49th IEEE Conference on Decision and Control
December 15-17, 2010
Hilton Atlanta Hotel, Atlanta, GA, USA

Optimal Control in the Presence of an Intelligent Jammer with Limited Actions

Abhishek Gupta Cédric Langbort Tamer Başar

Abstract— **We consider a dynamic zero-sum game between two players. The first player acts as a controller for a discrete time LTI plant, while the second player acts to jam the communication between the controller and the plant. The number of jamming actions is limited. We determine saddle-point equilibrium control and jamming strategies for this game under the full state, total recall information structure for both players, and show that the jammer acts according to a threshold policy at each decision step. Various properties of the threshold functions are derived and complemented by numerical simulation studies.**

I. Introduction

One of the main challenges associated with networked control systems is the presence of a communication channel between the controller and the plant. This channel introduces a number of limitations in the forward path of the control loop, which restricts the controller's ability to stabilize the system or achieve optimality in closed-loop.

Examples of such limitations include finite rate and channel capacity, stochastic packet drops and delays, and bounded signal-to-noise ratio. In addition, in cases where the controller receives no acknowledgment from the channel, and hence has no information as to whether or not the input has reached the plant, additional difficulties arise in the form of so-called non-standard information patterns (see, e.g., [1], [2] for recent special issues devoted to these issues). The effects of such communication channel-induced limitations on control have been intensively studied in the past decade. For example, a number of papers have considered the minimum channel capacity or rate necessary for stabilization (see, e.g. [3], [4]) or achieving optimal quadratic closed-loop performance [5], [6], [7].

In all the works mentioned above, the channel effects are treated as exogenous in the sense that the channel behavior is independent from the controller's actions and control task (for example, delays are typically not assumed to result from the rate at which the controller communicates with the plant). In particular, the degradation of the controller's output signal is not strategic or malicious. While this type of model is appropriate for problems of control over the Internet, it may not be so when the channel is a dedicated wireless link that is susceptible to jamming and/or in military scenarios where an adversary might be specifically interested in disrupting the communication between, say, a base-station and a remote vehicle.

In this paper, we consider the case of a strategic jammer, whose goal is to actively and optimally perturb the control process by using a finite number of jamming actions over a horizon of N time steps. This constraint on the number of jamming actions is similar to that introduced in [8], [9] in the case of optimal control (without an adversary), and in [10] in the case of estimation (again without an adversary). It is introduced in the present problem to capture the fact that, since jamming is a power intensive activity and available power on-board a jammer is typically limited, continuous action throughout the entire decision horizon is impossible.

Our formulation, detailed in Section II, naturally results in a dynamic zero-sum game between the jammer and the controller. We show that saddle-point equilibrium strategies exist and use dynamic programming to compute them. In particular, we show that the jammer saddle-point equilibrium strategy is threshold-based, which means that at every time step, the jammer jams if and only if the plant's state is larger than an off-line computable and time-varying threshold. We start by investigating a simple situation in Section III, in which the jammer can only act once over a 3-steps horizon. We derive the threshold functions analytically in this case. The case of general N is then treated in Section IV. Finally, in Section V, we indicate further directions for research on the problem of optimal control in the presence of a strategic jammer, both within and beyond the framework and models introduced in this paper.

II. Problem Formulation

The class of problems considered in this paper can be viewed as the standard discrete-time linear-quadratic-Gaussian (LQG) control problem with state feedback, but with one major difference: as a networked control system, the link connecting the output of the controller to the plant is unreliable due to the presence of adversarial jamming, with a possibility of the control signal being intercepted by the jammer and not reaching the plant. Instead of limiting the jammer's action through an energy constraint, we allow the jammer only M possibilities of interception on problem of horizon N, where $M < N$. We further assume that if the control signal is intercepted, the input to the plant is *zero*.[1]

Using scalar system dynamics, the scenario above can be captured through the following mathematical formulation:

All authors are with the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801; emails: {gupta54, langbort, basar1}@illinois.edu. Research was partially supported by AFOSR Grant FA9550-09-1-0249.

[1]It would be possible to adopt an alternate formulation where whenever the control signal is intercepted, the actuator generates an input that is based on the most recently received control signal, but this will not be pursued in this paper.

The state equation under adversarial jamming evolves as

$$x_{k+1} = Ax_k + \alpha_k u_k + w_k \,, \qquad k = 0, 1, ..., N-1 \,, \quad (1)$$

where $x_k \in \mathbb{R}$ is the state of the plant, $u_k \in \mathbb{R}$ is the control signal, $\{w_k\}$ is a discrete-time zero mean Gaussian white noise process with variance σ_w^2 (i.e. $w_k \sim \mathscr{N}(0, \sigma_w^2)$, and x_0 is also a zero mean Gaussian variable, with variance σ_0^2, and independent of $\{w_k\}$. The sequence $\{\alpha_k \in \{0, 1\}\}$ is the control of the jammer, where $\alpha_k = 0$ means that the jammer is active at time k, whereas $\alpha_k = 1$ means that the jammer is inactive and the control signal reaches the plant. The assumption that the jammer is allowed to intercept at most M times (in a horizon of N), is captured by the jammer constraint $\sum_{k=0}^{N-1}(1-\alpha_k) = M$. Note that here we actually use an equality rather than an inequality because, as will become clear later in our analysis, there is no incentive for the jammer not to use all M allotments for interception, since it does not incur any cost during each jamming instance.

The cost function associated with this problem is

$$J = E\left\{ \sum_{k=0}^{N-1} (x_k^2 + \alpha_k u_k^2) + x_N^2 \right\} \qquad (2)$$

which is to be minimized by the controller and maximized by the jammer. Note that when the control signal is intercepted (that is, $\alpha_k = 0$), the controller accrues no cost for control.

This is clearly a zero-sum dynamic game, but to make the problem precise we have to specify the underlying information structure, and the equilibrium solution concept to be adopted. Toward this end, let $x_{[0,k]} := \{x_0, \dots, x_k\}$,with a similar definition applying to $\alpha_{[0,k]}$, and let us introduce

$$I_0 := \{x_0\}\,, \;\; I_k := \{x_{[0,k]}, \alpha_{[0,k-1]}\} \;\text{ for } k \geq 1$$

as the information available to both the controller and the jammer at time k. We introduce control policies (strategies) for the controller and the jammer as measurable mappings, $\{\gamma_k\}$ and $\{\mu_k\}$, respectively, from their information sets (which are the same for both) to their action sets; more precisely, $u_k = \gamma_k(I_k)$ and $\alpha_k = \mu_k(I_k)$, where

$$\gamma_k : \mathbb{R}^{k+1} \times \{0,1\}^k \to \mathbb{R} \text{ and } \mu_k : \mathbb{R}^{k+1} \times \{0,1\}^k \to \{0,1\}\,.$$

We further restrict $\mu := \{\mu_0, \dots, \mu_{N-1}\}$ to those maps that satisfy the jammer constraint, with $\alpha_k = \mu_k(I_k)$; let us denote the class of all such policies for the jammer by $\mathcal{M}$ and for controller by Γ. At each point in time, the controller has access to the current value of the state, recalls the past values, and also has full memory on whether any of the previous control signal transmissions were intercepted or not. This latter information could be made available to the controller through acknowledgement messages sent from the plant, as in TCP of the Internet. Likewise, the jammer has access to full state information, and recalls its past actions. There could, of course, be various variations of this information structure. Given the information structure introduced above, and the feasible policies of the controller and the jammer, we rewrite the cost function as $J(\gamma, \mu)$, in terms of the policies γ and μ, and seek a pair $(\gamma^* \in \Gamma, \mu^* \in \mathcal{M})$ with the property:

$$J(\gamma^*, \mu) \leq J(\gamma^*, \mu^*) \leq J(\gamma, \mu^*) \quad \forall \gamma \in \Gamma, \mu \in \mathcal{M}\,.$$

This is a saddle-point solution for the underlying game, where the controller is the minimizer and the jammer the maximizer, and the order in which they determine their policies is immaterial (that is, the upper and lower values are equal). Of course, this has not been established as yet, and one of the goals of the paper is to show that this is indeed the case, and also to obtain the saddle-point solution.

When $M = 0$, this is precisely the standard LQG problem with perfect state measurements, and for $M = N$, the controller signal is always intercepted and hence any pair of the form $(\gamma, 0)$ with $\gamma \in \Gamma$ is trivially a saddle-point solution; it is the *intermediate* case that is of interest.

A. Extended game and solution approach

In order to establish the existence of and compute saddle-point equilibrium strategies, it is easiest to extend the game's state space so as to keep track of the jammer's options at a particular time step, and redefine the dynamics on this state space. An extented state of the dynamic zero-sum game defined by cost function J, information sets $\{I_k\}$ and evolution equation (1) is a triple $(x, s, t) \in \mathcal{E} := \mathbb{R} \times \{0, ..., M\} \times \{0, ..., N\}$, where x can be thought of as the state of the controlled plant, $t = N - k$ can be thought of as the number of remaining decision steps, and s can be thought of as the number of remaining jamming instances available to the jammer. We will also say that "x is the state of the plant at stage (s, t)" and sometimes write $x_{(s,t)}$ to denote this. We will denote the jammer's action space at stage (s, t) by $\mathcal{A}_{(s,t)} \subset \{0, 1\}$. From an extended state $(x, s, t) \in \mathcal{E}$ such that $\mathcal{A}_{(s,t)} = \{0, 1\}$, the system can transition to two extended states, depending on the jammer's and controller's action at that state: $(Ax + u + w, s, t - 1)$ or $(Ax + w, s - 1, t - 1)$. The first state is reached when the controller is applying input u and the jammer is inactive ($\alpha = 1$), while the second is reached when the jammer is active ($\alpha = 0$), *regardless of the controller's action*. When $\mathcal{A}_{(s,t)}$ is a strict subset of $\{0, 1\}$, only one of those two transitions is possible. The projection of the extended state space onto the $(s, t)-$ space thus has the structure of the graph of Figure 1.

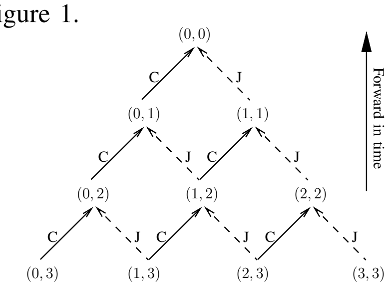

Fig. 1. A portion of the extended state space. Here 'J' denotes that the jammer is active in that stage and 'C' denotes that the jammer is idle (and control signal is received by the plant). Depending on the value of M, some of the depicted transitions may not be possible.

The original zero-sum dynamic game introduced previously naturally induces a zero-sum dynamic game on the extended state space $\mathcal{E}$ by keeping the same cost function J as in (2) and using the extended state transition rule defined above. A controller *feedback* policy on $\mathcal{E}$ is a map $\tilde{\gamma} : \mathcal{E} \to \mathbb{R}$ and, likewise, a jammer feedback policy on $\mathcal{E}$ is a map $\tilde{\mu} : \mathcal{E} \to \{0,1\}$. Given a controller policy $\tilde{\gamma}$ on $\mathcal{E}$, we can define a feasible policy $\gamma \in \Gamma$ for the original game by

$$\begin{aligned}&\gamma_k(x_{[0,k]}, \alpha_{[0,k-1]}) := \\ &\tilde{\gamma}(x_k, M - card\{i \in [0, k-1] \,|\alpha_i = 0\,\}, N-k)\end{aligned}$$

for all k. Similarly, we can associate a jammer policy $\mu \in \mathcal{M}$ to a feedback jammer policy on $\mathcal{E}$. As a result, if the zero-sum game defined on the extended state space has a saddle-point equilibrium *in feedback strategies*, the original zero-sum game admits a saddle-point equilibrium $(\gamma^* \in \Gamma, \mu^* \in \mathcal{M})$. Note, however, that the converse may not be true, as a feasible strategy γ for the original game does not always uniquely correspond to a feedback strategy $\tilde{\gamma}$ on $\mathcal{E}$ (since, e.g., some γ_k could depend on the exact jamming sequence $\alpha_{[0,k-1]}$ instead of just the number of jamming events so far). As a first approach to the problem of control in the presence of an intelligent jammer, we focus exclusively on saddle-point equilibrium strategies corresponding to feedback strategies defined on $\mathcal{E}$ in this paper. A more complete characterization of all saddle-point strategies in $\Gamma \times \mathcal{M}$ will be the subject of future work.

A straightforward generalization of Corollary 6.2 on page 282 of [11], establishes that strategies $\tilde{\gamma}^*$ and $\tilde{\mu}^*$ are feedback saddle-point equilibrium strategies defined on $\mathcal{E}$ if and only if, for all $(s,t) \in \{0,...,M\} \times \{0,...,N\}$ there exist functions $V_{(s,t)} : \mathbb{R} \to \mathbb{R}$ such that the following recursive equations hold for all $x \in \mathbb{R}$:

$$\begin{aligned}&V_{(0,0)}(x) = x^2,\\ &V_{(s,t)}(x) = \inf_u \max_{\alpha \in \mathcal{A}_{(s,t)}} \big(E\{x^2 + \alpha u^2 + \\ &\qquad\qquad V_{(s^+(\alpha),t-1)}(Ax + \alpha u + w)\}\big). \qquad (3)\end{aligned}$$

In (3), we have let $s^+(\alpha) = \begin{cases} s & \text{if } \alpha = 1 \\ s-1 & \text{if } \alpha = 0 \end{cases}$.

In the remaining sections of this paper, we explicitly compute such functions $V_{(s,t)}$, thus effectively and constructively proving the existence of feedback saddle-point equilibrium strategies defined on $\mathcal{E}$, and, in turn, of saddle-point equilibrium strategies in $\Gamma \times \mathcal{M}$ for the original game. At this point, it is worth emphasizing that the equality between inf-max and max-inf does indeed hold in (3), i.e., that the game has a value. This follows directly from the fact that the function $u \mapsto E\{x^2 + V_{(s-1,t-1)}(Ax + w)\}$ (which appears in the right hand-side of (3) when $\alpha = 0$) is a constant and the following lemma.

Lemma 1: Let f be a function and M be a constant. Then $\inf_u \max(f(u), M) = \max(\inf_u f(u), M)$.

Proof: Let $\mathcal{U} := \{u | f(u) < M\}$. Then we have two cases: $\mathcal{U} = \emptyset$ and $\mathcal{U} \neq \emptyset$. When $\mathcal{U} = \emptyset$,

$$f(u) \geq M \text{ for all } u. \qquad (4)$$

hence $\max(f(u), M) = f(u)$ for all u and $\inf_u \max(f(u), M) = \inf_u f(u)$. Besides, inequality (4) also implies that $\inf_u f(u) \geq M$, so that $\max(\inf_u f(u), M) = \inf_u f(u)$. Now, if $\mathcal{U} \neq \emptyset$, then $\inf_u f(u) < M$ and $\max(\inf_u f(u), M) = M$. On the other hand, by definition, $\max(f(u), M) \geq M$ for all u and, since $\mathcal{U} \neq \emptyset$, there exists u_0 such that $\max(f(u_0), M) = M$. Hence, $\inf_u \max(f(u), M) = M$. ■

B. Notation

We now introduce some notations. We denote the jammer's and controller's best response costs at stage (s,t), respectively as

$$\begin{aligned}J_{(s,t)}(x,u,\alpha) :=\;& E\{x^2 + \alpha u^2 + \\ & V_{(s^+(\alpha),t-1)}(Ax + \alpha u + w)\},\\ J^J_{(s,t)}(x) :=\;& x^2 + E\{V_{(s-1,t-1)}(Ax + w)\},\\ J^C_{(s,t)}(x) :=\;& \inf_u E\{x^2 + u^2 + V_{(s,t-1)}(Ax + u + w)\}.\end{aligned}$$

With these notations, feedback saddle-point equilibrium strategies defined on $\mathcal{E}$ are characterized by the fact that, when the plant state is x at stage (s,t), the controller's action minimizes $J_{(s,t)}(x,u,\alpha)$ over u, while the jammer is choosing the action corresponding to the largest of the two costs between $J^C_{(s,t)}(x)$ and $J^J_{(s,t)}(x)$ when $\mathcal{A}_{(s,t)} = \{0,1\}$. As we will see, this results in a threshold-based policy in which the action of the jammer at (s,t) depends on the sign of the quantity $|x| - \tau_{(s,t)}(x)$ for an off-line computable threshold function $\tau_{(s,t)}(x)$.

Another object that we will make frequent use of in the subsequent sections is the conditional probability density function of the state at a given stage. When a transition from stage (s,t) to stage (s',t') is possible in Figure 1, and control action u is applied at stage (s,t), we denote this conditional probability density function of the state $x_{(s',t')}$ given the state $x_{(s,t)}$ and u by $f(x_{(s',t')}|x_{(s,t)}, u)$. If the jammer was inactive during the stage (s,t), then $s' = s$, $t' = t-1$, and $x_{(s',t')} = Ax_{(s,t)} + u + w_k$. Since the noise w_k is an i.i.d. Gaussian random variable, the conditional probability density function follows a normal distribution, given by

$$f\left(x_{(s,t-1)}|x_{(s,t)}, u\right) = \mathscr{N}\left(Ax_{(s,t)} + u, \sigma_w^2\right). \qquad (5)$$

If the jammer is active at stage (s,t), $s' = s-1$, $t' = t-1$, and $x_{(s',t')} = Ax_{(s,t)} + w_k$ so that the conditional probability density function is

$$f\left(x_{(s-1,t-1)}|x_{(s,t)}, u\right) = \mathscr{N}\left(Ax_{(s,t)}, \sigma_w^2\right). \qquad (6)$$

Note that it does not depend on control action u in this case.

III. The $M = 1, N = 3$ case

In order to illustrate the main steps of our derivations while keeping notation to a minimum, we start by computing feedback saddle-point equilibrium strategies $(\tilde{\gamma}^*, \tilde{\mu}^*)$ for the extended game in the simple case where $N = 3$ and $M = 1$ (i.e., the jammer can only jam once in three time steps). By definition, $V_{(0,0)}(x) = x^2$. At the next step, we can be in either of the two stages $(0,1)$ and $(1,1)$, depending upon whether the jammer was active in the last decision period or not (see Figure 1). At stage $(0,1)$, the jammer has no chance left to jam and his action space is reduced to $\mathcal{A}_{(0,1)} = \{1\}$. The jammer best response cost is thus

$$J_{(0,1)}(x,u) = E\{(Ax+u+w_2)^2 + x^2 + u^2\}. \quad (7)$$

Using the first order necessary condition for optimality, we find that the optimal control action $\tilde{\gamma}^*(x,0,1)$ satisfies

$$\frac{\partial J_{(0,1)}}{\partial u} = 2(Ax + \tilde{\gamma}^*(x,0,1)) + 2\tilde{\gamma}^*(x,0,1) = 0,$$

i.e., $\tilde{\gamma}^*(x,0,1) = -\frac{A}{2}x$. The value function at this stage is

$$V_{(0,1)}(x) = \left(1 + \frac{A^2}{2}\right)x^2 + \sigma_w^2. \quad (8)$$

In stage $(1,1)$, the jammer must always jam, otherwise the jammer constraint is violated. The value function at $(1,1)$ is

$$V_{(1,1)}(x) = J^J_{(1,1)}(x) = (1+A^2)x^2 + \sigma_w^2. \quad (9)$$

Let us now move on to stages $(0,2)$ and $(1,2)$. Note that the noise w_1 in these stages is independent from the noise w_2 occurring in the next stage. Applying the same approach as above, we find that the optimal control for stage $(0,2)$ is

$$\tilde{\gamma}^*(x,0,2) = -A\left(\frac{1+\frac{A^2}{2}}{2+\frac{A^2}{2}}\right)x \quad (10)$$

and that the corresponding value function is given by

$$V_{(0,2)}(x) = \left(1 + A^2 - \frac{2A^2}{4+A^2}\right)x^2 + \left(2 + \frac{A^2}{2}\right)\sigma_w^2. \quad (11)$$

Define $\kappa^{1,C}_{(0,2)} = \left(1 + A^2 - \frac{2A^2}{4+A^2}\right)$ and $\kappa^{2,C}_{(0,2)} = \left(2 + \frac{A^2}{2}\right)$. The case of stage $(1,2)$ requires more effort since the jammer has two options, i.e., $\mathcal{A}_{(1,2)} = \{0,1\}$. The controller's best response costs are found to be

$$J^J_{(1,2)}(x) = \kappa^{1,J}_{(1,2)}x^2 + \kappa^{2,J}_{(1,2)}\sigma_w^2$$
$$J^C_{(1,2)}(x) = \kappa^{1,C}_{(1,2)}x^2 + \kappa^{2,C}_{(1,2)}\sigma_w^2,$$

where $\kappa^{1,J}_{(1,2)} = 1 + A^2\left(1 + \frac{A^2}{2}\right), \quad \kappa^{2,J}_{(1,2)} = 2 + \frac{A^2}{2}$

$$\kappa^{1,C}_{(1,2)} = 1 + A^2 - \frac{A^2}{2+A^2}, \quad \kappa^{2,C}_{(1,2)} = 2 + A^2$$

If the difference between these costs $J^J_{(1,2)}(x) - J^C_{(1,2)}(x) \geq 0$, then the jammer must jam. Hence, the value function is

$$V_{(1,2)}(x) = \begin{cases} J^J_{(1,2)}(x) & \text{if } |x| \geq \tau_{(1,2)} \\ J^C_{(1,2)}(x) & \text{if } |x| < \tau_{(1,2)} \end{cases} \quad (12)$$

where we defined $\tau_{(1,2)} := \sqrt{\left(\frac{2+A^2}{A^4+2A^2+2}\right)}\sigma_w$. The feedback saddle-point equilibrium strategies $(\tilde{\gamma}^*, \tilde{\mu}^*)$ is

$$\tilde{\mu}^*(x,1,2) = \begin{cases} 0 & \text{if } |x| \geq \tau_{(1,2)} \\ 1 & \text{if } |x| < \tau_{(1,2)} \end{cases}, \quad (13)$$

and $\tilde{\gamma}^*(x,1,2) = -A\left(\frac{1+A^2}{2+A^2}\right)x \quad \forall\, x. \quad (14)$

Let us now consider stage $(1,3)$, the initial stage. The controller's cost if the jammer decides to jam at this stage is

$$J^J_{(1,3)}(x) = \kappa^{1,J}_{(1,3)}x^2 + \kappa^{2,J}_{(1,3)}\sigma_w^2$$

where $\kappa^{1,J}_{(1,3)} = \left(1 + A^2\kappa^{1,C}_{(0,2)}\right)$, $\kappa^{2,J}_{(1,3)} = \kappa^{1,C}_{(0,2)} + \kappa^{2,C}_{(0,2)}$

To compute the controller's best response cost when the jammer is idle, $J^C_{(1,3)}$, we need to calculate $E(V_{(1,2)}(x_1))$, where $x_1 = Ax + u + w_0$ for a given controller action u. According to (12), and recalling the definition of $f(.|.)$ introduced in Section II-B, we see that

$$E(V_{(1,2)}(x_1)) = \int_{|x_1| \geq \tau_{1,2}} f(x_1|x,u)J^J_{(1,2)}(x_1)dx_1 + \int_{|x_1| < \tau_{1,2}} f(x_1|x,u)J^C_{(1,2)}(x_1)dx_1. \quad (15)$$

Let us introduce $P_{(1,3)}(x,u,\tau_{(1,2)})$ as the conditional probability that $|x_{(1,2)}|$ lies above the threshold $\tau_{1,2}$, given that $x_{(1,3)} = x$ and the control action at stage $(1,3)$ is u,

$$P_{(1,3)}(x,u) = \int_{|x_1| \geq \tau_{(1,2)}} f(x_1|x,u)dx_1. \quad (16)$$

Let us also write $\overline{P}_{(1,3)}(x,u) = 1 - P_{(1,3)}(x,u)$ for the conditional probability that $|x_{(1,2)}| < \tau_{(1,2)}$, and introduce the following two second moments of x_1

$$R_{(1,3)}(x,u) = \frac{\int_{|x_1| \geq \tau_{(1,2)}} x_1^2 f(x_1|x,u)dx_1}{(Ax+u)^2 + \sigma_w^2} \quad (17)$$

and $\overline{R}_{(1,3)}(x,u) := 1 - R_{(1,3)}(x,u)$. The cost at stage $(1,3)$ with control is

$$J^C_{(1,3)}(x) = \inf_u \left[x^2 + u^2 + E(V_{(1,2)}(Ax+u+w_0))\right]. \quad (18)$$

The infimum in (18) is attained and the corresponding minimum point is $\tilde{\gamma}^*(x,1,3)$. To compute this minimum[2],

$$0 = H(x) = \frac{\partial}{\partial u}\left[x^2 + u^2 + E(V_{(1,2)}(Ax+u+w_0))\right] \quad (19)$$

which gives an implicit equation characterizing $\tilde{\gamma}^*(x,1,3)$. Now, letting $L_{(1,3)}(x) := -\tilde{\gamma}^*(x,1,3)/(Ax)$ and plugging the obtained value of $\tilde{\gamma}^*(x,1,3)$ back into (18), yields

$$J^C_{(1,3)}(x) = \kappa^{1,C}_{(1,3)}(x)x^2 + \kappa^{2,C}_{(1,3)}(x)\sigma_w^2 \quad (20)$$

where $\kappa^{1,C}_{(1,3)}(x)$ and $\kappa^{2,C}_{(1,3)}(x)$ is obtained by putting $u = -L_{(1,3)}(x)Ax$ in (18). Once both functions $J^J_{(1,3)}$ and $J^C_{(1,3)}$ have been determined, the value function at stage $(1,3)$ is

$$V_{(1,3)}(x) = \begin{cases} J^J_{(1,3)}(x) & \text{if } |x| \geq \tau_{(1,3)}(x) \\ J^C_{(1,3)}(x) & \text{if } |x| < \tau_{(1,3)}(x) \end{cases}, \quad (21)$$

where the threshold function $\tau_{(1,3)}(x)$ is defined such that $J^J_{(1,3)}(x) - J^C_{(1,3)}(x) \geq 0$ if and only if $|x| \geq \tau_{(1,3)}(x)$. Analytically, we find that

$$\tau_{(1,3)}(x) = \sqrt{\frac{\kappa^{2,C}_{(1,3)}(x) - \kappa^{2,J}_{(1,3)}}{\kappa^{1,J}_{(1,3)} - \kappa^{1,C}_{(1,3)}(x)}}\sigma_w. \quad (22)$$

[2]The second order condition for infimum is difficult to evaluate at the optimal control $\tilde{\gamma}^*(x,1,3)$ analytically. However, in all our simulations, we observed that $\tilde{\gamma}^*(x,1,3)$ is the unique minimum.

Note that, unlike $\tau_{(1,2)}$, threshold function $\tau_{1,3}$ is not constant, and that its computation requires determining $\tilde{\gamma}(.,1,3)$. Also note that $\kappa^{1,C}_{(1,3)}$ and $\kappa^{2,C}_{(1,3)}$ are even functions, i.e., that $\kappa^{1,C}_{(1,3)}(-x) = \kappa^{1,C}_{(1,3)}(x)$ and $\kappa^{2,C}_{(1,3)}(-x) = \kappa^{2,C}_{(1,3)}(x)$. This is because $P_{(1,3)}(-x,-u) = P_{(1,3)}(x,u)$ and the same property holds for $R_{(1,3)}(x,u)$. As a result, $\tau_{(1,3)}(x)$ is even.

We are now in a position to prove two results, which give us an insight into the nature of threshold function $\tau_{(1,3)}$.

Proposition 1: $L_{(1,3)}(x) \neq 1 \; \forall \, x \neq 0$

Proof: The proof is by contradiction. If $L_{(1,3)}(x) = 1$, then the right hand-side of (19) vanishes when $u = -Ax$. But, the derivative of $R_{(1,3)}$ and $P_{(1,3)}$ with respect to u is zero at $u = -Ax$. Using this relation in (19) for $x \neq 0$, we get

$$H(x) = 2u = -2Ax \neq 0$$

Therefore, $\tilde{\gamma}^*(x,1,3) \neq -Ax$ and $L_{(1,3)}(x) \neq 1$ for all x. ■

Proposition 2: $\lim_{|x|\to\infty} \tau_{(1,3)}(x)$ exists and is finite.

Proof: From Lemma 1, we know that $L_{(1,3)}(x) \neq 1$ for all $x \neq 0$. Therefore, $\lim_{|x|\to+\infty} |Ax + \tilde{\gamma}(x,1,3)| = +\infty$ also holds. Taking the limit as $|x| \to \infty$ in (16)-(17), we obtain

$$\lim_{|x|\to\infty} P_{(1,3)}(x,u) = 1, \qquad \lim_{|x|\to\infty} R_{(1,3)}(x,u) = 1,$$

Also, the derivative terms in (19) vanish as $|x| \to \infty$. Taking $\lim_{|x|\to\infty} \frac{1}{Ax} H(x)$ in (19) and rearranging, gives $\lim_{|x|\to\infty} L_{(1,3)}(x) = \frac{\kappa^{1,J}_{(1,2)}}{1+\kappa^{1,J}_{(1,2)}}$. Substituting this in (22),

$$\lim_{|x|\to\infty} \tau_{(1,3)}(x) = \sqrt{\frac{\kappa^{1,J}_{(1,2)} + \kappa^{2,J}_{(1,2)} - \kappa^{2,J}_{(1,3)}}{\kappa^{1,J}_{(1,3)} - \left(1 + A^2 \frac{\kappa^{1,J}_{(1,2)}}{1+\kappa^{1,J}_{(1,2)}}\right)}}\,\sigma_w \tag{23}$$

which proves the proposition. ■

Figure 2 shows the graph of the threshold function $\tau_{(1,3)}(x)$. As predicted by Lemma 2, we observe that the threshold $\tau_{(1,3)}(x)$ reaches the limiting value given by (23) when the state value x is sufficiently large. Also notice that when the state is sufficiently large, the value of $|x| - \tau_{(1,3)}(x)$ is greater than 0, and it is beneficial for jammer to jam in this region.

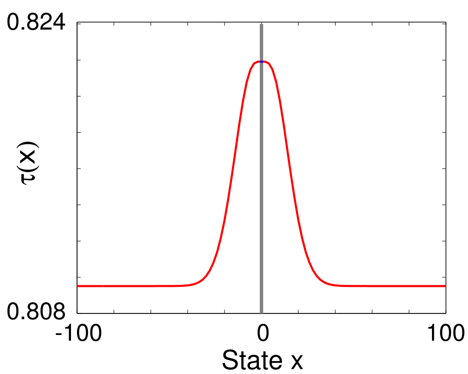

Fig. 2. Graph of function $\tau_{(1,3)}(x)$ for $A = 2.5$ and $\sigma_w = 1$. In the dark-colored narrow strip, absolute value of state $|x|$ is less than the threshold $\tau_{(1,3)}(x)$, while the reversed inequality holds in the white region. The jammer is active at stage $(1,3)$ if x belongs to the white region.

IV. General Case

Building on the intuition drawn from the results of Section III, we can prove the following theorem regarding the existence and characterization of feedback saddle-point equilibrium strategies defined on $\mathcal{E}$ in the case where $M = 1$ and N is arbitrary. The result is proved by induction on t. While the proof is omitted for reasons of space, the steps and rationale are identical to those used in Section III.

Theorem 1: Let $M = 1$ and $N > 1$. Let coefficients be defined according to the following recursion:

$$\kappa^{1,C}_{(0,0)} = 1, \; \kappa^{2,C}_{(0,0)} = 0,$$

$$\kappa^{1,C}_{(0,t)} = 1 + A^2 \frac{\kappa^{1,C}_{(0,t-1)}}{1+\kappa^{1,C}_{(0,t-1)}}; \; \kappa^{2,C}_{(0,t)} = \kappa^{1,C}_{(0,t-1)} + \kappa^{2,C}_{(0,t-1)},$$

$$\kappa^{1,J}_{(1,t)} = 1 + A^2 \kappa^{1,C}_{(0,t-1)}; \qquad \kappa^{2,J}_{(1,t)} = \kappa^{1,C}_{(0,t-1)} + \kappa^{2,C}_{(0,t-1)},$$

$$\kappa^{1,C}_{(1,t)}(x) = 1 + A^2 \Big(L^2_{(1,t)}(x) + (1 - L_{(1,t)}(x))^2 \psi^1_{(1,t)}(x, \tilde{\gamma}^*(x,1,t))\Big)$$

$$\kappa^{2,C}_{(1,t)}(x) = \psi^1_{(1,t)}(x, \tilde{\gamma}^*(x,1,t)) + \psi^2_{(1,t)}(x, \tilde{\gamma}^*(x,1,t)),$$

for all $t \geq 1$ and all x, where, the set $\mathcal{X}_{(1,t)}$ is defined as

$$\mathcal{X}_{(1,t)} = \left\{ x_{(1,t)} \in \mathbb{R} : x^2_{(1,t)} - \tau^2_{(1,t)}(x_{(1,t)}) \geq 0 \right\},$$

the threshold $\tau_{(1,t)}(x_{(1,t)})$ and $\psi(x,u)$'s are defined as

$$\tau_{(1,t)}(x_{(1,t)}) = \sqrt{\frac{\kappa^{2,C}_{(1,t)}(x_{(1,t)}) - \kappa^{2,J}_{(1,t)}}{\kappa^{1,J}_{(1,t)} - \kappa^{1,C}_{(1,t)}(x_{(1,t)})}}\,\sigma_w,$$

$$\psi^1_{(1,t)}(x,u) = \int_{\mathcal{X}^c_{(1,t-1)}} \frac{\kappa^{1,C}_{(1,t-1)}(\bar{x})\bar{x}^2}{(Ax+u)^2 + \sigma^2_w} f(\bar{x}|x,u)\,d\bar{x} + R_{(1,t)}(x,u)\kappa^{1,J}_{(1,t-1)}$$

$$\psi^2_{(1,t)}(x,u) = \int_{\mathcal{X}^c_{(1,t-1)}} \kappa^{2,C}_{(1,t-1)}(\bar{x}) f(\bar{x}|x,u)\,d\bar{x} + P_{(1,t)}(x,u)\kappa^{2,J}_{(1,t-1)},$$

conditioned probability and second moment defined as

$$P_{(1,t)}(x_{(1,t)}, u_{(1,t)}) = \Pr\{x_{(1,t-1)} \in \mathcal{X}_{(1,t-1)} | x_{(1,t)}, u_{(1,t)}\}$$

$$R_{(1,t)}(x_{(1,t)}, u_{(1,t)}) = \frac{\int_{\mathcal{X}_{(1,t-1)}} x^2 f(x|x_{(1,t)}, u_{(1,t)})\,dx}{(Ax_{(1,t)} + u_{(1,t)})^2 + \sigma^2_w}$$

and optimal control $\tilde{\gamma}^*(x,1,t)$ is

$$\tilde{\gamma}^*(x,1,t) = \arg\inf_u \Big[x^2 + u^2 + (Ax+u)^2 \psi^1_{(1,t)}(x,u) + \sigma^2_w \left(\psi^1_{(1,t)}(x,u) + \psi^2_{(1,t)}(x,u)\right)\Big].$$

Then, the strategies $(\tilde{\gamma}^*, \tilde{\mu}^*)$ given below are feedback saddle-point equilibrium strategies defined on $\mathcal{E}$:

$$\tilde{\gamma}^*(x,0,t) = -\left(\frac{A\kappa^{1,C}_{(0,t-1)}}{1+\kappa^{1,C}_{(0,t-1)}}\right)x; \; \tilde{\mu}^*(x,0,t) = 1 \; \forall \, t, x$$

$$\tilde{\mu}^*(x,1,t) = \begin{cases} 0 & \text{if } x \in \mathcal{X}_{(1,t)} \\ 1 & \text{if } x \in \mathcal{X}^c_{(1,t)} \end{cases}$$

and $\tilde{\gamma}^*(x, 1, t)$ as obtained above.

Figure 3 shows the set $\mathcal{X}_{(1,t)}$ for t ranging from 2 to 12 for an unstable system with $A = 2.5$ and $\sigma_w = 1$.

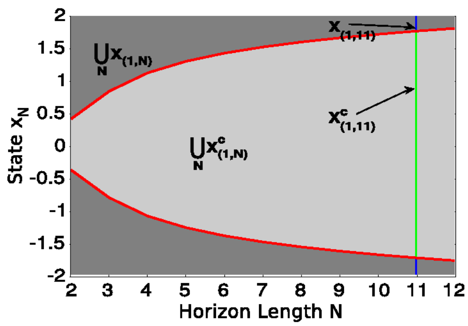

Fig. 3. Region showing the union of the sets in which the jammer jams and does not jam as a function of horizon length N. At a fixed integer N, the darker region indicates the set $\mathcal{X}_{(1,N)}$ and lighter region indicates the set $\mathcal{X}^C_{(1,N)}$. The bold line in the graph is the boundary of the set $\mathcal{X}_{(1,N)}$. The vertical line extends to ∞ above and $-\infty$ below.

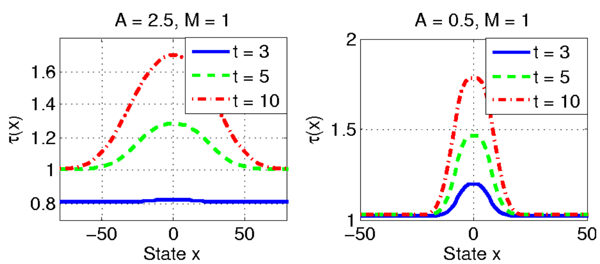

Fig. 4. Variations in $\tau_{(1,t)}(x_{(1,t)})$ as a function of state $x_{(1,t)}$ for an unstable system with $A = 2.5$, $\sigma_w = 1$ and for a stable system with $A = 0.5$, $\sigma_w = 1$ for $t = 3, 5, 10$.

The two sets of graph of Figure 4 show the values of $\tau_{(1,t)}$ for various values of t and $x_{(1,t)}$. It should be noted that for an N-stage problem, the threshold function $\tau_{(1,t)}(x)$, $t < N$ is the same as $\tau_{(1,t)}(x)$ for a t-stage problem. As can be observed from these curves, and similarly to the $M = 1, N = 3$ case, these threshold functions have a limit as the state goes to infinity. The following proposition can be proved, in complete analogy to Proposition 2. The proof is omitted for reasons of space.

Proposition 3: $\lim_{|x|\to\infty} \tau_{(1,t)}(x)$ exists and is equal to

$$\sqrt{\frac{\kappa^{1,J}_{(1,t-1)} + \kappa^{2,J}_{(1,t-1)} - \kappa^{2,J}_{(1,t)}}{\kappa^{1,J}_{(1,t)} - \left(1 + A^2 \frac{\kappa^{1,J}_{(1,t-1)}}{1+\kappa^{1,J}_{(1,t-1)}}\right)}}\sigma_w.$$

Expanding the present analysis to $M \geq 2$ seems to be quite challenging. It is unlikely that one can find closed form solution for the thresholds in this case. However, the results obtained for the $M = 1$ case can be extended to the $M \geq 2$ case. It is seen that the jammer's policy is to not jam when the state is small, and to jam when the state is large. This result holds true for $M \geq 2$ case too. For intermediate values of state, the accurate value of threshold is required.

In multi-dimensional state space, the regions of jamming and not jamming is separated by hyperplanes, and their analysis involves integrating the Gaussian distribution in a multi-dimensional space. This makes the computation and analysis of jammer's policy even more difficult than the scalar case. Extending the result from single dimension to multi-dimensional state space, the jammer must jam if the state is large, and not jam if the state is close to zero.

V. Conclusions and Future Work

We have considered the problem of optimal control of a scalar discrete time LTI system in the presence of a strategic but action-limited jammer potentially disrupting the communication between the controller and the plant. This led to a zero-sum dynamic game for which we established the existence of saddle-point equilibrium strategies. In the case where the jammer can only act once over the decision horizon, we proved that its strategy is threshold-based, and characterized the behavior of the threshold in the large state limit. Since the analyses for general cases are difficult, efficient computational methods will have to be employed to compute the approximate policy. One possible approach is to use rolling horizon control, such that the total number of the instances when the jammer acts in the entire horizon is still limited by M. One can also switch to efficient computational methods to compute or approximate the set $\mathcal{X}_{(s,t)}$, possibly using ideas from approximate dynamic programming.

References

[1] P. Antsaklis and J. Baillieul, "Special issue on technology of networked control systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 5–8, 2007.

[2] C. Hadjicostis, C. Langbort, N. Martins, and S. Yüksel, "Special issue on information processing and decision making in distributed control systems," *Int. J. Systems, Control, and Communications*, vol. 2, no. 1/2/3, 2010.

[3] S. Tatikonda and S. Mitter, "Control under communication constraints," *IEEE Transactions on Automatic Control*, vol. 49, no. 7, pp. 1056–1068, 2004.

[4] G. N. Nair, F. Fagnani, S. Zampieri, and R. J. Evans, ""feedback control under data rate constraints: an overview," *Proceedings of the IEEE, Special issue on Technology of Networked Control Systems*, vol. 95, pp. 108–37, january 2007.

[5] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. Sastry, "Foundations of control and estimation over lossy networks," *Proceedings of the IEEE*, vol. 95, no. 1, p. 163, 2007.

[6] O. Imer, S. Yüksel, and T. Başar, "Optimal control of LTI systems over unreliable communication links," *Automatica*, vol. 42, no. 9, pp. 1429–1439, 2006.

[7] E. Garone, B. Sinopoli, and A. Casavola, "Lqg control over lossy tcp-like networks with probabilistic packet acknowledgements," *Int. J. Systems, Control, and Communications (Special Issue on Information Processing and Decision Making in Distributed Control Systems)*, vol. 2, no. 1/2/3, pp. 55–81, 2010.

[8] P. Bommannavar and T. Başar, "Optimal control with limited control actions and lossy transmissions," in *47th IEEE Conference on Decision and Control, 2008. CDC 2008*, pp. 2032–2037, 2008.

[9] O. Imer and T. Başar, "Optimal control with limited controls," in *Proceedings of the 2006 American Control Conference*, pp. 298–303, Citeseer, 2006.

[10] O. Imer and T. Başar, "Optimal estimation with limited measurements," *Int. J. Systems, Control, and Communications (Special Issue on Information Processing and Decision Making in Distributed Control Systems)*, vol. 2, no. 1/2/3, pp. 5–29, 2010.

[11] T. Başar and G. Olsder, *Dynamic Noncooperative Game Theory*. Series in Classics in Applied Mathematics, Philadelphia: SIAM, January 1999.