On Time-Varying Bit-Allocation Maintaining Stability and Performance: A Convex Parameterization

Sridevi V. Sarma, *Member, IEEE*, Munther A. Dahleh, *Fellow, IEEE*, and Srinivasa Salapaka, *Member, IEEE*

***Abstract*—In this paper, we analyze and derive conditions for stability of a feedback system in which the plant and feedback controller are separated by a noiseless finite-rate communication channel. We allow for two deterministic classes of reference inputs to excite the system, and derive sufficient conditions for *input–output (IO) stability* as a function of the encoding strategy and controller. We first construct an encoder as a quantizer that can have *infinite memory* and can be *time-varying*, in that the strategy it follows to allocate a total of R bits to its inputs, is a function of time. This construction of the quantizer leads to the result that the set of allocation strategies that maintains stability for each class of reference signals is *convex*, allowing the search for the most efficient strategy to ensure stability to be formulated as a convex optimization problem. We then synthesize quantizers and time-varying controllers to minimize the rate required for stability and to track commands. Examples presented in this paper demonstrate how this framework enables computationally efficient methods for simultaneously designing quantizers and controllers for given plants. Furthermore, we observe that our finite memory quantizers that minimize the rate required for stability do not reduce to trivial memoryless bit-allocation strategies.**

***Index Terms*—Bit-allocation strategy, input–output (IO) stability, quantized control, transmission rate.**

I. Introduction

THE classical control paradigm addresses problems where communication between the plant and the controller is essentially perfect. Recently, problems in control over networked systems, whose components are connected via noisy communication links that may also induce delays and have finite rate constraints, are emerging. Applications include remote navigation systems (deep-space and sea exploration) and multirobot control systems (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.

Manuscript received March 1, 2006; revised June 12, 2007 and October 23, 2007. Published August 27, 2008 (projected). Recommended by Associate Editor E. Jonckheere.

S. V. Sarma and M. A. Dahleh are with the Massachusetts Institute of Technology, Cambridge, MA 02139-4307 USA (e-mail: sree@mit.edu; dahleh@mit.edu).

S. M. Salapaka is with the Mechanical Science and Engineering Department, University of Illinois, Urbana-Champaign, Urbana, IL 61801-2906 USA (e-mail: salapaka@uiuc.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2008.923660

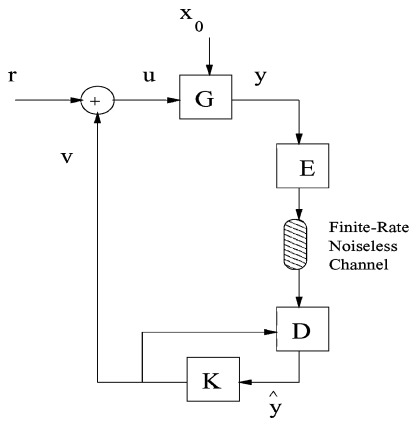


Fig. 1. Simple feedback network.

We consider the control system shown in Fig. 1. In this system, the output of the plant G is separated from the controller K by a finite-rate noiseless channel. We assume that the plant output is encoded by some operator E before entering the channel, and decoded by another operator D after exiting the channel. E and D work together to reduce the deleterious effects of the channel. The system shown in Fig. 1 has the following model:

$$\begin{aligned} x_{t+1} &= Ax_t + B(v_t + r_t) \qquad \forall t \geq 0 \\ y_t &= Cx_t \\ v_t &= K(\hat{y}_t) \end{aligned} \tag{1}$$

where $t \in \mathbf{Z}_+, x_t \in \mathbb{R}^n$, and $r_t, y_t, v_t \in \mathbb{R}$.

Previous works mainly focus on some notion of state stability under finite-rate (or countable) feedback control, where the only excitation to the system is an unknown bounded initial state condition ($r_t = 0$) [1], [3]–[5], [7]–[9], [11], [12], [14], [15], [19], [20], [22]–[24]. This research aims at finding conditions on the channel rate that will guarantee that the state of the system (or some function of the state) approach the origin or remain bounded as time goes to infinity. More recent works address synthesis of quantizers and/or performance limitations under finite-rate and finite-capacity feedback control [13], [23], [10], [16], [17].

In contrast, we allow for certain classes of inputs r_t to excite the system and study *input-output (IO)* stability under finite-rate feedback. Our approach introduces a computational methodology for analysis and synthesis of quantizers and controllers (E, D, and K) to meet various control objectives.

Another important distinction between our work and previous studies is that we only consider encoders that have memory but do not have access to the plant input and cannot compute this control signal. When encoders have access to the control signal or can compute the control signal exactly, the plant output signal that must be communicated down to the decoder can be described by a finite number of parameters. More concretely, $y_t = CA^t x_0 + C\sum_{i=0}^{t-1} A^{t-1-i}Bu_i$. Therefore, the only quantity unknown to E and D is x_0, which is just a vector of n numbers that belong to a bounded set in $\mathbb{R}^n$. Therefore, assuming G is observable, the encoder can compute x_0 after n time steps and start transmitting x_0 through the channel down to the decoder, at a rate of R bits per time step. Consider an encoder that allocates R_i bits to the ith component of x_0, such that $\sum_i R_i = R$, while the decoder continues to update its approximation of x_0 and x_t. The error vector then evolves as follows:

$$e_t = x_t - \hat{x}_t = A^t(x_0 - \hat{x}_{0,t})$$

where $\hat{x}_{0,t}$ is the estimate of x_0 at time t. If we assume that A is diagonal,[1] we get the following upper bound on the magnitude of each error component:

$$|e_t(i)| \leq L\left\{\frac{|\lambda_i(A)|}{2^{R_i}}\right\}^t, \qquad i = 1,2,\ldots,n$$

where $|x_0(i)| \leq L$, for $i = 1,2,\ldots,n$. It is easy to see that if $R_i > \max(0, \log(|\lambda_i(A)|)$, for $i = 1,\ldots,n$, which implies that $R > \sum_i \max(0, \log(|\lambda_i(A)|)$, then the system is asymptotically stable, because K is assumed to be stabilizing.

On the other hand, when the encoder does not have access to any signal in the loop except for the plant output, the unknown quantities characterizing y_t are x_0 and $u_0, u_1, u_2, \ldots$, which are an infinite number of parameters. From the encoder and decoder's perspective, the output of the plant is suddenly very "rich." For any fixed t, the decoder must approximate $y_0, y_1, \ldots, y_t$. These approximations cannot converge to their actual values, and the best strategy that E and D can employ is to improve the approximations over time by allowing E to allocate more and more bits to them (this motivates our construction of the quantizer presented in Section II-A). The system boils down to a quantized feedback system, which is difficult to analyze, and where the usual tradeoff of delay versus accuracy holds.

We conclude this introduction with an overview of this paper. In Section II, we describe our setup, introduce a new parameterized class of encoders that have memory but no access to the control signal, and formulate a stability problem. In Section III, we derive sufficient conditions for finite-gain stability for bounded and decaying classes of reference inputs. In Section IV, we show that the set of stable bit-allocation strategies implemented by our parameterized quantizers is convex, enabling synthesis of such quantizers that achieve desired objectives. In Sections V and VI, we design finite-memory bit-allocation strategies and controllers that minimize the rate required for stability and that track commands for different plants. Finally, we conclude in Section VII.

[1]All results hold for general A matrices as shown in [22].

II. Problem Formulation

We study system (1), in which the channel encoder has memory and no access to the control input. E is a limited-rate quantizer that has *infinite memory* and is *time-varying* in that the strategy it follows in allocating a total of R bits to all of the inputs up to time t is a function of t (see Section II-A for details). We assume that the channel can transmit R bits instantaneously with each use. The channel decoder D computes updates on the current and past values of y and sends these to the controller. K is a causal linear time-varying system described in Section II-B.

We define the closed-loop system to be IO stable if for all $r \in \mathcal{C}_r$, there exists a finite positive constant α and a finite constant β such that $\|y\|_\infty \leq \alpha\|r\|_\infty + \beta$. Here, we investigate IO stability with respect to the following classes of reference inputs $\mathcal{C}_r$.

1) **Bounded Signals**: $\mathcal{C}_r = l_{\infty,\overline{r}}$, where $l_{\infty,\overline{r}}$ is the class of all signals that are bounded in magnitude by $\overline{r}$.
2) **Decaying Signals**: $\mathcal{C}_r = \mathcal{C}_{\gamma,\overline{r}}$, where $\mathcal{C}_{\gamma,\overline{r}}$ is the class of all signals that are bounded by the positive decaying function $\overline{r}\gamma^k$ for all $k \geq 0$ and $0 < \gamma < 1$, i.e., if $r \in \mathcal{C}_{\gamma,\overline{r}}$, then $|r_k| \leq \overline{r}\gamma^k$ for all $k \geq 0$.

A. Limited-Rate Time-Varying $(\mathcal{R}, M)$-Quantizers

Before stating the problems that we are interested in solving, we first define and model the parameterized class of time-varying infinite-memory $(\mathcal{R}, M)$-quantizers.

We view the quantizer as a module that approximates its input, which, in general, requires an infinite number of bits, with a finite number of bits. Formally speaking, an $(\mathcal{R}, M)$-quantizer with bit rate R, is a sequence of causal time-varying operators, parameterized by an infinite-dimensional *rate matrix* $\mathcal{R}$ of the form

$$\begin{bmatrix} R_{00} & 0 & 0 & \cdots & \cdots \\ R_{01} & R_{11} & 0 & 0 & \cdots \\ R_{02} & R_{12} & R_{22} & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{bmatrix}$$

where $\sum_j R_{ij} = R - 1$ for all i, and an infinite-dimensional positive-definite diagonal *scale matrix* $M = \mathrm{diag}(M_{00}, M_{11}, M_{22}, \ldots,)$. The $(\mathcal{R}, M)$-quantizer saturates to output M_{kk}, the $(k+1)$th diagonal of M, when its input y_k has magnitude greater than or equal to M_{kk}, i.e., when $|y_k| \geq M_{kk}$. However, we denote the quantizer "valid" only when $|y_k| \leq M_{kk}$ for all $k \geq 0$, and define what the quantizer does in this case below.

Let $\hat{y}_i(j)$ be the quantized estimate of y_i at time j. Then, $\mathcal{R}$ determines that at time $t = 0$, 1 bit is used to denote the sign of y_0, and R_{00} bits are used to quantize the magnitude of y_0 to produce $\hat{y}_0(0)$. At time $t = 1$, an *additional* R_{01} bits are used to quantize the magnitude of y_0 to produce $\hat{y}_0(1)$; 1 bit is used to denote the sign of y_1, and R_{11} bits are used to quantize the magnitude of y_1 to produce $\hat{y}_1(1)$, and so on. The accuracy of $\hat{y}_i(j)$ is within $\pm M_{ii}2^{-(\sum_{k=i}^{j} R_{ik}+1)}$ of y_i for all $i \geq 0$.

For the sake of analysis, the quantizer, the channel, and the decoder can be broken down into the five steps at any instant t as shown in Fig. 2 when $|y_k| \leq M_{kk}$ for $0 \leq k \leq t$. First,

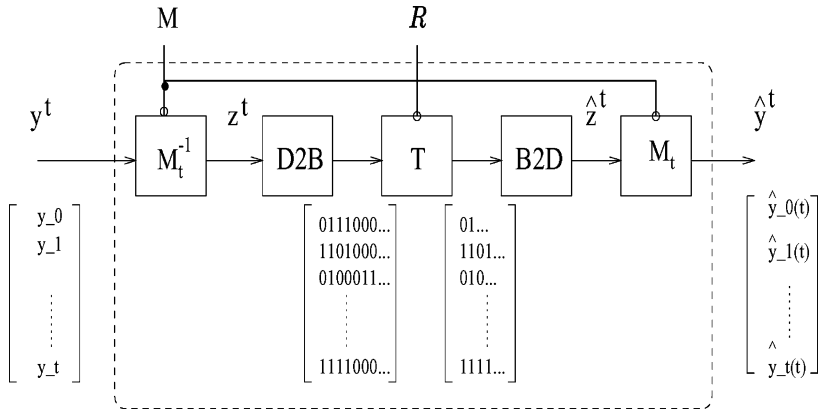


Fig. 2. Quantizer–channel–decoder operator at time t.

the ith component of the vector y^t is scaled by $(1)/(M_{ii})$ for $i = 1, 2 \ldots, t$, to produce z^t, where $y^t = [y_0\ y_1\ \cdots\ y_t]'$, and $M_t = \mathrm{diag}(M_{00}, M_{11}, \ldots, M_{tt})$. The scaling by M_t^{-1} ensures that $\|z^t\|_\infty \leq 1$. Then, each element of z^t is converted into its binary representation, i.e., a string of "0s" and "1s" in the decimal-to-binary (D2B) converter. Next, each binary string is truncated according to the bit-allocation strategy induced by $\mathcal{R}$. Specifically, the binary string representing $z^t(i) = (y_i)/(M_{ii})$ is truncated to contain only its first $(\sum_{j=i}^{t} R_{ij})$ bits. Note that this truncation induces an error of at most $2^{-(\sum_{j=i}^{t} R_{ij}+1)}$ in magnitude for $z^t(i)$, i.e., $|z^t(i) - \hat{z}^t(i)| \leq 2^{-(\sum_{j=i}^{t} R_{ij}+1)}$.

As shown in Fig. 2, the truncated binary string is converted back into its decimal representation, via the binary-to-decimal module (B2D), to produce $\hat{z}^t$. Finally, $\hat{z}^t$ is scaled by M_t to produce $\hat{y}^t$, where $\hat{y}^t = [\hat{y}_0(t)\ \hat{y}_1(t)\ \hat{y}_2(t)\ \cdots\ \hat{y}_t(t)]'$. An upper bound on the error between each input component and its approximate output is

$$|y_k - \hat{y}_k(t)| \leq M_{kk} 2^{-(\sum_{i=k}^{t} R_{ki}+1)}$$

$\forall k \leq t$. Stated differently, if $|y_k| \leq M_{kk}$ for $0 \leq k \leq t$, then there exists a $w_k(t)$ with $\mathrm{sign}(w_k(t)) = \mathrm{sign}(y_k)$ and $|w_k(t)| \leq 1\ \forall t \geq 0$, such that

$$\hat{y}_k(t) = y_k + M_{kk} 2^{-(\sum_{i=k}^{t} R_{ki}+1)} w_k(t)$$

for all $k \leq t$. For analysis, we augment the output of the quantizer at time t to be the vector of *all* estimates of $\mathbf{y}^t$ from time 0 to time t. We denote the augmented vector as $\hat{\mathbf{y}}_a^t$ as shown below

$$\hat{\mathbf{y}}_a^t = \begin{bmatrix} \underline{\hat{y}_0(0)} \\ \hat{y}_0(1) \\ \underline{\hat{y}_1(1)} \\ \vdots \\ \hline \hat{y}_0(t) \\ \hat{y}_1(t) \\ \hat{y}_2(t) \\ \vdots \\ \hat{y}_t(t) \end{bmatrix}.$$

We can then model the quantizer in its "valid" region as the following sequence of time-varying operators:

$$Q(\mathcal{R}, M) = \Big\{ Q_t : \mathbb{R}^{t+1} \to \mathbb{R}^{\frac{(t+1)(t+2)}{2}} | Q_t(\mathbf{y}^t) = \hat{\mathbf{y}}_a^t = I_t \mathbf{y}^t + F_t(\mathcal{R}) \overline{M}_t(M) \mathbf{w}_a^t,\ t \geq 0 \Big\}$$

where

$$I_t = \begin{bmatrix} \underline{1} \\ 1 \quad 0 \\ \underline{0 \quad 1} \\ \vdots \\ \hline I^{t \times t} \end{bmatrix}$$

and

$$F_t(\mathcal{R}) = \begin{bmatrix} f_{00} & & & & \\ & f_{01} & & & \\ & & f_{11} & & \\ & & & \ddots & \\ & & & & f_{0t} \\ & & & & & \ddots \\ & & & & & & f_{tt} \end{bmatrix}$$

$$\overline{M}_t(M) = \begin{bmatrix} M_{00} & & & & \\ & M_{00} & & & \\ & & M_{11} & & \\ & & & \ddots & \\ & & & & M_{00} \\ & & & & & \ddots \\ & & & & & & M_{tt} \end{bmatrix}$$

with $f_{ks} = 2^{-(\sum_{i=k}^{s} R_{ki}+1)}$ for $s = 0, 1, \ldots, t$, and $k = 0, 1, \ldots, s$.

Also

$$\mathbf{w}_a^t = \begin{bmatrix} \underline{w_0(0)} \\ \underline{w_0(1)} \\ \underline{w_1(1)} \\ \vdots \\ \overline{w_0(t)} \\ w_1(t) \\ w_2(t) \\ \vdots \\ w_t(t) \end{bmatrix}$$

where $\mathbf{w}_a \in \mathbf{l}_\infty$ such that $\|\mathbf{w}_a\|_\infty \leq 1$.

B. Plant and Controller

We represent the linear time-invariant (LTI) causal system G and the causal linear time-varying controller K as the following matrix multiplication operators at any time instance t:

$$G_t = \begin{bmatrix} g_0 & & & & \\ g_1 & g_0 & & & \\ g_2 & g_1 & g_0 & & \\ \vdots & & \ddots & \ddots & \\ g_t & \cdots & \cdots & g_1 & g_0 \end{bmatrix}$$

$$K_t = \begin{bmatrix} k_0 & & & & & & \\ & k_1 & k_0 & & & & \\ & & & k_2 & k_1 & k_0 & \\ \vdots & & & & & & \ddots \end{bmatrix}.$$

Note that, with matrix I_t as defined above, K_t satisfies $K_t I_t = K_t^{\mathrm{LTI}}$ where K_t^{LTI} represents a linear time invariant system given by

$$K_t^{\mathrm{LTI}} = \begin{bmatrix} k_0 & & & & \\ k_1 & k_0 & & & \\ k_2 & k_1 & k_0 & & \\ \vdots & & \ddots & \ddots & \\ k_t & \cdots & \cdots & k_1 & k_0 \end{bmatrix}.$$

Fig. 3 illustrates the closed-loop system at time t when the quantizer is modeled as an endogenous disturbance as described in Section II-A, with

$$\mathbf{r}^t = \begin{bmatrix} r(0) \\ r(1) \\ \vdots \\ r(t) \end{bmatrix} \qquad \mathbf{u}^t = \begin{bmatrix} u(0) \\ u(1) \\ \vdots \\ u(t) \end{bmatrix}.$$

From here onwards, we refer to $F_t(\mathcal{R})$ as F_t and $\overline{M}_t(M) = \overline{M}_t$ for an easier read.

C. Problem Statement

We are interested in solving the following problems.

1) Given $G, K, \mathcal{C}_r$, and a rate matrix $\mathcal{R}$, determine whether there exists a set of scale matrices M that maintain IO stability and quantizer validity.

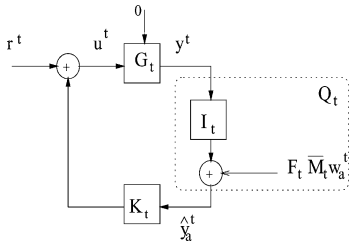


Fig. 3. Control system at time t.

2) Given G, K, and $\mathcal{C}_r$, characterize the set of all rate matrices $\mathcal{R}$, such that the system is IO-stable and the quantizer is valid.
3) Within the set of stabilizing rate matrices, find the minimum transmission rate R of the channel.
4) For a given G and $\mathcal{C}_r$, design K and $\mathcal{R}$ to minimize the rate required for stability and to track commands in $\mathcal{C}_r$.

III. Stability Analysis

A. Bounded Signals

In this section, we derive sufficient conditions for IO stability when $\mathcal{C}_r = l_{\infty,\overline{r}}$. Let $T_t \triangleq (I - G_t K_t^{\mathrm{LTI}})^{-1} G_t$, then it is straightforward to show that $y^t = T_t r^t + T_t K_t F_t \overline{M}_t w_a^t$. The following theorem then gives sufficient conditions for IO stability and quantizer validity.[2] Note that for a matrix A, $\|A\|_1 = \max_i \sum_j |a_{ij}|$.

Theorem 3.1: Consider system (1) with $x_0 = 0$. Let $E = Q(\mathcal{R}, M)$, for a given rate matrix $\mathcal{R}$, and let $r \in l_{\infty,\overline{r}}$. If

1) $\|T\|_1 < \infty$;
2) $\|TKF\|_1 < 1$;

then there exists a constant scale matrix $M = mI$, such that

- (IO stability) $\|y\|_\infty \leq \|T\|_1 \|r\|_\infty + m\|TKF\|_1$;
- (quantizer validity) $\|y\|_\infty \leq m$.

Proof: Choose $m \geq (\|T\|_1 \overline{r})/(1 - \|TKF\|_1) \geq 0$, which is possible given the norm bounds on r, T, and TKF. Then

$$\begin{aligned} \|y^t\|_\infty &= \sup_t \{\|T_t r^t + T_t K_t F_t \overline{M}_t w^t\|_\infty\} \\ &\leq \sup_t \|T_t\|_1 \|r^t\|_\infty + \sup_t \|T_t K_t F_t \overline{M}_t\|_1 \\ &\leq \|T\|_1 \|r\|_\infty + m\|TKF\|_1 \\ &\leq m. \end{aligned}$$

The last inequality comes from our choice of m. ■

The stability condition in Theorem 3.1 is *sufficient* as we have not yet proven that $|y_k| > m$, for any $k \geq 0$, renders the system unstable. We note that memoryless, time-invariant quantizers are represented by an identity rate matrix multiplied by the value of the fixed rate $R - 1$, which leads to the following corollary.

Corollary 3.1: Consider system (1) with $x_0 = 0$. Let $E = Q(\mathcal{R}, M)$, for a *diagonal* rate matrix $\mathcal{R} = (R-1)I$, and let $r \in l_{\infty,\overline{r}}$. If

1) $\|T\|_1 < \infty$;
2) $\|TK^{\mathrm{LTI}}\|_1 < 2^R$;

[2]One can add an exogenous input d at the input of the controller and derive sufficient conditions for external stability by computing transfer functions from r and d to y and v (output of the controller). We omit the details here as the analysis is straightforward.

then there exists a constant scale matrix $M = mI$ such that

- (IO stability) $\|y\|_\infty \le \|T\|_1\|r\|_\infty + m2^{-R}\|TK^{\mathrm{LTI}}\|_1$;
- (quantizer validity) $\|y\|_\infty \le m$.

B. Decaying Signals

We now consider the case where $r \in \mathcal{C}_{\gamma,\overline{r}}$, and derive sufficient conditions for IO stability. We show that if G and K^{LTI} are finite-dimensional systems, the conditions can guarantee that the output signal y_t decays exponentially over time. In this section, both G and K^{LTI} are assumed to be finite-dimensional systems.

We know that with perfect feedback and a stabilizing K^{LTI}, exponentially decaying reference signals generate system outputs that exponentially decay over time. We would like to generate the same types of decaying responses with finite-rate feedback, and thus consider the quantizer scales M_{tt} to be a decaying function of t, i.e., $M_t = \mathrm{diag}(m, m\beta, m\beta^2, \ldots, m\beta^t)$, where $m < \infty$ and $0 \le \beta < 1$. The matrix $(F\overline{M})_t$ will have kth diagonal equal to $\beta^k 2^{-\sum_{i=k}^{t} R_{ki}}$ for $0 \le k \le t$, and the quantizer is valid only if $|y_t| \le m\beta^t$ for all $t \ge 0$. Below we state a theorem that states sufficient conditions for IO stability. The details that lead to the derivation of the theorem and proof are shown in part A of the Appendix.

Theorem 3.2: Assume that T and TK^{LTI} are finite-dimensional *stable* LTI systems, whose corresponding A matrices (of state-space descriptions) have spectral radii ρ and ν, respectively (both ρ and ν have magnitudes less than 1). Denote $\{T\}_i$ and $\{TK^{\mathrm{LTI}}\}_i$ as the ith components $(i = 0, 1, \ldots)$ of the impulse responses of T and TK, respectively. Given system (1), $r \in \mathcal{C}_{\gamma,\overline{r}}$, and a rate matrix $\mathcal{R}$, if

1) $\{T\}_i \le \eta_1\rho^i$ $(0 \le \eta_1 < \infty)$ for all $i \ge 0$;
2) $\{TK^{\mathrm{LTI}}\}_i \le \eta_2\nu^i$ $(0 \le \eta_2 < (1)/(\|F\|_1))$ for all $i \ge 0$;
3) $\max(\nu, \rho, \gamma)(1 - \eta_2\|F\|_1) > \nu$;

then there exists a decaying scale matrix $M(M_{kk} = m\beta^k)$ such that

- (IO stability) $\|y\|_\infty \le \|T\|_1\|r\|_\infty + m\|TK\|_1$;
- (quantizer validity) $|y_t| \le m\beta^t \ \forall t \ge 0$.

IV. Characterization of Stable Rate Matrices

We have shown that if $\|TKF\|_1 < \eta$ for $\eta > 0$, then the system is IO-stable for bounded and decaying inputs. This inequality can be written as a set *convex* constraints on the rate matrix parameters. The following theorem shows this result.

Theorem 4.1: Let $X = \{\mathrm{vec}(\mathcal{R})\} = [R_{00}\ R_{01}\ R_{02}\ \cdots\ R_{11}\ R_{12}\ \cdots]'$,[3] then for any infinite-dimensional matrix P, the condition $\|PF(X)\|_1 < \eta$ is convex in X for any $\eta > 0$.

Proof: $\|PF(X)\|_1 > \eta \Leftrightarrow \sum_{j=0}^{\infty} f_j(X)|P_{ij}| < \eta$, $i = 0, 1, \ldots$, where $f_j(X) = 2^{-\sum_{i=j}^{\infty} R_{ji}}$. We now show that $f_j(X)$ is convex in X, and thus, any nonnegative combination of $f_j(X)$ is convex. First, we recall that 2^{-a} is a convex function in a. Let $2^{-(\sum_{i=j}^{\infty} R_{ji}+1)} = 2^{-(c_j'X+1)}$, where c_j' is an appropriate row vector for $j = 0, 1, \ldots$. Note that

$$\begin{aligned} &2^{-\lambda(c_j'X_1+1)-(1-\lambda)(c_j'X_2+1)} \\ &\quad = 2^{-\lambda(a_1+1)-(1-\lambda)(a_2+1)} \\ &\quad \le \lambda 2^{-(a_1+1)} + (1-\lambda)2^{-(a_2+1)} \\ &\quad = \lambda 2^{-(c_j'X_1+1)} \\ &\quad + (1-\lambda)2^{-(c_j'X_2+1)}. \end{aligned}$$

■

If we let $P = TK$, then we get that the stability condition $\|TKF\|_1 < \eta$ is a set of convex constraints on the infinite-dimensional vector $\mathrm{vec}(\mathcal{R})$. This result enables the search for the most efficient quantizer to be formulated as a convex optimization problem.

V. Synthesis: Minimizing Channel Rate

In this section, we synthesize (R, M)-quantizers and time-varying controllers for different plants to minimize the rate required for IO stability for $\mathcal{C}_r = l_{\infty,\overline{r}}$. In particular, we set out to solve the following problem:

$$\min_{K^{\mathrm{LTI}},\mathcal{R}} R \tag{2}$$

$$\text{s.t. } \|T\|_1 < \infty \tag{3}$$

$$\|TKF(\mathcal{R})\|_1 < 1 \tag{4}$$

$$\sum_j R_{ij} = R - 1, \qquad i = 0, 1, \ldots$$

$$R_{ij} \ge 0, \qquad j \le i = 0, 1, \ldots$$

where $T = (I - GK^{\mathrm{LTI}})^{-1}G$ and (3) and (4) are stability conditions. We make a few comments regarding our approach to solve (2).

1) The $(\mathcal{R}, M)$-quantizer is described by an infinite number of parameters (R_{ij}) as it has infinite memory. To make things easily computable and more practical, we restrict ourselves to finite-memory $(\mathcal{R}, M)$-quantizers, defined in Section V-A, each of which is described by a finite number of parameters.
2) The optimization problem (2) has constraints that are nonconvex in both K^{LTI} and $\mathcal{R}$, and therefore, is not efficiently solvable. We propose an iterative algorithm, described in Section V-B, that alternates between computing a controller and quantizer. In each computation, we solve a convex optimization problem subject to convex constraints. We show that our iterative algorithm has a nonincreasing cost, and therefore, converges to a local minimum. This iteration is reminiscent of the D-K iteration used to compute μ when modeling systems with structured uncertainty [2], [21]. Numerical examples are given in Section V-C.

A. Finite-Memory $(\mathcal{R}, M)$-Quantizers

We consider a special class of practical quantizers that have finite memory and are periodic. Specifically, each value of y gets approximated by the quantizer for *at most* N consecutive time steps. In fact, for any $t \ge 0$, y_{tN+j} gets approximated

[3]The "vec" operator on a matrix simply concatenates all the column to form one large column vector.

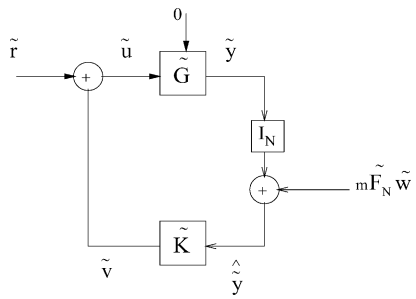
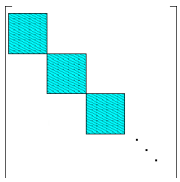


Fig. 4. Lifted closed-loop system.

for $N-j$ time steps, for $j=0,1,\ldots,N-1$. Moreover, the bit-allocation strategy repeats every N time steps. We call this class of quantizers, "repeated-block" (RB) quantizers because the structure of the rate matrix is block diagonal as shown in the following:



Each block is the following $N\times N$ matrix:

$$R_{\text{block}}=\begin{bmatrix} R_{00} & & & \\ R_{01} & R_{11} & & \\ \vdots & \vdots & \ddots & \\ R_{0,N-1} & R_{1,N-1} & \cdots & R_{N-1,N-1}\end{bmatrix}.$$

It is useful to highlight that RB quantizers are time-invariant operators in "lifted" coordinates, where each time step in the lifted coordinates is equivalent to N time steps in original coordinates. We define the following lifted signals for $t\geq 0$:

$$\tilde{\mathbf{r}}_t=\begin{bmatrix} r_{tN}\\ r_{tN+1}\\ \vdots\\ r_{(t+1)N-1}\end{bmatrix}$$

$$\tilde{\mathbf{y}}_t=\begin{bmatrix} y_{tN}\\ y_{tN+1}\\ \vdots\\ y_{(t+1)N-1}\end{bmatrix}$$

$$\tilde{\mathbf{w}}_t=\begin{bmatrix} w_{tN}((t+1)N-1)\\ w_{tN+1}((t+1)N-1)\\ \vdots\\ w_{(t+1)N-1}((t+1)N-1)\end{bmatrix}.$$

The model for a repeated-block quantizer Q_{RB} in the lifted coordinates, denoted $\tilde{Q}_{\text{RB}}$, is

$$Q_{\text{RB}}(\mathcal{R},M)=\Big\{\tilde{Q}_{\text{RB}}:\mathbb{R}^N\to\mathbb{R}^N \mid \tilde{Q}_{\text{RB}}(\tilde{\mathbf{y}})=\hat{\tilde{\mathbf{y}}}=I_N\tilde{\mathbf{y}}+m\tilde{F}_N\tilde{\mathbf{w}}\Big\}$$

where, written as a matrix multiplication operator, $\tilde{F}_N=\text{diag}(F_{N-1},F_{N-1},\ldots)$ and $I_N=I_t$ evaluated at $t=N$. The closed-loop system in lifted coordinates is shown in Fig. 4.

After lifting by a factor of N, the plant $\tilde{G}$ is LTI with N inputs and N outputs and the controller $\tilde{K}$ becomes LTI with $(N(N+1))/(2)$ inputs and N outputs. This enables us to use existing control tools for LTI systems to synthesize $\tilde{K}$.

We state sufficient conditions for IO stability for bounded signals in the lifted coordinate space with arbitrary linear time-varying controllers, but first state the following Lemma whose proof is straightforward and left to the reader.

Lemma 5.1: Let P be any causal LTI single-input–single-output (SISO) system, and $\tilde{P}$ is a lifted version of P with lift factor N. Then, for any $N\geq 1, \|\tilde{P}\|_1\geq\|P\|_1$.

Theorem 5.1: Consider system (1) with an arbitrary time-varying controller lifted by a factor of N, with $x_0=0$. Let $E=Q_{\text{RB}}(\mathcal{R},M)$, for a given repeated-block rate matrix $\mathcal{R}$, and let $r\in l_{\infty,\overline{r}}$. If

1) $\|(I-\tilde{G}\tilde{K}^{\text{LTI}})^{-1}\tilde{G}\|_1<\infty$;
2) $\|(I-\tilde{G}\tilde{K}^{\text{LTI}})^{-1}\tilde{G}\tilde{K}\tilde{F}_N\|_1<1$;

then the original system is IO-stable.

Proof: If conditions (1) and (2) hold, then by invoking Lemma 5.1, we get that $\|(I-GK^{\text{LTI}})^{-1}G\|_1<\infty$, and $\|(I-GK^{\text{LTI}})^{-1}GKF\|_1<1$. From Theorem 3.1, we then get that the original system (unlifted) is IO stable. ■

B. $\mathcal{R}-K$ Iteration

The $\mathcal{R}-K$ iteration algorithm proposed here efficiently computes a locally optimal controller and finite-memory RB quantizer that solves (2). The algorithm is described as follows. Note that $\tilde{T}=(I-\tilde{G}\tilde{K}_0^{\text{LTI}})^{-1}\tilde{G}$.

$\mathcal{R}-K$ Iteration Algorithm

1) For a given plant G, pick any stabilizing controller K_0^{LTI}. Pick a convergence threshold $\epsilon>0$ and set $c=0$.
2) Solve the following problem to construct $\mathcal{R}_c$:

$$\begin{aligned}&\min_{\mathcal{R}}\ R\\ &\|\tilde{T}\tilde{K}_c\tilde{F}_N(\mathcal{R})\|_1>1,\\ &\text{s.t. } R=1+\sum_j R_{ij},\qquad i=0,1,\ldots,N-1\\ &R_{ij}\geq 0,\qquad j\leq i=0,1,\ldots,N-1.\end{aligned}\tag{5}$$

Note that the above problem has constraints that are convex in all unknown parameters R_{ij} as shown in Section IV and is, therefore, efficiently computable using any standard convex optimization software.
If $c>=1$ and if $R_{c-1}-R_c\leq\epsilon$ then STOP. Else goto Step 3.
3) Fix $\mathcal{R}_c$ and compute new controller K_{c+1} by solving

$$\begin{aligned}&\min_{K^{\text{LTI}}}\quad \|\tilde{T}\tilde{K}\tilde{F}_N(\mathcal{R}_c)\|_1\\ &\text{s.t.}\quad K^{\text{LTI}}\text{ stabilizing}\end{aligned}$$

Apply "Q-parameterization" or "Youla parameterization" to parameterize all stabilizing

controllers and convert the above problem to a convex optimization problem of the form

$$\min_{Q} \quad \|H_0 - H_1 Q H_2\|_1$$
$$\text{s.t.} \quad Q \text{ LTI stable rational and proper} \tag{6}$$

where H_0, H_1, and H_2 are derived from the system to equate $\|\tilde{T}\tilde{K}\tilde{F}_N(\mathcal{R}_0)\|_1$ to $\|H_0 - H_1 Q H_2\|_1$. Note that there is a one-to-one relationship with $\tilde{K}_{c+1}$ and Q given by the parameterization. We omit details as this is a common procedure and refer the reader to [6] and [18] for details. Note that (6) is again efficiently computable using any standard convex optimization software.

4) Increment c by 1 and goto Step 2.

We demonstrate how we use MATLAB's "fmincon.m" function to solve (5) in part B of the Appendix.

The above algorithm continues to iterate between the two optimization problems until both costs fail to change more than a given ϵ. This iteration will converge to a local minimum rate required for stability as described in Proposition 5.1.

Proposition 5.1: The $\mathcal{R} - K$ iteration algorithm has a nonincreasing cost in the channel rate R, and therefore, it converges to a local minimum.

Proof: For a given plant and controller, suppose we have just computed rate matrix $\mathcal{R}_c$ with rate R_c for some positive integer c. Given $\mathcal{R}_c$, the iteration algorithm then computes a controller K_{c+1} that minimizes the bias $\|\tilde{T}\tilde{K}\tilde{F}_N(\mathcal{R}_c)\|_1$ by solving (6) and converting Q back to $\tilde{K}$. Then, we get that $\|\tilde{T}\tilde{K}_{c+1}\tilde{F}_N(\mathcal{R}_c)\|_1 \leq \|\tilde{T}\tilde{K}\tilde{F}_N(\mathcal{R}_c)\|_1$ for all controllers $\tilde{K}$. Next, we compute a rate matrix $\mathcal{R}_{c+1}$ that minimizes the channel rate required for stability subject to the constraint that $\|\tilde{T}\tilde{K}_{c+1}\tilde{F}_N(\mathcal{R})\|_1 < 1$. Note that $\mathcal{R}_c$ is a feasible solution, therefore $R_{c+1} \leq R_c$. ∎

C. Examples

We now execute one iteration of the $\mathcal{R} - K$ iteration algorithm for three different unstable plants with $\overline{r} = 1$ and $N = 4$ (memory size of quantizer). We summarize the results for each plant in separate boxes. In each case, we give the initial controller K_0 used to compute repeated-block quantizer $\mathcal{R}_0$, which is entirely characterized by $R_{\text{block},0}$. We also compare the rate R_0 to the minimum rate R_{ml} obtained if we restrict the quantizer to be memoryless and time-invariant (a diagonal rate matrix). Note that the minimum channel rate R_0 required for stability for each corresponding closed-loop system is $R_{\text{block},0}(1,1)+1$. We then show the controller K_1 computed by fixing the quantizer to be parameterized by $\mathcal{R}_0$ and the corresponding quantizer bias term $\|\tilde{T}\tilde{K}_0\tilde{F}_N(\mathcal{R})\|_1$. Finally, we fix the controller to K_1 and compute $\mathcal{R}_1$ to complete one iteration.

Example 1

- **Step 1:**

$$G = \frac{10}{z+1} \quad K_0 = \frac{-0.075(z^2+z+1)}{z^3} \quad \epsilon = 1.5$$

- **Step 2:**

$$R_{\text{block},0} = \begin{bmatrix} 1.4171 & & & \\ 0.2621 & 1.1551 & & \\ & 0.4907 & 0.9265 & \\ & & 0.2665 & 1.1507 \end{bmatrix}$$
$$R_0 = 2.4171 \quad R_{ml} = 2.6792$$

- **Step 3:**

$$K_1 = \frac{-0.04375z^2 - 0.01074z + 0.001745}{z^3}$$
$$\|\tilde{T}\tilde{K}_0\tilde{F}_N(\mathcal{R}_0)\|_1 = 1$$
$$\|\tilde{T}\tilde{K}_1\tilde{F}_N(\mathcal{R}_0)\|_1 = 0.2252$$

- **Step 2:**

$$R_{\text{block},1} = \mathbf{0} \quad R_1 = 1 \quad (R_0 - R_1) = 1.4171 < \epsilon$$

Example 2

- **Step 1:**

$$G = \frac{10}{z+1.5} \quad K_0 = \frac{-0.072z^2 - 0.0216z - 0.00648}{z^3} \quad \epsilon = 1$$

- **Step 2:**

$$R_{\text{block},0} = \begin{bmatrix} 2.1630 & & & \\ 0.1830 & 1.9800 & & \\ & 0.3580 & 1.8050 & \\ & & 0.1847 & 1.9783 \end{bmatrix}$$
$$R_0 = 3.163 \quad R_{ml} = 3.3461$$

- **Step 3:**

$$K_1 = \frac{-0.1173z^2 - 0.05309z - 0.01148}{z^3}$$
$$\|\tilde{T}\tilde{K}_0\tilde{F}_N(\mathcal{R}_0)\|_1 = 1$$
$$\|\tilde{T}\tilde{K}_1\tilde{F}_N(\mathcal{R}_0)\|_1 = 0.4731$$

- **Step 2:**

$$R_{\text{block},1} = \begin{bmatrix} 1.2131 & & & \\ 0.2875 & 0.9255 & & \\ & 0.5750 & 0.6380 & \\ & & 0.2876 & 0.9255 \end{bmatrix}$$
$$R_1 = 2.213 \quad (R_0 - R_1) = 0.95 < \epsilon$$

Example 3

- **Step 1:**

$$G = \frac{10}{z+2} \quad K_0 = \frac{-0.144z^2 - 0.0576z - 0.02304}{z^3} \quad \epsilon = 1.1$$

- **Step 2:**

$$R_{\text{block},0} = \begin{bmatrix} 4.4267 & & & \\ 0.2284 & 4.1982 & & \\ & 0.4329 & 3.9938 & \\ & & 0.2271 & 4.1996 \end{bmatrix}$$

$$R_0 = 5.4267 \quad R_{ml} = 5.6551$$

- **Step 3:**

$$K_1 = \frac{-0.1841z^2 - 0.08769z - 0.01779}{z^3}$$

$$\|\tilde{T}\tilde{K}_0\tilde{F}_N(\mathcal{R}_0)\|_1 = 1$$

$$\|\tilde{T}\tilde{K}_1\tilde{F}_N(\mathcal{R}_0)\|_1 = 0.2915$$

- **Step 2:**

$$R_{\text{block},1} = \begin{bmatrix} 3.4036 & & & \\ 0.2692 & 3.1344 & & \\ & 0.5298 & 2.8738 & \\ & & 0.2683 & 3.1353 \end{bmatrix}$$

$$R_1 = 4.4036 \quad (R_0 - R_1) = 1.0231 < \epsilon$$

From Examples 1–3, we see that $R_0 < R_{ml}$ in all cases indicating that allowing the quantizer to have memory and to allocate bits to the past maintains stability for channels with smaller rates than in the case where the quantizer is memoryless and does not allocate to the past. In addition, the more unstable the plant is, the more rate is required for closed-loop stability. Finally, one iteration shows a marked improvement in rate required for stability as $R_1 < R_0$ in all cases.[4]

VI. Synthesis: Tracking Commands

In this section, we synthesize limited-rate finite-memory (R, M)-quantizers and time-varying controllers for a given plant to track a family of bounded reference commands $\mathcal{C}_r = l_{\infty,\overline{r}}$ such that the rate is limited by a given $\overline{R}$. From Fig. 3, we get that the tracking error is $y - r = (T - I)r + TKF_N(\mathcal{R})\overline{M}w_a$. Therefore, $\|y - r\|_\infty \leq \|T - I\|_1 + m\|TKF_N(\mathcal{R})\|_1$, where m satisfying $\|T\|_1 + m\|TKF_N(\mathcal{R})\|_1 \leq m$ ensures quantizer validity. We set up the following optimization problem:

$$\min_{K^{\text{LTI}},\mathcal{R}} \quad \|T - I\|_1 + m\|TKF_N(\mathcal{R})\|_1 \tag{7}$$

$$\text{s.t.} \quad \|T\|_1 < \infty \tag{8}$$

$$\|T\|_1 + m\|TKF_N(\mathcal{R})\|_1 \leq m \tag{9}$$

$$\sum_j R_{ij} = R - 1, \qquad i = 0, 1, \ldots$$

$$R_{ij} \geq 0, \qquad j \leq i = 0, 1, \ldots$$

$$R \leq \overline{R}$$

where (8) and (9) are the IO stability conditions. Again, we restrict ourselves to repeated-block quantizers and apply an iteration algorithm to minimize tracking cost. We outline the iteration algorithm below.

[4]Note that third-order finite impulse response (FIR) controllers were designed to simplify computation and are shown in Examples 1–3. Searching over such FIR controllers is not optimal, yet we are still able to demonstrate the validity of our synthesis algorithms. The same was done for the tracking examples.

We fix m in Step 2 of the tracking iteration algorithm and check to make sure that the quantizer remains valid for that chosen value of m in both (10) and (12). Keeping m as a function of $\mathcal{R}$ and K would not enable convexity of the subproblems (10) and (12). One can find the smallest m over all $\mathcal{R}$ and K such that the quantizer remains valid and minimizes the tracking cost by applying the above iteration algorithm to different values of m, changing m via a bisection algorithm.

It is straightforward to see that the tracking iteration algorithm described above always results in a nonincreasing tracking cost as both subproblems (10) and (12) minimize the same cost function.

$\mathcal{R} - K$ Iteration Tracking Algorithm

1) For a given plant G and rate limit $\overline{R}$, pick any stabilizing controller K_0, such that $\|\tilde{T}\tilde{K}_0\tilde{F}(\mathcal{R})\|_1 < 1$. If no such controller can be found, increase $\overline{R}$ and return to Step 1.
2) Pick an $m \geq (\|\tilde{T}\|_1)/(\|\tilde{T}\tilde{K}_0\tilde{F}(\mathcal{R})\|_1)$ and a convergence threshold $\epsilon > 0$. Set $c = 0$.
3) Solve the following problem to construct $\mathcal{R}_c$:

$$\min_{\mathcal{R}} \quad \|I - \tilde{T}\|_1 + m\|\tilde{T}\tilde{K}_c\tilde{F}_N(\mathcal{R})\|_1$$

$$\text{s.t.} \quad R = 1 + \sum_j R_{ij}, \qquad i = 0, 1, \ldots, N-1$$

$$R_{ij} \geq 0, \qquad j \leq i = 0, 1, \ldots, N-1$$

$$R \leq \overline{R}. \tag{10}$$

Note that the above problem has cost and constraints that are convex in all unknown parameters R_{ij} and is, therefore, efficiently computable using any standard convex optimization software. Denote the resulting optimal tracking cost as γ_c.
If $c >= 1$ and if $\gamma_{c-1} - \gamma_c \leq \epsilon$, then STOP. Else goto Step 4.
4) Fix $\mathcal{R}_c$ and compute new controller K_{c+1} by solving

$$\min_{K^{\text{LTI}}} \quad \|I - \tilde{T}\|_1 + m\|\tilde{T}\tilde{K}\tilde{F}_N(\mathcal{R}_c)\|_1$$

$$\text{s.t.} \quad K^{\text{LTI}} \text{ stabilizing}$$

$$\|\tilde{T}\|_1 + m\|\tilde{T}\tilde{K}\tilde{F}_N(\mathcal{R}_c)\|_1 \leq m. \tag{11}$$

Apply "Q-parameterization" or "Youla parameterization" to parameterize all stabilizing controllers and convert the above problem to a convex optimization problem of the form

$$\min_{Q} \quad \|V_0 - V_1 Q V_2\|_1 \tag{12}$$

$$\text{s.t.} \quad Q \text{ LTI stable, rational, proper}$$

$$\|f_1(Q)\|_1 + m\|f_2(Q)\|_1 \leq m$$

where V_0, V_1, and V_2 are derived from the system to equate $\|I - \tilde{T}\|_1 + m\|\tilde{T}\tilde{K}\tilde{F}_N(\mathcal{R}_c)\|_1$ to $\|V_0 - V_1 Q V_2\|_1$ and f_1 and f_2 are affine functions in Q such that $\|f_1(Q)\|_1 + m\|f_2(Q)\|_1$ equals $\|\tilde{T}\|_1 + m\|\tilde{T}\tilde{K}\tilde{F}_N(\mathcal{R}_c)\|_1$. Note that (12) is again

efficiently computable using any standard convex optimization software.

5) Increment c by 1 and goto Step 3.

A. Examples

We now execute one iteration of the $\mathcal{R}-K$ iteration tracking algorithm for one unstable plant and different $(\overline{R}, m)$ pairs. We set $\overline{r}=1$ and summarize the results for each example. In each case, we give the initial controller K_0 used to compute RB quantizer $\mathcal{R}_0$, which is entirely characterized by $R_{\text{block},0}$ and its tracking cost. We then show the controller K_1 computed by fixing the quantizer to be parameterized by $\mathcal{R}_0$, and the corresponding bound on the 1-norm of the tracking error. Finally, we fix the controller to K_1 and compute $\mathcal{R}_1$ to complete one iteration.

Tracking Example 1

- **Steps 1-2**:

$$G=\frac{10}{z+1},\quad (\overline{R}=10, m=30)$$
$$K_0=\frac{-0.0368z^2-0.0006z+0.0026}{z^3}$$

- **Step 3:**

$$R_{\text{block},0}=\begin{bmatrix}9.0000 & & & \\ 0.0500 & 8.9500 & & \\ & 0.0975 & 8.9025 & \\ & & 0.0532 & 8.9468\end{bmatrix}$$
$$R_0=10\quad \gamma_0=30.9563$$

- **Step 4:**

$$K_1=\frac{-0.06532z^2-0.03061z-0.007916}{z^3}$$
$$||I-\tilde{T}||_1+m||\tilde{T}\tilde{K}_0\tilde{F}_N(\mathcal{R}_0)||_1=30.9563$$
$$||I-\tilde{T}||_1+m||\tilde{T}\tilde{K}_1\tilde{F}_N(\mathcal{R}_0)||_1=25.0942$$

- **Step 3:**

$$R_{\text{block},1}=\begin{bmatrix}9.0000 & & & \\ & 9.0000 & & \\ & & 9.0000 & \\ & & 0.0899 & 8.9101\end{bmatrix}$$
$$R_1=10\quad \gamma_1=25.0917\quad (\gamma_0-\gamma_1)=5.8647$$

Tracking Example 2

- **Steps 1-2**:

$$G=\frac{10}{z+1}\quad (\overline{R}=15, m=30)$$
$$K_0=\frac{-0.0368z^2-0.0006z+0.0026}{z^3}$$

- **Step 3:**

$$R_{\text{block},0}=\begin{bmatrix}14.0000 & & & \\ 8.6722 & 5.3278 & & \\ & 5.9975 & 8.0025 & \\ & & 3.0004 & 10.9996\end{bmatrix}$$
$$R_0=15\quad \gamma_0=30.9326$$

- **Step 4:**

$$K_1=\frac{-0.06527z^2-0.03054z-0.007872}{z^3}$$
$$||I-\tilde{T}||_1+m||\tilde{T}\tilde{K}_0\tilde{F}_N(\mathcal{R}_0)||_1=30.9326$$
$$||I-\tilde{T}||_1+m||\tilde{T}\tilde{K}_1\tilde{F}_N(\mathcal{R}_0)||_1=25.0594$$

- **Step 3:**

$$R_{\text{block},1}=\begin{bmatrix}14.0000 & & & \\ & 14.0000 & & \\ & 1.8799 & 12.1201 & \\ & & 0.9405 & 13.0595\end{bmatrix}$$
$$R_1=15\quad \gamma_1=25.0474\quad (\gamma_0-\gamma_1)=5.8875$$

Tracking Example 3

- **Steps 1-2**:

$$G=\frac{10}{z+2}\quad (\overline{R}=15, m=50)$$
$$K_0=(-0.0368z^2-0.0006z+0.0026)/(z^3)$$

- **Step 3:**

$$R_{\text{block},0}=\begin{bmatrix}14.000 & & & \\ 5.1680 & 8.8320 & & \\ & 3.5761 & 10.4239 & \\ & & 1.7834 & 12.2166\end{bmatrix}$$
$$R_0=15\quad \gamma_0=30.9305$$

- **Step 4:**

$$K_1=\frac{-0.06416z^2-0.02832z-0.005319}{z^3}$$
$$||I-\tilde{T}||_1+m||\tilde{T}\tilde{K}_0\tilde{F}_N(\mathcal{R}_0)||_1=30.9305$$
$$||I-\tilde{T}||_1+m||\tilde{T}\tilde{K}_1\tilde{F}_N(\mathcal{R}_0)||_1=25.4943$$

- **Step 3:**

$$R_{\text{block},1}=\begin{bmatrix}14.000 & & & \\ & 14.000 & & \\ & 0.4685 & 13.5315 & \\ & & 0.4673 & 13.5327\end{bmatrix}$$
$$R_1=15\quad \gamma_1=25.4895\quad (\gamma_0-\gamma_1)=5.4410$$

From Examples 1–3, we make a few observations.

- If we fix m and increase $\overline{R}$, the tracking error cost improves. This makes sense as the channel becomes less restrictive as $\overline{R}$ increases. Fig. 5 below plots each of the two terms in the tracking cost versus the rate limit $\overline{R}$ for a fixed

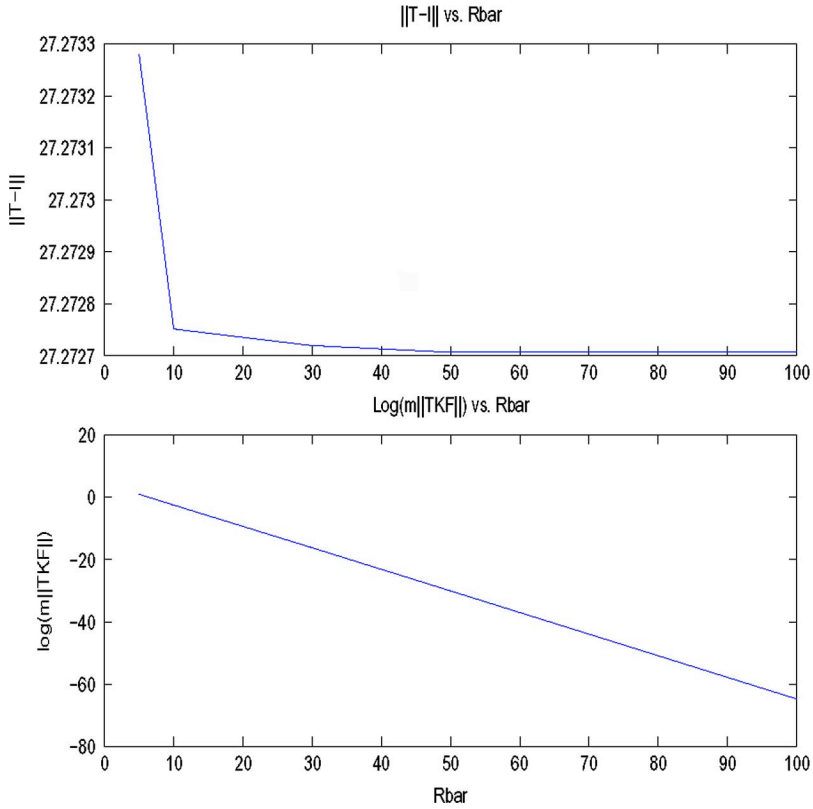


Fig. 5. (a) $||T-I||_1$ decays rapidly as $\overline{R}$ increases and settles to the ideal tracking cost when $R \to \infty$ (b) $m||TKF||_1$ is so small in comparison to $||T-I||_1$ that we plot $\log(m||TKF||_1)$ as a function of $\overline{R}$ to see its rapid decay to 0 as $R \to \infty$.

$m = 45$ large enough to be feasible for all $\overline{R}$s plotted for a given plant $G = (10)/(z + 1.1)$.

- If we fix $\overline{R}$ and increase m, the tracking error cost increases. This also makes sense because as m increases, the cost $||I - \tilde{T}||_1 + m||\tilde{T}\tilde{K}_0\tilde{F}_N(\mathcal{R}_0)||_1$ also increases (for a fixed controller and quantizer).
- The tracking cost in just one iteration is decreasing as $\gamma_0 - \gamma_1$ is positive for all examples.
- Although we do not show results for memoryless quantizers, we observed that quantizers with finite memory perform better than memoryless quantizers in tracking.

VII. Conclusion

In summary, we consider a plant and feedback controller separated by a finite-rate noiseless channel and we study a new class of encoders that have memory but do not have access to the control input to the plant. If the encoder is not local to the plant and if it does not know what controller will receive the signal it sends through the channel, then it can be modeled as belonging to this class. We have constructed a parameterization of time-varying quantizers that belong to this encoder class, and that leads to a convex characterization of bit-allocation strategies that maintain finite-gain stability. For finite memory quantizers, the convex characterization of stabilizing quantizers allows for efficient and nontrivial bit-allocation strategies and controllers to be synthesized for a given plant to meet various performance objectives.

In the future, we would like to obtain necessary conditions that maintain IO stability.

Appendix

A. Proof of Theorem 3.2

Below, we make some observations that lead to Theorem 3.2 and its proof. We assume the following norm conditions:

1) $||T||_1 < \infty$;
2) $||TK^{\mathrm{LTI}}||_1 < \infty$, and break down y_t as follows:

$$y_t = (Tr)_t + (TKF\overline{M}w)_t = yr_t + yq_t.$$

Under the first assumption, we show that there exists a positive finite constant C_1 and an $\alpha_1 (0 \leq \alpha_1 < 1)$, such that $|yr_t| \leq C_1\alpha_1^t$, for all $t \geq 0$.

Assuming that $||T||_1 < \infty$, there exists a ρ, with $0 \leq \rho < 1$, and a positive finite constant η_1, such that $|\{T\}_i| \leq \eta_1\rho^i$, for all $i \geq 0$. $\{T\}_i$ is the ith value of the impulse response of T. In fact, ρ can be chosen to be the spectral radius of the stable A matrix of a state-space description for T. Because T is LTI, $yr_t = \sum_{i=0}^{t}\{T\}_{t-i}r_i$, therefore

$$\begin{aligned} |yr_t| &\leq \sum_{i=0}^{t} |\{T\}_{t-i}||r_i| \\ &\leq \overline{r}\sum_{i=0}^{\infty} \eta_1\rho^{t-i}\gamma^i \\ &= \eta_1\overline{r}\rho^t \sum_{i=0}^{\infty}\left(\frac{\gamma}{\rho}\right)^i \\ &= \frac{\eta_1\overline{r}\rho}{\rho-\gamma}\rho^t \end{aligned}$$

where the last equality holds if $\gamma < \rho < 1$. Similarly, $|yr_t| = |\sum_{i=0}^{t}\{T\}_i r_{t-i}| \le (\eta_1\overline{r}\gamma)/(\gamma-\rho)\gamma^t$, if $\rho < \gamma < 1$. Putting both cases together, we get that

$$|yr_t| \le C_1\alpha_1^t$$

for all $t \ge 0$, where

$$C_1 = \max\left(\frac{\eta_1\overline{r}\gamma}{\gamma-\rho}, \frac{\eta_1\overline{r}\rho}{\rho-\gamma}\right) \quad \alpha_1 = \max(\rho,\gamma).$$

Under the second assumption $\|TK^{\mathrm{LTI}}\|_1 < \infty$, we show that there exists a positive finite constant C_2 and a constant $\alpha_2 (0 \le \alpha_2 < 1)$, such that $|yq_t| \le C_2\alpha_2^t$, for all $t \ge 0$.

Because $\|TK^{\mathrm{LTI}}\|_1 < \infty$, there exists a constant ν, with $0 \le \nu < 1$, and a positive constant η_2, such that $|\{TK^{\mathrm{LTI}}\}_i| \le \eta_2\nu^i$, for all $i \ge 0$. We now look at the magnitude of the response due to the exogenous disturbance induced by the quantizer yq_t

$$\begin{aligned}
|yq_t| &= \left|\sum_{j=0}^{t} T_{t-j}\sum_{i=0}^{j} K_{j-i}^{\mathrm{LTI}} m\beta^i 2^{-\sum_{l=0}^{j} R_{il}} w_i(j)\right| \\
&\le m\|F\|_1 \left|\sum_{j=0}^{t}\sum_{i=0}^{j} T_{t-j}K_{j-i}^{\mathrm{LTI}}\beta^i\right| \\
&\le m\|F\|_1 \sum_{j=0}^{t} |TK_{t-j}^{\mathrm{LTI}}|\beta^j \\
&\le \eta_2 m\|F\|_1 \sum_{j=0}^{\infty} \nu^{t-j}\beta^j \\
&= \eta_2 m\|F\|_1 \nu^t \sum_{j=0}^{\infty}\left(\frac{\beta}{\nu}\right)^j \\
&= \frac{\|F\|_1\eta_2 m\nu}{\nu-\beta}\nu^t
\end{aligned}$$

where the last equality holds if $\beta < \nu < 1$. If $\nu < \beta < 1$, then it is easy to show that $|yq_t| \le (\|F\|_1\eta_2 m\beta)/(\beta-\nu)\beta^t$. We then get that

$$|yq_t| \le C_2\alpha_2^t$$

for all $t \ge 0$, where

$$C_2 = \max\left(\frac{\|F\|_1 m\eta_2\beta}{\beta-\nu}, \frac{\|F\|_1 m\eta_2\nu}{\nu-\beta}\right)$$
$$\alpha_2 = \max(\nu,\beta).$$

Recall that β is a parameter we are looking for to ensure that $|y_t| \le m\beta^t$. Therefore, $\beta \ge \nu$, which gives us

$$C_2 = \frac{\|F\|_1 m\eta_2\beta}{\beta-\nu} \quad \alpha_2 = \beta.$$

Putting everything together, we get that for all $t \ge 0$

$$|y_t| \le |yr_t| + |yq_t| \le C_1\alpha_1^t + \frac{\|F\|_1 m\eta_2\beta}{\beta-\nu}\beta^t.$$

For the quantizer to be valid, we require $|y_t| \le m\beta^t$, for all $t \ge 0$. The above observations lead us to the following proof.

Proof: Define $C_1 = \max((\eta_1\overline{r}\gamma)/(\gamma-\rho), (\eta_1\overline{r}\rho)/(\rho-\gamma))$, as computed in Section III-B. Then, choose $\beta = \max(\rho,\nu,\gamma)$ and $m = (C1)/(1-(\eta_2\beta\|F\|_1)/(\beta-\nu))$. It is straightforward to show that m is finite and positive due to condition (3). To show IO stability, we have

$$\begin{aligned}
\|y^t\|_\infty &= \sup_t\{\|T_tG_t\mathbf{r}^t + T_tG_tK_tF_tM_t\mathbf{w}^t\|_\infty\} \\
&\le \|T\|_1\|r\|_\infty + m\|TKF\|_1 \\
&\le \|T\|_1\|r\|_\infty + m\|TK\|_1
\end{aligned}$$

where the last inequality comes from the fact that $\|F\|_1 \le 1$. To show quantizer validity, we computed in Section III-B that for $\alpha_1 = \max(\rho,\gamma)$

$$\begin{aligned}
|y_t| &\le C_1\alpha_1^t + \frac{\|F\|_1 m\eta_2\beta}{\beta-\nu}\beta^t \\
&\le m\beta^t
\end{aligned} \tag{13}$$

where the last inequality comes from our choices of β and m.

B. Using Matlab's "fmincon.m" to Synthesize Quantizers

In this section, we show how we use Matlab to solve the following optimization problem:

$$\min R$$
$$\text{s.t.} \sum_j R_{ij} = R-1, \qquad i = 0,2,\ldots,N-1 \tag{14}$$
$$\|(I-\tilde{G}\tilde{K}^{\mathrm{LTI}})^{-1}\tilde{G}\tilde{K}\tilde{F}_N\|_1 < 1 \tag{15}$$
$$R_{ij} \ge 0, \qquad j \le i = 0,1,\ldots,N-1 \tag{16}$$

For a given G and K^{LTI}, we lift the original system by a factor of N and denote the state-space description of $(I-\tilde{G}\tilde{K}^{\mathrm{LTI}})^{-1}\tilde{G}\tilde{K}$ by $(\tilde{A}_{cl}, \tilde{B}_{cl}, \tilde{C}_{cl}, \tilde{D}_{cl})$. Let $X \triangleq \vec{(}R_{\mathrm{block}})$ and rewrite the above optimization problem as

$$\begin{aligned}
&\min X_1 \\
&\text{s.t.} \quad A_{eq}X = B_{eq} \\
&P(X) < 0 \\
&X \ge 0
\end{aligned}$$

where $A_{eq}X = B_{eq}$ captures the equality constraints (2), $P(X) < 0$ captures constraints (3), and $X \ge 0$ are equivalent to constraints (4). For example, if $N = 4$, then the equation shown at the top of the next page holds, with

$$\begin{aligned}
c_1' &= [1\ \ 1\ \ 1\ \ 1\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0] \\
c_2' &= [0\ \ 0\ \ 0\ \ 0\ \ 1\ \ 1\ \ 1\ \ 0\ \ 0\ \ 0] \\
c_3' &= [0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 1\ \ 1\ \ 0] \\
c_4' &= [0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 1]
\end{aligned}$$

and

$$\tilde{F}_4(X) = \frac{1}{2}\begin{bmatrix} f_1 & & & \\ & f_2 & & \\ & & f_3 & \\ & & & f_4 \end{bmatrix}$$

$$R_{\text{block}} = \begin{bmatrix} R_{00} & & & \\ R_{01} & R_{11} & & \\ R_{02} & R_{12} & R_{22} & \\ R_{03} & R_{13} & R_{23} & R_{33} \end{bmatrix}$$

$$X = \begin{bmatrix} R_{00} & R_{01} & R_{02} & R_{03} & R_{11} & R_{12} & R_{13} & R_{22} & R_{23} & R_{33} \end{bmatrix}'$$

$$A_{eq} = \begin{bmatrix} -1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

$$B_{eq} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}'$$

and

$$P(X) = \|\tilde{D}_{cl} F_4(X)\|_1 + \sum_{k=1}^{n} \|\tilde{C}_{cl} \tilde{A}_{cl}^{k-1} \tilde{B}_{cl} F_4(X)\|_1 - 1$$

where

$$f_1 = 2^{-X_1}$$

$$f_2 = \begin{bmatrix} 2^{-\sum_{i=1}^{2} X_i} & \\ & 2^{-X_5} \end{bmatrix}$$

$$f_3 = \begin{bmatrix} 2^{-\sum_{i=1}^{3} X_i} & & \\ & 2^{-\sum_{i=5}^{6} X_i} & \\ & & 2^{-X_8} \end{bmatrix}$$

$$f_4 = \begin{bmatrix} 2^{-\sum_{i=1}^{4} X_i} & & & \\ & 2^{-\sum_{i=5}^{7} X_i} & & \\ & & 2^{-\sum_{i=8}^{9} X_i} & \\ & & & 2^{-X_{10}} \end{bmatrix}.$$

Note that we approximate the 1-norm of $(I - \tilde{G}\tilde{K}^{\text{LTI}})^{-1}\tilde{G}\tilde{K}\tilde{F}_N$ to guarantee that we can numerically compute the solution to the constrained minimization problem.[5]

REFERENCES

[1] J. Baillieul, "Feedback coding for information-based control: Operating near the data-rate limit," in *Proc. 41st IEEE Conf. Decision Control*, Dec. 2002, pp. 3229–3236.

[2] G. J. Balas, J. C. Doyle, K. Glover, A. Packard, and R. Smith, μ-analysis and synthesis toolbox The MathWorks, Inc., Natick, MA, 1993.

[3] W. S. Wong and R. Brockett, "Systems with finite communication bandwidth constraints-I: State estimation problems," *IEEE Trans. Autom. Control*, vol. 42, no. 9, pp. 1294–1299, Sep. 1997.

[4] R. Brockett, "Quantized feedback stabilization of linear systems," *IEEE Trans. Autom. Control*, vol. 45, no. 7, pp. 1279–1289, Jul. 2000.

[5] W. S. Wong and R. Brockett, "Systems with finite communication bandwidth constraints-II: Stabilization with limited information feedback," *IEEE Trans. Autom. Control*, vol. 44, no. 5, pp. 1049–1053, May 1999.

[6] M. A. Dahleh and I. Diaz-Bobillo, *Control of Uncertain Systems: A Linear Programming Approach*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[5]To guarantee that the approximate 1-norm satisfies our stability condition, we require a finite-sum upper bound on the 1-norm of $(I - \tilde{G}\tilde{K})^{-1}\tilde{G}\tilde{K}\tilde{F}_N$, to satisfy the condition. See [6] for details on such an upper bound.

[7] D. Delchamps, "Stabilizing a linear system with quantized state feedback," *IEEE Trans. Autom. Control*, vol. 35, no. 8, pp. 916–924, Aug. 1990.

[8] N. Elia and S. Mitter, "Stabilization of linear systems with limited information," *IEEE Trans. Autom. Control*, vol. 46, no. 9, pp. 1384–1400, Sep. 2001.

[9] F. Fagnani and S. Zampieri, "Stability analysis and synthesis for scalar linear systems with a quantized feedback," *IEEE Trans. Autom. Control*, vol. 48, no. 9, pp. 1569–1584, Sep. 2003.

[10] F. Fagnani and S. Zampieri, "Quantized stabilization of linear systems: Complexity versus performance," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1534–1548, Sep. 2004.

[11] H. Ishii and B. Francis, *Limited Data Rate in Control Systems With Networks*. Berlin, Germany: Springer-Verlag, 2002.

[12] Q. Ling and M. D. Lemmon, "Stability of quantized control systems under dynamic bit assignment," *IEEE Trans. Autom. Control*, vol. 50, no. 5, pp. 734–740, May 2005.

[13] Q. Ling and M. D. Lemmon, "Optimal dropout compensation in networked control systems," in *Proc. 42nd IEEE Conf. Decision Control*, Dec. 2003, pp. 670–675.

[14] D. Liberzon, "A note on stabilization of linear systems using coding and limited communication," in *Proc. 41st IEEE Conf. Decision Control*, Dec. 2002, pp. 836–841.

[15] N. C. Martins, N. Elia, and M. A. Dahleh, "Stabilization of uncertain systems in the presence of a direct link," *IEEE Trans. Autom. Control*, vol. 51, no. 3, pp. 438–447, Mar. 2006.

[16] N. C. Martins and M. A. Dahleh, "Feedback control in the presence of noisy channels: Bode-like fundamental limitations of performance," *IEEE Trans. Autom. Control*, accepted for publication.

[17] N. C. Martins, M. A. Dahleh, and J. C. Doyle, "Fundamental limitations of disturbance attenuation with side information," *IEEE Trans. Autom. Control*, vol. 52, no. 1, pp. 56–66, Jan. 2007.

[18] A. Megretski, "Lecture notes for 6.245: Multivariable control systems, Q-parameterization chapter," Massachusetts Inst. Technol., Cambridge, MA, 2004.

[19] G. N. Nair and R. J. Evans, "Stabilization with data-rate-limited feedback: Tightest attainable bounds," *Syst. Control Lett.*, vol. 41, pp. 49–56, 2000.

[20] G. N. Nair and R. J. Evans, "Stabilizability of stochastic linear systems with finite feedback data rates," *SIAM J. Control Optimization*, vol. 43, no. 2, pp. 413–436, Jul. 2, 2004.

[21] A. Packard, "Gain scheduling via linear fractional transformations," *Syst. Control Lett.*, vol. 22, pp. 79–92, 1994.

[22] S. Tatikonda, "Control under communication constraints," *IEEE Trans. Autom. Control*, vol. 49, no. 7, pp. 1056–1068, Jul. 2004.

[23] S. Tatikonda, A. Sahai, and S. Mitter, "Control of LQG systems under communication constraints," in *Proc. Amer. Control Conf.*, Jun. 1999, pp. 2778–2782.

[24] S. Yuksel and T. Basar, "State estimation and control for linear systems over communication networks," in *Proc. 2003 IEEE Conf. Control Appl.*, Jun. 2003, vol. 1, pp. 409–414.

Sridevi V. Sarma (M'07) received the B.S. (1994) from Cornell University and the M.S. (1997) and Ph.D. degrees (2006) from Massachusetts Institute of Technology in Electrical Engineering and Computer Science. She is now a postdoctoral fellow at Harvard Medical School. Her research interests include control of constrained and defective systems (applications in neuroscience) and large-scale optimization. She is president and cofounder of Infolenz Corporation, a Marketing Analytics company.

Dr. Sarma is a recipient of the GE faculty for the future scholarship, a National Science Foundation graduate research fellow, a recipient of the L'Oreal USA fellowships for Women in Science, and a recipient of the Burroughs Wellcome Fund Careers at the Scientific Interface Award.

Munther A. Dahleh (S'84–M'87–SM'97–F'01) was born in 1962. He received the B.S. degree from Texas A&M University, College Station, in 1983, and the Ph.D. degree from Rice University, Houston, TX, in 1987, all in electrical engineering.

Since then, he has been with the Department of Electrical Engineering and Computer Science at MIT, where he is now a full Professor. He is interested in problems at the interface of robust control, filtering, information theory, and computation. He is also interested in model reduction of discrete-alphabet hidden Markov models, universal learning approaches for systems with both continuous and discrete alphabets, and problems at the interface between systems theory and neurobiology.

Srinivasa M. Salapaka (M'03) received the B.Tech. degree from Indian Institute of Technology in 1995, the M.S. and the Ph.D. degrees from the University of California at Santa Barbara in 1997 and 2002 respectively in Mechanical Engineering. During 2002–2003, he was a postdoctoral associate in the Laboratory for Information and Decision Systems at Massachusetts Institute of Technology. In January 2004, he joined the Mechanical Science and Engineering department at the University of Illinois, Urbana-Champaign. His areas of current research interest include nanotechnology, combinatorial resource allocation, and numerical analysis of integral equations. He is the CAREER award recipient for the year 2005.