UNIVERSITE DE GENEVE
Faculté des Sciences Economiques et Sociales
Section des Hautes Etudes Commerciales

Stochastic Optimization: Numerical Methods

Thèse présentée par

Francesco Moresino

pour l'obtention du grade de
Docteur ès Sciences Economiques et Sociales
mention gestion d'entreprise

Membres du jury:
Professeur Manfred Gilli, Département d'Econométrie, Université de Genève,
Professeur Alain Haurie, Section HEC, Université de Genève, directeur de thèse,
Professeur Yves Smeers, Université de Louvain-La-Neuve,
Professeur Jean-Philippe Vial, Section HEC, Université de Genève, président du jury.

Thèse no 491
Genève, 1999

La Faculté des sciences économiques et sociales, sur préavis du jury, a autorisé l'impression de la présente thèse, sans entendre, par là, émettre aucune opinion sur les propositions qui s'y trouvent énoncées et qui n'engagent que la responsabilité de leur auteur.

Genève, le 29 octobre 1999

Le doyen
Beat BÜRGENMEIER

Impression d'après le manuscrit de l'auteur

Contents

Abstract

In the present dissertation, we study three numerical methods, each adapted to a special stochastic optimization problem. More precisely, we propose and implement two numerical methods and prove their convergence property. Both methods are based on a time discretization. The first method is adapted to the resolution of piecewise deterministic control systems and will be used to compute an approximation of the optimal scheduling rule in a flexible workshop. The second method is adapted to the resolution of piecewise deterministic games where the information structure is S-adapted and will be used to compute the equilibrium strategies in a piecewise deterministic oligopoly.

In addition to these two methods, we implement a decomposition method adapted to hybrid stochastic models with two time scales. This method has been proposed by Filar and Haurie. The originality of this approach lies in the coupling of a linear programming method with a policy improvement algorithm. Another valuable contribution is the implementation of a parallel version of the method.

Remerciements

Je remercie infiniment le Professeur Alain Haurie pour son soutien et pour la confiance qu'il m'a accordée durant ces années de travail passées sous sa direction. Qu'il sache que ce fût pour moi un immense plaisir et un enrichissement, tant du point de vue personnel que scientifique. Je lui exprime toute ma reconnaissance pour la patience avec laquelle il a lu et relu ce travail.

Je tiens à remercier le Professeur Jean-Philippe Vial pour les nombreux conseils et encouragements prodigués dans le cadre de ce travail et tout particulièrement lors de mon initiation à la méthode ACCPM. Mes remerciements s'adressent également à Messieurs les professeurs Manfred Gilli et Yves Smeers pour avoir accepté de juger cette thèse et de faire partie du jury.

Je ne remercierai jamais assez le Professeur Emmanuel Fragnière qui m'a fait connaître le monde de la recherche opérationnelle. Il est l'instigateur de cette thèse, car sans son implication personnelle je ne serais jamais devenu l'assistant du Professeur Haurie.

J'aimerais tout particulièrement remercier Robert Sarkissian pour son aide précieuse et permanente, aussi bien dans les domaines informatique que mathématique.

Un immense merci à Odile Pourtallier et Mabel Tidball pour les nombreuses aides qu'elles m'ont apportées. J'ai particulièrement apprécié leur accueil, lors de mes deux séjours à Sophia Antipolis.

J'adresse un grand merci au Professeur Jacek Gondzio pour les innombrables conseils qu'il m'a prodigués. Je remercie également Cathy Wolfheim pour avoir lu et corrigé l'introduction de cette thèse. Mes remerciements vont aussi à Roman Kanala et Christian Condevaux-Lanloy sans lesquels mon ordinateur n'aurait jamais fonctionné. Enfin, je remercie Olivier Péton, pour avoir partagé le même bureau durant plus de trois années sans le moindre conflit, et tous les membres du LOGILAB qui m'ont accompagné durant ces années de recherche et d'assistanat.

La recherche présentée dans cette thèse a été, en grande partie, subventionnée par le Fonds National Suisse de la Recherche Scientifique.

List of abbreviations

We give here the abbreviations used throughout the present dissertation.

ACCPM	Analytic Center Cutting Plane Method
DP	Dynamic Programming
HJB	Hamilton-Jacobi-Bellman
LP	Linear Program
MDP	Markov Decision Problem
MFCP	Manufacturing Flow Control Problem
PDCS	Piecewise Deterministic Control System
POL	Piecewise Open Loop
SP	Stochastic Programming

Chapter 1

Introduction

1.1 Motivations

Many interesting models in economic sciences involve dynamic systems subject to uncertainties. This can be shown by a brief survey of the literature in, for example, production management (see [47]), investment planning (see [8] and [91]) or finance (see [71] and [93]). The theory of stochastic optimization offers tools to handle a broad variety of those models. It happens that, for most real life problems, economists have to utilize complex models for which an analytical solution cannot be obtained. In such circumstances, numerical methods offer an interesting alternative to approximate the solution of these complex models.

Unfortunately, in many applications, the size of the optimization problem is such that numerical methods reach their limit. This is the so-called "curse of dimensionality" already pointed out by Bellman [11]. The resolution of such problems is therefore only possible if they can be simplified. When the model allows it, time-scale decomposition is a possibility to achieve this simplification. This permits one to split the initial problem into many smaller weakly linked problems. The decomposition can be achieved by applying the theory of singularly perturbed systems. In addition to the simplification obtained, the decomposition permits a parallel implementation which decreases the computation time.

Many economic models also involve competition between several agents, and cannot, in general, be reduced to an optimization problem. However, as in the case of classical optimization, numerical methods may offer a way to approximate the solution to these problems, as long as the curse of dimensionality does not strike again.

The aim of the present work is threefold. Firstly, we propose to exploit mathematical programming techniques in stochastic optimization. These techniques are expected to be more efficient than the usual dynamic programming methods, but, they are valid only when the stochastic process does not depend on the actions. Secondly, when the stochastic process does depend on the actions and hence when only dynamic programming approaches can be used, we explore a time-scale decomposition method in order to simplify the initial problem. Thirdly, we show how to extend the stochastic programming approach to multi-agent models.

1.1.1 Stochastic optimization and manufacturing flow control

Manufacturing engineering is a domain where researchers have been particularly active in applying and developing the theory of stochastic optimization. Nowadays the market calls for rapid changes, due mainly to the fast evolution of technology and the rapid changes of consumer desires. In addition, the globalization of markets increases the competition pressure on firms. Consequently, to survive, firms have to adapt their production quickly to the changes in demand. Moreover, an increasing number of firms have adopted just-in-time production which requires sufficient flexibility. These facts show that the cornerstone of the competitiveness of a firm is its flexibility.

To achieve flexibility, firms invest in machines which can perform multiple operations. Such flexible machines are equipped with magazines containing all the tools the machine requires. When, in a production sequence, the machine has to perform a different operation than the preceding one, the tools are changed automatically. Flexible machines are efficient but also very expensive. Good scheduling is therefore important in order to utilize the machine to its maximal capacity so as to recover the high cost of investment.

The scheduling has to take into account the uncertainty concerning the availability of the machines or the level of the demand. One is confronted with a complex class of optimization problems under uncertainty for which there exists an analytical solution only for the simplest cases, for example a single machine producing a single part. In the more general context the challenge still remains to implement efficient scheduling rules for real-life production systems.

One possibility to cope with the complexity of the problem is to compute suboptimal scheduling rules (see, *e.g.*, [45] and [22]). Although appealing, such an approach is obviously not totally satisfactory, as one cannot guarantee the performance level of the suboptimal scheduling rules. Another approach consists of applying numerical methods to approximate, as closely as desired, the optimal scheduling rules. This approach has been used by several authors (see, *e.g.*, [15] and [16]) who implemented dynamic programming techniques. But these methods suffers from the curse of dimensionality and tend to become ineffective for stochastic models with more than two parts.

The first part of this thesis is a contribution to the modeling of stochastic manufacturing systems. We propose a new approach combining stochastic programming and simulation that will be valid when the disturbance jump process does not depend on the control.

1.1.2 Singularly perturbed systems and time-scale decomposition

Numerous interesting cases in economic science can be cast in the formalism of stochastic control problems that are amenable to the dynamic programming approach. It is well known that dynamic programming is a numerical method which attains rapidly its limits when the state space of the control system is too large. In this context, the resolution of these problems is only possible if they can be simplified. To achieve this simplification, a *decomposition* of the initial problem into many smaller weakly linked problems can be a powerful tool. The two main possibilities are *spatial* and *time-scale* decompositions.

Spatial decomposition is possible, for example, for a model of a firm with one parent company and many subsidiary companies. Each subsidiary company can be seen as an isolated system linked to the parent company: once the general strategy of the parent company is known, each subsidiary company has its own autonomy.

Time scale decomposition arises when events occur at very different time scales. Each time scale is associated with a hierarchical level of decomposition. Roughly speaking, this decomposition concept is based on two approximations: given a level, events at higher levels occur with lower frequencies and are considered constant whereas event at lower levels occur with higher frequencies and are approximated with their average rate.

In manufacturing engineering the concept of hierarchical decomposition has been intensively used with success (see Gershwin [47]). Moreover, decomposition permits parallel implementation which can drastically decrease the computation time.

It has been shown in [35] and [36] that time scale decomposition can be achieved via the methodology developed for singularly perturbed system. In the second part of this thesis, we will implement this theory using an example of a production system with two time scales.

1.1.3 Stochastic games

In economic sciences, one is very often depicting competition between several agents and, in this context, the classical theory of optimization is inappropriate. However, Game theory, which is a mathematical theory of conflicts, offers tools to handle multi-agent optimization. Game theory was initiated in the early 19th century with Cournot's work on oligopolistic competition (see [24]). In the middle of the 20th century von Neumann and Morgenstern [120] proposed a first general theory of games. Since then, Game theory has become a very useful and popular tool for economists, as testified by the growing literature in this field and the recent attribution of the Nobel prize in economics to J. Nash, R. Selten and J. Harsanyi for their pioneering works in this field.

Good methods are available to compute equilibria in a *static* context. Unfortunately, interesting economic problems occur often in *dynamic* context. For dynamic games there exist only a few methods to compute equilibria; these encompass a more restricted class of problems than those for static games. A second difficulty is often associated with the dynamic structure: for an important part of economic reality, *uncertainty* about the future plays an important role. To take into account uncertainty in his models, the economist has then to deal with stochastic dynamic games. In this class of games there are few cases for which there exists an efficient numerical method.

The third part of the thesis is a contribution to the computation of stochastic Nash-Cournot equilibrium: we propose a numerical method to compute the equilibrium of a class of dynamic stochastic game, namely the piecewise deterministic nonzero-sum games.

1.2 Literature review

1.2.1 Manufacturing flow control

Manufacturing control models with random events can be modeled as systems with jump Markov disturbances. These systems are defined over an hybrid state where a discrete stochastic state is described by a controlled Markov chain while a continuous state is governed by a set of deterministic differential equations depending on the value of the discrete state. Sworder [116] and Wonham [121] were the first to study this class of stochastic control systems in the linear-quadratic context. For more general cases, the most relevant publication is Rishel's paper [102], wherein the author established stochastic minimum principles that have to be satisfied at optimality.

Olsder & Suri [97] proposed a stochastic control model based on Rishel's formalism for production planning in a flexible manufacturing system. In their model, each machine is subject to random failure according to an homogeneous Markov process. They obtained a solution for a simple problem but described the derivation of the solution for real-life problems as practically impossible.

In [77], Kimemia and Gershwin proposed a three level[1] decomposition method for approximating the optimal policy in complex manufacturing systems. The top level is dedicated to evaluate the value function. To obtain an approximation of the value function, the authors suggested decomposing the nth-order Bellman partial differential equation that has to be satisfied by the value function in n first order ordinary differential equation. This is achieved by approximating the polyhedral production capacity sets by inscribed or circumscribed hypercubes.

[1]The decomposition proposed in this paper has four levels. However, Maimon and Gershwin [89] found that the two intermediate levels should not be separated. Consequently the decomposition method has three levels.

Once the value function is estimated, the middle level can compute the instantaneous production rates by solving a linear program. At the lower level the production rates are translated into actual part dispatch actions. The authors also showed that the optimal policy is a so-called hedging point policy, where a safety stock is maintained when the capacity is sufficient in order to hedge against future capacity shortages.

Gershwin, Akella and Choong proposed in [45] a new implementation of the hierarchical policy proposed by Kimemia and Gershwin. They postulated a quadratic value function which permitted them to reduce the complexity of the top level problem and simplified the middle-level computations. At the lower level they devised a simple rule to dispatch parts to the work stations so as to reduce the accumulation of material into the system.

In [5], Akella and Kumar studied the problem of controlling the production rate over an infinite horizon for a manufacturing system, producing a single commodity with one unreliable machine, so as to minimize the discounted inventory cost. They show that the optimal solution is a hedging stock policy and give an explicit formula for the hedging stock level.

The same problem, where *discounted* inventory cost are replaced with *average* inventory cost, has been considered by Bielecki and Kumar [12]. They studied the steady-state probability distribution of the inventory level under a hedging stock strategy and used it to obtain the optimal hedging stock value. They showed that zero-inventory is optimal when the system has sufficient excess capacity.

The case where the machine can be in one of many possible states of wear, has been studied by Sharifnia in [109]. In this paper, the author derived equations for the steady-state probability distribution of the surplus level, when a hedging stock strategy is used. Once this distribution function is determined the average surplus cost is easily calculated in terms of the value of the hedging points. This average cost is then minimized to find the optimal hedging points.

Extending the work of Bielecki & Kumar [12] and Sharifnia [109], Algoet [6] derived, for the multiple part-type and multiple machine model, a system of linear first-order partial differential equation that has to be satisfied by the steady-state probability distribution of the surplus level. However, analytical solutions are known only for the single machine single part-type model under a hedging point strategy.

For complex manufacturing systems where stochastic events occur at very different time scales, Gershwin [46] proposed a hierarchical approach with many levels. This approach is based on two approximations: given a level of hierarchy, events at higher levels occur with lower frequencies and are considered constant whereas event at lower levels occur with higher frequencies and are approximated with their average rate.

In [15], Boukas and Haurie presented a model of flexible manufacturing system including age-dependent machine failure rates and allowing preventive maintenance. By using an adaptation of the approximation technique initially proposed by Kushner (see [83]), they computed the optimal control for a single part-type two-machine system.

Caramanis and Sharifnia [22] proposed a near optimal controller whose design is computationally feasible for realistic size systems. The design uses a decomposition of the multiple part-type problem to many analytically tractable one part-type problems. The decomposition is achieved by replacing the polyhedra production capacity sets with inscribed hypercubes, as suggested by Kimemia and Gershwin in [77]. They presented computational results for a three-machine-state two-part-type flexible manufacturing system.

In [90], Malhamé and Boukas showed that the transient statistical evolution of a single product manufacturing system under hedging point control policies is characterized via a system of coupled partial differential equations. They demonstrated the Markov renewal nature of the controlled process. This allowed them to characterize the ergodicity, the steady state (if it exists) and the speed of convergence to steady state.

Haurie and van Delft [66] established a relationship between hedging points and the turnpike property. In this paper they studied a general class of piecewise deterministic systems that includes manufacturing flow control models. They used the Markov renewal decision process formalism to characterize optimal policies via a discrete event dynamic programming approach. Then, they associated with these optimality conditions a family of control problems with random stopping time. Finally, they showed that these problems can be reformulated as infinite horizon deterministic control problems, for which the turnpike property holds under convexity assumptions.

The theoretical results of this paper permitted Boukas, Haurie and van Delft [16] to propose a numerical technique, called “turnpike improvement”, for approximating the solution of manufacturing flow control problems.

Sharifnia, Caramanis and Gershwin [111] investigated setup scheduling problem in the context of a multilevel hierarchy of discrete events with distinct frequencies. They considered a feedback setup scheduling policy which uses corridors in surplus space to determine the timing of the setup changes. They also determined conditions for linear corridors which result in a stable limit cycle of the trajectory of the surplus levels.

In [21], Caramanis and Liberopoulos proposed a near optimal controller design technique. This technique estimates the parameters of a quadratic approximation of the value function that characterizes the optimal policy. A sample trajectory of the system's performance under a trial controller is simulated to provide a measure of its performance. From this sample run, infinitesimal perturbation analysis extracts gradient information with respect to controller parameters. Gradient and performance estimates are then employed to optimize the parameters in order to obtain a near optimal controller.

In [81] Krichagina, Lou, Sethi and Taksar considered the problem of a single unreliable machine producing a single product with a stochastic demand. More precisely the demand has two components: one is deterministic with a constant rate whereas the other is stochastic with random demand batches. Under heavy loading conditions they showed that the control problem can be approximated by a singular stochastic control problem, which can be solved explicitly.

Following the work of Caramanis and Liberopoulos [21], Haurie, L'Ecuyer and van Delft [63] used stochastic approximation coupled with infinitesimal perturbation analysis method to approximate optimal control in a manufacturing control model. They considered an *a priory* fixed class of feedback control laws depending on a (small) finite set of parameters and proved that, under appropriate conditions, stochastic approximation coupled with infinitesimal perturbation analysis converges to the best policy in this class.

Using a similar approach to that used by Akella and Kumar [5], Hu, Vakili and Yu [69] analyzed the one machine one part-type infinite horizon discounted cost problem, where the failure rate of the machine depends on the rate of production. They showed that the linearity of the failure rate function is both necessary and sufficient for the optimality of the hedging point policy.

Hu and Xiang [70] studied the same one part-type one machine model as Akella and Kumar in [5] but replaced the *exponential* failure and repair times with *general* times. They showed that if the failure rate is increasing (respectively decreasing) with the age of the machine, the hedging stock will also increase (respectively decrease) with the age of the machine.

In [114], Srivatsan and Dallery studied the performance of hedging point policies in a single-unreliable-machine two-part-type system. They first generalized known stability results for hedging point policies in single-part-type systems and extended them to two-part-type systems. They used an average cost analysis of trajectories to *partially* characterize an optimal policy belonging to this class. However, they obtained a *complete* characterization for special parameter values.

We end this survey of the literature on manufacturing flow control by recommending the works of Gershwin [47] and Sethi & Zhang [107], where most of the above quoted papers are nicely summarized.

1.2.2 Singularly perturbed systems

Schweitzer [106] seems to be the first who studied singularly perturbed Markov chains. In the last section of his paper he showed that a perturbation formalism exists in the multiple subchain case under some assumptions.

Although Schweitzer's paper deals with uncontrolled Markov chains, it inspired numerous applications concerning controlled Markov chains. For example, an interesting application in the domain of hydropower production has been conduced by Delebecque and Quadrat in [29]. In this paper, they used averaging and singular perturbation techniques to approach the optimal policy for the management of a system of hydroelectric dams.

In [30], Delebecque and Quadrat studied the optimal control of Markov chains with an almost block diagonal structure, where the generator is composed of N diagonal blocks linked by a small additive perturbation. They proposed a policy improvement algorithm involving only decentralized computations within the N blocks and computations relative to an aggregate N states Markov chain.

In [13], Bielecki and Filar considered a singularly perturbed MDP with the limiting average cost criterion. They proved that an optimal solution to the perturbed MDP can be approximated by an optimal solution of a limit Markov control problem. They also demonstrated that the limit Markov control problem is equivalent to a suitably constructed nonlinear program in the space of long-run state-action frequencies.

In [2], Abbad and Filar proposed a unified approach to the asymptotic analysis of a Markov decision process disturbed by ϵ-additive perturbation. They studied the limiting average and the discounted cost criterion and obtained results for a general additive perturbation.

In [3], Abbad, Filar and Bielecki considered a singularly perturbed MDP with the limiting average cost criterion. They assumed that the underlying process is composed of several separable irreducible processes, which are "united" into a single irreducible process by a small perturbation. They proposed two algorithms for the solution of the limit Markov control problem. The first is a linear program possessing the primal block-diagonal structure inherited from the underlying process, whereas the second is an aggregation-disaggregation policy improvement algorithm.

Filar, Gaitsgory and Haurie [33] studied, in the context of singular perturbation theory, the stochastic control of hybrid systems involving two different time scales. The fast mode of the system is represented by deterministic state equations whereas the slow mode of the system corresponds to a jump disturbance process. They considered both the finite and infinite discounted horizon cases and showed how an approximate optimal control law can be constructed from the solution of the limit control problem.

In [35], Filar and Haurie studied singularly perturbed systems composed of a fast mode, described as a deterministic or stochastic diffusion subsystem and a slow mode described as a jump process. They dealt with an infinite time horizon and limit average cost criterion and showed the convergence of the optimal average cost of the perturbed system toward the optimal average cost of a limit-control problem. They hinted to the connection which can be established with the theory of singularly perturbed controlled Markov chains ([13], [2], [3]) when a numerical method is implemented as indicated in [83]. In a second paper [36], they detailed this numerical method for a production system and established a link between the theory of singularly perturbed controlled switching diffusions and the theory of singularly perturbed controlled Markov chains.

1.2.3 Continuous time stochastic games

The third part of the present thesis deals with piecewise deterministic differential games with S-adapted[2] information structure, a special case of the class of dynamic stochastic games played in continuous time. Before reviewing the literature on this specific topic, let us briefly survey the three classes of games, which are close to our interest, namely the differential games, the stochastic Markov games and the stochastic differential games.

Introduced by Isaacs in [73], the study of differential games was motivated by pursuit evasion problems (typically a jet fighter vs. a bomber). These game are played in continuous-time and the state variables obey differential equations. Four major monographs in this field are Isaacs [74] and Krassovski & Soubbotine [80] for two-player zero-sum differential games, Friedman [40] for more general differential games, and the book written by Başar and Olsder [9] which encompasses the wider class of dynamic games.

Stochastic Markov games were first introduced by Shapley [108] in 1957. These games are played in stages as follows. At each stage, the game is in one of a finite number states. Each player chooses his action in a set which depends on the current state of the game and receives a reward depending on the current state of the game and the actions chosen by all the players. Finally, the state of the game makes a random transition, also depending on the players' actions, to another state and the game continuous at the next stage. Actually, the state is a stochastic process which is described by a controlled Markov chain. Each player tries to maximize his expected total reward. Since the pioneering work of Shapley, there have been numerous publications in this field and a compilation of the most valuable contributions can be found in the monograph of Filar and Vrieze [34].

Friedman [41] seems to be the first who studied stochastic differential games. These games are played in continuous-time and the state obeys a stochastic differential equation. In [32], Elliott considered a two-person zero sum game and proved, when the Isaacs condition holds, the existence of a saddle point in feedback strategies. Uchida [117] considered the N-person nonzero-sum case and showed that if the Nash condition (generalized Isaacs condition) holds there is a Nash equilibrium point in feedback strategies.

We can now turn to the class of games that we will study in this dissertation, namely the piecewise deterministic differential games. This class of dynamic games has been considered by Haurie in [58] and [57]. These games are characterized by an hybrid state equation where a continuous state is governed by a deterministic differential equation, while a discrete stochastic state is described by a controlled Markov chain.

In [65], Haurie and Roche proposed a turnpike adjustment algorithm for a piecewise deterministic differential game played with a Piecewise Open Loop

[2] S-adapted for "sample path adapted".

(POL) information structure. In the POL information structure the players observe, at each jump time of the discrete state, both the discrete and the continuous states and then choose open-loop controls to be implemented until the next jump occurs. The algorithm [65] has been implemented for the computation of a POL Cournot equilibrium in a duopoly with random market condition.

In this thesis we will focus our interest on games played in the S-adapted information structure. S-adapted information structure was first introduced by Haurie, Smeers, Zaccour and Legrand[3] in two papers, [68] and [67]. It corresponds to an information structure where the players adapt their actions to an observation of the realization of the random disturbances affecting the game dynamics. These disturbances are supposed to take the form of a jump process. In [68] they exhibit some properties of the S-adapted information structure, whereas in [67] a detailed example dealing with the modeling of the competition in the European gas market is fully developed.

In [55], Gürkan, Özge and Robinson proposed a simulation-based method for solving stochastic variational inequality. This method, called sample-path optimization, was then applied on a slightly modified version of the model of European gas market proposed in [67].

In [99] Pineau and Murto proposed an example of a game with S-adapted information structure to analyze the Finnish electricity market. This market has the property of being strongly influenced by the price of gas. This motivated the authors to introduce in their model a stochastic jump process describing the price of the gas.

1.3 Thesis organization and contributions

In the present dissertation, we study three numerical methods, each adapted to a special stochastic optimization problem. More precisely, we propose and implement two numerical methods and prove their convergence property. Both methods are based on a time discretization. The first method is adapted to the resolution of piecewise deterministic control systems and will be used to compute an approximation of the optimal scheduling rule in a flexible workshop. The second method is adapted to the resolution of piecewise deterministic games where the information structure is S-adapted and will be used to compute the equilibrium strategies in a piecewise deterministic oligopoly.

In addition to these two methods, we implement a decomposition method adapted to hybrid stochastic models with two time scales. This method has been proposed by Filar and Haurie in [35] and [36]. The originality of this work is the coupling of a linear programming method with a policy improvement algorithm.

[3] Haurie, Zaccour and Smeers for the first paper; Haurie, Zaccour, Legrand and Smeers for the second.

Another valuable contribution is the implementation of a parallel version of the method.

Note that, for the sake of simplifying the reading of the thesis, the convergence proofs of each method is placed in an appendix at the end of the corresponding part. Finally let us mention that the three parts are self-contained and each one can therefore be read independently of the others.

1.3.1 Manufacturing flow control via stochastic programming methods

In the first part of the thesis we propose a new approach for the resolution of piecewise deterministic control systems. This method combines optimization and simulation and is valid when the disturbance jump process is not controlled. This method is used for the resolution of a flexible manufacturing system where the machines are subject to random failures and repairs.

In Chapter 3, we present the formulation of the manufacturing flow control problem we propose to solve. We consider a flexible workshop consisting of unreliable machines producing several part types and we wish to minimize work-in-process, inventory and backlog costs on a finite time horizon. This problem is a special instance of the class of piecewise deterministic control systems for which the dynamic programming solution leads to a system of coupled Hamilton-Jacobi-Bellman equations.

In Chapter 4, using a discretization of the continuous time scale, we reformulate the initial continuous time control problem as an approximating discrete time control problem which has the structure of a stochastic linear program. The convergence proof of the approximating value function to the value function of the initial problem, when the discretization step tends to zero, calls for techniques of approximation of viscosity solutions and is proposed in Appendix A.

An event tree describing the stochasticity of the system is associated to the stochastic program. In Chapter 5, we show how the large size of the event tree, and hence the large size of the stochastic program, can be reduced using a Monte-Carlo sampling method. The convergence, when the sample size increases to infinity, of the solution of the stochastic program defined on the randomly sampled tree toward the solution of the discrete time control problem follows from the strong law of large numbers.

In Chapter 6, we illustrate the convergence of the method on a single-part-type manufacturing flow control problem. For this small example a direct dynamic programming method is more efficient than the method presented in this work and will be taken as bench mark. We can show that for larger models, the method proposed in this presentation is more efficient than a direct dynamic programming approach.

Finally, in Chapter 7, we apply the numerical method to two examples that are close to real life implementation. The first example is a two-machine two-part-type flexible workshop whereas the second example is a six-machine four-part-type flexible workshop.

1.3.2 Decomposition method in a singularly perturbed hybrid stochastic model

The second part of the thesis is dedicated to the study of the optimal control of hybrid stochastic systems with two time scales. When the time scale ratio tends to zero, following the works of Filar and Haurie [35, 36], we obtain an approximation of the problem that can be solved as a block-diagonal linear programming problem, where each sub-block can be identified as an MDP. We propose the implementation of a decomposition approach coupling a linear programming method with a policy improvement algorithm. This coupling permits us to exploit optimally both the primal block-diagonal structure and the special structure of each sub-block.

In Chapter 10, we expose the two-time-scale hybrid stochastic control problem. The system is driven by two dynamics, fast and slow. The fast mode of the system is characterized by a continuous stochastic variable which takes the form of a controlled jump and diffusion process. The slow mode of the system is described by a discrete stochastic variable which takes the form of a controlled Markov jump process. One looks for an optimal control law such that the expected average reward per unit of time is maximized subject to the dynamics of the fast and slow processes.

In Chapter 11, following Filar and Haurie [35, 36], we obtain an approximation of the problem that can be solved as a structured linear programming problem. This approximation is obtained in two steps. Firstly, using the numerical method proposed by Kushner and Dupuis [83], we derive an approximation of the initial problem that can be solved as a Linear Program (LP). This LP does not exhibit a particular structure and for large models would be difficult to solve. Secondly, using the theory of singularly perturbed systems developed by Abbad, Bielecki and Filar [2] [3], we obtain a second LP which exhibits a primal block-diagonal structure. The solution of this second LP is an approximation of the solution of the first LP, when the time scale ratio tends to zero.

In Chapter 12, we explain how a decomposition method using the Analytic Center Cutting Plane Method (ACCPM) can be implemented for the resolution of the structured LP.

Finally, in Chapter 13, we apply the decomposition method to an example of production problem. The model concerns a firm producing one good with two different human resources. The size of each employee category is described by a continuous vector variable, whose dynamics is represented by a controlled jump and diffusion process. The control describes the enrollment effort of new employees, whereas the jump and diffusion term describes the uncontrolled departure or arrival of employees. The profit function depends on the number of employees, the enrollment effort and the random state of the market. The state of the market, *i.e.* the selling price, can be in one of four different states and is described by a controlled Markov jump process whose transition rates depend on the production level, which, in turn, depends on the number of employees. The time scale for the uncontrolled departure or arrival of employees is supposed to be much "faster" than the changes of market state. The objective of the firm is to control the number of different employees, in order to maximize its expected profit over an infinite horizon. With this example, we can show the advantages of the decomposition method coupled with a policy improvement algorithm compared with a frontal method. In particular we put in evidence the reduction of the RAM memory utilized, the reduction of the execution time and the accuracy of the solution concerning the policies. We also show that we can obtain a good speed-up in parallelizing the decomposition method.

1.3.3 Computation of S-adapted equilibria in piecewise deterministic games via stochastic programming methods

In the third part of the thesis we propose a new method to compute equilibria in a piecewise deterministic differential game played with S-adapted information structure. This method is based on a discretization of the continuous time scale and is valid when the disturbance jump process does neither depend on the continuous states nor on the controls.

Before studying the piecewise deterministic game, in Chapter 17, we investigate a deterministic version of the game. Briefly described, we consider an oligopoly where several competing firms supply a market for an homogeneous good. Each firm can control through investments its production capacity, which depreciates with time. The profit rate of a given firm is a function of the total production capacity of the different firms, *i.e.* the total supply, and the investment effort of this firm. Finally, the game is played with an open-loop information structure over a finite horizon. The uniqueness of the equilibrium is proved under concavity properties of the reward functions.

A discretization of the time scale permits one to reformulate the initial continuous time game as an approximating discrete time game. The open-loop equilibrium of this approximating game can be computed via the solution of a variational inequality, for which efficient methods exist. Under concavity properties of the reward functions it is shown that the equilibrium of the approximating game is unique. We then show that the equilibrium strategies of the approximating game permit one to construct ϵ-equilibrium strategies for the initial game.

In Chapter 18 we study a stochastic extension of this game, where the profit rate functions depend on the random market condition. The random market condition is described by an uncontrolled continuous-time discrete-state Markov chain taking value in a finite set. The game is played with an S-adapted information structure, which permits the players to adapt their actions to the realization of the random process but not to the action of the other players. This information structure has been introduced by Haurie, Zaccour and Smeers in Ref. [68] and Haurie, Zaccour, Legrand and Smeers in Ref. [67].

As in the deterministic game, we discretize the continuous time scale in order to obtain a discrete time approximating game. The S-adapted equilibrium of this approximating game can be computed via the solution of a variational inequality and is proven to be unique under concavity properties of the reward functions. We then show the proximity existing between the equilibrium strategies of the approximating game and those of the initial game.

Finally, in Chapter 19, we apply the method to a stochastic duopoly model already studied by Haurie and Roche in [65]. Haurie and Roche considered the Piecewise Open-Loop (POL) information structure, where the players observe, at each jump time, the market condition and the production capacity of each firm and then choose open-loop controls to be implemented until the next jump occurs. Numerical results confirm the conjecture that, when the jump Markov process is uncontrolled, S-adapted and POL equilibria are likely to yield quite close outcomes.

Part I

Manufacturing Flow Control via Stochastic Programming Methods

Chapter 2

Introduction to Part I

Piecewise deterministic control systems (PDCS) offer an interesting paradigm for the modeling of many industrial and economic processes. The theory developed by Wonham [121] or Sworder [115] for linear quadratic systems, Davis [27], Rishel [102, 103] and Vermes [119] for more general cases, has established the foundations of a dynamic programming (DP) approach for the solution of this class of problems. There are two possible types of DP equations that can be associated with a PDCS: (i) the Hamilton-Jacobi-Bellman (HJB) equations defined as a set of coupled (functional) partial differential equations (see *e.g.* [102, 103]); (ii) the discrete event dynamic programming equations based on a fixed-point operator à la Denardo [31] for a value function defined at jump times of the disturbance process (see *e.g.* [15]).

The modeling of manufacturing flow control processes has greatly benefited from the use of PDCS paradigms. Olsder & Suri [97] have first introduced this model for a flexible manufacturing cell where the deterministic system represents the evolution of parts surpluses and the random disturbances represent the machine failures and repairs. This modeling framework has been further studied and developed by many others (we cite Gershwin et al. [45], [44] and Akella & Kumar [5], Bielecki & Kumar [12] as a small sample of the large literature on these models, nicely summarized in the books of Gershwin [47] and Sethi & Zhang [107]). When the model concerns a single part system and the failure process does not depend on the part surplus and production control, an analytic solution of the HJB equations can be obtained as shown in [5]. As soon as the number of parts is two or more, an analytic solution is difficult to obtain and one has to rely on a numerical approximation technique. A solution of the HJB equations via the approximation scheme introduced by Kushner and Dupuis [83] has been proposed by Boukas & Haurie [15]. A solution of the discrete event dynamic programming equations via an approximation of the Denardo fixed-point operator has been proposed in Boukas, Haurie & Van Delft [16]. Both methods suffer from the *curse of dimensionality* and tend to become ineffective as the number of parts is three or over. Caramanis & Liberopoulos [21] have proposed an interesting approach based on the use of a sub-optimal class of controls, depending on a finite set of parameters, these parameters being optimized via an infinitesimal perturbation technique. Haurie, L'Ecuyer & Van Delft [60] have further studied and experimented such a method based on a combination of optimization and simulation.

In the first part of the present thesis we propose another approach combining optimization and simulation that will be valid when the disturbance jump process does not depend on the continuous state and control. The approach exploits the formal proximity which exists, under this assumption, between the PDCS formalism and the Stochastic Programming (SP) paradigm introduced in the realm of mathematical programming by Dantzig [26] and further developed by many others (see the survey book of Kall and Wallace [75] or the book of Infanger [72] as representatives of a vast list of contributions). The proposed method is based on a two-step approximation: (i) the state equations for the finite horizon continuous time stochastic control problem are discretized over a set of sampled times; this defines an associated discrete time stochastic control problem which, due to the finiteness of the sample path set for the Markov disturbance process, can be written as a SP problem. (ii) The very large event tree representing the sample path set is replaced with a reduced tree obtained by randomly sampling over this sample path set. It will be shown that the solution of the stochastic program defined on the randomly sampled tree converges toward the solution of the discrete time control problem when the sample size tends to infinity. The solution of the discrete time control problem converges to the solution of the flow control problem when the discretization mesh decreases. Therefore SP methods can be implemented to solve this class of PDCS and the recent advances in the numerical solution of very large scale stochastic programs can be exploited to obtain insight for problems that fall out of reach of standard dynamic programming techniques.

The first part of this thesis is organized as follows. In Chapter 3 we recall a formulation of the manufacturing flow control problem proposed by Sharifnia [110] with the PDCS formalism and the HJB equations one has to solve in order to characterize the optimal control. In Chapter 4 we construct the discrete time approximation leading to a SP problem which will be characterized, usually, by a very large event tree representing the uncertainties. In Chapter 5 we show how to use a random sampling of scenarios to reduce the size of the event tree and we prove convergence of this Monte-Carlo method. In Chapter 6 we compare different approaches on a simple single part model. In Chapter 7 we experiment the SP approximation method on two instances of a more realistic multi-part model. Finally, in Appendix A, we prove the convergence of the discrete time approximation using the theory of viscosity solutions.

Chapter 3

The manufacturing flow control problem

In this chapter we recall the model of a flexible manufacturing system which was proposed by Sharifnia in [110]. We have chosen this model since it was already linked to linear programming in a discrete time approximation of the solution of the manufacturing flow control problem (MFCP), in the absence of random disturbances. The random disturbances introduced in many formulations of the MFCP are represented as an uncontrolled Markov chain that describes the evolution of the operational state of the machines. Under these conditions, the discrete time approximation proposed in [110] will easily lend itself to a formulation as a stochastic linear programming problem.

3.1 The continuous flow formulation

We consider a flexible workshop consisting of M unreliable machines, and producing P part types. We use a continuous flow approximation to represent the production process. Each part, to be produced, has to visit some machines in a given sequence. We call this sequence a route. For a given part the route may not be unique, therefore there are R routes with $R \geq P$. An input buffer is associated with each machine. Set-up times are assumed to be negligible and processing times are supposed to be deterministic. An instance of this type of organization is represented in Figure 3.1. Assume that the machines are unreliable, the repair and failure times are exponentially distributed random variables. The demand is supposed to be known in advance. The objective is to minimize the expected cost associated with the work-in-process and finished parts' inventory.

For a more formal description of the model we introduce the following vari-

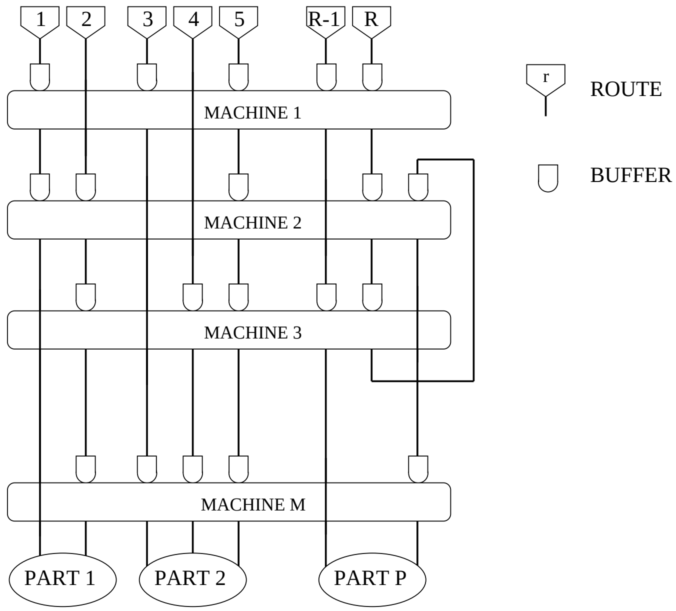


Figure 3.1: Flexible workshop producing P parts with M machines

ables.

$$
\begin{array}{rcll}
v(t) & = & (v_1(t), \ldots, v_B(t))^T & : \text{ buffer processing rates} \\
w(t) & = & (w_1(t), \ldots, w_R(t))^T & : \text{ part release rates} \\
q(t) & = & (q_1(t), \ldots, q_B(t))^T & : \text{ buffer levels} \\
d(t) & = & (d_1(t), \ldots, d_P(t))^T & : \text{ finished parts demand rates} \\
y(t) & = & (y_1(t), \ldots, y_P(t))^T & : \text{ finished parts surplus levels.}
\end{array}
$$

The state variables are $q(t)$ and $y(t)$, while $w(t)$, $v(t)$ are the control variables. The state equations are

$$\dot{q}(t) = A_1 v(t) + A_2 w(t) \tag{3.1}$$
$$\dot{y}(t) = A_3 v(t) - d(t) \tag{3.2}$$

where the term $A_1 v(t)$ in Eq. (3.1) represents the internal material flows among buffers, the term $A_2(t)w(t)$ in Eq. (3.1) represents external arrival into the system and the term $A_3 v(t)$ represents the arrival of finished parts in the last buffer. The i-th line of A_1 is composed of a -1 in cell (i, i), and a $+1$ in cell (i, j) if the buffer j is upstream to buffer i. All other entries of row i are 0 valued. The incidence matrix A_2 is of dimension $B \times R$, with $\{0, 1\}$ entries that determine which buffers receive the new arrivals. Eq. (3.2) represents the dynamics of finished parts surplus. The $P \times B$ matrix A_3 has a $+1$ in entry (i, j) if buffer j is the last buffer of a route for part i. Otherwise, this entry is 0.

Let τ_j denote the processing time of parts in buffer j and $\mathcal{B}^{(m)}$ be the set of buffers for machine m. The capacity constraints on the control are defined as follows:

$$\sum_{j \in \mathcal{B}^{(m)}} \tau_j v_j(t) \leq \xi^m(t) \quad m = 1, \ldots, M, \tag{3.3}$$

where $\{\xi^m(t) : t \geq 0\}$ is a continuous time Markov jump process taking the values 0 or 1. $\xi^m(t) = 1$ indicates that the machine m is operational (up) at time t, $\xi^m(t) = 0$ that it is not operational (down) at time t.

The following inequality constraints have to be satisfied

$$\begin{aligned} v(t) &\geq 0 && (3.4)\\ w(t) &\geq 0 && (3.5)\\ q(t) &\geq 0. && (3.6) \end{aligned}$$

Notice that (3.6) represents a state constraint.

Let's call

$$x(t) = (q(t), y(t))$$

the *continuous state* of the system and

$$\Xi(t) = (\xi^m(t))_{m=1,\ldots,M}$$

its *operational state* while

$$u(t) = (w(t), v(t))$$

is the control at time t. The operational state $\Xi(t)$ evolves as a continuous time Markov jump process with transition probabilities that are easily computed from the failure and repair rates of each machine

$$\begin{aligned} P[\Xi(t+dt) = j | \Xi(t) = i] &= q_{ij} dt + \circ(dt) \quad (i \neq j)\\ P[\Xi(t+dt) = i | \Xi(t) = i] &= 1 + q_{ii} dt + \circ(dt)\\ \lim_{dt \to 0} \frac{o(dt)}{dt} &= 0 \end{aligned}$$

for $i, j \in I = \{0,1\}^M$. As usual we define $q_{ii} = -\sum_{i \neq j} q_{ij}$.

A production policy γ can be viewed either as

- a piecewise open-loop control $u^{\Xi(t)}(t) : t \geq 0$ that is adapted to the vector jump process $\{\Xi(t) = (\xi^m(t))_{m=1,\ldots M} : t \geq 0\}$ and satisfies the constraints (3.3-3.6), when one uses a discrete event dynamic programming formalism

- a feedback law $u(t) = \gamma(t, x(t), \Xi(t))$, when one uses the coupled HJB dynamic programming equations formalism.

The variable $y^+(t) = (\max\{y_j(t), 0\})_{j=1,\dots,p}$ represents the inventory of finished parts while $y^-(t) = (\max\{-y_j(t), 0\})_{j=1,\dots,p}$ represents the backlog of finished parts. The objective is to find a policy γ^* which minimizes the expected total cost

$$E_\gamma[\int_0^T \{hq(t) + g^+y^+(t) + g^-y^-(t)\}\, dt], \tag{3.7}$$

where h, g^+ and g^- are cost-rate vectors for the work-in-process and finished parts inventory/backlog respectively.

3.2 The system of coupled HJB equations

To summarize, the optimal operation of the flexible workshop is a particular instance of a stochastic control problem

$$\begin{aligned}
J^i(0, x_0) &= \min_\gamma E_\gamma[\int_0^T L(x(t))\, dt] && (3.8)\\
\text{s.t.} & && \\
\dot{x}(t) &= f(x(t), u(t)) && (3.9)\\
P[\Xi(t+dt) &= j|\Xi(t) = i] = q_{ij}dt + o(dt) \quad (i \neq j) && (3.10)\\
P[\Xi(t+dt) &= i|\Xi(t) = i] = 1 + q_{ii}dt + o(dt) \quad i, j \in I, && (3.11)\\
\lim_{dt\to 0} \frac{o(dt)}{dt} &= 0 && (3.12)\\
u(t) &\in U^{\Xi(t)} && (3.13)\\
& \Xi(0) = i, x(0) = x_0 && (3.14)\\
& && (3.15)
\end{aligned}$$

where $L(x)$ and $f(x, u)$ satisfy the usual regularity assumptions for control problems.

Define the value functions

$$J^i(t, x) = \min_\gamma E_\gamma[\int_t^T L(x(s))\, ds|x(t) = x \text{ and } \Xi(t) = i], \quad i \in I. \tag{3.16}$$

If these functions are differentiable in x, then the optimal policy is characterized by a system of coupled HJB equations

$$\begin{aligned}
-\frac{\partial}{\partial t}J^i(t, x) &= \min_{u\in U^i}\{+L(x) + \frac{\partial}{\partial x}J^i(t, x)f(u) + \sum_{j\neq i} q_{ij}[J^j(t, x) - J^i(t, x)]\},\\
& \quad i \in I \quad t \in [0, T[&& (3.17)\\
J^i(T, x) &= 0 \quad \forall x. && (3.18)
\end{aligned}$$

When the value functions $J^i(t,x)$ is known, the optimal strategy $u^*(x,t,i)$ is obtained as the solution of a set of "static" optimization problems

$$\min_{u \in U^i} \frac{\partial}{\partial x} J^i(t,x) f(u). \tag{3.19}$$

In the case of our MFCP these problems reduce to simple linear programs.

The value function differentiability issue can be addressed through the use of the so-called viscosity solution (see Appendix A).

Theorem 1. *The optimal value function is obtained as the unique viscosity solution to the system of coupled HJB equations.*

Proof. The proof follows directly from theorems 4, 6 and 7 given in Appendix A. These theorems assume that the assumptions 1 to 4, stated in Appendix A, hold. Clearly we can easily show that assumptions 1, 2 and 4 are satisfied by the MFCP. Unfortunately, Assumption 3 is not satisfied because of Equation 3.6 which forbids negative buffer levels. However we can construct an auxiliary MFCP which allows negative buffer levels but penalizes this with a huge cost. This auxiliary MFCP satisfies assumptions 1 to 4 and therefore the conclusion of the present theorem are valid for its value function. If the penalty cost is big enough, the auxiliary MFCP has, at optimality, the same property as our MFCP, *i.e.* the value function and the optimal strategies are the same for both models. Consequently the conclusion of the present theorem are also valid for the original MFCP. □

Chapter 4

A stochastic linear programming reformulation

In this chapter we define a stochastic programming problem that will be used to approximate the solution of the MFCP under study.

4.1 A discrete time reformulation

We discretize time as in [110]. This permits us to replace the continuous time state equation with a difference equation and to approximate the continuous time Markov chain by a discrete time Markov chain. Let t_k denote the k-th sampled time point $k = 0, 1, \ldots, K$ with $t_0 = 0$ and $t_K = T$, $\delta t_k = t_k - t_{k-1}$, $\tilde{q}(k) := q(t_k)$, $\tilde{y}(k) := y(t_k)$, $\tilde{w}(k) := w(t_k)$, $\tilde{v}(k) := v(t_k)$, $\tilde{\xi}^m(k) := \xi^m(t_k)$ and replace the differential state equations with the difference equations:

$$\begin{aligned}
\tilde{q}(k) &= \tilde{q}(k-1) + \delta t_k A_1 \tilde{v}(k) + \delta t_k A_2 \tilde{w}(k) \\
\tilde{y}(k) &= \tilde{y}(k-1) + \delta t_k A_3 \tilde{v}(k) - \delta t_k \tilde{d}(k),
\end{aligned}$$

for $k = 1, \ldots, K$. The control and state constraints become

$$\begin{aligned}
\sum_{j \in \mathcal{B}^{(m)}} \tau_j \tilde{v}_j(k) &\leq \tilde{\xi}^m(k) \quad m = 1, \ldots, M, \\
\tilde{v}(k) &\geq 0 \\
\tilde{w}(k) &\geq 0 \\
\tilde{q}(k) &\geq 0 \quad , \quad k = 1, \ldots, K \\
\tilde{q}(0) &= \tilde{q}_0 \\
\tilde{x}(0) &= \tilde{x}_0.
\end{aligned}$$

Denote $\tilde{x}(k) = (\tilde{q}(k), \tilde{y}(k))^T$ the continuous state variables, $\tilde{u}(k) = (\tilde{w}(k), \tilde{v}(k))^T$ the control variables and $\tilde{\Xi}(k) = (\tilde{\xi}^m(k)_{m=1 \ldots M})$ the discrete state variable that evolves according to a Markov chain with transitions probabilities

$$\begin{aligned} P[\tilde{\Xi}(k+1)=j|\tilde{\Xi}(k)=i] &= q_{ij}\delta t_k \quad (i \neq j) \\ P[\tilde{\Xi}(k+1)=i|\tilde{\Xi}(k)=i] &= 1+q_{ii}\delta t_k. \end{aligned}$$

This time discretization can be envisioned when the average times to repair and failure are much greater than δt_k. The solution of the associated discrete time stochastic control problem can be obtained through the solution of the following discrete time DP equations:

$$\begin{aligned} &\tilde{J}^i(k-1,\tilde{x}(k-1)) \\ &= \min_{\tilde{u}(k)\in U^i}\{L(\tilde{x}(k))\delta t_k + \sum_{j\neq i} q_{ij}\delta t_k \tilde{J}^j(k,\tilde{x}(k)) + (1+q_{ii}\delta t_k)\tilde{J}^i(k,\tilde{x}(k))\} \end{aligned} \tag{4.1}$$

for $i \in I$ and $k = 1 \dots K$; with terminal conditions:

$$\tilde{J}^i(K,\tilde{x}(K)) = 0. \tag{4.2}$$

The following result can be established, using techniques of approximation of viscosity solutions (see Appendix A).

Theorem 2. *The solution of the discrete time DP equation (4.1,4.2) converges uniformly when $\delta t_k \to 0$ to the viscosity solution of the system of coupled HJB equation of the continuous time model (3.17,3.18).*

Proof. The proof follows directly from Theorem 10 given in Appendix A. Theorem 10 assumes that the assumptions 1 to 4, stated in Appendix A, hold. As we mentioned in the proof of Theorem 1, our MFCP does not, in general, satisfy Assumption 3. Consequently, to prove the present theorem, the same auxiliary MFCP and the same arguments as in the proof of Theorem 1 need to be used. ☐

4.2 The scenario concept

Since the disturbance Markov jump process is uncontrolled, the solution of the discrete time stochastic control problem can also be obtained via the so-called *stochastic programming* technique. This is a mathematical programming technique based on the concept of a scenario. For our problem we call *scenario* ω a sample path $\{(\tilde{\xi}^1_\omega(1),\dots,\tilde{\xi}^M_\omega(1)),\dots,(\tilde{\xi}^1_\omega(K),\dots,\tilde{\xi}^M_\omega(K))\}$ of the $\tilde{\Xi}(\cdot)$ process. On a time horizon of K periods, as the state in the first period is identical for all scenarios, the discrete time Markov chain will generate $(2^M)^{K-1}$ different scenarios. We denote $u_{\omega_i}(k)$ the control for period k when the realized scenario is ω_i.

For two scenarios ω_i and ω_j that satisfy

$$(\tilde{\xi}^1_{\omega_i}(k),\dots,\tilde{\xi}^M_{\omega_i}(k)) = (\tilde{\xi}^1_{\omega_j}(k),\dots,\tilde{\xi}^M_{\omega_j}(k)) \quad \forall k \leq l \tag{4.3}$$

the controls $u_{\omega_i}(k)$ and $u_{\omega_j}(k)$ must be equal for all $k \leq l$. These conditions are called the *nonanticipativity constraints.*

There are two possible ways to take these constraints into account in the optimization problem

(i) introduce as many subproblems as there are scenarios and couple them through the nonanticipativity constraints explained above,

(ii) handle the scenario tree on a node by node basis with the nonanticipativity constraint taken into account implicitly.

The second approach is usually preferable because it reduces the number of constraints in the associated mathematical program. Let $\mathcal{N}(k) = \{\mathcal{N}_1(k), \ldots, \mathcal{N}_{i_k}(k)\}$ be the set of the nodes at period k. For each scenario ω and for each period k, ω passes through one and only one node $\mathcal{N}_i(k)$ (that we denote $\omega \hookrightarrow \mathcal{N}_i(k)$). If ω_i and ω_j are indistinguishable until the period l, that is if (4.3) holds, then they share the same node $\mathcal{N}_i(k)$ at all periods $k \leq l$. Note that since all scenarios are indistinguishable in the first period, we have only one node for this period, *e.g.* $\mathcal{N}(1) = \{\mathcal{N}_1(1)\}$. Each node n, except $\mathcal{N}_1(1)$ noted n_1, has a direct ancestor, denoted $\mathcal{A}(n)$, in the set of the nodes of the previous period. If ω passes through $\mathcal{N}_i(k)$ at period $k > 1$, then it passes through the ancestor of $\mathcal{N}_i(k)$ at period $k-1$. The set of all scenarios passing through the node $\mathcal{N}_i(k)$ is denoted by $N_i(k)$. The probability of the node $\mathcal{N}_i(k)$ is then

$$p(\mathcal{N}_i(k)) = \sum_{\omega \hookrightarrow N_i(k)} p(\omega)$$

where $p(\omega)$ denotes the probability of the scenario ω. We must then index each variable on the node set: $\tilde{q}_n(k)$, $\tilde{y}_n(k)$, $\tilde{i}_n(k)$, $\tilde{v}_n(k)$, $\tilde{\xi}_n^m(k)$ for all $n \in \mathcal{N}(k)$.

To illustrate this representation, consider a workshop of one machines with an horizon of 3 periods. In the first period the machine is up. There exist 4 scenarios which are listed in Figure 4.1. In the scenario ω_1 the machine is up during all periods. In the scenario ω_2 (resp. ω_3) the machine is up during all periods except period 3 (resp. period 2). In the last scenario ω_4, the machine is down during all periods except period 1. For example, the scenario ω_2 is defined by $(\tilde{\xi}(1), \tilde{\xi}(2), \tilde{\xi}(3)) = (1, 1, 0)$. $\mathcal{N}(2)$, the set of nodes at period 2, contains two nodes: $\mathcal{N}_1(2)$ and $\mathcal{N}_2(2)$. The direct ancestor of $\mathcal{N}_2(3)$ is $\mathcal{N}_1(2)$.

4.3 A linear stochastic programming problem

To summarize, we have to solve a stochastic linear program with the objective function

$$\tilde{J}^{0,\tilde{\Xi}}(\tilde{x}_0) = \min \sum_{k=1}^{K} \sum_{n \in \mathcal{N}(k)} p(n)\{h\tilde{q}_n(k) + g^+\tilde{y}_n^+(k) + g^-\tilde{y}_n^-(k)\}\delta t_k. \quad (4.4)$$

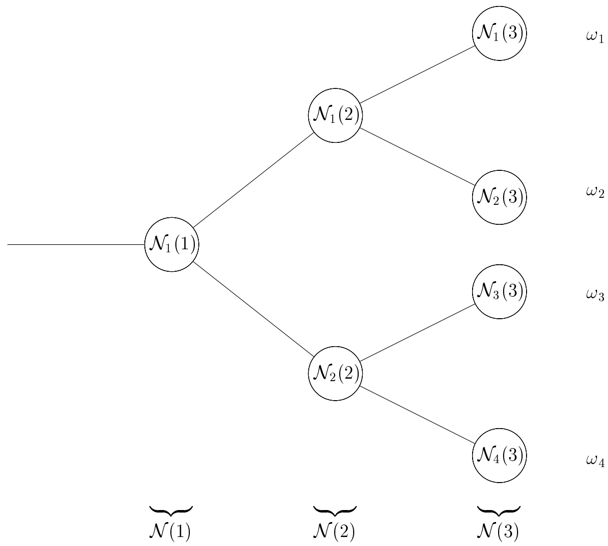


Figure 4.1: The scenario tree of a workshop with 1 machine and 3 periods

For the first period the constraints are

$$\begin{aligned}
\tilde{q}_{n_1}(1) &= \tilde{q}_0 + \delta t_1 A_1 \tilde{v}_{n_1}(1) + \delta t_1 A_2 \tilde{w}_{n_1}(1) && (4.5)\\
\tilde{y}_{n_1}(1) &= \tilde{y}_0 + \delta t_1 A_3 \tilde{v}_{n_1}(1) - \delta t_1 \tilde{d}_{n_1}(1), && (4.6)\\
\sum_{j \in \mathcal{B}(m)} \tau_j (\tilde{v}_{n_1})_j(1) &\leq \tilde{\xi}^m_{n_1}(1) \quad m = 1, \ldots, M, && (4.7)
\end{aligned}$$

with the initial conditions

$$\begin{aligned}
(\tilde{\xi}^m_{n_1}(1))_{m=1,\ldots,M} &= \tilde{\Xi}\\
(\tilde{q}_0, \tilde{y}_0) &= \tilde{x}_0.
\end{aligned}$$

For each period $k = 2 \ldots K$ the following constraints must hold for $n \in \mathcal{N}(k)$:

$$\begin{aligned}
\tilde{q}_n(k) &= \tilde{q}_{\mathcal{A}(n)}(k-1) + \delta t_k A_1 \tilde{v}_n(k) + \delta t_k A_2 \tilde{w}_n(k) && (4.8)\\
\tilde{y}_n(k) &= \tilde{y}_{\mathcal{A}(n)}(k-1) + \delta t_k A_3 \tilde{v}_n(k) - \delta t_k \tilde{d}_n(k), && (4.9)\\
\sum_{j \in \mathcal{B}(m)} \tau_j (\tilde{v}_n)_j(k) &\leq \tilde{\xi}^m_n(k) \quad m = 1, \ldots, M. && (4.10)
\end{aligned}$$

For $k = 1 \ldots K$ and $n \in \mathcal{N}(t)$ the following non-negativity constraints must hold.

$$\tilde{v}_n(k), \quad \tilde{w}_n(k), \quad \tilde{q}_n(k) \geq 0. \tag{4.11}$$

The optimal policy $\tilde{\gamma}^*$ is then described by the controls $\tilde{u}_n(k) = (\tilde{w}_n(k), \tilde{v}_n(k))$ for $n \in \mathcal{N}(t)$.

4.4 Identification of hedging points

The stochastic programming formulation will be used primarily for a computation of the control law at the initial time 0. Using parametric analysis we will be able to identify a suboptimal policy for running the flexible manufacturing system in a stationary (ergodic) environment. The optimal control for an MFCP is often an hedging point policy. In the continuous time, infinite horizon HJB formulation the hedging point corresponds to the minimum of the (potential) value function. In a finite time horizon formulation, the minimum of the value function at time 0 will tend to approximate the optimal hedging point when the horizon increases. In our discrete time, finite horizon formulation, if we let the initial stocks $\tilde{q}_0$ and $\tilde{y}_0$ be free variables, their optimal values will therefore give an indication of the hedging points. Actually, the discretization of time will often eliminate the uniqueness of the hedging points defined as the minimum of the value function. It will be then useful to identify the hedging point as the initial state for which the actual optimal production rate is exactly equal to the demand rate.

Chapter 5

Approximating the stochastic linear program by sampling

In this chapter we propose a sampling technique to reduce the size of the stochastic programming problem one has to solve to approximate the control policy.

5.1 The approximation scheme

To solve the linear stochastic program introduced in chapter 4, we have to consider the event tree representing the $(2^M)^{K-1}$ different possible scenarios. This number of possible scenarios increases exponentially with the number of periods and the problem becomes rapidly intractable. To reduce the size of the problem we extract a smaller event tree composed of randomly sampled scenarios.

Only the control for the first period is really relevant and we want to find the optimal policy $\gamma^*(t, x(t), \Xi(t))$ for $t = 0$. We will solve the sampled stochastic programming model for different initial states $\tilde{x}_0$ on a given finite grid G. Then the control $\gamma^*(0, x(0), \Xi(0))$ is approximated by $\tilde{u}^*(1)$, the solution for the first period in the sampled stochastic programming model when $\tilde{\Xi} = \Xi(0)$ and where $\tilde{x}_0$ is the nearest point to $x(0)$ in G.

5.2 Convergence of the sampled problems solutions

Let us introduce a few simplifying notations. Consider a discrete probability space $(\Omega, \mathcal{B}, P)$, where Ω is the finite set of possible realizations ω of the uncertain parameters and P the corresponding probability distribution. As Ω is finite, the event set is $\mathcal{B} = 2^\Omega$. Let $S = |\Omega|$ be the number of different scenarios. The elements of Ω are denoted $\Omega = \{\omega_1, \ldots, \omega_S\}$. Let $p(\omega_i)$ denote the probability of the realization ω_i. A generic stochastic optimization problem can be represented as a convex optimization problem (here x and y are used to represent generic variables in an optimization problem; they don't have the signification given to them in the MFCP)

$$\begin{aligned} z &= \min \sum_{\omega\in\Omega} f(x,\omega)p(\omega) && (5.1)\\ \text{s.t.} & & \\ x &\in C \subseteq \mathbf{R}^n && (5.2) \end{aligned}$$

We assume that $f(x,\omega)$ is convex in x on the convex set C but not necessarily differentiable. This formulation (5.1, 5.2) encompasses the classical two-period stochastic program with recourse

$$f(x,\omega) = cx + \min_y C(\omega)y$$

$$\begin{aligned} s.t \quad & D(\omega)y = d(\omega) + B(\omega)x \\ & y \geq 0 \end{aligned}$$

$$C = \{x \in \mathbf{R}^n | \quad Ax = b, \quad x \geq 0\}.$$

In this formulation the variable x represents the decision in the first period and y is the recourse in second period. Once the optimization w.r.t. y has been done for each possible realization ω, the problem is reduced to the form (5.1, 5.2).

The stochastic programming problem obtained from the time discretization of the MFCP can also be put in the general form (5.1, 5.2) through a nested reduction of a sequence of two stage stochastic programming problems. The variable x will then represent the decision variables for the initial period (the one we are particularly interested in).

We now formulate an approximation of the generic problem obtained through a random sampling scheme. A sampled problem, with sample size m, is obtained, if we draw randomly m scenarios among the S possible. A specific scenario ω_i is selected at a given draw with probability $p(\omega_i)$. We denote $\varpi^m = \{\varpi_j, j = 1,\ldots,m\}$, the scenario sample thus obtained. The sampled SP problem is defined as

$$\begin{aligned} z^{\varpi^m} &= \min_x \frac{1}{m} \sum_{j=1}^{m} f(x,\varpi_j) && (5.3)\\ \text{s.t.} & & \\ x &\in C \subseteq \mathbf{R}^n. && (5.4) \end{aligned}$$

Let ν_i be the observed frequency of scenario ω_i in the sample ϖ^m. If we denote by $w_i = \frac{\nu_i}{m}$ the observed proportion of scenario ω_i, the problem (5.3, 5.4) can also be reformulated as

$$z^{\varpi^m} = \min_x \sum_{i=1}^{S} f(x, \omega_i) w_i \tag{5.5}$$

s.t.

$$x \in C \subseteq \mathbf{R}^n. \tag{5.6}$$

The convergence of the sampled problem solution to the original solution is stated in the following theorem.

Theorem 3. *When $m \to \infty$ the solution z^{ϖ^m} of the sampled stochastic optimization problem (5.3, 5.4)converges almost surely to the solution z of the original stochastic optimization problem (5.1, 5.2).*

Proof. According to the strong law of large numbers we know that the observed proportions $(w_i)_{i=1,\ldots,S}$ converge almost surely to the probabilities $p(\omega_i)_{i=1,\ldots,S}$ when the sample size m tends to infinity. Furthermore, one can easily show that the function $\min_{x \in C} \sum_{i=1}^{S} f(x, \omega_i) p_i$ is convex, and therefore continuous, in $(p_i)_{i=1,\ldots,S} \in \mathbf{R}^S$. These two properties lead to the desired result. $\square$

Chapter 6

Empirical verification of convergence

In this chapter we illustrate the convergence of the SP method on a single-machine single-part-type MFCP. The solution of this MFCP in the infinite horizon case, obtained in [12], is recalled in the first section. A solution for the finite horizon case has been proposed in [122] under the rather strong assumption that once the machine fails it will never be repaired. In the general case with finite horizon there is no analytical solution available, however a direct numerical solution of the HJB equations can be obtained with good accuracy, using the weak convergence technique proposed by Kushner and Dupuis [83]. This alternative numerical solution will be used to control the convergence of our sampled SP models.

Indeed for this example the direct solution of the dynamic programming equations is more efficient than the sampled SP method. However, when there are two or more part-types we expect the sampled SP method to be more efficient than the direct dynamic programming method.

6.1 The infinite horizon case

We consider as a test problem the single-machine single-part-type example proposed by Bielecki and Kumar [12]. The problem is:

$$\min_{\gamma} E_{\gamma} \left[\lim_{T \to \infty} \frac{1}{T} \int_0^T L(x(t))\, dt \right]$$

$$\begin{aligned} s.t. \quad & \dot{x}(t) = u(t) - d \\ & P[\Xi(t+dt) = j | \Xi(t) = i] = q_{ij} dt + \circ(dt) \quad (i \neq j) \\ & P[\Xi(t+dt) = i | \Xi(t) = i] = 1 + q_{ii} dt + \circ(dt) \\ & u(t) \in U^{\Xi(t)} \quad U^0 = \{0\} \quad U^1 = [0, u_{\max}] = [0, \frac{1}{\tau}] \\ & \Xi(0) = i \in \{0,1\} \\ & x(0) = x_0 \end{aligned}$$

With $L(x(t)) = g^+ x^+(t) + g^- x^-(t)$.

The HJB system of equations is:

$$g = L(x) - \frac{\partial}{\partial x} W^0(x) d + q_{01}[W^1(x) - W^0(x)]$$

$$g = \min_{u \in U^1} \{L(x) + \frac{\partial}{\partial x} W^1(x)(u - d) + q_{10}[W^0(x) - W^1(x)]\}$$

Where g is the minimum expected cost growth rate and $W^i(x)$ is the differential cost to go function at initial state i.

Bielecki and Kumar have shown that the optimal policy is defined by

$$u^*(x) = \begin{cases} u_{\max} & \text{if } x < Z \\ d & \text{if } x = Z \\ 0 & \text{if } x > Z \end{cases}$$

where Z is the so-called hedging point given by

$$Z = \begin{cases} \frac{\ln(ab(1+\frac{g^-}{g^+}))}{b} & \text{if } g^+ - b(g^+ + g^-) < 0 \\ 0 & \text{otherwise} \end{cases} \tag{6.1}$$

with

$$b = \frac{q_{01}}{d} - \frac{q_{10}}{u_{\max} - d}$$

and

$$a = \frac{u_{\max} q_{10}}{b(q_{01} + q_{10})(u_{\max} - d)} \quad .$$

the minimal average cost per unit of time g is given by

$$g = g^+(Z - a) + a(g^+ + g^-) \exp(-bZ).$$

6.2 The finite horizon case

The same model with a finite time horizon T is a particular case of the model developed earlier. Although an analytical solution is not available, an accurate numerical solution can be obtained via a direct solution of the dynamic programming equations. This numerical solution shows that for the finite-time horizon the optimal control is still an hedging point policy but with a safety stock that decreases when one gets closer to the end of horizon T, i.e.

$$u^*(x,t) = \begin{cases} u_{\max} & \text{if } x < Z(T-t) \\ d & \text{if } x = Z(T-t) \\ 0 & \text{if } x > Z(T-t), \end{cases}$$

where $Z(\cdot)$ is an increasing function called the *hedging curve*.

6.3 Accuracy of the SP solution

We solve the finite horizon model with the following data ([47] p.292): $g^+ = 1$, $g^- = 10$, $d = 0.5$, $q_{01} = 0.09 = -q_{11}$, $q_{10} = 0.01 = -q_{00}$ and $\tau = 1$.

To control the convergence of our SP solution, we implemented the method of Kushner and Dupuis ([83] Chapter 12) on the x-state space grid

$$G = \{-30, -29.99, -29.98, \ldots, 70\}$$

and with a time step 0.001. The hedging point computed according to (6.1) is $Z = 4.9279$. The solid line in Figure 6.1 is the hedging curve obtained via the Kushner and Dupuis numerical technique. One notices that, as expected, the hedging curve tends asymptotically to the hedging point value 4.9279 when the horizon increases.

The size of the associated stochastic programming model increases exponentially with the number of periods K. The largest possible value K permitted by the memory on our machine (IBM RISC 6000, with 128 Mb memory, running SP/OSL software) was equal to 13 corresponding to 4096 different scenarios. For the computations concerning a model with more than 13 periods, we applied the following recursive method. We first compute the value functions $\tilde{J}^0(0,\cdot)$ and $\tilde{J}^1(0,\cdot)$, defined in Equation (4.4), for 13 periods. Then, in the objective function, a piecewise linear approximation of each value functions is introduced as a terminal cost penalty. The value functions of this new model are computed and a piecewise linear approximation of each of this new value function is introduced in the objective function. We can repeat this recursive procedure as often as desired. This corresponds to a value iteration on a two stage dynamic programming process.

In the SP approach we have identified the initial stock for which the optimal policy in the first period is to produce the same amount as the demand d. As

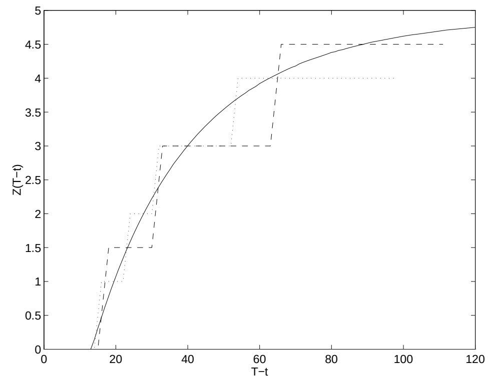


Figure 6.1: The hedging curve $Z(T-t)$

noticed previously these values correspond to the hedging points. We notice that the time discretization yields an approximation of the exact hedging curve by a discontinuous function which remains however quite close to $Z(T-t)$.

Figure 6.1 compares the value of $Z(T-t)$ obtained via three different methods

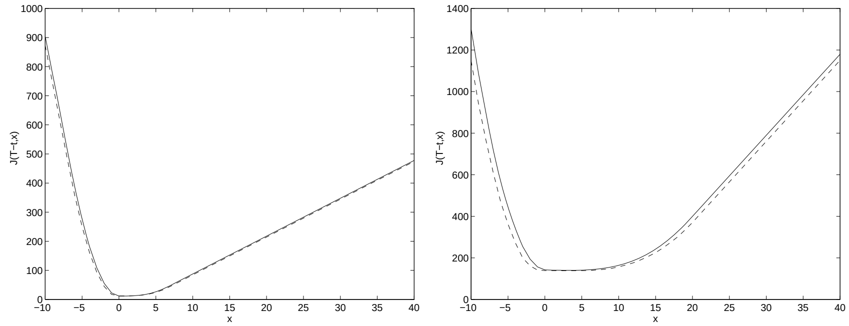
- the solid line corresponds to the solution of the dynamic programming equations obtained via the Kushner and Dupuis method.
- the dashed line shows the solution obtained with the SP method where $\delta t = 3.0$.
- the dotted line shows the solution obtained with the SP method where $\delta t = 2.0$.

It can be observed that the hedging curve $Z(T-t)$ is approximated in the SP approach by a discontinuous function with values $\delta t \cdot d \cdot I$ where I is an integer and d is the demand rate.

In Figure 6.2 we have represented the value function $J^1(T-t,x)$ with a solid line, when evaluated by a direct solution of the DP equations and a dashed line when evaluated through the SP approach with $K = 13$.

6.4 Accuracy of the SP solution with sampling

We investigate now the convergence of the solution of the SP method with sampling to the solution of the SP method when the whole scenario tree is taken into account. For all numerical experiments of this section we use $\delta t = 3$. As we noticed in the previous section, the approximation of the hedging curve $Z(T-t)$ obtained with the SP reformulation is a step function. Consequently there are times to go $T-t$ at which $Z(T-t)$ is discontinuous. Therefore we investigated the SP method with sampling for the computation of $Z(T-t)$ at two possible



(a) $T - t = 13$ (b) $T - t = 39$

Figure 6.2: $J^1(T-t,x)$ versus x. SP dashed.

	sample size: 500	sample size: 5000	sample size: 50000
sample no.1	3	3	3
sample no.2	0	3	3
sample no.3	0	1.5	3
sample no.4	1.5	3	3
sample no.5	3	3	3
sample no.6	1.5	3	3

Figure 6.3: $Z(T-t)$ for $T-t=33$.

values of the time to go $T-t$: one near a discontinuity ($T-t=33$) and one far from a discontinuity ($T-t=39$). We took different sample sizes to construct the approximating event tree and the results are shown in Figure 6.3 for the time to go $T-t=33$ and in Figure 6.4 for the time to go $T-t=39$. We see that a sample size of 500 is not sufficient for both cases. A sample size of 5000 is sufficient for $T-t=39$ but not for $T-t=33$. However a sample size equal to 50000 is sufficient for $T-t=33$.

	sample size: 500	sample size: 5000
sample no.1	3	3
sample no.2	3	3
sample no.3	3	3
sample no.4	6	3
sample no.5	3	3
sample no.6	1.5	3

Figure 6.4: $Z(T-t)$ for $T-t=39$

Chapter 7

Numerical experiments

In this chapter we apply the numerical method presented in this part of the dissertation to two examples that are closer to a real life implementation. In the first section we approximate the optimal strategy for a flexible workshop with two machines and two part types. As the size of the model is not too big, we display the optimal strategy in full details and discuss the results. In the second section we study a larger system, namely a flexible workshop with six machines and four part types. Due to the size of the model, the optimal strategy cannot be fully displayed in a simple figure and therefore only the optimal hedging stocks are given.

7.1 Implementation

Our approximation scheme leads to the solution of a stochastic program. To generate and solve the stochastic program we coupled two softwares: AMPL and SP/OSL. AMPL [38] is a modeling language for mathematical programming, which is designed to help formulate models, communicate with a variety of solvers, and examine the solutions. SP/OSL [78] is an interface library of C-language subroutines that supports the modeling, construction and solution of stochastic programs.

We obtain the solution of the stochastic program in four steps.

(i) We describe the flexible workshop topology using the algebraic facilities of AMPL. First we model the flexible workshop without the stochasticity on the machine availability (all machines are always up). This corresponds to a single scenario from the scenario tree which is from now on called the *base case scenario*.

(ii) The base case scenario is passed, in an MPS file, to SP/OSL and the whole stochastic program is constructed by specifying for every possible scenario the difference with the base case scenario and its probability or its sampled

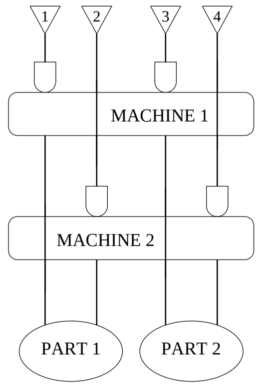


Figure 7.1: Flexible workshop producing two parts with two machines

frequency. All scenarios with null probability are discarded. The sampled scenarios are then aggregated into a scenario tree.

(iii) The stochastic program is solved with SP/OSL routines, which implement a Benders decomposition.

(iv) The results are graphically displayed using MATLAB.

7.2 Two-machine two-part-type example

The example considered is a flexible workshop composed of two machines producing two parts. One operation has to be performed on each part on either the first machine or the second one, thus there are four routes. The first machine is specialized on the first part and the second machine is specialized on the second part. The processing time τ for each part, is equal to 0.004 for the specialized machine and 0.008 for the other machine. The penalty for work-in-process, for finished part inventory and backlog are the following:

$$\begin{aligned} h &= (1,1,1,1) \\ g^{+} &= (5,5) \\ g^{-} &= (8,8). \end{aligned}$$

The failure rate is equal to 0.3 for the first machine and 0.1 for the second one. The repair rate is equal to 0.8 for the first machine and 0.5 for the second one. The demand is supposed to be constant at 200 units per period for each part. The flexible workshop is represented in Figure 7.1. We consider a time horizon $T = 8$ with $K = 8$ periods. The total number of possible scenarios is about 16000 and we took as sample size $m = 10000$.

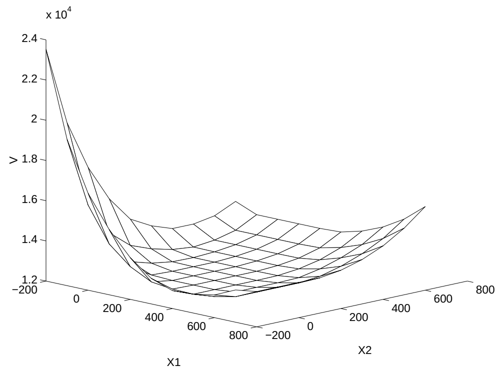


Figure 7.2: The function $z^*(\tilde{y}_0, \tilde{\Xi})$. In the initial period the two machines are up

For this simple example, as only one operation has to be performed on each part, it is penalizing to have non-zero inventory in the internal buffers. So the state $\tilde{x}(k)$ is reduced to $\tilde{y}(k)$ and the policy $\tilde{u}(k)$ is fully determined by $\tilde{v}(k)$. For the finite grid G approximating $\tilde{x}_0$ we took the following values:

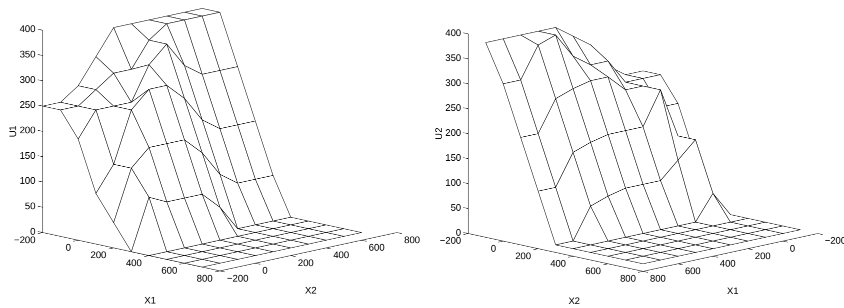
$$\tilde{y}_0 \in G = \{(\tilde{y}_1(0), \tilde{y}_2(0)) | \quad \tilde{y}_1(0), \tilde{y}_2(0) \in \{-200, -100, 0, 100, \ldots, 700\}\}.$$

The value function $\tilde{J}^{\tilde{\Xi}}(0, \tilde{y}_0)$ is shown in Figure 7.2 for $\tilde{\Xi} = (1, 1)$. In this figure, we see that the value function attains a minimum on a plateau. The values of $\tilde{y}_0$ that minimize this function can be regarded as hedging points. Due to the time discretization, the set of hedging points is not, as in the continuous time case, a curve or a point, but a surface. For other values of Ξ, the value function presents the same general shape.

For convenience, the optimal policy in the first period is rearranged as follows: the total amount of part 1 produced during the first period is denoted by $U1(\tilde{y}_0, \tilde{\Xi})$, and the total amount of part 2 produced during the first period is denoted by $U2(\tilde{y}_0, \tilde{\Xi})$. The functions $U1(\tilde{y}_0, \tilde{\Xi})$ and $U2(\tilde{y}_0, \tilde{\Xi})$ are shown in Figure 7.3 for $\tilde{\Xi} = (1, 1)$ and in Figure 7.4 for $\tilde{\Xi} = (1, 0)$.

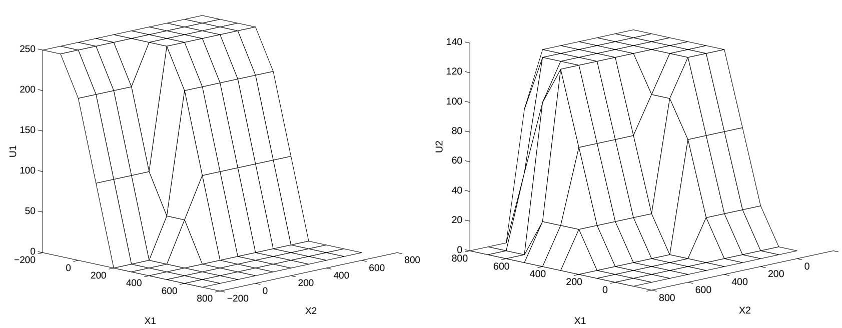
Here again we see a difference between the optimal policy of our discrete-time approximation and a typical "bang-bang" optimal policy of the continuous time model. It can be explained as follows. Suppose that for the continuous time model the optimal "bang-bang" policy is to produce at minimum rate from $t = 0$ to $t = t^*$ and then produce at maximum rate (Figure 7.5 top). Suppose that we discretize the time scale the same way as in chapter 4 with $t_{k-1} < t^* < t_k$. This optimal policy will translate on the discrete time scale as follows: produce at minimum for the periods 1 to $k-1$, produce at maximum for the periods $k+1$ to K and produce between minimum and maximum for the period k (Figure 7.5 bottom). This is clearly not a "bang-bang" policy.

An interesting result is displayed in Figure 7.6 which gives a cross-section of the surface shown in Figure 7.3 for $\tilde{y}_1(0) = 0$. We see that the priority is



(a) $U1(\tilde{y}_0, \tilde{\Xi})$. (b) $U2(\tilde{y}_0, \tilde{\Xi})$.

Figure 7.3: Optimal policy for $\Xi = (1, 1)$.



(a) $U1(\tilde{y}_0, \tilde{\Xi})$. (b) $U2(\tilde{y}_0, \tilde{\Xi})$.

Figure 7.4: Optimal policy for $\Xi = (1, 0)$.

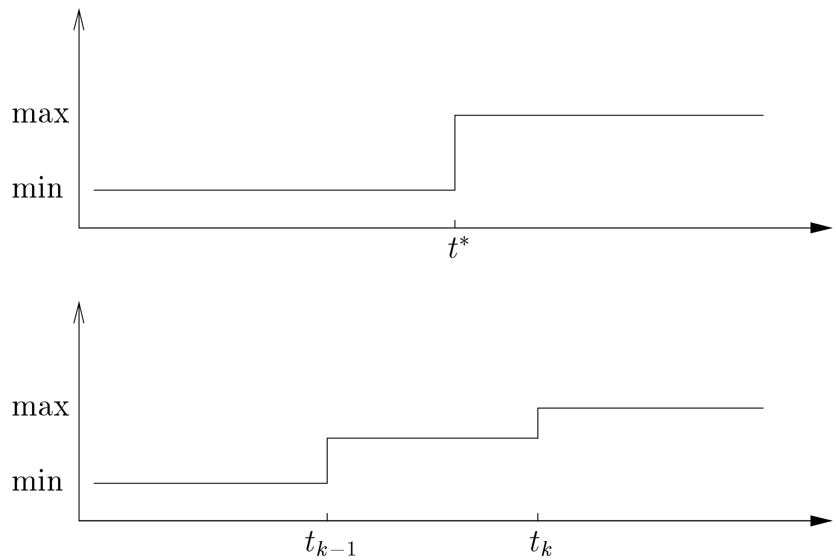


Figure 7.5: Effects of a time discretization on a "bang-bang" policy

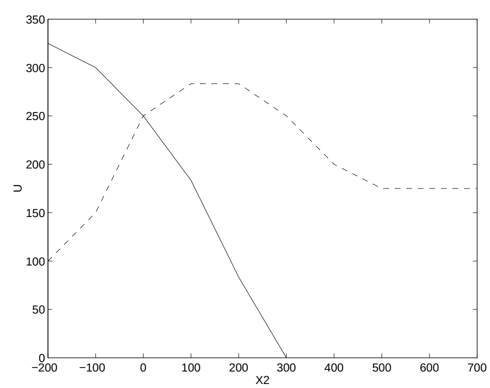


Figure 7.6: The functions $U1(\tilde{y}_2(0), \tilde{\Xi})$ [dotted line] and $U2(\tilde{y}_2(0), \tilde{\Xi})$ [solid line] for $\tilde{y}_1(0) = 0$. In the initial period the two machines are up

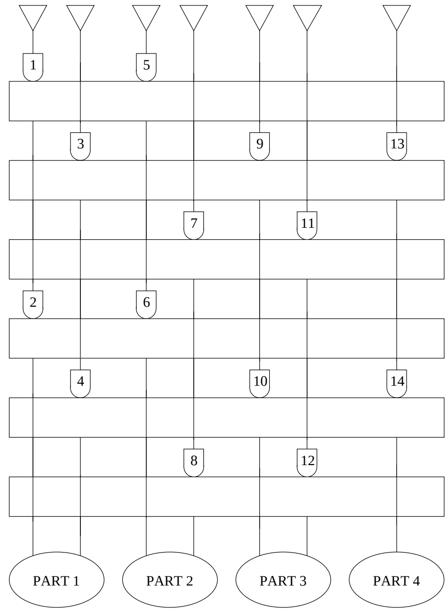


Figure 7.7:

given to the part with the highest backlog. We see also that a high surplus of part 2 (above 300) hedges also for part 1. However this cross-hedging reaches a saturation point: a surplus of part 2 higher than 500 has the same effect as a surplus of 500.

7.3 Six-machine four-part-type example

The larger example considered here is a flexible workshop composed of six machines, among which 3 are unreliable, and producing four parts. The workshop topology is pictured in Figure 7.7. The processing time vector is given by

$$\begin{aligned}\tau \;=\;& (0.005, 0.005, 0.01, 0.015, 0.006, 0.006, 0.006,\\ & 0.006, 0.01, 0.01, 0.005, 0.005, 0.006, 0.006).\end{aligned}$$

For machines 1 and 3, the failure rate is equal to 0.1 and the repair rate is equal to 0.4. The failure rate for Machine 2 is equal to 0.2 and the repair rate is equal to 0.7. The other machines are reliable. The penalty for work-in-process equals 1 in each internal buffer; the penalty for finished part inventory (resp. backlog) equals 5 (resp. 50) for each part type. We considered a time horizon $T = 5$. The demand is supposed to be constant at 100 units for each part type.

We solved the model with $K = 5$ periods and a sample of 10000 scenarios. Given the size of the state space, it is impossible to describe the optimal policy with a simple picture. However we give in Figure 7.8 the hedging stocks when the

Internal buffer	Hedging stock	Internal buffer	Hedging stock
buffer 1	0	buffer 2	93
buffer 3	0	buffer 4	0
buffer 5	0	buffer 6	67
buffer 7	0	buffer 8	67
buffer 9	0	buffer 10	0
buffer 11	0	buffer 12	40
buffer 13	0	buffer 14	233

Finished part buffer	Hedging stock
Part 1	100
Part 2	100
Part 3	100
Part 4	0

Figure 7.8:

six machines are operational. Since upstream from each route there is a fictive infinite buffer, we obtain, as expected, a zero hedging stock for the first buffer on each route. Although we do not show the complete optimal strategy for this model, we must emphasize that it is possible to do so.

Chapter 8

Concluding remarks for Part I

We have shown in this first part of the thesis that a stochastic programming approach could be used to approximate the solution of the associated stochastic control problem in relatively large scale MFCPs. As this approach combines simulation and optimization, it can be considered as another possible method for gaining some insight on the shape of the optimal value functions that will ultimately define the optimal control. In fact, the strength of the proposed numerical method is that it is simulation based although no assumption on the nature of the optimal policy are made. Consequently the numerical approximation of the optimal strategy gives insight on the true nature of the optimal strategy. The stochastic programming approach exploits the fact that the disturbance Markov jump process is uncontrolled. It also allows the use of advanced mathematical programming techniques like decomposition and parallel processing.

Appendix A

Convergence of the stochastic programming approach

The scope of this appendix is to prove, using viscosity techniques, the convergence of the discretization scheme proposed in Chapter 4. More precisely, the theory developed throughout this appendix is needed to prove theorems 1 and 2. Let us first recall that the approximation scheme is based on a time discretization which reformulates the original Piecewise Deterministic Control System (PDCS) as a stochastic program. Therefore, as the value function of the PDCS is a vector valued function, it is necessary to extend the classical viscosity solution first introduced by Crandall, Ishii and Lions (see *e.g.* [25]). Such an extension was already done in [20] and [85] in a slightly different context. The scheme of proof used in this appendix is an extension of some techniques first used for deterministic control system by Capuzzo-Dolcetta in [19].

The theory developed in this appendix encompasses a wider class of problems than the manufacturing flow control problems presented in the previous chapters. Thus, for the sake of generality, we were obliged to change slightly some notations used in the previous chapters. However, this is not too unpleasant as the appendix is self-contained and can therefore be read independently of the rest of the dissertation.

This appendix is organized as follows. In Section A.1 we formulate the PDCS and the admissible strategies and give the hypotheses needed throughout the appendix. We recall the dynamic programming principle stated in [102] and prove regularity properties satisfied by the vector value function of the PDCS. In Section A.2 we define the vector viscosity solution and show that it is consistent with the classical solution. We prove, using fixed point arguments, that the vector value function of the PDCS is the unique Lipschitz continuous vector viscosity solution of a system of coupled Hamilton-Jacobi-Bellman's (HJB) type equations. In Section A.3 we recall briefly the approximation scheme studied in Chapter 4 and its interpretation in terms of stochastic program. We prove that the sequence of vector value functions associated to the approximating stochastic programs converges to the vector viscosity solution of the HJB system, and consequently to the vector value function of the initial PDCS, as the discretization's step tends to zero.

A.1 The piecewise deterministic control system

A.1.1 Dynamics of the PDCS

We consider a piecewise deterministic control system, with hybrid state $(y(t), \xi(t))$, where $y(t) \in I\!R^m$ denotes the continuous part of the state while $\xi(t)$ denotes the discrete part. The discrete state, $\xi(t)$, belongs to a finite set $\mathcal{I} = \{1, 2, \ldots, I\}$, and evolves according to a continuous time Markov jump process with transition rates defined by

$$\begin{aligned} P[\xi(t+dt) = j | \xi(t) = i] &= q_{ij} dt + o(dt) \quad i, j \in \mathcal{I}, \quad i \neq j \\ P[\xi(t+dt) = i | \xi(t) = i] &= 1 + q_{ii} dt + o(dt)\,, \end{aligned} \tag{A.1}$$

with

$$q_{ii} = -\sum_{i \neq j} q_{ij} < 0 \tag{A.2}$$

and

$$\lim_{dt \to 0} \frac{o(dt)}{dt} = 0 \tag{A.3}$$

The continuous state evolves according to a differential equation that depends on the value taken by the discrete state. More precisely, if the discrete state at time t is $\xi(t) = i$, then the continuous state evolves according to the following state equation from time t on until the next jump of the discrete state occurs

$$\dot{y}(t) = f^i(y(t), u(t)), \quad u(t) \in U^i, \tag{A.4}$$

where the control $u(t)$ takes value in a set U^i.

Assumption 1. *U^i is a closed compact set for all i in $\mathcal{I}$.*

Assumption 2. *The functions f^i, $i \in \mathcal{I}$ are Lipschitz continuous in x, continuous in (x, u) and bounded, i.e.*

- $||f^i(x,u) - f^i(y,u)|| \leq C_f ||x - y||, \quad \forall i \in \mathcal{I}, x, y \in I\!R^m, u \in U^i,$
- *$f^i(x, u)$ is continuous in (x, u),*
- $||f^i(x,u)|| \leq M_f, \quad \forall (x,u) \in I\!R^m \times U^i.$

Assumptions 1 and 2 insure the existence and uniqueness of the solution of (A.4) for each possible initial point (x, i) at time t.

Denote $X_0 \subset I\!R^m$ the compact set of possible initial continuous state. Let $\mathcal{X} \subset I\!R^m \times [0, T]$ be the set of points (y, t) such that there exists a trajectory starting at time $t = 0$ from a point $x \in X_0$ and reaching the continuous state y at time t. It is clear from Assumptions 1 and 2 that the reachable set $\mathcal{X}$ is a closed bounded subset of $I\!R^m \times [0, T]$.

Define λ_0 as follows:

$$\lambda_0 = \sup_{\substack{u \in \mathcal{U}\, i \in \mathcal{I} \\ x \neq y}} \frac{\langle f^i(x, u) - f^i(y, u), x - y \rangle}{|x - y|^2}. \tag{A.5}$$

Assumption 3. $\lambda_0 \leq C_\lambda, \quad C_\lambda > 0.$

Lemma 1. *Let $(x(s), \xi(t))$ and $(y(s), \xi(t))$ be two trajectories starting respectively at time $t = 0$ at (x_0, ξ_0) and (y_0, ξ_0) associated with the same control $u(\cdot)$ and the same realization of the jump process. Then we have for all $s > 0$*

$$|x(s) - y(s)| \leq |x_0 - y_0| e^{\lambda_0 s}.$$

Proof. Since the evolution of the discrete state does not depend on the value of the continuous state, the lemma follows directly from Gronwald theorem (see *e.g.* [4]). □

A.1.2 Admissible controls and piecewise open loop strategies

For a given initial hybrid state of the system (x, i) at time τ, an admissible open loop control, will be a measurable mapping

$$u(\cdot) : [\tau, T[\rightarrow U^i$$

such that the solution of (A.4), with initial condition $y(\tau) = x$ exists and is unique. Let us denote $\mathcal{U}^i$ the set of such mappings.

In this context a strategy of the controller is described in the following way. At each jump time of the system, *i.e.* at each instant τ when a jump of the discrete state occurs, the controller observes the new hybrid state $(y(\tau), \xi(\tau)) = (x, i)$ and chooses an admissible open loop control in the set $\mathcal{U}^i$. This open loop control will be applied, until time τ', when either a new jump of the system occurs, or the final time is reached, *i.e.* $\tau' = T$. During the interval of time $[\tau, \tau'[$, the associated trajectory $y(\cdot)$ of the continuous state is solution of

$$\dot{y}(t) = f^i(y(t), u(t)), \text{ with } y(\tau) = x.$$

To sum up, a strategy γ is thus a mapping from $[\tau, T[\times I\!R^m \times \mathcal{I}$ to $\cup_{i\in\mathcal{I}}\mathcal{U}^i$.

Notice that since we have a deterministic control problem between two successive jumps, open loop and closed loop strategies are equivalent between these two jumps. So in this case piecewise open loop and piecewise closed loop strategies are equivalent.

A complete and precise description of admissible strategies involves the use of a concept of solution of an ordinary differential equation with discontinuous right-hand side, since the control $u(\cdot)$ and consequently the functions $f^i(y(t), u(t))$ can be discontinuous. This can be found in [102] in a more general setup.

A.1.3 Value function and optimality equations

Suppose that, at time t, the state of the system is $(x, i) \in X_t \times \mathcal{I}$, where $X_t = \{x | (x, t) \in \mathcal{X}\}$, and that the controller uses a given strategy γ. The trajectory $y(\cdot)$ of the continuous state together with the trajectory of the control $u(\cdot)$ is a stochastic processes with measure P induced by γ. We define the evaluation function associated to the strategy γ and initial state (x, i) as

$$J^i_\gamma(x, t) = E_P[\int_t^T L^{\xi(s)}(y(s), u(s))\, ds \mid y(t) = x, \xi(t) = i], \tag{A.6}$$

Assumption 4. *The instantaneous cost is Lipschitz continuous in x, continuous in (x, u) and bounded, i.e*

- $|L^i(x, u) - L^i(y, u)| \leq C_L||x - y||, \quad \forall x, y \in \mathcal{X}, u \in U^i$,
- *$L^i(x, u)$ is continuous in (x, u),*
- $|L^i(x, u)| \leq M_L\, \forall x \in I\!R^m, u \in U^i$.

Throughout the chapter we assume that Assumptions 1 to 4 are satisfied.

We are interested to find the optimal vector value function $(\mathcal{V}^i(x, t))_{i=1,\ldots,I}$ where

$$\mathcal{V}^i(x, t) = \inf_\gamma J^i_\gamma(x, t), \quad i = 1, \ldots, I. \tag{A.7}$$

Now we state the dynamic programming principle verified by $\mathcal{V}^i$. This results can be found, for example, in [33], in a slightly different formulation, or in a more general set up in [102].

Proposition 1. *For any (x, t) in $\mathcal{X}$ and positive ϵ we have*

$$\begin{aligned} \mathcal{V}^i(x, t) = \inf_{u(\cdot)\in\mathcal{U}^i} \Bigg\{ & \mathcal{V}^i(y(t + \epsilon), t + \epsilon) \\ & + \int_t^{t+\epsilon} L^i(y(s), u(s)) + \sum_{j\in\mathcal{I}} q_{ij}\mathcal{V}^j(y(s), s) ds \Bigg\} \end{aligned} \tag{A.8}$$

where $y(\cdot)$ is given by (A.4) and $i = 1, \dots, I$.

Theorem 4. *If Assumptions 2 to 4 hold then the functions $\mathcal{V}^i(x,t)$ are bounded and Lipschitz continuous in (x,t).*

Proof. The boundedness is straightforward from the definition (A.6) of the cost function together with Assumption 4. Let us compute the difference $\mathcal{V}^i(\bar{x},\bar{t}) - \mathcal{V}^i(x,t)$. We have

$$\begin{aligned}&|\mathcal{V}^i(\bar{x},\bar{t}) - \mathcal{V}^i(x,t)| \\ &\leq |\mathcal{V}^i(\bar{x},\bar{t}) - \mathcal{V}^i(\bar{x},t)| + |\mathcal{V}^i(\bar{x},t) - \mathcal{V}^i(x,t)|\end{aligned} \tag{A.9}$$

On the one hand, by definition (A.6) of the cost function, we have

$$\begin{aligned}&|\mathcal{V}^i(\bar{x},t) - \mathcal{V}^i(x,t)| \\ &\quad = \Bigg|\inf_{\gamma} E_P\left[\int_t^T L^{\xi(s)}(\bar{y}(s),u(s))ds|\xi(t)=i,\bar{y}(t)=\bar{x}\right] \\ &\qquad -\inf_{\gamma} E_P\left[\int_t^T L^{\xi(s)}(y(s),u(s))ds|\xi(t)=i,y(t)=x\right]\Bigg| \\ &\quad = \sup_{\gamma} E_P\left[\int_t^T |L^{\xi(s)}(\bar{y}(s),u(s))ds - L^{\xi(s)}(y(s),u(s))|ds\ |\xi(t)=i,\bar{y}(t)=\bar{x},y(t)=x\right].\end{aligned}$$

The Lipschitz property of L and lemma 1 imply

$$|\mathcal{V}^i(\bar{x},t) - \mathcal{V}^i(x,t)| \leq ||\bar{x}-x||C_L C_1 \leq ||(\bar{x},\bar{t})-(x,t)||C_L C_1 \tag{A.10}$$

where $C_1 = \frac{e^{C_\lambda T}}{C_\lambda}$ is a constant.

On the other hand,

$$\begin{aligned}|\mathcal{V}^i(\bar{x},\bar{t}) - \mathcal{V}^i(\bar{x},t)| \quad = \quad &\Bigg|\inf_{\gamma} E_P\left[\int_{\bar{t}}^T L^{\bar{\xi}(s)}(\bar{y}(s),u(s))ds \mid \bar{\xi}(\bar{t})=i,\bar{y}(\bar{t})=\bar{x}\right] \\ &- \inf_{\gamma} E_P\left[\int_t^T L^{\xi(s)}(y(s),u(s))ds|\xi(t)=i,y(t)=\bar{x}\right]\Bigg|.\end{aligned}$$

Without loss of generality, we may suppose that $\bar{t} > t$. Using the fact that the instantaneous cost L is bounded it comes that

$$\begin{aligned}&|\mathcal{V}^i(\bar{x},\bar{t}) - \mathcal{V}^i(\bar{x},t)| \\ &\quad \leq \Bigg|\inf_{\gamma} E_P\left[\int_{\bar{t}}^{T-(\bar{t}-t)} L^{\bar{\xi}(s)}(\bar{y}(s),u(s))ds \mid \bar{\xi}(\bar{t})=i,\bar{y}(\bar{t})=\bar{x}\right] + \int_{T-(\bar{t}-t)}^T M_L ds \\ &\qquad -\inf_{\gamma} E_P\left[\int_t^T L^{\xi(s)}(y(s),u(s))ds|\xi(t)=i,y(t)=\bar{x}\right]\Bigg|.\end{aligned}$$

The first and the last terms in the previous expression involve stochastic replicas and are therefore equal. We thus obtain

$$|\mathcal{V}^i(\bar{x},\bar{t}) - \mathcal{V}^i(\bar{x},t)| \leq M_L|\bar{t}-t| \leq M_L||(\bar{x},\bar{t})-(x,t)||. \tag{A.11}$$

Taking together inequalities (A.9), (A.10) and (A.11) leads to the following result

$$|\mathcal{V}^i(\bar{x},\bar{t}) - \mathcal{V}^i(x,t)| \leq (M_L + C_L C_1)||(\bar{x},\bar{t})-(x,t)||, \tag{A.12}$$

which concludes the proof. □

A.2 Vectorial viscosity solution

In this section we extend the notion of classical viscosity solution for a system of first order partial differential equations.

A.2.1 Definition and properties

We consider the following system of coupled first order differential equations with boundary conditions

$$\begin{cases} H^1(x,t,V^1(x,t),\ldots,V^I(x,t),\nabla V^1(x,t)) = 0 \\ H^2(x,t,V^1(x,t),\ldots,V^I(x,t),\nabla V^2(x,t)) = 0 \\ \vdots \\ H^I(x,t,V^1(x,t),\ldots,V^I(x,t),\nabla V^I(x,t)) = 0, \quad \forall (x,t) \in \mathcal{X} \\ \\ V^1(x,T) = V^2(x,T) \cdots = V^I(x,T) = 0 \quad \forall x, \end{cases} \tag{A.13}$$

where H^i are continuous functions from $I\!R^m \times I\!R \times I\!R^I \times I\!R^{m+1}$ to $I\!R$.

Definition 1. *The continuous vector function of $\mathcal{X}$ $V(x,t) = (V^1(x,t),\ldots,V^I(x,t))$ is said to be a continuous vector viscosity solution of the system (A.13) if:*

- *for all $\phi \in C^1(\mathcal{X})$, if (y,s) is a local maximum of $V^i - \phi$ for some $i \in \mathcal{I}$, we have*

$$H^i(y,s,V^1(y,s),\ldots,V^I(y,s),\nabla\phi(y,s)) \leq 0, \tag{A.14}$$

and

- *for all $\phi \in C^1(\mathcal{X})$, if (y,s) is a local minimum of $V^i - \phi$ for some $i \in \mathcal{I}$, we have*

$$H^i(y,s,V^1(y,s),\ldots,V^I(y,s),\nabla\phi(y,s)) \geq 0. \tag{A.15}$$

If $V(x,t)$ verifies only (A.14) (respectively (A.15)), we call it a vector viscosity sub-solution (resp. super-solution).

A similar definition was proposed in [20] and in [85]. This definition is a straightforward extension of the continuous viscosity solution for first order partial differential equation first developed by Crandall, Ishii and Lions (see *e.g.* [25]).

The following theorem links classical solution and viscosity solution of (A.13), in the case where (A.13) admits a classical solution (*i.e.* a continuously differentiable solution).

Theorem 5. *Suppose $V^i(x,t)$ is C^1 in $\mathcal{X}$, for all i in $\mathcal{I}$, then $V(x,t)$ is a classical solution of (A.13) if and only if it is a vector viscosity solution of (A.13).*

Proof. Let V be a viscosity solution of (A.13). We take $\phi = V^i$ in the definition 1. Since each point of $\mathcal{X}$ is a maximum and a minimum of $V^i - \phi$ we obtain the two inequalities $H^i(x,t,V^1(x,t),\ldots,V^I(x,t),\nabla V^i(x,t)) \leq 0$ and $H^i(x,t,V^1(x,t),\ldots,V^I(x,t),\nabla V^i(x,t)) \geq 0$ over $\mathcal{X}$, and conclude that V is a classical solution.
Conversely, if V is a classical solution of (A.13) and if y is a local maximum (*resp.* minimum) of $V^i - \phi$, we have $\nabla V^i(y) = \nabla\phi(y)$ and so $H(y,V(y),\nabla\phi(y)) = H(y,V(y),\nabla V^i(y)) = 0$, from which we conclude that V is a viscosity subsolution (*resp.* super-solution). □

A.2.2 The viscosity solution of the PDCS

The dynamic programming principle applied to PDCS defines a system of coupled HJB equations.

Theorem 6. *The vector value function of the PDCS, $\mathcal{V} = (\mathcal{V}^1,\ldots,\mathcal{V}^I)$, defined by (A.7) with (A.6) is a viscosity solution of the equation (A.13) where the Hamiltonians are defined by*

$$H^i(x,t,V^1(x,t),\ldots,V^I(x,t),\nabla V^i(x,t)) =$$
$$\min_{u\in U^i}\{-L^i(x,u) - \sum_{j\in\mathcal{I}} q_{ij}V^j(x,t) - \frac{\partial}{\partial t}V^i(x,t) - \nabla V^i(x,t)f^i(x,u)\} \quad i=1,\ldots,I. \tag{A.16}$$

Proof. We first prove that the value function $\mathcal{V}$ is sub-solution. Consider φ a $C^1(\mathcal{X})$ function. Suppose that a local minimum of $\mathcal{V}^i - \varphi$ is attained on $(\tilde{x},\tilde{t})$. Without any loss of generality we can assume that $\mathcal{V}^i(\tilde{x},\tilde{t}) = \varphi(\tilde{x},\tilde{t})$. Consequently, there exists a real positive number, r, such that for all (y,s) such that $||(\tilde{x},\tilde{t}) - (y,z)|| \leq r$ we have

$$\mathcal{V}^i(y,s) \leq \varphi(y,s),$$

and then for ϵ sufficiently small, using proposition 1 we can write,

$$\begin{aligned}
&\varphi(\tilde{x},\tilde{t}) - \varphi(\tilde{y}(\tilde{t}+\epsilon),\tilde{t}+\epsilon) \leq \mathcal{V}^i(\tilde{x},\tilde{t}) - \mathcal{V}^i(\tilde{y}(\tilde{t}+\epsilon),\tilde{t}+\epsilon) \\
&\leq \int_{\tilde{t}}^{\tilde{t}+\epsilon} L^i(\tilde{y}(s),u)\, ds + \sum_j \int_{\tilde{t}}^{\tilde{t}+\epsilon} q_{ij}\mathcal{V}^j(\tilde{y}(s),s)\, ds,
\end{aligned} \tag{A.17}$$

where $\tilde{y}(\cdot)$ is the continuous state trajectory corresponding to initial conditions $(\tilde{x},i)$ at time $\tilde{t}$, when the constant control $u(\cdot) = u \in U^i$ is applied. Dividing equation (A.17) by ϵ and letting ϵ tend to zero we obtain:

$$-\frac{\partial}{\partial t}\varphi(\tilde{x},\tilde{t}) - \nabla\varphi(\tilde{x},\tilde{t})f^i(\tilde{x},u) - L^i(\tilde{x},u) - \sum_{j\neq i} q_{ij}[\mathcal{V}^j(\tilde{x},\tilde{t}) - \mathcal{V}^i(\tilde{x},\tilde{t})] \leq 0.$$

This inequality is true for all u, and in particular for the control u that minimizes the left hand side of the last inequality. We thus can conclude that $\mathcal{V} = (\mathcal{V}^1,\ldots,\mathcal{V}^I)$ is a sub-solution of (A.13) with (A.16), in the sense of the definition 1.

We now prove that $\mathcal{V}$ is a super-solution. Let $\varphi \in C^1(\mathcal{X})$ and $(\tilde{x},\tilde{t})$ be a local maximum of $\mathcal{V}^i - \varphi$. Again we assume that $\mathcal{V}^i(\tilde{x},\tilde{t}) = \varphi(\tilde{x},\tilde{t})$. So, there exists $r > 0$ such that, for any (y,s) satisfying $||(y,s) - (\tilde{x},\tilde{t})|| \leq r$, we have :

$$\mathcal{V}^i(y,s) \geq \varphi(y,s).$$

From (A.8) we obtain the following equality:

$$\begin{aligned}
\mathcal{V}^i(\tilde{x},\tilde{t}) \quad = \quad & \min_{u(\cdot)\in\mathcal{U}^i}\{\mathcal{V}^i(\tilde{y}(\tilde{t}+\epsilon),\tilde{t}+\epsilon) + \\
& \int_{\tilde{t}}^{\tilde{t}+\epsilon} L^i(\tilde{y}(s),u(s)) + \sum_{j\in\mathcal{I}} q_{ij}\mathcal{V}^j(\tilde{y}(s),s)ds\}
\end{aligned}$$

and consequently

$$\begin{aligned}
\varphi(\tilde{x},\tilde{t}) \quad \geq \quad & \min_{u(\cdot)\in\mathcal{U}^i}\{\varphi(\tilde{y}(\tilde{t}+\epsilon),\tilde{t}+\epsilon) + \\
& \int_{\tilde{t}}^{\tilde{t}+\epsilon} L^i(\tilde{y}(s),u(s)) + \sum_{j\in\mathcal{I}} q_{ij}\mathcal{V}^j(\tilde{y}(s),s)\, ds\}.
\end{aligned}$$

Using the equality

$$\begin{aligned}
&\varphi(\tilde{y}(\tilde{t}+\epsilon),\tilde{t}+\epsilon) = \varphi(\tilde{x},\tilde{t}) \\
&+ \int_{\tilde{t}}^{\tilde{t}+\epsilon} \frac{\partial}{\partial t}\varphi(\tilde{y}(s),s) + \nabla\varphi(\tilde{y}(s),s)f^i(\tilde{y}(s),u(s))ds,
\end{aligned}$$

we rewrite the last inequality

$$\begin{aligned} 0 \;\geq\; & \min_{u(\cdot)\in\mathcal{U}^i}\{\int_{\tilde{t}}^{\tilde{t}+\epsilon} L^i(\tilde{y}(s),u(s)) + \sum_{j\in\mathcal{I}} q_{ij}\mathcal{V}^j(\tilde{y}(s),s) \\ & +\frac{\partial}{\partial t}\varphi(\tilde{y}(s),s) + \nabla\varphi(\tilde{y}(s),s)f^i(\tilde{y}(s),u(s))ds\} \end{aligned}$$

which implies

$$\begin{aligned} 0 \;\geq\; & \min_{u(\cdot)\in\mathcal{U}^i}\{\int_{\tilde{t}}^{\tilde{t}+\epsilon} \min_{w\in\mathcal{U}^i}[L^i(\tilde{y}(s),w) + \sum_{j} q_{ij}\mathcal{V}^j(\tilde{y}(s),s) \\ & +\frac{\partial}{\partial t}\varphi(\tilde{y}(s),s) + \nabla\varphi(\tilde{y}(s),s)f^i(y(s),w)]ds\} \end{aligned}$$

or again

$$0 \geq \min_{u(\cdot)\in\mathcal{U}^i}\{\int_{\tilde{t}}^{\tilde{t}+\epsilon} -H^i(\tilde{y}(s),s,\mathcal{V}(\tilde{y}(s),s),\nabla\varphi(\tilde{y}(s),s))\,ds\}.$$

Dividing by $-\epsilon$ and letting ϵ tend to zero we obtain

$$0 \leq H^i(\tilde{x},\tilde{t},\mathcal{V}(\tilde{x},\tilde{t}),\nabla\varphi(\tilde{x},\tilde{t})), \tag{A.18}$$

which concludes the proof of the fact that $\mathcal{V}$ is a super-solution and finishes the proof of the theorem. □

Let us denote $C^{0,1}(\mathcal{X})$ the set of the Lipschitz continuous functions defined on $\mathcal{X}$.

Theorem 7. *In $\Pi_{i\in\mathcal{I}}C^{0,1}(\mathcal{X})$ there exists a unique viscosity solution of (A.13,A.16).*

To prove theorem 7 we need a result that links "classical" viscosity solution with the solution of an optimal control problem, and a result that gives the uniqueness of "classical" viscosity solution. These results can be found for example in [10] (theorem 3.6) that we recall below for the sake of completeness :

Theorem 8. *Consider the finite horizon optimal control problem given by its dynamics*

$$\dot{y} = f(y(t),u(t)), \quad y(0) = x$$

and the cost function

$$J(x,u(\cdot)) = \int_{t=0}^{T} L(y(s),u(s))ds.$$

Suppose that the dynamics and instantaneous cost functions f and L are Lipschitz continuous with respect to their first argument, bounded and continuous in (y, u) and that V_0 is bounded uniformly continuous. Then the value function of the optimal control problem is the unique bounded uniformly continuous viscosity solution of the Bellman equation

$$\frac{\partial V}{\partial t} + H(x, V, \nabla V) = 0$$

associated to the boundaries conditions

$$V(x, 0) = V_0(x)$$

with

$$H(x, V, p) = \sup_{u \in U} \left\{ -pf(x, u) - L(x, u) \right\}.$$

Proof of Theorem 7. Existence follows directly from theorem 6, since we have exhibited such a solution.

In order to prove uniqueness, we introduce an operator $\mathcal{T}$ from the set $\Pi_{i \in \mathcal{I}} C^{0,1}(\mathcal{X})$ to itself, and will show that it is contractive and consequently admits a unique fixed point.

Define the operator $\mathcal{T}$ in the following way :

$$\begin{array}{rl} \mathcal{T} : & \Pi_{i \in \mathcal{I}} C^{0,1}(\mathcal{X}) \to \Pi_{i \in \mathcal{I}} C^{0,1}(\mathcal{X}) \\ & (\mathcal{W}^1, \ldots, \mathcal{W}^I) \longrightarrow (\mathcal{V}^1, \ldots, \mathcal{V}^I), \end{array}$$

where $\mathcal{V}^i$ is the standard viscosity solution of the following equation

$$\begin{aligned} 0 \quad = \quad & \min_{u \in U^i} \{ -\tilde{L}^i(x, u) \\ & -\frac{\partial}{\partial t}\mathcal{V}^i(x, t) - \nabla \mathcal{V}^i(x, t) f^i(x, u) - q_{ii} \mathcal{V}^i(x, t) \}, \end{aligned} \tag{A.19}$$

where

$$\tilde{L}^i(x, u) = L^i(x, u) + \sum_{j \neq i} q_{ij} \mathcal{W}^j(x, t) \tag{A.20}$$

with terminal conditions $\mathcal{V}^i(x, T) = 0$ for all x.

Notice that if for $i \neq j$, the functions $\mathcal{W}^j$ are Lipschitz continuous, then the instantaneous cost functions $\tilde{L}^i$ satisfy Assumption 4.
Notice also that, according to theorem 8, the viscosity solution of this equation can be interpreted as the solution of a problem of control with finite horizon T,

where the dynamics is given by the function f^i, the instantaneous cost is given by the function $\tilde{L}$ and the discount rate is given by q_{ii}.

Let us now prove that $\mathcal{T}$ is well defined, *i.e.* that $\mathcal{T}(\mathcal{W}^1,\ldots,\mathcal{W}^I)$ exists, is unique, and belongs to $\Pi_{i\in\mathcal{I}}C^{0,1}(\mathcal{X})$ for $(\mathcal{W}^1,\ldots,\mathcal{W}^I)\in\Pi_{i\in\mathcal{I}}C^{0,1}(\mathcal{X})$. The existence follows directly from the previous remark. $\mathcal{T}(\mathcal{W}^1,\ldots,\mathcal{W}^I)$ can be interpreted as a vector of classical viscosity solutions for decoupled optimal control problems. According to Theorem 8 we have existence and uniqueness. The fact that each component of $\mathcal{T}(\mathcal{W}^1,\ldots,\mathcal{W}^I)$ belongs to $C^{0,1}$, follows directly from the fact that each component can be interpreted as the solution of a finite horizon optimal control problem without final cost and Lipschitz continuous instantaneous cost function.
Now, from the definition of a vector viscosity solution and "classical" viscosity solution, it is straightforward to see that a fixed point of the operator $\mathcal{T}$ is a vector viscosity solution of equation (A.13-A.16), and conversely, any viscosity solution of (A.13-A.16) is a fixed point of the operator $\mathcal{T}$.

Again we use theorem 8 to interpret each component of $\mathcal{T}(\mathcal{W}^1,\ldots,\mathcal{W}^I)$ as the value function of an optimal control problem. Now to prove that $\mathcal{T}$ is contractive let us compute a upper bound for $||\mathcal{T}\mathcal{W}-\mathcal{T}\tilde{\mathcal{W}}||$, where the norm is defined as

$$||\mathcal{W}|| = ||\mathcal{W}^1,\ldots,\mathcal{W}^I|| = \max_{i\in\mathcal{I}}\max_{(x,t)\in\mathcal{X}}|\mathcal{W}^i(x,t)|.$$

We have then

$$\begin{aligned}
&\mathcal{T}\tilde{\mathcal{W}}^i(x,t) - \mathcal{T}\mathcal{W}^i(x,t)\\
&\quad= \inf_{u(\cdot)}\int_t^T e^{q_{ii}(s-t)}L^i(x(s),u(s))\,ds + \sum_{j\neq i} q_{ij}\int_t^T e^{q_{ii}(s-t)}\tilde{\mathcal{W}}^j(x(s),s)\,ds\\
&\quad -\inf_{u(\cdot)}\int_t^T e^{q_{ii}(s-t)}L^i(x(s),u(s))\,ds + \sum_{j\neq i} q_{ij}\int_t^T e^{q_{ii}(s-t)}\mathcal{W}^j(x(s),s)\,ds\\
&\quad\le \sup_{u(\cdot)}\sum_{j\neq i} q_{ij}\int_t^T e^{q_{ii}(s-t)}[\tilde{\mathcal{W}}^j(x(s),s)-\mathcal{W}^j(x(s),s)]\,ds\\
&\quad\le \int_t^T \sum_{j\neq i} q_{ij}e^{q_{ii}(s-t)}\,ds \quad ||\tilde{\mathcal{W}}-\mathcal{W}||\\
&\quad= \int_t^T -q_{ii}e^{q_{ii}(s-t)}\,ds \quad ||\tilde{\mathcal{W}}-\mathcal{W}||\\
&\quad= (1-e^{q_{ii}(T-t)}) \quad ||\tilde{\mathcal{W}}-\mathcal{W}||\\
&\quad= (1-e^{q_{ii}T})||\tilde{\mathcal{W}}-\mathcal{W}|| = \theta\ \ ||\tilde{\mathcal{W}}-\mathcal{W}||,
\end{aligned}$$

with $0 < \theta < 1$ as $q_{ii} < 0$ and $0 < T < \infty$. In the same way we could have proved that

$$\mathcal{T}\tilde{\mathcal{W}}^i(x,t) - \mathcal{T}\mathcal{W}^i(x,t) \leq \theta \ \ ||\tilde{\mathcal{W}} - \mathcal{W}||.$$

The property of contraction follows since the previous inequalities are true for any (x,t) in $\mathcal{X}$.

□

A.3 Approximation of the value function

A.3.1 The discrete time problem

We now turn to the approximation of the viscosity solution of the coupled system of HJB equations (A.13,A.16). We use time discretization. Let us denote $\delta_K = T/K$ the time step of the time interval $[0,T]$. In order to obtain an approximation $\mathcal{V}_K^i$ of the function $\mathcal{V}^i$ associated to the time interval discretization, K, we approximate the dynamics $\dot{x}(t) = f^i(x,u)$ by

$$x(t+\delta_K) = x(t) + \delta_K f^i(x,u),$$

and the time derivative of the value functions

$$\frac{d}{dt}\mathcal{V}^i(x(t),t) = \frac{\partial}{\partial t}\mathcal{V}^i(x(t),t) + \nabla\mathcal{V}^i(x(t),t)f^i(x,u),$$

by

$$\frac{\mathcal{V}_K^i(x+f^i(x,u)\delta_K, t+\delta_K) - \mathcal{V}_K^i(x,t)}{\delta_K}.$$

Plugging these approximations in the system of equations (A.13-A.16), we obtain for any time $t \in \{0, \delta_K, 2\delta_K, (K-1)\delta_K\}$:

$$\begin{aligned}\mathcal{V}_K^i(x,t) \;=\; \min_{u\in U^i}\Bigg\{ & L^i(x+f^i(x,u)\delta_K, u)\delta_K \\ & + \sum_{j\neq i} q_{ij}\delta_K\ \mathcal{V}_K^j(x+f^i(x,u)\delta_K, t+\delta_K) + (1+q_{ii}\delta_K)\mathcal{V}_K^i(x+f^i(x,u)\delta_K, t+\delta_K)\Bigg\}.\end{aligned} \tag{A.21}$$

and terminal conditions

$$\mathcal{V}_K^i(x,T) = 0. \tag{A.22}$$

The approximating HJB system of equations (A.21-A.22) can be interpreted as the dynamic programming equations for a stochastic discrete time problem starting at time $t_0 = 0$, and given by

- the dynamics of the continuous state x

$$x((k+1)\delta_K) = x(k\delta_K) + f^{\xi^K(k\delta_K)}(x(k\delta_K), u(k\delta_K))\delta_K,\\ x(0) = x_0 \in X_0 \tag{A.23}$$

- the discrete time Markov jump process ξ^K,

$$P[\xi^K((k+1)\delta_K) = j|\xi^K(k\delta_K) = i] = q_{ij}\delta_K \quad i \neq j,\\ P[\xi^K((k+1)\delta_K) = i|\xi^K(k\delta_K) = i] = 1 + q_{ii}\delta_K, \tag{A.24}$$

- the evaluation function

$$J_K^i(x,0) = E[\sum_{k=0}^{K-1} L^{\xi^K(k\delta_K)}(x(k\delta_K), u(k\delta_K))\delta_K]. \tag{A.25}$$

Analogously to a strategy in the continuous problem, a strategy of this discrete problem is defined as a mapping from $\{0, \delta_K, 2\delta_K, \ldots, (K-1)\delta_K\} \times I\!R^m \times \mathcal{I}$ to the set $\cup_{i\in\mathcal{I}} U^i$.

A way to solve numerically this problem now is the following. Since the stochastic process does not depend on the continuous state or on the control it is possible to construct an event tree associated to this process. To each arc of this tree (A.24) gives a probability. Equations (A.25-A.23) can thus be interpreted as a stochastic program, where (A.25) is the function to be minimized and (A.23) is the set of constraints. In the previous chapters we used this technique with a linear problem.

A.3.2 Interpolation of the discrete time value function

In order to get convergence results we need to define the value function $\mathcal{V}_K^i$ for any t in the interval $[0,T]$. We define $\mathcal{V}_K^i(x,t)$ for any t in $[0,T]$ by using equation (A.21) and (A.22) together with the following terminal conditions for any t in the last segment $](K-1)\delta_K, T[$

$$\mathcal{V}_K^i(x,t) = \min_{u\in U^i} \left\{ L^i(x + f^i(x,u)(T-t), u)(T-t) \right\}. \tag{A.26}$$

A most natural way to get an interpolation of $\mathcal{V}_K^i$ on any t would have been to use linear interpolation techniques. Nevertheless the proofs of the results given on the last section are much simpler using the interpolation introduced above.

Theorem 9. *The functions $(\mathcal{V}_K^i(x,t))_K$ are bounded by a constant M, and Lipschitz continuous in (x,t) with the same Lipschitz constant for any K.*

Proof. The proof is similar to the proof of theorem 4. The fact that the functions are bounded comes from boundedness of the instantaneous cost, which is independent of K. □

A.3.3 Convergence result

We now turn to the approximation theorem :

Theorem 10. *For any $i \in \mathcal{I}$, $\mathcal{V}_K^i(x,t)$ converges to $\mathcal{V}^i(x,t)$, locally uniformly in $\mathcal{X}$ as K tends to infinity, where $\mathcal{V}$ is the viscosity solution of (A.13,A.16).*

Proof. By Theorem 9 the family $(\mathcal{V}_K^i(x,t))_K$ is equicontinuous. Therefore, by Ascoli Theorem, there exists a subsequence $\{\mathcal{V}_p^i\}_p$ of the sequence $\{\mathcal{V}_K^i\}_K$ that converges locally uniformly to some function V. Let us prove that the vector function $V = (V^1, V^2, \ldots, V^I)$ is a vector viscosity solution of system (A.13-A.16). This will prove the theorem, since $\mathcal{V}$ has been proved to be the unique Lipschitz continuous vector viscosity solution of (A.13-A.16).

Let φ be a $C^1(\mathcal{X})$ function, and $(\tilde{x},\tilde{t})$ a local maximum of $\mathcal{V}^i - \varphi$ for some i in $\mathcal{I}$. Without loss of generality we can suppose that this maximum is strict, Hence there exists a positive integer $r > 0$ such that for any (x,t) in the ball B centered in $(\tilde{x},\tilde{t})$ and of radius r we have

$$(\mathcal{V}^i - \varphi)(\tilde{x},\tilde{t}) > (\mathcal{V}^i - \varphi)(x,t).$$

Define now the point $\{\tilde{x}_p,\tilde{t}_p\}_p$, such that $(\tilde{x}_p,\tilde{t}_p)$ is a maximum of $\mathcal{V}_p^i - \varphi$ on the closed set B. Since $\mathcal{V}_p^i$ converges uniformly to V^i on B, it follows that $(\tilde{x}_p,\tilde{t}_p)$ converges to $(\tilde{x},\tilde{t})$. Since $f^i(x,u)$ is bounded, for p large enough, $(\tilde{x}_p + f^i(\tilde{x}_p,u)\delta_p, \tilde{t}_p + \delta_p)$, belongs to B, and consequently from the definition of $(\tilde{x}_p,\tilde{t}_p)$ we have

$$\begin{aligned}&\mathcal{V}_p^i(\tilde{x}_p,\tilde{t}_p) - \varphi(\tilde{x}_p,\tilde{t}_p) \geq \\ &\quad \mathcal{V}_p^i(\tilde{x}_p + f^i(\tilde{x}_p,u)\delta_p, \tilde{t}_p + \delta_p) - \varphi(\tilde{x}_p + f^i(\tilde{x}_p,u)\delta_p, \tilde{t}_p + \delta_p).\end{aligned} \tag{A.27}$$

Using the expression (A.21), and keeping in mind the equation (A.2), we obtain

$$\begin{aligned}0 \;=\;& \min_{u\in\mathcal{U}^i}\{-\mathcal{V}_p^i(\tilde{x}_p,\tilde{t}_p) + L^i(\tilde{x}_p + f^i(\tilde{x}_p,u)\delta_p, u)\delta_p \\ &+ \sum_{j\neq i} q_{ij}\delta_p \mathcal{V}_p^j(\tilde{x}_p + f^i(\tilde{x}_p,u)\delta_p, \tilde{t}_p + \delta_p) + (1 + q_{ii}\delta_p)\mathcal{V}_p^i(\tilde{x}_p + f^i(\tilde{x}_p,u)\delta_p, \tilde{t}_p + \delta_p)\} \\ \leq\;& \min_{u\in\mathcal{U}^i}\{\varphi(\tilde{x}_p + f^i(\tilde{x}_p,u)\delta_p, \tilde{t}_p + \delta_p) - \varphi(\tilde{x}_p,\tilde{t}_p) + L^i(\tilde{x}_p + f^i(\tilde{x}_p,u)\delta_p, u)\delta_p \\ &+ \sum_{j\neq i} q_{ij}\delta_p[\mathcal{V}_p^j(\tilde{x}_p + f^i(\tilde{x}_p,u)\delta_p, \tilde{t}_p + \delta_p) - \mathcal{V}_p^i(\tilde{x}_p + f^i(\tilde{x}_p,u)\delta_p, \tilde{t}_p + \delta_p)].\}\end{aligned} \tag{A.28}$$

Since φ is a $C^1(\mathcal{X})$ function we have for some θ in $[0,1]$,

$$\begin{aligned}&\varphi(\tilde{x}_p + f^i(\tilde{x}_p,u)\delta_p, \tilde{t}_p + \delta_p) - \varphi(\tilde{x}_p,\tilde{t}_p) = \\ &\delta_p \nabla\varphi(\tilde{x}_p + \theta f^i(\tilde{x}_p,u)\delta_p, \tilde{t}_p + \theta\delta_p) f^i(\tilde{x}_p,u) + \delta_p \frac{\partial}{\partial t}\varphi(\tilde{x}_p + \theta f^i(\tilde{x}_p,u)\delta_p, \tilde{t}_p + \theta\delta_p).\end{aligned}$$

Substituting this last expression in (A.28) and dividing by δ_p we obtain

$$\begin{aligned} 0 \;\leq\; & \min_{u \in U^i}\{\nabla\varphi(\tilde{x}_p + \theta f^i(\tilde{x}_p, u)\delta_p, \tilde{t}_p + \theta\delta_p) f^i(\tilde{x}_p, u) + \frac{\partial}{\partial t}\varphi(\tilde{x}_p + \theta f^i(\tilde{x}_p, u)\delta_p, \tilde{t}_p + \theta\delta_p) \\ & + L^i(\tilde{x}_p + f^i(\tilde{x}_p, u)\delta_p, u) \\ & + \sum_{j \neq i} q_{ij}[\mathcal{V}_p^j(\tilde{x}_p + f^i(\tilde{x}_p, u)\delta_p, \tilde{t}_p + \delta_p) - \mathcal{V}_p^i(\tilde{x}_p + f^i(\tilde{x}_p, u)\delta_p, \tilde{t}_p + \delta_p)]\}. \end{aligned}$$

Now we let p tend to infinity to get

$$0 \leq \min_{u \in U^i}\{\nabla\varphi(\tilde{x}, \tilde{t}) \cdot f^i(\tilde{x}_p, u) + \frac{\partial}{\partial t}\varphi(\tilde{x}, \tilde{t}) + L^i(\tilde{x}, u) + \sum_{j \neq i} q_{ij}[\mathcal{V}^j(\tilde{x}, \tilde{t}) - \mathcal{V}^i(\tilde{x}, \tilde{t})]\},$$

which establishes that V is a viscosity subsolution of (A.13,A.16). The same arguments can be used to prove that V is also a viscosity super-solution. This ends the proof of the theorem. □

Part II

Decomposition Method in a Singularly Perturbed Hybrid Stochastic Model

Chapter 9

Introduction to Part II

Hybrid stochastic control models offer a nice paradigm for the modeling of manufacturing and economic production systems (see [37],[47], [63], [62], [64] and [107] for a small sample of the abundant literature in this area). Unfortunately, these models are complex and the optimal control law is difficult to compute. Although an analytical expression of the optimal control law can sometimes be obtained for a few very simple models, in general, numerical methods are the only possibility to compute the optimal control law.

Often, these systems have the property that the stochastic events occur at very different time scales. In the present work we will study a class of hybrid stochastic control problems with two time scales. The "fast" mode of the system is characterized by a continuous stochastic variable which takes the form of a controlled jump and diffusion process. The "slow" mode of the system is described by a discrete stochastic variable which takes the form of a controlled Markov jump process. By applying the numerical techniques developed by Kushner and Dupuis [83], the hybrid stochastic control problem is then approximated by a singularly perturbed controlled Markov chain. Due to the two different time scales, this controlled Markov chain is ill-conditioned and the optimal control law is difficult to compute because of numerical instabilities. However, the two-time-scale structure permits a hierarchical approach between the slow and the fast modes. When the time scale ratio tends to zero, the hierarchical approach leads to an approximation of the initial control problem by a structured control problem, called *limit control problem.*

Taking the limit, when the time scale ratio tends to zero, of the singularly perturbed controlled Markov chain would not lead, in general, to an approximation of the limit control problem since the ergodic structure of that chain will be altered. This phenomenon can be illustrated by the following example due to Schweitzer [106]. Let

$$M_\varepsilon = \begin{pmatrix} 1-\varepsilon & \varepsilon \\ \varepsilon & 1-\varepsilon \end{pmatrix}$$

be the perturbed Markov chain whose stationary distribution matrix is

$$M_\varepsilon^\infty = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}$$

for all $\varepsilon \in]0,1[$. Taking the limit, when ε tends to zero, of the stationary distribution matrix leads to

$$\lim_{\varepsilon \downarrow 0} M_\varepsilon^\infty = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix},$$

which is different from

$$M_0^\infty = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

the stationary distribution matrix of the unperturbed Markov chain M_0.

Fortunately, the theory of singularly perturbed systems permits one to obtain a controlled Markov chain approximating the limit control problem (see [35] and [36]). The linear program associated with this controlled Markov chain has a primal bloc-diagonal structure permitting a decomposition approach and, moreover, the ill-conditioning encountered before has disappeared.

The main contribution of this work is the implementation of a decomposition approach coupling a linear programming method with a policy improvement algorithm. This coupling permits one to exploit optimally both the primal block-diagonal structure and the special structure of each sub-block which can be identified as Markov Decision Problem (MDP). Another valuable contribution is the implementation of a parallel version of the decomposition method.

This presentation is organized as follows. In Chapter 10 we expose a two-time-scale hybrid stochastic control problem. In Chapter 11, following [83], we derive an approximation for the control problem. Then, following [3], we derive an approximation for the limit control problem, when ε tends to zero. In Chapter 12, following [35] and [36], we derive a decomposition approach for the limit-control problem which exploits the bloc-diagonal structure. We then explain how this decomposition can be implemented using the Analytic Center Cutting Plane Method (ACCPM). Finally, in Chapter 13, we apply the decomposition method to an example of a production plant with two types of human resources and four different market states. We then compare the performance of the decomposition approach with the performance of a frontal method. We also show the speed-up resulting from a parallel implementation of the decomposition method.

Chapter 10

Two-time-scale hybrid stochastic systems

In this chapter we expose the hybrid stochastic system that will be used in this presentation. It has two time scales, where a continuous state variable $x \in \mathbb{R}^K$ is "moving fast" according to a jump and diffusion process while a discrete state variable ξ is "moving slowly" according to a continuous time stochastic jump process.

10.1 The fast dynamics

Let $x = (x_1, x_2, \ldots, x_K)$ be a controlled random variable, whose state equation is indexed over a finite set ($i \in E$) which describes the different possible values taken by the discrete state variable ξ. More precisely, we assume that the dynamics of the continuous state variable is represented by a controlled jump and diffusion process

$$
\begin{aligned}
&dx(t) = f^i(x(t), u(t))dt + \sigma dz(t) + dS \\
&x(0) = x^0 \\
&u(t) \in U^i \subset \mathbb{R}^K,
\end{aligned}
$$

where U^i is a compact set, $f^i(x, u)$ satisfies the usual smoothness conditions for optimal control problems (C^1 in x and continuous in u) and $\{z(t) : t \geq 0\}$ is a K-dimensional Wiener process. The jump process $S(\cdot)$ is given by

$$
S(t) = \int_0^t \int_\Gamma r(x(s^-), \rho) N(ds d\rho),
$$

and $N(\cdot)$ is a Poisson measure of intensity $\Lambda dt \times \Pi(d\rho)$, where $\Pi(\cdot)$ has compact support Γ. For simplicity, we consider the case where the volatility σ is a $\mathbb{R}^K \times \mathbb{R}^K$ diagonal matrix with diagonal elements given by $\boldsymbol{\sigma} = (\sigma_1, \cdots, \sigma_K)$.

If the continuous state variable is restricted to a given subset of $\mathbb{R}^K$, one has to impose reflecting boundary conditions (see [83] for a detailed description).

10.2 The slow dynamics

We assume that the discrete state variable is described by a controlled jump process $\xi(\cdot)$ taking values in the set E. The ξ-process transition rates, which depend on the state x and the control u, are defined by

$$\varepsilon q_{ij}(x,u)dt = P[\xi(t+dt) = j|\xi(t) = i,\, x(t) = x,\, u(t) = u] + o(dt) \quad i,j \in E$$

where

$$\lim_{dt\to 0} \frac{o(dt)}{dt} = 0$$

uniformly in x, u. The parameter ε is the time-scale ratio that will, eventually, be considered as very small.

10.3 Admissible policies and performance criterion

One looks for an optimal control, that is an $(x(\cdot), \xi(\cdot))$-adapted process $u(\cdot)$ such that the expected average reward per unit of time

$$J^\varepsilon(u(\cdot)) = \lim\inf_{T\to\infty} \frac{1}{T} E_u(\cdot) \left[\int_0^T \pi(x(t), \xi(t), u(t))\, dt \right] \tag{10.1}$$

is maximized subject to the fast dynamics of the $x(\cdot)$-process and the slow dynamics of the $\xi(\cdot)$-process.

Chapter 11

Numerical approximation scheme

In the first section of this chapter, following [83], we derive an approximation for the control problem for the case where the fast dynamics is a pure diffusion process. In the second section, following [3], we derive an approximation for the limit control problem, when ε tends to zero, for the case where the fast dynamics is a pure diffusion process. The case where the fast dynamic is a jump and diffusion process is discussed in the third section.

11.1 The control problem

In this section we assume that the fast dynamics is a pure diffusion process. The ergodic cost stochastic control problem identified in the previous section is an instance of the class of controlled switching diffusion studied by Arapostatis, Gosh and Markus in [48]. The dynamic programming equations, established in the previous reference as a necessary optimality condition take the form

$$\begin{aligned} J &= \max_{u \geq 0} \{ \pi(x,i,u) + \varepsilon \sum_{j \neq i} q_{ij}(x,u)[V(x,j) - V(x,i)] \\ &\qquad + \frac{\partial}{\partial x} V(x,i) f^i(x,u) + \frac{\sigma^2}{2} \frac{\partial^2}{\partial x^2} V(x,i) \}, \quad i \in E \end{aligned} \tag{11.1}$$

where $V(x,\cdot)$ is C^2 in x for each i in E and represents a potential value function and J is the maximal expected reward growth rate.

This system of Hamilton-Jacobi-Bellman (HJB) equations cannot, in general, be solved analytically. However a numerical approximation technique can be implemented following a scheme proposed by Kushner and Dupuis [83]. The space of the continuous state is discretized with mesh h. That means that the variable x_k belongs to the grid $\mathcal{X}_k = \{x_k^{\min}, x_k^{\min} + h, x_k^{\min} + 2h, \ldots, x_k^{\max}\}$. Denote e_k the unit vector on the x_k axis and $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2 \times \cdots \times \mathcal{X}_K$. We approximate the first partial derivatives by finite differences, taken "in the direction of the flow", as follows:

$$\frac{\partial}{\partial x_k}V(x) \to \begin{cases} \frac{V(x+e_x h)-V(x)}{h} & \text{if } \dot{x}_k \geq 0 \\ \frac{V(x)-V(x-e_x h)}{h} & \text{if } \dot{x}_k < 0. \end{cases} \tag{11.2}$$

The second partial derivatives are approximated by

$$\frac{\partial^2}{\partial x_k^2}V(x) \to \frac{V(x+e_k h)+V(x-e_k h)-2V(x)}{h^2}. \tag{11.3}$$

We define the interpolation interval as

$$\Delta t_h = \frac{h^2}{\tilde{Q}_h},$$

where

$$Q_h(x,i,u) = \varepsilon q^i(x,u)h^2 + \sum_{k=1}^{K}\{\sigma_k^2 + h|f_k^i(x,u)|\}, \quad q^i(x,u) = \sum_{j\neq i} q_{ij}(x,u)$$

and

$$\tilde{Q}_h = \max_{x,i,u} Q_h(x,i,u).$$

We define transitions probabilities to neighboring grid points as follows

$$p_h[(x,i),(x\pm e_k h,,i)|u] = \frac{\frac{\sigma_k^2}{2} + hf^i(x,u)^{\pm}}{\tilde{Q}_h}, \tag{11.4}$$

$$p_h[(x,i),(x,j)|u] = \varepsilon h^2 \frac{q_{ij}(x,u)}{\tilde{Q}_h} \qquad i \neq j, \tag{11.5}$$

$$p_h[(x,i),(x,i)|u] = 1 - \frac{Q(x,i,u)}{\tilde{Q}_h}. \tag{11.6}$$

The other transitions probabilities are equal to zero. For the case where the fast dynamics is a pure diffusion process, the possible transitions are represented in Figure 11.1, for an example with card(E)=3 and $K = 2$.

If we substitute in the HJB-equations (11.1) the finite differences (11.2) and (11.3) to the partial derivatives, after regrouping terms and using the transition probabilities (11.4), (11.5) and (11.6), we can formulate the following associated discrete state MDP dynamic programming equation:

$$\begin{aligned} g_h\Delta t_h + W(x,i) \quad &= \quad \max_{u\geq 0}\{\sum_{x'} p_h[(x,i),(x',i)|u]W(x',i) + \\ &\qquad \sum_{j\neq i} p_h[(x,i),(x,j)|u]W(x,j) + \Delta t_h \pi(x,i,u)\}, \\ &x \in \mathcal{X}, \qquad i \in E. \end{aligned} \tag{11.7}$$

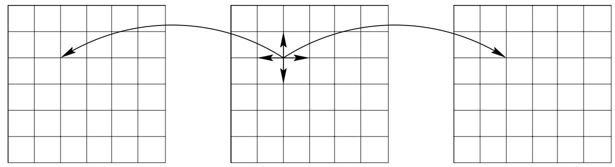


Figure 11.1: Transitions in the grid set.

In this discrete state MDP, the term g approximates the maximal expected reward growth rate J and the functions $W(x, j)$ approximate, in the sense of weak convergence, the potential value functions $V(x, j)$. Solving this MDP gives thus a numerical approximation to the solution of the HJB-equation (11.1).

If we discretize the space of the control with mesh h_u ($u_k \in \mathcal{U}_k = \{u_k^{\min}, u_k^{\min} + h_u, u_k^{\min} + 2h_u, \dots, u_k^{\max}\}$), we obtain an MDP with finite state and action spaces. The optimal control law of this MDP can be obtained through the solution of the following linear program (see [28] and [92]):

$$\max \sum_i \sum_x \sum_u \pi(x, i, u) Z^i(x, u) \tag{11.8}$$

s.t.

$$\sum_i \sum_x \sum_u G_h^\varepsilon[(x, i), (x', j)|u] Z^i(x, u) = 0 \quad x' \in \mathcal{X}, j \in E \tag{11.9}$$

$$\sum_i \sum_x \sum_u Z^i(x, u) = 1 \tag{11.10}$$

$$Z^i(x, u) \geq 0, \tag{11.11}$$

where $G_h^\varepsilon[(x, i), (x', j)|u]$ denotes the generator of the MDP, defined as follows:

$$G_h^\varepsilon[(x, i), (x', j)|u] = \begin{cases} p_h[(x, i), (x, i)|u] - 1 & \text{if } (x, i) = (x', j) \\ p_h[(x, i), (x', j)|u] & \text{otherwise.} \end{cases}$$

Then the steady state probabilities will be defined as

$$P[x, i] = \sum_u Z^i(x, u)$$

and the conditional steady-state probabilities, given a mode i are

$$P[x|i] = \frac{\sum_u Z^i(x, u)}{\sum_x \sum_u Z^i(x, u)}.$$

One should notice that the linear program (11.8-11.11) will tend to be ill-conditioned when ε tends to be small since coefficient with difference of an order of magnitude appear in the same constraints.

11.2 The limit control problem

The generator of the MDP can be written (see Appendix B)

$$G_h^\varepsilon[(x,i),(x',j)|u] = B_h[(x,i),(x',j)|u] + \varepsilon D_h[(x,i),(x',j)|u] + o(\varepsilon),$$

where $B_h[(x,i),(x',j)|u]$ is the generator of a completely decomposable MDP, with card(E) subprocesses which don't communicate one with the other[1] and $\varepsilon D_h[(x,i),(x',j)|u]$ is a perturbation that links together these card(E) sub-blocks.

We are interested in the case where the fast dynamics is much faster than the slow dynamics, in other words we want to study the limit control problem when ε tends to zero. For singularly perturbed systems, the optimal solution of the limit control problem is, in general, different from the optimal solution of the initial problem where ε has been replaced by zero. However, the theory developed by Abbad, Filar and Bielecki in [1] and [3] offers tools to handle the limit of singularly perturbed MDP. Concretely, when ε tends to zero the optimal control law of the MDP (11.7) can be obtained through the solution of the following linear program (see [3]):

$$\max \sum_i \sum_x \sum_u \pi(x,i,u) Z^i(x,u) \tag{11.12}$$

s.t.

$$\sum_x \sum_u B_h[(x,i),(x',i)|u] Z^i(x,u) = 0 \quad x' \in \mathcal{X}, i \in E \tag{11.13}$$

$$\sum_i \sum_{x'} \sum_x \sum_u D_h[(x,i),(x',j)|u] Z^i(x,u) = 0 \quad j \in E \tag{11.14}$$

$$\sum_i \sum_x \sum_u Z^i(x,u) = 1 \tag{11.15}$$

$$Z^i(x,u) \geq 0 \tag{11.16}$$

Indeed this linear program exhibits a typical bloc-diagonal structure in the constraints (11.13). The constraints (11.14-11.15) are the so-called coupling constraints. In Chapter 12 we will apply a decomposition technique to exploit this structure. It should be noticed that the ill-conditioning has vanished since the variable ε doesn't appear in the linear program.

[1] If $i \neq j \quad B_h[(x,i),(x',j)|u] \equiv 0 \quad \forall x, x'$.

11.3 Fast dynamics: jump and diffusion process

An approximation for the model where the fast dynamics follows a jump and diffusion process can be found following the method described in [83] (p. 127-133). This method proceeds in two steps. Firstly we consider an auxiliary system where the fast dynamics is a pure diffusion process. This auxiliary system has the same properties as the initial system except that the jump term is suppressed. As the fast dynamics of this auxiliary system is a pure diffusion process, one can compute the transition probabilities for its associated MDP as in equations (11.4-11.6). Secondly, the jump term is added to these transition probabilities. The transitions probabilities of the associated MDP for the system with jump and diffusion are given by

$$\begin{aligned} P_h[(x,i),(y,i)|u] &= [1-\Lambda\Delta t_h(x,u)]p_h[(x,i),(y,i)|u] \\ &\quad +[\Lambda\Delta t_h(x,u)]\Pi\{\rho : r(x,\rho)=y-x\}. \end{aligned} \tag{11.17}$$

This can be interpreted as follows:

- with probability $1-\Lambda\Delta t_h(x,u)$ the fast dynamics follows a pure diffusion process ;
- with probability $\Lambda\Delta t_h(x,u)$ the fast dynamics makes a jump with intensity $r(x,\rho)$, where ρ has the distribution $\Pi(\cdot)$.

Using the transition probabilities (11.17) in place of transition probabilities (11.4-11.6), we obtain, for the model where the fast dynamics is a jump and diffusion process, the same decomposition principle as in section 11.2.

Chapter 12

A decomposition approach for the limit-control problem

In this chapter, following [35] and [36], we derive for the limit-control problem (11.12-11.16) a decomposition approach which exploits the bloc-diagonal structure. We then explain how this decomposition can be implemented using AC-CPM.

12.1 The decomposition

The dual problem of the limit-control problem (11.12-11.16) writes

$$\begin{aligned}
&\min_{\psi,\phi,\Upsilon} \Upsilon && (12.1)\\
&\text{s.t.}\\
&\Upsilon \geq \pi(x,i,u) - \sum_{x'} B_h[(x,i),(x',i)|u]\phi(x',i)\\
&\qquad - \sum_{j}\sum_{x'} D_h[(x,i),(x',j)|u]\psi(j) && (12.2)\\
&\qquad i \in E,\ x \in \mathcal{X}, u \in \mathcal{U}.
\end{aligned}$$

In this formulation we recognize the second approach proposed in [3], under the name *Aggregation-Disaggregation.* Indeed, if we define the modified costs

$$\Pi(\psi,x,i,u) = \pi(x,i,u) - \sum_{j}\sum_{x'} D_h[(x,i),(x',j)|u]\psi(j) \qquad (12.3)$$

then the expression (12.2) corresponds to a set of card(E) decoupled MDPs. More precisely, the dual problem can be rewritten as

$$\min_{\psi,\phi,\Upsilon} \Upsilon \tag{12.4}$$

$$\text{s.t.}$$

$$\Upsilon \geq \Pi(\psi, x, i, u) - \sum_{x'} B_h[(x,i),(x',i)|u]\phi(x',i)$$

$$i \in E,\ x \in \mathcal{X}, u \in \mathcal{U} \tag{12.5}$$

Now, for each $i \in E$, (12.5) defines a decoupled MDP with modified transition cost (12.3).

12.2 ACCPM

The problem (12.4-12.5) can be solved using ACCPM, which is an interior point cutting plane algorithm for convex optimization problems. It is beyond the scope of this work to give a detailed description of ACCPM. So, readers interested in the details of the theory and the implementation of ACCPM should refer, for example, to Sarkissian's thesis [105].

The solution of the problem (12.4-12.5) is identical to the solution of the convex optimization problem

$$\min_{\psi \in \mathbb{R}^{\mathrm{card}(E)}} \chi(\psi) \tag{12.6}$$

where $\chi(\psi)$ is a convex function defined as the maximum of the value functions of the card(E) different MDPs defined by (12.5), *i.e.*

$$\chi(\psi) = \max_{i \in E} \chi_i(\psi) \tag{12.7}$$

and

$$\chi_i(\psi) = \min_{\phi,\Upsilon} \Upsilon \tag{12.8}$$

$$\text{s.t.}$$

$$\Upsilon \geq \Pi(\psi, x, i, u) - \sum_{x'} B_h[(x,i),(x',i)|u]\phi(x',i) \quad x \in \mathcal{X}, u \in \mathcal{U} \tag{12.9}$$

Since the problem is convex, the epigraph of χ can be approximated by intersections of half-spaces. The procedure called *oracle*, given $\tilde{\psi}$ in $\mathbb{R}^{\mathrm{card}(E)}$, generates a subgradient $X(\tilde{\psi}) \in \partial\chi$ at $\tilde{\psi}$ with the property

$$\chi(\psi) \geq \chi(\tilde{\psi}) + \langle X(\tilde{\psi}), \psi - \tilde{\psi}\rangle \tag{12.10}$$

This inequality defines a supporting hyperplane for the function to be optimized; we call it an optimality cut[1].

Suppose the oracle has been called at a given sequence of points $\{\psi^l\}$, $l \in L$. The oracle has therefore generated a set of optimality cuts defining a piecewise linear approximation $\underline{\chi} : \mathbb{R}^{\mathrm{card}(E)} \to \mathbb{R}$ to the convex function χ

$$\underline{\chi}(\psi) = \max_{l \in L} \{\chi(\psi^l) + \langle X(\psi^l), \psi - \psi^l \rangle\}. \tag{12.11}$$

This permits us to write the following linear program

$$\begin{array}{ll} \min & \zeta \\ \text{s. t.} & \zeta \geq \chi(\psi^l) + \langle X(\psi^l), \psi - \psi^l \rangle, \quad \forall l \in L, \end{array}$$

the solution of which gives a lower bound θ_l for the convex problem (12.6). Observe also that the best feasible solution in the generated sequence provides an upper bound θ_u for the convex problem (12.6), *i.e.*

$$\theta_u = \min_{l \in L} \{\chi(\psi^l)\}. \tag{12.12}$$

For a given upper bound θ, we call *localization set* the following polyhedral approximation

$$\mathcal{L}(\theta) = \{(\zeta, \psi) : \quad \theta \geq \zeta, \quad \zeta \geq \chi(\psi^l) + \langle X(\psi^l), \psi - \psi^l \rangle, \quad \forall l \in L\}. \tag{12.13}$$

It is the best (outer) approximation of the optimal set in (12.6).

We can now summarize the ACCPM algorithm for our special case.

1. Compute the analytical center[2] $(\overline{\zeta}, \overline{\psi})$ of the localization set $\mathcal{L}(\theta_u)$ and an associated lower bound $\underline{\theta}$.

2. Call the oracle at $(\overline{\zeta}, \overline{\psi})$. The oracle returns one or several cuts and an upper bound $\chi(\overline{\psi})$

3. Update the bounds:

   (a) $\theta_u = \min\{\chi(\overline{\psi}), \theta_u\}$

   (b) $\theta_l = \max\{\underline{\theta}, \theta_l\}$

4. Update the upper bound θ in the definition of the localization set (12.13) and add the new cuts.

[1]We must emphasize that in the general theory of ACCPM the oracle can generate two types of cuts: *optimality* cuts and *feasibility* cuts. However, for our special model we encounter only optimality cuts and therefore leave the interested reader to consult *e.g.* Ref. [53] for the description of a feasibility cut.

[2]It is beyond the scope of this thesis to give the definition of an analytical center. However, the interested readers can refer to the paper [113] for a detailed description.

These steps are repeated until a point is found such that $\theta_u - \theta_l$ falls below a prescribed optimality tolerance.

In our case, as the function χ is the maximum of card(E) functions, *i.e.*

$$\chi(\psi) = \max_{i \in E} \chi_i(\psi), \tag{12.14}$$

the oracle may generate multiple cuts, one for each i in E. The single cut (12.10) is replaced with the following card(E) cuts:

$$\chi(\psi) \geq \chi_i(\tilde{\psi}) + \langle X_i(\tilde{\psi}), \psi - \tilde{\psi} \rangle, \tag{12.15}$$

where $X_i(\tilde{\psi})$ is a subgradient of the function χ_i at $\tilde{\psi}$. This multiple cut approach is more efficient than a single cut approach since the computation time to introduce a cut is negligible and the work of the oracle is the same in both cases. Indeed, in both cases the oracle has to compute, at each iteration, $\chi_i(\psi)$ and $X_i(\tilde{\psi})$ for all subproblems i in E. In the single cut approach one selects the cut touching the epigraph of χ and one doesn't use the other cuts, contrarily to the multiple cut approach where all cuts are used.

At this point we must emphasize that the oracle can benefit from a parallel implementation, since the card(E) MDPs defined by (12.8-12.9) are totally decoupled and, therefore, can be solved on different computers.

Chapter 13

Numerical experiments

In this chapter we apply the decomposition method presented in the previous chapters to an example of a production plant with two types of human resources and four different market states. We then compare the performance of this decomposition approach with the performance of a frontal method. We also show that a good speed-up can be obtained with a parallel implementation of the decomposition method.

13.1 The model

We propose to study an example of a plant producing one good with the help of two production factors and subject to random changes of the market price. This example is a special instance of the class of the two-time-scale hybrid stochastic systems we presented in Chapter 10. The discrete variable ξ describes the state of the market, which influences the profit derived from the produced good. We suppose that we have four different market states, so the ξ-process takes value in the set $E = \{1, 2, 3, 4\}$. The continuous variable $x \in (\mathbb{R}^+)^2$ describes the state of the 2 different factors of production. More precisely, x_1 corresponds to the number of skilled employees while x_2 corresponds to the number of unskilled employees.

The output is determined by a CES production function[1]

$$Y(x_1, x_2) = \left(\eta[x_1]^{-\beta} + (1-\eta)[x_2]^{-\beta}\right)^{-\frac{1}{\beta}},$$

where $-1 < \beta < \infty$ is the substitution parameter ($\beta \neq 0$) and $0 < \eta < 1$ is the distribution parameter. The profit rate structure is described by the function

$$\begin{aligned}\pi(x_1(t), x_2(t), \xi(t), u_1(t), u_2(t)) \quad = \quad & c(\xi(t))Y(x_1(t), x_2(t)) \\ & -a_1x_1(t) - a_2x_2(t) - A_1x_1^2(t) - A_2x_2^2(t) \\ & -b_1u_1(t) - b_2u_2(t) - B_1u_1^2(t) - B_2u_2^2(t),\end{aligned}$$

[1] See, for example, [84] Section 9-4

where $c(i)$ is the selling price, given the market is in state $i \in E$, $a_k x_k(t) + A_k x_k^2(t)$ is a cost function, related to the holding of a stock $x_k(t)$ of employees and $b_k u_k(t) + B_k u_k^2(t)$ is a cost function related to the enrollment effort, $u_k(t)$, of new employees.

We assume that the the price is influenced by the level of production of the firm. We rank the 4 market states by increasing selling price and we suppose that only jumps to neighboring market states can occur. More precisely, the ξ-process transition rates are defined by

$$\begin{aligned} \varepsilon q_{i(i+1)}(x_1, x_2) &= \varepsilon \left(E_i - e_i Y(x_1, x_2)\right) \\ \varepsilon q_{i(i-1)}(x_1, x_2) &= \varepsilon \left(D_i + d_i Y(x_1, x_2)\right). \end{aligned}$$

The parameter ε is the time-scale ratio that will, eventually, be considered as very small. The positive terms e_i, E_i, d_i and D_i are parameters which depend on the market state $i \in E$. We see that the transition rate toward a highest market price is negatively correlated to the production level, whereas the transition rate toward a lowest market price is positively correlated to the production level.

The state x_1, which corresponds to the skilled employees, is described by a jump and diffusion process whereas the state x_2, which corresponds to the unskilled employees, is described by a pure diffusion process. The jumps correspond to massive departures of skilled employees (*e.g.* for launching a start-up company). When a jump occurs, given we are in state (x, i), the distribution of the jump is uniform in the interval $[0, x_1]$ and zero outside. In other words, the amount of the leaving employees is between zero and the total number of skilled employees. Let Λ be the jump's intensity. The dynamics of the employees is described by

$$dx_2(t) = [u_2(t) - \alpha_2 x_2(t)]dt + \sigma_2 d\omega_2(t),$$

$$dx_1(t) = [u_1(t) - \alpha_1 x_1(t)]dt + \sigma_1 d\omega_1(t) + dS,$$

where

$$S(t) = \int_0^t \int_{[0,1]} r(x_1(s^-), \rho) N(ds d\rho).$$

$N(\cdot)$ is a Poisson measure of intensity $\Lambda dt \times \Pi(d\rho)$. $\Pi(\cdot)$ has a uniform distribution concentrated on the segment $[0, 1]$ and $r(x_1, \rho) = -x_1 \rho$.

13.2 Implementation

We implemented two methods for the resolution of the limit model: the decomposition method presented in Chapter 12 and the frontal method which consists to solve the linear program (11.12-11.16) directly.

For the decomposition method, we use ACCPM (see [49], [50], [51] and [53]) with a policy improvement (PI) algorithm[2] for the oracle. The oracle is written in C and uses the sparse linear equation solver developed by Kunder and Sangiovanni-Vincentelli [82]. In addition, a parallel implementation of the decomposition method has been realized using MPI, a library of C-callable routine (see MPI's reference book [112]). The reader can find in Appendix C technical details of the implementation.

For the frontal method, the modeling was done with the software AMPL (see AMPL's reference book [38]). We solved the model with the commercial software CPLEX.

13.3 Numerical results

We consider the model described in the previous section with the set of parameter values given in Table 13.1. We solved the limit model, when ε tends to zero, with the decomposition method described in Chapter 12.

$\Lambda = 0.1$	$c(1) = 1.3$		$\alpha_1 = \alpha_2 = 0.05$
$\nu = 1.0$	$c(2) = 1.6$		$\sigma_1 = \sigma_2 = 3.0$
$\eta = 0.5$	$c(3) = 1.9$		$x_1^{\max} = x_2^{\max} = 100$
$\beta = -0.6$	$c(4) = 2.2$		$x_1^{\min} = x_2^{\min} = 0$
$a_1 = a_2 = 0.4$	$e_i = 0.002$	$\forall i \in E$	$h = 4$
$A_1 = A_2 = 0.004$	$E_i = 0.4$	$\forall i \in E$	$u_1^{\max} = u_2^{\max} = 10$
$b_1 = b_2 = 0$	$d_i = 0.004$	$\forall i \in E$	$u_1^{\min} = u_2^{\min} = 0$
$B_1 = B_2 = 0.05$	$D_i = 0.15$	$\forall i \in E$	$h_u = 2$

Table 13.1: List of parameter values for the numerical experiments.

The steady state probabilities are shown in Figure 13.1. The possibility of jumps in the x_1 axis are the cause of the tail in the probability distributions. As expected, the higher the selling price the higher is the production level. For comparison, we considered also the model associated with a fixed ξ-process, that is, the model where the selling price stays the same forever. For the fixed ξ-process, the steady state probabilities are shown in Figure 13.2. Given a market state, the production level is higher for the model associated with the fixed ξ-process than for the limit model. This comes from the fact that, when the price can change, the probability that it will increase, resp. decrease, is negatively, resp. positively, correlated with the production level. The effect of the production level on the price can be seen in Figure 13.3. In this Figure, we displayed, for the limit

[2] See [101] for a description of the policy improvement algorithm.

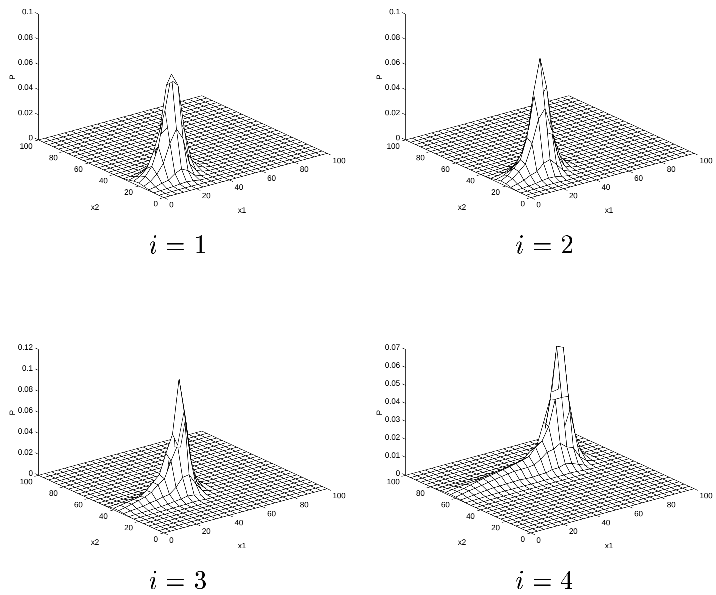


Figure 13.1: Steady state probabilities for the limit control problem, given the market state i.

model, the steady state probabilities as a function of the state for two policies, namely the optimal policy and the optimal policy of the model with fixed ξ-process[3]. We see distinctly that the price tends to be higher in the first case (where the production level is lower) than in the second case.

The maximal expected reward growth rate J equals 25.92. The value function is shown in Figure 13.4, for the case when the market is in the state $i = 2$. For the other states, the value functions are similar and therefore not displayed.

The optimal policy for the enrollment of new employees is shown in Figure 13.5, when the market is in state $i = 2$. For the other states, the optimal policies are similar and therefore not displayed.

13.4 Computational performance

In this section we compare the computational performance of the frontal method with the decomposition method presented in the previous chapters. It is beyond the scope of this presentation to study in details the performance of both method on a large range of different models; this has been done in Ref. [52]. However, on the same example as in the previous section, we show the advantages of the decomposition method compared with the direct method, namely

[3]Note that this second policy is, in general, not optimal for the limit model.

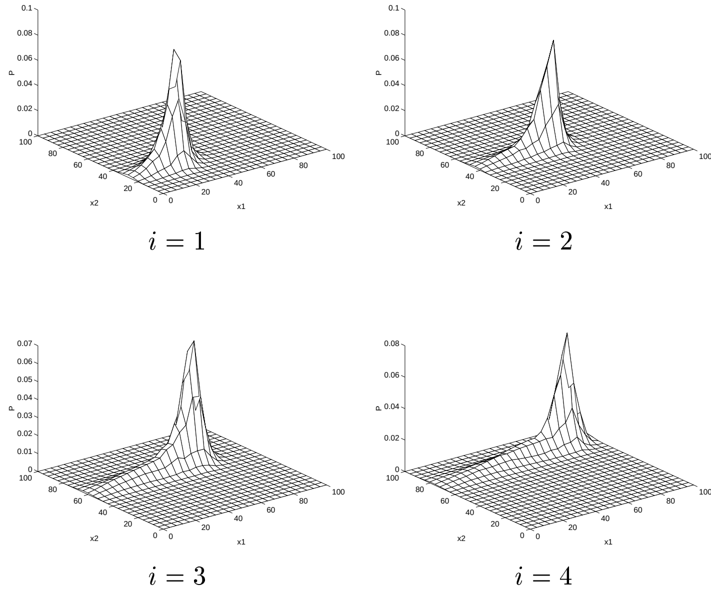


Figure 13.2: Steady state probabilities for the fixed ξ-process, given $\xi(t) = i \quad \forall t \geq 0$.

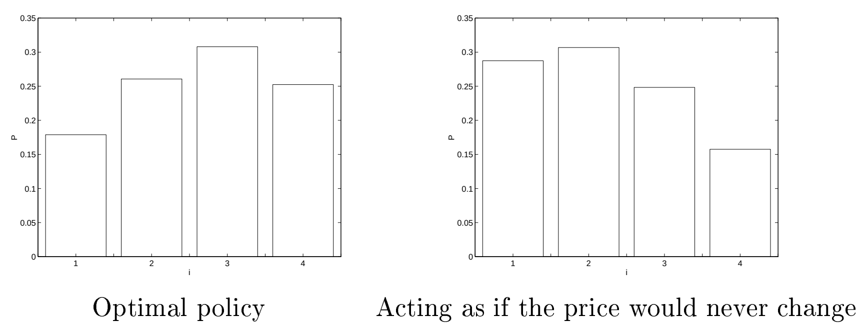


Figure 13.3: Steady state probabilities as a function of the state i.

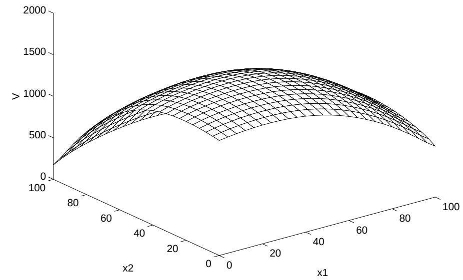
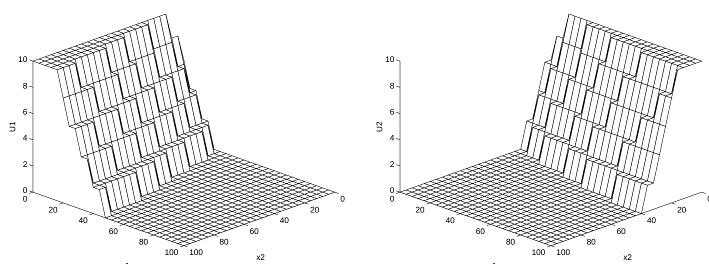


Figure 13.4: Value function $V(x, 2)$



$u_1(x)$ $u_2(x)$

Figure 13.5: Optimal policy $u(x)$, $i = 2$.

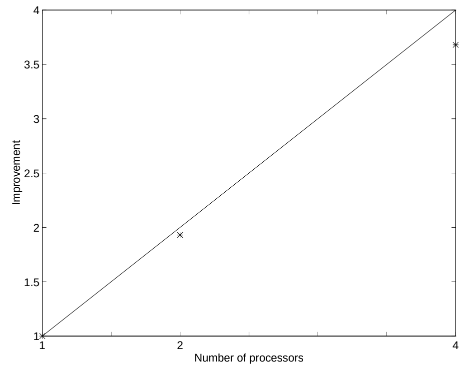


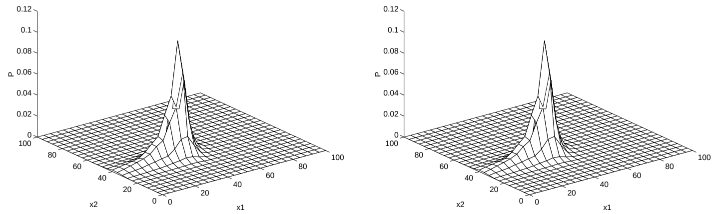
Figure 13.6: Speed-up as a function of the number of processors.

- reduction of the execution time,
- accuracy of the solution concerning the policies,
- reduction of the RAM memory utilized.

To illustrate the reduction in execution time, we considered the same model as in the previous section. The corresponding linear program has 2709 rows 94884 columns and 1873866 non-zero elements. To solve this problem, the direct method needs 1410[4] seconds, whereas the decomposition method needs 497 seconds. In addition, if we run on four processors the parallel version of the decomposition method, the execution time drops to 135 seconds. In Figure 13.6 we display, for the parallel implementation of the decomposition method, the speed-up as a function of the number of processors.

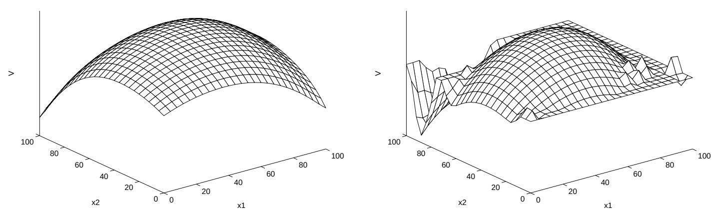
Before showing the inaccuracy of the optimal controls computed with the direct approach, let us show that the steady state probability obtained with this method are exact. Figure 13.7 shows the steady state probabilities, when the market is in state $i = 3$, for both methods. We see distinctly that both methods give the same (correct) result. In addition, the maximal expected reward growth rate J equals 25.92, for both methods. Although the linear programming direct approach gives an accurate solution concerning the steady state probabilities and the maximal expected reward growth rate, this method gives, in most cases, a wrong solution concerning the controls and the value function. Figure 13.8 shows the value function and Figure 13.9 shows the optimal policy, when the market is in state $i = 3$, for both methods. We see that the direct approach

[4]CPLEX offers three methods, namely the simplex, the dual simplex and an interior point method. The solver took 1410 seconds with the dual simplex, 4431 seconds with the primal simplex. The interior point method of CPLEX stopped after 445 seconds and proposed an infeasible solution with an objective value close to the optimal value. Running the crossover to obtain a feasible solution took 3040 seconds more.



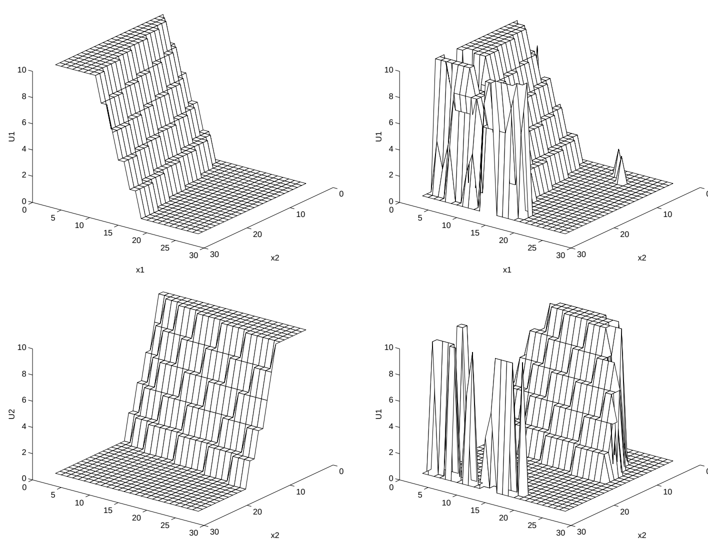
ACCPM-PI CPLEX

Figure 13.7: Steady state probability.



ACCPM-PI CPLEX

Figure 13.8: Value function.



ACCPM-PI CPLEX

Figure 13.9: Optimal policy.

gives an accurate result in the middle of the grid but a wrong result near the boundaries. This can be explained as follows. The direct method consists in solving the linear program (11.12-11.16), where the objective function (11.12) is a sum weighted with the steady states probabilities. But, as we can see in Figure 13.7, near the boundaries, the steady state probabilities are very close to zero. Therefore, near the boundaries, an imprecision in the optimal control has negligible effects on the objective function (11.12). We can see distinctly that this problem is not present for the decomposition method with an oracle using a policy improvement algorithm. Notice that this problem is inherent to a linear programming approach. A decomposition method using an oracle based on a linear programming method instead of a policy improvement (PI) algorithm would meet the same problem.

Finally, to compare the RAM memory utilized by the two methods we run the largest possible model for both methods. Again, we considered the same model as in the previous section, but, in order to allow modifications of the model's size, we took different values for the continuous state discretization mesh h. The results are significant: the size of the largest model is about 50% higher when using the decomposition approach rather than the direct approach. Indeed, the largest model solved with the decomposition approach has a mesh value of 2.5, which corresponds to a grid of 40×40; whereas for the frontal approach, the largest model solved has a mesh value of 2.0, which corresponds to a grid of 50×50. The sizes of the corresponding linear programs are given in the following table:

Mesh value	grid size	rows	columns	nonzeros
2.5	40×40	6729	242064	6639048
2.0	50×50	10404	374544	≈10000000

Chapter 14

Concluding remarks for Part II

In this part of the thesis we have implemented a decomposition method for the resolution of hybrid stochastic models with two time scales. This method, which was proposed by Filar and Haurie in [35] and [36], reformulates the initial problem as an approximating singularly perturbed MDP that can be solved as a structured linear programming problem. The originality of this work was the coupling of ACCPM with a policy improvement algorithm to achieve a decomposition in order to exploit the special bloc-diagonal structure. We showed the impact of such an implementation compared with a frontal method, on the reduction of the RAM memory utilized, the reduction of the execution time and the accuracy of the solution concerning the policies. We also showed that a good speed-up could be obtained with a parallel implementation of the method.

Appendix B

Decomposition of the generator of the Markov decision problem

The value of $B_h[(x,i),(x',j)|u]$ and $\varepsilon D_h[(x,i),(x',j)|u]$ can be calculated as follows. Define

$$R_h(x,i,u) = \sum_{k=1}^{K}\{\sigma_k^2 + h|f_k^i(x,u)|\}$$

and

$$\tilde{R}_h = \max_{x,i,u} R_h(x,i,u).$$

Recall that we have already defined

$$\tilde{Q}_h = \varepsilon q^{\tilde{i}}(\tilde{x},\tilde{u})h^2 + \sum_{k=1}^{K}\{\sigma_k^2 + h|f_k^i(\tilde{x},\tilde{u})|\}$$

with $(\tilde{x},\tilde{i},\tilde{u}) = \arg\max_{(x,i,u)} Q_h(x,i,u)$. As, for fixed h and sufficient small ε we have

$$\tilde{R}_h = \sum_{k=1}^{K}\{\sigma_k^2 + h|f_k^i(\tilde{x},\tilde{u})|\} = \tilde{Q}_h - \varepsilon q^{\tilde{i}}(\tilde{x},\tilde{u})h^2,$$

we can apply the following approximation

$$\frac{1}{\tilde{Q}_h} = \frac{1}{\tilde{R}_h[1+\frac{\varepsilon q^{\tilde{i}}(\tilde{x},\tilde{u})h^2}{\tilde{R}_h}]} = \frac{1}{\tilde{R}_h} - \frac{\varepsilon q^{\tilde{i}}(\tilde{x},\tilde{u})h^2}{\tilde{R}_h^2} + o(\varepsilon)$$

to obtain

$$B_h[(x,i),(x \pm e_k h, i)|u] = \frac{\frac{\sigma_k^2}{2} + hf_k^i(x,u)^{\pm}}{\tilde{R}_h},$$

$$B_h[(x,i),(x,i)|u] = -\sum_{k=1}^{K}\{B_h[(x,i),(x+e_kh,i)|u] + B_h[(x,i),(x-e_kh,i)|u]\},$$

$$D_h[(x,i),(x\pm e_kh,i)|u] = -\frac{q^{\tilde{i}}(\tilde{x},\tilde{u})h^2}{\tilde{R}_h}B_h[(x,i),(x\pm e_kh,i)|u],$$

$$D_h[(x,i),(x,j)|u] = h^2\frac{q_{ij}(x,u)}{\tilde{R}_h} \qquad i\neq j,$$

and

$$\begin{aligned} D_h[(x,i),(x,i)|u] &= -\sum_{k=1}^{K}\{D_h[(x,i),(x+e_kh,i)|u] + \\ &\qquad D_h[(x,i),(x-e_kh,i)|u]\} \\ &\quad -\sum_{j\neq i} D_h[(x,i),(x,j)|u]. \end{aligned} \tag{B.1}$$

Appendix C

Implementation

In this appendix we present the main points relative to the implementation of the decomposition method. In the first section, we show how the dimension of the space of the optimization problem (12.6) has to be reduced from card(E) to card(E) $-$ 1. In the second section we show how to compute the subgradients needed to characterize the cuts. Finally in the third section we give the pseudo-code of the parallel implementation.

C.1 Reduction of the dimension of the optimization problem

Recall that for a fixed i in E, $\chi_i(\psi)$ is the value function of the MDP defined by (12.8-12.9). The dual problem of (12.8-12.9) writes

$$\chi_i(\psi) = \max \sum_x \sum_u \left(\pi(x,i,u) - \sum_j \sum_{x'} D_h[(x,i),(x',j)|u]\psi(j) \right) Z^i(x,u) \tag{C.1}$$

s.t.

$$\sum_x \sum_u B_h[(x,i),(x',i)|u] Z^i(x,u) = 0 \quad x' \in \mathcal{X}$$

$$\sum_x \sum_u Z^i(x,u) = 1$$

$$Z^i(x,u) \geq 0.$$

Using property (B.1) we can rewrite the modified cost function (C.1) as

$$\chi_i(\psi) = \sum_x \sum_u \left(\pi(x,i,u) + \sum_{j \neq i} D_h[(x,i),(x,j)|u](\psi(i) - \psi(j)) \right) Z^i(x,u).$$

In this expression only the difference $\psi(i) - \psi(j)$ is relevant and one can, without loss of generality, fix one variable, say $\psi(\text{card}(E))$, to an arbitrary value (zero for example). The $\text{card}(E)$ functions $(\chi_1, \ldots, \chi_{\text{card}(E)})$ are therefore totally characterized by the $\text{card}(E) - 1$ variables $(\psi(1), \ldots, \psi(\text{card}(E) - 1))$.

In the model presented in Chapter 13, an other issue is possible. For this model, the ξ-process can only jump to neighboring states (*i.e.* transitions from i to $i \pm 1$). Therefore we can rewrite the modified cost function (C.1) as follows:

$$\begin{aligned}\sum_x \sum_u \{&\pi(x,i,u) + D_h[(x,i),(x,i+1)|u][\psi(i) - \psi(i+1)] \\ &+ D_h[(x,i),(x,i-1)|u][\psi(i) - \psi(i-1)]\} \, Z^i(x,u).\end{aligned}$$

In this expression only the difference $\psi(i \pm 1) - \psi(i)$ is relevant and we can, without lost of generality, use, in place of the $\text{card}(E)$ variables $\psi(i)$, the $\text{card}(E) - 1$ variables $\Delta\psi(i)$ defined as follows:

$$\Delta\psi(i) = \psi(i) - \psi(i+1) \qquad i \in \{1, \ldots, \text{card}(E) - 1\}.$$

The modified cost function (C.1) can be rewritten

$$\begin{aligned}\chi_i(\Delta\psi) \;=\; \sum_x \sum_u \{&\pi(x,i,u) + D_h[(x,i),(x,i+1)|u]\Delta\psi(i) \\ &- D_h[(x,i),(x,i-1)|u]\Delta\psi(i-1)\} \, Z^i(x,u).\end{aligned}$$

The convex optimization problem (12.6) writes, for the model presented in Chapter 13

$$\min_{\Delta\psi \in \mathbb{R}^{\text{card}(E)-1}} \chi(\Delta\psi).$$

We must emphasize that if one uses the $\text{card}(E)$ variables $\psi(i)$, ACCPM can encounter problems of convergence. Therefore, the uses of $\text{card}(E) - 1$ variables (for example $\Delta\psi$ for the model presented in Chapter 13) is highly recommended.

C.2 Computation of the subgradients

Remind that at each query point, the oracle has to return for all i in E the value of χ_i and a subgradient of χ_i. For the general case, a subgradient of $\chi_i(\psi)$, denoted $X_i = \left(X_i^1, \ldots, X_i^{\mathrm{card}(E)}\right)$, can be computed as follow

$$X_i^j = \begin{cases} \sum_x \sum_u \sum_{j \neq i} D_h[(x,i),(x,j)|u]\tilde{Z}^i(x,u) & \text{if } i = j \\ -\sum_x \sum_u D_h[(x,i),(x,j)|u]\tilde{Z}^i(x,u) & \text{otherwise,} \end{cases}$$

where $\tilde{Z}^i(x,u)$ are the optimal values of $Z^i(x,u)$.

For the model presented in Chapter 13, we showed in the previous section that one can use the variables $\Delta\psi(i)$. For this case, a subgradient of $\chi_i(\Delta\psi)$, denoted $X_i = \left(X_i^1, \ldots, X_i^{\mathrm{card}(E)-1}\right)$ can be computed as follow

$$X_i^j = \begin{cases} \sum_x \sum_u D_h[(x,i),(x,i+1)|u]\tilde{Z}^i(x,u) & \text{if } j = i \\ -\sum_x \sum_u D_h[(x,i),(x,i-1)|u]\tilde{Z}^i(x,u) & \text{if } j = i-1 \\ 0 & \text{otherwise.} \end{cases}$$

C.3 Parallel implementation

Below, we give the pseudo-code for the parallel implementation. The names of the communication routines of the MPI library are written in bold type[1].

[1] The routine **broadcast** is utilized for both sending and receiving communication. To avoid confusion we write "**broadcast**" when the communication is sent and "receive **broadcast**ed" when the communication is received.

Initialize MPI.
Compute $E(p) \in E$, the set of subproblems distributed to the current processor p.

IF the current processor is the master processor
  THEN
    Initialize ACCPM.
    Initialize the value of $\overline{\psi}$.
    Set continue=true.
    Broadcast to all processors the current value of $\overline{\psi}$ and S.
    WHILE continue
      DO
      Compute the value of $\chi_i(\overline{\psi})$ and $X_i(\overline{\psi})$ for $i \in E(p)$,
        where p is the current processor.
      Receive from all other processors p' the value of $\chi_i(\overline{\psi})$ and $X_i(\overline{\psi})$
        for all $i \in E(p')$.
      Add the card(E) cuts.
      Compute the new analytic center.
      Update the bounds.
      IF the optimality tolerance is achieved
        THEN Set $S =$ stop.
      Broadcast to all processors the current value of $\overline{\psi}$ and S.
      DONE

IF the current processor is not the master processor
  THEN
    Receive from the master processor the **broadcast**ed value of $\overline{\psi}$ and S.
    WHILE continue
      DO
      Compute the value of $\chi_i(\overline{\psi})$ and $X_i(\overline{\psi})$ for $i \in E(p)$,
        where p is the current processor.
      Send to the master processor the value of $\chi_i(\overline{\psi})$ and $X_i(\overline{\psi})$
        for all $i \in E(p)$.
      Receive from the master processor the **broadcast**ed
        value of $\overline{\psi}$ and S.
      DONE

IF the current processor is the master processor
  THEN
    Print the results
    Free ACCPM.

Free MPI.

Part III

Computation of S-adapted Equilibria in Piecewise Deterministic Games via Stochastic Programming Methods

Chapter 15

Introduction to Part III

The aim of this third part of the presentation is to propose a numerical technique for the approximation of a class of equilibria in a stochastic game of oligopoly. These equilibria, called S-adapted in [68] correspond to an information structure where the players adapt their actions to an observation of the realization of the random disturbances affecting the game dynamics. These disturbances are supposed to take the form of an uncontrolled jump process. Recently this class of problems has received a renewed attention from researchers in Mathematical Programming circles (see e.g. [55] and [54]) who extended the numerical experiments reported in [68]. The present dissertation complements these previous works in the following way:

1. The oligopoly model is formulated in continuous time as in [65];
2. the S-adapted information structure is compared with the Piecewise Open Loop (POL) information structure used in [65];
3. an approximation to the S-adapted equilibrium is obtained through the solution of a sequence of variational inequality problems defined via a discretization over time of the game dynamics and perturbing jump process;
4. the approximating S-adapted equilibrium is proved to be unique under strict diagonal concavity of the total reward function;
5. convergence results are proved for the approximating S-adapted equilibrium;
6. a numerical example, consistent with the one given in [65] is fully detailed and shows the proximity of the equilibria under S-adapted and piecewise open-loop information structure.

This part of the dissertation is organized as follows. In Chapter 16 we recall the concepts of Nash equilibrium and variational inequality and show the link between them. In Chapter 17 we formulate a differential game of oligopoly with an open-loop information structure and we show that the Nash-equilibrium can be approximated via a variational inequality solution using mathematical programming techniques. This provides another efficient way to approximate an open-loop equilibrium in a differential game of oligopoly. In Chapter 18 a piecewise deterministic game version of the same oligopoly model is proposed and the concept of S-adapted information structure is discussed. S-adapted equilibria are compared with POL equilibria and one conjectures that these equilibria could coincide in many cases. An approximation of the S-adapted equilibria through a sequence of variational inequality solutions is proposed. In Chapter 19 some numerical experiments are reported and a comparison with the POL information structure is made on the basis of the numerical solutions obtained, which tends to confirm the conjecture.

Chapter 16

Nash equilibrium and variational inequality

The study of oligopoly is closely linked to the concept of Nash-Cournot equilibrium. Oligopoly theory dates back to Cournot (see [24]), who investigated competition between two producers in a noncooperative behavior. In his book, the decisions made by the producers are said to be in equilibrium if no one can increase his income by unilateral action, given that that the other producer does not alter his decision. Nash [95, 96] subsequently generalized the Cournot equilibrium concept to a noncooperative game, and proved existence of mixed-strategy equilibria for N-player matrix games.

In this chapter we recall the concept of Nash-Cournot equilibrium and the possible characterization of an equilibrium in a continuous game through variational inequalities. A complete theory on the subject can be found in Nagurney's book [94].

16.1 Nash equilibrium

Let us first introduce some notations that will be used throughout the rest of the dissertation. Let $\mathcal{J}$ be the set of indices $\{1, 2, \ldots, J\}$ representing the players. We denote u_j the control of player j, $\mathbf{u}$ the J-dimensional vector formed with the control of each players and $(v, \mathbf{u}_{-j})$, the J-dimensional vector formed with the components of $\mathbf{u}$, where the j-th component has been replaced with v. Denote $V_j(\mathbf{u})$ the payoff for player j assuming each player j' has chosen the control $u_{j'}$. Finally, let $U = (U_1, \ldots, U_J)$, where U_j denote the set of the admissible controls for player j. Each player tries to maximize his own payoff.

The controls $\mathbf{u}^*$ are said to be in Nash equilibrium if

$$V_j(\mathbf{u}^*) \geq V_j(v_j, \mathbf{u}^*_{-j}) \quad \forall j \in \mathcal{J} \quad \forall v_j \in U_j.$$

In other words, we have a Nash equilibrium when each player has no incentive to change unilaterally his control.

16.2 Variational inequality

Recall first the definition of a variational inequality problem. Let $\mathbf{u}$ and $\mathbf{u}^*$ be, as in the previous section, J- dimensional real vectors. We say that $\mathbf{u}^* \in U$ is solution of the variational inequality associated with the function $F : R^n \to R^n$ and the convex set $U \in R^n$ if it satisfies the following inequality

$$\langle F(\mathbf{u}^*), \mathbf{u}^* - \mathbf{u} \rangle \geq 0 \quad \forall \mathbf{u} \in U.$$

The variational inequality problem is a general formulation that encompasses a plethora of mathematical problem, including, among others, nonlinear equations, optimization problems, complementary problems, fixed point problems and equilibrium problems.

It has been shown (see [56] and [43]) that, under enough regularity conditions, Nash equilibria satisfy variational inequalities. In the present context, under the assumption that the payoff function $V_j(\cdot)$ of each player j is continuously differentiable on U and concave with respect to u_j, $\mathbf{u}^*$ is a Nash equilibrium if and only if $\mathbf{u}^*$ is a solution of the variational inequality associated with the function

$$F(\mathbf{u}) = \begin{pmatrix} \nabla_{u_1} V_1(\mathbf{u}) \\ \vdots \\ \nabla_{u_j} V_j(\mathbf{u}) \\ \vdots \\ \nabla_{u_J} V_J(\mathbf{u}) \end{pmatrix}$$

and the set $U = (U_1, \ldots, U_J)$.

This can be shown as follows. For a fixed player j, the first order necessary and sufficient optimality conditions (see *e.g.* [88]) state that no feasible ascent direction exist at the optimum, *i.e.*,

$$\nabla_{u_j} V_j(\mathbf{u}^*)(u_j^* - u_j) \geq 0 \quad \forall u_j \in U_j. \tag{16.1}$$

Aggregating (16.1) for all players yields the desired result.

This link between Nash equilibria and variational inequalities will be used intensively throughout the rest of the dissertation.

Chapter 17

The deterministic dynamic oligopoly

In this chapter we consider a deterministic differential game model of oligopoly, propose an approximating discrete time model, show that the open-loop equilibrium in the discrete time model can be computed via the solution of a variational inequality and prove that the equilibrium of this approximating game converges to the equilibrium of the initial game as the dicretizations step tends to zero.

17.1 The formulation of the oligopoly

The model is similar to those studied in [17], [61] and [65]. There are J competing firms (also called players) supplying a market for an homogeneous good. Let $\mathcal{J} = \{1, 2, \ldots, J\}$ be the set of all players. The control variables are such that $u_j \in [u_j^{\min}, u_j^{\max}]$, $j \in \mathcal{J}$; they represent the investment in production capacity by each firm. The state variables are $x_j \in \mathbb{R}^+$, $j \in \mathcal{J}$; they represent the accumulated production capacity of each firm.

The state equation for player j is given by

$$\dot{x}_j(t) = u_j(t) - \mu_j x_j(t) \quad j \in \mathcal{J}, \tag{17.1}$$

$$x_j(0) = x_j^0. \tag{17.2}$$

With $u_j(\cdot)$ a measurable function over $[0, T]$, bounded above by $u_j^{\max}$ and below by $u_j^{\min}$, such that the generated trajectory $x_j(t)$ is not negative. μ_j is the capacity depreciation rate for firm j. Given the initial state $x_j(0) = x_j^0$, the solution $x_j(\cdot)$ of (17.1) can be expressed as a function of the control $u_j(\cdot)$:

$$x_j(t) = e^{-\mu_j t} x_j^0 + \int_0^t e^{-\mu_j (t-s)} u_j(s)\, ds. \tag{17.3}$$

The information structure is open-loop; hence each player j knows the initial states $(x_1^0, \ldots, x_J^0) = x^0$ and chooses a control function $u_j(\cdot) : [0, T] \to [u_j^{\min}, u_j^{\max}]$ which generates a positive trajectory. Let U_j be the set of admissible control functions and X_j the set of trajectories $x_j(\cdot)$ generated by admissible controls $u_j(\cdot)$ in U_j.

The profit functions are thus defined by

$$V_j(x^0; u_1(\cdot), \ldots, u_J(\cdot)) = \int_0^T e^{-\rho_j t} L_j(x(t), u_j(t))\, dt, \tag{17.4}$$

where ρ_j is the discount rate for player j, $x(t) = (x_1(t), \ldots, x_J(t))$ and $L_j(x, u_j)$ is a profit rate function which is assumed to be C^1 in x and in u_j.

Definition 2. *Let $\epsilon \geq 0$. The J-tuple $(u_1^*(\cdot), \ldots, u_J^*(\cdot))$ is an ϵ-Nash open-loop equilibrium if we have for all j in $\mathcal{J}$ and for all $u_j(\cdot)$ in U_j*

$$V_j(x^0; u_1^*(\cdot), \ldots, u_j(\cdot), \ldots, u_J^*(\cdot)) \leq V_j(x^0; u_1^*(\cdot), \ldots, u_J^*(\cdot)) + \epsilon.$$

If $\epsilon = 0$ we obtain a Nash equilibrium.

17.2 An equilibrium principle

The Hamiltonian of player j is defined as usual by

$$\mathcal{H}_j(t, x, u_j, p_j) = L_j(x, u_j) + p_j(u_j - \mu_j x_j).$$

The optimized Hamiltonian is

$$H_j(t, x, p_j) = \max_{u_j \in [u_j^{\min}, u_j^{\max}]} \mathcal{H}_j(t, x, u_j, p_j).$$

In order for $(x_j(\cdot),\, u_j(\cdot))_{j \in \mathcal{J}}$ to be an open-loop equilibrium it is necessary that there exist absolutely continuous costate trajectories $(p_j(\cdot))_{j \in \mathcal{J}}$ such that

$$\dot{p}_j(t) = -\frac{\partial H_j(t, x(t), p_j(t))}{\partial x_j},$$

$$\dot{x}_j(t) = \frac{\partial H_j(t, x(t), p_j(t))}{\partial p_j},$$

and

$$u_j(t) = \arg \max_{u_j \in [u_j^{\min}, u_j^{\max}]} \mathcal{H}_j(t, x(t), u_j, p_j(t)).$$

With the transversality condition $p_j(T) = 0$.

Definition 3. *The combined Hamiltonian* $\sum_j H_j(t,x,p)$ *is strictly diagonally concave in* x*, convex in* p *if for all* t, x, $\tilde{x}$, p *and* $\tilde{p}$

$$\sum_j (p_j - \tilde{p}_j)(\nabla_{p_j} H_j(t,x_1,\ldots,x_J,p_j) - \nabla_{p_j} H_j(t,\tilde{x}_1,\ldots,\tilde{x}_J,\tilde{p}_j))$$
$$- (x_j - \tilde{x}_j)(\nabla_{x_j} H_j(t,x_1,\ldots,x_J,p_j) - \nabla_{x_j} H_j(t,\tilde{x}_1,\ldots,\tilde{x}_J,\tilde{p}_j)) > 0.$$

Definition 4. *The total reward function* $\sum_j L_j(x_1,\ldots,x_J,u_j)$ *is strictly diagonally concave in* (x,u) *if for all* x, $\tilde{x}$, p *and* $\tilde{p}$

$$\sum_j (u_j - \tilde{u}_j)(\nabla_{u_j} L_j(x_1,\ldots,x_J,u_j) - \nabla_{u_j} L_j(\tilde{x}_1,\ldots,\tilde{x}_J,\tilde{u}_j))$$
$$+ (x_j - \tilde{x}_j)(\nabla_{x_j} L_j(x_1,\ldots,x_J,u_j) - \nabla_{x_j} L_j(\tilde{x}_1,\ldots,\tilde{x}_J,\tilde{u}_j)) < 0.$$

The strict diagonal concavity in x, convexity in p of the combined Hamiltonian can be verified by applying the following Lemma which we borrow from [23].

Lemma 2. *Assume* $L_j(x,u_j)$ *is concave in* (x,u_j) *and assume that the total reward function* $\sum_j L_j(x,u_j)$ *is strictly diagonally concave in* (x,u) *then the combined Hamiltonian* $\sum_j H_j(t,x,p)$ *is strictly diagonally concave in* x*, convex in* p*.*

The following uniqueness result can then be proved as in [23].

Theorem 11. *If the combined Hamiltonian is strictly diagonally concave in* x*, convex in* p*, then the open-loop equilibrium is unique.*

17.3 A discrete time approximation

We shall now explore a method for approximating the open-loop equilibrium of the duopoly game that uses a sequence of solutions of variational inequalities to get an ϵ-equilibrium of the continuous time game.

We use a discrete time approximation of the dynamic oligopoly model. The approximating game of order K is defined as follows: Let $t_k = \delta k$ with $k = 0,\ldots,K$ and $\delta = T/K$. The discrete time state and control variables are $\mathbf{x}_j^K = (x_j^K(k))_{k=1}^K$ and $\mathbf{u}_j^K = (u_j^K(k))_{k=1}^K$ respectively. Using a slight abuse of notation we call $x_j^K(k)$ the discrete time state at time t_k and similarly for $u_j^K(k)$. The state equations are the difference equations:

$$x_j^K(k) = u_j^K(k)\frac{1-e^{-\mu_j\cdot\delta}}{\mu_j} + e^{-\mu_j\cdot\delta}x_j^K(k-1) \tag{17.5}$$

$$= x_j^0 e^{-\mu_j t_k} + \frac{1-e^{-\mu_j\cdot\delta}}{\mu_j}\sum_{l=1}^{k} u_j^K(l)e^{-\mu_j(t_k-t_l)} \tag{17.6}$$

and the profit functions are given by

$$V_j^K(x^0; \mathbf{u}_1^K, \dots, \mathbf{u}_J^K) = \sum_{k=1}^{K} e^{-\rho_j t_k} L_j(x_1^K(k), \dots, x_J^K(k), u_j^K(k))\delta. \tag{17.7}$$

Where $(x_1^K(k), \dots, x_J^K(k))_{k=1,\dots,K}$ is the trajectory J-tuple emanating from x^0 and generated by the controls as shown in(17.6).

An admissible open-loop strategy for player j in the approximating game of order K is a vector $\mathbf{u}_j^K \in [u_j^{\min}, u_j^{\max}]^K$ such that the generated trajectory $\mathbf{x}_j^K$ remains positive. Let U_j^K be the set of strategies and X_j^K the set of the corresponding trajectories for player j. We have thus defined a game where the strategies are elements of an Euclidean space. The equilibrium $\mathbf{u}^{K*} = (\mathbf{u}_1^{K*}, \dots, \mathbf{u}_J^{K*})$ is a solution of the following variational inequality:

$$\langle F(\mathbf{u}^{K*}), \mathbf{u}^{K*} - \mathbf{u}^K \rangle \geq 0 \quad \forall \mathbf{u}^K \in U_1^K \times \cdots \times U_J^K = U^K, \tag{17.8}$$

where $\langle \cdot, \cdot \rangle$ denotes the scalar product and

$$\mathbf{u}^K = \begin{pmatrix} \mathbf{u}_1^K \\ \vdots \\ \mathbf{u}_J^K \end{pmatrix},$$

$$F(\mathbf{u}^K) = \begin{pmatrix} \nabla_{\mathbf{u}_1^K} V_1^K(x^0; \mathbf{u}_1^K, \dots, \mathbf{u}_J^K) \\ \vdots \\ \nabla_{\mathbf{u}_J^K} V_J^K(x^0; \mathbf{u}_1^K, \dots, \mathbf{u}_J^K) \end{pmatrix}. \tag{17.9}$$

The gradients of the reduced profit functions can easily be obtained from Eq.(17.7) once one expresses

$$\frac{\partial x_j^K(k)}{\partial u_j^K(l)} = \begin{cases} \frac{1-e^{-\mu_j \cdot \delta}}{\mu_j} e^{-\mu_j (t_k - t_l)} & \text{if } k \geq l \\ 0 & \text{if } k < l. \end{cases}$$

Let us first recall the definition of monotony.

Definition 5. *$G(\cdot) : U^K \to \left(I\!R^K\right)^J$ is a monotone operator in U^K if it satisfies*

$$\langle G(\mathbf{u}^K) - G(\tilde{\mathbf{u}}^K), \mathbf{u}^K - \tilde{\mathbf{u}}^K \rangle \geq 0 \quad \forall \mathbf{u}^K, \tilde{\mathbf{u}}^K \in U^K$$

Theorem 12. *Assume the total reward function $\sum_j L_j(x_1, \dots, x_J, u_j)$ is strictly diagonally concave in (x, u) then the operator $-F(\cdot)$ defined in Eq. (17.9) is monotone.*

Proof. By straightforward verification. ∎

Theorem 13. *Under the assumptions of Theorem 12, there exists a unique equilibrium for the approximating game of order K.*

Proof. This theorem is a special case of Theorem 17, to be proved later on for the stochastic case. □

Assumption 5. *We suppose the following*

- *There is state and control separation in the profit rate functions, i.e.:*

$$L_j(x_1, \ldots, x_J, u_j) = \mathcal{L}_j(x_1, \ldots, x_J) + \mathcal{M}_j(u_j)$$

  where $\mathcal{L}_j$ and $\mathcal{M}_j$ are Lipschitz continuous functions.

- *$\sum_j \mathcal{L}_j(x)$ is strictly diagonally concave in x and $\mathcal{M}_j(u_j)$ is strictly concave in u_j.*

We now address the question of approximating the solution of the continuous time game through the solution of approximating games. Our approach is inspired from [100] and [7]. To establish a correspondence between the continuous time game and its approximation of order K let us define the mappings $\phi_j^K : U_j^K \to U_j$ and $\sigma_j^K : U_j \to U_j^K$ as follows:

$$\phi_j^K(\mathbf{u}_j^K)(t) = u_j^K(k) \quad \text{where } k \text{ is such that } t_k = \min_s \{t_s | t_s \geq t\}$$

$$\sigma_j^K(u_j)(k) = \frac{\mu}{1 - e^{-\mu\delta}} \int_{t_{k-1}}^{t_k} u_j(s) e^{-\mu(t_k - s)}\, ds.$$

With each control for the discrete time game of order K, the mapping ϕ_j^K associates a piecewise constant control for the continuous time game. With each control for the continuous time game, the mapping σ_j^K associates a control for the discrete time game of order K. One can verify that these mappings satisfy the following property $\sigma_j^K \circ \phi_j^K = \mathbb{1}$. Furthermore the mappings preserve the property of non-negativity of the generated trajectory.

The convergence of the discrete time equilibrium toward the continuous time equilibrium is stated in the two following theorems.

Theorem 14. *Suppose assumption 5 holds. Let $u_1^*, \ldots, u_J^*$ be the equilibrium controls of the continuous time oligopoly. Then for all positive ϵ there exists K_ϵ such that for all $K > K_\epsilon$ the control vector $(\sigma_1^K(u_1^*), \ldots, \sigma_J^K(u_J^*))$ is an ϵ-Nash equilibrium for the discrete time oligopoly of order K.*

Theorem 15. *Suppose assumption 5 holds. Let $\mathbf{u}_1^{K*}, \ldots, \mathbf{u}_J^{K*}$ be the equilibrium controls of the discrete time oligopoly of order K. Then for all positive ϵ there exists K_ϵ such that for all $K > K_\epsilon$ the controls vector $(\phi_1^K(\mathbf{u}_1^{K*}), \ldots, \phi_J^K(\mathbf{u}_J^{K*}))$ is an ϵ-Nash equilibrium for the continuous time oligopoly.*

Proof. The proofs of these two theorems can be found in Appendix D.2. □

Chapter 18

The stochastic dynamic oligopoly

A stochastic oligopoly model has been proposed, in a discrete time setting, by Haurie, Smeers, Zaccour and Legrand in [68] and [67]. In the proposed model the random disturbances were uncontrolled. The information structure used in these papers has been called S-adapted, for "sample path adapted", and the equilibrium has been computed via the solution of a variational inequality. These papers extended, in some sense, the stochastic programming technique to the case of Nash-Cournot equilibria. Recently, this discrete time stochastic equilibrium framework has been further studied in [55].

In [65], Haurie and Roche have studied a stochastic oligopoly model, in a continuous time setting and with uncontrolled random jump disturbances. These authors used the information structure called *Piecewise Open-Loop* (POL). The POL-equilibrium was characterized and approximated through the solution of a discrete event dynamic programming equation.

In this chapter we revisit the stochastic oligopoly model presented in [65] but with the S-adapted information structure. We compare the S-adapted equilibrium with the POL equilibrium and we conjecture that, in many instances, these equilibrium solutions will coincide. We propose an approximation via a discrete time model, show that the S-adapted equilibrium in the discrete time model can be computed via the solution of a variational inequality. We prove that under strict diagonal concavity of the total reward function, there exists a unique S-adapted equilibrium for the approximating game. Finally we show that the equilibria of the approximating games converge to an equilibrium of the continuous time game.

18.1 A system with jump Markov disturbances

The state equations are still given by

$$\dot{x}_j(t) = u_j(t) - \mu_j x_j(t) \quad j \in \mathcal{J}$$

With $u_j^{\min} \leq u_j(t) \leq u_j^{\max}$ and $x_j(0) = x_j^0$. Let $\{\xi(\cdot) : [0,T] \to \mathcal{I}\}$ be a continuous-time discrete-state Markov chain taking values in the finite set $\mathcal{I} = \{1,\ldots,I\}$ which describes random changes in the market condition. The dynamics of $\xi(t)$ is defined by the transition rate matrix $Q = [q_{ih}] \quad i,h \in \mathcal{I}$. The profit rate functions $L_j^i(x,u_j)$ now also depend on the market condition $i \in \mathcal{I}$.

18.2 The S-adapted information structure

Let $(\Omega, \mathcal{B}, P)$ be the probability space for the $\xi(\cdot)$ process. We call $\xi(\omega,\cdot) : [0,T) \to \mathcal{I}$, $\omega \in \Omega$, a sample path of the ξ-process and $\xi_\theta(\omega,\cdot) : [0,\theta] \to \mathcal{I}$ its history up to time θ. We assume that the players know the initial state $x^0 = (x_1^0,\ldots,x_J^0)$ and, at each instant t, the ξ-process history up to time t, *i.e.* $\xi_t(\omega,\cdot)$. The game is played as follows.

- Let $\tau_0 = 0$, τ_1, τ_ν, $\nu = 0,1,2,\ldots$ be the successive jump times of the ξ-process and $\xi_0 = \xi(0)$, $\xi_1 = \xi(\tau_1)$, $\xi_\nu = \xi(\tau_\nu)$ the visited states;
- Call $h_\nu = \{\tau_0,\xi_0,\tau_1,\xi_1,\ldots,\tau_\nu,\xi_\nu\}$ the jump process history up to jump ν;
- at any jump time τ_ν, each player $j \in \mathcal{J}$, knowing x^0 and h_ν chooses an open-loop control $u_j^\nu(\cdot) : [\tau_\nu,T) \to [u_j^{\min},u_j^{\max}]$ that will be used until the next jump $\tau_{\nu+1}$ occurs;
- This information is called S-adapted as the control of each player is progressively adapted to the sample path $\xi_t(\omega,\cdot)$.

Denote by Γ_j the set of the S-adapted strategies for player j. Let $\gamma_j(\xi_t(\cdot),t) : t \in [0,T]$ be the control used at time t by player j according to the S-adapted strategy γ_j. The strategic J-tuple is denoted $\gamma \in \Gamma$. The profit functions are then defined by

$$V_j(x^0;\gamma_1,\ldots,\gamma_J) = \mathrm{E}_\gamma\left[\int_0^T e^{-\rho_j t} L_j^{\xi(t)}(x_1(t),\ldots,x_J(t),u_j(t))\,dt\right], \tag{18.1}$$

where $u_j(t) = \gamma_j(\xi_t(\cdot),t)$ and $x_j(t)$ is the resulting state-trajectory.

18.3 Comparison between S-adapted and POL information structures

In the POL information structure studied and used e.g. in [59] and [65] the players observe, at each jump time τ_ν, the discrete state ξ_ν and the continuous state $x^\nu = (x_1(\tau_\nu),\ldots,x_J(\tau_\nu))$ and then choose open-loop controls to be implemented until the next jump occurs. Therefore, in the POL information structure, the players are confronted with a sequential game where decisions are made at each jump time with full state information.

In the S-adapted information structure the players cannot observe the current continuous state $x(t)$, even at jump time. We can immediately notice the following

Remark 1. *In the POL information structure, with perfect recall, one could as well assume that the players know, at any jump time τ_ν, the whole $(\xi(\cdot), x(\cdot))$-history $H_\nu = \{\tau_0, \xi_0, x^0, \tau_1, \xi_1, x^1 \ldots, \tau_\nu, \xi_\nu, x^\nu\}$. Therefore, clearly, the S-adapted information structure corresponds to a coarser information structure than POL with perfect recall.*

It appears that the S-adapted information structure is close to the open-loop information structure whereas the POL one is closer to the feedback information structure. It is appropriate to recall here the definition of the concepts of *time consistency* and *subgame perfectness at jump times.*

Definition 6. *An equilibrium is time consistent, if, at any time θ, given that the equilibrium has been played up to that time and given the state $s^*(\theta) = (\xi(\theta), x^*(\theta))$ that has been reached, the same strategies would remain an equilibrium if one restarts the game at that time θ with initial state s^*.*

Definition 7. *An equilibrium is subgame perfect at jump times, if, at any jump time τ_ν, whatever has been played up to that time, given the state $s_\nu = (\xi_\nu, x^\nu)$ reached, the same strategies would remain an equilibrium if one restarts the game at that time τ_ν, with initial state s_ν.*

Now the difference between the two information structures can be seen in the following remarks.

Remark 2. *A Markov strategy in the POL information structure uses only the information available at τ_ν to update the control on the next random time interval $[\tau_\nu, \tau_{\nu+1})$. An equilibrium based on Markov strategy will indeed be time consistent and subgame perfect at jump times.*

The time consistency and subgame perfectness at jump times is a direct consequence of the characterization of the equilibrium via a discrete event dynamic programming equation (see [65] and [59]).

Remark 3. *In an open-loop information structure, for a deterministic game, each player knows the initial state x^0 and chooses a control $u_j(\cdot) : [0,T] \to [u_j^{\min}, u_j^{\max}]$. Given a J-tuple $u(\cdot) = (u_j(\cdot))_{j\in\mathcal{J}}$, the players know at each instant of time the state $x(t)$ that has been reached. In an open-loop equilibrium $u^*(\cdot) = (u_j^*(\cdot))_{j\in\mathcal{J}}$, the trajectory generated $x^*(\cdot)$ is therefore known to each player. Time consistency is observed along $x^*(\cdot)$.*

Remark 4. *In the S-adapted information structure each player knows x^0 and chooses a control adapted to the history of the $\xi(\cdot)$ process. Given a J-tuple of $\xi(\cdot)$-adapted controls and the history h_ν, where $\tau_\nu = \sup\{\tau_l : \tau_l \leq t\}$, each player is able to find out what is the current state $x(t)$. In an S-adapted equilibrium, there is a set of possible trajectories, indexed over the sample space Ω of the $\xi(\cdot)$-process. Call $x^*_\omega(\cdot)$ the equilibrium trajectory associated with the sample value $\omega \in \Omega$. For any $\omega \in \Omega$, $t \in [0,T]$, the S-adapted equilibrium strategy will still be an equilibrium for the game starting at time t, with initial state $s^*_\omega(t) = (\xi_\omega(t), x^*_\omega(t))$. let us call $X^*_\Omega(t) = \{x^*_\omega(t) : \omega \in \Omega\}$ the reachable set at time t, associated with the S-adapted equilibrium strategy J-tuple and the set of all possible sample paths of the ξ-process. Hence we propose as conjecture that the S-adapted equilibrium is subgame perfect at any jump time τ_ν for any x^ν contained in the set $X^*_\Omega(\tau_\nu)$.*

The time consistency and subgame perfectness at jump time on the reachable set of S-adapted equilibria is a consequence of the dynamic optimality conditions.

It would be interesting to compare the set $X^*_\Omega(t)$ with the reachable set $\hat{X}^*_\Omega(t)$ associated at time t with a POL equilibrium. We conjecture that they may coincide for many games, in particular the oligopoly game considered in the present dissertation. If these sets coincide and if they are locally convex then the POL and the S-adapted information structure will yield the same equilibrium values. A more precise study of this conjecture should be the object of further investigation.

18.4 A discrete time stochastic game approximation

We shall again proceed to an approximation of the S-adapted equilibrium of the continuous time game through a sequence of variational inequality solutions. Each variational inequality corresponds to the equilibrium of an associated approximating game of order K. We proceed as follows:

- We discretize time. Let $t_k = \delta \cdot k$ with $k = 0 \ldots K$ and $\delta = T/K$.
- We consider the discrete-time discrete-state Markov chain $\tilde{\xi}^K(k)$ with state set $\mathcal{I}$ and transitions probabilities:

$$P[\tilde{\xi}^K(k+1) = h | \tilde{\xi}^K(k) = i] = \begin{cases} e^{q_{ii}\delta} & \text{if } i = h \\ (1 - e^{q_{ii}\delta})\frac{-q_{ih}}{q_{ii}} & \text{otherwise .} \end{cases}$$

- We represent the set of all sample paths of the Markov chain $\tilde{\xi}^K(k)$ as an event tree. Let $\mathcal{N} = \{1, 2, \ldots, N\}$ be the set of the nodes of this tree and $\mathcal{N}(k)$ the set of nodes associated with period k. In this representation, each node n_k at period k corresponds to a whole history of the Markov chain from period 1 to period k. A complete path along the event tree is also often called a *scenario*. Let $\mathcal{A}(n_k)$ denote the unique predecessor of n_k along the unique path going from n_1 to n_k; let $\mathcal{S}(n_k)$ denote the set of nodes $n_{k+1} \in \mathcal{N}(k+1)$ that can be successors of n_k along a sample path; denote also $\mathcal{U}(n_k)$ the set of all the upstream nodes w.r.t. n_k (including n_k) and $\mathcal{D}(n_k)$ the set of downstream nodes w.r.t. n_k (including n_k), respectively. The number of nodes in the event tree is given by $N = \frac{I^K - 1}{I-1}$.

- We index the state and control variables over the set of nodes of the event tree.

$$\mathbf{x}_j^K = (x_j^K(n_k))_{n_k \in \mathcal{N}},$$

$$\mathbf{u}_j^K = (u_j^K(n_k))_{n_k \in \mathcal{N}},$$

and we introduce the state equations

$$\begin{aligned} x_j^K(n_k) &= u_j^K(n_k)\frac{1-e^{-\mu_j \cdot \delta}}{\mu_j} + e^{-\mu_j \cdot \delta} x_j^K(\mathcal{A}(n_k)) & (18.2) \\ &= x_j^0 e^{-\mu_j t_k} + \frac{1-e^{-\mu_j \cdot \delta}}{\mu_j} \sum_{l=1}^{k} u_j^K(\mathcal{A}^{k-l}(n_k)) e^{-\mu_j (t_k - t_l)}, & (18.3) \end{aligned}$$

where $\mathcal{A}^{k-l}$ means $\mathcal{A}$ to the power $k-l$, *i.e.* the $k-l$ step predecessor. In the S-adapted information structure a strategy for player j in the approximating game of order K is a vector $\mathbf{u}_j^K \in [u_j^{\min}, u_j^{\max}]^N$ such that the generated trajectory $\mathbf{x}_j^K$ remains positive. Let U_j^K be the set of such strategies and X_j^K the set of the corresponding trajectories.

Associated with a strategy J-tuple we have the payoff functions[1]

$$V_j^K(x^0; \mathbf{u}_1^K, \dots, \mathbf{u}_J^K) = \sum_{k=1}^{K} \sum_{n_k \in \mathcal{N}(k)} L_j^{n_k}(x^K(n_k), u_j^K(n_k)) p(n_k) e^{-\rho_j t_k}, \quad (18.4)$$

where $p(n_k)$ denotes the probability of the node $n(k)$ and is given by

$$p(n_k) = \begin{cases} 1 & \text{if } k = 1 \\ p(\mathcal{A}(n_k)) \mathrm{Prob}(\tilde{\xi}^K(n_k) | \tilde{\xi}^K(\mathcal{A}(n_k))) & \text{otherwise}, \end{cases}$$

while $x^K(n_k)$ is the state reached at period k, given the history summarized by n_k and the controls $u_1, \dots, u_J$ as indicated in (18.3).

[1] We use the notation $L_j^{n_k}$ in place of the more correct notation $L_j^{i(n_k)}$.

We have thus defined a game in normal form with strategies in an Euclidean space. An equilibrium $\mathbf{u}^{K*} = (\mathbf{u}_1^{K*}, \ldots, \mathbf{u}_J^{K*})$ is a solution of the following variational inequality

$$\langle F(\mathbf{u}^{K*}), \mathbf{u}^{K*} - \mathbf{u}^K \rangle \geq 0 \quad \forall \mathbf{u}^K \in U_1^K \times \cdots \times U_J^K = U^K. \tag{18.5}$$

With

$$\mathbf{u}^K = \begin{pmatrix} \mathbf{u}_1^K \\ \vdots \\ \mathbf{u}_J^K \end{pmatrix},$$

$$F(\mathbf{u}^K) = \begin{pmatrix} \nabla_{\mathbf{u}_1^K} V_1^K(x^0; \mathbf{u}_1^K, \ldots, \mathbf{u}_J^K) \\ \vdots \\ \nabla_{\mathbf{u}_J^K} V_J^K(x^0; \mathbf{u}_1^K, \ldots, \mathbf{u}_J^K) \end{pmatrix}.$$

The partial derivatives of the reduced profit functions can be calculated from (18.4) once one expresses

$$\frac{\partial x_j^K(n_k)}{\partial u_j^K(n_l)} = \begin{cases} \frac{1-e^{-\mu_j \cdot \delta}}{\mu_j} e^{-\mu_j(t_k - t_l)} & \text{if } n_l \in \mathcal{U}(n_k) \\ 0 & \text{otherwise .} \end{cases}$$

Theorem 16. *Assume the total reward function $\sum_j L_j^i(x_1, \ldots, x_J, u_j)$ is strictly diagonally concave in (x, u) for all i in $\mathcal{I}$ then the operator $-F(\mathbf{u}^K)$ is monotone.*

Proof. By direct verification. □

Theorem 17. *Under the assumptions of Theorem 16, there exists a unique equilibrium for the approximating game of order K.*

Proof. As $-F$ is monotone, we have that $V_j^K(\mathbf{u}_1^K, \ldots, \mathbf{u}_J^K)$ is concave in $\mathbf{u}_j^K$. Moreover the set $U = U_1 \times \cdots \times U_J$ is convex, so the game is a concave game and according to Theorem 1 in [104] we know that there exists an equilibrium. From the monotony of $-F$ and Theorem 2 in [104] we know that the equilibrium is unique. □

For the approximation results that we shall establish, it is convenient to have a property of continuity of the strategies w.r.t. sample paths. This requires first a definition of a distance on the space of sample paths of the $\xi(\cdot)$ process. Let Λ be the class of strictly increasing, continuous mapping of $[0, T]$ into itself.

Definition 8. *Let $\xi(\omega_1, \cdot)$ and $\xi(\omega_2, \cdot)$ be two sample paths of the continuous time Markov chain. Define the distance*[2] *$d(\xi(\omega_1, \cdot), \xi(\omega_2, \cdot))$ as the infimum of those positive D for which there exists a λ in Λ such that* $\sup_t |\lambda(t) - t| \leq D$ *and* $\sup_t |\xi(\omega_1, t) - \xi(\omega_2, \lambda(t))| \leq D$.

[2] Also called Skorohod topology; see [14].

Assumption 6. *We suppose the following*

- *state and control separation:* $L^i_j(x,u_j) = \mathcal{L}^i_j(x) + \mathcal{M}^i_j(u_j)$ *with* $\mathcal{L}^i_j$ *and* $\mathcal{M}^i_j$ *Lipschitz continuous,*
- *monotony:* $\sum_j \mathcal{L}^i_j(x)$ *is strictly diagonally concave in* x *and* $\mathcal{M}^i_j(u_j)$ *strictly concave in* u_j *for all* j *in* $\mathcal{J}$ *and* i *in* $\mathcal{I}$.
- *stability: the admissible strategies* γ *verify, except perhaps on a null measure set, the following property: for all* ϵ *and for all realizations* $\xi(\omega_1,\cdot)$ *and* $\xi(\omega_2,\cdot)$ *there exists a* η *such that*[3] $d(\xi(\omega_1,\cdot),\xi(\omega_2,\cdot)) < \eta \Rightarrow ||\gamma(\omega_1,\cdot) - \gamma(\omega_2,\cdot)|| < \epsilon$ *where the norm is the* L_1*-norm:* $||v|| = \int_0^T |v(t)|\,dt$.

The last assumption says that the control must be continuous with respect to the random disturbance trajectory; it insures a stability of the control with respect to the random process.

In order to establish convergence results it is convenient to modify the representation of the approximating game and of the continuous time game in order to explicit the dependence on the sample paths. In the approximating game of order K, the control, previously denoted $u^K_j(n(k))$, will now be written $u^K_j(\tilde{\omega}^K,k)$, where $\tilde{\omega}^K$ indexes the sample paths for the discrete time Markov chain and k is the current period. The same representation is used for the state variables.

Let $\xi(\omega,\cdot)$ be a sample path of the continuous-time Markov chain and $\tilde{\xi}^K(\tilde{\omega}^K,\cdot) = (\tilde{\xi}^K(\tilde{\omega}^K,1),\ldots\tilde{\xi}^K(\tilde{\omega}^K,K))$ be a sample path of the discrete-time Markov chain. We define the projection $\pi_K : \{\xi(\omega,\cdot) : [0,T] \to \mathcal{I}\} \to \mathcal{I}^K$ as:

$$\pi_K(\xi(\omega,\cdot)) = (\xi(\omega,t_0),\ldots,\xi(\omega,t_{K-1})). \tag{18.6}$$

We define, as in the deterministic case, $\phi^K_j : U^K_j \to U_j$ and $\sigma^K_j : U_j \to U^K_j$ by

$$\phi^K_j(\mathbf{u}^K_j)(\omega,t) = u^K_j(\tilde{\omega}^K,k)$$

where k is such that $t_k = \min_s\{t_s|t_s \geq t\}$ and $\tilde{\xi}^K(\tilde{\omega}^K,\cdot) = \pi_K(\xi(\omega,\cdot))$

and

$$\sigma^K_j(\gamma_j)(\tilde{\omega}^K,k) = \frac{\mu}{1-e^{-\mu\delta}}\int_{t_{k-1}}^{t_k} \gamma_j(\omega,s)e^{-\mu(t_k-s)}\,ds$$

where $\xi(\omega,t) = \tilde{\xi}^K(\tilde{\omega}^K,k)$ and $k = \arg\min_s\{t_s|t_s \geq t\}$, respectively.

Let $x^K_{j,\sigma^K_j(\gamma)}(\tilde{\omega}^K,\cdot)$ be the discrete time trajectory generated by $\sigma^K_j(u_j)(\tilde{\omega}^K,\cdot)$ and $x_{j,\phi^K_j(\mathbf{u}^K_j)}(\omega,\cdot)$ the continuous time trajectory generated by $\phi^K_j(\mathbf{u}^K_j)(\omega,\cdot)$. One can verify that we have again the property $\sigma^K_j \circ \phi^K_j = \mathbb{1}$. The mappings preserve the property of non-negativity of the generated trajectory and the non-anticipativity with respect to the S-adapted information structure.

[3]We prefer the shorter notation $\gamma(\omega,\cdot)$ to the notation $\gamma(\xi_t(\omega,\cdot),\cdot)$.

The convergence of the discrete time oligopoly toward the continuous time oligopoly is shown in the two following theorems.

Theorem 18. *Suppose assumption 6 holds. Let $\gamma_1^*, \ldots, \gamma_J^*$ be the equilibrium strategies of the continuous time oligopoly. Then for all positive ϵ there exists K_ϵ such that for all $K > K_\epsilon$ the strategies vector $(\sigma_1^K(\gamma_1^*), \ldots, \sigma_J^K(\gamma_J^*))$ is an ϵ-Nash equilibrium for the discrete time oligopoly of order K.*

Theorem 19. *Suppose assumption 6 holds. Let $\mathbf{u}_1^{K*}, \ldots, \mathbf{u}_J^{K*}$ be the equilibrium controls of the discrete time oligopoly of order K. Then for all j in $\mathcal{J}$, all γ_j in Γ_j and all positive ϵ, there exists $K_\epsilon(\gamma_j)$ such that for all $K > K_\epsilon(\gamma_j)$ we have*

$$V_j(\phi_1^K(\mathbf{u}_1^{K*}), \ldots, \gamma_j, \ldots, \phi_J^K(\mathbf{u}_J^{K*})) \leq V_j(\phi_1^K(\mathbf{u}_1^{K*}), \ldots, \phi_j^K(\mathbf{u}_j^{K*}), \ldots, \phi_J^K(\mathbf{u}_J^{K*})) + \epsilon.$$

Proof. The proofs of these two theorems can be found in the appendix D.1. $\square$

Chapter 19

Numerical experiments

In this chapter we recall the stochastic duopoly model presented in [65] and show that the total reward function is strictly diagonally concave. We then propose to use an algorithm due to Konnov [79] to solve the variational inequalities. Finally the numerical results are presented and discussed.

19.1 A stochastic duopoly model

We take the same duopoly model as in [65]. The depreciation rates are $\mu_1 = 0.08$ and $\mu_2 = 0.06$ respectively. Assume that the firms supply, according to their production capacity, a market characterized by an inverse demand law depending on the market condition i

$$D^i(x_1(t) + x_2(t)) = \frac{a_i}{x_1(t) + x_2(t) + b_i} - c_i.$$

Here D^i is the market clearing price, given the total supply $x_1(t) + x_2(t)$. The continuous-time Markov chain, describing the market condition, takes three different values, corresponding to three different demand functions $D^i(x_1 + x_2)$, $i \in \mathcal{I} = \{1, 2, 3\}$. The coefficient are: $a_1 = 120$, $a_2 = 100$, $a_3 = 80$, $b_i = 20$, $i = 1, 2, 3$, $c_1 = 3$, $c_2 = 2.5$, $c_3 = 2$. The dynamics of the continuous-time Markov chain is described by the following transition rate matrix:

$$Q = \begin{pmatrix} -0.2 & 0.2 & 0.0 \\ 0.01 & -0.05 & 0.04 \\ 0.1 & 0.0 & -0.1 \end{pmatrix}.$$

We also assume that each firm has a quadratic maintenance and investment cost. So the reward functions are given by

$$L_j^i(x_1, x_2, u_j) = D^i(x_1 + x_2)x_j - (x_j)^2 - (u_j)^2.$$

The discount rate is $\rho = 0.09$, for both firms. The time horizon is $T = 10$. For the time discretization we take the number of periods $K = 10$. So the number of nodes is given by

$$N = \frac{I^K - 1}{I - 1} = 29524.$$

Applying Theorem 6 from [104], it can be checked that the total reward function $\sum_j L_j(x_1, x_2, u_j)$ is strictly diagonally concave in (x, u) for all i in $\mathcal{I}$. So according to Theorems 16 and 17 the function $-F(\mathbf{u}^K)$ is monotone and there exists a unique equilibrium for the approximating game.

19.2 Implementation

If $-F(\mathbf{u})$ is a monotone function that is Lipschitz with constant L the solution $\mathbf{u}^*$ of the following variational inequality

$$\langle F(\mathbf{u}^*), \mathbf{u}^* - \mathbf{u}\rangle \geq 0 \quad \forall \mathbf{u} \in U \subseteq R^{2N}$$

can be obtained via the following algorithm given by Konnov [79]:

Step 0 (initialization)
Choose λ, $0 < \theta < 1/(\lambda(1 + L))$, ε and $\mathbf{u}(0)$.
Set $n = 0$.

Step 1 (computation of the next point and stopping criterion)
$p(n+1) = \mathrm{proj}_U(\mathbf{u}(n) + \lambda F(\mathbf{u}(n)))$
$\mathbf{v}(n+1) = \mathbf{u}(n) + \theta(p(n) - \mathbf{u}(n))$

IF $F(\mathbf{v}(n+1)) = 0$ THEN STOP $\mathbf{u}^* = \mathbf{v}(n+1)$

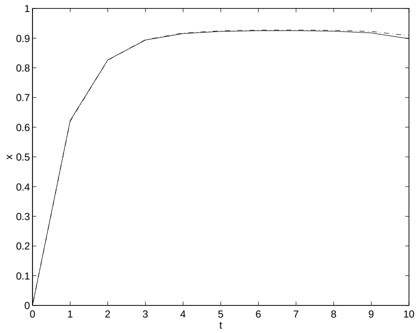
$\alpha = \langle F(\mathbf{v}(n+1)), \mathbf{v}(n+1) - \mathbf{u}(n)\rangle / ||F(\mathbf{v}(n+1))||$
$\mathbf{u}(n+1) = \mathrm{proj}_U(\mathbf{u}(n)) + \alpha F(\mathbf{v}(n+1))$

Evaluate $g(\mathbf{u}(n+1)) = \min_{\mathbf{u}\in U}\langle F(\mathbf{u}(n+1)), \mathbf{u}(n+1) - \mathbf{u}\rangle/(2N)$
IF $g(\mathbf{u}(n+1)) \geq -\varepsilon$ THEN STOP $\mathbf{u}^* = \mathbf{u}(n+1)$
ELSE Increment n
GO TO Step 1 .

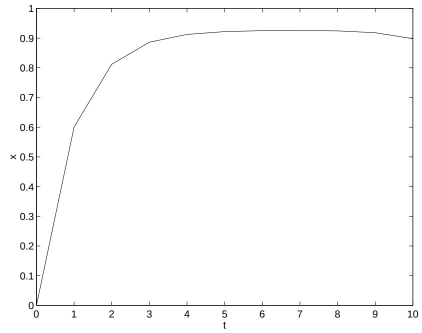
We implemented this algorithm and computed the S-adapted open-loop equilibrium of the stochastic duopoly taking as stopping criterion $\varepsilon = 0.001$.

19.3 Numerical results

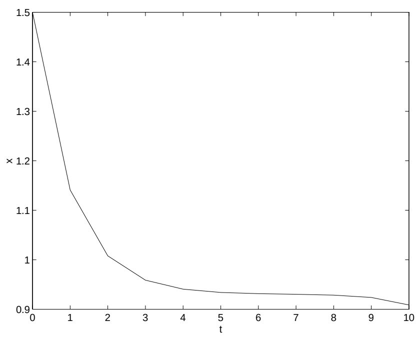
All the equilibrium trajectories illustrated in this report are for the state $i = 2$. They are similar for the other states. Figure 19.1(a) shows the equilibrium



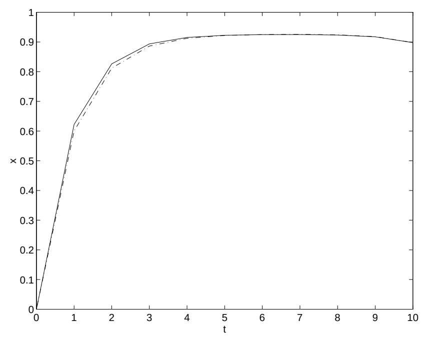
(a) $x_1(t)$ and $x_2(t)$ (dotted) for $x_1(0) = x_2(0) = 0$.



(b) $x_1(t)$ for $x_1(0) = 0$, $x_2(0) = 1.5$.



(c) $x_2(t)$ for $x_1(0) = 0$, $x_2(0) = 1.5$.



(d) $x_1(t)$ for $x_1(0) = 0$, $x_2(0) = 0$ and for $x_1(0) = 0$, $x_2(0) = 1.5$ (dotted).

Figure 19.1: Equilibrium trajectories in state i=2.

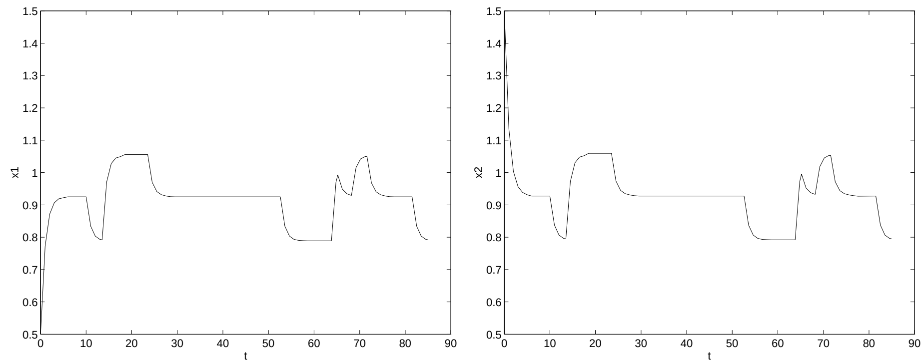
trajectories for the initial state $x_1(0) = 0$ and $x_2(0) = 0$. As the model is nearly symmetric (the only difference between player 1 and player 2 is the depreciation rate) it is not surprising to have the equilibrium trajectories nearly similar for the two players. Figures 19.1(b) and 19.1(c) show the equilibrium trajectories for the initial state $x_1(0) = 0$ and $x_2(0) = 1.5$. We see that they are consistent with the results in [65]. Figure 19.1(d) compares the equilibrium trajectories for player 1 when player 2 has two different initial states ($x_2(0) = 0$ and $x_2(0) = 1.5$). We see that the the turnpikes[1] are identical and, as expected, the first trajectory lies above the second one.

In Table 19.1 we recall the value of the turnpikes computed under the POL information structure and the corresponding turnpike values obtained with the

[1]See [65] for the definition.

Information	$\tilde{x}_1^1$	$\tilde{x}_2^1$	$\tilde{x}_1^2$	$\tilde{x}_2^2$	$\tilde{x}_1^3$	$\tilde{x}_2^3$
POL	1.04644	1.05067	0.92472	0.92854	0.80014	0.80353
S-adapted	1.055671	1.059330	0.924880	0.927263	0.788885	0.792057

Table 19.1: Turnpikes



(a) Player 1 (b) Player 2

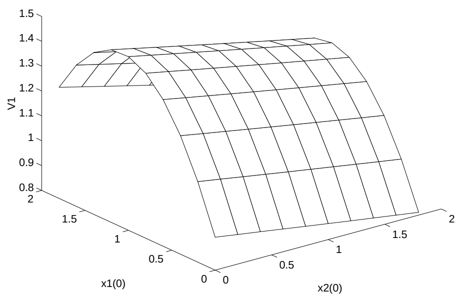
Figure 19.2: Equilibrium trajectories for a given realization of the random process.

S-adapted information structure. We can observe that these values are very close.

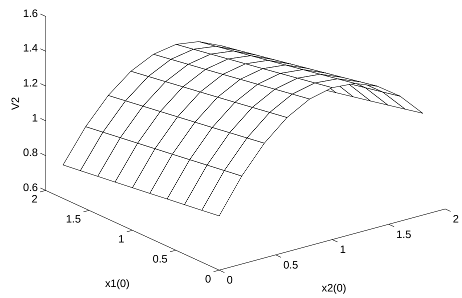
As in [65] we consider the following realization of the random process: the sequence of the modal change is 2,3,1,2,3,1,2,1,2,3 and the jump times are 10, 13.5, 23.5, 52.6, 63.8, 65.2, 68.2, 71.6, 81.5. For this realization of the random process the equilibrium trajectories are displayed in Figure 19.2. One sees that after each jump the equilibrium trajectories are attracted by the turnpikes associated with the current state i and remain close to them until the next jump occurs.

The value functions in state $i = 2$ for different initial states $x_i(0)$ are pictured in Figure 19.3.

We can see that the results for S-adapted information structure are really close to the ones for the POL information structure.



(a) Player 1



(b) Player 2

Figure 19.3: Value functions in state $i = 2$ for different initial states $x_i(0)$

Chapter 20

Concluding remarks for Part III

In this part of the thesis we have proposed to use a numerical technique based on the solution of an approximating variational inequality to compute a continuous time Nash-Cournot ϵ-equilibrium. We have considered first the open-loop information structure for a deterministic differential game model and then the S-adapted information structure when the dynamic system is subject to jump Markov random disturbances. We have compared the S-adapted information structure with the POL information structure introduced in [59]-[65] and conjectured that, when the jump Markov process is uncontrolled, the two types of equilibria are likely to yield similar outcomes. We have verified this conjecture on the model already considered in [65], with the POL information structure. Finally we have proved the proximity existing between the equilibrium strategies of the approximating game and those of the continuous time game.

In all these developments we have used the dynamic oligopoly formalism of [17],[23] and [65]. Future extensions of this work should deal with a more general class of differential games and explore in greater detail the proximity of equilibria under POL and S-adapted information structures.

Appendix D

Convergence of the stochastic programming approach

D.1 Convergence results for the stochastic case

D.1.1 Proof outline

To simplify the notations the proofs are done for a duopoly; they can be easily extended to the case of an oligopoly. We shall state a list of Propositions to be proved in chapter D.1.2.

Proposition 2. *For any $\epsilon > 0$, there exists K_ϵ such that, for any approximating game of order $K > K_\epsilon$ the following holds*

$$\forall (\mathbf{u}_1^K, \mathbf{u}_2^K) \in U_1^K \times U_2^K \quad |V_j^K(\mathbf{u}_1^K, \mathbf{u}_2^K) - V_j(\phi_1^K(\mathbf{u}_1^K), \phi_2^K(\mathbf{u}_2^K))| < \epsilon.$$

Proposition 3. *Given γ_1 and γ_2, for any $\epsilon > 0$, there exists K_ϵ such that, for any approximating game of order $K > K_\epsilon$ the following holds*

$$|V_j^K(\sigma_1^K(\gamma_1), \sigma_2^K(\gamma_2)) - V_j(\gamma_1, \gamma_2)| < \epsilon.$$

Proposition 4. *Given γ_2, for any $\epsilon > 0$, there exists K_ϵ such that, for any approximating game of order $K > K_\epsilon$ the following holds*

$$\forall \mathbf{u}_1^K \in U_1^K \quad |V_1^K(\mathbf{u}_1^K, \sigma_2^K(\gamma_2)) - V_1(\phi_1^K(\mathbf{u}_1^K), \gamma_2)| < \epsilon$$

Proposition 5. *Given γ_1, for any $\epsilon > 0$, there exists K_ϵ such that, for any approximating game of order $K > K_\epsilon$ the following holds*

$$\forall \mathbf{u}_2^K \in U_2^K \quad |V_1^K(\sigma_1^K(\gamma_1), \mathbf{u}_2^K) - V_1(\gamma_1, \phi_2^K(\mathbf{u}_2^K))| < \epsilon$$

Proof of Theorem 18. For a game of order K, consider $V_1^K(\mathbf{u}_1^K, \sigma_2^K(\gamma_2^*))$ for any admissible $\mathbf{u}_1^K$. If K is large enough ($K > K_\epsilon^1$), according to Proposition 4 one has

$$V_1^K(\mathbf{u}_1^K, \sigma_2^K(\gamma_2^*)) \leq V_1(\phi_1^K(\mathbf{u}_1^K), \gamma_2^*) + \frac{\epsilon}{2}$$

and by the equilibrium property of γ^*

$$V_1(\phi_1^K(\mathbf{u}_1^K), \gamma_2^*) \leq V_1(\gamma_1^*, \gamma_2^*).$$

Finally, according to Proposition 3, we have

$$V_1(\gamma_1^*, \gamma_2^*) \leq V_1^K(\sigma_1^K(\gamma_1^*), \sigma_2^K(\gamma_2^*)) + \frac{\epsilon}{2}$$

if K is large enough ($K > K_\epsilon^2$). Therefore if $K > K_\epsilon = \sup\{K_\epsilon^1, K_\epsilon^2\}$ one has

$$V_1^K(\mathbf{u}_1^K, \sigma_2^K(\gamma_2^*)) \leq V_1^K(\sigma_1^K(\gamma_1^*), \sigma_2^K(\gamma_2^*)) + \epsilon.$$

The same property holds for Player 2. The proof is complete. □

Proof of Theorem 19. Let $(\mathbf{u}_1^{K*}, \mathbf{u}_2^{K*})$ be an S-adapted equilibrium of the approximating game of order K. Consider, for any admissible policy γ_1 of the continuous time game, the payoff $V_1(\gamma_1, \phi_2^K(\mathbf{u}_2^{K*}))$. According to Proposition 5, if K is large enough ($K > K_\epsilon^1$), one has

$$V_1(\gamma_1, \phi_2^K(\mathbf{u}_2^{K*})) \leq V_1^K(\sigma_1^K(\gamma_1), \mathbf{u}_2^{K*}) + \frac{\epsilon}{2}.$$

Since $(\mathbf{u}_1^{K*}, \mathbf{u}_2^{K*})$ is an S-adapted equilibrium pair for the approximating game of order K, one has also

$$V_1^K(\sigma_1^K(\gamma_1), \mathbf{u}_2^{K*}) \leq V_1^K(\mathbf{u}_1^{K*}, \mathbf{u}_2^{K*}).$$

Finally, according to Proposition 2 one has

$$V_1^K(\mathbf{u}_1^{K*}, \mathbf{u}_2^{K*}) \leq V(\phi_1^K(\mathbf{u}_1^{K*}), \phi_2^K(\mathbf{u}_2^{K*})) + \frac{\epsilon}{2}$$

if K is large enough ($K > K_\epsilon^2$). Therefore if $K > K_\epsilon = \sup\{K_\epsilon^1, K_\epsilon^2\}$ one has

$$V_1(\gamma_1, \phi_2^K(\mathbf{u}_2^{K*})) \leq V(\phi_1^K(\mathbf{u}_1^{K*}), \phi_2^K(\mathbf{u}_2^{K*})) + \epsilon.$$

The same property holds for the other player. □

D.1.2 Proof of the Propositions

We will first prove the following lemmas which will be used later on.

Lemma 3. *Suppose assumption 6 holds. Then* $\forall t \in [0,T]$, $\forall x, y \in X_1 \times \cdots \times X_J$, $\forall u_j, v_j \in U_j$ *and* $\forall i, l \in I$, *we have*

$$|L_j^i(x(t), u_j(t)) - L_j^l(y(t), v_j(t))| \leq C$$

Proof. This comes from the Lipschitz property L_j^i and the boundness of the controls and state trajectories. ☐

Let $(\Omega, \mathcal{B}, P)$ be the probability space associated with the continuous time stochastic process and $(\tilde{\Omega}^K, \tilde{P}_K)$ the probability space associated with the discrete time stochastic process of order K. Let $\pi_K^{-1}(\tilde{\xi}^K(\tilde{\omega}^K, \cdot)) = \{\xi(\omega, \cdot) : \pi_K(\xi(\omega, \cdot)) = (\tilde{\xi}^K(\tilde{\omega}^K, 1), \ldots, \tilde{\xi}^K(\tilde{\omega}^K, K))\}$ with π_K defined as in Eq.(18.6). Notice that the two Markov chain $\tilde{\xi}^K(\cdot)$ and $\pi_K(\xi(\cdot))$ associated with the projection π_K have the same sample path set. The convergence of the first Markov chain to the second one is stated in the following Lemma.

Lemma 4. $\tilde{P}_K(\tilde{\xi}^K(\tilde{\omega}^K, \cdot))$*tends to* $P(\pi_K^{-1}(\tilde{\xi}^K(\tilde{\omega}^K, \cdot)))$, *uniformly in* $\tilde{\omega}^K$, *as* K *tends to infinity.*

Proof. As before let $\delta = T/K$. Denote by S^K the transition probability matrix $[\text{Prob}(\xi(t+\delta) = h|\xi(t) = i)] \quad i, h \in \mathcal{I}$ and by $\tilde{S}^K$ the transition probability matrix $[\text{Prob}(\tilde{\xi}^K(k+1) = h|\tilde{\xi}^K(k) = i)] \quad i, h \in \mathcal{I}$. The only thing to prove is that $\tilde{S}^K$, the generator of the Markov chain $\tilde{\xi}^K(\cdot)$, tends to S^K, the generator of the Markov chain $\pi_K(\xi(\cdot))$, as K tends to infinity.

Define $||M||$ the norm of an $I \times I$ matrix as follows:

$$||M|| = \sup_i \sum_l |m_{il}|.$$

Applying the Taylor expansion we have:

$$\tilde{S}^K = \mathbf{1} + Q\delta + \tilde{R}$$

With $|\tilde{r}_{il}| \leq \delta^2 |q_{ii}| q_{il}/2$ which implies

$$||\tilde{R}|| \leq \frac{||Q||^2 \delta^2}{2}.$$

From [39] p.149, we have the following:

$$S^K = e^{Q\delta} = \mathbf{1} + Q\delta + R.$$

With

$$||R|| \leq \frac{||Q||^2\delta^2}{2}.$$

So we have

$$|\tilde{s}_{ij}^K(\delta) - s_{ij}^K(\delta)| \leq ||Q||^2\delta^2.$$

□

To prove Proposition 2 we will need the following Lemma:

Lemma 5. *For any $\tilde{\omega}^K \in \tilde{\Omega}^K$ and any $\omega \in \Omega$ such that $\pi_K(\xi(\omega,\cdot)) = \tilde{\xi}^K(\tilde{\omega}^K,\cdot)$ we have*

$$\begin{aligned}
&|\sum_{k=1}^{K} L_j^{\tilde{\xi}^K(\tilde{\omega}^K,k)}(x_1^K(\tilde{\omega}^K,k), x_2^K(\tilde{\omega}^K,k), u_j^K(\tilde{\omega}^K,k))e^{-\rho\delta k}\delta \\
&-\sum_{k=1}^{K}\int_{t_{k-1}}^{t_k} L_j^{\tilde{\xi}^K(\tilde{\omega}^K,k)}(x_{1,\phi_1^K(\mathbf{u}_1^K)}(\omega,t), x_{2,\phi_2^K(\mathbf{u}_2^K)}(\omega,t), \phi_j^K(\mathbf{u}_j^K)(\omega,t))e^{-\rho t}\,dt|
\end{aligned}$$

tends uniformly in $\mathbf{u}_j^K$ to zero as K tends to infinity. Recall that $x_{j,\phi_j^K(\mathbf{u}_j^K)}(\omega,\cdot)$ is the continuous time trajectory generated by $\phi_j^K(\mathbf{u}_j^K)(\omega,\cdot)$ as introduced in Section 18.

Proof. Let $\tilde{\omega}^K \in \tilde{\Omega}^K$ and $\omega \in \Omega$ such that $\pi_K(\xi(\omega,\cdot) = \tilde{\xi}^K(\tilde{\omega}^K,\cdot)$. Let us first check that the term in $\mathcal{M}_j^i$ has a uniform convergence in $\mathbf{u}_j^K$:

$$\begin{aligned}
&|\sum_{k=1}^{K} \mathcal{M}_j^{\tilde{\xi}^K(\tilde{\omega}^K,k)}(u_j^K(\tilde{\omega}^K,k))e^{-\rho\delta k}\delta - \sum_{k=1}^{K}\int_{t_{k-1}}^{t_k} \mathcal{M}_j^{\tilde{\xi}^K(\tilde{\omega}^K,k)}(\phi_j^K(\mathbf{u}_j^K)(\omega,t))e^{-\rho t}\,dt| \\
=&|\sum_{k=1}^{K} \mathcal{M}_j^{\tilde{\xi}^K(\tilde{\omega}^K,k)}(u_j^K(\tilde{\omega}^K,k))e^{-\rho\delta k}\delta[1 - \frac{e^{\rho\delta}-1}{\rho\delta}]| \\
\leq& T\cdot C\cdot[1 - \frac{e^{\rho\delta}-1}{\rho\delta}] = T\cdot C\cdot[1 - \frac{e^{\rho\frac{T}{K}}-1}{\rho\frac{T}{K}}].
\end{aligned}$$

This tends to zero (independently of $\mathbf{u}_j^K$) as K tends to infinity.

The convergence of the term in $\mathcal{L}_j^i$ comes from the properties of the x-trajectories. We defined in (17.6) the dynamics of the system in such a way that for $k = 0\ldots K$ $x_{j,\sigma_j^K(\gamma)}^K(\omega,t_k)$ equals $x_j^K(\tilde{\omega}^K,k)$. As the controls are bounded, the trajectories are Lipschitz and so the convergence is uniform in $\mathbf{u}_j^K$. □

Proof of Proposition 2.

$$
\begin{aligned}
&|V_j^K(\mathbf{u}_1^K,\mathbf{u}_2^K) - V_j(\phi_1^K(\mathbf{u}_1^K),\phi_2^K(\mathbf{u}_2^K))| \\
=&|\sum_{\tilde\omega^K} \tilde P_K(\tilde\xi^K(\tilde\omega^K,\cdot))\sum_{k=1}^K L_j^{\tilde\xi^K(\tilde\omega^K,k)}(x_1^K(\tilde\omega^K,k),x_2^K(\tilde\omega^K,k),u_j^K(\tilde\omega^K,k))e^{-\rho\delta k}\delta \\
&- E_\xi[\int_0^T L_j^{\xi(t)}(x_{1,\phi_1^K(\mathbf{u}_1^K)}(\omega,t),x_{2,\phi_2^K(\mathbf{u}_2^K)}(\omega,t),\phi_j^K(\mathbf{u}_j^K)(\omega,t)))e^{-\rho t}\,dt]|.
\end{aligned}
\tag{D.1}
$$

Let F_l be the set of all realizations $\xi(\omega,\cdot)$ that have more than l jumps over $[0,T]$. We set $l=\lfloor\sqrt{K}\rfloor$, the largest integer smaller than the squared root of K. The second term can be rewritten as:

$$
\begin{aligned}
&E_{\xi|\xi(\omega,\cdot)\in F_l}[\sum_{k=1}^K\int_{t_{k-1}}^{t_k}\{L_j^{\xi(t_{k-1})}(x_{1,\phi_1^K(\mathbf{u}_1^K)}(\omega,t),x_{2,\phi_2^K(\mathbf{u}_2^K)}(\omega,t),\phi_j^K(\mathbf{u}_j^K)(\omega,t)) - \\
&\qquad L_j^{\xi(t)}(x_{1,\phi_1^K(\mathbf{u}_1^K)}(\omega,t),x_{2,\phi_2^K(\mathbf{u}_2^K)}(\omega,t),\phi_j^K(\mathbf{u}_j^K)(\omega,t))\}e^{-\rho t}\,dt] + \\
&E_{\xi|\xi(\omega,\cdot)\notin F_l}[\sum_{k=1}^K\int_{t_{k-1}}^{t_k}\{L_j^{\xi(t_{k-1})}(x_{1,\phi_1^K(\mathbf{u}_1^K)}(\omega,t),x_{2,\phi_2^K(\mathbf{u}_2^K)}(\omega,t),\phi_j^K(\mathbf{u}_j^K)(\omega,t)) - \\
&\qquad L_j^{\xi(t)}(x_{1,\phi_1^K(\mathbf{u}_1^K)}(\omega,t),x_{2,\phi_2^K(\mathbf{u}_2^K)}(\omega,t),\phi_j^K(\mathbf{u}_j^K)(\omega,t))\}e^{-\rho t}\,dt] + \\
&E_\xi[\sum_{k=1}^K\int_{t_{k-1}}^{t_k} L_j^{\xi(t_{k-1})}(x_{1,\phi_1^K(\mathbf{u}_1^K)}(\omega,t),x_{2,\phi_2^K(\mathbf{u}_2^K)}(\omega,t),\phi_j^K(\mathbf{u}_j^K)(\omega,t))e^{-\rho t}\,dt].
\end{aligned}
$$

According to Lemma 3 and as $P(F_l)$ tends to zero as l tends to infinity, the norm of the first term is smaller than ϵ for K big enough. For the second term we know that the realizations have less than l jumps. If in the interval $[t_{k-1},t_k[$ we have no jumps, the two terms under the integral are identical. If there is one or more jump in the interval $[t_{k-1},t_k[$, the integral over the interval is, according Lemma 3, less than C/K. As we have less than $l\leq\sqrt{K}$ jumps, the norm of the second term is smaller than $C/\sqrt{K}$ and so smaller than ϵ for K big enough. For the third term we can see that for two realizations ω_1 and ω_2 such that $\pi_K(\xi(\omega_1,\cdot))=\pi_K(\xi(\omega_2,\cdot))$, the integral is the same. With any $\tilde\omega^K$ in Ω^K we associate $\varphi(\tilde\omega^K)$ in Ω such that $\xi(\omega,t)=\tilde\xi^K(\tilde\omega^K,k)$ and $k=\arg\min_s\{t_s|t_s\geq t\}$.

The third term can thus be rewritten as:

$$\sum_{\tilde{\omega}^K} P(\pi_K^{-1}(\tilde{\xi}^K(\tilde{\omega}^K,\cdot)) \sum_{k=1}^{K} \int_{t_{k-1}}^{t_k} L_j^{\tilde{\xi}^K(\tilde{\omega}^K,k)}(x_{1,\phi_1^K(\mathbf{u}_1^K)}(\varphi(\tilde{\omega}^K),t), x_{2,\phi_2^K(\mathbf{u}_2^K)}(\varphi(\tilde{\omega}^K),t),$$

$$\phi_j^K(\mathbf{u}_j^K)(\varphi(\tilde{\omega}^K),t))e^{-\rho t}\,dt.$$

To summarize, for K big enough, Eq. (D.1) becomes:

$$|V_j^K(\mathbf{u}_1^K,\mathbf{u}_2^K) - V_j(\phi_1^K(\mathbf{u}_1^K),\phi_2^K(\mathbf{u}_2^K))| \leq \epsilon + \epsilon +$$

$$|\sum_{\tilde{\omega}^K} \tilde{P}_K(\tilde{\xi}^K(\tilde{\omega}^K,\cdot)) \sum_{k=1}^{K} L_j^{\tilde{\xi}^K(\tilde{\omega}^K,k)}(x_1(\tilde{\xi}^K(\tilde{\omega}^K,\cdot),k), x_2(\tilde{\xi}^K(\tilde{\omega}^K,\cdot),k),$$

$$u_j(\tilde{\xi}^K(\tilde{\omega}^K,\cdot),k))e^{-\rho\delta k}\delta +$$

$$\sum_{\tilde{\omega}^K} P(\pi_K^{-1}(\tilde{\xi}^K(\tilde{\omega}^K,\cdot)) \sum_{k=1}^{K} \int_{t_{k-1}}^{t_k} L_j^{\tilde{\xi}^K(\tilde{\omega}^K,k)}(x_{1,\phi_1^K(\mathbf{u}_1^K)}(\varphi(\tilde{\omega}^K),t), x_{2,\phi_2^K(\mathbf{u}_2^K)}(\varphi(\tilde{\omega}^K),t),$$

$$\phi_j^K(\mathbf{u}_j^K)(\varphi(\tilde{\omega}^K),t))e^{-\rho t}\,dt|.$$

For K big enough, Lemmas 3,4 and 5 asserts that the term in absolute value tends to zero as K tends ton infinity. Thus we have proved that for all $\epsilon > 0$ there exist a K_ϵ (uniform in $\mathbf{u}_1^K$ and $\mathbf{u}_2^K$) such that for all K greater than K_ϵ we have the following

$$|V_j^K(\mathbf{u}_1^K,\mathbf{u}_2^K) - V_j(\phi_1^K(\mathbf{u}_1^K),\phi_2^K(\mathbf{u}_2^K))| \leq 3\epsilon.$$

□

To prove Proposition 3 we will need the following Lemma:

Lemma 6. *For any admissible strategy* γ_j *the norm* $||\phi_j^K \circ \sigma_j^K(\gamma_j) - \gamma_j||$ *tends almost surely to zero as* K *tends to infinity.*

Proof. We have to prove that for any ω in Ω and any admissible strategy γ_j, $||\phi_j^K \circ \sigma_j^K(\gamma_j)(\omega,\cdot) - \gamma_j(\omega,\cdot)||$ tends to zero as K tends to infinity, except perhaps on a set of probability 0.

With any ω in Ω we associate $\psi(\omega,K)$ in Ω as follows: $\xi(\psi(\omega,K),t) = \xi(\omega,t_k)$ where $t_k = \min_s\{t_s : t_s \geq t\}$. We thus obtain an approximation of the sample path $\xi(\omega,\cdot)$ through a step function with jumps at times t_k only. Suppose that $\xi(\omega,t)$ has finitely many jumps, then for all ϵ positive there exists

K_ϵ such that $K > K_\epsilon$ implies $d(\omega, \psi(\omega, K)) < \epsilon$. If assumption 6 holds this implies the following: for all ϵ positive there exists K_ϵ such that $K > K_\epsilon$ implies $||\gamma_j(\psi(\omega, K), \cdot) - \gamma_j(\omega, \cdot)|| < \epsilon$. That means that the control for any realization ω, except perhaps the ones with infinitely many jump (which is a null set), can be approximated as closely as desired by the control for a realization where the jump times are multiples of T/K_ϵ. This means that for fixed ω the proof is complete if one prove that for all γ_j we have $||\phi_j^K \circ \sigma_j^K(\gamma_j)(\psi(\omega, K), \cdot) - \gamma_j(\psi(\omega, K), \cdot)||$ tends to zero as K tends to infinity. To simplify the notation we denote by $w(\cdot)$ the control $\gamma_j(\psi(\omega, K), \cdot)$ associated with the realization $\psi(\omega, K)$.

$$||\phi_j^K \circ \sigma_j^K(w_j)(\cdot) - w_j(\cdot)|| = \sum_{k=1}^{K} \int_{t_{k-1}}^{t_k} \left|\sigma_j^K(w_j)(k) - w_j(t)\right| \, dt.$$

We know that for all ϵ there exists a step function g_j^ϵ with $u_j^{\min} \leq g_j^\epsilon \leq u_j^{\max}$ such that $||g_j^\epsilon - w_j||_1 < \epsilon$

$$\begin{aligned}
&\sum_{k=1}^{K} \int_{t_{k-1}}^{t_k} \left|\sigma_j^K(w_j)(k) - w_j(t)\right| \, dt \\
\leq &\sum_{k=1}^{K} \int_{t_{k-1}}^{t_k} \left|\sigma_j^K(w_j)(k) - g_j^\epsilon(t)\right| \, dt + \sum_{k=1}^{K} \int_{t_{k-1}}^{t_k} \left|g_j^\epsilon(t) - w_j(t)\right| \, dt \\
\leq &\sum_{k=1}^{K} \int_{t_{k-1}}^{t_k} \left|\sigma_j^K(w_j)(k) - g_j^\epsilon(t)\right| \, dt + \epsilon.
\end{aligned}$$

Define $u^m = \max_{j \in \mathcal{J}}\{\max(u_j^{\min}, u_j^{\max})\}$. Let ζ be the number of jumps of g_j^ϵ. Then we choose K greater than $\zeta T u^m/\epsilon$. So the total length of the intervals where we have a jump is $\zeta\delta = \zeta T/K$ which is less than ϵ/u^m. For an interval were we have no jump we have the following:

$$\begin{aligned}
\int_{t_{k-1}}^{t_k} \left|\sigma_j^K(w_j)(k) - g_j^\epsilon(t)\right| \, dt &= \left|\sigma_j^K(w_j)(k) - g_j^\epsilon(t_{k-1})\right| \delta \\
&= \left| \frac{\mu}{1 - e^{-\mu\delta}} \int_{t_{k-1}}^{t_k} w(s) e^{-\mu(t_k - s)} \, ds - g_j^\epsilon(t_{k-1}) \right| \delta \\
&= \left| \frac{\mu}{1 - e^{-\mu\delta}} \int_{t_{k-1}}^{t_k} [w(s) - g_j^\epsilon(t_{k-1})] e^{-\mu(t_k - s)} \, ds \right| \delta \\
&= \frac{\mu\delta}{1 - e^{-\mu\delta}} \left| \int_{t_{k-1}}^{t_k} [w(s) - g_j^\epsilon(s)] e^{-\mu(t_k - s)} \, ds \right| \\
&\leq \frac{\mu\delta}{1 - e^{-\mu\delta}} \int_{t_{k-1}}^{t_k} \left|w(s) - g_j^\epsilon(s)\right| e^{-\mu(t_k - s)} \, ds \\
&\leq \frac{\mu\delta}{1 - e^{-\mu\delta}} \int_{t_{k-1}}^{t_k} \left|w(s) - g_j^\epsilon(s)\right| \, ds.
\end{aligned}$$

Therefore the following holds true

$$\begin{aligned}\sum_{k=1}^{K}\int_{t_{k-1}}^{t_k}\left|\sigma_j^K(w_j)(k)-g_j^\epsilon(t)\right|\,dt &\leq \frac{\mu\delta}{1-e^{-\mu\delta}}\int_0^T\left|w(t)-g_j^\epsilon(t)\right|\,dt+\epsilon\\ &\leq \frac{\mu\delta}{1-e^{-\mu\delta}}\epsilon+\epsilon\\ &\leq \frac{\mu T}{1-e^{-\mu T}}\epsilon+\epsilon.\end{aligned}$$

For the first inequality, the term ϵ is the majoration for the intervals where we have a jump. The last inequality is valid as the fraction is an increasing function for δ positive and that $\delta \leq T$. So the proof is complete. $\square$

Proof of Proposition 3.

$$\begin{aligned}&|V_j^K(\sigma_1^K(u_1),\sigma_2^K(u_2))-V_j(u_1,u_2)| \quad \text{(D.2)}\\ \leq\ &|V_j^K(\sigma_1^K(u_1),\sigma_2^K(u_2))-V_j(\phi_1^K\circ\sigma_1^K(u_1),\phi_2^K\circ\sigma_2^K(u_2))|+\\ &|V_j(\phi_1^K\circ\sigma_1^K(u_1),\phi_2^K\circ\sigma_2^K(u_2))-V_j(u_1,u_2)|. \quad \text{(D.3)}\end{aligned}$$

The first term in (D.3) tends to zero according to Proposition 2. The profit functions can be rewritten as

$$V_j(\gamma_1,\gamma_2)=\mathrm{E}_\gamma\left[\int_0^T e^{-\rho_j t}L_j^{\xi(t)}(x_1(t),\ldots,x_J(t),u_j(t))\,dt\right]=\mathrm{E}_\xi[W_j^{\gamma_1,\gamma_2}(\xi)].$$

As L_j^i are Lipschitz continuous functions and as the strategies satisfy the stability requirement of assumption 6, Lemma 6 implies that the random variable $W_j^{\gamma_1,\gamma_2}(\cdot)$ converges almost surely to the random variable $W_j^{\phi_1^K\circ\sigma_1^K(\gamma_1),\phi_2^K\circ\sigma_2^K(\gamma_2)}(\cdot)$. This implies that the second term in (D.3) tends to zero. $\square$

Proof of Proposition 4.

$$\begin{aligned}&|V_1^K(\mathbf{u}_1^K,\sigma_2^K(u_2))-V_1(\phi_1^K(\mathbf{u}_1^K),u_2)| \quad \text{(D.4)}\\ \leq\ &|V_1^K(\mathbf{u}_1^K,\sigma_2^K(u_2))-V_1(\phi_1^K(\mathbf{u}_1^K),\phi_2^K\circ\sigma_2^K(u_2))|+\\ &|V_1(\phi_1^K(\mathbf{u}_1^K),\phi_2^K\circ\sigma_2^K(u_2))-V_1(\phi_1^K(\mathbf{u}_1^K),u_2)|. \quad \text{(D.5)}\end{aligned}$$

The first term in (D.5) tends to zero according to Proposition 2. The same argumentation as in the previous proof shows that the second term in (D.5) tends to zero. The uniformity in $\{\mathbf{u}_1^K\}_{K=1}^\infty$ is guaranteed by the Lipschitz property of the functions L_j^i. $\square$

The proof of Proposition 5 is similar to the last proof.

D.2 Convergence results for the deterministic case

As the deterministic oligopoly is a special case of the stochastic one, all the theorems and lemmas in the appendix D.1 are valid for the deterministic oligopoly. In particular Theorem 14 is a special case of Theorem 18 which has been proved in appendix D.1. However, we have to prove Theorem 15 since it is stronger than Theorem 19 when it is applied to the deterministic oligopoly.

We will endow the control and state spaces with norms defined as follows $||u_j|| = \int_0^T |u_j(t)|\, dt$, a norm on U_j, $||\mathbf{u}_j^K|| = \sum_k |u^K(k)|\delta$, a norm on U_j^K, $||x_j|| = \max_{t\in[0,T]} |x_j^K(t)|$, a norm on X_j and $||\mathbf{x}_j^K|| = \max_k |x(k)|$, a norm on X_j^K. Before proving Theorem 15 we will state and prove the following Lemma.

Lemma 7. *Assume that, for all j, $L_j(x, u_j)$ is Lipschitz continuous in x and u_j. Then $V_j^K(\mathbf{u}_1^K, \mathbf{u}_2^K)$ is uniformly continuous in $\mathbf{u}_j^K \in U_j^K$.*

Proof. First we show that two controls $\mathbf{u}_j^K$ and $\mathbf{v}_j^K$ which are close to each other generate two trajectories $\mathbf{x}_j^K$ and $\mathbf{y}_j^K$ which are close to each other.

$$\begin{aligned}
|x_j^K(k) - y_j^K(k)| &= \left| \sum_{l=0}^{k} \sum_{l=1}^{k} [u_j^K(l) - v_j^K(l)] \cdot e^{-\mu_j(t_k - t_l)} \frac{1 - e^{-\mu_j\cdot\delta}}{\mu_j} \right| \\
&\leq \left| e^{-\mu_j(t_k - t_l)} \frac{1 - e^{-\mu_j\cdot\delta}}{\mu_j \delta} \right| \sum_{l=0}^{k} \left| u_j^K(l) - v_j^K(l) \right| \delta \\
&= \alpha_K \sum_{l=0}^{k} \left| u_j^K(l) - v_j^K(l) \right| \delta \\
&\leq \alpha_K ||\mathbf{u}_j^K - \mathbf{v}_j^K||.
\end{aligned}$$

So we have

$$||\mathbf{x}_j^K - \mathbf{y}_j^K|| = \max_k |x_j^K(k) - y_j^K(k)| \leq \alpha_K ||\mathbf{u}_j^K - \mathbf{v}_j^K||.$$

We now prove the continuity of $V_1^K(\mathbf{u}_1^K, \mathbf{u}_2^K)$ in $\mathbf{u}_1^K$. The proof of the continuity in $\mathbf{u}_2^K$ is similar. We have

$$
\begin{aligned}
&|V_1^K(\mathbf{u}_1^K,\mathbf{u}_2^K) - V_1^K(\mathbf{v}_1^K,\mathbf{u}_2^K)| \\
&= |\sum_{k=1}^K e^{-\rho_j t_k}[L_1(x_1^K(k),x_2^K(k),u_1^K(k)) - L_1(y_1^K(k),x_2^K(k),v_1^K(k))]\,\delta| \\
&\le \sum_{k=1}^K |L_1(x_1^K(k),x_2^K(k),u_1^K(k)) - L_1(y_1^K(k),x_2^K(k),v_1^K(k))|\,\delta \\
&\le \sum_{k=1}^K |L_1(x_1^K(k),x_2^K(k),u_1^K(k)) - L_1(y_1^K(k),x_2^K(k),u_1^K(k))|\,\delta + \\
&\quad \sum_{k=1}^K |L_1(y_1^K(k),x_2^K(k),u_1^K(k)) - L_1(y_1^K(k),x_2^K(k),v_1^K(k))|\,\delta \\
&\le \sum_{k=1}^K C|x_1^K(k) - y_1^K(k)|\delta + \sum_{k=1}^K C|u_1^K(k) - v_1^K(k)|\delta \\
&= CT||\mathbf{x}_1^K - \mathbf{y}_1^K|| + C||\mathbf{u}_1^K - \mathbf{v}_1^K|| \\
&\le C(T\alpha_K + 1)||\mathbf{u}_1^K - \mathbf{v}_1^K||.
\end{aligned}
$$

□

Proof of Theorem 15. The set X_j of the trajectories with the norm $||x_j|| = \max|x_j(t)|$ is compact. Let $M_j(u_j) = \int_0^T \mathcal{M}_j(u_j(t))\,dt$ and $N_j = \{M_j(u_j) : u_j \in U_j\}$. As M_j is Lipschitz and U_j is compact N_j is compact. Therefore we can extract from the sequence $(\phi_1^K(\mathbf{x}_1^{K*}),\phi_2^K(\mathbf{x}_2^{K*}),M_1(\phi_1^K(\mathbf{u}_1^{K*})),M_2(\phi_2^K(\mathbf{u}_2^{K*})))$ a converging subsequence $(\phi_1^{K_p}(\mathbf{x}_1^{K_p*}),\phi_2^{K_p}(\mathbf{x}_2^{K_p*}),M_1(\phi_1^{K_p}(\mathbf{u}_1^{K_p*})),M_2(\phi_2^{K_p}(\mathbf{u}_2^{K_p*})))$. Denote by $(\tilde{x}_1,\tilde{x}_2,\tilde{m}_1,\tilde{m}_2)$ the limit of this subsequence. Denote by $(\tilde{u}_1,\tilde{u}_2)$ a control which generates the trajectory $(\tilde{x}_1,\tilde{x}_2)$. Here we must emphasize that we could have $M_j(\tilde{u}_i) \neq \tilde{m}_i$.

We first show that the case where $M_j(\tilde{u}_i) < \tilde{m}_i$ is impossible. Suppose $M_j(\tilde{u}_i) < \tilde{m}_i$. We rewrite the function $V_1(u_1,u_2)$ as $\tilde{V}_1(x_1,x_2,M_1(u_1))$ to emphasize that it is a sum of two terms, one depending on (x_1,x_2) and the other depending on $M_1(u_1)$. Thus we have $\tilde{V}_1(\tilde{x}_1,x_2,\tilde{m}_1) < \tilde{V}_1(\tilde{x}_1,x_2,M_1(u_1))$. So there exist $\Delta > 0$ such that

$$\tilde{V}_1(\tilde{x}_1,x_2,\tilde{m}_1) = \tilde{V}_1(\tilde{x}_1,x_2,M_1(u_1)) - \Delta. \tag{D.6}$$

For K_p sufficiently big we have the following:

$$
\begin{aligned}
V_1^{K_p}(\mathbf{u}_1^{K_p*},\mathbf{u}_2^{K_p*}) &\le V_1(\phi_1^{K_p}(\mathbf{u}_1^{K_p*}),\phi_2^{K_p}(\mathbf{u}_2^{K_p*})) + \epsilon \\
&\le V_1(\tilde{u}_1,\phi_2^{K_p}(\mathbf{u}_2^{K_p*})) + 2\epsilon - \Delta \\
&\le V_1^{K_p}(\sigma_1^{K_p}(\tilde{u}_1),\sigma_2^{K_p}\circ\phi_2^{K_p}(\mathbf{u}_2^{K_p*})) + 3\epsilon - \Delta \\
&= V_1^{K_p}(\sigma_1^{K_p}(\tilde{u}_1),\mathbf{u}_2^{K_p*}) + 3\epsilon - \Delta.
\end{aligned}
$$

The first inequality is implied by Proposition 2. The second inequality is implied by the convergence of $\phi_1^{K_p}(\mathbf{x}_1^{K_p *})$ to $\tilde{x}_1$, the continuity of $\phi_j^{K_p}$, the continuity of $V_j^{K_p}$ (Lemma 7) and equation (D.6). The third inequality comes from Proposition 3. The last inequality is valid as $\sigma_j^K \circ \phi_j^K = \mathbb{1}$. As ϵ can be chosen as small as desired for K_p big enough and as $\Delta > 0$ we conclude that

$$V_1^{K_p}(\mathbf{u}_1^{K_p *}, \mathbf{u}_2^{K_p *}) < V_1^{K_p}(\sigma_1^{K_p}\tilde{u}_1, \mathbf{u}_2^{K_p *})$$

which contradicts the fact that $(\mathbf{u}_1^{K_p *}, \mathbf{u}_2^{K_p *})$ is an equilibrium for the approximating game.

So we can assume that $M_j(\tilde{u}_i) \geq \tilde{m}_i$. Then, for all u_1 in U_1 and K_p big enough, the following holds:

$$\begin{aligned} V_1(u_1, \tilde{u}_2) &\leq V_1^{K_p}(\sigma_1^{K_p}(u_1), \sigma_2^{K_p}(\tilde{u}_2)) + \epsilon \\ &\leq V_1^{K_p}(\sigma_1^{K_p}(u_1), \mathbf{u}_2^{K_p *}) + 2\epsilon \\ &\leq V_1^{K_p}(\mathbf{u}_1^{K_p *}, \mathbf{u}_2^{K_p *}) + 2\epsilon \\ &\leq V_1(\phi_1^{K_p}(\mathbf{u}_1^{K_p *}), \phi_2^{K_p}(\mathbf{u}_2^{K_p *})) + 3\epsilon \\ &\leq V_1(\tilde{u}_1, \tilde{u}_2) + 4\epsilon. \end{aligned}$$

The first inequality is implied by Proposition 3. The second inequality is implied by the convergence[1] of $\phi_j^{K_p}(\mathbf{x}_j^{K_p *})$ to $\tilde{x}_j$, the continuity of $\phi_j^{K_p}$, the continuity of $V_j^{K_p}$ (Lemma 7) and the property $\sigma_j^K \circ \phi_j^K = \mathbb{1}$. The third inequality is valid as $(\mathbf{u}_1^{K_p *}, \mathbf{u}_2^{K_p *})$ is the equilibrium for the discrete time oligopoly of order K_p. The fourth inequality is obtained applying Proposition 2. The last inequality is valid as $(\phi_1^{K_p}(\mathbf{x}_1^{K_p *}), \phi_2^{K_p}(\mathbf{x}_2^{K_p *}))$ converges to $(\tilde{x}_1, \tilde{x}_2)$, when $M_1(\phi_1^{K_p}(\mathbf{u}_1^{K_p *}))$ converges to something smaller than or equal to $M_1(\tilde{u}_1)$ and since V_j is a continuous function in all is arguments. A similar property can be obtained for any other player j. As ϵ can be taken as small as desired, $(\tilde{x}_1, \tilde{x}_2)$ is the equilibrium trajectory of the continuous time oligopoly. As the equilibrium is unique, the whole sequence $(\phi_1^K(\mathbf{x}_1^{K*}), \phi_2^K(\mathbf{x}_2^{K*}))$ converges to $(\tilde{x}_1, \tilde{x}_2) = (x_1^*, x_2^*)$. We will show now that we can exclude the case $M_j(u_i^*) = M_j(\tilde{u}_i) > \tilde{m}_i$. By the same argumentation as before we can show that there exist $\Delta > 0$ such that

$$\tilde{V}_1(x_1^*, x_2^*, M_j(u_1^*)) = \tilde{V}_1(x_1^*, x_2^*, \tilde{m}_1) - \Delta.$$

As $(\phi_1^{K_p}(\mathbf{x}_1^{K_p *}), M_1(\phi_1^{K_p}(\mathbf{u}_1^{K_p *})))$ converges to $(x_1^*, \tilde{m}_1)$, as $\phi_j^{K_p}$ is a continuous function, and as $\tilde{V}_1$ is a continuous function in all is arguments, for all ϵ there exists K_p big enough such that

$$\tilde{V}_1(x_1^*, x_2^*, \tilde{m}_1) \leq \tilde{V}_1(\phi_1^{K_p}(\mathbf{x}_1^{K_p *}), x_2^*, M_1(\phi_1^{K_p}(\mathbf{u}_1^{K_p *}))) + \epsilon.$$

[1] Recall that $V_1^{K_p}$ does not depend on $M_j(\phi_j^{K_p}(\mathbf{u}_j^{K_p *}))$ $j \neq 1$.

Thus, from the last two statements, we obtain

$$V_1(u_1^*, u_2^*) < V_1(\phi_1^{K_p}(\mathbf{u}_1^{K_p*}), u_2^*)$$

which is a contradiction hence $M_j(u_i^*) = M_j(\tilde{u}_i) = \tilde{m}_i$. $\square$

Bibliography

[1] M. Abbad, *Perturbation and Stability Theory for Markov Control Problems*, PhD thesis, University of Maryland at Baltimore County, Dept. of Mathematics, 1991.

[2] M. Abbad and J. A. Filar, *Perturbation and stability theory for markov control problems*, IEEE Transactions on Automatic Control, 37 (1992), pp. 1415–1420.

[3] M. Abbad, J. A. Filar, and T. R. Bielecki, *Algorithms for singularly perturbed limiting average markov control problems*, IEEE Transactions on Automatic Control, 37 (1992), pp. 1421–1425.

[4] R. Abraham and J. Marsden, *Foundation of Mechanics*, Addison-Wesley, Redwood City, California, 1987.

[5] R. Akella and P. Kumar, *Optimal control of production rate in a failure prone manufacturing system*, IEEE Transactions on Automatic Control, 31 (1986), pp. 1421–1425.

[6] P. H. Algoet, *Flow balence equations for steady-sate distribution of a flexible manufacturing system*, IEEE Transactions on Automatic Control, 34 (1989), pp. 917–921.

[7] E. Altman, O. Pourtallier, and M. Tidball, *Approximations in dynamic zero-sum games, II*, SIAM J. Control optim., 35 (1997), pp. 2101–2117.

[8] K. J. Arrow and M. Kurz, *Public Investment, the Rate of Return and Optimal Fiscal Policy*, The Johns Hopkins Press, Baltimore, 1971.

[9] T. Başar and G. Olsder, *Dynamic Noncooperative Game Theory*, Academic Press, London, 89.

[10] G. Barles, *Solutions de viscosité des équations de Hamilton-Jacobi*, Springer-Verlag, Berlin, 1994.

[11] R. E. Bellman, *Dynamic programming*, Princeton University Press, Princeton N.J., 1957.

[12] T. Bielecki and P. Kumar, *Optimality of zero inventory policies for unreliable manufacturing systems*, Operations Research, 36 (1988), pp. 523–541.

[13] T. R. Bielecki and J. A. Filar, *Singularly perturbed markov control problem: Limiting average cost*, Annals of Operations Research, 28 (1991), pp. 153–168.

[14] P. Billingsley, *Convergence of Probability Measures*, John Wiley and Sons, New-York, 1968.

[15] E. K. Boukas and A. Haurie, *Manufacturing flow control and preventive maintenance: A stochastic control approach*, IEEE Transactions on Automatic Control, 35 (1990), pp. 1024–1031.

[16] E. K. Boukas, A. Haurie, and C. van Delft, *A turnpike improvement algorithme for piecewise deterministic control*, Optimal Control Application & Methods, 12 (1991), pp. 1–17.

[17] W. Brock, *Differential games with active and passive variables*, in Mathematical Economics and Game Theory: Essays in Honor of Oskar Morgenstern, Henn and Moeschlin, eds., Springer Verlag, Berlin, 1977, pp. 34–52.

[18] A. E. Bryson and Y.-C. Ho, *Applied Optimal Control*, Ginn and Company, Waltham, 1969.

[19] I. Capuzzo Dolcetta, *On a discrete approximation of the hamilton-jacobi equation of dynamic programming*, Applied Mathematics and Optimization, 10 (1983), pp. 367–377.

[20] I. Capuzzo Dolcetta and L. C. Evans, *Optimal switching for ordinary differential equations*, SIAM Journal Control and Optimization, 22 (1984), pp. 143–161.

[21] M. Caramanis and G. Liberopoulos, *Perturbation analysis for the design of flexible manufacturing system flow controllers*, Operations Research, 40 (1992), pp. 1107–1125.

[22] M. Caramanis and A. Sharifnia, *Near optimal manufacturing flow controller design*, International Journal of Flexible Manufacturing Systems, 3 (1991), pp. 321–336.

[23] D. Carlson and A. Haurie, *A turnpike theory for infinite-horizon open-loop competitive processes*, Siam J. Control and Optimization, 34 (1996), pp. 1405–1419.

[24] A. Cournot, *Recherches sur les Principes Mathématiques de la Théorie des Richesses*, Hachette, Paris, 1838.

[25] M. G. Crandall, H. Ishii, and P. L. Lions, *User's guide to viscosity solutions of second order partial differential equation*, Bulletin of American Society, 27 (1992), pp. 1–67.

[26] G. Dantzig, *Linear programming under uncertainty*, Management Science, 1 (1955), pp. 197–206.

[27] M. Davis, *Piecewise deterministic markov processes: A general class of non-diffusion stochastic models*, J. of R. Stat. Soc., 46 (1984), pp. 353–388.

[28] G. T. de Ghellinck, *Les problèmes de décision séquentielles*, Cahier du Centre d'Etudes de Recherche Operationnelle, 2 (1960), pp. 161–179.

[29] F. Delebecque and J.-P. Quadrat, *Contribution of stochastic control, singular perturbation averaging and team theory to an example of large scale system: Management of hydropower production*, IEEE Transactions on Automatic Control, AC-23 (1978), pp. 209–222.

[30] F. Delebecque and J.-P. Quadrat, *Optimal control of Markov chains admitting strong and weak interactions*, Automatica, 17 (1981), pp. 281–296.

[31] E. V. Denardo, *Contractions mappings in the theory underlying dynamic programming*, SIAM Review, 9 (1967), pp. 165–177.

[32] R. Elliott, *The existence of the value in stochastic differential games*, SIAM Journal of Control and Optimization, 14 (1976), pp. 85–94.

[33] J. Filar, V. Gaitsgory, and A. Haurie, *Control of singularly perturbed hybrid stochastic systems*, in Proc. of the 35th IEEE Conference on Decison and Control, December 1996.

[34] J. Filar and K. Vrieze, *Competitive Markov decision Processes*, Springer, New-York, 1997.

[35] J. A. Filar and A. Haurie, *Optimal ergodic control of singularly perturbed hybrid stochastic systems*, Lectures in Applied Mathematics, 33 (1997), pp. 101–126.

[36] ——, *A two factor stochastic production model with two time scales*, tech. rep., 1998.

[37] W. Fleming, S. Sethi, and H. Soner, *An optimal stochastic production planing problem with randomly fluctuating demand*, SIAM Journal on Optimization and Control, 25 (1987), pp. 1494–1502.

[38] R. Fourer, D. M. Gay, and B. W. Kernighan, *AMPL: A modeling language for mathematical programming*, The Scientific Press, San Francisco, 1993.

[39] D. Freedman, *Markov Chains*, Springer-Verlag, New-York, 1983.

[40] A. Friedman, *Differential Games*, Wiley-Interscience, New-York, 1971.

[41] ——, *Stochastic differential games*, Journal of Differential Equations, 11 (1972), pp. 79–108.

[42] D. Fudenberg and J. Tirole, *Game Theory*, The MIT Press, Cambridge MA, 1995.

[43] D. Gabay and H. Moulin, *On the uniqueness and stability of Nash-equilibria in noncooperative games*, in Applied Stochastic Control in Econometrics and Management Sciences, A. Bensoussan, P. Kleindorfer, and C. Tapiero, eds., North-Holland, 1980, pp. 271–294.

[44] S. Gershwin, R. Akella, and Y. Choong, *Performance of hierarchical production scheduling policy*, IEEE Transaction on Components, Hybrids, and Manufacturing Technology, 7 (1984), pp. 225–248.

[45] ——, *Short term scheduling of an automated manufacturing facility*, IBM Journal of Research and Development, 29 (1985), pp. 393–400.

[46] S. B. Gershwin, *Hierarchical flow control: A framework for scheduling and planning discrete events in manufacturing systems*, Proceedings of the IEEE, 77 (1989), pp. 195–209.

[47] ——, *Manufacturing Systems Engineering*, PTR Prentice Hall, Englewood Cliffs, 1994.

[48] M. K. Ghost, A. Arapostathis, and S. I. Marcus, *Ergodic control of switching diffusions*, SIAM Journal of Control and Optimization, 35 (1997), pp. 1952–1988.

[49] J.-L. Goffin, A. Haurie, and J.-P. Vial, *Decomposition and nondifferentiable optimization with the projective algorithm*, Management Science, 38 (1992), pp. 284–302.

[50] J.-L. Goffin and J.-P. Vial, *Cutting planes and column generation techniques with the projective algorithm*, Journal of Optimization Theory and Applications, 65 (1990), pp. 409–429.

[51] J. Gondzio, O. du Merle, R. Sarkissian, and J.-P. Vial, *ACCPM - a library for convex optimization based on an analytic center cutting plane method*, European Journal of Operational Research, 94 (1996), pp. 206–211.

[52] J. Gondzio, A. Haurie, F. Moresino, and J.-P. Vial, *Parallel implementation of a decomposition method in a two time scale hybrid stochastic model.* In preparation.

[53] J. Gondzio, R. Sarkissian, and J.-P. Vial, *Parallel implementation of a central decomposition method for solving large scale planning problems*, tech. rep., HEC University of Geneva, 102 Bd Carl-Vogt, 1211 Genève, Switzerland, January 1998.

[54] G. Gürkan, A. Y. Özge, and S. M. Robinson, *Sample-path solutionis for simulation optimization problems and stochastic variational inequalities*, in Interfaces in Computer Science and Operations Research, D. Woodruff, ed., Kluwer, Boston, 1998, pp. 169–188.

[55] ——, *Sample-path solution of stochastic variational inequalities*, Mathematical Programming, 84 (1999), pp. 313–333.

[56] P. Hartman and G. Stampacchia, *On some nonlinear elliptic differential functional equations*, Acta Mathematica, 115 (1966), pp. 271–310.

[57] A. Haurie, *Duopole et percée technologiques: Un modèle de jeu différentiel déterministe par morceaux*, L'actualité Economique, (1989), pp. 105–118.

[58] ——, *Piecewise deterministic differential games*, in Differential Games and Applications, T. Başar and P. Bernhard, eds., Springer, 1989, pp. 114–127.

[59] ——, *Stochastic differential games in economic modeling*, in System Modelling and Optimization, J. Henry and J.-P. Yvon, eds., vol. 197 of Lecture Notes in Control and Information, Springer Verlag, 1994, pp. 90–108.

[60] A. Haurie, P. L'Ecuyer, and C. Van Delft, *Convergence of stochastic approximation coupled with perturbation analysis in a class of manufacturing flow control models*, Discrete Event Dynamic Systems: Theory and Applications, 4 (1994), pp. 87–111.

[61] A. Haurie and G. Leitmann, *On the global stability of equilibrium solutions for open-loop differential games*, Large Scale Systems, 6 (1984), pp. 107–122.

[62] A. Haurie, A. Leizarowitz, and C. van Delft, *Turnpike properties for a class of piecewise deterministic systems arising in manufacturing flow control*, Annals of Operations Research, 29 (1991), pp. 351–357.

[63] ——, *Boundedly optimal control of piecewise deterministic systems*, European Journal of Operational Research, 73 (1994), pp. 237–251.

[64] ——, *Turnpike in flow control modesl of unreliable manufacturing systems*, European Journal of Operational Research, 82 (1995), pp. 338–351.

[65] A. Haurie and M. Roche, *Turnpikes and computation of piecewise open-loop equilibria in stochastic differential games*, Journal of Economic Dynamics and Control, 18 (1994), pp. 317–344.

[66] A. Haurie and C. van Delft, *Turnpike properties for a class of piecewise deterministic systems arising in manufacturing flow control*, Annals of Operations Research, 29 (1991), pp. 351–374.

[67] A. Haurie, G. Zaccour, J. Legrand, and Y. Smeers, *A stochastic dynamic Nash-Cournot model for European gas market*, tech. rep., HEC Montréal, 1987.

[68] A. Haurie, G. Zaccour, and Y. Smeers, *Stochastic equilibrium programming for dynamic oligopolistic markets*, Journal of Optimization Theory and Applications, 66 (1990), pp. 243–253.

[69] J. Q. Hu, P. Vakili, and G. X. Yu, *Optimality of hedging point policy for failure prone manufacturing systems*, IEEE Transactions on Automatic Control, 39 (1994), pp. 1875–1880.

[70] J. Q. Hu and D. Xiang, *Monotonicity of optimal flow controls for failure prone production systems*, Journal of Optimization Theory and Applications, 86 (1995), pp. 57–71.

[71] J. C. Hull, *Options, Futures, and Other Derivatives*, Prentice Hall, Upper Saddle River, 1997.

[72] G. Infanger, *Planning under Uncertainty*, The Scientific Press Series, Danvers, 1994.

[73] R. Isaacs, *Differential games,I,II,III,IV*, Tech. Rep. RM-1391,1399,1411 and 1486, Rand Corporation, 1954.

[74] ——, *Differential Games*, Wiley, New-York, 1965.

[75] P. Kall and S. Wallace, *Stochastic Programming*, Wiley - Interscience, New York, 1995.

[76] J. G. Kimemia, *Hierarchical Control of production in Flexible Manufacturing Systems*, PhD thesis, MIT, 1982.

[77] J. G. Kimemia and S. B. Gershwin, *An algorithm for the computer control of production in flexible manufacturing systems*, IIE Transactions, 15 (1983), pp. 352–362.

[78] A. King, *Stochastic programming interface library*, tech. rep., T. J. Watson Research Center, IBM Research Center, Yorktown Heights, NY 10598, May 1994.

[79] I. Konnov, *Combined relaxation methods for finding equilibrium points and solving related problems*, Izvestiya VUZ, Mathematika, 37 (1993), pp. 44–51.

[80] N. Krassovski and A. Soubbotine, *Jeux différentiels*, Editions Mir, Moscou, 1974. Traduit du russe par D. Embarek.

[81] E. V. Krichagina, S. X. C. Lou, S. P. Sethi, and M. I. Taksar, *Production control in a failure-prone manufacturing system: Diffusion approximation and asymptotic optimality*, The Annals of Applied Probability, 3 (1993), pp. 421–453.

[82] K. S. Kundert and A. Sangiovanni-Vincentelli, *Sparse user's guide. A sparse linear equation solver*, tech. rep., Department of Engineering and Computer Science, University of California, Berkeley, June 1988.

[83] H. J. Kushner and P. G. Dupuis, *Numerical Methods for Stochastic Control Problems in Continuous Time*, Springer Verlag, New-York, 1992.

[84] P. R. G. Layard and A. A. Walters, *Microeconomic Theory*, McGraw-Hill, New-York, 1978.

[85] S. Lenhart, *Viscosity solutions associated with switching control problems for piecewise-deterministic processes*, Houston journal of mathematics, 13 (1987), pp. 405–426.

[86] G. Liberopoulos, *Flow Control of failure Prone Manufacturing Systems: Controller Design Thory and Applications*, PhD thesis, Boston University, 1993.

[87] R. Luce and H. Raiffa, *Games and Decisions*, Wiley and Sons, New-York, 1957.

[88] D. G. Luenberger, *Introduction to Linear and Nonlinear Programming*, Addison-Wesley, Reading, MA, 1984.

[89] O. Z. Maimon and S. B. Gershwin, *Dynamic scheduling and routing for flexible manufacturing systems that have unreliable machines*, Operations Research, 36 (1988), pp. 279–292.

[90] R. Malhamé and K. Boukas, *A renewal theoretic analysis of a class of manufacturing systems*, IEEE Transactions on Automatic Control, 36 (1991), pp. 580–587.

[91] A. G. Malliaris and W. A. Brock, *Stochastic Methods in Economics and Finance*, Elsevier Science Bublishers B. V., Amsterdam, 1984.

[92] A. Manne, *Linear programming and sequential decisions*, Management Science, 6 (1960), pp. 259–267.

[93] R. C. Merton, *Continuous-Time Finance*, Blackwell, Cambridge USA, 1990.

[94] A. Nagurney, *Network Economics: A Variational Inequality Approach*, Kluwer, Dordrecht, 1993.

[95] J. F. Nash, *Equilibrium point in n-person games*, Proceeding of the National Academy of Sciences, 36 (1950), pp. 48–49.

[96] ———, *Noncooperative games*, Annals of Mathematics, 54 (1951), pp. 286–298.

[97] G. Olsder and R. Suri, *Time optimal control of parts-routing in a manufacturing system with failure prone machines*, in Proc. IEEE International Conference on Decision and Control, 1980, pp. 722–727.

[98] G. Owen, *Game Theory*, Saunders, Philadelphia, 1968.

[99] P.-O. Pineau and P. Murto, *Production and investment dynamics in the Finnish electricity market.* 1998.

[100] O. Pourtallier, *Approximations in static non-zero-sum games.* Working paper INRIA, 1998.

[101] M. L. Puterman, *Markov Decision Processes*, Wiley, New-York, 1994.

[102] R. Rishel, *Dynamic programming and minimum principles for systems with jump markov disturbances*, SIAM Journal on Control, 13 (1975), pp. 338–371.

[103] R. Rishel, *Controlled wear process: Modeling optimal control*, IEEE Transactions on Automatic Control, 36 (1988), pp. 1100–1102.

[104] J. B. ROSEN, *Existence and uniqueness of equilibrium points for concave n-person games*, Econometrica, 33 (1965), pp. 520–534.

[105] R. SARKISSIAN, *Telecomunications Networks: Routing and Survivability Optimization Using a Central Cutting Plane Method*, PhD thesis, EPFL, 1998.

[106] P. J. SCHWEITZER, *Perturbation theory and finite Markov chains*, Journal of Applied Probability, 5 (1968), pp. 401–413.

[107] S. SETHI AND Q. ZHANG, *Hierarchical Decision Making in Stochastic Manufacturing Systems*, Birkhäuser, Boston, 1995.

[108] L. S. SHAPLEY, *Stochastic games*, Proceedings National Academy od Science, 39 (1953), pp. 1095–1100.

[109] A. SHARIFNIA, *Production control of a manufacturing system with multiple machine states*, IEEE Transactions on Automatic Control, 33 (1988), pp. 620–625.

[110] ——, *Stability and performance of distributed production control methods based on continuous-flow models*, IEEE Transactions on Automatic Control, 39 (1994), pp. 725–737.

[111] A. SHARIFNIA, M. CARAMANIS, AND S. B. GERSHWIN, *Dynamic setup scheduling and flow control in manufacturing systems*, Discrete Event Dynamic Systems: Theory and Applications, 1 (1991), pp. 149–175.

[112] M. SNIR, S. W. OTTO, S. HUSS-LEDERNAN, D. W. WALKER, AND J. DONGARRA, *MPI: The Complete Reference*, The MIT Press, Cambrige, Massachusetts, 1996.

[113] G. SONNEVEND, *New algorithms in convex programming based on a notion of "centre" (for systems of analytic inequalities) and on rational extrapolation*, in Trends in Mathematical Optimization : Proceedings of the 4th French–German Conference on Optimization in Irsee, Germany, April 1986, K. H. Hoffmann, J. B. Hiriat-Urruty, C. Lemarechal, and J. Zowe, eds., vol. 84 of International Series of Numerical Mathematics, Birkhäuser Verlag, Basel, Switzerland, 1988, pp. 311–327.

[114] M. SRIVATSAN AND Y. DALLERY, *Partial characterization of optimal hedging point policies in unreliable two-part-type manufacturing system*, Operations Research, 48 (1998), pp. 36–45.

[115] D. SWORDER, *Feedback control of a class of linear systems with jump parameters*, IEEE, Trans. Aut. Control, AC-14 (1969), pp. 9–14.

[116] D. D. Sworder, *Feedback control of a class of linear systems with jump parameters*, IEEE Transactions on Automatic Control, 14 (1969), pp. 9–14.

[117] K. Uchida, *On existence of n-person nonzero sum stochastic differential games: An existence theorem*, SIAM Journal of Control and Optimization, 16 (1989), pp. 153–177.

[118] C. van Delft, *Commande optimale de système déterministe par morceaux: Application à la gestion de la production d'un atelier flexible*, PhD thesis, Université de Geneève, 1991.

[119] D. Vermes, *Optimal control of piecewise deterministic markov process*, Stochastics, 14 (1985), pp. 165–208.

[120] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behaviour*, Princeton University Press, 1944.

[121] W. Wonham, *Random differential equations in control theory*, in Probabilistic Methods in Applied Mathematics, A. Barucha-Reid, ed., Academic Press, NY, 1970, pp. 131–217.

[122] Q. Zhang and G. Yin, *Turnpike sets in stochastic manufacturing systems with finite time horizon*, Stochastics and Stochastic Reports, 51 (1994), pp. 11–40.