Preprints of the 20th World Congress
The International Federation of Automatic Control
Toulouse, France, July 9-14, 2017

Linear Quadratic Stochastic Differential Games under Asymmetric Value of Information ⋆

Dipankar Maity and John S. Baras

Department of Electrical & Computer Engineering, Institute for Systems Research
University of Maryland, USA (e-mail: {dmaity, baras}@umd.edu).

Abstract: This paper considers a variant of two-player linear quadratic stochastic differential games. In this framework, none of the players has access to the state observations for all the time, which restricts the possibility of continuous feedback strategies. However, they can observe the state intermittently at discrete time instances by paying some finite cost. Having on demand costly measurements ensure that open-loop strategy is not the only strategy for this game. The individual cost functions for each player explicitly incorporate the value of information and the asymmetry that comes along with different costs of state observation for different players.
We study the structural properties of the Nash equilibrium for this particular class of problems when the cost of observation is finite and positive. We show that the game problem simplifies into two decoupled game problems: one for deciding the control strategies, and the other for deciding the observation acquisition times. The study also reveals that under two extreme cases -cost of observation being 0 or ∞- the strategies coincide with feedback and open-loop strategies respectively.

Keywords: game theory, linear-quadratic games, stochastic games, non-cooperative game, event-based game

1. INTRODUCTION

Game theory has been an active topic for research in control for its wide applicability in stochastic control, robust control; and it has been studied extensively by the community as can be found in Basar and Olsder (1995), James and Baras (1996), Engwerda (2005), Başar and Bernhard (2008), Fleming and Hernández-Hernández (2011) and many others. A differential stochastic game encompasses many aspects of a control problem such as optimality, stochasticity and filtering, and estimation; hence the results can reveal several properties related to those. Linear-quadratic-Gaussian is a subclass of such differential game problems that attain a closed form analytical solution for the Nash strategy as reported in Cruz Jr. and Chen (1971), Jacobson (1973), Weeren et al. (1999). The solution of linear-quadratic differential games are generally constructed by certain Riccati equation; for details, see Jacobson (1973), Weeren et al. (1999), and the references therein. Studies on the necessary and sufficient conditions for a strategy to be a Nash strategy for a linear-quadratic game can be found in the work of Foley and Schmitendorf (1971), and Bernhard (1979). Basar (1976) studied the uniqueness property of a Nash strategy. The work of Weeren et al. (1999) studies the asymptotic behavior of the Nash strategy over an infinite horizon.

⋆ Research partially supported by ARO grants W911NF-14-1-0384 and W911NF-15-1-0646, and by National Science Foundation (NSF) grant CNS-1544787.

Unlike the well perceived fact about linear control laws being optimal for a linear-quadratic-Gaussian problem, Basar (1974) provided a counterexample showing that the optimality is achieved by some nonlinear Nash strategy for a linear-quadratic game problem. A non-cooperative game is inherently a joint problem on the control and the decision making process and hence the solution relies on the knowledge of the behavior of the opponent.

In the vast majority of the past work in the community, the studied problems either assume that the state information is available to the players for all time or only the initial state information is available. The former situation results in a $feedback$-type Nash strategy whereas the latter exhibits an *openloop* Nash strategy. To the best of our knowledge, the problem of having multiple discrete state measurements for this problem class remained unaddressed. In this scenario, the players have less information about the state of the system since the measurements are available at only certain discrete time instances, not for all time; however, they have more information than merely having the knowledge of the initial state. In this work we address this linear-quadratic game problem under discrete measurements where the players are given the freedom to select their time instances to acquire the measurements of the state. Moreover, it is imposed that each such query about the state information requires some finite cost. This new framework introduces certain changes in the well known behavior of the Nash strategy since the feedback strategy is not plausible, and the open loop is not necessarily optimal. Given the fact that each observation requires finite cost, the players must decide optimal time instances for observing the state. Therefore, the problem includes designing a sampling policy to measure the state and synthesizing a controller such that they constitute a Nash equilibrium for the game.

In this work, we assume that the sampling is done instantaneously and there is no delay or noise in communicating the sampled value to the controller. We consider asynchronous switching i.e. players can choose their switching policy irrespective of the policy chosen by their opponent. We also assume that whenever a player receives a sample, the opponent is notified about that but the opponent does not get the value of the sample.

In this study, we show that given a switching policy, there always exists a Nash strategy for controller synthesis and the controller is a dynamic controller that resets its value in an optimal way every time the switch is closed. The problem is decomposed into two decoupled sub-problems for designing the switching policy and designing the controller. The studied game is asymmetric since the parameters associated with the players are different (e.g. cost per sample is different for different players) and that essentially leads to different strategy for them.

2. NOTATION

$x(t)$: state of the game, $\mathcal{C}_i$: controller of player-i, $\mathcal{S}_i$: switching policy of player-i, $\|a\|_B^2 = a'Ba$ for matrices a and b of proper dimensions, $\mathcal{T}_i(t)$: set of sampling times until t for player-i, $\mathcal{X}_i(t)$: set of sampled state values for player-i until t, $\mathcal{I}_i(t)$: total information available to player-i that includes $\mathcal{X}_i(t)$, $\mathcal{T}_1(t)$ and $\mathcal{T}_2(t)$. For any matrix M, $\Phi_M(t,s)$ denotes the associated state transition matrix.

3. PROBLEM FORMULATION

Let us consider the following stochastic linear differential game dynamics:

$$dx = (Ax + B_1u_1 + B_2u_2)dt + GdW_t \tag{1}$$

where $x \in \mathbb{R}^n$, $u_1 \in \mathbb{R}^{m_1}$, $u_2 \in \mathbb{R}^{m_2}$ and W_t is a p dimensional Wiener process noise, independent of the initial state $x(0)$, acting on the system. The associated quadratic cost is:

$$J(u_1,u_2) = \mathbb{E}\left[\int_0^T (x'Lx + u_1'R_1u_1 - u_2'R_2u_2)dt\right] \tag{2}$$

where $L, R_i \succ 0$. All the matrices A, B_i, L, R_i, G are time varying unless or otherwise mentioned in the paper.

The objective of player-1 (or player-2) is to minimize (or maximize) the cost functional (2) with the knowledge of $x(t)$ at some finite number of discrete time instances. Let us consider the schematic presented in Figure 1 where player-i has to design its controller $\mathcal{C}_i$ and the optimal switching policy $\mathcal{S}_i$. The switch $\mathcal{S}_i$ closes only for a time instance and opens immediately so that the controller gets the state value only at a single time instance. We assume there is no delay in the switching action or in the noise-less channel so that the controller $\mathcal{C}_i$ gets the state information precisely at the switching time instance.

Prior works on linear-quadratic differential games either consider the switches $\mathcal{S}_i$ are closed for all t or open for

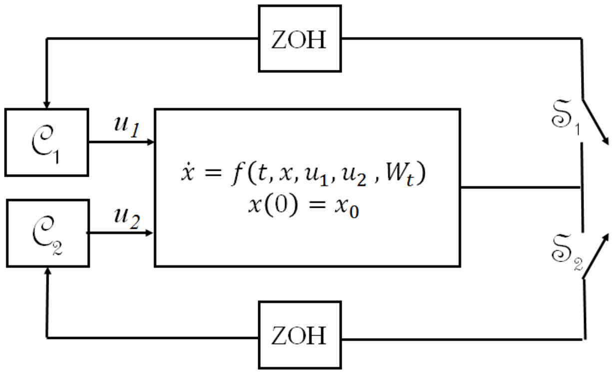

Fig. 1. Schematic of the game: $\mathcal{C}_i$ represents the controller (dynamic) of player-i and $\mathcal{S}_i$ implements a switch that samples the state x at some optimal instances. ZOH is a zero order hold circuit. The switches $\mathcal{S}_i$ are initially closed and they open at $t = 0^+$ so that each player has the knowledge of x_0.

all $t > 0$. We study the characteristics of the game and the associated Nash strategy(s) for this special set up of the game. Maity and Baras (2016a) studies the problem when the switches $\mathcal{S}_1$ and $\mathcal{S}_2$ operate synchronously i.e. they samples the state at the same time instances.

In this work we assume that the state x is fully observable, however the study easily extends under partial observation framework along the lines of Maity et al. (2017). Moreover, the players are given the freedom to select their switching instances for sampling the state by incurring a finite cost.

Let $\mathcal{T}_i(t) = \{\tau_1^i, \tau_2^i, \cdots, \tau_{n_i(t)}^i\}$ be the set of selected time instances for closing the switch $\mathcal{S}_i$ of player-i till time t, where $\tau_k^i < \tau_{k+1}^i < \cdots < \tau_{n_i(t)}^i < t$. Let $\mathcal{T}(t) = \mathcal{T}_1(t) \cup \mathcal{T}_2(t)$ denote the set of all instances for observing the state. The state information available to player-i at time t is denoted as $\mathcal{X}_i(t) = \{x(\tau) \mid \tau \in \mathcal{T}_i(t)\}$. The total information available to player-i at any time instance t is $\mathcal{I}_i(t) = \mathcal{X}_i(t) \cup \mathcal{T}(t)$. However as mentioned, the information acquisition is not free and in order to construct $\mathcal{I}_i(t)$, player-i needs to pay $\lambda_i (> 0)$ for each sample of the state and $c_{ij} (> 0)$ for each element in $\mathcal{T}_j$ (player-j is the opponent of player-i). Thus, it should be noted that the cost function $J(u_1,u_2)$ is implicitly a function of the information set $\mathcal{I}_1$ and $\mathcal{I}_2$ since the strategies u_1 and u_2 are $\mathcal{I}_1$ and $\mathcal{I}_2$ measurable functions respectively ($J(u_1,u_2) \equiv J(u_1,u_2,\mathcal{I}_1,\mathcal{I}_2)$).

Therefore, player-1 (P1) needs to minimize:

$$J_1(u_1,u_2,\mathcal{I}_1,\mathcal{I}_2) = \mathbb{E}\Big[\int_0^T (\|x\|_L^2 + \|u_1\|_{R_1}^2 - \|u_2\|_{R_2}^2)dt + \lambda_1 N_1 + c_{12}N_2\Big], \tag{3}$$

and player-2 (P2) should maximize:

$$J_2(u_1,u_2,\mathcal{I}_1,\mathcal{I}_2) = \mathbb{E}\Big[\int_0^T (\|x\|_L^2 + \|u_1\|_{R_1}^2 - \|u_2\|_{R_2}^2)dt - \lambda_2 N_2 - c_{21}N_1\Big]. \tag{4}$$

where $N_i = n_i(T)$ is the number of samples in $\mathcal{T}_i(T)$. In their respective cost functions, appropriate terms have been added to account for the cost of sampling (or could be thought as the cost of communication over the channel). It should be noted right away that the new cost functions do not allow infinite number of switching and hence the Zeno behavior is not possible for an optimal switching policy.

The game is studied under asymmetric information structure and remarks are made when the game is symmetric (i.e. $c_{12} = c_{21}$, $\lambda_1 = \lambda_2$, $B_1 = B_2$, $R_1 = R_2$).

4. NASH STRATEGY

In this section we aim to study the existence of Nash strategy for the proposed game framework. The Nash strategy includes designing a Nash switching policy ($\mathcal{S}_i$) and designing a Nash controller ($\mathcal{C}_i$) for both the players. We seek for Nash equilibrium solution $(u_1^*, u_2^*, \mathcal{I}_1^*, \mathcal{I}_2^*)$ such that

$$J_1^* = J_1(u_1^*, u_2^*, \mathcal{I}_1^*, \mathcal{I}_2^*) = \min_{u_1, \mathcal{I}_1} J_1(u_1, u_2^*, \mathcal{I}_1, \mathcal{I}_2^*) \quad (5)$$

$$J_2^* = J_2(u_1^*, u_2^*, \mathcal{I}_1^*, \mathcal{I}_2^*) = \max_{u_2, \mathcal{I}_2} J_2(u_1^*, u_2, \mathcal{I}_1^*, \mathcal{I}_2) \quad (6)$$

Note that for player-i, choosing an $\mathcal{I}_i^*$ essentially means to choose a $\mathcal{T}_i^*$. It can be easily shown that

$$J_1^* = \min_{\mathcal{I}_1} \left[\min_{u_1} J(u_1, u_2^*, \mathcal{I}_1, \mathcal{I}_2^*) + \lambda_1 N_1 \right] + c_{12} N_2^* \quad (7)$$

$$J_2^* = \max_{\mathcal{I}_2} \left[\max_{u_2} J(u_1^*, u_2, \mathcal{I}_1^*, \mathcal{I}_2) - \lambda_2 N_2 \right] + c_{21} N_1^* \quad (8)$$

Equations (7) and (8) decouple the problem that is the minimization (or maximization) is performed in two stages rather than in a single stage. Therefore, as a first step towards the proof, the Nash strategies $(u_1^*, u_2^*) \equiv (u^*(\mathcal{I}_1, \mathcal{I}_2), u_2^*(\mathcal{I}_1, \mathcal{I}_2))$ are found for the cost function $J(u_1, u_2, \mathcal{I}_1, \mathcal{I}_2)$ for a given $(\mathcal{I}_1, \mathcal{I}_2)$. Let us denote $J^{\#}(\mathcal{I}_1, \mathcal{I}_2)$ to be the value of $J(u_1, u_2, \mathcal{I}_1, \mathcal{I}_2)$ at the Nash equilibrium (u_1^*, u_2^*). That is, for a fixed $(\mathcal{I}_1, \mathcal{I}_2)$, and for any u_1, u_2,

$$J(u_1^*, u_2, \mathcal{I}_1, \mathcal{I}_2) \le J^{\#}(\mathcal{I}_1, \mathcal{I}_2) \le J(u_1, u_2^*, \mathcal{I}_1, \mathcal{I}_2). \quad (9)$$

Theorem 2 characterizes the Nash strategy (u_1^*, u_2^*).

Using completion of squares, it can be shown that (cf. Maity and Baras (2016b)):

$$J(u, u_2, \mathcal{I}_1, \mathcal{I}_2) = \mathbb{E}\big[\|x(0)\|^2_{P(0)}\big] + \int_0^T tr(PGG')dt \quad (10)$$
$$+ \mathbb{E}\left[\int_0^T (\|u_1 + R_1^{-1}B_1'Px\|^2_{R_1} - \|u_2 - R_2^{-1}B_2'Px\|^2_{R_2})dt\right]$$

where $P(t)$ satisfies the Riccati equation:

$$\dot{P} + A'P + PA + L + P(B_2R_2^{-1}B_2' - B_1R_1^{-1}B_1')P = \mathbf{0}$$
$$P(T) = \mathbf{0} \quad (11)$$

Assumption 1. In order to ensure the existence and well-definedness of the solution of the Riccati equation (11), we assume that $B_2R_2^{-1}B_2' - B_1R_1^{-1}B_1' \preceq \mathbf{0}$

The admissible strategy $u_i(t)$ has to be $\mathcal{I}_i(t)$ measurable.

Lemma 1. For a given switching $\mathcal{I}_i$, the optimal control strategy for player-i is of the form,

$$u_i^*(t) = (-1)^i R_i^{-1}B_i'P\hat{x}_i(t) \quad (12)$$

for some $\mathcal{I}_i(t)$ measurable optimal $\hat{x}_i(t)$.

A proof this lemma can be found in (Maity et al., 2017, Proposition 3.2). Therefore, the goal is to find the Nash equilibrium of

$$\mathcal{J}(\hat{x}_1, \hat{x}_2) = \mathbb{E}\left[\int_0^T (\|\hat{x}_1 - x\|^2_{Q_1} - \|\hat{x}_2 - x\|^2_{Q_2})dt\right]$$

for a given $(\mathcal{T}_1, \mathcal{T}_2)$ (or equivalently $(\mathcal{I}_1, \mathcal{I}_2)$), where $Q_i = PB_iR_i^{-1}B_i'P$.

Theorem 2. $\mathcal{J}$ has a unique saddle point at $(\hat{x}_1^*, \hat{x}_2^*)$ such that:

$$\mathcal{J}(\hat{x}_1^*, \hat{x}_2) \le \mathcal{J}(\hat{x}_1^*, \hat{x}_2^*) \le \mathcal{J}(\hat{x}_1, \hat{x}_2^*) \quad (13)$$

for all $\hat{x}_1, \hat{x}_2$. The optimal $\hat{x}_1^*$ and $\hat{x}_2^*$ satisfy the following differential equations:

$$\dot{\hat{x}}_1^* = (A - P^{-1}Q_1 + P^{-1}Q_2)\hat{x}_1^* \quad (14)$$
$$\hat{x}_1^*(\tau_1) = x(\tau_1)$$

$$\dot{\hat{x}}_2^* = (A - P^{-1}Q_1 + P^{-1}Q_2)\hat{x}_2^* \quad (15)$$
$$\hat{x}_2^*(\tau_2) = x(\tau_2)$$

for all $\tau_1 \in \mathcal{T}_1$ and $\tau_2 \in \mathcal{T}_2$.

Proof: The proof of this theorem is presented in the Appendix A.

Therefore, Theorem 2 ensures that, for a fixed switching pair $(\mathcal{T}_1, \mathcal{T}_2)$, there exists a Nash controller pair $(\mathcal{C}_1^*, \mathcal{C}_2^*)$. Using the results from Theorem 2, we can write (the $*$ are removed from $\hat{x}_i^*$ for brevity),

$$d\hat{x}_1 - dx = (A + P^{-1}Q_2)(\hat{x}_1 - x)dt + P^{-1}Q_2(x - \hat{x}_2)dt - GdW_t \quad (16)$$
$$d\hat{x}_2 - dx = (A - P^{-1}Q_1)(\hat{x}_2 - x)dt - P^{-1}Q_1(x - \hat{x}_1)dt - GdW_t$$

with the resetting conditions $\hat{x}_1(\tau_1) = x(\tau_1)$ and $\hat{x}_2(\tau_2) = x(\tau_2)$ for all $\tau_1 \in \mathcal{T}_1$ and $\tau_2 \in \mathcal{T}_2$. Denoting $z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} \hat{x}_1 - x \\ \hat{x}_2 - x \end{bmatrix}$, we can write

$$dz = \begin{bmatrix} A + P^{-1}Q_2 & -P^{-1}Q_2 \\ P^{-1}Q_1 & A - P^{-1}Q_1 \end{bmatrix} zdt - \begin{bmatrix} GdW_t \\ GdW_t \end{bmatrix}$$

with $z_i(\tau) = 0$ for all $\tau \in \mathcal{T}_i$, $i = 1, 2$.

Let us denote $\Sigma(t) = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} = \mathbb{E}[z(t)z'(t)]$. Therefore,

$$\dot{\Sigma} = \bar{A}\Sigma + \Sigma\bar{A}' + \bar{G}\bar{G}' \quad (17)$$

where $\bar{G} = \begin{bmatrix} G \\ G \end{bmatrix}$, and $\bar{A} = \begin{bmatrix} A + P^{-1}Q_2 & -P^{-1}Q_2 \\ P^{-1}Q_1 & A - P^{-1}Q_1 \end{bmatrix}$. At $\tau \in \mathcal{T}_1$, $\Sigma_{11}(\tau) = \Sigma_{12}(\tau) = \Sigma_{21}(\tau) = 0$ and at $\tau \in \mathcal{T}_2$, $\Sigma_{22}(\tau) = \Sigma_{12}(\tau) = \Sigma_{21}(\tau) = 0$.

The solution of (17) is given by,

$$\Sigma(t) = \Phi_{\bar{A}}(t, s)\Sigma(s)\Phi_{\bar{A}}(t, s)' + \int_s^t \Phi_{\bar{A}}(t, r)\bar{G}\bar{G}'\Phi_{\bar{A}}'(t, r)dr$$

where $\forall\sigma \in (s, t]$, $\sigma \notin \mathcal{T}(t)$.

One can show that, $\Phi_{\bar{A}}(t, s) = \begin{bmatrix} \Phi_1(t, s) & \Phi_2(t, s) \\ \Phi_3(t, s) & \Phi_4(t, s) \end{bmatrix}$ has the following expression (we leave out the details here).

$$\Phi_1(t, s) = \Phi_A(t, s) + \int_s^t \Phi_A(t, \sigma)P^{-1}(\sigma)Q_2(\sigma)\Phi_{\tilde{A}}(\sigma, s)d\sigma$$

where $\tilde{A} = A - P^{-1}Q_1 + P^{-1}Q_2$.

$$\Phi_2(t, s) = -\int_s^t \Phi_A(t, \sigma)P^{-1}(\sigma)Q_2(\sigma)\Phi_{\tilde{A}}(\sigma, s)d\sigma,$$

$$\Phi_3(t,s) = \int_s^t \Phi_A(t,\sigma)P^{-1}(\sigma)Q_1(\sigma)\Phi_{\tilde{A}}(\sigma,s)d\sigma$$

and,

$$\Phi_4(t,s) = \Phi_A(t,s) - \int_s^t \Phi_A(t,\sigma)P^{-1}(\sigma)Q_1(\sigma)\Phi_{\tilde{A}}(\sigma,s)d\sigma.$$

Since $\Sigma_{11} = [I\ O]\,\Sigma \begin{bmatrix} I \\ O \end{bmatrix}$ and $\Sigma_{11} = [O\ I]\,\Sigma \begin{bmatrix} O \\ I \end{bmatrix}$ we get,

$$\Sigma_{11}(t) = \int_s^t \Phi_A(t,\sigma)G(\sigma)G(\sigma)'\Phi_A'(\sigma,s)d\sigma + \\ \Phi_1(t,s)\Sigma_{11}(s)\Phi_1'(t,s) + \Phi_1(t,s)\Sigma_{12}(s)\Phi_2'(t,s) \\ + \Phi_2(t,s)\Sigma_{21}(s)\Phi_1'(t,s) + \Phi_2(t,s)\Sigma_{22}(s)\Phi_2'(t,s) \tag{18}$$

Similar expression can be found for Σ_{22} as well.

Therefore, the game problem at this stage is represented with the objective

$$\mathcal{J}(\hat{x}_1^*, \hat{x}_2^*) = J^{\#}(\mathcal{T}_1, \mathcal{T}_2) = \int_0^T tr(Q_1\Sigma_{11} - Q_2\Sigma_{22})dt \\ = \int_0^T tr(\bar{Q}\Sigma)dt \tag{19}$$

where $\bar{Q} = \begin{bmatrix} Q_1 & 0 \\ 0 & -Q_2 \end{bmatrix}$. Σ satisfies the dynamics:

$$\dot{\Sigma} = \bar{A}\Sigma + \Sigma\bar{A}' + \bar{G}\bar{G}' \tag{20}$$

At $\tau \in \mathcal{T}_i$, $\Sigma_{ii}(\tau) = \Sigma_{ij}(\tau) = \Sigma_{ji}(\tau) = 0$ for all $i, j = 1, 2$.

The dynamic game at this point has a linear dynamics (20) with a linear cost criterion (19), however, the actions of the players are switching actions i.e. the actions partially reset the value of Σ. Clearly, when player-1 strategically selects a switching instance τ_1, it resets Σ_{11} to zero and consequently the cost is reduced, however this reduction in the cost comes with an additional switching cost of λ_1.

The objective of player-1 is to minimize $J^{\#}(\mathcal{T}_1, \mathcal{T}_2) + \lambda_1 n_1$ (since the other term, $c_{12}n_2$, depends solely on the opponent's action) and player-2 aims to maximize $J^{\#}(\mathcal{T}_1, \mathcal{T}_2) - \lambda_2 n_2$, where n_i is the cardinality of $\mathcal{T}_i(T)$.

It should be noted at this point that the game is totally characterized by $\Sigma(t)$ which can uniquely be determined whenever the switching instances are known. This game subproblem is decoupled from the game subproblem seeking the Nash control strategies (Theorem 2). Due to space constraints, solving for the Nash switching strategy is beyond the scope of this paper. Instead, we will assume a solution to this switching game problem exists and we will provide some characterization of the solution in the rest of the paper.

If $(\mathcal{T}_1^*, \mathcal{T}_2^*)$ is an equilibrium strategy for (19) with optimal number of switching being (n_1^*, n_2^*), then we have

$$J^{\#}(\mathcal{T}_1^*, \mathcal{T}_2) - \lambda_2(n_2 - n_2^*) \le J^{\#}(\mathcal{T}_1^*, \mathcal{T}_2^*) \\ \le J^{\#}(\mathcal{T}_1, \mathcal{T}_2^*) + \lambda_1(n_1 - n_1^*) \tag{21}$$

for all $\mathcal{T}_1$ and $\mathcal{T}_2$ with cardinality being n_1 and n_2 respectively. For all $(\mathcal{T}_1, \mathcal{T}_2)$ such that $n_i = n_i^*$, we obtain:

$$J^{\#}(\mathcal{T}_1^*, \mathcal{T}_2) \le J^{\#}(\mathcal{T}_1^*, \mathcal{T}_2^*) \le J^{\#}(\mathcal{T}_1, \mathcal{T}_2^*). \tag{22}$$

It can be shown that for any switching strategy played by Player-2, if player-1 selects no-switching strategy (i.e. does not attempt to reset $\Sigma(t)$), denoted by $\mathcal{T}_1^0$, then the following inequality holds:

$$J^{\#}(\mathcal{T}_1, \mathcal{T}_2) \le \int_0^T tr(Q_1\Sigma_0)dt - \lambda_1 n_1 \tag{23}$$

for all $(\mathcal{T}_1, \mathcal{T}_2)$ with cardinalities n_1 and n_2 respectively, and $\Sigma_0(t) = \int_0^t \Phi_A(t,s)GG'\Phi_A'(s,0)ds$. Similarly for all $(\mathcal{T}_1, \mathcal{T}_2)$, it can be shown that

$$J^{\#}(\mathcal{T}_1, \mathcal{T}_2) \ge -\int_0^T tr(Q_2\Sigma_0)dt + \lambda_2 n_2. \tag{24}$$

Combining (23) and (24),

$$-\int_0^T tr(Q_1\Sigma_0)dt \le J^{\#}(\mathcal{T}_1, \mathcal{T}_2) \le \int_0^T tr(Q_1\Sigma_0)dt \tag{25}$$

for all $(\mathcal{T}_1, \mathcal{T}_2)$.

Proposition 3. If n_i^* is the number of switchings of player-i at equilibrium, then

$$n_i^* \le \frac{1}{\lambda_i}\int_0^T tr\big((Q_1 + Q_2)\Sigma_0\big)dt.$$

Proof: From (23), $J^{\#}(\mathcal{T}_1^*, \mathcal{T}_2^*) \le \int_0^T tr(Q_1\Sigma_0)dt - \lambda_1 n_1^*$. Using (25), we obtain $-\int_0^T tr(Q_1\Sigma_0)dt \le \int_0^T tr(Q_1\Sigma_0)dt - \lambda_1 n_1^*$. Hence,

$$n_1^* \le \frac{1}{\lambda_i}\int_0^T tr\big((Q_1 + Q_2)\Sigma_0\big)dt.$$

Similarly we can proceed for n_2^*.

Proposition 3 provides an upper bound on the number of equilibrium switchings and it is inversely proportional to the cost of switching λ_i, as expected. Also notice that as $\lambda_i \to 0$, upper bound on $n_i^* \to \infty$, resembling the continuous-closed-loop strategy. Also, when $\lambda_i \to \infty$, $n_i^* \to 0$ resembling the open-loop strategy.

Let $(\mathcal{T}_1^*, \mathcal{T}_2^*)$ and $(\mathcal{T}_3^*, \mathcal{T}_4^*)$ two distinct equilibrium strategies with number of switchings equal to (n_1^*, n_2^*) and (n_3^*, n_4^*) respectively. Therefore,

$$J^{\#}(\mathcal{T}_1^*, \mathcal{T}_2^*) \le J^{\#}(\mathcal{T}_3^*, \mathcal{T}_2^*) + (n_3^* - n_1^*)\lambda_1 \\ \le J^{\#}(\mathcal{T}_3^*, \mathcal{T}_4^*) + (n_3^* - n_1^*)\lambda_1 - (n_4^* - n_2^*)\lambda_2 \\ \le J^{\#}(\mathcal{T}_1^*, \mathcal{T}_4^*) - (n_4^* - n_2^*)\lambda_2 \\ \le J^{\#}(\mathcal{T}_1^*, \mathcal{T}_2^*)$$

Therefore, all the inequalities in the above equation should be equalities, and that results into

$$J^{\#}(\mathcal{T}_1^*, \mathcal{T}_2^*) + n_1^*\lambda_1 - n_2^*\lambda_2 = J^{\#}(\mathcal{T}_3^*, \mathcal{T}_4^*) + n_3^*\lambda_1 - n_4^*\lambda_2.$$

If there exist two equilibria $(\mathcal{T}_1^*, \mathcal{T}_2^*)$ and $(\mathcal{T}_3^*, \mathcal{T}_4^*)$ such that $n_1^* = n_3^*$ and $n_2^* = n_4^*$, then $J^{\#}(\mathcal{T}_1^*, \mathcal{T}_2^*) = J^{\#}(\mathcal{T}_3^*, \mathcal{T}_4^*)$. Moreover, the costs incurred by both the players at these two different equilibria remain the same for both the players. We conclude this section by citing some results for a symmetric game.

4.1 Symmetric Games

In this section, we extend the results for a game where $\lambda_1 = \lambda_2 = \lambda$ and $B_1R_1^{-1}B_1' = B_2R_2^{-1}B_2'$. In this case, $\tilde{A} = A$, and $Q_1 = Q_2 = Q$.

One can verify that $J^{\#}(\mathcal{T}, \mathcal{T}) = 0$ for all switching strategy $\mathcal{T}$ (with n number of switchings). In this case, the cost of player-1 is $n\lambda$ and the same for player-2 is $-n\lambda$.

Let $(\mathcal{T}_1^*, \mathcal{T}_2^*)$ be a equilibrium switching strategy for under this symmetric situation. Let n_i^* be the number of elements in $\mathcal{T}_i^*$. Therefore,

$$J^{\#}(\mathcal{T}_1^*, \mathcal{T}_2^*) + n_1^*\lambda \leq J^{\#}(\mathcal{T}_2^*, \mathcal{T}_2^*) + n_2^*\lambda = n_2^*\lambda \quad (26)$$

and

$$J^{\#}(\mathcal{T}_1^*, \mathcal{T}_2^*) - n_2^*\lambda \geq J^{\#}(\mathcal{T}_1^*, \mathcal{T}_1^*) - n_1^*\lambda = -n_1^*\lambda \quad (27)$$

Combining the above two inequalities,

$$J^{\#}(\mathcal{T}_1^*, \mathcal{T}_2^*) = (n_2^* - n_1^*)\lambda \quad (28)$$

Under this situation, the cost incurred by player-1 is $n_2^*\lambda$ (≥ 0) and by player-2 is $-n_1^*\lambda$ (≤ 0).

Remark 4. If the players cooperate, then no-switching for all $t > 0$ produces the best cost for both the players. In this situation $J^{\#} = 0$ and the costs incurred by both the players are 0.

5. CONCLUSION

In this work we have considered a variant of two-player linear quadratic games where we restricted the possibility of feedback strategies by putting a finite cost for accessing the state. This work studies the structure of the Nash controllers of the players under this asymmetric game setup. We show that the game problem can be solved by independently solving two simpler game problems.

The costly switching behavior makes this problem challenging and interesting. The results show that for a high enough cost ($\lambda_i > \int_0^T tr\big((Q_1 + Q_2)\Sigma_0\big)dt$) player-$i$ opts for openloop strategy (i.e. $\mathcal{S}_i$ is always open). Thus, the results show the 'value of information' of a sample obtained by performing the switching. It is a trade-off between the reduction in cost (J) by sampling the state, and the cost incurred to obtain the state information.

REFERENCES

Basar, T. (1976). On the uniqueness of the Nash solution in linear-quadratic differential games. *International Journal of Game Theory*, 5(2-3), 65–90.

Basar, T. (1974). A counterexample in linear-quadratic games: Existence of nonlinear Nash solutions. *Journal of Optimization Theory and Applications*, 14(4), 425–430.

Başar, T. and Bernhard, P. (2008). *H-infinity optimal control and related minimax design problems: a dynamic game approach*. Springer Science & Business Media.

Basar, T. and Olsder, G.J. (1995). *Dynamic noncooperative game theory*, volume 200. SIAM.

Bernhard, P. (1979). Linear-quadratic, two-person, zero-sum differential games: necessary and sufficient conditions. *Journal of Optimization Theory and Applications*, 27(1), 51–69.

Cruz Jr., J. and Chen, C. (1971). Series Nash solution of two-person, nonzero-sum, linear-quadratic differential games. *Journal of Optimization Theory and Applications*, 7(4), 240–257.

Engwerda, J. (2005). *LQ dynamic optimization and differential games*. John Wiley & Sons.

Fleming, W.H. and Hernández-Hernández, D. (2011). On the value of stochastic differential games. *Commun. Stoch. Anal*, 5(2), 341–351.

Foley, M. and Schmitendorf, W. (1971). On a class of nonzero-sum, linear-quadratic differential games. *Journal of Optimization Theory and Applications*, 7(5), 357–377.

Jacobson, D.H. (1973). Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games. *Automatic Control, IEEE Transactions on*, 18(2), 124–131.

James, M.R. and Baras, J. (1996). Partially observed differential games, infinite-dimensional hamilton-jacobi-isaacs equations, and nonlinear h_∞ control. *SIAM Journal on Control and Optimization*, 34(4), 1342–1364.

Maity, D. and Baras, J.S. (2016a). Optimal strategies for stochastic linear quadratic differential games with costly information. In *Decision and Control (CDC), 2016 IEEE 55th Conference on*, 276–282. IEEE.

Maity, D. and Baras, J.S. (2016b). Strategies for two-player differential games with costly information. In *Discrete Event Systems (WODES), 2016 13th International Workshop on*, 211–216. IEEE.

Maity, D., Raghavan, A., and Baras, J.S. (2017). Stochastic differential linear-quadratic games with intermittent asymmetric observations. In *American Control Conference (ACC), 2017*, (to appear). IEEE.

Weeren, A., Schumacher, J., and Engwerda, J. (1999). Asymptotic analysis of linear feedback Nash equilibria in nonzero-sum linear-quadratic differential games. *Journal of Optimization Theory and Applications*, 101(3), 693–722.

Appendix A. PROOF OF THEOREM 2:

Player-1 wants to minimize

$$\mathcal{J}(\hat{x}_1, \hat{x}_2) = \int_0^T \mathbb{E}\left[\|\hat{x}_1 - x\|^2_{Q_1} - \|\hat{x}_2 - x\|^2_{Q_2}| \ \mathcal{I}_1\right] dt \quad \text{(A.1)}$$

whereas, player-2 wants to maximize (with slight abuse of notation)

$$\mathcal{J}(\hat{x}_1, \hat{x}_2) = \int_0^T \mathbb{E}\left[\|\hat{x}_1 - x\|^2_{Q_1} - \|\hat{x}_2 - x\|^2_{Q_2} \mid \mathcal{I}_2\right] dt \quad \text{(A.2)}$$

Let us denote the solution of (1) as:

$$x(t) = \Phi_A(t, t_0)x(t_0) + K_1^{t,t_0}[\hat{x}_1](t) + K_2^{t,t_0}[\hat{x}_2](t) + K_3^{t,t_0}[W](t) \quad \text{(A.3)}$$

for $t \geq t_0$. K_i^{t,t_0} for $i = 1, 2, 3$ are linear operators defined as follows:

$$K_i^{t,t_0}[f](t) = (-1)^i \int_{t_0}^t \Phi_A(t, s)P^{-1}(s)Q_i(s)f(s)ds, \quad \text{(A.4)}$$

for $i = 1, 2$ and

$$K_3^{t,t_0}[W](t) = \int_{t_0}^t \Phi_A(t, s)G(s)dW(s). \quad \text{(A.5)}$$

We seek the Nash equilibrium $(\hat{x}_1^*, \hat{x}_2^*)$ of (A.1). We proceed using calculus of variation technique by studying the first and second order Gateaux differentials of the cost functional $\mathcal{J}$ and finally we look for a saddle point of $\mathcal{J}$.

Let us calculate the Gateaux differential of the functional $\mathcal{J}$:

$$\delta\mathcal{J}[\hat{x}_1,\hat{x}_2](h_1,h_2) = \lim_{\epsilon\to 0}\frac{\mathcal{J}(\hat{x}_1+\epsilon h_1,\hat{x}_2+ah_2)-\mathcal{J}(\hat{x}_1,\hat{x}_2)}{\epsilon} \quad \text{(A.6)}$$

where the notation $\delta\mathcal{J}[\hat{x}_1,\hat{x}_2](h_1,h_2)$ means the Gateaux differential of $\mathcal{J}$ evaluated at the point $(\hat{x}_1,\hat{x}_2)$ in the direction (h_1,h_2). Note that $\mathcal{J}[\hat{x}_1,\hat{x}_2](\cdot,\cdot)$ is a linear functional parameterized by $\hat{x}_1$ and $\hat{x}_2$.

Let us denote $x^{h_1,h_2}(t)$ to be the perturbed solution of the following dynamics:

$$\begin{aligned} x^{h_1,h_2}(t) &= \Phi_A(t,0)x(0) + K_1^{t,0}[\hat{x}_1+h_1](t) \\ &+ K_2^{t,0}[\hat{x}_2+h_2](t) + K_3^{t,0}[W](t), \end{aligned} \quad \text{(A.7)}$$

Therefore,

$$\begin{aligned} &\frac{1}{2}\delta\mathcal{J}[\hat{x}_1,\hat{x}_2](h_1,h_2) = \\ &\int_0^T \mathbb{E}\Big[\big(\langle\hat{x}_1-x,\ h_1-K_1^{t,0}[h_1]-K_2^{t,0}[h_2]\rangle_{Q_1} - \\ &\langle h_2-K_1^{t,0}[h_1]-K_2^{t,0}[h_2],\ \hat{x}_2-x\rangle_{Q_2}\big)\mid\mathcal{I}_1\Big]dt \end{aligned} \quad \text{(A.8)}$$

where $\langle a,b\rangle_C = a'Cb$ and $a,b,\ C$ are matrices (or vectors) of compatible dimensions.

In order for $(\hat{x}_1,\hat{x}_2)$ to be a Nash Equilibrium (saddle point of $\mathcal{J}$), the necessary condition is $\delta\mathcal{J}[\hat{x}_1,\hat{x}_2](h_1,h_2)=0$ for all (h_1,h_2).

Let us consider

$$h_2(t) = -\int_0^t \Phi_{\tilde{A}_2}(t,s)P^{-1}Q_1h_1ds \quad \text{(A.9)}$$

where $\Phi_{\tilde{A}_2}$ is the state transition matrix corresponding to the drift matrix $A+P^{-1}Q_2$ i.e. $\dot{\Phi}_{\tilde{A}_2}(t,s) = (A(t)+P^{-1}(t)Q_2(t))\Phi_{\tilde{A}_2}(t,s)$ and $\Phi_{\tilde{A}_2}(s,s)=I_{n\times n}$

This choice of (h_1,h_2) implies,

$$\begin{aligned} &K_2^{t,0}[h_2](t) \\ &= -\int_0^t \Phi_A(t,s)P^{-1}Q_2\int_0^s \Phi_{\tilde{A}_2}(s,\sigma)P^{-1}Q_1\ h_1\ d\sigma ds \\ &= -\int_0^t\left[\int_\sigma^t \Phi_A(t,s)P^{-1}Q_2\Phi_{\tilde{A}_2}(s,\sigma)ds\right]P^{-1}Q_1\ h_1\ d\sigma \\ &= -\int_0^t\left[\int_\sigma^t \frac{d}{ds}(\Phi_A(t,s)\Phi_{\tilde{A}_2}(s,\sigma))ds\right]P^{-1}Q_1\ h_1\ d\sigma \\ &= h_2(t) - \int_0^t \Phi_A(t,s)(-P^{-1}Q_1\ h_1\ )ds \\ &= h_2(t) - K_1^{t,0}[h_1](t). \end{aligned}$$

Substituting these (h_1,h_2), we obtain

$$\begin{aligned} &\frac{1}{2}\delta\mathcal{J}[\hat{x}_1,\hat{x}_2](h_1,h_2) = \\ &\int_0^T \mathbb{E}\Big[\big(\langle\hat{x}_1-x,\ h_1-h_2\rangle_{Q_1}\big)\mid\mathcal{I}_1\Big]dt = 0 \end{aligned} \quad \text{(A.10)}$$

Equation (A.10) holds true for all choices of h_1. Thus the necessary condition becomes,

$$\mathbb{E}[x(t)-\hat{x}_1(t)\mid\mathcal{I}_1(t)] = 0 \quad \text{(A.11)}$$

or $\mathbb{E}[x(t)\mid\mathcal{I}_1(t)] = \hat{x}_1(t)$ for all t.

Similarly,

$$h_1(t) = \int_0^t \Phi_{\tilde{A}_1}(t,s)P^{-1}Q_2h_2ds \quad \text{(A.12)}$$

implies $K_1^{t,0}[h_1](t) + K_2^{t,0}[h_2](t) = h_1(t)$. With this pair of (h_1,h_2),

$$\begin{aligned} &\frac{1}{2}\delta\mathcal{J}[\hat{x}_1,\hat{x}_2](h_1,h_2) = \\ &\int_0^T \mathbb{E}\Big[\big(\langle\hat{x}_2-x,\ h_1-h_2\rangle_{Q_2}\big)\mid\mathcal{I}_1\Big]dt \end{aligned} \quad \text{(A.13)}$$

Therefore, another necessary condition is:

$$\mathbb{E}[x(t)-\hat{x}_2(t)\mid\mathcal{I}_1(t)] = 0 \quad \text{(A.14)}$$

or $\mathbb{E}[\hat{x}_2(t)\mid\mathcal{I}_1(t)] = \hat{x}_1(t)$.

Similarly, one can show using (A.2) that the following relations hold:

$$\hat{x}_2(t) = \mathbb{E}[x(t)\mid\mathcal{I}_2(t)] \quad \text{(A.15)}$$

$$\hat{x}_2(t) = \mathbb{E}[\hat{x}_1(t)\mid\mathcal{I}_2(t)] \quad \text{(A.16)}$$

From (A.11) and (A.14), and using the fact $\mathbb{E}[W_t\mid\mathcal{I}_1(t)] = 0$ for all t, we obtain $\forall t$

$$\begin{aligned} \dot{\hat{x}}_1 &= (A-P^{-1}Q_1+P^{-1}Q_1)\hat{x}_1 \\ \hat{x}_1(\tau_1) &= x(\tau_1) \end{aligned} \quad \text{(A.17)}$$

for all $\tau_1\in\mathcal{T}_1$.

Similarly by considering (A.2), one can show that $\forall t$,

$$\begin{aligned} \dot{\hat{x}}_2 &= (A-P^{-1}Q_1+P^{-1}Q_1)\hat{x}_2 \\ \hat{x}_2(\tau_2) &= x(\tau_2) \end{aligned} \quad \text{(A.18)}$$

for all $\tau_2\in\mathcal{T}_2$.

Therefore, equations (A.17) and (A.18) are necessary conditions for $\hat{x}_1$ and $\hat{x}_2$ to be a Nash Equilibrium.

To prove that $\hat{x}_1$ and $\hat{x}_2$ satisfying (A.17),(A.18) are a saddle point pair (hence Nash Equilibrium) for $\mathcal{J}$, we need to evaluate the second order Gateaux differential of $\mathcal{J}$. We do not present the details of this derivation due to space limitation, but one can check that:

$$\frac{1}{2}\delta^2\mathcal{J}[\hat{x}_1,\hat{x}_2](h_1,h_2) = D_1 - D_2 \quad \text{(A.19)}$$

where for $i=1,2$,

$$D_i = \int_0^T \|h_i - K_1^{t,0}[h_1] - K_2^{t,0}[h_2]\|_{Q_i}^2. \quad \text{(A.20)}$$

We need to prove that $\delta^2\mathcal{J}[\hat{x}_1,\hat{x}_2]$ is indefinite i.e. depending on the direction (h_1,h_2), $\delta^2\mathcal{J}$ can be positive as well as negative. Let us consider a (h_1,h_2) pair such that $h_2\neq 0$ identically for all t and

$$h_1(t) = \int_0^t \Phi_{\tilde{A}_1}(t,s)P^{-1}(s)Q_2(s)h_2(s)ds \quad \text{(A.21)}$$

Therefore for some h_2 (say h_2= constant) we have $D_2>0$ and $D_1=0$ implying that $\delta^2\mathcal{J}[\hat{x}_1,\hat{x}_2](h_1,h_2)<0$. Also in a similar fashion, by choosing

$$h_2(t) = -\int_0^t \Phi_{\tilde{A}_2}(t,s)P^{-1}(s)Q_1(s)h_1(s)ds \quad \text{(A.22)}$$

one can show that $\delta^2\mathcal{J}[\hat{x}_1,\hat{x}_2](h_1,h_2)>0$.

This proves that the pair $(\hat{x}_1,\hat{x}_2)$ satisfying (A.17)-(A.18) is a saddle point of $\mathcal{J}$.

The uniqueness of the saddle point is due to the uniqueness property of the solution of a linear differential equation. □