

Optimality of Walrand-Varaiya Type Policies and Approximation Results for Zero Delay Coding of Markov Sources

Richard G. Wood, Tamás Linder, and Serdar Yüksel

Abstract—Optimal zero-delay coding (quantization) of a finite-state Markov source is considered. Building on our earlier work and previous literature, using a stochastic control problem formulation, the existence and structure of optimal quantization policies are studied. Our main result establishes, for infinite horizon problems, the optimality of deterministic and stationary (Walrand-Varaiya type) Markov coding policies. In addition, the ϵ -optimality of finite-memory quantizers is established and the dependence between the memory length and ϵ is quantified. Numerical results are also presented.

Keywords: Zero-delay source coding, causal coding, quantization, stochastic control

I. INTRODUCTION

We consider a zero-delay (sequential) encoding problem where a sensor encodes an observed information source without delay. It is assumed that the information source $\{x_t\}_{t \geq 0}$ is an \mathbb{X} -valued discrete-time Markov process where \mathbb{X} is a finite set. The encoder encodes (quantizes) the source samples and transmits the encoded versions to a receiver over a discrete noiseless channel with input and output alphabet $\mathcal{M} := \{1, 2, \dots, M\}$, where M is a positive integer.

Formally, the encoder (also called the *quantization policy*) Π is a sequence of functions $\{\eta_t\}_{t \geq 0}$ with $\eta_t : \mathcal{M}^t \times (\mathbb{X})^{t+1} \rightarrow \mathcal{M}$. At time t , the encoder transmits the M -valued message $q_t = \eta_t(I_t)$ with $I_0 = x_0$, $I_t = (q_{[0,t-1]}, x_{[0,t]})$ for $t \geq 1$, where we have used the notation $q_{[0,t-1]} = (q_0, q_1, \dots, q_{t-1})$ and $x_{[0,t]} = (x_0, x_1, \dots, x_t)$. The collection of all such zero-delay encoders is called the set of admissible quantization policies and is denoted by Π_A .

Observe that for fixed $q_{[0,t-1]}$ and $x_{[0,t-1]}$, as a function of x_t the encoder $\eta_t(q_{[0,t-1]}, x_{[0,t-1]}, \cdot)$ is a *quantizer*, i.e., a (naturally Borel measurable) mapping of \mathbb{X} to the finite set \mathcal{M} . Thus any quantization policy at each time $t \geq 0$ selects a quantizer $Q_t : \mathbb{X} \rightarrow \mathcal{M}$ based on past information $(q_{[0,t-1]}, x_{[0,t-1]})$, and then “quantizes” x_t as $q_t = Q_t(x_t)$. Upon receiving q_t , the receiver generates its reconstruction u_t , also without delay. A zero-delay receiver policy is a sequence of functions $\gamma = \{\gamma_t\}_{t \geq 0}$ of type $\gamma_t : \mathcal{M}^{t+1} \rightarrow \mathbb{U}$, where \mathbb{U} denotes the finite reconstruction alphabet. Thus

$$u_t = \gamma_t(q_{[0,t]}), \quad t \geq 0.$$

For the finite horizon setting the goal is to minimize the

average cumulative cost (distortion)

$$J_{\pi_0}(\Pi, \gamma, T) := E_{\pi_0}^{\Pi, \gamma} \left[\frac{1}{T} \sum_{t=0}^{T-1} c_0(x_t, u_t) \right], \quad (1)$$

for some $T \geq 1$, where $c_0 : \mathbb{X} \times \mathbb{U} \rightarrow \mathbb{R}$ is a nonnegative cost (distortion) function, and $E_{\pi_0}^{\Pi, \gamma}$ denotes expectation with initial distribution π_0 for x_0 and under the quantization policy Π and receiver policy γ .

We primarily consider the more challenging infinite-horizon average cost problem where the objective is to minimize

$$J_{\pi_0}(\Pi, \gamma) := \limsup_{T \rightarrow \infty} E_{\pi_0}^{\Pi, \gamma} \left[\frac{1}{T} \sum_{t=0}^{T-1} c_0(x_t, u_t) \right]. \quad (2)$$

Our main assumption on the Markov source $\{x_t\}$ is the following.

Assumption 1. $\{x_t\}$ is an irreducible and aperiodic finite state Markov chain.

A. Revisiting structural results for finite-horizon problems

Theorem 1 (Witsenhausen [1]). *For the finite horizon problem, any zero-delay quantization policy $\Pi = \{\eta_t\}$ can be replaced, without any loss in performance, by a policy $\hat{\Pi} = \{\hat{\eta}_t\}$ which only uses $q_{[0,t-1]}$ and x_t to generate q_t , i.e., such that $q_t = \hat{\eta}_t(q_{[0,t-1]}, x_t)$ for all $t = 1, \dots, T-1$.*

Let $\mathcal{P}(\mathbb{X})$ denote the space of probability measures on \mathbb{X} . Given a quantization policy Π , for all $t \geq 1$ let $\pi_t \in \mathcal{P}(\mathbb{X})$ be the conditional probability defined by

$$\pi_t(A) := P(x_t \in A | q_{[0,t-1]})$$

for any $A \subset \mathbb{X}$.

The following result is due to Walrand and Varaiya [2].

Theorem 2. *For the finite horizon problem, any zero-delay quantization policy can be replaced, without any loss in performance, by a policy which at any time $t = 1, \dots, T-1$ only uses the conditional probability measure $\pi_t = P(dx_t | q_{[0,t-1]})$ and the state x_t to generate q_t . In other words, at time t such a policy uses π_t to select a quantizer $Q_t : \mathbb{X} \rightarrow \mathcal{M}$ and then q_t is generated as $q_t = Q_t(x_t)$.*

Due to space restrictions, we are unable to provide a detailed account of the literature; see [3], [4] and [5] for a review of some representative work.

In [6], we provided existence results for finite and infinite horizon average-cost problems when the source is \mathbb{R}^d -valued. However, even though the existence of optimal quantizers for finite horizons and the existence of optimal deterministic

The authors are with the Department of Mathematics and Statistics, Queen's University, Kingston, Ontario, Canada, K7L 3N6. Email: richard.wood@queensu.ca, linder@mast.queensu.ca, yuksel@mast.queensu.ca

This research was partially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

non-stationary policies for infinite horizons were established for stationary Markov sources in [6], the optimality of *stationary* and *deterministic* policies was not established. In [6], for technical reasons, it was assumed that the quantizers applied have convex codecells, an assumption we will not need since we deal with a finite source alphabet. Related recent work in the literature that investigates infinite horizon problems includes [7] and [8].

The main contribution here is to establish the optimality of *stationary* and *deterministic* Walrand-Varaiya type policies, among all admissible policies, for optimal zero-delay source coding for a class of Markov sources in the infinite time horizon case. In addition, we establish the ϵ -optimality of periodic zero-delay finite-length codes and obtain explicit bounds on how ϵ and the length of the coding period are related. We note that optimality of Witsenhausen and Walrand-Varaiya type coding schemes are based on dynamic-programming principles and thus critically depend on the finiteness of the time horizons. We wish to highlight the fact that our results for the infinite horizon case do not directly follow from existing results on partially observable Markov decision processes, because in a partially observable Markov decision process the state and the actions as well as the probability measure-valued *expanded state* and the action always constitute a *controlled Markov chain*. This is not the case in the problem considered; a quantizer and the state constructed in this paper do not form a controlled Markov chain under an arbitrary quantizer policy; see [15] for further discussions on this topic.

II. QUANTIZER ACTIONS AND CONTROLLED MARKOV PROCESS CONSTRUCTION

We build on the construction in [9]. Recall the notation $M = \{1, \dots, M\}$.

Definition 1. An M -cell quantizer Q on \mathbb{X} is a mapping $Q : \mathbb{X} \rightarrow M$. We let \mathcal{Q} denote the collection of all M -cell quantizers on \mathbb{X} .

Note that each $Q \in \mathcal{Q}$ is uniquely characterized by its *quantization cells* (or bins) $B_i = Q^{-1}(i) = \{x : Q(x) = i\}$, $i = 1, \dots, M$ which form a partition of \mathbb{X} . We also allow for the possibility that some of the cells of the quantizer are empty. In view of Theorem 2, any admissible quantization policy in a finite-horizon scheme can be replaced by a Walrand-Varaiya-type policy. We also refer to such policies as Markov policies. The class of all such policies is denoted by Π_W and is formally defined in [6] as follows.

Definition 2. An (admissible) quantization policy $\Pi = \{\eta_t\}$ belongs to Π_W if there exist a sequence of mappings $\{\hat{\eta}_t\}$ of the type $\hat{\eta}_t : \mathcal{P}(\mathbb{X}) \rightarrow \mathcal{Q}$ such that for $Q_t = \hat{\eta}_t(\pi_t)$ we have $q_t = Q_t(x_t) = \eta_t(I_t)$.

A policy in Π_W is called *stationary* if $\hat{\eta}_t$ does not depend on t . Suppose we use a quantizer policy $\Pi = \{\eta_t\}$ in Π_W . Let $P(x_{t+1}|x_t)$ denote the transition kernel of the process $\{x_t\}$ and note that $P(q_t|\pi_t, x_t)$ is determined by the quantizer policy as $P(q_t|\pi_t, x_t) = 1_{\{Q_t(x_t)=q_t\}}$, where $Q_t =$

$\hat{\eta}_t(\pi_t)$. Then standard properties of conditional probability can be used to obtain the following “filtering equation” for the evolution of π_t :

$$\begin{aligned} \pi_{t+1}(x_{t+1}) &= \frac{\sum_{x_t} \pi_t(x_t) P(q_t|\pi_t, x_t) P(x_{t+1}|x_t)}{\sum_{x_t} \sum_{x_{t+1}} \pi_t(x_t) P(q_t|\pi_t, x_t) P(x_{t+1}|x_t)} \\ &= \frac{1}{\pi_t(Q^{-1}(q_t))} \sum_{Q^{-1}(q_t)} P(x_{t+1}|x_t) \pi_t(x_t). \end{aligned} \quad (3)$$

Clearly, given π_t and Q_t , π_{t+1} is conditionally independent of $(\pi_{[0,t-1]}, Q_{[0,t-1]})$. Thus $\{\pi_t\}$ can be viewed as $\mathcal{P}(\mathbb{X})$ -valued controlled Markov process [10] with \mathcal{Q} -valued control $\{Q_t\}$ and average cost up to time $T-1$ given by

$$E_{\pi_0}^{\Pi} \left[\frac{1}{T} \sum_{t=0}^{T-1} c(\pi_t, Q_t) \right] = \min_{\gamma} J_{\pi_0}(\Pi, \gamma, T),$$

where

$$c(\pi_t, Q_t) := \sum_{i=1}^M \min_{u \in \mathbb{U}} \sum_{Q_t^{-1}(i)} \pi_t(dx) c_0(x, u). \quad (4)$$

In this context, Π_W corresponds to the class of deterministic Markov control policies [10].

III. FINITE HORIZON AND INFINITE HORIZON DISCOUNTED COSTS

A. Finite horizon cost

For any quantization policy Π in Π_W and any $T \geq 1$ we define

$$J_{\pi_0}(\Pi, T) := \inf_{\gamma} J_{\pi_0}(\Pi, \gamma, T) = E_{\pi_0}^{\Pi} \left[\frac{1}{T} \sum_{t=0}^{T-1} c(\pi_t, Q_t) \right],$$

where $c(\pi_t, Q_t)$ is defined in (4). The following statements follow from results in [6], but also can be easily derived since, in contrast to [6], in our case here are only finitely many M -level quantizers on \mathbb{X} .

Theorem 3. For any $T \geq 1$, there exists a policy Π in Π_W such that

$$J_{\pi_0}(\Pi, T) = \inf_{\Pi' \in \Pi_W} J_{\pi_0}(\Pi', T). \quad (5)$$

Letting $J_T^T(\cdot) := 0$, $J_0^T(\pi_0) := \min_{\Pi \in \Pi_W} J_{\pi_0}(\Pi, T)$, the dynamic programming recursion

$$T J_t^T(\pi) = \min_{Q \in \mathcal{Q}} (c(\pi, Q) + T E[J_{t+1}^T(\pi_{t+1}) | \pi_t = \pi, Q_t = Q])$$

holds for $t = T-1, T-2, \dots, 0$, $\pi \in \mathcal{P}(\mathbb{X})$.

B. Infinite horizon discounted cost

Consider the so-called discounted cost problem where the goal is to find policies that achieve

$$J^{\beta}(\pi_0) = \inf_{\Pi \in \Pi_W} J_{\pi_0}^{\beta}(\Pi) \quad (6)$$

for some $\beta \in (0, 1)$, where

$$J_{\pi_0}^{\beta}(\Pi) = \inf_{\gamma} \lim_{T \rightarrow \infty} E_{\pi_0}^{\Pi, \gamma} \left[\sum_{t=0}^{T-1} \beta^t c_0(x_t, u_t) \right].$$

From a source coding point of view the discounted cost of much less importance than the average cost; we will use it to derive results for the average cost problem.

The analysis for Theorem 3 builds first on Witsenhausen's result Theorem 1, which only considers finite horizon problems. Now, going from a finite horizon to an infinite horizon follows from a change of order of limit and infimum as we discuss in the following. Observe that

$$\begin{aligned} & \inf_{\{\Pi\}} \lim_{T \rightarrow \infty} E_{\pi_0}^{\Pi, \gamma} \left[\sum_{t=0}^{T-1} \beta^t c_0(x_t, u_t) \right] \\ & \geq \limsup_{T \rightarrow \infty} \inf_{\Pi \in \Pi_W} E_{\pi_0}^{\Pi, \gamma} \left[\sum_{t=0}^{T-1} \beta^t c_0(x_t, u_t) \right] \end{aligned} \quad (7)$$

On the other hand, due to the discounted nature of the problem, the right hand side can be studied through the dynamic programming (Bellman) iteration algorithms and is guaranteed to lead to an admissible policy, which further provides a lower bound on the discounted cost for any causal coding policy. This leads to the fixed point equation

$$\begin{aligned} J^\beta(\pi) &= \min_Q \left(c(\pi, Q) \right. \\ & \quad \left. + \beta \int P(d\pi_{t+1} | \pi_t = \pi, Q_t = Q) J^\beta(\pi_{t+1}) \right). \end{aligned}$$

We state this result formally.

Theorem 4. *There exists an optimal deterministic quantization policy in Π_W among all policies in Π_A that solves (6).*

IV. AVERAGE COST AND THE AVERAGE COST OPTIMALITY EQUATION (ACOE)

The more challenging case is the average cost problem where one considers

$$J_{\pi_0}(\Pi, \gamma) = \limsup_{T \rightarrow \infty} E_{\pi_0}^{\Pi, \gamma} \left[\frac{1}{T} \sum_{t=0}^{T-1} c_0(x_t, u_t) \right] \quad (8)$$

and the goal is to find an optimal policy attaining

$$J_{\pi_0} := \inf_{\Pi \in \Pi_A} \inf_{\gamma} J_{\pi_0}(\Pi, \gamma). \quad (9)$$

A. Optimality of non-stationary deterministic policies in Π_W

For the infinite horizon setting the structural results in Theorems 1 and 2 are not available in the literature, due to the fact that the proofs are based on dynamic programming which starts at a finite terminal time stage and optimal policies are computed backwards. However, as in [6], we can prove an infinite-horizon analog of Theorem 2 assuming that an invariant measure π^* for $\{x_t\}$ exists and $\pi_0 = \pi^*$.

Theorem 5 ([6]). *If $\{x_t\}$ starts from π^* , then there exists an optimal policy in Π_W solving the minimization problem (9), i.e., there exists $\Pi \in \Pi_W$ such that*

$$\limsup_{T \rightarrow \infty} E_{\pi^*}^{\Pi} \left[\frac{1}{T} \sum_{t=0}^{T-1} c(\pi_t, Q_t) \right] = J_{\pi^*}.$$

The proof of the theorem relies on a construction that pieces together policies from Π_W that on time segments of

appropriately large lengths increasingly well approximate the infimum of the infinite-horizon cost achievable by policies in Π_A . We can however, for the setup considered here, also establish the optimality of deterministic and stationary policies. For this, we first revisit a result which is useful in its own right given its practical implications.

B. ϵ -optimality of finite-memory policies

Define

$$J_{\pi^*}(T) := \min_{\Pi \in \Pi_A} \min_{\gamma} E_{\pi^*}^{\Pi, \gamma} \left[\frac{1}{T} \sum_{t=0}^{T-1} c_0(x_t, u_t) \right]$$

and note that by [6, p. 5988], $\limsup_{T \rightarrow \infty} J_{\pi^*}(T) \leq J_{\pi^*}$. Thus there exists an *increasing* sequence of time indices $\{T_k\}$ such that for all $k = 1, 2, \dots$,

$$J_{\pi^*}(T_k) \leq J_{\pi^*} + \frac{1}{k}. \quad (10)$$

Observe that by Theorem 2 for all k there exists $\Pi_k = \{\hat{\eta}_t^{(k)}\} \in \Pi_W$ such that

$$J_{\pi^*}(\Pi_k, T_k) := E_{\pi^*}^{\Pi_k} \left[\frac{1}{T_k} \sum_{t=0}^{T_k-1} c(\pi_t, Q_t) \right] \leq J_{\pi^*}(T_k) + \frac{1}{k}. \quad (11)$$

The above shows that for every ϵ , there exists a finite-memory encoding policy which is ϵ -optimal. This is a practically important result since most of the practical encoding schemes are finite-memory schemes. ϵ -optimal. In particular, for every ϵ , there exists N_ϵ such that, an encoder of the form: $\eta_t(x_t, m_t)$ with $m_t = g_t(m_{t-1}, q_{t-1})$ and $m_t \in \mathcal{N}$, $|\mathcal{N}| = N_\epsilon$ is ϵ -optimal. Different from [6], later in the paper we will make this relation numerically explicit.

Remark 1. Even when the initial distribution of the source is not stationary, the above policy achieves the optimal average cost corresponding to the stationary source. This follows from similar arguments as in Theorem 10 of [6].

C. Optimality of deterministic stationary quantization policies in Π_W and the Average Cost Optimality Equation (ACOE)

First note that without loss of generality the evolution of $\{x_t\}$ can be represented as

$$x_{t+1} = F(x_t, w_t), \quad t = 0, 1, 2, \dots, \quad (12)$$

where $F : \mathbb{X} \times \mathbb{X} \rightarrow \mathbb{X}$ and $\{w_t\}$ is an i.i.d. vector noise sequence which is independent of x_0 .

Lemma 1. *Under Assumption 1, there exists a representation of the form (12) such that for any initial conditions $a \in \mathbb{X}$ and $b \in \mathbb{X}$, two Markov chains x'_t with $x'_0 = a$ and x''_t with $x''_0 = b$, driven by the same noise process, we have $E[\inf\{t > 0 : x'_t = x''_t\}] \leq K_1 = \frac{n}{\bar{p}_n} < \infty$, where*

$$\begin{aligned} n &= \inf\{m > 0, m \in \mathbb{Z} : P(x'_{i+m} = a | x'_i = b) > 0, \\ & \quad P(x''_{i+m} = a | x''_i = b) > 0, \forall a, b \in \mathbb{X}\} \\ \bar{p}_k &= \min_{a, b \in \mathbb{X}} P(x'_{i+k} = x''_{i+k} | x'_i = a, x''_i = b). \end{aligned}$$

The following is one of the main results of this paper.

Theorem 6. (i) Under Assumption 1, there exists an optimal quantization policy in Π_W , which is furthermore stationary, among all policies in Π_A .

(ii) If $x_0 \sim \pi$, where π is the invariant probability measure, then for every ϵ , there exists a finite-memory non-stationary but periodic quantization policy with period less than $\frac{2K_1\|c\|_\infty}{\epsilon}$ that achieves an ϵ -optimal performance.

D. Proof of Theorem 6

1) Part (i): First, we provide a brief overview of the *Average Cost Optimality Equation* (or ACOE), see e.g., Chapter 5 in [10]. The collection of functions g, h, f is a canonical triplet if for all π ,

$$\begin{aligned} g(\pi) &= \inf_Q \int g(\pi_{t+1}) P(d\pi_{t+1} | \pi, Q) \\ g(\pi) + h(\pi) &= \inf_Q \left(c(\pi, Q) + \int h(\pi_{t+1}) P(d\pi_{t+1} | \pi, Q) \right) \end{aligned} \quad (13)$$

with $g(\pi) = \int g(\pi_{t+1}) P(d\pi_{t+1} | \pi, f(\pi))$ and $g(\pi) + h(\pi) = c(\pi, Q) + \int h(\pi_{t+1}) P(d\pi_{t+1} | \pi, f(\pi))$

Theorem 7. Let g, h, f be a canonical triplet. If g is a constant (in which case (13) is known as the *Average Cost Optimality Equation*) and $\limsup_{n \rightarrow \infty} \frac{1}{n} E_\pi^\Pi[h(\pi_n)] = 0$ for all π and under every policy Π , then the stationary deterministic policy $\Pi^* = \{f\}$ is optimal so that

$$g = J(\pi, \Pi^*) = \inf_{\Pi \in \Pi_A} J(\pi, \Pi)$$

where $J(\pi, \Pi) = \limsup_{T \rightarrow \infty} \frac{1}{T} E_\pi^\Pi[\sum_{k=0}^{T-1} c(\pi_k, Q_k)]$. Furthermore,

$$\begin{aligned} & \left| \frac{1}{n} E_\pi^{\Pi^*} \sum_{t=1}^n [c(\pi_{t-1}, Q_{t-1})] - g \right| \\ & \leq \frac{\max(|h(\pi_0)|, |E_\pi^\Pi[h(\pi_n)]|)}{n}. \end{aligned} \quad (14)$$

Sufficient conditions for the ACOE to hold have been reported in [10] (see Assumptions 4.2.1 and 5.5.1 in [10]). More applicable to our context are Assumptions 1 and 2 in [11]; since the state and the action sets are compact, weak continuity of the transition kernel is directly established from the corresponding results in [6], and the fact that the action space is finite. We need to establish the boundedness and the equicontinuity of the collection of functions $V_\beta(\mu_0) = J^\beta(\mu_0) - J^\beta(\zeta_0)$, $\beta \in (0, 1)$ for some arbitrary but fixed ζ_0 . This follows from an argument below, which essentially builds on a related approach by Borkar [12] (see also [13] and [14]), but due to the control-free evolution of the source process and that the quantization outputs under a deterministic policy are specified once the source values are, the argument is more direct here.

We will establish the conditions by, as in [14], enlarging the space of coding policies to allow for randomization at the encoder. Since for a discounted infinite horizon optimal encoding problem, optimal policies are deterministic even among possibly randomized policies, the addition of

common randomness does not benefit the encoder for optimal performance. The randomization is achieved through a parallel-path simulation argument, similar to (but different from) the one in [14].

Suppose that we have a randomization device which allows for the following *parallel paths* simulation: let $x'_0 \sim \mu_0$ and $x''_0 \sim \zeta_0$ be two possibly dependent initial state variables for our path, generated through a simulation device so that both systems are driven by *identical* noise realizations. As a result, we have that $x'_t = F_t(x'_0, w_{[0,t-1]})$ and $x''_t = F_t(x''_0, w_{[0,t-1]})$ for some sequence of functions F_t . This allows the decoder outputs for x'_t to be applied as suboptimal decoder outputs for x''_t which result from a randomized quantization procedure for x'_t , resulting in bounds on the difference between the value functions corresponding to different initial probability measures on the state process. The explicit construction of the simulation is as follows: (i) Let $x'_{t+1} = F(x'_t, w_t)$ be a realization of the Markov chain, where w_t is an i.i.d. noise process, with probability measure ν . (ii) Let $x'_0 \sim \mu_0$ and $x''_0 \sim \zeta_0$. We can write $x''_0 = G(x'_0, w'_0)$. (iii) For every $a, b \in \mathbb{X}$, let $F_a^{-1}(b) = \{c : (a, c) \in F^{-1}(b)\}$. Then, generate \hat{w}_t according to the simulation law $P(\hat{w}_t = c | x'_t = a, x'_{t+1} = b)$ so that the following holds:

$$\begin{aligned} \nu(c) &= \sum_{a,b} P(\hat{w}_t = c | x'_t = a, x'_{t+1} = b) P(x'_t = a) \\ &\quad \times P(x'_{t+1} = b | x'_t = a) \end{aligned} \quad (15)$$

with the property that \hat{w}_t and x_t are independent for any given t . An explicit construction which satisfies the relation (15) is

$$P(\hat{w}_t = c | x'_t = a, x'_{t+1} = b) = \nu(c) \frac{1_{\{c \in F_a^{-1}(b)\}}}{\nu(F_a^{-1}(b))}.$$

(iv) With the realized \hat{w}_t , generate $x''_{t+1} = F(x''_t, \hat{w}_t)$.

Note that by construction, we also have $x'_{t+1} = F(x'_t, \hat{w}_t)$ for all t .

Lemma 2. With the simulation described as above, the distribution of x''_t is as desired, i.e., it is given by

$$P(x''_{[0,t]}) = \zeta_0(x''_0) \prod_{i=0}^{t-1} P(x''_i | x''_{i-1})$$

Observe that for $\{x'_t\}$, we can obtain the q''_t values under the optimal policy corresponding to the process with initial distribution ζ_0 . The receiver applies the quantizer policy corresponding to x'_t ; i.e., to generate u'_t we take $u'_t = \gamma'_t(q''_{[0,t]})$. This is clearly a suboptimal policy for the coding of x'_t . Thus, we are able to generate a *randomized encoder* and a *deterministic decoder* for the original problem.

Now, we are in a position to bound the expected costs uniformly for $\beta \in (0, 1)$. Consider the difference

$$\left| \inf_{\Pi} E_{\mu_0}^\Pi \left[\sum_{k=0}^{\infty} \beta^k c_0(x'_k, u'_k) \right] - \inf_{\Pi} E_{\zeta_0}^\Pi \left[\sum_{k=0}^{\infty} \beta^k c_0(x''_k, u''_k) \right] \right|.$$

Suppose without loss of generality that the second term is not greater than the first one. Note that by the previous

construction, the encoder for x'_t can simulate the encoder outputs for x''_t (under the initial probability measure ζ_0) with the optimal encoder Π'' , and apply the decoder functions $\{\gamma''_t\}$ to compute $u''_t = \gamma''_t(q''_{[0,t]})$. Thus we obtain an upper bound for the first term. Define now $\tau_{\mu_0, \zeta_0} = \min(k > 0 : x'_k = x''_k, x'_0 \sim \mu_0, x''_0 \sim \zeta_0)$. Then

$$\begin{aligned} & \left| \inf_{\Pi} E_{\mu_0}^{\Pi} \left[\sum_{k=0}^{\infty} \beta^k c_0(x'_k, u'_k) \right] - \inf_{\Pi} E_{\zeta_0}^{\Pi} \left[\sum_{k=0}^{\infty} \beta^k c_0(x''_k, u''_k) \right] \right| \\ & \leq \left| E_{P(x'_0, x''_0): x'_0 \sim \mu_0, x''_0 \sim \zeta_0}^{\Pi''} \left[\sum_{k=0}^{\infty} \beta^k (c_0(x'_k, u''_k) - c_0(x''_k, u''_k)) \right] \right| \\ & \leq 2E[\tau_{\mu_0, \zeta_0}] \|c_0\|_{\infty} \end{aligned} \quad (16)$$

where the last inequality follows from the coupling of the two chains. We can in addition show that for processes that satisfy Assumption 1, $E[\tau_{\mu_0, \zeta_0}] < \infty$.

As a result, by (16) the boundedness condition required for the existence of the ACOE is obtained. We also need to establish the equicontinuity of $V_{\beta}(\mu_0) - V_{\beta}(\zeta_0)$ for fixed ζ_0 over all β . For this, we use the Wasserstein metric to select the simulation distribution P , where we consider without loss of generality $\mathbb{X} = \{1, 2, \dots, |\mathbb{X}|\}$, viewed as a subset of \mathbb{R} :

$$W_1(\mu_0, \zeta_0) = \inf_{\substack{\mu \in \mathcal{P}(\mathbb{X} \times \mathbb{X}) \\ \mu(x, \mathbb{X}) = \mu_0(x) \\ \mu(\mathbb{X}, x) = \zeta_0(x)}} \sum \mu(x, y) |x - y|.$$

Lemma 3. For some $K < \infty$, $E[\tau_{\mu_0, \bar{\mu}_0}] < KW_1(\mu_0, \bar{\mu}_0)$.

To complete the proof of Theorem 6, note that Lemma 2 and Lemma 3 imply that $V_{\beta}(\mu_0) = J^{\beta}(\mu_0) - J^{\beta}(\zeta_0)$ is bounded from above and below, and is equicontinuous. Thus, Theorem 7 (the ACOE) holds.

2) (ii) *Rates of convergence in ϵ -optimality of finite-memory encoders:* Using the fact that $J_{\pi^*}(T) \geq J(\pi^*)$ for all T ,

$$|J_{\pi^*}(T) - J(\pi^*)| \leq \frac{K}{T},$$

where $K \leq 2K_1 \|c\|_{\infty}$.

E. A numerical implementation

To illustrate the ϵ -optimality of periodic finite-memory quantizers from Section IV-B, in this section we present implementation results for optimal zero-delay coding of a stationary Markov source over a finite time horizon T . The algorithm uses dynamic programming results to solve for the optimal sequence of quantizers.

For the simulation study, the system settings are $|\mathbb{X}| = 4$, $|\mathcal{M}| = 2$, making $|\mathcal{Q}| = 14$ (quantizers with empty cell are not counted). The transition kernel for the source is given by

$$P = \begin{bmatrix} \frac{8}{10} & \frac{1}{15} & \frac{1}{15} & \frac{1}{15} \\ \frac{1}{15} & \frac{8}{10} & \frac{1}{15} & \frac{1}{15} \\ \frac{1}{15} & \frac{1}{15} & \frac{8}{10} & \frac{1}{15} \\ \frac{1}{15} & \frac{1}{15} & \frac{1}{15} & \frac{8}{10} \end{bmatrix}$$

The P matrix as defined above admits a uniform invariant distribution. In Fig. 1, we plot the performance of the optimal

zero-delay policies for various time horizons T under the mean-squared difference distortion measure.

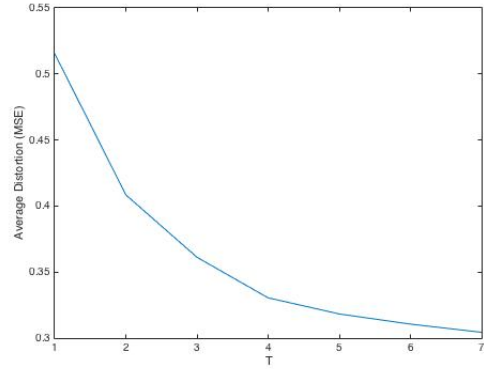


Fig. 1. Average mean square error for the finite horizon problem, with $1 \leq T \leq 7$

REFERENCES

- [1] H. S. Witsenhausen, "On the structure of real-time source coders," *Bell Syst. Tech. J.*, vol. 58, pp. 1437–1451, July/August 1979.
- [2] J. C. Walrand and P. Varaiya, "Optimal causal coding-decoding problems," *IEEE Transactions on Information Theory*, vol. 19, pp. 814–820, November 1983.
- [3] S. Yüksel and T. Başar, *Stochastic Networked Control Systems: Stabilization and Optimization under Information Constraints*. Boston, MA: Birkhäuser, 2013.
- [4] D. Teneketzis, "On the structure of optimal real-time encoders and decoders in noisy communication," *IEEE Transactions on Information Theory*, vol. 52, pp. 4017–4035, September 2006.
- [5] A. Mahajan and D. Teneketzis, "On the design of globally optimal communication strategies for real-time noisy communication with noisy feedback," *IEEE Journal on Selected Areas in Communications*, vol. 26, pp. 580–595, May 2008.
- [6] T. Linder and S. Yüksel, "On optimal zero-delay quantization of vector Markov sources," *IEEE Transactions on Information Theory*, vol. 60, pp. 2975–2991, October 2014.
- [7] H. Asnani and T. Weissman, "On real time coding with limited lookahead," *Information Theory, IEEE Transactions on*, vol. 59, no. 6, pp. 3582–3606, 2013.
- [8] T. Javidi and A. Goldsmith, "Dynamic joint source-channel coding with feedback," in *Information Theory Proceedings (ISIT), 2013 IEEE International Symposium on*. IEEE, 2013, pp. 16–20.
- [9] S. Yüksel and T. Linder, "Optimization and convergence of observation channels in stochastic control," *SIAM J. on Control and Optimization*, vol. 50, pp. 864–887, 2012.
- [10] O. Hernandez-Lerma and J. Lasserre, *Discrete-time Markov control processes*. Springer, 1996.
- [11] N. Saldi, S. Yüksel, and T. Linder, "Near optimality of quantized policies in stochastic control under weak continuity conditions," *arXiv preprint arXiv:1410.6985*, 2014.
- [12] V. S. Borkar, "Average cost dynamic programming equations for controlled markov chains with partial observations," *SIAM Journal on Control and Optimization*, vol. 39, no. 3, pp. 673–681, 2000.
- [13] —, "Dynamic programming for ergodic control of markov chains under partial observations: a correction," *SIAM Journal on Control and Optimization*, vol. 45, no. 6, pp. 2299–2304, 2007.
- [14] V. S. Borkar, S. K. Mitter, and S. Tatikonda, "Optimal sequential vector quantization of Markov sources," *SIAM J. Control and Optimization*, vol. 40, pp. 135–148, 2001.
- [15] S. Yüksel, "On optimal causal coding of partially observed Markov sources in single and multiterminal settings," *IEEE Transactions on Information Theory*, vol. 59, pp. 424–437, January 2013.