Real-time Binary Posterior Matching

Anusha Lalitha, Anatoly Khina, Tara Javidi, and Victoria Kostina

***Abstract*—We consider the problem of communications over the binary symmetric channel with feedback, where the information sequence is made available in a causal, possibly random, fashion. We develop a real-time variant of the renowned Horstein scheme and provide analytical guarantees for its error-probability exponential decay rate. We further use the scheme to stabilize an unstable control plant over a binary symmetric channel and compare the analytical guarantees with its empirical performance as well as with those of anytime-reliable codes.**

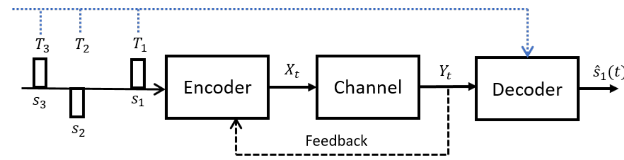

Fig. 1. Figure shows the transmission of a stream of bits which arrive at the encoder at random times T_i. The arrival times are available at both at the encoder and decoder.

I. Introduction

While feedback cannot increase the capacity of memoryless channels [1, Ch. 7.12], it can dramatically reduce the probability of error and the complexity of the communication schemes that achieve them. For the binary symmetric channel (BSC), a *horizon-free* sequential scheme was proposed by Horstein [2]; it was rigorously proved to attain capacity by Shayevitz and Feder [3] for this and other channels, via its generalization—the *posterior matching* (PM) scheme. Exponential error-probability guarantees, for the *finite-horizon* setting, were constructed in [4]–[7]. An exponential bound on the error probability in the horizon-free case has been devised by Waeber et al. [8], although this bound becomes trivial for rates much below the capacity.

The availability of instantaneous noiseless feedback obviates the need of transmitting long error-correcting codes across long epochs, and enables instead the use of sequential communication schemes, by providing full knowledge of the receiver's state to the transmitter. A class of problems where this may have powerful implications is that of stabilizing an unstable control plant over a noisy channel. In particular, in the presence of feedback, the structure of the horizon-free PM decoder seems to match the structure of anytime reliable decoders (proposed for stabilizing unstable linear plants over noisy channel [9]–[11]).

However, the classical PM schemes assume that the entire information (possibly infinite bit) sequence is available essentially non-causally to the transmitter, prior to the beginning of transmission. That is, they are sequential with respect to the transmitted sequence (codeword) but not with respect to the information sequence. Consequently, the non-causal knowledge assumption precludes the use of the classical PM scheme for real-time and control scenarios, in which the data to be transmitted is determined in a causal fashion.

In the current work, we consider a real-time setting, described in detail in Sec. II, in which the bits arrive to the transmitter one-by-one at random times, under the assumption that the inter-arrival times (time-arrival differences) have a known finite support. We construct, in Sec. III, a causal (horizon-free) PM scheme for this setting, i.e., a scheme that is sequential with respect to both the information and the transmitted sequences. We provide exponential guarantees for the error probability akin to those of [8], in Sec. IV.

We apply the proposed scheme, in Sec. V, for control over a BSC with feedback and compare its analytic and empirical stabilization performance with those of the anytime-reliable codes of Sahai and Mitter [9] that use no feedback but are computationally demanding, as well as with those of Simsek et al. [12],[1] in Sec. V-A. We conclude the paper with a discussion, in Sec. VI.

<u>*Notation:*</u> $\mathbb{N}$ denotes the set of natural numbers. For $k, t \in \mathbb{N}$, $k < t$, the sequence $\{s_k, s_{k+1}, \ldots, s_t\}$ is denoted as s_k^t. For $M \in \mathbb{N}$, the sequence of integers $\{1, 2, \ldots, M\}$ is denoted $[M]$. The binary entropy of probability p is denoted by $h(p) = -p \log p - \bar{p} \log \bar{p}$ with $\bar{p} := 1-p$; all logarithms in this work are to the base 2. For any probability mass function (pmf) p, let $p^{\bigotimes i}$ denote p convolved with itself i times.

II. Problem Formulation

The transmitter wishes to transmit an infinite stream of bits over a BSC with cross over probability $p \in (0, 1/2)$. We assume the bits are revealed to the transmitter causally at arbitrary (possibly random) times, as follows. Let $\{N_i\}_{i \geq 1}$

A. Lalitha and T. Javidi are with the Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92093, USA (e-mails: {alalitha, tjavidi}@eng.ucsd.edu).

A. Khina is with the School of Electrical Engineering, Tel Aviv University, Tel Aviv, Israel 6997801 (e-mail: anatolyk@eng.tau.ac.il).

V. Kostina is with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125, USA (e-mail: vkostina@caltech.edu).

The work of A. Lalitha and T. Javidi was supported in part by the National Science Foundation (NSF) under Grant CIF-1719133, CIF-1513883 and CIF 1448237. This work has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 708932. The work of V. Kostina was supported in part by the NSF under Grant CCF-1566567 and Grant CCF-1751356. This work was done in part while A. Khina and V. Kostina were visiting the Simons Institute for the Theory of Computing.

[1] Analytic guarantees for the scheme of [12] exist only for the case in which the entire information sequence is known in advance, which corresponds, to the case of stabilizing an unstable linear system with possibly unknown initial conditions but with no system disturbance.

be an i.i.d. random process where each $N_i \in [n_{\min}, n_{\max}]$, and has a pmf p_N. Then the i-th bit arrives at time $T_i := \sum_{j=1}^{i-1} N_j + 1$ for all $i \geq 2$ with $T_1 = 1$. Notationally, we consider the infinite bit sequence as the binary expansion of a single message point Θ uniformly distributed over the unit interval i.e., $\Theta \sim \mathrm{Unif}[0, 1)$.

Remark 1 (Periodic arrival times)**.** An important special instance of this framework is the case of deterministic and periodic arrival times with, in which a new information bit is revealed every (fixed) $n \in \mathbb{N}$ time steps.

We now define the feedback communication scheme of an information bit sequence that is made available causally to the encoder with random inter bit-arrival times, depicted in Fig. 1. We assume the times at which the bits are revealed to the transmitter are known at the receiver. The encoder $\mathcal{E}$ is described by a sequence of (causal) functions $\{\mathcal{E}_t\}_{t\geq 1}$. For any $i \geq 1$, let s_1^i denote the first i bits of binary expansion of message point $\Theta \in [0, 1)$. Assuming the first bit is available at the beginning, the encoder has access to the first i bits at time t[2], for $t \geq n_{\min}(i-1)+1$ with a non-zero probability and for $t \geq n_{\max}(i-1)+1$ with probability 1. Furthermore, after $t-1$ channel uses, the encoder has access to past channel outputs y_1^{t-1} due to the availability of feedback. Provided that the first i bits are available, at time t, a causal encoder with feedback emits a channel input symbol $x_t \in \{0, 1\}$:

$$x_t = \mathcal{E}_t\left(s_1^i, y_1^{t-1}\right).$$

The decoder $\mathcal{D}$ is described by the sequence of functions $\{\mathcal{D}_t\}_{t\geq 1}$. After observing t channel outputs, if the first i bits are available, the decoder outputs a vector of estimates of all the bits available at the encoder thus far, $\hat{s}_1^i(t) = [\hat{s}_1(t), \hat{s}_2(t), \ldots, \hat{s}_i(t)] \in \{0, 1\}^i$:

$$\hat{s}_1^i(t) = \mathcal{D}_t\left(y_1^t\right).$$

For any $i \in \mathbb{N}$ at any time instant $t \geq n_{\min}(i-1)+1$, we aim to analyze the probability of error in decoding the first j bits $\mathsf{P}\left(\hat{s}_1^j(t) \neq s_1^j\right)$ for $1 \leq j \leq i$. Since the bits that arrive early get encoded for longer it is natural to expect that the probability of error in decoding the older bits is smaller than that in decoding the newer bits.

III. Causal Posterior Matching Strategy

In this section, we propose a causal PM based encoding and decoding strategy to transmit a causally available message where the inter bit-arrival times are random.

First, we provide an overview of the strategy. At time t, suppose the first i bits are available to the encoder. Consider a unit interval $[0, 1]$ and divide it into bins of equal length 2^{-i}. The message point Θ is located on the unit interval, whose the first i bits s_1^i provide the index of the bin containing Θ, where an index belongs to $\{0, 1, \ldots, 2^i - 1\}$, containing Θ. The encoder and decoder maintain a posterior probability of the message point Θ belonging to each bin after observing the past channel outputs. For the next N_i channel uses, we use causal posterior matching (described in detail in Sections III-B and III-C below) to encode the first i bits and perform a Bayesian update to the posterior probability of Θ given the received channel outputs. After these N_i channel uses, a new bit arrives. Then, we divide each bin from the previous 2^i bins into 2 bins, resulting in 2^{i+1} bins in total. Furthermore, we equally divide the posterior probability to accommodate the new bit. Now, the first $i+1$ bits provide the index of the bin containing Θ on a grid with $2^{(i+1)}$ bins. This process of dividing the existing bins and the posterior probability to accommodate a new bit continues in a horizon-free manner. At any time $t \geq 1$, the bits which provide the index of the bin that contains the median of the posterior distribution are declared as estimates of the bits available at the encoder.

A. Preliminaries

Let $\mathrm{BSC}(p)$ denote a BSC with cross-over probability $p \in (0, 1/2)$ with input $X \in \{0, 1\}$, output $Y \in \{0, 1\}$:

$$\mathsf{P}(Y = y | X = x) = \begin{cases} p & \text{if } y \neq x, \\ \bar{p} & \text{if } y = x. \end{cases}$$

Let $C(p) := 1 - h(p)$ denote the capacity of $\mathrm{BSC}(p)$.

Suppose after t channel uses, the encoder has access to i bits. The decoder maintains a posterior distribution of Θ after observing t channel outputs y_1^t, i.e., $\mathsf{P}_{\Theta|Y_1^t}\left(\Theta \in [(k-1)2^{-i}, k2^{-i}) \middle| y_1^t\right)$ for all $k \in [2^i]$. Let denote by $F_{\Theta|Y_1^t}$ the corresponding posterior cumulative distribution function (CDF) of posterior probability distribution. Due to the presence of feedback, the posterior distribution maintained by the decoder is available to the encoder as well. We refer to the point $F_{\Theta|Y_1^t}^{-1}(1/2|y_1^t)$ as the median of the posterior probability distribution at time t.

The following definitions will be useful, as we shall see, in describing the causal PM strategy. For every $n \in \mathbb{N}$, let $\beta(n)$ denote the solution of the following equation

$$\beta = \psi^*(\beta) - \frac{1}{n}, \tag{1}$$

where $\psi(\lambda) := -\log\left\{(2p)^\lambda + (2\bar{p})^\lambda\right\} + 1$, and

$$\psi^*(\beta) := \sup_{\lambda > 0}\left(\psi(\lambda) - \lambda\beta\right) \tag{2}$$

denotes its Legendre–Fenchel transform. Further denote by $\lambda^*(n) \in [0, 1]$ the λ that achieves the supremum in (2) when $\psi^*(\beta)$ satisfies (1). We are now ready to describe the causal PM strategy in detail.

B. Encoder

Fix a parameter $\lambda \in \{\lambda^*(n_{\min}), \lambda^*(n_{\max})\}$. Suppose only the first i bits of the message Θ are available, the encoding is performed as a function of the first i bits. Let $k_i^{(t)} \in [2^i]$ denote the index of the bin containing the median $F_{\Theta|Y_1^t}^{-1}(1/2)$ in the grid with resolution 2^{-i}, i.e,

$$\left(k_i^{(t)} - 1\right)2^{-i} < F_{\Theta|Y^{t-1}}^{-1}\left(1/2\middle|y_1^t\right) \leq k_i^{(t)} 2^{-i}.$$

[2]This time t depends on the index of the last bit revealed to the transmitter.

Let $d_1^{(t)}$ and $d_2^{(t)}$ denote the interval lengths to the left and to the right of the median in bin $k_i^{(t)}$, respectively.

$$d_1^{(t)} := F^{-1}_{\Theta|Y^{t-1}}(1/2) - \left(k_i^{(t)} - 1\right)2^{-i},$$
$$d_2^{(t)} := k_i^{(t)}2^{-i} - F^{-1}_{\Theta|Y^{t-1}}(1/2).$$

Note that $d_1^{(t)}, d_2^{(t)} \leq 1/2$. Define further, for any $\lambda \in [0,1]$,

$$\pi_1^{(t)}(\lambda) := \frac{h(\lambda, d_2^{(t)})}{h(\lambda, d_1^{(t)}) + h(\lambda, d_2^{(t)})}, \quad \pi_2^{(t)}(\lambda) := 1 - \pi_1^{(t)}(\lambda),$$

where $h(\lambda, d) := (1-2(\bar{p}-p)d)^{-\lambda} - (1+2(\bar{p}-p)d)^{-\lambda}$. Conditioned on the past observations y_1^{t-1} and the first i bits that are available, with probability $\pi_1^{(t)}(\lambda)$ (resp. $\pi_2^{(t)}(\lambda)$), the encoding is

$$X_t = \begin{cases} 0 & \text{if } 0.s_1^i \leq \left(k_i^{(t)}-1\right)2^{-i} \text{ (resp. } k_i^{(t)}2^{-i}\text{)}, \\ 1 & \text{if } 0.s_1^i > \left(k_i^{(t)}-1\right)2^{-i} \text{ (resp. } k_i^{(t)}2^{-i}\text{)}. \end{cases}$$

Whenever a new bit arrives, the encoder divides each of the previous bins into two equal-length bins with equal posterior probabilities. That is, for all $i \geq 1$, after $t = \sum_{j=1}^{i} N_j + 1$ channel uses, when the $(i+1)$-th bit arrives, the encoder sets

$$\begin{aligned} &\mathsf{P}_{\Theta|Y_1^t}\left(\Theta \in \left[(2k-1)2^{-i-1}, (2k)2^{-i-1}\right) \middle| y_1^t\right) \\ &\quad = \mathsf{P}_{\Theta|Y_1^t}\left(\Theta \in \left[(2k-2)2^{-i-1}, (2k-1)2^{-i-1}\right) \middle| y_1^t\right) \\ &\quad = \frac{1}{2}\mathsf{P}_{\Theta|Y_1^t}\left(\Theta \in \left[(k-1)2^{-i}, k2^{-i}\right) \middle| y_1^t\right), \qquad \forall k \in [2^i]. \end{aligned}$$

C. Decoder

Upon receiving the channel output at time instant t, the decoder performs a Bayesian update to the posterior of Θ. For $k < k_i^{(t)}$, since $F_{\Theta|Y_1^{t-1}}\left(k2^{-i}|y_1^{t-1}\right) \leq 1/2$, the Bayesian update is given as follows. With probability $\pi_1^{(t)}(\lambda)$:

$$\frac{F_{\Theta|Y_1^t}\left(k2^{-i}|y_1^t\right)}{F_{\Theta|Y_1^{t-1}}\left(k2^{-i}|y_1^{t-1}\right)} = \begin{cases} \frac{p}{\frac{1}{2}+(\bar{p}-p)d_1^{(t)}} & \text{if } Y_t = 1, \\ \frac{q}{\frac{1}{2}-(\bar{p}-p)d_1^{(t)}} & \text{if } Y_t = 0, \end{cases}$$

and with probability $\pi_2^{(t)}(\lambda)$:

$$\frac{F_{\Theta|Y_1^t}\left(k2^{-i}|y_1^t\right)}{F_{\Theta|Y_1^{t-1}}\left(k2^{-i}|y_1^{t-1}\right)} = \begin{cases} \frac{p}{\frac{1}{2}-(\bar{p}-p)d_2^{(t)}} & \text{if } Y_t = 1, \\ \frac{q}{\frac{1}{2}+(\bar{p}-p)d_2^{(t)}} & \text{if } Y_t = 0. \end{cases}$$

For $k \geq k_i^{(t)}$, since $1 - F_{\Theta|Y^{t-1}}(k2^{-i}) < 1/2$, the Bayesian update rule can be specified similarly.

At any time instant t, the decoder generates an estimate $\widehat{\Theta}_t = F^{-1}_{\Theta|Y^t}(1/2|y_1^t)$ of Θ. The estimates $\hat{s}_1^i(t)$ of the first i bits are the bits associated with the bin containing the median. Furthermore, when a new bit arrives, similarly to the encoder, the decoder divides each bin into two equal-length bins with equal posterior probabilities.

Remark 2. In the special case where the median coincides with the (right) end point of the bin, $k_i^{(t)}2^{-i}$, the encoder transmits 1 if s_1^i bits are to the right of the median and—0 otherwise. Furthermore, the decoder's update reduces to the update of non-causal PM, where each $F_{\Theta|Y^{t-1}}(k2^{-i})$, $k \leq k_i^{(t)} - 1$ and similarly $1 - F_{\Theta|Y^{t-1}}(k2^{-i})$, $k \geq k_i^{(t)}$, expands by $2\bar{p}$ or shrinks by $2p$.

Remark 3. For any $i \geq 1$, for all $t \geq T_i$, the encoder has access to the first i bits and hence the number of bins is at least 2^i. In other words, from time T_i to t, the causal PM strategy operates on a grid whose resolution is finer than 2^{-i} and hence updates the bin end points of $F_{\Theta|Y^t}$ according to the grid of resolution 2^{-i}. This implies that, for all $t \geq T_i$, we always encode the first i bits along with the newly available bits. Furthermore, although we assume bits arrive one at a time, the strategy and our analysis can be extended to the case where $k \geq 1$ bits arrive at a time.

IV. Main Results

In this section, we provide our main result on the error exponent attained by the causal PM strategy.

Theorem 1. *Consider the causal PM strategy with parameter λ over a $BSC(p)$. The i-th bit arrives at the encoder at a random time T_i, whose pmf is $p_N^{\bigotimes i}$, where the inter-arrival times satisfy $N \in [n_{\min}, n_{\max}]$. Then,*

1) *For $\lambda = \lambda^*(n_{\min})$, the probability of error in decoding the first $j \in [\lfloor t/n_{\min} \rfloor]$ message bits after t channel uses is bounded by*

$$\begin{aligned} &\mathsf{P}\left(\hat{s}_1^j(t) \neq s_1^j \middle| T_{j-1} < t\right) \\ &\leq \kappa \mathbb{E}\left[2^{-\beta(n_{\min})(t-T_j)} \middle| T_{j-1} < t\right], \end{aligned} \tag{3}$$

where $\beta(n_{\min})$ is the solution of (1) *for $n = n_{\min}$ and where κ is a finite positive constant.*

2) *For $\lambda = \lambda^*(n_{\max})$, the probability of error in decoding the first $j \in [\lfloor t/n_{\max} \rfloor]$ message bits after t channel uses is bounded by*

$$\mathsf{P}\left(\hat{s}_1^j(t) \neq s_1^j\right) \leq \kappa\left(2^{-\beta(n_{\max})(t-n_{\max}(j-1))}\right), \tag{4}$$

where $\beta(n_{\max})$ is the solution of (1) *for $n = n_{\max}$ and where κ is a finite positive constant.*

Theorem 1 shows that the causal PM strategy can operate in two regimes based on how the randomization $\pi_1^{(t)}$ and $\pi_2^{(t)}$ are shown given the past observations and the number of bits available at the encoder is chosen, i.e., based on the parameter λ. In regime 1, causal PM can be thought of as operating in a "*high-rate regime*", since it decodes all the arrived information bits, but with a lower error exponent of (3), corresponding to $\beta(n_{\min})$. In contrast, in regime 2, causal PM can be thought of as operating in a "*low-rate regime*", as it decodes only the first $\lfloor t/n_{\max} \rfloor$ bits, but with a higher error exponent of (4), corresponding to $\beta(n_{\min})$. A similar analysis can be done for $\lambda \in (\lambda^*(n_{\min}), \lambda^*(n_{\max}))$, however we leave that for future work.

The proof above theorem is available in the supplementary material [13]. The proof relies on the analyzing the tails of the posterior probabilities $\min\{F_{\Theta|Y_1^t}(\theta|y_1^t), 1 - F_{\Theta|Y_1^t}(\theta|y_1^t)\}$

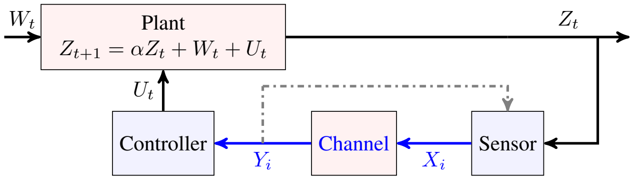

Fig. 2. A scalar linear plant that is controlled over a noisy channel. The Sensor transmits to the controller over a noisy channel with feedback; n channel uses $\{X_i\}$ per control sample Z_t are assumed.

for $\theta \in (0,1)$ inspired by the analysis in [8]. However, the analysis of the expected value of decay of the tails is based on the analysis of Burnashev and Zigangirov in [4].

Corollary 1 (Periodic arrival times)**.** *Consider the causal PM strategy with parameter λ over a $BSC(p)$. The i-th bit arrives at the encoder at time $T_i = n(i-1)+1$, i.e., the inter-arrival time is constant $n \geq 1$. Then, for $\lambda = \lambda^*(n)$, the probability of error in decoding the first $j \in [\lfloor t/n \rfloor]$ bits of a message after t channel uses is bounded by*

$$\mathsf{P}\left(\hat{s}_1^j\,(t) \neq s_1^j\right) \leq \kappa \left(2^{-\beta(n)(t-n(j-1))}\right),$$

where $\beta(n)$ is the solution of (1) *for n, and where $0 \leq \kappa < \infty$.*

V. Application to Control over Noisy Channels

Consider the problem of stabilizing an unstable scalar plant,

$$Z_{t+1} = \alpha Z_t + W_t + U_t, \tag{5}$$

where $\alpha > 1$, the initial state is a random variable $Z_0 \in [-\Delta, \Delta]$, the disturbances $\{W_t\}_{t \geq 0}$ are i.i.d. with a bounded support $W_t \in [-W, W]$ and U_t is a control signal applied by the controller at time t. The controller, that generates U_t, is separated from the sensor that measures Z_t by a BSC(p) with feedback, i.e., n channel uses per each control sample Z_t are available. For $\eta \geq 1$, we want to stabilize the η-th moment, i.e., $\sup_t \mathbb{E}[|Z_t|^\eta] < \infty$. To that end, suppose the observer quantizes the plant measurements into 1 bit, which implies a new bit arrives after every n channel uses. This is a special case of our strategy where the inter-arrival time of the bits is fixed (recall Rem. 1). This model is depicted in Fig. 2.

Remark 4. For the ease of exposition, we consider a 1 bit quantizer but the strategy can be extended to a k-bit quantizer for any $1 \leq k \leq n$.

To stabilize the plant it suffices to apply a control signal $U_t = -\alpha \hat{Z}_t$, where $\{\hat{Z}_t\}_{t \geq 1}$ satisfies $\sup_t \mathbb{E}[|Z_t - \hat{Z}_t|^\eta] < \infty$. The following corollary provides the values α for which the plant can be stabilized.

Corollary 2. *Consider the plant of* (5) *for $\alpha > 1$ observed through a BSC(p) with feedback with a budget of n channel uses. Then, for all $\eta \geq 1$, the plant is η-stabilizable, i.e., $\sup_t \mathbb{E}\left[\left|Z_t - \hat{Z}_t\right|^\eta\right] < \infty$, for*

$$\log \alpha \leq \min\left\{1/n, \beta(n)/\eta\right\}, \tag{6}$$

where $\beta(n)$ is the solution of (1).

Proof: We use the causal PM strategy to transmit the quantized plant measurements over a BSC(p) with feedback. This is a special case of our causal PM strategy where the inter-arrival time of the bits is a constant n, hence we set $\lambda = \lambda^*(n)$. For each step of the plant evolution we convey one bit over n channel uses. Corollary 1 provides the following guarantees on the estimates generated by the causal PM strategy

$$\mathsf{P}\left(\hat{s}_1^j\,(nt) \neq s_1^j\right) \leq \kappa \left(2^{-\beta(n)n(t-j)}\right),$$

for all $j \in [t]$, where $\beta(n)$ is the solution of (1). Hence, using [9, Theorem 4.1] we have that the plant is η-stabilizable if (6) holds. ∎

Remark 5. The constraint $\log \alpha < 1/n \leq 1$ is due to a 1-bit quantization requirement that we implicitly impose by assuming that a single bit arrives at a time. This requirement can be lifted by allowing higher quantization rates, along with the appropriate adaptation of the proposed scheme, at the price of reducing the error exponent β. In other words, two conflicting effects can be seen in the problem of stabilizing an unstable plant over a noisy channel: (i) Source quantization: we wish to maximize the quantization resolution to allow for finer source approximation, however this results in higher channel-coding rate since more bits have to be sent over a given channel budget n (ii) Channel coding: we wish to minimize the channel-coding rate to minimize the error due to decoding, i.e., to maximize the error exponent. These two effects are manifested by the two minimands in (6).

Remark 6. As a consequence of Corol. 2 , for a given $\eta \geq 1$ and $p \in (0, 1/2)$, we obtain a lower bound $R(p)$ on the maximum rate (i.e., minimum channel budget $\lceil 1/R(p) \rceil$) at which the communication channel BSC(p) can be operated such that the plant (5) is η-stabilizable for some $\alpha > 1$. Using (1), note that we have

$$\min\left\{\frac{1}{n}, \frac{\beta(n)}{\eta}\right\} \geq \max_{\beta > 0} \min\left\{\frac{\beta}{\eta}, \frac{1}{\eta}\left(\psi^*(\beta) - \frac{1}{n}\right), \frac{1}{n}\right\}$$

Hence, using Corol. 2, this implies that $R(p)$ is the largest $R > 0$ that satisfies the following equation:

$$\psi^*(\eta R) = (\eta + 1)R.$$

In other words, we obtain that $2^{R(p)}$ is a lower bound on the largest α for which the plant (5) can be η-stabilized over a BSC(p) for any channel budget $n > 1$.

A. Simulations for Control over Noisy Channels

For various values of channel budget $n \in \mathbb{N}$ we numerically compute the bound provided by Corollary 2 on largest eigenvalue α for which a plant is stabilizable. Fig. 3 shows the largest eigenvalue α as a function of inverse of the channel budget, i.e. rate, for different values of crossover probability of a BSC. This illustrates that the causal PM-based scheme can stabilize the plant for α values that are strictly greater than one for the considered crossover probabilities by choosing channel budget appropriately.

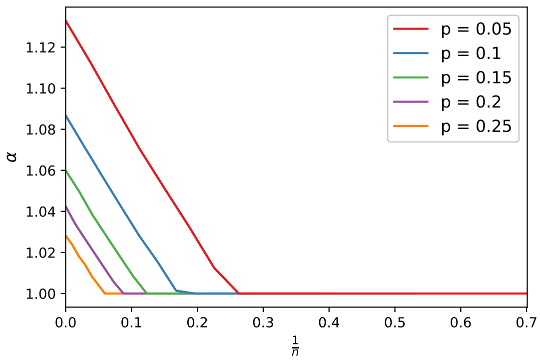

Fig. 3. The maximum eigenvalue α of a plant that is stabilizable using the causal PM strategy over a BSC(p) as a function of inverse of channel budget $1/n$ for various values of p.

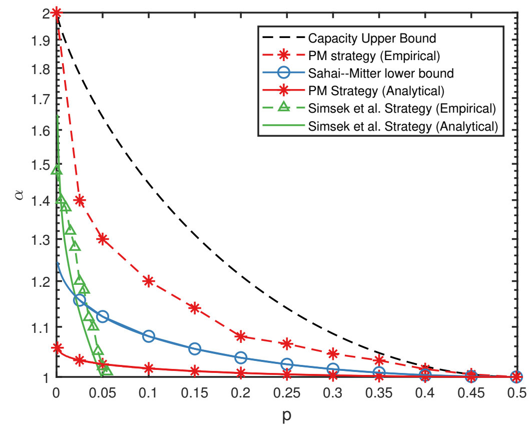

Fig. 4. The maximum eigenvalue α of a plant that is stabilizable over a BSC(p) as a function of p using: the causal PM strategy (analytically and empirically), the Simsek et al.strategy [12] (analytically and empirically), the Sahai–Mitter tree-code lower bound [9], and the capacity upper bound.

We compare the performance of the proposed causal PM strategy with previously proposed upper and lower bounds for the maximal value of α for which the plant (5) can be stabilized. Fig. 4 compares the stabilizability of a system as a function of the crossover probability of a BSC. The empirical as well as the theoretical performance of both the causal PM-based strategy and a strategy proposed by Simsek et al. [12] (albeit for the interference-free case: $W_t \equiv 0$), as well as the Sahai–Mitter lower bound without feedback (anytime-reliable tree codes) of [9] and the capacity upper bound are illustrated. From Fig. 4 we see that the bound $2^{R(p)}$ on α, provided by our analysis of the causal PM-based scheme, is rather conservative in comparison to its empirical performance. The latter clearly outperforms the Simsek et al. strategy [12] (for which the analysis is rather tight) and exceeds the Sahai–Mitter lower bound. This demonstrates that the causal PM-based strategy provides better performance both in terms of stability and complexity. We further note that the causal PM-based scheme can stabilize the plant for α values that are strictly greater than one for all crossover probabilities $p \in [0, 1/2)$, even under the provided conservative analysis. This is in stark contrast to the strategy of Simsek et al., which can stabilize unstable plants only below a certain threshold crossover probability.

VI. Conclusions and Future Work

We considered the problem of transmitting an infinite stream of bits over a BSC where the bits are revealed to the transmitter causally and the inter bit-arrival time may be random. We proposed a causal PM strategy and provided guarantees for the error exponent of the decoded bits using this strategy. The causal PM is parameterized by $\lambda(n)$ which decides the randomization of the encoding functions. Hence, it implicitly decides the number of bits decoded and their error exponent. We derived explicit results for two extremes of $\lambda(n)$. An interesting area of future work would be to extend our analysis to any λ between these two extremes. Another important future direction is to extend our analysis to the case where the bit arrival times are unknown at the receiver.

Furthermore, we applied our strategy to the problem of stabilizing a control plant over a BSC. We provided analytical guarantees on the maximal plant eigenvalue for which the plant can be stabilized using causal posterior matching. Closing the gap between our analysis and the empirical performance is an important area of future.

References

[1] T. M. Cover and J. A. Thomas, *Elements of Information Theory, Second Edition*. New York: Wiley, 2006.
[2] M. Horstein, "Sequential transmission using noiseless feedback," *IEEE Trans. Inf. Theory*, vol. 9, no. 3, pp. 136–143, Jul. 1963.
[3] O. Shayevitz and M. Feder, "Optimal feedback communication via posterior matching," *IEEE Trans. Inf. Theory*, vol. 57, no. 3, pp. 1186–1222, Mar. 2011.
[4] M. V. Burnashev and K. Zigangirov, "An interval estimation problem for controlled observations," (in Russian) *Problemy Pered. Info. (Problems of Info. Trans.)*, vol. 10, no. 3, pp. 51–61, 1974.
[5] A. G. D'yachkov, "Upper bounds on the error probability for discrete memoryless channels with feedback," (in Russian) *Problemy Pered. Info. (Problems of Info. Trans.)*, vol. 11, no. 4, pp. 13–28, 1975.
[6] C. T. Li and A. El Gamal, "An efficient feedback coding scheme with low error probability for discrete memoryless channels," *IEEE Trans. Inf. Theory*, vol. 61, no. 6, pp. 2953–2963, Jun. 2015.
[7] M. Naghshvar, T. Javidi, and M. Wigger, "Extrinsic Jensen–Shannon divergence: Applications to variable-length coding," *IEEE Trans. Inf. Theory*, vol. 61, no. 4, pp. 2148–2164, Apr. 2015.
[8] R. Waeber, P. I. Frazier, and S. G. Henderson, "Bisection search with noisy responses," *SIAM Journal on Control and Optimization*, vol. 51, no. 3, pp. 2261–2279, May 2013.
[9] A. Sahai and S. K. Mitter, "The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link—part I: Scalar systems," *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3369–3395, Aug. 2006.
[10] R. T. Sukhavasi and B. Hassibi, "Linear time-invariant anytime codes for control over noisy channels," *IEEE Trans. Auto. Cont.*, vol. 61, no. 12, pp. 3826–3841, Dec. 2016.
[11] A. Khina, W. Halbawi, and B. Hassibi, "(Almost) practical tree codes," in *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, Barcelona, Spain, Jul. 2016, pp. 2404–2408.
[12] T. Simsek, R. Jain, and P. Varaiya, "Scalar estimation and control with noisy binary observations," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1598–1603, 2004.
[13] A. Lalitha, A. Khina, T. Javidi, and V. Kostina, "Supplementary material," 2019. [Online]. Available: https://sites.google.com/a/eng.ucsd.edu/alalitha/publications?authuser=0#h.p_KIAtCcS6_mzb