Proceedings of the 13th International Workshop on Discrete Event Systems, Xi'an, China, May 30 - June 1, 2016

Strategies for Two-Player Differential Games with Costly Information

Dipankar Maity and John S. Baras

Abstract— **In this work, a two players nonzero-sum differential game is considered, where one player tries to minimize some predefined cost and the other tries to maximize the same. The game is described by a stochastic differential system and the actions of the players serve as the control inputs to the dynamical system. The cost being a function of the actions chosen by the players and the state of the dynamical system, the players aim to control the state in order to optimize the cost functional. However in this problem the players do not have the access to the states for every time, rather the states are available at discrete time instances after some finite costs are paid by the players. The inclusion of the information-cost makes the structure of the cost functional non-classical. The work presents the strategies for the players under no-cost information access as well as under costly information access. Explicit time instances for the information access are also derived by solving certain finite dimensional optimization problems.**

I. INTRODUCTION

To our best knowledge, differential games with the zero-sum framework were first introduced in [1]. [2] studied an special case of nonzero-sum framework extending the results of zero-sum games. Soon after that, nonzero-sum games were formally introduced and studied in the works of Starr and Ho [3], [4]. Differential games were not merely studied for the sake of game theory rather the study reveals many facts about robust control, minimax stochastic control and general stochastic control problems [5], [6], [7], [8], [9]. Nash equilibrium strategies and their properties for different game frameworks have been studied for decades. Existence and uniqueness of such strategies are of paramount interests to [10], [11], [12] and many others.

Linear-Quadratic differential games (LQDG) are one of the most important and most studied class of problems among the various classes of differential games. In this framework, the state of the game -i.e. the state of the underlying dynamics- depends linearly on the strategies of the players and the optimization criteria for the players are quadratic. Most game problems either lack a closed form expression for the Nash strategy or it is very intricate to calculate one. However, for LQDGs, the closed form expression for the Nash strategy is available and it can be obtained by solving certain coupled Riccati type equations [11], [13], [14]. While [10], [15] and many others studied the necessary and sufficient conditions for a strategy to be a Nash strategy, [12] addressed the problem of uniqueness of a Nash strategy. While a vast portion of the previous works address LQDG of fixed, finite time, [11] studies the asymptotic behavior of the Nash strategy when the duration of the game extends to infinity. Alongside, the notion of Stackelberg strategies [16] has been developed to model games where one of the players does not know the performance criterion of the opponent.

In principle a non-cooperative game is a decision making problem and thus the information structures available to the players play a great role in determining the strategies for the players. In the class of Linear-Quadratic games (LQG), there are notions of open-loop, close-loop, feedback information patterns where the players can have the knowledge of the initial state of the game only, or all present and past states, or only the present state respectively. This limited information determines the different structures of the corresponding Nash equilibrium strategies [17]. [18] provided a counterexample showing the existence of a nonlinear Nash strategy for some LQG, startled the understanding that linear strategies are optimal for LQGs. A very recent work [19] considered a differential game with the noise depending linearly on the state. The optimal strategy is shown to be a linear state feedback for the game [19], however it requires solving a stochastic Riccati-like equation for the players rather than a deterministic one. The information pattern for this game is feedback type and hence the players can access the current state without paying any cost.

In this work, we consider a nonclassical LQDG framework where we add a cost for acquiring state information. This new structure of the cost introduces notable changes in the well known structure of the Nash strategy for a similar game with no-cost information acquisition. In the long history of differential games, the problem of including information cost somehow lacked attention. This work aims to address the question of 'value of information' from an LQDG point of view. The players can access state information only at discrete time instances and each such access query carries a finite cost. In addition to the costly and limited availability of information, the players do not have any information about the opponent's information space.

In the cost function, we take an integral cost plus pointwise quadratic costs at several intermediate time instances. Pointwise cost at final time (T) is mostly considered in the framework of Linear-Quadratic control, however multiple pointwise costs allow the possibility of putting extra emphasis on some discrete time instances. This unusual structure of the cost changes the structure of the optimal strategy even for no-cost information access case. The computation of

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 National Institute of Standards and Technology (NIST) grant 70NANB11H148.

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

978-1-5090-4190-9/16/$31.00 ©2016 IEEE

the optimal strategy requires solving a Riccati-type equation semi-coupled with a linear ordinary differential equation.

Since the information acquisition is costly, the players should ask for information only at some optimal time instances. One of our intentions is to study the relation (if any) between the time instances of acquiring state information and that of incurring pointwise cost. Our study shows that there is no relation between them. To our surprise, the optimization problems for finding the optimal strategy and that of finding the optimal instances to acquire state information turn out to be decoupled and the later can be solved offline.

The paper starts by finding the optimal strategy for the players under no-cost information acquisition in Section III. The analysis in Section III provides an indication on the possible structures of the strategies for the costly information access case. In the next section we add the cost for information acquisition and solve the two decoupled optimization problems to describe the structure of the optimal strategy and the optimal way to construct the information sets.

II. Problem Formulation

Let us consider a two-player game $\mathcal{G}(x, J, u, v)$ as defined below,

Definition 2.1: A two-player game $\mathcal{G}(x, J, u, v)$ consists of four components: $x(t)$: state of the underlying dynamical system at time t, $u(t), v(t)$: actions chosen by the players at time t, and J: the cost associated with the game. In a non-cooperative game set up, the objective of one player is to minimize the cost, whereas the other player maximizes the cost.

The state of the game obeys the following linear stochastic dynamics (1).

$$dx = Axdt + Budt + Cvdt + GdW_t \tag{1}$$

where W_t is an r dimensional standard Wiener process and the above dynamics should be treated in the sense of Ito. $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^{m_1}$ and $v(t) \in \mathbb{R}^{m_2}$. The model parameters A, B, C and G are time variant however, to maintain notational brevity, we suppress the time argument.

The state of the game is linearly affected by the actions taken by the players and also by a Wiener process noise on which none of the players has any control.
The game is played for a finite time interval $[0, T]$ and the cost function is considered to be (2).

$$J(u,v) = \mathbf{E}\Big[\int_0^T (\|x\|_L^2 + \|u\|_R^2 - \|v\|_S^2)dt + \sum_{i=1}^N \|x(t_i) - r_i\|_{\alpha_i}^2\Big] \tag{2}$$

where $\|p\|_Q^2 = p'Qp$ for any two matrices p and Q of proper dimensions, and $L, R, S > 0$. The cost function J consists of a running cost (integral term) and pointwise costs at some predefined time instances $0 \le t_1 < t_2 < \cdots < t_N \le T$.

The cost function has to be minimized for player-1 (P1) who selects the action u, while player-2 (P2) aims to maximize the cost by selecting the action v. In classical LQG control, there is (if any) only terminal state cost however, the cost which we have considered is more general and it allows for penalization if the state at time t_i deviates from some given reference value r_i.

Each player has only limited information about the state of the game and they can ask for the current state of the game by paying some finite cost. Let us denote the number of times state information is requested by P1 up to time t be $n_1(t)$ and that of P2 be $n_2(t)$. The cost associated with these information acquisitions are $\lambda_1 n_1(t)$ and $\lambda_2 n_2(t)$. We add these costs to the cost function J and can derive two separate cost functions for the players ($k = 1, 2$):

$$\begin{aligned} J_k(u,v) =& \mathbf{E}\Big[\int_0^T (\|x\|_L^2 + \|u\|_R^2 - \|v\|_S^2)dt \\ &+ \sum_{i=1}^N \|x(t_i) - r_i\|_{\alpha_i}^2 - (-1)^k \lambda_k n_k(T)\Big]. \end{aligned} \tag{3}$$

$-\lambda_2 n_2(T)$ was added in the last equation for the fact that P2 will be maximizing that cost function.

The objective of P1 is to determine the time instances τ_k^1 to construct its information set $\mathcal{I}_1(t) = \{x(\tau_k^1)\}_{k=1}^{n_1(t)}$. P1 selects the control u, as a function of this information set, so as to minimize $J_1(u, v)$. On the other hand, P2 forms its information set $\mathcal{I}_2(t)$ and selects v, with the limited knowledge accumulated in $\mathcal{I}_2(t)$, to maximize $J_2(u, v)$ given in (3). The information sets are ordered i.e. $\tau_k^i < \tau_{k+1}^i$ and non-anticipative i.e. for any t, $\tau_{n_i(t)}^i \le t$. The set of admissible controls for P1 (or P2) is adapted to the sigma-field generated by the information set $\mathcal{I}_1(t)$ (or $\mathcal{I}_2(t)$).

III. Optimal Strategies with No-cost Information Access

Before considering the actual problem, let us first consider the situation where $\lambda_1 = \lambda_2 = 0$ i.e. full state information is available without any cost. Under this situation, let us study the optimal strategies for the players of this game.

Theorem 3.1: The optimal strategies for the players are:

$$u^*(t) = -R^{-1}B'\left(\frac{\eta}{2} + Px\right) \tag{4}$$

$$v^*(t) = S^{-1}C'\left(\frac{\eta}{2} + Px\right). \tag{5}$$

The optimal cost is

$$\begin{aligned} J^* =& \mathbf{E}[\|x(0)\|_{P(0)}^2 + x(0)'\eta(0)] + \sum_{i=1}^N \|r_i\|_{\alpha_i}^2 \\ &+ \int_0^T \big(tr(PGG') + \frac{1}{4}\|\eta\|_{CS^{-1}C'}^2 - \frac{1}{4}\|\eta\|_{BR^{-1}B'}^2\big)dt], \end{aligned} \tag{6}$$

where

$$\dot{P} + A'P + PA + L + P(CS^{-1}C' - BR^{-1}B')P = \mathbf{0} \tag{7}$$

$$P(T) = \mathbf{0}$$

$$P(t_i^-) - P(t_i^+) = \alpha_i$$

and

$$\dot{\eta} = -\big(P(CS^{-1}C' - BR^{-1}B') + A'\big)\eta \tag{8}$$

$$\eta(T) = 0$$

$$\eta(t_i^-) - \eta(t_i^+) = -2\alpha_i r_i$$

Proof: Let $\psi(t,x) = x(t)'P(t)x(t)$ and using Ito rule, $\psi(t_i^-)-\psi(t_{i-1}^+) = \int_{t_{i-1}^+}^{t_i^-} \Big(\big(\frac{\partial \psi}{\partial t} + \frac{1}{2}tr(G'\nabla^2\psi G)\big)dt + \nabla\psi dx \Big)$, where ∇ and ∇^2 denote the gradient and Hessian operators respectively. Therefore,
$\mathbf{E}(\psi(t_i^-) - \psi(t_{i-1}^+)) = \mathbf{E}\big[\int_{t_{i-1}^+}^{t_i^-} [x'(\dot{P} + A'P + PA)x + x'P(Bu + Cv) + (Bu + Cv)'Px + tr(PGG')]dt\big]$
Thus,

$$\mathbf{E}[\psi(T) - \psi(0) + \sum_{i=1}^{N}(\psi(t_i^-) - \psi(t_i^+))] =$$
$$\mathbf{E}\Big[\int_0^T [x'(\dot{P} + A'P + PA)x + x'P(Bu + Cv) + (Bu + Cv)'Px + tr(PGG')]dt\Big]. \quad (9)$$

Similarly, let $\phi(t,x) = x(t)'\eta(t)$; following the above steps,

$$\mathbf{E}[\phi(T) - \phi(0) + \sum_{i=1}^{N}(\phi(t_i^-) - \phi(t_i^+))] =$$
$$\mathbf{E}[\int_0^T [x'(\dot{\eta} + A'\eta) + (Bu + Cv)'\eta]dt]. \quad (10)$$

Using (9), (10) and (2), we obtain

$$\begin{aligned} J = \mathbf{E}\big[&\|x(0)\|^2_{P(0)} - \|x(T)\|^2_{P(T)} + x(0)'\eta(0) - x(T)'\eta(T) \\ &+ \sum_{i=1}^{N}\Big(\|x(t_i)\|^2_{\Delta P(t_i)} + x(t_i)'\big(\Delta\eta(t_i)\big) + \|x(t_i) - r_i\|^2_{\alpha_i}\Big) \\ &+ \int_0^T \big(\|x\|^2_{\dot{P}+A'P+PA+L} + \|u\|^2_R - \|v\|^2_S + \\ &x'P(Bu + Cv) + (Bu + Cv)'Px + x'(\dot{\eta} + A'\eta) + \\ &\eta'(Bu + Cv) + tr(PGG')\big)dt\big] \end{aligned} \quad (11)$$

where $\Delta P(t_i) = P(t_i^+) - P(t_i^-)$ and $\Delta\eta(t_i) = \eta(t_i^+) - \eta(t_i^-)$. Rearranging the above terms,

$$\begin{aligned} J = \mathbf{E}\big[&\|x(0)\|^2_{P(0)} - \|x(T)\|^2_{P(T)} + x(0)'\eta(0) - x(T)'\eta(T) \\ &+ \sum_{i=1}^{N}\Big(\|x(t_i)\|^2_{\Delta P(t_i)+\alpha_i} + x(t_i)'\big(\Delta\eta(t_i) - 2\alpha_i r_i\big) + \|r_i\|^2_{\alpha_i}\Big) \\ &+ \int_0^T \big(\|u + R^{-1}B'(Px + \frac{\eta}{2})\|^2_R - \|v - S^{-1}C'(Px + \frac{\eta}{2})\|^2_S \\ &+ x'\big(\dot{\eta} + A'\eta - P(BR^{-1}B' - CS^{-1}C')\eta\big) \\ &+ \|x\|^2_{\dot{P}+A'P+PA+L-P(BR^{-1}B'-CS^{-1}C')P} \\ &- \frac{1}{4}\|\eta\|^2_{BR^{-1}B'-CS^{-1}C'} + tr(PGG')\big)dt\big] \end{aligned} \quad (12)$$

With the assumptions that equations (7) and (8) have well defined solutions[1] in $[0,T]$, we have:

$$\begin{aligned} J = \mathbf{E}\big[&\|x(0)\|^2_{P(0)} + x(0)'\eta(0) + \sum_{i=1}^{N}\|r_i\|^2_{\alpha_i} \\ &+ \int_0^T \big(\|u + R^{-1}B'(Px + \frac{\eta}{2})\|^2_R - \|v - S^{-1}C'(Px + \frac{\eta}{2})\|^2_S \\ &- \frac{1}{4}\|\eta\|^2_{BR^{-1}B'-CS^{-1}C'} + tr(PGG')\big)dt\big] \end{aligned} \quad (13)$$

[1]a sufficient condition would be $BR^{-1}B' - CS^{-1}C' > 0$.

With this completion of squares, it is clear that the optimal strategies for the players are

$$u^* = -R^{-1}B'(Px + \frac{\eta}{2}).$$
$$v^* = S^{-1}C'(Px + \frac{\eta}{2}).$$

This is the only Nash Equilibrium for this game. Therefore, the cost incurred for this optimal strategy is,

$$\begin{aligned} J^*(u^*, v^*) = &\mathbf{E}\big[\|x(0)\|^2_{P(0)} + x(0)'\eta(0) + \sum_{i=1}^{N}\|r_i\|^2_{\alpha_i} \\ &+ \int_0^T \big(tr(PGG') - \frac{1}{4}\|\eta\|^2_{BR^{-1}B'-CS^{-1}C'}\big)dt\big] \end{aligned} \quad (14)$$

■

Remark 3.2: *The optimal strategy consists of an open loop term to optimize the pointwise cost and a closed loop term to optimize both the pointwise cost and the integral cost.*

Remark 3.3: *For the case when $t_N = T$, the jump and boundary conditions for (7) and (8) are given as follows:*

$$P(t_i^-) - P(t_i^+) = \alpha_i$$
$$\eta(t_i^-) - \eta(t_i^+) = -2\alpha_i r_i \quad \forall\ 1 \le i < N,$$

and $P(T) = \alpha_N$, $\eta(T) = -2\alpha_N r_N$.

Remark 3.4: *If the game parameters for both the players are the same i.e. $B = C$ and $R = S$, the game exhibits some interesting properties:*

- *The optimal strategies can be computed by solving two decoupled linear equations of P and η with the same boundary and jump conditions as in (7) and (8).*
- *The optimal cost in this case will be*

$$\begin{aligned} J^* = &\mathbf{E}\big[\|x(0)\|^2_{\sum_{i=1}^{N}\|\Phi_A(t_i,0)\|^2_{\alpha_i} + \int_0^T \|\Phi_A(t,0)\|^2_L dt} \\ &- 2x(0)'\sum_{i=1}^{N}\Phi_A(t_i,0)'\alpha_i r_i\big] + \sum_{i=1}^{N}\|r_i\|^2_{\alpha_i} \\ &+ \int_0^T tr(PGG')dt \end{aligned} \quad (15)$$

- *Any pair of strategies (u, v) of the form $(\gamma, -\gamma)$ will achieve the optimal cost. However, the pair $(-R^{-1}B'Px, R^{-1}B'Px)$ is the Nash equilibrium strategy.*

IV. Strategies with Costly Information

In the previous section, the explicit formulae for the optimal strategies of both the players are obtained and they require the knowledge of the state $x(t)$ for all time $t \in [0,T]$. In this section we will investigate how the strategies change for both the players when they have state information only at finite number of time instances; caused by the cost of acquiring information. Before attacking this problem, we consider a simpler problem where the Information sets $\mathcal{I}_1(t)$ and $\mathcal{I}_2(t)$ are given to P1 and P2 respectively for all t and they have to determine the strategies based on the given information. In the subsequent sections 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.

From (13), we can divide the cost in two parts J_0, J_d -one, (J_0), being independent of the actions of the players and the other, (J_d), that depends on the choice of u and v. Thus,

$$J_0 = \mathbf{E}\big[\|x(0)\|^2_{P(0)} + x(0)'\eta(0) + \sum_1^N \|r_i\|^2_{\alpha_i} + \int_0^T \big(tr(PGG') - \frac{1}{4}\|\eta\|^2_{BR^{-1}B'-CS^{-1}C'}\big)dt\big] \quad (16)$$

and,

$$J_d(u,v) = \mathbf{E}[\int_0^T \big(\|u + R^{-1}B'(Px + \frac{\eta}{2})\|^2_R - \|v - S^{-1}C'(Px + \frac{\eta}{2})\|^2_S\big)dt]. \quad (17)$$

Since the players know the game structures i.e. A, B, C, G, S and R, they can compute the open-loop terms without having any knowledge of the current state. Thus, without loss of generality, it is sufficient for the players to optimize (18).

$$J_d(u,v) = \mathbf{E}[\int_0^T \big(\|u + R^{-1}B'Px\|^2_R - \|v - S^{-1}C'Px\|^2_S\big)dt]. \quad (18)$$

Let us denote $\tilde{R} = PBR^{-1}B'P$ and $\tilde{S} = PCS^{-1}C'P$. Without loss of generality one can show that optimal controllers for P1 and P2 are of the forms $u = -R^{-1}B'P\hat{x}_1$ and $v = S^{-1}C'P\hat{x}_2$, where $\hat{x}_i(t) = f(t, \mathcal{I}_i(t), \hat{x}_i(\tau)|_{\tau\in[0,t)})$. Thus rewriting J_d, we have,

$$J_d(\hat{x}_1, \hat{x}_2) = \mathbf{E}[\int_0^T (\|x - \hat{x}_1\|^2_{\tilde{R}} - \|x - \hat{x}_2\|^2_{\tilde{S}})dt] \quad (19)$$

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

$$dx = (Ax + Bu_0 + Cv_0 - P^{-1}\tilde{R}\hat{x}_1 + P^{-1}\tilde{S}\hat{x}_2)dt + GdW_t. \quad (20)$$

From equation (20), we can write the solution to be:

$$x(t) = \Phi_A(t,t_0)x(t_0) + K^{t,t_0}[Bu_0 + Cv_0](t) - K^{t,t_0}[P^{-1}\tilde{R}\hat{x}_1 - P^{-1}\tilde{S}\hat{x}_2](t) + K_1^{t,t_0}[GW](t) \quad (21)$$

where K^{t,t_0} and K_1^{t,t_0} are linear operators defined as follows:

$$K^{t,t_0}[f](t) = \int_{t_0}^t \Phi_A(t,s)f(s)ds \quad (22)$$

$$K_1^{t,t_0}[fW](t) = \int_{t_0}^t \Phi_A(t,s)f(s)dW(s) \quad (23)$$

If P1 receives the state information at t_0, then the first four terms of (21) are known to P1. Let P1 select its strategy as:

$$\hat{x}_1 = \Phi_A(t,t_0)x(t_0) + K^{t,t_0}[Bu_0 + Cv_0 - P^{-1}\tilde{R}\hat{x}_1](t) + f_1(t) \quad (24)$$

Similarly, P2 may select its strategy in the following way:

$$\hat{x}_2 = \Phi_A(t,t_0)x(t_0) + K^{t,t_0}[Bu_0 + Cv_0 + P^{-1}\tilde{S}\hat{x}_2](t) - f_2(t) \quad (25)$$

where f_1 and f_2 are to be determined by the players.

It should be noted at this point that the proposed structures of $\hat{x}_i$ in (24) and (25) do not restrict the choice of $\hat{x}_i$ since f_is can be chosen freely. Using (24) for any interval $[t_0, t_1]$, we get:

$$J_d(\hat{x}_1, \hat{x}_2) = \mathbf{E}[\int_{t_0}^{t_1} \|K^{t,t_0}P^{-1}\tilde{S}\hat{x}_2 - f_1\|^2_{\tilde{R}}dt - \int_{t_0}^{t_1} \|K^{t,t_0}P^{-1}\tilde{R}\hat{x}_1 - f_2\|^2_{\tilde{S}}dt + \int_{t_0}^{t_1} \|K_1^{t,t_0}GW\|^2_{\tilde{R}-\tilde{S}}dt]. \quad (26)$$

The objective of P1 would be to choose an optimal f_1 to optimize (26) without having any knowledge about f_2 and the noise W_s.

$$J_d(\hat{x}_1, \hat{x}_2) \le J_d^1(\hat{x}_1, \hat{x}_2) = \mathbf{E}[\int_{t_0}^{t_1} \|K^{t,t_0}P^{-1}\tilde{S}\hat{x}_2 - f_1\|^2_{\tilde{R}}dt + \int_{t_0}^{t_1} \|K_1^{t,t_0}GW\|^2_{\tilde{R}}dt], \quad (27)$$

where the equality holds when P2 accesses the state at every time. P1 must minimize (27) in order to ensure that for any choice f_2 that P2 makes, the cost is always upper bounded by the r.h.s. of (27). Similar arguments from the point of view of P2 lead to the fact that f_2 must be chosen in order to maximize (28).

$$J_d^2(\hat{x}_1, \hat{x}_2) = -\mathbf{E}[\int_{t_0}^{t_1} \|K^{t,t_0}P^{-1}\tilde{R}\hat{x}_1 - f_2\|^2_{\tilde{S}}dt + \int_{t_0}^{t_1} \|K_1^{t,t_0}GW\|^2_{\tilde{S}}dt] \quad (28)$$

The second term in both equations (27) and (28) are constant and does not play any role in the solution of f_1 and f_2 that optimize the equations.

Claim 4.1 (Main Result): *For a given interval $[t_0, t_1]$, the optimal f_1 satisfies the following differential equation:*

$$\dot{f}_1 = Af_1 + P^{-1}\tilde{S}\xi \quad (29)$$
$$f_1(t_0) = 0$$
$$\dot{\xi} = (A + P^{-1}\tilde{S})\xi + Bu_0 + Cv_0 - P^{-1}\tilde{R}\hat{x}_1 \quad (30)$$
$$\xi(t_0) = x(t_0)$$

Proof: The proof of the above claim is presented in Appendix section. ■

Remark 4.2 (Main Result): *$\hat{x}_1$ satisfies the following system of equations for the time interval $[t_0, t_1]$:*

$$\dot{\hat{x}}_1 = (A - P^{-1}\tilde{R})\hat{x}_1 + Bu_0 + Cv_0 + P^{-1}\tilde{S}\xi \quad (31)$$
$$\dot{\xi} = (A + P^{-1}\tilde{S})\xi + Bu_0 + Cv_0 - P^{-1}\tilde{R}\hat{x}_1 \quad (32)$$
$$\hat{x}_1(t_0) = \xi(t_0) = x(t_0)$$

Comparing (20) and (31), ξ serves as the estimate of $\hat{x}_2$ for the evolution of $\hat{x}_1$. The difference is that the latest state information available to ξ is $x(t_0)$ whereas $\hat{x}_2$ may have an old or updated state information $x(t_0')$.

Similarly for P2, equivalent results to Claim 4.1 and Remark 4.2 are stated below.

Claim 4.3: *For a given interval $[t_0, t_1]$, the optimal f_2 satisfies the following differential equation:*

$$\dot{f}_2 = Af_1 - P^{-1}\tilde{R}\zeta \quad (33)$$
$$f_2(t_0) = 0$$
$$\dot{\zeta} = (A - P^{-1}\tilde{R})\zeta + Bu_0 + Cv_0 + P^{-1}\tilde{S}\hat{x}_2 \quad (34)$$
$$\zeta(t_0) = x(t_0)$$

Proof: The proof is similar to the proof of Claim 4.1. ■

Remark 4.4: *$\hat{x}_2$ satisfies the following system of equations for the time interval $[t_0, t_1]$:*

$$\dot{\hat{x}}_2 = (A + P^{-1}\tilde{S})\hat{x}_2 + Bu_0 + Cv_0 - P^{-1}\tilde{R}\zeta \quad (35)$$
$$\dot{\zeta} = (A - P^{-1}\tilde{R})\zeta + Bu_0 + Cv_0 + P^{-1}\tilde{S}\hat{x}_2 \quad (36)$$
$$\hat{x}_2(t_0) = \zeta(t_0) = x(t_0)$$

Comparing (20) and (35), ζ serves as the estimate of $\hat{x}_1$ for the evolution of $\hat{x}_2$. The difference is that the latest state information available to ζ is $x(t_0)$ whereas $\hat{x}_1$ may have an old or updated state information $x(t_0')$.

In this problem, none of the players has any idea about the information set of the other player and hence the structure of the dynamics of $\hat{x}_1$ and $\hat{x}_2$ are as given in (31) and (35). If any prior knowledge about the information set of the opponent were known, the dynamics (31) and (35) would possibly be different. Study of the later problem is beyond the scope of this paper and will be studied elsewhere.

From (27), the cost incurred by P1 having the information set $\mathcal{I}_1(T) = x(\tau_k^1)_{k=1}^{n_1(T)}$ is

$$\mathbf{E}[\int_0^T \|x - \hat{x}_1\|_{\tilde{R}}^2 dt] = \sum_{i=0}^{n_1(T)+1} \mathbf{E}[\int_{\tau_i^1}^{\tau_{i+1}^1} \|K^{t,\tau_i^1}[P^{-1}\tilde{S}\hat{x}_2](t) - f_1(t)\|_{\tilde{R}}^2 dt + \quad (37)$$
$$\int_{\tau_i^1}^{\tau_{i+1}^1} \|K_1^{t,\tau_i^1}[GW](t)\|_{\tilde{R}}^2 dt + \lambda_1 n_1(T)]$$

where $\tau_0^1 = 0$ and $\tau_{n_1(T)+1}^1 = T$.

In the optimal selection of $\mathcal{I}_i(T)$, the first term in the right hand side of (37) does not play any role since all the extractable information from $\hat{x}_2$ is captured in f_1 through the term ξ (Claim 4.1). Therefore the Information set $\mathcal{I}_1$ is formed such that $\sum_{i=0}^{n_1(T)+1} \mathbf{E}[\int_{\tau_i^1}^{\tau_{i+1}^1} \|K_1^{t,\tau_i^1}[GW](t)\|_{\tilde{R}}^2 dt + \lambda_1 n_1(T)]$ is minimized. Similar arguments reveal that the information set $\mathcal{I}_2$ for P2 is chosen such that $\sum_{i=0}^{n_2(T)+1} \mathbf{E}[\int_{\tau_i^2}^{\tau_{i+1}^2} \|K_1^{t,\tau_i^2}[GW](t)\|_{\tilde{S}}^2 dt + \lambda_2 n_2(T)]$ is minimized.

At this point we are ready to relax the assumption that we made at the beginning of this section that $\mathcal{I}_1$ and $\mathcal{I}_2$ are given. The following optimization problems should be solved by P1 and P2 in order to construct their information sets.

For P1:

$$\min_{n_1(T),\tau_1^1,\cdots,\tau_{n_1(T)}^1} \sum_{i=0}^{n_1(T)+1} \mathbf{E}\Big[\int_{\tau_i^1}^{\tau_{i+1}^1} \|K_1^{t,\tau_i^1}[GW](t)\|_{\tilde{R}}^2 dt + \lambda_1 n_1(T)\Big] \quad (38)$$

For P2:

$$\min_{n_2(T),\tau_1^2,\cdots,\tau_{n_2(T)}^2} \sum_{i=0}^{n_2(T)+1} \mathbf{E}\Big[\int_{\tau_i^2}^{\tau_{i+1}^2} \|K_1^{t,\tau_i^2}[GW](t)\|_{\tilde{S}}^2 dt + \lambda_2 n_2(T)\Big] \quad (39)$$

Remark 4.5: *The parameters of the information set i.e. the number of elements in the set ($n_i(T)$) and the instances of sampling the state (τ_k^i) can be determined offline by solving the above finite dimensional optimization problems.*

Remark 4.6: *The optimal choice of information set and the optimal strategy selection are two decoupled problems for each player.*

A. Characteristics of τ_k^1, τ_k^2

For a fixed $n_1(T) = N$, τ_k^1s are determined by minimizing

$$\sum_{i=0}^{N+1} \int_{\tau_i^1}^{\tau_{i+1}^1} \mathbf{E}\big[\|K_1^{t,\tau_i^1}[GW](t)\|_{\tilde{R}}^2 dt\big]$$

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

From the properties of the Wiener process,

$$\mathbf{E}[\|K_1^{t,\tau_i^1}[GW](t)\|_{\tilde{R}}^2] = \int_{\tau_i^1}^t tr(\|\Phi_A(t,s)G(s)\|_{\tilde{R}(s)}^2)ds \quad (40)$$

Thus, to select τ_k^1, the optimization is performed on the objective function:

$$H(\tau_1^1, \cdots, \tau_N^1) = \sum_{i=0}^{N+1} \int_{\tau_i^1}^{\tau_{i+1}^1} \int_{\tau_i^1}^t tr(\|\Phi_A(t,s)G(s)\|_{\tilde{R}(s)}^2)ds. \quad (41)$$

If the initial state is not given, the players need to have that information in the first place. The optimal choice of τ_i^{1*} should satisfy the following necessary conditions:

For all $i = 1, 2, \cdots, N$

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

Similarly for P2, the optimal choice of the τ_i^{2*}s needs to satisfy

$$\int_{\tau_{i-1}^{2*}}^{\tau_i^{2*}} tr(\|\Phi_A(\tau_i^{2*}, t)G(t)\|_{\tilde{S}(t)}^2)dt = \quad (43)$$
$$\int_{\tau_i^{2*}}^{\tau_{i+1}^{2*}} tr(\|\Phi_A(t, \tau_i^{2*})G(\tau_i^{2*})\|_{\tilde{S}(\tau_i^{2*})}^2)dt$$

Remark 4.7: *For a game where $BR^{-1}B' = CS^{-1}C'$, the optimal choices of τ_k^{i*} for both the players are the same.*

Remark 4.8: *For a deterministic game ($G \equiv 0$) with the Information structure considered here (i.e. no prior knowledge about the information set of the opponent), the results show that the players do not need any more information than $x(0)$. This fact might change if any of (or both) the players have some knowledge about the opponent's information set.*

Remark 4.9: *The set of state information acquisition times $\{\tau_k\}$ selected by the players is independent of the time instances of incurring the pointwise cost i.e. $\{t_i\}$.*

V. Conclusions

In this work a two-player stochastic Linear-Quadratic differential game is considered. The Nash equilibrium strategy for both the players under full state information is derived. With costly state information, the optimal instances for obtaining the state information for the players have been studied. It turns out that the optimization problems for the selection of the optimal strategy (u^* or v^*) and the optimal information acquisition instances (τ_k^{1*} or τ_k^{2*}) are two decoupled problems for each player. The optimal triggering instances τ_k^{i*} can be found offline.

This framework can easily be extended to related problems such as the cost function being linear-exponential-quadratic or quadratic with exponential forgetting factor.

This work does not assume any prior knowledge to the players about the structure of the information sets of their opponents. With some prior knowledge (may be in a probabilistic sense or so), the structure of the optimal strategies u^* and v^* will possibly be different from what derived here.

VI. Appendix

Proof of Claim 4.1.

For any $t \in [t_0, t_1]$, let the latest element in the Information set $\mathcal{I}_2(t)$ be $x(t')$ (generally, t' depends on t). Therefore from (25),

$$\hat{x}_2(t) = \int_{t'}^{t} \Phi_{A+P^{-1}\tilde{S}}(t,s)(B(s)u_0(s) + C(s)v_0(s) - \dot{f}_2 + Af_2)ds + \Phi_{A+P^{-1}\tilde{S}}(t,t')x(t') \quad (44)$$

Let us denote $\Phi_S = \Phi_{A+P^{-1}\tilde{S}}$ and replacing Φ_A with Φ_S in (22) and (23), we define two new operators K_S and K_{S1} respectively. From the dynamics of x, we can write:

$$x(t') = K_S(t',t_0)[Bu_0 + Cv_0 - P^{-1}\tilde{R}\hat{x}_1 - P^{-1}\tilde{S}e_2](t') + K_{S1}^{t',t_0}[GW](t) + \Phi_S(t',t_0)x(t_0) \quad (45)$$

where $e_2 = x - \hat{x}_2$. Substituting (45) in (44) and rearranging the terms,

$$\hat{x}_2(t) = \Phi_S(t,t_0)x(t_0) + K_S^{t,t_0}[Bu_0 + Cv_0 - P^{-1}\tilde{R}\hat{x}_1](t) - \int_{t_0}^{t'} \Phi_S(t,s)[P^{-1}(s)\tilde{S}(s)e_2(s))ds - G(s)dW(s)] + \int_{t'}^{t} \Phi_S(t,s)[P^{-1}(s)\tilde{R}(s)\hat{x}_1(s) - \dot{f}_2 + Af_2]ds \quad (46)$$

In (46), the last two integrals cannot be estimated by P1 since knowledges of $W(s)$, f_2 and t', which are associated with $\mathcal{I}_2(t)$, are required.

Thus, at any $t \in [t_0, t_1]$, the extractable information from $\hat{x}_2(t)$ is the amount $\Phi_S(t,t_0)x(t_0) + K_S^{t,t_0}[Bu_0 + Cv_0 - P^{-1}\tilde{R}\hat{x}_1](t)$.

Therefore, to optimize $\int_{t_0}^{t_1} \|K^{t,t_0}P^{-1}\tilde{S}\hat{x}_2 - f_1\|_{\tilde{R}}^2 dt$, the choice of f_1 should be of the form $f_1 = K^{t,t_0}P^{-1}\tilde{S}\xi$, where for $t \geq t_0$

$$\xi(t) = \Phi_S(t,t_0)x(t_0) + K_S^{t,t_0}[Bu_0 + Cv_0 - P^{-1}\tilde{R}\hat{x}_1](t) \quad (47)$$

These definitions of f_1 and ξ are equivalent to the dynamics proposed in Claim 4.1.

References

[1] R. Isaacs, "Differential games III," 1954.
[2] J. H. Case, "Equilibrium points of N-person differential games." DTIC Document, Tech. Rep., 1967.
[3] 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.
[4] ——, "Further properties of nonzero-sum differential games," *Journal of Optimization Theory and Applications*, vol. 3, no. 4, pp. 207–219, 1969.
[5] T. Basar and G. J. Olsder, *Dynamic noncooperative game theory*. SIAM, 1999, vol. 23.
[6] M. R. James and J. Baras, "Partially observed differential games, infinite-dimensional Hamilton-Jacobi-Isaacs equations, and nonlinear H_∞ control," *SIAM Journal on Control and Optimization*, vol. 34, no. 4, pp. 1342–1364, 1996.
[7] T. Başar and P. Bernhard, *H-infinity optimal control and related minimax design problems: a dynamic game approach*. Springer Science & Business Media, 2008.
[8] J. Engwerda, *LQ dynamic optimization and differential games*. John Wiley & Sons, 2005.
[9] 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.
[10] 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.
[11] 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.
[12] 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.
[13] 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.
[14] D. H. Jacobson, "Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games," *Automatic Control, IEEE Transactions on*, vol. 18, no. 2, pp. 124–131, 1973.
[15] P. Bernhard, "Linear-quadratic, two-person, zero-sum differential games: necessary and sufficient conditions," *Journal of Optimization Theory and Applications*, vol. 27, no. 1, pp. 51–69, 1979.
[16] M. Simaan and J. B. Cruz Jr, "On the Stackelberg strategy in nonzero-sum games," *Journal of Optimization Theory and Applications*, vol. 11, no. 5, pp. 533–555, 1973.
[17] T. Basar, "Lecture notes on non-cooperative game theory," *Game Theory Module of the Graduate Program in Network Mathematics*, 2010.
[18] ——, "A counterexample in linear-quadratic games: Existence of nonlinear Nash solutions," *Journal of Optimization Theory and Applications*, vol. 14, no. 4, pp. 425–430, 1974.
[19] T. E. Duncan and B. Pasik-Duncan, "Some stochastic differential games with state dependent noise," in *Decision and Control (CDC), 2015 IEEE 54th Annual Conference on*. IEEE, 2015, pp. 3773–3777.