A Single-Letter Upper Bound on the Feedback Capacity of Unifilar Finite-State Channels

Oron Sabag, *Student Member, IEEE*, Haim H. Permuter, *Senior Member, IEEE*, Henry D. Pfister, *Senior Member, IEEE*

***Abstract*— An upper bound on the feedback capacity of unifilar finite-state channels (FSCs) is derived. A new technique, called the Q-context mapping, is based on a construction of a directed graph that is used for a sequential quantization of the receiver's output sequences to a finite set of *contexts*. For any choice of Q-graph, the feedback capacity is bounded by a single-letter expression, $C_{\text{fb}} \leq \sup I(X, S; Y|Q)$, where the supremum is over $p(x|s, q)$ and the distribution of (S, Q) is their stationary distribution. It is shown that the bound is tight for all unifilar FSCs, where feedback capacity is known: channels where the state is a function of the outputs, the trapdoor channel, Ising channels, the no-consecutive-ones input-constrained erasure channel, and the memoryless channel. Its efficiency is also demonstrated by deriving a new capacity result for the dicode erasure channel; the upper bound is obtained directly from the above-mentioned expression and its tightness is concluded with a general sufficient condition on the optimality of the upper bound. This sufficient condition is based on a fixed point principle of the BCJR equation and, indeed, formulated as a simple lower bound on feedback capacity of unifilar FSCs for arbitrary Q-graphs. This upper bound indicates that a single-letter expression might exist for the capacity of finite-state channels with or without feedback based on a construction of auxiliary random variable with specified structure, such as the Q-graph, and not with i.i.d distribution. The upper bound also serves as a non-trivial bound on the capacity of channels without feedback, a problem that is still open.**

***Index Terms*— Converse, dicode erasure channel, feedback capacity, finite state channels, trapdoor channel, unifilar channels, upper bound.**

I. Introduction

A FINITE-STATE channel (FSC) is a mathematical model for channels with memory that has been applied to wireless communications [1], [2] and magnetic recording [3]. In this model, the channel memory is encapsulated in a state which takes values from a finite set. In this paper, we focus on unifilar FSCs with feedback, as described in Fig. 1, where the new channel state is a time-invariant function of the previous state, the current input and the current output.

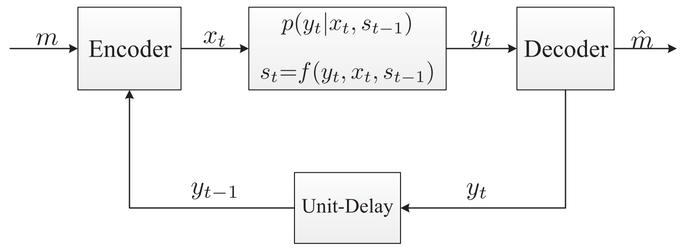

Fig. 1. Unifilar FSC with feedback.

The feedback capacity of FSCs has been investigated in [4]–[6] and still has no simple closed-form expression. For the special case of unifilar FSCs, it was shown in [7] that the feedback capacity is:

$$C_{\text{fb}} = \lim_{N\to\infty} \sup_{\{p(x_t|s_{t-1},y^{t-1})\}_{t=1}^{N}} \frac{1}{N}\sum_{i=1}^{N} I(X_i, S_{i-1}; Y_i|Y^{i-1}). \tag{1}$$

As can be seen from the capacity formula, this capacity expression is very hard to compute in a straightforward manner. However, it was shown in [6] and [7] that the capacity can be formulated as a dynamic programming (DP) optimization problem; this has benefits such as efficient algorithms for estimating the capacity and analytical tools for calculating capacity.

The relationship between the feedback capacity of FSCs and DP first appeared in Tatikonda's thesis [8]. The need for this formulation arises from difficulties in the computability of the capacity expression as can be seen in (1). In [9], a DP formulation of a sub-family of unifilar FSCs was given, where the state can be computed at the decoder. It was shown that the DP can be analytically solved under mild conditions on the channel, resulting in a computable capacity expression. DP formulations of feedback capacities appeared also for channels where the state is determined by the inputs [10], Markov channels [6] and Gaussian channels with stationary noise [11].

A typical approach for solving DP problems is the well-known Bellman equation. Loosely speaking, one should find a constant and a function which satisfy some fixed point equation; the constant is then the optimal reward (equivalent to the feedback capacity). This approach led to explicit capacity expressions for the trapdoor channel [7], the Ising channel [12], [13], the input-constrained erasure channel [14]

Manuscript received April 8, 2016; accepted October 5, 2016. Date of publication December 23, 2016; date of current version February 14, 2017.

O. Sabag and H. H. Permuter were supported in part by European Research Council, European Unions Seventh Framework Program (FP7/2007-2013)/ERC under Grant 337752 and in part by Joint UGC-ISF Research Grant. This paper was presented at the 2016 IEEE International Symposium on Information Theory.

O. Sabag and H. H. Permuter are with the Department of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Beersheba 8410501, Israel (e-mail: oronsa@post.bgu.ac.il; haimp@bgu.ac.il).

H. D. Pfister is with the Department of Electrical and Computer Engineering, Duke University, Durham 27708, NC USA (e-mail: henry.pfister@duke.edu).

Communicated by T. Liu, Associate Editor for Shannon Theory.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2016.2636851

and the input-constrained binary symmetric channel [15]. The difficulty in the Bellman equation based approach lies in the fact that the state space of the DP problem is uncountably infinite, so finding a function which satisfies this equation is a challenging task.

Nevertheless, computer-based simulations of DP provide crucial insights into the feedback capacity of a particular channel. Specifically, implementation of the value-iteration algorithm together with DP simulation can be used to derive an analytic expression for the feedback capacity. The numerical results naturally provide a lower bound on the feedback capacity, so it remains to provide an upper bound. One option is solving the Bellman equation but this is quite challenging so an alternative method to calculate tight upper bounds on the feedback capacity is desirable. Our main result is a derivation of a simple upper bound on the feedback capacity of unifilar FSCs.

The derivation of the upper bound is initiated with an almost trivial inequality which, for a sequence of deterministic mappings $\{\Phi_i\}_{i\geq 1}$, is given by

$$\begin{aligned} H(Y_i|Y^{i-1}) &= H(Y_i|Y^{i-1}, \Phi_{i-1}(Y^{i-1})) \\ &\leq H(Y_i|\Phi_{i-1}(Y^{i-1})). \end{aligned}$$

The equality follows from the fact that $\Phi_{i-1}(Y^{i-1})$ is a deterministic function of Y^{i-1}, and the inequality follows from the fact that conditioning reduces entropy. A naive choice of the mappings $\{\Phi_i\}_{i\geq 1}$ can result in an equality in this upper bound; for example, if the mapping Φ_{i-1} returns a constant then the above inequality becomes $H(Y_i|Y^{i-1}) \leq H(Y_i)$ which is known to be a tight upper bound for the memoryless channel where i.i.d. outputs process is optimal. For non-trivial FSCs, this naive choice results an upper bound that is not tight. Throughout this paper, it will be shown that a structured mapping may improve the upper bound performance.

Our derivation is based on a new technique called the Q-context mapping which is defined by the sequence of functions $\{\Phi_i\}_{i\geq 1}$ that transform the history of the output process into a Markov chain. Specifically, a time-invariant Q-context mapping is described by a directed graph (called the Q-graph) and it has the property that the outgoing edges per node are labelled with all possible channel outputs. Then, given an initial node on the graph, any-length sequence can be mapped into a unique node by walking along the labelled edges until the sequence ends. Thus, the Q-graph describes a *sequential quantization* of output sequences to the set of nodes which are called contexts.

Combining the Q-context mapping technique with the capacity expression in (1) leads to our main result, a single-letter upper bound on the feedback capacity of unifilar FSCs:

$$C_{\text{fb}} \leq \sup_{p(x|s,q)} I(X, S; Y|Q), \tag{2}$$

for all Q-graphs, where the distribution of (S, Q) is given by the stationary distribution induced by the time-invariant Q-context mapping. It is shown that the upper bound is tight for all unifilar FSCs where the feedback capacity is already known and, obviously, for the memoryless channel with feedback. Therefore, the derived upper bound also provides a unified expression for all feedback capacities known so far. This result provides hope that the feedback capacity of general FSCs might be characterized by a single-letter expression, which would be quite surprising.

Throughout the paper, we demonstrate that the bound is tight for a proper choice of the Q-graph for any channel where the state is computable at the decoder [9], the trapdoor channel [7], the input-constrained binary erasure channel (BEC) [14] and Ising channels [12], [13]. These derivations also serve as an easily implemented and alternative converse proof for these capacity results.

It is also demonstrated that the upper bound can be used to derive new results, such as the feedback capacity of the dicode erasure channel (DEC). The DEC is a simplified model of the known dicode channel with additive white Gaussian noise (AWGN), which was studied in [16] and [17]. DP-based simulations for this channel show that the optimal policy only visits a finite subset of the state space and the actions associated with those states are unconstrained. Since actions are unconstrained, the solution of the Bellman equation is very challenging, if not impossible. However, since the set of visited states is finite, it is possible to extract a Q-graph from this simulation and to derive a simple upper bound on the capacity. Its tightness follows from checking a sufficient condition that implies the upper bound in (2) is achievable.

The sufficient condition is based on an invariant-property of the BCJR equation for the channel state estimation, $p(s_t|y^t)$. Roughly speaking, the condition states that, for a Q-graph and some input distribution $p(x|s,q)$, if the state estimate $p(s_t|y^t)$ depends only on the Q-context $\Phi_t(y^t)$ and not on the entire sequence y^t, then $I(X, S; Y|Q)$ is an achievable rate. This condition is easy to verify using the BCJR forward-recursive equation for unifilar FSCs and may be exploited in two ways. The first is to verify that the upper bound is tight, as is done for the DEC. The second is to provide a computable lower bound for an arbitrary Q-graph and input $p(x|s,q)$ that satisfy the condition.

The remainder of the paper is organized as follows. Section II defines notation and provides mathematical background. Section III states our main result on the upper bound and the sufficient condition for the tightness of this bound. In Section IV, several examples of unifilar FSCs are studied and it is shown that the upper bound is tight. In Section V, we provide a detailed proof of the main result. Finally, the paper is concluded in Section VI.

II. Notation and Preliminaries

Random variables are denoted by upper-case letters, such as X, while realizations are denoted by lower-case letters, e.g., x and calligraphic letters, e.g., $\mathcal{X}$, denote sets. We use X^n to denote the n-tuple $(X_1, \ldots, X_n)$ and x^n to denote vectors of n elements, i.e., $x^n = (x_1, x_2, ..., x_n)$. The binary entropy is denoted by $H_2(\alpha) = -\alpha \log_2 \alpha - (1-\alpha)\log_2(1-\alpha)$, where $\alpha \in [0, 1]$. Finally, $H_3(\alpha_1, \alpha_2) = -\alpha_1 \log_2 \alpha_1 - \alpha_2 \log_2 \alpha_2 - (1-\alpha_1-\alpha_2)\log_2(1-\alpha_1-\alpha_2)$ denotes the ternary entropy function for scalars $\alpha_1, \alpha_2 \in [0, 1]$ satisfying $\alpha_1 + \alpha_2 \leq 1$. The quaternary entropy function, $H_4(\alpha_1, \alpha_2, \alpha_3)$, is defined in a similar manner. Probability mass functions (PMFs) are denoted by $p(X = x)$, the conditional probability of $X = x$ given $Y = y$ is denoted by $p(X = x|Y = y)$ and the joint distribution is denoted as $p(X = x, Y = y)$; when the random variable is clear from the context we use the short hands $p(x)$, $p(x|y)$ and $p(x, y)$, respectively. The $(|\mathcal{S}| - 1)$-dimensional simplex is denoted by $\mathcal{Z}$.

A. Unifilar Finite-State Channels

A *finite state channel* is defined by the triplet $(\mathcal{X} \times \mathcal{S}, p(y, s|x, s'), \mathcal{Y} \times \mathcal{S})$ where X is the channel input, Y is the channel output, S' is the channel state at the beginning of the transmission and S is the channel state at the end of the transmission. The cardinalities $\mathcal{X}, \mathcal{Y}, \mathcal{S}$ are assumed to be finite. At each time t, the channel has the property

$$p(s_t, y_t|x^t, s^{t-1}, y^{t-1}) = p(s_t, y_t|x_t, s_{t-1}).$$

An FSC is called *unifilar* if the new channel state, s_t, is given by a time-invariant function $f(x_t, y_t, s_{t-1})$ of the input, output, and previous state.

The input to the channel at time t, x_t, depends both on the message m and on the output tuple y^{t-1}. A unifilar channel is *strongly connected* if for all $s, s' \in \mathcal{S}$, there exist T and $\{p(x_t|s_{t-1})\}_{t=1}^{T}$ such that $\sum_{t=1}^{T} p(S_t = s|S_0 = s') > 0$. It is also assumed that the initial state, s_0, is available to both the encoder and the decoder.

The capacity of the unifilar FSC is given by the following theorem:

Theorem 1 ([7, Th. 1]): The feedback capacity of a strongly connected unifilar FSC, where s_0 is available to both to the encoder and the decoder, can be expressed by

$$C_{\text{fb}} = \lim_{N\to\infty} \sup_{\{p(x_t|s_{t-1}, y^{t-1})\}_{t=1}^{N}} \frac{1}{N}\sum_{i=1}^{N} I(X_i, S_{i-1}; Y_i|Y^{i-1}).$$

B. Labelled Directed Graphs

A labelled directed graph is defined by a set of nodes, a set of directed edges, and a label function mapping edges to labels. Node i is said to *communicate* with node j if there exists a path from i to j. This definition leads to an equivalence relation: two nodes i, j lie in the same *communicating class* if i communicates with j and vice versa. A communicating class is said to be closed if there are no outgoing edges from this class. A graph is *irreducible* if all nodes in the graph communicate.

For a closed communicating class, the *period* of a node is defined as the gcd of all natural numbers, n, such that there is a loop to this node with length n. It can be shown that the period is a class property, that is, all nodes in a closed communicating class have equal periods. A closed class is *aperiodic* if it has a period of 1.

One useful property of irreducible graphs with period D is that the graph can be partitioned uniquely into D disjoint subsets $A_0, A_1, \ldots, A_{D-1}$ on a cycle, i.e., all edges from A_i lead to $A_{(i+1) \bmod D}$.

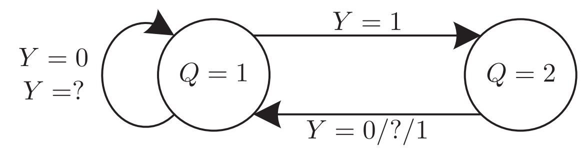

Fig. 2. An example for a Q-graph with $|\mathcal{Q}| = 2$, and $\mathcal{Y} = \{0, 1, ?\}$.

C. Q-Context Mapping

The upper bound in this paper is based on the inequality:

$$H(Y_i|Y^{i-1}) \le H(Y_i|\Phi_{i-1}(Y^{i-1})), \quad i \in \mathbb{N},$$

which holds for any set of mappings $\Phi_{i-1} : \mathcal{Y}^{i-1} \to \mathcal{Q}$. The *context* of the sequence y^{i-1} is defined as $q_{i-1} \triangleq \Phi_{i-1}(y^{i-1})$.

Our interest is limited to the set of mappings which can be described by a time-invariant function $g : \mathcal{Q} \times \mathcal{Y} \to \mathcal{Q}$, where $\Phi_i(y^i) = g(\Phi_{i-1}(y^{i-1}), y_i)$ for all i. The Q-context mapping is defined by a function $g(\cdot, \cdot)$ or, equivalently, by a Q-graph with $|\mathcal{Q}|$ nodes, each taking a realization $q \in \mathcal{Q}$; an edge $q \to q'$ with label y exists if $q' = g(q, y)$. It is assumed that the Q-graph is finite and irreducible. These definitions imply that each node in the Q-graph has $|\mathcal{Y}|$ outgoing edges. An example for a Q-graph is illustrated in Fig. 2.

The next step is to embed the FSC characterization into the Q-graph. This is done by constructing a new directed graph which includes the entire information on the Q-graph and the channel states evolution. An (S, Q)*-graph* is constructed as follows:

1) Each node in the Q-graph is split into $|\mathcal{S}|$ new nodes, which are represented by the pairs $(s, q) \in \mathcal{S} \times \mathcal{Q}$.
2) An edge $(s, q) \to (s', q')$ with a label (x, y) exists if and only if there exists a pair (x, y) such that $s' = f(s, x, y)$, $q' = g(q, y)$, and $p(y|x, s) > 0$.

The (S, Q)-graph might have more than a single closed communicating class. It is then clear that if the initial pair (s_0, q_0) lies in a closed communicating class, then all other classes will never be reached. Recall that s_0 is given by the problem, while the initial context q_0 can be chosen. The following lemma formalizes a few properties of the (S, Q)-graph that will be used to simplify our analysis.

Lemma 1: There exists at least one closed communicating class in the (S, Q)-graph. For every $s \in \mathcal{S}$ (or $q \in \mathcal{Q}$) and for every closed communicating class, $\mathcal{C}$, there exists $q \in \mathcal{Q}$ (or $s \in \mathcal{S}$) such that $(s, q) \in \mathcal{C}$.

The proof of Lemma 1 appears in Appendix A. The freedom of choosing q_0, together with Lemma 1, verifies that for a given s_0 there always exists q_0 such that (s_0, q_0) lies within any of the closed classes. Therefore, we will assume throughout this paper that the (S, Q)-graph has a single closed communicating class only. There is no concrete example where the initial closed class effects the upper bound, but one should be aware that if different closed classes give different upper bounds, then each value is a valid upper bound.

In order to present the (S, Q)-graph as a Markov chain on $\mathcal{S} \times \mathcal{Q}$, probabilities should be assigned on the edges. For a given input matrix $p(x|s, q)$, an outgoing edge from (s, q) that is labelled by (x, y) will have a probability of

$p(y|x,s)p(x|s,q)$. This assignment might effect the structure of the (S,Q)-graph; specifically, if an edge has $p(x|s,q)=0$ then it can be removed. Denote by $\mathcal{A}(p(x|s,q))$ the (S,Q)-graph after edge removal and define

$$\mathcal{P}_\pi \triangleq \{p(x|s,q) : \mathcal{A}(p(x|s,q)) \text{ has a single closed class}\} \quad (3)$$

The subscript π emphasizes that for all $p(x|s,q)\in\mathcal{P}_\pi$ there exists a unique stationary distribution on the (S,Q)-graph that is denoted by $\pi(S=s,Q=q)$, or for short as $\pi(s,q)$. This stationary distribution can be calculated directly on the induced single closed class, as all other nodes are inessential and have zero probability.

III. Main Result

The following theorem is our main result.

Theorem 2: If the initial state s_0 is available to both the encoder and decoder, then the feedback capacity C_{fb} of a strongly connected unifilar FSC satisfies

$$C_{\text{fb}} \leq \sup_{p(x|s,q)\in\mathcal{P}_\pi} I(X,S;Y|Q), \quad (4)$$

for all irreducible Q-graphs with q_0 such that (s_0,q_0) lies in an aperiodic closed communicating class. The random variables Y,X,S,Q are associated with the time-invariant system and the joint distribution is $p(y,x,s,q)=p(y|x,s)p(x|s,q)\pi(s,q)$ where $\pi(s,q)$ is the stationary distribution of the (S,Q)-graph associated with the initial vertex (s_0,q_0).

Remark 1: The random variable Q is an auxiliary random variable (RV) representing the comon knowledge that is shared by the encoder and decoder. Here, the implied sequence of auxiliary RVs has memory induced by the structure of the chosen Q-graph and is encapsulated in its stationary distribution. In contrast, auxiliary RVs are typically chosen to be independent and identically distributed (i.i.d.). For example, consider the Wyner-Ziv and Gelfand-Pinsker models.

The auxiliary RV in (4) can be thought of as a dual version of the conventional auxiliary RV in capacity expressions; mostly, in capacity expressions, any choice of auxiliary RV distribution results an achievable rate (i.e., a lower bound), while in (4) any choice of an auxiliary Q-graph results an upper bound.

In standard derivations of upper bounds on capacities, it is shown that auxiliary RVs exist, while in (4) the upper bound holds for all irreducible Q-graphs. Indeed, if it could be shown that there is always an optimal Q-graph with a finite number of graph nodes, then minimizing over all possible Q-graphs would transform (4) into a single-letter capacity formula for any unifilar FSC with feedback.

Remark 2: The restriction on the input distributions in $\mathcal{P}_\pi$ implies that a unique stationary distribution exists. Note that the stationary distribution $\pi(s,q)$ depends on the value of $p(x|s,q)$, and can be found as the unique solution of

$$\pi(s',q') = \sum_{(s,q)\in G}\pi(s,q)\sum_{y:g(q,y)=q'}\sum_{x:f(x,y,s)=s'} p(x|s,q)p(y|x,s).$$

Remark 3: As discussed in Section II, if the (S,Q)-graph contains more than a single closed class, then the upper bound holds for all closed communicating classes which are aperiodic. This fact follows from Lemma 1, where it is shown that each closed class contains all $s_0\in\mathcal{S}$ and all $q_0\in\mathcal{Q}$.

Remark 4: Since the transmitter is free to ignore the feedback, the feedback capacity is greater than or equal to the non-feedback capacity. Thus, Theorem 2 also provides a computable and non-trivial upper bound on the non-feedback capacity of a unifilar FSC, which remains an open problem.

Remark 5: An efficient method for finding the optimal Q-graph is to study the corresponding DP. Standard simulations (see [7], [12], [14]) produce a histogram of the DP states that are visited under an estimated optimal policy. The inaccuracy of such simulations follows from the required quantization of the DP parameters.

When the resulting histogram of the DP states is discrete, i.e., only a finite number of DP states are visited, then a Q-graph can be extracted from the DP simulation. Specifically, each visited DP state is taken as a node in the Q-graph and the labelled edges are the evolution of the DP states as a function of the outputs.
In the following section, a sufficient condition for the optimality of the upper bound is provided.

A. Lower Bound on Capacity

Before presenting the lower bound, the BCJR recursive equation of the channel state estimation is derived. For the outputs tuple y^t, and $s_t\in\mathcal{S}$, this gives

$$\begin{aligned} p(s_t|y^t) &= \frac{p(s_t,y_t|y^{t-1})}{p(y_t|y^{t-1})} \\ &= \frac{\sum_{x_t,s_{t-1}} p(s_t,y_t,x_t,s_{t-1}|y^{t-1})}{\sum_{s_t,x_t,s_{t-1}} p(s_t,y_t,x_t,s_{t-1}|y^{t-1})}. \end{aligned} \quad (5)$$

This is a forward-recursive equation in the sense that with a set of scalars $\{p(s_{t-1}|y^{t-1})\}_{s_{t-1}\in\mathcal{S}}$ and an output symbol, y_t, one can compute the set $\{p(s_t|y^t)\}_{s_t\in\mathcal{S}}$. Note that the collection of scalars $p(s_{t-1}|y^{t-1})$ is an element from $\mathcal{Z}$ (recall that $\mathcal{Z}$ denotes the $(|\mathcal{S}|-1)$-dimensional simplex). For each input, $p(x|s,q)$, one can write the BCJR equation in (5) as a mapping $B_s:\mathcal{Z}\times\mathcal{Y}\to[0,1]$ where the subscript s stands for the state whose probability is being estimated.

Given an irreducible Q-graph, an input distribution $p(x|s,q)\in\mathcal{P}_\pi$ is said to be an *aperiodic input* if its corresponding (S,Q)-graph is aperiodic. Each aperiodic input induces a stationary distribution, $\pi(s,q)$, and we say that an aperiodic input is *BCJR-invariant* if its state probability vector, $\pi(S|Q=q)\triangleq\{\pi(S=s|Q=q)\}_{s\in\mathcal{S}}$, satisfies

$$\pi(S=s|Q=g(q,y)) = B_s(\pi(S|Q=q),y)$$

for all $s,q\in\mathcal{S}\times\mathcal{Q}$ and $y\in\mathcal{Y}$, where $g(q,y)$ is the context that is calculated from the node q and the output y.

The following theorem provides a lower bound on feedback capacity.

Theorem 3: The feedback capacity of unifilar FSCs is bounded by

$$C_{\text{fb}} \geq I(X,S;Y|Q), \quad (6)$$

for all aperiodic inputs, $p(x|s,q) \in \mathcal{P}_\pi$, that are BCJR-invariant.

Remark 6: Theorem 3 acts as a complementary tool for the upper bound in Theorem 2. One application is to evaluate the upper bound from Theorem 2 for some Q-graph and then to verify its optimality by the BCJR-invariant property. However, there are cases where the upper bound is tight and the corresponding BCJR-invariant property is not satisfied; therefore, this property is a sufficient but not necessary for the optimality of the upper bound. Nevertheless, the above statement suggests a lower bound for all aperiodic inputs, so it can be exploited as a lower bound with an arbitrary aperiodic input, as we will see in Section IV-B.

IV. Examples

This section covers several examples from the literature where the capacity of a unifilar FSC is known.

A. Channel State is a Function of the Outputs

In [9], a unifilar FSC where the channel state is available to all parties and evolves according to $p(s_{t+1}|x_t, s_t)$ was studied. It was shown that this class of FSCs is, indeed, equivalent to a unifilar FSC where the channel state is the last output, i.e., $s_i = y_i$. The authors showed that for channels with strongly irreducible and aperiodic states,[1] the capacity is given by $C_{\text{fb}} = \max_{p(x|s)} I(X; Y|S)$, where $p(y, x, s) = p(y|x, s)p(x|s)\pi(s)$.

To apply Theorem 2 for this case, the Q-graph is taken as the states graph since states can be computed from outputs. If the channel states form an aperiodic graph, then Theorem 2 gives

$$C_{\text{fb}} \le \sup_{p(x|s)\in\mathcal{P}_\pi} I(X; Y|S). \quad (7)$$

Note that the derived bound (7) holds for outputs that are strongly connected and form an aperiodic graph, while in [9], the outputs are assumed to be strongly irreducible and aperiodic, which is a stronger property.

B. Input-Constrained BEC

The setting consists of a BEC, where inputs must admit the $(1,\infty)$-RLL constraint, i.e., the input sequence does not contain consecutive ones. This setting does not fall into the classical definition of unifilar FSCs. However, it is possible to convert the input constraint into a channel state, $s_i = x_i$, and to derive the upper bound in Theorem 2, when the maximization is over constrained inputs. The feedback capacity of this channel was found in [14] using an explicit and tedious solution for the Bellman equation.

The following result is a re-statement of the known feedback capacity.

Theorem 4 ([14, Th. 1]): The feedback capacity of the input-constrained BEC is

$$C_{\text{icBEC}} = \max_{0\le p\le 0.5} \frac{H_2(p)}{\frac{1}{1-\epsilon}+p}. \quad (8)$$

[1]A channel is strongly irreducible if the graph with $|\mathcal{S}|$ nodes (each corresponds to an output) and the edge $s' \to s$ exists if $p(s|x,s') > 0$ for all x, is irreducible. Strong aperiodicity is defined in a similar manner.

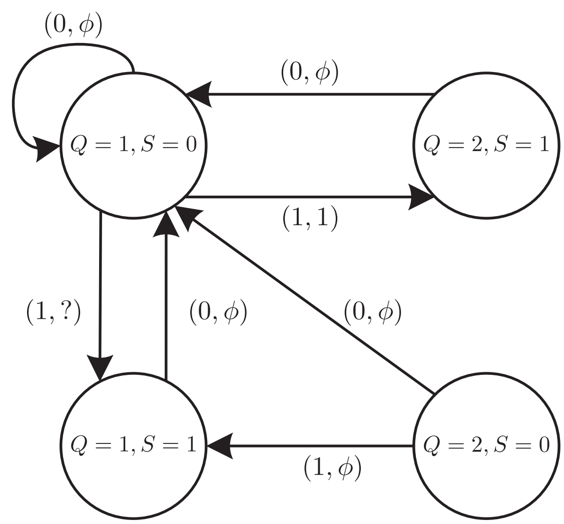

Fig. 3. The (S, Q)-graph for the input-constrained BEC for the Q-graph in Fig. 2. Each edge is labelled by a pair (x, y), where ϕ stands for the "don't care" symbol, i.e., all possible outputs.

Here, we will provide an alternative proof for Theorem 4: the upper bound is shown by applying Theorem 2 with the Q-graph presented in Fig. 2, while the lower bound is achieved by applying Theorem 3 with a new Q-graph that is presented in Fig. 4.

Remark 7: In this example, the BCJR-invariant property is not satisfied for the graph that is presented in Fig. 2 and, therefore, this is a sufficient but not a necessary condition for the tightness of the upper bound. On the other hand, calculation of the upper bound with the Q-graph in Fig. 4 results in a tight upper bound as well; however, it is preferable to calculate the upper bound using a Q-graph with the fewest nodes.

Upper Bound: The (S, Q)-graph for the Q-graph from Fig. 2 is presented in Fig. 3. There is a single closed class in this graph consisting of all nodes but $(Q = 2, S = 0)$, and this class is aperiodic since there is a loop of length 1. Since inputs are constrained, we have $p(X = 1|S = 1, Q = q) = 0$ for all q and, therefore, the matrix $p(x|s, q)$ can be parameterized with a single parameter $a \triangleq p(X = 1|S = 0, Q = 1)$.

Calculation of the stationary distribution for the (S, Q)-graph gives $[\pi_{0,1}, \pi_{1,1}, \pi_{1,2}] = \left[\frac{1}{1+a}, \frac{\epsilon a}{1+a}, \frac{(1-\epsilon)a}{1+a}\right]$, where we use the shorthand $\pi_{i,j} = \pi(S = i, Q = j)$. Then, one can calculate the conditional distribution,

$$\begin{aligned} p(Y=1|Q=1) &= p(Y=1, X=1, S=0|Q=1) \\ &= (1-\epsilon)a\frac{\pi_{0,1}}{\pi_{0,1}+\pi_{1,1}} \\ &= \frac{(1-\epsilon)a}{1+\epsilon a}. \end{aligned} \quad (9)$$

By Theorem 2, we have:

$$\begin{aligned} &C_{\text{icBEC}} \\ &\le \sup_{p(x|s,q)\in\mathcal{P}_\pi} I(X,S;Y|Q) \\ &= \sup_{p(x|s,q)\in\mathcal{P}_\pi} H(Y|Q) - H_2(\epsilon) \end{aligned}$$

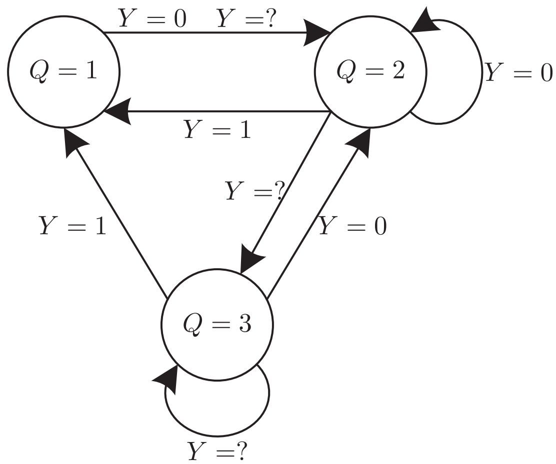

Fig. 4. Q-graph for the input-constrained BEC with an output alphabet $\mathcal{Y} = \{0, 1, ?\}$. Each node in the Q-graph can be interpreted by the corresponding decoder's knowledge: when $Q = 1$ (or $Q = 2$), the decoder knows that the state is $S = 1$(or $S = 0$), while in $Q = 3$ the decoder does not know the channel state.

$$\overset{(a)}{=} \max_{0\le a\le 1} [\pi_{0,1} + \pi_{1,1}]H_3\left(\frac{(1-\epsilon)a}{1+\epsilon a}, \epsilon\right) + \pi_{1,2}H_2(\epsilon) - H_2(\epsilon)$$
$$\overset{(b)}{=} \max_{0\le a\le 1} \frac{1+\epsilon a}{1+a}(1-\epsilon)H_2\left(\frac{a}{1+\epsilon a}\right)$$
$$\overset{(c)}{\le} \max_{0\le p\le 1} \frac{H_2(p)}{\frac{1}{1-\epsilon}+p},$$

where (a) follows from $\pi_{0,2} = 0$ and substituting (9), (b) follows from the identity $H_3((1-\epsilon)x, \epsilon) = H_2(\epsilon) + (1-\epsilon)H_2(x)$, for all $x \in [0, 1]$, and $\sum_{i,j} \pi_{i,j} = 1$. Next, (c) follows by changing the maximization variable to $p \triangleq \frac{a}{1+\epsilon a}$ and then expanding the new maximization domain to $[0, 1]$. Finally, the given expression holds because the function is decreasing for $p > 0.5$. ■

Lower Bound: consider the Q-graph that is presented in Fig. 4 with inputs that are given by $p(X = 1|S = 0, Q = 2) = p$ and $p(X = 1|S = 0, Q = 3) = \frac{p}{1-p}$ where $p \in [0, 0.5]$.

Construction of the (S, Q)-graph reveals that the pairs $(S = 0, Q = 1)$ and $(S = 1, Q = 2)$ cannot be reached, while the stationary distribution of the other pairs is positive and equals:

$$\pi_{1,1} = \frac{\bar{\epsilon}p}{1+\bar{\epsilon}p}$$
$$\pi_{0,2} = \frac{\bar{\epsilon}}{1+\bar{\epsilon}p}$$
$$\pi_{0,3} = \frac{\epsilon(1-p)}{1+\bar{\epsilon}p}$$
$$\pi_{1,3} = \frac{\epsilon p}{1+\bar{\epsilon}p}, \tag{10}$$

where $\pi_{i,j} = \pi(S = i, Q = j)$. By (10), it can be calculated that

$$\pi(S = 0|Q = 1) = 0$$
$$\pi(S = 0|Q = 2) = 1$$
$$\pi(S = 0|Q = 3) = 1 - p.$$

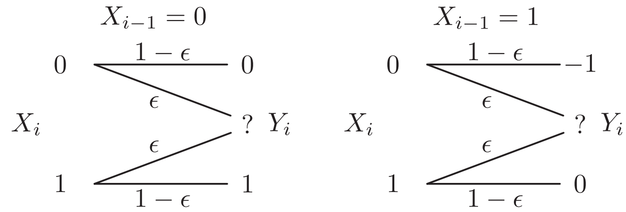

Fig. 5. DEC with erasure probability ϵ.

Since $|\mathcal{S}| = 2$, the value of $\pi(S = 0|Q = i)$ determines uniquely the value of $\pi(S = 1|Q = i)$ and it is sufficient to show the BCJR-invariant property for $\pi(S = 0|Q = i)$. The BCJR equation can then be written as:

$$p(S = 0|Q = g(i, y)) = \begin{cases} 1 & \text{if } Y = 0, \\ 1 - p(X = 1, S = 0|Q = i) & \text{if } Y = ?, \\ 0 & \text{if } Y = 1, \end{cases}$$

where $p(X = 1, S = 0|Q = i) = \pi(S = 0|Q = i)p(X = 1|S = 0, Q = i)$. We show the BCJR-invariant property for each node: the node $Q = 1$ has input edges that are labeled by $Y = 1$ only and, therefore, $p(S = 0|Q = g(i, 1)) = \pi(S = 0|Q = 1) = 0$ for $i = 2, 3$, as required. For the node $Q = 2$, all incoming edges are labelled by $Y = 0$ except for the edge $(Q = 1) \to (Q = 2)$ that is labelled with $Y = ?$. For this edge, calculation gives that $1 - \pi(S = 0|Q = 1)p(X = 1|S = 0, Q = 1) = 1$ and it can be concluded that the node $Q = 2$ is BCJR-invariant as well. Finally, $Q = 3$ has two incoming edges that satisfy $1 - \pi(S = 0|Q = 2)p(X = 1|S = 0, Q = 2) = 1 - \pi(S = 0|Q = 3)p(X = 1|S = 0, Q = 3) = 1 - p$.

By Theorem 3,

$$C_{\text{icBEC}} \ge I(X, S; Y|Q) = (1-\epsilon)\left[\frac{\bar{\epsilon}}{1+\bar{\epsilon}p}H_2(p) + \frac{\epsilon}{1+\bar{\epsilon}p}H_2(p)\right] = \frac{H_2(p)}{\frac{1}{1-\epsilon}+p}. \tag{11}$$

Since the lower bound (11) holds for all $p \in [0, 0.5]$, maximization on this parameter can be performed and concludes the proof of this theorem. ■

C. Dicode Erasure Channel (DEC)

The DEC [16], [18], as described in Fig. 5, is a simplified version of the well-known dicode channel with AWGN. Specifically, a binary input goes through a discrete-time linear filter described by $1 - D$, i.e., the filter outputs $x_i - x_{i-1}$ on the real line, and this is then transmitted on an erasure channel.

Inputs are taken from $\mathcal{X} = \{0, 1\}$, while outputs take values in $\mathcal{Y} = \{-1, 0, 1, ?\}$. The channel output is $y_i = x_i - x_{i-1}$ with probability $1 - \epsilon$, and equals $y_i = ?$ with probability ϵ, where ϵ is a parameter in $[0, 1]$. It is evident that the DEC is a unifilar FSC if the channel state is taken as the previous input, i.e., $s_i = x_i$.

The following theorem encapsulates the feedback capacity for the DEC.

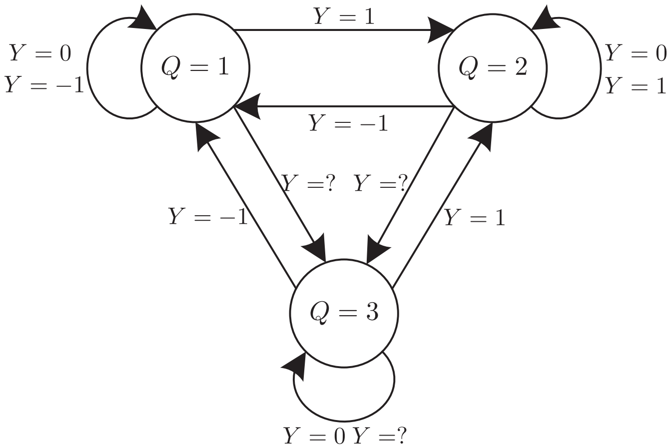

Fig. 6. Q-graph for the DEC with an output alphabet $\mathcal{Y} = \{0, 1, -1, ?\}$.

Theorem 5 (DEC Capacity): The feedback capacity of the DEC is:

$$C_{\text{DEC}} = \max_{0\le p\le 1}(1-\epsilon)\frac{p+\epsilon H_2(p)}{\epsilon+(1-\epsilon)p}. \tag{13}$$

Theorem 5 is obtained by calculating the upper bound with the Q-graph from Fig. 6, and the lower bound follows from the sufficient condition provided in Theorem 3. Indeed, this Q-graph has a nice interpretation as $Q = 1$ and $Q = 2$ correspond to a perfect knowledge of the channel state by the decoder, while $Q = 3$ implies that the decoder does not know the channel state.

Proof of Theorem 5: For the Q-graph in Fig. 6, its corresponding (S, Q)-graph can be described with a matrix. Each input to the matrix corresponds to a pair (x, y) of input and output, and we use ϕ as a notation for all possible channel outputs.

(S,Q)	$(0,1)$	$(1,1)$	$(0,2)$	$(1,2)$	$(0,3)$	$(1,3)$
$(0,1)$	$(0,0)$	$-$	$-$	$(1,1)$	$(0,?)$	$(1,?)$
$(1,1)$	$(0,-1)$	$(1,0)$	$-$	$-$	$(0,?)$	$(1,?)$
$(0,2)$	$-$	$-$	$(0,0)$	$(1,1)$	$(0,?)$	$(1,?)$
$(1,2)$	$(0,-1)$	$-$	$-$	$(1,0)$	$(0,?)$	$(1,?)$
$(0,3)$	$-$	$-$	$-$	$(1,1)$	$(0,\phi)$	$(1,?)$
$(1,3)$	$(0,-1)$	$-$	$-$	$-$	$(0,?)$	$(1,\phi)$

From the matrix above, it can be noted that the nodes $(S = 1, Q = 1)$ and $(S = 0, Q = 2)$ are not in the single closed communicating class that is formed by all other states. This closed class is aperiodic since there is a loop with length 1 for the node $(S = 0, Q = 1)$.

By exploiting the symmetry of the channel and the (S, Q)-graph, the maximization on input distributions can be limited to:

$$p(X=0|S=0,Q=1) = p(X=1|S=1,Q=2) = a$$
$$p(X=1|S=0,Q=3) = p(X=0|S=1,Q=3) = p,$$

where $a, p \in [0, 1]$ are the parameters of the input distribution. It follows that the stationary distribution is:

$$\pi_{0,1} = \frac{(1-\epsilon)p}{2\epsilon+2(1-\epsilon)p}$$
$$\pi_{1,2} = \frac{(1-\epsilon)p}{2\epsilon+2(1-\epsilon)p}$$
$$\pi_{0,3} = \frac{\epsilon}{2\epsilon+2(1-\epsilon)p}$$
$$\pi_{1,3} = \frac{\epsilon}{2\epsilon+2(1-\epsilon)p}, \tag{14}$$

where $\pi_{i,j} = \pi(S = i, Q = j)$.

Consider the following chain of equalities:

$$\begin{aligned} &H(Y|Q)\\ &= \sum_{q=1}^{3}(\pi_{0,q}+\pi_{1,q})H(Y|Q=q)\\ &\overset{(a)}{=} \frac{(1-\epsilon)p}{\epsilon+(1-\epsilon)p}H_3((1-\epsilon)a,(1-\epsilon)(1-a))\\ &\quad + \frac{\epsilon}{\epsilon+(1-\epsilon)p}H_4\left(\epsilon,(1-\epsilon)\frac{p}{2},(1-\epsilon)\frac{p}{2}\right)\\ &\overset{(b)}{=} (1-\epsilon)\left[\frac{(1-\epsilon)pH_2(a)}{\epsilon+(1-\epsilon)p} + \frac{\epsilon H_3\left(\frac{p}{2},\frac{p}{2}\right)}{\epsilon+(1-\epsilon)p}\right] + H_2(\epsilon)\\ &\overset{(c)}{=} (1-\epsilon)\left[\frac{(1-\epsilon)pH_2(a)}{\epsilon+(1-\epsilon)p} + \frac{\epsilon(p+H_2(p))}{\epsilon+(1-\epsilon)p}\right] + H_2(\epsilon), \end{aligned} \tag{15}$$

where (a) is obtained by substituting the stationary distribution from (14), (b) follows from the identity $H_3((1-\delta)\gamma, (1-\delta)(1-\gamma)) = H_2(\delta) + (1-\delta)H_2(\gamma)$ by choosing $\delta = \epsilon, \gamma = a$. Finally, (c) follows from the above identity by choosing $1-\delta = p$ and $\gamma = \frac{1}{2}$.

The upper bound on the capacity can then be established:

$$\begin{aligned} C_{\text{DEC}} &\le \sup_{p(x|s,q)\in\mathcal{P}_\pi} I(X,S;Y|Q)\\ &\overset{(a)}{=} \max_{(a,p)\in[0,1]^2} H(Y|Q) - H(Y|X,S,Q)\\ &\overset{(b)}{=} \max_{(a,p)\in[0,1]^2}(1-\epsilon)\left[\frac{(1-\epsilon)pH_2(a)}{\epsilon+(1-\epsilon)p} + \frac{\epsilon(p+H_2(p))}{\epsilon+(1-\epsilon)p}\right]\\ &\overset{(c)}{=} \max_{p\in[0,1]}(1-\epsilon)\frac{p+\epsilon H_2(p)}{\epsilon+(1-\epsilon)p}, \end{aligned}$$

where (a) follows from $H(Y|X, S, Q) = H_2(\epsilon)$, (b) follows from (15) and $H(Y|X, S, Q) = H_2(\epsilon)$ and (c) follows from $H_2(a) \le 1$.

For the lower bound on feedback capacity, we simply take the maximizing distribution from the upper bound $p(X = 0|S = 0, Q = 1) = p(X = 1|S = 1, Q = 2) = 0.5$ and $p(X = 1|S = 0, Q = 3) = p(X = 0|S = 1, Q = 3) = p$ for some $p \in [0, 1]$ and show that the BCJR-invariant property is satisfied. This input distribution is an aperiodic input since the (S, Q)-graph has a loop with length 1. The stationary distribution that is given in (14) gives that $[\pi(S = 0|Q = 1), \pi(S = 0|Q = 2), \pi(S = 0|Q = 3)] = [1, 0, 0.5]$. The BCJR equation is then calculated in (12), as shown at the top the next page.

To show the BCJR-invariant property, it is convenient to treat each output observation separately. First, it is easy to note that all edges with $Y = -1$ or $Y = 1$ lead to $Q = 1$ and $Q = 2$, respectively, which approves the invariant property since $\pi(S = 0|Q = 1) = 1$ and $\pi(S = 0|Q = 2) = 0$. For

$$p(S=0|Q=g(q,y)) = \begin{cases} 1 & \text{if } y=-1, \\ 0 & \text{if } y=1, \\ \pi(S=0|Q=q)p(X=0|S=0,Q=q)+\pi(S=1|Q=q)p(X=1|S=0,Q=q) & \text{if } y=?, \\ \dfrac{\pi(S=0|Q=q)p(X=0|S=0,Q=q)}{\pi(S=0|Q=q)p(X=0|S=0,Q=q)+\pi(S=1|Q=q)p(X=1|S=1,Q=q)} & \text{if } y=0. \end{cases} \tag{12}$$

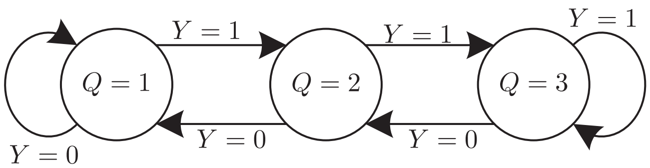

Fig. 7. Q-graph for the trapdoor channel.

the output $Y=?$, one can show that $\pi(S=0\|Q=i)p(X=0|S=0,Q=i)+\pi(S=1|q_i)p(X=1|S=0,Q=i)=0.5$, for $i=1,2,3$. For the output $Y=0$, the BCJR-invariant property can be established in a similar manner and this concludes that the input is BCJR-invariant. Since we used the Q-graph and the maximizer of the upper bound, there is no need to calculate the expression $I(X,S;Y|Q)$ since, obviously, it equals the upper bound. ■

D. Trapdoor Channel

The trapdoor channel was invented by Blackwell [19] in 1961. The capacity of this channel has been investigated in several papers and still remained an open problem. The channel has $\mathcal{S}=\mathcal{X}=\mathcal{Y}=\{0,1\}$. The output of the channel y_t is equal to s_{t-1} with probability p and equals x_t with probability $1-p$. Here, p is the channel parameter and can take any value in $[0,1]$. Finally, the channel state is $s_t=s_{t-1}\oplus x_t\oplus y_t$, where $\oplus$ is the XOR operation.

In [7], the feedback capacity for the trapdoor channel with parameter $p=0.5$ was shown to be $\log_2\phi$, where ϕ is the known golden ratio. The solution relied on a DP formulation of the problem and finding a solution to the Bellman equation. Here, we provide an alternative converse for $p=0.5$ that simplifies the proof in [7].

Applying Theorem 2 with the Q-graph in Fig. 7 gives the following.

Theorem 6 (Upper Bound): The feedback capacity of the trapdoor channel is bounded by

$$\begin{aligned} C_{\text{Trap}}(p) \le &\max_{(\alpha_1,\alpha_2,\alpha_3)\in[0,1]^3} \\ &2(\kappa_1+\kappa_2)H_2\left(\frac{\kappa_1(1-\alpha_1(1-p))+\kappa_2(1-p)\alpha_2}{\kappa_1+\kappa_2}\right) \\ &-2H_2(p)(\kappa_1\alpha_1+\kappa_2\alpha_2-0.5\alpha_3)+2\kappa_3, \end{aligned} \tag{16}$$

where

$$\begin{aligned} \delta &= 4\alpha_1 p-2\alpha_3+2 \\ &\quad +2(1-p)[\alpha_1-\alpha_2+\alpha_1\alpha_3-\alpha_1\alpha_2+\alpha_2\alpha_3] \\ \kappa_1 &= \frac{(1-\alpha_3)(1-\alpha_2(1-p))}{\delta} \\ \kappa_2 &= \frac{\alpha_1(p+\alpha_3(1-p))}{\delta} \\ \kappa_3 &= \frac{\alpha_1(1-\alpha_2(1-p))}{\delta}. \end{aligned}$$

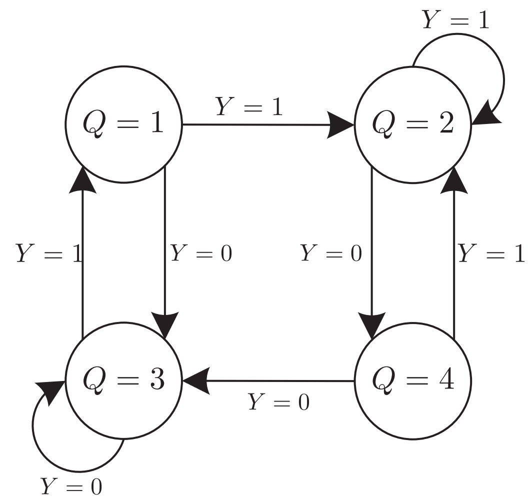

Fig. 8. Expanded Q-graph for the trapdoor channel.

The proof of Theorem 6 is omitted and follows by direct application of Theorem 2 with the Q-graph from Fig. 7. A special case of Theorem 6 is when $p=0.5$ and careful calculation gives

Corollary 1 (Upper Bound, p=0.5): The feedback capacity of the trapdoor channel with $p=0.5$ is bounded by

$$C_{\text{Trap}}(0.5)\le\log_2\phi. \tag{17}$$

Therefore, it follows that the upper bound from Theorem 6 is tight for $p=0.5$. Note that, at this point, the tightness of the upper bound follows from our previous knowledge of the feedback capacity in [7]. Next, we use Theorem 3 to show that $\log_2\phi$ is achievable not only for $p=0.5$ but for all $p\in[0,1]$.

Theorem 7 (Lower Bound): The feedback capacity of the trapdoor channel is bounded by

$$C_{\text{Trap}}(p)\ge\log_2\phi, \tag{18}$$

for all p.

Corollary 1 and Theorem 7 provide an alternative proof for the feedback capacity presented in [7]. The proofs of Corollary 1 and Theorem 7 appear in Appendix C and Appendix D, respectively.

Let us extend our realm of interest to a general parameter; numerical evaluation of Theorem 6 and a lower bound that is obtained from DP simulations give the plotted results in

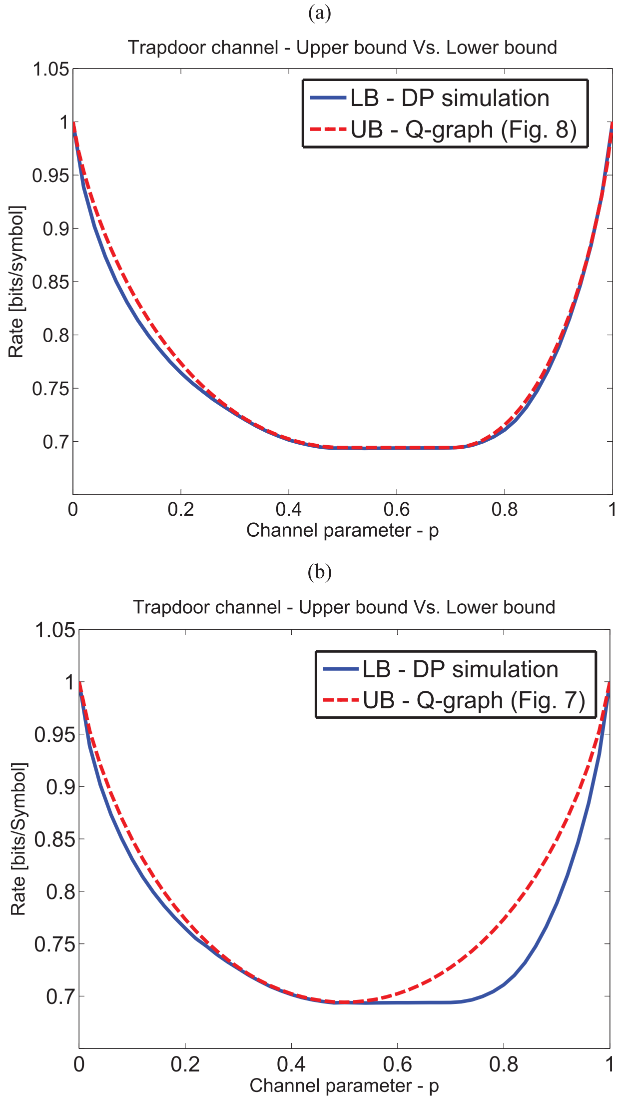

Fig. 9. A comparison between a lower bound (LB) on the feedback capacity that is achieved from DP simulation, and two upper bounds that were obtained from Theorem 2. In (a), the upper bound is calculated with the Q-graph (Fig. 7), while in (b) the upper bound is calculated with the expanded Q-graph (Fig. 8).

Fig. 9. Coarse inspection shows that the upper bound and the lower bound do not coincide in general, except for when $p = 0.5$. Now, an *expanded Q-graph* is introduced in Fig. 8 and is plotted in Fig. 9 with the same lower bound from DP simulations. It can be seen that the new upper bound shows a significant improvement in comparison with the upper bound in Fig. 9.

V. Proof of Theorem 2

An outline of the proof of Theorem 2 is given here and comprises three building blocks appearing in Lemmas 2 - 4. The first step expresses the essence of our bound and is encapsulated in the following lemma:

Lemma 2 (Step 1): For a strongly connected unifilar FSC, where s_0 is available to both the encoder and the decoder,

$$C_{\text{fb}} \leq \sup_{\{p(x_t|s_{t-1},q_{t-1})\}_{t\geq 1}} \liminf_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} I(X_i, S_{i-1}; Y_i|Q_{i-1}), \tag{19}$$

for all Q-graphs. The joint distribution is calculated with respect to $p(s_0,q_0)\prod_{i=1}^{N} p(x_i|s_{i-1},q_{i-1})p(y_i|x_i,s_{i-1})\mathbb{1}\{s_i = f(s_{i-1},x_i,y_i)\}\mathbb{1}\{q_i = g(q_{i-1},y_i)\}$.

The proof of Lemma 2 appears in Section V-A.

The upper bound in Lemma 2 is still difficult to compute since it is given by a limiting expression. The second step of the proof is tedious but necessary for our derivation, as we show that it is sufficient to restrict our maximization domain to the stationary inputs distribution. This step relies heavily on the DP formulation of the upper bound in (19); then, a known result from the literature is used to show the existence of an optimal stationary policy (equivalent to stationary inputs distribution). This second step is precisely outlined as follows:

Lemma 3 (Step 2): It is sufficient to maximize the upper bound in (19) *over stationary input distributions, i.e.,*

$$\begin{aligned}&\sup_{\{p(x_t|s_{t-1},q_{t-1})\}_{t\geq 1}} \liminf_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} I(X_i, S_{i-1}; Y_i|Q_{i-1})\\ &\quad= \sup_{p(x|s,q)} \liminf_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} I(X_i, S_{i-1}; Y_i|Q_{i-1}),\end{aligned} \tag{20}$$

for all irreducible Q-graphs where q_0 lies in an aperiodic closed class. The input distribution in the RHS of (20) *is $p(x|s,q)$ at all times.*

The proof of Lemma 3 appears in Section V-B.

Finally, the calculation of the upper bound with stationary inputs can be made; a minor restriction on the maximization domain verifies the existence of a stationary distribution on the (S,Q)-graph and, then,

Lemma 4 (Step 3):

$$\begin{aligned}&\sup_{p(x|s,q)} \liminf_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} I(X_i, S_{i-1}; Y_i|Q_{i-1})\\ &\quad\leq \sup_{p(x|s,q)\in\mathcal{P}_\pi} I(X,S;Y|Q),\end{aligned} \tag{21}$$

where $\mathcal{P}_\pi$ is defined in (3)*. If the supremum is attained with an aperiodic input then* (21) *holds with equality.*

The proof of Lemma 4 appears in Section V-C.

A. Proof of Lemma 2 (Step 1):

The proof comprises of the following steps:

$$\begin{aligned}C_{\text{fb}} &\overset{(a)}{=} \lim_{N\to\infty} \sup_{\{p(x_t|s_{t-1},y^{t-1})\}_{t=1}^{N}} \frac{1}{N}\sum_{i=1}^{N} I(X_i, S_{i-1}; Y_i|Y^{i-1})\\ &= \lim_{N\to\infty} \sup_{\{p(x_t|s_{t-1},y^{t-1})\}_{t=1}^{N}} \frac{1}{N}\sum_{i=1}^{N} \Big[H(Y_i|Y^{i-1})\\ &\quad - H(Y_i|X_i,S_{i-1})\Big]\\ &\overset{(b)}{\leq} \lim_{N\to\infty} \sup_{\{p(x_t|s_{t-1},y^{t-1})\}_{t=1}^{N}} \frac{1}{N}\sum_{i=1}^{N} I(X_i, S_{i-1}; Y_i|Q_{i-1})\\ &\overset{(c)}{\leq} \lim_{N\to\infty} \sup_{\{p(x_t|s_{t-1},q_{t-1})\}_{t=1}^{N}} \frac{1}{N}\sum_{i=1}^{N} I(X_i, S_{i-1}; Y_i|Q_{i-1})\end{aligned}$$

$$\stackrel{(d)}{=} \sup_{\{p(x_t|s_{t-1},q_{t-1})\}_{t\geq 1}} \liminf_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} I(X_i, S_{i-1}; Y_i|Q_{i-1}).$$

where

(a) follows from [7, Th. 1, eq. (18)];
(b) follows from the fact that conditioning reduces entropy and from the Markov chain of the channel $Y_i - (X_i, S_{i-1}) - Y^{i-1} - \Phi_{i-1}(Y^{i-1}) \triangleq Q_{i-1}$;
(c) follows from Lemma 5;
(d) follows from the arguments in [7, Lemma 4].

■

Lemma 5: Given (s_0, q_0)*, the maximization domain can be restricted as*

$$\sup_{\{p(x_t|s_{t-1},y^{t-1})\}_{t=1}^N} \sum_{i=1}^{N} I(X_i, S_{i-1}; Y_i|Q_{i-1}) = \sup_{\{p(x_t|s_{t-1},q_{t-1})\}_{t=1}^N} \sum_{i=1}^{N} I(X_i, S_{i-1}; Y_i|Q_{i-1}), \quad (22)$$

for all N.

Proof of Lemma 5: It is shown that the same objective is achieved when exchanging the domain $\{p(x_t|s_{t-1}, y^{t-1})\}_{t=1}^N$ with the domain $\{p(x_t|s_{t-1}, q_{t-1})\}_{t=1}^N$; the second domain is calculated as the marginal distribution of $\{p(x_t, s_{t-1}, y^{t-1})\}_{t=1}^N$ that is induced by the first domain. To this end, we show that two distributions $\{p_1(x_t|s_{t-1}, y^{t-1})\}_{t=1}^N$ and $\{p_2(x_t|s_{t-1}, y^{t-1})\}_{t=1}^N$ with the same induced marginal distribution $\{\tilde{p}(x_t|s_{t-1}, q_{t-1})\}_{t=1}^N$ have the same objective. The objective is determined by $\{p(y_t, x_t, s_{t-1}, q_{t-1})\}_{t=1}^N$ since the mutual information at each time is a function of one instance from this set.

This is shown using induction: for $n = 1$, write $p(x_1, y_1, s_0, q_0) = p(y_1|x_1, s_0)\tilde{p}(x_1|s_0, q_0)p(s_0, q_0)$, indicating that reward depends on the marginal $\tilde{p}$ only. Assume that $\{p(x_t, y_t, s_{t-1}, q_{t-1})\}_{t=1}^N$ is induced by both input distributions and, thus, induce the same N-th objective. Let us show that $p(x_{N+1}, y_{N+1}, s_N, q_N)$ depends on the marginal $\tilde{p}$ only. First, note that $p(s_N, q_N)$ is determined by the N-th step since q_N is a function of (q_{N-1}, y_N) and s_N is a function of (x_N, y_N, s_{N-1}). Furthermore, $p(x_{N+1}|s_N, q_N)$ is identical under both input distributions and $p(y_{N+1}|s_N, x_{N+1}, q_N)$ is given by the channel specification. Thus, $p(x_{N+1}, y_{N+1}, s_N, q_N)$ is equal under both input distributions. ■

B. Proof of Lemma 3 (Step 2):

In this section, the goal is to show that stationary inputs are optimal for the upper bound derived in Lemma 2. The first stage is to formulate the upper bound as a DP problem. We then present a known result from the DP literature [20] that states sufficient conditions for the existence of an optimal stationary policy. Finally, it is proved that these conditions are satisfied in our DP problem and, thus, the existence of an optimal stationary policy is established.

1) DP Formulation: The DP definitions presented here follow the formulation in [21]; similar formulations can also be found in [6], [7], and [14].

Define the DP state at time t (prior to the t-th action) as the probability vector $z_{t-1} \triangleq p(s_{t-1}, q_{t-1})$ which takes values in $(|\mathcal{S}| \times |\mathcal{Q}| - 1)$-dimensional simplex; throughout this section we denote the state space by $\mathcal{Z}$. As the initial state, (s_0, q_0), lies in a closed communicating class, A, the state space is taken as the $(|A| - 1)$-dimensional unit simplex. Actions are valid conditional distributions $p(x|s, q)$ and, specifically, the action at time t is $u_t \triangleq p(x_t|s_{t-1}, q_{t-1})$. The reward gained at time t is taken to be $I(X_t, S_{t-1}; Y_t|Q_{t-1})$. Note that this is a deterministic DP as no disturbance is defined.

To show that the above definitions hold for the DP properties, we must verify that there exists a deterministic next-state function and that the reward at time t is a function of (z_{t-1}, u_t):

Dynamics: We show that there exists a deterministic next-state function, denoted by $F(\cdot)$, such that $z_t = F(z_{t-1}, u_t)$. Each state, z_t, is a collection of the probabilities $p(s_t, q_t)$, and can be calculated as follows:

$$\begin{aligned}
p(s_t, q_t) &= \sum_{y_t,x_t,s_{t-1},q_{t-1}} p(s_t, q_t, y_t, x_t, s_{t-1}, q_{t-1}) \\
&= \sum_{y_t,x_t,s_{t-1},q_{t-1}} p(s_t, q_t|x_t, y_t, s_{t-1}, q_{t-1})p(y_t|x_t, s_{t-1}) \\
&\quad \times p(x_t|s_{t-1}, q_{t-1})p(s_{t-1}, q_{t-1}) \\
&\stackrel{(a)}{=} \sum_{y_t,x_t,s_{t-1},q_{t-1}} \mathbb{1}\{s_t = f(x_t, y_t, s_{t-1})\}\mathbb{1}\{q_t = g(y_t, q_{t-1})\} \\
&\quad \times p(y_t|x_t, s_{t-1})p(x_t|s_{t-1}, q_{t-1})p(s_{t-1}, q_{t-1}), \qquad (23)
\end{aligned}$$

where step (a) follows from the facts that the state in a unifilar channel is a function of the triplet (x_t, y_t, s_{t-1}), and the Q-graph function $g : \mathcal{Q} \times \mathcal{Y} \to \mathcal{Q}$. Recall that z_{t-1} consists of all entries $p(s_{t-1}, q_{t-1})$ and the actions are $u_t = p(x_t|s_{t-1}, q_{t-1})$; therefore, each entry in z_t is a time-invariant function of the pair (z_{t-1}, u_t).

Reward: Let us show that each reward is a function of the current state and the chosen actions. The reward at time t is $I(X_t, S_{t-1}; Y_t|Q_{t-1})$ and is a function of $p(y_t, x_t, s_{t-1}, q_{t-1})$, which can be written as $p(y_t|x_t, s_{t-1})p(x_t|s_{t-1}, q_{t-1})p(s_{t-1}, q_{t-1})$. The latter factorization of the joint distribution is a function of the state $z_{t-1} = p(s_{t-1}, q_{t-1})$, the action $u_t = p(x_t|s_{t-1}, q_{t-1})$ and the channel $p(y_t|x_t, s_{t-1})$. From now on, we use the notation $R(z, u)$ for the mutual information that is achieved with a state z and action u.

The DP formulation above implies that the optimal average reward is

$$\rho^* = \sup_{\omega^\infty} \liminf_{N\to\infty} \frac{1}{N}\sum_{t=1}^{N} I(X_t, S_{t-1}; Y_t|Q_{t-1}),$$

where ω^∞ corresponds to a policy, i.e., an infinite sequence of actions. Note that ρ^* is equal to the upper bound in (19), so this is an equivalent DP problem for the upper bound calculation.

In addition, we define for $\beta < 1$ and initial state $\xi \in \mathcal{Z}$ their optimal discounted reward as

$$\nu_\beta(\xi) = \sup_{\omega^\infty} \sum_{t=1}^{\infty} \beta^t I(X_t, S_{t-1}; Y_t | Q_{t-1}).$$

2) Sufficient Conditions for the Existence of an Optimal Stationary Policy:

(C1) The transition kernel is continuous with respect to weak convergence in $P(\mathcal{Z})$. In our case, the transition kernel is defined by the function $F(\cdot, \cdot)$ in (23).

(C2) The state space, $\mathcal{Z}$, is a locally compact space with a countable base.

(C3) The multifunction $\mathcal{U}(z)$ is upper semi-continuous. The notation $\mathcal{U}(z)$ stands for allowed actions at state z. In our case, $\mathcal{U}(z)$ is the set of all conditional distributions of the form $p(x|s, q)$, i.e., the set of allowed actions equals the full set of actions for all z.

(C4) The reward function $R(z, u)$ is lower semi-continuous in (z, u).

(C5) Let $m_\beta = \sup_z \nu_\beta(z)$; then $\sup_{\beta<1}\{m_\beta - \nu_\beta(z)\} < \infty$ for all $z \in \mathcal{Z}$.

In [20], the above conditions were presented for a model where the optimal average reward is defined as the minimization over all policies. Since our model is defined as a maximization problem, trivial modifications should be made in **(C3)-(C4)**; however, we will show that in our problem these conditions are satisfied in both the upper and lower cases.

*Theorem 8 ([20, Th. 3.8]): If **(C1)-(C5)** are satisfied then there exists an optimal stationary policy for the average reward DP problem.*

Returning to our problem, we will show that **(C1)-(C5)** are satisfied and this leads to the conclusion that there exists an optimal stationary policy.

*Conditions **(C1)-(C5)** are satisfied in our problem:*

(C1) The transition kernel is continuous with respect to weak convergence if the following is satisfied: for all $v(\cdot) \in \mathcal{C}_b(\mathcal{Z})$(continuous and bounded functions on $\mathcal{Z}$),

$$\int_{\mathcal{Z}} v(y) F(dy|\cdot, \cdot) \in \mathcal{C}_b(\mathcal{Z} \times \mathcal{U}). \tag{24}$$

The transition kernel, $F(dy|z, u)$, is a dirac measure and, therefore, integration over y (24) returns $v(F(z, u))$.

The function $v(F(z, u))$ is bounded since $v(\cdot)$ is bounded. For the continuity, note that by (23) each element in $F(z, u)$ is continuous with respect to (z, u) (in any norm) since it is a finite sum of elements in (z, u). Since the function $v(\cdot)$ is continuous, the composition of $F(z, u)$ into $v(\cdot)$ is continuous in (z, u). To conclude, the composition $v(F(z, u))$ is bounded and continuous with respect to (z, u).

(C2) The state space is the $(|A| - 1)$-dimensional unit simplex. As the $(|A| - 1)$ dimensional simplex is a closed subset of the $|A|$-dimensional unit cube, it is locally compact with a countable base.

(C3) The general scenario is where the action space can depend on z; however, in our problem $\mathcal{U}(z)$ is constant in z and, thus, trivially continuous in z.

(C4) The mutual information can be written as a sum of entropies, where each entropy is continuous in the joint distribution of (y, x, s, q) that is induced by (z, u); therefore, it is both lower and upper semi-continuous.

(C5) By [20, Proposition 2.1], conditions (C1)-(C4) imply that there exists an optimal stationary policy for the discounted problem, which is denoted here as $f_\beta = p^*_\beta(x|s, q)$. The policy f_β implies a structure on the (S, Q)-graph and might result in several closed communicating classes in the case where there are edges with probability zero. Denote by $\mathcal{A}$ the (S, Q)-graph after removing edges with $p^*_\beta(x|s, q) = 0$. It is convenient to partition the analysis for two cases based on the structure of $\mathcal{A}$:

Case A: The graph induced by the policy f_β, $\mathcal{A}$, has a single closed communicating class.

Case B: The graph induced by the policy f_β, $\mathcal{A}$, has more than one closed communicating class.

Case A: With some abuse of terminology, we will refer to $\mathcal{A}$ as the closed communicating class in the (S, Q)-graph, since all nodes outside this class are inessential in the infinite-horizon regime. Denote by T the transfer matrix induced by f_β on the single closed class, and by D its period. Since $\mathcal{A}$ is irreducible, the graph can be partitioned into $A_0, A_1, \ldots A_{D-1}$ disjoint classes on a cycle based on a period equivalence. The stationary distribution of the Markov chain on $\mathcal{A}$ is denoted by π_{f_β}.

Consider the D-blocks Markov chain and, specifically, a Markov chain with transition matrix T^D. Since D is the period of the original graph, the new Markov chain implies D separate aperiodic and irreducible Markov chains. Each Markov chain is on a class A_d and we denote by $\pi(A_d)$ the stationary distribution of each class A_d where $d \in [0 : D - 1]$.

For initial state $\xi \in \mathcal{P}(Z)$, denote its weights vector as $W(\xi)$ with d inputs, where the d-th input is $w_d(\xi) = \sum_{(s,q)\in A_d} \xi(s, q)$. Define for all k:

$$\pi_k(\xi) \triangleq [w_{[k]}(\xi)\pi(A_0), w_{[k+1]}(\xi)\pi(A_1), \ldots, w_{[k+D-1]}(\xi)\pi(A_{D-1})], \tag{25}$$

where the indices with $[\cdot]$ are taken modulo D. Finally, the vectors in (25) are used to define

$$\nu^\pi_\beta(\xi) = \sum_{n=1}^{\infty} \beta^n R(\pi_n(\xi), f_\beta)$$
$$\nu^*_\beta = \sum_{n=1}^{\infty} \beta^n R(\pi_{f_\beta}, f_\beta).$$

The expression $\nu^\pi_\beta(\xi)$ corresponds to the discounted reward that is achieved with states that are moved periodically through all possibilities in (25). The first step is to show that for a fixed initial state, the actual reward and its corresponding periodic reward, $\nu^\pi_\beta(\xi)$ are bounded as follows:

Lemma 6: For all initial states, ξ,

$$\sup_\beta |\nu_\beta(\xi) - \nu^\pi_\beta(\xi)| < \infty.$$

The second step of the proof is to show that the achieved periodic reward (which is a function of the initial state) is bounded with respect to some constant quantity, which does not depend on ξ:

Lemma 7: For all initial states, ξ,

$$\sup_{\beta} |\nu_{\beta}^{*} - \nu_{\beta}^{\pi}(\xi)| < \infty.$$

A direct conclusion from the above two lemmas is the required condition **(C5)**:

$$\begin{aligned}
&\sup_{\beta} |\nu_{\beta}(\xi) - \nu_{\beta}(\xi')| \\
&\stackrel{(a)}{\leq} 2 \sup_{\beta} \max_{\xi} |\nu_{\beta}(\xi) - \nu_{\beta}^{*}| \\
&\stackrel{(a)}{\leq} 2 \sup_{\beta} \max_{\xi} \left(|\nu_{\beta}(\xi) - \nu_{\beta}^{\pi}(\xi)| + |\nu_{\beta}^{\pi}(\xi) - \nu_{\beta}^{*}| \right) \\
&\stackrel{(b)}{<} \infty,
\end{aligned}$$

where (a) follows from the triangle inequality and (b) follows from Lemma 6 and Lemma 7.

The proofs of Lemma 6 and Lemma 7 appear in Appendix E and Appendix F, respectively.

Their proof requires the following preliminaries on total-variation distance and Markov chains.

Definition 1: For finite alphabet, $\mathcal{X}$, and two PMFs, p_X and q_X, the total-variation distance is

$$||p_X - q_X||_{TV} \triangleq \tfrac{1}{2} \sum_x |p(x) - q(x)|.$$

The following Lemma summarizes two properties of the total-variation distance:

Lemma 8 ([22, Lemma V.I–V.II]): For two joint PMFs, $p_{X,Y}$ and $q_{X,Y}$ on a finite alphabet $\mathcal{X} \times \mathcal{Y}$, their total variation distance satisfies

$$\|p_X - q_X\|_{TV} \leq \|p_{X,Y} - q_{X,Y}\|_{TV},$$

and the equality holds if $p(y|x) = q(y|x)$.

The following is an upper bound on the convergence rate of aperiodic Markov chains.

Lemma 9 ([23, Th. 4.9]): Let T be a transfer matrix of an irreducible and aperiodic Markov chain on a finite space $\mathcal{X}$ with a stationary distribution π. Then there exist constants $\alpha \in (0, 1)$ and $C > 0$ such that

$$\max_{\xi \in \mathcal{P}(X)} \|\xi T^n - \pi\|_{TV} \leq C\alpha^n.$$

The Markov chain in our problem is not necessarily aperiodic; therefore, a slight adaptation of Lemma 9 for the periodic case is now given.

Lemma 10 (Convergence of Periodic Markov Chains): Let T be a transfer matrix of an irreducible Markov chain with period D on a space $\mathcal{X}$. Then there exist constants $\alpha \in (0, 1)$ and $C > 0$ such that for all ξ

$$\|\xi T^{nD+k} - \pi_k(\xi)\|_{TV} \leq C\alpha^n,$$

where $\pi_k(\xi)$ are defined in (25)*.*

Proof of Lemma 10: For some $k \in [0 : D - 1]$ and for all ξ, consider

$$\begin{aligned}
&||\xi T^{nD+k} - \pi_k(\xi)||_{TV} \\
&= \frac{1}{2} \sum_{(s,q)} |\xi T^{nD+k}(s, q) - \pi_k(\xi)(s, q)| \\
&= \sum_{d} \frac{1}{2} \sum_{(s,q)\in A_d} |\xi T^{nD+k}(s, q) - \pi_k(\xi)(s, q)| \\
&\stackrel{(a)}{=} \sum_{d} \frac{1}{2} \sum_{(s,q)\in A_d} |\xi T^{nD+k}(s, q) - w_{[k+d]}(\xi)\pi(A_d)(s, q)| \\
&= \sum_{d} w_{[k+d]}(\xi) \frac{1}{2} \sum_{(s,q)\in A_d} \left| \frac{\xi T^{nD+k}(s, q)}{w_{[k+d]}(\xi)} - \pi(A_d)(s, q) \right| \\
&\stackrel{(b)}{=} \sum_{d} w_{[k+d]}(\xi) \left|\left| \frac{\xi T^{nD+k}}{w_{[k+d]}(\xi)} - \pi(A_d) \right|\right|_{TV} \\
&\stackrel{(c)}{\leq} \sum_{d} w_{[k+d]}(\xi) C_d \alpha_d^n \\
&\leq \sum_{d} w_{[k+d]}(\xi) \max_{d} C_d \alpha_d^n \\
&\stackrel{(d)}{\triangleq} C\alpha^n,
\end{aligned}$$

where (a) follows by substituting Eq. (25), (b) follows by the total-variation distance definition when conditioned on the class A_d, (c) follows from Lemma 9 and (d) follows by $\sum_d w_{[k+d]}(\xi) = 1$. ■

Case B: We give an outline of the proof for case B, as it essentially follows the same steps used for Case A. In this scenario, there are several closed communicating classes, denoted by $A_1, \ldots, A_k$, and their corresponding periods are $D_1, \ldots, D_k$. The technique used in Case A is composed of two steps: the first is to show that the reward is bounded with a reward that has some periodic behavior, as argued in Lemma 6, and the second step is to show that this periodic reward is bounded with respect to some constant quantity (with respect to the initial state).

The first step is addressed by studying the periodic behavior of each closed class, as was done in Case A. Clearly, the initial weight of each closed class is time-invariant since weight cannot move between closed classes. It follows that the common period of all classes is simply the multiplication of all periods, i.e., $D = \prod_i D_i$. This concludes the analysis that is required for the first part of the proof. The second part of the proof follows the lines used for the proof of Lemma 7; specifically, the upper bound derivation can be followed with the defined D, and the ϵ-policy construction is identical. ■

C. Proof of Lemma 4 (Step 3)

Before presenting the proof of Lemma 4, we impose a restriction on the maximization domain:

Lemma 11: It is sufficient to take the supremum in (21) *over $p(x|s, q)$ which lies in $\mathcal{P}_{\pi}$.*

Proof of Lemma 11: In this proof, we will take the maximizer of the LHS in (21), and show that there exists a distribution from $\mathcal{P}_{\pi}$ that induces the maximal reward.

Let $p^*(x|s,q)$ be a maximizer which implies two closed communicating classes, A_1 and A_2, with average rewards, R_1 and R_2, respectively. Construct $\tilde{p}(x|s,q)$ exactly as $p^*(x|s,q)$, but where positive probabilities are given for edges from A_1 to A_2. This modification is legitimate since the (S,Q)-graph is irreducible. This modification did not effect the rewards for initial states in A_2, while the rewards of initial states in A_1 might be changed.

By the optimality of $p^*(x|s,q)$, the reward R_1 cannot be increased and, therefore, $R_1 = R_2$. Since $R_1 = R_2$, the policy induced by $\tilde{p}(x|s,q) \in \mathcal{P}_\pi$ achieves the same optimal rewards as the maximizer. The above argument can be extended to any number of closed communicating classes since the graph is finite. ■

By the definition of the set $\mathcal{P}_\pi$, there is a single closed communicating class for each input distribution $p(x|s,q)$. Let A denote the graph which describes the closed class and let T_A be its transition probability matrix. Since A is irreducible, there exists a stationary distribution $\pi = [\pi_1, \pi_2, \dots, \pi_{|A|}]$ which is the unique solution for the equation $\pi T_A = \pi$. Here, the stationary distribution is in the Cesàro sum sense, i.e.,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{m=1}^{n} T_A^m = \begin{pmatrix} \pi_1 & \pi_2 & \dots & \pi_{|A|} \\ \vdots & & & \vdots \\ \pi_1 & \pi_2 & \dots & \pi_{|A|} \end{pmatrix}.$$

Let $|\mathcal{D}|$ be the period of the graph A, and let $A_1, \dots A_{|\mathcal{D}|}$ be the disjoint subsets of nodes based on a period equivalence. The dependence of A, $|\mathcal{D}|$ and A_i on $p(x|s,q)$ is omitted.

Proof of Lemma 4: Consider the following chain of equalities:

$$\begin{aligned}
&\sup_{p(x|s,q)} \liminf_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} I(X_i, S_{i-1}; Y_i|Q_{i-1}) \\
&\overset{(a)}{=} \sup_{p(x|s,q)\in\mathcal{P}_\pi} \liminf_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} I(X_i, S_{i-1}; Y_i|Q_{i-1}) \\
&\overset{(b)}{\leq} \sup_{p(x|s,q)\in\mathcal{P}_\pi} \liminf_{N\to\infty} \frac{1}{N|\mathcal{D}|}\sum_{i=1}^{N|\mathcal{D}|} I(X_i, S_{i-1}; Y_i|Q_{i-1}) \\
&\overset{(c)}{=} \sup_{p(x|s,q)\in\mathcal{P}_\pi} \liminf_{N\to\infty} \frac{1}{N|\mathcal{D}|}\sum_{d=1}^{|\mathcal{D}|}\sum_{i=0}^{N-1} \\
&\qquad I(X_{i|\mathcal{D}|+d}, S_{i|\mathcal{D}|+d-1}; Y_{i|\mathcal{D}|+d}|Q_{i|\mathcal{D}|+d-1}) \\
&\overset{(d)}{=} \sup_{p(x|s,q)\in\mathcal{P}_\pi} \frac{1}{|\mathcal{D}|}\sum_{d=1}^{|\mathcal{D}|} \liminf_{N\to\infty} \frac{1}{N}\sum_{i=0}^{N-1} \\
&\qquad I(X_{i|\mathcal{D}|+d}, S_{i|\mathcal{D}|+d-1}; Y_{i|\mathcal{D}|+d}|Q_{i|\mathcal{D}|+d-1}) \\
&\overset{(e)}{=} \sup_{p(x|s,q)\in\mathcal{P}_\pi} \frac{1}{|\mathcal{D}|}\sum_{d\in\mathcal{D}} I(X, S_d; Y|Q_d) \\
&\overset{(f)}{=} \sup_{p(x|s,q)\in\mathcal{P}_\pi} I(X,S;Y|Q,D) \\
&= \sup_{p(x|s,q)} H(Y|Q,D) - H(Y|X,S,Q,D) \\
&\overset{(g)}{\leq} \sup_{p(x|s,q)\in\mathcal{P}_\pi} H(Y|Q) - H(Y|X,S,Q) \\
&= \sup_{p(x|s,q)\in\mathcal{P}_\pi} I(Y;X,S|Q),
\end{aligned} \tag{26}$$

where

(a) follows from Lemma 11;
(b) follows by taking the limit on a subsequence of N, i.e., the sequence $|\mathcal{D}|, 2|\mathcal{D}|, \dots$;
(c) follows by re-indexing the summation in blocks of $\mathcal{D}$ elements;
(d) follows by exchanging the order of the limit and sum due to the limit existence of each term in the sum;
(e) follows by calculating the limit for a fixed d. Specifically, the value of d determines a class A_d. The distribution of $p(s_{i|\mathcal{D}|+d-1}, q_{i|\mathcal{D}|+d-1})$ tends to the stationary distribution since the chain is aperiodic and irreducible. This, in turn, gives that the distribution for each d is $p_d(y,x,s,q) = p(y,x|s,q)\pi_d(s,q)$, where

$$\pi_d(s,q) = \begin{cases} \dfrac{\pi(s,q)}{\sum_{(s,q)\in A_d}\pi(s,q)} & \text{if } (s,q)\in A_d; \\ 0 & \text{otherwise.}\end{cases} \tag{27}$$

(f) follows by defining a uniform RV, D, on $[1 : |\mathcal{D}|]$. The joint distribution is $p(y,x,s,q,d) = p(y,x|s,q)\pi_d(s,q)p(d)$;
(g) follows from the Markov $Y - (X,S) - D$ and the fact that conditioning reduces entropy. This expression is calculated with respect to $p(s,q,x,y)$, which is the marginal distribution of $p(d)\pi_d(s,q)p(x|s,q)p(y|x,s)$. Explicit calculation gives that

$$\begin{aligned}
p(s,q) &= \sum_d p(d)\pi_d(s,q) \\
&\overset{(*)}{=} \frac{1}{|\mathcal{D}|}\frac{\pi(s,q)}{\sum_{(s,q)\in A_d}\pi(s,q)} \\
&\overset{(**)}{=} \frac{1}{|\mathcal{D}|}\frac{\pi(s,q)}{\frac{1}{|\mathcal{D}|}} \\
&= \pi(s,q),
\end{aligned}$$

where $(*)$ is obtained by substituting the expression from (27) and $(**)$ follows from $\sum_{(s,q)\in A_d}\pi(s,q) = \frac{1}{|\mathcal{D}|}$, for all d, since each class is on a cycle.

To conclude the proof, we have shown in (26) that $I(X,S;Y|Q)$ with $\pi(s,q)p(x|s,q)p(y|x,s)$ is an upper bound. ■

VI. Conclusions

An upper bound on the feedback capacity of unifilar FSCs was derived. The upper bound is expressed as a computable single-letter expression and it was shown how the bound can be computed for known capacity results. Calculation of the upper bound for the DEC resulted a new capacity result together with the sufficient condition for the optimality of the upper bound. For all studied channels, the optimal Q-graph was obtained from DP simulations. A further direction that is under investigation is a structured method for finding such an optimal Q-graph without DP simulations.

The upper bound gives a useful insight into the structure of optimal output processes. Specifically, as the bound is tight for all known capacities of unifilar FSCs, this provides a unified structure for the optimal output processes. The technique used in this paper might also be applied to any entropy rate of a random process. Specifically, for the n-th instance, $H(Y_n|Y^{n-1})$, the process history can be quantized using a Q-graph. However, even for FSCs without feedback, the obtained upper bound is not computable since the contexts are not revealed to the encoder.

APPENDIX A
PROOF OF LEMMA 1

Each node has an outgoing edge since for each s there exists (x, y) such that $p(y|x, s) > 0$. Therefore, each node (s, q) has an outgoing edge $(s, q) \to (g(q, y), f(s, x, y))$. It should be clear that each node has at least one outgoing edge to another node since, if a node (s, q) has edges to itself only, this means that for all (x, y), $s = f(x, y, s)$, implying $|\mathcal{S}| = 1$. Therefore, by the pigeonhole principle, there exists at least one closed communicating class.

To show that each closed class has all $q \in \mathcal{Q}$, recall that the Q-graph is irreducible and, therefore, for each pair (q_0, q_n), there exists a path $q_0 \to q_n$ labelled by $y_1 \dots y_n$ such that $q_i = g(y_i, q_{i-1})$. For the first label, y_1, there exists (x_1, s_1) such that $p(y_1|x_1, s_1) > 0$. Then, for a node (s, q_1) in the closed class there is an edge to $(f(y_1, x_1, s_1), q_2)$; this argument can be repeated until a node of the form $(\cdot, q_n)$ is reached. Since it is a closed communicating class, each path leads to a node in this class.

The proof that each closed class has all $s \in S$ is similar to the previous argument, but using the strongly connected property of states, i.e., that the states graph is irreducible. For each s, s', there exists a path labelled by $(x_1, y_1), (x_2, y_2), \dots$ with probabilities $p(y_i|x_i, s_{i-1}) > 0$ such that s reaches s'. Therefore, for each (s, q) there is a path to $(s', \cdot)$ for all s'. ■

APPENDIX B
PROOF OF THEOREM 3

In this proof, we show that BCJR-invariant inputs induce the Markov chain $Y_t - Q_{t-1} - Y^{t-1}$ for all t. This Markov chain gives, in turn, that the feedback capacity expression with the chosen input is $I(X, S; Y|Q)$.

Since inputs $p(x|s, q)$ are assumed to be aperiodic inputs, it may be assumed that the initial distribution is $\pi(s, q)$ since it is reached with high probability. We will show by induction that the value of the BCJR estimator is determined by a context of sequence, i.e., $p(S_t = s|Y^t = y^t) = \pi(S = s|Q = \Phi(y^t))$. At time $t - 1$, assume that $p(S_{t-1} = s|Y^{t-1} = y^{t-1}) = \pi_{S=s|Q=q}$ for all $s \in \mathcal{S}$, where $q \triangleq \Phi(y^{t-1})$. Then, one can calculate at time t,

$$\begin{aligned} p(S_t = s|Y^t = y^t) &\overset{(a)}{=} B_s(\pi(S|Q = q, y_t)) \\ &\overset{(b)}{=} \pi(S = s|Q = g(q, y_t)), \end{aligned}$$

where (a) follows from the induction hypothesis and the forward-recursive relation, (5), and (b) follows from the BCJR-invariant property.

From the induction proof above, we have that $p(S_{t-1} = s|y^{t-1}) = \pi(S = s|Q = q)$, so we can show the Markov chain $Y_t - Q_{t-1} - Y^{t-1}$:

$$\begin{aligned} &p(y_t|y^{t-1}, q_{t-1}) \\ &= \sum_{s_{t-1}, x_t} p(y_t, x_t, s_{t-1}|y^{t-1}, q_{t-1}) \\ &= \sum_{s_{t-1}, x_t} p(y_t|x_t, s_{t-1}) p(x_t|s_{t-1}, q_{t-1}, y^{t-1}) \\ &\quad \times p(s_{t-1}|y^{t-1}, q_{t-1}) \\ &\overset{(a)}{=} \sum_{s_{t-1}, x_t} p(y_t|x_t, s_{t-1}) p(x_t|s_{t-1}, q_{t-1}) \pi(s_{t-1}|q_{t-1}) \\ &= p(y_t|q_{t-1}), \end{aligned} \tag{28}$$

where (a) follows from the fact that the input x_t depends on the pair (s_{t-1}, q_{t-1}) only, and the from the above inductive argument which shows that $p(s_{t-1}|y^{t-1}, q_{t-1}) = \pi(s_{t-1}|q_{t-1})$.

Finally, the theorem can be proved by the following chain of inequalities

$$\begin{aligned} C_{\text{fb}} &\overset{(a)}{=} \sup_{\{p(x_t|s_{t-1}, y^{t-1})\}_{t\geq 1}} \liminf_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} I(X_i, S_{i-1}; Y_i|Y^{i-1}) \\ &\overset{(b)}{=} \sup_{\{p(x_t|s_{t-1}, y^{t-1})\}_{t\geq 1}} \liminf_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \\ &\quad I(X_i, S_{i-1}; Y_i|Q_{i-1}) - I(Y_i; Y^{i-1}|Q_{i-1}) \\ &\overset{(c)}{\geq} \liminf_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} I(X_i, S_{i-1}; Y_i|Q_{i-1}) - I(Y_i; Y^{i-1}|Q_{i-1}) \\ &\overset{(d)}{=} \liminf_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} I(X_i, S_{i-1}; Y_i|Q_{i-1}) \\ &\overset{(e)}{=} I(X, S; Y|Q), \end{aligned}$$

where

(a) follows from the capacity formula from Theorem 1;
(b) follows from re-writing $I(X_i, S_{i-1}; Y_i|Y^{i-1}) = H(Y_i|Y^{i-1}) - H(Y_i|X_i, S_{i-1})$ and adding $H(Y_i|Q_{i-1}) - H(Y_i|Q_{i-1})$;
(c) follows by taking the input distribution to be $p(x_t|s_{t-1}, y^{t-1}) = p(x|s, q)$ for all t;
(d) follows from the Markov chain $Y_i - Q_{i-1} - Y^{i-1}$ in (28);
(e) follows from the aperiodic Markov chain on the state space $\mathcal{S} \times \mathcal{Q}$) which induces its corresponding stationary distribution.

■

APPENDIX C
PROOF OF COROLLARY 1

In this section we show that $C_{\text{Trap}}(0.5) \leq \log_2 \phi$. The upper bound on the capacity of the trapdoor channel with $p = 0.5$ from Theorem 6 is:

$$\begin{aligned} C_{\text{Trap}}(0.5) \leq \max_{(\alpha_1, \alpha_2, \alpha_3) \in [0,1]^3} & \lambda_1(\alpha_1, \alpha_2, \alpha_3) \\ & + \lambda_2(\alpha_1, \alpha_2, \alpha_3), \end{aligned}$$

where

$$\lambda_1(\alpha_1,\alpha_2,\alpha_3) = 2(\kappa_1+\kappa_2)H_2\left(\frac{\kappa_1(1-0.5\alpha_1)+0.5\kappa_2\alpha_2}{\kappa_1+\kappa_2}\right)$$
$$\lambda_2(\alpha_1,\alpha_2,\alpha_3) = 2(\kappa_3-\kappa_1\alpha_1-\kappa_2\alpha_2-0.5\alpha_3)$$

and

$$\begin{aligned}\delta &= 3\alpha_1-\alpha_2+\alpha_1\alpha_3-\alpha_1\alpha_2+\alpha_2\alpha_3-2\alpha_3+2\\ \kappa_1 &= \frac{(1-\alpha_3)(1-0.5\alpha_2)}{\delta}\\ \kappa_2 &= \frac{0.5\alpha_1(1+\alpha_3)}{\delta}\\ \kappa_3 &= \frac{\alpha_1(1-0.5\alpha_2)}{\delta}.\end{aligned}$$

The proof will follow from the facts that $\lambda_2(\cdot)\le 0$ and $\lambda_1(\cdot)\le\log\phi$. Let us begin with $\lambda_2(\cdot)$ that is equal to

$$\frac{\alpha_2\alpha_3-\alpha_1\alpha_2-\alpha_1\alpha_3-2\alpha_3-\alpha_1\alpha_3^2-\alpha_2\alpha_3^2+2\alpha_3^2-\alpha_1\alpha_2\alpha_3}{\delta}. \tag{30}$$

Since $\delta>0$, it is sufficient to verify that the numerator is always negative; to this end, we write the numerator of (30) as a polynomial of α_3 when α_1,α_2 are some parameters:

$$\alpha_3^2(2-\alpha_1-\alpha_2)+\alpha_3(-2-\alpha_1+\alpha_2-\alpha_1\alpha_2)-\alpha_1\alpha_2.$$

The coefficient of α_3^2 is always positive and, therefore, it is a convex function. It can also be noted that the function is negative at both boundaries, $\alpha_3=0$ and $\alpha_3=1$, thus, for $\alpha_3\in[0,1]$ the function $\lambda_2(\cdot)$ is upper bounded with zero.

Let us provide an upper bound for $\lambda_1(\alpha_1,\alpha_2,\alpha_3)$:

$$\begin{aligned}\lambda_1(\alpha_1,\alpha_2,\alpha_3) &= 2(\kappa_1+\kappa_2)H_2\left(\frac{\kappa_1(1-0.5\alpha_1)+\kappa_2 0.5\alpha_2}{\kappa_1+\kappa_2}\right)\\ &\overset{(a)}{=} 2(\kappa_1+\kappa_2)H_2\left(\frac{\kappa_1 0.5\alpha_1+\kappa_2(1-0.5\alpha_2)}{\kappa_1+\kappa_2}\right)\\ &\overset{(b)}{=} 2(\kappa_1+\kappa_2)H_2\left(\frac{\kappa_3}{\kappa_1+\kappa_2}\right)\\ &\overset{(c)}{=} 2(\kappa_1+\kappa_2)H_2\left(\frac{0.5-(\kappa_1+\kappa_2)}{\kappa_1+\kappa_2}\right)\\ &\overset{(d)}{=} 2\frac{1}{2(p+1)}H_2(p)\\ &\overset{(e)}{\le} \max_{0\le p\le 1}\frac{H_2(p)}{1+p}\\ &= \log_2\phi,\end{aligned}$$

where

(a) follows from the symmetry of the binary entropy function, i.e., $H_2(p)=H_2(1-p)$;
(b) follows from $\kappa_1 0.5\alpha_1+\kappa_2(1-0.5\alpha_2)=\kappa_3$;
(c) follows from $\kappa_1+\kappa_2+\kappa_3=0.5$;
(d) follows by defining a new variable $p(\alpha_1,\alpha_2,\alpha_3)=\frac{0.5-(\kappa_1+\kappa_2)}{\kappa_1+\kappa_2}$;
(e) follows by taking the maximum over p, which is obviously restricted to $[0,1]$.

Finally, we can show that

$$\begin{aligned}C_{\text{Trap}}(p) &\le \max_{(\alpha_1,\alpha_2,\alpha_3)\in[0,1]^3}\lambda_1(\cdot)+\lambda_2(\cdot)\\ &\le \max_{(\alpha_1,\alpha_2,\alpha_3)\in[0,1]^3}\lambda_1(\cdot)+\max_{(\alpha_1,\alpha_2,\alpha_3)\in[0,1]^3}\lambda_2(\cdot)\\ &\overset{(a)}{\le} \log_2\phi,\end{aligned}$$

where (a) follows from the derived upper bounds on each function separately. ∎

APPENDIX D
PROOF OF THEOREM 7

The proof is based on Theorem 3 with the Q-graph from Fig. 8 and the following input distribution:

$$\begin{aligned}p(X=0|S=0,Q=1) &= 1\\ p(X=0|S=0,Q=2) &= 1\\ p(X=0|S=0,Q=3) &= \frac{zp}{1-(1-p)z}\\ p(X=0|S=0,Q=4) &= \frac{zp}{1-(1-p)z}\\ p(X=1|S=1,Q=1) &= \frac{zp}{1-(1-p)z}\\ p(X=1|S=1,Q=2) &= \frac{zp}{1-(1-p)z}\\ p(X=1|S=1,Q=3) &= 1\\ p(X=1|S=1,Q=4) &= 1,\end{aligned}$$

where z is a parameter in $[0,1]$ and p is the channel parameter. Straightforward calculation gives that $[\pi(S=0|Q=1), \pi(S=0|Q=2), \pi(S=0|Q=3), \pi(S=0|Q=4)] = [(1-p)z, 1-z, z, 1-(1-p)z]$.

The BCJR equation can be written as (29), as shown at the bottom of this page. where $\delta_i=\pi(S=0|Q=i)p(X=0|S=0,Q=i)$ and $\gamma_i=\pi(S=1|Q=i)p(X=1|S=1,Q=i)$. The explicit calculation of the BCJR-invariant property is omitted here as it is identical to the calculations for the DEC and the input-constrained BEC.

For simplicity, we denote $\alpha\triangleq\frac{zp}{1-(1-p)z}$ which can take any value on $[0,1]$, and then we have the stationary vector of the Q-graph:

$$\begin{aligned}&[\pi(Q=1),\pi(Q=2),\pi(Q=3),\pi(Q=4)]\\ &=\left[\frac{1-\alpha}{4-2\alpha},\frac{1}{4-2\alpha},\frac{1}{4-2\alpha},\frac{1-\alpha}{4-2\alpha}\right],\end{aligned}$$

$$p(S=0|Q=g(i,y)) = \begin{cases}\dfrac{\delta_i}{(1-p)(\delta_i-\gamma_i)+p\pi(S=0|Q=i)+(1-p)\pi(S=1|Q=i)} & \text{if } y=0,\\ \dfrac{(1-p)(\pi(S=0|Q=i)-\delta_i)+p(\pi(S=1|Q=i)-\gamma_i)}{(1-p)(\pi(S=0|Q=i)-\delta_i)+\pi(S=1|Q=i)+(1-p)(\gamma_i-\pi(S=1|Q=i))} & \text{if } y=1,\end{cases} \tag{29}$$

and the per node rewards:

$$I(X,S;Y|Q=i) = zH_2(p)$$
$$I(X,S;Y|Q=j) = H_2(\alpha) - (1-\alpha)zH_2(p),$$

for $i = 1,4$ and $j = 2,3$.

Then, the lower bound can be computed:

$$\begin{aligned} C_{\text{Trap}}(p) &\geq I(X,S;Y|Q) \\ &= 2\frac{(1-\alpha)zH_2(p)}{4-2\alpha} + 2\frac{H_2(\alpha)-(1-\alpha)zH_2(p)}{4-2\alpha} \\ &= \frac{H_2(\alpha)}{2-\alpha}. \end{aligned}$$

By taking a maximum over α, we obtain that the capacity is lower bounded with $\log\phi$. ∎

APPENDIX E
PROOF OF LEMMA 6

Before presenting the proof, we recall a known upper bound on the difference between two entropies for different PMFs is presented.

Lemma 12 ([24, Th. 3]): For two joint PMFs, p_X and q_X, on a finite set $\mathcal{X}$,

$$|H_p(X) - H_q(X)| \leq \|p_X - q_X\|_{TV}\log(|\mathcal{X}|-1) + H_2(\|p_X - q_X\|_{TV}).$$

Proof of Lemma 6: For initial state ξ, the real distribution on $\mathcal{S}\times\mathcal{Q}$ at time $nD+k$ is ξT^{nD+k}, while $\pi_k(\xi)$ is the distribution that was defined in (25). Recall that the distribution $p(y,x|s,q) = p(y|x,s)p(x|s,q)$ is determined by the policy and the channel. With some loss of accuracy, the dependence on the initial state ξ might be omitted and the derivations hold for all ξ.

Consider the rewards difference at time $nD+k$:

$$\begin{aligned} &|R(\xi T^{nD+k}, f_\beta) - R(\pi_k(\xi), f_\beta)| \quad (31)\\ &= |\, I_{T^{nD+k}}(X,S;Y|Q) - I_{\pi^k}(X,S;Y|Q)| \\ &\overset{(a)}{\leq} |H_{T^{nD+k}}(Y|Q) - H_{\pi^k}(Y|Q)| \\ &\quad + |H_{T^{nD+k}}(Y|X,S,Q) - H_{\pi^k}(Y|X,S,Q)|, \quad (32) \end{aligned}$$

where (a) follows by the triangle inequality. The first term in (31) can be bounded by

$$\begin{aligned} &|H_{T^{nD+k}}(Y|Q) - H_{\pi^k}(Y|Q)| \quad (33)\\ &\overset{(a)}{\leq} |\mathcal{Q}| \max_q |H_{T^{nD+k}}(Y|Q=q) - H_{\pi^k}(Y|Q=q)| \\ &\overset{(b)}{\leq} |\mathcal{Q}| \max_q \Big\{ ||T^{nD+k}_{Y|Q=q} - \pi^k_{Y|Q=q}||_{TV}\log(|\mathcal{Y}|-1) \\ &\quad + H_2(||T^{nD+k}_{Y|Q=q} - \pi^k_{Y|Q=q}||_{TV})\Big\}, \quad (34) \end{aligned}$$

where (a) follows by the triangle inequality and (b) follows from Lemma 12.

Consider for all $q \in \mathcal{Q}$

$$\begin{aligned} ||T^{nD+k}_{Y|Q=q} - \pi^k_{Y|Q=q}||_{TV} &\overset{(a)}{\leq} ||\xi T^{nD+k}_{Y,Q} - \pi^k_{Y,Q}(\xi)||_{TV} \\ &\overset{(b)}{\leq} ||T^{nD+k}_{S,Q,X,Y} - \pi^k_{S,Q,X,Y}||_{TV} \\ &\overset{(b)}{=} ||T^{nD+k}_{S,Q} - \pi^k_{S,Q}||_{TV} \\ &\overset{(c)}{\leq} C\alpha^n, \end{aligned}$$

where (a) follows by adding terms to the sum of total-variation, (b) follows from Lemma 8 and (c) follows from Lemma 10. Since $C\alpha^n \to 0$, there exists some N' for which $||T^{nD+k}_{S,Q} - \pi^k_{S,Q}||_{TV} \leq 0.5$ for all $n > N'$.

Therefore, (33) can be bounded for all $n > N'$ as follows:

$$\begin{aligned} &|H_{T^{nD+k}}(Y|Q) - H_{\pi^k}(Y|Q)| \\ &\overset{(a)}{\leq} |\mathcal{Q}|\Big\{||T^{nD+k}_{S,Q} - \pi^k_{S,Q}||_{TV}\log(|\mathcal{Y}|-1) \\ &\qquad + H_2(||T^{nD+k}_{S,Q} - \pi^k_{S,Q}||_{TV})\Big\} \\ &\leq |\mathcal{Q}|\left\{C\alpha^n\log(|\mathcal{Y}|-1) + H_2(C\alpha^n)\right\}, \end{aligned}$$

where (a) follows from (33) and Lemma 8. The same derivation can be repeated for the second term in (31) resulting in the same convergence rate. To summarize, there exist some constants $C' > 0$ and $\alpha \in (0,1)$ such that

$$|R(\xi T^{nD+k}, f_\beta) - R(\pi_k(\xi), f_\beta)| \leq C'\alpha^n + 2H_2(C\alpha^n), \quad (35)$$

for all $n > N'$.

For all ξ, consider

$$\begin{aligned} &|\nu_\beta(\xi) - \nu^\pi_\beta(\xi)| \\ &= \left|\sum_{n=1}^{\infty}\beta^{nD}\sum_{k=0}^{D-1}\beta^k\left[R(\xi T^{nD+k}, f_\beta) - R(\pi_k(\xi), f_\beta)\right]\right| \\ &\overset{(a)}{\leq} \sum_{n=1}^{\infty}\beta^{nD}\sum_{k=0}^{D-1}\beta^k|R(\xi T^{nD+k}, f_\beta) - R(\pi_k(\xi), f_\beta)| \\ &\overset{(a)}{\leq} N'D\log|\mathcal{Y}| + \sum_{n=N'+1}^{\infty}\beta^{nD} \\ &\quad \times \sum_{k=0}^{D-1}\beta^k|R(\xi T^{nD+k}, f_\beta) - R(\pi_k(\xi), f_\beta)| \\ &\overset{(b)}{\leq} N'D\log|\mathcal{Y}| + \sum_{n=N'+1}^{\infty}\beta^{nD}\sum_{k=0}^{D-1}\beta^k C'\alpha^n + 2H_2(C\alpha^n) \\ &\overset{(c)}{\leq} N'D\log|\mathcal{Y}| + \sum_{n=1}^{\infty}C'\alpha^n + 2H_2(C\alpha^n), \quad (36) \end{aligned}$$

where (a) follows by the triangle inequality, (b) follows from (35) and (c) follows from $\beta \leq 1$. Finally, by verifying that $\sum_{n=1}^{\infty}H_2(C\alpha^n) < \infty$ and by taking the supremum on both sides of (36) we have that $\sup_\beta|\nu_\beta(\xi) - \nu^\pi_\beta(\xi)| < \infty$. ∎

APPENDIX F
PROOF OF LEMMA 7

The proof of Lemma 7 comprises two main steps. First, we derive an upper bound on $\nu^\pi_\beta(\xi)$ which does not depend on the initial state ξ. Secondly, we construct a new policy that can achieve a reward that is arbitrarily close to the provided upper bound. From the optimality of f_β, the two steps taken imply that all $\nu_\beta^\pi(\xi)$ are, indeed, close "enough" to the upper bound.

Proof of Lemma 7: Let us derive an upper bound on the average of D consecutive rewards for some initial state ξ:

$$\begin{aligned}\frac{1}{D}\sum_{k=0}^{D-1}\beta^k R(\pi_k(\xi), f_\beta) &\overset{(a)}{\leq} \frac{1}{D}\sum_{k=0}^{D-1} R(\pi_k(\xi), f_\beta)\\ &\overset{(b)}{=} \frac{1}{D}\sum_{k=0}^{D-1} I(X, S_k(\xi); Y|Q_k(\xi))\\ &\overset{(c)}{=} I(X, S; Y|Q, K(\xi))\\ &\overset{(d)}{\leq} I_\xi(X, S; Y|Q)\\ &\overset{(e)}{=} R(\pi_{f_\beta}, f_\beta), \end{aligned} \tag{37}$$

where (a) follows from $\beta \leq 1$, (b) follows from the notation $p(s_k(\xi), q_k(\xi)) = \pi_k(\xi)$, (c) follows by defining a uniform RV, K, on $[0 : D-1]$ and (d) follows from the fact that conditioning reduces entropy and from the Markov chain $Y-(X,S)-K(\xi)$, where the subscript ξ is added to emphasize the dependence on the initial state. Finally, step (e) shows that the marginal distribution does not depend on ξ; the marginal distribution of $p(s,q)$ for some $(s,q) \in A_i$ is

$$\begin{aligned}p(s,q) &= \sum_k p(k,s,q)\\ &= \frac{1}{D}\sum_k p(s_k(\xi), q_k(\xi))\\ &\overset{(a)}{=} \frac{1}{D}\sum_k w_k(\xi)\pi(A_i)(s,q)\\ &\overset{(b)}{=} \frac{1}{D}\pi(A_i)(s,q)\\ &\overset{(c)}{=} \pi_{f_\beta}(s,q),\end{aligned}$$

where (a) follows from (25), (b) follows from $\sum_k w_k(\xi) = 1$ and the notation $\pi(A_i)(s,q)$ as the stationary distribution of the state (s,q). Finally, (c) follows from the property that in a periodic Markov chain each class has a uniform distribution.

The derivation above is used to provide an upper bound on $\nu_\beta^\pi(\xi)$:

$$\begin{aligned}\nu_\beta^\pi(\xi) &= \sum_{n=1}^{\infty}\beta^{nD}\sum_{k=0}^{D-1}\beta^k R(\pi_k(\xi), f_\beta)\\ &= \sum_{n=1}^{\infty}\beta^{nD} D\frac{1}{D}\sum_{k=0}^{D-1}\beta^k R(\pi_k(\xi), f_\beta)\\ &\overset{(a)}{\leq} \sum_{n=1}^{\infty}\beta^{nD} D R(\pi_{f_\beta}, f_\beta)\\ &= \frac{DR(\pi_{f_\beta}, f_\beta)}{1-\beta^D}\\ &\overset{(b)}{=} \nu_\beta^* + R(\pi_{f_\beta}, f_\beta)\frac{D-(1+\beta+\cdots+\beta^{D-1})}{1-\beta^D}\\ &\overset{(c)}{=} \nu_\beta^* + K_\beta,\end{aligned} \tag{38}$$

where (a) follows from (37), (b) follows from the fact that $\nu_\beta^* = \frac{R(\pi_{f_\beta}, f_\beta)}{1-\beta}$ and (c) is just a notation K_β; by using L'Hopital's rule it can be noted that $\sup_\beta K_\beta < \infty$.

A new stationary policy, $f_\beta(\epsilon)$, is constructed by taking the policy f_β and letting a path be with $\epsilon > 0$ weights, so that the resultant graph is aperiodic. This modification is possible due to the aperiodicity assumption in Theorem 2. Moreover, ϵ is chosen to be small enough such that a node with modified outgoing edges still has positive probabilities for all other outgoing edges. The stationary distribution of this modified policy is denoted by $\pi(\epsilon)$, satisfying $\pi(0) = \pi_{f_\beta}$. The reward gained by the policy $f_\beta(\epsilon)$ is denoted by $\nu_\beta^\epsilon(\xi)$.

For $\epsilon \geq 0$, the stationary distribution exists and is unique, since it is a solution of linear equations. The stationary distribution is continuous with respect to ϵ since each entry in this vector is a rational function of ϵ and, clearly, $\epsilon = 0$ is not a pole. We also know that mutual information is continuous w.r. to $\pi(\epsilon)$ and $f_\beta(\epsilon)$ and, therefore, the composition $I_{\pi(\epsilon)} \triangleq R(\pi(\epsilon), f_\beta(\epsilon))$ is continuous with respect to the parameter ϵ.

By repeating the arguments in Lemma 6 with Lemma 9 on the convergence rate of aperiodic Markov chains, it can be deduced that

$$\sup_\beta |\nu_\beta^\epsilon(\xi) - \nu_\beta^\epsilon(\xi')| < \infty, \tag{39}$$

for all ξ, ξ'. Note that (39) holds for all states and, specifically, for $\xi' = \pi(\epsilon)$.

For a fixed β, the continuity of each instantaneous reward in ϵ_β assures that there exists ϵ_β^* such that the difference between $|I_{\pi(0)} - I_{\pi(\epsilon_\beta)}| < 1-\beta$ for all $\epsilon_\beta < \epsilon_\beta^*$. By combining this continuity and (39), we have

$$\sup_\beta |\nu_\beta^{\epsilon_\beta^*}(\xi) - \nu^*| < \infty, \tag{40}$$

for all ξ.

The reward $\nu_\beta^{\epsilon_\beta^*}(\xi)$ is achievable for initial state ξ by the following trivial policy: for some initial state ξ use policy $f_\beta(\epsilon)$ and, otherwise, use f_β. Clearly, the constructed policy does not change rewards for initial states other than ξ, and the optimality of f_β gives that $\nu_\beta^{\epsilon_\beta^*}(\xi) \leq \nu_\beta(\xi)$.

For all initial states,

$$\begin{aligned}\nu_\beta^\pi(\xi) &\overset{(a)}{\leq} \nu_\beta^* + K_\beta\\ &\overset{(b)}{=} \nu_\beta^{\epsilon_\beta^*}(\xi) + K_{\beta,\epsilon_\beta^*}(\xi) + K_\beta\\ &\overset{(c)}{\leq} \nu_\beta(\xi) + K_{\beta,\epsilon_\beta^*}(\xi) + K_\beta\\ &\overset{(d)}{=} \nu_\beta^\pi(\xi) + K'_\beta(\xi) + K_{\beta,\epsilon_\beta^*}(\xi) + K_\beta,\end{aligned} \tag{41}$$

where (a) follows from inequality (38), (b) follows from (40) and the notation $K_{\beta,\epsilon_\beta^*}(\xi)$ for the difference between the rewards, (c) follows by the optimaility of the policy f_β and (d) follows from Lemma 6 and the notation $\nu_\beta(\xi) = \nu_\beta^\pi(\xi) + K'_\beta(\xi)$, where $\sup_\beta |K'_{\beta,\epsilon}(\xi)| < \infty$.

By subtracting $\nu_\beta^\pi(\xi)$ from (41), one can conclude that for all β

$$|\nu_\beta^* - \nu_\beta^\pi(\xi)| \leq \max\{|K'_\beta(\xi) + K_{\beta,\epsilon_\beta^*}(\xi)|, K_\beta\}. \tag{42}$$

By taking a supremum on both sides of (42), the proof of Lemma 7 is concluded. ■

ACKNOWLEDGMENT

The authors would like to thank the Associate Editor and the anonymous reviewer, for their valuable and constructive comments, which helped to improve this paper.

REFERENCES

[1] P. Sadeghi, R. A. Kennedy, P. B. Rapajic, and R. Shams, "Finite-state Markov modeling of fading channels—A survey of principles and applications," *IEEE Signal Process. Mag.*, vol. 25, no. 5, pp. 57–80, Sep. 2008.
[2] Q. Zhang and S. A. Kassam, "Finite-state Markov model for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 47, no. 11, pp. 1688–1692, Nov. 1999.
[3] K. A. Schouhamer Immink, P. H. Siegel, and J. K. Wolf, "Codes for digital recorders," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2260–2299, Oct. 1998.
[4] H. H. Permuter, T. Weissman, and A. J. Goldsmith, "Finite state channels with time-invariant deterministic feedback," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 644–662, Feb. 2009.
[5] R. Dabora and A. J. Goldsmith, "On the capacity of indecomposable finite-state channels with feedback," *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 193–203, Jan. 2013.
[6] S. Tatikonda and S. Mitter, "The capacity of channels with feedback," *IEEE Trans. Inf. Theory*, vol. 55, no. 1, pp. 323–349, Jan. 2009.
[7] H. Permuter, P. Cuff, B. Van Roy, and T. Weissman, "Capacity of the trapdoor channel with feedback," *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 3150–3165, Jul. 2009.
[8] S. C. Tatikonda, "Control under communication constraints," Ph.D. dissertation, Dept. Electr. Eng. Comput. Sci., Massachusetts Inst. Technol., Cambridge, MA, USA, 2000.
[9] J. Chen and T. Berger, "The capacity of finite-state Markov channels with feedback," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 780–789, Mar. 2005.
[10] S. Yang, A. Kavčić, and S. Tatikonda, "Feedback capacity of finite-state machine channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 799–810, Mar. 2005.
[11] S. Yang, A. Kavčić, and S. Tatikonda, "On the feedback capacity of power-constrained Gaussian noise channels with memory," *IEEE Trans. Inf. Theory*, vol. 53, no. 3, pp. 929–954, Mar. 2007.
[12] O. Elishco and H. Permuter, "Capacity and coding for the Ising channel with feedback," *IEEE Trans. Inf. Theory*, vol. 60, no. 9, pp. 5138–5149, Sep. 2014.
[13] A. Sharov and R. Roth, "On the capacity of generalized ising channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2015, pp. 2256–2260.
[14] O. Sabag, H. H. Permuter, and N. Kashyap, "The feedback capacity of the binary erasure channel with a no-consecutive-ones input constraint," *IEEE Trans. Inf. Theory*, vol. 62, no. 1, pp. 8–22, Jan. 2016.
[15] O. Sabag, H. H. Permuter, and N. Kashyap, "The feedback capacity of the binary symmetric channel with a no-consecutive-ones input constraint," in *Proc. 53rd Annu. Allerton Conf. Commun., Control, Comput.*, Sep. 2015, pp. 160–164.
[16] H. D. Pfister and P. H. Siegel, "Joint iterative decoding of LDPC codes for channels with memory and erasure noise," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 2, pp. 320–337, Feb. 2008.
[17] C. Wang and H. D. Pfister, "Upper bounds on the MAP threshold of iterative decoding systems with erasure noise," in *Proc. 5th Int. Symp. Turbo Codes Related Topics*, Sep. 2008, pp. 7–12.
[18] H. D. Pfister, "On the capacity of finite state channels and the analysis of convolutional accumulate-*m* codes," Ph.D. dissertation, Dept. Elect. Comput. Eng., Univ. California, San Diego, CA, USA, 2003.
[19] D. Blackwell, *Information Theory. Modern Mathematics for the Engineer: Second Series*, 1961, pp. 182–193.
[20] M. Schäl, "Average optimality in dynamic programming with general state space," *Mathematics of Operations Research*, vol. 18, no. 1, pp. 163–172, 1993.
[21] A. Arapostathis, V. S. Borkar, E. Fernández-Gaucherand, M. K. Ghosh, and S. I. Marcus, "Discrete-time controlled Markov processes with average cost criterion: A survey," *SIAM J. Control Optim.*, vol. 31, no. 2, pp. 282–344, 1993.
[22] P. Cuff, "Distributed channel synthesis," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7071–7096, Nov. 2013.
[23] D. A. Levin, E. Wilmer, and E. L. Wilmer, *Markov Chains and Mixing Times*. Providence, RI, USA: American Mathematical Society, 2009.
[24] I. Sason, "Entropy bounds for discrete random variables via maximal coupling," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7118–7131, Nov. 2013.

Oron Sabag (S'14) received the B.Sc. (*cum laude*) degree in electrical and computer engineering from the Ben-Gurion University of the Negev, Israel, in 2013, where he is currently pursuing the Ph.D. degree in electrical and computer engineering. He is a recipient of the Lachish Fellowship for honor students in the direct Ph.D. Program and Excellence Scholarship from Electrical and Computer Engineering Department, Ben-Gurion University of the Negev.

Haim H. Permuter (M'08–SM'13) received the B.Sc. (*summa cum laude*) and M.Sc. (*summa cum laude*) degrees in electrical and computer engineering from the Ben-Gurion University of the Negev, Israel, in 1997 and 2003, respectively, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, USA, in 2008.

From 1997 to 2004, he was an Officer with Research and Development Unit, Israeli Defense Forces. Since 2009, he has been with the Department of Electrical and Computer Engineering, Ben-Gurion University of the Negev, where he is currently an Associate Professor.

Prof. Permuter is a recipient of several awards, such as the Fullbright Fellowship, the Stanford Graduate Fellowship, the Allon Fellowship, and the U.S.-Israel Binational Science Foundation Bergmann Memorial Award. He is currently serving on the editorial boards of the IEEE TRANSACTIONS ON INFORMATION THEORY.

Henry D. Pfister (S'99–M'03–SM'09) received the Ph.D. degree in electrical engineering from the University of California at San Diego, San Diego, CA, USA, in 2003. He was a Professor with Texas A&M University (2006–2014), a Post-Doctoral Fellow with the École Polytechnique Fédérale de Lausanne (2005–2006), and a Senior Engineer with Qualcomm Corporate R&D in San Diego (2003–2004). He is currently an Associate Professor with Electrical and Computer Engineering Department, Duke University.

He received the NSF Career Award in 2008 and the Texas A&M ECE Department Outstanding Professor Award in 2010. He is a co-author of the 2007 IEEE COMSOC Best Paper in Signal Processing and Coding for Data Storage and a co-author of the 2016 Symposium on the Theory of Computing Best Paper. He is currently an Associate Editor for the Coding Theory of the IEEE TRANSACTIONS ON INFORMATION THEORY (2013–2016) and a Distinguished Lecturer of the IEEE Information Theory Society (2015–2016).