JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 114, No. 1, pp. 171–188, July 2002 (©2002)

Stochastic Discrete-Time Nash Games with Constrained State Estimators

A. R. Kian,[1] J. B. Cruz, Jr.,[2] and M. A. Simaan[3]

Abstract. In this paper, we consider stochastic linear-quadratic discrete-time Nash games in which two players have access only to noise-corrupted output measurements. We assume that each player is constrained to use a linear Kalman filter-like state estimator to implement his optimal strategies. Two information structures available to the players in their state estimators are investigated. The first has access to one-step delayed output and a one-step delayed control input of the player. The second has access to the current output and a one-step delayed control input of the player. In both cases, statistics of the process and statistics of the measurements of each player are known to both players. A simple example of a two-zone energy trading system is considered to illustrate the developed Nash strategies. In this example, the Nash strategies are calculated for the two cases of unlimited and limited transmission capacity constraints.

Key Words. Stochastic linear-quadratic systems, nonzero-sum discrete-time Nash games, optimal Nash strategies, estimator-controller, stochastic dynamic programming, Nash equilibrium.

1. Introduction

Stochastic Nash games in which the players have access only to noise-corrupted output measurements have been of interest since the late 1960s. Rhodes and Luenberger (Refs. 1, 2) studied linear-quadratic Gaussian (LQG) continuous-time zero-sum stochastic games with output measurements corrupted by additive independent white Gaussian noise. They considered a saddle-point solution to this problem under the constraint that each player is limited to a linear state estimator for generating his optimal

[1]Vice President of Engineering and Development, Genscape, Louisville, Kentucky.
[2]Howard D. Winbigler Chair in Engineering and Professor, Department of Electrical Engineering, Ohio State University, Columbus, Ohio.
[3]Bell of PA/Bell Atlantic Professor, Department of Electrical Engineering, University of Pittsburgh, Pittsburgh, Pennsylvania.

controls. Saksena and Cruz (Ref. 3) extended these results to nonzero-sum Nash games and derived a method for obtaining a near-equilibrium limiting solution for the Nash strategies under a multimodel situation. They also studied stochastic control of singularly perturbed linear systems with multiple decision-makers possessing different observations under a multimodel situation. In all of the above-mentioned research works (Ref. 1–3), the authors assumed that the separation principle holds.

In this paper, we investigate discrete-time nonzero-sum LQG Nash games with constrained state estimators and two different information structures. In the first information structure, we assume that, at each stage, each player has access to his output measurements and his own control input at the previous stage. In the second information structure, we assume that, at each stage, each player has access to his output measurements at the current stage and his own control input at the previous stage. We show that, for these two information structures, the separation principle does not hold; therefore, the controller and estimator are interrelated and cannot be determined separately. In each case, we derive an algorithm for calculating the optimum estimators and controllers for the closed-loop Nash solution. The results are applied to a simple energy trading system with two generating firms as players.

2. Problem Formulation

Consider a two-player discrete linear game described by the following difference equation:

$$x_{k+1} = A_k x_k + B_k^1 u_k^1 + B_k^2 u_k^2 + w_k, \qquad k = 0, 1, \ldots, N-1, \tag{1}$$

where x_k is the state vector and u_k^1, u_k^2 are the control vectors of players 1, 2 respectively. We will assume that players 1 and 2 have noisy measurements of the form

$$y_k^1 = C_k^1 x_k + v_k^1, \tag{2a}$$

$$y_k^2 = C_k^2 x_k + v_k^2. \tag{2b}$$

The dimensions of $x_k, u_k^1, u_k^2, y_k^1, y_k^2$ are $n \times 1, m^1 \times 1, m^2 \times 1, p^1 \times 1, p^2 \times 1$ respectively and the matrices $A_k, B_k^1, B_k^2, C_k^1, C_k^2$ all have proper dimensions. The random disturbances w_k, v_k^1, v_k^2 in (2) are assumed to be independent, white, zero-mean, and Gaussian with covariances

$$\operatorname{cov}\{w_k\} = W_k, \tag{3a}$$

$$\operatorname{cov}\{v_k^1\} = N_k^1, \tag{3b}$$

$$\operatorname{cov}\{v_k^2\} = N_k^2. \tag{3c}$$

For simplicity, it is also assumed that both players know that the initial state x_0 is a Gaussian random vector (uncorrelated with w_k, v_k^1, v_k^2) with mean $\bar{x}_0$ and covariance

$$\operatorname{cov}\{x_0\} = \Gamma_0. \tag{4}$$

Consider the following discrete-time additive quadratic cost functions for players 1 and 2:

$$J_k^1(x_k, u_k^1, u_k^2) = E_{w_k}\left\{x_N' Q_N^1 x_N + \sum_{i=k}^{N-1} (x_i' Q_i^1 x_i + u_i^{1'} R_i^{11} u_i^1 + u_i^{2'} R_i^{12} u_i^2)\right\}, \tag{5a}$$

$$J_k^2(x_k, u_k^1, u_k^2) = E_{w_k}\left\{x_N' Q_N^2 x_N + \sum_{i=k}^{N-1} (x_i' Q_i^2 x_i + u_i^{1'} R^{21_i} u_i^1 + u_i^{2'} R_i^{22} u_i^2)\right\}, \tag{5b}$$

where the matrices Q_j^1, Q_j^2, R_j^{11}, R_j^{22}, R_j^{12}, R_j^{21} are all symmetric and positive semidefinite for $j \in [0, N]$. These represent costs-to-go functions from an arbitrary initial state x_k at stage k. We assume that the admissible controls are linear in the state estimates and that each player is restricted to use an n-dimensional linear estimator to generate his state estimate at each time step. We consider the following two information structures:

Type A: $X_k^i = \{y_{k-1}^i, u_{k-1}^i\}$ for player i, $i = 1, 2$ and $k = 1, \ldots, N-1$.

Type B: $X_k^i = \{y_k^i, u_{k-1}^i\}$ for player i, $i = 1, 2$ and $k = 1, \ldots, N-1$.

In both cases, there are no observations at $k = 0$. As mentioned earlier, both players know all the expected values and covariances of all the noises and the initial state.

3. Closed-Loop Nash for Information Structure Type A

For the information structure $X_k^1 = \{y_{k-1}^1, u_{k-1}^1\}$ for player 1 and $X_k^2 = \{y_{k-1}^2, u_{k-1}^2\}$ for player 2, we consider a state estimator for player 1 in the form

$$\hat{x}_{k+1}^1 = F_k^1 \hat{x}_k^1 + G_k^1[y_k^1 - C_k^1 \hat{x}_k^1] + D_k^1 u_k^1, \tag{6a}$$

and for player 2 in the form

$$\hat{x}_{k+1}^2 = F_k^2 \hat{x}_k^2 + G_k^2[y_k^2 - C_k^2 \hat{x}_k^2] + D_k^2 u_k^2. \tag{6b}$$

The objective function of player 1 is the conditional expectation $E\{J_k^1(x_k, u_k^1, u_k^2)|X_k^1\}$, and the objective function of player 2 is the conditional expectation $E\{J_k^2(x_k, u_k^1, u_k^2)|X_k^2\}$. The matrices F_k^1, G_k^1, D_k^1, the initial condition $\hat{x}_0^1$, and the closed-loop control law $u_k^1(\hat{x}_k^1)$ are to be selected by player 1, and the matrices F_k^2, G_k^2, D_k^2, the initial condition $\hat{x}_0^2$, and the closed-loop control law $u_k^2(\hat{x}_k^2)$ are to be selected by player 2, such that

$$E\{J_k^1(x_k, u_k^{1^*}, u_k^{2^*})|X_k^1\} \leq E\{J_k^1(x_k, u_k^1, u_k^{2^*})|X_k^1\}, \tag{7a}$$

$$E\{J_k^2(x_k, u_k^{1^*}, u_k^{2^*})|X_k^2\} \leq E\{J_k^2(x_k, u_k^{1^*}, u_k^2)|X_k^2\}. \tag{7b}$$

The pair of inequalities in (7) defines the closed-loop Nash equilibrium pair $\{u_k^{1^*}(\hat{x}_k^1), u_k^{2^*}(\hat{x}_k^2)\}$ for the information structure

$$X_k^i = \{y_{k-1}^i, u_{k-1}^i\}, \qquad i = 1, 2.$$

The admissible controls in (7) are assumed to be linear in the state estimates for all $k \in [0, N-1]$,

$$u_k^i = L_k^i \hat{x}_k^i, \qquad i = 1, 2, \tag{8}$$

where $\hat{x}_k^i$ for $i = 1, 2$ are the state estimates obtained from (6) and L_k^i for $i = 1, 2$ are $m^i \times n$ bounded closed-loop gain matrices to be determined so that (6) and (7) are satisfied.

We provide algorithms for computing the Nash-optimal feedback gains $L_k^i, i = 1, 2$, the matrices $F_k^1, G_k^1, D_k^1, F_k^2, G_k^2, D_k^2$, and the initial conditions $\hat{x}_0^1, \hat{x}_0^2$ in the state estimators equations (6). Consider the following set of coupled recursive equations in K_k^i, L_k^i, M_k, S_k^i:

$$K_N^i = Q_N^i, \qquad i = 1, 2, \tag{9a}$$

$$\begin{aligned} K_k^i = {} & Q_k^i + L_k^{j'} R_k^{ij} L_k^j + L_k^{i'} R_k^{ii} L_k^i \\ & + (\mathrm{A}_k + \mathrm{B}_k^i \mathrm{L}_k^i + \mathrm{B}_k^j \mathrm{L}_k^j)' \, \mathrm{K}_{k+1}^i (\mathrm{A}_k + \mathrm{B}_k^i \mathrm{L}_k^i + \mathrm{B}_k^j \mathrm{L}_k^j), \\ & \mathrm{i}, \mathrm{j} = 1, 2, \mathrm{i} \neq \mathrm{j}, \text{ and } k = 0, \ldots, N-1, \end{aligned} \tag{9b}$$

where

$$L_k^i = -(R_k^{ii} + B_k^{i'} K_{k+1}^i B_k^i)^{-1} B_k^{i'} K_{k+1}^i (A_k + B_k^j L_k^j S_k^j) \tag{10}$$

and

$$S_k^j = \begin{cases} I + (M_k^{ji} - M_k^{j0}) \times (M_k^{00} - M_k^{0i})^{-1}, & i, j = 1, 2, i \neq j, \text{ and } k = 1, \ldots, N-1, \\ I, \quad j = 1, 2 \text{ and } k = 0. \end{cases} \tag{11}$$

In the above expression, M_k is a symmetric nonnegative-definite matrix defined as follows:

$$M_k = E\{m_k m_k'\}, \qquad m_k' = [x_k', \tilde{x}_k^{1'}, \tilde{x}_k^{2'}] \tag{12}$$

$$\tilde{x}_k^i = x_k - \hat{x}_k^i, \qquad i = 1, 2. \tag{13}$$

and satisfying the following recursive equation:

$$M_{k+1} = \bar{A}_k M_k + M_k \bar{A}_k' + \bar{B}_k \bar{N}_k \bar{B}_k' + \bar{W}_k, \tag{14a}$$

with boundary conditions

$$M_0 = \begin{bmatrix} \bar{x}_0 \bar{x}_0' + \Gamma_0 & \Gamma_0 & \Gamma_0 \\ \Gamma_0 & \Gamma_0 & \Gamma_0 \\ \Gamma_0 & \Gamma_0 & \Gamma_0 \end{bmatrix}, \tag{14b}$$

where

$$M_k = \begin{bmatrix} M_k^{00} & M_k^{01} & M_k^{02} \\ M_k^{10} & M_k^{11} & M_k^{12} \\ M_k^{20} & M_k^{21} & M_k^{22} \end{bmatrix}, \tag{15}$$

$$\bar{A}_k = \begin{bmatrix} A_k + B_k^1 L_k^1 + B_k^2 L_k^2 & -B_k^1 L_k^1 & -B_k^2 L_k^2 \\ A_k + B_k^2 L_k^2 - F_k^1 & F_k^1 - G_k^1 C_k^1 & -B_k^2 L_k^2 \\ A_k + B_k^1 L_k^1 - F_k^2 & -B_k^1 L_k^1 & F_k^2 - G_k^2 C_k^2 \end{bmatrix}, \tag{16}$$

$$\bar{B}_k = \begin{bmatrix} 0 & 0 \\ G_k^1 & 0 \\ 0 & G_k^2 \end{bmatrix}, \tag{17}$$

$$\bar{N}_k = \begin{bmatrix} N_k^1 & 0 \\ 0 & N_k^2 \end{bmatrix}, \tag{18}$$

$$\bar{W}_k = \begin{bmatrix} W_k & 0 & 0 \\ 0 & W_k & 0 \\ 0 & 0 & W_k \end{bmatrix}. \tag{19}$$

Assume that the solution matrices $K_k^1, K_k^2, L_k^1, L_k^2, M_k, S_k^1, S_k^2$ for the above set of equations exist and are bounded for all $k \in [0, N-1]$. Furthermore, assume that the matrices $M_k^{00} - M_k^{01}, M_k^{00} - M_k^{02}$ are invertible for all $k \in [1, N-1]$, and that each player knows the objective function and the system dynamics of the other player. In other words, the matrices Q_k^1, Q_k^2,

R_k^{11}, R_k^{22}, R_k^{12}, R_k^{21}, A_k, B_k^1, B_k^2, C_k^1, C_k^1, N_k^1, N_k^2, W_k are known to both players. We have the following theorem.

Theorem 3.1. Given the linear system described by (1)–(4), the quadratic cost functions given by (5), the estimates of x_{k+1} at time $k+1$ as given by (6), and the definition of a Nash equilibrium solution as given by (7), if there exist sequences of bounded matrices K_k^i, L_k^i, M_k, S_k^i, for $i=1,2$ and $k=0,1,\ldots,N-1$, satisfying (9)–(19), then the feedback controls in (8) are Nash-optimal strategies where the matrices L_k^i are given by (10) and the optimal matrices F_k^1, F_k^2, G_k^1, G_k^2, D_k^1, D_k^2, M_k and initial conditions $(\hat{x}_0^1, \hat{x}_0^2)$ in the state estimators equations (6) are given by

$$F_k^1 = A_k + B_k^2 L_k^2 S_k^2, \tag{20a}$$

$$G_k^1 = M_k^{11} C_k^{1'} (N_k^1)^{-1}, \tag{20b}$$

$$D_k^1 = B_k^1, \tag{20c}$$

$$F_k^2 = A_k + B_k^1 L_k^1 S_k^1, \tag{21a}$$

$$G_k^2 = M_k^{22} C_k^{2'} (N_k^2)^{-1}, \tag{21b}$$

$$D_k^2 = B_k^2, \tag{21c}$$

$$\hat{x}_0^1 = \hat{x}_0^2 = \bar{x}_0. \tag{22}$$

Furthermore, the optimal conditional expectations of the cost-to-go functions at time k are given by

$$\begin{aligned} &E\{J_k^i(x_k, u_k^{i*}, u_k^{j*}) | X_k^i\} \\ &= E\{x_k' K_k^i x_k | X_k^i\} - \sum_{m=k}^{N-1} (\mathrm{tr}\{P_m^{ii} M_m^{ii}\} + \mathrm{tr}\{P_m^{ij} M_m^{jj}\}) \\ &\quad + \sum_{m=k}^{N-1} E\{w_m' K_{m+1}^i w_m\}, \qquad i,j=1,2,\ i \neq j, \text{ and } k=0,\ldots,N-1, \end{aligned} \tag{23a}$$

with terminal condition

$$E\{J_N^i(x_N) | X_N^i\} = E\{x_N' Q_N^i x_N | X_N^i\}, \qquad i=1,2, \tag{23b}$$

where

$$P_k^{ii} = L_k^{i'} (R_k^{ii} + B_k^{i'} K_{k+1}^i B_k^i) L_k^i, \qquad i=1,2, \tag{24}$$

$$P_k^{ij} = L_k^{j'} (R_k^{ij} + B_k^{j'} K_{k+1}^i B_k^j) L_k^j, \qquad i,j=1,2, \qquad i \neq j. \tag{25}$$

Remark 3.1. We note that the closed-loop Nash control laws

$$u_k^{1*} = L_k^1 \hat{x}_k^1, \tag{26a}$$

$$u_k^{2*} = L_k^2 \hat{x}_k^2 \tag{26b}$$

do not satisfy the separation principle. As it can be seen from equation (10), the calculations of the optimal control gains (L_k^1, L_k^2) depend on the components of the matrix M_k, which is related to the estimation error covariance matrix. Also, the estimator equations (21)–(22) depend on the controller gains. Thus, the controller and estimator cannot be separated. However, we should note that, if the measurements $y_k^1 = y_k^2$ in (2), then the matrices S_k^1 and S_k^2 in (11) will converge to the identity matrix and the controller gains will not depend on the state estimation errors. However, the estimator equations will still depend on the controller gains.

Proof of Theorem 3.1. First, we show that the estimator-controller for player 1 given by expressions (6a), (20), (22), (26a) is unbiased for all $k \in [0, N-1]$,

$$E\{x_k | X_k^1\} = \hat{x}_k^1. \tag{27}$$

A similar argument applies for the estimator-controller of player 2. By substituting (26) in the system difference equation (1), we have

$$x_{k+1} = (A_k + B_k^1 L_k^1 + B_k^2 L_k^2) x_k - B_k^1 L_k^1 \tilde{x}_k^1 - B_k^2 L_k^2 \tilde{x}_k^2 + w_k. \tag{28}$$

By subtracting (6a) from (28), using (20c) [note that, using (20c), the estimation error dynamics will not be a function of optimal control u_k^1], and after some algebraic manipulations, we have

$$\tilde{x}_{k+1}^1 = (A_k + B_k^2 L_k^2 - F_k^1) x_k + (F_k^1 - G_k^1 C_k^1) \tilde{x}_k^1 - B_k^2 L_k^2 \tilde{x}_k^2 - G_k^1 v_k^1 + w_k. \tag{29}$$

Similarly by subtracting (6b) from (28), using (21c) [note that, using (21c), the estimation error dynamics will not be a function of optimal control u_k^2], and after some algebraic manipulations, we have

$$\tilde{x}_{k+1}^2 = (A_k + B_k^1 L_k^1 - F_k^2) x_k + (F_k^2 - G_k^2 C_k^2) \tilde{x}_k^2 - B_k^1 L_k^1 \tilde{x}_k^1 - G_k^2 v_k^2 + w_k. \tag{30}$$

Now, we form the composite vectors m_k, v_k defined by

$$m_k' = [x_k', \tilde{x}_k^{1'}, \tilde{x}_k^{2'}], \tag{31}$$

$$v_k' = [v_k^{1'}, v_k^{2'}]. \tag{32}$$

From (28)–(30) and (16)–(19), we have

$$m_{k+1} = \bar{A}_k m_k - \bar{B}_k v_k + I \cdot w_k. \tag{33}$$

From the above expression, we can derive equation (14a) with boundary condition (14b) when M_k is defined by (12). If we consider (14)–(19) to write the components M^{01}_{k+1} and M^{11}_{k+1}, after some algebra and the use of (20a) and (20b), we have that, for all $k \in [0, N-1]$,

$$M^{11}_{k+1} - M^{01}_{k+1} = \Delta_k [M^{11}_k - M^{01}_k] + [M^{11}_k - M^{01}_k]\Lambda_k, \tag{34}$$

where Δ_k and Λ_k are known matrices. By choice of the initial condition (14b), $M^{11}_0 = M^{01}_0$, we conclude that

$$M^{11}_k = M^{01}_k, \qquad \text{for all } k \in [0, N-1].$$

Moreover, since M_k is a symmetric matrix, we have

$$M^{10}_k = M^{01}_k = M^{11}_k. \tag{35}$$

From the definition (12) of M_k and equation (35) for all $k \in [0, N-1]$, we have

$$E\{x_k(x'_k - \hat{x}^{1'}_k)\} = E\{(x_k - \hat{x}^1_k)(x'_k - \hat{x}^{1'}_k)\} = E\{(x_k - \hat{x}^1_k)x'_k\}, \tag{36}$$

which implies

$$E\{(x_k - \hat{x}^1_k)\hat{x}^{1'}_k\} = M^{01}_k - M^{11}_k = 0. \tag{37}$$

Thus, each component of the error $x_k - \hat{x}^1_k$ is orthogonal to each component of $\hat{x}^1_k$. Therefore, using standard estimation theory, given M_k [obtained from (14)] and X^1_k, the vector $\hat{x}^1_k$ can be regarded as the best estimate of x_k for all $k \in [0, N-1]$. That is,

$$E\{x_k | X^1_k\} = \hat{x}^1_k.$$

Similarly, it can be shown that

$$M^{20}_k = M^{02}_k = M^{22}_k, \qquad \text{for all } k \in [0, N-1],$$

which implies that

$$E\{x_k | X^2_k\} = \hat{x}^2_k.$$

By a procedure analogous to the above, given M_k [obtained from (14)] and X^1_k, we seek the matrix S^2_k such that the best estimate of $\hat{x}^2_k$ (given X^1_k) for all $k \in [0, N-1]$ is

$$E\{\hat{x}^2_k | X^1_k\} = S^2_k \hat{x}^1_k. \tag{38}$$

By standard estimation theory, we require that

$$E\{(\hat{x}^2_k - S^2_k \hat{x}^1_k)(S^2_k \hat{x}^1_k)'\} = 0, \tag{39a}$$

or

$$E\{\hat{x}^2_k \hat{x}^{1'}_k\} S^{2'}_k - S^2_k \cdot E\{\hat{x}^1_k \hat{x}^{1'}_k\} S^{2'}_k = [E\{\hat{x}^2_k \hat{x}^{1'}_k\} - S^2_k \cdot E\{\hat{x}^1_k \hat{x}^{1'}_k\}] \cdot S^{2'}_k = 0, \tag{39b}$$

or

$$E\{[(\hat{x}_k^2 - x_k) + x_k] \cdot \hat{x}_k^{1'}\} - S_k^2 \cdot E\{\hat{x}_k^1 \hat{x}_k^{1'}\} = 0. \tag{39c}$$

Let us calculate

$$\begin{aligned} E\{(\hat{x}_k^2 - x_k)\hat{x}_k^{1'}\} &= E\{(x_k - \hat{x}_k^2)(x_k - \hat{x}_k^1)'\} - E\{(x_k - \hat{x}_k^2)x_k'\} \\ &= M_k^{21} - M_k^{20}, \end{aligned} \tag{40}$$

$$\begin{aligned} E\{x_k \hat{x}_k^{1'}\} &= E\{[(x_k - \hat{x}_k^1) + \hat{x}_k^1]\hat{x}_k^{1'}\} \\ &= E\{(x_k - \hat{x}_k^1)\hat{x}_k^{1'}\} + E\{\hat{x}_k^1 \hat{x}_k^{1'}\} = E\{\hat{x}_k^1 \hat{x}_k^{1'}\}, \end{aligned} \tag{41}$$

$$\begin{aligned} &E\{\hat{x}_k^1 \hat{x}_k^{1'}\} \\ &= E\{(x_k - \hat{x}_k^1)(x_k - \hat{x}_k^1)'\} - E\{(x_k - \hat{x}_k^1)x_k'\} - E\{x_k(x_k - \hat{x}_k^1)'\} + E\{x_k x_k'\} \\ &= M_k^{11} - M_k^{10} - M_k^{01} + M_k^{00} = M_k^{00} - M_k^{01}. \end{aligned} \tag{42}$$

At $k = 0$,

$$E\{x_0 | X_0^1\} = \hat{x}_0^1 = \bar{x}_0, \tag{43a}$$

$$E\{\hat{x}_0^2 | X_0^1\} = \hat{x}_0^1 = \bar{x}_0. \tag{43b}$$

Thus,

$$S_0^2 = I.$$

From (39)–(42), we have

$$S_k^j = \begin{cases} I + (M_k^{ii} - M_k^{j0}) \times (M_k^{00} - M_k^{0i})^{-1}, & i, j = 1, 2, i \neq j \text{ and } k = 1, \ldots, N-1, \\ I, \quad j = 1, 2 \text{ and } k = 0. \end{cases} \tag{44}$$

Using (20)–(22), (26), (38), the difference equation (6a) for $\hat{x}_k^1$ may be written as

$$\begin{aligned} E\{x_{k+1} | X_k^1\} = A_k E\{x_k | X_k^1\} + B_k^1 u_k^1 + B_k^2 E\{u_k^{2*} | X_k^1\} \\ + M_k^{11} C_k^{1'} (N_k^1)^{-1} [y_k^1 - C_k^1 E\{x_k | X_k^1\}], \end{aligned} \tag{45}$$

with initial condition given by (43a). Similarly, we can show that, given M_k [obtained from (14)] and X_k^2, the best estimate of $\hat{x}_k^1$ (given X_k^2) for all $k \in [0, N-1]$ is given by

$$E\{\hat{x}_k^1 | X_k^2\} = S_k^1 \hat{x}_k^2. \tag{46}$$

Using stochastic dynamic programming and using mathematical induction, the closed-loop Nash control gains for players 1 and 2 are calculated as follows:

$$\min_{u_k^i} E\{J_k^i(x_k, u_k^i, u_k^j)|X_k^i\}$$
$$= \min_{u_k^i} E\{J_{k+1}^i(x_{k+1}) + x_k' Q_k^i x_k + u_k^{i'} R_k^{ii} u_k^i + u_k^{j'} R_k^{ij} u_k^j | X_k^i\},$$
$$i, j = 1, 2, i \neq j, \text{ and } k = 0, \ldots, N-1. \tag{47}$$

We will present the proof for player1. A similar procedure can be followed for player 2. We have that

$$E\{J_k^1(x_k, u_k^1, u_k^{2*})|X_k^1\}$$
$$= E\{x_k' Q_k^1 x_k + u_k^{1'} R_k^{11} u_k^1 + \hat{x}_k^{2'} L_k^{2'} R_k^{12} L_k^2 \hat{x}_k^2$$
$$+ (A_k x_k + B_k^1 u_k^1 + B_k^2 L_k^2 \hat{x}_k^2 + w_k)' K_{k+1}^1 (A_k x_k + B_k^1 u_k^1 + B_k^2 L_k^2 \hat{x}_k^2 + w_k)|X_k^1\}$$
$$+ \sum_{m=k+1}^{N-1} E\{w_m' K_{m+1}^1 w_m\} - \sum_{m=k+1}^{N-1} (\text{tr}\{P_m^{11} M_m^{11}\} + \text{tr}\{P_m^{12} M_m^{22}\}), \tag{48}$$

$$\min_{u_k^1} E\{J_k^1(x_k, u_k^1, u_k^{2*})|X_k^1\}$$
$$= \min_{u_k^1} E\{u_k^{1'}(R_k^{11} + B_k^{1'} K_{k+1}^1 B_k^1) u_k^1 + u_k^{1'} B_k^{1'} K_{k+1}^1 (A_k x_k + B_k^2 L_k^2 \hat{x}_k^2)$$
$$+ (A_k x_k + B_k^2 L_k^2 \hat{x}_k^2)' K_{k+1}^1 B_k^1 u_k^1 | X_k^1\}$$
$$+ E\{(A_k x_k + B_k^2 L_k^2 \hat{x}_k^2)' K_{k+1}^1 (A_k x_k + B_k^2 L_k^2 \hat{x}_k^2) + \hat{x}_k^{2'} L_k^{2'} R_k^{12} L_k^2 \hat{x}_k^2 | X_k^1\}$$
$$+ \sum_{m=k}^{N-1} E\{w_m' K_{m+1}^1 w_m\} - \sum_{m=k+1}^{N-1} (\text{tr}\{P_m^{11} M_m^{11}\} + \text{tr}\{P_m^{12} M_m^{22}\}). \tag{49}$$

Now, setting

$$\partial E\{J_k^1(x_k, u_k^1, u_k^{2*})|X_k^1\}/\partial u_k^1 = 0$$

leads to

$$u_k^{1*} = -(R_k^{11} + B_k^{1'} K_{k+1}^1 B_k^1)^{-1} B_k^{1'} K_{k+1}^1 \times [A_k E\{x_k|X_k^1\} + B_k^2 L_k^2 E\{\hat{x}_k^2|X_k^1\}]$$
$$= -(R_k^{11} + B_k^{1'} K_{k+1}^1 B_k^1)^{-1} B_k^{1'} K_{k+1}^1 (A_k + B_k^2 L_k^2 S_k^2) \hat{x}_k^1$$
$$= L_k^1 \hat{x}_k^1. \tag{50}$$

Similarly, for player 2, setting

$$\partial E\{J_k^2(x_k, u_k^{1*}, u_k^2)|X_k^2\}/\partial u_k^2 = 0$$

leads to

$$\begin{aligned} u_k^{2*} &= -(R_k^{22} + B_k^{2'} K_{k+1}^2 B_k^2)^{-1} B_k^{2'} K_{k+1}^2 (A_k + B_k^1 L_k^1 S_k^1) \hat{x}_k^2 \\ &= L_k^2 \hat{x}_k^2. \end{aligned} \tag{51}$$

Denote

$$u_k^{1*} = L_k^1 \hat{x}_k^1 = L_k^1 (x_k - \tilde{x}_k^1), \tag{52}$$

$$u_k^{2*} = L_k^2 \hat{x}_k^2 = L_k^2 (x_k - \tilde{x}_k^2). \tag{53}$$

By substitution of (52)–(53) into (48) and using (24)–(25), we have

$$\begin{aligned} &E\{J_k^1(x_k, u_k^{1*}, u_k^{2*}) | X_k^1\} \\ &= E\{x_k'(Q_k^1 + L_k^{1'} R_k^{11} L_k^1 + L_k^{2'} R_k^{12} L_k^2 \\ &\quad + (A_k + B_k^1 L_k^1 + B_k^2 L_k^2)' K_{k+1}^1 (A_k + B_k^1 L_k^1 + B_k^2 L_k^2)) x_k | X_k^1\} \\ &\quad + E\{\tilde{x}_k^{1'} L_k^{1'} (B_k^{1'} K_{k+1}^1 B_k^1 + R_k^{11}) L_k^1 \tilde{x}_k^1\} \\ &\quad + E\{\tilde{x}_k^{2'} L_k^{2'} (B_k^{2'} K_{k+1}^1 B_k^2 + R_k^{12}) L_k^2 \tilde{x}_k^2\} \\ &\quad - 2 \cdot E\{x_k' L_k^{1'} (B_k^{1'} K_{k+1}^1 B_k^1 + R_k^{11}) L_k^1 \tilde{x}_k^1\} \\ &\quad - 2 \cdot E\{x_k' L_k^{2'} (B_k^{2'} K_{k+1}^1 B_k^2 + R_k^{12}) L_k^2 \tilde{x}_k^2\} \\ &\quad + \sum_{m=k}^{N-1} E\{w_m' K_{m+1}^1 w_m\} \\ &\quad - \sum_{m=k+1}^{N-1} (\mathrm{tr}\{P_m^{11} M_m^{11}\} + \mathrm{tr}\{P_m^{12} M_m^{22}\}) \\ &= E\{x_k' K_k^1 x_k | X_k^1\} + \mathrm{tr}\{P_k^{11} M_k^{11}\} + \mathrm{tr}\{P_k^{12} M_k^{22}\} \\ &\quad - 2\,\mathrm{tr}\{P_k^{11} M_k^{01}\} - 2\,\mathrm{tr}\{P_k^{11} M_k^{02}\} \\ &\quad + \sum_{m=k}^{N-1} E\{w_m' K_{m+1}^1 w_m\} \\ &\quad - \sum_{m=k+1}^{N-1} (\mathrm{tr}\{P_m^{11} M_m^{11}\} + \mathrm{tr}\{P_m^{12} M_m^{22}\}), \end{aligned} \tag{54}$$

which completes the proof for player 1. A symmetrical argument proves the results for player 2. □

Although the optimum cost-to-go at time k separates into three groups of terms, the first group is not just the conditional expectation of a term that is analogous to the deterministic case, because the Riccati equation for K_{k+1}^i depends on the estimation error covariances. The second group of terms looks like the contribution of the estimation error covariances, but the controller gains affect the estimation errors. The third group relates to the covariance of the process noise, but the weighting matrix K^i_{k+1} is not the same as in the deterministic case. The separation principle does not hold in this case.

The linear strategy pair (u^{1*}_k, u^{2*}_k) [given by (26)] will be the unique closed-loop Nash strategy for the above stochastic discrete-time linear-quadratic game if the matrices $M_k, F^1_k, F^2_k, G^1_k, G^2_k$ exist and are bounded for all $k \in [0, N-1]$. This requires that the matrices L^1_k, L^2_k, K^1_k, K^2_k exist and are bounded for all $k \in [0, N-1]$. Moreover, the matrices $M^{00}_k - M^{01}_k$, $M^{00}_k - M^{02}_k$ should be invertible for all $k \in [1, N-1]$.

4. Closed-Loop Nash for Information Structure Type B

We now consider the second type of information structure, $X^i_k = \{y^i_k, u^i_{k-1}\}$ for player i, $i = 1, 2$. In this case, instead of (6), we use the following form of the state estimators:

$$\hat{x}^1_{k+1} = (F^1_k - G^1_{k+1} C^1_{k+1} A_k)\hat{x}^1_k + G^1_{k+1} y^1_{k+1} + (D^1_k - G^1_{k+1} C^1_{k+1} B^1_k)u^1_k, \tag{55a}$$

$$\hat{x}^2_{k+1} = (F^2_k - G^2_{k+1} C^2_{k+1} A_k)\hat{x}^2_k + G^2_{k+1} y^2_{k+1} + (D^2_k - G^2_{k+1} C^2_{k+1} B^2_k)u^2_k, \tag{55b}$$

where the matrices A_k, B_k are the same as given in equation (1), the matrices F^1_k, F^2_k, G^1_{k+1}, G^2_{k+1}, D^1_k, D^2_k are the same as given in equations (20)–(21), and the matrices C^1_{k+1}, C^2_{k+1} are the same as given in equation (2). The closed-loop Nash strategies for this information structure lead to the same algorithms for the controller and estimation matrices as the ones obtained earlier. However, note that the actual optimal estimator equations resulting from (55) are not the same as the optimal estimator equations resulting from (6).

5. Numerical Computation

A suggested iteration for the calculation of matrices K^1_k, K^2_k, L^1_k, L^2_k, S^1_k, S^2_k, M_k, $k = 0, \ldots, N-1$, in (9)–(19) is given as follows:

Step 1. Set $S^1_k = S^2_k = I, k = 0, \ldots, N-1$.

Step 2. Calculate K^1_k, K^2_k, L^1_k, L^2_k, $k = N-1, \ldots, 0$, backward in time using equations (9) and (10).

Step 3. Start Iteration 1. Calculate the matrix $\bar{A}_k, k = 0, \ldots, N-1$, forward in time using equation (16).

Step 4. Calculate M_k, S_k^1, S_k^2, $k=0,\dots,N-1$, forward in time using equations (14) and (11), respectively.

Step 5. Solve for K_k^1, K_k^2, L_k^1, L_k^2, $k=N-1,\dots,0$, backward in time using equations (9) and (10). This is the end of Iteration 1.

Step 6. Start Iteration 2 by repeating Steps 3 to 5.

Step 7. Compare the matrices K_k^1, K_k^2, L_k^1, L_k^2, S_k^1, S_k^2, M_k, $k=0,\dots,N-1$, of Iteration 2 with those of Iteration 1.

Step 8. If the differences are zero (or smaller than a predetermined error criterion), stop the iteration loop.
Else, repeat the above procedure for Iterations 3, 4, ..., m, where the results of Iteration m satisfy the condition of Step 8, or m is the maximum number of iterations allowed.

6. Simple Application

In this section, we present a simple example of a two-zone energy trading system, shown in Fig. 1, as an application of the proposed Nash solution. In this example, two electric energy-generating firms, labeled G1 and G2, are trying to sell energy in a day-ahead (24 hour ahead) energy market. They try to minimize their electricity production costs in the Nash sense. The firms should satisfy the system demand D_k at each trading hour k. The market rules require both firms to bid for all 24 trading hours at once. Therefore, the optimization horizon for both firms is 24 hours ($N=24$). Itis assumed that, at each hour k, both firms have quadratic cost-to-go functions of the form

$$J_k^1(x_k, q_k^1) = 0.1x_N^2 + \sum_{i=k}^{N-1} \{0.1(q_i^1)^2\},$$

$$J_k^2(x_k, q_k^2) = 0.1x_N^2 + \sum_{i=k}^{N-1} \{0.15(q_i^2)^2\},$$

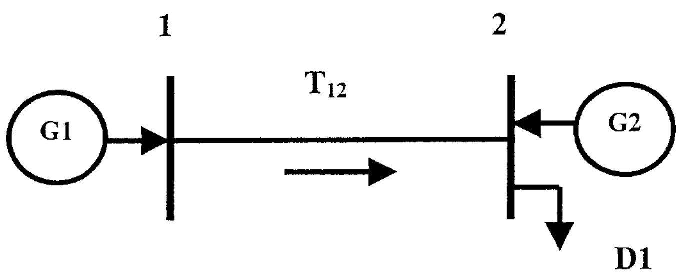

Fig. 1. A two-zone energy trading system.

where x_k represents the residual demand at hour k, and where q_k^1, and q_k^2 represent the bid quantities for firms G1 and G2 respectively at hour k. The term $0.1x_N^2$ represents the no-load cost for each firm. The system dynamics is described by the energy balance equation

$$x_{k+1} = x_k + D_k - q_k^1 - q_k^2 + w_k, \qquad k = 1, 2, \ldots, 24.$$

We will assume that each firm has a noisy measurement of the system demand at each trading hour k. That is,

$$y_k^1 = x_k + v_k^1,$$

$$y_k^2 = x_k + v_k^2.$$

In the above two expressions, w_k, v_k^1, v_k^2 are independent, white, zero-mean, and Gaussian noises with covariances

$$W_k = 1.0, \qquad N_k^1 = 0.5, \qquad N_k^2 = 0.8,$$

respectively. Furthermore, the initial demand x_0 at $k = 0$ is assumed to be a Gaussian random process with mean $\bar{x}_0 = 0.5$ and covariance $\Gamma_0 = 0.085$. Finally, assuming that the bid prices (marginal costs) p_k^1 and p_k^2 for both firms are given by the expressions

$$p_k^1 = 0.2q_k^1, \qquad p_k^2 = 0.3q_k^2,$$

the total profits for each firm over the 24 hour time horizon can be calculated as

$$\pi_{G1} = \sum_{k=0}^{N-1} (p_k^1 \cdot q_k^{1*}) - J_0^1(\bar{x}_0, q_k^{1*}),$$

$$\pi_{G2} = \sum_{k=0}^{N-1} (p_k^2 \cdot q_k^{2*}) - J_0^2(\bar{x}_0, q_k^{2*}).$$

The system parameters that are needed to calculate the Nash bid quantities and prices for each firm for $k \in [0, N-1]$ are

$$A_k = 1, \qquad B_k^1 = B_k^2 = -1, \qquad C_k^1 = C_k^2 = 1,$$

$$Q_N^1 = Q_N^2 = 0.1, \qquad Q_k^1 = Q_k^2 = 0,$$

$$R_k^{11} = 0.1, \qquad R_k^{22} = 0.15, \qquad R_k^{12} = R_k^{21} = 0.$$

The Nash strategies are calculated for two different cases:

Case 1. No transmission capacity constraint,
Case 2. Transmission line T_{12} has a limited capacity.

The system demand, player 1 estimate of the system demand, player 2 estimate of the system demand, and the Nash strategies for both firms in Case 1 (no transmission constraint) are shown in Fig. 2. The Nash strategies include bid quantities (in MW) and bid prices (in $ per MW) for both firms. Similar results are shown in Fig. 3 for Case 2 (transmission line capacity limit of 80 MW). The simulation results of Case 2 show that, for the peak demand hours (8 to 20), firm G1 cannot sell more that 80 MW, due to the

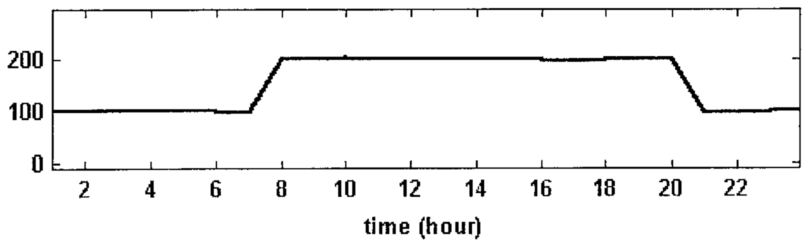

Fig. 2a. Simulation results for Case 1: System demand (solid line), Player 1's estimate of system demand (dashed line) and Player 2's estimate of system demand (dot-dashed line) all in MW. (Note that the three plots are almost identical, so they appear superimposed as one plot on the figure).

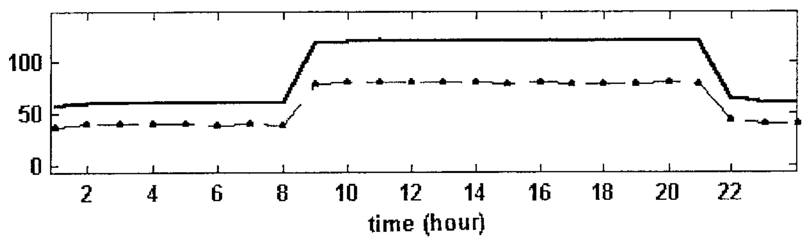

Fig. 2b. Simulation results for Case 1: Player 1 bid quantity (q_k^{1*}, solid line) and Player 2 bid quantity (q_k^{2*}, dashed line) in MW.

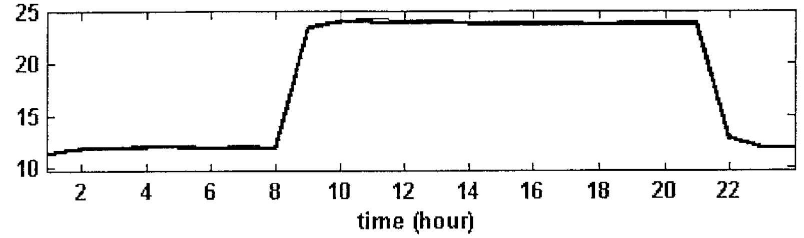

Fig. 2c. Simulation results for Case 1: Player 1 bid price (Bid_k^1, solid line) and Player 2 bid price (Bid_k^2, dashed line) in $ per MW.

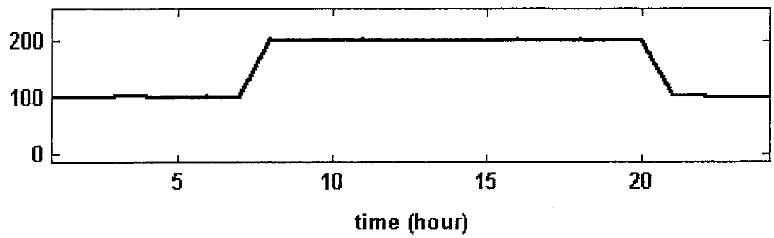

Fig. 3a. Simulation results for Case 2: System demand (solid line), Player 1's estimate of system demand (dashed line) and Player 2's estimate of system demand (dot-dashed line) all in MW. (Note that the three plots are almost identical, so they appear superimposed as one plot on the figure).

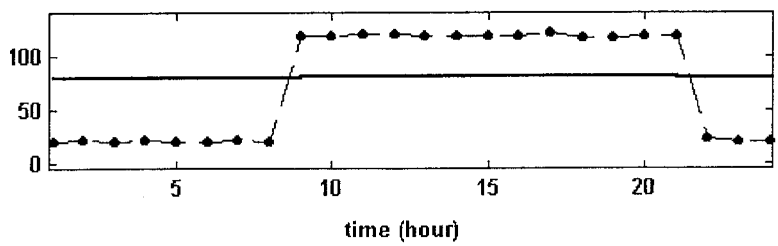

Fig. 3b. Simulation results for Case 2: Player 1 bid quantity (q_k^{1*}, solid line) and Player-2's bid quantity (q_k^{2*}, dashed line) in MW.

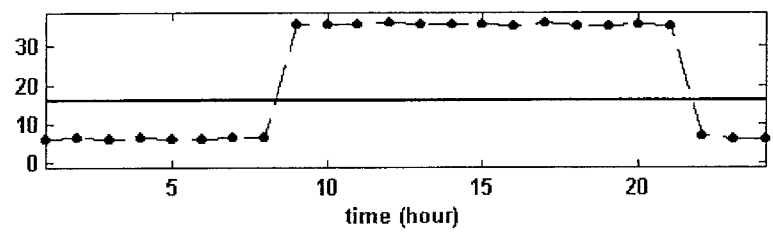

Fig. 3c. Simulation results for Case 2: Player 1 bid price (Bid_k^1, solid line) and Player 2 bid price (Bid_k^2, dashed line) in $ per MW.

transmission line capacity limit. This will cause an increase in the average energy price in zone 2 during the peak demand hours. The average energy price goes up from $24.00 per MW in Case 1 (no transmission constraint) to $36.00 per MW in Case 2 (with transmission constraints) during peak demand hours. On the other hand, the average energy price in zone 1 will drop to \$16.00 per MW. Therefore, the average congestion price of transmission line T_{12} can be calculated as the difference between the zone 1 and zone 2 average energy prices (\$36.00 − \$16.00 = \$20.00 per MW).

The expected profits for the two firms in Case 1 are

$$\pi_{G1} = \$21{,}684.00, \qquad \pi_{G2} = \$13{,}761.00.$$

The ratios of total cost to no-load cost for both firms in Case 1 are

$$r_{G1} = 21.7, \qquad r_{G2} = 14.5.$$

The expected profits for the two firms in Case 2 are

$$\pi_{G1} = \$15{,}222.00, \qquad \pi_{G2} = \$25{,}899.00.$$

The ratios of total cost to no-load cost for both firms in Case 2 are:

$$r_{G1} = 16.9, \qquad r_{G2} = 32.0.$$

7. Conclusions

In this paper, a solution for the stochastic LQG nonzero-sum Nash games with constrained state information is presented. It is assumed that each player is limited to a linear Kalman filter-like estimator to generate his optimal controls. The optimal control laws (u_k^{1*}, u_k^{2*}) that solve the above stochastic discrete-time Nash game do not satisfy the separation principle. As it can be seen from equation (10), the calculation of the optimal control gains (L_k^1, L_k^2) depends on the components of matrix M_k, which is related to the estimation error covariance matrix. Therefore, the controller and estimator cannot be separated. However, we should note that, if the measurements $y_k^1 = y_k^2$ in (2), then the matrices S_k^1, S_k^2 in (11) will converge to the identity matrix and the controller gains will not depend on the state estimation errors. However, the estimator equations will still depend on the controller gains.

A simple example of a two-zone energy trading system is considered to illustrate the developed Nash strategies. In this example, the Nash strategies are calculated for the two cases of unlimited and limited transmission capacity constraints.

The simulation results show that transmission capacity constraint (Case 2) could cause a decrease in expected profits of player 1. On the other hand, the expected profits of player 2 would increase during peak demand hours upon the existence of transmission congestion. The linear strategy pair (u_k^{1*}, u_k^{2*}) [given by (26)] will be the unique Nash strategy for the above stochastic discrete-time linear-quadratic game if the solutions for the matrices $M_k, F_k^1, F_k^2, G_k^1, G_k^2$ exist and are bounded for all $k \in [0, N-1]$. This requires that the solutions for the matrices L_k^1, L_k^2, K_k^1, K_k^2 exist and are bounded for all $k \in [0, N-1]$. Moreover, the matrices $\mathrm{M}_k^{00} - \mathrm{M}_k^{01}$, $M_k^{00} - M_k^{02}$ should be invertible for all $k \in [1, N-1]$.

References

1. Rhodes, I. B., and Luenberger, D. G., *Differential Games with Imperfect State Information*, IEEE Transactions on Automatic Control, Vol. 14, No. 1, pp. 29–38, 1969.
2. Rhodes, I. B., and Luenberger, D. G., *Stochastic Differential Games with Constrained State Estimators*, IEEE Transactions on Automatic Control, Vol. 14, No. 5, pp. 476–481, 1969.
3. Saksena, V. R., and Cruz, J. R., *A Multimodel Approach to Stochastic Nash Games*, Automatica, Vol. 18, No. 3, pp. 295–305, 1982.
4. Basar, T., and Olsder, G. J., *Dynamic Noncooperative Game Theory*, 2nd Edition, Academic Press, London, England, 1995.