STOCHASTIC CONTROL FOR DISTRIBUTED SYSTEMS WITH APPLICATIONS TO WIRELESS COMMUNICATIONS

Minyi Huang

Departement of Electrical and Computer Engineering
McGill University, Montréal

June 2003

A Thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

ABSTRACT

This thesis investigates control and optimization of distributed stochastic systems motivated by current wireless applications. In wireless communication systems, power control is important at the user level in order to minimize energy requirements and to maintain communication Quality of Service (QoS) in the face of user mobility and fading channel variability. Clever power allocation provides an efficient means to overcome in the uplink the so-called near-far effect, in which nearby users with higher received powers at the base station may overwhelm signal transmission of far away users with lower received powers, and to compensate for the random fluctuations of received power due to combined shadowing and possibly fast fading (multipath interference) effects.

With the wireless uplink power control problem for dynamic lognormal shadow fading channels as an initial paradigm, a class of stochastic control problems is formulated which includes a fading channel model and a power adjustment model. For optimization of such a system, a cost function is proposed which reflects the QoS requirements of mobile users in wireless systems. For the resulting stochastic control problem, existence and uniqueness of the optimal control is established.

By dynamic programming, a Hamilton-Jacobi-Bellman (HJB) equation is derived for the value function associated with the stochastic power control problem. However, due to the degenerate nature of the HJB equation, the value function cannot be interpreted as a classical solution, which hinders the solution of explicit control laws or even the reliance on numerical methods. In the next step, a perturbation technique is applied to the HJB equation and a suboptimal control law using a classical solution to the perturbed HJB equation is derived. Control computation via numerical methods becomes possible and indicates an interesting equalization phenomenon for the dynamic power adjustment under an i.i.d. channel dynamics assumption. Analysis of the suboptimal control reveals an interesting bang-bang control structure which indicates simple manipulation in power adjustment. However, in view of the partial differential equations involved, implementation for systems with more than two users appears elusive.

The above stochastic power control problem suggests an investigation of a wider class of degenerate stochastic control problems which are characterized both by a weak coupling condition for the components of the involved diffusion process, and by a particular rapid growth condition in the cost function. We analyze viscosity solutions to the resulting HJB equations. We develop a localized semiconvex/semiconcave approximation technique to deal with the rapid growth condition. A maximum principle is established for the viscosity subsolution/supersolution of the HJB equation and it is used to prove uniqueness of the viscosity solution. The theoretical tools thus developed serve as a mathematical foundation for our stochastic power control problem.

At this point, with the aim of constructing an analytically more tractable solution to the wireless power control problem, we consider a linear quadratic optimization approach in which the power attenuation is treated as a random parameter. In this setup, the value function is expressed as a quadratic form of the vector of individual user powers, and the optimal feedback control is proved to be affine in the power. Unfortunately, the resulting control law remains too formidable to compute in large systems. However, based on the obtained analytic solution, we are able to develop local polynomial approximations for the value function and seek approximate solutions to the HJB equation by an algebraic approach under small noise conditions. Suboptimal control laws are also constructed using the approximate solutions. Remarkably, here the scheme for approximation solutions can be combined with a single user based design to construct a localized control law for each user in systems with large populations. The single user based design substantially reduces the complexity of determining the power control law.

It is of significant interest to consider the asymptotics of power optimization for large population systems. In such systems, it may be unrealistic to apply the standard stochastic optimal control approach due to the complexity of implementing the centralized control law. Suboptimal but distributed control laws may be more desirable. Before proceeding to investigate this challenging issue, we first consider a large-scale linear quadratic Gaussian (LQG) model for which the agents contained in the system interact with each other either via a global cost or via related individual costs. We study both the optimal control problem based on the global cost, and the LQG game based on individual costs. For the LQG game, we develop an aggregation technique based on examining individual and mass behaviour; highly localized control strategies for all agents are obtained and a so-called ε-Nash equilibrium property for these strategies is proved. Finally, we evaluate the loss incurred by opting for the the distributed game theoretic solution, versus the centralized optimal control solution, as measured by the associated costs differential.

For the large population power control problem, apart from the centralized stochastic control approach, we also consider optimization in a game theoretic context by generalizing the techniques in the large-scale LQG problem. The combination of the individual costs and state aggregation leads to decentralized power control.

RÉSUMÉ

Cette thèse investigue la commande et l'optimisation de système stochastique distribué, ces derniers étant motivés par les applications sans fil actuelles. Dans les systèmes de communication sans fil, il est important de régler la puissance au niveau de l'utilisateur dans le but de minimiser l'énergie requise et pour maintenir la qualité de service (QoS) en présence de déplacement de l'usager et de variabilité dans les voies sujettes à évanouissement. Une allocation de puissance intelligente fournit un moyen efficace de surmonter, dans la liaison montante, l'effet dit de proximité-éloignement pour lequel les usagers avoisinant et disposant d'une puissance de réception plus élevée peuvent submerger le signal de transmisson d'usagers éloignés du point d'accès sans fil. De plus, cette allocation permet de compenser pour les fluctuations aléatoires de la puissance reçue résultant d'ombrages combinés et, possiblement, d'effets d'évanouissements rapides (i.e. interférences par trajet multiple).

Avec le problème de régulation de puissance dans la liaison montante sans fil appliqué aux voies log normales dynamiques d'évanouissement d'ombrage comme paradigme de départ, une classe de problème de commande stochastique est formulée en incluant un modèle de voie dévanouissement et un modèle d'ajustement de puissance. Pour l'optimisation de tels sytèmes, une fonction de coût est proposée reflétant les demandes de QoS des usagers mobiles des systèmes sans fil. Pour le problème de commande stochastique résultant, l'existence et l'unicité de la commande optimale sont démontrées.

Par programmation dynamique, une équation de Hamilton-Jacobi-Bellman (HJB) est dérivée pour la fonction de valeur associée avec le problème stochastique de régulation de puissance. Toutefois, en raison de la nature dégénérée de l'équation

HJB, la fonction de valeur ne peut pas être interprétée comme une solution classique, ceci entrave la solution explicite de lois de contrôle et affecte même la confiance accordée aux méthodes numériques. À l'étape suivante, une technique de perturbation est appliquée à l'équation HJB et une loi de contrôle suboptimale utilisant une solution classique de l'équation HJB perturbée est dérivée. Des calculs contrôlés par méthode numérique deviennent possibles et indiquent un intéressant phénomène d'égalisation de l'ajustement dynamique de puissance sous la supposition d'une voie dynamique i.i.d. L'analyse de la commande suboptimale révèle une intéressante structure de commande de type bang-bang, i.e. indiquant une simple manipulation de l'ajustement de puissance. Néanmoins, en raison des équations différentielles aux dérivées partielles impliquées, l'implémentation de système avec plus que deux utilisateurs apparait illusoire.

Le problème stochastique de régulation de puissance ci-dessus suggère une investigation d'une classe plus large de problème de commande stochastique dégénéré caractérisé à la fois par une faible condition de couplage des composants impliqués dans le processus de diffusion et par une condition particulière de croissance rapide de la fonction de coût. Nous analysons les solutions de viscosité résultant des équations HJB. Nous développons une technique d'approximation localisée semiconvexe/semiconcave pour traiter la condition de croissance rapide. Un principe de maximisation est établi pour la sous-solution/super-solution de viscosité de l'équation HJB et celui-ci est utilisé pour prouver l'unicité de la solution de viscosité. Les outils théoriques ainsi développés sont utilisés comme fondement mathématique de notre problème stochastique de régulation de puissance.

À ce point, dans le but de construire une solution analytique avec une tractabilitée accrue pour le problème de régulation de puissance sans fil, nous considérons une approche d'optimisation quadratique linéaire dans laquelle l'atténuation de puissance est traitée comme un paramètre aléatoire. Dans cette configuration, la fonction de valeur est exprimée comme un vecteur de forme quadratique des puissances individuelles des utilisateurs, et la commande optimdal d'asservissement est prouvée être affine en puissance. Malheureusement, la loi de contrôle résultante demeure trop complexe pour le calcul de système de grande dimension. Toutefois, sur la base des solutions analytiques obtenues, nous somme capable de développer des approximations polynomiales locales de la fonction de valeur et de rechercher des solutions approximatives de l'équation HJB par une méthode algébrique soumise à des conditions de faible bruit. Des lois de contrôle suboptimales sont aussi construites en utilisant les solutions approximatives. Remarquablement, le mécanisme de solution approximative peut aussi être combiné avec un design basé sur un usager unique pour construire une loi de commande locale pour chacun des usagers dans les systèmes avec une population importante. Le design basé sur un usager unique réduit substantiellement la complexité pour déterminer la loi de commande de puissance.

Il est d'intérêt significatif de considérer les asymptotes de l'optimisation de puissance pour les systèmes avec une population importante. Pour de tels systèmes, il peut être irréaliste d'appliquer l'approche de la commande stochastique optimale en raison de la complexité de l'implémentation d'une loi de contrôle centralisée. Des lois de contrôle suboptimales mais distribuées peuvent être davantages désirables. Avant de débuter l'investigation de ce stimulant problème, nous devons d'abord considérer le cas d'un modéle linéaire quadratique gaussien (LQG) de grande dimension pour lequel les agents contenus dans le système interagissent entre eux soit via a un coût global ou via des coûts reliés entre les individus. Nous étudions à la fois le problème de commande optimale basé sur le coût global et le jeux LQG basé sur les coûts individuels. Pour le jeux LQG, nous développons une technique d'agrégation basée sur l'examen des individus et le comportements de masse; des stratégies de commande hautement localisées pour tous les agents sont obtenues et une propriété dite équilibre ϵ-Nash est prouvée pour ces stratégies. Finallement, nous évaluons la perte induite par le choix de solutions distribuées par théorie des jeux versus la solution centralisée optimal sur la base de la mesure du coût différentiel associé.

Pour le problème de commande de population importante, mis à part l'approche de la commande stochastique centralisée, nous considérons aussi l'optimisation dans le contexte de la théorie des jeux en généralisant les techniques du problème LQG de grande dimension. La combinaison des coûts individuels et l'agrégration des états mènent à la commande de puissance décentralisée.

CLAIMS OF ORIGINALITY

The following original contributions are presented in this thesis:

- Formulation of code division multiple access (CDMA) uplink wireless power adjustment as a stochastic control problem including: (1) a dynamic lognormal fading channel model, (2) a bounded rate based power control model and (3) the signal to interference based performance measure.
- Proof of existence and uniqueness of the optimal control.
- For computability, perturbation of the associated degenerate Hamilton-Jacobi-Bellman (HJB) equation and synthesis of resulting suboptimal control laws via numerical methods.
- Consideration of a related class of degenerate stochastic control problems with weakly coupled dynamics and rapid growth conditions; viscosity solution analysis; localized semiconvex/semiconcave approximation technique proposed for the proof of an associated maximum principle.
- For analytic tractability, reformulation of power allocation as a linear quadratic optimization problem; analysis of the classical solutions; suboptimal approximation methods by local polynomial equation systems; a one against the mass scheme for partially decentralized power control in systems with large populations.
- Isolation of a new class of large-scale stochastic control problems; formulation of a related linear quadratic Gaussian (LQG) optimal control and dynamic game for large population systems, namely, dynamically independent

and cost-coupled systems of significance in communications and economics, etc.

- Investigation of these large-scale LQG systems in the context of centralized and distributed (or decentralized) control. Explicit expression of the feedback control law for the centralized optimal control problem. Dynamic LQG game solution; state aggregation techniques for extracting the dynamics of the mass influence on a given agent; individual-mass behaviour analysis and approximate Nash equilibria. Discrepancy between the optimal control and decentralized game in terms of a cost gap, state trajectories as well as population behaviour.
- Formulation of power control for large population systems; the optimal control approach; initial investigation of decentralized control via a generalization of the state aggregation technique in the LQG game framework to the nonlinear power control context.

N.B. Almost all of the work above appears in articles which have been published or are currently under review and revision for publication; see page xi.

Published and Submitted Articles:

[A1] M. Huang, P.E. Caines, C.D. Charalambous, and R.P. Malhamé. Power control in wireless systems: a stochastic control formulation. *Proc. American Contr. Confer.*, Arlington, Virginia, pp.750-755, June, 2001.

[A2] M. Huang, P.E. Caines, C.D. Charalambous, and R.P. Malhamé. Stochastic power control for wireless systems: classical and viscosity solutions. *Proc. 40th IEEE Conf. Decision and Control*, Orlando, Florida, pp.1037-1042, December, 2001.

[A3] M. Huang, P.E. Caines, and R.P. Malhamé. On a class of singular stochastic control problems arising in communications and their viscosity solutions. *Proc. 40th IEEE Conf. Dec. Contr.*, Orlando, Florida, pp.1031-1037, Dec., 2001.

[A4] M. Huang, P.E. Caines, and R.P. Malhamé. Stochastic power control for wireless systems: centralized dynamic solutions and aspects of decentralized control. *Proc. 15th IFAC World Congress on Automatic Control*, Barcelona, Spain, CDROM, July, 2002.

[A5] M. Huang, P.E. Caines, and R.P. Malhamé. Uplink power adjustment in wireless communication systems: a stochastic control analysis. *Under Revision for IEEE Trans. Automatic Control.* First submission to IEEE in September, 2002.

[A6] M. Huang, P.E. Caines, and R.P. Malhamé. Degenerate stochastic control problems with exponential costs and weakly coupled dynamics: viscosity solutions and a maximum principle. *Under Revision for SIAM J. Contr. Optim.* First submission to SIAM in Oct. 2002.

[A7] M. Huang, P.E. Caines, and R.P. Malhamé. Individual and mass behaviour in large population stochastic wireless power control problems: centralized and Nash equilibrium solutions. To be presented at *the 42nd IEEE Conf. Decision and Control*, Hawaii, December, 2003.

[A8] M. Huang, R.P. Malhamé, and P.E. Caines. Quality of service control for wireless systems: minimum power and minimum energy solutions. *Proceedings of the American Control Conference*, Anchorage, Alaska, pp.2424-2429, May, 2002.

[A9] M. Huang, R.P. Malhamé, and P.E. Caines. Stochastic power control in wireless communication systems with an infinite horizon discounted cost. *Proceedings of the American Control Conference*, Denver, Colorado, pp.963-968, June, 2003.

[A10] M. Huang, R.P. Malhamé, and P.E. Caines. Stochastic power control in wireless systems and state aggregation. *Under Revision for IEEE Trans. Automat. Contr.* First submission to IEEE in January, 2003.

[A11] M. Huang, R.P. Malhamé, and P.E. Caines. Stochastic power control in wireless communication systems: analysis, approximate control algorithms and state aggregation. To be presented at *the 42nd IEEE Conf. Decision and Control*, December, 2003.

ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest gratitude to my supervisor Professor Peter E. Caines for his guidance, encouragement and support over the past years. During my research at CIM, I have benefited immensely from his extremely profound understanding of the area of system and control, and his broad knowledge on various related disciplines. His scientific foresight is a key factor for a perfect choice of my research topics. He is a superb scholar, and also a great teacher by whom I am deeply moved.

I am deeply indebted to my co-supervisor Professor Roland P. Malhamé of École Polytechnique de Montréal for his advice, encouragement and support which are crucial to the steady progress of my research. His insightful suggestions and constant flow of novel ideas have significant contributions to my work.

This thesis is reaped from the close collaboration of our mini-team of three people. I must thank Peter and Roland together for their efforts in the past four years in arranging our "Friday Poly Café Sessions" which are intellectually stimulating and very fruitful. I am sincerely grateful to both of them. I would also thank Peter and his family for their kindness.

I thank Professor Charalambos D. Charalambous of Ottawa University for his encouragement and useful advice during the early stage of my research. I thank Professor Lorne G. Mason for learning from him wired networks during the "networks seminar series".

I would thank my friends and fellow students at CIM; they have created a pleasant atmosphere for research and life. Many thanks to Shahid Shaikh, Miguel Torres-Torriti, Iakov Romanovski, Mu-Chiao Lu, Ying Xuan Duan, Ammar Haurani for their encouragement, support, and interesting conversations. I am very grateful to Miguel for very kindly teaching me how to use the sophisticated CIM computer system in the long duration of my study. My special thanks go to Dany Dionne for translating the abstract of the thesis into French.

I wish to thank the CIM staff, in particular, Jan Binder and Cynthia Davidson for their excellent management of the research facilities.

The financial support of Clifford Wong McGill Major Fellowship (2001–2002) and NCE-MITACS Program (1999–2002) is sincerely acknowledged.

I am very grateful to Professor Lei Guo, Professor Jifeng Zhang, and Professor Han-Fu Chen — my teachers at Academia Sinica, Beijing, for what I have learned from them and for their moral support.

Finally, deeply from my heart, I would thank my family — my parents, sister and brother for their endless patience, understanding, and support for many years. I am most grateful to what they have done and what they have sacrificed during my career pursuit; words would never be enough.

Minyi Huang

Montréal, Quebéc, June 2003

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

Introduction

There is a significant body of literature on stochastic control theory, which has been developed for the analysis and optimization of various physical and social systems experiencing random disturbances from their ambient environment. The existing stochastic control theory finds its applications to a vast range of areas, including industrial control systems, inventory theory, dynamic resource allocation, production planning, queuing networks, mathematical finance, and many others; see [**4, 5, 8, 47, 48, 65, 78**] among others.

In addition to the well known stochastic system models in the above mentioned areas, the emergence and advances of new technology give rise to new opportunities for formulating the associated optimal control problems within the powerful framework of stochastic control. Specifically, the rapid development of modern wireless technology has unveiled a world of characteristically complex wireless networks with inherent statistical properties concerning their dynamic behaviour. Typically, this kind of systems involve service providers as well as a great number of clients, which may be conveniently termed as agents in future analysis. Sometimes, in order to gain a more concrete sense, we will also feel free to term various variables or objects for the involved more general stochastic control systems by their counterparts in the wireless communication networks.

In this thesis we study the control and optimization of a class of distributed stochastic systems as well as their generalization where current mobile communication systems serve as a motivating technological background. In such wireless systems, a large number of mobile users are distributed in large areas and communicate with each other through one or more base stations, and the transmitted signals are subject to random fading. Modelling and optimization of such systems naturally resorts to stochastic system theory.

We set out to investigate the stochastic wireless power control problem and then investigate control problems of a more general form which are well motivated by the underlying power control problem. A feature shared by all systems considered in this thesis is that they involve multiple dynamic agents which can act based on individual interests, while their dynamics interact weakly through the utility function they seek to optimize.

Dynamic Modelling of Radio Propagation and Stochastic Power Control

There has been an extensive literature on modelling of radio propagation. Generally, radio channels experience both small-scale (short-term) fading and large-scale (long-term) fading, and various statistical models have been proposed to model the resulting random fluctuation of received signal power. In general, the two different fading effects are understood as superimposed and can be treated separately due to the different mechanisms from which they are generated. Indeed, small-scale (with a time scale of millisecond) fading is caused by multipath replicas of the same signal which in view of their respective phase shifts, can interact either constructively, or destructively. It is a problem which can be addressed via the so-called diversity techniques (see [**43, 63**]). Large-scale (with a time scale of hundreds of milliseconds) fading is caused by shadowing effects due to buildings and moving obstacles, such as trucks, partially blocking or deflecting mobile or base station signals.

In this thesis we only consider the modelling of the large-scale fading and investigate effective methods for mitigating its impairments on the channel; this will be achieved by transmission power control. In a static context, for any fixed positioning for the user and the base station, the large scale fading can be accurately modelled by a lognormal random variable, or a normal random variable measured in decibels (dB). Due to the idealized assumption of no relative motion between the transmitter and the receiver, the static lognormal modelling is inadequate for applications. To get realistic modelling for the channel condition, one has to take into account the user mobility and environment variations in the vicinity of the user in a communication scenario. This dictates the use of dynamic channel models able to capture the spatio-temporal correlation properties of fading channels. In some early research, a first order auto-regressive (AR) innovation model was proposed for modelling the large scale fading for mobile users [**28, 75**]. In this thesis, we adopt the continuous time modelling for the lognormal fading by use of stochastic differential equations introduced by Charalambous et. al. [**17**]. The dynamics is intended to model the fading channels for both outdoor and indoor users where the fading effect exhibits spatial and temporal variations.

Using the above modelling framework, in Chapter 2 we formulate the distributed stochastic control problem. A primary issue here is to determine the way the power should be adjusted. In this Chapter, a bounded rate based control model is proposed for power adjustment. It is motivated by the way power control is achieved through a sequence of fixed steps in current wireless technology. The next issue in approaching such a problem is to set the criteria for system optimization. To this end, a cost function is introduced which measures the performance of different control strategies. The cost function adopted here aims at achieving the required signal to interference ratio while limiting power usage as far as possible. The existence and uniqueness of the optimal control is investigated in Chapter 2. The analysis is complicated by the fact that one faces a degenerate stochastic control system.

In general, the degenerate Hamilton-Jacobi-Bellman (HJB) equation derived in Chapter 2 admits no classical solutions, and hence it is difficult to explicitly specify the optimal control law. To circumvent this difficulty, a meaningful approach is to consider approximating the HJB equation or the value function for the optimal control problem, via a perturbed HJB equation. This program is carried out in Chapter 3, whereby approximated value functions which are classical solutions to perturbed HJB equations can be obtained. Numerical simulations are performed to verify the satisfactory performance of the resulting suboptimal controller. In this setup for the suboptimal control law, the value function can be approximated off-line and the suboptimal control law in real time can be determined by some simple rules. However, computations are prohibitively complex for multiuser systems.

Viscosity Solutions for Systems with Rapid Growth Conditions

Chapter 4 is in itself, a contribution to the mathematics of stochastic control. we study a general class of degenerate stochastic control problems which includes the system in Chapter 2 as a special case. A viscosity solution analysis is presented in this Chapter. We develop a certain semiconvex/semiconcave approximation technique for functions with rapid growth. The approximation is achieved by use of a pair of localized envelope functions and it is proved that the envelope functions have semiconvex/semiconcave properties on a compact set when the parameters involved in the definition of the envelope functions are appropriately set. Further we apply this approximation to establish a maximum principle for the degenerate HJB equation under a weak coupling condition on the dynamics, and uniqueness of the viscosity to the HJB equation follows as a corollary. Uniqueness of the viscosity solution is an important aspect to the stochastic control problem both for understanding the nature of the optimal cost function, and developing numerical solutions, since a multiple solution situation may cause additional difficulty in finding a desired numerical approximation.

Linear Quadratic Power Optimization

The initial formulation of the stochastic power control problem, although realistic, is hindered by significant mathematical difficulties, and does not scale up easily in terms of computations.

The analysis of Chapter 5 employs instead a quadratic type cost function with the specific input bound constraint replaced by a penalty term for the input in the cost. The control problem is analyzed in terms of classical solutions. The optimal control law can be expressed analytically. We then address the important issue of the computability of the solutions to certain Riccati equations which stem from analysis of the problem. For a significant number of users, an analysis of local expansions of solutions around a steady state is useful in the small noise case because the system state is expected to spend a disproportionate of time in a small neighborhood of the steady state. The nearly optimal control law thus obtained enjoys a simple structure which enables efficient implementation in a simple feedback form. Extensive numerical approximations are developed to construct nearly optimal control laws.

Finally, we give a thorough analysis for the single user system and then apply the results to systems of large populations via a relatively coarse approximation relying on state aggregation. In the treatment of large systems, a certain scaling technique is adopted in the definition of the cost function. This is necessary in order to get a meaningful mass behaviour in a context where the number of users is allowed to increase to a significant level by assuming sufficiently large cell capacity. In this setup, after the scaling step the impact received by an individual from all the other agents can be approximated by a deterministic process which is then substituted into the control law. It turns out that the individual user can effectively adapt to the behaviour of the mass and the total population will gradually settle down to a steady state behaviour.

Large-scale Linear Quadratic Gaussian Systems and ε-Nash Equilibria

The next two Chapters of the thesis are specifically focussed on the asymptotics (as the number of agents increases) of centralized, versus distributed (or decentralized) control and the potential system performance degradation as measured by the corresponding cost differences. Game theoretic concepts play a central role. Chapter 6 is a strictly linear quadratic version of the problem, inspired in its structure by the wireless power control problem, but interesting in its own right. In Chapter 7, we consider approximations to the nonlinear power control problem, based on the analysis of Chapter 6.

In Chapter 6 we investigate a special class of Linear Quadratic Gaussian (LQG) optimization problems. In this context, the system in question consists of many players which are governed by independent dynamics subject to individual controls. All the players are linked by a global cost function with an additive structure. By a simple splitting, one can derive a set of individual costs from this global cost function. Thus the system can be optimized either based on the global cost, or starting from the individual costs. The global cost based optimization can be approached by the standard LQG method, while for the latter individual cost based dynamic game, the solution is sought in the Nash equilibrium framework. Specifically, for the individual cost based optimization, we study decentralized approximate Nash equilibria, or so-called ε-Nash equilibria. It is shown that such decentralized ε-Nash equilibria possess an inherent stability feature, which is interesting in a large population system since in this solution framework the involved individual strategies will lead the players to eventually reach a stable mass behaviour. Also, a cost gap is evidenced between the cost associated with the global cost derived control, and that associated with the individual cost derived control.

Large Population Power Control

For wireless systems accommodating a large population of users, the standard stochastic control approach suffers from high computational complexity. In additional to the heavy computational load, this approach also requires close coordination between all users in order to maximize the overall interest of the population. In a large distributed network, such coordination is a highly demanding task. In addition, the centralized optimal control approach lacks robustness in the face of misbehaviour of individuals and possibly unreliable transmission of commands.

These facts naturally suggest we consider individual cost based control for the large population power control. In Chapter 7 we make initial investigation of this approach by assigning a cost function to each mobile user. In the reformulation of the large population power control problem, it is also recognized that in the presence of a large number of users, their collective impact on a given user coalesces into a largely deterministic but time-varying signal. For optimization of the given user, the source of uncertainty reduces mainly to its own channel variation. We then apply a heuristic argument to generalize the method developed for the LQG problem to the nonlinear power control problem aiming at decentralized power allocation strategies. In this manner, we can extract the dynamics of the mass behaviour by a deterministic approximation. By a combination of the individual dynamics and the mass evolution, we obtain highly localized control laws for each user.

CHAPTER 2

Distributed Stochastic Systems Arising in Wireless Communications

2.1. The Background Systems

In the past decades, stochastic control theory has been developed and successfully applied to various areas including industrial process control, inventory theory, dynamic resource allocation, production planning, queuing networks, mathematical finance, and many others [**4, 5, 8, 47, 48, 65**]. In the framework of stochastic control and optimization, typically the dynamic behavior of the object under consideration is described by a random process. In many applications, the modelling and control of evolution of the object are relatively simple in that the underlying physical system is located and operates in a small region, the system state is of low dimension and the way the system experiences random disturbances is simple.

In this Chapter, we introduce a class of stochastic systems which differs from the traditional ones mentioned above. First, the object we are concerned with in this research consists of many sub-objects, which may be called individuals or agents with their own control objectives when participating in the system's evolution; second, the sub-objects are geographically distributed in large areas. Indeed, this highly distributed feature does not increase the complexity of the system dynamics for the individual's activity; however it may give rise to challenging issues in the design of control strategies since, in this case, it is significantly more difficult to exchange information among different components of the system for determination of individual control actions. From the specific distributed feature of such systems arises the need of designing control laws with relatively simple structure. Finally, to get a more precise modelling for practical systems, we need to take into account the effect of human behaviour which is less predictable. The human factor makes the modelling aspect more difficult.

In this Chapter the considered class of control problems is motivated by the current theory and technological implementations in wireless communications. We will use the underlying wireless communication model as the workhorse for a general theoretic analysis. In particular, we focus on the power control problem for dynamic lognormal fading channels.

In Chapter 4 we treat a more general system model of which the lognormal power control model of this Chapter is a special case. To begin with, we give a brief overview of the power control problem in the literature under various frameworks.

2.2. The Power Control Problem

In current digital communication systems, the mobile users are partitioned into different cells and each mobile user accesses the network through the base station able to provide service with lowest power requirements. Power control in cellular telephone systems is important at the user level both in order to minimize energy requirements, and to guarantee constant or adaptable Quality of Service (QoS) in the face of telephone mobility and fading channels. This is particularly crucial in code division multiple access (CDMA) systems where individual users are identified not by a particular frequency carrier and a particular frequency content, but by a wideband signal associated with a given pseudo-random number code. In such a context, the received signal of a given user at the base station views all other user signals within the same cell, as well as other cell signals arriving at the base station, as interference or noise, because both degrade the decoding process of identifying and extracting a given user's signal. Thus, it becomes crucial that individual mobiles emit power at a level which will insure adequate signal to noise ratio at the base station. More specifically, excess levels of signalling from a given mobile will act as interference on other mobile signals and contribute to an accelerated depletion of cellular phone batteries. Conversely, low levels of signalling will result in inadequate QoS. In fact, tight power control is indirectly related to the ability of the CDMA base station to accommodate as many users as possible while maintaining a required QoS [**76**].

There has been a rich literature on the topic of power control. Previous attempts at capacity determination in CDMA systems have been based on a "load balancing" view of the power control problem [**76**]. This reflects an essentially static or at best quasi-static view of the power control problem which largely ignores the dynamics of channel fading as well as user mobility. In essence, in this formulation power control at successive sampling time points is viewed as a pointwise optimization problem with total statistical independence assumed between the variables (control or signal) at distinct time points. In a deterministic framework, [**68, 69, 70**] present an attempt at reintroducing dynamics into the analysis, at least insofar as convergence analysis to the static pointwise optimum is concerned. This is achieved by recognizing that in current technological implementations, power level set points dictated by the base station to the mobile can only increase or decrease by fixed amounts. In [**1**], power control is considered for a CDMA system in which a signal to interference ratio (SIR) based utility function is assigned to each individual user; this gives rise to a multi-objective power optimization formulation. In the stochastic framework, attempts at recognizing the time correlated nature of signals are made in [**56**], where blocking is defined, not as an instantaneous reaching of a global interference level but via the sojourn time of global interference above a given level which, if sufficiently long, induces blocking. The resulting analysis employs the theory of level crossings. Stochastic approximation algorithms are proposed in [**73**] for distributed power control with constant channel gains, and mean square convergence to the optimum is proved. In [**46**], the authors proposed power control methods for Rayleigh fading channels based on outage probability. At each time snapshot, the power is computed by minimizing total power subject to outage probability constraints or by minimizing outage probability subject to power constraints. Down link power control for fading channels is studied in [**13**] by heavy traffic limit where averaging methods are used. In [**16**], the authors consider decentralized dynamic power control for a finite state Markovian channel, the power control law is determined by the so-called single-user policy where the intercell and incell interferences are approximated by a constant on the overall time duration of power control.

In contrast to the above research, the modelling and analysis of power control strategies investigated in this thesis employ continuous time wireless models which are time-varying and subject to fading. In particular, the dynamic model for power loss expressed in dBs is a linear stochastic differential equation whose properties model the long-term fading effects due to (i) reflection power loss, and (ii) power loss due to long distance transmission of electromagnetic waves over large areas [**17, 19**]. This gives rise to power loss trajectories which are log-normally distributed. Lognormal power loss models are justified by experimental data [**61, 63**]. Recently, there is an increasing interest in the effect of lognormal fading on communication quality of service; see [**81, 2, 29**].

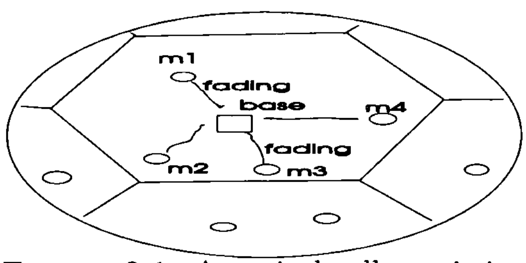


FIGURE 2.1. A typical cell consisting of a base station and many users

Motivated by the current technology in use [**62**], in this Chapter we propose a (bounded) rate based power control model for the power adjustment of log-normal fading channels and then a performance function is introduced. An important consequence of the existence of a bound on the rate of change of mobile power, is that successive uplink power adjustments can no longer be considered as a sequence of independent pointwise optimization problems (currently prevailing telecommunications view). The structure of the performance function is related to the system signal to interference ratio (SIR) requirements. We do not make direct use of the SIR or other related quantities such as the bit error rates (BER) or outage probabilities in the definition of the performance function [**22**]; instead we use a loss function integrated over time which depends upon the factors determining the SIRs and the power levels. By this means we will be able to avoid certain technical difficulty in the analysis and computation of the control laws. Our current analysis of the optimal control law of each individual user involves centralized information, i.e., the control input of each user depends on the state variable of all the users. It would be of significant interest to investigate the feasibility of decentralized control under fading channels since this would potentially reduce the system complexity for practical implementation of the control laws. This important issue will be addressed in Chapter 7.

2.3. Dynamic Modelling for Radio Propagation under User Mobility

2.3.1. Traditional Modelling for Fading. In mobile communication systems, the signal delivered from the transmitter to the receiver experiences two types of fading — small scale (short-term) fading and large-scale (long-term) fading [**61, 63**]. For the uplink of a mobile communication system our convention is that the transmitter and the receiver shall refer to the mobile user and the base station, respectively.

Small-scale fading is characterized as deep and rapid fluctuation of the amplitude of the received signal over a very short time duration or over very short travelling distances (up to a few wavelengths). This kind of rapid fading is caused by the multipath effect in which the received signal is the superposition of multiple replicas of the transmitted signal arriving at the receiver with slightly different delays [**59**]. Small-scale fading is typically modelled by Rayleigh distributions. Specifically, the received signal envelope (with amplitude $r \geq 0$) is Rayleigh distributed and is described by the density function

$$f_{ray}(r) = \begin{cases} \frac{r}{\sigma_{ray}^2} \exp(-\frac{r^2}{2\sigma_{ray}^2}), & r \geq 0, \\ 0, & r < 0, \end{cases}$$

where $\sigma_{ray}^2 > 0$ is the time average power of the received signal before envelope detection [**63**]. It can be shown that subject to Rayleigh fading the received signal power is exponentially distributed [**46**]. In certain circumstances, apart from the multipath effect there exists a dominant signal component reaching the receiver from the transmitter (for instance, due to a line-of-sight (LOS) propagation path). In this case the small-scale fading envelope has a Rician distribution. This situation is termed as Rician fading. The density function of a Rician distribution is expressed by modifying the Rayleigh density with a modified Bessel function of the first kind and of zero-order; the interested reader is referred to [**63**] for details. When the dominant component vanishes, the Rician distribution degenerates to a Rayleigh distribution.

In contrast, large-scale fading is used to characterize signal attenuation over long distances caused by shadowing effects due to variations of the terrain profile and the surroundings of the transmitter and the receiver. Large-scale fading is conveniently described in terms of large-scale path loss (PL) (simply called path loss), which measures the amount of amplitude decrease by decibels (dB) when the power is delivered from the transmitter to the receiver. Extensive experiments and their statistical analysis indicate that expressed in dB the path loss is the sum of two terms: the power-law distance loss and a zero mean random variable with a normal distribution [**63**]. The power-law distance loss is determined by the distance between the user and the base station and a power-law loss exponent. Quantitatively, the path loss is represented as

$$\begin{aligned} PL(m,B) &= [\overline{PL}(d_0) + 10\gamma \log_{10}(\frac{d(m,B)}{d_0})] + \xi_{\sigma^2}(m,B) \\ &\triangleq \overline{PL}(m,B) + \xi_{\sigma^2}(m,B), \end{aligned}$$

where d_0 is a reference distance from the base station B and γ is the power-law loss exponent which is in the range $[2, 4]$ [**63**], $d(m, B)$ denotes the distance between the position m of the mobile user and the base station, and $\overline{PL}(d_o)$ is a deterministic value representing the average large-scale path loss for a transmitter-receiver separation distance of d_0. We shall call $\overline{PL}(m, B)$ the power-law distance loss. $-PL(m, B)$ is usually called the power attenuation at m with respect to the base station B. The variance σ^2 of the spatially indexed normal random variable $\xi_{\sigma^2}(m, B)$ will also be called the standard deviation of the lognormal fading. Lognormal fading is also commonly called lognormal shadow fading due to the role of shadowing effects in generating large-scale fading. For a large suburban area (or an urban area), the standard deviation σ^2 of the lognormal fading is a constant depending on the near ground geography of large areas. The spatial correlation of the lognormal fading can be determined by experiments and is shown to decay with separation distance at an exponential rate [**28**].

In the following table we list the three frequently used models (Rayleigh, Rician and lognormal) for a comparison.

distribution	category	time-scale	caused by
Rayleigh	small-scale	millisecond	multipath
Rician	small-scale	millisecond	multipath dominant paths
Lognormal	large-scale	hundreds of milliseconds	shadowing

For radio propagation, the large-scale fading and small-scale fading are considered as superimposed and can be treated separately due to the independence assumption of the two phenomena [**49, 63**]. Also, the methods to mitigate the impairments of large-scale fading and small-scale fading are quite different. In general, practical power control algorithms can efficiently compensate for large-scale fading but cannot effectively cope with small-scale fading [**30**]; the more effective techniques to combat small-scale fading include antenna arrays, coding, etc. [**81**].

For these reasons, in the subsequent analysis we only deal with lognormal fading, and the small-scale fading will not be in the scope of our research. We note that in certain environments the small-scale component may play an increasingly important role for channel modelling.

2.3.2. Channel Variation Due to Outdoor User Mobility. In [**49**], systematic experimental investigations are carried out on an integrated simulation platform. In the experiments, power control is applied under signal propagation conditions for travelling mobile users. The radio power loss is modelled as a spatially correlated lognormal stochastic process. This also naturally gives a time correlated lognormal stochastic process for a real time power control when the spatial location of the user changes from time to time. This illustrates the apparent rationality of modelling the lognormal fading of a user by a random process under an outdoor mobility condition.

We use the following example to illustrate the spatial variation of the large-scale path loss and show the necessity of dynamic modelling of lognormal fading for mobile users. We consider a large coverage cellular system (macrocell) which is typically used in suburban areas. Let B be the location of the base station at the center of a 15 km$\times$15 km service area. Denote by $d(m_1, m_2)$ the distance between two user locations m_1 and m_2. Suppose $d(m_1, B) = 5$ km, $d(m_2, B) = 5.1$ km and $d(m_1, m_2) = 0.1$ km, i.e., m_2 is on the straight line determined by B and m_1. In the macrocell case the reference distance d_0 can be taken as 1 km [**63**]. Using the representation of the previous subsection, we write

$$PL(m_1, B) = \overline{PL}(d_0) + 10\gamma \log_{10}(\frac{d(m_1, B)}{d_0}) + \xi_{\sigma^2}(m_1, B),$$

$$PL(m_2, B) = \overline{PL}(d_0) + 10\gamma \log_{10}(\frac{d(m_2, B)}{d_0}) + \xi_{\sigma^2}(m_2, B).$$

The spatial correlation $R(m_1, m_2)$ between $\xi_{\sigma^2}(m_1, B)$ and $\xi_{\sigma^2}(m_2, B)$ can be accurately described by $R(m_1, m_2) = \sigma^2 \exp(-\frac{d(m_1,m_2)}{d^*})$ [**28, 49**].

For illustration, we take $\sigma^2 = 10$ dB, $\gamma = 2.7$, and $d^* = 0.5$ km. With this selection of d^*, the correlation for a separation distance of 0.1 km is 0.8187; see [**28**] for determination of d^* from experimental measurements. Experimental data as well as systematic statistical analysis on determining these parameters for the lognormal shadowing effect can be found in [**63, 28**]. We have

$$\overline{PL}(m_1, B) - \overline{PL}(m_2, B) = 10\gamma \log_{10}\left(\frac{d(m_1, B)}{d(m_2, B)}\right),$$

$$E|\xi_{\sigma^2}(m_1, B) - \xi_{\sigma^2}(m_1, B)| = \frac{2\sqrt{\sigma^2 - R(m_1, m_2)}}{\sqrt{\pi}}.$$

We have $\overline{PL}(m_1, B) - \overline{PL}(m_2, B) = -0.2895$ and $E|\xi_{\sigma^2}(m_1, B) - \xi_{\sigma^2}(m_2, B)| = 1.5143$. This indicates that in the process of successive user position changes, the lognormal shadowing effect actually causes a much greater fluctuation in the path loss than the increase or decrease of distance does. This clearly shows the necessity of capturing the spatial variations of the lognormal fading in a mobile communication situation.

2.3.3. Spatio-Temporal Correlation of Indoor Fading. In the classic lognormal modelling of large-scale fading, each location is assigned a lognormal random variable and experimental verification is performed with fixed transmitter-receiver positioning. In this modelling irregular human disturbances around the transmitter and receiver are neglected. For an indoor environment (consisting of walls, indoor obstacles,etc) such a simplification is not acceptable; this is due to the extreme sensitivity of propagation patterns with respect to source and obstacle motion; this motion consequently becomes a very significant aspect of the modelling exercise.

It is shown by experiments in [**26, 27**] that the local movement of personnel near the terminal (i.e. transmitter or receiver) and the local movement of the terminal around a give location (for instance, slightly shaking the terminal by the user) have drastic effect on the received power. Under such conditions the lognormal fading model still fits with the measurements, however the channel exhibits observable short time and spatial variations. So in a practical indoor communication scenario, it is infeasible to model the channel condition by a static lognormal distribution,

Well justified by the above facts, for a practical indoor communication scenario, although the user is physically confined in a small area it is more realistic and potentially more precise to introduce dynamic modelling of lognormal fading which is likely to capture the spatio-temporal variations of the channel and will be able to further characterize the underlying spatio-temporal correlation feature of the lognormal process.

2.3.4. The SDE modelling of Dynamic Channel Characteristics. For both outdoor and indoor scenarios, taking into account user mobility and variations of the surroundings of the user, the lognormal fading the user experiences can be modelled as a lognormal random process with certain statistical properties. In [**28**] a first order autoregressive (AR) innovation model was used to model the evolution of the lognormal fading for mobile users along an evenly sampled time sequence; see also [**75**]. As a natural generalization to the continuous time case, Charalambous et. al. [**17**] employ a linear stochastic differential equation (SDE) in the modelling of the channel characteristic. In both [**28, 75**] and [**17**], the basic modelling hypothesis is a Markovian assumption concerning the property of the lognormal fading process. More general but more complex modelling can be obtained by considering inhomogeneous Markovian modelling in contrast to the homogeneous (or time-invariant) Markovian models in [**17, 75, 28**]. For mobile users, the variation associated with the power-law distance loss can also be explicitly incorporated into the modelling. In general, for indoor users and outdoor users moving in a small area, the power-law distance loss can be approximated by a constant. For users travelling in a large area within the duration of service, the situation is more complicated; several factors including travelling speed, cell size and handover should be taken into account for realistic channel modelling. In the stochastic control formulation of this Chapter, we will follow the fading channel model in [**17**]. These more complicated inhomogeneous models involving high speed travelling conditions will not be considered here.

2.4. A Stochastic Optimal Control Formulation

2.4.1. The Dynamic Lognormal Fading Channel Model. Let $x_i(t)$, $1 \leq i \leq n$, denote the attenuation (expressed in dBs and scaled to the natural logarithm basis) at the instant t of the power of the i-th mobile user of a network and let $\alpha_i(t) = e^{x_i(t)}$ denote the corresponding power loss. Based on the work in [**17**], we model the power attenuation dynamics by

$$dx_i = -a_i(x_i + b_i)dt + \sigma_i dw_i, \quad t \geq 0, \quad 1 \leq i \leq n, \tag{2.1}$$

where n denotes the number of mobiles, $\{w_i,\ 1 \leq i \leq n\}$ are n independent standard Wiener processes, and the initial states $x_i(0)$, $1 \leq i \leq n$, are mutually independent Gaussian random variables which are also independent of the Wiener processes. In (2.1), $a_i > 0, b_i > 0, \sigma_i > 0,\ 1 \leq i \leq n$. The first term in (2.1) implies a long-term adjustment of x_i towards the long-term mean $-b_i$, and a_i is the speed of the adjustment. Correspondingly, the i-th power loss α_i has a long-term adjustment toward its long-term mean, which is the average large-scale path loss [**17**].

The model (2.1) corresponds to a stable diffusion process due to the positivity of a_i, and the process x_i is referred to as a mean reverting Ornstein-Uhlenbeck process [**17**], where the mean $-b_i$ of the power attenuation is explicitly incorporated into the dynamics.

2.4.2. Rate Based Power Control. Currently, the power control algorithms employed in the mobile telephone domain use gradient type algorithms with bounded step size [**62**]. This is motivated by the fact that cautious algorithms are sought which behave adaptively in a communications environment in which the actual position of the mobile and its corresponding channel properties are unknown and varying.

We model the adaptive step-wise adjustments of the (sent) power p_i (i.e., that sent in practice by the i-th mobile) by the so-called rate adjustment model [**31, 32**]

$$dp_i = u_i dt, \quad t \geq 0, \quad |u_i| \leq u_{i max}, \quad 1 \leq i \leq n, \tag{2.2}$$

where the bounded input u_i controls the size of increment dp_i at the instant t. Without loss of generality, u_{imax} will be set equal to one. The adaptive nature of practical rate adjustment control laws is replaced here by an optimal control calculation based on full knowledge of channel parameters a_i, b_i, and σ_i, $1 \leq i \leq n$. In the intended practical implementation of our solution these parameters would be replaced by on-line estimates. We write

$$x = [x_1, \cdots, x_n]^\tau, \quad p = [p_1, \cdots, p_n]^\tau, \quad u = [u_1, \cdots, u_n]^\tau.$$

Notice that the above rate adjustment model (2.2) may be compared with the up/down power control scheme proposed in [**67**] where the power of the next time step is calculated from the current power level and an additive adjustment which is optimized by a statistical linearization technique. The algorithm in [**67**] is in discrete time, and the required information for updating power includes the current power, the channel state and a target SIR.

2.4.3. Quality of Service Requirements and Criteria for Optimization.

Let $\eta > 0$ be the constant system background noise intensity which is assumed to be the same for all n mobile users in a network. Then, in terms of the power levels $p_i \geq 0$, $1 \leq i \leq n$, and the channel power attenuations α_i, $1 \leq i \leq n$, the so-called signal to interference ratio (SIR) for the i-th mobile is given by

$$\Gamma_i = \frac{\alpha_i p_i}{\sum_{j \neq i}^n \alpha_j p_j + \eta}, \qquad 1 \leq i \leq n. \tag{2.3}$$

A standard communications Quality of Service (QoS) constraint is to require that

$$\Gamma_i \geq \gamma_i > 0, \qquad 1 \leq i \leq n, \tag{2.4}$$

where γ_i, $1 \leq i \leq n$, is a prescribed set of individual target signal to interference ratios. We note that the constraints (2.4) are equivalent to the linear constraints

$\alpha_i p_i \geq \gamma_i(\sum_{j\neq i}^{n} \alpha_j p_j + \eta)$, $1 \leq i \leq n$, which, in turn, are equivalent to

$$(1+\gamma_i)\alpha_i p_i \geq \gamma_i(\sum_{j=1}^{n} \alpha_j p_j + \eta), \qquad 1 \leq i \leq n,$$

and hence to

$$\Gamma_i' = \frac{\alpha_i p_i}{\sum_{j=1}^{n} \alpha_j p_j + \eta} \geq \mu_i, \qquad 1 \leq i \leq n, \tag{2.5}$$

where $\mu_i \triangleq \frac{\gamma_i}{1+\gamma_i} > 0$, $1 \leq i \leq n$. Further, since

$$\sum_{i=1}^{n} \Gamma_i' = \frac{\sum_{j=1}^{n} \alpha_j p_j}{\sum_{j=1}^{n} \alpha_j p_j + \eta}, \tag{2.6}$$

it neccessarily follows that

$$\sum_{i=1}^{n} \mu_i < 1, \tag{2.7}$$

if (2.5) is solvable with $p_i \geq 0$, $1 \leq i \leq n$.

A plausible power allocation would be satisfying the generalized SIR requirements (2.5) with as low power consumption as possible. In a real time power allocation senario, a straightforward way to formulate the optimization problem would be to seek control functions which yield the minimization of the integrated power $\int_0^T \sum_{i=1}^{n} p_i(t)dt$ subject to the constraints (2.5)-(2.7) at each instant t, $0 \leq t \leq T$.

Here we begin by considering the pointwise global minimization of the summed power $\sum_{i=1}^{n} p_i$ under the inequality constraints (2.5)-(2.7) and the constraints $p_i \geq 0$, $1 \leq i \leq n$. Setting n inequalities in (2.5) as equalities and taking into account the constraint (2.7), we get a positive power vector $p^0 = (p_1^0, \cdots, p_n^0)$ given by

$$p_i^0 = \frac{\mu_i \eta}{\alpha_i(1 - \sum_{i=1}^{n} \mu_i)}, \qquad 1 \leq i \leq n. \tag{2.8}$$

It turns out that p^0 is the unique positive vector which minimizes $\sum_{i=1}^{n} p_i$ under constraints (2.5)-(2.7). Furthermore, it can be shown [**70**] that any nontrivial local perturbation of p^0 to a vector p which also satisfies the constraints results in a strict increase of each component p_i^0. Hence, such a p^0 is a local (linear inequality constrained) minimum which is also a global (linear inequality constrained) minimum. In other words, provided (2.7) holds, the solution to

$$\text{minimize} \quad \sum_{i=1}^{n} p_i, \quad p_i \geq 0, \tag{2.9}$$

subject to the constraints (2.5) is the unique solution to

$$\frac{\alpha_i p_i}{\sum_{j=1}^{n} \alpha_j p_j + \eta} = \mu_i, \qquad 1 \leq i \leq n. \tag{2.10}$$

(See [**70**]). Hence it is well motivated to replace the above pointwise constrained deterministic optimization problem with the corresponding unconstrained deterministic penalty function optimization problem:

$$\text{minimize} \quad \sum_{i=1}^{n} [\alpha_i p_i - \mu_i (\sum_{j=1}^{n} \alpha_j p_j + \eta)]^2 + \lambda \sum_{i=1}^{n} p_i, \quad \lambda \geq 0, \tag{2.11}$$

over $p_i \geq 0$, $1 \leq i \leq n$. However, because the power vector is a part of the stochastic channel-power system state with dynamics (2.1)-(2.2) and full state (α, p), it is impossible to instantaneously minimize (2.11) via $u(t)$ at all times t. Hence, over the interval $[0, T]$, we employ the following averaged integrated cost function:

$$E \int_0^T \{\sum_{i=1}^{n} [\alpha_i p_i - \mu_i (\sum_{j=1}^{n} \alpha_j p_j + \eta)]^2 + \lambda \sum_{i=1}^{n} p_i\} dt \tag{2.12}$$

subject to (2.1) and (2.2), where $\lambda \geq 0$. Here the small positive parameter λ is used to adjust the power level and to avoid potential power overshoot.

In the cost function (2.12), the first term of the integrand is related to the instantaneous SIR in an indirect way. If the SIR term defined by (2.4) is directly applied in the cost function, this will cause a potential zero division problem and present more analytic difficulties since in our current formulation we do not add hard constraints to ensure positivity of the powers.

In a practical implementation, the power of each user should remain positive. To meet such a requirement, one can choose appropriate control models and associated cost functions. For example, one might choose the control model

$$dp_i = u_i p_i dt, \qquad t \geq 0, \qquad 1 \leq i \leq n,$$

with a positive initial power for each individual mobile; then all power trajectories will remain positive with probability 1 on $[0, T]$. However, this and related setups may deviate significantly from the current technology in that the power adjustment is done in an additive way in practice. Instead, we use the rate based control model and the cost function introduced above. By choosing a small weight coefficient λ and increasing the upper bound u_{imax} for the control input, we can guarantee that the optimally controlled power process $\widehat{p}$ obtained below in the stochastic optimal control framework takes non-positive values with only a small probability. For a better understanding of this point, we consider the ideal powers for minimizing the integrand of (2.12). For a fixed time, we take the attenuations as constants and write

$$\begin{pmatrix} \widetilde{p}_1 \\ \widetilde{p}_2 \\ \cdots \\ \widetilde{p}_n \end{pmatrix} = \begin{pmatrix} 1-\mu_1 & -\mu_1 & \cdots & -\mu_1 \\ -\mu_2 & 1-\mu_2 & -\mu_2 & \cdots \\ \cdots & & & \cdots \\ -\mu_n & -\mu_n & \cdots & 1-\mu_n \end{pmatrix} \begin{pmatrix} \alpha_1 p_1 \\ \alpha_2 p_2 \\ \cdots \\ \alpha_n p_n \end{pmatrix}. \tag{2.13}$$

And we write the integrand in (2.12) as

$$\sum_{i=1}^{n} (\widetilde{p}_i - \mu_i \eta)^2 + \lambda \sum_{i=1}^{n} \beta_i \widetilde{p}_i, \tag{2.14}$$

where the coefficients β_i are determined from (2.13). The minimum of (2.14) is attained at

$$\widetilde{p}_i = \frac{2\mu_i \eta - \lambda \beta_i}{2}, \qquad 1 \leq i \leq n. \tag{2.15}$$

Thus, when the attenuations are fixed and $0 \leq \lambda << 1$, (2.15) gives a positive vector $\widetilde{p}$. By a straightforward algebraic calculation it can be further shown that under assumption (2.7), the coefficient matrix in (2.13) has an inverse with all positive entries and therefore we can obtain a positive power vector p^0 from $\widetilde{p}$. Although p^0

cannot be realized by a control input, the optimal control will try to track p^0. Once the actual power is deviating from p^0, a greater penalty results. In such a manner the optimal control makes efforts to steer the optimally controlled power to be positive with a large probability. We remark that it is of interest to consider power adjustment using the rate based power control model (2.2) with positive power constraints. This issue will be addressed in Chapter 4 in a more general context.

We introduce the assumption:

(H2.1) The positive constants $\mu_i, 1 \le i \le n$, in (2.12) satisfy $\sum_{i=1}^n \mu_i < 1$. □

Throughtout Chapters 2-3 we assume **(H2.1)** holds for the formulation of the power control problem. However we note that technically **(H2.1)** is used only in the proof of Theorem 2.1 below.

2.5. Optimal Control and the HJB Equation

In the following we will analyze the optimal control problem in terms of the state vector (x, p); this facilitates the definition of the value function v since x_i is defined on R, while α_i is only defined on R^+, $1 \le i \le n$. Clearly the results in terms of x can be re-expressed in terms of the power loss α by substitution of variables. Further define

$$f(x) = \begin{pmatrix} -a_1(x_1 + b_1) \\ \vdots \\ -a_n(x_n + b_n) \end{pmatrix}, \qquad H = \begin{pmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_n \end{pmatrix},$$

$$z = \begin{pmatrix} x \\ p \end{pmatrix}, \qquad \psi = \begin{pmatrix} f \\ u \end{pmatrix}, \qquad G = \begin{pmatrix} H \\ 0 \end{pmatrix},$$

where the second block of G is an $n \times n$ zero matrix. Now we write the equations (2.1) and (2.2) together in the vector form

$$dz = \psi dt + G dw, \qquad t \ge 0, \tag{2.16}$$

where w is an $n \times 1$ standard Wiener process determined by (2.1). In further analysis we will denote the state variable either by (x, p) or by z, or in a mixing form, when it is convenient. We also rewrite the integrand in (2.12) in terms of (x, p) as

$$l(z) = l(x,p) = \sum_{i=1}^{n} [e^{x_i} p_i - \mu_i (\sum_{j=1}^{n} e^{x_j} p_j + \eta)]^2 + \lambda \sum_{i=1}^{n} p_i, \qquad \lambda \geq 0.$$

The admissible control set is specified as

$$\mathcal{U} = \{u(\cdot) \mid u(t) \text{ is adapted to } \sigma(x_s, p_s, s \leq t), \text{ and } u(t) \in U \triangleq [-1,1]^n, \ \forall\, 0 \leq t \leq T\}.$$

As is stated in Section 2.4.1, the initial state vector is independent of the $n \times 1$ Wiener process; we make the additional assumption that p has a deterministic initial value $p(0)$ at $t = 0$. Then it is evident that $\sigma(x_s, p_s, s \leq t) = \sigma(x_0, w_s, s \leq t)$. We also introduce

$$\mathcal{L} = \{u(\cdot) \mid u \text{ is adapted to } \sigma(x_s, p_s, s \leq t), \text{ and } E\textstyle\int_0^T |u_t|^2 dt < \infty, \}.$$

If $\mathcal{L}$ is endowed with an inner product $\langle u, u' \rangle = E \int_0^T u^\tau u' ds$, for $u, u' \in \mathcal{L}$, then $\mathcal{L}$ constitutes a Hilbert space. By the above inner product we can induce a norm $\|\cdot\|$ on $\mathcal{L}$. Under this norm $\mathcal{U}$ is a bounded, closed and convex subset of $\mathcal{L}$. Finally, the cost associated with the system (2.16) and a control $u(\cdot)$ is specified to be

$$J(s, x, p, u) = E[\int_s^T l(x_t, p_t) dt | x_s = x, p_s = p],$$

where $s \in [0, T]$ is taken as the initial time of the system; further we set the value function

$$v(s, x, p) = \inf_{u \in \mathcal{U}} J(s, x, p, u), \tag{2.17}$$

and simply write $J(0, x, p, u)$ as $J(x, p, u)$.

Theorem 2.1. If Assumption **(H2.1)** holds, there exists a unique optimal control $\widehat{u} \in \mathcal{U}$ such that

$$J(x_0, p_0, \widehat{u}) = \inf_{u \in \mathcal{U}} J(x_0, p_0, u),$$

where (x_0, p_0) is the initial state at time $s = 0$, and uniqueness holds in the following sense: if $\widetilde{u} \in \mathcal{U}$ is another control such that $J(x_0, p_0, \widetilde{u}) = J(x_0, p_0, \widehat{u})$, then $P_\Omega(\widetilde{u}_s \neq \widehat{u}_s) > 0$ only on a set of times $s \in [0, T]$ of Lebesgue measure zero, where Ω is the underlying probability sample space.

PROOF. The existence of the optimal control can be established by a typical approximation argument on the subset $\mathcal{U}$ of the Hilbert space $\mathcal{L}$, and the details are omitted here (see, e.g., [**78**]).

Uniqueness: Assume there is $\widetilde{u} \in \mathcal{U}$ such that $J(x_0, p_0, \widetilde{u}) = J(x_0, p_0, \widehat{u})$, and denote the power corresponding to $\widetilde{u}$ by $\widetilde{p}$. For any fixed $x \in \mathrm{R}^n$, by **(H2.1)** it can be verified that $\frac{\partial^2 l(x,p)}{\partial p^2}$ is strictly positive definite, and therefore $l(x, p)$ is strictly convex with respect to p. So we have

$$l(x_s, \frac{1}{2}(\widehat{p}_s + \widetilde{p}_s)) \leq \frac{1}{2}[l(x_s, \widehat{p}_s) + l(x_s, \widetilde{p}_s)], \tag{2.18}$$

and a strict inequality holds on the set $A^0 \triangleq \{(s, \omega) \in [0, T] \times \Omega, \widetilde{p}_s \neq \widehat{p}_s\}$. Now we assume that $E \int_0^T 1_{(\widehat{p}_s \neq \widetilde{p}_s)} ds > 0$, i.e., A^0 has a strictly positive measure, and then the control $\frac{1}{2}(\widehat{u} + \widetilde{u}) \in \mathcal{U}$ yields

$$J(x_0, p_0, \frac{1}{2}(\widehat{u} + \widetilde{u})) < \frac{1}{2}[J(x_0, p_0, \widehat{u}) + J(x_0, p_0, \widetilde{u})] = \inf_{u \in \mathcal{U}} J(x_0, p_0, u),$$

by integrating and taking expectation on both sides of (2.18), which is a contradiction, and therefore

$$E \int_0^T 1_{(\widehat{p}_s \neq \widetilde{p}_s)} ds = 0. \tag{2.19}$$

Since with probability 1 the trajectories of p_s are continuous, by (2.19) we have

$$\widetilde{p}_s - \widehat{p}_s \equiv 0 \qquad \text{on } [0, T]$$

with probability 1. By (2.2) we have

$$\int_0^s (\widetilde{u}_t - \widehat{u}_t)dt = \widehat{p}_s - \widetilde{p}_s,$$

for all $s \in [0,T]$, so that with probability 1, $\widetilde{u}_s - \widehat{u}_s = 0$ a.e. on $[0,T]$; or equivalently,

$$E\int_0^T 1_{(\widetilde{u}_s \neq \widehat{u}_s)} ds = \int_0^T P_\Omega(\widetilde{u}_s \neq \widehat{u}_s)ds = 0.$$

So that $P_\Omega(\widetilde{u}_s \neq \widehat{u}_s) > 0$ only on a set of times $s \in [0,T]$ of Lebesgue measure zero. This proves uniqueness. □

Proposition 2.1. The value function v is continuous on $[0,T] \times \mathrm{R}^{2n}$, and furthermore,

$$v(t,x,p) \leq C(1 + \sum_{i=1}^n p_i^4 + \sum_{i=1}^n e^{4x_i}), \tag{2.20}$$

where $C > 0$ is a constant independent of (t,x,p).

PROOF. The continuity of v can be established by continuous dependence of the cost on the initial condition of the system (2.16). The inequality (2.20) is obtained by a direct estimate of the cost function. □

The above growth estimate of the value function will be used in Chapter 3 to define a function class in which approximate solutions are sought.

CHAPTER 3

The HJB Equation and its Approximation

3.1. The Dynamic Programming Equation

In this Chapter we continue the investigation of the stochastic control problem introduced in Chapter 2. The notation used below is consistent with that of Section 2.5 of the previous Chapter. Formally applying dynamic programming to the stochastic optimal control problem formulated in Section 2.5, Chapter 2, we may write the Hamilton-Jacobi-Bellman (HJB) equation for the value function v defined by (2.17) as follows:

$$\begin{aligned} 0 &= -\frac{\partial v}{\partial t} + \sup_{u \in U}\{-\frac{\partial^{\tau} v}{\partial z}\psi\} - \frac{1}{2}\mathrm{tr}(\frac{\partial^2 v}{\partial z^2}GG^{\tau}) - l, \qquad (3.1) \\ v(T, x, p) &= 0, \end{aligned}$$

where $z = (x^{\tau}, p^{\tau})^{\tau}$. It is seen that in (3.1) the covariance matrix GG^{τ} is not of full rank. In general, under such a condition the corresponding stochastic optimal control problem does not admit classical solutions due to the degenerate nature of the resulting HJB equations. The solution for such an HJB equation can be formulated in the *viscosity solution* framework. The definition of a viscosity solution will be given in Chapter 4. After the existence of a viscosity solution is proved for the HJB equation, an interesting issue arises as to whether the value function is the unique viscosity solution in a certain function class. For such degenerate HJB equations proving the uniqueness of viscosity solutions is not only of apparent mathematical interest, but is also important for analyzing convergence to the viscosity solution for certain approximation schemes [**7, 24**].

In order to prove uniqueness of the viscosity solution to the above HJB equation, we introduce the function class $\mathcal{G}$ such that each $v(t,x,p) \in \mathcal{G}$ satisfies

(i) $v \in C([0,T] \times \mathrm{R}^{2n})$ and

(ii) there exist $C, k_1, k_2 > 0$ such that $|v| \le C[1 + \sum_{i=1}^n e^{k_1|x_i|} + \sum_{i=1}^n (|x_i|^{k_2} + |p_i|^{k_2})]$, where the constants C, k_1, k_2 can vary with each v.

Theorem 3.1. The value function v defined by (2.17) is a viscosity solution to the HJB equation (3.1), and moreover, the value function v is a unique viscosity solution to (3.1) in the class $\mathcal{G}$.

PROOF. It is easy to verify that the stochastic control problem formulated in Section 2.5 is a special case of the class of stochastic control problems in Section 4.2. Specifically, the system (2.16) satisfies Assumptions **(H4.1)-(H4.2)** of Chapter 4. Hence by Theorems 4.1 we see that v defined by (2.17) is a viscosity solution to (3.1).

Obviously the value function v is in the class $\mathcal{G}$ by Proposition 2.1. By Theorem 4.3 it follows that v is a unique viscosity solution to (3.1) in the class $\mathcal{G}$. □

3.2. Perturbation of the HJB Equation

As is pointed out in Section 3.1, in general, one cannot prove the existence of a classical solution to the HJB equation (3.1) due to the lack of uniform parabolicity. Now we modify (3.1) by adding a perturbing term $\frac{1}{2}\sum_{i=1}^n \varepsilon^2 \frac{\partial^2 v}{\partial {p_i}^2}$ and formally carrying out the minimization to get

$$0 = \frac{\partial v^\varepsilon}{\partial t} + \frac{1}{2}\sum_{i=1}^n \sigma_i^2 \frac{\partial^2 v^\varepsilon}{\partial x_i^2} + \frac{1}{2}\sum_{i=1}^n \varepsilon^2 \frac{\partial^2 v^\varepsilon}{\partial p_i^2} - \sum_{i=1}^n \frac{\partial v^\varepsilon}{\partial x_i} a_i(x_i + b_i) - \sum_{i=1}^n \left|\frac{\partial v^\varepsilon}{\partial p_i}\right| + l, \quad (3.2)$$

where we use v^ε to indicate the dependence of the solution on $\varepsilon > 0$. We will seek a classical solution v^ε in the class $\mathcal{F}$:

(i) $v^\varepsilon \in C^{1,2}((0,T) \times \mathrm{R}^{2n}) \cap C([0,T] \times \mathrm{R}^{2n})$ and

(ii) $|v^\varepsilon| \le C(1+|p|^{k_1}+e^{k_2|x|})$, where $C, k_1, k_2 > 0$ can vary with each v^ε, and

(iii) $v^\varepsilon(T,x,p) = 0$.

We prove the existence of a solution to (3.2) in $\mathcal{F}$ by an approximation approach. First we fix $0 < \varepsilon < 1$. For integer $d \ge 1$, we introduce a cut-off function $h^d(x,p) = h^d(z)$ such that $h^d(z) = 1$ for $|z| \le d$, $h^d(z) = 0$ for $|z| \ge d+1$, and $|{h^d}_{z_i}| \le 2$, $1 \le i \le 2n$. Write the auxiliary equation

$$\begin{aligned} 0 = {} & \frac{\partial v^d}{\partial t} + \frac{1}{2}\sum_{i=1}^n \sigma_i^2 \frac{\partial^2 v^d}{\partial x_i^2} + \frac{1}{2}\sum_{i=1}^n \varepsilon^2 \frac{\partial^2 v^d}{\partial p_i^2} - \sum_{i=1}^n \frac{\partial v^d}{\partial x_i} a_i(x_i+b_i)h^d(z) \\ & - \sum_{i=1}^n \left|\frac{\partial v^d}{\partial p_i}\right| h^d(z) + l(x,p)h^d(z), \\ & v^d(T,x,p) = 0. \end{aligned} \tag{3.3}$$

Theorem 3.2. The equation (3.2) has a unique classical solution in the class $\mathcal{F}$ for all $\varepsilon > 0$.

PROOF. The existence of a classical solution can be proved in a way similar to the proof of Theorem VI6.2 [**23**], and it can be shown first that (3.3) admits a classical solution v^d in the class $\mathcal{F}$. Fix any $d_0 > 1$. We take $D = (0,T) \times (|(x,p)| < d_0)$. Then for any $d \ge d_0$, $v^d(t,x,p)$ in (3.3) satisfies (3.2) for $|z| < d_0$, and moreover v^d, $v^d_{x_i}$, $v^d_{p_i}$ are uniformly bounded on D with respect to d. For any $Q = (0,T) \times (|z| < d')$, $0 < d' < d_0$, by local estimates it can be shown that

$$|v^d|^{(2)}_{\lambda,Q} \triangleq |v^d|_{\lambda,Q} + |v^d_s|_{\lambda,Q} + \sum_i |v^d_{z_i}|_{\lambda,Q} + \sum_{i,j} |v^d_{z_i z_j}|_{\lambda,Q}$$

is uniformly bounded with respect to d, where $|\cdot|_{\lambda,Q}$ denotes the $L^\lambda(Q)$ norm. In the above we can take $\lambda > n+2$, and therefore by the Hölder estimates, $v^d_{z_i}$ satisfies a uniform Hölder condition on Q. We can further use the Hölder estimates to show that v^d_s, $v^d_{z_i z_j}$, $d = d_0+1, d_0+2, \cdots$, satisfy a uniform Hölder condition on Q. Finally we use Arzela-Ascoli theorem [**64**] to take a subsequence $\{d_{k_q}, q \ge 1\}$ of $\{d_k \triangleq d_0 + k, k \ge 1\}$ such that $v^{d_{k_q}}$, $v_s^{d_{k_q}}$, $v_{z_i}^{d_{k_q}}$, $v_{z_i z_j}^{d_{k_q}}$ converge uniformly to v^ε, v^ε_s, $v^\varepsilon_{z_i}$, $v^\varepsilon_{z_i z_j}$ on Q, respectively, as $q \to \infty$, where v^ε satisfies (3.2) and is in the class $\mathcal{F}$. By the growth condition of v^ε, we can use Itö's formula to show that v^ε is the value function to a related stochastic control system, and thus it is a unique solution to (3.2) in the class $\mathcal{F}$. □

Theorem 3.3. For $0 < \varepsilon < 1$ and B a compact subset of R^{2n}, if v^ε is the solution of (3.2) in class $\mathcal{F}$, then $v^\varepsilon \to v$ uniformly on $[0,T] \times B$, where v is the value function of the system (2.16).

PROOF. Suppose $\{w_i,\ \nu_i,\ 1 \le i \le n\}$ are mutually independent standard Wiener processes. Write

$$dp_i^\varepsilon = u_i dt + \varepsilon d\nu_i, \quad 1 \le i \le n. \tag{3.4}$$

Here we use $\mathcal{U}^{w,\nu}$ to denote $\sigma(w_i, \nu_i)$-adapted controls satisfying $|u_i| \le 1,\ 1 \le i \le n$. It can be shown that the optimal cost of the system (2.16) does not change when in (2.17) $\mathcal{U}$ is replaced by $\mathcal{U}^{w,\nu}$. In fact, in both cases of admissible control set $\mathcal{U}$ and $\mathcal{U}^{w,\nu}$ we can prove by dynamic programming that the resulting value functions are a viscosity solution to the associated HJB equation (3.1) in the class $\mathcal{F}$ and the viscosity solution is unique; see the viscosity solution analysis in Chapter 4 or [**36**]. Hence in the following proof we always take controls from $\mathcal{U}^{w,\nu}$. And in fact, $v^\varepsilon \in \mathcal{F}$ determined by (3.2) is the value function to the stochastic control problem (2.1)-(3.4), i.e.,

$$v^\varepsilon(s,x,p) = \inf_{u \in \mathcal{U}^{w,\nu}} J(s,x,p,u) = \inf_u E\Big[\int_s^T l(x_t, p_t^\varepsilon) dt | x_s = x, p_s = p\Big].$$

For a fixed $u \in \mathcal{U}^{w,\nu}$, we have $P\{\lim_{\varepsilon \to 0} \sup_{0 \le t \le T} |p_i^\varepsilon - p_i| = 0\} = 1$, and using Lebesgue's dominated convergence theorem [**64**] we obtain

$$|J^\varepsilon(s,x,p,u) - J(s,x,p,u)| \to 0, \qquad \text{as } \varepsilon \to 0,$$

and therefore, $v^\varepsilon(s,x,p) \to v(s,x,p)$, as $\varepsilon \to 0$. It is easy to verify that $v^\varepsilon(s,x,p)$ is uniformly bounded on $[0,T] \times B$ for $0 < \varepsilon < 1$. Furthermore, by taking two different initial conditions we can show that on $[0,T] \times B$, v^ε is equicontinuous with respect to

$0 < \varepsilon < 1$. By Arzela-Ascoli theorem, $v^\varepsilon(s,x,p) \longrightarrow v(s,x,p)$ uniformly on $[0,T] \times B$, as $\varepsilon \longrightarrow 0$. □

3.3. Interpretation of Bounded Rate Control

In the HJB equation (3.1), the value function is specified by the formal use of its first and second order derivatives, and then the equation is interpreted in a viscosity solution sense. Evidently the optimal control is not specified as a function of time and the state variable globally due to the nondifferentiable points of the value function. However, by checking the Dynamic Programming Principle at any point (t,x,p) such that the value function v is differentiable in a neighbourhood of (t,x,p), we see that locally the optimal control can be specified by the derivative information of the value function around such a point and the control input is a bang-bang control.

After the perturbation of HJB equation, the associated suboptimal cost function is differentiable everywhere. Then the suboptimal control law is constructed by the rule:

$$u = \arg\min_{u\in U} \psi^T \frac{\partial v^\varepsilon}{\partial z}, \tag{3.5}$$

which also gives a bang-bang control. We note that the suboptimal control law (3.5) resembles the up/down power control algorithms in [**67**] where at each discrete time instant the power is increased or decreased by a fixed amount and the increment is determined by the current power, the observed random channel gain and a target SIR level. But our method here differs from [**67**] since the fading dynamics modelled by (2.1) are incorporated into the calculation of the control law (3.5). Clearly for (3.5), $u_i = -\text{sgn}\frac{\partial v^\varepsilon}{\partial p_i}$ since $u_i \in [-1,1]$. In a discrete time implementation, we assume the time axis is evenly sampled by a period of ΔT. At time $k\Delta T$, $k = 0,1,2,\cdots$, the i-th user only needs to increase or decrease its power by ΔT in the case $\frac{\partial v^\varepsilon}{\partial p_i}|_{t=k\Delta T} < 0$ or $\frac{\partial v^\varepsilon}{\partial p_i}|_{t=k\Delta T} > 0$, respectively; if $\frac{\partial v^\varepsilon}{\partial p_i}|_{t=k\Delta T} = 0$, the power increment for p_i is set as 0. The significance of the suboptimal control law is that it gives a very simple scheme (i.e., increase or decrease the power by a fixed amount or keep the same power level) for updating the power of users by requiring limited information exchange between the base station and the users (in the current technology, the base station sends the power adjustment command to the users based on its information on the operating status of each user), and thus reduces implementational complexity.

On the other hand, from the structure of the suboptimal control law we see that each user should use centralized information, i.e., the current powers and attenuations of all the users, to determine its own power adjustment. In general, to implement the centralized control law requires more information exchange between the base station and the individual users than in the case of static channels [**68, 69, 82**].

3.4. Numerical Implementation of ε-Perturbation Suboptimal Control

From the above analysis it is seen that for a numerical implementation, we only need to choose a small positive constant $\varepsilon > 0$ and solve equation (3.2) and the suboptimal control is determined in a feedback control form. Consider the case of two users with i.i.d. channel dynamics

$$dx_i = -a(x_i + b)dt + \sigma dw_i, \qquad i = 1, 2, \quad 0 \le t \le 1.$$

We take the time interval [0, 1] and use a performance function $E\int_0^1 l(x_t, p_t)dt$ with

$$\begin{aligned} l =& [e^{x_1}p_1 - 0.4(e^{x_1}p_1 + e^{x_2}p_2 + 0.25)]^2 \\ &+ [e^{x_2}p_2 - 0.4(e^{x_1}p_1 + e^{x_2}p_2 + 0.25)]^2 + \lambda(p_1 + p_2). \end{aligned}$$

In order to compute the suboptimal control law, we need to solve the approximation equation numerically,

$$\begin{aligned} 0 =& v_t + \frac{1}{2}\sigma^2(v_{x_1x_1} + v_{x_2x_2}) + \frac{1}{2}\varepsilon^2(v_{p_1p_1} + v_{p_2p_2}) \\ &- a(x_1 + b)v_{x_1} - a(x_2 + b)v_{x_2} - |v_{p_1}| - |v_{p_2}| + l, \\ v(1, x, p) =& 0. \end{aligned} \tag{3.6}$$

The above equation is solved by a standard difference scheme [**3**] in a bounded region

$$S = \{(t, x, p):\ 0 \le t \le 1,\ -4 \le x_1, x_2 \le 3,\ |p_1|, |p_2| \le 3\}.$$

An additional boundary condition is added such that $v(t, x, p)|_{\overline{\partial}} = 0$, where $\overline{\partial} = \partial S - \{(t, x, p), t = 0\}$. We let δt, $h > 0$ be the step sizes, and denote $z = (x_1, x_2, p_1, p_2)^\tau$, $e_i = (0, \cdots, 1, \cdots, 0)^\tau$ where 1 is the i-th entry in the row. We discretize (3.6) to get the difference equation

$$\begin{aligned}
0 =& \frac{1}{\delta t}[v(t + \delta t, z) - v(t, z)] \\
&+ \frac{\sigma^2}{2h^2}[v(t, z + e_1 h) + v(t, z - e_1 h) - 2v(t, z)] \\
&+ \frac{\sigma^2}{2h^2}[v(t, z + e_2 h) + v(t, z - e_2 h) - 2v(t, z)] \\
&+ \frac{\varepsilon^2}{2h^2}[v(t, z + e_3 h) + v(t, z - e_3 h) - 2v(t, z)] \\
&+ \frac{\varepsilon^2}{2h^2}[v(t, z + e_4 h) + v(t, z - e_4 h) - 2v(t, z)] \\
&- \frac{a(x_1 + b)}{h}[v(t, z + e_1 h) - v(t, z)]1_{\{a(x_1+b) \le 0\}} \\
&- \frac{a(x_1 + b)}{h}[v(t, z) - v(t, z - e_1 h)]1_{\{a(x_1+b) > 0\}} \\
&- \frac{a(x_2 + b)}{h}[v(t, z + e_2 h) - v(t, z)]1_{\{a(x_2+b) \le 0\}} \\
&- \frac{a(x_2 + b)}{h}[v(t, z) - v(t, z - e_2 h)]1_{\{a(x_2+b) > 0\}} \\
&+ \frac{u_1}{2h}[v(t, z + e_3 h) - v(t, z - e_3 h)] \\
&+ \frac{u_2}{2h}[v(t, z + e_4 h) - v(t, z - e_4 h)] + l(z),
\end{aligned} \tag{3.7}$$

where

$$u_1 = -\text{sgn}[v(t, z + e_3 h) - v(t, z - e_3 h)], \tag{3.8}$$

$$u_2 = -\text{sgn}[v(t, z + e_4 h) - v(t, z - e_4 h)]. \tag{3.9}$$

With the boundary condition and an initial approximate solution, we can determine the variables u_1 and u_2 (the control variables) by the rules (3.8)-(3.9), and update the numerical solution. The iterations converge to the exact solution to the difference equation (3.7), as can be proved by the method in [**53**]. We remark that there are general results concerning the convergence of this type of difference scheme to the solution of the original partial differential equation. The interested reader is referred to the literature (see, e.g., [**24, 50, 51**]).

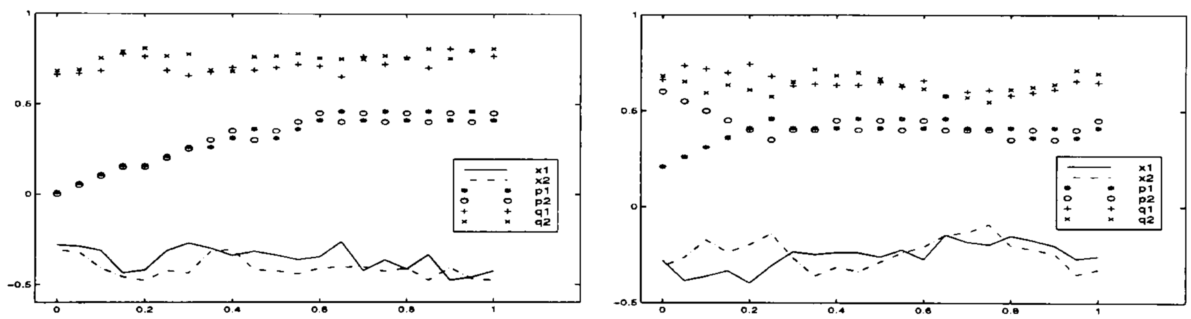
3.4.1. Numerical Examples. In the numerical simulations, we consider the system with parameters $a = 4$, $b = 0.3$, $\sigma^2 = 0.09$, $\varepsilon^2 = 0.15$, and three cases for λ: (1) $\lambda = 0.01$; (2) $\lambda = 0.001$; (3) $\lambda = 0$. In the difference scheme the step size is 0.1 for t, x_i, p_i, $i = 1, 2$. To improve the approximation we may reduce ε, and at the same time we should reduce h to guarantee convergence of iterations of the difference scheme [**53**]. In the simulation, the value function will be further interpolated to get a step size of 0.05 which will help reduce overshoot in the power adjustment. The power loss processes are also discretized with a time step size of 0.05. In the control determination, a current time space vector (t, x_1, x_2, p_1, p_2) is mapped to a grid point. Then the control is determined by the descent direction of the value function with respect to the control input u_i, $i = 1, 2$. If either increasing or decreasing the power level does not result in an evident decrease of the value function, we set the control to be 0. Figures 3.1-3.3 present the numerical simulation results where x_1, x_2 denote the attenuations, p_1, p_2 denote the powers for two users, and q_1, q_2 are the pointwise optimal powers obtained from (2.8). Figures 3.1, 3.2, 3.3 correspond to cases 1, 2, 3, respectively. It can be seen in all of the cases that after a certain period of time, the two power levels are very close to each other.

When at the initial time one mobile has a significantly different power level than the other, we see that an interesting equalization phenomenon takes place; this is shown in Figures 3.1 (b), 3.2 (b) and 3.3 (b). Starting from the initial instant the controller will first make the mobile with a high power level reduce power and the other increase power; after a certain period however both mobiles will increase their power together. This happens because a large difference in the two powers induces a large penalty in the performance function.

Figures 3.1, 3.2 and 3.3 show that the power is rather sensitive to the weight factor λ. When the cost function places a small emphasis upon power saving the optimal power trajectories are seen to be close to the pointwise optimal powers.

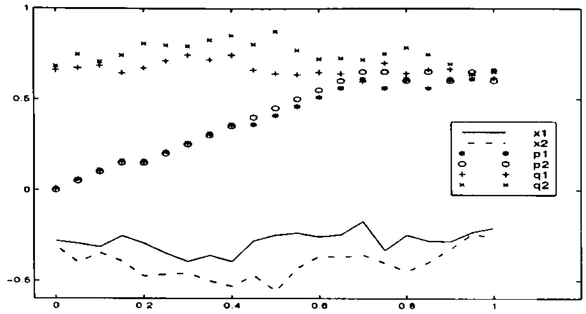
Figures 3.4 and 3.5 demonstrate the the trajectories of the attenuation, the controlled power and the associated control input for two users. The control for each user has a bang-bang feature where for most of time it takes -1 or 1, and the control is set as 0 for some rare cases when the calculated gradient of the value function w.r.t. the control is very small and thus treated as 0 derivative.

Figure 3.6 shows two surfaces of the value function at different times when the attenuations are fixed. It illustrates the variation of the value function w.r.t. different power levels.

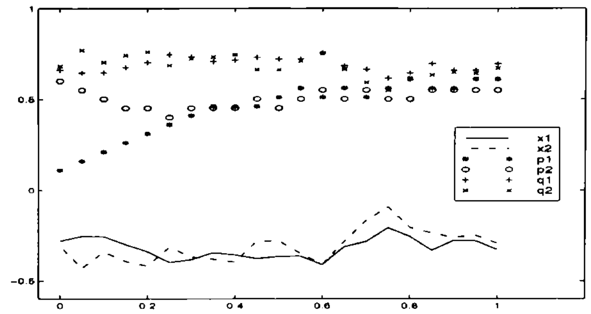


(a) $p_1(0) = 0.01, p_2(0) = 0$ (b) $p_1(0) = 0.21, p_2(0) = 0.6$

FIGURE 3.1. The trajectories for the attenuation x_i and power p_i; Different initial powers are used in (a) and (b); The power weight $\lambda = 0.01$

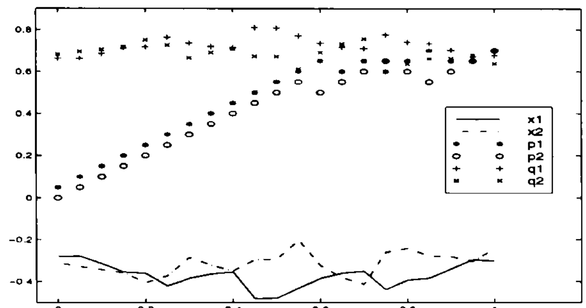


(a) $p_1(0) = 0.01, p_2(0) = 0$

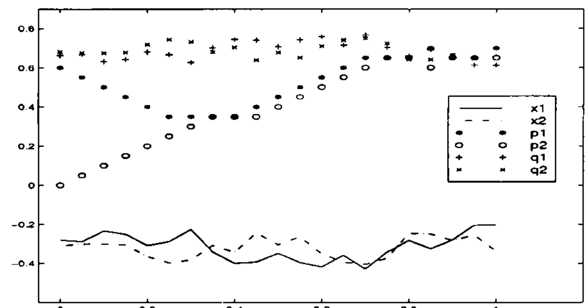


(b) $p_1(0) = 0.11, p_2(0) = 0.6$

FIGURE 3.2. The trajectories for the attenuation x_i and power p_i; The power weight $\lambda = 0.001$

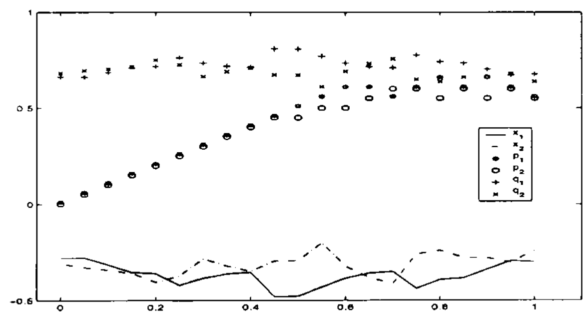


(a) $p_1(0) = 0.05, p_2(0) = 0$

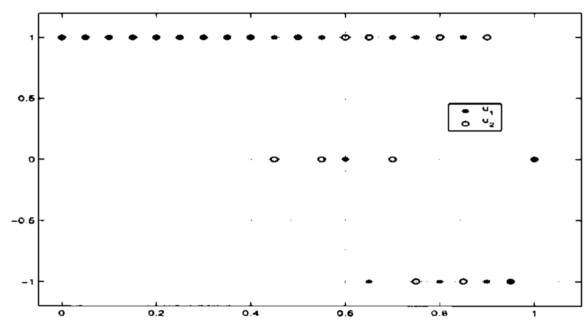


(b) $p_1(0) = 0.6, p_2(0) = 0$

FIGURE 3.3. The trajectories for the attenuation x_i and power p_i; The power weight $\lambda = 0$

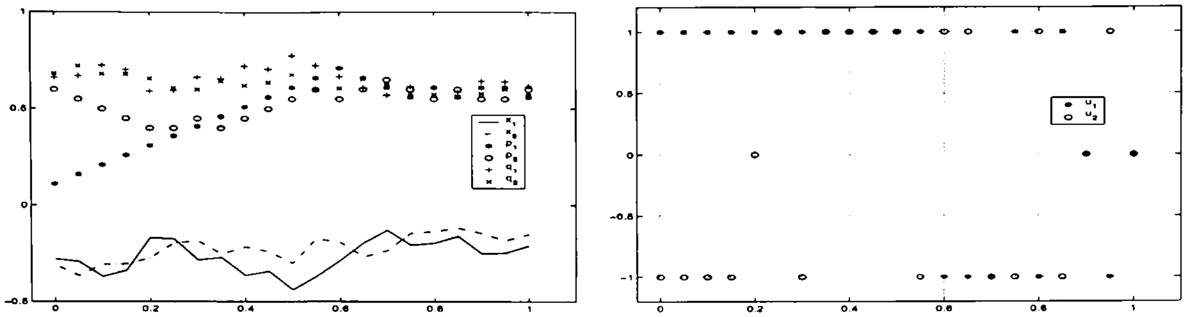


(a) $p_1(0) = 0.01, p_2(0) = 0, \lambda = 0.001$



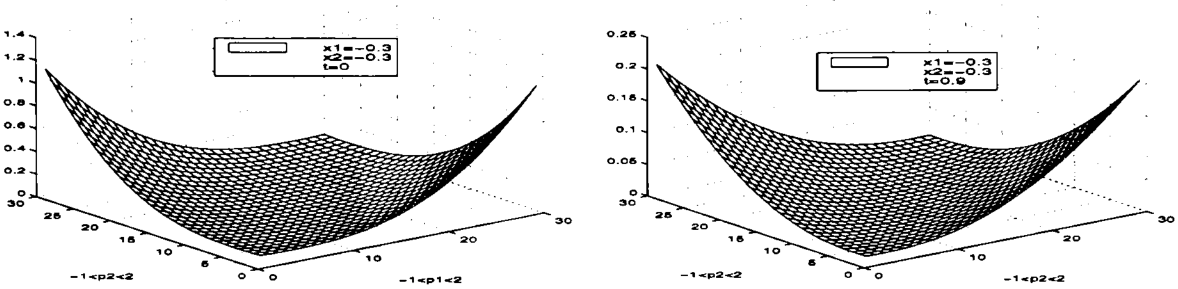
(b) u_1, u_2, $\lambda = 0.001$

FIGURE 3.4. Left: the trajectories for the attenuation x_i and power p_i; Right: the control input u_i of two users; The power weight $\lambda = 0.001$



(a) $p_1(0) = 0.11, p_2(0) = 0.6, \lambda = 0.001$ (b) $u_1,\ u_2,\ \lambda = 0.001$

FIGURE 3.5. Left: the trajectories for the attenuation x_i and power p_i; Right: the control input u_i of two users; The power weight $\lambda = 0.001$



(a) $x_1, x_2 = -0.3, t = 0$ (b) $x_1, x_2 = -0.3, t = 0.9$

FIGURE 3.6. The surfaces of the value function for fixed x and varying p; The power weight $\lambda = 0.001$

CHAPTER 4

Viscosity Solution Analysis for a Class of Degenerate Stochastic Control Problems

4.1. Introduction

This Chapter is concerned with a class of optimization problems arising from the power control problem for wireless communication systems and forms a mathematical foundation for the results in Chapters 2 and 3 and the papers [**32, 35**]. The material in this Chapter follows the papers [**33, 36**]. We will first formulate a class of degenerate stochastic control problems which take the form of the regulation the state of a controlled process where an exogenous random parameter process is involved in the performance function, and then we show that the communication application reduces to a special case for the general formulation. A mathematical finance problem will also be introduced for illustration of the general case.

The random parameter process and the controlled process are denoted by $x_t \in \mathrm{R}^n$ and $p_t \in \mathrm{R}^n$, $t \in \mathrm{R}_+$, respectively. Suppose x is modelled by the stochastic differential equation

$$dx = f(t,x)dt + \sigma(t,x)dw, \qquad t \geq 0, \tag{4.1}$$

where f and σ are the drift and diffusion coefficients, respectively, w is an n dimensional standard Wiener process with covariance $Ew_t w_t^T = tI$ and the initial state x_0 is independent of $\{w_t,\ t \geq 0\}$ with finite exponential moment, i.e., $Ee^{2|x_0|} < \infty$.

The process p is governed by the model

$$dp = g(t,p,u)dt, \qquad t \geq 0, \tag{4.2}$$

where the component $g_i(t,p,u), 1 \leq i \leq n$, controls the size of the increment dp_i at the time instant t, $u \in \mathrm{R}^n$, $|u_i| \leq u_{imax}$, $1 \leq i \leq n$. Without loss of generality we set $u_{imax} = 1$, and we shall write

$$x = [x_1, \cdots, x_n]^\tau, \quad p = [p_1, \cdots, p_n]^\tau, \quad u = [u_1, \cdots, u_n]^\tau.$$

In the regulation of p, we introduce the following cost function

$$E\int_0^T [p^\tau C(x)p + 2D^\tau(x)p]dt, \tag{4.3}$$

where $T < \infty$, $C(x), D(x)$ are $n \times n$ positive definite matrix (for any $x \in \mathrm{R}^n$), $n \times 1$ vector, respectively, and the components of $C(x)$ and $D(x)$ are exponential functions of linear combinations of x_i, $1 \leq i \leq n$. For simplicity, in this Chapter we take $C_{ij}(x) = c_{ij}e^{x_i+x_j}$, $D_i(x) = d_ie^{x_i} + s_i$ for $1 \leq i,j \leq n$, where $c_{ij}, d_i, s_i \in \mathrm{R}$ are constants. This particular structure of the weight coefficients indicates that each p_i is directly associated with the parameter component x_i through the cost function for $1 \leq i \leq n$, when expanding the integrand in (4.3) into its components. The more general case of expressing the components of $C(x)$ and $D(x)$ as exponential function of general linear combinations of x_i, $1 \leq i \leq n$, can be considered without further difficulty. We will give the complete optimal control formulation in Section 4.2, where the technical assumptions of *weak coupling* for the dynamics (4.1)-(4.2) will be introduced.

4.1.1. The Stochastic Power Control Example.

We now briefly review the motivating stochastic power control problem for lognormal fading channels. In an urban or suburban environment, the power attenuations of wireless networks are described by lognormal random processes. Let $x_i(t)$, $1 \leq i \leq n$, denote the power attenuation (expressed in dBs and scaled to the natural logarithm basis) at the instant

t of the i-th mobile and let $\alpha_i(t) = e^{x_i(t)}$ denote the actual attenuation. Based on the work in [**17**], the power attenuation dynamics adopted in Chapters 2, 3 and in the papers [**32, 35**] are given as a special form of (4.1):

$$dx_i = -a_i(x_i + b_i)dt + \sigma_i dw_i, \quad t \geq 0, \quad 1 \leq i \leq n. \tag{4.4}$$

In (4.4) the constants $a_i, b_i, \sigma_i > 0$, $1 \leq i \leq n$. (See [**17**] for a physical interpretation of the parameters in (4.4)). In a network, at time t the i-th mobile sends its power $p_i(t)$ and the resulting received power at the base station is $e^{x_i(t)}p_i(t)$. The mobile has to adjust its power p_i in real time so that a certain Quality of Service (QoS) is maintained. In Chapters 2, 3 and [**31, 32, 38**] the adjustment of the (sent) power vector p for the n users is modelled by simply taking $g(t,p,u) = u$ in (4.2) which is called the rate adjustment model. Subsequently, based on the system signal to interference ratio (SIR) requirements, the following averaged integrated performance function

$$E\int_0^T \{\sum_{i=1}^n [e^{x_i}p_i - \mu_i(\sum_{j=1}^n e^{x_j}p_j + \eta)]^2 + \lambda \sum_{i=1}^n p_i\}dt \tag{4.5}$$

was employed, where $\eta > 0$ is the system background noise intensity, $\lambda \geq 0$, and $\mu_i, 1 \leq i \leq n$, is a set of positive numbers determined from the SIR requirements. The resulting power control problem is to adjust u as a function of the system state (x,p) so that the above performance function is minimized.

4.1.2. A Mathematical Finance Example. In the area of mathematical finance we take a special form of (4.1) in which

$$dx_i = f_i(t)dt + \sigma_i(t)dw_i, \qquad 1 \leq i \leq n, \tag{4.6}$$

where $f_i(t)$ and $\sigma_i(t)$ are continuous on $[0,T]$. Taking $\alpha_i = e^{x_i}$ we obtain from (4.6)

$$\begin{aligned} \frac{d\alpha_i}{\alpha_i} &= [f_i(t) + \frac{\sigma^2(t)}{2}]dt + \sigma_i(t)dw_i \\ &\triangleq b_i(t)dt + \sigma_i(t)dw_i, \qquad 1 \leq i \leq n, \end{aligned} \tag{4.7}$$

which is the so-called geometric Brownian motion (GBM) model and is well known in mathematical finance for modelling prices of risky assets, for instance, stocks [**57, 48**]. (4.7) is also the fundamental stock price model in the celebrated Black-Scholes theory [**12, 48**]. We now suppose (4.7) models the prices for n stocks. A shareholder's decisions are usually made by means of adjusting the fraction of wealth invested on the n stocks while consideration is also given to possibly other investment (for instance, savings) as well as to consumption. This leads to utility based portfolio optimization. Let p_i, $1 \leq i \leq n$, stand for the number of shares of the i-th stock. In the process of asset management, at time t the value carried by the stock shares of the investor is the sum of the terms $e^{x_i(t)}p_i(t)$, $1 \leq i \leq n$. The share number p_i varies with time according to the investing strategy of the shareholder and since this is a controlled quantity a connection with the power control problem in this thesis is revealed which will be studied in future research.

4.1.3. Organization of the Analysis. The analysis in this Chapter treats a general class of performance functions that have an exponential growth rate with respect to $x_i, 1 \leq i \leq n$; hence this analysis covers the loss function in (4.5) and it differs from that appearing in most stochastic control problems in the literature, where linear or polynomial growth conditions usually pertain [**23, 78**]. Two novel features of the class of models (4.1)-(4.2) are (i) neither the drift nor the diffusion of the state subprocess x are subject to control and hence x can be regarded as an exogenous signal, and (ii) further, the controlled state subprocess p has no diffusion part. Hence (4.1)-(4.2) gives rise to degenerate stochastic control systems. As is well known, the optimization of such systems leads to degenerate Hamiltonian-Jacobi-Bellman (HJB) equations which in general do not admit classical solutions [**24, 78**].

This Chapter deals with the mathematical control theoretic questions arising from the class of stochastic optimal control problems considered in Chapter 3 and [**32**] where some approximation and numerical methods are proposed for implementation of the control laws. For the resulting degenerate HJB equations, we adopt viscosity solutions and show that the value function of the optimal control is a viscosity solution. To prove uniqueness of the viscosity solution, we develop a localized semiconvex/semiconcave approximation technique. Specifically, we introduce particular localized envelope functions in the unbounded domain to generate semiconvex/semiconcave approximations on any compact set. Compared to previous works [**24, 78**], by use of the set of envelope functions, we can treat very rapid growth conditions, and we note that no Lipschitz or Hölder type continuity assumption is required for the function class involved. We also consider the optimal control subject to state constraints which leads to the formulation of constrained viscosity solutions to the associated second order HJB equations; this part is parallel to [**66**], where a first order HJB equation is investigated. This Chapter is organized as follows: in Section 4.2 we state existence and uniqueness of the optimal control, and show that the value function is a viscosity solution to a degenerate HJB equation; we then give two theorems as the main results about the solution of the HJB equation. Section 4.3 is devoted to introducing a class of semiconvex/semiconcave approximations for continuous functions; this technique permits us to treat viscosity solutions with rapid growth. In Section 4.4, we analyze the HJB equation and prove a maximum principle by which it follows that the HJB equation has a unique viscosity solution in a certain function class. Section 4.5 considers the control problem subject to state constraints.

4.2. Optimal Control and the HJB Equations

Define

$$z = \begin{pmatrix} x \\ p \end{pmatrix}, \qquad \psi = \begin{pmatrix} f \\ g \end{pmatrix}, \qquad G = \begin{pmatrix} \sigma \\ 0_{n\times n} \end{pmatrix}.$$

We now write the equations (4.1) and (4.2) together in the vector form

$$dz = \psi dt + G dw, \qquad t \geq 0. \tag{4.8}$$

In the following analysis we will denote the state variable by (x, p) or z, or in a mixing form; As we do in Section 4.5, we may also write the functions in (4.1)-(4.2) in a unifying way in terms of (t, z). We write the integrand in (4.3) as

$$l(z) = l(x, p) = p^{\tau} C(x) p + 2D^{\tau}(x) p, \tag{4.9}$$

where $C(x) > 0$ for all $x \in \mathrm{R}^n$. The admissible control set is specified as

$$\mathcal{U} = \{u(\cdot) \mid u_t \text{ is adapted to } \sigma(z_s, s \leq t) \text{ and } u_t \in U \triangleq [-1,1]^n, \forall\, 0 \leq t \leq T\}.$$

As is stated in the introduction, the initial state vector is independent of the $n \times 1$ Wiener process; we make the additional assumption that p has a deterministic initial value p_0 at $t = 0$. Then it is easily verified that $\sigma(z_s, s \leq t) = \sigma(x_s, s \leq t)$. Define $\mathcal{L} = \{u(\cdot) \mid u \text{ is adapted to } \sigma(z_s, s \leq t), u_t \in \mathrm{R}^n \text{ and } E\int_0^T |u|^2 ds < \infty\}$. If we endow $\mathcal{L}$ with an inner product $\langle u, u'\rangle = E\int_0^T u^{\tau} u' ds,\ u, u' \in \mathcal{L}$, then $\mathcal{L}$ constitutes a Hilbert space with the induced norm $\|u\| = \langle u, u\rangle^{\frac{1}{2}} \geq 0,\ u \in \mathcal{L}$. Under this norm, $\mathcal{U}$ is a bounded, closed and convex subset of $\mathcal{L}$. Finally, the cost associated with the system (4.8) and a control $u \in \mathcal{U}$ is specified to be

$$J(s, z, u) = E[\int_s^T l(z_t) dt | z_s = z],$$

where $s \in [0, T]$ is taken as the initial time of the system; further we set the value function $v(s, z) = \inf_{u \in \mathcal{U}} J(s, z, u)$, and simply write $J(0, z, u)$ as $J(z, u)$.

The following assumptions on the time interval $[0, T]$ will be used in our further analysis:

(H4.1) In (4.1)-(4.2), $f \in C([0,T] \times \mathrm{R}^n, \mathrm{R}^n)$, $\sigma \in C([0,T] \times \mathrm{R}^n, \mathrm{R}^{n \times n})$, $g \in C([0,T] \times \mathrm{R}^{2n}, \mathrm{R}^n)$ and f, σ, g satisfy a uniform Lipschitz condition, i.e., there exists a constant $C_1 > 0$ such that $|f(t,x) - f(s,y)| \leq C_1(|t-s| + |x-y|)$, $|\sigma(t,x) - \sigma(s,y)| \leq C_1(|t-s| + |x-y|)$, $|g(t,p,u) - g(s,q,u)| \leq C_1(|t-s| + |p-q|)$ for all $t, s \in [0,T]$, $x, y, p, q \in \mathrm{R}^n$, $u \in U$. In addition, there exists a constant C_σ such that $|\sigma_{ij}(t,x)| \leq C_\sigma$ for $1 \leq i, j \leq n$ and $(t, x) \in [0,T] \times \mathrm{R}^n$. □

(H4.2) For $1 \le i \le n$, $f_i(x)$ can be written as $f_i(x) = -a_i(t)x_i + f_i^0(t,x)$, where $a_i(t) \ge 0$ for $t \in [0,T]$, and $\sup_{[0,T]\times \mathrm{R}^n} |f_i^0(t,x)| \le C_{f^0}$ for a constant $C_{f^0} \ge 0$. □

Throughout this Chapter we assume **(H4.1)** holds. **(H4.2)** is used only in Theorems 4.3 and 4.2.

Remark 4.1. Assumption **(H4.1)** ensures existence and uniqueness of the solution to (4.8) for any fixed $u \in \mathcal{U}$. In **(H4.1)**, the Lipschitz condition with respect to t will be used to obtain certain estimates in the proof of uniqueness of the viscosity solution. Here σ is assumed to be bounded to get a finite cost for any initial state and admissible control. Obviously **(H4.2)** covers the lognormal fading channel model. From **(H4.2)** it is seen that the evolution of x_i does not receive strong influence from the other state components $x_j, j \neq i$, in the sense that the cross term $f_i^0(t,x)$ is bounded by a constant. **(H4.2)** shall be called the *weak coupling condition* which will be used to establish uniqueness of the viscosity solution. □

Proposition 4.1. Assuming in the control model (4.2), $g(t,p,u)$ is linear in p and u, i.e., there exist continuous functions A_t, B_t such that $g(t,p,u) = A_t p + B_t u$, there exists an optimal control $\widehat{u} \in \mathcal{U}$ such that $J(x_0, p_0, \widehat{u}) = \inf_{u\in\mathcal{U}} J(x_0,p_0,u)$, where (x_0,p_0) is the initial state at time $s=0$; if in addition, B_t is invertible for all $t \in [0,T]$, then the optimal control $\widehat{u}$ is unique and uniqueness holds in the following sense: if $\widetilde{u} \in \mathcal{U}$ is another control such that $J(x_0,p_0,\widetilde{u}) = J(x_0,p_0,\widehat{u})$, then $P_\Omega(\widetilde{u}_s \neq \widehat{u}_s) > 0$ only on a set of times $s \in [0,T]$ of Lebesgue measure zero, where Ω is the underlying probability sample space. □

The proof of Proposition 4.1 can be given in the same say as the proof of Theorem 2.1 and is omitted here.

Proposition 4.2. The value function $v(s,z)$ is continuous on $[0,T] \times \mathrm{R}^{2n}$, and furthermore

$$v(s,z) \leq C[1 + \sum_{i=1}^{n} e^{4z_i} + \sum_{i=n+1}^{2n} z_i^4], \tag{4.10}$$

where $C > 0$ is a constant independent of (s,z).

PROOF. The continuity of v can be established by continuous dependence of the cost on the initial condition of the system (4.8). For an initial state $z_s = z$ and any fixed input u, from the equation (4.8), using the structure of $C(x)$ and $D(x)$ in the cost integrand we have the estimates

$$\begin{aligned} J(s,z,u) = E\int_s^T l(z_t)dt &\leq E\int_s^T C_0[1 + \sum_{i=1}^{n} e^{4z_i(t)} + \sum_{i=n+1}^{2n} z_i^4(t)]dt \\ &\leq C[1 + \sum_{i=1}^{n} e^{4z_i} + \sum_{i=n+1}^{2n} z_i^4], \end{aligned}$$

for some constants C_0, C independent of (s,z), and (4.10) follows. □

We see that in (4.8) the noise covariance matrix GG^{τ} is not of full rank. In general, under such a condition the corresponding stochastic optimal control problem does not admit classical solutions due to the degenerate nature of the arising HJB equations. Here we analyze viscosity solutions.

Definition 4.1. $\underline{v}(t,z) \in C([0,T] \times \mathrm{R}^{2n})$ is called a **viscosity subsolution** to the HJB equation

$$\begin{aligned} 0 = -\frac{\partial v}{\partial t} + \sup_{u \in U}\{-\frac{\partial^{\tau} v}{\partial z}\psi\} - \frac{1}{2}\mathrm{tr}(\frac{\partial^2 v}{\partial z^2}GG^{\tau}) - l, \\ v\mid_{t=T} = h(z), \quad z \in \mathrm{R}^{2n}, \end{aligned} \tag{4.11}$$

if $\underline{v}\mid_{t=T} \leq h$, and for any $\varphi(t,z) \in C^{1,2}([0,T] \times \mathrm{R}^{2n})$, whenever $\underline{v} - \varphi$ takes a local maximum at $(t,z) \in [0,T) \times \mathrm{R}^{2n}$, we have

$$-\frac{\partial \varphi}{\partial t} + \sup_{u \in U}\{-\frac{\partial^{\tau} \varphi}{\partial z}\psi\} - \frac{1}{2}\mathrm{tr}(\frac{\partial^2 \varphi}{\partial z^2}GG^{\tau}) - l \leq 0, \quad z \in \mathrm{R}^{2n} \tag{4.12}$$

at (t, z). $\overline{v}(t, z) \in C([0,T] \times \mathrm{R}^{2n})$ is called a **viscosity supersolution** to (4.11) if $\overline{v}\mid_{t=T} \geq h$, and in (4.12) we have an opposite inequality at (t, z), whenever $\overline{v} - \varphi$ takes a local minimum at $(t, z) \in [0, T) \times \mathrm{R}^{2n}$. $v(t, z)$ is called a **viscosity solution** if it is both a viscosity subsolution and a viscosity supersolution. □

Theorem 4.1. The value function v is a viscosity solution to the HJB equation

$$0 = -\frac{\partial v}{\partial t} + \sup_{u \in U}\{-\frac{\partial^\tau v}{\partial z}\psi\} - \frac{1}{2}\mathrm{tr}(\frac{\partial^2 v}{\partial z^2}GG^\tau) - l, \tag{4.13}$$

$$v(T, z) = 0.$$

PROOF. The value function v is continuous (by Proposition 4.2) and it satisfies the boundary condition in (4.13). Now, for any $\varphi(t, z) \in C^{1,2}([0, T] \times \mathrm{R}^{2n})$, suppose $v - \varphi$ has a local maximum at (s, z_0), $s < T$. We denote by $z^{(1)}, z^{(2)}$ the first n and last n components of z, respectively. In the following proof, we assume that $\varphi(t, z) = 0$ for all $z^{(1)}$ such that $|z^{(1)} - z_0^{(1)}| \geq C > 0$; otherwise we can multiply $\varphi(t, z)$ by a C^∞ function $\zeta(z^{(1)})$ with compact support and $\zeta(z^{(1)}) = 1$ for $|z^{(1)} - z_0^{(1)}| \leq \frac{C}{2}$. We take a constant control $u \in [-1, 1]$ on $[s, T]$ to generate z_u with initial state $z_s = z_0$ and write $\Delta(t, z) = v(t, z) - \varphi(t, z)$. Since (s, z_0) is a local maximum point of $\Delta(t, z)$, we can find $\epsilon > 0$ such that $\Delta(s_1, z) \leq \Delta(s, z_0)$ for $|s_1 - s| + |z - z_0| \leq \epsilon$. For $s_1 \in (s,\ T]$, $z_s = z_0$, write $1_{A^\epsilon} = 1_{(|s_1 - s| + |z_{s_1} - z_0| \geq \epsilon)}$. Then

$$
\begin{aligned}
&E[\Delta(s, z_0) - \Delta(s_1, z_{s_1})] \\
&= E[\Delta(s, z_0) - \Delta(s_1, z_{s_1})](1 - 1_{A^\epsilon}) + E[\Delta(s, z_0) - \Delta(s_1, z_{s_1})]1_{A^\epsilon} \\
&\geq E[\Delta(s, z_0) - \Delta(s_1, z_{s_1})]1_{A^\epsilon} = O(Ee^{2|z_{s_1}^{(1)}|}1_{A^\epsilon}) && (4.14) \\
&= O(Ee^{2|z_{s_1}^{(1)}|}1_{(|z_{s_1}^{(1)} - z_0^{(1)}| \geq \epsilon/2)} && (4.15) \\
&= O(|s - s_1|^2) && (4.16)
\end{aligned}
$$

when $s_1 \downarrow s$. Here we get (4.14) by basic estimates for the change of optimal cost with respect to different initial states, obtain (4.15) by $z_{s_1}^{(2)} \to z_0^{(2)}$ uniformly as $s_1 \downarrow s$, and obtain the bound (4.16) using basic moment estimates for $z_{s_1}^{(1)}$. It follows from

(4.16) that

$$\lim_{s_1 \downarrow s} \frac{1}{s_1 - s} E[\Delta(s, z_0) - \Delta(s_1, z_{s_1})] \geq 0. \tag{4.17}$$

But for $s_1 \in (s, T]$, we also have

$$\begin{aligned} &\frac{1}{s_1 - s} E[\Delta(s, z_0) - \Delta(s_1, z_{s_1})] \leq \frac{1}{s_1 - s} E[\int_s^{s_1} l(z_t)dt - \varphi(s, z_0) + \varphi(s_1, z_{s_1})] \\ &\to l(s, z_0) + \frac{\partial \varphi}{\partial s} + \frac{\partial^\tau \varphi}{\partial z} \psi\big|_u + \frac{1}{2} \mathrm{tr}(\frac{\partial^2 \varphi}{\partial z^2} G G^\tau), \quad \forall u \in U, \end{aligned} \tag{4.18}$$

as $s_1 \downarrow s$, where we get the inequality by the principle of optimality, and obtain the last line by using Ito's formula to express $\varphi(s_1, z_{s_1})$ near (s, z_0) and then taking expectations. In the above since v satisfies the growth condition in Proposition 4.2, $\varphi(t, z) = 0$ for $|z^{(1)} - z_0^{(1)}| \geq C$, all the expectations are finite. Therefore, for $z \in \mathrm{R}^{2n}$, by (4.17) and (4.18)

$$\frac{\partial \varphi}{\partial s} + \min_{u \in U} \{\frac{\partial^\tau \varphi}{\partial z} \psi\} + \frac{1}{2} \mathrm{tr}(\frac{\partial^2 \varphi}{\partial z^2} G G^\tau) + l \geq 0,$$

at (s, z_0). On the other hand, if $v - \varphi$ has a local minimum at (s, z_0), $s < T$, then for any small $\varepsilon > 0$, we can choose sufficiently small $s_1 \in (s, T]$ and find a control $u \in \mathcal{U}$ generating z_u such that

$$\begin{aligned} &E\{v(s, z_0) - \varphi(s, z_0) - v(s_1, z_{s_1}) + \varphi(s_1, z_{s_1})\} \\ \geq &E\{\int_s^{s_1} l(z_t)dt + \varphi(s_1, z_{s_1}) - \varphi(s, z_0)\} - \varepsilon(s_1 - s). \end{aligned} \tag{4.19}$$

Similar to (4.16), we also have

$$E[\Delta(s, z_0) - \Delta(s_1, z_{s_1})] \leq O(|s - s_1|^2),$$

which together with (4.19) and Ito's formula gives

$$\frac{\partial \varphi}{\partial s} + \min_{u \in U} \{\frac{\partial^\tau \varphi}{\partial z} \psi\} + \frac{1}{2} \mathrm{tr}(\frac{\partial^2 \varphi}{\partial z^2} G G^\tau) + l \leq 0,$$

at (s, z_0), so that the value function v is a viscosity solution. □

To analyze uniqueness of the viscosity solution, we introduce the function class $\mathcal{G}$ such that each $W(t,z) \in \mathcal{G}$ satisfies:

(i) $W \in C([0,T] \times \mathrm{R}^{2n})$ and

(ii) there exist $C, k_1, k_2 > 0$ such that $|W| \leq C[1 + \sum_{i=1}^{n} e^{k_1|z_i|} + \sum_{i=1}^{2n} |z_i|^{k_2}]$, where the constants C, k_1, k_2 can vary with each W.

Here we state a general maximum principle in an unbounded domain for the HJB equation (4.13). The proof of the maximum principle is postponed to Section 4.4.

Theorem 4.2. Assuming **(H4.1)** and **(H4.2)** hold, if $\underline{v}$, $\overline{v} \in \mathcal{G}$ are viscosity subsolution and supersolution to (4.13), respectively, and $\sup_{\partial^* Q_0}(\underline{v} - \overline{v}) < \infty$, then

$$\sup_{Q_0}(\underline{v} - \overline{v}) = \sup_{\partial^* Q_0}(\underline{v} - \overline{v}), \tag{4.20}$$

where $Q_0 = [0,T] \times \mathrm{R}^{2n}$, $\partial^* Q_0 = \{(T,z) : \ z \in \mathrm{R}^{2n}\}$. □

Theorem 4.3. Assuming **(H4.1)** and **(H4.2)** hold, there exists a unique viscosity solution to the equation (4.13) in the class $\mathcal{G}$.

PROOF. By considering two possibly distinct viscosity solutions v_1 and v_2 both in $\mathcal{G}$, and setting respectively $(v_1, v_2) = (\underline{v}, \overline{v})$ and $(v_2, v_1) = (\underline{v}, \overline{v})$ in Theorem 4.2, we obtain Theorem 4.3 as a corollary. □

4.3. Semiconvex and Semiconcave Approximations over Compact Sets

To facilitate our analysis, write the Hamiltonian

$$\widetilde{H}(t,z,u,\xi,V) = -\xi^{\tau}\psi(t,z,u) - \frac{1}{2}\mathrm{tr}\{VG(t,z)G^{\tau}(t,z)\} - l(z), \tag{4.21}$$

$$H(t,z,\xi,V) = \sup_{u \in U} \widetilde{H}(t,z,u,\xi,V),$$

where $\xi \in \mathrm{R}^{2n}$, V is a $2n \times 2n$ real symmetric matrix, and the other terms are defined in Section 4.2. Then the HJB equation (4.13) can be written as

$$0 = -v_t + H(t, z, v_z, v_{zz}), \tag{4.22}$$

$$v(T, z) = 0. \tag{4.23}$$

Definition 4.2. [78] A function $\varphi(x)$ defined on a convex set $Q \subset \mathrm{R}^m$ is said to be semiconvex on Q, if there exists a constant $C > 0$ such that $\varphi(x) + C|x|^2$ is convex. $\varphi(x)$ is semiconcave on Q if $-\varphi(x)$ is semiconvex on Q. □

Definition 4.3. A function $\varphi(x)$ defined on a convex set $Q \subset \mathrm{R}^m$ is said to be locally semiconvex on Q, if for any $y \in Q$, there exists a convex neighborhood N_y (relative to Q) of y such that $\varphi(x)$ is semiconvex on N_y. □

Proposition 4.3. If $\varphi(x)$ is locally semiconvex on a convex compact set Q, then $\varphi(x)$ is semiconvex on Q.

PROOF. For any $y \in Q$ there exists a convex N_y open relative to Q such that $y \in N_y$ and $\varphi(x)$ is semiconvex on N_y. So there exists $C_y > 0$ such that $\varphi(x)+C_y|x|^2$ is convex on N_y. Since $\{N_y,\ y \in Q\}$ is an open cover of Q, there exists a finite subcover $\{N_{y_i},\ 1 \le i \le k\}$. Take $C = \max_{1\le i\le k} C_{y_i}$ and then obviously $\varphi(x) + C|x|^2 \triangleq \widehat{\varphi}(x)$ is convex on each N_{y_i}, $1 \le i \le k$. Now for any $x_1,\ x_2 \in Q,\ 0 \le \lambda \le 1$, we prove that $\widehat{\varphi}(\lambda x_1 + (1-\lambda)x_2) \le \lambda\widehat{\varphi}(x_1) + (1-\lambda)\widehat{\varphi}(x_2)$. We only need to consider the case $0 < \lambda < 1$. First, from the collection $\{N_{y_i},\ 1 \le i \le k\}$ we select open sets, without loss of generality, denoted as $\mathcal{N} \triangleq \{N_{y_i},\ i = 1, \cdots, m \le k\}$ such that $L \triangleq \{x :\ x = \lambda x_1+(1-\lambda)x_2,\ 0 \le \lambda \le 1\} \subset \cup_{N_{y_i}\in\mathcal{N}} N_{y_i}$. For simplicity we consider the case $m = 2$ and $x_1 \in N_{y_1},\ x_2 \in N_{y_2}$. The general case can be treated inductively. To avoid triviality, we assume neither N_{y_1} nor N_{y_2} covers L individually, and then we can find $x_a \in L,\ x_a \neq x_\lambda$ such that $x_a \in N_{y_1} \cap N_{y_2}$ and $x_a = c_1x_1+(1-c_1)x_2,\ 0 < c_1 < 1$. Without loss of generality we assume x_λ is between x_1 and x_a. Then we further choose $x_b \in N_{y_1} \cap N_{y_2}$ such that $x_b = c_2x_1 + (1-c_2)x_2$ and x_b is between x_a and x_2. Now it is obvious that $0 < c_2 < c_1 < \lambda < 1$. It is straightforward to verify that

$$x_\lambda = \frac{\lambda - c_1}{1 - c_1} x_1 + \frac{1-\lambda}{1-c_1} x_a, \quad x_a = \frac{c_1 - c_2}{\lambda - c_2} x_\lambda + \frac{\lambda - c_1}{\lambda - c_2} x_b, \quad x_b = \frac{c_2}{c_1} x_a + \frac{c_1 - c_2}{c_1} x_2.$$

Hence we have

$$\begin{aligned}
\widehat{\varphi}(x_\lambda) &\le \frac{\lambda - c_1}{1 - c_1} \widehat{\varphi}(x_1) + \frac{1-\lambda}{1-c_1} \widehat{\varphi}(x_a), \\
\widehat{\varphi}(x_a) &\le \frac{c_1 - c_2}{\lambda - c_2} \widehat{\varphi}(x_\lambda) + \frac{\lambda - c_1}{\lambda - c_2} \widehat{\varphi}(x_b), \\
\widehat{\varphi}(x_b) &\le \frac{c_2}{c_1} \widehat{\varphi}(x_a) + \frac{c_1 - c_2}{c_1} \widehat{\varphi}(x_2),
\end{aligned}$$

where we get the first two inequalities and the last one by the local convexity of $\widehat{\varphi}(x)$ on N_{y_1} and N_{y_2}, respectively. By a simple transformation with the above inequalities to eliminate $\widehat{\varphi}(x_a)$ and $\widehat{\varphi}(x_b)$ we obtain

$$\widehat{\varphi}(x_\lambda) \le \lambda \widehat{\varphi}(x_1) + (1-\lambda)\widehat{\varphi}(x_2).$$

By arbitrariness of x_1, x_2 in Q it follows that $\widehat{\varphi}(x)$ is convex on Q. This completes the proof. □

We adopt the semiconvex/semiconcave approximation technique of [**78, 20, 42, 44, 45**], but due to the highly nonlinear growth condition of the class $\mathcal{G}$, we apply a particular localized technique to construct envelope functions to generate semiconvex/semiconcave approximations on any bounded domain. For any $W \in \mathcal{G}$, define the upper/lower envelope functions with $\eta \in (0, 1]$,

$$W^\eta(t,z) = \sup_{(s,w)\in B^\eta(t,z)} \{W(s,w) - \frac{1}{2\eta^2}(|t-s|^2 + |z-w|^2)\}, \tag{4.24}$$

$$W_\eta(t,z) = \inf_{(s,w)\in B^\eta(t,z)} \{W(s,w) + \frac{1}{2\eta^2}(|t-s|^2 + |z-w|^2)\}, \tag{4.25}$$

where $B^\eta(t,z)$ denotes the closed ball (relative to $[0,T] \times \mathrm{R}^{2n}$) centering (t,z) with radius η. As will be shown in the following lemma, our construction above will generate semiconvex/semiconcave approximations to a given continuous function on a compact set for small η.

Lemma 4.1. For any fixed $W \in \mathcal{G}$ and compact convex set $Q \subset [0,T] \times \mathrm{R}^{2n}$, there exists a constant $\eta_Q \leq 1$ depending only on Q so that for all $\eta \leq \eta_Q$, $W^\eta(t,z)$ is semiconvex on Q and $W_\eta(t,z)$ is semiconcave on Q.

PROOF. Since any fixed $W \in \mathcal{G}$ is uniformly continuous and bounded on any compact set Q, there exists $\eta_Q > 0$ depending only on Q, so that for all $\eta \leq \eta_Q$ and $(t,z) \in Q$,

$$W^\eta(t,z) = \sup_{(s,w)\in B^{\eta/2}(t,z)} \{W(s,w) - \frac{1}{2\eta^2}[|t-s|^2 + |z-w|^2]\}. \tag{4.26}$$

Indeed, we can find $\eta_Q > 0$ such that for all $\eta \leq \eta_Q$, $|W(s,w) - W(t,z)| \leq \frac{1}{16}$ for $(s,w) \in B^\eta(t,z)$, where $(t,z) \in Q$. Then for any (s,w) satisfying $\frac{\eta^2}{4} \leq |s-t|^2 + |w-z|^2 \leq \eta^2$, we have

$$W(s,w) - \frac{1}{2\eta^2}(|s-t|^2 + |w-z|^2) \leq W(t,z) + \frac{1}{16} - \frac{1}{2\eta^2}\frac{\eta^2}{4} < W(t,z),$$

and (4.26) follows. In the following we assume $\eta \leq \eta_Q$. Next we show that for any $(t_0,z_0) \in Q$, $W^\eta(t,z)$ is semiconvex on $B^{\eta/4}(t_0,z_0)$. It suffices to show that $W^\eta(t,z) + \frac{1}{2\eta^2}(t^2 + |z|^2)$ is convex on $B^{\eta/4}(t_0,z_0)$. Denote

$$R(s,w,t,z) = W(s,w) - \frac{1}{2\eta^2}(|t-s|^2 + |z-w|^2) + \frac{1}{2\eta^2}(t^2 + |z|^2).$$

If $(t_1,z_1), (t_2,z_2) \in B^{\eta/4}(t_0,z_0)$, we have $(t_2,z_2) \in B^{\eta/2}(t_1,z_1)$. For any $\lambda \in [0,1]$, denote $(t_\lambda, z_\lambda) = (\lambda t_1 + (1-\lambda)t_2, \lambda z_1 + (1-\lambda)z_2)$. It is obvious that $B^{\eta/2}(t_\lambda,z_\lambda) \subset B^\eta(t_1,z_1) \cap B^\eta(t_2,z_2)$. Then it follows

$$\begin{aligned}
& W^\eta(t_\lambda,z_\lambda) + \frac{1}{2\eta^2}[t_\lambda^2 + |z_\lambda|^2] \\
= & \sup_{(s,w)\in B^\eta(t_\lambda,z_\lambda)} R(s,w,t_\lambda,z_\lambda) = \sup_{(s,w)\in B^{\eta/2}(t_\lambda,z_\lambda)} R(s,w,t_\lambda,z_\lambda) \\
= & \sup_{(s,w)\in B^{\eta/2}(t_\lambda,z_\lambda)} [\lambda R(s,w,t_1,z_1) + (1-\lambda)R(s,w,t_2,z_2)] \\
\leq & \sup_{(s,w)\in B^{\eta/2}(t_\lambda,z_\lambda)} \lambda R(s,w,t_1,z_1) + \sup_{(s,w)\in B^{\eta/2}(t_\lambda,z_\lambda)} (1-\lambda)R(s,w,t_2,z_2)
\end{aligned}$$

$$\leq \sup_{(s,w)\in B^{\eta}(t_1,z_1)} \lambda R(s,w,t_1,z_1) + \sup_{(s,w)\in B^{\eta}(t_2,z_2)} (1-\lambda)R(s,w,t_2,z_2)$$
$$=\lambda[W^{\eta}(t_1,z_1) + \frac{1}{2\eta^2}(t_1^2+|z_1|^2)] + (1-\lambda)[W^{\eta}(t_2,z_2) + \frac{1}{2\eta^2}(t_2^2+|z_2|^2)].$$

So that $W^{\eta}(t,z)$ is semiconvex on $B^{\eta/4}(t_0,z_0)$. And by Proposition 4.3, $W^{\eta}(t,z)$ is semiconvex on Q. Similarly we can prove $W_{\eta}(t,z)$ is semiconcave on Q. □

We use an example to illustrate the construction of the semiconvex approximation to a given function.

Example 4.1. Consider a continuous function W: R → R defined as follows:

$$W(x) = \begin{cases} (x-1)^3+1 & \text{for } x \leq 0, \\ -(x+1)^3+1 & \text{for } x > 0. \end{cases}$$

We take $0 < \eta \leq 0.125$ and write

$$\theta(x) = 1 - x + \frac{1}{6\eta^2} - \sqrt{[1-x+\frac{1}{6\eta^2}]^2 - (1-x)^2}, \qquad x \leq 0.$$

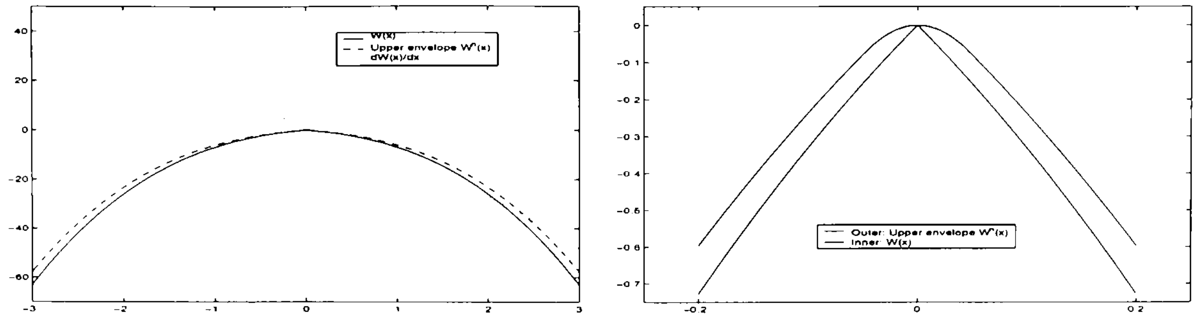
It is evident that the upper envelope function $W^{\eta}(x)$ is even on R and its value on $(-\infty, 0]$ is determined by

$$W^{\eta}(x) = \begin{cases} W(x+\eta) - \frac{1}{2} & \text{for } x \leq 1 - \eta - \frac{1}{\sqrt{3\eta}}, \\ W(x+\theta(x)) - \frac{\theta^2(x)}{2\eta^2} & \text{for } 1-\eta-\frac{1}{\sqrt{3\eta}} < x \leq -3\eta^2, \\ W(0) - \frac{x^2}{2\eta^2} & \text{for } -3\eta^2 < x \leq 0. \end{cases} \tag{4.27}$$

□

From Figure 4.1 it is seen that at $x = 0$ the first order derivative of $W(x)$ has a negative jump, which corresponds to a sharp turn at $x = 0$ on the function curve. After the semiconvexifying procedure, the sharp turn at $x = 0$ vanishes as shown by the curve of $W^{\eta}(x)$.

We give a lemma which is parallel to the one in [**78**]. But here we do not make Lipschitz or Hölder type continuity assumptions on W. For completeness we give the details.



(a) (b)

FIGURE 4.1. Semiconvex approximation with $\eta = 0.125$, (a) The curves in a large range, (b) The curves in the local region

Lemma 4.2. For $W \in \mathcal{G}$ and $\eta \in (0, 1]$, W^η and W_η are equicontinuous (w.r.t. η) on any compact set $Q \subset [0, T] \times \mathrm{R}^{2n}$ and

$$W^\eta(t, z) \leq C[1 + \sum_{i=1}^{n} e^{k_1|z_i|} + \sum_{i=1}^{2n} |z_i|^{k_2}], \tag{4.28}$$

$$W^\eta(t, z) = W(t_0, z_0) - \frac{1}{2\eta^2}(|t - t_0|^2 + |z - z_0|^2), \quad \text{for some } (t_0, z_0) \in B^\eta(t, z), \tag{4.29}$$

$$\frac{1}{2\eta^2}(|t - t_0|^2 + |z - z_0|^2) \to 0 \quad \text{uniformly on } Q, \text{ as } \eta \to 0, \text{ and} \tag{4.30}$$

$$0 \leq W^\eta(t, z) - W(t, z) \to 0 \quad \text{uniformly on } Q, \text{ as } \eta \to 0, \tag{4.31}$$

where C is a constant independent of η. (4.28)-(4.30) also hold when W^η is replaced by W_η, and

$$0 \leq W(t, z) - W_\eta(t, z) \to 0 \quad \text{uniformly on } Q, \text{ as } \eta \to 0. \tag{4.32}$$

PROOF. (4.28) follows from the definition of $\mathcal{G}$, and (4.29) is obvious. Moreover, by (4.29) we have

$$\frac{1}{2\eta^2}(|t - t_0|^2 + |z - z_0|^2) = W(t_0, z_0) - W^\eta(t, z) \leq W(t_0, z_0) - W(t, z). \tag{4.33}$$

Since $|t - t_0| + |z - z_0| \to 0$ as $\eta \to 0$, by (4.33) and the uniform continuity of W on Q, (4.30) follows. (4.31) follows from (4.29) and (4.30). The equicontinuity of W^η

(w.r.t. η) on Q can be established by (4.31) and the continuous dependence of W^η on $(\eta, t, z) \in [\varepsilon, 1] \times Q$ for any $0 < \varepsilon \le 1$. The case of W_η can be treated similarly. □

We define

$$H^\eta(t, z, \xi, V) = \inf_{(s,w)\in B^\eta(t,z)} \sup_{u\in U} \widetilde{H}(s, w, u, \xi, V), \tag{4.34}$$

$$H_\eta(t, z, \xi, V) = \sup_{(s,w)\in B^\eta(t,z)} \sup_{u\in U} \widetilde{H}(s, w, u, \xi, V). \tag{4.35}$$

Then it can be shown that H^η and H_η converge to $H(t, z, \xi, V)$ uniformly on any compact subset of $[0, T] \times \mathrm{R}^{2n} \times \mathrm{R}^{2n} \times S^{2n}$ as $\eta \to 0$, where S^{2n} denotes the set of $2n \times 2n$ real symmetric matrices. The following lemma can be proved by a similar method as in [**24, 42, 45**]. The proof is omitted here. Notice that the viscosity sub/supersolution properties hold on a domain smaller than $[0, T] \times \mathrm{R}^{2n}$.

Lemma 4.3. If $\underline{v}$ ($\overline{v}$, respectively) is a viscosity subsolution (supersolution, respectively) to (4.22) on $[0, T] \times \mathrm{R}^{2n}$, then $\underline{v}^\eta$ ($\overline{v}_\eta$, respectively) is a viscosity subsolution (supersolution, respectively) to HJB equation A (B, respectively) on $[0, T - \eta] \times \mathrm{R}^{2n}$, where the HJB equations A and B are given by

$$A : \begin{cases} -v_t + H^\eta(t, z, v_z, v_{zz}) = 0, \\ v(T - \eta, z) = \underline{v}^\eta(T - \eta, z), \end{cases} \qquad B : \begin{cases} -v_t + H_\eta(t, z, v_z, v_{zz}) = 0, \\ v(T - \eta, z) = \overline{v}_\eta(T - \eta, z). \end{cases}$$

In the above $\underline{v}^\eta$ and $\overline{v}_\eta$ are defined by (4.24)-(4.25). □

4.4. Proof of the Maximum Principle

In the Section we give a proof of Theorem 4.2. We note that certain technical but standard arguments are not included here for reasons of economy of exposition; complete references to the detailed versions of these parts of the proof are supplied at appropriate places in the proof.

We follow the method in [**78, 24**] employing the particular structure of the system dynamics and will make necessary modifications. For the viscosity subsolution and supersolution $\underline{v}$, $\overline{v} \in \mathcal{G}$, we prove that

$$\sup_{Q_1}(\underline{v} - \overline{v}) = \sup_{\partial^* Q_0}(\underline{v} - \overline{v}) \triangleq c_0 \quad \text{for } Q_1 = [T_1, T] \times \mathrm{R}^{2n}, \tag{4.36}$$

where $T_1 = T - \frac{1}{4\Delta}$, $\Delta = 25n(C_g + C_\sigma) + 10C_{f^0}$, C_g is a finite constant such that $|g_i(t, p, u)| \leq C_g(1 + \sum_{k=1}^n |p_k|)$ for $t \in [0, T]$, $p \in \mathrm{R}^n$, $u \in U$, $1 \leq i \leq n$, for g given in (4.2), and C_σ, C_{f^0} are given in Assumptions **(H4.1)-(H4.2)** in Section 4.2. The maximum principle (4.2) follows by repeating the above procedure backward with time. Our proof by contradiction starts with observation that if (4.36) is not true, there exists $(\hat{t}, \hat{z}) \in (T_1, T) \times \mathrm{R}^{2n}$ such that

$$\underline{v}(\hat{t}, \hat{z}) - \overline{v}(\hat{t}, \hat{z}) = c_0^+ > c_0. \tag{4.37}$$

We break the proof into several steps: (1) we construct a comparison function Λ depending on positive parameters $\alpha, \beta, \varepsilon, \lambda$, and based upon (4.37), Λ is used to induce a certain interior maximum, (2) using the viscosity sub/supersolutions conditions, we get a set of inequalities at the interior maximum, and (3) we establish an inequality relation between α and β by taking appropriate vanishing subsequences of $\varepsilon, \lambda, \eta$, and this inequality relation is shown to lead to a contradiction. The weak coupling condition is used to obtain estimates used in Step 3 below.

Step 1: Constructing a comparison function and the interior maximum. To avoid introducing too many constants, we assume $\underline{v}$ and $\overline{v}$ belong to the class $\mathcal{G}$ with associated constants $k_1 = k_2 = 4$. The more general case can be treated in exactly the same way. Now we define the comparison function

$$\begin{aligned}\Lambda(t, z, s, w) =& \frac{\alpha(2\mu T - t - s)}{2\mu T}\{\sum_{i=1}^{n}[e^{5\sqrt{z_i^2+1}} + e^{5\sqrt{w_i^2+1}}] + \sum_{i=1}^{2n}(z_i^6 + w_i^6)\} - \beta(t + s)\\ &+ \frac{1}{2\varepsilon}|t - s|^2 + \frac{1}{2\varepsilon}|z - w|^2 + \frac{\lambda}{t - T_1} + \frac{\lambda}{s - T_1},\end{aligned}$$

where $\alpha, \beta, \varepsilon, \lambda \in (0, 1]$, $\mu = 1 + \frac{1}{4T\Delta}$, $z, w \in \mathrm{R}^{2n}$ and $t, s \in (T_1, T]$. We write $\Phi(t, z, s, w) = \underline{v}^\eta(t, z) - \overline{v}_\eta(s, w) - \Lambda(t, z, s, w)$, where $\underline{v}^\eta$ and $\overline{v}_\eta$ are also in $\mathcal{G}$ by Lemma 4.2. Noticing that $\Phi \to -\infty$ as $t \wedge s \to T_1$ or $|z| + |w| \to \infty$, there exists (t_0, z_0, s_0, w_0)

such that $\Phi(t_0, z_0, s_0, w_0) = \sup_{Q_1\times Q_1} \Phi(t, z, s, w)$. By $\Phi(t_0, z_0, s_0, w_0) \geq \Phi(T, 0, T, 0)$, one can find a constant C_α depending only on α such that (see Remark 4.2)

$$|z_0| + |w_0| + \frac{1}{2\varepsilon}|t_0 - s_0|^2 + \frac{1}{2\varepsilon}|z_0 - w_0|^2 \leq C_\alpha \quad \text{and} \quad t_0, s_0 \in [T_1 + \frac{\lambda}{C_\alpha}, T]. \quad (4.38)$$

Combining $2\Phi(t_0, z_0, s_0, w_0) \geq \Phi(t_0, z_0, t_0, z_0) + \Phi(s_0, w_0, s_0, w_0)$, (4.38) and Lemma 4.2 we get (see Remark 4.3)

$$\frac{1}{2\varepsilon}|t_0 - s_0|^2 + \frac{1}{2\varepsilon}|z_0 - w_0|^2 \to 0, \quad \text{as } \varepsilon \to 0. \quad (4.39)$$

In this Section, we take $\beta \in (0, \frac{c_0^+ - c_0}{4T})$. We further show that there exists $\alpha_0 > 0$ such that for $\alpha < \alpha_0$ and for sufficiently small r_0 (which may depend upon α) and $\eta \leq r_0, \varepsilon \leq r_0, \lambda \leq r_0$, the maximum of Φ on Q_1 is attained at an interior point (t_0, z_0, s_0, w_0) of the set

$$Q_\alpha = \{(t, z, s, w): \; T_1 + \frac{\lambda}{2C_\alpha} \leq t, s \leq T - \eta, \text{ and } |z|, |w| \leq 2C_\alpha\}, \quad (4.40)$$

where C_α is determined in (4.38).

We begin by observing that $\Phi(t_0, z_0, s_0, w_0) \geq \Phi(\hat{t}, \hat{z}, \hat{t}, \hat{z})$ yields

$$\begin{aligned} & \underline{v}^\eta(\hat{t}, \hat{z}) - \overline{v}_\eta(\hat{t}, \hat{z}) \leq \underline{v}^\eta(t_0, z_0) - \overline{v}_\eta(s_0, w_0) - \Lambda(t_0, z_0, s_0, w_0) + \Lambda(\hat{t}, \hat{z}, \hat{t}, \hat{z}) \\ \leq \; & \underline{v}^\eta(t_0, z_0) - \overline{v}_\eta(s_0, w_0) + 2\beta T + \frac{2\lambda}{\hat{t} - T_1} + 2\alpha[\sum_{i=1}^{n} e^{5\sqrt{\hat{z}_i^2+1}} + \sum_{i=1}^{2n} \hat{z}_i^6]. \quad (4.41) \end{aligned}$$

Let $\mathbb{H}^\beta$ stand for the assertion that there exists α_0 such that when $\alpha \leq \alpha_0$ and $\max\{\eta, \varepsilon, \lambda\} \leq r_0$ for sufficiently small r_0, (t_0, z_0, s_0, w_0) is an interior point of Q_α in (4.40).

If $\mathbb{H}^\beta$ is not true, then there exists an arbitrarily small $\alpha \in (0, 1]$ such that for this fixed α we can select $\eta^{(k)}, \varepsilon^{(k)}, \lambda^{(k)} \to 0$ for which the resulting $(t_0^{(k)}, z_0^{(k)}, s_0^{(k)}, w_0^{(k)}) \notin \text{Int}(Q_\alpha)$. By (4.38) it necessarily follows that $t_0^{(k)} \vee s_0^{(k)} \geq T - \eta^{(k)} \to T$ and (4.39) gives $|t_0^{(k)} - s_0^{(k)}| + |s_0^{(k)} - w_0^{(k)}| \to 0$. It is also clear that $(t_0^{(k)}, z_0^{(k)}, s_0^{(k)}, w_0^{(k)})$ is contained in a compact set determined by α. Then by selecting an appropriate subsequence of

$(t_0^{(k)}, z_0^{(k)}, s_0^{(k)}, w_0^{(k)})$ and taking the limit in (4.41) along this subsequence, we get

$$\underline{v}(\hat{t},\hat{z}) - \overline{v}(\hat{t},\hat{z}) \leq \underline{v}(T, z^\alpha) - \overline{v}(T, z^\alpha) + \frac{c_0^+ - c_0}{2} + 2\alpha[\sum_{i=1}^{n} e^{5\sqrt{\hat{z}_i^2+1}} + \sum_{i=1}^{2n} \hat{z}_i^6]$$

$$\leq \frac{c_0^+ + c_0}{2} + 2\alpha[\sum_{i=1}^{n} e^{5\sqrt{\hat{z}_i^2+1}} + \sum_{i=1}^{2n} \hat{z}_i^6], \tag{4.42}$$

where z^α denotes the common limit of the selected subsequences of $z_0^{(k)}$ and $w_0^{(k)}$. Sending $\alpha \to 0$, we get $\underline{v}(\hat{t},\hat{z}) - \overline{v}(\hat{t},\hat{z}) < c_0^+$, which contradicts (4.37), hence $\mathbb{H}^\beta$ holds. From the argument leading to (4.42) it is seen that α_0 can be chosen independently of β.

Step 2: Applying Ishii's Lemma. Hereafter we assume $\beta < \frac{c_0^+ - c_0}{4T}$, $\alpha < \alpha_0$ and $\max\{\eta, \varepsilon, \lambda\} \leq r_0$ are always satisfied and thus $\mathbb{H}^\beta$ holds. We assume Φ attains a strict maximum at (t_0, z_0, s_0, w_0); otherwise we replace Λ by $\Lambda + |t - t_0|^2 + |s - s_0|^2 + |z - z_0|^4 + |w - w_0|^4$. Following the derivations in [**78, 42, 24**], and using the interior maximum obtained in Step 1, the semiconvexity of $\underline{v}^\eta$, and the semiconcavity of $\overline{v}_\eta$ for $\eta \leq \eta_{Q_\alpha}$ by Lemma 4.1, and by Lemma 4.3, we obtain the so-called Ishii's lemma, i.e., there exist $2n \times 2n$ symmetric matrices M_k, $k = 1, 2$ such that

$$-\Lambda_t(t_0, z_0, s_0, w_0) + H^\eta(t_0, z_0, \Lambda_z(t_0, z_0, s_0, w_0), M_1) \leq 0, \tag{4.43}$$

$$\Lambda_s(t_0, z_0, s_0, w_0) + H_\eta(s_0, w_0, -\Lambda_w(t_0, z_0, s_0, w_0), M_2) \geq 0, \tag{4.44}$$

$$\begin{pmatrix} M_1 & 0 \\ 0 & -M_2 \end{pmatrix} \leq \begin{pmatrix} \Lambda_{zz} & \Lambda_{zw} \\ \Lambda_{zw}^\tau & \Lambda_{ww} \end{pmatrix} \Bigg|_{(t_0, z_0, s_0, w_0)}, \tag{4.45}$$

We note that it is important to have $t_0 \vee s_0 < T - \eta$ in order to establish (4.43)-(4.44) by Lemma 4.3 and an approximation procedure (see e.g. [**24**] for the case of a bounded domain). Now (4.43) and (4.44) yield

$$\begin{aligned} &-\Lambda_t(t_0, z_0, s_0, w_0) - \Lambda_s(t_0, z_0, s_0, w_0) \\ \leq &H_\eta(s_0, w_0, -\Lambda_w(t_0, z_0, s_0, w_0), M_2) - H^\eta(t_0, z_0, \Lambda_z(t_0, z_0, s_0, w_0), M_1). \end{aligned} \tag{4.46}$$

Step 3: Estimates for LHS and RHS of (4.46). The final stage in our deduction of a contradiction from (4.37) involves estimates of the LHS and RHS of (4.46). The estimates for both sides of (4.46) are taken at (t_0, z_0, s_0, w_0), but for brevity we omit the subscript 0 for each variable. We have

$$\begin{aligned}\text{LHS of (4.46)} =& \frac{\alpha}{\mu T}[\sum_{i=1}^{n}(e^{5\sqrt{z_i^2+1}} + e^{5\sqrt{w_i^2+1}}) + \sum_{i=1}^{n}(z_i^6 + w_i^6)] \\ & + 2\beta + \frac{\lambda}{(t-T_1)^2} + \frac{\lambda}{(s-T_1)^2} \\ \geq & \frac{\alpha}{\mu T}[\sum_{i=1}^{n}(e^{5\sqrt{z_i^2+1}} + e^{5\sqrt{w_i^2+1}}) + \sum_{i=1}^{n}(z_i^6 + w_i^6)] + 2\beta, \end{aligned} \tag{4.47}$$

and

$$\begin{aligned}\text{RHS of (4.46)} =& \sup_{u\in U}[\Lambda_w^\tau(t,z,s,w)\psi(\widehat{s},\widehat{w},u)] - \sup_{u\in U}[-\Lambda_z^\tau(t,z,s,w)\psi(\widehat{t},\widehat{z},u)] \\ & + \frac{1}{2}\text{tr}[G(\widehat{t},\widehat{z})G^\tau(\widehat{t},\widehat{z})M_1] - \frac{1}{2}\text{tr}[G(\widehat{s},\widehat{w})G^\tau(\widehat{s},\widehat{w})M_2] + l(\widehat{z}) - l(\widehat{w}) \\ \leq & \sup_{u\in U}[\Lambda_w^\tau(t,z,s,w)\psi(\widehat{s},\widehat{w},u) + \Lambda_z^\tau(t,z,s,w)\psi(\widehat{t},\widehat{z},u)] \\ & + \frac{1}{2}\text{tr}[G(\widehat{t},\widehat{z})G^\tau(\widehat{t},\widehat{z})M_1] - \frac{1}{2}\text{tr}[G(\widehat{s},\widehat{w})G^\tau(\widehat{s},\widehat{w})M_2] - l(\widehat{z}) - l(\widehat{w}),\end{aligned}$$

which together with (4.45), (4.34)-(4.35) leads to

$$\begin{aligned}\text{RHS of (4.46)} \leq & \sup_{u\in U}[\Lambda_w^\tau(t,z,s,w)\psi(\widehat{s},\widehat{w},u) + \Lambda_z^\tau(t,z,s,w)\psi(\widehat{t},\widehat{z},u)] \quad (\triangleq A_1) \\ & + \frac{1}{2\varepsilon}\text{tr}\{[G(\widehat{t},\widehat{z}) - G(\widehat{s},\widehat{w})]^\tau[G(\widehat{t},\widehat{z}) - G(\widehat{s},\widehat{w})]\} \quad (\triangleq A_2) \\ & + \frac{\alpha(2\mu T - t - s)}{2\mu T}\sum_{i,k=1}^{n}\frac{1}{2}[\sigma_{ik}^2(\widehat{t},\widehat{z})(\Gamma''(z_i) + 30z_i^4) + \sigma_{ik}^2(\widehat{s},\widehat{w})(\Gamma''(w_i) + 30w_i^4)] \quad (\triangleq A_3) \\ & + [l(\widehat{z}) - l(\widehat{w})] \quad (\triangleq A_4) \\ = & A_1 + A_2 + A_3 + A_4,\end{aligned} \tag{4.48}$$

where $\Gamma(r) \triangleq e^{5\sqrt{r^2+1}}$, $\Gamma'' = \frac{d^2\Gamma}{dr^2}$ and $(\widehat{t}_0, \widehat{z}_0) \in B^\eta(t_0, z_0)$, $(\widehat{s}_0, \widehat{w}_0) \in B^\eta(s_0, w_0)$. Notice that the set $\mathcal{S}_{\eta,\varepsilon} = \{(t_0, z_0), (\widehat{t}_0, \widehat{z}_0), (s_0, w_0), (\widehat{s}_0, \widehat{w}_0)\}$ is contained in a compact set Q^*_α determined by α. For $0 < \varepsilon \le 1$ appearing in $\Lambda(t,z,s,w)$, there exists $\eta_\varepsilon > 0$ such that for all $0 < \eta \le \eta_\varepsilon$,

$$\text{RHS of (4.46)} \le A_1^0 + A_2^0 + A_3^0 + A_4^0 + \varepsilon, \tag{4.49}$$

where, without writing the subscript 0 for (t_0, z_0, s_0, w_0), we denote

$$\begin{aligned}
A_1^0 &= \sup_{u\in U}[\Lambda_w^\tau(t,z,s,w)\psi(s,w,u) + \Lambda_z^\tau(t,z,s,w)\psi(t,z,u)],\\
A_2^0 &= \frac{1}{2\varepsilon}\text{tr}\{[G(t,z)-G(s,w)]^\tau[G(t,z)-G(s,w)]\},\\
A_3^0 &= \frac{\alpha(2\mu T - t - s)}{2\mu T}\sum_{i,k=1}^{n}\frac{1}{2}[\sigma_{ik}^2(t,z)(\Gamma''(z_i)+30z_i^4) + \sigma_{ik}^2(s,w)(\Gamma''(w_i)+30w_i^4)],\\
A_4^0 &= l(z) - l(w).
\end{aligned}$$

Since $\mathcal{S}_{\eta,\varepsilon}$ is contained in Q^*_α and the diameter of $\mathcal{S}_{\eta,\varepsilon}$ tends to 0 as $\eta, \varepsilon \to 0$, by taking an appropriate sequence $(\eta^{(k)}, \varepsilon^{(k)}, \lambda^{(k)}) \to 0$ satisfying $\eta^{(k)} \le \eta_{\varepsilon^{(k)}}$, we get a convergent sequence $(t_0^{(k)}, z_0^{(k)})$, $(t_0^{(k)}, \widehat{z}_0^{(k)})$, $(s_0^{(k)}, w_0^{(k)})$, $(s_0^{(k)}, \widehat{w}_0^{(k)}) \to (\tilde{t}, \tilde{z})$, as $k \to \infty$. In the following we use the same C to denote different constants which are independent of α. We have the three relations

$$\limsup_{k\to\infty} \text{ LHS of (4.46) } (\eta^{(k)}, \varepsilon^{(k)}, \lambda^{(k)}) \ge \frac{2\alpha}{\mu T}[\sum_{i=1}^{n} e^{5\sqrt{\tilde{z}_i^2+1}} + \sum_{i=1}^{2n}|\tilde{z}_i|^6] + 2\beta, \tag{4.50}$$

$$\lim_{k\to\infty} (A_2^0 + A_4^0)\ (\eta^{(k)}, \varepsilon^{(k)}, \lambda^{(k)}) = 0, \tag{4.51}$$

$$\limsup_{k\to\infty} A_3^0\ (\eta^{(k)}, \varepsilon^{(k)}, \lambda^{(k)}) \le \frac{n\alpha C_\sigma(\mu T - \tilde{t})}{\mu T}\sum_{i=1}^{n}(25e^{5\sqrt{\tilde{z}_i^2+1}} + 30|\tilde{z}_i|^4), \tag{4.52}$$

where (4.50) follows from (4.47), and (4.51) follows from continuity of $l(z)$, Lipschitz continuity of $G(t,z)$ by **(H4.1)**, and (4.39). Now we analyze A_1^0.

$$\begin{aligned}
A_1^0 \le &\sup_{u\in U}\sum_{i=n+1}^{2n}[\Lambda_{z_i}(t,z,s,w)\psi_i(t,z,u) + \Lambda_{w_i}(t,z,s,w)\psi_i(s,w,u)]\\
&+\sum_{i=1}^{n}[\Lambda_{z_i}(t,z,s,w)f_i(t,z) + \Lambda_{w_i}(t,z,s,w)f_i(s,w)] \triangleq A_{11}^0 + A_{12}^0.
\end{aligned}$$

Then it follows that

$$\limsup_{k\to\infty} A_{11}^0(\eta^{(k)}, \varepsilon^{(k)}, \lambda^{(k)}) \le \frac{\alpha(\mu T - \tilde{t})}{\mu T} \sum_{i=n+1}^{2n} 12C_g[2n|\tilde{z}_i|^6 + |\tilde{z}_i|^5]. \tag{4.53}$$

We employ $a_i(t) \ge 0$ for $t \in [0, T]$ in the weak coupling condition **(H4.2)**, and the Lipschitz continuity property of $f_i(t, z) = a_i(t)z_i + f_i^0(t, z)$ by **(H4.1)** to obtain

$$\begin{aligned} A_{12}^0 = & \frac{\alpha(2\mu T - t - s)}{2\mu T} \sum_{i=1}^{n} \{[\tfrac{5z_i}{\sqrt{z_i^2+1}} e^{5\sqrt{z_i^2+1}} + 6z_i^5 + \frac{z_i - w_i}{\varepsilon}][-a_i(t)z_i + f_i^0(t, z)] \\ & + [\tfrac{5w_i}{\sqrt{w_i^2+1}} e^{5\sqrt{w_i^2+1}} + 6w_i^5 + \frac{w_i - z_i}{\varepsilon}][-a_i(s)w_i + f_i^0(s, w)]\} \\ \le & \frac{\alpha(2\mu T - t - s)}{2\mu T} \sum_{i=1}^{n} \{[\tfrac{5z_i}{\sqrt{z_i^2+1}} e^{5\sqrt{z_i^2+1}} + 6z_i^5] f_i^0(t, z) \\ & + [\tfrac{5w_i}{\sqrt{w_i^2+1}} e^{5\sqrt{w_i^2+1}} + 6w_i^5] f_i^0(s, w)\} + O(\frac{|t-s|^2}{\varepsilon} + \frac{|z-w|^2}{\varepsilon}). \end{aligned} \tag{4.54}$$

Hence invoking (4.39), it follows that

$$\limsup_{k\to\infty} A_{12}^0(\eta^{(k)}, \varepsilon^{(k)}, \lambda^{(k)}) \le \frac{\alpha C_{f^0}(\mu T - \tilde{t})}{\mu T} \sum_{i=1}^{n} [10 e^{5\sqrt{\tilde{z}_i^2+1}} + 12|\tilde{z}_i|^5], \tag{4.55}$$

which together with (4.51)-(4.53) gives

$$\begin{aligned} & \limsup_{k\to\infty} \text{ RHS of (4.46) } (\eta^{(k)}, \varepsilon^{(k)}, \lambda^{(k)}) \\ \le & \frac{[10C_{f^0} + 25n(C_\sigma + C_g)]\alpha(\mu T - \tilde{t})}{\mu T} [\sum_{i=1}^{n} e^{5\sqrt{\tilde{z}_i^2+1}} + \sum_{i=1}^{2n} |\tilde{z}_i|^6 + C] \\ \le & \frac{\alpha}{2\mu T} [\sum_{i=1}^{n} e^{5\sqrt{\tilde{z}_i^2+1}} + \sum_{i=1}^{2n} |\tilde{z}_i|^6 + C]. \end{aligned} \tag{4.56}$$

Hence it follows from (4.46), (4.50) and (4.56) that

$$2\beta \le -\frac{3\alpha}{2\mu T} \{\sum_{i=1}^{n} e^{5\sqrt{\tilde{z}_i^2+1}} + \sum_{i=1}^{2n} |\tilde{z}_i|^6\} + \alpha C \le \alpha C. \tag{4.57}$$

We recall from Step 1 that $\beta \leq 1$ can take a strictly positive value from the interval $(0, \frac{c_0^+ - c_0}{4T})$ and $\alpha \in (0, \alpha_0)$. Letting $\alpha \to 0$ in (4.57) yields $\beta \leq 0$ which is a contradiction to $\beta \in (0, \frac{c_0^+ - c_0}{4T})$, and this completes the proof. □

Remark 4.2. By $\Phi(t_0, z_0, s_0, w_0) \geq \Phi(T, 0, T, 0)$ and $|\underline{v} - \overline{v}| = o([\sum_{i=1}^n (e^{5|z_i|} + e^{5|w_i|}) + \sum_{i=1}^{2n}(z_i^6 + w_i^6)])$, there exists $\delta_\alpha > 0,\ C > 0$ such that

$$\frac{1}{2\varepsilon}|t_0 - s_0|^2 + \frac{1}{2\varepsilon}|z_0 - w_0|^2 + \frac{\lambda}{t_0 - T_1} + \frac{\lambda}{s_0 - T_1}$$
$$+\delta_\alpha[\sum_{i=1}^n (e^{5\sqrt{1+z_{0,i}^2}} + e^{5\sqrt{1+w_{0,i}^2}}) + \sum_{i=1}^{2n}(z_{0,i}^6 + w_{0,i}^6)] \leq C.$$

Then (4.38) follows readily. □

Remark 4.3. By expanding $2\Phi(t_0, z_0, s_0, w_0) \geq \Phi(t_0, z_0, t_0, z_0) + \Phi(s_0, w_0, s_0, w_0)$ using all the individual terms, it is found that $\frac{1}{2\varepsilon}|t_0 - s_0|^2 + \frac{1}{2\varepsilon}|z_0 - w_0|^2$ is dominated by a continuous function $F(t_0, z_0, s_0, w_0)$ which goes to zero as $|t_0 - s_0| + |z_0 - w_0| \to 0$, which also follows from (4.38) when $\varepsilon \to 0$. □

Remark 4.4. The proof of the theorem is based on the methods in [**78, 42, 45, 20**]. Since here we deal with the function class $\mathcal{G}$ with a highly nonlinear growth condition on an unbounded domain, a localized semiconvex/semiconcave approximation technique is devised. The particular structure of the system dynamics also plays an important role in the proof of uniqueness, and in general it is more difficult to obtain uniqueness results under more general dynamics and the above fast growth condition. It is seen that the weak coupling feature of the dynamics of the state subprocess x is crucial for the above proof. When there exists an $a_i < 0$ (see Assumption **(H4.2)**), the estimate (4.54) would not be valid. □

4.5. Control with State Constraints

In this Section we consider the case when the state subprocess p is subject to constraints, i.e., the trajectory of each p_i must be maintained to be in a certain range. We term this situation as optimization under hard constraints. In [**66**] the author considered a deterministic model and obtained a constrained viscosity solution formulation for a first order HJB equation. Now due to the exogenous attenuation we come up with a second order HJB equation and we will develop a similar formulation. Suppose that $u \in U$ where U is a compact convex set in R^n, and p should satisfy $p_i \in [0, \overline{P}_i]$, where $\overline{P}_i$ is the upper bound. For simplicity we take $U = [-1,1]^n$ and $\overline{P}_i = \infty$. For any fixed initial value $p_0 \geq 0$ (i.e. each $(p_0)_i \geq 0$), define the admissible control set

$$\begin{aligned}\mathcal{U}^{p_0} = \{u(\cdot) \mid\ & u \text{ is adapted to } \sigma(z_s, s \leq t), \text{ and with probability } 1 \\ & (p_i(t) \geq 0 \text{ for all } 0 \leq t \leq T) \text{ holds, and } u(t) \in U,\ E\textstyle\int_0^T |u_t|^2 ds < \infty\}.\end{aligned}$$

In this Section we consider the simple case of

$$g(t,p,u) = u.$$

Under the admissible control set $\mathcal{U}^{p_0}$, we will use the notation of Section 4.2 for which the interpretation should be clear, and in the following we also use $\mathcal{U}^{p_0}$ with any initial time $s \leq T$. It is evident that $\mathcal{U}^{p_0}$ is a convex set. Under the norm $\|\cdot\|$ on $\mathcal{L}$ defined in Section 4.2, $\mathcal{U}^{p_0}$ is also closed. Indeed, if $\|u^{(k)} - u\| \to 0$ as $k \to \infty$, where $u^{(k)} \in \mathcal{U}^{p_0}$, one can show that u will also generate positive p trajectories with probability 1 with initial value p_0. So that $u \in \mathcal{U}^{p_0}$. As in the state unconstrained case, one can prove existence and uniqueness of the optimal control. Write

$$\begin{aligned} Q_T^0 &= [0,T) \times \mathrm{R}^n \times (0,\infty)^n, \\ Q_T &= [0,T) \times \mathrm{R}^n \times [0,\infty)^n, \\ \overline{Q}_T &= [0,T] \times \mathrm{R}^n \times [0,\infty)^n. \end{aligned}$$

We consider the HJB equation

$$\begin{aligned} 0 &= -\frac{\partial v}{\partial t} + \sup_{u \in U}\{-\frac{\partial^\tau v}{\partial z}\psi\} - \frac{1}{2}\mathrm{tr}(\frac{\partial^2 v}{\partial z^2} G G^\tau) - l, \qquad (4.58) \\ v\ &|_{t=T} = 0, \end{aligned}$$

where $(t, z) = (t, x, p) \in \overline{Q}_T$.

Definition 4.4. $v(t, z) \in C(\overline{Q}_T)$ is called a **constrained viscosity solution** to (4.58) if i) $v\,|_{t=T} = 0$, and for any $\varphi(t, z) \in C^{1,2}(\overline{Q}_T)$, whenever $v - \varphi$ takes a local maximum at $(t, z) \in Q_T^0$, we have

$$-\frac{\partial \varphi}{\partial t} + \sup_{u \in U}\{-\frac{\partial^\tau \varphi}{\partial z}\psi\} - \frac{1}{2}\mathrm{tr}(\frac{\partial^2 \varphi}{\partial z^2}GG^\tau) - l \leq 0, \quad z \in \mathrm{R}^{2n} \tag{4.59}$$

at (t, z), and ii) for any $\varphi(t, z) \in C^{1,2}(\overline{Q}_T)$, whenever $v - \varphi$ takes a local minimum at $(t, z) \in Q_T$, in (4.59) we have an opposite inequality at (t, z). For short, we term the constrained viscosity solution $v(t, z) \in C(\overline{Q}_T)$ as a viscosity subsolution on Q_T^0, and a viscosity supersolution on Q_T. □

Remark 4.5. Conditions i) and ii) hold on Q_T^0 and Q_T, respectively. Here we give a heuristic interpretation on how the state constrains are captured by Condition ii). Suppose $v - \varphi$ attains a minimum at $(\bar{t}, \bar{x}, \bar{p})$, where v is the value function and satisfies equation (4.58) at $(\bar{t}, \bar{x}, \bar{p})$ with classical derivatives, i.e.,

$$0 = -\frac{\partial v}{\partial t} + \{-\frac{\partial^\tau v}{\partial z}\psi\}|_{u=\widehat{u}} - \frac{1}{2}\mathrm{tr}(\frac{\partial^2 v}{\partial z^2}GG^\tau) - l. \tag{4.60}$$

In addition, we assume $\widehat{u}$ is admissible w.r.t. $(\bar{x}, \bar{p})$. Here $\bar{t} \in [0, T)$ and $\bar{p}$ lies on the boundary of $[0, \infty)^n$. By the necessary condition for a minimum, at $(\bar{t}, \bar{x}, \bar{p})$, we have

$$v_t - \varphi_t \geq 0, \quad v_{x_i} - \varphi_{x_i} = 0, \quad v_{x_i x_i} - \varphi_{x_i x_i} \geq 0, \quad 1 \leq i \leq n, \tag{4.61}$$

where the first inequality becomes equality when $\bar{t} \in (0, T)$. Since $\bar{p}$ is on the boundary of $[0, T)^n$, we can find an index set I such that $\bar{p}_i = 0$ when $i \in I$, and $\bar{p}_i > 0$ when $i \in \{1, \cdots, n\} \backslash I$. Again, by the minimum property at $(\bar{t}, \bar{x}, \bar{p})$ we get

$$v_{p_i} - \varphi_{p_i} \geq 0 \ \text{ for } \ i \in I, \quad v_{p_i} - \varphi_{p_i} = 0 \ \text{ for } \ i \in \{1, \cdots, n\} \backslash I, \tag{4.62}$$

at $(\bar{t}, \bar{x}, \bar{p})$. Since we assume $\widehat{u}$ is admissible w.r.t. $(\bar{x}, \bar{p})$, then we have $\widehat{u}_i \geq 0$ for $i \in I$, and therefore by (4.62), at $(\bar{t}, \bar{x}, \bar{p})$

$$(v_p - \varphi_p)^\tau \widehat{u} \geq 0. \tag{4.63}$$

Hence (4.61) and (4.63) lead to

$$-\frac{\partial\varphi}{\partial t}+\{-\frac{\partial^\tau\varphi}{\partial z}\psi\}|_{u=\hat{u}}-\frac{1}{2}\mathrm{tr}(\frac{\partial^2\varphi}{\partial z^2}GG^\tau)-l\geq 0,$$

and therefore Condition ii) holds at $(\bar{t},\bar{x},\bar{p})$. □

Lemma 4.4. For any initial pair (s_0,x_0,p_0) with each $(p_0)_i\geq 0$, and any $u\in\mathcal{U}$, there exists $\widetilde{u}\in\mathcal{U}^{p_0}$ such that

$$P_\Omega\{\int_{s_0}^T|\widetilde{u}-u|ds\leq 4\varepsilon\}=1, \tag{4.64}$$

where with probability 1 and for all $1\leq i\leq n$, the constant $\varepsilon>0$ satisfies

$$\sup_{t\in[s_0,T]}\max\{-p_i(t,s_0,p_0,u),\ 0\}\leq\varepsilon, \tag{4.65}$$

and $p(t,s_0,p_0,u)$ denotes the value of p at t corresponding to initial condition (s_0,p_0) and control u.

PROOF. We only need to modify each component u_i of u in the following way. Define $\tau_i^0=s_0$, and for $k\geq 1$,

$$\tau_i^k=\inf\{t>\tau_i^{k-1},\quad p_i(t,s_0,p_0,\widetilde{u})=0\}, \tag{4.66}$$

$$\tau_i^k=T\quad\text{if }\ p_i(t,\tau_i^{k-1}+\varepsilon,p_i(\tau_i^{k-1}+\varepsilon),u)>0\text{ for all }t\geq\tau_i^{k-1}+\varepsilon, \tag{4.67}$$

$$\widetilde{u}_i(t)=1\quad\text{ on }[\tau_i^{k-1},\ \tau_i^{k-1}+\varepsilon), \tag{4.68}$$

$$\widetilde{u}_i(t)=u_i(t)\quad\text{ on }[\tau_i^{k-1}+\varepsilon,\ \tau_i^k). \tag{4.69}$$

Then it is obvious that $\widetilde{u}\in\mathcal{U}^{p_0}$. Suppose (4.64) is not true, and then there exist i and a set A^0 with $P_\Omega(A^0)>0$, such that on A^0

$$\int_{s_0}^T|\widetilde{u}_i-u_i|ds>4\varepsilon. \tag{4.70}$$

For any fixed $\omega\in A^0$, if $\tau_i^{k_0}$ is the last stopping time defined by (4.66), then by (4.70) we can easily show that $p_i(\tau_i^{k_0-1},s_0,p_0,u)<-2\varepsilon$, which is a contradiction. □

Using Lemma 4.4, we can further show that the value function $v(t, z)$ is continuous on $\overline{Q}_T$ by a comparison method as in the unconstrained case [**23**]. The details are omitted here. The growth condition of Proposition 4.2 also holds in the constrained case.

Proposition 4.4. The value function v is a constrained viscosity solution to the HJB equation (4.58).

PROOF. We verify condition i) first. For an initial condition pair (s, z) with $z \in Q_T^0$, and any $u \in U$ we construct control $\widetilde{u} = u$ on $[s, s+\varepsilon]$ and $\widetilde{u} = 0$ on $(s+\varepsilon, T]$. We see that when ε is sufficiently small, $\widetilde{u}$ is in the admissible control set w.r.t. (s, z) since each $p_i \in [0, \infty)$. All the remaining part and the verification of condition ii) can be done as in Theorem 4.1. □

CHAPTER 5

Linear Quadratic Optimization for Wireless Power Control

5.1. Introduction

In Chapters 2 and 3, subject to bounded admissible controls, the value function of the optimal control problem is described by viscosity solutions, and in general, it is not possible to get an analytic optimal control law. To obtain optimal power allocation explicitly utilizing information on the channel states and power levels, in this Chapter we adopt another approach. The following analysis will be based on the lognormal fading model (2.1) and the power control model (2.2). For convenience of reading, here we also write the vector system model. As in Section 2.5 of Chapter 2, setting $f_i(x) = -a_i(x_i + b_i), 1 \leq i \leq n$, $H = \text{Diag }(\sigma_i)_{i=1}^n$ and $z^\tau = (x^\tau, p^\tau)$, $\psi^\tau = (f^\tau, u^\tau)$, $G^\tau = (H, 0_{n\times n})$, we write

$$dz = \psi dt + G dw, \qquad t \geq 0, \tag{5.1}$$

where all the variables are consistent with those in Chapter 2.

In the cost function introduced below, the cost integrand takes a quadratic form in terms of the power p and the control u while the attenuation x is regarded as a random parameter subject to no control. For this reason, we shall term the power control of this Chapter as "linear quadratic optimization".

In this quadratic cost based optimization framework, we impose no bound constraint on the control input u and introduce a penalty term for u in the cost function. We write

$$E\int_0^\infty e^{-\rho t}\{\sum_{i=1}^{n}[e^{x_i}p_i - \mu_i(\sum_{j=1}^{n} e^{x_j}p_j + \eta)]^2 + u^\tau Ru\}dt, \tag{5.2}$$

where R is a positive definite weight matrix, and the positive coefficients μ_i, $1 \le i \le n$, satisfy $\sum_{i=1}^{n}\mu_i < 1$ (i.e., Assumption **(H2.1)** also holds throughout this Chapter). This cost function includes a discount factor $\rho > 0$ and an infinite horizon, which will lead to an elliptic partial differential equation system describing the value function. In the above integral, the first term is based on the SIR requirements and the second term is added to penalize abrupt change of powers since in practical systems there exist basic physical limits for power adjustment rate. Another fact is that in real systems the operating conditions of a mobile are only estimated approximately, and it is generally preferred to avoid very rapid power change and hence the power of users is adjusted in a cautious manner. In (5.2), the weight matrix R should be chosen in accordance with power change rate requirements. After subtracting the constant component from the integrand in (5.2) we get the cost function

$$J(x,p,u) = E[\int_0^\infty e^{-\rho t}\{p^\tau C(x_t)p + 2D^\tau(x_t)p + u_t^\tau Ru_t\}dt|x,p], \tag{5.3}$$

where $C(x_t)$, $D(x_t)$ are $n \times n$ positive definite matrix and $n \times 1$ vector, respectively, which are determined from (5.2), and (x,p) denotes the initial state at $t = 0$. In this Chapter we adopt (5.3) as our cost function; also see [**39, 35**]. We remark that another possible way to approach the above power optimization problem is to modify the cost function (5.2) in a suitable form so that the power is adjusted to track an exogenous random signal based on stochastic pointwise optimum. Define the admissible control set

$$\mathcal{U}_2 = \{u|u \text{ adapted to } \sigma(x_s,p_s,s\le t), \text{ and } E\textstyle\int_0^\infty e^{-\rho t}|u_t|^2dt < \infty\}.$$

As in Chapter 2, in this Chapter we also assume that the initial value of p_s at $s = 0$ is deterministic; then one simply has $\sigma(x_s, p_s, s \le t) = \sigma(x_s, s \le t)$.

We define the value function associated with the the cost (5.3) as

$$v(x,p) = \inf_{u \in \mathcal{U}_2} J(x,p,u). \tag{5.4}$$

Notice that certain controls from $\mathcal{U}_2$ may result in an infinite cost due to the presence of the e^{x_i} process, $1 \le i \le n$. However the optimal control problem is still well defined subject to the new admissible control set $\mathcal{U}_2$ and one can show the existence of an optimal control by standard approximation techniques [**78**].

We investigate the infinite horizon optimal control problem and its associated HJB equation. The merit of minimization of the infinite horizon cost is that the resulting optimal control law is in a steady state form and various suboptimal control laws can be constructed by an algebraic approach based on this HJB equation. In order to analyze the infinite horizon optimal cost by a discrete approximation technique, we will also study the finite horizon cost case to obtain some auxiliary results in Section 5.2 below.

In the cost (5.3), the weight matrix $C(x_t)$ is related to the unbounded random processes e^{x_i}, $1 \le i \le n$. In Section 5.2, a certain truncation technique is used to deal with $C(x_t)$ and then obtain a structure for the optimal cost function in the finite horizon case. Section 5.2 is quite technical. The reader may *skip* the long sequence of lemmas and simply refer to the main result in Theorem 5.1 which will be used in Section 5.3.

5.2. The Finite Horizon Control Problem and Some Auxiliary Results

Subject to the system dynamics (5.1), for $0 < T < \infty$, we define the finite horizon version of the cost function (5.3) as

$$J^T(x,p,u) = E\left[\int_0^T e^{-\rho t}\{\Phi(x_t,p_t) + u_t^\tau R u_t\} dt | x, p\right], \tag{5.5}$$

where $\Phi(x_t, p_t) = p_t^\tau C(x_t)p_t + 2D^\tau(x_t)p_t$ and (x, p) denotes the initial state at $t = 0$. It is evident that

$$\Phi(x_t, p_t) + \eta^2(\sum_{i=1}^{n} \mu_i)^2 \triangleq \Phi(x_t, p_t) + \eta_\mu \geq 0, \tag{5.6}$$

for all $x_t, p_t \in \mathrm{R}^n$. For integer $N > 0$, we also define the truncated version of $J^T(x, p, u)$ as

$$J_N^T(x, p, u) = E\left[\int_0^T e^{-\rho t}\{\Phi(x_t, p_t)1_{(|x_t|\leq N)} + u_t^\tau R u_t\}dt|x, p\right], \tag{5.7}$$

where $1_{(|x_t|\leq N)}$ is the indicator function. In both (5.5) and (5.7), p_t is generated by the control u_t through the dynamics (5.1) on $[0, T]$. Define the admissible control set as $\mathcal{U}_2^T = \{u|u$ adapted to $\sigma(x_s, p_s, s \leq t)$, and $E\int_0^T e^{-\rho t}|u_t|^2 dt < \infty\}$. Set $v^T(x, p) = \inf_{u\in\mathcal{U}_2^T} J^T(x, p, u)$. For developing discrete approximation schemes in below, we need to define a subset of $\mathcal{U}_2^T$ as $\mathcal{U}_2^{T,k} = \{u|u \in \mathcal{U}_2^T$ and is a stepwise random process specified by times $t = \frac{iT}{2^k}, 0 \leq i \leq 2^k\}$. Write $v^{T,k}(x, p) = \inf_{u\in\mathcal{U}_2^{T,k}} J^T(x, p, u)$. We give the following lemmas.

Lemma 5.1. For a sequence of R^n-valued controls $u_k \in \mathcal{U}_2^T$, $k = 1, 2, \cdots$, assume there exists $u_\infty \in \mathcal{U}_2^T$ such that $\lim_{k\to\infty} E\int_0^T |u_k - u_\infty|^2 dt = 0$ and denote by p_k the solution to $dp = udt$ corresponding to $u = u_k$, $k = 1, \cdots, \infty$, and the same initial condition $p|_{t=0}$. Then $\lim_{k\to\infty} E\int_0^T |p_k - p_\infty|^2 dt = 0$, and $\lim_{k\to\infty} J_N^T(x, p, u_k) = J_N^T(x, p, u_\infty)$ where p stands for the initial value $p|_{t=0}$.

PROOF. It is obvious that $|p_k - p_\infty|_t \leq \int_0^t |u_k - u_\infty|_s ds$, which yields

$$\begin{aligned} &E\int_0^T |p_k - p_\infty|_t^2 dt \leq E\int_0^T (\int_0^t |u_k - u_\infty|_s ds)^2 dt \\ \leq &E\int_0^T \int_0^t ds \cdot \int_0^t |u_k - u_\infty|_s^2 ds dt \leq E\int_0^T t\int_0^T |u_k - u_\infty|_s^2 ds dt \\ = &\frac{T^2}{2} E\int_0^T |u_k - u_\infty|_s^2 ds \to 0, \quad \text{as } k \to \infty. \end{aligned} \tag{5.8}$$

To prove the last part of the lemma, by Schwartz inequality we have

$$\begin{aligned}
&E\left|\int_0^T [p_k^\tau C(x_t)p_k + 2D^\tau(x_t)p_k]1_{(|x_t|\le N)}dt - \int_0^T [p_\infty^\tau C(x_t)p_\infty + 2D^\tau(x_t)p_\infty]1_{(|x_t|\le N)}dt\right| \\
&= E\left|\int_0^T (p_k - p_\infty)^\tau C(x_t)p_k 1_{(|x_t|\le N)}dt + \int_0^T (p_k - p_\infty)^\tau C(x_t)p_\infty 1_{(|x_t|\le N)}dt\right. \\
&\quad \left. +2\int_0^T D^\tau(x_t)(p_k - p_\infty)1_{(|x_t|\le N)}dt\right| \\
&\le E\int_0^T |p_k - p_\infty|^2 dt \cdot \left[E\int_0^T |C(x_t)p_k|^2 1_{(|x_t|\le N)}dt + \right. \\
&\qquad \left. E\int_0^T |C(x_t)p_\infty|^2 1_{(|x_t|\le N)}dt + E\int_0^T |D(x_t)|^2 1_{(|x_t|\le N)}dt\right] \qquad (5.9) \\
&\to 0, \quad \text{as } k\to\infty, \qquad (5.10)
\end{aligned}$$

where (5.10) follows from the L_2 convergence of u_k, p_k, $k \ge 1$, on $[0,T]\times\Omega$ and boundedness of $C(x_t)1_{(|x_t|\le N)}$, $D(x_t)1_{(|x_t|\le N)}$ in (5.9). Similarly, we have

$$E|\int_0^T u_k^\tau R u_k dt - \int_0^T u_\infty^\tau R u_\infty dt| \to 0, \quad \text{as } k\to\infty, \qquad (5.11)$$

and therefore it follows from (5.10)-(5.11) that $\lim_{k\to\infty} J_N^T(x,p,u_k) = J_N^T(x,p,u_\infty)$. □

Lemma 5.2. For $J_N^T(x,p,u)$ defined by (5.7), we have

$$\lim_{N\to\infty}\inf_{u\in\mathcal{U}_2^T} J_N^T(x,p,u) = v^T(x,p), \qquad (5.12)$$

for any $x\in\mathrm{R}^n$ and $p\in\mathrm{R}^n$.

PROOF. We define

$$J_{N,\eta_\mu}^T(x,p,u) \triangleq J_N^T(x,p,u) + \eta_\mu E[\int_0^T e^{-\rho t}1_{(|x_t|\le N)}dt|x],$$

$$v_{\eta_\mu}^T(x,p) \triangleq v^T(x,p) + \eta_\mu\int_0^T e^{-\rho t}dt,$$

where $\eta_\mu = \eta^2(\sum_{i=1}^n \mu_i)^2$. Recalling (5.6) and the definition of $J_N^T(x,p,u)$, $v^T(x,p)$, for any (x,p) and $N > 0$, it is obvious that

$$\inf_{u \in \mathcal{U}_2^T} J_{N,\eta_\mu}^T(x,p,u) \leq v_{\eta_\mu}^T(x,p), \tag{5.13}$$

and

$$J_{N_1,\eta_\mu}^T(x,p,u) \leq J_{N_2,\eta_\mu}^T(x,p,u), \tag{5.14}$$

for $N_1 < N_2$. We will show that

$$\lim_{N\to\infty} \inf_{u \in \mathcal{U}_2^T} J_{N,\eta_\mu}^T(x,p,u) = v_{\eta_\mu}^T(x,p), \tag{5.15}$$

for all $x, p \in \mathrm{R}^n$. We prove (5.15) by contradiction. If (5.15) is not true, then by (5.13)-(5.14), there exist a pair (x,p) and $\varepsilon > 0$ such that

$$\inf_{u \in \mathcal{U}_2^T} J_{N,\eta_\mu}^T(x,p,u) < v_{\eta_\mu}^T(x,p) - \varepsilon, \tag{5.16}$$

for all $N > 0$. Then for each fixed N, there exists $u_N \in \mathcal{U}_2^T$ such that

$$J_{N,\eta_\mu}^T(x,p,u_N) < v_{\eta_\mu}^T(x,p) - \frac{\varepsilon}{2}, \tag{5.17}$$

which together with (5.7) implies $\sup_N E\int_0^T |u_N|^2 dt < \infty$; hence by well known results in functional analysis [**79, 78**], there exists a subsequence $\{u_{N_i}, i = 1, 2, \cdots\}$ of $\{u_N, N = 1, 2, 3\cdots\}$ such that u_{N_i} converges weakly to a limit $\widehat{u} \in \mathcal{U}_2^T$. For simplicity in the following we still denote $\{u_{N_i}\}$ by $\{u_N\}$. Furthermore, by Mazur's theorem [**79**], there exist $\lambda_{ik} \geq 0, \sum_{k=1}^\infty \lambda_{ik} = 1$ such that

$$\lim_{i\to\infty} E\int_0^T |\sum_{k=1}^\infty \lambda_{ik} u_{i+k} - \widehat{u}|^2 dt = 0, \tag{5.18}$$

Now by (5.18) and Lemma 5.1 it follows that

$$
\begin{aligned}
J^T_{N,\eta_\mu}(x,p,\widehat{u}) &= \lim_{N\leq i\to\infty} J^T_{N,\eta_\mu}(x,p,\sum_{k=1}^{\infty}\lambda_{ik}u_{i+k}) \\
&\leq \lim_{N\leq i\to\infty} \sum_{k=1}^{\infty}\lambda_{ik}J^T_{N,\eta_\mu}(x,p,u_{i+k}) \qquad (5.19)\\
&\leq \lim_{N\leq i\to\infty} \sum_{k=1}^{\infty}\lambda_{ik}J^T_{i+k,\eta_\mu}(x,p,u_{i+k}) \qquad (5.20)\\
&\leq v^T_{\eta_\mu}(x,p) - \frac{\varepsilon}{2}. \qquad (5.21)
\end{aligned}
$$

Here (5.19) is obtained by convexity with respect to (p,u) of the integrand in (5.7) and the linear dynamics $dp = u$; (5.20) follows by (5.14) and we get (5.21) from (5.17). On the other hand, by Lebesgue's monotone convergence theorem [**64**] we have

$$
J^T_{N,\eta_\mu}(x,p,\widehat{u}) \uparrow J^T(x,p,\widehat{u}) + \eta_\mu \int_0^T e^{-\rho t}dt, \qquad \text{as } N\uparrow\infty, \qquad (5.22)
$$

since $[\Phi(x_t,p_t)+\eta_\mu]1_{(|x_t|\leq N)} \uparrow \Phi(x_t,p_t)+\eta_\mu$ a.e. on $[0,T]\times\Omega$. Hence by (5.22) and (5.21), we have

$$
J^T(x,p,\widehat{u}) = \lim_{N\to\infty} J^T_{N,\eta_\mu}(x,p,\widehat{u}) - \eta_\mu \int_0^T e^{-\rho t}dt \leq v^T(x,p) - \frac{\varepsilon}{2}, \qquad (5.23)
$$

which is a contradiction. Thus we have proved that (5.15) holds, and therefore

$$
\begin{aligned}
\lim_{N\to\infty}\inf_{u\in\mathcal{U}_2^T} J^T_N(x,p,u) &= \lim_{N\to\infty}\inf_{u\in\mathcal{U}_2^T}[J^T_{N,\eta_\mu}(x,p,u) - \eta_\mu E\int_0^T e^{-\rho t}1_{(|x_t|\leq N)}dt] \\
&= \lim_{N\to\infty}\inf_{u\in\mathcal{U}_2^T} J^T_{N,\eta_\mu}(x,p,u) - \eta_\mu\int_0^T e^{-\rho t}dt \\
&= v^T_{\eta_\mu}(x,p) - \eta_\mu\int_0^T e^{-\rho t}dt = v^T(x,p)
\end{aligned}
$$

and the lemma follows. □

Lemma 5.3. For any $u \in \mathcal{U}_2^T$ there exists a sequence of stepwise random processes $u^k \in \mathcal{U}_2^{T,k}$, $k \geq 1$, such that

$$\lim_{k\to\infty} E\int_0^T |u_t - u_t^k|^2 dt = 0.$$

PROOF. This follows easily from the proof of Lemma 4.4 in [**55**], and in fact each u^k can be chosen to be bounded by a deterministic constant. □

Lemma 5.4. The finite horizon optimal cost function $v^T(x,p)$ can be represented in the form

$$v^T(x,p) = \lim_{N\to\infty}\lim_{k\to\infty}\inf_{u\in\mathcal{U}_2^{T,k}} J_N^T(x,p,u). \tag{5.24}$$

PROOF. By Lemma 5.2 it suffices to prove that

$$\lim_{k\to\infty}\inf_{u\in\mathcal{U}_2^{T,k}} J_N^T(x,p,u) = \inf_{u\in\mathcal{U}_2^T} J_N^T(x,p,u), \tag{5.25}$$

for which the left hand side exists since $\mathcal{U}_2^{T,k} \subset \mathcal{U}_2^{T,k+1}$ and the sequence of optimal costs (relative to $\mathcal{U}_2^{T,k}$) $\inf_{u\in\mathcal{U}_2^{T,k}} J_N^T(x,p,u)$, $k \geq 1$, monotonely decreases as $k \uparrow \infty$. Since for all k,

$$\inf_{u\in\mathcal{U}_2^{T,k}} J_N^T(x,p,u) \geq \inf_{u\in\mathcal{U}_2^T} J_N^T(x,p,u),$$

it follows that

$$\lim_{k\to\infty}\inf_{u\in\mathcal{U}_2^{T,k}} J_N^T(x,p,u) \geq \inf_{u\in\mathcal{U}_2^T} J_N^T(x,p,u). \tag{5.26}$$

Now we only need to prove an opposite inequality for (5.26). For any $\varepsilon > 0$, take $\widehat{u} \in \mathcal{U}_2^T$ such that $J_N^T(x,p,\widehat{u}) \leq \inf_{u\in\mathcal{U}_2^T} J_N^T(x,p,u) + \frac{\varepsilon}{3}$. By Lemmas 5.3 and 5.1 there exists a sufficiently large k_0 and $u^{k_0} \in \mathcal{U}_2^{T,k_0}$ such that $|J_N^T(x,p,u^{k_0}) - J_N^T(x,p,\widehat{u})| \leq \frac{\varepsilon}{3}$. Then it follows that

$$\begin{aligned}\lim_{k\to\infty}\inf_{u\in\mathcal{U}_2^{T,k}} J_N^T(x,p,u) &\leq \inf_{u\in\mathcal{U}_2^{T,k_0}} J_N^T(x,p,u) \leq J_N^T(x,p,u^{k_0})\\ &\leq J_N^T(x,p,\widehat{u}) + \frac{\varepsilon}{3} \leq \inf_{u\in\mathcal{U}_2^T} J_N^T(x,p,u) + \frac{2\varepsilon}{3}.\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, it follows that

$$\lim_{k\to\infty} \inf_{u\in\mathcal{U}_2^{T,k}} J_N^T(x,p,u) \le \inf_{u\in\mathcal{U}_2^T} J_N^T(x,p,u), \tag{5.27}$$

which completes the proof. □

Lemma 5.5. For any fixed pair N and k, $v_N^{T,k}(x,p,u) \triangleq \inf_{u\in\mathcal{U}_2^{T,k}} J_N^T(x,p,u)$ is a quadratic form in terms of p with coefficients depending on x.

PROOF. We decompose the cost function to get

$$\begin{aligned} &J_N^T(x,p,u) \\ =&E\left[\sum_{i=0}^{2^k-1}\int_{\frac{iT}{2^k}}^{\frac{(i+1)T}{2^k}} [p_t^\tau C(x_t)p_t 1_{(|x_t|\le N)} + 2D^\tau(x_t)p_t 1_{(|x_t|\le N)} + u_t R u_t]dt|x,p\right], \end{aligned} \tag{5.28}$$

where $u \in \mathcal{U}_2^{T,k}$. We minimize $J_N^T(x,p,u)$ backward by applying dynamic programming. The last term in the sum is given by

$$s(2^k-1) = \int_{\frac{(2^k-1)T}{2^k}}^{T} [p_t^\tau C(x_t)p_t 1_{(|x_t|\le N)} + 2D^\tau(x_t)p_t 1_{(|x_t|\le N)} + u_t R u_t]dt. \tag{5.29}$$

Denote $r_0 = \frac{(2^k-1)T}{2^k}$. In the above integral the initial condition for the state variable is (x_{r_0}, p_{r_0}). For $t \in [r_0, T]$, $u_t = u_{r_0}$ and $p_t = p_{r_0} + u_{r_0}(t-r_0)$. Then there exist $F_1(\cdot)$, $F_2(\cdot)$, $F_3(\cdot)$, $F_4(\cdot)$ such that

$$\begin{aligned} &E[s(2^k-1)|\sigma(x_t,p_t,t\le r_0)] \\ =&p_{r_0}^\tau F_1(x_{r_0})p_{r_0} + F_2^\tau(x_{r_0})p_{r_0} + u_{r_0}^\tau F_3(x_{r_0})u_{r_0} + F_4^\tau(x_{r_0})u_{r_0}, \end{aligned} \tag{5.30}$$

where the matrices $F_1(x_{r_0}) \ge 0$ and $F_4(x_{r_0}) \ge R$. By dynamic programming we minimize (5.30) and obtain the minimum as a quadratic function of p_{r_0} with coefficients depending on x_{r_0}. Repeating the above LQ minimization procedure in (5.28) and by induction we see that $v_N^{T,k}(x,p,u)$ is a quadratic form in terms of p with coefficients depending on x. □

Lemma 5.6. Suppose a sequence of continuous functions $v_k(x, p)$ mapping $\mathrm{R}^n \times \mathrm{R}^n$ into R, $k \geq 1$, is represented as

$$v_k(x, p) = p^\tau K_k(x)p + 2p^\tau S_k(x) + q_k(x), \tag{5.31}$$

where $K_k(x) = K_k^\tau(x)$, and in addition, there exists a finite $\lim_{k\to\infty} v_k(x, p)$ for all $x, p \in \mathrm{R}^n$. Then there exist $K_\infty(x) = K_\infty^\tau(x)$, $S_\infty(x)$, $q_\infty(x)$, all continuous in x, such that $(K_k(x), S_k(x), q_k(x)) \to (K_\infty(x), S_\infty(x), q_\infty(x))$, as $k \to \infty$.

PROOF. We have

$$\lim_{k\to\infty} v_k(x, p) = \lim_{k\to\infty} [p^\tau K_k(x)p + 2p^\tau S_k(x) + q_k(x)].$$

In particular,

$$\lim_{k\to\infty} v_k(x, 0) = \lim_{k\to\infty} q_k(x) \triangleq q_\infty(x), \tag{5.32}$$

which further implies that $\lim_{k\to\infty}[p^\tau K_k(x)p + 2p^\tau S_k(x)]$ also exists and is finite for all (x, p). We take $p = [0, \cdots, p_i, \cdots, 0]^\tau$, then $\lim_{k\to\infty} K_{kii}(x)p_i^2 + 2S_{ki}p_i =$ finite. By taking $p_i = 1, 2$, respectively, we get $\lim_{k\to\infty} K_{kii}(x) + 2S_{ki} =$ finite, $\lim_{k\to\infty} 4K_{kii}(x) + 4S_{ki} =$ finite. Then it follows easily that both $\lim_{k\to\infty} K_{kii}(x)$ and $\lim_{k\to\infty} S_{ki}(x)$ have finite values. Repeating this procedure for each entry of the matrices, we see that there exist $K_\infty(x)$, $S_\infty(x)$ such that

$$\lim_{k\to\infty} K_k(x) = K_\infty(x), \quad \lim_{k\to\infty} S_k(x) = S_\infty(x).$$

For any fixed x, using a similar argument as above, we can show that K_∞, S_∞, q_∞ are continuous at x. This completes the proof. □

We conclude this Section with the following theorem concerning the structure of $v^T(x, p)$.

Theorem 5.1. There exist $K^T(x), S^T(x)$ and $q^T(x)$ of suitable dimension such that the finite horizon optimal cost $v^T(x, p)$ can be represented as

$$v^T(x, p) = p^\tau K^T(x)p + 2p^\tau S^T(x) + q^T(x), \tag{5.33}$$

where $K^T(x) = (K^T)^\tau(x)$, and the superscript T is used to indicate the time horizon.

PROOF. This follows from Lemmas 5.4, 5.5 and 5.6. □

5.3. The Infinite Horizon Optimal Cost and the HJB Equation

We proceed to analyze the infinite horizon optimal control problem formulated in Section 5.1; formally applying dynamic programming, we may write the HJB equation for the value function v defined by (5.4) as follows:

$$0 = \rho v - f^\tau \frac{\partial v}{\partial x} - \frac{1}{2}\mathrm{tr}(\frac{\partial^2 v}{\partial z^2} GG^\tau) + \sup_{u\in \mathrm{R}^n}\{-u^\tau \frac{\partial v}{\partial p} - u^\tau R u\} - p^\tau C(x)p - 2D^\tau(x)p,$$

which gives

$$\begin{aligned}\rho v =& f^\tau \frac{\partial v}{\partial x} + \frac{1}{2}\mathrm{tr}[GG^\tau \frac{\partial^2 v}{\partial z^2}] - \frac{1}{4}(\frac{\partial v}{\partial p})^\tau R^{-1}\frac{\partial v}{\partial p} + p^\tau C(x)p + 2D^\tau(x)p \\ =& -\sum_{i=1}^n a_i(x_i+b_i)\frac{\partial v}{\partial x_i} + \frac{1}{2}\sum_{i=1}^n \sigma_i^2 \frac{\partial^2 v}{\partial {x_i}^2} - \frac{1}{4}v_p^\tau R^{-1} v_p + p^\tau C(x)p + 2D^\tau(x)p.\end{aligned} \tag{5.34}$$

Before v is determined as a classical solution to the HJB equation (5.34), we need some local Lipschitz estimates. To determine the range of x and p for the following comparison method, we define a subset of $\mathrm{R}^n \times \mathrm{R}^n$ by taking $Q_B = \{(x,p)|x,p \in \mathrm{R}^n \text{ and } |x| \le B, |p| \le B\}$, where the constant $B > 0$.

Since $-\frac{\eta^2}{\rho}\sum_{i=1}^n \mu_i^2 \le v(x,p) \le J(x,p,0)$, there exists a constant $C_B > 0$ depending on Q_B such that $\sup_{(x,p)\in Q_B}|v(x,p)| \le C_B$. For each $(x,p) \in Q_B$ define the subset of the admissible control set $\mathcal{U}_2$ as $\mathcal{U}_2^{(x,p)} = \{u \in \mathcal{U}_2 | J(x,p,u) \le 2C_B\}$ and take the union $\mathcal{U}_{2,B} = \cup_{(x,p)\in Q_B}\mathcal{U}_2^{(x,p)}$. By explicitly substituting any initial condition $(x,p) \in Q_B$ into the integrand of $J(x,p,\cdot)$ and by basic bound estimates for the entries $e^{x_i}p_i(t)$ and their products involved in $p_t^\tau C(x_t)p_t$, it can be further verified that there exists constants $\widehat{C}_B^1$, $\widehat{C}_B^2$ such that

$$|J(x,p,u)| \le \widehat{C}_B^1, \qquad |J(x,p,u) - J(x',p',u)| \le \widehat{C}_B^2(|x-x'| + |p-p'|), \tag{5.35}$$

where $u \in \mathcal{U}_{2,B}$ and $(x,p),(x',p') \in Q_B$.

On the other hand, for $(x,p),(x',p') \in Q_B$ we have

$$|v(x,p) - v(x',p')| \leq \sup_{u \in \mathcal{U}_{2,B}} |J(x,p,u) - J(x',p',u)|. \tag{5.36}$$

Indeed, for any $\varepsilon > 0$ there exist $\widehat{u}, \widehat{u}' \in \mathcal{U}_{2,B}$ such that

$$J(x,p,\widehat{u}) < v(x,p) + \frac{\varepsilon}{4}, \quad J(x',p',\widehat{u}') < v(x',p') + \frac{\varepsilon}{4} \tag{5.37}$$

Without loss of generality, we assume $v(x,p) > v(x',p')$ and ε has been chosen sufficiently small such that $v(x,p) - v(x',p') \geq \varepsilon$. Then it follows that

$$\begin{aligned} |v(x,p) - v(x',p')| &\leq J(x,p,\widehat{u}) - J(x',p',\widehat{u}') + \frac{\varepsilon}{4} \\ &\leq J(x,p,\widehat{u}') - J(x',p',\widehat{u}') + \frac{\varepsilon}{2} \\ &\leq \sup_{u \in \mathcal{U}_{2,B}} |J(x,p,u) - J(x',p',u)| + \frac{\varepsilon}{2} \end{aligned}$$

which leads to (5.36) since $\varepsilon > 0$ can be arbitrarily small.

Combining (5.36) and (5.35) we obtain the following proposition:

Proposition 5.1. There exists a constant $\widehat{C}_B > 0$ such that

$$|v(x,p) - v(x',p')| \leq \widehat{C}_B(|x - x'| + |p - p'|),$$

where $(x,p),(x',p') \in Q_B \triangleq \{(x,p) | x,p \in \mathrm{R}^n, \text{ and } |x| \leq B, \ |p| \leq B\}, \ B > 0.$ □

Theorem 5.2. The value function v defined by (5.4) is a continuous function of (x,p) and can be written as

$$v(x,p) = p^\tau K(x)p + 2p^\tau S(x) + q(x) \tag{5.38}$$

where $K(x) = K(x)^\tau$, $S(x)$ and $q(x)$ are continuous functions of x, and are all of order $O(1 + \sum_{i=1}^n e^{2x_i})$.

PROOF. The continuity of v follows from Proposition 5.1 since B can be taken as any positive constant. We approximate $v(x,p)$ by a sequence of costs $v^T(x,p)$, $T > 0$, each of which is the optimal cost of the stochastic control problem with the time horizon $[0,T]$. Applying a similar argument as in the proof of Lemma 5.2 to the current case by extending the time horizon to infinite, it can be verified that $v(x,p) = \lim_{T\to\infty} v^T(x,p)$ for all (x,p), and therefore (5.38) holds by Theorem 5.1 and Lemma 5.6. The upper bound for $K(x), S(x), q(x)$ is obtained by a direct estimate of the growth rate of v. □

Theorem 5.3. The value function v is a classical solution to the HJB equation (5.34), i.e., $\frac{\partial v}{\partial x_i}$, $\frac{\partial^2 v}{\partial {x_i}^2}$, $\frac{\partial v}{\partial p_i}$, $1 \le i \le n$, exist and are continuous on R^{2n}.

PROOF. By a vanishing viscosity technique [**23, 78**] one can show that the value function v is a generalized solution to (5.34) in terms of weak derivatives with respect to (x,p). By Theorem 5.2, we see that $\frac{\partial v}{\partial p}$ exists and is continuous. Now (5.34) can be looked at as a partial differential equation parametrized by p. Then one can further show by use of smooth test functions of the form $\varphi_1(x)\varphi_2(p)$ with compact support that v is a generalized solution with respect to x for each fixed p.

By Proposition 5.1 one can verify that $K(x), S(x)$ satisfy a local Lipschitz condition, and hence for each fixed p, the term $\Psi^p(x) \triangleq -\frac{1}{4}v_p^\tau R^{-1} v_p + p^\tau C(x)p + 2D^\tau(x)p$ in (5.34) also satisfies a local Lipschitz condition w.r.t. x.

For each fixed p, (5.34) can be written as follows:

$$-\rho v - \sum_{i=1}^{n} a_i(x_i + b_i)\frac{\partial v}{\partial x_i} + \frac{1}{2}\sum_{i=1}^{n}\sigma_i^2 \frac{\partial^2 v}{\partial {x_i}^2} + \Psi^p(x) = 0. \tag{5.39}$$

Since (5.39) is uniformly elliptic and Ψ^p is locally Lipschitz continuous w.r.t. x, the generalized solution v (w.r.t. x) has the first and second order classical derivatives with respect to x [**25**], i.e., $\frac{\partial v}{\partial x_i}$, $\frac{\partial^2 v}{\partial {x_i}^2}$, $1 \le i \le n$, exist and are continuous. Hence v has all the classical derivatives appearing in the HJB equation (5.34). □

5.3.1. Associated PDE's and the Control Law. By Theorems 5.2 and 5.3, we have

$$
\begin{aligned}
& p^\tau \rho K(x)p + 2p^\tau \rho S(x) + \rho q(x) \\
=& f^\tau(x)\frac{\partial}{\partial x}[p^\tau K(x)p + 2p^\tau S(x) + q(x)] \\
& + \frac{1}{2}\mathrm{tr}\{GG^\tau \frac{\partial^2}{\partial z^2}[p^\tau K(x)p + 2p^\tau S(x) + q(x)]\} \\
& - [K(x)p + S]^\tau R^{-1}[K(x)p + S] + p^\tau C(x)p + 2D^\tau(x)p.
\end{aligned}
$$

This gives

$$
\begin{aligned}
& p^\tau \rho K(x)p + 2p^\tau \rho S(x) + \rho q(x) \\
=& p^\tau(\sum_{k=1}^n f_k \frac{\partial K}{\partial x_k})p + 2p^\tau(\sum_{i=1}^n f_k \frac{\partial S}{\partial x_k}) + f^\tau \frac{\partial q}{\partial x} \\
& + p^\tau(\sum_{k=1}^n \frac{\sigma_k^2}{2}\frac{\partial^2 K}{\partial {x_k}^2})p + p^\tau \sum_{k=1}^n \sigma_k^2 \frac{\partial^2 S}{\partial {x_k}^2} + \sum_{k=1}^n \frac{\sigma_k^2}{2}\frac{\partial^2 q}{\partial {x_k}^2} \\
& - p^\tau K R^{-1} K p - S^\tau R^{-1} S - 2p^\tau K R^{-1} S.
\end{aligned}
$$

Hence we get the partial differential equation system

$$
\rho K = \frac{1}{2}\sum_{k=1}^n \sigma_k^2 \frac{\partial^2 K}{\partial {x_k}^2} + \sum_{k=1}^n f_k \frac{\partial K}{\partial x_k} - KR^{-1}K + C, \tag{5.40}
$$

$$
\rho S = \frac{1}{2}\sum_{k=1}^n \sigma_k^2 \frac{\partial^2 S}{\partial {x_k}^2} + \sum_{k=1}^n f_k \frac{\partial S}{\partial x_k} - KR^{-1}S + D, \tag{5.41}
$$

$$
\rho q = \frac{1}{2}\sum_{k=1}^n \sigma_k^2 \frac{\partial^2 q}{\partial {x_k}^2} + f^\tau \frac{\partial q}{\partial x} - S^\tau R^{-1} S, \tag{5.42}
$$

where we shall refer to (5.40) as the Riccati equation of the system. Finally the optimal control law is given by

$$
u = [u_1, \cdots, u_n]^\tau = -R^{-1}[K(x)p + S(x)], \tag{5.43}
$$

where p denotes the power vector. This gives the control law for all users. From the above it is seen that the optimal control is determined as a feedback which uses the measurement of the current power and attenuation processes.

5.3.2. Simulations. In the numerical experiments below, we study two different systems with the same cost function. Each system includes two users and the channel dynamics of both systems are given by

$$\text{System A:} \qquad dx_i = -4(x_i + 0.3)dt + 0.3dw_i, \quad i = 1, 2,$$

and

$$\text{System B:} \qquad \begin{cases} dx_1 = -4(x_1 + 0.3)dt + 0.3dw_1, \\ dx_2 = -3.5(x_2 + 0.2)dt + 0.2dw_2. \end{cases}$$

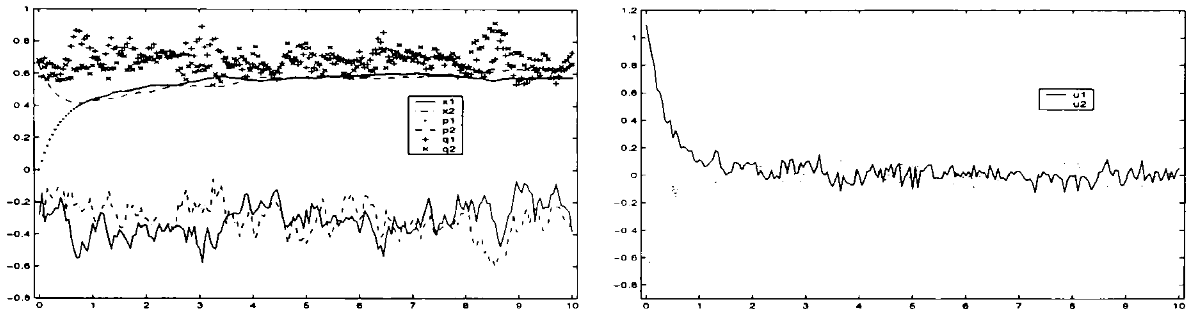
In the quadratic type cost function (5.3), the discount factor $\rho = 0.5$, the weight matrix $R = 0.03I_2$, and (5.3) is derived from (5.2) where $\mu_1 = \mu_2 = 0.4, \eta = 0.25$. In the simulation the time step size is 0.05. We use a similar difference scheme as in the bounded control case of Chapter 3 to compute the value function approximately and the control law is determined by a quadratic type minimization based calculation, i.e.

$$u = \text{argmin}_u \{u^T \frac{\partial v}{\partial p} + u^T Ru\}.$$

Figures 5.1 and 5.2 correspond to system A and system B, respectively, where $x_i, p_i, q_i, i = 1, 2$, denote the attenuation, the controlled power, and the static pointwise optimum obtained from (2.10), respectively. Figures 5.1 (b) and 5.2 (b) indicate the control inputs corresponding to Figures 5.1 (a) and 5.2 (a), respectively.

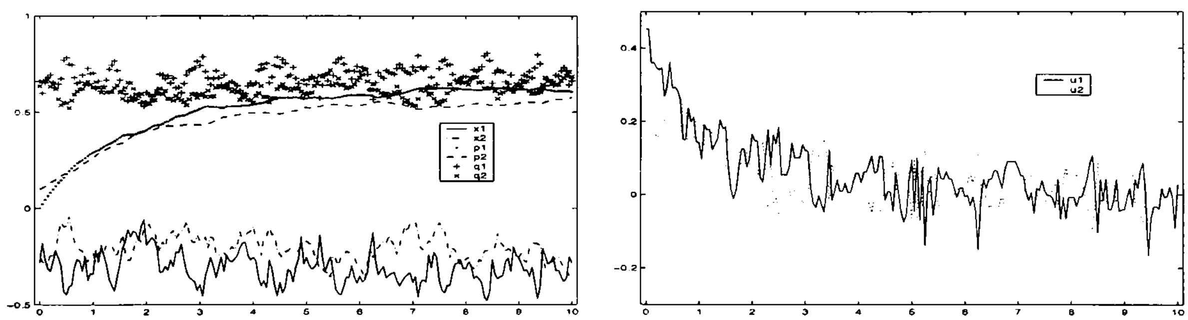
For system A, as shown by Figure 5.1, whenever the two users have significantly different initial powers there is an initial convergence of the power levels to a common level and then subsequent approximately equal behavior which converges toward a steady level. In the long term, the two controlled power trajectories are very close to each other; this happens because the two mobiles have i.i.d. channel dynamics. For system B, after a fast adjustment, the powers of two mobiles will be maintained at stable levels and the power trajectories of the first user will generally stay above that of the second user due to the asymmetry of the channel dynamics of the two users, which is different from the case of system A.

From Figures 5.1 (b) and 5.2 (b) it is seen that in contrast to their initial behavior, after both powers settle down in a small neighborhood of the optimum, at each step only a minor effort is required for each mobile to adjust its power, which also differs from the bounded control case in Chapter 3 where the power adjustment takes the form of bang-bang controls.



(a) $p_1(0) = 0,\ p_2(0) = 0.65$ (b) u_1 and u_2

FIGURE 5.1. Simulation for system A



(a) $p_1(0) = 0,\ p_2(0) = 0.1$ (b) u_1 and u_2

FIGURE 5.2. Simulation for system B

5.4. The Classical Solution and its Small Noise Approximation

In this section we address the important issue of the computability of solutions to the equations in Section 5.3.1. In general, constructing a control law by solving the partial differential equation systems in the high dimensional case is a formidable or even impossible task. For a significant number of users, an analysis of local expansions of solutions around a steady state attenuation value $\overline{x}$ is useful in the small noise case because the attenuation trajectory $x(t)$ will be expected to spend a disproportionate amount of time in a small neighborhood of $\overline{x}$. In this way it is possible to construct suboptimal controller with ideal performance but at significantly reduced complexity.

For simplicity we make the *Symmetry Assumption*: all the users have i.i.d. channel dynamics with $a_i = a$, $b_i = b$, $\sigma_i = \sigma$ and $\gamma_i = \gamma$, $R = rI_n$. We use $K(x) = (K_{ij}(x))_{i,j=1}^n$ to denote the solution of the Riccati equation (5.40) and write

$$\begin{aligned} K_{ij}(x) &= K_{ij}(\overline{x}) + (x-\overline{x})^\tau \frac{\partial K_{ij}(\overline{x})}{\partial x} + \frac{1}{2}(x-\overline{x})^\tau \frac{\partial^2 K_{ij}(\overline{x})}{\partial x^2}(x-\overline{x}) + o(|x-\overline{x}|^2) \\ &\triangleq K_{ij}(\overline{x}) + (x-\overline{x})^\tau K_{ij}^{(1)}(\overline{x}) + \frac{1}{2}(x-\overline{x})^\tau K_{ij}^{(2)}(\overline{x})(x-\overline{x}) + o(|x-\overline{x}|^2), \quad (5.44) \end{aligned}$$

where $\overline{x} = (b, \cdots, b)^\tau$ is the steady state mean of the attenuation vector.

5.4.1. Complexity of the Local Expansion of the Matrix $K(x)$. In the following we will show that in the symmetric case, when n grows the complexity of the local polynomial approximation does not increase with the dimension, i.e., the total number of distinct entries of the three coefficient matrices in (5.44) does not increase with n. To this end, we first show an important property of the entries of the Riccati matrix $K(x)$. For the ordered integers $I = (1, 2, \cdots, n)$, let $\overline{I} = (i_1, i_2, \cdots, i_n)$ be an arbitrary permutation of I. For $1 \le j \le n$, suppose j is the $s(j)$-th element in the row $\overline{I}$.

Proposition 5.2. Under the *Symmetry Assumption*, for the Riccati matrix $K(x)$, we have

$$K_{ij}(x_1, x_2, \cdots, x_n) = K_{s(i)s(j)}(x_{i_1}, x_{i_2}, \cdots, x_{i_n}). \quad (5.45)$$

PROOF. By the symmetry assumption on the channel dynamics and the cost function, we have

$$v(x_1, x_2, \cdots, x_n, p_1, p_2, \cdots, p_n) = v(x_{i_1}, x_{i_2}, \cdots, x_{i_n}, p_{i_1}, p_{i_2}, \cdots, p_{i_n}). \quad (5.46)$$

Using Theorem 5.2 and comparing the coefficients of the quadratic terms of p, we get the above relation (5.45). □

Using Proposition 5.2 repeatedly, we can verify that the computation of each K_{ij} and its first, second order partial derivatives at $\overline{x} = (b, \cdots, b)^{\tau}$ can be reduced to K_{11}, K_{12} and their derivatives at $\overline{x}$. In fact, these first and second order derivatives at $\overline{x}$ can be represented by 13 variables. Here we list all the distinct entries in the three matrices of (5.44) as follows:

$$\begin{aligned}
&K_{11},\ K_{12},\ \frac{\partial K_{11}}{\partial x_1},\ \frac{\partial K_{11}}{\partial x_2},\ \frac{\partial^2 K_{11}}{\partial {x_1}^2},\ \frac{\partial^2 K_{11}}{\partial {x_2}^2},\ \frac{\partial^2 K_{11}}{\partial x_1 \partial x_2},\ \frac{\partial^2 K_{11}}{\partial x_2 \partial x_3},\\
&\frac{\partial K_{12}}{\partial x_1} = \frac{\partial K_{12}}{\partial x_2},\ \frac{\partial K_{12}}{\partial x_3},\ \frac{\partial^2 K_{12}}{\partial {x_3}^2},\ \frac{\partial^2 K_{12}}{\partial x_1 \partial x_2},\ \frac{\partial^2 K_{12}}{\partial {x_1}^2} = \frac{\partial^2 K_{12}}{\partial {x_2}^2},\\
&\frac{\partial^2 K_{12}}{\partial x_1 \partial x_3} = \frac{\partial^2 K_{12}}{\partial x_2 \partial x_3},\ \frac{\partial^2 K_{12}}{\partial x_3 \partial x_4},
\end{aligned}$$

where all the quantities are computed at $\overline{x} = (b, \cdots, b)^{\tau}$. Here we verify the last identity. In fact, by (5.45) and the transpose symmetry of K, i.e., $K(x) = K^{\tau}(x)$, we have

$$\begin{aligned}
\left.\frac{\partial^2 K_{12}}{\partial x_2 \partial x_3}\right|_{\overline{x}} &= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon^2}[K_{12}(\overline{x}_1, \overline{x}_2 + \varepsilon, \overline{x}_3 + \varepsilon, \overline{x}_4, \cdots) - K_{12}(\overline{x}_1, \overline{x}_2, \overline{x}_3 + \varepsilon, \overline{x}_4, \cdots)\\
&\qquad - K_{12}(\overline{x}_1, \overline{x}_2 + \varepsilon, \overline{x}_3, \overline{x}_4, \cdots) + K_{12}(\overline{x}_1, \overline{x}_2, \overline{x}_3, \overline{x}_4, \cdots)]\\
&= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon^2}[K_{12}(\overline{x}_1 + \varepsilon, \overline{x}_2, \overline{x}_3 + \varepsilon, \overline{x}_4, \cdots) - K_{12}(\overline{x}_1, \overline{x}_2, \overline{x}_3 + \varepsilon, \overline{x}_4, \cdots)\\
&\qquad - K_{12}(\overline{x}_1 + \varepsilon, \overline{x}_2, \overline{x}_3, \overline{x}_4, \cdots) + K_{12}(\overline{x}_1, \overline{x}_2, \overline{x}_3, \overline{x}_4, \cdots)]\\
&= \left.\frac{\partial^2 K_{12}}{\partial x_1 \partial x_3}\right|_{\overline{x}},
\end{aligned}$$

where $\overline{x} = (\overline{x}_1, \cdots, \overline{x}_n)^{\tau} = (b, \cdots, b)^{\tau}$. The remarkable feature of limited complexity for $K(x)$ as well as its first and second order derivatives at $\overline{x}$ is potentially useful to obtain simple and efficient local approximation for $K(x)$ by determining the above 15 unknowns.

5.4.2. The Approximating Equation System. We write the Riccati equation (5.40) in terms of its components to obtain

$$\rho K_{ij}(x) = \frac{\sigma^2}{2}\sum_k \frac{\partial^2 K_{ij}(x)}{\partial {x_k}^2} + \sum_k f_k \frac{\partial K_{ij}(x)}{\partial x_k} - \sum_k \frac{1}{r} K_{ik}(x) K_{kj}(x) + C_{ij}(x). \quad (5.47)$$

And using (5.44) we write the system of approximate equations (up to second order)

$$\begin{aligned}
&\rho K_{ij}(\overline{x}) + \rho(x-\overline{x})^\tau K_{ij}^{(1)}(\overline{x}) + \frac{\rho}{2}(x-\overline{x})^\tau K_{ij}^{(2)}(\overline{x})(x-\overline{x}) \\
&= \frac{\sigma^2}{2}\sum_k K_{ij,k}^{(2)}(\overline{x}) + \sum_k [-a(x_k-\overline{x}_k)]\ [K_{ij,k}^{(1)}(\overline{x}) + K_{ij,k(\cdot)}^{(2)}(\overline{x})(x-\overline{x})\ ] \\
&\quad - \sum_k \frac{1}{r}[\ K_{ik}(\overline{x}) + (x-\overline{x})^\tau K_{ik}^{(1)}(\overline{x}) + \frac{1}{2}(x-\overline{x})^\tau K_{ik}^{(2)}(\overline{x})(x-\overline{x})]\cdot \\
&\qquad [K_{kj}(\overline{x}) + (x-\overline{x})^\tau K_{kj}^{(1)}(\overline{x}) + \frac{1}{2}(x-\overline{x})^\tau K_{kj}^{(2)}(\overline{x})(x-\overline{x})] \\
&\quad + C_{ij}(\overline{x}) + (x-\overline{x})^\tau C_{ij}'(\overline{x}) + \frac{1}{2}(x-\overline{x})^\tau C_{ij}''(\overline{x})(x-\overline{x}),
\end{aligned} \quad (5.48)$$

where $K_{ij,k}^{(2)}(\overline{x})$, $K_{ij,k(\cdot)}^{(2)}(\overline{x})$ are the k-th diagonal entry and the k-th row of the matrix $K_{ij}^{(2)}(\overline{x})$, respectively, and $K_{ij,k}^{(1)}(\overline{x})$ is the k-th entry of $K_{ij}^{(1)}(\overline{x})$. Notice that in writing the equation (5.48) only the first three terms in (5.44) are formally substituted into (5.47) and the higher order terms are neglected. When the higher order terms are taken into account, additional terms of the order $|\frac{\sigma^2}{2}K_{ij}^{(3)}|$ and $|\frac{\sigma^2}{4}K_{ij}^{(4)}|$ will appear in equations (5.50) and (5.51) below, respectively, where $K_{ij}^{(3)}$ and $K_{ij}^{(4)}$ denote the third and fourth order mixing partial derivatives of $K_{ij}(x)$ at $\overline{x}$ assuming their existence. Here in order to avoid an infinitely coupled equation system we omit the additional terms but maintain sufficiently close approximation to the exact solution since we are considering the small noise case.

However we write an exact equation corresponding to the zero order term since it has more weight in the suboptimal control law when the state stays in a small neighborhood of $\overline{x}$. Grouping terms with zero power of $(x-\overline{x})$ in (5.48), we obtain the equation system

$$\rho K_{ij}(\overline{x}) = \frac{\sigma^2}{2}\sum_k K^{(2)}_{ij,k}(\overline{x}) - \sum_k \frac{1}{r}K_{ik}(\overline{x})K_{kj}(\overline{x}) + C_{ij}(\overline{x}), \tag{5.49}$$

or equivalently, in the matrix form

$$\rho K(\overline{x}) = \frac{\sigma^2}{2}\Big(\ \mathrm{tr}\{K^{(2)}_{ij}(\overline{x})\}\ \Big)^n_{i,j=1} - \frac{1}{r}K(\overline{x})K(\overline{x}) + C(\overline{x}),$$

which takes the form of a perturbed algebraic Riccati equation. By (5.48) we also have

$$\begin{aligned}(x-\overline{x})^\tau \rho K^{(1)}_{ij}(\overline{x}) &= (x-\overline{x})^\tau\Big(-aK^{(1)}_{ij,k}(\overline{x})\Big)^n_{k=1}\\ &- \sum_k (x-\overline{x})^\tau \frac{1}{r}[\ K^{(1)}_{ik}(\overline{x})K_{kj}(\overline{x}) + K_{ik}(\overline{x})K^{(1)}_{kj}(\overline{x})\ ] + (x-\overline{x})^\tau C'_{ij}(\overline{x}),\end{aligned}$$

which gives

$$(\rho + a)K^{(1)}_{ij}(\overline{x}) = -\frac{1}{r}\sum_k [\ K^{(1)}_{ik}(\overline{x})K_{kj}(\overline{x}) + K_{ik}(\overline{x})K^{(1)}_{kj}(\overline{x})\ ] + C'_{ij}(\overline{x}). \tag{5.50}$$

Finally, by inspecting the second order terms in (5.48) we get

$$\begin{aligned}(\frac{\rho}{2}+a)K^{(2)}_{ij}(\overline{x}) = &-\frac{1}{2r}\sum_k [\ K_{ik}(\overline{x})K^{(2)}_{kj}(\overline{x}) + K^{(2)}_{ik}(\overline{x})K_{kj}(\overline{x})\ ]\\ &-\frac{1}{r}\sum_k K^{(1)}_{ik}(\overline{x})[K^{(1)}_{kj}(\overline{x})]^\tau + \frac{1}{2}C''_{ij}(\overline{x}).\end{aligned} \tag{5.51}$$

It would be of interest to investigate the procedure to solve the above equation system numerically, which is an important step toward implementing the suboptimal control law in a simple and efficient manner. We will design the numerical procedure to solve the equation system (5.49)-(5.51) below.

5.4.3. A Recursive Algorithm. In this part we design a recursive algorithm to solve the equation system (5.49)-(5.51). To achieve good numerical stability and convergence for the recursive algorithm the coupled polynomial equations are rewritten. First, we write (5.50) in the following form:

$$[\ \rho + a + \frac{K_{ii} + K_{jj}}{r}\ ]K_{ij}^{(1)}(\overline{x})$$
$$= -\frac{1}{r}\sum_{k\neq i,j}[\ K_{ik}^{(1)}(\overline{x})K_{kj}(\overline{x}) + K_{ik}(\overline{x})K_{kj}^{(1)}(\overline{x})\ ] + C'_{ij}(\overline{x}), \tag{5.52}$$

and (5.51) is written as

$$[\ \frac{\rho}{2} + a + \frac{K_{ii} + K_{jj}}{2r}\ ]K_{ij}^{(2)}(\overline{x}) = -\frac{1}{2r}\sum_{k\neq i,j}[\ K_{ik}(\overline{x})K_{kj}^{(2)}(\overline{x}) + K_{ik}^{(2)}(\overline{x})K_{kj}(\overline{x})\ ]$$
$$-\frac{1}{r}\sum_{k} K_{ik}^{(1)}(\overline{x})[K_{kj}^{(1)}(\overline{x})]^{\tau} + \frac{1}{2}C''_{ij}(\overline{x}). \tag{5.53}$$

where K_{ii} and K_{jj} are computed at $\overline{x}$. Denote

$$A_{ij} = \rho + a + \frac{K_{ii} + K_{jj}}{r}, \quad B_{ij} = \frac{\rho}{2} + a + \frac{K_{ii} + K_{jj}}{2r}.$$

Assuming $A_{ij} \neq 0$ and $B_{ij} \neq 0$, from (5.52)-(5.53) we write

$$K_{ij}^{(1)}(\overline{x}) = -\frac{1}{rA_{ij}}\sum_{k\neq i,j}[\ K_{ik}^{(1)}(\overline{x})K_{kj}(\overline{x}) + K_{ik}(\overline{x})K_{kj}^{(1)}(\overline{x})\ ] + \frac{1}{A_{ij}}C'_{ij}(\overline{x})$$
$$\triangleq \Psi_{ij}(K, K^{(1)}), \tag{5.54}$$
$$K_{ij}^{(2)}(\overline{x}) = -\frac{1}{2rB_{ij}}\sum_{k\neq i,j}[\ K_{ik}(\overline{x})K_{kj}^{(2)}(\overline{x}) + K_{ik}^{(2)}(\overline{x})K_{kj}(\overline{x})\ ]$$
$$-\frac{1}{rB_{ij}}\sum_{k} K_{ik}^{(1)}(\overline{x})[K_{kj}^{(1)}(\overline{x})]^{\tau} + \frac{1}{2B_{ij}}C''_{ij}(\overline{x}) \triangleq \Phi_{ij}(K, K^{(1)}, K^{(2)}), \tag{5.55}$$

where $K^{(1)}$, $K^{(2)}$ denote the sequence of $K_{ij}^{(1)}(\overline{x})$, $K_{ij}^{(2)}(\overline{x})$, $1 \leq i, j \leq n$, with any predetermined order, respectively. For (5.49) we assume a unique positive definite solution K exists for $K^{(2)}$ varying in a certain range and we indicate the dependence by

$$K = \Lambda(K^{(2)}). \tag{5.56}$$

To approximate the solution to the equation system (5.49)-(5.51), we first set the initial condition $K_0, K_0^{(1)}, K_0^{(2)}$ where the subscript is used to indicate the time instant, and then we update the solution by the following scheme:

$$K_{t+1} = \Lambda(K_t^{(2)}), \tag{5.57}$$

$$(K_{ij}^{(1)})_{t+1} = \Psi_{ij}(K_{t+1}, K_t^{(1)}), \qquad 1 \leq i, j \leq n, \tag{5.58}$$

$$(K_{ij}^{(2)})_{t+1} = \Phi_{ij}(K_{t+1}, K_{t+1}^{(1)}, K_t^{(2)}), \qquad 1 \leq i, j \leq n, \tag{5.59}$$

where $t \geq 0$.

To illustrate the efficiency of the recursive algorithm we apply (5.57)-(5.59) to system A of Section 5.3.2 with the same cost function as specified in Section 5.3.2. Since the two mobiles have i.i.d. dynamics, we adopt a certain symmetry with the initialization of the algorithm. Specifically, we take $K_0 = I_{2\times 2}$ and all the first and second order derivatives of K take the initial condition of zero matrices or vectors of suitable dimension. Figure 5.3 below demonstrates the asymptotic behavior of the iteration of (5.57)-(5.59) from step 2 to step 20. Interpretation of the entries in the plot can be found in Section 5.4.2.

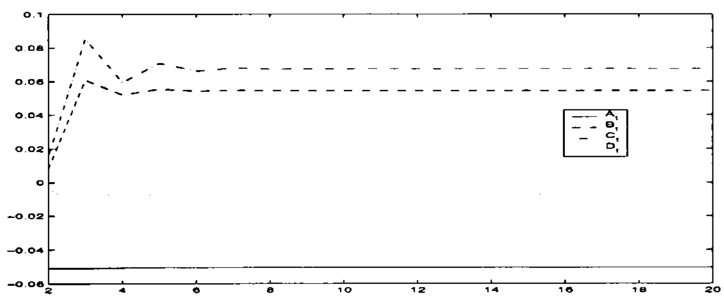


FIGURE 5.3. $A_t = K_{12}(t)$, $B_t = K_{11,1}^{(1)}(t)$, $C_t = K_{11,1}^{(2)}(t)$, $D_t = K_{12,1}^{(2)}(t)$

Remark 5.1. For this two dimensional example, it is worth mentioning an important feature for the operators on the right hand side of (5.58) and (5.59). To begin with, by removing the leading term on the right hand side of (5.49) we get a usual Riccati equation for which we obtain a so-called nominal solution $\widehat{K}$. Let $K_t^{(1)}$ denote the following composite vector

$$K_t^{(1)} \triangleq [K_{11}^{(1)^\tau}(t), K_{12}^{(1)^\tau}(t), K_{21}^{(1)^\tau}(t), K_{22}^{(1)^\tau}(t)]^\tau.$$

We replace K_{t+1} by its nominal value $\widehat{K}$ in (5.58). Now (5.58) can be written as

$$K_{t+1}^{(1)} = \Psi_1 K_t^{(1)} + \Psi_2, \tag{5.60}$$

where $\Psi_1 \in \mathrm{R}^{8\times 8}$ is determined by $\widehat{K}$ and $\Psi_2 \in \mathrm{R}^8$ is a constant vector depending on $\widehat{K}$ and $C'(\overline{x})$. For systems A and the associated cost function we can verify that the absolute value of each eigenvalue of Ψ_1 is less than 0.37 and hence Ψ_1 is stable. We can also verify a similar stabilizing property for the map (5.59) when K_{t+1} is replaced by its nominal value $\widehat{K}$. □

We write the results of the recursion in the following compact form where the determination of each block is evident.

$$K = \begin{bmatrix} 0.070945 & -0.050494 \\ -0.050494 & 0.070945 \end{bmatrix}, \tag{5.61}$$

$$\begin{bmatrix} K_{11}^{(1)} & K_{12}^{(1)} \\ K_{21}^{(1)} & K_{22}^{(1)} \end{bmatrix} = \begin{bmatrix} 0.054577 & -0.019913 \\ -0.007263 & -0.019913 \\ -0.019913 & -0.007263 \\ -0.019913 & 0.054573 \end{bmatrix}, \tag{5.62}$$

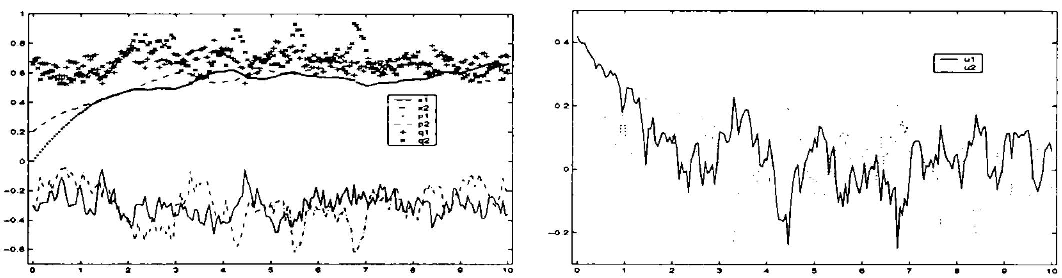
$$\begin{bmatrix} K_{11}^{(2)} & K_{12}^{(2)} \\ K_{21}^{(2)} & K_{22}^{(2)} \end{bmatrix} = \begin{bmatrix} 0.067471 & -0.004126 & -0.007098 & -0.010009 \\ -0.004126 & -0.004070 & -0.022420 & -0.007098 \\ -0.007098 & -0.022420 & -0.004070 & -0.004126 \\ -0.010009 & -0.007098 & -0.004126 & 0.067471 \end{bmatrix}. \tag{5.63}$$

Notice that the resulting matrices $K_{12}^{(2)}$ and $K_{21}^{(2)}$ are not symmetric. This is not unusual since the equations (5.49)-(5.51) are obtained by a certain approximation technique (with a noise level $\sigma^2 = 0.09$) and it may not provide tight approximation to certain entries in the second order Taylor coefficient matrices for $K(x)$ in (5.44). With K, $K^{(1)}$, $K^{(2)}$ given by (5.61)-(5.63), applying the local expansion method of Section 5.4.2 to the equation (5.41) we can also determine the approximate values of the associated coefficients $S(\overline{x})$, $S^{(1)}(\overline{x})$, $S^{(2)}(\overline{x})$, the computation of which is simple since they are described by a set of linear equations. We have

$$S = \begin{bmatrix} -0.012523 \\ -0.012523 \end{bmatrix}, \qquad [S_1^{(1)}\ S_2^{(1)}] = \begin{bmatrix} -0.000485 & -0.001771 \\ -0.001771 & -0.000485 \end{bmatrix},$$

$$[S_1^{(2)}\ S_2^{(2)}] = \begin{bmatrix} 0.000834 & -0.000209 & -0.000438 & -0.001290 \\ -0.001290 & -0.000438 & -0.000209 & 0.000834 \end{bmatrix}.$$

Now we substitute the local second order approximation of $K(x)$ and $S(x)$ into the controller (5.43) to get a nearly optimal controller. The following Figure 5.4 is the simulation of this local expansion based controller for system A in Section 5.3.2. The variables x_i (power attenuation), p_i (controlled power), q_i (static pointwise optimum) in Figure 5.4 (a) are specified in the same way as in Section 5.3.2. Figure 5.4 (b) demonstrates the control inputs of two mobiles. In Figures 5.4 and 5.1, the basic behaviour of the power adjustment is quite similar and in both cases the powers of the two users will gradually be brought to certain stable level.



(a) $p_1(0) = 0,\ p_2(0) = 0.2$ (b) u_1 and u_2

FIGURE 5.4. Simulation for system A using the nearly optimal control law

5.4.4. The Single User Case. In this part we consider the very simple example of $n = 1$. This corresponds to the case of a single mobile in service under the effect of a fading channel and the background noise. However we mention that the solution is useful to construct nearly optimal control laws in systems with large populations where a particular mobile M is singled out for analysis and the scaled interference generated from all the other users can be approximated by a deterministic quantity due to the effect of the law of large numbers. For $n = 1$, we have

$$C(\overline{x}) = (1-\mu)^2 e^{-2b}, \quad C'(\overline{x}) = 2(1-\mu)^2 e^{-2b}, \quad C''(\overline{x}) = 4(1-\mu)^2 e^{-2b},$$

and the equations (5.49)-(5.51) reduce to

$$\rho K = \frac{\sigma^2}{2} K^{(2)} - \frac{1}{r} K^2 + C, \tag{5.64}$$

$$(\rho + a) K^{(1)} = -\frac{2}{r} K K^{(1)} + C', \tag{5.65}$$

$$(\frac{\rho}{2} + a) K^{(2)} = -\frac{1}{r} K K^{(2)} - \frac{1}{r} K^{(1)} K^{(1)} + \frac{1}{2} C'', \tag{5.66}$$

where C, C' and C'' take their values at $\overline{x}$. In the following we seek a solution for the small noise case satisfying $K \geq 0$.

Proposition 5.3. There exists $\sigma_0^2 > 0$ such that for any finite $\sigma^2 \leq \sigma_0^2$ the equation system (5.64)-(5.66) has a solution $(\overline{K}, \overline{K}^{(1)}, \overline{K}^{(2)})$ and $\overline{K} \geq 0$.

PROOF. Rewriting the system (5.64)-(5.66) yields

$$K = \frac{-r\rho + \sqrt{r^2\rho^2 + 2r(\sigma^2 K^{(2)} + 2C)}}{2} \triangleq G_0(K^{(2)}), \tag{5.67}$$

$$K^{(1)} = \frac{rC'}{r(\rho + a) + 2K} \triangleq G_1(K), \tag{5.68}$$

$$K^{(2)} = \frac{rC'' - 2K^{(1)}K^{(1)}}{r(\rho + 2a) + 2K} \triangleq G_2(K, K^{(1)}). \tag{5.69}$$

We introduce four constants

$$c_1 \triangleq \frac{C'}{\rho + a}, \qquad c_2 \triangleq \frac{C''}{\rho + 2a},$$

$$c_0 \triangleq \frac{-r\rho + \sqrt{r^2\rho^2 + 2r(\sigma^2 c_2 + 2C)}}{2},$$

$$c_2^- \triangleq \inf_{0 \leq s \leq G_0(c_2)} \frac{rC'' - 2c_1^2}{r(\rho + 2a) + 2s},$$

and a convex compact subset of R^3 denoted by

$$\mathcal{K} \triangleq \{(x_0, x_1, x_2) : \ 0 \leq x_i \leq c_i, \ i = 0, 1 \text{ and } c_2^- \leq x_2 \leq c_2\}. \tag{5.70}$$

Set $\sigma_0^2 = \sup\{\sigma^2 : \sigma^2 c_2^- + 2C \geq 0\}$. Then for any $\sigma^2 \leq \sigma_0^2$, the square root in (5.67) is always no less than $r\rho$ for $c_2^- \leq K^{(2)} \leq c_2$. We define the continuous map G on $\mathcal{K}$ such that

$$G(K, K^{(1)}, K^{(2)}) = (\ G_0(K^{(2)}),\ G_1(K),\ G_2(K, K^{(1)})\ ). \tag{5.71}$$

Then it is readily verified that $G(\mathcal{K}) \subseteq \mathcal{K}$ and therefore, by Brouwer's fixed point theorem G has a fixed point $(\overline{K}, \overline{K}^{(1)}, \overline{K}^{(2)})$. From (5.67) it is seen that $\overline{K} \geq 0$. Thus we have proved that the system (5.64)-(5.66) has a solution $(\overline{K}, \overline{K}^{(1)}, \overline{K}^{(2)})$ and $\overline{K} \geq 0$. □

We proceed to consider the local approximation of $S(x)$ in Section 5.3.1. We write

$$S(x) = S(\overline{x}) + S^{(1)}(\overline{x})(x - \overline{x}) + \frac{1}{2}S^{(2)}(\overline{x})(x - \overline{x})^2 + o(|x - \overline{x}|^2).$$

Then similar to the treatment for $K(x)$, from (5.41) we obtain a system of algebraic equations

$$(\rho + \frac{\overline{K}}{r})S - \frac{\sigma^2}{2}S^{(2)} = D, \tag{5.72}$$

$$\frac{\overline{K}^{(1)}}{r}S + (a + \rho + \frac{\overline{K}}{r})S^{(1)} = D', \tag{5.73}$$

$$\frac{\overline{K}^{(2)}}{2r}S + \frac{\overline{K}^{(1)}}{r}S^{(1)} + (\frac{\rho}{2} + a + \frac{\overline{K}}{2r})S^{(2)} = \frac{1}{2}D'', \tag{5.74}$$

where $D = D' = D'' = -\mu\eta(1-\mu)e^{-b}$.

Example 5.1. For $n = 1,\ a = 2,\ b = 0.3,\ \sigma^2 = 0.01,\ \mu = 0.6,\ \eta = 0.25,\ \rho = 0.5,\ r = 0.1$, we have

$$(\overline{K}, \overline{K}^{(1)}, \overline{K}^{(2)}) = (\ 0.072120,\ 0.044547,\ 0.052429\ ),$$

$$(\overline{S}, \overline{S}^{(1)}, \overline{S}^{(2)}) = (\ -0.036412,\ -0.008763,\ -0.003362\ ).$$

□

Example 5.2. For $n = 1,\ a = 2,\ b = 0.4,\ \sigma^2 = 0.01,\ \mu = 0.6,\ \eta = 0.25,\ \rho =$ 0.5, $r = 0.1$, we have

$$(\overline{K}, \overline{K}^{(1)}, \overline{K}^{(2)}) = (\ 0.063525,\ 0.038134,\ 0.044794\ ),$$
$$(\overline{S}, \overline{S}^{(1)}, \overline{S}^{(2)}) = (\ -0.035443,\ -0.008517,\ -0.003475\ ).$$

□

Remark 5.2. For $x = (x_1, x_2, x_3) \in \mathrm{R}^3$, we define $\|x\| = \max_i |x_i|$. By examining the upper bounds for $|\frac{\partial G_j}{\partial x_i}|$ on $\mathcal{K}$, $j = 0, 1, 2,\ i = 1, 2, 3$, where $\mathcal{K}$ and G are defined by (5.70) and (5.71), we can show that in Examples 5.1 and 5.2, the map G is a contraction on $\mathcal{K}$ under the norm $\|\cdot\|$. In this case the unique solution of (5.64)-(5.66) can be approximated iteratively. □

By substituting the local second order polynomial approximation of $K(x)$ and $S(x)$ into the feedback control (5.43), the suboptimal control law for the single user is determined as

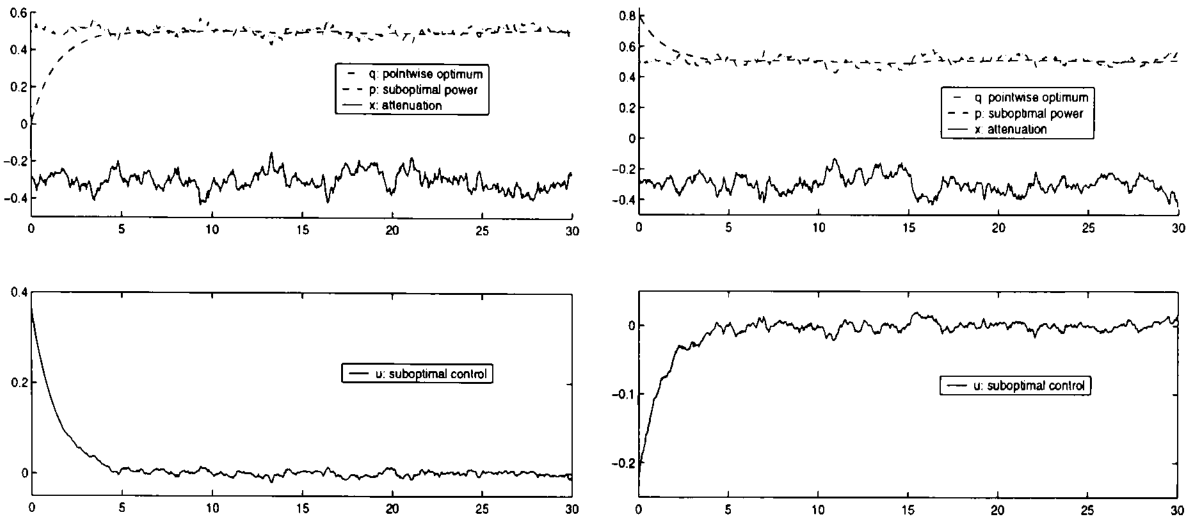
$$\begin{aligned} u = &-\frac{1}{r}[\overline{K} + \overline{K}^{(1)}(x+b) + \frac{1}{2}\overline{K}^{(2)}(x+b)^2]p \\ &-\frac{1}{r}[\overline{S} + \overline{S}^{(1)}(x+b) + \frac{1}{2}\overline{S}^{(2)}(x+b)^2]. \end{aligned} \tag{5.75}$$

From (5.75) we write the 0-th order approximation of the optimal control law as $u^{(0)} = -\frac{\overline{K}}{r}p - \frac{\overline{S}}{r}$, for which the steady state power is $p^\infty = -\frac{\overline{S}}{\overline{K}}$. On other other hand, we determine the nominal power level $\overline{p}$ by setting $e^{\overline{x}}\overline{p} - \mu(e^{\overline{x}}\overline{p} + \eta) = 0$, and define the relative error between p^∞ and $\overline{p}$ by $Err(p^\infty, \overline{p}) \triangleq \frac{|p^\infty - \overline{p}|}{\overline{p}}$. For Examples 5.1 and 5.2, we have

Example	p^∞	$\overline{p}$	$Err(p^\infty, \overline{p})$
1	0.504881	0.506197	$< 0.3\%$
2	0.557938	0.559434	$< 0.3\%$

The following is the simulation of the suboptimal control law given by (5.75). The pointwise optimum q is determined by setting $\frac{qe^x}{qe^x+\eta}(t) = \mu$ for each $t \geq 0$. Figure 5.5

demonstrates the dynamic behavior of the system in Example 5.1 with two different initial powers.



(a): $p(0) = 0$ (b): $p(0) = 0.8$

FIGURE 5.5. Left (a) and Right (b): The trajectories of attenuation x, power p and control u with initial power 0 and 0.8, respectively

5.5. Application of the Single User Based Design to Systems with Large Populations

In this section we study the power control problem in a large population context. In this case the Quality of Service (QoS) measure needs to be suitably scaled. For instance, in (2.7), to insure solvability of the static problem (i.e., there exists at least one positive power vector satisfying (2.5)), one can diminish μ_i toward 0 as $n \longrightarrow \infty$. However, here we shall not follow this scaling. The QoS measure of this Section is based on the SIR of users after matched filtering in CDMA systems. Specifically we seek to have

$$\frac{\widehat{p}_i}{\sum_{j\neq i}^{n} \beta_{ji}\widehat{p}_j + \eta} \approx \frac{\widehat{p}_i}{\sum_{j=1}^{n} \beta_{ji}\widehat{p}_j + \eta} \approx \mu_i, \quad 1 \leq i \leq n,$$

in some sense, where the received power $\widehat{p}_i = e^{x_i}p_i$, and $\beta_{ji} = (\mathbf{s}_j^\tau \mathbf{s}_i)^2, j \neq i$, is the crosscorrelation between the signature sequences $\mathbf{s}_j, \mathbf{s}_i$ of length n_s for users j, i, respectively. Following [**72, 74, 80**], we make the standard assumption that $\frac{n}{n_s} \to \alpha > 0$, as $n \to \infty$. By appropriately choosing random signature sequences of length n_s, one can get $\beta_{ji} \approx \frac{1}{n_s}$ [**72, 80**], and hence $\beta_{ji} \approx \frac{\alpha}{n}$. Here for simplicity we take $\beta_{ji} = \frac{1}{n} \triangleq \beta_n$.

Hence we write the following static nominal equation

$$e^{x_i} p_i = \mu_i (\beta_n \sum_{j=1}^{n} e^{x_j} p_j + \eta), \tag{5.76}$$

which is equivalent to

$$e^{x_i} p_i = \mu_i [e^{x_i} p_i + (1 - \mu_i)(\beta_n \sum_{j=1}^{n} e^{x_j} p_j + \eta)].$$

We set $\widetilde{\eta}(x,p) = \beta_n \sum_{j=1}^{n} e^{x_j} p_j + \eta$ and term $\widetilde{\eta}(x,p)$ as the *network interference index*. Notice that under mild conditions on the distribution of x_i, p_i, $1 \leq i \leq n$, $\widetilde{\eta}(x,p)$ has a small variance compared to its mean, and thus can be approximated by a deterministic constant at each fixed time instant. In particular, when most users in the system are in stable working conditions, for analysis of a small group of newly admitted users the variation of the network interference index with respect to time is negligible.

The above analysis suggests we write an individual cost function for the i-th mobile

$$J_i = E \int_0^\infty e^{-\rho t} \{ [e^{x_i} p_i - \mu_i (e^{x_i} p_i + (1 - \mu_i) \widetilde{\eta})]^2 + r u_i^2 \} dt. \tag{5.77}$$

In this setup the i-th mobile is singled out from the large population to deal with an interference generated by all other mobiles and the true background noise. In essense, the i-th user is treated as the only user of a "virtual system" with an "equivalent time-varying background noise intensity". Then on a time interval $[kT, (k+1)T)$, $T > 0$, $k = 0, 1, 2, \cdots$, the control law u_i of the i-th user is determined by (5.75) using $(1 - \mu_i)\widetilde{\eta}$ as a time-varying parameter for the equivalent background noise intensity. In the implementation, $\widetilde{\eta}$ is replaced by its measurement at $t = kT$ and updated at

$t = (k+1)T$. We construct the control law of the i-th mobile as

$$u_i = -\frac{1}{r}[\overline{K} + \overline{K}^{(1)}(x_i + b_i) + \frac{1}{2}\overline{K}^{(2)}(x_i + b_i)^2]p_i - \frac{(1-\mu_i)\widetilde{\eta}_0}{r\eta}[\overline{S} + \overline{S}^{(1)}(x_i + b_i) + \frac{1}{2}\overline{S}^{(2)}(x_i + b_i)^2], \quad (5.78)$$

where $(\overline{S}, \overline{S}^{(1)}, \overline{S}^{(2)})$ is the solution of (5.72)-(5.74) corresponding to the constant η, and $\widetilde{\eta}_0 = \widetilde{\eta}(x,p)(kT)$ for $t \in [kT, (k+1)T)$. Compared to (5.75), the second term in (5.78) contains the factor $\frac{(1-\mu_i)\widetilde{\eta}_0}{\eta}$ since in the present case the original η is replaced by $(1-\mu_i)\widetilde{\eta}_0$ in (5.77) and $(\overline{S}, \overline{S}^{(1)}, \overline{S}^{(2)})$ depends on η linearly as indicated by (5.72)-(5.74) . Here $(\overline{K}, \overline{K}^{(1)}, \overline{K}^{(2)})$ is independent of η.

Notice that the control law (5.77) is partially decentralized since for each user it depends only on its own state and the network interference index $\widetilde{\eta}$ to be measured by the base station. In fact, $\widetilde{\eta}$ is the sum of the scaled total received power and the background noise intensity.

Remark 5.3. Assuming all the users start from zero power, an initial increase of powers of all users leads to a higher network interference index, which in turn requires a further increase of individual powers. This gives rise to the question whether there would be an unlimited growth of individual powers. To a large extent, this question is related to stability of the power updating scheme. In fact, by examining the 0-th order approximate control law induced from (5.78) we see that under very mild conditions on the coefficients the corresponding network interference index has a stable behavior after successive iterations of powers along the steady state $\overline{x}$. □

5.5.1. Simulation Examples. The following simulation shown in Figure 5.6 is for a system of 140 users. For all users we take $\mu_i = 0.6$ in (5.78). Parameters for each half (i.e., 70) of the users are as specified in Examples 5.1 and 5.2, respectively. User 1 has an initial power 0. The initial powers of other users are distributed in a small neighborhood of 0.2.

For this simulation we label the users specified in Examples 5.1 and 5.2, respectively, by 1-70 (Population 1) and 71-140 (Population 2), respectively. We take

$n = 140$ and use the attenuation, power and control variables x, p, $u \in \mathrm{R}^n$ to describe all users. Here we give an informal stability analysis for the individual cost based control law (5.78) for which we write the so-called nominal equation as follows:

$$u_i = -\frac{\overline{K}_i}{r}p_i - \frac{(1-\mu_i)\overline{S}_i}{r\eta}\left(\frac{1}{n}\sum_{i=1}^{n} e^{-b_i}p_i + \eta\right), \tag{5.79}$$

where $1 \leq i \leq 70$ for Population 1 and $71 \leq i \leq 140$ for Population 2. For instance, if user i is in Population 1 (i.e., $1 \leq i \leq 70$) then all the parameters in (5.79) are determined from Example 5.1; similarly for Population 2 in Example 5.2. We further write the nominal equation for the closed-loop power adjustment as

$$\frac{dp}{dt} = u = \Lambda_1 p + \Lambda_2, \tag{5.80}$$

where $\Lambda_1 \in \mathrm{R}^{n\times n}$, $\Lambda_2 \in \mathrm{R}^n$ are easily determined by (5.79) and Λ_1 can be written in the form

$$\Lambda_1 = \begin{pmatrix} \alpha+\delta_1 & \cdots & \alpha & \alpha & \cdots & \alpha \\ \vdots & & & \vdots & & \\ \alpha & \cdots & \alpha+\delta_1 & \alpha & \cdots & \alpha \\ \beta & \cdots & \beta & \beta+\delta_2 & \cdots & \beta \\ \vdots & & & \vdots & & \\ \beta & \cdots & \beta & \beta & \cdots & \beta+\delta_2 \end{pmatrix}_{n\times n},$$

where $n = 140$, $\alpha = 0.003083$, $\beta = 0.002715$, $\delta_1 = -0.7212$ and $\delta_2 = -0.6352$. The eigenvalues of Λ_1 are given by

$$\lambda_3 = \lambda_4 = \cdots = \lambda_{71} = \delta_1,$$

$$\lambda_{72} = \lambda_{73} = \cdots = \lambda_{140} = \delta_2,$$

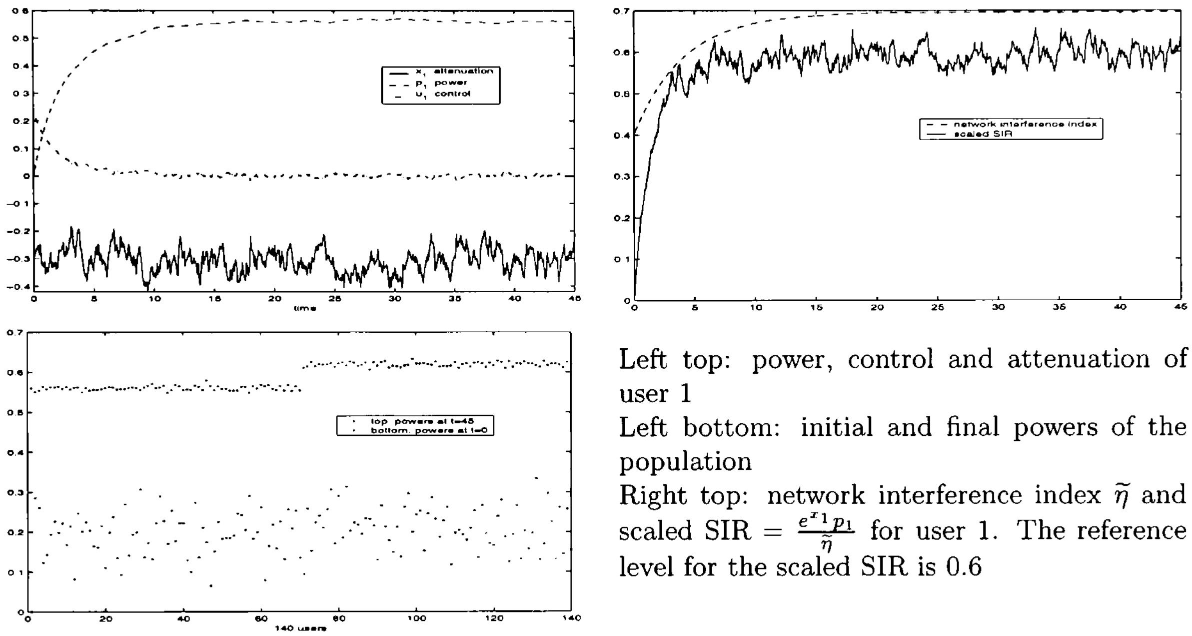
and λ_1, λ_2 are the two roots of

$$\lambda^2 - [\frac{n}{2}(\alpha+\beta) + \delta_1 + \delta_2]\lambda + [\frac{n}{2}(\alpha\delta_2 + \beta\delta_1) + \delta_1\delta_2] = 0,$$

from which we take

$$\lambda_1 = -0.6800, \quad \lambda_2 = -0.2705.$$

It is seen that all $\lambda_i, 1 \leq i \leq 140$, lie in the left half plane with a strictly positive stability margin. This fact reveals the stabilizing feature of the power adjustment scheme (5.78) in the small noise situation.



Left top: power, control and attenuation of user 1
Left bottom: initial and final powers of the population
Right top: network interference index $\widetilde{\eta}$ and scaled SIR = $\frac{e^{x_1}p_1}{\widetilde{\eta}}$ for user 1. The reference level for the scaled SIR is 0.6

FIGURE 5.6. The power adjustment of user 1 and the behaviour of the population with the single user based control law

It is shown by Left bottom in Figure 5.6 that at $t = 45$, the powers of the 140 users stay in small neighbourhoods of two different levels due to different long-term means of the attenuations for the two groups (i.e., Populations 1 and 2) of users.

5.5.2. Investigation of Population Behaviour in Large Systems. As indicated by the above numerical example, under a large population condition the network interference index exhibits a largely deterministic behavior in its evolution with respect to time (see Right top in Figure 5.6); this fact suggests the feasibility of modelling the system's evolution by a certain deterministic dynamics and associated initial conditions. This may potentially lead to completely decentralized control laws since each mobile uses only its own state and a deterministic process (subject to the aggregated influence of the individuals) to determine its control input.

Concerning power control, the important issues of large populations, decentralization, and the nature of the associated control laws will be addressed in Chapter 7. Also within the context of large population systems, we will investigate a class of large-scale linear models in Chapter 6 and develop the general methodology for analyzing decentralized control and for studying individual-mass behaviour.

Although in Chapter 6 the structure of the large-scale cost-coupled linear quadratic Gaussian (LQG) systems is inspired by that of the power control problem, it is of interest in its own right. In this linear context, we shall develop both the centralized optimal control, and more importantly, a decentralized game theoretic solution with an individual playing against the mass of other players. In turn, the solution for the large-scale linear systems will serve as a paradigm for approaching the large population power control problem in Chapter 7.

5.6. Adaptation with Unknown Parameters in Channel Dynamics

We rewrite the lognormal fading channel model of Section 2.4 as follows:

$$dx_i = -a_i(x_i + b_i)dt + \sigma_i dw_i, \qquad 1 \le i \le n. \tag{5.81}$$

In this model, the channel variation is characterized by the parameters $a_i > 0, b_i > 0, \sigma_i > 0$. For practical implementations, a_i, b_i, σ_i may not be known a priori, but x_i can be measured, for instance, with the aid of pilot signals [**18, 60**]. In CDMA systems, the power of users is updated with a period close to 1 millisecond (for instance, by 800Hz [**71**]) while the time scale of lognormal fading is much larger. Hence the channel may be regarded as varying in a very slow rate. In such a case one expects to have estimation of the channel state at high accuracy. In the following analysis, we shall assume perfect knowlege on the channel state x_i.

Consider an estimation algorithm for a_i and b_i via the measurement of x_i. For the i-th mobile, the parameters are estimated by the least squares algorithm where $\widehat{a}_i(t)$, $\widehat{b}_i(t)$ denote the estimate of a_i, b_i at $t \geq 0$, respectively. Define

$$\widehat{b}_i(t) = -\frac{1}{t}\int_0^t x_i(s)ds, \qquad t > 0, \tag{5.82}$$

$$dP_i = -P_i(x_i + \widehat{b}_i)(x_i + \widehat{b}_i)P_i dt, \qquad t \geq 0, \tag{5.83}$$

$$d\widehat{a}_i = -P_i(x_i + \widehat{b}_i)[dx_i + \widehat{a}_i(x_i + \widehat{b}_i)dt], \qquad t \geq 0, \tag{5.84}$$

where the initial conditions are given by $\widehat{b}_i(0)$, $P_i(0) > 0$, $\widehat{a}_i(0)$, respectively. The estimates are strongly consistent as stated by the following proposition.

Proposition 5.4. The estimates $\widehat{b}_i(t)$ and $\widehat{a}_i(t)$ converge to the true parameters with probability one as $t \to \infty$, i.e.,

$$\lim_{t\to\infty} \widehat{b}_i(t) = b_i, \qquad a.s. \tag{5.85}$$

$$\lim_{t\to\infty} \widehat{a}_i(t) = a_i, \qquad a.s. \tag{5.86}$$

with initial conditions $\widehat{b}_i(0)$, $\widehat{a}_i(0)$, $P_i(0) > 0$.

PROOF. Since $a_i > 0, \sigma_i > 0$, it follows that x_i is an ergodic diffusion process satisfying

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t x_i(s)ds = \lim_{t\to\infty} Ex_i(t) = -b_i \qquad a.s.$$

and (5.85) follows. Write $\widetilde{a}_i = \widehat{a}_i - a_i$, $\widetilde{b}_i = \widehat{b}_i - b_i$ and $y_i = x_i + \widehat{b}_i$. It is easy to verify that

$$\begin{aligned} d\widetilde{a}_i &= -P_i(x_i + \widehat{b}_i)[dx_i + \widehat{a}_i(x_i + \widehat{b}_i)dt] \\ &= -P_i\widetilde{a}_i y_i^2 dt - a_i\widetilde{b}_i P_i y_i dt - \sigma_i P_i y_i dw_i, \end{aligned} \tag{5.87}$$

$$dP_i^{-1} = y_i^2 dt. \tag{5.88}$$

By (5.88) it follows that

$$
\begin{aligned}
&|\frac{1}{t}P_i^{-1}(t) - \frac{1}{t}\int_0^t (x_i + b_i)^2 ds| \\
=&|\frac{1}{t}P_i^{-1}(0) + \frac{1}{t}\int_0^t (b_i - \widehat{b}_i)^2 ds - \frac{2}{t}\int_0^t (x_i + b_i)(b_i - \widehat{b}_i)dt| \\
\leq&\frac{1}{t}P_i^{-1}(0) + \frac{1}{t}\int_0^t \widetilde{b}_i^2 ds + 2(\frac{1}{t}\int_0^t (x_i + b_i)^2 ds)^{\frac{1}{2}}(\frac{1}{t}\int_0^t \widetilde{b}_i^2 ds)^{\frac{1}{2}}.
\end{aligned} \tag{5.89}
$$

Since $\lim_{t\to\infty}\frac{1}{t}\int_0^t (x_i + b_i)^2 ds = \lim_{t\to\infty} E[x_i(t) - Ex_i(t)]^2$ a.s. by ergodicity of x_i, and $\lim_{t\to\infty}\frac{1}{t}\int_0^t \widetilde{b}_i^2 ds = 0$ a.s., it follows that

$$
\lim_{t\to\infty}\frac{1}{t}P_i^{-1}(t) = \lim_{t\to\infty} E[x_i(t) - Ex_i(t)]^2 > 0, \qquad a.s. \tag{5.90}
$$

By (5.87) and (5.88) we obtain

$$
dP_i^{-1}\widetilde{a}_i^2 = -\widetilde{a}_i^2 y_i^2 dt - 2a_i\widetilde{b}_i\widetilde{a}_i y_i dt + \sigma_i^2 P_i y_i^2 dt - 2\sigma_i\widetilde{a}_i y_i dw_i. \tag{5.91}
$$

Applying the technique in [**15**], from (5.91) we get

$$
\begin{aligned}
&P_i^{-1}(t)\widetilde{a}_i^2(t) - P_i^{-1}(0)\widetilde{a}_i^2(0) \\
=& -\int_0^t \widetilde{a}_i^2 y_i^2 ds + \int_0^t \sigma_i^2 P_i y_i^2 dt - 2\int_0^t a_i\widetilde{b}_i\widetilde{a}_i y_i dt - 2\int_0^t \sigma_i\widetilde{a}_i y_i dw_i \\
=& -\int_0^t \widetilde{a}_i^2 y_i^2 ds + O(\log P_i^{-1}(t)) + O\left([\int_0^t \widetilde{b}_i^2 ds]^{\frac{1}{2}} \cdot [\int_0^t \widetilde{a}_i^2 y_i^2 ds]^{\frac{1}{2}}\right) \\
&+ O\left((\int_0^t \widetilde{a}_i^2 y_i^2 ds)^{\frac{1}{2}+\varepsilon}\right), \qquad a.s.
\end{aligned} \tag{5.92}
$$

where $0 < \varepsilon < \frac{1}{2}$. From (5.92) it follows that

$$
\int_0^t \widetilde{a}_i^2 y_i^2 ds \leq O(\log P_i^{-1}(t)) + O\left([\int_0^t \widetilde{b}_i^2 ds]^{\frac{1}{2}} \cdot [\int_0^t \widetilde{a}_i^2 y_i^2 ds]^{\frac{1}{2}}\right) + O\left((\int_0^t \widetilde{a}_i^2 y_i^2 ds)^{\frac{1}{2}+\varepsilon}\right),
$$

which yields

$$
\int_0^t \widetilde{a}_i^2 y_i^2 ds = O(\log t) + O(\int_0^t \widetilde{b}_i^2 ds), \qquad a.s.
$$

Since $\widetilde{b}_i(t) \to 0$, a.s., as $t \to \infty$, it follows that

$$\int_0^t \widetilde{a}_i^2 y_i^2 ds = o(t) \qquad a.s. \tag{5.93}$$

By (5.90), (5.92) and (5.93), we get $\lim_{t\to\infty} \widetilde{a}_i(t) = 0$, a.s., and (5.86) follows. □

The estimation of σ^2 is more complicated than that of a_i and b_i. In the following we employ a discrete time prediction error term to construct the empirical variance. We first take a sampling step $h > 0$ to discretize (5.81) to write

$$x_i[(k+1)h] + b_i = e^{-a_i h}[x_i(kh) + b_i] + \sigma_i \int_{kh}^{(k+1)h} e^{-a_i[(k+1)h-s]} dw_i(s), \quad k \geq 0. \tag{5.94}$$

Setting $A_i = e^{-a_i h}$ and $\nu_i(kh) = \int_{kh}^{(k+1)h} e^{-a_i[(k+1)h-s]} dw_i(s)$, (5.94) can be written in the form

$$x_i[(k+1)h] + b_i = A_i[x_i(kh) + b_i] + \sigma_i \nu_i(kh).$$

It is easy to verify that $\mathrm{Var}(\nu_i(kh)) = \frac{1-e^{-2a_i h}}{2a_i} \triangleq \Sigma_{\nu_i}$. Denote $\widehat{A}_i(kh) = e^{-\widehat{a}_i(kh)h}$, $\widehat{\Sigma}_{\nu_i}(kh) = \frac{1-e^{-2\widehat{a}_i(kh)h}}{2\widehat{a}_i(kh)}$ and

$$\widehat{\sigma}_i^2(nh) = \frac{1}{n\widehat{\Sigma}_{\nu_i}(kn)} \sum_{k=0}^{n-1} \left(x_i[(k+1)h] + \widehat{b}_i(kh) - \widehat{A}_i(kh)[x_i(kh) + \widehat{b}_i(kh)] \right)^2. \tag{5.95}$$

It is straightforward to show that (5.95) can be written in a recursive form. We have the following proposition:

Proposition 5.5. For $\widehat{\sigma}_i^2(nh), n \geq 1$, defined by (5.95), we have

$$\lim_{n\to\infty} \widehat{\sigma}_i^2(nh) = \sigma_i^2, \qquad a.s. \tag{5.96}$$

where $\sigma_i^2 > 0$ is determined by (5.81).

PROOF. For notational brevity, in the following proof we write $x_i(kh)$, A_i, b_i, $\widehat{A}_i(kh)$, $\widehat{b}_i(kh)$, $\nu_i(kh)$ as $x_i(k)$, A, b, $\widehat{A}(k)$, $\widehat{b}(k)$, $\nu(k)$, respectively. Setting $\widetilde{A}(k) =$

$\widehat{A}_i(kh) - A_i$ and $\widetilde{b}(k) = \widehat{b}_i(kh) - b_i$, we have

$$\sum_{k=0}^{n-1}\Big(x_i[(k+1)h] + \widehat{b}_i(kh) - \widehat{A}_i(kh)[x_i(kh) + \widehat{b}_i(kh)]\Big)^2$$

$$= \sum_{k=0}^{n-1}[\widetilde{A}(k)x_i(k)]^2 + \sum_{k=0}^{n-1}[Ab - \widehat{A}(k)\widehat{b}(k) - \widehat{b}(k)]^2 + \sum_{k=0}^{n-1}\sigma_i^2\nu^2(k)$$

$$+ 2\sum_{k=0}^{n-1}[\widetilde{A}(k)x_i(k)][\widehat{A}(k)\widehat{b}(k) + \widetilde{b}(k) - Ab] + 2\sum_{k=0}^{n-1}[-\widetilde{A}(k)x_i(k)][\sigma_i\nu(k)]$$

$$+ 2\sum_{k=0}^{n-1}[Ab - \widehat{A}(k)\widehat{b}(k) - \widetilde{b}(k)][\sigma_i\nu(k)] \triangleq S_1 + S_2 + S_3 + S_{12} + S_{13} + S_{23}. \quad (5.97)$$

Since $\widehat{A}(k) \to A$, $\widehat{b}(k) \to b$ a.s., as $k \to \infty$, and $\sum_{k=0}^{n-1} x_i^2(k) = O(n)$ a.s., it follows that

$$|S_1| + |S_2| + |S_{12}| = o(n), \qquad a.s. \quad (5.98)$$

On the other hand, by the estimates in [**54**], we have

$$S_{13} = O(S_1^{\frac{1}{2}+\varepsilon}), \quad S_{23} = O(S_2^{\frac{1}{2}+\varepsilon}), \qquad a.s.$$

for any $0 < \varepsilon < \frac{1}{2}$, and therefore,

$$|S_{13}| + |S_{23}| = o(n), \qquad a.s. \quad (5.99)$$

By (5.97)-(5.99), it follows that

$$\lim_{n\to\infty} \widehat{\sigma}_i^2(nh) = \lim_{n\to\infty} \frac{1}{n\widehat{\Sigma}_{\nu_i}(kn)} \sum_{k=0}^{n-1} \sigma_i^2\nu^2(k) = \sigma_i^2, \quad a.s. \quad (5.100)$$

which completes the proof. □

It is also possible to apply certain well known discrete time parameter estimation algorithms (see, e.g., [**14**]) to (5.94) for estimating a_i and b_i.

The estimation algorithm given above will potentially remove certain obstacles in applying the stochastic control framework developed in this Chapter as well as in

Chapters 2-3 to power adjustment with unknown channel parameters. Notably, using the above algorithm one can obtain an adaptive version of the suboptimal control law of Section 5.5 in a straightforward way, and the adaptive control law is derived by solving (5.64)-(5.66) and (5.72)-(5.74) when the parameters a_i, b_i, σ_i are replaced by their on-line estimates.

CHAPTER 6

LQG Optimization for Large-Scale Cost-Coupled Systems

6.1. Introduction

In this Chapter, we investigate optimization of large-scale linear quadratic Gaussian (LQG) control systems. We intend to develop implementationally simple control scheme for such large-scale systems. At the present stage, the system dynamics under consideration are not in their most general form; instead, the dynamics proposed are those for describing the dynamical behaviour of many agents (also to be equivalently called individuals or players) evolving in a similar manner. Roughly speaking, these agents have independent dynamics when state regulation is not included. The optimization of such systems is based on quadratic costs including two cases: the global cost and the individual cost.

To facilitate our exposition, the global cost based optimization shall be termed as the (*centralize*) *optimal control* and the individual cost based optimization shall be called the *dynamic LQG game* (*solution*), or simply *LQG game* (*solution*).

For the LQG system we first examine the optimal control problem and analyze the resulting algebraic Riccati equation as well as the feedback control.

Subsequently, we turn to the LQG game; in this part we analyze the resulting ε-Nash equilibrium properties for the control law. In this case, each agent is weakly coupled with the other agents in the sense that it is only connected to the other agents through its cost function. We view this to be the characteristic property of a class of problems we call *cost coupled* (*distributed*) control problems. The connection to economic problems is immediately evident.

Due to the particular structure of the individual cost, the mass formed by all other agents imposes its impact on a given agent as a nearly deterministic quantity; for any known mass influence, any given individual will seek its best way to respond to the mass so that its cost is minimized. In a practical situation, the mass influence cannot be assumed known a priori; this, however, does not present any difficulty for applying the individual-mass interplay methodology. In this noncooperative game setup, by assuming that all agents are at the same level of rationality and adopting basically the same mode of reasoning for any presumed known mass behaviour, one can find the natural response of all individuals, which in turn determine a corresponding mass behaviour. Thus a meaningful solution to the underlying problem is to find a fixed point for this procedure, i.e., find a mass behaviour so that the optimal response in a certain sense will exactly generate the aforementioned mass behaviour. We note that this LQG control problem is closely related to the stochastic power control problem. The framework presented in this part is particularly suitable for optimization of large-scale systems where individuals in the system seek to optimize for their own return and where, moreover, it is more difficult to achieve global optimality through close coordination between all agents. The general methodology of noncooperative games provides a feasible methodology for building simple optimization rules which under appropriate conditions can lead to stable population behaviour.

At the end of this Chapter we give a general analysis comparing the centralized cost based optimal control with the individual cost based (decentralized) control.

6.2. Dynamically Independent and Cost-Coupled Systems

Suppose in a linear stochastic system, the state evolution of each of the n individuals or agents is described by

$$dz_i = (az_i + bu_i)dt + \sigma dw_i, \qquad 1 \leq i \leq n, \quad t \geq 0, \tag{6.1}$$

where $\{w_i, 1 \leq i \leq n\}$ denotes n independent standard scalar Wiener processes. The initial state $z_i(0)$ are mutually independent and are also independent of $\{w_i, 1 \leq i \leq n\}$. In addition, $E|z_i(0)|^2 < \infty$ and $b \neq 0$. For a given individual in the system, apart from the control input, its state is not subject to direct influence from the other individuals. In the following we will investigate the behaviour of the agents when they interact with each other through coupling costs. Thus we term this class of models as dynamically independent and cost-coupled systems. Concerning the costs, we will consider two scenarios, i.e., a global cost function and individual costs.

In *the first scenario*, i.e., the centralized optimal control problem, the n individuals interact with each other through a *global cost function*

$$\begin{aligned} J = J(u_1, v_1; u_2, v_2; \cdots; u_n, v_n) &= \sum_{i=1}^{n} J_i(u_i, v_i) \\ &\triangleq E \int_0^\infty \sum_{i=1}^{n} e^{-\rho t}[(z_i - v_i)^2 + ru_i^2]dt; \end{aligned} \tag{6.2}$$

and in particular we set in the cost-coupled case $v_i = \gamma(\frac{1}{n}\sum_{k\neq i}^{n} z_k + \eta)$. Here we assume $\rho, r, \gamma, \eta > 0$.

In *the second scenario*, i.e., the dynamic LQG game problem, each agent is assigned a cost function J_i defined as above, and we study the large system behaviour in the dynamic noncooperative game framework.

In the rest of this Section we give a production planning example for illustration of the cost J_i in (6.2), and we will derive a link term v_i^0 which is different from but closely resembles $v_i = \gamma(\frac{1}{n}\sum_{k\neq i}^{n} z_k + \eta)$. Most of the analysis in this Chapter is related to the case with the link term $v_i = \gamma(\frac{1}{n}\sum_{k\neq i}^{n} z_k + \eta)$; and we reserve the phrase "cost-coupled case" only for this v_i. The methodology of this Chapter can be applied to the production planning example without difficulty.

We stress that *throughout* this Chapter z_i is described by the dynamics (6.1).

6.2.1. A Production Planning Example.

The example below is motivated by the work of Basar and Ho [**9**] where a quadratic nonzero-sum game was considered for a static duopoly model. In their work it was assumed that the price of the commodity decreases linearly as the overall production level of the two firms increases. We will study here a dynamic model consisting of many players.

Consider n firms F_i, $1 \leq i \leq n$, supplying the same product to the market. First, let x_i be the production level of firm F_i and suppose x_i is subject to adjustment by the following model:

$$dx_i = u_i dt + \sigma dw_i, \qquad t \geq 0, \tag{6.3}$$

which is a special form of (6.1). Here u_i denotes the action of increasing or decreasing the production level x_i, and σdw_i denotes uncertainty in the change of x_i.

Second, by generalizing the affine linear price model of [**9**] to the case of many players, we assume the price of the product is given by

$$p = \overline{\eta} - \overline{\gamma}(\frac{1}{n}\sum_{i=1}^{n} x_i), \tag{6.4}$$

where $\overline{\eta}, \overline{\gamma} > 0$. In (6.4) the overall production level $\sum_{i=1}^{n} x_i$ is scaled by the factor $\frac{1}{n}$ instead of a straight summation. A justification for doing so is that we are modelling an expanding market in which an increasing number of firms are allowed to compete together to serve an increasing number of consumers in large geographical areas. So $\frac{1}{n}\sum_{i=1}^{n} x_i$ is a reasonable choice to measure the average production level in an expanding market. Following [**9**], $\overline{\gamma}$ may be intepreted as the "slope of the demand curve".

Third, we assume that the firm F_i adjusts its production level x_i by referring to the current price of the product. When the price goes up, if the firm increases x_i too rapidly, it may creat excessive inventory load. In the case of a price drop, if the firm overly decreases x_i, it may face high risk of supply shortage in the near future. Hence it is critical to have an appropriate planning for the production level in order to be in a better position with respect to profit-making. In the following we take a mild adjustment rate for the production level by seeking

$$x_i \approx \beta p = \beta[\overline{\eta} - \overline{\gamma}(\frac{1}{n}\sum_{i=1}^{n} x_i)], \tag{6.5}$$

where $\beta > 0$ is a constant. Based on (6.5) we write a penalty term

$$\{x_i - \beta[\overline{\eta} - \overline{\gamma}(\frac{1}{n}\sum_{i=1}^{n} x_i)]\}^2 \triangleq (x_i - v_i^0)^2. \tag{6.6}$$

On the other hand, in the adjustment of x_i the control u_i corresponds to actions of shutting down or restarting production lines, or even the construction of new ones. Each of these actions will incurr certain costs to the firm; for simplicty we denote the instantaneous cost of the adjustment by ru_i^2, where $r > 0$. We now write the infinite horizon discounted cost for the firm F_i as follows:

$$J_i^x(u_i, v_i^0) = E\int_0^\infty e^{-\rho t}[(x_i - v_i^0)^2 + ru_i^2]dt, \tag{6.7}$$

where $\rho > 0$ and we use the supscript in J_i^x to indicate that the associated dynamics is (6.3). Here obviously $v_1^0 = \cdots = v_n^0$. Notice that $v_i^0 = \beta[\overline{\eta} - \overline{\gamma}(\frac{1}{n}\sum_{i=1}^{n} x_i)]$ and $v_i = \gamma(\frac{1}{n}\sum_{k\neq i}^{n} z_k + \eta)$ share the common feature of taking an average over a mass.

In this production planning example, each firm has its individual dynamics and all the firms interact with each other through the market and their cost function.

6.3. The Global Cost Based Optimal Control

Assuming complete information for the system state, we determine the optimal control law minimizing the cost function (6.2) with $v_i = \gamma(\frac{1}{n}\sum_{k\neq i}^{n} z_k + \eta)$, subject to the dynamics (6.1). Writing the optimal cost v associated with (6.2) in the form $v(z) = z^T P z + 2s^T z + s_0$, and invoking the standard results of LQG control (see, e.g., [**52, 10, 11**]), we have

$$2(a - \frac{\rho}{2})P - \frac{b^2}{r}P^2 + Q = 0, \tag{6.8}$$

$$\rho s = as - \frac{b^2}{r}Ps + (\frac{n-1}{n}\gamma^2\eta - \gamma\eta)[1, \cdots, 1]^T,$$

$$\rho s_0 = -\frac{b^2}{r}s^T s + \sigma^2 \mathrm{tr} P + n\gamma^2\eta^2. \tag{6.9}$$

where $Q, P \in \mathrm{R}^{n\times n}$. Q is seen to have the form below

$$Q = \begin{pmatrix} \alpha & \beta & \cdots & \beta \\ \beta & \alpha & \cdots & \beta \\ \vdots & & \ddots & \vdots \\ \beta & \beta & \cdots & \alpha \end{pmatrix}, \tag{6.10}$$

where $\alpha = 1 + \frac{(n-1)\gamma^2}{n^2}$, $\beta = -\frac{2\gamma}{n} + \frac{(n-2)\gamma^2}{n^2}$, which results in P taking the form

$$P = \begin{pmatrix} p & q & \cdots & q \\ q & p & \cdots & q \\ \vdots & & \ddots & \vdots \\ q & q & \cdots & p \end{pmatrix}. \tag{6.11}$$

The eigenvalues of Q are give by $\lambda_1 = \alpha + (n-1)\beta = [1 - \frac{(n-1)\gamma}{n}]^2$, $\lambda_2 = \lambda_3 = \cdots = \lambda_n = \alpha - \beta = (1 + \frac{\gamma}{n})^2$. In the following we consider two cases.

Case 1: $\gamma > 0$ is chosen such that $\lambda_1 > 0$. Then clearly $Q > 0$ (i.e., strictly positive definite) and the pair $[(a-\frac{\rho}{2})I_n, Q^{\frac{1}{2}}]$ is observable, so that the Riccati equation (6.8) has a unique solution $P > 0$.

Case 2: $\gamma > 0$ is chosen such that $\lambda_1 = 0$. In this case the solution P to (6.8) is only positive semidefinite. Indeed, using an orthogonal transform T such that

$T^\tau QT = \text{Diag}(\lambda_i) \triangleq \Lambda_Q$, from (6.8) we obtain

$$2(a - \frac{\rho}{2})T^\tau PT - \frac{b^2}{r}(T^\tau PT)^2 + \Lambda_Q = 0,$$

Obviously, to the above equation there exists a unique positive semidefinite solution $T^\tau PT$ of rank $n-1$. So that there exists a unique positve semidefinite solution P of rank $n-1$ to the equation (6.8).

Substituting P into (6.8) and denoting $\overline{a} = a - \frac{\rho}{2}$, $\overline{b} = \frac{b}{\sqrt{r}}$, we get the following equations

$$2\overline{a}p - \overline{b}^2[p^2 + (n-1)q^2] + \alpha = 0, \tag{6.12}$$

$$2\overline{a}q - \overline{b}^2[2pq + (n-2)q^2] + \beta = 0, \tag{6.13}$$

which further give

$$\overline{b}^2(p-q)^2 - 2\overline{a}(p-q) - (\alpha - \beta) = 0. \tag{6.14}$$

Under the positive semidefinite condition of P (including both case 1 and 2), solving (6.12)-(6.14) yields

$$p = \frac{\overline{a} + \sqrt{\overline{a}^2 + \overline{a}^2(\alpha - \beta)}}{\overline{a}^2} + \frac{\sqrt{\overline{a}^2 + \overline{b}^2(\alpha-\beta) + n\overline{b}^2\beta} - \sqrt{\overline{a}^2 + \overline{b}^2(\alpha-\beta)}}{n\overline{b}^2}, \tag{6.15}$$

$$q = \frac{\sqrt{\overline{a}^2 + \overline{b}^2(\alpha-\beta) + n\overline{b}^2\beta} - \sqrt{\overline{a}^2 + \overline{b}^2(\alpha-\beta)}}{n\overline{b}^2}. \tag{6.16}$$

Sumarizing the above analysis we state the following proposition:

Proposition 6.1. There exists a unique solution $P \geq 0$ to (6.8), where Q is defined by (6.10), and P is given by (6.11), (6.15), (6.16). If $\lambda_1 = [1 - \frac{(n-1)\gamma}{n}]^2 > 0$, then $P > 0$; if $\lambda_1 = 0$, $P \geq 0$ is of rank $n-1$. □

6.3.1. The Optimal Control and its Asymptotic Properties: the Individual and the Collective. By well known results in LQG optimal control, the feedback control for the i-th player is determined as

$$u_i = -\frac{b}{r}(Pz)_i - \frac{b}{r}s_i = -\frac{b}{r}pz_i - \frac{b}{r}q\sum_{k\neq i} z_k - \frac{b}{r}s_i. \tag{6.17}$$

For large n, we have the expressions

$$p(n) = \frac{\bar{a} + \sqrt{\bar{a}^2 + \bar{b}^2}}{\bar{b}^2} + o(1), \quad q(n) = \frac{\sqrt{\bar{a}^2 + (1-\gamma)^2\bar{b}^2} - \sqrt{\bar{a}^2 + \bar{b}^2}}{n\bar{b}^2} + o(\frac{1}{n}),$$

$$s_i(n) = \frac{\gamma\eta(\gamma - 1)}{\frac{\rho}{2} + \sqrt{\bar{a}^2 + (1-\gamma)^2\bar{b}^2}} + o(1), \quad i = 1, 2, \cdots, n,$$

where we use $p(n)$, $q(n)$, and $s_i(n)$ to indicate their explicit dependence on n.

The main feature of the control u_i is characterized by the first two terms in the far right side of (6.17). The first term $-\frac{b}{r}pz_i$ contains information on the state of the i-th individual itself. Noticing the asymptotic property of q, in the second term $-\frac{b}{r}q\sum_{k\neq i} z_k$ an averaging takes place. This averaged quantity measures the collective effect of all the other agents.

Assuming the initial state $z|_{t=0}$ is always 0, then the limit average cost incurred by each agent is given as

$$\lim_{n\to\infty}\frac{v}{n} = \lim_{n\to\infty}\frac{s_0(n)}{n} = \frac{\gamma^2\eta^2}{\rho}\left[1 - \frac{\bar{b}^2(\gamma-1)^2}{(\frac{\rho}{2} + \sqrt{\bar{a}^2 + (1-\gamma)^2\bar{b}^2})^2}\right] + \frac{\sigma^2(\bar{a} + \sqrt{\bar{a}^2 + \bar{b}^2})}{\rho\bar{b}^2}. \tag{6.18}$$

6.4. The Linear Tracking Problem

In this Chapter, one of our main interests is in the large population behaviour when each agent pursues its optimization strategy based on its individual cost. With the above particular set of individual costs associated with the population, the essence of an individual's participation in the play is to adjust its state process so that it can follow the average effect of the mass in a certain sense. We seek to develop simplified control strategies; in doing so, a major step in our analysis will be to construct a certain deterministic approximation of the average effect which a given player receives from the mass.

To begin with, for large n, assume $z^*_{-i} \triangleq \gamma(\frac{1}{n}\sum_{k\neq i}^n z_k + \eta)$ in Section 6.2 is approximated by a *deterministic* continuous function $z^*(t)$. Now assuming z^* is known, we construct the *individual cost* for the i-th player as follows:

$$J_i(u_i, z^*) = E\int_0^\infty e^{-\rho t}\{[z_i - z^*(t)]^2 + ru_i^2\}ds, \tag{6.19}$$

which can be regarded as the i-th component of the global cost (6.2) with $v_i(t) \triangleq z^*(t)$. We note that for large n, it is reasonable to use a single $z^*(t)$ to approximate all $z^*_{-i}, 1 \le i \le n$. By use of such a deterministic function we have a natural splitting of the centralized cost such that the i-th player's cost is isolated from direct interaction with the other players.

Before we develop the approximation technique for individual cost based optimization problems, we first examine a general tracking problem where we consider bounded z^*. For the tracking problem itself, the boundedness condition on z^* can be relaxed, but we shall not do so since the results obtained here is sufficient for the individual mass behaviour analysis in the following Sections of this Chapter.

In the following for two functions $\varphi_1(t), \varphi_2(t) > 0$, both defined on $[0,\infty)$, $\varphi_1(t) = o(\varphi_2(t))$ means $\lim_{t\to\infty}\frac{\varphi_1(t)}{\varphi_2(t)} = 0$, and $\varphi_1(t) = O(1)$ means $\sup_{t\ge 0}|\varphi(t)| < \infty$.

Let Π be the positive solution to the algebraic Riccati equation

$$\rho\Pi = 2a\Pi - \frac{b^2}{r}\Pi^2 + 1. \tag{6.20}$$

It is well known that $a - \frac{b^2\Pi}{r} - \frac{\rho}{2} < 0$, or equivalently, $-a + \frac{b^2\Pi}{r} + \frac{\rho}{2} > 0$. Denote

$$\beta_1 = -a + \frac{b^2}{r}\Pi, \quad \beta_2 = -a + \frac{b^2}{r}\Pi + \rho. \tag{6.21}$$

It is obvious that $\beta_2 > 0$ since $-a + \frac{b^2\Pi}{r} + \frac{\rho}{2} > 0$.

We have the following proposition.

Proposition 6.2. Suppose i) the initial condition $z_i|_{t=0}$ has finite second order moment, $z^* \in C_b[0,\infty) \triangleq \{x \in C[0,\infty),\ \sup_{t\in[0,\infty)} |x(t)| < \infty\}$; ii) $\Pi > 0$ is the solution to the Riccati equation (6.20) and $\beta_1 = -a + \frac{b^2}{r}\Pi > 0$; and iii) $s \in C_b[0,\infty)$ is determined by the following differential equation

$$\rho s = \frac{ds}{dt} + as - \frac{b^2}{r}\Pi s - z^*, \tag{6.22}$$

Then $\widehat{u}_i = -\frac{b}{r}(\Pi z_i + s)$ is an optimal control minimizing $J_i(u_i, z^*)$, for all u_i adapted to $\sigma(w_i(r), r \leq t)$. □

Before proving the proposition, we first have a brief discussion about the assumptions introduced above. For minimization of J_i, the admissible control set can be taken as $\mathcal{U}_i \triangleq \{u_i | u_i$ adapted to $\sigma(w_i(r), r \leq t)$, and $\int_0^\infty e^{-\rho t}(z_i^2 + u_i^2)dt < \infty\}$, where the process z_i is subject to the control u_i. In fact, $\mathcal{U}_i$ is nonempty due to controllability of (6.1); Condition i) ensures that J_i has a finite minimum with respect to u_i adapted to $\sigma(w_i(s), s \leq t)$, and this minimum is attained in $\mathcal{U}_i$. Condition ii) is needed in a technical step of the proof to establish an auxiliary equality; it means that the resulting closed-loop system has a stable pole. Condition ii) will also be used later for asymptotic analysis in the large population game context. In Condition iii) instead of an initial condition $s|_{t=0}$, only a boundedness condition for s is specified. It turns out this boundedness condition can uniquely determine s on $[0,\infty)$. This point will be illustrated after the proof of the proposition.

Proof of Proposition 6.2. The proof will be done following an algebraic approach as in **[10]**, but in the current infinite horizon case, we need to estimate the growth rate of the stochastic processes involved. First we define the auxiliary process y with initial condition $y_0 = z_i|_{t=0}$ as follows:

$$dy = \{ay + b[-\frac{b}{r}(\Pi y + s)]\}dt + dw_i. \tag{6.23}$$

For any $u_i \in \mathcal{U}_i$, the resulting state evolution of z_i is described by

$$dz_i = (az_i + bu_i)dt + \sigma dw_i.$$

Denote $u_i = -\frac{b}{r}(\Pi z_i + s) + \widetilde{u}$ and $z_i - y = \widetilde{z}$. Then it is obvious

$$d\widetilde{z} = (a - \frac{b^2}{r}\Pi)\widetilde{z}dt + b\widetilde{u}dt, \quad \widetilde{z}|_{t=0} = 0. \tag{6.24}$$

Since $E\int_0^\infty e^{-\rho t}[z_i^2 + u_i^2]dt < \infty$, it follows that

$$E\int_0^\infty e^{-\rho t}\widetilde{u}^2 dt < \infty. \tag{6.25}$$

Now the cost $J_i(u_i, z^*)$ can be written as

$$\begin{aligned} J_i(u_i, z^*) =& E\int_0^\infty e^{-\rho t}[(y - z^*)^2 + \frac{b^2}{r}(\Pi y + s)^2]dt \quad (\triangleq I_1) \\ &+ E\int_0^\infty e^{-\rho t}[\widetilde{z}^2 + r(\widetilde{u} - \frac{b}{r}\Pi\widetilde{z})^2]dt \quad (\triangleq I_2) \\ &+ 2E\int_0^\infty e^{-\rho t}[\widetilde{z}(y - z^*) - b(\Pi y + s)(\widetilde{u} - \frac{b}{r}\Pi\widetilde{z})]dt \quad (\triangleq 2I_3). \end{aligned} \tag{6.26}$$

For $T > 0$, using Ito's formula and taking expectation we get

$$\begin{aligned} Ee^{-\rho T}\widetilde{z}(\Pi y + s)(T) = E\int_0^T & e^{-\rho t}\{-\rho\widetilde{z}(\Pi y + s) + [(a - \frac{b^2}{r}\Pi)\widetilde{z} + b\widetilde{u}](\Pi y + s) \\ &+ \widetilde{z}[\Pi(a - \frac{b^2}{r}\Pi)y - \frac{b^2}{r}\Pi s + (\rho - a + \frac{b^2}{r}\Pi)s + z^*]\}dt \end{aligned} \tag{6.27}$$

$$\triangleq E\int_0^T e^{-\rho t}K(\widetilde{z}, \widetilde{u}, y, s, z^*)dt. \tag{6.28}$$

It can be verified that $Ey^2 = O(1)$ by Condition ii); and moreover, since $\beta_1 > 0$ it follows that

$$E\widetilde{z}_t^2 \le \int_0^t e^{-2\beta_1(t-\tau)}b^2 E\widetilde{u}_\tau^2 d\tau \le \int_0^t b^2 E\widetilde{u}_\tau^2 d\tau = \int_0^t b^2 e^{\rho\tau}e^{-\rho\tau}E\widetilde{u}_\tau^2 d\tau = o(e^{\rho t}),$$

where we get the estimate $o(e^{\rho t})$ by the fact (6.25), and therefore $Ee^{-\rho T}\widetilde{z}(\Pi y + s)(T) \to 0$, as $T \to \infty$. Taking limit on both sides of (6.27) gives

$$E\int_0^\infty e^{-\rho t}K(\widetilde{z}, \widetilde{u}, y, s, z^*)dt = 0. \tag{6.29}$$

Denoting the integrand in I_3 of (6.26) as $I(\tilde{z}, \tilde{u}, y, s, z^*)$, then it is straightforward to verify that $I(\tilde{z}, \tilde{u}, y, s, z^*) + K(\tilde{z}, \tilde{u}, y, s, z^*)$ is identical to zero, and consequently,

$$J_i(u_i, z^*) = I_1 + I_2. \tag{6.30}$$

Taking into account the initial condition and dynamics of $\tilde{z}$, it follows that $u_i = -\frac{b}{r}(\Pi z_i + s)$ minimizes J_i and the proof is complete. □

The optimal cost for the deterministic tracking problem is gived as follows.

Proposition 6.3. Assume Assumptions i)-iii) in Proposition 6.2 hold and $q \in C_b[0, \infty)$ is a solution to the equation

$$\rho q = \frac{dq}{dt} - \frac{b^2}{r}s^2 + (z^*)^2 + \sigma^2 \Pi, \tag{6.31}$$

Then the cost for the optimal control $\hat{u}_i = -\frac{b}{r}(\Pi z_i + s)$ is given by $J_i(\hat{u}_i, z^*) = \Pi E z_i^2(0) + 2s(0)Ez_i(0) + q(0)$, where $z_i(0)$ is the initial state.

PROOF. First, we write the closed-loop system for the control $\hat{u}_i$ as

$$dz_i = [(a - \frac{b^2}{r}\Pi)z_i - \frac{b^2}{r}s^2]dt + \sigma dw_i. \tag{6.32}$$

For any $T > 0$, by (6.32) and Ito's formula it follows that

$$\begin{aligned} &E\int_0^T d[e^{-\rho t}(\Pi z_i^2 + 2sz_i + q)] \\ =&E\int_0^T (-\rho)e^{-\rho t}(\Pi z_i^2 + 2sz_i + q)dt + E\int_o^T e^{-\rho t}\Phi dt \end{aligned} \tag{6.33}$$

where

$$\begin{aligned} \Phi \triangleq & 2\Pi z_i[(a - \frac{b^2}{r}\Pi)z_i - \frac{b^2}{r}s^2] + \Pi\sigma^2 + 2z_i(\rho s - as + \frac{b^2}{r}\Pi s + z^*) \\ &+ 2s[(a - \frac{b^2}{r}\Pi)z_i - \frac{b^2}{r}s^2] + [\rho q + \frac{b^2}{r}s^2 - (z^*)^2 - \sigma^2\Pi]. \end{aligned}$$

Then it can be verified that

$$-\rho(\Pi z_i^2 + 2sz_i + q) + \Phi = -(z_i - z^*)^2 - \frac{b^2}{r}(\Pi z_i + s)^2 \tag{6.34}$$

By (6.33) and (6.34) it follows that

$$J(\widehat{u}_i, z^*) = E\int_0^\infty [(z_i - z^*)^2 + \frac{b^2}{r}(\Pi z_i + s)^2]dt$$

$$= - E\lim_{T\to\infty}\int_0^T d[e^{-\rho t}(\Pi z_i^2 + 2sz_i + q)] = \Pi E z_i^2(0) + 2s(0)Ez_i(0) + q(0) \qquad (6.35)$$

since $\lim_{T\to\infty} Ee^{-\rho T}(\Pi z_i^2 + 2sz_i + q)(T) = 0$ by the growth condition of Ez_i^2, s and q. This completes the proof. □

Now we show how s can be uniquely determined subject to the boundedness condition specified in Proposition 6.2, i.e., $z^*(t) = C_b[0,\infty)$, $s(t) = C_b[0,\infty)$. With an initial condition s_0 and recalling (6.22), s can be expressed as

$$s(t) = s_0 e^{\beta_2 t} + e^{\beta_2 t}\int_0^t e^{-\beta_2 \tau} z^*(\tau)d\tau. \qquad (6.36)$$

Since $\beta_2 > 0$, the integral $\int_0^\infty e^{-\beta_2\tau} z^*(\tau)d\tau$ exists and is finite. We take initial condition $s_0 = -\int_0^\infty e^{-\beta_2\tau} z^*(\tau)d\tau$, so that

$$s(t) = -e^{\beta_2 t}\int_t^\infty e^{-\beta_2\tau} z^*(\tau)d\tau \in C_b[0,\infty),$$

and any other initial condition will yield a solution which is not in the set $C_b[0,\infty)$ and is excluded for our problem.

6.5. Competitive Behaviour and Mass Behaviour

We now return to the system (6.1) when each agent is assigned an individual cost $J_i(u_i, v_i)$ which is the i-th component of J defined by (6.2). As a first step, after applying the optimal tracking control law

$$u_i = -\frac{b}{r}(\Pi z_i + s), \qquad (6.37)$$

with respect to a deterministic function z^*, where Π and s are determined by (6.20) and (6.22), the closed loop for the i-th player is

$$dz_i = (a - \frac{b^2}{r}\Pi)z_i dt - \frac{b^2}{r}s dt + \sigma dw_i. \qquad (6.38)$$

Denoting $\bar{z}_i(t) = Ez_i(t)$ and taking expectation on both sides of (6.38) yields

$$\frac{d\bar{z}_i}{dt} = (a - \frac{b^2}{r}\Pi)\bar{z}_i - \frac{b^2}{r}s, \tag{6.39}$$

where the initial condition is $\bar{z}_i|_{t=0} = Ez_i(0)$.

We further define the population average of means (simply called population mean) as $\bar{z} \triangleq \frac{1}{n}\sum_{i=1}^n \bar{z}_i$; then it is clear that $\bar{z}$ satisfies the same equation as $\bar{z}_i$, i.e.,

$$\frac{d\bar{z}}{dt} = (a - \frac{b^2}{r}\Pi)\bar{z} - \frac{b^2}{r}s, \tag{6.40}$$

where the initial condition is given by $\bar{z}|_{t=0} = \frac{1}{n}\sum_{i=1}^n Ez_i(0)$.

Here one naturally comes up with the important questions how the deterministic process z^* is chosen when it is applied to system (6.1) to approximate the influence of all other players on the given player, and in what way it captures the dynamic behaviour of the collection of many individuals. Since we wish to have $z^*(t) \approx z^*_{-i} = \gamma(\frac{1}{n}\sum_{k\neq i}^n z_k + \eta)$, for large n it is reasonable to express z^* in terms of the population mean $\bar{z}$ as

$$z^*(t) = \gamma(\bar{z}(t) + \eta), \tag{6.41}$$

whenever an equality for all time t is possible. We note that z^* defined above is used to approximate $z_i^* = \gamma(\frac{1}{n}\sum_{k\neq i}^n z_k + \eta)$ in the context of a fixed large n. As n increases, accuracy of this approximation is expected to improve. Subject to such an equality constraint, a dynamic interaction is built up between the individual and the mass. Specifically, based on the population mean $\bar{z}$ a tracking level z^* is determined by the rule (6.41) which is then used to compute the individual control law; in turn the individual control will lead to a corresponding mass behaviour. In the following we will address certain stability issues associated with such interactions.

Combining (6.22), (6.40) and (6.41) together and setting the derivatives as zero, we write a set of steady state equations as follows

$$\begin{cases} \gamma\bar{z}_\infty - z^*_\infty = -\gamma\eta, \\ z^*_\infty - (a - \frac{b^2}{r}\Pi - \rho)s_\infty = 0, \\ (a - \frac{b^2}{r}\Pi)\bar{z}_\infty - \frac{b^2}{r}s_\infty = 0, \end{cases} \tag{6.42}$$

Example 6.1. $a = 1$, $b = 1$, $\sigma = 0.3$, $\rho = 0.5$, $\gamma = 0.6$, $r = 0.1$, $\eta = 0.25$. We get

$$\Pi = 0.4, \quad (\bar{z}_\infty, z^*_\infty, s_\infty) = (0.333333, 0.35, -0.1).$$

□

We make the following key assumptions:

(H6.1) $\beta_1 > 0$, and $\frac{M}{\beta_1\beta_2} < 1$, where $M = \frac{b^2\gamma}{r}$, and β_1, β_2 are defined by (6.21). □

(H6.2) $z_i(0), 1 \le i \le n$, are mutually independent and $Ez_i^2(0) < C$ for C independent of n. □

Notice that the condition $\beta_1 > 0$ has been used in Proposition 6.2. It can be verified that **(H6.1)** holds for Example 6.1. Under Assumption **(H6.1)**, (6.42) is a nonsingular linear equation and has a unique solution $(\bar{z}_\infty, z^*_\infty, s_\infty)$.

Eliminating s in (6.40) by (6.22) and (6.41), we get the equation for the population mean

$$\frac{d\bar{z}}{dt} = (a - \frac{b^2}{r}\Pi)\bar{z} + \frac{b^2\gamma}{r}\int_t^\infty e^{\beta_2(t-\tau)}\bar{z}(\tau)d\tau + \frac{b^2\gamma\eta}{r\beta_2}. \tag{6.43}$$

For bounded $\bar{z}$ on $[0, \infty)$, the integral in (6.43) is well defined since $\beta_2 > 0$.

Theorem 6.1. Under Assumption **(H6.1)**, the integral-differential equation (6.43) subject to any initial condition $\bar{z}_0$ and the terminal condition $\lim_{t\to\infty}\bar{z}(t) = \bar{z}_\infty$ admits a unique solution.

PROOF. By (6.42) we have

$$(a - \frac{b^2}{r}\Pi)\bar{z}_\infty + \frac{b^2\gamma\bar{z}_\infty}{r\beta_2} + \frac{b^2\gamma\eta}{r\beta_2} = (a - \frac{b^2}{r}\Pi)\bar{z}_\infty - \frac{b^2}{r}s_\infty = 0. \tag{6.44}$$

Taking $\tilde{z} = \bar{z} - \bar{z}_\infty$ we rewrite (6.43) in the equivalent form

$$\frac{d\tilde{z}}{dt} = (a - \frac{b^2}{r}\Pi)\tilde{z} + \frac{b^2\gamma}{r}\int_t^\infty e^{\beta_2(t-\tau)}\tilde{z}(\tau)d\tau + (a - \frac{b^2}{r}\Pi)\bar{z}_\infty + \frac{b^2\gamma\bar{z}_\infty}{r\beta_2} + \frac{b^2\gamma\eta}{r\beta_2}$$

$$= -\beta_1\tilde{z} + M\int_t^\infty e^{\beta_2(t-\tau)}\tilde{z}(\tau)d\tau. \tag{6.45}$$

where $\beta_1 = -a + \frac{b^2}{r}\Pi$, $\beta_2 = -a + \frac{b^2}{r}\Pi + \rho$, $M = \frac{b^2\gamma}{r}$ and $\tilde{z}(t)$ satisfies the terminal condition $\lim_{t\to\infty}\tilde{z}(t) = 0$.

We write $\tilde{z} = e^{-\beta_1 t}\hat{z}$ and use a change of variable to obtain from (6.45)

$$\frac{d\hat{z}}{dt} = M\int_t^\infty e^{(\beta_1+\beta_2)(t-\tau)}\hat{z}(\tau)d\tau, \tag{6.46}$$

where $\hat{z}(t)$ satisfies the growth condition $\hat{z}(t) = o(e^{\beta_1 t})$. It is easily seen that the initial condition is $\hat{z}(0) = \bar{z}(0) - \bar{z}_\infty$. We write (6.46) in the equivalent form of a double integral equation

$$\hat{z}(t) = \hat{z}(0) + M\int_0^t\int_s^\infty e^{(\beta_1+\beta_2)(s-\tau)}\hat{z}(\tau)d\tau ds. \tag{6.47}$$

For analyzing existence and uniqueness of the solution to (6.47) we introduce the function class $\mathcal{C} = \{x \in C[0,\infty),\ \lim_{t\to\infty}e^{-\beta_1 t}x(t) = 0\}$, and set $\|x\| \triangleq \sup_{t\in[0,\infty)} e^{-\beta_1 t}|x(t)|$. Then it is straightforward to verify that under the norm $\|\cdot\|$, $\mathcal{C}$ is a Banach space. Define the map

$$F(x) = \bar{z}(0) - \bar{z}_\infty + M\int_0^t\int_s^\infty e^{(\beta_1+\beta_2)(s-\tau)}x(\tau)d\tau ds, \tag{6.48}$$

for $x \in \mathcal{C}$. It is obvious $F(x) \in C[0,\infty)$ for $x \in \mathcal{C}$. We verify that we also have $\lim_{t\to\infty} e^{-\beta_1 t}F(x) = 0$. For any fixed $\varepsilon > 0$ and $x \in \mathcal{C}$, there exists $T > 0$ such that

$e^{-\beta_1 t}|x(t)| < \varepsilon$ for all $t \geq T$. We denote $c = \sup_{t\in[0,\infty)} e^{-\beta_1 t}|x(t)|$. For $t > T$, we have

$$\begin{aligned}
&|e^{-\beta_1 t}\int_0^t\int_s^\infty e^{(\beta_1+\beta_2)(s-\tau)}x(\tau)d\tau ds| \\
=&|e^{-\beta_1 t}\int_0^T\int_s^\infty e^{(\beta_1+\beta_2)(s-\tau)}x(\tau)d\tau ds + e^{-\beta_1 t}\int_T^t\int_s^\infty e^{(\beta_1+\beta_2)(s-\tau)}x(\tau)d\tau ds| \\
\leq& e^{-\beta_1 t}\int_0^T\int_s^\infty e^{(\beta_1+\beta_2)s}e^{-\beta_2\tau}e^{-\beta_1\tau}|x(\tau)|d\tau ds + e^{-\beta_1 t}\int_T^t\int_s^\infty e^{(\beta_1+\beta_2)(s-\tau)}\varepsilon e^{\beta_1\tau}d\tau ds \\
\leq& e^{-\beta_1 t}\int_0^T\int_s^\infty e^{(\beta_1+\beta_2)s}e^{-\beta_2\tau}c\,d\tau ds + e^{-\beta_1 t}\int_T^t\int_s^\infty e^{(\beta_1+\beta_2)(s-\tau)}\varepsilon e^{\beta_1\tau}d\tau ds \\
\leq& \frac{ce^{\beta_1 T}-1}{\beta_1\beta_2}e^{-\beta_1 t} + \frac{\varepsilon}{\beta_1\beta_2}(1-e^{\beta_1(T-t)}).
\end{aligned} \tag{6.49}$$

It follows from (6.48)-(6.49) that for sufficiently large t we have $e^{-\beta_1 t}|F(x)| \leq \frac{2M\varepsilon}{\beta_1\beta_2}$. Since $\varepsilon > 0$ is arbitrary, we have $\lim_{t\to\infty} e^{-\beta_1 t}F(x) = 0$ and hence $F(x) \in \mathcal{C}$. Next we establish a contractive property for F. For $x_1, x_2 \in \mathcal{C}$, we have

$$\begin{aligned}
&|e^{-\beta_1 t}[F(x_1)-F(x_2)]| = M|e^{-\beta_1 t}\int_0^t\int_s^\infty e^{(\beta_1+\beta_2)(s-\tau)}[x_1(\tau)-x_2(\tau)]d\tau ds| \\
=&M|e^{-\beta_1 t}\int_0^t\int_s^\infty e^{(\beta_1+\beta_2)(s-\tau)}e^{\beta_1\tau}e^{-\beta_1\tau}[x_1(\tau)-x_2(\tau)]d\tau ds| \\
\leq& M\|x_1-x_2\|\cdot e^{-\beta_1 t}\int_0^t\int_s^\infty e^{(\beta_1+\beta_2)(s-\tau)}e^{\beta_1\tau}d\tau ds \\
=&\frac{M}{\beta_1\beta_2}(1-e^{-\beta_1 t})\|x_1-x_2\|,
\end{aligned} \tag{6.50}$$

and therefore

$$\|F(x_1)-F(x_2)\| \leq \frac{M}{\beta_1\beta_2}\|x_1-x_2\|. \tag{6.51}$$

Since $\frac{M}{\beta_1\beta_2} < 1$, F is a contraction on $\mathcal{C}$ so that (6.48) has a unique solution in $\mathcal{C}$. Hence the integral-differential equation (6.43) has a unique solution satisfying $\lim_{t\to\infty}\overline{z}(t) = \overline{z}_\infty$. This completes the proof. □

6.5.1. An Analytic Solution to the Equation System. We sketch computing analytic expressions for $\overline{z}$ and s as follows: The calculation will first be stated in terms of $\widetilde{z}$ and $\widetilde{s}$, where $\widetilde{z}(t) = \overline{z}(t) - \overline{z}_\infty$ and $\widetilde{s}(t) = s(t) - s_\infty$.

Taking differentiation on both sides of (6.45) gives

$$\frac{d^2\tilde{z}}{dt^2} = -\beta_1\frac{d\tilde{z}}{dt} + \beta_2 M\int_t^\infty e^{\beta_2(t-\tau)}\tilde{z}(\tau)d\tau - M\tilde{z}(t)$$

which combined again with (6.45) yields

$$\frac{d^2\tilde{z}}{dt^2} - \rho\frac{d\tilde{z}}{dt} + (M - \beta_1\beta_2)\tilde{z} = 0. \tag{6.52}$$

The characteristic equation of (6.52) is $\lambda^2 - \rho\lambda + (M - \beta_1\beta_2)$ with two distinct egein-values:

$$\lambda_1 = \frac{\rho - \sqrt{\rho^2 + 4(\beta_1\beta_2 - M)}}{2} < 0, \quad \lambda_2 = \frac{\rho + \sqrt{\rho^2 + 4(\beta_1\beta_2 - M)}}{2} > 0.$$

Recalling the growth condition for $\overline{z}$ (i.e., we are interested only in bounded $\overline{z}$) and hence for $\tilde{z}$, we have

$$\tilde{z} = \tilde{z}(0)e^{\lambda_1 t} = (\overline{z}(0) - \overline{z}_\infty)e^{\lambda_1 t}, \tag{6.53}$$

and it is readily verified that $\tilde{z}$ is a solution to (6.45).

On the other hand, from (6.22) and (6.41) we get

$$\frac{d\tilde{s}}{dt} = \beta_2\tilde{s} + \gamma\tilde{z}. \tag{6.54}$$

Assuming initial condition $\tilde{s}(0) = s(0) - s_\infty$, we obtain

$$\begin{aligned}\tilde{s}(t) &= \tilde{s}(0)e^{\beta_2 t} + e^{\beta_2 t}\int_0^t e^{-\beta_2\tau}\gamma[\overline{z}(0) - \overline{z}_\infty]e^{\lambda_1\tau}d\tau \\ &= \tilde{s}(0)e^{\beta_2 t} + e^{\beta_2 t}\int_0^\infty e^{-\beta_2\tau}\gamma[\overline{z}(0) - \overline{z}_\infty]e^{\lambda_1\tau}d\tau \\ &\quad - e^{\beta_2 t}\int_t^\infty e^{-\beta_2\tau}\gamma[\overline{z}(0) - \overline{z}_\infty]e^{\lambda_1\tau}d\tau\end{aligned} \tag{6.55}$$

Setting the initial condition in (6.55) as

$$\tilde{s}(0) = -\int_0^\infty e^{-\beta_2\tau}\gamma[\overline{z}(0) - \overline{z}_\infty]e^{\lambda_1\tau}d\tau, \tag{6.56}$$

we get

$$\widetilde{s}(t) = \frac{\gamma}{\beta_2 - \lambda_1}(\overline{z}_\infty - \overline{z}(0))e^{\lambda_1 t}. \tag{6.57}$$

Notice that any initial condition $\widetilde{s}|_{t=0}$ other than (6.56) yields $\widetilde{s}$ and s with a growth rate of $e^{\beta_2 t}$ which is excluded here.

We summarize the above calculation to get the following proposition:

Proposition 6.4. If **(H6.1)** holds, the unique asymptotically convergent solution $(\overline{z}, s)$ determined by (6.22), (6.40) and (6.41), is given by

$$\overline{z}(t) = \overline{z}_\infty + (\overline{z}(0) - \overline{z}_\infty)e^{\lambda_1 t},$$
$$s(t) = s_\infty + \frac{\gamma}{\beta_2 - \lambda_1}(\overline{z}_\infty - \overline{z}(0))e^{\lambda_1 t},$$

where $\beta_2 = -a + \frac{b^2}{r}\Pi + \rho$, and $\lambda_1 = \frac{\rho - \sqrt{\rho^2 + 4(\beta_1\beta_2 - M)}}{2} < 0$. □

6.5.2. The Decentralized ε-Nash Equilibrium. In the current context we give the definition of Nash equilibrium.

Definition 6.1. [6] A set of controls $u_k \in \mathcal{U}_k, 1 \leq k \leq n$, for n players where $\mathcal{U}_k$ is a specified class of measurable functions of the state processes $z_1(\cdot), \cdots, z_n(\cdot)$, such that the resulting v_k is adapted to some subfiltration of the underlying Brownian motion, is called a *Nash equilibrium* with respect to the costs $J_k(u_k, v_k), 1 \leq k \leq n$, if for any fixed $1 \leq i \leq n$, we have

$$J_i(u_i, v_i(u_1, \cdots, u_{i-1}, u_{i+1} \cdots, u_n)) \leq J_i(u_i', v_i(u_1, \cdots, u_{i-1}, u_{i+1}, \cdots, u_n)),$$

when any alternative control $u_i' \in \mathcal{U}_i$ is applied by the i-th player. □

Definition 6.2. A set of controls $u_k \in \mathcal{U}_k, 1 \leq k \leq n$, for n players is called an *ε-Nash equilibrium* with respect to the costs $J_k, 1 \leq k \leq n$, if there exists $\varepsilon > 0$ such that for any fixed $1 \leq i \leq n$, we have

$$J_i(u_i, v_i(u_1, \cdots, u_{i-1}, u_{i+1} \cdots, u_n)) \leq J_i(u_i', v_i(u_1, \cdots, u_{i-1}, u_{i+1}, \cdots, u_n)) + \varepsilon,$$

when any alternative control $u_i' \in \mathcal{U}_k$ is applied by the i-th player. □

In the following we use $J_i(u_i, v_i(u_1, \cdots, u_{i-1}, \cdots, u_{i+1}, \cdots, u_n))$ to denote the individual cost with respect to the coupled reference trajectory $v_i = \gamma(\frac{1}{n}\sum_{k\neq i}^{n} z_k(u_k) + \eta)$ for the i-th player when player k applies control u_k, $1 \le k \le n$, and n is the population size. Let

$$J_i(u_i, v_i(u_1^0, \cdots, u_{i-1}^0, u_{i+1}^0 \cdots, u_n^0))$$
$$\triangleq E\int_0^\infty e^{-\rho t}\{[z_i(u_i) - \gamma(\frac{1}{n}\sum_{k\neq i}^{n} z_k(u_k^0) + \eta)]^2 + ru_i^2\}dt, \tag{6.58}$$

where $z_k(u_k^0) = z_k(u_k^0(z^*, z_k))$. Here we use u_i^0 to denote the optimal tracking based control law for the i-th player, i.e.,

$$u_i^0 = -\frac{b}{r}(\Pi z_i + s), \tag{6.59}$$

where s and the associated z^* are derived from (6.22), (6.40) and (6.41). In particular,

$$J_i(u_i^0, v_i(u_1^0, \cdots, u_{i-1}^0, u_{i+1}^0 \cdots, u_n^0)) = J_i(u_i, v_i(u_1^0, \cdots, u_{i-1}^0, u_{i+1}^0 \cdots, u_n^0))|_{u_i = u_i^0}.$$

Notice that the initial condition of $\overline{z}$ is take as $\frac{1}{n}\sum_{k=1}^{n} Ez_k(0)$, which further induces the initial condition of $z^*(t) = \gamma(\overline{z}(t) + \eta)$.

Denote $\sigma_0^2 = \sup_{1\le i\le n} E[z_i(0) - Ez_i(0)]^2$, $\sigma_0 \ge 0$. In the case all $z_i(0)$ become deterministic, we simply have $\sigma_0 = 0$.

Lemma 6.1. Under **(H6.1)-(H6.2)**, for z^* determined by (6.22), (6.40) and (6.41), we have

$$E\int_0^\infty e^{-\rho t}[z^* - \gamma(\frac{1}{n}\sum_{k\neq i}^{n} z_k(u_k^0) + \eta)]^2 dt = O(\frac{\sigma^2 + \sigma_0^2}{n} + \frac{1}{n^2}),$$

where the state $z_k(u_k^0)$ of player k, $k \neq i$, is generated by the optimal tracking based control law u_k^0 given by (6.59) for the k-th player.

PROOF. By equations (6.39), (6.40) and their initial conditions $\bar{z}_i|_{t=0} = Ez_i(0)$, $\bar{z}|_{t=0} = \frac{1}{n}\sum_{i=1}^n Ez_i(0)$, it follows that

$$z^* - \gamma(\frac{1}{n}\sum_{k\neq i}^n z_k + \eta) = \gamma(\frac{1}{n}\sum_{k=1}^n Ez_k + \eta) - \gamma(\frac{1}{n}\sum_{k\neq i}^n z_k + \eta)$$

$$= \frac{\gamma}{n}Ez_i - \frac{\gamma}{n}\sum_{k\neq i}^n (z_k - Ez_k), \tag{6.60}$$

where we simply write $z_k(u_k^0)$ as z_k. Writing $\tilde{z}_{n,i} \triangleq \frac{1}{n}\sum_{k\neq i}^n (z_k - Ez_k)$, we have

$$d\tilde{z}_{n,i} = -\beta_1 \tilde{z}_{n,i}dt + \frac{\sigma}{n}\sum_{k\neq i}^n dw_k, \quad t \geq 0. \tag{6.61}$$

By directly solving (6.61) and recalling Assumptions **(H6.1)-(H6.2)**, it follows that there exists a constant C_1 independent of i and n such that $\sup_{t\geq 0} E\tilde{z}_{n,i}^2(t) \leq C_1 \cdot \frac{\sigma^2+\sigma_0^2}{n}$ and moreover, $\sup_{t\geq 0}|Ez_i(t)| \leq C_1$. Consequently, from (6.60) we get

$$E[z^* - \gamma(\frac{1}{n}\sum_{k\neq i}^n z_k + \eta)]^2 = O(\frac{\sigma^2+\sigma_0^2}{n} + \frac{1}{n^2}), \tag{6.62}$$

and the lemma follows. □

Theorem 6.2. Under **(H6.1)-(H6.2)**, we have

$$|J_i(u_i^0, \gamma(\frac{1}{n}\sum_{k\neq i}^n z_k(u_k^0) + \eta)) - J_i(u_i^0, z^*)| = O(\frac{\sigma+\sigma_0}{\sqrt{n}} + \frac{1}{n}), \tag{6.63}$$

where $J_i(u_i^0, z^*)$ is the individual cost with respect to z^*, and u_i^0 is given by (6.59). □

The proof is postponed until after Theorem 6.3.

Theorem 6.3. Under **(H6.1)-(H6.2)**, the set of controls $u_i^0, 1 \leq i \leq n$, for the n players is an ε-Nash equilibrium with respect to the costs $J_i(u_i, \gamma(\frac{1}{n}\sum_{k\neq i}^n z_k(u_k) +$

$\eta)), 1 \le i \le n$, with $\varepsilon = O(\frac{\sigma+\sigma_0}{\sqrt{n}} + \frac{1}{n})$, i.e., for any i, we have

$$J_i(u_i^0, \gamma(\frac{1}{n}\sum_{k\neq i}^{n} z_k(u_k^0) + \eta)) - O(\frac{\sigma+\sigma_0}{\sqrt{n}} + \frac{1}{n}) \tag{6.64}$$

$$\le \inf_{u_i} J_i(u_i, \gamma(\frac{1}{n}\sum_{k\neq i}^{n} z_k(u_k^0) + \eta)) \tag{6.65}$$

$$\le J_i(u_i^0, \gamma(\frac{1}{n}\sum_{k\neq i}^{n} z_k(u_k^0) + \eta)), \tag{6.66}$$

where u_k^0 is the optimal tracking based control law given by (6.59) for the k-th player, and u_i is any alternative control which depends on $(t, z_1, \cdots, z_n)$.

PROOF. The inequality (6.66) is obviously true. We prove the inequality (6.65). For any full state dependent u_i satisfying

$$J_i(u_i, \gamma(\frac{1}{n}\sum_{k\neq i}^{n} z_k(u_k^0) + \eta)) \le J_i(u_i^0, \gamma(\frac{1}{n}\sum_{k\neq i}^{n} z_k(u_k^0) + \eta)) \tag{6.67}$$

we can find a fixed constant C independent of n such that

$$J_i(u_i, \gamma(\frac{1}{n}\sum_{k\neq i}^{n} z_k(u_k^0) + \eta))$$

$$= E\int_0^\infty e^{-\rho t}\{[z_i(u_i) - \gamma(\frac{1}{n}\sum_{k\neq i}^{n} z_k(u_k^0) + \eta)]^2 + r u_i^2\} dt \le C. \tag{6.68}$$

Here and hereafter in the proof, $(z_i(u_i), u_i), (z_k(u_k^0), u_k^0), k \neq i$, denote the corresponding state-control pairs. For notational brevity in the following we omit the associated control in $z_i(u_i)$, $z_k(u_k^0)$, $k \neq i$ and simply write z_i, z_k without causing confusion. Since all $z_k, k \neq i$, are fixed after the control u_k^0 is selected, for (u_i, z_i) satisfying (6.68) there exists $C > 0$ independent of n such that

$$E\int_0^\infty e^{-\rho t} z_i^2 dt \le C, \quad E\int_0^\infty e^{-\rho t}(z_i - z^*(t))^2 dt \le C. \tag{6.69}$$

On the other hand we have

$$
\begin{aligned}
&E\int_0^\infty e^{-\rho t}\{[z_i-\gamma(\frac{1}{n}\sum_{k\neq i}^{n} z_k(u_k^0)+\eta)]^2+ru_i^2\}dt\\
=&E\int_0^\infty e^{-\rho t}\{[(z_i-z^*)+(z^*-\gamma(\frac{1}{n}\sum_{k\neq i}^{n} z_k+\eta))]^2+ru_i^2\}dt\\
=&E\int_0^\infty e^{-\rho t}[(z_i-z^*)^2+ru_i^2]dt+E\int_0^\infty e^{-\rho t}[z^*-\gamma(\frac{1}{n}\sum_{k\neq i}^{n} z_k+\eta)]^2dt\\
&+2E\int_0^\infty e^{-\rho t}(z_i-z^*)[z^*-\gamma(\frac{1}{n}\sum_{k\neq i}^{n} z_k+\eta)]dt \triangleq I_1+I_2+I_3. \qquad (6.70)
\end{aligned}
$$

Then we have

$$
\begin{aligned}
I_1 &= J(u_i,z^*)\geq J_i(u_i^0,z^*)\\
&\geq J_i(u_i^0,\gamma(\frac{1}{n}\sum_{k\neq i}^{n} z_k(u_k^0)+\eta))-O(\frac{\sigma+\sigma_0}{\sqrt{n}}+\frac{1}{n}), \qquad (6.71)\\
I_2 &= O(\frac{\sigma^2+\sigma_0^2}{n}+\frac{1}{n^2}) \qquad (6.72)
\end{aligned}
$$

where (6.71) follows from Theorem 6.2 and (6.72) follows from Lemma 6.1. Moreover

$$
\begin{aligned}
|I_3| &\leq 2\int_0^\infty e^{-\rho t}[E(z_i-z^*)^2]^{\frac{1}{2}}\{E[z^*-\gamma(\frac{1}{n}\sum_{k\neq i}^{n} z_k+\eta)]^2\}^{\frac{1}{2}}dt\\
&\leq 2[\int_0^\infty e^{-\rho t}E(z_i-z^*)^2dt]^{\frac{1}{2}}\{\int_0^\infty e^{-\rho t}[z^*-\gamma(\frac{1}{n}\sum_{k\neq i}^{n} z_k+\eta)]^2dt\}^{\frac{1}{2}}\\
&= O(\sqrt{I_2})=O(\sqrt{\frac{\sigma^2+\sigma_0^2}{n}+\frac{1}{n^2}})=O(\frac{\sigma+\sigma_0}{\sqrt{n}}+\frac{1}{n}). \qquad (6.73)
\end{aligned}
$$

Hence it follows from the above estimates that there exists $c>0$ such that

$$
J_i(u_i,\gamma(\frac{1}{n}\sum_{k\neq i}^{n} z_k(u_k^0)+\eta))\geq J_i(u_i^0,\gamma(\frac{1}{n}\sum_{k\neq i}^{n} z_k(u_k^0)+\eta))-c(\frac{\sigma+\sigma_0}{\sqrt{n}}+\frac{1}{n}), \qquad (6.74)
$$

where c is independent of σ^2 and n. This completes the proof. □

In other words, when all the players $k = 1, \cdots i-1, i+1, \cdots, n$, retain their decentralized controls u_k^0 and the i-th player is allowed to use a full state based control u_i, it can reduce its cost at most by $O(\frac{\sigma+\sigma_0}{\sqrt{n}} + \frac{1}{n})$.

Proof of Theorem 6.2. Similar to (6.70) we have

$$
\begin{aligned}
&J_i(u_i^0, \gamma(\frac{1}{n}\sum_{k\neq i}^{n} z_k(u_k^0) + \eta)) \\
=&E\int_0^\infty e^{-\rho t}\{[z_i(u_i^0) - \gamma(\frac{1}{n}\sum_{k\neq i}^{n} z_k(u_k^0) + \eta)]^2 + r(u_i^0)^2\}dt \\
=&E\int_0^\infty e^{-\rho t}\{[(z_i(u_i^0) - z^*) + (z^* - \gamma(\frac{1}{n}\sum_{k\neq i}^{n} z_k(u_k^0) + \eta))]^2 + r(u_i^0)^2\}dt \\
=&J_i(u_i^0, z^*) + E\int_0^\infty e^{-\rho t}[z^* - \gamma(\frac{1}{n}\sum_{k\neq i}^{n} z_k(u_k^0) + \eta)]^2 dt \\
&+ 2E\int_0^\infty e^{-\rho t}(z_i(u_i^0) - z^*)[z^* - \gamma(\frac{1}{n}\sum_{k\neq i}^{n} z_k(u_k^0) + \eta)]dt \\
\triangleq&J_i(u_i^0, z^*) + I_2' + I_3'.
\end{aligned}
\tag{6.75}
$$

Finally, similar to (6.72) and (6.73), we have

$$|I_2' + I_3'| = O(\frac{\sigma+\sigma_0}{\sqrt{n}} + \frac{1}{n}), \tag{6.76}$$

and this completes the proof. □

6.5.3. The Virtual Agent, Policy Iteration and Attraction to Mass Behaviour. In this subsection we investigate certain asymptotic properties on the interaction between the individual and the mass, and the formulation shall be interpreted in the large population limit (i.e., an infinite population) context.

Assume each agent is assigned a cost according to (6.19), i.e.,

$$J_i(u_i, z^*) = E\int_0^\infty e^{-\rho t}\{[z_i - z^*(t)]^2 + ru_i^2\}ds, \qquad i \geq 1. \tag{6.77}$$

We now introduce a *virtual central agent* (or simply *virtual agent*) (VA) to represent the mass effect and use the function z^* to describe the behaviour of the virtual agent. Here the virtual agent acts as a passive player in the game in the sense that it does not have the freedom to change z^* by its own will. Instead, after each selection of the individual control laws, a new z^* will be induced as specified below; subsequently, the individual shall consider its optimal policy to respond to this new z^*. Thus, the interplay between a given individual and the virtual agent may be described in terms of a series of asynchronous plays. In the following policy iteration analysis, we take the virtual agent as a passive leader and the individual agents as active followers.

Suppose that there is a priori $z^{*(k)} \in C_b[0,\infty) \triangleq \{x \in C[0,\infty),\ \sup_{t\in[0,\infty)} |x(t)| < \infty\}$, $k \geq 0$; then by Proposition 6.2 the optimal control for the i-th agent with respect to $z^{*(k)}$ is given as

$$u_i^{(k+1)} = -\frac{b}{r}(\Pi z_i + s^{(k+1)}) \tag{6.78}$$

where

$$\rho s^{(k+1)} = \frac{ds^{(k+1)}}{dt} + as^{(k+1)} - \frac{b^2}{r}\Pi s^{(k+1)} - z^{*(k)}, \quad s^{(k+1)}(t) \in C_b[0,\infty). \tag{6.79}$$

The unique solution $s^{(k+1)} \in C_b[0,\infty)$ to (6.79) can be represented by the map

$$s^{(k+1)} = -e^{\beta_2 t}\int_t^\infty e^{-\beta_2\tau} z^{*(k)}(\tau)d\tau. \tag{6.80}$$

Subsequently, by the control law $u_i^{(k+1)}$ the corresponding population mean is described by

$$\frac{d\overline{z}^{(k+1)}}{dt} = (a - \frac{b^2}{r}\Pi)\overline{z}^{(k+1)} - \frac{b^2}{r}s^{(k+1)}, \tag{6.81}$$

where the inital condition is $\overline{z}^{(k+1)}|_{t=0} = \overline{z}_0$ for all k, and z_0 is the initial value of the population mean. (6.81) shall be interpreted as the limiting version of (6.40) in the infinite population case.

Then the virtual agent's state z^* corresponding to $u_i^{(k+1)}$ is determined as

$$z^{*(k+1)} = \gamma(\overline{z}^{(k+1)} + \eta) \tag{6.82}$$

which is a recursive version of (6.41). From (6.81)-(6.82) we have

$$\frac{dz^{*(k+1)}}{dt} = -\beta_1 z^{*(k+1)} - \frac{\gamma b^2}{r} s^{(k+1)} + \beta_1 \gamma \eta. \tag{6.83}$$

Combining (6.80) and (6.83) gives

$$\frac{dz^{*(k+1)}}{dt} = -\beta_1 z^{*(k+1)} + \frac{\gamma b^2}{r} e^{\beta_2 t} \int_t^\infty e^{-\beta_2 \tau} z^{*(k)}(\tau) d\tau + \beta_1 \gamma \eta, \tag{6.84}$$

where the initial condition is $z^{*(k+1)} = z_0^*$ for all k. In addition, $z_0^* = \gamma(\overline{z}_0 + \eta)$.

For $x \in C_b[0, \infty)$, define the norm $\|x\|_b = \sup_{t \in [0,\infty)} |x(t)|$.

We now suppose Assumption **(H6.1)** holds; then $\beta_1 > 0$. Obviously, for any $z^{*(k)} \in C_b[0, \infty)$ there exists a unique solution $z^{*(k+1)}$ to (6.84) which is also in $C_b[0, \infty)$. From (6.84) we induce a map $\mathcal{L}_0 : C_b[0, \infty) \to C_b[0, \infty)$ such that the unique solution can be represented by

$$z^{*(k+1)} \triangleq \mathcal{L}_0 z^{*(k)}. \tag{6.85}$$

Theorem 6.4. If **(H6.1)** holds, the map $\mathcal{L}_0$ is a contraction on $C_b[0, \infty)$.

PROOF. We take $z^{*(k)}, y^{*(k)} \in C_b[0, \infty)$ and set

$$z^{*(k+1)} = \mathcal{L}_0 z^{*(k)} \tag{6.86}$$

$$y^{*(k+1)} = \mathcal{L}_0 y^{*(k)} \tag{6.87}$$

Denote $\Delta_1 = z^{*(k+1)} - y^{*(k+1)}$, $\Delta_0 = z^{*(k)} - y^{*(k)}$. We have

$$\frac{d\Delta_1}{dt} = -\beta_1 \Delta_1 + \frac{\gamma b^2}{r} e^{\beta_2 t} \int_t^\infty e^{-\beta_2 \tau} \Delta_0(\tau) d\tau \tag{6.88}$$

Since $\Delta_1|_{t=0} = z^{*(k+1)}|_{t=0} - y^{*(k+1)}|_{t=0} = z_0^* - z_0^* = 0$, it follows that

$$|\Delta_1(t)| = |\int_0^t e^{-\beta_1(t-s)}\frac{\gamma b^2}{r}e^{\beta_2 s}\left(\int_s^\infty e^{-\beta_2\tau}\Delta_0(\tau)d\tau\right)ds|$$
$$\leq \frac{\gamma b^2}{r}\|\Delta_0\|_b\int_0^t e^{-\beta_1(t-s)}e^{\beta_2 s}\left(\int_s^\infty e^{-\beta_2\tau}d\tau\right)ds \leq \frac{\gamma b^2}{r\beta_1\beta_2}\|\Delta_0\|_b, \qquad (6.89)$$

so that

$$\|\Delta_1\|_b \leq \frac{\gamma b^2}{r\beta_1\beta_2}\|\Delta_0\|_b \qquad (6.90)$$

where $\frac{\gamma b^2}{r\beta_1\beta_2} < 1$ by Assumption **(H6.1)**, and therefore $\mathcal{L}_0$ is a contraction. □

By the asynchronous updation of the individual strategies against the virtual agent, we induce the mass behaviour by a sequence of functions $z^{*(k)} = \mathcal{L}_0 z^{*(k-1)} = \mathcal{L}_0^k z^{*(0)}$. We have the proposition.

Proposition 6.5. Under **(H6.1)**, $\lim_{k\to\infty} z^{*(k)} = z^*$ where z^* is determined by (6.22), (6.40) and (6.41).

PROOF. This follows from Theorem 6.4. □

The above proposition reveals certain stability and attraction feature of the evolution of the individual and mass behaviour.

6.6. A Cost Gap between the Centralized Optimal Control and Decentralized Tracking

As shown by the analysis in the foregoing Sections of this Chapter, for the underlying large population system the global cost based optimal control (6.17) and the individual cost based control (tracking) (6.59) have very different nature, which may be further illustrated by means of the resulting costs and the state trajectories in the two cases.

For a comparison of the costs associated with the two different methods, we assume the initial state $z_i(0)$ of all agents is 0 in the two cases. Let n be the cardinality of the population. We scale the global optimal cost (with 0 intial state for all players)

$v(0) \stackrel{\Delta}{=} \inf J|_{z_i=0,1\le i\le n} = \inf(\sum_{i=1}^n J_i)|_{z_i(0)=0,1\le i\le n}$ in Section 6.3 by n to get $\overline{v}_n(0) = \frac{v(0)}{n}$, and set $\overline{v}(0) = \lim_{n\to\infty} \overline{v}_n(0)$. Here $\overline{v}(0)$ may be intepretated as the optimal cost incurred per agent with identically 0 intial state. By (6.18), we have

$$\overline{v}(0) = \frac{\gamma^2\eta^2}{\rho}\left[1 - \frac{\overline{b}^2(\gamma-1)^2}{(\frac{\rho}{2} + \sqrt{\overline{a}^2 + (1-\gamma)^2\overline{b}^2})^2}\right] + \frac{\sigma^2(\overline{a} + \sqrt{\overline{a}^2 + \overline{b}^2})}{\rho\overline{b}^2}. \tag{6.91}$$

We now consider the case of the LQG game. In the large population limit, when each agent applies the optimal tracking based control law $u_i^0 = -\frac{b}{r}(\Pi z_i + s)$, let $v_i(0)$ be the resulting individual cost, where again we assume the initial state is 0 for all agents. Write $v_{ind}(0) = v_i(0)$ for any i since all agents have 0 initial state.

With s and z^* determined from Proposition 6.4, one can get from (6.31) a solution $q \in C_b[0,\infty)$ if and only if the initial condition is given by

$$q(0) = \left\{(\beta_2^2 - \overline{b}^2)\frac{1}{\rho} + \frac{2\overline{b}^2\gamma}{(\rho-\lambda_1)\beta_1}(\frac{\overline{b}^2}{\beta_1+\lambda_2} - \beta_2) + \frac{\gamma^2\overline{b}^4}{(\rho-2\lambda_1)\beta_1^2}[1 - \frac{\overline{b}^2}{(\beta_1+\lambda_2)^2}]\right\} s_\infty^2 + \frac{\Pi\sigma^2}{\rho} \tag{6.92}$$

and it is clear $v_{ind}(0) = q(0)$ by Proposition 6.3.

By the fact $\Pi = \frac{\overline{a}+\sqrt{\overline{a}^2+\overline{b}^2}}{\overline{b}^2}$ and $s_\infty = -\frac{\beta_1\beta_2\eta}{\beta_1\beta_2 - \overline{b}^2\gamma}\gamma$ given by (6.42), we derive from (6.91) and (6.92) that

$$|\overline{v}(0) - v_{ind}(0)| = O(\gamma^2).$$

The *gap* between $\overline{v}(0)$ and $v_{ind}(0)$ is demonstrated in Figure 6.1.

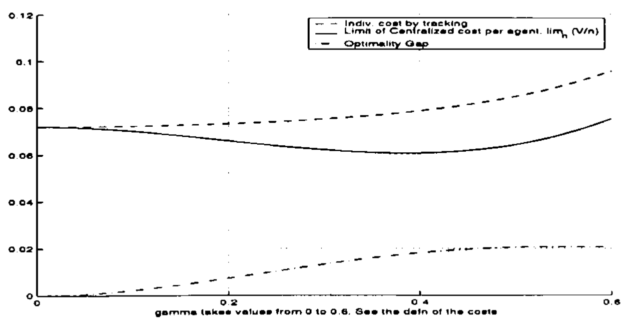


FIGURE 6.1. Top: Individual tracking based cost $v_{ind}(0)$; Middle: Scaled global cost $\overline{v}(0)$; Bottom: The cost gap $|\overline{v}(0) - v_{ind}(0)|$.

We have the following observations: If each agent applies the global cost based optimal control (6.17), all of them will be in a better situation compared to the case of everyone applying the optimal tracking based control law (6.59). However this universal well-being requires a strong coordination between all the agents, and greedy attempts from individuals easily destroy the global optimality. This means that when all agents are applying the global cost based control law, any individual player should be restrained from taking advantage of the other agents' presumably fixed control strategies by selfishly moving to a new strategy for reducing its own cost. In contrast, the individual cost based control is robust under greedy individual strategies as indicated by its ε-Nash equilibrium property.

Subsequently, we examine the state trajectories of the two control designs.

Suppose in a large population system $\mathcal{S}$, the dynamics for the agents is given by: $a = b = 1, \sigma = 0.05, \rho = 0.5, \gamma = 0.6, r = 0.1, \eta = 0.25$. The population mean $\frac{1}{n}\sum_{i=1}^{n} Ez_i(0) = 0.1$ (this will be used to set the initial condition $\overline{z}|_{t=0} = \frac{1}{n}\sum_{i=1}^{n} Ez_i(0) = 0.1$ for the mass).

Figure 6.2 shows the behaviour of two agents, labelled by 1 and 2. Both agents 1 and 2 are sampled from the above system $\mathcal{S}$ and apply the tracking control law (6.59). Both agents have different initial conditions but eventually their trajectories merge together.

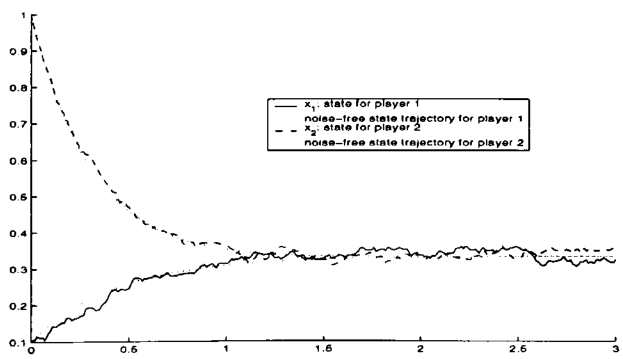


FIGURE 6.2. Trajectories of players 1 and 2

Now we analyze the optimal control. Recalling (6.17), in the global cost based control law (i.e., the centralized information optimal control law) the i-th agent's control is

$$u_i = -\frac{b}{r}pz_i - \frac{b}{r}q\sum_{k\neq i}^{n} z_k - \frac{b}{r}s_i, \tag{6.93}$$

where p, q and s_i depend on n which is the population cardinality. We analyze the asymptotic behaviour of the closed-loop system as $n \to \infty$. Set $\overline{z}_n = \frac{1}{n}\sum_{i=1}^{n} z_i$, and $\widetilde{z}_n = \frac{1}{n}\sum_{i=1}^{n}(z_i - Ez_i)$. Then we have

$$d\widetilde{z}_n = [a - \overline{b}^2 p(n) - \overline{b}^2 q(n)]\widetilde{z}_n dt + \frac{\sigma}{n}\sum_{i=1}^{n} dw_i. \tag{6.94}$$

We assume that in the increasing population context the initial state of all agents is deterministic and $\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} z_i(0)$ has a finite limit. Then using (6.94) it is easy to verify that for any $T > 0$

$$\sup_{0\le t\le T}|\widetilde{z}_n(t)| = \sup_{0\le t\le T}|\overline{z}_n - E\overline{z}_n| \xrightarrow{P.} 0, \tag{6.95}$$

as $n \to \infty$, where $\xrightarrow{P.}$ means convergence in probability.

Thus in the control of agent 1, with (6.95) in mind we approximate $\frac{1}{n}\sum_{k\neq 1} z_k$ by the limit $\overline{z} \triangleq \lim_{n\to\infty}\overline{z}_n$ which satisfies

$$\frac{d\overline{z}}{dt} = (\frac{\rho}{2} - \sqrt{\overline{a}^2 + (1-\gamma)^2\overline{b}^2})\overline{z} - \frac{\overline{b}^2\gamma\eta(\gamma - 1)}{\frac{\rho}{2} + \sqrt{\overline{a}^2 + (1-\gamma)^2\overline{b}^2}}. \tag{6.96}$$

Here (6.96) is derived from the closed-loop equtation for $z_i, 1 \le i \le n$, by first summing over z_i to write the equation for $\sum_{i=1}^{n} z_i$ and then taking $\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} z_i$.

We combine (6.93) with (6.96) and take a large population limit with the n agent based optimal control law to write the closed-loop dynamics for player 1 in the following form.

$$\begin{aligned} dz_1 =& (\frac{\rho}{2} - \sqrt{\overline{a}^2 + \overline{b}^2})z_1 dt + \left[\sqrt{\overline{a}^2 + \overline{b}^2} - \sqrt{\overline{a}^2 + (1-\gamma)^2\overline{b}^2}\right]\overline{z}dt \\ &+ \frac{\overline{b}^2\gamma\eta(1-\gamma)}{\frac{\rho}{2} + \sqrt{\overline{a}^2 + (1-\gamma)^2\overline{b}^2}} + \sigma dw_1. \end{aligned} \tag{6.97}$$

Figure 6.3 compares typical trajectories for two control laws, where the lower state trajectory is generated by the dynamics (6.97). It is seen that when the large population limit version of the global cost based optimal control law (6.93) is applied, the resulting state trajectory is generally below the one generated from the optimal tracking based control law (6.59).

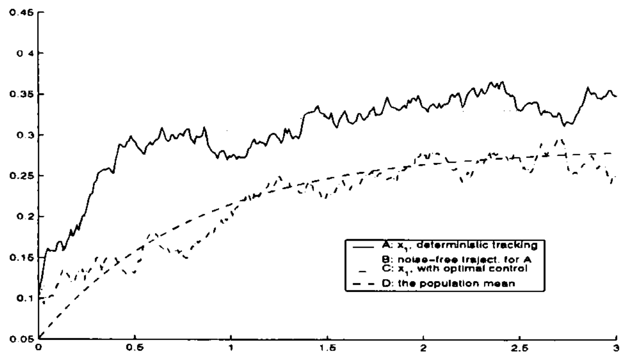


FIGURE 6.3. Trajectories of player 1 generated by two control laws

CHAPTER 7

Individual and Mass Behaviour in Large Population Wireless Systems: Centralized and Nash Equilibrium Solutions

7.1. Introduction

In Chapters 2, 3 and 5, we have analyzed power control for lognormal fading channels by a stochastic control approach. This leads to determining the control law of the users by dynamic programming equations, i.e., HJB equations. To obtain implementable control laws, we developed approximation techniques and numerical methods for computing various suboptimal versions of the control law. However, for systems with large populations, there exists the basic limitation of computational complexity associated with this approach. Hence it is desirable to develop new techniques for obtaining simplified yet efficient control laws.

Based on the work in the previous Chapters, in this Chapter we make an attempt at analyzing the properties of systems operating in large population conditions. The system includes the lognormal fading channel and a rate based uplink power control model associated with each user. Our interest is in investigating the feasibility of localized or decentralized control under fading channels since this would potentially reduce the system complexity for practical implementation of the control laws. We analyze the effect of large population sizes on the controller structure. As a first step, we examine the structure of the optimal control law. The feedback control is affine in the system power with a random gain matrix (called the Riccati matrix) which carries the channel information. Under the assumptions of i.i.d. channel dynamics and equal Quality of Service (QoS) requirements, it turns out that the Riccati matrix exhibits a certain symmetry; furthermore, the power adjustment rate for any given individual mobile is determined by its own channel state, power level and an average of the impact of all other mobiles. Intuitively, when the population size is big, the network interference should exhibit a statistically stable behaviour whereby the impact of a single specific mobile becomes negligible. Based on the above facts, it is possible to develop a system configuration for network optimization which is less complex than the full state system.

In reality, for a system as complex as a large-scale mobile communication network, a centralized optimization approach may face fundamental limitation in implementation since it generally requires efficient coordination and huge amount of information exchange between different parts of the system. Hence, in contrast to the highly complex centralized optimal control, in the next step we consider simplified but efficient control design utilizing new optimization criteria. For the control determination of a fixed individual user, we group the effect of all other users into a single term and consider its approximation. This is reasonable due to the particular structure of the cost function reflecting the QoS measurement. By this means we can capture the interaction between the behaviour of any single user and the statistical behaviour of the overall system.

Subsequently we introduce the individual cost based optimization approach to the power control problem and give a game theoretic formulation. Concerning game theoretic approach for power control of lognormal fading channels, some initial investigation was presented in [**34, 37**]. In practical systems, it is important to implement control strategy in a decentralized manner, i.e., each mobile user adjusts its power based on its local information concerning the network. This can significantly reduce information exchange efforts among users and base stations and thus reduce system running costs. And based on these aspects, it makes sense to place emphasis on decentralized games. The interested reader is referred to [**58, 77, 70, 21**] and references therein for the game theoretic approach to rate allocation, power control, and other network service allocation for various static models on wired or wireless networks.

The method developed for the LQG problem in Chapter 6, combined with some reasonable hypotheses, enables us in the power control problem to get an approximation for the collective effect of all the other individuals on a given individual mobile. The procedure has connections with the single user based control design in Section 5.5, Chapter 5 (also see [**40, 41**]) where we appropriately scaled the total interference generated by all the other mobiles and treated this scaled quantity as a slowly time-varying process. In this Chapter, a particular form of the loss function is used which leads to a separation of the control law into a sum of two terms where the first term involves the given individual's channel and power state, and the second is a function of the its channel attenuation and time. Here the time dependence of the second term reflects the average effect of all other individuals, particularly during the transient phase of the power adjustment. In this framework, due to the specific decentralized information structure for individual's power adjustment, we may feel free to call the resulting control by distributed control.

We emphasize that the above state aggregation technique leads to highly localized control configurations in contrast to the full state based optimal control. Specifically, the control of of a particular *individual* mobile can be formulated in terms of its own channel dynamics, its own state, the aggregated system dynamics and the average of the interference the mobile receives from a *mass* or *collective* representing all other users.

7.2. The Problem Statement

In this Section we reformulate the stochastic power control problem in the large population context. Let $x_i(t), 1 \leq i \leq n$, denote the attenuation (expressed in dBs and scaled to the natural logarithm basis) at the instant t of the power of the i-th mobile of a network and let $\alpha_i(t) = e^{x_i(t)}$ denote the actual attenuation. The power attenuation dynamics of n mobile users are given by

$$dx_i = -a_i(x_i + b_i)dt + \sigma_i dw_i, \quad 1 \le i \le n, \quad t \ge 0, \tag{7.1}$$

where $\{w_i, 1 \le i \le n\}$ are n independent standard Wiener processes, and the initial states $x_i(0)$, $1 \le i \le n$ are mutually independent Gaussian random variables which are also independent of the Wiener processes. In (7.1) $a_i > 0$, $b_i > 0$, $\sigma_i > 0$, $1 \le i \le n$.

As in Chapter 5, We model the step-wise adjustments **[62]** of the transmitted power p_i (i.e., the uplink power control for the i-th mobile) by the so-called rate adjustment model

$$dp_i = u_i dt, \qquad 1 \le i \le n. \tag{7.2}$$

We write $x = [x_1, \cdots, x_n]^\tau$, $p = [p_1, \cdots, p_n]^\tau$, $u = [u_1, \cdots, u_n]^\tau$.

In a CDMA context, the signal to interference ratio (SIR) for the users achieved after matched filtering is given by

$$\Gamma_i = \frac{\widehat{p}_i}{\sum_{k \ne i} \beta_{k,i} \widehat{p}_k + \eta}, \qquad 1 \le i \le n, \tag{7.3}$$

where $\widehat{p}_i$ denotes the received power at the based station for user i, $\beta_{k,i} = (s_k^\tau s_i)^2$, $k \ne i$, is the crosscorrelation between the (normalized) signature sequences s_k, s_i of users k, i, respectively, and η is the constant background noise intensity. We denote the dimension (i.e., the spreading gain) of s_i by n_s. In the uplink, these signature sequences are assumed being not strictly orthogonal to each other.

Following **[72, 74, 80]**, we consider the mobile system in the context of a large number of users and make the standard assumption that $\frac{n}{n_s} \to \alpha > 0$ as $n \to \infty$, i.e., the signature length n_s increases in proportion to the system population, which is necessary in order to suppress the inter-user interference (i.e., reduce the crosscorrelation) such that the system can accommodate an increasing number of users. Here α is called the number of users per degree of freedom. By appropriately choosing random signature sequences of length n_s, one can have $\beta_{k,i} \approx \frac{1}{n_s}$ **[72, 80]**, and hence $\beta_{k,i} \approx \frac{\alpha}{n}$. For simplicity, here we take $\beta_{k,i} = \frac{1}{n}$ for all $1 \le k \ne i \le n$. Moreover, we wish Γ_i to be staying around a target SIR level $\gamma_i \in (0,1)$, i.e.,

$$\Gamma_i = \frac{\widehat{p}_i}{\frac{1}{n}\sum_{k \ne i} \widehat{p}_k + \eta} \approx \gamma_i, \qquad 1 \le i \le n. \tag{7.4}$$

under the condition of lognormal fading we have $\widehat{p}_i = e^{x_i} p_i, 1 \le i \le n$, where the power attenuation x_i is described by (7.1).

Following Chapter 5 and taking into account the SIR requirement (7.4), we introduce the following modified loss function:

$$E\int_0^\infty e^{-\rho t}\{\sum_{i=1}^n [e^{x_i} p_i - \gamma_i(\frac{1}{n}\sum_{k \ne i}^n e^{x_k} p_k + \eta)]^2 + u^\tau R u\}dt. \tag{7.5}$$

where $\rho > 0$ is the discount factor and R is a positive definite weight matrix, and $\eta > 0$ is the constant system background noise intensity. For simplicity we take a diagonal weight matrix $R = \text{Diag}(r_i)_{i=1}^n > 0$. In the above integral, the first term is based on the SIR requirements (7.4) and the second term is added to penalize abrupt change of powers since in practical systems there are basic limits for power adjustment rate. In practice, avoiding rapid change of power levels has more to do with caution in an environment where channel characteristics are estimated and are possibly time-varying. After subtracting the constant component from the integrand in (7.5) we get the cost function to be employed:

$$J(u) = E\int_0^\infty e^{-\rho t}[p^\tau C(x) p + 2D^\tau(x) p + u^\tau R u]dt, \tag{7.6}$$

where $C(x)$, $D(x)$ are $n \times n$ positive definite matrix, $n \times 1$ vector, respectively, which are determined from (7.5).

To facilitate further analysis, we set $f_i(x) = -a_i(x_i + b_i), 1 \le i \le n$, $H = \text{Diag}\ (\sigma_i)_{i=1}^n$ and $z^\tau = (x^\tau, p^\tau)$, $\psi^\tau = (f^\tau, u^\tau)$, $G^\tau = (H, 0_{n\times n})$. We write (7.1) and (7.2) in the vector form

$$dz = \psi dt + G dw, \qquad t \ge 0, \tag{7.7}$$

We take the admissible control set

$$\mathcal{U} = \{u|u \text{ is adapted to } \sigma(x_s, p_s, s \leq t), \text{ and } E\int_0^\infty e^{-\rho t}|u_t|^2 dt < \infty\}.$$

Assume that p has a deterministic initial value $p(0)$ at $s = 0$; then clearly $\sigma(x_s, p_s, s \leq t) = \sigma(x_0, w_s, s \leq t)$. Let $\Phi(x, p, u) = p^\tau C(x)p + 2D^\tau(x)p + u^\tau Ru$. The cost associated with (7.7) and a control u is $J(x, p, u) = E[\int_0^\infty e^{-\rho t}\Phi(x_t, p_t, u_t)dt | x_{t=0} = x, p_{t=0} = p]$, where (x, p) is taken as the initial state; further we set the value function $v(x, p) = \inf_{u\in\mathcal{U}} J(x, p, u)$.

7.3. The Value Function and HJB Equation

In this Section we restate some of the results in Section 5.3, Chapter 5 in the current large population context. We write the HJB equation for the value function v as follows:

$$\begin{aligned}\rho v &= f^\tau \frac{\partial v}{\partial x} + \frac{1}{2}\mathrm{tr}(\frac{\partial^2 v}{\partial z^2} GG^\tau) + \inf_{u\in\mathrm{R}^n}\{u^\tau \frac{\partial v}{\partial p} + u^\tau Ru\} + p^\tau C(x)p + 2D^\tau(x)p, \\ &= -\sum_{i=1}^n a_i(x_i + b_i)\frac{\partial v}{\partial x_i} + \frac{1}{2}\sum_{i=1}^n \sigma_i^2 \frac{\partial^2 v}{\partial x_i{}^2} - \frac{1}{4}v_p^\tau R^{-1} v_p + p^\tau C(x)p + 2D^\tau(x)p. \quad (7.8)\end{aligned}$$

Proposition 7.1. The value function v is a classical solution to the HJB equation (7.8) and can be written as

$$v(x, p) = p^\tau K(x)p + 2p^\tau S(x) + q(x) \quad (7.9)$$

where $K(x) = K^\tau(x)$, $S(x)$, $q(x)$ are continuous in x, and are all of order $O(1 + \sum_{i=1}^n e^{2x_i})$. □

We note that elliptic HJB equations such as (7.8) may admit multiple classical solutions when there is no boundary condition. In general, additional growth conditions are required in order to determine the value function by the HJB equation. See [**24**] for a general discussion.

Substituting (7.9) into the HJB equation (7.8) and comparing powers of p, we obtain the partial differential equation system

$$\rho K = \frac{1}{2}\sum_{k=1}^{n}\sigma_k^2\frac{\partial^2 K}{\partial {x_k}^2} + \sum_{k=1}^{n} f_k\frac{\partial K}{\partial x_k} - KR^{-1}K + C, \tag{7.10}$$

$$\rho S = \frac{1}{2}\sum_{k=1}^{n}\sigma_k^2\frac{\partial^2 S}{\partial {x_k}^2} + \sum_{k=1}^{n} f_k\frac{\partial S}{\partial x_k} - KR^{-1}S + D, \tag{7.11}$$

$$\rho q = \frac{1}{2}\sum_{k=1}^{n}\sigma_k^2\frac{\partial^2 q}{\partial {x_k}^2} + f^\tau\frac{\partial q}{\partial x} - S^\tau R^{-1}S, \tag{7.12}$$

where we shall refer to (7.10) as the Riccati equation of the system. Finally the optimal control law for the n users is given by

$$u = [u_1, \cdots, u_n]^\tau = -R^{-1}[K(x)p + S(x)], \tag{7.13}$$

and for user i the control is

$$u_i = -\frac{1}{r_i}K_{ii}(x)p_i - \frac{1}{r_i}\sum_{k\neq i}^{n} K_{ik}(x)p_k - \frac{1}{r_i}S_i(x). \tag{7.14}$$

It is seen from (7.14) that for user i, the control mainly relies on its own current power level and a weighted sum of other users' powers. Since all the coefficients involved in this individual control law depend on the attenuations of all users, this optimal control law in highly centralized. A practical implementation of the optimal control law systems with large populations is unfeasible due to its high complexity concerning channel conditions. To simplify our analysis, we make the following assumptions:

(H7.1) All users have i.i.d. dynamics, i.e., $a_i = a$, $b_i = b$, $\sigma_i = \sigma$, $1 \leq i \leq n$. □

(H7.2) All users have equal QoS requirements, i.e., $\gamma_i = \gamma$, $1 \leq i \leq n$, and in addition, $R = rI_n$. □

To analyze the control law in a large population situation, we first consider the case of a static channel, i.e., $\sigma_i = \sigma = 0$, $a_i = a = 0$ for all i, and assume $0 < \gamma < 1$; denote the corresponding constant solution to (7.10) by K^0. Using the method of Section 6.3, it can be verified that $K^0_{ii} = O(1)$, as $n \to \infty$, and $K^0_{ik} = O(\frac{1}{n})$, for $i \neq k$, as $n \to \infty$. It is an interesting issue to estimate the magnitude of $K_{ii}(x)$ and $K_{ik}(x)$, $i \neq k$, in the general case with $a > 0, \sigma > 0$.

If the magnitude of all $K_{ik}(x), i \neq k$, is significantly smaller than that of $K_{ii}(x)$ in a certain sense, the randomness associated with the second term in (7.14) should be small due to the scaling effect of $K_{ik}(x), i \neq k$, and hence the actual interference from all the other users to a given user is in the form of an averaged effect.

7.4. Game Theoretic Approach and State Aggregation

In this Section we assume that **(H7.1)-(H7.2)** hold. The notation used in this Section is consistent with that in Sections 7.2-7.3, and some notion of Sections 6.2-6.4 will be extended to the power control context. The ε-Nash equilibrium can also be defined here in an obvious way.

We will generalize the method of Chapter 6 to the current nonlinear case by a heuristic argument. Specifically, under certain assumptions we approximate the power control problem for large population systems by a tracking problem with an exogenous random process associated with each player. We set the individual cost for the i-th player with respect to the mass as

$$J_i(u_i, \gamma(\frac{1}{n}\sum_{k\neq i}^{n} e^{x_k}p_k + \eta)) \triangleq \int_0^\infty e^{-\rho t}\{[e^{x_i}p_i - \gamma(\frac{1}{n}\sum_{k\neq i}^{n} e^{x_k}p_k + \eta)]^2 + ru_i^2\}dt, \quad (7.15)$$

i.e., the i-th component in the centralized cost function (7.5) in Section 7.2. We also define the i-th individual cost with respect to a deterministic process z^* as

$$J_i(u_i, z^*) = \int_0^\infty e^{-\rho t}\{[e^{x_i}p_i - z^*(t)]^2 + ru_i^2\}dt, \quad (7.16)$$

where $z^*(t) \in C_b[0,\infty) \triangleq \{x | x \in C[0,\infty), \text{ and } \sup_{t\in[0,\infty)} |x(t)| < \infty\}$. When the individual cost $J_i(u_i, z^*)$ is applied, assuming sufficient differentiability of the optimal cost function we can write the equation system

$$\rho K(x_i) = \frac{\sigma^2}{2}\frac{\partial^2 K}{\partial x_i^2} - a(x_i + b)\frac{\partial K}{\partial x_i} - \frac{1}{r}K^2 + e^{2x_i}, \tag{7.17}$$

$$\rho s(t, x_i) = \frac{ds}{dt} + \frac{\sigma^2}{2}\frac{\partial^2 s}{\partial x_i^2} - a(x_i + b)\frac{\partial s}{\partial x_i} - \frac{1}{r}Ks - z^* e^{x_i}. \tag{7.18}$$

We assume $K(x_i) = O(1 + e^{2x_i})$, and $s(t, x_i) = O(1 + e^{x_i})$ uniformly with respect to t. Here $K(x_i)$ is a function of a single variable in contrast to the centralized optimal control case. By an argument using the verification theorem one can show that the control law minimizing (7.16) for the i-th user is determined as

$$u_i = -\frac{1}{r}[K(x_i)p_i + s(t, x_i)], \tag{7.19}$$

and hence we have the closed-loop equation for p_i in the form

$$dp_i = u_i dt = -\frac{1}{r}[K(x_i)p_i + s(t, x_i)]dt. \tag{7.20}$$

As in the linear quadratic case analyzed in Section 6.4, here we also have the issue of determining the function z^* which is to be tracked by individual players. With the original SIR based cost function (7.15) in mind, we consider taking

$$z^* \approx \gamma(\frac{1}{n}\sum_{k\neq i}^{n} e^{x_k}p_k + \eta), \tag{7.21}$$

for large n. To further simplify our analysis, in addition to independence between any pair of processes $x_i, x_k, i \neq k$, we assume that each x_i has initial condition $x_i|_{t=0}$ such that x_i is a stationary Gaussian process. We also assume that powers of all mobile users have identical deterministic initial conditions p_0. The generalization to more general initial conditions for the attenuations and powers will present no technical difficulty. For large n, the scaled sum in (7.21) may be approximated by the the mean of a single term under mild conditions for $p_i(t)$, $i \geq 1$. Thus we write

$$z^*(t) = \gamma(Ee^{x_i}p_i + \eta), \tag{7.22}$$

where the right hand side depends only on time t and the initial power p_0 after the feedback is determined by (7.19) for all individuals. We make the Hypothesis:

(H7.3) The equation system (7.17), (7.18) and (7.22) has a solution $(K(x_i), s(x_i), z^*(t))$ where $z^* \in C_b[0, \infty)$, $K \in C^2(R)$, $S \in C^{1,2}(R_+ \times R)$; in addition $K(x_i) = O(1 + e^{2x_i})$, and $s(t, x_i) = O(1 + e^{x_i})$ uniformly with respect to t. □

Proposition 7.2. Under **(H7.3)**, the control law determined by (7.19) is an ε-Nash equilibrium for the costs (7.15) subject to full information for individual controls, where $\varepsilon = O(\frac{1}{\sqrt{n}})$.

PROOF. The proof is similar to that of Theorem 6.3 and the details are omitted here. □

It is of significant interest to study the dynamic behaviour of z^*. A possible approach is to introduce a controlled Fokker-Planck equation for the joint distribution or density of (x_i, p_i) and then describe z^* in terms of the Fokker-Planck equation. The challenging issue of existence of a solution to the resulting equation system will be investigated in future work.

7.5. Concluding Remarks

In this Chapter we have investigated stochastic power control subject to lognormal fading in a large population context. Two different methods are considered: the global cost based centralized information control and the individual cost based decentralized control.

In general, the global cost based approach emphasizes a certain coordination between individuals to achieve global optimality; in this approach for large population systems, assuming that the feedback gain satisfies a certain condition on its magnitude, the information used by a given individual exhibits a certain separation in that its control law mainly depends its own channel-power condition and another quantity reflecting the average effect of the collective of other users which is close to a deterministic process (for a large population). It should be noted that in this centralized framework, each individual does not make direct efforts to optimize against this roughly deterministic process, which differs from the dynamic game theoretic scenario.

On the other hand, noticing the scaling nature involved in the cost function, we consider approximation and splitting of the global cost function which naturally induces individual costs. This leads to a game theoretic framework. In such an individual cost based optimization framework, there is also a roughly deterministic process generated by the mass or collective. In contrast to the global cost case, here each individual determines its control law by optimizing against the mass. Thus there is an intrinsic clash of interest between different users. But individual and the mass can still reach a certain stable behaviour under certain conditions.

CHAPTER 8

Future Research

Suggested Research on Adaptation with Channel Dynamics

- In the current stochastic control framework for the power adjustment, it is assumed that the dynamics for lognormal fading are known. In a more realistic setup, one may assume that the parameters of the fading channel are unknown and consider adaptive implementation of the control. In practice when the channel attenuation is measured, for instance, by means of pilot signals, one can identify the parameters of the fading channel model by well established identification algorithms for linear stochastic models as shown in Chapter 5. Then the estimated parameters may be combined with the stochastic control approach to give adaptive versions of the control laws.
- An even more challenging issue is to develop a stochastic adaptive control scheme for power adjustment by assuming that only indirect measurements for the channel state are available.

Relaxation of the Dynamics Assumption in the LQG Game

- In the large-scale LQG game of Chapter 6, all agents essentially have the same dynamics which is a strong assumption. A possible generalization is to consider randomized coefficients in the dynamic which can be described by a certain distribution.

Suggested Research on Modelling Mass Behaviour

- To solve the large population power control problem in Chapter 7, a crucial step is to develop an efficient modelling methodology for the mass behaviour. For any given agent in the system, it is useful to further investigate the evolution of the joint distribution of its own state and the mass subject to any fixed control law.

Indiscipline of Sub-Population

- We have solved the large-scale LQG problem in the noncooperative game theoretic context where a state aggregation technique is applied to construct ε-Nash equilibria. In this setup, each agent has the task to determine the behaviour of other individuals and estimate the mass influence it may receives. Thus the feasibility of the resulting localized strategy relies heavily on certain universal rationality of the population.
- In further generalization of the state aggregation technique it is appealing to consider tolerating misbehaviour of a sub-population. In reality, it is possibly for some agents to take irrational actions due to their own way of reasoning or because of receiving unreliable information from the system. We term this situation as indiscipline of sub-population.
- Important issues concerning indiscipline of sub-population include what is a tolerable size of this sub-population with misbehaviour and to what extent they are allowed to act on their own will. The two aspects may be related to each other.

Systems with Varying Populations

- In the methodology proposed for power control in this thesis, the population size is assumed to be constant. In a real system, over time new users will join the user population while others may leave upon completion of their service. Taken into account this fact, we may model the population variation by a birth-death process, i.e., the population is modelled as a jumping Markov process. The main issue then would be to design localized control configuration allowing population variation.

REFERENCES

[1] T. Alpcan, T. Basar, R. Srikant, and E. Altman. CDMA uplink power control as a noncooperative game. *Proc. 40th IEEE Conf. Decision and Contr.*, Orlando, Florida, pp.197-202, Dec. 2001.

[2] M-S. Alouini and M.K. Simon. Dual diversity over correlated lognormal fading channels. *IEEE Trans. Communications*, vol.50, no.12, pp.1946-1959, 2002.

[3] W.F. Ames. *Numerical Methods for Partial Differential Equations*, 3rd Ed., Acad. Press, New York, 1992.

[4] M. Aoki. Stochastic control in economic theory and economic systems. *IEEE Trans. Automat. Control*, vol.21, no.2, pp.213-220, 1976.

[5] K.J. Astrom. *Introduction to Stochastic Control Theory*, Academic Press, New York, 1970.

[6] J.-P. Aubin. *Optima and Equilibria*, Springer, 2nd Ed., 1998.

[7] G. Barles and P.E. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *J. Asymptotic Analysis.* 4, pp.271-283, 1991.

[8] Y. Bar-Shalom, R. Larson, and M. Grossberg. Application of stochastic control theory to resource allocation under uncertainty. *IEEE Trans. Automat. Control*, vol.19, no.1, pp.1-7, 1974.

[9] T. Basar and Y.-C. Ho. Informational properties of the Nash solutions of two stochastic nonzero-sum games. *J. Econ. Theory*, vol.7, pp.370-387, 1974.

[10] A. Bensoussan. *Stochastic Control of Partially Observable Systems*, Cambridge University Press, 1992.

[11] D.P. Bertsekas. *Dynamic Programming and Optimal Control, Vol. I, II*, Athena Scientific, Belmont, Mass., 1995.

[12] F. Black and M. Scholes. The pricing of options and corporate liabilities. *J. Political Economy*, vol.81, no.3, pp.637-654, 1973.

[13] R. Buche and H.J. Kushner. Control of mobile communications with time-varying channels in heavy traffic. *IEEE Trans. on Automatic Control*, vol.47, no.6, 2002.

[14] P.E. Caines. *Linear Stochastic Systems*, Wiley, New York, 1988.

[15] P.E. Caines. Continuous time stochastic adaptive control: non-explosion, ε-consistency and stability. *Systems and Control Letters*, vol.19, no.3, pp.169-176, 1992.

[16] J.-F. Chamberland and V.V. Veeravalli. Decentralized dynamic power control for cellular CDMA systems. *IEEE Trans. Wireless Commun.*, vol.2, no.3, pp. 549-559, 2003.

[17] C.D. Charalambous and N. Menemenlis. Stochastic models for long-term multipath fading channels. *Proceedings of the 38th IEEE Conference on Decision and Control*, Phoenix, Arizona, pp.4947-4952, December 1999.

[18] P. Chaudhury, W. Mohr, and S. Onoe. The 3GPP proposal for IMT-2000. *IEEE Commun. Mag.*, pp.72-81, Dec. 1999.

[19] A.J. Coulson, G. Williamson, and R.G. Vaughan. A statistical basis for lognormal shadowing effects in multipath fading channels. *IEEE Trans. on Commun.*, vol.46, no.4, pp.494-502, 1998.

[20] M.G. Crandall, H. Ishii, and P.L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc.*, 27, pp.1-67, 1992.

[21] Z. Dziong and L.G. Mason. Fair-efficient call admission control policies for broadband networks — a game theoretic framework. *IEEE/ACM Trans. Networking*, vol.4, no.1, pp.123-136, 1996.

[22] J. Evans and D.N.C. Tse. Large system performance of linear multiuser receivers in multipath fading channels. *IEEE Trans. Information Theory*, vol.46, no.6, pp.2059-2077, 2000.

[23] W.H. Fleming and R.W. Rishel. *Deterministic and Stochastic Optimal Control*, Springer-Verlag, 1975.

[24] W.H. Fleming and H.M. Soner. *Controlled Markov Processes and Viscosity Solutions*, Springer-Verlag, 1993.

[25] A. Friedman. *Partial Differential Equations of Parabolic Type*, Prentice-Hall Inc., NJ, 1964.

[26] R. Ganesh and K. Pahlavan. Effects of traffic and local movements on multipath characteristics of an indoor radio channel. *Electronics Letters*, vol.26, no.12, pp.810-812, 1990.

[27] R. Ganesh and K. Pahlavan. Statistics of short time and spatial variations measured in wideband indoor radio channels. *IEE Proceedings-H*, vol.140, no.4, pp.297-302, 1993.

[28] M. Gudmundson. Correlation model for shadow fading in mobile radio systems. *Electronics Letters*, vol.27, no.23, pp.2145-2146, 1991.

[29] M.O. Hasna and M.-S. Alouini. Optimum power allocation for soft handoff algorithms over lognormal shadowing channels. *IEEE Internat. Commun. Conf.*, New York, NY, pp.3207-3211, April, 2002.

[30] M.L. Honig and H.V. Poor. *Adaptive interference suppression*, in *Wireless Communications: Signal Processing Perspective*, H.V. Poor and G.W. Wornell, Eds. Englewood Cliffs, NJ: Prentice-Hall, pp.64-128, 1998.

[31] M. Huang, P.E. Caines, C.D. Charalambous, and R.P. Malhamé. Power control in wireless systems: a stochastic control formulation. *Proc. of Amer. Contr. Confer.*, Arlington, Virginia, pp.750-755, June, 2001.

[32] M. Huang, P.E. Caines, C.D. Charalambous, and R.P. Malhamé. Stochastic power control for wireless systems: classical and viscosity solutions. *Proc. 40th IEEE Conf. Decision and Control*, Orlando, Florida, pp.1037-1042, December, 2001.

[33] M. Huang, P.E. Caines, and R.P. Malhamé. On a class of singular stochastic control problems arising in communications and their viscosity solutions. *Proc. of the 40th IEEE Conf. Dec. Contr.*, Orlando, Florida, pp.1031-1037, Dec., 2001.

[34] M. Huang, P.E. Caines, and R.P. Malhamé. Stochastic power control for wireless systems: centralized dynamic solutions and aspects of decentralized control. *Proc. 15th IFAC World Congress on Automatic Control*, Barcelona, Spain, CDROM, July, 2002.

[35] M. Huang, P.E. Caines, and R.P. Malhamé. Uplink power adjustment in wireless communication systems: a stochastic control analysis. *Under Revision for IEEE Trans. Automat. Control.* First submission to IEEE in September, 2002.

[36] M. Huang, P.E. Caines, and R.P. Malhamé. Degenerate stochastic control problems with exponential costs and weakly coupled dynamics: viscosity solutions and a maximum principle. *Under Revision for SIAM J. Contr. Optim.* First submission to SIAM in Oct. 2002.

[37] M. Huang, P.E. Caines, and R.P. Malhamé. Individual and mass behaviour in large population stochastic wireless power control problems: centralized and Nash equilibrium solutions. *Submitted to IEEE Conf. Decision and Control*, Mar. 2003.

[38] M. Huang, R.P. Malhamé, and P.E. Caines. Quality of service control for wireless systems: minimum power and minimum energy solutions. *Proceedings*

of the American Control Conference, Anchorage, Alaska, pp.2424-2429, May, 2002.

[39] M. Huang, R.P. Malhamé, and P.E. Caines. Stochastic power control in wireless communication systems with an infinite horizon discounted cost. *Proc. of the American Control Conference*, Denver, Colorado, pp.963-968, June, 2003.

[40] M. Huang, R.P. Malhamé, and P.E. Caines. Stochastic power control in wireless systems and state aggregation. *Under Revision for IEEE Trans. Automat. Control.* First submission to IEEE in January, 2003.

[41] M. Huang, R.P. Malhamé, and P.E. Caines. Stochastic power control in wireless communication systems: analysis, approximate control algorithms and state aggregation. *Submitted to IEEE Conf. Decision and Control*, Jan. 2003.

[42] H. Ishii. On uniqueness and existence of viscosity solutions of fully nonlinear second-order elliptic PDE's. *Communications on Pure and Applied Mathematics*, vol.42, pp.15-45, 1989.

[43] W.C. Jakes. *Microwave Mobile Communications*, Wiley, New York, 1974.

[44] R. Jensen. The maximum principle for viscosity solutions of second order fully nonlinear partial differential equations. *Archive for Rational Mechanics and Analysis*, vol.101, pp.1-27, 1988.

[45] R. Jensen, P.L. Lions, and P.E. Souganidis. A uniqueness result for viscosity solutions of second order fully nonlinear partial differential equations. *Proceedings of the American Mathematical Society*, vol.102, no.4, pp.975-978, April, 1988.

[46] S. Kandukuri and S. Boyd. Optimal power control in interference-limited fading wireless channels with outage-probability specifications. *IEEE Trans. Wireless Commun.*, vol.1, no.1, pp.46-55, 2002.

[47] I. Karatzas. *Lectures on the Mathematics of Finance*, AMS, Providence, RI, 1997.

[48] I. Karatzas and S.E. Shreve. *Methods of Mathematical Finance*, Springer-Verlag, New York, 1998.

[49] O.E. Kelly, J. Lai, N.B. Mandayam, A.T. Ogielski, J. Panchal, R.D. Yates. Scalable parallel simulations of wireless networks with WIPPEP: modelling of radio propagation, mobility and protocols. *Mobile Networks and Applications*, no.5, pp.199-208, 2000.

[50] N.V. Krylov. Approximating value functions for controlled degenerate diffusion processes by using piece-wise constant policies, Electronic J. of Probability, 4(2), pp.1-19, 1999. http://www.math.washington.edu/ejpecp/EjpVol4/paper2.abs.html.

[51] N.V. Krylov. On the rate of convergence of finite-difference approximations for Bellman's equations with variable coefficients. *Probab. Theory Relat. Fields*, 117, pp.1-16, 2000.

[52] H.J. Kushner. *Introduction to stochastic control*, New York: Holt, Rinehart & Winston, 1971.

[53] H.J. Kushner and A.J. Kleinman. Numerical methods for the solution of the degenerated nonlinear elliptic equations arising in optimal stochastic control theory. *IEEE Trans. Automat. Contr.*, vol.13, no.4, pp.344-353, 1968.

[54] T.L. Lai and C.Z. Wei. Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems. *Ann. Statist.*, vol.10, no.1, pp.154-166, 1982.

[55] R.S. Liptser and A.N. Shiryayev. *Statistics of Random Processes, Vol.I*, Springer-Verlag, 1977.

[56] N.B. Mandayam, P.C. Chen, and J.M. Holtzman. Minimum duration outage for cellular systems: a level crossing analysis. *Proc. 46th IEEE Conference on Vehicular Technology*, Atlanta, Georgia, pp.879-883, April, 1996.

[57] R.C. Merton. Lifetime portfolio selection under uncertainty: the continuous-time case. *Rev. Econom. Statist.*, 51, pp.247-257, 1969.

[58] A. Orda and N. Shimkin. Incentive pricing in multiclass systems. *Telecomm. Syst.*, 13, pp.241-267, 2000.

[59] J.D. Parsons. *The Mobile Radio Propagation Channel*, 2nd, Ed., John Wiley & Sons, Inc., 2000.

[60] W.G. Phoel and M.L. Honig. Performance of coded DS-CDMA with pilot-assisted channel estimation and linear interference suppression. *IEEE Trans. Commun.*, vol.50, no.50, pp.822-832, 2002.

[61] K. Pahlavan and A.H. Levesque. *Wireless Information Networks*, Wiley-Interscience, New York, 1995.

[62] QUALCOMM Inc. *An overview of the application of Code Division Multiple Access (CDMA) to digital cellular systems and personal cellular networks.* (Document no. EX60-10010), 1992.

[63] T.S. Rappaport. *Wireless Communications: Principles and Practice*, 2nd Ed., Prentice Hall, N.J., 2002.

[64] W. Rudin. *Real and Complex Analysis*, 3rd Ed., McGraw-Hill, NY, 1987.

[65] S.P. Sethi and Q. Zhang. *Hierarchical Decision Making in Stochastic Manufacturing Systems*, Boston, Birkhauser, 1994.

[66] H.M. Soner. Optimal control with state-space constraint I, II. *SIAM Journal on Control and Optimization*, vol.24, pp.552-562, pp.1110-1121, 1986.

[67] L. Song, N.B. Mandayam, and Z. Gajic. Analysis of an up/down power control algorithm for the CDMA reverse link under fading. *IEEE J. Select. Areas Commun.*, vol.19, no.2, pp.277-286, 2001.

[68] C.W. Sung and W.S. Wong. A distributed fixed-step power control algorithm with quantization and active link quality protection. *IEEE Transactions on Vehicular Technology*, vol.48, no.2, pp.553-562, 1999.

[69] C.W. Sung and W.S. Wong. The convergence of an asynchronous cooperative algorithm for distributed power control in cellular systems. *IEEE Transactions on Vehicular Technology*, vol.48, no.2, pp.563-570, 1999.

[70] C.W. Sung and W.S. Wong. Mathematical aspects of the power control problem in mobile communication systems. In *Lectures on systems, control and information: Lectures at the Morningside Center of Mathematics*, L. Guo and S.S.-T. Yau Eds. Providence, RI: AMS/IP Studies in Advanced Mathematics, vol.17, 2000.

[71] TIA/EIA/IS-95-A. *Mobile station-base station compatibility standard for dual-mode wideband spread spectrum cellular system.* Tech. Report, Telecommunications Industry Association, 1995.

[72] D.N.C. Tse and S.V. Hanly. Linear multiuser receivers: effective interference, effective bandwidth and user capacity. *IEEE Trans. Inform. Theory*, vol.45, no.2, pp.641-657, 1999.

[73] S. Ulukus and R.D. Yates. Stochastic power control for cellular radio systems. *IEEE Trans. Commun.*, vol.46, no.6, pp.784-798, 1998.

[74] S. Verdú and S. Shamai. Spectral efficiency of CDMA with random spreading. *IEEE Trans. Inform. Theory*, vol.45, no.2, pp.622-640, 1999.

[75] H. Viswanathan. Capacity of Markov channels with receiver CSI and delayed feedback. *IEEE Trans. Inform. Theory*. vol.45, no.2, pp.761-771, 1999.

[76] A.M. Viterbi and A.J. Viterbi. Erlang capacity of a power-controlled CDMA system. *IEEE Selected Areas in Communications*, pp.892-900, August, 1993.

[77] H. Yaiche, R.R. Mazumdar, and C. Rosenberg. A game theoretic framework for bandwidth allocation and pricing in broadband networks. *IEEE/ACM Trans. Network.*, vol.8, pp.667-678, 2000.

[78] J. Yong and X.Y. Zhou. *Stochastic Controls: Hamiltonian Systems and HJB Equations*, Springer, 1999.

[79] K. Yosida. *Functional Analysis*, 6th Ed., Springer-Verlag, 1980.

[80] J. Zhang and E.K.P. Chong. CDMA systems in fading channels: admissibility, network capacity and power control. *IEEE Trans. Inform. Theory*, vol.46, no.3, pp.962-981, 2000.

[81] J. Zhang, E.K.P. Chong, and I. Kontoyianis. Unified spatial diversity combining and power allocation for CDMA systems in multiple time-scale fading channels. *IEEE J. Selected Areas in Communications*, vol.19, no.7, pp.1276-1288, 2001.

[82] H. Zhang, W.S. Wong, W. Ge, and P.E. Caines. A stochastic approximation approach to the robust power control problem, preprint, 2003.

Document Log:

Manuscript Version 1 — 30 May 2003
Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-LaTeX — 6 June 2003

MINYI HUANG

CENTER FOR INTELLIGENT MACHINES, MCGILL UNIVERSITY, 3480 UNIVERSITY STREET, MONTRÉAL (QUÉBEC) H3A 2A7, CANADA

E-mail address: mhuang@cim.mcgill.ca

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-LaTeX