ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC

MÉMOIRE PAR ARTICLES PRÉSENTÉ À
L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

COMME EXIGENCE PARTIELLE
À L'OBTENTION DE LA
MAÎTRISE EN GÉNIE MÉCANIQUE

PAR
Thang DIEP-THANH

COMMANDE OPTIMALE STOCHASTIQUE APPLIQUÉE AUX SYSTÈMES MANUFACTURIERS AVEC DES SAUTS SEMI-MARKOVIENS

MONTRÉAL, LE 14 OCTOBRE 2011

CE MÉMOIRE A ÉTÉ ÉVALUÉ

PAR UN JURY COMPOSÉ DE :

M. Thien-My Dao, directeur de mémoire
Département de Génie mécanique à l'École de technologie supérieure

M. Saad Maarouf, président du jury
Département de Génie électrique à l'École de technologie supérieure

M. Jean-Pierre Kenné, membre du jury
Département de Génie mécanique à l'École de technologie supérieure

ELLE FAIT L'OBJECT D'UNE SOUTENANCE DEVANT JURY

LE 03 OCTOBRE 2011

À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

À mes parents,
Mes frères, mes sœurs et toute ma famille
Je vous aime

REMERCIEMENTS

Tout d’abord, j’aimerais sincèrement remercier Monsieur Thien-My Dao qui m’a offert son aide tout au long de ce projet même si je faisais face à des débuts difficiles. Toutes ces années passées sous sa supervision m’ont appris à être autonome dans le travail et à accéder à une grande connaissance de mon sujet. Je le remercie de ses nombreux conseils, ainsi de la confiance qu’il a su m’accorder.

Je tiens à remercier Monsieur Jean-Pierre Kenné qui avait accepté en premier de réaliser ce travail de mémoire. Je lui suis reconnaissant du temps qu’il a pu me consacrer, ainsi que de son soutien qui m’a stimulé pour mener à bien mon projet.

Je tiens à remercier Monsieur Saad Maarouf d'avoir accepté de présider ce jury de mémoire et pour avoir accepté d'évaluer ce travail.

J’adresse mes remerciements au département de Génie Mécanique et Industriel de l’Université du Minnesota à Duluth (UMD) aux États-Unis, ainsi qu’au Professeur Séraphin Abou qui m’a accordé un séjour de recherche à l’UMD.

J’adresse mes remerciements à tous mes collègues et amis du Laboratoire d’intégration des technologies de production de l’ETS, qui m’ont apporté leur aide et m’ont appuyé tout au long de ma recherche.

Je voudrais aussi remercier toute ma famille de mon pays natal pour leur soutien et leur support, sans lesquels je n’aurais pas pu atteindre cette étape finale.

COMMANDE OPTIMALE STOCHASTIQUE APPLIQUÉE AUX SYSTÈMES MANUFACTURIERS AVEC DES SAUTS SEMI-MARKOVIENS

Thang DIEP-THANH

RÉSUMÉ

Les travaux de ce mémoire sont constitués de deux parties principales. La première partie tente de formuler un nouveau modèle du problème de commande optimale stochastique de systèmes sur un horizon fini. Les systèmes considérés sont soumis à des phénomènes aléatoires dits sauts de perturbation qui sont modélisés par un processus semi-Markovien. Ces sauts de perturbation traduits par des taux de transition dépendent de l'état du système et du temps. Par conséquent, le problème de commande est formulé comme un problème d'optimisation dans un environnement stochastique. La deuxième partie vise à modéliser des systèmes de production flexible (SPF). Dans ce mémoire, ces SPF se composent de plusieurs machines en parallèles, ou en série, ou d'une station de travail (une machine représentative). Ces machines sont sujettes à des pannes et à des réparations aléatoires. L'objectif de la modélisation est de déterminer les taux de production $\boldsymbol{u}(t)$ de ces machines en satisfaisant les fluctuations de demande $\boldsymbol{d}(t)$ sur un horizon fini.

Dans ce mémoire, nous avons :

(a) proposé un nouveau modèle du problème d'optimisation dans un environnement stochastique sur un horizon fini pour deux cas; avec taux d'actualisation ($\rho > 0$) et sans taux d'actualisation ($\rho = 0$);

(b) modélisé des SPF en déterminant une stratégie de commande plus réaliste incluant stratégie de production;

(c) présenté des exemples numériques à l'aide d'une méthode de Kushner et Dupuis (2001).

Mots-clés : Systèmes manufacturiers, commande optimale stochastique, processus Markovien et semi-Markovien, planification production.

OPTIMAL STOCHASTIC CONTROL APPLIED TO MANUFACTURING SYSTEMS WITH SEMI-MARKOVIAN JUMPS

Thang DIEP-THANH

ABSTRACT

In this work, we present a new model for optimal production control of manufacturing systems. The new model is formulated as an optimal control problem in random environment in finite horizon for two cases; with discount rate and without it. The systems are subject to random events which are modeled by a semi-Markov process. The lifetime of each random event obeys non-exponential distribution instead of being exponential in the Markov framework. By using this new model, the modeling of the manufacturing systems aims to find the production rate $\boldsymbol{u}(t)$ in real-time in which the arrival of demand is considered as a random event. The manufacturing systems considered are constituted of several interconnected machines. These machines are subject to random breakdowns and repairs, and their functioning distributions depend on the time (the age).

Consequently, in this work, our contributions are:

(a) development of a new model for an optimization problem in random environment in finite horizon for two cases; with discount rate ($\rho > 0$) and without discount rate ($\rho = 0$);

(b) modeling of manufacturing systems whose objective is to determine the strategies of production;

(c) using numerical approach of Kushner and Dupuis (2001) is to represent numerial exemples.

Keywords: Manufacturing systems, Optimal stochastic control, Markov and semi-Markov processes, Production planning.

SOMMAIRE DU MÉMOIRE

L'étude actuelle de la performance d'un système de production doit tenir compte des différents aléas tels que les fluctuations de la demande, les pannes et les réparations des machines, les actions de la maintenance préventive, du setup, etc. Ces phénomènes aléatoires sont donc non-prévisibles lorsque le système est en opération. Pour augmenter la performance du système en présence des imprévisibilités, une bonne façon est d'optimiser les inventaires du système par la recherche des taux de production. Ce contexte nous permet d'aborder le problème de commande optimale stochastique de système de production par modélisation mathématique.

Dans le contexte de la commande optimale stochastique de système où les variables sont construites à partir de l'état continu et de l'état discret, le développement théorique est en évolution depuis le premier formalisme de Rishel (1975), de Davis (1984), de Boukas (1987), de Boukas and Haurie (1990), ainsi que de Sethi et al. (2005). Dans sa théorie du formalisme, Rishel (1975) a considéré des systèmes en temps continu et des états discrets caractérisés par des processus Markoviens homogènes. En tenant compte d'une extension du formalisme de Rishel (1975), Boukas (1987) a considéré les mêmes systèmes, mais a rejeté des états discrets caractérisés par les processus Markoviens non-homogènes. Que ce soit les processus Markoviens homogènes ou les processus Markoviens non-homogènes, ils sont modélisés par une chaîne de Markov dont les taux de transition sont indépendants du temps.

Les propositions de ce mémoire sont les suivantes :

1. Une généralisation du formalisme de Rishel en utilisant un processus semi-Markovien pour caractériser des états discrets de systèmes au cas où les taux de transitions (de sauts aléatoires) entre des états dépendraient du temps. Ce processus stochastique modélise non

seulement une discontinuité de la partie continue de l'état du système $\boldsymbol{x}^{\alpha}(t)$, mais une continuité stochastique du système $\xi(t)$ par rapport au temps[1].

2. Une application de ce nouveau modèle pour la modélisation des systèmes de production constitués de plusieurs machines parallèles de même qu'une application de modélisation sur une ligne de production. Pour ces systèmes considérés, les taux de demandes sont considérés comme des variables aléatoires par rapport au temps.

Le problème de commande optimale stochastique sur un horizon fini (horizon déterministe), que nous considérons dans ce mémoire, consiste à minimiser l'espérance mathématique du coût envisagé. Les conditions d'optimum établies satisfont le principal d'optimum de Pontryagin.

---

[1] $\boldsymbol{x}^{\alpha}(t)$ est le niveau d'inventaire (variable d'état du système) dans le mode α à l'instant t; lorsque le système est en panne cette variable est discontinue. $\xi(t)$ est le processus stochastique à temps continu et à état fini.

PROBLÉMATIQUE GÉNÉRALE

Comme nous l'avons mentionné dans le sommaire précédent, ce mémoire aborde le problème d'optimisation stochastique appliqué aux systèmes de production. Nous commençons à introduire un exemple simple [2] qui permet de mieux comprendre la dynamique, les phénomènes aléatoires et la demande aléatoire du système considéré. Prenons comme exemple, une boutique de café brut qui vend seulement du type de café A et B. Les clients y arrivent de façon aléatoire et le stock de café est limité, mais le nombre de fournisseur en café brut est illimité.

Un client arrive pour acheter cinq boîtes de café A placées sur un rayon; le commis constate qu'il en manque et va chercher dans l'entrepôt un nouveau lot de cinq boîtes pour les mettre en place. À ce moment-là, il n'y a plus de café de type A en stock. Un autre client arrive immédiatement après et commande cinq boîtes de type A et cinq boîtes de type B. Le magasinier contacte le fournisseur qui les livre immédiatement. Le délai de livraison est normalement d'une heure, mais malheureusement il y a un encombrement de la circulation et la boutique ne les reçoit qu'après deux heures. Un troisième client arrive une heure après le deuxième client pour acheter cinq boîtes de café de type A; il doit la quitter pour chercher une autre boutique. Supposons que chaque boîte de café vendue rapporte cinq (5) dollars de revenu, tandis qu'on doit payer 1,50$ de plus pour en stocker. La boutique perd alors 5*(5,0$-1,5$) = 17,50$ de revenu pendant une heure à cause de cet embouteillage.

Cet exemple se compose de :

- deux phénomènes aléatoires : l'arrivée des clients et leurs demandes, l'embouteillage et le délai de livraison;
- une chaîne de deux stations : le stock maintenu dans la boutique et le fournisseur;
- le stock maintenu dans la boutique est actif; il est variable par rapport au temps;
- la perte de 17,50$ est le coût de pénalité pour une heure;

---

[2] Cet exemple est tiré de l'ouvrage de Blondel, F. (2007): *Gestion de la prodution*. Dunod., page 291.

- les deux types de café A et B sont les nombres de types de pièces que la ‘‘machine’’ peut fabriquer.

Cet exemple correspond au problème de commande stochastique des systèmes de production où l’arrivée des clients est la demande aléatoire, l’embouteillage correspond aux pannes et aux réparations des machines, une chaîne de deux stations est la ligne de production à deux machines en série. L’action de commande au fournisseur est le Kanban (étiquette). Il faut alors considérer que le café manquant dans le stock est la commande de rétroaction (feedback control) en temps réel, et la quantité de pièces de commande est la variable de décision (taux de production).

De la même façon, nous considérons les systèmes de production flexible (SPF) qui sont soumis à des événements aléatoires tels que des pannes et des réparations de machines, des fluctuations de la demande, des actions de la maintenance, du setup, etc. Ces systèmes considérés sont constitués de plusieurs machines en parallèle ou en série. Le but est de trouver une bonne performance du système par minimisation des niveaux d’inventaire (stock), du temps de cycle (Lead Time), maximisation de la valeur de la machine. En considérant ceci, le problème de la commande optimale stochastique d’un SPF est le suivant: étant donné un SPF dont les états discrets sont caractérisés par un processus de saut dépendant du temps et traduisant l’évolution dynamique de la structure du système, un taux de demande $d(t)$ est aléatoire par rapport au temps. La loi de commande consiste à :

- déterminer le taux de production en satisfaisant la demande sur un horizon fini;
- d’établir l’ordonnancement des pièces dans le système pour satisfaire la politique de production.

Dans ce mémoire, nous nous intéressons principalement au sous-problème qui consiste à déterminer les taux de production.

ÉTAT DE L'ART ET OBJECTIF DU MÉMOIRE

État de l'art

Du point de vue pratique, l'opération du SPF est variable par rapport au temps. Les employés dans l'entreprise doivent donc respecter tous les petits changements du système en satisfaisant des conditions d'opération. Bien qu'ils traitent instantanément des opérations du système, ils peuvent parvenir à de bons résultats.

Du point de vue théorique, l'opération du SPF est dynamique; elle obéit, sur un horizon complété, à des lois quelconques telles qu'un comportement (béhaviorisme), un modèle mathématique, des théories de management et modes, etc.

Qu'elle soit théorique ou pratique, l'opération du SPF ne peut éviter les risques tels que les événements aléatoires. Les praticiens peuvent éliminer les riques dès qu'elles se sont produites, tandis que les théoriciens les considèrent comme des états physiques du système qui sont non-préventifs. Malgré l'existence d'un dilemme entre la théorie et la pratique sur la gestion de la production, les théories ont contribué à plusieurs liaisons entre la gestion et l'opération qui sont basées sur chaque objectif et chaque fonction de l'entreprise, fondées sur les systèmes de production, d'approvisionnements et de demandes. La pratique et la théorie doivent se compléter mutuellement dans la recherche opérationnelle. Par conséquent, cette recherche n'est qu'une continuation du développement théorique. Bien qu'elle nous donne des résultats intéressants, les données du système, dans les exemples numériques, ne correspondent pas à un système de données réel.

Objectif général

L'objectif général de ce mémoire tient compte de deux aspects :

- Aspect théorique. Une généralisation du formalisme de Rishel (1975) dans le cas où le processus de saut serait caractérisé par le processus semi-Markovien et ses taux de transition dépendraient du temps. En effet, les conditions d'optimum obtenues sont décrites par des équations différentielles par rapport au temps et aux états du système considéré.

- Aspect pratique. Sous ce nouveau modèle, la modélisation du système de production flexible est de déterminer les taux de production en satisfaisant les demandes aléatoires.

TABLE DES MATIÈRES

LISTE DES FIGURES

LISTE DES TABLEAUX

LISTE DES ABRÉVIATIONS PRINCIPALES

x_0	Stock initial du produit
x^+	Stock positive du produit
x^-	Stock négative du produit
z^*	Seuil critique ou Hedging Point (Niveau du stock optimal)
$u(t,.)$	Taux de production à l'instant t (variable de contrôle)
$u^*(t,.)$	Taux de production optimal à l'instant t
r	Taux de production maximal
$d(t)$	Taux de demande à l'instant t
$\xi(t)$	Processus semi-Markovien en temps continu et état fini décrivant le système
ρ	Taux d'actualisation
$\mathcal{I}$	Ensemble d'espace d'états du système considéré : $\mathcal{I} = \{0, 1, 2,\ldots\}$
α,β	Modes de la machine : $\alpha,\beta \in \mathcal{I} = \{0, 1, 2,\ldots\}$
λ, μ	Paramètres d'échelle et de forme de la distribution de Weibull
λ_0, σ	Paramètres de la loi log-normal, qui sont respectivement la moyenne et

	l'écart-type du logarithme de la variable
λ_1, μ_1	Paramètres de forme et d'intensité de la distribution gamma
$P_{\alpha\beta}(t)$	Probabilité de transition de la machine du mode (étate) α au mode β (état)
$p_{\alpha\beta}(t)$	Densité de la probabilité de transition de la machine du mode α au mode β
$p(t)$	Taux de panne de la machine à l'instant t
$q(t)$	Taux de réparation de la machine à l'instant t
MTTF	Temps moyen de fonctionement de la machine (Mean Time To Fail)
MTTR	Temps moyen de réparation de la machine (Mean Time To Repair)
$J^{\alpha}(t,x;.)$	Fonction coût total sur un horizon déterministe
$g(x(t),u(t))$	Fonction coût instantané
$v^{\alpha}(t,x)$	Fonction valeur
c^{+}	Coût de stock
c^{-}	Coût de pénurie
m	Nombre de machines
n	Nombre de types de pièces
T	Temps final de l'intervalle d'optimisation considéré

$\mathcal{A}(.)$	Ensemble des commandes admissibles pour $u(t,.$

CHAPITRE 1

INTRODUCTION GÉNÉRALE

Ce chapitre présente une revue de la littérature générale, la motivation de recherche, la méthodologie et les contributions du travail du mémoire. Il est composé de différents sujets n'ayant pas été traités formellement dans chacun des articles de cette étude (chapitres 2 à 3), mais qui permettent néanmoins d'introduire certaines notions complémentaires et d'avoir une meilleure compréhension de la problématique soulevée dans ce projet. Pour terminer, une brève description de l'organisation du mémoire sera présentée.

1.1 Revue de la littérature

Il existait jusqu'ici plusieurs contributions en ce qui a trait à la recherche d'optimisation du système de production flexible (SPF) à partir des modèles mathématiques généraux servant à la modélisation du SPF. Les modèles mathématiques seront introduits dans un premier temps. La deuxième partie présentera la modélisation du SPF en s'inspirant de ces modèles mathématiques.

1.1.1 Problème de commande optimale stochastique

Tout d'abord, nous voulons présenter le problème de commande optimale. Ce problème est né depuis la Seconde Guerre mondiale et est basé sur une méthode mathématique d'optimisation. Le principe du maximum avait été introduit par Pontryagin[3] (1903-1988), qui présentait aussi une commande de type bang-bang, une commande restrictive par des conditions de bord. Dans l'ouvrage publié en 1954 (voir Bellman (1954)), Bellman a présenté une nouvelle technique dite programmation dynamique qui permet de résoudre une

---

[3] L'information de Pontryagin est présentée dans l'encyclopédie en ligne Wikipédia (2010).

classe de problèmes d'optimisation sous contraintes. Cette méthode s'applique à des problèmes d'optimisation dont la fonction objectif se décrit comme suit :

$$J(t_0) = \int_{t_0}^{T} F(.)ds = \int_{t_0}^{t_0+h} F(s)ds + \int_{t_0+h}^{T} F(s)ds, \tag{1.1}$$

où h est suffisamment petit. On définit le terme $O(h)$ comme suit :

$$\lim_{h \to 0} \frac{O(h)}{h} = 0,$$

l'expression (1.1) devient :

$$J(t_0) \cong F(t_0)h + J(t_0 + h) + O(h). \tag{1.2}$$

Ces développements mathématiques permettent d'obtenir des équations différentielles du problème considéré.

Par la suite, les autres contributions relatives à une autre classe de systèmes à états hybridés ont été introduites par Krasovskii et Lidskii (1961) et Lidskii (1963). Ces auteurs ont présenté des problèmes sur un horizon infini dans le cas où l'état du système serait constitué de deux différentes parties; $\boldsymbol{x}(t)$ est l'état continue du système, $\xi(t)$ est l'état discret du système qui est caractérisé par une chaîne de Markov. Puis, Sworder (1969) a traité d'un problème linéaire sur un horizon fini en présentant une nouvelle version du principe du maximum. Le développement de ce principe du maximum a été présenté par Rishel (1975); il a étudié le problème dont les stratégies optimales sont décrites par des lois de commande de rétroaction (feedback control) où les états du système sont dynamiques. La contribution de Rishel a été très importante. Il a bien établi les conditions d'optimum du problème qui sont décrites par des équations différentielles admettant une solution unique. Pour le formalisme de Rishel, le principe de Pontryagin a impliqué les équations différentielles dites équations d'Hamilton-Jacobi-Bellman. Ensuite, les travaux de Davis (1984) ont abordé le même problème que ceux de Rishel. En revanche, dans la contribution de Davis (1984), les états ainsi que les politiques de commande du système sont divisés dans l'échelle du temps (time scale). Ce là pour transformer le problème stochastique à celui déterministe dit processus Markovien déterministe par des morceaux (Piecewise Deterministic Markov process).

Contrairement aux travaux de Rishel (1975) et de Davis (1984) où les sauts de transition du système considéré sont perturbés par des processus Markoviens homogènes; dans la thèse de doctorat de Boukas (1987), les processus Markovien non-homogènes sont utilisés afin d'introduire une variable commandée dans l'état discret du système. Cette extension a contribué, dans la classe de commande optimale stochastique, à une nouvelle version. Dans tous les travaux de Rishel (1975), Davis (1984) et Boukas (1987), les conditions d'optimum, incluant celles nécessaires et suffisamment, sont établies par la méthode de programmation dynamique. Nous demandons au lecteur de se référer à l'ouvrage de Fleming et Soner (2006) pour plus de détails en ce qui a trait le problème de commande avec la chaîne de Markov. Dans Fleming et Soner (2006), on peut aussi trouver une méthode de résolution des équations d'Hamilton-Jacobi-Bellman dite solution de viscosité.

Devant des complexités des structures des conditions d'optimum dans le modèle Markovien proposé dans la littérature, lorsque la taille du système est plus grande, une nouvelle approche dite hiérarchique est présentée par Sethi et al. (1994). En fait, cette approche des perturbations singulières existait depuis les travaux de Lehoczky et al. (1991). Avec cette hiérarchie, la solution des équations d'*Hamilton-Jacobi-Bellman* (HJB) devient plus simple en ce qui a trait le problème simplifié équivalent. Récemment, une autre contribution du problème d'optimalisation avec des processus stochastiques a été présentée par Cao (2007) intitulé *Stochastic Learning and Optimization* (SLO). Dans l'ouvrage de Cao, les lois de commande sont appréciées par l'analyse de sensibilité du système au lieu des stratégies optimales. Malgré l'impossibilité, pour que l'on puisse appliquer directement aux systèmes à états discrets dans le temps réel, l'étude de problèmes d'optimisation des systèmes stochastique (SLO) est appliquée dans les domaines dans le cas où les sauts du système pourraient être considérés comme des pas continus tels que la communication, le mouvement de robots, etc.

1.1.2 Application au SPF

Considérant les modèles mathématiques présentés dans la littérature, depuis plus de trente ans, l'application au SPF visait à mesurer la performance du système telle que les taux de production, les stratégies de la maintenance, la stabilité ainsi que le temps de réparation ou de réglage (setup time), etc. Cette sous-section représente l'application relative à l'ordonnancement de la production, de la maintenance préventive et de la stabilité du SPF. En fait, nous n'avons pas présenté le problème du setup du système parce que l'on considère le SPF dans les cas où la flexibilité serait adaptée parfaitement par la gestion de production assistée par ordinateur (Computer Aided Manufacturing-CAM).

A) Commande optimale des SPF

L'application des théories mathématiques à la commande optimale stochastique des SPF a débuté dans les années 80, notamment dans les travaux d'Older et Suri (1980) dans les cas où les pannes des machines seraient perturbées par des chaînes de Markov; le problème posé est alors l'ordonnancement d'un SPF. Puis, le problème de commande optimale stochastique appliqué aux grands SPF constitués de plusieurs stations de travail (plusieurs machines) a été introduit par les travaux de Kimemia et Gershwin (1983); l'objectif était de trouver des taux de production par minimisation des coûts d'inventaire. Kimemia et Gershwin ont proposé une solution générale associée au concept du seuil critique (hedging point), c'est le point auquel la production du système est optimale. En utilisant la théorie de fille d'attente, Tsitsiklis (1984) a étudié le même problème appliqué au système à trois machines en série. Il a montré que les politiques obtenues sont optimales si et seulement si la fonction coût, l'ensemble des commandes admissibles, ainsi que les espaces d'état sont convexe.

Par la suite, en utilisant un processus Markovien, Akella et Kumar (1986), pour un système à une machine traitant un seul type de produit et soumise à des pannes, ont obtenu la politique optimale de type seuil critique exacte par minimisation d'un critère de coût actualisé sur un horizon infini. Cette politique dépend d'un niveau optimal de l'inventaire z^*. Le même problème a été étudié par Bielecki et Kumar (1988). En fait, Bielecki et Kumar ont utilisé le modèle de file d'attente de la forme *M/M/1* et de la fonction coût linéaire pour déterminer le taux de production. En obtenant une politique optimale, ces auteurs ont montré que le seuil critique est stable si le taux d'utilisation du système est inférieur à 1.

Les extensions de la stratégie de commande de type seuil critique sont également discutées par Sharifnia (1988), Malhamé et Boukas (1991). Srivatsant (1993) a présenté une solution exacte du problème pour un SPF à une machine traitant deux types de pièces. En ce qui a trait le problème du seuil critique du SPF à plusieurs types de pièces, celui-ci a été présenté par Perkins et Srikant (1997), Shu et Perkins (2001), et Perkins (2004). Une autre contribution au problème du seuil critique a été présentée par Ciprut et al. (1998); ces auteurs ont directement utilisé les modèles de file d'attente de types $M/G/\infty$, et $G/M/\infty$ de Miller (1963) pour résoudre le problème d'optimisation. Même si on avait obtenu des résultats intéressants au problème du seuil critique, celui-ci devient très complexe lorsque les SPF traitent plusieurs types de pièces. Par conséquent, le problème pourrait devenir celui des polynômes non-déterministes NP; l'opération du SPF pourrait alors être chaotique[4].

Comme nous l'avons mentionné ci-dessus (sous-section 1.1.1), Boukas (1987) a contribué à une nouvelle version du problème de commande optimale stochastique par l'extension du modèle de Rishel. Cette version permet de modéliser un SPF dans le cas où les taux de transition incluraient l'usure des machines associées aux actions de maintenance préventive de la machine. Dans les travaux de Boukas, la distribution des probabilités de pannes d'une machine dépend de l'état du système (âge de la machine). Par la suite, Boukas et al. (1995) ont montré que le seuil critique dépend de l'âge de la machine contrairement aux modèles d'Akella et Kumar (1986) et de Bielecki et Kumar (1988). Les prochains paragraphes discutent du problème de maintenance préventive.

[4] Tous les théoriciens et les praticiens veulent éviter l'opération conduisant en chao.

En ce qui concerne les systèmes complexes et larges, afin de simplifier le problème de commande, chaque événement aléatoire est considéré comme un sous-problème. En conséquence, le problème de commande considéré devient plusieurs sous-problèmes selon le nombre d'événements aléatoires dit commande hiérarchisée. Selon cette technique proposée par Gershwin (1989), la fréquence est une proportion inverse du temps de vie entre deux événements consécutifs et le classement de niveau est à proportion de la fréquence. Soit f_1 la fréquence du niveau 1, f_k ($k > 1$) celle du niveau k, alors on a : $0 \leq f_1 << f_2 <<\ldots f_k << \infty$. Exemple : (1) une machine est en panne une fois par 1000 heures, (2) l'action de maintenance préventive est prise une fois par 50 heures, (3) taux de production est 1 unité/heure; on a alors $f_1 = 1/1000 << f_2 = 1/50 << f_3 = 1$. La prise de décision est consécutive à partir du niveau 1 jusqu'au niveau k. Cette idée fut adoptée afin de réduire la taille du problème de commande des systèmes larges dans les travaux de Sethi et al. (1994-1997) et de Kenné et Boukas (2003). Gershwin (2002) s'inscrit dans une démarche similaire.

Dans les mêmes systèmes complexes et larges, le problème de commande est considéré sur un horizon infini sans taux d'actualisation ρ, la fonction objectif (coût total) peut être infinie. Pour aboutir aux conditions d'optimum, une autre approche est introduite dans certains travaux, par exemple Bertsekas (2001), Sethi et al. (2005), ainsi que Presman et al. (2002). Cette approche offre la possibilité d'optimiser l'espérance des coûts moyens, dite commande optimale avec coût moyen par étage (en anglais, *Average cost control*). Même si celle-ci est fondée sur la méthode de programmation dynamique, les résultats obtenus du problème de commande optimale ne peuvent plus s'appliquer au système réel car les variables de décision sont quantitatives.

B) La performance des lignes de production

Dans les lignes de production, le système permet de fabriquer une masse de production en tenant compte de la haute performance. La performance est mesurée par le taux de production et les quantités de stocks dans le système. L'opération du système donne des flux physiques desquels les états du système dépendent. De surcroit, les machines peuvent tomber en panne de façon aléatoire, ce qui perturbe la dynamique de ces flux. Pour modéliser cette dynamique, on doit considérer des états du système comme des événements discrets qui sont modélisés à l'aide de la théorie des files d'attente. Cette théorie est présentée de façon détaillé dans l'ouvrage de Kleinrock (1975) et son application est décrite dans Altiok (1997). Une des méthodes importantes pour étudier des lignes de production à grande taille est celle de la décomposition. Cette méthode a été introduite pour la première fois par Gershwin (1987) pour le système synchrone. Par la suite, Dallery et al. (1989) l'ont amélioré pour appliquer le système asynchrone. Une solution exacte du système à deux machines en série est présentée dans Gershwin (2002). Le même problème est étudié par Kim et Gershwin (2005,2008), Tan et Gershwin (2009), ainsi que Ciprut et al. (2000).

Quant aux politiques de contrôle des lignes de production, il existe certaines méthodes telles que Kanban, CONWIP (*constant work-in-process*), blocage minimal (en anglais, *minimal blocking*), politiques hybridées (combinaison de Kanban et CONWIP), etc. Dans la mémoire doctorale de Bonvik (1996), celui-ci a montré que la politique hybridée est plus efficace que les autres. Mais Bonvik l'a appréciée par la méthode de simulation, qui n'est pas un modèle exact. Par contre, en formulant un modèle mathématique, Sethi et al. (1997) ont montré que la politique de rétroaction (en anglais, *feedback policy*) est beaucoup plus efficace que celle de Kankan. Nous ne voulons pas présenter en détail les lignes de production car elles ne constituent qu'une partie de l'objet de notre recherche.

C) Problème de maintenance préventive

En plus de planifier la production, on peut gérer la maintenance préventive dans le cas des systèmes vieillissants à cause du temps, de la fatigue, de la corrosion, des pannes, de l'instabilité, etc. La théorie de la fiabilité sur la maintenance préventive était développée depuis la Seconde Guerre mondiale durant laquelle elle était appliquée aux armes militaires[5]. Elle s'applique actuellement à plusieurs systèmes tels que ceux de la communication, de la technique de l'information et de la production.

---

[5] La théorie de la fiabilité est apparue à la même époque que celle de la commande optimale. On peut trouver une description détaillée de l'histoire de ces théories dans l'ouvrage de Nakagawa, T. (2005): *Maintenance theory of reliability*. Springer.

En ce qui a trait au système de production, le problème de maintenance préventive est de réduire les risques nuisant au système par élimination des pannes de la machine, et par l'augmentation des fonctionnements de la machine. L'objectif est de trouver une politique afin que la machine soit remplacée par une nouvelle ou qu'elle soit entretenue le plus tôt possible. La formation de ce problème consiste à optimiser les coûts totaux, incluant le coût d'inventaire, le coût d'investissement et la valeur de la machine. En conséquence, le problème de maintenance préventive est considéré comme un problème de commande stochastique du système considéré. Dans la littérature, une première version de ce problème a été présentée dans les travaux de Boukas (1987) et de Boukas et Haurie (1990); ces auteurs ont considéré la distribution des probabilités de panne d'une machine selon son âge. Ce formalisme fut développé par la suite par Boukas et Yang (1996), Boukas et Kenné (1997), Kenné et al. (2007) et Dehayem et al. (2009).

Dans la contribution de Kamien et Schwartz (1971), les auteurs proposent un modèle de maintenance préventive dans le cas où le taux de panne de la machine dépendrait de son âge dans le domaine temporel. L'objectif est d'éliminer le taux de panne si nécessaire par maximisation de la valeur de la machine. En effet, la politique optimale permet de renouveler ou de remplacer par une nouvelle machine lorsque la valeur de la machine est très faible. Malgré le modèle de Kamien et Schwartz (KS) soit en temps continu, il ne permet plus de répondre quand l'action de maintenance est prise. Afin de résoudre ce problème, Bensoussan et Sethi (2007) ont révisé le modèle de KS et proposé une solution du problème de temps d'arrêt. Cette extension offre la possibilité de trouver la stratégie de commande qui peut répondre aux deux importantes questions suivantes : quand la machine doit-elle être remplacée ou renouvelée? et quelle situation doit-on mettre en place pour que l'action de la maintenance soit prise? Nous ne voulons pas présenter en détail le problème de maintenace préventive car il ne constitue pas l'objet de notre recherche.

Nous avons brièvement présenté dans cette section des travaux sur l'ensemble de la théorie et de son application aux systèmes manufacturiers. Cela fait l'objet de notre recherche, qui s'articule autour du problème de commande optimale stochastique, de la modélisation des systèmes manufacturiers incluant des lignes de production.

1.2 Motivation de la recherche

Nous observons que dans les modèles Markoviens homogènes et non-homogènes, les taux de transition entre des états sont indépendants du temps et qui sont caractérisés par des distributions de probabilité exponentielles. Ces distributions sont construites par des hypothèses simplificatrices [6]. En outre, les distributions de probabilité exponentielles possèdent le coefficient de variation égale un ($CV = 1$), qui d'une part influe inutilement la performance du système, et d'autre part limite immédiatement une réponse aux fluctuations de la demande. Ce sont les limites des modèles classiques. Il est donc indispensable de formuler un modèle dans les cas où les sauts aléatoires seraient perturbés par le processus semi-Markovien. Un nouveau modèle de ce type s'oppose aux modèles classiques de la littérature. Pour suivre le format de mémoire exigé, le détail de la motivation de notre recherche est présenté au fil des chapitres, avec pour lien conducteur l'objectif suivant : la motivation générale concerne la physique du système considéré; l'espace et le temps sont toujours des quantitatives dynamiques lorsque le système est en opération.

1.3 Méthodologie

Cette section présente une synmémoire de la méthodologie théorique et pratique qui est utilisée dans le cadre de ce projet. Cette dernière peut se diviser en deux grandes étapes qui sont : (1) formuler un modèle mathématique du problème de commande optimale stochastique et (2) procéder à la modélisation des systèmes manufacturiers s'inspirant de ce nouveau modèle. La première étape est de formuler le problème considéré sur un horizon fini (horizon déterministe) pour deux cas: l'un est le problème de commande en absence du taux d'actualisation ($\rho = 0$); l'autre est le problème de commande en présence du taux

---

[6] Cette conclusion a été présentée dans l'ouvrage de Ross (2003).

d'actualisation ($\rho > 0$). La deuxième étape est de modéliser des systèmes manufacturiers associés aux problèmes de commande de production satisfaisant les fluctuations de la demande.

Ainsi, les principaux développements de notre démarche sont les suivants :

Formulation mathématique (deux étapes)

E.1. Formuler le problème de commande optimale stochastique sur un horizon fini en absence du taux d'actualisation ($\rho = 0$) au cas où la dynamique discrète du système serait caractérisée par un processus semi-Markovien en temps continu. En utilisant les hypothèses dans les travaux de Rishel (1975), l'approche consiste à construire un nouveau modèle basé sur la *programmation dynamique*[7]. Cette étape nous permet d'établir les conditions d'optimum en présence du temps, lesquelles garantissent l'existence et l'unicité des lois de commande optimales associées à la commande de type rétroaction (feedback control).

E.2. Extension du modèle de l'étape E.1 vers le problème de commande optimale stochastique en présence du taux d'actualisation ($\rho > 0$).

Modélisation des SPF (trois étapes)

E.3. Modéliser un SPF à plusieurs machines en parallèle traitant un type de pièce en s'inspirant du nouveau modèle dans l'étape E.1. L'objectif est de trouver les taux de production en satisfaisant les fluctuations de la demande.

E.4. Modéliser un SPF à deux-machine en tandem (en série) traitant un type de pièce, et un SPF à une machine traitant deux types de pièces, en s'inspirant du nouveau modèle dans l'étape E.2. L'objectif est de trouver les taux de production en satisfaisant les fluctuations de la demande.

---

[7] L'avantage de cette méthode sera présenté dans l'annexe VII.

E.5. Appliquer l'approche [8] de Kushner et Dupuis (2001) au problème déterministe équivalent dont les conditions d'optimum issues des étapes E.3, et E.4. Cette résolution permettra d'obtenir des décisions semi-Markoviennes dans l'espace d'états et le domaine temporel.

Comme nous l'avons mentionné dans la problématique de recherche, le but de cette recherche est de répondre aux questions suivantes :

Q.1. Pouvons-nous généraliser le formalisme de Rishel dans les cas où le processus de perturbation serait commandé et où ses taux de transition de transitions dépendraient du temps?

Q.2. Pouvons-nous résoudre numériquement des conditions d'optimum par l'approche de Kushner et Dupuis?

1.4 Contribution du travail de mémoire

À partir de la motivation de la recherche traitant des aspects théoriques et appliqués pour les deux propositions mentionnées (E.1-E.2), nous pouvons préciser que la contribution de cette recherche correspond à la réponse des deux questions précédentes (Q.1-Q.2). Cette contribution se compose de deux aspects: l'un est théorique, l'autre pratique.

L'aspect théorique correspond aux étapes E.1-E2 pour répondre à la question Q.1. Il consiste à formuler le problème de commande optimale stochastique dans les cas où le processus de perturbation serait caractérisé par un processus semi-Markovien et où ses taux de transition dépendraient du temps. L'aspect appliqué correspond aux étapes E.3-E.5 pour répondre à la question Q.2. Cet aspect est constitué de deux parties : La première partie est l'étude des applications du nouveau modèle en ce qui concerne les problèmes de planification de la production des SPF. La deuxième partie se propose de résoudre des conditions d'optimum à

---

[8] L'algorithme de cette approche sera présenté dans les Annexes III et IV.

l'aide des méthodes numériques appliquées au problème de commande optimale stochastique proposée.

En général, la contribution de ce travail de mémoire peut représenter un nouveau modèle dont les conditions d'optimum sont décrites par des équations d'Hamilton-Jacobi-Bellman :

$$\rho v^{\alpha}(t,x) = \min_{u(t,x)\in \mathcal{A}(t,\alpha)} \left\{ g(.) + v_t^{\alpha} + v_x^{\alpha}\left(u(t)-d(t)\right) - p_{\alpha\alpha}(t) v^{\alpha}(t,x) + \sum_{\beta\neq\alpha} p_{\alpha\beta}(t) v^{\beta}(t,x) \right\}, \quad (1.3)$$

où le taux de demande $d(t)$ fluctue au cours du temps, les dérivées de probabilité de transition de l'état α au β, noté $p_{\alpha\beta}(t)$, dépend du temps t et sont non-exponentielles contrairement aux modèles classiques Markoviens. Le détail de cette expression sera présenté dans la section 3.3 du Chapitre 3.

Les contributions de ce mémoire sont composées de deux (2) conférences et de la rédaction de (2) articles de revues (chapitres 2 et 3). Les articles de conférences sont référencés par :

1. Thang T. Diep, J.P. Kenné and T.M Dao (2007), ``*Queuing theory based hedging point for non-Markovian manufacturing system*``. Conference on Systems and Control. May 16-18 Marrakech (Maroc). (6 pages).
2. Thang T. Diep, T.M Dao, S. Abou (2009), ‘ *Real-time control of stochastic manufacturing systems with jump semi-Markov*''. Conference on Identification, Control and Applications, IASTED, Aug 17-19, Honolulu, Hawaii, USA. (6 pages).

1.5 Organisation du mémoire

Ce mémoire a ainsi été orienté autour de deux articles scientifiques qui sont inclus intégralement aux chapitres 2, et 3. Chaque chapitre est accompagné de résumé en français qui est placé au tout début du chapitre.

Le premier article, intitulé «**Optimal Control of Production on a failure-prone machine system with jump semi-Markov**», a été soumis à la revue *IIE Transaction* en Mars 2010 avec la référence **UIIE – 1979**. L'objet principal de cet article est d'une part d'intégrer la formation du problème de commande optimale stochastique sur un horizon fini sans taux d'actualisation ($\rho = 0$); et de présenter d'autre part la modélisation du système de production flexible à plusieurs machines identiques en parallèle traitant un type de pièce.

1. La formation du problème considéré a été utilisée dans les hypothèses de Rishel (1975) pour modéliser des éventualités du système par des sauts semi-Markoviens nommés la dynamique discrète du système. Cette formulation a été adoptée par la méthode de programmation dynamique pour établir des conditions d'optimum. Celles-ci sont décrites par les équations d'Hamilton-Jacobi-Bellman. Le problème de commande obtenu est la commande de rétroaction.
2. La modélisation du SPF à plusieurs machines en parallèle est de déterminer des taux de production en satisfaisant les demandes aléatoires. Selon la politique de type seuil critique, nous avons analysé la production cumulative[9] sur un court-terme, ce qui correspond à la productivité dans le temps réel.

Le deuxième article s'intitule «**Feedback optimal Control for two-machine flowshop**» a été publié dans la revue *International Journal of Industrial Engineering Computations,* Juin 2010, pp. 95-120. L'objet principal de cet article est une extension du modèle du première article au cas où le taux d'actualisation serait non-zéro ($\rho > 0$). En s'inspirant de ce nouveau modèle, la modélisation du système de production flexible à deux-machine en tandem (en série) traitant un type de pièce est de déterminer les taux de production de deux machines.

Nous concluons ce chapitre par la présentation des relations entre les chapitres, les étapes, la théorie et l'application dans la figure 1.1. Notons que l'annexe IV présente la méthode numérique basée sur l'approche de Kushner et Dupuis (2001); l'annexe V présente l'exemple

---

[9] En anglais, cette analyse s'appelle « *target production with hedging point policy* ».

numérique d'un SPF à une machine traitant deux types de pièces; l'annexe VI présente la relation entre les modèles Markovien et semi-Markovien ainsi que les méthodes d'optimisation appliquées au problème de commande; et l'annexe VII présente brièvement la structure du système manufacturier en tenant compte de l'étude de notre mémoire.

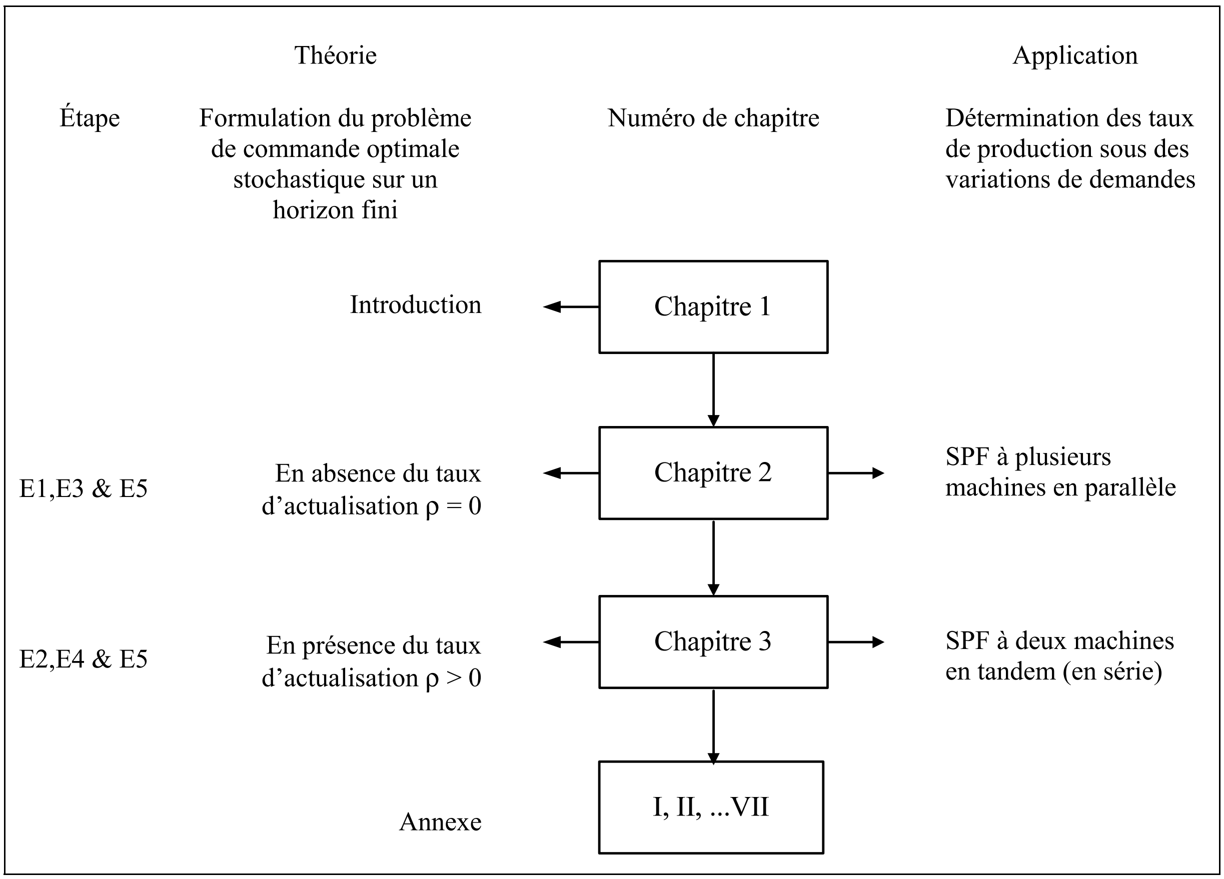


Figure 1.1 Relations entre des chapitres, des étapes, la théorie et l'application.

CHAPITRE 2

ARTICLE # 1 OPTIMAL CONTROL OF PRODUCTION ON FAILURE-PRONE MACHINE SYSTEMS WITH SEMI-MARKOV JUMP

Thang Diep [1], Seraphin C. Abou [2], Thien-My Dao [1]

[1] Mechanical Engineering Department, Ecole de Techologie Supeieure,

1100 Notre-Dame St. West, Montreal, (Quebec), Canada H3C1K3

[2] Department of Mechanical and Industrial Engineering, University of Minnesota Duluth,

1305 Ordean Court, Duluth, MN 55812, USA

Résumé

L'objet principal de cet article est d'une part d'intégrer la formation du problème de commande optimale stochastique sur un horizon fini sans taux d'actualisation ($\rho = 0$), et de présenter d'autre part la modélisation du système de production flexible à plusieurs machines identiques en parallèle traitant un type de pièce.

Nous avons formulé le problème considéré en utilisant les hypothèses de Rishel (1975) et des sauts semi-Markoviens. Cette formulation a été adoptée par la méthode de programmation dynamique pour établir des conditions d'optimum. Celles-ci sont décrites par les équations d'Hamilton-Jacobi-Bellman (HJB). Le problème de commande obtenu est la commande de rétroaction. L'application de ce nouveau modèle au SPF à plusieurs machines en parallèle traitant un seul type de pièce est de déterminer des taux de production en satisfaisant les demandes aléatoires. Selon la politique de type seuil critique, nous avons analysé la production cumulative sur un court-terme, ce qui correspond à la productivité dans le domain temporel. La résolution des équations HJB est proposée à l'aide des méthodes numériques basées sur l'approche de Dupuis et Kushner (2001).

Abstract

In this paper, we formulate an analytical model for an optimal production problem of multiple machines in parallel producing single part-type systems. The formulation is a generation of the formalism of Rishel - which is the Markov framework. In the considered production system, each machine is subject to random failures and repairs. The proposed model assumes that each machine's times to failure and times to repair are non-exponential distribution, i.e., the random failures and repairs are characterized by semi-Markov jumps.

The objective of the control problem is to find the production rates $u(t)$ that meet the demand rates $d(t)$ by minimization of the expected cost of inventory/shortage. Based on Bellman principle, the optimality conditions obtained satisfy the Hamilton-Jacobi-Bellman equation, which leads to a feedback control. The new proposed model whose coefficient of variation in the case of non-exponential distributions is less than one ($CV_{up/down} < 1$) can improve the Markov's model with ($CV_{up/down} = 1$). Numerical methods are used to solve the optimality conditions, and result analysis show that by using the proposed models, machines can satisfy the varying demands in time.

Keywords: Semi-Markov process, Manufacturing system, Hamilton-Jacobi-Bellman equation, Numerical method.

2.1 Introduction

In this paper, we consider a manufacturing system consisting of multiple identical machines in parallel, producing a single part-type. Each machine is subject to random failure and a repair, i.e., each machine has two states: available (up) and unavailable (down), with non-exponential distributions of up- and downtimes. The objective of this study is to formulate an optimal stochastic control problem of production systems in order to determine the production rates $u(t)$ with minimizing the expected cost of inventory/shortage in finite horizon. This section presents a review of the literature, the motivation for using the semi-Markov process, and the study's contribution.

2.1.1 Literature review

The most difficult problem encountered with production planning for a manufacturing system is optimal stochastic production control; the manufacturing system is subject to random events, such as operations, failures, as well as fluctuations in raw material supply and customer demand. In fact, over the last few decades, the optimal production control of stochastic manufacturing systems has been considered under the Markov process. This approach proposed by many authors in the literature uses a special class of *piecewise deterministic system* (PDS). A class of stochastic models was developed in the 1960s for optimal control problems (Krasovskii and Lidskii (1961), Lidskii (1963), and Sworder (1969)), after which Rishel (1975) formulated a stochastic optimal control problem using Markov jumps. Based on Rishel pioneering work, Davis (1984) presented solutions for control problems whose dynamics describe the characteristics of processes between jumps. Then Kimemia and Gershwin (1983), Akella and Kumar (1986), Bielecki and Kumar (1988), Sharifnia (1988), Liberopoulos and Caramanis (1994), and Sethi et al. (2005) solved optimal production control problems in manufacturing systems without preventive maintenance. Further, various methods for developing a suitable production solution, such as minimizing of inventories and lead time versus the Just-in-Time (JIT) concept, were introduced. Another contribution to the area of preventive maintenance was introduced by Boukas and Haurie (1990). As the natural probability of machine failure increases with age these works constitute an extension of Rishel's formalism to non-homogeneous Markov jumps in order to deal with the non-smoothness of the value function of maintenance modes.

It should be noted that the transition rates between different states of the above-mentioned Markov models are constant, i.e., they do not depend on time. Clearly, at any given point, the dynamics of physical systems is only influenced at time t, but however, it is not influenced by its previous state at ($t-\delta t$) or by the future state at ($t+\delta t$) (Williams (2007)). Moreover, in the Markov model, the up- and downtimes of the machine obey exponential distributions that lead to the coefficient of variation, (CV_{up}, CV_{down}), being equal to one (Enginarlar et al. (2005), Li and Meerkov (2005a,b)). As a result, the exponential distributions approach may be replaced by a more appropriate method, such as the non-exponential distribution function proposed in the semi-Markov jump model, for instance. The problem of the suitability of the semi-Markov process for production planning and preventive maintenance has been addressed by Love and Zitron (1998), Dimitrakos and Kyriakidis (2008), and Nodem et al. (2009). They considered the problem of optimal preventive maintenance policies under conditions in which the machine is subject to failure within the framework of a discrete semi-Markov decision as the machine's failure rate depends both on its real age and on the number of failures.

Both Markov and semi-Markov framework, the optimality conditions established lead to a feedback control. This is because the feedback control is indispensable to handle the inaccuracies and uncertainties (including stochastic phenomena) that are present in design process, and to make full use of the capacity of the equipment (see e.g Engell (2007)).

In this study, which examines the optimal production problem, we assume that the up- and-downtimes follow known continuous distribution functions (e.g., Weibull, Log-normal, and Gamma). Thus, we make use of the semi-Markov definition given in Becker et al. (2000), the assumptions in Rishel (1975), and dynamic programming methods, to formulate the proposed optimal control model.

2.1.2 Motivation for using semi-Markov process

This paper is motivated by two factors:

1. From a practical point of view, a more general random process model provides computational evidence that *suitably* describes the lifetime of a machine (Grabsky (2003)). That means the impact of machine aging can be presented by assuming a non-

exponential distribution for the machine's up- and downtimes with a coefficient of variation ($CV_{up/down}$) less than one (see Enginarlar et al. (2005), and Li and Meerkov (2005a,b)). Thus, the machines may be said to be *aging over time without any restriction.*

2. Non-exponential up-and downtime distribution issues can be considered by extending the dynamic programming method using semi-Markov jumps. As a result, a unified model which includes the production planning strategy can be developed.

2.1.3 Contribution of this paper

This paper uses the semi-Markov jumps approach without discount rate to develop a new model for the optimal stochastic control problem of failure-prone machine in finite time horizon. Under assumptions stated by Rishel (1975), this proposed model relies on the dynamic programming approach to develop the semi-Markov jumps model, in which the transition rates and transition probabilities depend on time. The application of the new model and related optimality conditions to (i) a single part-type and single-machine system and (ii) a single part-type and two-machine system use log-normal, Weibull, and gamma distributions to present the distributions of the uptime and downtime of the machine.

The next sections are organized as follows: Section 2.2 presents the problem statement for a general system of *m* identical machines in parallel. Section 2.3 describes the optimality conditions, while Section 2.4 presents the optimal feedback control in real time. Section 2.5 presents the application of the proposed model for a single-machine and single part-type production system. A practical case study on a two-machine in parallel, single part-type system is presented in Section 2.6. Finally, concluding remarks are presented in Section 2.7.

2.2 Problem statement

Consider a manufacturing system consisting of *m* identical machines in parallel (the configuration of a parallel machine has been studied by some authors, for example Kimemia and Gershwin (1983), Kang and Shin (2010)...) The machines are subject to random failures and repairs, and can produce a single-part type, as shown in Figure 2.1. Each machine has two possible operational states: at any given time, the operational mode of the machine *j* is described by a stochastic process $\{\xi_j(s)\colon 0 \leq t \leq s,\ j = 1,...,m\}$; an *up* state in which the machine is fully functional is defined as ($\xi_j(s) = 1$), while the *down* state in which it cannot produce any usable output is (($\xi_j(s) = 0$)).

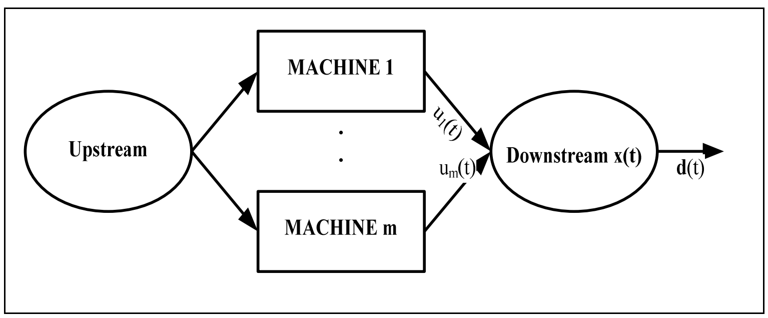


Figure 2.1 Multiple-machine, single part-type system.

Let the machine state variable be: $\xi(s) = \{\xi_1(s),..., \xi_m(s)\} \in \mathcal{I} = \{0, 1\}^m$ for $t \leq s \leq T$. At a given state, the machine operates in a distributed manner before jumping to another state. For a large time interval, the average time between failures of the machine is quantified as the *Mean Time To Fail* (MTTF) at the *up state*, while the average time required to repair a failure when the machine is at the *down state* is termed the *Mean Time To Repair* (MTTR). In the sequel of this analysis, we define $x_j(s) \in \Re = (-\infty, +\infty)$, $u_j(s) \in \Re^+ = [0, +\infty)$ and $d(s) \in \Re^+$ as the buffer level, the production rate, and the demand rate on machine *j* (denoted M_j) at time *s*, respectively. Thus, using the pull model, the following differential equation is used to characterize the system dynamics:

$$\sum_{j=1}^{m}\frac{dx_j(s)}{ds}=\sum_{j=1}^{m}u_j(s)-d(s);\ x_j(t)=x_j,\ \xi_j(t)=\alpha_j;\alpha_j=0,1, \tag{2.1}$$

where $j = 1,..,m$; $d(s)$ is a random variable at time s, α_j is initial state of machine at time t, x_j is also initial conditions of $x_j(.)$ at time t.

Let ν_j, $j = 1,\ldots,m$ be the time unit that each type (product) requires at machine j before it leaves the system (measured in sec, minute, hour, or day...). This is termed the *processing time* of M_j. Thus, the *contraints on controls* are give by:

$$\nu_j.u_j(s)\le \text{Ind}\left\{\xi_j(s)=1: t\le s\le T\right\},\ j = 1,\ldots,m, \tag{2.2}$$

$$u_j(s)\le r,\ j = 1,\ldots,m, \tag{2.3}$$

where $\text{Ind}\left\{\xi_j(s)=1: t\le s\le T\right\}$ is indicator for the machine j in *up state* (available), and r is the maximum production rate on any machine.

Since these processes $\{\xi(s)$: $t \le s \le T\}$ consist of 2^m states, and their duration is arbitrarily distributed within each state, let us define the semi-Markov process with the jump by machine j from state α_j to state β_j by transition probabilities (Becker et al. (2000)) as follows:

$$P_{\alpha_j\beta_j}(s)=\Pr\left[S_n-S_{n-1}\le s\cap\xi_j(S_n)=\beta_j \mid \xi_j(S_{n-1})=\alpha_j\right], \tag{2.4}$$

where S_n is the time of next transition and S_{n-1} the time of last transition ($S_0 = 0$) with respect to $t \le s \le T$.

Consider the following case: if the system enters a state γ, a number of independent times T_γ with distribution functions $F_\gamma(s)$, and probability density function $f_\gamma(s)$ for $\gamma = 0, 1, 2...$, the system will go to state β, if the realization of T_γ is the smallest of all these variables, and the sojourn time in state α will be this smallest realization. The derivative of the semi-Markov transition probability $p_{\alpha\beta}(s)$ will then be given by:

$$p_{\alpha\beta}(s)=\frac{dP_{\alpha\beta}(s)}{ds}=\prod_{\gamma\neq\beta}^{2^m}\left(1-F_{\gamma}(s)\right)f_{\beta}(s);\gamma=0,1,...,\alpha,...,2^m-1. \tag{2.5}$$

Let $g(.)$ denote the running cost function of surplus and production. The cost function J, in the deterministic time interval $[t_o, T]$ (i.e., finite horizon, $0 \le t_o < T$), is defined as follows:

$$J^{\alpha}\left(t,x;u(.)\right)=E_u\left[\int_t^T g\left(x(s),u(s)\right)ds \mid x(t)=x,\xi(t)=\alpha\right], \tag{2.6}$$

where $E_{\boldsymbol{u}}$ stands for the mathematical expectation with respect to the measure induced by the control law $u(s,x) = (u_1 + u_2 +...+ u_m) \in \Re^+$, $x(s) = (x_1 + x_2 +...+ x_m) \in \Re$ for $t \le s \le T$. Note that $g\left(x(s),u(s)\right)=\sum_{j=1}^{m}\left(c_j^+x_j^+(s)+c_j^-x_j^-(s)\right)$ is the running surplus cost with c_j^+ the unit surplus cost and c_j^- the unit backlog penalty of the product at M_j, where $x^+(s) = \max(x(s), 0)$ and $x^-(s) = \max(0,-x(s))$.

The function (2.6) is called the *surplus cost function* and given that we start it at time t in state x. For the sake of simplicity, we make the following assumptions in describing the manufacturing system:

Hypothesis 1. *The manufacturing Lead Times are considered only with respect to the processing time while the setup time, the move time, and the queue time are neglected. All operating machines start their operations at the same time.*

Hypothesis 2. *The demand rate is considered for both varying and constant variables.*

We assume that the planning horizon ($T - t$) is wide enough such that the lead time of the system must be more than or equal to the sum of the *Mean Time to Fail* (MTTF) and the *Mean Time to Repair* (MTTR).

Definitions 2.2.1.(1) *A control* $\mathcal{U}(t,x)=\{u(t,x) \geq 0,\ v_j u_j(t,x) \leq 1\}$, *for* $j = 1,\ldots,m$ *is admissible*;

(2) *A control* $\mathcal{A}$ *is the set of admissible controls with initial value* $x(t) = x$..

Our aim is to obtain an admissible control $u(.)\in \mathcal{A}(t,\alpha)$ that minimizes the cost function (2.6) in which the characteristic lifetime of machines obey the non-exponential distribution. In Section 2.3, we will build the optimal production control model that satisfies constraints in (2.1) and minimization of (2.6).

2.3 Optimality conditions

Based on the production problem defined above, we assume that the production rate can be adjusted during production runs. Under appropriate conditions, the optimal production policy aims to satisfy (2.1), (2.2), (2.3), (2.5), and (2.6) in order to determine the optimal production rate $u(t,x)$ which minimizes the cost function described in (2.6). This policy is characterized by a target production level, and is subject to capacity constraints. We denote the value function $v^{\alpha}(t, x)$ as follows:

$$v^{\alpha}(t,x) = \min_{\substack{u(s,x)\in\mathcal{A}(s,\alpha)\\ t\leq s\leq T}} J^{\alpha}(t,x;u) = \min_{\substack{u(s,x)\in\mathcal{A}(s,\alpha)\\ t\leq s\leq T}} E_u\left[\int_t^T g\big(x(s),u(s,x)\big)ds \mid x(t)=x, \xi(t)=\alpha\right] \quad (2.7)$$

where $t \leq s \leq T$, α is initial state of machine at time t for $\alpha = 0,1,\ldots,2^m - 1$, x is also initial conditions of $x(t)$.

Based on the dynamic programming principle, the following theorem is used for the generalization of the value function in (2.7):

Theorem 2.3.1 *The stochastic control problem satisfies the system of partial differential equations:*

$$0=\min_{\boldsymbol{u}(t,\boldsymbol{x})\in\mathcal{A}(t,\alpha)}\left\{g\left(x(.),u(.)\right)+v_t^{\alpha}+v_x^{\alpha}f^{\alpha}\left(t,x,u\right)-p_{\alpha\alpha}(t)v^{\alpha}\left(t,x\right)+\sum_{\beta\neq\alpha}p_{\alpha\beta}(t)v^{\beta}\left(t,x\right)\right\}. \tag{2.8}$$

At time t, the initial and boundary conditions are satisfied:

$$\begin{cases}\left(x(t),\xi(t)\right)=\left(x,\alpha\right) \quad \text{for} \quad \left(t,x\right)\in Q,\alpha=0,1,...,2^m-1,\\ v^{\alpha}\left(T,x(T)\right)=0,\end{cases} \tag{2.9}$$

where $f^{\alpha}\left(t,x,u\right)=\dot{x}=u\left(t,x\right)-d\left(t\right)$*, α is the state index, and the terms* $v_t^{\alpha}(.)$ *and* $v_x^{\alpha}(.)$ *denote the gradient of the value function with respect to time t and the state variables x, respectively, and* $Q=\left[t_0,T\right]\times\Re$.

Proof. The proof of this theorem is presented in Appendix II.

The partial differential equations (2.8) system is well known as the Hamilton-Jacobi-Bellman (HJB) equation. It depends not only on the state variable of the system *x*(.) but also on its time variation, due to the characteristics of the decision process and the derivative of the semi-Markov transition probability $p_{\alpha\beta}(t)$. In addition, it describes an optimal feedback control process.

Obviously, many different performance indices can be considered as the objective functions in the dynamic programming method formulation in the problem, and all these dynamic programming problems may be solved by taking into account only the constraints related to enable transitions between states. This is because we know that the speed of transitions that are not enabled is zero. More importantly, $v^{\alpha}(t,x)$ provides a unique solution for all *x* and *t* in the relevant ranges; that is equal to the minimum expected total cost among appropriately defined classes of the admissible control law of the system:

Theorem 2.3.2. *Let* $v^{\alpha}(t,x) \in \mathcal{I} \times Q$ *be a solution to* (2.8). *Then for all* $(t,x) \in Q$:

(i) $v^{\alpha}(t,x) \leq J^{\alpha}(t,x;u)$ *for every admissible control system* $u(t,x)$;

(ii) if there exists an admissible system $u^{*}(t,x)$ *such that*

$$u^{*}(t,x) \in \arg \min_{u(t,x) \in \mathcal{A}(t,\alpha)} \left\{ g(x,u) + v_{t}^{\alpha} + v_{x}^{\alpha} f^{\alpha}(t,x,u) \right\}, \tag{2.10}$$

almost anywhere in t with the probability 1, then

$$v^{\alpha}(t,x) = J^{\alpha}(t,x;u^{*}). \tag{2.11}$$

Proof. The proof of this theorem is presented in Appendix II.

2.4 Optimal Feedback Control

As described above, the system (2.8) presents an optimal feedback control as the closed loop control satisfies the Bellman principle. Note that the solution of the first-order partial derivative in (2.8) is not simple. In order to interpret the value function $v^{\alpha}(t,x)$, the concept of a viscosity solution is often used. For more information and discussion on this concept, the reader is referred to Sethi et al. (2005), and Fleming and Soner (2006).

To overcome certain difficulties related to the dynamics of the system, such the number of machine modes as well as the number of part types, a heuristic method is used to solve the problem, assuming that the optimal control problem in real time is similar to the feedback control problem. The idea is to subdivide the multivariable problem described in (2.8) into different sets of problems, with each problem having a single variable control, and corresponding to Alkella and Kumar's optimal solution as seen in the hedging point policy. A further simplification of equation (2.8) in determining a control $u(t,x)$ is addressed by the following linear program:

$$\min_{u(t,x)\in\mathcal{A}(t,\alpha)} \left\{ \sum_{i=1}^{m} \frac{\partial v^{\alpha}(t,x)}{\partial x_i} \left(u_i(t,x) - d(t)\right) \right\}, \tag{2.12}$$

subject to equations (2.1), (2.2), and (2.3) above.

The feedback optimal control ((2.12) or (2.8)) is designed to lead the system to the hedging point. If the system state is $\xi(t) = 0$, where all machines are down, we must have $u(t,x) = 0$. Whenever the system state is $\xi(t) = \alpha$, the linear program of (2.12) presents a real-time feedback controller and the production rate is calculated at every time instant with $\xi(t) \neq 0$ according to varying demand. The point-x space at which the gradient of $v^{\alpha}(t,x)$ is equal to zero ($v_x^{\alpha}(t,x) = 0$) is called the *Hedging point* z^*.

The optimal control problem (2.12), for the first time was established by Kimemia and Gershwin (1983), after which Akella and Kumar (1986) established the optimal production rate as the *bang-bang control* for a *single-part type and signle machine system* $u^*(t,x)$ as follows:

$$u_j^*(.) = \begin{cases} r & \text{if } \; x_j(t) < z^* \;\&\; \xi_j(t) = 1 \\ d(t) & \text{if } \; x_j(t) = z^* \;\&\; \xi_j(t) = 1 \\ 0 & \text{if } \; x_j(t) > z^* \;\&\; \text{otherwise} \end{cases}, \tag{2.13}$$

where the hedging point, z^* is determined by minimizing the following objective function:

$$J(t,x) = \lim_{T\to\infty} \frac{1}{T} \left[\int_t^T g\left(x(s), u(s)\right) ds \right], \tag{2.14}$$

with respect to $x(t)$, $0 \leq t < T$ (see Akella and Kumar (1986) and Bielecki and Kumar (1988)).

2.5 Application to single-machine, single part-type

In this section, using the proposed model, we focus on a small manufacturing system which is a single unreliable machine capable of producing a single part-type with varying demand rates.

2.5.1 System description

This case considers a single-machine, single-part production system as shown in Figure 2.1 with $m = 1$, $d(t) \geq 0$ is the demand rate, $x(t) \in (-\infty,+\infty)$ is the buffer level (*state variable*), defined as the difference between the cumulative production of finished goods and the cumulative demand, and $u(t) = u(t,x)$ is the production rate (*control variable*). The dynamical buffer level $\dot{x}(t)$ and the buffer level $x(t)$ are defined as follows:

$$\begin{cases} \dot{x}(t) = u(t) - d(t) \\ x(t) = x_0 + \int_0^t \big(u(\tau) - d(\tau)\big) d\tau \end{cases}, t \geq 0, x(t_0 = 0) = x_0, \qquad (2.15)$$

which is the number of parts produced in time interval $[0, t]$ after satisfied the demand rate $d(t)$. In equation (2.15), if t_0 is equal to any time t, the initial condition $x(t) = x$.

Let us assume that r is the maximum production rate of the machine such that $r > d(t)$. Thus, for time $t \leq T$, the production rate obeys to the following relationship:

$$0 \leq u(t) \leq r, \quad \text{and} \quad \xi(t) \in \{0,1\}. \qquad (2.16)$$

The machine is available (up) for production if ($\xi(t)$ = 1) and unavailable (down) when it is under repair ($\xi(t)$ = 0). As a result, equation (2.16) satisfies the following:

$$\begin{gathered} \xi(t)=0 \Rightarrow u(t)=0, \\ \xi(t)=1 \Rightarrow 0 \leq u(t) \leq r. \end{gathered} \tag{2.17}$$

The hybrid state $\{x(t), \xi(t)\}$ consists of the continuous component $x(t)$ and the discrete component $\xi(t)$. The stochastic process provides an alternative expression of the transition probability for the system under consideration (including sojourn time in state α for α = 0,1). Below, we will use $F(t)$ and $G(t)$ to denote the general stationary probability distribution functions for the up state and the down state, respectively. Specifically, the mean time to fail (MTTF) and the mean time to repair (MTTR) are expressed as follows:

$$MTTF=\int_0^\infty tf(t)\,dt, MTTR=\int_0^\infty tg(t)\,dt, \tag{2.18}$$

where $f(t)=\dfrac{dF(t)}{dt}; g(t)=\dfrac{dG(t)}{dt}$.

The failure rate $p(t)$ and repair rate $q(t)$ are given by

$$p(t)=\frac{f(t)}{1-F(t)}; q(t)=\frac{g(t)}{1-G(t)}.$$

Production prevails if the machine has sufficient average capacity to meet demands, and therefore the following relationship holds:

$$\frac{MTTF}{MTTF+MTTR}r-d>0. \tag{2.19}$$

The optimal production rate $u^*(t,x)$ which minimizes the expected long run average shortage and inventory cost combined is found by solving the following objective function:

$$v^{\alpha}(t,x) = \min_{u(s,x)\in \mathcal{A}([t,T],\alpha)} E\left\{\int_t^T \left(c^+x^+(s)+c^-x^-(s)\right)ds \,|x(t)=x, \xi(t)=\alpha\right\}, \qquad (2.20)$$

where $x^+(s) = \max(0, x(s))$ and $x^-(s) = \max(0, -x(s))$ are the inventory and shortage levels at time $s \geq t$, respectively, and c^+ and c^- are the cost per unit of inventory and shortage cost per time unit, respectively.

Let $g(x,u) = x^+(t)c^+ + x^-(t)c^-$ be a convex function. A production policy that satisfies (2.20) must consider the function $u^{\alpha}(t,x)$ for every $(t,x)\in Q=[0,T]\times(-\infty,+\infty), \alpha\in\{0,1\}$:

$$u(t,x)\in \mathcal{A}(t,\alpha)=\{u(.):0\leq u(.)\leq r, \xi(t)=\alpha\}. \qquad (2.21)$$

The production policies are set before the feedback control laws, which allow the determination of a feasible production rate for each buffer and machine state, at a time interval of length T, $[0, T]$.

2.5.2 HJB equations for the Optimal Control Problem

The Hamilton Jacobi Bellman (HJB) equation was used as the candidate equation described in expression (2.8). In this problem, the *cost-to-go* functions are time-dependent, and the HJB equation is expressed as follows:

$$\min_{\substack{u(t,x)\in[0,r]\\ \alpha\in[0,1]}} \left\{g(.)+\frac{\partial v^{\alpha}(.)}{\partial t}+\frac{\partial v^{\alpha}(.)}{\partial x}\left(u(t)-d(t)\right)-p_{\alpha\alpha}(t)v^{\alpha}(.)+p_{\alpha\beta}(t)v^{\beta}(.)\right\}=0. \qquad (2.22)$$

In order to calculate the optimal cost for a given downtime distribution, equation (2.22) can be further simplified as follows:

- *machine is down* ($\alpha = 0$), *set* $u(.) = 0$

$$g(.) + v_t^0(.) - v_x^0(.)d(t) + p_{01}(t)v^1(.) - p_{00}(t)v^0(.) = 0, \tag{2.23}$$

- *machine is up* ($\alpha = 1$), $u(.) \geq 0$

$$\min_{0 \leq u \leq r} \left\{ g(.) + v_t^1(.) + v_x^1(.)\left(u(t) - d(t)\right) + p_{10}(t)v^0(.) - p_{11}(t)v^1(.) \right\} = 0. \tag{2.24}$$

In the case of $\xi(t) = 0$, since the machine is down, we have $u = 0$. Since there is only one continuous state variable, the HJB equation leads to a set of determining functions of two variables, $v^0(t,x)$ and $v^1(t,x)$. The control $u(.)$ is determined in the equation (2.24) by:

$$\min_{u(t,x) \in [0,r]} \frac{\partial v^1(x,t)}{\partial x}\left(u(t,x) - d(t)\right). \tag{2.25}$$

2.5.3 Numerical example

A) Parameters of the system

To illustrate the proposed technique, an example is used with a real-life practical project. Although the example problem is drawn specially from a small manufacturing system, the generic version of the problem and the challenges presented are common to many manufacturing systems. We proceeded by first giving a description of the system, including data such as the varying demand rate. In this example, the up- and downtimes are distributed according to one of the following probability density functions, referred to as reliability models:

i. Weibull, i.e.,

$$f^W(t) = \lambda^{\mu} \mu t^{\mu-1} \exp\left(-(\lambda t)^{\mu}\right).$$

This distribution is denoted as $W(\mu, \lambda)$.

ii. Log-normal, i.e.,

$$f^L(t) = \frac{1}{t\sigma\sqrt{2\pi}} \exp\left(-\left\{\frac{[\ln(t) - \lambda_o]^2}{2\sigma^2}\right\}\right).$$

This distribution is denoted as $L(\lambda_o, \sigma)$.

iii. Gamma, i.e.,

$$f^G(t) = \frac{\lambda_1^{\mu_1} t^{\mu_1 - 1} \exp\{-\lambda_1 t\}}{\Gamma(\mu_1)}.$$

This distribution is denoted as $G(\mu_1, \lambda_1)$.

We have to be certain that the failure rate $p(t)$ and the repair rate $q(t)$ are in function of time, and that the coefficient of variation, CV, is less than one. The set of up-and-down times used in this example is shown in Table 2.1. The coefficients of variation, CV, which take values less than one, are: $CV_W = 0.76$, $CV_L = 0.95$, and $CV_G = 0.60$.

The system has the following parameters:
- the varying demand, rate with time t, is given in Table 2.2,
- maximum production rate $r = 0.25$,
- $c^+ = 1$ unit-cost, $c^- = 2$ unit-cost,
- interval time and stock level take values in [0, 100] and [-80, 80], respectively.

Table 2.1 Up-Downtime distributions considered

Case	Uptime	Downtime	MTTF	MTTR
A	$L(4.06, 0.80)$	$W(1.25, 0.115)$	80	8
B	$G(2.80, 0.035)$	$W(1.25, 0.115)$	80	8

Table 2.2 Varying demand rate data

T	0-20	20-40	40-60	60-80	80-100
$d(t)$	0.10	0.15	0.135	0.165	0.125

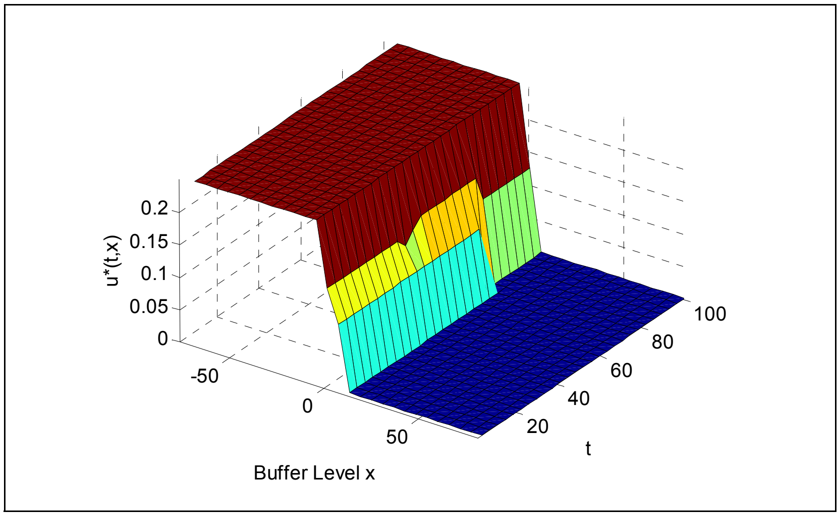
To solve the HJB equations, we use Kushner and Dupuis approximation scheme based on the method proposed by Kushner and Dupuis (2001). Let $\Delta x > 0$ and $\Delta t > 0$ denote the lengths of the finite difference interval of the variables x and t, respectively. The value first-order partial derivative functions (equations (2.23) and (2.24)) are replaced by the following expressions:

$$v_x^\alpha(t,x) = \begin{cases} \dfrac{v^\alpha(t, x+\Delta x) - v^\alpha(t,x)}{\Delta x} & \text{if } \dot{x} \geq 0 \\ \dfrac{v^\alpha(t,x) - v^\alpha(t, x-\Delta x)}{\Delta x} & \text{otherwise} \end{cases},$$

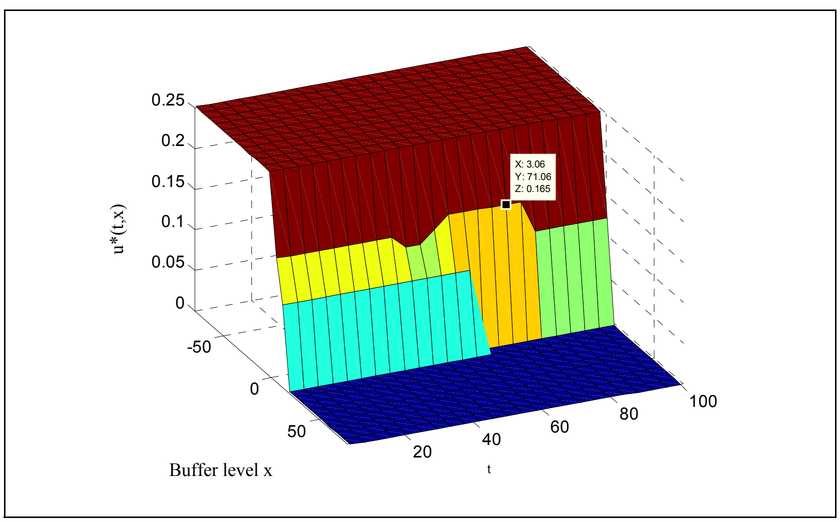
$$v_t^\alpha(t,x) = \frac{v^\alpha(t+\Delta t, x) - v^\alpha(t,x)}{\Delta t}.$$

B) Result analysis

The results of case study A and B are illustrated in Figure 2.2 as the optimal production rates for varying demand rates, as shown in Table 2.2. Results obtained are in agreement with the control law with respect to the optimal production rate, as in expression (2.13), called the *hedging point policy*. In both cases A and B in Figure 2.2, maximum production rates are obtained at (r = 0.25), while the stock level is less than zero. This maximum value decreases to zero when the stock level rises above the hedging point, z^* = 3.06 parts.



a) Case A



b) Case B

Figure 2.2 Optimal Production Rate versus *t* and *x* in case A and B.

The result of case B is illustrated in Figure 2.2 (b). This is in fact similar to case A, with the optimal control law obeying the bang-bang control problem as in the expression (2.13). Although the up- and downtime distribution are different from what we have in case *A*, the results obtained in both cases are almost similar. This is because the same MTTF and MTTR in Table 2.1 are used for both cases.

Through the computing process and by analyzing the results, it appears that the developed semi-Markov processing model, which bears the usual Markov assumptions, is to consider non-exponential machine and varying demand in time. That is significantly better than the baseline Markov model. The semi-Markov model is an extension of the Markov model, in which each state has an explicit state duration probability distribution.

The results above show that the optimal control law is similar to the bang-bang control problem when the stock level variation in time ranges between $0 < t < T$ (i.e., bang[0, r; $dv^1(t, x)/dt$] as in the expression (2.13)). The optimal policy $u^*(.)$ is in agreement with the analysis model developed in Akella and Kumar (1986), and is equal to $d(t)$, where the hedging point $z^*(t)$ exists. Results indicate that the semi-Markov model performs very well as compared to the Markov model.

C) Hedging point policy analysis of case B

Since the results of two cases were similar to each other, we analyzed the hedging point policy with respect to its short-term behaviour for the case B. The hedging point policy and production targets have presented the results in Figures 2.3 and 2.4. Using a MATLAB code for HJB equations, the average steady-state cost was calculated in terms of hedging points, after which we found the minimum values for the hedging points. The maximum production rate of the machine, the penalty costs per unit of surplus and the backlog provide clear information and help understand why (and under what condition) a hedging point is optimal. In addition, this provides useful information at the system design level for choosing system parameters, such as machine capacity and reliability, to ensure that the hedging point is optimal. The analysis was performed over 100 time units, and the hedging point was determined by minimizing $J(t,x)$ (see Eq.(2.14)).

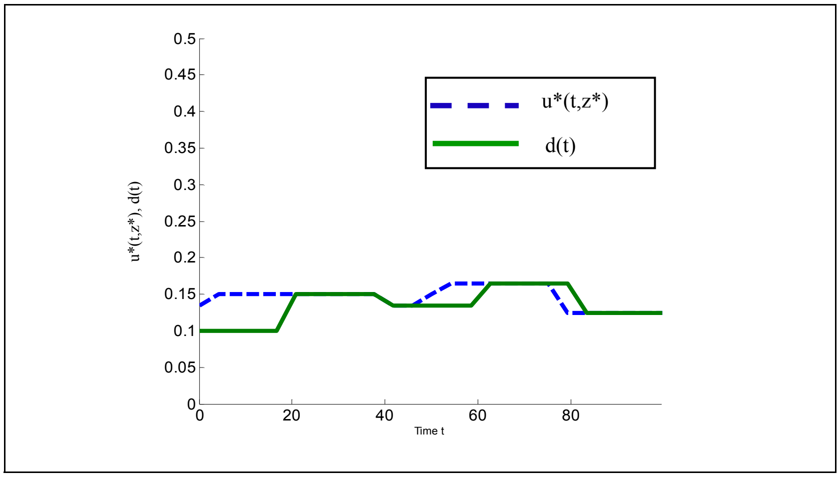


Figure 2.3 Hedging point policy with varying demand.

In Figure 2.3, the solid and dashed lines present the demand rate $d(t)$ and the optimal policy $u^*(t,z^*)$, respectively.

As can be seen in Figure 2.3, the optimal policy at the hedging point is found at time interval (0, 100), and becomes asymptotic to time-varying demand. The hedging point is at $z^* = 3.06$, and the z^* value is computed as follows:

$$z^* = \arg \min_{x \subset (-20,20)} \left(v^1 \left(t, x \right) \right). \tag{2.26}$$

Short-term behaviour

Figure 2.4 presents the production targets with a hedging point policy in which the process is balanced with varying demand, and service time variations are considered. The production rate was recorded for every parameter combination, and by $t = 20$ time units, the production rate was higher than the demand rate $u^*(t,z^*) = 0.15 > d(t) = 0.1$, i.e., when $z^*(t) = 3.06$ parts, which is higher than the cumulative production of part $\int_0^{20} \left(u^*(t, z^*) - d(t) \right) dt = 1$ part. When the production rate coincides with the demand rate, then the production rate $u^*(t,z^*)$ should be exactly equal to $z^*(t) + t.d(t)$. Within the time interval [20, 40], the production reached the hedging point with $u^*(t) = 0.15$, and before that, was equal to $z^*(t) + t.d(t) = 3.56$ parts at $t =$

23.75 time units. The hedging point time interval, when $u^*(t,z^*)$ was equal to $d(t)$, lasted between 23.75 and 38 time units.

Note that within the interval (40, 60], the production did not reach the hedging points because of the change in the demand rate from $d(t)$ = 0.15 to 0.135. However, the total sum of $z^*(t) + t.d(t) = 6.27$ parts was produced at t = 38 time units. The result is similar for the (60, 80] and (80,100) time periods.

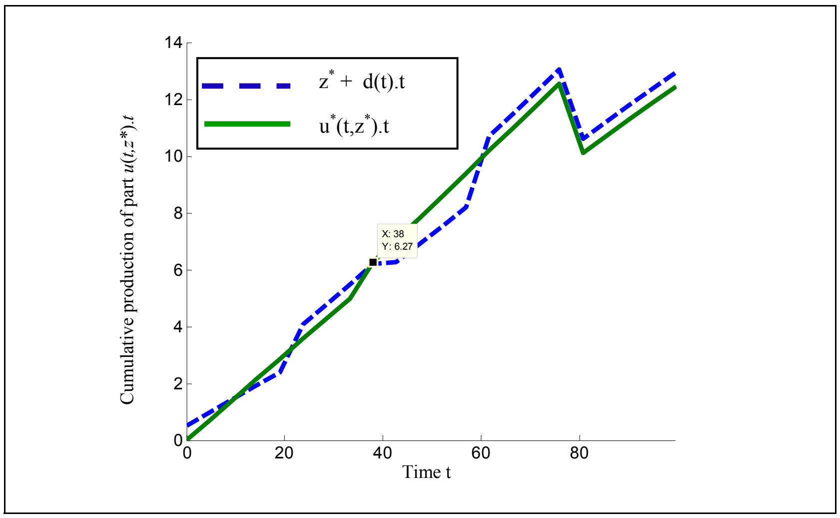


Figure 2.4 Production Targets with hedging point policy.

The results obtained indicate that the semi-Markov model works well for small systems. In the real world, the demand is not constant, so this model can be applied to the modeling of manufacturing systems in real-time, and with a varying demand rate.

2.6 Application analysis to two-machine in parallel

2.6.1 Production system description

Consider a manufacturing system as shown in Fig 2.1 with $m = 2$. It is a parallel arrangement of two identical machines producing a single part-type. Each machine is subject to random failures and repairs. Assume that each machine has two finite states $\xi_i(t)$; i =1, 2 (i.e., available $\xi_i(t) = 1$ and unavailable $\xi_i(t) = 0$) with probability distributions $F(t)$ and $G(t)$ respectively applicable to each state. A computation of the value function $v^{\xi(t)}(t,x)$ and the production rate $u^{\xi(t)}(t,x)$ includes all states.

Let $\xi(t) = (\xi_i(t), t \geq 0, i = 1,2)$ be defined by the following expression:

$$\xi(t) = \begin{cases} 2 & \text{If All two machines are operational, i.e., } \xi_1(t) = 1 \;\&\; \xi_2(t) = 1 \\ 1 & \text{If One of them is operational, i.e., } \xi_i(t) = 1 \;\&\; \xi_j(t) = 0; i \neq j\,. \\ 0 & \text{If None of them is operational, i.e., } \xi_1(t) = 0 \;\&\; \xi_2(t) = 0 \end{cases} \qquad (2.27)$$

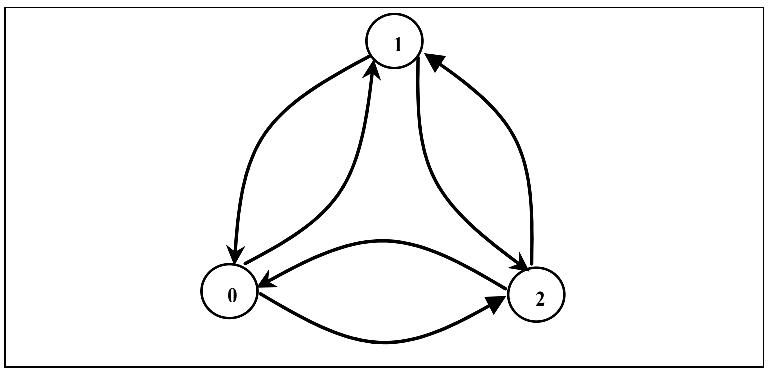


Figure 2.5 Transition graph of two-machine system with three states.

Thus, we have the transition graph of finite states as show in Figure 2.5. Performance measures of the two-state machines (state 0 and state 1) are evaluated using the steady-state probability distribution and the transition probability distributions of the semi-Markov process. We must admit that only the total time of the system operation influences its reliability, since the probability distribution for the system at each next step only depends on the current state of the system. The configuration of three-state machines (state 0, state 1, and state 2), proposed by Kim and Gershwin (2005,2008), is illustrated in Figure 2.6.

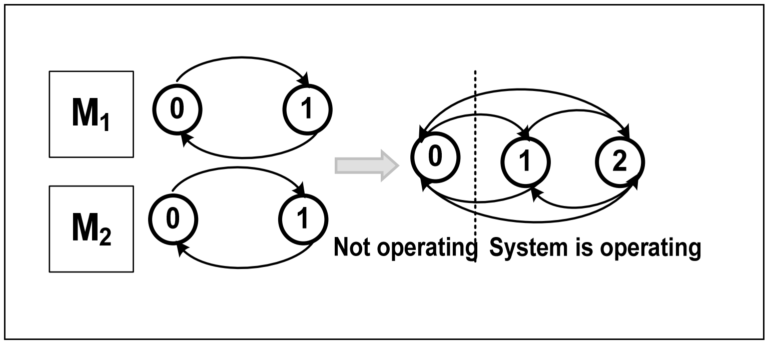


Figure 2.6 Two-state machines and the concurrent three state-machines schematic.

The conventional implementation of *state machines* is based on the selection of successor states and the execution of *related* actions. It is natural to use state machines for modeling these applications, since an application that must carry out a series of actions, or handle inputs (responses and indications), differently depending on what state it is in, is often best implemented as a state machine. The machine states are described as follows:

State 0: the system is not operating (down); *State* 1: the system is operating and producing bad parts (up/bad), with only one machine; State 2: the system is operating and producing good parts, with both machines. In this example, we assume that all transition times are not exponentially distributed.

Using the reliability theory (Ross (2003)), the dynamics of the two parallel machines model is described as similar to that of an equivalent single machine, as follows:

$$\begin{cases} F_{\xi(t)}(t) = \Pr\{\text{System is functioning at time } t,\ t \geq 0\} \\ F_{\xi(t)}(t) = 1 - \Pr\{\text{All machines are "down" at } t,\ t \geq 0\} \end{cases},$$

$$F_{\xi(t)}(t) = 1 - \prod_{i=1}^{2} \left(1 - \Pr\{\text{machine } i \text{ is functional at } t \geq 0\}\right).$$

Let $F_2(t)$, $F_1(t)$, and $F_0(t)$ be the probability that all two machines are operational, only one of them is operational, and none of them is operational, respectively. Six transition probabilities are shown in Figure 2.5, for three states. The derivative of the semi-Markov transition probability $p_{\alpha\beta}(t)$ is defined by equation (2.5), where α, β = 0,1,2. A calculation of $F_2(t)$, $F_1(t)$, $F_0(t)$, and $p_{\alpha\beta}(t)$ is presented in Appendix II.

In Figure 2.6, *at state* 2: the future state of the system is as only one machine is "up" (2→1) or none of them is "down" (2→0). A*t state* 1: the future state of the system is as all two machines are "up" (1→2) or none of them is "up" (1→0). Thus, *at state* 0: the future state of the system is as all two machines are "up" (0→2) or one of them is "up" (0→1). Another measure of performance that may be considered is the one similar to one machine. The value function of two machines system is described by the HJB equation with three states as follows:

$$\begin{cases} \text{at state 0:} \\ g\left(x\right)+v_t^0(t)-d.v_x^0(t)+\sum_{\beta=0}^{2}\delta(t)p_{0\beta}(t)v^{\beta}(t)=0, \end{cases}$$

$$\begin{cases} \text{at state 1:} \\ \min_{0\le u\le r}\left\{g(x)+v_t^1(t)+(\frac{u}{2}-d)v_x^1(t)+\sum_{\beta=0}^{2}\delta(t)p_{1\beta}(t)v^{\beta}(t)\right\}=0, \end{cases}$$

$$\begin{cases} \text{at state 2:} \\ \min_{0\le u\le r}\left\{g(x)+v_t^2(t)+(u-d)v_x^2(t)+\sum_{\beta=0}^{2}\delta(t)p_{2\beta}(t)v^{\beta}(t)\right\}=0, \end{cases}$$

where $\delta(t)$ = 1 if $\alpha \neq \beta$; $\delta(t)$ = - 1 if $\alpha = \beta$; α, β = 0,1,2.

2.6.2 Numerical example

A) Parameters of the system

As in the case of a single-machine system, we use the downtime and uptime distributions whose parameters are shown in Table 2.1 for case B. Other parameters are given as:

- the maximum production rate r = 0.25,

- $c^+ = 1$ unit-cost, $c^- = 3$ unit-cost,
- constant demand rate $d(t) = 0.135$.

B) Result Analysis

In state 1, only one machine is functioning, and the production rate is equal to half of the control variable $u(.)$ (*i.e.,* $u_1(.) = u_2(.) = 0.5\ u(.)$). In state 2 mode, the two machines are functioning, with the production rate $u_1(.) + u_2(.) = u(.)$, and producing good parts. Obviously, the production rate needs to be zero when the system is located in state 0.

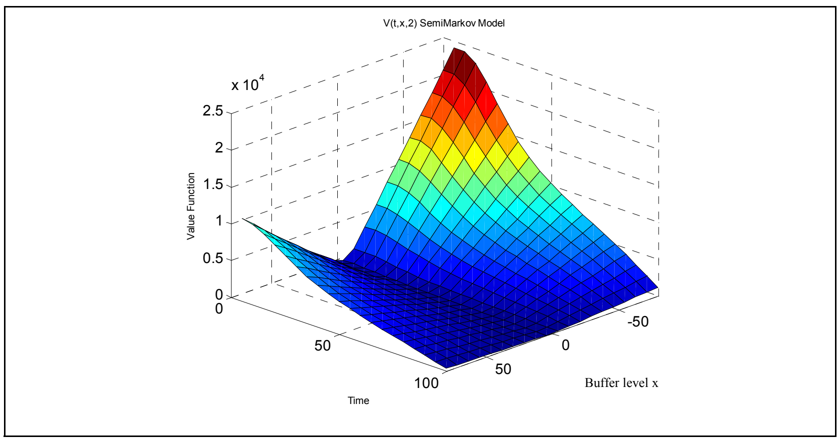


Figure 2.7 Value function $v^2(t,x)$.

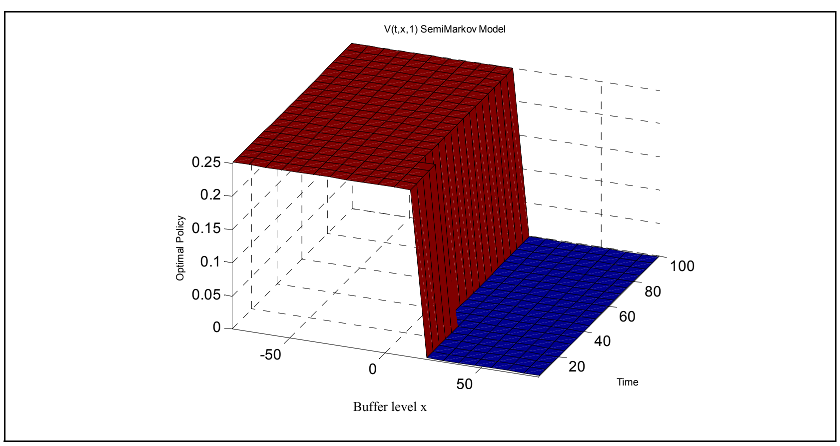


Figure 2.8 Production rate $u(t,x)$.

The results of this example are illustrated in Figures 2.7 - 2.8. Figures 2.7 and 2.8 present the value functions $v^2(t,x)$ the production rate *u(t,x)* versus time, and the stock level, which are closely associated with given internal values of $x \subset [-80, 80]$ and $t \subset [0, 100]$, respectively. As in the single-machine system case presented in Section 2.5, the results obtained corresponds to the control law in the optimal production rate (2.13) called the *hedging point policy*. In Figure 2.8, the production rate reaches its maximum value at ($r = 0.25$), when the stock levels are below zero.

The characteristics of the two-machine system correspond to those of the single-machine model that has three states, for example, in Kim and Gershwin (2005)-(2008). *State* 0: the system is not operating (down); *State* 1: the system is operating and producing bad parts (up/bad) because only one machine can produce; and State 2: the system is operating and producing good parts with both machines. Thus at state 0, the system is not operating, and is only operating at state 1 & 2.

2.7 Conclusion

In this paper, we have provided an analysis of an optimal production problem for multiple identical parallel machines, which underlies the well-known Markov properties. The discrete machine states are characterized by semi-Markov processes, and so machines are subject to random failures and repairs, with non-exponential distributions of up and down times. Using semi-Markov processes in continuous time, simultaneously with Rishel's assumptions and a dynamic programming approach, a new model of the stochastic control problem in a deterministic horizon is formulated without a discount rate. While the dynamic programming approach is used to make decisions in stages over time, Rishel assumptions are used to model the discrete machine states. In our model, we used the definitions of semi-Markov processes found in Becker et al. (2000) and the assumptions of Rishel (1975) to characterize the discrete events of the system such a breakdowns and machine repairs.

We applied our proposed model to small manufacturing system: single machine and single product system, two parallel machine and single product system with machine having Weibull, log-normal, and gamma distributions. The HJB equation was involved by using numerical approach. To valid our proposed model, a sensitivity evaluation has been implemented on a sample path.

Acknowledgements

The authors wish to express their gratitude to Professor J.P. Kenné for his suggestions.

CHAPITRE 3

ARTICLE # 2 FEEDBACK OPTIMAL CONTROL OF DYNAMIC STOCHASTIC TWO-MACHINE FLOWSHOP

Thang Diep [1], Jean-Pierre Kenné [1], and Thien-My Dao [1]

[1] Mechanical Engineering Department, Laboratory of Integrated Production Technologies, University of Quebec, Ecole de technologie supérieure, 1100, Notre Dame West, Montreal (Quebec), Canada, H3C 1K3

Résumé

L'objet principal de cet article est une extension du modèle du première article (dans le chapitre 2) dans le cas où le taux d'actualisation est non-zéro ($\rho > 0$). Sous ce nouveau modèle, la modélisation du système de production flexible à deux-machines en tandem traitant un type de pièce est de déterminer les taux de production de ces deux machines. L'utilisation des méthodes heuristiques est de résoudre le problème de commande à plusieurs états discrets, et de nous permettre de sélectionner les lois de contrôle pour chaque état. La résolution des équations HJB est proposée à l'aide des méthodes numériques basées sur l'approche de Dupuis et Kushner (2001).

Abstract

This paper examines the optimization of production involving a tandem two-machine system producing a single part-type, with each machine being subject to breakdowns and repairs. An analytical model is formulated with a view to solving an optimal stochastic production problem of the system with machines having up-downtime non-exponential distributions. The model developed is obtained by using a dynamic programming approach and a semi-Markov process. The control problem aims to find the production rates needed by the

machines to meet the demand rate, through a minimization of the inventory/shortage cost. Using the Bellman principle, the optimality conditions obtained satisfy the Hamilton-Jacobi-Bellman equation, which depends on time and system states, and ultimately, leads to a feedback control. Consequently, the new model enables us to improve the coefficient of variation ($CV_{\text{up/down}}$) to be less than one while it is equal to one in Markov model. Heuristics methods are used to involve the problem because of the difficulty of the analytical model using several states, and to show what control law should be used in each system state (i.e., including Kanban, feedback and CONWIP control). Numerical methods are used to solve the optimality conditions and to show how a machine should produce.

Keywords: Semi-Markov process, Optimality conditions, Flow-shop system, Numerical method, Heuristic method.

3.1 Introduction

In this paper, we examine a tandem two-machine system producing a single part-type with finite-size internal buffers. Each machine is subject to random breakdowns and repairs, and the system is deterministic if both machines do not fail; otherwise, it is stochastic. The system can have four different states: two machines fail; two machines work simultaneously; the upstream machine fails while downstream machine is working, and the downstream machine fails while upstream machine is working. The goal of the study is to formulate a model which consists in minimizing the expected discounted cost of inventory/shortage in deterministic horizon in order to find the production rates of a stochastic system. This section presents a literature review, the motivation for using the semi-Markov process, and the contribution of this paper.

3.1.1 Literature review

At a decision making level for the operation of manufacturing system, one of the most a configuration studied is a *flow-shop system* or *transfer line* (i.e., including a specified number of machines in series). Naturally, each machine is subject to random breakdowns and repairs (making them failure-prone machines). Other states characterising a machine include setup time, changing demands, preventive maintenance, etc. Thus, the number of discrete states of a system will grow as the number of machines increases. Indeed, consider a flowshop system consisting of *m* machines in which each machine can be in two states (up and down); we therefore have 2^m distinct states, so it is difficult to determine the performance of the system when it is modeled as a discrete-space Markov process with large state spaces. In practice, the optimal production planning of stochastic manufacturing lines (i.e., with failure-prone machines) constitutes an extremely difficult problem Fong and Zhou (2000). Obviously, no exact analytical model could be obtained for a system with the length of the machines. On other hand, one of the characteristic features of a stochastic dynamic of a flow shop is the fact that the inventory of semi-processed parts in buffers between any two machines, known as internal buffers, must be nonnegative (see Dallery et al. (1989) and Sethi et al. (1994)). Some papers, such that Kimemia and Gershwin (1983), Akella and Kumar (1986), Bielecki and Kumar (1988), Perkins and Srikant (1997), Shu and Perkins (2001), have covered these features.

The first version of the problem of production planning for a single machine producing a single part-type with two states (up and down) was studied in Akella and Kumar (1986) and Bielecki and Kumar (1988). Both these papers presented an exact solution which could find the production rate and a hedging point. The hedging point is a buffer level at which each part type must be produced with a rate equal to its demand rate ($u(t) = d(t)$). This agreement is a threshold-type that could be considered as a Just-In-Time (JIT) method for solving the stochastic problem by maintaining an inventory level equal to the hedging point. In Perkins and Srikant (1997) and Shu and Perkins (2001), they then went on to consider the problem of a single machine system producing multiple part-types. They used the decomposition method, in which the multiple part-types problem is decomposed into a two part-type problem, as well as a graph technique for a linear switching curve problem.

In the case of a two-machine flowshop, some authors have conducted studies on both deterministic and stochastic problems. Since the optimal production planning of a stochastic manufacturing system is difficult, Sethi et al. (1997) have studied a system with a single part-type using a hierarchical approach: the idea is to carry out the uncertainty in the machine's capacity which is averaged, and replace the more general stochastic problem with a limiting problem. In that paper, they show that the performance in the feedback control is better than in the Kanban control. In the literature, a hierarchical control approach was introduced in Gershwin (1987) and Lechocky et al. (1991), and was based on the frequency of occurrence of different types of events (also called the *time-scale control*). On the other hand, a deterministic problem of two-machine flowshop was studied in Fong and Zhou (2000), these authors although gave an exact solution whose optimality conditions satisfy the Hamilton-Jacobi-Bellman equation, they did not involve the real problem of manufacturing systems that is stochastic rather than deterministic.

Bai and Gershwin (1996) used the heuristic method to obtain sub-solution controls in *N*-machines and in a single part-type system with the objective of long-term average cost minimization. Presman et al. (2002) and Sethi et al. (2005) studied the *N*-machine flow shop whose profit function is minimized by the average cost. As stated above, so far, there is no exact solution for failure-prone machine systems with the large of the transfer lines. The simulation method therefore represents a significant advantage in terms of analysis of the performance of the system, as can be seen in Kenné and Gharbi (2001) and Lavoie et al. (2009). Other papers focus on the performance parameters of transfer lines (i.e., lines including production rate and average buffer levels); that is the case in Dallery et al. (1989), Ciprut et al. (2000), Kim and Gershwin (2005-2008), and Tan and Gershwin (2009), where long lines are decomposed into two-machine lines (flowshop systems) in the case of identical machines. This technique is called the *decomposition method*, and through it, the system becomes simpler and behaves like a buffer or work-in-process inventory between upstream and downstream machines.

In most of the works mentioned above, the state machines are characterized by *Markov processes* and the demand rate is constant. This feature was developed from the formalism of several pioneers such as Rishel (1975) and Davis (1984) used the Markov chain to formulate a stochastic model in continuous time as *a Piecewise deterministic system* (PDS). Moreover, the Markov framework with machines having exponential distributions of uptime and downtime has a coefficient of variation (CV_{up} and CV_{down}) equal to 1 and breakdown and repair rates equal to constant. As the results in Li and Meerkov (2005a) and Enginarlar et al. (2005), the performance of the average number of parts produced (*PP*) by the last machine depends mostly on the $CV_{up/down}$: if the $CV_{up/down}$ decreases to less than 1, the performance *PP* does increase and the sensibility of the *PP* assumes values within the 6% range. Indeed, the $CV_{up/down}$ is less than 1 if the breakdown and repair rates are functions of time, as indicated in Li and Meerkov (2005b). That means the machine lifetime must obey the non-exponential distribution as in Grabsky (2003).

3.1.2 Motivation for using the Semi-Markov process

The simultaneous use of the semi-Markov process and the two-machine flowshop system is motivated by the following three factors:

1. From a practical point of view, the lifetime of a machine is described by a more general random process, as stated in Grabsky (2003). That means machines often have up-down time distributions which could be non-exponential, and characterized by a coefficient of variation ($CV_{up/down}$), often less than 1 (see Li and Meerkov (2005a,b) and Enginarlar et al. (2005)). Thus, the machines may be referred to as aging over time without any restriction while using exponential distributions, as can be seen in the literature.

2. The study of transfer lines is based on a two-machine line (flowshop system) because no exact analytical solution exists for longer lines, and brute force numerical techniques are unsatisfactory with sizes of the state spaces (Gershwin (2002)). The performance parameters of a two-machine system, such as the production rate and the average buffer

levels, depend on the work-in-process (WIP) inventory and throughput time (lead time or cycle time), and the performance is good if the WIP inventory and lead time are optimized (Bai and Gershwin (1996)). That leads to an optimal production problem respecting the minimizing of the total cost of inventory/backlog over deterministic time. Moreover, the appreciation of the performance of the system influences the coefficients of variation, $CV_{up/down}$ (see Enginarlar et al. (2005)) and as a result, an optimal production control problem with semi-Markov jumps should be formulated.

3. The time and non-exponential distributions issues can be considered by extending the dynamic programming method using semi-Markov jumps. Hence, a unified model, including production, is developed in this paper and the optimality conditions obtained are then solved to obtain the optimal control policy.

3.1.3 Contribution of this paper

The purpose of this paper is to present a new model for the optimal stochastic control of a failure-prone two-machine system in a finite horizon, with semi-Markov jumps and a discount rate. This model is based on the dynamic programming approach, and adopts the assumptions of Rishel (1975). However, unlike Rishel, who generated an optimal control with Markov jumps with constant transition rates, we use semi-Markov jumps, whose transition rates and probabilities are time-dependent. Using the Bellman principle, the optimality conditions satisfy the Hamilton-Jacobi-Bellman equation which appears in this paper. The new model and related optimality conditions are applied to a real world manufacturing system involving log-normal, Weibull, and gamma distributions, which are in turn used to represent the machine's up (operating) and down times with a $CV_{up/down}$ of less than one. This paper also proposes a solution for the new model with heuristic and numerical approaches.

The next sections are organized as follows: Section 3.2 presents the problem formulation, and Section 3.3, the optimality conditions. The dynamics of the system is given in Section

3.4. The hedging point policy is analyzed in Section 3.5. Section 3.6 presents the heuristic method for optimal feedback control, while Section 3.7 and 3.8 present two practical case studies. Finally, Section 3.9 presents the conclusion.

3.2 Problem formulation

We consider a dynamic stochastic flow shop consisting of a tandem two-machine system devoted to producing a single product, as shown in Figure 3.1. The machines are subject to random breakdowns and repairs. Each machine has a finite number of states (modes), denoted as $\alpha_j \in \{0, 1\}$ for $j = 1, 2$; machine j is in up state (available) with $\alpha_j = 1$, or in down state (unvavailable) with $\alpha_j = 0$. Consider the number of parts in the buffer between the first and the second machines, called the work-in-process (WIP), as $x_1(t)$, and the surplus level of the finished goods as $x_2(t)$ for $t \geq 0$. The number of parts in WIP cannot be negative and the buffers usually have limited storage capacities such as $0 \leq x_1(t) \leq B$; B is the upper bound on the WIP. If the surplus level $x_2(t) > 0$, we have inventories; however, if $x_2(t) < 0$, then we have backlogs.

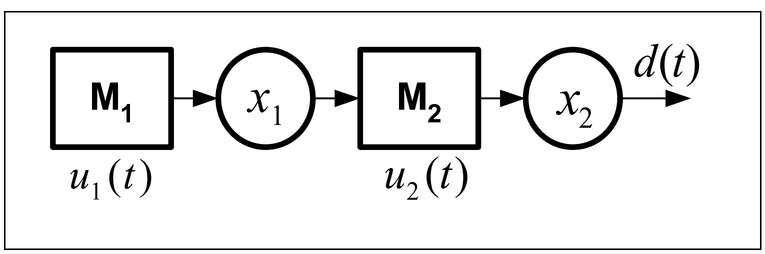


Figure 3.1 Two-machine flow-shop system.

Let $u_1(t)$ and $u_2(t)$ be the production rates of the first and second machines, respectively. Accordingly, the maximum production capacities of these two machines are denoted as r_1 and r_2. We assume a varying demand $d(t)$, which is random variable, as the input. Let $\xi(t)$ be the mode of a given machine system at time t. The set of possible value of $\xi(t)$ can be determined from the value of α_i as illusatred in Table 3.1.

Table 3.1 Mode of two-machine system

α_1	0	1	0	1	Machine 1
α_2	0	0	1	1	Machine 2
$\xi(t)$	0	1	2	3	System

It is described by a semi-Markov process with the state space $\mathcal{I}$ = {0,1,2,3} and their transition probabilities from state α to state β, as follows (Becker et al. (2000) :

$$P_{\alpha\beta}(s) = \Pr\left[S_n - S_{n-1} \le s \cap \xi(S_n) = \beta \mid \xi(S_{n-1}) = \alpha\right], \tag{3.1}$$

where S_n is the time of the next transition and S_{n-1}, the time of the last transition (with $S_0 = 0$) with respect to $s \ge t$.

The dynamics of system contains two different parts, the first being the continuous part, and the second, the stochastic part. The dynamics of the continuous part of the process is described as follows:

$$\begin{cases} \dfrac{dx_1(s)}{ds} = u_1(s) - u_2(s) \\ \dfrac{dx_2(s)}{ds} = u_2(s) - d(s) \end{cases}, s \ge t; x_1(t) = x_1; x_2(t) = x_2. \tag{3.2}$$

Let $S = [0, B] \times \Re^1$ be a state constraint domain. Then, let $\boldsymbol{x}(s) = (x_1(s), x_2(s))' \in S$ and $0 \le u_j(s) \le r_j$, $j = 1, 2$. For simplicity: $x_j(t) = x_j$ for $j = 1,2$; $s \ge t$, and $\boldsymbol{x}(t) = \boldsymbol{x}$ (initial conditions) for $t \ge 0$. Let $\mathcal{U}(t, \boldsymbol{x})$ be the constraint domain of control as follows:

$$\mathcal{U}(t, \boldsymbol{x}) = \left\{\boldsymbol{u}(.) = (u_1(.), u_2(.))' : 0 \le u_j(.) \le r_j, j = 1, 2\right\}. \tag{3.3}$$

The equation (3.2) can be written as follows:

$$\dot{\boldsymbol{x}} = f^{\xi(s)}(s, \boldsymbol{x}, \boldsymbol{u}), \text{ for } s \geq t \geq 0, \boldsymbol{x}(t) = \boldsymbol{x}, \xi(t) = \alpha. \tag{3.4}$$

This stochastic differential equation (3.4) is the hybrid system. If the system enters a state γ *then* a number of independent times is T_γ with distribution functions $F_\gamma(s)$, and probability density function $f_\gamma(s)$ for $\gamma \in \mathcal{I}$. The system will go to state β, if the realization of T_γ is the smallest of all these variables, and the sojourn time in state α will just be the smallest realisation. Then, the derivative of the semi-Markov transition probability $p_{\alpha\beta}(s)$ for $s \geq t \geq 0$ is given by

$$p_{\alpha\beta}(s) = \frac{dP_{\alpha\beta}(s)}{ds} = \prod_{\gamma \neq \beta}^{3} \left(1 - F_\gamma(s)\right) f_\beta(s). \tag{3.5}$$

Let $\mathcal{F}_k$ denote the σ-algebra generated by the random process and the number of independent random times τ_k as follows:

$$\mathcal{F}_k = \sigma\Big(\big(\boldsymbol{x}(\tau_k), \xi(\tau_k)\big) : 0 \leq \tau_k \leq t\Big), \big\{\boldsymbol{x}(t), \xi(t) : t \geq 0\big\}.$$

We now define the concept of admissible controls.

Definition 3.2.1 *A control* $\mathcal{U}(t, \boldsymbol{x}) = \big(u_1(t, \boldsymbol{x}) = u_1(t), u_2(t, \boldsymbol{x}) = u_2(t)\big) \in \Re_+^2$ *is admissible with respect to the initial state vector* $\boldsymbol{x} = (x_1, x_2)' \in S$ *if:*

(i) $\mathcal{U}(t, \boldsymbol{x})$ *is an* $\mathcal{F}_k$ *-adapted measurable process;*

(ii) $\boldsymbol{u}(t,\boldsymbol{x}) \in \mathcal{U}(t, \boldsymbol{x})$ *for all* $t \geq 0$.

For more information and a discussion of this concept, the reader is referred to Sethi et al. (1997) and reference therein.

Let $\mathcal{A}(t,\alpha)$ be the set of admissible feedback controls with the initial vector $\boldsymbol{x}(t) = \boldsymbol{x}$. Let $0 \leq t_o < T$ and consider initial times t_o in interval $[t_o, T]$.
Let

$$g\left(\boldsymbol{x}(t), \boldsymbol{u}(t)\right) = c_1 x_1(t) + c_2^+ x_2^+(t) + c_2^- x_2^-(t) \tag{3.6}$$

be the surplus cost, c_1 is the unit inventory cost of the internal buffer, c_2^+ the unit surplus cost of the finished product in the external buffer, and c_2^- the unit backlog penalty of the finished product. Our objective is to find an admissible control $\boldsymbol{u}(t, \boldsymbol{x}) \in \mathcal{A}(t,\alpha)$ at time t in state $\boldsymbol{x}$ that minimizes the following cost function:

$$J^{\alpha}(t, \boldsymbol{x}, \boldsymbol{u}) = E_u \left[\int_t^T e^{-\rho(s-t)} g(\boldsymbol{x}(s), \boldsymbol{u}(s)) ds \mid \boldsymbol{x}(t) = \boldsymbol{x}, \xi(t) = \alpha \right], \tag{3.7}$$

where $\rho > 0$ is the discount factor, $E_{\boldsymbol{u}}$ is the mathematical expectation taken with respect to the measure induced by the control law $\boldsymbol{u}(t, \boldsymbol{x})$, T is deterministic horizon (also *deterministic time*), α is initial state of system, and $\boldsymbol{x}$ is initial value at time t. The function (3.7) is called *the surplus cost function*.

For the manufacturing system, the following assumptions are made in developing the control strategy:

H.1. *Assume $c_1 > 0$ and $c_1 \leq c_2^+ \leq c_2^-$. This means that holding costs typically increase as the "value added" increases.*

H.2. *The manufacturing Lead Times are considered only on processing time while setup time, transfer time, and queue time are neglected. All operating machines start their operations at the same time.*

H.3. *The demand rate is considered for both varying and constant variables*.

This optimization problem falls within the framework of the optimization system with semi-Markov jumps called *stochastic optimal control problem*, in which machines' life times obey the non-exponential distribution. In the next section, we establish the optimality conditions described by the Hamilton-Jacobi-Bellman (HJB) equation as candidate of the optimal control problem.

3.3 Optimal feedback control

In this section, our analysis covers the construction of an optimal feedback control structure that satisfies (3.2), (3.3) and (3.7), and determines the optimal production rate $\boldsymbol{u}(t,\boldsymbol{x})$ with the minimum cost function described in (3.7). Moreover, it is closely related to the idea of a feedback control in which the control variable $\boldsymbol{u}(t,\boldsymbol{x})$ is chosen based not only on the time t but also on the state $\boldsymbol{x}(t)$. Let $Q=[t_0,T)\times S\in[t_0,T)\times[0,B]\times\Re^1$ and $(t,\boldsymbol{x})\in Q$ be the initial date. Let $v^{\alpha}(t,\boldsymbol{x})$ denote the value function, i.e.:

$$v^{\alpha}(t,\boldsymbol{x})=\min_{\boldsymbol{u}(.)\in\mathcal{A}(.)}J^{\alpha}(t,\boldsymbol{x};\boldsymbol{u})=\min_{\substack{\boldsymbol{u}(s,\boldsymbol{x})\in\mathcal{A}(s,\alpha)\\ t\le s\le T}}E_u\left[\int_t^T e^{-\rho(s-t)}g(\boldsymbol{x}(s),\boldsymbol{u}(s))ds \mid \boldsymbol{x}(t)=\boldsymbol{x},\xi(t)=\alpha\right]. \quad (3.8)$$

Using dynamic programming, the value function in (3.8) is generalized to the following theorem:

Theorem 3.3.1 *The stochastic control problem satisfies the system of partial differential equations:*

$$\rho v^{\alpha}(t,\boldsymbol{x})=\min_{\boldsymbol{u}(t,\boldsymbol{x})\in\mathcal{A}(t,\alpha)}\left\{g(\boldsymbol{x},\boldsymbol{u})+v_t^{\alpha}+v_{\boldsymbol{x}}^{\alpha}f^{\alpha}(t,\boldsymbol{x},\boldsymbol{u})-p_{\alpha\alpha}(t)v^{\alpha}(t,\boldsymbol{x})+\sum_{\beta\neq\alpha}p_{\alpha\beta}(t)v^{\beta}(t,\boldsymbol{x})\right\}, \quad (3.9)$$

at time t, the initial and boundary conditions are satisfied:

$$\begin{cases} (\boldsymbol{x}(t),\xi(t)) = (\boldsymbol{x},\alpha) \ \text{ for } \ (t,\boldsymbol{x}) \in Q, \alpha \in \mathcal{I} = \{0,1,2,3\}, \\ v^{\alpha}(T, x(T)) = 0. \end{cases} \tag{3.10}$$

In equation (3.9), the terms $v_t^{\alpha}(t,\boldsymbol{x})$ *and* $v_{\boldsymbol{x}}^{\alpha}(t,\boldsymbol{x})$ *denote the gradient of the value function with respect to time t and state variables, respectively.*

Proof. The proof of this theorem is presented in Appendix III.

Remark 3.1 (i) The system of partial differential equations (3.9) is the well-known *HJB equation*; (ii) It depends not only on the state variable of the system, but also on their time variation because of the dynamics of semi-Markov decision processes such as $p_{\alpha\beta}(s)$.

Hence, in order to characterize the optimal control, we review the concepts, and the following results represent some properties of the value function $v^{\alpha}(t, \boldsymbol{x})$ that are needed in order to address the main results on feedback control analysis.

Theorem 3.3.2 (i) *For each* $\alpha \in \mathcal{I}$, *there exists a constant* C_1, *such that the value function satisfies:*

$$\left|v^{\alpha}(t,\boldsymbol{x})\right| \le C_1 (T-t) \ \textit{ for every } (t,\boldsymbol{x}) \in Q. \tag{3.11}$$

(ii) *For each* $\alpha \in \mathcal{I}$, *there exists a constant* C_M, *for every* $(t,\boldsymbol{x}),(\bar{t},\bar{\boldsymbol{x}}) \in Q$, *such that the value function satisfies the following condition:*

$$| v^{\alpha}(t,\boldsymbol{x}) - v^{\alpha}(\bar{t},\bar{\boldsymbol{x}}) | \le C_M \left(| t-\bar{t} | + | \boldsymbol{x}-\bar{\boldsymbol{x}} | \right). \tag{3.12}$$

Proof. The proof of this theorem is presented in Appendix III.

As we defined important measures, we finally considered the stochastic optimal control of a two-machine flowshop in (3.8) with the initial condition ($\boldsymbol{x}(t)$, $\xi(t)$) = ($\boldsymbol{x}$, α). For this, we established the following verification theorem and requirements which meet the HJB equation (3.9). In addition, $v^{\alpha}(t, \boldsymbol{x})$ has a unique solution for all $\boldsymbol{x}$ and t in the relevant ranges; that is equal to the minimum expected total cost among appropriately defined classes of the admissible control law of the system.

Theorem 3.3.3. *Let* $v^{\alpha}(t, \boldsymbol{x}) \in \mathcal{I} \times Q$ *be a solution to* (3.9). *Then for all* $(t, \boldsymbol{x}) \in Q$

(i) $v^{\alpha}(t, \boldsymbol{x}) \le J^{\alpha}(t, \boldsymbol{x}; \boldsymbol{u})$ *for every admissible control system* $\boldsymbol{u}(t, \boldsymbol{x})$.

(ii) If there exists an admissible system $\boldsymbol{u}^{*}(t, \boldsymbol{x})$ *such that*

$$\boldsymbol{u}^{*}(t, \boldsymbol{x}) \in \arg \min_{\boldsymbol{u}(t, \boldsymbol{x}) \in \mathcal{A}(t, \alpha)} \left\{ g(\boldsymbol{x}, \boldsymbol{u}) + v_t^{\alpha} + v_x^{\alpha} f^{\alpha}(t, \boldsymbol{x}, \boldsymbol{u}) \right\}, \tag{3.13}$$

almost everywhere in t with the probability 1, then

$$v^{\alpha}(t, \boldsymbol{x}) \le J^{\alpha}\left(t, \boldsymbol{x}, \boldsymbol{u}^{*}\right). \tag{3.14}$$

Proof. The proof of this theorem is presented in Appendix III.

The optimality conditions established in (3.9) lead to a feedback control. In practice, the feedback control is indispensable to handle the inaccuracies and uncertainties (including stochastic phenomena) that are present in design process, and to make full use of the capacity of the equipment (see Engell (2007)).

3.4 Dynamic system

This section describes the dynamics of the manufacturing problem and explicitly relates the HJB equation to the control structure.

Let us assume that each machine is subject to random failures and repairs and has two finite states α_i; i = 1,2 (i.e., available $\alpha_i = 1$ and unavailable $\alpha_i = 0$) with probability distributions

$F_i(t)$ and $G_i(t)$ respectively applicable to each state. Computing of the value function $v^{\alpha}(t, \boldsymbol{x})$ and the production rate $\boldsymbol{u}^{\alpha}(t, \boldsymbol{x})$ includes all states. Let $\xi(t) = \{\alpha_i, t \geq 0, i = 1, 2\}$ be defined by the following expression:

$$\xi(t) = \begin{cases} 3 & \text{if Both machines are operational,} \\ 2 & \text{if } \mathrm{M}_2 \text{ is operational, and } \mathrm{M}_1 \text{ down,} \\ 1 & \text{if } \mathrm{M}_1 \text{ is operational, and } \mathrm{M}_2 \text{ down,} \\ 0 & \text{if None of them is operational.} \end{cases} \tag{3.15}$$

Thus, the transition graph of finite states is shown in Figure 3.2.

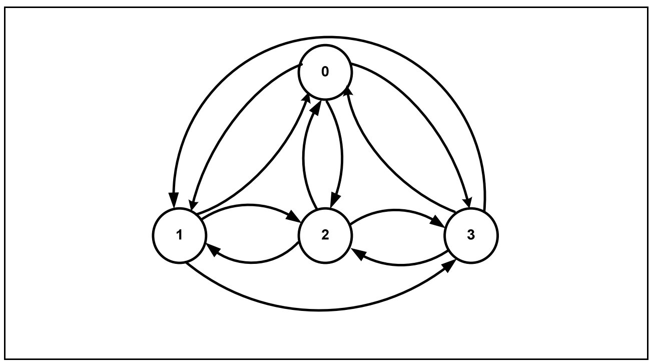


Figure 3.2 Transition graph of two-machine system with four states.

To determine the probability and transition probability distributions of the two-machine system, we will replace two-state machines (state 0 and state 1) by four-state machines (state 0, state 1, state 2 and state 3).

The conventional implementation of *state machines* is based on the selection of successor states and the execution of each *related* action. The machine states are described as follows: *State* 0: the system is not operating (down); *State* 1: the system corresponds to the first machine, which is operating and available to produce under the limited buffer level $0 \leq x_1(t) \leq B$; *State* 2: the system corresponds to the second machine, which is operating and available to produce if the buffer level $x_1(t) > 0$; and *State* 3: the system is equivalent to the deterministic two-machine flowshop. In this case, we assume that all transition times are not exponentially distributed.

Using the reliability theory in Ross (2003), the dynamics of the two-parallel-machine model is described as similar to that of one equivalent machine as follows:

$$P_{\xi(t)}(t) = \Pr\{\text{System is functioning at time } t,\ t \geq 0\},$$

$$P_{\xi(t)}(t) = \prod_{i=1}^{2} \Pr\{\text{Machine } i \text{ is functioning at time } t \geq 0\}. \qquad (3.16)$$

Let $P_3(t)$, $P_2(t)$, $P_1(t)$, and $P_0(t)$ be the probability that both machines are operational, only the second machine is operational, only the first machine is operational, and neither of them is operational, respectively. These results in twelve transition probabilities, as shown in Figure 3.2. The derivative of the semi-Markov transition probability $p_{\alpha\beta}(t)$ is defined by equation (3.5) where $\alpha, \beta = 0,1,2,3$. The calculation of $P_\alpha(t)$, and $p_{\alpha\beta}(t)$ is presented in Appendix III.

In Figure 3.2, *at state* 3: the future state of the system may be that both machines are "down" (3→0) or may be that either the first machine is "up" while the second is "down" (3→1) or the second is "up" while the first is "down" (3→2). *At state* 2: the future state of the system may be that both machines are "down" (2→0) or may be that either the first machine is "up" while the second is "down" (2→1) or both of them are "up" (2→3). *At state* 1: the future state of the system may be that both machines are "down" (1→0) or may be that either the first machine is "down" while the second is "up" (1→2) or both of them are "up" (1→3). *At state* 0: the future state of the system may be that both machines are "up" (0→3) or may be either that the first machine is "up" while the second is "down" (0→1) or the second is "up" while the first is "down" (0→2).

The value function of two-machine flowshop is described by the HJB equation (3.9) with four states as follows:

$$\begin{cases} \text{at state 0: set } u_1(.) = 0, u_2(.) = 0 \\ \rho v^0(.) = \left\{ g(.) + v_t^0(t) - d(t).v_{x_2}^0(.) + \sum_{\beta=0}^{3} \delta(t) p_{0\beta}(t) v^\beta(.) \right\}, \end{cases} \quad (3.17)$$

$$\begin{cases} \text{at state 1: set } u_1(.) \geq 0, u_2(.) = 0 \\ \rho v^1(.) = \min_{0\leq u\leq r} \left\{ g(.) + v_t^1(.) - d(t) v_{x_2}^1(.) + u_1(.) v_{x_1}^1(.) + \sum_{\beta=0}^{3} \delta(t) p_{1\beta}(t) v^\beta(.) \right\}, \end{cases} \quad (3.18)$$

$$\begin{cases} \text{at state 2:set } u_1(.) = 0, u_2(.) \geq 0 \\ \rho v^2(.) = \min_{0\leq u_2(t)\leq r_2} \left\{ g(.) + v_t^2(.) + (u_2(.) - d(t)) v_{x_2}^2(.) - u_2(.) v_{x_1}^2(.) + \sum_{\beta=0}^{3} \delta(t) p_{2\beta}(t) v^\beta(.) \right\}, \end{cases} \quad (3.19)$$

$$\begin{cases} \text{at state 3:set } u_1(.) \geq 0, u_2(.) \geq 0 \\ \rho v^3(.) = \min_{\substack{0\leq u_2(t)\leq r_2 \\ 0\leq u_1(t)\leq r_1}} \left\{ g(.) + v_t^3(.) + (u_1 - u_2) v_{x_1}^3(.) + (u_2 - d) v_{x_2}^3(.) + \sum_{\beta=0}^{3} \delta(t) p_{3\beta}(t) v^\beta(.) \right\}, \end{cases} \quad (3.20)$$

where $\delta(t) = 1$ if $\alpha \neq \beta$; $\delta(t) = -1$ if $\alpha = \beta$; $\alpha, \beta = 0,1,2,3$.

3.5 Hedging point policy with feedback control

In this section, we describe the hedging point policy whose solution leads to the deterministic problem and bang-bang control characteristics. It is based on the HJB equation (3.9), which is linear in production rates and satisfies Bellman principle of optimality. The solution of the first-order partial derivative in (3.9) is not simple. To interpret the value function $v^\alpha(t, \boldsymbol{x})$, the concept of viscosity solution is often used. For more information and discussion of the concept of viscosity solution, the reader is referred to Sethi et al. (2005) and Fleming and Soner (2006). However, in this paper, we use a heuristic method in order to overcome the solution of the multivariable problem in (3.9). The idea is to divide the multivariable problem in (3.9) into two different problems with each one having a single variable control, and corresponding to Akella and Kumar's optimal solution as in hedging point policy (Akella and Kumar (1986)).

Further simplification of the equation (3.9) is addressed by determining a control $\boldsymbol{u}(t,\boldsymbol{x})$ through the following linear program:

$$\min_{\boldsymbol{u}(.)\in\mathcal{A}(t,\alpha)}\left\{\frac{\partial v^{\alpha}(.)}{\partial x_1}\big(u_1(.)-u_2(.)\big)+\frac{\partial v^{\alpha}(.)}{\partial x_2}\big(u_2(.)-d(.)\big)\right\}, \tag{3.21}$$

subject to equations (3.2) and (3.3) above.

The optimal feedback control (3.21) is designed to drive the system to the hedging point. If the system state is $\xi(t) = 0$, at which all machines are down, we must have $\boldsymbol{u}(t,\boldsymbol{x}) = 0$. Whenever the system state is $\xi(t) = \alpha$, the linear program in (3.21) presents a real-time feedback controller, and the production rate is calculated at every time instant with $\xi(t) \neq 0$ either according to varying demand or to constant demand. We assume that if the buffer is neither empty nor full, the choice of u_1 should be independent of x_2 and α_2; the choice of u_2 should be independent of x_1 and α_1 (see Gershwin (2002), page 543). Obviously, since (3.21) is linear in $\boldsymbol{u}(t,\boldsymbol{x})$, we obtain the following systems:

$$\min_{u_1(.)\in\mathcal{A}(t,\alpha)}\left[\frac{\partial v^{\alpha}(.)}{\partial x_1}\big(u_1(.)-u_2(.)\big)\right], \tag{3.22}$$

$$\min_{u_2(.)\in\mathcal{A}(t,\alpha)}\left[\frac{\partial v^{\alpha}(.)}{\partial x_2}\big(u_2(.)-d(.)\big)\right]. \tag{3.23}$$

The point-$\boldsymbol{x}$ space at which the gradient of $v^{\alpha}(t, \boldsymbol{x})$ is equal to zero is called the *Hedging point* $\boldsymbol{z}^*(.)$. To use the result of a single machine system in Kimenia and Gershwin (1983) as well as in Alkella and Kumar (1986), the equation (3.23) is first solved, and then the equation

(3.22) is solved. The optimal control problem (3.23) was established by Kimemia and Gershwin (1983) established the optimal production rate $u_2^*(t,x_2)$ as follows:

$$u_2^*(.) = \begin{cases} r_2 & \text{if} \quad x_2(t) < z_2^* \ \& \ \xi_2(t) = 1 \\ d(t) & \text{if} \quad x_2(t) = z_2^* \ \& \ \xi_2(t) = 1 \\ 0 & \text{if} \quad x_2(t) > z_2^* \ \& \ \text{otherwise} \end{cases}, \tag{3.24}$$

where the hedging point, z_2^* is determined by minimizing the following objective function:

$$J(t,\boldsymbol{x}) = \lim_{T\to\infty} \frac{1}{T}\left[\int_t^T g(\boldsymbol{x}(s),\boldsymbol{u}(s))\,ds\right], \tag{3.25}$$

with respect to $x_2(t)$, $0 \le t < T$ (see Akella and Kumar (1986) and Bielecki and Kumar (1988)).

Applying a similar analysis to the optimal control problem (3.22), the optimal production rate $u_1^*(t,x_1)$ is established using the following proposition:

Proposition 3.5.1. *Let y^* be a solution to* (3.25) *respect to such that $y(t) = x_1(t)+ x_2(t)$. Then the optimal production rate* $u_1^*(t,x_1)$ may be *given as follows:*

$$u_1^*(.) = \begin{cases} r_1 & \text{if} \quad x_1(t) < z_1^* \ \& \ \xi_1(t) = 1, \\ d(t) & \text{if} \quad x_1(t) = z_1^* \ \& \ \xi_1(t) = 1, \\ 0 & \text{if} \quad x_1(t) > z_1^* \ \& \ \text{otherwise}. \end{cases} \tag{3.26}$$

where

$$z_1^* = y^* - z_2^* \tag{3.27}$$

Proof. The proof of this proposition is presented in Appendix III.

In the expression (3.26) the variable z_1^* is the hedging point of the WIP on the first machine. From the maximum production capacities ($r_1 > r_2$), we assume that the hedging point of WIP z_1^* is a finite non-negative value. Figure 3.3 presents the hedging point policy. It has been extended to the hedging point policy from results in Sethi et al. (1997) and contains four different zones which discuss in what follows.

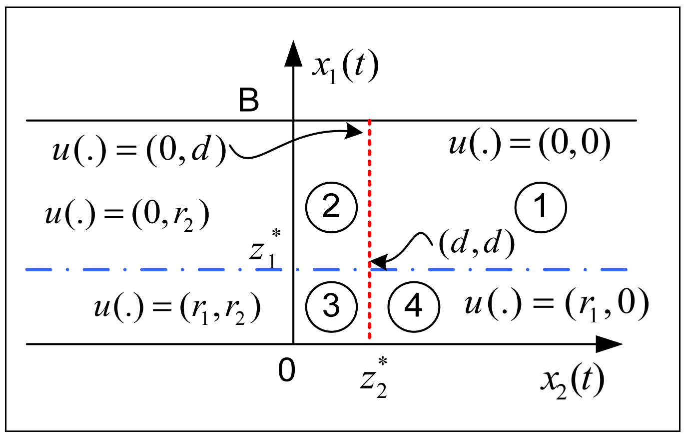


Figure 3.3 Hedging point policy.

The control variable in the first zone is (0,0) when $z_1^* < x_1 \leq B$ and $x_2 > z_2^*$; here, there is no need to produce because the buffer level is high enough. However, when $z_1^* < x_1 \leq B$, $x_2 < z_2^*$, there is need to produce on the second machine only. As a result, the control variable in this second zone is (0, r_2). The control strategy at the hedging point (z_1^*, z_2^*) is (d, d), which results in minimum objective function (3.25) being used. In zone three ($0 \leq x_1 < z_1^*$ and $x_2 < z_2^*$), there is need to produce with $\boldsymbol{u}^* = (r_1, r_2)$; here, the control policy should be set to rapidly reduce the shortage while keeping the upper-zero inventory in the buffer of the first machine. When $z_1^* < x_1 \leq B$ and $x_2 = z_2^*$, the control strategy is (0, d) because of the condition $c_1 < c_2^+$. Finally, in zone four, the control is (r_1, 0).

3.6 System behaviour under the optimal policy

In practice, it is difficult to determine an optimal control with all four discrete states. While the Kanban control is only considered when the system is deterministic, the CONWIP control is applied to systems with constant buffer levels (see Bonvik (1996)). In the system presented in this paper, we have both stochastic dynamics and a finite buffer level, and so we therefore intend to apply the heuristic control for each state as shown above.

3.6.1 Analysis of state 1 of the system

The control structure of the system is conditioned by machine 1. Only the first machine is operational while the second one is down. The behaviour of the surplus trajectory depends on $x_1(t)$ at time t. The first machine is blocked when $x_1(t) = B$. When $x_1(t) < B,$ the first machine is ready to produce and the characteristic of the time saved is as follows:

$$ST_1(u_1) = \frac{B - x_1(.)}{u_1(.)}. \tag{3.28}$$

This characteristic time depends on the control policy u_1 and the current WIP $x_1(t)$, as in Figure 3.4. It means the *time saved* in which the first machine can only produce a number of parts under-bound B. Using (3.28), when $x_1(t) = 0$, the minimum time saved is given by:

$$ST_{1B}^{\min} = \min_{u_1(.)}\left\{\frac{B}{u_1}\right\} = \frac{B}{r_1} \tag{3.29}$$

If we decide on the control at time t (t is current time), the real time of the minimum time saved is determined by:

$$ST_{1B}^{real} = t + ST_{1B}^{\min}\,,\ 0 \le t \le T. \tag{3.30}$$

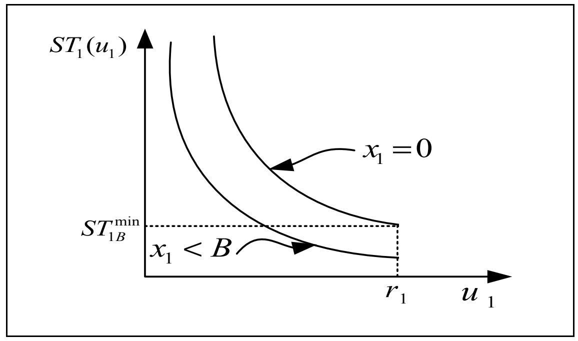


Figure 3.4 Characteristic of the time saved of M_1.

Example 6.1 Consider the system at time t; assume B = 10 parts, r_1 = 0.2 part/time unit, $x_1(t)$ = 0, then: $ST_{1B}^{\min}$ = 10/0.2 = 50 time units and ST_{1B}^{real} = t + 50. It means that after 50 time units from t, the first machine had produced 10 parts with maximal capacity $u_1 = r_1$ and it stops at 50 + t because the WIP is equal to upper bound B = 10.

At the hedging point z_1^*, the production rate is equal to the demand rate. The time saved at the hedging point is also called the Hedging time $ST_{z_1}^* = z_1^* / d$, i.e., it is the time unit that machine requires to produce parts z_1^*. For this policy $u_1^* = d$, the time saved at the hedging point is determined by:

$$ST_1^* = \frac{B - z_1^*}{d} = \frac{B}{d} - ST_{z_1}^*. \tag{3.31}$$

3.6.2 Analysis of state 2 of the system

The behaviour of this state is presented in Figure 3.5. The control structure of the system is conditioned by machine 2. Only the second machine is operational, while the first is down. The behaviour of the surplus trajectory depends on $x_1(t)$ at time t. Because the first machine is down, the second machine is starved when $x_1(t)$ = 0. When $x_1(t)$ > 0 the second machine is available to produce and satisfy the demand $d(t)$ at time t. Its time saved characteristic can be written as follows:

$$ST_2(u_2) = \frac{x_1(t)}{u_2(.)}. \tag{3.32}$$

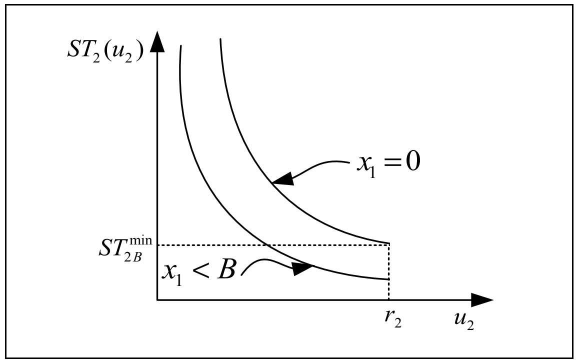


Figure 3.5 Characteristic of the time saved of M_2.

The production with the hedging point policy in (3.24) may be adopted, but it must depend on its characteristic of the time saved. Under a similar condition as in state 1, when $x_1 = B$, the minimum time saved and its real time are given by:

$$ST_{2B}^{\min} = \min_{u_2(t)} \left\{ \frac{B}{u_2} \right\} = \frac{B}{r_2}, \tag{3.33}$$

$$ST_{2B}^{real} = t + ST_{2B}^{\min}, 0 \le t \le T. \tag{3.34}$$

The time saved at the hedging point is:

$$ST_{z_2}^{*} = \frac{z_2^{*}}{d}. \tag{3.35}$$

Remark 6.1 (*i*) The hedging times, ST_i^* for i = 1,2, refer to the time saved in which the machine can produce a number of parts $ST_i^*.d$; (*ii*) The components $ST_i\ (u_i)$, for i = 1, 2, refer to the time saved in which the machine can only produce in interval [0, $ST_i(.)$] with any policy u_i; (*iii*). Meanwhile, the terms $ST_i^{\min}(.)$, for i = 1, 2, refer to the minimum time saved that the machine can produce a number of parts $ST_i^{\min}.r_i$; (*iv*) The processing time of states 1

& 2 depends implicitly on the size of the buffer between the two machines (B) and the current WIP $x_1(t)$.

3.6.3 Analysis of state 3 of the system

At this state, both machines are up, and the optimal control problem becomes a deterministic problem. Consequently, the optimal feedback control is considered. We then have the following three cases.

a) If the buffer $x_1 = B$ is full, the choice of u_1 depends on the capacity of the second machine. Here, the way to go is to produce with $u_1 = 0$ because the second machine has an amount of time saved $ST_{2B}^{\min}$ to produce $ST_{2B}^{\min}.u_2(.)$ parts. We can approximate the optimal policy for the second machine by using the model as in Akella and Kumar (1986):

$$u_2(.) = \begin{cases} 0 & \text{if} \quad x_2(t) > z_2^* \\ d(t) & \text{if} \quad x_2(t) = z_2^* \\ r_2 & \text{if} \quad x_2(t) < z_2^* \end{cases} ,t \in \left[0, ST_2\left(u_2\right)\right].$$

This policy corresponds to the Kanban control, where the first machine is instructed to stop production.

b) If the buffer $x_1 = 0$ is empty, it must produce $u_1(.) = r_1$ because the second machine needs $1/r_1$ time units before it can continue to produce with any policy $u_2(.)$, then $dx_2(t)/dt = -d(t) < 0$ for $t \in [0, 1/r_1]$. In this time interval $[0, 1/r_1]$, the second machine is called a *starved machine*, and the production is only determined after $1/r_1$ time units.

c) When the buffer $x_1(t)$ is neither empty nor full, the feedback optimal control in equations (3.24) & (3.26) is applied. Here, the policy may correspond to the CONWIP control, and the proper production path consists of responding to actual demands $u_1(.) = u_2(.) = d(.)$ (i.e., corresponding to hedging point policy).

3.6.4 Analysis of state 0 of the system

This state corresponds to the case where both machines are down: $dx_1(t)/\mathrm{d}t = 0$, $dx_2(t)/dt = -d(t) < 0$ and the buffer level becomes constant.

3.7 Result analysis with constant demand

This section aims to illustrate the validity of the results of the proposed model by using the numerical method with constant demand rate $d(t)$. To that end, we consider a problem with a two-machine flowshop system producing a single part-type (in Section 3.2). The dynamics of the system was described in Section 3.4, and it has four discrete states.

The up-downtimes are distributed according to one of the following three probability density functions, referred to as reliability models:

i. Weibull, i.e.,

$$f^W(t) = \lambda^\mu \mu t^{\mu-1} \exp\left(-(\lambda t)^\mu\right).$$

This distribution is denoted as $W(\mu, \lambda)$.

ii. Log-normal, i.e.,

$$f^L(t) = \frac{1}{t\sigma\sqrt{2\pi}} \exp\left(-\left\{\frac{\left[\ln(t) - \lambda_o\right]^2}{2\sigma^2}\right\}\right).$$

This distribution is denoted as $L(\lambda_o, \sigma)$.

iii. Gamma, i.e.,

$$f^G(t) = \frac{\lambda_1^{\mu_1} t^{\mu_1 - 1} \exp\{-\lambda_1 t\}}{\Gamma(\mu_1)}.$$

This distribution is denoted as $G(\mu_1, \lambda_1)$.

We ensure that the failure rate and the repair rate are function of time, and the coefficient of variation, CV, is less than one. We then present the set of up-down times used in this example in Table 3.2, which also shows their coefficients of variation, CV (which take values less than one and are equal to CV_W, CV_L, and CV_G with $CV_W = 0.93$, $CV_L = 0.95$, and $CV_G = 0.57$), the MTTF (Mean Time to Fail) and the MTTR (Mean Time to Repair). The parameters of the system are as follows:

- maximal production rate $r_1 = 0.25$, $r_2 = 0.225$,
- unit inventory cost of the internal buffer $c_1 = 0.5$,
- unit surplus cost of the finished product in the external buffer $c_2^+ = 1$,
- unit backlog penalty of the finished product and $c_2^- = 2$,
- discount rate $\rho = 0.65$,
- demand rate $d = 0.145$.

Table 3.2 Up-Downtime distributions considered on Machines 1 & 2

Case	**Uptime**	**Downtime**	**MTTF**	**MTTR**
A	L (4.3, 0.80)	W (1.08, 0.009)	100	10
B	G (3.15,0.0315)	W (1.08, 0.009)	100	10

We use the numerical method based on the Kushner and Dupuis (2001) approach because it is very difficult to solve the HJB equation with an analysis model. Let $\Delta x_k > 0$ for $k = 1,2$, and $\Delta t > 0$ denote the lengths of the finite difference intervals of the variables x_k and t, respectively. The first-order partial derivatives of the value function in equations (3.17)-(3.20) are replaced by the following expressions:

$$v_{x_k}^{\alpha}(t,\boldsymbol{x}) = \begin{cases} \dfrac{v^{\alpha}(t, x_k + \Delta x_k) - v^{\alpha}(t,\boldsymbol{x})}{\Delta x_k} & \text{if } \dot{x}_k \geq 0 \\ \dfrac{v^{\alpha}(t,\boldsymbol{x}) - v^{\alpha}(t, x_k - \Delta x_k)}{\Delta x_k} & \text{otherwise} \end{cases}, \tag{3.36}$$

$$v_t^{\alpha}(t,\boldsymbol{x}) = \frac{v^{\alpha}(t+\Delta t,\boldsymbol{x}) - v^{\alpha}(t,\boldsymbol{x})}{\Delta t}. \tag{3.37}$$

For details of this method, the reader is referred to Kushner and Dupuis (2001). The results of this example are illustrated in Figures 3.6-3.11 with given internal values of $x_1 \subset [0, 20]$, $x_2 \subset [-20, 20]$, and $t \subset (0, 500)$.

3.7.1 Interpretation of the results for case A in Table 3.2

This corresponds to case A in Table 3.2 above. Figures 3.6 and 3.7 represent the production rates $u_1(t,x_2)$ and $u_2(t,x_2)$ versus surplus level x_2 and time t at $x_1 = B = 20$ parts. Figures 3.8 and 3.9 represent the production rates $u_1(x_1, x_2)$ and $u_2(x_1, x_2)$ versus WIP x_1 and surplus level x_2 at $t = 205$ time units, where this time t is chosen arbitrarily from within the time interval (0, 500). Simulation results correspond to the hedging point policy as in the expressions (3.24) and (3.26).

In Figure 3.6, the production rate of M_1 $u_1(t,x_2)$ is equal to zero, which corresponds to zones 1 and 2 in Figure 3.3 (i.e., $x_1 \geq z_1^*$). In Figure 3.7, the production rate of M_2 $u_2(.)$ is equal to maximum at ($r_2 = 0.225$), while the surplus level x_2 is less than zero (zones 2 and 3 in Figure 3.3), and is equal to $d = 0.145$ when $x_2 = z_2^*$=0.95 parts. However, this rate is equal to zero when the surplus level is more than 0.95 parts over time (0, 500).

In Figure 3.8, the production rate $u_1(.)$ is equal to maximum ($u_1 = r_1 = 0.25$) when the WIP x_1 = 0, and is equal to zero if $z_1^* < x_1 \leq B$, which corresponds to zones 1 and 2 in Figure 3.3. Figure 3.9 presents the production rate of M_2 ($u_2(x_1,x_2)$) versus x_1 and x_2. The optimal policy at the hedging point z_2^* is found within the whole interval (0,500) with $u_2^* = d = 0.145$ and $z_2^* = 0.95$ parts. The value of z_k^* is computed as follows:

$$z_k^*(t) = \arg \min_{\substack{x_1 \subset [0,20) \\ x_2 \subset (-20,20)}} \left(v^3 \left(t, \boldsymbol{x} \right) \right) \text{ for } k = 1,2. \tag{3.38}$$

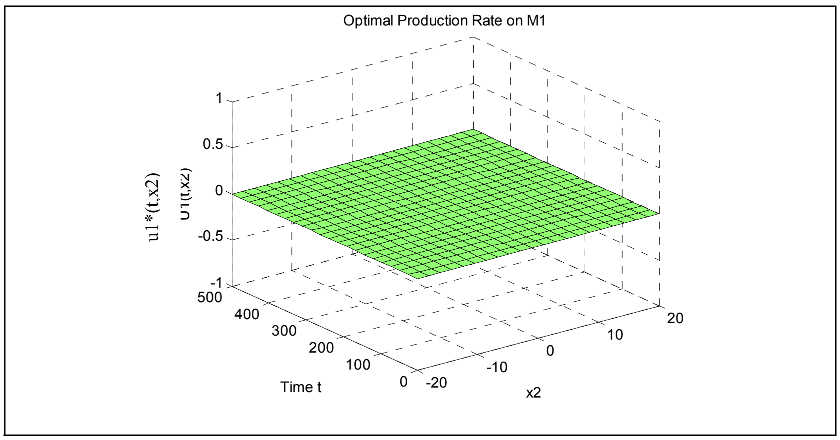


Figure 3.6 Optimal Production Rate for u_1(.) on M_1 versus t and x_2 at $x_1 = B$.

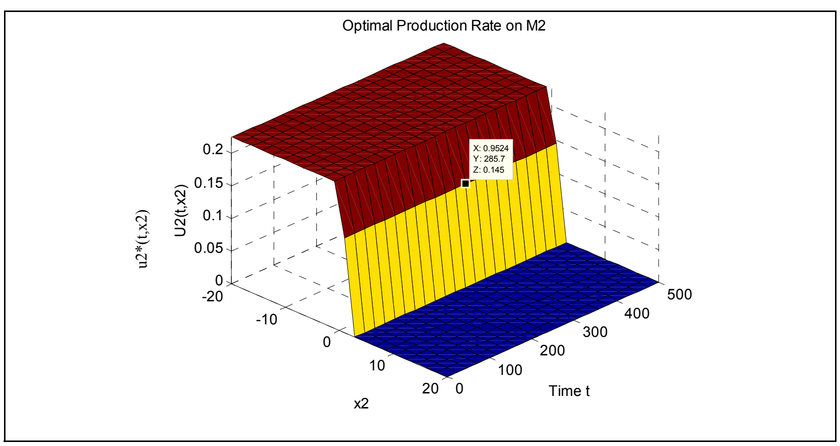


Figure 3.7 Optimal Production Rate for u_2(.) on M_2 versus t and x_2 at $x_1 = 0$.

We also obtain the hedging point on the first machine $z_1^* = 0$, and then the value $y^* = z_2^* = 0.95$ parts.

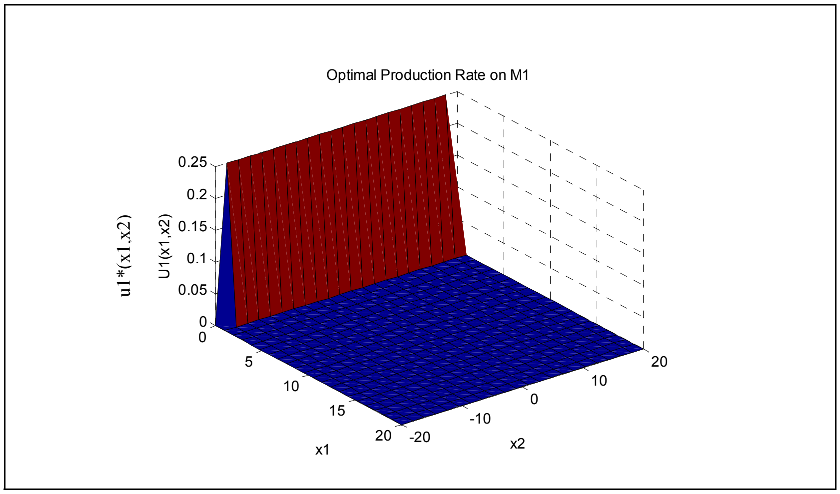


Figure 3.8 Optimal Production Rate for u_1(.) on M_1 versus x_1 and x_2 at $t = 205$ time units.

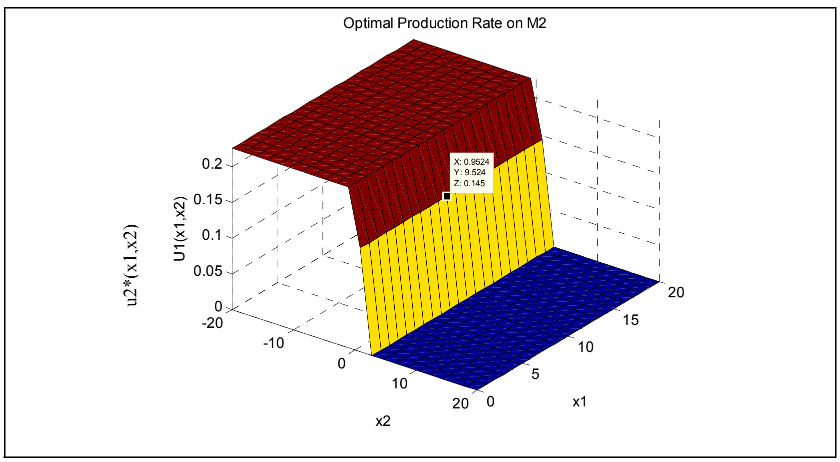


Figure 3.9 Optimal Production Rate for u_2(.) on M_2 versus x_1 and x_2 at $t = 205$.

Results obtained show that the optimal control law is similar to the bang-bang control problem when the surplus level varies in time $t < T < \infty$. Note that the hedging point policy $u_2^*\left(t, z^*\right) = d$ is valid over time. The value of the optimal production rate u_2(.) is greater than zero when the system is in state 3 (i.e., both machines are up), and is equal to zero when the system in states 0 and 1 (i.e., machine M_2 is down). On other hand, the value of u_1(.) is expressed analogically as u_2(.), but it is equal to zero when the system in states 0 and 2. When the system is in state 0, 1, or 2 the heuristic policy described in sub-Section 3.6.1 and

3.6.2 is used. As results, the optimal policy agrees with the analytical model developed in Akella and Kumar (1986).

3.7.2 Interpretation of the results for case B in Table 3.2

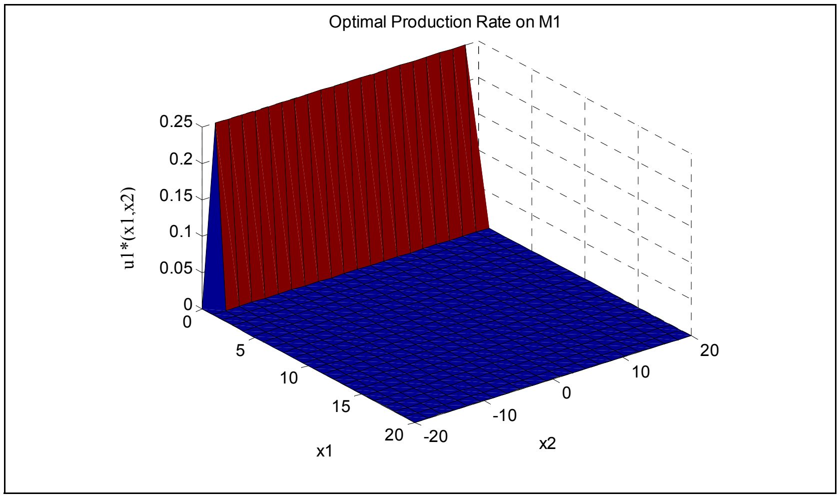


Figure 3.10 Optimal Production Rate for u_1(.) on M_1 versus x_1 and x_2 at t = 205 time units.

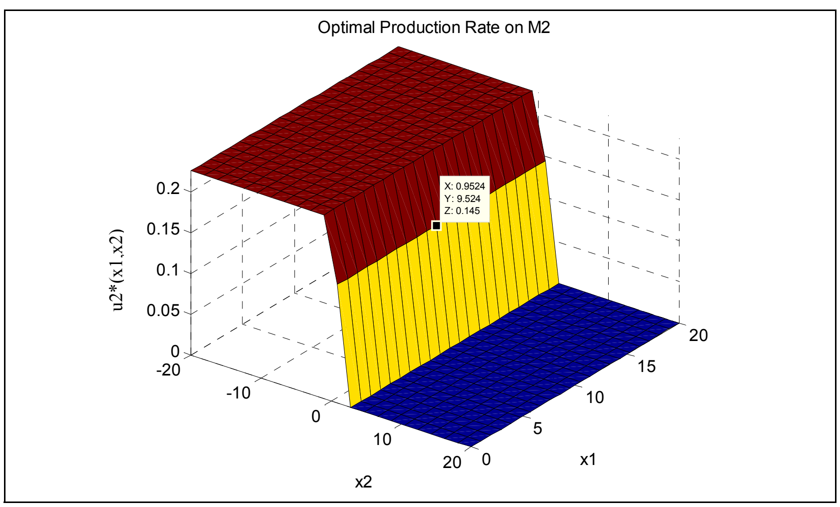


Figure 3.11 Optimal Production Rate for u_2(.) on M_2 versus x_1 and x_2 at t = 205 time units.

The results of this case are similar to the case when the up-downtime obeys log-normal and Weibull distributions, as in Figures 3.10 and 3.11 because the use the same MTTF and

MTTR. Both cases A and B in Table 3.2 (i.e., where log-normal and gamma distributions are used for machine uptimes) and in Figures 3.8 and 3.10 show that the machine must produce at a maximum production rate $u_1 = r_1$ when the WIP level is equal to zero.

3.8 Result analysis with varying demand

This example is to illustrate the optimum cost values for different demand scenarios with the varying demand rate with time *t* for which data is given in Table 3.3. The parameters of the system are the same as in constant demand rate with case A in Table 3.2.

Table 3.3 Data of the varying demand rate

t	0-100	100-200	200-300	300-400	400-500
$d(t)$	0.125	0.125	0.145	0.175	0.175

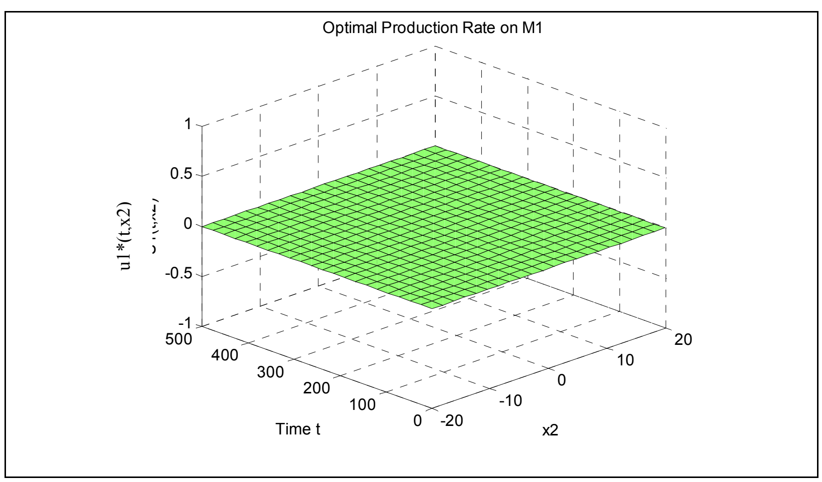


Figure 3.12 Optimal Production Rate for $u_1(.)$ on M_1 versus t and x_2 at $x_1 = B$.

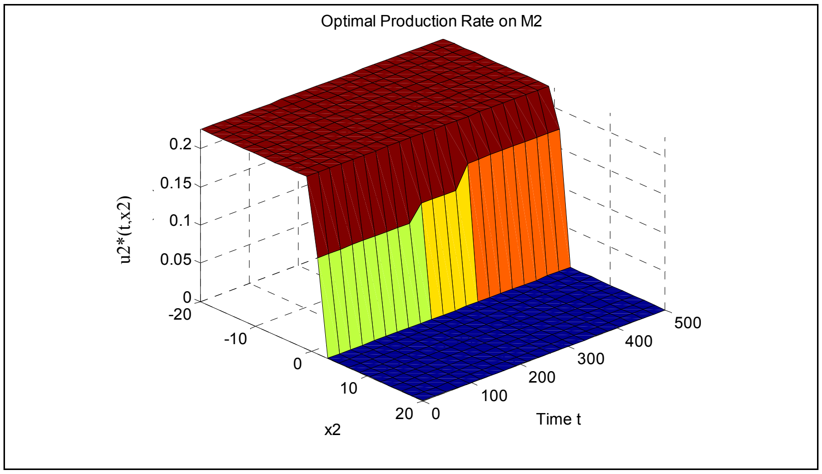


Figure 3.13 Optimal Production Rate for u_2(.) on M_2 versus t and x_2 at $x_1 = B$.

Results of this example are shown in Figures 3.12-3.13 as the optimal production rates for M_1 and M_2, u_1(.), u_2(.) versus the time and the surplus level at $x_1 = B = 20$. At $x_1 = B = 20$ parts, the production rate for the first machine shown in Figure 3.12 is equal to zero for every time t and x_2. In Figure 3.13, the production rate of the second machine u_2(.) is equal to the maximum value when the surplus level x_2 is less than zero, and is equal to zero when the surplus level is greater than the hedging point value $z_2(t)$. This value z_2(.) is equal to 0.95 when $u_2(.)= d(t)$ over time (0, 500). The production rate u_2(.) fluctuates in *time* because of the varying demand $d(t)$ with the bang-bang control, as in Figure 3.13.

3.9 Conclusion

3.9.1 Summary and extensions

In this paper, the optimal stochastic production problem for a two-machine flowshop, single-product manufacturing system has been considered, with machines subject to random breakdowns and repairs. Using Markov properties, we have formulated a new model in the form of a stochastic control problem by adopting Rishel's assumptions to model discrete machine states, which are characterized by semi-Markov jumps, and by using a dynamic programming approach to make decisions at different stages over time. The objective was to find the production rate for upstream and downstream machines while minimizing surplus costs by using a semi-Markov process (i.e., Markov properties). The optimality conditions were established using Pontryagain's principle, and led to the HJB equation. In effect, the production control is the feedback control, for the control variables of two production rates are linear over time.

The heuristic approach presented seeks to improve the complexity of the HJB equation when the system has stochastic control variables and it makes the problem deterministic. We also provided an analysis of the hedging point policy for the feedback controller. We applied our proposed model to a real word manufacturing system with machines having Weibull, log-normal, and gamma distributions. In what follows, we discuss the other extensions in our model, which is very important.

While the classical Markov model (see Rishel (1975) and Davis (1984)) has been considered as a stochastic optimal control with homogenous Markov jumps, the Boukas proposition (see Boukas (1987)) gave an analysis of stochastic problems with non-homogenous Markov jumps. Obviously, both homogenous and non-homogenous Markov jumps have constant transitions. That led to a model that is not time-dependent even through the control problem is considered in continuous time. However, we have extended to the stochastic problem in the continuous-time optimal problem with semi-Markov jumps, i.e., the transition rates between the machine states as well as the transition probabilities are time-dependent. Hence, the optimality conditions do not depend only on the system states, but also on the time t. This first extension can enable us to consider the failure and repair rates as being functions of time instead, and to thereby improve the coefficient of variation CV for the up-downtime distribution which is less than one. This is very different from the Markov framework, and can lead to a high system performance (see Li and Meerkov (2005a,b)). A rich body of works exists in the literature, examining semi-Markov processes, and these go back fifty years. They include the following: A detailed theoretical analysis of semi-Markov processes is described in Howard (1971); Glynn (1989) considered a *generalized semi-Markov process* (GSMP) of discrete events. Abbad and Filar (1991) presented the *semi-Markov control* *problem* (SMCP) using an infinite horizon approach with discounted rewards and showed that the SMCP and the Markov control problem (MCP) are of the same ergodic class. Recently, D'Amico et al. (2005) and Janssen and Manca (2007) presented a generation of applied semi-Markov processes which can apply to economic and financial issues. They showed that a semi-Markov process is the renewal process. In our model, we used definitions of semi-Markov processes found in Becker et al. (2000), and the assumptions of Rishel (1975) to characterize the discrete events of the system, such as machine breakdown and repairs.

The second extension considers the control problem in deterministic horizon with discount rate. This extension is neither similar to the Rishel formulism nor to the Boukas extension; Rishel considered the problem in a finite horizon, without a discount rate while Boukas considered it in an infinite horizon, with a discount rate. Consequently, the proposed model looks at a control of the system in order to meet either constant or varying demand.

3.9.2 Future Research

This new model can be applied to a large-scale system (job shop) with machine maintenance and setup problems, using the Just in Time (JIT) concept.

CONCLUSIONS GÉNÉRALES

Pratiquement, l'opération de systèmes de production réels est toujours active par rapport au temps; la dynamique de ces systèmes peut être influencée par une nature aléatoire telle que les fluctuations des demandes, des pannes et des réparations des machines. L'étude et la modélisation de cette dynamique constitue le cadre de ce mémoire.

Concernant la méthodologie, l'analyse mathématique permet d'aborder complémentairement cette dynamique. En réalité, l'analyse mathématique peut être considérée comme un modèle exact parce qu'elle peut décrire profondément la dépendance de la dynamique de systèmes de production réels qui fait intervenir des paramètres de contrôle.

L'objectif général de ce mémoire est d'une part de formuler un nouveau modèle mathématique, et d'autre part de l'appliquer aux systèmes de production réels afin de répondre aux deux questions que nous avons présentées à la section 1.3. Ce mémoire comporte donc deux parties. Dans la première partie, l'on s'est intéressé à la théorie qui consiste en la formulation du problème de commande optimale des systèmes à structure variable. La deuxième partie est l'application de ce nouveau modèle dans les systèmes manufacturiers.

Dans l'aspect théorique, nous avons étudié le problème de commande optimale stochastique par la généralisation du formalisme de Rishel dans les cas où le processus de sauts et ses taux de transition dépendraient du temps. Les conditions d'optimum obtenues dans ce cas satisfont le principe du minimum de Pontryagin, en ayant une extension de celles obtenues par Rishel (1975).

Sur le plan pratique, ce nouveau modèle est utilisé pour modéliser des systèmes de production flexible (SPF) dans certain cas.

(1) La modélisation du SPF à plusieurs machines identiques en parallèles, lors de taux d'actualisation égal zéro, a pour but de déterminer les taux de production en satisfaisant les fluctuations de demande par rapport au temps. Selon la politique du seuil critique, nous avons analysé la production cumulative sur un court-terme.

(2) La modélisation du SPF à deux-machine en tandem avec taux d'actualisation $\rho > 0$. L'objectif de cette modélisation est aussi de déterminer les taux de production en satisfaisant les fluctuations de demande par rapport au temps. Dans ce système de production, la politique de contrôle est plus complexe que dans le système à plusieurs machines identiques en parallèles.

Nous avons aussi appliqué la méthode numérique basée sur l'approche de Kushner et Dupuis (2001) pour résoudre des problèmes de commande considérée.

Dans le cadre d'une extension de ce nouveau modèle, comme future recherche, nous pensons qu'il est souhaitable :

1. de traiter le problème de commande optimale des systèmes de type jop-shop traitant plusieurs types de pièces;
2. de traiter le problème de commande optimale d'approvisionnement en matière première;
3. d'adapter la demande qui obéit à des distributions de probabilité quelconque;
4. d'introduire le problème de *setup* (mis en course), et celui de maintenance préventive dans le modèle;

ANNEXE I

RISHEL'S ASSUMPTIONS

In this appendix, we present the Rishel's assumptions in order to prove all three articles simultaneously. We make use of Rishel's assumptions from the result of Rishel (1975) in order to formulate our problems in finite horizon.

H.1. According to Rishel (1975), there are two events which can occur in a stochastic system. The two events represent two jump types as follows :

(1) given $\xi(t) = \alpha$, the probability that there is no event that the system remains in state α in the interval [0, T] is:

$$1-\int_{t}^{T} p_{\alpha\alpha}(s-t)\,ds, \tag{I.1}$$

Where the term $\int_{t}^{T} p_{\alpha\alpha}(s-t)\,ds$ is the probability that the system remains in state α in [0, T].

(2) given $\xi(t) = \alpha$, the probability that there is the first jump of $\xi(t)$ from α to β at time t in the interval [0, T] is:

$$\int_{t}^{T} p_{\alpha\beta}(s-t)\,ds. \tag{I.2}$$

H.2. We consider the process in the finite interval [0, T] and consider events that have $\xi(t)$ exactly $\mathcal{M}$ jumps, for all $t \leq T$. Assume that T is bounded, with probability one; thus, event $\xi(t)$ has more than a finite number of jumps in [0, T] and has probability of zero. Let $(\Omega, \mathcal{F}, P)$ be the probability space. Let $\eta_l(t,\omega)$ be the characteristic function of the set of $\omega \in \Omega$, in which $\xi(t,\omega)$ has exactly $\mathcal{M}$ jumps in $[0, T(\omega)]$ for $l \in \mathcal{M}$. Ω is sample space, ω is sample point.

Definition I.1. Let O(h) be the zero function if

$$\lim_{h\to 0}\frac{O(h)}{h} = 0. \tag{I.3}$$

ANNEXE II

APPENDIX OF THE ARTICLE 1

II.1. Proof of the theorem 2.3.1

To prove this Theorem 2.3.1, we consider two different parts; the first comprised of the construction of the integral equations; and the second - the building of the partial differential equations.

II.1.1 Integral equation terms

Using the assumption H.2, we have the value function in finite horizon as follows:

$$v^{\alpha}(t,x) = \min_{\substack{u(s,x)\in\mathcal{A}(s,\alpha)\\ t\le s\le T}} \sum_{l=0}^{\mathcal{M}} E[\eta_l(t)\int_t^T g\left(x(s),u(s)\right)ds \mid x(t)=x;\xi(t)=\alpha], \tag{II.1}$$

where $\eta_l(t) = 1$, it means that given at state α, the system can go to another state at any t.

Using equation (I.1) for the probability of no jump from α to α at s and for $s \le \tau < T$:

$$\min_{\substack{u(s,x)\in\mathcal{A}(s,\alpha)\\ t\le s\le T}} E\left[\eta_0(t)\int_t^T g\left(x(s),u(s)\right)ds \mid x(t)=x;\xi(t)=\alpha\right]$$

$$= \min_{\substack{u(s,x)\in\mathcal{A}(s,\alpha)\\ t\le s\le T}} \left[\int_t^T g(x^{\alpha}(s),u(s))ds - \left(\int_t^T p_{\alpha\alpha}(s-t)E\int_s^{\tau} g(x(z),u(z))dz \mid x(s)=x(s);\xi(s)=\alpha\right)ds\right]$$

$$= \min_{\substack{u(s,x)\in\mathcal{A}(s,\alpha)\\ t\le s\le T}} \left[\int_t^T g(x^{\alpha}(s),u(s))ds - \left(\int_t^T p_{\alpha\alpha}(s-t)v^{\alpha}\left(s,x^{\alpha}(s)\right)\right)ds\right]. \tag{II.2}$$

Using equation (I.2) for the probability of the other jumps from α to β, the terms in equation (II.1) can be written by induction, starting with:

$$\min_{\substack{u(s,x)\in\mathcal{A}(s,\alpha)\\ t\le s\le T}} \sum_{\beta\ne\alpha}\int_t^T p_{\alpha\beta}(s-t)E[\eta_{l-1}(s)\int_s^{\tau} g(x(z),u(z))dz \mid x(s)=x^{\alpha};\xi(s)=\beta]ds$$

$$= \min_{\substack{u(s,x)\in\mathcal{A}(s,\alpha)\\ t\le s\le T}} \sum_{\beta\ne\alpha}^{m}\int_t^T p_{\alpha\beta}(s-t)v^{\beta}\left(s,x^{\alpha}(s)\right)ds\,. \tag{II.3}$$

Combining equations (II.2) and (II.3) gives the following:

$$v^{\alpha}(t,x) =$$

$$\min_{\substack{u(s,x)\in\mathcal{A}(s,\alpha)\\ t\le s\le T}} \left\{ \int_t^T g(x(s),u(s))ds - \int_t^T p_{\alpha\alpha}(s-t)v^{\alpha}(s,x(s))ds + \sum_{\beta\neq\alpha}^{m}\int_t^T p_{\alpha\beta}(s-t)v^{\beta}(s,x^{\alpha}(s))ds \right\}$$

$$= \min_{\substack{u(s,x)\in\mathcal{A}(s,\alpha)\\ t\le s\le T}} \left\{ \int_t^T g\left(x(s),u(s)\right)ds + \sum_{\beta}^{m}\delta\left(t\right)\int_t^T p_{\alpha\beta}(s-t)v^{\beta}(s,x^{\alpha}(s))ds \right\}, \tag{II.4}$$

where $\delta(t)$ is the indicator function

$$\delta(t) = \begin{cases} -1 & \text{if} \quad \alpha = \beta \\ +1 & \text{if} \quad \alpha \neq \beta \end{cases}. \tag{II.5}$$

The expression (II.4) is the integral equation.

II.1.2 Partial differential equation terms

Let *h(s,x,u)* be a continuous function defined by:

$$h^{\alpha}(s,x,u) = g\left(x(s),u(s)\right) + \sum_{\beta}^{m}\delta(s)p_{\alpha\beta}(s-t)v^{\beta}(s,x^{\alpha}(s)). \tag{II.6}$$

Substituting (II.6) in (II.4), we obtain:

$$v^{\alpha}(t,x) = \min_{\substack{u(s,x)\in\mathcal{A}(s,\alpha)\\ t\le s\le T}} \int_t^T h^{\alpha}\left(s,x(s),u(s)\right)ds, \tag{II.7}$$

where the minimization is over all functions $u(s,x)$, $t \le s \le T$, such that $u(s,x) \in \mathcal{U}\left(s,x\right)$ in Definition 2.2.1, and such that $\xi(s)$ satisfies (2.4), (2.5) and (2.7). Note that $v^{\alpha}(T,x(T)) = 0$ is no Cauchy condition.

We may split the integral at any small increment value $\Delta t > 0$ to obtain:

$$v^{\alpha}(t,x) = \min_{\substack{u(s,x)\in\mathcal{A}(s,\alpha)\\ t\le s\le T}} \left\{ \int_t^{t+\Delta t} h^{\alpha}\left(s,x(s),u(s)\right)ds + \int_{t+\Delta t}^{T} h^{\alpha}\left(s,x(s),u(s)\right)ds \right\}$$

$$= \min_{\substack{u(s,x)\in\mathcal{A}(s,\alpha)\\ t\le s\le t+\Delta t}} \left\{ \int_t^{t+\Delta t} h^{\alpha}(.)ds + v^{\alpha}(t+\Delta t, x(t+\Delta t)) \right\}. \tag{II.8}$$

Since $h^{\alpha}(s,x(s),u(s))$ is a continuous function, using the Taylor series on its first variable, the integral in (II.8) is approximately $h^{\alpha}(s,x(s),u(s))\Delta t$ so that we can rewrite

$$v^{\alpha}(t,x)=\min_{u(t,x)\in\mathcal{A}(t,\alpha)}\left\{h^{\alpha}(t,x,u)\Delta t+v^{\alpha}(t+\Delta t,x(t+\Delta t))+O(\Delta t)\right\}, \tag{II.9}$$

where $O(\Delta t)$ denotes a collection of higher-order terms in Δt and defined in Definition I.1.

For small Δt, (II.9) becomes, approximately

$$0=\min_{u(t,x)\in\mathcal{A}(t,\alpha)}\left\{h^{\alpha}(t,x,u)+v_t^{\alpha}(t,x)+v_{\boldsymbol{x}}^{\alpha}(t,x)\dot{x}\right\}. \tag{II.10}$$

Combining (II.10) and (II.6), equation (2.8) applies.

Note that, the function $u(s,x)$ for $t\le s\le T$ in eq. (2.7) and (II.8) have become $u(t,x)$ $0\le t\le T$ because the dynamic programming approach had transferred the multi-stage problem in the time interval $[0, T]$ into a single stage problem at any time t.

II.2. Proof of the theorem 2.3.2
II.2.1. Proof of condition (ii)

Let $\hat{x}^{\alpha}(t)$ and $\hat{u}^{\alpha}(t,x)$ be the optimal states and the optimal control variables for $\alpha\in\mathcal{I}$ and $t\in D=[0,T]$, respectively.

Using assumptions in ***Theorem 2.3.1***, we have:

$$g(\hat{x}^{\alpha}(t),\hat{u}^{\alpha}(t,.))+v_t^{\alpha}\left(t,\hat{x}^{\alpha}(t)\right)+v_{\boldsymbol{x}}^{\alpha}f^{\alpha}(\hat{x}^{\alpha}(t),\hat{u}(t,.),t)+\sum_{\beta}\delta(t)\,p_{\alpha\beta}(t)v^{\beta}(t,\hat{x}^{\alpha}(t))=0. \tag{II.11}$$

Let $H[t,x^{\alpha}(t),u^{\alpha}(t,x)]$ be the Hamiltonian:

$$H[t,x^{\alpha}(t),u^{\alpha}(t,x)]=g(x^{\alpha}(t),u^{\alpha}(t,x))+v_x^{\alpha}f^{\alpha}(x^{\alpha}(t),u^{\alpha}(t,x),t). \tag{II.12}$$

Let $H[t,\hat{x}^{\alpha}(t),\hat{u}^{\alpha}(t,x)]$ be the Hamiltonian minimizing condition.

$$H[t,\hat{x}^{\alpha}(t),\hat{u}^{\alpha}(t,x)]=\min_{\substack{u(t,x)\in\mathcal{A}(t,\alpha)\\ \alpha\in\mathcal{I}}}H[t,x^{\alpha}(t),u^{\alpha}(t,x)]. \tag{II.13}$$

We can write the following:

$$H[t,\hat{x}^{\alpha},\hat{u}^{\alpha},t]\le H[t,\hat{x}^{\alpha},u^{\alpha}],\ \text{for all } u(t,x)\in\mathcal{U}(t,\boldsymbol{x}). \tag{II.14}$$

Introducing the following term $v_t^{\alpha}\left(t,\hat{x}^{\alpha}\right)+\sum_{\beta}\delta(t)\,p_{\alpha\beta}(t)v^{\beta}(t,\hat{x}^{\alpha})$ on both sides of equation (II.14) we obtain:

$$H[t,\hat{x}^{\alpha},u^{\alpha}]+v_{t}^{\alpha}\left(t,\hat{x}^{\alpha}\right)+\sum_{\beta}\delta\left(t\right)p_{\alpha\beta}(t)v^{\beta}(t,\hat{x}^{\alpha})\geq$$
$$H[t,\hat{x}^{\alpha},\hat{u}^{\alpha}]+v_{t}^{\alpha}\left(t,\hat{x}^{\alpha}\right)+\sum_{\beta}\delta\left(t\right)p_{\alpha\beta}(t)v^{\beta}(t,\hat{x}^{\alpha}). \quad \text{(II.15)}$$

We use properties of equation (II.11) to cancel the right side of equation (II.15) and as a result, we have:

$$H[t,\hat{x}^{\alpha},u^{\alpha}]+v_{t}^{\alpha}\left(t,\hat{x}^{\alpha}\right)+\sum_{\beta}\delta\left(t\right)p_{\alpha\beta}(t)v^{\beta}(t,\hat{x}^{\alpha})\geq 0. \quad \text{(II.16)}$$

Combining equation (II.16) with ***Theorem 2.3.1*** gives (ii) in ***Theorem 2.3.2***, where $\hat{u}^{\alpha}\left(t,x\right)=u^{*}\left(t,x\right)$.

II.2.2 Proof of condition (i)

Let $Q=[0,T]\times\Re^{m}$. Let $\hat{u}^{\alpha}\left(t,x\right)$ be optimal values of the control variable for $\alpha\in\mathcal{I}$, $\left(t,x\right)\in Q$. Consider $v^{\alpha}(t,\boldsymbol{x})$ and $w^{\alpha}(t,\boldsymbol{x})$, the conditional expected cost in equation (2.8) as corresponding to the optimal control $\hat{u}^{\alpha}(t,x)$ and $u^{\alpha}(t,x)$, respectively.

To prove this sufficient condition, we use the deterministic optimal problem. Then, let $V^{\alpha}\left(t,x\right)$, $W^{\alpha}\left(t,x\right)\in Q\times\mathcal{I}$, $\partial V^{\alpha}\left(t,x\right)/\partial t$, $\partial V^{\alpha}\left(t,x\right)/\partial t$ exist and be continuous for $\left(t,x\right)\in Q$. We define:

$$V^{\alpha}(t,x)=\int_{t}^{T}g(x^{\alpha}(s),\hat{u}(s))ds, \quad \text{(II.17)}$$
$$W^{\alpha}(t,x)=\int_{t}^{T}g(x^{\alpha}(s),u(s))ds. \quad \text{(II.18)}$$

Since (II.4), we can rewrite:

$$V^{\alpha}(t,x)=v^{\alpha}(t,x)-\sum_{\beta}\int_{t}^{T}p_{\alpha\beta}(s-t)\delta(s)v^{\beta}(s,x^{\alpha})ds, \quad \text{(II.19)}$$
$$W^{\alpha}(t,x)=w^{\alpha}(t,x)-\sum_{\beta}\int_{t}^{T}p_{\alpha\beta}(s-t)\delta(s)w^{\beta}(s,x^{\alpha})ds. \quad \text{(II.20)}$$

Because $v^{\alpha}(T,x(T))=0$, therefore $V^{\alpha}(T,x(T))=0$, which is similar to $W^{\alpha}(T,x(T))$.

The derivative of these two functions is given as follows:

$$\mathcal{D}^{+}V^{\alpha}(t,x)=\frac{dV^{\alpha}(t,x)}{dt}=v_{t}^{\alpha}(t,x)+v_{x}^{\alpha}(t,x).f^{\alpha}(t,x,\hat{u}^{\alpha})+\sum_{\beta}p_{\alpha\beta}(t)\delta(t)v^{\beta}(t,x^{\alpha}), \quad \text{(II.21)}$$

$$\mathcal{D}^{+}W^{\alpha}(t,x)=\frac{dW^{\alpha}(t,x)}{dt}=w_t^{\alpha}(t,x)+w_x^{\alpha}(t,x).f^{\alpha}(t,x,u^{\alpha})+\sum_{\beta}p_{\alpha\beta}(t)\delta(t)w^{\beta}(t,x^{\alpha}). \quad \text{(II.22)}$$

Using the infinitesimal operator $\mathcal{L}_{\hat{u}}$ for $v^{\alpha}(t,x)$ and $\mathcal{L}_{u}$ for $w^{\alpha}(t,x)$ at time *t*, we can write:

$$\mathcal{L}_{\hat{u}}v^{\alpha}(t,x)=v_t^{\alpha}(t,x)+v_x^{\alpha}(t,x).f^{\alpha}(t,x,\hat{u}^{\alpha})+\sum_{\beta}p_{\alpha\beta}(t)\delta(t)v^{\beta}(t,x^{\alpha}), \quad \text{(II.23)}$$

$$\mathcal{L}_{u}w^{\alpha}(t,x)=w_t^{\alpha}(t,x)+w_x^{\alpha}(t,x).f^{\alpha}(t,x,u^{\alpha})+\sum_{\beta}p_{\alpha\beta}(t)\delta(t)w^{\beta}(t,x^{\alpha}). \quad \text{(II.24)}$$

Introducing the term: $g(x(t), u(t))$ on both sides of equations (II.21) and (II.22), respectively, and then integrating them from *t* to *T*, we obtain the following:

$$\int_t^T g(x(s),u(s))ds - V^{\alpha}(t,x)=\int_t^T \mathcal{L}_{\hat{u}}v^{\alpha}(s,x)\,ds+\int_t^T g(x(s),\boldsymbol{u}(s))ds, \quad \text{(II.25)}$$

$$\int_t^T g(x(s),u(s))ds - W^{\alpha}(t,x)=\int_t^T \mathcal{L}_{u}w^{\alpha}(s,x)\,ds+\int_t^T g(x(s),u(s))ds. \quad \text{(II.26)}$$

Since (II.17) & (II.18), (II.25) & (II.26) become:

$$0\geq\int_t^T \mathcal{L}_{\hat{u}}v^{\alpha}(s,x)\,ds+\int_t^T g(x(s),u(s))ds, \quad \text{(II.27)}$$

$$0=\int_t^T \mathcal{L}_{u}w^{\alpha}(s,x)\,ds+\int_t^T g(x(s),u(s))ds. \quad \text{(II.28)}$$

Using (II.23) & (II.24), equations (II.27) and (II.28) imply:

$$v^{\alpha}(t,x)\leq w^{\alpha}(t,x). \quad \text{(II.29)}$$

Indeed, $\hat{u}^{\alpha}(t,\boldsymbol{x})$ is optimal control variable. As a result, equation (2.7) implies:

$$v^{\alpha}(t,x)=\min_{\substack{u(s,x)\in\mathcal{A}(s,\alpha)\\ t\leq s\leq T}} E_u\left[\int_t^T g(x(s),u(s))ds \mid x(t)=x,\xi(t)=\alpha\right].$$

This completes the proof.

II.3 Dtermination of $F_0(t)$, $F_1(t)$, $F_2(t)$ and $p_{\alpha\beta}(t)$

II.3.1 Determination of $F_0(t)$, $F_1(t)$, $F_2(t)$

We should recall that the probability distribution of a two-machine system in state $\xi(t)$ is given by

$$F_{\xi(t)}(t) = 1 - \prod_{i=1}^{2}\left(1 - \Pr\left\{\text{machine } i \text{ is functioning at } t\right\}\right). \tag{II.30}$$

Then, we have

$$F_0(t) = 1 - \left(1 - G(t)\right)\left(1 - G(t)\right) = G(t)\left[2 - G(t)\right], \tag{II.31}$$

$$F_1(t) = 1 - \left(1 - G(t)\right)\left(1 - F(t)\right) = F(t)\left[1 - G(t)\right] + G(t), \tag{II.32}$$

$$F_2(t) = 1 - \left(1 - F(t)\right)\left(1 - F(t)\right) = F(t)\left[2 - F(t)\right]. \tag{II.33}$$

Let $f_\alpha(t)$ be density functions of $F_\alpha(t)$ for $\alpha = 0,1,2$, respectively; we have:

$$f_\alpha(t) = \frac{dF_\alpha(t)}{dt}. \tag{II.34}$$

II.2.2. Determination of $p_{\alpha\beta}(t)$

Based on equation (8) in Becker et al. (2000), we have:

$$p_{\alpha\alpha}(t) = \left(1 - F_\beta(t)\right)\left(1 - F_\gamma(t)\right) f_\alpha(t). \tag{II.35}$$

The other terms, $p_{\alpha\beta}(t)$ for $\alpha \neq \beta$ are also determined in a similar manner. There are nine terms to calculte such as

$$\begin{gathered} p_{00}(t), p_{01}(t), p_{02}(t), \\ p_{10}(t), p_{11}(t), p_{12}(t), \\ p_{10}(t), p_{21}(t), p_{22}(t), \end{gathered}$$

ANNEXE III

APPENDIX OF THE ARTICLE 2

III.1 Proof of the theorem 3.1.1

By using the H.2 in the Appendix I, we consider a stochastic optimal control problem in finite horizon with discount rate to obtain:

$$v^{\alpha}(t,\boldsymbol{x}) = \min_{\substack{\boldsymbol{u}(s,\boldsymbol{x})\in\mathcal{A}(s,\alpha)\\ t\le s\le T}} \sum_{l=0}^{\mathcal{M}} E\left[\eta_l(t)\int_t^T e^{-\rho(s-t)}g(\boldsymbol{x}(s),\boldsymbol{u}(s))ds \mid X^{\xi(t)}(t)=X^{\alpha}\right]. \quad \text{(III.1)}$$

Using eq.(I.1) for the probability of no jump from α to α at s:

$$\min_{\substack{\boldsymbol{u}(s,\boldsymbol{x})\in\mathcal{A}(s,\alpha)\\ t\le s\le T}} E\left[\eta_0(t)\int_t^T e^{-\rho(s-t)}g(\boldsymbol{x}(s),\boldsymbol{u}(s))ds \mid X^{\xi(t)}(t)=X^{\alpha}\right]=$$

$$\min_{\substack{\boldsymbol{u}(s,\boldsymbol{x})\in\mathcal{A}(s,\alpha)\\ t\le s\le T}}\left[\int_t^T e^{-\rho(s-t)}g\left(\boldsymbol{x}(s),\boldsymbol{u}(s)\right)ds - \int_t^T\left(p_{\alpha\alpha}\left(s-t\right)E\int_s^{\tau} e^{-\rho(z-s)}g(.)dz \mid X^{\xi(s)}(s)=X^{\alpha}\right)ds\right]$$

$$= \min_{\substack{\boldsymbol{u}(s,\boldsymbol{x})\in\mathcal{A}(s,\alpha)\\ t\le s\le T}}\left[\int_t^T e^{-\rho(s-t)}g(\boldsymbol{x}(s),\boldsymbol{u}(s))ds - \int_t^T p_{\alpha\alpha}\left(s-t\right)v^{\alpha}\left(s,\boldsymbol{x}\right)ds\right]. \quad \text{(III.2)}$$

Using eq.(I.2) for the probability of the other jumps from α to β, the terms in eq. (III.1) can be written by induction starting with:

$$\min_{\substack{\boldsymbol{u}(s,\boldsymbol{x})\in\mathcal{A}(\alpha,s)\\ t\le s\le T}} \sum_{\beta\ne\alpha}^{\mathcal{M}}\int_t^T p_{\alpha\beta}(s-t)E[\eta_{l-1}(s)\int_s^{\tau} e^{-\rho(s-t)}g(.)dz \mid \boldsymbol{x}(s)=\boldsymbol{x}^{\alpha};\xi(s)=\beta]ds$$

$$= \min_{\substack{\boldsymbol{u}(s,\boldsymbol{x})\in\mathcal{A}(s,\alpha)\\ t\le s\le T}} \sum_{\beta\ne\alpha}^{m}\int_t^T p_{\alpha\beta}(s-t)v^{\beta}(s,\boldsymbol{x}^{\alpha}(s))ds. \quad \text{(III.3)}$$

Combining eq. (III.2) and eq. (III.3) gives as follows

$$v^{\alpha}(t,\boldsymbol{x}) = \min_{\substack{\boldsymbol{u}(s,\boldsymbol{x})\in\mathcal{A}(s,\alpha)\\ t\le s\le T}}\left\{\int_t^T e^{-\rho(s-t)}g(.)ds + \sum_{\beta}^{m}\int_t^T \delta(s)p_{\alpha\beta}(s-t)v^{\beta}\left(s,\boldsymbol{x}^{\alpha}(s)\right)ds\right\}, \quad \text{(III.4)}$$

where $\delta(s)$ is indicator function as in (II.5).

Let $h^{\alpha}(s,\boldsymbol{x},\boldsymbol{u})$ be a continuous function defined by

$$h^{\alpha}(s,\boldsymbol{x},\boldsymbol{u}) = g(\boldsymbol{x}(s),\boldsymbol{u}(s)) + \sum_{\beta}^{m} \delta(s) e^{-\rho(t-s)} p_{\alpha\beta}(s-t) v^{\beta}(s,\boldsymbol{x}^{\alpha}(s)). \qquad \text{(III.5)}$$

Substituting (III.5) in (III.4) to rewrite

$$v^{\alpha}(t,\boldsymbol{x}) = \min_{\substack{\boldsymbol{u}(s,\boldsymbol{x})\in\mathcal{A}(s,\alpha)\\ t\le s\le T}} \int_{t}^{T} e^{-\rho(s-t)} h^{\alpha}(s,\boldsymbol{x}(s),\boldsymbol{u}(s))\,ds, \qquad \text{(III.6)}$$

where the minimization is over all functions $\boldsymbol{u}(s,\boldsymbol{x})$, $t \le s \le T$, such that $\boldsymbol{u}(s,\boldsymbol{x}) \in \mathcal{A}(t, \alpha)$ in Definition 3.2.1 and such that $\boldsymbol{x}(s)$ and $\xi(s)$ satisfy (3.4), (3.5) and (3.7). We may split the integral at any value of small increment $\Delta t > 0$ to obtain:

$$\begin{aligned} v^{\alpha}(t,\boldsymbol{x}) &= \min_{\substack{\boldsymbol{u}(s,\boldsymbol{x})\in\mathcal{A}(s,\alpha)\\ t\le s\le T}} \left\{ \int_{t}^{t+\Delta t} e^{-\rho(s-t)} h^{\alpha}\left(s,\boldsymbol{x}(s),\boldsymbol{u}(s)\right) ds + \int_{t+\Delta t}^{T} e^{-\rho(s-t)} h^{\alpha}\left(s,\boldsymbol{x}(s),\boldsymbol{u}(s)\right) ds \right\} \\ &= \min_{\substack{\boldsymbol{u}(s,\boldsymbol{x})\in\mathcal{A}(s,\alpha)\\ t\le s\le t+\Delta t}} \left\{ \int_{t}^{t+\Delta t} e^{-\rho(s-t)} h^{\alpha}\left(s,\boldsymbol{x}(s),\boldsymbol{u}(s)\right) ds + e^{-\rho\Delta t} v^{\alpha}(t+\Delta t,\boldsymbol{x}(t+\Delta t)) \right\}. \end{aligned} \qquad \text{(III.7)}$$

Since $h^{\alpha}(s,\boldsymbol{x},\boldsymbol{u})$ is a continuous function, using the Taylor series on its first variable, the integral in (III.7) is approximately $h^{\alpha}(t,\boldsymbol{x},\boldsymbol{u})\Delta t$ so that we can rewrite

$$v^{\alpha}(t,\boldsymbol{x}) = \min_{\substack{\boldsymbol{u}(s,\boldsymbol{x})\in\mathcal{A}(\alpha,s)\\ t\le s\le T}} \left\{ h^{\alpha}(t,\boldsymbol{x},\boldsymbol{u})\Delta t + e^{-\rho\Delta t} v^{\alpha}(t+\Delta t,\boldsymbol{x}(t+\Delta t)) + O(\Delta t) \right\}. \qquad \text{(III.8)}$$

In (III.8) the term $O(\Delta t)$ is defined in Definition I.1.

Now using the following Taylor's expression:

$$e^{-y} = 1 - y + \frac{y^2}{2!} - \frac{y^3}{3!} + \ldots, \qquad \text{(III.9)}$$

and substituting (III.9) into (III.8), then for small Δt, (III.8) becomes, approximately

$$\rho v^{\alpha}(t,\boldsymbol{x}) = \min_{\substack{\boldsymbol{u}(s,\boldsymbol{x})\in\mathcal{A}(\alpha,s)\\ t\le s\le T}} \left\{ h^{\alpha}(t,\boldsymbol{x},\boldsymbol{u}) + v_t^{\alpha}(t,\boldsymbol{x}) + v_x^{\alpha}(t,\boldsymbol{x})\dot{\boldsymbol{x}} \right\}. \qquad \text{(III.10)}$$

Remark III.1. (a) The term $v^{\alpha}(t+\Delta t, \boldsymbol{x}(t+\Delta t))$ in (III.8) is the current value function at time $t+\Delta t$. (b) The result of (III.10) holds for present value function $v^{\alpha}(t,\boldsymbol{x})$ for deterministic horizon optimal control problems defined by Eq. (3.8). (c) The HJB equation (III.10) derived from Bellman principle of optimality obtains on time scales.

III.2 Proof of the theorem 3.3.2

This proof has been extended toward the optimal control problem with discount cost ($\rho > 0$) from results of Chapter II in Fleming and Soner (2006). Therefore, to prove that $v^{\alpha}(t,\boldsymbol{x})$ is convex in $(t, \boldsymbol{x})$, it suffices to show that $J^{\alpha}(t,\boldsymbol{x},\boldsymbol{u})$ is jointly convex in $(t, \boldsymbol{x}, \boldsymbol{u})$. Let

$$c^{+} = \begin{cases} c_1 & \text{if } \boldsymbol{x}^{+}(s) = x_1(s) \\ c_2^{+} & \text{if } \boldsymbol{x}^{+}(s) = x_2^{+}(s) \end{cases} \text{ and } c^{-} = c_2^{-}.$$

From (3.7), we get

$$\begin{aligned} J^{\alpha}(t,\boldsymbol{x},\boldsymbol{u}) &= E_{\boldsymbol{u}}\int_t^T e^{-\rho(s-t)} g(\boldsymbol{x}(s),\boldsymbol{u}(s))ds \\ &= E_{\boldsymbol{u}}\int_t^T e^{-\rho(s-t)}\left\{\left(c^{+}+c^{-}\right)\left(\boldsymbol{x}^{+}(s)+\boldsymbol{x}^{-}(s)\right)-c^{-}\boldsymbol{x}^{+}(s)-c^{+}\boldsymbol{x}^{-}(s)\right\}ds \\ &\le \left(c^{+}+c^{-}\right)E_{\boldsymbol{u}}\int_t^T e^{-\rho(s-t)}\left(\boldsymbol{x}^{+}(s)+\boldsymbol{x}^{-}(s)\right)ds \\ &\le \left(c^{+}+c^{-}\right)E_{\boldsymbol{u}}\int_t^T e^{-\rho(s-t)}\left|\boldsymbol{x}(s)\right|ds, \end{aligned}$$

where $$\boldsymbol{x}(s) = \boldsymbol{x}(t) + \int_t^s f^{\xi(t)}\left(\tau,\boldsymbol{x}(\tau),\boldsymbol{u}(\tau)\right)d\tau.$$

Note that: $$\left|\boldsymbol{x}(s)\right| \le \left|\boldsymbol{x}(t)\right| + \left|\int_t^s f^{\xi(t)}\left(\tau,\boldsymbol{x}(\tau),\boldsymbol{u}(\tau)\right)d\tau\right|.$$

Chosen $\boldsymbol{u}(t) = d(t)$ such that $\boldsymbol{x}(s) = \boldsymbol{x}(t)$ to obtain:

$$\begin{aligned} J^{\alpha}(t,\boldsymbol{x},\boldsymbol{u}) &\le \left(c^{+}+c^{-}\right)E_{\boldsymbol{u}}\int_t^T e^{-\rho(s-t)}\left|\boldsymbol{x}(t)\right|ds \\ &\le \left(c^{+}+c^{-}\right)E_{\boldsymbol{u}}\int_t^T \left|\boldsymbol{x}(t)\right|ds = C_1\left(T-t\right). \end{aligned} \tag{III.11}$$

The coefficient C_1 depends on initial value $\boldsymbol{x}(s)$ such that $C_1 = \left(c^{+}+c^{-}\right)\left|\boldsymbol{x}(t)\right|$.
From (III.11), it is shown that:

$$v^{\alpha}(t,\boldsymbol{x}) = \min_{\substack{\boldsymbol{u}(s,\boldsymbol{x})\in\mathcal{A}(s,\alpha) \\ t\le s\le T}} J^{\alpha}(t,\boldsymbol{x},\boldsymbol{u}) \le C_1\left(T-t\right)$$

is convex function for every $(t,\boldsymbol{x})$. This proves (i).

Now we proceed to prove that the value function $v^{\alpha}(t,\boldsymbol{x})$ satisfies the Lipschitz condition for every (*t*,***x***). First of all, the $v^{\alpha}(t,\boldsymbol{x})$ must satisfy the variable ***x***. For any control $\boldsymbol{u}(t,\boldsymbol{x})\in \mathcal{A}(t,\alpha)$, let :

$$\boldsymbol{x}(s)=\boldsymbol{x}(t)+\int_t^s f^{\xi(\tau)}\left(\tau,\boldsymbol{x}(\tau),\boldsymbol{u}(\tau)\right)d\tau,$$

$$\begin{aligned}\overline{\boldsymbol{x}}(s)&=\overline{\boldsymbol{x}}(t)+\int_t^s f^{\xi(\tau)}\left(\tau,\boldsymbol{x}(\tau),\boldsymbol{u}(\tau)\right)d\tau\\&=\overline{\boldsymbol{x}}(t)+\boldsymbol{x}(s)-\boldsymbol{x}(t).\end{aligned}$$

Since $\boldsymbol{x}(t)=\boldsymbol{x},\ \overline{\boldsymbol{x}}(t)=\overline{\boldsymbol{x}}$, we can rewrite:

$$\boldsymbol{x}(s)=\boldsymbol{x}+\int_t^s f^{\xi(t)}\left(\tau,\boldsymbol{x}(\tau),\boldsymbol{u}(\tau)\right)d\tau, \tag{III.12}$$

$$\overline{\boldsymbol{x}}(s)=\overline{\boldsymbol{x}}-\boldsymbol{x}+\boldsymbol{x}(s). \tag{III.13}$$

Then,

$$\begin{aligned}\left|J^{\alpha}(t,\boldsymbol{x},\boldsymbol{u})-J^{\alpha}(t,\overline{\boldsymbol{x}},\boldsymbol{u})\right|&=\left|\int_t^T e^{-\rho(s-t)}g(\boldsymbol{x}(s),\boldsymbol{u}(s))ds-\int_t^T e^{-\rho(s-t)}g(\overline{\boldsymbol{x}}(s),\boldsymbol{u}(s))ds\right|\\&\le\left(c^{+}+c^{-}\right)\int_t^T e^{-\rho(s-t)}\left|\boldsymbol{x}(s)-\overline{\boldsymbol{x}}(s)\right|ds\\&\le\left(c^{+}+c^{-}\right)\left|\boldsymbol{x}-\overline{\boldsymbol{x}}\right|\left(T-t\right).\end{aligned} \tag{III.14}$$

Since this is true for every $\boldsymbol{u}(t,\boldsymbol{x})\in \mathcal{A}(t,\alpha)$, it follows that:

$$\begin{aligned}\left|v^{\alpha}(t,\boldsymbol{x})-v^{\alpha}(t,\overline{\boldsymbol{x}})\right|&\le\left(c^{+}+c^{-}\right)|\,\overline{\boldsymbol{x}}-\boldsymbol{x}\,|\left(T-t\right)\\&=C_2\,|\,\overline{\boldsymbol{x}}-\boldsymbol{x}\,|\left(T-t\right).\end{aligned} \tag{III.15}$$

Next, let $t<\overline{t}<T$. Let $\boldsymbol{x}^{*}(.)$ be the optimal trajectory for initial date (*t*,***x***). Since expression (3.8), we can write:

$$\begin{aligned}v^{\alpha}\left(t,\boldsymbol{x}\right)&=\min_{\substack{\boldsymbol{u}(s,\boldsymbol{x})\in\mathcal{A}(s,\alpha)\\ t\le s\le T}}\left[\int_t^T e^{-\rho(s-t)}g(\boldsymbol{x}(s),\boldsymbol{u}(s))ds\right]\\&=\left[\int_t^{\overline{t}} e^{-\rho(s-t)}g(\boldsymbol{x}^{*}(s),\boldsymbol{u}^{*}(s))ds+e^{-\rho(\overline{t}-t)}v^{\alpha}\left(\overline{t},\boldsymbol{x}^{*}(\overline{t})\right)\right]\\&\le\left[\int_t^{\overline{t}} e^{-\rho(s-t)}g(\boldsymbol{x}^{*}(s),\boldsymbol{u}^{*}(s))ds+v^{\alpha}\left(\overline{t},\boldsymbol{x}^{*}(\overline{t})\right)\right].\end{aligned} \tag{III.16}$$

Using (III.11), the convex function of $J^{\alpha}\left(t,\boldsymbol{x}^{*},\boldsymbol{u}^{*}\right)$ becomes:

$$J^{\alpha}(t,\boldsymbol{x}^{*},\boldsymbol{u}^{*})=E_{\boldsymbol{u}}\int_{t}^{\bar{t}}e^{-\rho(s-t)}g\left(\boldsymbol{x}^{*}(s),\boldsymbol{u}^{*}(s)\right)ds\le C_{3}\left(\bar{t}-t\right), \tag{III.17}$$

where $C_{3}=\left(c^{+}+c^{-}\right)|\,\boldsymbol{x}^{*}\,|$, $\boldsymbol{x}^{*}$ is initial value of $\boldsymbol{x}^{*}(s)$ for $t\le s\le\bar{t}$.

Substituting (III.17) into (III.16) to obtain:

$$\left|v^{\alpha}\left(t,\boldsymbol{x}\right)-v^{\alpha}\left(\bar{t},\boldsymbol{x}\right)\right|\le C_{3}\left|\bar{t}-t\right|+\left|v^{\alpha}\left(\bar{t},\boldsymbol{x}^{*}(\bar{t})\right)-v^{\alpha}\left(\bar{t},\boldsymbol{x}\right)\right|. \tag{III.18}$$

Using (III.15), the following term is given:

$$\begin{aligned}\left|v^{\alpha}\left(\bar{t},\boldsymbol{x}^{*}(\bar{t})\right)-v^{\alpha}\left(\bar{t},\boldsymbol{x}\right)\right| &\le C_{2}\,|\,\boldsymbol{x}^{*}(\bar{t})-\boldsymbol{x}\,|\left(T-\bar{t}\right)\\ &\le C_{2}\,|\,\boldsymbol{x}^{*}(\bar{t})-\boldsymbol{x}\,|\left(T-t\right)\\ &= C_{2}\left(|\int_{t}^{\bar{t}}f^{\alpha}\left(s,\boldsymbol{x}^{*},\boldsymbol{u}^{*}\right)|\right)\left(T-t\right)\\ &\le C_{2}.C_{4}\,|\,\bar{t}-t\,|\left(T-t\right),\end{aligned} \tag{III.19}$$

where $C_{4}=\max\left(|\,f^{\alpha}\left(s,\boldsymbol{x}^{*},\boldsymbol{u}^{*}\right)|\right)$, that means choosing $\boldsymbol{u}(t) = r$ if $(r - d) > d$. If not, $\boldsymbol{u}(t) = 0$.

Therefore, from (III.18) & (III.19) we have:

$$\left|v^{\alpha}\left(t,\boldsymbol{x}\right)-v^{\alpha}\left(\bar{t},\boldsymbol{x}\right)\right|\le\left(C_{3}+C_{2}.C_{4}\left(T-t\right)\right)\left|\bar{t}-t\right|. \tag{III.20}$$

Note that inequalities (III.15) and (III.20) are in Lipchitz form for $v^{\alpha}\left(t,\boldsymbol{x}\right)$. In the sequel, let us consider $\left(t,\boldsymbol{x}\right),\left(\bar{t},\bar{\boldsymbol{x}}\right)\in[0,T]\times[0,B]\times\Re^{1}$, for any control $\boldsymbol{u}(t,\boldsymbol{x})$:

$$\begin{aligned}v^{\alpha}(t,\boldsymbol{x})-v^{\alpha}(\bar{t},\bar{\boldsymbol{x}}) &= \min_{\substack{\boldsymbol{u}(s,\boldsymbol{x})\in\mathcal{A}(s,\alpha)\\ t\le s\le T}}E_{\boldsymbol{u}}\left\{\int_{t}^{T}e^{-\rho(s-t)}g(\boldsymbol{x}(s),\boldsymbol{u}(s))ds-\int_{\bar{t}}^{T}e^{-\rho(s-t)}g(\bar{\boldsymbol{x}}(s),\boldsymbol{u}(s))ds\right\}\\ &= \min_{\substack{\boldsymbol{u}(s,\boldsymbol{x})\in\mathcal{A}(s,\alpha)\\ t\le s\le T}}E_{\boldsymbol{u}}\left\{\int_{t}^{T}e^{-\rho(s-t)}\left(g(\boldsymbol{x}(.),\boldsymbol{u}(.))-g(\bar{\boldsymbol{x}}(.),\boldsymbol{u}(.))\right)ds+\right.\\ &\qquad\left.\int_{t}^{T}e^{-\rho(s-t)}g(\bar{\boldsymbol{x}}(.),\boldsymbol{u}(.))ds-\int_{\bar{t}}^{T}e^{-\rho(s-t)}g(\bar{\boldsymbol{x}}(.),\boldsymbol{u}(.))ds\right\}.\end{aligned} \tag{III.21}$$

By combining (III.14),(III.15) and (III.20), expression (III.21) becomes:

$$\left|v^{\alpha}(t,\boldsymbol{x})-v^{\alpha}(\bar{t},\bar{\boldsymbol{x}})\right|\leq\left|v^{\alpha}(t,\boldsymbol{x})-v^{\alpha}(t,\bar{\boldsymbol{x}})\right|+\left|v^{\alpha}(t,\bar{\boldsymbol{x}})-v^{\alpha}(\bar{t},\bar{\boldsymbol{x}})\right|$$
$$\leq C_M\left(|\,\bar{t}-t\,|+|\,\bar{\boldsymbol{x}}-\boldsymbol{x}\,|\right),$$

where: $C_M = C_3 + C_2\left(1+C_4\right)\left(T-t\right)$.

Therefore the Lipchitz condition is satisfied for every $(t, \boldsymbol{x})$. This also proves (ii).

III.3 Proof of the theorem 3.3.3

To prove this theorem, we recall the definition 3.2.1. Let $(\Omega, \mathcal{F}, P)$ be a probability space for $t \le s \le T$, and $(t, \boldsymbol{x})$ be initial date. Let $\Phi \in \left[0,T\right]\times\Re^2$, the following assumptions hold:

(a) $\Phi\left(t,\boldsymbol{x}\right)=e^{-\rho t}\varphi(\boldsymbol{x})$ for $\varphi\in\Re^2$ and $0<T<\infty$;

(b) $E_x\int_0^{\infty} e^{-\rho t} g(x(t),u(t))dt<\infty$..

(i) Let $\mathcal{A}(t,\alpha)$ be any admissible feedback control with the initial vector $\boldsymbol{x}(t) = \boldsymbol{x}$. Since $\boldsymbol{u}(t,\boldsymbol{x})\in\mathcal{A}(t,\alpha)$, equation (3.9) becomes:

$$v_t^{\alpha}(t,\boldsymbol{x})+v_{\boldsymbol{x}}^{\alpha}(t,\boldsymbol{x})f^{\alpha}\left(t,\boldsymbol{x},\boldsymbol{u}\right)-\rho v^{\alpha}(t,\boldsymbol{x})+\sum_{\beta}\delta(t)p_{\alpha\beta}(t)v^{\beta}(t,\boldsymbol{x})+g\left(\boldsymbol{x}(t),\boldsymbol{u}(t)\right)\geq 0. \quad \text{(III.22)}$$

Put $g(t,\boldsymbol{x},\boldsymbol{u})$ on right side to obtain:

$$v_t^{\alpha}(t,\boldsymbol{x})+v_{\boldsymbol{x}}^{\alpha}(t,\boldsymbol{x})f^{\alpha}\left(t,\boldsymbol{x},\boldsymbol{u}\right)-\rho v^{\alpha}(t,\boldsymbol{x})+\sum_{\beta}\delta(t)p_{\alpha\beta}(t)v^{\beta}(t,\boldsymbol{x})\geq -g\left(\boldsymbol{x}(t),\boldsymbol{u}(t)\right). \quad \text{(III.23)}$$

Since Assumption (a), by applying Dynkin's formula to $v^{\alpha}\left(t,x\right)$ in (III.23) (see chapter III in Fleming and Soner (2006)):

$$e^{-\rho T}E\left[v^{\alpha}\left(T,\boldsymbol{x}(T)\right)\right]-v^{\alpha}\left(t,\boldsymbol{x}\right)\geq-\int_t^T e^{-\rho(s-t)}g\left(\boldsymbol{x}(s),\boldsymbol{u}(s)\right)ds. \quad \text{(III.24)}$$

Here, there are two ways to get the results: by using the boundary condition $v^{\alpha}\left(T,x(T)\right)=0$ which is our finite time problem, or by sending a limit $\lim_{T\to\infty}e^{-\rho T}E\left[v^{\alpha}\left(T,x(T)\right)\right]$ that tends to zero for the infinite time problem. Then, the equation (III.24) applies:

$$v^{\alpha}\left(t,\boldsymbol{x}\right)\leq J^{\alpha}(s,\boldsymbol{x},\boldsymbol{u}). \quad \text{(III.25)}$$

(ii) In the proof of (i), equality now replaces inequality in (III.25).

These complete the Proof of Theorem 3.2.

III.4 Proof of the theorem 3.5.1

Let $y(t) \in \Re^1$ be the state variable of (3.25). By using the dynamic programming, the optimal control in (3.25) becomes the following linear program:

$$\min_{u(.)\in A(t,\alpha)} \left[\frac{\partial v^{\alpha}(t,y)}{\partial y} \frac{\partial y}{\partial t} \right]. \tag{III.26}$$

Since the relation $y(t) = x_1(t) + x_2(t)$, its derivative is given by:

$$\frac{\partial y(t)}{\partial t} = \frac{\partial x_1(t)}{\partial t} + \frac{\partial x_2(t)}{\partial t} = u_1(t) - d(t). \tag{III.27}$$

Substituting (III.27) into (III.26) we obtain:

$$\min_{u_1(t,y)\in A(t,\alpha)} \left[\frac{\partial v^{\alpha}(t,y)}{\partial y}(u_1 - d) \right]. \tag{III.28}$$

Because $y(t) \in \Re^1$ the value function $v^{\alpha}(.)$ is convex in y. As results, the optimal control problem (III.28) is established by Akella and Kumar (1986), in which the optimal production rate is $u_1^*(t)$ with hedging point $y^*(t)$. Both two machines are simultaneously controlled, and their optimal policies are validated at time t together. Therefore, combining the optimal solutions in (3.23) and (III.28) with the relationship $y(t) = x_1(t) + x_2(t)$ we can obtain the hedging point $z_1^*(t) = y^*(t) - z_1^*(t)$ using $u_1^*(t)$.

This completes the Proof of Proposition 4.1.

III.5 Determination of $P_{\gamma}(t)$ & $p_{\alpha\beta}(t)$
III.5.1 Determination of probability distribution $P_{\gamma}(t)$

Since the probability distribution of two-machine system in state $\xi(t)$ (see Ross (2003)) is given by

$$P_{\xi(t)}(t) = \prod_{i=1}^{2} \Pr\{\text{machine } i \text{ is functioning at } t\}. \tag{III.29}$$

Then, we have

$$P_0(t) = G_1(t).G_2(t), \tag{III.30}$$
$$P_1(t) = F_1(t).G_2(t), \tag{III.31}$$
$$P_2(t) = G_1(t).F_2(t), \tag{III.32}$$
$$P_3(t) = F_1(t).F_2(t). \tag{III.33}$$

Let $p_\gamma(t)$, $\gamma = 0, 1, 2, 3$ be the density functions of $P_\gamma(t)$ as follows:

$$p_\gamma(t) = \frac{dP_\gamma(t)}{dt}, \gamma = 0,1,2,3. \tag{III.34}$$

III.5.2. Determination of derivative of transition probability $p_{\alpha\beta}(t)$

Based on the eq. 8 in Becker et al. (2000), we have:

$$p_{00}(t) = (1 - P_1(t))(1 - P_2(t))(1 - P_3(t))\, p_0(t),$$
$$p_{11}(t) = (1 - P_0(t))(1 - P_2(t))(1 - P_3(t))\, p_1(t),$$
$$p_{22}(t) = (1 - P_0(t))(1 - P_1(t))(1 - P_3(t))\, p_2(t),$$
$$p_{33}(t) = (1 - P_0(t))(1 - P_1(t))(1 - P_2(t))\, p_3(t).$$

It is similar to determiner other terms $p_{\alpha\beta}(s)$, $\alpha \neq \beta$.

III.5 Numerical approach

As shown above, we use the Kushner and Dupuis's approach to construct an approximating discrete-time discrete state Markov chain to the original continuous stochastic control problem. This is because semi-Markov processes and continuous-time Markov chain have the same discrete event dynamic system (see Glynn (1989)). The idea is to approximate the value function $v^\alpha(t,\boldsymbol{x})$ by $(v^\alpha(t,\boldsymbol{x}))_\Delta$ via a finite difference method. Thus the first-order partial derivatives of the value function are approximated by the expressions (36) and (37) above.
Then the discrete version of HJB equation is given by:

$$v_\Delta^\alpha(t,\boldsymbol{x}) = \min_{\substack{0 \le u_1 \le r_1 \\ 0 \le u_2 \le r_2}} \left\{ \frac{1}{\rho + p_{\alpha\alpha}(t) + \frac{1}{\Delta t} + \frac{|u_2 - d|}{\Delta x_2} + \frac{|u_1 - u_2|}{\Delta x_1}} \times \left(\frac{|u_2 - d|}{\Delta x_2} \begin{bmatrix} v_\Delta^\alpha(t, x_1, x_2 + \Delta x_2) I_{(u_2 - d) \ge 0} + \\ v_\Delta^\alpha(t, x_1, x_2 - \Delta x_2) I_{(u_2 - d) < 0} \end{bmatrix} + \right.\right.$$

$$\frac{|u_1-u_2|}{\Delta x_1}\begin{bmatrix} v_\Delta^\alpha\left(t,x_1+\Delta x_1,x_2\right)I_{(u_1-u_2)\geq 0}+ \\ v_\Delta^\alpha\left(t,x_1-\Delta x_1,x_2\right)I_{(u_1-u_2)<0}\end{bmatrix}+\frac{v_\Delta^\alpha\left(t+\Delta t,x_1,x_2\right)}{\Delta t}+\sum_{\beta\neq\alpha}p_{\alpha\beta}(t)v_\Delta^\beta\left(t,x_1,x_2\right)\Bigg)\Bigg\}. \quad \text{(III.35)}$$

III.5.1. Boundary conditions

Consider the boundary conditions according the time t and the field $\boldsymbol{x}(t)$, we have:

$$v^\alpha(t,\boldsymbol{x})=\min_{\boldsymbol{u}(.)\in\mathcal{A}(.)}\int_t^T e^{-\rho(s-t)}g\left(\boldsymbol{x}(s),\boldsymbol{u}(s)\right)ds, \quad \text{(III.36)}$$

$$v_t^\alpha(t,\boldsymbol{x})=\min_{\boldsymbol{u}(.)\in\mathcal{A}(.)}\frac{d}{dt}\int_t^T e^{-\rho(s-t)}g(\boldsymbol{x}(s),\boldsymbol{u}(s))ds=\min_{\boldsymbol{u}(.)\in\mathcal{A}(.)}\left\{g\left(\boldsymbol{x}(t),\boldsymbol{u}(t)\right)\right\}, \quad \text{(III.37)}$$

$$v_x^\alpha(t,\boldsymbol{x})=\int_t^T e^{-\rho(s-t)}\frac{d}{dx}g\left(\boldsymbol{x}(s),\boldsymbol{u}(s)\right)ds=\begin{cases}\frac{c_2^+}{\rho}(1-e^{-\rho(T-t)}) & \text{if } \boldsymbol{x}(t)=x_2(t)\geq 0\\ \frac{c_2^-}{\rho}(1-e^{-\rho(T-t)}) & \text{if } \boldsymbol{x}(t)=x_2(t)<0.\\ \frac{c_2}{\rho}(1-e^{-\rho(T-t)}) & \text{if } \boldsymbol{x}(t)=x_1(t)\end{cases} \quad \text{(III.38)}$$

We use the computation domain as $\{t_0\leq t\leq T\},\{0\leq x_1\leq a\},\{-b\leq x_2\leq b\}$ for positive values of a and b**.** Therefore, we can use the following constraints as the boundary conditions:

$$\begin{cases} v^\alpha(T+\Delta t,x)=v^\alpha(T,x)-\Delta t.g(.)\\ v^\alpha(t,b+\Delta x_2)=v^\alpha(t,b)+\Delta x_2.\frac{c_2^+}{\rho}\left(1-e^{-\rho(T-t)}\right)\\ v^\alpha(t,-b-\Delta x_2)=v^\alpha(t,-b)+\Delta x_2.\frac{c_2^-}{\rho}\left(1-e^{-\rho(T-t)}\right).\\ v^\alpha(t,0-\Delta x_1)=v^\alpha(t,0)+\Delta x_1.\frac{c_1}{\rho}\left(1-e^{-\rho(T-t)}\right)\\ v^\alpha(t,a+\Delta x_1)=v^\alpha(t,a)+\Delta x_1.\frac{c_1}{\rho}\left(1-e^{-\rho(T-t)}\right)\end{cases} \quad \text{(III.39)}$$

ANNEXE IV

MÉTHODE NUMÉRIQUE

La méthode consiste à approximer les dérivées partielles du premier ordre de la fonction valeur $v_x^\alpha(t,x)$ dans les équations d'HJB par l'expression suivante :

$$v_x^\alpha(t,x) = \begin{cases} \dfrac{v^\alpha(t,x+\Delta t) - v^\alpha(t,x)}{\Delta t} & \text{si} \quad \dot{x} \geq 0 \\ \dfrac{v^\alpha(t,x) - v^\alpha(t,x-\Delta t)}{\Delta t} & \text{si} \quad \dot{x} < 0 \end{cases}. \tag{IV.1}$$

Pour le terme $v_t^\alpha(t,x)$, on a :

$$v_t^\alpha(t,x) = \frac{v^\alpha(t+\Delta t,x) - v^\alpha(t,x)}{\Delta t}. \tag{IV.2}$$

Pour obtenir une approximation de cette solution on utilise la technique d'itération de la commande avec son algorithme présenté de la façon suivante :

1. Initiation : Choisir $\delta \in \Re^+$, poser $k := 1$ et $\left(v_\alpha^\Delta\right)^0(t,x) = 0$, $\forall t \in G_t^h, \forall x \in G_x^h$, $\forall \alpha \in \mathcal{I}$. Supposer une commande stationnaire $u^k \in G_u^h$.

2. Évaluation de la commande : Avec la commande stationnaire $u^k \in G_u^h$ et $0 \leq u \leq u_{\max}$, calculer la fonction valeur correspondante $v_h^k(.)$ en utilisant : $\left(v_\alpha^\Delta\right)^{k-1}(t,x) = \left(v_\alpha^\Delta\right)^k(t,x)$, $\forall \alpha \in \mathcal{I}$, $\forall t \in G_t^h, et$ $\forall x \in G_x^t$

3. Raffinement de la commande : Obtenir une nouvelle commande stationnaire $u^{k+1} \in G_u^h$ et $0 \leq u \leq u_{\max}$, $\forall \alpha \in \mathcal{I}$, $\forall t \in G_t^h, et$ $\forall x \in G_x^t$.

4. Test de h convergence:

$$C_{\min} = \min_{x \in G_x^h, t \in G_t^h} \left\{ \left(v_\alpha^\Delta\right)^k(t,x) - \left(v_\alpha^\Delta\right)^{k-1}(t,x) \right\}, C_{\max} = \max_{x \in G_x^h, t \in G_t^h} \left\{ \left(v_\alpha^\Delta\right)^k(t,x) - \left(v_\alpha^\Delta\right)^{k-1}(t,x) \right\}.$$

Si $|C_{\min} - C_{\max}| \leq \delta$ pour $\rho = 0$ ou $\rho(1-\rho)^{-1}|C_{\min} - C_{\max}| \leq \delta$ pour $\rho > 0$, alors arrêter l'exécution et poser $u^* = u^k$, sinon incrémenter k (c-t-d $k = k + 1$) et retourner à l'étape 2.

Pour des définitions détaillées de cette méthode, nous référons le lecteur à Kushner et Dupuis (2001) et Boukas (1995).

ANNEXE V

SYSTÈME DE PRODUCTION À UNE MACHINE TRAITANT DEUX TYPES DE PIÈCES

Les résultats de cette annexe ont été présentés dans la " IASTED International Conference on Signal and Image Processing " qui a eu lieu du 17 au 19 August 2009 à Honolulu-Hawaii (États-Unis) sous le titre «Real-time control of stochastic manufacturing systems with semi-Markov jump».

L'objet de cette annexe est de présenter la loi de commande des taux de production pour un système à une seule machine traitant deux types de pièces. Le problème de commande de production du système est considéré sur un horizon fini avec le taux d'actualisation $\rho > 0$. Comme dans les chapitres 2, et 3, la solution des conditions d'optimum stochastiques du problème de planification de la production pour des équations d'HJB est d'une résolution numérique. En utilisant les résultats dans le chapitre 3, nous avons les équations d'HJB pour un système à une machine traitant deux types de pièces comme suit :

$$\begin{cases} \text{à l'état 0:} \\ \rho v^0(.) = \left\{ g(x_1, x_2) + v_t^0(t) - d_1(t).v_{x_1}^0(.) - d_2(t).v_{x_2}^0(.) - p_{00}(t)v^0(.) + p_{01}(t)v^1(.) \right\}, \end{cases} \tag{V.1}$$

$$\begin{cases} \text{à l'état 1:} \\ \rho v^1(.) = \min\limits_{\substack{0 \le u_1 \le r_1 \\ 0 \le u_2 \le r_2}} \left\{ g(.) + v_t^1(.) + \sum_{k=1}^{2} \left(u_k(t) - d_k(t) \right) v_{x_k}^1(.) - p_{11}(t)v^1(.) + p_{10}(t)v^0(.) \right\}. \end{cases} \tag{V.2}$$

L'algorithme d'itération de la commande, présenté à l'annexe IV, a été programmé en utilisant Matlab avec des données de simulations fixées. Ces données sont présentées ci-dessous :

- Taux d'actualisation et de demande

$$\rho = 0.1. \tag{V.3}$$

t	0-200	200-400	400-600	600-800	800-1000
d_1, d_2	0.10	0.15	0.135	0.170	0.125

- Variables de commande

$$U_{1\max} = r_1 = 0.25,\ U_{2\max} = r_2 = 0.225. \tag{V.4}$$

$$0 \le u_1 \le r_1 ; 0 \le u_2 \le r_2 . \tag{V.5}$$

- Constantes du coût instantané

$$c_{1,2}^{+} = 1, c_{1,2}^{-} = 3. \quad \text{(V.6)}$$

- Variable d'états et de temps

$$x_{1,2} \in [-25, 25], t \in (0, 1000]. \quad \text{(V.7)}$$

En utilisant les mêmes données qu'à la section 3.8 pour machine M_2 et le cas A dans la table 3.1, nous avons résolu numériquement les équations d'HJB (V.1) et (V.2). Les figures V.1-V.3 représentent les taux de production pour produit 1 et 2. Les figures V.4 à V.5 représentent les fonctions valeurs dans le mode où la machine est opérationnelle.

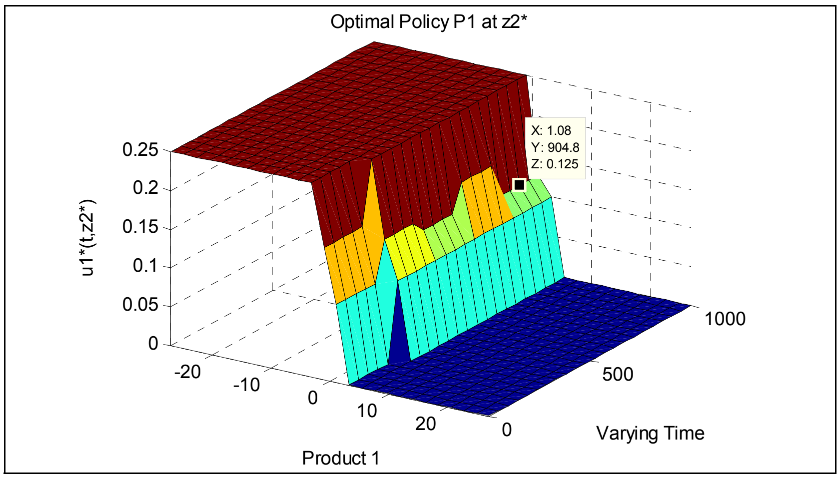


Figure V.1 Taux de production pour produit 1 versus x_1 et t.

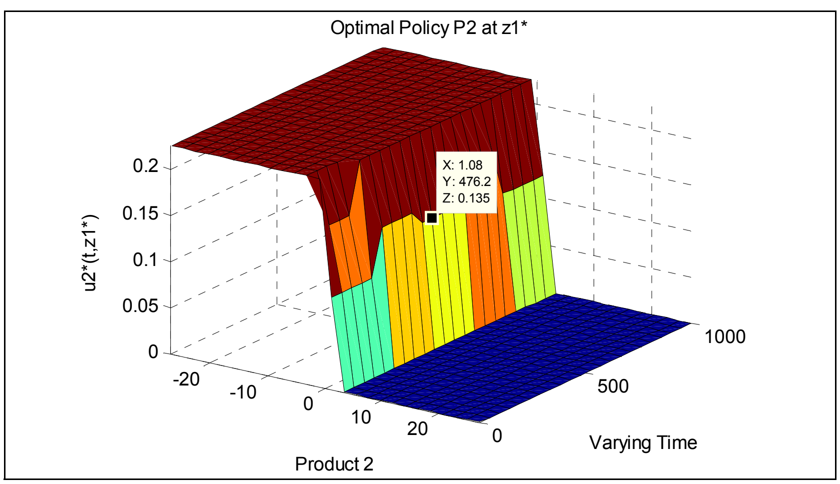


Figure V.2 Taux de production pour produit 2 versus x_2 et t.

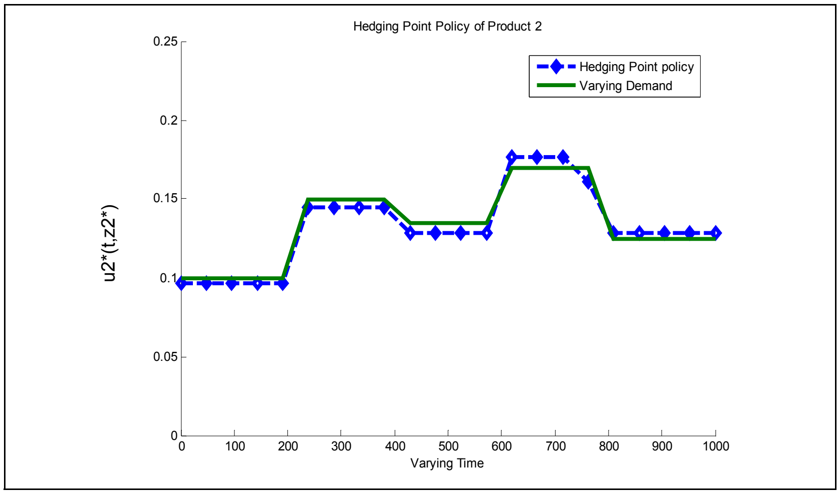


Figure V.3 Taux de production aux seuils critiques versus *t*.

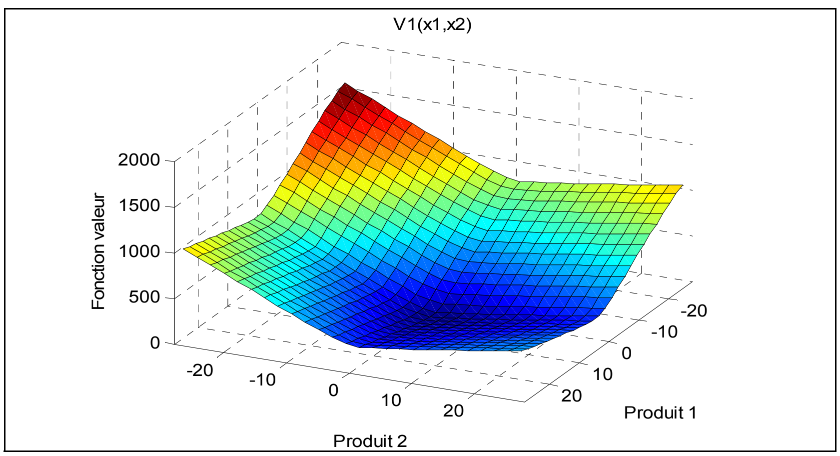


Figure V.4 Allure de la fonction valeur dans le mode 1.

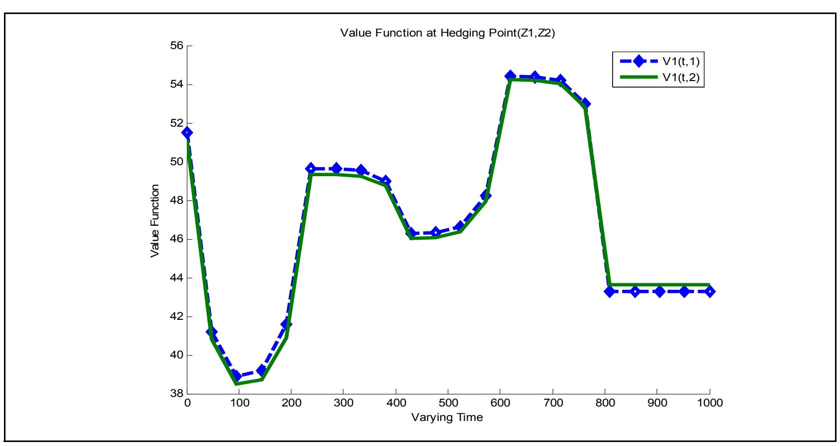


Figure V.5 Fonction valeur pour types de pièce P_1 et P_2 versus *t*.

La figure V.4 présente la fonction convexe de la fonction valeur versus les variables d'états du système x_1 et x_2 (stock) dans le mode 1 de la machine à l'instant t = 428.6 unités de temps. Cet instant est choisi de façon arbitraire dans l'intervalle (0, 1000]. La figure V.5 présente la fonction valeur aux seuils critiques $\left(z_{1,2}^{*}\right)$ versus le temps t. La fonction valeur dans la Figure V.5 fluctue parce que les demandes sont des variables aléatoires dans l'intervalle (0, 1000].

ANNEXE VI

MODÈLE MATKOVIEN VS SEMI-MAKOVIEN ET MÉTHODES D'ÉTUDE

VI.1 Modèle Markovien

Le but de cette section est de présenter le modèle de commande optimale stochastique de type Markovien homogène. Ce modèle est construit par Rishel (1975) et est formulé selon le processus Markovien homogène dont les taux de transition sont constants. Rishel a considéré un système stochastique décrit par l'équation d'état en temps continu :

$$\frac{d\boldsymbol{x}(t)}{dt} = f^{\xi(t)}\left(t, \boldsymbol{x}, \boldsymbol{u}\right),\ \boldsymbol{x}(0) = \boldsymbol{x}^{o}, \qquad \text{(VI.1)}$$

où $\{\xi(t), t \geq 0\}$ est le **processus Markovien en temps continu** et **à état fini**, $\boldsymbol{x}(t) \in \Re^n$ est le vecteur d'état continu du système, $\boldsymbol{u}(t) \in \Re^m$ est le vecteur de commande, $\boldsymbol{x}^o$ est le vecteur de la valeur initiale. Soit $\mathcal{I}$ l'espace d'état est discret et fini. À l'instant *t*, chaque mode $\alpha \in \mathcal{I} = \{0,1,...,m\}$ nous avons $f^{\alpha}\left(t, \boldsymbol{x}(t), \boldsymbol{u}(t)\right) : \Re^n \times \Re^m \times [0,T] \rightarrow \Re^m$ une équation bornée, continue par rapport à $\boldsymbol{x}(t)$ et $\boldsymbol{u}(t)$.
Le processus Markovien $\{\xi(t), t \geq 0\}$ est caractérisé par la matrice des taux de transition *Q*(.) = $\left[\lambda_{\alpha\beta}\right]$ telle que $\forall \alpha, \beta \in \mathcal{I}$

$$\lambda_{\alpha\beta} \geq 0, \lambda_{\alpha\alpha} = -\sum_{\beta \neq \alpha} \lambda_{\alpha\beta}, \sum_{\beta} \lambda_{\alpha\beta} = 0 \qquad \text{(VI.2)}$$

Les probabilités de transition du mode α au mode β de chaque machine sont décrites par :

$$P\{\xi(t+\Delta t) = \beta \mid \xi(t) = \alpha\} = \begin{cases} \lambda_{\alpha\beta}\delta t + O(\Delta t), & si\ \alpha \neq \beta \\ 1 + \lambda_{\alpha\beta}\delta t + O(\Delta t) & si\ \alpha = \beta \end{cases}. \qquad \text{(VI.3)}$$

Soit $\mathcal{A}(t,\alpha)$ est l'ensemble des lois de commandes admissibles et défini par :

$$\mathcal{A}(t,\alpha) = \left\{\boldsymbol{u}(t) \in \Re^m : 0 \leq \boldsymbol{u}(t) \leq \boldsymbol{u}^{\max}, \xi(t) = \alpha, t \geq 0\right\},$$

où $\boldsymbol{u}^{\max}$ est le taux de commande maximal.

Pour appliquer le SPF, nous avons utilisé le développement du modèle de Rishel dans l'ouvrage de Gershwin (2002) page 289. Soit $g(\boldsymbol{x}(t),\boldsymbol{u}(t))$ la fonction coût instantanée. Elle est une fonction continue qui admet des dérivées partielles par rapport à $\boldsymbol{x}$. Le problème d'optimisation consiste à choisir une loi de commande $\boldsymbol{u} \in \mathcal{A}(.)$ qui minimise la fonction coût comme suit :

$$v^{\alpha}\left(t,\boldsymbol{x}\right)=\min_{\substack{\boldsymbol{u}(s)\in\mathcal{A}(s,\alpha)\\ t\le s\le T}} E\left\{\int_t^T g\left(\boldsymbol{x}(s),\boldsymbol{u}(s)\right)ds \mid \boldsymbol{x}(t)=\boldsymbol{x},\xi(t)=\alpha\right\}. \qquad \text{(VI.4)}$$

En utilisant la méthode de programmation dynamique, Rishel a établi les conditions nécessaires et suffisantes du problème d'optimisation stochastique. Ces conditions sont écrites par des équations aux dérivées partielles dites équations d'HJB :

$$\min_{u(.)\in\mathcal{A}(t,\alpha)}\left\{g\left(\boldsymbol{x},\boldsymbol{u}\right)+v_t^{\alpha}(t,\boldsymbol{x})+v_{\boldsymbol{x}}^{\alpha}(t,\boldsymbol{x})f^{\alpha}(t,\boldsymbol{x},\boldsymbol{u})+\sum_{\beta}\lambda_{\alpha\beta}v^{\beta}(t,\boldsymbol{x}^{\alpha})\right\}=0. \quad \text{(VI.5)}$$

VI.2 Relation entre modèles Markovien et semi-Markovien

Le but de cette section est de comparer le modèle Markovien (modèle de Rishel (1975)) avec celui semi-Markovien que nous avons construit. Pour simplifier la comparaison, nous considérons la formulation du problème de commande optimale stochastique sur un horizon fini sans taux d'actualisation dans le chapitre 2. Pour celle-ci, la comparaison est basée sur des conditions d'optimum (équations d'Hamilton-Jacobi-Bellman). Cette comparaison est la suivante :

En multipliant un terme Δt et puis en introduisant le terme $v^{\alpha}(t,\boldsymbol{x}^{\alpha})$ dans les deux expressions (VI.5) et (2.8), on peut obtenir :

<u>Modèle Markovien :</u>

$$\min_{\boldsymbol{u}(t)\in\mathcal{A}(t,\alpha)}\Big\{g\left(\boldsymbol{x},\boldsymbol{u}\right)\Delta t+v_t^{\alpha}(t,\boldsymbol{x})\Delta t+v_{\boldsymbol{x}}^{\alpha}(t,\boldsymbol{x})f^{\alpha}(t,\boldsymbol{x},\boldsymbol{u})\Delta t+\left(1-\lambda_{\alpha\alpha}\Delta t\right)v^{\alpha}(t,\boldsymbol{x}^{\alpha})$$
$$+\sum_{\beta\neq\alpha}\lambda_{\alpha\beta}\Delta t v^{\beta}(t,\boldsymbol{x}^{\alpha})\Big\}=v^{\alpha}(t,\boldsymbol{x}^{\alpha}).$$

(VI.6)

<u>Modèle semi-Markovien :</u>

$$\min_{\boldsymbol{u}(t)\in\mathcal{A}(t,\alpha)}\Big\{g\left(\boldsymbol{x},\boldsymbol{u}\right)\Delta t+v_t^{\alpha}(t,\boldsymbol{x})\Delta t+v_{\boldsymbol{x}}^{\alpha}(t,\boldsymbol{x})f^{\alpha}(t,\boldsymbol{x},\boldsymbol{u})\Delta t+\left(1-p_{\alpha\alpha}(t)\Delta t\right)v^{\alpha}(t,\boldsymbol{x}^{\alpha})$$
$$+\sum_{\beta\neq\alpha}p_{\alpha\beta}(t)\Delta t v^{\beta}(t,\boldsymbol{x}^{\alpha})\Big\}=v^{\alpha}(t,\boldsymbol{x}^{\alpha}). \qquad \text{(VI.7)}$$

À gauche, on peut remarquer qu'il y a des termes différents entre ces deux expressions :

$$\left(1-\lambda_{\alpha\alpha}\Delta t\right)v^{\alpha}(t,\boldsymbol{x}^{\alpha})+\sum_{\beta\neq\alpha}\lambda_{\alpha\beta}\Delta t v^{\beta}(t,\boldsymbol{x}^{\alpha}) \quad \text{pour le modèle Markovien,}$$

$$\left(1-p_{\alpha\alpha}(t)\Delta t\right)v^{\alpha}(t,\boldsymbol{x}^{\alpha})+\sum_{\beta\neq\alpha}p_{\alpha\beta}(t)\Delta t v^{\beta}(t,\boldsymbol{x}^{\alpha}) \quad \text{pour le modèle semi-Markovien.}$$

(i) Les termes $\left(1-\lambda_{\alpha\alpha}\Delta t\right)$ et $\left(1-p_{\alpha\alpha}(t)\Delta t\right)$ sont les mêmes probabilités pour que le processus $\xi(s)$ ne saute dans l'intervalle (t, $t + \Delta t$). (ii) Les termes $\sum_{\beta\neq\alpha}\lambda_{\alpha\beta}\Delta t$ et $\sum_{\beta\neq\alpha}p_{\alpha\beta}(t)\Delta t$ sont les mêmes probabilités pour que le processus $\xi(s)$ puisse subir un certain nombre fini de sauts dans l'intervalle (t, $t + \Delta t$).

Dans le modèle Markovien, le processus stochastique est présenté par les taux de transition $\lambda_{\alpha\beta}$ qui sont indépendants du temps. Par contre, dans le modèle semi-Markovien, il est présenté par des densités de probabilité de transition $p_{\alpha\beta}(t)$ qui dépendent du temps.

Ces densités de probabilité de transition $p_{\alpha\beta}(t)$, nous permettent de modéliser des systèmes fiables dans le domaine temporel.

VI.3 Méthodes d'études

Pour l'étude du problème de commande optimale du système, il y a certaines méthodes telles que Monté Carlo, la programmation dynamique, l'étude des différences temporelles (Temporal Difference), l'étude du facteur Q (Q-learning), etc. Voir Bertkesas (2001), Sutton et Barto (2000). La méthode de programmation dynamique est utilisée pour modéliser fidèlement des systèmes dans l'environnement dynamique, ce qui est convenable pour des systèmes réels avec des incrémentaux de changement par rapport au temps. Par contre, d'autres méthodes ne font qu'approximer des fonctions objectifs $v^{\alpha}(t,x)$ comme des valeurs moyennes sur un horizon infini. Ce ne sont pas des modèles exacts; d'ailleurs, ils consument le temps.

L'avantage de la méthode de programmation dynamique est de donner un modèle analytique qui peut décrire non seulement la dynamique du système, mais aussi l'interaction des paramètres de contrôle.

Malgré l'inconvénient de cette méthode (qui est limitée pour modéliser des systèmes à grande dimension, ce que Bellman avait appelé malédiction de dimension (Bertkesas (2001), Sutton et Barto (2000))), cette dernière est assez puissante pour nous permettre de comprendre le comportement des systèmes considérés. De plus, un large système donnerait un complexe problème de commande, ce qu'il faudrait éviter d'aborder.

ANNEXE VII

MODÈLE DU SYSTÈME MANUFACTURIER

Dans cette annexe, nous allons présenter brièvement deux structures représentatives du système manufacturier du point de vue de commande optimale: l'une est un système à plusieurs machines en parallèle traitant plusieurs types de pièces, l'autre est à plusieurs machines en série. Ces structures sont présentées respectivement dans la section VII.1 et VII.2.

VII.1 Plusieurs machines en parallèle

Considérons un système constitué de plusieurs machines (identiques ou différentes). Chaque machine peut produire plusieurs types de pièces ; elle ne peut produire qu'un seul type produit à la fois et est soumise à des pannes. La structure physique en nature est représentée dans la figure VII.1.

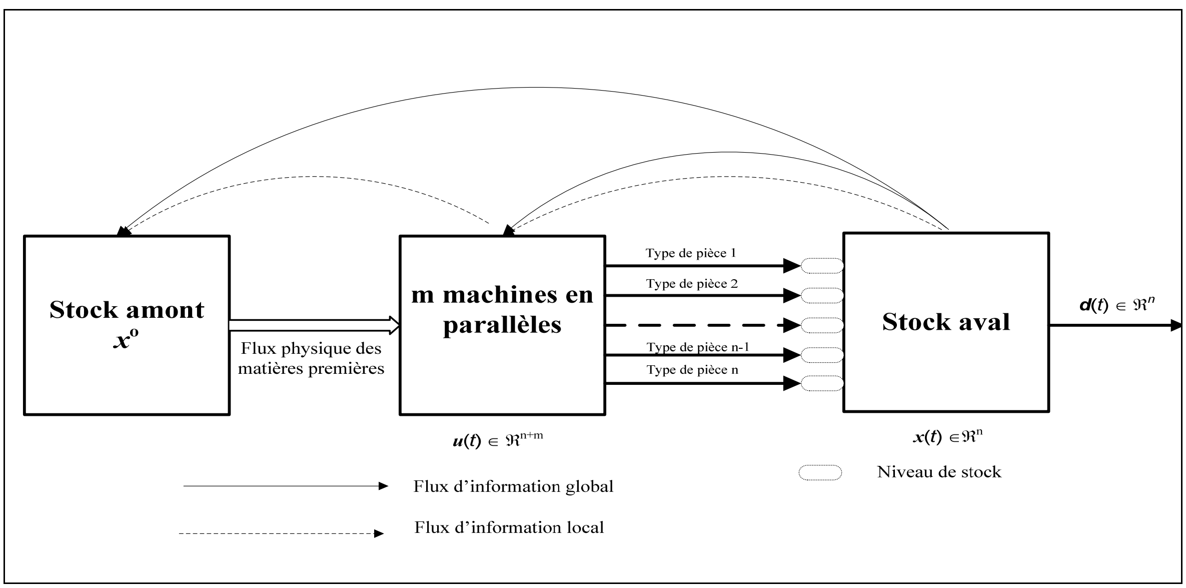


Figure VII.1 SPF à m machines en parallèles.

Pour la figure VII.1 où $\boldsymbol{d} = \{d_1, d_2,..., d_n\}$ est le vecteur de taux de demande, $\boldsymbol{u}(t) = \{u_{ij}(t),\ i = 1,2,...,n,\ j = 1,2,...,m\}$ est le vecteur de taux de production à l'instant t, $\boldsymbol{x}(t) = (x_i(t),\ i = 1,..., n)'$ est le vecteur des différents stocks à l'instant t.

VII.2 Plusieurs machines en série

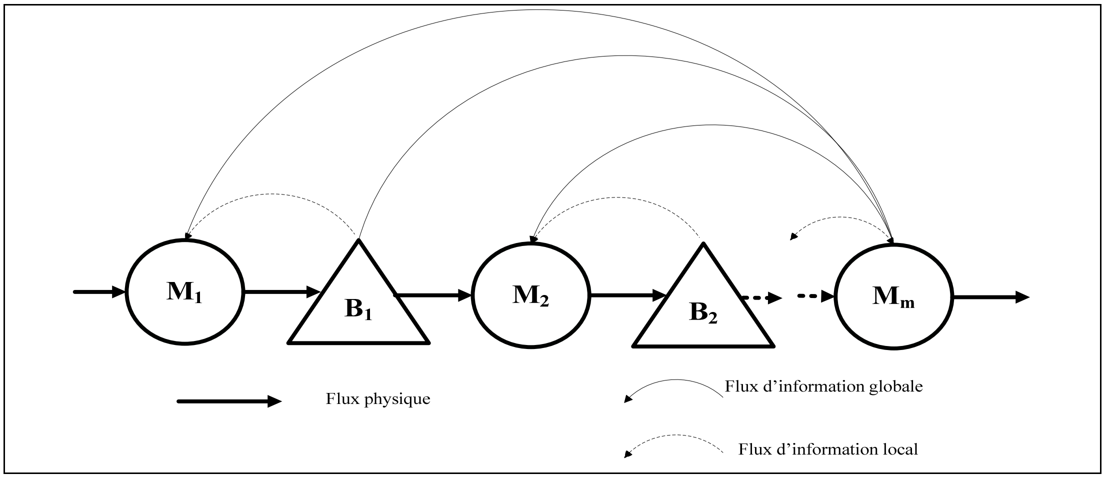


Figure VII.2 Ligne de production à m machines.

Pour la figure VII.2, le système est constitué d'un réseau à plusieurs stations de services ou machines séparées par des zones tampons ou de stockage B_i. Selon la figure VII.2, le produit circule de la machine M_1 vers la zone tampon B_1, en suite à la machine M_2 et ainsi de suite jusqu'à la dernière machine, puis enfin vers le consommateur. Ce réseau est appelé ligne de production ou ligne de transfert. Lorsque le système est opérationnel, les machines peuvent être soumises à des pannes et à des réparations de façon régulière, alors les zones tampons ne sont pas nécessaires. Cependant, les machines peuvent tomber en panne de façon irrégulière.

Dans chacune des deux figures VII.1 et VII.2, les lignes continues vers la droite sont des flux physiques du système. Les arcs continus et discontinus vers la gauche sont des flux d'informations globaux et locaux, respectivement.

BIBLIOGRAPHIE

Abbad, M., and Filar, J. A. (1991): *Perturbation theory for semi-Markov control problems*. *IEEE Conference on Decision and Control*, pp.489-493.

Akella, R., and Kumar, P. R. (1986): *Optimal control of production rate in a failure prone manufacturing system*. *IEEE Transaction on Automation Control* **31**, pp.116-126.

Altiok, T. (1997): *Performance analyse of namufacturing systems*. Pringer.

Bai, S. X., and Gershwin, S. B. (1996): *Scheduling manufacturing systems with work-in-process inventoty control: reentrant systems*. *OR Spektrum* **18**, pp.187-195.

Becker, G., Camarinopoulos, L., and Zioutas, G. (2000): *A semi-Markov model allowing for inhomogenities with respect to process time*. *Reliability Engineering and System Safety* **70**, pp. 41-48.

Bellman, R. E. (1954): *Dynamic programming of continuous processes*. Santa Monica, Calif. Rand Corp.

Bensoussan, A., and Sethi, S. P. (2007): *The machine maintenance and sale age model of Kamien and Schwart revisited*. *Management Science* **53**, pp.1964-1976.

Bertsekas, D. P. (2001): *Dynamic programming and Optimal control*. Athena Scientific.

Bielecki, T., and Kumar, P. R. (1988): *Optimality of zero-inventory policies for unreliable manufacturing systems*. *Operation Research* **36**, pp.532-541.

Blondel, F. (2007): *Gestion de la prodution*. Dunod.

Bonvik, A. M. (1996): *'Performance analysis of manufacturing systems under hybrid control policies'*, pp. 179: *Electrical Engineering and Computer Science*, Massachusetts Institute of Technology.

Boukas, E. K. (1987): *Commande optimale stochastique appliquée aux systèmes de production*: *Génie Électrique*, École Polytechnique de Montréal, Montréal.

Boukas, E. K. (1995): *Numerical methods for HJB equations of optimization problems for piecewise deterministic systems*. *Optimal Control Applications & Methods* **16**, pp.41-58.

Boukas, E. K., and Haurie, A. (1990): *Manufacturing flow control and preventive maintenance: A stochastic control approach*. *IEEE Transaction on Automation Control* **35**, pp. 1024-1031.

Boukas, E. K., and Kenne, J. P. (1997): *Maintenance and production control of manufacturing systems with setups*. *Lectures in Applied Mathematics* **33**, pp.55-70.

Boukas, E. K., Kenne, J. P., and Zhu, Q. (1995): *Age dependent hedging point policies in manufacturing systems*. *Proceedings of the ACC Conference, Seatle, Washington*.

Boukas, E. K., and Yang, H. (1996): *Optimal control of manufacturing flow and preventive maintenance*. *IEEE Transaction on Automation Control* **41**, pp.881-885.

Cao, X. R. (2007): *Stochastic learning and optimization*. Springer.

Ciprut, P., Hongler, O. M., and Salama, Y. (1998): *Hedging point for non-Markovian piecewise deterministic production processes*. *Discrete Event Dynamic Systems: Theory and Applications* **8**, pp.365-375.

Ciprut, P., Hongler, O. M., and Salama, Y. (2000): *Fluctuations of the production output of transfer lines*. *Journal of Intelligent Manufacturing* **11**, pp.183-189.

D'Amico, G., Janssen, J., and Manca, R. (2005): *Homogeneous semi-Markov reliability models for credit risk management*. *Decisions in Economics and Finance* **28**, pp.79-93.

Dallery, Y., David, R., and Xie, X. L. (1989): *Approximate Analysis of transfer lines with unreliable machines and finite buffers*. *IEEE Transaction on Automation Control* **34**, pp.943-953.

Davis, H. M. A. (1984): *Piecewise-deterministic Markov processes: a general class of non-diffusion stochastic models (with discussion)*. *Journal of the Royal Statistical Society S.B. Methodological* **46**, pp.353-388.

Dehayem, F. I. N., Kenne, J. P., and Gharbi, A. (2009): *Hierarchical decision making in production and repair/replacement planning with imperfect repairs under uncertainties*. *European Journal of Operational Research* **198**, pp.173-189.

Dimitrakos, T. D., and Kyriakidis, E. G. (2008): *A semi-Markov decision algorithm for the maintenance of a production system with buffer capacity and continuous repair times*. *International Journal of Production Economics* **111**, pp.752-762.

Dogramaci, A. and Fraiman, N.A. (2004) ‘Replacement decision with maintenance under uncertainty: An imbedded optimal control model’, *Operations Research*. Vol.52, No.5, pp.785-794.

Engell, S. (2007): *Feedback control for optimal process operation*. *Journal of Process Control* **17**, pp. 203-219.

Enginarlar, E., Li, J., and Meerkov, S. M. (2005): *Lean buffering in serial production lines with non-exponential machines*. *OP Spectrum* **27**, pp.195-219.

Fleming, W. H., and Soner, H. M. (2006): *Controlled Markov processes and viscosity solutions*. Spinger.

Fong, N. T., and Zhou, X. Y. (2000): *Optimal feedback controls in deterministic two-machine flowshops with finite buffers*. *IEEE Transaction on Automation Control* **45**, pp.1198-1203.

Gershwin, S. B. (1987): *An efficient decomposition method for approximate evaluation of temdem queues with finite storage space and blocking*. *Operation Research* **35**, pp. 291-305.

Gershwin, S. B. (1989): *Hierarchical flow control: A framework for scheduling and planning discrete events in manufacturing systems. Proceedings of the IEEE* **77**, pp..195-209

Gershwin, S. B. (2002): *Manufacturing systems engineering*. Massachusetts Institute of Technology.

Ghosh, D. and Roy, S. (2010) 'A decision-making framework for process plant maintenance', *European Journal of Industrial Engineering*. Vol.4, No.1, pp.78-98.

Glynn, P. (1989): *A GSMP formalism for discrete event systems*. *IEEE Transaction on Automation Control* **77**, pp.14-23.

Grabsky, F. (2003): *The reliability of an object with semi-Markov failure rate*. *Applied Mathematics and Computation* **135**, pp.1-16.

Howard, R. A. (1971): *Dynamic probabilitic systems*. Wiley. New York.

Janssen, J., and Manca, R. (2007): *Semi-Markov risk models for finance, insurance, and reliability*. Springer.

Jellouli, O., Chatelet, E. and Lallement, P. (2007) 'Optimal (s,S) production policies with delivery time guarantees and breakdowns', *European Journal of Industrial Engineering*. Vol.1, No. 3 pp.266-279.

Kamien, M.I. and Schwart, N.L. (1971) 'Optimal maintenance and sale age for a machine subject to failure', *Management Science*. Vol.17, No.8, pp. B495-B504.

Kang, Y. H., and Shin, H. J. (2010): *An adaptive scheduling algorithm for a parallel machine problem with rework processes*. *International Journal of Production Research* **48**, pp.95-115.

Kenne, J. P., and Boukas, E. K. (2003): *Hierarchical control of production and maintenance rates in manufacturing systems*. *Journal of Quality in Maintenance Engineering* **9**, pp.66-82.

Kenne, J. P., and Gharbi, A. (2001): *A simulation optimization approach in production planning of failure prone manufacturing systems*. *Journal of Intelligent Manufacturing* **12**, pp.421-431.

Kenne, J. P., Gharbi, A., and Beit, M. (2007): *Age-dependent production planning and maintenance strategies in unreliable manufacturing systems with lost sale*. *European Journal of Operational Research* **178**, pp. 408-420.

Kim, J., and Gershwin, S. B. (2005): *Integrated quality and quantity modeling of production lines*. *OP Spectrum* **27**, pp.287-314.

Kim, J., and Gershwin, S. B. (2008): *Analysis of long flow lines with quality and operational failures*. *IIE Transactions* **40**, pp.284-296.

Kimemia, J., and Gershwin, S. B. (1983): *An algorithm for the computer control of a flexible manufacturing system*. *IIE Transaction* **15**, pp.353-362.

Kleinrock, L. (1975): *Queueing systems*. John Wiley & Sons. New York.

Krasovskii, N. N., and Lidskii, E. A. (1961): *Analytical design of controllers in stochastic systems with velocity limited controlling action*. *Journal of Apllied Mathematics and Mechanics* **25**, pp.627-643.

Kushner, H. J., and Dupuis, P. (2001): *Numerical method for stochastic control problems in continuous time*. Springer. New York.

Lavoie, P., Kenne, J. P., and Gharbi, A. (2009): *Optimization of production control policies in failure-prone homogenous transfer lines*. *IIE Transaction* **41**, pp.209-222.

Lehoczky, J., Sethi, S. P., Soner, H. M., and Taksar. (1991): *An asymptotic analysis of hierarchical control of manufacturing systems under uncertainty. Math Operation Research* **16**, pp.596-608.

Li, J., and Meerkov, S. M. (2005a): *Evaluation of throughput in serial production lines with non-exponential machines*. *Chapter book - Springer*, pp.61-90.

Li, J., and Meerkov, S. M. (2005b): *On the coefficiens of variation of uptime and downtime in manufacturing equipment*. *Mathematical Problems in Engineering* **1**, pp.1-6.

Liberopoulos, G., and Caramanis, M. (1994): *Production control of manufacturing systems with production rate-dependent failure rates*. *IEEE Transaction on Automation Control* **39**, pp.889-895.

Lidskii, E. A. (1963): *Optimal control of systems with random properties*. *Journal of Apllied Mathematics and Mechanics* **27**, pp.42-59.

Love, C. E., and Zitron, M. A. (1998): *An SMDP approach to optimal repair/replacement decisions for systems experiencing imperfect repairs*. *Journal of Quality in Maintenance Engineering* **4**, pp.131-149.

Malhamé, R. P., and Boukas, E. K. (1991): *A renewal theoretic analysis of a class of manufacturing systems*. *IEEE Transaction on Automation Control* **36**, pp.580-587.

Miller, R. G. (1963): *Continuous time stochastic storage processes with random linear inputs and outputs*. *Journal of Mathematics and Mechanics* **12**, pp.275-291.

Nakagawa, T. (2005): *Maintenance theory of reliability*. Springer.

Older, G. J., and Suri, R. (1980): *Time optimal control of parts-routing in a manufacturing system with failure-prone machines*. *IEEE*, pp.722-727.

Perkins, J. R. (2004): *Optimal control of failure-prone manufacturing systems with constant repair times*. *Annals of Operation Research* **125**, pp.233-261.

Perkins, J. R., and Srikant, R. (1997): *Scheduling multiple part-types in an unrelibale single-machine manufacturing system*. *IEEE Transaction on Automation Control* **42**, pp.364-376.

Presman, E., Sethi, R. P., Zhang, H., and Zhang, H. (2002): *On optimality of stochastic N-Machine flowshop with long-run average cost*. *Stochastic Theory and Control* **280**, pp.399-417.

Rishel, R. (1975): *Dynamic programming and minimum principles for systems with jump Markov disturbances*. *SIAM Journal on Control* **13**, pp.338-371.

Ross, S. M. (2003): *Introduction to probability models*. Academic Press.

Sethi, S.P. and Thomson, G.L. (2000) ‘Optimal control theory: Application to management science and economics’, *Springer.*

Sethi, S. P., Zhang, H., and Zhang, Q. (2005): *Average-cost control of stochastic manufacturing systems*. Springer. New York.

Sethi, S. P., Zhang, Q., and Zhou, X. Y. (1994): *Hierarchical controls in stochastic manufacturing systems with convex costs*. *Journal of Optimization Theory and Applications* **80**, pp.299-317.

Sethi, S. P., Zhang, Q., and Zhou, X. Y. (1997): *Hierarchical production controls in a stochastic two-machine flowshop with a finite internal buffer*. *IEEE Transaction on Automation Control* **13**, pp.1-12.

Sharifnia, A. (1988): *Production control of a manufacturing system with multiple machine states*. *IEEE Transaction on Automation Control* **33**, pp.620-625.

Shu, C., and Perkins, J. R. (2001): *Optimal PHP production of multiple part-types on a failure-prone machine with quadratic buffer costs*. *IEEE Transaction on Automation Control* **46**, pp.541-549.

Srivatsant, N. (1993): *Synthesis of optimal policies for stochastic manufacturing systems*, pp. 221: *Electrical Enineering*, Massachusetts Institute of Technology.

Sutton, R. S., and Barto, A. G. (2000): *Reinforcement Learning: An Introduction*. MIT Press Cambridge

Sworder, D. D. (1969): *Feedback control of a class of linear systems with jump parameters*. *IEEE Transaction on Automation Control* **14**, pp.9-14.

Tan, B., and Gershwin, S. B. (2009): *Analysis of a general Markovian two-stage continuous-flow production system with a finite buffer*. *International Journal of Production Economics* **120**, pp. 327-339.

Tsitsiklis, J. N. (1984): *Convexity and characterization of optimal policies in a dynamic routing problem*. *Journal of Optimization Theory and Applications* **44**, pp.105-135.

Williams, B. K. (2007): *Optimal management of non-Markovian biological populations*. *Ecological Modelling* **200**, pp.234-242.