CHANNEL-MATCHED TRELLIS CODES FOR FINITE-STATE INTERSYMBOL-INTERFERENCE CHANNELS

Thomas Lang, Amine Mezghani and Josef A. Nossek

Institute for Circuit Theory and Signal Processing, Technische Universität München
Theresienstraße 90, 80290 Munich, Germany
Email: {Lang, Mezghani, Nossek}@nws.ei.tum.de

ABSTRACT

This paper addresses the optimization of constrained stationary Markov input processes which achieve high information rates on intersymbol-interference (ISI) channels. The considered Markov processes define an optimized subset of equiprobable input symbols for each channel state and are thus uniquely described by the set of branches in a trellis section. We propose an iterative algorithm that efficiently solves the optimization problem. The algorithm successively removes the worst branches and thereby constructs a reduced trellis code that is matched to the channel. The second contribution of this paper are tight lower bounds for the mutual information rate (MIR) of Markov sources on ISI channels with finite input and output alphabets. The bounds can be evaluated within a small trellis window without using Monte Carlo methods.

1. INTRODUCTION

Recently, several methods for the construction of capacity-achieving codes for channels with memory were proposed [1, 2, 3]. In [1], Kavčić *et al.* show that the capacity of channels with ISI and finite input alphabets is greater than the i.u.d. capacity, which is the MIR induced by independently and uniformly distributed input. Moreover, they demonstrate that rates above the i.u.d. capacity cannot be achieved by random linear codes. To overcome this limitation, they introduce matched information rate codes. Their idea is to combine an inner trellis code which mimics an information rate maximizing input Markov process with interleaved outer low-density parity-check (LDPC) codes.

Channel-matched coding requires that the transmitter and receiver exchange information about the channel, codebook etc. Also, many communication applications do not tolerate long delays. Hence, the code length is bounded in practice.

In this paper, we present a novel method for constructing channel-matched inner trellis codes that when combined with outer (random linear) codes enable reliable communication at information rates above the i.u.d. capacity of finite-state (FS) ISI channels. The method is a generalization of the approach in [4, 5] to ISI channels. The key idea is to consider the input Markov process as a means to virtually transform asymmetric channels into almost symmetric ones with lower conditional entropy. Given the reduction in conditional entropy, comparatively short outer codes are sufficient to communicate reliable at high information rates and at a relatively low delay. In addition, we present an efficient method for the computation of tight lower bounds for the MIR of Markov sources on ISI channels with finite input and output alphabets. Finally, we provide simulation results for Gaussian MIMO channels with ISI, quaternary-phase shift-keyed (QPSK) input symbols and single-bit outputs.

Please note, that the use of single-bit digital-to-analog and analog-to-digital converters allows for a low complexity and power efficient transceiver implementation. This is particularly beneficial in high-speed MIMO communication systems, e.g. ultra-wideband communications. Also, the power penalty with respect to the mutual information due to 1-bit quantization is approximately equal to $\pi/2$ at low SNR [6].

Notation: *, $^{\mathrm{T}}$ and $^{\mathrm{H}}$ denote the conjugate, transpose and Hermitian transpose operators. Scalars, vectors and matrices are written in lower case, lower case bold and upper case bold italic letters, respectively. A random variable X and its realization x are written in upper and lower case italic letters, respectively. Random variables in a random sequence are marked with a time index t, e.g., X_t. A random sequence $[X_t, X_{t+1}, \ldots, X_n]$ is shortly denoted by X_t^n. $|\mathcal{X}|$ is the size of the alphabet $\mathcal{X}$. We use the short notations $P_t(i,j|Y_1^n) = \Pr(S_{t-1} = i, S_t = j|Y_1^n)$, $P_t(i|Y_1^n) = \Pr(S_{t-1} = i|Y_1^n)$, $P_t(Y_1^n|i,j) = \Pr(Y_1^n|S_{t-1} = i, S_t = j)$ and $P_t(Y_1^n|i) = \Pr(Y_1^n|S_{t-1} = i)$.

2. SOURCE AND CHANNEL MODEL

We consider a stationary Markov source of order M, which generates a random process X_t, whose realizations x_t take values from a finite-size source alphabet $\mathcal{X}$. The corresponding Markov source distribution satisfies

$$\Pr(X_t|X_{1-M}^{t-1}) = \Pr(X_t|X_{t-M}^{t-1}),\ t = 1, 2, \ldots . \quad (1)$$

Besides, we consider a time-invariant ISI channel with input alphabet $\mathcal{X}$, output alphabet $\mathcal{Y}$ and memory length $L \leq M$.

The joint source/channel state is given by

$$S_t = X_{t-M+1}^t = [X_{t-M+1}, X_{t-M+2}, \dots, X_t], \quad (2)$$

and the corresponding realization s_t is an element of the state index set $\mathcal{S} = \{1, 2, \dots, |\mathcal{X}|^M\}$. From (1) we have

$$\Pr(S_t|S_0^{t-1}) = \Pr(S_t|S_{t-1}),\ t = 1, 2, \dots\ . \quad (3)$$

We assume that the state sequence S_t forms an irreducible and aperiodic Markov chain. Given an initial state $s_0 = [x_{-M+1}, x_{-M+2}, \dots, x_0]$ the input sequence x_1^t and state sequence s_1^t determine each other uniquely. The first order Markov process S_t can also be represented by the following set of transition probabilities

$$P_{ij} = \Pr(S_t = j|S_{t-1} = i),\ t > 0,\ (i, j) \in \mathcal{T}, \quad (4)$$

which fulfills $\sum_{j:(i,j)\in\mathcal{T}} P_{ij} = 1,\ \forall i \in \mathcal{S}$. Here, $\mathcal{T}$ is the set of all possible state pairs $(S_{t-1} = i, S_t = j)$. From (2) it is clear, that not all state transitions from state i to state j are possible. Each state $i \in \mathcal{S}$ exhibits a steady-state probability

$$\mu_i = \lim_{l\to\infty} \Pr(S_{t+l} = i|S_t = j),\ t > 0,\ \forall j \in \mathcal{S}. \quad (5)$$

The steady-state probability of every state $i \in \mathcal{S}$ satisfies

$$\mu_i = \sum\nolimits_{j:(i,j)\in\mathcal{T}} \mu_j P_{ji} \geq 0, \quad (6)$$

and the set of steady-state probabilities fulfills $\sum_{i\in\mathcal{S}} \mu_i = 1$. The state sequence S_t induces a hidden Markov sequence Y_t with realizations $y_t \in \mathcal{Y}$. The channel output satisfies

$$\Pr(Y_t|S_0^\infty, Y_1^{t-1}, Y_{t+1}^\infty) = \Pr(Y_t|S_{t-1}^t),\ t > 0. \quad (7)$$

The MIR of the considered FS ISI channel is defined as [7]

$$\mathcal{I}(X_t; Y_t) = \mathcal{I}(S_t; Y_t) = \lim_{n\to\infty} \frac{1}{n} I(S_1^n; Y_1^n|S_0), \quad (8)$$

where $I(S_1^n; Y_1^n|S_0)$ denotes the mutual information between S_1^n and Y_1^n given the initial state S_0. The maximal information rate for a Markov source of order M is defined as

$$C_M = \max_{P_{ij}:\ (i,j)\in\mathcal{T}} \mathcal{I}(S_t; Y_t), \quad (9)$$

where $S_t = X_{t-M+1}^t$. The i.u.d. capacity is given by

$$C_{\text{i.u.d.}} = \mathcal{I}(S_t; Y_t) \quad \text{s.t.}\quad P_{ij} = |\mathcal{X}|^{-1},\ \forall(i, j) \in \mathcal{T}. \quad (10)$$

3. A NEW EXPRESSION FOR THE MUTUAL INFORMATION RATE

In this section, we derive a new expression for the MIR, which will be used in the following section to construct lower bounds for the MIR. The MIR in (8) can be rewritten as [7]

$$\begin{aligned}\mathcal{I}(S_t; Y_t) &= \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^n I(S_t; Y_1^n|S_{t-1}) \\ &= \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^n H(S_t|S_{t-1}) - H(S_t|Y_1^n, S_{t-1}). \quad (11)\end{aligned}$$

Due to the stationarity the first term in (11) simplifies to

$$H(S_t|S_{t-1}) = -\sum\nolimits_{(i,j)\in\mathcal{T}} \mu_i P_{ij} \log_2(P_{ij}). \quad (12)$$

The second term in (11) can be expressed as [7]

$$\begin{aligned}-H(S_t|S_{t-1}, Y_1^n) &= \mathrm{E}_{Y_1^n, S_{t-1}^t}[\log_2(\Pr(S_t|S_{t-1}, Y_1^n))] \\ &= \sum_{i,j:(i,j)\in\mathcal{T}} \mu_i P_{ij}\, \mathrm{E}_{Y_1^n|S_{t-1}=i,S_t=j}[\log_2(P_t(i, j|Y_1^n))] \\ &\quad - \sum_{i\in\mathcal{S}} \mu_i\, \mathrm{E}_{Y_1^n|S_{t-1}=i}[\log_2(P_t(i|Y_1^n))]. \quad (13)\end{aligned}$$

The last line in (13) can be expanded as follows

$$\begin{aligned}&\sum_{i\in S} \mu_i\, \mathrm{E}_{Y_1^n|S_{t-1}=i}[\log_2(P_t(i|Y_1^n))] \\ &= \sum_{i\in S} \mu_i \sum_{Y_1^n\in\mathcal{Y}^n} \Pr(Y_1^n|S_{t-1} = i) \cdot [\log_2(P_t(i|Y_1^n))] \\ &= \sum_{i\in S} \mu_i \sum_{Y_1^n\in\mathcal{Y}^n} \sum_{j:(i,j)\in\mathcal{T}} \Pr(Y_1^n, S_t = j|S_{t-1} = i) \cdot [\dots] \\ &= \sum_{i,j:(i,j)\in\mathcal{T}} \mu_i P_{ij}\, \mathrm{E}_{Y_1^n|S_{t-1}=i,S_t=j}[\log_2(P_t(i|Y_1^n))]. \quad (14)\end{aligned}$$

Inserting (14) into (13) yields

$$-H(S_t|S_{t-1}, Y_1^n) = \sum_{i,j:(i,j)\in\mathcal{T}} \mu_i P_{ij}\, \mathrm{E}_{Y_1^n|S_{t-1}=i,S_t=j} \left[\log_2\left(\frac{P_t(i, j|Y_1^n)}{P_t(i|Y_1^n)}\right)\right] \quad (15)$$

Hence, the MIR in (11) is also given by

$$\mathcal{I}(S_t; Y_t) = \sum\nolimits_{i,j:(i,j)\in\mathcal{T}} \mu_i P_{ij} \cdot (U_{ij} - \log_2(P_{ij}))\,, \quad (16)$$

where the expectation U_{ij} is defined as

$$U_{ij} = \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^n \mathrm{E}_{Y_1^n|S_{t-1}=i,S_t=j} \left[\log_2\left(\frac{P_t(i, j|Y_1^n)}{P_t(i|Y_1^n)}\right)\right]. \quad (17)$$

Combining the relations

$$\frac{P_t(i, j|Y_1^n)}{P_t(i|Y_1^n)} = \frac{P_t(Y_1^n|i, j) \cdot P_{ij}}{\sum_{u:(i,u)\in\mathcal{T}} P_t(Y_1^n|i, u) \cdot P_{iu}}, \quad (18)$$

$$P_t(Y_1^n|i, j) = P(Y_t^n|i, j) \cdot P_t(Y_1^{t-1}|i), \quad (19)$$

as well as $\sum_{Y_1^{t-1}\in\mathcal{Y}^{t-1}} \Pr(Y_1^{t-1}|S_{t-1} = i) = 1$, the expression for U_{ij} in (17) simplifies to

$$\begin{aligned}U_{ij} &= \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^n \sum_{Y_t^n\in\mathcal{Y}^{n-t+1}} P_t(Y_t^n|i, j) \\ &\quad \cdot \log_2\left(\frac{P_t(Y_t^n|i, j) \cdot P_{ij}}{\sum_{u:(i,u)\in\mathcal{T}} P_t(Y_t^n|i, u) \cdot P_{iu}}\right). \quad (20)\end{aligned}$$

4. TIGHT LOWER BOUNDS FOR THE MUTUAL INFORMATION RATE

In this section, we derive tight lower bounds for the MIR of ISI channels with finite input and finite output alphabets.
As the frame length n goes to infinity and as the Markov chain S_t is irreducible and aperiodic, the MIR does not depend on the initial or terminal state. Moreover, as conditioning reduces uncertainty and as the considered Markov processes are stationary, (11) can be lower bounded with

$$\begin{aligned} \mathcal{I}(S_t;Y_t) &\geq \lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} I(S_t;Y_t^{t+v}|S_{t-1}) \\ &= I(S_t;Y_t^{t+v}|S_{t-1}), \end{aligned} \tag{21}$$

assuming $v < n$. Replacing Y_t^n with Y_t^{t+v} in the derivation of U_{ij} in Section 3 gives rise to the lower bound

$$\begin{aligned} \mathcal{I}_{\hat{U}_{ij}(v)}(S_t;Y_t) &= \sum_{i,j:(i,j)\in\mathcal{T}} \mu_i P_{ij}\cdot\left(\hat{U}_{ij}(v) - \log_2(P_{ij})\right) \\ &= I(S_t;Y_t^{t+v}|S_{t-1}), \end{aligned} \tag{22}$$

where the truncated expectation $\hat{U}_{ij}(v)$ is defined by

$$\begin{aligned} \hat{U}_{ij}(v) &= \sum\nolimits_{Y_t^{t+v}\in\mathcal{Y}^{v+1}} P_t(Y_t^{t+v}|i,j) \\ &\quad\cdot \log_2\left(\frac{P_t(Y_t^{t+v}|i,j)\cdot P_{ij}}{\sum_{u:(i,u)\in\mathcal{T}} P_t(Y_t^{t+v}|i,u)\cdot P_{iu}}\right). \end{aligned} \tag{23}$$

Clearly, for $v_1 > v_2$ we have

$$I(S_t;Y_t^{t+v_1}|S_{t-1}) \geq I(S_t;Y_t^{t+v_2}|S_{t-1}). \tag{24}$$

Thus, by increasing the truncation parameter v, the MIR can be tightly lower bounded. The computation of $\mathcal{I}_{\hat{U}_{ij}(v)}(S_t;Y_t)$ is straightforward, because (23) is exclusively composed of channel and state transition probabilities, that is

$$P_t(Y_t^{t+v}|i,j) = P_{t+1}(Y_{t+1}^{t+v}|j)\cdot P_t(Y_t|i,j), \tag{25}$$

and $P_{t+1}(Y_{t+1}^{t+v}|j)$ can be computed recursively

$$P_{t+1}(Y_{t+1}^{t+v}|j) = \sum_{l:(j,l)\in\mathcal{T}} P_{t+1}(Y_{t+1}|j,l)\cdot P_{jl}\cdot P_{t+2}(Y_{t+2}^{t+v}|l). \tag{26}$$

Clearly, the complexity of the expression (23) grows exponentially with v. However, the simulation results show that choosing small $v > L$ results in tight lower bounds.

The MIR can also be estimated numerically using [7]

$$\mathcal{I}(S_t;Y_t) = \sum\nolimits_{i,j:(i,j)\in\mathcal{T}} \mu_i P_{ij}\cdot(T_{ij} - \log_2(P_{ij})), \tag{27}$$

where

$$\hat{T}_{ij}(n) = \frac{1}{n}\sum_{t=1}^{n}\left[\log_2\left(\frac{P_t(i,j|Y_1^n)^{\frac{P_t(i,j|Y_1^n)}{\mu_i P_{ij}}}}{P_t(i|Y_1^n)^{\frac{P_t(i|Y_1^n)}{\mu_i}}}\right)\right], \tag{28}$$

and $\lim_{n\to\infty}\hat{T}_{ij}(n) = T_{ij}$. It is worth mentioning, that the forward recursion of the Baum-Welch/BCJR algorithm is sufficient to estimate the MIR [8].

5. CHANNEL-MATCHED REDUCED TRELLIS CODES

In this section, we propose a new method for constructing channel-matched trellis codes. The method is a generalization of the approach in [4, 5] to ISI channels.
The code construction is based on the following optimization

$$\begin{aligned} C_M^{\text{uniform}} &= \max_{P_{ij}:\,(i,j)\in\mathcal{T}} \mathcal{I}(S_t;Y_t) \\ \text{s.t. } & P_{ij}\in\{0,K^{-1}\},\ \sum\nolimits_{j:(i,j)\in\mathcal{T}} P_{ij} = 1. \end{aligned} \tag{29}$$

Here, C_M^{uniform} is the maximal information rate, which can be achieved with a Markov source of order M that is constrained to a subset of K equiprobable input symbols x_t for each channel state $s_{t-1}\in\mathcal{S}$. Such a Markov input process is completely characterized by a reduced set of branches in a trellis section. Hence, the name *reduced trellis code* (RTC). The RTC induces a MIR of

$$\mathcal{I}(S_t;Y_t) = \sum\nolimits_{i,j:(i,j)\in\mathcal{T}_K} \mu_i P_{ij}\cdot(T_{ij} - \log_2(P_{ij})), \tag{30}$$

where $\mathcal{T}_K$ denotes the set of selected state pairs

$$\mathcal{T}_K = \{(i,j)|(i,j)\in\mathcal{T}, P_{ij} = K^{-1}\}. \tag{31}$$

The *iterative information rate maximization* (IR-max) algorithm in Algorithm 1 efficiently computes near-optimal solutions of (29). The IR-max algorithm uses either $\hat{T}_{ij}(n)$ or $\hat{U}_{ij}(v)$ as a measure for the quality Q_{ij} of trellis branches $(i,j)\in\mathcal{T}$. In every iteration of the k-loop, the algorithm determines the worst branch with the lowest Q_{ij} for each state $i\in\mathcal{S}$ and removes it, that is $P_{id} = 0$. After K iterations the algorithm returns the optimized transition probabilities P_{ij}, i.e., a channel-matched RTC.

Algorithm 1 Iterative Information Rate Maximization

1: **initialization:** $P_{ij} = \begin{cases} 1/|\mathcal{X}| & \text{if } (i,j)\in\mathcal{T} \\ 0 & \text{otherwise} \end{cases}$
   Choose K and decide whether to use $Q_{ij} = \hat{T}_{ij}(n)$ or $Q_{ij} = \hat{U}_{ij}(v)$ for all $(i,j)\in\mathcal{T}$
2: **for** $k = |\mathcal{X}|$ **down to** K **do**
3:   **estimate all** Q_{ij}
4:   **for all** $i\in\mathcal{S}$ **do**
5:     $d = \arg\min_{j:P_{ij}>0}\{Q_{ij}\}$
6:     $P_{id} = 0$
7:   **for all** $(i,j)\in\mathcal{T}$ **do**
8:     $P_{ij} = P_{ij}\cdot k/(k-1)$

6. GAUSSIAN MIMO CHANNELS WITH ISI AND SINGLE-BIT OUTPUTS

Consider the MIMO channel with ISI and single-bit output quantization shown in Fig. 1. The channel has a memory of

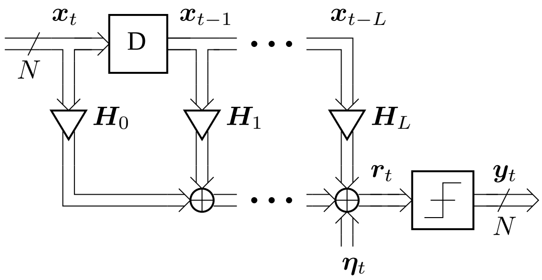

Fig. 1. MIMO channel with ISI and single-bit outputs.

length L and it is governed by the channel law

$$\boldsymbol{y}_t = \mathcal{Q}\{\boldsymbol{r}_t\} = \mathcal{Q}\left\{\sum_{k=0}^{L} \boldsymbol{H}_k \boldsymbol{x}_{t-k} + \boldsymbol{\eta}_t\right\}. \tag{32}$$

Here, $\boldsymbol{H}_k \in \mathbb{C}^{N\times N}$ is the k-th channel matrix. $\boldsymbol{x}_k \in \mathcal{X}$, $\boldsymbol{\eta}_k \in \mathbb{C}^N$, $\boldsymbol{r}_k \in \mathbb{C}^N$ and $\boldsymbol{y}_k \in \mathcal{Y} = \{\alpha + \mathrm{j}\beta | \alpha, \beta \in \{+1, -1\}\}^N$ denote the channel input vector, the noise vector, the unquantized receive vector and the channel output vector, at the k-th time instant, respectively. The single-bit quantization operator $\mathcal{Q}$ returns the sign of the real and imaginary part of each component of the unquantized received signal $\boldsymbol{r}_t$, i.e.,

$$\mathcal{Q}\{\boldsymbol{r}_t\} = \mathrm{sign}(\mathrm{real}\{\boldsymbol{r}_t\}) + \mathrm{j}\cdot \mathrm{sign}(\mathrm{imag}\{\boldsymbol{r}_t\}). \tag{33}$$

The conditional probability of the channel output satisfies

$$\Pr(\boldsymbol{y}_t | \boldsymbol{x}_{-L}^{\infty}, \boldsymbol{y}_1^{t-1}, \boldsymbol{y}_{t+1}^{\infty}) = \Pr(\boldsymbol{y}_t | \boldsymbol{x}_{t-L}^{t}),\ t \geq 0. \tag{34}$$

Here, $\boldsymbol{x}_0^{\infty}$ and $\boldsymbol{y}_1^{t-1}$ stand for the sequences $[\boldsymbol{x}_{-L}, x_{-L+1}, \ldots]$ and $[\boldsymbol{y}_1, \boldsymbol{y}_2 \ldots, \boldsymbol{y}_{t-1}]$, respectively. The noise is additive white Gaussian with covariance matrix $\mathrm{E}[\boldsymbol{n}_t \boldsymbol{n}_t^{\mathrm{H}}] = \sigma_\eta^2 \boldsymbol{I}_N$. The transmit signal energy is normalized to 1, that is $\|\boldsymbol{x}_t\|_2 = 1$. The (i,j)-th real and imaginary component of the t-th channel matrix $[\boldsymbol{H}_t]_{i,j}$ is zero-mean Gaussian with variance $\sigma_t^2/2$. On average the channel taps are normalized pursuant to

$$\sum_{t=0}^{L} \sigma_t^2 = 1,\ \forall (i,j). \tag{35}$$

The signal-to-noise ratio is given by

$$\mathrm{SNR} = 1/\sigma_\eta^2. \tag{36}$$

The channel transition probabilities can be calculated via

$$\Pr(\boldsymbol{y}_t | \boldsymbol{x}_{t-L}^t) = \prod_{c\in\{\mathbb{R},\mathbb{I}\}} \prod_{i=1}^{N} \Phi\left(\frac{[\boldsymbol{y}_t]_{c,i}\cdot[\bar{\boldsymbol{r}}]_{c,i}}{\sigma_\eta/\sqrt{2}}\right). \tag{37}$$

Here, $[\boldsymbol{x}_t]_{\mathbb{R},i}$ ($[\boldsymbol{x}_t]_{\mathbb{I},i}$) denotes the i-th real (imaginary) component of the input vector $\boldsymbol{x}_t$, $\bar{\boldsymbol{r}} = \sum_{k=0}^{L} \boldsymbol{H}_k \boldsymbol{x}_{t-k}$ is the noise-free unquantized receive vector and $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$ is the cumulative normal distribution function. The input symbols are modulated using QPSK. Consequently, we consider FS ISI channels with $|\mathcal{X}| = |\mathcal{Y}| = 4^N$ and $|\mathcal{T}| = (4^N)^{M+1}$.

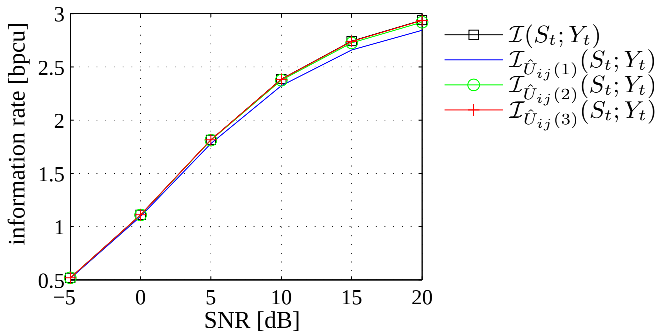

Fig. 2. Tightness of the lower bounds in (22) for $v \in \{1, 2, 3\}$.

7. SIMULATION RESULTS

We consider the FS ISI channels of Section 6 with $N = 2$ transmit and receives antennas and channel memory length $L = 1$. On average, two thirds of the transmit energy is conveyed over the undelayed channel tap $\boldsymbol{H}_0$, that is $\sigma_0^2 = 2/3$ and $\sigma_1^2 = 1/3$. We choose a Markov source of order $M = L$, because C_L is noticeably larger compared to the i.u.d. capacity, whereas memory lengths above L yield only a small additional gain in MIR. Thus, we consider FS ISI channels with $|\mathcal{X}| = |\mathcal{Y}| = 16$ and $|\mathcal{T}| = 256$.

Fig. 2 demonstrates that the derived lower bounds for the MIR in (22) are very tight even for small $v > L$. The results are averaged of 50 channels. The MIR was evaluated using randomly generated (and properly normalized) state transition probabilities P_{ij}. Please note, that the evaluation of $\hat{U}_{ij}(v = 2)$ is computational inexpensive compared to a (precise) Monte Carlo simulation of $\hat{T}_{ij}(n)$.

Fig. 3 shows how the information rate and conditional entropy evolve during the first 12 iterations of the IR-max algorithm and the generalized Blahut-Arimoto (GBA) algorithm [9] at an SNR of 10dB. The GBA is initialized with i.u.d. state transition probabilities P_{ij}. The results are averaged over 10 channels. For $K \leq 7$ iterations, the IR-max algorithm decreases the conditional entropy while increasing the information rate. This result reflects our code design methodology. In contrast to an unconstrained information rate maximization, the enforced reduction of entropy in (29) is compensated as much as possible by decreasing the conditional entropy.

Finally, Fig. 4 investigates the performance of the IR-max algorithm over the whole SNR range. The results are averaged over 10 channels. In order to avoid the use of distribution shapers, the target rate is fixed at $\log_2(K) =$ 3bits per channel use (bpcu). Hence, the code bits of outer standard codes can be mapped directly to the input symbols of the RTC and vice versa. The plots indicate, that RTCs achieve information rates well above $C_{\text{i.u.d.}}$ as long as $C_M \leq \log_2(K) - 1$. The complexity of the IR-max algorithm can be adjusted by choosing smaller n (or rather smaller v). Fortunately, the IR-max algorithm exhibits a good performance with as few

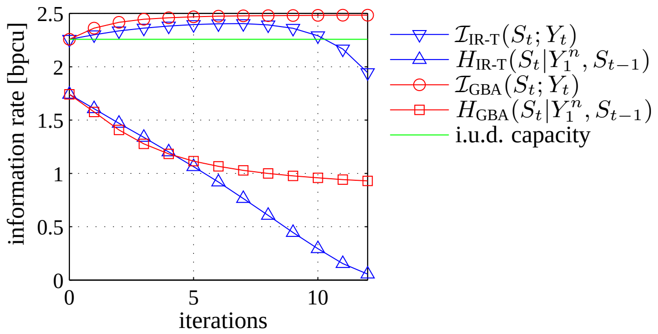

Fig. 3. Comparison of the GBA algorithm (labeled 'GBA') and IR-max algorithm (labeled 'IR-T') over the iterations at an SNR of 10dB using $\hat{T}_{ij}(n = 800 \cdot |\mathcal{T}|)$.

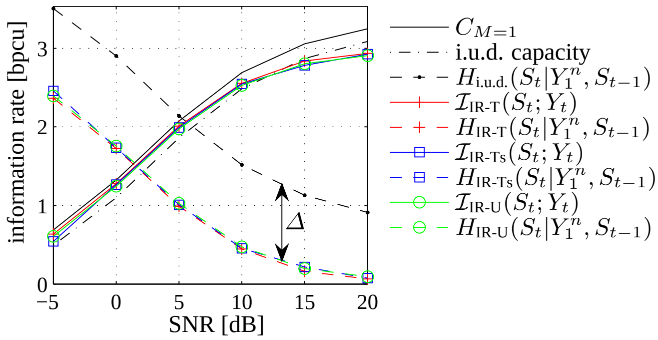

Fig. 4. Performance of the IR-max algorithm using $\hat{T}_{ij}(n = 800 \cdot |\mathcal{T}|)$ (labeled 'IR-T'), $\hat{T}_{ij}(n = 20 \cdot |\mathcal{T}|)$ (labeled 'IR-Ts') and $\hat{U}_{ij}(v = 2)$ (labeled 'IR-U') for $K = 8$.

as 20 estimates (on average) per state transition $(i, j) \in \mathcal{T}$. This property might be crucial, if the channel is time-varying. Moreover, RTCs are delay-free and completely describable with $|\mathcal{T}|$ feedback bits. Though, most importantly, RTCs provide a reduction of conditional entropy of approximately $\Delta \approx$ 1bpcu for SNR $<$ 13dB compared to i.u.d. input. Thus, if the SNR is below 13dB, the reduction from $H_{\text{i.u.d.}}(S_t|Y_1^n, S_{t-1})$ down to $H_{\text{IR-T}}(S_t|Y_1^n, S_{t-1})$ would render it possible to communicate reliable at information rates close to $C_{\text{i.u.d.}}$ using considerably shorter, less complex outer standard codes. In addition, the computation of APPs for the outer code is less complex on a reduced trellis. Unfortunately, a brute-force computation of C_M^{uniform}, which is the maximum of (29), is not feasible.

8. CONCLUSION

In this paper, we presented a novel method for constructing channel-matched trellis codes, which induce information rates that surpass the i.u.d. capacity of FS ISI channels.
The proposed trellis codes can be easily combined with any standard code on any FS ISI channel. The basic idea is to consider the input Markov process as a means to virtually transform asymmetric channels into almost symmetric ones with lower conditional entropy.
In addition, we derived tight lower bounds for the mutual information rate of FS ISI channels with finite input and output alphabets. The bounds can be evaluated within a small trellis window without using Monte Carlo simulation techniques. In the future, we will apply the reduced trellis codes to other communication channels.

9. REFERENCES

[1] A. Kavčić, X. Ma, and N. Varnica, "Matched information rate codes for partial response channels," *IEEE Trans. on Information Theory*, vol. 51, pp. 973–989, Mar. 2005.

[2] J. B. Soriaga and P. H. Siegel, "On near-capacity coding systems for partial-response channels," in *Proc. of IEEE International Symposium on Information Theory*, June 2004.

[3] J. H. Bae and A. Anastasopoulos, "Capacity-achieving codes for finite-state channels with maximum-likelihood decoding," *IEEE Journal on Selected Areas in Communications*, vol. 27, pp. 974–984, Aug. 2009.

[4] A. Mezghani, M. T. Ivrlac, and J. A. Nossek, "Achieving near-capacity on large discrete memoryless channels with uniform distributed selected input," *Proc. of International Symposium on Information Theory and its Applications*, Dec. 2008.

[5] T. Lang, A. Mezghani, and J. A. Nossek, "Channel-adaptive coding for coarsely quantized MIMO systems," in *Proc. of International ITG Workshop on Smart Antennas*, Feb. 23–24, 2010.

[6] A. Mezghani and J. A. Nossek, "On ultra-wideband MIMO systems with 1-bit quantized outputs: Performance analysis and input optimization," in *Proc. of IEEE International Symposium on Information Theory*, Nice, June 24–29, 2007, pp. 1286–1289.

[7] A. Kavčić, "On the capacity of markov sources over noisy channels," in *Proc. of IEEE Global Communications Conference*, San Antonio, TX, USA, 2001, vol. 5, pp. 2997–3001.

[8] D. M. Arnold and H.-A. Loeliger, "On the information rate of binary-input channels with memory," in *Proc. of IEEE International Conference on Communications*, Helsinki, Finland, 2001, vol. 9, pp. 2692–2695.

[9] P. O. Vontobel, A. Kavčić, D. M. Arnold, and H.-A. Loeliger, "A generalization of the blahut-arimoto algorithm to finite-state channels," *IEEE Trans. on Information Theory*, vol. 54, pp. 1887–1918, May 2008.