Coding and Control over Discrete Noisy Forward and Feedback Channels

Serdar Yüksel[1] and Tamer Başar[1]

***Abstract*— We consider the problem of stabilizability of remote LTI systems where both the forward (from the sensor to the controller) and the feedback (from the controller to the plant) channels are noisy, discrete, and memoryless. Information theory and the theory of Markov processes are used to obtain necessary and sufficient conditions (both structural and operational) for stabilizability, with the conditions being on error exponents, delay and source-channel codes. These results generalize some of the existing results in the literature which assume either the forward or the reverse channel to be noise-free. We observe that unlike continuous alphabet channels, discrete channels entail a substantial complexity in encoding the unbounded state and control spaces for control of noisy plants. We introduce a state-space encoding scheme utilizing the dynamic evolution. We also present variable-length coding through variable-sampling to transmit countably infinite symbols over a finite channel.**

I. Introduction

A. Problem Formulation

We consider in this paper a remote control problem with communication constraints, as depicted in Fig. 1. The system

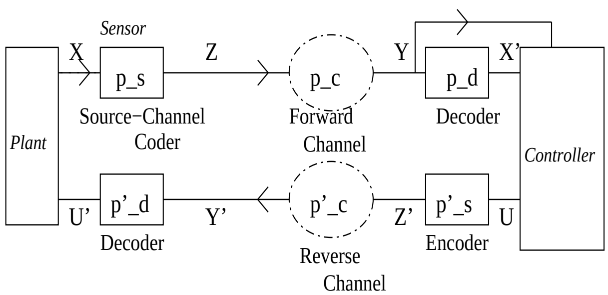

Fig. 1: Control over discrete noisy channels.

to be controled is a sampled version of an LTI continuous-time plant with the scalar dynamics

$$dx_t = (\xi x_t + b' u'_t)dt + dB_t, \tag{1}$$

where B_t is the standard Brownian motion process, u'_t is the (applied) control which is assumed to be piecewise constant (zeroth order hold) over intervals of length T_s, the initial state x_0 is a second-order random variable; and $\xi > 0$, which means that the system is unstable without control. After sampling, with period T_s, we have the discrete-time system

$$x_{t+1} = a x_t + b u'_t + d_t \tag{2}$$

where x_t is the state at time t, $\{d_t\}$ is a zero-mean i.i.d. Gaussian process. Here $a = e^{\xi T_s}$, $b = b'(e^{\xi T_s} - 1)/\xi$ and $E[d_t^2] = e^{2\xi T_s} - 1/2\xi$.

In the remote control setting, we refer to the channel which connects the sensor to the controller as the *forward channel*, and the channel which connects the controller to the plant as the *reverse channel* (see Fig. 1).

The timeline of the events is as follows: The state is sampled at discrete time instants $kT_s, \quad k \geq 0$. It takes $\alpha(N_f)N_f$ seconds to use the forward channel N_f times[2], and $\beta(N_r)N_r$ seconds to use the reverse channel N_r times.

The coding rate for the forward channel is defined as $R_f = \log(|M_f|)/N_f$, where M_f is the set of sensor symbols, and N_f is the number of channel uses. Likewise, for the reverse channel, the coding rate is $R_r = \log(|M_r|)/N_r$. Our goal in this paper is to obtain:

1) Outer and inner bounds for the set of forward and reverse rates which lead to a finite state variance in the limit, that is bounds for
$$\{R_f, R_r : \lim_{T\to\infty} E[x_T^2] < \infty\}.$$
2) Encoding schemes for both state and control symbols, with infinite size codebooks in both and dynamic evolution only in the former.

We focus here on stabilizability, since this is a necessary condition for the more comprehensive problem of controllability for linear systems.

B. Connections with the Literature

Works most relevant to this one in the literature are [2], [3] [4], [5], [6], [7], and [8]. References [2] and [3] are among the first to consider noisy channels. Reference [3] also introduces various problems which have had a significant impact on the emerging field of remote control. Reference [4] adopts a Lyapunov-based approach to stabilize a system over noiseless channels and shows that the coarsest time-invariant stabilizing quantizer is logarithmic and that the design has the same base for construction regardless of the sampling interval. We will show that this property regarding sampling carries over to stochastic systems as well. Reference [9] has shown that capacity does not have much relevance in a control context, and has introduced *anytime capacity* as a necessary and sufficient measure using noiseless feedback; furthermore, any-time decoding uses only finite delay with probability one. Unlike [9], the encoder and decoder in this

[1]Coordinated Science Laboratory, University of Illinois, 1308 West Main Street, Urbana, IL 61801-2307 USA. E-mail: (yuksel,tbasar)@decision.csl.uiuc.edu. Research supported in part by the NSF ITR grant CCR 00-85917.

[2]It might be possible to causally encode more recent information considering the delay in transmission; here we assume block coding and encode the state at time kT_s, $k \geq 0$.

work are not only causal but also of *zero-delay* type, i.e., the encoding and decoding are done symbol by symbol. Furthermore we do not allow for any feedback in communication and take the reverse channel also noisy. Another related reference, [6], studies stability over noisy discrete channels. There, the plant is noise-free, the reverse channel is noiseless, and for such a system it is argued that capacity is a sufficient measure. In our case, however, the plant and the reverse channels are also noisy. We will observe that for noiseless plants, if the number of codelengths is penalized, then capacity is not a sufficient measure; except for noiseless discrete channels. Furthermore we provide structural results on coding and decoding schemes for stabilizability.

Most of the studies in the literature have considered at least one noiseless channel connecting the controller and the plant and have not touched upon the effects when both channels are unreliable. Regarding noisy feedback channels, there have been just a few studies: Reference [10] has addressed the Gaussian channel case, with no encoding in the reverse channel in the relaxation of the noiseless feedback; reference [11] studies optimal control policies with packet losses in the feedback channel as well as the forward one. In a parallel work, [12], we study control over Gaussian channels for scalar systems and provide the optimal linear coder and controllers. In [13], communication with a noisy feedback channel has been considered in the context of estimation.

To recapitulate, in this paper we consider systems where both channels are noisy and discrete. The presence of a noisy channel with no explicit feedback leads to a *non-classical* information structure [2], since the agents (controller, encoders and decoders) do not have nested information. Furthermore the *dual effect* of control is present. Due to these difficulties, we will use indirect methods, information theory, and Markov stability theory, to arrive at necessary and sufficient conditions.

C. Notation and the System Model

In our setup, both the sensor and the controller act as both transmitters and receivers because of the closed-loop structure. We model the forward source-channel encoder as a mapping $p_s(z_t|x_t), x_t \in \mathcal{R},\ z_t \in \mathcal{Z}$, between the source output and channel input. The forward channel is a memoryless stochastic mapping between the channel input and output, $p_c(y_t|z_t), y_t \in \mathcal{Y}$, and the decoder is a mapping between the channel output, the information available at the control, I_{t-1}, and the output, i.e., $p_d(x'_t|I_{t-1}, y_t), x'_t \in \mathcal{X}'$, and $I_t = \{I_{t-1}, y_t, u_{t-1}\}$. The control, $u_t \in \mathcal{U}$, is generated using I_t. The reverse channel also has a source-channel encoder, $p'_s(z'_t|u_t), z'_t \in \mathcal{Z}'$, channel mapping $p'_c(y'_t|z'_t), y'_t \in \mathcal{Y}'$, and a channel decoder $p'_d(u'_t|y'_t), u'_t \in \mathcal{U}'$ (see Fig. 1), where the $p(\cdot|\cdot)$'s are all conditional probability densities or mass functions.

Definition 1.1: A Discrete Memoryless Channel (DMC) is characterized by an input alphabet $\mathcal{X}$, an output alphabet $\mathcal{Y}$, and a mapping $p_{y|x}(y|x)$, from $\mathcal{X}$ to $\mathcal{Y}$, which satisfies: $p_{y^n|x^n}(y_1^n|x_1^n) = \prod_{i=1}^{n} p_{y_i|x_i}(y_i|x_i), \forall x^n \in \mathcal{X}^n, y^n \in \mathcal{Y}^n$.

The source-coder is the quantizer, and the channel encoder generates the bit stream for each of the corresponding quantization symbols, thus generating the joint-source channel encoder.

We say the controller has memory of order m if the information available to it at time t is

$$I_t^m = \{y_{t-m}, \dots, y_t; u_{t-\max(m,1)}, \dots, u_{t-1}\}.$$

In case $m = 0$, we will have a memoryless controller; i.e., $I_t^0 = y_t$, which we will study in detail. In this case we will lump the forward source-channel encoder, the forward channel and the decoder mappings into a single mapping $p(x'|x)$, and likewise the reverse source-channel encoder, reverse channel and decoder mappings into $p'(u'|u)$.

A quantizer Q is constructed by corresponding bins $\{\mathcal{B}_i\}$ and their reconstruction levels q_i such that $\forall i,\ Q(x) = q_i <=> x \in \mathcal{B}_i$. We have $\forall i, q_i \in \mathcal{B}_i$. For scalar quantization, $x \in \mathbb{R}$ and $\mathcal{B}_i = (\delta_i, \delta_{i+1}]$, where $\{\delta_i\}$ are termed as "bin edges" and w.l.o.g. we assume the monotonicity on bin edges: $\forall i, \delta_i < \delta_{i+1}$. In this paper we consider "symmetric quantizers", which are defined as: If $\exists$ a quantization bin $(\delta_i, \delta_{i+1}]$, where $0 < \delta_i < \delta_{i+1}$, then $\mathcal{B}_{-i} = [-\delta_{i+1}, -\delta_i)$ is also a quantization bin.

We define the encodable state set $S_x \in \mathcal{R}$ as the set of elements which are represented by some codeword, $S_x := \bigcup_i \mathcal{B}_i$. Such a definition applies to the encodable control set, S_c, as well. Suppose the state is within the encodable set and is in the ith bin of the quantizer. The source coding output at the plant sensor will represent this state as q_i and send the ith index over the channel. After a joint mapping of the channel and the channel decoder, the controller will receive the index i as index j with probability $p(j|i)$. The controller will apply its control over index j, computing Q'_j -thus the controller decoder, controller, and encoder can be regarded as a single mapping- and send it over the reverse channel, which would interpret this value as Q'_l with probability $p'(l|j)$, by a mapping through the reverse channel. Given that the state is in the ith bin, the plant will receive the control Q'_l with probability $\sum_j p'(l|j)p(j|i)$. Thus, the applied control will be $u'_t = Q'_l$ with probability $\sum_j p'(l|j)p(j|i)$, and the probability of the state to be in the ith bin is $p(i) = p(x \in B_i)$.

In the study of stability of a Markovian system, an appropriate approach is to use drift conditions [14] (in particular see Chapters 8 and 14); we will use these conditions to first characterize and then construct state encoders. We will need the following two definitions [14] regarding Markov chains in the development to follow.

Definition 1.2: A Markov Chain, Φ, in a state space X, is $\Psi-$irreducible if for some measure Ψ, for any $B \in X$ with $\Psi(B) > 0$, $\forall x \in X$, there exists some integer $n > 0$, possibly depending on B and x, such that $P^n(x, B) > 0$, where $P^n(x, B)$ is the transition probability in n stages.

Definition 1.3: A probability measure π is invariant on $(X, \mathcal{B}_X)$ if $\pi(D) = \int_X P(x, D)\pi(dx), \quad \forall D \in \mathcal{B}_{\mathcal{X}}$.

We close this section with a brief outline of the organization of the paper. We study necessary conditions on the rates and the structures of the codes in section II, and then sufficiency results and code constructions in section III. We discuss the variable length coding for side channels in section IV, and conclude with comments on extensions to multi-dimensional systems in section V.

II. NECESSITY CONDITIONS

A. Conditions on Capacities

We note that the problem of minimizing $E[x_{t+1}^2]$ is identical to the minimization of

$$E[a^2(b/au'_t - (-x_t))^2 + d_t^2],$$

which can be regarded as a state estimation cost. Thus, we can approach the control problem as a problem of information transmission over a degraded relay channel, and the problem can be regarded as a state estimation problem over such a channel.

Theorem 2.1: For the existence of an invariant density with finite variance, channels should satify

$$\min(C_f, C_r) > \log_2(|a|),$$

where C_f and C_r are respectively the forward and reverse channel capacities.

Proof: An invariant density with a finite variance implies a finite invariant entropy (which is bounded by the entropy of the Gaussian density with the same variance). Since $x_{t+1} = a(x_t - b/au'_t) + d_t$, and conditioning does not increase entropy, and D_t is an independent noise process, we have

$$\begin{aligned} H(x_{t+1}) &\geq H(x_{t+1}|u'_t) = H(a(x_t - b/au'_t) + d_t|u'_t) \\ &= H(ax_t + d_t|u'_t) > H(ax_t + d_t|u'_t, d_t) \\ &= H(ax_t|u'_t) = \log_2(|a|) + H(x_t|u'_t), \end{aligned} \tag{3}$$

which implies $H(x_{t+1}) - H(x_t|u'_t) > \log_2(|a|)$. Since $I(x_t; u'_t) = H(x_t) - H(x_t|u'_t)$, we have

$$I(x_t; u'_t) > H(x_t) + \log_2(|a|) - H(x_{t+1}).$$

But $\lim_{t\to\infty}(H(x_{t+1}) - H(x_t)) = 0$, which leads to $\lim_{t\to\infty} I(x_t; u'_t) > \log_2(|a|)$. Now, from the data processing inequality [15] and the definition of capacity we have $\min(C_f, C_r) > \log_2(|a|)$. $\diamond$

We will observe in the next section that the capacity constraints are far from being sufficient as long as the delays in transmission due to longer codelengths are penalized.

B. Structural Conditions

An important observation in the development of this paper is now the following.

Theorem 2.2: For a linear system with $|a| > 1$, with channel transitions forming an irreducible Markov chain, if the encodable control set is bounded, the chain is transient.

Proof: Let $|b'u'_t| < M, \forall t \geq 0$. Define a process, $dv_t = \gamma v_t + dB_t$, with $v_0 = x_0 \in T_k := (R, 2^k R)$, where $x_0 > R > M/(\xi - \gamma)$, and $\xi > \gamma > 0$. Define $\tau := \inf\{t : x_t \leq R\}$,$\tau' := \inf\{t : v_t \leq R\}$, $\tau'_k := \inf\{t : v_t \notin T_k\}$. We have $\tau' \leq \tau$ almost surely. Let $f(x) = e^{-2\gamma x}$. Using Dynkin's formula [16], we have

$$E_{x_0}[f(v_{\tau'_k})] = f(x_0) + E_{x_0}[\int_0^{\tau'_k} Af(v_s)ds],$$

where A is the generator function, given by $Af(x) = \gamma(\partial f/\partial x) + 1/2(\partial^2 f/\partial x^2)$. Let p_R be the probability of exiting at R. Thus, we have $p_R e^{-2\gamma R} + (1-p_R)e^{-\gamma 2^{k+1}R} = e^{-2\gamma x_0}$. Since p_R is bounded, and $\gamma > 0$, we obtain:

$$\lim_{k\to\infty} p_R = e^{-2\gamma x_0}/e^{-2\gamma R} < 1.$$

Thus $p(\tau' < \infty) < 1$ and $p(\tau < \infty) < 1$. Hence, the chain is transient. $\diamond$

The counterpart of this result for the encodable state set is the following.

Theorem 2.3: For a linear system with $|a| > 1$, with discrete channel transitions forming an irreducible Markov chain, if the encodable state set is bounded, the Markov chain is transient.

The above results show that for the noisy discrete channel case one needs to encode the entire state space. Unlike a continuous alphabet channel, this restriction entails significant complexity on encoding for control over a discrete noisy channel, for there needs to be a matching between the entire state space which requires a countably infinite number of codewords and a finite-symbol channel. We will observe that using a dynamic structure, this problem can be overcome in some cases.

We now study the conditions for the existence of am invariant density with a finite second moment for the state for systems connected over DMCs.

III. STABILIZING RATE REGIONS

A. Stability Through Drift Conditions

We now consider the original system (2), and study the stochastic evolution of the state. Consider symmetric quantizers studied before. Suppose a time invariant decoding policy is used by the controller.

Theorem 3.1: Let $S \subset X$ be a closed and bounded interval around the origin, $L < \infty$, and let $\delta_i > 0, \quad \forall i$ (positive portion of the symmetric quantizer). Finally, let $1_{x\in S}$ be the indicator function for x being in S. Then, for a discrete channel, if the following drift condition holds for some sufficiently small $\epsilon > 0$, and for all bins:

$$\begin{aligned} -\delta_i + \max\Bigg(&|[\textstyle\sum_l \sum_j p(j|i)p'(l|j)[a\delta_i + bQ'_l]]|, \\ &|\textstyle\sum_l \sum_j p(j|i)p'(l|j)[a\delta_{i+1} + bQ'_l]|\Bigg) < -\epsilon + L1_{x\in S} \end{aligned}$$

then there exists an invariant probability distribution. Furthermore, if the following condition holds for all bins:

$$\begin{aligned} \max\Bigg(&\textstyle\sum_l \sum_j p(j|i)p'(l|j)[a\delta_i + bQ'_l]^2, \\ &\textstyle\sum_l \sum_j p(j|i)p'(l|j)[a\delta_{i+1} + bQ'_l]^2\Bigg) - \delta_i^2 \\ &< -\epsilon\delta_{i+1}^2 + L1_{\in S}\,, \end{aligned} \tag{4}$$

then $\lim_{t\to\infty} E[x_t^2]$ exists and is finite. The limit distribution is independent of the initial distribution.

Proof. See [1]. $\diamond$

For the case when the channels are noiseless, this leads to a logarithmic quantizer (with $\epsilon = 0, L = 0$), which was, in a control context, first introduced in [4].

Proposition 3.1: Let the forward and the reverse channels be noiseless. Consider a symmetric quantizer. For a scalar system to satisfy a drift towards the origin, for the non-negative quantizer values, quantizer bin edges have to satisfy

$$\delta_{i+1} \le (1 + 2/|a|)\delta_i \tag{5}$$

B. Trade-off Between Reliability and Delay

Although longer block codes improve the channel reliability, long delays and larger sampling periods are undesirable in control. The explicit dependence of error probability on the length is characterized by the *error exponents* [15]. The probability of error between two different codewords (i.e., $p(m|m'), m \neq m'; m, m' \in \mathcal{X}'$) can be upper bounded using the largest value of the minimum Bhattacharyya distance in a codebook ([15], Chapter 12). For any two codewords m, m',

$$d(m, m') \ge N[E_L(R) - o(N)/N], \quad m \neq m',$$

where R is the coding rate and $\lim_{n\to\infty} o(n)/n = 0$, and $E_L(.)$ is the Gilbert lower bound on the error exponent [17]. Thus, the probability of error between any two (different) codewords ($p(m|m')$) will be upper bounded by $e^{-NE_L(R)+o(N)}$. Likewise the average probability of error $p_e := 1/M_f \sum_i p_{e|i}(e|i)$ can be lower bounded; here, $p(e|i)$ denotes the probability of error given that the ith message is transmitted. By the sphere-packing bound ([18], Chp. 5), $p_e \ge e^{N(E_{sp}(R)-o(N)/N)}$. We will use the sphere packing exponent $E_{sp}(R)$ to obtain negative results.

Let us fix the forward and reverse channel rates, $R_f = \log_2(M_f)/N_f$ and $R_r = \log_2(M_r)/N_r$. Thus the error exponent will not change as N_f and N_r increase. We penalize the codelengths in the forward and reverse channels by a possibly linear term in the sampling period; it then takes longer to send more bits; reliability competes with delay.

First, the case where the system (2) is noiseless is considered. Later, the noisy case will be considered.

C. Asymptotic Stability

The following theorem indicates that if the controller waits long enough, stability can be achieved.

Theorem 3.2: Suppose a scalar continuous-time system $\dot{x}_t = \xi x_t + b' u'_t$, with a bounded initial state x_0, is remotely controlled. Let the sampling period be a function of block lengths: $T_s = \alpha N_f + \beta N_r$; α, β be possibly depending on the codelengths, and the number of symbols in the state and control be $K = |\mathcal{X}'| = |\mathcal{U}| = |\mathcal{U}'|$. Let the rates $R_f = \log_2(K)/N_f$ and $R_r = \log_2(K)/N_r$ be kept constant as N_f, N_r grow. If the system and channel parameters satisfy

$$\begin{aligned}
&(2\xi\alpha - E_L^f(R_f))N_f + (2\xi\beta N_r) < 0,\\
&(2\xi\beta - E_L^r(R_f))N_r + (2\xi\alpha N_f) < 0,\\
&K = e^{N_f R_f} = e^{N_r R_r} > e^{\xi(\alpha N_f + \beta N_r)},
\end{aligned} \tag{6}$$

then, $\lim_{T_s\to\infty} E[x_{T_s}^2] = 0$. Further, let the minimum distance between two codes in $\mathcal{X}'$ be positive. Then, if any of the following holds

$$\begin{aligned}
&(2\xi\alpha - E_{sp}^f(R_f))N_f + (2\xi\beta N_r) > 0,\\
&(2\xi\beta - E_{sp}^r(R_f))N_r + (2\xi\alpha N_f) > 0,\\
&K = e^{N_f R_f} = e^{N_r R_r} < e^{\xi(\alpha N_f + \beta N_r)},
\end{aligned} \tag{7}$$

then, $\lim_{T_s\to\infty} E[x_{T_s}^2] = \infty$.

Proof: See [1]. $\diamond$

Now we make the following observations:

1) If there is no channel noise, the condition is the well-studied quantization condition: $K \ge |a|$.
2) Theorem 3.2 shows that the error exponents being positive (which is the case when rate is less than the capacity, $R < C$) does not directly lead to stability, and there needs to be a positive lower bound on the exponent. Thus the accurate measure is the reliability of the channel, not necessarily the capacity. This had been observed in [5].
3) Capacity is a sufficient measure if: (i) the error exponent is infinite, as in a digital noiseless channel, so long as $R < C$; (ii) there is no cost associated with the number of channel uses per sampling period, i.e., αN_f and βN_r are kept constant with growing N_f, N_r.
4) The set of stabilizing rates could be empty. For instance, in case there is no noise in the reverse channel, we need $\log_2(|a|) < N E_{sp}^{f\,-1}(2\log_2(|a|))$.

In view of the above, the achievable rates satisfy the following inequalities:

$$\frac{1}{N_f}\xi(\alpha N_f + \beta N_r) < R_f < (E_L^f)^{-1}([2\xi\alpha] + [2\xi\beta N_r]/N_f)$$

$$\frac{1}{N_r}\xi(\alpha N_f + \beta N_r) < R_r < (E_L^r)^{-1}([2\xi\alpha]N_f/N_r + 2\xi\beta)$$

As an illustration of the rate regions, we use binary symmetric channels with cross-over probabilities 0.01, for which the Gilbert exponent is $E_L(R) = H^{-1}(\ln 2 - R)\ln(2\sqrt{p(1-p)})$, where H is the binary entropy function. We plot the achievable rate region in Fig. 2, where we take $\alpha = \beta = 0.1$.

D. Asymptotic Stability with Delay Restricted Codes

We now consider the original system (2) driven by i.i.d. noise, where the sampling period is finite, and further the amount of data to be sent over a sampling period is finite. In this case the asymptotic analysis becomes inapplicable, and we need a scheme with finite length codes sent per time stage. We know from Proposition 2.2 that the encodable set has to be unbounded, and we need to represent this with a finite (in an expectation or a deterministic sense) number of codewords. The controller has access to the plant dynamics; therefore, there is some side information available at the controller about the next value of the state, but this side information is not available at the sensor as in the Slepian-Wolf coding context. We introduce a new coding scheme for dynamic systems using this interpretation. The scheme

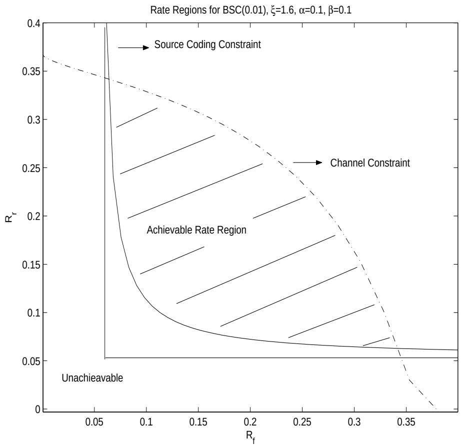

Fig. 2: Achievable rates over Binary Symmetric Channels. The capacities $C_f = C_r = 0.59$ are ineffective in the achievable rate analysis.

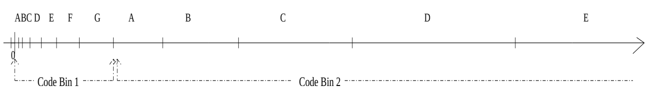

Fig. 3: Illustration of the binning approach to the joint source channel code; the symbols in a given CodeBin are represented by the same channel code -letters A, B, C, ... -; the mode symbol -1,2,3 ... -is carried by the side channel.

is based on binning [19], where we partition the state space into cosets, and transmit the coset of the symbol. We also assume that there is a possibly noisy side channel carrying the indices of the cosets. In [20] uniform binning was used in a decentralized linear system context. Here the system is centralized but the channel is noisy. Instead of uniform binning, we apply here logarithmic binning to satisfy the drift requirements. We quantify below the requirements needed by this scheme.

Suppose we have $K = 2^{N_f R_f}$ symbols that we will transmit over each sampling interval. We will partition the entire state space into bins and group K adjacent elements into one larger bin, indexed by I, and represent them by a single channel codebook. We refer to this ensemble of bins as a *CodeBin*. Hence, a total of $2^{N_f R_f}$ codewords are used to represent the entire state space (see Figure 3). Thus, CodeBin(I) $:= \{x : \delta_{I N_f R_f} \leq x < \delta_{(I+1) N_f R_f}\}$. This applies to the controller also, with the bin edges, δ, being scaled by $-a/b$.

We denote the bin indices by $\delta_{I,i}$, which means the edge belongs to CodeBin I and is represented by the ith channel codeword. We say the source code is in mode I, if the state is in CodeBin I (see Figure 4). The reconstruction value of each bin is assumed to be the midpoint, such that $Q_i = (1/2)(\delta_i + \delta_{i+1})$. We define $p^m(J|I)$ as the probability of error of CodeBin (mode of the quantizer) transmission from mode I to mode J through the side channel. This means the mode is erroneously transmitted from the plant to the

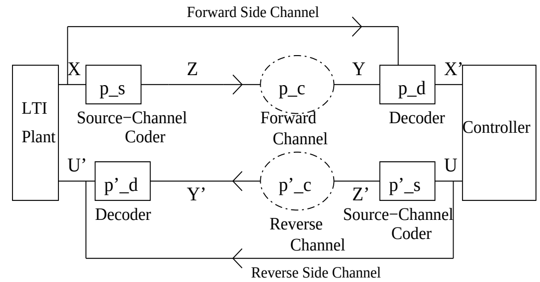

Fig. 4: Side channels carry the mode of the source mapping; the system is tolerant to the errors in the side channel as well.

controller if $J \neq I$. Likewise for the feedback channel we have $p^{m'}(L|J)$ as the side channel mapping.

Theorem 3.3: Suppose the scalar continuous-time system $dx_t = (\xi x_t + b' u'_t)dt + dB_t$, is remotely controlled. Let T_s be a sampling period which is function of block lengths: $T_s = \alpha N_f + \beta N_r$, α, β be possibly depending on the codelengths, and the number of symbols in the state and control be $K = |\mathcal{X}'| = |\mathcal{U}| = |\mathcal{U}'|$. Suppose the forward and reverse channel codes are of N_f and N_r bits long, and let the rates be $R_f = \log_2(K)/N_f$ and $R_r = \log_2(K)/N_r$. Define $T(\gamma, p^m, I)$, $U(\gamma, I)$, and Z respectively as

$$T := \sum_{L, L \neq I} \sum_J p^{m'}(L|J) p^m(J|I) 4 \gamma^{2(\max(0, N_f R_f(|L|-I))+1)}$$

$$U := \gamma^2 \sum_L \sum_J p^{m'}(L|J) p^m(J|I) \gamma^{2\max(0, N_f R_f(|L|-I+1))}$$

$$Z := K e^{-N_f E_L^f(R_f) - N_r E_L^r(R_r)} + e^{-N_f E_L^f(R_f)} + e^{-N_r E_L^r(R_r)}$$

If for some $\gamma > 1$, the forward, reverse and side channels satisfy the following

$$\limsup_{I \to \infty} U(\gamma, I) =: \bar{U}(\gamma) < \infty$$

$$\limsup_{I \to \infty} T(\gamma, p^m, I) =: T(\gamma, p^m) < 1$$

$$\gamma < 1 + 2(e^{-\xi})^{\alpha N_f + \beta N_r}$$

$$\cdot \sqrt{[(1-\epsilon) - 4Z 2^{N_f R_f} \bar{U}(\gamma) - T(\gamma, p^m)]},$$

then drift conditions are satisfied, and there exists a coding scheme leading to a finite second moment. The source coder is a symmetric logarithmic quantizer with expansion ratio γ, i.e., $|\delta_{i+1}| < \gamma |\delta_i|$.

Proof: See [1] $\diamond$

IV. Variable Length Encoding for Side Channels

In case we have noiseless side channels, there is a restriction on the number of channel uses for the side channels. If the restriction is only on the average number of channel uses, Huffman coding can be used to obtain a finite expected codelength, since the entropy of the invariant process is finite. However, in practice, there is a bound on the actual number of channel uses and not only on its average. Markovian stability theory can be used to show that even with such a restriction, stability can be achieved. In case an invariant density exists, the occurrence of high magnitude signals will be rare. We build on this in the following.

A. Variable Length Encoding for Side Channels

The controller and the sensor can send *side channel information* over variable periods by using variable length codes (such as uniquely decodable prefix codes). To achieve this, Codebins are generated according to the number of sampling periods required to send the side channel information, thus the effective sampling period will vary. However, in this case the drift analysis we employed earlier becomes inapplicable and one ought to use state-dependent drift conditions [21] to study stability. If the effective sampling period is $kT_s, k \in Z^+$, then the system will be open loop during kT_s seconds. These considerations lead to the following counterpart of Theorem 3.3 in the case of variable-length encoding.

Theorem 4.1: Let $U(\gamma) := \gamma^{2(N_f R_f + 1)}$, and

$$Z(k) := e^{-kN_f E_L^f (R_f/k) - kN_r E_L^r (R_r/k)} 2^{N_f R_f} + e^{-kN_f E_L^f (R_f/k)} + e^{-kN_r E_L^r (R_r/k)}.$$

If for some $\gamma > 1$, $\exists k_0 > 0$ such that, $\forall k > k_0$, $\forall x \in \{x : |x| > \gamma^{kN_f R_f}\delta_1\}$, the following holds

$$\gamma < 1 + 2(e^{-k\xi})^{\alpha N_f + \beta N_r} \cdot \sqrt{[(1-\epsilon) - 4Z(k)2^{N_f R_f}U(\gamma) - \frac{e^{2\xi k T_s} - 1}{2\xi \gamma^{2kN_f R_f}\delta_1^2}},$$

then drift conditions hold and there exists a coding scheme leading to a finite second moment. The source coder is a symmetric logarithmic quantizer with expansion ratio γ, i.e., $|\delta_{i+1}| < \gamma|\delta_i|$. There exists a solution for small enough ξ.

B. Relaxation of the Forward Side Channel

The control has access to the plant dynamics and there is already some side information available to it. This information might be useful in relaxing the conditions on the forward side channel.

Proposition 4.1: Suppose the forward channel error has a bound of Δ_f, the system noise has a bound of Δ_s, and the reverse channel is bounded. To achieve $\limsup_{t\to\infty} |x_t|^2 < \infty$, there is no need for a forward side channel.

Proof: The total uncertainty will be bounded by $\Delta =: |a|(\Delta_f + \Delta_r) + \Delta_s$. Clearly for large enough x, using a logarithmic quantizer, there exists a k_0 such that $\forall k > k_0$, $\gamma^{kN_r R_f}\delta_1 > 2\Delta$. Using binning [20], there will be no error in distinguishing between two codewords with the same coset. In this case if the two nearest bins sharing the same coset are spread out with a distance greater than $|a|(\Delta_f + \Delta_r) + \Delta_s$, then with only the coset information, the controller can find out the exact value of the bin, see [20]. $\diamond$

Unlike the transmissions from the controller, in general the plant cannot predict the control signal it will receive, since it does not have access to the decision policy at the controller. Therefore such a relaxation does not apply to the reverse channel.

Remark: The essential difficulty in code construction is to transmit sufficient information in a finite time over a finite symbol channel. In a continuous alphabet channel, this is not an issue, as is studied in [12], since for instance in the Gaussian channel case, arbitrary values can be coded in one channel use, and one can use high magnitude signals provided that the expected power remains finite. $\diamond$

V. Higher Dimensional Systems

We have not discussed here the multi-dimensional case, first because of page limitations, and second because the analysis in this case would be quite tedious. If the system can be transformed to a first-order Markovian system, since the drift conditions apply to any finite dimensional space, the coding schemes can be readily applied.

References

[1] S. Yüksel and T. Başar, "Memoryless Coding and Control for Stability of Linear Systems over Noisy Forward and Feedback Channels," preprint, DCL, UIUC, Sept. 2005.

[2] R. Bansal and T. Başar, "Simultaneous design of measurement and control strategies for stochastic systems with feedback." *Automatica,* 25:679-694, Sept. 1989

[3] V. S. Borkar and S. K. Mitter, "LQG Control with Communication Constraints" in *Kailath Festschrift*, pp. 365-373, Kluwer Academic Publishers, Boston 1997.

[4] N. Elia and S. Mitter, "Stabilization of linear systems with limited information," *IEEE Trans. Aut. Control*, 46:1384-1400, Sept. 2001.

[5] A. Sahai, *Any-Time Information Theory*, PhD thesis, MIT, 2000.

[6] A.V Savkin and I. R. Petersen, "An analogue of Shannon information theory for networked control systems: Stabilization via a noisy discrete channel", in Proc. *IEEE CDC*, Bahamas, Dec. 2004, pp. 4491-4496.

[7] F. Fagnani and S. Zampieri, "Stability analysis and synthesis for scalar linear systems with a quantized feedback", *IEEE Trans. Automatic Control*, 48: 1569-1584, Sept. 2003.

[8] N. Martins and M. Dahleh, "Fundamental limitations of disturbance attenuation in the presence of finite capacity feedback," in Proc. *Allerton Conference* 2004.

[9] A. Sahai and S. K. Mitter "The necessity and sufficiency of anytime capacity for control over a noisy communication link Part I: scalar systems", submitted to *IEEE Trans. Information Theory*.

[10] S. Tatikonda, A. Sahai, and S. K. Mitter, "LQG Control Problems Under Communication Constraints", in *Proc. IEEE CDC*, December 1998, FL, pp. 1165-1170.

[11] O.C. Imer, S. Yüksel, and T. Başar, "Optimal control of dynamical systems over unreliable communication links", in *Proc. NOLCOS 2004*, Stuttgart, Germany, 2004

[12] S. Yüksel and T. Başar, "Achievable Rates for Stability of LTI systems over noisy forward and feedback channels" in *Proc. CISS*, March 2005.

[13] A. Sahai, T. Simsek, "On the variable-delay reliability function of discrete memoryless channels with access to noisy feedback," in *Proc. Information Theory Workshop*, San Antonio, Texas, October 2004.

[14] S. Meyn and R. Tweedie, *Markov Chains and Stochastic Stability*, Springer Verlag, London, 1993.

[15] R. E. Blahut *Principles of Information Theory*. Course Notes at the UIUC, August 2002.

[16] B. Oksendal, *Stochastic Differential Equations*, Springer, 2003.

[17] J. K. Omura "On general Gilbert bounds," *IEEE Trans. Inform. Theory*, 19:661-666, Sept. 1973.

[18] R. G. Gallager, *Information Theory and Reliable Communication,* John Wiley Sons, 1968.

[19] S. S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (DISCUS): design and construction,", *IEEE Trans. Information Theory*, 49:626-643, March 2003.

[20] S. Yüksel and T. Başar, "Communication constraints for stability in decentralized multi-sensor control systems" submitted to *IEEE Trans. Automatic Control*, April 2005.

[21] S. Meyn and R. Tweedie, "State-dependent criteria for convergence of Markov chains," *Ann. Appl. Probab.* 4:149-168, 1994.