August, 1997

LIDS- P 2395

Research Supported By:

Homi Bhabha Fellowship
US Army Grant PAAL03-92-G-0115
CICS

Optimal Sequential Vector Quantization of Markov Sources

Borkar, V.S.
Mitter, S.K.
Tatikonda, S.

Optimal Sequential Vector Quantization of Markov Sources

V. S. Borkar*
Department of Computer Science and Automation
Indian Institute of Science
Bangalore 560012
India

Sanjoy K. Mitter†
Department of Electrical Engineering and Computer Science
and Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
35–308
Cambridge, Massachusetts 02139
U.S.A.

Sekhar Tatikonda ‡
Department of Electrical Engineering and Computer Science
and Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139
U.S.A.

*This work done while visiting the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology. This research supported by a Homi Bhabha Fellowship and U.S. Army Grant PAAL03-92-G-0115, Center for Intelligent Control Systems.

†This research supported by U.S. Army Grant PAAL03-92-G-0115, Center for Intelligent Control Systems.

‡This research support by U.S. Army Grant PAAL03-92-G-0115, Center for Intelligent Control Systems.

Submitted for publication to IEEE Transactions on Information Theory, June 1997.

Abstract

Two formulations of the problem of optimal finite horizon sequential vector quantization of a Markov source are studied. The first aims to minimize a weighted sum of average entropy of the quantized process and average distortion, while the second aims to minimize the former with a hard constraint on the latter. These are converted to equivalent stochastic control problems. Existence results and dynamic and linear programming formulations are studied, leading to a 'verification theorem' for the first case that gives sufficient conditions for optimality.

Keywords: Optimal vector quantization, Sequential source coding, Markov processes

1 Introduction

In this paper, we consider the problem of optimal sequential quantization of stationary Markov processes. 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 sequential encoding and decoding schemes.

Witsenhausen [18] looked at the optimal finite horizon sequential quantization problem for finite state encoders and decoders. His encoder, however, had a fixed number of levels. He showed that the optimal encoder for a k-th order Markov source depends on at most the last k source symbols and the present state of decoder's memory. Walrand and Varaiya [17] 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 [10] looked at sequential quantization over classes of finite state encoders and decoders. Though they laid down many definitions, their results, by their own admission, are incomplete.

In this paper, we do not impose a fixed number of levels on the quantizer. The aim is to somehow jointly optimize the entropy of the quantized process in order to obtain a better compression rate, and a distortion measure. The traditional rate distortion framework calls for minimization of the former with a hard constraint on the latter. This is one of the criteria we consider. For the most part, however, we consider the analytically more tractable problem of minimizing a weighted sum of the two. Chou, Lookabaugh and Gray [8] have looked at the entropy-constrained quantization problem for one-shot coding problems. Neuhoff and Gilbert [16] showed that for a memoryless source, the optimal encoder and decoder is also memoryless. Thus the optimal sequential quantizer for an independent process is just the entropy-constrained quantizer of [8]. The situation, however, is much more complex in the Markov case.

We approach the problem from a control point of view, treating the choice of the sequential quantizer as a control choice. The correct 'state space' can then be shown to be the space of conditional laws of the process given the quantizer outputs. As already mentioned, we consider two versions of the control problem. The first is a traditional average cost control problem which seeks to minimize the asymptotic time average of expected running cost given by a weighted sum of entropy and distortion terms. In this case, we obtain a 'verification theorem' using dynamic programming which spells out sufficient conditions for optimality. The second formulation is the more traditional one, viz., as a constrained control problem wherein the entropy cost is to be minimized under a hard constraint on the average distortion. Here we have only an 'existence' result. In both cases, an important conclusion is that for Markov sources, the optimal quantization scheme may be taken to be 'Markov' in an appropriate sense.

There are limitations to our model: we assume that the state space of the source is compact. Also, a technical condition leads indirectly to an upper bound on the number of output words any quantizer might have. Finally, we have only abstract existence results and sufficient conditions. These do, however, provide a starting point for potential approximate schemes for good quantizer design. We comment on this aspect again in the concluding section.

The structure of the paper is as follows. In Section 2, we describe the sequential quantization problem and introduce the formalism. Section 3 reduces it to a pair of control problems. Section 4 recalls some relevant facts from nonlinear filtering. Section 5 establishes the abstract existence results for optimal controls for both our formulations. Section 6 introduces the Hilbert metric on the space of probability measures and derives its consequences in the present context. Section 7 approaches the average cost problem using dynamic programming, leading to the aforementioned 'verification theorem.' Section 8 comments briefly on the constrained control problem. Section 9 concludes with some relevant remarks.

2 Sequential Quantization

In this section, we formulate the sequential vector quantization problem. We shall use the following notation throughout: for a Polish space X (i.e., a separable Hausdorff space metrizable with a complete metric), $P(X)$ will denote the Polish space of probability measures on X with the Prohorov topology ([7], Ch. 2). Also, for any random process $\{Y_m\}$, set $Y^n = \{Y_i; i \leq n\}$, its 'past' up to time n.

Let $\{X_n\}$ be a stationary Markov process taking values in a compact subset S of R^d, $d \geq 1$. We assume S to be such that for any relatively open $O \subset S$, $\lambda(O) > 0$, λ being the Lebesgue measure on R^d. The transition kernel of $\{X_n\}$ will be denoted by $p(x, dy)$, viewed as a measurable map $S \to P(S)$. We assume that $p(x, dy)$ has a density $p(y|x) > 0$ w.r.t. λ on S, which is continuous in x. The strict positivity of $p(y|x)$ also implies that the process is ergodic. This is because it forces any two ergodic invariant probability measures to be mutually absolutely continuous w.r.t. λ, hence w.r.t. each other, and therefore they coincide.

Let $\{T_n\}$ denote the transition semigroup associated with $\{X_n\}$ and $\bar{\mu}$ its unique invariant probability measure. We impose the following 'strong ergodicity' condition [14]:

$$\lim_{n\to\infty} T_n f(x) = \int f(y) d\bar{\mu}(y) \quad \forall f \in L_1(\bar{\mu}), \quad x \in \text{ support } (\bar{\mu}).$$

Finally, we impose on $p(y|x)$ the following restriction: these exist $\sigma_2 > \sigma_1 > 0$ and $\varphi \in C(S)$, $\varphi(\cdot) > 0$, such that

$$\sigma_1 \varphi(y) \leq p(y|x) \leq \sigma_2 \varphi(y) \quad \forall x, y. \tag{2.1}$$

We shall comment on a possible relaxation of this condition later, in the last section.

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

Let D denote the set of finite nonempty subsets of S (more generally, of a compact subset $\hat{S}$ of R^d containing S — this won't affect our analysis), satisfying the following condition:

(†) There exists a fixed $\Delta > 0$ such that for any $A \in D$ and any distinct $[x_1, \ldots x_d]$, $[y_1, \ldots, y_d] \in A$, $|x_i - y_i| > \Delta \ \forall i$.

Note that this restricts the maximum cardinality of A, $A \in D$, by (say) $N \geq 1$. We endow D with the Hausdorff metric which renders it compact metric and therefore Polish. For $A \in D$, let $l_A : S \to A$ denote the map that maps $x \in R^d$ to the element of A nearest to it w.r.t. $\|\cdot\|$, any tie being resolved as per some fixed priority rule. Let $i_A : A \to \Sigma$ 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.

At each time n, a measurable map $\eta_n : \Sigma^\infty \to D$ is chosen. With $Q_n \triangleq \eta_n(q^n)$, one sets $q_{n+1} = i_{Q_n} \circ l_{Q_n}(X_{n+1})$, which defines $\{q_n\}$ recursively. This is the process that will be encoded and transmitted across a communication channel.

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

Our aim will be to jointly optimize over the choice of $\{\eta_n\}$, the average entropy of $\{q_n\}$ ($\approx$ the average codelength, if encoding is done optimally) and average distortion. We shall consider two formulations of this problem. The first is the 'penalty function' model wherein we minimize a weighted sum of the two, i.e.,

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n-1} E\left[H(q_{m+1}/q^m) + \mathcal{V}\|X_m - \bar{q}_m\|^2\right] \tag{2.2}$$

where $\bar{q}_m = i^{-1}_{Q_{m-1}}(q_m) = l_{Q_{m-1}}(X_m)\ \forall m$ and $\mathcal{V} > 0$ is a prescribed constant. The second is the 'constraint' model wherein we minimize average entropy with a hard constraint on the distortion, i.e., minimize

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n-1} E\left[H(q_{m+1}/q^m)\right] \tag{2.3}$$

subject to

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n-1} E\left[\|X_m - \bar{q}_m\|^2\right] \ \leq\ C \tag{2.4}$$

for a prescribed $C > 0$.

We shall primarily be concerned with the former problem, though we shall also comment on the latter.

3 Reduction to Control Problems

In this section we reduce the two formulations of the optimal vector quantization problem to equivalent control problems.

Let $\pi_n(dx)$ denote the regular conditional law of X_n given q^n. (Thus $\{\pi_n\}$ is a random process taking values in the compact Polish space $P(S)$.) For $a \in \Sigma$, we have

$$\begin{aligned} P(q_{n+1} = a/q^n) &= E\left[E\left[I\{q_{n+1} = a\}/q^n, X^n\right]/q^n\right] \\ &= E\left[\int p(y|X_n) I\{i_{Q_n} \circ l_{Q_n}(y) = a\}\lambda(dy)/q^n\right] \\ &= \int \pi_n(dx) \int p(y|x) I\left\{i_{\eta_n(q^n)} \circ l_{\eta_n(q^n)}(y) = a\right\}\lambda(dy) \qquad (3.1) \\ &\triangleq h_a(\pi_n, Q_n) \end{aligned}$$

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

$$h_a(\pi, A) = \int \pi(dx) \int p(y|x) I\{i_A \circ l_A(y) = a\}\lambda(dy)\,. \qquad (3.2)$$

Define $k, g : P(S) \times D \to R$ by:

$$k(\pi, A) = -\Sigma_a h_a(\pi, A) \log h_a(\pi, A)\,, \qquad (3.3)$$

$$g(\pi, A) = \int \pi\,(dx) \int p(y|x)(y - l_A(y))^2 \lambda(dy)\,, \qquad (3.4)$$

where the logarithm is to the base 2. Now (2.2), (2.3), (2.4) can be rewritten, respectively, as

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n-1} \quad E[k(\pi_m, Q_m) + \mathcal{V} g(\pi_m, Q_m)]\,, \tag{3.5}$$

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n-1} \quad E[k(\pi_m, Q_m)]\,, \tag{3.6}$$

and,

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n-1} \quad E[g(\pi_m, Q_m)] \leq C\,. \tag{3.7}$$

A standard application of the Bayes rule shows that $\{\pi_n\}$ is given recursively by the nonlinear filter (see, e.g., [4], Ch. VIII)

$$\pi_{n+1}(dy) = \frac{\int I\{i_{Q_n} \circ l_{Q_n}(y) = q_{n+1}\} p(y|x) \lambda(dy) \pi_n(dx)}{\int\int I\{i_{Q_n} \circ l_{Q_n}(z) = q_{n+1}\} p(z|x) \lambda(dz) \pi_n(dx)}\,. \tag{3.8}$$

By (†), $l_A^{-1} \circ i_A^{-1}(a)$ contains a relatively open set of S for all a, A. Combining this with our conditions on S and $p(y|x)$, we see that the denominator of (3.8) is strictly positive and thus (3.8) is well-defined.

In view of (3.1), one may now view $\{\pi_n\}$ as a $P(S)$-valued controlled Markov process controlled by the D-valued control sequence $\{Q_n\}$. The 'penalty function' model then reduces to the control problem of minimizing the average cost (3.5). The 'constraint' model in turn leads to the constrained control problem of minimizing the average cost (3.6) subject to the hard constraint (3.7).

Having cast the problem in the framework of controlled Markov processes, we can borrow certain standard concepts from the latter. To start with, we enlarge the class of admissible $\{Q_n\}$. So far we considered Q_n of the form $\eta_n(q^n)$ for prescribed $\eta_n(\cdot)$. More generally, we allow any $\{Q_n\}$ that satisfies: Q_{n+1}, X^∞ are conditionally independent given q^{n+1}, Q^n. This is in tune with the class of possibly randomized but nonanticipative controls of stochastic control theory, the largest class of control processes one usually admits a priori. We shall call such $\{Q_n\}$ admissible. One then redefines π_n as the regular conditional law of X_n given q^n, Q^n. This does not affect (3.1)–(3.8).

A word of caution is warranted here. Though randomization of controls is standard in stochastic control theory, it does not quite make sense in our original communications problem. For one thing, the decoder will not know the exact quantizer used even on receiving $\{q_n\}$ in an error-free manner. Worse, a little thought shows that (3.6) is not the correct expression for long run average entropy if randomization is used, making our formulation faulty. We shall, however, ignore these issues, treating randomization purely as a technical convenience. Our aim will be to provide sufficient conditions for nonrandomized optimal policies (a 'verification theorem' in control parlance) which we do for the penalty function model.

We also identify two important subclasses of admissible $\{Q_n\}$. The first is that of stationary policies wherein $Q_n = v(\pi_n)\ \forall n$ for a measurable map $v : P(S) \to D$. The second is the stationary randomized policies wherein each Q_n is conditionally independent of q^n, π^n, Q^{n-1} given π_n, with a regular conditional law $u(\pi_n, dz)$ for a measurable $u : P(S) \to P(D)$ that is independent of n. We identify the stationary policy (resp., the stationary randomized policy) with the map $v(\cdot)$ (resp., $u(\cdot)$) by abuse of notation. Under either, $\{\pi_n\}$ becomes a time-homogeneous Markov process.

We take up these control problems again in Section 5, following some additional results on the nonlinear filter in the next section.

4 The Nonlinear Filter

This section establishes some key results concerning $\{\pi_n\}$. Let $(\pi, A) \in P(S) \times D \to r(\pi, A, dy) \in P(P(S))$ denote the measurable map that is the transition kernel of the controlled Markov process $\{\pi_n\}$.

Lemma 4.1 *The map $r(\cdot, \cdot, dy)$ is continuous.*

Proof: It suffices to check that for $\tilde{f} \in C(P(S))$, the map $\int \tilde{f}(y) r(\cdot, \cdot, dy)$ is continuous. By the Stone-Weierstrass theorem, any $\tilde{f} \in C(P(S))$ can be uniformly approximated by functions of the form $f(\pi) = F\left(\int f_1 \, d\pi, \ldots, \int f_m \, d\pi\right)$ for some $m \geq 1$, $f \in C_b(R^m)$, $f_i \in C(S)$ $\forall i$. Thus it suffices to consider such an f. Let

$$v_{ai}(\pi, A) = \iint f_i(y) I\{i_A \circ l_A(y) = a\} p(y|x) \lambda(dy) \pi(dx), \quad a \in \Sigma, \quad 1 \leq i \leq m.$$

Direct verification gives

$$\int f(\nu) r(\pi, A, d\nu) = \sum_a h_a(\pi, A) F\left(\frac{v_{a1}(\pi, A)}{h_a(\pi, A)}, \cdots, \frac{v_{am}(\pi, A)}{h_a(\pi, A)}\right) \tag{4.1}$$

Let $(x_n, A_n) \to (x_\infty, A_\infty)$ in $S \times D$. Since $A \to i_A(A)$ is continuous,

$$I\{i_{A_n} \circ l_{A_n}(y) = a\} \to I\{i_{A_\infty} \circ l_{A_\infty}(y) = a\} \quad \text{a.e.} \quad \forall a\,.$$

(The convergence fails only on the boundaries of the Voronoi regions $l_{A_\infty}^{-1}(b)$, $b \in A$, which have zero Lebesgue measure.) Thus

$$\forall j,\ a,\ f_j(y) I\left\{i_{A_n} \circ l_{A_n}(y) = a\right\} \to f_j(y) I\left\{i_{A_\infty} \circ l_{A_\infty}(y) = a\right\} \text{ a.e.}$$

As $n \to \infty$, $p(y|x_n) \to p(y|x_\infty)$ by continuity, $\forall y$. By Scheffe's theorem ([7], pp. 26), $p(x_n, dy) \to p(x_\infty, dy)$ in total variation. Hence $\forall j, a$,

$$\int f_j(y) I\{i_{A_n} \circ l_{A_n}(y) = a\} p(x_n, dy) \to \int f_j(y) I\{i_{A_\infty} \circ l_{A_\infty}(y) = a\} p(x_\infty, dy)\,.$$

That is, $\forall j, a$, the map

$$(x, A) \to \int f_j(y) I\{i_A \circ l_A(y) = a\} p(x, dy)$$

is continuous. It is clearly bounded. The continuity of $v_{ai}(\cdot,\cdot)$ follows. The continuity of $h_a(\cdot,\cdot)$ follows similarly. The continuity of the sum in (4.1) now follows by one more application of Scheffe's theorem. □

We also have:

Lemma 4.2 *Under a stationary or stationary randomized policy, $\{\pi_n\}$ is an ergodic Markov process.*

Proof: Under our hypotheses on $\{X_n\}$, this follows exactly as in Theorem 2.2, pp. 242–244, [14]. □

5 Existence of Optimal Controls

As a prelude to our search for optimal quantizers, we establish here existence results for optimal control policies for the control problems introduced in Section 3. For this purpose, define 'empirical measures' $\mu_n \in P(P(S) \times D)$, $\psi_n \in P(P(S))$ by:

$$\begin{aligned} \mu_n(A \times B) &= \frac{1}{n} \sum_{m=0}^{n-1} P(\pi_m \in A, Q_m \in B)\,, \\ \psi_n(A) &= \frac{1}{n} \sum_{m=0}^{n-1} P(\pi_m \in A)\,, \end{aligned}$$

for A, B Borel in $P(S)$, D respectively. Thus ψ_n is the marginal of μ_n under the projection $P(S) \times D \to P(S)$. By compactness of D and $P(S)$, it follows (cf. Prohorov's theorem, [4], pp. 25) that $\{\mu_n\}, \{v_n\}$ are tight and hence relatively compact. Thus they each converge to a compact set of limit points in the respective spaces. Our first task will be to characterize the same.

Let $u : P(S) \to P(D)$ be a stationary randomized policy and $m_u(\cdot)$ the unique invariant probability measure of the corresponding ergodic Markov process $\{\pi_n\}$. Define an associated 'ergodic occupation measure' $\phi \in P(P(S) \times D)$ by

$$\phi(d\pi, dA) = m_u(d\pi)u(\pi, dA)\,. \tag{5.1}$$

This satisfies: $\forall f \in C(P(S))$,

$$\iint \left[f(\pi) - \int f(\pi')r(\pi, z, d\pi') \right] \phi(d\pi, dz) = 0\,. \tag{5.2}$$

This is simply a restatement of the invariance of $m_u(\cdot)$ under $u(\cdot)$. Conversely, any ϕ satisfying (5.2) may be disintegrated as in (5.1), whereby the $m_u(\cdot)$ therein will be the unique invariant probability measure for the ergodic Markov process corresponding to the stationary randomized policy $u(\cdot)$ in (5.1), specified m_u-a.s. uniquely. This establishes a correspondence between stationary randomized policies and their associated ergodic occupation measures. Let $H = \{\phi \in P(P(S) \times D) | (5.2) \text{ holds}\}$ and H_0 its subset with the further restriction: for m_u as in (5.1),

$$\int \left(\int f(x)\pi(dx) \right) m_u(d\pi) = E[f(X_0)]\,. \tag{5.3}$$

That is, the barycenter of m_u coincides with the law of X_0 (= the law of X_n for any n).

Lemma 5.1 *$H_0 = H$, is nonempty, convex, compact and furthermore, $\mu_n \to H$.*

Proof: Let $\mu(d\pi, dz) = m_u(d\pi)u(\pi, dz) \in H$. By Lemma 4.2, $\{\pi_n\}$ is ergodic under the stationary randomized policy $u(\cdot)$. Thus the law of π_n converges in the Cesaro sense to the unique invariant distribution m_u, regardless of the initial condition. In particular, if we pick

$\pi_0 =$ the law of X_0, then $\{\pi_n\}$ has the interpretation of being the process of conditional laws of X_n given q^n, Q^n resp. and therefore the barycenter of the law of π_n is the law of $X_n =$ the law of X_0 $\forall n$. The same must then be true for m_u. Thus $\mu \in H_0$. That is, $H_0 = H$. By Lemma 4.1, the expression in the square brackets in (5.2) is continuous in (π, z). Thus if $\Phi_n \to \Phi$ in $P(P(S) \times D)$ and Φ_n satisfies (5.2) for each n, so will Φ. Thus H is closed. Since (5.2) is preserved under convex combinations, it is also convex. For $f \in C(P(S))$, the strong law of large numbers for martingales ([11], pp. 35–36) leads to

$$\frac{1}{n}\sum_{m=0}^{n-1}\left(f(\pi_m) - \int f(\pi')r(\pi_m, Q_m, d\pi')\right) \to 0 \quad \text{a.s.}$$

By the dominated convergence theorem, we may take expectations to conclude that

$$\int \left(f(\pi) - \int f(\pi')r(\pi, z, d\pi')\right)\mu_n(d\pi, dz) \to 0\,.$$

Thus every limit point of $\{\pi_n\}$ satisfies (5.2). That is, $\mu_n \to H$ and H is nonempty. $\square$

The limsups in (3.5)–(3.7) are recognized as $\lim_{n\to\infty}\int(k + \mathcal{V}g)d\mu_n$, $\lim_{n\to\infty}\int kd\mu_n$ and $\lim_{n\to\infty}\int gd\mu_n$ respectively. But by the above lemma, all limit points of the pairs $(\int kd\mu_n, \int gd\mu_n)$ as $n \to \infty$, are of the type $(\int kd\mu, \int gd\mu)$ for some $\mu \in H$. Thus it suffices to restrict our attention to stationary randomized policies whereby the above limsups become limits, equal to the integrals of $k + \mathcal{V}g$, k, g resp. w.r.t the associated ergodic occupation measures. The penalty function problem then reduces to minimizing $\mu \to \int(k + \mathcal{V}g)d\mu$ on H. The constraint problem becomes: minimize $\mu \to \int kd\mu$ on H subject to $\int gd\mu \leq C$. By the usual compactness-continuity arguments, we then have:

Lemma 5.2 *Both problems above have optimal stationary randomized policies.*

6 The Hilbert Metric

This section introduces the Hilbert metric on $\mathcal{M}(S)$ and uses it to establish an important contractivity property of the unnormalized filter, along the lines of [2].

The Hilbert metric ρ on $\mathcal{M}(S)$ is defined by

$$\rho(\mu,\tilde{\mu}) = \ln\left[\sup \frac{\mu(A)\tilde{\mu}(B)}{\tilde{\mu}(A)\mu(B)}\right]$$

where the supremum is over all Borel A, B for which $\tilde{\mu}(A) > 0$, $\mu(B) > 0$. Note that this is $+\infty$ if $\mu, \tilde{\mu}$ are not mutually absolutely continuous. Also, it does not change if $\mu, \tilde{\mu}$ are scaled by positive scalars.

For each n, define $K_n : \mathcal{M}(S) \to \mathcal{M}(S)$ by

$$\begin{aligned} (K_n\mu)(dy) &= \int I\left\{i_{Q_n} \circ l_{Q_n}(y) = q_{n+1}\right\} p(y|x)\lambda(dy)\mu(dx) \\ &\triangleq \int k_n(x,y)\mu(dx)\lambda(dy) \end{aligned}$$

for a suitably defined $k_n(\cdot,\cdot)$. Also define

$$\begin{aligned} \bar{\rho}(K_n) &= \sup_{\mu,\mu'} \rho(K_n\mu, K_n\mu') \\ &= \ln\left[\sup_{y,y'} \operatorname{esssup}_{x,x'} \frac{k_n(x,y)k_n(x',y')}{k_n(x,y')k_n(x',y)}\right]. \end{aligned}$$

Lemma 6.1 $\rho(K_n\mu, K_n\mu') \leq \tanh(\bar{\rho}(K_n)/4)\rho(\mu,\mu')$.

For a proof, see Theorem 1.1, pp. 242–243, [15].

Corollary 6.1 *$\rho(K_n\mu, K_n\mu') \leq \delta\rho(\mu,\mu')$, where $\delta = \tanh(\ln(\sigma_2/\sigma_1)/2) < 1$.*

Proof: Immediate from Lemma 6.1 and (2.1). □

Lemma 6.2 *For $\mu,\mu' \in P(S)$, $\|\mu-\mu'\|_{TV} \leq \frac{2}{\ln 2}\rho(\mu,\mu')$, where $\|\cdot\|_{TV}$ is the total variation norm.*

Proof: Without loss of generality, assume that μ,μ' are mutually absolutely continuous. Write the Hahn-Jordan decomposition of the signed measure $\mu-\mu'$ as $\mu-\mu' = \mu^+ - \mu^-$, where μ^+, μ^- are mutually singular nonnegative measures, and let $A \subset S$ be a Borel set such that A (resp. A^c) carries the entire mass of μ^+ (resp. μ^-). Since $\mu^+(A) \geq 0$, $\mu(A) \geq \mu'(A)$. Similarly, $\mu'(A^c) \geq \mu(A^c)$. Clearly, $\mu^+,\mu^- > 0$. Then $\mu(A^c),\mu'(A) > 0$. Hence

$$\begin{aligned} 0 &\leq \mu(A)-\mu'(A) \leq \mu'(A)\left(\frac{\mu(A)}{\mu'(A)} - 1\right) \\ &\leq \mu'(A)\left(\frac{\mu(A)}{\mu'(A)}\cdot\frac{\mu'(A^c)}{\mu(A^c)} - 1\right) \\ &\leq \mu'(A)\left(e^{\rho(\mu,\mu')} - 1\right) \end{aligned}$$

Similarly,

$$0 \leq \mu'(A^c)-\mu(A^c) \leq \mu(A^c)\left(e^{\rho(\mu,\mu')} - 1\right)$$

Thus

$$\begin{aligned} \|\mu-\mu'\|_{TV} &= (\mu(A)-\mu'(A)) + (\mu'(A^c)-\mu(A^c)) \\ &\leq (\mu'(A)+\mu(A^c))\left(e^{\rho(\mu,\mu')} - 1\right) \end{aligned}$$

$$\leq \; 2\left(e^{\rho(\mu,\mu')}-1\right)$$

Since we also have

$$\|\mu-\mu'\|_{TV} \leq 2\,,$$

we have

$$\|\mu-\mu'\|_{TV} = \min\left(2, 2\left(e^{\rho(\mu,\mu')}-1\right)\right)\,.$$

On the interval $[0, \ln 2]$, the graph of the function $x \to 2(e^x - 1)$ lies below that of the line $x \to \frac{2}{\ln 2}x$. Combining this fact with the last inequality above, the claim follows.

□

Let $\{\pi_n\}, \{\pi'_n\}$ denote the outputs of the nonlinear filter "differing only in the initial conditions" in the following sense: Write (3.8) as

$$\pi_{n+1} = F(\pi_n, Q_n, q_{n+1}), n \geq 0,$$

for a suitably defined map F. Also, it can be easily verified that the conditional law of q_{n+1} given $\{\pi_m, q_m, Q_m, m \leq n\}$ depends only on π_n, Q_n. Thus standard stochastic realization theoretic arguments [5] allow us to write

$$q_{n+1} = G(\pi_n, Q_n, \xi_n), n \geq 0,$$

for a measurable map G, $\{\xi_n\}$ being i.i.d. random variables uniformly distributed on $[0, 1]$ such that ξ_n is independent of $\{\pi_m, q_m, Q_m, \xi_{m-1}, m \leq n\}$ $\forall n$. (This may need an enlargement of the underlying probability space, but this technicality does not affect our analysis.) Thus

$$\pi_{n+1} = F(\pi_n, Q_n, G(\pi_n, Q_n, \xi_n)), n \geq 0.$$

We say that $\{\pi_n\}, \{\pi'_n\}$ differ only in the initial conditions if both are generated by the above recursion for the same processes $\{Q_n\}, \{\xi_n\}$, on a common probability space, but π_o, π'_o are allowed to differ.

Lemma 6.3 *There exists a constant $C_1 > 0$ such that $\|\pi_n - \pi'_n\|_{TV} \leq C_1\delta^n$.*

Proof:

In view of the foregoing, and the fact that ρ is invariant under scaling by positive scalars,

$$\|\pi_n - \pi'_n\|_{TV} \leq \frac{2}{\ln 2}\rho(\pi_n, \pi'_n\} \leq \frac{2}{\ln 2}\delta^{n-1}\rho(K_1\pi_o, K_1\pi'_o) \leq \frac{2}{\ln 2}\delta^{n-1}\bar{\rho}(K_1).$$

□

7 The Penalty Function Model

We shall approach this problem by adapting the 'vanishing discount' argument of Markov decision theory [12]. But first we need some technical lemmas.

To start with, note that by (†), each Voronoi region (i.e., the polytope of points that map to any given element of A under the nearest neighbor rule) corresponding to any

$A \in D$ must contain a d-dimensional cube with sides of length $\Delta/3$. By (2.1), this cube has a minimum probability of some $\bar{a} > 0$ where '$\bar{a}$' depends only on $\varphi(\cdot)$, σ, and Δ. Thus $h_a(\pi, z)$ defined by (3.2), when nonzero, is at least $\bar{a}$. Also, it is zero for $a = \alpha_i$, $i > N$. To summarize, the vector $[h_{\alpha_1}(\pi, z), h_{\alpha_2}(\pi, z), \ldots]$ takes values in

$$G = \left\{ p = [p_1, p_2, \ldots] \mid p_i \in [0,1], \sum_i p_i = 1, \ p_i = 0 \text{ for } i > N, \right. \\ \left. \text{and for } i \leq N, \text{ either } p_i = 0 \text{ or } p_i \geq \bar{a} \right\}.$$

The function $p \to -\sum_i p_i \ln p_i$ is Lipschitz on $G \cdot A$ straightforward computation then leads to the following: Let $\{\pi_n\}$, $\{\pi'_n\}$ be as in Corollary 6.1.

Lemma 7.1 *For a suitable constant $C_3 > 0$,*

$$\left| k(\pi_n, z) - k(\pi'_n, z) \right| \leq C_3 \|\pi_n - \pi'_n\|_{TV} \quad \forall z,$$

$$\left| g(\pi_n, z) - g(\pi'_n, z) \right| \leq C_3 \|\pi_n - \pi'_n\|_{TV} \quad \forall z.$$

Let $\beta > 0$ and consider the 'discounted cost control problem' of minimizing over admissible $\{Q_n\}$ the discounted cost

$$J_\beta(\{Q_n\}, \pi_0) = E\left[\sum_{n=0}^{\infty} \beta^n \left(k(\pi_n, Q_n) + \mathcal{V} g(\pi_n, Q_n) \right) \right].$$

Define the 'value function' $V_\beta : P(S) \to R$ by

$$V_\beta(\pi) = \inf_{\{Q_n\} \text{ admissible}} J_\beta(\{Q_n\}, \pi).$$

Lemma 7.2 *$V_\beta(\cdot)$ is the unique least solution in $C_b(P(S))$ to the following 'dynamic programming' equations:*

$$V_\beta(\pi) = \min_Q \left(k(\pi, Q)) + \mathcal{V}g(\pi, Q) + \beta \int r(\pi, Q, d\pi')V_\beta(\pi') \right) . \tag{7.1}$$

Furthermore, these exists an optimal stationary policy $v(\cdot)$ optimal for any initial law, satisfying

$$v(\pi) \in \text{Argmin} \left(k(\pi, \cdot) + \mathcal{V}g(\pi, \cdot) + \beta \int r(\pi, \cdot, d\pi')V_\beta(\pi') \right) . \tag{7.2}$$

Conversely, any stationary policy satisfying (7.2) is optimal under any initial law.

This follows from standard results in Markov decision theory [12].

Lemma 7.3 *There exists a constant $\bar{C} > 0$ such that*

$$\left|V_\beta(\pi) - V_\beta(\pi')\right| \leq \bar{C} \quad \forall \pi, \pi' \in P(S), \quad \forall \beta > 0 .$$

Proof: We have

$$\begin{aligned} \left|V_\beta(\pi) - V_\beta(\pi')\right| &\leq \left| \sup_{\{Q_n\}} J_\beta(\{Q_n\}, \pi) - \sup_{\{Q_n\}} J_\beta(\{Q_n\}, \pi') \right| \\ &\leq \sup_{\{Q_n\}} \left| J_\beta(\{Q_n\}, \pi) - J_\beta(\{Q_n\}, \pi') \right| \end{aligned}$$

where in the r.h.s. we consider costs under processes $\{\pi_n\}, \{\pi'_n\}$ differing only in the initial conditions π, π' resp. By lemma 6.3, the r.h.s. is bounded by

$$\sup_{\{Q_n\}} E \left[\sum_{n=0}^{\infty} \beta^n \left| k(\pi_n, Q_n) + \mathcal{V}g(\pi_n, Q_n) - k(\pi'_n, Q'_n) - \mathcal{V}g(\pi'_n, Q'_n) \right| \right]$$

$$\leq CE\left[\sum_{n=0}^{\infty}\beta^n\delta^n\right] \leq \frac{C}{1-\delta\beta} \leq \frac{1}{1-\delta}$$

for a suitable constant $C > 0$.

□

From Lemma 5.2, we already know that the problem has an optimal stationary randomized policy $u^*(\cdot)$. Let Γ denote the corresponding (optimal) cost and $\bar{\beta} = 1 - \beta$. Write $J(\{Q_n\}, v)$ as $J_\beta(\{Q_n\}, v)$ to make explicit its dependence on β.

Lemma 7.4 $\lim_{\beta\to 1}\bar{\beta}V_\beta(v) = \Gamma \ \forall v \in P(S)$.

Proof: Since $\bar{\beta}V_\beta(v) \leq \bar{\beta}J_\beta(\{Q_n^*\}, v)$ for $\{Q_n^*\}$ generated according to $u^*(\cdot)$, we get

$$\liminf_{\beta\to 1}\bar{\beta}V_\beta(v) \leq \limsup_{\beta\to 1}\bar{\beta}V_\beta(v) \leq \lim_{\beta\to 1}\bar{\beta}J_\beta(\{Q_n^*\}, v) = \Gamma\,.$$

If either inequality is strict, we can find $\epsilon < 0$ and $\{\beta_m\}$ in $(0, 1)$ with $\beta_m \to 1$ such that $\forall m$,

$$\bar{\beta}_m J_{\beta_m}(\{Q_n^*\}, v) > \Gamma - \epsilon \geq \bar{\beta}_m V_{\beta_m}(v)\,.$$

For each m, let $v_m(\cdot)$ be an optimal stationary policy under β_m and $(\pi_n^m, Q_n^m = v_m(\pi_n^m))$, $n \geq 0$, the corresponding optimal processes for initial condition v. That is, $V_{\beta_m}(v) = J_{\beta_m}(\{Q_n^m\}, v)$. Define $\varphi_m \in P(P(S) \times D)$, $m \geq 1$, as follows: For $f \in C(P(S) \times D)$,

$$\int f\, d\varphi_m = \bar{\beta}_m E\left[\sum_{n=0}^{\infty}\beta_m^n f(\pi_n^m, Q_n^m)\right]\,.$$

Then φ_m satisfies:

$$\int f\, d\varphi_m = \beta_m \iint f(\pi')r(\pi, z, d\pi')\varphi_m(d\pi, dz) + \bar{\beta}_m f(v, v_m(v)) .$$

Let φ be a limit point of $\{\varphi_n\}$ in $P(P(S) \times D)$ as $m \to \infty$. Then φ must satisfy: $\forall f$ as above,

$$\int f\, d\varphi = \iint f(\pi')r(\pi, z, d\pi')\varphi(d\pi, dz) .$$

Disintegrating $\varphi(d\pi, dz)$ as $\eta(d\pi)\tilde{u}(\pi, dz)$, it follows that $\eta(\cdot)$ is the unique invariant probability measure under the stationary randomized policy $\tilde{u}(\cdot)$, i.e., φ is the ergodic occupation measure under $\tilde{u}(\cdot)$. But then $\int(k + \mathcal{V}g)d\varphi = \lim \int(k + \mathcal{V}g)d\varphi_m$ (under an appropriate subsequence) $\leq \Gamma - \epsilon$, which contradicts the definition of Γ. This proves the claim. □

Now fix $v_0 \in P(S)$ and let $\bar{V}_\beta(v) = V_\beta(v) - V_\beta(v_0)$ for $v \in P(S)$. Let $\bar{V}(v) = \limsup_{\beta \to 1} \bar{V}_\beta(v)$, bounded by virtue of Lemma 7.3. Rewrite (7.1) as

$$\bar{V}_\beta(\pi) = \min_Q \left[k(\pi, Q) + \mathcal{V}g(\pi, Q) - \bar{\beta}V_\beta(v_0) + \beta \int r(\pi, Q, d\pi')\bar{V}_\beta(\pi')\right] .$$

Thus

$$\begin{aligned} \sup_{1>\beta'>\beta} \bar{V}_\beta(\pi) \;&\leq\; \sup_{1>\tilde{\beta}>\beta} \min_Q \Bigg[k(\pi, Q) + \mathcal{V}g(\pi, Q) - \bar{\tilde{\beta}}V_{\tilde{\beta}}(v_0) \\ &\qquad + \tilde{\beta} \int r(\pi, Q, d\pi')\big(\sup_{1>\beta'>\beta} \bar{V}_{\beta'}(\pi')\big)\Bigg] \\ &\leq\; \min_Q \Bigg[k(\pi, Q) + \mathcal{V}g(\pi, Q) - \inf_{1>\beta'>\beta} \bar{\beta}'V_{\beta'}(v_0) \end{aligned}$$

$$+\left(\sup_{1>\tilde{\beta}>\beta}\tilde{\beta}\int r(\pi,Q,d\pi')\left(\sup_{1>\beta'>\beta}\bar{V}_{\beta'}(\pi')\right)\right)\Bigg] \tag{7.3}$$

Letting $\beta\to 1$, we have:

Lemma 7.5

$$\bar{V}(\pi)\le\inf_Q\left[k(\pi,Q)+\mathcal{V}g(\pi,Q)-\Gamma+\int r(\pi,Q,d\pi')\bar{V}(\pi')\right] \tag{7.4}$$

Proof: First note that the r.h.s. of (7.3) differs from

$$\inf_Q\left[k(\pi,Q)+\mathcal{V}g(\pi,Q)-\inf_{1>\beta'>\beta}\bar{\beta}'V_{\beta'}(v_0)+\int r(\pi,Q,d\pi')\sup_{1>\beta'>\beta}\bar{V}_{\beta}(\pi')\right]$$

by a term that goes to zero in a bounded fashion as $\beta\to 1$. The claim then follows by letting $\beta\to 1$ in (7.3) and observing that when a bounded sequence of real-valued functions monotonically decreases pointwise to a function, the corresponding infima converge to the infimum of the limiting function. □

Theorem 7.1 *For any optimal stationary randomized policy $u(\cdot)$ with associated invariant probability measure $m_u(\cdot)$, the following holds:*

(*) For m_u- a.s. π, $u(\pi)$ is supported in

$$\mathrm{Argmin}\left(k(\pi,\cdot)+\mathcal{V}g(\pi,\cdot)+\int r(\pi,\cdot,d\pi')\bar{V}(\pi')\right),$$

and (7.4) holds with equality.

Conversely, if (∗) holds for a stationary randomized policy, it must be optimal.

Proof: Let $u(\cdot)$ be an optimal stationary randomized policy and $\{\pi_n\}$ the corresponding stationary ergodic process. By (7.4),

$$\begin{aligned} \bar{V}(\pi_n) &\leq \min_Q \left(k(\pi_n, Q) + \mathcal{V}g(\pi_n, Q) - \Gamma + \int r(\pi_n, Q, d\pi')\bar{V}(\pi') \right) \\ &\leq k(\pi_n, Q_n) + \mathcal{V}g(\pi_n, Q_n) - \Gamma + E\left[\bar{V}(\pi_{n+1})/\pi_n\right] \end{aligned} \tag{7.5}$$

where $\{Q_n\}$ is the control sequence governing $\{\pi_n\}$. Since $\{\pi_n\}$ is optimal, $E[k(\pi, n, Q_n) + \mathcal{V}g(\pi_n, Q_n)] = \Gamma$. Taking expectations in (7.5), we have $E[\bar{V}(\pi_n)]$ on the l.h.s. and $E[\bar{V}(\pi_{n+1})] + E[k(\pi_n, Q_n) + \mathcal{V}g(\pi_n Q_n)] - \Gamma = E[\bar{V}(\pi_{n+1})] = E[\bar{V}(\pi_n)]$ on the r.h.s. Thus equality must prevail throughout, establishing the first claim. For the converse, if the said conditions hold, an argument similar to the above shows that the corresponding cost must equal Γ, the optimum. □

Corollary 7.1 *Suppose $v(\cdot)$ is a stationary policy such that the inf in (7.4) is attained at $Q = v(\pi)\ \forall\pi$, with equality. Then $v(\cdot)$ is optimal for any initial condition.*

Corollary 7.1 and the converse part of Theorem 7.1 specialized to stationary policies serve as sufficient conditions for a stationary policy to be optimal. This is in the spirit of the 'verification theorem' of classical optimal control. Note, however, that in absence of proven lower semicontinuity of $\bar{V}(\cdot)$, it is not guaranteed that the infimum in (7.4) is attained for each π. One does, however, have the following weaker claim:

Theorem 7.2 *For any $\epsilon > 0$, there exists an ϵ-optimal stationary policy.*

Proof: By Lemma 7.3, $\left|\bar{\beta}\bar{V}_\beta(\cdot)\right| \to 0$ uniformly as $\beta \to 1$, thus $\left|\bar{\beta}V_\beta(\cdot) - \Gamma\right| \to 0$ uniformly. Let $v_\beta(\cdot)$ be an optimal stationary policy for β and $m_\beta(\cdot)$ the corresponding unique invariant probability distribution of $\{\pi_n\}$. Then $\left|\bar{\beta}\int V_\beta(\pi)m_\beta(d\pi) - \Gamma\right| \to 0$. A straightforward computation shows that

$$\bar{\beta}\int V_\beta(\pi)m_\beta(d\pi) = \int \left(k(\pi, v_\beta(\pi)) + \mathcal{V}g\left(\pi, v_\beta(\pi)\right)\right) m_\beta(d\pi)\,,$$

i.e., the cost (3.5) under $v(\pi)$. The claim follows. □

Note that this also gives a recipe for finding near-optimal stationary policies.

If $(\tilde{V}, \tilde{\Gamma})$ is another solution to (7.4) with $\tilde{V}(\cdot)$ bounded measurable, we have

$$\tilde{V}(\pi_n) \leq k(\pi_n, Q_n) + \mathcal{V}g(\pi_n, Q_n) - \tilde{\Gamma} + E[\tilde{V}(\pi_{n+1})/\pi_n]$$

under an optimal stationary ergodic $\{\pi_n\}$ controlled by an optimal stationary randomized policy realized as $\{Q_n\}$. Taking expectations, we have

$$\tilde{\Gamma} \leq E\left[k(\pi_n, Q_n) + \mathcal{V}g(\pi_n, Q_n)\right] = \Gamma\,.$$

Now recall the 'occupation measure' formulation of the control problem from Section 5. It can be recast as: minimize $\int (k + \mathcal{V}g)\, d\mu$ over $\mu \in P(S \times D)$ satisfying

$$\iint f(y)r(x,u,dy)\mu(dx,du) = \int f(x)\mu(dx, D) \qquad \forall f \in C(S)\,.$$

This can be viewed as an infinite dimensional linear programming problem in the space of measures. As in [12], Ch. 6, the dual program is to find the pair $(b, V(\cdot))$, $V(\cdot)$ bounded measurable, that maximizes b subject to

$$b \leq \int r(x,u,dy)V(y) - V(x) + k(x,u) + \mathcal{V}g(x,u) \qquad \forall x \in S\,,\ u \in D\,.$$

The foregoing discussion then implies that both linear programs have a solution and that there is no duality gap.

8 The Constraint Model

The aim of this brief section is to underscore the difficulties in analyzing the constraint model. From Section 5, we already know the existence of an optimal stationary randomized policy for this problem. It is, however, difficult to go beyond that. If we pursue the convex programming approach of Section 5, standard Lagrange multiplier theory would tell us that the $\mu \in H_0$ that is optimal for the constrained problem is also optimal for the penalty function problem for a specific choice of $\mathcal{V}$, viz., the associated Lagrange multiplier. But this is not necessarily much help even if we knew $\mathcal{V}$, because it is just one of the minimizers of the latter cost, among possibly others. Thus, this constraint — penalty function model equivalence does not let us conclude the existence of an optimal stationary policy.

Even in the analytically more accessible discrete state/time stochastic control problems with constraints, these very problems remain [1], [6]. In fact, randomization may be unavoidable in some cases. Also, dynamic programming is not found to be a very tractable approach to constrained problems. One usually goes for the linear programming formulation where the constraint leads to another additional inequality without altering the nature of the linear program significantly [1]. In the present framework, it would be interesting to use approximation to this linear program to come up with 'good' suboptimal quantization rules.

9 Concluding Remarks

We shall comment on some potential extensions of the foregoing.

(i) The condition (2.1) can be relaxed by replacing $p(y|x)$ by n-stage transition probability density for some prescribed n. One then has to apply the arguments of Section 6 to the process sampled at times $kn + i$, $k \geq 0$, separately for each i, $0 \leq i < n$, and combine the results.

(ii) The extension to the whole space, i.e., $S = R^d$, can be managed by considering $\bar{R}^d =$ the one point compactification of R^d and then imposing conditions (†) and (2.1) on this compactified state space, in terms of an appropriate metric on $\bar{R}^d$ for the former. These conditions, however, become quite restrictive and artificial. It would be interesting to extend these results to $S = R^d$ without having to impose such restrictions. The main difficulty lies in extending the results of Section 6.

(iii) The abstract dynamic/linear programming formulations we have arrived at may be made a basis of approximation scheme for deriving 'good' vector quantization schemes. One promising approach is the recently developed techniques of neuro-dynamic programming [3].

(iv) Though stationary randomized policies are untenable for reasons already discussed, one may be able to achieve the same effect through time-sharing deterministic stationary policies. This needs further study.

(v) In this paper we do not take into account channel statistics. It is conjectured the separation theorem of information theory that states that source coding and channel coding can be designed separately without loss of optimality will no longer hold in this formulation. A future paper will discuss the effects of noisy channels on sequential lossy coding.

Finally, a concluding remark from a stochastic control perspective: What we have here is what is known as the 'separated control problem' associated with a control problem with partial observations ([4], Ch. XIII). For the average cost case, the derivation of dynamic programming conditions via the 'vanishing discount' argument is known to be hard in this case and only limited results are available [12]. We have here a special case where the situation is more fortunate, a key role being played by the fact that the 'control' affects only the observation process and not the original 'state' process $\{X_n\}$.

Acknowledgments

The authors would like to thank Dr. Rami Atar for pointing out an error in an earlier draft of this paper.

References

[1] E. Altman, *Constrained Markov decision processes*, INRIA Research Report **No. 2574**, INRIA, Sophia-Antipolis, 1995.

[2] R. Atar, O. Zeitouni, Exponential stability of nonlinear filtering, preprint, Dept. of Elec. Eng., Technion, Haifa, 1996.

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

[4] V. Borkar, *Topics in Controlled Markov Chains*, Pitman Research Notes in Maths. **No. 240**, Longman Scientific and Technical, Harlow, England, 1991.

[5] V. Borkar, White-noise representation in stochastic realization theory, *SIAM J. Control and Optim.* **31 (5)**, 1993, pp. 1093–1102.

[6] V. Borkar, Ergodic control of Markov chains with constraints — the general case, *SIAM J. Control and Optim.* **32 (1)**, 1994, pp. 176–186.

[7] V. Borkar, *Probability Theory: An Advanced Course*, Springer-Verlag, New York, 1996.

[8] P. Chou, T Lookabaugh, R. Gray, Entropy-constrained vector quantization, *IEEE Trans. on Acoustics, Speech and Signal Processing* **37 (1)**, 1989, pp. 31–42.

[9] T. Cover, J. Thomas, *Elements of Information Theory*, John Wiley, New York, 1991.

[10] N. T. Gaarder, D. Slepian, On optimal finite-state digital transmission systems, *IEEE Trans. on Information Theory* **28 (2)**, 1982, pp. 167–186.

[11] P. Hall, C. C. Heyde, *Martingale Limit Theory and Its Applications*, Academic Press, New York-London, 1980.

[12] O. Hernández-Lerma, J. B. Lasserre, *Discrete-Time Markov Control Processes: Basic Optimality Criteria*, Springer-Verlag, New York, 1996.

[13] J. Kieffer, A survey of the theory of source coding with fidelity criterion, *IEEE Trans. on Information Theory* **39 (5)**, 1993, pp. 1473–1490.

[14] H. Kunita, Ergodic properties of nonlinear filtering processes, in *"Spatial Stochastic Processes,"* K. C. Alexander and J. C. Watkins (eds.), Birkhäuser, Boston, 1991.

[15] C. Liverani, Decay of correlations, *Annals of Mathematics* **142 (1)**, 1995, pp. 239–301.

[16] D. Neuhoff, R. K. Gilbert, Causal source codes, *IEEE Trans. on Information Theory*, **28 (5)**, 1982, pp. 701–713.

[17] J. Walrand, P. Varaiya, Optimal causal coding-decoding problems, *IEEE Trans. on Information Theory*, **29 (6)**, 1983, pp. 814-820.

[18] H. Witsenhausen, On the structure of real-time source coders, *The Bell System Technical Journal*, **58 (6)**, 1979, pp. 1437–1451.