Preprint of chapter in Encyclopedia of Systems and Control, Springer, 2015, pp. 1115-1121.
Latest version available at link.springer.com

Quantized Control and Data-Rate Constraints

Girish N. Nair

Abstract This article briefly describes the topic of quantized control with limited data rates. The focus is on the problem of stabilizing a linear time-invariant plant over a digital channel, and the associated *data rate theorems*. It is shown that the deepest results in this area require a unified treatment of its communications and control aspects.

Keywords quantised control · quantization · control under communication constraints

Introduction

One of the standard assumptions of classical control theory is that the signals sent from sensors to controllers and from controllers to actuators take continuous values with infinite precision. The advent of computer-based and digitally networked control systems challenged this assumption, since the analog plant outputs or control variables in such systems must be reduced to finite bit-strings or discrete symbols for storage, manipulation and transmission. This process of converting a continuous-valued variable into a finite-valued one is called *quantization*, and entails a potentially significant loss of resolution and closed-loop performance. Quantized control is concerned with the analysis and design of control systems which feature such analog-to-digital conversions in the feedback loop.

There is a vast literature on this topic and the aim of this article is to briefly explain some of its key ideas. For reasons of space, the discussion is largely confined to the question of how to stabilize a linear time-invariant plant over a digital channel. It is shown that the deepest results here emerge from treating the communications and control aspects jointly, instead of separately. The reader is referred to the survey (Nair et al, 2007) and the references therein for a discussion of other issues such as optimality and transient performance.

Supported by Australian Research Council grant DP110102401.

Dept. Electrical & Electronic Engineering, University of Melbourne, VIC 3010, Australia, E-mail: gnair@unimelb.edu.au

Quantization

Quantization has long been an object of study in communications and information theory - see (Gersho and Gray, 1993) and the references therein. In its simplest form, a signal $x(\cdot) : \mathbb{R} \to \mathbb{R}^n$ is first *sampled* at regular time intervals $t = 0, \tau, 2\tau, \dots$ to yield a discrete-time signal $(x(k\tau))_{k\in\mathbb{Z}}$, with the sampling frequency $1/\tau$ chosen to be greater than the Nyquist frequency of x (i.e. twice its bandwidth). Each sample $x_k := x(k\tau)$ is then passed through a static, memoryless quantizer Q to yield a quantized discrete-time signal

$$x_k^q = Q(x_k) \in \{q^1, \dots q^M\} \subset \mathbb{R}^n, \;\; k \in \mathbb{Z}_{\geq 0}. \quad (1)$$

which can take M distinct values in $\mathbb{R}^n$. If the quantizer is known to both transmitter and receiver, each of these M values can be represented by a binary string with $\lceil \log_2 M \rceil$ bits. When the input dimension $n = 1$, the quantizer is called scalar; otherwise it is a vector quantizer. The regions $R^i := Q^{-1}(q^i)$, $1 \leq i \leq M$, are called the quantizer cells and together form a partition of $\mathbb{R}^n$. Thus an M-valued quantizer is fully defined by its quantizer cells R^i and associated quantizer points q^i, $1 \leq i \leq M$.

The quantization error or *quantizer noise* is defined as $n_k := x_k^q - x_k$. When the inputs x_k are identically distributed random variables, then a standard goal is to design Q so as to minimize the mean-square quantizer noise

$$D := \mathrm{E}[\|Q(x_k) - x_k\|^2], \quad (2)$$

where $\mathrm{E}[\cdot]$ is the expectation functional. This yields an optimal quantizer Q_* with cells that satisfy the nearest-neighbour property, i.e.

$$x \in R_*^i \Rightarrow \|Q_*(x) - q_*^i\| \leq \|Q_*(x) - q_*^j\|, \;\; \forall j \neq i.$$

When $\|\cdot\|$ is the Euclidean norm (possibly weighted), the quantizer cells R_*^i, $1 \le i \le m$, are convex polygons and form a *Voronoi partition* of $\mathbb{R}^n$, and furthermore q_*^i is the centroid of R_*^i with respect to the stationary distribution F_X of x_k, i.e. $q_*^i = \mathrm{E}[x_k | x_k \in R_*^i]$. As a consequence, the optimal quantizer is statistically unbiased, i.e. $\mathrm{E}[n_k] = 0$, and furthermore x_k and the quantizer noise n_k are uncorrelated at time k, i.e. $\mathrm{E}[x_k n_k^{\mathrm{T}}] = 0$. However, note that n_k and x_j may be correlated for $j \neq k$, and (n_k) may itself be a correlated process.

If Q is not optimal but M is large (i.e. the quantizer is *high resolution* or *fine*), then $\mathrm{E}[x_k n_k^{\mathrm{T}}] = \mathrm{o}(1/M)$, provided that q^i is the centroid of R^i with respect to Lebesgue measure μ and x_k has a probability density function (pdf) f_X with suitable continuity properties. The reasoning here is that each region R_i will typically be very small, so that f_X will not vary much on each R^i, yielding a conditional pdf of x_k given R_i that is approximately uniform on R_i.

When Q is a scalar uniform quantizer on an interval $[a,b]$, these considerations yield the asymptotic formula

$$D \approx (b-a)^2/(12M^2), \tag{3}$$

provided that the overload regions - i.e. the tails of $f_X(x)$ on the regions $x < a$ or $x > b$ - make negligible contributions to D. Note that this expression does not depend on the distribution of the input. For large M, it can be shown that the optimal vector quantizer has a normalized point density proportional to $f_X^{1/3}$ and yields

$$D_{\min} \approx \frac{c}{M^2}\left(\int f_X(x)^{1/3}\mathrm{d}\mu(x)\right)^3, \tag{4}$$

where the constant c depends only on n.

Quantized Control - Basic Formulation

Much of the theory of quantized control concerns finite-dimensional linear time-invariant (LTI) plants. A formulation is provided in this section to help fix ideas, for the case of a single feedback loop containing a single errorless digital channel.

Consider the discrete-time plant

$$x_{k+1} = Ax_k + Bu_k + v_k, \;\; y_k = Fx_k + w_k, \tag{5}$$

where at every time $k \in \mathbb{Z}_{\ge 0}$, $x_k \in \mathbb{R}^n$ is the state with x_0 unknown, $u_k \in \mathbb{R}^m$ is the control input, $y_k \in \mathbb{R}^p$ is the measured output, $v_k \in \mathbb{R}^n$ is unknown process noise, $w_k \in \mathbb{R}^p$ is unknown measurement noise, and A, B, and F are constant known matrices of appropriate dimensions. For the problem to be well-posed, assume that the matrix pairs (A,B) and (F,A) are respectively reachable and observable. Suppose that the output sensors communicate with the controller over a digital channel that can carry one symbol s_k from a finite, possibly time-varying alphabet $\mathscr{S}_k$ of cardinality $M_k \ge 1$ during the $(k+1)$-th sampling interval. Assume for simplicity that the channel is errorless, with negligible propagation delay. The asymptotic average rate at which the channel transports data may then be defined as

$$R := \varliminf_{k\to\infty} \frac{1}{k}\sum_{j=0}^{k-1} \log_2 M_j \text{ (bits/sample)}. \tag{6}$$

Note that if the channel alphabet $\mathscr{S}_k$ is constant or varies periodically with k, the inferior limit reduces to a straight limit.

In full generality, each transmitted symbol may depend on all past and present measurements and past symbols,

$$s_k = \gamma_k(y_0^k, s_0^{k-1}) \in \mathscr{S}, \quad \forall k \in \mathbb{Z}_{\ge 0}, \tag{7}$$

where γ_k is the coder mapping at time k. At time k the controller has $s_0, \ldots, s_k$ available and then applies a control law of the general form

$$u_k = \delta_k(s_0^k) \in \mathbb{R}^m, \quad \forall k \in \mathbb{Z}_{\ge 0}, \tag{8}$$

where δ_k is the controller mapping at time k.

In practice, additional memory or structural constraints are usually placed on the general coding and control rules (7)–(8). For instance, if a static quantizer of the form (1) is used then the coding alphabet $\mathscr{S}_k \equiv \mathscr{S}$ will be constant and $s_k \equiv \gamma(y_k)$ will represent the index of the quantizer cell that contains y_k. Similarly, a static, memoryless controller is captured by setting $u_k = \delta(s_k)$ in (8).

Finite-dimensional coding and control laws may be formulated by defining internal coder and controller states ψ_k^γ and ψ_k^δ with local updates of the form

$$s_k = \gamma\left(y_k, \psi_{k-1}^\gamma\right), \;\; \psi_k^\gamma = \phi\left(s_k, \psi_{k-1}^\gamma\right), \tag{9}$$

$$\psi_k^\delta = \eta\left(s_k, \psi_{k-1}^\delta\right), \;\; u_k = \delta\left(\psi_k^\delta\right). \tag{10}$$

If the states ψ_k^γ and ψ_k^δ are finite-valued, then the coding and control laws are called *finite-state*.

Additive Noise Model

Early approaches to control modelled quantization errors as additive noise in order to allow the use of well-developed tools from linear stochastic control (Curry, 1970). While this was reasonable at high quantizer resolution, it failed to capture two key properties.

A simple example illustrates this. Consider a scalar, noiseless, fully observed, unstable LTI plant - i.e. (5) with $n = 1$, $A = a$ with $|a| > 1$, $B, C = 1$, and $w_k, v_k = 0$ - where x_0 is a random variable. Under static, high-resolution uniform quantization, the data available to the controller is expressed as a noisy linear measurement

$$y_k' := Q(x_k) = x_k + n_k, \;\; k \in \mathbb{Z}_{\ge 0},$$

where the quantizer error process (n_k) is treated as zero mean white noise uncorrelated with (x_k) and having constant variance given by (3).

The first shortcoming of this approach is that it precludes the possibility of asymptotic mean-square stability, which would effectively require the controller to estimate the initial state x_0 with a mean-square error diminishing strictly faster than a^{-2k}. This turns out to be impossible under the uncorrelatedness assumption and the constraint $|a| > 1$.

However, in the seminal paper (Delchamps, 1990), it was shown that asymptotic stability could in fact be achieved, by using a nonlinear controller that exploited the correlation between successive quantizer errors. To see this, suppose that the unknown initial state x_0 is confined to a known interval $[-l_0, l_0]$ At time $k \geq 0$, suppose that $l_k > 0$, $k = 1, 2, \ldots$ represent bounds to be determined on the future states x_k. Let Q be a static one-bit quantizer - i.e. with $M = 2$ - such that $Q(x) = 1$ if $x \geq 0$ and $Q(x) = -1$ if $x < 0$. At time k let $u_k = -0.5al_kQ(x)$ so that

$$x_{k+1} = \begin{cases} a(x_k - 0.5l_k) & \text{if } 0 \leq x_k \leq x_k \\ a(x_k + 0.5l_k) & \text{if } -l_k \leq x_k < 0 \end{cases}.$$

$\Rightarrow |x_{k+1}| \leq 0.5|a|l_k =: l_{k+1}$.

If $|a| < 2$ then the bound $l_k \to 0$ and asymptotic stability is achieved, uniformly and with exponential convergence.

However, the main drawback of the additive white noise model is that it does not predict the loss of closed-loop stability that can result when the quantizer resolution is too coarse. This is because the number M of quantizer points only serves to determine the variance of the additive noise n_k: reducing M increases the variance of n_k and the mean-square states, but they remain bounded over time. In contrast, a rigorous analysis reveals that stability is impossible by any means, linear or nonlinear, when M drops below a certain threshold.

Numerous proofs of this loss of stability exist. In a stochastic setting, the argument is based on fixing the coder and controller, and expanding out the closed-loop dynamics of the scalar LTI plant to write

$$x_k = a^k x_0 - a^k z_k \tag{11}$$

where $z_k := -a^{-k}\sum_{j=0}^{k-1} a^{k-j-1}u_j$. As z_k is a function of $s_0^{k-1} \in \mathscr{S}^k$, it can take at most M^k values. Furthermore, in the absence of noise it is fully determined by x_0, for a given coding and control policy (7)–(8). Thus z_k can be regarded as the output $Q'_k(x_0)$ of an M^k-valued quantizer. Substituting this into (11) yields

$$x_k = a^k \left(x_0 - Q'_k(x_0)\right).$$

From the asymptotic quantizer result (4), it then follows that for large k,

$$\mathrm{E}[x_k^2] \geq c\frac{a^{2k}}{M^{2k}}\left(\int f_{X_0}(x)^{1/3}\mathrm{d}\mu(x)\right)^3.$$

Thus a necessary condition for asymptotic mean-square stabilizability is that $M > |a|$ - see (Nair and Evans, 2000) for details.

The Data Rate Theorem

The discussions above emphasized the need for a more rigorous approach to quantized control. In the literature, the necessary condition $M > |a|$ was first derived in a non-random setting, where it was shown to be both sufficient and necessary to be able to ensure uniform stability (Wong and Brockett, 1999; Baillieul, 1999).

The sufficiency argument is constructive. Let Q be an M-level uniform quantizer on $[-1, 1]$, with cells formed by partitioning $[-1, 1]$ into M subintervals $R^1, \ldots, R^M$ of equal length and setting $Q(z)$ to be the midpoint of R^i when $z \in R^i$. Suppose that at time k the unknown state x_k lies in a known interval $[-l, l]$, and set $u_k = -alQ(x_k/l)$. Thus

$$|x_{k+1}| = |a||x_k - lQ(x_k/l)| = |a|l\left|\frac{x}{l} - Q\left(\frac{x}{l}\right)\right| \leq |a|\frac{l}{M}.$$

When $M > |a|$, the right-hand side $< l$. Thus $x_{k+1} \in [-l, l]$ as well, and boundedness is achieved. Uniform asymptotic stability can be achieved by replacing the constant parameter l in the argument above with a time-varying bound l_k, updated as $l_{k+1} = |a|l_k/M \to 0$.

The necessity argument is based on volume-partitioning. The basic idea is to fix an arbitrary coding and control policy and let m_k be the Lebesgue measure of the set of values that x_k can take at time $k \in \mathbb{Z}_{\geq 0}$. After k time steps, the plant dynamics expand this uncertainty volume m_0 by a factor $|a|^k$. However, the coder effectively divides this region into M^k disjoint, exhaustive pieces, each of which is shifted by the controller. As Lebesgue measure is translation-invariant, it then follows that $m_k \geq |a|^k m_0/M^k$. Consequently M must exceed $|a|$ if the closed loop is uniformly asymptotically stable.

The tight criterion $M > |a|$, or equivalently $R > \log_2|a|$, was the first instance of the *data rate theorem*. Volume-partitioning arguments and Jordan canonical forms can be used to generalize it to LTI plants with vector-valued states, yielding the necessary and sufficient condition

$$R > \sum_{i:|\lambda_i|\geq 1} \log_2|\lambda_i| =: H, \tag{12}$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A. This criterion is remarkably universal, having been shown to be tight for a variety of settings and objectives: e.g. for asymptotic r-th moment stabilizability with random, unbounded x_0 and no process or measurement noise (Nair and Evans, 2003); uniform stabilizability with bounded x_0 and no process or measurement noise (Baillieul, 2002); uniform stabilizability with bounded initial state, process and measurement noise (Hespanha et al, 2002; Tatikonda and Mitter, 2004); and mean-square stabilizability with random, unbounded initial state, process and measurement noise (Nair and Evans, 2004).

The deep nature of (12) becomes even clearer when it is noted that the right-hand side of (12) coincides with the intrinsic entropy generation rate H of the (open-loop) plant, in both the Kolmogorov-Sinai and topological senses; that is, it describes the growth rate of the number of distinguishable state trajectories. Thus the data rate theorem states that stability is possible iff the communication rate in the feedback loop exceeds the rate at which the plant generates uncertainty. This interpretation leads to the notion of *feedback entropy* (see cross-reference to article by C. Kawan).

Zooming Quantized Control

When the plant noise and initial state of the plant (5) are bounded, stability (in a uniform sense) can be guaranteed by applying a linear observer to track the plant states with bounded error, and then applying a suitable static, memoryless coding and control policy on the observer states x_k^o.

However, if the noise or initial state has unbounded support - e.g. when they are Gaussian or when prior bounds on them are not known - then stability cannot be achieved by any such static memoryless scheme, or indeed by any scheme where the control inputs (8) are bounded (Nair and Evans, 2004). The explanation is simple: due to the infinite support, there is a nonzero probability that the propagated state Ax_t will be beyond reach of the control input at some time t. The unstable plant dynamics then amplify this short-fall, causing the same phenomenon to occur with increasing probability at subsequent times, and inevitably leading to instability.

One solution is to use a *zooming quantizer*, i.e. having a dynamic range $l_k > 0$ that is not bounded *a priori* but expands or contracts according to the most recent symbol (Brockett and Liberzon, 2000). In the noiseless case, if this symbol corresponds to the 'overload region' of the quantizer (as indicated by a special symbol) then the range is updated as $l_{k+1} := \phi_{\text{out}} l_k$, where $\phi_{\text{out}} > 1$ is the 'zoom-out' factor. Otherwise $l_{k+1} := \phi_{\text{in}} l_k$, where $\phi_{\text{in}} < 1$ is the 'zoom-in' coefficient.

In the communications literature such schemes are called *adaptive quantizers* (Goodman and Gersho, 1974). If ϕ_{out} is sufficiently large compared to the unstable open-loop eigenvalues, and if ϕ_{in} is not too small, then global asymptotic stability ensues. With unbounded noise in the plant, variants of this scheme guarantee mean-square stability at any data rate satisfying (12) (Nair and Evans, 2004) or *input-to-state stability* (Liberzon and Nesic, 2007).

Zooming quantization is an important example of a finite-dimensional coder-controller (9)–(10), with l_k playing the role of an internal state variable. As the range update are driven by the symbols, both coder and controller can each generate identical copies of l_k, provided that there are no errors in the channel and they both start from the same initial range l_0. The important issue of how to design a scheme that can cope with mismatched initial internal states or a small level of channel errors is as yet largely unexplored.

Erroneous Digital Channels

The information-theoretic aspects of quantized control become especially pronounced when the channel is not error-free. In this case, the data rate theorem (12) can be extended, but in ways that are highly dependent on the precise setting and stability objective.

A common figure-of-merit for a stochastic discrete memoryless channel (DMC) is its *ordinary capacity* C. This is defined operationally as the largest block-code bit rate that can be transmitted across the channel with negligible probability of decoding error, and also coincides with the largest rate of Shannon information across the channel (Shannon, 1948). For a noiseless LTI plant with random initial state is controlled over a DMC, the condition $C > H$ is a tight criterion for almost (a.s.) asymptotic stabilizability (Matveev and Savkin, 2007a). This is a natural generalization of (12).

On the other hand, if the objective is to bound the state moments of a scalar LTI plant subject to bounded process noise, then the achievability of this goal is determined by the *anytime capacity* C_{any} (Sahai and Mitter, 2006): this is essentially given by the fastest decay rate of the decoding error probability.

However, if the aim is a.s. boundedness of an LTI plant with random initial state and bounded, nonstochastic process noise, then the stabilizability criterion changes again to $C_{0f} > H$ (Matveev and Savkin, 2007b). Here C_{0f} is the *zero-error feedback capacity* of the channel, defined as the largest block-code bit rate that can be transmitted across the channel with *exactly zero* probability of decoding error and with perfect channel feedback (Shannon, 1956).

As $C_{0f} < C_{\text{any}} < C$ for most channels, these conditions do not coincide. This suggests that there is no universal, operationally relevant information theory for feedback control over error-prone channels: such a theory must instead be tailored to match the underlying objectives and assumptions. For systems with nonstochastic disturbances, preliminary steps in this direction have been taken in (Nair, 2013, 2012). The reader is also referred to (You and Xie, 2011; Minero et al, 2013) for information-theoretic analyses of stochastic linear systems controlled via Markov channels.

Summary and Future Directions

This article described the key elements of quantized control with finite data rates, emphasising the interplay between coding and control. A great deal is now known about the fundamental limitations on stability in quantized control systems consisting a single feedback loop. Two major directions for future research suggest themselves:

- Little work has been done on designing optimal coding and control schemes or determining optimal costs at a given rate, apart from one or two special cases and structural results - see (Nair et al, 2007) and the references therein. It is very unlikely that explicit, closed-form solutions will be possible. However, numerical approaches based on the Lloyd-Max algorithm, particle filtering and model-predictive control may prove fruitful.
- Networked control systems usually consist of a number of subsystems interconnected over a network. Furthermore, in multi-agent systems the main objective may not be stability, but rather coordination or consensus to a common state. Comparatively little is known about the data rate requirements and information-theoretic aspects of these problems.

Cross References

"Data-rate of nonlinear control systems and feedback entropy"

References

Baillieul J (1999) Feedback designs for controlling device arrays with communication channel bandwidth constraints. In: ARO Workshop on Smart Structures, Penn. State U., USA

Baillieul J (2002) Feedback designs in information-based control. In: Pasik-Duncan B (ed) Stochastic Theory and Control. Proceedings of a Workshop held in Lawrence, Kansas, Springer, pp 35–57

Brockett RW, Liberzon D (2000) Quantized feedback stabilization of linear systems. IEEE Trans Autom Contr 45(7):1279–89

Curry RE (1970) Estimation and Control with Quantized Measurements. MIT Press, Cambridge, Massachusetts

Delchamps DF (1990) Stabilizing a linear system with quantized state feedback. IEEE Trans Autom Contr 35:916–24

Gersho A, Gray RM (1993) Vector Quantization and Signal Compression. Kluwer

Goodman DJ, Gersho A (1974) Theory of an adaptive quantizer. IEEE Trans Comms 22:1037–45

Hespanha J, Ortega A, Vasudevan L (2002) Towards the control of linear systems with minimum bit-rate. In: Proc. 15th Int. Symp. Math. The. Netw. Sys. (MTNS), U. Notre Dame, USA

Liberzon D, Nesic D (2007) Input-to-state stabilization of linear systems with quantized state measurements. IEEE Trans Autom Contr 52:767–81

Matveev AS, Savkin AV (2007a) An analogue of Shannon information theory for detection and stabilization via noisy discrete communication channels. SIAM J Contr Optim 46(4):1323–67

Matveev AS, Savkin AV (2007b) Shannon zero error capacity in the problems of state estimation and stabilization via noisy communication channels. Int Jour Contr 80:241–55

Minero P, Coviello L, Franceschetti M (2013) Stabilization over Markov feedback channels: the general case. IEEE Trans Autom Contr 58(2):349–62

Nair GN (2012) A nonstochastic information theory for feedback. In: Proc. IEEE Conf. Decision and Control, Maui, USA, pp 1343–48

Nair GN (2013) A nonstochastic information theory for communication and state estimation. IEEE Trans Autom Contr 58(6):1497–510

Nair GN, Evans RJ (2000) Stabilization with data-rate-limited feedback: tightest attainable bounds. Sys Contr Lett 41(1):49–56

Nair GN, Evans RJ (2003) Exponential stabilisability of finite-dimensional linear systems with limited data rates. Automatica 39:585–93

Nair GN, Evans RJ (2004) Stabilizability of stochastic linear systems with finite feedback data rates. SIAM J Contr Optim 43(2):413–36

Nair GN, Fagnani F, Zampieri S, Evans RJ (2007) Feedback control under data rate constraints: an overview. Proceedings of the IEEE 95(1):108–37, in special issue on Technology of Networked Control Systems

Sahai A, Mitter S (2006) The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link part 1: scalar systems. IEEE Trans Info Theory 52(8):3369–95

Shannon CE (1948) A mathematical theory of communication. Bell Syst Tech Jour 27:379–423, 623–56, reprinted in 'Claude Elwood Shannon Collected Papers', IEEE Press, 1993

Shannon CE (1956) The zero-error capacity of a noisy channel. IRE Trans Info Theory 2:8–19

Tatikonda S, Mitter S (2004) Control under communication constraints. IEEE Trans Autom Contr 49(7):1056–68

Wong WS, Brockett RW (1999) Systems with finite communication bandwidth constraints II: stabilization with limited information feedback. IEEE Trans Autom Contr 44:1049–53

You K, Xie L (2011) Minimum data rate for mean square stabilizability of linear systems with Markovian packet losses. IEEE Trans Autom Contr 56(4):772–85