

REAL-TIME FINITE-RATE FEEDFORWARD CONTROL

*Sridevi V. Sarma, Munther A. Dahleh **

Massachusetts Institute of Technology
Laboratory of Information and Decision Systems
sree@mit.edu, dahleh@mit.edu

ABSTRACT

In this paper, we consider a set up in which the plant and controller are local to each other, but are together driven by a remote reference signal that is transmitted through a finite-rate noiseless channel. When control must be done over a communication channel, there is a fundamental tradeoff between allowing enough time for reconstruction of signals over the channel and achieving performance in finite-time. Most work in the area of control under communication constraints have addressed infinite-horizon control objectives (eg. stability, disturbance rejection). In this paper, we study two finite-horizon control tasks. The first task is to minimize a finite-horizon weighted tracking error between the remote system output and a reference command, which belongs to a class of signals. The second task is to navigate the state of the remote system from a nonzero initial condition, which lies in a bounded set, to as close to the origin as possible in finite-time. We compute lower and upper bounds on the worst-case performance for the two objectives as a function of the channel rate, time horizon, and system parameters. We achieve both lower bounds with noncausal coding schemes and show that imposing causality on the coding scheme limits achievable performance in both cases, but more severely for tracking. We illustrate how the bounds behave under various scenarios and show tradeoffs between time and performance accuracy.

I. INTRODUCTION

The classical control paradigm addresses problems where communication between one plant and one controller is essentially perfect. Today new problems in control over networked systems, whose components are connected via communication links that can be very noisy, induce delays, and have finite rate constraints, are emerging. Applications include remote navigation systems (deep-space and sea exploration) and multi-robot control systems (eg. aircraft and spacecraft formation flying control, coordinated control of land robots, control of multiple surface and underwater vehicles), where robots exchange data through communication channels that impose constraints on the design of coordination strategies.

*This research has been supported by AFOSR: 6892167.

In communication systems, problems entail designing channel encoders and decoders to reconstruct signals sent through noisy channels. Questions about *asymptotic* reconstruction are typically addressed. In control systems, problems often entail designing controllers to generate *real-time* desirable responses from a system. Therefore, when control must be done over a communication channel, there is a fundamental tradeoff between allowing enough time for reconstruction of signals and achieving stability and performance in finite-time.

Control under communication constraints is a research area of growing interest. Much work has focused on stability under finite-rate (or countable) feedback control, where the only excitation to the system is an unknown (but bounded) initial state condition [2], [3], [4], [5], [7], [10], [14], [15], [16], [18], [19], [20], [22], [23]. The questions posed involve conditions on the channel rate that will guarantee that the state of the system (or some function of the state) approach the origin/remains bounded as time goes to infinity. More recently, disturbance rejection limitations were derived for the same setting, assuming stochastic exogenous signals entering the system [17]. Although these studies have contributed greatly to our understanding of the interplay between communication and control, few studies address finite-horizon performance limitations under communication constraints.

A handful of recent studies explore the tradeoffs between finite-horizon performance and control complexity for linear systems and finite automata systems [6], [8], [11], [12]. In [12], Fagnani et al. consider the closed-loop system shown in Figure 1.

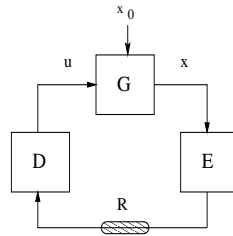


Fig. 1. Equivalent Closed-Loop System

G is a single-input multi-output discrete-time causal LTI system with unknown initial condition $x_0 \in \mathbb{R}^n$, which is a random vector with uniform probability density over a given bounded set $W \subset \mathbb{R}^n$. The feedback control law, $u \in \mathbb{R}$, must be generated over a finite-rate link that transmits exactly R bits per time step. Fagnani et al. ask the following question:

Given a subset V of W , find the minimum expected time, $E\{T_{(W,V)}\}$ that “traps” the state x_t in V for all $t \geq T$.

Fagnani et al. show that for any given $\beta > 0$,

$$\frac{E\{T_{(w,v)}\}}{\ln(C)} \leq \beta \Rightarrow \frac{LN}{\ln(C)} \geq \delta(\beta),$$

where $C = \frac{\mu[W]}{\mu[V]}$ (μ is the Lebesgue measure in \mathbb{R}^n) is a contraction rate that describes how small the target set is with respect to the starting set. L is a measure of the complexity of the coding scheme (E, D) and $\delta(\beta) = H_1 \beta w^{\frac{1}{\beta}}$, for some $w > 1$ and constant H_1 , which depends on the plant dynamics. See [12] for details. This result shows that demanding

smaller values of the expected minimum time to reach set V , results in requiring more complicated coding schemes.

In this paper we compute lower bounds for finite-horizon tracking and navigation objectives under finite-rate *feedforward* control. In our tracking problem, we compute the smallest allowable worst-case performance over a class of reference signals. In our navigation problem, we compute the smallest allowable ball around the origin that the state of the system can reach in T time steps, given that the initial condition lies in an ellipsoid. In turn, we also compute the minimum number of time steps, T , for the state of the system to reach a given ball around the origin. We also construct quantization/coding schemes to derive upper bounds on both performance objectives, and illustrate how imposing causality on the quantizer limits achievable performance. Our framework is deterministic and our lower bounds are independent of the complexity of the coding scheme.

II. TRACKING

In this section, we are interested in tracking a class of reference commands, r , over a finite-horizon and under finite-rate constraints. We consider the cascade of SISO discrete-time systems shown in Figure 2.

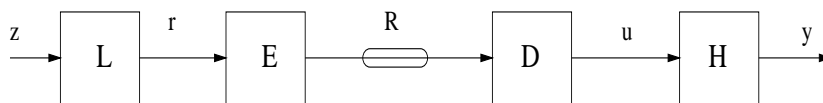


Fig. 2. Finite Horizon Tracking Set Up

Specifically,

- $z \in \mathbb{R}^T$ s.t. $\|z\|_2 \leq 1$,
- $L : \mathbb{R}^T \rightarrow \mathbb{R}^T$ is an invertible linear operator,
- $E : \mathbb{R}^T \rightarrow \{0, 1\}^{RT}$ is an arbitrary operator (encoder) that maps a real vector to a sequence of 2^{RT} binary symbols,
- R is the channel rate for the finite-rate noiseless channel that maps $\{0, 1\}^{RT} \rightarrow \{0, 1\}^{RT}$,
- $D : \{0, 1\}^{RT} \rightarrow \mathbb{R}^T$ is an arbitrary operator (decoder) that maps a sequence of 2^{RT} binary symbols to a real vector¹, and
- $H : \mathbb{R}^T \rightarrow \mathbb{R}^T$ is an invertible linear operator (model of remote plant and controller system at rest).

Note that L defines a class of signals, \mathcal{C}_r , that is generated from a unit ball in \mathbb{R}^T . Since L is linear, it maps the unit ball to a bounded ellipsoid. We set out to minimize a tracking error over all signals, r , in this class (worst-case analysis). Since the input and output signals have finite length, the following performance metric is computed over a finite-horizon: $\|W(y - r)\|_2^2$, where $W \in \mathbb{R}^T \times \mathbb{R}^T$ is a given full-rank weight matrix.

It is worth commenting that in the ideal case of perfect communication ($R = \infty$), it is possible to construct an encoder and decoder ($E = L^{-1}$ and $D = H^{-1}$) such that $\|W(y - r)\|_2^2 = 0 \quad \forall r \in \mathcal{C}_r$. However, with a finite-rate constraint, the control, u , can only take 2^{RT} values over a horizon of T time steps. Furthermore, with H being a one-to-one mapping,

¹Note that we denote the reconstructed signal as u , which is standard notation for a control input to a system.

the output, y , can only take 2^{RT} values over a horizon of T time steps. Therefore, it is not clear what level of performance is achievable over \mathcal{C}_r .

To understand tracking limitations under finite-rate feedforward control, we compute γ_{LB} and γ_{UB} , such that

$$\gamma_{LB} \leq \min_{(E,D)} \sup_{r \in \mathcal{C}_r} \|W(y - r)\|_2^2 \leq \gamma_{UB}.$$

Knowledge of γ_{LB} tells us that regardless of the encoder and decoder that we select, we can do no better than this lower bound. Therefore, we expect it to be independent of E and D . The upper bound tells us that there exists a coding scheme such that the worst case performance is always less than or equal to γ_{UB} . Therefore, to compute γ_{UB} , we need to construct an encoder and decoder and compute the corresponding worst-case performance. We compute γ_{LB} and γ_{UB} in the following sections.

II-A. A Lower Bound

In this section we derive the lower bound on worst-case performance.

Theorem II.1. *Given the tracking set up defined above, assume that $\det(W) \neq 0$, $\det(L) \neq 0$. Then,*

$$\gamma_{LB} = 2^{-2R} \{|\det(L)| |\det(W)|\}^{\frac{2}{T}}.$$

Proof.

The set of all possible commands, $\mathcal{C}_r \triangleq \{r \in \mathbb{R}^T | r = Lz, z'z \leq 1\} = \{r \in \mathbb{R}^T | (L^{-1}r)'(L^{-1}r) \leq 1\}$. \mathcal{C}_r is a bounded ellipsoid in \mathbb{R}^T centered at the origin with volume $\eta \det\{((L^{-1})'(L^{-1}))^{-0.5}\} = \eta |\det(L)|$, where η is the volume of a unit ball in \mathbb{R}^T .

Over a horizon T , the channel sends a total of RT bits which limits the control signal to take on no more than 2^{RT} values; and, since H is a one-to-one mapping, the channel limits the output to take on no more than 2^{RT} values. Consider a selection of outputs $y_1, y_2, \dots, y_{2^{RT}}$, which correspond to inputs $u_1, u_2, \dots, u_{2^{RT}}$, respectively. We must then map each $r \in \mathcal{C}_r$ to exactly one y_i $i = 1, 2, \dots, 2^{RT}$. Such a mapping induces a partition on \mathcal{C}_r . In particular, define $P_i = \{r \in \mathcal{C}_r | r \rightarrow y_i\}$ for $i = 1, 2, \dots, 2^{RT}$. Now, suppose that the selection $y_1, y_2, \dots, y_{2^{RT}}$ were chosen such that $\|W(y_i - r)\|_2^2 \leq \gamma$ for all $r \in P_i$, and for all i . Then necessarily $P_i \subseteq S_{y_i}^\gamma \triangleq \{r \in \mathbb{R}^T | (r - y_i)'W'W(r - y_i) \leq \gamma\}$. Note that $S_{y_i}^\gamma$ is a bounded ellipsoid in \mathbb{R}^T centered at point y_i with volume $\eta(\sqrt{\gamma})^T \det\{(W'W)^{-0.5}\} = \frac{\eta\sqrt{\gamma}^T}{|\det(W)|}$. See Figure 3 for an illustration.

Since $P_i \subseteq S_{y_i}^\gamma$ for each $i = 1, 2, \dots, 2^{RT}$, it is necessary that 2^{RT} bounded ellipsoids ($S_{y_i}^\gamma$) cover the set \mathcal{C}_r . This implies that $2^{RT} \times \text{volume}(S_{y_i}^\gamma) \geq \text{volume}(\mathcal{C}_r)$. Equivalently,

$$2^{RT} \geq \frac{\text{volume}(\mathcal{C}_r)}{\text{volume}(S_{y_i}^\gamma)} = \frac{|\det(L)| |\det(W)|}{(\sqrt{\gamma})^T}.$$

After rearranging terms, we get that $\gamma \geq 2^{-2R} \{|\det(L)| |\det(W)|\}^{\frac{2}{T}}$. ■

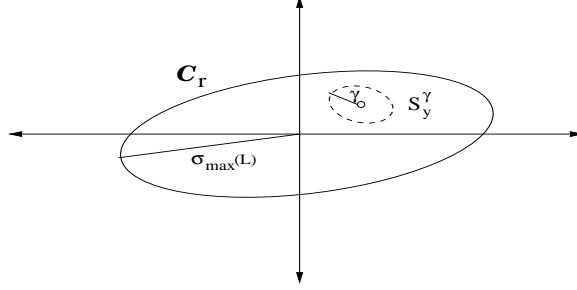


Fig. 3. Bounded Ellipsoids C_r and S_y^γ

Since we often consider classes of inputs generated from LTI systems, *i.e.*, L is LTI, we compute the lower bound for this case in the following corollary.

Corollary II.1. *Assume that $\det(W) \neq 0$, $\det(L) \neq 0$, and H is a one-to-one mapping. If L is a causal SISO LTI system with state-space description $L = ss(A_l, B_l, C_l, D_l)$, then*

$$\|W(y - r)\|_2^2 \geq 2^{-2R} (D_l)^2 \{|\det(W)|\}^{\frac{2}{T}}.$$

Proof. *If L is a SISO causal LTI with state-space description $L = ss(A_l, B_l, C_l, D_l)$, then for T time steps, it can be represented as a $T \times T$ lower triangular Toeplitz matrix operator, with all T eigenvalues equal to D_l . This implies that the $\{\det(L)\}^{\frac{2}{T}} = (D_l)^2$. ■*

As expected, γ_{LB} depends on L (class of reference commands), W (performance weights), T (performance horizon), and R (channel rate). It is helpful (as we will see when we compute upper bounds) to rewrite the lower bound in terms of the singular and eigenvalues of the matrix WL as follows: $\gamma_{LB} = 2^{-2R} \{\prod_{i=0}^{T-1} \sigma_i(WL)\}^{\frac{2}{T}} = 2^{-2R} \{\prod_{i=0}^{T-1} |\lambda_i(WL)|\}^{\frac{2}{T}}$.

When computing the lower bound, we made no assumptions on whether the encoder and decoder are causal or noncausal operators. If E and D are both noncausal, then our tracking problem reduces vector quantization problem with a deterministic error metric [1], where time need not enter the picture. At time $t = 0$, a noncausal decoder “knows” the future, that is, it can compute u_k for $k = 0, 1, \dots, T - 1$, which are represented by TR bits over a horizon of T steps. On the other hand, if E and D are causal, then the decoder only can compute u_k at time k and u_k is represented by at most $(k + 1)R$ bits.

In the following sections, we compute two upper bounds. One bound is computed by constructing a noncausal encoder and decoder, and the second upper bound is computed by constructing a causal coding scheme, which is more practical.

II-B. A Noncausal Upper Bound

In this section, we derive an upper bound on worst-case performance assuming that the encoder and decoder are noncausal. The upper bound is derived using a coding scheme that transmits information about the signal r in terms of a basis derived from the singular value decomposition (SVD) of the matrix WL .

Consider Figure 4 below. The encoder first uses the SVD of $WL = U\Sigma V^*$ to write $Wr = \sum_{i=0}^{T-1} \sigma_i \alpha_i u_i$, where σ_i is the i th singular value of WL , $\alpha_i = v_i^* w$ where v_i^* is the i th row vector of V^* , and u_i is the i th column vector of U . The α_i 's are then each converted into their binary representations and truncated according to the bit-allocation strategy denoted in $\mathcal{R} = (R_0, R_1, \dots, R_{T-1})$. In particular, a total of R_k bits are allocated to α_k , for $k = 0, 1, \dots, T$, and the only restriction is that $\sum_{k=0}^{T-1} R_k = TR$.

The decoder uses the bit-allocation strategy \mathcal{R} to reconstruct α and then uses the SVD of WL to compute \hat{r} from $\hat{\alpha}$. Finally, the decoder applies H^{-1} to \hat{r} to generate u . We call this $E - D$ construction the ‘‘SVD Coding Scheme.’’

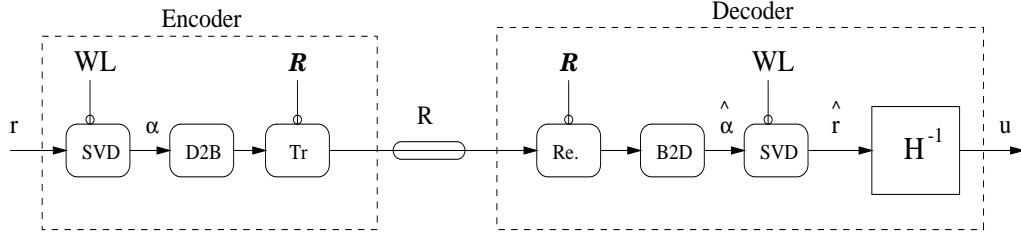


Fig. 4. SVD Coding Scheme

Note that with the above SVD coding scheme,

$$\begin{aligned}
\sup_{r \in \mathcal{C}_r} \|W(y - r)\|_2^2 &= \sup_{r \in \mathcal{C}_r} \|W(\hat{r} - r)\|_2^2 \\
&= \sup_{\{w \mid \|w\|_2 \leq 1\}} \|WL(\hat{w} - w)\|_2^2 \\
&= \sup_{\{\alpha \mid \|\alpha\|_2 \leq 1\}} \sum_{i=0}^{T-1} \sum_{j=0}^{T-1} (\hat{\alpha}_i - \alpha_i)(\hat{\alpha}_j - \alpha_j) \sigma_i \sigma_j (u_i' u_j) \\
&\leq \sup_{\{\alpha \mid \|\alpha\|_2 \leq 1\}} \sum_{i=0}^{T-1} |\alpha_i|^2 2^{-2R_i} \sigma_i^2 \\
&\leq \sigma_{max}(S)
\end{aligned}$$

where S is the diagonal matrix $diag(2^{-R_0} \sigma_0, 2^{-R_1} \sigma_1, \dots, 2^{-R_{T-1}} \sigma_{T-1})$.

To derive the upper bound using the above SVD coding scheme, we construct $\mathcal{R} = (R_0, R_1, \dots, R_{T-1})$ to solve the following optimization problem:

$$\begin{aligned}
&\min_{\mathcal{R}} \max_i 2^{-2R_i} \sigma_i^2 \\
&\text{s.t. } \sum_{i=0}^{T-1} R_i = TR \\
&\quad R_i \geq 0 \quad \forall i.
\end{aligned}$$

We allow the rates to take on non-integer values to solve for an optimal bit-allocation strategy. The resulting non-integer valued rates can be interpreted as average rates over time. The solution to the above problem forces

$2^{-2R_i^*} \sigma_i^2 = 2^{-2R_j^*} \sigma_j^2 \quad \forall i \neq j$, where $R_i^* = R + \log(\sigma_i) - \frac{1}{T} \sum_{j=0}^{T-1} \log(\sigma_j)$ for $i = 0, 1, \dots, T-1$. This gives us $\gamma_{UB} = 2^{-2R} \{\log(\prod_{i=0}^{T-1} \sigma_i)\}^{\frac{2}{T}} = \gamma_{LB}$.

Therefore, we achieve the lower bound with the noncausal SVD coding scheme above. However, this is not a practical implementation as the encoder must have access to the entire signal, r , at time 0.

II-C. A Causal Upper Bound

In this section, we derive an upper bound assuming that the encoder and decoder are causal and implement a practical coding scheme illustrated in Figure 5. In this scheme, the encoder is a quantizer parameterized by a rate matrix which dictates how bits are allocated to each component in the encoder's memory at each time step. Specifically, the rate matrix has the following form.

$$\mathcal{R} = \begin{bmatrix} R_{00} & 0 & 0 & \dots & \dots \\ R_{01} & R_{11} & 0 & 0 & \dots \\ R_{02} & R_{12} & R_{22} & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots & \ddots \\ R_{0,T-1} & R_{1,T-1} & R_{2,T-1} & \dots & R_{T-1,T-1} \end{bmatrix},$$

such that $\sum_j R_{ij} = R$ for $i = 0, 1, \dots, T-1$.

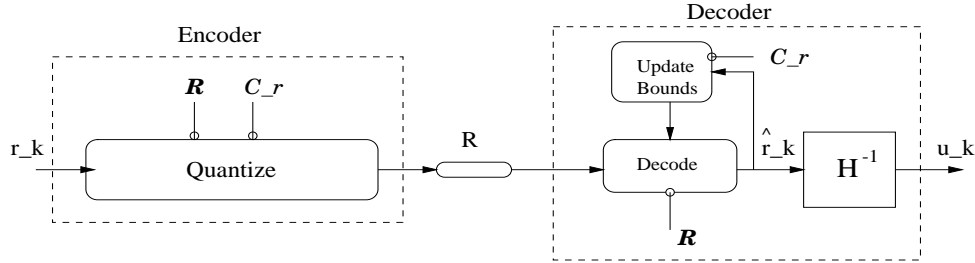


Fig. 5. Causal Coding Scheme

To understand how the above rate matrix dictates a bit-allocation strategy, let $\hat{r}_i(j)$ be the quantized estimate of r_i at time j . Then, \mathcal{R} determines that at time $t = 0$, R_{00} bits are used to quantize r_0 to produce $\hat{r}_0(0)$. At time $t = 1$, an *additional* R_{01} bits are used to quantize r_0 to produce $\hat{r}_0(1)$, and R_{11} bits are used to quantize r_1 to produce $\hat{r}_1(1)$, and so on. The accuracy of $\hat{r}_i(j)$ is within $\pm M_i 2^{-\sum_{k=i}^j R_{ik}}$ of r_i for all $i \geq 0$, where $2M_i$ is the length of the interval in which r_i belongs to as computed by the decoder at time step j . The decoder computes M_i from its past inputs and from constraints imposed on $r \in \mathcal{C}_r$. These computations will become clear below.

Consider the following example for horizon length $T = 2$,

$$Q \triangleq (L^{-1})'(L^{-1}) = \begin{bmatrix} q_{00} & q_{01} \\ q_{01} & q_{11} \end{bmatrix}, \quad W'W = \begin{bmatrix} w_{00} & w_{01} \\ w_{01} & w_{11} \end{bmatrix},$$

for some q_{01} and $w_{01} \in \mathbb{R}$.

At time $t = 0$, the encoder receives r_0 which has magnitude less than or equal to M_0 ,² and its quantization region is on the interval $\{-M_0, M_0\}$ which is divided into $2^{R_{00}}$ equal intervals as shown in Figure 6. The union of the representatives for each region comprise the range of E at time 0. Specifically,

$$E(r_0; t = 0) = nM_02^{(1-R_{00})}$$

for $nM_02^{(1-R_{00})} < r_0 \leq (n+2)M_02^{(1-R_{00})}$, $n = -2^{(R_{00}-1)}, -2^{(R_{00}-1)}+2, \dots, 0, 2, \dots, 2^{(R_{00}-2)}$. Thus, when the encoder receives r_0 it outputs the right boundary value of the interval in which r_0 falls, which is represented by R_{00} bits. The decoder then receives $\hat{r}_0(0)$ and updates its bounds for both r_0 and r_1 to prepare for its next input.

The bounds for r_1 are generated from the constraints imposed by the set $\mathcal{C}_r \triangleq \{r \in \mathbb{R}^T | r = Lz, z'z \leq 1\}$. For $T = 2$, \mathcal{C}_r forces $|r_1 + \frac{q_{01}}{q_{11}}r_0| \leq \frac{\sqrt{1+\mu r_0^2}}{\sqrt{q_{11}}}$ given r_0 , where $\mu = \frac{q_{01}^2}{q_{11}} - q_{00}$. Therefore, after time step 0, we have that $l_0(0) \leq r_0 \leq u_0(0)$, $l_1(0) \leq r_1 \leq u_1(0)$,³ where

$$\begin{aligned} l_0(0) &= \hat{r}_0(0) - 2^{-R_{00}}M_0, \\ u_0(0) &= \hat{r}_0(0) + 2^{-R_{00}}M_0, \\ l_1(0) &= \min\left\{\frac{-q_{01}}{q_{11}}l_0(0) - \frac{\sqrt{1+\mu(l_0(0))^2}}{\sqrt{q_{11}}}, \frac{-q_{01}}{q_{11}}u_0(0) - \frac{\sqrt{1+\mu(u_0(0))^2}}{\sqrt{q_{11}}}\right\}, \\ u_1(0) &= \max\left\{\frac{-q_{01}}{q_{11}}l_0(0) + \frac{\sqrt{1+\mu(l_0(0))^2}}{\sqrt{q_{11}}}, \frac{-q_{01}}{q_{11}}u_0(0) + \frac{\sqrt{1+\mu(u_0(0))^2}}{\sqrt{q_{11}}}\right\}. \end{aligned}$$

At time $t = 1$, the encoder further quantizes r_0 by dividing the interval $[l_0(0), u_0(0)]$ into $2^{R_{01}}$ equal length intervals, and sends the representative of the new interval in which r_0 lies, denoted $\hat{r}_0(1)$. The encoder then uses the remainder R_{11} bits to quantize r_1 by dividing the updated interval $[l_1(1), u_1(1)]$ into $2^{R_{11}}$ intervals, and then sends the center of the interval in which r_1 falls. It is straightforward to compute

$$\begin{aligned} l_0(1) &= \hat{r}_0(1) - 2^{-(R_{00}+R_{01})}M_0, \\ u_0(1) &= \hat{r}_0(1) + 2^{-(R_{00}+R_{01})}M_0, \\ l_1(1) &= \min\left\{\frac{-q_{01}}{q_{11}}l_0(1) - \frac{\sqrt{1+\mu l_0(1)^2}}{\sqrt{q_{11}}}, \frac{-q_{01}}{q_{11}}u_0(1) - \frac{\sqrt{1+\mu u_0(1)^2}}{\sqrt{q_{11}}}\right\}, \\ u_1(1) &= \max\left\{\frac{-q_{01}}{q_{11}}l_0(1) + \frac{\sqrt{1+\mu l_0(1)^2}}{\sqrt{q_{11}}}, \frac{-q_{01}}{q_{11}}u_0(1) + \frac{\sqrt{1+\mu u_0(1)^2}}{\sqrt{q_{11}}}\right\}. \end{aligned}$$

It is important to note that past allocation of bits to r_0 at time $t = 1$ allows for a tradeoff of a smaller known interval in which r_1 lies, $[l_1(1), u_1(1)]$, and finer quantization of the interval itself.

Since \mathcal{C}_r is an ellipsoid in \mathbb{R}^T , knowledge of r_j impacts the lower and upper bounds on r_k for $k \neq j$. Therefore, it appears that allocating bits to past signal components may be advantageous. It turns out however, that when \mathcal{C}_r is any ellipsoid, it is always optimal in our worst-case setting to allocate all R bits to the current value r_k at time k , *i.e.*, it is *never* optimal to allocate bits to past values r_0, r_1, \dots, r_{k-1} to quantize r_k .

Theorem II.2. *Consider the tracking problem that implements the causal coding scheme above parameterized by a rate*

²The encoder can compute $M_0 = \{|\{\Sigma_l\}_{11}\{U_l\}_{11}\}$ from the SVD composition of $L = U_l \Sigma_l V_l^*$

³Our notation of $l_j(k)$ is the lower bound for component r_j right after time step k as computed by the decoder. The upper bound is similarly defined.

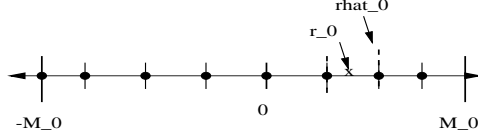


Fig. 6. Quantization region at time $t = 0$

matrix \mathcal{R} . Then, the optimal solution to $\min_{\mathcal{R}} \sup_{r \in \mathcal{C}_r} \|W(r - \hat{r})\|_2^2$ is a diagonal rate matrix, $\mathcal{R}^* = RI$ and $r^* = 0$, the centroid of \mathcal{C}_r .

Proof.

We prove the theorem assuming L is diagonal for a simpler read, however, the general case holds and can be proven using an identical argument outlined below.

- 1) We compute an upper bound, U , for $\sup_{r \in \mathcal{C}_r} \|W(r - \hat{r})\|_2^2$.
- 2) We show that $r^* = \mathbf{0}$ achieves U and compute $\|W(r^* - \hat{r}^*)\|_2^2$ as a function of the rate matrix.
- 3) We show that minimizing $\sup_{r \in \mathcal{C}_r} \|W(r - \hat{r})\|_2^2 = \|W(r^* - \hat{r}^*)\|_2^2$ over all possible rate matrices gives rise to a diagonal rate matrix.

Consider the set up for where L (and hence Q) is diagonal.

$$Q \triangleq (L^{-1})'(L^{-1}) = \begin{bmatrix} q_0 & & & & \\ & q_1 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & q_{T-1} \end{bmatrix},$$

$$W'W = \begin{bmatrix} w_{00} & w_{01} & \dots & w_{0,T-1} \\ w_{01} & w_{11} & \dots & w_{1,T-1} \\ \vdots & & & \vdots \\ w_{T-1,1} & w_{T-1,1} & \dots & w_{T-1,T-1} \end{bmatrix}.$$

Then, $\mathcal{C}_r = \{r \in \mathbb{R}^2 \mid \sum_{i=0}^{T-1} q_i r_i^2 \leq 1\}$. Now, since we transmit the signal r in a causal manner, the class of signals forces $|r_i| \leq M_i$, where

$$M_i = \sqrt{\frac{1 - \sum_{k=0}^{i-1} q_k r_k^2}{q_i}}, \quad i = 1, 2, \dots, T-1,$$

and $M_0 = |\{\Sigma_l\}_{11}\{U_l\}_{11}|$ from the SVD composition of $L = U_l \Sigma_l V_l^*$. At the decoder end, at time i , the transmitted signal (r_0, r_1, \dots, r_i) is not perfectly known. Therefore, it computes each bound as

$$M_i = \sqrt{\frac{1 - \sum_{k=0}^{i-1} q_k (\hat{r}_k(i) - M_k 2^{(1-R_k(i))})^2}{q_i}},$$

where $R_k(i) = \sum_{j=k}^i R_{kj}$. Note that M_i is a function of $\hat{r}_0(i), \hat{r}_1(i), \dots, \hat{r}_i(i)$ and \mathcal{R}_i , which is the first $i \times i$ elements of the full $T \times T$ rate matrix. We suppress these dependencies for an easier read. The above bound assumes, without loss of generality, the following encoder operator at time j :

$$E(r_i; j) = n_i M_i 2^{R_i(j)}$$

$$n_i \in S_i \triangleq \{-2^{R_i(j)-1}, -2^{R_i(j)-1} + 2, \dots, 0, 2, \dots, 2^{R_i(j)-2}\},$$

for $n_i M_i 2^{R_i(j)} \leq r_i < (n_i + 2) M_i 2^{R_i(j)}$. To see where the above expression for M_i comes from, consider a simple example where $T = 2$ and $R = 3$ illustrated in Figure 7. Given $\hat{r}_0(1)$ at $t = 1$, the uncertainty set as computed by the decoder for r shrinks from \mathcal{C}_r to one of $2^{(R_{00}+R_{01})}$ “strips” of the 2D-ellipsoid. For example, for $r_0 \in (4M_0 2^{-(R_{00}+R_{01})}, 6M_0 2^{-(R_{00}+R_{01})}]$, the uncertainty computed by the decoder is the shaded strip shown in Figure 7, and the corresponding $\hat{r}_0(1) = 6M_0 2^{-(R_{00}+R_{01})}$. As shown in Figure 7, the bound $M_1 = \sqrt{\frac{1-q_0(\hat{r}_0(1)-M_0 2^{(1-(R_{00}+R_{01}))})^2}{q_1}}$.

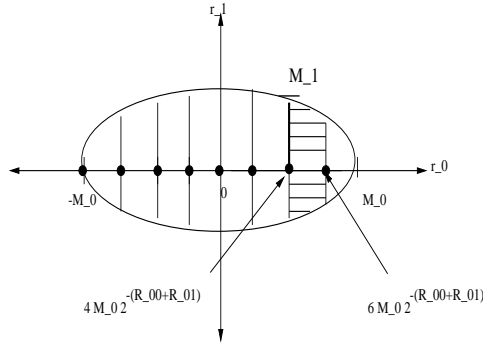


Fig. 7. Intervals for r_1 when $r_0 \in (4M_0 2^{-(R_{00}+R_{01})}, 6M_0 2^{-(R_{00}+R_{01})}]$.

Now, we can begin computing an upper bound on the cost function. Note that

$$\begin{aligned} \|W(r - \hat{r})\|_2^2 &\leq \sum_{i=0}^{T-1} w_{ii} (M_i 2^{-R_{ii}})^2 \\ &+ \sum_{i=0}^{T-1} \sum_{j=i+1}^{T-1} w_{ij} [M_i 2^{-R_{ii}}] [M_j 2^{-R_{jj}}]. \end{aligned}$$

Then, the following is an upper bound on the cost.

$$\begin{aligned} \sup_{r \in \mathcal{C}_r} \|W(r - \hat{r})\|_2^2 &\leq \sup_{r \in \mathcal{C}_r} \sum_{i=0}^{T-1} w_{ii} (M_i 2^{-R_{ii}})^2 \\ &+ 2 \sum_{i=1}^{T-1} \sum_{j=i+1}^{T-1} w_{ij} M_i M_j 2^{-(R_{ii}+R_{jj})} \\ &\leq \end{aligned}$$

$$\begin{aligned} & \sup_{n_i \in S_i} \sum_{i=0}^{T-1} w_{ii} \left(\sqrt{\frac{1 - \sum_{k=0}^{i-1} q_k ((n_k - 1) M_k 2^{(1-R_k(i))})^2}{q_i}} 2^{-R_{ii}} \right)^2 \\ & + 2 \sum_{i=1}^{T-1} \sum_{j=i+1}^{T-1} w_{ij} \sqrt{\frac{1 - \sum_{k=0}^{i-1} q_k ((n_k - 1) M_k 2^{(1-R_k(i))})^2}{q_i}} \dots \\ & \sqrt{\frac{1 - \sum_{k=0}^{j-1} q_k ((n_k - 1) M_k 2^{(1-R_k(j))})^2}{q_j}} 2^{-(R_{ii} + R_{jj})}. \end{aligned}$$

We rewrite the above inequality as

$$\begin{aligned} \sup_{r \in \mathcal{C}_r} \|W(r - \hat{r})\|_2^2 & \leq \sup_{n_i \in S_i} \left\{ \sum_{i=0}^{T-1} C_{1,i} M_i^2(n_i) + \right. \\ & \left. 2 \sum_{i=1}^{T-1} \sum_{j=i+1}^{T-1} C_{2,ij} M_i(n_i) M_j(n_j) \right\}. \end{aligned}$$

where $M_0(n_i) = |\{\Sigma_l\}_{11}\{U_l\}_{11}|$, and $M_i(n_i) = \sqrt{\frac{1 - \sum_{k=0}^{i-1} q_k ((n_k - 1) M_k 2^{(1-R_k(i))})^2}{q_i}}$, for $i = 1, \dots, T-1$, $C_{1,i} = w_{ii} 2^{-2R_{ii}}$, and $C_{2,ij} = 2w_{ij} 2^{-(R_{ii} + R_{jj})}$. We can then take partial derivatives of $f(n_0, \dots, n_{T-1}) = \sum_{i=1}^{T-1} C_{1,i} M_i^2(n_i) + 2 \sum_{i=1}^{T-1} \sum_{j=i+1}^{T-1} C_{2,ij} M_i(n_i) M_j(n_j)$ with respect to n_i assuming n_i is continuous for $i = 0, 1, \dots, T-1$, and set each to zero to see if the corresponding n_i^* 's are integers. When we take partial derivatives, we get

$$\frac{\delta f}{\delta n_i} = 2 \sum_{i=1}^{T-1} C_{1,i} M_i(n_i) \frac{\delta M_i(n_i)}{\delta n_i} + 2 \sum_{j=i+1}^{T-1} C_{2,ij} \frac{\delta M_i(n_i)}{\delta n_i} M_j(n_j).$$

After some algebra, one can show that $n_i^* = 1$ results in $\frac{\delta f}{\delta n_i} = 0$ and $\frac{\delta^2 f}{\delta n_i^2} |_{n_i^*} < 0$, for $i = 0, 1, \dots, T-1$. Since the n_i^* 's are integer maxes, $\hat{r}_i^*(i) = M_i 2^{-R_{ii}}$, and

$$\begin{aligned} \sup_{r \in \mathcal{C}_r} \|W(r - \hat{r})\|_2^2 & \leq \sum_{i=0}^{T-1} w_{ii} \left(\frac{2^{-R_{ii}}}{\sqrt{q_i}} \right)^2 \\ & + 2 \sum_{i=1}^{T-1} \sum_{j=i+1}^{T-1} w_{ij} \sqrt{\frac{1}{q_i q_j}} 2^{-(R_{ii} + R_{jj})} \triangleq U. \end{aligned}$$

Finally, it is straightforward to show that when $r_i^* = 0$, for $i = 0, 1, \dots, T-1$, then $\hat{r}_i^*(i) = M_i 2^{-R_{ii}}$, and $\|W(r^* - \hat{r}^*)\|_2^2 = U$. Therefore, $\sup_{r \in \mathcal{C}_r} \|W(r - \hat{r})\|_2^2$ for any given rate matrix occurs at the centroid of the ellipsoid.

Now, we minimize $\sup_{r \in \mathcal{C}_r} \|W(r - \hat{r})\|_2^2$ over all rate matrices as follows:

$$\begin{aligned} \min_{\mathcal{R}} \quad & \sum_{i=1}^{T-1} w_{ii} \frac{2^{-2R_{ii}}}{q_i} + 2 \sum_{i=0}^{T-1} \sum_{j=i+1}^{T-1} w_{ij} \frac{2^{-(R_{ii} + R_{jj})}}{q_i q_j} \\ \text{s.t.} \quad & \sum_{i=0}^j R_{ij} = R \quad \text{for } j = 0, 1, \dots, T-1 \\ & R_{ij} > 0 \quad \text{for all } i, j. \end{aligned}$$

By inspection, the optimal rate matrix is diagonal. ■

One can show that past allocation may be useful for signal sets that are nonsymmetric or finite or for different performance metrics. Also, if the encoder has access to the entire signal r at time $t = 0$, but still transmits R bits per time step, past allocation improves performance as shown in our navigation problem (section III).

II-D. Comparison of Bounds

We now compare the lower bound and causal upper bound to each other for different LTI causal systems

$L = ss(A_l, B_l, C_l, D_l)$, and for different time horizons T . We consider diagonal weight matrices

$W = diag(\lambda_0, \lambda_1, \dots, \lambda_{T-1})$ with $|\lambda_i| \leq 1, \forall i$, and fix the rate $R = 10$. Under such conditions, we note that $\gamma_{LB} = 2^{-2R} (D_l)^2 \{ \prod_{i=0}^{T-1} |\lambda_i| \}^{\frac{2}{T}}$.

Figures 8 and 9 illustrate the bounds for various scenarios, and we make a few observations.

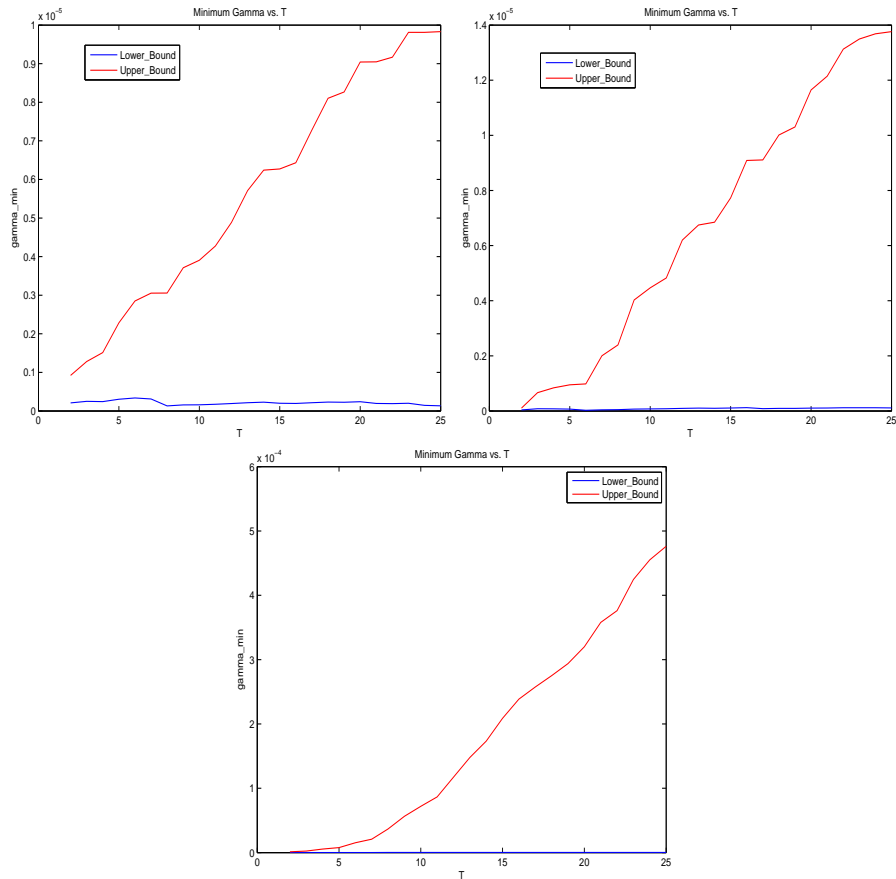


Fig. 8. Top: Bounds for $L = ss(0.01, 0.01, 1, 1)$ and random weights, Middle: Bounds for $L = ss(0.3, 0.3, 1, 1)$ and random weights, Bottom: Bounds for $L = ss(0.9, 0.9, 1, 1)$ and random weights.

- When the eigenvalues of W are chosen randomly from an i.i.d. process, then we see that the lower bound plateaus for large T . To see why this makes sense, we compute the expected value and variance of γ_{LB} with $L = ss(0.9, 0.9, 1, 1)$

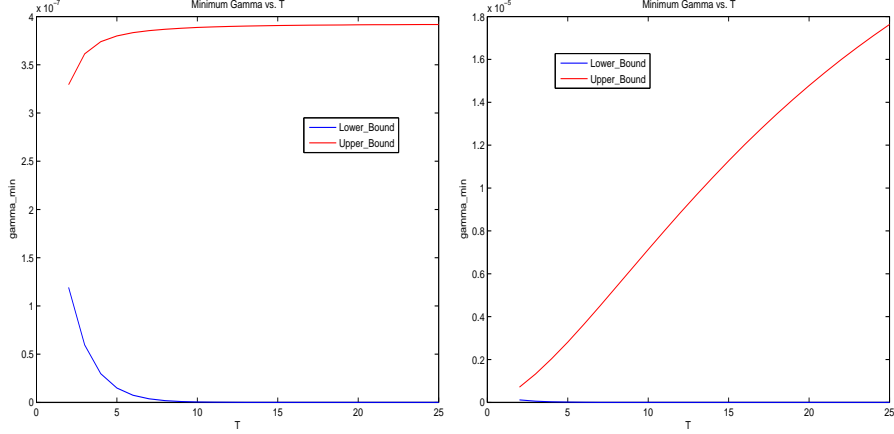


Fig. 9. Top: Bounds for $L = ss(0.1, 0.1, 1, 1)$ and decaying weights, Bottom: Bounds for $L = ss(0.9, 0.9, 1, 1)$ and decaying weights

and observe their behaviour as T grows large.

$$E\{\gamma_{LB}\} = 2^{-2R} E\left\{\prod_{i=0}^{T-1} |\lambda_i|^{\frac{2}{T}}\right\}.$$

Since the λ_i 's are all independent, we get that the expectation of the product is the product of the expectations. In addition, since the λ_i 's are identically distributed their expectations are all equal, and we get the following.

$$E\{\gamma_{LB}\} = 2^{-2R} \prod_{i=0}^{T-1} E\{|\lambda_i|^{\frac{2}{T}}\} = 2^{-2R} E^T\{|\lambda|^{\frac{2}{T}}\}.$$

Now, as $T \rightarrow \infty$, we get that $E\{\gamma_{LB}\} \rightarrow 2^{-2R}$ since $|\lambda_i| \leq 1$. Next, we compute the variance of γ_{LB} by first computing $E\{\gamma_{LB}^2\}$.

$$E\{\gamma_{LB}^2\} = 2^{-4R} E\left\{\prod_{i=0}^{T-1} |\lambda_i|^{\frac{4}{T}}\right\} = 2^{-4R} E^T\{|\lambda|^{\frac{4}{T}}\}.$$

We see that as $T \rightarrow \infty$, we get that $E\{\gamma_{LB}^2\} \rightarrow 2^{-4R}$, which implies that $var(\gamma_{LB}) = E\{\gamma_{LB}^2\} - E^2\{\gamma_{LB}\} \rightarrow 0$. Therefore, for large T , we expect that the lower bound approaches 2^{-2R} .

- When the eigenvalues of W are exponentially decaying, *i.e.*, $\lambda_i = (\beta)^i$ for $i = 0, 1, \dots, T - 1$, and for some $0 < \beta < 1$, then the lower bound approaches 0 as $T \rightarrow \infty$. This can be verified by showing that the ratio $\frac{\gamma_{LB}(T+1)}{\gamma_{LB}(T)} = \frac{\{\prod_{i=0}^T \beta^i\}^{\frac{2}{T+1}}}{\{\prod_{i=0}^{T-1} \beta^i\}^{\frac{2}{T}}} = \beta < 1$. The causal upper bound increases but plateaus for large T , but at a much slower rate when the pole of L is close to the unit disk.
- The causal upper bound is closer to the lower bound when the pole of the LTI system of L or that which generates a noncausal L is close to the origin than if the pole is close to the unit disk.

III. NAVIGATION

In this section, we assume that the remote system has some unknown initial condition x_0 which lies in a known bounded ellipsoid in \mathbb{R}^n . We want to steer the state of the remote system as close to the origin as possible under the constraint that

the control input can take on at most 2^{RT} values after T time steps, *i.e.*, the command is transmitted through a finite-rate noiseless channel as shown in Figure 10.

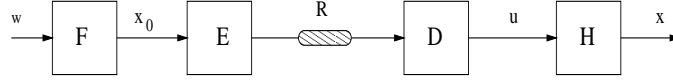


Fig. 10. Finite Horizon Navigation Set Up

Specifically,

- $w \in \mathbb{R}^n$ *s.t.* $\|w\|_2 \leq 1$,
- $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a linear operator,
- $E : \mathbb{R}^n \rightarrow \{0, 1\}^{RT}$ is an arbitrary operator (encoder) that maps a real vector to a sequence of 2^{RT} binary symbols,
- R is the channel rate for the finite-rate noiseless channel that maps $\{0, 1\}^{RT} \rightarrow \{0, 1\}^{RT}$,
- $D : \{0, 1\}^{RT} \rightarrow \mathbb{R}^T$ is an arbitrary operator (decoder) that maps a sequence of 2^{RT} binary symbols to a real vector, and
- H is a causal SISO LTI system with state-space representation $H = ss(A, B, I, 0)$ with (A, B) reachable and A is full rank.

We investigate two performance metrics for our navigation problem. The first metric fixes the horizon T and looks for the smallest ball around the origin that the state can reach. In particular, we compute lower and upper bounds such that

$$\gamma_{LB} \leq \min_{(E,D)} \sup_{x_0 \in \mathcal{C}_{x_0}} \|x_T\|_2^2 \leq \gamma_{UB},$$

where $\mathcal{C}_{x_0} \triangleq \{x_0 \in \mathbb{R}^n | x_0 = Fw, w'w \leq 1\} = \{x_0 \in \mathbb{R}^n | (F^{-1}x_0)'(F^{-1}x_0) \leq 1\}$. Knowledge of γ_{LB} tells us that regardless of the encoder and decoder that we select, we can do no better than this lower bound. Therefore, we expect it to be independent of E and D . The upper bound tells us that there exists a coding scheme (and encoder and decoder) such that the worst case performance is always less than or equal to γ_{UB} . Therefore, to compute γ_{UB} , we need to construct an encoder and decoder and compute the corresponding worst-case performance.

The second metric fixes γ and looks for the minimum time it takes for the state vector to reach a ball of size γ . Therefore, we fix γ and then look for T_{min} such that $\min_{(E,D)} \sup_{x_0 \in \mathcal{C}_{x_0}} \|x_T\|_2^2 \leq \gamma$ for all $T \geq T_{min}$.

We compute the bounds in the following sections.

III-A. A Lower Bound

In this section we derive the lower bound on worst-case navigation performance.

Theorem III.1. *Consider a reachable SISO LTI causal DT system, $H = ss(A, B, I, 0)$ where A is full-rank and with initial condition $x_0 \in \{x \in \mathbb{R}^n | x = Fw, w \in \mathbb{R}^n, w'w \leq 1\}$, for a given full-rank F . If the control input is constrained to take on at most 2^{RT} values after $T \geq n$ time steps, then $\|x_T\|_2^2 \leq \gamma$ only if*

$$\gamma \geq 2^{-2RT/n} \{|\det(F)| |\det(A^T)|\}^{\frac{2}{n}}.$$

Proof.

We first observe that the set of all possible initial conditions, $\mathcal{C}_{x_0} \triangleq \{x_0 \in \mathbb{R}^n | x_0 = Fw, w'w \leq 1\} = \{x_0 \in \mathbb{R}^n | (F^{-1}x_0)'(F^{-1}x_0) \leq 1\}$. \mathcal{C}_{x_0} is a bounded ellipsoid in \mathbb{R}^n centered at the origin with volume $\eta \det\{((F^{-1})'(F^{-1}))^{-0.5}\} = \eta|\det(F)|$, where η is the volume of a unit ball in \mathbb{R}^n . Since A is full rank, the set created by multiplying x_0 on the left by A^T is also a bounded ellipsoid in \mathbb{R}^n . Denote this new set as $\mathcal{C}_{A^T x_0} \triangleq \{y \in \mathbb{R}^n | y = A^T x_0, x_0 \in \mathcal{C}_{x_0}\}$.

Next, note that $\|x_T\|_2^2 = \|A^T x_0 + Mu\|_2^2$, where A^T is the $(T)^{th}$ power of the matrix A , and $M = \begin{bmatrix} A^{T-2}B & A^{T-3}B & \dots & AB & B \end{bmatrix}$ is the reachability matrix of system H . Since M is full row rank, one can pick any value $y \in \mathbb{R}^n$ and find a $u \in \mathbb{R}^T$ such that $y = Mu$. Over a horizon T , the channel sends a total of RT bits which limits the control signal to take on no more than 2^{RT} values. Therefore, $y = Mu$ can only take on 2^{RT} values.

Consider a selection of signals $y_1, y_2, \dots, y_{2^{RT}}$, which correspond to inputs $u_1, u_2, \dots, u_{2^{RT}}$, respectively. We must then map each $A^T x_0 \in \mathcal{C}_{A^T x_0}$ to exactly one y_i $i = 1, 2, \dots, 2^{RT}$. Such a mapping induces a partition on $\mathcal{C}_{A^T x_0}$. In particular, define $P_i = \{v \in \mathcal{C}_{A^T x_0} | v \rightarrow y_i\}$ for $i = 1, 2, \dots, 2^{RT}$.

Now, suppose that the selection $y_1, y_2, \dots, y_{2^{RT}}$ were chosen such that $\|A^T x_0 + y_i\|_2^2 \leq \gamma$ for all $A^T x_0 \in P_i$, and for all i . Then necessarily $P_i \subseteq \mathcal{B}_\gamma$, where \mathcal{B}_γ is a ball in \mathbb{R}^n of size γ with volume $\sqrt{\gamma}^n$.

Since $P_i \subseteq \mathcal{B}_\gamma$ for each $i = 1, 2, \dots, 2^{RT}$, it is necessary that 2^{RT} γ -balls cover the set $\mathcal{C}_{A^T x_0}$. This implies that $2^{RT} \times \text{volume}(\mathcal{B}_\gamma) \geq \text{volume}(\mathcal{C}_{A^T x_0})$. Equivalently,

$$2^{RT} \geq \frac{\text{volume}(\mathcal{C}_{x_0})}{\text{volume}(\mathcal{B}_\gamma)} = \frac{|\det(F)| |\det(A^T)|}{(\sqrt{\gamma})^n}.$$

After rearranging terms, we get that $\gamma \geq 2^{-2RT/n} \{|\det(F)| |\det(A^T)|\}^{\frac{2}{n}}$. ■

Since we often consider classes of inputs generated from LTI systems, i.e., F is LTI, we compute the lower bound for this case in the following corollary.

Corollary III.1. *Given the navigation set up defined above, if F is generated by a causal SISO LTI system with state-space description $ss(A_f, B_f, C_f, D_f)$, then*

$$\|x_T\|_2^2 \geq 2^{-2RT/n} (D_f)^2 \{|\det(A^T)|\}^{\frac{2}{n}}.$$

Proof. *If F is generated by a SISO causal LTI system with state-space description $F = ss(A_f, B_f, C_f, D_f)$, then it can be represented as a $n \times n$ lower triangular Toeplitz matrix operator, with all n eigenvalues equal to D_f . This implies that the $\{|\det(F)|\}^{\frac{2}{n}} = (D_f)^2$.* ■

We now comment on γ_{LB} .

- γ_{LB} depends on F (class of initial conditions), n (the dimension of the system state), T (performance horizon), and R (channel rate).

- If A is stable, then necessarily the lower bound approaches 0 and T grows large since A^T is approaching 0. In this case, for large T one does not need to apply a control input u and the state will approach the origin.
- It is helpful (as we will see when we compute upper bounds) to rewrite the lower bound in terms of the singular and eigenvalues of the matrix $A^T F$ as follows:

$$\gamma_{LB} = 2^{-2RT/n} \left\{ \prod_{i=0}^{n-1} \sigma_i(A^T F) \right\}^{\frac{2}{n}} = 2^{-2RT/n} \left\{ \prod_{i=0}^{n-1} |\lambda_i(A^T F)| \right\}^{\frac{2}{n}}.$$

Again, when computing the lower bound, we made no assumptions on whether the encoder and decoder are causal or noncausal. In the following sections, we compute noncausal and causal upper bounds.

III-B. A Noncausal Upper Bound

In this section, we derive an upper bound on worst-case performance assuming that the encoder and decoder are noncausal. The upper bound is derived using a coding scheme that transmits information about the vector x_0 in terms of a basis derived from the singular value decomposition (SVD) of the matrix $A^T F$ similar to the scheme describes in section II-B.

Consider Figure 11 below. The encoder first uses the SVD of $A^T F = U \Sigma V^*$ to write $A^T x_0 = A^T F w = \sum_{i=0}^{n-1} \sigma_i \alpha_i u_i$, where σ_i is the i th singular value of $A^T F$, $\alpha_i = v_i^* w$ where v_i^* is the i th row vector of V^* , and u_i is the i th column vector of U . The α_i 's are then each converted into their binary representations and truncated according to the bit-allocation strategy denoted in $\mathcal{R} = (R_1, R_2, \dots, R_n)$. In particular, a total of R_k bits are allocated to α_k , for $k = 1, \dots, n$, and the only restriction is that $\sum_{k=1}^n R_k = TR$.

The decoder uses the bit-allocation strategy \mathcal{R} to reconstruct α and then uses the SVD of $A^T F$ to compute \hat{x}_0 from $\hat{\alpha}$. Finally, the decoder computes the minimum 2-norm control signal u such that $A^T \hat{x}_0 + M u = 0$, which is the least squares solution to an under-constrained set of linear equations.

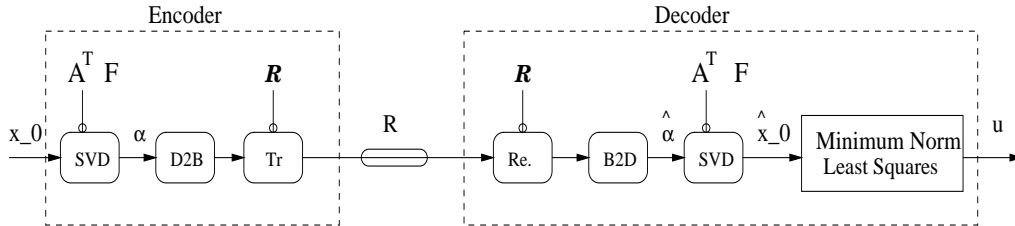


Fig. 11. SVD Coding Scheme

Note that with the above SVD coding scheme,

$$\begin{aligned}
\sup_{x_0 \in \mathcal{C}_{x_0}} \|x_T\|_2^2 &= \sup_{x_0 \in \mathcal{C}_{x_0}} \|A^T(\hat{x}_0 - x_0)\|_2^2 \\
&= \sup_{\{w \mid \|w\|_2 \leq 1\}} \|A^T F(\hat{w} - w)\|_2^2 \\
&= \sup_{\{\alpha \mid \|\alpha\|_2 \leq 1\}} \sum_{i=1}^n \sum_{j=1}^n (\hat{\alpha}_i - \alpha_i)(\hat{\alpha}_j - \alpha_j) \sigma_i \sigma_j (u_i' u_j) \\
&\leq \sup_{\{\alpha \mid \|\alpha\|_2 \leq 1\}} \sum_{i=1}^n |\alpha_i|^2 2^{-2R_i} \sigma_i^2 \\
&\leq \max_i 2^{-2R_i} \sigma_i^2.
\end{aligned}$$

To derive the upper bound using the above SVD coding scheme, we construct $\mathcal{R} = (R_1, R_1, \dots, R_n)$ to solve the following optimization problem:

$$\begin{aligned}
&\min_{\mathcal{R}} \max_i 2^{-2R_i} \sigma_i^2 \\
&\text{s.t. } \sum_{i=1}^n R_i = TR \\
&\quad R_i \geq 0 \quad \forall i.
\end{aligned}$$

We allow the rates to take on non-integer values to solve for an optimal bit-allocation strategy. The resulting non-integer valued rates can be interpreted as average rates over time. The above problem is computable and it is easy to verify (using Lagrange multipliers) that the optimal solution is $R_i^* = (\frac{RT}{n} - \frac{1}{n} \sum_{i=1}^n \log_2(\sigma_i)) + \log_2(\sigma_i) \quad \forall i$. Therefore, the larger the singular value σ_i , the more bits are allocated to α_i . It is straightforward to show that the resulting upper bound achieves the lower bound.

III-C. A Causal Upper Bound

In this section, we derive an upper bound by constructing a modified SVD Coding Scheme, where the encoder has access to the entire vector $x_0 \in \mathcal{C}_{x_0}$ at time $t = 0$, but the decoder is restricted to only process R bits of information per time step. This is a more practical implementation.

The causal coding scheme transmits information about x_0 in terms of a basis derived from the singular value decomposition (SVD) of the matrix F . Consider Figure 12. Here, the encoder first uses the SVD of $F = U\Sigma V^*$ to write $x_0 = \sum_{i=0}^{n-1} \sigma_i \alpha_i u_i$, where σ_i is the i th singular value of F , $\alpha_i = v_i^* w$ where v_i^* is the i th row vector of V^* , and u_i is the i th column vector of U . The α_i 's are then each converted into their binary representations and truncated according to the bit-allocation strategy dictated by a $T \times n$ rate matrix

$$\mathcal{R} = \begin{bmatrix} R_{01} & R_{12} & \dots & R_{0n} \\ R_{11} & R_{12} & & R_{1n} \\ R_{21} & R_{22} & & R_{2n} \\ \vdots & \vdots & & \vdots \\ R_{T-1,1} & R_{T-1,2} & R_{T-1,3} & \dots & R_{T-1,n} \end{bmatrix},$$

such that $\sum_j R_{ij} = R$ for $i = 0, 1, \dots, T-1$. Let $R_i(t) = \sum_{j=0}^t R_{ji}$ for $i = 1, \dots, n$ and $t = 0, 1, \dots, T-1$. Then, a total of $R_i(t)$ bits are allocated to α_i , for $i = 1, \dots, n$ at time t .

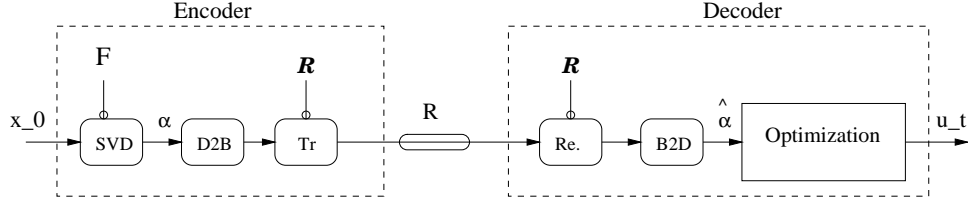


Fig. 12. SVD Coding Scheme for Navigation

At time t , the decoder uses the bit-allocation strategy \mathcal{R} (as described in section II-C) to reconstruct α as $\hat{\alpha}(t)$. The decoder then computes the scalar control input u_t that minimizes $\|x_t\|_2$, by solving the following optimization problem

$$\begin{aligned} \min_{u_t} \sup_{\alpha \in S_t} \|AU\Sigma\alpha + \sum_{i=0}^{t-2} A^{t-1-i}Bu_i + Bu_t\|_2 \\ \text{s.t. } S_t = \{\alpha \in \mathbb{R}^n \mid |\alpha_i - \hat{\alpha}_i(t)| \leq 2^{-R_i(t)} \quad i = 1, 2, \dots, n\}. \end{aligned}$$

The solution to the above min-sup problem can be computed exactly. First, we use the property that the supremum of a convex function over a bounded interval occurs at the boundaries, which gives us that $\alpha^*(t)$, the solution to $\sup_{\alpha \in S_t} \|AU\Sigma\alpha + \sum_{i=0}^{t-2} A^{t-1-i}Bu_i + Bu_t\|_2$, as one of the following vectors:

$$\alpha^*(t) = \begin{bmatrix} \hat{\alpha}_1(t) \pm 2^{-R_1(t)} \\ \hat{\alpha}_2(t) \pm 2^{-R_2(t)} \\ \vdots \\ \hat{\alpha}_n(t) \pm 2^{-R_n(t)} \end{bmatrix}.$$

Then, we minimize $\|AU\Sigma\alpha^*(t) + \sum_{i=0}^{t-2} A^{t-1-i}Bu_i + Bu_t\|_2$ by taking its derivative with respect to u_t and then setting it to 0. We get that the optimal control input at time t is $u_t^* = \frac{(AU\Sigma\alpha^*(t) + \sum_{i=0}^{t-2} A^{t-1-i}Bu_i)'B}{(B'B)}$. Finally, we can show that $\sup_{C_{x_0}} \|x_T\|_2$ occurs at $x_0 = 0$ and correspondingly $\alpha = 0$, which gives

$$\alpha^*(t) = \begin{bmatrix} 2^{-R_1(t)} \\ 2^{-R_2(t)} \\ \vdots \\ 2^{-R_n(t)} \end{bmatrix}.$$

Therefore, the decoder computes $u_t^* = \frac{(AU\Sigma\alpha^*(t) + \sum_{i=0}^{t-2} A^{t-1-i}Bu_i)'B}{(B'B)}$ using the above $\alpha^*(t)$.

The critical difference between the above coding scheme and that introduced in section II-B is the restriction of allocating a total of R bits each time step, which is captured by the rate matrix.

III-D. Comparison of Bounds

We now compare the lower and upper navigation bounds on γ to each other for different LTI causal systems $H = ss(A, B, C, D)$, and for different time horizons T . We consider diagonal 4×4 ($n = 4$) state-transition matrices $A = diag(a_0, a_1, a_2, a_3)$, an F that is generated by LTI system $ss(A_f, B_f, C_f, D_f)$, and we fix the rate $R = 5$. Under such conditions,

$$\gamma_{LB} = 2^{-2RT/n} (D_f)^2 \left\{ \prod_{i=0}^{n-1} |a_i|^T \right\}^{\frac{2}{n}}.$$

Figures 13 and 14 illustrate the bounds for the following scenarios.

- 1) $F = ss(0.99, 0.99, 1, 1)$, $A = diag([0.2 \ 0.8 \ 0.9 \ 0.8])$, and $B = [1 \ 1 \ 1 \ 1]'$.
- 2) $F = ss(0.01, 0.01, 1, 1)$, and A and B are the same matrices as those generated in 1.
- 3) $F = ss(0.99, 0.99, 1, 1)$, $A = diag([1.2 \ 1.8 \ 1.9 \ 1.8])$, and $B = [1 \ 1 \ 1 \ 1]'$.
- 4) $F = ss(0.01, 0.01, 1, 1)$, and A and B are the same matrices as those generated in 3.

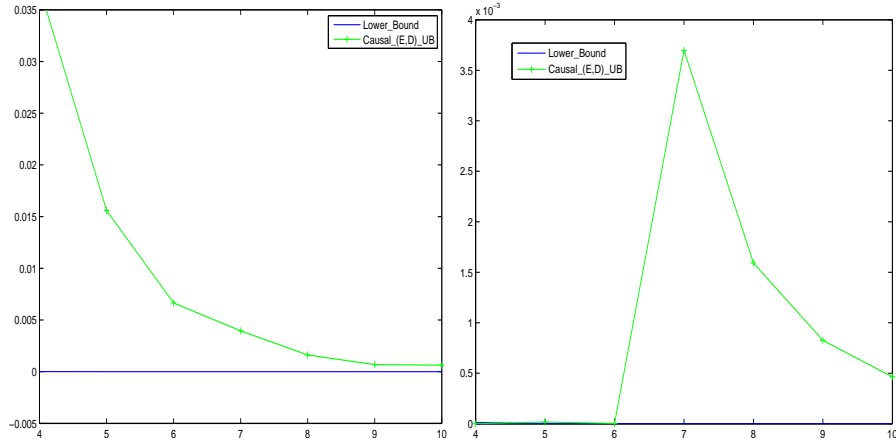


Fig. 13. Left: Bounds for $F = ss(0.99, 0.99, 1, 1)$ and A stable, Right: Bounds for $F = ss(0.01, 0.01, 1, 1)$ and A stable

We make a few observations.

- *Stability of A:* All bounds decay when A is stable as T grows. When A is unstable, then the causal upper bound only decays when the pole of the system that generates F is closer to the unit disk.
- *Pole of F:* When the pole of the system that generates F is closer to the origin, the singular values of F are all comparable and therefore using the SVD basis to represent x_0 in the causal coding scheme is less helpful. Therefore, we expect causal coding performance to deteriorate, which it does. Put another way, when F has a pole close to the unit disk, then the ellipsoid set \mathcal{C}_{x_0} has more structure, that is knowing some components of x_0 give a lot of information about the remaining component of x_0 . When the pole of F is closer to 0, then \mathcal{C}_{x_0} looks more and more like an n -dimensional sphere.

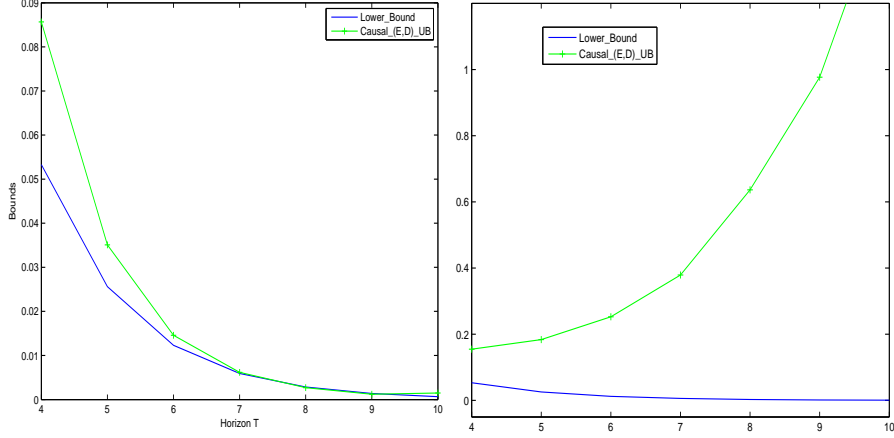


Fig. 14. Left: Bounds for $F = ss(0.99, 0.99, 1, 1)$ and A unstable, Right: Bounds for $F = ss(0.01, 0.01, 1, 1)$ and A unstable,

III-E. A Lower Bound on Time Horizon

In this section, we seek to minimize the time it takes for the state vector to reach a ball of size γ . Therefore, we fix γ and then look for the smallest horizon T to meet performance.

Theorem III.2. Consider a reachable SISO LTI causal DT system, $H = ss(A, B, I, 0)$ with A being full rank and with initial condition $x_0 \in \{x \in \mathbb{R}^n | x = Fw, w \in \mathbb{R}^n, w^T w \leq 1\}$, for a given F . If the control input is constrained to take on at most 2^{RT} values after $T \geq n$ time steps, then $\|x_T\|_2^2 \leq \gamma$ (for any given $\gamma > 2^{-2R}|\det(A)|^2$) only if

$$T \geq \max\left(\left\lceil \frac{2(\log_2(|\det(F)|) - \log_2(|\det(A)|))}{\log_2(\gamma) + 2R - 2\log_2(|\det(A)|)} \right\rceil, n\right).$$

Proof.

As shown in the proof for Theorem III.1, an equivalent expression for navigation performance $\|Mu + A^T Fw\|_2^2 \leq \gamma$ is $\|U_m \alpha + A^T Fw\|_2^2 \leq \gamma$. We apply the same counting argument as that given in the proof for Theorem III.1, but isolate T on one side of the inequality instead of γ . We then get that the lower bound on the number of time steps as $\left\lceil \frac{2(\log_2(|\det(F)|) - \log_2(|\det(A)|))}{\log_2(\gamma) + 2R - 2\log_2(|\det(A)|)} \right\rceil$ as long as $\log_2(\gamma) + 2R - 2\log_2(|\det(A)|) > 0$, which is equivalent to $\gamma > 2^{-2R}|\det(A)|^2$. Finally, since we assumed that $T \geq n$ (M full row rank), this lower bound is only valid if it is greater than n . If $\left\lceil \frac{2(\log_2(|\det(F)|) - \log_2(|\det(A)|))}{\log_2(\gamma) + 2R - 2\log_2(|\det(A)|)} \right\rceil < n$, then it is possible to reach a ball of size γ in n time steps. ■

Note that the lower bound depends on R , F , γ , and the system dynamics A .

IV. CONCLUSIONS & FUTURE WORK

In this paper, we present a deterministic worst-case approach for reconstructing signals that are transmitted through a finite-rate noiseless channel and then applied to linear dynamical systems. We define two finite-horizon system performance

criteria, discuss performance limitations when coding is restricted to be causal, and quantify achievable performance levels for each.

In this paper, we ignore feedback and only consider how the finite-rate channel impacts control in a feedforward setting. This makes sense when one is controlling “slowly”. That is, a command is sent to a remote system when it is at rest and then it “moves” accordingly. Feedback as to where the system has moved to is given, but at a slower time-scale. In future work, we will allow for feedback and study finite-horizon objectives in this case.

V. ACKNOWLEDGEMENTS

The authors would like to thank Alexander Megretski for insightful discussions and reviewers of previous related work, which helped us complete our results.

VI. REFERENCES

- [1] Allen, Gersho, Gray, Robert M., “Vector Quantization and Signal Compression,” *Boston : Kluwer Academic Publishers*, c1992.
- [2] Baillieul, J., “Feedback Coding for Information-based Control: Operating Near the Data-Rate Limit,” *Proceedings of the 41st IEEE Conference on Decision and Control*, December 2002.
- [3] Wong, Wing Shing, Brockett, Roger, “Systems with Finite Communication Bandwidth Constraints-I: State Estimation Problems,” *IEEE Transactions on Automatic Control*: vol. 42, no. 9, September 1997.
- [4] Brockett, Roger W., “Quantized Feedback Stabilization of Linear Systems,” *IEEE Transactions on Automatic Control*: vol. 45, pp. 1279-1289, 2000.
- [5] Wong, Wing Shing, Brockett, Roger, “Systems with Finite Communication Bandwidth Constraints-II: Stabilization with Limited Information Feedback,” *IEEE Transactions on Automatic Control*: vol. 44, no. 5, May 1999.
- [6] Delevenne, Jean-Charles, Blondel, Vincent, “Complexity of Control of Finite Automata,” *IEEE Transactions on Automatic Control*: preprint.
- [7] Delchamps, David, “Stabilizing a Linear System with Quantized State Feedback,” *IEEE Transactions on Automatic Control*: vol. 35, no. 8, August 1990.
- [8] Delevenne, Jean-Charles, “An optimal quantized feedback strategy for scalar linear systems,” *IEEE Transactions on Automatic Control*: accepted in 2005.
- [9] Doyle, John, Francis, Bruce, Tannenbaum, Allen “Feedback Control Theory,” *Macmillan Publishing Company; New York*, c1992.
- [10] Elia, Nicola, Mitter, Sanjoy “Stabilization of Linear Systems With Limited Information,” *IEEE Transactions on Automatic Control*: vol. 46, no. 9, September 2001.
- [11] Fagnani, Fabio, Zampieri, Sandro, “Stability Analysis and Synthesis for Linear Systems With Quantized Feedback,” *IEEE Transactions on Automatic Control*: vol. 48, no. 9, 2003.
- [12] Fagnani, Fabio, Zampieri, Sandro, “Quantized Stabilization of Linear Systems: Complexity Versus Performance,” *IEEE Transactions on Automatic Control*: vol. 49, no. 9, 2004.

- [13] Gallager, Robert, "Information Theory and Reliable Communication," *John Wiley and Sons, Inc.; New York*, c1968.
- [14] Ishii, Hideaki, Francis, Bruce, "Limited data rate in control systems with networks," *Berlin; New York: Springer*, c2002.
- [15] Liberzon, Daniel, "A note on stabilization of linear systems using coding and limited communication," *Proceedings of the 41st IEEE Conference on Decision and Control*, December 2002.
- [16] Martins, N.C., Dahleh M.A., and Elia N. "Stabilization of Uncertain Systems in the Presence of a Stochastic Digital Link", *43rd IEEE Conference on Decision and Control*, (Extended version submitted to the *IEEE Transactions in Automatic Control*).
- [17] Martins, N.C., Dahleh M.A., "Bounded Input Bounded Output Stabilization is Impossible, Under Finite Capacity Feedback," *pre-print*.
- [18] Nair, G.N., Evans, R.J., "Stabilization with Data-rate-limited Feedback: tightest attainable bounds," *Systems and Control Letters*, volume 41, 2000, pp. 49-56.
- [19] Nair, G.N., Evans, R.J. "Stabilizability of stochastic linear systems with finite feedback data rates", *SIAM Journal on Control and Optimization, Society for Industrial and Applied Mathematics, USA*, vol. 43, no. 2, pp. 413 - 436, July 2, 2004.
- [20] Sarma, Sridevi, Dahleh, Munther A., Salapaka, Srinivasa, "On Time-Varying Bit-Allocation Maintaining Stability: A Convex Parameterization," *Proceedings of the 43rd IEEE Conference on Decision and Control*, December 2004.
- [21] Sarma, Sridevi, Dahleh, Munther A., "Real-Time Finite-Rate Control," *Preprint to be submitted to IEEE TAC*, January 2006.
- [22] Tatikonda, Sekhar, "Control Under Communication Constraints," *IEEE Transactions on Automatic Control* : Volume 49, Issue 7, July 2004 Page(s): 1056 - 1068.
- [23] Yuksel, Serder, Basar, Tamer, "State Estimation and Control for Linear Systems over Communication Networks," *Proceedings of 2003 IEEE Conference on Control Applications*, Volume 1, June 2003.