# Joint Optimization of Communication Rates and Linear Systems

Lin Xiao[1] Mikael Johansson[2] Haitham Hindi[3]
Stephen Boyd[4] Andrea Goldsmith[5]

Information Systems Laboratory
Stanford University

June 20, 2001

[1]lxiao@stanford.edu (contact author)
[2]mikaelj@stanford.edu
[3]hhindi@stanford.edu
[4]boyd@stanford.edu
[5]andrea@systems.stanford.edu
Research supported in part by NSF Grant ECS-9707111, by AFOSR Grant F49620-98-1-0147, by DARPA Contract F33615-99-C-3014, by NSF Grant CCR-9714002, by SRI (DARPA) Contract: 50-000018, by Caltrans PATH program under contract MOV4212, by the Swedish Foundation for International Cooperation in Higher Education (STINT), and by a Stanford Graduate Fellowship.

**Abstract**

We consider a linear system, such as a controller or estimator, in which several signals are transmitted over communication channels with bit rate limitations.We focus on finding the allocation of communication resources such as transmission powers, bandwidths, or time-slot fractions, that yields optimal system performance.

Assuming conventional uniform quantization and a standard white-noise model for quantization errors, we consider two specific problems. In the first, we assume that the linear system is fixed and address the problem of allocating communication resources to optimize system performance. We observe that this problem is often convex (at least, when we ignore the constraint that individual quantizers have an integral number of bits), hence readily solved. We describe a general dual decomposition method for solving these problems that exploits the special structure often found in network resource allocation problems. This method reduces to the standard waterfilling techniques used in problems with only one coupling constraint. We briefly describe how the integer bit constraints can be handled, and give a bound on how suboptimal these heuristics can be.

The second problem we consider is that of jointly allocating communication resources and designing the linear system in order to optimize system performance. This problem is in general not convex, but can be solved heuristically in a way that exploits the special structure of the communication resource allocation problems, and appears to work well in practice.

We demonstrate these ideas and methods on two numerical examples. In the first, we consider a networked estimator in which sensors transmit measurements over a multiple access channel, and we optimize bandwidth, power allocation, and bit rates to the sensors. In the second example, we consider a networked LQG controller, in which the sensor signals are transmitted over a multiple access channel and the actuator signals are transmitted over a broadcast channel. The sensor and actuator channels have separate power limits, but share a common bandwidth constraint. Here we allocate power and bandwidth to each actuator and sensor channel, as well as the total bandwidth available to the sensors and actuators, and in addition optimize the controller itself.

# 1 Introduction

We consider a linear system in which several signals are transmitted over wireless communication links, as shown in figure 1. All signals are vector-valued: $w$ is a vector of exogenous signals (such as disturbances or noises acting on the system); $z$ is a vector of performance signals (including error signals and actuator signals); and $y$ and $y_r$ are the signals transmitted over the communication network, and received, respectively. This general arrangement can represent a variety of systems, for example a controller or estimator in which sensor, actuator, or command signals are sent over wireless links. It can also represent a distributed controller or estimator, in which some signals (*i.e.*, inter-process communication) are communicated across a network.

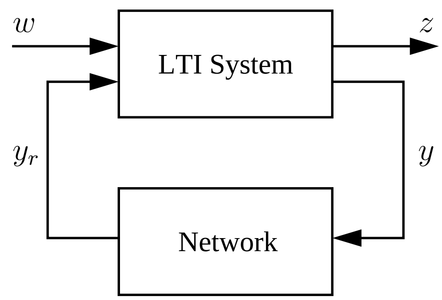


**Figure 1:** System set-up. The exogenous signal $w$ includes noises, disturbances, and commands; the output signal $z$ contains the critical performance variables. The signal $y$ is transmitted over a network communication system; $y_r$ is the received, decoded version of $y$.

Many issues arise in the design of networked controllers and the associated communication systems, including bit rate limitations, communication delays, packet loss, transmission errors, and asynchronicity (see, for example, [WB97, NBW98, XHH00, SSK99, Özg89]). In this paper we consider only the first issue, *i.e.*, bit rate limitations. In other words, we assume that each communication link has a fixed and known delay (which we model as part of the LTI system), does not drop packets, transfers bits without error, and operates (at least for purposes of analysis) synchronously with the discrete-time linear system.

Our focus is on the optimal allocation of the underlying communication resources, which in turn limits the achievable bit rates. For a fixed sampling frequency $f_s$ of the linear system this translates into a constraint on the number of bits that can be transmitted over each communication channel during one sampling period. We will assume that the individual signals $y_i$ are coded using conventional memoryless uniform quantizers, as shown in figure 2. This coding scheme is certainly not optimal (see, *e.g.*, [WB97, NE98]), but it is conventional, easily implemented, and allows us to use a simple and standard model for the loss of system performance due to the network communication constraints.

Much work has been done in the control and signal processing literature on the allocation of bits in linear systems with quantizers. The main effort has been to derive analysis and design methods for fixed-point filter and controller implementations, for which there is a vast literature (see [Wil85, WK89, SW90]). In that case, the bit constraints arise due to hardware

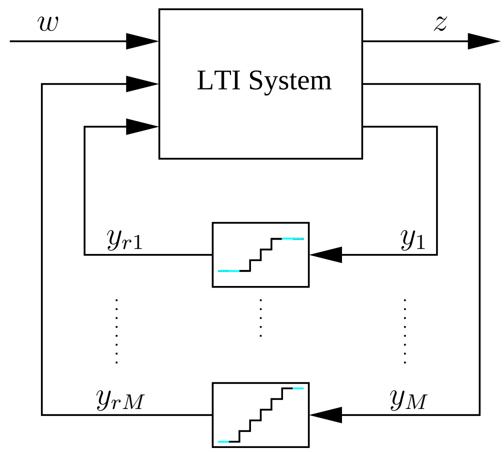


**Figure 2:** The delays in the network are assumed known and fixed, and modeled as part of the LTI system. The communication channels are modeled as uniform quantizers, with (possibly) different numbers of bits and scale factors.

or software limitations. In our case, the limitations arise due to the bit rate constraints of the communication network.

In this paper we assume that the coding scheme and medium access control is fixed. We concentrate on the selection of certain critical communication parameters such as individual channel transmission powers, and the bandwidths or time-slot fractions allocated to the channels (or groups of channels). We refer to these critical communication parameters collectively as *communication variables.* The communication variables, which are the design variables for us, indirectly limit the number of bits allocated to each quantizer and hence affect overall system performance.

We consider two specific problems in this paper. First, we assume the linear system is fixed and consider the problem of choosing the communication variables to optimize the overall system performance. We observe that this problem is often convex, provided we ignore the constraint that the number of bits for each quantizer is an integer. This means that (ignoring the integrality constraint) these communication resource allocation problems can be solved efficiently, using a variety of convex optimization techniques. We describe a general approach for solving these problems based on dual decomposition. The method results in very efficient procedures for solving for many communication resource allocation problems, and reduces to well known water-filling in simple cases. We also show several methods that can be used to handle the integrality constraint. The simplest is to round down the number of bits for each channel to the nearest integer. We show that this results in an allocation of communication resources that is feasible, and at most a factor of two suboptimal in terms of the RMS (root-mean-square) value of critical variable $z$. We also describe a simple and effective heuristic that often achieves performance close to the bound obtained by solving the convex problem, ignoring the integrality constraints.

The second problem we consider is the problem of jointly allocating communication resources and designing the linear system in order to optimize performance. Here we have two sets of design variables: the communication variables (which indirectly determine the number of bits assigned to each quantizer), and the controller variables (such as estimator or controller gains in the linear system). Clearly the two are strongly coupled, since the effect of quantization errors depends on the linear system, and similarly, the choice of linear system will affect the choice of communication resource allocation. We show that this joint problem is in general not convex. We propose an alternating optimization method that exploits problem structure and appears to work well in practice.

In §2, we describe the linear system and our model for the effect of uniform quantization error on overall system performance. In §3, we describe a generic convex model for the bit rate limitations imposed by communication systems, and describe several examples. In §4, we formulate the communication resource allocation problem for fixed linear systems, describe the dual decomposition method which exploits the separable structure, and give a heuristic rounding method to deal with the integrality of bit allocations. In §5, we demonstrate the nonconvexity of the joint design problem, and give a iterative heuristic to solve such problems. Two examples, a networked linear estimator and a LQG control system over communication networks, are used to illustrate the optimization algorithms in §4 and §5. We conclude the paper in §6.

# 2 Linear system and quantizer model

## 2.1 Linear system model

To simplify the presentation we assume a synchronous, single-rate discrete-time system. The linear time-invariant (LTI) system can be described as

$$z = G_{11}(\varphi)w + G_{12}(\varphi)y_r, \qquad y = G_{21}(\varphi)w + G_{22}(\varphi)y_r, \tag{1}$$

where $G_{ij}$ are LTI operators (*i.e.*, convolution systems described by transfer or impulse matrices). Here, $\varphi \in \mathbf{R}^q$ is the vector of design parameters in the linear system that can be tuned or changed to optimize performance. To give lighter notation, we suppress the dependence of $G_{ij}$ on $\varphi$ except when necessary. We assume that $y(t),\ y_r(t) \in \mathbf{R}^M$, *i.e.*, the $M$ scalar signals $y_1, \ldots, y_M$ are transmitted over the network during each sampling period.

We assume that the signals sent (*i.e.*, $y$) and received (*i.e.*, $y_r$) over the communication links are related by memoryless scalar quantization, which we describe in detail in the next subsections. This means that all communication delays are assumed constant and known, and included in the LTI system model.

## 2.2 Quantization model

### 2.2.1 Unit uniform quantizer

We first describe a unit range uniform $b$-bit quantizer. Such a quantizer partitions the range $[-1, 1]$ into $2^b$ intervals of uniform width $2^{1-b}$. To each quantization interval a codeword of $b$ bits is assigned. Given the associated codeword, the value is approximated by (or reconstructed as) $u_r$, which is the midpoint of the interval corresponding to the codeword.

The relationship between the original and reconstructed values can be expressed as

$$Q_b(u) = \frac{\mathbf{round}(2^{b-1}u)}{2^{b-1}} \tag{2}$$

for $|u| < 1$. Here, $\mathbf{round}(z)$ is the integer nearest to $z$ (with ties rounded down).

The behavior of the quantizer when $u$ overflows (*i.e.*, $|u| \geq 1$) is not specified. One approach is to introduce two more codewords, corresponding to negative and positive overflow, respectively, and to extend $Q_b$ to saturate for $|u| \geq 1$. The details of the overflow behavior will not affect our analysis or design, since we assume by appropriate scaling (described below) that overflow does not occur, or occurs rarely enough to not affect overall system performance.

The associated quantization error, $u_r - u$, can be expressed as

$$E_b(u) = u_r - u = \frac{\mathbf{round}(2^{b-1}u) - 2^{b-1}u}{2^{b-1}}.$$

As long as the quantizer does not overflow, the numerator in the right-hand expression lies between $\pm 1/2$ and the quantization error $E_b(u)$ lies in the interval $\pm 2^{-b}$.

### 2.2.2 Scaling

To avoid overflow, each signal $y_i(t)$ is scaled by the factor $s_i^{-1} > 0$ prior to encoding with a unit uniform $b_i$-bit quantizer, and re-scaled by the factor $s_i$ after decoding (figure 3), so that

$$y_{ri}(t) = s_i Q_{b_i}(y_i(t)/s_i).$$

The associated quantization error is given by

$$q_i(t) = y_{ri}(t) - y_i(t) = s_i E_{b_i}(y_i(t)/s_i),$$

which lies in the interval $\pm s_i 2^{-b_i}$, provided $|y_i(t)| < s_i$.

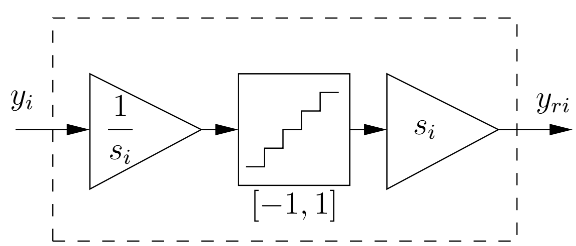


**Figure 3:** Scaling before and after the quantizer.

To minimize quantization error while ensuring no overflow (or ensuring that overflow is rare) the scale factors $s_i$ should be chosen as the maximum possible value of $|y_i(t)|$, or as a value that with very high probability is larger than $|y_i(t)|$. For example, we can use the so-called $3\sigma$-*rule*,

$$s_i = 3\,\mathbf{rms}(y_i),$$

where $\mathbf{rms}(y_i)$ denotes the root-mean-square value of $y_i$,

$$\mathbf{rms}(y_i) = \left(\lim_{t\to\infty} \mathbf{E}\, y_i(t)^2\right)^{1/2}.$$

If $y_i$ has a Gaussian amplitude distribution, this choice of scaling ensures that overflow occurs only about 0.3% of the time.

### 2.2.3 White-noise quantization error model

We adopt the standard stochastic quantization noise model introduced by Widrow (see, *e.g.*, [FPW90, Chapter 10]). Assuming that overflow is rare, we model the quantization errors $q_i(t)$ as independent random variables, uniformly distributed on the interval

$$s_i[-2^{-b_i}, 2^{-b_i}].$$

(This model is reasonable when the quantized signals $y_i(t)$ change by at least several levels from sample to sample.) In other words, we model the effect of quantizing $y_i(t)$ as an additive white noise source $q_i(t)$ with zero mean and variance $\mathbf{E}\, q_i(t)^2 = (1/3)s_i^2 2^{-2b_i}$.

Using the white noise quantization noise model, we obtain the system shown in figure 4.

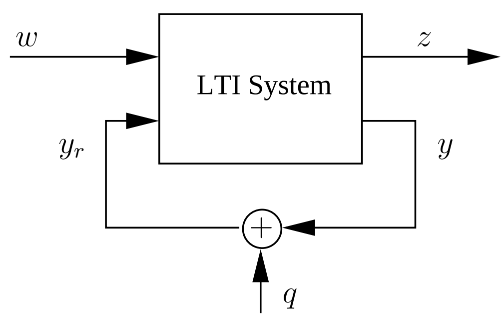


**Figure 4:** LTI system with white noise quantization noise model. The LTI system is driven by two exogenous inputs, $w$ and $q$.

## 2.3 Performance of the closed-loop system

We can express $z$ and $y$ in terms of the inputs $w$ and $q$ as

$$z = G_{zw}w + G_{zq}q, \qquad y = G_{yw}w + G_{yq}q,$$

where

$$\begin{aligned} G_{zw} &= G_{11} + G_{12}(I - G_{22})^{-1}G_{21}, & G_{zq} &= G_{12} + G_{12}(I - G_{22})^{-1}G_{22} \\ G_{yw} &= (I - G_{22})^{-1}G_{21}, & G_{yq} &= (I - G_{22})^{-1}G_{22} \end{aligned}$$

are the closed-loop transfer matrices from $w$ and $q$ to $z$ and $y$, respectively.

From the expression for $z$, we see that it consists of two terms: $G_{zw}w$, which is what $z$ would be if the quantization were absent, and $G_{zq}q$, which is the component of $z$ due to the quantization. The variance of $z$ induced by the quantization is given by

$$V_q = \mathbf{E}\,\|G_{zq}q\|^2 = \sum_{i=1}^{M} \|G_{zqi}\|^2 (1/3) s_i^2 2^{-2b_i},$$

where $G_{zqi}$ is the $i$th column of the transfer matrix $G_{zq}$, and $\|\cdot\|$ denotes the $\mathbf{L}^2$ norm (see [BB91, §5.2.3]). We can use $V_q$ as a measure of the effect of quantization on the overall system performance. If $w$ is also modeled as a stationary stochastic process, the overall variance of $z$ is given by

$$V = \mathbf{E}\,\|z\|^2 = V_q + \mathbf{E}\,\|G_{zw}w\|^2. \tag{3}$$

For future use, we express $V_q$ in the form

$$V_q = \sum_{i=1}^{M} a_i 2^{-2b_i}, \tag{4}$$

where $a_i = (1/3)\|G_{zqi}\|^2 s_i^2$. This expression shows how $V_q$ depends on the allocation of quantizer bits $b_1, \ldots, b_M$, as well as the scalings $s_1, \ldots, s_M$ and LTI system (which affect the $a_i$'s).

We remind the reader of our assumption that overflow is rare, *i.e.*, that $|y_i(t)| < s_i$ with high probability. This requirement can be expressed by finding the variance of $y_i$ from $y = G_{yw}w + G_{yq}q$, and imposing the $3\sigma$-rule. We also note that while the formula (4) was derived assuming that $b_i$ are integers, it makes sense for $b_i \in \mathbf{R}$.

# 3 Communications model and assumptions

## 3.1 A generic model for bit rate constraints

In this section we describe our model of the communication system. The capacity of the individual links depend on the media access scheme and the selection of certain critical parameters, such as transmission powers and bandwidths or time-slot fractions allocated to individual channels (or groups of channels). We refer to these critical communication parameters collectively as *communication variables*, and denote the vector of communication variables by $\theta$. The communication variables are themselves limited by various resource constraints, such as limits on the total power or total bandwidth available. We will assume that the medium access methods and coding and modulation schemes are fixed, but that we can optimize over the underlying communication variables $\theta$.

We let $b \in \mathbf{R}^M$ denote the vector of bits allocated to each quantized signal. The associated communication rate $r_i$ (in bits per second) can be expressed as $b_i = \alpha r_i$, where the constant $\alpha$ has the form $\alpha = c_s/f_s$. Here $f_s$ is the sample frequency, and $c_s$ is the channel coding efficiency in source bits per transmission bit. This relationship will allow us to express capacity constraints in terms of bit allocations rather than communication rates.

We will use the following general model to relate the vector of bit allocations $b$, and the vector of communication variables $\theta$:

$$\begin{array}{rcll} f_i(b,\theta) & \leq & 0, & i=1,\ldots,m_f \\ h_i^T\theta & \leq & d_i, & i=1,\ldots,m_h \\ \theta_i & \geq & 0, & i=1,\ldots,m_\theta \\ \underline{b}_i \ \leq \ b_i & \leq & \overline{b}_i, & i=1,\ldots,M \end{array} \tag{5}$$

We make the following assumptions about this generic model.

- The functions $f_i$ are convex functions of $(b,\theta)$, monotone increasing in $b$ and monotone decreasing in $\theta$. These inequalities describe capacity constraints on individual links or groups of links. The monotonicity conditions on $f_i$ mean that, roughly speaking, the capacity of the channels increase with increasing resources.

- The second set of constraints describes resource limitations, such as a total available power or bandwidth for a group of channels. We assume the vectors $h_i$ have nonnegative entries. We assume that $d_i$, which represent resource limits, are positive.

- The third constraint specifies that the communication resource variables (which represent powers, bandwidths, time-slot fractions) are nonnegative.

- The last group of inequalities specify lower and upper bounds for each bit allocation. We assume that $\underline{b}_i$ and $\overline{b}_i$ are (nonnegative) integers. The lower bounds are imposed to ensure that the white noise model for quantization errors is reasonable. The upper bounds can arise from hardware limitations.

This generic model will allow us to formulate the communication resource allocation problem, *i.e.*, the problem of choosing $\theta$ to optimize overall system performance, as a convex optimization problem.

There is also one more important constraint on $b$ not included in the model above:

$$b_i \text{ is an integer}, \quad i=1\ldots,M. \tag{6}$$

For the moment, we ignore this constraint. We will return to it in §4.2.

## 3.2 Capacity constraints

In this section, we describe some simple channel models, showing how they fit the generic model (5) given above. More detailed descriptions of these channel models, as well as derivations, can be found in, *e.g.*, [CT91, Gol99].

### 3.2.1 Gaussian channel

We start by considering a single Gaussian channel. The communication variables are the bandwidth $W>0$ and transmission power $P>0$. Let $N$ be the power spectral density of the additive white Gaussian noise at the front-end of the receiver. The channel capacity is given by ([CT91])

$$R = W \log_2 \left(1 + \frac{P}{N\,W}\right)$$

(in bits per second). The achievable communication rate $r$ is bounded by this channel capacity, *i.e.*, we must have $r \leq R$. Expressed in terms of $b$, we have

$$b \leq \alpha W \log_2 \left(1 + \frac{P}{NW}\right). \tag{7}$$

We can express this in the form

$$f(b, W, P) = b - \alpha W \log_2 \left(1 + \frac{P}{NW}\right) \leq 0,$$

which fits the generic form (5). To see that the function $f$ is jointly convex in the variables $(b, W, P)$, we note that the function $g(P) = -\alpha \log_2(1 + P/N)$ is a convex function of $P$ and, therefore its *perspective function* (see [BV98])

$$W g(P/W) = -\alpha W \log_2 \left(1 + \frac{P}{NW}\right)$$

is a convex function of $(P, W)$. Adding the linear (hence convex) function $b$ establishes convexity of $f$. It is easily verified that $f$ is monotone increasing in $b$, and monotone decreasing in $W$ and $P$.

### 3.2.2 Parallel Gaussian channels with total power constraint

We now consider a set of $n$ independent Gaussian channels, with a total power limit $P_{\text{tot}}$. The communication variables are the transmission powers $P_i$, which are limited by the total power constraint

$$P_1 + \cdots + P_n \leq P_{\text{tot}},$$

which has the generic form of a resource limit, *i.e.*, a linear inequality in the communication variables. We could also impose upper bounds on the transmission powers.

The capacity is constrained by

$$b_i \leq \alpha W \log_2 \left(1 + \frac{P_i}{N_i W}\right), \qquad i = 1, \ldots, n, \tag{8}$$

where $W$ is the bandwidth, and $N_i$ is spectral density of the receiver noise.

### 3.2.3 Gaussian broadcast channel with FDMA

In the Gaussian broadcast channel with frequency-domain multiple access (FDMA), a transmitter sends information to $n$ receivers over disjoint frequency bands with bandwidths $W_i > 0$. The communication parameters are the bandwidths $W_i$ and the transmit powers $P_i > 0$ for each individual channel. The communication variables are constrained by a total power limit

$$P_1 + \cdots + P_n \leq P_{\text{tot}}$$

and a total available bandwidth limit

$$W_1 + \cdots + W_n \leq W_{\rm tot},$$

which have the generic form for communication resource limits.

The receivers are subject to independent white Gaussian noises with power spectral densities $N_i$. The transmitter assigns power $P_i$ and bandwidth $W_i$ to the $i$th receiver. The achievable bit rates $b$ are constrained by

$$b_i \leq \alpha W_i \log_2\left(1 + \frac{P_i}{N_i W_i}\right), \qquad i = 1, \ldots, n. \tag{9}$$

Again, the constraints relating $b$ and $\theta = (P, W)$ have the generic form (5).

### 3.2.4 Gaussian broadcast channel with TDMA

We consider a Gaussian broadcast channel with time-division multiplexing (TDMA). The transmitter assigns a fraction $\tau_i$ of the sampling period of the linear system to the $i$th receiver; during each time slot, it transmits with power $P$ and bandwidth $W$ to a single receiver. Here the communication variables are $\tau_1, \ldots, \tau_n$, which satisfy the constraint

$$\tau_1 + \cdots + \tau_n \leq 1.$$

(usually we have equality here for an optimal resource allocation). We assume that the power $P$ and bandwidth $W$ are fixed, *i.e.*, they are not communication variables.

The achievable rates satisfy

$$b_i \leq \tau_i \alpha W \log_2\left(1 + \frac{P}{N_i W}\right), \quad i = 1, \ldots, n,$$

which is readily expressed in the generic form. Indeed, these constraints are simple linear inequalities relating $b_i$ and $\tau_i$.

### 3.2.5 Gaussian multiple access channel with CDMA

In the Gaussian multiple access channel (MAC) $n$ transmitters, each with power $P_i$, send information to a common receiver which is corrupted by additive white Gaussian noise of power density $N$. For code-division multiplexing, the achievable rates $b$ satisfy the set of constraints

$$\sum_{i\in Z} b_i \leq \alpha W \log_2\left(1 + \frac{\sum_{i\in Z} P_i}{NW}\right) \quad \text{for all} \quad Z \subseteq \{1, 2, \ldots, n\}. \tag{10}$$

The communication variables here are the transmission powers $P_i$, which satisfy $0 \leq P_i \leq \bar{P}_i$ where $\bar{P}_i$ is the upper bound for $P_i$, or a total power limit. These inequalities also have the generic convex form (5).

With separate power constraints $0 \leq P_i \leq \bar{P}_i$, the $2^n - 1$ constraints in (10), together with $b_i \geq 0$, give the rate region (set of achievable $b_i$'s, more precisely of $r_i = b_i/\alpha$) a *polymatroid* structure. Tse [TH98] exploits this polymatroid structure (in the more general fading channel context) and shows that the optimal power allocation can be explicitly obtained in a greedy manner.

When the powers $P_i$ are not constrained except through their sum, the rate region is simply a simplex:

$$\left\{ b \;\middle|\; \sum_{i=1}^{n} b_i \le \alpha W \log_2\left(1 + \frac{P_{\rm tot}}{NW}\right),\ b_i \ge 0,\ i = 1, \ldots, n \right\},$$

which can directly serve as the constraints in the bit allocation problem in §4.

### 3.2.6 Gaussian multiple access channel with FDMA

In a Gaussian multiple access channel with FDMA, each transmitter sends information in disjoint frequency bands with bandwidth $W_i$. The achievable bit rates are determined by the constraints

$$b_i \le \alpha W_i \log_2\left(1 + \frac{P_i}{NW_i}\right), \qquad i = 1, \ldots, n.$$

Here the communication variables are the powers $P_i$ and bandwidths $W_i$, which are limited by separate or total power constraints, and a total bandwidth constraint.

For multiple access channel, the total power constraint may not represent a physical constraint as it does in the broadcast channel, since the transmitters might not use the same power source.

### 3.2.7 Variations and extensions

Channels with time-varying gain variations (fading) as well as rate constraints based on bit error rates (with or without coding) can be formulated in a similar manner; see, *e.g.*, [LG01, CG].

### 3.2.8 Modeling complex communication systems

We can combine the channel models described above to model more complex communication systems. As a simple example, suppose that we have $M = 6$, *i.e.*, six signals transmitted over channels. The first three signals $y_1,\ y_2,\ y_3$ are transmitted over a broadcast channel, while the last three signals $y_4$, $y_5$, and $y_6$, are transmitted over a multiple access channel. The communication variables are the six associated powers $P_1, \ldots, P_6$, and the associated bandwidths $W_1, \ldots, W_6$. We have the capacity constraints

$$b_i \le \alpha W_i \log_2\left(1 + \frac{P_i}{N_i W_i}\right), \qquad i = 1, \ldots, 6,$$

which relate the bit allocation, power, and bandwidth for each channel. We might have a total power constraint for the broadcast channel,

$$P_1 + P_2 + P_3 \le P_{\rm bc,tot},$$

and a separate total power constraint for the multiple access channel,

$$P_4 + P_5 + P_6 \le P_{\text{mac,tot}}.$$

The bandwidth allocations could also be separate for the multiple access and the broadcast channels, or, possibly, combined:

$$W_4 + W_5 + W_6 \le W_{\text{tot}}.$$

In this case we not only allocate bandwidth among the individual channels; we also allocate bandwidth between the broadcast and the multiple access channel. Finally, we might impose bounds on the number of bits allocated to each signal, as in

$$5 \le b_i \le 12, \qquad i = 1, \dots, 6.$$

# 4 Optimal resource allocation for fixed linear system

In this section, we assume that the linear system is fixed and consider the problem of choosing the communication variables to optimize the system performance. We take as the objective (to be minimized) the variance of the performance signal $z$, given by (3). Since this variance consists of a fixed term (related to $w$) and the variance induced by the quantization, we can just as well minimize the variance of $z$ induced by the quantization error, *i.e.*, the quantity $V_q$ defined in (4). This leads to the optimization problem

$$\begin{array}{lll} \text{minimize} & \sum_{i=1}^{M} a_i 2^{-2b_i} & \\ \text{subject to} & f_i(b, \theta) \le 0, & i = 1, \dots, m_f \\ & h_i^T \theta \le d_i, & i = 1, \dots, m_h \\ & \theta_i \ge 0, & i = 1, \dots, m_\theta \\ & \underline{b}_i \le b_i \le \overline{b}_i, & i = 1, \dots, M \end{array} \tag{11}$$

where the optimization variables are $\theta$ and $b$. For the moment we ignore the constraint that $b_i$ must be integers.

Since the objective function, and each constraint function in the problem (11) is a convex function, this is a convex optimization problem. This means that it can be solved globally and efficiently using a variety of methods, *e.g.*, interior-point methods (see, *e.g.*, [BV98]). In many cases, we can solve the problem (11) more efficiently than by applying general convex optimization methods by exploiting its special structure. This is explained in the next subsection.

## 4.1 The dual decomposition method

The objective function in the communication resource allocation problem (11) is separable, *i.e.*, a sum of functions of each $b_i$. In addition, the constraint functions $f_k(b, \theta)$ usually involve only one $b_i$, and a few components of $\theta$, since the channel capacity is determined by the bandwidth, power, or time-slot fraction, for example, allocated to that channel. In other words, the resource allocation problem (11) is almost separable; the small groups of variables (that relate to a given link or channel) are coupled mostly through the resource limit constraints $h_i^T\theta \le d_i$. These are the constraints that limit the total power, total bandwidth, or total time-slot fractions.

This almost separable structure can be efficiently exploited using a technique called dual decomposition (see, *e.g.*, [BV98, Ber99]). We will explain the method for a simple FDMA system to keep the notation simple, but the method applies to any communication resource allocation problem with almost separable structure. We consider an FDMA system with $M$ channels, and variables $P \in \mathbf{R}^M$ and $W \in \mathbf{R}^M$, with a total power and a total bandwidth constraint. We will also impose lower and upper bounds on the bits. This leads to

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^M a_i 2^{-2b_i} \\ \text{subject to} & b_i \le \alpha W_i \log_2(1+P_i/N_iW_i), \quad i=1,\dots,M \\ & P_i \ge 0, \quad i=1,\dots,M \\ & \sum_{i=1}^M P_i \le P_{\text{tot}} \\ & W_i \ge 0, \quad i=1,\dots,M \\ & \sum_{i=1}^M W_i \le W_{\text{tot}} \\ & \underline{b}_i \le b_i \le \overline{b}_i, \quad i=1,\dots,M. \end{array} \tag{12}$$

Here $N_i$ is the receiver noise spectral density of the $i$th channel, and $\underline{b}_i$ and $\overline{b}_i$ are the lower and upper bounds on the number of bits allocated to each channel. Except for the total power and total bandwidth constraint, the constraints are all local, *i.e.*, involve only $b_i$, $P_i$, and $W_i$.

We first form the Lagrange dual problem, by introducing Lagrange multipliers but *only* for the two coupling constraints. The Lagrangian has the form

$$L(b,P,W,\lambda,\mu) = \sum_{i=1}^M a_i 2^{-2b_i} + \lambda\left(\sum_{i=1}^M P_i - P_{\text{tot}}\right) + \mu\left(\sum_{i=1}^M W_i - W_{\text{tot}}\right).$$

The dual function is defined as

$$\begin{aligned} g(\lambda,\mu) &= \inf\left\{L \mid P_i \ge 0,\ W_i \ge 0,\ \underline{b}_i \le b_i \le \overline{b}_i,\ b_i \le \alpha W_i \log_2(1+P_i/N_iW_i)\right\} \\ &= \sum_{i=1}^M g_i(\lambda,\mu) - \lambda P_{\text{tot}} - \mu W_{\text{tot}} \end{aligned}$$

where

$$\begin{aligned} g_i(\lambda,\mu) = \inf\Big\{ & a_i 2^{-2b_i} + \lambda P_i + \mu W_i \ \Big| \\ & P_i \ge 0,\ W_i \ge 0,\ \underline{b}_i \le b_i \le \overline{b}_i,\ b_i \le \alpha W_i \log_2(1+P_i/N_iW_i)\Big\}. \end{aligned}$$

Finally, the Lagrange dual problem associated with the communication resource allocation problem (12) is given by

$$\begin{array}{ll} \text{maximize} & g(\lambda,\mu) \\ \text{subject to} & \lambda \ge 0, \quad \mu \ge 0. \end{array} \tag{13}$$

This problem has only two variables, namely the variables $\lambda$ and $\mu$ associated with the total power and bandwidth limits, respectively. It is a convex optimization problem, since $g$ is a concave function (see [BV98]). Assuming that Slater's condition holds, the optimal value of the dual problem (13) and the primal problem (12) are equal. Moreover, from the optimal solution of the dual problem, we can recover the optimal solution of the primal. Suppose $(\lambda^\star, \mu^\star)$ is the solution to the dual problem (13), then the primal optimal solution is the minimizer $(b^\star, P^\star, W^\star)$ when evaluating the dual function $g(\lambda^\star, \mu^\star)$. In other words, we can solve the original problem (12) by solving the dual problem (13).

The dual problem can be solved using a variety of methods, for example, cutting-plane methods. To use these methods we need to be able to evaluate the dual objective function, and also obtain a subgradient for it (see [BV98]), for any given $\mu \geq 0$ and $\lambda \geq 0$. To evaluate $g(\lambda, \mu)$, we simply solve the $M$ separate problems,

$$\begin{array}{ll} \text{minimize} & a_i 2^{-2b_i} + \lambda P_i + \mu W_i \\ \text{subject to} & P_i \geq 0, \ W_i \geq 0, \\ & \underline{b}_i \leq b_i \leq \overline{b}_i, \\ & b_i \leq \alpha W_i \log_2(1 + P_i/N_i W_i), \end{array}$$

each with three variables, which can be carried out separately or in parallel. Many methods can be used to very quickly solve these small problems.

A subgradient of the concave function $g$ at $(\lambda, \mu)$ is a vector $h \in \mathbf{R}^2$ such that

$$g(\tilde{\lambda}, \tilde{\mu}) \leq g(\lambda, \mu) + h^T \left[ \begin{array}{c} \tilde{\lambda} - \lambda \\ \tilde{\mu} - \mu \end{array} \right]$$

for all $\tilde{\lambda}$ and $\tilde{\mu}$. To find such a vector, let the optimal solution to the subproblems be denoted

$$b_i^\star(\lambda, \mu), \quad P_i^\star(\lambda, \mu), \quad W_i^\star(\lambda, \mu).$$

Then, a subgradient of the dual function $g$ is readily given by

$$\left[ \begin{array}{c} \sum_{i=1}^M P_i^\star(\lambda, \mu) - P_{\text{tot}} \\ \sum_{i=1}^M W_i^\star(\lambda, \mu) - W_{\text{tot}} \end{array} \right].$$

This can be verified from the definition of the dual function.

Putting it all together, we find that we can solve the dual problem in time linear in $M$, which is far better than the standard convex optimization methods applied to the primal problem, which require time proportional to $M^3$.

The same method can be applied whenever there are relatively few coupling constraints, and each link capacity is dependent on only a few communication resource parameters. In fact, when there is only one coupling constraint, the subproblems that we must solve can be solved analytically, and the master problem becomes an explicit convex optimization problem with only one variable. It is easily solved by bisection, or any other one-parameter search method. This is the famous water-filling algorithm (see, *e.g.*, [CT91]).

## 4.2 Integrality of bit allocations

We now come back to the requirement that the bit allocations must be integers. The first thing we observe is that we can always round down the bit allocations found by solving the convex problem to the nearest integers. Let $b_i$ denote the optimal solution of the convex resource allocation problem (11), and define $\tilde{b}_i = \lfloor b_i \rfloor$. Here, $\lfloor b_i \rfloor$ denotes the *floor* of $b_i$, *i.e.*, the largest integer smaller than or equal to $b_i$. First we claim that $\tilde{b}$ is feasible. To see this, recall that $f_k$ and $h_k$ are monotone decreasing in $b$, so since $b$ is feasible and $\tilde{b} \leq b$, we have $\tilde{b}$ feasible.

We can also obtain a crude performance bound for $\tilde{b}$. Clearly the objective value obtained by ignoring the integer constraint, *i.e.*,

$$J_{\text{cvx}} = \sum_{i=1}^{M} a_i 2^{-2b_i},$$

is a lower bound on the optimal objective value $J_{\text{opt}}$ of the problem with integer constraints. The objective value of the rounded-down feasible bit allocation $\tilde{b}$ is

$$J_{\text{rnd}} = \sum_{i=1}^{M} a_i 2^{-2\tilde{b}_i} \leq \sum_{i=1}^{M} a_i 2^{-2(b_i-1)} = 4J_{\text{cvx}} \leq 4J_{\text{opt}},$$

using the fact that $\tilde{b}_i \geq b_i - 1$. Putting this together we have

$$J_{\text{opt}} \leq J_{\text{rnd}} \leq 4J_{\text{opt}},$$

*i.e.*, the performance of the suboptimal integer allocation obtained by rounding down is never more than a factor of four worse than the optimal solution. In terms of RMS, the rounded-down allocation is never more than a factor of two suboptimal.

### Variable threshold rounding

Of course, far better heuristics can be used to obtain better integer solutions. Here we give a simple method based on a variable rounding threshold.

Let $0 < t \leq 1$ be a threshold parameter, and round $b_i$ as follows:

$$\tilde{b}_i = \begin{cases} \lfloor b_i \rfloor, & \text{if } b_i - \lfloor b_i \rfloor \leq t, \\ \lceil b_i \rceil, & \text{otherwise.} \end{cases} \tag{14}$$

Here, $\lceil b_i \rceil$ denotes the *ceiling* of $b_i$, *i.e.*, the smallest integer larger than or equal to $b_i$. In other words, we round $b_i$ down if its remainder is smaller than or equal to the threshold $t$, and round up otherwise. When $t = 1/2$, we have standard rounding, with ties broken down. When $t = 1$, all bits are rounded down, as in the scheme described before. This gives a feasible integer solution, which we showed above has a performance within a factor of four of optimal. For $t < 1$ feasibility of the rounded bits $\tilde{b}$ is not guaranteed, since bits can be rounded up.

For a given fixed threshold $t$, we can round the $b_i$'s as in (14), and then solve a convex feasibility problem over the remaining continuous variables $\theta$:

$$\begin{array}{rcl} f_i(\tilde{b}, \theta) & \leq & 0 \\ h_i^T \theta & \leq & d_i \\ \theta_i & \geq & 0 \end{array} \tag{15}$$

The upper and lower bound constraints $\underline{b}_i \leq \tilde{b}_i \leq \overline{b}_i$ are automatically satisfied because $\underline{b}_i$ and $\overline{b}_i$ are integers. If this problem is feasible, then the rounded $\tilde{b}_i$'s and the corresponding $\theta$ are suboptimal solutions to the integer constrained bit allocation problem.

Since $f_i$ is monotone increasing in $b$, hence in $t$, and monotone decreasing in $\theta$, there exists a $t^\star$ such that (15) is feasible if $t \geq t^\star$ and infeasible if $t < t^\star$. In the variable threshold rounding method, we find $t^\star$, the smallest $t$ which makes (15) feasible. This can be done by bisection over $t$: first try $t = 1/2$. If the resulting rounded bit allocation is feasible, we try $t = 1/4$; if not, we try $t = 3/4$, etc.

Roughly speaking, the threshold $t$ gives us a way to vary the conservativeness of the rounding procedure. When $t$ is near one, almost all bits are rounded down, and the allocation is likely to be feasible. When $t$ is small, we round many bits up, and the bit allocation is unlikely to be feasible. But if it is, the performance (judged by the objective) will be better than the bit allocation found using more conservative rounding (*i.e.*, with a larger $t$). A simple bisection procedure can be used to find a rounding threshold close to the aggressive one that yields a feasible allocation.

## 4.3 Example: networked linear estimator

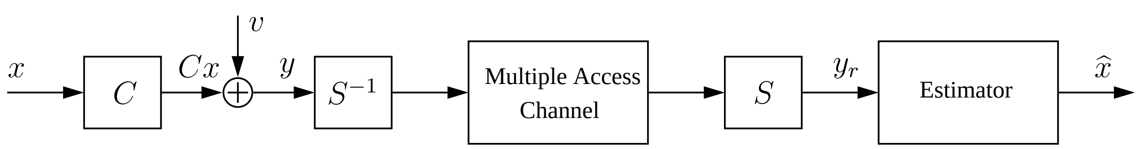


**Figure 5:** Networked linear estimator over a multiple access channel

To illustrate the idea of this section, we consider the problem of designing a networked linear estimator. We want to estimate an unknown point $x \in \mathbf{R}^{20}$, where we assume that $\|x\| \leq 1$. There are $M = 200$ linear sensors,

$$y_i = c_i^T x + v_i, \qquad i = 1, \ldots, M.$$

The sensor noises $v_i$ are IID (independent identically distributed) with covariance matrix $R_v = 10^{-6} I$, where $I$ is the identity matrix. The sensor coefficients $c_i$ for this example were chosen as uniformly distributed in direction, with $\|c_i\|$ uniformly distributed on $[0, 5]$.

By our assumption that $\|x\| \leq 1$, the range of the $i$th sensor is (ignoring the noise) $\pm\|c_i\|$. Based on this, we choose the scaling factors $s_i = \|c_i\|$.

Each sensor transmits its measurement to a central estimator through a Gaussian multiple access channel (as in figure 5). The $i$th sensor is assigned $b_i$ bits to code the measurement, so the received signals are

$$y_{ri} = c_i^T x + v_i + q_i, \qquad i = 1, \dots, M.$$

The covariance matrix of the quantization noise $q$ is given by

$$R_q = \mathbf{diag}\left(\frac{s_1^2}{3}2^{-2b_1}, \dots, \frac{s_M^2}{3}2^{-2b_M}\right). \tag{16}$$

The estimate of $x$ is formed using a linear unbiased estimator, *i.e.*,

$$\widehat{x} = K y_r,$$

where $KC = I$, with $C = [c_1, \cdots, c_M]^T$. For example, the minimum variance estimator is given by

$$K = (C^T Q C)^{-1} C^T Q \tag{17}$$

where the weighting matrix $Q$ is given by

$$Q = (R_v + R_q)^{-1}.$$

(Note that the optimal estimator depends on the bit allocations; we will address this issue in the next section.)

The performance of the estimator is evaluated by the estimation error variance:

$$J_K(b) = \mathbf{E}\,\|\widehat{x} - x\|^2 = \mathbf{Tr}(K(R_v + R_q)K^T) = \frac{1}{3}\sum_{i=1}^{M} s_i^2 \|k_i\|^2 2^{-2b_i} + \mathbf{Tr}(K R_v K^T).$$

where $k_i$ is the $i$th column of the matrix $K$. Clearly, $J_K(b)$ is in the form of (3), and will serve as the objective function for the optimization problem (11).

In this example, the noise power density of the Gaussian multiple access channel is $N = 0.1$, the total available power is $P = 300$, and the total available bandwidth is $W = 200$ (using FDMA). The coding constant is $\alpha = 2$, and the common lower and upper bounds for all bits are $\underline{b} = 5$ and $\overline{b} = 12$.

| RMS values | equal allocation | relaxed optimization | variable threshold rounding |
|---|---|---|---|
| $\mathbf{rms}(x - \widehat{x})$ | $3.6760 \times 10^{-3}$ | $3.1438 \times 10^{-3}$ | $3.2916 \times 10^{-3}$ |

**Table 1:** RMS estimation errors of the networked linear estimator.

First we allocate power and bandwidth evenly to all sensors, which results in $b_i = 8$ for each sensor. Based on this allocation, we compute the covariance matrix $R_q$ as in (16) and design a least-squares estimator as in (17). The RMS estimation error of this estimator is shown in Table 1. Then we fix the estimator gain $K$, and find the optimal resource allocation to minimize the estimation error variance by solving the relaxed optimization problem (12), and then performing variable threshold rounding, which gives $t^\star = 0.4211$ in this example. Figure 6 shows the distribution of rounded bit allocation. The RMS estimation errors are listed in Table 1.

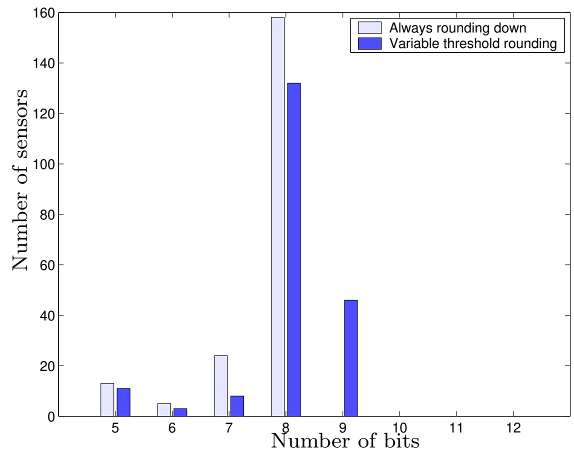


**Figure 6:** Bit allocation for networked least-squares estimator.

We can see that the allocation obtained from optimization and variable threshold rounding give a 10% improved performance compared to the uniform resource allocation, and is not very far from the performance bound given by the relaxed convex optimization problem.

Note that with the new bit allocations, the quantization covariance changes — it is not the one that was used to design $K$. In the next section, we address this issue of the coupling between the choice of the communication variables and the estimator.

# 5 Joint optimization of communication and linear system variables

We have seen that when the linear system is fixed, the problem of optimally allocating communication resources is convex (when we ignore integrality of bit allocations), and can be efficiently solved. In order to achieve the optimal system performance, however, one should optimize the parameters of the linear system *and* the communication system *jointly*. Unfortunately, this joint design problem is in general not convex. In some cases, however, the joint design problem is bi-convex: for fixed resource allocation the controller design problem is convex, and for fixed controller design and scalings the resource allocation problem is convex. This special structure can be exploited to develop a heuristic method for the joint design problem, that appears to work well in practice.

## 5.1 Nonconvexity of the joint design problem

To illustrate that the joint design problem is nonconvex, we consider the problem of designing a simple networked least-squares estimator for an example small enough that we can solve the joint problem globally.

An unknown scalar parameter $x \in \mathbf{R}$ is measured using two sensors that are subject to measurement noises:

$$y_1 = x + v_1, \qquad y_2 = x + v_2.$$

We assume that $v_1$ and $v_2$ are independent zero-mean Gaussian random variables with variances $\mathbf{E}\, v_1^2 = \mathbf{E}\, v_2^2 = 0.001$. The sensor measurements are coded and sent over a communication channel with a constraint on the total bit rate. With a total of $b_{\rm tot}$ bits available we allocate $b_1$ bits to the first sensor and the $b_2 = b_{\rm tot} - b_1$ remaining bits to the second sensor. For a given bit allocation, the minimum-variance unbiased estimate can be found by solving a weighted least-squares problem. Figure 7 (left) shows the optimal performance as function of $b_1$ when $b_{\rm tot} = 8$. The relationship is clearly not convex.

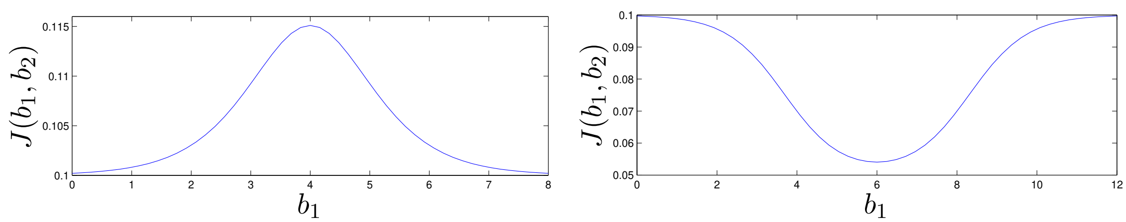


**Figure 7:** Performance as function of bit-allocation for $b_1 + b_2 = 8$ (left) and $b_1 + b_2 = 12$ (right). The nonconvexity is apparent in the left figure.

These figures, and the optimal solutions, make perfect sense. When $b_{\rm tot} = 8$, the quantization noise is the dominant noise source, so one should allocate all 8 bits to one sensor and disregard the other. When $b_{\rm tot} = 12$, the quantization noises are negligible in comparison with the sensor noise. It is then advantageous to use both sensors (*i.e.*, assign each one 6 bits), since it allows us to average out the effect of the measurement noises.

## 5.2 Alternating optimization for joint design

The fact that the joint problem is convex in certain subsets of the variables while others are fixed can be exploited. For example (and ignoring the integrality constraints) the globally optimal communication variables can be computed very efficiently, sometimes even semi-analytically, when the linear system is fixed. Similarly, when the communication variables are fixed, we can (sometimes) compute the globally optimal variables for the linear system. Finally, when the linear system variables and the communication variables are fixed, it is straightforward to compute the quantizer scalings using the $3\sigma$-rule. This makes it natural to apply an approach where we sequentially fix one set of variables and optimize over the others:

**given** initial linear system variables $\phi^{(0)}$, communication variables $\theta^{(0)}$, scalings $s^{(0)}$
$k = 0$
**repeat**

1. Fix $\phi^{(k)}$, $s^{(k)}$, and optimize over $\theta$. Let $\theta^{(k+1)}$ be the optimal value.
2. Fix $\theta^{(k+1)}$, $s^{(k)}$, and optimize over $\phi$. Let $\phi^{(k+1)}$ be the optimal value.
3. Fix $\phi^{(k+1)}$, $\theta^{(k+1)}$. Let $s^{(k+1)}$ be appropriate scaling factors.

**until** convergence

Many variations on this basic heuristic method are possible. We can, for example, add trust region constraints to each of the optimization steps, to limit the variables changes in each step. Another variation is to convexify (by, for example, linearizing) the jointly nonconvex problem, and solve in each step using linearized versions for the constraints and objective terms in the remaining variables; see, *e.g.*, [HHB99] and the references therein. . We have already seen how the optimization over $\theta$ can be carried out efficiently. In many cases, the optimization over $\phi$ can also be carried efficiently, using, *e.g.*, LQG or some other controller or estimator design technique.

Since the joint problem is not convex, there is no guarantee that this heuristic converges to the global optimum. On the other hand the heuristic method appears to work well in practice.

### 5.3 Example: networked linear estimator

To demonstrate the heuristic method for joint optimization described above, we apply it to the networked linear estimator described in §4.3. The design of the linear system and the communication system couple through the weighting matrix $Q$ in (17). The alternating procedure for this problem becomes

> **given** initial estimator gain $K^{(0)}$ and resource allocations $(P^{(0)}, W^{(0)}, b^{(0)})$
> **repeat**
> 1. Fix estimator gain $K^{(k)}$ and solve the problem (12) to obtain resource allocation $(P^{(k+1)}, W^{(k+1)}, b^{(k+1)})$.
> 2. Update the covariance matrix $R_q^{(k+1)}$ using (16) and compute new estimator gain $K^{(k+1)}$ as in (17) using the weighting matrix $Q^{(k+1)} = (R_v + R_q^{(k+1)})^{-1}$.
>
> **until** bit allocation converges.

Note that the scaling factors are fixed in this example, since neither the bit allocations nor the estimator gain affect the signals that are quantized, hence the scaling factors.

When we apply the alternating optimization procedure to the example from §4.3, the algorithm converges in six iterations, and we obtain very different resource allocation results from before. Figure 8 shows the distribution of rounded bit allocation. This result is intuitive: try to assign as much resources as possible to the best sensors, and the bad sensors only get minimum number of bits. The RMS estimation error of the joint design is reduced significantly, 80%, as shown in Table 2. In this table, $\mathbf{rms}(e)$ is the total RMS error, $\mathbf{rms}(e_q)$ is the RMS error induced by quantization noise, and $\mathbf{rms}(e_v)$ is the RMS error induced by sensor noise.

| RMS values | equal allocation | joint optimization | variable threshold rounding |
|---|---|---|---|
| $\mathbf{rms}(e_q)$ | $3.5193 \times 10^{-3}$ | $0.3471 \times 10^{-3}$ | $0.3494 \times 10^{-3}$ |
| $\mathbf{rms}(e_v)$ | $1.0617 \times 10^{-3}$ | $0.6319 \times 10^{-3}$ | $0.6319 \times 10^{-3}$ |
| $\mathbf{rms}(e)$ | $3.6760 \times 10^{-3}$ | $0.7210 \times 10^{-3}$ | $0.7221 \times 10^{-3}$ |

**Table 2:** RMS estimation errors of the networked LS estimator.

We can see that joint optimization reduces the estimation errors due to both quantization and sensor noise. In the case of equal resource allocation, the RMS error due to quantization is much larger than that due to sensor noise. After the final iteration of the alternating convex optimization, the RMS error due to quantization is at the same level as that due to sensor noise. Also, because the in the relaxed problem, most bits are integers (either $\underline{b} = 5$ or $\overline{b} = 12$; see Figure 8), variable threshold rounding (which gives $t^\star = 0.6797$) does not change the solution, or the performance, much.

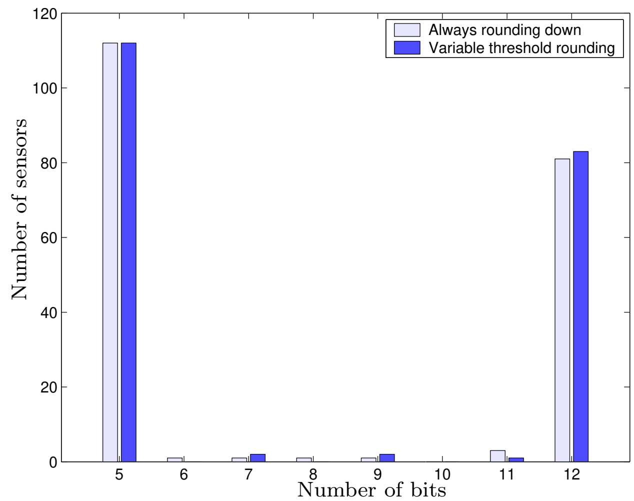


**Figure 8:** Joint optimization of bit allocation and least-squares estimator

## 5.4 Example: LQG control over communication networks

We now give a more complex example than the simple static, open-loop estimator described above. The situation is more complicated when the linear system is dynamic and involves feedback loops closed over the communication links. In this case, the RMS values of both control signals and output signals change when we re-allocate communication resources or adjust the controller. Hence, the alternating optimization procedure needs to include the step that modifies the scalings.

### Basic system setup

First we consider the system setup in figure 9, where no communication links are included. The linear system has a state-space model

$$\begin{aligned} x(t+1) &= Ax(t) + B\left(u(t) + w(t)\right) \\ y(t) &= Cx(t) + v(t) \end{aligned}$$

where $u(t) \in \mathbf{R}^{M_u}$ and $y(t) \in \mathbf{R}^{M_y}$. Here $w(t)$ is the process noise and $v(t)$ is the sensor noise. Assume that $w(t)$ and $v(t)$ are independent zero-mean white noises with covariance matrices $R_w$ and $R_v$ respectively.

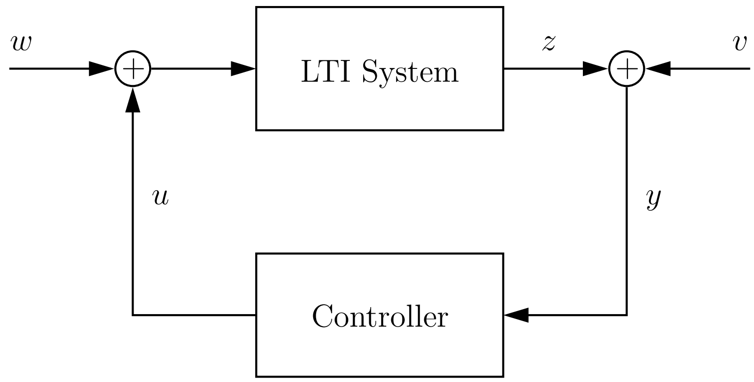


**Figure 9:** Closed-loop control system without communication links.

Our goal is to design the controller that minimizes the RMS value of $z = Cx$, subject to some upper bound constraints on the RMS values of the control signals:

$$\begin{array}{ll} \text{minimize} & \mathbf{rms}(z) \\ \text{subject to} & \mathbf{rms}(u_i) \leq \beta_i, \quad i = 1, \ldots, M_u \end{array} \tag{18}$$

The limitations on the RMS values of the control signals are added to avoid actuator saturation.

It can be shown that the optimal controller for this problem has the standard estimated state feedback form,

$$\begin{aligned} \widehat{x}(t+1|t) &= A\widehat{x}(t|t-1) + Bu(t) + L\left(y(t) - C\widehat{x}(t|t-1)\right) \\ u(t) &= -K\widehat{x}(t|t-1) \end{aligned}$$

where $K$ is the state feedback control gain and $L$ is the estimator gain, found by solving the algebraic Riccati equations associated with an appropriately weighted LQG problem. Finding the appropriate weights, for which the LQG controller solves the problem (18), can be done via the dual problem; see, *e.g.*, [TM89, BB91].

### Communications setup

We now describe the communications setup for the example. The sensors send their measurements to a central controller through a Gaussian multiple access channel, and the controller sends control signals to the actuators through a Gaussian broadcast channel, as shown in figure 10.

The linear system can be described as

$$\begin{aligned} x(t+1) &= Ax(t) + B\left(u(t) + w(t) + p(t)\right) \\ y_r(t) &= Cx(t) + v(t) + q(t), \end{aligned}$$

where $p$ and $q$ are quantization noises due to the bit rate limitations of the communication channels. Since these are modeled as white noises, we can include the quantization noises

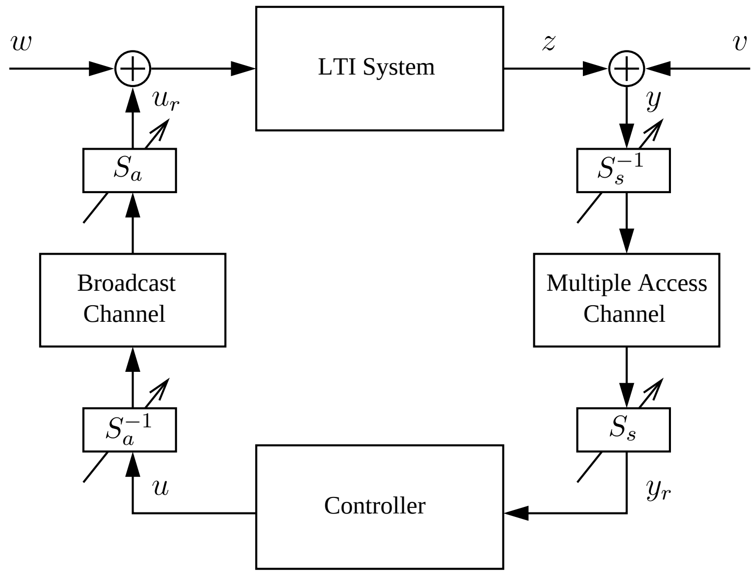


**Figure 10:** Closed-loop control system over communication networks.

in the process and measurement noises, by introducing the equivalent process noise and measurement noise

$$\widetilde{w}(t) = w(t) + p(t), \qquad \widetilde{v}(t) = v(t) + q(t),$$

with covariance matrices

$$\begin{aligned} R_{\widetilde{w}} &= R_w + \mathbf{diag}\left(\frac{s_{a,1}^2}{3}2^{-2b_{a,1}}, \dots, \frac{s_{a,M_u}^2}{3}2^{-2b_{a,M_u}}\right), \\ R_{\widetilde{v}} &= R_v + \mathbf{diag}\left(\frac{s_{s,1}^2}{3}2^{-2b_{s,1}}, \dots, \frac{s_{s,M_y}^2}{3}2^{-2b_{s,M_y}}\right). \end{aligned} \tag{19}$$

Here $b_a$ and $b_s$ are number of bits allocated to the actuators and sensors.

The scaling factors can be found from the $3\sigma$-rule, by computing the variance of the sensor and actuator signals. Hence, given the signal ranges and numbers of quantization bits, we can calculate $R_{\widetilde{w}}$ and $R_{\widetilde{v}}$, and then design a controller by solving (18). Notice that the signal ranges are determined by the RMS values, which in turn depend on the controller design. This intertwined relationship will show up in the iterative design procedures.

### Iterative procedure to design a controller with uniform bit allocation

First we allocate an equal number of bits to each actuator and sensor. This means that we assign power and bandwidth (in the case of FDMA) uniformly across all channels. We design a controller for such uniform resource allocation via the following iterative procedure (iterate on the scaling factors and the controller):

**given** $\beta_i = \mathbf{rms}(u_i)$ and estimated $\mathbf{rms}(z_j)$.
**repeat**
1. Let $s_{a,i} = 3\,\mathbf{rms}(u_i)$ and $s_{s,j} = 3\,\mathbf{rms}(z_j)$, and compute $R_{\tilde{w}}$ and $R_{\tilde{v}}$ as in (19).
2. Solve problem (18) and compute $\mathbf{rms}(u_i)$ and $\mathbf{rms}(z_j)$ of the closed-loop system.

**until** stopping criterion is satisfied.

If the procedure converges, the resulting controller variables $K$ and $L$ of this iterative design procedure will satisfy the constraints on the control signals.

### The alternating optimization procedure

Our goal here is to do joint optimization of bit allocation and controller design. This involves an iteration procedure over controller design, scaling matrices update and bit allocation. The controller and scaling matrices designed for uniform bit allocation by the above iteration procedure can serve as a good starting point. Here is the alternating optimization procedure:

**given** $R_w, R_v$, $\beta_i = \mathbf{rms}(u_i)$ and $\mathbf{rms}(z_j)$ from the above iteration design procedure.
**repeat**
1. Allocate bit rates $b_{a,i}$, $b_{s,j}$ and communication resources by solving a convex optimization problem of the form (11).
2. Compute $R_{\tilde{w}}$ and $R_{\tilde{v}}$ as in (19), and find controller variables $K$ and $L$ by solving (18).
3. Compute closed-loop system RMS values $\mathbf{rms}(u_i)$ and $\mathbf{rms}(z_j)$, then determine the signal ranges $s_{a,i}$ and $s_{s,j}$ by the $3\sigma$ rule.

**until** the RMS values $\mathbf{rms}(z_j)$ and bit allocation converges.

The convex optimization problem to be solved in step 1 depends on the communication system setup and resource constraints.

### Numerical example: control of a mass-spring system

Now we consider the specific example shown in figure 11. The position sensors on each mass send measurements $y_i = x_i + v_i$, where $v_i$ is the sensor noise, to the controller through a Gaussian multiple access channel using FDMA. The controller receives data $y_{ri} = x_i + v_i + q_i$, where $q_i$ is the quantization error due to bit rate limitation of the multiple access channel. The controller sends control signals $u_j$ to actuators on each mass through a Gaussian broadcast channel using FDMA. The actual force acting on each mass is $u_{rj} = u_j + w_j + p_j$, where $w_j$ is the exogenous disturbance force, and $p_j$ is the quantization disturbance due to bit rate limitation of the broadcast channel. The mechanical system parameters are

$$m_1 = 10, \quad m_2 = 5, \quad m_3 = 20, \quad m_4 = 2, \quad m_5 = 15, \quad k = 1$$

The discrete-time system dynamics is obtained using a sampling frequency which is 5 times faster than the fastest mode of the continuous-time dynamics. The independent zero mean noises $w$ and $v$ have covariance matrices $R_w = 10^{-6}I$ and $R_v = 10^{-6}I$ respectively. The actuators impose RMS constraints on the control signals:

$$\mathbf{rms}(u_i) \leq 1, \quad i = 1, \ldots, 5.$$

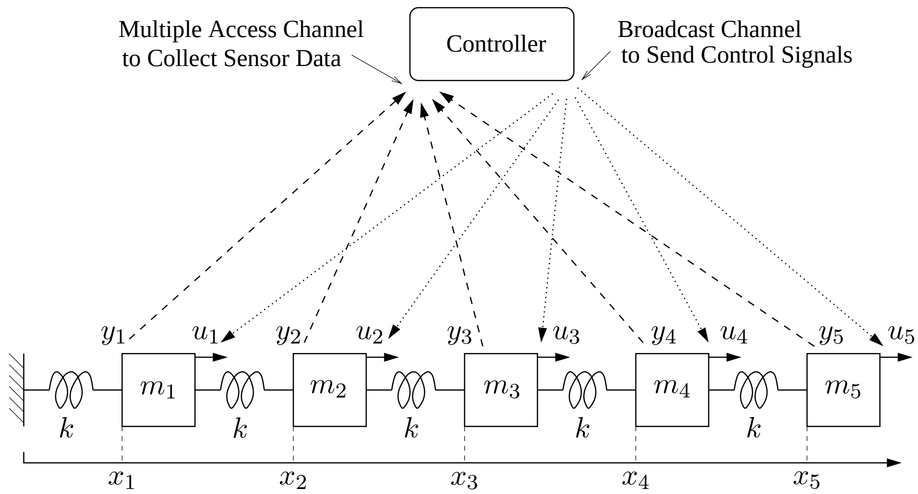


**Figure 11:** Series-connected mass-spring system controlled over network.

For the Gaussian multiple access channel, the noise power density is $N = 0.1$, and the total power available is $P_{\mathrm{mac,tot}} = 7.5$. For the Gaussian broadcast channel, the noise power density at each user is $N_i = 0.1$ for all $i$'s, and the total power available for all users is $P_{\mathrm{bc,tot}} = 7.5$. All users of the multiple access channel and the broadcast channel share a total bandwidth of $W = 10$. Other constants are $\alpha = 2$, $\underline{b} = 5$ and $\overline{b} = 12$.

| RMS values | equal allocation | joint optimization | variable threshold rounding |
|---|---|---|---|
| $\mathbf{rms}(z_1)$ | 0.1487 | 0.0424 | 0.0438 |
| $\mathbf{rms}(z_2)$ | 0.2602 | 0.0538 | 0.0535 |
| $\mathbf{rms}(z_3)$ | 0.0824 | 0.0367 | 0.0447 |
| $\mathbf{rms}(z_4)$ | 0.4396 | 0.0761 | 0.0880 |
| $\mathbf{rms}(z_5)$ | 0.1089 | 0.0389 | 0.0346 |
| $\mathbf{rms}(z)$ | 0.5493 | 0.1155 | 0.1258 |

**Table 3:** RMS-values of the output signal.

First we allocate power and bandwidth evenly to all sensors and actuators, which results in a uniform allocation of 8 bits for each channel. We then designed a controller using the first iteration procedure based on this uniform resource allocation. This controller yields $\mathbf{rms}(u_i) = 1$ for all $i$'s, and the RMS-values of the output signal $z$ are listed in Table 3.

Finally, we used the second iteration procedure to do joint optimization of bit allocation and controller design. The resulting resource allocation after four iterations is shown in figure 12. It can be seen that more bandwidth, and hence more bits are allocated to the broadcast channel than to the multiple access channel. This means that the closed-loop performance is more sensitive to the equivalent process noises than to the equivalent sensor noises. The joint optimization resulted in $\mathbf{rms}(u_i) = 1$ for all $i$'s, and the RMS-values of the output signal $z$ are listed in Table 3. At each step of the variable threshold rounding, we check the feasibility of the resource allocation problem. The optimal threshold found is $t^\star = 0.6150$. Then we fix the integer bit allocation obtained with this threshold, and used the

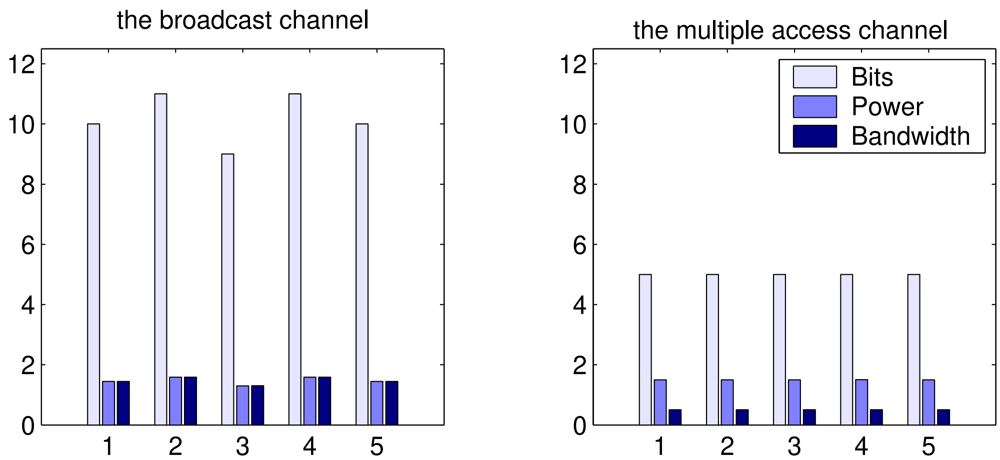


**Figure 12:** Joint optimization of bit rates and linear control system.

first iteration procedure to design the controller. We see a 77% reduction in RMS value over the result for uniform bit allocation, and the performance obtained by variable threshold rounding is quite close to that of the relaxed non-integer joint optimization.

# 6 Conclusions and extensions

We have considered the problem of jointly optimizing the parameters of a linear system and the associated communication system. A typical application is the simultaneous design of networked controllers and estimators, and optimal resource allocation for the communication system that supports them.

Many issues arise in the design of controllers and estimators for use in a networked environment. These issues include bit rate limitations, communication delays, packet loss, transmission errors, and asynchronous operation of subsystems. Our focus has been on the optimization of the underlying communication resources when the communication channel has a rate constraint. To model the influence of rate allocations on the performance of the linear system, we assumed conventional uniform quantization and used a simple white noise model of the quantization errors.

We have considered two specific problems. In the first problem, we assumed the linear system to be fixed and considered the problem of choosing the communication variables to optimize the overall system performance. We observed that this problem is often convex (ignoring the integrality constraint) hence readily solved. Moreover, for many important channel models, the communication resource allocation problem is separable except for a small number of constraints on the total communication resources. We showed how dual decomposition can be used to solve this class of problems very efficiently. The approach was demonstrated on an example, in which we allocated powers, bandwidths, and communication rates for a networked linear estimator. We also give a variable threshold rounding method to deal with the integrality of bit allocations.

In the second problem, we considered joint allocation of communication resources and design of the linear system in order to optimize the performance. We showed that this problem is in general not convex. However, it is often convex in subsets of variables while others are fixed. We gave an iterative heuristic method for the joint design problem that exploits this special structure. Applied to the networked linear estimator, the joint optimization method re-allocated resources and greatly improved the performance. We also gave an example of LQG control over communication networks, where modifying the scalings of transmitted signals must be included in the iterative design procedure.

There are many possible extensions of the work in this paper. Notice that the objective function of the bit allocation problem in §4 has the same form as the distortion-rate function of Gaussian sources in rate distortion theory (see, *e.g.*, [CT91]), so the same optimization model and solution techniques can be applied to minimize the total distortion of Gaussian signals transmitted over a group of communication channels with resource limited capacities. Another extension of the resource allocation problem is to simultaneously optimize routing tables and resource allocation in a wireless communication network.

# Acknowledgements

The authors are grateful to Wei Yu and Xiangheng Liu for helpful discussions.

# References

[BB91] S. Boyd and C. Barratt. *Linear Controller Design: Limits of Performance.* Prentice-Hall, 1991.

[Ber99] D. P. Bertsekas. *Nonlinear Programming.* Athena Scientific, second edition, 1999.

[BV98] S. P. Boyd and L. Vandenberghe. *Course reader for EE364: Introduction to Convex Optimization with Engineering Applications.* Stanford University, 1998.

[CG] S. T. Chung and A. J. Goldsmith. Degrees of freedom in adaptive modulation: A unified view. to appear in IEEE Transactions on Communications, 2001.

[CT91] T. Cover and J. Thomas. *Elements of Information Theory.* John Wiley & Sons, 1991.

[FPW90] G. F. Franklin, J. D. Powell, and M. L. Workman. *Digital Control of Dynamic Systems.* Addison Wesley, 3rd edition, 1990.

[Gol99] A. Goldsmith. *Course reader for EE359: Wireless Communications.* Stanford University, 1999.

[HHB99] A. Hassibi, J. P. How, and S. P. Boyd. A path-following method for solving BMI problems in control. In *Proceedings of American Control Conference*, volume 2, pages 1385–9, June 1999.

[LG01] L. Li and A. J. Goldsmith. Capacity and optimal resource allocation for fading broadcast channels: Part I: Ergodic capacity. *IEEE Transactions on Information Theory*, 47(3):1103–1127, March 2001.

[NBW98] J. Nilsson, B. Bernhardsson, and B. Wittenmark. Stochastic analysis and control of real-time systems with random time delays. *Automatica*, 34(1):57–64, 1998.

[NE98] G. N. Nair and R. J. Evans. State estimation under bit-rate constraints. In *Proc. IEEE Conference on Decision and Control*, pages 251–256, Tampa, Florida, 1998.

[Özg89] Ü. Özgüner. Decentralized and distributed control approaches and algorithms. In *Proceedings of the 28th Conference on Decision and Control*, pages 1289–1294, Tampa, Florida, December 1989.

[SSK99] K. Shoarinejad, J. L. Speyer, and I. Kanellakopoulos. An asymptotic optimal design for a decentralized system with noisy communication. In *Proceedinngs of the 38th Conference on Decision and Control*, Phoenix, Arizona, December 1999.

[SW90] R. E. Skelton and D. Williamson. Guaranteed state estimation accuracies with roundoff error. In *Proceedings of the 29th Conference on Decision and Control*, pages 297–298, Honolulu, Hawaii, December 1990.

[TH98] D. N. C. Tse and S. V. Hanly. Multiaccess fading channels — part I: Polymatroid structure, optimal resource allocation and throughput capacities. *IEEE Transactions on Information Theory*, 44(7):2796–2815, November 1998.

[TM89] H. T. Toivonen and P. M. Mäkilä. Computer-aided design procedure for multiobjective LQG control problems. *Int. J. Control*, 49(2):655–666, February 1989.

[WB97] W. S. Wong and R. W. Brockett. Systems with finite communication bandwidth constraints I: state estimation problems. *IEEE Transactions on Automatic Control*, 42:1294–1299, 1997.

[Wil85] D. Williamson. Finite wordlength design of digital Kalman filters for state estimation. *IEEE Transactions on Automatic Control*, 30(10):930–939, October 1985.

[WK89] D. Williamson and K. Kadiman. Optimal finite wordlength linear quadratic regulation. *IEEE Transactions on Automatic Control*, 34(12):1218–1228, December 1989.

[XHH00] L. Xiao, A. Hassibi, and J. P. How. Control with random communication delays via a discrete-time jump system approach. In *Proc. American Control Conf.*, Chicago, IL, June 2000.