

# Bounds on the Information Rate for Sparse Channels with Long Memory and i.u.d. Inputs

Andreja Radosevic, *Student Member, IEEE*, Dario Fertonani, *Member, IEEE*, Tolga M. Duman, *Fellow, IEEE*, John G. Proakis, *Life Fellow, IEEE*, and Milica Stojanovic, *Fellow, IEEE*

**Abstract**— In this paper we propose new bounds on the achievable information rate for discrete-time Gaussian channels with intersymbol interference (ISI) and independent and uniformly distributed (i.u.d.) channel input symbols drawn from finite-order modulation alphabets. Specifically, we are interested in developing new bounds on the achievable rates for sparse channels with long memory. We obtain a lower bound which can be achieved by practical receivers, based on MMSE channel shortening and suboptimal symbol detection for a reduced-state channel. An upper bound is given in the form of a semi-analytical solution derived using basic information theoretic inequalities, by a grouping of the channel taps into several clusters resulting in a newly defined single-input multiple-output (SIMO) channel. We show that the so obtained time-dispersive SIMO channel can be represented by an equivalent single-input single-output (SISO) channel with a significantly shorter channel memory. The reduced computational complexity allows the use of the BCJR algorithm for the newly defined channel. The proposed bounds are illustrated through several sparse channel examples and i.u.d. input symbols, showing that the upper bound significantly outperforms existing bounds. Performance of our lower bound strongly depends on the channel structure, showing best results for minimum-phase and maximum-phase systems.

**Index Terms**— Bounds, channel capacity, information rates, intersymbol interference, sparse channel.

## I. INTRODUCTION

CHANNELS with sparse inter-symbol interference (ISI) are commonly found in a wide range of communication systems, such as underwater acoustic, aeronautical and satellite systems. For instance, underwater acoustic (UWA) communication systems are inherently wideband since the acoustic propagation is best supported at low frequencies (e.g. up to several tens of kHz for a range of several km). They also exhibit extensive multipath spreads, resulting in ISI spans of tens or even hundreds of symbols for single-carrier systems. However, very few propagation paths carry significant energy even though severe multipath propagation spans

Paper approved by G. M. Vitetta, the Editor for Equalization and Fading Channels of the IEEE Communications Society. Manuscript received July 19, 2010; revised January 26, 2011 and May 24, 2011.

A. Radosevic and J. G. Proakis are with the Department of Electrical and Computer Engineering, Univ. of California, San Diego, La Jolla, CA 92093 (e-mail: {aradosevic, jproakis}@ucsd.edu).

D. Fertonani and T. M. Duman are with the School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ 85287-5706 (e-mail: dario.fertonani@gmail.com, duman@asu.edu).

M. Stojanovic is with the Department of Electrical and Computer Engineering, Northeastern University, Boston, MA 02115 (e-mail: militsa@ece.neu.edu).

Digital Object Identifier 10.1109/TCOMM.2011.120611.100435

This work is supported by the Multidisciplinary University Research Initiative (MURI) ONR Grants N00014-07-1-0738 and N00014-07-1-0739.

a large number of symbol intervals, resulting in sparse ISI channels. For discrete-time ISI channel models with additive white Gaussian noise there is no single-letter expression for the achievable information rate that can be supported with a particular modulation alphabet [1]. Under an independent and uniformly distributed (i.u.d.) or Markov input constraints, the achievable information rate can be estimated as accurately as desired by means of a simulation-based algorithm described in [2][3]. This approach requires simulation of forward recursion of the full-complexity Bahl–Cocke–Jelinek–Raviv (BCJR) algorithm [2] with a trellis whose complexity increases exponentially with channel memory in order to estimate the joint probability of the output sequence. The so obtained probability is used to compute an estimate of the mutual information between the input and output sequences. Analytical simplifications are not available, and due to the long channel memory, this simulation-based analysis is basically infeasible for practical UWA channels.

The only tools for the characterization of the information rate of channels with long memory are upper and lower bounds on the achievable information rates. Shamai *et al.* [1][4] investigated bounds on the capacity and the achievable information rate of the channel with memory excited by identically distributed (not necessary independent) inputs. They provided lower and upper bounds which can be interpreted as the average mutual information that corresponds to the output of an ideal decision feedback equalizer (DFE) with errorless past decisions (power degradation argument),<sup>1</sup> and a memoryless channel with independent and identically distributed (i.i.d.) inputs that collect the overall energy content of the channel impulse response (power enhancement argument),<sup>2</sup> respectively. While these bounds are obtained in the form of an integral, Arnold *et al.* [5] presented a simulation-based upper bound computed by means of a reduced-state recursion of the BCJR algorithm. Even though this bound is very tight for high SNR, it is very loose for low SNR, due to the fact that the probability mass of the output sequence is spread over all states. In [5] and [6], lower bounds are obtained under mismatched decoding, that is, when the BCJR algorithm assumes that the channel is governed by a transition law that differs from the

<sup>1</sup>In [1] a lower bound on the achievable information rate is obtained based on the ideal post-cursor cancellation where power degradation results from the channel memory introduced by ISI.

<sup>2</sup>In [1] an upper bound on the achievable information rate corresponds to single symbol transmission. For uncoded communications this leads to matched filter lower bound on the error probability as a result of power enhancement.

actual channel law. Unfortunately, these bounds are loose, since minimum-phase representation of a sparse channel with a long memory doesn't collect a significant portion of the energy in the first few channel taps. The same observation holds for the matched-filter representation of the ISI channel, considered in [6]. A more advanced method is developed in [7], where the auxiliary channel for mismatched decoding is constrained to be an arbitrary finite-state machine, and the conditional probabilities between states are optimized at each stage of the algorithm. In a recent work [8], a proposed method for the computation of a lower bound requires evaluation of the magnitude sum of the precursor ISI terms as well as identification of the dominant terms as seen at the output of the MMSE-DFE. The computational load required to obtain a tight bound is high, since the number of dominant terms of the precursor ISI for the sparse channels with large delay spreads is potentially very large. These bounds can be easily extended to multiple-input multiple-output (MIMO) systems based on the results given in [9] and [10]. None of the existing bounds seem to be suitable for ISI channels with long delay spreads and a sparse multipath structure.

In this paper, we investigate bounds on the information rate of a deterministic channel with long memory and sparse multipath structure, assuming i.i.d. inputs drawn from a finite-order modulation alphabet. We propose a lower bound achievable by practical receivers, extending the initial analysis in [11] where a particular reduced-complexity soft-output detection scheme is employed based on the application of the sum-product algorithm to the factor graphs [12]. We analyze the performance of the receiver composed of a finite length impulse response (FIR) filter designed to reduce the number of channel states by applying a channel shortening that does not completely eliminate the ISI. We apply the modified BCJR algorithm of the trellis size matched to the memory of the reduced-state channel. We propose a simulation-based approach for the computation of an upper bound. We group channel taps into several clusters resulting in a newly defined single-input multiple-output (SIMO) channel with a reduced channel memory. From the so obtained SIMO channel we derive an equivalent single-input single-output channel based on maximal ratio combining (MRC) using the maximum likelihood (ML) criterion. Using the BCJR algorithm we estimate the achievable information rate for the derived single-input single-output (SISO) channel and use it to obtain an upper bound on the achievable information rate of the sparse channel of interest. Therefore, we develop our lower and upper bounds by utilizing the modified and the classical BCJR algorithm, respectively, and operating on a trellis of reduced computational complexity.

The paper is organized as follows. In the next section, we describe the system model and the general framework under which new bounds are derived. In Sections III and IV, lower and upper bounds on the achievable information rate of ISI channels are presented, respectively. Section V presents the performance results of the proposed bounds based on different channel examples. Concluding remarks are given in Section VI.

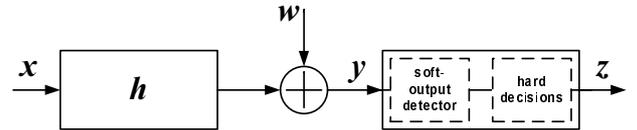


Fig. 1. System model.

## II. SYSTEM MODEL

We consider the transmission scheme depicted in Fig. 1, where it is assumed that the channel input is a sequence of i.i.d. complex-valued symbols drawn from finite-order modulation alphabet of cardinality  $M$ , and transmission is from time index  $n = 0$  to  $n = N - 1$ . This symbol sequence  $\mathbf{x} = \{x_n\}_{n=0}^{N-1}$  is obtained by a direct mapping of information bits to symbols of the modulation alphabet. Assuming observation intervals that are shorter than the coherence time of the channel, and ideal synchronization, the physical propagation environment can be modeled as a fixed discrete-time channel, so that the received sample at time index  $n$  is given as [13]

$$y_n = \sum_{l=0}^L h_l x_{n-l} + w_n, \quad (1)$$

where  $\{w_n\}_{n=0}^{N-1}$  are i.i.d. complex Gaussian random variables with zero-mean and variance  $\sigma^2$  per dimension,  $\mathbf{h} = \{h_n\}_{n=0}^{L-1}$  represents the discrete-time equivalent channel impulse response, and  $L$  is the channel memory. We focus on channels with long memory  $L$ , and very few non-zero channel taps  $h_l$  in order to emphasize the channel *sparseness*. We assume that  $\mathbf{h}$  and  $\sigma^2$  are known at the receiver.

The achievable information rate supported by the system (1) can be expressed as [14]

$$I(X; Y) = \lim_{N \rightarrow \infty} \frac{1}{N} I(\mathbf{x}; \mathbf{y}), \quad (2)$$

which can be estimated as accurately as desired by using the simulation-based BCJR algorithm. Note that  $\mathbf{x}$  is not defined for negative indices, for which we assume that  $x_n = 0$ . This fact will lead to an "edge effect" with a diminishing effect on the term from (2), as  $N$  tends to infinity. This approach requires simulation of a full-complexity forward recursion of the BCJR algorithm that processes a trellis with  $M^L$  states in order to estimate the joint probability of the output sequence. Due to the large values of  $L$ , this simulation-based analysis is practically infeasible. Therefore, upper and lower bounds on  $I(X; Y)$  are required to characterize the achievable information rate for channels with long memory.

### A. Lower Bound on $I(X; Y)$ – Preliminaries

A lower bound can be obtained based on the data-processing inequality [14], by considering an arbitrary receiver that processes the received samples  $\mathbf{y}$ , and produces the decisions  $\mathbf{z} = \{z_n\}_{n=0}^{N-1}$ . In the case of hard-output detection,  $z_n$  is the estimate of the symbol  $x_n$ , and thus belongs to the signal constellation. For both hard and soft-output detection, the data-processing inequality guarantees that

$$\lim_{N \rightarrow \infty} \frac{1}{N} I(\mathbf{x}; \mathbf{z}) = I(X; Z) \leq I(X; Y) \quad (3)$$

The mutual information in (3) still involves infinite-length sequences, which makes it impractical to compute when the channel and the receiver have memory. However, the chain rule for the mutual information guarantees that  $I(X; Z)$  is lower bounded by

$$I_{LB} = I(x_n; z_n) \quad (4)$$

for each  $n$  at which the system is not affected by transient effects. Note that the evaluation of  $I_{LB}$  requires computation of the mutual information between elements of the sequences, and not between the entire sequences as in (3). Hence, similar to [11] where a different soft-output detector is employed, we can evaluate  $I_{LB}$  by first estimating the joint statistics of  $x_n$  and  $z_n$  through long simulations of the channel and the receiver, and then employing numerical methods. The specific receiver adopted for the computation of the lower bound in (4) does not affect its validity, but does affect the tightness of the inequality. In fact, the value of  $I_{LB}$  gives the ultimate information rate supported by a system adopting that specific receiver, when concatenated with a fully-interleaved outer code [15]. Hence, the better the receiver, the tighter the lower bound. Note that iterating between the outer channel decoder and the soft-output detector will increase the information rate at the output of the iterative receiver. Unfortunately, there is no information theoretic support that this side information will not violate the lower bound condition, since data processing inequality is no longer applicable.

Optimal detection is based on the computation of the maximum *a posteriori* probability (APP)  $P(x_n | \mathbf{y})$  of each detected modulation symbol  $x_n$  for each time instance  $n$ , given the received sequence  $\mathbf{y} = \{y_n\}_{n=0}^{N-1}$ . Unfortunately, providing APP as soft-output information is infeasible, since the optimal soft-output BCJR detector is impractical for implementation due to the long memory  $L$  of the channel. Therefore, suboptimal soft-output detection schemes are employed. This topic will be addressed in Section III.

### B. Upper Bound on $I(X; Y)$ – Preliminaries

The upper bound in [1] is obtained based on the information rate computation of a memoryless channel with i.i.d. inputs that collects the overall energy content of the channel impulse response (power enhancement argument). The framework under which this bound is derived motivate us to propose a new class of upper bounds with the tightness governed by the multipath structure of the channel.

In general, we can re-write the mutual information  $I(\mathbf{x}; \mathbf{y})$ , by applying the chain rule [1]:

$$\begin{aligned} I(\mathbf{x}; \mathbf{y}) &= I(\mathbf{x}_0^{N-1}; \mathbf{y}_0^{N-1}) \\ &= \sum_{n=0}^{N-1} I(x_n; \mathbf{y}_0^{N-1} | \mathbf{x}_0^{n-1}) \\ &\leq \sum_{n=0}^{N-1} I(x_n; \mathbf{y}_0^{N-1} | \mathbf{x}_0^{n-1}, x_{n+1}, x_{n+2}, \dots, x_{N-1}), \end{aligned} \quad (5)$$

where the inequality follows from the fact that we consider i.u.d. channel inputs  $\mathbf{x}$ . An abbreviated form  $\mathbf{x}_i^j$  is used to denote the sequence  $\{x_n\}_{n=i}^j$ . This argument, following the derivation in [1], will lead to the well known upper bound on

the achievable information rate for i.i.d. inputs based on MRC (power enhancement argument):

$$I(X; Y) = \lim_{N \rightarrow \infty} \frac{1}{N} I(\mathbf{x}; \mathbf{y}) \leq I(x_n; \rho x_n + w_n), \quad (6)$$

where

$$\rho = \sqrt{\sum_{l=0}^L |h_l|^2}. \quad (7)$$

Generally, this bound is not tight for channels with large delay spreads and strong ISI, meaning that the channel energy  $\rho$  is more or less evenly distributed among all channel taps. The reason is that collecting of the channel energy  $\rho$  into a single channel tap maximizes the achievable information rate among all ISI channels with the same energy constraint. This fact indicates that different channels with the same energy constraint support rates that are between the rates of the original channel (1) and the single-tap channel (6). In fact, we will see in Section IV that a tighter bound on the achievable information rate can be obtained by grouping the channel taps into clusters as in (5), and exploiting the similar method as MRC. This will transform our channel into the newly defined SIMO channel with a reduced size memory, which leads to a less demanding computation of the achievable information rate of the modified channel, i.e., an upper bound on the information rate of the original one.

### III. LOWER BOUND ON INFORMATION RATE WITH I.U.D. INPUTS

In order to compute the lower bound on the achievable information rate as in (4) suboptimal detectors are required that exploit the knowledge of the channel sparseness.

We have considered several receivers in the literature and found that the best performance/complexity trade-off is provided by that proposed in [12], a reduced-complexity soft-output detector based on the application of the sum-product algorithms to factor graphs representing joint APP of the transmitted symbols. The most attractive features of this receiver are its computational complexity, which increases exponentially not with the channel memory  $L$  (as in the BCJR algorithm), but with the number of non-zero taps. The advantage of this algorithm is very clear in the case of UWA channels characterized by long delay spreads and a sparse multipath structure. It was verified that the sum-product algorithm approaches the performance to the optimal BCJR algorithm [12] that computes exact marginal APPs of the transmitted symbols when the factor graph does not have cycles of length less than 6. The proposed lower bound significantly improves the lower bound from [4] as demonstrated by numerical examples previously reported in [11]. However, the receiver used in [11] requires multiple iterations that are critical for the proper functionality of the detector, and calibration of various parameters, which is very challenging for typical UWA channels with large delay spreads. There are also other suboptimal detectors that exploit knowledge of the channel sparseness [16][17].

We consider an alternative receiver structure consisting of a channel shortening filter, and a modified soft-output BCJR detector employed with the objective of obtaining a suboptimal

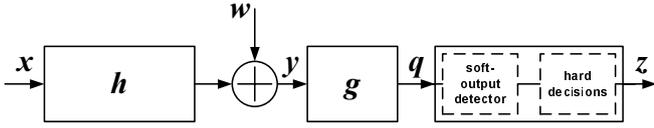


Fig. 2. Reduced-complexity receiver structure.

soft-output detector with a reasonable performance/complexity trade-off, and computational complexity that is lower than the receiver proposed in [12]. Essentially, bounds from this paper and [11] originate from the same idea – to exploit the detector which is suitable for a sparse multipath structure and long channel delays. The receiver shown in Fig. 2 is discussed in the rest of this section.

### A. Prefiltering

Suboptimal detectors can be derived by reducing the memory of the channel at the input of the detector. For this purpose we consider linear channel shortening filters. Prefiltering will usually result in loss of the sparse structure of the channel, what makes our approach applicable even to general (non-sparse) ISI channels. The channel shortening filter in conjunction with the classical BCJR algorithm is no longer the optimal detector for soft-output detection and modifications to the classical BCJR are required in order to improve its performance. Nevertheless, the filter coefficients can be computed efficiently based on various design methods available in the literature, leading to significant complexity reduction.

In this paper, the MMSE method similar to the one described in [18][19] is used to design linear channel shortening filters. Let us define a target channel impulse response  $\mathbf{f} = \{f_n\}_{n=0}^{L'}$  where  $L'$  is the memory of the target channel impulse response, and a MMSE equalizer filter  $\mathbf{g} = \{g_n\}_{n=0}^{\bar{L}}$  with finite memory  $\bar{L}$ . The objective of the MMSE-based approach is to find the target function  $\mathbf{f}$  and the equalizer filter  $\mathbf{g}$  simultaneously by minimizing the mean-squared error between the equalizer output and the desired target. By defining error sequence  $\epsilon_n = \mathbf{f}^H \mathbf{x}_{n-L'}^n - \mathbf{g}^H \mathbf{y}_{n-\bar{L}}^n$ , and under the assumption of independent inputs and independent noise samples, the MSE can be written as [19]

$$E\{|\epsilon_n|^2\} = \epsilon^2 = \mathbf{f}^H R_{xx} \mathbf{f} + \mathbf{g}^H R_{yy} \mathbf{g} - \mathbf{f}^H R_{xy} \mathbf{g} - \mathbf{g}^H R_{yx} \mathbf{f}, \quad (8)$$

where  $R_{xx} = E\{\mathbf{x}_{n-L'}^n (\mathbf{x}_{n-L'}^n)^H\}$ ,  $R_{yy} = E\{\mathbf{y}_{n-\bar{L}}^n (\mathbf{y}_{n-\bar{L}}^n)^H\}$  and  $R_{xy} = E\{\mathbf{x}_{n-L'}^n (\mathbf{y}_{n-\bar{L}}^n)^H\}$  are  $(L'+1) \times (L'+1)$  correlation,  $(\bar{L}+1) \times (\bar{L}+1)$  correlation and  $(L'+1) \times (\bar{L}+1)$  cross-correlation matrices, respectively.

Minimization of the MSE given by (8) with respect to both  $\mathbf{f}$  and  $\mathbf{g}$ , is done under the constraint that  $g_0 = 1$  to avoid the trivial solution  $\mathbf{f} = 0$  and  $\mathbf{g} = 0$ . This constraint corresponds to the minimum-phase solution to the decision feedback equalizer design problem for large  $L'$ . In [18] it is shown that this constraint introduces less correlation in the noise at the output of the equalizer filter than other constraints such as fixed energy or partial response targets. In order to minimize the MSE subject to  $g_0 = 1$ , we write the Lagrangian in the following form:

$$\epsilon^2 = \mathbf{f}^H R_{xx} \mathbf{f} + \mathbf{g}^H R_{yy} \mathbf{g} - \mathbf{f}^H R_{xy} \mathbf{g} - \mathbf{g}^H R_{yx} \mathbf{f} - 2\lambda(\mathbf{e}_0^H \mathbf{f} - 1), \quad (9)$$

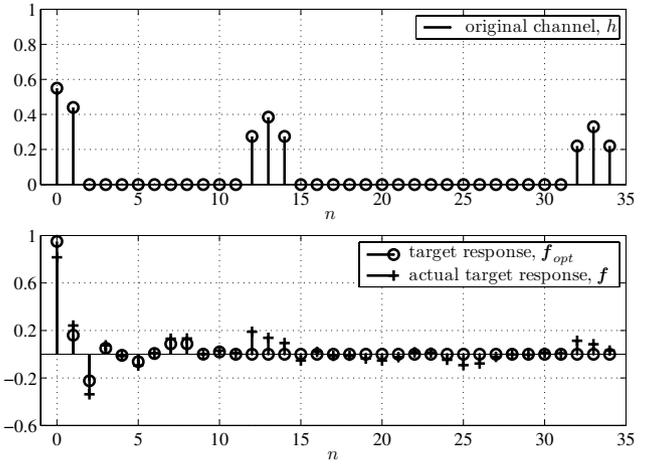


Fig. 3. Example of the optimal and actual target response with  $L' = 9$  for an underwater channel impulse response with  $L = 34$ .

where  $\mathbf{e}_0$  is a unit vector with first element having value 1 and all other elements set to 0. Taking the derivative of MSE with respect to  $\mathbf{f}$ , we obtain the following solution for the optimal target response at the output of the equalizer filter:

$$\mathbf{f}_{opt} = \frac{(R_{xx} - R_{xy} R_{yy}^{-1} R_{yx})^{-1} \mathbf{e}_0}{\mathbf{e}_0^H (R_{xx} - R_{xy} R_{yy}^{-1} R_{yx})^{-1} \mathbf{e}_0}. \quad (10)$$

Consequently, the optimal solution for the equalizer filter is obtained as

$$\mathbf{g}_{opt}^H = \mathbf{f}_{opt}^H R_{xy} R_{yy}^{-1}. \quad (11)$$

We note that other constraints may result in a lower MSE in (8), but at the expense of increased correlation among noise samples at the output of the equalizer filter  $\mathbf{g}$ .

In Fig. 3, we provide an example of the channel shortening filter applied to a shallow water acoustic channel impulse response (measured in KAM08 experiment [20]). The optimal target channel impulse response  $\mathbf{f}_{opt}$  is assumed to be of length  $L' = 9$ . We observe that the actual target channel impulse response  $\mathbf{f}$  (obtained by convolving  $\mathbf{h}$  and  $\mathbf{g}_{opt}$ ) is close to  $\mathbf{f}_{opt}$  (i.e. for  $n \geq L' + 1$ , the channel taps  $f_n$  are sufficiently small).

We assume that the channel output (1) is filtered by the linear equalizer  $\mathbf{g}_{opt}$  from (11) and produces output sequence  $\mathbf{q} = \{q_n\}_{n=0}^{N-1}$ . We consider equalizer filters which are nonsingular (this is guaranteed from the MMSE design [21]), or filter approximations [22], for which there exists a stable invertible filter. Under this condition, we can write

$$I(X; Y) = I(X; Q) = \lim_{N \rightarrow \infty} \frac{1}{N} I(\mathbf{x}; \mathbf{q}), \quad (12)$$

where the sequence  $\mathbf{q}$  preserves the information rate supported by the original channel output sequence  $\mathbf{y}$ . The achievable information rate  $I(X; Q)$  is bounded as in (3) by employing a soft-output detector to the prefiltered sequence. Therefore, following the channel shortening filter  $\mathbf{g}_{opt}$  given by (11), the classical BCJR algorithm with a  $M^{L'}$  state trellis can be applied to the prefiltered sequence  $\mathbf{q}$ . Note that the classical BCJR is mismatched by coloring of the noise at the output of the equalizer, and a difference between the actual and the optimal target impulse responses as illustrated in Fig. 3.

### B. Modified BCJR

The noise correlation induced by  $\mathbf{g}_{opt}$  degrades the performance of the classical BCJR algorithm significantly, and modifications of the classical algorithm are required. Therefore, we propose a modified BCJR detector based on noise prediction. For this trellis-based algorithm, let us define, at each time instance  $n$ , the state  $s_n$  as

$$s_n = (x_{n-L'}, x_{n-L'+1}, \dots, x_{n-2}, x_{n-1}). \quad (13)$$

The branch metric of the classical BCJR algorithm is given as [23][24]

$$\gamma_n(x_n, s_n) = \frac{1}{M} \exp \left( -\frac{1}{2\sigma^2} \left| y_n - \sum_{l=0}^L h_l x_{n-l} \right|^2 \right). \quad (14)$$

Ideally, the branch metrics (14) can be modified by whitening the noise with infinite length impulse response filters. Because the linear filter design based on the MMSE criterion presented in the previous subsection introduces less noise correlation than other similar methods, we observed that an ideal infinite-length invertible filter can be approximated reliably by an FIR filter  $\mathbf{g}_{inv} = \{g_{inv,n}\}_{n=0}^{\hat{L}}$ , where  $\hat{L}$  is the memory of the finite length approximation. Therefore, following similar arguments as in [25], the branch metric in (14) can be modified as

$$\gamma_n^*(x_n, s_{n-\hat{L}}^n) = \frac{1}{M} \exp \left( -\frac{\left| \mathbf{g}_{inv}^H (\mathbf{q}_{n-\hat{L}}^n - \mathbf{p}(x_n, s_{n-\hat{L}}^n)) \right|^2}{2\sigma^2} \right), \quad (15)$$

where  $\mathbf{p}(x_n, s_{n-\hat{L}}^n) = \{p_k\}_{k=n-\hat{L}}^{k=n}$  is a sequence of noiseless outputs of the target channel impulse response,  $p_k = \sum_{l=0}^{L'} f_l x_{k-l}$ . Since our modified BCJR algorithm has a trellis of size  $M^{L'}$ , the channel inputs  $x_{n-i}$ , for  $L'+1 \leq i \leq L'+\hat{L}$ , are not available at time instant  $n$ , forcing us to use tentative decisions  $\hat{x}_{n-i}$ . These decisions can be obtained during a modified forward recursion of the BCJR algorithm as described in the remainder of this section.

Besides the branch metric formulation, an important part of the classical BCJR algorithm is given by two recursive equations [24]:

$$\alpha_{n+1}(s_{n+1}) = \sum_{x_n, s_n} I(x_n, s_n, s_{n+1}) \gamma_n(x_n, s_n) \alpha_n(s_n), \quad (16)$$

$$\beta_n(s_n) = \sum_{x_n, s_{n+1}} I(x_n, s_n, s_{n+1}) \gamma_n(x_n, s_n) \beta_{n+1}(s_{n+1}), \quad (17)$$

where  $\alpha_n(s_n)$  and  $\beta_n(s_n)$  are the state metrics of *forward recursion* (16) and *backward recursion* (17), respectively.  $I(x_n, s_n, s_{n+1})$  is the indicator function equal to 1 if  $x_n$ ,  $s_n$  and  $s_{n+1}$  satisfy the trellis constraints, and 0 otherwise. We will assume that the first state  $\bar{s}_0$  and the last state  $\bar{s}_{N+1}$  of the BCJR algorithm are known. Therefore, recursions (16) and (17) are initialized by  $\alpha_0(\bar{s}_0) = 1$  and 0 for other state metrics, and  $\beta_{N+1}(\bar{s}_{N+1}) = 1$  and 0 for other state metrics, respectively. The modification of the forward recursion given in equation (16) is based on the concept of survivor paths from the classical Viterbi algorithm [26] and it is given by

$$\alpha_{n+1}^*(s_{n+1}) = \max_{x_n, s_n} I(x_n, s_n, s_{n+1}) \gamma_n^*(x_n, s_n) \alpha_n^*(s_n). \quad (18)$$

Once the tentative decisions  $\hat{x}_{n-i}$  for  $L'+1 \leq i \leq L'+\hat{L}$  are obtained, we can compute the modified backward recursion in the following form:

$$\beta_n^*(s_n) = \sum_{x_n, s_{n+1}} I(x_n, s_n, s_{n+1}) \gamma_n^*(x_n, s_n) \beta_{n+1}^*(s_{n+1}). \quad (19)$$

Note that the concept of survivor paths is not needed for backward recursion since tentative decisions are already available. We should point out that determining tentative decisions from the forward recursion (18) is more reliable in the case of a minimum-phase channel, suggesting that without any loss in the achievable information rate, every channel of interest should be transformed to a minimum-phase representation before the modified BCJR detector is employed. For maximum-phase systems, the alternative is to use the backward recursion to determine tentative decisions on transmitted symbols. The lower bound (4) is computed numerically by estimating the first order joint statistics of the channel input  $x_n$  and the soft-output  $z_n$  of the modified BCJR algorithm, as done in [15].

Note that the proposed receiver structure is applicable to non-sparse ISI patterns as well, but the effectiveness of the proposed approach is strictly limited by the design of the channel shortening filter. The equalizer filter from (11) would have to suppress most of the ISI which is outside the scope of the memory length  $L'$  of the target channel impulse response (11). We conducted extensive simulations which indicate that channel shortening is not effective for the channel with long memory and non-sparse structure, if the memory of the target channel impulse response is reasonably short. The reason is a high correlation among the noise samples at the output of the equalizer, which seriously limits the performance of the modified BCJR algorithm.

### IV. UPPER BOUND ON INFORMATION RATE WITH I.U.D. INPUTS

Let us consider a SIMO ISI channel with the channel outputs given by

$$y_{k,n} = h_{k(L'+1)} x_n + h_{k(L'+1)+1} x_{n-1} + \dots + h_{k(L'+1)+L'} x_{n-L'} + w_{k,n}, \quad (20)$$

where  $k = 0, 1, \dots, K-1$ ,  $K = \lceil (L+1)/(L'+1) \rceil$ ,  $n = 0, 1, \dots, N-1$  and  $w_{k,n}$  is i.i.d. complex Gaussian random variable in both indices  $k$  and  $n$ , with zero-mean and variance  $\sigma^2$  per dimension. The number of outputs (clusters) of the SIMO ISI channel is bounded by  $K \leq L$ , and depends on the sparse structure of the original channel (1). The achievable information rate of the SIMO channel from (20) can be expressed as

$$I(X; Y_1, Y_2, \dots, Y_K) = \lim_{N \rightarrow \infty} \frac{1}{N} I(\mathbf{x}; \mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_K), \quad (21)$$

where  $\mathbf{y}_k = \{y_{k,n}\}_{n=0}^{N-1}$ . Note from the expression (20) that the computational complexity of the achievable information rate of the derived SIMO channel is driven by the cluster with the longest memory  $L_{eq} = \max_k L_k \leq L'$ , where  $L_k$  is the memory of  $k$ -th cluster, and  $k = 1, 2, \dots, K$ . The SIMO ISI channel of interest is illustrated in Fig. 4. We note that noise in each cluster is i.i.d. with the variance equal to  $\sigma^2$ .

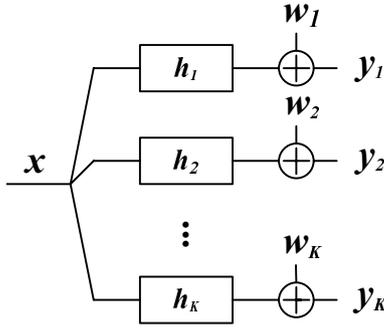


Fig. 4. The derived SIMO ISI channel with  $K$  clusters of memory  $L_k$ ,  $k = 1, 2, \dots, K$ .

Instead of the inequality in (6), we can provide a tighter bound on the information rate by noting that we can group channel inputs (for simplicity) into blocks:

$$\begin{aligned} I(\mathbf{x}; \mathbf{y}) &= I(\mathbf{x}_0^{N-1}; \mathbf{y}_0^{N-1}) \\ &= I(\mathbf{a}_0^{N-1}; \mathbf{y}_0^{N-1}) \\ &= \sum_{n=0}^{N-1} I(\mathbf{a}_n; \mathbf{y}_0^{N-1} | \mathbf{a}_0^{n-1}) \\ &\leq \sum_{n=0}^{N-1} I(\mathbf{a}_n; \mathbf{y}_{1,0}^{N-1}, \mathbf{y}_{2,0}^{N-1}, \dots, \mathbf{y}_{K,0}^{N-1} | \mathbf{a}_0^{n-1}), \end{aligned} \quad (22)$$

where  $\mathbf{a}_n = \{x_i\}_{i=n-L}^n$ . The inequality follows from the fact that we consider i.u.d. channel inputs  $\mathbf{x}$  and multiple channel observations. This bound is strictly tighter than the one given by the inequality in (6) since we have excluded from the conditioning in (5) the future blocks  $\mathbf{a}_{n+L'+1}^{N-1}$ . Hence, we can bound the achievable information rate by that of the SIMO channel obtained by grouping the channel taps into no more than  $L$  clusters (by setting  $K = L+1$  and conditioning on the future blocks we obtain the bound in (5)). At this point, we should indicate that grouping the channel inputs into blocks  $\mathbf{a}_n$  of different size in (22), will organize the channel taps differently in order to exploit the knowledge of the sparse structure.

Since this approach for deriving an upper bound is motivated by the desire to avoid the long channel memory, we prefer to group the channel inputs such that the channel taps in all the clusters are either equally spaced, so that the channel is potentially compressible, or all  $L_k$ 's (and consequently  $L_{eq}$ ) are relatively small. From an information theoretical viewpoint we emphasize that relative delays among clusters in the derived SIMO ISI channel in (20) are irrelevant. Therefore, for simplicity, we assume that the first tap in each cluster has zero delay and the last tap is at delay  $L_k$ , i.e. the taps in each cluster in the expression (20) are advanced by  $n$  time instances. Now, the  $k$ -th received element at time instant  $n$  can be written as

$$y_{k,n} = \sum_{l=0}^{L_k} h_{k,l} x_{n-l} + w_{k,n}, \quad (23)$$

where  $h_{k,l}$  are determined from (20). It is useful to re-write the model in the following matrix notation:

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x} + \mathbf{w}_k, \quad (24)$$

where  $\mathbf{y}_k$ ,  $\mathbf{x}$ , and  $\mathbf{w}_k$  are column vectors collecting  $\{y_{k,n}\}$ ,  $\{x_n\}$ , and  $\{w_{k,n}\}$ , respectively, while  $\mathbf{H}_k$  is a matrix constructed from the ISI vector  $\{h_{k,n}\}$ . In the matrix model (24), the dimension of the column vectors is  $N \times 1$ , while the dimension of the matrix  $\mathbf{H}_k$  is  $N \times N$ . The information rate of the system is completely defined by the conditional probability density function of the received vector  $(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_K)$  given  $\mathbf{x} = \mathbf{a}$ , denoted by  $p(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_K | \mathbf{a})$ , and can be written as

$$\begin{aligned} p(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_K | \mathbf{a}) &= \prod_{k=1}^K p(\mathbf{y}_k | \mathbf{a}) \\ &= \prod_{k=1}^K (2\pi\sigma^2)^{-\frac{K}{2}} \exp\left\{-\frac{\|\mathbf{y}_k - \mathbf{H}_k \mathbf{a}\|^2}{2\sigma^2}\right\} \end{aligned} \quad (25)$$

By expanding the expression on the right-hand side of (25) and neglecting factors which do not depend on  $\mathbf{a}$ , we can write

$$\begin{aligned} p(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_K | \mathbf{a}) &\propto \prod_{k=1}^K \exp\left\{\frac{2\mathbf{y}_k^H \mathbf{H}_k \mathbf{a} - \mathbf{a}^H \mathbf{H}_k^H \mathbf{H}_k \mathbf{a}}{2\sigma^2}\right\} \\ &= \prod_{k=1}^K \exp\left\{\left(\mathbf{x}^H - \frac{1}{2}\mathbf{a}^H\right) \mathbf{G}_k \mathbf{a} + \mathbf{N}_k \mathbf{a}\right\}, \end{aligned} \quad (26)$$

where we defined

$$\mathbf{G}_k = \frac{\mathbf{H}_k^H \mathbf{H}_k}{\sigma^2}, \quad (27)$$

$$\mathbf{N}_k = \frac{\mathbf{w}_k^H \mathbf{H}_k}{\sigma^2}. \quad (28)$$

Finally, we can re-write (26) as

$$p(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_K | \mathbf{a}) \propto \exp\left\{\left(\mathbf{x}^H - \frac{1}{2}\mathbf{a}^H\right) \mathbf{G}_e \mathbf{a} + \mathbf{N}_e \mathbf{a}\right\}, \quad (29)$$

where we defined

$$\mathbf{G}_e = \sum_{k=1}^K \mathbf{G}_k, \quad (30)$$

$$\mathbf{N}_e = \sum_{k=1}^K \mathbf{N}_k. \quad (31)$$

It can be easily shown that  $\mathbf{G}_e$  is a valid autocorrelation matrix and that

$$E\left\{\mathbf{N}_e^H \mathbf{N}_e\right\} = \mathbf{G}_e. \quad (32)$$

In practice, the relationship in (29) indicates that the information rate of the system is completely defined by the matched-filter representation of the ISI channel, scaled by the noise variance. Particularly, different noise-whitened representations are equivalent when their matched-filter representations are equal, which is well known. Hence, the SIMO channel is equivalent to a SISO channel with matched-filter representation  $\mathbf{G}_e$ . Practically, the outputs  $\mathbf{z}_e = \{z_e\}_{n=0}^{N-1}$  of the equivalent SISO channel in a matched-filter representation can be obtained by matched-filtering and scaling (by the noise variance) of the cluster outputs,  $\mathbf{y}_k$ 's, followed by the summing operator. Therefore, the relationship in (29) implies identical achievable information rates of the SIMO channel in (23) and

the equivalent SISO channel in (29). From the inequality (22), it follows that

$$\begin{aligned}
 I(X; Y) &= \lim_{N \rightarrow \infty} \frac{1}{N} I(\mathbf{x}; \mathbf{y}) \\
 &\leq \lim_{N \rightarrow \infty} \frac{1}{N} I(\mathbf{x}; \mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_K) \\
 &= \lim_{N \rightarrow \infty} \frac{1}{N} I(\mathbf{x}; \mathbf{z}_e) \\
 &= I_{UB},
 \end{aligned} \tag{33}$$

where we denote the output sequence of the equivalent SISO channel as  $\mathbf{z}_e = \{z_e\}_{n=0}^{N-1}$ . The upper bound  $I_{UB}$  is thus obtained on the achievable information rate for channels with memory. Our first observation is that unlike the bound derived in [1], which depends only on the overall power contained in the channel impulse response, the upper bound (33) depends on the channel structure as well. Note that the computational complexity of this bound depends exponentially on  $L_{eq} \leq L'$ . It is feasible to optimize  $I_{UB}$ , and it can be accomplished through two degrees of freedom inherently present in our approach: the number of clusters  $K$  in (23) and the distribution of the channel taps among the  $K$  clusters. Considering the number of clusters  $K$ , our preliminary results indicate that by choosing fewer clusters, better performance is obtained. Intuitively, this is expected, because in this way we minimize the number of observation sequences ( $\mathbf{y}_k$ 's) at the receiver. With regard to the distribution of the channel taps among clusters, we rely on our observations which suggest that we should not keep the strongest path alone in the ISI-free cluster (if such a cluster exists), but combine it with other taps. Another observation is that the bound is tighter if clusters have a power ratio that tends to a small value, i.e. we should place the weakest tap into a separate cluster, if feasible.

For non-sparse ISI channels, the idea of clustering the channel taps is also applicable. However, computational complexity of the corresponding upper bound is very high if we are aiming at obtaining a tight bound, i.e. if we split the channel taps into two clusters only, the memory of the SIMO channel will be half of the memory of the original channel, and it cannot be further compressed due to the channel non-sparseness. Increasing the number of clusters will degrade the tightness of the bound for a general ISI channel.

## V. EXAMPLES

In this section, the performance of the simulation-based bounds that were developed in Section III and Section IV are presented. For each of the channel examples, we will compare the performance of the proposed bounds with those reported in [1], [4] and [5]. The channel impulse responses considered in the computer simulation are shown in Table I.

In Fig. 5 we consider the achievable information rates with BPSK ( $M = 2$ ) transmission over the minimum-phase channel  $C_1$ . Note that the channel is sparse with just three non-zero taps, and its delay spread is  $L = 7$ . All the zeros of channel  $C_1$  in the  $Z$ -domain are inside the unit circle, making it a minimum-phase system. Consequently, the simulation-based upper bound [5] is very loose as it can be seen from Fig. 5. In order to obtain the proposed upper bound, we group the channel taps  $C_1(0) = 1.0$  and  $C_1(4) = 0.6$  into one cluster,

TABLE I  
THE CHANNEL IMPULSE RESPONSE EXAMPLES.

Channel	$L$	$f$
$C_1$	7	$\frac{1}{\sqrt{1.45}} [1 \ 0 \ 0 \ 0 \ .6 \ 0 \ 0 \ .3]$
$C_2$	5	$\frac{1}{\sqrt{1.45}} [1 \ 0 \ 0 \ .6 \ 0 \ .3]$
$C_3$	7	$\frac{1}{\sqrt{1.45}} [.3 \ 0 \ 0 \ .6 \ 0 \ 0 \ 0 \ 1]$
$C_4$	7	$\frac{1}{\sqrt{1.49}} [.7 \ 0 \ 0 \ 0 \ .8 \ 0 \ 0 \ .6]$
$C_5$	50	$\frac{1}{\sqrt{3.31}} [1 \ .8 \ 0 \ (25) \ .5 \ .7 \ .5 \ 0 \ (18) \ .4 \ .6 \ .4]$

and the weak tap  $C_1(7) = 0.3$  into the second cluster. The proposed upper bound now outperforms the simulation-based bound [5], and as it was shown in Section IV, performs better than the power enhancement bound [1] over the entire SNR region. The former approach is illustrated in Fig. 5 as bound  $I_{UBnew, sub}$  where we keep the first channel tap  $C_1(0)$  alone. This is clearly poor choice of channel clustering and the bound is very close to the power enhancement bound. The latter approach is trivial for the case of the channel with only three non-zero taps, since the memory of our equivalent SISO channel  $\mathbf{z}_e$  reduces to  $L' = 1$ . We also note that our lower bound is tighter than bounds [4] achievable by practical receivers, where we assumed that the memory of the target channel response is  $L' = 3$ . This also indicates that tentative channel decisions obtained through the modified forward recursion of the BCJR algorithm in (18) are reliable, since the filter  $\mathbf{g}_{opt}$  induces noise correlation that is mainly suppressed through the perfectly known channel inputs from the state definition in (13). We also compare our bound with the lower bound we presented previously in [11], which utilizes a detector based on the application of the sum-product algorithm to the factor graph of the channel impulse response. The performance of this detector converges to that of the optimal BCJR detector, since  $C_1$  does not have cycles of length less than 6. We observe that our bound, based on a channel shortening filter and the modified BCJR detector, converges to the optimal detector performance for high SNR, while slight degradation is observed for low SNR, where tentative decisions are less reliable.

In Fig. 6 we consider a minimum-phase channel  $C_2$  with a channel structure similar to  $C_1$ . The span of the ISI is shorter, and for  $L' = 3$  the simulation-based upper bound [5], based on the reduced-search on the full-complexity trellis, is tighter than the one for  $C_1$ . The reason for such behavior is that more channel states carrying significant portion of the estimated probability mass of the channel output sequence are included into reduced-search. Since we use the same strategy of grouping channel taps into the clusters, our upper bound is the same as in the case of the previous channel  $C_1$ . We also observe that the tightness of our lower bound is slightly improved, since prefiltering will introduce less correlation in the noise by canceling out only the weak ISI components for tap indices strictly greater than  $L' = 3$ . We note that our bound performs almost as well as the one previously proposed [11] for factor graph-based detector.

In Fig. 7 we consider a maximum-phase representation of  $C_1$ . The tightness of our upper bound is similar to the one for the minimum-phase channel, since the power ratio among

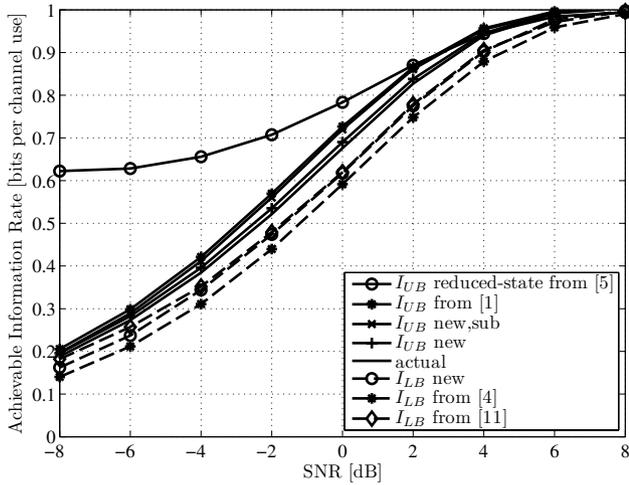


Fig. 5. Bounds on the achievable information rate for minimum-phase channel  $C_1$ . Target memory  $L' = 3$  is assumed for the computation of both,  $I_{UBnew}$  and  $I_{LBnew}$ .

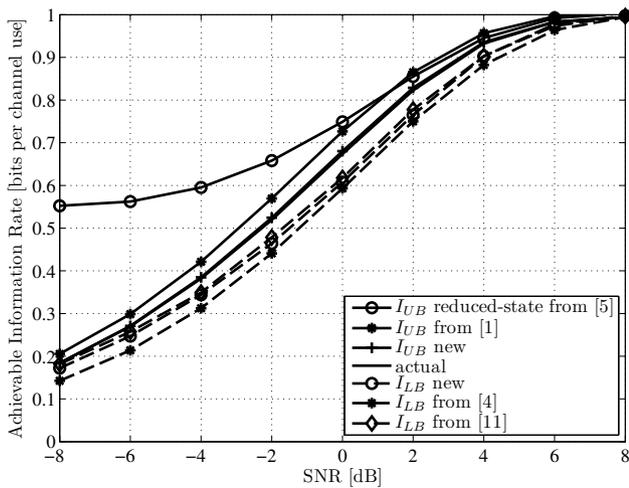


Fig. 6. Bounds on the achievable information rate for minimum-phase channel  $C_2$ . Target memory  $L' = 3$  is assumed for the computation of both,  $I_{UBnew}$  and  $I_{LBnew}$ .

channel taps is the same as in the previous examples. The simulation-based upper bound [5] is significantly looser for the entire SNR region, and therefore, it is omitted from the figure. As for the lower bound, the performance is seriously degraded unless we redefine our approach to increase the reliability of tentative decisions. This can be accomplished by running the modified backward recursion in (19) as in (18), and then the modified forward recursions with these decisions. With this approach, we observe significant improvement in the receiver performance as shown in Fig. 7 by comparing  $I_{LBnew,backward}$  with  $I_{LBnew,forward}$ . We note that the lower bound based on the power degradation argument is sensitive to the phase of the channel. Since  $C_3$  and  $C_1$  are equivalent channels, identical performance comparison as in Fig. 5 between our bound and the one previously proposed [11] is observed.

In Fig. 8 we consider a mixed-phase channel  $C_4$ . For this channel the upper bound is similar to those from the previous examples, and obtained by grouping  $C_4(0)$  and  $C_4(4)$

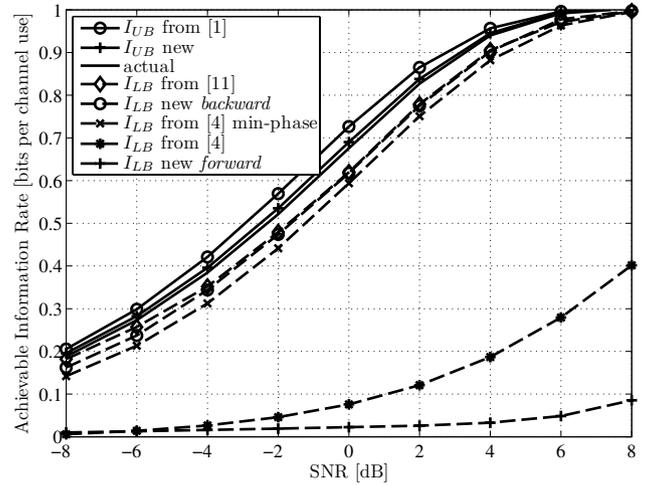


Fig. 7. Bounds on the achievable information rate for maximum-phase channel  $C_3$ . Target memory  $L' = 3$  is assumed for the computation of  $I_{UBnew}$ ,  $I_{LBnew,forward}$  and  $I_{LBnew,backward}$ .

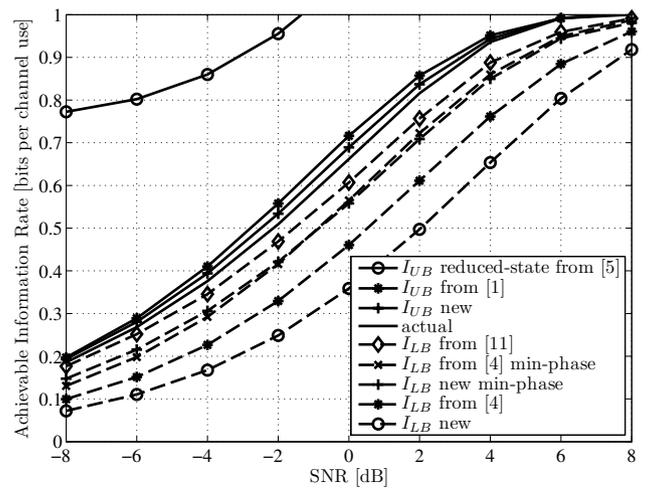


Fig. 8. Bounds on the achievable information rate for mixed-phase channel  $C_4$ . Target memory  $L' = 3$  is assumed for the computation of  $I_{UBnew}$ ,  $I_{LBnew}$  and  $I_{LBnew,min-phase}$ .

into one cluster and  $C_4(7)$  into a second one, whereas the lower bound shows a degradation due to unreliable tentative decisions, obtained either by means of forward or backward recursion. We also observe that the bound based on factor graphs [11] outperforms our bound  $I_{LB}$  significantly. This result suggests that a mixed-phase type of channel may be transformed either to a minimum-phase or a maximum-phase representation ( $I_{LBnew,min-phase}$ ) prior to the application of the reduced-state modified BCJR. However, this method will certainly provide a looser lower bound than the one treated in our approach, because the modified BCJR shows the best performance for the minimum-phase channel representation. The validity of the bound obtained from the minimum-phase representation, follows from the fact that the achievable information rate of the channel is insensitive to the channel phase.

In Fig. 9 we illustrate the proposed bounds for the channel  $C_5$  (with long memory of  $L = 50$ ), and clustered non-zero tap pattern, where few consecutive channel taps carry

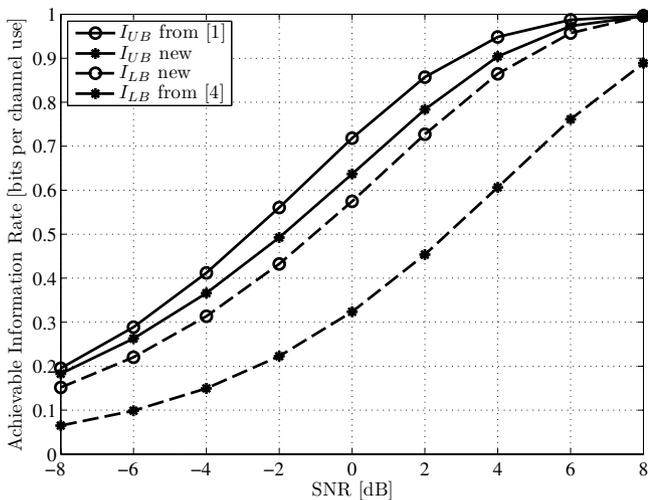


Fig. 9. Bounds on the achievable information rate for minimum-phase channel  $C_5$ . Target memory  $L' = 9$  is assumed for the computation of both,  $I_{UBnew}$  and  $I_{LBnew}$ .

significant energy. The simulation-based upper bound [5] is omitted since it is loose for the entire SNR region. We derive the proposed upper bound by assigning each of the channel clusters to the clusters of the SIMO channel (23). As we observe, the proposed upper bound shows a significant improvement with respect to the bound [1] based on the power enhancement argument. The new lower bound on the channel impulse response shows significant improvement with respect to the one based on the power degradation argument due to strong ISI components present in the large delay spread of the channel  $C_5$ . We emphasize that  $L' = 9$  is considered to be sufficiently large in order for the actual target channel impulse response  $f$  to be relatively close to the optimal one  $f_{opt}$  for this example. Since the receiver used in [11] requires calibration of various parameters that are critical for the proper functionality of the detector, this bound was not considered in the case of the realistic UWA channel with the large delay spread, as in the case of  $C_5$ . The newly proposed detector has slight performance degradation, and significantly lower implementation complexity, which is the main reason why we favor it.

## VI. CONCLUDING REMARKS

Upper and lower bounds on the achievable information rate of the discrete-time additive white Gaussian channel with long memory and sparse structure are examined. The proposed lower bound is attainable by practical receivers, with varying degrees of tightness depending upon the specific receiver used. Particularly we consider a receiver structure consisting of a channel shortening filter and a modified BCJR detector with the objective of obtaining a suboptimal soft-output detector with reasonable performance/complexity trade-off. The lower bound improves upon usefulness of the existing bounds for minimum-phase and maximum-phase channels, and indicates the utility of the considered reduced-complexity receiver. For mixed-phase systems, the lower bound shows poor performance due to the unreliable tentative decisions made by the modified BCJR algorithm. The proposed upper

bound is based on grouping of the channel taps into clusters. Performance of the upper bound depends on the strategy for channel clustering, and it can be optimized further by computer search. For sparse channels with a long memory, this approach leads to a significant complexity reduction in the simulation of the newly defined SISO channel, which also yields the upper bound on the achievable information rate of the actual channel. As illustrated by examples, the upper bound significantly outperforms the bounds previously published in the literature.

## REFERENCES

- [1] S. Shamai, L. H. Ozarow, and A. D. Wyner, "Information rates for a discrete-time Gaussian channel with intersymbol interference and stationary inputs," *IEEE Trans. Inf. Theory*, vol. 37, pp. 1527–1539, Nov. 1991.
- [2] D. Arnold and H.-A. Loeliger, "On the information rate of binary-input channels with memory," *Proc. 2001 IEEE Int. Conf. Commun.*, vol. 9, pp. 2692–2695.
- [3] H. D. Pfister, J. B. Soriaga, and P. H. Siegel, "On the achievable information rates of finite state ISI channels," in *Proc. 2001 IEEE GLOBECOM*, pp. 2992–2996.
- [4] S. Shamai (Shitz) and R. Laroia, "The intersymbol interference channel: lower bounds on capacity and channel precoding loss," *IEEE Trans. Inf. Theory*, vol. 42, no. 5, pp. 1388–1404, Sep. 1996.
- [5] D. M. Arnold, H. Loeliger, P. O. Vontobel, A. Kavcic, and W. Zeng, "Simulation-based computation of information rates for channels with memory," *IEEE Trans. Inf. Theory*, vol. 52, pp. 3498–3508, Aug. 2006.
- [6] F. Rusek and D. Fertonani, "Lower bounds on the information rate of intersymbol interference channels based on the Ungerboeck observation model," in *Proc. 2009 IEEE International Symp. Inf. Theory*.
- [7] P. Sadeghi, P. O. Vontobel, and R. Shams, "Optimization of information rate upper and lower bounds for channels with memory," *IEEE Trans. Inf. Theory*, vol. 55, pp. 663–688, Feb. 2009.
- [8] S. Jeong and J. Moon, "Computing lower bounds on the information rate of the intersymbol interference channels," arXiv:1001.3911v1 [cs.IT], Jan. 2010.
- [9] W. J. Higley, "Self-adaptive processes for the mitigation of coherent multipath in ocean acoustics," Ph.D. dissertation, University of California at San Diego, 2007.
- [10] Z. Zhang, T. M. Duman, and E. M. Kurtas, "Achievable information rates and coding for MIMO systems over ISI channels and frequency-selective fading channels," *IEEE Trans. Commun.*, vol. 52, pp. 1698–1710, Oct. 2004.
- [11] A. Radosevic, D. Fertonani, T. M. Duman, J. G. Proakis, and M. Stojanovic, "Capacity of MIMO systems in shallow water acoustic channels," in *Proc. 2009 IEEE Asilomar Conf. Signals, Syst. Comput.*
- [12] G. Colavolpe and G. Germini, "On the application of factor graphs and the sum-product algorithm to ISI channels," *IEEE Trans. Commun.*, vol. 53, pp. 818–825, May 2005.
- [13] J. G. Proakis, *Digital Communications*, 4th edition. McGraw-Hill, 2001.
- [14] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, Inc., 1991.
- [15] S. ten Brink, G. Kramer, and A. Ashikhmin, "Design of low-density parity-check codes for modulation and detection," *IEEE Trans. Commun.*, vol. 52, no. 4, pp. 670–678, Apr. 2004.
- [16] N. Benvenuto and R. Marchesani, "The Viterbi algorithm for sparse channels," *IEEE Trans. Commun.*, vol. 44, no. 3, pp. 287–289, Mar. 1996.
- [17] N. C. McGinty, R. A. Kennedy, and P. Hoeher, "Parallel trellis Viterbi algorithm for sparse channels," *IEEE Commun. Lett.*, vol. 2, no. 5, pp. 143–145, May 1998.
- [18] J. Moon and W. Zeng, "Equalization for maximum likelihood detectors," *IEEE Trans. Magn.*, vol. 31, pp. 1083–1087, Mar. 1995.
- [19] N. Al-Dhahir and J. M. Cioffi, "Efficiently computed reduced-parameter input-aided MMSE equalizers for ML detection: a unified approach," *IEEE Trans. Inf. Theory*, vol. 42, pp. 903–915, May 1996.
- [20] A. Radosevic, J. G. Proakis, and M. Stojanovic, "Statistical characterization and capacity of shallow water acoustic channels," in *Proc. 2009 IEEE Oceans Conf.*
- [21] J. Chow and J. M. Cioffi, "A cost-effective maximum likelihood receiver for multicarrier systems," in *Proc. 1992 Int. Conf. Commun.*, pp. 948–952.

- [22] N. Al-Dhahir, A. H. Sayed, and J. M. Cioffi, "A high-performance cost-effective pole-zero MMSE-DFE," in *Proc. 1993 Allerton Conf. Commun., Control, Comput.*, pp. 1166–1175.
- [23] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inf. Theory*, vol. 20, pp. 284–287, Mar. 1974.
- [24] D. Fertonani, A. Barbieri, and G. Colavolpe, "Reduced-complexity BCJR algorithm for turbo equalization," *IEEE Trans. Commun.*, vol. 55, no. 12, pp. 2279–2287, Dec. 2007.
- [25] A. Kavcic and J. M. F. Moura, "The Viterbi algorithm and Markov noise memory," *IEEE Trans. Commun.*, vol. 46, no. 1, pp. 291–301, Jan. 2000.
- [26] A. J. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. Inf. Theory*, vol. IT-13, pp. 260–269, Apr. 1967.



**Andreja Radosevic** (S'09) received the B.S. degree in electrical engineering from the University of Belgrade, Belgrade, Serbia, in 2007, and the M.S. degree in electrical and computer engineering from the University of California, San Diego, in 2009, where he is currently pursuing the Ph.D. degree in electrical and computer engineering.

His current research interests include channel capacity analysis, adaptive modulation, channel estimation, and orthogonal frequency division multiplexing (OFDM) techniques mainly in the context

of underwater acoustic communications.



**Dario Fertonani** was born in Mantua, Italy, in 1980. He received the M.S. degree in telecommunications engineering in 2004 and the Ph.D. degree in information technology in 2008, both from the University of Parma, Parma, Italy. He held a postdoctorate position at Arizona State University, Tempe, AZ, from May 2008 to December 2009. Currently, he works at Qualcomm Inc., San Diego, CA.

His research interests include various topics in digital communications and information theory, with a particular focus on detection and coding for chan-

nels with memory. He is the coauthor of several publications in international journals and conferences.



**Tolga M. Duman** (S'97 - M'98 - SM'03 - F'11) is a Professor of Electrical Engineering with the School of ECEE at Arizona State University. He received the B.S. degree from Bilkent University in Turkey in 1993, M.S. and Ph.D. degrees from Northeastern University, Boston, in 1995 and 1998, respectively, all in electrical engineering. Since August 1998, he has been with the Electrical Engineering Department of Arizona State University first as an Assistant Professor (1998–2004), then as an Associate Professor (2004–2008), and starting August 2008 as a Profes-

sor. Dr. Duman's current research interests are in systems, with particular focus on communication and signal processing, including wireless and mobile communications, coding/modulation, coding for wireless communications, data storage systems and underwater acoustic communications.

Dr. Duman has served as an editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS (2003–08) and IEEE COMMUNICATIONS SURVEYS AND TUTORIALS (2002–07). He is currently an editor for IEEE TRANSACTIONS ON COMMUNICATIONS in the area of coding and communication theory (2007–present), and for *Elsevier Physical Communications Journal* (2010–present). Dr. Duman is a Fellow of IEEE, a recipient of the National Science Foundation CAREER Award and IEEE Third Millennium medal.



**John Proakis** (S'58-M'62-F'84-LF'99) received the BSEE from the University of Cincinnati in 1959, the MSEE from MIT in 1961 and the Ph.D. from Harvard University in 1967. He is an Adjunct Professor at the University of California at San Diego and a Professor Emeritus at Northeastern University. He was a faculty member at Northeastern University from 1969 through 1998 and held the following academic positions: Associate Professor of Electrical Engineering, 1969–1976; Professor of Electrical Engineering, 1976–1998; Associate Dean

of the College of Engineering and Director of the Graduate School of Engineering, 1982–1984; Interim Dean of the College of Engineering, 1992–1993; Chairman of the Department of Electrical and Computer Engineering, 1984–1997. Prior to joining Northeastern University, he worked at GTE Laboratories and the MIT Lincoln Laboratory.

His professional experience and interests are in the general areas of digital communications and digital signal processing. He is the co-author of the books *Digital Communications* (New York: McGraw-Hill, 2008, fifth edition), *Introduction to Digital Signal Processing* (Upper Saddle River: Prentice Hall, 2007, fourth edition); *Digital Signal Processing Laboratory* (Englewood Cliffs: Prentice Hall, 1991); *Advanced Digital Signal Processing* (New York: Macmillan, 1992); *Algorithms for Statistical Signal Processing* (Upper Saddle River: Prentice Hall, 2002); *Discrete-Time Processing of Speech Signals* (New York: Macmillan, 1992, IEEE Press, 2000); *Communication Systems Engineering*, (Upper Saddle River: Prentice Hall, 2002, second edition); *Digital Signal Processing Using MATLAB V.4* (Boston: Brooks/Cole-Thomson Learning, 2007, second edition); *Contemporary Communication Systems Using MATLAB* (Boston: Brooks/Cole-Thomson Learning, 2004, second edition); *Fundamentals of Communication Systems* (Upper Saddle River: Prentice Hall, 2005)



**Milica Stojanovic** (SM'08, F'10) graduated from the University of Belgrade, Serbia, in 1988, and received the M.S. and Ph.D. degrees in electrical engineering from Northeastern University, Boston, MA, in 1991 and 1993. After a number of years with the Massachusetts Institute of Technology, where she was a Principal Scientist, she joined the faculty of Electrical and Computer Engineering Department at Northeastern University in 2008. She is also a Guest Investigator at the Woods Hole Oceanographic Institution, and a Visiting Scientist at MIT. Her

research interests include digital communications theory, statistical signal processing and wireless networks, and their applications to underwater acoustic communication systems. Milica is an Associate Editor for the IEEE JOURNAL OF OCEANIC ENGINEERING and the IEEE TRANSACTIONS ON SIGNAL PROCESSING.