t^*(z)]$ définies au paragraphe 4.2.4. La simulation de l’approximation $\hat{v}_0(\hat{Z}_0)$ pour chaque discrétisation se fait avec un certain nombre de paramètres fixés :

- L’ensemble \mathbb{U}_{fini} est défini à partir des grilles $G = H = \{2 \text{ ans}, 4 \text{ ans}, 6 \text{ ans}, 8 \text{ ans}\}$, soit une discrétisation par pas de deux ans pour les temps de fonctionnements des éléments de la caméra et pour le temps calendaire. L’ensemble \mathbb{U}_{fini} a alors 286 éléments.

| pas de discrétisation | $t_{approx,U}$ | t_{approx,x_0} | $\hat{v}_0(\hat{Z}_0)$ |
|-----------------------|----------------|------------------|------------------------|
| Δ_0 | 1h13 | 3h06 | 11445 |
| Δ_1 | 25min | 1h17 | 11445 |
| Δ_2 | 6min10 | 2min57 | 11447 |
| Δ_5 | 1min24 | 16s | 11455 |
| Δ_{10} | 43s | 4s | 11648 |
| Δ_{25} | 32s | 1s | 12007 |
| Δ_{50} | 30s | 0.5s | 12149 |
| Δ_{100} | 30s | 0.4s | 12231 |

TAB. 4.4 – Approximation de la fonction valeur, à paramètres et grilles de quantification fixes, pour différentes discrétisations définies au paragraphe 4.2.4.

- Les grilles de quantification $(\Gamma_n^{i,\Theta})_{0 \leq n \leq N_0-1}$ et leurs matrices associées $(M_n^i)_{0 \leq n \leq N_0-2}$ sont simulées une seule fois avec $N_U = 100$ points dans chaque grille $\Gamma_n^{i,\Theta}$, $1 \leq n \leq N_0 - 1$ et $S_U = 500N_U$ répliques.
- Les grilles de quantification $(\Gamma_n^\Theta)_{0 \leq n \leq N_0}$ et leurs matrices associées $(M_n)_{0 \leq n \leq N_0-1}$ sont simulées une seule fois avec $N_{x_0} = 10^3$ points dans chaque grille Γ_n^Θ , $1 \leq n \leq N_0$ et $S_{x_0} = 10^3 N_{x_0}$ répliques.
- La fonction coût est celle définie par la formule (2.24), où toutes les politiques d'interventions répondant aux critères (C_1) à (C_3) sont autorisées.

On note $t_{approx,U}$ le temps d'exécution de l'algorithme 1 et t_{approx,x_0} celui de l'algorithme 2. Le tableau 4.4 récapitule les résultats obtenus d'approximation de la fonction valeur pour les différentes discrétisations. La figure 4.1 montre l'évolution de l'approximation de la fonction valeur $\hat{v}_0(\hat{Z}_0)$ en fonction du pas de discrétisation choisi.

On constate en premier lieu que l'approximation de la fonction valeur $\hat{v}_0(\hat{Z}_0)$ est la même pour la discrétisation 0 et pour la discrétisation 1. Ceci nous confirme que le pas de discrétisation $\Delta(z)$ choisi dans le cas d'un point z ayant un régime autre qu'un régime atelier n'influe pas sur le résultat. En revanche, les temps d'exécution des deux parties de l'algorithme sont nettement plus importants dans le cas de la discrétisation 0 : environ 4h20 contre 1h45 environ dans le cas de la discrétisation avec le pas de temps Δ_1 . Il y a donc tout intérêt à distinguer le pas de discrétisation pour les points ayant un régime de panne, du pas de discrétisation pour les autres points, ce dernier pouvant être très grossier.

L'évolution de l'approximation de la fonction valeur $\hat{v}_0(\hat{Z}_0)$ en fonction du pas Δ_k , $1 \leq k \leq 70$ choisi, est montrée par la figure 4.1. On peut voir que la valeur $\hat{v}_0(\hat{Z}_0)$ est globalement décroissante quand n tend vers 1, c'est-à-dire quand le pas de discrétisation est de plus en plus fin. Plusieurs choses sont à noter : pour $51 \leq n \leq 70$, l'approximation de la fonction valeur est constante égale à 12231. Ceci peut s'expliquer par le fait que le pas de discrétisation devient trop grand et qu'une variation de quelques jours dans le choix du pas de discrétisation n'a plus d'influence sur la fonction valeur. La seule possibilité raisonnable d'intervention est alors celle ayant lieu au temps $t = 0$ de la discrétisation.

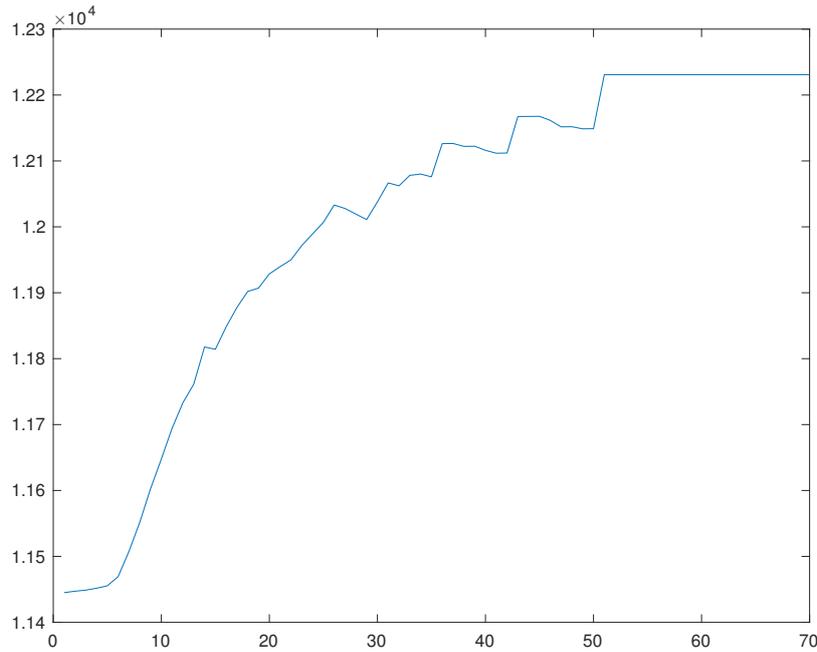


FIGURE 4.1 – Évolution de l’approximation de la fonction valeur $\hat{v}_0(\hat{Z}_0)$ en fonction du numéro k de la discrétisation, pour k allant de 1 à 70.

D’autre part, il est intéressant de constater que l’approximation de la fonction valeur est de moins en moins sensible au fur et à mesure que le pas de discrétisation diminue. Pour des pas de discrétisation de l’ordre de 1 à 5 jours, l’approximation de la fonction valeur évolue peu. En raison des différences de temps d’exécution des programmes donnés par le tableau 4.4, il peut être préférable de choisir la discrétisation Δ_2 , dont le temps d’exécution de l’ordre de 10 minutes est dix fois moins important que celui de la discrétisation Δ_1 . On privilégiera donc par la suite le pas de discrétisation Δ_2 , pour lequel l’approximation de la fonction valeur est sensiblement proche de celle obtenue avec le pas de discrétisation Δ_1 .

4.3.3 Influence de la quantification

Nous avons vu dans les tableaux 4.1 et 4.2 une estimation de l’erreur du coût moyen calculé sur des trajectoires créées à partir des grilles de quantification par rapport au coût moyen des trajectoires simulées normalement. Dans les deux tableaux, nous nous sommes intéressés uniquement à des trajectoires partant de l’état neuf de la caméra.

Néanmoins, la méthode d’approximation de la fonction valeur utilisée dans ce travail nécessite également la quantification des trajectoires partant de tous les points de contrôle, ce qui nous amène à utiliser u séries de N_0 grilles de quantification, en plus de la série de $N_0 + 1$ grilles pour les trajectoires partant de l’état neuf.

Il est donc important d’étudier l’impact de la quantification sur les résultats d’approximation de la fonction valeur. Pour ce faire, nous nous intéressons d’abord à l’influence de la quantification des trajectoires partant de l’ensemble de contrôle, à l’aide de simulations réalisées avec différentes valeurs du nombre N_U de points dans les grilles et différentes valeurs du facteur multiplicatif ω_U qui détermine le nombre de répliques S_U utilisées pour la création des grilles. Les autres paramètres sont fixés tels que :

- L’ensemble \mathbb{U}_{fini} est défini à partir des grilles $G = H = \{2 \text{ ans}, 4 \text{ ans}, 6 \text{ ans}, 8 \text{ ans}\}$, soit

| | | | | | |
|--|--------|-----------------|--------|-----------------|--------|
| $N_{\mathbb{U}}$ | 100 | 100 | 100 | 500 | 1000 |
| $\omega_{\mathbb{U}}$ | 100 | 500 | 10^3 | 100 | 100 |
| $S_{\mathbb{U}} = \omega_{\mathbb{U}}N_{\mathbb{U}}$ | 10^4 | 5×10^4 | 10^5 | 5×10^4 | 10^5 |
| $t_{quantif,\mathbb{U}}$ | 11min | 1h12 | 1h47 | 1h28 | 4h41 |
| $t_{approx,\mathbb{U}}$ | 6min24 | 6min20 | 6min19 | 55min | 5h20 |
| t_{approx,x_0} | 3min13 | 3min19 | 3min17 | 3min14 | 3min13 |
| $\hat{v}_0(\hat{Z}_0)$ | 11457 | 11446 | 11459 | 11462 | 11466 |

TAB. 4.5 – Approximation de la fonction valeur $\hat{v}_0(\hat{Z}_0)$, à paramètres et grilles de quantification $(\Gamma_k^\Theta)_{0 \leq k \leq N_0}$ fixés, pour des grilles variables de quantification $(\Gamma_k^{i,\Theta})_{0 \leq i \leq N_0-1}$ ayant $N_{\mathbb{U}}$ points et créées à partir de $S_{\mathbb{U}}$ réplifications.

un ensemble à 286 éléments.

- Les grilles de quantification $(\Gamma_n^\Theta)_{0 \leq n \leq N_0}$ et leurs matrices associées $(M_n)_{0 \leq n \leq N_0-1}$ sont simulées une seule fois et sont valables pour toutes les simulations. Il y a $N_{x_0} = 10^3$ points dans chaque grille Γ_n^Θ , $1 \leq n \leq N_0$ et $S_{x_0} = 10^3 N_{x_0}$ réplifications sont utilisées pour la création de ces grilles.
- La discrétisation du temps se fait avec le pas de temps Δ_2 défini au paragraphe 4.2.4.
- La fonction coût est celle définie par la formule (2.24) avec pour matrice de coût C . On autorise ainsi toutes les politiques d'interventions répondant aux critères (C_1) à (C_3) .

Les résultats obtenus sont présentés dans le tableau 4.5. Nous pouvons constater que les résultats de l'approximation de la fonction valeur sont assez semblables : ils sont tous compris dans l'intervalle [11446; 11466]. La variabilité de l'erreur due à la quantification ne nous permet pas de comparer un à un ces résultats. En effet, le tableau 4.6 montre que si l'approximation de la fonction valeur est faite $n = 10$ fois de suite pour $N_{\mathbb{U}} = 100$ et $\omega_{\mathbb{U}} = 100$, la moyenne obtenue est de 11460 avec un écart-type de 21.25, représentant environ 0.19% de la valeur moyenne de $\hat{v}_0(\hat{Z}_0)$.

Au regard de l'écart entre les résultats d'approximation de la fonction valeur obtenus dans le tableau 4.5 et de l'écart-type du tableau 4.6, on peut raisonnablement en conclure que l'influence de la quantification des trajectoires partant des points de contrôle sur l'approximation de la fonction valeur n'est pas très importante sur les couples $(N_{\mathbb{U}}, S_{\mathbb{U}})$ testés.

Nous nous intéressons maintenant à l'impact de la quantification des trajectoires partant du point x_0 correspondant à l'état neuf d'une caméra. On fait varier les paramètres N_{x_0} représentant le nombre de points dans les grilles Γ_k^Θ pour $1 \leq k \leq N_0$ et ω_{x_0} le facteur multiplicatif pour le nombre de réplifications S_{x_0} . Les autres paramètres sont fixés tels que :

- L'ensemble \mathbb{U}_{fini} est défini à partir des grilles $G = H = \{2 \text{ ans}, 4 \text{ ans}, 6 \text{ ans}, 8 \text{ ans}\}$. L'ensemble \mathbb{U}_{fini} contient 286 éléments.
- Les grilles de quantification $(\Gamma_n^{i,\Theta})_{0 \leq n \leq N_0-1}$ et leurs matrices associées $(M_n^i)_{0 \leq n \leq N_0-2}$ sont simulées une seule fois avec $N_{\mathbb{U}} = 100$ points dans chaque grille $\Gamma_n^{i,\Theta}$, $1 \leq n \leq N_0 - 1$ et $S_{\mathbb{U}} = 500N_{\mathbb{U}}$ réplifications.

| | |
|--|--------|
| $N_{\mathbb{U}}$ | 100 |
| $\omega_{\mathbb{U}}$ | 100 |
| $S_{\mathbb{U}} = \omega_{\mathbb{U}}N_{\mathbb{U}}$ | 10^4 |
| n | 10 |
| $\hat{v}_0(\hat{Z}_0)$ moyenne | 11460 |
| écart-type | 21.25 |

TAB. 4.6 – Valeur moyenne de l'approximation de la fonction valeur obtenue par une méthode de Monte-Carlo avec un nombre n de réplifications, pour des grilles de quantification des trajectoires partant des points de l'ensemble de contrôle simulées à nouveau à chaque approximation de la fonction valeur.

| | | | | | | |
|---------------------------------|--------|--------|--------|--------|--------|--------|
| N_{x_0} | 100 | 100 | 100 | 1000 | 1000 | 1000 |
| ω_{x_0} | 10^2 | 10^4 | 10^6 | 10^2 | 10^3 | 10^4 |
| $S_{x_0} = \omega_{x_0}N_{x_0}$ | 10^4 | 10^6 | 10^8 | 10^5 | 10^6 | 10^7 |
| $t_{quantif,x_0}$ | 2s | 3min16 | 5h25 | 47s | 9min | 1h34 |
| t_{approx,x_0} | 6s | 6s | 6s | 3min12 | 3min19 | 3min12 |
| $\hat{v}_0(\hat{Z}_0)$ | 11554 | 11439 | 11428 | 11441 | 11446 | 11443 |

TAB. 4.7 – Approximation de la fonction valeur $\hat{v}_0(\hat{Z}_0)$, à paramètres et grilles de quantification $(\Gamma_k^{i,\Theta})_{0 \leq k \leq N_0, 1 \leq i \leq u}$ fixés, pour des grilles variables de quantification $(\Gamma_k^{\Theta})_{0 \leq i \leq N_0}$ ayant N_{x_0} points et créées à partir de S_{x_0} réplifications.

- La discrétisation du temps se fait avec le pas de temps Δ_2 défini au paragraphe 4.2.4.
- La fonction coût est celle définie par la formule (2.24) avec pour matrice de coût C . Toutes les politiques d'interventions répondant aux critères (C_1) à (C_3) sont autorisées.

Le tableau 4.7 présente les résultats obtenus. Nous pouvons constater qu'un nombre faible de points dans les grilles ($N_{x_0} = 100$) couplé à un nombre faible de réplifications ($S_{x_0} = 10^4$) implique une approximation de la fonction valeur supérieure aux autres cas traités dans le tableau, qui ont des approximations de la fonction valeur comprises entre 11428 et 11446. On ne peut pas comparer les résultats suivants obtenus à cause de la variabilité de la quantification.

En revanche, en fixant les paramètres $N_{x_0} = 10^3$ et $\omega_{x_0} = 10^3$, on peut exécuter plusieurs fois l'approximation de la fonction valeur et effectuer une moyenne sur les résultats obtenus et calculer l'écart-type. Le tableau 4.8 montre que l'on obtient une approximation moyenne de la fonction valeur à 11443 avec un écart-type de 10.51, ce qui représente moins de 0.01% de $\hat{v}_0(\hat{Z}_0)$.

On en conclut qu'en évitant un nombre trop faible de points N_{x_0} couplé à un nombre faible de réplifications S_{x_0} , l'impact de la quantification des trajectoires partant de l'état initial x_0 est très modéré dans les résultats d'approximation de la fonction valeur $\hat{v}_0(\hat{Z}_0)$, sur les valeurs des

| | |
|---|--------|
| N_{x_0} | 1000 |
| ω_{x_0} | 10^3 |
| $S_{\mathbb{U}} = \omega_{\mathbb{U}} N_{\mathbb{U}}$ | 10^6 |
| n | 10 |
| $\hat{v}_0(\hat{Z}_0)$ moyenne | 11443 |
| écart-type | 10.51 |

TAB. 4.8 – Valeur moyenne de l'approximation de la fonction valeur obtenue par une méthode de Monte-Carlo avec un nombre n de réplifications pour des grilles de quantification des trajectoires partant de l'état neuf de la caméra simulées à nouveau pour chaque approximation de la fonction valeur.

couples (N_{x_0}, S_{x_0}) testées.

Finalement, la quantification, au travers du choix des nombres N_{x_0} et $N_{\mathbb{U}}$ des points dans les grilles de quantification et des nombres S_{x_0} et $S_{\mathbb{U}}$ de réplifications pour la création des grilles, a peu d'impact sur le résultat d'approximation de la fonction valeur. On peut cependant constater que la quantification des trajectoires partant de l'état neuf de la caméra semble avoir un impact plus important sur le résultat d'approximation de la fonction valeur quand les nombres N_{x_0} et S_{x_0} sont faibles. En conclusion, le choix des couples de paramètres (N_{x_0}, S_{x_0}) et $(N_{\mathbb{U}}, S_{\mathbb{U}})$ égaux respectivement à $(10^3, 10^3)$ et $(10^2, 10^4)$ permet d'avoir une faible influence sur la variabilité de la quantification et par conséquent sur le résultat de l'approximation de la fonction valeur.

4.3.4 Approximation de la fonction valeur pour différents ensembles de contrôle

Dans le cadre de la méthode numérique d'approximation de la fonction valeur, on impose que l'ensemble de contrôle soit fini. Nous avons proposé et discuté au paragraphe 3.4 du choix d'un tel ensemble, selon un critère basé sur la différence du coût des trajectoires suivant que l'ensemble de contrôle était infini ou fini. Le tableau 3.7 récapitule les résultats obtenus. On en avait conclu que le choix des grilles G et H toutes deux égales à l'ensemble $\{2 \text{ ans}, 4 \text{ ans}, 6 \text{ ans}, 8 \text{ ans}\}$ pour construire l'ensemble de contrôle était un bon compromis entre l'erreur induite par la projection sur l'ensemble de contrôle à l'entrée de l'atelier et le cardinal de cet ensemble de contrôle.

On propose dans ce paragraphe d'étudier les résultats d'approximation des fonctions valeurs associées à différents ensembles de contrôle. On choisit de fixer les entrées suivantes des algorithmes d'approximation :

- Pour la quantification des trajectoires partant des points de l'ensemble de contrôle, on choisit $N_{\mathbb{U}} = 100$ et $S_{\mathbb{U}} = 500N_{\mathbb{U}}$ réplifications.
- Les grilles de quantification $(\Gamma_n^{\Theta})_{0 \leq n \leq N_0}$ et leurs matrices associées $(M_n)_{0 \leq n \leq N_0-1}$ sont simulées une seule fois avec $N_{x_0} = 10^3$ points dans chaque grille Γ_n^{Θ} , $1 \leq n \leq N_0$ et $S_{x_0} = 10^3 N_{x_0}$ réplifications.
- La discrétisation du temps est celle correspondant au pas de temps Δ_1 défini au paragraphe 4.2.4.
- La fonction coût est celle définie par la formule (2.24) avec pour matrice de coût C .

| | | | | |
|---------------------------------|--------------------------|------------------|---------------------|--------------------|
| G | {1000; 2000; 3000} jours | {2, 4, 6, 8} ans | {0, 2, ..., 10} ans | {1, 2, ..., 9} ans |
| H | {1, 2, ..., 9} ans | {2, 4, 6, 8} ans | {0, 2, ..., 10} ans | {1, 2, ..., 9} ans |
| cardinal de \mathbb{U}_{fini} | 246 | 286 | 934 | 3426 |
| $t_{quantif, \mathbb{U}}$ | 48min | 1h | 2h45 | 11h36 |
| $t_{approx, \mathbb{U}}$ | 18min | 25min | 1h11 | 9h |
| t_{approx, x_0} | 1h57 | 2h17 | 2h20 | 1h54 |
| $\hat{v}(\hat{Z}_0)$ | 10947 | 11445 | 11520 | 12030 |

TAB. 4.9 – Approximation des fonctions valeurs $\hat{v}(\hat{Z}_0)$, à paramètres et grilles de quantification fixés, pour différents ensembles de contrôle \mathbb{U}_{fini} construit à partir des grilles G et H .

L'ensemble de contrôle \mathbb{U}_{fini} n'étant pas fixe, les grilles de quantification $(\Gamma_n^{i, \Theta})_{0 \leq n \leq N_0 - 1}$ sont refaites à chaque approximation de la fonction valeur. Il faudra donc tenir compte de la variabilité due à cette quantification dans les résultats obtenus.

En notant $t_{quantif, \mathbb{U}}$ le temps d'exécution de la quantification des trajectoires partant des points de l'ensemble de contrôle, $t_{approx, \mathbb{U}}$ le temps d'exécution de l'algorithme 1 et t_{approx, x_0} celui de l'algorithme 2, on construit le tableau 4.9 donnant les résultats de l'approximation des fonctions valeurs associées aux ensembles de contrôle \mathbb{U}_{fini} .

D'une part, on peut constater que plus le cardinal de l'ensemble de contrôle \mathbb{U}_{fini} est grand, plus les temps d'exécution des deux algorithmes de l'approximation de la fonction valeur sont élevés. En effet, plus le cardinal de \mathbb{U}_{fini} est grand, plus il y a de grilles de quantification et d'approximations de fonctions valeurs sur les points de contrôle à calculer.

D'autre part, on peut observer que les approximations des fonctions valeurs respectent un ordre pour les simulations effectuées : plus le cardinal de l'ensemble de contrôle \mathbb{U}_{fini} est grand, plus l'approximation de la fonction valeur associée augmente. On retrouve ce phénomène dans le tableau 3.7 : quand le cardinal de l'ensemble de contrôle \mathbb{U}_{fini} augmente, l'estimation du coût de la politique d'intervention n°3 augmente également, vers la valeur du coût moyen de cette même politique quand l'ensemble de contrôle est \mathbb{U}_∞ .

Pour conclure, comme l'algorithme de l'approximation de la fonction valeur associée à l'ensemble de contrôle \mathbb{U}_{fini} construit à l'aide des grilles $G = H = \{2, 4, 6, 8\}$ ans, a des temps d'exécution raisonnables, nous continuerons d'utiliser cet ensemble de contrôle pour la suite des simulations.

4.3.5 Approximation de la fonction valeur pour différents coûts d'intervention

La fonction valeur dépend des coûts d'intervention choisis. Ces coûts sont entièrement définis par une matrice C indiquant les coûts de passage d'un régime à un autre. Par défaut, la matrice C est définie dans le paragraphe 2.4.2. Cependant elle peut être modifiée. On présente dans le paragraphe 3.1 quatre matrices C_1 à C_4 présentant des coûts différents. On étudie maintenant les résultats d'approximation de la fonction valeur associée à chaque matrice C_i , pour $1 \leq i \leq 4$.

Pour pouvoir effectuer une comparaison pertinente, on choisit de fixer les paramètres suivants :

- L'ensemble \mathbb{U}_{fini} est défini à partir des grilles $G = H = \{2 \text{ ans}, 4 \text{ ans}, 6 \text{ ans}, 8 \text{ ans}\}$, soit un ensemble \mathbb{U}_{fini} à 286 éléments.

| matrice de coût | $\hat{v}_0(\hat{Z}_0)$ | politique | coût moyen avec horizon $N_0 = 5$ |
|-----------------|------------------------|-----------|-----------------------------------|
| C_1 | 19147 | n°1 | 19138 |
| C_2 | 11484 | n°2 | 10589 |
| C_3 | 11455 | n°3 | 10553 |
| C_4 | 15676 | n°4 | 16057 |

TAB. 4.10 – Dans le tableau de gauche, approximation de la fonction valeur $\hat{v}_0(\hat{Z}_0)$ associée à chaque matrice de coûts. Dans celui de droite, coût moyen obtenu par une méthode de Monte-Carlo pour des trajectoires suivant une politique d'interventions de référence, avec horizon de contrôle $N_0 = 5$.

- Les grilles de quantification $(\Gamma_n^{i,\Theta})_{0 \leq n \leq N_0-1}$ et leurs matrices associées $(M_n^i)_{0 \leq n \leq N_0-2}$ sont simulées une seule fois avec $N_U = 100$ points dans chaque grille $\Gamma_n^{i,\Theta}$, $1 \leq n \leq N_0 - 1$ et $S_U = 500N_U$ répliques.
- Les grilles de quantification $(\Gamma_n^\Theta)_{0 \leq n \leq N_0}$ et leurs matrices associées $(M_n)_{0 \leq n \leq N_0-1}$ sont simulées une seule fois avec $N_{x_0} = 10^3$ points dans chaque grille Γ_n^Θ , $1 \leq n \leq N_0$ et $S_{x_0} = 10^3 N_{x_0}$ répliques.
- La discrétisation du temps est celle correspondant au pas de temps Δ_2 défini au paragraphe 4.2.4.

Le tableau 4.10 est un double tableau montrant d'une part les résultats d'approximation de la fonction valeur $\hat{v}_0(\hat{Z}_0)$ associée à chaque matrice de coût C_i , $1 \leq i \leq 4$ et d'autre part, les valeurs des coûts moyens calculés à partir des trajectoires simulées suivant la politique n°1 à n°4, avec horizon de contrôle $N_0 = 5$. On entend par trajectoire avec horizon de contrôle N_0 , une trajectoire pour laquelle on ne peut appliquer la politique de contrôle que durant les cinq premiers sauts du processus. Après ce cinquième saut (s'il a lieu), le processus a exactement la même dynamique que le processus non contrôlé ; il n'y a donc plus de sauts correspondant à des interventions à partir du sixième saut du processus.

Les fonctions valeurs étant calculées avec un horizon de contrôle $N_0 = 5$, elles correspondent également à des trajectoires qui sont telles qu'à partir du sixième saut, il n'y a plus d'interventions possibles. On peut ainsi comparer les fonctions valeurs obtenues avec les coûts moyen des politiques de référence avec horizon de contrôle $N_0 = 5$. En effet, chaque politique n° i fait appel à la matrice de coût C_i .

On devrait donc avoir en théorie une approximation de la fonction valeur, associée à la matrice de coût C_i , qui devrait être inférieure au coût moyen de la politique n° i , avec horizon de contrôle. Cependant, nous pouvons constater que ce n'est pas le cas ici. L'erreur d'approximation de la fonction valeur n'est pas négligeable puisque les approximations des fonctions valeurs sont au-dessus du coût moyen des politiques de référence avec horizon de contrôle, d'environ 8.5% pour les politiques d'intervention n°2 et n°3, et d'environ 2.4% pour la politique d'intervention n°4. Néanmoins, nous pouvons tirer de ces résultats des informations intéressantes.

En premier lieu, la politique n°1, correspondant à une politique sans intervention, est la seule politique d'intervention dont la matrice de coût associée est C_1 . En effet, la matrice C_1 est définie comme étant partout égale à la constante de pénalisation afin de n'avoir aucune intervention possible sur les trajectoires. La fonction valeur devrait donc être théoriquement égale au coût moyen des trajectoires suivant la politique d'intervention n°1. En pratique, nous pouvons voir

que ces deux valeurs sont très proches; il y a moins de 0.05% de différence. Cet écart peut provenir aussi bien de l'erreur de quantification dans l'approximation de la fonction valeur que de l'erreur induite par la méthode de Monte-Carlo pour calculer le coût moyen des trajectoires sans intervention.

Par ailleurs, pour les matrices de coût C_2 à C_4 , nous pouvons constater que les écarts entre les approximations des fonctions valeurs sont du même ordre de grandeur que ceux entre les coûts moyens des politiques d'intervention n°2 à n°4 avec horizon de contrôle. De manière plus détaillée, le coût moyen de la politique n°3 avec horizon de contrôle est le plus faible (10553) et est relativement proche du coût moyen de la politique n°2 avec horizon de contrôle (10589). En revanche, la politique n°4 présente un coût relativement éloigné des politiques n°2 et n°3 (16057). Ces constatations se retrouvent pour les approximations des fonctions valeurs : la valeur $\hat{v}_0(\hat{Z}_0)$ associée à la matrice C_3 est la plus faible (11455), puis vient juste après la valeur $\hat{v}_0(\hat{Z}_0)$ associée à la matrice C_2 (11484) et enfin la valeur $\hat{v}_0(\hat{Z}_0)$ associée à la matrice C_4 , beaucoup plus éloignée (15676). Au final, les ordres de grandeur entre les approximations des fonctions valeurs associées aux matrices de coût sont similaires aux coûts de politiques dont les matrices associées sont les mêmes.

4.3.6 Évolution de l'horizon de contrôle

Jusqu'à présent, l'horizon de contrôle pour l'approximation numérique de la fonction valeur a été pris égal à 5, afin de limiter les temps d'exécution des programmes. On propose dans ce paragraphe d'étudier l'évolution de l'approximation de la fonction valeur quand on augmente l'horizon de contrôle N_0 . Il est en effet logique d'augmenter cet horizon; nous avons pu voir dans le tableau 3.3 que les trajectoires du processus contrôlé par les politiques n°2 à n°4 présentaient un nombre de sauts compris entre 1 et 31 pour 10^7 trajectoires simulées.

Afin d'observer uniquement l'évolution de l'approximation de la fonction valeur, les paramètres du modèles sont fixés tels que :

- L'ensemble \mathbb{U} est défini à partir des grilles $G = H = \{2 \text{ ans}, 4 \text{ ans}, 6 \text{ ans}, 8 \text{ ans}\}$, soit un ensemble \mathbb{U} à 286 éléments.
- Pour la quantification des trajectoires partant des points de l'ensemble de contrôle, on choisit $N_{\mathbb{U}} = 100$ et $S_{\mathbb{U}} = 100N$ réplifications. La quantification n'est effectuée qu'une seule fois.
- Les grilles de quantification $(\Gamma_n^{\Theta})_{0 \leq n \leq N_0}$ et leurs matrices associées $(M_n)_{0 \leq n \leq N_0-1}$ sont simulées une seule fois avec $N_{x_0} = 10^3$ points dans chaque grille Γ_n^{Θ} , $1 \leq n \leq N_0$ et $S_{x_0} = 10^3N$ réplifications.
- La discrétisation du temps est celle correspondant au pas de temps Δ_2 défini au paragraphe 4.2.4.
- La fonction coût est celle définie par la formule (2.24) avec pour matrice de coût C .

Les quantifications ne sont effectuées qu'une seule fois avec pour horizon de quantification le plus grand horizon de contrôle choisi pour l'étude; $N_{quantif} = 24$ pour la quantification des chaînes $(\Theta_n^i)_{1 \leq i \leq u}$ et $N_{quantif} = 25$ pour la quantification de la chaîne (Θ_n) . Ainsi, la variabilité de la quantification n'a pas d'impact sur les résultats.

Les résultats obtenus d'approximation de la fonction valeur en fonction de l'horizon N_0 sont représentés par le graphique 4.2. Le tableau 4.11 donne plus précisément quelques valeurs obtenues. On constate que plus l'horizon de contrôle N_0 augmente, plus l'approximation de la fonction valeur diminue et tend vers la valeur d'environ 9160. La dernière diminution de la fonction valeur se fait pour le passage de l'horizon de contrôle $N_0 = 18$ à l'horizon de contrôle

| N_0 | approximation fonction valeur |
|-------|-------------------------------|
| 3 | 19064 |
| 5 | 11435 |
| 7 | 10550 |
| 10 | 9390 |
| 15 | 9163.6 |
| 18 | 9160.1 |
| 19 | 9159.8 |
| 25 | 9159.8 |

TAB. 4.11 – Approximation des fonctions valeurs, à paramètres et grilles de quantification fixés, pour différents horizons de contrôle N_0 .

$N_0 = 19$. Pour des horizons de contrôle supérieurs à 19, l'approximation de la fonction valeur reste identique quand elle s'exprime avec cinq chiffres significatifs.

Le tableau 3.3 représentant la répartition des trajectoires contrôlées par les politiques n°2 à n°4 en fonction de leur nombre de sauts, peut donner un aperçu de la raison pour laquelle l'approximation de la fonction valeur évolue de cette manière. Bien qu'il ne soit question que de trois politiques d'intervention données, ce tableau montre que plus de 95% des trajectoires contrôlées ont un nombre de saut inférieur ou égal à 12. Les trajectoires ayant au plus 18 sauts représentent quant à elles plus de 99.86% des trajectoires contrôlées par les politiques d'intervention n°2 à n°4. L'horizon de contrôle N_0 autorise les intervention jusqu'au N_0 -ième saut du processus contrôlé. Or passer d'un horizon de contrôle N_0 de 18 à 19 donne peu de possibilité d'interventions supplémentaires, étant donné que les trajectoires potentiellement concernées par ces interventions représentent (pour les politiques n°2 à n°4) moins de 0.14% des trajectoires.

Par ailleurs, le processus de Markov déterministe par morceaux construit dans notre modélisation est tel qu'après 10 ans, plus aucun saut n'est possible car il se retrouve nécessairement dans l'état cimetièr. Par l'hypothèse 1 associée aux PDMP, on sait donc que le nombre moyen de sauts du PDMP modélisant l'évolution de la caméra contrôlée par une politique d'intervention donnée est fini. Il paraît donc logique qu'il existe un seuil pour l'horizon de contrôle N_0 au-delà duquel l'approximation de la fonction valeur est stabilisée.

Pour le modèle mis en place au chapitre 2, l'approximation numérique de la fonction valeur atteint ce seuil pour un horizon $N_0 = 19$ et la valeur de cette approximation est d'environ 9160. Cette valeur est inférieure aux coûts des politiques d'intervention n°1 à 4 (tableau 3.6). De manière plus spécifique, le coût de la politique n°3 étant de 9403, on peut maintenant confirmer que la politique n°3 est une bonne politique de maintenance, au sens où celle-ci n'est éloignée du coût optimal que d'environ 2.65%. Néanmoins, ce résultat laisse à penser qu'il y a probablement d'autres politiques de maintenance se rapprochant plus du coût optimal, donc meilleures que la politique d'intervention n°3. Ces politiques peuvent être complexes à décrire et donc difficiles à trouver. C'est la raison pour laquelle, on poursuit le développement théorique du contrôle impulsif pour des processus de Markov déterministes par morceaux dans la partie II de ce travail. On y propose une construction explicite d'une famille de stratégies ε -optimales, c'est-à-

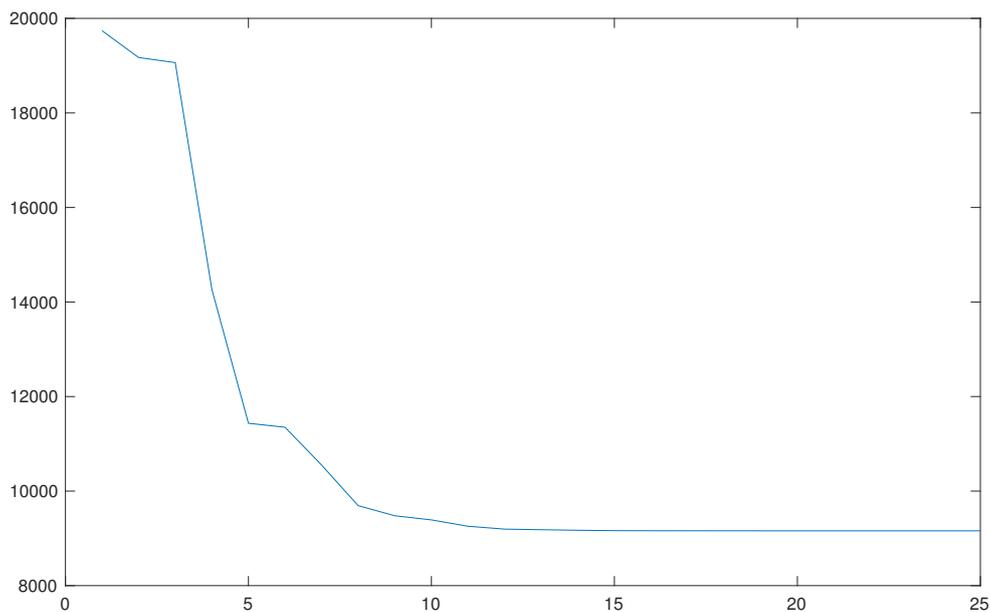


FIGURE 4.2 – Évolution de l'approximation de la fonction valeur en fonction de l'horizon de contrôle, à grilles de quantification fixées.

dire des politiques d'intervention dont le coût est proche à ε près du coût optimal.

4.4 Nouveau modèle

On propose d'élargir le critère (\mathcal{C}_1) qui autorise les interventions uniquement lorsque la caméra est en panne. On souhaite maintenant permettre les interventions dès lors que la caméra est dans l'état dégradé ou l'état de panne. Pour cela, il est nécessaire de redéfinir le modèle en introduisant notamment de nouveaux régimes. Ce travail est réalisé dans la partie 4.4.1. On s'intéresse ensuite dans le paragraphe 4.4.2 à deux politiques explicites d'intervention. Enfin, dans le paragraphe 4.4.3, on présente et discute des résultats d'approximation de la fonction valeur.

4.4.1 Définition du modèle

Régimes de la caméra

Par rapport au modèle défini dans le chapitre 2, nous numérotons l'état cimetièrre à 30 et nous introduisons trois nouveaux régimes d'atelier, à savoir :

- régime 27 : seule la machine à froid est réparée,
- régime 28 : seul le biafocal est réparé,
- régime 29 : seuls la machine à froid et le biafocal sont réparés.

Les régimes de ce nouveau modèle sont récapitulés dans le tableau 4.12. Ils sont au nombre de 30. On notera comme précédemment \mathcal{M} l'ensemble contenant tous les régimes du modèle. De même, on notera \mathcal{S} , \mathcal{D} , \mathcal{P} , \mathcal{A} et \mathcal{C} les ensembles contenant respectivement le régime stable, les régimes dégradés, les régimes de panne, les régimes d'atelier et le régime cimetièrre. Comparé au

modèle du chapitre 2, seuls les ensembles \mathcal{A} des régimes d'atelier et \mathcal{C} du régime cimetièrè sont changés ; ils sont maintenant tels que $\mathcal{A} = \{13, 14, \dots, 29\}$ et $\mathcal{C} = \{30\}$.

Processus de Markov déterministe par morceaux

En considérant les nouveaux ensembles de régimes, on note maintenant $\Delta := (30, \delta)$ l'état cimetièrè. Le PDMP décrivant l'évolution de la caméra (sans intervention) se définit à l'aide des mêmes formules que précédemment, à savoir :

- l'espace d'état E par la formule (2.1),
- le flot Φ par la formule (2.2),
- le taux de saut λ par la formule (2.4),
- le noyau de Markov par les formules (2.5) à (2.10).

Ensemble de contrôle

L'ensemble de contrôle infini est défini comme précédemment par :

$$\mathbb{U}_\infty = \mathcal{A} \times \mathbb{R}_+^3 \times [0; d_{étude}[\times\{0\}.$$

En revanche, l'ensemble de contrôle fini, noté \mathbb{U}_{fini} , demande quelques modifications par rapport à celui défini dans le chapitre 2. En effet, l'ensemble des régimes d'atelier possède trois nouveaux régimes. On redéfinit les ensembles $(\mathcal{A}_i)_{1 \leq i \leq 6}$ en tenant compte de ces régimes :

$$\begin{aligned} \mathcal{A}_1 &= \{13; 22\} \\ \mathcal{A}_2 &= \{14; 19; 23; 28\} \\ \mathcal{A}_3 &= \{15\} \\ \mathcal{A}_4 &= \{16, 20, 24, 29\} \\ \mathcal{A}_5 &= \{17, 27\} \\ \mathcal{A}_6 &= \{18, 21, 25, 26\}. \end{aligned}$$

Ainsi, l'ensemble de contrôle fini est défini toujours par :

$$\mathbb{U}_{fini} = \bigcup_{i=1}^6 \mathbb{U}_i,$$

où les ensembles $(\mathbb{U}_i)_{1 \leq i \leq 6}$ sont définis par les formules (2.11) et (2.12).

Coûts

Le coût d'indisponibilité reste identique dans ce modèle ; il est défini par la formule (2.13). En revanche, les coûts des interventions ne sont plus les mêmes. En effet, on souhaite permettre des interventions quand la caméra est dans l'état dégradé ou dans l'état de panne. On modifie donc le premier des trois critères que doivent satisfaire les interventions. On définit

(\mathcal{D}_1) Lors d'une intervention, la caméra passe d'un régime dégradé ou de panne à un régime d'atelier.

On impose que toutes les interventions satisfassent les critères (\mathcal{D}_1), (\mathcal{C}_2) et (\mathcal{C}_3). Ainsi, par rapport au modèle précédemment défini, il est possible d'intervenir quand la caméra est dans l'état dégradé et non pas uniquement quand celle-ci se trouve dans l'état de panne.

| régime | E | RB | MAF | BF | état de la caméra |
|--------|------------|------------------------|------------|------------|---------------------------------|
| 1 | - | - | - | - | \mathcal{S} : stable |
| 2 | - | - | dégradé | - | \mathcal{D} : dégradé |
| 3 | - | - | - | dégradé | |
| 4 | - | - | dégradé | dégradé | |
| 5 | - | panne du bloc (MAF+RB) | | - | \mathcal{P} : caméra en panne |
| 6 | - | - | - | panne | |
| 7 | panne | - | - | - | |
| 8 | - | - | dégradé | panne | |
| 9 | panne | - | dégradé | - | |
| 10 | - | panne du bloc (MAF+RB) | | dégradé | |
| 11 | panne | - | - | dégradé | |
| 12 | panne | - | dégradé | dégradé | |
| 13 | - | changement | changement | - | \mathcal{A} : atelier |
| 14 | - | - | - | changement | |
| 15 | changement | - | - | - | |
| 16 | - | - | réparation | changement | |
| 17 | changement | - | réparation | - | |
| 18 | - | changement | changement | réparation | |
| 19 | changement | - | - | réparation | |
| 20 | changement | - | réparation | réparation | |
| 21 | - | changement | changement | changement | |
| 22 | changement | changement | changement | - | |
| 23 | changement | - | - | changement | |
| 24 | changement | - | réparation | changement | |
| 25 | changement | changement | changement | réparation | |
| 26 | changement | changement | changement | changement | |
| 27 | - | - | réparation | - | |
| 28 | - | - | - | réparation | |
| 29 | - | - | réparation | réparation | |
| 30 | arrêté | arrêté | arrêté | arrêté | \mathcal{C} : cimetière |

TAB. 4.12 – Numérotation et description des régimes de la caméra pour le nouveau modèle. Le symbole – signifie que l'élément est dans un état stable. Les éléments E , RB , MAF et BF correspondent respectivement à l'électronique, au roulement à billes, à la machine à froid et au biafocal.

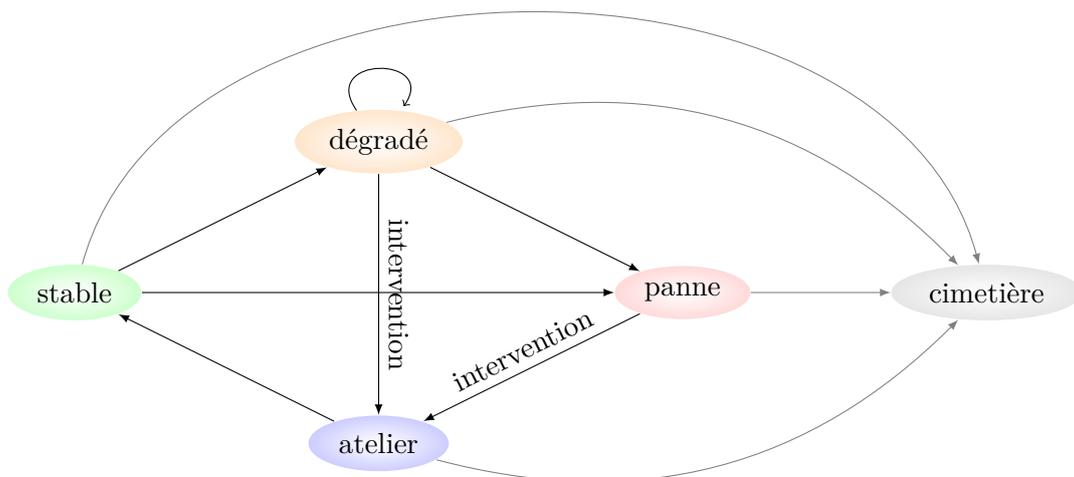


FIGURE 4.3 – Graphique d'évolution générale de la caméra Sophie si celle-ci est contrôlée à l'aide d'interventions satisfaisant les critères (\mathcal{D}_1) , (\mathcal{C}_2) et (\mathcal{C}_3) .

Le graphique d'évolution de la caméra contrôlée avec les interventions répondant aux critères (\mathcal{D}_1) , (\mathcal{C}_2) et (\mathcal{C}_3) est donné par la figure 4.3.

La fonction du coût des interventions est définie par les formules (2.22) à (2.24). Cependant, la matrice C n'est plus identique à celle du chapitre 2. En effet, pour m et m' appartenant à \mathcal{M} , $C(m, m')$ représente le coût d'un passage du régime m au régime m' . Or les régimes 27 à 29 sont nouveaux. La nouvelle matrice C est donc de taille 30×30 .

Comme les interventions doivent répondre au critère (\mathcal{D}_1) , on a :

$$\forall m \notin \mathcal{D} \cup \mathcal{P}, \forall m' \in \mathcal{M}, \quad C(m, m') = 0.$$

Les formules (2.14) à (2.21) définissant le coût des interventions d'un régime de panne à un régime atelier restent valables. À celles-ci s'ajoutent les trois formules suivantes permettant de passer d'un état dégradé à un état d'atelier :

$$\begin{aligned} \forall m' \in \mathcal{M}, \quad C(2, m') &= 3000 \times \mathbb{1}_{\{m'=27\}} + 4000 \times \mathbb{1}_{\{m'=29\}} + 6000 \times \mathbb{1}_{\{m'=17\}} + 7000 \times \mathbb{1}_{\{m'=20\}} \\ &+ 8000 \times \mathbb{1}_{\{m'=13\}} + 9000 \times \mathbb{1}_{\{m' \in \{16,18\}\}} + 11000 \times \mathbb{1}_{\{m'=22\}} \\ &+ 12000 \times \mathbb{1}_{\{m' \in \{24,25\}\}} + 14000 \times \mathbb{1}_{\{m'=21\}} + 17000 \times \mathbb{1}_{\{m'=26\}}, \end{aligned} \quad (4.2)$$

$$\begin{aligned} C(3, m') &= 3000 \times \mathbb{1}_{\{m'=28\}} + 4000 \times \mathbb{1}_{\{m'=29\}} + 6000 \times \mathbb{1}_{\{m'=19\}} + 7000 \times \mathbb{1}_{\{m'=20\}} \\ &+ 8000 \times \mathbb{1}_{\{m'=14\}} + 9000 \times \mathbb{1}_{\{m' \in \{16,18\}\}} + 11000 \times \mathbb{1}_{\{m'=23\}} \\ &+ 12000 \times \mathbb{1}_{\{m' \in \{24,25\}\}} + 14000 \times \mathbb{1}_{\{m'=21\}} + 17000 \times \mathbb{1}_{\{m'=26\}}, \end{aligned} \quad (4.3)$$

$$\begin{aligned} C(4, m') &= 4000 \times \mathbb{1}_{\{m'=29\}} + 7000 \times \mathbb{1}_{\{m'=20\}} + 9000 \times \mathbb{1}_{\{m' \in \{16,18\}\}} \\ &+ 12000 \times \mathbb{1}_{\{m' \in \{24,25\}\}} + 14000 \times \mathbb{1}_{\{m'=21\}} + 17000 \times \mathbb{1}_{\{m'=26\}}. \end{aligned} \quad (4.4)$$

La définition du processus de Markov déterministe par morceaux modélisant l'évolution de la caméra et le problème de contrôle impulsif étant mis en place, nous pouvons maintenant nous intéresser à des politiques d'intervention particulières.

4.4.2 Politiques d'intervention

Dans ce paragraphe, nous définissons deux politiques d'intervention. La première, appelée politique n°5, impose de réparer les éléments dans l'état dégradé et de changer les éléments dans l'état de panne lorsque la caméra revient à l'atelier. Ce retour s'effectue lorsque la caméra a passé une durée fixée dans l'état dégradé ou dans l'état de panne. La seconde politique définie dans ce paragraphe, appelée politique n°6, impose de changer tous les éléments n'étant pas dans l'état stable lors du passage à l'atelier de la caméra, qui a lieu lorsque celle-ci a passé une durée fixée dans l'état dégradé ou dans l'état de panne.

Après avoir défini précisément ces deux politiques d'intervention, on calcule leur coût et on les compare.

Politique d'intervention n°5

La politique d'intervention n°5 est définie telle qu'elle satisfait les critères (C_2'') , (C_3) et le critère suivant :

(D_1') L'intervention fait passer la caméra d'un régime dégradé ou d'un régime de panne à un régime d'atelier après un temps d_{panne} passé soit dans le régime dégradé, soit dans le régime de panne.

La matrice C_5 des coûts associée à cette politique est une matrice de taille 30 sur 30. Elle est définie par :

$$\begin{aligned} \forall m, m' \in \mathcal{M}, \quad C_5(m, m') = & 3000 \times \mathbb{1}_{\{m \in \{2,3\} \} \cap \{m' = m+25\}} + 4000 \times \mathbb{1}_{\{(m,m') = (4,29)\}} \\ & + \mathbb{1}_{\{m \in \mathcal{P}\} \cap \{m' = m+8\}} (5000 \times \mathbb{1}_{\{m=7\}} + 6000 \times \mathbb{1}_{\{m \in \{9,11\}\}} \\ & + 7000 \times \mathbb{1}_{\{m=12\}} + 8000 \times \mathbb{1}_{\{m \in \{5,6\}\}} + 9000 \times \mathbb{1}_{\{m \in \{8,10\}\}}). \end{aligned} \quad (4.5)$$

Politique d'intervention n°6

La politique d'intervention n°6 diffère de la politique d'intervention n°5 par le second critère. On change les éléments qui se trouvent dans l'état dégradé au lieu de les réparer, ce qui impacte le coût des trajectoires. La politique d'intervention n°6 satisfait les critères (D_1') , (C_2') et (C_3) .

La matrice des coûts C_6 correspondant à cette politique d'intervention est définie pour $(m, m') \in \mathcal{M}^2$, par :

$$\begin{aligned} C_6(m, m') = & 8000 \times \mathbb{1}_{\{(m,m') \in \{(2,13), (3,14), (5,13), (6,14)\}\}} + 5000 \times \mathbb{1}_{\{(m,m') = (7,15)\}} \\ & + 14000 \times \mathbb{1}_{\{(m,m') \in \{(4,21), (8,21), (10,21)\}\}} \\ & + 11000 \times \mathbb{1}_{\{(m,m') = (9,22)\} \cup \{(m,m') = (11,23)\}} + 17000 \times \mathbb{1}_{\{(m,m') = (12,26)\}}. \end{aligned} \quad (4.6)$$

La politique n°6 privilégie les changements des éléments de la caméra, au contraire de la politique d'intervention n°5 qui remplace les éléments dans l'état dégradé. Par conséquent, la politique n°5 devrait être meilleure que la politique n°6 en terme de coût. Pour le reste comme le nombre moyen de pannes, il n'y a pas de différences entre ces deux politiques. En effet, une réparation revient à une remise à neuf de l'élément.

Tailles des trajectoires contrôlées

On considère l'ensemble de contrôle \mathbb{U}_∞ et les variables suivantes : $d_{étude} = 10$ ans, $d_{atelier} = 35$ jours, $d_{panne} = 1$ jour. On simule 10^7 trajectoires contrôlées par la politique n°5, puis 10^7 trajectoires contrôlées par la politique n°6.

Le nombre de sauts des trajectoires est compris entre 1 et 31 pour les deux politiques. Le tableau 4.13 détaille la répartition en pourcentage du nombre de sauts pour les trajectoires

| nombre de sauts des trajectoires | politique n°5 | politique n°6 |
|----------------------------------|---------------|---------------|
| 1 | 13.91% | 13.92% |
| 2 | 0.49% | 0.49% |
| 3 | 0.21% | 0.21% |
| 4 | 25.03% | 25.03% |
| 5 | 16.93% | 16.89% |
| 6 | 0.65% | 0.64% |
| 7 | 10.74% | 10.73% |
| 8 | 13.46% | 13.49% |
| 9 | 5.49% | 5.49% |
| 10 | 2.79% | 2.79% |
| 11 | 3.98% | 3.99% |
| 12 | 2.87% | 2.87% |
| 13 | 1.28% | 1.28% |
| 14 | 0.75% | 0.76% |
| 15 | 0.66% | 0.66% |
| 16 | 0.38% | 0.38% |
| 17 | 0.17% | 0.17% |
| 18 | 0.10% | 0.10% |
| 19 | 0.07% | 0.06% |
| 20 | 0.03% | 0.03% |
| 21 | 0.015% | 0.015% |
| de 22 à 31 | < 0.01% | < 0.01% |

TAB. 4.13 – Répartition exprimée en pourcentage, de 10^7 trajectoires simulées pour des caméras partant de l'état neuf suivant les politiques d'intervention n°5 et n°6, en fonction de leur nombre de sauts.

contrôlées par les deux politiques d'intervention. Les résultats sont similaires pour les deux politiques. En effet, la dynamique des processus contrôlés par ces deux politiques est semblable ; seuls les coûts sont plus élevés pour la politique n°6 car cette dernière impose de changer toutes les pièces dégradées de la caméra, au lieu de les réparer.

Coût des politiques d'intervention

Le coût des politiques d'intervention n°5 et n°6 s'estime par une méthode de Monte-Carlo, de la même manière qu'au chapitre 2, avec $\mathbb{U}_\infty = \mathcal{A} \times \mathbb{R}_+^3 \times [0; d_{\text{étude}}[\times\{0\}$. Les résultats sont regroupés dans le tableau 4.14.

Le coût de la politique d'intervention n°5 (7002) est très inférieur au coût de la politique n°6 (9880), alors que les coûts d'indisponibilité sont identiques. En effet, les politiques n°5 et n°6 ne diffèrent que par les opérations effectuées sur les éléments dans l'état dégradés ; pour la politique n°5, ils sont réparés alors que pour la politique n°6, ils sont changés à neuf. La dynamique du processus contrôlé par ces politiques en exceptant les interventions est donc la même, d'où un coût d'indisponibilité identique. En revanche, le changement d'un élément coûtant sensiblement plus cher que la réparation de cet élément, il est normal d'obtenir un coût pour la politique

| politique d'intervention n° | 5 | 6 |
|-----------------------------|--------------|--------------|
| coût d'indisponibilité | 9.93 | 9.93 |
| coût interventions | 6992 | 9870 |
| coût moyen | 7002 | 9880 |
| intervalle de confiance | [7000; 7003] | [9878; 9882] |

TAB. 4.14 – Coût des politiques d'intervention n°5 et n°6, estimé par une méthode de Monte-Carlo avec 10^7 trajectoires simulées et pour l'ensemble de contrôle \mathbb{U}_∞ .

d'intervention n°5 inférieur à celui de la politique n°6.

Par rapport aux résultats du modèle défini au chapitre 2, on peut constater que le coût de la politique d'intervention n°5 est inférieur de plus de 25% au coût des politiques de référence n°1 à n°4 pour le modèle du chapitre 2. Il est également inférieur à l'approximation numérique de la fonction valeur (9160). Il est donc très intéressant d'élargir le critère (\mathcal{C}_1) pour permettre des interventions lorsque la caméra est dans l'état dégradé.

4.4.3 Approximation de la fonction valeur

On s'intéresse maintenant à l'approximation numérique de la fonction valeur. Les paramètres sont fixés tels que :

- L'ensemble \mathbb{U}_{fini} est défini à partir des grilles $G = H = \{2 \text{ ans}, 4 \text{ ans}, 6 \text{ ans}, 8 \text{ ans}\}$, soit un ensemble \mathbb{U}_{fini} à 356 éléments.
- Pour la quantification des trajectoires partant des points de l'ensemble de contrôle, on choisit $N_{\mathbb{U}} = 100$ et $S_{\mathbb{U}} = 100N$ réplifications. La quantification n'est effectuée qu'une seule fois.
- Les grilles de quantification $(\Gamma_n^\Theta)_{0 \leq n \leq N_0}$ et leurs matrices associées $(M_n)_{0 \leq n \leq N_0-1}$ sont simulées une seule fois avec $N_{x_0} = 10^3$ points dans chaque grille Γ_n^Θ , $1 \leq n \leq N_0$ et $S_{x_0} = 10^3N$ réplifications.
- La discrétisation du temps est celle correspondant au pas de temps Δ'_2 défini par :

$$\forall z = (m, \zeta) \in E, \quad \Delta'_2 := \begin{cases} 2 & \text{si } m \in \mathcal{D} \cup \mathcal{P} \\ \frac{t^*(z)}{2} & \text{sinon.} \end{cases}$$

- La fonction coût est celle associée à la matrice C définie dans le paragraphe 4.4.1.

On étudie l'évolution de l'approximation numérique de la fonction valeur en fonction de l'horizon, afin de déterminer la valeur pour laquelle l'approximation trouvée se stabilise. Pour cela, on quantifie les trajectoires du processus (Θ_n) sur un horizon $N_{quantif}$ correspondant au maximum des horizons de contrôle testé, à savoir 25 et on quantifie les trajectoires des processus (Θ_n^i) , pour $1 \leq i \leq u = 356$, avec un horizon $N_{quantif} = 24$. Ainsi, la variabilité des résultats de l'approximation numérique due à la quantification n'interfère pas dans les résultats obtenus dans cette partie.

Le graphique d'évolution de l'approximation numérique de la fonction valeur en fonction de l'horizon de contrôle N_0 est donné par la figure 4.4. Quelques valeurs (à cinq chiffres significatifs près) sont détaillées dans le tableau 4.15.

| N_0 | approximation fonction valeur |
|-------|-------------------------------|
| 3 | 12593 |
| 5 | 7664 |
| 7 | 5673 |
| 10 | 5405 |
| 15 | 5374.6 |
| 18 | 5373.4 |
| 19 | 5373.3 |
| 25 | 5373.3 |

TAB. 4.15 – Approximation des fonctions valeurs, à paramètres et grilles de quantification fixés, pour différents horizons de contrôle, pour le modèle à 30 régimes.

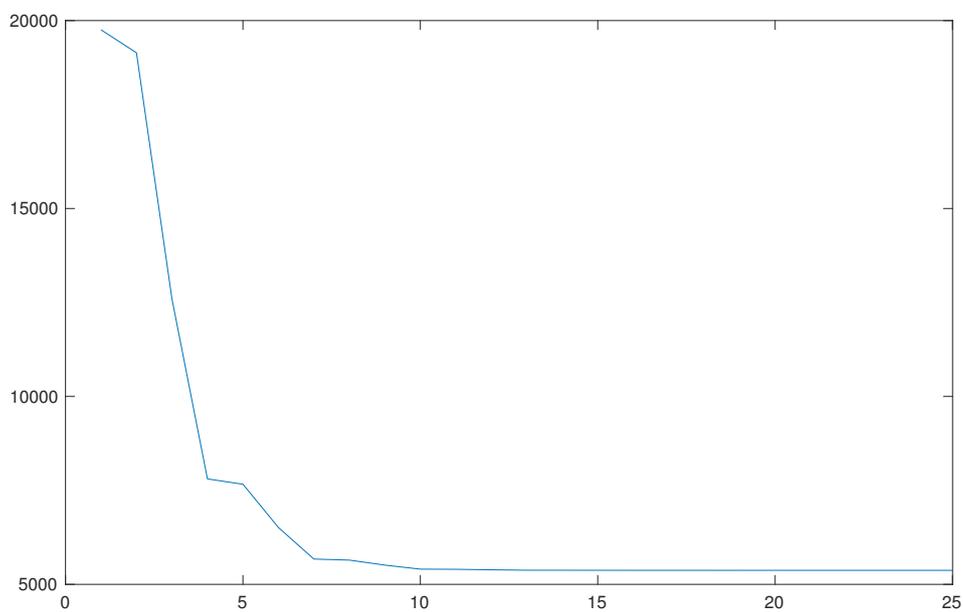


FIGURE 4.4 – Évolution de l'approximation de la fonction valeur en fonction de l'horizon de contrôle, à grilles de quantification fixées, pour le modèle à 30 régimes.

On constate que l'approximation numérique de la fonction valeur décroît quand l'horizon de contrôle augmente. À partir de l'horizon $N_0 = 19$, la valeur obtenue est constante jusqu'à l'horizon $N_0 = 25$. On peut donc considérer que l'approximation numérique de la fonction valeur pour le modèle défini dans le paragraphe 4.4.1 est d'environ 5373. Ce comportement peut se justifier de la même manière que précédemment, au vu du tableau 4.13.

Par rapport au coût moyen de la politique de référence n°5 (7002), l'approximation de la fonction valeur (5373) est bien inférieure. Il devient donc très intéressant de rechercher des politiques d'intervention dont le coût se rapproche de cette valeur.

4.5 Bilan

Notre objectif initial était de trouver une approximation de la fonction valeur du problème de contrôle impulsif liée à notre application d'optimisation de la maintenance de la caméra, en mettant en place la méthode numérique développée dans de Saporta et Dufour (2012).

Dans ce chapitre, nous avons effectué plusieurs tests pour déterminer l'influence de paramètres sur le résultat de la fonction valeur. Dans un premier temps, nous avons pu constater que le pas de discrétisation de l'intervalle $[0; t^*(z)]$, où $z \in E$, avait un impact sur le résultat d'approximation de la fonction valeur : plus le pas était grand, plus l'approximation de la fonction valeur augmentait. On a néanmoins pu observer que l'augmentation de cette approximation restait faible quand le pas de discrétisation était compris entre 1 jour et 5 jours. Il pourrait être possible de diminuer encore le pas de discrétisation, c'est-à-dire de prendre un pas inférieur à un jour. Cependant, les temps d'exécution des programmes augmenteraient beaucoup et on peut s'attendre à un résultat d'approximation de la fonction valeur proche de celle associée au pas de discrétisation à un jour, étant donné le graphique 4.1.

Par ailleurs, nous avons testé l'influence de la quantification dans l'approximation de la fonction valeur. Pour les valeurs testées, nous avons constaté que le nombre de points dans les grilles de quantification et le nombre de répliques pour créer ces grilles influent peu sur le résultat de la fonction valeur (tableaux 4.5 et 4.7). D'autre part, l'algorithme permettant de créer les grilles de quantification présente une variance. En effet, les grilles créées dépendent des simulations obtenues pour l'initialisation et la phase de réplique. Les résultats des tableaux 4.6 et 4.8 ont montré qu'a priori, la variance des grilles de quantification avait un impact très modéré sur le résultat d'approximation de la fonction valeur, de l'ordre de 0.2%. Une piste d'amélioration pourrait donc être d'augmenter encore le nombre de points dans les grilles et le nombre de répliques pour les créer. Néanmoins, l'espace mémoire utilisé par les programmes est tel qu'il demande des moyens spécifiques (optimisation des algorithmes, plateforme de calculs, ...). Cependant, l'augmentation du nombre de points dans les grilles de quantification pourrait permettre une meilleure gestion des régimes rares, présents dans notre modèle.

Concernant le choix des grilles de discrétisation pour la construction de l'ensemble de contrôle \mathbb{U}_{fini} , nous nous sommes basés sur le critère développé dans le paragraphe 3.4. Nous avons opté pour un ensemble \mathbb{U}_{fini} avec une discrétisation des temps de fonctionnement et du temps calendaire égale à l'ensemble $\{2, 4, 6, 8\}$ ans. Cette discrétisation avait l'avantage d'offrir un bon compromis entre l'erreur créée par la projection sur cet ensemble de contrôle et le cardinal de cet ensemble. Au paragraphe 4.3.4, nous avons calculé une approximation de la fonction valeur associée à différents choix de discrétisation du temps calendaire et des temps de fonctionnement. Nous avons constaté que les approximations obtenues pouvaient présenter un écart entre elles non négligeable. Néanmoins, à un ensemble de contrôle \mathbb{U}_{fini} correspond un problème de contrôle impulsif spécifique et donc une unique fonction valeur. Il reste donc délicat de les comparer.

Dans les paragraphes 4.3.6 et 4.4.3, nous avons obtenu les résultats suivants. En considérant les mêmes paramètres que dans le paragraphe 4.3.5 et en prenant pour matrice de coût celle définie dans la partie 2.4.2, c'est-à-dire en autorisant toutes les interventions répondant aux

critères (\mathcal{C}_1) à (\mathcal{C}_3) , on obtient une approximation de la fonction valeur égale à 9160. En adaptant le modèle initialement choisi et en autorisant toutes les interventions répondant aux critères (\mathcal{D}_1) , (\mathcal{C}_2) et (\mathcal{C}_3) , on évalue l'approximation de la fonction valeur à environ 5373. Dans les deux modèles, nous avons pu constater que pour un horizon de contrôle N_0 supérieur à 19, l'approximation numérique obtenue de la fonction valeur était constante.

Des pistes d'amélioration des résultats d'approximation de la fonction existent. Comme il a été suggéré précédemment, la discrétisation de l'intervalle $[0; t^*(z)]$ peut être affinée, le nombre de points dans les grilles de quantification et le nombre de répliques faites pour la création de ces grilles peuvent être augmentés. Il est également possible de changer l'ensemble de contrôle, en explorant les multiples possibilités de discrétisation des temps de fonctionnement et de discrétisation du temps calendaire. Il est important de noter que ces pistes d'amélioration demandent une bonne gestion de l'espace mémoire utilisé par les programmes.

Concernant l'optimisation de la caméra Sophie, les approximations numériques obtenues nous donnent des informations très intéressantes par rapport aux politiques de référence proposées dans cette thèse. Dans le cas des interventions répondant aux critères (\mathcal{C}_1) à (\mathcal{C}_3) , la politique présentant le coût le plus faible est la politique n°3. Celle-ci correspond à la réparation des éléments dégradés et au changement des blocs en panne lorsque la caméra est en panne. Son coût est de 9403. En comparaison avec l'approximation numérique de la fonction valeur trouvée (9160), on peut considérer que la politique n°3 est une très bonne politique. En effet, elle présente un coût proche de la fonction valeur et peut être décrite facilement. Une politique dont le coût est inférieur au coût de la politique n°3 pourrait être trouvée. Néanmoins, il n'est pas évident que celle-ci soit facile à mettre en oeuvre. La politique n°3 offre donc une solution performante pour l'entreprise Thales, quand seuls les maintenances lors des pannes sont possibles.

Si nous considérons maintenant le modèle élargi où les interventions répondent aux critères (\mathcal{D}_1) , (\mathcal{C}_2) et (\mathcal{C}_3) , c'est-à-dire que les interventions sont rendues possibles lorsque la caméra est en panne ou dans l'état dégradé, la meilleure politique explicitée dans ce travail est la politique n°5 dont le coût est 7002. Cette politique nous permet de voir le gain possible pour l'entreprise de passer de maintenances uniquement lors des pannes à des maintenances dès lors que la caméra entre dans un état dégradé. Par ailleurs, le coût de la politique n°5 est assez éloigné de l'approximation numérique de la fonction valeur qui est alors de 5373. Il est donc très intéressant d'étudier d'autres politiques d'intervention afin d'en trouver une plus performante que la politique n°5. Pour cela, la partie II de cette thèse réalise la première étape de cet objectif, à savoir la construction théorique de stratégies ε -optimales.

Deuxième partie

Stratégies optimales de contrôle impulsionnel pour des processus Markoviens déterministes par morceaux

Chapitre 5

Construction de stratégies ε -optimales

Sommaire

| | |
|---|------------|
| 5.1 Construction d'approximation des fonctions valeurs | 116 |
| 5.1.1 Propriété de monotonie des opérateurs K et J | 117 |
| 5.1.2 Approximation des fonctions valeurs | 118 |
| 5.2 Construction de stratégies d'intervention explicites | 120 |
| 5.3 Coût des trajectoires contrôlées | 127 |
| 5.3.1 Majoration du temps éventuel d'intervention | 127 |
| 5.3.2 Distributions de la première position et du premier temps de saut pour les processus contrôlés et non contrôlés | 128 |
| 5.3.3 Propriété de Markov des temps d'intervention | 129 |
| 5.3.4 Coût associé aux stratégies explicites construites | 131 |
| 5.4 Stratégies ε-optimales | 138 |

Dans ce chapitre, on s'intéresse à un problème de contrôle impulsif avec facteur d'oubli et horizon infini, pour des processus de Markov déterministes par morceaux (PDMP). Ce type de problème est défini au paragraphe 1.6.2. Notre objectif est de proposer une nouvelle famille de stratégies ε -optimales pour cette catégorie de contrôle impulsif. En effet, la construction de stratégies ε -optimales proposée dans Costa et Davis (1989) présente l'inconvénient de devoir non seulement résoudre plusieurs problèmes d'arrêt optimal, mais également de connaître, en tout point de l'espace, la fonction valeur associée à un problème de contrôle impulsif où le processus considéré est une version discrétisée du PDMP initial, à valeurs dans un espace fini de points de l'espace.

Nous proposons une construction de stratégies ε -optimales ne nécessitant pas la résolution au préalable de problèmes techniques. La construction de ces stratégies est basée sur l'itération d'un opérateur, dit de simple-saut-ou-intervention, associé au PDMP. L'idée de base de cette construction repose sur le procédé utilisé pour la construction des temps d'arrêt ε -optimaux, développé par U.S. Gugerli dans Gugerli (1986) pour le problème d'arrêt optimal. Néanmoins, le fait de devoir chercher la meilleure position après chaque intervention pour faire redémarrer le processus complique significativement la construction de stratégies ε -optimales. Il est en effet nécessaire d'ajouter un nouvel opérateur qui sera associé à la sélection du nouveau point de départ du processus après une intervention. Le contenu de ce chapitre a fait l'objet de l'article de Saporta *et al.* (2017).

Le chapitre est organisé de la manière suivante. Dans le paragraphe 5.1, on construit une suite d'approximation des fonctions valeurs. On construit ensuite dans la partie 5.2, un processus contrôlé sur un espace d'état élargi par rapport au processus initial. Puis on montre qu'il correspond à une famille explicite de stratégies. Dans la partie 5.3, après avoir énoncé et démontré plusieurs résultats nécessaires à la suite de notre travail, on calcule le coût des trajectoires contrôlées par la famille de stratégies précédemment construite, ce qui nous permet de montrer que ce coût correspond exactement à l'approximation des fonctions valeurs construites au paragraphe 5.1. Il s'agit de la partie la plus importante de ce chapitre. Il nous permet en effet de montrer que les stratégies construites au paragraphe 5.2 sont ε -optimales. La partie 5.4 détaille ce dernier résultat et donne des perspectives concernant les applications numériques qui peuvent découler de ce travail.

5.1 Construction d'approximation des fonctions valeurs

On commence par construire une approximation de la suite des fonctions valeurs (\mathcal{V}_n) . Quand l'opérateur \mathcal{L} est itéré, le contrôleur sélectionne l'option qui génère le coût le plus faible entre : attendre le prochain temps de saut (l'opérateur K gagne la minimisation) et intervenir (l'opérateur J gagne la minimisation). Dans ce dernier cas, le meilleur temps d'intervention est donné par le temps qui réalise l'infimum de J . Pour la construction des approximations des fonctions valeurs, on utilisera une approximation de cet infimum.

On pose $V_0 = h$ et on fixe $\varepsilon > 0$. On définit tout d'abord pour tout x dans l'espace d'état E :

$$r_\varepsilon^1(x) = \begin{cases} t^*(x) & \text{si } KV_0(x) < \inf_{t \in \mathbb{R}^+} J(MV_0, V_0)(x, t), \\ \inf\{s \in \mathbb{R}^+ \mid J(MV_0, V_0)(x, s) < \inf_{t \in \mathbb{R}^+} J(MV_0, V_0)(x, t) + \varepsilon\} & \text{sinon,} \end{cases} \quad (5.1)$$

et

$$V_1(x) = \begin{cases} KV_0(x) & \text{si } KV_0(x) < \inf_{t \in \mathbb{R}^+} J(MV_0, V_0)(x, t), \\ J(MV_0, V_0)(x, r_\varepsilon^1(x)) & \text{sinon.} \end{cases} \quad (5.2)$$

Par récurrence, on définit pour tout $k \in \mathbb{N}^*$

$$r_\varepsilon^k(x) = \begin{cases} t^*(x) & \text{si } KV_{k-1}(x) < \inf_{t \in \mathbb{R}^+} J(MV_{k-1}, V_{k-1})(x, t), \\ \inf\{s \in \mathbb{R}^+ \mid J(MV_{k-1}, V_{k-1})(x, s) < \inf_{t \in \mathbb{R}^+} J(MV_{k-1}, V_{k-1})(x, t) + \varepsilon\} & \text{sinon} \end{cases} \quad (5.3)$$

et

$$V_k(x) = \begin{cases} KV_{k-1}(x) & \text{si } KV_{k-1}(x) < \inf_{t \in \mathbb{R}^+} J(MV_{k-1}, V_{k-1})(x, t), \\ J(MV_{k-1}, V_{k-1})(x, r_\varepsilon^k(x)) & \text{sinon.} \end{cases} \quad (5.4)$$

Remarque 2

Étant donné que l'ensemble $\mathbf{L}_\Phi(E)$ est stable par les opérateurs K et $J(M\cdot, \cdot)$ (de Saporta et Dufour, 2012, lemmes A.3 à A.6 et A.9) et que $h \in \mathbf{L}_\Phi(E)$, on obtient l'appartenance des fonctions V_k à l'ensemble $\mathbf{L}_\Phi(E)$ pour tout $k \in \mathbb{N}$.

Les fonctions $(r_\varepsilon^k)_{k \in \mathbb{N}^*}$ donnent le temps éventuel d'intervention. En effet, deux cas peuvent se produire : si l'opérateur K gagne la minimisation dans la définition de la fonction V_k , alors il est plus rentable de ne pas intervenir avant le prochain saut. La fonction r_ε^k est égale au temps d'atteinte à la frontière de l'espace d'état t^* , ce qui permet au processus d'évoluer librement jusqu'au saut suivant. En revanche, si l'opérateur J gagne la minimisation dans la définition de la fonction V_k , il est préférable d'intervenir là où le coût est le plus faible, c'est-à-dire pour le temps qui réalise l'infimum $\inf_{t \in \mathbb{R}^+} J(MV_{k-1}, V_{k-1})(\cdot, t)$ s'il existe. Cependant, pour être

sûr de trouver un temps possible d'intervention, on prend une approximation de cet infimum, donnée par la deuxième ligne de la définition 5.3 de la fonction r_ε^k . Ainsi, on permet au contrôleur d'intervenir sur le processus au temps donné par la fonction r_ε^k .

On propose dans ce paragraphe de montrer que les fonctions $(V_k)_{k \in \mathbb{N}}$ sont bien des approximations des fonctions valeurs $(\mathcal{V}_k)_{k \in \mathbb{N}}$. On commence par démontrer au paragraphe 5.1.1 un résultat de monotonie des opérateurs K et J , nécessaire pour montrer ensuite l'encadrement des fonctions valeurs $(\mathcal{V}_k)_{k \in \mathbb{N}}$ par les fonctions $(V_k)_{k \in \mathbb{N}}$ au paragraphe 5.1.2.

5.1.1 Propriété de monotonie des opérateurs K et J

Le lemme suivant nous est utile pour monter que les fonction valeurs $(\mathcal{V}_k)_{k \in \mathbb{N}}$ peuvent être approchées par les fonctions $(V_k)_{k \in \mathbb{N}}$, définies ci-dessus. Ce lemme provient de la monotonie de l'espérance.

Lemme 5.1.1

Soit $\beta > 0$. On suppose que

$$\forall x \in E, \quad \mathcal{V}_k(x) \leq V_k(x) \leq \mathcal{V}_k(x) + \beta. \quad (5.5)$$

Alors pour tout $x \in E$ et pour tout $t \in \mathbb{R}^+$, on a

$$K\mathcal{V}_k(x) \leq KV_k(x) \leq K\mathcal{V}_k(x) + \beta \quad (5.6)$$

et

$$J(M\mathcal{V}_k, \mathcal{V}_k)(x, t) \leq J(MV_k, V_k)(x, t) \leq J(M\mathcal{V}_k, \mathcal{V}_k)(x, t) + \beta. \quad (5.7)$$

Remarque 3

On a vu (remarque 2) que pour tout $k \in \mathbb{N}$, les fonctions V_k et les fonctions valeurs \mathcal{V}_k appartiennent à $\mathbf{L}_\Phi(E)$. Par conséquent, les fonctions $K\mathcal{V}_k$, KV_k , $J(M\mathcal{V}_k, \mathcal{V}_k)$ et $J(MV_k, V_k)$ existent et sont bien définies.

Démonstration : On suppose avoir l'inégalité (5.5) pour tout $x \in E$. Commençons par montrer l'inégalité (5.6). La variable aléatoire S_1 est positive, $\alpha > 0$ et Z_1 est une variable aléatoire à valeurs dans l'espace E . Par conséquent, on a

$$\begin{aligned} e^{-\alpha S_1} \mathcal{V}_k(Z_1) &\leq e^{-\alpha S_1} V_k(Z_1) \leq e^{-\alpha S_1} (\mathcal{V}_k(Z_1) + \beta) \\ e^{-\alpha S_1} \mathcal{V}_k(Z_1) &\leq e^{-\alpha S_1} V_k(Z_1) \leq e^{-\alpha S_1} \mathcal{V}_k(Z_1) + \beta. \end{aligned}$$

En utilisant la monotonie de l'espérance, on obtient

$$\mathbb{E}_x[e^{-\alpha S_1} \mathcal{V}_k(Z_1)] \leq \mathbb{E}_x[e^{-\alpha S_1} V_k(Z_1)] \leq \mathbb{E}_x[e^{-\alpha S_1} \mathcal{V}_k(Z_1)] + \beta.$$

Soit $x \in E$. On a

$$F(x, t^*(x)) + \mathbb{E}_x[e^{-\alpha S_1} \mathcal{V}_k(Z_1)] \leq F(x, t^*(x)) + \mathbb{E}_x[e^{-\alpha S_1} V_k(Z_1)] \leq F(x, t^*(x)) + \mathbb{E}_x[e^{-\alpha S_1} \mathcal{V}_k(Z_1)] + \beta.$$

La définition (1.8) de l'opérateur K implique

$$K\mathcal{V}_k(x) \leq KV_k(x) \leq K\mathcal{V}_k(x) + \beta,$$

ce qui prouve l'inégalité (5.6).

Montrons maintenant l'inégalité (5.7). Soient $x \in E$ et $t \in \mathbb{R}^+$. D'une part, par la définition (1.9) de l'opérateur M , l'inégalité (5.5) entraîne

$$M\mathcal{V}_k(x) \leq MV_k(x) \leq M\mathcal{V}_k(x) + \beta,$$

et la monotonie de l'espérance nous donne

$$\begin{aligned} \mathbb{E}_x[e^{-\alpha(t \wedge t^*(x))} M\mathcal{V}_k(\Phi(x, t \wedge t^*(x))) \mathbb{1}_{\{S_1 \geq t \wedge t^*(x)\}}] \\ \leq \mathbb{E}_x[e^{-\alpha(t \wedge t^*(x))} MV_k(\Phi(x, t \wedge t^*(x))) \mathbb{1}_{\{S_1 \geq t \wedge t^*(x)\}}] \\ \leq \mathbb{E}_x[e^{-\alpha(t \wedge t^*(x))} M\mathcal{V}_k(\Phi(x, t \wedge t^*(x))) \mathbb{1}_{\{S_1 \geq t \wedge t^*(x)\}}] + \beta \mathbb{E}_x[e^{-\alpha(t \wedge t^*(x))} \mathbb{1}_{\{S_1 \geq t \wedge t^*(x)\}}]. \end{aligned}$$

D'autre part, en utilisant encore l'inégalité (5.5) et la monotonie de l'espérance, on a

$$\begin{aligned} \mathbb{E}_x[e^{-\alpha S_1} \mathcal{V}_k(Z_1) \mathbb{1}_{\{S_1 < t \wedge t^*(x)\}}] &\leq \mathbb{E}_x[e^{-\alpha S_1} V_k(Z_1) \mathbb{1}_{\{S_1 < t \wedge t^*(x)\}}] \\ &\leq \mathbb{E}_x[e^{-\alpha S_1} \mathcal{V}_k(Z_1) \mathbb{1}_{\{S_1 < t \wedge t^*(x)\}}] + \beta \mathbb{E}_x[e^{-\alpha S_1} \mathbb{1}_{\{S_1 < t \wedge t^*(x)\}}]. \end{aligned}$$

Avec les deux précédentes inégalités et la définition (1.7) de l'opérateur J , on obtient

$$\begin{aligned} J(M\mathcal{V}_k, \mathcal{V}_k)(x, t) &\leq J(MV_k, V_k)(x, t) \leq J(M\mathcal{V}_k, \mathcal{V}_k)(x, t) \\ &\quad + \beta \mathbb{E}_x[e^{-\alpha(t \wedge t^*(x))} \mathbb{1}_{\{S_1 \geq t \wedge t^*(x)\}} + e^{-\alpha S_1} \mathbb{1}_{\{S_1 < t \wedge t^*(x)\}}] \\ &\leq J(M\mathcal{V}_k, \mathcal{V}_k)(x, t) + \beta \mathbb{E}_x[\mathbb{1}_{\{S_1 \geq t \wedge t^*(x)\}} + \mathbb{1}_{\{S_1 < t \wedge t^*(x)\}}] \\ &\leq J(M\mathcal{V}_k, \mathcal{V}_k)(x, t) + \beta, \end{aligned}$$

ce qui correspond bien au résultat voulu. ■

5.1.2 Approximation des fonctions valeurs

On montre maintenant que les fonctions V_k sont proches des fonctions valeurs \mathcal{V}_k de notre problème de contrôle impulsif.

Théorème 5.1.1

On a l'encadrement suivant

$$\forall k \in \mathbb{N}, \forall x \in E, \quad \mathcal{V}_k(x) \leq V_k(x) \leq \mathcal{V}_k(x) + k\varepsilon. \quad (5.8)$$

Démonstration : Le cas $k = 1$ est immédiat car $\mathcal{V}_0 = V_0 = h$.

Pour $k \geq 1$, on procède par récurrence. On suppose que pour un certain $k \in \mathbb{N}^*$ fixé, on a

$$\forall x \in E, \quad \mathcal{V}_k(x) \leq V_k(x) \leq \mathcal{V}_k(x) + k\varepsilon.$$

En appliquant le lemme 5.1.1 à cette inégalité, on obtient :

$$\forall x \in E, \quad K\mathcal{V}_k(x) \leq KV_k(x) \leq K\mathcal{V}_k(x) + k\varepsilon \quad (5.9)$$

et

$$\forall x \in E, \forall t \in \mathbb{R}^+, \quad J(M\mathcal{V}_k, \mathcal{V}_k)(x, t) \leq J(MV_k, V_k)(x, t) \leq J(M\mathcal{V}_k, \mathcal{V}_k)(x, t) + k\varepsilon. \quad (5.10)$$

On distingue deux cas.

Cas 1 : $K\mathcal{V}_k(x) < \inf_{t \in \mathbb{R}^+} J(M\mathcal{V}_k, \mathcal{V}_k)(x, t)$

Par la définition des fonctions valeurs $(\mathcal{V}_k)_{k \in \mathbb{N}}$ (proposition 1.6.1), on a alors

$$\mathcal{V}_{k+1}(x) = K\mathcal{V}_k(x) < \inf_{t \in \mathbb{R}^+} J(M\mathcal{V}_k, \mathcal{V}_k)(x, t). \quad (5.11)$$

L'inégalité (5.9) entraîne

$$\mathcal{V}_{k+1}(x) \leq KV_k(x) \leq \mathcal{V}_{k+1}(x) + k\varepsilon. \quad (5.12)$$

On doit distinguer maintenant les deux sous-cas. Ces derniers surviennent dans la définition des approximations $(V_k)_{k \in \mathbb{N}}$ des fonctions valeurs.

- Si $KV_k(x) < \inf_{t \in \mathbb{R}^+} J(MV_k, V_k)(x, t)$, alors par la définition (5.4) des fonctions $(V_k)_{k \in \mathbb{N}}$, on a $V_{k+1}(x) = KV_k(x)$. Par conséquent, en utilisant l'inégalité (5.12), on obtient le résultat voulu :

$$\mathcal{V}_{k+1}(x) \leq V_{k+1}(x) \leq \mathcal{V}_{k+1}(x) + k\varepsilon.$$

- Si $\inf_{t \in \mathbb{R}^+} J(MV_k, V_k)(x, t) \leq KV_k(x)$, alors la définition (5.4) des fonctions valeurs $(V_k)_{k \in \mathbb{N}}$ entraîne

$$V_{k+1}(x) = J(MV_k, V_k)(x, r_\varepsilon^{k+1}(x)). \quad (5.13)$$

D'une part, l'inégalité (5.10) avec $t = r_\varepsilon^{k+1}(x)$ nous donne

$$J(M\mathcal{V}_k, \mathcal{V}_k)(x, r_\varepsilon^{k+1}(x)) \leq J(MV_k, V_k)(x, r_\varepsilon^{k+1}(x)),$$

ce qui implique, en utilisant l'égalité (5.13)

$$J(M\mathcal{V}_k, \mathcal{V}_k)(x, r_\varepsilon^{k+1}(x)) \leq V_{k+1}(x).$$

On obtient ainsi

$$\inf_{t \in \mathbb{R}^+} J(M\mathcal{V}_k, \mathcal{V}_k)(x, t) \leq V_{k+1}(x)$$

et l'inégalité (5.11) entraîne

$$\mathcal{V}_{k+1}(x) \leq \inf_{t \in \mathbb{R}^+} J(M\mathcal{V}_k, \mathcal{V}_k)(x, t) \leq V_{k+1}(x). \quad (5.14)$$

D'autre part, par la définition (5.3) des fonctions $(r_\varepsilon^k)_{k \in \mathbb{N}}$ et l'égalité (5.13), on a

$$\begin{aligned} V_{k+1}(x) &= J(MV_k, V_k)(x, r_\varepsilon^{k+1}(x)) \\ &\leq \inf_{t \in \mathbb{R}^+} J(MV_k, V_k)(x, t) + \varepsilon \\ &\leq KV_k(x) + \varepsilon, \end{aligned}$$

et la seconde partie de l'inégalité (5.12) implique

$$V_{k+1}(x) \leq \mathcal{V}_{k+1}(x) + (k+1)\varepsilon. \quad (5.15)$$

En combinant les inégalités (5.14) et (5.15), on obtient le résultat souhaité :

$$\mathcal{V}_{k+1}(x) \leq V_{k+1}(x) \leq \mathcal{V}_{k+1}(x) + (k+1)\varepsilon.$$

Cas 2 : $K\mathcal{V}_k(x) \geq \inf_{t \in \mathbb{R}^+} J(M\mathcal{V}_k, \mathcal{V}_k)(x, t)$

Alors la définition des fonctions valeurs $(\mathcal{V}_k)_{k \in \mathbb{N}}$ (proposition 1.6.1) nous donne

$$\mathcal{V}_{k+1}(x) = \inf_{t \in \mathbb{R}^+} J(M\mathcal{V}_k, \mathcal{V}_k)(x, t) \leq K\mathcal{V}_k(x). \quad (5.16)$$

De la même manière que précédemment, on distingue les deux sous-cas intervenant dans la définition des fonctions $(V_k)_{k \in \mathbb{N}}$.

- Si $KV_k(x) < \inf_{t \in \mathbb{R}^+} J(MV_k, V_k)(x, t)$, alors par la définition (5.4) des fonctions $(V_k)_{k \in \mathbb{N}}$, on a

$$V_{k+1}(x) = KV_k(x) < \inf_{t \in \mathbb{R}^+} J(MV_k, V_k)(x, t). \quad (5.17)$$

Puis en utilisant les équations (5.10) et (5.16), on obtient

$$V_{k+1}(x) < \inf_{t \in \mathbb{R}^+} J(M\mathcal{V}_k, \mathcal{V}_k)(x, t) + k\varepsilon = \mathcal{V}_{k+1}(x) + k\varepsilon. \quad (5.18)$$

Par ailleurs, la combinaison des équations (5.16), (5.9) et (5.17) implique

$$\mathcal{V}_{k+1}(x) \leq K\mathcal{V}_k(x) \leq KV_k(x) = V_{k+1}(x). \quad (5.19)$$

Finalement, les inégalités (5.18) et (5.19) amènent au résultat désiré :

$$\mathcal{V}_{k+1}(x) \leq V_{k+1}(x) \leq \mathcal{V}_{k+1}(x) + k\varepsilon.$$

- Si $KV_k(x) \geq \inf_{t \in \mathbb{R}^+} J(MV_k, V_k)(x, t)$, l'équation (5.13) est encore valable. D'une part, la définition (5.3) des fonctions $(r_\varepsilon^k)_{k \in \mathbb{N}}$ nous donne

$$V_{k+1}(x) \leq \inf_{t \in \mathbb{R}^+} J(MV_k, V_k)(x, t) + \varepsilon.$$

Puis en utilisant les équations (5.10) et (5.16), on obtient

$$V_{k+1}(x) \leq \mathcal{V}_{k+1}(x) + (k+1)\varepsilon. \quad (5.20)$$

D'autre part, en repartant de l'égalité (5.16) et en utilisant à nouveau l'inégalité (5.10), on a

$$\begin{aligned} \mathcal{V}_{k+1}(x) &= \inf_{t \in \mathbb{R}^+} J(M\mathcal{V}_k, \mathcal{V}_k)(x, t) \leq \inf_{t \in \mathbb{R}^+} J(MV_k, V_k)(x, t) \\ &\leq J(MV_k, V_k)(x, r_\varepsilon^{k+1}(x)). \end{aligned}$$

et l'équation (5.13) nous donne :

$$\mathcal{V}_{k+1}(x) \leq J(MV_k, V_k)(x, r_\varepsilon^{k+1}(x)) = V_{k+1}(x).$$

Si on complète cette dernière inégalité avec l'inégalité (5.20), on obtient

$$\mathcal{V}_{k+1}(x) \leq V_{k+1}(x) \leq \mathcal{V}_{k+1}(x) + (k+1)\varepsilon.$$

Finalement, dans tous les cas, on obtient le résultat attendu.

$$\mathcal{V}_{k+1}(x) \leq V_{k+1}(x) \leq \mathcal{V}_{k+1}(x) + (k+1)\varepsilon$$

ce qui prouve le théorème. ■

5.2 Construction de stratégies d'intervention explicites

Dans ce paragraphe, on définit un PDMP auxiliaire sur un espace d'état élargi. Ce processus peut être vu comme une version du PDMP original, contrôlée par une famille de stratégies explicites indexées par un certain entier n , où les interventions sont autorisées jusqu'au n -ième temps de saut du processus. Pour toute la suite de ce travail, on considère qu'un saut du processus contrôlé correspond soit à un saut dit *naturel* car issu du processus non contrôlé, soit à une intervention.

Partant du processus $\{X_t\}$, on construit un nouveau PDMP $\{\tilde{X}_t\}$ sur un espace d'état élargi, ayant un flot identique mais un mécanisme de saut différent de $\{X_t\}$. Le nouveau processus a quatre variables : $\tilde{X}_t = (m_t, \zeta_t, n_t, \theta_t)$, pour $t \geq 0$. On considère que $\tilde{m}_t = (m_t, n_t)$ est la variable de mode et $\tilde{\zeta}_t = (\zeta_t, \theta_t)$ est la variable euclidienne du PDMP $\{\tilde{X}_t\}$. L'espace d'état de ce processus est défini par

$$\tilde{E} = \left\{ (m, \zeta, n, \theta) : (m, \zeta) \in E, n \in \mathbb{N}, \theta \in [0; r_\varepsilon^n(m, \zeta) \wedge t^*(m, \zeta)] \right\} \cup \{\Delta\},$$

où Δ est un point isolé, appelé état cimetièr. De la même manière que pour l'espace E , on définit une distance sur $\tilde{E} \setminus \{\Delta\}$ comme suit. Soit (m, ζ, n, θ) et $(m', \zeta', n', \theta')$ appartenant à $\tilde{E} \setminus \{\Delta\}$. On pose $\tilde{m} = (m, n)$, $\tilde{m}' = (m', n')$, $\tilde{\zeta} = (\zeta, \theta)$ et $\tilde{\zeta}' = (\zeta', \theta')$. On définit la distance par

$$|(\tilde{m}, \tilde{\zeta}) - (\tilde{m}', \tilde{\zeta}')| = |\tilde{\zeta} - \tilde{\zeta}'| \mathbb{1}_{\{\tilde{m}=\tilde{m}'\}} + \infty \mathbb{1}_{\{\tilde{m} \neq \tilde{m}'\}}.$$

Soit $(m, \zeta, n, \theta) \in \tilde{E} \setminus \{\Delta\}$. En posant $r_\varepsilon^0(m, \zeta) = +\infty$ et en rappelant que la définition de r_ε^n est donnée par les formules (5.1) et (5.3), on pose $\tilde{r}_\varepsilon(\Delta) = +\infty$ et

$$\tilde{r}_\varepsilon(m, \zeta, n, \theta) = r_\varepsilon^n(m, \zeta). \quad (5.21)$$

Pour $n \in \mathbb{N}^*$, on définit

$$y_\varepsilon^n(m, \zeta) = \arg \min_{y \in \mathbb{U}} \{c(\Phi((m, \zeta), r_\varepsilon^n(m, \zeta)), y) + V_{n-1}(y)\} \quad (5.22)$$

et on pose pour $n \in \mathbb{N}$

$$\tilde{y}_\varepsilon(m, \zeta, n, \theta) = \begin{cases} (y_\varepsilon^n(m, \zeta), n-1, 0) & \text{si } n \neq 0 \\ \Delta & \text{sinon.} \end{cases} \quad (5.23)$$

Remarque 4

Si on suppose que l'ensemble de contrôle \mathbb{U} est un compact et non plus un ensemble fini, il devient difficile d'assurer l'existence de la fonction y_ε . En effet, la continuité des fonctions $(V_k)_{k \in \mathbb{N}^}$ n'est pas prouvée et il s'agit à l'heure actuelle d'un problème ouvert. Par conséquent, on choisit de garder l'ensemble de contrôle \mathbb{U} comme étant simplement un ensemble fini. Ce choix se justifie d'autant plus qu'une des perspectives de ce travail consiste à construire une approximation numérique des stratégies ε -optimales.*

Sans entrer dans les détails, on peut considérer que la fonction \tilde{r}_ε donne le temps éventuel avant la prochaine intervention du processus contrôlé $\{\tilde{X}_t\}$, à partir d'une position donnée, et la fonction \tilde{y}_ε donne la meilleure position de redémarrage du processus contrôlé, après une intervention. Il s'agit là de la nouveauté par rapport à la construction de temps d'arrêt ε -optimaux dans Gugerli (1986). En effet, dans l'arrêt optimal, seul le temps d'arrêt du processus nous importe. On ne souhaite pas faire repartir le processus après l'arrêt. Or dans le contrôle impulsif, le processus continue d'évoluer après une intervention. Il est donc nécessaire d'introduire une nouvelle fonction \tilde{y}_ε , liée à la position après l'intervention et dépendant du temps éventuel \tilde{r}_ε avant la prochaine intervention.

Les caractéristiques locales du processus contrôlé $\{\tilde{X}_t\}$ sont les suivantes. Soit (m, ζ, n, θ) un point de l'espace d'état $\tilde{E} \setminus \{\Delta\}$.

- Le flot $\tilde{\Phi}$ est défini par

$$\tilde{\Phi}((m, \zeta, n, \theta), t) = (m, \Phi_m(\zeta, t), n, \theta + t) \quad (5.24)$$

et on pose $\tilde{\Phi}(\Delta, t) = 0$. Le temps d'atteinte de la frontière de l'espace \tilde{E} par le flot $\tilde{\Phi}$ est noté \tilde{t}^* et est défini par

$$\tilde{t}^*(m, \zeta, n, \theta) = t^*(m, \zeta) \wedge r_\varepsilon^n(m, \zeta). \quad (5.25)$$

- L'intensité de saut $\tilde{\lambda}$ est donnée par

$$\tilde{\lambda}(m, \zeta, n, \theta) = \lambda(m, \zeta), \quad (5.26)$$

et $\tilde{\lambda}(\Delta) = 0$. De plus, en posant,

$$\tilde{\Lambda}(m, \zeta, n, \theta, t) := \int_0^t \tilde{\lambda}(m, \Phi_m(\zeta, s), n, \theta + s) ds,$$

on obtient

$$\tilde{\Lambda}(m, \zeta, n, \theta, t) = \Lambda(m, \zeta, t). \quad (5.27)$$

- Soit $A \in \mathfrak{B}(\tilde{E})$, $B \in \mathfrak{B}(\mathbb{N})$ et $C \in \mathfrak{B}(\mathbb{R}^+)$. Le noyau de Markov \tilde{Q} est défini par

$$\begin{aligned} \tilde{Q}((m, \zeta, n, \theta), A \times B \times C) &= Q((m, \zeta), A) \mathbb{1}_{\{0\}}(C) \\ &\times \left[\mathbb{1}_{\{n=0\}} \mathbb{1}_{\{0\}}(B) + \mathbb{1}_{\{n \geq 1\}} \mathbb{1}_{\{n-1\}}(B) \right] \\ &\times \left(\mathbb{1}_{\{\theta < r_\varepsilon^n(m, \zeta)\}} + \mathbb{1}_{\{\theta = r_\varepsilon^n(m, \zeta) = t^*(m, \zeta)\}} \right) \\ &+ \mathbb{1}_{\{n \geq 1\}} \mathbb{1}_{\{\theta = r_\varepsilon^n(m, \zeta) \neq t^*(m, \zeta)\}} \times \mathbb{1}_{\{\tilde{y}_\varepsilon(m, \zeta, n, \theta)\}}(A \times B \times C), \end{aligned} \quad (5.28)$$

et $\tilde{Q}(\Delta, \Delta) = 1$.

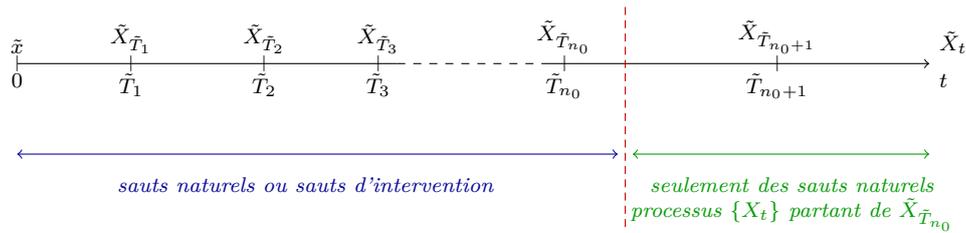


FIGURE 5.1 – Configuration des trajectoires contrôlées avec horizon de contrôle $n_0 \in \mathbb{N}$; les n_0 premiers sauts du processus sont des sauts naturels ou des interventions, alors que les sauts suivants, c'est-à-dire à partir du $(n_0 + 1)$ -ième saut, sont des sauts naturels uniquement.

Il existe un espace de probabilité filtré $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}})_{t \geq 0}, (\tilde{\mathbb{P}}_{\tilde{x}})_{\tilde{x} \in \tilde{E}})$ dans lequel le processus $\{\tilde{X}_t\}$ est un processus de Markov fort (Davis, 1993, sections 24-25).

Les composantes additionnelles du processus $\{\tilde{X}_t\}$ peuvent être interprétées de la manière suivante : θ_t est le temps écoulé depuis le dernier saut ayant eu lieu avant t , et n_t peut être vu comme l'horizon de contrôle restant, ou de manière équivalente comme un compteur décroissant du nombre de sauts qui peuvent être des sauts dits naturels car issus du processus non contrôlé $\{X_t\}$ ou des interventions. Le processus $\{\tilde{X}_t\}$ peut être interprété comme une version contrôlée du processus $\{X_t\}$, où les interventions sont possibles jusqu'au n_0 -ième saut, comme présenté dans la figure 5.1. Autrement dit, les n_0 premiers sauts du processus contrôlé $\{\tilde{X}_t\}$ peuvent être des interventions ou des sauts naturels. Après le n_0 -ième saut qui correspond à un saut naturel ou à une intervention, les interventions ne sont plus permises et le processus suit la même dynamique que le processus initial $\{X_t\}$: le comportement du processus contrôlé à partir du temps T_{n_0} correspond au comportement du processus non contrôlé partant du point $\tilde{X}_{T_{n_0}}$. Ainsi, l'horizon de contrôle n_t décroît d'une unité à chaque saut (naturel ou intervention) jusqu'à atteindre 0, où il reste constant.

Quelques exemples de configurations possibles entre sauts naturels et interventions pour des trajectoires contrôlées avec horizon de contrôle $n_0 \in \mathbb{N}$ sont donnés par la figure 5.2. La figure (a) montre l'exemple d'une trajectoire où le premier saut et le troisième saut sont des interventions, le reste étant des sauts naturels du processus. La figure (b) présente l'exemple d'une trajectoire n'ayant qu'une intervention. Cette dernière a lieu au n_0 -ième saut. Tous les autres sauts du processus contrôlé sont des sauts naturels. Enfin, la figure (c) expose un exemple pour lequel tous les sauts du temps 0 et au temps T_{n_0} compris, sont des interventions. À partir du $(n_0 + 1)$ -ième saut, le processus contrôlé a le même comportement que le processus non contrôlé partant du point $\tilde{X}_{T_{n_0}}$.

L'interprétation du noyau de Markov \tilde{Q} est la suivante. Dans tous les cas, le temps θ écoulé depuis le dernier saut ayant eu lieu avant t , est remis à zéro. De même, le compteur décroissant n du nombre de sauts naturels et d'intervention décroît d'une unité si ce dernier est strictement positif, et reste égal à 0 sinon. Si $n \geq 1$, deux cas de figure peuvent se présenter. Soit le temps inter-sauts est égal au temps éventuel d'intervention $r_\varepsilon^n(m, \zeta)$ mais différent du temps d'atteinte à la frontière de l'espace d'état E associé au processus non contrôlé $t^*(m, \zeta)$. Dans ce cas, on intervient sur le processus et ce dernier redémarre au point $\tilde{y}(m, \zeta, n, \theta)$. Soit le temps inter-sauts est strictement inférieur au temps éventuel d'intervention $r_\varepsilon^n(m, \zeta)$ ou le temps de saut est égal au temps $r_\varepsilon^n(m, \zeta)$ et au temps d'atteinte à la frontière $t^*(m, \zeta)$. Dans ce cas, il n'y a pas d'intervention et le processus évolue librement comme le processus non contrôlé : la nouvelle position du processus contrôlé après le saut suit la distribution donnée par le noyau de Markov Q du processus initial.

On peut décrire la dynamique de saut de manière plus détaillée comme suit. On note π la projection de l'espace $E \times \mathbb{N} \times \mathbb{R}$ sur l'espace E . On note également $(\tilde{S}_n)_{n \in \mathbb{N}^*}$ les temps de séjour

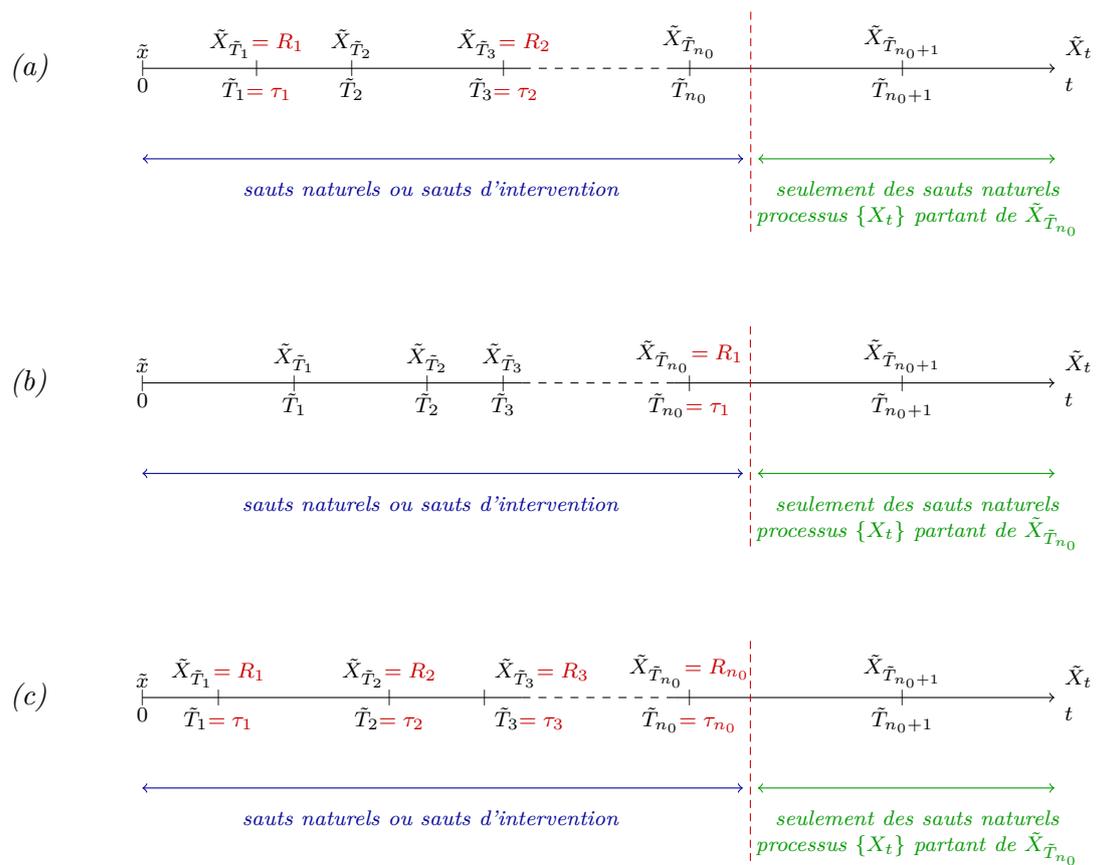


FIGURE 5.2 – Exemples de configurations de trajectoires contrôlées avec horizon de contrôle $n_0 \in \mathbb{N}$.

entre deux sauts consécutifs du processus contrôlé $\{\tilde{X}_t\}$, $(\tilde{Z}_n)_{n \in \mathbb{N}}$ les positions après les sauts et $(\tilde{T}_n)_{n \in \mathbb{N}}$ les temps de sauts du processus $\{\tilde{X}_t\}$. On distingue les deux types de sauts (naturel ou intervention) de la façon suivante. Pour $k \in \mathbb{N}^*$:

- si $\tilde{S}_k < \tilde{r}_\varepsilon(\tilde{Z}_{k-1})$ ou si $\tilde{S}_k = \tilde{r}_\varepsilon(\tilde{Z}_{k-1}) = t^*(\pi(\tilde{Z}_{k-1}))$, c'est-à-dire si le temps inter-saut \tilde{S}_k se produit avant le temps éventuel d'intervention $\tilde{r}_\varepsilon(\tilde{Z}_{k-1})$ ou si le saut a lieu à la frontière de l'espace d'état E , alors \tilde{Z}_k est déterminé par le noyau de Markov Q du processus non contrôlé (équation (5.28)). On interprète le k -ème saut comme étant un *saut naturel* car issu du processus non contrôlé $\{X_t\}$.
- si $\tilde{S}_k = \tilde{r}_\varepsilon(\tilde{Z}_{k-1}) \neq t^*(\pi(\tilde{Z}_{k-1}))$, c'est-à-dire si le saut a lieu au temps éventuel d'intervention $T_{k-1} + \tilde{r}_\varepsilon(\tilde{Z}_{k-1})$ mais n'a pas lieu à la frontière de l'espace d'état E , alors par l'équation (5.28), on a $\tilde{Z}_k = \tilde{y}_\varepsilon(\tilde{Z}_{k-1})$. On interprète le k -ème saut comme une *intervention*. Quand une intervention survient, la nouvelle position de départ du processus $\{\tilde{X}_t\}$ est choisie par les formules (5.28) et (5.23), où l'ensemble \mathbb{U} est l'ensemble de contrôle.

On explicite maintenant la stratégie de contrôle correspondant au processus $\{\tilde{X}_t\}$. On introduit le processus de comptage défini par

$$p^*(t) := \sum_{i=1}^{\infty} \mathbb{1}_{\{\tilde{T}_i \leq t\}} \mathbb{1}_{\{\tilde{X}_{\tilde{T}_i} \in \partial \tilde{E}\}} \mathbb{1}_{\{\tilde{S}_i \neq t^*(\pi(\tilde{Z}_{i-1}))\}}. \quad (5.29)$$

Ce processus compte le nombre de sauts correspondant à des interventions. Posons

$$\tilde{\tau}_i = \inf\{t \in \mathbb{R}^+ \mid p^*(t) = i\} \quad (5.30)$$

le temps de la i -ème intervention et définissons la position de redémarrage comme étant la suivante

$$\tilde{R}_i = \begin{cases} \tilde{Z}_{\tilde{\tau}_i} & \text{si } \tilde{\tau}_i < \infty \\ \Delta & \text{sinon.} \end{cases}$$

Ainsi, la stratégie associée au processus contrôlé $\{\tilde{X}_t\}$, démarrant au point $(x, n_0, 0)$ où $x \in E$ et $n_0 \in \mathbb{N}$, correspond à la suite

$$\tilde{\mathcal{S}}_\varepsilon^{n_0} = (\tilde{\tau}_i, \tilde{R}_i)_{i \in \mathbb{N}^*}.$$

La figure 5.3 montre les premières étapes de la construction d'un processus contrôlé par la stratégie $\tilde{\mathcal{S}}_\varepsilon^{n_0}$. Au temps $t = 0$ (figure (a)), on calcule le temps éventuel d'intervention $r_\varepsilon^{n_0}(x_0)$. Le processus $\{\tilde{X}_t\}$ suit le flot $\tilde{\Phi}$ jusqu'au temps \tilde{T}_1 (figure (b)). Comme ce temps est strictement inférieur au temps possible d'intervention, le premier saut du processus est un saut naturel. Le nouveau point de départ du processus est alors déterminé par le noyau Markovien \tilde{Q} (figure (c)). La nouvelle position est notée $\tilde{X}_{\tilde{T}_1} = (x_1, n_0 - 1, 0)$ et le processus se trouve dans un nouveau mode m' . Au temps \tilde{T}_1 , on calcule le nouveau temps éventuel d'intervention $\tilde{T}_1 + r_\varepsilon^{n_0-1}(x_1)$ (figure (d)). Le processus contrôlé suit une fois de plus le flot $\tilde{\Phi}$ jusqu'au temps de saut \tilde{T}_2 . Dans notre exemple, ce dernier correspond au temps éventuel d'intervention mais est différent du temps d'atteinte à la frontière de l'espace d'état $t^*(x)$ (figure (e)). Par conséquent, on intervient au temps \tilde{T}_2 et la nouvelle position du processus contrôlé est choisie de manière déterministe par la fonction $y_\varepsilon^{n_0-1}$ en (x_1) (figure (f)).

S'assurer que la stratégie $\tilde{\mathcal{S}}_\varepsilon^{n_0}$ ainsi définie est admissible est un problème très difficile, qui ne sera pas discuté dans cette thèse. En effet, d'un point de vue théorique, l'admissibilité d'une stratégie requiert la vérification de cinq conditions, données dans l'article [Costa et Davis \(1989\)](#), section 2.3. Ces conditions sont très difficiles à vérifier et demandent un travail conséquent, comme souligné dans ce même article (remarque précédant la proposition 6). D'un point de vue pratique, sans entrer dans les détails, l'admissibilité d'une stratégie est relative au fait qu'elle doit être non anticipative et doit dépendre uniquement de variables calculables. Les stratégies $(\tilde{\mathcal{S}}_\varepsilon^n)_{n \in \mathbb{N}}$ construites dans ce paragraphe vérifient bien ces deux conditions.

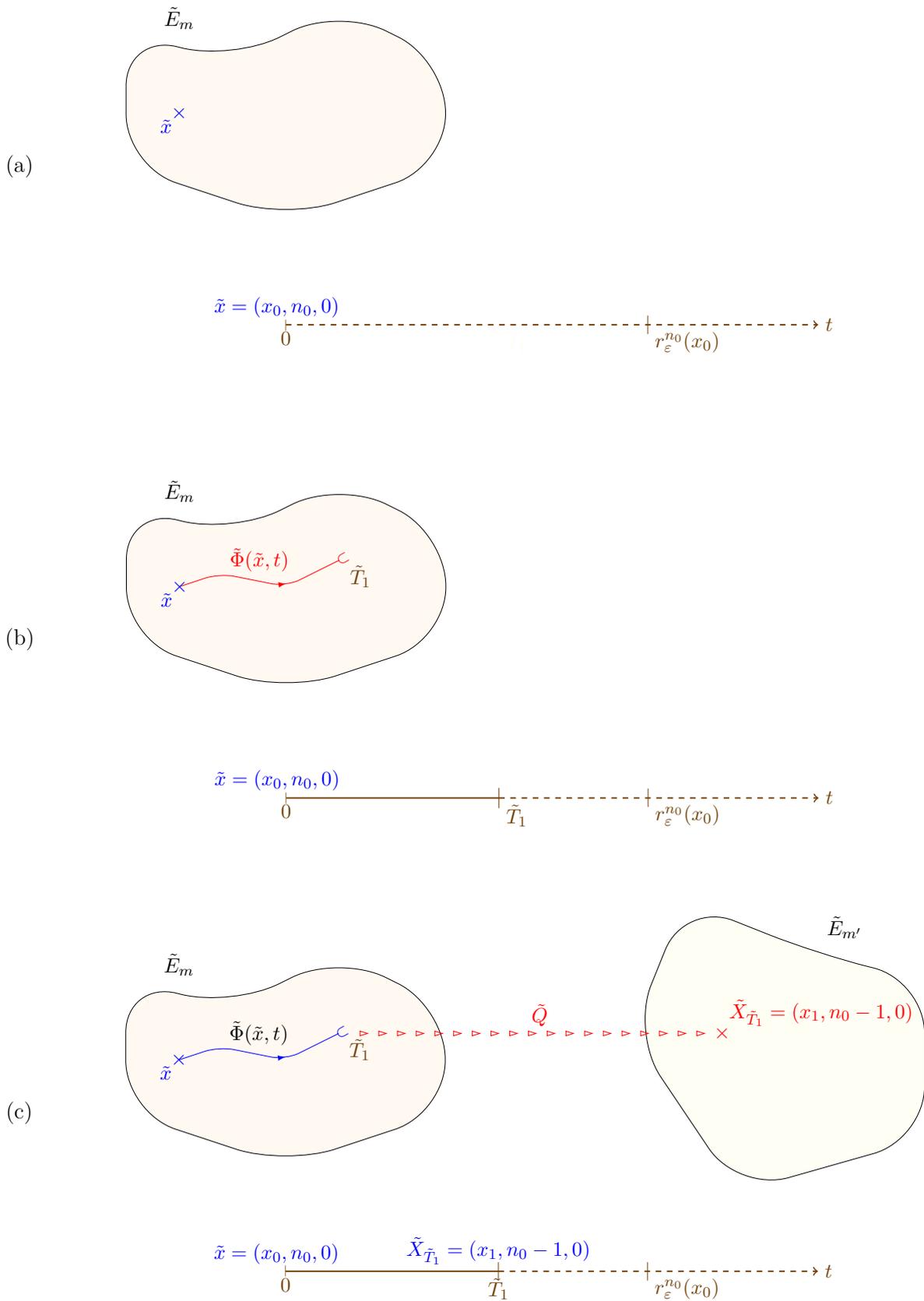


FIGURE 5.3 – (Partie 1) Soit $n_0 \in \mathbb{N}^*$. La figure présente un exemple des premières étapes de la construction d'un processus $\{\tilde{X}_t\}$ partant du point $\tilde{x} = (x_0, n_0, 0) \in \tilde{E}$ et contrôlé par la stratégie $\tilde{S}_\varepsilon^{n_0}$.

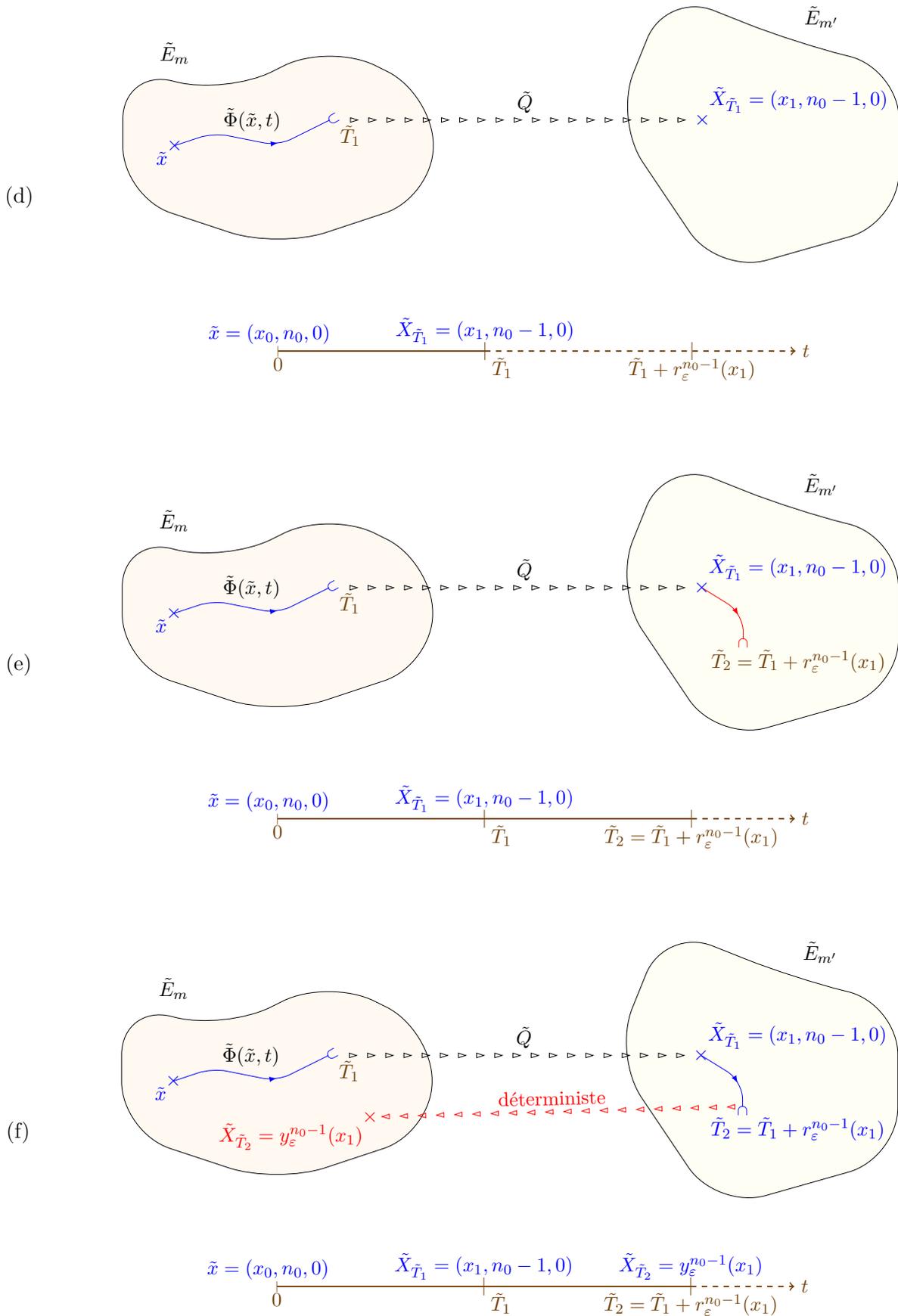


FIGURE 5.3 – (Partie 2) Soit $n_0 \in \mathbb{N}^*$. La figure présente un exemple des premières étapes de la construction d'un processus $\{\tilde{X}_t\}$ partant du point $\tilde{x} = (x_0, n_0, 0) \in \tilde{E}$ et contrôlé par la stratégie $\tilde{S}_\varepsilon^{n_0}$.

5.3 Coût des trajectoires contrôlées

On établit dans ce paragraphe le résultat le plus important de ce chapitre. On démontre que le coût associé aux stratégies $(\tilde{\mathcal{S}}_\varepsilon^n)_{n \in \mathbb{N}}$ correspond aux approximations $(V_n)_{n \in \mathbb{N}^*}$ des fonctions valeurs. Pour cela, on procède en plusieurs étapes. Tout d'abord, on montre au paragraphe 5.3.1 que sous une certaine condition, le temps éventuel d'intervention donné par la fonction r_ε^k est toujours strictement inférieur au temps d'atteinte à la frontière t^* . On exprime ensuite au paragraphe 5.3.2 la distribution de la première position et du premier temps inter-sauts du processus contrôlé avec ceux du processus non contrôlé. Au paragraphe 5.3.3, on démontre un résultat important qui concerne la propriété de Markov des temps d'intervention de nos stratégies $(\tilde{\mathcal{S}}_\varepsilon^n)_{n \in \mathbb{N}}$. Enfin, ces trois résultats nous permettent de prouver au paragraphe 5.3.4 que le coût associé aux stratégies $(\tilde{\mathcal{S}}_\varepsilon^n)_{n \in \mathbb{N}}$ est égal aux fonctions $(V_n)_{n \in \mathbb{N}^*}$.

5.3.1 Majoration du temps éventuel d'intervention

À $k \in \mathbb{N}^*$, par la définition 1.7 de l'opérateur J et par la définition 5.3 de la fonction r_ε^k , le temps éventuel d'intervention est toujours inférieur ou égal au temps d'atteinte à la frontière :

$$\forall x \in E, \quad r_\varepsilon^k \leq t^*(x).$$

Cependant, pour prouver que le coût associé aux stratégies $(\tilde{\mathcal{S}}_\varepsilon^n)_{n \in \mathbb{N}}$ est égal aux fonctions $(V_n)_{n \in \mathbb{N}^*}$, il nous est nécessaire de ne pas avoir le cas d'égalité quand la condition $KV_{k-1}(x) \geq \inf_{t \in \mathbb{R}^+} J(MV_{k-1}, V_{k-1})(x, t)$ est réalisée. Ce résultat fait l'objet de la propriété 5.3.1. Sa démonstration requiert un résultat de continuité sur \mathbb{R}^+ de la fonction $J(MV_k, V_k)(x, \cdot)$ qui correspond au lemme suivant.

Lemme 5.3.1

Pour tout $k \in \mathbb{N}$ et pour tout $x \in E$, la fonction $J(MV_k, V_k)(x, \cdot)$ est continue sur \mathbb{R}^+ .

Démonstration : On fixe $k \in \mathbb{N}$ et $x \in E$. Par définition (1.7) de J , on a, pour tout $t \in \mathbb{R}^+$,

$$\begin{aligned} J(MV_k, V_k)(x, t) &= F(x, t) + \mathbb{E}_x \left[e^{-\alpha(t \wedge t^*(x))} MV_k(\Phi(x, t \wedge t^*(x))) \mathbb{1}_{\{S_1 \geq t \wedge t^*(x)\}} \right] \\ &\quad + \mathbb{E}_x \left[e^{-\alpha S_1} V_k(Z_1) \mathbb{1}_{\{S_1 < t \wedge t^*(x)\}} \right]. \end{aligned}$$

Pour montrer la continuité de $J(MV_k, V_k)(x, \cdot)$ sur \mathbb{R}^+ , on décompose J en trois parties que l'on étudie indépendamment. La première partie de J est la fonction $F(x, \cdot)$ définie par (1.6), qui est clairement continue sur \mathbb{R}^+ . La seconde partie de J correspond à la première espérance qu'elle contient. Par la formule (1.3), elle est égale à

$$\begin{aligned} \mathbb{E}_x \left[e^{-\alpha(t \wedge t^*(x))} MV_k(\Phi(x, t \wedge t^*(x))) \mathbb{1}_{\{S_1 \geq t \wedge t^*(x)\}} \right] \\ &= e^{-\alpha(t \wedge t^*(x))} MV_k(\Phi(x, t \wedge t^*(x))) \mathbb{P}_x(\mathbb{1}_{\{S_1 \geq t \wedge t^*(x)\}}) \\ &= e^{-\alpha(t \wedge t^*(x))} MV_k(\Phi(x, t \wedge t^*(x))) e^{-\Lambda(x, t \wedge t^*(x))} \end{aligned}$$

D'une part, la fonction $t \mapsto e^{-\alpha(t \wedge t^*(x))} e^{-\Lambda(x, t \wedge t^*(x))}$ est continue sur \mathbb{R}^+ par définition (1.2) de Λ . D'autre part, en utilisant la définition (1.9) de l'opérateur M , comme $\Phi(x, \cdot)$ est continue sur \mathbb{R}^+ et c est continue sur $E \times \mathbb{U}$, la fonction $t \mapsto MV_k(\Phi(x, t \wedge t^*(x)))$ est continue sur \mathbb{R}^+ . Donc la seconde partie de J est continue sur \mathbb{R}^+ .

La dernière partie de J est continue sur \mathbb{R}^+ par le théorème de continuité d'une intégrale à paramètres. On en déduit que la fonction $t \mapsto J(MV_k, V_k)(x, \cdot)$ est continue sur \mathbb{R}^+ . ■

Remarque 5

On peut remarquer que l'indépendance de l'ensemble de contrôle \mathbb{U} par rapport à la position x du processus, simplifie la preuve de continuité de la fonction $t \mapsto MV_k(\Phi(x, t \wedge t^*(x)))$

sur \mathbb{R}^+ . Si l'ensemble de contrôle \mathbb{U} dépendait de x , il serait encore possible de montrer la continuité de cette fonction en ajoutant de nouvelles hypothèses mais cela rendrait les choses plus sophistiquées.

Proposition 5.3.1

Soit $k \in \mathbb{N}^*$ et $x \in E$ tels que $KV_{k-1}(x) \geq \inf_{t \in \mathbb{R}^+} J(MV_{k-1}, V_{k-1})(x, t)$. Alors on a

$$r_\varepsilon^k(x) < t^*(x).$$

Démonstration : Comme vu précédemment, l'inégalité $r_\varepsilon^k(x) \leq t^*(x)$ est toujours vraie. Supposons que $r_\varepsilon^k(x) = t^*(x)$. Alors la définition de r_ε^k entraîne

$$J(MV_{k-1}, V_{k-1})(x, t^*(x)) < \inf_{t \in \mathbb{R}^+} J(MV_{k-1}, V_{k-1})(x, t) + \varepsilon, \quad (5.31)$$

et pour tout $s \in [0; t^*(x)[$,

$$J(MV_{k-1}, V_{k-1})(x, s) \geq \inf_{t \in \mathbb{R}^+} J(MV_{k-1}, V_{k-1})(x, t) + \varepsilon. \quad (5.32)$$

On définit $d = \inf_{t \in \mathbb{R}^+} J(MV_{k-1}, V_{k-1})(x, t) + \varepsilon - J(MV_{k-1}, V_{k-1})(x, t^*(x))$. L'inéquation (5.31) montre que $d > 0$. Par continuité de $J(MV_{k-1}, V_{k-1})(x, \cdot)$ sur \mathbb{R}^+ (lemme 5.3.1), il existe $\eta > 0$ tel que pour tout $s \in [t^*(x) - \eta; t^*(x)[$

$$|J(MV_{k-1}, V_{k-1})(x, s) - J(MV_{k-1}, V_{k-1})(x, t^*(x))| < d. \quad (5.33)$$

Considérons $s \in [t^*(x) - \eta; t^*(x)[$. Deux cas sont possibles.

- Si $J(MV_{k-1}, V_{k-1})(x, s) \leq J(MV_{k-1}, V_{k-1})(x, t^*(x))$, alors par l'inéquation (5.31) on a

$$J(MV_{k-1}, V_{k-1})(x, s) < \inf_{t \in \mathbb{R}^+} J(MV_{k-1}, V_{k-1})(x, t) + \varepsilon,$$

contredisant l'inéquation (5.32).

- Si $J(MV_{k-1}, V_{k-1})(x, s) > J(MV_{k-1}, V_{k-1})(x, t^*(x))$, alors par (5.33) et par définition de d , on obtient

$$J(MV_{k-1}, V_{k-1})(x, s) < \inf_{t \in \mathbb{R}^+} J(MV_{k-1}, V_{k-1})(x, t) + \varepsilon,$$

contredisant encore l'inéquation (5.32).

On en conclut que $r_\varepsilon^k(x) \neq t^*(x)$ dans les deux cas. ■

5.3.2 Distributions de la première position et du premier temps de saut pour les processus contrôlés et non contrôlés

On établit une relation entre la distribution de la position du processus contrôlé $\{\tilde{X}_t\}$ après le premier saut et son premier temps de séjour $(\tilde{Z}_1, \tilde{S}_1)$ et la distribution de la position du processus non contrôlé $\{X_t\}$ après le premier saut et son premier temps de séjour (Z_1, S_1) . Ce résultat nous servira dans la démonstration des théorèmes 5.3.2 et 5.3.3.

Lemme 5.3.2

Soit $\tilde{x} \in \tilde{E}$ tel que $\tilde{x} = (x, n_0, 0)$, où $x \in E$ et $n_0 \in \mathbb{N}$. Soit $\tilde{v} \in \mathbf{B}(\tilde{E} \times \mathbb{R}^+)$. On note :

$$\forall x \in E, \forall n \in \mathbb{N}, \forall t \in \mathbb{R}^+, \quad \tilde{v}(x, n, 0, t) = v(x, n, t).$$

Alors on a

$$\begin{aligned} \tilde{\mathbb{E}}_{\tilde{x}}[\tilde{v}(\tilde{Z}_1, \tilde{S}_1)] &= \mathbb{E}_x \left[\mathbb{1}_{\{n_0=0\}} v(Z_1, 0, S_1) + (\mathbb{1}_{\{S_1 < t^*(x) \wedge r_\varepsilon^{n_0}(x)\}} \mathbb{1}_{\{n_0 \neq 0\}} + \mathbb{1}_{\{S_1 = t^*(x) = r_\varepsilon^{n_0}(x)\}}) \right. \\ &\quad \left. \times v(Z_1, n_0 - 1, S_1) + \mathbb{1}_{\{S_1 \geq r_\varepsilon^{n_0}(x) \neq t^*(x)\}} v(y_\varepsilon^{n_0}(x), n_0 - 1, r_\varepsilon^{n_0}(x)) \right]. \end{aligned}$$

Démonstration : La distribution des variables \tilde{Z}_1 et \tilde{S}_1 implique

$$\begin{aligned} \tilde{\mathbb{E}}_{\tilde{x}}[\tilde{v}(\tilde{Z}_1, \tilde{S}_1)] &= \int_{\tilde{E}} \int_0^{\tilde{t}^*(\tilde{x})} \tilde{v}(\tilde{z}, \tilde{s}) \tilde{\lambda}(\tilde{\Phi}(\tilde{x}, \tilde{s})) e^{-\tilde{\Lambda}(\tilde{x}, \tilde{s})} \tilde{Q}(\tilde{\Phi}(\tilde{x}, \tilde{s}), d\tilde{z}) d\tilde{s} \\ &\quad + \int_{\tilde{E}} \tilde{v}(\tilde{z}, \tilde{t}^*(\tilde{x})) e^{-\tilde{\Lambda}(\tilde{x}, \tilde{t}^*(\tilde{x}))} \tilde{Q}(\tilde{\Phi}(\tilde{x}, \tilde{t}^*(\tilde{x})), d\tilde{z}). \end{aligned}$$

Par construction du processus $\{\tilde{X}\}$, la variable aléatoire \tilde{Z}_1 peut s'écrire presque sûrement par : $\tilde{Z}_1 = (Z, (n_0 - 1) \wedge 0, 0)$, où Z est une variable aléatoire à valeurs dans E . En utilisant les définitions (5.24) à (5.28) de $\tilde{\Phi}$, \tilde{t}^* , $\tilde{\lambda}$, $\tilde{\Lambda}$ et \tilde{Q} , on obtient

$$\begin{aligned} \tilde{\mathbb{E}}_{\tilde{x}}[\tilde{v}(\tilde{Z}_1, \tilde{S}_1)] &= \int_E \int_0^{t^*(x) \wedge r_\varepsilon^{n_0}(x)} v(z, (n_0 - 1) \wedge 0, s) \lambda(\Phi(x, s)) e^{-\Lambda(x, s)} Q(\Phi(x, s), dz) ds \\ &\quad + \mathbb{1}_{\{r_\varepsilon^{n_0}(x) \geq t^*(x)\}} e^{-\Lambda(x, t^*(x))} \int_E v(z, (n_0 - 1) \wedge 0, t^*(x)) Q(\Phi(x, t^*(x)), dz) \\ &\quad + \mathbb{1}_{\{r_\varepsilon^{n_0}(x) < t^*(x)\}} e^{-\Lambda(x, r_\varepsilon^{n_0}(x))} v(y_\varepsilon^{n_0}(x), n_0 - 1, r_\varepsilon^{n_0}(x)). \end{aligned}$$

La formule (1.3) pour $t = r_\varepsilon^{n_0}$ et la continuité de la loi de S_1 sur $[0; t^*(x)[$ nous donnent

$$e^{-\Lambda(x, r_\varepsilon^{n_0}(x))} \mathbb{1}_{\{r_\varepsilon^{n_0}(x) < t^*(x)\}} = \mathbb{P}_x(S_1 \geq r_\varepsilon^{n_0}(x) \neq t^*(x)).$$

De plus, la loi de S_1 est telle que $e^{-\Lambda(x, t^*(x))} = \mathbb{P}_x(S_1 = t^*(x))$. Ainsi

$$\begin{aligned} \tilde{\mathbb{E}}_{\tilde{x}}[\tilde{v}(\tilde{Z}_1, \tilde{S}_1)] &= \mathbb{E}_x \left[\mathbb{1}_{\{S_1 < t^*(x) \wedge r_\varepsilon^{n_0}(x)\}} v(Z_1, (n_0 - 1) \wedge 0, S_1) \right. \\ &\quad + \mathbb{1}_{\{r_\varepsilon^{n_0}(x) \geq t^*(x)\}} \mathbb{1}_{\{S_1 = t^*(x)\}} v(Z_1, (n_0 - 1) \wedge 0, S_1) \\ &\quad \left. + \mathbb{1}_{\{S_1 \geq r_\varepsilon^{n_0}(x) \neq t^*(x)\}} v(y_\varepsilon^{n_0}(x), n_0 - 1, r_\varepsilon^{n_0}(x)) \right]. \end{aligned}$$

Par la définition (5.3) de $r_\varepsilon^{n_0}$, on a toujours $r_\varepsilon^{n_0}(x) \leq t^*(x)$ pour $n_0 \geq 1$. Pour $n_0 = 0$, on a posé $r_\varepsilon^{n_0}(x) = +\infty$ dans la section 5.2. Par ailleurs, $t^*(x) < +\infty$ car on a supposé que le temps de sortie t^* est borné (hypothèse 2). On a donc l'égalité $\mathbb{1}_{\{r_\varepsilon^{n_0}(x) \geq t^*(x)\}} = \mathbb{1}_{\{r_\varepsilon^{n_0}(x) = t^*(x)\}} + \mathbb{1}_{\{n_0=0\}}$, qui nous permet d'écrire

$$\begin{aligned} \tilde{\mathbb{E}}_{\tilde{x}}[\tilde{v}(\tilde{Z}_1, \tilde{S}_1)] &= \mathbb{E}_x \left[\mathbb{1}_{\{S_1 < t^*(x) \wedge r_\varepsilon^{n_0}(x)\}} v(Z_1, (n_0 - 1) \wedge 0, S_1) + \mathbb{1}_{\{S_1 = t^*(x) = r_\varepsilon^{n_0}(x)\}} v(Z_1, n_0 - 1, S_1) \right. \\ &\quad + \mathbb{1}_{\{n_0=0\}} \mathbb{1}_{\{S_1 = t^*(x)\}} v(Z_1, 0, S_1) \\ &\quad \left. + \mathbb{1}_{\{S_1 \geq r_\varepsilon^{n_0}(x) \neq t^*(x)\}} v(y_\varepsilon^{n_0}(x), n_0 - 1, r_\varepsilon^{n_0}(x)) \right] \\ &= \mathbb{E}_x \left[\mathbb{1}_{\{n_0=0\}} v(Z_1, 0, S_1) + \mathbb{1}_{\{S_1 < t^*(x) \wedge r_\varepsilon^{n_0}(x)\}} \mathbb{1}_{\{n_0 \neq 0\}} v(Z_1, n_0 - 1, S_1) \right. \\ &\quad + \mathbb{1}_{\{S_1 = t^*(x) = r_\varepsilon^{n_0}(x)\}} v(Z_1, n_0 - 1, S_1) \\ &\quad \left. + \mathbb{1}_{\{S_1 \geq r_\varepsilon^{n_0}(x) \neq t^*(x)\}} v(y_\varepsilon^{n_0}(x), n_0 - 1, r_\varepsilon^{n_0}(x)) \right]. \end{aligned}$$

Le théorème est ainsi démontré. ■

5.3.3 Propriété de Markov des temps d'intervention

On montre maintenant un résultat important de propriété de Markov pour les temps d'interventions associées aux stratégies $(\tilde{S}_\varepsilon^n)_{n \in \mathbb{N}}$. Ce résultat servira lors de l'application de la propriété de Markov dans la démonstration du théorème 5.3.3.

Proposition 5.3.2

Soient $f \in \mathbf{B}(\mathbb{R}^+)$ et $\tilde{x} \in \tilde{E}$ tel que $\tilde{x} = (x, n, 0)$, où $x \in E$ et $n \in \mathbb{N}$. Alors pour tout $i \in \mathbb{N}^*$, on a

$$\begin{aligned} \tilde{\mathbb{E}}_{\tilde{x}}[f(\tilde{\tau}_i)|\tilde{\mathcal{F}}_{\tilde{S}_1}] &= (\mathbb{1}_{\{\tilde{S}_1 < r_\varepsilon^n(x) \wedge t^*(x)\}} + \mathbb{1}_{\{\tilde{S}_1 = r_\varepsilon^n(x) = t^*(x)\}}) \tilde{\mathbb{E}}_{\tilde{Z}_1}[f(\tilde{\tau}_i + \tilde{S}_1)] \\ &\quad + \mathbb{1}_{\{\tilde{S}_1 = r_\varepsilon^n(x) \neq t^*(x)\}} \tilde{\mathbb{E}}_{\tilde{Z}_1}[f(\tilde{\tau}_{i-1} + r_\varepsilon^n(x))]. \end{aligned}$$

Démonstration : La distribution du premier temps inter-saut \tilde{S}_1 du processus contrôlé, implique que $\tilde{S}_1 \leq r_\varepsilon^n(x) \wedge t^*(x)$. On a ainsi :

$$\begin{aligned} \tilde{\mathbb{E}}_{\tilde{x}}[f(\tilde{\tau}_i)|\tilde{\mathcal{F}}_{\tilde{S}_1}] &= (\mathbb{1}_{\{\tilde{S}_1 < r_\varepsilon^n(x) \wedge t^*(x)\}} + \mathbb{1}_{\{\tilde{S}_1 = r_\varepsilon^n(x) \wedge t^*(x)\}}) \times \tilde{\mathbb{E}}_{\tilde{x}}[f(\tilde{\tau}_i)|\tilde{\mathcal{F}}_{\tilde{S}_1}] \\ &= (\mathbb{1}_{\{\tilde{S}_1 < r_\varepsilon^n(x) \wedge t^*(x)\}} + \mathbb{1}_{\{\tilde{S}_1 = r_\varepsilon^n(x) = t^*(x)\}} + \mathbb{1}_{\{\tilde{S}_1 = r_\varepsilon^n(x) \neq t^*(x)\}}) \times \tilde{\mathbb{E}}_{\tilde{x}}[f(\tilde{\tau}_i)|\tilde{\mathcal{F}}_{\tilde{S}_1}], \end{aligned}$$

car le cas $\tilde{S}_1 = t^*(x) \neq r_\varepsilon^n(x)$ est impossible. En effet, on a toujours l'inégalité $r_\varepsilon^n(x) \leq t^*(x)$. La définition des temps d'intervention $(\tilde{\tau}_i)_{i \in \mathbb{N}^*}$ (formules (5.29) et (5.30)) permet d'écrire alors

$$\begin{aligned} \tilde{\mathbb{E}}_{\tilde{x}}[f(\tilde{\tau}_i)|\tilde{\mathcal{F}}_{\tilde{S}_1}] &= (\mathbb{1}_{\{\tilde{S}_1 < r_\varepsilon^n(x) \wedge t^*(x)\}} + \mathbb{1}_{\{\tilde{S}_1 = r_\varepsilon^n(x) = t^*(x)\}} + \mathbb{1}_{\{\tilde{S}_1 = r_\varepsilon^n(x) \neq t^*(x)\}}) \\ &\quad \times \tilde{\mathbb{E}}_{\tilde{x}} \left[f \left(\inf \left\{ t \in \mathbb{R}^+; \sum_{k=1}^{\infty} \mathbb{1}_{\{\tilde{T}_k \leq t\}} \mathbb{1}_{\{\tilde{X}_{\tilde{T}_k} \in \partial \tilde{E}\}} \mathbb{1}_{\{\tilde{S}_k \neq t^*(\pi(\tilde{Z}_{k-1}))\}} = i \right\} \right) \middle| \tilde{\mathcal{F}}_{\tilde{S}_1} \right]. \end{aligned}$$

On s'intéresse au premier terme de la somme dans l'espérance ci-dessus. En premier lieu, on peut remarquer que dans tous les cas, comme i est non nul, on a toujours $\tilde{\tau}_i \geq \tilde{S}_1$. Par conséquent, on obtient $\mathbb{1}_{\{\tilde{T}_1 \leq t\}} = 1$. On distingue maintenant trois cas :

- Si $\tilde{S}_1 < r_\varepsilon^n(x) \wedge t^*(x)$, alors $\tilde{S}_1 \neq \tilde{t}^*(\tilde{x})$, donc $\tilde{X}_{\tilde{T}_1} \notin \partial \tilde{E}$ et on peut conclure que le premier terme de la somme est nul.
- Si $\tilde{S}_1 = r_\varepsilon^n(x) = t^*(x)$, alors $\tilde{S}_1 = t^*(\pi(\tilde{Z}_0))$, donc le premier terme de la somme est nul.
- Si $\tilde{S}_1 = r_\varepsilon^n(x) \neq t^*(x)$, alors $\tilde{S}_1 \neq t^*(\pi(\tilde{Z}_0))$. De plus, $\tilde{S}_1 = \tilde{t}^*(\tilde{x})$ car on a toujours $r_\varepsilon^n(x) \leq t^*(x)$. Donc $\tilde{X}_{\tilde{T}_1} \in \partial \tilde{E}$ et on peut conclure que le premier terme dans la somme est égal à 1.

Ainsi on a

$$\begin{aligned} \tilde{\mathbb{E}}_{\tilde{x}}[f(\tilde{\tau}_i)|\tilde{\mathcal{F}}_{\tilde{S}_1}] &= (\mathbb{1}_{\{\tilde{S}_1 < r_\varepsilon^n(x) \wedge t^*(x)\}} + \mathbb{1}_{\{\tilde{S}_1 = r_\varepsilon^n(x) = t^*(x)\}}) \\ &\quad \times \tilde{\mathbb{E}}_{\tilde{x}} \left[f \left(\inf \left\{ t \in \mathbb{R}^+ \middle| \sum_{k=2}^{\infty} \mathbb{1}_{\{\tilde{T}_k \leq t\}} \mathbb{1}_{\{\tilde{X}_{\tilde{T}_k} \in \partial \tilde{E}\}} \mathbb{1}_{\{\tilde{S}_k \neq t^*(\pi(\tilde{Z}_{k-1}))\}} = i \right\} \right) \middle| \tilde{\mathcal{F}}_{\tilde{S}_1} \right] \\ &\quad + \mathbb{1}_{\{\tilde{S}_1 = r_\varepsilon^n(x) \neq t^*(x)\}} \\ &\quad \times \tilde{\mathbb{E}}_{\tilde{x}} \left[f \left(\inf \left\{ t \in \mathbb{R}^+ \middle| \sum_{k=2}^{\infty} \mathbb{1}_{\{\tilde{T}_k \leq t\}} \mathbb{1}_{\{\tilde{X}_{\tilde{T}_k} \in \partial \tilde{E}\}} \mathbb{1}_{\{\tilde{S}_k \neq t^*(\pi(\tilde{Z}_{k-1}))\}} = i - 1 \right\} \right) \middle| \tilde{\mathcal{F}}_{r_\varepsilon^n(x)} \right]. \end{aligned}$$

Comme le processus $\{(\tilde{Z}_n, \tilde{S}_n)_n\}$ est une chaîne de Markov, la propriété de Markov nous donne

$$\begin{aligned} \tilde{\mathbb{E}}_{\tilde{x}}[f(\tilde{\tau}_i)|\tilde{\mathcal{F}}_{\tilde{S}_1}] &= (\mathbb{1}_{\{\tilde{S}_1 < r_\varepsilon^n(x) \wedge t^*(x)\}} + \mathbb{1}_{\{\tilde{S}_1 = r_\varepsilon^n(x) = t^*(x)\}}) \\ &\quad \times \tilde{\mathbb{E}}_{\tilde{x}} \left[f \left(\inf \left\{ t \in \mathbb{R}^+ \middle| \sum_{k=2}^{\infty} \mathbb{1}_{\{\tilde{T}_{k-1} \leq t\}} \mathbb{1}_{\{\tilde{X}_{\tilde{T}_{k-1}} \in \partial \tilde{E}\}} \mathbb{1}_{\{\tilde{S}_{k-1} \neq t^*(\pi(\tilde{Z}_{k-2}))\}} = i \right\} + \tilde{S}_1 \right) \right] \\ &\quad + \mathbb{1}_{\{\tilde{S}_1 = r_\varepsilon^n(x) \neq t^*(x)\}} \tilde{\mathbb{E}}_{\tilde{x}} \left[f \left(\inf \left\{ t \in \mathbb{R}^+ \middle| \sum_{k=2}^{\infty} \mathbb{1}_{\{\tilde{T}_{k-1} \leq t\}} \right. \right. \right. \\ &\quad \left. \left. \times \mathbb{1}_{\{\tilde{X}_{\tilde{T}_{k-1}} \in \partial \tilde{E}\}} \mathbb{1}_{\{\tilde{S}_{k-1} \neq t^*(\pi(\tilde{Z}_{k-2}))\}} = i - 1 \right\} + r_\varepsilon^n(x) \right) \right]. \end{aligned}$$

En changeant l'indexation dans les sommes, on obtient

$$\begin{aligned} \tilde{\mathbb{E}}_{\tilde{x}}[f(\tilde{\tau}_i)|\tilde{\mathcal{F}}_{\tilde{S}_1}] &= (\mathbb{1}_{\{\tilde{S}_1 < r_\varepsilon^n(x) \wedge t^*(x)\}} + \mathbb{1}_{\{\tilde{S}_1 = r_\varepsilon^n(x) = t^*(x)\}}) \\ &\times \tilde{\mathbb{E}}_{\tilde{x}} \left[f \left(\inf \left\{ t \in \mathbb{R}^+ \mid \sum_{k=1}^{\infty} \mathbb{1}_{\{\tilde{T}_k \leq t\}} \mathbb{1}_{\{\tilde{X}_{\tilde{T}_k} \in \partial \tilde{E}\}} \mathbb{1}_{\{\tilde{S}_k \neq t^*(\pi(\tilde{Z}_{k-1}))\}} = i \right\} + \tilde{S}_1 \right) \right] \\ &+ \mathbb{1}_{\{\tilde{S}_1 = r_\varepsilon^n(x) \neq t^*(x)\}} \tilde{\mathbb{E}}_{\tilde{x}} \left[f \left(\inf \left\{ t \in \mathbb{R}^+ \mid \sum_{k=1}^{\infty} \mathbb{1}_{\{\tilde{T}_k \leq t\}} \right. \right. \right. \\ &\left. \left. \left. \times \mathbb{1}_{\{\tilde{X}_{\tilde{T}_k} \in \partial \tilde{E}\}} \mathbb{1}_{\{\tilde{S}_k \neq t^*(\pi(\tilde{Z}_{k-1}))\}} = i - 1 \right\} + r_\varepsilon^n(x) \right) \right]. \end{aligned}$$

Pour finir, on utilise la définition (5.30) des temps d'intervention, ce qui nous donne le résultat souhaité. \blacksquare

5.3.4 Coût associé aux stratégies explicites construites

Dans cette partie, on s'intéresse au calcul du coût de la stratégie $\tilde{\mathcal{S}}_\varepsilon^{n_0}$, à $n_0 \in \mathbb{N}$ fixé, et on le compare à la fonction valeur \mathcal{V}_{n_0} . De manière plus précise, on montre que le coût associé à cette stratégie est égal à la fonction V_{n_0} . On commence par définir les fonctions coûts associés aux stratégies $(\tilde{\mathcal{S}}_\varepsilon^{n_0})_{n_0 \in \mathbb{N}}$. Pour $n_0 = 0$, on pose :

$$\mathcal{J}_0(x) = \tilde{\mathbb{E}}_{(x,0,0)} \left[\int_0^\infty e^{-\alpha s} \tilde{f}(\tilde{X}_s) ds \right], \quad (5.34)$$

le coût de la stratégie $\tilde{\mathcal{S}}_\varepsilon^0$ qui correspond à une stratégie sans intervention. Pour $n_0 \geq 1$, on pose

$$\mathcal{J}_{n_0}(x) = \tilde{\mathbb{E}}_{(x,n_0,0)} \left[\int_0^\infty e^{-\alpha s} \tilde{f}(\tilde{X}_s) ds + \sum_{i=1}^{\infty} \mathbb{1}_{\{\tilde{\tau}_i < \infty\}} e^{-\alpha \tilde{\tau}_i} \tilde{c}(\tilde{X}_{\tilde{\tau}_i-}, \tilde{X}_{\tilde{\tau}_i}) \right], \quad (5.35)$$

le coût de la stratégie $\tilde{\mathcal{S}}_\varepsilon^{n_0}$, où les fonctions \tilde{f} et \tilde{c} sont définies par

$$\forall t \in \mathbb{R}^+, \quad \begin{cases} \tilde{c}(\tilde{X}_{t-}, \tilde{X}_t) = c(\pi(\tilde{X}_{t-}), \pi(\tilde{X}_t)) \\ \tilde{f}(\tilde{X}_t) = f(\pi(\tilde{X}_t)), \end{cases} \quad (5.36)$$

Pour rappel, π est la projection de l'espace \tilde{E} sur l'espace E .

On prouve maintenant par récurrence que $\mathcal{J}_n = V_n$, pour tout $n \geq 0$. La preuve se décompose en trois théorèmes. On prouve tout d'abord que $\mathcal{J}_0 = h = V_0$. Ensuite, on montre que la suite (\mathcal{J}_n) peut s'exprimer de la même manière que la suite (V_n) , au rang $n = 1$ pour l'initialisation de la récurrence, puis pour tout $n \geq 1$.

Théorème 5.3.1

Soit $x \in E$. On a l'égalité suivante :

$$\mathcal{J}_0(x) = h(x) = \mathbb{E}_x \left[\int_0^\infty e^{-\alpha s} f(X_s) ds \right]. \quad (5.37)$$

Démonstration : Soit $x = (m, \zeta) \in E$. Par la définition (5.34) de \mathcal{J}_0 et (5.36) de \tilde{f} , on a

$$\mathcal{J}_0(x) = \tilde{\mathbb{E}}_{(x,0,0)} \left[\int_0^\infty e^{-\alpha s} f(\pi(\tilde{X}_s)) ds \right].$$

On montre que $\{\pi(\tilde{X}_t)\}$ partant de $\tilde{x} = (x, 0, 0)$ est un PDMP qui a les mêmes caractéristiques que le processus $\{X_t\}$ partant de x . Le processus $\{\pi(\tilde{X}_t)\}$ part du point $\pi(\tilde{X}_0) = (m, \zeta) = x$. Il suit alors le flot (voir (5.24))

$$\pi(\tilde{\Phi}((m, \zeta, 0, \theta), t)) = (m, \Phi_m(\zeta, t)) = \Phi(x, t)$$

jusqu'au premier temps de saut, qui est déterminé par l'équation de survie :

$$\tilde{\mathbb{P}}_{\tilde{x}}(\tilde{T}_1 > t) = \exp(-\tilde{\Lambda}(\tilde{x}))\mathbb{1}_{\{t < \tilde{t}^*(\tilde{x})\}}.$$

Par la définition (5.25) de \tilde{t}^* et par la définition (5.21) de \tilde{r}_ε , on a $\tilde{t}^*(\tilde{x}) = t^*(x)$. De plus, l'expression (5.27) de $\tilde{\Lambda}$ nous donne

$$\tilde{\mathbb{P}}_{\tilde{x}}(\tilde{T}_1 > t) = \exp(-\Lambda(x))\mathbb{1}_{\{t < t^*(x)\}} = \mathbb{P}_x(T_1 > t).$$

À ce temps aléatoire, un saut survient et le processus $\{\pi(\tilde{X}_t)\}$ redémarre au point $\pi(\tilde{X}_{\tilde{T}_1}) = (m_{\tilde{T}_1}, \zeta_{\tilde{T}_1})$ selon la distribution (voir (5.28))

$$\begin{aligned} \tilde{Q}(\tilde{\Phi}(\tilde{x}, \tilde{T}_1), (\cdot, 0, 0)) &= \tilde{Q}((\Phi(x, \tilde{T}_1), 0, \theta_{\tilde{T}_1}), (\cdot, 0, 0)) \\ &= Q(\Phi(x, \tilde{T}_1), \cdot) \\ &= Q(\Phi(x, T_1), \cdot). \end{aligned}$$

On peut alors continuer la construction du processus $\{\pi(\tilde{X}_t)\}$ de la même manière. On en conclut que $\{\pi(\tilde{X}_t)\}$ est un PDMP avec les mêmes caractéristiques que le PDMP $\{X_t\}$. La conséquence en est que

$$\mathcal{J}_0(x) = \mathbb{E}_x \left[\int_0^\infty e^{-\alpha s} f(X_s) ds \right],$$

ce qui correspond exactement à la définition (1.5) de la fonction h . ■

Théorème 5.3.2

Soit $x \in E$. On a l'égalité suivante :

$$\mathcal{J}_1(x) = \begin{cases} K\mathcal{J}_0(x) & \text{si } K\mathcal{J}_0(x) < \inf_{t \in \mathbb{R}^+} J(M\mathcal{J}_0, \mathcal{J}_0)(x, t), \\ J(M\mathcal{J}_0, \mathcal{J}_0)(x, r_\varepsilon^1(x)) & \text{sinon,} \end{cases}$$

impliquant $\mathcal{J}_1 = V_1$ car $\mathcal{J}_0 = h$.

Démonstration : Soit $x \in E$ et $\tilde{x} := (x, 1, 0)$. On rappelle que $h = Kh$ par la proposition 1.6.2. On distingue deux cas.

Cas 1 : $K\mathcal{J}_0(x) < \inf_{t \in \mathbb{R}^+} J(M\mathcal{J}_0, \mathcal{J}_0)(x, t)$

Par le théorème 5.3.1, on a $h(x) = Kh(x) < \inf_{t \in \mathbb{R}^+} J(Mh, h)(x, t)$. Donc la formule (5.1) donne $r_\varepsilon^1(x) = t^*(x)$, ce qui entraîne que le saut suivant est un saut naturel. Nous allons montrer que $\{\pi(\tilde{X}_t)\}$ partant de $\tilde{x} = (x, 1, 0)$ a la même distribution que $\{X_t\}$ partant de x .

La formule (5.35) du coût associé à la stratégie S_ε^1 et la définition (5.28) de \tilde{Q} nous donnent

$$\mathcal{J}_1(x) = \tilde{\mathbb{E}}_{(x,1,0)} \left[\int_0^{\tilde{S}_1} e^{-\alpha s} \tilde{f}(\tilde{\Phi}((x, 1, 0), s)) ds + \int_{\tilde{S}_1}^\infty e^{-\alpha s} \tilde{f}(\tilde{X}_s) ds \right].$$

En utilisant maintenant les définitions (5.36) de \tilde{f} et (5.24) de $\tilde{\Phi}$, et en effectuant un changement de variable dans la seconde intégrale, on a

$$\mathcal{J}_1(x) = \tilde{\mathbb{E}}_{(x,1,0)} \left[\int_0^{\tilde{S}_1} e^{-\alpha s} f(\Phi(x, s)) ds + e^{-\alpha \tilde{S}_1} \tilde{\mathbb{E}}_{(x,1,0)} \left[\int_0^\infty e^{-\alpha t} \tilde{f}(\tilde{X}_{t+\tilde{S}_1}) dt \middle| \tilde{\mathcal{F}}_{\tilde{S}_1} \right] \right].$$

La propriété de Markov pour le processus $\{\tilde{X}_t\}$ implique

$$\mathcal{J}_1(x) = \tilde{\mathbb{E}}_{(x,1,0)} \left[\int_0^{\tilde{S}_1} e^{-\alpha s} f(\Phi(x, s)) ds \right] + \tilde{\mathbb{E}}_{(x,1,0)} \left[e^{-\alpha \tilde{S}_1} \tilde{E}_{\tilde{Z}_1} \left[\int_0^\infty e^{-\alpha t} \tilde{f}(\tilde{X}_t) dt \right] \right].$$

Comme $n_0 = 1$ et $t^*(x) = r_\varepsilon^1(x)$, on a soit $S_1 < t^*(x) \wedge r_\varepsilon^1(x)$, soit $S_1 = t^*(x) = r_\varepsilon^1(x)$. Alors par le lemme 5.3.2, on obtient

$$\mathcal{J}_1(x) = \mathbb{E}_x \left[\int_0^{S_1} e^{-\alpha s} f(\Phi(x, s)) ds \right] + \mathbb{E}_x \left[e^{-\alpha S_1} \tilde{\mathbb{E}}_{(Z_1,0,0)} \left[\int_0^\infty e^{-\alpha t} \tilde{f}(\tilde{X}_t) dt \right] \right].$$

La définition (5.34) de \mathcal{J}_0 et la définition (1.6) de F entraînent

$$\mathcal{J}_1(x) = F(x, t^*(x)) + \mathbb{E}_x \left[e^{-\alpha S_1} \mathcal{J}_0(Z_1) \right].$$

Finalement, la définition (1.8) de l'opérateur K et le théorème 5.3.1 nous donnent $\mathcal{J}_1(x) = K\mathcal{J}_0(x) = Kh(x)$.

Cas 2 : $K\mathcal{J}_0(x) \geq \inf_{t \in \mathbb{R}^+} J(M\mathcal{J}_0, \mathcal{J}_0)(x, t)$

Par le théorème 5.3.1, on a

$$Kh(x) \geq \inf_{t \in \mathbb{R}^+} J(Mh, h)(x, t).$$

Donc la proposition 5.3.1 entraîne que $r_\varepsilon^1(x) < t^*(x)$. En particulier, on a $r_\varepsilon^1(x) \neq t^*(x)$. L'expression (5.28) de \tilde{Q} montre qu'une intervention est possible.

Premièrement, on peut noter que $\tilde{S}_1 = r_\varepsilon^1(x)$ est équivalent à $\tilde{S}_1 \geq r_\varepsilon^1(x)$ car on a toujours $\tilde{S}_1 \leq \tilde{t}^*(\tilde{x}) = r_\varepsilon^1(x)$ par la définition (5.25) de $\tilde{t}^*(\tilde{x})$. Alors, en partant de la définition (5.35) de \mathcal{J}_{N_0} pour $n_0 = 1$ et en utilisant les définitions (5.36) de \tilde{c} et (5.24) de $\tilde{\Phi}$, on a

$$\mathcal{J}_1(x) = \phi_1(x) + \phi_2(x) + \phi_3(x). \quad (5.38)$$

où

$$\begin{aligned} \phi_1(x) &:= \tilde{\mathbb{E}}_{(x,1,0)} \left[\int_0^{\tilde{S}_1 \wedge r_\varepsilon^1(x)} e^{-\alpha s} \tilde{f}(\tilde{X}_s) ds \right] \\ \phi_2(x) &:= \tilde{\mathbb{E}}_{(x,1,0)} \left[\mathbf{1}_{\{\tilde{S}_1 < r_\varepsilon^1(x)\}} \int_{\tilde{S}_1}^{+\infty} e^{-\alpha s} \tilde{f}(\tilde{X}_s) ds \right] \\ \phi_3(x) &:= \tilde{\mathbb{E}}_{(x,1,0)} \left[\mathbf{1}_{\{\tilde{S}_1 \geq r_\varepsilon^1(x)\}} \left\{ \int_{r_\varepsilon^1(x)}^{+\infty} e^{-\alpha s} \tilde{f}(\tilde{X}_s) ds + e^{-\alpha r_\varepsilon^1(x)} c(\Phi(x, r_\varepsilon^1(x)), y_\varepsilon^1(x)) \right\} \right]. \end{aligned}$$

Pour la première partie de (5.38), on a

$$\phi_1(x) = \tilde{\mathbb{E}}_{(x,1,0)} \left[\int_0^{\tilde{S}_1 \wedge r_\varepsilon^1(x)} e^{-\alpha s} f(\Phi(x, s)) ds \right].$$

Comme $n_0 = 1$ et $r_\varepsilon^1(x) < t^*(x)$, le lemme 5.3.2 dit que pour toute fonction \tilde{v} satisfaisant les hypothèses du lemme, on a :

$$\tilde{\mathbb{E}}_{\tilde{x}}[\tilde{v}(\tilde{Z}_1, \tilde{S}_1)] = \mathbb{E}_x \left[\mathbf{1}_{\{S_1 < r_\varepsilon^1(x)\}} \times v(Z_1, 0, S_1) + \mathbf{1}_{\{S_1 \geq r_\varepsilon^1(x)\}} \times v(y_\varepsilon^1(x), 0, r_\varepsilon^1(x)) \right]. \quad (5.39)$$

Appliqué à Φ_1 , on obtient :

$$\begin{aligned} \phi_1(x) &= \mathbb{E}_x \left[\mathbf{1}_{\{S_1 < r_\varepsilon^1(x)\}} \int_0^{S_1 \wedge r_\varepsilon^1(x)} e^{-\alpha s} f(\Phi(x, s)) ds + \mathbf{1}_{\{S_1 \geq r_\varepsilon^1(x)\}} \int_0^{r_\varepsilon^1(x)} e^{-\alpha s} f(\Phi(x, s)) ds \right] \\ &= \mathbb{E}_x \left[\int_0^{S_1 \wedge r_\varepsilon^1(x)} e^{-\alpha s} f(\Phi(x, s)) ds \right]. \end{aligned}$$

Finalement, la définition (1.6) de F entraîne

$$\phi_1(x) = F(x, r_\varepsilon^1(x)). \quad (5.40)$$

Pour la seconde partie de (5.38), on a

$$\phi_2(x) = \tilde{\mathbb{E}}_{(x,1,0)} \left[\mathbf{1}_{\{\tilde{S}_1 < r_\varepsilon^1(x)\}} e^{-\alpha \tilde{S}_1} \tilde{\mathbb{E}}_{(x,1,0)} \left[\int_0^{+\infty} e^{-\alpha t} \tilde{f}(\tilde{X}_{t+\tilde{S}_1}) dt \middle| \tilde{\mathcal{F}}_{\tilde{S}_1} \right] \right].$$

La propriété de Markov implique

$$\phi_2(x) = \tilde{\mathbb{E}}_{(x,1,0)} \left[\mathbf{1}_{\{\tilde{S}_1 < r_\varepsilon^1(x)\}} e^{-\alpha \tilde{S}_1} \tilde{\mathbb{E}}_{\tilde{Z}_1} \left[\int_0^{+\infty} e^{-\alpha t} \tilde{f}(\tilde{X}_t) dt \right] \right].$$

La formule (5.39) est encore valable. Appliquée à Φ_2 , elle donne :

$$\begin{aligned} \phi_2(x) = & \mathbb{E}_x \left[\mathbf{1}_{\{\tilde{S}_1 < r_\varepsilon^1(x)\}} \left(\mathbf{1}_{\{S_1 < r_\varepsilon^1(x)\}} e^{-\alpha \tilde{S}_1} \mathbb{E}_{Z_1} \left[\int_0^{+\infty} e^{-\alpha t} f(X_t) dt \right] \right) \right. \\ & \left. + \mathbf{1}_{\{S_1 \geq r_\varepsilon^1(x)\}} \left(\mathbf{1}_{\{r_\varepsilon^1(x) < r_\varepsilon^1(x)\}} e^{-\alpha r_\varepsilon^1(x)} \mathbb{E}_{y_\varepsilon^1(x)} \left[\int_0^{+\infty} e^{-\alpha t} f(X_t) dt \right] \right) \right]. \end{aligned}$$

On en conclut

$$\phi_2(x) = \mathbb{E}_x \left[\mathbf{1}_{\{S_1 < r_\varepsilon^1(x)\}} e^{-\alpha S_1} \tilde{\mathbb{E}}_{(Z_1, 0, 0)} \left[\int_0^{+\infty} e^{-\alpha t} \tilde{f}(\tilde{X}_t) dt \right] \right].$$

La définition (5.34) de \mathcal{J}_0 puis l'application du théorème 5.3.1 entraînent

$$\phi_2(x) = \mathbb{E}_x \left[\mathbf{1}_{\{S_1 < r_\varepsilon^1(x)\}} e^{-\alpha S_1} h(Z_1) \right]. \quad (5.41)$$

Pour la troisième partie de (5.38), on a

$$\begin{aligned} \phi_3(x) = & \tilde{\mathbb{E}}_{(x, 1, 0)} \left[\mathbf{1}_{\{\tilde{S}_1 \geq r_\varepsilon^1(x)\}} e^{-\alpha r_\varepsilon^1(x)} \right. \\ & \left. \times \left\{ \tilde{\mathbb{E}}_{(x, 1, 0)} \left[\int_0^{+\infty} e^{-\alpha t} \tilde{f}(\tilde{X}_{t+r_\varepsilon^1(x)}) dt \middle| \tilde{\mathcal{F}}_{r_\varepsilon^1(x)} \right] + c(\Phi(x, r_\varepsilon^1(x)), y_\varepsilon^1(x)) \right\} \right]. \end{aligned}$$

Par la propriété de Markov, on obtient

$$\phi_3(x) = \tilde{\mathbb{E}}_{(x, 1, 0)} \left[\mathbf{1}_{\{\tilde{S}_1 \geq r_\varepsilon^1(x)\}} e^{-\alpha r_\varepsilon^1(x)} \left\{ \tilde{\mathbb{E}}_{(y_\varepsilon^1(x), 0, 0)} \left[\int_0^{+\infty} e^{-\alpha t} \tilde{f}(\tilde{X}_t) dt \right] + c(\Phi(x, r_\varepsilon^1(x)), y_\varepsilon^1(x)) \right\} \right].$$

Premièrement, en utilisant la définition (5.34) de \mathcal{J}_0 et le théorème 5.3.1, puis en utilisant la définition (1.9) de l'opérateur M et la définition (5.23) de y_ε^1 , on a

$$\begin{aligned} \phi_3(x) = & \tilde{\mathbb{E}}_{(x, 1, 0)} \left[\mathbf{1}_{\{\tilde{S}_1 \geq r_\varepsilon^1(x)\}} e^{-\alpha r_\varepsilon^1(x)} \left\{ h(y_\varepsilon^1(x)) + c(\Phi(x, r_\varepsilon^1(x)), y_\varepsilon^1(x)) \right\} \right] \\ = & \tilde{\mathbb{E}}_{(x, 1, 0)} \left[\mathbf{1}_{\{\tilde{S}_1 \geq r_\varepsilon^1(x)\}} e^{-\alpha r_\varepsilon^1(x)} Mh(\Phi(x, r_\varepsilon^1(x))) \right]. \end{aligned}$$

On applique la formule (5.39), ce qui donne :

$$\begin{aligned} \phi_3(x) = & \mathbb{E}_x \left[\mathbf{1}_{\{\tilde{S}_1 < r_\varepsilon^1(x)\}} \mathbf{1}_{\{S_1 \geq r_\varepsilon^1(x)\}} e^{-\alpha r_\varepsilon^1(x)} Mh(\Phi(x, r_\varepsilon^1(x))) \mathbf{1}_{\{S_1 \geq r_\varepsilon^1(x)\}} \right. \\ & \left. + \mathbf{1}_{\{\tilde{S}_1 \geq r_\varepsilon^1(x)\}} \mathbf{1}_{\{r_\varepsilon^1(x) \geq r_\varepsilon^1(x)\}} e^{-\alpha r_\varepsilon^1(x)} Mh(\Phi(x, r_\varepsilon^1(x))) \mathbf{1}_{\{S_1 \geq r_\varepsilon^1(x)\}} \right], \end{aligned}$$

d'où

$$\phi_3(x) = \mathbb{E}_x \left[e^{-\alpha r_\varepsilon^1(x)} Mh(\Phi(x, r_\varepsilon^1(x))) \mathbf{1}_{\{S_1 \geq r_\varepsilon^1(x)\}} \right]. \quad (5.42)$$

Finalement, les équations (5.38), (5.40), (5.41) et (5.42) impliquent

$$\mathcal{J}_1(x) = F(x, r_\varepsilon^1(x)) + \mathbb{E}_x \left[e^{-\alpha r_\varepsilon^1(x)} Mh(\Phi(x, r_\varepsilon^1(x))) \mathbf{1}_{\{S_1 \geq r_\varepsilon^1(x)\}} + e^{-\alpha S_1} h(Z_1) \mathbf{1}_{\{S_1 < r_\varepsilon^1(x)\}} \right]$$

et la définition (1.7) de l'opérateur J nous donne le résultat : $\mathcal{J}_1(x) = J(Mh, h)(x, r_\varepsilon^1(x))$. \blacksquare

Théorème 5.3.3

Soit n_0 un entier naturel tel que $n_0 \geq 1$. Soit $x \in E$. On a l'expression suivante

$$\mathcal{J}_{n_0}(x) = \begin{cases} K \mathcal{J}_{n_0-1}(x) & \text{si } K \mathcal{J}_{n_0-1}(x) < \inf_{t \in \mathbb{R}^+} J(M \mathcal{J}_{n_0-1}, \mathcal{J}_{n_0-1})(x, t), \\ J(M \mathcal{J}_{n_0-1}, \mathcal{J}_{n_0-1})(x, r_\varepsilon^{n_0}(x)) & \text{sinon.} \end{cases}$$

De plus, on a $\mathcal{J}_{n_0} = V_{n_0}$.

Démonstration : On procède par récurrence. Le cas $n_0 = 1$ a été montré précédemment dans le théorème 5.3.2. Soit $n_0 \geq 1$ et $x \in E$. Supposons que

$$\mathcal{J}_{n_0}(x) = \begin{cases} K\mathcal{J}_{n_0-1}(x) & \text{si } K\mathcal{J}_{n_0-1}(x) < \inf_{t \in \mathbb{R}^+} J(M\mathcal{J}_{n_0-1}, \mathcal{J}_{n_0-1})(x, t), \\ J(M\mathcal{J}_{n_0-1}, \mathcal{J}_{n_0-1})(x, r_\varepsilon^{n_0}(x)) & \text{sinon.} \end{cases}$$

et $\mathcal{J}_{n_0} = V_{n_0}$. On distingue deux cas.

Cas 1 : $K\mathcal{J}_{n_0}(x) < \inf_{t \in \mathbb{R}^+} J(M\mathcal{J}_{n_0}, \mathcal{J}_{n_0})(x, t)$

Par hypothèse de récurrence, on obtient $KV_{n_0}(x) < \inf_{t \in \mathbb{R}^+} J(MV_{n_0}, V_{n_0})(x, t)$. Par la définition (5.3) de $r_\varepsilon^{n_0+1}$, cette inégalité implique $r_\varepsilon^{n_0+1}(x) = t^*(x)$. La définition (5.28) de \tilde{Q} montre alors que le premier saut n'est pas une intervention. Par conséquent, on a

$$\mathcal{J}_{n_0+1}(x) = \psi_1(x) + \psi_2(x), \quad (5.43)$$

où

$$\begin{aligned} \psi_1(x) &:= \tilde{\mathbb{E}}_{(x, n_0+1, 0)} \left[\int_0^{\tilde{S}_1} e^{-\alpha s} \tilde{f}(\tilde{\Phi}((x, n_0+1, 0), s)) ds \right] \\ \psi_2(x) &:= \tilde{\mathbb{E}}_{(x, n_0+1, 0)} \left[\int_{\tilde{S}_1}^\infty e^{-\alpha s} \tilde{f}(\tilde{X}_s) ds + \sum_{i=1}^\infty e^{-\alpha \tilde{\tau}_i} \tilde{c}(\tilde{X}_{\tilde{\tau}_i-}, \tilde{X}_{\tilde{\tau}_i}) \right]. \end{aligned}$$

Pour la première partie de \mathcal{J}_{n_0+1} , on applique le lemme 5.3.2 avec $n_0 \geq 1$ et $r_\varepsilon^{n_0+1}(x) = t^*(x)$: pour toutes fonctions \tilde{v} et v répondant aux hypothèses du lemme 5.3.2, on a

$$\tilde{\mathbb{E}}_{(x, n_0+1, 0)} = \mathbb{E}_x[v(Z_1, n_0, S_1)]. \quad (5.44)$$

D'où

$$\psi_1(x) = \mathbb{E}_x \left[\int_0^{S_1} e^{-\alpha s} f(\Phi(x, s)) ds \right],$$

et par la définition (1.6) de l'opérateur F , on obtient finalement l'égalité suivante :

$$\psi_1(x) = F(x, t^*(x)). \quad (5.45)$$

Pour la seconde partie \mathcal{J}_{n_0+1} , on a

$$\psi_2(x) = \tilde{\mathbb{E}}_{(x, n_0+1, 0)} \left[e^{-\alpha \tilde{S}_1} \tilde{\mathbb{E}}_{(x, n_0+1, 0)} \left[\int_0^\infty e^{-\alpha t} \tilde{f}(\tilde{X}_{t+\tilde{S}_1}) dt + \sum_{i=1}^\infty e^{-\alpha(\tilde{\tau}_i - \tilde{S}_1)} \tilde{c}(\tilde{X}_{\tilde{\tau}_i-}, \tilde{X}_{\tilde{\tau}_i}) \Big| \tilde{\mathcal{F}}_{\tilde{S}_1} \right] \right].$$

Étant donné que dans notre cas, $n_0 \geq 1$ et $r_\varepsilon^{n_0+1}(x) = t^*(x)$, la proposition 5.3.2 nous donne

$$\begin{aligned} \forall g \in \mathbf{B}(\mathbb{R}^+), \quad \tilde{\mathbb{E}}_{\tilde{x}}[f(\tilde{\tau}_i) | \tilde{\mathcal{F}}_{\tilde{S}_1}] &= (\mathbb{1}_{\{\tilde{S}_1 < r_\varepsilon^{n_0+1}(x) = t^*(x)\}} + \mathbb{1}_{\{\tilde{S}_1 = r_\varepsilon^{n_0+1}(x) = t^*(x)\}}) \tilde{\mathbb{E}}_{\tilde{Z}_1}[f(\tilde{\tau}_i + \tilde{S}_1)] \\ &= \tilde{\mathbb{E}}_{\tilde{Z}_1}[f(\tilde{\tau}_i + \tilde{S}_1)]. \end{aligned}$$

La propriété de Markov appliquée à ψ_2 permet alors d'écrire

$$\begin{aligned} \psi_2(x) &= \tilde{\mathbb{E}}_{(x, n_0+1, 0)} \left[e^{-\alpha \tilde{S}_1} \tilde{\mathbb{E}}_{\tilde{Z}_1} \left[\int_0^\infty e^{-\alpha t} \tilde{f}(\tilde{X}_t) dt + \sum_{i=1}^\infty e^{-\alpha \tilde{\tau}_i} \tilde{c}(\tilde{X}_{(\tilde{\tau}_i - \tilde{S}_1) - \tilde{S}_1}, \tilde{X}_{(\tilde{\tau}_i + \tilde{S}_1) - \tilde{S}_1}) \right] \right] \\ &= \tilde{\mathbb{E}}_{(x, n_0+1, 0)} \left[e^{-\alpha \tilde{S}_1} \tilde{\mathbb{E}}_{\tilde{Z}_1} \left[\int_0^\infty e^{-\alpha t} \tilde{f}(\tilde{X}_t) dt + \sum_{i=1}^\infty e^{-\alpha \tilde{\tau}_i} \tilde{c}(\tilde{X}_{\tilde{\tau}_i-}, \tilde{X}_{\tilde{\tau}_i}) \right] \right]. \end{aligned}$$

La formule (5.44) étant toujours valable, elle implique

$$\psi_2(x) = \mathbb{E}_x \left[e^{-\alpha S_1} \tilde{\mathbb{E}}_{(Z_1, n_0, 0)} \left[\int_0^\infty e^{-\alpha t} \tilde{f}(\tilde{X}_t) dt + \sum_{i=1}^\infty e^{-\alpha \tilde{\tau}_i} \tilde{c}(\tilde{X}_{\tilde{\tau}_i-}, \tilde{X}_{\tilde{\tau}_i}) \right] \right].$$

Par la définition (5.35) de \mathcal{J}_{n_0} , on a finalement

$$\psi_2(x) = \mathbb{E}_x \left[e^{-\alpha S_1} \mathcal{J}_{n_0}(Z_1) \right]. \quad (5.46)$$

Donc les équations (5.43), (5.45) et (5.46), ainsi que la définition (1.8) de l'opérateur K nous donnent

$$\mathcal{J}_{N_0+1}(x) = K \mathcal{J}_{N_0}(x).$$

Cas 2 : $K \mathcal{J}_{n_0}(x) \geq \inf_{t \in \mathbb{R}^+} J(M \mathcal{J}_{n_0}, \mathcal{J}_{n_0})(x, t)$

Par hypothèse de récurrence, on a $KV_{n_0}(x) \geq \inf_{t \in \mathbb{R}^+} J(MV_{n_0}, V_{n_0})(x, t)$. Par conséquent, la proposition 5.3.1 entraîne $r_\varepsilon^{n_0+1}(x) < t^*(x)$. En particulier, on a $r_\varepsilon^{n_0+1}(x) \neq t^*(x)$.

Pour calculer le coût associé à la stratégie $\mathcal{S}_\varepsilon^{n_0+1}$, on part de la formule (5.35). La définition (5.28) de \tilde{Q} montre que le premier saut du processus contrôlé $\{\tilde{X}_t\}$ peut être une intervention. Cela nous donne

$$\mathcal{J}_{N_0+1}(x) = \Psi_1(x) + \Psi_2(x) + \Psi_3(x) + \Psi_4(x) \quad (5.47)$$

où

$$\begin{aligned} \Psi_1(x) &:= \tilde{\mathbb{E}}_{(x, n_0+1, 0)} \left[\int_0^{\tilde{S}_1 \wedge r_\varepsilon^{n_0+1}(x)} e^{-\alpha s} \tilde{f}(\tilde{X}_s) ds \right] \\ \Psi_2(x) &:= \tilde{\mathbb{E}}_{(x, n_0+1, 0)} \left[\mathbb{1}_{\{\tilde{S}_1 < r_\varepsilon^{n_0+1}(x)\}} \left\{ \int_{\tilde{S}_1}^\infty e^{-\alpha s} \tilde{f}(\tilde{X}_s) ds + \sum_{i=1}^{+\infty} e^{-\alpha \tilde{\tau}_i} \tilde{c}(\tilde{X}_{\tilde{\tau}_i-}, \tilde{X}_{\tilde{\tau}_i}) \right\} \right] \\ \Psi_3(x) &:= \tilde{\mathbb{E}}_{(x, n_0+1, 0)} \left[\mathbb{1}_{\{\tilde{S}_1 = r_\varepsilon^{n_0+1}(x)\}} e^{-\alpha r_\varepsilon^{n_0+1}(x)} \tilde{c}(\tilde{\Phi}(\tilde{x}, r_\varepsilon^{n_0+1}(x)), (y_\varepsilon^{n_0+1}(x), n_0, 0)) \right] \\ \Psi_4(x) &:= \tilde{\mathbb{E}}_{(x, n_0+1, 0)} \left[\mathbb{1}_{\{\tilde{S}_1 = r_\varepsilon^{n_0+1}(x)\}} \left\{ \int_{r_\varepsilon^{n_0+1}(x)}^\infty e^{-\alpha s} \tilde{f}(\tilde{X}_s) ds + \sum_{i=2}^{+\infty} e^{-\alpha \tilde{\tau}_i} \tilde{c}(\tilde{X}_{\tilde{\tau}_i-}, \tilde{X}_{\tilde{\tau}_i}) \right\} \right]. \end{aligned}$$

Pour la première partie de \mathcal{J}_{n_0+1} , en utilisant le lemme 5.3.2 avec $n_0 \geq 1$ et $r_\varepsilon^{n_0+1}(x) < t^*(x)$, on obtient pour toutes fonctions \tilde{v} et v satisfaisant les hypothèses du lemme,

$$\tilde{\mathbb{E}}_{\tilde{x}}[\tilde{v}(\tilde{Z}_1, \tilde{S}_1)] = \mathbb{E}_x \left[\mathbb{1}_{\{S_1 < r_\varepsilon^{n_0}(x)\}} v(Z_1, n_0 - 1, S_1) + \mathbb{1}_{\{S_1 = r_\varepsilon^{n_0}(x)\}} v(y_\varepsilon^{n_0}(x), n_0 - 1, r_\varepsilon^{n_0}(x)) \right], \quad (5.48)$$

ce qui nous donne

$$\begin{aligned} \Psi_1(x) &= \mathbb{E}_x \left[\mathbb{1}_{\{S_1 < r_\varepsilon^{n_0}(x)\}} \int_0^{S_1 \wedge r_\varepsilon^{n_0+1}(x)} e^{-\alpha s} f(X_s) ds + \mathbb{1}_{\{S_1 = r_\varepsilon^{n_0}(x)\}} \int_0^{r_\varepsilon^{n_0+1}(x)} e^{-\alpha s} f(X_s) ds \right] \\ &= \mathbb{E}_x \left[\int_0^{S_1 \wedge r_\varepsilon^{n_0+1}(x)} e^{-\alpha s} f(X_s) ds \right]. \end{aligned}$$

D'où par la définition (1.6) de l'opérateur F ,

$$\Psi_1(x) = F(x, r_\varepsilon^{n_0+1}(x)). \quad (5.49)$$

Pour la seconde partie de \mathcal{J}_{n_0+1} , on a

$$\begin{aligned} \Psi_2(x) &= \tilde{\mathbb{E}}_{(x, n_0+1, 0)} \left[\mathbb{1}_{\{\tilde{S}_1 < r_\varepsilon^{n_0+1}(x)\}} e^{-\alpha \tilde{S}_1} \right. \\ &\quad \left. \times \tilde{\mathbb{E}}_{(x, n_0+1, 0)} \left[\int_0^\infty e^{-\alpha t} \tilde{f}(\tilde{X}_{t+\tilde{S}_1}) dt + \sum_{i=1}^{+\infty} e^{-\alpha(\tilde{\tau}_i - \tilde{S}_1)} \tilde{c}(\tilde{X}_{\tilde{\tau}_i-}, \tilde{X}_{\tilde{\tau}_i}) \middle| \tilde{\mathcal{F}}_{\tilde{S}_1} \right] \right]. \end{aligned}$$

La proposition 5.3.2 pour $n_0 \geq 1$ et $r_\varepsilon^{n_0+1}(x) < t^*(x)$ donne

$$\forall g \in \mathbf{B}(\mathbb{R}^+), \quad \tilde{\mathbb{E}}_{\tilde{x}}[g(\tilde{\tau}_i) | \tilde{\mathcal{F}}_{\tilde{S}_1}] = \mathbb{1}_{\{\tilde{S}_1 < r_\varepsilon^n(x)\}} \tilde{\mathbb{E}}_{\tilde{Z}_1}[g(\tilde{\tau}_i + \tilde{S}_1)] + \mathbb{1}_{\{\tilde{S}_1 = r_\varepsilon^n(x)\}} \tilde{\mathbb{E}}_{\tilde{Z}_1}[g(\tilde{\tau}_{i-1} + r_\varepsilon^n(x))]. \quad (5.50)$$

On applique cette formule et la propriété de Markov à Ψ_2 . On obtient ainsi

$$\Psi_2(x) = \tilde{\mathbb{E}}_{(x, n_0+1, 0)} \left[\mathbf{1}_{\{\bar{S}_1 < r_\varepsilon^{n_0+1}(x)\}} e^{-\alpha \bar{S}_1} \tilde{\mathbb{E}}_{\bar{Z}_1} \left[\int_0^\infty e^{-\alpha t} \tilde{f}(\tilde{X}_t) dt + \sum_{i=1}^{+\infty} e^{-\alpha \bar{\tau}_i} \tilde{c}(\tilde{X}_{\bar{\tau}_i-}, \tilde{X}_{\bar{\tau}_i}) \right] \right].$$

La formule (5.48), issue du lemme 5.3.2, implique

$$\Psi_2(x) = \mathbb{E}_x \left[\mathbf{1}_{\{S_1 < r_\varepsilon^{n_0+1}(x)\}} e^{-\alpha S_1} \tilde{\mathbb{E}}_{(Z_1, n_0, 0)} \left[\int_0^\infty e^{-\alpha t} \tilde{f}(\tilde{X}_t) dt + \sum_{i=1}^{+\infty} e^{-\alpha \bar{\tau}_i} \tilde{c}(\tilde{X}_{\bar{\tau}_i-}, \tilde{X}_{\bar{\tau}_i}) \right] \right].$$

Finalement, la définition (5.35) de \mathcal{J}_{n_0} nous donne

$$\Psi_2(x) = \mathbb{E}_x \left[\mathbf{1}_{\{S_1 < r_\varepsilon^{n_0+1}(x)\}} e^{-\alpha S_1} \mathcal{J}_{n_0}(Z_1) \right]. \quad (5.51)$$

Pour la troisième partie de \mathcal{J}_{n_0+1} , on a

$$\Psi_3(x) := \tilde{\mathbb{E}}_{(x, n_0+1, 0)} \left[\mathbf{1}_{\{\bar{S}_1 = r_\varepsilon^{n_0+1}(x)\}} e^{-\alpha r_\varepsilon^{n_0+1}(x)} \tilde{c}(\tilde{\Phi}(\tilde{x}, r_\varepsilon^{n_0+1}(x)), (y_\varepsilon^{n_0+1}(x), n_0, 0)) \right].$$

À l'aide de la formule (5.48), issue du lemme 5.3.2, et en remarquant que l'évènement $\{S_1 = r_\varepsilon^{n_0+1}(x)\}$ est le même évènement que $\{S_1 \geq r_\varepsilon^{n_0+1}(x)\}$ car on a toujours $\{S_1 \leq t^*(x) \wedge r_\varepsilon^{n_0+1}(x)\}$, on peut conclure

$$\Psi_3(x) = \mathbb{E}_x \left[\mathbf{1}_{\{S_1 \geq r_\varepsilon^{n_0+1}(x)\}} e^{-\alpha r_\varepsilon^{n_0+1}(x)} c(\Phi(x, r_\varepsilon^{n_0+1}(x)), y_\varepsilon^{n_0+1}(x)) \right]. \quad (5.52)$$

Pour la quatrième partie de \mathcal{J}_{n_0+1} , on a

$$\begin{aligned} \Psi_4(x) = & \tilde{\mathbb{E}}_{(x, n_0+1, 0)} \left[\mathbf{1}_{\{\bar{S}_1 = r_\varepsilon^{n_0+1}(x)\}} e^{-\alpha r_\varepsilon^{n_0+1}(x)} \tilde{\mathbb{E}}_{(x, n_0+1, 0)} \left[\int_0^\infty e^{-\alpha t} \tilde{f}(\tilde{X}_{t+r_\varepsilon^{n_0+1}(x)}) dt \right. \right. \\ & \left. \left. + \sum_{i=2}^{+\infty} e^{-\alpha(\bar{\tau}_i - r_\varepsilon^{n_0+1}(x))} \tilde{c}(\tilde{X}_{\bar{\tau}_i-}, \tilde{X}_{\bar{\tau}_i}) \Big| \tilde{\mathcal{F}}_{r_\varepsilon^{n_0+1}(x)} \right] \right]. \end{aligned}$$

La formule (5.50), issue de la proposition 5.3.2, et la propriété de Markov impliquent

$$\begin{aligned} \Psi_4(x) = & \tilde{\mathbb{E}}_{(x, n_0+1, 0)} \left[\mathbf{1}_{\{\bar{S}_1 = r_\varepsilon^{n_0+1}(x)\}} e^{-\alpha r_\varepsilon^{n_0+1}(x)} \tilde{\mathbb{E}}_{\bar{Z}} \left[\int_0^\infty e^{-\alpha t} \tilde{f}(\tilde{X}_t) dt \right. \right. \\ & \left. \left. + \sum_{i=2}^{+\infty} e^{-\alpha \bar{\tau}_{i-1}} \tilde{c}(\tilde{X}_{\bar{\tau}_{i-1}-}, \tilde{X}_{\bar{\tau}_{i-1}}) \right] \right] \\ = & \tilde{\mathbb{E}}_{(x, n_0+1, 0)} \left[\mathbf{1}_{\{\bar{S}_1 = r_\varepsilon^{n_0+1}(x)\}} e^{-\alpha r_\varepsilon^{n_0+1}(x)} \tilde{\mathbb{E}}_{\bar{Z}} \left[\int_0^\infty e^{-\alpha t} \tilde{f}(\tilde{X}_t) dt + \sum_{i=1}^{+\infty} e^{-\alpha \bar{\tau}_i} \tilde{c}(\tilde{X}_{\bar{\tau}_i-}, \tilde{X}_{\bar{\tau}_i}) \right] \right]. \end{aligned}$$

On applique la formule (5.48), issue du lemme 5.3.2, ce qui nous donne

$$\Psi_4(x) = \mathbb{E}_x \left[\mathbf{1}_{\{S_1 = r_\varepsilon^{n_0+1}(x)\}} e^{-\alpha r_\varepsilon^{n_0+1}(x)} \tilde{\mathbb{E}}_{(Z_1, n_0, 0)} \left[\int_0^\infty e^{-\alpha t} \tilde{f}(\tilde{X}_t) dt + \sum_{i=1}^{+\infty} e^{-\alpha \bar{\tau}_i} \tilde{c}(\tilde{X}_{\bar{\tau}_i-}, \tilde{X}_{\bar{\tau}_i}) \right] \right].$$

En utilisant la définition (5.35) de \mathcal{J}_{n_0} et par le même argument que dans la troisième partie de \mathcal{J}_{n_0+1} , on obtient finalement

$$\Psi_4(x) = \mathbb{E}_x \left[\mathbf{1}_{\{S_1 \geq r_\varepsilon^{n_0+1}(x)\}} e^{-\alpha r_\varepsilon^{n_0+1}(x)} \mathcal{J}_{n_0}(y_\varepsilon^{n_0+1}(x)) \right]. \quad (5.53)$$

En combinant maintenant les équations (5.47), (5.49), (5.51), (5.52) et (5.53), l'expression du coût \mathcal{J}_{n_0+1} devient

$$\begin{aligned} \mathcal{J}_{n_0+1}(x) &= F(x, r_\varepsilon^{n_0+1}(x)) + \mathbb{E}_x \left[\mathbb{1}_{\{S_1 < r_\varepsilon^{n_0+1}(x)\}} e^{-\alpha S_1} \mathcal{J}_{n_0}(Z_1) \right. \\ &\quad \left. + \mathbb{1}_{\{S_1 \geq r_\varepsilon^{n_0+1}(x)\}} e^{-\alpha r_\varepsilon^{n_0+1}(x)} \left\{ c(\Phi(x, r_\varepsilon^{n_0+1}(x)), y_\varepsilon^{n_0+1}(x)) + \mathcal{J}_{n_0}(y_\varepsilon^{n_0+1}(x)) \right\} \right]. \end{aligned}$$

La définition (1.9) de l'opérateur M et la définition (1.7) de l'opérateur J entraîne

$$\mathcal{J}_{n_0+1}(x) = J(M\mathcal{J}_{n_0}, \mathcal{J}_{n_0})(x, r_\varepsilon^{n_0+1}(x)).$$

Conclusion :

Par hypothèse de récurrence, on a : $\mathcal{J}_{n_0} = V_{n_0}$. Par conséquent, le coût \mathcal{J}_{n_0+1} peut s'écrire

$$\mathcal{J}_{n_0+1}(x) = \begin{cases} KV_{n_0}(x) & \text{si } KV_{n_0}(x) < \inf_{t \in \mathbb{R}^+} J(MV_{n_0}, V_{n_0})(x, t), \\ J(MV_{n_0}, V_{n_0})(x, r_\varepsilon^{n_0+1}(x)) & \text{sinon.} \end{cases}$$

ce qui correspond exactement à la définition (5.4) de V_{n_0+1} . On a donc bien $\mathcal{J}_{n_0+1} = V_{n_0+1}$. ■

5.4 Stratégies ε -optimales

Le regroupement des théorèmes 5.1.1 et 5.3.3 nous donne le théorème suivant :

Théorème 5.4.1

Soit n_0 un entier naturel. Alors

$$\forall x \in E, \quad \mathcal{V}_{n_0}(x) \leq \mathcal{J}_{n_0}(x) \leq \mathcal{V}_{n_0}(x) + n_0\varepsilon.$$

À n_0 fixé, le coût associé à la stratégie $\tilde{\mathcal{S}}_\varepsilon^n$ donné par la fonction \mathcal{J}_{n_0} est proche de la fonction valeur \mathcal{V}_{n_0} du problème de contrôle impulsif où l'horizon de contrôle est n_0 , d'autant que ε est proche de 0.

Ainsi, la famille de stratégies construite au paragraphe 5.2 est ε -optimale pour le problème de contrôle impulsif avec facteur d'oubli et horizon infini. La construction de cette famille de stratégies est explicite et ne requiert pas la résolution de problème complexe au préalable mais nécessite la connaissance du coût moyen associé à la stratégie de non-intervention, ainsi que la connaissance des opérateurs J et K et leurs itérations.

Une perspective de ce travail est de construire une approximation numérique des stratégies ainsi construites afin de résoudre des problèmes tels que l'optimisation de maintenance. Pour ce faire, il peut être intéressant de se baser sur l'article de Saporta et Dufour (2012), dans lequel les auteurs proposent une approximation numérique de la fonction valeur. En effet, l'idée de cette approximation est la discrétisation des opérateurs J , K et M , à l'aide d'une grille de discrétisation adaptée pour le temps et à l'aide de la quantification pour la discrétisation en espace du processus. La construction des stratégies ε -optimales décrites dans ce chapitre utilisant également les opérateurs K , J et M , une approche possible pour construire une approximation de ces stratégies pourrait être d'utiliser la discrétisation présentée dans de Saporta et Dufour (2012).

Conclusion et perspectives

L'objectif de cette thèse était d'apporter une contribution à la théorie du contrôle impulsif pour des processus de Markov déterministes par morceaux (PDMP) et de modéliser un problème concret par des PDMP afin de mettre en application des méthodes numériques de contrôle impulsif, dans le but de répondre à une problématique industrielle.

Dans la première partie, nous avons étudié la maintenance d'une caméra thermique développée par Thales. Nous avons mis en place une modélisation des évolutions de la caméra par des PDMP. Ce travail a été conséquent et a été le fruit de très nombreuses discussions et échanges avec l'entreprise Thales, afin de permettre l'obtention d'un modèle simple mais proche de la réalité et adaptable aux conditions futures de maintenance de la caméra. Nous avons alors défini un problème de contrôle impulsif qui modélise les décisions de maintenance prises durant une période de temps défini. Nous avons cherché à optimiser un coût comprenant le coût des maintenances et le coût d'indisponibilité de la caméra. Cependant, le fonctionnement de la caméra étant assez complexe et le nombre de possibilités de décisions étant très grand, nous nous sommes restreints à une seule catégorie de maintenances, répondant à trois critères fixés.

Une des perspectives naturelles de ce travail est donc d'élargir les maintenances possibles à d'autres catégories, notamment en modifiant le premier critère ((\mathcal{C}_1) ou (\mathcal{D}_1)) pour permettre des maintenances réalisées lorsque la caméra est dans l'état stable. Cette piste est très prometteuse : l'élargissement du critère (\mathcal{C}_1) au critère (\mathcal{D}_1) a permis de réduire les coûts de plus de 25% si l'on considère les meilleures politiques explicitées (politique n°3 et n°5) et de réduire d'environ 41% l'approximation de la fonction valeur. On peut également envisager des maintenances préventives lorsque la caméra est en état de fonctionnement. Les possibilités de maintenances sont multiples et permettent d'envisager de nombreuses suites à ce travail.

La mise en place de la méthode numérique d'approximation de la fonction valeur, issue de de Saporta et Dufour (2012), dans le cadre de notre modélisation (chapitre 4), a permis d'obtenir des informations quant à l'efficacité des politiques d'intervention de référence. Elle a en effet montré que la politique d'intervention n°3 est proche de ce que l'on peut faire de meilleur lorsque les interventions sont autorisées uniquement quand la caméra est en panne. En revanche, si l'on considère des maintenances lorsque la caméra est dans l'état dégradé ou dans l'état de panne, la meilleure politique d'intervention (n°5) reste éloignée de l'approximation de la fonction valeur. Il peut donc être très intéressant de développer d'autres politiques de référence et de les tester afin de trouver éventuellement une politique dont le coût est plus proche de l'approximation de la fonction valeur, si elle existe.

La seconde partie de ce travail constitue une première étape dans le but de trouver une politique d'interventions dont le coût est proche de la fonction valeur. Nous avons construit explicitement une famille de stratégies ε -optimales pour un problème de contrôle impulsif avec facteur d'oubli. Pour cela, nous avons d'abord construit des approximations des fonctions valeurs en reprenant l'idée de Gugerli (1986) pour l'arrêt optimal. Nous avons ensuite construit explicitement une famille de stratégies, dont nous avons pu montrer que son coût associé correspondait aux fonctions d'approximations des fonctions valeurs définies au début de notre étude. Cela nous a permis de montrer que ces stratégies étaient ε -optimales.

Le principal avantage de la famille de stratégies construites dans cette thèse est le fait qu'elle ne requiert pas la résolution de problèmes complexes. En effet, à notre connaissance, seul l'article [Costa et Davis \(1989\)](#) propose une construction d'une famille de stratégies ε -optimales pour ce type de problème de contrôle impulsionnel. Cependant, la construction proposée par O.L.V. Costa et M.H.A. Davis nécessite non seulement la résolution d'une suite de problèmes d'arrêt optimal, mais également la connaissance d'une fonction valeur associée au problème, et ce en tout point de l'espace. La construction proposée dans cette thèse ne nécessite que la connaissance de la fonction du coût pour une stratégie sans intervention, ainsi que les opérateurs K et J et leurs itérations.

Une suite naturelle à ce travail est de développer la méthode numérique associée pour implémenter la construction de telles stratégies. L'idée est de se baser sur le travail réalisé dans [de Saporta *et al.* \(2010\)](#) dans le cadre d'un problème d'arrêt optimal. En utilisant notamment la quantification et une discrétisation adaptée de l'espace pour les opérateurs K et J , il est possible de proposer une version approchée des fonctions r et y introduites au paragraphe [5.2](#).

Une fois la méthode numérique définie pour la construction de stratégies ε -optimales, une perspective plus lointaine serait d'appliquer cette méthode dans le cas de l'optimisation de la caméra Sophie. On pourrait alors mettre en œuvre une politique d'intervention ε -optimale sur la caméra et ainsi, obtenir un bon compromis entre la disponibilité de la caméra auprès du client et le coût de maintenance de la caméra.

Annexe : simulations de loi exponentielle et de loi de Weibull

Les simulations de loi exponentielle et de loi de Weibull se font par la méthode de la fonction de répartition inverse (voir Coccozza-Thivent (1997) p.15), qui est basée sur la proposition suivante :

Proposition .0.1

Soit F la fonction de répartition d'une loi de probabilité pour laquelle F est inversible, et G son inverse. Soit U une variable aléatoire de loi uniforme sur $[0; 1]$. Alors

$$X = G(U)$$

est une variable aléatoire de fonction de répartition F .

La fonction de répartition d'une loi de Weibull de paramètres (α, β) , où $\alpha, \beta > 0$, est issue de la proposition 2.2.1 :

$$\forall t > 0, \quad F(t) = 1 - \exp \left\{ - \left(\frac{t}{\alpha} \right)^\beta \right\}.$$

Cette fonction est inversible, d'inverse :

$$\forall u \in [0; 1[, \quad F^{-1}(u) = \alpha [-\log(1 - u)]^{1/\beta}.$$

Par conséquent, une loi de Weibull de paramètres (α, β) peut être simulée par l'algorithme suivant :

$$X = \alpha * (-\log(\text{rand}))^{1/\beta}.$$

Cependant, on aura besoin de simuler un loi de Weibull de paramètres (α, β) sachant qu'elle a dépassé un seuil $s > 0$ fixé. Cette simulation sera nécessaire pour déterminer les futurs temps de panne pour des éléments ayant déjà été utilisés. La fonction de répartition de cette loi s'obtient par le calcul suivant, pour $t > s$:

$$\begin{aligned} \mathbb{P}(X > t | X > s) &= \frac{\mathbb{P}(X > t)}{\mathbb{P}(X > s)} \\ &= \exp \left(- \left(\frac{t}{\alpha} \right)^\beta - \left(\frac{s}{\alpha} \right)^\beta \right) \\ &= \exp \left(- \frac{t^\beta - s^\beta}{\alpha^\beta} \right), \end{aligned}$$

ce qui donne

$$\forall t > s, \quad F_s(t) = 1 - \exp \left\{ - \left(\frac{t^\beta - s^\beta}{\alpha^\beta} \right) \right\}.$$

Cette fonction est inversible et afin de déterminer cet inverse, on fixe $t > s$ et pour $u \in [0; 1[$, on a les équivalences suivantes :

$$\begin{aligned} F_s(t) \leq u &\Leftrightarrow 1 - \exp \left\{ - \left(\frac{t^\beta - s^\beta}{\alpha^\beta} \right) \right\} \leq u \\ &\Leftrightarrow \ln(1 - u) \leq - \frac{t^\beta - s^\beta}{\alpha^\beta} \\ &\Leftrightarrow s^\beta - \alpha^\beta \ln(1 - u) \geq t^\beta \\ &\Leftrightarrow (s^\beta - \alpha^\beta \ln(1 - u))^{\frac{1}{\beta}} \geq t \end{aligned}$$

La fonction de répartition inverse, notée G_s , est alors donnée par :

$$\forall u \in [0; 1[, \quad G_s(u) = (s^\beta - \alpha^\beta \ln(1 - u))^{1/\beta}.$$

Ainsi, pour simuler une loi de Weibull de paramètres (α, β) sachant qu'elle a dépassé le seuil $s > 0$, on utilise l'algorithme 4.

Algorithme 4 simulation d'une loi de Weibull ayant pour paramètres (a, b) et ayant dépassé le seuil s

Entrées : a premier paramètre de la loi de Weibull

b second paramètre de la loi de Weibull

s seuil dépassé

Sortie : X simulation de la loi $\mathcal{W}(\alpha, \beta)$ ayant dépassé le seuil s

$$X = (s^b - a^b * \log(rand))^{1/b}$$

La loi exponentielle $\mathcal{E}(1/\lambda)$ étant un cas particulier de la loi de Weibull, on peut utiliser l'algorithme 4 pour la simuler en mettant en entrées $a = \lambda$, $b = 1$ et $s = 0$, ou bien créer un deuxième algorithme qui la génère (algorithme 5).

Algorithme 5 simulation d'une loi exponentielle de paramètre $1/a$

Entrée : a inverse du paramètre de la loi exponentielle

Sortie : X simulation de la loi $\mathcal{E}(1/a)$

$$X = -a * \log(rand)$$

Bibliographie

- AMAMI, R. (2012). *Contrôle impulsif appliqué à la gestion de changement de technologie dans une entreprise*. PhD thesis, University of Toulouse III.
- AZAÏS, R. (2013). *Estimation non paramétrique pour les processus Markoviens déterministes par morceaux*. PhD thesis, University of Bordeaux 1.
- AZAÏS, R. (2014). A recursive nonparametric estimator for the transition kernel of a piecewise-deterministic Markov process. *ESAIM Probab. Stat.*, 18:726–749.
- AZAÏS, R., BARDET, J.-B., GÉNADOT, A., KRELL, N. et ZITT, P.-A. (2014a). Piecewise deterministic Markov process—recent results. In *Journées MAS 2012*, volume 44 de *ESAIM Proc.*, pages 276–290. EDP Sci., Les Ulis.
- AZAÏS, R., DUFOUR, F. et GÉGOUT-PETIT, A. (2014b). Non-parametric estimation of the conditional distribution of the interjumping times for piecewise-deterministic Markov processes. *Scand. J. Stat.*, 41(4):950–969.
- BALLY, V. et PAGÈS, G. (2003). A quantization algorithm for solving multi-dimensional discrete-time optimal stopping problems. *Bernoulli*, 9(6):1003–1049.
- BARDET, J.-B., CHRISTEN, A., GUILLIN, A., MALRIEU, F. et ZITT, P.-A. (2013). Total variation estimates for the TCP process. *Electron. J. Probab.*, 18:no. 10, 21.
- BÄUERLE, N. et RIEDER, U. (2011). *Markov decision processes with applications to finance*. Universitext. Springer, Heidelberg.
- BAYSSE, C. (2013). *Analyse et optimisation de la fiabilité d'un équipement opto-électronique équipé de HUMS*. PhD thesis, University of Bordeaux 1.
- BENSOUSSAN, A. et LIONS, J. L. (1975). Nouvelle méthode en contrôle impulsif. *Applied Math. Optim.*, 1:289–312.
- BERMAN, S. (1994). Optimal timing of antiviral therapy in hiv infection. *J. Appl. Probab.*, 31A:3–15.
- BERTAIL, P., CLÉMENÇON, S. et TRESSOU, J. (2008). A storage model with random release rate for modeling exposure to food contaminants. *Math. Biosci. Eng.*, 5(1):35–60.
- BOUGUET, F. (2015). Quantitative speeds of convergence for exposure to food contaminants. *ESAIM Probab. Stat.*, 19:482–501.
- BRANDEJSKY, A., de SAPORTA, B. et DUFOUR, F. (2012a). Numerical method for expectations of piecewise deterministic Markov processes. *Commun. Appl. Math. Comput. Sci.*, 7(1):63–104.
- BRANDEJSKY, A., DE SAPORTA, B. et DUFOUR, F. (2012b). Numerical methods for the exit time of a piecewise-deterministic Markov process. *Adv. in Appl. Probab.*, 44(1):196–225.

- BRANDEJSKY, A., de SAPORTA, B. et DUFOUR, F. (2013). Optimal stopping for partially observed piecewise-deterministic Markov processes. *Stochastic Process. Appl.*, 123(8):3201–3238.
- CHAFAI, D., MALRIEU, F. et PAROUX, K. (2010). On the long time behavior of the TCP window size process. *Stochastic Process. Appl.*, 120(8):1518–1534.
- CHARLIER, I. (2015). *Conditional quantile estimation through optimal quantization*. PhD thesis, Université de Bordeaux 1, Université libre de Bruxelles.
- COCOZZA-THIVENT, C. (1997). *Processus stochastiques et fiabilité des systèmes*, volume 28 de *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer-Verlag, Berlin.
- COSTA, M. (2016). A piecewise deterministic model for a prey-predator community. *arXiv preprint*, arXiv :1503.02492v4.
- COSTA, O. L. V. (1991). Impulse control of piecewise-deterministic processes via linear programming. *IEEE Trans. Automat. Control*, 36(3):371–375.
- COSTA, O. L. V. et DAVIS, M. H. A. (1988). Approximations for optimal stopping of a piecewise-deterministic process. *Math. Control Signals Systems*, 1(2):123–146.
- COSTA, O. L. V. et DAVIS, M. H. A. (1989). Impulse control of piecewise-deterministic processes. *Math. Control Signals Systems*, 2(3):187–206.
- COSTA, O. L. V. et DUFOUR, F. (2008). Stability and ergodicity of piecewise deterministic Markov processes. *SIAM J. Control Optim.*, 47(2):1053–1077.
- COSTA, O. L. V., RAYMUNDO, C. et DUFOUR, F. (2000). Optimal stopping with continuous control of piecewise deterministic Markov processes. *Stochastics Stochastics Rep.*, 70(1-2):41–73.
- COSTA, O. L. V. et RAYMUNDO, C. A. B. (2000). Impulse and continuous control of piecewise deterministic Markov processes. *Stochastics Stochastics Rep.*, 70(1-2):75–107.
- CRUDU, A., DEBUSSCHE, A., MULLER, A. et RADULESCU, O. (2012). Convergence of stochastic gene networks to hybrid piecewise deterministic processes. *Ann. Appl. Probab.*, 22(5):1822–1859.
- DASSIOS, A. et EMBRECHTS, P. (1989). Martingales and insurance risk. *Comm. Statist. Stochastic Models*, 5(2):181–217.
- DAVIS, M. H. A. (1984). Piecewise-deterministic Markov processes : a general class of nondiffusion stochastic models. *J. Roy. Statist. Soc. Ser. B*, 46(3):353–388. With discussion.
- DAVIS, M. H. A. (1993). *Markov models and optimization*, volume 49 de *Monographs on Statistics and Applied Probability*. Chapman & Hall, London.
- DAVIS, M. H. A., DEMPSTER, M. A. H., SETHI, S. P. et VERMES, D. (1987). Optimal capacity expansion under uncertainty. *Adv. in Appl. Probab.*, 19(1):156–176.
- de SAPORTA, B. et DUFOUR, F. (2012). Numerical method for impulse control of piecewise deterministic Markov processes. *Automatica J. IFAC*, 48(5):779–793.
- de SAPORTA, B., DUFOUR, F. et GEERAERT, A. (2017). Optimal strategies for impulse control of piecewise deterministic Markov processes. *Automatica J. IFAC*, 77:219–229.

- de SAPORTA, B., DUFOUR, F. et GONZALEZ, K. (2010). Numerical method for optimal stopping of piecewise deterministic Markov processes. *Ann. Appl. Probab.*, 20(5):1607–1637.
- de SAPORTA, B., DUFOUR, F. et ZHANG, H. (2015). *Numerical methods for simulation and optimization of piecewise deterministic Markov processes*. Mathematics and Statistics Series. ISTE, London ; John Wiley & Sons, Inc., Hoboken, NJ. Application to reliability.
- DEMPSTER, M. A. H. et YE, J. J. (1995). Impulse control of piecewise deterministic Markov processes. *Ann. Appl. Probab.*, 5(2):399–423.
- DOUMIC, M., HOFFMANN, M., KRELL, N. et ROBERT, L. (2015). Statistical estimation of a growth-fragmentation model observed on a genealogical tree. *Bernoulli*, 21(3):1760–1799.
- DUFOUR, F. et COSTA, O. L. V. (1999). Stability of piecewise-deterministic Markov processes. *SIAM J. Control Optim.*, 37(5):1483–1502 (electronic).
- ERBAN, R. et OTHMER, H. G. (2004). From individual to collective behavior in bacterial chemotaxis. *SIAM J. Appl. Math.*, 65(2):361–391 (electronic).
- FAGGIONATO, A., GABRIELLI, D. et CRIVELLARI, M. R. (2010). Averaging and large deviation principles for fully-coupled piecewise deterministic Markov processes and applications to molecular motors. *Markov Process. Related Fields*, 16(3):497–548.
- GĄTAREK, D. (1992). Optimality conditions for impulsive control of piecewise-deterministic processes. *Math. Control Signals Systems*, 5(2):217–232.
- GONZALEZ, K. (2010). *Contribution à l'étude des processus markoviens déterministes par morceaux : étude d'un cas-test de la sûreté de fonctionnement et problème d'arrêt optimal à horizon aléatoire*. PhD thesis, University of Bordeaux 1.
- GRAF, S. et LUSCHGY, H. (2000). *Foundations of quantization for probability distributions*, volume 1730 de *Lecture Notes in Mathematics*. Springer-Verlag, Berlin.
- GUGERLI, U. S. (1986). Optimal stopping of a piecewise-deterministic Markov process. *Stochastics*, 19(4):221–236.
- HILL, T. (2009). Savoir quand s'arrêter. *Pour la Science*, (381):24–30.
- LENHART, S. et LIAO, Y.-C. (1985). Integro-differential equations associated with optimal stopping time of a piecewise-deterministic process. *Stochastics and Stochastics Reports*, 15:183–207.
- LENHART, S. M. (1989). Viscosity solutions associated with impulse control problems for piecewise-deterministic processes. *Internat. J. Math. Math. Sci.*, 12(1):145–157.
- LEPELTIER, J.-P. et MARCHAL, B. (1984). Théorie générale du contrôle impulsif markovien. *SIAM J. Control Optim.*, 22(4):645–665.
- NIVOT, C. (2016). *Analyse et étude des processus markoviens décisionnels*. PhD thesis, University of Bordeaux 1.
- PAGÈS, G., PHAM, H. et PRINTEMS, J. (2004). Optimal quantization methods and applications to numerical problems in finance. In *Handbook of computational and numerical methods in finance*, pages 253–297. Birkhäuser Boston, Boston, MA.
- PAGÈS, G. et PRINTEMS, J. (2003). Optimal quadratic quantization for numerics : the Gaussian case. *Monte Carlo Methods Appl.*, 9(2):135–165.

- PAKDAMAN, K., THIEULLEN, M. et WAINRIB, G. (2010). Fluid limit theorems for stochastic hybrid systems with application to neuron models. *Adv. in Appl. Probab.*, 42(3):761–794.
- PHAM, H. (1997). Optimal stopping, free boundary, and American option in a jump-diffusion model. *Appl. Math. Optim.*, 35(2):145–164.
- PHAM, H. (2006-2007). Méthodes de quantification optimale et applications en finance. *In Poly-copié de cours Master 2, Paris 7, Statistiques et Modèles Aléatoires en Economie et Finance.*
- PHAM, H., RUNGALDIER, W. et SELLAMI, A. (2005). Approximation by quantization of the filter process and applications to optimal stopping problems under partial observation. *Monte Carlo Methods Appl.*, 11(1):57–81.
- ROBIN, M. (1978). *Contrôle impulsif des Processus de Markov*. Thèse d'état, Paris IX.
- SCHÄL, M. (1998). On piecewise deterministic Markov control processes : control of jumps and of risk processes in insurance. *Insurance Math. Econom.*, 22(1):75–91. The interplay between insurance, finance and control (Aarhus, 1997).