2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

One-Stage Control Over an Adversarial Channel with Finite Codelength

Abhishek Gupta Cédric Langbort Tamer Başar

Abstract— **We consider a model of stealthy attack on a networked control system by formulating a static zero-sum game among four players. Three of the players constitute a team of encoder, decoder and controller for a scalar discrete-time linear plant, while the fourth player is a jammer, who acts to flip the bits of the binary encoded observation signal of the communication channel between the plant and the controller. We assume that the observation and control signals have finite codelength. We characterize the encoding/decoding/control defense strategies available to the controller, and for simplicity in conveying the main ideas, we model it for a scalar discrete-time system with only one time step. We prove that there is no loss of generality in restricting our attention to binning-based encoding and control strategies. We determine the control and jamming strategies that are in saddle-point equilibrium for this game and show that the saddle-point value does not depend on the jamming policy. We also provide a necessary and sufficient condition on the minimum number of bits that are required to drive the cost to zero for this one-stage control problem in the presence of a jammer.**

I. Introduction

Adverse effects of communication channel-induced limitations on control have been intensively studied in the past decade. For example, a number of papers have considered the information theoretic minimum rate necessary for stabilization (see, e.g. [1], [2], [3], [4], [5]) or achieving optimal quadratic closed-loop performance [6], [7], [8]. In these papers, the channel behavior is assumed independent of the controller's action or plant's state, and uncorrelated across time. This, however, does not capture the scenarios in which the channel is adversarial.

In the absence of appropriate security measures, control systems are highly vulnerable to attack. Two types of attack on such systems have been considered in the past, namely denial of service (DoS) attack and deception (or integrity) attack [9], [10], [11]. Under DoS attack, the communication link is jammed in order to break the information exchange between the subsystems, while in a deception attack, the data of the subsystems are altered in order to deceive the controller and harm the system. Such attacks on control systems are hard to analyze, since the channel behavior is dependent on the state of the system, as well as correlated across time.

In our earlier work [12] and [13], we considered a DoS attack, in which an intelligent jammer jams the communication link between a controller and a scalar discrete-time linear plant. The jammer's goal was to optimally block the control signal by using a finite number of jamming actions over a horizon of N time steps. The restriction on the number of jamming actions captures the energy constraint of the jammer. We formulated this problem as a zero-sum game between the controller and the jammer. We showed that, when playing optimally, the jammer uses a threshold-based policy, i.e., jammer jams if the state is above a certain threshold, which is dependent on the jamming history and the current time step.

In [12], we modeled the communication as an analog channel, which can pass real numbers over the network without any error. However, in digital systems, real numbers need to be quantized and binary codewords are sent across a channel. Limited bandwidth also prohibits the controller to send large amount of data over the network within a short span of time. This means that one cannot have a large number of quantization bins with arbitrarily small size, since this requires long codelength. Hence, in this work, we consider the scenario where the observation and control signals are sent in binary codewords with limited codelength. The jammer, instead of blocking the signal completely, can only flip a limited number of bits in the codeword to corrupt the data. The jammer's role is in a sense similar to a binary symmetric channel, but is also different in the sense that it flips the bits deterministically and strategically to alter the data. This study therefore falls in the category of deception attack as described above.

Our main goal is to compute the saddle-point encoding and control strategies for a general dynamic multi-agent system under an adversarial attack like that of a jammer. However, to gain insight and for simplicity of notation, we restrict our attention to scalar systems where a static game is played with the jammer. We also use concepts from non-cooperative game theory to study this class of attacks. We formulate the precise problem in Section II. Section III is the main part of this paper, in which we obtain encoding and control policies to attain zero cost. We also discuss relevant tools from error correcting coding theory for the current problem. Next, we derive the saddle-point value and corresponding strategies for the jammer-controller pair in Section IV. We analyze a dynamic case in Section V. Finally, we conclude our discussion in Section VI.

II. Problem Formulation

The state equation evolves as (note that we have a one-stage problem)

$$x^+ = Ax + u + w, \tag{1}$$

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 supported in part by AFOSR Grant FA9550-09-1-0249, and part by AFOSR MURI Grant FA9550-10-1-0573.

978-1-61284-799-3/11/$26.00 ©2011 IEEE

where $x \in \mathbb{R}$ is the initial state of the plant, $x^+ \in \mathbb{R}$ is the final state of the plant, $u \in \mathbb{R}$ is the control signal and w is a random variable with support $[-\Delta, \Delta], \Delta < 1$. Initial state x is a uniformly distributed random variable in the interval $\mathcal{I} := [-1, 1]$ and independent of the process noise w. We study only the case of $A > 1 - \Delta$, since for $|A| \leq 1 - \Delta$, a trivial control strategy is $u = 0$. The case of $A < -1 + \Delta$ can be analyzed in a similar fashion as in this paper.

What we consider is a prototype of a scenario where the controller and the plant are far from each other, such that the plant sends the state information to the controller and the controller sends the control signal to the plant via a communication channel. Figure 1 provides a schematic description of the interconnections and the flow of information in the system. For the analysis, the channel is assumed to be perfect (but unsecured). It does not induce any error on the received bit at the controller or the plant end as only the jammer can induce errors. The plant and the controller can send at most n bits across the channel. The codeword sent over the channel from the plant to the controller is intercepted by a jammer, which can flip at most t bits in the codeword of the observation signal. We assume that the jammer jams the channel from the plant to the controller, while the channel from the controller to the plant is not intercepted by the jammer. Furthermore, we assume that the jammer knows the codebook of the encoder.

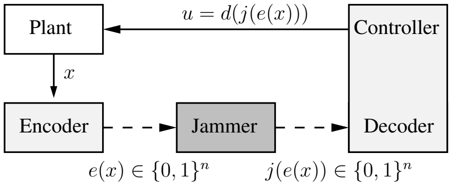

Fig. 1. Control in the presence of an intelligent jammer. The lightly shaded blocks belong to one player (referred to as controller) and the darker shaded block is the other player (the jammer). See text for details.

In the problem described above, it is required that the state of the plant does not deviate too much from a desirable set point. Hence, if the state of the system starts within a bounded set, we would like it to remain in the same set with high probability at the next time step. Thus, the cost function associated with this problem is

$$J = P\left\{x^+ \notin \mathcal{I} | x \in \mathcal{I}\right\} \tag{2}$$

which is to be minimized by the encoder-decoder-controller team and maximized by the jammer. We will henceforth refer to this team as controller, while in fact, it comprises three players.

For a given codelength $n > 0$ and jamming parameter $t \geq 0$, we denote the set of all measurable maps from $\mathcal{I}$ to $\{0,1\}^n$ by $\mathcal{E}_n$ and the set of all measurable maps from $\{0,1\}^n$ to $\mathbb{R}$ by $\mathcal{D}_n$.

Let $\epsilon_1, \epsilon_2 \in \{0,1\}^n$ be two codewords of length n. Then, the Hamming distance $H(\cdot,\cdot)$ between the two codewords is given by

$$H(\epsilon_1, \epsilon_2) = \sum_{i=1}^{n} \epsilon_{1i} \oplus \epsilon_{2i},$$

where ϵ_{1i} and ϵ_{2i} are the i^{th} bits of the codewords ϵ_1 and ϵ_2 respectively, and $\oplus$ is the XOR operator. We denote the set of all jamming strategies from $\{0,1\}^n$ to $\{0,1\}^n$ with Hamming distance less than or equal to t by $\mathcal{J}_{(t,n)}$, i.e.,

$$\begin{aligned} \mathcal{J}_{(t,n)} \quad = \quad & \{j : \{0,1\}^n \to \{0,1\}^n : \\ & H(\epsilon, j(\epsilon)) \leq t, \forall \epsilon \in \{0,1\}^n\}. \end{aligned}$$

The set $\mathcal{E}_n \times \mathcal{D}_n$ can be thought of as the strategy space for the team composed of the encoder, which communicates the plant's observation in n bits, and the decoder/controller, which maps the possibly corrupted message at the end of the channel into a control input. Similarly, $\mathcal{J}_{(t,n)}$ is the jammer's strategy space, which flips at most t bits in the encoded sequence. To every choice $(e, d) \in \mathcal{E}_n \times \mathcal{D}_n$ of the encoder-decoder/controller team and $j \in \mathcal{J}_{(t,n)}$ of the jammer corresponds the cost which, by a slight abuse of notation, we denote by $J(e, d; j)$.

We are interested in computing the quantity

$$\gamma(n) := \inf_{(e,d)\in\mathcal{E}_n\times\mathcal{D}_n} \sup_{j\in\mathcal{J}_{(t,n)}} J(e, d; j) \tag{3}$$

as a function of the codelength n and, in particular, in determining the smallest codelength $n^\star$ for which $\gamma(n) = 0$ for all $n \geq n^\star$.

We also consider a zero-sum game between the controller and the jammer with the cost function in (2), which the controller strives to minimize and the jammer tries to maximize. Towards this end, we wish to compute the saddle-point equilibrium control strategy $(e^*, d^*) \in \mathcal{E}_n \times \mathcal{D}_n$ and jamming strategy $j^* \in \mathcal{J}_{(t,n)}$ such that

$$J(e^*, d^*; j) \leq J(e^*, d^*; j^*) \leq J(e, d; j^*),$$

which holds for all $(e, d) \in \mathcal{E}_n \times \mathcal{D}_n$ and $j \in \mathcal{J}_{(t,n)}$. It should be noted that $\gamma(n)$ as defined above, is equal to the value of the game $J(e^*, d^*; j^*)$ if the saddle-point equilibrium exists [14, pp. 19]. If the equilibrium does not exist, then $\gamma(n)$ is an upper bound for the minimum cost achievable by the controller in the presence of a jammer, and is called *security level* of the controller. Our main results are as follows :

1) There is no loss of generality in restricting the strategy space of the controller to binning-based control strategies. We show this by constructing a binning-based control strategy given any other control strategy, such that the binning-based control strategy achieves a lower or equal cost.
2) We show that the saddle-point equilibrium exists, and hence $\gamma(n) = J(e^*, d^*; j^*)$.
3) We show that the saddle-point equilibrium strategy (e^*, d^*) of the controller can be restricted to a subset of $\mathcal{E}_n \times \mathcal{D}_n$. Moreover, we show that if the saddle-point value of the game is V, then

$$J(e^*, d^*; j) = V \qquad \forall j \in \mathcal{J}_{(t,n)}.$$

4) We obtain a necessary and sufficient condition on $n^\star$ such that $\gamma(n) = 0$ for all $n \geq n^\star$.

III. Binning-based Encoding and Control Strategies

When there is a finite length encoding of a real number, it is natural to think of using quantization and a binning-based strategy for the controller. Towards this end, we partition the interval $\mathcal{I}$ into N sub-intervals (henceforth, termed as bins). The encoder takes x as the input, determines which bin x belongs to, and outputs the codeword corresponding to that bin. Note that our problem formulation in the previous section does not enforce this structure on the controller.

A. Notation

We now introduce some additional notation, which will be used later in the paper. We say that an encoding strategy $e \in \mathcal{E}_n$ is N bin-based if there exists a mutually disjoint, non-empty set of partitions $(\mathcal{B}_1, ..., \mathcal{B}_N)$ of the interval $\mathcal{I}$ such that

$$\mathcal{I} \;\; = \;\; \bigcup_{i \in \mathscr{J}} \mathcal{B}_i, \quad \mathscr{J} := \{1, \ldots, N\},$$

and corresponding to each bin $\mathcal{B}_i$, there exists a codeword $\epsilon_i \in \{0,1\}^n$ such that

$$e(x) \;\; = \;\; \epsilon_i \text{ for all } x \in \mathcal{B}_i, i \in \mathscr{J}.$$

If, for $i, k \in \mathscr{J}$, the codewords satisfy

$$H(\epsilon_i, \epsilon_k) \geq 2s + 1 \text{ for all } i \neq k, \quad s \in \mathbb{N} \cup \{0\}, \tag{4}$$

we say that the encoding strategy $e \in \mathcal{E}_n$ is s-error free. Let $\epsilon \in \{0,1\}^n$ be the codeword received by the decoder and define $h : \{0,1\}^n \to \mathscr{J}$ by the following relation:

$$h(\epsilon) := \underset{i \in \mathscr{J}}{\arg\min} \;\; H(\epsilon_i, \epsilon). \tag{5}$$

We say that a control strategy $d \in \mathcal{D}_n$ is N bin-based if there exist N codewords $\epsilon_i \in \{0,1\}^n$ ($i \in \mathscr{J}$) and N control inputs $u_1, ..., u_N \in \mathbb{R}$ such that

$$d(\epsilon) = u_{h(\epsilon)} \text{ for all } \epsilon \in \{0,1\}^n.$$

The set of points $\mathcal{T}_i^d \subset \mathcal{I}$ where the control u_i keeps the state within the interval $\mathcal{I}$ is given by

$$\mathcal{T}_i^d = \left[\frac{-1 + \Delta - u_i}{A}, \frac{1 - \Delta - u_i}{A} \right] \cap \mathcal{I}. \tag{6}$$

We let $\mathcal{N}_i^e \subseteq \mathscr{J}$ denote the set of bin indices defined as:

$$\mathcal{N}_i^e = \{k \in \mathscr{J} : \exists j \in \mathcal{J}_{(t,n)}, k = h(j(e(x))) \, \forall x \in \mathcal{B}_i\}, \tag{7}$$

where $h(\cdot)$ is defined in (5). Clearly, the set $\mathcal{N}_i^e$ for each bin index $i \in \mathscr{J}$ is dependent on the encoding strategy. Since the jammer flips at most t bits, the set $\mathcal{N}_i^e$ consists of all bin indices corresponding to the nearest neighbors of all 0 to t flips in the codeword for the i^{th} bin. For all $m \in \mathcal{N}_i^e$, $H(e(x), e(y)) \leq 2t \;\; \forall \; x \in \mathcal{B}_i, \; \forall \; y \in \mathcal{B}_m$. Also, it is easy to see that $i \in \mathcal{N}_i^e$ for all $i \in \mathscr{J}$.

We say that N bin-based encoding and control strategies are adapted if the codewords $\{\epsilon_1, ..., \epsilon_N\}$ are the same for both strategies, and refer to the pair (e, d) as s-error free when e is s-error free. The set of all adapted N bin-based, s-error free encoding and control strategy pairs with codelength n is denoted by $\mathcal{S}_{(n,N,s)}$.

Although an N bin-based decoding strategy may not be well-defined over $\{0,1\}^n$ (because there may be more than one index satisfying (5)), the expression $d(j(e(x)))$ is well defined for every $x \in \mathcal{I}$, every N bin-based encoding strategy e *that is adapted to* d *and* t*-error free*, and every $j \in \mathcal{J}_{(t,n)}$. Indeed, in this case, condition (4) ensures that every codeword used by the encoding and control strategies is uniquely recovered by the nearest neighbor rule (5), regardless of which t out of its n bits are flipped. In addition, for every $(e, d) \in \mathcal{S}_{(n,N,t)}$,

$$d(j(e(x))) = u_i \;\Leftrightarrow\; x \in \mathcal{B}_i \tag{8}$$

for all $j \in \mathcal{J}_{(t,n)}$ and all $x \in \mathcal{I}$.

In words, when the encoder-decoder/controller team uses a pair of strategies in $\mathcal{S}_{(n,N,t)}$, it is guaranteed that system (1) will receive the input signal corresponding to the actual bin in which the state lies, regardless of the action of the jammer. The goal of the team is to achieve a cost $\sup_{j \in \mathcal{J}_{(t,n)}} J(e, d; j)$ of zero by appropriately choosing control inputs corresponding to each bin.

B. Achieving Zero Cost

The lemma below states a necessary and sufficient condition on the number of bins that are required to keep the state in the set $\mathcal{I}$ at the next time step.

Lemma 1: Let n, N, and t be such that N bin-based, t-error free encoding strategies exist. Then there exists $(\bar{e}, \bar{d}) \in \mathcal{S}_{(n,N,t)}$ such that $J(\bar{e}, \bar{d}; j) = 0$ for all $j \in \mathcal{J}_{(t,n)}$ if and only if

$$N \geq \left\lceil \frac{|A|}{1 - \Delta} \right\rceil. \tag{9}$$

In addition, when (9) holds, $\bar{e}$ and $\bar{d}$ can be constructed with the following choice of bins $(\mathcal{B}_1, ..., \mathcal{B}_N)$ and control inputs $u_1, ..., u_N$:

- if N is odd:
$$\mathcal{B}_k = \left[\frac{2k-1}{N}, \frac{2k+1}{N} \right), \; u_k = -\frac{2kA}{N}, \tag{10}$$
for all $-\frac{N-1}{2} \leq k \leq \frac{N-1}{2}$
- if N is even,
$$\mathcal{B}_k = \left[\frac{2(k-1)}{N}, \frac{2k}{N} \right), \; u_k = -\frac{(2k-1)A}{N}, \tag{11}$$
for all $-\frac{N}{2} + 1 \leq k \leq \frac{N}{2}$.

In both cases, any set of codewords $\epsilon_1, ..., \epsilon_N \in \{0,1\}^n$ which renders $\bar{e}$ t-error free can be chosen.

Proof: The reader is referred to [13, pp. 52]. ■

For the no-noise case, $\Delta = 0$, and the length of the codewords is $\log_2 \lceil |A| \rceil$. This value is the same as obtained in [1], [2], [3], since it has been shown that the codelength has to be (strictly) greater than $\log_2 |A|$ for the system to be stabilizable. However, if the goal is to keep the state bounded in the same interval and in the presence of process noise with bounded support, the required codelength increases. If the process noise has unbounded support (as in the case of Gaussian noise), then there is no control strategy with the finite codelength scheme which can keep the state within the given bound with probability one, as has been observed in [1], [2] among many others.

C. Jamming and Error Correcting Code

In Lemma 1, we noticed that a certain minimum number of bins is necessary to be able to keep the state in the set $\mathcal{I}$ even without a jammer. If the number of bins is less than that, then there is no hope of being able to keep the state within the interval $\mathcal{I}$ at the next time step. In the absence of a jammer, the required codelength for encoding N bins is bounded below by $\lceil \log_2(N) \rceil$. In the presence of the jammer, if the control strategy lies in the set $\mathcal{S}_{(n,N,t)}$ with $N \geq \left\lceil \frac{|A|}{1-\Delta} \right\rceil$, then the cost achieved by the controller is zero.

The following result, which is classical in the theory of error correcting codes, provides conditions for the (non) existence of t-error free N bin based encoding strategies with a codelength n.

Lemma 2 (Gilbert & Hamming bounds [15]): If

$$N \leq \frac{2^n}{\sum_{j=0}^{2t} \begin{pmatrix} n \\ j \end{pmatrix}}, \tag{12}$$

then there exists a t-error free N bin-based encoding strategy. However, if

$$N > \frac{2^n}{\sum_{j=0}^{t} \begin{pmatrix} n \\ j \end{pmatrix}}, \tag{13}$$

there does not exist any t-error free N bin-based encoding strategy. ■

Lemma 2 implies that, for every N and t, the set of codelengths for which there exists a t-error free N-bin based encoding strategy is non-empty. In addition, the minimum codelength for which there exists a control strategy in $\mathcal{S}_{(n,N,t)}$, denoted by $n_{ecc}(N,t)$, satisfies

$$n_1(N,t) \leq n_{ecc}(N,t) \leq n_2(N,t), \tag{14}$$

where $n_i(N,t), i = 1, 2$ is given by

$$n_i(N,t) = \min \left\{ n \in \mathbb{N} : N \leq \frac{2^n}{\sum_{j=0}^{it} \begin{pmatrix} n \\ j \end{pmatrix}} \right\}$$

Currently, it is not known how close $n_{ecc}(N,t)$ is to $n_1(N,t)$ (known as the Hamming bound) or to $n_2(N,t)$ (known as the Gilbert bound) [15]. There are only a few coding strategies for which the Hamming bound is tight, and they are known as perfect codes [15].

We now look into the case when the codelength $n < n_{ecc}(N,t)$. If $n < n_{ecc}(N,t)$, then there is no encoding-decoding strategy which can correct all t flips by the jammer. Using nearest neighbor decoding rule, the decoder may obtain a wrong bin index, and then a wrong control input is sent to the plant. Thus, the cost of $zero$ cannot be achieved by a binning-based control strategy if $n < n_{ecc}(N,t)$. We prove this claim in Theorems 3 and 7 below.

Among the many ways by which an encoding and a control strategy can be designed for $n < n_{ecc}(N,t)$, we focus our attention on two. One possible way (and rather naive way) is to encode in such a way that the codewords are t-error free, and compute the control strategy which minimizes the cost to the controller. Another possible way is to consider a game between the controller team and the jammer, and compute the saddle-point equilibrium strategy for the team and the jammer. The surprising result is that the naive way of encoding and controlling the plant is also the saddle-point strategy for the team and the jammer.

For a fixed codelength n and the number of flips t, and define N_{cr} to be

$$N_{cr} = \max\{N \in \mathbb{N} : n_{ecc}(N,t) \leq n\}. \tag{15}$$

In this case, we can obtain N_{cr} number of codewords, each of codelength n, which are $2t+1$ bits apart. Using this set of codewords, we can obtain an upper bound on the cost function and as a consequence, an upper bound on $\gamma(n)$. The following theorem establishes an upper bound on the cost as a function of n.

Theorem 3: Let N_{cr} be given by (15). If $N_{cr} < |A|/(1-\Delta)$, then for a pair $(e,d) \in \mathcal{S}_{(n,N_{cr},t)}$ such that $\mathcal{T}_i^d \subseteq \mathcal{N}_i^e$ for all $i \in \{1,\ldots,N_{cr}\}$, the cost to the controller is

$$J(e,d;j) = \left(1 - \frac{N_{cr}(1-\Delta)}{|A|}\right), \quad \forall j \in \mathcal{J}_{(t,n)}. \tag{16}$$

Moreover, this is the best cost achievable by the controller in the class of strategies in the set $\mathcal{S}_{(n,N_{cr},t)}$.

Proof: For a proof, refer to [13, pp. 57]. ■

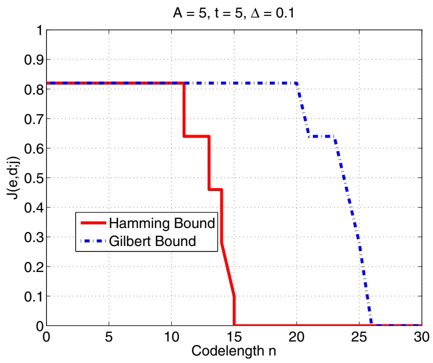

Fig. 2. The change in cost to the controller $J(e,d;j)$ with increase in the codelength n as obtained from Theorem 3 using the Hamming bound and the Gilbert bound. The actual cost lies between the two curves and depends on $n_{ecc}(N,t)$.

The following theorem establishes the fact that there is no loss of generality in restricting the encoding strategy to t-error free binning-based strategies. For proving this, we need the following lemma which simplifies the cost function.

Lemma 4: Let $(e, d) \in \mathcal{E}_n \times \mathcal{D}_n$ be an N-bin based control strategy. The maximum cost function over $j \in \mathcal{J}_{(t,n)}$ is equivalent to

$$\max_{j\in\mathcal{J}_{(t,n)}} J(e,d;j) = \sum_{i=1}^{N} \max_{k\in\mathcal{N}_i^e} P\{x \in \mathcal{B}_i \backslash \mathcal{T}_k^d\}. \tag{17}$$

Proof: This can be proved using Bayes' theorem. The detailed proof is given in [13, pp. 56, 61]. ■

Using the lemma above, we prove the following theorem.

Theorem 5: For a fixed n and t, let $(e, d) \in \mathcal{E}_n \times \mathcal{D}_n$ be an $\tilde{N}$-bin based control strategy. Then, there exists a corresponding t-error free N bin-based encoding and control strategy $(e_N, d_N) \in \mathcal{S}_{(n,N,t)}$ with $N \leq \tilde{N}$ bins such that $\sup_{j\in\mathcal{J}_{(t,n)}} J(e_N, d_N; j) \leq \sup_{j\in\mathcal{J}_{(t,n)}} J(e, d; j)$.

Proof: Let $\{\tilde{\mathcal{B}}_i\}_{i=1}^{\tilde{N}}$ be the set of mutually disjoint bins and $\{\tilde{\epsilon}_i\}_{i=1}^{\tilde{N}}$ be the set of corresponding codewords used in encoding of the bins with the control strategy (e, d). From the set $\{\tilde{\epsilon}_i\}_{i=1}^{\tilde{N}}$ of codewords, extract the maximal subset of codewords $\{\tilde{\epsilon}_{i_l}\}_{l=1}^{N} \subset \{\tilde{\epsilon}_i\}_{i=1}^{\tilde{N}}$ such that $H(\tilde{\epsilon}_{i_l}, \tilde{\epsilon}_{i_k}) \geq 2t+1$ for $l, k \in \{1, \ldots, N\}, l \neq k$.

Now, we construct t-error free N bin-based control strategy $(e_N, d_N) \in \mathcal{S}_{(n,N,t)}$. Towards this end, we define $\mathcal{N}_1^{e_N} = \mathcal{N}_{i_1}^e$ and if $N \geq 2$, define $\mathcal{N}_l^{e_N} = \mathcal{N}_{i_l}^e \backslash \bigcup_{k=1}^{l-1} \mathcal{N}_{i_k}^e$ for $l \in \{2, \ldots, N\}$. Then, the set $\{\mathcal{N}_l^{e_N}\}_{l=1}^N$ has mutually disjoint sets. Note that since $i_l \in \mathcal{N}_m^e$ for all $m \in \mathcal{N}_l^{e_N}$, we have $i_l \in \mathcal{N}_l^{e_N}$. Moreover, $\bigcup_{l=1}^N \mathcal{N}_l^{e_N} = \bigcup_{i=1}^{\tilde{N}} \mathcal{N}_i^e$.

Define $\mathcal{B}_l = \bigcup_{m\in\mathcal{N}_l^{e_N}} \tilde{\mathcal{B}}_m$ for $l = 1, \ldots, N$ and set $\epsilon_l = \tilde{\epsilon}_{i_l}$ such that $e_N(x) := \epsilon_l$ for all $x \in \mathcal{B}_l$. Since $\{\tilde{\epsilon}_{i_l}\}_{l=1}^N$ is a set of t-error free codewords, the encoding strategy e_N is t-error free. Also, since the set $\{\mathcal{N}_l^{e_N}\}_{l=1}^N$ has mutually disjoint sets, we have $\mathcal{B}_l \bigcap \mathcal{B}_k = \emptyset$ for all $l, k \in \{1, \ldots, N\}, l \neq k$. Define $d_N(\epsilon_l) := d(\tilde{\epsilon}_{i_l})$ for all $l \in \{1, \ldots, N\}$ to be the N bin-based control strategy.

Using Lemma 4, we know that the maximum value of the cost function with control strategy (e, d) over $j \in \mathcal{J}_{(t,n)}$ is

$$\max_{j\in\mathcal{J}_{(t,n)}} J(e,d;j) = \sum_{i=1}^{\tilde{N}} \max_{k\in\mathcal{N}_i^e} P\{x \in \tilde{\mathcal{B}}_i \backslash \mathcal{T}_k^d\}.$$

This can be rewritten and bounded below by

$$\begin{aligned}
\max_{j\in\mathcal{J}_{(t,n)}} J(e,d;j) &= \sum_{l=1}^{N} \sum_{m\in\mathcal{N}_l^{e_N}} \max_{k\in\mathcal{N}_m^e} P\{x \in \tilde{\mathcal{B}}_m \backslash \mathcal{T}_k^d\}, \\
&\geq \sum_{l=1}^{N} \sum_{m\in\mathcal{N}_l^{e_N}} P\{x \in \tilde{\mathcal{B}}_m \backslash \mathcal{T}_{i_l}^d\}, \\
&= \sum_{l=1}^{N} P\{x \in \mathcal{B}_l \backslash \mathcal{T}_{i_l}^d\}.
\end{aligned}$$

Here, the first inequality holds because we select $k = i_l$ for each $l \in \{1, \ldots, N\}$ instead of taking maximum over each $\mathcal{N}_m^e$. The second equality holds since

$$\begin{aligned}
P\left\{x \in \mathcal{B}_l \backslash \mathcal{T}_{i_l}^d\right\} &= P\left\{x \in \left(\bigcup_{m\in\mathcal{N}_l^{e_N}} \tilde{\mathcal{B}}_m\right) \backslash \mathcal{T}_{i_l}^d\right\}, \\
&= P\left\{x \in \bigcup_{m\in\mathcal{N}_l^{e_N}} \left(\tilde{\mathcal{B}}_m \backslash \mathcal{T}_{i_l}^d\right)\right\}, \\
&= \sum_{m\in\mathcal{N}_l^{e_N}} P\left\{x \in \tilde{\mathcal{B}}_m \backslash \mathcal{T}_{i_l}^d\right\}.
\end{aligned}$$

Since $(e_N, d_N) \in \mathcal{S}_{(n,N,t)}$, the corresponding cost $J(e_N, d_N; j)$ remains the same for all $j \in \mathcal{J}_{(t,n)}$. It follows that $\sup_{j\in\mathcal{J}_{(t,n)}} J(e_N, d_N; j) \leq \sup_{j\in\mathcal{J}_{(t,n)}} J(e, d; j)$.

Therefore, the t-error free binning-based control strategy $(e_N, d_N) \in \mathcal{S}_{(n,N,t)}$ achieves a lower or the same cost, and this completes the proof of Theorem 5. ■

As a result of Theorem 5, we have the following result.

Theorem 6: For the problem formulated in Section II, let N_{cr} be given by (15). If $N_{cr} < |A|/(1-\Delta)$, then

$$\gamma(n) = \left(1 - \frac{N_{cr}(1-\Delta)}{|A|}\right). \tag{18}$$

Also, $\gamma(n) = 0$ if and only if $n \geq n^\star := n_{ecc}\left(\left\lceil \frac{|A|}{(1-\Delta)} \right\rceil, t\right)$.

Proof: Using the definition of $\gamma(n)$, we have

$$\gamma(n) \leq \inf_{(e_N,d_N)\in\mathcal{S}_{(n,N,t)}} \sup_{j\in\mathcal{J}_{(t,n)}} J(e,d;j),$$

since we are infimizing over a smaller set $\mathcal{S}_{(n,N,t)}$, which is a subset of $\mathcal{E}_n \times \mathcal{D}_n$. From Theorem 5, we get

$$\inf_{(e_N,d_N)\in\mathcal{S}_{(n,N,t)}} \sup_{j\in\mathcal{J}_{(t,n)}} J(e_N,d_N;j) \leq \gamma(n).$$

Both inequalities, together with the result of Theorem 3 yields the equality in (18).

Next, if $n \geq n_{ecc}(\lceil |A|/(1-\Delta) \rceil, t)$, then the construction in Lemma 1 guarantees zero cost. This implies that $\gamma(n) = 0$. Conversely, let $n < n_{ecc}(\lceil |A|/(1-\Delta) \rceil, t)$. From (15), we know that if $n < n_{ecc}(\lceil |A|/(1-\Delta) \rceil, t)$, then $N_{cr} < |A|/(1-\Delta)$. As a result of (18), $\gamma(n) > 0$. This completes the proof of Theorem 6. ■

IV. Zero-sum Game Problem

In the last section, we obtained the worst cost $\gamma(n)$ the controller can achieve in the presence of a jammer, which by definition, is the security level of the controller for the zero-sum game. We now have the following theorem, which says that the security level of jammer in the zero-sum game considered in Section II is the same as the cost in (16) given in Theorem 3. Since the security levels of both the players are equal, the saddle-point value of the zero-sum game exists [14]. The following theorem summarizes and proves this fact.

Theorem 7: For a given n and t, if $n < n_{ecc}\left(\left\lceil \frac{|A|}{(1-\Delta)} \right\rceil, t\right)$, then the saddle-point encoding and control strategy is $(e^*, d^*) \in \mathcal{S}_{(n,N_{cr},t)}$ such that $\mathcal{T}_i^{d^*} \subseteq \mathcal{N}_i^{e^*}$ for all $i \in \{1, \ldots, N_{cr}\}$ and the saddle-point jamming strategy j^* is any jamming strategy in $\mathcal{J}_{(t,n)}$. The corresponding saddle-point value of the game is given by (16).

Proof: Let $(e,d) \in \mathcal{E}_n \times \mathcal{D}_n$ be an arbitrary control strategy and let $\tilde{j}^* \in \mathcal{J}_{(t,n)}$ be the jamming strategy which achieves the maximum of $J(e,d;j)$ over all $j \in \mathcal{J}_{(t,n)}$. Then, using the construction in Theorem 5, we can obtain $(e_N, d_N) \in \mathcal{S}_{(n,N,t)}$, such that $J(e_N, d_N; j) \leq J(e,d;\tilde{j}^*)$ $\forall j \in \mathcal{J}_{(t,n)}$. Using this fact and Theorem 3, we get

$$J(e^*, d^*; j^*) \leq J(e_N, d_N; j) \leq J(e,d;\tilde{j}^*) \quad \forall j \in \mathcal{J}_{(t,n)}.$$

From Theorem 3, we know that if $(e^*, d^*) \in \mathcal{S}_{(n,N_{cr},t)}$ such that $\mathcal{T}_i^{d^*} \subseteq \mathcal{N}_i^{e^*}$ for all $i \in \{1, \ldots, N_{cr}\}$, then

$$J(e^*, d^*; j) = J(e^*, d^*; j^*) \qquad \forall j \in \mathcal{J}_{(t,n)}.$$

Hence, (e^*, d^*) and j^* are a saddle-point control and jamming strategy, respectively, for the zero-sum game. Since the saddle-point control and jamming strategies are independent of each other, by ordered interchangeability property of multiple saddle-point strategies in a zero-sum game, any other saddle-point control and jamming strategy incurs the same value. Therefore, we conclude that the saddle-point value of the game is given by (16). ■

V. The Dynamic Case

Due to non-linear control strategy and additive noise, the probability distribution of x^+ is not uniform. However, we analyze the support of the probability distribution of the state at every time step $k \in \mathbb{N}$ for a dynamic case in the next theorem.

Theorem 8: Let the initial state have a symmetric probability distribution with support of length $2\lambda_0$ with $\lambda_0 = 1$, and $2\lambda_k$ denote the length of the support of the probability distribution of the state at time instant k. Then, λ_{k+1} grows as

$$\lambda_{k+1} = \frac{|A|}{N_{cr}} \lambda_k + \Delta, \quad k = 0, 1, \ldots. \tag{19}$$

If $N_{cr} \geq |A|/(1-\Delta)$, then the sequence $\{\lambda_i\}_{i=0}^{\infty}$ converges, and it satisfies $\lambda_k \leq \lambda_0 = 1 \;\forall\, k \in \mathbb{N}$.

Proof: Let $\mathcal{B}_i$ be a bin such that $x_1 = \inf \mathcal{B}_i$, $x_2 = \sup \mathcal{B}_i$, u be the control action corresponding to this bin and assume $A > 1$. Then, $\lambda_{k+1} \geq Ax_2 + u + \Delta$ and $-\lambda_{k+1} \leq Ax_1 + u - \Delta$. Both these inequalities give $x_2 - x_1 \leq \frac{2(\lambda_{k+1} - \Delta)}{A}$, which means $N_{cr} = \frac{2\lambda_k}{(x_2 - x_1)} \geq \frac{A\lambda_k}{(\lambda_{k+1} - \Delta)}$. This yields the sequence in (19) when $x_2 - x_1 = \frac{2(\lambda_{k+1} - \Delta)}{A}$, i.e., when the bins are of equal length. Clearly, this sequence converges if $N_{cr} > |A|$. However, if $N_{cr} \geq |A|/(1-\Delta)$, a simple induction argument shows that $\lambda_k \leq 1$ for all time steps $k \in \mathbb{N}$. ■

VI. Conclusion

We considered a model of deception attack on a scalar linear control system for a single time step, when the process noise has bounded support. We analyzed binning and control strategies for the case where the jammer can flip a limited number of bits in codewords of fixed length. We showed that there is no loss of generality in restricting the controller's strategy space to binning-based control strategies. We formulated a zero-sum game between the controller and the jammer, and derived the saddle-point value and corresponding strategies for the controller and the jammer.

An immediate extension of the results of the paper is to consider a finite horizon or an infinite horizon dynamic game between the jammer and the controller under codelength constraints. Besides jamming, if there is channel noise, then the problem is non-trivial as well as challenging due to non-classical information pattern at both the controller and the encoder sides. The case of channel without jammer but with channel noise has been investigated in [16], [17] using concepts from Markov chains. One of our goals in the future is to extend the framework in [16], [17] to adversarial channels.

References

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

[2] G. Nair, F. Fagnani, S. Zampieri, and R. Evans, "Feedback control under data rate constraints: an overview," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 108–137, 2007.

[3] S. Yüksel and T. Başar, "Minimum rate coding for LTI systems over noiseless channels," *IEEE Transactions on Automatic Control*, vol. 51, no. 12, pp. 1878–1887, 2006.

[4] P. Minero, M. Franceschetti, S. Dey, and G. Nair, "Data rate theorem for stabilization over time-varying feedback channels," *IEEE Transactions on Automatic Control*, vol. 54, no. 2, pp. 243–255, 2009.

[5] N. Martins, M. Dahleh, and N. Elia, "Feedback stabilization of uncertain systems in the presence of a direct link," *IEEE Transactions on Automatic Control*, vol. 51, no. 3, pp. 438–447, 2006.

[6] 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, pp. 163–187, 2007.

[7] 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.

[8] E. Garone, B. Sinopoli, and A. Casavola, "LQG control over lossy TCP-like networks with probabilistic packet acknowledgements," *International Journal of Systems, Control and Communications*, vol. 2, no. 1, pp. 55–81, 2010.

[9] S. Amin, A. Cárdenas, and S. Sastry, "Safe and secure networked control systems under denial-of-service attacks," *Hybrid Systems: Computation and Control*, pp. 31–45, 2009.

[10] H. Sandberg, A. Teixeira, and K. Johansson, "On security indices for state estimators in power networks," in *Preprints of the First Workshop on Secure Control Systems, CPSWEEK*, 2010.

[11] A. Cárdenas, S. Amin, and S. Sastry, "Research challenges for the security of control systems," in *Proceedings of the 3rd Conference on hot topics in security*, pp. 1–6, 2008.

[12] A. Gupta, C. Langbort, and T. Başar, "Optimal control in the presence of an intelligent jammer with limited actions," in *Proc. of 49th IEEE Conference on Decision and Control (CDC)*, pp. 1096–1101, December 2010.

[13] A. Gupta, "Control in the presence of an intelligent jammer with limited actions," Master's thesis, University of Illinois at Urbana-Champaign, Illinois, USA, 2011.

[14] T. Başar and G. Olsder, *Dynamic Noncooperative Game Theory*. Society for Industrial Mathematics (SIAM) Series in Classics in Applied Mathematics, Philadelphia, 1999.

[15] S. Wicker, *Error Control Systems for Digital Communication and Storage*. Prentice Hall, New Jersey, 1995.

[16] S. Yüksel and T. Başar, "Control over noisy forward and reverse channels," *IEEE Transactions on Automatic Control*, vol. 56, pp. 1014 –1029, May 2011.

[17] S. Yüksel and T. Linder, "Optimization and convergence of observation channels in stochastic control," *Arxiv preprint arXiv:1009.3824*, 2010.