SIAM J. Control Optim.
Vol. 40, No. 1, pp. 135–148

OPTIMAL SEQUENTIAL VECTOR QUANTIZATION OF MARKOV SOURCES*

VIVEK S. BORKAR†, SANJOY K. MITTER‡, AND SEKHAR TATIKONDA§

Abstract. The problem of sequential vector quantization of a stationary Markov source is cast as an equivalent stochastic control problem with partial observations. This problem is analyzed using the techniques of dynamic programming, leading to a characterization of optimal encoding schemes.

Key words. optimal vector quantization, sequential source coding, Markov sources, control under partial observations, dynamic programming

AMS subject classifications. 94A29, 90E20, 90C39

PII. S0363012999365261

1. Introduction.

In this paper, we consider the problem of optimal sequential vector quantization of stationary Markov sources. In the traditional rate distortion framework, the well-known result of Shannon shows that one can achieve entropy rates arbitrarily close to the rate distortion function for suitably long lossy block codes [9]. Unfortunately, long block codes imply long delays in communication systems. In particular, control applications require causal coding and decoding schemes.

These concerns are not new, and there is a sizable body of literature addressing these issues. We shall briefly mention a few key contributions. Witsenhausen [24] looked at the optimal finite horizon sequential quantization problem for finite state encoders and decoders. His encoder had a fixed number of levels. He showed that the optimal encoder for a kth order Markov source depends on at most the last k symbols and the present state of the decoder's memory. Walrand and Varaiya [23] looked at the infinite horizon sequential quantization problem for sources with finite alphabets. Using Markov decision theory, they were able to show that the optimal encoder for a Markov source depends only on the current input and the current state of the decoder. Gaarder and Slepian [12] look at sequential quantization over classes of finite state encoders and decoders. Though they lay down several useful definitions, their results, by their own admission, are incomplete. Other related works include a neural network based scheme [17] and a study of optimality properties of codes in specific cases [3], [10]. Some abstract theoretical results are given in [19].

*Received by the editors December 20, 1999; accepted for publication (in revised form) December 6, 2000; published electronically May 31, 2001. A preliminary version of this paper appeared in the *Proceedings of the* 1998 *IEEE International Symposium on Information Theory*, IEEE Information Theory Society, Piscataway, NJ, 1998, p. 71.

http://www.siam.org/journals/sicon/40-1/36526.html

†School of Technology and Computer Science, Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400005, India (borkar@tifr.res.in). This work was done while this author was visiting the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology. The research of this author was supported by a Homi Bhabha fellowship, NSF KDI: Learning, Adaptation, and Layered Intelligent Systems grant 6756500, and grant III 5(12)/96-ET of Department of Science and Technology, Government of India.

‡Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Room 35-403, Cambridge, MA 02139 (mitter@lids.mit.edu). The research of this author was supported by NSF KDI: Learning, Adaptation, and Layered Intelligent Systems grant ECS-9873451.

§University of California at Berkeley, Soda Hall, Room 485, Berkeley, CA 94720 (tatikond@eecs.berkeley.edu). The research of this author was supported by U.S. Army grant PAAL03-92-G-0115.

A formulation similar in spirit to ours (insofar as it aims to minimize a "Lagrangian distortion measure" described below) is studied in [7], [8]. They show empirically that one can make gains in performance by entropy coding the codewords. In [7] the entropy constrained vector quantization problem for a block is formulated and a Max–Lloyd-type algorithm is introduced. In [8] they introduce the conditional entropy constrained vector quantization problem and show that one should use conditional entropy coders when the codewords are not independent from block to block. In these papers there is more emphasis on synthesizing algorithms and less emphasis on proving rigorously the optimality of the schemes proposed. Along with this work there is a large literature on differential predictive coding, where one encodes the innovation. Other than the Gauss–Markov case, though, it is not apparent how one may prove the optimality of such innovation coding schemes. Herein we emphasize, through the dynamic programming formulation, the optimality properties of the sequential quantization scheme. This leads the way for the application of many powerful approximate dynamic programming tools.

In this paper we do not impose a fixed number of levels on the quantizer. The aim is to somehow jointly optimize the entropy rate of the quantized process (in order to obtain a better compression rate) as well as a suitable distortion measure. The traditional rate distortion framework [9] calls for the minimization of the former with a hard constraint on the latter. We shall, however, consider the analytically more tractable Lagrangian distortion measure of [7], [8], which is a weighted combination of the two. We approach the problem from a stochastic control viewpoint, treating the choice of the sequential quantizer as a control choice. The correct "state space" then turns out to be the space of conditional laws of the underlying process given the quantizer outputs, these conditional laws serving as the "state" or "sufficient statistics." The "state dynamics" is then given by the appropriate nonlinear filter. While this is very reminiscent of the finite state quantizers studied, e.g., in [16], the state space here is not finite, and the state process has the familiar stochastic control interpretation as the output of a nonlinear filter. We then consider the "separated" or "certainty equivalent" control problem of controlling this nonlinear filter so as to minimize an appropriately transformed Lagrangian distortion measure. This problem can be analyzed in the traditional dynamic programming framework. This in turn can be made a basis for computational schemes for near-optimal code design.

To summarize, the main contributions of this paper are as follows.

(i) We formulate a stochastic control problem equivalent to the optimal vector quantization problem. In the process, we make precise the passage from the source output to its encoded version in a manner that ensures the well-posedness of the control problem.

(ii) We underscore the crucial role of the process of conditional laws of the source given the quantized process as the correct "sufficient statistics" for the problem.

(iii) We analyze the equivalent control problem by using the methodology of Markov decision theory. This opens up the possibility of using the computational machinery of Markov decision theory for code design.

Specifically, we consider a pair of a "state process" $\{X_n\}$ and an associated "observation process" $\{Y_n\}$, given by the dynamics

$$X_{n+1} = g(X_n, \xi_n),\ Y_{n+1} = h(X_n, \xi'_n),$$

where $\{\xi_n\}, \{\xi'_n\}$ are independently and identically distributed (i.i.d.) driving noise processes. We quantize Y_{n+1} into its quantized version q_{n+1} that has a finite range and is selected based on the "history" $q^n \triangleq [q_0, q_1, \ldots, q_n]$. The aim then is to minimize the long run average of the Lagrangian distortion measure $R_n = E[H(q_{n+1}/q^n) + \lambda||Y_n - \bar{q}_n||^2]$, where $\lambda > 0$ is a prescribed constant, $H(\cdot/\cdot)$ is the conditional entropy, and $\bar{q}_n$ is the best estimate of Y_n given q_n.

Let π_n be the regular conditional law of X_n given q^n for $n \geq 0$. From π_n, one can easily derive the regular conditional law of Y_{n+1} given q^n. Using Bayes's rule, $\{\pi_n\}$ can be evaluated recursively by a nonlinear filter. Furthermore, one can express R_n as the expected value of a function of π_n and a "control" process Q_n alone. ($\{Q_n\}$ is, in fact, the finite set depicting the range of the vector quantization of Y_{n+1} prior to its encoding into a fixed finite alphabet.) This allows us to consider the equivalent problem of controlling $\{\pi_n\}$ with the aim of minimizing the long run average of the R_n recast as above. This then fits the framework of traditional Markov decision theory and can be approached by dynamic programming. As usual, one has to derive the dynamic programming equations for the average cost control problem by a "vanishing discount" argument applied to the associated infinite horizon discounted control problem for which the dynamic programming equation is easier to justify.

The structure of the paper is as follows. In section 2, we describe the sequential quantization problem and introduce the formalism. Section 3 derives the equivalent control problem. This is analyzed in section 4 using the formalism of Markov decision theory.

2. Sequential quantization. This section formulates the sequential vector quantization problem. In particular, it describes the passage from the observation process to its quantized version, which in turn gets mapped into its encoding with respect to a fixed alphabet. We also lay down our key assumptions which, apart from making the coding scheme robust, also make its subsequent control formulation well-posed. The section concludes with a precise statement of this "long run average cost" control problem with partial observations that is equivalent to our original vector quantization problem.

Throughout, for a Polish (i.e., complete separable metric) space $X, P(X)$ will denote the Polish space of probability measures on X with Prohorov topology [6, Chapter 2]. For a random process $\{Z_m\}$, set $Z^n = \{Z_m, 0 \leq m \leq n\}$, its past up to time n. Finally, K will denote a finite positive constant, depending on the context.

Let $\{X_n\}$ be an ergodic Markov process taking values in $R^s, s \geq 1$, with an associated "observation process" $\{Y_n\}$ taking values in $R^d, d \geq 1$. ($\{Y_n\}$ thus is the actual process being observed.) Their joint evolution is governed by a transition kernel $x \in R^s \to p(x, dz, dy) \in P(R^s \times R^d)$, as described below. We assume this map to be continuous and further, that $p(x, dz, dy) = \varphi(y, z|x)dzdy$ for a density $\varphi(\cdot, \cdot|\cdot) : R^d \times R^s \times R^s \to R^+$ that is continuous and strictly positive, and furthermore, $\varphi(y, z|\cdot)$ is Lipschitz uniformly in y, z.

The evolution law is as follows. For $A \subset R^s, B \subset R^d$ Borel,

$$\begin{aligned} P(X_{n+1} \in A, Y_{n+1} \in B/X^n, Y^n) &= \int_{A \times B} p(X_n, dx, dy) \\ &= \int_A \int_B \varphi(y, z|X_n)dydz. \end{aligned}$$

Following [13], we call the pair $(\{X_n\}, \{Y_n\})$ a Markov source, though the terminology "hidden Markov model" is more common nowadays. We impose on $(\{X_n\}, \{Y_n\})$ the condition of "asymptotic flatness" described next. We assume that these processes are given recursively by the dynamics

$$X_{n+1} = g(X_n, \xi_n), \tag{2.1}$$

$$Y_{n+1} = h(X_n, \xi'_n), \tag{2.2}$$

where $\{\xi_n\}, \{\xi'_n\}$ are i.i.d. R^m-valued (say) random variables independent of each other and of X_0, and $g : R^s \times R^m \to R^s$, $h : R^s \times R^m \to R^d$ are prescribed measurable maps satisfying

$$||g(x,y)||, ||h(x,y)|| \leq K(1 + ||x||) \ \ \forall\, y.$$

Equations (2.1) and (2.2) and the laws of $\{\xi_n\}, \{\xi'_n\}$ completely specify $p(x, dz, dy)$, and therefore the conditions we impose on the latter will implicitly restrict the choice of the former.

Let $(\{X_n(x)\}, \{Y_n(x)\}), (\{X_n(y)\}, \{Y_n(y)\})$ denote the solutions to (2.1), (2.2) for $X_0 = x$, respectively, y with the *same* driving noises $\{\xi_n\}, \{\xi'_n\}$. The assumption of asymptotic flatness then is that there exist $K > 0, 0 < \beta < 1$, such that

$$E[||X_n(x) - X_n(y)||] \leq K\beta^n ||x - y||, n \geq 0.$$

A simple example would be the case when $g(x,u) = \bar{g}(x) + u, h(x,u) = \bar{h}(x) + u$ for all x, u, where $\bar{g} : R^s \to R^s$ is a contraction with respect to some equivalent norm on R^s. This covers, e.g., the usual linear quadratic Gaussian (LQG) case when the state process is stable. Another example would be a discretization of continuous time asymptotically flat processes considered in [1], where a Lyapunov-type sufficient condition for asymptotic flatness is given. This assumption, one must add, is not required for our formulation of the optimization problem per se but will play a key role in our derivation of the dynamic programming equations in section 4.

Let $\sum = \{\alpha_1, \alpha_2, \ldots, \alpha_N\}$ be an ordered set that will serve as the alphabet for our vector quantizer. Let $\{q_n\}$ denote the $\sum$-valued process that stands for the "vector quantized" version of $\{Y_n\}$. The passage from $\{Y_n\}$ to $\{q_n\}$ is described below.

Let D denote the set of finite nonempty subsets of R^d with cardinality at most $N \geq 1$, satisfying the following.

(†) There exist $M > 0$ ("large") and $\triangle > 0$ ("small") such that

(i) $x \in A \in D$ implies $||x|| \leq M$,

(ii) $x = [x_1, \ldots, x_d], y = [y_1, \ldots, y_d]$ for $x, y \in A \in D$, $x \neq y$, implies $|x_i - y_i| > \triangle$ for all i.

We endow D with the Hausdorff metric which renders it a compact Polish space. For $A \in D$, let $l_A : R^d \to A$ denote the map that maps $x \in R^d$ to the element of A nearest to it with reference to the Euclidean norm $||\cdot||$, any tie being resolved according to some fixed priority rule. Let $i_A : A \to \sum$ denote the map that first orders the elements $\{a_1, \ldots, a_m\}$ of A lexicographically and then maps them to $\{\alpha_1, \ldots, \alpha_m\}$ preserving the order.

Let $\sum^\infty = \sum \times \sum \times \cdots$ (i.e., a one-sided countably infinite product. Analogous notation will be used elsewhere.) At each time n, a measurable map $\eta_n : \sum^{n+1} \to D$ is chosen. With $Q_n \triangleq \eta_n(q^n)$, one sets

$$q_{n+1} = i_{Q_n}(l_{Q_n}(Y_{n+1})).$$

This defines $\{q_n\}$ recursively as the quantized process that is to be encoded and transmitted across a communication channel.

The explanation of this scheme is as follows. In case of a fixed quantizer, the finite subset of R^d to which the signal gets mapped can itself be identified with the alphabet $\sum$. In our case, however, this set will vary from one instant to another and therefore must be mapped to a fixed alphabet $\sum$ in a uniquely invertible manner. This is achieved through the map i_A. Assuming that the receiver knows ahead of time the deterministic maps $\{n_n(\cdot)\}$ (later on we argue that a single fixed $\eta(\cdot)$ will suffice), she can reconstruct Q_n as $\eta_n(q^n)$ on having received q^n by time n. In turn, she can reconstruct $i_{Q_n}^{-1}(q_{n+1}) = l_{Q_n}(Y_{n+1})$ as the vector quantized version of Y_{n+1}. The main contribution of the condition (†) is to render the map $A = \{a_1, \ldots, a_m\} \in D \to \{i_A(a_1), \ldots, i_A(a_m)\} \in \sum^*$ continuous. Not only does this make sense from the point of view of robust decoding, but it also makes the control problem we formulate later well-posed.

As mentioned in the introduction, our aim will be to jointly optimize over the choice of $\{\eta_n(\cdot)\}$ the average entropy rate of $\{q_n\}$ ($\approx$ the average code length if the encoding is done optimally) and the average distortion. The conventional rate distortion theoretic formulation would be to minimize the average entropy rate

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n-1} E[H(q_{m+1}/q^m)],$$

$H(\cdot)$ being the (conditional) Shannon entropy, subject to a hard constraint on the distortion

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n-1} E[||Y_m - \bar{q}_m||^2] \leq K,$$

where $\bar{q}_m = i_{Q_{m-1}}^{-1}(q_m) = l_{Q_{m-1}}(Y_m)$. We shall, however, consider the simpler problem of minimizing the Lagrangian distortion measure

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n-1} E[H(q_{m+1}/q^m) + \lambda ||Y_m - \bar{q}_m||^2], \tag{2.3}$$

where $\lambda > 0$ is a prescribed constant. One may think of λ as a Lagrange multiplier, though, strictly speaking, such an interpretation is lacking given our arbitrary choice thereof.

3. Reduction to the control problem. This section derives the "completely observed" optimal stochastic control problem equivalent to the optimal vector quantization problem described above. In this, we follow the usual "separation" idea of stochastic control by identifying the regular conditional law of state given past observations (in our case, past encodings of the actual observations) as the new state process for the completely observed control problem. The original cost function is rewritten in an equivalent form that displays it as a function of the new state and control processes alone. Under the assumptions of the previous section on the permissible vector quantization schemes (as reflected in our definition of D), the above controlled Markov process is shown to have a transition kernel continuous in the initial state and control. Finally, a relaxation of this control problem is outlined, which allows for a larger class of controls. This is purely a technical convenience required for the proofs of the next section and does not affect our control problem in any essential manner.

Let $\pi_n(dx) \in P(R^s)$ denote the conditional law of X_n given $q^n, n \geq 0$. A standard application of the Bayes rule shows that $\{\pi_n\}$ is given recursively by the nonlinear filter

$$\pi_{n+1}(dx') = \frac{\int\int I\{i_{Q_n}(l_{Q_n}(y)) = q_{n+1}\}\varphi(y, x'|x)dydx'\pi_n(dx)}{\int\int\int I\{i_{Q_n}(l_{Q_n}(y)) = q_{n+1}\}\varphi(y, z|x)dydz\pi_n(dx)}. \tag{3.1}$$

By (†), $l_A^{-1}(i_A^{-1}(a))$ contains an open subset of R^d for any a, A. Given this fact and the condition that $\varphi(\cdot, \cdot|\cdot) > 0$, it follows that the denominator above is strictly positive, and hence the ratio is well defined. The initial condition for the recursion (3.1) is $\pi_0 =$ the conditional law of X_0 given q_0. We assume q_0 to be the trivial quantizer, i.e., $q_0 \equiv 0$, say, so that $\pi_0 =$ the law of X_0. Thus defined, $\{\pi_n\}$ can be viewed as a $P(R^s)$-valued controlled Markov process with a D-valued "control" process $\{Q_n\}$. To complete the description of the control problem, we need to define our cost (2.3) in terms of $\{\pi_n\}, \{Q_n\}$. For this purpose, let $\bar{\varphi}(y|x) \triangleq \int \varphi(y, z|x)dz$ for all $(x, y) \in R^s \times R^d$. Note that for $a \in \sum$,

$$\begin{aligned} P(q_{n+1} = a/q^n) &= E[E[I\{q_{n+1} = a\}/q^n, X^n]/q^n] \\ &= E\left[\int p(X_n, R^s, dy)I\{q_{n+1} = a\}/q^n\right] \\ &= \int \pi_n(dx) \int \bar{\varphi}(y|x)I\{i_{\eta_n(q^n)}(l_{\eta_n(q^n)}(y)) = a\}dy \\ &\triangleq h_a(\pi_n, Q_n), \end{aligned}$$

where $h_a : P(R^s) \times D \to R$ is defined by

$$h_a(\pi, A) = \int \pi(dx)f_a(x, A)$$

with

$$f_a(x, A) = \int \bar{\varphi}(y|x)I\{i_A(l_A(y)) = a\}dy.$$

Also define

$$\begin{aligned} \hat{f}(x, A) &= \int \bar{\varphi}(y|x)||y - l_A(y)||^2dy, \\ k(\pi, A) &= -\sum_a h_a(\pi, A)\log h_a(\pi, A), \\ r(\pi, A) &= \int \pi(dx)\hat{f}(x, A), \end{aligned}$$

where the logarithm is to the base 2. We assume $f_a(\cdot, A), \hat{f}(\cdot, A)$ to be Lipschitz uniformly in a, A. This would be implied in particular by the condition that $\bar{\varphi}(y/\cdot)$ be Lipschitz uniformly in y. Now (2.3) can be rewritten as

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{m=0}^{n-1} E[k(\pi_m, Q_m) + \lambda r(\pi_m, Q_m)]. \tag{3.2}$$

Strictly speaking, we should consider the problem of controlling $\{\pi_n\}$ given by (3.1) so as to minimize the cost (3.2). We shall, however, introduce some further simplifications, thereby replacing (3.2) by an approximation of the same. Let $\frac{1}{N} > \epsilon^\star > 0$ be a small positive constant. For $n \geq 1$, let $P_n^\star$ denote the simplex of probability vectors in R^n which have each component bounded from below by $\epsilon^\star$. That is,

$$P_n^\star = \left\{ x = [x_1, \dots, x_n] \in R^n : x_i \in [\epsilon^\star, 1] \ \ \forall\, i, \quad \sum_i x_i = 1 \right\}.$$

Similarly, let

$$P_n = \left\{ x = [x_1, \dots, x_n] \in R^n : x_i \in [0, 1] \ \ \forall\, i, \quad \sum_i x_i = 1 \right\}$$

denote the entire simplex of probability vectors in R^n. Let $\Pi_n : P_n \to P_n^\star$ denote the projection map. Let $h(\pi, A) = [h_{a_1}(\pi, A), \dots, h_{a_m}(\pi, A)]$ for $A = \{a_1, \dots, a_m\}$ and

$$\begin{aligned} \tilde{h}(\pi, A) &= \Pi_{|A|}(h(\pi, A)) \\ &\triangleq [\tilde{h}_{a_1}(\pi, A), \dots, \tilde{h}_{a_m}(\pi, A)]. \end{aligned}$$

Note that

$$|\log \tilde{h}_a(\pi, A)| \leq -\log \epsilon^\star < \infty \quad \forall\, a, \pi, A. \tag{3.3}$$

Finally, let

$$\tilde{k}(\pi, A) = -\sum_a \tilde{h}_a(\pi, A) \log \tilde{h}_a(\pi, A).$$

The control problem we consider is that of controlling $\{\pi_n\}$ so as to minimize the cost

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n-1} E[\tilde{k}(\pi_n, Q_n) + \lambda r(\pi_n, Q_n)]. \tag{3.4}$$

Replacing $k(\cdot,\cdot)$ by $\tilde{k}(\cdot,\cdot)$ is a purely technical convenience to suit the needs of the developments to come in section 4. We believe that it should be possible to obtain the same results directly for (3.2), though possibly at the expense of a considerable additional technical overhead.

We shall analyze this problem using techniques of Markov decision processes. With this in mind, call $\{Q_n\}$ a stationary control policy if $Q_n = v(\pi_n)$ for all n for a measurable $v : P(R^s) \to D$. The map $v(\cdot)$ itself may be referred to as the stationary control policy by a standard abuse of notation. Let $(\pi, A) \in P(R^s) \times D \to \phi(\pi, A, d\pi') = P(P(R^s))$ denote the transition kernel of the controlled Markov process $\{\pi_n\}$.

LEMMA 3.1. *The map $\phi(\cdot,\cdot,d\pi')$ is continuous.*

Proof. It suffices to check that for $f \in C_b(P(R^s))$, the map $\int f(y)\phi(\cdot,\cdot,dy)$ is continuous. Let $(\mu_n, A_n) \to (\mu_\infty, A_\infty)$ in $P(R^s) \times D$. Then $\{\mu_n\}$ are tight, and therefore, for any $\epsilon > 0$, we can find a compact $S_\epsilon \subset R^s$ such that $\mu_n(S_\epsilon) > 1 - \epsilon$ for $n = 1, 2, \dots, \infty$. Fix $\epsilon > 0$ and $S_\epsilon \subset R^s$. By the Stone–Weierstrass theorem, any $f \in C_b(P(R^s))$ can be approximated uniformly on S_ϵ by $\bar{f} \in C_b(P(R^s))$ of the form

$$\bar{f}(\mu) = F\left(\int f_1 d\mu, \dots, \int f_l d\mu \right)$$

for some $l \geq 1, f_1, \ldots, f_l \in C_b(R^s)$ and $F \in C_b(R^l)$. Then

$$\left| \int f(y)\phi(\mu_n, A_n, dy) - \int f(y)\phi(\mu_\infty, A_\infty, dy) \right|$$
$$\leq 4\epsilon K + \sup_{\mu \in S_\epsilon} |f(\mu) - \bar{f}(\mu)| + \left| \int \bar{f}(y)\phi(\mu_n, A_n, dy) - \int \bar{f}(y)\phi(\mu_\infty, A_\infty, dy) \right|. \tag{3.5}$$

Let

$$\nu_{ai}(\pi, A) = \int\int f_i(y) I\{i_A(l_A(y)) = a\} \bar{\varphi}(y|x) dy \pi(dx)$$

for $a \in \sum, 1 \leq i \leq l$. Direct verification leads to

$$\int \bar{f}(y)\phi(\pi, A, dy) = \sum_a h_a(\pi, A) F\left(\frac{\nu_{a1}(\pi, A)}{h_a(\pi, A)}, \ldots, \frac{\nu_{al}(\pi, A)}{h_a(\pi, A)} \right). \tag{3.6}$$

Note that for all a,

$$I\{i_{A_n}(l_{A_n}(y)) = a\} \to I\{i_{A_\infty}(l_{A_\infty}(y)) = 0\} \text{ almost everywhere (a.e.),}$$

because this convergence fails only on the boundaries of the regions $l_{A_\infty}^{-1}(b)$, $b \in A_\infty$, which have zero Lebesgue measure. (These are the so called *Voronoi* regions in vector quantization literature, viz., sets in the partition generated by the quantizer $l_{A_\infty}(\cdot)$.) Therefore, for all a, j,

$$f_j(y) I\{i_{A_n}(l_{A_n}(y)) = a\} \to f_j(y) I\{i_{A_\infty}(l_{A_\infty}(y)) = a\} \text{ a.e.}$$

If $x_n \to x_\infty$ in R^s, $\bar{\varphi}(y|x_n) \to \bar{\varphi}(y|x_\infty)$ for all y. Then by Scheffe's theorem [6, p. 26],

$$\bar{\varphi}(y|x_n)dy \to \bar{\varphi}(y|x_\infty)dy$$

in total variation. Hence for any a, j,

$$\int f_j(y) I\{i_{A_n}(l_{A_n}(y)) = a\} \bar{\varphi}(y|x_n) dy \to \int f_j(y) I\{i_{A_\infty}(l_{A_\infty}(y)) = a\} \bar{\varphi}(y|x_\infty) dy.$$

That is, the map

$$(x, A) \to \int f_j(y) I\{i_A(l_A(y)) = a\} \bar{\varphi}(y|x) dy$$

is continuous. It is clearly bounded. The continuity of $\nu_{ia}(\cdot, \cdot)$ follows. That of $h_a(\cdot, \cdot)$ follows similarly. The continuity of the sum in (3.6) then follows by one more application of Scheffe's theorem. Thus the last term on the right-hand side (RHS) of (3.5) tends to zero as $n \to \infty$. Since $\epsilon > 0$ was arbitrary and the second term on the RHS of (3.5) can be made arbitrarily small by a suitable choice of $\bar{f}$, the claim follows. ∎

We conclude this section with a description of a certain relaxation of this control problem wherein we permit a larger class of control policies, the so-called wide sense admissible controls used in [11]. Let $(\Omega, \mathcal{F}, P)$ denote the underlying probability space, where, without loss of generality, we may suppose that $\mathcal{F} = V_n \mathcal{F}_n$ for $\mathcal{F}_n = \sigma(X_i, Y_i, \xi_i, \xi_i', Q_i, i \leq n)$, $n \geq 0$. Define a new probability measure P_0 on $(\Omega, \mathcal{F})$ as follows. Let $\psi_n : \sum^{n+1} \times R^m \to P(\sum)$ denote the regular conditional law of q_{n+1} given $(q^n,\ Y_{n+1})$ for $n \geq 0$. (Thus we are now allowing for a randomized choice of Q_n, i.e., Q_n is not necessarily a deterministic function of (q^n, Y_{n+1}).) Let $\Gamma \in P(\sum)$ be any fixed probability measure with full support. If, for $n \geq 0, P_n, P_{0n}$, we denote the restrictions of P, P_0 to $(\Omega, \mathcal{F}_n)$, respectively, then $P_n << P_{0n}$ with

$$\frac{dP_n}{dP_{0n}} = \prod_{m=0}^{n-1} \frac{\psi_n(q^m, Y_{m+1})(\{q_{m+1}\})}{\Gamma(\{q_{m+1}\})}, \quad n \geq 1.$$

Then, under $P_0, \{q_n\}$ are independent of $\{X_n, Y_n, \xi_n, \xi'_n\}$ and are i.i.d. with law Γ. We say that $\{Q_n\}$ is a wide sense admissible control if under $P_0, (q_{n+1}, q_{n+2}, \ldots)$ is independent of (q^n, Q^n) for $n \geq 0$. Note that this includes $\{Q_n\}$ of the type $Q_n = \eta_n(q^n)$ for suitable maps $\{\eta_n(\cdot)\}$.

It should be kept in mind that this allows explicit randomization in the choice of $\{Q_n\}$, whence the entropy rate expression in (3.2) or (3.4) is no longer valid. Nevertheless, we continue with wide sense admissible controls in the context of (3.1)–(3.4) because, for us, this is strictly a temporary technical device to facilitate proofs. The dynamic programming formulation that we shall finally arrive at in section 4 will permit us to return without any loss of generality to the apparently more restrictive class of $\{Q_n\}$ we started out with.

4. The vanishing discount limit. This section derives the dynamic programming equations for the equivalent "separated control problem" by extending the traditional "vanishing discount" argument to the present setup. Deriving the dynamic programming equations for the long run average cost control of the separated control problem has been an outstanding open problem in the general case. We solve it here by using in a crucial manner the asymptotic flatness assumption introduced earlier. It should be noted that this assumption was not required at all in the development thus far and is included purely for facilitating the vanishing discount limit argument that follows. In particular, it could be dispensed with altogether were we to consider the finite horizon or infinite horizon discounted cost. For an alternative set of conditions (also strong) under which the dynamic programming equations for the average cost control under partial observations have been derived, see [21].

Our first step will be to modify the construction at the end of section 3 so as to construct on a common probability space two controlled nonlinear filters with a common control process but differing in their initial condition. This allows us to compare discounted cost value functions for two different initial laws. In turn, this allows us to show that their difference, with one of the two initial laws fixed arbitrarily, remains bounded and equicontinuous with respect to a certain complete metric on the space of probability measures, as the discount factor approaches unity. (This is where one uses the condition of asymptotic flatness.) The rest of the derivation mimics the classical arguments in this field.

For $\alpha \in (0, 1)$, consider the discounted control problem of minimizing

$$J_\alpha(\pi_0, \{Q_n\}) = E\left[\sum_{n=0}^{\infty} \alpha^n (\tilde{k}(\pi_n, Q_n) + \lambda r(\pi_n, Q_n))\right] \tag{4.1}$$

over $\Phi \triangleq$ the set of all wide sense admissible controls, with the prescribed π_0. Define the associated value function $V_\alpha : P(R^s) \to R$ by

$$V_\alpha(\pi_0) = \inf_{\Phi} J(\pi_0, \{Q_n\}).$$

Standard dynamic programming arguments show that $V_\alpha(\cdot)$ satisfies

$$V_\alpha(\pi) = \min_A \left[k(\pi, A) + \lambda r(\pi, A) + \beta \int \phi(\pi, A, d\pi') V_\alpha(\pi') \right] \tag{4.2}$$

for $\pi \in P(R^s)$. We shall arrive at the dynamic programming equation for our original problem by taking a "vanishing discount" limit of a variant of (4.2). For this purpose, we need to compare $V_\alpha(\cdot)$ for two distinct values of its argument. In order to do so, we first set up a framework for comparing (4.1) for two choices of π_0 but with a "common" wide sense admissible control $\{Q_n\}$. This will be done by modifying the construction at the end of the preceding section. Let $(\Omega, \mathcal{F}, P_0)$ be a probability space on which we have (i) R^s-valued, possibly dependent random variables $\hat{X}_0, \tilde{X}_0$, with laws π_0, π_0', respectively; (ii) R^m-valued i.i.d. random processes $\{\xi_m\}, \{\xi_m'\}$, independent of each other and of $[\hat{X}_0, \tilde{X}_0]$ with laws as in (2.1), (2.2); and (iii) $\sum$-valued i.i.d. random sequences $\{\hat{q}_m\}, \{\tilde{q}_m\}$ with law Γ. Also defined on $(\Omega, \mathcal{F}, P_0)$ is a D-valued process $\{Q_n\}$ independent of $([\hat{X}_0, \tilde{X}_0], \{\xi_n\}, \{\xi_n'\}, \{\tilde{q}_n\})$ and satisfying the following. For $n \geq 0, (\hat{q}_{n+1}, \hat{q}_{n+2}, \ldots)$ is independent of $Q^n, \hat{q}^n$. Let $(\hat{X}_n, \hat{Y}_n), (\tilde{X}_n, \tilde{Y}_n)$ be solutions to (2.1), (2.2) with $\hat{X}_0, \tilde{X}_0$ as above. Without loss of generality, we may suppose that $\mathcal{F} = V_n \mathcal{F}_n$ with $\mathcal{F}_n = \sigma(\hat{X}^n, \tilde{X}^n, \hat{Y}^n, \tilde{Y}^n, \hat{q}^n, \tilde{q}^n, Q^n), n \geq 0$. Define a new probability measure P on $(\Omega, \mathcal{F})$ as follows. If P_n, P_{0n} denote the restrictions of P, P_0, respectively, to $(\Omega, \mathcal{F}_n), n \geq 0$, then $P_n << P_{0n}$ with

$$\frac{dP_n}{dP_{0n}} = \prod_{m=0}^{n-1} \frac{\psi_n(\hat{q}^n, \hat{Y}_{n+1})(\{\hat{q}_{n+1},\}) \psi_n'(\tilde{q}^n, \tilde{Y}_{n+1})(\{\tilde{q}_{n+1}\})}{\Gamma(\{\hat{q}_{n+1}\})\Gamma(\{\tilde{q}_{n+1}\})},$$

where the ψ_n (respectively, ψ_n') are the regular conditional laws of $Q_n(\hat{Y}_{n+1})$ given $(\hat{q}^n, \hat{Y}_{n+1})$ (respectively, of $Q_n(\tilde{Y}_{n+1})$ given $(\tilde{q}^n, \tilde{Y}_{n+1})$) for $n \geq 0$.

What this construction achieves is the identification of each wide sense admissible control $\{Q_n\}$ for initial law $\hat{\pi}_0$ with one wide sense admissible control for $\tilde{\pi}_0$. (This identification can be many-one.) By a symmetric argument that interchanges the roles of $\hat{\pi}_0$ and $\tilde{\pi}_0$, we can identify each wide sense admissible control for $\tilde{\pi}_0$ with one for $\hat{\pi}_0$. Now suppose that $V_\alpha(\hat{\pi}_0) \leq V_\alpha(\tilde{\pi}_0)$. Then for a wide sense admissible control $\{Q_n\}$ that is optimal for $\hat{\pi}_0$ (existence of this follows by standard dynamic programming arguments), we have

$$\begin{aligned} |V_\alpha(\hat{\pi}_0) - V_\alpha(\tilde{\pi}_0)| &= V_\alpha(\tilde{\pi}_0) - V_\alpha(\hat{\pi}_0) \\ &\leq J_\alpha(\tilde{\pi}_0, \{Q_n\}) - J_\alpha(\hat{\pi}_0, \{Q_n\}) \\ &\leq \sup_\Phi |J_\alpha(\tilde{\pi}_0, \{Q_n\}) - J_\alpha(\hat{\pi}_0, \{Q_n\})|, \end{aligned}$$

where we use the above identification. If $V_\alpha(\hat{\pi}_0) \geq V_\alpha(\tilde{\pi}_0)$, a symmetric argument applies. Thus we have proved the following lemma.

Lemma 4.1.

$$|V_\alpha(\hat{\pi}_0) - V_\alpha(\tilde{\pi}_0)| \leq \sup_\Phi |J_\alpha(\hat{\pi}_0, \{Q_n\}) - J_\alpha(\tilde{\pi}_0, \{Q_n\})|.$$

Next, let $P_1(R^s) = \{\mu \in P(R^s) : \int ||x|| \mu(dx) < \infty\}$, topologized by the (complete) Vasserstein metric [20]

$$\rho(\mu_1, \mu_2) = \inf E[||X - Y||],$$

where the infimum is over all joint laws of (X, Y) such that the law of X (respectively, Y) is μ_1 (respectively, μ_2). We shall assume from now on that $\pi_0 \in P_1(R^s)$. Given the linear growth condition on $g(\cdot, y), h(\cdot, y)$ of (2.1), (2.2), uniformly in y, it is then easily deduced that $E[||X_n||] < \infty$ for all n and therefore $\pi_n \in P_1(R^s)$ almost surely (a.s.) for all n. Thus we may and do view $\{\pi_n\}$ as a $P_1(R^s)$-valued process. We then have the following lemma.

LEMMA 4.2. *For* $\hat{\pi}_0, \tilde{\pi}_0 \in P_1(R^s)$ *and* $\alpha > 0, |V_\alpha(\hat{\pi}_0) - V_\alpha(\tilde{\pi}_0)| \leq K\rho(\hat{\pi}_0, \tilde{\pi}_0)$.

Proof. Let $\{\hat{\pi}_n\}, \{\tilde{\pi}_n\}$ be solutions to (3.1) with initial conditions $\hat{\pi}_0, \tilde{\pi}_0$, respectively, and a "common" wide sense admissible control $\{Q_n\} \in \Phi$. Then for $\{\hat{X}_n\}, \{\tilde{X}_n\}$ as above (with K denoting a generic positive constant that may change from step to step)

$$\begin{aligned}
&|E[r(\hat{\pi}_n, Q_n)] - E[r(\tilde{\pi}_n, Q_n]| \\
&\quad = |E[\hat{f}(\hat{X}_n, Q_n)] - E[\hat{f}(\tilde{X}_n, Q_n)]| \\
&\quad \leq E[|\hat{f}(\hat{X}_n, Q_n) - \hat{f}(\tilde{X}_n, Q_n)|] \\
&\quad \leq KE[||\hat{X}_n - \tilde{X}_n||]
\end{aligned}$$

(by the Lipschitz condition on $\hat{f}$)

$$\leq K\beta^n E[||\hat{X}_0 - \tilde{X}_0||]$$

(by asymptotic flatness).

Now consider

$$|E[\tilde{k}(\hat{\pi}_n, Q_n)] - E[\tilde{k}(\tilde{\pi}_n, Q_n]|.$$

Suppose that $E[\tilde{k}(\hat{\pi}_n, Q_n)] \geq E[\tilde{k}(\tilde{\pi}_n, Q_n)]$. Then

$$\begin{aligned}
&|E[\tilde{k}(\hat{\pi}_n, Q_n)] - E[\tilde{k}(\tilde{\pi}_n, Q_n)]| \\
&\quad = E[\tilde{k}(\hat{\pi}_n, Q_n)] - E[\tilde{k}(\tilde{\pi}_n, Q_n)] \\
&\quad = E\left[\sum_a \tilde{h}_a(\tilde{\pi}_n, Q_n) \log \tilde{h}_a(\tilde{\pi}_n, Q_n)\right] - E\left[\sum_a \tilde{h}_a(\hat{\pi}_n, Q_n) \log \tilde{h}_a(\hat{\pi}_n, Q_n)\right] \\
&\quad = E\left[\sum_a \left(\tilde{h}_a(\tilde{\pi}_n, Q_n) \log \tilde{h}_a(\tilde{\pi}_n, Q_n) - \tilde{h}_a(\hat{\pi}_n, Q_n) \log \tilde{h}_a(\tilde{\pi}_n, Q_n)\right.\right. \\
&\qquad \left.\left. + \tilde{h}_a(\hat{\pi}_n, Q_n) \log \frac{\tilde{h}_a(\tilde{\pi}_n, Q_n)}{\tilde{h}_a(\hat{\pi}_n, Q_n)}\right)\right] \\
&\quad \leq E\left[\sum_a (\tilde{h}_a(\tilde{\pi}_n, Q_n) - \tilde{h}_a(\hat{\pi}_n, Q_n)) \log \tilde{h}_a(\tilde{\pi}_n, Q_n))\right]
\end{aligned}$$

(by Jensen's inequality)

$$\begin{aligned}
&\quad \leq E\left[\sum_a (f_a(\tilde{X}_n, Q_n) - f_a(\hat{X}_n, Q_n)) \log \tilde{h}_a(\tilde{\pi}_n, Q_n))\right] \\
&\quad \leq KE[||\tilde{X}_n - \hat{X}_n||] \\
&\quad \leq K\beta^n E[||\tilde{X}_0 - \hat{X}_0||],
\end{aligned}$$

where we use (3.3) to arrive at the second to last inequality. A symmetric argument works if $E[\tilde{k}(\hat{\pi}_n, Q_n)] \leq E[\tilde{k}(\tilde{\pi}_n, Q_n)]$, leading to the same conclusion. Combining everything, we have

$$|E[\tilde{k}(\tilde{\pi}_n, Q_n) + \lambda r(\tilde{\pi}_n, Q_n)] - E[\tilde{k}(\hat{\pi}_n, Q_n) + \lambda r(\hat{\pi}_n, Q_n)]| \\ \leq K\beta^n E[||\hat{X}_0 - \tilde{X}_0||].$$

Therefore, by Lemma 4.1,

$$|V_\alpha(\hat{\pi}_0) - V_\alpha(\tilde{\pi}_0)| \leq K \sum_n \beta^n \alpha^n E[||\hat{X}_0 - \tilde{X}_0||] \\ \leq \frac{K}{1-\beta} E[||\hat{X}_0 - \tilde{X}_0||].$$

For any $\epsilon > 0$, we can render

$$E[||\hat{X}_0 - \tilde{X}_0||] \leq \rho(\hat{\pi}_0, \tilde{\pi}_0) + \epsilon$$

by suitably choosing the joint law of $(\hat{X}_0, \tilde{X}_0)$. Since $\epsilon > 0$ is arbitrary, the claim follows. □

Fix $\pi^\star \in P(R^s)$ and define $\bar{V}_\alpha(\pi) = V_\alpha(\pi) - V_\alpha(\pi^\star)$ for $\pi \in P(R^s), \alpha \in (0,1)$. By the above lemma, $\bar{V}_\alpha(\cdot)$ is bounded equicontinuous. Letting $\alpha \to 1$, we use the Arzela–Ascoli theorem to conclude that $\bar{V}_\alpha(\cdot)$ converges in $C(P_1(R^s))$ to some $V(\cdot)$ along a subsequence $\{\alpha(n)\}, \alpha(n) \to 1$. By dropping to a further subsequence if necessary, we may also suppose that $\{(1-\alpha(n))V_{\alpha(n)}(\pi^*)\}$, which is clearly bounded, converges to some $\gamma \in R$ as $n \to \infty$. These $V(\cdot), \gamma$ will turn out to be, respectively, the value function and optimal cost for our original control problem.

Our main result is the following theorem.

THEOREM 4.3.

(i) $(V(\cdot), \gamma)$ *solve the dynamic programming equation*

$$V(\pi) = \min_u \left(\tilde{k}(\pi, u) + \lambda r(\pi, u) + \int \phi(\pi, u, d\pi')V(\pi') - \gamma \right). \tag{4.3}$$

(ii) γ *is the optimal cost, independent of the initial condition. Furthermore, a stationary policy* $v(\cdot)$ *is optimal for any initial condition if*

$$v(\pi) \in \text{Argmin} \left(\tilde{k}(\pi, \cdot) + \lambda r(\pi, \cdot) + \int \phi(\pi, \cdot, d\pi')V(\pi') \right) \quad \forall\ \pi.$$

In particular, an optimal stationary policy exists.

(iii) *If* $v(\cdot)$ *is an optimal stationary policy and* μ *is a corresponding ergodic probability measure for* $\{\pi_n\}$, *then*

$$V(\pi) = \tilde{k}(\pi, v(\pi)) + \lambda r(\pi, v(\pi)) + \int \phi(\pi, v(\pi), d\pi')V(\pi') - \gamma, \ \mu\text{-a.s.}$$

Proof. For (i) rewrite (4.2) as

$$\bar{V}_\alpha(\pi) = \min_u \left(\tilde{k}(\pi, u) + \lambda r(\pi, u) + \alpha \int \phi(\pi, u, d\pi')\bar{V}_\alpha(\pi') - (1-\alpha)V_\alpha(\pi^\star) \right).$$

Let $\alpha \to 1$ along $\{\alpha(n)\}$ to obtain (4.3).

For (ii) note that the first two statements follow by a standard argument which may be found, e.g., in [15, Theorem 5.2.4, pp. 80–81]. The last claim follows from a standard measurable selection theorem—see, e.g., [22].

For (iii) note that the claim holds if "=" is replaced by "≤". If the claim is false, we can integrate both sides with respect to μ to obtain

$$\gamma < \int (\tilde{k}(\pi, v(\pi)) + \lambda r(\pi, v(\pi)))\mu(d\pi).$$

The RHS is the cost under $v(\cdot)$, whereby this inequality contradicts the optimality of $v(\cdot)$. The claim follows. □

This result opens up the possibility of exploiting the computational machinery of Markov decision theory (see, e.g., [2], [18], [21]) for code design.

Finally, we briefly consider the decoder's problem. If transmission is error free, the decoder can construct $\{\pi_n\}$ recursively given $\{q_n\}$ and the stationary policy $v(\cdot)$. Then $\{X_n\}, \{Y_n\}$ may be estimated by the maximum a posteriori (MAP) estimates:

$$\hat{X}_n = \text{argmax } \pi_n(\cdot),$$
$$\hat{Y}_n = \text{argmax} \left(\iint I\{i_{Q_{n-1}}(l_{Q_{n-1}}(\cdot)) = q_{n+1}\}\varphi(\cdot, z|x)dz\ \pi_{n-1}(dx) \right).$$

Suppose the decoder receives $\{q_n\}$ through a noisy but memoryless channel with input alphabet $\sum$ and output alphabet another finite set O, with transition probabilities $\tilde{p}(i,j), i \in D, j \in O$. Thus $\tilde{p}(i,j) \geq 0, \sum_l \tilde{p}(i,l) = 1$ for all i, j. Let d_n be the channel output at time n.

The decoder can estimate (X_n, Y_n) given $d^n, n \geq 0$, but this is no longer easy because we cannot reconstruct $\{Q_n\}$ exactly in absence of his knowledge of $\{\pi_n\}, \{q_n\}$. Thus he should estimate $\{q_n\}$ by $\{\hat{q}_n\}$, say (e.g., by maximum likelihood), given $\{d_n\}$ and use these estimates in place of $\{q_n\}$ in the nonlinear filter for $\{\pi_n\}$, giving an approximation $\{\hat{\pi}_n\}$ to $\{\pi_n\}$. The guess for Q_n then is $v(\hat{\pi}_n), n \geq 0$.

5. Conclusions and extensions. In this paper we have considered the problem of optimal sequential vector quantization of a stationary Markov source. We have formulated the problem as a stochastic control problem. We have used the methodology of Markov decision theory. Further, we have shown that the conditional law of the source given the quantized past is a sufficient statistic for the problem. Thus the optimal encoding scheme has a separated structure. The conditional laws are given recursively by the nonlinear filter described in (3.1). The optimal policy is characterized by Theorem 4.3.

The next step is to apply traditional Markov decision problem approximation techniques to compute approximate schemes. If we have access to training data, then we can use the tools of reinforcement learning. Here the idea is to parametrize the value function space or the control law itself and apply stochastic approximation techniques to optimize those parameters.

In general, the nonlinear filter recursion is very complicated. In the literature people have approximated this by a linear prediction of the mean. These linear predictive methods can be considered an approximation to the general nonlinear filter.

REFERENCES

[1] G. Basak and R. N. Bhattacharya, *Stability in distribution for a class of singular diffusions*, Ann. Probab., 20 (1992), pp. 312–321.

[2] D. Bertsekas and J. Tsitsiklis, *Neurodynamic Programming*, Athena Scientific, Belmont, MA, 1996.

[3] A. Bist, *Differential state quantization of high order Gauss-Markov processes*, in Proceedings of the IEEE Data Compression Conference, Snowbird, UT, 1994, pp. 62–71.
[4] V. S. Borkar, *Optimal Control of Diffusion Processes*, Pitman Lecture Notes in Math. 203, Longman Scientific and Technical, Harlow, UK, 1989.
[5] V. S. Borkar, *Topics in Controlled Markov Chains*, Pitman Lecture Notes in Math. 240, Longman Scientific and Technical, Harlow, UK, 1991.
[6] V. S. Borkar, *Probability Theory: An Advanced Course*, Springer-Verlag, New York, 1995.
[7] P. Chou, T. Lookabaugh, and R. Gray, *Entropy-constrained vector quantization*, IEEE Trans. Acoust. Speech Signal Process., 37 (1989), pp. 31–42.
[8] P. Chou and T. Lookabaugh, *Conditional entropy-constrained vector quantization of linear predictive coefficients*, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Vol. 1, Albuquerque, NM, 1990, pp. 197–200.
[9] T. Cover and J. Thomas, *Elements of Information Theory*, John Wiley, New York, 1991.
[10] J. G. Dunham, *An iterative theory for code design*, in Proceedings of the IEEE International Symposium on Information Theory, St. Jovite, QC, Canada, 1983, pp. 88–90.
[11] W. Fleming and E. Pardoux, *Optimal control for partially observed diffusions*, SIAM J. Control Optim., 20 (1982), pp. 261–285.
[12] N. T. Gaarder and D. Slepian, *On optimal finite-state digital transmission systems*, IEEE Trans. Inform. Theory, 28 (1982), pp. 167–186.
[13] R. E. Gallager, *Information Theory and Reliable Communication*, John Wiley, New York, 1968.
[14] P. Hall and C. C. Heyde, *Martingale Limit Theory and Its Applications*, Academic Press, New York, London, 1980.
[15] O. Hernandez-Lerma and J. B. Lasserre, *Discrete-Time Markov Control Processes*, Springer-Verlag, New York, 1996.
[16] J. C. Kieffer, *Stochastic stability of feedback quantization schemes*, IEEE Trans. Inform. Theory, 28 (1982), pp. 248–254.
[17] E. Levine, *Stochastic vector quantization and stochastic VQ with state feedback using neural networks*, in Proceedings of the IEEE Data Compression Conference, Snowbird, UT, 1996, pp. 330–339.
[18] S. P. Meyn, *Algorithms for optimization and stabilization of controlled Markov chains*, in SADHANA: Indian Academy of Sciences Proceedings in Engineering Sciences 24, Bangalore, 1999, pp. 339–368.
[19] D. Neuhoff and R. K. Gilbert, *Causal source codes*, IEEE Trans. Inform. Theory, 28 (1982), pp. 701–713.
[20] S. T. Rachev, *Probability Metrics and the Stability of Stochastic Models*, John Wiley, Chichester, UK, 1991.
[21] W. J. Runggaldier and L. Stettner, *Approximations of Discrete Time Partially Observed Control Problems*, Applied Maths. Monographs 6, Giardini Editori e Stampatori, Pisa, Italy, 1994.
[22] D. H. Wagner, *Survey of measurable selection theorems*, SIAM J. Control Optim., 15 (1977), pp. 859–903.
[23] J. Walrand and P. P. Varaiya, *Optimal causal coding-decoding problems*, IEEE Trans. Inform. Theory, 29 (1983), pp. 814–820.
[24] H. Witsenhausen, *On the structure of real-time source coders*, The Bell System Technical Journal, 58 (1979), pp. 1437–1451.