

 **COORDINATED SCIENCE LABORATORY**

DECISION AND CONTROL LABORATORY

SELF-TUNING METHODS FOR MULTIPLE-CONTROLLER SYSTEMS

Y.M. CHAN

APPROVED FOR PUBLIC RELEASE. DISTRIBUTION UNLIMITED.

REPORT R-915

UIU-ENG 81-2246

UNIVERSITY OF ILLINOIS - URBANA, ILLINOIS

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) SELF-TUNING METHODS FOR MULTIPLE-CONTROLLER SYSTEMS		5. TYPE OF REPORT & PERIOD COVERED Technical Report
7. AUTHOR(s) YICK MAN CHAN		6. PERFORMING ORG. REPORT NUMBER R-915(DC-47); UILU-ENG-2246
9. PERFORMING ORGANIZATION NAME AND ADDRESS Coordinated Science Laboratory University of Illinois at Urbana-Champaign Urbana, Illinois 61801		8. CONTRACT OR GRANT NUMBER(s) N00014-79-C-0424 NSF-ECS-79-19396 AFOSR-78-3633
11. CONTROLLING OFFICE NAME AND ADDRESS Joint Services Electronics Program		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		12. REPORT DATE August 1981
		13. NUMBER OF PAGES 105
		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Stochastic adaptive control Stochastic adaptive games Decentralized adaptive Nash games		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) The optimization of stochastic systems with unknown parameters and multiple decision-makers or controllers each having his own objective is considered. Based on a centralized information pattern, steady-state solutions are obtained for the stochastic adaptive Nash game and Leader-Follower game problems. These adaptive solutions, after a judicious transformation, resemble closely the implicit self-tuning solution for the single-controller single-objective case, and thus preserve the salient and advantageous features of self-tuning methods--simplicity and easy implementation. In addition, due to this close resemblance.		

20. Abstract (continued)

convergence results for the game problems are established by extending the convergence result from the single-controller single-objective case. The decentralized stochastic adaptive Nash game problem is also considered. Two explicit self-tuning type algorithms are proposed. The first algorithm is an ad hoc constraint on the policy form while the second one is based on extension from static Nash game theory. Simulation results indicate all these self-tuning methods are capable of stabilizing a system along targeted paths.

SELF-TUNING METHODS FOR MULTIPLE-CONTROLLER SYSTEMS

by

Yick Man Chan

This work was supported in part by the Joint Services Electronics Program under Contract N00014-79-C-0424, in part by the National Science Foundation under Grant ECS-79-19396, and in part by the U. S. Air Force under Grant AFOSR-78-3633.

Reproduction in whole or in part is permitted for any purpose of the United States Government.

Approved for public release. Distribution unlimited.

SELF-TUNING METHODS FOR MULTIPLE-CONTROLLER SYSTEMS

BY

YICK MAN CHAN

B.S., Rose-Hulman Institute of Technology, 1977
M.S., University of Illinois, 1979

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 1981

Thesis Adviser: Professor J. B. Cruz, Jr.

Urbana, Illinois

SELF-TUNING METHODS FOR MULTIPLE-CONTROLLER SYSTEMS

Yick Man Chan, Ph.D.

Department of Electrical Engineering
University of Illinois at Urbana-Champaign, 1981

The optimization of stochastic systems with unknown parameters and multiple decision-makers or controllers each having his own objective is considered. Based on a centralized information pattern, steady-state solutions are obtained for the stochastic adaptive Nash game and Leader-Follower game problems. These adaptive solutions, after a judicious transformation, resemble closely the implicit self-tuning solution for the single-controller single-objective case, and thus preserve the salient and advantageous features of self-tuning methods—simplicity and easy implementation. In addition, due to this close resemblance, convergence results for the game problems are established by extending the convergence result from the single-controller single-objective case. The decentralized stochastic adaptive Nash game problem is also considered. Two explicit self-tuning type algorithms are proposed. The first algorithm is an ad hoc constraint on the policy form while the second one is based on extension from static Nash game theory. Simulation results indicate all these self-tuning methods are capable of stabilizing a system along targeted paths.

ACKNOWLEDGEMENT

The author would like to extend his gratitude and appreciation towards his advisor, Professor J. B. Cruz, Jr. for his invaluable guidance and support. He would also like to thank Professors W. R. Perkins, D. P. Looze, and H. V. Poor for serving on his dissertation committee.

Sincere thanks are also given to Mrs. Dee Wrather for her excellent typing.

TABLE OF CONTENTS

CHAPTER		Page
1	INTRODUCTION	1
	1.1 Motivation	1
	1.2 Self-Tuning Strategy	2
	1.3 Thesis Outline	3
2	SELF-TUNING PRINCIPLE	5
	2.1 Introduction	5
	2.2 Self-Tuning Regulator (STR)	6
	2.2.1 Minimum Variance Strategy	6
	2.2.2 Regulator for System with Unknown Parameters	8
	2.3 Self-Tuning Controller (STC)	11
	2.3.1 Controller Design with Known Parameters	12
	2.3.2 Controller for System with Unknown Parameters	17
	2.4 Convergence Analysis for Self-Tuning	19
	2.5 Simulation Example	21
3	STOCHASTIC ADAPTIVE NASH GAMES	33
	3.1 Introduction	33
	3.2 Problem Formulation	34
	3.3 Self-Tuning Nash Game	36
	3.3.1 N-Person Nash Equilibrium Strategy	36
	3.3.2 N-Person Self-Tuning Nash Equilibrium Strategy	39
	3.3.3 Convergence	41
	3.4 Simulation Example	42
4	STOCHASTIC ADAPTIVE STACKELBERG GAMES	47
	4.1 Introduction	47
	4.2 Problem Formulation	49

CHAPTER		Page
4.3	Self-Tuning Leader-Follower Game	50
4.3.1	Two-Person Leader-Follower Equilibrium Strategy	50
4.3.2	Two-Person Self-Tuning Stackelberg Strategy	53
4.3.3	Convergence	55
4.4	Simulation Study	55
5	DECENTRALIZED STOCHASTIC ADAPTIVE NASH GAMES	60
5.1	Introduction	60
5.2	Problem Formulation	61
5.3	Constrained Decentralized Nash Game	64
5.3.1	Nash Game with Constrained Policy	65
5.3.2	Self-Tuning Constrained Decentralized Nash Game	66
5.4	Extended Static Decentralized Nash Game	69
5.4.1	Extended Static Decentralized Nash Solution	71
5.4.2	Self-Tuning Extended Static Decentralized Nash Games	74
5.5	Simulation Studies and Conclusion	76
6	CONCLUSION	85
	REFERENCES	87
	APPENDIX	
A	TRANSFORMATION OF SYSTEMS	92
B	PROOF OF THEOREM 3.1	93
C	ECONOMIC MODEL	95
D	PROOF OF THEOREM 4.1	97
E	PROOF OF THEOREM 5.1	100
	VITA	105

CHAPTER 1

INTRODUCTION

1.1 Motivation

Many systems we encounter in our daily routines have these dominant features: i) unknown or partial knowledge of system dynamics; ii) presence of multiple decision-makers or controllers each of whom has his different objective; iii) presence of unmeasurable disturbances. Examples of such systems include distributed industrial systems, power and energy systems, transportation systems, environmental systems, biological systems and socio-economic systems, just to name a few. Optimization of such systems falls naturally into the framework of stochastic adaptive games.

A dynamic game is a system characterized by the presence of multiple decision-makers. The theory of games first attained its formalism due to the publication of the book "Theory of Games and Economic Behavior" by [43]. A majority of the work for game theory has been done for systems with known parameters [20, 21]. In this thesis, we propose an adaptive procedure to tackle the game problem when we have no information or just partial knowledge of the system parameters. This particular adaptive algorithm, which incorporates a minimum variance control strategy and a least squares identification scheme, is the Self-Tuning Strategy [1, 5, 45]. The reason for using the Self-Tuning Strategy in tackling the Stochastic Adaptive Game problem is primarily due to the simplicity of the algorithm and proven success in industrial applications [2, 6].

1.2 Self-Tuning Strategy

The Self-Tuning Strategy is basically a suboptimal control scheme because the design of the control signal does not take into consideration the effect of the control signal on the estimation of the system dynamics. In the design of stochastic adaptive controllers, the role of the control signal is two-fold: i) the attainment of the control objective; ii) the identification of the system parameters or dynamics [4, 8, 9, 56]. This dual nature of the control signal was first pointed out in [22]. Controllers which take into account of the dual nature of the control system are classified as dual controllers. By this definition, the Self-Tuning Strategy is a non-dual type algorithm because it approaches the estimation problem and control problem independently and assumes no interaction exists between the two problems. Even though the Self-Tuning Strategy is a non-dual adaptive control method, it has received wide attention and generated a substantial amount of results on both the theoretical level and practical applications primarily due to its simplicity and ease of implementation. The advent of microprocessor, with its falling cost and rising computing power, has allowed a prototype portable self-tuner to be constructed and tested on site for various industrial processes [19]. These self-tuners are particularly appealing under the following situations:

- i) frequent manual retuning needed for the traditional three term PID (Proportional, Integral, Derivative) control scheme;
- ii) frequent changes in set point for linearized system dynamics;
- iii) presence of noise in the system;
- iv) presence of slowly time-varying system parameters.

We hope the proven applications of the self-tuner will provide us with a practical tool for solving the stochastic adaptive game problem and ultimately enable us to implement, with ease, the theory of games to the numerous systems we encounter daily.

1.3 Thesis Outline

In this thesis, we will utilize the self-tuning method to solve the stochastic adaptive Nash game and Stackelberg game problems. Our objective is to seek steady state game solutions that can be practically implemented with ease. Indeed, by restricting the cost functions of each decision-maker to a certain class, we obtain solutions for the game problem, which resembles closely, after certain transformation, the solution of the self-tuning control problem with only one decision-maker. This close resemblance implies that the computation for the game solution can be carried out using similar methods that are used for the self-tuners. Microprocessor implementation, naturally, is a desirable goal.

Since our approach is based on the self-tuning principle, we will briefly review the various aspects of this theory in Chapter 2. We will concentrate on the original self-tuning regulator [6] and a generalized self-tuning method proposed in [16, 17]. Extension and new convergence result for the method in [16, 17] are also presented in this chapter.

In Chapters 3 and 4, we will define and formulate the stochastic adaptive Nash game and Stackelberg game problems respectively. We will assume a centralized information pattern, that is, the game problems will be solved with the assumption that every decision-maker has the same

input-output data about the system. Convergence for the game problem will be shown by extending the convergence results of the self-tuning controller with one decision-maker.

In Chapter 5, decentralized stochastic adaptive Nash games will be considered. Specifically, we will consider a "one-step-delay information sharing pattern". By restricting the cost functions for the decision-makers to single-stage, an adaptive games solution is obtained by extending the results of static games with known parameters. We also obtain similar adaptive solution by a straightforward constraint on the form of each decision-maker's control law. Simulation results using these procedures are presented.

CHAPTER 2

SELF-TUNING PRINCIPLE

2.1 Introduction

In order to control processes where there are unknown parameters and unmeasurable disturbances, the self-tuning method has been proposed to overcome these problems. In this chapter, we will review the underlying idea of self-tuning for the single decision-maker single criterion case. New convergence result and extension are also presented.

In Section 2.2, the Self-Tuning Regulator (STR) of [5] will be reviewed. The STR basically combines a minimum variance control law and a least squares estimator to deal with the unknown parameter and noisy system. A variation of the STR, the Self-Tuning Controller (STC) in [16, 17, 18, 25, 26], will be reviewed and extended in Section 2.3. Properties of these controllers are discussed.

In Section 2.4, convergence results for the STR will be presented and we will show how the convergence results for the STR can be carried over to the STC. A remark that is worth mentioning at this point is that a similar procedure will be used in obtaining convergence results for the game problems. In other words, we will show how the convergence results for the STR can be carried over to the Nash game and Stackelberg problems.

Finally, in Section 2.5, an example based on a paper making machine in [14] is simulated using the two different self-tuners.

2.2 Self-Tuning Regulator (STR)

The process to be regulated is formulated in an input-output model form. The objective of the control action is to minimize the output variances of the process. To review the STR concept, we will first present the minimum variance strategy for the system assuming complete knowledge of the system parameters. Then, the adaptive minimum variance control law to deal with unknown system parameters is presented. Further details can be found in [6, 14].

2.2.1 Minimum Variance Strategy

The process to be controlled is governed by

$$A(q^{-1})y(t) = B(q^{-1})u(t-k-1) + C(q^{-1})e(t), \quad k \geq 0 \quad (2.1)$$

where q^{-1} denotes the backward shift operator, k is time delay, y is the output vector, u is the input vector, and $\{e(t)\}$ is a sequence of independent, identically distributed random vectors with zero mean and finite covariance. The vectors y , u , and e are all the same dimension p . The polynomial matrices $A(z)$, $B(z)$, and $C(z)$ are all of dimension $p \times p$ given by

$$A(z) = I + A_1 z + \dots + A_n z^n, \quad (2.2a)$$

$$B(z) = B_0 + B_1 z + \dots + B_{n-1} z^{n-1}, \quad B_0 \text{ non-singular} \quad (2.2b)$$

$$C(z) = I + C_1 z + \dots + C_n z^n, \quad (2.2c)$$

with $\det B(z)$ and $\det C(z)$ all have their zeros strictly outside the unit circle.

The objective of the control action is to minimize, given the input-output data up to time t , with respect to $u(t)$, a cost function J given by

$$J = E\{y^T(t+k+1)Qy(t+k+1)\} \quad (2.3)$$

where E denotes the expectation operation and Q is a $p \times p$ symmetric positive semidefinite matrix. The minimum variance strategy minimizes J over all admissible controls $u(t)$, specifically, all $u(t)$ which consists of functions of all current and past outputs $y(t), y(t-1), \dots$ and past inputs $u(t-1), u(t-2), \dots$.

It can be shown that the minimum variance strategy is given by

$$\bar{G}(q^{-1})y(t) + \bar{F}(q^{-1})B(q^{-1})u(t) = 0 \quad (2.4)$$

where $\bar{F}(z)$ and $\bar{G}(z)$ satisfy

$$C(z) = A(z)F(z) + z^{k+1}G(z) \quad (2.5a)$$

$$\bar{F}(z)G(z) = \bar{G}(z)F(z) \quad (2.5b)$$

$$\det \bar{F}(z) = \det F(z), \quad \bar{F}(0) = I \quad (2.5c)$$

and $F(z), G(z)$ are polynomial matrices given by

$$F(z) = I + F_1 z + \dots + F_k z^k \quad (2.6a)$$

$$G(z) = G_0 + G_1 z + \dots + G_{n-1} z^{n-1}. \quad (2.6b)$$

Derivation of the minimum variance strategy can be found in [14].

The closed loop system with this strategy being applied becomes

$$\bar{C}(q^{-1})y(t) = \bar{C}(q^{-1})F(q^{-1})e(t) \quad (2.7)$$

where

$$\bar{F}(z)C(z) = \bar{C}(z)F(z) . \quad (2.8)$$

Since $\det \bar{F}(z) = \det F(z)$, $\det C(z) = \det \bar{C}(z)$. Thus, the closed loop system is stable as $\det C(z)$ is assumed to have all zeros strictly outside the unit circle.

The control error with this strategy is asymptotically given by

$$y(t) = F(q^{-1})e(t) \quad (2.9)$$

which is a moving average of order k of the noise $e(t)$.

2.2.2 Regulator for System with Unknown Parameters

In order to control the process given by (2.1) with unknown parameters, the following model is used for representation of the process

$$y(t) + \mathcal{A}(q^{-1})y(t-k-1) = \mathcal{B}(q^{-1})u(t-k-1) + e(t) \quad (2.10)$$

where

$$\mathcal{A}(z) = \mathcal{A}_0 + \mathcal{A}_1 z + \dots + \mathcal{A}_{n-1} z^{n-1} \quad (2.11a)$$

$$\mathcal{B}(z) = \mathcal{B}_0 + \mathcal{B}_1 z + \dots + \mathcal{B}_{n+k-1} z^{n+k-1} \quad (2.11b)$$

and $e(t)$ is the error to be minimized in the least squares sense.

The minimum variance strategy for the process (2.10) is given by

$$\mathcal{B}(q^{-1})u(t) = \mathcal{A}(q^{-1})y(t) . \quad (2.12)$$

For the STR, at each instance of time, it performs a least squares estimation for the model given by (2.10). The estimates $\hat{\mathcal{A}}(z)$ and $\hat{\mathcal{B}}(z)$ for $\mathcal{A}(z)$ and $\mathcal{B}(z)$ respectively are then substituted into (2.12) to obtain

the optimal control $u(t)$. The certainty equivalence principle is invoked during the control calculation procedure as we have assumed the optimal control signal can be obtained even with the estimates substituting the true parameters. That is, we have assumed

$$\hat{B}(q^{-1})u(t) = \hat{A}(q^{-1})y(t) \quad (2.13)$$

will yield the same optimal control as (2.4).

In the adaptive control literature, this method is classified as an implicit method or direct method since the parameters of the system are not estimated explicitly (thus implicit method) and that the parameters for the regulator are estimated directly (thus direct method). If an explicit estimation of the system parameters is being done, that is, the estimates $\hat{A}(z)$, $\hat{B}(z)$, and $\hat{C}(z)$ are obtained for the process (2.1), polynomial matrix factorizations and computations will have to be carried out before arriving at the optimal control signal $u(t)$. The direct method here allows simpler and faster computations for the optimal control.

To estimate the parameters of the regulator recursively, the following least squares procedure may be used [3, 6]. Introduce the parameter matrix Θ given by

$$\Theta = [\theta_1 \ \theta_2 \ \theta_3 \ \dots \ \theta_p] = \begin{bmatrix} \mathcal{A}_0^T \\ \vdots \\ \mathcal{A}_{n-1}^T \\ \mathcal{B}_0^T \\ \vdots \\ \mathcal{B}_{n+k-1}^T \end{bmatrix} \quad (2.14)$$

The following recursions are carried out at each step of time to estimate Θ for $i = 1, 2, \dots, p$:

$$\theta_i(t) = \theta_i(t-1) + K(t-1)[y_i(t) - \eta(t-k-1)\theta_i(t-1)] \quad (2.15)$$

$$K(t-1) = P(t-1)\eta^T(t-k-1)[1 + \eta(t-k-1)P(t-1)\eta^T(t-k-1)]^{-1} \quad (2.16)$$

$$P(t) = P(t-1) - K(t-1)[1 + \eta(t-k-1)P(t-1)\eta^T(t-k-1)]K^T(t-1) \quad (2.17)$$

with y_i being the i -th component of y and

$$\eta(t-k-1) = [-y^T(t-k-1) \dots -y^T(t-k-1-n_A)u^T(t-k-1) \dots u^T(t-k-1-n_B)] \quad (2.18)$$

where $n_A = n-1$ and $n_B = n+k-1$. It can be observed that if the initial values of $P(t)$ is the same for all of the p steps of the estimation at each step of time, the corresponding gain matrix $K(t)$ will remain constant at each step of time for all of the parameter vectors θ_i ($i = 1, 2, \dots, p$). In this manner, considerable computational efforts can be saved. Another remark is that other types of least squares estimation schemes may also be used to deal with slowly-varying parameters or to enhance numerical stability of the estimation computations.

Results for convergence of this adaptive algorithm have been reported [27, 28, 38, 39]. In [38, 39], convergence of optimal control, based on convergence of estimates, is guaranteed for single input single output system if the system input-output remains bounded and if a certain positive real condition of the noise dynamic $c(z)$ is satisfied. In [27, 28], the boundedness of the system variable is removed. Further discussion will be presented in Section 2.4. What should be kept in mind at this point

is that these convergence results developed for the STR will be used to determine convergence, first of the self-tuning controller (STC) which is presented in the next section, and eventually the stochastic adaptive game problems.

2.3 Self-Tuning Controller (STC)

The single input single output STC in [17] is basically the same as the STR except for a penalty on the control signal in the cost function. The presence of a penalty may reduce the excessive control signal magnitude that is common for the STR. It may also offer a simpler method to deal with nonminimum phase processes [25, 29]. We will generalize the single input single output (SISO) case to multiple input multiple output (MIMO) case here. See also [33] for another approach.

The cost function for the STC is given by

$$J = E\{y^T(t+k+1)Qy(t+k+1) + u^T(t)Ru(t)\} \quad (2.19)$$

where Q is a symmetric positive semidefinite matrix and R is a symmetric positive definite matrix.

To facilitate our analysis in the latter part of the report, we will consider the process under consideration to be governed by

$$a(q^{-1})y(t) = B(q^{-1})u(t-k-1) + C(q^{-1})e(t) \quad (2.20)$$

where k is known time delay and $a(z)$ is a scalar polynomial and $B(z)$, $C(z)$ are polynomial matrices given by

$$a(z) = 1 + a_1 z + \dots + a_n z^n \quad (2.21a)$$

$$B(z) = B_0 + B_1 z + \dots + B_{m-1} z^{m-1} \quad (2.21b)$$

$$C(z) = I + C_1 z + \dots + C_\ell z^\ell, \quad (2.21c)$$

and $C(z)$ has all its zeros outside the unit circle. The process given by (2.1) can readily be converted to (2.20) as shown in Appendix A.

2.3.1 Controller Design with Known Parameters

We will, again, first consider the system with known parameters and derive a strategy that minimizes (2.19) and then the adaptive algorithm for unknown parameters will be presented in the next section.

Theorem 2.1. The control law that minimizes (2.19) for the system (2.20) satisfies

$$MG(q^{-1})C(q^{-1})y(t) + [MF(q^{-1})C(q^{-1})B(q^{-1}) + \bar{c}(q^{-1})H]u(t) = 0 \quad (2.22)$$

where

$$M = B_0^T Q \quad (2.23a)$$

$$H = R \quad (2.23b)$$

$$\mathcal{C}(z) = \text{adjoint } C(z) \quad (2.23c)$$

$$\bar{c}(z) = \det C(z) \quad (2.23d)$$

$$G(z) = G_0 + G_1 z + \dots + G_{n-1} z^{n-1} \quad (2.23e)$$

$$F(z) = I + F_1 z + \dots + F_k z^k \quad (2.23f)$$

and $G(z)$, $F(z)$ satisfies

$$C(z) = a(z)F(z) + z^{k+1}G(z). \quad (2.24)$$

Proof. Premultiply (2.20) by $z^{k+1}F(z)C(z)$ to obtain

$$\begin{aligned} F(q^{-1})C(q^{-1})a(q^{-1})y(t+k+1) &= F(q^{-1})C(q^{-1})B(q^{-1})u(t) \\ &+ F(q^{-1})C(q^{-1})e(t+k+1) . \end{aligned}$$

Using (2.24), the above equation becomes

$$\begin{aligned} [C(q^{-1}) - q^{-(k+1)}G(q^{-1})]C(q^{-1})y(t+k+1) \\ = F(q^{-1})C(q^{-1})B(q^{-1})u(t) + F(q^{-1})C(q^{-1})C(q^{-1})e(t+k+1) \end{aligned}$$

or

$$\begin{aligned} \bar{c}(q^{-1})[y(t+k+1) - F(q^{-1})e(t+k+1)] \\ = G(q^{-1})C(q^{-1})y(t) + F(q^{-1})C(q^{-1})B(q^{-1})u(t) \end{aligned} \quad (2.25)$$

where the fact $\bar{c}(z)I = C(z)C(z)$ has been used. Denote

$$y^*(t+k+1|t) = y(t+k+1) - F(q^{-1})e(t+k+1) , \quad (2.26)$$

that is, y^* is the least squares optimal predictor of y given the data up to time t , which is uncorrelated with $F(q^{-1})e(t+k+1)$. Combining (2.25) and (2.26) yields

$$\begin{aligned} \bar{c}(q^{-1})y^*(t+k+1|t) &= G(q^{-1})C(q^{-1})y(t) \\ &+ F(q^{-1})C(q^{-1})B(q^{-1})u(t) . \end{aligned} \quad (2.27)$$

Substituting (2.26) into the cost function (2.19) yields

$$J = E\{[F(q^{-1})e(t+k+1)]^T Q [F(q^{-1})e(t+k+1)]\} \\ + E\{[y^*(t+k+1|t)]^T Q y^*(t+k+1|t) + u^T(t) R u(t)\} . \quad (2.28)$$

The first term of (2.28) is related to future noise covariance which cannot be optimized. Hence, the optimization is concentrated on the second term. Assuming the existence of $\frac{\partial J}{\partial u(t)}$, the necessary condition for a minimum yields

$$\frac{\partial J}{\partial u(t)} = 0 \\ = B_0^T Q y^*(t+k+1|t) + R^T u(t)$$

or

$$0 = M y^*(t+k+1|t) + H u(t) . \quad (2.29)$$

Premultiply (2.29) by $\bar{c}(q^{-1})$ and combining the resulting equation with (2.27), we have

$$M G(q^{-1}) C(q^{-1}) y(t) + [M F(q^{-1}) C(q^{-1}) B(q^{-1}) + \bar{c}(q^{-1}) H] u(t) = 0$$

as stated in (2.22).

Q.E.D.

A remark that is noteworthy is that in the STR where there is an absence of penalty on control, the optimal control is obtained by setting the predictor y^* to zero. This setting of the least squares optimal predictor to zero to compute the optimal control is the underlying factor that enables the parameters for the STR to be directly estimated in the adaptive situation. It will be most convenient if a similar direct method can be used for the STC. To accomplish this ultimate goal, we continue

our analysis on the optimal control equation (2.22) for known parameters to see if a suitable optimal least squares predictor function can be obtained.

Define a function ϕ^* such that

$$\phi^*(t+k+1|t) = My^*(t+k+1|t) + Hu(t) \quad (2.30)$$

where M and H are as defined in Theorem 2.1. From (2.29), the optimal control is obtained by setting ϕ^* to zero; thus, ϕ^* seems to be a possible candidate for the predictor function.

Let the function ϕ be defined by

$$\phi(t+k+1) = My(t+k+1) + Hu(t) . \quad (2.31)$$

Equations (2.20), (2.27), (2.30) and (2.31) yields the following system:

$$\begin{aligned} a(q^{-1})\phi(t) &= (MB(q^{-1}) + Ha(q^{-1}))u(t-k-1) \\ &\quad + MC(q^{-1})e(t) \end{aligned} \quad (2.32)$$

$$\begin{aligned} \bar{c}(q^{-1})\phi^*(t+k+1|t) &= MG(q^{-1})C(q^{-1})y(t) \\ &\quad + [MF(q^{-1})C(q^{-1})B(q^{-1}) + \bar{c}(q^{-1})H]u(t) \end{aligned} \quad (2.33)$$

$$\phi(t+k+1) = \phi^*(t+k+1|t) + MF(q^{-1})e(t+k+1) \quad (2.34)$$

since $y^*(t+k+1|t)$ is uncorrelated with $F(q^{-1})e(t+k+1)$; thus, $\phi^*(t+k+1|t)$ and $MF(q^{-1})e(t+k+1)$ are also uncorrelated, which implies ϕ^* is the least squares optimal predictor for ϕ . Furthermore if we define a new cost function I given by

$$I = E\{\phi^T(t+k+1)\phi(t+k+1)\} \quad (2.35)$$

then the minimum variance strategy for the system governed by (2.32) with cost function (2.35) is obtained by setting the optimal predictor ϕ^* to zero.

To summarize, the original system given by

$$a(q^{-1})y(t) = B(q^{-1})u(t-k-1) + C(q^{-1})e(t)$$

$$\bar{c}(q^{-1})y^*(t+k+1|t) = G(q^{-1})C(q^{-1})y(t) + F(q^{-1})C(q^{-1})B(q^{-1})u(t)$$

$$y(t+k+1) = y^*(t+k+1|t) - F(q^{-1})e(t+k+1)$$

has been transformed to an equivalent system

$$a(q^{-1})\phi(t) = (MB(q^{-1}) + Ha(q^{-1}))u(t-k-1) + MC(q^{-1})e(t)$$

$$\bar{c}(q^{-1})\phi^*(t+k+1|t) = MG(q^{-1})C(q^{-1})y(t)$$

$$+ [MF(q^{-1})C(q^{-1})B(q^{-1}) + \bar{c}(q^{-1})H]u(t)$$

$$\phi(t+k+1) = \phi^*(t+k+1|t) - MF(q^{-1})e(t+k+1)$$

with

$$\phi(t) = My(t) + Hu(t-k-1) .$$

An additional advantage to be gained in transforming the original system into a system which is similar to a STR structure is the possibility of applying directly the convergence results for the STR to the STC. In the latter part of the report, our convergence analysis for the game problem will also be based on this approach of transforming the original game problem to a system governed by (2.32-2.34).

The closed-loop system with (2.22) being applied becomes

$$\begin{aligned} & \bar{c}(q^{-1})(M + a(q^{-1})HB^{-1}(q^{-1}))y(t) \\ &= \bar{c}(q^{-1})(MF(q^{-1}) + HB^{-1}(q^{-1})C(q^{-1}))e(t) . \end{aligned} \quad (2.36)$$

Since $\bar{c}(z)$ is assumed to have all its roots outside the unit circle, the stability of the system is thus dependent on the roots of the system

$$\det(MB(z) + Ha(z)) \neq 0 . \quad (2.37)$$

Hence by choosing M and/or H properly, the system can be stabilized even with a $B(z)$ that does not have all its zeros outside the unit disc.

2.3.2 Controller for System with Unknown Parameters

In order to control the process given by (2.20) with unknown parameters, the following model is used for representation of the system

$$\phi(t) + \mathcal{A}(q^{-1})y(t-k-1) = \mathcal{B}(q^{-1})u(t-k-1) + e(t) \quad (2.38)$$

where ϕ is defined in (2.31) and

$$\begin{aligned} \mathcal{A}(z) &= \mathcal{A}_0 + \mathcal{A}_1 z + \dots + \mathcal{A}_{n-1} z^{n-1} \\ \mathcal{B}(z) &= \mathcal{B}_0 + \mathcal{B}_1 z + \dots + \mathcal{B}_{m+k-1} z^{m+k-1} \end{aligned}$$

and $e(t)$ is the error to be minimized in the least squares sense.

The certainty equivalent minimum variance control law for (2.38) is given by

$$\hat{\mathcal{B}}(q^{-1})u(t) = \hat{\mathcal{A}}(q^{-1})y(t) \quad (2.39)$$

where $\hat{\beta}(z)$ and $\hat{\alpha}(z)$ denote the least squares estimates for $\beta(z)$ and $\alpha(z)$ respectively.

The following recursive estimation scheme may be used [27, 28]. Introduce a parameter matrix Θ as defined in (2.14). Then at each step of time, the following recursions with $k = 0$ are carried out to estimate Θ for $i = 1, 2, \dots, p$:

$$\theta_i(t) = \theta_i(t-1) + \frac{\bar{a}}{r_i(t-1)} \eta_i(t-1) [\phi_i(t) - \eta_i^T(t-1)\theta_i(t-1)] \quad (2.40)$$

$$r_i(t-1) = r_i(t-2) + \eta_i^T(t-1)\eta_i(t-1), \quad r_i(0) = 1 \quad (2.41)$$

where η_i is given by

$$\eta_i^T(t) = [-y^T(t) \dots -y^T(t-n+1) u^T(t) \dots u^T(t-m+1)] \quad (2.42)$$

and $\bar{a} > 0$ is a constant. See [27] for general delay $k \neq 0$.

Using ϕ as the input to the controller, it is possible to avoid some of the complex matrix calculations and determine the controller parameters directly. However, this may present some problem since knowledge of the B_0 parameter is required in computing the signal ϕ . In the present case for a single decision-maker, this may not be extremely annoying since an arbitrary choice for the matrix M results only in a change of the penalty on the output variances. Certainly, one method to overcome the problem is to estimate the system parameters explicitly and go through all the matrix computations. We will, however, at this point assume that the B_0 parameter in the process is known as this does not appear to be a very stringent requirement in practical applications.

It should be noted that similar derivation of the self-tuning controller strategy is carried out in [33] using different system representations. However, in our approach, by adhering to the system representation (2.20), convergence and stability results in [27] can be readily established for the STC as shown in the next section.

2.4 Convergence Analysis for Self-Tuning

It has been shown in [38, 39] for the SISO STR that the estimated parameters of the regulator will converge and yield the optimal control based on true system parameters if the following conditions are satisfied:

- i) the sequences $\{y(t)\}$, $\{u(t)\}$ are uniformly bounded;
- ii) the polynomial $(\frac{1}{c(q^{-1})} - \frac{1}{2})$ is strictly positive real; and
- iii) there is no factor common to $A(z)$, $B(z)$ and $C(z)$ in (2.1).

Similar analysis for the MIMO case has been reported in [14].

In another approach using Martingale theory in [27], convergence of parameters is not explicitly required and the boundedness of the system input-output is removed. Their result is stated in the following theorem.

Theorem 2.2 [27, Theorem 5.1]. Consider the cost function (2.3) and the system (2.20) which satisfies the following assumptions:

- i) the number of inputs p equals the number of outputs;
- ii) the delay $k = 0$;
- iii) upper bounds for the orders of the scalar polynomials appearing in $\{a(q^{-1}), B(q^{-1}), C(q^{-1})\}$ are known;

- iv) $\det B(z) \neq 0 \quad |z| \leq 1$
 $\det C(z) \neq 0 \quad |z| \leq 1$;
- v) $(\bar{c}(z) - \frac{\bar{a}}{2})$ is strictly positive real,

then with probability one,

$$(1) \quad \sup_N \frac{1}{N} \sum_1^N \|y(t)\|^2 < \infty$$

$$(2) \quad \sup_N \frac{1}{N} \sum_1^N \|u(t)\|^2 < \infty$$

$$(3) \quad \lim_N \frac{1}{N} \sum_1^N E\{y_i^2(t)\} = \Gamma_{ii}, \quad i = 1, 2, \dots, p$$

where Γ_{ii} is the minimum mean square error for any causal linear feedback (including the one designed using true system parameters), if the STR algorithm in Section (2.2.2) is applied to the system.

Notice the presence of \bar{a} , which is present in (2.40), allows a certain degree of freedom to ensure condition v) is satisfied.

The result of Theorem 2.2 will be applied to the equivalent transformed system when the STC algorithm is used. Hence, the process will be governed by (2.32) with the cost function (2.35). By modifying certain assumptions on the system, we see that the input-output will remain bounded and the control error of the transformed system will achieve its global minimum. Hence, we have the following theorem.

Theorem 2.3. Consider the cost function (2.19) and the transformed system (2.32) which satisfies assumptions i), ii), v) of Theorem 2.2 and the following conditions:

- iii) upper bounds for the order of the scalar polynomials appearing in $\{a(q^{-1}), MB(q^{-1}) + Ha(q^{-1}), C(q^{-1})\}$ are known;

- iv) $\det(MB(z) + Ha(z)) \neq 0$, $|z| \leq 1$
 $\det C(z) \neq 0$, $|z| \leq 1$;
vi) B_0 is known.

If the STC algorithm in Section 2.3.2 is applied to the system, then the input-output will remain bounded and the output error of $\phi(t)$ will achieve a minimum achievable by any causal linear feedback.

Theorem 2.3 will be used in showing convergence of the game problems that is considered in the latter part of the report.

2.5 Simulation Example

To illustrate some of the features of the STR and STC, an example based on a paper making machine is simulated and evaluated. See [14] for details of the model.

The plant is governed by

$$y(t) + A_1 y(t-1) = B_0 u(t-1) + e(t)$$

where

$$A_1 = \begin{bmatrix} -0.99101 & 8.80512 \times 10^{-3} \\ -0.80610 & -0.77089 \end{bmatrix}$$

$$B_0 = \begin{bmatrix} 0.89889 & -4.59328 \times 10^{-3} \\ 19.390 & 0.88052 \end{bmatrix}$$

$$E\{e(t)e^T(t)\} = \begin{bmatrix} 0.02 & 0.35 \\ 0.35 & 7.6 \end{bmatrix} .$$

The cost function J is given by

$$J = E\{y^T(t+1)Qy(t+1) + u^T(t)Ru(t)\}$$

with

$$Q = \begin{bmatrix} 6 & 0 \\ 0 & 10 \end{bmatrix}$$

$$R = \begin{cases} 0 & , \text{ for the STR} \\ \begin{bmatrix} 0.05 & 0 \\ 0 & 0.05 \end{bmatrix} & , \text{ for the STC} \end{cases}$$

All initial parameters except $\hat{\mathcal{B}}_0$ are set to zero. $\hat{\mathcal{B}}_0$ is set to the identity matrix to prevent control saturation.

Self-Tuning Regulator: The STR algorithm estimates recursively the parameters of the model

$$y(t) + \hat{\mathcal{A}}_0 y(t-1) = \hat{\mathcal{B}}_0 u(t-1) + \epsilon(t) ,$$

and the control is given by

$$u(t) = \hat{\mathcal{B}}_0^{-1} \hat{\mathcal{A}}_0 y(t) .$$

The input-output of a typical run will be presented and discussed later.

Self-Tuning Controller: The STC algorithm estimates recursively the parameters of the model

$$\phi(t) + \tilde{\mathcal{A}}_0 y(t-1) = \tilde{\mathcal{B}}_0 u(t-1) + \epsilon(t) ,$$

and the control is given by

$$u(t) = \hat{\beta}_0^{-1} \hat{a}_0 y(t) .$$

The input to the controller $\phi(t)$ is given by

$$\phi(t) = y(t) + M^{-1} H u(t-1)$$

with $M = B_0^T Q$, $H = R^T$. Notice since M^{-1} exists in this case, our operation in premultiplying (2.31) by M^{-1} is justified. Such transformation of ϕ will not affect the stabilizing property of the algorithm.

Comparisons of STR and STC: In order to compare the different features of the two self-tuners, we have carried out the simulation in two parts. We will first assume B_0 is known, thus β_0 and M are known. The input-output of a typical run is shown in Figures 2.1a-2.4a for the STR and the corresponding trajectories for the STC are shown in Figures 2.1b-2.4b. In the second part, B_0 is unknown and M is arbitrarily chosen as the identity matrix for the STC. The input-output of a typical run (with other conditions the same as the first part) is shown in Figures 2.5a-2.8a for the STR and the corresponding trajectories for the STC is shown in Figures 2.5b-2.8b.

The simulation results indicate that both self-tuners indeed perform satisfactorily regardless of the knowledge of B_0 . However, from Figure 2.3 and Figure 2.4, we notice that u_1 and u_2 are substantially reduced especially during start up if the STC is used. The prevention of excessive control action is attained in this instance. From Figures 2.5-2.8, it can be observed again the reduction of excessive control and thus excessive output variance is achieved even with unknown B_0 . The STC does seem to have an edge in terms of smoother control action. However, we

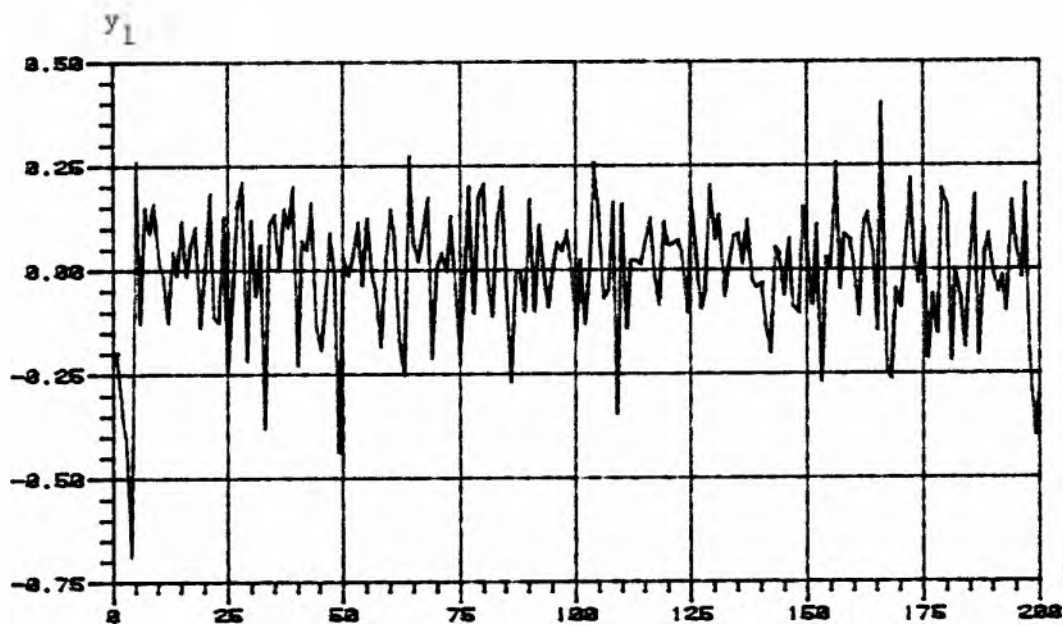


Figure 2.1a. Time response of y_1 using STR with known B_0 .

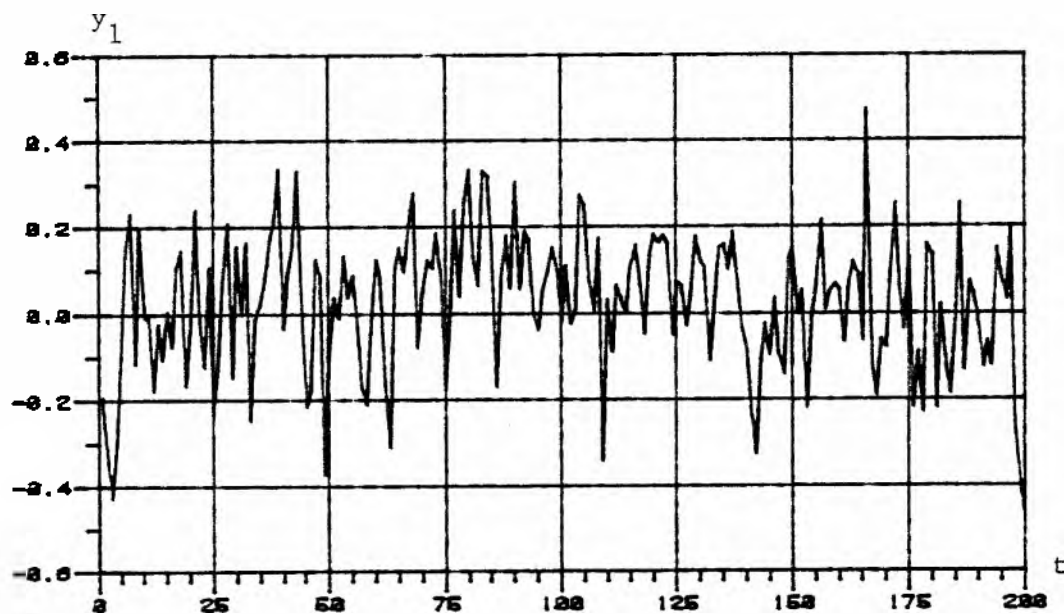


Figure 2.1b. Time response of y_1 using STC with known B_0 .

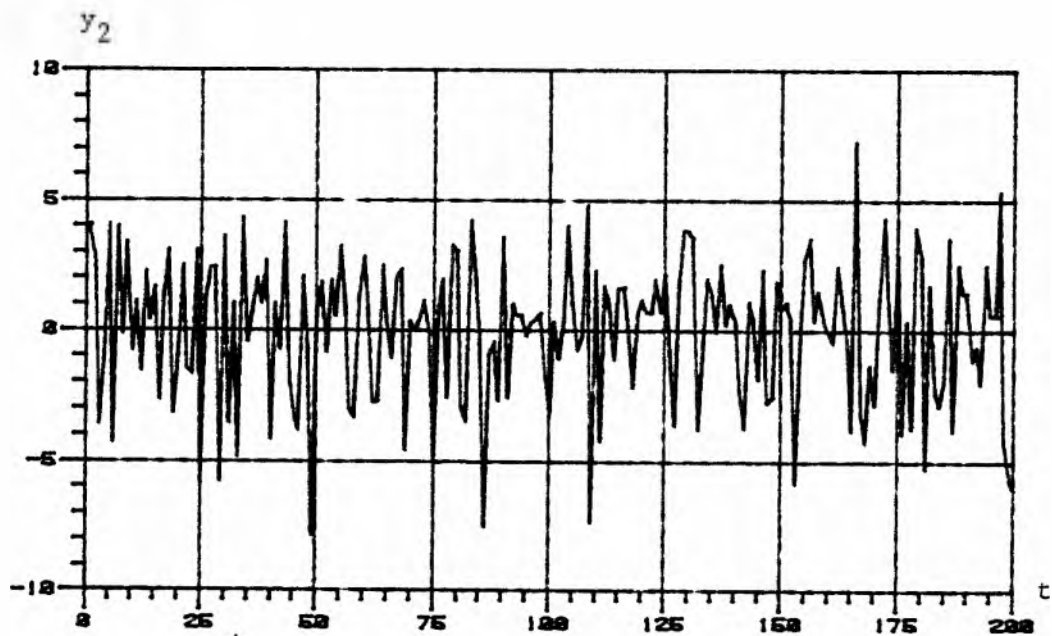


Figure 2.2a. Time response of y_2 using STR with known B_0 .

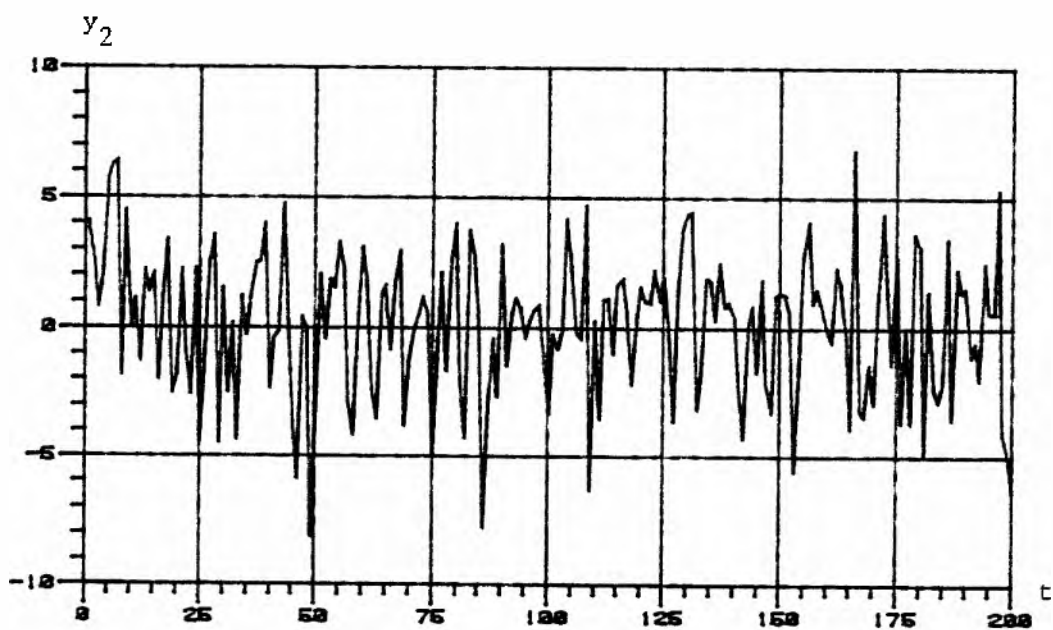


Figure 2.2b. Time response of y_2 using STC with known B_0 .

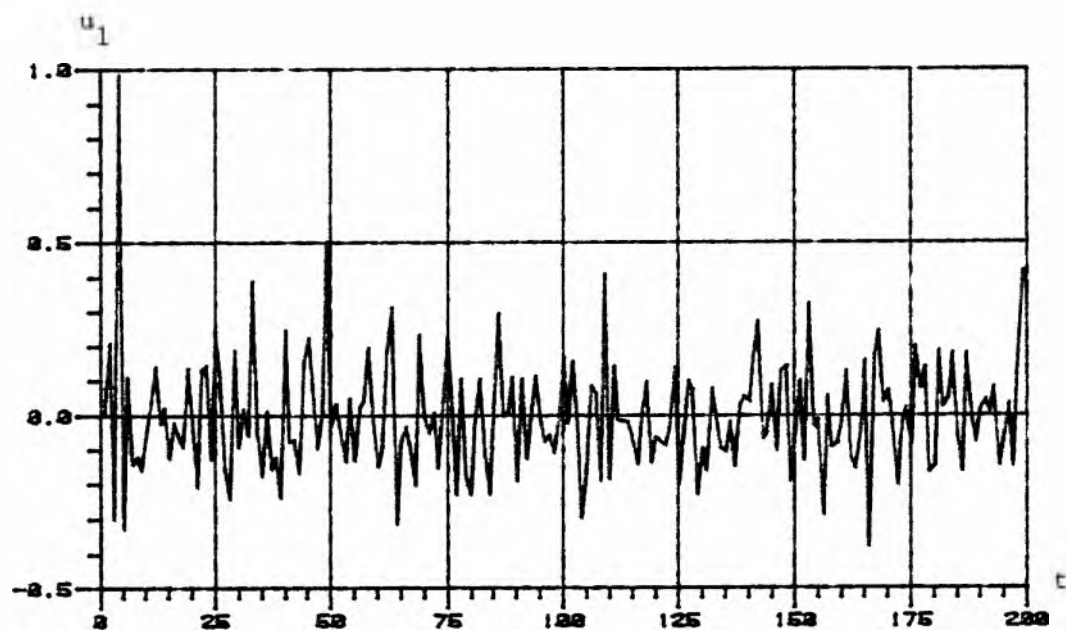


Figure 2.3a. Time response of u_1 using STR with known B_0 .

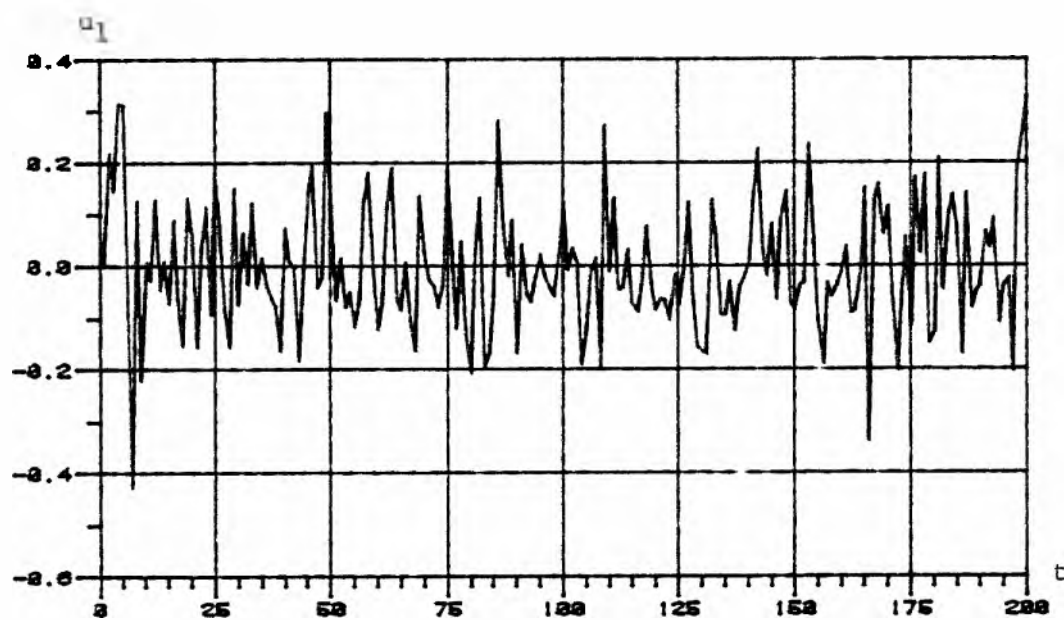


Figure 2.3b. Time response of u_1 using STC with known B_0 .

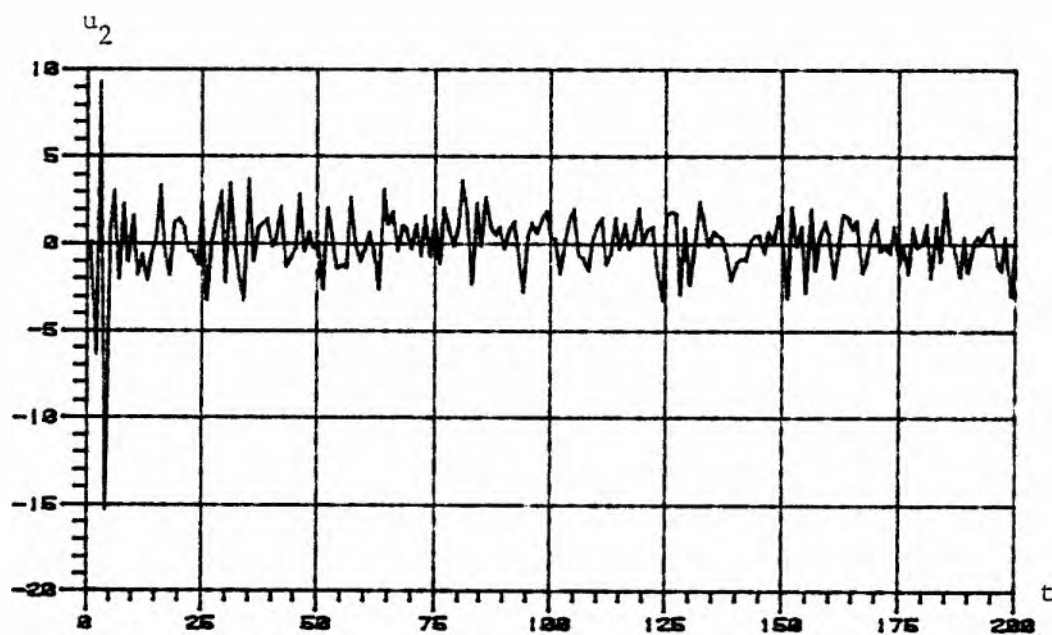


Figure 2.4a. Time response of u_2 using STR with known B_0 .

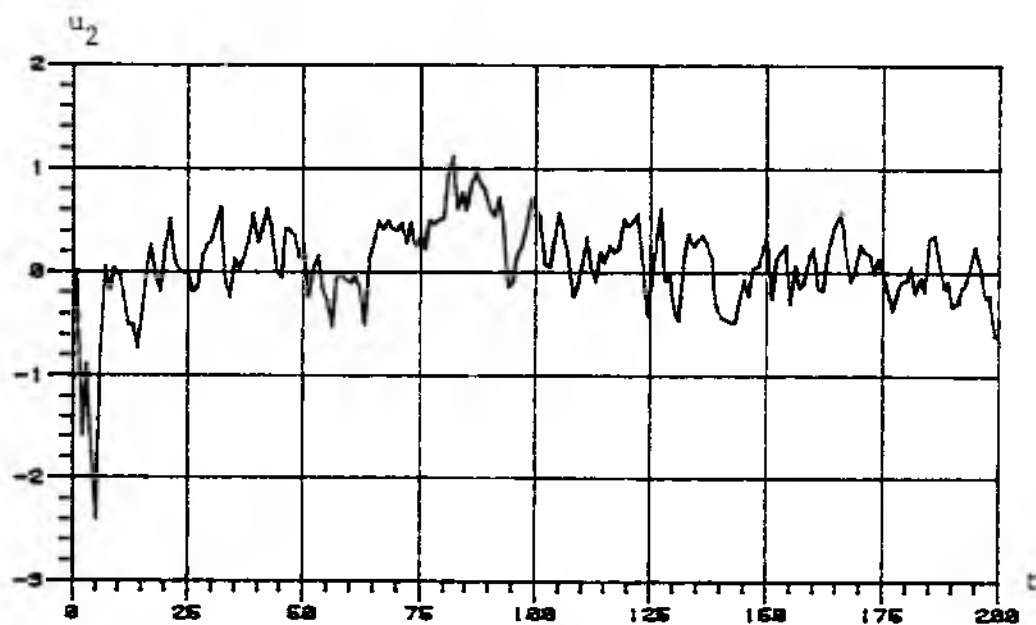


Figure 2.4b. Time response of u_2 using STC with known B_0 .

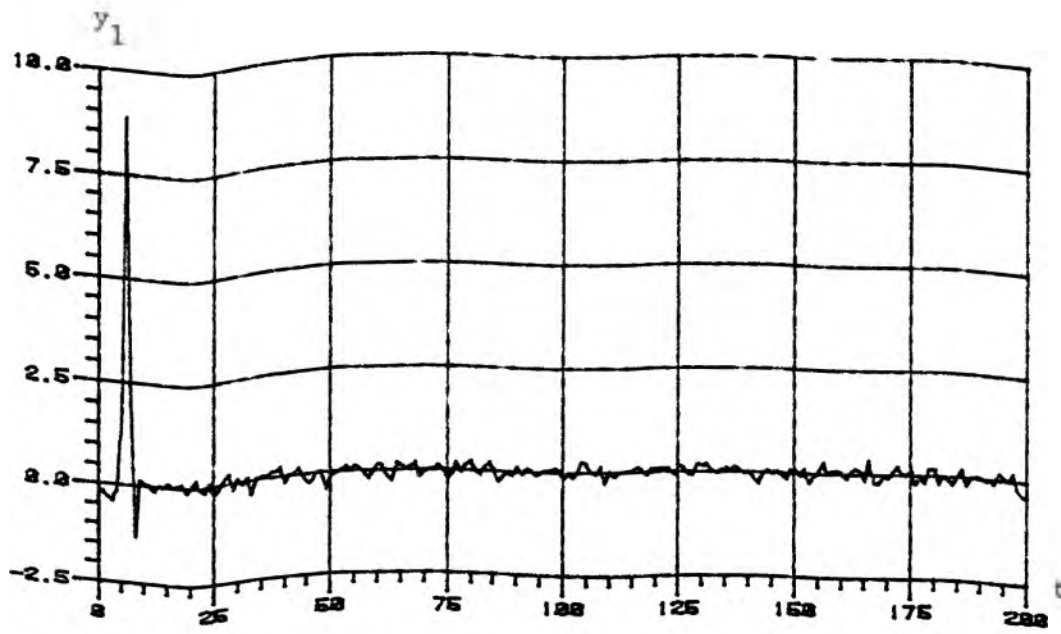


Figure 2.5a. Time response of y_1 using STR with unknown B_0 .

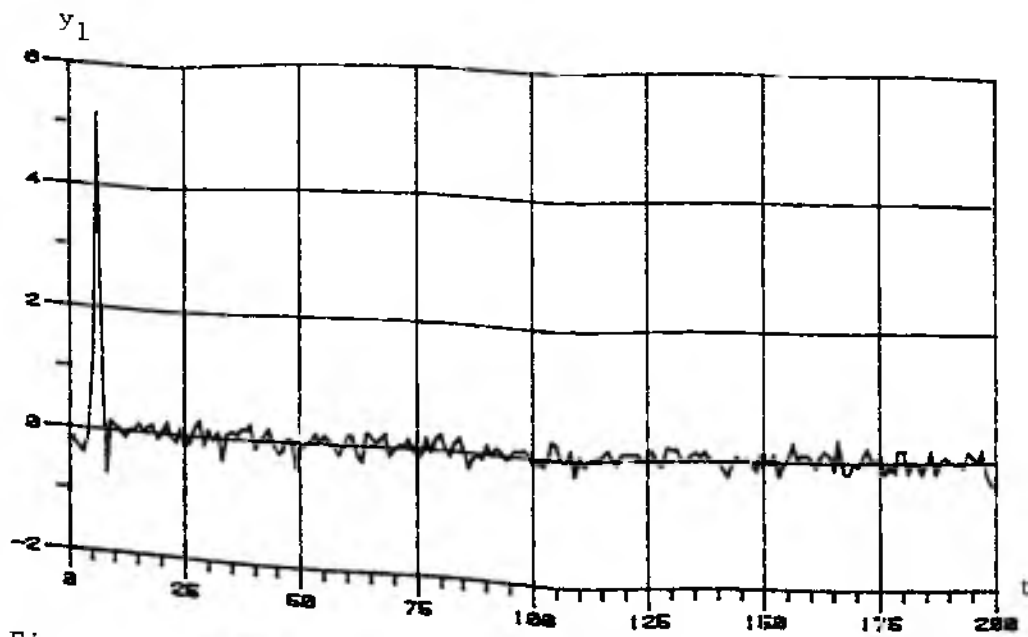


Figure 2.5b. Time response of y_1 using STC with unknown B_0 .

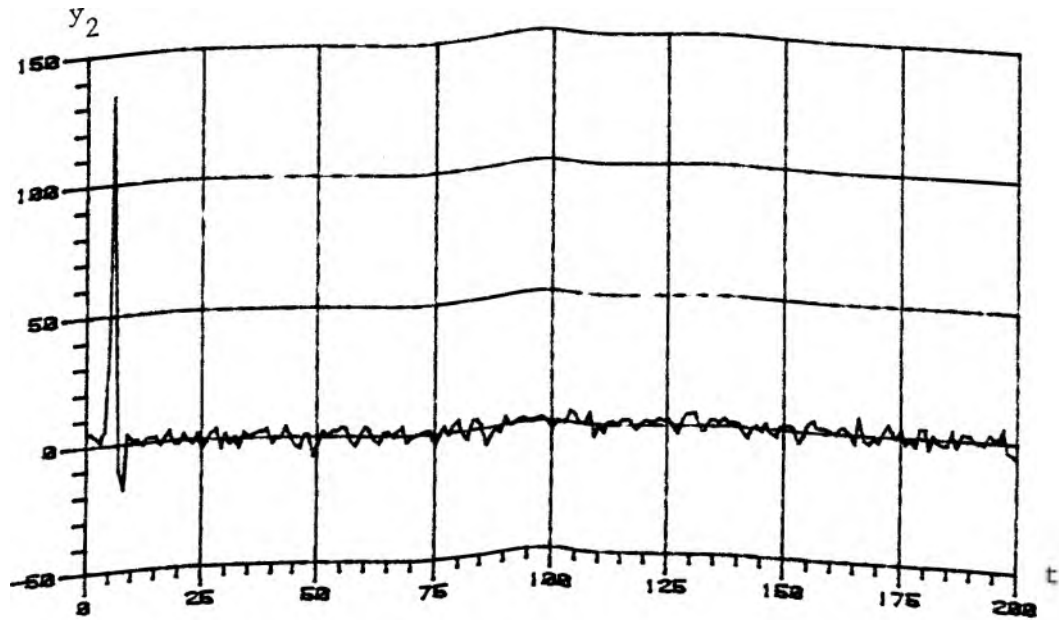


Figure 2.6a. Time response of y_2 using STR with unknown B_0 .

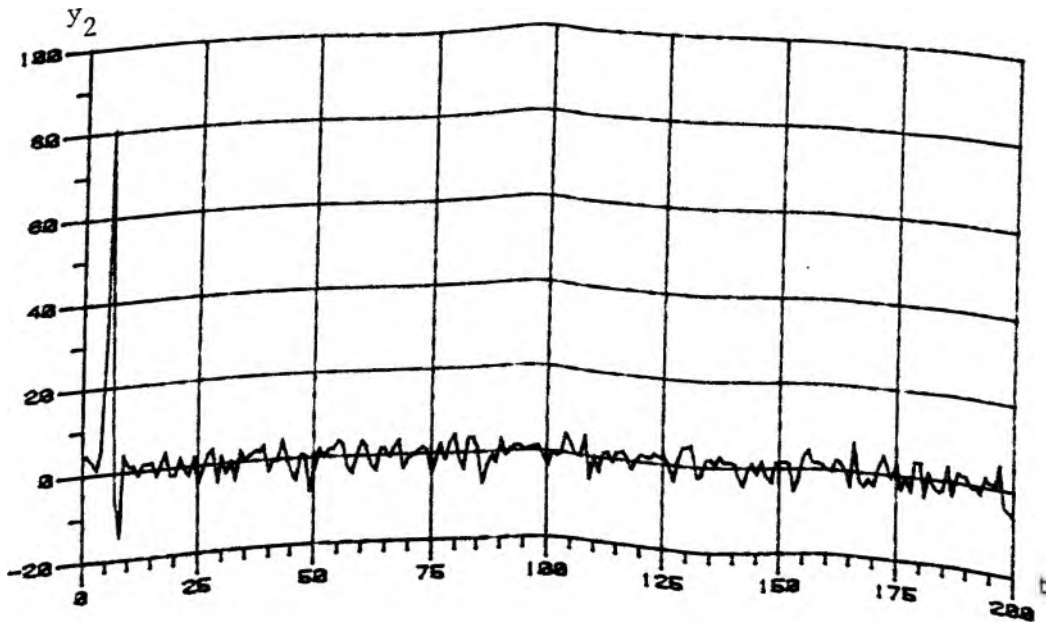


Figure 2.6b. Time response of y_2 using STC with unknown B_0 .

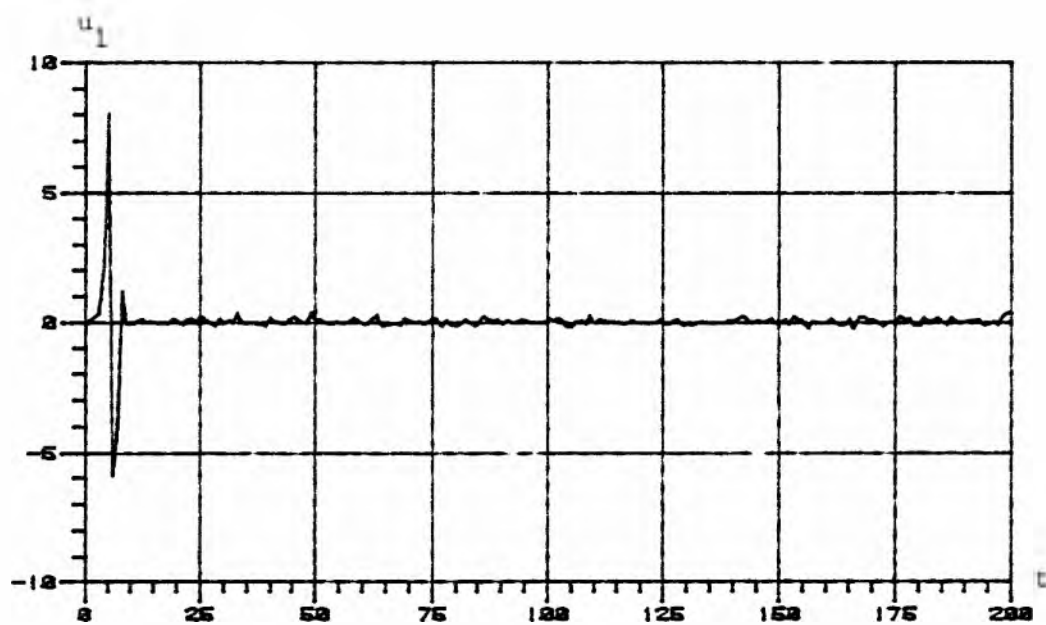


Figure 2.7a. Time response of u_1 using STR with unknown B_0 .

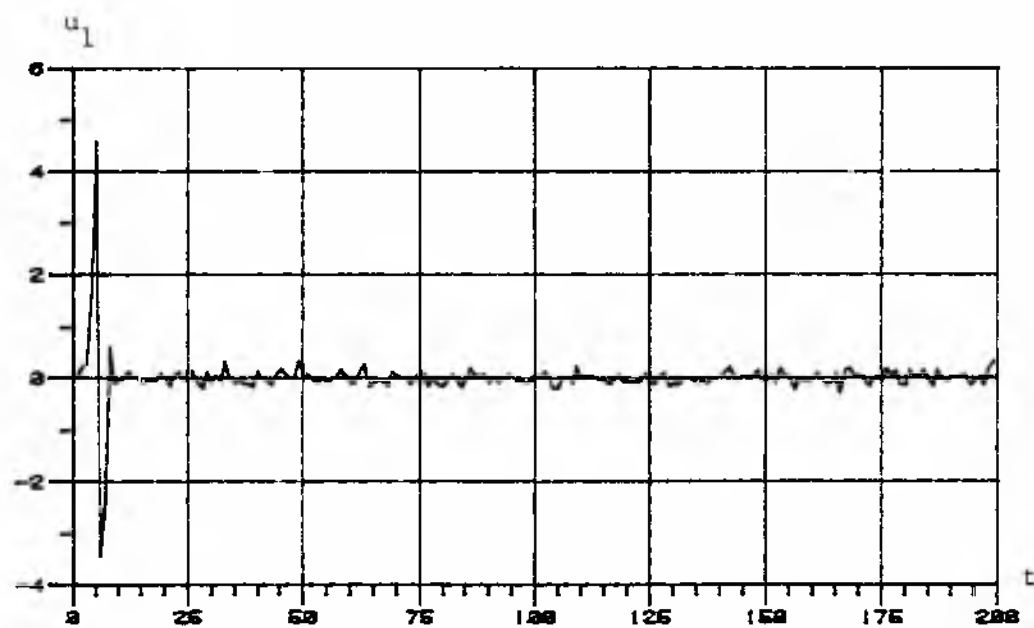


Figure 2.7b. Time response of u_1 using STR with unknown B_0 .

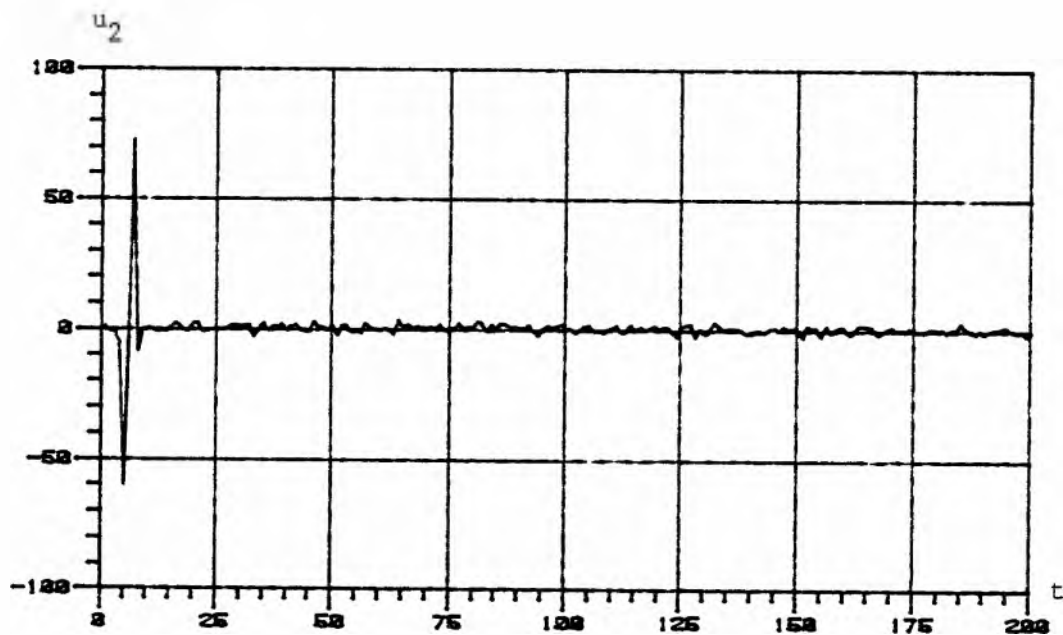


Figure 2.8a. Time response of u_2 using STR with unknown B_0 .

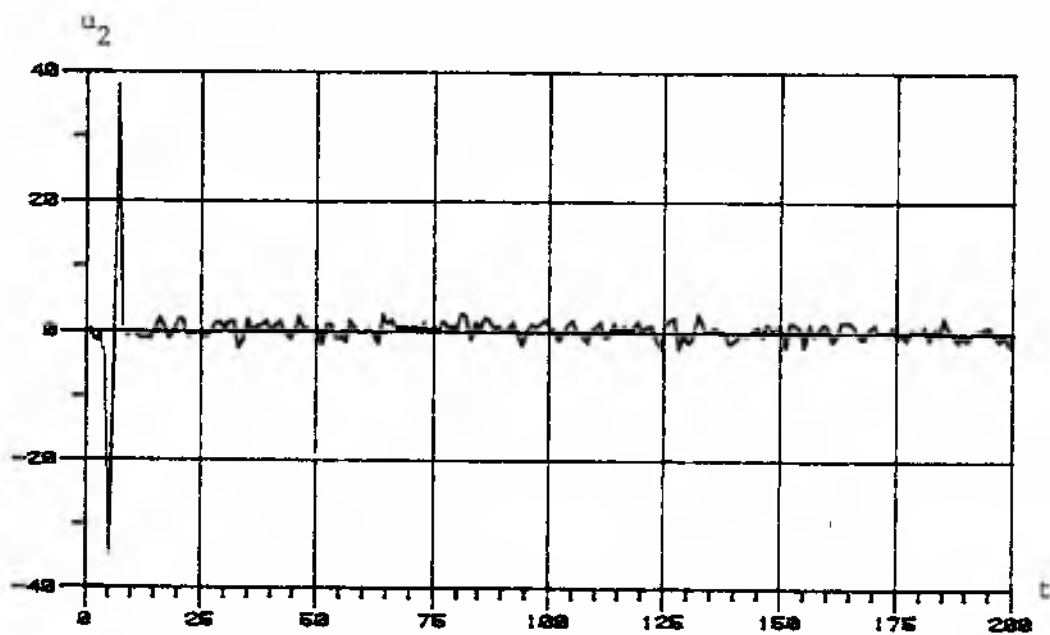


Figure 2.8b. Time response of u_2 using STC with unknown B_0 .

should take caution not to penalize the control signal excessively as this may result in inadequate probing of the system to yield estimates that can stabilize the system.

CHAPTER 3

STOCHASTIC ADAPTIVE NASH GAMES

3.1 Introduction

In this chapter we will consider the stochastic adaptive Nash game problem using the self-tuning controller (STC) approach. Nash games were first introduced and investigated in a static framework in [42]. They were later extended to the dynamic case [53, 54]. The decision-makers in a Nash game simultaneously minimize their respective cost functions with respect to their individual controls. The resulting optimal strategy is called the Nash equilibrium strategy. This strategy has the property that if one decision-maker deviates from it, he cannot improve his performance. However, it may be possible for some or all of the decision-makers to improve their performance when more than one decision-maker deviates from the equilibrium strategy. That is, the Nash equilibrium strategy is secure against unilateral deviation but not necessarily collusion. The Nash game framework, thus, is very appealing to large scale systems or distributed industrial systems where there are a host of noncooperative decision-makers or controllers each trying to minimize his own cost functional.

Definition 3.1. A strategy set $\{u_1^*, u_2^*, \dots, u_N^*\}$ is a Nash equilibrium strategy set if

$$J_i(u_1^*, \dots, u_{i-1}^*, u_i^*, u_{i+1}^*, \dots, u_N^*) \leq J_i(u_1^*, \dots, u_{i-1}^*, u_i, u_{i+1}^*, \dots, u_N^*) \quad , \quad i = 1, 2, \dots, N \quad (3.1)$$

for all admissible controls u_i of decision-maker i ; and J_i is the cost function for decision-maker i for which that decision-maker is trying to minimize.

As in stochastic optimal control problems, there are different solution concepts, namely, open-loop and closed-loop solutions, to the Nash game problem [53]. In general, the open-loop and closed-loop solution of a game problem is different. However, by restricting the cost function of the decision-maker to single-stage, the distinction between the different types of solution cease to exist as we have essentially reduced the problem to a static framework. Moreover, the restriction also enables us to seek steady state solutions to the game problem using the self-tuning approach.

In Section 3.2, formulation of the Nash game problem is presented. The solution to the stochastic adaptive Nash game problem will be discussed in Section 3.3. It turns out that the game solution closely resembles, after a judicious transformation, a minimum variance control problem with one decision-maker. It is this resemblance that enables the established results for the STR to be applied to the game problem. Finally, in Section 3.4, a simulation study on an economic system is presented to illustrate the proposed adaptive solution.

3.2 Problem Formulation

Previously, dynamic games have mostly been analyzed for system in state-space representation. We will formulate the game problem in the input-output form so that the self-tuning algorithm can readily be applied. Consider a system given by the equations

$$x(t+1) = Fx(t) + G_1 u_1(t) + G_2 u_2(t) + \dots + G_N u_N(t) + Ke(t) \quad (3.2)$$

$$y(t) = Tx(t) + e(t) \quad (3.3)$$

where each u_i , which is of dimension m_i , represents a controller that tries to minimize a cost function given by

$$J_i = E\{[y(t+k+1) - y^r(t+k+1)]^T Q_i [y(t+k+1) - y^r(t+k+1)] + [u(t) - u(t-1)]^T R_i [u(t) - u(t-1)]\}, \quad (3.4)$$

$$i = 1, 2, \dots, N$$

where Q_i is a symmetric positive semidefinite matrix and y^r is the desired value of y . The matrix R_i is a symmetric matrix with its i, i -th block, denoted by $(R_i)_{ii}$, being positive definite. The vector u is formed by stacking all the u_i 's ($i = 1, 2, \dots, N$) in a column. The reason for penalizing the term $[u(t) - u(t-1)]$ is to avoid finding the reference control signal u^r that corresponds to a nonzero y^r . The state vector $x(t)$ is n -dimensional. The input $u(t)$, output $y(t)$ and the noise sequence $\{e(t)\}$ are all of dimension p (that is, $p = \sum_{i=1}^N m_i$). Furthermore, $\{e(t)\}$ is assumed to be an independent equally distributed zero mean random vector with finite covariance. Let $G = [G_1, G_2, \dots, G_N]$, then (3.2) becomes

$$x(t+1) = Fx(t) + Gu(t) + Ke(t). \quad (3.5)$$

It can be shown that (3.5) and (3.3) can be transformed to an input-output representation [27], which is given by

$$a(q^{-1})y(t) = B(q^{-1})u(t-k-1) + C(q^{-1})e(t), \quad k \geq 0 \quad (3.6)$$

where k is a known time delay and $a(z)$ is the scalar characteristic polynomial for the system (3.5) and $a(z)$, $B(z)$, $C(z)$ are in the form as given in (2.21).

To allow more flexibility, we will consider the system to be governed by

$$a(q^{-1})y(t) = B(q^{-1})u(t-k-1) + C(q^{-1})e(t) + D \quad (3.7)$$

where D is a p -dimensional offset vector.

3.3 Self-Tuning Nash Game

The self-tuning approach is adopted to seek steady state solutions for the stochastic adaptive Nash game problem. As in the usual analysis for such an approach, the control strategy is first derived assuming all the parameters are known, then an adaptive procedure is incorporated to deal with unknown parameters.

3.3.1 N-Person Nash Equilibrium Strategy

The derivation of the Nash equilibrium strategy is very similar to that of the STC. We will summarize the result in a theorem.

Theorem 3.1. Let $L^{(i)}$ represent the i -th column block (of dimension $p \times m_i$) of the $p \times p$ matrix L . The Nash Equilibrium Strategy $u^*(t)$ for the system (3.7) with cost functions (3.4) is given by

$$\begin{aligned}
& MG(q^{-1})C(q^{-1})y(t) + [MF(q^{-1})C(q^{-1})B(q^{-1}) \\
& + (1-q^{-1})\bar{c}(q^{-1})H]u^*(t) + MF(q^{-1})C(q^{-1})D \\
& - \bar{c}(q^{-1})My^r_{(t+k+1)} = 0
\end{aligned} \tag{3.8}$$

with

$$C(z) = \text{adjoint } C(z) \tag{3.9a}$$

$$\bar{c}(z) = \det C(z) \tag{3.9b}$$

$$G(z) = G_0 + G_1 z + \dots + G_{n-1} z^{n-1} \tag{3.9c}$$

$$F(z) = I + F_1 z + \dots + F_k z^k \tag{3.9d}$$

$$M = \begin{bmatrix} T(1) \\ B_0 Q_1 \\ \vdots \\ T(N) \\ B_0 Q_N \end{bmatrix} \tag{3.9e}$$

$$H = \begin{bmatrix} T(1) \\ R_1 \\ \vdots \\ T(N) \\ R_N \end{bmatrix} \tag{3.9f}$$

and $G(z)$, $F(z)$ satisfy the following identity

$$C(z) = a(z)F(z) + z^{k+1}G(z) . \tag{3.9g}$$

Proof. See Appendix B.

Notice that (3.8) is just a system of linear equations, which is extremely convenient in the computational aspects over other game solutions

that usually involves Riccati equations. Another observation of (3.8) is that it closely resembles the optimal control law for the STC given in Theorem 2.1 except for the definition of the matrices M and H . Hence, the original system (3.7) with $y(t)$ as output can be transformed into an equivalent system with $\phi(t)$ as output as in the case for the STC.

Let the function $\phi(t)$ be defined by

$$\phi(t) = M(y(t) - y^r(t)) + H(u(t-k-1) - u(t-k-2)) . \quad (3.10)$$

The equivalent transformed system is then given by

$$\begin{aligned} a(q^{-1})\phi(t) = & (MB(q^{-1}) + (1-q^{-1})a(q^{-1})H)u(t-k-1) \\ & + MC(q^{-1})e(t) + MD - a(q^{-1})My^r(t) . \end{aligned} \quad (3.11)$$

Furthermore, as in the STC, by defining a new cost function I given by

$$I = E\{\phi^T(t+k+1)\phi(t+k+1)\} , \quad (3.12)$$

it is possible to obtain the Nash strategy (3.8) by considering the transformed system (3.11) with the cost function (3.12) as a control problem with only one decision-maker. That is, instead of optimizing J_i with respect to $u_i(t)$ for the i -th decision-maker, every decision-maker can determine the Nash strategy (3.8) by optimizing I with respect to $u(t)$.

Another interesting property of the solution (3.8) is that if the penalty on control in the J_i 's is zero, and assuming the matrix M is non-singular, the resulting Nash strategy is equivalent to the MIMO minimum variance strategy developed for the STR in which there is only one decision-maker. Moreover, premultiplying (3.8) by M^{-1} when H is zero,

we see that this Nash strategy is independent of the weighting matrices Q_i ($i = 1, 2, \dots, N$). Essentially, the game flavor of the problem will not arise if every decision-maker does not penalize the control effort. On the other hand, even when there are penalties on controls, and if all the Q_i 's and R_i 's are identical for $i = 1, 2, \dots, N$, the Nash strategy collapses to the optimal strategy of the STC in Theorem 2.1. Situations in which every controller has the same cost functional are analyzed in the realm of team theory [48].

3.3.2 N-Person Self-Tuning Nash Equilibrium Strategy

Basically, the same approach utilized in the MIMO STC will be used to deal with unknown parameters. However, further restrictions have to be placed on each decision-maker. From this point on, we assume every controller agrees to use the same estimation scheme and identical initial conditions. These restrictions ensure every controller has the same model for the system and rid us of the complications of multimodeling. With these restrictions, each decision-maker is essentially a complete STC by himself. That is, there are N identical STC doing identical computations to compute the Nash equilibrium strategy. In other words, in order to arrive at the Nash strategy for the system (3.7), every decision-maker uses the following model for representation of the process

$$\begin{aligned} \phi(t) = & \alpha(q^{-1})y(t-k-1) + \beta(q^{-1})u(t-k-1) + \delta \\ & + \gamma(q^{-1})y^r(t) + \epsilon(t) \end{aligned} \quad (3.13)$$

with $\phi(t)$ defined in (3.10) and

$$\alpha(z) = \alpha_0 + \alpha_1 z + \dots + \alpha_{n-1} z^{n-1} \quad (3.14a)$$

$$\beta(z) = \beta_0 + \beta_1 z + \dots + \beta_{m+k-1} z^{m+k-1} \quad (3.14b)$$

$$\bar{\beta}(z) = \bar{\beta}_0 + \bar{\beta}_1 z + \dots + \bar{\beta}_{n_{\bar{\beta}}} z^{n_{\bar{\beta}}}, \quad n_{\bar{\beta}} = \text{degree } \bar{c}(z) \quad (3.14c)$$

$$\beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}_{p \times 1} \quad (3.14d)$$

and $e(t)$ is the error to be minimized in the least squares sense.

The certainty equivalent Nash equilibrium strategy for the system (3.13) is given by

$$0 = \hat{\beta}(q^{-1})u(t) + \hat{\alpha}(q^{-1})y(t) + \hat{\beta}(q^{-1})y^r(t+k+1) + \hat{\delta} \quad (3.15)$$

where \hat{W} denotes the estimates of W . The parameters can be estimated using (2.40) and (2.41) with

$$\eta_1^T(t) = [y^T(t)y^T(t-1)\dots u^T(t)u^T(t-1)\dots y^{rT}(t+k+1)\dots 1] \quad (3.16)$$

and the parameter matrix Θ defined by

$$\Theta = [\theta_1 \dots \theta_p] = \begin{bmatrix} \alpha_0^T \\ \vdots \\ \alpha_0^T \\ \vdots \\ \alpha_1^T \\ \vdots \\ \alpha_1^T \\ \vdots \\ \alpha_{H-1}^T \end{bmatrix} \quad (3.17)$$

The following recursions (for $k = 0$) are then carried out at each step of time:

$$\theta_i(t) = \theta_i(t-1) + \frac{\bar{a}}{r_i(t-1)} \eta_i(t-1) [\phi_i(t) - \eta_i^T(t-1)\theta_i(t-1)] \quad (3.18)$$

$$r_i(t-1) = r_i(t-2) + \eta_i^T(t-1)\eta_i(t-1), \quad r_i(0) = 1. \quad (3.19)$$

3.3.3 Convergence

Convergence of the estimated prediction $\hat{\phi}^*$ to the true prediction ϕ^* can be analyzed using Theorem 2.3. By defining the matrix M and H according to (3.9e) and (3.9f) respectively, we can apply Theorem 2.3 to show convergence of the game problem, which is stated in the following theorem.

Theorem 3.2. Consider the cost functions J_i ($i = 1, 2, \dots, N$) in (3.4) for the system (3.13) which is assumed to satisfy condition i), ii) and v) of Theorem 2.3 and the following conditions:

- iii) upper bounds for the order of all the scalar polynomials appearing in $\{a(q^{-1}), (MB(q^{-1}) + (1-q^{-1})a(q^{-1})H), C(q^{-1})\}$ are known;
- iv) $\det(MB(z) + (1-z)a(z)H) \neq 0, \quad |z| \leq 1$
 $\det C(z) \neq 0, \quad |z| \leq 1$
 with M and H as defined in (3.9).

If the Self-Tuning Nash algorithm (3.15)-(3.19) is applied to the system, then the system input-output will remain bounded and the prediction error for $\phi(t)$ will tend to its global minimum with probability one.

3.4 Simulation Example

During recent years, optimal control theory has been widely used in the field of economic analysis [7, 10, 31, 46, 55]. Optimal control theory seems to provide an extremely versatile tool for the economist to determine tradeoffs between policies, economic stabilization policies, long term investment policies and other functions alike. Hence, the Nash strategy proposed here is applied to a rather simple minded quarterly economic model with two inputs and two outputs to illustrate the ability of the algorithm to stabilize the system. The two outputs are the consumption expenditure $C(t)$ and private investment $I(t)$. The two inputs are government expenditure $G(t)$ and money supply $M(t)$. All variables are measured in constant 1958 dollars. Further details can be found in [15]. In the United States, the formulation of the monetary policy is in the domain of the Federal Reserve System (FRS) while the formulation of the fiscal policy is primarily in the hands of the Congress and the President [47]. There have been many instances during which the two "controllers" hold different objectives. This certainly falls naturally into a game framework. We will assume that the two controllers (FRS and the federal government) want to stabilize this system along certain target paths or growth patterns. However, they have different views on where the emphasis should be placed, which is manifested by having different cost functionals. Let J_1 and J_2 be the cost functions of the federal government (Congress and the President) and FRS respectively. The J_i 's are given by

$$J_1 = E\{[y(t+1) - y^r(t+1)]^T Q_1 [y(t+1) - y^r(t+1)] \\ + [u(t) - u(t-1)]^T R_1 [u(t) - u(t-1)]\}$$

where $y(t) = [C(t) \ I(t)]^T$, $u(t) = [G(t) \ M(t)]^T$ and y^r is the desired output. We assume $y^r(t)$ grows at an annual rate of 4% from $y^r(0) = [300 \ 75]^T$.

The system is governed by

$$a(q^{-1})y(t) = B(q^{-1})u(t-1) + Ce(t) + D$$

where the numerical values of $a(z)$, $B(z)$, $C(z)$, and D are listed in Appendix C. The weighting matrices and the covariance of the noise is also included in Appendix C. We assume B_0 is known during the simulation.

Simulation results indicate that the system can indeed be stabilized along the targeted growth path. The input-output time responses of a typical run are shown in Figures 3.1-3.4. Figure 3.1b and Figure 3.2b shows the output responses with expanded ordinate after the algorithm has settled. We notice that there are extreme fluctuations during the start up. In practical applications, these may not be permitted and can be avoided by starting with initial estimates that yield a satisfactory response. See [18], for instance, for practical considerations.

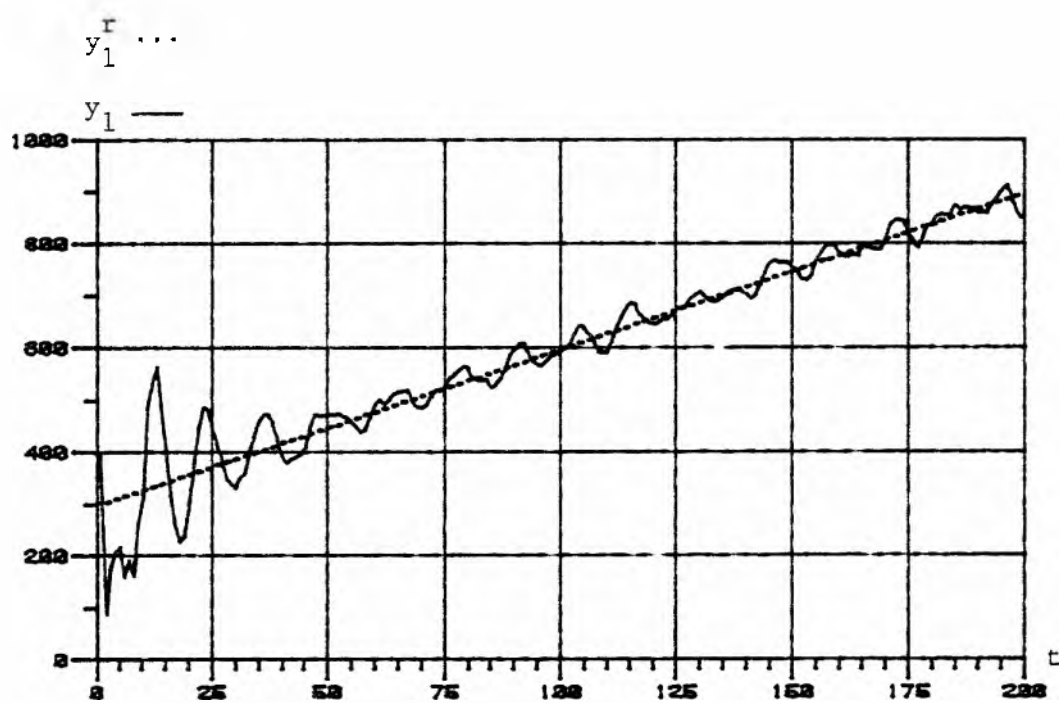


Figure 3.1a. Time response of y_1 for Nash game.

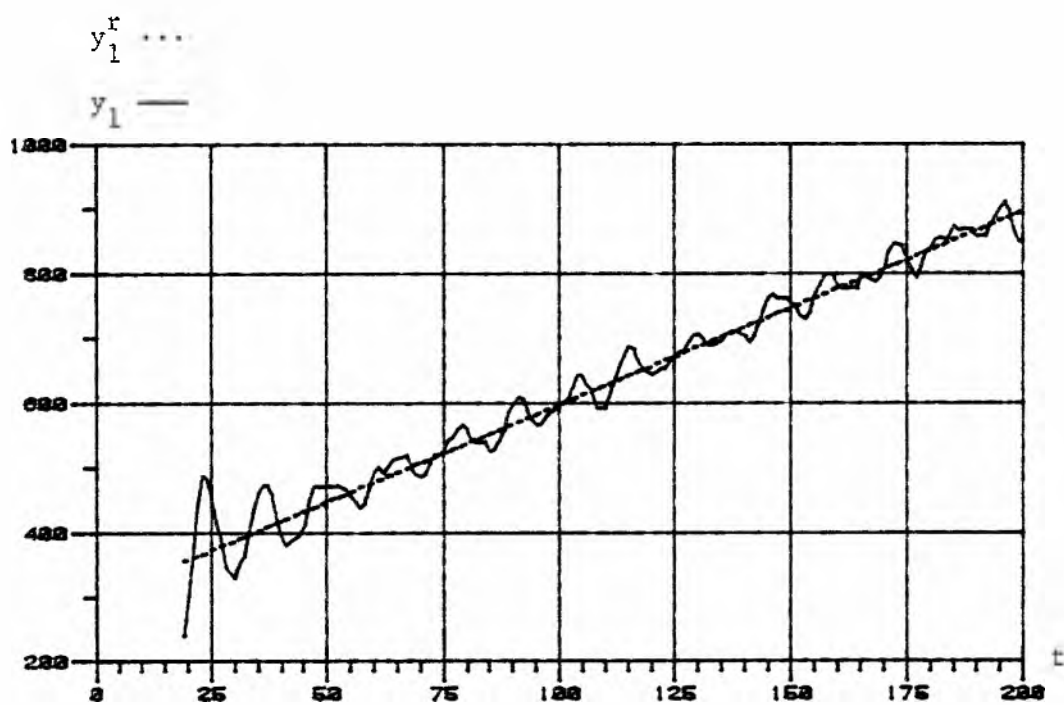


Figure 3.1b. Time response of y_1 for Nash game with expanded ordinate.

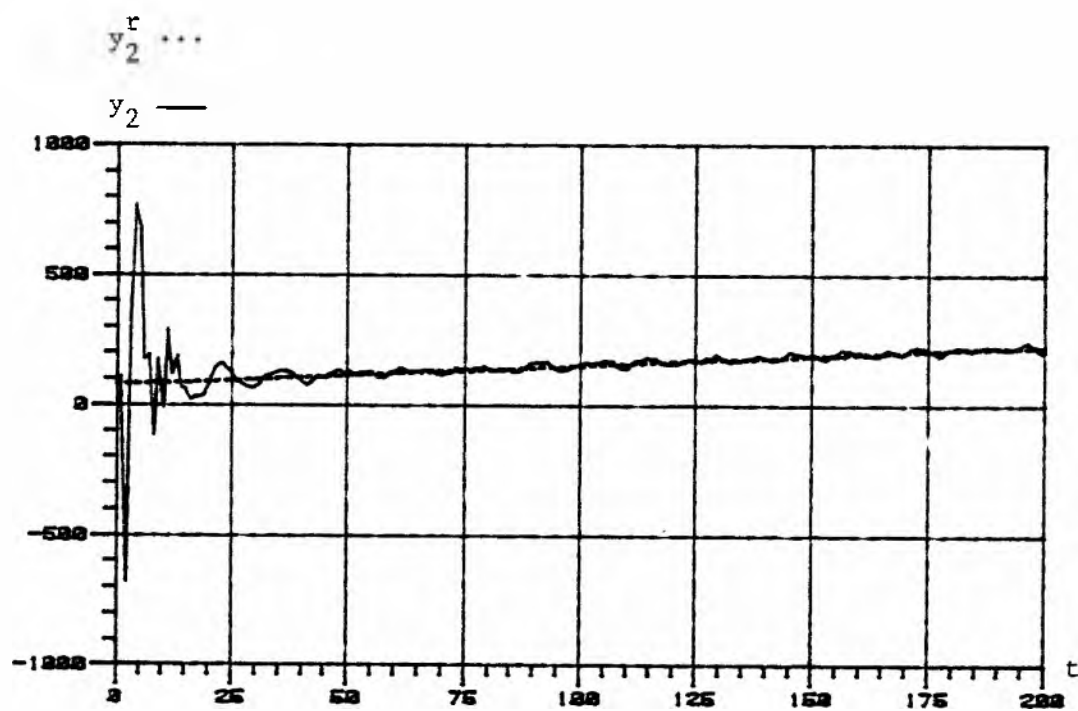


Figure 3.2a. Time response of y_2 for Nash game.

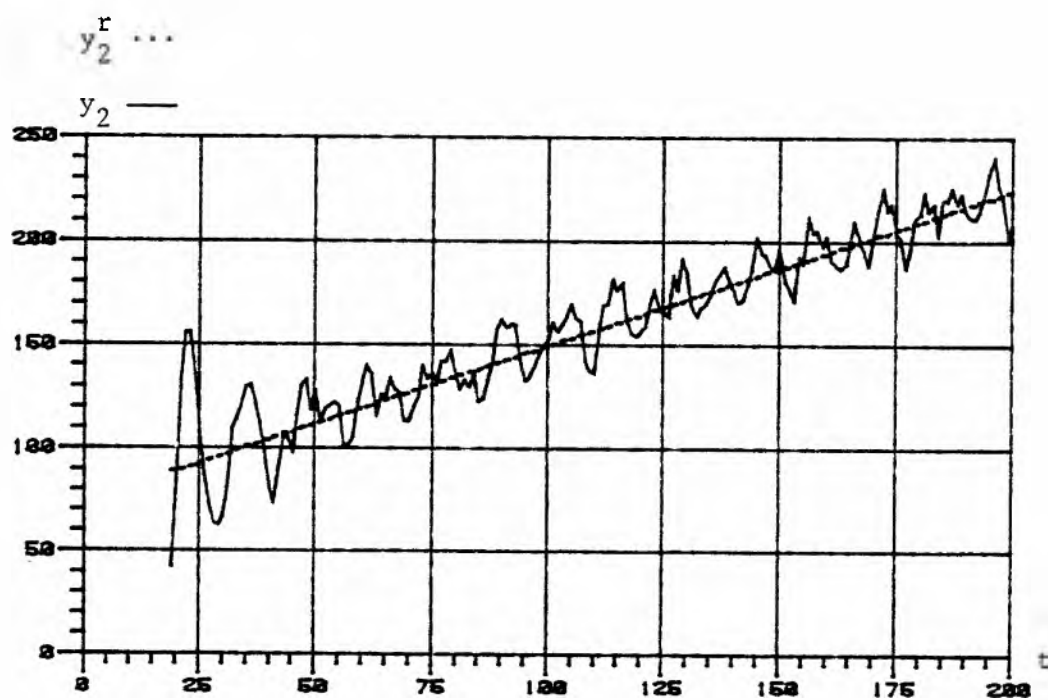


Figure 3.2b. Time response of y_2 for Nash game with expanded ordinate.

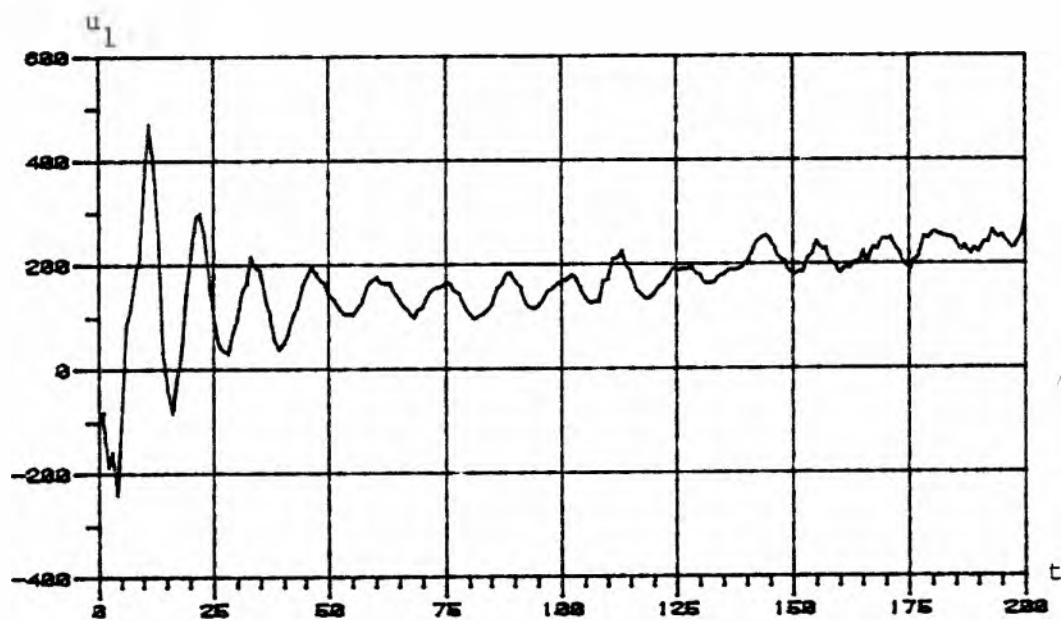


Figure 3.3. Time response of u_1 for Nash game.

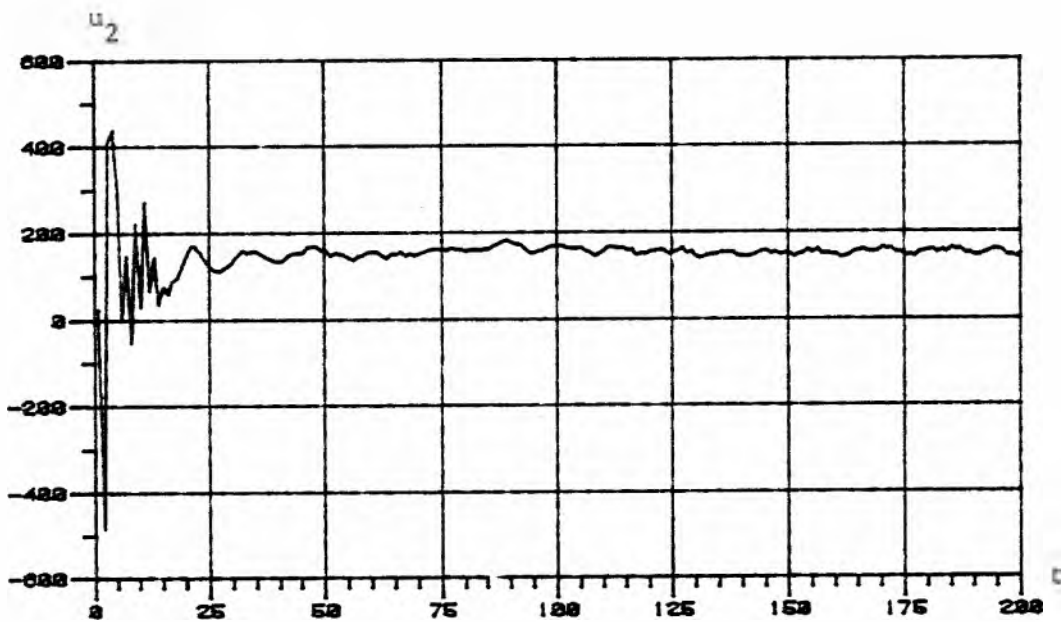


Figure 3.4. Time response of u_2 for Nash game.

CHAPTER 4

STOCHASTIC ADAPTIVE STACKELBERG GAMES

4.1 Introduction

In this chapter, the self-tuning principle will be called upon to solve the stochastic adaptive Stackelberg game problem. The Stackelberg game, or the Leader-Follower (L-F) game, was first introduced in the context of a static economic problem with two decision-makers [52]. It has been extended to dynamic cases in [49, 50, 51]. In the L-F game, one or more group of decision-makers, which will be called the follower. For information than the other group, which is called the follower. For instance, the leader knows the cost function of the follower but the follower may not know the leader's cost function. Equipped with the knowledge of the follower's cost function (thus the possible rational decision of the follower), the leader will perform his optimization taking into account the possible reaction of the follower. In the L-F game, the leader will announce his strategy first or act first. The follower then performs his optimization subject to his knowledge of the leader's action; that is, he is reacting to the leader's decision. Even though the computation for the leader may be more complicated than the Nash game case, he will do no worse, in terms of cost, and in general will do better using the L-F strategy rather than the Nash strategy. In general, however, nothing can be stated regarding the cost for the follower compared to his cost in the Nash game case. The L-F game framework is particularly appealing to optimization of hierarchical or multilevel systems where the follower or lower level controllers may have limited access to certain

information or they may have limited computing capability. In an economic system, for example, the government may be the leader over the business community because of its vast data base. Another example is in distributed control system in which the local process control computers may have limited computing capacity compared to the central computer.

There are a host of variations in the L-F game. For instance, the group of leader and/or follower may elect to use the Nash strategy instead of conforming to one single objective among their respective group. There may also be N groups of decision-makers with a hierarchical structure such that the higher level controller is a leader to the succeeding controller [13, 24]. In this report, we will concentrate on 2-Person Stackelberg games with DM_1 denoting the leader and DM_2 denoting the follower.

Definition 4.1. Let DM_1 , the leader, choose control $u_1 \in U_1$ and DM_2 , the follower, choose control $u_2 \in U_2$ where U_1 and U_2 are the sets of admissible controls. The cost function associated with DM_i is J_i ($i = 1, 2$). Assume there exists a mapping $T: U_1 \rightarrow U_2$. For each control u_1 chosen by DM_1 , DM_2 chooses $u_2 = T(u_1)$ such that

$$J_2(u_1, T(u_1)) \leq J_2(u_1, u_2) \quad , \quad \forall u_2 \in U_2 \quad . \quad (4.1)$$

The leader, DM_1 , chooses u_1^* such that

$$J_1(u_1^*, T(u_1^*)) \leq J_1(u_1, T(u_1)) \quad , \quad \forall u_1 \in U_1 \quad . \quad (4.2)$$

The strategy pair $(u_1^*, u_2^* = T(u_1^*))$ is called the Stackelberg equilibrium strategy pair.

As in the Nash game problem, there are different solution concepts to the L-F game. The solutions to open-loop, feedback, and closed-loop L-F games are in general different [21, 23, 44]. However, if we restrict the cost functions of the decision-makers to single-stage, we reduce the problem to a static one and circumvent the problem of different types of solution.

In Section 4.2, the L-F game problem will be formulated. The solution to the stochastic adaptive Stackelberg game problem is presented in Section 4.3. The same basic approach used in the analysis for the Nash game problem is found to be quite appropriate. A simulation example of an economic system is presented in Section 4.4.

4.2 Problem Formulation

Consider a system given by an input-output description

$$a(q^{-1})y(t) = B(q^{-1})u(t-k-1) + C(q^{-1})e(t) + D, \quad k \geq 0 \quad (4.3)$$

where $a(z)$, $B(z)$, $C(z)$ and D are as defined in (3.7). The vector $u(t) = [u_1^T(t) u_2^T(t)]^T$ with u_1 being the control of the leader and u_2 being the control of the follower. The output $y(t)$, input $u(t)$, and noise sequence $\{e(t)\}$ are all of dimension p . $\{e(t)\}$ is assumed to be an independently equally distributed zero mean random vector with finite covariance. The cost function associated with the i -th decision-maker is given by

$$J_i = E\{[y(t+k+1) - y^r(t+k+1)]^T Q_i [y(t+k+1) - y^r(t+k+1)] \\ + [u(t) - u(t-1)]^T R_i [u(t) - u(t-1)]\}, \quad i = 1, 2 \quad (4.4)$$

where y^r is the desired output of the system and Q_i and R_i are symmetric positive semidefinite matrices. The reason for the penalty of control change between each time step is, as mentioned before, to avoid the problem of calculating the reference control signal associated with non-zero reference output.

4.3 Self-Tuning Leader-Follower Game

The L-F equilibrium strategy will first be derived for the system assuming all the parameters are known. Then an adaptive scheme similar to the one used in the Nash game problem is presented to deal with unknown parameters. For ease of derivation, we will limit the controls u_1 and u_2 to be scalar valued. That is, we will consider a two input two output system in this chapter. The results, however, can easily be extended to vector valued controls.

4.3.1 Two-Person Leader-Follower Equilibrium Strategy

The derivation of the L-F equilibrium strategy from the necessary conditions is basically the same as that of the STC except for certain modifications. The result is summarized in the following theorem.

Theorem 4.1. Let $L^{(i)}$ represent the i -th column of a matrix L and let $(R)_{ij}$ denote the i,j -th entry of the matrix R . The Leader-Follower equilibrium strategy $u^*(t) = (u_1^*(t) u_2^*(t))^T$ for the system (4.3) with cost functions (4.4) is given by

$$\begin{aligned} & MG(q^{-1})C(q^{-1})y(t) + [MF(q^{-1})C(q^{-1})B(q^{-1}) \\ & + (1-q^{-1})\bar{c}(q^{-1})H]u^*(t) + MF(q^{-1})C(q^{-1})D \\ & - \bar{c}(q^{-1})My^r(t+k+1) = 0 \end{aligned} \quad (4.5)$$

with

$$\bar{C}(z) = \text{adjoint } C(z) \quad (4.6a)$$

$$\bar{c}(z) = \det C(z) \quad (4.6b)$$

$$G(z) = G_0 + G_1 z + \dots + G_{n-1} z^{n-1} \quad (4.6c)$$

$$F(z) = 1 + F_1 z + \dots + F_k z^k \quad (4.6d)$$

$$M = \begin{bmatrix} M_1 \\ M_2 \end{bmatrix} = \begin{bmatrix} (k_2 B_0^{(1)T} - k_1 B_0^{(2)T}) Q_1 \\ B_0^{(2)T} Q_2 \end{bmatrix} \quad (4.6e)$$

$$H = \begin{bmatrix} H_1 \\ H_2 \end{bmatrix} = \begin{bmatrix} k_2 R_1^{(1)T} - k_1 R_1^{(2)T} \\ R_2^{(2)T} \end{bmatrix} \quad (4.6f)$$

where

$$k_i = B_0^{(2)T} Q_2 B_0^{(1)} + (R_2)_{i2} \quad (4.6g)$$

and $G(z)$, $F(z)$ satisfy

$$C(z) = a(z)F(z) + z^{k+1}G(z) \quad (4.7)$$

Proof. See Appendix D.

It can readily be observed that the L-F strategy resembles the Nash strategy (3.8) and STC strategy (2.22) in every aspect except for the definition of the matrices M and H . Another remark is that if there has been no penalty on the control ($H = 0$) and assuming that M is non-singular, the L-F strategy will reduce to the minimum variance strategy for the STR

and this resulting L-F strategy is independent of the weights Q_i ($i = 1, 2$). That is, the game aspects of the problem will not arise if there is no penalty on the control action.

Notice that the leader will have to solve (4.5) to obtain the L-F strategy. The follower, on the other hand, will just need to go through part of (4.5) to obtain u_2^* since the leader acts first and thus u_1^* is known to the follower. Also, notice that the follower only needs to know M_2 and H_2 in order to solve u_2^* . The fact that the follower may not know the leader's cost function is again very transparent in this instance.

To utilize the self-tuning approach, the original system is transformed to an equivalent system as in the Nash game analysis. Let $\phi(t)$ be defined by

$$\phi(t) = M(y(t) - y^r(t)) + H(u(t-k-1) - u(t-k-2)) . \quad (4.8)$$

The transformed system in ϕ becomes

$$\begin{aligned} a(q^{-1})\phi(t) &= (MB(q^{-1}) + (1-q^{-1})a(q^{-1})H)u(t-k-1) \\ &\quad + MC(q^{-1})e(t) + MD - a(q^{-1})My^r(t) . \end{aligned} \quad (4.9)$$

If we define a cost function I given by

$$I = E\{\phi^T(t+k+1)\phi(t+k+1)\} , \quad (4.10)$$

it is possible, as in the Nash game case, to obtain the L-F equilibrium strategy by considering the transformed system as a minimum variance control problem with one decision-maker.

4.3.2 Two-Person Self-Tuning Stackelberg Strategy

In order to deal with unknown parameters, we assume all the decision-makers agree to use the same estimation scheme and identical initial conditions. The parameter B_0 is assumed to be known so that the matrix M can be computed.

To control the process (4.3) with cost functions given by (4.4), each decision-maker will use the following model as representation of the system

$$\begin{aligned} \phi(t) = & \mathcal{A}(q^{-1})y(t-k-1) + \mathcal{B}(q^{-1})u(t-k-1) \\ & + \mathcal{F}(q^{-1})y^r(t) + \mathcal{J} + \epsilon(t) \end{aligned} \quad (4.11)$$

where $\mathcal{A}(z)$, $\mathcal{B}(z)$, $\mathcal{F}(z)$ and \mathcal{J} are as defined in (3.14) and $\epsilon(t)$ is the error to be minimized in the least squares sense.

The certainty equivalent L-F strategy u^* is then given by

$$0 = \hat{\mathcal{B}}(q^{-1})u^*(t) + \hat{\mathcal{A}}(q^{-1})y(t) + \hat{\mathcal{F}}(q^{-1})y^r(t+k+1) + \hat{\mathcal{J}} \quad (4.12)$$

where \hat{L} denotes the estimate for L . The parameters can be estimated using stochastic approximation scheme given by (3.16)-(3.19).

To further appreciate the structure in the self-tuning L-F game, we will go into some interesting properties of this adaptive procedure. The leader in the game will have to estimate all the controller parameters in order to compute u^* . On the other hand, the follower, who acts after the leader has acted, has a simpler estimation computation. Specifically, the follower's estimation computation is part of the leader's computations. Let us elaborate by further considering the case where u_1 and u_2 are

scalar-valued. For simplicity, assume $D = 0$ and $y^r = 0$. Let $A(z)$, $B(z)$, $\phi(t)$ be given by

$$A(z) = \begin{bmatrix} a_{11}(z) & a_{12}(z) \\ a_{21}(z) & a_{22}(z) \end{bmatrix} \quad (4.13a)$$

$$B(z) = \begin{bmatrix} b_{11}(z) & b_{12}(z) \\ b_{21}(z) & b_{22}(z) \end{bmatrix} \quad (4.13b)$$

$$\phi(t) = \begin{bmatrix} \phi_1(t) \\ \phi_2(t) \end{bmatrix} = \begin{bmatrix} M_1 \\ M_2 \end{bmatrix} y(t) + \begin{bmatrix} H_1 \\ H_2 \end{bmatrix} u(t-k-1) \quad (4.14)$$

From (4.14) $\phi_2(t)$ is given by

$$\phi_2(t) = M_2 y(t) + H_2 u(t-k-1) \quad (4.15)$$

and M_2 , H_2 are functions of the follower's cost function only. The follower, in fact, only requires ϕ_2 for his controller. Consider the following equation which is part of (4.11)

$$\begin{aligned} \phi_2(t+k+1) - \epsilon_2(t+k+1) &= a_{21}(q^{-1})y_1(t) + a_{22}(q^{-1})y_2(t) \\ &\quad + b_{21}(q^{-1})u_1(t) + b_{22}(q^{-1})u_2(t) . \end{aligned} \quad (4.16)$$

The follower's optimal strategy u_2^* , with u_1^* available after the leader's action, is given by

$$\begin{aligned} \hat{b}_{22}(q^{-1})u_2^*(t) &= -\hat{a}_{21}(q^{-1})y_1(t) - \hat{a}_{22}(q^{-1})y_2(t) \\ &\quad - \hat{b}_{21}(q^{-1})u_1^*(t) . \end{aligned} \quad (4.17)$$

Hence, the follower will save some computational effort compared to the leader.

4.3.3 Convergence

Convergence results for the L-F game problem is exactly the same as stated in Theorem 3.2 except for the change in the definition of the matrices M and H .

Theorem 4.2. Let the matrices M and H be defined by (4.6e) and (4.6f) respectively and assume the conditions on the system in Theorem 3.2 are satisfied. If the self-tuning L-F strategy (4.12) is applied to the system (4.3), then the system input-output will remain bounded and the prediction error for $\phi(t)$ will tend to its global minimum achievable by any causal linear feedback with probability one.

4.4 Simulation Study

The economic model presented in Section 3.4 is again used to study the performance of the self-tuning strategy. In this case, we assume the federal government (Congress and the President) to be the leader and the FRS as the follower. The same Q_i and R_i used in the Nash game simulation are used in this case.

Simulation results indicate that the algorithm can indeed stabilize the system along the targeted 1% quarterly growth. The input-output time responses of a typical run are shown in Figures 4.1-4.4. Figure 4.1b and Figure 4.2b shows the output responses with expanded ordinate after the algorithm has settled. Again, there are extreme fluctuations during the start up period as the controller is trying to learn the characteristics

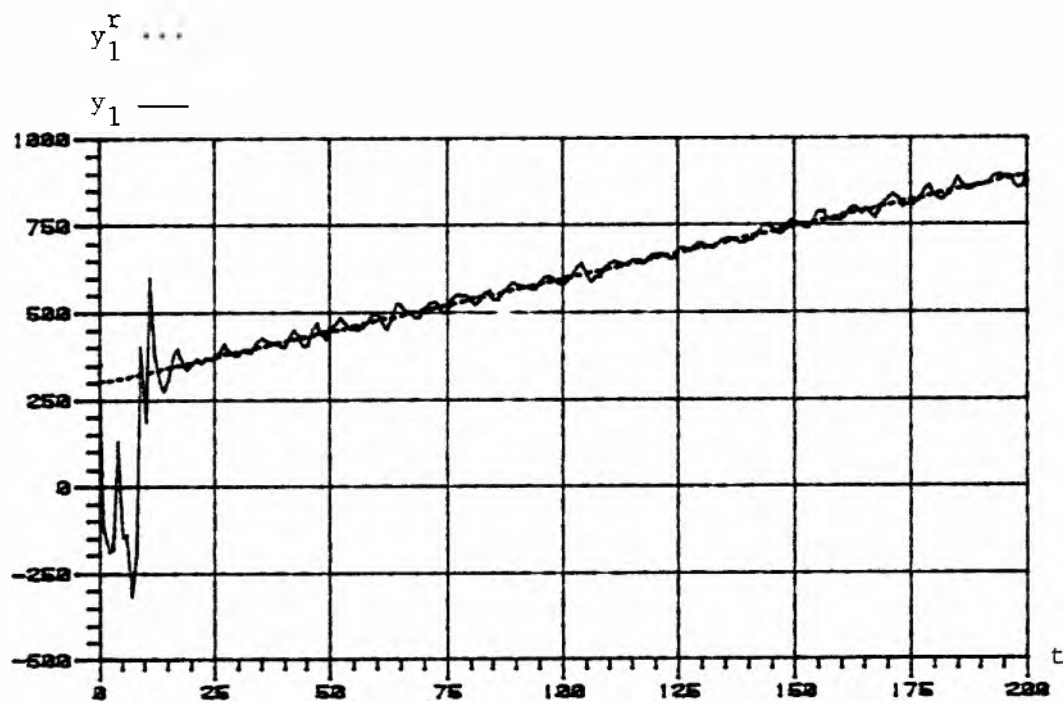


Figure 4.1a. Time response of y_1 for Stackelberg game.

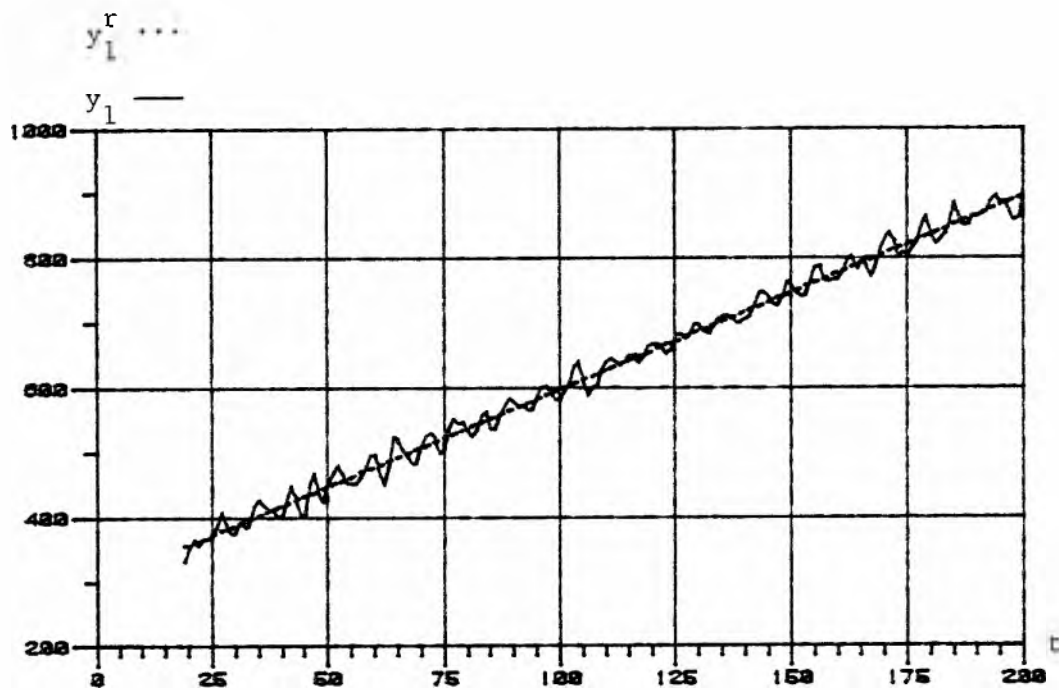


Figure 4.1b. Time response of y_1 for Stackelberg game with expanded ordinate.

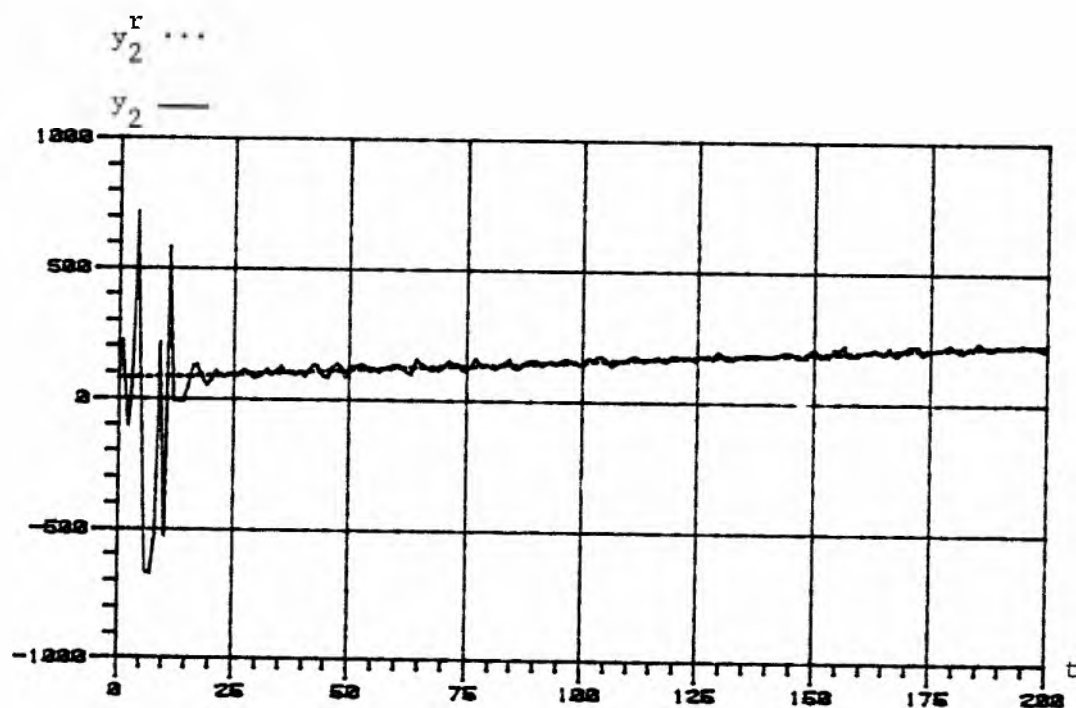


Figure 4.2a. Time response of y_2 for Stackelberg game.

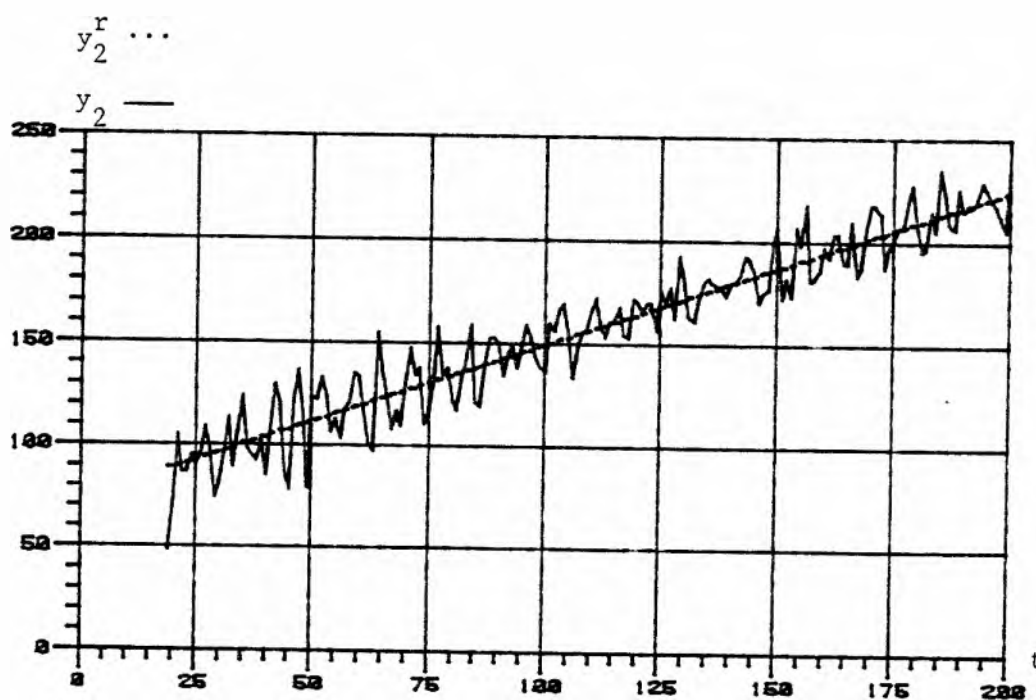


Figure 4.2b. Time response of y_2 for Stackelberg game with expanded ordinate.

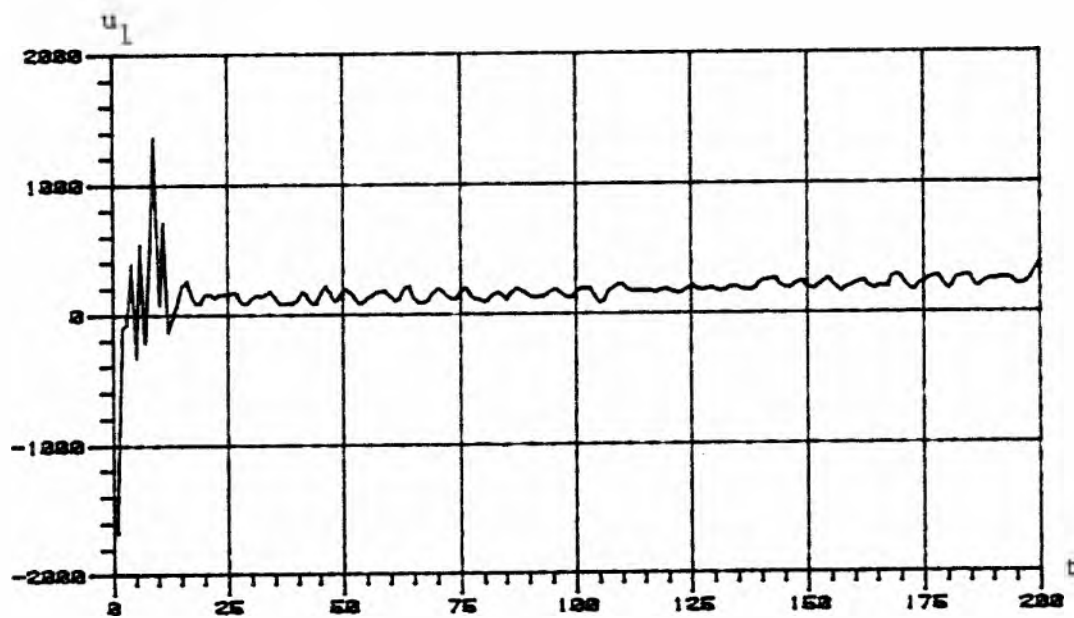


Figure 4.3. Time response of u_1 for Stackelberg game.

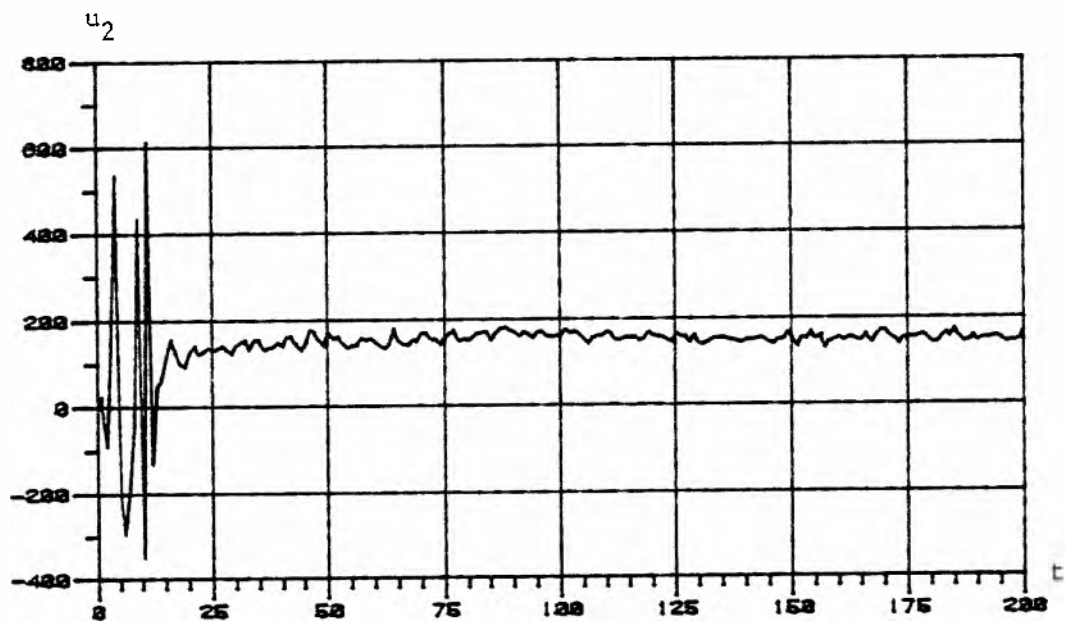


Figure 4.4. Time response of u_2 for Stackelberg game.

of the process. These excessive fluctuations may be undesirable and may be reduced as discussed earlier in Section 3.4.

CHAPTER 5

DECENTRALIZED STOCHASTIC ADAPTIVE NASH GAMES

5.1 Introduction

In this chapter, an explicit self-tuning method is utilized to develop an algorithm for systems with unknown parameters and multiple controllers each, besides having a different objective, has a different set of information about the system. This decentralized system framework is suitable for analyzing large scale interconnected systems in which the communication and/or computational costs involved may prohibit the implementation of a centralized control policy. Decentralized information among decision-makers was first studied in the framework of static team theory in [48] and was further extended in [11, 12, 34, 35, 36].

In this report, we will confine our analysis to Two-Person decentralized stochastic adaptive Nash games with an information structure termed "one-step-delay sharing pattern" [57]. We will restrict the cost function of each decision-maker to a single-stage, thus, turning the dynamic situation into a static Nash game framework. In Section 5.2, the formulation of the decentralized Nash game problem is presented. In Section 5.3, we approach the known parameter problem by a straightforward constraint on the form of the control policy as done similarly for the single controller problem in [32, 37, 41]. In Section 5.4, another approach is used to tackle the known parameter decentralized game problem. Specifically, we extend results of static Nash games in [11, 12] to our problem.

To deal with the unknown parameter case, a recursive estimator is used to determine the system parameters explicitly. We force the certainty equivalence condition upon the system and substitute the true parameters by the estimates into the control law. The proposed algorithm can be classified as an explicit self-tuning strategy as the systems parameters are estimated explicitly and then manipulated to determine the optimal policy. Even though convergence for this procedure is not guaranteed, our simulation studies for an economic system, which is presented in Section 5.5, do show that the algorithm is capable of stabilizing the system along a desired path asymptotically. Furthermore, our simulation results indicate that the two different decentralized approaches will generate the same optimal policy hinting that the two methods may actually be equivalent.

5.2 Problem Formulation

Consider a system with multiple decision-makers each has u_i ($i = 1, 2, \dots, N$) as his control. The system is governed by

$$y(t+1) = a(q^{-1})y(t) + B(q^{-1})u(t) + e(t+1) + D \quad (5.1)$$

where $u(t)$ is formed by stacking up all the $u_i(t)$. The dimension of the i -th component of y , y_i , is assumed to be of the same dimension as u_i . The sequences $\{y(t)\}$, $\{u(t)\}$, $\{e(t)\}$ are all of dimension p . The disturbance sequence $\{e(t)\}$ is an independent identically distributed zero mean white noise with finite covariance given by $E\{e(t)e^T(t)\} = W$. $B(z)$ is a matrix polynomial and $a(z)$ is a scalar polynomial as given by

$$a(z) = a_0 + a_1 z + \dots + a_n z^n \quad (5.2)$$

$$B(z) = B_0 + B_1 z + \dots + B_n z^n. \quad (5.3)$$

The D in (5.1) is an p -dimensional offset vector with i -th row block D_i , which is also the same dimension as u_i . A steady state decentralized Nash equilibrium strategy for the system is to be sought. The cost function of each decision-maker is given by

$$\begin{aligned} J_i^0 = E\{[y_i(t+1) - y_i^r(t+1)]^T Q_i [y_i(t+1) - y_i^r(t+1)] \\ + [u_i(t) - u_i(t-1)]^T R_i [u_i(t) - u_i(t-1)]\}, \\ i = 1, 2, \dots, N \quad (5.4) \end{aligned}$$

where Q_i is symmetric positive semidefinite, R is symmetric positive definite and y_i^r is the desired value of the i -th output y_i .

In our problem, at every step of time t , the i -th decision-maker is assumed to have: i) $y_i(t)$ and past outputs $y(t-1), y(t-2), \dots$; ii) past inputs $u(t-1), u(t-2), \dots$ as his information. This class of information pattern is called "one-step-delay sharing pattern" [57]. The i -th decision-maker attempts, under this information structure, to minimize (5.4) with respect to $u_i(t)$ with the assumption that the other decision-makers use the Nash equilibrium strategy as well. In this report, the number of decision-makers is limited to two. However, the algorithm can easily be generated for more than two controllers once the methodology of the solution is understood.

To facilitate our analysis, the cost functional (5.4) will be decomposed to a form in which only the part that directly affects the

optimization result is kept. It can be shown by straightforward substitution that (5.4) can be written in the following form:

$$\begin{aligned}
 J_i^0 = & E\{u_i^T(t)D_{ii}u_i(t) + 2u_i^T(t)D_{ij}u_j(t) \\
 & + 2a_0y_i^T(t)Q_i(B_0)_{ii}u_i(t) + 2[\tilde{a}(q^{-1})y_i(t) \\
 & + \tilde{B}_{ij}(q^{-1})u_j(t) + \tilde{B}_{ii}(q^{-1})u_i(t) \\
 & + D_i - y_i^T(t+1)]^T Q_i(B_0)_{ii}u_i(t) \\
 & + 2[u_i^T(t-1)R_i u_i(t)]\} \\
 & + \text{terms not involving } u_i(t), \quad i, j = 1, 2 \\
 & \quad \quad \quad i \neq j
 \end{aligned} \tag{5.5}$$

with

$$D_{ii} = (B_0)_{ii}^T Q_i (B_0)_{ii} + R_i \tag{5.6}$$

$$D_{ij} = (B_0)_{ii}^T Q_i (B_0)_{ij} \tag{5.7}$$

and $(B_0)_{ij}$ denotes the i, j -th block of the zero-th order element, B_0 , of the matrix polynomial $B(z)$. $\tilde{B}_{ij}(z)$ denotes the i, j -th block of $B(z)$ with $(B_0)_{ij}$ taken out, that is,

$$\tilde{B}_{ij}(z) = B_{ij}(z) - (B_0)_{ij} . \tag{5.8}$$

The scalar polynomial $\tilde{a}(z)$ is similarly defined as

$$\tilde{a}(z) = a(z) - a_0 . \tag{5.9}$$

We will let J_i denote the "active" part of J_i^0 in (5.5), that is, the part that involves $u_i(t)$. Hence, we have

$$J_i^0 = J_i + \text{terms not involving } u_i(t) . \tag{5.10}$$

5.3 Constrained Decentralized Nash Game

Consider a system governed by (5.1) in which each controller has a cost function given by

$$\begin{aligned}
 J_i = E \{ & u_i^T(t) D_{ii} u_i(t) + 2u_i^T(t) D_{ij} u_j(t) \\
 & + 2a_0 y_i^T(t) Q_i(B_0)_{ii} u_i(t) + 2[\tilde{a}(q^{-1}) y_i(t) \\
 & + \tilde{B}_{ii}(q^{-1}) u_i(t) + \tilde{B}_{ij}(q^{-1}) u_j(t) \\
 & + D_i - y_i^T(t+1)]^T Q_i(B_0)_{ii} u_i(t) \\
 & + 2u_i^T(t-1) R_i u_i(t) \} , \quad i, j = 1, 2 \quad i \neq j
 \end{aligned} \tag{5.11}$$

Let the matrix C_i be defined by

$$C_i = a_0 Q_i(B_0)_{ii} , \tag{5.12}$$

and let the function x_i be defined by

$$\begin{aligned}
 x_i(t) = & (B_0)_{ii}^T Q_i [\tilde{a}(q^{-1}) y_i(t) + \tilde{B}_{ii}(q^{-1}) u_i(t) \\
 & + \tilde{B}_{ij}(q^{-1}) u_j(t) + D_i - y_i^T(t+1)] \\
 & + R_i u_i(t-1) , \quad i = 1, 2 .
 \end{aligned} \tag{5.13}$$

Notice that at time t , the value of $x_i(t)$ is known as it does not depend on any future data.

Now we can rewrite J_i as

$$\begin{aligned}
 J_i = E \{ & u_i^T(t) D_{ii} u_i(t) + 2u_i^T(t) D_{ij} u_j(t) + 2y_i^T(t) C_i u_i(t) \\
 & + 2x_i^T(t) u_i(t) \} , \quad i, j = 1, 2 \quad i \neq j
 \end{aligned} \tag{5.14}$$

The constrained decentralized Nash equilibrium strategy for the system (5.1) with cost functional J_i in (5.14) is first presented in Section 5.3.1 assuming all the system parameters are known. Then, the certainty equivalence is invoked heuristically and a stochastic approximation type estimation scheme is used to obtain estimates that are substituted into the optimal policy in place of the true parameters.

5.3.1 Nash Game with Constrained Policy

Let the control u_i of the i -th decision-maker be of dimension m_i , $i = 1, 2$. Thus, the associated output measurement y_i for the i -th controller is also m_i -dimensional. The i -th decision-maker tries to minimize J_i with respect to u_i which is of the form

$$u_i(t) = G_i y_i(t) + g_i, \quad i = 1, 2 \quad (5.15)$$

where G_i is a $m_i \times m_i$ matrix and g_i is a m_i -dimensional vector. The constrained policy is stated in the following theorem.

Theorem 5.1. Let the characteristic root of a matrix A with maximum absolute value be denoted by $\lambda_m(A)$, then the condition

$$|\lambda_m(D_{11}^{-1}D_{12}D_{22}^{-1}D_{21})| < 1 \quad (5.16)$$

is sufficient for the system (5.1) with cost functions (5.14) to admit a unique Nash solution. The gains G_i, g_i of (5.15) satisfy the following

$$\begin{aligned} G_i &= D_{ii}^{-1}D_{ij}D_{jj}^{-1}D_{ji}G_jW_{ij}W_{jj}^{-1}W_{ji}W_{ii}^{-1} \\ &= -D_{ii}^{-1}C_i^T + D_{ii}^{-1}D_{ij}D_{jj}^{-1}C_jW_{ji}W_{ii}^{-1}, \quad i, j = 1, 2 \\ &\quad i \neq j \end{aligned} \quad (5.17)$$

$$g_i + D_{ii}^{-1} D_{ij} g_j = -D_{ii}^{-1} [D_{ii} G_i \bar{y}_i(t) + D_{ij} G_j \bar{y}_j(t) + C_i^T \bar{y}_i(t) + x_i(t)] , \quad \begin{matrix} i, j = 1, 2 \\ i \neq j \end{matrix} \quad (5.18)$$

where \bar{y}_i denotes the expectation of y_i and W_{ij} is the i, j -th block of the noise covariance matrix W .

Proof. See Appendix E.

Notice that the gain G_i is independent on a_0 , B_0 , Q_i and R_i only. Hence, in the case when a_0 , B_0 parameters are known, once G_i is determined, it does not require further computation.

5.3.2 Self-Tuning Constrained Decentralized Nash Game

In order to obtain the Nash strategy for the system with unknown parameters, we propose an ad hoc method of certainty equivalence. In this procedure we will assume a_0 and B_0 are known to avoid possibilities of non-existence of solutions for (5.17). Furthermore, we will allow a unit delay in the estimation scheme, that is, at time t , the system parameter estimates used for the control computation are based on past input-output data only. In addition, we assume each decision-maker uses the same estimation scheme and initial conditions so that the problem of multi-modelling can be avoided.

The recursive procedure in [40], which is a stochastic approximation type algorithm, can be used to estimate the system parameters explicitly. Introduce the parameter matrix Θ defined by

$$\Theta = [\theta_1 \theta_2 \dots \theta_p] = \begin{bmatrix} \mathcal{A}_0 \\ \mathcal{A}_1 \\ \vdots \\ \mathcal{A}_n \\ \mathcal{B}_0^T \\ \vdots \\ \mathcal{B}_n^T \\ \mathcal{C}_1^T \dots \mathcal{C}_p^T \end{bmatrix} \quad (5.19)$$

where each \mathcal{A}_i ($i = 0, 1, 2, \dots, n$) is a diagonal matrix. The following recursions are carried out at each step of time to estimate θ_j , $j = 1, 2, \dots, p$:

$$\theta_j(t) = \theta_j(t-1) + \frac{\gamma(t)}{r_j(t)} \eta_j(t-1) [y_j(t) - \eta_j^T(t-1) \theta(t-1)] \quad (5.20)$$

$$r_j(t) = r_j(t-1) + \gamma(t) [\eta_j^T(t-1) \eta_j(t-1) - r_j(t-1)] , \quad (5.21)$$

$$r_j(0) = 1$$

$$\eta_j(t-1) = [y_j^T(t-1) \dots y_j^T(t-n) u^T(t-1) \dots u^T(t-n) 1]^T \quad (5.22)$$

with $\gamma(t)$ being a decreasing sequence in t . Notice that the assumptions a_0, B_0 are known and will lead to the setting of $\mathcal{A}_0 = a_0 I$ and $\mathcal{B}_0 = B_0$. A block diagram of the closed-loop system is shown in Figure 5.1.

Convergence of the estimator will certainly lead to the convergence of the Nash strategy. The condition for convergence for the filter equations (5.20) and (5.21) has been investigated [40]. It is shown that

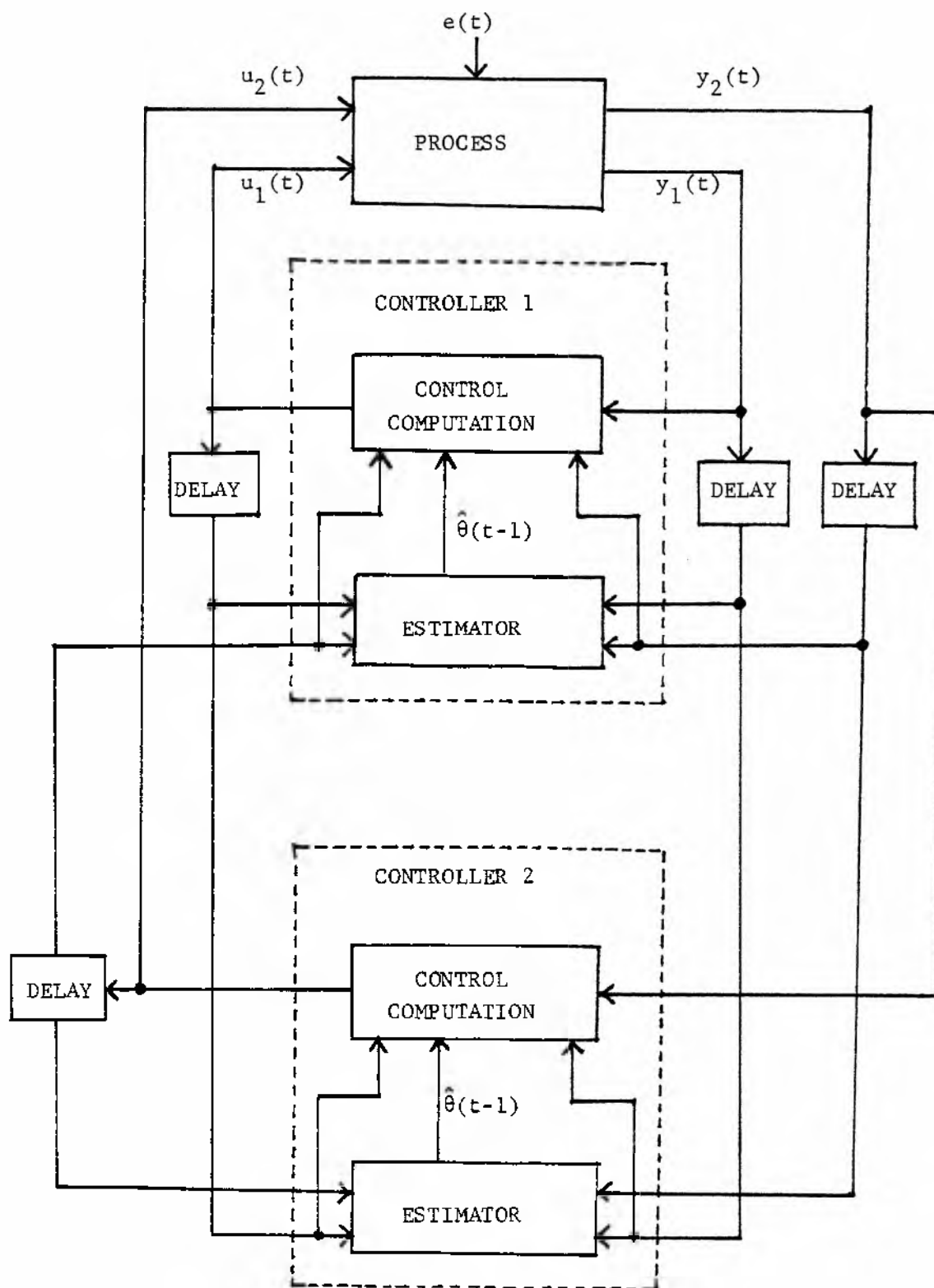


Figure 5.1. Decentralized Nash game with constrained policy.

if $u(t)$ is a white noise process, then the estimates will yield a correct description of the input-output data. Conceivably, in a multivariable system, there may be different sets of estimates that yield the same description of the system. Hence, suitable identifiability condition of the system is required to ensure proper convergence of the adaptive scheme. Identifiability conditions for multivariable systems has been investigated in [30, 55].

5.4 Extended Static Decentralized Nash Game

In this section, we will solve the known parameter case by applying results in [11, 12] to our present problem. Then the estimation scheme (5.18)-(5.22) is used to obtain explicitly the system parameters which are then substituted into the optimal policy derived from the known parameter case.

Before utilizing the results in static Nash games, reformulation of the problem into the appropriate setting is required. The cost function in (5.11) is rewritten in the form

$$J_i = E\{u_i^T(t)D_{ii}u_i(t) + 2u_i^T(t)D_{ij}u_j(t) + 2x^T(t)C_{i1}^T u_i(t)\}, \quad (5.23)$$

$$\begin{aligned} i, j &= 1, 2 \\ i &\neq j \end{aligned}$$

with D_{ij} as defined in (5.6) and (5.7) and

$$C_1^T = \begin{matrix} \overbrace{\begin{bmatrix} I \\ 0 \end{bmatrix}}^{m_1} \end{matrix} \left. \begin{matrix} \end{matrix} \right\}^{m_1} \quad (5.24a)$$

$$C_2^T = \begin{matrix} \begin{bmatrix} 0 \\ I \end{bmatrix} \end{matrix} \left. \begin{matrix} \end{matrix} \right\}^{m_2} \quad (5.24b)$$

where m_i = dimension u_i and

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \quad (5.25)$$

where

$$\begin{aligned} x_i(t) = & (B_0)_{ii}^T Q_i [a_0 y_i(t) + \tilde{a}(q^{-1}) y_i(t) + \tilde{B}_{ii}(q^{-1}) u_i(t) \\ & + \tilde{B}_{ij}(q^{-1}) u_j(t) + D_i - y_i^r(t+1)] + R_i u_i(t-1), \end{aligned} \quad (5.26)$$

$i, j = 1, 2$
 $i \neq j$

Notice that in $x_i(t)$, the term $(B_0)_{ii}^T Q_i a_0 y_i(t)$ is the current measurement that is available only to the i -th decision-maker at time t . The other terms in $x(t)$ are dependent on past input-output data that are available to every decision-maker under the "one-step-delay sharing pattern". Hence, $x_i(t)$ can be considered as information that is privileged to the i -th controller only.

5.4.1 Extended Static Decentralized Nash Solution

Consider $x(t)$ as the state vector in a state-space representation of a system in which the i -th controller has $z_i(t)$ as his measurement. The measurement $z_i(t)$ is given by

$$z_i(t) = H_i x(t) + v_i(t), \quad i = 1, 2 \quad (5.27)$$

where

$$H_1 = \left[\begin{array}{c|c} \overbrace{I}^{m_1} & \overbrace{0}^{m_2} \end{array} \right] m_1 \quad (5.28a)$$

$$H_2 = \left[\begin{array}{c|c} \overbrace{0}^{m_1} & \overbrace{I}^{m_2} \end{array} \right] m_2 \quad (5.28b)$$

and $v_i(t)$ is zero mean white noise with positive semidefinite covariances T_i , $i = 1, 2$. To utilize results in [12], the mean value of $x(t)$, $\bar{x}(t)$ and the covariance of $x(t)$, $\text{cov}\{x(t)\}$ are also required. We illustrate here how $\bar{x}(t)$ and $\text{cov}\{x(t)\}$ can be computed. At time t , past input-output data are known, thus

$$\begin{aligned} \bar{x}_i(t) &= E\{x_i(t)\} \\ &= (B_0)_{ii}^T Q_i [a_0 \bar{y}_i(t) + \tilde{a}(q^{-1}) y_i(t) + \bar{B}_{ii}(q^{-1}) u_i(t) \\ &\quad + \bar{B}_{ij}(q^{-1}) u_j(t) + D_i - y_i^r(t+1)] + R_i u_i(t-1), \\ &\quad i, j = 1, 2 \\ &\quad i \neq j \end{aligned} \quad (5.29)$$

where $\bar{y}_i(t) = E\{y_i(t)\}$ and \bar{y}_i is given by

$$\begin{aligned} \bar{y}_i(t) = & a(q^{-1})y_i(t-1) + B_{ii}(q^{-1})u_i(t-1) \\ & + B_{ij}(q^{-1})u_j(t-1) + D_i, \quad i, j = 1, 2, \\ & i \neq j \end{aligned} \quad (5.30)$$

Let the $\text{cov}\{x(t)\} = Q$, then

$$Q = E\{[x(t) - \bar{x}(t)][x(t) - \bar{x}(t)]^T\}$$

or

$$Q = E \left\{ \begin{bmatrix} (B_0)_{11}^T Q_1 a_0 (y_1(t) - \bar{y}_1(t)) \\ (B_0)_{22}^T Q_2 a_0 (y_2(t) - \bar{y}_2(t)) \end{bmatrix} \begin{bmatrix} (B_0)_{11}^T Q_1 a_0 (y_1(t) - \bar{y}_1(t)) \\ (B_0)_{22}^T Q_2 a_0 (y_2(t) - \bar{y}_2(t)) \end{bmatrix}^T \right\}$$

or

$$Q = a_0^2 \begin{bmatrix} (B_0)_{11}^T Q_1 W_{11} Q_1 (B_0)_{11} & (B_0)_{11}^T Q_1 W_{12} Q_2 (B_0)_{22} \\ (B_0)_{22}^T Q_2 W_{21} Q_1 (B_0)_{11} & (B_0)_{22}^T Q_2 W_{22} Q_2 (B_0)_{22} \end{bmatrix} \quad (5.31)$$

where W_{ij} denotes the i, j -th block of the noise covariance matrix W and the fact that

$$E\{(y_i(t) - \bar{y}_i(t))(y_j(t) - \bar{y}_j(t))^T\} = W_{ij} \quad (5.32)$$

has been used.

We are now ready to apply the result in [12, Theorem 2] to our game problem.

Theorem 5.2. The condition (5.16) in Theorem 5.1,

$$|\lambda_m(D_{11}^{-1} D_{12} D_{22}^{-1} D_{21})| < 1$$

is sufficient for the system (5.1) with cost functions (5.23) to admit a unique Nash solution. The control law of each decision-maker is given by

$$u_i(t) = G_i \bar{x}(t) + F_i (\hat{x}_i(t) - \bar{x}(t)) , \quad i = 1, 2 \quad (5.33)$$

with

$$G_i = -[I - D_{ii}^{-1} D_{ij} D_{jj}^{-1} D_{ji}]^{-1} D_{ii}^{-1} [C_i - D_{ij} D_{jj}^{-1} C_j] , \quad (5.34)$$

$$\begin{aligned} i, j &= 1, 2 \\ i &\neq j \end{aligned}$$

$$\begin{aligned} \hat{x}_i(t) &= E\{x(t) | z_i(t)\} \\ &= \bar{x}(t) + QH_i^T (H_i QH_i^T + T_i)^{-1} (z_i(t) - H_i \bar{x}(t)) , \\ &\quad i = 1, 2 \end{aligned} \quad (5.35)$$

and F_1 is the unique solution to

$$F_1 + P F_1 L = M \quad (5.36)$$

where

$$P = -D_{11}^{-1} D_{12} D_{22}^{-1} D_{21} \quad (5.37)$$

$$L = QH_1^T (H_1 QH_1^T + T_1)^{-1} H_1 QH_2^T (H_2 QH_2^T + T_2)^{-1} H_2 \quad (5.38)$$

$$M = -D_{11}^{-1} C_1 + D_{11}^{-1} D_{12} D_{22}^{-1} C_2 QH_2^T (H_2 QH_2^T + T_2)^{-1} H_2 \quad (5.39)$$

and

$$F_2 = -D_{22}^{-1} D_{21} F_1 QH_1^T (H_1 QH_1^T + T_1)^{-1} H_1 - D_{22}^{-1} C_2 . \quad (5.40)$$

Proof. See proof of Theorem 2 in [12].

There is a close resemblance of Theorems 5.1 and 5.2. In both cases, the control $u_i(t)$ is affine in $y_i(t)$. Moreover, as in Theorem 5.1, the gains for the information, F_i , is dependent upon the system parameters a_0 and B_0 but not the rest of $a(z)$ or $B(z)$.

Notice that in the formulation of the problem in (5.27), white noise sequence $\{v_i(t)\}$ have been introduced into the system. We may consider this disturbance as measurement error of $y_i(t)$ and/or noise in transmitting past input-output data. On the other hand, if no such noise is allowed into the system ($T_i = 0$), it is, intuitively, reasonable to expect Theorems 5.1 and 5.2 to generate the same optimal policy. In fact, the presence of a positive definite T_i during the formulation stage is to ensure that the matrices $(H_i Q H_i^T + T_i)$, $i = 1, 2$ are non-singular. If we assume Q is positive definite, by nature of the definition of H_i , the problem of singularity can be avoided.

5.4.2 Self-Tuning Extended Static Decentralized Nash Games

In order to avoid the possibility of non-existence of solutions, we assume the system parameters a_0 and B_0 are known while the rest of $a(z)$, $B(z)$ and D are unknown. As in the previous approach, the system parameters are estimated recursively using equations (5.19)-(5.22) assuming identical algorithm and initial conditions for all decision-makers. The estimates are then substituted into the equations of Theorem 5.2 in place of the true parameters to obtain the optimal strategy. Hence, convergence of this Nash policy depends on the convergence of the estimates, as commented previously in Section 5.3.2.

A block diagram of the closed-loop system is shown in Figure 5.2. It is obvious that the two different approaches are almost identical

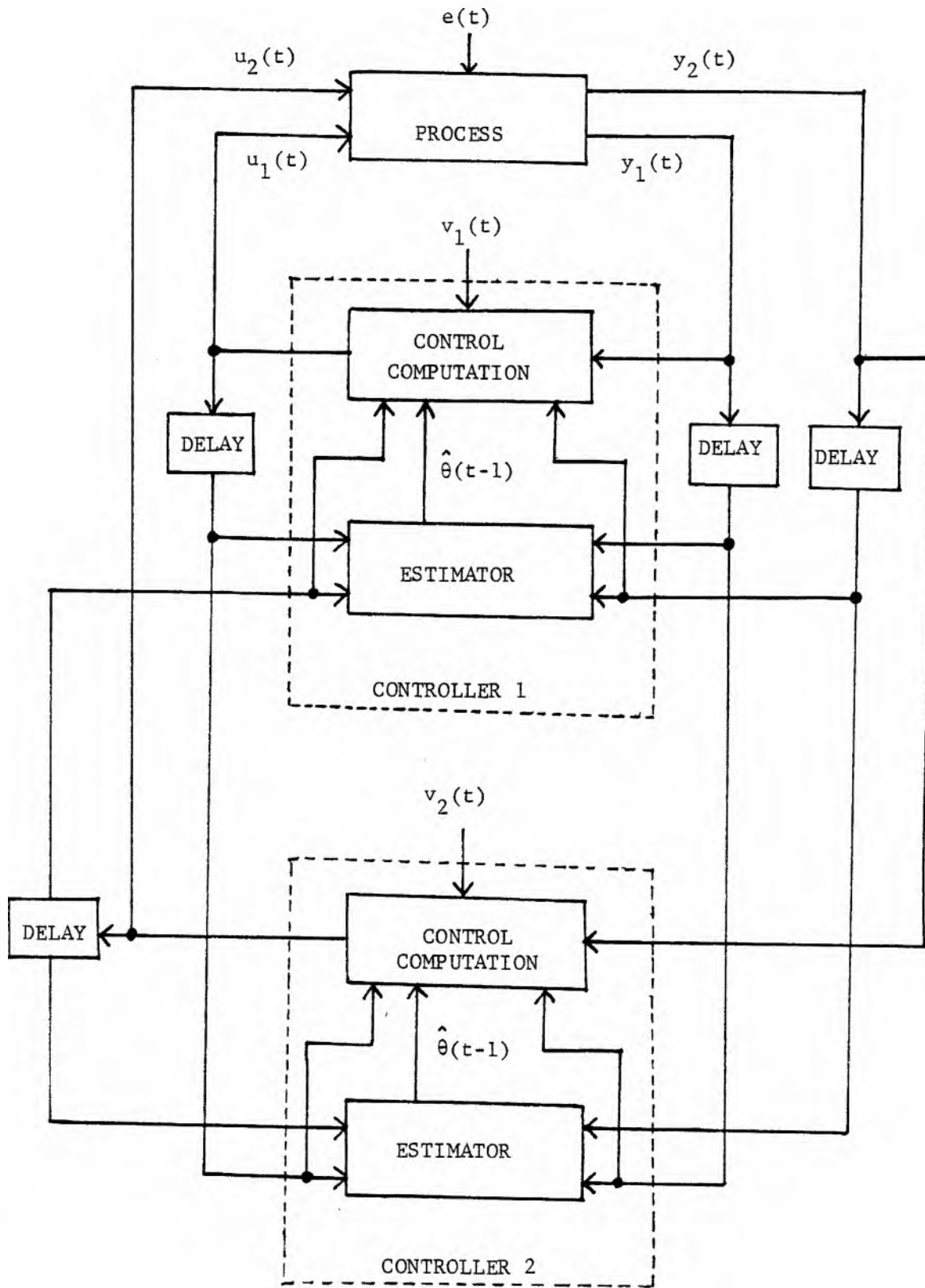


Figure 5.2. Extended static decentralized Nash game.

except for the noise sequences $\{v_i(t)\}$ ($i = 1, 2$) that are introduced in the second method.

5.5 Simulation Studies and Conclusion

The economic system presented in Section 3.4 is used to demonstrate the performance of the algorithm. We assume the government who controls $u_1(t)$ has the consumption expenditure $y_1(t)$ as its measurement and the Federal Reserve System who controls $u_2(t)$ has the private investment $y_2(t)$ as its measurement. Although this phenomenon may not be entirely realistic, we can interpret this case as a situation in which the government places a strong emphasis on the current consumption while the FRS focuses its entire energy on ensuring the targeted path of current investment is followed. The cost functions weighting matrices are given by

$$\begin{aligned} Q_1 &= 5 & R_1 &= 0.02 \\ Q_2 &= 10 & R_2 &= 0.08 \end{aligned}$$

The same noise covariance used in previous simulations is used. The same nominal y^r in previous cases is used.

Simulation results indicate that both adaptive procedures can indeed stabilize the system along the targeted growth path. The input-output time responses of a typical run using the extended static Nash game approach (with $T_i = 0$) are shown in Figures 5.3a-5.6a and the corresponding trajectories using the constrained policy approach are shown alongside in Figures 5.3b-5.6b. The two sets of input-output responses indicate the two methods generate exactly the same optimal policy. Hence, it is reasonable to assume the two methods are equivalent. There may be situations in

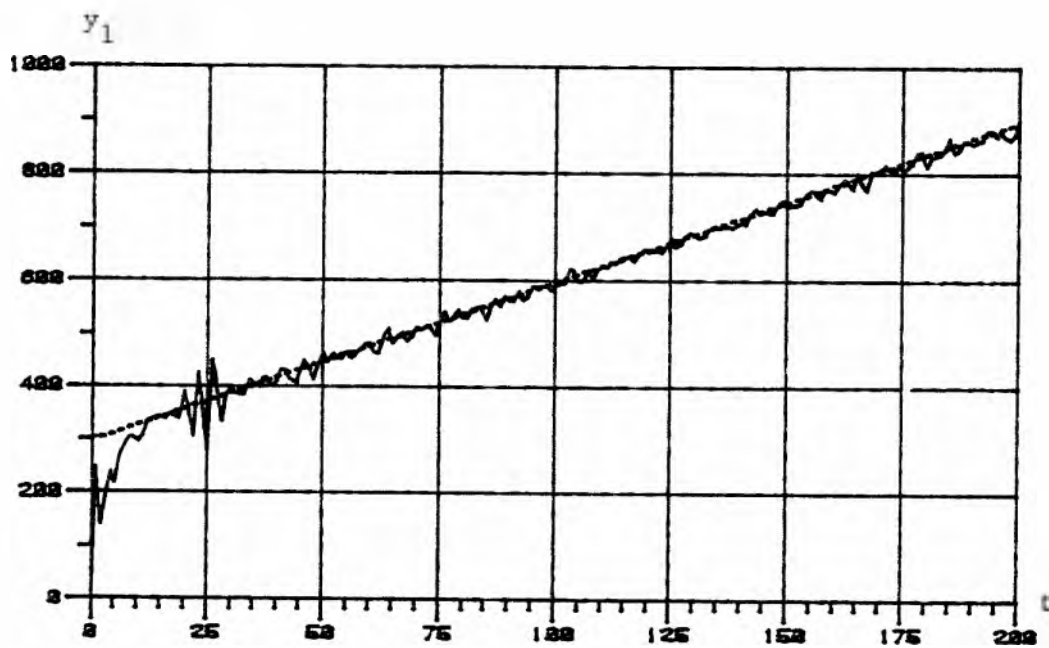


Figure 5.3a. Time response of y_1 using extended static game.

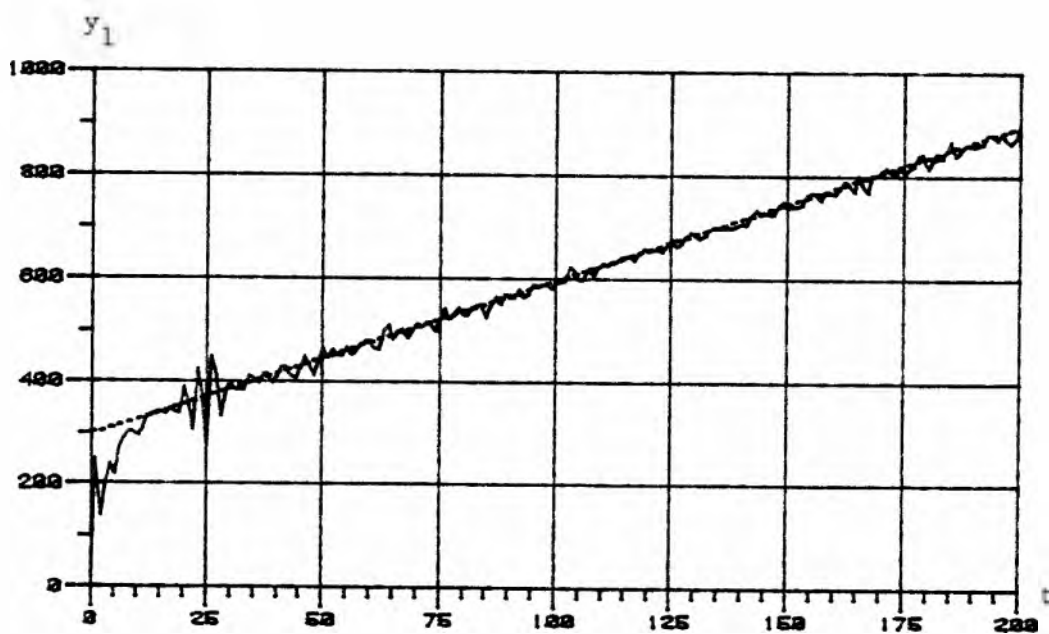


Figure 5.3b. Time response of y_1 using constrained policy.

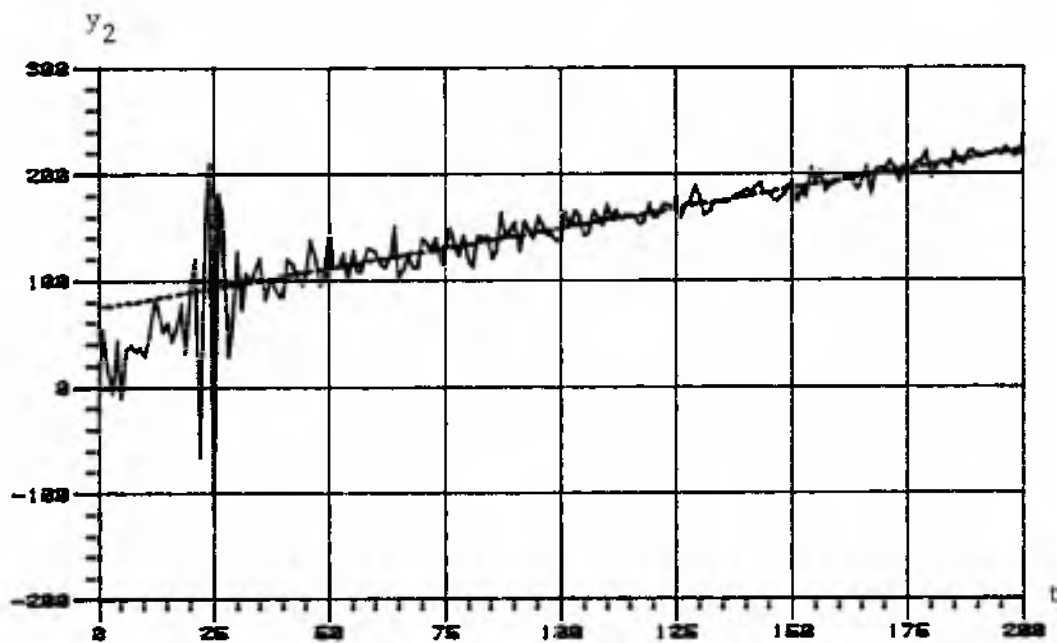


Figure 5.4a. Time response of y_2 using extended static game.

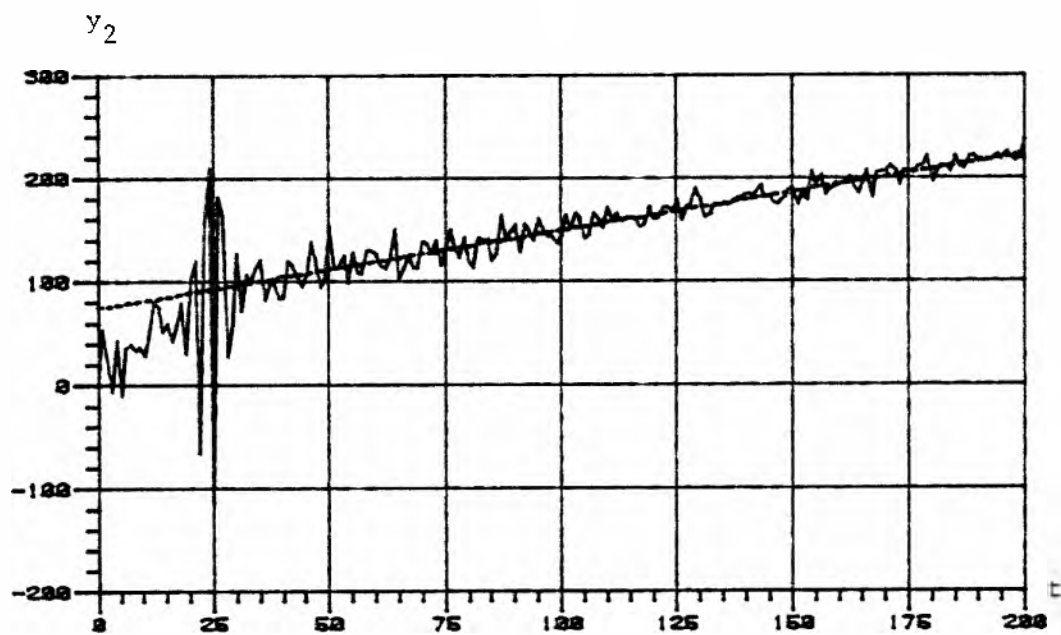


Figure 5.4b. Time response of y_2 using constrained policy.

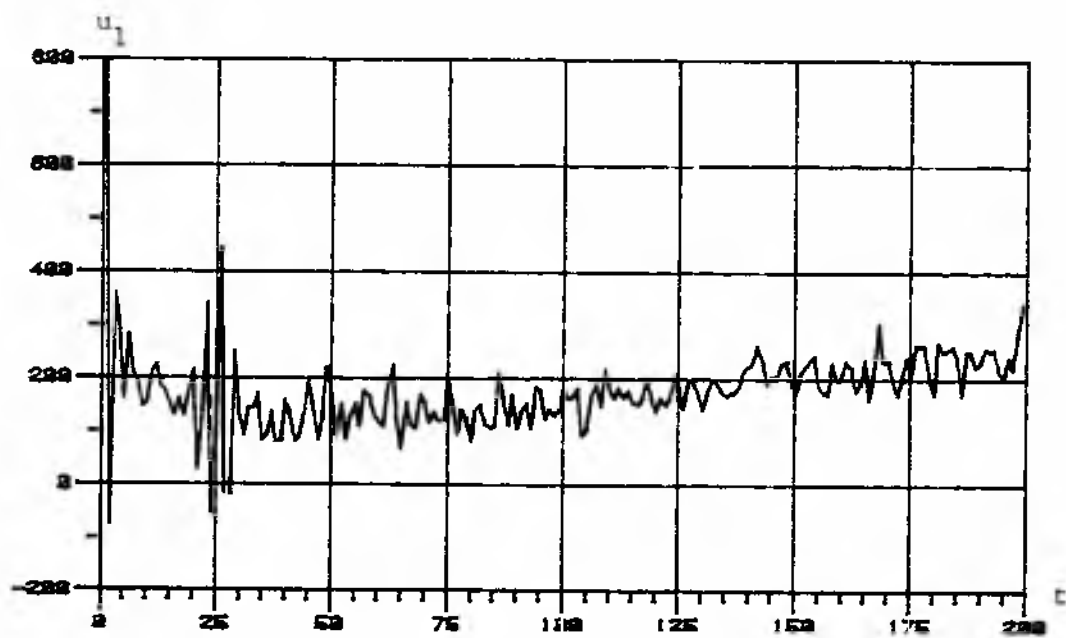


Figure 5.5a. Time response of u_1 using extended static game.

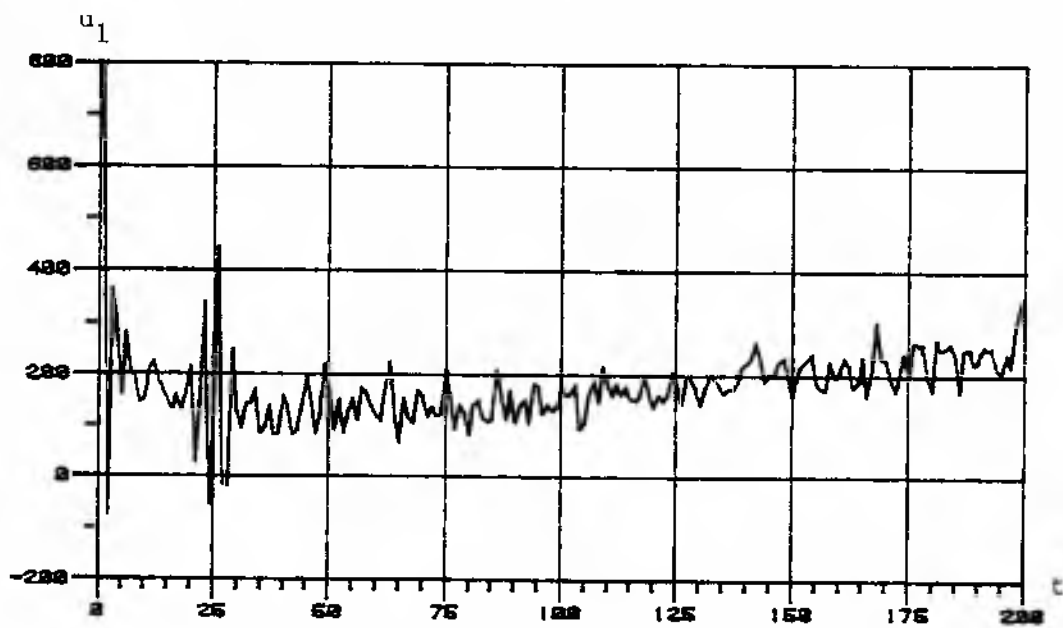


Figure 5.5b. Time response of u_1 using constrained policy.

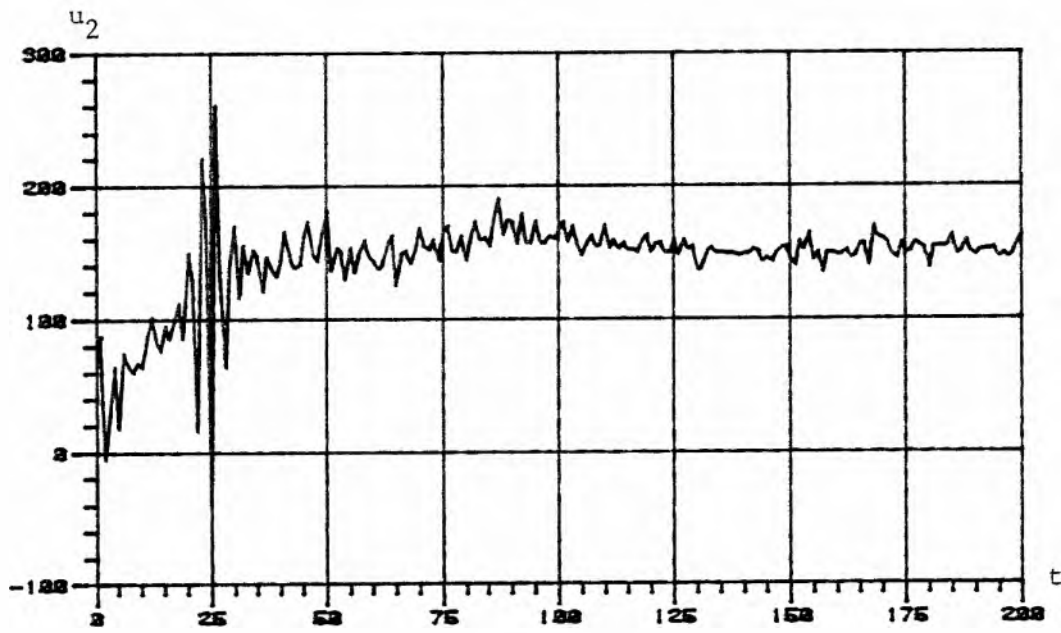


Figure 5.6a. Time response of u_2 using extended static game.

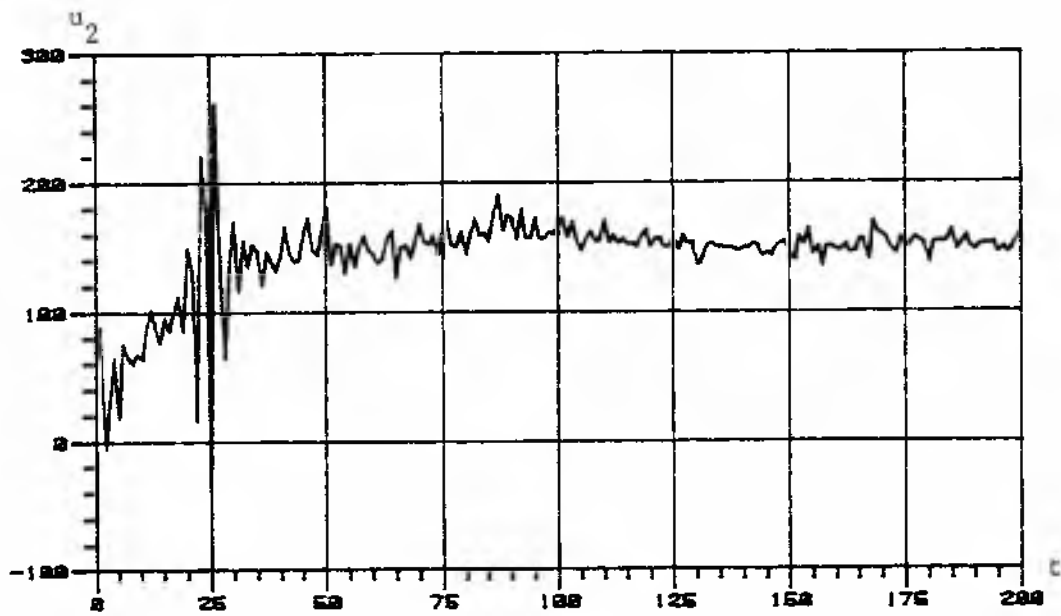


Figure 5.6b. Time response of u_2 using constrained policy.

which the constrained approach may offer simpler computations, and thus hold an edge over the theoretical by solid but cumbersome extended static game approach. We can use this constrained approach with peace of mind if we know there does exist a theoretical basis for the policy structure.

The input-output responses of the same run with all the parameters known are shown in Figures 5.7 and 5.8. The algorithms perform satisfactorily when all the parameters are known, which is particularly evident during the start up period. To compare the error in the optimal policy, we let $u^k(t)$ denote the controls obtained with known parameters and $u^n(t)$ denote the policy obtained with unknown parameters. The quantity $e^u(t) = u^k(t) - u^n(t) = [e_1^u(t)e_2^u(t)]^T$ is shown in Figure 5.9. We see that the policy error $e^u(t)$ seems to be a zero mean quantity, which indicates the algorithms are providing good controls even though the parameter estimates in the simulation are far from converging.

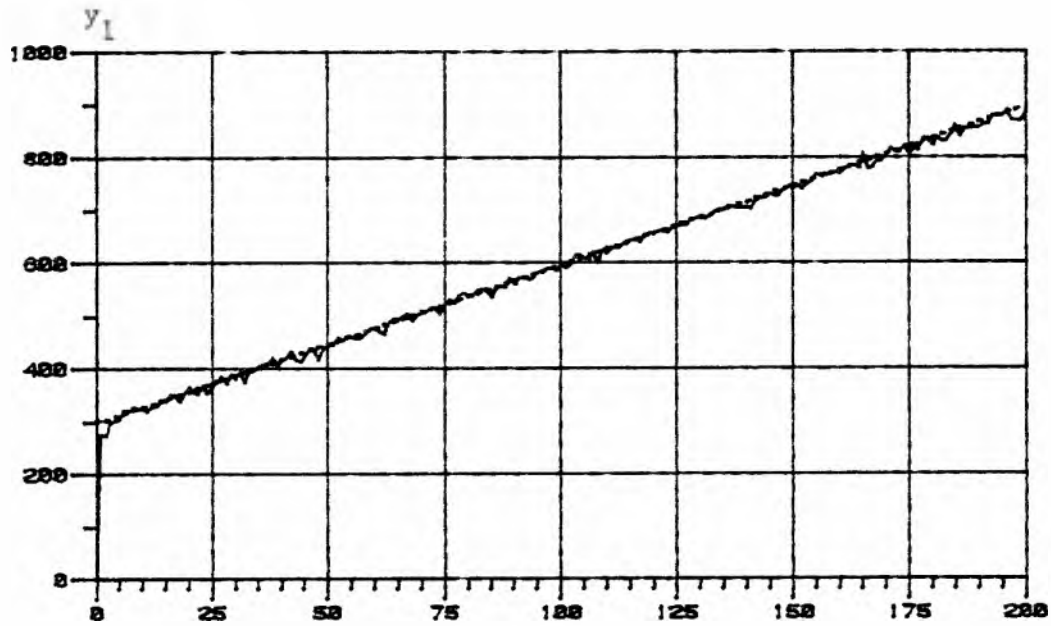


Figure 5.7a. Time response of y_1 with known parameter.

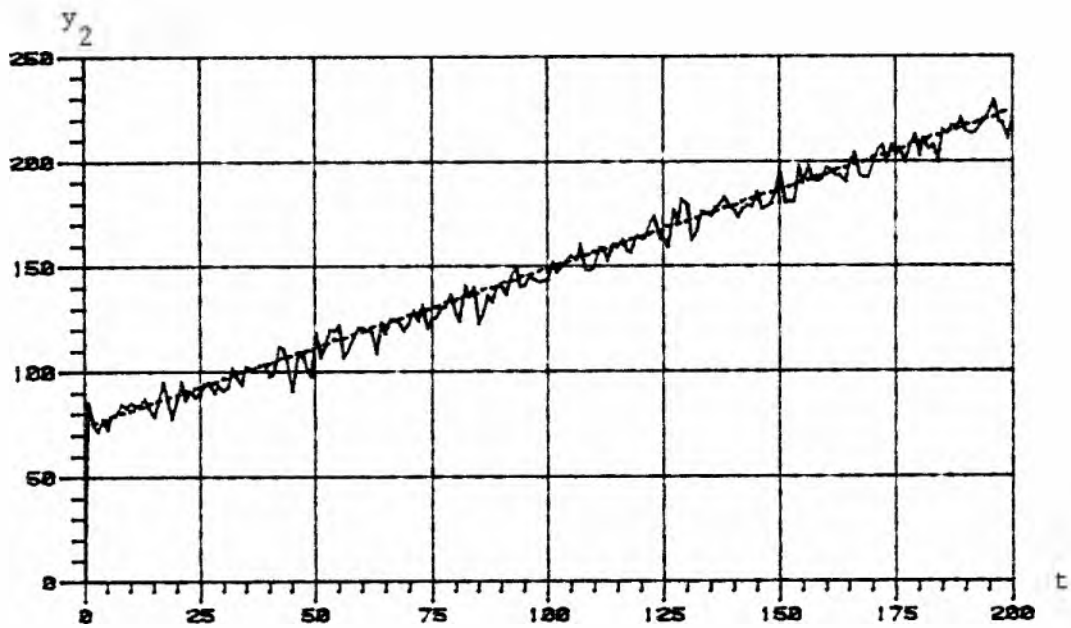


Figure 5.7b. Time response of y_2 with known parameter.

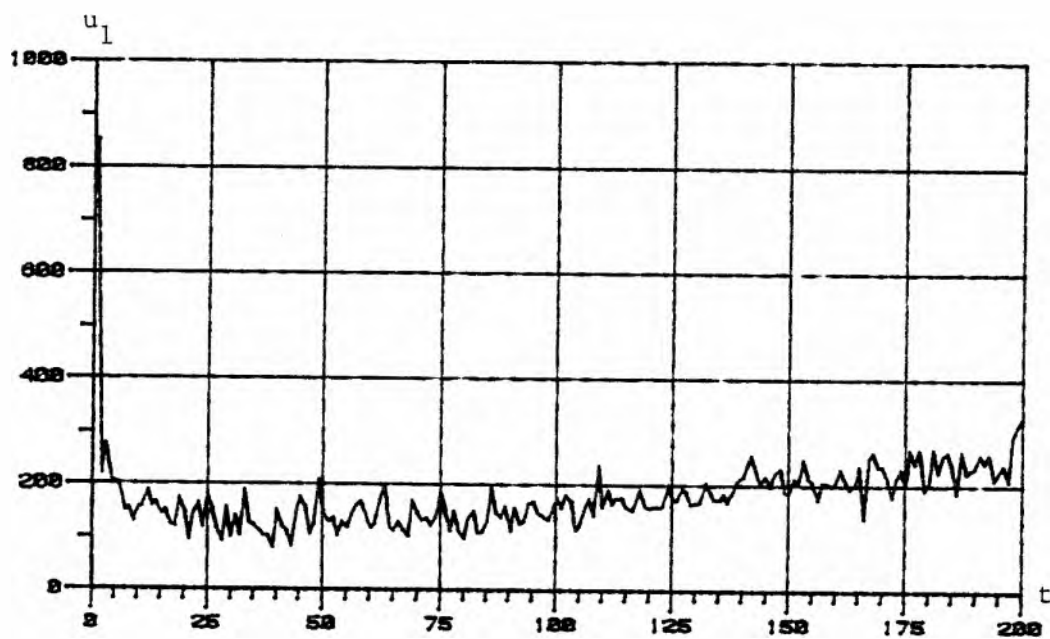


Figure 5.8a. Time response of u_1 with known parameter.

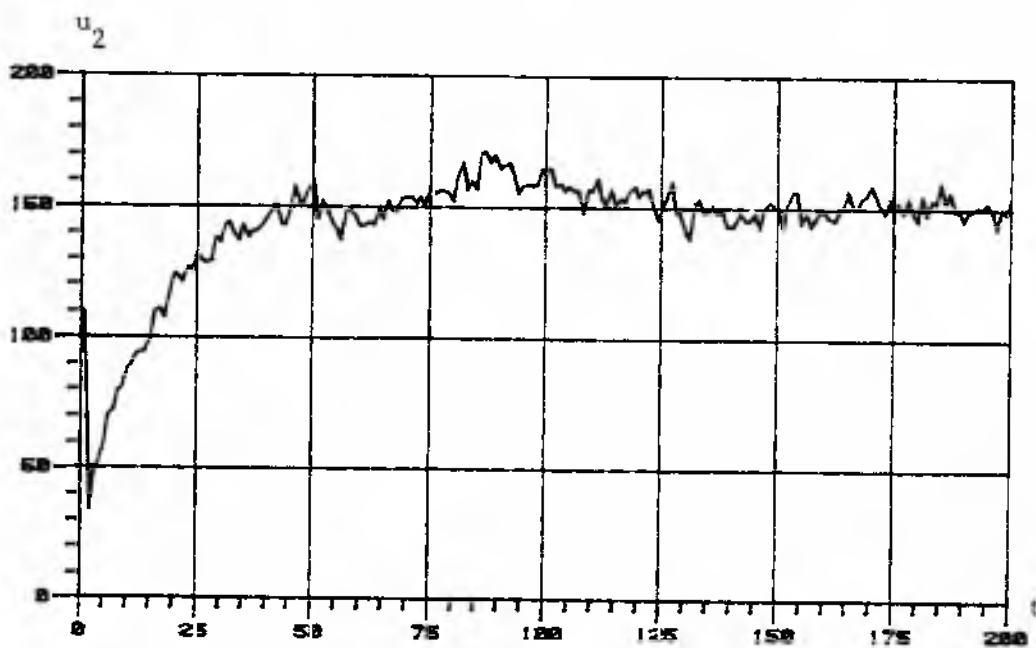


Figure 5.8b. Time response of u_2 with known parameter.

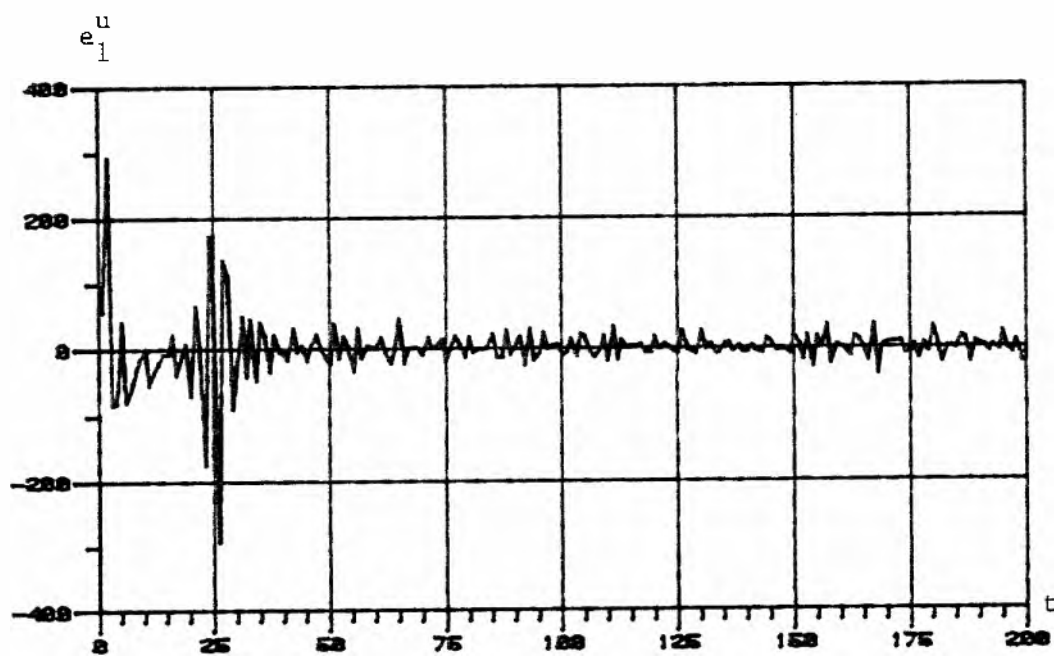


Figure 5.9a. Time response of policy error e_1^u .

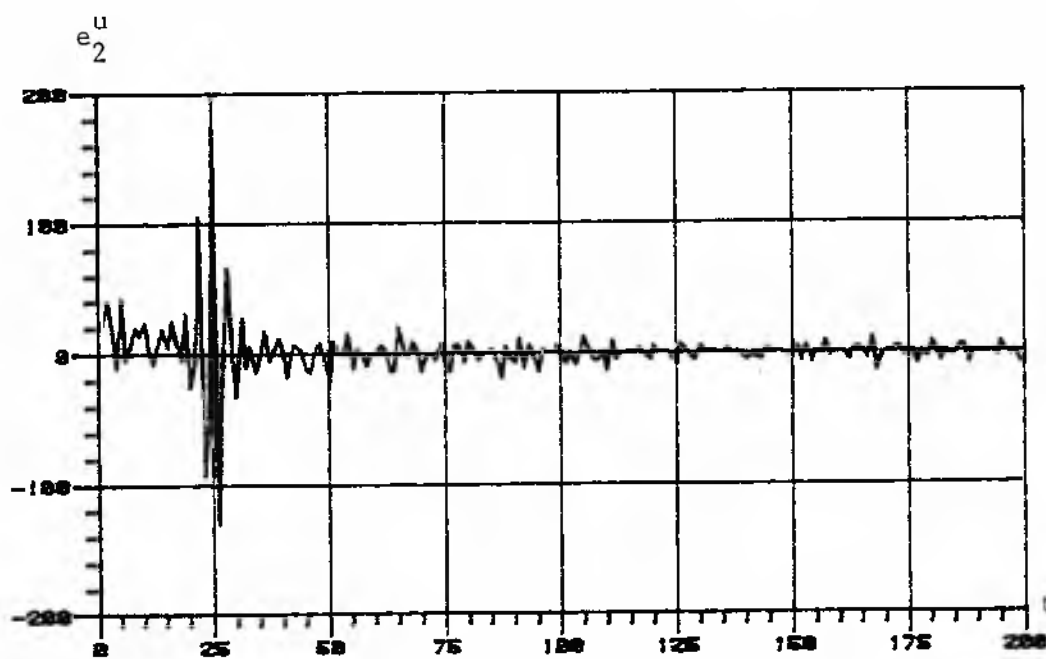


Figure 5.9b. Time response of policy error e_2^u .

CHAPTER 6

CONCLUSION

In this report, steady-state solutions are obtained for the optimization of stochastic systems with unknown parameters and multiple decision-makers each having his own objective. The solutions obtained for these systems, or games, have the advantage of simplicity and easy implementation and thus lend themselves to possible applications in a variety of actual systems.

Two types of centralized stochastic adaptive games are considered: the Nash game problem and the Leader-Follower game problem. The resulting adaptive solutions for these games can be classified as those of the implicit self-tuning type. It is established in this report that by a judicious transformation, these game solutions can be made to resemble closely the implicit self-tuning solution for the single-controller single-objective case, thus endowing them with the desirable property of simple implementation. In addition, convergence of these game problems is established utilizing this close resemblance.

In Chapter 5, we proposed two explicit self-tuning type methods for decentralized stochastic adaptive Nash games under the "one-step-delay information sharing pattern". The first method is an ad hoc constraint on the policy form while the second one is an extension of static Nash game theory. Simulation results show that both methods generate identical optimal policy and indicate that the two algorithms may be equivalent. Even though results from simulation are satisfactory, a theoretical basis for convergence of the decentralized Nash game problem still needs to be established.

All the methods used in this report to deal with unknown parameters have been of the certainty equivalent types. Future research into this area may include combination of the present approach with some other methods that will take into account the inaccuracies of the estimates. Another area related to the estimates is the use of different estimation schemes by different controllers, thus leading to the problem of multi-modeling.

REFERENCES

- [1] Astrom, K. J., Introduction to Stochastic Control Theory, Academic Press, New York, 1970.
- [2] Astrom, K. J., "Computer Control of a Paper Machine—an Application of Linear Stochastic Control Theory," IBM Journal, July 1967, pp. 389-405.
- [3] Astrom, K. J. and Eykhoff, P., "System Identification—A Survey," Automatica, Vol. 7, pp. 123-162, 1971.
- [4] Astrom, K. J. and Wittenmark, B., "Problems of Identification and Control," JOTA, Vol. 34, 1971, pp. 90-113.
- [5] Astrom, K. J. and Wittenmark, B., "On Self Tuning Regulators," Automatica, Vol. 9, pp. 185-199, 1973.
- [6] Astrom, K. J., Borisson, U., Ljung, L., Wittenmark, B., "Theory and Applications of Self-Tuning Regulators," Automatica, Vol. 13, pp. 457-476, 1977.
- [7] Athans, M. and Kendrick, D., "Control Theory and Economics: A Survey, Forecast, and Speculations," IEEE Trans. on Automatic Control, Vol. AC-19, No. 5, Oct. 1974, pp. 518-524.
- [8] Bar-Shalom, Y. and Tse, E., "Dual Effect, Certainty Equivalence and Separation in Stochastic Control," IEEE Trans. on Automatic Control, Vol. AC-19, No. 5, Oct. 1974, pp. 494-500.
- [9] Bar-Shalom, Y. and Tse, E., "Concepts and Methods in Stochastic Control," in Control and Dynamic Systems, Vol. 12, edited by C. T. Leondes, Academic Press, New York, 1976.
- [10] Bar-Shalom, Y. and Walls, K. D., "Dual Adaptive Control and Uncertainty Effects in Macroeconomic Systems Optimization," Automatica, Vol. 16, pp. 147-156, 1980.
- [11] Basar, T., "Equilibrium Solutions in Two-Person Quadratic Decision Problems with Static Information Structure," IEEE Trans. on Automatic Control, Vol. AC-20, No. 3, June 1975, pp. 320-328.
- [12] Basar, T., "Decentralized Multicriteria Optimization of Linear Stochastic Systems," IEEE Trans. on Automatic Control, Vol. AC-23, No. 2, April 1978, pp. 233-243.
- [13] Basar, T. and Selbuz, H., "Closed-Loop Stackelberg Strategies with Applications in the Optimal Control of Multilevel Systems," IEEE Trans. on Automatic Control, Vol. AC-24, No. 2, April 1979, pp. 166-179.

- [14] Borison, U., "Self-Tuning Regulators for a Class of Multivariable Systems," Automatica, Vol. 15, pp. 209-215, 1979.
- [15] Chow, G. C., Analysis and Control of Dynamic Economic System, John Wiley and Sons, Toronto, 1975.
- [16] Clarke, D. W. and Hastings-James, R., "Design of Digital Controllers for Randomly Disturbed Systems," Proc. IEE, Vol. 118, No. 10, Oct. 1971, pp. 1503-1506.
- [17] Clarke, D. W. and Gawthrop, P. J., "Self-tuning Controller," Proc. IEE, Vol. 122, No. 9, Sept. 1975, pp. 929-934.
- [18] Clarke, D. W. and Gawthrop, P. J., "Self-Tuning Control," Proc. IEE, Vol. 126, No. 6, June 1979, pp. 633-640.
- [19] Clarke, D. W. and Gawthrop, P. J., "Implementation and Application of Microprocessor-Based Self-Tuners," Automatica, Vol. 17, No. 1, pp. 233-244, 1981.
- [20] Cruz, J. B., Jr., "Survey of Nash and Stackelberg Equilibrium Strategies in Dynamic Games," Annals of Economic and Social Measurement, 4/2, 1975, pp. 339-344.
- [21] Cruz, J. B., Jr., "Leader-Follower Strategies for Multilevel Systems," IEEE Trans. on Automatic Control, Vol. AC-23, No. 2, April 1978, pp. 244-255.
- [22] Feldbaum, A., "Dual Control Theory, I-IV," Automation and Remote Control, Vol. 21, pp. 1240-1249, 1960; Vol. 21, pp. 1453-1464, 1960; Vol. 22, pp. 3-16, 1961; Vol. 22, pp. 139-143, 1961.
- [23] Gardner, B. F. and Cruz, J. B., Jr., "Feedback Stackelberg Strategy for a Two Player Game," IEEE Trans. on Automatic Control, Vol. AC-22, No. 2, April 1977, pp. 270-271.
- [24] Gardner, B. F. and Cruz, J. B., Jr., "Feedback Stackelberg Strategy for M-Level Hierarchical Games," IEEE Trans. on Automatic Control, Vol. AC-23, No. 3, June 1978, pp. 489-491.
- [25] Gawthrop, P. J., "Some Interpretations of the Self-Tuning Controller," Proc. IEE, Vol. 124, No. 10, Oct. 1977, pp. 889-894.
- [26] Gawthrop, P. J., "On the Stability and Convergence of a Self-Tuning Controller," Int. J. Control, Vol. 31, No. 5, 1980, pp. 973-998.
- [27] Goodwin, G. C., Ramadge, P. J. and Caines, P. E., "Discrete Time Stochastic Adaptive Control," Proc. IEEE 18th Conference on Decision and Control, pp. 736-739, Florida, Dec. 1979.

- [28] Goodwin, G. C., Sin, K. S. and Saluja, K. K., "Stochastic Adaptive Control and Prediction—The General Delay-Colored Noise Case," IEEE Trans. on Automatic Control, Vol. AC-25, No. 5, Oct. 1980, pp. 946-950.
- [29] Grimble, M. J., "A Control Weighted Minimum-Variance Controller for Non-Minimum Phase Systems," Int. J. Control, Vol. 33, No. 4, 1981, pp. 751-762.
- [30] Hannan, E. J., "The Identification Problem for Multiple Equation Systems with Moving Average Errors," Econometrica, Vol. 39, No. 5, Sept. 1971, pp. 751-765.
- [31] Holly, S., Rustem, B. and Zarrop, B., Optimal Control for Econometric Models: An Approach to Economic Policy Formulation, St. Martin's Press, New York, 1979.
- [32] Johnson, T. L. and Athans, M., "On the Design of Optimal Constrained Dynamic Compensators for Linear Constant Systems," IEEE Trans. on Automatic Control, Vol. AC-15, No. 6, Dec. 1970, pp. 658-660.
- [33] Keviczky, L. and Kumar, K. S. P., "Multivariable Self-Tuning Regulator with Generalized Cost Functions," Int. J. Control, Vol. 33, No. 5, May 1981, pp. 913-922.
- [34] Kurtaran, B., "Decentralized Stochastic Control with Delayed Sharing Information Pattern," IEEE Trans. on Automatic Control, Vol. AC-21, No. 4, August 1976, pp. 576-581.
- [35] Kurtaran, B., "Corrections and Extensions to Decentralized Stochastic Control with Delayed Sharing Information Pattern," IEEE Trans. on Automatic Control, Vol. AC-24, No. 4, August 1979, pp. 656-657.
- [36] Kurtaran, B., "Dynamic Two-Person Two-Objective Control Problems with Delayed Sharing Information Pattern," IEEE Trans. on Automatic Control, Vol. AC-22, No. 4, August 1977, pp. 659-661.
- [37] Levine, W. S. and Athans, M., "On the Determination of the Optimal Constant Output Feedback Gains for Linear Multivariable Systems," IEEE Trans. on Automatic Control, Vol. AC-15, No. 1, Feb. 1970, pp. 44-48.
- [38] Ljung, L., "On Positive Real Transfer Functions and the Convergence of Some Recursive Schemes," IEEE Trans. on Automatic Control, Vol. AC-22, No. 4, August 1977, pp. 539-551.
- [39] Ljung, L., "Analysis of Recursive Stochastic Algorithms," IEEE Trans. on Automatic Control, Vol. AC-22, No. 4, August 1977, pp. 551-575.

- [40] Ljung, L., "Convergence of an Adaptive Filter Algorithm," Int. J. Control, Vol. 27, No. 5, 1979, pp. 673-693.
- [41] Mendel, J. M., "A Concise Derivation of Optimal Constant Limited State Feedback Gains," IEEE Trans. on Automatic Control, Vol. AC-19, No. 4, August 1974, pp. 447-448.
- [42] Nash, J., "Noncooperative Games," Annals of Mathematics, Vol. 54, No. 2, Sept. 1951.
- [43] von Neumann, J. and Morgenstern, O., Theory of Games and Economic Behavior, Princeton University Press, 1944.
- [44] Papavassilopoulos, G. P. and Cruz, J. B., Jr., "Nonclassical Control Problems and Stackelberg Games," IEEE Trans. on Automatic Control, Vol. AC-24, No. 2, April 1979, pp. 155-165.
- [45] Peterka, V., "Adaptive Digital Regulation of Noisy Systems," Preprint IFAC Symposium on Identification and Process Parameter Estimation, Prague, 1970.
- [46] Pindyck, R. S., Optimal Planning for Economic Stabilization, North-Holland, Amsterdam, 1973.
- [47] Pindyck, R. S., "Optimal Economic Stabilization Policies Under Decentralized Control and Conflicting Objectives," IEEE Trans. on Automatic Control, Vol. AC-22, No. 4, August 1977, pp. 517-530.
- [48] Radner, R., "Team Decision Problems," Annals of Math. Statistics, Vol. 33, Sept. 1962, pp. 857-881.
- [49] Simaan, M. and Cruz, J. B., Jr., "On the Stackelberg Strategy in Nonzero-Sum Games," JOTA, Vol. 11, No. 5, May 1973, pp. 533-555.
- [50] Simaan, M. and Cruz, J. B., Jr., "Additional Aspects of the Stackelberg Strategy in Nonzero-Sum Games," JOTA, Vol. 11, No. 6, June 1973, pp. 613-626.
- [51] Simaan, M. and Cruz, J. B., Jr., "A Stackelberg Solution for Games with Many Players," IEEE Trans. on Automatic Control, Vol. AC-18, No. 3, June 1973, pp. 322-324.
- [52] von Stackelberg, H., The Theory of the Market Economy, Oxford University Press, Oxford, England, 1952.
- [53] Starr, A. W. and Ho, Y. C., "Nonzero-Sum Differential Games," JOTA, Vol. 3, No. 3, 1969, pp. 184-206.
- [54] Starr, A. W. and Ho, Y. C., "Further Properties of Nonzero-Sum Differential Games," JOTA, Vol. 3, No. 4, 1969, pp. 207-219.

- [55] Wall, K. D. and Westcott, J. H., "Macroeconomic Modeling for Control," IEEE Trans. on Automatic Control, Vol. AC-19, No. 6, Dec. 1974, pp. 862-873.
- [56] Wittenmark, B., "Stochastic Adaptive Control Methods: A Survey," Int. J. Control, Vol. 21, No. 5, 1975, pp. 705-730.
- [57] Witsenhausen, H. S., "Separation of Estimation and Control for Discrete Time Systems," Proc. of the IEEE, Vol. 59, No. 11, Nov. 1971, pp. 1557-1566.

APPENDIX A

TRANSFORMATION OF SYSTEMS

Given a system governed by

$$A(q^{-1})y(t) = B(q^{-1})u(t-k-1) + C(q^{-1})e(t) + D \quad (A.1)$$

where D is a constant offset vector and $A(z)$, $B(z)$, $C(z)$ are matrix polynomials. The vectors $y(t)$, $u(t)$, $e(t)$ and D are all of the same dimension. Let $a(z)$ be the scalar polynomial formed by taking the determinant of $A(z)$ and let $\tilde{A}(z)$ represent the adjoint of $A(z)$. Hence, the inverse of $A(z)$, A^{-1} is given by

$$A^{-1}(z) = \frac{1}{a(z)} \tilde{A}(z) \quad (A.2)$$

Premultiplying (A.1) by $a(z)A^{-1}(z)$ yields

$$\begin{aligned} a(q^{-1})y(t) &= \tilde{A}(q^{-1})B(q^{-1})u(t-k-1) + \tilde{A}(q^{-1})C(q^{-1})e(t) \\ &\quad + \tilde{A}(q^{-1})D \end{aligned}$$

or

$$a(q^{-1})y(t) = \bar{B}(q^{-1})u(t-k-1) + \bar{C}(q^{-1})e(t) + \bar{D} \quad (A.3)$$

where $\bar{B}(z) = \tilde{A}(z)B(z)$, $\bar{C}(z) = \tilde{A}(z)C(z)$ and $\bar{D} = \tilde{A}(1)D$. The resulting system (A.3) has a scalar polynomial operating on $y(t)$.

APPENDIX B

PROOF OF THEOREM 3.1

Consider the system governed by

$$a(q^{-1})y(t) = B(q^{-1})u(t-k-1) + C(q^{-1})e(t) + D \quad (B.1)$$

The cost function J_i associated with the i -th controller is given by

$$J_i = E\{[y(t+k+1) - y^r(t+k+1)]^T Q_i [y(t+k+1) - y^r(t+k+1)] \\ + [u(t) - u(t-1)]^T R_i [u(t) - u(t-1)]\}, \quad (B.2)$$

$$i = 1, 2$$

From the proof in Theorem 2.1, we can transform (B.1) into the following prediction model form

$$\bar{c}(q^{-1})y^*(t+k+1|t) = G(q^{-1})C(q^{-1})y(t) + F(q^{-1})C(q^{-1})B(q^{-1})u(t) \\ + F(q^{-1})C(q^{-1})D \quad (B.3)$$

where

$$y^*(t+k+1|t) = y(t+k+1) - F(q^{-1})e(t+k+1) \quad (B.4)$$

Assuming the existence of $\frac{\partial J_i}{\partial u_i(t)}$, we substitute y^* into (B.2) and set

$\frac{\partial J_i}{\partial u_i(t)}$ to zero to obtain

$$0 = B_0^T(i) Q_i (y^*(t+k+1|t) - y^r(t+k+1)) \\ + R_i^T(i) (u(t) - u(t-1)), \quad i = 1, 2, \dots, N \quad (B.5)$$

Stacking up the N equations in (B.5), we have

$$0 = M(y^*(t+k+1|t) - y^r(t+k+1)) + H(u(t) - u(t-1)) \quad (B.6)$$

where M and H are defined in (3.9e) and (3.9f) respectively. Multiply (B.6) by $\bar{c}(z)$ and combining the resulting equation with (B.3), we have

$$\begin{aligned} MG(q^{-1})C(q^{-1})y(t) + [MF(q^{-1})C(q^{-1})B(q^{-1}) + (1-q^{-1})\bar{c}(q^{-1})H]u(t) \\ + MF(q^{-1})C(q^{-1})D - \bar{c}(q^{-1})My^r(t+k+1) = 0 \end{aligned}$$

as stated in the theorem.

APPENDIX C

ECONOMIC MODEL

The economic model used for the simulation study is taken from [15, pp. 272]. It is given in the following form

$$C_t = 0.9266C_{t-1} - 0.0203I_{t-1} + 0.3190G_t + 0.4206M_t - 63.2386 \quad (C.1)$$

$$I_t = 0.1527C_{t-1} + 0.3806I_{t-1} - 0.0735G_t + 1.538M_t - 210.8994 \quad (C.2)$$

where C_t denotes consumption expenditures, I_t is private investment expenditure, G_t is government expenditure and M_t is the money supply.

Let $y_1(t) = C_t$, $y_2(t) = I_t$, $u_1(t-1) = G_t$, and $u_2(t-1) = M_t$. We assume the current G_t and M_t are the result of, and equal to, the desired levels that were specified in the previous time step, thus the time lag in the definition of u_1 and u_2 [46]. The resulting model in terms of y and u is given in the following matrix polynomial form

$$\left(I + \begin{bmatrix} 0.9266 & -0.0203 \\ 0.1527 & 0.3806 \end{bmatrix} q^{-1} \right) y(t) = \begin{bmatrix} 0.3190 & 0.4206 \\ -0.0735 & 1.538 \end{bmatrix} u(t-1) - \begin{bmatrix} 63.2386 \\ 210.8994 \end{bmatrix} \quad (C.3)$$

Using the transformation technique given in Appendix A, we have the transformed system

$$\begin{aligned}
 (1 - 1.3072q^{-1} + 0.3596q^{-2})y(t) = & \begin{bmatrix} 0.3190 & 0.4206 \\ -0.0735 & 1.5380 \end{bmatrix} \\
 + & \begin{bmatrix} -0.1199 & -0.1913 \\ 0.1168 & -1.3609 \end{bmatrix} q^{-1} u(t-1) + \begin{bmatrix} -34.8887 \\ -25.1366 \end{bmatrix} \quad (C.4)
 \end{aligned}$$

or

$$a(q^{-1})y(t) = B(q^{-1})u(t-1) + D. \quad (C.5)$$

For simulation, we assume the system is perturbed by zero mean white noise $e(t)$, that is,

$$a(q^{-1})y(t) = B(q^{-1})u(t-1) + D + e(t) \quad (C.6)$$

with

$$E\{e(t)e^T(t)\} = \begin{bmatrix} 54 & 12 \\ 12 & 26 \end{bmatrix}.$$

The weighting matrices for simulation in Chapters 3 and 4 are

$$\begin{aligned}
 Q_1 &= \begin{bmatrix} 5 & 0 \\ 0 & 15 \end{bmatrix}, & R_1 &= \begin{bmatrix} 0.2 & 0 \\ 0 & 0.4 \end{bmatrix} \\
 Q_2 &= \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix}, & R_2 &= \begin{bmatrix} 0.1 & 0 \\ 0 & 0.5 \end{bmatrix}.
 \end{aligned}$$

All input-output variables are in billions of dollars and each time step t is one quarter.

APPENDIX D

PROOF OF THEOREM 4.1

Consider the system

$$a(q^{-1})y(t) = B(q^{-1})u(t-k-1) + C(q^{-1})e(t) + D. \quad (D.1)$$

The cost function J_i associated with the i^{th} controller is

$$J_i = E\{[y(t+k+1) - y^r(t+k+1)]^T Q_i [y(t+k+1) - y^r(t+k+1)] \\ + [u(t) - u(t-1)]^T R_i [u(t) - u(t-1)]\}, \quad (D.2)$$

$$i = 1, 2$$

From proof in Theorem 2.1, it is possible to arrive at the following prediction model form of (D.1),

$$\bar{c}(q^{-1})y^*(t+k+1|t) = G(q^{-1})C(q^{-1})y(t) + F(q^{-1})C(q^{-1})B(q^{-1})u(t) \\ + F(q^{-1})C(q^{-1})D \quad (D.3)$$

where

$$y^*(t+k+1|t) = y(t+k+1) - F(q^{-1})e(t+k+1). \quad (D.4)$$

The y^* is substituted into the cost function (D.2) and the necessary condition for minimum is then derived as follows.

Follower. J_2 is the cost function associated with the follower.

The necessary condition for the follower is

$$\begin{aligned}
\frac{\partial J_2}{\partial u_2(t)} &= 0 \\
&= B_0^T(2) Q_2 [y^*(t+k+1|t) - y^r(t+k+1)] \\
&\quad + R_2^T(2) [u(t) - u(t-1)] \quad (D.5)
\end{aligned}$$

$$= M_2 [y^*(t+k+1|t) - y^r(t+k+1)] + H_2 [u(t) - u(t-1)] . \quad (D.6)$$

Leader. J_1 is the cost function associated with the leader. The leader optimizes J_1 taking into consideration the possible reaction of the follower. Thus, he will append the follower's optimization equation to that of his, that is, he will minimize, with respect to u_1 and u_2 , a cost function J given by

$$J = J_1 + \lambda_t^T \frac{\partial J_2}{\partial u_2(t)} \quad (D.7)$$

where λ_t is a Lagrange multiplier. The necessary conditions of minimum for the leader is $\frac{\partial J}{\partial u_1(t)} = 0$, for $i = 1, 2$. For simplicity of derivation, we will assume u_1 and u_2 to be scalar valued. Hence, we have

$$\begin{aligned}
\frac{\partial J}{\partial u_1(t)} &= 0 \\
&= B_0^T(1) Q_1 [y^*(t+k+1|t) - y^r(t+k+1)] + \lambda_t^T (B_0^T(2) Q_2 B_0^{(1)} + (R_2)_{12}) \\
&\quad + R_1^T(1) [u(t) - u(t-1)] \quad (D.8)
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial J}{\partial u_2(t)} &= 0 \\
&= B_0^{T(2)} Q_2 [y^*(t+k+1|t) - y^r(t+k+1)] + \lambda_t (B_0^{T(2)} Q_2 B_0^{(1)} + (R_2)_{22}) \\
&\quad + R_1^{T(2)} [u(t) - u(t-1)] \quad . \quad (D.9)
\end{aligned}$$

Equations (D.6), (D.8), and (D.9) yield

$$M[y^*(t+k+1|t) - y^r(t+k+1)] + H[u(t) - u(t-1)] = 0 \quad (D.10)$$

where M and H are as defined in (4.6e) and (4.6f) respectively.

Combining (D.3) and (D.10), we have

$$\begin{aligned}
&MG(q^{-1})C(q^{-1})y(t) + [MF(q^{-1})C(q^{-1})B(q^{-1}) + (1-q^{-1})\bar{c}(q^{-1})H]u(t) \\
&\quad + MF(q^{-1})C(q^{-1})D - \bar{c}(q^{-1})My^r(t+k+1) = 0
\end{aligned}$$

as given in (4.5) of Theorem 4.1.

APPENDIX E

PROOF OF THEOREM 5.1

Consider the system governed by (5.1)

$$y(t+1) = a(q^{-1})y(t) + B(q^{-1})u(t) + e(t+1) + D \quad (E.1)$$

with the cost functions (5.14)

$$\begin{aligned} J_i = E\{ & u_i^T(t)D_{ii}u_i(t) + 2u_i^T(t)D_{ij}u_j(t) + 2y_i^T(t)C_i u_i(t) \\ & + 2x_i^T(t)u_i(t) \} , \quad i = 1, 2 . \end{aligned} \quad (E.2)$$

The policy u_i is constrained to be of the form

$$u_i(t) = G_i y_i(t) + g_i . \quad (E.3)$$

Without loss of generality, let us consider J_1 . Substituting u_1 in (E.3) into J_1 yields

$$\begin{aligned} J_1 = E\{ & y_1^T(t)G_1^T D_{11} G_1 y_1(t) + g_1^T D_{11} g_1 + 2y_1^T(t)G_1^T D_{11} g_1 \\ & + 2[y_1^T(t)G_1^T D_{12} G_2 y_2(t) + y_1^T(t)G_1^T D_{12} g_2 \\ & + g_1^T D_{12} G_2 y_2(t) + g_1^T D_{12} g_2] + 2[y_1^T(t)C_1 G_1 y_1(t) \\ & + y_1^T(t)C_1 g_1 + x_1^T(t)G_1 y_1(t) + x_1^T(t)g_1] \} . \end{aligned} \quad (E.4)$$

Denote

$$E\{y_i(t)y_j^T(t)\} = P_{ij}(t) \quad (E.5)$$

$$E\{y_i(t)\} = \bar{y}_i(t) . \quad (E.6)$$

Then, taking expectation of the terms in (E.4), we have

$$\begin{aligned}
 J_1 = & \text{trace}[G_1^T D_{11} G_1 P_{11}(t) + g_1^T D_{11} g_1 + 2\bar{y}_1^T(t) G_1^T D_{11} g_1 \\
 & + 2G_1^T D_{12} G_2 P_{21}(t) + 2\bar{y}_1^T(t) G_1^T D_{12} g_2 + 2g_1^T D_{12} G_2 \bar{y}_2(t) \\
 & + 2g_1^T D_{12} g_2 + 2C_1 G_1 P_{11}(t) + 2\bar{y}_1^T(t) C_1 g_1 \\
 & + 2x_1^T(t) G_1 \bar{y}_1(t) + 2x_1^T(t) g_1] . \quad (E.7)
 \end{aligned}$$

The following formulae are then used to evaluate the necessary conditions for minimum $\frac{\partial J}{\partial G_1} = 0$ and $\frac{\partial J}{\partial g_1} = 0$:

$$\frac{\partial}{\partial Z} \text{tr}[NZ] = N^T$$

$$\frac{\partial}{\partial Z} \text{tr}[NZ^T] = N$$

$$\frac{\partial}{\partial Z} \text{tr}[NZL] = N^T L^T$$

$$\frac{\partial}{\partial Z} \text{tr}[Z^T L Z N] = L^T Z N^T + L Z N .$$

Hence, we have

$$\begin{aligned}
 0 = & \frac{\partial J_1}{\partial G_1} \\
 = & D_{11}^T G_1 P_{11}^T(t) + D_{11} G_1 P_{11}(t) + 2D_{11} g_1 \bar{y}_1^T(t) + 2D_{12} G_2 P_{21}(t) \\
 & + 2D_{12} g_2 \bar{y}_1^T(t) + 2C_1^T P_{11}^T(t) + 2x_1(t) \bar{y}_1^T(t)
 \end{aligned}$$

or

$$\begin{aligned}
 0 = & D_{11} G_1 P_{11}(t) + D_{11} g_1 \bar{y}_1^T(t) + D_{12} G_2 P_{21}(t) + D_{12} g_2 \bar{y}_1^T(t) \\
 & + C_1^T P_{11}(t) + x_1(t) \bar{y}_1^T(t) \quad (E.8)
 \end{aligned}$$

and

$$\begin{aligned}
 0 &= \frac{\partial J_1}{\partial g_1} \\
 &= 2D_{11} + 2D_{11}^T G_1 \bar{y}_1(t) + 2D_{12} G_2 \bar{y}_2(t) + 2D_{12} g_2 \\
 &\quad + 2C_1^T \bar{y}_1(t) + 2x_1(t)
 \end{aligned}$$

or

$$\begin{aligned}
 0 &= D_{11} g_1 + D_{11}^T G_1 \bar{y}_1(t) + D_{12} G_2 \bar{y}_2(t) + D_{12} g_2 \\
 &\quad + C_1^T \bar{y}_1(t) + x_1(t) . \quad (E.9)
 \end{aligned}$$

Similarly, from $\frac{\partial J_2}{\partial G_2} = 0$ and $\frac{\partial J_2}{\partial g_2} = 0$ we obtain

$$\begin{aligned}
 0 &= D_{22} G_2 P_{22}(t) + D_{22} g_2 \bar{y}_2^T(t) + D_{21} G_1 P_{12}(t) + D_{21} g_1 \bar{y}_2^T(t) \\
 &\quad + C_2^T P_{22}(t) + x_2(t) \bar{y}_2^T(t) \quad (E.10)
 \end{aligned}$$

$$\begin{aligned}
 0 &= D_{22} g_2 + D_{22} G_2 \bar{y}_2(t) + D_{21} G_1 \bar{y}_1(t) + D_{21} g_1 \\
 &\quad + C_2^T \bar{y}_2(t) + x_2(t) . \quad (E.11)
 \end{aligned}$$

Since $\bar{y}(t)$ and $E\{y(t)y^T(t)\}$ are required to solve for G_i, g_i ($i = 1, 2$) in (E.8)-(E.11), we show how these terms are computed.

At time t , past u, y are known, then

$$\begin{aligned}
E\{y_i(t)\} &= E\{a(q^{-1})y_i(t-1) + B_{ii}(q^{-1})u_i(t-1) \\
&\quad + B_{ij}(q^{-1})u_j(t-1) + e(t) + D\} \\
&= a(q^{-1})y_i(t-1) + B_{ii}(q^{-1})u_i(t-1) \\
&\quad + B_{ij}(q^{-1})u_j(t-1) + D \quad .
\end{aligned} \tag{E.12}$$

Let

$$P(t) = \begin{bmatrix} P_{11}(t) & P_{12}(t) \\ P_{21}(t) & P_{22}(t) \end{bmatrix} = E\{y(t)y^T(t)\} = \text{variance}\{y(t)\} \quad .$$

Since $\text{variance}\{y(t)\} = \text{covariance}\{y(t)\} + \bar{y}(t)\bar{y}^T(t)$, thus

$$\begin{aligned}
P(t) &= E\{(y(t) - \bar{y}(t))(y(t) - \bar{y}(t))^T\} + \bar{y}(t)\bar{y}^T(t) \\
&= E\{e(t)e^T(t)\} + \bar{y}(t)\bar{y}^T(t) \\
&= W + \bar{y}(t)\bar{y}^T(t) \quad .
\end{aligned} \tag{E.13}$$

Postmultiplying (E.9) by the term $\bar{y}_1^T(t)$ and subtracting the resulting equation from (E.8) yields

$$\begin{aligned}
D_{11}G_1[P_{11}(t) - \bar{y}_1(t)\bar{y}_1^T(t)] + D_{12}G_2[P_{21}(t) - \bar{y}_2(t)\bar{y}_1^T(t)] \\
= -C_1^T(P_{11}(t) - \bar{y}_1(t)\bar{y}_1^T(t))
\end{aligned}$$

or

$$D_{11}G_1W_{11} + D_{12}G_2W_{21} = -C_1^TW_{11} \quad . \tag{E.14}$$

Similarly, if $\bar{y}_2^T(t)$ is postmultiplied to (E.11) and then subtracted from (E.10), the following is obtained

$$D_{21}G_1W_{12} + D_{22}G_2W_{22} = -C_2^TW_{22} \quad (E.15)$$

After some manipulations with (E.14) and (E.15), we have

$$G_i - L_i G_i N_i = S_i, \quad i = 1, 2 \quad (E.16)$$

with

$$L_i = D_{ii}^{-1}D_{ij}D_{jj}^{-1}D_{ji}, \quad i, j = 1, 2, \quad i \neq j \quad (E.17)$$

$$N_i = W_{ij}W_{jj}^{-1}W_{ji}W_{ii}^{-1}, \quad i, j = 1, 2, \quad i \neq j \quad (E.18)$$

$$S_i = -D_{ii}^{-1}C_i^T + D_{ii}^{-1}D_{ij}D_{jj}^{-1}C_jW_{ji}W_{ii}^{-1}, \quad i, j = 1, 2, \quad i \neq j \quad (E.19)$$

which is stated in (5.17) of Theorem 5.1. After the G_i 's are determined, g_1 and g_2 can be solved from (E.9) and (E.11) respectively

$$\begin{aligned} g_i + D_{ii}^{-1}D_{ij}g_j = & -D_{ii}^{-1}[D_{ii}G_i\bar{y}_i(t) + D_{ij}G_j\bar{y}_j(t) \\ & + C_i^T\bar{y}_i(t) + x_i(t)], \quad i, j = 1, 2, \quad i \neq j \end{aligned} \quad (E.20)$$

as stated in (5.18) of Theorem 5.1. Sufficient conditions for existence of solution to (E.16) is discussed in the proof of Theorem 3 in [11] and Corollary 1.1 in [12] and is stated in (5.16).

VITA

Yick Man Chan was born on October 15, 1956 in Hong Kong. He received his Bachelor of Science degree in Mathematics from Rose-Hulman Institute of Technology, Terre Haute, Indiana in 1977 and the Master of Science degree in Electrical Engineering from the University of Illinois at Urbana-Champaign in 1979. From August 1977 until August 1981, he was a research assistant at the Coordinated Science Laboratory, University of Illinois, Urbana.

Mr. Chan is a member of the Institute of Electrical and Electronic Engineers and the honor societies of Tau Beta Pi, Eta Kappa Nu, and Pi Mu Epsilon.