A Deterministic Algorithm for the Capacity of Finite-State Channels

Chengyu Wu, Guangyue Han, *Senior Member, IEEE*, Venkat Anantharam, *Fellow, IEEE*, and Brian Marcus, *Life Fellow, IEEE*

***Abstract*— We propose two modified versions of the classical gradient ascent method to compute the capacity of finite-state channels with Markovian inputs. For the case that the channel mutual information rate is strongly concave in a parameter taking values in a compact convex subset of some Euclidean space, our first algorithm proves to achieve polynomial accuracy in polynomial time and, moreover, for some special families of finite-state channels our algorithm can achieve exponential accuracy in polynomial time under some technical conditions. For the case that the channel mutual information rate may not be strongly concave, our second algorithm proves to be at least locally convergent.**

***Index Terms*— Channel capacity, finite-state channels, gradient ascent, hidden Markov processes.**

I. Introduction

As opposed to a discrete memoryless channel (DMC), which can be characterized by the conditional distribution of the output given the input, in a finite-state channel (FSC) this conditional distribution depends on an underlying state variable which evolves with time. Encompassing DMCs as special cases, FSCs have long been used in a wide range of communication scenarios where the current behavior of the channel may be affected by its past. Among many others, conventional examples of FSCs include inter-symbol interference channels [10], partial response channels [28], [31] and Gilbert-Elliott channels [24].

Manuscript received August 1, 2020; revised July 24, 2021; accepted November 16, 2021. Date of publication December 3, 2021; date of current version February 17, 2022. The work of Guangyue Han was supported in part by the Research Grants Council of the Hong Kong Special Administrative Region, China, under Project 17301017 and Project 17304121; and in part by the National Natural Science Foundation of China, under Project 61871343. The work of Venkat Anantharam was supported in part by NSF under Grant CNS-1527846, Grant CCF-1618145, Grant CCF-1901004, and Grant CIF-2007965; in part by the NSF Science & Technology Center (Science of Information) under Grant CCF-0939370; and in part by the William and Flora Hewlett Foundation supported Center for Long Term Cybersecurity at Berkeley. An earlier version of this paper was presented at 2019 International Symposium on Information Theory [DOI: 10.1109/ISIT.2019.8849433]. *(Corresponding author: Chengyu Wu.)*

Chengyu Wu and Guangyue Han are with the Department of Mathematics, The University of Hong Kong, Hong Kong (e-mail: chengyuw@connect.hku.hk; ghan@hku.hk).

Venkat Anantharam is with the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, Berkeley, CA 94720 USA (e-mail: ananth@berkeley.edu).

Brian Marcus is with the Department of Mathematics, The University of British Columbia, Vancouver, BC V6T 1Z2, Canada (e-mail: marcus@math.ubc.ca).

Communicated by M. Dalai, Associate Editor for Shannon Theory.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TIT.2021.3132675.

Digital Object Identifier 10.1109/TIT.2021.3132675

While it is well-known that the Blahut-Arimoto algorithm [2], [4] can be used to efficiently compute the capacity of a DMC, the computation of the capacity of a general FSC has long been a notoriously difficult problem and it has been open for decades. The difficulty of this problem may be justified by the widely held (yet not proven) belief that typically the capacity of an FSC may not be achieved by any finite-order Markovian input, and an increase of the memory of the input may lead to an increase of the channel capacity. For recent works studying the computability of the capacity of FSCs, see [9] and [5].

We are mainly concerned with FSCs with Markov processes of a fixed order as their inputs. Possibly an unavoidable compromise we have to make in exchange for progress in computing the capacity, the extra fixed-order assumption imposed on the input process is also necessary for the situation where the channel input has to satisfy certain constraints, notably finite-type constraints [22] that are commonly used in magnetic and optical recording [18], [23], [30]. On the other hand, the focus on Markovian inputs can also be justified by the known fact that the Shannon capacity of an indecomposable FSC [11] can be approximated by the Markov capacity with increasing orders (see Theorem 2.1 of [21]).

Recently, there has been some progress in computing the capacity of FSCs with such input constraints. Below we only list the most relevant work in the literature, and we refer the reader to [14] for a comprehensive list of references. In [19], the Blahut-Arimoto algorithm was reformulated into a stochastic expectation-maximization procedure and a similar algorithm for computing a lower bound on the capacity of FSCs was proposed, which led to a generalized Blahut-Arimoto algorithm [32] that proves to compute the capacity under some concavity assumptions. More recently, inspired by ideas in stochastic approximation, a randomized algorithm was proposed in [14] to compute the capacity under weaker concavity assumptions, which can be verified to hold true for several families of practical channels [16], [20].

Both of the above-mentioned algorithms, however, are of a randomized nature (a feasible implementation of the generalized Blahut-Arimoto algorithm will necessitate a randomization procedure). By comparison, our algorithms, which are deterministic in nature, can be used to derive accurate estimates on the channel capacity, as evidenced by the tight bounds in Section III-B.

In this paper, we first deal with the case that the mutual information rate of the FSC is strongly concave in a parameter taking values in a compact convex subset of some Euclidean space, for which we propose our first algorithm that proves to converge to the channel capacity exponentially fast. This algorithm largely follows the spirit of the classical gradient ascent method. However, unlike the classical case, the lack of an explicit expression for our target function and the boundedness of the variable domain (without an explicit description of the boundary) pose additional challenges. To overcome the first issue, a convergent sequence of approximating functions (to the original target function) is used instead in our treatment; meanwhile, an additional check condition is also added to ensure that the iterates stay inside the given variable domain. A careful convergence analysis has been carried out to deal with the difficulties caused by such modifications. This algorithm is efficient in the sense that, for a general FSC (satisfying the above-mentioned concavity condition and some additional technical conditions), it achieves polynomial accuracy in polynomial time (see Theorem III.12), and for some special families of FSCs it achieves exponential accuracy in polynomial time (see Section III-B).

It is well known that the mutual information rate of an FSC may not be concave under the natural parametrization in several examples; see, e.g., [16], [20]. Another modification of the classical gradient ascent method is proposed to handle this challenging scenario. Similar to our first algorithm, our second one replaces the original target function with a sequence of approximating functions, which unfortunately renders conventional methods such as the Frank-Wolfe method (see, e.g., [3]) or methods using the Łojasiewicz inequality (see, e.g., [1]) inapplicable. To address this issue, among other subtle modifications, we impose an extra check in the algorithm to slow down the pace "a bit" to avoid a premature convergence to a non-stationary point but "not too much" to ensure the local convergence.

As variants of the classical gradient ascent method, our algorithms can be applied to any sequence of convergent functions, so they can be of particular interest in information theory since many information-theoretic quantities are defined as the limit of their finite-block versions. We would also like to add that our algorithms are actually stated in a much more general setting and may have potential applications in optimization scenarios where the target functions are difficult to compute but amenable to approximations.

The remainder of this paper is organized as follows. In Section II, we describe our channel model in greater detail. Then, we present our first algorithm (Algorithm III.3) in Section III and analyze its convergence behavior in Section III-A under some strong concavity assumptions. Applications of this algorithm for computing the capacity of FSCs under concavity assumptions will be discussed in Section III-B. In particular, in this section, we show that the estimation of the channel capacity can be improved by increasing the Markov order of the input process in some examples. In Section IV, our second algorithm (Algorithm IV.2) is presented, which proves to be at least locally convergent. Finally, in Section IV-B, our second algorithm is applied to two FSCs where the concavity of the channel mutual information rates in the natural parametrization are not known, and yet fast convergence behaviors are observed.

In the remainder of this paper, the base of the logarithm is assumed to be e.

II. Channel Model and Problem Formulation

In this section, we introduce the channel model considered in this paper, which is essentially the same as that in [14], [32].

As mentioned before, we are concerned with a discrete-time FSC with a Markovian channel input. Let $X = \{X_n : n = 1, 2, \dots\}$ denote the channel input process, which is often assumed to be a first-order stationary Markov chain[1] over a finite alphabet $\mathcal{X}$, and let $Y = \{Y_n : n = 1, 2, \dots\}$ and $S = \{S_n : n = 0, 1, \dots\}$ denote the channel output and state processes over finite alphabets $\mathcal{Y}$ and $\mathcal{S}$, respectively.

Let Π be the set of all the stochastic matrices of dimension $|\mathcal{X}| \times |\mathcal{X}|$, where $|\mathcal{X}|$ denotes the cardinality of $\mathcal{X}$. For any finite set $F \subseteq \mathcal{X}^2$ and any $\delta > 0$, define

$$\Pi_{F,\delta} \triangleq \{A \in \Pi : A_{ij} = 0, \text{ for } (i, j) \in F \text{ and } A_{ij} \geq \delta \text{ otherwise}\}.$$

It can be easily verified that if one of the matrices from $\Pi_{F,\delta}$ is primitive, then all matrices from $\Pi_{F,\delta}$ will be primitive, in which case, as elaborated on in [14], F gives rise to a so-called mixing finite-type constraint. Such a constraint has been widely used in data storage and magnetic recording [23], [30], the best known example being the so-called (d, k)-run length limited ((d, k)-RLL) constraint over the alphabet $\{0, 1\}$, which forbids any sequence with fewer than d or more than k consecutive zeros in between two successive 1's.

The following conditions will be imposed on the FSC described above:

(II.a) The channel is stationary and characterized by

$$p(y_n|x_1^N, s_0^N, y_1^{n-1}) = p(y_n|x_n, s_{n-1})$$

for any $1 \leq n \leq N$, where $p(y_n|x_n, s_{n-1}) > 0$ for any positive integer n and any x_n, s_{n-1}, y_n.

(II.b) $\{X_n, S_{n-1}\}_{n=1}^{\infty}$ is a first-order stationary Markov chain whose transition probabilities satisfy

$$p(x_{n+1}, s_n|x_n, s_{n-1}) = p(x_{n+1}|x_n)p(s_n|x_n, s_{n-1}),$$

for any positive integer n, where $p(s_n|x_n, s_{n-1}) > 0$ for any s_{n-1}, s_n, x_n.

(II.c) The input process X is a first-order stationary Markov chain, and there exist a set $F \subseteq \mathcal{X}^2$ and $\delta > 0$ such that the transition probability matrix of X belongs to $\Pi_{F,\delta}$, each element of which is a primitive matrix.

It follows from Theorem 4.6.3 of [11] that an FSC specified as above is indecomposable. Therefore, assuming that the input X (or, more precisely, the transition probability matrix of X) is analytically parameterized by a finite-dimensional parameter θ in the interior of a compact convex subset Θ of some Euclidean

[1] The assumption that X is a first-order Markov chain is for notational convenience only: through a usual "reblocking" technique, the higher-order Markov case can be boiled down to the first-order case.

space, the parametrization being continuous at the boundary (such a parameterization exists thanks to the stationarity of X, and we will simply say that X is analytically parametrized by θ, for convenience), we can express the capacity of the above channel as

$$C = \max_{\theta\in\Theta} I(X(\theta);Y(\theta)) = \max_{\theta\in\Theta}\lim_{k\to\infty} I_k(X(\theta);Y(\theta)), \quad (1)$$

where

$$I_k(X(\theta);Y(\theta)) \triangleq \frac{H(X_1^k(\theta)) + H(Y_1^k(\theta)) - H(X_1^k(\theta), Y_1^k(\theta))}{k}. \quad (2)$$

Moreover, it can be shown as in [14] that $I_k(X(\theta);Y(\theta))$ (*resp.*, its derivatives) converges to $I(X(\theta);Y(\theta))$ (*resp.*, the corresponding derivatives) exponentially fast in k under Assumptions (II.a), (II.b) and (II.c). Hence, although the value of the target function $I(X(\theta);Y(\theta))$ cannot be exactly computed, it can be approximated by the function $I_k(X(\theta);Y(\theta))$, which has an explicit expression, within an error exponentially decreasing in k.

Instead of merely solving (1), we will deal with the following slightly more general problem

$$\max_{\theta} f(\theta) \triangleq \max_{\theta}\lim_{k\to\infty} f_k(\theta)$$
$$\text{subject to } \theta\in\Theta, \quad (3)$$

under the following assumptions:

(A1) Θ is a compact convex subset of $\mathbb{R}^d$ for some $d\in\mathbb{N}$ with nonempty interior Θ° and boundary $\partial\Theta$;

(A2) $f(\theta)$ and all $f_k(\theta)$, $k\geq 0$, are continuous on Θ and twice continuously differentiable in Θ°;

(A3) there exist $M_0 > 0$, $N > 0$ and $0<\rho<1$ such that for all $k\geq 1$, $\theta\in\Theta^\circ$ and $\ell = 0,1,2$, it holds true that $\|f_0^{(\ell)}(\theta)\|_2 \leq M_0$ and

$$\|f_k^{(\ell)}(\theta) - f_{k-1}^{(\ell)}(\theta)\|_2 \leq N\rho^k, \; \|f_k^{(\ell)}(\theta) - f^{(\ell)}(\theta)\|_2 \leq N\rho^k, \quad (4)$$

where $\|\cdot\|_2$ denotes the Frobenius norm (see, for example, Section 5.6 of [17]) of a vector/matrix and the superscript $^{(\ell)}$ denotes the ℓ-th order derivative. Here, for any twice continuously differentiable function $f(\theta)$, $f^{(\ell)}(\theta)$ denotes the gradient of $f(\theta)$ for $\ell=1$ and the Hessian matrix of $f(\theta)$ for $\ell=2$.

Obviously, if we set $f_k(\theta) = I_k(X(\theta);Y(\theta))$ and assume that $X(\theta)$ is analytically parameterized by some $\theta\in\Theta$, then (3) boils down to (1).

When the target function $f(\theta)$ has an explicit expression and Θ is specified by finitely many inequalities with twice differentiable terms, the optimization problem (3) can be effectively solved via, for example, the classical gradient ascent method [6] or the Frank-Wolfe method [3] or their numerous variants. However, feasible implementations and executions of these algorithms usually hinge on explicit descriptions of Θ and ∇f, both of which can be rather intricate in our setting.

Before moving to the next two sections to present our algorithms, we make some observations about the sequence $\{f_k(\theta)\}_{k=0}^\infty$. From the boundedness of $\|f_0^{(\ell)}(\theta)\|_2$ and the inequalities in (4), the uniform boundedness of $\{f_k^{(\ell)}(\theta)\}_{k=0}^\infty$ follows, i.e., there exists some $0<M<\infty$ such that for all $k\geq 0$, $\ell = 0,1,2$ and for all $\theta\in\Theta^\circ$ we have

$$\left\|f_k^{(\ell)}(\theta)\right\|_2 \leq M. \quad (5)$$

In particular, for any $\theta\in\Theta^\circ$, when $\ell=2$, $f_k^{(\ell)}(\theta) = \nabla^2 f_k(\theta)$ is a symmetric matrix whose spectral norm is given by

$$|||\nabla^2 f_k(\theta)|||_2 \triangleq \sup_{\mathbf{x}\neq 0}\frac{\|\nabla^2 f_k(\theta)\cdot\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = |\lambda_1(\theta)|,$$

where λ_1 denotes the largest (in modulus) eigenvalue of $\nabla^2 f_k(\theta)$. Hence, (5) and the easily verifiable fact $|||\nabla^2 f_k(\theta)|||_2 \leq \|\nabla^2 f_k(\theta)\|_2$ imply that

$$-M\mathbb{I}_d \preceq \nabla^2 f_k(\theta) \preceq M\mathbb{I}_d \quad (6)$$

for any k and any $\theta\in\Theta^\circ$, where $\mathbb{I}_d$ denotes the $d\times d$ identity matrix, and for two matrices A, B of the same dimension, by $A\preceq B$, we mean that $B-A$ is a positive semidefinite matrix. The existence of the constant M in (6) will be crucial for implementing our algorithms.

III. The First Algorithm: With Concavity

Throughout this section, we assume that $f(\theta)$ is strongly concave, i.e., there exists an $m>0$ such that for all $\theta\in\Theta^\circ$,

$$\nabla^2 f(\theta) \preceq -m\mathbb{I}_d. \quad (7)$$

We also assume in this section that

$$f \text{ achieves its unique maximum in } \Theta^\circ. \quad (8)$$

We will present our first algorithm to solve the optimization problem (3) under the assumptions (A1), (A2), (A3) in Section II. As mentioned before, the algorithm is in fact a modified version of the classical gradient ascent algorithm, whereas its convergence analysis is more intricate than the classical one. To overcome the issue that the target function $f(\theta)$ may not have an explicit expression we capitalize on the fact that it can be well approximated by $\{f_k(\theta)\}_{k=0}^\infty$, which will be used instead to compute how to move in each iteration.

Before presenting our algorithm, we need the following lemma, which, as will be evidenced later, is important in initializing and analyzing our first algorithm.

Lemma III.1: There exists a non-negative integer k_0 such that

(a) $\dfrac{(N+M)M\rho^{k_0+1} + 2N\rho^{k_0+1}}{1-\rho} \leq \dfrac{\delta}{8}$ and $N\rho^{k_0} \leq \dfrac{\delta}{8}$, where $\delta \triangleq \max_{\theta\in\Theta} f(\theta) - \max_{\theta\in\partial\Theta} f(\theta) > 0$.

(b) For any $k\geq k_0$, $f_k(\theta)$ is strongly concave and has a unique maximum in Θ°; and moreover, we have

$$\sup_{k\geq k_0}\|\theta_k^* - \theta^*\|_2 + \frac{d^{1/2}\rho^{k_0}}{1-\rho} < \text{dist}(\theta^*, \partial\Theta), \quad (9)$$

where θ^* denotes the unique maximum point of $f(\theta)$, θ_k^* denotes the unique maximum point of $f_k(\theta)$ and $\text{dist}(x, A)$ is defined as

$$\text{dist}(x, A) \triangleq \inf_{y\in A}\|x-y\|_2$$

for any point $x\in\mathbb{R}^d$ and any set $A\subseteq\mathbb{R}^d$.

(c) There exists a $y_0 \in \mathbb{R}$ such that for any integer $k \geq k_0$,

$$\emptyset \subsetneq B_k \subseteq C_k \subseteq \Theta^\circ \quad \text{and} \quad \text{dist}(C_k, \partial\Theta) > 0,$$

where

$$B_k \triangleq \{x \in \Theta : f_k(x) \geq y_0\},$$
$$C_k \triangleq \left\{x \in \Theta : f_k(x) \geq y_0 - \frac{\delta}{8}\right\}$$

and $\text{dist}(A, B)$ is defined as

$$\text{dist}(A, B) \triangleq \inf_{x \in A, y \in B} \|x - y\|_2$$

for any set $A, B \subseteq \mathbb{R}^d$.

Proof: Since (a) trivially holds for sufficiently large k_0, we will omit its proof and proceed to prove (b). Towards this end, note that according to (4) and (7), it holds true that for sufficiently large k, each f_k is strongly concave. Noting that $f(\theta^*) - \max_{\theta \in \partial\Theta} f(\theta) = \delta$, we deduce from (a) and (4) that for k large enough,

$$\begin{aligned}
&\max_{\theta \in \Theta} f_k(\theta) - \max_{\theta \in \partial\Theta} f_k(\theta) \\
&\quad \geq f_k(\theta^*) - \max_{\theta \in \partial\Theta} f(\theta) - \frac{\delta}{8} \\
&\quad \geq f(\theta^*) - \max_{\theta \in \partial\Theta} f(\theta) - \frac{\delta}{4} \\
&\quad = \frac{3\delta}{4} > 0. \qquad (10)
\end{aligned}$$

Hence, for k sufficiently large, f_k achieves its unique maximum at $\theta_k^* \in \Theta^\circ$.

We now prove that $\theta_k^* \to \theta^*$ as $k \to \infty$. To see this, observe that (4) implies the uniform convergence of f_k to f, i.e., for any $\varepsilon > 0$, there exists a non-negative integer K such that for any $k > K$ and any $\theta \in \Theta$, $f(\theta) - \varepsilon \leq f_k(\theta) \leq f(\theta) + \varepsilon$. In particular, if $k > K$, we have

$$f(\theta^*) - \varepsilon \leq f_k(\theta^*) \leq f_k(\theta_k^*) \leq f(\theta_k^*) + \varepsilon \leq f(\theta^*) + \varepsilon,$$

which further implies that $f_k(\theta_k^*) \to f(\theta^*)$ as $k \to \infty$. It then follows from the triangle inequality that

$$f(\theta_k^*) \to f(\theta^*), \quad \text{as } k \to \infty. \qquad (11)$$

Now, by the Taylor series expansion, there exists some $\tilde{\theta} \in \Theta^\circ$ such that

$$f(\theta_k^*) - f(\theta^*) = \nabla f(\theta^*)^T(\theta_k^* - \theta^*) + (\theta_k^* - \theta^*)^T \nabla^2 f(\tilde{\theta})(\theta_k^* - \theta^*). \qquad (12)$$

Since $\nabla f(\theta^*) = 0$ and $\nabla^2 f(\tilde{\theta}) \preceq -m\mathbb{I}_d$ according to (7), it follows from (11) and (12) that $\theta_k^* \to \theta^*$ as $k \to \infty$, as desired.

It then immediately follows that $\|\theta_k^* - \theta^*\|_2 + d^{1/2}\rho^k/(1-\rho) \to 0$ as $k \to \infty$. Observing that $\text{dist}(\theta^*, \partial\Theta) > 0$ (since $\theta^* \in \Theta^\circ$), we infer that (9) holds for sufficiently large k. Hence, (b) will be satisfied as long as k_0 is sufficiently large.

We now show that (c) also holds for sufficiently large k_0. From the definition of δ, there exists a y_0 such that

$$\max_{\theta \in \partial\Theta} f(\theta) + \frac{\delta}{4} < y_0 < \max_{\theta \in \Theta} f(\theta) - \frac{\delta}{4}.$$

Recalling (4) and using the same logic as that used to derive (10), we infer that for all sufficiently large k,

$$\max_{\theta \in \partial\Theta} f_k(\theta) < y_0 - \frac{\delta}{8} < y_0 < \max_{\theta \in \Theta} f_k(\theta). \qquad (13)$$

According to (b) and the fact that $\theta_k^* \in \Theta^\circ$, which follows from (13), we deduce that $\emptyset \subsetneq B_k \subseteq C_k \subseteq \Theta^o$ and $\text{dist}(C_k, \partial\Theta) > 0$ with

$$C_k \triangleq \left\{x : f_k(x) \geq y_0 - \frac{\delta}{8}\right\} \quad \text{and} \quad B_k \triangleq \{x : f_k(x) \geq y_0\}.$$

Therefore, (c) is valid as long as k_0 is sufficiently large. Finally, choosing a larger k_0 if necessary, we conclude that there exists a non-negative integer k_0 such that (a), (b) and (c) are all satisfied. ■

Remark III.2: We remark that, for any $k \geq k_0$, each B_k specified as above has a non-empty interior, which is due to the rightmost strict inequality in (13) and the continuity of f_k.

We are now ready to present our first algorithm, which modifies the classical gradient ascent method in the following manner: instead of using ∇f to find a feasible direction, we use ∇f_k as the ascent direction in the k-th iteration and then pose additional check conditions for a careful choice of the step size. Note that such modifications make the convergence analysis more difficult compared to the classical case, as elaborated on in the next subsection.

Algorithm III.3 (The First Modified Gradient Ascent Algorithm):

Step 0. Choose k_0 such that Lemma III.1 (a)-(c) hold. Set $k = 0$, $g_0 = f_{k_0}$ and choose $\alpha \in (0, 0.5)$, $\beta \in (0, 1)$ and $\theta_0 \in \Theta^\circ$ such that $\theta_0 \in B_{k_0}$ and $\nabla g_0(\theta_0) \neq 0$.

Step 1. Increase k by 1, and set $t = 1$, $g_k = f_{k_0+k}$.

Step 2. If $\nabla g_{k-1}(\theta_{k-1}) = 0$, set

$$\tau = \theta_{k-1} + t\nabla g_{k-1}(\theta_{k-1} + \rho^{k+k_0}\mathbf{1}),$$

where $\mathbf{1}$ denotes the all-one vector in $\mathbb{R}^d$; otherwise, set

$$\tau = \theta_{k-1} + t\nabla g_{k-1}(\theta_{k-1}).$$

If $\tau \notin \Theta$ or

$$g_k(\tau) < g_k(\theta_{k-1}) + \alpha t \|\nabla g_{k-1}(\theta_{k-1})\|_2^2 - (N+M)Mt\rho^{k+k_0},$$

set $t = \beta t$ and go to Step 2, otherwise set $\theta_k = \tau$ and go to Step 1.

Remark III.4: N and ρ are chosen such that (4) holds and the choices of them can be derived from Section IV of [15] in practice. Furthermore, it is obvious from the definition of g_k that as k tends to infinity, g_k (*resp.*, its first and second order derivatives) converges to f (*resp.*, its first and second order derivatives) exponentially fast with the same constants N and ρ as in (4).

Remark III.5: According to Lemma III.1, the choice of k_0 depends on practical evaluations of constants N, ρ, M, and is different from case to case. Moreover, the existence of θ_0 can also be justified by Lemma III.1 (c).

Remark III.6: We point out that for any $k \geq 1$ in Step 2 of Algorithm III.3, when $\nabla g_{k-1}(\theta_{k-1}) = 0$, the point $\theta_{k-1} + \rho^{k+k_0}\mathbf{1}$ will always lie in Θ°. To see this, note that if θ_{k-1} is the maximum point of $g_{k-1} = f_{k+k_0-1}$, then

$\theta_{k-1} = \theta^*_{k+k_0-1}$. However, by Lemma III.1 (b), $\theta^*_{k+k_0-1}$ satisfies (9), which immediately implies that $\theta_{k-1}+\rho^{k+k_0}\mathbf{1} \in \Theta^\circ$ when $\nabla g_{k-1}(\theta_{k-1}) = 0$, for any $k \geq 1$.

Remark III.7: For technical reasons that will be made clear in the next section, α is chosen within $(0, 0.5)$ to ensure the convergence of the algorithm. In Step 2 of Algorithm III.3, the case that $\nabla g_{k-1}(\theta_{k-1}) = 0$ is singled out for special treatment to prevent the algorithm from getting trapped at the maximum point of f_{k-1} for a fixed k, which may be still far away from the maximum point of f.

A. Convergence Analysis

As mentioned earlier, compared to the classical gradient ascent method, Algorithm III.3 poses additional challenges for convergence analysis. The main difficulties come from the two check conditions in Step 2: the "perturbed" Armijo condition (see, e.g., Chapter 2 of [3] for more details)

$$g_k(\tau) \geq g_k(\theta_{k-1})+\alpha t\,\|\nabla g_{k-1}(\theta_{k-1})\|_2^2-(N+M)Mt\rho^{k+k_0}$$

may break the monotonicity of the sequence $\{g_k(\theta_k)\}_{k=0}^{\infty}$ which would have been used to simplify the convergence analysis in the classical case; and the extra check condition $\tau \in \Theta$ (τ depends on k) forces us to seek uniform control (over all k) of the time used to ensure the validity of this condition in each iteration. In the remainder of this section, we deal with these problems and examine the convergence behavior of Algorithm III.3. In a nutshell, we will prove that our algorithm converges exponentially fast under some strong concavity assumptions.

Note that the variable k in Algorithm III.3 actually records the number of times that Step 1 has been executed at the present moment. To facilitate the analysis of our algorithm, we will put it into an equivalent form, where an additional variable n is used to record the number of times that Step 2 has been executed.

Below is Algorithm III.3 rewritten with the additional variable n.

Algorithm III.8 (An Equivalent Form of Algorithm III.3):

Step 0. Choose k_0 such that Lemma III.1 (a)-(c) hold. Set $n = 0, k = 0, \hat{g}_0 = g_0 = f_{k_0}$, and choose $\alpha \in (0, 0.5), \beta \in (0, 1)$ and $\hat{\theta}_0 \in \Theta^\circ$ such that $\hat{\theta}_0 \in B_{k_0}$ and $\nabla\hat{g}_0(\hat{\theta}_0) \neq 0$.

Step 1. Increase k by 1, and set $t = 1$, $g_k = f_{k_0+k}$.

Step 2. Increase n by 1. If $\nabla\hat{g}_{n-1}(\hat{\theta}_{n-1}) = 0$, set

$$\tau = \hat{\theta}_{n-1} + t\nabla\hat{g}_{n-1}(\hat{\theta}_{n-1} + \rho^{k+k_0}\mathbf{1}); \tag{14}$$

otherwise, set

$$\tau = \hat{\theta}_{n-1} + t\nabla\hat{g}_{n-1}(\hat{\theta}_{n-1}). \tag{15}$$

If $\tau \notin \Theta^\circ$ or

$$g_k(\tau) < g_k(\hat{\theta}_{n-1})+\alpha t\|\nabla\hat{g}_{n-1}(\hat{\theta}_{n-1})\|_2^2-(N+M)Mt\rho^{k+k_0}, \tag{16}$$

then set $\hat{\theta}_n = \hat{\theta}_{n-1}, \hat{g}_n = \hat{g}_{n-1}, t = \beta t$ and go to Step 2; otherwise, set $\hat{\theta}_n = \tau, \hat{g}_n = g_k$ and go to Step 1.

Remark III.9: Let $n_0 = 0$, and for any $k \geq 1$, recursively define

$$n_k \triangleq \inf\{n > n_{k-1} : \hat{\theta}_n \neq \hat{\theta}_{n-1}\}.$$

Then, it can be verified in a straightforward manner that $\hat{\theta}_{n_k} = \theta_k$, $\hat{g}_{n_k} = g_k = f_{k+k_0}$ for any $k \geq 0$ and $\hat{\theta}_l = \hat{\theta}_{l+1}$, $\hat{g}_l = \hat{g}_{l+1}$ for any l with $n_{k-1} \leq l \leq n_k-1$. This justifies the equivalence between Algorithm III.3 and Algorithm III.8.

The following theorem establishes the exponential convergence of Algorithm III.8 with respect to n.

Theorem III.10: Suppose, as in (7) and (8), that the strongly concave function f achieves its unique maximum at $\theta^* \in \Theta^\circ$. Then there exist an $\hat{M} > 0$ and a $\hat{\xi} \in (0, 1)$ such that for all $n \geq 0$,

$$|\hat{g}_n(\hat{\theta}_n) - f(\theta^*)| \leq \hat{M}\hat{\xi}^n, \tag{17}$$

where $\hat{g}_n(\hat{\theta}_n)$ is obtained by executing Algorithm III.8.

Proof: For simplicity, we only deal with the case that $\nabla\hat{g}_{n-1}(\hat{\theta}_{n-1}) \neq 0$ in Step 2 of Algorithm III.8 (and therefore (15) is actually executed), since the opposite case follows from a similar argument by replacing $\hat{\theta}_{n-1}$ with $\hat{\theta}_{n-1} + \rho^{k+k_0}\mathbf{1}$.

Let $T_1(k)$ denote the smallest non-negative integer p such that

$$\hat{\theta}_{n_{k-1}} + \beta^p\nabla\hat{g}_{n_{k-1}}(\hat{\theta}_{n_{k-1}}) \in \Theta^\circ, \tag{18}$$

$T(k)$ denote the smallest non-negative integer q such that $q \geq T_1(k)$ and

$$\begin{aligned} &g_k(\hat{\theta}_{n_{k-1}} + \beta^q\nabla\hat{g}_{n_{k-1}}(\hat{\theta}_{n_{k-1}})) \\ &\quad\geq g_k(\hat{\theta}_{n_{k-1}}) + \alpha\beta^q\|\nabla\hat{g}_{n_{k-1}}(\hat{\theta}_{n_{k-1}})\|_2^2 - (N+M)M\beta^q\rho^{k+k_0}. \end{aligned}$$

Note that the well-definedness of $T_1(k)$ and $T(k)$ follows from the observation that if (18) holds for some non-negative integer p, then it also holds for any integer $p' > p$. Adopting these definitions, we can immediately verify that $T(k) = n_k - n_{k-1}$, which corresponds to the number of times Step 2 (of Algorithm III.3) has been executed to obtain $\hat{\theta}_{n_k}$ from $\hat{\theta}_{n_{k-1}}$.

The remainder of the proof consists of the following three steps.

Step 1: Uniform boundedness of $T(k)$. In this step, we show that there exists an $A \geq 0$ such that $T(k) \leq A$ for all k.

Since Θ° is open and $\hat{\theta}_0 \in \Theta^\circ$, we have $T_1(k) < \infty$ for any $k \geq 0$. Note that we have not shown that $T_1(k)$ is uniformly bounded at this stage.

For any $q \geq T_1(k)$, letting

$$\tau = \hat{\theta}_{n_{k-1}} + \beta^q\nabla\hat{g}_{n_{k-1}}(\hat{\theta}_{n_{k-1}}),$$

we have $\tau \in \Theta^\circ$ and so both $f_k(\tau)$ and $f(\tau)$ are well-defined. Recalling from (6) that

$$\nabla^2 g_k(\theta) = \nabla^2 f_{k+k_0}(\theta) \succeq -M\mathbb{I}_d$$

for any $k \geq 0$ and any $\theta \in \Theta^\circ$, we derive from the Taylor series expansion that

$$\begin{aligned} g_k(\tau) =& g_k(\hat{\theta}_{n_{k-1}}) + \beta^q \nabla g_k(\hat{\theta}_{n_{k-1}})^T \nabla \hat{g}_{n_{k-1}}(\hat{\theta}_{n_{k-1}}) \\ &+ \frac{\beta^{2q}}{2} \nabla \hat{g}_{n_{k-1}}(\hat{\theta}_{n_{k-1}})^T \nabla^2 g_k(\tilde{\theta}_k) \nabla \hat{g}_{n_{k-1}}(\hat{\theta}_{n_{k-1}}) \\ \geq& g_k(\hat{\theta}_{n_{k-1}}) + \beta^q \nabla g_k(\hat{\theta}_{n_{k-1}})^T \nabla \hat{g}_{n_{k-1}}(\hat{\theta}_{n_{k-1}}) \\ &- \frac{M\beta^{2q}}{2} \|\nabla \hat{g}_{n_{k-1}}(\hat{\theta}_{n_{k-1}})\|_2^2, \end{aligned} \tag{19}$$

where $\tilde{\theta}_k \in \Theta^\circ$. According to (4), we have

$$\begin{aligned} &\nabla g_k(\hat{\theta}_{n_{k-1}})^T \nabla \hat{g}_{n_{k-1}}(\hat{\theta}_{n_{k-1}}) \\ &= \nabla \hat{g}_{n_{k-1}}(\hat{\theta}_{n_{k-1}})^T \nabla \hat{g}_{n_{k-1}}(\hat{\theta}_{n_{k-1}}) \\ &\quad + (\nabla g_k(\hat{\theta}_{n_{k-1}})^T \nabla \hat{g}_{n_{k-1}}(\hat{\theta}_{n_{k-1}}) \\ &\quad - \nabla \hat{g}_{n_{k-1}}(\hat{\theta}_{n_{k-1}})^T \nabla \hat{g}_{n_{k-1}}(\hat{\theta}_{n_{k-1}})) \\ &\geq \|\nabla \hat{g}_{n_{k-1}}(\hat{\theta}_{n_{k-1}})\|_2^2 - N\rho^{k+k_0} \|\nabla \hat{g}_{n_{k-1}}(\hat{\theta}_{n_{k-1}})\|_2. \end{aligned}$$

This, together with (19), implies that

$$\begin{aligned} g_k(\tau) \geq& g_k(\hat{\theta}_{n_{k-1}}) + \beta^q \|\nabla \hat{g}_{n_{k-1}}(\hat{\theta}_{n_{k-1}})\|_2^2 \\ &- \frac{M\beta^{2q}}{2} \|\nabla \hat{g}_{n_{k-1}}(\hat{\theta}_{n_{k-1}})\|_2^2 \\ &- N\beta^q \rho^{k+k_0} \|\nabla \hat{g}_{n_{k-1}}(\hat{\theta}_{n_{k-1}})\|_2 \\ \geq& g_k(\hat{\theta}_{n_{k-1}}) + \beta^q \|\nabla \hat{g}_{n_{k-1}}(\hat{\theta}_{n_{k-1}})\|_2^2 \\ &- \frac{M\beta^{2q}}{2} \|\nabla \hat{g}_{n_{k-1}}(\hat{\theta}_{n_{k-1}})\|_2^2 - NM\beta^q \rho^{k+k_0}, \end{aligned}$$

where the last inequality follows from (5). Note that for any non-negative integer $q \geq -\log M/\log \beta$, we have

$$\beta^q - \frac{M\beta^{2q}}{2} \geq \frac{1}{2}\beta^q > \alpha\beta^q,$$

which immediately implies that (16) fails; in other words, for any non-negative integer $q \geq T_1(k)$, we have

$$\begin{aligned} &g_k(\hat{\theta}_{n_{k-1}} + \beta^q \nabla \hat{g}_{n_{k-1}}(\hat{\theta}_{n_{k-1}})) \\ &\quad \geq g_k(\hat{\theta}_{n_{k-1}}) + \alpha\beta^q \|\nabla \hat{g}_{n_{k-1}}(\hat{\theta}_{n_{k-1}})\|_2^2 \\ &\quad - (N+M)M\beta^q \rho^{k+k_0} \end{aligned}$$

as long as $q \geq -\log M/\log \beta$. It then follows that for any integer $k \geq 1$, $T(k)$ can be bounded as

$$T(k) \leq \begin{cases} A_2 & \text{if } T_1(k) \leq A_2 \\ T_1(k) & \text{if } T_1(k) > A_2, \end{cases} \tag{20}$$

where $A_2 \triangleq \max\{0, -\log M/\log \beta + 1\}$ is a constant independent of k. Now, to prove the uniform boundedness of $T(k)$, what remains is to show that there exists an $A_1 \geq 0$ such that for all k, $T_1(k) \leq A_1$.

From the definition of $T(k)$, we have

$$\begin{aligned} g_k(\hat{\theta}_{n_k}) \geq& g_k(\hat{\theta}_{n_{k-1}}) + \alpha\beta^{T(k)} \|\nabla \hat{g}_{n_{k-1}}(\hat{\theta}_{n_{k-1}})\|_2^2 \\ &- (N+M)M\beta^{T(k)}\rho^{k+k_0}. \end{aligned} \tag{21}$$

Note that (4) and (21) imply that

$$g_k(\hat{\theta}_{n_k}) \geq g_{k-1}(\hat{\theta}_{n_{k-1}}) - (N+M)M\rho^{k+k_0} - N\rho^{k+k_0}.$$

By summation, we obtain that

$$\begin{aligned} g_k(\hat{\theta}_{n_k}) &\geq g_0(\hat{\theta}_0) - \sum_{i=0}^{k-1} \left[(N+M)M\rho^{i+k_0+1} + N\rho^{i+k_0+1}\right] \\ &\geq g_0(\hat{\theta}_0) - \left[\frac{(N+M)M\rho^{k_0+1}}{1-\rho} + \frac{N\rho^{k_0+1}}{1-\rho}\right]. \end{aligned}$$

It then follows from (4) that for all $k \geq 0$ we have

$$\begin{aligned} &g_0(\hat{\theta}_{n_k}) \\ &\geq g_0(\hat{\theta}_0) - \left[\frac{(N+M)M\rho^{k_0+1}}{1-\rho} + \frac{N\rho^{k_0+1}}{1-\rho}\right] - \sum_{i=1}^{k} N\rho^{i+k_0} \\ &\geq g_0(\hat{\theta}_0) - \left[\frac{(N+M)M\rho^{k_0+1}}{1-\rho} + \frac{2N\rho^{k_0+1}}{1-\rho}\right] \\ &\geq g_0(\hat{\theta}_0) - \frac{\delta}{8}, \end{aligned} \tag{22}$$

where the last inequality follows from Lemma III.1 (a). Now, letting y_0, B_{k_0} and C_{k_0} be defined as in Lemma III.1, we infer from (22) and Lemma III.1 (c) that $\{\hat{\theta}_{n_k}\}_{k=0}^\infty \subseteq C_{k_0} \subseteq \Theta^\circ$. Hence, for any non-negative integer $p \geq \log(\text{dist}(C_{k_0}, \partial\Theta)/M)/\log \beta$, we have $\hat{\theta}_{n_{k-1}} + \beta^p \nabla \hat{g}_{n_{k-1}}(\hat{\theta}_{n_{k-1}}) \in \Theta^\circ$ and it then follows that $T_1(k) \leq A_1$, where A_1 is defined as

$$A_1 \triangleq \max\left\{0, \frac{\log(\text{dist}(C_{k_0}, \partial\Theta)/M)}{\log \beta} + 1\right\}. \tag{23}$$

Finally, it follows from (20) and (23) that

$$T(k) \leq A \triangleq \max\{A_1, A_2\}, \tag{24}$$

as desired.

Step 2: Exponential convergence of $\{f(\hat{\theta}_{n_k})\}$. Using (4), (5) and the definition of $\{\hat{g}_{n_k}\}_{k=0}^\infty$, we deduce From (21) that

$$\begin{aligned} f(\hat{\theta}_{n_k}) \geq& f(\hat{\theta}_{n_{k-1}}) + \alpha\beta^{T(k)} \|\nabla f(\hat{\theta}_{n_{k-1}})\|_2^2 \\ &- [(N+M)M\beta^{T(k)} + 2N + 2NM\rho]\rho^{k+k_0}. \end{aligned}$$

On the other hand, we infer from (7) that

$$f(\theta^*) \leq f(\hat{\theta}_{n_{k-1}}) + \nabla f(\hat{\theta}_{n_{k-1}})^T(\theta^* - \hat{\theta}_{n_{k-1}}) - \frac{m}{2}\|\theta^* - \hat{\theta}_{n_{k-1}}\|_2^2,$$

which, coupled with some straightforward estimates, yields

$$2m(f(\theta^*) - f(\hat{\theta}_{n_{k-1}})) \leq \|\nabla f(\hat{\theta}_{n_{k-1}})\|_2^2.$$

It then follows that

$$\begin{aligned} &f(\theta^*) - f(\hat{\theta}_{n_k}) \\ &\quad \leq f(\theta^*) - f(\hat{\theta}_{n_{k-1}}) - \alpha\beta^{T(k)} \|\nabla f(\hat{\theta}_{n_{k-1}})\|_2^2 \\ &\quad\quad + [(N+M)M\beta^{T(k)} + 2N + 2NM\rho]\rho^{k+k_0} \\ &\quad \leq (1 - 2m\alpha\beta^{T(k)})(f(\theta^*) - f(\hat{\theta}_{n_{k-1}})) \\ &\quad\quad + [(N+M)M + 2N + 2NM\rho]\rho^{k+k_0} \\ &\quad \overset{(d)}{\leq} \left(1 - \min\left\{2\, m\alpha\beta^{A_1}, 2\, m\alpha\beta^{A_2}\right\}\right)(f(\theta^*) - f(\hat{\theta}_{n_{k-1}})) \\ &\quad\quad + \left(\frac{NM + M^2 + 2N}{\rho} + 2NM\right)\rho^{k+k_0+1} \\ &\quad = \eta(f(\theta^*) - f(\hat{\theta}_{n_{k-1}})) + \gamma_k, \end{aligned} \tag{25}$$

where

$$\eta \triangleq 1 - \min\left\{2m\alpha, \frac{\text{dist}(C_{k_0}, \partial\Theta)}{M}2m\alpha\beta, \frac{2m\alpha\beta}{M}\right\},$$
$$\gamma_k \triangleq \left(\frac{NM + M^2 + 2N}{\rho} + 2NM\right)\rho^{k+k_0+1}$$

and (d) follows from (24). Recursively applying (25) and noting that $0 < \eta < 1$, we infer that there exist an $M' > 0$ and a $\xi \in (0,1)$ such that

$$f(\theta^*) - f(\hat{\theta}_{n_k}) \leq M'\xi^k. \tag{26}$$

Step 3: Exponential convergence of $\{\hat{g}_n(\hat{\theta}_n)\}$. In this step, we establish (17) and thereby finish the proof.

First, note that for any positive integer $n \geq 0$, there exists an integer $k' \geq 0$ such that

$$n_{k'} \leq n \leq n_{k'+1}, \quad n \leq (k'+1)A,$$
$$\hat{\theta}_n = \hat{\theta}_{n_{k'}}, \quad \hat{g}_n(\hat{\theta}_n) = \hat{g}_{n_{k'}}(\hat{\theta}_{n_{k'}}),$$

where A is defined in (24). These four inequalities, together with (4) and (26), imply that there exists an $\hat{M} > 0$ and a $\hat{\xi} \in (0,1)$ such that for any $n \geq 0$,

$$\begin{aligned}
&|\hat{g}_n(\hat{\theta}_n) - f(\theta^*)| \\
&\quad\leq |\hat{g}_{n_{k'}}(\hat{\theta}_{n_{k'}}) - f(\hat{\theta}_{n_{k'}})| + |f(\hat{\theta}_{n_{k'}}) - f(\theta^*)| \\
&\quad\leq N\rho^{k'+k_0} + M'\xi^{k'} \\
&\quad\leq N\rho^{k_0}\rho^{\lfloor n/A \rfloor - 1} + M'\xi^{\lfloor n/A \rfloor - 1} \\
&\quad\leq \hat{M}\hat{\xi}^n,
\end{aligned}$$

which completes the proof of the theorem. ■

Theorem III.10, together with the uniform boundedness of $T(k)$ established in its proof, immediately implies that Algorithm III.3 converges exponentially in k. More precisely, we have the following theorem.

Theorem III.11: For a strongly concave function f whose unique maximum is achieved at $\theta^* \in \Theta^\circ$, given in terms of the approximating sequence of functions $\{f_k\}_{k=0}^{\infty}$ as in (3), satisfying assumptions (A1), (A2) and (A3) in Section II, there exist an $\tilde{M} > 0$ and a $\tilde{\xi} \in (0,1)$ depending on m, M, N and ρ such that for all k,

$$|g_k(\theta_k) - f(\theta^*)| \leq \tilde{M}\tilde{\xi}^k, \tag{27}$$

where $g_k(\theta_k)$ is defined as in Algorithm III.3.

B. Applications of Algorithm III.3

In this section, we discuss some applications of Algorithm III.3 in information theory.

Consider an FSC satisfying (II.a)-(II.c) and assume that all the matrices in $\Pi_{F,\delta}$ are analytically parameterized by $\theta \in \Theta^\circ$, where Θ is a compact convex subset of $\mathbb{R}^d$, $d \in \mathbb{N}$. Setting

$$f(\theta) = I(X(\theta); Y(\theta))$$

and

$$\begin{aligned}
f_k(\theta) = &H(X_2(\theta)|X_1(\theta)) + H(Y_{k+1}(\theta)|Y_1^k(\theta)) \\
&- H(X_{k+1}(\theta), Y_{k+1}(\theta)|X_1^k(\theta), Y_1^k(\theta)),
\end{aligned} \tag{28}$$

we derive from [13] that (4) holds (the way of choosing N and ρ in (4) for practical channels can also be derived from Section IV of [15]). So, when $f(\theta)$ is strongly concave with respect to θ (this may hold true for some special channels, see, for example, [16] and [20]) as in (7), our algorithm applied to $\{f_k(\theta)\}_{k=0}^{\infty}$ converges exponentially fast in the number of steps to the maximum value of $f(\theta)$. This, and the easily verifiable fact that the computational complexity of $f_k(\theta)$ is at most exponential in k, leads to the conclusion that Algorithm III.3, when applied to $\{f_k(\theta)\}_{k=0}^{\infty}$ as above, achieves exponential accuracy in exponential time. We now trade exponential time for polynomial time at the expense of accuracy (see, e.g., Section 5 of [25] for a similar idea). For any fixed $r \in \mathbb{R}_+$ and any $k \geq \lceil r\log 2 \rceil$, choose the largest $l \in \mathbb{N}$ such that $k = \lceil r \log l \rceil$. Substituting this into (27), we have

$$|g_{\lceil r\log l \rceil}(\theta_{\lceil r\log l \rceil}) - f(\theta^*)| \leq \tilde{M} l^{r\log\tilde{\xi}}.$$

In other words, as summarized in the following theorem, we have shown that Algorithm III.3, when used to compute the channel capacity as above, achieves polynomial accuracy in polynomial time.

Theorem III.12: For a general FSC satisfying (II.a)-(II.c) and parameterized as above, if $I(X(\theta); Y(\theta))$ is strongly concave with respect to $\theta \in \Theta$ and achieves its unique maximum in Θ°, then for any fixed $r \in \mathbb{R}_+$, there exists an algorithm computing its fixed-order Markov capacity that achieves polynomial accuracy $O(l^{r\log\tilde{\xi}})$ in polynomial time $O(l^{r\log c})$, where $\tilde{\xi}$ is the constant defined in (27) and c is the cardinality of the output alphabet of the channel.

In the following, we show that for certain special families of FSCs, we get a stronger convergence result than that in Theorem III.12. In particular, for the following two examples, Algorithm III.3 can be applied to the sequence $\{f_k(\theta)\}_{k=0}^{\infty}$ (or its variants) defined in (28) to compute the channel capacity, achieving exponential accuracy in polynomial time.

1) BEC With the $(1,\infty)$-RLL Constraint: In this section, we consider the Markov capacity of a binary erasure channel (BEC) under the $(1,\infty)$-RLL constraint.

A BEC with parameter ε (denoted by BEC(ε)) is a channel with binary input process $\{X_n\}_{n=1}^{\infty}$ and output process $\{Y_n\}_{n=1}^{\infty}$ such that for each n, Y_n equals the transmitted symbol X_n with probability $1-\varepsilon$, and equals the erasure symbol with probability ε. When there is no input constraint, it is well-known that the capacity of this channel equals $1-\varepsilon$. In this section, we focus on the case when $\{X_n\}_{n=1}^{\infty}$ is required to satisfy the $(1,\infty)$-RLL constraint, i.e., $\{11\}$ is a *forbidden set* for $\{X_n\}_{n=1}^{\infty}$. Under this constraint, the BEC(ε) can be viewed as an FSC with only one state. It is worth noting that although the feedback capacity of this channel was found in [29], its non-feedback capacity remains unknown. In the following, we will consider the Markov capacity of the BEC(ε) under the $(1,\infty)$-RLL constraint when there is no feedback. For such a channel, we will use Algorithm III.3 to evaluate the first-order Markov capacity, which, compared to a lower bound for the second-order Markov capacity, will lead to the conclusion that higher-order memory in the channel input may increase the Markov capacity, even though the input of the channel forms a 1-step shift of finite type [22].

We first suppose that $\{X_n\}_{n=1}^{\infty}$ is a first-order stationary Markov chain with the transition probability matrix (indexed by 0, 1)

$$\Pi = \begin{bmatrix} 1-\theta & \theta \\ 1 & 0 \end{bmatrix}$$

for $0 < \theta < 1$. It has been established in [20] that the mutual information rate $I(X(\theta);Y(\theta))$ of a BEC(ε) with this input can be computed as

$$I(X(\theta);Y(\theta)) = (1-\varepsilon)^2 \sum_{l=0}^{\infty} H(X_{l+2}(\theta)|X_1(\theta))\varepsilon^l$$

and $I(X(\theta),Y(\theta))$ is strictly concave with respect to θ. Now, setting $f(\theta) = I(X(\theta);Y(\theta))$, one verifies, through straightforward computation, that

$$f(\theta) = \lim_{k\to\infty} f_k(\theta),$$

where

$$\begin{aligned} f_0(\theta) = & f_1(\theta) \triangleq (1-\varepsilon)^2 \frac{H(\theta)}{1+\theta}, \\ f_k(\theta) \triangleq & (1-\varepsilon)^2 \left(\frac{H(\theta)}{1+\theta} + \sum_{l=2}^{k} \left(\frac{1}{1+\theta} H\left(\frac{1-(-\theta)^{l+1}}{1+\theta} \right) \right) \varepsilon^{l-1} \right. \\ & \left. + \sum_{l=2}^{k} \left(\frac{\theta}{1+\theta} H\left(\frac{1-(-\theta)^{l}}{1+\theta} \right) \right) \varepsilon^{l-1} \right) \end{aligned}$$

for $k \geq 2$ and $H(p) \triangleq -p\log p - (1-p)\log(1-p)$ is the binary entropy function. In what follows, assuming $\varepsilon = 0.1$, we will show that Algorithm III.3 can be applied to compute the first-order Markov capacity of the BEC(ε) under the $(1,\infty)$-RLL constraint, i.e., the maximum of $f(\theta)$ over all $\theta \in [0,1]$.

From now on, we suppose $\varepsilon = 0.1$ in the remainder of this section.

First of all, we claim that $f(\theta)$ achieves its unique maximum within the interval $[0.25, 0.55]$ and therefore in the interior of $\Theta \triangleq [0.2, 0.6]$. To see this, noting that $f_k(\theta) \leq f(\theta)$ for any θ and through evaluating the elementary function $f_{100}(\theta)$, we have

$$0.442239 < \max_{\theta\in[0.25,0.55]} f_{100}(\theta) < 0.442240$$

and therefore

$$\max_{\theta\in[0.25,0.55]} f(\theta) \geq 0.442239, \tag{29}$$

where (29) follows from the fact that $f_k(\theta)$ is monotonically increasing in k. On the other hand, using the stationarity of $\{Y_n\}_{n=1}^{\infty}$ and the fact that conditioning can not increase entropy, we have

$$\begin{aligned} f(\theta) &= I(X(\theta);Y(\theta)) \\ &= H(Y) - H(\varepsilon) \\ &\leq H(Y_3(\theta)|Y_1(\theta),Y_2(\theta)) - H(\varepsilon), \end{aligned}$$

where $H(Y)$ is the entropy rate of $\{Y_n\}_{n=1}^{\infty}$. Then, by straightforward computation, we deduce that (recall that $\varepsilon = 0.1$),

$$\begin{aligned} &\max_{\theta\in[0,0.25]\cup[0.55,1]} f(\theta) \\ &\quad \leq \max_{\theta\in[0,0.25]\cup[0.55,1]} H(Y_3(\theta)|Y_1(\theta),Y_2(\theta)) - H(\varepsilon) \\ &\quad < 0.414483, \end{aligned}$$

which, together with (29), yields

$$\max_{\theta\in[0,0.25]\cup[0.55,1]} f(\theta) < \max_{\theta\in[0.25,0.55]} f(\theta),$$

as desired.

Next, we will show that (4), (5) and (7) are satisfied for all $\theta \in [0.2, 0.6]$. Note that for $k \geq 2$ we have

$$\begin{aligned} &f_k(\theta) - f_{k-1}(\theta) \\ &\quad = \left(\frac{1}{1+\theta} H\left(\frac{1-(-\theta)^{k+1}}{1+\theta} \right) + \frac{\theta}{1+\theta} H\left(\frac{1-(-\theta)^{k}}{1+\theta} \right) \right) \\ &\quad \times (1-\varepsilon)^2 \varepsilon^{k-1}. \end{aligned}$$

Recalling that we require $\varepsilon = 0.1$, this implies that for any $k \geq 5$ and any $\theta \in [0.2, 0.6]$,

$$|f_k(\theta) - f_{k-1}(\theta)| \leq (1-\varepsilon)^2 \varepsilon^{k-1} = 8.1 \times 0.1^k.$$

This, together with the easily verifiable fact that $0.378 \leq f_5(\theta) \leq 0.443$ for any $\theta \in [0.2, 0.6]$, further implies that

$$|f_k(\theta) - f(\theta)| \leq 0.9 \times 0.1^k \quad \text{and} \quad 0.37 \leq f_k(\theta) \leq 0.45$$

for any $k \geq 5$ and any $\theta \in [0.2, 0.6]$.

Going through similar arguments, we obtain that when $\varepsilon = 0.1$, for any $k \geq 13$ and any $\theta \in [0.2, 0.6]$,

$$\begin{aligned} |f_k'(\theta) - f_{k-1}'(\theta)| &\leq 72.9 \times 0.1^k, \\ |f_k'(\theta) - f'(\theta)| &\leq 8.1 \times 0.1^k, \end{aligned}$$

and

$$-0.44 \leq f_k'(\theta) \leq 0.76,$$

and, for any $k \geq 18$ and any $\theta \in [0.2, 0.6]$,

$$\begin{aligned} |f_k''(\theta) - f_{k-1}''(\theta)| &\leq 370.575 \times 0.1^k, \\ |f_k''(\theta) - f''(\theta)| &\leq 41.175 \times 0.1^k, \end{aligned}$$

and

$$-5.81 \leq f_k''(\theta) \leq -1.88.$$

To sum up, we have shown that (4) is satisfied with $N = 371$ and $\rho = 0.1$, (5) is satisfied with $M = 5.81$ and (7) is satisfied with $m = 1.88$. Under these choices of the constants, direct calculation shows that $k_0 = 18$ is sufficient for Lemma III.1. As a result, Algorithm III.3 is applicable to find the maximum of the function f. Observing that, by its definition, the computational complexity of $f_k(\theta)$ is polynomial in k, we conclude that Algorithm III.3 achieves exponential accuracy in polynomial time.

Now, applying Algorithm III.3 to the sequence $\{f_k(\theta) : k \geq 18\}$ over $\Theta = [0.2, 0.6]$ with $\alpha = 0.4$, $\beta = 0.9$ and the initial point $\theta_0 = 0.5$, we obtain that

$$\theta_{110} \approx 0.395485, \quad f_{110}(\theta_{110}) \approx 0.442239.$$

Furthermore, under the settings given above, ξ and η can be chosen such that $\xi = \eta < 0.767$. It now follows from (4), (26) and $\hat{\theta}_{n_k} = \theta_k$ (see Remark III.9) that

$$|f_{110}(\theta_{110}) - f(\theta^*)| \\ \leq |f_{110}(\theta_{110}) - f(\theta_{110})| + |f(\theta_{110}) - f(\theta^*)| \leq 2.621 \times 10^{-7},$$

which further implies that when the input is a first-order Markov chain and $\varepsilon = 0.1$, the capacity of the BEC(ε) under the $(1, \infty)$-RLL constraint can be bounded as

$$0.4422382 \leq f(\theta^*) \leq 0.4422398. \quad (30)$$

Note that according to [29], the feedback capacity of a BEC(ε) under the $(1, \infty)$-constraint can be computed as

$$C_{\text{fb}}(\varepsilon) = \max_{0 \leq p \leq 0.5} \frac{(1-\varepsilon)H(p)}{1 + (1-\varepsilon)p}.$$

When $\varepsilon = 0.1$, this value is 0.445502, which is greater than the upper bound we obtain in (30).

We now consider the case when the input is a second-order stationary Markov chain, whose transition probability matrix (indexed by $00, 01$ and 10 only since 11 is prohibited by the $(1, \infty)$-RLL constraint) is given by

$$\begin{bmatrix} p & 1-p & 0 \\ 0 & 0 & 1 \\ q & 1-q & 0 \end{bmatrix},$$

where $0 < p, q < 1$. For this case, from the Birch lower bound (see, e.g., Lemma 4.5.1 of [8]), we have

$$H(Y_6|Y_5, Y_4, Y_3, X_2, X_1) - H(\varepsilon) \leq H(Y) - H(\varepsilon) = I(X;Y).$$

It can then be checked by direct computation that for the case $\varepsilon = 0.1$, when $p \approx 0.597275$ and $q \approx 0.614746$ we have

$$H(Y_6|Y_5, Y_4, Y_3, X_2, X_1) - H(\varepsilon) \approx 0.442329,$$

which is a lower bound on the second-order Markov capacity yet strictly larger than the upper bound on the first-order Markov capacity given in (30). Hence we can draw the conclusion that for the BEC channel with Markovian inputs under the $(1, \infty)$-RLL constraint, an increase of the Markov order of the input process from 1 to 2 does increase the channel capacity.

2) A Noiseless Channel With Two States: In this section, we consider a noiseless FSC with two channel states, for which we show that Algorithm III.3 can be applied to show that higher-order memory can yield larger Markov capacity.

More precisely, the channel input $\{X_n\}_{n=1}^{\infty}$ is a stationary process taking values from the alphabet $\mathcal{A} = \{0, 1\}$ and, except at time 0, the channel state $\{S_n\}_{n=1}^{\infty}$ is determined by the channel input, that is, $S_n = X_n$, $n = 1, 2, \ldots$. The channel is characterized by the following input-output equation:

$$Y_n = \phi(S_{n-1}, X_n), \quad n = 1, 2, \ldots, \quad (31)$$

where ϕ is a deterministic function with $\phi(0,0) = 1, \phi(0,1) = 0, \phi(1,0) = 0$ and $\phi(1,1) = 0$. We remark that this channel is a unifilar FSC, whose feedback capacity corresponds to the optimal reward of a dynamical program (see [27] for more details). In this section, we consider the case when there is no feedback. In this case, ϕ naturally induces a sliding block code that maps the full $\mathcal{A}$-shift $\mathcal{S}$ onto the shift of finite type $\mathcal{S}_{\mathcal{F}}$, where the forbidden set $\mathcal{F}$ is $\{101\}$. It can be readily verified that the Shannon capacity of (31) is equal to its stationary capacity[2] [12], which can be computed as the largest eigenvalue of the adjacency matrix of the 3rd higher block shift of $\mathcal{S}_{\mathcal{F}}$ and is approximately equal to 0.562399 (see Chapter 4 and 13 of [22] for more details). In what follows, we will focus on the Markov capacity of (31); more specifically, we will compute the Markov capacity when the input $\{X_n\}_{n=1}^{\infty}$ is an i.i.d. process and a first-order stationary Markov chain, which will be compared with the Shannon capacity.

It can be easily verified that the mutual information rate of (31) can be computed as

$$\begin{aligned} I(X;Y) &= \lim_{k\to\infty} H(Y_{k+1}|Y_1^k) - \frac{1}{k} H(Y_1^k|X_1^k) \\ &= \lim_{k\to\infty} H(Y_{k+1}|Y_1^k) = H(Y). \end{aligned}$$

When $\{X_n\}_{n=1}^{\infty}$ is a stationary Markov chain, the output $\{Y_n\}_{n=1}^{\infty}$ is a hidden Markov chain with an unambiguous symbol whose entropy rate can be computed by the following formula [15]:

$$H(Y) = \sum_{n=1}^{\infty} P(Y_1^n = (1, \underbrace{0, \ldots, 0}_{n-1})) H(Y_{n+1}|Y_1^n = (1, \underbrace{0, \ldots, 0}_{n-1})). \quad (32)$$

This formula will play a key role in our analysis detailed below.

We first consider the degenerate case that $\{X_n\}_{n=1}^{\infty}$ is an i.i.d. process. Letting θ denote $P(X_1 = 0)$, we note that the Markov chain $\{(X_{n-1}, X_n)\}_{n=2}^{\infty}$ has the following transition probability matrix (indexed by $00, 01, 10, 11$)

$$\begin{bmatrix} \theta & 1-\theta & 0 & 0 \\ 0 & 0 & \theta & 1-\theta \\ \theta & 1-\theta & 0 & 0 \\ 0 & 0 & \theta & 1-\theta \end{bmatrix},$$

whose left eigenvector corresponding to the largest eigenvalue is

$$(\pi_1(\theta), \pi_2(\theta), \pi_3(\theta), \pi_4(\theta)) = (\theta^2, \theta(1-\theta), \theta(1-\theta), (1-\theta)^2).$$

Using (32), we have

$$\begin{aligned} H(Y) = &- \sum_{l=0}^{\infty} \pi_1(\theta) \mathbf{r}(B_\theta)^l \mathbf{1} \log \frac{\mathbf{r}(B_\theta)^l \mathbf{1}}{\mathbf{r}(B_\theta)^{l-1} \mathbf{1}} \\ &- \sum_{l=0}^{\infty} \pi_1(\theta) \mathbf{r}(B_\theta)^{l-1} \mathbf{c} \log \frac{\mathbf{r}(B_\theta)^{l-1} \mathbf{c}}{\mathbf{r}(B_\theta)^{l-1} \mathbf{1}}, \end{aligned}$$

where $\mathbf{r} = (1-\theta, 0, 0)$, $\mathbf{c} = (0, \theta, 0)^T$, $\mathbf{1} = (1, 1, 1)^T$,

$$B_\theta = \begin{bmatrix} 0 & \theta & 1-\theta \\ 1-\theta & 0 & 0 \\ 0 & \theta & 1-\theta \end{bmatrix},$$

and both $\mathbf{r}(B_\theta)^{-1}\mathbf{1}, \mathbf{r}(B_\theta)^{-1}\mathbf{c}$ should be interpreted as 1.

[2]According to our ongoing research, the stationary capacity of this channel is indeed achieved by a second-order Markov input.

Setting $f(\theta) \triangleq H(Y)$, we note that

$$f(\theta) = \lim_{k\to\infty} f_k(\theta),$$

where

$$f_k(\theta) \triangleq -\sum_{l=0}^{k} \pi_1(\theta)\mathbf{r}(B_\theta)^l \mathbf{1} \log \frac{\mathbf{r}(B_\theta)^l \mathbf{1}}{\mathbf{r}(B_\theta)^{l-1}\mathbf{1}} - \sum_{l=0}^{k} \pi_1(\theta)\mathbf{r}(B_\theta)^{l-1}\mathbf{c} \log \frac{\mathbf{r}(B_\theta)^{l-1}\mathbf{c}}{\mathbf{r}(B_\theta)^{l-1}\mathbf{1}}, \quad k \geq 0.$$

Similarly as in the previous example, we can show that

$$\max_{\theta\in[0,0.41]\cup[0.89,1]} f(\theta) < \max_{\theta\in[0.41,0.89]} f(\theta),$$

which means that $f(\theta)$ will achieve its maximum within the interior of $[0.4, 0.9]$. Moreover, through tedious but similar evaluations as in the previous example, we can choose (below, rather than a constant, N is a polynomial in k, but the proof of Theorem III.10 carries over almost verbatim)

$$k_0 = 120, \quad N = (374.945k^2 + 6207.73k + 46587.2), \\ \rho = 0.875, \quad m = 1.2, \quad M = 10.37.$$

Though the function $f(\theta)$ is not concave near $\theta = 0$, tedious yet straightforward computation indicates that $f''(\theta) \leq f''_{120}(\theta) + N\rho^{120} < 0$ for any $\theta \in [0.4, 0.9]$, which immediately implies that $f(\theta)$ is strongly concave within the interior of the interval $[0.4, 0.9]$. Then, similarly as in Section III-B1, one verifies that, when applied to the channel in (31), Algorithm III.3 achieves exponential accuracy in polynomial time.

Letting $\alpha = 0.4, \beta = 0.9$, we apply our algorithm to the sequence $\{f_k(\theta) : k \geq 120\}$ with $\Theta \triangleq [0.4, 0.9]$, $\theta_0 = 0.5$, $\eta = \xi = 0.901061$, and we obtain that

$$\theta_{450} \approx 0.6257911, \quad f_{450}(\theta_{450}) \approx 0.4292892.$$

Now from (4), (26) and the fact that $\hat{\theta}_{n_k} = \theta_k$, we conclude

$$|f_{450}(\theta_{450}) - f(\theta^*)| \\ \leq |f_{450}(\theta_{450}) - f(\theta_{450})| + |f(\theta_{450}) - f(\theta^*)| \leq 0.0001745,$$

which further implies that

$$0.4291146 \leq f(\theta^*) \leq 0.4294638 \tag{33}$$

for the i.i.d. case.

Now, we consider the case that $\{X_n\}_{n=1}^{\infty}$ is a genuine first-order stationary Markov process, and assume the Markov chain $\{(X_{n-1}, X_n)\}_{n=2}^{\infty}$ has the following transition probability matrix (indexed by $00, 01, 10, 11$)

$$\begin{pmatrix} p & 1-p & 0 & 0 \\ 0 & 0 & q & 1-q \\ p & 1-p & 0 & 0 \\ 0 & 0 & q & 1-q \end{pmatrix},$$

where $0 < p, q < 1$. Again, straightforward computation shows that for $p \approx 0.674521, q \approx 0.595176$, $H(Y_4|Y_3, X_2, X_1)$ is approximately 0.513259, which gives a lower bound on $H(Y)$. Comparing this lower bound with the upper bound in (33), we conclude that the capacity is increased when increasing the Markov order of the input from 0 to 1. Finally, we also point out that direct evaluation of a trivial upper bound (for the first-order Markov capacity of (31)) gives

$$\max_{p,q} H(Y_6|Y_5, Y_4, Y_3, Y_2, Y_1) \approx 0.548481$$

for $p \approx 0.629902, q \approx 0.734121$. Comparing this upper bound with 0.562399, the Shannon capacity given at the beginning of this section, we also conclude that the Shannon capacity of (31) cannot be achieved by any first-order Markovian input.

IV. The Second Algorithm: Without Concavity

In this section, we consider the optimization problem (3) for the case when f may not be concave.

For a non-convex optimization problem with a continuously differentiable target function f and a bounded domain, conventionally there are two major methods for finding its solution: the Frank-Wolfe method [3] and the method through the Łojasiewicz inequality (see, e.g., [1]). However, both of these methods in general tend to fail in our setting: for the Frank-Wolfe method, the computation for finding the feasible ascent direction and the verification of the relevant gradient condition (which is necessary for the convergence of this method) both depend on the existence of an exact formula for ∇f and a tractable description of Θ, which is however not available in our case; on the other hand, due to the fact that our target function is the limit of a sequence of approximating functions, the method through the Łojasiewicz inequality necessitates a "uniform" version of the Łojasiewicz inequality over all sequences of approximating functions, which does not seem to hold true in our setting.

Motivated by Algorithm III.3, we propose in the following our second algorithm to efficiently solve the optimization problem (3) when the target function may not be concave. Except for using the sequence $\{\nabla f_k\}_{k=0}^{\infty}$ as the ascent direction in each iteration, an additional check condition is proposed for the choice of the step size. This check condition is chosen carefully to ensure an appropriate pace for the decay of ∇f_k, which turns out to be crucial for the convergence of this algorithm.

Similarly as in Section III, we need the following lemma before presenting our second algorithm.

Lemma IV.1: Assume the function $f(\theta)$ has s stationary points $\{\theta_i^*\}_{i=1}^{s}$ which are all contained in Θ°, and that $f(\theta)$ achieves its maximum in Θ°. If, for each k, $f_k(\theta)$ also has finitely many stationary points which are all contained in Θ°, then for any fixed $b \in \mathbb{R}$ with $0 < b < 1$, there exists a non-negative integer k_0 such that

(a) $\rho^{1/3} + \rho^{2k_0/3} < 1$ and $\frac{2N\rho^{k_0}}{1-\rho} \leq \frac{\delta}{8}$, where $\delta \triangleq \max_{\theta_i^*:1\leq i\leq s} f(\theta_i^*) - \max_{\theta\in\partial\Theta} f(\theta) > 0$;

(b) There exists a $y_0 \in \mathbb{R}$ such that

$$\emptyset \subsetneq B_{k_0} \subseteq C_{k_0} \subseteq \Theta^\circ, \; A_{k_0} \cap B_{k_0} \neq \emptyset, \; \mathrm{dist}(C_{k_0}, \partial\Theta) > 0,$$

where

$$A_{k_0} \triangleq \left\{\theta \in \Theta^\circ : \|\nabla f_{k_0}(\theta)\|_2 \geq \frac{2N\rho^{k_0/3}}{1-b}\right\},$$

$$B_{k_0} \triangleq \{\theta \in \Theta : f_{k_0}(\theta) \geq y_0\},$$
$$C_{k_0} \triangleq \left\{\theta \in \Theta : f_{k_0}(\theta) \geq y_0 - \frac{\delta}{8}\right\}.$$

Note that A_{k_0} depends on b, whereas B_{k_0} and C_{k_0} do not.

Proof: By replacing what was assumed to be the unique maximum of f with $\max_{\theta_i^*:1\leq i\leq s} f(\theta_i^*)$, a similar argument as in the proof of Lemma III.1(c) yields that there exists a $y_0 < y^* - \frac{\delta}{4}$ such that for all sufficiently large k, $\emptyset \subsetneq B_k \subseteq C_k \subseteq \Theta^\circ$ and $\mathrm{dist}(C_k, \Theta^c) > 0$, where

$$y^* = \max_{\theta_i^*:1\leq i\leq s} f(\theta_i^*),$$
$$B_k \triangleq \{\theta \in \Theta : f_k(\theta) \geq y_0\},$$
$$C_k \triangleq \left\{\theta \in \Theta : f_k(\theta) \geq y_0 - \frac{\delta}{8}\right\}.$$

Now, for any k and any fixed $0 < b < 1$, let

$$A_k \triangleq \left\{\theta \in \Theta^\circ : \|\nabla f_k(\theta)\|_2 \geq \frac{2N\rho^{k/3}}{1-b}\right\}.$$

We claim that for all large enough k, $A_k \cap B_k \neq \emptyset$. To see this, define

$$D_k \triangleq \left\{\theta \in \Theta^\circ : \|\nabla f(\theta)\|_2 \geq \frac{2N\rho^{k/3}}{1-b} + N\rho^k\right\},$$
$$B' \triangleq \left\{\theta \in \Theta : f(\theta) \geq y_0 + \frac{\delta}{8}\right\}.$$

It then follows from (4), the continuity of f and the fact $y_0 + \delta/8 < y^*$ that $D_k \subseteq A_k, B' \subseteq B_k$ for all large enough k and B' has a non-empty interior. Observing that D_k^c converges to the finite set consisting of all stationary points of f, we deduce that $D_k \cap B' \neq \emptyset$ and therefore $A_k \cap B_k \neq \emptyset$ for sufficiently large k and therefore establish the claim. Finally, it immediately follows from this claim and the observation that (a) trivially holds for k_0 sufficiently large that there exists a non-negative integer k_0 such that (a) and (b) are both satisfied. ■

Recalling that f and each f_k are assumed to have finitely many stationary points in Θ°, we now present our second algorithm.

Algorithm IV.2 (The Second Modified Gradient Ascent Algorithm):

Step 0. Choose $0 < b < 1$. Choose k_0, y_0 and such that the conditions in Lemma IV.1 are satisfied. Set $k = 0$, $g_0 = f_{k_0}$ and choose $\alpha \in (0, 0.5)$, $\beta \in (0, 1)$, $\theta_0 \in A_{k_0} \cap B_{k_0}$ where A_{k_0} and B_{k_0} are defined as in Lemma IV.1.

Step 1. Increase k by 1. Set $t = 1$ and $g_k = f_{k+k_0}$.

Step 2. Set

$$\tau = \theta_{k-1} + t\nabla g_{k-1}(\theta_{k-1}).$$

If $\tau \notin \Theta^\circ$ or

$$\|\nabla g_k(\tau)\|_2 < \frac{2N\rho^{k/3}}{1-b}$$

or

$$g_k(\tau) < g_k(\theta_{k-1}) + \alpha t \|\nabla g_{k-1}(\theta_{k-1})\|_2^2,$$

set $t = \beta t$ and go to Step 2, otherwise set $\theta_k = \tau$ and go to Step 1.

Remark IV.3: The constants in Step 0 are chosen to ensure the convergence of the algorithm. The existence of θ_0 follows from Lemma IV.1 (b). Furthermore, as stated in Remark III.4, N and ρ are chosen such that (4) holds and the choices of them can be derived from Section IV of [15] in practice.

Remark IV.4: In Step 2, for any feasible k, one of the necessary conditions for updating the value of θ_k is

$$\|\nabla g_k(\tau)\|_2 \geq \frac{2N\rho^{k/3}}{1-b}.$$

This is a key condition imposed to make sure that $\|\nabla g_k(\tau)\|_2$ is not too small and thereby the algorithm will not prematurely converge to a non-stationary point.

A. Convergence Analysis

To conduct the convergence analysis of Algorithm IV.2, we need to reformulate the algorithm via possible relabelling of the functions $\{g_k\}_{k=0}^\infty$ and iterates $\{\theta_k\}_{k=0}^\infty$ similarly as in Section III-A. For ease of presentation only, we assume in the remainder of this section that such a relabelling is not needed and thereby k actually records the number of times that Step 2 has been executed.

The following theorem asserts the convergence of Algorithm IV.2 under some regularity conditions.

Theorem IV.5: Under the same assumptions as in Lemma IV.1,

$$\lim_{k\to\infty} g_k(\theta_k) \text{ exists and } \|\nabla g_k(\theta_k)\|_2 \to 0,$$

where $g_k(\theta_k)$ is defined in Algorithm IV.2.

Proof: Similarly as in Section III-A, define

$$T_1(k) \triangleq \inf\{p \in \mathbb{Z} : \theta_{k-1} + \beta^p \nabla g_{k-1}(\theta_{k-1}) \in \Theta^\circ\},$$
$$\hat{T}(k) \triangleq \inf\left\{q \in \mathbb{Z} : q \geq T_1(k), \right.$$
$$\left. \|\nabla g_k(\theta_{k-1} + \beta^q \nabla g_{k-1}(\theta_{k-1}))\|_2 \geq \frac{2N\rho^{(k+k_0)/3}}{1-b}\right\},$$
$$T(k) \triangleq \inf\{r \in \mathbb{Z} : r \geq \hat{T}(k),$$
$$g_k(\theta_{k-1}+\beta^r \nabla g_{k-1}(\theta_{k-1})) \geq g_k(\theta_{k-1}) + \alpha\beta^r \|\nabla g_{k-1}(\theta_{k-1})\|_2^2\},$$

and

$$T_2(k) \triangleq \hat{T}(k) - T_1(k), \quad T_3(k) \triangleq T(k) - \hat{T}(k).$$

In other words, for each k, $T_1(k)$ can be regarded as the number of times that Step 2 of Algorithm IV.2 has been executed before the condition $\tau \in \Theta^\circ$ is met; $T_2(k)$ can be regarded as the number of additional times that Step 2 of Algorithm IV.2 has been executed before the condition

$$\|\nabla g_k(\theta_{k-1} + \beta^q \nabla g_{k-1}(\theta_{k-1}))\|_2 \geq \frac{2N\rho^{(k+k_0)/3}}{1-b}$$

is also met, and $T_3(k)$ can be regarded as the number of additional times that Step 2 of Algorithm IV.2 has been executed before the Armijo condition

$$g_k(\theta_{k-1} + \beta^r \nabla g_{k-1}(\theta_{k-1})) \geq g_k(\theta_{k-1}) + \alpha\beta^r \|\nabla g_{k-1}(\theta_{k-1})\|_2^2$$

is also met. The well-definedness of $\hat{T}(k)$ is based on the fact that if $\theta_{k-1}+\beta^p\nabla g_{k-1}(\theta_{k-1})\in\Theta^\circ$ for some non-negative integer p, then the same inequality also holds for any integer $p'>p$; and the well-definedness of $T(k)$ will be postponed to Step 2 of this proof detailed below.

The remainder of the proof consists of 5 steps, with the first three devoted to establishing the uniform boundedness of $T_1(k)$, $T_2(k)$ and $T_3(k)$ and thus that of $T(k)$.

Step 1: Uniform boundedness of $T_2(k)$. As in the proof of Theorem III.10, it can be readily verified that $T_1(k)<\infty$ for all $k\geq 0$. Hence, when considering $T_2(k)$, we assume that $\tau=\theta_{k-1}+\beta^q\nabla g_{k-1}(\theta_{k-1})$ is already in Θ°.

In order to prove the uniform boundedness of $T_2(k)$, we proceed by way of induction. First of all, by the definition of g_0 and the choice of θ_0, we have $\|\nabla g_0(\theta_0)\|_2\geq 2N\rho^{k_0/3}/(1-b)$. Now, assuming that

$$\|\nabla g_{k-1}(\theta_{k-1})\|_2\geq\frac{2N\rho^{(k_0+k-1)/3}}{1-b} \tag{34}$$

holds for some non-negative integer k, we will derive a sufficient condition on β^q such that $\|\nabla g_k(\tau)\|_2\geq 2N\rho^{(k_0+k)/3}/(1-b)$, where we recall that τ is defined by

$$\tau=\theta_{k-1}+\beta^q\nabla g_{k-1}(\theta_{k-1}). \tag{35}$$

To this end, we first note that by the Taylor series expansion, there exist ξ and $\hat{\xi}$ in Θ° such that

$$\begin{aligned}&g_k(\tau)-g_k(\theta_{k-1})\\&=\nabla g_k(\tau)^T(\tau-\theta_{k-1})-(\theta_{k-1}-\tau)^T\frac{\nabla^2 g_k(\xi)}{2}(\theta_{k-1}-\tau)\end{aligned}$$

and

$$\begin{aligned}&g_k(\tau)-g_k(\theta_{k-1})\\&=\nabla g_k(\theta_{k-1})^T(\tau-\theta_{k-1})+(\tau-\theta_{k-1})^T\frac{\nabla^2 g_k(\hat{\xi})}{2}(\tau-\theta_{k-1}),\end{aligned}$$

which immediately imply that

$$\begin{aligned}&\nabla g_k(\tau)^T(\tau-\theta_{k-1})-(\theta_{k-1}-\tau)^T\frac{\nabla^2 g_k(\xi)}{2}(\theta_{k-1}-\tau)\\&=\nabla g_k(\theta_{k-1})^T(\tau-\theta_{k-1})+(\tau-\theta_{k-1})^T\frac{\nabla^2 g_k(\hat{\xi})}{2}(\tau-\theta_{k-1}).\end{aligned} \tag{36}$$

Noting that $\left\|\nabla^2 g_k(\xi)\right\|_2\leq M$ for all $\xi\in\Theta^\circ$ and

$$\begin{aligned}&\|\nabla g_k(\theta)-\nabla g_{k-1}(\theta)\|_2\\&=\|\nabla f_{k+k_0}(\theta)-\nabla f_{k+k_0-1}(\theta)\|_2\leq N\rho^{k+k_0}\end{aligned} \tag{37}$$

for all $\theta\in\Theta^\circ$, we deduce from (36) that

$$\begin{aligned}&\|\nabla g_k(\tau)\|_2\|\tau-\theta_{k-1}\|_2\\&\geq\nabla g_k(\theta_{k-1})^T(\tau-\theta_{k-1})-M\|\tau-\theta_{k-1}\|_2^2\\&\geq\nabla g_{k-1}(\theta_{k-1})^T(\tau-\theta_{k-1})\\&\quad-N\rho^{k+k_0}\|\tau-\theta_{k-1}\|_2-M\|\tau-\theta_{k-1}\|_2^2.\end{aligned} \tag{38}$$

Clearly, it follows from (35) that the vectors $\nabla g_{k-1}(\theta_{k-1})$ and $\tau-\theta_{k-1}$ have the same direction, which means that (38) can be rewritten as

$$\begin{aligned}&\|\nabla g_k(\tau)\|_2\|\tau-\theta_{k-1}\|_2\\&\geq\|\nabla g_{k-1}(\theta_{k-1})\|_2\|\tau-\theta_{k-1}\|_2\\&\quad-N\rho^{k+k_0}\|\tau-\theta_{k-1}\|_2-M\|\tau-\theta_{k-1}\|_2^2.\end{aligned}$$

Simplifying this inequality and recalling that $\tau-\theta_{k-1}=\beta^q\nabla g_{k-1}(\theta_{k-1})$, we have

$$\begin{aligned}\|\nabla g_k(\tau)\|_2&\geq(1-M\beta^q)\|\nabla g_{k-1}(\theta_{k-1})\|_2-N\rho^{k+k_0}\\&\geq(1-M\beta^q)\frac{2N\rho^{(k_0+k-1)/3}}{1-b}-N\rho^{k+k_0}.\end{aligned} \tag{39}$$

Now, using the fact $1-\rho^{1/3}-\rho^{2k_0/3}>0$ (see Lemma IV.1 (a)), (34) and (39), we conclude that the condition

$$\beta^q\leq\frac{1-\rho^{1/3}-\rho^{2k_0/3}}{M} \tag{40}$$

is sufficient for $\|\nabla g_k(\tau)\|_2\geq 2N\rho^{(k+k_0)/3}/(1-b)$. In other words, the induction argument successfully proceeds as long as (40) holds, and therefore $T_2(k)$ can be uniformly bounded as below:

$$T_2(k)\leq\max\left\{0,\frac{\log\left(\left(1-\rho^{1/3}-\rho^{2k_0/3}\right)/M\right)}{\log\beta}+1\right\}. \tag{41}$$

Step 2: Uniform boundedness of $T_3(k)$. First note from (40) that if the inequality

$$\|\nabla g_k(\theta_{k-1}+\beta^q\nabla g_{k-1}(\theta_{k-1}))\|_2\geq\frac{2N\rho^{(k+k_0)/3}}{1-b}$$

holds for some non-negative integer q, then it remains true for any integer $q'>q$. This observation justifies the well-definedness of $T_3(k)$. Moreover, due to the boundedness of $T_2(k)$ for each k (in fact, it is uniformly bounded), we can assume without loss of generality that $\|\nabla g_k(\tau)\|_2\geq 2N\rho^{(k+k_0)/3}/(1-b)$ is already satisfied, where $\tau=\theta_{k-1}+\beta^r\nabla g_{k-1}(\theta_{k-1})$, before we proceed to establish the uniform boundedness of $T_3(k)$.

By the Taylor series expansion formula and (37), we have

$$\begin{aligned}g_k(\tau)&\geq g_k(\theta_{k-1})+\beta^r\nabla g_k(\theta_{k-1})^T\nabla g_{k-1}(\theta_{k-1})\\&\quad-\frac{M\beta^{2r}}{2}\|\nabla g_{k-1}(\theta_{k-1})\|_2^2\\&\geq g_k(\theta_{k-1})+\beta^r\|\nabla g_{k-1}(\theta_{k-1})\|_2^2\\&\quad-\frac{M\beta^{2r}}{2}\|\nabla g_{k-1}(\theta_{k-1})\|_2^2-N\rho^{k+k_0}\beta^r\|\nabla g_{k-1}(\theta_{k-1})\|_2,\end{aligned}$$

where $\tau=\theta_{k-1}+\beta^r\nabla g_{k-1}(\theta_{k-1})$. It then follows that the condition

$$\beta^r\leq\frac{1}{M}-\frac{2N\rho^{k+k_0-1}}{M\|\nabla g_{k-1}(\theta_{k-1})\|_2}$$

is sufficient to ensure that

$$g_k(\tau)\geq g_k(\theta_{k-1})+\alpha\beta^r\|\nabla g_{k-1}(\theta_{k-1})\|_2^2. \tag{42}$$

Recalling that

$$\|\nabla g_{k-1}(\theta_{k-1})\|_2\geq\frac{2N\rho^{(k+k_0-1)/3}}{1-b}\geq\frac{2N\rho^{k+k_0-1}}{1-b}$$

with $0 < b < 1$ fixed, we deduce that the condition $\beta^r \leq b/M$ is sufficient for (42). In other words, we have

$$T_3(k) \leq \max\left\{0, \frac{\log b - \log M}{\log \beta} + 1\right\}. \quad (43)$$

Step 3: Uniform boundedness of $T_1(k)$ and $T(k)$. In this step, we will show that $T_1(k)$ is uniformly bounded over all k. This, together with the established fact that $T_2(k)$ and $T_3(k)$ are both uniformly bounded, immediately implies the uniform boundedness of $T(k)$ over all k.

From Algorithm IV.2,

$$g_k(\theta_k) \geq g_k(\theta_{k-1}) + \alpha t \|\nabla g_{k-1}(\theta_{k-1})\|_2^2$$

for all $k \geq 0$, where $\theta_k = \theta_{k-1} + \beta^{T(k)} \nabla g_{k-1}(\theta_{k-1})$. Using (4), we have

$$\begin{aligned} g_0(\theta_k) &\geq g_k(\theta_k) - \frac{N\rho^{k_0+1}}{1-\rho} \geq g_k(\theta_{k-1}) - \frac{N\rho^{k_0+1}}{1-\rho} \\ &\geq g_{k-1}(\theta_{k-1}) - N\rho^{k+k_0} - \frac{N\rho^{k_0+1}}{1-\rho}, \end{aligned}$$

from which we arrive at

$$g_0(\theta_k) \geq g_0(\theta_0) - \sum_{k=1}^{\infty} N\rho^{k+k_0} - \frac{N\rho^{k_0+1}}{1-\rho} \geq g_0(\theta_0) - \frac{2N\rho^{k_0}}{1-\rho}, \quad (44)$$

for all $k \geq 0$. Recalling from Lemma IV.1 and Step 0 of Algorithm IV.2 that

$$\theta_0 \in B_{k_0} = \{x \in \Theta : f_{k_0}(x) \geq y_0\} = \{x \in \Theta : g_0(x) \geq y_0\},$$

we deduce from (44) and Lemma IV.1 that for all $k \geq 0$,

$$\theta_k \in \left\{x : g_0(x) \geq y_0 - \frac{2N\rho^{k_0}}{1-\rho}\right\} \subseteq C_{k_0} \subseteq \Theta^\circ,$$

where C_{k_0} is defined in Lemma IV.1 (b) and $\mathrm{dist}(C_{k_0}, \partial\Theta) > 0$. Hence, for any non-negative integer p such that $p \geq \log(\mathrm{dist}(C_{k_0}, \partial\Theta)/M)/\log\beta$, we have $\theta_{k-1} + \beta^p \nabla g_k(\theta_{k-1}) \in \Theta^\circ$, establishing the following uniform bound

$$T_1(k) \leq \max\left\{0, \frac{\log(\mathrm{dist}(C_{k_0}, \Theta^c)/M)}{\log\beta} + 1\right\}. \quad (45)$$

Finally, it is clear from (41), (43), (45) and the definition of $T(k)$ that there exists a non-negative integer B such that, for all k,

$$T(k) \leq B. \quad (46)$$

Step 4: Convergence of $g_k(\theta_k)$. It follows from (4), (46) and the fact $\|\nabla g_{k-1}(\theta_{k-1})\|_2 \geq 2N\rho^{(k+k_0-1)/3}/(1-b)$ that

$$\begin{aligned} g_k(\theta_k) &\geq g_k(\theta_{k-1}) + \alpha\beta^{B+1} \|\nabla g_{k-1}(\theta_{k-1})\|_2^2 \\ &\geq g_{k-1}(\theta_{k-1}) + \alpha\beta^{B+1} \|\nabla g_{k-1}(\theta_{k-1})\|_2^2 - N\rho^{k+k_0} \\ &\geq g_{k-1}(\theta_{k-1}) + \frac{4\alpha\beta^{B+1}N^2\rho^{2(k+k_0-1)/3}}{(1-b)^2} - N\rho^{k+k_0}. \end{aligned}$$

Observing that if k is large enough,

$$\frac{4\alpha\beta^{B+1}N^2\rho^{2(k+k_0-1)/3}}{(1-b)^2} \geq N\rho^{k+k_0},$$

we deduce that $g_k(\theta_k) \geq g_{k-1}(\theta_{k-1})$ for sufficiently large k. Noting that (4) and the definition of g_k imply that there exists a $C > 0$ such that $g_k(\theta_k) \leq C$ for all k, we conclude that $\lim_{k\to\infty} g_k(\theta_k)$ exists.

Step 5: $\|\nabla g_k(\theta_k)\|_2 \to 0$. Since

$$g_k(\theta_k) \geq g_{k-1}(\theta_{k-1}) + \alpha\beta^{B+1} \|\nabla g_{k-1}(\theta_{k-1})\|_2^2 - N\rho^{k+k_0},$$

we have

$$\sum_{k=1}^{n-1} \alpha\beta^{B+1} \|\nabla g_{k-1}(\theta_{k-1})\|_2^2 \leq g_n(\theta_n) - g_0(\theta_0) + \sum_{k=1}^{n-1} N\rho^{k+k_0},$$

which, together with the uniform boundedness of $\{g_k(\theta_k)\}_{k=0}^{\infty}$, yields

$$\sum_{k=1}^{\infty} \alpha\beta^{B+1} \|\nabla g_{k-1}(\theta_{k-1})\|_2^2 < \infty.$$

Hence, $\lim_{n\to\infty} \|\nabla g_{k-1}(\theta_{k-1})\|_2 = 0$. The proof of the theorem is thus complete. ■

B. Applications of Algorithm IV.2

1) Gilbert-Elliott Channel: In this section, we consider a Gilbert-Elliott channel with a first-order Markovian input under the $(1,\infty)$-RLL constraint.

A Gilbert-Elliott channel [24] is a special FSC with two states: a good state g and a bad state b. The state process $\{S_n\}_{n=0}^{\infty}$ is a stationary first-order Markov chain and the channel alternates between two binary symmetric channels (BSCs) according to the channel state. More precisely, let $\{S_n\}_{n=0}^{\infty}$ be the state process which is a stationary Markov chain with alphabet $\{g, b\}$ and parameters

$$p_g \triangleq P(S_l = g | S_{l-1} = b), \quad p_b \triangleq P(S_l = b | S_{l-1} = g).$$

Then the Gilbert-Elliott channel is characterized by the input-output equation

$$Y_n = X_n \oplus E_n, \quad n = 1, 2, \cdots, \quad (47)$$

where $\oplus$ denotes binary addition, $\{X_n\}_{n=1}^{\infty}$ is the input, $\{Y_n\}_{n=1}^{\infty}$ is the output and $\{E_n\}_{n=1}^{\infty}$ is the noise process given by

$$E_n = \begin{cases} 0, & \text{with probability } 1 - \varepsilon_g, \\ 1, & \text{with probability } \varepsilon_g, \end{cases}$$

when $S_{n-1} = g$ and

$$E_n = \begin{cases} 0, & \text{with probability } 1 - \varepsilon_b, \\ 1, & \text{with probability } \varepsilon_b, \end{cases}$$

when $S_{n-1} = 1$, for $0 \leq \varepsilon_g \leq \varepsilon_b \leq 1$. In other words, at time n, if the previous channel state S_{n-1} takes the value 0, the channel is a BSC with crossover probability ε_g, and if S_{n-1} takes the value 1, it is a BSC with crossover probability ε_b. It is worth noting that the channel is characterized by

$$p(y_n, s_n | x_n, s_{n-1}) = p(y_n | x_n, s_{n-1}) p(s_n | s_{n-1})$$

and the mutual information rate can be computed as

$$I(X;Y) = \lim_{k\to\infty} H(Y_k | Y_1^{k-1}) - H(E_k | E_1^{k-1}).$$

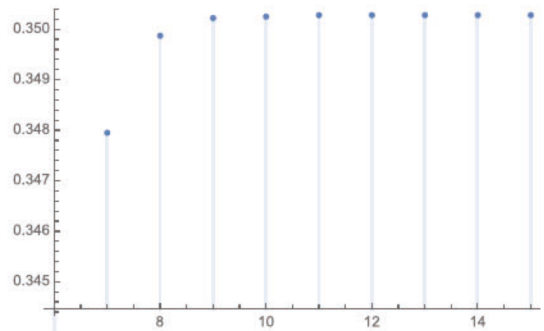

Fig. 1. Values of $f_k(\theta_k)$.

Although the capacity of the Gilbert-Elliott without the input constraint can be iteratively computed as in [24], we consider here its first-order Markov capacity under the $(1,\infty)$-RLL constraint. In particular, we consider the case for $p_g = p_b = 0.3$, $\varepsilon_g = 0.01$, $\varepsilon_b = 0.1$ and assume that $\{X_n\}_{n=1}^{\infty}$ is a first-order binary Markov chain with the transition probability matrix

$$\begin{bmatrix} 1-\theta & \theta \\ 1 & 0 \end{bmatrix}, \quad 0 \le \theta \le 1.$$

In this case, the mutual information rate $I(X(\theta);Y(\theta))$ is a function of θ and finding the channel capacity boils down to maximizing $I(X(\theta);Y(\theta))$ over $0 \le \theta \le 1$. To the best of our knowledge, the concavity of $I(X(\theta);Y(\theta))$ for this channel is not known, yet our Algorithm IV.2 can be applied to effectively maximize it. Setting

$$f_k(\theta) = H(Y_k(\theta)|Y_1^{k-1}(\theta)) - H(E_k|E_1^{k-1}),$$

and applying Algorithm IV.2 with the initial point $\theta_{k_0} = 0.2$ and $k_0 = 6$, we have obtained the simulation results shown in Figure 1.

k	θ_k	$\nabla f_k(\theta_k)$	$f_k(\theta_k)$
6	0.200000	7.059197×10^{-1}	0.281366
7	0.288240	3.606449×10^{-1}	0.327527
8	0.378401	1.049006×10^{-1}	0.347958
9	0.404626	4.271872×10^{-2}	0.349884
10	0.415306	1.862974×10^{-2}	0.350211
11	0.417635	1.346518×10^{-2}	0.350248
12	0.421001	6.053556×10^{-3}	0.350281
13	0.422514	2.742047×10^{-3}	0.350288
14	0.423200	1.246199×10^{-3}	0.350289
15	0.423511	5.672211×10^{-4}	0.350289
16	0.423653	2.583526×10^{-4}	0.350289

The table in Figure 1 shows how the iterate θ_k, the gradient $\nabla f_k(\theta_k)$ and the function $f_k(\theta_k)$ behave when k increases. We observe that, as k becomes larger, $\nabla f_k(\theta)$ and $f_k(\theta_k)$ stabilize, both very quickly, and our algorithm converges very fast for this example.

Finally, as a comparison, we note from [24] that the unconstrained capacity of a general Gilbert-Elliott channel is given by

$$C = \ln 2(1 - \lim_{n\to\infty} E[H(P(E_n = 1|E_{n-1}, S_0))]),$$

where $E[H(P(E_n = 1|E_{n-1}, S_0))]$ increases in n. Hence, using the recursive relation of $P(E_n = 1|E_{n-1}, S_0)$ given in [24], we compute that the unconstrained capacity of the Gilbert-Elliott channel with $p_g = p_b = 0.3, \varepsilon_g = 0.01$ and $\varepsilon_b = 0.1$ is approximately 0.474806, which is strictly larger than the values of $f_k(\theta_k)$ in Figure 1, which we believe are close to the first-order Markov capacity of the channel under the $(1,\infty)$-RLL constraint.

2) POST Channel: Our second example is the so-called Previous Output is the STate (POST) channel [26] with a first-order Markovian input under the $(1,\infty)$-RLL constraint.

Let $\{X_n\}_{n=1}^{\infty}$ and $\{Y_n\}_{n=1}^{\infty}$ denote the binary channel inputs and outputs, respectively. A POST(α) channel is an FSC such that

$$\begin{cases} Y_n = X_n & \text{if } X_n = Y_{n-1}, \\ Y_n = X_n \oplus Z_n & \text{if } X_n \ne Y_{n-1}, \end{cases}$$

where $\{Z_n\}_{n=1}^{\infty}$ is a Bernoulli(α) process and $\oplus$ denotes binary addition. Alternatively, letting $\{E_n\}_{n=1}^{\infty}$ denote the noise process and $\{S_n\}_{n=1}^{\infty}$ denote the state process defined by $S_n = Y_{n-1}$ for any positive integer n, we can characterize the POST(α) channel by

$$Y_n = X_n \oplus E_n \qquad n = 1, 2, \cdots$$

where $E_n = 0$ if $X_n = S_{n-1}$ and E_n is a Bernoulli(α) random variable if $X_n \ne S_{n-1}$.

It has been shown in [26] that feedback does not increase the unconstrained capacity of a POST channel, which can be computed as

$$C = \ln(1 + (1-\alpha)\alpha^{\frac{\alpha}{1-\alpha}}).$$

In this section, we focus on the Markov capacity of this channel under the $(1,\infty)$-RLL constraint and we show that our Algorithm IV.2 is applicable in this case. More precisely, we consider the case $\alpha = 0.01$ and suppose $\{X_n\}_{n=1}^{\infty}$ is a first-order Markov chain with the transition probability matrix

$$\begin{bmatrix} 1-\theta & \theta \\ 1 & 0 \end{bmatrix}, \quad 0 \le \theta \le 1.$$

In this case, similarly as in Section $IV - B1$, we have

$$\begin{aligned} I(X;Y) &= \lim_{k\to\infty} \frac{1}{k} I(X_1^k; Y_1^k) \\ &= \lim_{k\to\infty} \frac{(H(Y_1^k) - H(E_1^k, X_1^k) + H(X_1^k)}{k} \\ &= H(X_2|X_1) + \lim_{k\to\infty} \big(H(Y_k|Y_1^{k-1}) \\ &\quad - H(X_k, E_k|X_1^{k-1}, E_1^{k-1})\big), \end{aligned}$$

where the last equality follows from the fact that $\{X_n\}_{n=1}^{\infty}$ is assumed to be a first-order Markov chain. Now letting

$$f_k(\theta) = H(X_2|X_1) + H(Y_k|Y_1^{k-1}) - H(X_k, E_k|X_1^{k-1}, E_1^{k-1}),$$

we apply Algorithm IV.2 to this example with the initial point $k_0 = 4$ and $\theta_{k_0} = 0.2$ and observe the following simulation result:

k	θ_k	$\nabla f_k(\theta_k)$	$f_k(\theta_k)$
4	0.200000	7.731777×10^{-1}	0.40568718788544
5	0.362147	4.758119×10^{-2}	0.46534645257927
6	0.372126	1.648384×10^{-2}	0.46566463030699
7	0.376447	3.338922×10^{-3}	0.46570737390734
8	0.377322	6.990975×10^{-4}	0.46570913902315
9	0.377505	1.473396×10^{-4}	0.46570921652954

Again, fast convergence is observed from our simulation results.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers and the associate editor for their constructive comments which have greatly helped to improve this paper.

REFERENCES

[1] P. A. Absil, R. Mahony, and B. Andrews, "Convergence of the iterates of descent methods for analytic cost functions," *SIAM J. Optim.*, vol. 16, no. 2, pp. 531–547, Jan. 2005.

[2] S. Arimoto, "An algorithm for computing the capacity of arbitrary discrete memoryless channels," *IEEE Trans. Inf. Theory*, vol. IT-18, no. 1, pp. 14–20, Jan. 1972.

[3] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, MA, USA: Athena Scientific, 1999.

[4] R. E. Blahut, "Computation of channel capacity and rate distortion functions," *IEEE Trans. Inf. Theory*, vol. IT-18, no. 1, pp. 460–473, Jan. 1972.

[5] H. Boche, R. F. Schaefer, and H. V. Poor, "Shannon meets Turing: Non-computability and non-approximability of the finite state channel capacity," *Commun. Inf. Syst.*, vol. 20, no. 2, pp. 81–116, 2020.

[6] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge Univ. Press, 2004.

[7] J. Chen and P. H. Siegel, "Markov processes asymptotically achieve the capacity of finite-state intersymbol interference channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 1295–1303, Mar. 2008.

[8] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. New York, NY, USA: Wiley, Jul. 2006.

[9] D. Elkouss and D. Pérez-García, "Memory effects can make the transmission capability of a communication channel uncomputable," *Nature Commun.*, vol. 9, no. 1, p. 1149, Mar. 2018.

[10] G. Forney, "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Trans. Inf. Theory*, vol. IT-18, no. 3, pp. 363–378, May 1972.

[11] R. Gallager, *Information Theory and Reliable Communication*. New York, NY, USA: Wiley, 1968.

[12] R. M. Gray, *Entropy and Information Theory*. New York, NY, USA: Springer, 2011.

[13] G. Han, "Limit theorems in hidden Markov models," *IEEE Trans. Inf. Theory*, vol. 59, no. 3, pp. 1311–1328, Mar. 2013.

[14] G. Han, "A randomized algorithm for the capacity of finite-state channels," *IEEE Trans. Inf. Theory*, vol. 61, no. 7, pp. 3651–3669, Jul. 2015.

[15] G. Han and B. Marcus, "Analyticity of entropy rate of hidden Markov chains," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5251–5266, Nov. 2006.

[16] G. Han and B. H. Marcus, "Concavity of the mutual information rate for input-restricted memoryless channels at high SNR," *IEEE Trans. Inf. Theory*, vol. 58, no. 3, pp. 1534–1548, Mar. 2012.

[17] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2012.

[18] K. A. S. Immink, "EFMplus: The coding format of the multimedia compact disc," *IEEE Trans. Consum. Electron.*, vol. 41, no. 3, pp. 491–497, Aug. 1995.

[19] A. Kavčić, "On the capacity of Markov sources over noisy channels," in *Proc. IEEE Global Telecommun. Conf.*, San Antonio, TX, USA, Nov. 2001, pp. 2997–3001.

[20] Y. Li and G. Han, "Asymptotics of input-constrained erasure channel capacity," *IEEE Trans. Inf. Theory*, vol. 64, no. 1, pp. 148–162, Jan. 2017.

[21] Y. Li, G. Han, and P. H. Siegel, "On NAND flash memory channels with intercell interference," Tech. Rep.

[22] D. Lind and B. Marcus, *An Introduction to Symbolic Dynamics and Coding*. Cambridge, U.K.: Cambridge Univ. Press, 1995.

[23] B. Marcus, R. Roth, and P. H. Siegel, "Constrained systems and coding for recording channels," in *Handbook Coding Theory*. Amsterdam, The Netherlands: Elsevier, 1998.

[24] M. Mushkin and I. Bar-David, "Capacity and coding for the Gilbert–Elliott channels," *IEEE Trans. Inf. Theory*, vol. 35, no. 6, pp. 1277–1290, Nov. 1989.

[25] E. Ordentlich and T. Weissman, "Bounds on the entropy rate of binary hidden Markov processes," in *Entropy of Hidden Markov Processes and Connections to Dynamical Systems* (London Mathematical Society Lecture Note), vol. 385. Cambridge, U.K.: Cambridge Univ. Press, 2011, pp. 117–171.

[26] H. H. Permuter, H. Asnani, and T. Weissman, "Capacity of a post channel with and without feedback," *IEEE Trans. Inf. Theory*, vol. 60, no. 10, pp. 6041–6057, Oct. 2014.

[27] H. 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. 2008.

[28] J. Proakis, *Digital Communications*, 4th ed. New York, NY, USA: McGraw-Hill, 2000.

[29] 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.

[30] K. A. S. 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.

[31] H. Thapar and A. Patel, "A class of partial response systems for increasing storage density in magnetic recording," *IEEE Trans. Magn.*, vol. MAG-23, no. 5, pp. 3666–3668, Sep. 1987.

[32] P. O. Vontobel, A. Kavčić, D. M. Arnold, and H.-A. Loeliger, "A generalization of the Blahut–Arimoto algorithm to finite-state channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 1887–1918, May 2008.

[33] C. Wu, G. Han, and B. Marcus, "A deterministic algorithm for the capacity of finite-state channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2019, pp. 1897–1901.

Chengyu Wu received the B.S. degree in mathematics from Zhengzhou University in 2015 and the Ph.D. degree in mathematics from The University of Hong Kong in 2020. He is currently a Research Associate with the Department of Mathematics, The University of Hong Kong. His research interests include information theory, probability theory, ergodic theory, and symbolic dynamics.

Guangyue Han (Senior Member, IEEE) received the B.S. and M.S. degrees in mathematics from Peking University, China, in 1997 and 2000, respectively, and the Ph.D. degree in mathematics from the University of Notre Dame, USA, in 2004. After three years with the Department of Mathematics, The University of British Columbia, Canada, he joined the Department of Mathematics, The University of Hong Kong, China, in 2007. His main research areas are coding and information theory.

Venkat Anantharam (Fellow, IEEE) received the B.Tech. degree in electronics from IIT Madras in 1980, and the M.S. degree in electrical engineering, the M.A. and C.Phil. degrees in mathematics, and the Ph.D. degree in electrical engineering from UC Berkeley in 1982, 1983, 1984, and 1986, respectively. From 1986 to 1994, he was a Faculty Member of the School of EE, Cornell University, before moving to the Department of Electrical Engineering and Computer Sciences, UC Berkeley, where he is currently a Faculty Member. His research interests include communication networking, game theory, information theory, probability theory, and stochastic control.

Brian Marcus (Life Fellow, IEEE) attended the Claremont McKenna College. He received the B.A. degree in mathematics from the Pomona College in 1971 and the Ph.D. degree in mathematics from the University of California at Berkeley, under the supervision of Rufus Bowen, in 1975. He was an IBM Watson Postdoctoral Fellow, an Associate Professor in mathematics with UNC Chapel Hill, and a Research Staff Member with the IBM Almaden Research Center. Since 2002, he has been a Professor in mathematics with The University of British Columbia, worked as the Department Head, from 2002 to 2007. He has held a visiting and an adjunct associate professor positions with UC Berkeley, UC Santa Cruz, and Stanford University and worked as a Principal Ph.D. Supervisor for students with UNC Chapel Hill, UC Santa Cruz, Stanford University, and UBC. He has published more than 70 research papers in ergodic theory, symbolic dynamics, and information theory; coauthored *An Introduction to Symbolic Dynamics and Coding* (Cambridge University Press, 1995, second edition 2021) with Doug Lind; and holds 12 U.S. patents. He shared the 1993 Leonard J. Abraham Prize Paper award of the IEEE Communications Society with Paul Siegel and Jack Wolf, was an Invited IEEE ISIT Plenary Speaker in 1995, and an AMS Fellow in 2018. He served on the AMS Council (2003–2006), the CMS Board of Director (2015–1019), and the Interim Deputy Director of the Pacific Institute for the Mathematical Sciences PIMS (2016–2018). He is currently the PIMS UBC Site Director.