Fixed-Memory Capacity Bounds for the Gilbert-Elliott Channel

Yutong Han
Technical University of Munich
yutong.han@tum.de

Albert Guillén i Fàbregas
University of Cambridge
Universitat Pompeu Fabra
guillen@ieee.org

***Abstract*—We derive finite-memory upper and lower bounds to the entropy rate of binary 2-state hidden Markov models. These directly provide upper and lower bounds to the capacity of the Gilbert-Elliott channel. As the memory increases, the bounds approach the capacity of the channel. Our numerical experiments suggest that even a simple memory-1 upper bound significantly improves over the current best upper bound by Mushkin and Bar-David.**

I. Introduction

A hidden Markov model (HMM) is a stochastic process that can be seen as a noisy observation of a Markov chain. It is characterized by a set of underlying Markovian states emitting observable symbols with certain probabilities. Unlike a Markov process whose next state only depends on a fixed number of previous states, the transition probabilities for an HMM depend on the entire history of the process, assuming that the underlying Markov chain is unknown. This makes the entropy rate calculation a fundamentally hard problem, and even for a binary symmetric HMM no closed-form expression seems to be known [1].

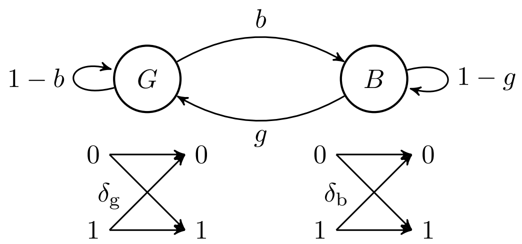

Fig. 1. Gilbert-Elliott channel model.

The Gilbert-Elliott channel [2], [3] is an elementary binary-input binary-output finite-state channel (FSC) described by the two-state Markov chain in Fig. 1. It can be interpreted as an additive noise channel where the noise sample at time i, Z_i, depends on all previous noise samples $Z^{i-1} \triangleq (Z_1, \ldots, Z_{i-1})$. This noisy process Z^i can be seen as a HMM with the transition matrix $\boldsymbol{P} \triangleq \begin{bmatrix} 1-b & b \\ g & 1-g \end{bmatrix}$ and the emission probability matrix $\boldsymbol{\Pi} \triangleq \begin{bmatrix} 1-\delta_{\mathrm{g}} & \delta_{\mathrm{g}} \\ 1-\delta_{\mathrm{b}} & \delta_{\mathrm{b}} \end{bmatrix}$. Reference [4] defined the *channel memory* as $\mu \triangleq 1 - g - b$. For $\mu > 0$, the channel has a *persistent* memory, whereas for $\mu < 0$ it has an *oscillatory* memory. When $\mu = 0$ the channel is *memoryless*, i.e. the current state is independent of all previous states. We denote the channel input, output, and noise sequences $x^n, y^n, z^n \in \{0,1\}^n$, where n is the length of the sequences.

Similarly to [4], we define $q_i \in [0,1]$ as the conditional probability

$$q_i(z^{i-1}) \triangleq \Pr(z_i = 1 | z^{i-1}) \tag{1}$$

under the assumption $0 < \delta_{\mathrm{g}} < \delta_{\mathrm{b}} < 0.5$. It was shown [2] that the following recursion holds

$$q_i \triangleq \begin{cases} \delta_{\mathrm{g}} + b(\delta_{\mathrm{b}} - \delta_{\mathrm{g}}) + \mu(q_{i-1} - \delta_{\mathrm{g}}) \frac{1-\delta_{\mathrm{b}}}{1-q_{i-1}}, & z_{i-1} = 0 \\ \delta_{\mathrm{g}} + b(\delta_{\mathrm{b}} - \delta_{\mathrm{g}}) + \mu(q_{i-1} - \delta_{\mathrm{g}}) \frac{\delta_{\mathrm{b}}}{q_{i-1}}, & z_{i-1} = 1 \end{cases} \tag{2}$$

The initial value of this recursion is

$$q_0 = \Pr(z_0 = 1) = \pi_G \delta_{\mathrm{g}} + \pi_B \delta_{\mathrm{b}} \tag{3}$$

where $[\pi_G, \pi_B] = \left[\frac{g}{g+b}, \frac{b}{g+b}\right]$ is the stationary distribution of the Markov chain that defines the channel (see Fig. 1).

In this paper, we develop a series of fixed-memory upper and lower bounds to the entropy rate of binary 2-state HMMs. These directly provide upper and lower bounds to the channel capacity of the Gilbert-Elliott channel C_{GE}. We first describe the idea using a simple memory-1 example and then generalize the result to arbitrary fixed memory. After presenting the algorithms to compute the series of bounds, we discuss through examples our observations for the persistent and oscillatory cases.

II. Memory-1 Bounds

Since both the underlying Markov chain and the emission probabilities are stationary, the hidden Markov noise process is also stationary [5]. Therefore, by using the recursion given in (2), we can compute

$$q_\infty(0) \triangleq \lim_{i\to\infty} q_i(0^{i-1}) \tag{4}$$

$$= \delta_{\mathrm{g}} + b(\delta_{\mathrm{b}} - \delta_{\mathrm{g}}) + \mu[q_\infty(0) - \delta_{\mathrm{g}}] \frac{1-\delta_{\mathrm{b}}}{1-q_\infty(0)} \tag{5}$$

which solving for $q_\infty(0)$ gives

$$q_\infty(0) = \frac{g + b + (1-g)\delta_{\mathrm{b}} + (1-b)\delta_{\mathrm{g}} - \sqrt{\Delta_0}}{2} \tag{6}$$

This work has been funded in part by the European Research Council under ERC grant agreement 725411 and by the Spanish Ministry of Economy and Competitiveness under grant PID2020-116683GB-C22.

where $\Delta_0 = [g+b+(1-g)\delta_\mathrm{b}+(1-b)\delta_\mathrm{g}]^2 - 4(g\delta_\mathrm{g}+b\delta_\mathrm{b}+\mu\delta_\mathrm{g}\delta_\mathrm{b})$. Similarly, we can write

$$q_\infty(1) \triangleq \lim_{i\to\infty} q_i(1^{i-1}) \tag{7}$$

$$= \frac{(1-b)\delta_\mathrm{g}+(1-g)\delta_\mathrm{b}+\sqrt{\Delta_1}}{2} \tag{8}$$

where $\Delta_1 = (1-b)^2\delta_\mathrm{g}^2+(1-g)^2\delta_\mathrm{b}^2+2(bg-\mu)\delta_\mathrm{g}\delta_\mathrm{b}$. For the oscillatory case it will be helpful to define the following

$$\tilde{q}_\infty(01) \triangleq \lim_{i\to\infty} q_i(0101\cdots01) \tag{9}$$

$$\tilde{q}_\infty(10) \triangleq \lim_{i\to\infty} q_i(1010\cdots10), \tag{10}$$

which can be computed in a similar manner. Observe that since $\delta_\mathrm{b} \le q_\infty \le \delta_\mathrm{g}$, all solutions admit only one solution each.

It can be observed in (2) that $q_i(z^{i-1})$ depends only on $q_{i-1}(z^{i-2})$ and the influence of earlier time instants is weaker. This loss of influence happens at an exponential rate [6], so the HMM modeling the Gilbert-Elliott channel can be efficiently simulated without having to keep track of the entire past. Since $q_i(z^{i-1}) < \delta_\mathrm{b}$, we have that

$$\frac{1-\delta_\mathrm{b}}{1-q_i(z^{i-1})} < \frac{\delta_\mathrm{b}}{q_i(z^{i-1})} \tag{11}$$

This observation together with the recursion in (2) implies that for a persistent channel ($\mu > 0$), we have that

$$q_i(z^{i-2}0) \le q_i(z^{i-2}1). \tag{12}$$

The inequality in (12) is reversed for oscillatory channels ($\mu < 0$).

Since both functions $\frac{q_i(z^{i-1})-\delta_\mathrm{g}}{1-q_i(z^{i-1})}$ and $\frac{q_i(z^{i-1})-\delta_\mathrm{g}}{q_i(z^{i-1})}$ increase monotonically with $q_i(z^{i-1})$, we know from (2) that $q_i(z^{i-1})$ increases monotonically with $q_{i-1}(z^{i-2})$ for a persistent channel whereas it decreases monotonically with $q_{i-1}(z^{i-2})$ for an oscillatory channel. Together with the observation in (12), we have that for a persistent channel, $q_i(z^{i-1})$ is bounded by the all-zero and all-one sequences, specifically

$$q_i(0^{i-1}) \le q_i(z^{i-1}) \le q_i(1^{i-1}) \tag{13}$$

and $q_i(z^{i-1})$ for the oscillatory channel is bounded by the alternating sequences, i.e.,

$$q_i(01\cdots01) \le q_i(z^{i-1}) \le q_i(10\cdots10) \tag{14}$$

In the persistent case, we know from (2) that

$$q_i(z^{i-2}0) = \delta_\mathrm{g} + b(\delta_\mathrm{b}-\delta_\mathrm{g}) + \mu(1-\delta_\mathrm{b})\frac{q_{i-1}(z^{i-2})-\delta_\mathrm{g}}{1-q_{i-1}(z^{i-2})} \tag{15}$$

$$\ge \delta_\mathrm{g} + b(\delta_\mathrm{b}-\delta_\mathrm{g}) + \mu(1-\delta_\mathrm{b})\frac{q_{i-1}(0^{i-2})-\delta_\mathrm{g}}{1-q_{i-1}(0^{i-2})} \tag{16}$$

$$= q_i(0^{i-1}) \tag{17}$$

where the inequality follows from (13) and the fact that $\frac{q_i(z^{i-1})-\delta_\mathrm{g}}{1-q_i(z^{i-1})}$ increases monotonically with $q_i(z^{i-1})$ within the domain. Similarly, for $z_{i-1}=1$ we can write

$$q_i(z^{i-2}1) = \delta_\mathrm{g} + b(\delta_\mathrm{b}-\delta_\mathrm{g}) + \mu\delta_\mathrm{b}\frac{q_{i-1}(z^{i-2})-\delta_\mathrm{g}}{q_{i-1}(z^{i-2})} \tag{18}$$

$$\ge \delta_\mathrm{g} + b(\delta_\mathrm{b}-\delta_\mathrm{g}) + \mu\delta_\mathrm{b}\left(1-\frac{\delta_\mathrm{g}}{q_{i-1}(0^{i-2})}\right) \tag{19}$$

$$= q_i(0^{i-2}1) \tag{20}$$

The upper bound can be derived similarly. By putting everything together, we have that for $a \in \{0,1\}$

$$q_i(0^{i-2}a) \le q_i(z^{i-2}a) \le q_i(1^{i-2}a). \tag{21}$$

Since the binary entropy function $h(p) \triangleq -p\log p - (1-p)\log(1-p)$ is monotonically increasing for $p \in [0, 0.5]$, we can thus write for a persistent channel that

$$h\big(q_i(0^{i-2}a)\big) \le h\big(q_i(z^{i-2}a)\big) \le h\big(q_i(1^{i-2}a)\big). \tag{22}$$

Using exactly the same arguments for oscillatory channels, we have that

$$h\big(q_{i+1}(01\cdots01a)\big) \le h\big(q_{i+1}(z^{i-1}a)\big) \le h\big(q_{i+1}(10\cdots10a)\big) \tag{23}$$

Armed with these observations, we can bound the probability of having active noise at time instant i by

$$P_{Z_i}(1) = \sum_{z^{i-1}} P_{Z^{i-2}Z_{i-1}}(z^{i-2}z_{i-1}) q_i(z^{i-2}z_{i-1}) \tag{24}$$

$$\ge \sum_{z_{i-1}} P_{Z_{i-1}}(z_{i-1}) q_i(0^{i-2}z_{i-1}) \tag{25}$$

$$= P_{Z_{i-1}}(0)q_i(0^{i-2}0) + P_{Z_{i-1}}(1)q_i(0^{i-2}1) \tag{26}$$

Since the HMM is stationary from previous discussion, we know that symbols Z_{i+1}^{i+k} have the same probability distribution as Z_1^k for all $i \ge 1$ and all $k \ge 1$. Specifically, we can write that $P_{Z_{i-1}}(1) = P_{Z_i}(1)$ for large i. Then by denoting $q_\infty(01) \triangleq \lim_{i\to\infty} q_i(0^{i-2}1)$, we express the stationary behavior of (26) as

$$p_1 \ge p_0 q_\infty(0) + p_1 q_\infty(01) \tag{27}$$

where

$$p_0 = \lim_{i\to\infty} P_{Z_i}(0), \qquad p_1 = \lim_{i\to\infty} P_{Z_i}(1) \tag{28}$$

Together with $p_0 + p_1 = 1$, we get

$$p_1 \ge \frac{q_\infty(0)}{1+q_\infty(0)-q_\infty(01)} \tag{29}$$

Similarly, by substituting the upper bound in (21) into (24), we get

$$p_1 \le \frac{q_\infty(10)}{1+q_\infty(10)-q_\infty(1)} \tag{30}$$

where $q_\infty(10) \triangleq \lim_{i\to\infty} q_i(1^{i-2}0)$.

We now proceed to use the memory-1 model to bound the entropy rate of the HMM, defined as $H_\infty(Z)$. For persistent channels, we have

$$\begin{aligned}
&H_\infty(Z)\\
&= \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} H(Z_i|Z^{i-1}) && (31)\\
&= \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\sum_{z^{i-1}} P_{Z^{i-1}}(z^{i-1})h\big(q_i(z^{i-1})\big) && (32)\\
&= \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\sum_{z^{i-2}}\sum_{z_{i-1}} P_{Z^{i-2}Z_{i-1}}(z^{i-2}z_{i-1})h\big(q_i(z^{i-2}z_{i-1})\big) && (33)\\
&\geq \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\sum_{z_{i-1}} P_{Z_{i-1}}(z_{i-1})h\big(q_i(0^{i-2}z_{i-1})\big) && (34)\\
&= \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} P_{Z_{i-1}}(0)h\big((q_i(0^{i-2}0)\big) + P_{Z_{i-1}}(1)h\big(q_i(0^{i-2}1)\big) && (35)\\
&= p_0 h\big(q_\infty(0)\big) + p_1 h\big(q_\infty(01)\big) && (36)\\
&= h\big(q_\infty(0)\big) + p_1\big[h\big(q_\infty(01)\big) - h\big(q_\infty(0)\big)\big] && (37)\\
&\geq h\big(q_\infty(0)\big) + \frac{q_\infty(0)}{1+q_\infty(0)-q_\infty(01)}\big[h\big(q_\infty(01)\big) - h\big(q_\infty(0)\big)\big] && (38)\\
&= \frac{1-q_\infty(01)}{1+q_\infty(0)-q_\infty(01)}h\big(q_\infty(0)\big)\\
&\qquad + \frac{q_\infty(0)}{1+q_\infty(0)-q_\infty(01)}h\big(q_\infty(01)\big) && (39)
\end{aligned}$$

where (32) follows from spelling out $H(Z_i|Z^{i-1})$, (34) follows from (22) and the inequality (38) is a result of (29). The corresponding upper bound can be computed using the same arguments, giving

$$\begin{aligned}
H_\infty(Z) \leq\ & \frac{1-q_\infty(1)}{1+q_\infty(10)-q_\infty(1)}h\big(q_\infty(10)\big)\\
&+ \frac{q_\infty(10)}{1+q_\infty(10)-q_\infty(1)}h\big(q_\infty(1)\big). \qquad (40)
\end{aligned}$$

Therefore, the memory-1 capacity bounds for a persistent Gilbert-Elliott channel can be written as

$$\log 2 - \overline{\boldsymbol{p}}^{\mathsf{T}}\begin{bmatrix} h\big(q_\infty(10)\big)\\ h\big(q_\infty(1)\big)\end{bmatrix} \leq C_{\mathrm{GE}} \leq \log 2 - \underline{\boldsymbol{p}}^{\mathsf{T}}\begin{bmatrix} h\big(q_\infty(0)\big)\\ h\big(q_\infty(01)\big)\end{bmatrix} \qquad (41)$$

where $\overline{\boldsymbol{p}}$ and $\underline{\boldsymbol{p}}$ are upper and lower bound of the distribution $\boldsymbol{p} \triangleq [p_0, p_1]^{\mathsf{T}}$, respectively given in (29) and (30). Likewise, we can show that the channel capacity for an oscillatory channel is bounded in the same way by replacing $q_\infty(0)$, $q_\infty(1)$, $q_\infty(01)$ and $q_\infty(10)$ by $\tilde{q}_\infty(01)$, $\tilde{q}_\infty(10)$, $\tilde{q}_\infty(011)$ and $\tilde{q}_\infty(100)$, respectively.

We illustrate the proposed bounds through an example of a persistent Gilbert-Elliott channel with parameters $b = 0.15,\ g = 0.3$ and $\delta_{\mathrm{b}} = 0.4$. The solid and dashed blue lines correspond to the upper and lower memory-1 capacity bounds respectively; the dotted line represents the simulated capacity plotted using the coin-tossing method [7] with $n = 10^6$; the dash-dotted curve stands for the generalized mutual information (GMI) of the channel [8] applied to the Gilbert-Elliott model with a memoryless decoding metric [9], and the red curve depicts the upper bound derived by Mushkin and Bar-David [4]. As was shown in [9], the GMI coincides with Mushkin and Bar-David's lower bound to the capacity. The figure shows that the GMI provides a much tighter lower bound whereas the memory-1 upper bound significantly outperforms Mushkin and Bar-David's upper bound [4]. As will be shown in Section III, lower bounds with higher memory orders outperform the GMI.

Figure 3 plots the memory-1 upper and lower bounds for an oscillatory channel with parameters $b = 0.7,\ g = 0.85$ and

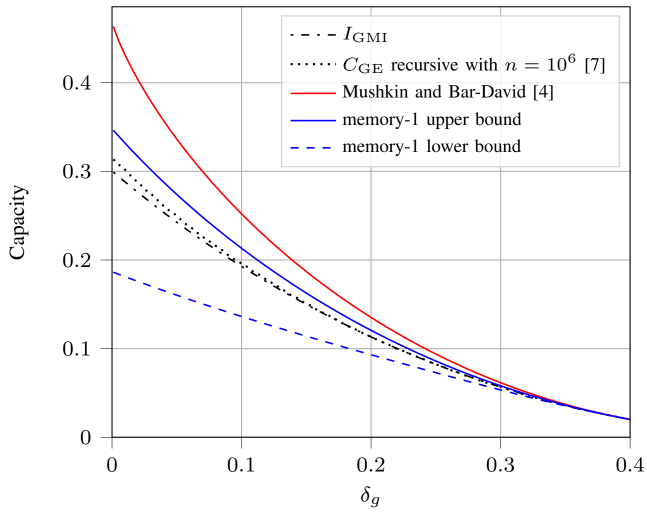

Fig. 2. Bounds for a Gilbert-Elliott channel with parameters $b = 0.15,\ g = 0.3,\ \delta_{\mathrm{b}} = 0.4$.

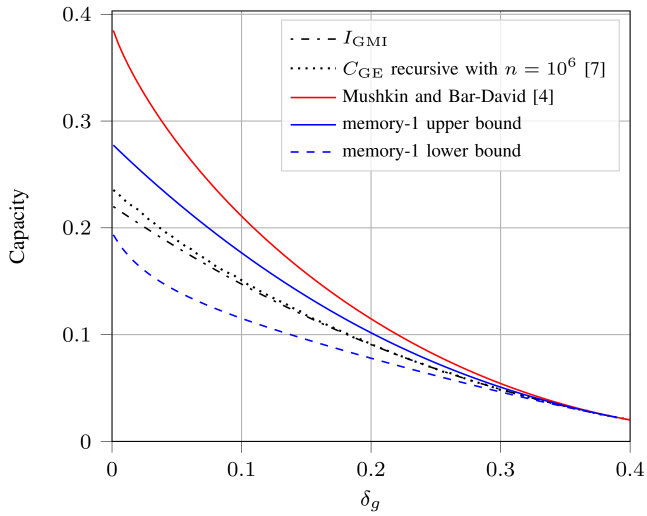

Fig. 3. Bounds for a Gilbert-Elliott channel with parameters $b = 0.7,\ g = 0.85,\ \delta_{\mathrm{b}} = 0.4$.

$\delta_{\mathrm{b}} = 0.4$. Again, although the GMI provides a better lower bound, the memory-1 upper bound significantly outperforms Mushkin and Bar-David's [4].

III. General Fixed-Memory Bounds

For persistent channels, one can easily show from (21) by induction that for any memory $m = 1, 2, 3, \ldots$

$$q_{i+m}(0^{i-1}a_1^m) \le q_{i+m}(z^{i-1}a_1^m) \le q_{i+m}(1^{i-1}a_1^m). \tag{42}$$

By defining $\bar{q}_i(z^{i-1}) \triangleq 1 - q_i(z^{i-1})$, we have

$$\bar{q}_{i+m}(1^{i-1}a_1^m) \le \bar{q}_{i+m}(z^{i-1}a_1^m) \le \bar{q}_{i+m}(0^{i-1}a_1^m). \tag{43}$$

We can then bound the joint distribution of the noise at time instant $i+1, i+2, \cdots, i+m$ using a similar approach to that described in Section II. For example, for $m = 2$, we have

$$\begin{aligned} & P_{Z_{i+1}^{i+2}}(01) \\ & = \sum_{z^i} P_{Z^i}(z^{i-2}z_{i-1}z_i)\bar{q}_{i+1}(z^{i-2}z_{i-1}z_i)q_{i+2}(z^{i-1}z_i0) && (44) \\ & \ge \sum_{z^i} P_{Z^i}(z^{i-2}z_{i-1}z_i)\bar{q}_{i+1}(1^{i-2}z_{i-1}z_i)q_{i+2}(0^{i-1}z_i0) && (45) \end{aligned}$$

Again, due to the stationarity of the HMM, we can write that

$$\begin{aligned} p_{01} \ge\ & p_{00}\bar{q}_\infty(100)q_\infty(0) + p_{01}\bar{q}_\infty(101)q_\infty(010) + \\ & p_{10}\bar{q}_\infty(110)q_\infty(0) + p_{11}\bar{q}_\infty(1)q_\infty(010) \end{aligned} \tag{46}$$

where for $a, b \in \{0, 1\}$

$$p_{ab} \triangleq \lim_{i\to\infty} P_{Z_{i+1}^{i+2}}(a, b). \tag{47}$$

Similar inequalities can be written for each pair $a, b \in \{0, 1\}$. Since $p_{00} + p_{01} + p_{10} + p_{11} = 1$, we will have 3 out of 4 independent inequalities. We can thus compute the lower bound $\underline{\boldsymbol{p}}$ by solving $(\boldsymbol{A} - \boldsymbol{I})\underline{\boldsymbol{p}} = \boldsymbol{b}$ where $\boldsymbol{A}$ is defined at the bottom of the page and $\underline{\boldsymbol{p}} = [\underline{p}_{00}, \underline{p}_{01}, \underline{p}_{10}, \underline{p}_{11}]^\mathsf{T}, \boldsymbol{b} = [0, 0, 0, 1]^\mathsf{T}$.

Similarly, the upper bound $\bar{\boldsymbol{p}}$ can be computed by substituting the upper bounds in (42) and (43) into (44). In general, the upper and lower bounds of the limiting joint distribution for the memory-m case can be computed by solving systems of linear equations.

By keeping track of the past m noise symbols, we generalize our previous memory-1 entropy rate bounds from (32) to

$$\begin{aligned} & H_\infty(Z) \\ & = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \sum_{z^{i-m-1}} \sum_{z_{i-m}^{i-1}} P_{Z^{i-1}}(z^{i-1}) h\big(q_i(z^{i-m-1}z_{i-m}^{i-1})\big) && (49) \\ & \ge \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \sum_{z_{i-m}^{i-1}} P_{Z_{i-m}^{i-1}}(z_{i-m}^{i-1}) h\big(q_i(0^{i-m-1}z_{i-m}^{i-1})\big) && (50) \\ & = \underline{\boldsymbol{p}}^\mathsf{T} \begin{bmatrix} h\big(q_\infty(00^m)\big) \\ \vdots \\ h\big(q_\infty(01^m)\big) \end{bmatrix} && (51) \end{aligned}$$

where the inequality follows from (43) and the monotonicity of the $h(p)$ for $p \in [0, 0.5]$. We have defined $q_\infty(0z^m) \triangleq \lim_{i\to\infty} q_i(0^{i-m-1}z^m)$ and we denote vector $\boldsymbol{p} \triangleq [p_{0^m}, \cdots, p_{1^m}]^\mathsf{T}$ with $p_{z^m} \triangleq P_{Z^m}(z^m)$.

Similarly, we obtain the upper bound by replacing $q_\infty(0z^m)$ with $q_\infty(1z^m) \triangleq \lim_{i\to\infty} q_i(1^{i-m-1}z^m)$, which leads to the following theorem.

Theorem 1. *For any fixed m, the channel capacity of a persistent Gilbert-Elliott channel $C_{\mathrm{GE}} = \log 2 - H_\infty(Z)$, can be bounded using the following bounds on the entropy rate of the induced HMM $H_\infty(Z)$*

$$\underline{\boldsymbol{p}}^\mathsf{T} \begin{bmatrix} h\big(q_\infty(00^m)\big) \\ \vdots \\ h\big(q_\infty(0z^m)\big) \\ \vdots \\ h\big(q_\infty(01^m)\big) \end{bmatrix} \le H_\infty(Z) \le \bar{\boldsymbol{p}}^\mathsf{T} \begin{bmatrix} h\big(q_\infty(10^m)\big) \\ \vdots \\ h\big(q_\infty(1z^m)\big) \\ \vdots \\ h\big(q_\infty(11^m)\big) \end{bmatrix}. \tag{52}$$

For any odd m, the entropy rate $H_\infty(Z)$ of an oscillatory Gilbert-Elliott channel can be bounded as

$$\underline{\boldsymbol{p}}^\mathsf{T} \begin{bmatrix} h\big(\tilde{q}_\infty(010^m)\big) \\ \vdots \\ h\big(\tilde{q}_\infty(01z^m)\big) \\ \vdots \\ h\big(\tilde{q}_\infty(011^m)\big) \end{bmatrix} \le H_\infty(Z) \le \bar{\boldsymbol{p}}^\mathsf{T} \begin{bmatrix} h\big(\tilde{q}_\infty(100^m)\big) \\ \vdots \\ h\big(\tilde{q}_\infty(10z^m)\big) \\ \vdots \\ h\big(\tilde{q}_\infty(101^m)\big) \end{bmatrix} \tag{53}$$

where we define $\tilde{q}_\infty(01z^m) \triangleq \lim_{i\to\infty} q_i(01\cdots01z^m)$ and $\tilde{q}_\infty(10z^m) \triangleq \lim_{i\to\infty} q_i(10\cdots10z^m)$. For even m, the upper and lower bounds are exchanged.

As we have seen previously, one can recover the channel capacity by setting the memory $m \to \infty$. Figure 4 plots

$$\boldsymbol{A} = \begin{bmatrix} \bar{q}_\infty(100)^2 & \bar{q}_\infty(101)\bar{q}_\infty(110) & \bar{q}_\infty(110)\bar{q}_\infty(100) & \bar{q}_\infty(1)\bar{q}_\infty(110) \\ \bar{q}_\infty(100)q_\infty(0) & \bar{q}_\infty(101)q_\infty(010) & \bar{q}_\infty(110)q_\infty(0) & \bar{q}_\infty(1)q_\infty(010) \\ q_\infty(0)\bar{q}_\infty(101) & q_\infty(001)\bar{q}_\infty(1) & q_\infty(010)\bar{q}_\infty(101) & q_\infty(011)\bar{q}_\infty(1) \\ 1 & 1 & 1 & 1 \end{bmatrix} \tag{48}$$

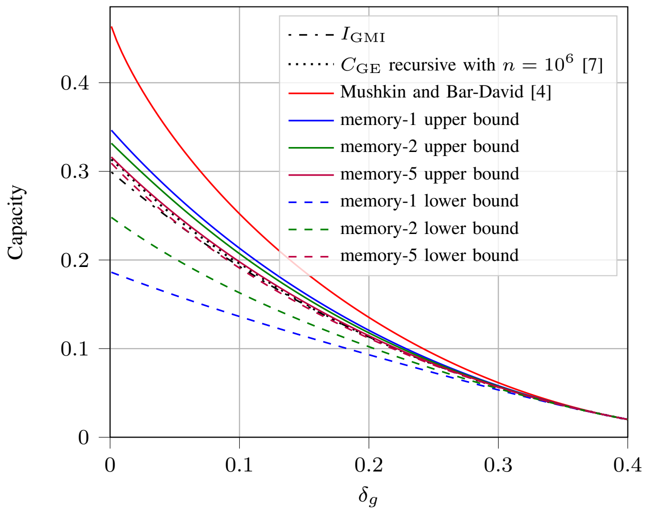

Fig. 4. Bounds for a persistent Gilbert-Elliott channel with parameters $b = 0.15,\ g = 0.3,\ \delta_{\mathrm{b}} = 0.4$.

the series of bounds for a persistent channel with the same parameters as in Fig. 2. The bounds get tighter as the memory increases. Our numerical simulations suggest that the fixed-memory upper bounds manifestly outperform the one derived by Mushkin and Bar-David, and the memory-5 lower bound can do better than the GMI for small δ_{g}. As is apparent from the figure, the memory-5 bounds already almost coincides with the simulated capacity curve.

Similarly, Figure 5 shows the bounds for an oscillatory Gilbert-Elliott channel with the parameters $b = 0.7,\ g = 0.85,\ \delta_{\mathrm{b}} = 0.4$. It is again clear from the plot that the fixed-memory bounds approach the capacity as the memory increases.

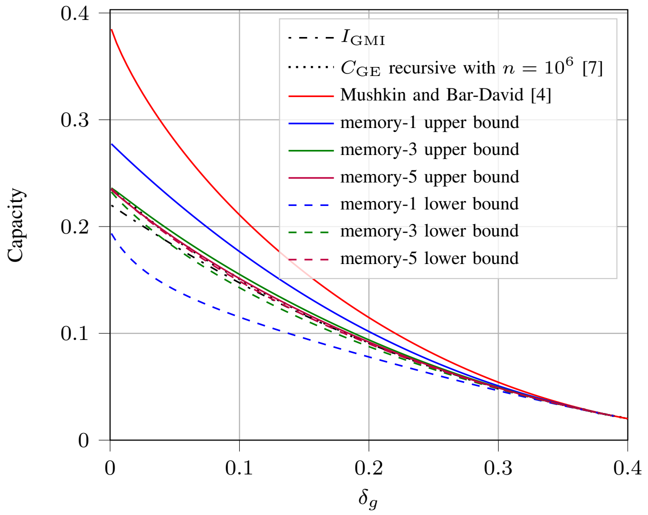

Fig. 5. Bounds for an oscillatory Gilbert-Elliott channel with parameters $b = 0.7,\ g = 0.85,\ \delta_{\mathrm{b}} = 0.4$.

IV. Conclusion

We have introduced a family of upper and lower bounds to the entropy rate of binary two-state HMMs. The proposed bounds only account for a finite memory of the process, and yield powerful bounds to the capacity of the Gilbert-Elliott channel.

References

[1] O. Zuk, E. Domany, I. Kanter, and M. Aizenman, "From finite-system entropy to entropy rate for a hidden Markov process," *IEEE Signal Processing Letters*, vol. 13, pp. 517 – 520, 10 2006.

[2] E. N. Gilbert, "Capacity of a burst-noise channel," *The Bell System Technical Journal*, vol. 39, no. 5, pp. 1253–1265, 1960.

[3] E. O. Elliott, "Estimates of error rates for codes on burst-noise channels," *The Bell System Technical Journal*, vol. 42, no. 5, pp. 1977–1997, 1963.

[4] 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, 1989.

[5] R. J. MacKay, "Estimating the order of a hidden Markov model," *The Canadian Journal of Statistics / La Revue Canadienne de Statistique*, vol. 30, no. 4, pp. 573–589, 2002. [Online]. Available: http://www.jstor.org/stable/3316097

[6] L. E. Baum and T. Petrie, "Statistical Inference for Probabilistic Functions of Finite State Markov Chains," *The Annals of Mathematical Statistics*, vol. 37, no. 6, pp. 1554 – 1563, 1966. [Online]. Available: https://doi.org/10.1214/aoms/1177699147

[7] M. Rezaeian, "Computation of capacity for Gilbert-Elliott channels, using a statistical method," in *2005 Australian Commun. Theory Workshop*, 2005, pp. 56–61.

[8] G. Kaplan and S. Shamai, "Information rates and error exponents of compound channels with application to antipodal signaling in a fading environment," *AEU. Archiv für Elektronik und Übertragungstechnik*, vol. 47, no. 4, pp. 228–239, 1993.

[9] Y. Han and A. Guillén i Fàbregas, "Coding for the Gilbert-Elliott channel: A mismatched decoding perspective," in *Proc. 2024 International Zürich Seminar, Zürich, Switzerland, Mar. 2024*, 2024.