

Asymptotic Policies for Stochastic Differential Linear Quadratic Games with Intermittent State Feedback

Dipankar Maity and John S. Baras

Abstract—In this paper, we consider infinite horizon linear quadratic stochastic differential games where the games are neither open-loop nor closed-loop. The state of the game dynamics is measured only when a certain switch is closed. The switch requires unanimous operation by the players, and continuum state measurements are not possible. There is an upper bound on the number of times the switch can be closed. Each player is given a quadratic cost function and the objective of each player is to design a switching strategy and a control strategy in order to optimize their respective cost function. We investigate the Nash control strategy and optimal switching policy for this game with two different cost structures: discounted cost and average-time cost.

I. INTRODUCTION

Differential games have been studied for a long time due to their wide applicability in robust control, minimax stochastic control and multi-agent systems [1], [2]. Among different forms of games, two-player linear quadratic (LQ) games have received ample attention for the last few decades. To mention a few, [3], [4], [5], [6] and the references therein show the immensity of the studies performed.

The stochastic dynamics of a general two-player LQ game can be expressed as

$$dx = (Ax + B_1u_1 + B_2u_2)dt + GdW_t \quad (1)$$

where $x \in \mathbb{R}^n$, $u_1 \in \mathbb{R}^{m_1}$, $u_2 \in \mathbb{R}^{m_2}$ and W_t is an m dimensional Wiener process noise. The associated quadratic cost for player- i is given by:

$$J_i(u_1, u_2) = \mathbb{E} \left[\int_0^T (x' L_i x + u_i' R_i u_i) dt \right]; \quad i = 1, 2 \quad (2)$$

where $L_i, R_i \succ \mathbf{0}$. More details on LQ game problems can be found in [2], [6].

Linear quadratic games were studied under the scenario of either open-loop or closed-loop. By open-loop we mean that the strategies depend only on the initial state of the system. On the other hand, in a closed-loop game, the strategies depend on the current value of the state $x(t)$.

However, in present days, it is not always possible to access the state information for a large system due to communication, sensing or energy constraints. In the control literature, there is a recent trend towards event-based [7],

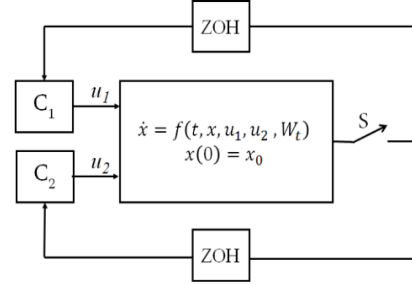


Fig. 1. The game dynamics is controlled by two independent controllers C_1 , C_2 operated by the players. The players receive perfect state measurements whenever the switch S is closed. The players unanimously decide the switching strategy for S .

event-triggered [8], self-triggered [9] or periodic control [10]. These control strategies do not require the state information $x(t)$ for all time t , rather they sample $x(t)$ intermittently depending on the systems' performance criterion [8], [11]. Event based techniques have proven to be efficient for large scale inter-connected systems to reduce communication and sensing operations [12]. At the same time, many properties of a large scale multi-agent system can be studied if formulated in a game theoretic framework as discussed in [13] and the references therein. Therefore, the game formulations of such multi-agent systems need to consider the strategies depending only on intermittent state information.

In this work, we consider a game formulation for a two-agents (henceforward called as two players) system where the system is equipped with a switch that enables the players to access the state information. Therefore, the players can have multiple state measurements (finite number of discrete measurements) over a horizon of $[0, T]$; i.e. they have the knowledge of $\{x(\tau_i)\}_{i=1}^N$ such that all $\tau_i \in [0, T]$. The players have more information than an open-loop setup where the only information is $x(0)$ and less information than a closed-loop game where players know $x(t)$ for all t . The schematic diagram, Fig 1, illustrates the game structure along with the switching mechanism to obtain the state information. In [14], [15], [16] a similar game problem has been studied for a finite duration $[0, T]$. This paper studies the game for an infinite horizon. The matrices A, B_i, L_i, R_i, G are time invariant for this infinite horizon game.

II. NOTATION

C_i : controller of player- i , S : switching strategy, $J_i^1(u_1, u_2, S)$: discounted infinite horizon cost function for player- i , $J_i^2(u_1, u_2, S)$: average infinite horizon cost function

Research partially supported by US Air Force Office of Scientific Research MURI grant FA9550-09-1-0538, by DARPA through ARO grant W911NF1410384, by National Science Foundation (NSF) grant CNS-1035655, and by 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, baras@umd.edu

for player- i , $J_i^j(t_0, t_1, \cdot)$: cost function $J_i^j(u_1, u_2, \mathcal{S})$ for a finite interval $(t_0, t_1]$, u_i^* : Nash controller strategy for player- i , $J_i^{j*}(\mathcal{S}) = J_i^j(u_1^*, u_2^*, \mathcal{S})$, $J_i^{j*}(t_0, t_1, \cdot)$: optimal cost function $J_i^{j*}(u_1^*, u_2^*, \mathcal{S})$ for an interval $(t_0, t_1]$, \mathbb{U}^{ad} : set of admissible control strategies, \mathbb{S} : set of admissible switching strategies, $\|a\|_b^2 \triangleq a'ba$ for any matrices a and b of compatible dimensions.

III. PROBLEM FORMULATION

As mentioned earlier, under the infinite time horizon, we will consider two versions of the problem. Firstly, we will consider discounted LQ-games with discount factor $\beta > 0$; and then we will consider the game where the objective is the time-average cost. In other words, we will consider the following two situations when :

$$J_i^1(u_1, u_2, \mathcal{S}) = \mathbb{E} \left[\int_0^\infty e^{-\beta t} (x' L_i x + u_i' R_i u_i) dt \right] \quad (3)$$

and

$$J_i^2(u_1, u_2, \mathcal{S}) = \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\int_0^T (x' L_i x + u_i' R_i u_i) dt \right] \quad (4)$$

for $i = 1, 2$.

The expectation in the above two costs are with respect to the probability measure \mathbb{P} induced by the controller strategies \mathcal{C}^i and switching strategy \mathcal{S} . In order to maintain notational brevity, we will not explicitly write the dependence of \mathbb{P} on \mathcal{C}^i and \mathcal{S} .

We assume the following features about the switch \mathcal{S} as it is mentioned in [14].

Assumption 3.1: A1: Switching provides noise-free instantaneous state information whenever, at t , the switch is closed.

A2: Switching is instantaneous i.e. if the switch \mathcal{S} is closed at τ then it is open for all $t > \tau$ until the next switching time.

A3: The switch is closed at $t = 0$, and x_0 is available to the players.

The switching element has constraints on the total number of times the switch can be closed. For these two different problems, we consider two different switching constraints as well. For the first problem, we consider an upper bound on the total number of switchings i.e. the number of state measurements available is finite even though the game continues for infinite horizon. On the other hand, for the second problem, the number of switchings over an arbitrary interval $(a, b]$ is upper bounded by the number $\lceil d(b - a) \rceil$ where $d > 0$ is given. By N_T^i , we will denote the number of switch closings (i.e. number of times the switch \mathcal{S} is closed) requested by player- i , for an arbitrary interval $T = (a, b]$. Let N_∞^i denote the number of switchings for the interval $(0, \infty)$ for the i -th player. Let us denote the set of all semi-open (left open, right closed) intervals on \mathbb{R}_+ by \mathcal{O} .

At this point, we formally state the two optimization problems. The discounted cost optimization problem can be expressed as:

Objective for Player- i

$$\min_{\mathcal{C}_i, \mathcal{S}} J_i^1(u_1, u_2, \mathcal{S}) \quad (5)$$

subject to $N_\infty^i \leq c$

The timed-average cost problem is written as:

Objective for Player- i

$$\min_{\mathcal{C}_i, \mathcal{S}} J_i^2(u_1, u_2, \mathcal{S}) \quad (6)$$

subject to $N_T^i \leq \lceil \mu(T)d \rceil \quad \forall T \in \mathcal{O}$

where $\mu(\cdot)$ is the Lebesgue measure on \mathbb{R} . In both the problem formulations (5) and (6), the dynamic equation of the game (1) is implicitly a constraint and we suppress it to maintain brevity.

IV. GAMES WITH DISCOUNTED COST

For a given discount factor, $\beta > 0$, the game problem has the following cost function for player- i :

$$J_i^1(u_1, u_2, \mathcal{S}) = \mathbb{E} \left[\int_0^\infty e^{-\beta t} (x' L_i x + u_i' R_i u_i) dt \right] \quad (7)$$

We aim to find the Nash controllers $(\mathcal{C}_1^*, \mathcal{C}_2^*)$ and the optimal switching strategy to minimize the cost function (7).

Let $J_i^1(t_0, t_1, \cdot)$ denote the following finite duration cost function:

$$\begin{aligned} J_i^1(t_0, t_1, \cdot) &= J_i^1(t_0, t_1, u_1, u_2, \mathcal{S}) \\ &= \mathbb{E} \left[\int_{t_0}^{t_1} e^{-\beta t} (x' L_i x + u_i' R_i u_i) dt \right] \end{aligned} \quad (8)$$

Since it is an infinite horizon game, the optimal cost functions J_i^1 remains finite if and only if $e^{-\beta t/2} x(t)$ converges to 0 as $t \rightarrow \infty$ and for any initial state x_0 . Therefore, without loss of generality, we may assume that the Nash strategies belong to the following set of admissible strategies:

$$\begin{aligned} \mathbb{U}_1^{ad}(\beta) &\triangleq \\ \{ (u_1(t, x), u_2(t, x)) \mid \forall t, \forall x_0 \in \mathbb{R}^n, \exists \mathcal{M}, \alpha < \beta \text{ s.t.} \\ \mathbb{E}[\|x(t)\|] &\leq \mathcal{M}e^{\alpha t/2}, \mathbb{E}[\|x(t)\|^2] \leq \mathcal{M}e^{\alpha t}; \\ \text{and } dx &= (Ax + B_1 u_1 + B_2 u_2)dt + GdW_t \} \end{aligned} \quad (9)$$

Let us take an arbitrary interval $(t_0, t_1]$ and consider the following filtering equation (10) for the quantity $y(t) = \mathbb{E}[e^{-\frac{\beta}{2}(t-t_0)} x(t) \mid x(t_0)]$ on the interval:

$$\begin{aligned} \dot{y}(t) &= \tilde{A}y(t) + B_1 \tilde{u}_1 + B_2 \tilde{u}_2 \\ y(t_0) &= x(t_0) \end{aligned} \quad (10)$$

where $\tilde{A} = A - \frac{\beta}{2}I$ and $\tilde{u}_i = e^{-\frac{\beta}{2}(t-t_0)} u_i$. Following these notations, we can write the cost function (7) for the arbitrary interval $(t_0, t_1]$ as given below:

$$\begin{aligned} \tilde{J}_i^1(t_0, t_1) &= e^{\beta t_0} J_i^1(t_0, t_1, \cdot) = \mathbb{E} \left[\int_{t_0}^{t_1} (y' L_i y + \tilde{u}_i' R_i \tilde{u}_i) dt \right] + \\ &\int_{t_0}^{t_1} \int_{t_0}^t e^{-\beta(t-t_0)} \text{tr}(\|\Phi_A(t, s) G\|_{L_i}^2) ds dt \end{aligned} \quad (11)$$

where Φ_A is the state transition matrix corresponding to the drift matrix A .

Due to the filtered equation (10), $y(t)$ is an $x(t_0)$ measurable function for all $t \in (t_0, t_1]$ and hence the expectation in the equation (11) is with respect to the distribution $\mathbb{P}(x(t_0))$. The Nash equilibrium of $\tilde{J}_i^1(t_0, t_1)$ for any arbitrary interval can be found by solving certain coupled Riccati equations [17].

Theorem 4.1: [17] Let P_1, P_2 be positive definite matrices such that the pair (P_1, P_2) is the solution of the following algebraic matrix Riccati equations (12)

$$P_1 \tilde{A} + \tilde{A}' P_1 + L_1 - P_1 B_1 R_1^{-1} B_1' P_1 - P_1 B_2 R_2^{-1} B_2' P_2 = 0 \quad (12a)$$

$$P_2 \tilde{A} + \tilde{A}' P_2 + L_2 - P_2 B_1 R_1^{-1} B_1' P_1 - P_2 B_2 R_2^{-1} B_2' P_2 = 0. \quad (12b)$$

The Nash open-loop strategy for the players within the interval $(t_0, t_1]$ is given by

$$\tilde{u}_i^*(t) = -R_i^{-1} B_i' P_i \Phi(t, t_0) x(t_0) \quad (13)$$

where $\Phi(t, s)$ satisfies the equations:

$$\begin{aligned} \frac{d}{dt} \Phi(t, s) &= (\tilde{A} - B_1 R_1^{-1} B_1' P_1 - B_2 R_2^{-1} B_2' P_2) \Phi(t, s) \\ \Phi(t, s) &= \Phi(t, r) \Phi(r, s) \\ \Phi(t, t) &= I \quad \forall t, s, r. \end{aligned} \quad (14)$$

Moreover, the optimal cost for player- i for the interval $(t_0, t_1]$ of the game is found to be:

$$\tilde{J}_i^1(t_0, t_1) = \mathbb{E}[y(t_0)' M_i y(t_0) - y(t_1)' M_i y(t_1)] + \int_{t_0}^{t_1} \int_{t_0}^t e^{-\beta(t-t_0)} \text{tr}(\|\Phi_A(t, s) G\|_{L_i}^2) ds dt. \quad (15)$$

M_i satisfies the Lyapunov equation:

$$A_c' M_i + M_i A_c + L_i + P_i B_i R_i^{-1} B_i' P_i = 0, \quad (16)$$

where $A_c = \tilde{A} - B_1 R_1^{-1} B_1' P_1 - B_2 R_2^{-1} B_2' P_2$.

From the relationship of $x(t)$ and $y(t)$ over the interval $(t_0, t_1]$, we can verify that $y(t)$ is independent of the noise W_s for all $s \in (t_0, t_1]$ and

$$e^{-\frac{\beta}{2}(t-t_0)} x(t) = y(t) + \int_{t_0}^t e^{-\frac{\beta}{2}(t-t_0)} \Phi_A(t, s) G dW_s. \quad (17)$$

Equation (17) leads to the fact that

$$\mathbb{E}[x(t)' M_i x(t)] = e^{\beta(t-t_0)} \mathbb{E}[y(t)' M_i y(t)] + \int_{t_0}^t \text{tr}(\|\Phi_A(t, s) G\|_{M_i}^2) ds. \quad (18)$$

Substituting (18) in (15), we obtain

$$\begin{aligned} \tilde{J}_i^1(t_0, t_1) &= \mathbb{E}[y(t_0)' M_i y(t_0) - e^{-\beta(t_1-t_0)} x(t_1)' M_i x(t_1)] + \\ &\int_{t_0}^{t_1} \int_{t_0}^t e^{-\beta(t-t_0)} \text{tr}(\|\Phi_A(t, s) G\|_{L_i}^2) ds dt + \\ &\int_{t_0}^{t_1} e^{-\beta(t_1-t_0)} \text{tr}(\|\Phi_A(t_1, s) G\|_{M_i}^2) ds. \end{aligned} \quad (19)$$

Therefore, the optimal cost $J_i^{1*}(t_0, t_1)$ for an arbitrary interval $(t_0, t_1]$ can be written as follows

$$\begin{aligned} J_i^{1*}(t_0, t_1, \cdot) &= \mathbb{E}[e^{-\beta t_0} x(t_0)' M_i x(t_0) \\ &\quad - e^{-\beta t_1} x(t_1)' M_i x(t_1)] \\ &\quad + \int_{t_0}^{t_1} \int_{t_0}^t e^{-\beta t} \text{tr}(\|\Phi_A(t, s) G\|_{L_i}^2) ds dt \\ &\quad + \int_{t_0}^{t_1} e^{-\beta t_1} \text{tr}(\|\Phi_A(t_1, s) G\|_{M_i}^2) ds. \end{aligned} \quad (20)$$

Now we may take $t_1 \rightarrow \infty$ in the analysis performed above and all the asymptotic results hold as shown in [17]. Since the Nash strategies are restricted to $\mathbb{U}_1^{ad}(\beta)$, we will have $\lim_{t_1 \rightarrow \infty} e^{-\beta t_1} x(t_1)' M_i x(t_1) = 0$. Therefore, for an interval (t_0, ∞) , we can write

$$\begin{aligned} J_i^{1*}(t_0, \infty, \cdot) &= \mathbb{E}[e^{-\beta t_0} y(t_0)' M_i y(t_0)] + \\ &\int_{t_0}^{\infty} \int_{t_0}^t e^{-\beta t} \text{tr}(\|\Phi_A(t, s) G\|_{L_i}^2) ds dt + \\ &\lim_{t_1 \rightarrow \infty} \int_{t_0}^{t_1} e^{-\beta t_1} \text{tr}(\|\Phi_A(t_1, s) G\|_{M_i}^2) ds. \end{aligned} \quad (21)$$

Since we restrict the controller strategies within the admissible set $\mathbb{U}_1^{ad}(\beta)$, we can conclude that

$$\lim_{t_1 \rightarrow \infty} \int_{t_0}^{t_1} e^{-\beta t_1} \text{tr}(\|\Phi_A(t_1, s) G\|_{M_i}^2) ds = 0.$$

Lemma 4.2: Let a given switching strategy S accesses the state at the time instances $\{\tau_i\}_{i=0}^N$ ($0 = \tau_0 < \tau_1 < \dots < \tau_N < \infty$) over the infinite horizon of the game. The optimal cost of player- i under switching S can be written as:

$$\begin{aligned} J_i^{1*}(S) &= \mathbb{E}[x_0' M_i x_0] + \\ &\sum_{k=1}^N \int_{\tau_{k-1}}^{\tau_k} \int_{\tau_{k-1}}^t e^{-\beta t} \text{tr}(\|\Phi_A(t, s) G\|_{L_i}^2) ds dt + \\ &\sum_{k=1}^N \int_{\tau_{k-1}}^{\tau_k} e^{-\beta \tau_k} \text{tr}(\|\Phi_A(\tau_k, s) G\|_{M_i}^2) ds + \\ &\int_{\tau_N}^{\infty} \int_{\tau_N}^t e^{-\beta t} \text{tr}(\|\Phi_A(t, s) G\|_{L_i}^2) ds dt \end{aligned} \quad (22)$$

Proof: The switching strategy S partitions the interval $(0, \infty)$ into a collection of subintervals $\{(\tau_i, \tau_{i+1}]\}_{i=0}^N$, where $\tau_{N+1} = \infty$. From Theorem (4.1), we can conclude for any such subintervals $(\tau_k, \tau_{k+1}]$

$$\begin{aligned} J_i^{1*}(\tau_k, \tau_{k+1}, \cdot) &= \mathbb{E}[e^{-\beta \tau_k} x(\tau_k)' M_i x(\tau_k) - \\ &\quad e^{-\beta \tau_{k+1}} x(\tau_{k+1})' M_i x(\tau_{k+1})] + \\ &\int_{\tau_k}^{\tau_{k+1}} \int_{\tau_k}^t e^{-\beta t} \text{tr}(\|\Phi_A(t, s) G\|_{L_i}^2) ds dt + \\ &\int_{\tau_k}^{\tau_{k+1}} e^{-\beta \tau_{k+1}} \text{tr}(\|\Phi_A(\tau_{k+1}, s) G\|_{M_i}^2) ds. \end{aligned} \quad (23)$$

Due to the optimality principle, the strategy u_i^* is optimal for $(0, \infty)$, if the strategy restricted to any arbitrary interval $(t_0, t_1) \subset (0, \infty)$ is an optimal strategy for that interval. Thus, the optimal strategy for the infinite horizon game can

be found by concatenating the optimal strategies over the intervals $(\tau_k, \tau_{k+1}]$. Therefore,

$$u_i^*(t) = -R_i^{-1}B_i'P_i\Phi(t, \tau_k)x(\tau_k) \quad (24)$$

$$\forall t \in (\tau_k, \tau_{k+1}], k = 1, \dots, N.$$

Consequently, the optimal cost over the entire interval $(0, \infty)$ is obtained by adding the costs over the intervals $(\tau_k, \tau_{k+1}]$, i.e.

$$J_i^{1*}(\mathcal{S}) = \sum_{k=0}^N J_i^{1*}(\tau_k, \tau_{k+1}, \cdot). \quad (25)$$

This leads to

$$J_i^{1*}(\mathcal{S}) = \sum_{k=1}^N \int_{\tau_{k-1}}^{\tau_k} \int_{\tau_{k-1}}^t e^{-\beta t} \text{tr}(\|\Phi_A(t, s)G\|_{L_i}^2) ds dt +$$

$$\sum_{k=1}^N \int_{\tau_{k-1}}^{\tau_k} e^{-\beta \tau_k} \text{tr}(\|\Phi_A(\tau_k, s)G\|_{M_i}^2) ds + \quad (26)$$

$$\int_{\tau_N}^{\infty} \int_{\tau_N}^t e^{-\beta t} \text{tr}(\|\Phi_A(t, s)G\|_{L_i}^2) ds dt + \mathbb{E}[x_0' M_i x_0].$$

Corollary 4.3: *Finiteness of the cost $J_i^{1*}(\mathcal{S})$ depends on the existence of a $\tau_N < \infty$ such that*

$$\int_{\tau_N}^{\infty} \int_{\tau_N}^t e^{-\beta(t-\tau_N)} \text{tr}(\|\Phi_A(t, s)G\|_{L_i}^2) ds dt < \infty. \quad (27)$$

Therefore, the necessary condition is that

$$\lim_{t \rightarrow \infty} \int_{\tau_N}^t e^{-\beta(t-\tau_N)} \text{tr}(\|\Phi_A(t, s)G\|_{L_i}^2) ds = 0, \quad (28)$$

or equivalently,

$$\lim_{t \rightarrow \infty} \int_{\tau_N}^t e^{-\beta s} \text{tr}(\|\Phi_{\tilde{A}}(t, s)G\|_{L_i}^2) ds = 0. \quad (29)$$

We will assume in this work that (29) holds and thus our problem is well defined.

Remark 4.4: A sufficient condition for (29) to hold is that $\tilde{A} = A - \frac{\beta}{2}I$ is Hurwitz.

A. Pareto Optimality of the Switching Strategy

Due to Lemma 4.2, we have the optimal cost of the game expressed explicitly as a function of the switching strategy \mathcal{S} . Therefore, player- i should optimize (22) with respect to the switching strategy \mathcal{S} , which is equivalent to a \mathbb{R}^N dimensional optimization problem.

In this game formulation, we have a single switch and thus the problem at this stage is a multi-objective (two objectives: $J_1^{1*}(\mathcal{S}), J_2^{1*}(\mathcal{S})$) optimization problem over the optimization variable \mathcal{S} .

Let us denote the set of feasible switching strategies by \mathbb{S} which is defined in the following way:

$$\mathbb{S} = \{ \{\tau_k\}_{k=0}^N \mid 0 = \tau_0 < \tau_1 < \dots < \tau_N < \infty, \text{ and } c \geq N \in \mathbb{N} \} \quad (30)$$

where \mathbb{N} is the set of natural numbers, and c is the maximum number of allowed switching. It is important to note here

that $\mathbb{S} \subset \mathbb{R}^c$. The optimization problem at this stage can be written as

$$\min_{\mathcal{S} \in \mathbb{S}} \{J_1^{1*}(\mathcal{S}), J_2^{1*}(\mathcal{S})\} \quad (31)$$

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

Definition 4.5: A feasible point $s \in \mathbb{S}$ is said to (Pareto) dominate another feasible point $s_1 \in \mathbb{S}$ if

1. $J_i^{1*}(s) \leq J_i^{1*}(s_1)$ for all $i = 1, 2$, and
2. $J_j^{1*}(s) < J_j^{1*}(s_1)$ for some $j \in \{1, 2\}$.

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

The set of Pareto optimal outcomes is often known as the Pareto frontier.

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

$$J(\mathcal{S}) = \theta J_1^{1*}(\mathcal{S}) + (1 - \theta) J_2^{1*}(\mathcal{S}), \quad (32)$$

where $\mathcal{S} \in \mathbb{S}$ is a feasible switching strategy. It is well known in the literature that a Pareto optimal solution of (31) has a one-to-one correspondence with the solution of (32) for some specific value of θ [18].

Proposition 4.6: For every $\theta \in [0, 1]$, (32) attains a minimum for some $\mathcal{S}(\theta) \in \mathbb{S}$.

Proof: [sketch] J_i^{1*} attains a minimum for every (M_i, L_i) pair such that $M_i, L_i \succ 0$. $J(\mathcal{S})$ is obtained by replacing L_i and M_i by $\theta L_1 + (1 - \theta)L_2$ and $\theta M_1 + (1 - \theta)M_2$ respectively. ■

Due to space limitation, we will conclude this section by pointing to the crucial fact that the optimization problem to find the Pareto optimal points are finite dimensional optimization problems and can be solved using classical techniques such as gradient descent.

V. GAMES WITH AVERAGE COST

In this section, we study the other variation of an infinite horizon game where the objective of is to optimize the average cost which is given by:

$$J_i^2(u_1, u_2, \mathcal{S}) = \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\int_0^T (x' L_i x + u_i' R_i u_i) dt \right] \quad (33)$$

Similar to the discounted cost problem, in this case also we want to constraint control and switching strategies so that the cost remains finite. We define the following set of admissible strategies for the controllers in (34).

$$\mathbb{U}_2^{ad} \triangleq$$

$$\{ (u_1(t, x), u_2(t, x)) \mid \forall t, \forall x_0 \in \mathbb{R}^n, \exists \mathcal{M} \text{ s.t.}$$

$$\mathbb{E}[\|x(t)\|] \leq \mathcal{M}, \mathbb{E}[\|x(t)\|^2] \leq \mathcal{M};$$

$$\text{and } dx = (Ax + B_1 u_1 + B_2 u_2)dt + GdW_t \} \quad (34)$$

Let us consider the filtration $z(t) = \mathbb{E}[x(t) \mid x(t_0)]$ for the stochastic dynamics for $t > t_0$. $z(t)$ satisfies the differential equation

$$\dot{z}(t) = Az(t) + B_1 u_1 + B_2 u_2 \quad (35)$$

$$z(t_0) = x(t_0).$$

Following similar steps as done for discounted cost problem, we consider an arbitrary interval $(t_0, t_1]$ and within this interval we define

$$J_i^2(t_0, t_1, \cdot) = J_i^2(t_0, t_1, u_1, u_2, \mathcal{S}) \quad (36)$$

$$= \mathbb{E} \left[\int_{t_0}^{t_1} (x' L_i x + u_i' R_i u_i) dt \right]$$

Therefore, we can write

$$J_i^2(t_0, t_1, \cdot) = \mathbb{E} \left[\int_{t_0}^{t_1} (z' L_i z + u_i' R_i u_i) dt \right] + \int_{t_0}^{t_1} \int_{t_0}^t \text{tr}(\|\Phi_A(t, s) G\|_{L_i}^2) ds dt \quad (37)$$

Similar to Theorem 4.1, we have the following theorem due to [17] that gives the open loop Nash strategy for the duration $(t_0, t_1]$.

Theorem 5.1: Let Q_1, Q_2 be positive definite matrices such that the pair (Q_1, Q_2) is the solution of the following algebraic matrix Riccati equation (38)

$$Q_1 A + A' Q_1 + L_1 - Q_1 B_1 R_1^{-1} B_1' Q_1 - Q_1 B_2 R_2^{-1} B_2' Q_2 = 0, \quad \mathbb{S}(N) = \{\{\tau_k\}_{k=0}^n \mid 0 = \tau_0 < \tau_1 < \dots < \tau_n < T; n \leq N\} \quad (38a)$$

$$Q_2 A + A' Q_2 + L_2 - Q_2 B_1 R_1^{-1} B_1' Q_1 - Q_2 B_2 R_2^{-1} B_2' Q_2 = 0. \quad (38b)$$

The open-loop Nash strategy for the players within the interval $(t_0, t_1]$ is given by

$$u_i^*(t) = -R_i^{-1} B_i' Q_i \Psi(t, t_0) x(t_0) \quad (39)$$

where $\Psi(t, s)$ satisfies the equations:

$$\begin{aligned} \frac{d}{dt} \Psi(t, s) &= (A - B_1 R_1^{-1} B_1' Q_1 - B_2 R_2^{-1} B_2' Q_2) \Psi(t, s) \\ \Psi(t, s) &= \Psi(t, r) \Psi(r, s) \\ \Psi(t, t) &= I \quad \forall t, s, r. \end{aligned} \quad (40)$$

Moreover, the optimal cost for player- i for the interval $(t_0, t_1]$ of the game is found to be:

$$J_i^{2*}(t_0, t_1, \cdot) = \mathbb{E}[z(t_0)' C_i z(t_0) - z(t_1)' C_i z(t_1)] + \int_{t_0}^{t_1} \int_{t_0}^t \text{tr}(\|\Phi_A(t, s) G\|_{L_i}^2) ds dt. \quad (41)$$

C_i satisfies the Lyapunov equation:

$$A_{cl}' C_i + C_i A_{cl} + L_i + Q_i B_i R_i^{-1} B_i' Q_i = 0, \quad (42)$$

where $A_{cl} = A - B_1 R_1^{-1} B_1' Q_1 - B_2 R_2^{-1} B_2' Q_2$.

Using the fact that $x(t) = z(t) + \int_{t_0}^t \Phi_A(t, s) G dW_s$, we can write

$$J_i^{2*}(t_0, t_1, \cdot) = \mathbb{E}[x(t_0)' C_i x(t_0) - x(t_1)' C_i x(t_1)] + \int_{t_0}^{t_1} \int_{t_0}^t \text{tr}(\|\Phi_A(t, s) G\|_{L_i}^2) ds dt + \int_{t_0}^{t_1} \text{tr}(\|\Phi_A(t_1, s) G\|_{C_i}^2) ds \quad (44)$$

Thus, for a switching strategy $\mathcal{S} = \{\tau_k\}_{k=1}^N$ over an interval $(0, T]$, the cost is

$$J_i^{2*}(0, T, \cdot) = \mathbb{E}[x_0' C_i x_0 - x(T)' C_i x(T)] + \sum_{k=1}^{N+1} \int_{\tau_{k-1}}^{\tau_k} \int_{\tau_{k-1}}^t \text{tr}(\|\Phi_A(t, s) G\|_{L_i}^2) ds dt + \sum_{k=1}^{N+1} \int_{\tau_{k-1}}^{\tau_k} \text{tr}(\|\Phi_A(\tau_k, s) G\|_{C_i}^2) ds \quad (45)$$

where $\tau_0 = 0$ and $\tau_{N+1} = T$.

For this game problem, we have the constraint that over an interval $(a, b]$, the maximum number of samples that can be acquired is $\lceil d(b-a) \rceil$. The following section describes the optimal switching strategy for an the infinite horizon game. Then we use the optimal switching strategy to explicitly calculate the cost (33).

A. Optimality of the Switching Policy

Let us consider an arbitrary interval $(0, T]$ and let us denote $N = \lceil Td \rceil$. Therefore, we aim to find a switching policy from the following set:

$$\mathbb{S}(N) = \{\{\tau_k\}_{k=0}^n \mid 0 = \tau_0 < \tau_1 < \dots < \tau_n < T; n \leq N\} \quad (46)$$

Proposition 5.2: For any $\tau_k > \tau_{k-1}$,

$$\begin{aligned} & \int_{\tau_{k-1}}^{\tau_k} \int_{\tau_{k-1}}^t \text{tr}(\|\Phi_A(t, s) G\|_{L_i}^2) ds dt + \\ & \int_{\tau_{k-1}}^{\tau_k} \text{tr}(\|\Phi_A(\tau_k, s) G\|_{C_i}^2) ds = \\ & \int_{\tau_{k-1}}^{\tau_k} \int_{\tau_{k-1}}^t \text{tr}(\|\Phi_A(t, s) G\|_{\tilde{L}_i}^2) ds dt + \text{tr}(G' C_i G)(\tau_k - \tau_{k-1}) \end{aligned} \quad (47)$$

where $\tilde{L}_i = L_i + A' C_i + C_i A \succ 0$.

The objective of player- i is to select the optimal switching strategy $\mathcal{S} \in \mathbb{S}(N)$ for the interval $(0, T]$ to minimize the cost. Let us define

$$H_i(n, \tau_0, \tau_1, \dots, \tau_n) = J_i^{2*}(0, T, \cdot) = \text{tr}(G' C_i G) T + \sum_{k=1}^{n+1} \int_{\tau_{k-1}}^{\tau_k} \int_{\tau_{k-1}}^t \text{tr}(\|\Phi_A(t, s) G\|_{\tilde{L}_i}^2) ds dt \quad (48)$$

where $\tau_0 = 0$ and $\tau_{n+1} = T$. For a fixed $n \leq N = \lceil Td \rceil$, we evaluate the first order necessary conditions $\frac{\partial H_i}{\partial \tau_k} \equiv 0$

$$\frac{\partial H_i}{\partial \tau_k} = \int_{\tau_{k-1}}^{\tau_k} \text{tr}(\|\Phi_A(\tau_k, s) G\|_{\tilde{L}_i}^2) ds - \int_{\tau_k}^{\tau_{k+1}} \text{tr}(\|\Phi_A(s, \tau_k) G\|_{\tilde{L}_i}^2) ds \quad (49)$$

Using the shift invariance property of $\Phi_A(t, s)$ along with the fact that $\Phi_A(t, s) = \Phi_A(-s, -t)$ we obtain

$$\begin{aligned} & \int_{\tau_k}^{\tau_{k+1}} \text{tr}(\|\Phi_A(s, \tau_k) G\|_{\tilde{L}_i}^2) ds = \\ & \int_{\tau_k}^{\tau_{k+1}} \text{tr}(\|\Phi_A(\tau_{k+1} - \tau_k, \tau_{k+1} - s) G\|_{\tilde{L}_i}^2) ds = \\ & \int_0^{\tau_{k+1} - \tau_k} \text{tr}(\|\Phi_A(\tau_{k+1} - \tau_k, s) G\|_{\tilde{L}_i}^2) ds, \end{aligned} \quad (50)$$

and similarly,

$$\int_{\tau_{k-1}}^{\tau_k} \text{tr}(\|\Phi_A(\tau_k, s)G\|_{L_i}^2)ds = \int_0^{\tau_k - \tau_{k-1}} \text{tr}(\|\Phi_A(\tau_k - \tau_{k-1}, s)G\|_{L_i}^2)ds.$$

Combining the above two results, we obtain

$$\frac{\partial H_i}{\partial \tau_k} = \int_0^{\tau_k - \tau_{k-1}} \text{tr}(\|\Phi_A(\tau_k - \tau_{k-1}, s)G\|_{L_i}^2)ds - \int_0^{\tau_{k+1} - \tau_k} \text{tr}(\|\Phi_A(\tau_{k+1} - \tau_k, s)G\|_{L_i}^2)ds. \quad (51)$$

Therefore, $\frac{\partial H_i}{\partial \tau_k} = 0$ requires $\tau_k - \tau_{k-1} = \tau_{k+1} - \tau_k$ for all k . This leads to the fact that the inter-sampling duration should be constant through out the interval $(0, T]$ and the sampling instances are given by

$$\tau_k^* = k \frac{T}{n+1}.$$

With this structure of τ_k , we can simplify $H_i(n, \tau_0, \tau_1, \dots, \tau_n)$ into

$$\mathbb{H}_i(n) = H_i(n, \tau_0^*, \tau_1^*, \dots, \tau_n^*) = \text{tr}(G' C_i G)T + (n+1) \int_0^{\frac{T}{n+1}} \int_0^t \text{tr}(\|\Phi_A(t, s)G\|_{L_i}^2)dsdt \quad (52)$$

One can check that $\mathbb{H}_i(n)$ is a decreasing function of n . Consequently $n^* = N = \lceil Td \rceil$.

Remark 5.3: The optimal τ_k^* are same for both the players in this case. The optimal switching strategy over an arbitrary interval $(0, T]$ is a periodic strategy with period $p(T) = \frac{T}{\lceil Td \rceil + 1}$. Therefore,

$$p^* = \lim_{T \rightarrow \infty} p(T) = \frac{1}{d}.$$

Thus, the optimal switching policy for the infinite horizon game is a periodic sampling policy where the period is the inverse of the bound d .

Therefore, the optimal cost of player- i is:

$$J_i^{2*}(u_1^*, u_2^*, S^*) = \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{H}_i(n) = \text{tr}(G' C_i G) + d \int_0^{\frac{1}{d}} \int_0^t \text{tr}(\|\Phi_A(t, s)G\|_{L_i}^2)dsdt \quad (53)$$

Thus, the switching strategy is similar to periodic control or time triggered control and hence this framework can be used to study game theoretic properties of large multi-agent systems under periodic or time-triggered control.

VI. CONCLUSIONS

We have considered an infinite horizon linear quadratic stochastic differential game with two different types of cost functions. The game under consideration is neither a closed-loop nor an open-loop game, rather there is a switch which can provide finite number of discrete time measurements. We show the existence of Nash control strategy of such a

game and then we address the question of optimal switching strategy.

The results illustrate that the optimal switching policy (or equivalently the optimal sampling times) for the discounted cost problem can be found by constructing a Pareto frontier of two optimization functions. On the other hand, the average cost problem shows that the optimal switching is a periodic sampling policy where the period can be uniquely determined from the given bound on the number of sampling in an unit interval.

REFERENCES

- [1] T. Basar and G. J. Olsder, *Dynamic noncooperative game theory*. SIAM, 1995, vol. 200.
- [2] J. Engwerda, *LQ dynamic optimization and differential games*. John Wiley & Sons, 2005.
- [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] 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.
- [5] J. Cruz Jr and C. Chen, "Series Nash solution of two-person, nonzero-sum, linear-quadratic differential games," *Journal of Optimization Theory and Applications*, vol. 7, no. 4, pp. 240–257, 1971.
- [6] R. Isaacs, *Differential games: a mathematical theory with applications to warfare and pursuit, control and optimization*. Courier Corporation, 1999.
- [7] 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.
- [8] D. Maity and J. S. Baras, "Event based control of stochastic linear systems," in *Event-based Control, Communication, and Signal Processing (ECCSP), 2015 International Conference on*. IEEE, 2015, pp. 1–8.
- [9] 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.
- [10] W. Bian and M. French, "General fast sampling theorems for nonlinear systems," *Systems & control letters*, vol. 54, no. 11, pp. 1037–1050, 2005.
- [11] D. Maity and J. S. Baras, "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.
- [12] D. V. Dimarogonas and K. H. Johansson, "Event-triggered control for multi-agent systems," in *Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on*. IEEE, 2009, pp. 7131–7136.
- [13] M. Wooldridge, *An introduction to multiagent systems*. John Wiley & Sons, 2009.
- [14] D. Maity and J. S. Baras, "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.
- [15] D. Maity, A. Raghavan, and J. S. Baras, "Stochastic differential linear-quadratic games with intermittent asymmetric observations," in *American Control Conference (ACC), 2017*. IEEE, 2017 (to appear).
- [16] D. Maity and J. S. Baras, "Linear quadratic stochastic differential games under asymmetric value of information," *IFAC World Congress*, 2017 (to appear).
- [17] J. C. Engwerda, "On the open-loop nash equilibrium in lq-games," *Journal of Economic Dynamics and Control*, vol. 22, no. 5, pp. 729–762, 1998.
- [18] 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.