

Signal Reconstruction in the Presence of Finite-Rate Measurements: Finite-Horizon Control Applications

Sridevi V. Sarma, Munther A. Dahleh*

*Massachusetts Institute of Technology, Laboratory of Information and Decision Systems, Cambridge,
Massachusetts, USA*

SUMMARY

In this paper, we study finite-length signal reconstruction over a finite-rate noiseless channel. We allow the class of signals to belong to a bounded ellipsoid and derive a universal lower bound on a *worst-case* reconstruction error. We then compute upper bounds on the error that arise from different coding schemes and under different causality assumptions. When the encoder and decoder are noncausal, we derive an upper bound that either achieves the universal lower bound or is comparable to it. When the decoder and encoder are both causal operators, we show that within a very broad class of causal coding schemes, memoryless coding prevails as optimal, imposing a hard limitation on reconstruction. Finally, we map our general reconstruction problem into two important control problems 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. The first problem is to minimize a finite-horizon weighted tracking error between the remote system output and a reference command. The second

*Correspondence to: sree@mit.edu

Contract/grant sponsor: AFOSR; contract/grant number: 6892167

problem is to navigate the state of the remote system from a nonzero initial condition to as close to the origin as possible in finite-time. Our analysis enables us to quantify the tradeoff between time horizon and performance accuracy which is not well-studied in the area of control with limited information as most works address infinite-horizon control objectives (eg. stability, disturbance rejection).

KEY WORDS: Signal Reconstruction, Finite-Rate, Finite-Horizon, Control

1. INTRODUCTION

Signal reconstruction over noisy channels has been well studied under stochastic settings, where performance criteria are typically characterized by asymptotic properties of the probability of error given *stochastic* descriptions of the input signal and channel. The main objective of signal reconstruction is to design computationally efficient coding schemes to optimize performance [18, 26]. Recent work by Voulgaris investigates reconstruction of *infinite-length discrete-valued* signals that are filtered via noisy channels using a *deterministic* framework [29]. In contrast, we study *finite-length real-valued* signal reconstruction filtered via finite-rate but otherwise noiseless channels using a deterministic framework. In particular, we are interested in minimizing reconstruction error in *finite-time*, whereas in most communication settings questions about *asymptotic* reconstruction are typically addressed. We study finite-time performance because we are ultimately interested in understanding how the reconstructed signal can be used to drive or control a system.

Control over noisy channels is a research area of growing interest. 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.

Much work in the area of control with limited information 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, 11, 17, 19, 21, 24, 25, 27, 28, 30]. 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/remain bounded as time goes to infinity. More recently, disturbance rejection limitations were derived for the same setting, assuming stochastic exogenous signals entering the system [22, 23]. Although these studies greatly contribute to our understanding of the interplay between communication and control, few studies have addressed 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, 12, 13]. A navigation problem similar in spirit to that which is presented here is described in [13]. In [13] Fagnani et al. consider the closed-loop system shown in Figure 1.

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

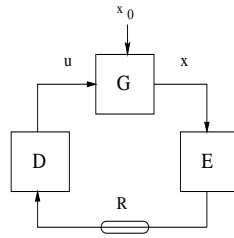


Figure 1. Equivalent Closed-Loop System

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 [13] for details. This result shows that demanding smaller values of the expected minimum time to reach set V requires more complex coding.

In this paper we introduce and analyze a general signal reconstruction framework which enables us to compute universal 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. 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.

2. GENERAL RECONSTRUCTION PROBLEM

In this section, we define a framework to study finite-length signal reconstruction under finite-rate measurements. We consider the cascade of SISO discrete-time systems shown in Figure 2.

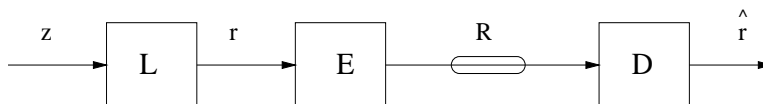


Figure 2. General Reconstruction Set Up

Specifically,

- $z \in \mathbb{R}^T$ s.t. $\|z\|_2 \leq 1$,
- $r \in \mathcal{C}_r \triangleq \{r \in \mathbb{R}^T, z \in \mathbb{R}^T | r = Lz, \|z\|_2 \leq 1\}$,
- $L : \mathbb{R}^T \rightarrow \mathbb{R}^T$ 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 $\{0, 1\}^{RT} \rightarrow \{0, 1\}^{RT}$,
- and
- $D : \{0, 1\}^{RT} \rightarrow \mathbb{R}^T$ is an arbitrary operator (decoder) that maps a sequence of 2^{RT} binary symbols to a real vector.

Note that L defines a class of finite-length 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 assume L and the

channel rate R are given, and we want to find an encoder E and decoder D to minimize a reconstruction 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: $\|W(r - \hat{r})\|_2^2$, 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 ($E = I$ and $D = I$) such that $\|W(r - \hat{r})\|_2^2 = 0 \quad \forall r \in \mathcal{C}_r$. However, with a finite-rate constraint the decoder output, \hat{r} , can only take 2^{RT} values over a horizon of T time steps. Therefore, it is not clear how well signals can be estimated under such constraints.

To understand reconstruction limitations under finite-rate measurements, we compute γ_{LB} and γ_{UB} , such that

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

Knowledge of γ_{LB} tells us that regardless of the encoder and decoder that we select, we can do no better than this universal lower bound. 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} . To compute γ_{UB} , we construct an encoder and decoder and compute the corresponding worst-case performance. In the following sections, we compute a universal lower bound and three upper bounds corresponding to three types of coding schemes.

3. UNIVERSAL LOWER BOUND

In this section we derive a universal lower bound on worst-case reconstruction using a standard counting or sphere-packing type argument.

Theorem 3.1. *Given the signal reconstruction 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 signals, $\mathcal{C}_r \triangleq \{r \in \mathbb{R}^T \mid r = Lz, z'z \leq 1\} = \{r \in \mathbb{R}^T \mid (L^{-1}r)'(L^{-1}r) \leq 1\}$ 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 decoder output, \hat{r} , to take on no more than 2^{RT}

Consider a selection of reconstructed estimates $\hat{r}^1, \hat{r}^2, \dots, \hat{r}^{2^{RT}}$. We must then map each $r \in \mathcal{C}_r$ to exactly one \hat{r}^i , $i = 1, 2, \dots, 2^{RT}$. Such a mapping induces a partition on \mathcal{C}_r . In particular, define $P_i = \{r \in \mathcal{C}_r \mid r \rightarrow \hat{r}^i\}$ for $i = 1, 2, \dots, 2^{RT}$. Now, suppose that the selection $\hat{r}^1, \hat{r}^2, \dots, \hat{r}^{2^{RT}}$ were chosen such that $\|W(r - \hat{r}^i)\|_2^2 \leq \gamma$ for all $r \in P_i$, and for all i . Then necessarily $P_i \subseteq S_{\hat{r}^i}^\gamma \triangleq \{r \in \mathbb{R}^T \mid (r - \hat{r}^i)'W'W(r - \hat{r}^i) \leq \gamma\}$. Note that $S_{\hat{r}^i}^\gamma$ is a bounded ellipsoid in \mathbb{R}^T centered at point \hat{r}^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.

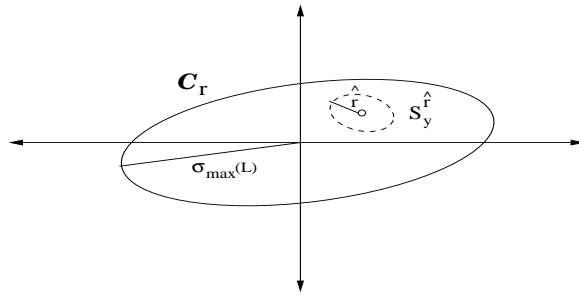


Figure 3. Bounded Ellipsoids \mathcal{C}_r and $S_{\hat{r}}^\gamma$

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

$$2^{RT} \geq \frac{\text{volume}(C_r)}{\text{volume}(S_{\hat{r}}^\gamma)} = \frac{|\det(L)| |\det(W)|}{(\sqrt{\gamma})^T}.$$

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

In control settings, we often consider classes of signals generated from LTI systems, *i.e.*, L is LTI. We compute the lower bound for this case in the following corollary.

Corollary 3.1. *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 $L = \text{ss}(A_l, B_l, C_l, D_l)$, then*

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

Proof. *If L is a SISO causal LTI with state-space description $L = \text{ss}(A_l, B_l, C_l, D_l)$, 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 signals), W (performance weights), T (signal length), 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:

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

When computing the lower bound, we made no assumptions on whether the encoder and decoder are causal or noncausal operators. If both the encoder and decoder are noncausal, then at time $t = 0$ the decoder “knows” the future. That is, at $t = 0$ it can compute \hat{r}_k for $k = 0, 1, \dots, T - 1$ which are represented by TR bits over a horizon of T steps and our

reconstruction problem reduces to a vector quantization (VQ) problem with a deterministic error metric [1, 15]. If we consider our signal r to be a correlated Gaussian random vector with covariance matrix M_r , then the lower bound on the minimum mean-squared error ($E(r - \hat{r})^T(r - \hat{r})$) is $2^{-2R} \det(M_r)^{\frac{1}{T}}$ [16], which is identical to our lower bound if we replace WL with $M_r^{\frac{1}{2}}$. This makes sense as the mean-squared error lower bound is for every possible realization \hat{r} and hence considers the worst case.

We still, however, find it useful to derive γ_{LB} in our deterministic setting and compare it to upper bounds computed under various coding schemes such as when coding must be done in a causal manner.

Recall that if E and D are causal, then at time k the encoder can only operate on r_0, r_1, \dots, r_k and decoder can only reconstruct $\hat{r}_0, \dots, \hat{r}_k$ where \hat{r}_k is represented by at most $(k + 1)R$ bits. In the following sections, we compute three upper bounds. The first bound is computed by constructing a noncausal encoder and decoder (γ_{NN}), the second is computed by constructing a noncausal encoder and a causal decoder (γ_{NC}), and the third is computed by constructing a causal encoder and decoder (γ_{CC}).

4. NONCAUSAL ENCODING AND DECODING

In this section, we derive an upper bound, γ_{NN} , on worst-case performance assuming that the encoder and decoder are both noncausal. The upper bound is computed 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 . This scheme is identical to that proposed for a Gaussian source signal in [16] and we repeat it here to compare to the coding scheme described later in section 6.

Consider Figure 4. Let the SVD decomposition of WL be defined as $WL = U\Sigma V^*$, where U is a $T \times T$ unitary matrix, Σ is a $T \times T$ diagonal matrix containing the singular values of WL , and V is a $T \times T$ unitary matrix. The encoder consists of rotator and quantizer operators, p and q , respectively which are defined as follows:

- $p : \mathbb{R}^T \rightarrow \mathbb{R}^T$ such that $p(r) = V^* L r$,
- $q : \mathbb{R}^T \rightarrow \{0, 1\}^{RT}$ is parameterized by the bit-allocation strategy denoted in $\mathcal{R} = (R_0, R_1, \dots, R_{T-1})$ as shown in Figure 4 where $\hat{\alpha}_i = q_i(\alpha_i)$ and $q_i(\alpha_i)$ is a uniform quantizer from $[-1, 1]$ with 2^{R_i} levels for $i = 0, 1, \dots, T-1$ ($\alpha = (\alpha_0 \ \alpha_1 \ \dots \ \alpha_{T-1})$).

The decoder first employs the bit-allocation strategy \mathcal{R} to reconstruct $\hat{\alpha}$ from the binary string that it receives. It then rotates the vector to compute \hat{r} . Specifically, $\hat{r} = L^{-1} V \hat{\alpha}$. We call this $E - D$ construction the ‘‘SVD Coding Scheme.’’

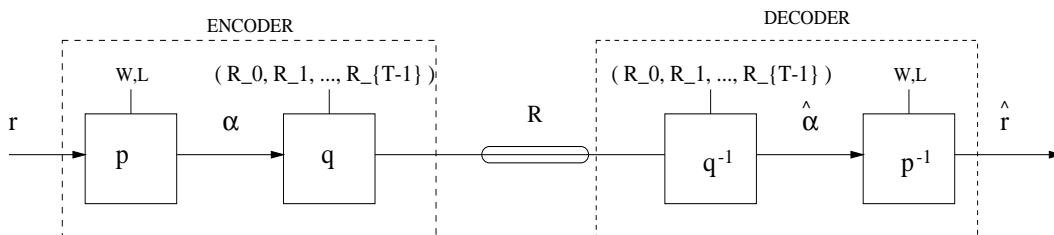


Figure 4. SVD Coding Scheme

Note that for the above SVD coding scheme,

$$\begin{aligned}
 \sup_{r \in \mathcal{C}_r} \|W(r - \hat{r})\|_2^2 &= \sup_{\{z \mid \|z\|_2 \leq 1\}} \|WL(\hat{z} - z)\|_2^2 \\
 &= \sup_{\{z \mid \|z\|_2 \leq 1\}} \|U\Sigma V^*(\hat{z} - z)\|_2^2 \\
 &= \sup_{\{\alpha \mid \|\alpha\|_2 \leq 1\}} \|\Sigma(\hat{\alpha} - \alpha)\|_2^2 \tag{1}
 \end{aligned}$$

$$= \sup_{\{\alpha \mid \|\alpha\|_2 \leq 1\}} \sum_{i=0}^{T-1} (\hat{\alpha}_i - \alpha_i)^2 \sigma_i^2 \tag{2}$$

$$\leq \sup_{\{\alpha \mid \|\alpha\|_2 \leq 1\}} \sum_{i=0}^{T-1} 2^{-2R_i} \sigma_i^2 |\alpha_i|^2 \tag{3}$$

$$\leq \max_i 2^{-2R_i} \sigma_i^2 \sup_{\{\alpha \mid \|\alpha\|_2 \leq 1\}} \sum_{i=0}^{T-1} |\alpha_i|^2 \tag{4}$$

$$= \max_i 2^{-2R_i} \sigma_i^2. \tag{5}$$

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

$$\begin{aligned}
 \min_{\mathcal{R}} \quad & \max_i 2^{-2R_i} \sigma_i^2 \tag{6} \\
 \text{s.t.} \quad & \sum_{i=0}^{T-1} R_i \leq TR \\
 & R_i \geq 0 \quad \forall i.
 \end{aligned}$$

Problem (6) is equivalent to the following optimization problem:

$$\begin{aligned}
\min_{\mathcal{R}, \gamma} \quad & \gamma & (7) \\
& 2^{-2R_i} \sigma_i^2 \leq \gamma \\
s.t. \quad & \sum_{i=0}^{T-1} R_i \leq TR \\
& R_i \geq 0 \quad \forall i. \\
& \gamma \geq 0.
\end{aligned}$$

† Now note that a lower bound to the optimal cost of (7) is the optimal solution to

$$\begin{aligned}
\min_{\mathcal{R}, \gamma} \quad & \gamma + \sum_{i=0}^{T-1} \lambda_i (-2R_i + 2\log(\sigma_i) - \log(\gamma)) + \mu (\sum_{i=0}^{T-1} R_i - TR) & (8) \\
& R_i \geq 0 \quad \forall i. \\
& \gamma \geq 0,
\end{aligned}$$

where $\mu \geq 0$ and $\lambda_i \geq 0$ for all i . We rearrange terms to get

$$\min_{\gamma \geq 0} [\gamma - \sum_{i=0}^{T-1} \lambda_i \log(\gamma)] + \min_{R_i \geq 0} [\sum_{i=0}^{T-1} R_i (\mu - 2\lambda_i)] + [2 \sum_{i=0}^{T-1} \lambda_i \log(\sigma_i) - \mu RT]. \quad (9)$$

† 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 minimization over nonnegative R_i (second term in (9)) is as follows

$$R_i^* = \begin{cases} 0 & \mu - 2\lambda_i \geq 0 \quad \forall i \\ -\infty & o.w. \end{cases}$$

which gives us

$$\begin{aligned} \min_{\gamma \geq 0} \quad & \gamma - \sum_{i=0}^{T-1} \lambda_i \log(\gamma) + 2 \sum_{i=0}^{T-1} \lambda_i \log(\sigma_i) - \mu RT. \\ \text{s.t.} \quad & 0 \leq \lambda_i \leq \frac{\mu}{2} \quad \forall i \end{aligned} \quad (10)$$

since we know a finite solution to (7) exists. Now, if we minimize over γ by taking the derivative of the cost function with respect to γ and setting it equal to 0, we get that

$$\gamma^* = \sum_i \frac{\lambda_i^*}{\ln(2)}, \quad (11)$$

and the dual to (7) is obtained by maximizing the lower bound as follows

$$\begin{aligned} \max_{\lambda_0, \lambda_1, \dots, \lambda_{T-1}, \mu} \quad & \sum_i \lambda_i \left(\frac{1}{\ln(2)} - \log\left(\sum_i \frac{\lambda_i}{\ln(2)}\right) \right) + \sum_i 2\lambda_i \log(\sigma_i) - \mu RT \\ \text{s.t.} \quad & 0 \leq \lambda_i \leq \frac{\mu}{2} \quad \forall i \\ & \mu \geq 0. \end{aligned} \quad (12)$$

It is fairly straightforward to show that the cost to (12) is a convex function of $\lambda_0, \lambda_1, \dots, \lambda_{T-1}$ over a bounded set, and therefore the optimal solution occurs at a boundary. That is,

$$\lambda_i^* = \begin{cases} \frac{\mu}{2} & i \in I \\ 0 & i \in I^c, \end{cases} \quad (13)$$

where $I \subset \{0, 1, \dots, T-1\}$. Plugging (13) into (12), the dual becomes

$$\begin{aligned} \max_{\mu} \quad & \frac{\mu|I|}{2} \left(\frac{1}{\ln(2)} - \log\left(\frac{\mu|I|}{2\ln(2)}\right) \right) + \mu \sum_{i \in I} \log(\sigma_i) - \mu RT \\ & \mu \geq 0. \end{aligned} \quad (14)$$

One can compute the solution to (14) as

$$\mu^* = \frac{2\ln(2)}{|I|} 2^{-\frac{2RT}{|I|}} \prod_{i \in I} \sigma_i^{\frac{2}{|I|}}.$$

Finally, we plug μ^* into (13) and then (9) to get

$$\gamma^* = \gamma_{NN} = \max \left\{ 2^{-\frac{2RT}{|I|}} \prod_{i \in I} \sigma_i^{\frac{2}{|I|}}, \max_{i \in I^c} \sigma_i^2 \right\}. \quad (15)$$

One can show that the dual cost equals the primal cost when

$$R_i^* = \begin{cases} \log(\sigma_i) + \frac{RT}{|I|} - \frac{1}{|I|} \sum_{i \in I} \log(\sigma_i) & \text{for } i \in I \\ 0 & \text{o.w.} \end{cases}$$

Since $\lambda_i^* = 0$ for all $i \in I^c$, necessarily $R_i^* = 0$ for $i \in I^c$ since the 2nd term in (4) must equal 0. As expected, the optimal rate allocation places more bits to components of α whose corresponding singular values are larger. Surprisingly, if $|I| = T$, then $R_i^* =$

$R + \log(\sigma_i) - \frac{1}{T} \sum_{j=0}^{T-1} \log(\sigma_j)$ for $i = 0, 1, \dots, T - 1$ and the resulting upper bound is $\gamma_{NN} = 2^{-2R} \{\prod_{i=0}^{T-1} \sigma_j\}^{\frac{T}{2}} = \gamma_{LB}$.

Finally, we comment on the construction of I that minimizes γ^* . Note that for the rates R_i to all be nonnegative, we require that

$$\begin{aligned}
 \log(\sigma_i) &\geq \frac{1}{|I|} \sum_{i \in I} \log(\sigma_i) - \frac{RT}{|I|} \quad \forall i \in I \\
 \Rightarrow \quad \sigma_i^2 &\geq 2^{\frac{-2RT}{|I|}} \prod_{i \in I} \sigma_i^{\frac{2}{|I|}} \quad \forall i \in I.
 \end{aligned} \tag{16}$$

Comparing the above inequality to the expression for γ^* in (10), we see that in order to minimize γ^* , we want to place all the indices corresponding to the larger singular values in the index set I until the positivity constraints on the rates are violated.

5. CAUSAL ENCODING AND DECODING

In this section, we derive an upper bound, γ_{CC} , assuming that the encoder and decoder are both causal. Recall that since the encoder is causal, at any time t , it can only operate on r_0, r_1, \dots, r_t (assume it has enough memory to hold past values of r). Similarly, at time t , the decoder can only recover r_0, r_1, \dots, r_t .

We consider the causal coding scheme illustrated in Figure 5. In this scheme, the encoder is a quantizer parameterized by a **rate matrix**, \mathcal{R} , 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 lower-triangular 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 $i = 0, 1, \dots, T-1$.

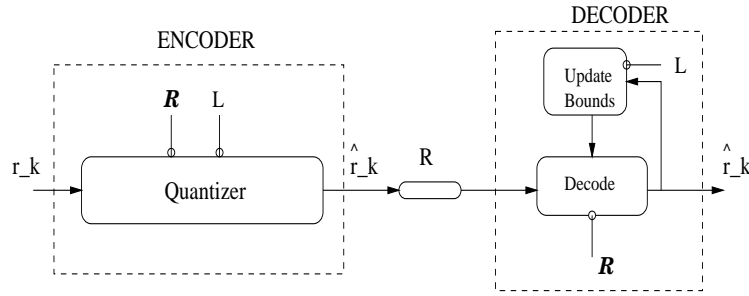


Figure 5. Causal Encoding and Decoding Scheme

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_j from its past inputs and from constraints imposed on $r \in \mathcal{C}_r$.

To clarify, we present the following example for horizon length $T = 2$,

$$Q \triangleq (L^{-1})'(L^{-1}) = \begin{bmatrix} q_0 & 0 \\ 0 & q_1 \end{bmatrix}, \quad W'W = \begin{bmatrix} w_{00} & w_{01} \\ w_{01} & w_{11} \end{bmatrix}.$$

Note that for this example the signal $r = (r_0 \ r_1)$ obeys the following inequality $q_0 r_0^2 + q_1 r_1^2 \leq 1$. At time $t = 0$, the encoder receives r_0 which has magnitude less than or equal to $M_0 = \frac{1}{\sqrt{q_0}}$. 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.

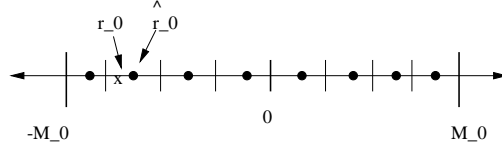


Figure 6. Quantization region at time $t = 0$

The union of the representatives for each region comprises the range of E at time 0. Specifically,

$$E(r_0; t = 0) = \hat{r}_0(0) = nM_02^{-R_{00}}$$

for $(n - 1)M_02^{-R_{00}} \leq r_0 < (n + 1)M_02^{-R_{00}}$, $n = \pm 1, \pm 3, \dots, \pm 2^{R_{00}} - 1$. Thus, when the encoder receives r_0 it outputs the centroid value of the interval in which r_0 falls which is represented by R_{00} bits. The decoder updates its uncertainty interval for r_0 as $r_0 \in [\hat{r}_0(0) - M_02^{-R_{00}}, \hat{r}_0(0) + M_02^{-R_{00}})$.

At time $t=1$, the encoder further quantizes r_0 by dividing the interval $[\hat{r}_0(0) - M_02^{-R_{00}}, \hat{r}_0(0) + M_02^{-R_{00}})$ 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 . The decoder first updates its uncertainty interval for r_0 as $r_0 \in [\hat{r}_0(1) - M_02^{-(R_{00}+R_{01})}, \hat{r}_0(1) + M_02^{-(R_{00}+R_{01})})$, and its uncertainty interval for r_1 as $r_1 \in [\hat{r}_1(1) - M_12^{-R_{11}}, \hat{r}_1(1) + M_12^{-R_{11}})$, where

$$\begin{aligned}
M_1 &= \max_{-1 \leq \rho \leq 1} \frac{\sqrt{1 - q_0(\hat{r}_0(1) + \rho 2^{-(R_{00} + R_{01})} M_0)^2}}{\sqrt{q_1}} \\
&= \begin{cases} \frac{\sqrt{1 - q_0(\hat{r}_0(0) - 2^{-(R_{00} + R_{01})} M_0)^2}}{\sqrt{q_1}} & \hat{r}_0(0) \geq 0 \\ \frac{\sqrt{1 - q_0(\hat{r}_0(0) + 2^{-(R_{00} + R_{01})} M_0)^2}}{\sqrt{q_1}} & \hat{r}_0(0) < 0 \end{cases}
\end{aligned}$$

Equivalently,

$$M_1 = \frac{\sqrt{1 - q_0(\hat{r}_0(0) - \text{sign}(\hat{r}_0(0)) 2^{-(R_{00} + R_{01})} M_0)^2}}{\sqrt{q_1}}.$$

Consider the illustration where $T = 2$ and $R = 3$ shown in Figure 7. At time $t = 0$, the uncertainty set as computed by the decoder for r shrinks from \mathcal{C}_r to one of $2^{R_{00}}$ “strips” of the 2D-ellipsoid. For example, for $r_0 \in [4M_0 2^{-R_{00}}, 6M_0 2^{-R_{00}})$, the uncertainty computed by the decoder is the shaded strip shown in Figure 7, and the corresponding $\hat{r}_0(0) = 5M_0 2^{-R_{00}}$.

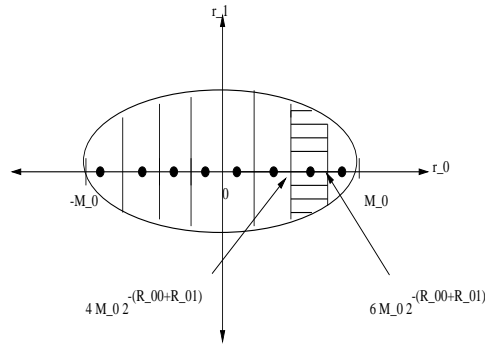


Figure 7. Intervals for r_1 when $r_0 \in [4M_0 2^{-R_{00}}, 6M_0 2^{-R_{00}})$.

At time $t = 1$, the encoder further quantizes r_0 by dividing the interval $[4M_0 2^{-R_{00}}, 6M_0 2^{-R_{00}})$ into $2^{R_{01}}$ equal length intervals as shown in Figure 8. The bound

$$M_1 = \sqrt{\frac{1 - q_0(\hat{r}_0(1) - M_0 2^{-(R_{00} + R_{01})})^2}{q_1}}.$$

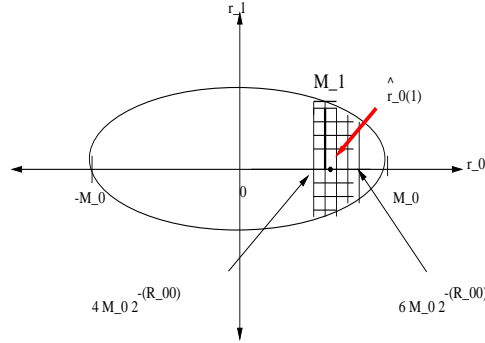


Figure 8. Intervals for r_1 when $\hat{r}_0(1)$ is received by the decoder.

In general, knowledge of r_j impacts the lower and upper bounds on r_k for $k \geq j$. Therefore, it appears that allocating bits to past signal components may be advantageous. It turns out however, that when \mathcal{C}_r is any bounded ellipsoid in \mathbb{R}^T , 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 . Proof of the following theorem is given in the appendix.

Theorem 5.1. *Consider the reconstruction problem that implements the causal coding scheme above parameterized by a rate matrix \mathcal{R} . Assume that L (and hence Q) is diagonal[‡] where*

$$Q \triangleq (L^{-1})'(L^{-1}) = \begin{bmatrix} q_0 & & & \\ & q_1 & & \\ & & \ddots & \\ & & & q_{T-1} \end{bmatrix}, \quad W'W = \begin{bmatrix} w_{00} & w_{01} & \dots & w_{0,T-1} \\ w_{01} & w_{11} & \dots & w_{1,T-1} \\ \vdots & & & \vdots \\ w_{T-1,1} & w_{T-1,1} & \dots & w_{T-1,T-1} \end{bmatrix}.$$

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

[‡]We prove the theorem assuming L is diagonal for a simpler read, however, the general case holds and can be proven using the argument outlined in the appendix.

\mathcal{C}_r) and $\mathcal{R}^* = RI$. Furthermore, the resulting upper bound is

$$\gamma_{CC} = \sum_{i=0}^{T-1} w_{ii} \frac{2^{-2R}}{q_i} + 2 \sum_{i=0}^{T-1} \sum_{j=i+1}^{T-1} w_{ij} \frac{2^{-2R}}{\sqrt{q_i q_j}}.$$

At first, it may seem surprising that allocating only to the present component of the signal (memoryless quantization) is optimal. However, recall that our metric is a worst-case deterministic measure and that the worst case occurs at $r = 0$. If one inputs reference signal $r = 0$ then no information can be extracted from each signal component. However, past allocation may be useful for signal sets that are non-symmetric or finite, or for different performance metrics and will be explored in future work.

6. NONCAUSAL ENCODING AND CAUSAL DECODING

In this section, we derive an upper bound, γ_{NC} , by constructing a modified SVD Coding Scheme in which the encoder is noncausal, *i.e.*, has access to the entire signal $r \in \mathcal{C}_r$ at time $t = 0$, but the decoder is causal. The scheme we propose, sketched in Figure 9, is similar to that described in section 4 with the restriction that the decoder can only process R bits of information at each time step. As a result, the rotator operator p remains unchanged from that defined in section 4 ($p : \mathbb{R}^T \rightarrow \mathbb{R}^T$ such that $p(r) = V^* L r$, where $W L = U \Sigma V^*$), but the quantizer operator q changes.

At each time step t , the quantizer has R bits that it allocates to the entire vector α . The bit-allocation is determined by the following rate matrix

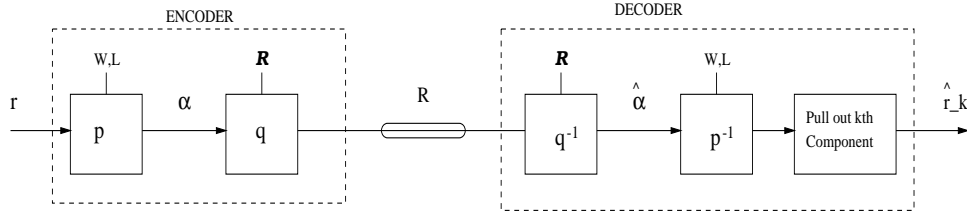


Figure 9. SVD Scheme for Noncausal Encoding and Causal Decoding

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

such that $\sum_j R_{ij} = R$ for $i = 0, 1, \dots, T - 1$. More specifically, let $R_i(t) = \sum_{j=0}^t R_{ij}$ for $i = 0, 1, \dots, T - 1$ and $t = 0, 1, \dots, T - 1$. Then, at time t , a total of $R_i(t)$ bits are allocated to α_i to produce $\hat{\alpha}_i(t)$ for $i = 0, 1, \dots, T - 1$. The decoder then produces an estimate of the entire signal $\hat{r}(t) = L^{-1}V\hat{\alpha}(t)$ and pulls out the k th component. Note that $\hat{r}(t) = (\hat{r}_0(t), \hat{r}_1(t), \dots, \hat{r}_{T-1}(t))$ and the worst-case cost is then

$$\begin{aligned}
\sup_{r \in \mathcal{C}_r} \|W(r - \hat{r})\|_2^2 &= \sup_{\{z \mid \|z\|_2 \leq 1\}} \|WL(\hat{z} - z)\|_2^2 \\
&= \sup_{\{z \mid \|z\|_2 \leq 1\}} \|U\Sigma V^*(\hat{z} - z)\|_2^2 \\
&= \sup_{\{\alpha \mid \|\alpha\|_2 \leq 1\}} \|\Sigma(\hat{\alpha} - \alpha)\|_2^2 \tag{17}
\end{aligned}$$

$$= \sup_{\{\alpha \mid \|\alpha\|_2 \leq 1\}} \sum_{i=0}^{T-1} (\hat{\alpha}_i - \alpha_i(i))^2 \sigma_i^2 \tag{18}$$

$$\leq \sup_{\{\alpha \mid \|\alpha\|_2 \leq 1\}} \sum_{i=0}^{T-1} 2^{-2R_i(i)} \sigma_i^2 |\alpha_i|^2 \tag{19}$$

$$\leq \max_i 2^{-2R_i(i)} \sigma_i^2 \sup_{\{\alpha \mid \|\alpha\|_2 \leq 1\}} \sum_{i=0}^{T-1} |\alpha_i|^2 \tag{20}$$

$$= \max_i 2^{-2R_i(i)} \sigma_i^2. \tag{21}$$

To derive the upper bound γ_{NC} using the above SVD coding scheme, we construct rate matrix \mathcal{R}_{NC} to solve the following optimization problem:

$$\min_{\mathcal{R}_{NC}} \max_i 2^{-2R_i(i)} \sigma_i^2 \tag{22}$$

$$s.t. \quad \sum_{j=1}^{T-1} R_{ij} = R \quad \text{for } i = 0, 1, \dots, T-1$$

$$R_{ij} \geq 0 \quad \forall i, j.$$

Note that the above optimization problem is similar to (6) which computes γ_{NN} . We note that there may exist singular values σ_i for $i = 0, 1, \dots, T-1$ and channel rates R that may result in $\gamma_{NC} = \gamma_{NN}$. In section 7.2, we illustrate a practical navigation control problem that gives rise to a noncausal encoder and causal decoder pair.

7. FINITE-HORIZON CONTROL APPLICATIONS

In this section, we show how the above analysis of general finite-length signal reconstruction enables us to quantify bounds on performance for two finite-horizon finite-rate feedforward control problems.

7.1. Finite-Horizon Tracking

A direct application for the case where the encoder and decoder are both causal is when one wants to track a finite-length reference command in real-time (finite-horizon cost) using finite-rate measurements. We consider the tracking set up shown in Figure 10 where H is a remote LTI system (invertible) with state-space description $ss(A, B, C, D)$ to be controlled, and would like to minimize $\sup_{r \in \mathcal{C}_r} \|W(r - y)\|_2^2$.

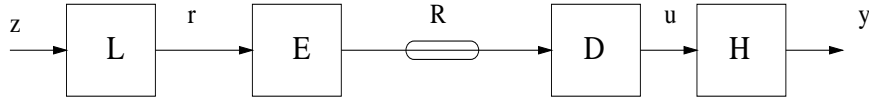


Figure 10. Finite Horizon Tracking Set Up

In section 8, we compute and compare the universal lower bound (γ_{LB}) to the non-causal upper bound (γ_{NN}) and the causal upper bound (γ_{CC}) for different system parameters R, L , and W .

7.2. Finite-Horizon Navigation

A less obvious application for the case where the encoder is noncausal and the decoder is causal is when one wants to navigate the state of a remote system under finite-rate measurements of its initial condition.

Assume that the remote system has some unknown initial condition x_0 which lies in a known bounded ellipsoid in \mathbb{R}^n and 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. This navigation problem can be analyzed as the cascade of systems shown in Figure 11.

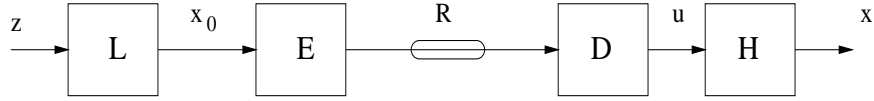


Figure 11. Finite Horizon Navigation Set Up

Specifically,

- $z \in \mathbb{R}^n$ *s.t.* $\|z\|_2 \leq 1$,
- $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$ 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; and the state vector at time t is x_t .

Our navigation metric is

$$\min_{(E, D)} \sup_{x_0 \in \mathcal{C}_{x_0}} \|x_T\|_2^2. \quad (23)$$

We next compute a universal lower bound for (23) in a similar fashion as that computed in section 3. First note that $x_T = A^T x_0 + Mu$, where $Mu = \sum_{i=0}^{T-1} A^{-(i+1)} Bu_i$. Since the system H is reachable, M is a $n \times T$ full rank matrix, and if we know x_0 exactly then we can construct a control input u such that $Mu = -A^T x_0$ which would make $x_T = 0$. Under finite-rate constraints, we must construct u such that $Mu = -A^T \hat{x}_0$, where \hat{x}_0 can take on at most 2^{RT} values in \mathcal{C}_{x_0} . The corresponding metric is then $\min_{(E,D)} \sup_{x_0 \in \mathcal{C}_{x_0}} \|A^T(x_0 - \hat{x}_0)\|_2^2$, which looks like our general reconstruction error metric $\min_{(E,D)} \sup_{r \in \mathcal{C}_r} \|W(r - \hat{r})\|_2^2$ with W replaced with A^T and r replaced with x_0 . The difference is that A^T is an $n \times n$ matrix while W is a $T \times T$ matrix and $r \in \mathbb{R}^T$ while $x_0 \in \mathbb{R}^n$. The time horizon over which the metric is computed for both problems is T . We state the following theorem.

Theorem 7.1. *Given the navigation set up defined above, assume that $\det(A) \neq 0$, $\det(L) \neq 0$. Then,*

$$\gamma_{LB}^{NAV} = 2^{-2RT/n} \{|\det(L)| |\det(A^T)|\}^{\frac{2}{n}}.$$

Proof. *The proof follows that given in Theorem 3.1. The set of all possible initial conditions, $\mathcal{C}_{x_0} \triangleq \{x_0 \in \mathbb{R}^n | x_0 = Lz, z'z \leq 1\} = \{x \in \mathbb{R}^n | (L^{-1}x_0)'(L^{-1}x_0) \leq 1\}$ is a bounded ellipsoid in \mathbb{R}^n 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}^n . Over a horizon T , the channel sends a total of RT bits which limits the decoder output to take on no more than 2^{RT} values. That is, u is chosen such that $Mu = -A^T \hat{x}_0$, where \hat{x}_0 can take on at most 2^{RT} values in \mathcal{C}_{x_0} . Note that is always a solution u since M is full rank.*

Consider a selection of reconstructed estimates $\hat{x}_0^1, \hat{x}_0^2, \dots, \hat{x}_0^{2^{RT}}$, which correspond to inputs $u^1, u^2, \dots, u^{2^{RT}}$, respectively. We must then map each $x_0 \in \mathcal{C}_{x_0}$ to exactly one \hat{x}_0^i ,

$i = 1, 2, \dots, 2^{RT}$. Such a mapping induces a partition on \mathcal{C}_{x_0} . In particular, define $P_i = \{x_0 \in \mathcal{C}_{x_0} \mid x_0 \rightarrow \hat{x}_0^i\}$ for $i = 1, 2, \dots, 2^{RT}$. Now, suppose that the selection $\hat{x}_0^1, \hat{x}_0^2, \dots, \hat{x}_0^{2^{RT}}$ were chosen such that $\|A^T(x_0 - \hat{x}_0^i)\|_2^2 \leq \gamma$ for all $x_0 \in P_i$, and for all i . Then necessarily $P_i \subseteq S_{\hat{x}_0^i}^\gamma \triangleq \{x_0 \in \mathbb{R}^n \mid (x_0 - \hat{x}_0^i)'(A^T)'A^T(x_0 - \hat{x}_0^i) \leq \gamma\}$. Note that $S_{\hat{x}_0^i}^\gamma$ is a bounded ellipsoid in \mathbb{R}^n centered at point \hat{x}_0^i with volume $\eta(\sqrt{\gamma})^n \det\{(A^T)'A^T\}^{-0.5} = \frac{\eta\sqrt{\gamma}^n}{|\det(A^T)|}$. Since $P_i \subseteq S_{\hat{x}_0^i}^\gamma$ for each $i = 1, 2, \dots, 2^{RT}$, it is necessary that 2^{RT} bounded ellipsoids ($S_{\hat{x}_0^i}^\gamma$) cover the set \mathcal{C}_{x_0} . This implies that $2^{RT} \times \text{volume}(S_{\hat{x}_0^i}^\gamma) \geq \text{volume}(\mathcal{C}_{x_0})$. Equivalently,

$$2^{RT} \geq \frac{\text{volume}(\mathcal{C}_{x_0})}{\text{volume}(S_{\hat{x}_0^i}^\gamma)} = \frac{|\det(L)| |\det(A^T)|}{(\sqrt{\gamma})^n}.$$

After rearranging terms, we get that $\gamma_{LB}^{NAV} = 2^{-2RT/n} \{|\det(L)| |\det(A^T)|\}^{\frac{2}{n}}$. ■

We note that this lower bound can be computed from the counting argument presented in [28] where asymptotic stability under finite-rate measurements is studied for the case $T \rightarrow \infty$ and A being unstable. We make a few observations regarding γ_{LB}^{NAV} :

- γ_{LB} depends on L (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 in section 8) to rewrite the lower bound in terms of the singular and eigenvalues of the matrix $A^T L$ as follows:

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

Now, we show how we implement the coding scheme described in section 6 to compute an upper bound, γ_{NC}^{NAV} , for (23). Specifically, we consider the following navigation error

$$\min_{\mathcal{R}_{NC}, u} \sup_{x_0 \in \mathcal{C}_{x_0}} \|x_T\|_2^2. \quad (24)$$

where \mathcal{R}_{NC} is a rate matrix introduced in section 6. Note that $x_T = A^T x_0 + Mu = A^T(x_0 - \hat{x}_0) + Mu + A^T \hat{x}_0$, where again $Mu = \sum_{i=0}^{T-1} A^{-(i+1)} B u_i$. Therefore,

$$\min_{\mathcal{R}_{NC}, u} \sup_{x_0 \in \mathcal{C}_{x_0}} \|x_T\|_2 = \min_{\mathcal{R}_{NC}, u} \sup_{x_0 \in \mathcal{C}_{x_0}} \|A^T(x_0 - \hat{x}_0) + Mu + A^T \hat{x}_0\|_2 \quad (25)$$

$$\leq \min_{\mathcal{R}_{NC}} \sup_{x_0 \in \mathcal{C}_{x_0}} \|A^T(x_0 - \hat{x}_0)\|_2 + \min_u \|Mu + A^T \hat{x}_0\|_2. \quad (26)$$

The two terms in the last inequality above is computed by first applying the coding scheme presented in section 6 to compute the $T \times n$ rate matrix, \mathcal{R}_{NC}^* , that minimizes $\sup_{x_0 \in \mathcal{C}_{x_0}} \|A^T(x_0 - \hat{x}_0)\|_2$, and then computing the control input u that minimizes $\|Mu + A^T \hat{x}_0\|_2$ given \mathcal{R}_{NC}^* . This would give the following upper bound on the navigation error (24)

$$\min_{\mathcal{R}_{NC}, u} \sup_{x_0 \in \mathcal{C}_{x_0}} \|x_T\|_2^2 \leq \gamma_{NC}^{NAV} + \min_u \|Mu + A^T \hat{x}_0\|_2,$$

where γ_{NC}^{NAV} is the optimal cost to

$$\min_{\mathcal{R}_{NC}} \max_i 2^{-2R_i(i)} \sigma_i^2 \quad (27)$$

$$s.t. \quad \sum_{j=1}^n R_{ij} = R \quad \text{for } i = 0, 1, \dots, T-1$$

$$R_{ij} \geq 0 \quad \forall i, j$$

and σ_i is the i 'th singular value of $A^T L$. Finally, $\min_u \|Mu + A^T \hat{x}_0\|_2$ is solved assuming the causal decoder computes the control input u_t for each time step t by solving the following optimization problem for $t \geq 0$:

$$\begin{aligned} \min_{u_t} \quad & \sup_{\alpha \in S_t} \|AU\Sigma\alpha + \sum_{i=0}^{t-2} A^{t-1-i} Bu_i + Bu_t\|_2 \\ \text{s.t.} \quad & S_t = \{\alpha \in \mathbb{R}^n \mid |\alpha_i - \hat{\alpha}_i(t)| \leq |\alpha_i| 2^{-R_i(t)} \quad i = 1, 2, \dots, n\}. \end{aligned} \quad (28)$$

The solution to (28) can be computed easily. First, we use the property that the supremum of a convex function over a bounded interval occurs at a boundary, which implies that $\alpha^*(t)$, the solution to $\sup_{\alpha \in S_t} \|AU\Sigma\alpha + \sum_{i=0}^{t-2} A^{t-1-i} Bu_i + Bu_t\|_2$, belongs to the following set of vectors:

$$\begin{bmatrix} \hat{\alpha}_1(t) \pm |\alpha_1| 2^{-R_1(t)} \\ \hat{\alpha}_2(t) \pm |\alpha_2| 2^{-R_2(t)} \\ \vdots \\ \hat{\alpha}_n(t) \pm |\alpha_n| 2^{-R_n(t)} \end{bmatrix}.$$

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. The optimal control input at time t is $u_t^* = \frac{(AU\Sigma\alpha^*(t) + \sum_{i=0}^{t-2} A^{t-1-i} Bu_i)' B}{(B' B)}$.

8. PERFORMANCE COMPARISON

In this section, we fix H to be an LTI system and quantify tradeoffs between time horizon and performance for tracking and navigation.

8.1. Tracking Performance

To analyze finite-horizon tracking performance we fix H and compute γ_{LB}, γ_{NN} , and γ_{CC} summarized in Table 1 for different rates (R), signal sets (L) and different time horizons T .

Specifically, we consider diagonal matrices $L = \text{diag}(\sigma_0, \sigma_1, \dots, \sigma_{T-1})$ when

1. $\sigma_i = \frac{(0.5)^i}{(0.5)^T}$ for $i = 1, 2, \dots, T$,
2. $\sigma_i = \frac{(0.9)^i}{(0.9)^T}$ for $i = 1, 2, \dots, T$,
3. $\sigma_i = (1.1)^i$ for $i = 1, 2, \dots, T$,

and vary the channel rate $R \in \{5, 10, 15\}$ and $T = 1, 2, \dots, 25$. The expressions for each bound for the case where $W = I$ are given in Table 1 below.

Encoder/Decoder	Bound
Universal	$\gamma_{LB} = 2^{-2R} \prod_{i=0}^{T-1} \sigma_i^{\frac{2}{T}}$
Noncausal/Noncausal	$\gamma_{NN} = \max \{2^{\frac{-2RT}{ I }} \prod_{i \in I} \sigma_i^{\frac{2}{ I }}, \max_i \sigma_i^2\}$.
Causal/Causal	$\gamma_{CC} = \sum_{i=0}^{T-1} 2^{-2R} \sigma_i^2$

Figure 12 illustrates the bounds for these different cases and we make the following

observations.

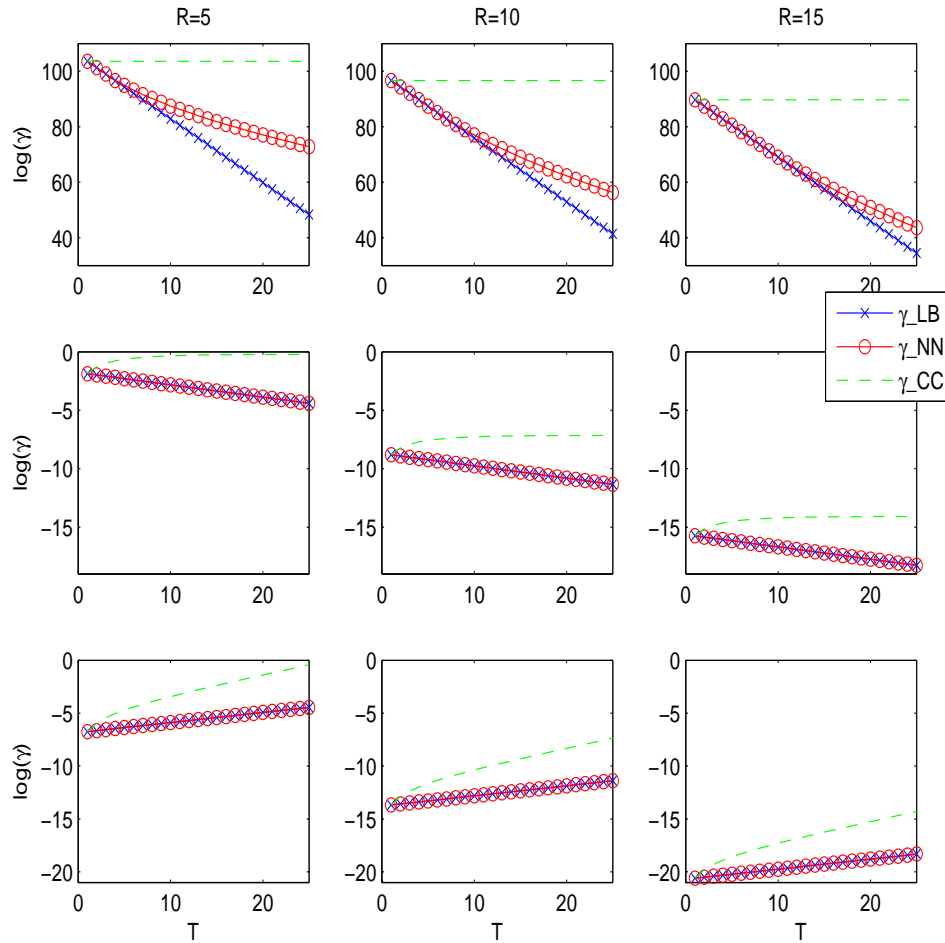


Figure 12. Top: Logarithm of bounds for $\sigma_i = \frac{(0.5)^i}{(0.5)^T}$ for $i = 1, 2, \dots, T$, Middle: Logarithm of bounds for $\sigma_i = \frac{(0.9)^i}{(0.9)^T}$ for $i = 1, 2, \dots, T$, Bottom: Logarithm of bounds for $\sigma_i = (1.1)^i$ for $i = 1, 2, \dots, T$.

- As expected, performance increases as the channel rate R increases since the bounds all decrease as R increases.
- When the eigenvalues of L 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 noncausal upper bound is larger than both the lower bound and the causal upper bound.

- When the eigenvalues of L are exponentially increasing a similar argument as above shows that the lower bound approaches ∞ as $T \rightarrow \infty$.
- The upper bound γ_{NN} is either equal to or comparable to the lower bound γ_{LB} . γ_{NN} deviates from the lower bound when T decays very rapidly (case 1 when the singular values decay by a factor of 10 as T increases) in our examples. The bounded ellipsoid signal set in this case is dominated by the major axis. Whereas, when the signal set looks more symmetric like a sphere, $\gamma_{NN} \approx \gamma_{LB}$.

8.2. Navigation Performance

We now compare the lower and upper navigation bounds on γ^{NAV} 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 L that is generated by LTI system $ss(A_l, B_l, C_l, D_l)$, and we fix the rate $R = 5$. Under such conditions,

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

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

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

We make a few observations.

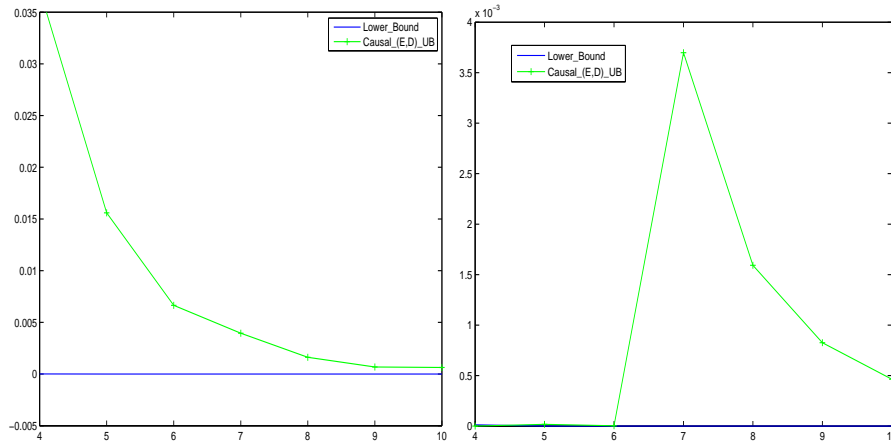


Figure 13. Left: Bounds for $L = ss(0.99, 0.99, 1, 1)$ and A stable, Right: Bounds for $L = ss(0.01, 0.01, 1, 1)$ and A stable

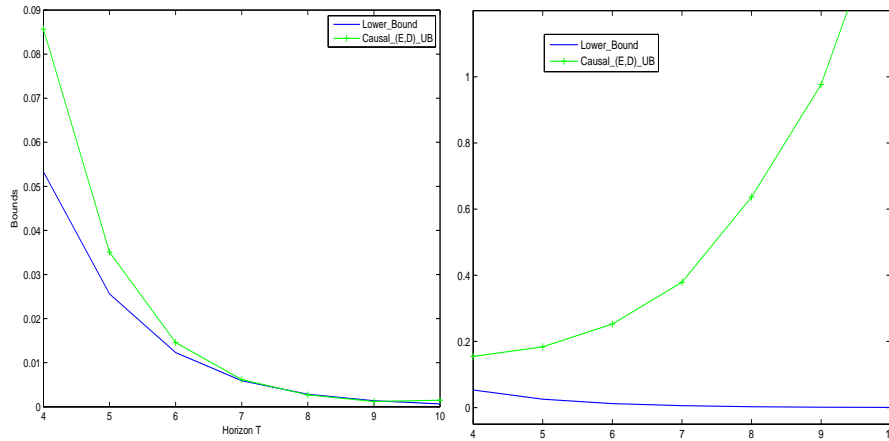


Figure 14. Left: Bounds for $L = ss(0.99, 0.99, 1, 1)$ and A unstable, Right: Bounds for $L = ss(0.01, 0.01, 1, 1)$ and A unstable,

- *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 L is closer to the unit disk.

- *Pole of L* : When the pole of the system that generates L is closer to the origin, the singular values of L 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 L 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 L is closer to 0, then \mathcal{C}_{x_0} looks more and more like an n -dimensional sphere.

In this paper, we first present a deterministic approach for reconstructing finite-length signals that are transmitted through a finite-rate noiseless channel. We allow the class of signals to belong to a bounded ellipsoid and derive a universal lower bound on a worst-case reconstruction error. We then compute upper bounds on the error that arise from different coding schemes. In particular, we study how imposing causality on the encoder and decoder impact the reconstruction error. We then map our general reconstruction problem into two important control problems 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. Our reconstruction framework allows us to study finite-horizon control and quantify tradeoffs between time horizon and performance accuracy.

9. CONCLUSIONS

10. ACKNOWLEDGEMENTS

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

APPENDIX

The proof for theorem 5.1 is given below.

Proof.

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

$$Q \triangleq (L^{-1})'(L^{-1}) = \begin{bmatrix} q_0 & & & & \\ & q_1 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & q_{T-1} \end{bmatrix}, \quad W'W = \begin{bmatrix} w_{00} & w_{01} & \dots & w_{0,T-1} \\ w_{01} & w_{11} & \dots & w_{1,T-1} \\ \vdots & & \ddots & \vdots \\ w_{T-1,1} & w_{T-1,1} & \dots & w_{T-1,T-1} \end{bmatrix}.$$

Then, $\mathcal{C}_r = \{r \in \mathbb{R}^2 \mid \sum_{i=0}^{T-1} q_i r_i^2 \leq 1\}$. Now, since r is transmitted in a causal manner, the class of signals forces $|r_i| \leq \sqrt{\frac{1 - \sum_{k=0}^{i-1} q_k r_k^2}{q_i}}$ for $i = 0, 1, 2, \dots, T-1$. At the decoder end, the

transmitted signal (r_0, r_1, \dots, r_i) at time $t = i$ is not perfectly known. Therefore, it computes each component's interval of uncertainty as $r_j \in [\hat{r}_j(i) - M_j 2^{-R_j(i)}, \hat{r}_j(i) + M_j 2^{-R_j(i)}]$, where

$$M_j = \sqrt{\frac{1 - \sum_{k=0}^{j-1} q_k (\hat{r}_k(i) - \text{sign}(\hat{r}_k(i)) M_k 2^{-R_k(i)})^2}{q_j}} \quad j \leq i, \quad (29)$$

and $R_k(i) = \sum_{j=k}^i R_{kj}$.[§] Next, we note that the encoder operator at time $t = i$ is

$$E(r_j; i) = \hat{r}_j(i) = n_j M_j 2^{-R_j(i)} \quad (30)$$

$$n_j \in S_j \triangleq \{\pm 1, \pm 3, \dots, \pm 2^{R_j(i)} - 1\},$$

for $(n_j - 1)M_j 2^{-R_j(i)} \leq r_j < (n_j + 1)M_j 2^{-R_j(i)}$. We now restrict ourselves to this case of $r \geq 0$ due to symmetry in \mathcal{C}_r and plug (16) into (15) to get

$$M_j = \sqrt{\frac{1 - \sum_{k=0}^{j-1} q_k ((n_k - 1) M_k 2^{-R_k(i)})^2}{q_j}} \quad j \leq i, \quad n_j \in S_j. \quad (31)$$

Note that M_j is a function of $(n_0, n_1, \dots, n_{j-1})$. Now, we can compute an upper bound on the cost function. Since the largest error $|r_j - \hat{r}_j(j)| \leq M_j 2^{-R_{jj}}$ for all j , we get that

[§]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.

$$\begin{aligned} \|W(r - \hat{r})\|_2^2 &\leq \sum_{i=0}^{T-1} w_{ii} (M_i 2^{-R_{ii}})^2 + \sum_{i=0}^{T-1} \sum_{j=i+1}^{T-1} w_{ij} [M_i 2^{-R_{ii}}] [M_j 2^{-R_{jj}}]. \quad (32) \\ &= \sum_{i=0}^{T-1} C_{1,i} M_i^2 + 2 \sum_{i=0}^{T-1} \sum_{j=i+1}^{T-1} C_{2,ij} M_i M_j, \end{aligned}$$

where $C_{1,i} = w_{ii} 2^{-2R_{ii}}$, and $C_{2,ij} = 2w_{ij} 2^{-(R_{ii}+R_{jj})}$. To maximize the upper bound over all $n_k \in S_k \quad \forall k$, we can take partial derivatives of $f(n_0, \dots, n_{T-1}) = \sum_{i=0}^{T-1} C_{1,i} M_i^2 + 2 \sum_{i=0}^{T-1} \sum_{j=i+1}^{T-1} C_{2,ij} M_i M_j$ with respect to n_k assuming n_k is continuous for $k = 0, 1, \dots, T-1$, and set each to zero to see if the corresponding n_k^* 's lie in the corresponding sets S_i 's. When we take partial derivatives, we get

$$\frac{\partial f}{\partial n_k} = 2 \sum_{i=k+1}^{T-1} C_{1,i} M_i \frac{\partial M_i}{\partial n_k} + 2 \sum_{i=k+1}^{T-1} \sum_{j=i+1}^{T-1} C_{2,ij} \left(\frac{\partial M_i}{\partial n_k} M_j + M_i \frac{\partial M_j}{\partial n_k} \right).$$

After some algebra, one can show that $n_k^* = 1 \in S_k$ is a unique solution to $\frac{\partial f}{\partial n_k} = 0$ and $\frac{\partial^2 f}{\partial n_k^2} |_{n_k^*} < 0$, for $k = 0, 1, \dots, T-1$. Alternatively, one can show that the M_j 's (17) are always positive and concave functions of n_0, n_1, \dots, n_{j-1} for all j . Furthermore, M_j is maximized for $n_k = 1 \quad k = 0, 1, \dots, j-1$. Since the upper bound (18) is a positive linear combination of such functions or positive linear combinations of products of such functions, the upper bound itself is concave maximizing at $n_k = 1 \quad k = 0, 1, \dots, j-1$.

We then get

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

Finally, it is straightforward to show that when $r_i^* = 0$, for $i = 0, 1, \dots, T-1$, $\hat{r}_i^*(i) = M_i 2^{-R_{ii}}$, 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:

$$\begin{aligned}
 \min_{\mathcal{R}} \quad & \sum_{i=0}^{T-1} w_{ii} \frac{2^{-2R_{ii}}}{q_i} + 2 \sum_{i=0}^{T-1} \sum_{j=i+1}^{T-1} w_{ij} \frac{2^{-(R_{ii}+R_{jj})}}{\sqrt{q_i q_j}} \\
 \text{s.t.} \quad & \sum_{j=0}^i R_{ij} = R \quad i = 0, 1, \dots, T-1 \\
 & R_{ij} \geq 0 \quad \forall i, j.
 \end{aligned} \tag{33}$$

Since the off diagonal entries in \mathcal{R} do not enter the cost and only enter in the constraints, we get that a lower bound to the optimal cost of (19) is the optimal cost to

$$\begin{aligned}
 \min_{R_{ii}} \quad & \sum_{i=0}^{T-1} w_{ii} \frac{2^{-2R_{ii}}}{q_i} + 2 \sum_{i=0}^{T-1} \sum_{j=i+1}^{T-1} w_{ij} \frac{2^{-(R_{ii}+R_{jj})}}{\sqrt{q_i q_j}} \\
 \text{s.t.} \quad & 0 \leq R_{ii} \leq R \quad \forall i.
 \end{aligned} \tag{34}$$

It is straightforward to show that the optimal solution to (20) is $R_{ii} = R$ for all i . Finally, $R_{ii} = R$ and $R_{ij} = 0$ for all i, j , results in the cost of (19) to achieve the lower bound cost of (19). Therefore, the optimal rate matrix is diagonal and

$$\gamma_{CC} = \sum_{i=1}^{T-1} w_{ii} \frac{2^{-2R}}{q_i} + 2 \sum_{i=0}^{T-1} \sum_{j=i+1}^{T-1} w_{ij} \frac{2^{-2R}}{\sqrt{q_i q_j}}.$$

■

REFERENCES

1. Allen, G, Gray, RM. *Vector Quantization and Signal Compression*. Kluwer Academic Publishers, 1992.
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* 2002.
3. Wong, WS, Brockett, R. Systems with Finite Communication Bandwidth Constraints-I: State Estimation Problems. *IEEE Transactions on Automatic Control* 1997; **42**(9).
4. Brockett, R. Quantized Feedback Stabilization of Linear Systems. *IEEE Transactions on Automatic Control* 2000; **45**: 1279-1289.
5. Wong, WS, Brockett, R. Systems with Finite Communication Bandwidth Constraints-II: Stabilization with Limited Information Feedback. *IEEE Transactions on Automatic Control* 1999; **44**(5).
6. Delevenne, JC, Blondel, V. Complexity of Control of Finite Automata. *IEEE Transactions on Automatic Control* 2006; **51**(6): 977- 986.
7. Delchamps, D. Stabilizing a Linear System with Quantized State Feedback. *IEEE Transactions on Automatic Control* 1990; **35**(8).
8. Delevenne, JC. An optimal quantized feedback strategy for scalar linear systems. *IEEE Transactions on Automatic Control* 2006; **51**(2): 298-303.
9. Doyle, J, Francis, B, Tannenbaum, A, *Feedback Control Theory* Macmillan Publishing Company: New York, 1992.
10. Dumer I., Pinsker M.S., and Prelov V.V. On Coverings of Ellipsoids in Euclidean Spaces *IEEE Transactions on Information Theory* **50**(10): October, 2004.
11. Elia, N, Mitter, S., Stabilization of Linear Systems With Limited Information *IEEE Transactions on Automatic Control* 2001; **46**(9).
12. Fagnani, F, Zampieri, S. Stability Analysis and Synthesis for Linear Systems With Quantized Feedback. *IEEE Transactions on Automatic Control* 2003; **48**(9).
13. Fagnani, F, Zampieri, S. Quantized Stabilization of Linear Systems: Complexity Versus Performance. *IEEE Transactions on Automatic Control* 2004; **49**(9).
14. Gallager, R., *Information Theory and Reliable Communication*. John Wiley and Sons, Inc.: New York, 1968.
15. Gersho, A., Gray R.M., *Vector Quantization and Signal Compression* Kluwer Academic Publishers: Boston, c1992.
16. Huang J.J.Y., Schultheiss, P.M., Block Quantization of Correlated Gaussian Random Variables *IEEE Transactions on Communication System* 1963; **11**(3), 289-296.

17. Ishii, H, Francis, B., *Limited data rate in control systems with networks*. Springer: New York, 2002.
18. Lee, EA, Messerschmitt, DG. *Digital Communication*. (2nd edn) Kluwer Academic Publishers, 1994.
19. Liberzon, D., A note on stabilization of linear systems using coding and limited communication. *Proceedings of the 41st IEEE Conference on Decision and Control* 2002;
20. Ling, Q, Lemmon MD. Stability of quantized control systems under dynamic bit assignment. *IEEE Transactions on Automatic Control* 2005; **50**(5): 734-740.
21. Martins, NC, Dahleh MA, and Elia N. Feedback stabilization of uncertain systems in the presence of a direct link. *IEEE Transactions on Automatic Control* 2006; **51**(3): 438-447
22. Martins, NC, Dahleh MA. Feedback Control in the Presence of Noisy Channels: 'Bode-Like' Fundamental Limitations of Performance. *Scheduled to appear in the IEEE Transactions on Automatic Control in May 2008* Abridged version in ACC 2005.
23. Martins, NC, Dahleh MA, Doyle JC, Fundamental Limitations of Disturbance Attenuation in the Presence of Side Information. *IEEE Transactions on automatic control* 2007; **52**(1).
24. Nair, GN, Evans, RJ. Stabilization with Data-rate-limited Feedback: tightest attainable bounds. *Systems and Control Letters* 2000; **41**: 49-56.
25. Nair, GN, Evans, RJ. Stabilizability of stochastic linear systems with finite feedback data rates. *SIAM Journal on Control and Optimization, Society for Industrial and Applied Mathematics, USA* 2004; **43**(2): 413 - 436.
26. Proakis, JG. *Digital Communications* (4th edn). McGraw Hill, 2000.
27. Sarma, SV, Dahleh, MA, Salapaka, S. On Time-Varying Bit-Allocation Maintaining Stability: A Convex Parameterization. *Proceedings of the 43rd IEEE Conference on Decision and Control* 2004.
28. Tatikonda, S. Control Under Communication Constraints. *IEEE Transactions on Automatic Control* 2004; **49**(7): 1056 - 1068.
29. Voulgaris, PG, Hadjicostis, CN, Touri, R. A Robust Control Approach to Perfect Reconstruction of Digital Signals. *IEEE Transactions on Signal Processing*: 2007; **55**(9): 4444-4457.
30. Yuksel, S, Basar, T. State Estimation and Control for Linear Systems over Communication Networks. *Proceedings of 2003 IEEE Conference on Control Applications* 2003.