SOME PARTIALLY OBSERVED MULTI-AGENT LINEAR EXPO- NENTIAL QUADRATIC STOCHASTIC DIFFERENTIAL GAMES

TYRONE E. DUNCAN
Mathematics Department
Snow Hall, 1460 Jayhawk Blvd
Lawrence, KS 66045, USA

Abstract. Some multi-agent stochastic differential games described by a stochastic linear system driven by a Brownian motion and having an exponential quadratic payoff for the agents are formulated and solved. The agents have either complete observations or partial observations of the system state. The agents act independently of one another and the explicit optimal feedback control strategies form a Nash equilibrium. In the partially observed problem the observations are the same for all agents which occurs in broadcast situations. The optimal control strategies and optimal payoffs are given explicitly. The method of solution for both problems does not require solving either Hamilton-Jacobi-Isaacs equations or backward stochastic differential equations.

1. Introduction. Two-person, zero-sum stochastic differential games developed as a natural generalization of (one-person) stochastic control problems and minimax control problems. Isaacs [16] obtained nonlinear partial differential equations that determine the lower and the upper values of the game. If these two equations for the game are equal then the game is said to have a value and the two nonlinear partial differential equations become one. Two well known general methods are available for solving stochastic differential games. The first method uses the Hamilton-Jacobi-Isaacs (HJI) equations which can be considered as the game analogue of the Hamilton-Jacobi-Bellman equations of stochastic optimal control and the Hamilton-Jacobi equations of deterministic optimal control. A number of results (e. g. Fleming and Hernandez-Hernandez [13], Fleming and Souganidis [14]) have provided conditions for the existence of solutions of HJI equations for stochastic differential games. The second method uses the solution of backward stochastic differential equations that generalizes the use of backward stochastic equations from stochastic optimal control [15]. This latter approach is more recent than the HJI equation approach and fewer results about these backward equations are available (e. g. Buckdahn and Li [5]). Two general references on differential games are the monographs [2] and [3].

2000 Mathematics Subject Classification. Primary: 91A15, 49N70; Secondary: 91A25.

Key words and phrases. Stochastic differential games, multi-agent games, partially observed games, risk sensitive games, exponential quadratic payoffs.

The author was supported by NSF grant DMS 1411412, AFOSR grant FA9550-17-1-0073, and ARO grant W911NF-14-10390.
Exponential quadratic cost functionals with linear stochastic systems have been considered for problems of optimal control with complete or partial observations (e.g. [4, 7, 17, 19, 21]). These problems are often called risk sensitive quadratic control. Two-person linear stochastic differential games with exponential quadratic payoffs and complete observations have been explicitly solved [9]. Many situations described by stochastic differential games require more than two agents so it is natural to consider multi-agent linear stochastic differential games with exponential quadratic payoffs. Furthermore it is important to study partially observed stochastic differential games. In this paper some multi-agent games are solved where the payoff is an exponential quadratic functional for each of the players and the observations are either complete or partial. The optimal control strategies for both types of problems form Nash equilibria [20]. The methods of solutions for the explicit optimal control strategies do not require solutions of Hamilton-Jacobi-Isaacs equations or backward stochastic differential equations. It seems that there is a limited amount of work on linear-quadratic n-person games [1] and no work for linear-exponential-quadratic n-person games particularly for partially observed problems.

2. Stochastic differential game formulation. A stochastic differential game for \( k_1 \) players who want to minimize the payoff and \( k_2 \) players who want to maximize the payoff is given now. The linear stochastic equation is the following

\[
\begin{align*}
    dX(t) &= A(t)X(t)dt + \sum_{i=1}^{k_1} B_i(t)U_i(t)dt + \sum_{i=1}^{k_2} C_i(t)V_i(t)dt \\
    X(0) &= X_0
\end{align*}
\]

where \( A(t) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n), B_i(t) \in \mathcal{L}(\mathbb{R}^{n_i}, \mathbb{R}^n) \) for \( i = 1, \ldots, k_1 \), \( C_j(t) \in \mathcal{L}(\mathbb{R}^{m_j}, \mathbb{R}^n) \) for \( j = 1, \ldots, k_2 \), \( U_i(t) \in \mathbb{R}^{n_i}, V_j(t) \in \mathbb{R}^{m_j} \), for \( i = 1, \ldots, k_1; j = 1, \ldots, k_2 \), the linear transformations \( (A, B_1, \ldots, B_{k_1}, C_1, \ldots, C_{k_2}) \) are continuous on \([0, T]\), the nonnegative integers \((k_1, k_2, n_1, \ldots, n_{k_1}, m_1, \ldots, m_{k_2})\) are arbitrary, \((W(t), t \in [0, T])\) is an \( \mathbb{R}^n \)-valued standard Brownian motion, \( F \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) \) and invertible, \( T > 0 \) is fixed, \( X_0 \in \mathbb{R}^n \) is a constant, the process \( W \) is defined on the complete probability space \((\Omega, \mathcal{F}, \mathbb{P})\), and \((\mathcal{F}(t), t \in [0, T])\) is the filtration for \( W \).

The payoff functional \( J_\mu \) is given as follows

\[
\begin{align*}
    J_{0,\mu}(U_1, \ldots, U_{k_1}, V_1, \ldots, V_{k_2}) &= \mu \exp \left[ \frac{\mu}{2} \int_0^T (\langle QX, X \rangle + \sum_{i=1}^{k_1} < R_iU_i, U_i > - \sum_{i=1}^{k_2} < S_iV_i, V_i > )dt \\
    &+ \frac{\mu}{2} < MX(T), X(T) > \right] \\
    J_\mu(U_1, \ldots, U_{k_1}, V_1, \ldots, V_{k_2}) &= \mathbb{E} J_{0,\mu}(U_1, \ldots, U_{k_1}, V_1, \ldots, V_{k_2})
\end{align*}
\]

where \( \mathbb{E} \) is expectation for the probability \( \mathbb{P} \) and \( \mu \) is a nonzero real number where subsequent conditions are imposed on it. The linear operators in (3) satisfy the conditions that \( Q \) is symmetric, uniformly positive and continuous on \([0, T]\) and \( R_1, \ldots, R_{k_1}, S_1, \ldots, S_{k_2} \) are symmetric, uniformly positive and continuous on \([0, T]\).

The family of admissible control strategies for each agent is generically described as satisfying

(A1) \( \{ U : [0, T] \rightarrow \mathbb{R}^i \text{ where } U \text{ is } (\mathcal{F}(t)) \text{ predictable and } \int_0^T |U|^2 dt < \infty \text{ a.s.} \} \)

where \( i \) is a suitable positive integer and \( J_{0,\alpha}(U_1, \ldots, U_{k_1}, 0, \ldots, 0) < \infty \) a.s. for some \( \alpha > 0 \) and \( \beta J_{0,\beta}(0, \ldots, 0, V_1, \ldots, V_{k_2}) < \infty \) a.s. for some \( \beta < 0 \).
Alternatively a sequence of stopping times \( (T_n, n \in \mathbb{N}) \) could be defined where the payoff is finite for each \( T_n \) and these stopping times converge to \( T \).

Agents with the control strategies \( (U_1, ..., U_{k_1}) \) seek to minimize the payoff \( J_\mu \) and agents with the control strategies \( (V_1, ..., V_{k_2}) \) seek to maximize the payoff \( J_\mu \). Both families of agents use feedback strategies determined by (A1).

The Riccati equation for the completely observed stochastic differential game is

\[
\frac{dp_\mu}{dt} = -p_\mu A - A^T p_\mu - Q
+ P_\mu (\sum_{i=1}^{k_1} B_i \Sigma_i^{-1} B_i^T - \sum_{i=1}^{k_2} C_i S_i^{-1} C_i^T - \mu FF^T) P_\mu
\]

\[
P_\mu(T) = M
\]

It is assumed that this Riccati equation has a unique, positive symmetric solution which follows for example if \( \sum_{i=1}^{k_1} B_i \Sigma_i^{-1} B_i^T - \sum_{i=1}^{k_2} C_i S_i^{-1} C_i^T - \mu FF^T > 0 \) uniformly in \( t \in [0, T] \) or equivalently there is a \( c > 0 \) that does not depend on \( t \in [0, T] \) such that \( \langle \sum_{i=1}^{k_1} B_i \Sigma_i^{-1} B_i^T - \sum_{i=1}^{k_2} C_i S_i^{-1} C_i^T - \mu FF^T \rangle x, x > c x, x > \) for each \( x \in \mathbb{R}^n \). The parameter \( \mu \) is constrained to ensure this Riccati solution uniqueness.

The solution of the Riccati equation arises naturally in this optimization problem. The natural geometric setting for the problem (1) and (3) is the Lagrangian Grassmannian that is usually denoted \( \Lambda(n) \) for an appropriate positive integer \( n \). It is a Grassmannian of \( n \) planes in \( 2n \) dimensional Euclidean space with a closed, nondegenerate two form denoted \( \omega \). The space \( \Lambda(n) \) has dimension \( \frac{n(n+1)}{2} \) and it can also be described as the homogeneous space \( \mathbb{U(n)}/O(n) \) where \( \mathbb{U(n)} \) is the group of unitary transformations on \( \mathbb{C}^n = \mathbb{R}^{2n} \) and \( O(n) \) is the orthogonal group on \( \mathbb{R}^n \). A (nonsingular) plane in \( \Lambda(n) \) can be described as \( \langle x, P x \rangle \) where \( P \) is a symmetric (nonsingular) matrix. This description of \( \Lambda(n) \) shows that \( dim \Lambda(n) = \frac{n(n+1)}{2} \) and \( \Lambda(n) \) is a natural description to determine optimal control strategies from the solution of a Riccati equation and the symplectic geometry of optimization.

3. Completely observed stochastic differential games. Initially a completely observed multi-agent stochastic differential game with an exponential quadratic payoff is solved. The solution can be considered as a natural generalization of the control problem [7] and the two-agent problem [9] with some additions. The following proposition provides optimal feedback control strategies for the multi-agents and these strategies are shown to form a Nash equilibrium for the game.

**Proposition 1.** The multi-agent stochastic differential game given by (1) and (3) with the families of admissible control strategies (A1) has a Nash equilibrium with the optimal control strategies that are given by

\[
U^*_i(t) = -R_i^{-1}(t)B_i^T(t)P_\mu(t)X(t)
\]

\[
V^*_j(t) = S_j^{-1}(t)C_j(t)P_\mu(t)X(t)
\]

for \( i = 1, ..., k_1 \) and \( j = 1, ..., k_2 \) where \( P_\mu \) is the unique symmetric, positive solution of the Riccati equation (5). The optimal payoff is

\[
J_\mu(U^*_1, ..., U^*_{k_1}, V^*_1, ..., V^*_{k_2}) = \mu exp\left[\frac{\mu}{2} (P_\mu(0)X_0, X_0) + \int_0^T tr(P_\mu F^TF)dt\right]
\]

**Proof.** The proof is motivated by a method that is used to solve a linear exponential quadratic Gaussian control problem [7] and a two person game [9]. In both of these publications the Brownian motions should have dimension of the state and the
linear transformations acting on these Brownian motions in the stochastic equations should be invertible. Let $Y(t) = \langle P_\mu(t)X(t), X(t) \rangle$ for each $t \in [0, T]$ where $X$ satisfies (1) and $P_\mu$ is the unique positive symmetric solution of (5). The dependence of $Y$ on $\mu$ is suppressed for notational simplicity. Apply the change of variables formula of Ito to the process $(Y(t), t \in [0, T])$ to obtain the following equality.

$$Y(T) - Y(0) = \int_0^T \left( 2 < P_\mu X, AX + \Sigma_{i=1}^k B_i U_i + \Sigma_{i=1}^k C_i V_i > dt + 2 < P_\mu X, FdW > + tr(P_\mu F^T F)dt ight)$$

$$+ < X, (-P_\mu A - A^T P_\mu - Q)X > dt + < X, (P_\mu (\Sigma_{i=1}^k B_i R_i^{-1} B_i^T - \Sigma_{i=1}^k C_i S_i^{-1} C_i^T - \mu F F^T) P_\mu)X > dt$$

Let $L_\mu$ be the quadratic functional that appears in the exponential of the payoff functional (3) as

$$L_\mu(U_1, ..., U_{k_1}, V_1, ..., V_{k_2}) = \frac{\mu}{2} \int_0^T \left( < Q X, X > + \Sigma_{i=1}^k R_i U_i, U_i > - \Sigma_{i=1}^k S_i V_i, V_i > dt + < MX(T), X(T) > \right)$$

so that the payoff $J_\mu$ can be expressed as

$$J_\mu(U_1, ..., U_{k_1}, V_1, ..., V_{k_2}) = \mu E \exp[L_\mu(U_1, ..., U_{k_1}, V_1, ..., V_{k_2})]$$

The following equality is determined from the definition of $L_\mu$ and (9).

$$L_\mu(U_1, ..., U_{k_1}, V_1, ..., V_{k_2}) - \frac{\mu}{2} < P_\mu(0)X_0, X_0 > = \frac{\mu}{2} \int_0^T \left( \langle \Sigma_{i=1}^k R_i U_i, U_i > - \Sigma_{i=1}^k S_i V_i, V_i > > 2 < P_\mu \Sigma_{i=1}^k B_i U_i, X > + 2 < P_\mu \Sigma_{i=1}^k C_i V_i, X > > < \Sigma_{i=1}^k P_\mu B_i R_i^{-1} B_i^T P_\mu X, X > - < \Sigma_{i=1}^k P_\mu C_i S_i^{-1} C_i^T P_\mu X, X > > - \mu < P_\mu F F^T P_\mu X, X > + tr(P_\mu F F^T)dt + 2 < FdW, P_\mu X > \right)$$

$$= \frac{\mu}{2} \int_0^T \left( \langle \Sigma_{i=1}^k R_i U_i, U_i > - \Sigma_{i=1}^k S_i V_i, V_i > > 2 < P_\mu \Sigma_{i=1}^k B_i U_i, X > + 2 < P_\mu \Sigma_{i=1}^k C_i V_i, X > > < \Sigma_{i=1}^k P_\mu B_i R_i^{-1} B_i^T P_\mu X, X > - < \Sigma_{i=1}^k P_\mu C_i S_i^{-1} C_i^T P_\mu X, X > > - \mu < P_\mu F F^T P_\mu X, X > + tr(P_\mu F F^T)dt + 2 < FdW, P_\mu X > \right)$$

To eliminate the stochastic integral term in (12) when it and the other terms are exponentiated, the stochastic integral and the term following it in (12) (an increasing process) are identified as a Radon-Nikodym derivative that is used to define a transformation of the Wiener measure by absolute continuity. The additional term in the Riccati equation (5) from the Riccati equation for a linear-quadratic payoff is used to determine the increasing process. Let $\tilde{P}$ be the probability measure that is obtained from the following Radon-Nikodym derivative as $d\tilde{P} = MdP$

$$M(t) = \exp[\mu \int_0^t < P_\mu X, FdW > - \frac{\mu^2}{2} \int_0^t |F^T P_\mu X|^2 dt]$$

(13)
where \( X \) is the solution of (1) and the linear feedback strategies from (12) are used. The process \( \bar{M}(\cdot) \) describes a Radon-Nikodym derivative by the strong dichotomy of the absolute continuity for Gaussian measures (Feldman-Hajek theorem) so it follows that \( \mathbb{E}[\bar{M}] = 1 \) and \( \bar{P} \) is indeed a probability measure. To verify the optimality of the strategies determined by making the quadratic integrands of the integrals of the agents’ strategies to be zero in (12), some variations of these strategies are made that preserve absolute continuity. It follows directly that suitably small variations of these strategies determine a solution \( X \) and a measure that is absolutely continuous with respect to Wiener measure so that a new measure is obtained from a process that corresponds to \( \bar{M}(\cdot) \) in (13). Thus let \( \bar{U}_i \) and \( \bar{V}_j \) for \( i \in \{1, ..., k_1\} \) and \( j \in \{1, ..., k_2\} \) be given as

\[
\bar{U}_i(t) = U^*(t) + U_{i,1}(t)
\]

where \( U_{i,1}(t) = \alpha_i \) for \( t \in [t_0, t_f] \) and the family \( \alpha_i \) are \( \mathcal{F}(t_0) \) measurable and bounded random variables for \( i = 1, ..., k_1 \) and similarly for \( V_j \), \( j = 1, ..., k_2 \). For the control strategies \( \bar{U}_i \) and \( \bar{V}_j \) \( i \in \{1, ..., k_1\} \) and \( j \in \{1, ..., k_2\} \) the exponential that corresponds to the control is also a Radon-Nikodym derivative so that it follows by comparing the two payoffs that the optimal strategies are determined by making the quadratic integrands in the agents’ strategies zero, that is, the optimal strategies are

\[
U^*_i(t) = -R_i^{-1}(t)B_i^T(t)P_\mu(t)X(t)
\]

\[
V^*_j(t) = S_j^{-1}(t)C_j^T(t)P_\mu(t)X(t)
\]

for \( i = 1, ..., k_1 \) and \( j = 1, ..., k_2 \) where \( P_\mu \) is the unique, symmetric, positive solution of the Riccati equation (5), that is,

\[
J_\mu(U_1^*, ..., U_{k_1}^*, V_1^*, ..., V_{k_2}^*) = J_\mu(U_1, ..., U_{k_1}, V_1^*, ..., V_{k_2}^*)
\]

\[
J_\mu(U_1^*, ..., U_{k_1}^*, V_1^*, ..., V_{k_2}^*) \geq J_\mu(U_1^*, ..., U_{k_1}^*, V_1, ..., V_{k_2})
\]

for \( \mu > 0 \) and the inequalities are reversed for \( \mu < 0 \). Clearly the optimal payoff is

\[
J_\mu(U_1^*, ..., U_{k_1}^*, V_1^*, ..., V_{k_2}^*) = \mu \exp\left[\frac{\mu}{2}(P_\mu(0)X_0, X_0) + \int_0^T tr(P_\mu F^T F)dt\right]
\]

4. Partially observed stochastic differential games. In this section some partially observed stochastic differential games are formulated and explicitly solved. The stochastic model is the linear stochastic differential equation (1) and the payoff is the exponential of the quadratic functional (3) as given in the previous section. However in this case the players have only noisy partial observations of the state. Again the solution is obtained directly without requiring the solutions of Hamilton-Jacobi-Isaacs equations or backward stochastic differential equations. This result can be viewed as a generalization of the partially observed linear exponential quadratic control problem [4] with a more direct proof that is motivated by [11].

It is assumed that the players have only noisy partial observations of the state that is the same for all players. This situation occurs in a broadcasting model or other models where all players have access to the same information. The common observation equation is the following

\[
dY(t) = H(t)X(t)dt + G(t)dV(t)
\]

\[
Y(0) = 0
\]
where \( Y(t) \in \mathbb{R}^p, H(t) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p), G \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^p) \) is invertible for each \( t \in [0, T] \), \( G, H \) are continuous on \([0, T]\), and \((V(t), t \in [0, T])\) is a standard \( \mathbb{R}^q \)-valued Brownian motion defined on the same probability space as \( W \). The pair of standard Brownian motions \((V, W)\) is independent. The filtration \((\mathcal{G}(t), t \in [0, T])\) is the observation process \((Y(t), t \in [0, T])\) where the control strategies are \( U_i(t) = -R_i^{-1}(t)B_i^T(t)\tilde{P}_\mu(t)E[X(t)|\mathcal{G}(t)] \) for \( i = 1, \ldots, k_1 \) and \( V_i(t) = S_i^{-1}(t)C_i^T(t)\tilde{P}_\mu(t)E[X(t)|\mathcal{G}(t)] \) for \( i = 1, \ldots, k_2 \) where \( \tilde{P}_\mu \) is defined below.

The appropriate estimation equation for the exponential quadratic payoff (3), often called the information filter (e.g. [12]), is given by

\[
dZ(t) = (A(t) - K_\mu(t)H^T(t)(G(t)G^T(t))^{-1}H(t) + \mu K_\mu(t)Q(t))Z(t)dt + \sum_{i=1}^{k_1}B_i(t)U_i(t)dt + \sum_{i=1}^{k_2}C_i(t)V_i(t)dt + K_\mu(t)H^T(t)(G(t)G^T(t))^{-1}dY(t)
\]

\[
Z(0) = X(0)
\]

and \((K_\mu(t), t \in [0, T])\) is the unique, positive symmetric solution of the following Riccati equation

\[
\frac{dK_\mu}{dt} = AK_\mu + K_\mu A^T - K_\mu(H^T(GG^T)^{-1}H - \mu Q)K_\mu - GG^T
\]

\[
K_\mu(0) = 0
\]

and \( \mu \) is suitably constrained to ensure a unique, positive, symmetric solution of the Riccati equation. The dependence of \( Z \) on \( \mu \) is suppressed for notational convenience. It is useful to note for some subsequent computations that \((\int_0^T dY - HZds), \mathcal{G}(t), t \in [0, T])\) is a Brownian motion with incremental covariance \( GG^T \).

The control strategies for the multi-agents are adapted to the observation filtration. The uniqueness assumption for the Riccati equation places a restriction on \( \mu \) which is satisfied if it is also assumed that \( \mu \) is constrained so that uniformly for \( t \in [0, T], H^T(t)(G(t)G^T(t))^{-1}H(t) - \mu Q(t) > 0 \), that is, there is a constant \( \tilde{c} \) that does not depend on \( t \in [0, T] \) such that \(< (H^T(t)(G(t)G^T(t))^{-1}H(t) - \mu Q(t))x, x > \geq \tilde{c} < x, x > \) for each \( t \in [0, T] \) and \( x \in \mathbb{R}^n \). This estimation process \( Z \) is defined as follows

\[
Z(t) = \arg\min_{h \in \mathcal{H}} \mathbb{E}[\mu \exp(\frac{\mu}{2} \int_0^t < Q(s)(X(s) - h(s)), X(s) - h(s) > ds) | \mathcal{G}(t)]
\]

where \( \mathcal{H} \) is the family of square integrable \( \mathcal{G}(\cdot) \) progressively measurable processes on \([0, T], \mu > 0 \), and \( h(\cdot) \) is a generic element of \( \mathcal{H} \) e.g. [18].

The family of admissible feedback strategies is generically assumed to satisfy

\[(B1) \{ U : [0, T] \rightarrow \mathbb{R}^i \} \text{ where } U \text{ is adapted to the observation filtration } (G(t), t \in [0, T]) \text{ and } \int_0^T |U|^2 dt < \infty \text{ a.s. } \] and \( i \) is a suitable positive integer as well as the finiteness of the payoff given in (A1) to prevent indeterminate forms.

Let \((\tilde{P}_\mu(t), t \in [0, T])\) be the unique, positive, symmetric solution of the following Riccati equation

\[
\frac{d\tilde{P}_\mu}{dt} = -\tilde{P}_\mu(A + \mu K_\mu Q) - (A^T + \mu QK_\mu)\tilde{P}_\mu - Q + \tilde{P}_\mu(S_{k_1}B_iR_i^{-1}B_i^T - \sum_{j=1}^{k_2}C_jS_j^{-1}C_j^T)
\]

\[
\tilde{P}_\mu(0) = 0
\]
where \( K_\mu \) is the solution of (23) and it is assumed that \( \mu \) is also chosen to ensure uniqueness of this Riccati equation, e.g. if the following inequality is satisfied
\[
\sum_{i=1}^{k_1} B_i(t) R_i^{-1}(t) B_i^T(t) - \sum_{j=1}^{k_2} C_j(t) S_j^{-1}(t) C_j^T(t) - \mu K_\mu(t) H^T(t)(G^T(t)G(t))^{-1} \times H(t) K_\mu(t) > 0 \quad \text{uniformly for } t \in [0,T] \quad \text{and } I - \mu M K_\mu(T) > 0.
\]

The following result describes explicit optimal control strategies for the agents for this partially observed problem and shows that these strategies form a Nash equilibrium.

**Theorem 4.1.** The partially observed, multi-agent stochastic differential game given by (1), (3), and (20) with the admissible control strategies (B1) has a Nash equilibrium given by the following optimal control strategies
\[
U_i^{**}(t) = -R_i^{-1}(t) B_i^T(t) \tilde{P}_\mu(t) Z(t)
\]
for \( i = 1,...,k_1 \) and \( j = 1,...,k_2 \) where \( Z \) is the solution of the information filter (22) and \( \tilde{P}_\mu \) is the unique symmetric, positive solution of the Riccati equation (26).

The optimal payoff, \( J_\mu(U_1^{**},...,U_{k_1}^{**},V_1^{**},...,V_{k_2}^{**}) \), is
\[
J_\mu(U_1^{**},...,U_{k_1}^{**},V_1^{**},...,V_{k_2}^{**}) = \mu \text{Eexp}[\frac{\mu}{2} < \tilde{P}_\mu(0) X_0, X_0 > + \int_0^T \text{tr}(\tilde{P}_\mu K_\mu H^T(GG^T)^{-1} H K_\mu) dt + \int_0^T \text{tr}(K_\mu Q) dt] \times (\text{det}(I - \mu M K_\mu(T)))^{-\frac{1}{2}}
\]
where \( K_\mu \) is the solution of the filtering Riccati equation (23).

**Proof.** The proof for the optimal control strategies uses a refinement of a technique from the solution of the completely observable stochastic differential game given above. Initially the evaluation of the payoff is restricted to computations about the process \( Z \) instead of the process \( X \). The \((\mathcal{G}()\)) progressively measurable version of the integral quadratic term of the state in the payoff functional is
\[
\mu \text{Eexp}[\frac{\mu}{2} \int_0^T < QZ, Z > dt + \int_0^T \text{tr}(K_\mu Q) dt]
\]
This expression follows from (24) where \((Z(t), t \in [0,T])\) satisfies (22) and is a result from Gaussian measures e.g. [18]. Now an expression for the conditional expectation of the final condition is given.
\[
\text{E}[\text{exp}[\frac{\mu}{2} < MX(T), X(T) >] | Z(T)]
\]
\[
= \text{exp}[\frac{\mu}{2} < (I - \mu M K_\mu(T))^{-1} Z(T), Z(T) >]
\]
Furthermore,
\[
\text{Eexp}[\frac{\mu}{2} < (I - \mu M K_\mu(T))^{-1} Z(T), Z(T) >] = |\text{det}(I - \mu M K_\mu(T))|^{\frac{1}{2}}
\]
Thus the problem is reduced to considering the process $Z$. Apply the Ito formula to the process, \( \frac{1}{2} < \tilde{P}_\mu(t)Z(t), Z(t) >, t \in [0, T] \) where \((Z(t), t \in [0, T])\) satisfies (22) and $\tilde{P}_\mu$ is the solution of the Riccati equation (26) to obtain

\[
\frac{1}{2} < \tilde{P}_\mu(T)Z(T), Z(T) > = \frac{1}{2} < \tilde{P}_\mu(0)Z(0), Z(0) > + \frac{1}{2} \int_0^T (2(< \tilde{P}_\mu(A + \mu K_\mu Q)Z, Z > + < \tilde{P}_\mu(\Sigma_{i=1}^k B_i U_i + \Sigma_{j=1}^k C_j V_j), Z >) dt + 2 < \tilde{P}_\mu K_\mu H^T (dY - HZ dt), Z > + \text{tr}(\tilde{P}_\mu K_\mu H^T (GG^T)^{-1} H K_\mu) dt - < (\tilde{P}_\mu(A + \mu K_\mu Q) + (A^T + \mu Q K_\mu)\tilde{P}_\mu + Q)Z, Z > dt + 2 < \Sigma_{i=1}^k B_i^T \tilde{P}_\mu Z, U_i > + < \Sigma_{j=1}^k C_j^T \tilde{P}_\mu Z, V_i >) dt + \int_0^T [< \tilde{P}_\mu K_\mu H^T (dY - HZ dt), Z > - \mu < \tilde{P}_\mu K_\mu H^T (GG^T)^{-1} H K_\mu Z, \tilde{P}_\mu Z, Z > dt + \text{tr}(\tilde{P}_\mu K_\mu H^T (GG^T)^{-1} H K_\mu) dt]
\]

Recall that $X(0) = Z(0)$ because $X(0)$ is a constant vector. Let $L_\mu(U_1, ..., U_k, V_1, ..., V_k)$ be the integral terms in the quadratic functional that appear in the exponential of the payoff functional (3) by replacing $X$ by $Z$, that is,

\[
L_\mu(U_1, ..., U_k, V_1, ..., V_k) = \frac{\mu}{2} \int_0^T (< QZ, Z > + \Sigma_{i=1}^k < R_i U_i, U_i > - \Sigma_{i=1}^k < S_i V_i, V_i >) dt
\]

This quadratic functional, $L_\mu$, and (33) can be combined as

\[
L_\mu(U_1, ..., U_k, V_1, ..., V_k) + \frac{\mu}{2} < \tilde{P}_\mu(T)Z(T), Z(T) > - \frac{\mu}{2} < \tilde{P}_\mu(0)Z(0), Z(0) > = \frac{\mu}{2} \int_0^T (\Sigma_{i=1}^k (< R_i U_i, U_i > + < \tilde{P}_\mu B_i R_i^{-1} B_i^T \tilde{P}_\mu Z, Z > + < \tilde{P}_\mu B_i U_i, Z >) dt - \Sigma_{i=1}^k (< S_i V_j, V_j > - < \tilde{P}_\mu C_j S_j^{-1} C_j^T \tilde{P}_\mu Z, Z > + < \tilde{P}_\mu C_j V_j, Z >) dt + 2 < \tilde{P}_\mu C_j V_j, Z > dt + 2 < \tilde{P}_\mu K_\mu H^T (dY - HZ dt), Z > - \mu < \tilde{P}_\mu K_\mu H^T (GG^T)^{-1} H K_\mu \tilde{P}_\mu Z, Z > dt + \int_0^T \text{tr}(\tilde{P}_\mu K_\mu H^T (GG^T)^{-1} H K_\mu) dt)
\]
Thus the $Z$ measurable payoff is

$$
\mu E \exp \{ L_\mu(U_1, ..., U_k, V_1, ..., V_k) + \frac{\mu}{2} < \tilde{P}_\mu(T)Z(T), Z(T) > \} \tag{36}
$$

$$
= \mu (E \exp \{ \frac{\mu}{2} < \tilde{P}_\mu(0), Z(0), Z(0) > 
+ \frac{\mu}{2} \int_0^T (\Sigma_{i=1}^{k_i} |R_i^{-\frac{1}{2}} (R_i U_i + B_i^T \tilde{P}_\mu Z)|^2 dt
- \Sigma_{j=1}^{k_2} |S_j^{-\frac{1}{2}} (S_j V_j - C_j^T \tilde{P}_\mu Z)|^2 dt
+ \mu \int_0^T < \tilde{P}_\mu K_\mu H^T (dY -HZdt), Z >
- \frac{\mu^2}{2} \int_0^T < \tilde{P}_\mu K_\mu H^T (GG^T)^{-1} HK_\mu \tilde{P}_\mu Z, Z > dt
+ \frac{\mu}{2} \int_0^T tr(\tilde{P}_\mu K_\mu H^T (GG^T)^{-1} HK_\mu) dt \} \tag{37}
$$

where $\tilde{E}$ is the expectation for $\tilde{P}$ given by

$$
d\tilde{P} = \exp \{ \mu \int_0^T < \tilde{P}_\mu K_\mu H^T (dY -HZdt), Z >
- \frac{\mu^2}{2} \int_0^T < \tilde{P}_\mu K_\mu H^T (GG^T)^{-1} HK_\mu \tilde{P}_\mu Z, Z > dt \} d\tilde{P} \tag{38}
$$

Recall from the observation equation (20) and the likelihood function result [6] that $(\int_0^t (dY -HZdt), G(t), t \in [0, T])$ is a Brownian motion with the incremental covariance $GG^T$. The fact that the exponential in (38) is a Radon-Nikodym derivative, that is, it integrates to one, follows from the strong dichotomy for the absolute continuity of Gaussian measures (Feldman-Hajek theorem).
It is claimed that the strategies \((U_i^{**}, V_j^{**}, 1 = 1, ..., k_1; j = 1, ..., k_2)\) that are defined by making the quadratic terms in the last equality of (36) to be zero are optimal. To verify the optimality of these strategies, let \((U_i, V_j, i = 1, ..., k_1; j = 1, ..., k_2)\) be defined as

\[
\bar{U}_i = U_i^{**} + U_{i,1} \tag{39}
\]

\[
\bar{V}_j = V_j^{**} + V_{j,1} \tag{40}
\]

where \(U_{i,1} = \alpha_i\) for \(t \in [t_0^1, t_1^1]\) and \(\alpha_i\) is \(\mathcal{G}(t_0^1)\) measurable and uniformly bounded for \(i = 1, ..., k_1\) and similarly for \(V_{j,1} = \beta_j\) for \(t \in [t_0^1, t_1^1]\) for \(j = 1, ..., k_2\). The corresponding solution \(\bar{Z}\) for the information filter has a Radon-Nikodym derivative (38) so it is only necessary to consider the quadratic forms in the exponential terms from the last equality of (36). From these quadratic terms it follows directly that \((U_i^{**}, V_j^{**}, i = 1, ..., k_1; j = 1, ..., k_2)\) are optimal control strategies and form a Nash equilibrium so that for \(\mu > 0\)

\[
\text{E} \exp[L_\mu(U_1^{**}, ..., U_{k_1}^{**}, V_1^{**}, ..., V_{k_2}^{**})] \leq \text{E} \exp[L_\mu(\bar{U}_1, ..., \bar{U}_{k_1}, \bar{V}_1^{**}, ..., \bar{V}_{k_2}^{**})] \tag{41}
\]

\[
\text{E} \exp[L_\mu(U_1^{**}, ..., U_{k_1}^{**}, V_1^{**}, ..., V_{k_2}^{**})] \geq \text{E} \exp[L_\mu(U_1^{**}, ..., U_{k_1}^{**}, \bar{V}_1, ..., \bar{V}_{k_2})] \tag{42}
\]

Thus the optimal payoff is achieved by choosing the (optimal) feedback control strategies

\[
U_i^{**}(t) = -R_i^{-1}(t)B_i^T(t)\bar{P}_\mu(t)\bar{Z}(t) \tag{43}
\]

\[
V_j^{**}(t) = S_j^{-1}(t)C_j^T(t)\bar{P}_\mu(t)\bar{Z}(t) \tag{44}
\]

for \(i = 1, ..., k_1\) and \(j = 1, ..., k_2\), \(\bar{P}_\mu\) is the unique symmetric, positive solution of the Riccati equation (26) and the optimal payoff is

\[
\mu \text{E} \exp[L_\mu(U_1^{**}, ..., U_{k_1}^{**}, V_1^{**}, ..., V_{k_2}^{**})] + \frac{\mu}{2} < MX(T), X(T) > \tag{45}
\]

\[
= \mu \text{exp}\left[< \frac{\mu}{2}\bar{P}_\mu(0)X(0), X(0) > \right] + \frac{\mu}{2} \left( \int_0^T \text{tr}(\bar{P}_\mu K\bar{H}(GG^T)^{-1}HK_\mu)dt \right. \\
+ \left. \int_0^T \text{tr}(K_\mu Q)dt \right) |\text{det}(I - \mu MK_\mu(T))|^{-\frac{1}{2}}
\]

The determinant term arises from (32). Thus the optimal control strategies are suitable linear functions of the solution of the information filter.

5. Concluding remarks. The explicit solutions of some multi-agent linear-exponential-quadratic stochastic differential games with complete or partial observations are obtained using a direct method that does not require solving either nonlinear partial differential equations (Hamilton-Jacobi-Isaacs equations) or backward stochastic differential equations. This direct method is naturally related to the method for a two-person, completely observed linear-quadratic stochastic differential game [9]. Furthermore, this direct method of verification is not limited to linear or finite dimensional stochastic differential equations e.g. [8], [10].
REFERENCES

[1] M. Bardi and F. S. Priuli, Linear-quadratic n-person and mean-field games with ergodic cost, *SIAM J. Control Optim.*, 52 (2014), 3022–3052.

[2] T. Basar and P. Bernhard, *H∞-Optimal Control and Related Minimax Design Problems*, Birkhauser, Boston, 1995.

[3] T. Basar and G. Olsder, *Dynamic Noncooperative Game Theory*, SIAM, 2nd Ed. 1999.

[4] A. Bensoussan and J. H. van Schuppen, Optimal control of partially observable stochastic systems with an exponential-of-integral performance index, *SIAM J. Control Optim.*, 23 (1995), 599–613.

[5] R. Buckdahn and J. Li, Stochastic differential games and viscosity solutions of Hamilton-Jacobi-Bellman-Isaacs equations, *SIAM J. Control Optim.*, 47 (2008), 444–475.

[6] T. E. Duncan, Evaluation of likelihood functions, *Information and Control*, 13 (1968), 62–74.

[7] T. E. Duncan, Linear-exponential-quadratic Gaussian control, *IEEE Trans. Autom. Control*, 58 (2013), 2910–2911.

[8] T. E. Duncan, Some linear-quadratic stochastic differential games for equations in a Hilbert space with fractional Brownian motions, *Discrete Cont. Dyn. Systems Ser. A*, 35 (2015), 5435–5445.

[9] T. E. Duncan, Linear exponential quadratic stochastic differential games, *IEEE Trans. Autom. Control*, 61 (2016), 2550–2552.

[10] T. E. Duncan and B. Pasik-Duncan, A solvable stochastic differential game in the two-sphere, *Proc. 52nd IEEE Conf. Decision and Control*, 7833–7837, Firenze, 2013.

[11] T. E. Duncan and B. Pasik-Duncan, Some results on optimal control for a partially observed linear stochastic system with an exponential quadratic cost, *Proc. IFAC World Congress*, Cape Town, 47 (2014), 8695–8698.

[12] R. J. Elliott, L. Aggoun and J. B. Moore, *Hidden Markov Models*, Springer, 1995.

[13] W. H. Fleming and D. Hernandez-Hernandez, On the value of stochastic differential games, *Commun. Stoch. Anal.*, 5 (2011), 341–351.

[14] W. H. Fleming and P. E. Souganidis, On the existence of value functions of two player, zero sum stochastic differential games, *Indiana Math. J.*, 38 (1989), 293–314.

[15] S. Hamadene, J. P. Lepeltier and S. Peng, BSDEs with continuous coefficients and stochastic differential games, *Backward Stochastic Differential Equations* (N. El Karoui et al., eds.), Pitman Res. Notes Math, 364 (1997), 115–128.

[16] R. Isaacs, *Differential Games*, J. Wiley, New York 1965.

[17] D. H. Jacobson, Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games, *IEEE Trans. Autom. Control*, AC-18 (1973), 124–131.

[18] M. L. Kleptsyna, A. Le Breton and M. Viot, On the linear-exponential filtering problem for general Gaussian processes, *SIAM J. control Optim.*, 47 (2008), 2886–2911.

[19] J. B. Moore, R. J. Elliott and S. Dey, Risk sensitive generalizations of minimum variance estimation and control, *J. Math. Syst., Estimation, Contr.*, (electronic), 7 (1997), 15 pp.

[20] J. Nash, *Non-cooperative games*. *Ann. Math.*, 54 (1951), 286–295.

[21] J. L. Speyer, J. Deyst and D. H. Jacobson, Optimization of stochastic linear systems with additive measurement and process noise using exponential performance criteria, *IEEE Trans. Autom. Control*, AC-19 (1974), 358–366.

Received January 2018; revised August 2018.

E-mail address: duncan@ku.edu