2018 ICSEE International Conference on the Science of Electrical Engineering

Coding for Noncausal Tracking

Robert Graczyk and Amos Lapidoth
Signal and Information Processing Laboratory
ETH Zurich, 8092 Zurich, Switzerland
Email: {graczyk, lapidoth}@isi.ee.ethz.ch

***Abstract*—A source generates a random sequence that is then described to a controller who wishes to employ feedback control on a given finite-state system in order for its output to closely resemble the sequence. The tension between the best achievable expected fidelity and the description length is studied in the asymptotic regime where the length of the sequence tends to infinity, with the description rate held fixed. The solution is the source-coding dual of coding for channels with states.**

I. Introduction

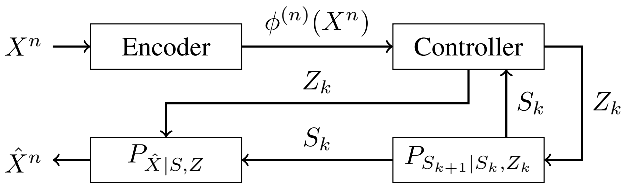

Fig. 1. The length-n source sequence X^n is described by $\phi^{(n)}(X^n)$ to a feedback controller for the system $(P_{S_{k+1}|S_k,Z_k}, P_{\hat{X}|S,Z})$, which produces the reconstruction $\hat{X}^n$.

This paper explores yet another facet of the interplay between control and information theory: it examines the extent to which a given system can be used for data compression. Concretely, we consider a setup consisting of an independent and identically distributed (IID) source, an encoder, and a time-invariant system with a feedback controller (see Figure 1). The encoder observes the source sequence X^n and describes it at a fixed rate to the controller, which–based on the encoder's description $\phi^{(n)}(X^n)$ and the causally-revealed state S_k–generates the input Z_k that minimizes the expected time-averaged distortion between the output sequence $\hat{X}^n$ produced by the system and X^n. In section III we provide a lower bound on the least description rate that allows a given time-averaged distortion and show that the bound is tight if at any time the system can be reset to its initial state. In section IV we apply the bound to a family of systems where the state evolves autonomously and the inputs only affect the system's output.

II. Problem Setup

The problem we address is how to describe a source sequence to a controller who, based on the description, wishes to employ feedback control on a given system to drive its output towards the source sequence. We model the n-length source sequence X^n as a sequence of n IID chance variables $X_1, \ldots, X_n$ that are drawn according to some given probability mass function (PMF) P_X from a finite set $\mathcal{X}$. The system is finite-state and time-invariant, with its states taking values in the set $\mathcal{S}$. It is specified by a triple

$$\mathbf{S} = (s_1, P_{S|S',Z}, P_{\hat{X}|S,Z}), \tag{1}$$

where $s_1 \in \mathcal{S}$ is its initial state; $P_{S|S',Z}$ is the conditional distribution of the time-$(k+1)$ state S_{k+1} given the time-k state S_k and the time-k input Z_k, with the latter taking values in a finite set $\mathcal{Z}$; and $P_{\hat{X}|S,Z}$ denotes the conditional PMF of the time-k output $\hat{X}_k$, which takes values in the finite set $\hat{\mathcal{X}}$, given the time-k state S_k and the time-k input Z_k.

A rate-R blocklength-n description is a mapping

$$\phi^{(n)} \colon \mathcal{X}^n \to \mathcal{I} \tag{2}$$

where

$$\mathcal{I} \triangleq \{1, 2, \ldots, 2^{nR}\} \tag{3}$$

is the index set. The controller $\psi^{(n)}$ associates with each index $i \in \mathcal{I}$ an n-tuple

$$\psi^{(n)}(i) = (\psi_1^{(n)}(i), \ldots, \psi_n^{(n)}(i)) \tag{4}$$

of n feedback rules, where the k-th rule $\psi_k^{(n)}(i)$ is a function mapping the states $S_1, \ldots, S_k$ to the time-k input Z_k. Rather than writing $Z_k = \psi_k^{(n)}(i)\big(S_1, \ldots, S_k\big)$, we sometimes prefer

$$Z_k = \psi_k^{(n)}\big(i, S_1, \ldots, S_k\big), \tag{5}$$

or

$$Z_k = \psi_k^{(n)}\big(i, S^k\big), \tag{6}$$

where we adopt the convention that S^k stands for $S_1, \ldots, S_k$, and S_k^n stands for $S_k, S_{k+1}, \ldots, S_n$. We also introduce $\mathcal{T}_k$ to denote the collection of mappings from $\mathcal{S}^k$ to $\mathcal{Z}$, i.e.,

$$\mathcal{T}_k = \mathcal{Z}^{(\mathcal{S}^k)}. \tag{7}$$

With this notation

$$\psi_k^{(n)}(i) \in \mathcal{T}_k, \quad k \in [1:n], \quad i \in \mathcal{I}. \tag{8}$$

The system's time-k output $\hat{X}_k$ is drawn conditionally on the time-k state S_k and input Z_k according to $P_{\hat{X}|S,X}(\cdot|S_k, Z_k)$, and the next state S_{k+1} according to $P_{\hat{X}|S,X}(\cdot|S_k, Z_k)$.

To quantify how closely the system's output sequence $\hat{X}^n$ resembles X^n, we fix a bounded distortion measure

$$d \colon \mathcal{X} \times \hat{\mathcal{X}} \to \mathbb{R}_{\geq 0} \tag{9}$$

and define the time-averaged distortion

$$\bar{D}_n \triangleq \frac{1}{n}\sum_{k=1}^{n} d(X_k, \hat{X}_k). \tag{10}$$

Its expectation over the source sequence and the system's behavior is denoted $\mathbb{E}[\bar{D}_n]$. It is determined by the description $\phi^{(n)}$ and feedback-controller $\psi^{(n)}$, the source distribution P_X, and the system $\mathbf{S}$. More concretely, $\mathbb{E}[\bar{D}_n]$ is computed with respect to the distribution

$$P_{S_1} P_X^n P_{T^n|X^n} P_{S_2^n|S_1,T^n} P_{\hat{X}^n|S^n,T^n}, \tag{11}$$

where P_{S_1} is the PMF of the initial state

$$P_{S_1} = \delta_{s_1}, \tag{12}$$

with δ_{s_1} denoting a PMF that places all the weight on s_1, i.e., $\mathbb{P}[S_1 = s_1] = 1$; where $T^n = (T_1, \ldots, T_n)$, with $T_k \in \mathcal{T}_k$ being the random feedback rule mapping S^k to $Z_k = \psi_k^{(n)}(\phi^{(n)}(X^n))(S^k)$, namely

$$T_k = \psi_k^{(n)}(\phi^{(n)}(X^n)); \tag{13}$$

and where the distribution of the state sequence S_2^n conditioned on the initial state S_1 and the feedback-controller T^n can be expressed as

$$P_{S_2^n|S_1,T^n}(s_2^n|s_1, \psi^{(n)}(i)) = \prod_{k=1}^{n-1} P_{S|S',Z}\big(s_{k+1}\big|s_k, \psi_k^{(n)}(i, s^k)\big) \tag{14}$$

with $\psi^{(n)}(i) = (\psi_1^{(n)}(i), \ldots, \psi_n^{(n)}(i))$ denoting the feedback rules when we substitute i for $\phi^{(n)}(x^n)$. The distribution of the output sequence $\hat{X}^n$ conditioned on (S^n, T^n) is

$$P_{\hat{X}^n|S^n,T^n}(\hat{x}^n|s^n, \psi^{(n)}(i)) = \prod_{k=1}^{n} P_{\hat{X}|S,Z}\big(\hat{x}_k\big|s_k, \psi_k^{(n)}(i, s^k)\big). \tag{15}$$

The source sequence X^n, the description $\phi^{(n)}$, the controller $\psi^{(n)}$, and the system $\mathbf{S}$ thus induce a joint distribution on the states S^n and the output sequence $\hat{X}^n$.

Given a maximal-allowed expected time-averaged distortion D, we seek the smallest rate R for which we can find a sequence of descriptions $(\phi^{(n)})$ and controllers $(\psi^{(n)})$ satisfying

$$\limsup_{n\to\infty} \mathbb{E}[\bar{D}_n] \le D, \tag{16}$$

where $\mathbb{E}[\bar{D}_n]$ is the expected time-averaged distortion induced by applying the description $\phi^{(n)}$ and the feedback-controller $\psi^{(n)}$ to $(X^n, \mathbf{S})$. This rate is denoted $\mathrm{RT}(P_X, \mathbf{S}, d, D)$ or $\mathrm{RT}(D; s_1)$, and we refer to the mapping $D \mapsto \mathrm{RT}(D; s_1)$ as the *rate-trackability function*.

Henceforth we assume that for every $\delta > 0$ there is a sequence of rate-$\log|\mathcal{X}|$ descriptions (allowing the sequence X^n to be described to the controller losslessly) with corresponding feedback-controllers for which

$$\limsup_{n\to\infty} \mathbb{E}[\bar{D}_n] \le \delta, \tag{17}$$

and hence $\mathrm{RT}(D; s_1)$ is well-defined for $D > 0$. If (17) cannot be satisfied for $\delta = 0$, we define $\mathrm{RT}(0; s_1) \triangleq \infty$. This assumption entails no loss of generality, since we can think of it as restricting the domain of $\mathrm{RT}(\cdot\,; s_1)$ to $(D_{\text{inf}}, +\infty]$, where D_{inf} denotes the infimum over all $D \ge 0$ for which there exists a sequence of rate-$\log|\mathcal{X}|$ descriptions with corresponding feedback-controllers satisfying

$$\limsup_{n\to\infty} \mathbb{E}[\bar{D}_n] \le D. \tag{18}$$

Characterizing the rate-trackability function is the objective of this paper. In section III we derive a lower bound on $\mathrm{RT}(D; s_1)$ and show that it holds with equality under the condition that to each state $s' \in \mathcal{S}$ there corresponds some "reset" input $r(s') \in \mathcal{Z}$ that drives the state to s_1 in the sense that

$$P_{S|S',Z}(s|s', r(s')) = \delta_{s_1}(s) \quad \forall s \in \mathcal{S}. \tag{19}$$

In section IV we establish a single-letter expression for $\mathrm{RT}(D; s_1)$ for systems where the time-k input Z_k only influences the output $\hat{X}_k$, and has no impact on the distribution of the next state S_{k+1}.

III. Characterization of $\mathrm{RT}(D; s_1)$

In order to characterize $\mathrm{RT}(D; s_1)$, it is useful to define the following auxiliary multi-letter R-D problem: A source sequence $\tilde{X}_1, \ldots, \tilde{X}_m$ is drawn IID with $\tilde{X}_j$ taking values in the set $\mathcal{X}^n$ according to the n-fold product distribution P_X^n, where P_X is the PMF governing the source law in the original problem. The sequence $\tilde{X}^m$ is described using a mapping

$$\tilde{\phi}^{(m,n)} \colon (\mathcal{X}^n)^m \to \{1, \ldots, 2^{nmR}\} \qquad \tilde{X}^m \mapsto i \tag{20}$$

and a reconstructor $\tilde{\psi}^{(m,n)}$ that produces the sequence

$$\tilde{\psi}^{(m,n)}(i) = \tilde{\psi}_1^{(m,n)}(i), \ldots, \tilde{\psi}_m^{(m,n)}(i), \tag{21}$$

where $\tilde{\psi}_j^{(m,n)}(i)$ is, for every $j \in [1:m]$, an n-tuple of the form

$$\tilde{\psi}_j^{(m,n)}(i) = \Big(\tilde{\psi}_{j,1}^{(m,n)}(i), \ldots, \tilde{\psi}_{j,n}^{(m,n)}(i)\Big) \tag{22}$$

with

$$\tilde{\psi}_{j,\nu}^{(m,n)}(i) \colon \mathcal{S}^\nu \to \mathcal{Z}, \quad \nu \in [1:n], \tag{23}$$

where $\mathcal{S}$ and $\mathcal{Z}$ are the state and input alphabets of the original problem. Recalling (7), we can rewrite (23) as

$$\tilde{\psi}_{j,\nu}^{(m,n)}(i) \in \mathcal{T}_\nu, \quad \nu \in [1:n] \tag{24}$$

and

$$\tilde{\psi}_j^{(m,n)}(i) \in \mathcal{T}_1 \times \cdots \times \mathcal{T}_n. \tag{25}$$

Note that in the auxiliary R-D problem the feedback-controller looks back at most n states.

We next define a distortion measure $\tilde{d}_n$ between $\tilde{x} \in \mathcal{X}^n$ and $\tilde{\psi} \in \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$. We imagine that the n-length source sequence is $\tilde{x}$ and that the first n feedback rules specified by $\tilde{\psi}$ are applied to $\mathbf{S}$ (which starts at s_1), and we define

$$\tilde{d}_n(\tilde{x}, \tilde{\psi}) \triangleq \frac{1}{n} \sum_{k=1}^{n} \mathbb{E}[d(X_k, \hat{X}_k) | X_k = \tilde{x}_k], \tag{26}$$

where the expectation is computed with respect to

$$P_{S_1} P_{S_2^n | S_1, \tilde{\psi}} P_{\hat{X}^n | S^n, \tilde{\psi}}, \tag{27}$$

with $P_{S_1}(s) = \delta(s_1)$, irrespective of the past controls and the system's behavior.

We denote this auxiliary R-D problem $\mathbf{P}_n(s_1)$ and its corresponding R-D function $R_n(\cdot\,; s_1)$, so normalized by n

$$R_n(D; s_1) = \min_{P_{T^n|X^n} : \mathbb{E}[\tilde{d}_n(X^n, T^n)] \le D} \frac{1}{n} I(X^n; T^n). \tag{28}$$

We extend the definition of this rate-distortion problem to systems where the initial state is not deterministic, but drawn from some PMF P_{S_1}, in which case the problem and its corresponding R-D function are denoted $\mathbf{P}_n(P_{S_1})$ and $R_n(\cdot\,; P_{S_1})$, respectively.

Theorem 1. *For all $D > 0$ the rate-trackability function satisfies*

$$\mathrm{RT}(D; s_1) \ge \liminf_{n \to \infty} R_n(D; s_1). \tag{29}$$

If the system has "reset" inputs satisfying (19)*, then for all $D > 0$*

$$\mathrm{RT}(D; s_1) = \liminf_{n \to \infty} R_n(D; s_1). \tag{30}$$

Proof. Proving (29) is conceptually simple. The proof is based on the observation that any blocklength-n pair $(\phi^{(n)}, \psi^{(n)})$ for the original problem can be viewed as "scalar quantizer" for the $\mathbf{P}_n(s_1)$ problem. Its rate must therefore be at least as large as that of the best (asymptotic) vector quantizer on $\mathbf{P}_n(s_1)$, namely $R_n(D; s_1)$.

To formalize this argument, note that by definition of the rate-trackability function we can, for every n, find some $m(n)$ with a description $\phi^{(m(n))}$ and a feedback-controller $\psi^{(m(n))}$ satisfying

$$\mathbb{E}[\bar{D}_{m(n)}] \le D + \frac{1}{n}. \tag{31}$$

Applying the pair $(\phi^{(m(n))}, \psi^{(m(n))})$ block-wise in the problem $\mathbf{P}_{m(n)}(s_1)$ shows that there is a rate-$\mathrm{RT}(D; s_1)$ coding scheme with a time-averaged distortion of at most $D + 1/n$. Observe that by this argument and since $\mathrm{RT}(D; s_1)$ is upper-bounded by $\log |\mathcal{X}|$ for every $D > 0$, so is $R_n(D; s_1)$ for all sufficiently large n. Next, by definition of the rate-distortion function,

$$\mathrm{RT}(D; s_1) \ge R_{m(n)}(D + 1/n; s_1), \tag{32}$$

and we note that monotonicity and continuity of $R_{m(n)}$ imply the existence of a small nonnegative constant δ_n satisfying

$$R_{m(n)}(D + 1/n; s_1) = R_{m(n)}(D; s_1) - \delta_n. \tag{33}$$

We now require the following lemma.

Lemma 1. *Consider a sequence of convex decreasing functions (f_n) on the domain $\mathbb{R}_{\ge 0}$, and assume that every f_n is upper-bounded by $C \in \mathbb{R}_{\ge 0}$. Then, for every sequence (δ_n) satisfying $\delta_n \downarrow 0$ as $n \to \infty$,*

$$\lim_{n \to \infty} f_n(x) - f_n(x + \delta_n) = 0 \quad \forall x > 0. \tag{34}$$

Proof. Suppose not. Then there is some $\epsilon > 0$ such that for infinitely many n

$$f_n(x) - f_n(x + \delta_n) \ge \epsilon \tag{35}$$

and hence, since f_n is assumed to be convex and decreasing,

$$f_n(x') \ge f(x) + (\epsilon/\delta_n)(x - x') \quad \forall x' \in [0, x], \tag{36}$$

contradicting $f_n(x') \le C$ for large n. □

With Lemma 1 at hand, we conclude the proof of (29) by noting that

$$\begin{aligned} \mathrm{RT}(D; s_1) &\ge \liminf_{n \to \infty} R_{m(n)}(D; s_1) - \delta_n && (37) \\ &= \liminf_{n \to \infty} R_{m(n)}(D; s_1) && (38) \\ &\ge \liminf_{n \to \infty} R_n(D; s_1), && (39) \end{aligned}$$

where we have applied Lemma 1 in the first equality to argue that $(R_{m(n)}(D; s_1) - R_{m(n)}(D + 1/n; s_1)) \to 0$.

We next prove (30) by showing that under the "reset assumption", for every $\delta > 0$ there exists a sequence of descriptions of rate at most $R' + \delta$, where

$$R' \triangleq \liminf_{n \to \infty} R_n(D; s_1), \tag{40}$$

with corresponding feedback-controllers satisfying

$$\limsup_{n \to \infty} \mathbb{E}[\bar{D}_n] \le D. \tag{41}$$

We first choose a positive integer n such that

$$R_n(D; s_1) \le R' + \delta/2, \tag{42}$$

and observe that monotonicity and continuity of the rate-distortion function $R_n(\cdot\,; s_1)$ imply the existence of a small positive constant δ_n satisfying

$$R_n(D-\delta_n;s_1) = R' + \delta. \tag{43}$$

Invoking Lemma 1, we see that δ_n cannot be too small for large n, in particular we can choose n such that in addition to satisfying (42), n will be large enough so that

$$\delta_n \geq \frac{2d_{\max}}{n}, \tag{44}$$

where $d_{\max}$ denotes the maximum element in the range of d.

We next observe that by definition of the rate-distortion function $R_n(D-\delta_n;s_1)$ we can find for all sufficiently large m a description

$$\tilde{\phi}^* : (\mathcal{X}^n)^m \to \{1,\ldots,2^{nm(R+\delta)}\} \tag{45}$$

and a reconstructor $\tilde{\psi}^*$, that given $i \in \{1,\ldots,2^{nm(R+\delta)}\}$, produces the feedback-controller

$$\tilde{\psi}^*(i) = (\tilde{\psi}_1^*(i),\ldots,\tilde{\psi}_m^*(i)) \tag{46}$$

with

$$\tilde{\psi}_j^*(i) \in \mathcal{T}_1 \times \cdots \times \mathcal{T}_n \quad \forall j \in [1:m] \tag{47}$$

and with the resulting expected time-averaged distortion between $(X^n)^m$ and its reconstruction $\tilde{\psi}_1^*(\tilde{\phi}^*(X^{nm})),\ldots,\tilde{\psi}_m^*(\tilde{\phi}^*(X^{nm}))$, namely,

$$\mathbb{E}[\tilde{D}_m] \triangleq \frac{1}{m}\sum_{j=1}^{m}\mathbb{E}[\tilde{d}_n(X_{(j-1)n+1}^{jn},\tilde{\psi}_j^*(\tilde{\phi}^*(X^{nm})))], \tag{48}$$

being smaller than $D-\delta_n/2$. We next derive a new feedback rule $\tilde{\psi}'(i)$ from $\tilde{\psi}^*(i)$ by replacing the last mapping in the tuple $\tilde{\psi}_j^*(i)$ with one producing the "reset" input, i.e.,

$$\tilde{\psi}_j'(i) \triangleq \left(\tilde{\psi}_{j,1}^*(i),\ldots,\tilde{\psi}_{j,n-1}^*(i), s^n \mapsto r(s_n)\right). \tag{49}$$

The idea is now to apply $\tilde{\phi}^*$ and $\tilde{\psi}'$ as a description and feedback-controller pair to the source sequence $(X^n)^m$ and the system $\mathbf{S}$. Observe that (49) guarantees that $S_k = s_1$ holds for $k = n+1, 2n+1, \ldots, (m-1)n+1$; and since

$$\mathbb{E}[\tilde{D}_m] \leq D - \frac{\delta_n}{2} \tag{50}$$

and the change in (49) increases the expected time-averaged distortion by no more than $d_{\max}/n$, we have demonstrated the existence of a description $\tilde{\phi}^*$ of rate $R'+\delta$ and a corresponding feedback-controller $\tilde{\psi}'$ that satisfies

$$\mathbb{E}[\bar{D}_{nm}] \leq \mathbb{E}[\tilde{D}_m] + \frac{d_{\max}}{d} \leq D - \frac{\delta_n}{2} + \frac{d_{\max}}{n} \leq D, \tag{51}$$

where the last inequality follows from (44).
Since (51) holds for every sufficiently large m, we conclude that there is a sequence of descriptions and feedback-controllers of rate at most $R'+\delta$ satisfying

$$\limsup_{l\to\infty}\mathbb{E}[\bar{D}_l] \leq D, \tag{52}$$

which follows because even if l is not of the form $l = nm$, we can express it as $l = nm + r$ with $r \leq n$, apply $\tilde{\phi}^*$ and $\tilde{\psi}'$ until time nm, and use an arbitrary feedback rule on the remaining r time steps. Since the distortion measure d is assumed to be bounded, this does asymptotically not change the time-averaged distortion.

□

In the following section we present an application of Theorem 1 by studying the rate-trackability function for autonomous, ergodic systems.

IV. Coding for Autonomous, Ergodic Systems

We conclude the paper by considering systems whose inputs do not influence their state evolution.

Theorem 2. *If the state-evolution law of* $\mathbf{S}$ *is of the form*

$$P_{S|S',Z} = P_{S|S'}, \tag{53}$$

and $P_{S|S'}$ *is irreducable and aperiodic, then for all* $D > 0$

$$\mathrm{RT}(D;s_1) = \min_{P_{T|X}:\mathbb{E}[d_S(X,T)]\leq D} I(T;X), \tag{54}$$

where $d_S(x,\psi)$ *is defined as*

$$d_S(x,\psi) \triangleq \sum_s \pi(s)\sum_{\hat{x}} P_{\hat{X}|S,Z}(\hat{x}|s,\psi(s))d(x,\hat{x}), \tag{55}$$

with π *denoting the stationary distribution induced by* $P_{S|S'}$ *and with the chance variable* T *taking values in* $\mathcal{T}_1$.

Proof. Intuitively speaking Theorem, 2 states that the following strategy is optimal: Wait until $\mathbb{P}[S_k = \cdot]$ is sufficiently close to the stationary distribution and then treat the situation as a single-letter R-D problem with a new distortion measure d_S and reconstruction alphabet $\mathcal{T}_1$.

To formalize this idea, let $\mathbf{P}^*$ denote the R-D problem for the source sequence $X^n \sim P_X^n$ and the single-letter distortion d_S of (55), and let $R^*(\cdot)$ be its R-D function. We now argue that for every $\delta > 0$ there exists a sequence of descriptions $(\phi^{(n)})$ and feedback-controllers $(\psi^{(n)})$ of rate $R^*(D)+\delta$ satisfying

$$\limsup_{n\to\infty}\mathbb{E}[\bar{D}_n] \leq D. \tag{56}$$

To see this, first observe that since $P_{S|S'}$ is irreducable and aperiodic, there exists for every $\epsilon > 0$ a positive integer g such that for every $s \in \mathcal{S}$ and all $k \geq g$

$$|\,\mathbb{P}[S_k = s] - \pi(s)| \leq \epsilon. \tag{57}$$

Next, let

$$\phi^{*(n)} : \mathcal{X}^n \to \{1,\ldots,2^{nR^*(D-\delta')}\}, \quad n \in \mathbb{N} \tag{58}$$

be a sequence of descriptions and

$$\varphi^{*(n)}\colon \{1,\ldots,2^{nR^*(D-\delta')}\} \to \mathcal{T}_1^n, \quad n \in \mathbb{N} \tag{59}$$

be a sequence of decoders of rate $R^*(D-\delta')$ for the rate-distortion problem P^* with $\delta' > 0$ chosen such as to satisfy $R^*(D-\delta') \leq R^*(D) + \delta$ (the existence of a suitable δ' follows from the continuity of the rate-distortion function). For n large enough, the expected time-averaged distortion as measured by d_S and induced by $(\phi^{*(n)}, \varphi^{*(n)})$ is bounded from above by $D - \delta'/2$. We now construct a new description $\phi^{(n)}$ of rate $R^*(D-\delta')$ that ignores X^g and encodes X_{g+1}^n according to $\phi^{*(n-g)}(X_{g+1}^n)$. The corresponding feedback-controller $\psi^{(n)}$ is constructed by applying an arbitrary feedback rule until time g and then applying the feedback-controller $\varphi^{*(n-g)}(\phi^{*(n-g)}(X_{g+1}^n))$. The expected distortion incurred until time g is constant and finite, and hence does not influence the asymptotic discussion. Due to (57), we further observe that computing the expected time-k distortion $\mathbb{E}[d(X_k, \hat{X}_k)]$ with respect to the true state distribution $\mathbb{P}[S_k = \cdot]$ instead of π for $k \geq g$ yields a value no larger than $\mathbb{E}[d_S(X_k; T_k)] + \epsilon d_{\max}$. Thus, by choosing g such that $\epsilon \leq \delta'/(2d_{\max})$, we can indeed construct a sequence $(\phi^{(n)})$ of rate-$(R^*(D)+\delta)$ descriptions with corresponding feedback-controllers $(\psi^{(n)})$ satisfying $\limsup_{n\to\infty} \mathbb{E}[\bar{D}_n] \leq D$. This concludes the achievability part of the proof.

For the converse we invoke Theorem 1 and show that $R^*(D)$ constitutes a lower bound on (29). To that end we first convice ourselves that in order to minimize (29), the support of T^n can be restricted to $\mathcal{T}_1^n$, i.e., the time-k feedback rule T_k takes into account only S_k (instead of S^k). To see this, consider any conditional distribution $P_{T^n|X^n}$ satisfying

$$\mathbb{E}[\tilde{d}_n(X^n, T^n)] \leq D. \tag{60}$$

From T^n we generate a new feedback-controller T'^n as follows: For every $k \in [2:n]$ and every state $s_k \in \mathcal{S}$ we independently generate a simulation of the past states S^{k-1}, denoted $S'^{k-1}_{s_k}$, according to $P_{S^{k-1}|S_k}(\cdot|s_k)$. Then the feedback rules

$$(s^{k-1}, s_k) \mapsto T_k(s^{k-1}, s_k) \tag{61}$$

are replaced by

$$(s^{k-1}, s_k) \mapsto T_k(S'^{k-1}_{s_k}, s_k) \tag{62}$$

for all $s^k \in \mathcal{S}^k$, $k \in [2:n]$. Observe that

$$\mathbb{E}[d(X_k, T'_k(S_k))] = \mathbb{E}[d(X_k, T_k(S'^{k-1}_{S_k}, S_k))] \tag{63}$$

$$= \mathbb{E}[d(X_t, T_k(S^{k-1}, S_k))], \tag{64}$$

where the second equality holds because the inputs to the system do not influence its state. This condition is necessary, as in general replacing the argument S^{k-1} of T_k with $S'^{k-1}_{S_k}$ changes the distribution of S_{k+1}.

From (64) we follow that $P_{T'^n|X^n}$ satisfies

$$\mathbb{E}[\tilde{d}_n(X^n, T'^n)] = \mathbb{E}[\tilde{d}_n(X^n, T^n)] \leq D, \tag{65}$$

and since $X^n \to T^n \to T'^n$ is a Markov chain, the data processing inequality implies

$$I(X^n; T^n) \geq I(X^n; T'^n), \tag{66}$$

and hence we may indeed assume that $T^n \in \mathcal{T}_1^n$ when considering (29). To wrap up the proof, we argue that assuming $T^n \in \mathcal{T}_1^n$,

$$\liminf_{n\to\infty} \min_{P_{T^n|X^n}:\mathbb{E}[\tilde{d}_n(X^n,T^n)]\leq D} \frac{1}{n} I(X^n; T^n) \geq R^*(D). \tag{67}$$

This will follow immediately from the lemma below that we state without proof.

Lemma 2. *Let $X^n \sim P_X^n$ for some PMF P_X with finite support $\mathcal{X}$, and let (d_n) be a sequence of distortion measures on the finite domain $\mathcal{X} \times \hat{\mathcal{X}}$. Suppose there is a distortion measure d such that $d_n \to d$ as $n \to \infty$ in the sense that*

$$\lim_{n\to\infty} \max_{x,\hat{x}} |d_n(x,\hat{x}) - d(x,\hat{x})| = 0. \tag{68}$$

Then, for $D \in \mathbb{R}_{\geq 0}$,

$$\liminf_{n\to\infty} \min_{P_{\hat{X}^n|X^n}:\frac{1}{n}\sum_{i=1}^n \mathbb{E}[d_i(X_i,\hat{X}_i)]\leq D} \frac{1}{n} I(X^n; \hat{X}^n) \tag{69}$$

$$\geq \min_{P_{\hat{X}|X}:\mathbb{E}[d(X,\hat{X})]\leq D} I(X; \hat{X})$$

We now observe that (67) follows from Lemma 2 with

$$d_n(x,\psi) = \sum_s \mathbb{P}[S_n = s] \sum_{\hat{x}} P_{\hat{X}|S,Z}(\hat{x}|s,\psi(s)) d(x,\hat{x}) \tag{70}$$

and $d = d_S$; the source and reconstruction alphabets are $\mathcal{X}$ and $\mathcal{T}_1$, respectively.

□

As a closing remark of this section, we see that for a system where $S^n \sim \pi^n$, the problem corresponds to an indirect R-D problem as studied by Witsenhausen in [3]; by Dobrushin and Tsybakov in [1]; and by Wolf and Ziv in [4]. We also mention that the channel-coding dual of our problem is considered by Gallager in [2, Chapter 4, p. 97 – 111].

References

[1] R. L. Dobrushin and B. S. Tsybakov. "Information transmission with additional noise". In: *IRE Trans. Inf. Theory* 18 (1962), pp. 293–304.

[2] R. Gallager. *Information Theory and Reliable Communication*. 1968.

[3] H. Witsenhausen. "Indirect rate distortion problems". In: *IEEE Trans. Inf. Theory* 26 (Sept. 1980), pp. 518–521.

[4] J. K. Wolf and J. Ziv. "Transmission of noisy information to a noisy receiver with minimum distortion". In: *IEEE Trans. Inf. Theory* 16 (July 1970), pp. 406–411.