2017 American Control Conference
Sheraton Seattle Hotel
May 24–26, 2017, Seattle, USA

Stochastic Differential Linear-Quadratic Games with Intermittent Asymmetric Observations

Dipankar Maity, Aneesh Raghavan and John S. Baras

***Abstract*— In this paper, we consider a two-players stochastic linear quadratic game framework. The game is partially observed and each player has their own private observation. The challenge is that none of the players has access to the continuum observations, rather they can access their respective observations at discrete time instances by operating a switch unanimously. The operation of the switch is costly and hence the gathering of the observations are costly. Each player is equipped with finite memory and she can only use the latest observation to construct the control strategy. The private observations of the players lead to a source of asymmetry in this game. Moreover, the players have different costs for operating the switch, which is another source of asymmetry. We study the structural properties of the Nash equilibrium for this particular class of problems and then we finally show that the switching problem simplifies to a bi-objective optimization problem.**

I. INTRODUCTION

Game theory has been viewed as one of the major topics of interest for control theorist since last few decades; see- [1], [2], [3], [4], [5], [6], [7], [8], [9], [10], [11]. Many useful results of stochastic control, robust control have been derived by studying differential games; [8], [10]. Among the various class of game problems, linear-quadratic (LQ) differential games has been studied extensively since a closed form solution for these games can be computed [10]. A subclass of LQ differential games deals with partially observed system where the strategies of the players depend on the corresponding observations. The general two-player LQ game consists of a linear dynamics and linear observations as presented below:

$$\begin{aligned} dx =& Axdt + Budt + Cvdt + GdW_t. \\ x(0) =& x_0. \end{aligned} \tag{1}$$

The observation equations are given by (2).

$$\begin{aligned} y^1(t) &= H^1 x(t) + R^1 W_t^1, \\ y^2(t) &= H^2 x(t) + R^2 W_t^2, \end{aligned} \tag{2}$$

where W_t, $\{W_t^i\}_{i=1,2}$ are standard mutually independent Wiener processes of dimensions r and $\{r^i\}_{i=1,2}$ respectively. The above dynamics should be treated in the sense of Îto where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^{m_1}$, $v(t) \in \mathbb{R}^{m_2}$, $y^i(t) \in \mathbb{R}^{p_i}$. For the sake of simplicity, the model parameters A, B, C, G, H^i, and R^i are kept time invariant, however a similar analysis could be carried out with time varying parameters.

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

The authors are with the Department of Electrical and Computer Engineering and The Institute for Systems Research, University of Maryland, College Park, USA. Email: dmaity@umd.edu, raghava@umd.edu, baras@umd.edu

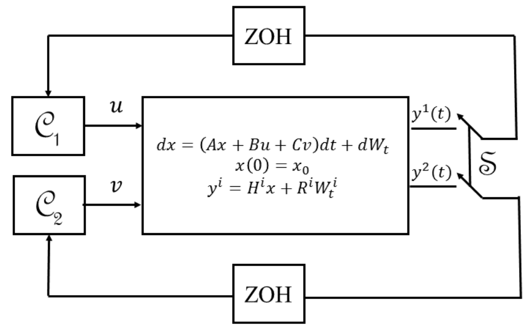

Fig. 1. The game dynamics is controlled by two independent controllers $\mathcal{C}_1$, $\mathcal{C}_2$ operated by the players. The players receive their respective observations y^i whenever the switch $\mathcal{S}$ is closed. The players unanimously decide the switching strategy for $\mathcal{S}$. ZOH is the zero-order hold circuit.

The game is equipped with a quadratic cost criterion for the players. In this paper, we consider the following cost criterion:

$$J(u,v) = \mathbb{E}\left[\int_0^T (\|x\|_L^2 + \|u\|_R^2 - \|v\|_S^2)dt\right] \tag{3}$$

where $\|p\|_Q^2 = p'Qp$ for any two matrices p and Q of proper dimensions. The game under consideration will be a zero-sum game. The cost function has to be minimized by player-1 (P1) who selects the action u, while player-2 (P2) aims to maximize the cost by selecting the action v. Therefore, for such a game there is notion of saddle-point that is represented by:

$$J(u^*, v) \leq J(u^*, v^*) \leq J(u, v^*) \tag{4}$$

The pair of strategies (u^*, v^*) is known as the Nash strategy for this game. The existance and uniqueness of such strategies have been presented in [3], [5], [6], and others.

In recent years of network-control and multi-agent systems, game theory has proven to be a very effective tool to study the interaction of the sub-systems to achieve global system optimality as presented in [12] and the references therein. However, often times in a large system the observations are not accessible continuously due to limitated communication, sensing or computing resources [13]. Some recent studies in control literature are oriented towards the variations of sampling based control which are known in many names such as event-based, event-triggered, self triggered, periodic control, Lebesgue sampling [14], [15], [16], [17] etc. These control strategies do not require the state information $x(t)$ (or observation $y(t)$) for all time

t, rather they sample $x(t)$ intermittently depending on the systems' performance criterion [18], [19]. Therefore, there is a need for a framework of game theoretic study where the continuous observations are not available and the Nash strategies will rely only on the intermittent available observations. This work aims to study LQ games with only discrete measurements that are accessed through a switch. The schematic of the game is presented in Figure 1. As shown in Figure 1, the players are equipped with a switch $\mathcal{S}$ which closes momentarily and opens immediately after closing. Therefore, the players can have finite number of measurements $\{y^i(\tau_k)\}_{k=1}^N$ over a finite interval of $[0,T]$. In this game we consider the switching has a finite cost i.e. to get a measurement $y^i(\tau_k)$, player-i has to pay $\lambda_i > 0$. Moreover, we assume that the players have a limited memory and hence they only remembers the most resent observation at any time t. This work is an extension of the work done in [20] where the game was studied with full observation.

The contribution of this work is to study the Nash strategies for the partially observed LQ games with finite number of intermittent observations where the players only remember the latest observation.

II. PROBLEM FORMULATION

Each player has only limited and asymmetric (non-shared observations y^i) information about the state of the game and they can ask for the current measurements by paying some cost. Let us denote the number of times the measurement information is requested by the players up to time t be $n(t)$. The cost associated with these information acquisitions are $\lambda_1 n(t)$ and $\lambda_2 n(t)$ for P1 and P2 respectively. We include these costs to the cost function J. Hence P1 should minimize:

$$J_1(u,v) = \mathbb{E}\big[\int_0^T (\|x\|_L^2 + \|u\|_R^2 - \|v\|_S^2)dt + \lambda_1 n(T)\big], \tag{5}$$

and P2 should maximize:

$$J_2(u,v) = \mathbb{E}\big[\int_0^T (\|x\|_L^2 + \|u\|_R^2 - \|v\|_S^2)dt - \lambda_2 n(T)\big]. \tag{6}$$

Though at this point we have two separate cost functions, one for each player, both optimization problems boil down to minimizing or maximizing (3) once the players finalize the number of times their respective measurements will be accessed. The objective of the players is to jointly determine the time instances $\{\tau_k\}$ to construct their respective information sets $\mathcal{I}^i(t)$. Let $\tau(t) \leq t$ be the latest time when the players received the observations. Then, we define $\mathcal{I}^1(t) = \{y^1(\tau(t))\}$ and $\mathcal{I}^2(t) = \{y^2(\tau(t))\}$ for all t.

P1 selects the control u, as an $\mathcal{I}^1(t)$ measurable function, to minimize $J_1(u,v)$. On the other hand, P2 constructs the strategy v, at time t, with the knowledge of $\mathcal{I}^2(t)$, to maximize $J_2(u,v)$ given in (6). The switching instances are ordered with probability 1 almost surely, i.e. $\tau_k < \tau_{k+1}$ and non-anticipative i.e. for any t, $\tau(t) = \tau_{n(t)} \leq t$.

There is another version of the same problem where instead of having cost for switching, there is a bound on total number of switching over the interval of $[0,T]$. The game under this scenario is presented as follows:

P1 should minimize:

$$J_1(u,v) = \mathbb{E}\big[\int_0^T (\|x\|_L^2 + \|u\|_R^2 - \|v\|_S^2)dt\big], \tag{7}$$
$$\text{subject to} \qquad n_1(T) < N$$

and P2 should maximize:

$$J_2(u,v) = \mathbb{E}\big[\int_0^T (\|x\|_L^2 + \|u\|_R^2 - \|v\|_S^2)dt\big], \tag{8}$$
$$\text{subject to} \qquad n_2(T) < N.$$

The objective functions (7) and (8) can be represented as (5) and (6) using suitable Lagrange multipliers (λ_1 and λ_2). Therefore, in this paper we will consider the problem posed in unconstrained form (5) and (6).

III. STRATEGIES WITH ASYMMETRIC INFORMATION

In this section, we study the Nash control strategies for the game under asymmetric information and asymmetric cost. Some structural properties of this game can be found in [10]. A recent work [20] has considered the similar game structure with symmetric (fully observed game) information. Preliminary results on these types of LQ games with costly full state information can be found in [20], [21]. The theorem below serves as a starting point of our analysis by providing an indication on the structure of the optimal control strategy of the problem in consideration.

Theorem 3.1 ([10], [20], [21]): The optimal strategies for the players with cost-less full state observations are:

$$u^*(t) = -R^{-1}B'Px \tag{9}$$
$$v^*(t) = S^{-1}C'Px. \tag{10}$$

The optimal cost is

$$J^* = \mathbb{E}[\|x(0)\|_{P(0)}^2 + \int_0^T \big(tr(PGG')dt],$$

where

$$\dot{P} + A'P + PA + L + P(CS^{-1}C' - BR^{-1}B')P = \mathbf{0} \tag{11}$$
$$P(T) = \mathbf{0}$$

The previous theorem gives explicit formulae for the optimal strategies of the players and those require the knowledge of the state $x(t)$ for all time $t \in [0,T]$. From this point onward, we will investigate how the strategies change for both the players when they have measurement information only at finite number of time instances. Before considering this problem, we attempt to solve the problem of finding the optimal strategies for both players for an arbitrary interval $(t_0, t_1] \subseteq [0,T]$.The strategies may depend on the measurement information at discrete time instances and the strategies should not ask for any future measurement. In calculating the strategies, we still do not consider the measurement query cost at this point. In the subsequent analyses, we will remove this assumption and comment on the original problem. Making this assumption for this section makes our problem tractable for this initial step.

By following standard completion of squares and using Itô calculus, it can be shown [20] that:

$$J = \mathbb{E}\big[\|x(0)\|^2_{P(0)} + \int_0^T \big(\|u + R^{-1}B'Px\|^2_R - \|v - S^{-1}C'Px\|^2_S + tr(PGG')\big)dt\big] \tag{12}$$

Equation (12) separates the cost that depends on the control strategy from the cost which remains invariant of the control strategy. Let us divide the cost J in two parts J_I, and J_d. J_I, being independent on the actions of the players and the other, J_d, depends on the choices of u and v. Thus,

$$J_I = \mathbb{E}\big[\|x(0)\|^2_{P(0)} + \int_0^T \big(tr(PGG')\big)dt\big],$$

$$J_d(u,v) = \mathbb{E}[\int_0^T \big(\|u + R^{-1}B'Px\|^2_R - \|v - S^{-1}C'Px\|^2_S\big)dt]. \tag{13}$$

Since the state (or observations y^i) is not available for all time t, the control strategies u and v will be different from what is presented in Theorem 3.1. The following proposition provides the structural properties of the strategies under partial discrete observations.

Proposition 3.2: The optimal u and v that optimizes J_d:

$$J_d(u,v) = \mathbb{E}[\int_0^T \big(\|u + R^{-1}B'Px\|^2_R - \|v - S^{-1}C'Px\|^2_S\big)dt]. \tag{14}$$

will be of the form

$$u^*(t) = -R^{-1}B'P(t)\hat{x}_1(t) \tag{15}$$

$$v^*(t) = S^{-1}C'P(t)\hat{x}_2(t) \tag{16}$$

for some $\hat{x}_1$ and $\hat{x}_2$ such that $u^*(t)$ is $\mathcal{I}^1(t)$ measurable and $v^*(t)$ is $\mathcal{I}^2(t)$ measurable.

Proof: For a fixed t, $R^{-1}B'P(t)$ is a linear mapping from $\mathbb{R}^n$ to $\mathbb{R}^{m_1}$. For all t, let us denote

$$u(t) = u_1(t) + u_2(t)$$

such that $u_1(t) \in \mathcal{R}ange(R^{-1}B'P(t))$ and $u_2(t) \in \mathcal{N}ull(P(t)BR^{-1})$ where $\mathcal{R}ange(\cdot)$ and $\mathcal{N}ull(\cdot)$ are the range space and null space of an operator respectively. Therefore,

$$J_d(u,v) = \mathbb{E}[\int_0^T \big(\|u_1 + R^{-1}B'Px\|^2_R + \|u_2\|^2_R - \|v - S^{-1}C'Px\|^2_S\big)dt] \tag{17}$$

since, $u_2 \in \mathcal{N}ull(P(t)BR^{-1})$ and $\mathcal{N}ull(\psi) = \mathcal{R}ange(\psi')^\perp$ for linear operator ψ and its adjoint ψ'. From (17), it is clear that the optimal choice is $u_2 \equiv 0$. Hence $u^*(t) = u_1(t) = R^{-1}B'P(t)\hat{x}_1(t)$ for some optimal $\hat{x}_1(t) \in \mathbb{R}^n$. Similarly it can also be proved that $v^*(t) = S^{-1}C'P(t)\hat{x}_2(t)$ for some optimal $\hat{x}_2(t) \in \mathbb{R}^n$. Since $u^*(t)$ has to be $\mathcal{I}^1(t)$ measurable, this implies $\hat{x}_1(t)$ needs to be $\mathcal{I}^1(t)$ measurable. Similar arguments show $\hat{x}_2(t)$ has to be $\mathcal{I}^2(t)$ measurable. ■

Let us denote $\tilde{R}(t) = P(t)BR^{-1}B'P(t)$ and $\tilde{S}(t) = P(t)CS^{-1}C'P(t)$, however the time argument will be suppressed frequently to maintain brevity. By Proposition 3.2, let the optimal controller for P1 be $u = -R^{-1}B'P\hat{x}_1$ and that for P2 be $v = S^{-1}C'P\hat{x}_2$, where $\hat{x}_i(t) = f(t, \mathcal{I}^i(t))$ is a $\mathcal{I}^i(t)$ measurable function. Thus the objective is to determine the optimal $\hat{x}_i(t)$ which will minimize the cost J_d.

Substituting u and v in J_d, we obtain:

$$J_d(\hat{x}_1, \hat{x}_2) = \mathbb{E}[\int_0^T (\|x - \hat{x}_1\|^2_{\tilde{R}} - \|x - \hat{x}_2\|^2_{\tilde{S}})dt]. \tag{18}$$

With these strategies, the state of the game evolves as:

$$dx = (Ax - P^{-1}\tilde{R}\hat{x}_1 + P^{-1}\tilde{S}\hat{x}_2)dt + GdW_t. \tag{19}$$

From equation (19), we can write the solution to be:

$$x(t) = \Phi_A(t,t_0)x(t_0) + K_1^{t,t_0}[\hat{x}_1](t) + K_2^{t,t_0}[\hat{x}_2](t) + K_3^{t,t_0}[GW](t) \tag{20}$$

for some $t_0 \le t$, where Φ_A is the state transition matrix for the drift matrix A. K_1^{t,t_0}, K_2^{t,t_0} and K_3^{t,t_0} are linear operators defined as follows:

$$K_1^{t,t_0}[f](t) = -\int_{t_0}^t \Phi_A(t,s)P^{-1}\tilde{R}f(s)ds \tag{21}$$

$$K_2^{t,t_0}[f](t) = \int_{t_0}^t \Phi_A(t,s)P^{-1}\tilde{S}f(s)ds \tag{22}$$

$$K_3^{t,t_0}[fW](t) = \int_{t_0}^t \Phi_A(t,s)f(s)dW(s). \tag{23}$$

Let us define a filtered variable $\bar{x}$ as given below:

$$\bar{x}(t) = \Phi_A(t,t_0)x(t_0) + K_1^{t,t_0}[\hat{x}_1](t) + K_2^{t,t_0}[\hat{x}_2](t). \tag{24}$$

Therefore, we can represent $x(t)$ as:

$$x(t) = \bar{x}(t) + K_3^{t,t_0}[GW](t) \tag{25}$$

Thus, for $(t_0, t_1] \subseteq [0, T]$, J_d is found to be:

$$J_d^{t_0,t_1}(u,v) = \mathbb{E}[\int_{t_0}^{t_1} (\|\bar{x} - \hat{x}_1\|^2_{\tilde{R}} - \|\bar{x} - \hat{x}_2\|^2_{\tilde{S}})dt] + \mathbb{E}[\int_{t_0}^{t_1} \|K_3^{t,t_0}[GW](t)\|^2_{\tilde{R}-\tilde{S}} dt] \tag{26}$$

From the properties of Wiener process it can be shown that,

$$\mathbb{E}[\|K_3^{t,t_0}GW\|^2_{\tilde{R}-\tilde{S}}] = \int_{t_0}^t tr(\|\Phi_A(t,s)G\|^2_{\tilde{R}(s)-\tilde{S}(s)})ds$$

Thus,

$$J_d^{t_0,t_1}(u,v) = \mathbb{E}[\int_{t_0}^{t_1} (\|\bar{x} - \hat{x}_1\|^2_{\tilde{R}} - \|\bar{x} - \hat{x}_2\|^2_{\tilde{S}})dt + \mathbb{E}[\int_{t_0}^{t_1}\int_{t_0}^{t} tr(\|\Phi_A(t,s)G\|^2_{\tilde{R}(s)-\tilde{S}(s)})dsdt]$$

Therefore, for any arbitrary interval $(t_0, t_1) \subseteq [0, T]$, it suffices to optimize the first term in $J_d^{t_0,t_1}$. Let us define:

$$J =$$
$$\mathbb{E}[\int_{t_0}^{t_1} (\|\Phi_A(t,t_0)x(t_0) + K_1^{t,t_0}[\hat{x}_1](t) + K_2^{t,t_0}[\hat{x}_2](t) - \hat{x}_1\|^2_{\tilde{R}}$$
$$- \|\Phi_A(t,t_0)x(t_0) + K_1^{t,t_0}[\hat{x}_1](t) + K_2^{t,t_0}[\hat{x}_2](t) - \hat{x}_2\|^2_{\tilde{S}})dt]$$

Here, we are looking for u and v (equivalently $\hat{x}_1$, $\hat{x}_2$) that is a saddle point for J. We will find the optimal $\hat{x}_i$ (saddle points) by considering the first order and second order Gâteaux differentials. The Gâteaux differentials of a functional $\mathcal{J}$ is given by:

$$\delta\mathcal{J}[f_1, f_2](h_1, h_2) = \lim_{a\to 0} \frac{\mathcal{J}(f_1 + ah_1, f_2 + ah_2) - \mathcal{J}(f_1, f_2)}{a}$$

where the notation $\mathcal{J}[f_1, f_2](h_1, h_2)$ means the Gateaux differential of $\mathcal{J}$ evaluated at the point (f_1, f_2) in the direction (h_1, h_2). Note that $\mathcal{J}[f_1, f_2](\cdot, \cdot)$ is a linear functional parameterized by f_1 and f_2. It can be shown that:

$$\frac{1}{2}\delta J[\hat{x}_1, \hat{x}_2](h_1, h_2) = \mathbb{E}\Big[\int_{t_0}^{t_1} \big(\langle \bar{x} - \hat{x}_1, K_1^{t,t_0}[h_1](t) + K_2^{t,t_0}[h_2](t) - h_1)\rangle_{\tilde{R}} - \langle \bar{x} - \hat{x}_1, K_1^{t,t_0}[h_1](t) + K_2^{t,t_0}[h_2](t) - h_2)\rangle_{\tilde{S}}\big)dt\Big] \tag{27}$$

where $\langle a, b\rangle_D = a'Db$ denotes the inner product w.r.t. D. In the following, we seek to find the extremal points $(\hat{x}_1, \hat{x}_2)$ where $\delta J[\hat{x}_1, \hat{x}_2](\cdot, \cdot)$ is a zero functional.

Lemma 3.3:

$$\delta J[\hat{x}_1, \hat{x}_2](h_1, h_2) = 0 \; \forall h_1, h_2$$
$$\Leftrightarrow \mathbb{E}[(\bar{x} - \hat{x}_i)|\mathcal{I}^j] = 0, \text{ for } i = 1, 2;\ j = 1, 2.$$

Proof: $\Leftarrow$ direction: $\mathbb{E}[\bar{x} - \hat{x}_i|\mathcal{I}^j] = 0 \Rightarrow \mathbb{E}[\bar{x} - \hat{x}_i] = 0 \Rightarrow \delta J[\hat{x}_1, \hat{x}_2](h_1, h_2) = 0 \; \forall h_1, h_2$.

$\Rightarrow$ direction: It is straightforward to show that $\mathbb{E}[\bar{x} - \hat{x}_i|\mathcal{I}^i] = 0$ due to the minimum mean square estimate (MMSE) argument. Suppose $\mathbb{E}[\bar{x} - \hat{x}_1|\mathcal{I}^2] \neq 0$. Let us choose $h_1 = 0$.

$$\frac{1}{2}\delta J[\hat{x}_1, \hat{x}_2](h_1, h_2) = \mathbb{E}\Big[\int_{t_0}^{t_1} \big(\langle \mathbb{E}[(\bar{x} - \hat{x}_1)|\mathcal{I}_t^2],\ K_2^{t,t_0}[h_2](t)\rangle_{\tilde{R}}\big)dt\Big] \tag{28}$$

If $\mathbb{E}[\bar{x} - \hat{x}_1|\mathcal{I}^2] \neq 0$, $\exists$ h_2, $\mathcal{I}_t^2$ measurable such that $\delta J[\hat{x}_1, \hat{x}_2](0, h_2) > 0$. Hence contradiction. ■

Lemma 3.4: $\delta J[\hat{x}_1, \hat{x}_2](h_1, h_2) = 0 \; \forall \; h_1$, h_2 only if H_1, H_2 have rank n.

Proof: We restrict ourselves to the space of linear estimates. Suppose at time t_0, a measurement is received.

$$\hat{x}_1(t_0) = \mathbb{E}[x(t_0)|y^1(t_0)] \coloneqq Q^1 y^1(t_0).$$

From Lemma 3.3,

$$\begin{aligned}\hat{x}_2(t_0) &= \mathbb{E}[\hat{x}_1(t_0)|y^2(t_0)]\\ &= \mathbb{E}[Q^1 y^1(t_0)|y^2(t_0)]\\ &= Q^1\mathbb{E}[H^1 x(t_0) + R^1 W^1(t_0)|y^2(t_0)]\\ &= Q^1 H^1 \hat{x}_2(t_0)\end{aligned}$$

Therefore, we must have

$$Q^1 H^1 = I \tag{29}$$

Similarly $Q^2 H^2 = I$. Since $H^i \in \mathbb{R}^{p_i\times n}$ and $Q^i H^i = I$, it is required that H^i has rank n. Therefore, $Q^i = ((H^i)^T H^i)^{-1}(H^i)^T$. ■

Therefore, in order to make $\delta J[\hat{x}_1, \hat{x}_2](h_1, h_2) \equiv 0$, we need $\mathbb{E}[\bar{x}(t) - \hat{x}_i(t)|\mathcal{I}^j] = 0$ for all $t \in (t_0, t_1]$ and for all $i, j = 1, 2$. The estimation $\hat{x}_i(t_0) = Q^i y^i(t_0)$ makes $\mathbb{E}[\bar{x}(t) - \hat{x}_i(t)|\mathcal{I}^j] = 0$ at time t_0.

Thus, differentiating $\mathbb{E}[\bar{x}(t) - \hat{x}_i(t)|\mathcal{I}^j]$ w.r.t. t and equating to zero will be necessary and sufficient to ensure $\mathbb{E}[\bar{x}(t) - \hat{x}_i(t)|\mathcal{I}^j] = 0$ for all $t \in (t_0, t_1]$. Therefore, we must have

$$\begin{aligned}\dot{\hat{x}}_1 =& A\hat{x}_1 - P^{-1}\tilde{R}\hat{x}_1 + P^{-1}\tilde{S}\mathbb{E}[\hat{x}_2|\mathcal{I}^1]\\ =& (A - P^{-1}\tilde{R} + P^{-1}\tilde{S})\hat{x}_1, \qquad (30)\\ \dot{\hat{x}}_2 =& (A - P^{-1}\tilde{R} + P^{-1}\tilde{S})\hat{x}_2 \qquad (31)\end{aligned}$$

with $\hat{x}_i(t_0) = Q^i y^i(t_0)$.

Now, let us consider the second Gateaux differential $\delta^2 J(\hat{x}_1, \hat{x}_2)(h_1, h_2)$. It can be shown that

$$\frac{1}{2}\delta J^2[\hat{x}_1, \hat{x}_2](h_1, h_2) = \mathbb{E}[\int_{t_0}^{t_1} \big(\|K_1^{t,t_0}[h_1](t) + K_2^{t,t_0}[h_2](t) - h_1\|^2_{\tilde{R}} - \|K_1^{t,t_0}[h_1](t) + K_2^{t,t_0}[h_2](t) - h_2\|^2_{\tilde{S}}\big)dt].$$

Let $h_1(t) = a \in \mathbb{R}^n \; \forall \; t \in (t_0, t_1)$. Let $\tilde{A_S} = A + P^{-1}S$ and also let,

$$h_2(t) = -\int_{t_0}^{t} \Phi_{\tilde{A}_S}(t,s)P^{-1}\tilde{R}\; a\; ds.$$

This implies,

$$\begin{aligned}&K_2^{t,t_0}[h_2](t)\\ &= -\int_{t_0}^{t} \Phi_A(t,s)P^{-1}\tilde{S}\int_{t_0}^{s} \Phi_{\tilde{A}_S}(s,\sigma)P^{-1}\tilde{R}\; a\; d\sigma ds\\ &= -\int_{t_0}^{t} \Big[\int_{\sigma}^{t} \Phi_A(t,s)P^{-1}\tilde{S}\Phi_{\tilde{A}_S}(s,\sigma)ds\Big] P^{-1}\tilde{R}\; a\; d\sigma\\ &= -\int_{t_0}^{t} \Big[\int_{\sigma}^{t} \frac{d}{ds}\Phi_A(t,s)\Phi_{\tilde{A}_S}(s,\sigma)ds\Big] P^{-1}\tilde{R}\; a\; d\sigma\\ &= h_2 - \int_{t_0}^{t} \Phi_A(t,s)(-P^{-1}\tilde{R}\; a\;)ds\\ &= h_2 - K_1^{t,t_0}[h_1](t).\end{aligned}$$

$$\Rightarrow K_1^{t,t_0}[h_1](t) + K_2^{t,t_0}[h_2](t) - h_2 = 0,$$
$$\text{and } K_1^{t,t_0}[h_1](t) + K_2^{t,t_0}[h_2](t) - h_1 = h_2 - h_1.$$

Thus,

$$\delta^2 J(\hat{x}_1, \hat{x}_2)(h_1, h_2) = \mathbb{E}[\int_{t_0}^{t_1} \|h_2 - h_1\|^2_{\tilde{R}} dt]$$

$$h_2 - h_1 = \int_{t_0}^{t} \Phi_{\tilde{A}_S}(t,\sigma)P^{-1}\tilde{R}\; a\; d\sigma ds - a \neq 0$$

for some $a \neq 0$. Thus $\delta^2 J(\hat{x}_1, \hat{x}_2)(h_1, h_2) > 0$. Similarly by choosing, $h_2 = b$ and $h_1(t) = \int_{t_0}^{t} \Phi_{\tilde{A}_R}(t,s) P^{-1}\tilde{S}\ b\ ds$, where $\tilde{A}_R = A - P^{-1}R$, we can show $\delta^2 J(\hat{x}_1, \hat{x}_2)(h_1, h_2) < 0$. Hence $\hat{x}_1, \hat{x}_2$ achieve the Nash equilibrium and the optimal strategies are given by :

$$u^*(t) = -R^{-1}B'P(t)\hat{x}_1(t),$$
$$v^*(t) = S^{-1}C'P(t)\hat{x}_2(t).$$

where $\hat{x}_i(t)$ satisfies the differential equations (30), (31).

IV. Optimal switching strategy

With this optimal strategy, the cost incurred is:

$$\mathbb{E}\left[\int_{t_0}^{t_1} (\|x - \hat{x}_1\|^2_{\tilde{R}} - \|x - \hat{x}_2\|^2_{\tilde{S}})\ dt\right] = \mathbb{E}\left[\int_{t_0}^{t_1} \|K_3^{t,t_0}[GW](t)\|^2_{\tilde{R}-\tilde{S}} dt\right] + \Gamma$$

where Γ (depends on R^i) is the cost incurred in estimating $\bar{x}$ by $\hat{x}_i$ and could found explicitly. In the following analyses, we neglect Γ since it would have a structure similar to the other term. Therefore, an identical analysis can be carried out by explicitly considering Γ. In the previous section, we have been able to find the Nash strategy for the players in an arbitrary interval $(t_0, t_1]$. Thus, if the players come to an agreement about the measurement information acquisition times $\{\tau_i\}_{i=1}^{n_T}$, they can find their strategies for the intervals $(\tau_i, \tau_{i+1}]$ where $\tau_0 = t_0$ and $\tau_{n_T+1} = T$. Since the strategies on different intervals are independent of each other, the strategy for the entire time horizon $[0, T]$ can be constructed by concatenating the individual strategies over the intervals $(\tau_i, \tau_{i+1}]$. Now the question that remains to be answered is how the players come to an agreement about the measurement information acquisition times. Let us analyse this from the point of view of both the players.

For P1:

$$\min_{n_1(T), \tau_1^1, \cdots, \tau^1_{n_1(T)}} \sum_{i=0}^{n_1(T)} \mathbb{E}\Big[\int_{\tau_i^1}^{\tau_{i+1}^1} \|K_3^{t,\tau_i^1}[GW](t)\|^2_{\tilde{R}-\tilde{S}} dt + \lambda_1 n_1(T)\Big]. \tag{32}$$

Whereas, P2 will seek to optimize the following function:

For P2:

$$\max_{n_2(T), \tau_1^2, \cdots, \tau^2_{n_2(T)}} \sum_{i=0}^{n_2(T)} \mathbb{E}\Big[\int_{\tau_i^2}^{\tau_{i+1}^2} \|K_3^{t,\tau_i^2}[GW](t)\|^2_{\tilde{R}-\tilde{S}} dt - \lambda_2 n_2(T)\Big] \tag{33}$$

where $\tau_0^1 = \tau_0^2 = 0$ and $\tau^2_{n_2(T)+1} = \tau^1_{n_1(T)+1} = T$.

Under the assumption that $\tilde{R} - \tilde{S} \succeq 0$ in order to ensure that the solution of the Riccati equation (11) is well defined, one can easily find out that for P2, the optimal choice would be to never access the state information. However, this is not the case for P1 and the choice of P1 depends on the value of λ_1 and the game parameters, namely A and G.

Remark 4.1: *If the game parameters for both the players satisfy the condition $CS^{-1}C' = BR^{-1}B'$, both optimization problems for the players have the same solution, and that solution does not ask for any observation except at initial time 0. Similar results were obtained in [20].*

Let us define,

$$\mathcal{L}_1(\tau_1^1, \cdots, \tau^1_{n_1(T)}) = \sum_{i=0}^{n_1(T)} \mathbb{E}\Big[\lambda_1 n_1(T) + \int_{\tau_i^1}^{\tau_{i+1}^1} \|K_3^{t,\tau_i^1}[GW](t)\|^2_{\tilde{R}-\tilde{S}} dt\Big], \tag{34}$$

and

$$\mathcal{L}_2(\tau_1^2, \cdots, \tau^2_{n_2(T)}) = \sum_{i=0}^{n_2(T)} \mathbb{E}\Big[\lambda_2 n_2(T) - \int_{\tau_i^2}^{\tau_{i+1}^2} \|K_3^{t,\tau_i^2}[GW](t)\|^2_{\tilde{R}-\tilde{S}} dt\Big]. \tag{35}$$

Let us denote the set of feasible switching strategies by:

$$\mathcal{S} = \{\{\tau_k\}_{k=0}^N \mid \quad 0 = \tau_0 < \tau_1 \cdots < \tau_N < T, \text{ and } N \in \mathbb{N}\} \tag{36}$$

where $\mathbb{N}$ is the set of natural numbers.

At this point, the problem is a multi-objective optimization problem where we have two objectives $\mathcal{L}_1$ and $\mathcal{L}_2$ that need to be minimized. This could be formulated as:

$$\min_{s \in \mathcal{S}} \ \{\mathcal{L}_1(s), \quad \mathcal{L}_2(s)\} \tag{37}$$

Let us denote the optimal value of $\mathcal{L}_i$ by c_i^*. Therefore, the optimization problem (37) can equivalently written as:

$$\min_{s \in \mathcal{S}} \{(\mathcal{L}_1 - c_1^*)^2, \quad (\mathcal{L}_2 - c_2^*)^2\}, \tag{38}$$

A multi-objective optimization problem has the notion of Pareto optimally and therefore we seek for Pareto optimal solution(s) of this problem.

Definition 4.2: A feasible point $s \in \mathcal{S}$ is said to (Pareto) dominate another feasible point $s_1 \in \mathcal{S}$ if

1. $\mathcal{L}_i(s) \leq \mathcal{L}_i(s_1)$ for all $i = 1, 2$, and
2. $\mathcal{L}_j(s) < \mathcal{L}_j(s_1)$ for some $j \in \{1, 2\}$.

A solution $s^* \in \mathcal{S}$ is called Pareto optimal point if there does not exist another solution $s \in \mathcal{S}$ that dominates it.

Let us take $\mu \in [0, 1]$ and define the weighted cost function

$$\mathcal{L}^\mu = \mu(\mathcal{L}_1 - c_1^*)^2 + (1 - \mu)(\mathcal{L}_2 - c_2^*)^2 \tag{39}$$

Let us define a weighted cost problem:

$$\min_{s \in \mathcal{S}} \ \mathcal{L}^\mu(s). \tag{40}$$

The Pareto optimal solution of (38) has one-to-one correspondence with the optimal solution of the weighted cost problem (40) for some μ, see [22]. Let us denote

$$\sum_{i=0}^{N} \int_{\tau_i}^{\tau_{i+1}} \mathbb{E}\|K_3^{t,\tau_i}[GW]\|^2_{\tilde{R}-\tilde{S}} dt = C(N, \tau_1, \cdots, \tau_N)$$
$$\Rightarrow \mathcal{L}^\mu(N, \tau_1, \cdots, \tau_N) = \mu(C + \lambda_1 N - c_1^*)^2 + (1 - \mu)(C - \lambda_2 N - c_2^*)^2$$

For a fixed N and μ, to select optimal τ_i we seek: $\frac{\partial \mathcal{L}^\mu}{\partial \tau_i} = 0$. Therefore the necessary conditions are:

$$(C + (\mu\lambda_1 - (1-\mu)\lambda_2)N - \mu c_1^* - (1-\mu)c_2^*)\frac{\partial C}{\partial \tau_i} = 0 \quad (41)$$

For all $i = 1, \cdots, N$ $\frac{\partial C}{\partial \tau_i} = 0$ implies

$$\int_{\tau_{i-1}}^{\tau_i} tr(\|\Phi_A(\tau_i, t)G\|^2_{\tilde{R}(t)-\tilde{S}(t)})dt = \int_{\tau_i}^{\tau_{i+1}} tr(\|\Phi_A(t, \tau_i)G\|^2_{\tilde{R}(\tau_i)-\tilde{S}(\tau_i)})dt, \quad (42)$$

which needs to be satisfied or $(C + (\mu\lambda_1 - (1-\mu)\lambda_2)N - \mu c_1^* - (1-\mu)c_2^*) = 0$ admits a solution.

Claim 4.3: *For a fixed N, $(C + (\mu\lambda_1 - (1-\mu)\lambda_2)N - \mu c_1^* - (1-\mu)c_2^*) = 0$ has a solution.*
The proof of the above claim follows directly from the fact that $C(N, \tau_1, \cdots, \tau_N)$ is a continuous function of τ_i and the maximum and minimum values of C are $c_2^* + \lambda_2 N$ and $c_1^* - \lambda_1 N$ respectively. Therefore, there is a point in $\mathcal{S}$ where the function attains a value equal to the convex combination (with parameter μ) of its maximum and minimum values.

It is straightforward to show that $\{\tau_i\}$ satisfying $C(N, \tau_1, \cdots, \tau_N) + (\mu\lambda_1 - (1-\mu)\lambda_2)N - \mu c_1^* - (1-\mu)c_2^* = 0$ is optimal. After this point, $\mathcal{L}^\mu$ will be a function of an integer variable N and can be solved.

Under this choice of sampling instances, the costs incurred by P1 and P2 are respectively $[(\mu c_1^* + (1-\mu)c_2^*) + (1-\mu)(\lambda_1 + \lambda_2)N]$ and $\frac{1}{2}[(\mu c_1^* + (1-\mu)c_2^*) - \mu(\lambda_1 + \lambda_2)N]$.

The above analysis shows how we can find the solutions for the weighted cost problem (40) for some μ and hence we can comment on the actual problem (38). Note that one could equivalently formulate a weighted cost $\tilde{\mathcal{L}}^\mu = \mu\mathcal{L}_1 + (1-\mu)\mathcal{L}_2$.

In the following we cite an interesting remark for this game and the remark is equivalent to one of the the remarks presented in [20]

Remark 4.4: *The noise matrix G plays a role in determining the instances when the observations are acquired, but the optimal strategies for the players do not rely on G. For a deterministic game ($G \equiv 0$) the results show that the player-i does not need any more information other than $y^i(0)$. In fact, the deterministic framework results in the same control and switching strategy as the symmetric (i.e. $B = C$, $S = R$) game framework.*

V. Conclusions

In this work, we have considered a two players partially observed stochastic differential LQ game where the players have intermittent but synchronised observations. The observations require finite costs. With costly information, we have derived the Nash strategies for the actions u and v, and also we have shown how to find the Pareto optimal instances for acquiring the observations for the players.

This framework can be used with event-based framework for a large systems. By design, there is a finite interval between two successive optimal observation acquisition times. Thus, this framework excludes the possibility of Zeno behavior.

References

[1] A. W. Starr and Y.-C. Ho, "Nonzero-sum differential games," *Journal of Optimization Theory and Applications*, vol. 3, no. 3, pp. 184–206, 1969.

[2] ——, "Further properties of nonzero-sum differential games," *Journal of Optimization Theory and Applications*, vol. 3, no. 4, pp. 207–219, 1969.

[3] M. Foley and W. Schmitendorf, "On a class of nonzero-sum, linear-quadratic differential games," *Journal of Optimization Theory and Applications*, vol. 7, no. 5, pp. 357–377, 1971.

[4] R. Isaacs, "Differential games iii," 1954.

[5] A. Weeren, J. Schumacher, and J. Engwerda, "Asymptotic analysis of linear feedback Nash equilibria in nonzero-sum linear-quadratic differential games," *Journal of Optimization Theory and Applications*, vol. 101, no. 3, pp. 693–722, 1999.

[6] T. Basar, "On the uniqueness of the Nash solution in linear-quadratic differential games," *International Journal of Game Theory*, vol. 5, no. 2-3, pp. 65–90, 1976.

[7] T. Basar and G. J. Olsder, *Dynamic noncooperative game theory*. SIAM, 1995, vol. 200.

[8] M. R. James and J. Baras, "Partially observed differential games, infinite-dimensional hamilton-jacobi-isaacs equations, and nonlinear h-infinity control," *SIAM Journal on Control and Optimization*, vol. 34, no. 4, pp. 1342–1364, 1996.

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

[10] J. Engwerda, *LQ dynamic optimization and differential games*. John Wiley & Sons, 2005.

[11] W. H. Fleming and D. Hernández-Hernández, "On the value of stochastic differential games," *Commun. Stoch. Anal*, vol. 5, no. 2, pp. 341–351, 2011.

[12] M. Wooldridge, *An introduction to multiagent systems*. John Wiley & Sons, 2009.

[13] L. Zhang, H. Gao, and O. Kaynak, "Network-induced constraints in networked control systems–a survey," *IEEE transactions on industrial informatics*, vol. 9, no. 1, pp. 403–416, 2013.

[14] W. Heemels, K. H. Johansson, and P. Tabuada, "An introduction to event-triggered and self-triggered control," in *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*. IEEE, 2012, pp. 3270–3285.

[15] K. J. Åström and B. Bernhardsson, "Comparison of Riemann and Lebesque sampling for first order stochastic systems," in *Proceedings of the 41st IEEE Conference on Decision and Control, 2002*, vol. 2. IEEE, 2002, pp. 2011–2016.

[16] M. Lemmon, T. Chantem, X. S. Hu, and M. Zyskowski, "On self-triggered full-information H-infinity controllers," in *Hybrid Systems: computation and control*. Springer, 2007, pp. 371–384.

[17] W. Bian and M. French, "General fast sampling theorems for nonlinear systems," *Systems & control letters*, vol. 54, no. 11, pp. 1037–1050, 2005.

[18] D. Maity and J. S. Baras, "Event based control of stochastic linear systems," in *Event-based Control, Communication, and Signal Processing (EBCCSP), 2015 International Conference on*. IEEE, 2015, pp. 1–8.

[19] ——, "Event based control for control affine nonlinear systems: A lyapunov function based approach," in *Decision and Control (CDC), 2015 IEEE 54th Annual Conference on*. IEEE, 2015, pp. 3767–3772.

[20] ——, "Optimal strategies for stochastic linear quadratic differential games with costly information," in *Decision and Control (CDC), 2016 IEEE 55th Conference on*. IEEE, 2016, pp. 276–282.

[21] ——, "Strategies for two-player differential games with costly information," in *2016 13th International Workshop on Discrete Event Systems (WODES)*. IEEE, 2016, pp. 211–216.

[22] C.-L. Hwang and A. S. M. Masud, *Multiple Objective Decision Making–Methods and Applications: a state-of-the-art survey*. Springer Science & Business Media, 2012, vol. 164.