DECENTRALIZED STOCHASTIC ADAPTIVE NASH GAMES

Y. M. CHAN AND J. B. CRUZ, JR.
Decision and Control Laboratory, Coordinated Science Laboratory, University of Illinois, 1101 W. Springfield Ave., Urbana, Illinois 61801, U.S.A.

SUMMARY

The optimization of stochastic systems with unknown parameters and multiple decision-makers or controllers each having his own objective is considered. Two explicit self-tuning type algorithms are proposed for decentralized stochastic adaptive Nash games under the 'one-step-delay sharing pattern'. The first algorithm is an ad hoc constraint on the policy form, whereas the second one is based on an extension from static Nash game theory. Simulation results on a simplified economic system indicate that these algorithms are capable of stabilizing a system along targeted paths.

KEY WORDS  Stochastic differential games  Decentralized control  Adaptive control  Self-tuning algorithms  Nash equilibria  Economic models

1. INTRODUCTION

Systems such as distributed industrial systems, power and energy systems, environmental systems and socioeconomic systems are large-scale systems which have the following features: (i) unknown or partial knowledge of system dynamics that may be time-varying; (ii) presence of unmeasurable disturbances; (iii) presence of multiple decision-makers or controllers, each of whom has his own objective; (iv) various decision-makers having different sets of information or measurements about the system. One solution concept for the optimization of these systems is the Nash strategy concept,\(^1\)\(^-\)\(^3\) and the situation where the concept is used is called a Nash game.

The decision-makers or controllers in a Nash game simultaneously minimize their respective cost functionals 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.

Definition 1

A strategy set \(\{u^*_1, u^*_2, \ldots, u^*_N\}\) is a Nash equilibrium strategy set if

\[
J_i(u^*_1, \ldots, u^*_{i-1}, u^*_i, u^*_{i+1}, \ldots, u^*_N) \leq J_i(u^*_1, \ldots, u^*_{i-1}, u_i, u^*_{i+1}, \ldots, u^*_N), \quad i = 1, 2, \ldots, N
\]

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

As in stochastic optimal control problems, there are different solution concepts, open-loop and closed-loop solutions, to the dynamic Nash game problem.\(^2\) In general, the open-loop and closed-loop solutions, to the dynamic Nash game problem.\(^2\) 2

© 1983 by John Wiley & Sons, Ltd.  
Received 6 September 1981  
Revised 26 July 1982
closed-loop solutions of a game problem are different. However, by restricting the cost functions of the decision-makers to a single-stage, the distinction between the different types of solutions ceases to exist, since we have essentially reduced the problem to a static framework. In fact, this type of cost functional has been widely adopted by various authors. Numerous industrial applications which use this concept of a single-stage cost function have been launched successfully. Moreover, the single-stage cost function is used repetitively to arrive at an asymptotic algorithm. This is our primary objective. We are essentially looking at an infinite-horizon problem one step at a time. This in turn enables us to decompose the potentially complex information structure to a simpler one.

In this paper, explicit self-tuning methods are used to develop algorithms for decentralized adaptive Nash games with an information structure termed 'one-step-delay sharing pattern'. The decentralized system framework is suitable for analysing 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 and was further extended later. In Section 2, we will formulate the decentralized Nash game problem. In Section 3 we approach the known-parameter problem by a straightforward constraint on the form of the control policy as has been done similarly for the single-controller problem. In Section 4, another approach is used to tackle the known-parameter decentralized game problem. Specifically, we extend results of static Nash games 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 estimates of the true parameters into the control law. The reason that the certainty equivalence method, as opposed to some dual control concept, is being used is that we are primarily interested in asymptotic solutions. Certainty equivalence is more appropriate in our context since the concept of dual control is usually applied to problems of finite horizon. If we are dealing with constant parameters, dual strategies and certainty equivalence strategies will result in the same asymptotic result. Finally, certainty equivalence does hold true for certain classes of problems such as the standard LQG system. The proposed algorithm can be classified as an explicit self-tuning strategy since the system 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 a simplified economic system, which is presented in Section 5, do show that the algorithm is capable of stabilizing the system along desired paths asymptotically. Furthermore, our simulation results indicate that the two different decentralized approaches will generate the same optimal policy. This hints that the two methods may actually be equivalent.

2. PROBLEM FORMULATION

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

\[
y(t+1) = a(q^{-1}) y(t) + B(q^{-1}) u(t) + e(t+1) + D
\]

(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 the same as that of \( u_i \). 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

$$a(z) = a_0 + a_1 z + \ldots + a_n z^n$$

$$B(z) = B_0 + B_1 z + \ldots + B_n z^n$$

The $D$ in (1) is a $p$-dimensional offset vector with $i$th row block $D_{oi}$, which is also the same dimension as $u_i$. A steady-state decentralized Nash equilibrium strategy for the system is sought. The cost function of each decision-maker is given by

$$J_i^0 = E\{[y_i(t+1) - y_i(t+1)]^T Q_i[y_i(t+1) - y_i(t+1)]$$

$$+ [u_i(t) - u_i(t-1)]^T R_i[u_i(t) - u_i(t-1)]\}, \quad i = 1, 2, \ldots, N$$

where $Q_i$ is symmetric positive semidefinite, $R_i$ is symmetric positive definite and $y_i$ is the desired value of the $i$th output $y_i$. Penalizing the control difference $[u_i(t) - u_i(t-1)]$ tends to favour non-abrupt changes in control. It eliminates steady-state errors in the output due to constant disturbances because of accumulator action.

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 $y(t-1), y(t-2), \ldots$; (ii) past inputs $u(t-1), u(t-2), \ldots$ as his information. This class of information pattern is called a 'one-step-delay sharing pattern'. Besides the theoretical implication of this information pattern (which is discussed in detail in Basar$^9$), this concept has practical significance. The one-step delay in information flow may allow, for example, simpler or less restrictive communication links between controllers. Another possibility is that we may use a slower central computer for parameter estimators, whereas the individual controllers may use faster computers for computation of their optimal policies (given the estimated parameters). This possibility arises in our formulation since we require every controller to use the same parameter estimation algorithm and to start with identical initial conditions. The $i$th decision-maker attempts, under this information structure, to minimize (4) with respect to $u_i(t)$ with the assumption that the other decision-makers use the Nash equilibrium strategy.

To facilitate our analysis, the cost functional (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 (4) can be written in the following form:

$$J_i^0 = E\{u_i^T(t) D_{ii} u_i(t) + 2u_i^T(t) D_{ij} u_j(t)$$

$$+ 2a_0 y_i^T(t) Q_i(B_0)_{ii} u_i(t) + 2[a(q^{-1})] y_i(t)$$

$$+ \tilde{B}_i(q^{-1}) u_i(t) + \tilde{B}_i(q^{-1}) u_i(t)$$

$$+ D_i - y_i(t+1)]^T Q_i(B_0)_{ii} u_i(t)$$

$$+ 2[u_i^T(t-1) R_i u_i(t)]\}$$

$$+ \text{terms not involving } u_i(t), \quad i, j = 1, 2, \quad i \neq j$$

with

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

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

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

$$\tilde{B}_i(z) = B_i(z) - (B_0)_{ii}$$
The scalar polynomial $\tilde{a}(z)$ is similarly defined as

$$\tilde{a}(z) = a(z) - a_0$$  \hfill (9)

We let $J_i$ denote the 'active' part of $J_i^0$ in (5), that is, the part that involves $u_i(t)$. Hence, we have

$$J_i^0 = J_i + \text{terms not involving } u_i(t)$$  \hfill (10)

3. CONSTRAINED DECENTRALIZED NASH GAME

Consider a system governed by (1) in which the $i$th controller tries to minimize $J_i$ given in (10), $i = 1, 2$. Notice that this formulation is equivalent to minimizing $J_i^0$ by the $i$th controller $u_i$ since $J_i^0$ has been decomposed into $J_i$ (which is the 'active' part of $J_i^0$) and terms not involving $u_i(t)$.

Let the matrix $C_i$ and the function $x_i$ be defined by

$$C_i = a_0 Q_i(B_0)B$$  \hfill (11)

$$x_i(t) = (B_0)^T_i Q_i[\tilde{a}(q^{-1})y_i(t) + \tilde{B}_i(q^{-1})u_i(t)$$

$$+ \tilde{B}_i(q^{-1})u_i(t) + D_i - y_i(t+1)]$$

$$+ R_i u_i(t-1), \quad i, j = 1, 2, \quad i \neq j$$  \hfill (12)

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

$$J_i = E\{u_i^T(t)D_{ii}u_i(t) + 2u_i^T(t)D_{ij}u_j(t) + 2y_i^T(t)C_iu_i(t) + 2x_i^T(t)u_i(t)\}, \quad i, j = 1, 2, \quad i \neq j$$  \hfill (13)

The constrained decentralized Nash equilibrium strategy for the system (1) with cost functionals $J_i(i = 1, 2)$ in (13) is first presented in Section 3.1. All system parameters are assumed to be known. Then, certainty equivalence is invoked heuristically, and a stochastic approximation-type estimation scheme is used to derive estimates that are substituted into the optimal policy in place of the true parameters.

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 according to our formulation. The $i$th decision-maker minimizes $J_i$ with respect to $u_i$ which is restricted to the form

$$u_i(t) = G_i y_i(t) + g_i, \quad i = 1, 2$$  \hfill (14)

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 1

Let the characteristic root of matrix $A$ with maximum absolute value be denoted by $\lambda_m(A)$. Then, if

$$|\lambda_m(D_{11}^{-1}D_{12}D_{22}^{-1}D_{21})| < 1$$  \hfill (15)

the system (1) with cost functions (13) will admit a unique Nash solution of the form (14). The
gains $G_i$ and $g_j$ satisfy the following.

$$G_i - D_{ii}^{-1}D_{ij}D_{jj}^{-1}D_{ji}G_i W_{ij}W_{jj}^{-1} W_{ii}^{-1}$$

$$= -D_{ii}^{-1}C_i^T + D_{ij}^{-1}D_{ij}C_j W_{ji}W_{jj}^{-1}, \quad i, j = 1, 2, \quad i \neq j$$

(16)

$$g_i + D_{ii}^{-1}D_{ij}g_j = -D_{ii}^{-1}[D_{ii}G_i \tilde{y}(t) + D_{ij}G_j \tilde{y}(t) + C_i^T \tilde{y}(t) + x_i(t)], \quad i, j = 1, 2, \quad i \neq j$$

(17)

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

Proof. See Appendix I.

Notice that the gains $G_i$ are dependent on $a_0$, $B_0$, $Q_i$ and $R_i$ only. Hence, when the parameters $a_0$, $B_0$ are known, once $G_i$ is determined, it does not require further recomputation.

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 (16). 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 multimodelling can be avoided.

The recursive procedure of Ljung,\(^{16}\) which is a stochastic approximation-type algorithm, can be used to estimate the system parameters explicitly. We now introduce the parameter matrix $\Theta$ defined by

$$\Theta = [\theta_1, \theta_2, \ldots, \theta_p] = \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_n \\ \beta_0^T \\ \vdots \\ \beta_p^T \\ \varphi_1^T, \ldots, \varphi_p^T \end{bmatrix}$$

(18)

where each $\alpha_i (i = 0, 1, 2, \ldots, n)$ is a diagonal matrix. The following recursions are carried out at each step of time to estimate $\theta_j$, $j = 1, 2, \ldots, p$:

$$\theta_j(t) = \theta_j(t-1) + \gamma(t) \frac{r_j(t)}{\gamma(t)} [y_j(t) - y_j(t-1) \theta(t-1)]$$

(19)

$$r_j(t) = r_j(t-1) + \gamma(t) [\eta_j(t-1) \eta_j(t-1) - r_j(t-1)], \quad r_j(0) = 1$$

(20)

$$\eta_j(t-1) = [y_j(t-1), \ldots, y_j(t-n), u(t-1), \ldots, u(t-n)]^T$$

(21)

with $\gamma(t)$ being a decreasing sequence in $t$. Notice that the assumption that $a_0$ and $B_0$ are known will lead to the setting of $\alpha_0 = a_0 I$ and $\beta_0 = B_0$. A block diagram of the closed-loop system is shown in Figure 1 with $v_i(t)$ deleted from the diagram.

Convergence of the algorithm is not guaranteed. Intuitively, convergence of the estimator (viz. $\tilde{\theta} \rightarrow \theta$) will result in the correct computation of the Nash strategy. Hence, one way to tackle the adaptive Nash game convergence problem may be first to establish the convergence condition for
the estimation scheme assuming that a form of separation principle exists between the controller and estimator. After the convergence of the estimator is established, the entire closed-loop system may be investigated for Nash convergence. The condition for convergence for the filtering equations (19)–(21) has been investigated. It is shown that if \( w(t) \) is a white noise process, then the estimates will yield a correct description of the input–output data. An additional problem that may be encountered in multivariable systems is that there may be different sets of estimates that yield the same description of the system. Hence, suitable identifiability conditions may be required to ensure proper convergence of the estimator. The convergence property of the \textit{ad hoc} algorithm is still to be investigated.
4. EXTENDED STATIC DECENTRALIZED NASH GAME

In this section, another ad hoc approach is taken to tackle the decentralized stochastic adaptive Nash game problem. We will first solve the known-parameter case by applying certain static Nash game results. Then, the estimation scheme (18)–(21) is used to obtain explicitly the system parameters which are then substituted into the optimal policy for the known parameter case.

In order to use the results from Basar, reformulation of our present problem will be required. Specifically, (i) we shall rewrite our cost function in another form, and (ii) we shall also inject Gaussian white noise into our system so that the solution of our problem will be more transparent. After we establish the necessary result, we shall indicate how this artificially introduced white noise can be eliminated from our system without affecting our established result.

The cost function (13) can be rewritten as

\[ J_i = E\{ u_i^T(t) D_{ii} u_i(t) + 2 u_i^T(t) D_{ij} u_j(t) + 2 x^T(t) C_i^T u_i(t) \}, \quad i, j = 1, 2, \quad i \neq j \]  

(22)

with \( D_{ij} \) as defined in (6) and (7) and

\[
 C_i^T = \begin{bmatrix} \bar{1} \\ \bar{0} \end{bmatrix} \} m_i, \quad C_2^T = \begin{bmatrix} 0 \\ 1 \end{bmatrix} m_2
\]

(23)

where \( m_i \) = the dimension of \( u_i \) and

\[ x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \]

(24)

where

\[ x_i(t) = (B_0^T Q_i [a_0 y_1(t) + \bar{a}(q^{-1}) y(t) + \bar{B}_i(q^{-1}) u_i(t) + \bar{B}_i(q^{-1}) u_j(t) + D_i - y_j(t+1)] + R_i u_i(t-1), \quad i, j = 1, 2, \quad i \neq j \]  

(25)

Notice that \( x_i(t) \) is a transformation of the current measurement \( y_i(t) \) (which is available only to the \( i \)th decision-maker at time \( t \) plus other past input–output data (which are available to every decision-maker under the 'one-step-delay sharing pattern').

4.1. Extended static decentralized Nash solution

To solve our present problem, we consider \( x(t) \) as the state vector in a state-space representation of a system in which the \( i \)th controller has \( z_i \) as his measurement. This measurement vector \( z_i(t) \) is given by

\[ z_i(t) = H_i x(t) + v_i(t), \quad i = 1, 2 \]

(26)

where

\[
 H_i = [ \begin{array}{cc} m_1 & m_2 \\ \bar{1} & \bar{0} \end{array} ] m_i, \quad H_2 = [ \begin{array}{cc} m_1 & m_2 \\ 0 & 1 \end{array} ] m_2
\]

(27)

and \( v_i(t) \) is zero-mean white Gaussian noise with positive semidefinite covariances \( T_i, i = 1, 2 \). This \( v_i(t) \) is our artificially-injected noise vector mentioned earlier. Its presence is just a means to bring us to our desired result. It will be eliminated from our formulation (that is, by letting \( T_i = 0, i = 1, 2 \)) once the result is established. This result is stated in the following theorem.
Theorem 2

Let the matrices \((B_0)_{ii}^T Q_i, i = 1, 2\), be non-singular. The condition (16) in Theorem 1,

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

is sufficient for the system (1) with cost functions (22) to admit a unique Nash solution. The control law of each decision-maker is given by

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

(28)

with

\[G_i = -[I - D_i(t)^{-1} D_{ij} D_{jj}^{-1} D_{ji}]^{-1} D_{ii}^{-1}[C_i - D_{ij} D_{jj}^{-1} C_j], \quad i, j = 1, 2, \quad i \neq j\]  

(29)

\[\dot{x}_i(t) = E[x(t) | z_i(t)] = \bar{x}(t) + QH_i^T (H_i QH_i^T + T_i)^{-1} (z_i(t) - H_i \bar{x}(t)), \quad i = 1, 2\]  

(30)

where

\[\bar{x}_i(t) = ([B_0]_{ii}^T Q_i [a_0 \bar{y}_i(t) + \bar{a}(q^{-1}) y_i(t) + \bar{B}_i(q^{-1}) u_i(t) + \bar{B}_i(q^{-1}) u_j(t) + D_i \bar{y}_i(t + 1)] + R_i u_i(t - 1), \quad i, j = 1, 2, \quad i \neq j\]  

(31)

\[Q = a_0^2 \begin{pmatrix} \begin{pmatrix} B_0 \end{pmatrix}_{11} Q_1 W_{11} Q_1(B_0)_{11} & \begin{pmatrix} B_0 \end{pmatrix}_{12} Q_1 W_{12} Q_2(B_0)_{22} \\ \begin{pmatrix} B_0 \end{pmatrix}_{22} Q_2 W_{21} Q_1(B_0)_{11} & \begin{pmatrix} B_0 \end{pmatrix}_{22} Q_2 W_{22} Q_2(B_0)_{22} \end{pmatrix} \end{pmatrix}\]  

(32)

\[T_i = 0, \quad i = 1, 2\]  

(33)

Also \(F_1\) is the unique solution to

\[F_1 + PF_1 L = M\]  

(34)

where

\[P = -D_{11}^{-1} D_{12} D_{22}^{-1} D_{21}\]  

(35)

\[L = QH_1^T (H_1 QH_1^T + T_1)^{-1} H_1 QH_2^T (H_2 QH_2^T + T_2)^{-1} H_2\]  

(36)

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

(37)

and

\[F_2 = -D_{22}^{-1} D_{21} F_1 QH_1^T (H_1 QH_1^T + T_1)^{-1} H_1 - D_{22}^{-1} C_2\]  

(38)

Proof. The proof of Theorem 2 is given by Basar. The non-singularity of \((B_0)_{ii}^T Q_i, i = 1, 2\) will ensure that \((H_i QH_i^T + T_i)\) is invertible.

There is a close resemblance between Theorems 1 and 2. In both cases, the control \(u_i(t)\) is affine in \(y_i(t)\). Moreover, as in Theorem 1, the gain for the information, \(F_i\), is dependent upon the system parameters \(a_0\) and \(B_0\) but not on the rest of \(a(z)\) or \(B(z)\).

4.2. Self-tuning extended static decentralized Nash games

In order to avoid the possibility of non-existence of solutions, we assume that the system parameters \(a_0\) and \(B_0\) are known, whereas the rest of \(a(z)\), \(B(z)\) and \(D\) are unknown. As in the previous approach, the system parameters are estimated recursively. We use equations (19)–(22) and assume identical algorithms and initial conditions for all decision-makers. The estimates are then substituted into the equations of Theorem 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 3.2.
A block diagram of the closed-loop system is shown in Figure 1. The structures for the two different approaches are almost identical except for the noise sequences \( \{v_i(t)\} \) \( (i = 1, 2) \) that are introduced in the second method.

5. SIMULATION EXAMPLE

The decentralized Nash strategies proposed in this paper are now applied to a simplified economic model with two inputs and two outputs. However, owing to the simplified nature of the model, this example is intended only to gain some insight into the numerical aspects of the algorithms; there is no intent to suggest that the policy obtained here be implemented. 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 Reference 23. In the United States, the formulation of the monetary policy is in the domain of the Federal Reserve System (FRS) whereas the formulation of the fiscal policy is primarily in the hands of the Congress and the President. 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, and this 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

\[
J_i = E\{[y_i(t+1) - y'_i(t+1)]^T Q_i[y_i(t+1) - y'_i(t+1)] + [u_i(t) - u'_i(t-1)]^T R_i[u_i(t) - u'_i(t-1)]\}, \quad i = 1, 2
\]

where \( y(t) = [C(t), I(t)]^T \), \( u(t) = [G(t), M(t)]^T \) and \( y' \) is the desired output. We assume \( y'(t) \) grows at an annual rate of 4 per cent from \( y'(0) = [300, 75]^T \).

With respect to the decentralized aspects of the system, we assume that the government, which controls \( u_1(t) \), has consumption expenditure \( y_1(t) \) as its measurement and the Federal Reserve System, which controls \( u_2(t) \), has private investment \( y_2(t) \) as its measurement. One-step-delayed input–output data are available to both decision-makers. 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 whereas the FRS focuses its entire attention on ensuring that a targeted path of current investment is followed.

The system is governed by

\[
a(q^{-1}) y(t) = B(q^{-1}) u(t-1) + C e(t) + D
\]

where the numerical values of \( a(z) \), \( B(z) \), \( C(z) \) and \( D \) are listed in Appendix II. The weighting matrices and the covariance of the noise are also included in Appendix II. We assume that the \( a_0 \) and \( B_0 \) parameters are known during the simulation.

Simulation results indicate that both adaptive procedures stabilize the system along the targeted growth paths. The output time responses of a typical run using the constrained policy approach and the corresponding trajectories using the extended static Nash game approach (with \( T_i = 0 \)) are shown in Figures 2 and 3. The two sets of responses indicate that the two methods generate exactly the same trajectories. We conjecture that the two methods are equivalent.

The output responses of the same run with all the parameters known are shown in Figures 4 and 5. The algorithms converge very fast for this instance. To compare the error in the optimal policy, we let \( u^i(t) \) denote the controls obtained with known parameters and \( u^n(t) \) denote
Figure 2. Time response of $y_1$ using constrained policy or extended static game

Figure 3. Time response of $y_2$ using constrained policy or extended static game
Figure 4. Time response of $y_1$ with known parameters

Figure 5. Time response of $y_2$ with known parameters
Figure 6. Time response of policy error $e_1^u$

Figure 7. Time response of policy error $e_2^u$
the policy obtained with unknown parameters. The time response of the quantity
\( e^u(t) = u^u(t) - u^u(t) = [e_1^u(t), e_2^u(t)]^T \) is shown in Figures 6 and 7. We observe that the policy error \( e^u(t) \) seems to be a zero-mean quantity, which indicates the adaptive algorithms are providing accurate controls even though the parameter estimates for the model are far from converging.

6. CONCLUSION

Two explicit self-tuning type methods have been proposed to deal with decentralized stochastic adaptive Nash games under the 'one-step-delay information sharing pattern'. Simulation results indicate that the two methods generate the same trajectories. This indicates that they may be equivalent algorithms. The simulation example also indicates the capability of the algorithms to stabilize certain systems along targeted paths. However, convergence of the algorithms have not been established as yet, and a theoretical basis for the convergence of the decentralized Nash game problem still needs to be investigated.

ACKNOWLEDGEMENTS

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

APPENDIX I: PROOF OF THEOREM 1

Consider the system governed by (1) with the cost functions (13). The policy \( u_i \) is constrained to be of the form (14). Without loss of generality, let us consider \( J_1 \). Substituting \( u_1 \) in (14) into \( J_1 \) yields

\[
J_1 = E\{y_1^2(t) G_1^T D_{11} G_1 y_1(t) + g_1^T D_{11} g_1 + 2y_1^2(t) G_1^T D_{12} g_1 \\
+ 2[y_1^2(t) G_1^T D_{12} G_2 y_2(t) + y_1^2(t) G_1 D_{12} g_2 \\
+ g_1^T D_{12} g_2 y_2(t) + g_1^T D_{12} g_2] + 2[y_1^2(t) C_1 G_1 y_1(t) \\
+ y_1^2(t) C_1 g_1 + x_1^2(t) G_1 y_1(t) + x_1^2(t) g_1]\}\tag{39}
\]

Denote

\[
E\{y_i(t) y_i^2(t)\} = P_i(t)
\]
\[
E\{y_i(t)\} = \bar{y}_i(t)
\]

Then, taking the expectation of the terms in (39), we have

\[
J_1 = \text{trace} \left[ G_1^T D_{11} G_1 P_{11}(t) + g_1^T D_{11} g_1 + 2\bar{y}_1^2(t) G_1^T D_{12} g_1 \\
+ 2G_1^T D_{12} G_2 P_{22}(t) + 2\bar{y}_1^2(t) G_1^T D_{12} g_2 + 2g_1^T D_{12} G_2 \bar{y}_2(t) \\
+ 2g_1^T D_{12} g_2 + 2C_1 G_1 P_{11}(t) + 2\bar{y}_1^2(t) C_1 g_1 \\
+ 2x_1^2(t) G_1 \bar{y}_1(t) + 2x_1^2(t) g_1 \right]
\]
The following formulae are then used to evaluate the necessary conditions for a minimum: 
\( \partial J / \partial \mathbf{G}_1 = 0 \) and \( \partial J / \partial \mathbf{g}_1 = 0 \):

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

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

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

\[
\frac{\partial}{\partial \mathbf{Z}} \text{tr} [\mathbf{Z}^T \mathbf{LZN}] = \mathbf{L}^T \mathbf{ZN}^T + \mathbf{LZN}
\]

Hence, we have

\[
0 = \frac{\partial J_1}{\partial \mathbf{G}_1}
\]

\[
= \mathbf{D}_{11}^T \mathbf{G}_1 \mathbf{P}_{11}^T(t) + \mathbf{D}_{11} \mathbf{G}_1 \mathbf{P}_{11}(t) + 2\mathbf{D}_{11} \mathbf{g}_1 \mathbf{y}_1^T(t) + 2 \mathbf{D}_{12} \mathbf{G}_2 \mathbf{P}_{21}(t)
\]

\[
+ 2 \mathbf{D}_{12} \mathbf{g}_2 \mathbf{y}_1^T(t) + 2 \mathbf{C}_1^T \mathbf{P}_{11}(t) + 2\mathbf{c}_1(t) \mathbf{y}_1^T(t)
\]

or

\[
0 = \mathbf{D}_{11} \mathbf{G}_1 \mathbf{P}_{12}(t) + \mathbf{D}_{11} \mathbf{g}_1 \mathbf{\bar{y}}_1^T(t) + \mathbf{D}_{12} \mathbf{G}_2 \mathbf{P}_{21}(t) + \mathbf{D}_{12} \mathbf{g}_2 \mathbf{\bar{y}}_1^T(t) + \mathbf{C}_1^T \mathbf{P}_{11}(t) + \mathbf{x}_1(t) \mathbf{\bar{y}}_1^T(t) \quad (43)
\]

and

\[
0 = \frac{\partial J_1}{\partial \mathbf{g}_1}
\]

\[
= 2 \mathbf{D}_{11} + 2 \mathbf{D}_{11}^T \mathbf{G}_1 \mathbf{\bar{y}}_1(t) + 2 \mathbf{D}_{12} \mathbf{G}_2 \mathbf{\bar{y}}_2(t) + 2 \mathbf{D}_{12} \mathbf{g}_2 + 2 \mathbf{C}_1^T \mathbf{\bar{y}}_1(t) + 2\mathbf{x}_1(t)
\]

or

\[
0 = \mathbf{D}_{11} \mathbf{g}_1 + \mathbf{D}_{11} \mathbf{G}_1 \mathbf{\bar{y}}_1(t) + \mathbf{D}_{12} \mathbf{G}_2 \mathbf{\bar{y}}_2(t) + \mathbf{D}_{12} \mathbf{g}_2 + \mathbf{C}_1^T \mathbf{\bar{y}}_1(t) + \mathbf{x}_1(t) \quad (44)
\]

Similarly, from \( \partial J_2 / \partial \mathbf{G}_2 = 0 \) and \( \partial J_2 / \partial \mathbf{g}_2 = 0 \), we obtain

\[
0 = \mathbf{D}_{22} \mathbf{G}_2 \mathbf{P}_{22}(t) + \mathbf{D}_{22} \mathbf{g}_2 \mathbf{\bar{y}}_2^T(t) + \mathbf{D}_{21} \mathbf{G}_1 \mathbf{P}_{12}(t) + \mathbf{D}_{21} \mathbf{g}_1 \mathbf{\bar{y}}_2^T(t) + \mathbf{C}_1^T \mathbf{P}_{22}(t) + \mathbf{x}_2(t) \mathbf{\bar{y}}_2^T(t) \quad (45)
\]

\[
0 = \mathbf{D}_{22} \mathbf{g}_2 + \mathbf{D}_{22} \mathbf{G}_2 \mathbf{\bar{y}}_2(t) + \mathbf{D}_{21} \mathbf{G}_1 \mathbf{\bar{y}}_1(t) + \mathbf{D}_{21} \mathbf{g}_1 + \mathbf{C}_2^T \mathbf{\bar{y}}_2(t) + \mathbf{x}_2(t) \quad (46)
\]

Since \( \mathbf{\bar{y}}(t) \) and \( E\{\mathbf{y}(t) \mathbf{y}^T(t)\} \) are required to solve for \( \mathbf{G}_i \) and \( \mathbf{g}_i \) \( (i = 1, 2) \) in (43)–(46), we show how these terms are computed.

At time \( t \), past \( \mathbf{u} \) and \( \mathbf{y} \) are known so that

\[
E\{\mathbf{y}(t)\} = E\{a(q^{-1}) \mathbf{y}(t-1) + \mathbf{B}_l(q^{-1}) \mathbf{u}(t-1) + \mathbf{B}_l(q^{-1}) \mathbf{u}(t-1) + \mathbf{e}(t) + \mathbf{D}\}
\]

\[
= a(q^{-1}) \mathbf{y}(t-1) + \mathbf{B}_l(q^{-1}) \mathbf{u}(t-1) + \mathbf{B}_l(q^{-1}) \mathbf{u}(t-1) + \mathbf{D} \quad (47)
\]

Let

\[
\mathbf{P}(t) = \begin{bmatrix} \mathbf{P}_{11}(t) & \mathbf{P}_{12}(t) \\ \mathbf{P}_{21}(t) & \mathbf{P}_{22}(t) \end{bmatrix} = E\{\mathbf{y}(t) \mathbf{y}^T(t)\} = \text{variance} \{\mathbf{y}(t)\}.
\]
Since variance \( \{y(t)\} = \text{covariance} \{y(t)\} + \bar{y}(t) \bar{y}^T(t) \), we have
\[
P(t) = E\{(y(t) - \bar{y}(t))(y(t) - \bar{y}(t))^T\} + \bar{y}(t) \bar{y}^T(t)
= E\{e(t)e^T(t)\} + \bar{y}(t) \bar{y}^T(t)
= W + \bar{y}(t) \bar{y}^T(t)
\] (48)

Postmultiplying (44) by the term \( \bar{y}^T_1(t) \) and subtracting the resulting equation from (43), we obtain
\[
D_{11} G_1 [P_{11}(t) - \bar{y}_1(t) \bar{y}_1^T(t)] + D_{12} G_2 [P_{21}(t) - \bar{y}_2(t) \bar{y}_2^T(t)] = -C_1^T (P_{11}(t) - \bar{y}_1(t) \bar{y}_1^T(t))
\]
\[
D_{11} G_1 W_{11} + D_{12} G_2 W_{21} = -C_1^T W_{11}
\] (49)

Similarly, if \( \bar{y}^T_2(t) \) is postmultiplied to (46) and then subtracted from (45), we obtain
\[
D_{21} G_1 W_{12} + D_{22} G_2 W_{22} = -C_2^T W_{22}
\] (50)

After some manipulations with (49) and (50), we have (16) and (17) as stated in Theorem 1. Sufficient conditions for the existence of a solution to (16) are discussed in the proof of Theorem 3 of Reference 6 and in Corollary 1.1 of Reference 7.

**APPENDIX II: ECONOMIC MODEL**

The economic model used for the simulation study is taken from Chow. It is given in the following form
\[
C_t = 0.9266C_{t-1} - 0.0203I_{t-1} + 0.3190G_t + 0.4206M_t - 63.2386
\] (51)
\[
I_t = 0.1527C_{t-1} + 0.3806I_{t-1} - 0.0735G_t + 1.538M_t - 210.8994
\] (52)

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 that the current \( G_t \) and \( M_t \) are the result of, and are equal to, the desired levels that were specified in the previous time step. This accounts for the time lag in the definition of \( u_1 \) and \( u_2 \). The resulting model in terms of \( y \) and \( u \) is given in the following matrix polynomial form
\[
\begin{pmatrix}
1 & \begin{pmatrix}
0.9266 & -0.0203 \\
0.1527 & 0.3806
\end{pmatrix} q^{-1}
\end{pmatrix} y(t) = \begin{pmatrix}
0.3190 & 0.4206 \\
-0.0735 & 1.538
\end{pmatrix} u(t-1) - \begin{pmatrix}
63.2386 \\
210.8994
\end{pmatrix}
\] (53)

We can transform the above system into an equivalent system with a scalar \( a(z) \) polynomial. Hence, we have
\[
(1 - 1.3072q^{-1} + 0.3596q^{-2}) y(t) = \begin{pmatrix}
0.3190 & 0.4206 \\
-0.0735 & 1.538
\end{pmatrix}
+ \begin{pmatrix}
-0.1199 & -0.1913 \\
0.1168 & -1.3609
\end{pmatrix} q^{-1} u(t-1) + \begin{pmatrix}
-34.8887 \\
-25.1366
\end{pmatrix}
\] (54)

or
\[
a(q^{-1}) y(t) = B(q^{-1}) u(t-1) + D
\] (55)

For simulation, we assume the system is perturbed by zero-mean white noise \( e(t) \), that is,
\[
a(q^{-1}) y(t) = B(q^{-1}) u(t-1) + D + e(t)
\] (56)
with
\[
E\{e(t)e'^T(t)\} = \begin{bmatrix} 54 & 12 \\ 12 & 26 \end{bmatrix}
\]

The weighting matrices for the simulation are
\[
Q_1 = 5 \quad R_1 = 0.02
\]
\[
Q_2 = 10 \quad R_2 = 0.08
\]

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

REFERENCES