

## Signal reconstruction in the presence of finite-rate measurements: finite-horizon control applications

Sridevi V. Sarma<sup>\*,†</sup> and Munther A. Dahleh

*Laboratory of Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA, U.S.A.*

### 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 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 (e.g. stability, disturbance rejection). Copyright © 2009 John Wiley & Sons, Ltd.

Received 8 February 2008; Revised 3 September 2008; Accepted 6 November 2008

KEY WORDS: signal reconstruction; finite-rate; finite-horizon; control; 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 [1, 2]. Recent work by Voulgaris investigates reconstruction of *infinite-length discrete-valued* signals that are filtered via noisy channels using a *deterministic* framework [3]. 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

\*Correspondence to: Sridevi V. Sarma, Laboratory of Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA, U.S.A.

†E-mail: sree@mit.edu

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

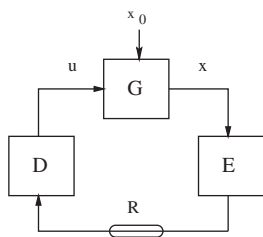


Figure 1. Equivalent closed-loop system.

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 (e.g. 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 [4–18]. 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 that stochastic exogenous signals entering the system [19, 20]. 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 trade-offs between finite-horizon performance and control complexity for linear systems and finite automata

systems [21–24]. A navigation problem similar in spirit to that which is presented here is described in [24]. In [24] Fagnani and Zampieri 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  $x_t$  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  $C = \mu[W]/\mu[V]$  ( $\mu$  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  $\delta(\beta) = H_1 \beta w^{1/\beta}$ , for some  $w > 1$  and constant  $H_1$ , which depends on the plant dynamics, see [24] 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.



Figure 2. General reconstruction setup.

## 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.

Specifically,

- $z \in \mathbb{R}^T$  s.t.  $\|z\|_2 \leq 1$ ,
- $r \in \mathcal{C}_r \triangleq \{r \in \mathbb{R}^T, z \in \mathbb{R}^T \mid r = Lz, \|z\|_2 \leq 1\}$ ,
- $L: \mathbb{R}^T \rightarrow \mathbb{R}^T$  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  $\mathbb{R}^T$ . As  $L$  is linear, it maps the unit ball to a bounded ellipsoid. We assume that  $L$  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  $\|W(r - \hat{r})\|_2^2 = 0 \forall r \in \mathcal{C}_r$ . 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  $\gamma_{\text{LB}}$  and  $\gamma_{\text{UB}}$ , such that

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

Knowledge of  $\gamma_{\text{LB}}$  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  $\gamma_{\text{UB}}$ . To compute  $\gamma_{\text{UB}}$ , 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 [25].

### Theorem 3.1

Given the signal reconstruction setup defined above, assume that  $\det(W) \neq 0$ ,  $\det(L) \neq 0$ . Then,

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

### Proof

The set of all possible signals,  $\mathcal{C}_r \triangleq \{r \in \mathbb{R}^T \mid r = Lz, \|z\|_2 \leq 1\} = \{r \in \mathbb{R}^T \mid (L^{-1}r)'(L^{-1}r) \leq 1\}$ , 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  $\eta$  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  $2^{RT}$  values.

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

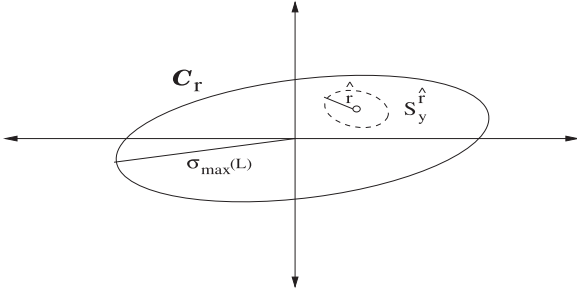


Figure 3. Bounded ellipsoids  $\mathcal{C}_r$  and  $S_r^\gamma$ .

at point  $\hat{r}^i$  with volume  $\eta(\sqrt{\gamma})^T \det\{(W'W)^{-0.5}\} = \eta\sqrt{\gamma}^T / |\det(W)|$ . See Figure 3 for an illustration.

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 ( $S_{\hat{r}_i}^\gamma$ ) cover the set  $\mathcal{C}_r$ . This implies that  $2^{RT} \times \text{volume}(S_{\hat{r}_i}^\gamma) \geq \text{volume}(\mathcal{C}_r)$ . Equivalently,

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

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

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

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

*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  $\{\det(L)\}^{2/T} = (D_l)^2$ .  $\square$

As expected,  $\gamma_{\text{LB}}$  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:

$$\begin{aligned} \gamma_{\text{LB}} &= 2^{-2R} \left\{ \prod_{i=0}^{T-1} \sigma_i(WL) \right\}^{2/T} \\ &= 2^{-2R} \left\{ \prod_{i=0}^{T-1} |\lambda_i(WL)| \right\}^{2/T} \end{aligned}$$

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 problem with a deterministic error metric [26, 27]. 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  $2^{-2R} \det(M_r)^{1/T}$  [28], which is identical to our lower bound if we replace  $WL$  with  $M_r^{1/2}$ . 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  $\gamma_{\text{LB}}$  in our deterministic setting and compare it with 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 ( $\gamma_{\text{NN}}$ ), the second is computed by constructing a noncausal encoder and a causal decoder ( $\gamma_{\text{NC}}$ ), and the third is computed by constructing a causal encoder and decoder ( $\gamma_{\text{CC}}$ ).

#### 4. NONCAUSAL ENCODING AND DECODING

In this section, we derive an upper bound,  $\gamma_{\text{NN}}$ , 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 [28] and we repeat it here to compare with 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,  $\Sigma$  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  $i = 0, 1, \dots, T-1$  ( $\alpha = (\alpha_0 \ \alpha_1 \ \alpha_{T-1})$ ).

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.’

Note that for the above SVD coding scheme,

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

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

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

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

$$= \max_i 2^{-2R_i} \sigma_i^2 \quad (5)$$

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

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

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

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

Now note that a lower bound to the optimal cost of (7) is the optimal solution to<sup>‡</sup>

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

<sup>‡</sup>We allow the rates to take on noninteger values to solve for an optimal bit-allocation strategy. The resulting noninteger-valued rates can be interpreted as average rates over time.

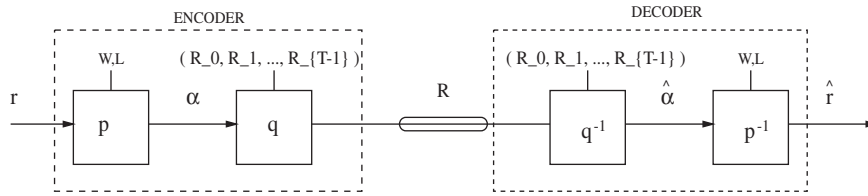


Figure 4. An SVD coding scheme.

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

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

The minimization over nonnegative  $R_i$  (second term in (9)) is as follows:

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

which gives us

$$\min_{\gamma \geq 0} \gamma - \sum_{i=0}^{T-1} \lambda_i \log(\gamma) + 2 \sum_{i=0}^{T-1} \lambda_i \log(\sigma_i) - \mu RT \quad (10)$$

s.t.  $0 \leq \lambda_i \leq \frac{\mu}{2} \quad \forall i$

since we know a finite solution to (7) exists. Now, if we minimize over  $\gamma$  by taking the derivative of the cost function with respect to  $\gamma$  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:

$$\max_{\lambda_0, \lambda_1, \dots, \lambda_{T-1}, \mu} \sum_i \lambda_i \left( \frac{1}{\ln(2)} - \log \left( \sum_i \frac{\lambda_i}{\ln(2)} \right) \right) + \sum_i 2\lambda_i \log(\sigma_i) - \mu RT \quad (12)$$

s.t.  $0 \leq \lambda_i \leq \frac{\mu}{2} \quad \forall i$   
 $\mu \geq 0$

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

$$\max_{\mu} \frac{\mu|I|}{2} \left( \frac{1}{\ln(2)} - \log \left( \frac{\mu|I|}{2\ln(2)} \right) \right) + \mu \sum_{i \in I} \log(\sigma_i) - \mu RT \quad (14)$$

$\mu \geq 0$

One can compute the solution to (14) as

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

Finally, we plug  $\mu^*$  into (13) and then (11) to get

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

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 second term in (9) must equal 0. As expected, the optimal rate allocation places more bits to components of  $\alpha$  whose corresponding singular values are larger. Surprisingly, if  $|I| = T$ , then

$$R_i^* = R + \log(\sigma_i) - \frac{1}{T} \sum_{j=0}^{T-1} \log(\sigma_j) \quad \text{for } i=0, 1, \dots, T-1$$

and the resulting upper bound is

$$\gamma_{\text{NN}} = 2^{-2R} \left\{ \prod_{i=0}^{T-1} \sigma_j \right\}^{T/2} = \gamma_{\text{LB}}$$

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

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

Comparing the above inequality to the expression for  $\gamma^*$  in (15), we see that in order to minimize  $\gamma^*$ , 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,  $\gamma_{\text{CC}}$ , 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 that 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

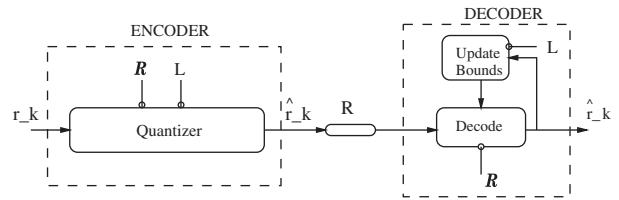


Figure 5. Causal encoding and decoding scheme.

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  $i=0, 1, \dots, T-1$ .

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  $r_0$  that has magnitude less than or equal to  $M_0 = 1/\sqrt{q_0}$ . Its quantization region is on the

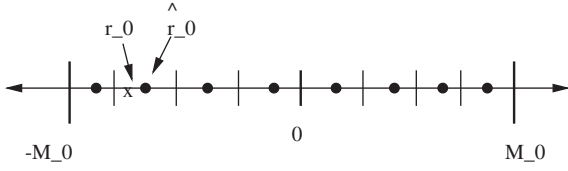


Figure 6. Quantization region at time  $t=0$ .

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 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

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

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  $r_0 \in [4M_02^{-R_{00}}, 6M_02^{-R_{00}})$ , the uncertainty computed by the decoder is the shaded strip shown in Figure 7, and the corresponding  $\hat{r}_0(0) = 5M_02^{-R_{00}}$ .

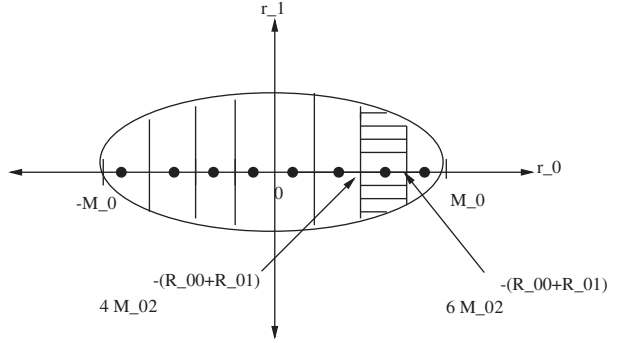


Figure 7. Intervals for  $r_1$  when  $r_0 \in [4M_02^{-R_{00}}, 6M_02^{-R_{00}})$ .

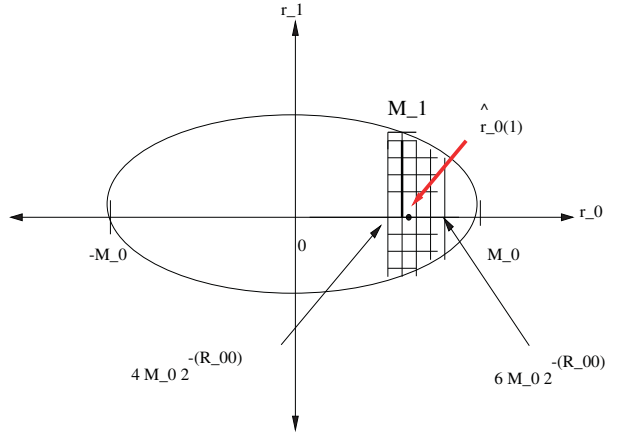


Figure 8. Intervals for  $r_1$  when  $\hat{r}_0(1)$  is received by the decoder.

At time  $t=1$ , the encoder further quantizes  $r_0$  by dividing the interval  $[4M_02^{-R_{00}}, 6M_02^{-R_{00}})$  into  $2^{R_{01}}$  equal length intervals as shown in Figure 8. The bound

$$M_1 = \sqrt{\frac{1 - q_0(\hat{r}_0(1) - M_02^{-(R_{00}+R_{01})})^2}{q_1}}$$

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<sup>§</sup> where

$$Q \triangleq (L^{-1})'(L^{-1}) = \text{diag}(q_0, q_1, \dots, q_{T-1})$$

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

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

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

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 nonsymmetric 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,  $\gamma_{\text{NC}}$ , by constructing a modified SVD coding scheme in which

<sup>§</sup>We prove the theorem assuming that  $L$  is diagonal for a simpler read, however, the general case holds and can be proven using the argument outlined in the appendix.

the encoder is noncausal, i.e. it 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  $\alpha$ . The bit-allocation is determined by the following rate matrix:

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

such that  $\sum_j R_{ij} = R$  for  $i = 0, 1, \dots, T-1$ . More specifically, let  $R_i(t) = \sum_{j=0}^t R_{ji}$  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  $\alpha_i$  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

$$\begin{aligned} \sup_{r \in \mathcal{C}_r} \|W(r - \hat{r})\|_2^2 &= \sup_{\{z \mid \|z\|_2 \leq 1\}} \|W L(\hat{z} - z)\|_2^2 \\ &= \sup_{\{z \mid \|z\|_2 \leq 1\}} \|U \Sigma V^*(\hat{z} - z)\|_2^2 \\ &= \sup_{\{\alpha \mid \|\alpha\|_2 \leq 1\}} \|\Sigma(\hat{\alpha} - \alpha)\|_2^2 \quad (17) \\ &= \sup_{\{\alpha \mid \|\alpha\|_2 \leq 1\}} \sum_{i=0}^{T-1} (\hat{\alpha}_i - \alpha_i)^2 \sigma_i^2 \quad (18) \\ &\leq \sup_{\{\alpha \mid \|\alpha\|_2 \leq 1\}} \sum_{i=0}^{T-1} 2^{-2R_i(i)} \sigma_i^2 |\alpha_i|^2 \quad (19) \end{aligned}$$

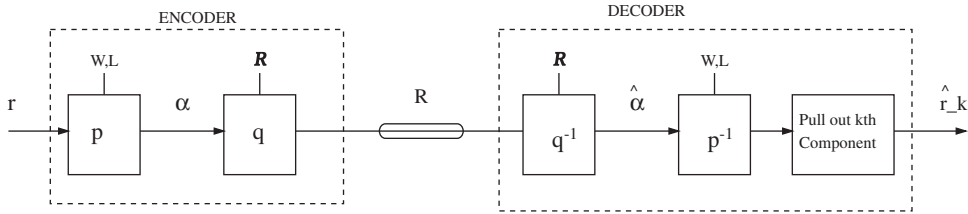


Figure 9. SVD scheme for noncausal encoding and causal decoding.

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

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

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

$$\begin{aligned} \min_{\mathcal{R}_{NC}} \quad & \max_i 2^{-2R_i(i)} \sigma_i^2 \\ \text{s.t.} \quad & \sum_{j=1}^{T-1} R_{ij} = R \quad \text{for } i=0, 1, \dots, T-1 \\ & R_{ij} \geq 0 \quad \forall i, j \end{aligned} \quad (22)$$

Note that the above optimization problem is similar to (6) which computes  $\gamma_{NN}$ . We note that there may exist singular values  $\sigma_i$  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

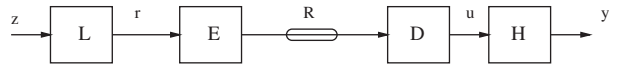


Figure 10. Finite-horizon tracking setup.

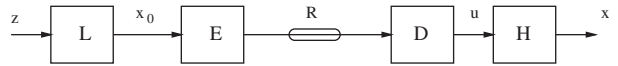


Figure 11. Finite-horizon navigation setup.

track a finite-length reference command in real-time (finite-horizon cost) using finite-rate measurements. We consider the tracking setup 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  $\sup_{r \in \mathcal{C}_r} \|W(r - y)\|_2^2$ .

In Section 8, we compute and compare the universal lower bound ( $\gamma_{LB}$ ) with the noncausal upper bound ( $\gamma_{NN}$ ) and the causal upper bound ( $\gamma_{CC}$ ) 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.

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  $H = \text{ss}(A, B, I, 0)$  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 manner as that computed in Section 3. First note that  $x_T = A^T x_0 + Mu$ , where  $Mu = \sum_{i=0}^{T-1} A^{-(i+1)} Bu_i$ . As the system  $H$  is reachable,  $M$  is an  $n \times T$  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 setup defined above, assume that  $\det(A) \neq 0$ ,  $\det(L) \neq 0$ . Then, a lower bound for (23) is

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

#### 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  $\eta$  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  $P_i = \{x_0 \in \mathcal{C}_{x_0} | x_0 \rightarrow \hat{x}_0^i\}$  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  $P_i \subseteq S_{\hat{x}_0^i}^\gamma \triangleq \{x_0 \in \mathbb{R}^n | (x_0 - \hat{x}_0^i)'(A^T)'A^T(x_0 - \hat{x}_0^i) \leq \gamma\}$ . 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

$$\eta(\sqrt{\gamma})^n \det\{((A^T)'A^T)^{-0.5}\} = \frac{\eta\sqrt{\gamma}^n}{|\det(A^T)|}$$

As  $P_i \subseteq S_{\hat{x}_0^i}^\gamma$  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  $\gamma_{\text{LB}}^{\text{NAV}} = 2^{-2RT/n} \{|\det(L)| |\det(A^T)|\}^{2/n}$ .  $\square$

We note that this lower bound can be computed from the counting argument presented in [16] where asymptotic stability under finite-rate measurements is studied for the case  $T \rightarrow \infty$  and  $A$  having all unstable

eigenvalues. We make a few observations regarding  $\gamma_{LB}^{NAV}$ :

- $\gamma_{LB}$  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:

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

Now, we show how we implement the coding scheme described in Section 6 to compute an upper bound,  $\gamma_{NC}^{NAV}$ , 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  $Mu = \sum_{i=0}^{T-1} A^{-(i+1)} Bu_i$ . Therefore,

$$\begin{aligned} &\min_{\mathcal{R}_{NC}, u} \sup_{x_0 \in \mathcal{C}_{x_0}} \|x_T\|_2 \\ &= \min_{\mathcal{R}_{NC}, u} \sup_{x_0 \in \mathcal{C}_{x_0}} \|A^T(x_0 - \hat{x}_0) + Mu + A^T \hat{x}_0\|_2 \quad (25) \\ &\leq \min_{\mathcal{R}_{NC}} \sup_{x_0 \in \mathcal{C}_{x_0}} \|A^T(x_0 - \hat{x}_0)\|_2 \\ &\quad + \min_u \|Mu + A^T \hat{x}_0\|_2 \quad (26) \end{aligned}$$

The two terms in the last inequality above are 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  $\gamma_{NC}^{NAV}$  is the optimal cost to

$$\begin{aligned} \min_{\mathcal{R}_{NC}} \quad &\max_i 2^{-2R_i(i)} \sigma_i^2 \\ \text{s.t.} \quad &\sum_{j=1}^n R_{ij} = R \quad \text{for } i=0, 1, \dots, T-1 \\ &R_{ij} \geq 0 \quad \forall i, j \end{aligned} \quad (27)$$

and  $\sigma_i$  is the  $i$ th singular value of  $A^T L$ . Finally,  $\min_u \|Mu + A^T \hat{x}_0\|_2$  is solved assuming that the causal decoder computes the control input  $u_t$  for each time step  $t$  by solving the following optimization problem for  $t \geq 0$ :

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

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  $\gamma_{LB}$ ,  $\gamma_{NN}$ , and  $\gamma_{CC}$  summarized in Table I 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.  $\sigma_i = (0.5)^i / (0.5)^T$  for  $i = 1, 2, \dots, T$ ,
2.  $\sigma_i = (0.9)^i / (0.9)^T$  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 I below.

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

- 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

$$\frac{\gamma_{LB}(T+1)}{\gamma_{LB}(T)} = \frac{\{\prod_{i=0}^T \beta^i\}^{2/(T+1)}}{\{\prod_{i=0}^{T-1} \beta^i\}^{2/T}} = \beta < 1$$

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  $\infty$  as  $T \rightarrow \infty$ .

- The upper bound  $\gamma_{NN}$  is either equal to or comparable to the lower bound  $\gamma_{LB}$ .  $\gamma_{NN}$  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  $\gamma^{\text{NAV}}$  with each other for different LTI causal systems  $H = \text{ss}(A, B, C, D)$ , and for different time horizons  $T$ . We consider diagonal  $4 \times 4$  ( $n = 4$ ) state-transition matrices  $A = \text{diag}(a_0, a_1, a_2, a_3)$ , an  $L$  that is generated by LTI system  $\text{ss}(A_l, B_l, C_l, D_l)$ , and we fix the rate  $R = 5$ . Under such conditions,

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

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

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

We make a few observations.

- *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

Table I. Summary of bounds for tracking performance.

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

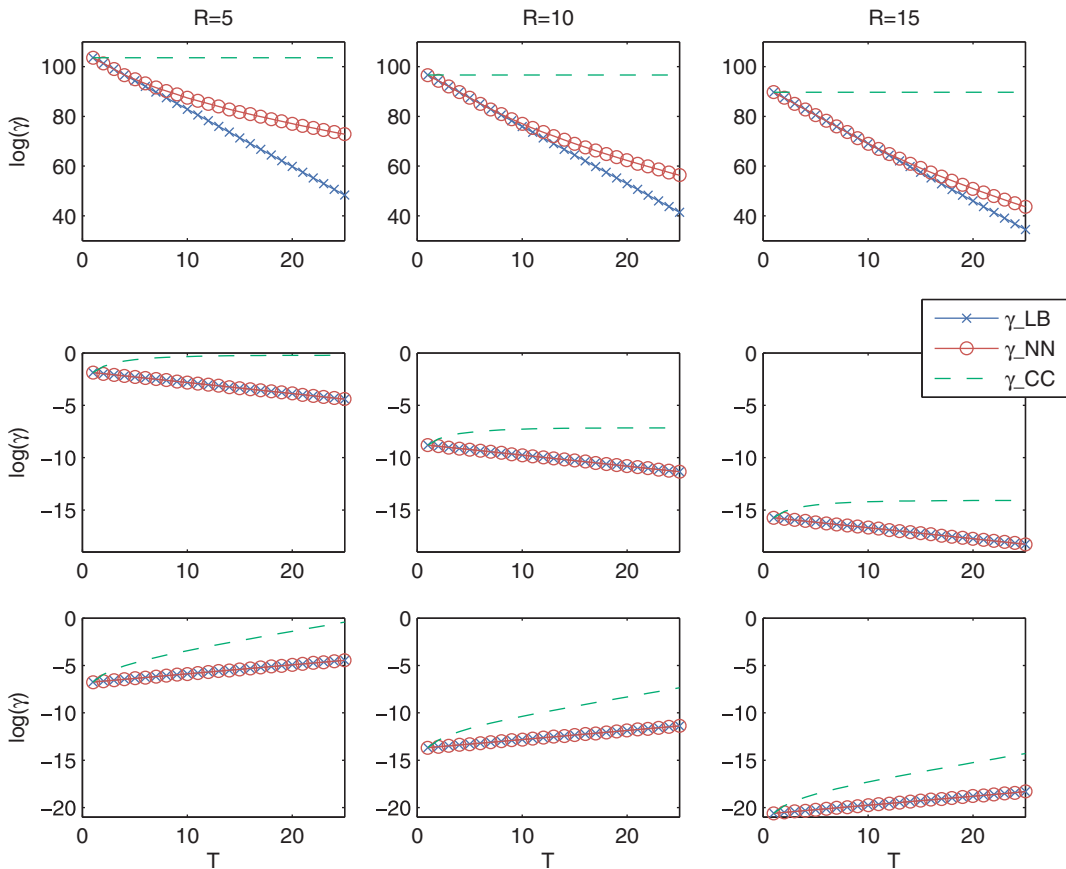


Figure 12. Top: logarithm of bounds for  $\sigma_i = (0.5)^i / (0.5)^T$  for  $i = 1, 2, \dots, T$ ; middle: logarithm of bounds for  $\sigma_i = (0.9)^i / (0.9)^T$  for  $i = 1, 2, \dots, T$ ; bottom: logarithm of bounds for  $\sigma_i = (1.1)^i$  for  $i = 1, 2, \dots, T$ .

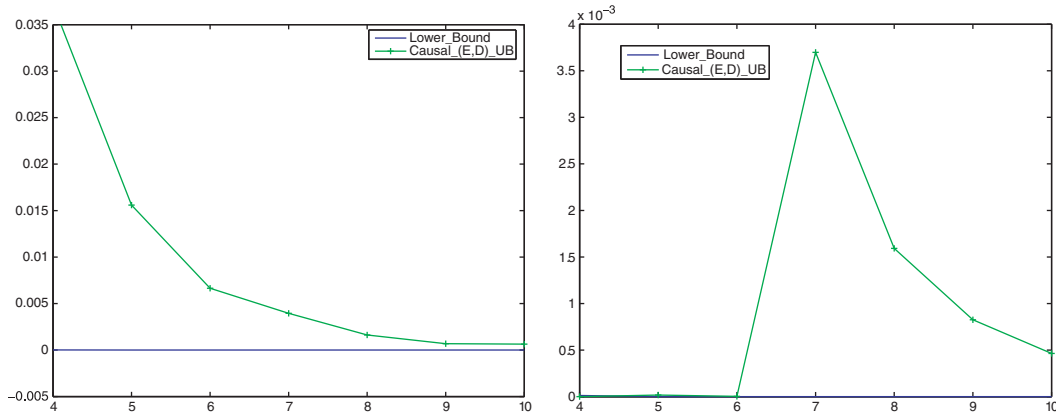


Figure 13. Left: bounds for  $L = ss(0.99, 0.99, 1, 1)$  and  $A$  stable; right: bounds for  $L = ss(0.01, 0.01, 1, 1)$  and  $A$  stable.

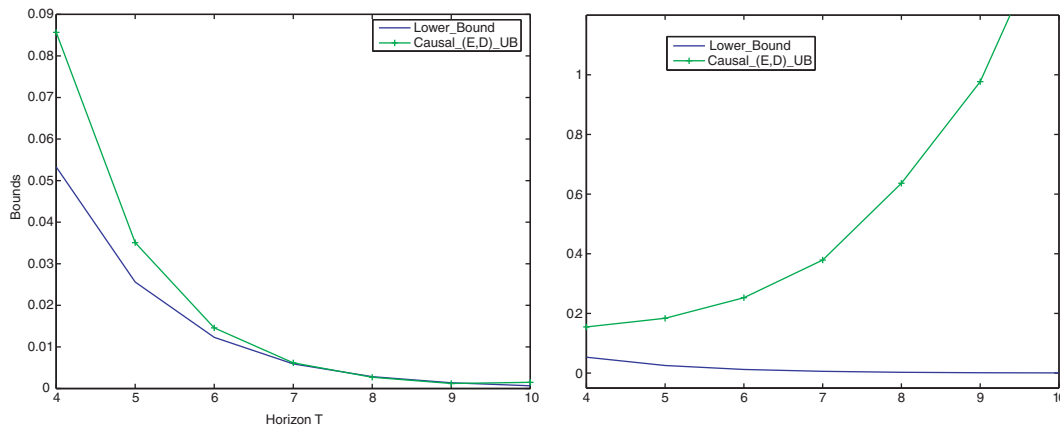


Figure 14. Left: bounds for  $L = ss(0.99, 0.99, 1, 1)$  and  $A$  unstable; right: bounds for  $L = ss(0.01, 0.01, 1, 1)$  and  $A$  unstable.

structure, that is knowing some components of  $x_0$  gives 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.

### 9. CONCLUSIONS

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 impacts 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.

APPENDIX A

The proof for Theorem 5.1 is given below.

*Proof*

We prove the theorem assuming that  $L$  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 setup for where  $L$  (and hence  $Q$ ) is diagonal.

$$Q \triangleq (L^{-1})'(L^{-1}) = \text{diag}(q_0, q_1, \dots, q_{T-1})$$

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

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

$$|r_i| \leq \sqrt{\frac{1 - \sum_{k=0}^{i-1} q_k r_k^2}{q_i}} \quad \text{for } i=0, 1, 2, \dots, T-1$$

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

$$M_j = \sqrt{\frac{1 - \sum_{k=0}^{j-1} q_k (\hat{r}_k(i) - \text{sign}(\hat{r}_k(i)) M_k 2^{-R_k(i)})^2}{q_j}} \quad j \leq i \tag{A1}$$

and  $R_k(i) = \sum_{j=k}^i R_{kj}$ .<sup>‡</sup> Next, we note that the encoder operator at time  $t=i$  is

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

$$n_j \in S_j \triangleq \{\pm 1, \pm 3, \dots, \pm 2^{R_j(i)} - 1\} \tag{A2}$$

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

$$M_j = \sqrt{\frac{1 - \sum_{k=0}^{j-1} q_k ((n_k - 1) M_k 2^{-R_k(i)})^2}{q_j}} \quad j \leq i, \quad n_j \in S_j \tag{A3}$$

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. As the largest error  $|r_j - \hat{r}_j(j)| \leq M_j 2^{-R_{jj}}$  for all  $j$ , we get that

$$\begin{aligned} \|W(r - \hat{r})\|_2^2 &\leq \sum_{i=0}^{T-1} w_{ii} (M_i 2^{-R_{ii}})^2 \\ &\quad + \sum_{i=0}^{T-1} \sum_{j=i+1}^{T-1} w_{ij} [M_i 2^{-R_{ii}}] [M_j 2^{-R_{jj}}] \\ &= \sum_{i=0}^{T-1} C_{1,i} M_i^2 \\ &\quad + 2 \sum_{i=0}^{T-1} \sum_{j=i+1}^{T-1} C_{2,ij} M_i M_j \end{aligned} \tag{A4}$$

<sup>‡</sup>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.

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  $n_k \in S_k \forall k$ , we can take partial derivatives of  $f(n_0, \dots, n_{T-1}) = \sum_{i=0}^{T-1} C_{1,i}M_i^2 + 2\sum_{i=0}^{T-1} \sum_{j=i+1}^{T-1} C_{2,ij}M_iM_j$  with respect to  $n_k$  assuming that  $n_k$  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

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

After some algebra, one can show that  $n_k^* = 1 \in S_k$  is a unique solution to  $\partial f / \partial n_k = 0$  and  $\partial^2 f / \partial n_k^2|_{n_k^*} < 0$ , 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  $n_k = 1 \ k = 0, 1, \dots, j-1$ . 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  $n_k = 1 \ k = 0, 1, \dots, j-1$ .

We then get

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

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

$$\begin{aligned} \min_{\mathcal{R}} \quad & \sum_{i=0}^{T-1} w_{ii} \frac{2^{-2R_{ii}}}{q_i} + 2 \sum_{i=0}^{T-1} \sum_{j=i+1}^{T-1} w_{ij} \frac{2^{-(R_{ii}+R_{jj})}}{\sqrt{q_i q_j}} \\ \text{s.t.} \quad & \sum_{j=0}^i R_{ij} = R_i = 0, 1, \dots, T-1 \\ & R_{ij} \geq 0 \quad \forall i, j \end{aligned} \quad (\text{A5})$$

As 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

$$\begin{aligned} \min_{R_{ii}} \quad & \sum_{i=0}^{T-1} w_{ii} \frac{2^{-2R_{ii}}}{q_i} + 2 \sum_{i=0}^{T-1} \sum_{j=i+1}^{T-1} w_{ij} \frac{2^{-(R_{ii}+R_{jj})}}{\sqrt{q_i q_j}} \\ \text{s.t.} \quad & 0 \leq R_{ii} \leq 0 \quad \forall i \end{aligned} \quad (\text{A6})$$

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

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

#### ACKNOWLEDGEMENTS

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

#### REFERENCES

1. Lee EA, Messerschmitt DG. *Digital Communication* (2nd edn). Kluwer Academic Publishers: Dordrecht, 1994.
2. Proakis JG. *Digital Communications* (4th edn). McGraw-Hill: New York, 2000.
3. 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.
4. Baillieul J. Feedback coding for information-based control: operating near the data-rate limit. *Proceedings of the 41st IEEE Conference on Decision and Control*, Las Vegas, NV, 2002.

5. Wong WS, Brockett R. Systems with finite communication bandwidth constraints—I: state estimation problems. *IEEE Transactions on Automatic Control* 1997; **42**(9).
6. Brockett R. Quantized feedback stabilization of linear systems. *IEEE Transactions on Automatic Control* 2000; **45**:1279–1289.
7. Wong WS, Brockett R. Systems with finite communication bandwidth constraints-II: stabilization with limited information feedback. *IEEE Transactions on Automatic Control* 1999; **44**(5).
8. Delchamps D. Stabilizing a linear system with quantized state feedback. *IEEE Transactions on Automatic Control* 1990; **35**(8).
9. Elia N, Mitter S. Stabilization of linear systems with limited information. *IEEE Transactions on Automatic Control* 2001; **46**(9).
10. Ishii H, Francis B. *Limited Data Rate in Control Systems with Networks*. Springer: New York, 2002.
11. Liberzon D. A note on stabilization of linear systems using coding and limited communication. *Proceedings of the 41st IEEE Conference on Decision and Control*, Las Vegas, NV, 2002.
12. Martins NC, Dahleh MA, Elia N. Feedback stabilization of uncertain systems in the presence of a direct link. *IEEE Transactions on Automatic Control* 2006; **51**(3):438–447.
13. Nair GN, Evans RJ. Stabilization with data-rate-limited feedback: tightest attainable bounds. *Systems and Control Letters* 2000; **41**:49–56.
14. Nair GN, Evans RJ. Stabilizability of stochastic linear systems with finite feedback data rates. *SIAM Journal on Control and Optimization* 2004; **43**(2):413–436.
15. 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*, Atlantis, Paradise Island, Bahamas, 2004.
16. Tatikonda S. Control under communication constraints. *IEEE Transactions on Automatic Control* 2004; **49**(7):1056–1068.
17. Yuksel S, Basar T. State estimation and control for linear systems over communication networks. *Proceedings of 2003 IEEE Conference on Control Applications*, Istanbul, Turkey, 2003.
18. Ling Q, Lemmon MD. Stability of quantized control systems under dynamic bit assignment. *IEEE Transactions on Automatic Control* 2005; **50**(5):734–740.
19. Martins NC, Dahleh MA. Feedback control in the presence of noisy channels: ‘bode-like’ fundamental limitations of performance. *IEEE Transactions on Automatic Control* 2008; **52**(7):1604–1615.
20. 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).
21. Delevenne JC, Blondel V. Complexity of control of finite automata. *IEEE Transactions on Automatic Control* 2006; **51**(6):977–986.
22. Delevenne JC. An optimal quantized feedback strategy for scalar linear systems. *IEEE Transactions on Automatic Control* 2006; **51**(2):298–303.
23. Fagnani F, Zampieri S. Stability analysis and synthesis for linear systems with quantized feedback. *IEEE Transactions on Automatic Control* 2003; **48**(9).
24. Fagnani F, Zampieri S. Quantized stabilization of linear systems: complexity versus performance. *IEEE Transactions on Automatic Control* 2004; **49**(9).
25. Dumer I, Pinsker MS, Prelov VV. On coverings of ellipsoids in euclidean spaces. *IEEE Transactions on Information Theory* 2004; **50**(9).
26. Allen G, Gray RM. *Vector Quantization and Signal Compression*. Kluwer Academic Publishers: Dordrecht, 1992.
27. Gersho A, Gray RM. *Vector Quantization and Signal Compression*. Kluwer Academic Publishers: Boston, 1992.
28. Huang JJY, Schultheiss PM. Block quantization of correlated Gaussian random variables. *IEEE Transactions on Communication System* 1963; **11**(3):289–296.