The Error Exponent of Variable-Length Codes Over Markov Channels With Feedback

Giacomo Como, Serdar Yüksel, and Sekhar Tatikonda, *Member, IEEE*

Abstract—The error exponent of Markov channels with feedback is studied in the variable-length block-coding setting. Burnashev's classic result is extended to finite-state ergodic Markov channels. For these channels, a single-letter characterization of the reliability function is presented, under the assumption of full causal output feedback, and full causal observation of the channel state both at the transmitter and at the receiver side. Tools from stochastic control theory are used in order to treat channels with intersymbol interference (ISI). Specifically, the convex-analytic approach to Markov decision processes is adopted in order to handle problems with stopping time horizons induced by variable-length coding schemes.

Index Terms—Channel coding with feedback, error exponents, finite-state Markov channels, Markov decision processes, variable-length block codes.

I. INTRODUCTION

THE role of feedback in channel coding is a long studied problem in information theory. In 1956, Shannon [28] proved that noiseless causal output feedback does not increase the capacity of a discrete memoryless channel (DMC). Feedback, though, can help in improving the tradeoff between reliability and delay of DMCs at rates below capacity. This tradeoff is traditionally measured in terms of error exponents; in fact, since Shannon's work, much research has focused on studying error exponents of channels with feedback. Burnashev [6] found a simple exact formula for the reliability function (i.e., the highest achievable error exponent) of a DMC with perfect causal output feedback in the variable-length block-coding setting. The present paper deals with a generalization of Burnashev's result to a certain class of channels with memory. Specifically, we shall provide a simple single-letter characterization of the reliability function of finite-state Markov channels (FSMCs), in the general case when intersymbol interference (ISI) is present. Under mild ergodicity assumptions, we will

G. Como was with the Department of Mathematics, Politecnico di Torino, Torino, Italy and the Department of Electrical Engineering, Yale University, New Haven, CT. He is now with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: giacomo@mit.edu).

S. Yüksel was with the Department of Electrical Engineering, Yale University, New Haven, CT. He is now with the Department of Mathematics and Statistics, Queen's University, Kingston, ON K7L 3N6, Canada (e-mail: yuksel@mast.queensu.ca).

S. Tatikonda is with the Department of Electrical Engineering, Yale University, New Haven, CT 06511 USA (e-mail: sekhar.tatikonda@yale.edu).

Communicated by G. Kramer, Associate Editor for Shannon Theory.

Digital Object Identifier 10.1109/TIT.2009.2015989

prove that, when one is allowed variable-length block coding with perfect causal output feedback and causal state knowledge both at the transmitter and at the receiver end, the reliability function has the form

$$E_B(R) = D\left(1 - \frac{R}{C}\right), \quad R \in (0, C). \tag{1}$$

In (1), R denotes the transmission rate, measured with respect to the average number of channel uses. The capacity C and the coefficient D are quantities which will be defined as solutions of finite-dimensional optimization problems involving the stochastic kernel describing the FSMC. The former will turn out to equal the maximum, over all choices of the channel input distributions as a function of the channel state, of the conditional mutual information between channel input and the pair of channel output and next channel state given the current state, whose marginal distribution coincides with the induced ergodic state measure (see (6)). The latter will instead equal the average, with respect to the induced ergodic state measure, of the Kullback–Leibler information divergence between the joint channel output and next state distributions associated to the pair of most distinguishable choices of a channel input symbol as a function of the current state (see (14)).

The problem of characterizing error exponents of memoryless channels with feedback has been addressed in the information theory literature in a variety of different frameworks. Particularly relevant are the choice of block versus continuous transmission, the possibility of allowing variable-length coding schemes, and the way delay is measured. In fact, much more than in the non-feedback case, these choices lead to very different results for the error exponent of DMCs, albeit not altering the capacity value. In continuous transmission systems information bits are introduced at the encoder, and later decoded, individually. Continuous transmission with feedback was considered by Horstein [19], who was probably the first showing that variable-length coding schemes can provide larger error exponents than fixed-length ones. Recently, continuous transmission with fixed delay has attracted renewed attention in the context of anytime capacity [27]. In this paper, however, we shall restrict ourselves to block transmission, which is the framework considered by the largest part of the previous literature.

In block transmission systems, the information sequence is partitioned into blocks of fixed length which are then encoded into channel input sequences. When there is no feedback, these sequences need to be of a predetermined, fixed length. When there is feedback, instead, the availability of common information shared between transmitter and receiver makes it possible to use variable-length schemes. Here, the transmission time is allowed to dynamically depend on the channel output

sequence. It is known that exploiting the possibility of using variable-length block-coding schemes guarantees high gains in terms of the attainable error exponent. In fact, Dobrushin [13] showed that the sphere-packing upper bound still holds for fixed-length block-coding schemes over symmetric DMCs even when perfect output feedback is causally available to the encoder (a generalization to nonsymmetric DMCs was addressed first in [18] and then in [31]). Even though fixed-length block-coding schemes with feedback have been studied (see [12], [39]), the aforementioned results pose severe constraints on the performance such schemes can achieve. Moreover, no closed form for the reliability function at all rates is known for fixed-length block coding with feedback, but for the very special class of symmetric DMCs with positive zero-error capacity (cf. [9, p. 199]). It is worth to mention that the situation can be different for infinite-input channels. For the additive white Gaussian noise channel (AWGNC) with average power constraint, Shalkwijk and Kailath [30] proved that a doubly exponential error rate is achievable by fixed-length block codes. However, when a peak power constraint to the input of an AWGNC is added, then this phenomenon disappears, as shown in [37]. At the same time it is also known that, if variable-length coding schemes are allowed, then the sphere-packing exponent can be beaten even when no output feedback is available but for a single bit used only once. This situation is traditionally referred to as decision feedback and was studied in [14] (see also [9, p. 201]).

A very simple exact formula was found by Burnashev [6] for the reliability function of DMCs with full causal output feedback in the case of variable-length block-coding schemes. Burnashev's analysis combined martingale theory arguments with more standard information-theoretic tools. It is remarkable that in this setting the reliability function is known, in a very simple form, at any rate below capacity, in sharp contrast to what happens in most channel-coding problems for which the reliability function can be exactly evaluated only at rates close to capacity. Another important point is that the Burnashev exponent of a DMC can dramatically exceed the sphere-packing exponent; in particular, it approaches capacity with nonzero slope.

Thus, variable-length block-coding appears to be a natural setting for transmission over channels with feedback. In fact, it has already been considered by many authors after [6]. A simple two-phase iterative scheme achieving the Burnashev exponent was introduced by Yamamoto and Itoh in [38]. More recently, low-complexity variable-length block-coding schemes with feedback have been proposed and analyzed in [25]. The works [33] and [34] dealt with universality issues, addressing the question whether the Burnashev exponent can be achieved without exact knowledge of the statistics of the channel but only knowing it belongs to a certain class of DMCs. In [2], a simplification of Burnashev's original proof [6] is proposed, while [23] is concerned with the characterization of the reliability function of DMCs with feedback and cost constraints. In [26], low-complexity schemes for FSMCs with feedback are proposed. However, to the best of our knowledge, no extension of Burnashev's theorem to channels with memory has been considered so far.

The present work deals with a generalization of Burnashev's result to FSMCs. As an example, channels with memory, and FSMCs in particular, model transmission problems where fading is an important component, like in wireless communication. Information-theoretic limits of FSMCs both with and without feedback have been widely studied in the literature: we refer to the classic textbooks [15], [36] and references therein for overview of the available literature (see also [16]). It is known that the capacity is strongly affected by the hypothesis about the nature of the channel state information (CSI) both available at the transmitter and at the receiver side. In particular, while output feedback does not increase the capacity when the state is causally observable both at the transmitter and at the receiver side (see [32] for a proof, first noted in [28]), it generally does so for different information patterns. In particular, when the channel state is not observable at the transmitter, it is known that feedback may help improving capacity by allowing the encoder to estimate the channel state [32]. However, in this paper only the case when the channel state is causally observed both at the transmitter and at the receiver end will be considered. Our choice is justified by the aim to separate the study of the role of output feedback in channel state estimation from its effect in allowing better reliability versus delay tradeoffs for variable-length block-coding schemes.

In [32], a general stochastic control framework for evaluating the capacity of channels with memory and feedback has been introduced. The capacity has been characterized as the solution of a dynamic-programming average-cost optimality equation. Existence of a solution to such an equation implies information stability [17]. Also lower bounds à la Gallager to the error exponents achievable with fixed-length coding schemes are obtained in [32]. In the present paper, we follow a similar approach in order to characterize the reliability function of variable-length block-coding schemes with feedback. Such an exponent will be characterized in terms of solutions to certain Markov decision processes (MDPs). The main new feature posed by variable-length schemes is that we have to deal with average cost optimality problems with a stopping time horizon, for which standard results in MDP theory cannot be used directly. We adopt the convex-analytic approach [4] and use Hoeffding–Azuma inequality in order to prove a strong uniform convergence result for the empirical measure process. (See [21] for results of a similar flavor in the finite-state finite-action setting.) This allows us to find sufficient conditions on the tails of a sequence of stopping times for the solutions of the average-cost optimality problems to asymptotically converge to the solution of the corresponding infinite-horizon problems, for which stationary policies are known to be optimal.

The rest of this paper is organized as follows. In Section II, causal feedback variable-length block-coding schemes for FSMCs are introduced, and capacity and reliability function are defined as solution of optimization problems involving the stochastic kernel describing the FSMC. The main result of the paper is then stated in Theorem 1. In Section III, we prove an upper bound to the reliability function of FSMCs with feedback and variable-length block coding. The main result of that section is contained in Theorem 2 which generalizes Burnashev's result [6]. Section IV is of a technical nature and deals with Markov decision processes with stopping time horizons. Some stochastic control techniques are reviewed and the main result

is contained in Theorem 3 which is then used to prove that the bound of Theorem 2 asymptotically coincides with the reliability function (1). In Section V, a sequence of simple iterative schemes based on a generalization of Yamamoto–Itoh's idea [38] is proposed and its performance is analyzed showing that this sequence is asymptotically optimal in terms of attainable error exponents. Finally, in Section VI, an explicit example is studied. Section VII presents some conclusions and points out to possible topics for future research.

II. STATEMENT OF THE PROBLEM AND MAIN RESULT

A. Stationary Ergodic Markov Channels

Throughout the paper $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{S}$ will, respectively, denote channel input, output, and state spaces. All are assumed to be finite.

Definition 1: A stationary *Markov channel* is described by:
- a stochastic kernel consisting of a family

$$\{P(\cdot, \cdot | s, x) \in \mathcal{P}(\mathcal{S} \times \mathcal{Y}) | s \in \mathcal{S}, \ x \in \mathcal{X}\}$$

of joint probability measures over $\mathcal{S} \times \mathcal{Y}$, indexed by elements of $\mathcal{S}$ and $\mathcal{X}$;
- an initial state distribution $\boldsymbol{\mu}$ in $\mathcal{P}(\mathcal{S})$.

As it will become clear, the quantity $P(s_+, y | s, x)$ corresponds to the conditioned joint probability that the next state is s_+ and the current output is y, given that the current state is s and the current input is x.

For a channel as in Definition 1, let

$$P_S(s_+ | s, x) := \sum_{y \in \mathcal{Y}} P(s_+, y | s, x)$$

$$P_Y(y | s, x) := \sum_{s_+ \in \mathcal{S}} P(s_+, y | s, x)$$

be the $\mathcal{S}$-marginals and the $\mathcal{Y}$-marginals, respectively. A Markov channel is said to have no ISI if, conditioned on the current state, the next state is independent from the current input and output, i.e., if the stochastic kernel factorizes as

$$P(s_+, y | s, x) = P_S(s_+ | s) P_Y(y | s, x). \tag{2}$$

We shall consider the associated stochastic kernels $\{Q_S(\cdot | s, \boldsymbol{u})\}$ and $\{Q(\cdot, \cdot | s, \boldsymbol{u})\}$ where, for the states $s_+, s \in \mathcal{S}$, the output $y \in \mathcal{Y}$, and the input distribution $\boldsymbol{u} \in \mathcal{P}(\mathcal{X})$

$$Q(s_+, y | s, \boldsymbol{u}) := \sum_{x \in \mathcal{X}} P(s_+, y | s, x) \boldsymbol{u}(x)$$

$$Q_S(s_+ | s, \boldsymbol{u}) := \sum_{x \in \mathcal{X}} P_S(s_+ | s, x) \boldsymbol{u}(x). \tag{3}$$

We shall use the notation

$$Q_{\boldsymbol{\pi}} := \big(Q_S(s_+ | s, \pi(s)) \big)_{s, s_+ \in \mathcal{S}} \tag{4}$$

for the state-transition stochastic matrix induced by a map $\pi : \mathcal{S} \to \mathcal{P}(\mathcal{X})$ (the latter will be referred to as a deterministic sta-

tionary policy). With a common abuse of notation, for any map $f : \mathcal{S} \to \mathcal{X}$ we shall write Q_f in place of $Q_{\delta_{f(\cdot)}}$.[1]

Throughout the paper we shall restrict ourselves to FSMCs satisfying the following ergodicity assumption.

Assumption 1: For every $f : \mathcal{S} \to \mathcal{X}$ the stochastic matrix Q_f is irreducible, i.e., for every $s_1, s_2 \in \mathcal{S}$ there exists some $t \geq 1$ such that $(Q_f^t)_{s_1, s_2} > 0$, where Q_f^t denotes the product of Q_f with itself t times.

Assumption 1 can be relaxed or replaced by other equivalent assumptions. Here, we limit ourselves to observe that it involves the S-marginals $\{P_S\}$ of the Markov channel only. Moreover, it is easily testable, since it requires a finite number of finite directed graphs to be strongly connected. Since taking a convex combination does not reduce the support, Assumption 1 guarantees that for every deterministic stationary policy $\pi : \mathcal{S} \to \mathcal{P}(\mathcal{X})$ the stochastic matrix Q_π is irreducible. Then, the Perron–Frobenius theorem [10, p. 58], guarantees that Q_π has a unique stationary distribution in $\mathcal{P}(\mathcal{S})$ which will be denoted by $\boldsymbol{\mu}_\pi$.

B. Capacity of Ergodic FSMCs

To any ergodic FSMC we associate the mutual information cost function $c : \mathcal{S} \times \mathcal{P}(\mathcal{X}) \to \mathbb{R}$

$$c(s, \boldsymbol{u}) := I(X; Y, S_+ | S = s)$$
$$= \sum_{x, y, s_+} \boldsymbol{u}(x) P(s_+, y | s, x) \log \frac{P(s_+, y | s, x)}{\sum_z \boldsymbol{u}(z) P(s_+, y | s, z)} \tag{5}$$

and define its capacity as

$$C := \max_\pi I(X; Y, S_+ | S)$$
$$= \max_{\pi: \mathcal{S} \to \mathcal{P}(\mathcal{X})} \sum_{s \in \mathcal{S}} \boldsymbol{\mu}_\pi(s) c(s, \pi(s)). \tag{6}$$

In the definitions (5) and (6), the terms $I(X; S_+, Y | S = s)$ and $I(X; S_+, Y | S)$, respectively, denote the mutual information between X and the pair (S_+, Y) when $S = s$, and the conditional mutual information (see [8]) between X and the pair (S_+, Y) given S, where S is an $\mathcal{S}$-valued random variable (r.v.) whose marginal distribution is given by the stationary measure $\boldsymbol{\mu}_\pi$, X is an $\mathcal{X}$-valued r.v. whose conditional distribution given S is described by the policy π, while S_+ and Y are, respectively, an $\mathcal{S}$-valued r.v. and a $\mathcal{Y}$-valued r.v. whose joint conditional distribution given X and S is described by the stochastic kernel $P(S_+, Y | S, X)$. Observe that the mutual information cost function c is continuous over $\mathcal{S} \times \mathcal{P}(\mathcal{X})^2$ and takes values in the bounded interval $[0, \log |\mathcal{X}|]$.

The quantity C defined above is known to equal the capacity of the ergodic Markov channel we are considering when perfect causal CSI is available at both transmission ends, with or

---

[1]Here and throughout the paper, for a measure space $(A, \mathcal{B})$ we shall denote Dirac's delta probability measure centered in a point $a \in A$ by δ_a, i.e., for $B \in \mathcal{B}$, $\delta_a(B) = 1$ if $a \in B$, and $\delta_a(B) = 0$ if $a \notin B$.

[2]Throughout the paper, finite sets will be considered equipped with the complete topology, finite-dimensional spaces equipped with the Euclidean topology, and product spaces with the product topology. Hence, for instance, the continuity of the function $c : \mathcal{S} \times \mathcal{P}(\mathcal{X}) \to \mathbb{R}$ is equivalent to the continuity of the functions $\boldsymbol{u} \mapsto c(s, \boldsymbol{u})$ over the simplex $\mathcal{P}(\mathcal{X})$, for all $s \in \mathcal{S}$.

without output feedback [32]. It is important to observe that, due to the presence of ISI the policy π plays a dual role in the optimization problem (5) since it affects both the mutual information cost $c(s, \pi(s)) = I(X; S_+, Y | S = s)$ as well as the ergodic channel state distribution $\boldsymbol{\mu}_\pi$ with respect to which the former is averaged.

In the case when there is no ISI, i.e., when (2) is satisfied, this phenomenon disappears. In fact, since the invariant measure $\boldsymbol{\mu}$ is independent from the policy π, we have that (6) reduces to

$$
\begin{aligned}
C &= \sum_s \boldsymbol{\mu}(s) \max_{p_X} c(s, p_X) \\
&= \sum_s \boldsymbol{\mu}(s) \max_{p_X} I(X; Y | S = s)
\end{aligned} \tag{7}
$$

where in the rightmost side of (7), $\max_{p_X} I(X; Y | S = s)$, coincides with the capacity of the DMC with input $\mathcal{X}$, output $\mathcal{Y}$, and transition probabilities $P_Y(\cdot | s, \cdot)$. The simplest case of FSMCs with no ISI is obtained when the state sequence is an independent and identically distributed (i.i.d.) exogenous process, i.e., when

$$
P(s_+, y | s, x) = \boldsymbol{\mu}(s_+) P_Y(y | s, x).
$$

In this case, (6) reduces to the capacity of a DMC with input space $\mathcal{X}' := \mathcal{X}^{\mathcal{S}}$—the set of all maps from $\mathcal{S}$ to $\mathcal{X}-$, output space $\mathcal{Y}' := \mathcal{S} \times \mathcal{Y}$—the Cartesian product of $\mathcal{S}$ and $\mathcal{Y}-$, and transition probabilities given by

$$
P'(y' | x') := \boldsymbol{\mu}(s) P_Y(y | s, x'(s)) \tag{8}
$$

where $y' = (s_+, y)$. Observe the difference with respect to the case when the state is causally observed at the transmitter only, whose capacity was first found in [29]. While the input spaces of the equivalent DMCs do coincide, the output space is larger, as we assume that the state is causally observable also at the receiver end.

Finally, notice that, when the state space is trivial (i.e., when $|\mathcal{S}| = 1$), (6) reduces to the well-known formula for the capacity of a DMC.

C. Burnashev Coefficient of FSMCs

Consider now the cost function $d : \mathcal{S} \times \mathcal{P}(\mathcal{X}) \rightarrow [0, +\infty]$

$$
\begin{aligned}
d(s, \boldsymbol{u}) &= D\left(Q(\cdot, \cdot | s, \boldsymbol{u}) \| Q(\cdot, \cdot | s, \boldsymbol{u}')\right) \\
&:= \sup_{\boldsymbol{u}'} \sum_{y, s_+} Q(s_+, y | s, \boldsymbol{u}) \log \frac{Q(s_+, y | s, \boldsymbol{u})}{Q(s_+, y | s, \boldsymbol{u}')}
\end{aligned} \tag{9}
$$

where $D(\boldsymbol{\nu}_1 \| \boldsymbol{\nu}_2)$ denotes the Kullback-Leibler information divergence between two probability measures $\boldsymbol{\nu}_1$ and $\boldsymbol{\nu}_2$. For each $s \in \mathcal{S}$, it is useful to consider the set

$$
\mathcal{Z}_s := \{(s_+, y) \in \mathcal{S} \times \mathcal{Y} | \exists x \in \mathcal{X} : P(s_+, y | s, x) > 0\} \tag{10}
$$

of all channel state and output pairs which can be achieved from the state s, and the quantity

$$
\lambda_s := \min\left\{ P(s_+, y | s, x) \big| x \in \mathcal{X}, (s_+, y) \in \mathcal{Z}_s \right\}. \tag{11}
$$

Observe that $\lambda_s = 0$ iff there exists a pair $(s_+, y) \in \mathcal{S} \times \mathcal{Y}$ which is reachable from s by some but not all possible inputs $x \in \mathcal{X}$. It follows that the cost function d is bounded and continuous over $\mathcal{S} \times \mathcal{P}(\mathcal{X})$ if and only if

$$
\lambda := \min\{\lambda_s | s \in \mathcal{S}\} \tag{12}
$$

is strictly positive, i.e.,

$$
\lambda > 0 \quad \Longleftrightarrow \quad d_{\max} := \sup_{s, \boldsymbol{u}} d(s, \boldsymbol{u}) < +\infty. \tag{13}
$$

Define the Burnashev coefficient of a Markov channel as

$$
D := \sup_{\pi : \mathcal{S} \longrightarrow \mathcal{P}(\mathcal{X})} \sum_{s \in \mathcal{S}} \boldsymbol{\mu}_\pi(s) d(s, \pi(s)). \tag{14}
$$

Notice that D is finite iff (13) holds.

Moreover, a standard convexity argument allows one to argue that both the suprema in (9) and in (14) are achieved in some corner points of the simplex $\mathcal{P}(\mathcal{X})$. More precisely, for (9), this follows immediately from the convexity of the Kullback–Leibler divergence $D(\boldsymbol{\nu}_1 \| \boldsymbol{\nu}_2)$ in $\boldsymbol{\nu}_2$ [8, Theorem 2.7.2], and the linearity of $Q(\cdot, \cdot \| s, \boldsymbol{u}')$ in $\boldsymbol{u}'$. For (14), one can invoke [5, Lemma 5.3] guaranteeing the convexity of $\sum_{s \in \mathcal{S}} \boldsymbol{\mu}_\pi(s) d(s, \pi(s))$ in $Q(\cdot, \cdot \| s, \boldsymbol{u})$, and again observe that $Q(\cdot, \cdot \| s, \boldsymbol{u})$ is linear in $\boldsymbol{u}$. Hence, we have

$$
\begin{aligned}
D &= \max_{f_a, f_b} \sum_{s, s_+, y} \boldsymbol{\mu}_{f_a}(s) P(s_+, y | s, f_a(s)) \log \frac{P(s_+, y | s, f_a(s))}{P(s_+, y | s, f_b(s))} \\
&= \max_{f_a, f_b} \sum_s \boldsymbol{\mu}_{f_a}(s) D\left(P(\cdot, \cdot | s, f_a(s)) \| P(\cdot, \cdot | s, f_b(s))\right)
\end{aligned} \tag{15}
$$

where the maximum is taken over all functions $f_a, f_b \in \mathcal{X}^{\mathcal{S}}$.

Similarly to what has been already noted for the role of the policy π in the optimization problem (6), it can be observed that, due to the presence of ISI, the map f_a has a dual effect in the maximization in (15) since it affects both the Kullback–Leibler information divergence cost

$$
D\left(P(\cdot, \cdot | s, f_a(s)) \| P(\cdot, \cdot | s, f_b(s))\right)
$$

and the ergodic state measure $\boldsymbol{\mu}_{f_a}$. Notice the asymmetry with the role of the map f_b whose associated invariant state measure instead does not come into the picture at all in the definition of the coefficient D. Once again, in the absence of ISI, (15) simplifies to

$$
D = \sum_{s \in \mathcal{S}} \boldsymbol{\mu}(s) \max_{x_a, x_b \in \mathcal{X}} D\left(P(\cdot, \cdot | s, x_a) \| P(\cdot, \cdot | s, x_b)\right).
$$

We observe that in the memoryless case (i.e., when $|\mathcal{S}| = 1$) the coefficient D coincides with the Kullback–Leibler information divergence between the output measures associated to the pair of most distinguishable inputs, the quantity originally denoted with the symbol C_1 in [6]. When the state space is nontrivial ($|\mathcal{S}| > 1$), and the channel state process forms an i.i.d. sequence independent from the channel input, then the Burnashev coefficient D reduces to that of the equivalent DMC with

enlarged input space $\mathcal{X}' = \mathcal{X}^{\mathcal{S}}$, output space $\mathcal{Y}' = \mathcal{S} \times \mathcal{Y}$, and transition probabilities defined in (8).

D. Causal Feedback Encoders, Sequential Decoders, and Main Result

We now introduce the class of coding schemes we wish to consider. A schematic representation of the information patterns is reported in Fig. 1.

Definition 2: A *causal feedback encoder* is the pair of a finite message set and a sequence of maps

$$\Phi = \left(\mathcal{W}, \left(\phi_t : \mathcal{W} \times \mathcal{S}^t \times \mathcal{Y}^{t-1} \to \mathcal{X} \right)_{t \in \mathbb{N}} \right). \quad (16)$$

With Definition 2, we are implicitly assuming that perfect channel state knowledge as well as perfect output feedback are available at the encoder side.

Given a stationary Markov channel and a causal feedback encoder as in Definition 2, we shall consider a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. The corresponding expectation operator will be denoted by $\mathbb{E}$, while, for an event $A \in \mathcal{A}$, $\overline{A} = \Omega \setminus A$ will denote the complementary event, and $\mathbb{1}_A : \Omega \to \{0, 1\}$ will denote its indicator function, defined by $\mathbb{1}_A(a) = 1$ if $a \in A$, $\mathbb{1}_A(a) = 0$ if $a \notin A$. We assume that the following r.v.'s are defined over $(\Omega, \mathcal{A}, \mathbb{P})$:

- $\mathcal{W}$-valued r.v. W describing the message to be transmitted;
- a sequence $\boldsymbol{X} = (X_t)$ of $\mathcal{X}$-valued r.v.'s (the channel input sequence);
- a sequence $\boldsymbol{Y} = (Y_t)$ of $\mathcal{Y}$-valued r.v.'s (the channel output sequence);
- a sequence $\boldsymbol{S} = (S_t)$ of $\mathcal{S}$-valued r.v.'s (the channel state sequence).

We shall consider the time ordering[3]

$$W, S_1, X_1, Y_1, S_2, X_2, Y_2, \ldots,$$

and assume that $\mathbb{P}$-a.s.

$$\mathbb{P}(W = w) = \frac{1}{|\mathcal{W}|}, \quad \mathbb{P}(S_1 = s | W) = \boldsymbol{\mu}(s),$$
$$\mathbb{P}(X_t = x | W, \boldsymbol{S}_1^t, \boldsymbol{X}_1^{t-1}, \boldsymbol{Y}_1^{t-1}) = \delta_{\phi_t(W, \boldsymbol{S}_1^t, \boldsymbol{Y}_1^{t-1})}(x),$$
$$\mathbb{P}(S_{t+1} = s, Y_t = y | W, \boldsymbol{S}_1^t, \boldsymbol{Y}_1^{t-1}, \boldsymbol{X}_1^t) = P\left(s, y | S_t, X_t\right).$$
$$(17)$$

It is convenient to introduce the following notation for the observation available at the encoder and decoder side. For every t we define the σ-fields $\mathcal{E}_t := \sigma\left(\boldsymbol{S}_1^t, \boldsymbol{Y}_1^{t-1}\right)$, describing the feedback observation available at the encoder side, and $\mathcal{F}_t := \sigma\left(\boldsymbol{S}_1^t, \boldsymbol{Y}_1^t\right)$, describing the observation available at the decoder. Notice that the full observation available at the encoder at time t is $\sigma(W, \boldsymbol{S}_1^t, \boldsymbol{Y}_1^{t-1})$. Clearly

$$\{\emptyset, \Omega\} = \mathcal{E}_0 = \mathcal{F}_0 \subseteq \mathcal{E}_1 \subseteq \mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{A}. \quad (18)$$

In particular, we end up with two nested filtrations: $\mathcal{F} := (\mathcal{F}_t)_{t \geq 0}$ and $\mathcal{E} := (\mathcal{E}_t)_{t \geq 0}$.

Fig. 1. Information patterns for variable-length block-coding schemes on an FSMC with causal feedback and CSI.

Observe that, while the space $(\Omega, \mathcal{A})$ and the filtrations $(\mathcal{F}_t)$ and $(\mathcal{E}_t)$ depend on the message set $\mathcal{W}$ and on the channel state, input and output sets $\mathcal{S}$, $\mathcal{X}$ and $\mathcal{Y}$ only,[4] the probability measure $\mathbb{P}$ does depend on the stochastic kernel P describing the channel, as well as on the encoder Φ. Many of the statements in this paper will be meant to hold $\mathbb{P}$-almost surely, thought this may not always be explicitly stated.

Definition 3: A *transmission time* T is a stopping time for the receiver filtration $\mathcal{F}$, i.e., it is $\{1, 2, \ldots, \infty\}$-valued r.v. such that the event $\{T \leq t\}$ is $\mathcal{F}_t$-measurable for each time t.

Definition 4: A *sequential decoder* for a causal feedback encoder Φ as in (16) is a sequence of maps

$$\Psi = \left(\psi_t : \mathcal{S}^t \times \mathcal{Y}^t \to \mathcal{M} \right)_{t \in \mathbb{N}}. \quad (19)$$

For a transmission time T and a sequential decoder Ψ, the estimated message is

$$\hat{W} := \psi_T \left(\boldsymbol{S}_1^T, \boldsymbol{Y}_1^T \right). \quad (20)$$

Notice that with Definitions 3 and 4 we are assuming that perfect causal state knowledge is available at the receiver. In particular, the fact that the transmitter's feedback and the receiver's observation patterns are nested allows one to use a variable-length scheme.

The triple (Φ, T, Ψ) consisting of a causal feedback encoder Φ, a transmission time T, and a sequential decoder Ψ, is called a variable-length block-coding scheme. Its error probability is given by

$$p_e(\Phi, T, \Psi) := \mathbb{P}\left(\hat{W} \neq W \right).$$

Following Burnashev's approach we shall consider the expected decoding time $\mathbb{E}[T]$ as a measure of the delay of the scheme (Φ, T, Ψ) and accordingly define its rate as

$$R(\Phi, T, \Psi) := \frac{\log |\mathcal{W}|}{\mathbb{E}[T]}.$$

We are now ready to state our main result. It is formulated in an asymptotic setting, considering infinite sequences of variable-length block-coding schemes with asymptotic average rate below capacity and vanishing error probability.

Theorem 1: For any R in $(0, C)$
1) any infinite sequence $(\Phi^{(n)}, T^{(n)}, \Psi^{(n)})$ of variable-length block-coding schemes[5] such that

$$\lim_n p_e\left(\Phi^{(n)}, T^{(n)}, \Psi^{(n)} \right) = 0$$
$$\liminf_n R\left(\Phi^{(n)}, T^{(n)}, \Psi^{(n)} \right) \geq R \quad (21)$$

---

[3]Different time orderings would lead to similar results: for instance, the time ordering $W, S_0, X_1, S_1, Y_1, X_2, S_2, Y_2, \ldots$ can be handled by considering the stochastic kernel $\{P(s, y | s_-, x)\}$ describing the joint probability distribution of the current state and output given the previous state and the current input.
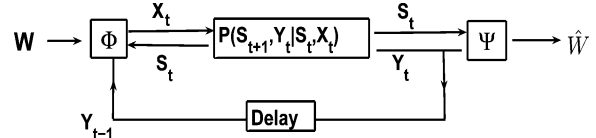
[4]Indeed, with no loss of generality Ω can be identified with $\mathcal{W} \times \mathcal{S}^{\mathbb{N}} \times \mathcal{X}^{\mathbb{N}} \times \mathcal{Y}^{\mathbb{N}}$, and the r.v.'s $W, \boldsymbol{S}, \boldsymbol{X}, \boldsymbol{Y}$ can be identified with the standard projections to $\mathcal{W}, \mathcal{S}^{\mathbb{N}}, \mathcal{X}^{\mathbb{N}}$, and $\mathcal{Y}^{\mathbb{N}}$ respectively.

[5]Here the sequence index n should not be confused with the time index t, nor with the average block length $\mathbb{E}[T^{(n)}]$.

satisfies

$$\limsup_n -\frac{1}{\mathbb{E}[T^{(n)}]} \log p_e\left(\Phi^{(n)}, T^{(n)}, \Psi^{(n)}\right) \le E_B(R);$$

(22)

2) there exists an infinite sequence $\left(\Phi^{(n)}, T^{(n)}, \Psi^{(n)}\right)$ of variable-length block-coding schemes satisfying (21) and such that:

- if $D < +\infty$

$$\lim_n -\frac{1}{\mathbb{E}[T^{(n)}]} \log p_e\left(\Phi^{(n)}, T^{(n)}, \Psi^{(n)}\right) = E_B(R)$$ (23)

- if $D = +\infty$

$$p_e\left(\Phi^{(n)}, T^{(n)}, \Psi^{(n)}\right) = 0, \quad \forall n \in \mathbb{N}.$$ (24)

Observe that Burnashev's original result [6] for DMCs can be recovered as a particular case of Theorem 2 when the state space is trivial, i.e., $|\mathcal{S}| = 1$.

Notice that, when $D = +\infty$, the first point of Theorem 2 becomes trivial, while the second point tells us that feedback coding schemes with zero-error probability exist. As it will become clear in Section V, the reason is that $D = +\infty$ iff there exist two states s and s_+, two inputs x_1 and x_2 and an output y such that $P(s_+, y|s, x_1) > 0$ and $P(s_+, y|s, x_2) = 0$: this makes it possible to build a sequence of binary coding schemes whose error probability conditioned on the transmission of one of the two codewords is identically zero, while the error probability conditioned on the transmission of the other codeword is asymptotically vanishing.

III. An Upper Bound on the Achievable Error Exponent

The aim of this section is to provide an upper bound on the error exponent of an arbitrary variable-length block-coding scheme. A first observation is that, without any loss of generality, we can restrict ourselves to the case when D is finite, since otherwise the claim (22) is trivially true. The main result of this section is contained in Theorem 2 whose proof will pass through a series of intermediate steps, contained in Lemmas 1–5. The results of this section generalize those in [2], [6], [23], [33], and [34] to Markov channels, and the proofs we present are close in spirit to the arguments developed in these references.

The main idea, borrowed from [6], is to obtain two different upper bounds for the error probability. Differently from [6], [23], [33], and [34], we will follow an approach similar to the one proposed in [2] and look at the behavior of the maximum *a posteriori* (MAP) error probability, rather than that of the *a posteriori* entropy. The aforementioned bounds correspond to two distinct phases which can be recognized in any sequential transmission scheme and will be the content of Sections III-A and III-B. The first one is provided in Lemma 2, whose proof is based on an application of Fano's inequality combined with a martingale argument invoking Doob's optional stopping theorem. The second bound is given by Lemma 4, whose proof combines the use of the log-sum inequality with another application of Doob's optional stopping theorem. In Section III-C, these two bounds will be combined obtaining Theorem 2.

A. A First Bound on the Error Probability

Suppose we are given a causal feedback encoder $\Phi = (\mathcal{W}, \{\phi_t\})$ as in (16) and a transmission time T as in Definition 3. The goal is to find a lower bound for the error probability $p_e(\Phi, T, \Psi)$, where Ψ is an arbitrary sequential decoder for Φ and T. Our arguments here closely parallel those developed in [2, Sec. IV] in the memoryless case.

It will be convenient to define for every time $t \ge 0$ the σ-field $\mathcal{G}_t := \mathcal{E}_{t+1}$ describing the encoder's feedback observation at time $t + 1$. $\mathcal{G} := (\mathcal{G}_t)$ will denote the corresponding filtration. Let

$$\boldsymbol{\theta}_t \in \mathcal{P}(\mathcal{W}), \qquad \boldsymbol{\theta}_t(w) := \mathbb{P}(W = w|\mathcal{G}_t)$$
$$\Pi_t := 1 - \max\{\boldsymbol{\theta}_t(w)|w \in \mathcal{W}\}$$

be, respectively, the conditioned probability distribution of the message W and the MAP error probability given the feedback observation $\mathcal{G}_t$ at time $t+1$. Clearly, both $\boldsymbol{\theta}_t$ and Π_t are $\mathcal{G}_t$-measurable r.v.'s.

For each time t, let us consider the classes of decoders $\mathcal{D}_t := \{\psi_t : \mathcal{S}^t \times \mathcal{Y}^t \to \mathcal{W}\}$, $\check{\mathcal{D}}_t := \{\psi_t : \mathcal{S}^{t+1} \times \mathcal{Y}^t \to \mathcal{W}\}$, differing because of the possible dependence on the state S_{t+1}. It is a well-known fact that the decoder minimizing the error probability over $\check{\mathcal{D}}_t$ is the maximum MAP one, defined by[6]

$$\check{\psi}_t(\boldsymbol{S}_1^{t+1}, \boldsymbol{Y}_1^t) := \arg\max_{w \in \mathcal{W}} \{\boldsymbol{\theta}_t(w)\}.$$ (25)

Since $\mathcal{D}_t \subseteq \check{\mathcal{D}}_t$, it follows that, for any decoder $\psi_t \in \mathcal{D}_t$, we have

$$p_e(\Phi, t, \psi_t) \ge p_e(\Phi, t, \check{\psi}_t) = \mathbb{E}[\Pi_t].$$

The preceding discussion naturally generalizes from the fixed-length setting to the sequential one. Given a transmission time T, observe that, since $\mathcal{F}_t \subseteq \mathcal{G}_t$ for every $t \ge 0$, T is also stopping time for the filtration $\mathcal{G}$ and $\mathcal{F}_T \subseteq \mathcal{G}_T$. It follows that the error probability of any variable-length block-coding scheme (Φ, T, Ψ) is lower-bounded by that of $(\Phi, T, \check{\Psi})$, where $\check{\Psi} := (\check{\psi}_t)$ is the sequential MAP decoder defined in (25). Therefore, we can conclude that

$$p_e(\Phi, T, \Psi) \ge \mathbb{E}[\Pi_T]$$ (26)

for any variable-length block-coding scheme (Φ, T, Ψ).

In the sequel, we shall obtain lower bounds for the right-hand side of (26). In particular, since W is uniformly distributed over the message set $\mathcal{W}$ and is independent from the initial state S_1, we have that $\boldsymbol{\theta}_0(w) = \mathbb{P}(W = w) = 1/|\mathcal{W}|$ for each message $w \in \mathcal{W}$, so that $\Pi_0 = (|\mathcal{W}| - 1)/|\mathcal{W}|$. Moreover, we have the following recursive lower bound for Π_t.

Lemma 1: Given any causal feedback encoder Φ, $\mathbb{P}$-a.s.

$$\Pi_t \ge \lambda \Pi_{t-1}, \quad t \ge 1.$$

Proof: See the Appendix. □

---

[6] We shall use the convention for the operator $\arg\max$ to arbitrarily assign one of the optimizing values in case of nonuniqueness.

For every ε in $(0, 1/2)$, we consider the r.v.

$$\tau_\varepsilon := \min\left\{T, \inf\left\{t \in \mathbb{N} : \Pi_t \le \varepsilon\right\}\right\} \tag{27}$$

describing the first time before T when the MAP error probability goes below ε. It is immediate to verify that τ_ε is a stopping time for the filtration $\mathcal{G}$. Moreover, the event $\{\Pi_{\tau_\varepsilon} > \varepsilon\}$ implies the event $\{\tau_\varepsilon = T\}$, so that an application of the Markov inequality and (26) give us

$$\begin{aligned}
\mathbb{P}\left(\Pi_{\tau_\varepsilon} > \varepsilon\right) &= \mathbb{P}\left(\{\Pi_{\tau_\varepsilon} > \varepsilon\} \cap \{\tau_\varepsilon = T\}\right) \\
&\le \mathbb{P}\left(\Pi_T > \varepsilon\right) \\
&\le \frac{1}{\varepsilon}\mathbb{E}\left[\Pi_T\right] \\
&\le \frac{1}{\varepsilon}p_e(\Phi, T, \Psi).
\end{aligned}$$

We introduce the following notation for the *a posteriori* entropy:

$$\Gamma_t := H(\boldsymbol{\theta}_t) = -\sum_{w \in \mathcal{W}} \boldsymbol{\theta}_t(w) \log \boldsymbol{\theta}_t(w), \quad t \ge 0.$$

Observe that, since the initial state S_1 is independent from the message W, then

$$\Gamma_0 = \log|\mathcal{W}|.$$

From Fano's inequality [8, Theorem 2.11.1, p. 39] it follows that, the event $A := \{\Pi_{\tau_\varepsilon} \le \varepsilon\}$ implies that

$$\Gamma_{\tau_\varepsilon} \le H(\varepsilon) + \varepsilon\log|\mathcal{W}|.$$

Hence, since $\Gamma_{\tau_\varepsilon} \le \log|\mathcal{W}|$, the expected value of $\Gamma_{\tau_\varepsilon}$ can be bounded from above as follows:

$$\begin{aligned}
\mathbb{E}\left[\Gamma_{\tau_\varepsilon}\right] &= \mathbb{E}\left[\Gamma_{\tau_\varepsilon}|A\right]\mathbb{P}(A) + \mathbb{E}\left[\Gamma_{\tau_\varepsilon}|\overline{A}\right]\mathbb{P}\left(\overline{A}\right) \\
&\le \mathbb{P}(A)\left(H(\varepsilon) + \varepsilon\log|\mathcal{W}|\right) + \mathbb{P}\left(\overline{A}\right)\log|\mathcal{W}| \\
&\le H(\varepsilon) + \left(\varepsilon + \frac{1}{\varepsilon}p_e(\Phi, T, \Psi)\right)\log|\mathcal{W}|. \tag{28}
\end{aligned}$$

We now introduce, for every time t, a $\mathcal{P}(\mathcal{X})$-valued r.v. $\boldsymbol{\Upsilon}_t$ describing the channel input distribution induced by the causal encoder Φ at time t

$$\begin{aligned}
\boldsymbol{\Upsilon}_t(x) &:= \mathbb{P}\left(X_t = x|\mathcal{E}_t\right) \\
&= \mathbb{P}\left(\phi_t(W, \boldsymbol{S}_1^t, \boldsymbol{Y}_1^{t-1}) = x|\mathcal{E}_t\right). \tag{29}
\end{aligned}$$

Notice that $\boldsymbol{\Upsilon}_t$ is an $\mathcal{E}_t$-measurable r.v., i.e., equivalently, it is a function of the pair $(\boldsymbol{S}_1^t, \boldsymbol{Y}_1^{t-1})$.

The following result relates three relevant quantities characterizing the performances of any variable-length block-coding scheme: the cardinality of the message set $\mathcal{W}$, the error probability of the coding scheme, and the mutual information cost $c(5)$ incurred up to the stopping time τ_ε

$$C_\varepsilon(\Phi, T) := \mathbb{E}\left[\sum_{1 \le t \le \tau_\varepsilon} c(S_t, \boldsymbol{\Upsilon}_t)\right]. \tag{30}$$

Lemma 2: For any variable-length block-coding scheme (Φ, T, Ψ) and any $0 < \varepsilon < 1/2$, we have

$$C_\varepsilon(\Phi, T) \ge \left(1 - \varepsilon - \frac{p_e(\Phi, T, \Psi)}{\varepsilon}\right)\log|\mathcal{W}| - H(\varepsilon). \tag{31}$$

Proof: See the Appendix . $\qquad\square$

B. A Lower Bound to the Error Probability of a Composite Binary Hypothesis Test

We now consider a particular binary hypothesis testing problem which will arise while proving the main result, and provide a lower bound on its error probability. The steps here are similar to those in [2, Sec.III] and [34, Sec.III].

Suppose we are given a causal feedback encoder $\Phi = (\mathcal{W}, (\phi_t))$. Consider a nontrivial binary partition of the message set

$$\mathcal{W} = \mathcal{W}_0 \cup \mathcal{W}_1, \quad \mathcal{W}_0 \cap \mathcal{W}_1 = \emptyset, \quad \mathcal{W}_0, \mathcal{W}_1 \ne \emptyset, \tag{32}$$

a stooping time T for the filtration $\mathcal{G}$, and a sequential binary hypothesis test $\tilde{\Psi} = (\tilde{\psi} : \mathcal{S}^{t+1} \times \mathcal{Y}^t \to \{0, 1\})$ between the two hypothesis $\{W \in \mathcal{W}_0\}$ and $\{W \in \mathcal{W}_1\}$. Following the common statistical terminology, we shall call $\tilde{\Psi}$ a composite test since it must decide between two classes of probability laws for the process $(\boldsymbol{S}, \boldsymbol{Y})$ rather than between two single laws. Define

$$\tilde{W} := \tilde{\psi}_T(\boldsymbol{S}_1^{T+1}, \boldsymbol{Y}_1^T).$$

For every t, we define the $\mathcal{P}(\mathcal{X})$-valued random variables $\boldsymbol{\Upsilon}_t^0$ and $\boldsymbol{\Upsilon}_t^1$ by

$$\boldsymbol{\Upsilon}_t^i(x) = \mathbb{P}\left(X_t = x|W \in \mathcal{W}_i, \mathcal{E}_t\right), \quad x \in \mathcal{X}, i = 0, 1.$$

The r.v. $\boldsymbol{\Upsilon}_t^0$ (resp., $\boldsymbol{\Upsilon}_t^1$) represents the channel input distribution at time t induced by the encoder Φ when restricted to the message subset $\mathcal{W}_0$ (resp., $\mathcal{W}_1$). Notice that

$$\boldsymbol{\Upsilon}_t = \boldsymbol{\theta}_{t-1}(\mathcal{W}_0)\boldsymbol{\Upsilon}_t^0 + \boldsymbol{\theta}_{t-1}(\mathcal{W}_1)\boldsymbol{\Upsilon}_t^1.$$

For $r \le t$ and $i = \{0, 1\}$, define the $\mathcal{G}_r$-conditioned probability distribution $\boldsymbol{\nu}_{r,t}^i \in \mathcal{P}\left(\mathcal{S}^{t+1} \times \mathcal{Y}^t\right)$ of the channel state and output pair $(\boldsymbol{S}_1^{t+1}, \boldsymbol{Y}_1^t)$ given $\{W \in \mathcal{W}_i\}$

$$\boldsymbol{\nu}_{r,t}^i(\boldsymbol{S}, \boldsymbol{Y}) = \mathbb{P}\left(\boldsymbol{S}_1^{t+1} = \boldsymbol{S}, \boldsymbol{Y}_1^t = \boldsymbol{Y}|W \in \mathcal{W}_i, \mathcal{G}_r\right).$$

Observe that both the random measures $\boldsymbol{\nu}_{r,t}^0$ and $\boldsymbol{\nu}_{r,t}^1$ put mass only on those sequences $(\tilde{\boldsymbol{S}}, \tilde{\boldsymbol{Y}})$ such that $\tilde{\boldsymbol{S}}_1^{r+1} = \boldsymbol{S}_1^{r+1}$ and $\tilde{\boldsymbol{Y}}_1^r = \boldsymbol{Y}_1^r$.

Let now τ be another stopping time for the filtration $\mathcal{G}$, such that $\tau \le T$. Then, $\boldsymbol{\nu}_{\tau,T}^0$ and $\boldsymbol{\nu}_{\tau,T}^1$ are well defined as $\mathcal{G}_\tau$-measurable random measures on the σ-field $\mathcal{G}_T$. Therefore, we can consider their Kullback–Leibler information divergences

$$\begin{aligned}
L_0 &:= D\left(\boldsymbol{\nu}_{\tau,T}^0 \| \boldsymbol{\nu}_{\tau,T}^1\right) \\
&= \mathbb{E}\left[\log\frac{\boldsymbol{\nu}_{\tau,T}^0\left(\boldsymbol{S}_1^{T+1}, \boldsymbol{Y}_1^T\right)}{\boldsymbol{\nu}_{\tau,T}^1\left(\boldsymbol{S}_1^{T+1}, \boldsymbol{Y}_1^T\right)}\bigg| W \in \mathcal{W}_0, \mathcal{G}_\tau\right]
\end{aligned}$$

$$L_1 := D\left(\boldsymbol{\nu}_{\tau,T}^1 \| \boldsymbol{\nu}_{\tau,T}^0\right)$$

$$= \mathbb{E}\left[\log \frac{\boldsymbol{\nu}_{\tau,T}^1(\boldsymbol{S}_1^{T+1}, \boldsymbol{Y}_1^T)}{\boldsymbol{\nu}_{\tau,T}^0(\boldsymbol{S}_1^{T+1}, \boldsymbol{Y}_1^T)}\Big| W \in \mathcal{W}_1, \mathcal{G}_\tau\right]. \quad (33)$$

Observe that both L_0 and L_1 are $\mathcal{G}_\tau$-measurable r.v.'s.

In the special case, when both τ and T are deterministic constants, an application of the log-sum inequality would show that, for $i = 0, 1$, L_i can be bounded from above by the $\mathcal{G}_\tau$-conditional expected value of the sum of the information divergence costs $d\left(S_t, \boldsymbol{\Upsilon}_t^i\right)$ incurred from time $\tau + 1$ to T. It turns out that the same is true in our setting where both τ and T are stopping times for the filtration $\mathcal{G}$, as stated in the following lemma, whose proof requires, besides an application of the log-sum inequality, a martingale argument invoking Doob's optional stopping theorem.

Lemma 3: Let τ and T be stopping times for the filtration $\mathcal{G}$ such that $\tau \leq T$, and consider a nontrivial binary partition of the message set as in (32). Then, for $i = 0, 1$

$$L_i \leq \mathbb{E}\left[\sum_{\tau < t \leq T} d\left(S_t, \boldsymbol{\Upsilon}_t^i\right)\Big| W \in \mathcal{W}_i, \mathcal{G}_\tau\right]. \quad (34)$$

Proof: See the Appendix. ☐

Suppose now that $\mathcal{W}_1$ is a $\mathcal{G}_\tau$-measurable random variable taking values in $2^{\mathcal{W}} \setminus \{\emptyset, \mathcal{W}\}$, the class of nontrivial proper subsets of the message set $\mathcal{W}$. In other words, we are assuming that $\mathcal{W}_1$ is a random subset of the message set $\mathcal{W}$, deterministically specified by the pair $(\boldsymbol{S}_1^{T+1}, \boldsymbol{Y}_1^\tau)$. The following result provides a lower bound on the error probability of the binary test $\tilde{\Psi}$ conditioned on the σ-field $\mathcal{G}_\tau$:

$$p_\tau := \mathbb{P}\left(\tilde{W} \neq \mathbf{1}_{\{W \in \mathcal{W}_1\}}\big| \mathcal{G}_\tau\right).$$

Lemma 4: Let Φ be any causal encoder, and τ and T be stopping times for the filtration $\mathcal{G}$ such that $\tau \leq T$. Then, for every $2^{\mathcal{W}}$-valued $\mathcal{G}_\tau$-measurable r.v. $\mathcal{W}_1$, we have

$$\mathbb{E}\left[\sum_{\tau < t \leq T} d\left(S_t, \boldsymbol{\Upsilon}_t^{\mathbf{1}_{\{W \in \mathcal{W}_1\}}}\right)\Big| \mathcal{G}_\tau\right] \geq \log \frac{Z}{4} - \log p_\tau \quad (35)$$

where $Z := \min\{\boldsymbol{\theta}_\tau(\mathcal{W}_0), \boldsymbol{\theta}_\tau(\mathcal{W}_1)\}$.

Proof: See the Appendix. ☐

C. Burnashev Bound for Markov Channels

Lemma 5: Let Φ be a causal feedback encoder and T a transmission time. Then, for every $0 < \varepsilon < 1/2$ there exists a $\mathcal{G}_{\tau_\varepsilon}$-measurable random subset $\mathcal{W}_1$ of the message set $\mathcal{W}$, whose *a posteriori* error probability satisfies

$$1 - \lambda\varepsilon \geq \boldsymbol{\theta}_{\tau_\varepsilon}(\mathcal{W}_1) \geq \lambda\varepsilon. \quad (36)$$

Proof: See the Appendix. ☐

To a causal encoder Φ and a transmission time T, for every $0 < \varepsilon < 1/2$ we associate the quantity

$$D_\varepsilon(\Phi, T) := \sup_{\substack{\mathcal{W}_1 \mathcal{G}_{\tau_\varepsilon}-meas. \\ 2^{\mathcal{W}}-valued\ r.v. \\ \lambda\varepsilon \leq \mathbb{P}(W \in \mathcal{W}_1 | \mathcal{G}_{\tau_\varepsilon}) \leq 1 - \lambda\varepsilon}} \mathbb{E}\left[\sum_{\tau_\varepsilon < t \leq T} d\left(S_t, \boldsymbol{\Upsilon}_t^{\mathbf{1}_{\{W \in \mathcal{W}_1\}}}\right)\right] \quad (37)$$

equal to the maximum, over all possible choices of a nontrivial partition of the message set $\mathcal{W}$ as a deterministic function of the joint channel state output process $(\boldsymbol{S}_1^{\tau_\varepsilon+1}, \boldsymbol{Y}_1^{\tau_\varepsilon})$ stopped at the intermediate time τ_ε, of the averaged sum of the information divergence costs $d\left(S_t, \boldsymbol{\Upsilon}_t^{\mathbf{1}_{\{W \in \mathcal{W}_1\}}}\right)$ incurred between times $\tau_\varepsilon + 1$ and T. Intuitively $D_\varepsilon(\Phi, T)$ measures the maximum error exponent achievable by the encoder Φ when transmitting a binary message between times τ_ε and T.

Based on Lemmas 2 and 4, we will now prove the main result of this section, consisting in an upper bound on the largest error exponent achievable by variable-length block-coding schemes with perfect causal state knowledge and output feedback.

Theorem 2: Consider a variable-length block-coding scheme (Φ, T, Ψ). Then, for every ε in $(0, 1/2)$

$$-\log p_e(\Phi, T, \Psi)$$

$$\leq \frac{D}{C}C_\varepsilon(\Phi, T) + D_\varepsilon(\Phi, T) - \frac{D}{C}\log|\mathcal{W}|(1-\alpha) - \beta \quad (38)$$

where

$$\alpha := \varepsilon + \frac{p_e(\Phi, T, \Psi)}{\varepsilon}, \quad \beta := \log\frac{\lambda\varepsilon}{4} - \frac{D}{C}H(\varepsilon). \quad (39)$$

Proof: Let $\mathcal{W}_1$ be a $\mathcal{G}_{\tau_\varepsilon}$-measurable subset of the message set $\mathcal{W}$ satisfying (36). We define the binary sequential decoder $\tilde{\Psi}_\varepsilon = (\tilde{\psi}_{\varepsilon,t})$, where

$$\tilde{\psi}_{\varepsilon,t}(\boldsymbol{s}, \boldsymbol{y}) := \mathbf{1}_{\mathcal{W}_1}(\psi(\boldsymbol{s}, \boldsymbol{y})), \quad \boldsymbol{s} \in \mathcal{S}^{t+1}, \ \boldsymbol{y} \in \mathcal{Y}^t.$$

We can lower-bound the error probability p_{τ_ε} of the composite hypothesis test $\tilde{\Psi}_\varepsilon$ conditioned on $\mathcal{G}_{\tau_\varepsilon}$ using Lemma 4 and (36), obtaining

$$-\log p_{\tau_\varepsilon} + \log\frac{\lambda\varepsilon}{4} \leq \mathbb{E}\left[\sum_{\tau_\varepsilon < t \leq T} d\left(S_t, \boldsymbol{\Upsilon}_t^{\mathbf{1}_{\{W \in \mathcal{W}_1\}}}\right)\Big| \mathcal{G}_\tau\right].$$

Observe that the error event of the decoder Ψ is implied by the error event of $\tilde{\Psi}_\varepsilon$, so that in particular

$$p_e(\Phi, T, \Psi) = \mathbb{E}\left[\mathbb{P}\left(\hat{W} \neq W | \mathcal{G}_{\tau_\varepsilon}\right)\right]$$

$$\geq \mathbb{E}\left[\mathbb{P}\left(\mathbf{1}_{\{\hat{W} \in \mathcal{W}_1\}} \neq \mathbf{1}_{\{W \in \mathcal{W}_1\}}\big| \mathcal{G}_{\tau_\varepsilon}\right)\right]$$

$$= \mathbb{E}[p_{\tau_\varepsilon}].$$

Since the function $x \mapsto -\log x$ is decreasing and convex on the interval $(0, 1]$, we get

$$D_\varepsilon(\Phi, T) \geq \mathbb{E}\left[\mathbb{E}\left[\sum_{\tau_\varepsilon < t \leq T} d\left(S_t, \boldsymbol{\Upsilon}_t^{\mathbf{1}_{\{W \in \mathcal{W}_1\}}}\right)\Big| \mathcal{G}_{\tau_\varepsilon}\right]\right]$$

$$\geq \mathbb{E}[-\log p_{\tau_\varepsilon}] + \log\frac{\lambda\varepsilon}{4}$$

$$\geq -\log \mathbb{E}\left[p_{\tau_\varepsilon}\right] + \log \frac{\lambda \varepsilon}{4}$$

$$= -\log p_e(\Phi, T, \Psi) + \log \frac{\lambda \varepsilon}{4} \tag{40}$$

the last inequality in (40) following from the Jensen inequality. The claim now follows by taking a linear combination of (40) and (31). ∎

In the memoryless case ($|\mathcal{S}| = 1$), Burnashev's original result (see [6, Sec. 4.1], or [2, eq. (12)]) can be recovered from (38) by optimizing over the channel input distributions $\boldsymbol{\Upsilon}_t$, $\boldsymbol{\Upsilon}_t^0$, and $\boldsymbol{v}_t^1$.

In order to prove Part 1 of Theorem 1 it remains to consider infinite sequences of variable-length coding schemes with vanishing error probability and to show that asymptotically the upper bound in (38) reduces to the Burnashev exponent $E_B(R)$. This involves new technical challenges which will be the object of next section.

IV. MARKOV DECISION PROCESSES WITH STOPPING TIME HORIZONS

In this section, we shall recall some concepts about Markov decision processes which will allow us to asymptotically estimate the terms $C_\varepsilon(\Phi, T)$ and $D_\varepsilon(\Phi, T)$, respectively, in terms of the capacity C, defined in (6), and the Burnashev coefficient D, defined in (14), of the FSMC.

The main idea consists in interpreting the maximization of $C_\varepsilon(\Phi, T)$ and $D_\varepsilon(\Phi, T)$ as stochastic control problems with average-cost criterion [1]. The control is the channel input distribution chosen as a function of the available feedback information and the controller is identified with the encoder. The main novelty these problems present with respect to those traditionally addressed by MDP theory consists in the fact that, as a consequence of considering variable-length coding schemes, we shall need to deal with the situation when the horizon is neither finite (in the sense of being a deterministic constant) nor infinite (in the sense of being concerned with the asymptotic normalized average running cost), but rather it is allowed to be a stopping time. In order to handle this case, we adopt the convex-analytic approach, a technique first introduced by Manne in [20] (see also [11]) for the finite-state finite-action setting, and later developed in great generality by Borkar [4].

In Section IV-A, we shall first reformulate the problem of optimizing the terms $C_\varepsilon(\Phi, T)$ and $D_\varepsilon(\Phi, T)$ with respect to the causal encoder Φ. Then, we present a brief review of the convex-analytic approach to Markov decision problems in Section IV-B, presenting the main ideas and definitions. In Section IV-C, we will prove a uniform convergence theorem for the empirical measure process and use this result to treat the asymptotic case of the average-cost problem with stopping time horizon. The main result of this section is contained in Theorem 3, which is then applied in Section IV-D together with Theorem 2 in order to prove Part 1 of Theorem 1.

A. Markov Decision Problems With Stopping Time Horizons

We shall consider a controlled Markov chain over $\mathcal{S}$, with compact control space $\mathcal{U} := \mathcal{P}(\mathcal{X})$, the space of channel input distributions. Let $g : \mathcal{S} \times \mathcal{U} \to \mathbb{R}$ be a continuous (and thus bounded, since $\mathcal{U} = \mathcal{P}(\mathcal{X})$ is compact) cost function; in our application g will coincide either with the mutual information cost c defined in (5) or with the information divergence cost d defined in (9). We prefer to consider the general case in order to deal with both problems at once.

The evolution of the system is described by a state sequence $\boldsymbol{S} = (S_t)$, an output sequence $\boldsymbol{Y} = (Y_t)$ and a control sequence $\boldsymbol{u} = (U_t)$. If at time t the system is in state $S_t = s$ in $\mathcal{S}$, and a control $U_t = u$ in $\mathcal{U}$ is chosen according to some policy, then a cost $g(s, u)$ is incurred and the system produces the output $Y_t = y$ in $\mathcal{Y}$ and moves to next state $S_{t+1} = s_+$ in $\mathcal{S}$ according to the stochastic kernel $Q(s_+, y|s, u)$, defined in (3). Once the transition into next state has occurred, a new action is taken and the process is repeated.

At time t, the control U_t is allowed to be an $\mathcal{E}_t$-measurable r.v., where $\mathcal{E}_t = \sigma(\boldsymbol{S}_1^t, \boldsymbol{Y}_1^{t-1})$ is the encoder's feedback observation at time t; in other words we are assuming that $U_t = \pi_t\left(\boldsymbol{S}_1^t, \boldsymbol{Y}_1^{t-1}\right)$ for some map

$$\pi_t : \mathcal{S}^t \times \mathcal{Y}^{t-1} \to \mathcal{U}.$$

We define a *feasible policy* $\boldsymbol{\pi}$ as an infinite sequence (π_t) of such maps. Once a feasible policy $\boldsymbol{\pi}$ has been chosen, a joint probability distribution $\mathbb{P}_{\boldsymbol{\pi}}$ for state, control and output sequences is well defined; we will denote by $\mathbb{E}_{\boldsymbol{\pi}}$ the corresponding expectation operator.

Let τ be a stopping time for the filtration $\mathcal{G} = (\mathcal{G}_t)$ (recall that $\mathcal{G}_t = \mathcal{E}_{t+1}$ describes the encoder's feedback and state information at time $t+1$), and consider the following optimization problem: maximize

$$\frac{1}{\mathbb{E}_{\boldsymbol{\pi}}[\tau]} \mathbb{E}_{\boldsymbol{\pi}}\left[\sum_{1 \leq t \leq \tau} g\left(S_t, \pi_t(\boldsymbol{S}_1^t, \boldsymbol{Y}_1^{t-1})\right)\right] \tag{41}$$

over all feasible policies $\boldsymbol{\pi} = (\pi_t)$ such that $\mathbb{E}_{\boldsymbol{\pi}}[\tau]$ is finite.

Clearly, in the special case when τ is almost surely constant (41) reduces to the standard finite-horizon problem which is usually solved with dynamic-programming tools. Another special case is when τ is geometrically distributed and independent of the processes $\boldsymbol{S}$, $\boldsymbol{U}$, and $\boldsymbol{Y}$. In this case, (41) reduces to the so-called discounted problem which has been widely studied in the stochastic control literature [1]. However, what makes the problem nonstandard is that in (41) τ is allowed to be an arbitrary stopping time for the filtration $\mathcal{G}$, typically dependent on the processes $\boldsymbol{S}$, $\boldsymbol{U}$, and $\boldsymbol{Y}$.

B. The Convex-Analytic Approach

We review some of the ideas of the convex-analytic approach following [4].

A feasible policy $\boldsymbol{\pi}$ is said to be stationary if the current control depends on the current state only and is independent from the past state and output history and of the time, i.e., there exists a map $\pi : \mathcal{S} \to \mathcal{U}$ such that $\pi_t(\boldsymbol{S}_1^t, \boldsymbol{Y}_1^{t-1}) = \pi(s_t)$ for all t. We shall identify a stationary policy as above with the map $\pi : \mathcal{S} \to \mathcal{U}$ itself. It has already been noted in Section II-A that, for every stationary policy π, the stochastic matrix Q_π describing the state transition probabilities under π (see (4)) is irreducible, so that existence and uniqueness of a stationary distribution $\boldsymbol{\mu}_\pi$

in $\mathcal{P}(\mathcal{S})$ are guaranteed. It follows from the Perron–Frobenious theorem [10] that, if a stationary policy π is used, then the normalized running cost $1/n \sum_{t=1}^{n} g(S_t, \pi(S_t))$ converges $\mathbb{P}_\pi$-almost surely to the ergodic average $\sum_{s \in \mathcal{S}} \boldsymbol{\mu}_\pi(s) g(s, \pi(s))$. Define

$$G := \max_{\pi: \mathcal{S} \longrightarrow \mathcal{U}} \sum_{s \in \mathcal{S}} \boldsymbol{\mu}_\pi(s) g(s, \pi(s)). \qquad (42)$$

Observe that the optimization in the right-hand side of (42) has the same form of those in the definitions (5) and (14) of the capacity and the Burnashev coefficient of an ergodic FSMC. Notice that compactness of the space $\mathcal{U}^{\mathcal{S}}$ of all stationary policies and continuity of the cost $g(s, \pi(s))$ and of the invariant measure $\boldsymbol{\mu}_\pi$ as functions of the stationary policy π guarantee the existence of maximizer in (42).

We now consider stationary randomized policies. These are defined as maps $\tilde{\pi} : \mathcal{S} \longrightarrow \mathcal{P}(\mathcal{U})$, where $\mathcal{P}(\mathcal{U})$ denotes the space of probability measures on $\mathcal{U}$, equipped with its Prohorov topology, i.e., the topology induced by weak convergence (see [3, Ch. 2]). If $\tilde{\pi}$ is a stationary randomized policy, we shall use the notation $\tilde{\pi}(\cdot|s)$ for the probability measure in $\mathcal{P}(\mathcal{U})$ associated by $\tilde{\pi}$ to the state $s \in \mathcal{S}$. To any stationary randomized policy $\tilde{\pi}$, the following control strategy is associated: if at time t the state is S_t, then the control U_t is randomly chosen in the control space $\mathcal{U}$ with conditional distribution given the available information $\mathcal{E}_t = \sigma(\boldsymbol{S}_1^t, \boldsymbol{Y}_1^{t-1})$ equal to $\tilde{\pi}(\cdot|S_t)$. Observe that there are two levels of randomization. The control space itself has already been defined as the space of channel input probability distributions $\mathcal{P}(\mathcal{X})$, while the strategy associated to the stationary randomized policy $\tilde{\pi}$ chooses a control at random with conditional distribution $\tilde{\pi}(\cdot|S_t)$ in $\mathcal{P}(\mathcal{U}) = \mathcal{P}(\mathcal{P}(\mathcal{X}))$. Of course, randomized stationary policies are a generalization of deterministic stationary policies, since to any deterministic stationary policy $\pi : \mathcal{S} \longrightarrow \mathcal{U}$ it is possible to associate the randomized policy $\tilde{\pi}(\cdot|s) = \delta_{\pi(s)}$. To any randomized stationary policy $\tilde{\pi} : \mathcal{S} \longrightarrow \mathcal{P}(\mathcal{U})$ we associate the stochastic matrix describing the associated state transition probabilities

$$Q_{\tilde{\pi}}(s_+|s) := \int_{\mathcal{U}} Q(s_+|s, u) \tilde{\pi}(\mathrm{d}u|s). \qquad (43)$$

Similarly to the case of stationary deterministic policies, it is not difficult to conclude that, since $Q_{\tilde{\pi}}$ can be written as a convex combination of a finite number of stochastic matrices Q_f, with $f : \mathcal{S} \longrightarrow \mathcal{X}$, all of which are irreducible, then $Q_{\tilde{\pi}}$ itself is irreducible and thus admits a unique invariant state distribution $\boldsymbol{\mu}_{\tilde{\pi}}$ in $\mathcal{P}(\mathcal{S})$.

Now, consider the space of joint probability measures $\mathcal{P}(\mathcal{S} \times \mathcal{U})$; we shall denote the action of $\boldsymbol{\eta} \in \mathcal{P}(\mathcal{S} \times \mathcal{U})$ on a continuous function $h : \mathcal{S} \times \mathcal{U} \to \mathbb{R}$ by

$$\langle \boldsymbol{\eta}, h \rangle := \int_{\mathcal{S} \times \mathcal{U}} h(s, u) \mathrm{d}\boldsymbol{\eta}(s, u).$$

The following definition of occupation measure is at the heart of the convex-analytical approach.

Definition 5: For every stationary (randomized) policy $\tilde{\pi} : \mathcal{S} \longrightarrow \mathcal{P}(\mathcal{U})$, the occupation measure of $\tilde{\pi}$ is $\boldsymbol{\eta}_{\tilde{\pi}}$ in $\mathcal{P}(\mathcal{S} \times \mathcal{U})$ defined by

$$\langle \boldsymbol{\eta}_{\tilde{\pi}}, h \rangle := \sum_{s \in \mathcal{S}} \boldsymbol{\mu}_{\tilde{\pi}}(s) \int_{\mathcal{U}} h(s, u) \tilde{\pi}(\mathrm{d}u|s), \quad \forall h \in \mathcal{C}_b(\mathcal{S} \times \mathcal{U})$$

where $\boldsymbol{\mu}_{\tilde{\pi}}$ in $\mathcal{P}(\mathcal{S})$ is the invariant measure of the stochastic matrix $Q_{\tilde{\pi}}$, while $\mathcal{C}_b(\mathcal{S} \times \mathcal{U})$ is the space of bounded continuous maps from $\mathcal{S} \times \mathcal{U}$ to $\mathbb{R}$.

The occupation measure $\boldsymbol{\eta}_\pi$ can be viewed as the long-time empirical frequency of the joint state–control process governed by the stationary (randomized) policy π. In fact, for every time n, we can associate to the controlled Markov process the empirical measure $\boldsymbol{v}_n$ which is a $\mathcal{P}(\mathcal{S} \times \mathcal{U})$-valued r.v. sample-path-wise defined by

$$\langle \boldsymbol{v}_n, h \rangle := \frac{1}{n} \sum_{1 \le t \le n} h(S_t, U_t), \quad \forall h \in \mathcal{C}_b(\mathcal{S} \times \mathcal{U}). \qquad (44)$$

Observe that $\boldsymbol{v}_n$ is a probability measure on the product space $\mathcal{S} \times \mathcal{U}$, and is itself an r.v. since it is defined as a function of the joint state control random process $(\boldsymbol{S}_1^t, \boldsymbol{U}_1^t)$.

If the process is controlled by a stationary (randomized) policy π and the initial state is distributed accordingly to $\boldsymbol{\mu}_\pi$, then, for any continuous function $h \in \mathcal{C}_b(\mathcal{S} \times \mathcal{U})$, the time average $\langle \boldsymbol{v}_n, h \rangle$ converges almost surely to the ergodic average

$$\langle \boldsymbol{\eta}_\pi, h \rangle = \int_{\mathcal{S} \times \mathcal{U}} h(s, u) \mathrm{d}\boldsymbol{\eta}_\pi(s, u)$$

the by the ergodic theorem. Therefore, at least in this case, we have

$$\lim_n \boldsymbol{v}_n = \boldsymbol{\eta}_\pi \qquad \mathbb{P}_\pi\text{-a.s.} \qquad (45)$$

where the convergence of the empirical measure sequence $(\boldsymbol{v}_n)$ to the occupation measure $\boldsymbol{\eta}_\pi$ is intended in the weak sense.[7]

We shall denote by K the set of the occupation measures associated to all the stationary randomized policies, i.e.,

$$K := \{\boldsymbol{\eta}_{\tilde{\pi}} | \tilde{\pi} : \mathcal{S} \longrightarrow \mathcal{P}(\mathcal{U})\} \subseteq \mathcal{P}(\mathcal{S} \times \mathcal{U}) \qquad (46)$$

and by K_e the set of all occupation measures associated to stationary deterministic policies

$$K_e := \{\boldsymbol{\eta}_\pi | \pi : \mathcal{S} \longrightarrow \mathcal{U}\} \subseteq \mathcal{P}(\mathcal{S} \times \mathcal{U}).$$

It is known (see [4]) that both K and K_e are closed subsets of $\mathcal{P}(\mathcal{S} \times \mathcal{U})$. Moreover, K is convex and K_e coincides with the set of extreme points of K (see Fig. 2). Furthermore, it is possible to characterize K as the set of zeros of the continuous linear functional $\boldsymbol{F} : \mathcal{P}(\mathcal{S} \times \mathcal{U}) \longrightarrow [-1, 1]^{\mathcal{S}}$

$$F_s(\boldsymbol{\eta}) := \boldsymbol{\eta}(\{s\}, \mathcal{U}) - \int_{\mathcal{S} \times \mathcal{U}} Q_S(s|j, u) \mathrm{d}\boldsymbol{\eta}(j, u)$$

[8] i.e. (see [4])

$$K = \{\boldsymbol{\eta} \in \mathcal{P}(\mathcal{S} \times \mathcal{U}) : \boldsymbol{F}(\boldsymbol{\eta}) = \boldsymbol{0}\}. \qquad (47)$$

---

[7]Recall that a sequence of probability measures $(\boldsymbol{\nu}_n)$ on a topological space $\mathcal{A}$ is said to be weakly convergent to some $\boldsymbol{\nu} \in \mathcal{P}(\mathcal{A})$ if $\lim_n \int_{\mathcal{A}} f(a) \mathrm{d}\boldsymbol{\nu}_n(a) = \int_{\mathcal{A}} f(a) \mathrm{d}\boldsymbol{\nu}(a)$ for all bounded and continuos test functions $f \in \mathcal{C}_b(\mathcal{A})$.

[8]Here $\boldsymbol{\eta}(\{s\}, \mathcal{U}) = \int_{\mathcal{U}} \boldsymbol{\eta}(\{s\}, u) \mathrm{d}u$ denotes the S-marginal of $\boldsymbol{\eta}$ evaluated in s.
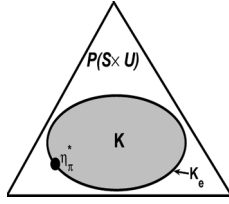
Fig. 2. A schematic representation of the optimization problem (3). The large triangular space is the infinite-dimensional Prohorov space $\mathcal{P}(\mathcal{S} \times \mathcal{U})$. Its gray-shaped subset represents the close convex set K of all occupation measures. The set of extreme points of K is K_e and corresponds to the set of all occupation measures associated to stationary deterministic policies. The optimal value of the linear functional $\boldsymbol{\eta} \mapsto \langle \boldsymbol{\eta}, g \rangle$ happens to be achieved on K_e and thus corresponds to the occupation measure $\boldsymbol{\eta}_\pi^*$ associated to an optimal deterministic stationary policy $\pi^* : \mathcal{S} \to \mathcal{P}(\mathcal{X})$.

In fact, it is possible to think of $\|\boldsymbol{F}(\boldsymbol{\eta})\|$[9] as a measure of how far the $\mathcal{S}$-marginal of a measure $\boldsymbol{\eta}$ in $\mathcal{P}(\mathcal{S} \times \mathcal{U})$ is from being invariant for the state process.

If one were interested in optimizing the infinite-horizon running average cost

$$\liminf_n \frac{1}{n} \mathbb{E}_\pi \left[\sum_{1 \le t \le n} g(S_t, U_t) \right] = \liminf_n \mathbb{E}_\pi \left[\langle \boldsymbol{v}_n, g \rangle \right]$$

over all (randomized) stationary policies π, then (45)–(47) would immediately lead to the following linear programming problem:

$$\max_{\boldsymbol{\eta} \in K} \langle \boldsymbol{\eta}, g \rangle = \max_{\substack{\boldsymbol{\eta} \in \mathcal{P}(\mathcal{S} \times \mathcal{U}): \\ \boldsymbol{F}(\boldsymbol{\eta}) = 0}} \langle \boldsymbol{\eta}, g \rangle.$$

We notice that, since $\mathcal{U}$ is compact and $\mathcal{S}$ is finite, $\mathcal{P}(\mathcal{S} \times \mathcal{U})$ is compact in the Prohorov metric [3] (i.e., sequentially compact under weak convergence). Thus, both K and K_e are compact. Hence, since the map

$$\mathcal{P}(\mathcal{S} \times \mathcal{U}) \ni \boldsymbol{\eta} \longmapsto \langle \boldsymbol{\eta}, g \rangle \in \mathbb{R}$$

is continuous (in the Prohorov topology), it achieves its maxima both on K and K_e. Moreover, such a map is linear so that these maxima do coincide, i.e., the maximum over K is achieved in an extreme point (see Fig. 3). Thus, we have the following chain of equalities:

$$\begin{aligned} G &= \max_{\pi: \mathcal{S} \to \mathcal{U}} \sum_{s \in \mathcal{S}} \boldsymbol{\mu}_\pi(s) g(s, \pi(s)) \\ &= \max_{\pi: \mathcal{S} \to \mathcal{U}} \langle \boldsymbol{\eta}_\pi, g \rangle \\ &= \max_{\boldsymbol{\eta} \in K_e} \langle \boldsymbol{\eta}, g \rangle \\ &= \max_{\substack{\boldsymbol{\eta} \in \mathcal{P}(\mathcal{S} \times \mathcal{U}): \\ \boldsymbol{F}(\boldsymbol{\eta}) = 0}} \langle \boldsymbol{\eta}, g \rangle. \end{aligned} \quad (48)$$

Hence, the optimal occupation measure $\boldsymbol{\eta}^*$ is induced by a stationary deterministic policy π^*, and is therefore given by

$$\langle \boldsymbol{\eta}^*, h \rangle = \sum_{s \in \mathcal{S}} \boldsymbol{\mu}_{\pi^*}(s) h(s, \pi^*(s)), \quad h \in \mathcal{C}_b(\mathcal{S} \times \mathcal{U})$$

[9]Here and throughout the rest of the paper $\|\boldsymbol{X}\| := \max_i |x_i|$ will denote the L^∞-norm of a vector $\boldsymbol{X}$.
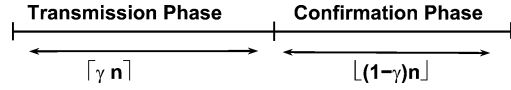

Fig. 3. One epoch in the generalized Yamamoto–Itoh scheme: A total length n is divided into two phases: a transmission one of length $\hat{n} = \lceil \gamma n \rceil$ and a confirmation one of length $\hat{n} = \lfloor (1 - \gamma) n \rfloor$.

where $\boldsymbol{\mu}_{\pi^*} \in \mathcal{P}(\mathcal{S})$ is the invariant state distribution induced by the policy π^*. Observe that in the last term in (48), both the constraints and the object functionals are linear. This indicates (infinite-dimensional) linear programming as a possible approach for computing G, alternative to the dynamic programming ones based on policy or value iteration techniques [1], [4]. Moreover, it points out to an easy way to generalize the theory taking into account average cost constraints (see [23] where the Burnashev exponent of DMCs with average cost constraints is studied). In fact, in the convex-analytic approach these merely translate into additional constraints for the linear program.

C. An Asymptotic Solution to Markov Decision Problems With a Stopping Time Horizon

It is known that, under the ergodicity and continuity assumptions we have made, G defined in (42) is the sample-path optimal value for the infinite horizon problem with cost g not only over the set of all stationary policies, but also over the larger set of all feasible policies (actually over all admissible policies, see [4]). This means that, for every feasible policy $\boldsymbol{\pi} = (\pi_t)$

$$\limsup_n \frac{1}{n} \sum_{1 \le t \le n} g\left(S_t, \pi_t(\boldsymbol{S}_1^t, \boldsymbol{Y}_1^{t-1})\right) \le G, \quad \mathbb{P}_{\boldsymbol{\pi}}\text{-a.s.} \quad (49)$$

For a sequence of admissible policies $(\boldsymbol{\pi}^{(n)})$, let $\mathbb{P}_{(n)}$ and $\mathbb{E}_{(n)}$ denote the probability and expectation operators induced by $\boldsymbol{\pi}^{(n)}$. It is known that

$$\begin{aligned} G &\ge \limsup_n \frac{1}{n} \mathbb{E}_{(n)} \left[\sum_{1 \le t \le n} g\left(S_t, \pi_t^{(n)}\left(\boldsymbol{S}_1^t, \boldsymbol{Y}_1^{t-1}\right)\right) \right] \\ &= \limsup_n \frac{1}{n} \mathbb{E}_{(n)} \left[\sum_{1 \le t \le n} g(S_t, U_t) \right] \end{aligned} \quad (50)$$

i.e., the limit of the optimal values of finite horizon problems coincides with infinite horizon optimal value. Inequality (50) can be proven by using dynamic programming arguments based on the Bellman principle of optimality. As shown in [32], (50) is useful in characterizing the capacity of channels with memory and feedback with fixed-length codes. Actually, a much more general result than (50) can be proved, as explained in the sequel.

In the convex-analytic approach, the key point in the proof of (49) consists in showing that, under any, not necessarily stationary, feasible policy $\boldsymbol{\pi}$, the empirical measure process $(\boldsymbol{v}_n)$ as defined in (44) converges $\mathbb{P}_{\boldsymbol{\pi}}$-almost surely to the set K. The way this is usually proven is by using a martingale central limit theorem in order to show that the finite-dimensional process $\boldsymbol{F}(\boldsymbol{v}_n)$ converges to $\boldsymbol{0}$ almost surely. The following is a stronger result, providing an exponential upper bound on the tails of the random sequence $(\|\boldsymbol{F}(\boldsymbol{v}_n)\|)$.

Lemma 6: For every $\varepsilon > 0$, and for every feasible policy $\boldsymbol{\pi}$

$$\mathbb{P}_{\boldsymbol{\pi}}\left(\|\boldsymbol{F}(\boldsymbol{v}_n)\| \geq \varepsilon + \frac{1}{n}\right) \leq 2|\mathcal{S}|\exp\left(\frac{-n\varepsilon^2}{2}\right). \quad (51)$$

Proof: See the Appendix. □

We emphasize the fact that the bound (51) is uniform with respect to the choice of the feasible policy $\boldsymbol{\pi}$. It is now possible to drive conclusions on the tails of the running average cost $1/n \sum_{t=1}^{n} g(S_t, U_t)$ based on (51). The core idea is the following. From (44), we can rewrite the normalized running cost as

$$\frac{1}{n}\sum_{1 \leq t \leq n} g(S_t, U_t) = \langle \boldsymbol{v}_n, g \rangle.$$

Since the map $\boldsymbol{\eta} \mapsto \langle \boldsymbol{\eta}, g \rangle$ is continuous over $\mathcal{P}(\mathcal{S} \times \mathcal{U})$, and $G = \max\{\langle \boldsymbol{\eta}, g \rangle | \boldsymbol{\eta} \in K\}$, we have that, whenever $\boldsymbol{\nu}_n$ is close to the set K, the quantity $\langle \boldsymbol{\nu}_n, g \rangle$ cannot be much larger than G. It follows that, if with high probability $\boldsymbol{v}_n$ is close enough to K, then with high probability $\langle \boldsymbol{v}_n, g \rangle$ cannot be much larger than G. In order to show that with high probability $\boldsymbol{v}_n$ is close to K, we want to use (51). In fact, if for some $\boldsymbol{\eta}$ in $\mathcal{P}(\mathcal{S} \times \mathcal{U})$ the quantity $\|\boldsymbol{F}(\boldsymbol{\eta})\|$ is very small, then $\boldsymbol{\eta}$ is necessarily close to G. More precisely, we define the function $\gamma : \mathbb{R}^+ \rightarrow \mathbb{R}$

$$\gamma(\varepsilon) := \sup\left\{\langle \boldsymbol{\eta}, g \rangle | \boldsymbol{\eta} \in \mathcal{P}(\mathcal{S} \times \mathcal{U}) : \|\boldsymbol{F}(\boldsymbol{\eta})\| \leq \varepsilon\right\}.$$

Clearly, γ is nondecreasing and $\gamma(0) = G$. Moreover, we have the following result.

Lemma 7: The map γ is upper semicontinuous (i.e., $\varepsilon_n \rightarrow \varepsilon$ $\Rightarrow \limsup_n \gamma(\varepsilon_n) \leq \gamma(\varepsilon)$)

Proof: See the Appendix. □

Let us now introduce the random process (G_n)

$$G_n := \sup_{t \geq n}\langle \boldsymbol{v}_t, g \rangle, \quad n \in \mathbb{N}.$$

Clearly, the process (G_n) is sample-path-wise nonincreasing in n.

Lemma 8: Let $(\boldsymbol{\pi}^{(n)})$ be a sequence of feasible policies, and $(\tau^{(n)})$ be a sequence of stopping times[10] such that for every n $\mathbb{E}_{(n)}[\tau^{(n)}] < \infty$, while

$$\lim_n \mathbb{P}_{(n)}\left(\tau^{(n)} \leq M\right) = 0, \quad \forall M \in \mathbb{N}. \quad (52)$$

Then

$$\lim_n \mathbb{P}_{(n)}\left(G_{\tau^{(n)}} > \gamma(\varepsilon)\right) = 0, \quad \forall \varepsilon > 0. \quad (53)$$

Proof: See the Appendix. □

The following result can be considered as an asymptotic estimate of (41). It consists in a generalization of (50) from a deterministic increasing sequence of time horizons to a sequence of stopping times satisfying (52).

[10]All the $\tau^{(n)}$ are assumed to be stopping times with respect to the filtration $\mathcal{G}$, i.e., we assume that the event $\{\tau^{(n)} \leq t\}$ is $\mathcal{G}_t$-measurable for all t and n, where we recall that $\mathcal{G}_t = \sigma(\boldsymbol{S}_1^{t+1}, \boldsymbol{Y}_1^t)$.

Theorem 3: Let $(\boldsymbol{\pi}^{(n)})$ be a sequence of feasible policies, and $(\tau^{(n)})$ be a sequence of stopping times such that for every n $\mathbb{E}_{(n)}[\tau^{(n)}] < \infty$, while (52) holds true. Then

$$\limsup_n \frac{1}{\mathbb{E}_{(n)}\left[\tau^{(n)}\right]}\mathbb{E}_{(n)}\left[\sum_{1 \leq t \leq \tau^{(n)}} g(S_t, U_t)\right] \leq G. \quad (54)$$

Proof: Let us fix an arbitrary $\varepsilon > 0$, and for $n \geq 1$ define the event $A_n := \{G_{\tau^{(n)}} \leq \gamma(\varepsilon)\}$.

By applying Lemma 8, we obtain

$$\begin{aligned}
\mathbb{E}_{(n)}\Big[\sum_{t=1}^{\tau^{(n)}} g(S_t, U_t)\Big] &= \mathbb{E}_{(n)}\left[\tau^{(n)}\langle \boldsymbol{v}_{\tau^{(n)}}, g \rangle\right]\\
&= \mathbb{E}_{(n)}\left[\tau^{(n)}\langle \boldsymbol{v}_{\tau^{(n)}}, g \rangle \mathbb{1}_{A_n}\right]\\
&\quad + \mathbb{E}_{(n)}\left[\tau^{(n)}\langle \boldsymbol{v}_{\tau^{(n)}}, g \rangle \mathbb{1}_{\overline{A_n}}\right]\\
&\leq \mathbb{E}_{(n)}\left[\tau^{(n)}\right]\left(\gamma(\varepsilon) + g_{\max}\mathbb{P}_{(n)}\left(\overline{A_n}\right)\right)
\end{aligned}$$

where $g_{\max} := \max\{g(s, u) | s \in \mathcal{S}, u \in \mathcal{U}\}$. From (53) we get

$$\begin{aligned}
\gamma(\varepsilon) &= \gamma(\varepsilon) + g_{\max}\limsup_n \mathbb{P}_{(n)}\left(\overline{A_n}\right)\\
&\geq \limsup_n \frac{\mathbb{E}_{(n)}\left[\sum_{1 \leq t \leq \tau^{(n)}} g(S_t, U_t)\right]}{\mathbb{E}_{(n)}\left[\tau^{(n)}\right]}.
\end{aligned}$$

Therefore, (54) follows from the arbitrariness of $\varepsilon > 0$, and the fact that, as a consequence of Lemma 7, we have $\lim_{\varepsilon \downarrow 0} \gamma(\varepsilon) = G$. ∎

D. An Asymptotic Upper Bound on the Error Exponent of a Sequence of Variable-Length Block-Coding Schemes

We are now ready to step back to the problem of estimating the error exponent of variable-length block-coding schemes over FSMCs. We want to combine the result in Theorem 2 with that in Theorem 3 in order to finally prove Part 1 of Theorem 1.

Let $(\Phi^{(n)}, T^{(n)}, \Psi^{(n)})$ be a sequence of variable-length block-coding schemes satisfying (21). Our goal is to prove that

$$\limsup_n \frac{-\log p_e\left(\Phi^{(n)}, T^{(n)}, \Psi^{(n)}\right)}{\mathbb{E}[T^{(n)}]} \leq D\left(1 - \frac{R}{C}\right). \quad (55)$$

A first simple conclusion which can be drawn from Theorem 2, using the crude bounds

$$c(S_t, \boldsymbol{\Upsilon}_t) \leq \log|\mathcal{X}|, \quad d(S_t, \boldsymbol{\Upsilon}_t^i) \leq d_{\max}, \quad i = 0, 1$$

is that

$$\begin{aligned}
&\limsup_n \frac{-\log p_e\left(\Phi^{(n)}, T^{(n)}, \Psi^{(n)}\right)}{\mathbb{E}[T^{(n)}]}\\
&\leq \frac{D}{C}\log|\mathcal{X}| + d_{\max} - R(1 - \varepsilon) - \log\frac{\lambda\varepsilon}{4} + \frac{D}{C}H(\varepsilon). \quad (56)
\end{aligned}$$

Thus, the error probability does not decay to zero faster than exponentially with the expected transmission time $\mathbb{E}[T^{(n)}]$.

The core idea to prove (55) consists in introducing a sequence (ε_n) of positive reals and showing that both

$$\tau^{(n)} := \min\left\{T^{(n)}, \inf\left\{t \in \mathbb{N} \middle| \Pi_t^{(n)} \leq \varepsilon_n\right\}\right\} \quad (57)$$

(where $\Pi_t^{(n)}$ denotes the MAP error probability of the encoder $\Phi^{(n)}$ given the observation $\mathcal{G}_t$) and $T^{(n)} - \tau^{(n)}$ diverge in the sense of satisfying (52). The sequence (ε_n) needs to be carefully chosen: we want it to be asymptotically vanishing in order to guarantee that $\tau^{(n)}$ diverges, but not too fast since otherwise $T^{(n)} - \tau^{(n)}$ would not diverge. It turns out that one possible good choice is

$$\varepsilon_n := \frac{-1}{\log p_e\left(\Phi^{(n)}, T^{(n)}, \Psi^{(n)}\right)}.$$

It is immediate to verify that, if $\lim_n p_e\left(\Phi^{(n)}, T^{(n)}, \Psi^{(n)}\right) = 0$, then

$$\lim_n \varepsilon_n = 0, \quad \lim_n \frac{1}{\varepsilon_n} p_e\left(\Phi^{(n)}, T^{(n)}, \Psi^{(n)}\right) = 0. \quad (58)$$

Lemma 9: Let $(\Phi^{(n)}, T^{(n)}, \Psi^{(n)})$ be a sequence of variable-length block-coding schemes satisfying (21). For every $n \in \mathbb{N}$, define $\tau^{(n)}$ as in (57). Then

$$\lim_n \mathbb{P}\left(\tau^{(n)} \leq M\right) = 0, \quad \forall M \in \mathbb{N}. \quad (59)$$

Moreover, for any choice of a $2^{\mathcal{W}}$-valued $\mathcal{G}_{\tau^{(n)}}$-measurable r.v. $\mathcal{W}^{(n)}$ such that

$$\lambda \varepsilon_n \leq \mathbb{P}(W \in \mathcal{W}^{(n)}) \leq 1 - \lambda \varepsilon_n,$$

we have

$$\lim_n \mathbb{P}\left(T^{(n)} - \tau^{(n)} \leq M \middle| W \in \mathcal{W}^{(n)}\right) = 0, \quad \forall M \in \mathbb{N}. \quad (60)$$

Proof: See the Appendix. $\qquad\square$

Thanks to (59), we can apply Theorem 3 to the mutual information cost c obtaining

$$\limsup_n \frac{C_{\varepsilon_n}(\Phi^{(n)}, T^{(n)})}{\mathbb{E}[\tau^{(n)}]}$$

$$= \limsup_n \frac{\mathbb{E}\left[\sum\limits_{1 \leq t \leq \tau^{(n)}} c(S_t, \boldsymbol{\Upsilon}_t)\right]}{\mathbb{E}[\tau^{(n)}]} \leq C. \quad (61)$$

Similarly, (60) allows us to apply Theorem 3 to the information divergence cost d, obtaining

$$\limsup_n \frac{D_{\varepsilon_n}(\Phi^{(n)}, T^{(n)})}{\mathbb{E}[T^{(n)} - \tau^{(n)}]} \leq D. \quad (62)$$

Therefore, by applying (61) and (62) first, and then Theorem 2, we get

$$D \geq \limsup_n \frac{1}{\mathbb{E}[T^{(n)}]}\left(\frac{D}{C}C_{\varepsilon_n}(\Phi^{(n)}, T^{(n)}) + D_{\varepsilon_n}(\Phi^{(n)}, T^{(n)})\right)$$

$$\geq \limsup_n -\frac{1}{\mathbb{E}[T^{(n)}]}\log p_e(\Phi^{(n)}, T^{(n)}, \Psi^{(n)})$$

$$+ \frac{D}{C}\frac{\log|\mathcal{W}^{(n)}|}{\mathbb{E}[T^{(n)}]}(1 - \alpha_n) + \frac{\beta_n}{\mathbb{E}[T^{(n)}]}$$

$$\geq \frac{D}{CR} + \limsup_n -\frac{1}{\mathbb{E}[T^{(n)}]}\log p_e\left(\Phi^{(n)}, T^{(n)}, \Psi^{(n)}\right)$$

where α_n and β_n are defined as in (39), and the last step follows from (58). Hence, we have proved (22).

V. AN ASYMPTOTICALLY OPTIMAL SCHEME

In this section, we propose and analyze a sequence of variable-length block-coding schemes with feedback asymptotically achieving the Burnashev exponent $E_B(R)$, thus proving Part 2 of Theorem 2.

The proposed scheme can be viewed as a generalization of the one introduced by Yamamoto and Itoh in [38] and consists of a sequence of epochs. Each epoch is made up of two distinct fixed-length transmission phases, respectively named communication and confirmation phase (see Fig. 3). In the communication phase, the message to be sent is encoded through a block code and transmitted over the channel. At the end of this phase, the decoder makes a tentative decision about the message sent based on the observation of the channel outputs and of the state sequence. As perfect causal feedback is available at the encoder, the result of this decision is known at the encoder. In the confirmation phase, a binary message, acknowledging the decoder's estimation if it is correct, or denying it if it is wrong, is sent by the encoder through a fixed-length repetition code function. The decoder performs a binary hypothesis test in order to decide whether a deny or an acknowledge message has been sent. If an acknowledge is detected, the transmission halts, while if a deny is detected the system restarts with a new epoch, transmitting the same message with the same protocol.

More precisely, we design our scheme as follows. Given a design rate R in $(0, C)$, let us fix an arbitrary γ in $(R/C, 1)$. For every n in $\mathbb{N}$, consider a message set $\mathcal{W}^{(n)}$ of cardinality $|\mathcal{W}^{(n)}| = \exp(\lfloor nR \rfloor)$ and two block lengths $\hat{n}$ and $\tilde{n}$, respectively defined as $\hat{n} = \lceil n\gamma \rceil$, $\tilde{n} := n - \hat{n}$.

Fixed-length block coding for the communication phase: It is known from previous works (see [32] for instance) that the capacity C of the stationary Markov channel we are considering is achievable by fixed-length block-coding schemes. Thus, since the rate of the communication phase is kept below capacity

$$\hat{R} := \lim_n \frac{\log|\mathcal{W}^{(n)}|}{\hat{n}} = \frac{R}{\gamma} < C$$

there exists a sequence of causal encoders $\hat{\Phi}^{(n)} = \left(\mathcal{W}^{(n)}, (\hat{\phi}_t^{(n)})\right)$ with $\hat{\phi}_t^{(n)} : \mathcal{W}^{(n)} \times \mathcal{S}^t \times \mathcal{Y}^{t-1} \to \mathcal{X}$, and a corresponding sequence of fixed-length-$\hat{n}$ decoders $(\hat{\Psi}^{(n)})$ (notice that n is the sequence index while $\hat{n}$ is the block length) with error probability asymptotically vanishing in n (see [32, Theorem 5.3, special case of Sec. 8.1.2]). More precisely, since the state space $\mathcal{S}$ is finite, the pair $\hat{\Phi}^{(n)}$ and $\hat{\Psi}^{(n)}$ can be designed in such a way that the probability $\mathbb{P}(\hat{\Psi}^{(n)} \neq W | W = w, S_1 = s)$ of error conditioned on the transmission of any message w in $\mathcal{W}_n$ and of an initial state s approaches zero uniformly with respect both to w and s, i.e.,

$$p(n) := \max_{w \in \mathcal{W}^{(n)}} \max_{s \in \mathcal{S}} \mathbb{P}\left(\hat{\Psi}^{(n)} \neq W | W = w, S_1 = s\right)$$

$$\xrightarrow{n \to \infty} 0. \quad (63)$$

The triple $(\hat{\Phi}^{(n)}, \hat{n}, \hat{\Psi}^{(n)})$ will be used in the first phase of each epoch of our iterative transmission scheme.

Binary hypothesis test for the confirmation phase: For the second phase, instead, we consider a causal binary input encoder $\tilde{\Phi}^{(n)}$ based on the optimal stationary policies of the maximization problem (15). More specifically, for $1 \leq t \leq n$, define

$$\tilde{\phi}_t^{(n)} : \{a,b\} \times \mathcal{S}^t \to \mathcal{X}, \quad \tilde{\phi}_t^{(n)}(m, \boldsymbol{S}) = f_m^*(s_t)$$

where $f_a^*, f_b^* : \mathcal{S} \to \mathcal{X}$ are such that

$$D = \sum_{s \in \mathcal{S}} \boldsymbol{\mu}_{f_a^*}(s) D\left(P(\cdot, \cdot | s, f_a^*(s)) \| P(\cdot, \cdot | s, f_b^*(s))\right).$$

Suppose that a confirmation message $m \in \{a,b\}$ is sent. Then it is easy to verify that the pair sequence $(S_{t+1}, Y_t)_{t=1}^{\tilde{n}}$ forms a Markov chain over the space of the achievable channel state output pairs

$$\mathcal{Z} := \bigcup_s \mathcal{Z}_s = \{(s_+, y) | \exists s, x : P(s_+, y | s, x) > 0\} \quad (64)$$

with transition probability matrix

$$P_m = \left(P_m(s_+, y | s, y_-) := P(s_+, y | s, f_m^*(s))\right).$$

It follows that a decoder for $\tilde{\Phi}^{(n)}$ performs a binary hypothesis test between two Markov chain hypotheses. Notice that for both chains the transition probabilities $P_m(s_+, y | s, y_-)$ do not depend on the second component y_- of the past state, but on its first component s only, as well as on the full future state (s_+, y).

When the coefficient D is finite, as a consequence of Assumption 1 and (13), we have that the stochastic matrix P_m is irreducible over $\mathcal{Z}$, with ergodic measure $\hat{\boldsymbol{\mu}}_m \in \mathcal{P}(\mathcal{Z})$ given by

$$\tilde{\boldsymbol{\mu}}_m(s_+, y) := \sum_{s \in \mathcal{S}} \boldsymbol{\mu}_{f_m}(s) P(s_+, y | s, f_m(s)), \quad m \in \{a,b\}.$$

Using known results on binary hypothesis tests for irreducible Markov chains (see [5], [24] and [10, pp. 72–82]) it is possible to show that a decoder

$$\tilde{\Psi}^{(n)} : (\mathcal{S} \times \mathcal{Y})^{\tilde{n}-1} \to \{a,b\}$$

can be chosen in such a way that, asymptotically in n, its type-b error probability achieves the exponent (recall (15))

$$\begin{aligned}
D &= \sum_{s \in \mathcal{S}} \boldsymbol{\mu}_a(s) D\left(P(\cdot, \cdot | s, f_a^*(s)) \| P(\cdot, \cdot | s, f_b^*(s))\right) \\
&= \sum_{z, z_+} \tilde{\boldsymbol{\mu}}_a(z) P_a(z_+ | z) \log \frac{\tilde{\boldsymbol{\mu}}_a(z) P_a(z_+ | z)}{\tilde{\boldsymbol{\mu}}_a(z) P_b(z_+ | z)} \\
&= \sum_{s, s_+, y_-, y} \tilde{\boldsymbol{\mu}}_a(s, y_-) P(s_+, y | s, f_a^*(s)) \\
&\quad \times \log \frac{P(s_+, y | s, f_a^*(s))}{P(s_+, y | s, f_b^*(s))}
\end{aligned}$$

while its type-a error probability is vanishing. More specifically, since the state space is finite, we have that, defining $p_m(n)$ as the maximum over all possible initial states of the error probability

of the pair $(\tilde{\Phi}^{(n)}, \tilde{\Psi}^{(n)})$ conditioned on the transmission of a confirmation message $m \in \{a,b\}$, i.e.,

$$p_m(n) := \max_{s \in \mathcal{S}} \mathbb{P}\left(\tilde{\Psi}^{(n)}\left(\boldsymbol{S}_2^{\tilde{n}}, \boldsymbol{Y}_1^{\tilde{n}-1}\right) \neq m | W = m, S_1 = s\right)$$

we have

$$\lim_n p_a(n) = 0, \quad \lim_n \frac{-\log p_b(n)}{\tilde{n}} = D. \quad (65)$$

When the coefficient D is infinite, then the stochastic matrix P_a is irreducible over the set $\mathcal{Z}$ of reachable state output pairs (this is because, by Assumption 1 all states $s \in \mathcal{S}$ are reachable, while by (10) every state output pair (s_+, y) in $\mathcal{Z}_s$ is reachable from s), and there exists at least two pairs (s, y_-) and (s_+, y) in $\mathcal{Z}$ such that $P_a(s_+, y | s, y_-) > 0$ while $P_b(s_+, y | s, y_-) = 0$. It follows that a sequence of binary tests $(\tilde{\Psi}^{(n)})$, with $\tilde{\Psi}^{(n)} : (\mathcal{S} \times \mathcal{Y})^{\tilde{n}-1} \to \{a,b\}$, can be designed such that

$$\lim_n p_a(n) = 0, \quad p_b(n) = 0, \ n \in \mathbb{N}. \quad (66)$$

Such a sequence of tests is given for instance by defining $\Psi^{(n)}(\boldsymbol{z})$ equal to a if and only if the $(\tilde{n}-1)$-tuple $\boldsymbol{z}$ contains a symbol z_- followed by a z_+. $\qquad \square$

Once fixed $\hat{\Phi}^{(n)}$, $\hat{\Psi}^{(n)}$, $\tilde{\Phi}^{(n)}$, and $\tilde{\Psi}^{(n)}$, the iterative protocol described above defines a variable-length block-coding scheme $(\Phi^{(n)}, T^{(n)}, \Psi^{(n)})$. As mentioned above, the scheme consists of a sequence of epochs, each of fixed length n; in particular, we have

$$T^{(n)} = n\zeta^{(n)}$$

where

$$\zeta^{(n)} := \inf\left\{k \geq 1 : \tilde{\Psi}^{(n)}\left(\boldsymbol{S}_{(k-1)n+\hat{n}+1}^{kn}, \boldsymbol{Y}_{(k-1)n+\hat{n}+1}^{kn}\right) = a\right\}$$

is a positive integer valued r.v. describing the number of epochs occurred until transmission halts.

The following result characterizes the asymptotic performance of the sequence of schemes $(\Phi^{(n)}, T^{(n)}, \Psi^{(n)})$. Its proof uses arguments similar to those in [34, Sec. III-B].

Theorem 4: For every design rate R in $(0, C)$, and every γ in $(R/C, 1)$, it holds that

$$\lim_n \frac{\log |\mathcal{W}^{(n)}|}{\mathbb{E}[T^{(n)}]} = R \quad (67)$$

and
- if $D < +\infty$

$$\lim_n \frac{-\log p_e(\Phi^{(n)}, T^{(n)}, \Psi^{(n)})}{\mathbb{E}[T^{(n)}]} = D(1 - \gamma), \quad (68)$$

- if $D = +\infty$

$$p_e(\Phi^{(n)}, T^{(n)}, \Psi^{(n)}) = 0, \quad n \in \mathbb{N}. \quad (69)$$

Proof: We introduce the following notation. First, for every $k \in \mathbb{N}$
- $\hat{e}_k := \left\{\hat{\Psi}\left(\boldsymbol{S}_{(k-1)n+1}^{(k-1)n+\hat{n}+1}, \boldsymbol{Y}_{(k-1)n+1}^{(k-1)n+\hat{n}}\right) \neq W\right\}$ is the error event in the communication phase of the kth epoch;

- $\tilde{e}_k := \left\{ \tilde{\Psi}\left(\boldsymbol{S}^{kn}_{(k-1)n+\hat{n}+2}, \boldsymbol{Y}^{kn-1}_{(k-1)n+\hat{n}+1} \right) \neq \mathbb{1}_{\hat{e}_k} \right\}$ is the error event in confirmation phase of the kth epoch.

Clearly, we have

$$\mathbb{P}(\hat{e}_k|\mathcal{F}_{(k-1)n}) \leq p(n)$$
$$\mathbb{P}(\tilde{e}_k|\mathcal{F}_{(k-1)n+\hat{n}}) \leq p_{\mathbb{1}_{\hat{e}_k}}(n).$$

The transmission halts at the end of the first epoch in which an acknowledge message a is detected at the end of the confirmation phase, i.e., the first time either a correct transmission in the communication phase is followed by the successful transmission of an acknowledge message in the confirmation phase, or an incorrect transmission in the communication phase is followed by an undetected transmission of a deny message b in the confirmation phase. It follows that we can rewrite $\zeta^{(n)}$ as

$$\zeta^{(n)} = \inf\left\{ k \in \mathbb{N} \text{ s.t. } (\hat{e}_k \cap \tilde{e}_k) \cup \overline{(\hat{e}_k \cup \tilde{e}_k)} \right\}.$$

We claim that

$$\mathbb{P}(\zeta^{(n)} \geq k) \leq (p(n) + p_a(n))^{k-1}. \tag{70}$$

Indeed, (70) can be shown by induction. It is clearly true for $k = 1$. Suppose it is true for some k in $\mathbb{N}$; then

$$\begin{aligned}
\mathbb{P}\left(\zeta^{(n)} > k\right) &= \mathbb{P}\left(\zeta^{(n)} > k|\zeta^{(n)} \geq k\right)\mathbb{P}\left(\zeta^{(n)} \geq k\right) \\
&= \mathbb{P}(\zeta^{(n)} \geq k)\mathbb{P}\left(\hat{e}_{k+1} \cap \overline{\tilde{e}_{k+1}}|\zeta^{(n)} \geq k\right) \\
&\quad + \mathbb{P}(\zeta^{(n)} \geq k)\mathbb{P}\left(\tilde{e}_{k+1} \cap \overline{\hat{e}_{k+1}}|\zeta^{(n)} \geq k\right) \\
&\leq (p(n) + p_a(n))\mathbb{P}(\zeta^{(n)} \geq k) \\
&\leq (p(n) + p_a(n))^k.
\end{aligned}$$

Thus, $\zeta^{(n)}$ is stochastically dominated by the sum of a constant 1 plus an r.v. with geometric distribution of parameter $p(n) + p_a(n)$. It follows that its expected value can be bounded

$$\begin{aligned}
1 &\leq \mathbb{E}[\zeta^{(n)}] \\
&= \sum_{t \geq 1} \mathbb{P}(\zeta^{(n)} \geq t) \\
&\leq \sum_{t \geq 1} (p(n) + p_a(n))^{t-1} \\
&= \frac{1}{1 - p(n) - p_a(n)}.
\end{aligned}$$

Hence, from (63) and (65) we have

$$\lim_n \mathbb{E}[\zeta^{(n)}] = 1. \tag{71}$$

From (71) it immediately follows that

$$\lim_n \frac{\log |\mathcal{W}^{(n)}|}{\mathbb{E}[T^{(n)}]} = \lim_n \frac{\log\left(\exp(\lceil nR \rceil)\right)}{n\mathbb{E}[\zeta^{(n)}]} = R.$$

Moreover, transmission ends with an error if and only if an error happens in the communication phase followed by a type-b error in the confirmation phase, so that, the error probability of the overall scheme $(\Phi^{(n)}, T^{(n)}, \Psi^{(n)})$ can be bounded as follows:

$$p_e\left(\Phi^{(n)}, T^{(n)}, \Psi^{(n)}\right) = \mathbb{P}\left(e_{\zeta^{(n)}} \cap \tilde{e}_{\zeta^{(n)}}\right)$$
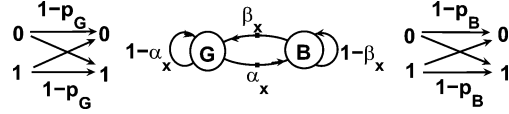

Fig. 4. an FSMC with binary state space $\mathcal{S} = \{G, B\}$ and binary input/output space $\mathcal{X} = \mathcal{Y} = \{0, 1\}$: notice that the state transition probabilities are allowed to depend on the current input (ISI).

$$\begin{aligned}
&= \sum_{t \geq 1} \mathbb{P}\left(e_t \cap \tilde{e}_t \cap \{\zeta^{(t)} = t\}\right) \\
&= \sum_{t \geq 1} \mathbb{P}\left(e_t \cap \tilde{e}_t \cap \{\zeta^{(n)} \geq t\}\right) \\
&= \sum_{t \geq 1} \mathbb{P}\left(e_t \cap \tilde{e}_t | \zeta^{(n)} \geq t\right)\mathbb{P}(\zeta^{(n)} \geq t) \\
&\leq p(n)p_b(n)\sum_{t \geq 1}\mathbb{P}(\zeta^{(n)} \geq t) \\
&\leq \frac{p(n)p_b(n)}{\kappa_n}
\end{aligned} \tag{72}$$

where $\kappa_n := 1 - p(n) - p_a(n)$. When D is infinite, (72) directly implies (69). When D is finite, from (63), (65), (71), and (72) it follows that

$$\begin{aligned}
D(1 - \gamma) &\geq \liminf_n \frac{1}{n}\kappa_n \log\frac{\kappa_n}{p_b(n)} \\
&= \liminf_n \frac{-\log p_e(\Phi^{(n)}, T^{(n)}, \Psi^{(n)})}{n\mathbb{E}[\zeta^{(n)}]} \\
&= \liminf_n \frac{-\log p_e(\Phi^n, T_n, \Psi^n)}{\mathbb{E}[T^{(n)}]}
\end{aligned}$$

which proves (68). ∎

It is clear that (23) follows from (68) and the arbitrariness of γ in $(R/C, 1)$, so that Part 2 of Theorem 2 is proved.

We end this section with the following observation. It follows from (70) that the probability that the proposed transmission scheme halts after more than one epoch is bounded by $p(n) + p_a(n)$, a term which is vanishing asymptotically with n. Then, even if the transmission time is variable, it exceeds n with probability which is asymptotically small in n. As also observed in [23] for memoryless channels, this is a desirable property from a practical viewpoint. Observe the difference with respect to the fixed-length block-coding setting, when the transmission time is required to be almost surely constant for all fixed $n \in \mathbb{N}$: in this case, as already mentioned in Section I, the error exponent with feedback is upper-bounded by the sphere-packing exponent (see [13], [31]).

VI. AN EXPLICIT EXAMPLE

We consider an FSMC as in Fig. 4, with state space $\mathcal{S} = \{G, B\}$, input and output spaces $\mathcal{X} = \mathcal{Y} = \{0, 1\}$, and stochastic kernel given by:

$$\begin{aligned}
P(s_+, y|s, x) &= P_S(s_+|x, s)P_Y(y|x, s) \\
P_S(B|G, 0) &= \alpha_0, \ P_S(B|G, 1) = \alpha_1 \\
P_S(G|B, 0) &= \beta_0, \ P_S(G|B, 1) = \beta_1 \\
P_Y(1|G, 0) &= P_Y(0|G, 1) = p_G
\end{aligned}$$

$$P_Y(1|B,0) = P_Y(0|B,1) = p_B$$

where $0 < p_G < p_B < 1/2$, and $\alpha_0, \alpha_1, \beta_0, \beta_1 \in (0,1)$. For any stationary policy $\pi : \mathcal{S} \to \mathcal{P}(\{0,1\})$, the ergodic state measure associated to π can expressed explicitly

$$\boldsymbol{\mu}_\pi(B) = \frac{\xi_\alpha}{\xi_\alpha + \xi_\beta}, \quad \boldsymbol{\mu}_\pi(G) = 1 - \boldsymbol{\mu}_\pi(B)$$

where $\xi_\alpha := \alpha_0[\pi(G)](0) + \alpha_1[\pi(G)](1)$ and $\xi_\beta := \beta_0[\pi(B)](0) + \beta_1[\pi(B)](1)$. The mutual information costs are given by

$$
\begin{aligned}
c(G,u) = {} & H\big(u(1)\alpha_1 + u(0)\alpha_0\big) \\
& + H\big(u(1)p_G + u(0)(1-p_G)\big) \\
& - H(p_G) - (u_G H(\alpha_1) + u(0)H(\alpha_0)) \\
c(B,u) = {} & H\big(u(1)\beta_1 + u(0)\beta_0\big) \\
& + H\big(u(1)p_B + u(0)(1-p_B)\big) \\
& - H(p_B) - (u(1)H(\beta_1) + u(0)H(\beta_0))
\end{aligned}
$$

H denoting the binary entropy function. The information divergence costs instead are given by

$$
\begin{aligned}
d\big(G, \delta_{f_0(G)}\big) &= D\big(p_G \| 1 - p_G\big) + D\big(\alpha_{f_0(G)} \| \alpha_{f_1(G)}\big) \\
d\big(B, \delta_{f_0(B)}\big) &= D\big(p_B \| 1 - p_B\big) + D\big(\alpha_{f_0(G)} \| \alpha_{f_1(G)}\big)
\end{aligned}
$$

where, for x, y in $[0,1]$

$$D(x\|y) := x \log \frac{x}{y} + (1-x) \log \frac{1-x}{1-y}.$$

In Figs. 5 and 6, the special case when $p_G = 0.001$, $p_B = 0.1$, $\alpha_0 = 1 - \beta_0 = 0.7$, and $\alpha_1 = 1 - \beta_1 = \gamma$ is studied as a function of the parameter γ in $(0,1)$. In particular, in Fig. 5 the capacity and the optimal policy $\pi : \mathcal{S} \to \mathcal{X}$ are plotted as a function of γ. Notice that for $\gamma = 0.7$, the channel has no ISI and actually coincides with a memoryless Gilbert–Elliot channel: for that value, the optimal policy chooses the uniform distribution both in the good state G as well as in the bad state B. For values of γ below 0.7 (resp., beyond 0.7), instead, the optimal policy puts more mass on the input symbol 1 (resp., the symbol 0) both in state G and state B, and it is more unbalanced in state B. In Fig. 6, the Burnashev coefficient of the channel is plotted as a function of the parameter γ, as well as the the values of the ergodic Kullback–Leibler cost corresponding to the four possible policies $f_0 : \{G,B\} \to \{0,1\}$. Observe as the minimum value of D is achieved for $\gamma = 0.7$; in that case, all the four nontrivial policies f_0, f_1 give the same value of the Kullbak–Leibler cost.

Finally, it is worth to consider the simple non-ISI case when $\alpha_0 = \alpha_1 = \beta_0 = \beta_1$. In this case, the state ergodic measure is the uniform one on $\{G,B\}$. Notice that by a basic convexity argument we get that its capacity C and Burnashev coefficient D satisfy

$$
\begin{aligned}
C &= 1 - \frac{1}{2}H(p_G) - \frac{1}{2}H(p_B) \\
&> 1 - H\left(\frac{1}{2}p_G + \frac{1}{2}p_B\right) =: \tilde{C} \quad (73) \\
D &= \frac{1}{2}D\left(p_G \| 1 - p_G\right) + \frac{1}{2}D(p_B \| 1 - p_B) \\
&> D\left(\frac{1}{2}p_G + \frac{1}{2}p_B \| 1 - \frac{1}{2}p_B - \frac{1}{2}p_G\right) =: \tilde{D}. \quad (74)
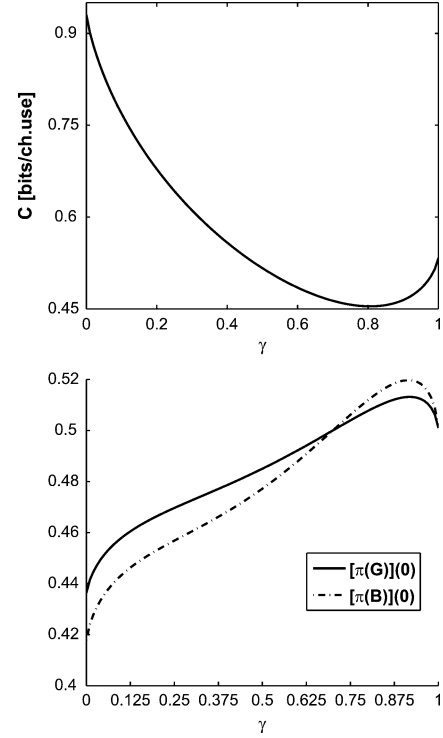\end{aligned}
$$

Fig. 5. In the top graph, the capacity of the FSMC of Fig. 4 for values of the parameters $p_G = 0.001$, $p_B = 0.1$, $\alpha_0 = 1 - \beta_0 = 0.7$, $\alpha_1 = 1 - \beta_1 = \gamma$, is plotted as a function of γ in $(0,1)$. In the bottom graph, for the same values of the parameters, the optimal policy $\boldsymbol{\pi}^* : \{G,B\} \to \mathcal{P}(\{0,1\})$ is plotted as a function of γ in $(0,1)$.
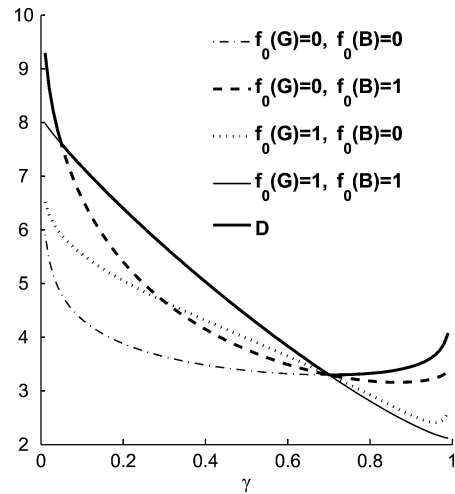

Fig. 6. The thick solid line is a plot of the Burnashev coefficient D (evaluated with natural log base) of the FSMC of Fig. 4 for the same values of the parameters as in Fig. 5.

In (73) and (74), $\tilde{C}$ and $\tilde{D}$ correspond, respectively, to the capacity and the Burnashev coefficient of memoryless binary symmetric channel with crossover probability equal to the ergodic average of the crossover probabilities p_B and p_G. Such a channel is introduced in practice when channel interleavers are used in order to apply to FSMCs coding techniques designed for DMCs. While this approach reduces the decoding complexity, it is well known that it reduces the achievable capacity (73)

(see [16]). Inequality (74) shows that this approach causes also a significant loss in the Burnashev coefficient of the channel.

VII. CONCLUSION

In this paper, we studied the error exponent of FSMCs with feedback. We have proved an exact single-letter characterization of the reliability function for variable-length block-coding schemes with perfect causal output feedback, generalizing the result obtained by Burnashev [6] for memoryless channels. Our assumptions are that the channel state is causally observable both at the encoder and the decoder and the stochatic kernel describing the channel satisfies some mild ergodicity properties.

As a first topic for future research, we would like to extend our result to the case when the state is either observable at the encoder only or it is not observable at either side. We believe that the techniques used in [32] in order to characterize the capacity of FSMCs with state not observable may be adopted to handle our problem as well. The main idea consists in studying a partially observable Markov decision process and reduce it to a fully observable one with a larger state space. However, an extension of the results of in Section IV is needed, as there we explicitly exploited the finiteness of the state space in our proofs. Finally, it would be interesting to consider the problem of finding universal schemes which do not require exact knowledge of the channel statistics but use feedback in order to estimate them.

APPENDIX

For the reader's convenience, all statements are repeated before their proof.

For $t \geq 1$ we will use the notation $\boldsymbol{\nu}_t \in \mathcal{P}(\mathcal{S} \times \mathcal{Y})$, $\boldsymbol{\nu}_t(s, y) := \mathbb{P}\left(S_{t+1} = s, Y_t = y | \mathcal{E}_t\right)$ for the conditioned probability distribution of the pair (S_{t+1}, Y_t) given the feedback observation $\mathcal{E}_t = \mathcal{G}_{t-1}$. Since, due to the assumption (17) on the causality of the channel and of the encoder, $(W, \boldsymbol{S}_1^t, \boldsymbol{Y}_1^{t-1})$ and (S_{t+1}, Y_t) are conditionally independent given (S_t, X_t), for all $w \in \mathcal{W}$ an application of the Bayes rule gives us

$$\boldsymbol{\theta}_t(w)\boldsymbol{\nu}_t(S_{t+1}, Y_t)$$
$$= \boldsymbol{\theta}_{t-1}(w)P(S_{t+1}, Y_t | S_t, \phi_t(w, \boldsymbol{S}_1^t, \boldsymbol{Y}_1^{t-1})). \quad (75)$$

Lemma 1: Given any causal feedback encoder Φ, for every $t \geq 1$

$$\Pi_t \geq \lambda \Pi_{t-1} \quad \text{a.s.}$$

Proof: From (17) it follows that, for channel state/output pair (S_{t+1}, Y_t) to be observed with nonzero probability after the state S_t, it is necessary that $(Y_t, S_{t+1}) \in \mathcal{Z}_{S_t}$, where the $\mathcal{Z}_{S_t}$ is the set of channel state and output pairs which are reachable from the state S_t—see (10). It follows that, almost surely, for all time t and for any message w in $\mathcal{W}$

$$\lambda \leq \lambda_{S_t}$$
$$= \min\left\{\min_x P\left(s_+, y | S_t, x\right) | (s_+, y) \in \mathcal{Z}_{S_t}\right\}$$

$$\leq \min_x P\left(S_{t+1}, Y_t | S_t, x\right)$$
$$\leq P\left(S_{t+1}, Y_t | S_t, \phi_t(w, \boldsymbol{S}_1^t, \boldsymbol{Y}_1^{t-1})\right).$$

Since $\boldsymbol{\nu}_t(S_{t+1}, Y_t) \leq 1$, using (75) and the inequality above, we have

$$\boldsymbol{\theta}_t(w) \geq \boldsymbol{\nu}_t(S_{t+1}, Y_t)\boldsymbol{\theta}_t(w)$$
$$= P\left(S_{t+1}, Y_t | S_t, \phi_t(w, \boldsymbol{S}_1^t, \boldsymbol{Y}_1^{t-1})\right)\boldsymbol{\theta}_{t-1}(w)$$
$$\geq \lambda\boldsymbol{\theta}_{t-1}(w).$$

Let $\check{W} := \check{\psi}_t(\boldsymbol{S}_1^{t+1}, \boldsymbol{Y}_1^t)$. It follows that

$$\Pi_t = \mathbb{P}\left(\check{W} \neq W | \mathcal{G}_t\right)$$
$$= \sum_{w \in \mathcal{W} \setminus \{\check{W}\}} \boldsymbol{\theta}_t(w)$$
$$\geq \sum_{w \in \mathcal{W} \setminus \{\check{W}\}} \lambda\boldsymbol{\theta}_{t-1}(w)$$
$$\geq \lambda\Pi_{t-1},$$

thus showing the claim. ∎

Lemma 2: For any variable-length block-coding scheme (Φ, T, Ψ) and any $0 < \varepsilon < 1/2$, we have

$$C_\varepsilon(\Phi, T) \geq \left(1 - \varepsilon - \frac{1}{\varepsilon}p_e\left(\Phi, T, \Psi\right)\right)\log|\mathcal{W}| - H(\varepsilon).$$

Proof: We introduce the r.v.'s

$$V_n := \Gamma_n + \sum_{1 \leq t \leq n} c\left(S_t, \boldsymbol{\Upsilon}_t\right), \quad n \geq 0.$$

First, we prove that $(V_n, \mathcal{G}_n)_{n \geq 0}$ is a martingale. Indeed, V_n is $\mathcal{G}_n$-measurable, since Γ_n is, and so do both S_t and $\boldsymbol{\Upsilon}_t$ for every $1 \leq t \leq n$. Using (75), it follows that

$$c\left(S_n, \boldsymbol{\Upsilon}_n\right) = \sum_{x,y,s_+} \boldsymbol{\Upsilon}_n(x)P(s_+, y | s, x)$$
$$\times \log\frac{P(s_+, y | s, x)}{\sum_z \boldsymbol{v}_n(z)P(s_+, y | s, z)}$$
$$= \mathbb{E}\left[\log\frac{P\left(S_{n+1}, Y_n | S_n, X_n\right)}{\boldsymbol{\nu}_n\left(S_{n+1}, Y_n\right)} | \mathcal{G}_{n-1}\right]$$
$$= \mathbb{E}\left[\log\frac{\boldsymbol{\theta}_n(W)}{\boldsymbol{\theta}_{n-1}(W)} | \mathcal{G}_{n-1}\right]$$
$$= \mathbb{E}\left[\Gamma_{n-1} - \Gamma_n | \mathcal{G}_{n-1}\right].$$

Hence

$$\mathbb{E}\left[V_n - V_{n-1} | \mathcal{G}_{n-1}\right] = \mathbb{E}\left[\Gamma_n - \Gamma_{n-1} + c\left(S_n, \boldsymbol{\Upsilon}_n\right) | \mathcal{G}_{n-1}\right] = 0.$$

Second, we observe that (V_n) has uniformly bounded increments since

$$|V_n - V_{n-1}| \leq |c\left(S_n, \boldsymbol{\Upsilon}_n\right)| + |\Gamma_n - \Gamma_{n-1}|$$
$$\leq \log|\mathcal{X}| + 2\log|\mathcal{W}|.$$

$$\sum_{t=1}^{n+1} Z_t = \log \frac{\boldsymbol{\nu}_{0,n}^0(\boldsymbol{S}_1^{n+1}, \boldsymbol{Y}_1^n)}{\boldsymbol{\nu}_{0,n}^1(\boldsymbol{S}_1^{n+1}, \boldsymbol{Y}_1^n)} + Z_{n+1}$$

$$= \log \frac{\boldsymbol{\nu}_{0,n}^0(\boldsymbol{S}_1^{n+1}, \boldsymbol{Y}_1^n) \sum_x P(S_{n+2}, Y_{n+1}|S_{n+1}, x)\boldsymbol{v}_{n+1}^0(x)}{\boldsymbol{\nu}_{0,n}^1(\boldsymbol{S}_1^{n+1}, \boldsymbol{Y}_1^n) \sum_x P(S_{n+2}, Y_{n+1}|S_{n+1}, x)\boldsymbol{v}_{n+1}^1(x)}$$

$$= \log \frac{\boldsymbol{\nu}_{0,n+1}^0(\boldsymbol{S}_1^{n+2}, \boldsymbol{Y}_1^{n+1})}{\boldsymbol{\nu}_{0,n+1}^1(\boldsymbol{S}_1^{n+2}, \boldsymbol{Y}_1^{n+1})}$$

Therefore, we can apply Doob's optional sampling theorem [35, Theorem 10.10, p. 100], concluding that

$$\log|\mathcal{W}| = \mathbb{E}\left[V_0|\mathcal{G}_0\right]$$
$$= \mathbb{E}\left[V_{\tau_\varepsilon}\right]$$
$$= \mathbb{E}\left[\Gamma_{\tau_\varepsilon}\right] + \mathbb{E}\left[\sum_{1\le t\le \tau_\varepsilon} c(S_t, \boldsymbol{\Upsilon}_t)\right]. \qquad (76)$$

Finally, combining (76) with (28), we obtain

$$C_\varepsilon(\Phi, T) = \mathbb{E}\left[\sum_{1\le t\le \tau_\varepsilon} c(S_t, \boldsymbol{\Upsilon}_t)\right]$$
$$\ge \left(1 - \varepsilon - \frac{p_e(\Phi, T, \Psi)}{\varepsilon}\right)\log|\mathcal{W}| - H(\varepsilon)$$

which completes the proof.   ∎

Lemma 3: Let τ and T be stopping times for the filtration $\mathcal{G}$ such that $\tau \le T$, and consider a nontrivial binary partition of the message set as in (32). Then, for $i = 0, 1$

$$L_i \le \mathbb{E}\left[\sum_{\tau < t\le T} d\left(S_t, \boldsymbol{\Upsilon}_t^i\right)\Big| W \in \mathcal{W}_i, \mathcal{G}_\tau\right]. \qquad (77)$$

Proof: We will prove the claim for $i = 0$. Define [11]

$$Z_t := \log \frac{\sum_x P(S_{t+1}, Y_t|S_t, x)\boldsymbol{\Upsilon}_t^0(x)}{\sum_x P(S_{t+1}, Y_t|S_t, x)\boldsymbol{\Upsilon}_t^1(x)}, \quad t \ge 0.$$

With probability one, the pair (S_{t+1}, Y_t) belongs to the achievable set $\mathcal{Z}_{S_t}$, so that, for $i = 0, 1$

$$\lambda \le \lambda_{S_t}$$
$$\le \min_x P(S_{t+1}, Y_t|S_t, x)$$
$$\le \sum_x P(S_{t+1}, Y_t|S_t, x)\boldsymbol{\Upsilon}_t^i(x) \le 1.$$

As a consequence we have

$$|Z_t| \le -\log\lambda \qquad (78)$$

Now, for $s \in \mathcal{S}$, $y \in \mathcal{Y}$, and $i = 0, 1$, define the r.v.

$$\Delta_{s,y}^i := \sum_{x\in\mathcal{X}} \boldsymbol{\Upsilon}_t^0(x)P(s, y|S_t, x).$$

[11]With the convention $\log 0/0 := 0$.

Then, by recalling the definition (9) of the cost d, applying the log–sum inequality [8, p. 29], we have, for $t \ge 1$

$$d(S_t, \boldsymbol{\Upsilon}_t^0) \ge \sum_{y,s,x} \boldsymbol{\Upsilon}_t^0(x)P(s, y|S_t, x)$$
$$\times \log \frac{P(s, y|S_t, x)\boldsymbol{\Upsilon}_t^0(x)}{P(s, y|S_t, x)\boldsymbol{\Upsilon}_t^1(x)}$$
$$\ge \sum_{y,s} \Delta_{s,y}^0 \log \frac{\Delta_{s,y}^0}{\Delta_{s,y}^1}$$
$$= \mathbb{E}\left[\log \frac{\Delta_{s,y}^0}{\Delta_{s,y}^1}\Big| W \in \mathcal{W}_0, \mathcal{G}_{t-1}\right]$$
$$= \mathbb{E}[Z_t|W \in \mathcal{W}_0, \mathcal{G}_{t-1}]. \qquad (79)$$

From (79), it follows that, if we define

$$V_n := \sum_{1\le t\le n} Z_t - \sum_{1\le t\le n} d(S_t, \boldsymbol{\Upsilon}_t^0), \quad n \ge 0$$

then $(V_n, \mathcal{G}_n)_{n\ge 0}$ is a submartingale with respect to the conditioned probability measure $\mathbb{P}(\cdot|W \in \mathcal{W}_0)$. Moreover, it follows from (78) (recall that we are assuming $\lambda > 0$ and that this is equivalent to the boundedness of the Burnashev coefficient D) that (V_n) has uniformly bounded increments

$$|V_{n+1} - V_n| \le |Z_{n+1}| + |d(S_{n+1}, \boldsymbol{\Upsilon}_{\Phi, n+1}^0)|$$
$$\le \log \frac{1}{\lambda} + d_{\max}.$$

Thus, since $\tau \le T$, Doob's optional stopping theorem [35, Theorem 10.10] can be applied yielding

$$\mathbb{E}\left[V_T - V_\tau|W \in \mathcal{W}_0, \mathcal{G}_\tau\right] \le 0. \qquad (80)$$

Then the claim follows from (80), upon showing that for every $n \ge 0$

$$\log \frac{\boldsymbol{\nu}_{0,n}^0(\boldsymbol{S}_1^{n+1}, \boldsymbol{Y}_1^n)}{\boldsymbol{\nu}_{0,n}^1(\boldsymbol{S}_1^{n+1}, \boldsymbol{Y}_1^n)} = \sum_{1\le t\le n} Z_t. \qquad (81)$$

In fact, (81) can be verified by induction. It holds true for $n = 0$, since [12] S_1 is independent from W and so $\boldsymbol{\nu}_{0,0}^0(S_t) = \boldsymbol{\nu}_{0,0}^1(S_t)$. Moreover, assume that (81) holds true for some n. Then, we have $\sum_{t=1}^{n+1} Z_t$ defined in the equation at the top of the page, thus proving (81).   ∎

[12]We use the convention for an empty summation to equal zero.

Lemma 4: Let Φ be any causal encoder, and τ and T be stopping times for the filtration $\mathcal{G}$ such that $\tau \leq T$. Then, for every $2^{\mathcal{W}}$-valued $\mathcal{G}_\tau$-measurable r.v. $\mathcal{W}_1$, we have

$$\mathbb{E}\left[\sum_{\tau < t \leq T} d\left(S_t, \boldsymbol{\Upsilon}_{\Phi,t}^{\mathbb{1}_{\{W \in \mathcal{W}_1\}}}\right) \Big| \mathcal{G}_\tau\right] \geq \log \frac{Z}{4} - \log p_\tau$$

where $Z := \min\{\boldsymbol{\theta}_\tau(\mathcal{W}_0), \boldsymbol{\theta}_\tau(\mathcal{W}_1)\}$.

Proof: First we will prove the statement when $\mathcal{W}_1$ is a fixed, nontrivial subset of the message set $\mathcal{W}$. For $i, j \in \{0, 1\}$, define $p_{i|j} := \mathbb{P}(\tilde{W} = i | W \in \mathcal{W}_j, \mathcal{G}_\tau)$. We shall now upper-bound L_0 defined in (33). From the log–sum inequality it follows that

$$
\begin{aligned}
L_0 &= \mathbb{E}\left[\log \frac{\boldsymbol{\nu}_{\tau,T}^0(\boldsymbol{S}_{\tau+2}^{T+1}, \boldsymbol{Y}_{\tau+1}^T)}{\boldsymbol{\nu}_{\tau,T}^1(\boldsymbol{S}_{\tau+2}^{T+1}, \boldsymbol{Y}_{\tau+1}^T)} \Big| W \in \mathcal{W}_0, \mathcal{G}_\tau\right] \\
&\geq p_{0|0} \log \frac{p_{0|0}}{p_{0|1}} + p_{1|0} \log \frac{p_{1|0}}{p_{1|1}} \\
&\geq -H(p_{1|0}) - p_{0|0} \log p_{0|1} \\
&\geq -\log 2 - p_{0|0} \log p_{0|1}.
\end{aligned}
$$

We have $p_\tau = \boldsymbol{\theta}_\tau(0) p_{1|0} + \boldsymbol{\theta}_\tau(1) p_{0|1} \geq Z p_{0|1}$. From Lemma 3 it follows that

$$
\begin{aligned}
\mathbb{E}\Bigg[\sum_{\tau < t \leq T} d\left(S_t, \boldsymbol{\Upsilon}_t^0\right) &| W \in \mathcal{W}_0, \mathcal{G}_\tau\Bigg] \\
&\geq L_0 \\
&\geq -\log 2 - p_{0|0} \log p_{0|1} \\
&\geq -\log 2 - p_{0|0} \log \frac{p_\tau}{Z}.
\end{aligned}
\tag{82}
$$

An analogous derivation leads to

$$
\begin{aligned}
\mathbb{E}\Bigg[\sum_{\tau < t \leq T} d\left(S_t, \boldsymbol{\Upsilon}_t^1\right) &| W \in \mathcal{W}_1, \mathcal{G}_\tau\Bigg] \\
&\geq -\log 2 - p_{1|1} \log \left(\frac{p_\tau}{Z}\right).
\end{aligned}
\tag{83}
$$

By averaging (82) and (83) with respect to the posterior distribution $\boldsymbol{\theta}_\tau$ of W given $\mathcal{G}_\tau$, we get

$$\mathbb{E}\left[\sum_{\tau < t \leq T} d\left(S_t, \boldsymbol{\Upsilon}_t^{\mathbb{1}_{\{W \in \mathcal{W}_1\}}}\right) | \mathcal{G}_\tau\right] \geq \log \frac{Z}{2} - (1 - p_\tau) \log p_\tau$$

and the claim follows upon observing that

$$p_\tau \log p_\tau \geq -H(p_\tau) \geq \log \frac{1}{2}. \qquad \blacksquare$$

Lemma 5: Let Φ be a causal feedback encoder and T a transmission time. Then, for every $0 < \varepsilon < 1/2$ there exists a $\mathcal{G}_{\tau_\varepsilon}$-measurable random subset $\mathcal{W}_1$ of the message set $\mathcal{W}$, whose *a posteriori* error probability satisfies

$$1 - \lambda\varepsilon \geq \boldsymbol{\theta}_{\tau_\varepsilon}(\mathcal{W}_1) \geq \lambda\varepsilon.$$

Proof: Suppose first that $\Pi_{\tau_\varepsilon} \leq \varepsilon$. Then, since clearly $\Pi_{\tau_\varepsilon - 1} \geq \varepsilon$, by Lemma 1 we have

$$\Pi_{\tau_\varepsilon} \geq \lambda\Pi_{\tau_\varepsilon - 1} \geq \lambda\varepsilon.$$

It follows that, if we define $\mathcal{W}_1 := \{\Psi_{\tau_\varepsilon}(\boldsymbol{S}_1^{\tau_\varepsilon + 1}, \boldsymbol{Y}_1^{\tau_\varepsilon})\}$ and $\mathcal{W}_0 := \mathcal{W} \setminus \mathcal{W}_1$, we have

$$
\begin{aligned}
\boldsymbol{\theta}_{\tau_\varepsilon}(\mathcal{W}_1) &= 1 - \Pi_{\tau_\varepsilon} \geq 1 - \varepsilon \geq \lambda\varepsilon \\
\boldsymbol{\theta}_{\tau_\varepsilon}(\mathcal{W}_0) &= \Pi_{\tau_\varepsilon} \geq \lambda\boldsymbol{F}\varepsilon\boldsymbol{F}.
\end{aligned}
$$

If instead $\Pi_{\tau_\varepsilon} > \varepsilon$, the *a posteriori* probability of any message w in $\mathcal{W}$ at time τ_ε satisfies $\boldsymbol{\theta}_{\tau_\varepsilon}(w) \leq 1 - \varepsilon$. Then it is possible to construct $\mathcal{W}_1$ in the following way. Introduce an arbitrary labeling of $\mathcal{W} = \{w_1, w_2, \ldots, w_{|\mathcal{W}|}\}$. For any $1 \leq i \leq |\mathcal{W}|$, define $\mathcal{W}_{(i)} = \{w_1, \ldots, w_i\}$. Set $k := \inf\{1 \leq i \leq |\mathcal{W}| | \boldsymbol{F} : \boldsymbol{F}\boldsymbol{\theta}_{\tau_\varepsilon}(\mathcal{W}_{(i)}) \geq \lambda\boldsymbol{F}\varepsilon\}$, and define $\mathcal{W}_1 = \mathcal{W}_{(k)}, \mathcal{W}_0 := \mathcal{W} \setminus \mathcal{W}_1$. Then, clearly $\boldsymbol{\theta}_{\tau_\varepsilon}(\mathcal{W}_1) \geq \lambda\boldsymbol{F}\varepsilon$, while

$$
\begin{aligned}
\boldsymbol{\theta}_{\tau_\varepsilon}(\mathcal{W}_0) &= 1 - \boldsymbol{\theta}_{\tau_\varepsilon}(\mathcal{W}_{(k)}) \\
&= 1 - \boldsymbol{\theta}_{\tau_\varepsilon}(\mathcal{W}_{(k-1)}) - \boldsymbol{\theta}_{\tau_\varepsilon}(w_k) \\
&\geq 1 - \lambda\boldsymbol{F}\varepsilon - (1 - \varepsilon) \\
&\geq \lambda\varepsilon\boldsymbol{F}. \qquad \blacksquare
\end{aligned}
$$

Lemma 6: For every $\varepsilon > 0$, and for every feasible policy $\boldsymbol{\pi}$

$$\mathbb{P}_{\boldsymbol{\pi}}\left(\|\boldsymbol{F}(\boldsymbol{v}_n)\| \geq \varepsilon + \frac{1}{n}\right) \leq 2|\mathcal{S}| \exp\left(\frac{-n\varepsilon^2}{2}\right)\boldsymbol{F}.$$

Proof: Let us fix an arbitrary admissible policy $\boldsymbol{\pi}$. For every s in $\mathcal{S}$ consider the following random process

$$
\begin{aligned}
Z_0^s &:= 0, \quad Z_1^s := 0, \\
Z_n^s &:= (n-1)F_s(\boldsymbol{v}_{n-1}) + \mathbb{1}_{\{S_n = s\}} - \mathbb{1}_{\{S_1 = s\}}, \quad n \geq 2.
\end{aligned}
$$

We have

$$
\begin{aligned}
Z_n^s &= (n-1)\boldsymbol{v}_{n-1}(\{s\}, \mathcal{U}) + \mathbb{1}_{\{S_n = s\}} - \mathbb{1}_{\{S_1 = s\}} \\
&\quad - (n-1)\int_{\mathcal{S} \times \mathcal{U}} Q_S(s|j, u)\mathrm{d}\boldsymbol{v}_{n-1}(j, u) \\
&= \sum_{t=2}^{n} \mathbb{1}_{\{S_t = s\}} - \sum_{t=2}^{n} Q(s|S_{t-1}, U_{t-1}) \\
&= \sum_{t=2}^{n} \left(\mathbb{1}_{\{S_t = s\}} - \mathbb{E}_{\boldsymbol{\pi}}\left[\mathbb{1}_{\{S_t = s\}} | \mathcal{E}_{t-1}\right]\right).
\end{aligned}
$$

It is immediate to check that Z_n^s is $\mathcal{E}_n$-measurable. Moreover

$$\mathbb{E}_{\boldsymbol{\pi}}[Z_{n+1}^s | \mathcal{E}_n] = Z_n^s, \quad \forall n \geq 0$$

so that $(Z_n^s, \mathcal{E}_n, \mathbb{P}_{\boldsymbol{\pi}})_{n \geq 0}$ is a martingale. Moreover, (Z_n^s) has uniformly bounded increments since $|Z_1^s - Z_0^s| = 0$, while

$$
\begin{aligned}
|Z_{n+1}^s - Z_n^s| &= \left|\mathbb{1}_{\{S_{n+1} = s\}} - \mathbb{E}_{\boldsymbol{\pi}}\left[\mathbb{1}_{\{S_{n+1} = s\}} | \mathcal{E}_n\right]\right| \\
&\leq a_{n+1},
\end{aligned}
$$

where $a_n = 1$ for $n \geq 2$. It follows that we can apply Hoeffding–Azuma inequality [22], obtaining

$$
\begin{aligned}
\mathbb{P}_{\boldsymbol{\pi}}\left(|Z_{n+1}^s| \geq \varepsilon n\right) &\leq 2\exp\left(\frac{-\varepsilon^2 n^2}{2\sum_{k=2}^{n+1} a_k}\right) \\
&= 2\exp\left(\frac{-\varepsilon^2}{2}n\right).
\end{aligned}
$$

By simply applying a union bound, we can argue that

$$
\mathbb{P}_{\boldsymbol{\pi}}\left(\|\boldsymbol{F}(\boldsymbol{v}_n)\| \geq \varepsilon + \frac{1}{n}\right)
$$
$$
= \mathbb{P}_{\boldsymbol{\pi}}\left(\max_s |Z_{n+1}^s + \mathbb{1}_{\{S_1=s\}} - \mathbb{1}_{\{S_{n+1}=s\}}| \geq \varepsilon n + 1\right)
$$
$$
\leq \mathbb{P}_{\boldsymbol{\pi}}\left(\bigcup_{s \in \mathcal{S}} \{|Z_{n+1}^s| \geq \varepsilon n\}\right)
$$
$$
\leq \sum_{s \in \mathcal{S}} \mathbb{P}_{\boldsymbol{\pi}}\left(|Z_{n+1}^s| \geq \varepsilon n\right)
$$
$$
\leq 2|\mathcal{S}| \exp\left(-\frac{\varepsilon^2}{2} n\right),
$$

which concludes the proof. ∎

Lemma 7: The map γ is upper semicontinuous (i.e., $\varepsilon_n \to \varepsilon \Rightarrow \limsup_n \gamma(\varepsilon_n) \leq \gamma(\varepsilon)$).

Proof: As γ is nondecreasing, with no loss of generality we can restrict ourselves to consider the case when $\varepsilon_n \downarrow \varepsilon$, so that $\lim_n \gamma(\varepsilon_n)$ exists. Since $\mathcal{S} \times \mathcal{U}$ is compact, the Prohorov space $\mathcal{P}(\mathcal{S} \times \mathcal{U})$ is compact as well [3]. Thus, since the map $\boldsymbol{\eta} \mapsto \|\boldsymbol{F}(\boldsymbol{\eta})\|$ is continuous, the sublevel $\{\|\boldsymbol{F}(\boldsymbol{\eta})\| \leq \varepsilon\}$ is compact. It follows that for every n there exists $\boldsymbol{\eta}_n$ in $\mathcal{P}(\mathcal{S} \times \mathcal{U})$ such that $\|F(\boldsymbol{\eta}_n)\| \leq \varepsilon_n$ and

$$
\gamma(\varepsilon_n) = \sup\left\{\langle \boldsymbol{\eta}, g \rangle \big| \boldsymbol{\eta} : \|\boldsymbol{F}(\boldsymbol{\eta})\| \leq \varepsilon_n\right\} = \langle \boldsymbol{\eta}_n, g \rangle.
$$

Since $\mathcal{P}(\mathcal{S} \times \mathcal{U})$ is compact, we can extract a converging subsequence (η_{n_k}); define $\overline{\boldsymbol{\eta}} := \lim_k \boldsymbol{\eta}_{n_k}$. Clearly

$$
\|\boldsymbol{F}(\overline{\boldsymbol{\eta}})\| = \lim_k \|\boldsymbol{F}(\boldsymbol{\eta}_{n_k})\| \leq \lim_k \varepsilon_{n_k} = \varepsilon.
$$

It follows that

$$
\gamma(\varepsilon) = \sup\left\{\langle \boldsymbol{\eta}, g \rangle \big| \boldsymbol{\eta} \in \mathcal{P}(\mathcal{S} \times \mathcal{U}) : \|\boldsymbol{F}(\boldsymbol{\eta})\| \leq \varepsilon\right\}
$$
$$
\geq \langle \overline{\boldsymbol{\eta}}, g \rangle
$$
$$
= \lim_k \langle \boldsymbol{\eta}_{n_k}, g \rangle
$$
$$
= \lim_n \gamma(\varepsilon_n)
$$

thus proving the claim. ∎

Lemma 8: Let $(\tau^{(n)})$ be a sequence of stopping times for the filtration $\mathcal{F}$ and $(\boldsymbol{\pi}^{(n)})$ be a sequence of feasible policies such that $\mathbb{E}_{(n)}[\tau^{(n)}] < \infty$ for every n and (52) holds true. Then

$$
\lim_n \mathbb{P}_{(n)}\left(G_{\tau^{(n)}} > \gamma(\varepsilon)\right) = 0, \quad \forall \varepsilon > 0.
$$

Proof: For every $\varepsilon > 0$, using a union bound estimation and (51) we get

$$
\mathbb{P}_{(n)}\left(G_t > \gamma\left(\varepsilon + \frac{1}{t}\right)\right)
$$
$$
= \mathbb{P}_{(n)}\left(\bigcup_{s \geq t}\left\{\langle \boldsymbol{v}_s, g \rangle > \gamma\left(\varepsilon + \frac{1}{t}\right)\right\}\right)
$$
$$
\leq \sum_{s \geq t} \mathbb{P}_{(n)}\left(\langle \boldsymbol{v}_s, g \rangle > \gamma\left(\varepsilon + \frac{1}{t}\right)\right)
$$
$$
\leq \sum_{s \geq t} \mathbb{P}_{(n)}\left(\langle \boldsymbol{v}_s, g \rangle > \gamma\left(\varepsilon + \frac{1}{s}\right)\right)
$$
$$
\leq 2|\mathcal{S}| \sum_{s \geq t} \exp\left(\frac{-s\varepsilon^2}{2}\right)
$$
$$
= 2|\mathcal{S}| \frac{\exp\left(\frac{-t\varepsilon^2}{2}\right)}{1 - \exp\left(\frac{-\varepsilon^2}{2}\right)}. \tag{84}
$$

It follows that for every M in $\mathbb{N}$ we have for the events

$$
A_n := \left\{G_{\tau^{(n)}} > \gamma\left(\varepsilon + \frac{1}{M}\right)\right\},
$$
$$
B_n := \left\{\tau^{(n)} \geq M\right\}
$$
$$
\mathbb{P}_{(n)}(A_n) = \mathbb{P}_{(n)}(A_n \cap B_n) + \mathbb{P}_{(n)}\left(A_n \cap \overline{B_n}\right)
$$
$$
\leq \sum_{s \geq M} \mathbb{P}_{(n)}\left(A_n \cap \{\tau^{(n)} = s\}\right) + \mathbb{P}_{(n)}(\overline{B_n})
$$
$$
\leq \sum_{s \geq M} \mathbb{P}_{(n)}\left(G_s > \gamma\left(\varepsilon + \frac{1}{M}\right)\right) + \mathbb{P}_{(n)}(\overline{B_n})
$$
$$
\leq \sum_{s \geq M} 2|\mathcal{S}| \frac{\exp\left(\frac{-s\varepsilon^2}{2}\right)}{1 - \exp\left(\frac{-\varepsilon^2}{2}\right)} + \mathbb{P}_{(n)}(\overline{B_n})
$$
$$
= \frac{2|\mathcal{S}| \exp\left(\frac{-M\varepsilon^2}{2}\right)}{\left(1 - \exp\left(\frac{-\varepsilon^2}{2}\right)\right)^2} + \mathbb{P}_{(n)}(\overline{B_n}).
$$

Thus, it follows from (52) that

$$
\limsup_n \mathbb{P}_{(n)}(A_n) \leq \frac{2|\mathcal{S}| \exp\left(\frac{-M\varepsilon^2}{2}\right)}{\left(1 - \exp\left(\frac{-\varepsilon^2}{4|\mathcal{S}|^2}\right)\right)^2}
$$

and by the arbitrariness of M in $\mathbb{N}$ we get the claim. ∎

Lemma 9: Let $(\Phi^{(n)}, T^{(n)}, \Psi^{(n)})$ be a sequence of variable-length block-coding schemes satisfying (21). For every $n \in \mathbb{N}$, define $\tau^{(n)}$ as in (57). Then

$$
\lim_n \mathbb{P}\left(\tau^{(n)} \leq M\right) = 0, \quad \forall M \in \mathbb{N}.
$$

Moreover, for any choice of a $2^{\mathcal{W}}$-valued $\mathcal{G}_{\tau^{(n)}}$-measurable r.v. $\mathcal{W}^{(n)}$ such that

$$
\lambda \varepsilon_n \leq \mathbb{P}(W \in \mathcal{W}^{(n)}) \leq 1 - \lambda \varepsilon_n
$$

we have

$$
\lim_n \mathbb{P}\left(T^{(n)} - \tau^{(n)} \leq M | W \in \mathcal{W}^{(n)}\right) = 0, \quad \forall M \in \mathbb{N}.
$$

Proof: From Lemma 1, we have that $\mathbb{P}$-a.s.

$$
\Pi_{T^{(n)}}^{(n)} \geq \Pi_{\tau^{(n)}}^{(n)} \lambda^{T^{(n)} - \tau^{(n)}} \geq \lambda \varepsilon_n \lambda^{T^{(n)} - \tau^{(n)}}.
$$

For $M, n \in \mathbb{N}$, define the events $B_n := \{T^{(n)} - \tau^{(n)} \leq M\}$, $F_n := \{W \in \mathcal{W}^{(n)}\}$,

$$
p_e\left(\Phi^{(n)}, T^{(n)}, \Psi^{(n)}\right) \geq \mathbb{E}\left[\Pi_{T^{(n)}}^{(n)} | B_n\right] \mathbb{P}(B_n)
$$

$$\geq \lambda \varepsilon_n \lambda^M \mathbb{P}(B_n)$$
$$\geq \lambda \varepsilon_n \lambda^M \mathbb{P}(B_n \cap F_n)$$
$$= \lambda \varepsilon_n \lambda^M \mathbb{P}(B_n|F_n) \mathbb{P}(F_n)$$
$$\geq \lambda^2 \varepsilon_n^2 \lambda^M \mathbb{P}(B_n|F_n).$$

It follows that

$$\mathbb{P}(B_n|F_n) \leq \lambda^{-M-2} \frac{p_e\left(\Phi^{(n)}, T^{(n)}, \Psi^{(n)}\right)}{\varepsilon_n^2} \xrightarrow[n \to \infty]{} 0$$

thus showing (60).

In order to show (59), suppose first that $\Pi_{\tau^{(n)}}^{(n)} \leq \varepsilon_n$. Then

$$\frac{|\mathcal{W}^{(n)}| - 1}{|\mathcal{W}^{(n)}|} \lambda^{\tau^{(n)}} \leq \Pi_{\tau^{(n)}}^{(n)} \leq \varepsilon_n. \tag{85}$$

For every fixed M in $\mathbb{N}$, define the event $F_n := \{\tau^{(n)} \leq M\}$. From (60) and (84), it follows that

$$\mathbb{P}(F_n) \leq \mathbb{P}\left(F_n \cap \left\{\Pi_{\tau^{(n)}}^{(n)} \leq \varepsilon_n\right\}\right) + \mathbb{P}\left(\Pi_{\tau^{(n)}}^{(n)} > \varepsilon_n\right)$$
$$\leq \mathbb{P}\left(\frac{|\mathcal{W}^{(n)}| - 1}{|\mathcal{W}^{(n)}|} \lambda^M \leq \varepsilon_n\right)$$
$$+ \mathbb{P}\left(\tau^{(n)} = T^{(n)}\right) \xrightarrow[n \to \infty]{} 0. \qquad \blacksquare$$

ACKNOWLEDGMENT

REFERENCES

[1] A. Arapostathis, V. S. Borkar, E. Fernandez-Gaucherand, M. K. Ghosh, and S. I. Marcus, "Discrete-time controlled Markov processes with average cost criterion: A survey," *SIAM J. Contr. Optim.*, vol. 31, pp. 282–344, 1993.

[2] P. Berlin, B. Nakiboglu, B. Rimoldi, and E. Telatar, "A Simple Converse of Burnashev's Reliability Function" 2006 [Online]. Available: http://arxiv.org/abs/cs/0610145

[3] V. S. Borkar, *Probability Theory: An Advanced Course*. New York: Springer, 1995.

[4] V. S. Borkar, "Convex analytic methods in Markov decision processes," in *Handbook of Markov Decision Processes*, E. Feinberg and A. Shwartz, Eds. Boston, MA: Kluwer, 2001.

[5] L. B. Boza, "Asymptotically optimal tests for finite Markov chains," *Ann. Math. Statist.*, vol. 42, pp. 1992–2007, 1971.

[6] M. V. Burnashev, "Data transmission over a discrete channel with feedback. Random transmission time," *Probl. Pered. Inf.*, vol. 12, pp. 10–30, 1976.

[7] M. V. Burnashev, "Sequential discrimination of hypotheses with control of observations," *Math. USSR—Izv.*, vol. 15, pp. 419–440, 1980.

[8] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley and Sons, 1991.

[9] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Budapest, Hunary: Akademiai Kiado, 1981.

[10] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd ed. New York: Springer, 1998.

[11] C. Derman, *Finite State Markovian Decision Processes*. New York: Academic, 1970.

[12] A. G. Djackov, "Upper bounds on the error probability for transmission with feedback in the case of memoryless discrete channels," *Probl. Pered. Inform.*, vol. 11, pp. 13–28, 1975.

[13] R. L. Dobrushin, "An asymptotic bound for the probability error of information transmission through a channel without memory using the feedback," *Probl. Kibern.*, vol. 8, pp. 161–168, 1962.

[14] G. D. Forney , Jr, "Exponential error bounds for erasure, list, and decision feedback schemes," *IEEE Trans. Inf. Theory*, vol. IT-14, no. 2, pp. 206–220, Mar. 1968.

[15] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[16] A. J. Goldsmith and P. P. Varaiya, "Capacity, mutual information, and coding for finite-state Markov channels," *IEEE Trans. Inf. Theory*, vol. 42, no. 3, pp. 868–886, May 1996.

[17] T. S. Han, *Information-Spectrum Method in Information Theory*. Tokio, Japan: Springer, 2003.

[18] E. A. Haroutunian, "Lower bound for error probability in channels with feedback," *Probl. Pered. Inform.*, vol. 13, pp. 36–44, 1977.

[19] M. Horstein, "Sequential transmission using noiseless feedback," *IEEE Trans. Inf. Theory*, vol. IT-9, no. 4, pp. 136–143, Jul. 1963.

[20] A. Manne, "Linear programming and sequential decision," *Manag. Sci.*, vol. 6, pp. 259–267, 1960.

[21] S. Mannor and J. Tsitsiklis, "On the empirical state-action frequencies in Markov decision processes under general policies," *Math. Oper. Res.*, vol. 30, pp. 545–561, 2005.

[22] C. McDiarmid, "Concentration," in *Probabilistic Methods in Algorithmic Discrete Mathematics*, M. Habib, Ed. *et al.* New York: Springer-Verlag, 1998.

[23] B. Nakiboglu and R. G. Gallager, "Error exponents for variable-length block codes with feedback and cost constraints," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 945–963, Mar. 2008.

[24] S. Natarajan, "Large deviations, hypotheses testing, and source coding for finite Markov chains," *IEEE Trans. Inf. Theory*, vol. IT-31, no. 3, pp. 360–365, May 1985.

[25] J. M. Ooi and G. W. Wornell, "Fast iterative coding techniques for feedback channels," *IEEE Trans. Inf. Theory*, vol. 44, no. 7, pp. 2960–2976, Nov. 1998.

[26] J. M. Ooi, *Coding for Channels With Feedback*. Boston, MA: Kluwer Academic Publishers, 1998.

[27] A. Sahai, "Why do block length and delay behave differently if feedback is present?," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 1860–1886, May 2008.

[28] C. E. Shannon, "The zero error capacity of a noisy channel," *IRE Trans. Inf. Theory*, vol. IT-2, no. 3, pp. 8–19, Sep. 1956.

[29] C. E. Shannon, "Channels with side information at the transmitter," *IBM J. Res. Devel.*, vol. 2, pp. 289–293, 1958.

[30] J. P. M. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback . I: No bandwidth constraint," *IEEE Trans. Inf. Theory*, vol. IT-12, no. 2, pp. 172–182, Apr. 1966.

[31] A. Y. Sheverdyaev, "Lower bound for error probability in a discrete memoryless channel with feedback," *Probl. Contr. Inform. Theory*, vol. 18, pp. 5–15, 1982.

[32] S. Tatikonda and S. Mitter, "The capacity of channels with feedback," *IEEE Trans. Inf. Theory*, vol. 55, no. 1, pp. 323–349, Jan. 2009.

[33] A. Tchamkerten and I. E. Telatar, "On the universality of Burnashev's error exponent," *IEEE Trans. Inf. Theory*, vol. 51, no. 8, pp. 2940–2944, Aug. 2005.

[34] A. Tchamkerten and İ. E. Telatar, "Variable length coding over an unknown channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2126–2145, May 2006.

[35] D. Williams, *Probability With Martingales*. Cambridge, U.K.: Cambridge Univ. Press, 1991.

[36] J. Wolfowitz, *Coding Theorems of Information Theory*, 2nd ed. Berlin, Germany: Springer-Verlag, 1964.

[37] A. D. Wyner, "On the Schalkwijk-Kailath coding scheme with a peak energy constraint," *IEEE Trans. Inf. Theory*, vol. IT-14, no. 1, pp. 129–134, Jan. 1968.

[38] H. Yamamoto and K. Itoh, "Asymptotic performance of a modified Schalkwijk-Barron scheme for channels with noiseless feedback," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 6, pp. 729–733, Nov. 1979.

[39] K. Zigangirov, "Upper bounds on the error probability for channels with feedback," *Probl. Pered. Inform.*, vol. 6, pp. 87–92, 1970.

Giacomo Como received the B.Sc., M.S., and Ph.D. degrees in applied mathematics from Politecnico di Torino, Torino, Italy, in 2002, 2004, and 2008, respectively.

During 2006–2007, he was Visiting Assistant in Research at the Department of Electrical Engineering, Yale University, New Haven, CT. He is currently a Postdoctoral Associate at the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA. His current research interests include coding theory, information theory, distributed estimation, and control.

Serdar Yüksel was born in Lice, Turkey, in 1979. He received the B.Sc. degree in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 2001 and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Illinois at Urbana-Champaign in 2003 and 2006, respectively.

He was a Postdoctoral researcher at Yale University, New Haven CT, for a year before joining Queen's University, Kingston, ON, Canada, as an Assistant Professor of Mathematics and Engineering at the Department of Mathematics and Statistics. His research interests are on stochastic control, information theory and decentralized control, and communications.

Sekhar Tatikonda (S'92–M'00) received the Ph.D. degree in electrical engineering and computer science from Massachusetts Institute of Technology, Cambridge, in 2000.

He is currently an Associate Professor of Electrical Engineering in the Department of Electrical Engineering, Yale University, New Haven, CT. From 2000 to 2002, he was a Postdoctoral Fellow in the Department of Computer Science, University of California, Berkeley. His current research interests include communication theory, information theory, stochastic control, distributed estimation and control, statistical machine learning, and inference.