Density Evolution Analysis of Braided Convolutional Codes on the Erasure Channel

Moloudi, Saeedeh; Lentmaier, Michael

Published in:
[Host publication title missing]

2014

Density Evolution Analysis of Braided Convolutional Codes on the Erasure Channel

Saeedeh Moloudi and Michael Lentmaier
Dept. of Electrical and Information Technology, Lund University, Sweden
Emails: {saeedeh.moloudi,michael.lentmaier}@eit.lth.se

***Abstract*—Braided convolutional codes (BCCs) are a class of spatially coupled turbo-like codes with a structure that is similar to product codes or generalized LDPC codes. We derive explicit input/output transfer functions of the component convolutional decoders for the binary erasure channel (BEC). These are then used to formulate exact density evolution equations for blockwise BCCs under belief propagation (BP) decoding with optimal component APP decoders. Thresholds are computed for the coupled and uncoupled case, which is equivalent to tailbiting. Due to the relatively high rate of the component codes a significant threshold improvement by spatial coupling can be observed.**

I. Introduction

Braided block codes (BBCs) [1] are a class of generalized low-density parity-check (LDPC) convolutional codes that can be viewed as a spatially coupled version of Elias' product codes [2]. Similar to LDPC codes, sparsity can be introduced into their structure, without changing the component codes, in order to construct codes of arbitrary length or memory. BBCs with BCH component codes were recently considered for high-speed optical communications in [3], where product-like codes are commonly applied [4]. Like the closely related staircase codes [5], they show an excellent performance together with low-complexity iterative hard decision decoding.

In this paper we consider a counterpart of BBCs called braided convolutional codes (BCCs) [6]. Again the encoding can be described by a two-dimensional sliding array in which each symbol is protected by a horizontal and a vertical component code. But now the component codes are convolutional codes, resulting in a class of spatially coupled turbo-like codes with a structure similar to generalized LDPC codes. Unlike parallel or serially concatenated convolutional codes all information and parity symbols are protected by both component codes in a symmetric fashion.

For a random ensemble of BCCs with Markov permutors it was shown in [6] that the minimum distance of typical codes grows linearly with their constraint length, i.e., BCCs are asymptotically good. Although a formal proof is still open, it is expected that this is also the case for the slightly different blockwise BCC construction that we consider here. Indeed, the simulation results for blockwise BCCs in [6, Fig. 12] indicate superior distance properties compared to parallel concatenated codes since no error floor is visible for comparable permutor sizes. At the same time, unlike serial concatenated codes, the BCCs can compete with the parallel concatenation in the waterfall region. Interestingly, the simulated codes in [6, Fig. 12] performed significantly better than the tailbiting case in [6, Fig. 13]. For the latter case the AWGN channel threshold was estimated by Monte Carlo techniques. It was conjectured that this performance improvement can be prescribed to a similar effect as the threshold saturation phenomenon known for coupled LDPC codes [7].

The aim of this paper is to confirm this conjecture by performing a threshold analysis of blockwise BCCs. After introducing BCCs and their decoding in Section II and Section III, we derive explicit input/output transfer functions that characterize the *a posteriori* probability (APP) decoders of their component codes in Section IV. Considering the binary erasure channel (BEC), these transfer functions can be computed analytically by means of a Markov chain analysis of the decoder metrics, as presented in [8] for rate $R = 1/2$ encoders. We apply the technique from [8] to rate $R = 2/3$ encoders with different input and output probabilities for each symbol type, resulting in a three-dimensional transfer function for each of the output symbols. These transfer functions are then used in Section V to formulate exact density evolution recursions for the blockwise BCCs and compute belief propagation (BP) thresholds for the coupled and tailbiting (or uncoupled) case. Due to the higher rate of the component codes the tailbiting/uncoupled threshold is worse than the thresholds of typical parallel concatenated codes. However, as expected, the coupled ensemble has a significantly better threshold.

II. Braided Convolutional Codes

Similar to turbo codes, BCCs have convolutional codes as component codes but the most important difference between turbo codes and braided codes is that, in BCCs, the parity symbols of one component encoder are used as future inputs of the other component encoder. Throughout this paper we limit ourselves to the example of rate $R = 1/3$ blockwise BCCs as illustrated in Fig.1. They consist of two systematic convolutional component encoders of rate $R = 2/3$. At time t, a block of N information symbols $\mathbf{u}_t$ and a block of N parity symbols $\mathbf{v}_{t-1}^{(2)}$ (there is a delay D^N of one block) enter Encoder A directly and through the block permutor $P^{(2)}$, respectively. Encoder B has permuted information symbols through block permutor $P^{(0)}$ and permuted parity symbols

This work was supported in part by the Swedish Research Council (VR) under grant 621-2013-5477.

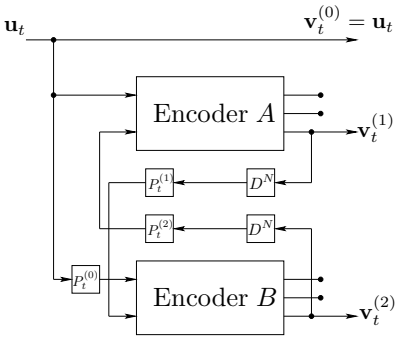

Fig. 1. Encoder of BCC

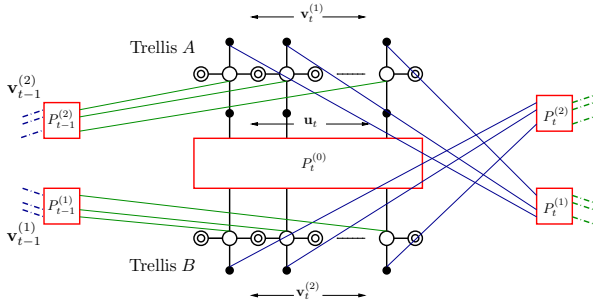

Fig. 3. Factor graph of a blockwise BCC at time t.

$\mathbf{v}_{t-1}^{(1)}$ from Encoder A through block permutor $P^{(1)}$ and delay block, as inputs. The output of the encoder at time t is $\mathbf{v}_t = (\mathbf{v}_t^{(0)}, \mathbf{v}_t^{(1)}, \mathbf{v}_t^{(2)})$, where $\mathbf{v}_t^{(0)} = \mathbf{u}_t$. It follows from the encoding procedure that BCCs are a class of spatially coupled codes because the encoded blocks $\mathbf{v}_t$ depend on blocks from previous time instants.

An uncoupled braided code can be defined by omitting the delay blocks. It is also possible to use block length $N = 1$ and not to use the permutors. In this case the codes are called tightly braided convolutional codes (TBCCs).

BCCs are closely related to classic product codes, in which the data is written in an infinite two-dimentional array and the rows and columns are encoded by separate component codes. Moreover, the horizontal and vertical encoders are linked for BCCs through parity feedback. The array of a TBCC is illustrated in Fig. 2. It consists of three diagonal ribbons and the information symbols are placed in the center ribbon. The parity symbols of the horizontal and vertical encoder are stored in the upper and lower ribbons, respectively. At time t, the output of the horizontal encoder, which is shown by $v_t^{(1)}$, depends on the current information symbol u_t, its left neighbor $v_{t-1}^{(2)}$ and the encoder state. $v_t^{(1)}$ will be placed as the right neighbor of u_t in the array. Shaded squares of the array contain the previous inputs and outputs, which are assumed to be known. The operation of the vertical encoder is analogous

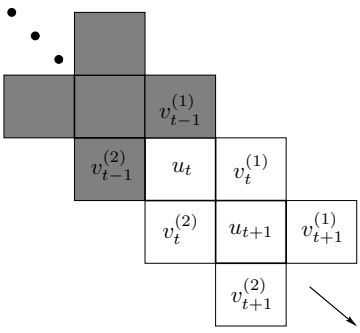

Fig. 2. Array representation of TBC codes

to the horizontal one. Finally, at time t, the coded symbols $\mathbf{v}_t = (u_t, v_t^{(1)}, v_t^{(2)})$ are sent over the channel.

Throughout this paper we consider transmission of a sequence of L coupled blocks $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_L$, and distinguish between encoders with termination or tailbiting. In the first case the encoder is terminated and the blocks at times $t < 1$ and $t > L$ are equal to $\mathbf{v}_t = 0$. In the second case we have a circular structure. Uncoupled BCCs can be defined by using tailbiting with $L = 1$.

III. Iterative Decoding

A factor graph representation of a blockwise BCC is shown in Fig. 3. We consider BP decoding, i.e., an iterative message passing decoder in which the trellises of the component codes are decoded by the BCJR algorithm. In every iteration each decoding block at time t, $t = 1, \ldots, L$ receives log-likelihood ratios (LLRs) from the channel and the decoders at the same time t and the neighboring blocks at time $t-1$ and $t+1$, resulting from previous iteration. Fig. 4 shows the connection of the decoders at different time instants and how they exchange LLRs between time slots for coupled BCCs. Note that we omit the permutations in the block diagram in order to simplify the illustration. LLRs coming from time $t < 1$ and $t > L$ are set to $+\infty$, since the corresponding symbols are equal to zero by definition.

Based on the input values $\mathbf{L}_{\text{in},t}^{(k)}$, $k = 0, 1, 2$, the BCJR decoders create new extrinsic output values $\mathbf{L}_{\text{out},t}^{(k)}$ that are passed back to the other BCJR decoder in the same and neighboring time instants. An illustration of the decoding block at time t is given in Fig. 5. Here $\mathbf{L}_{ch,t}^{(k)}$ denotes the channel LLRs of the kth symbol block $\mathbf{v}_t^{(k)}$ and $\mathbf{L}_{t,A\to B}^{(k)}$ defines the extrinsic outputs $\mathbf{L}_{\text{out},t}^{(k)}$ of Decoder A, which contributes to the input of Decoder B. $\mathbf{L}_{t,B\to A}^{(k)}$ is defined analogously. In the first decoding iteration the values from the previous iteration are initialized as erasures, which means that the respective LLRs are defined as zero.

In the BCC decoding block of Fig. 5, green lines show the LLRs that are exchanged with the past, and LLRs exchanged with the future are illustrated in brown lines. Some LLRs are produced at the current time instant and only used at the

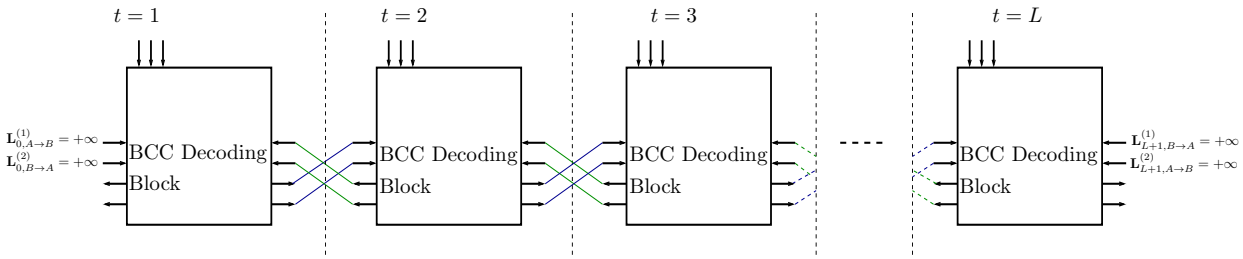

Fig. 4. Block diagram of the iterative message passing decoder of a blockwise BCC.

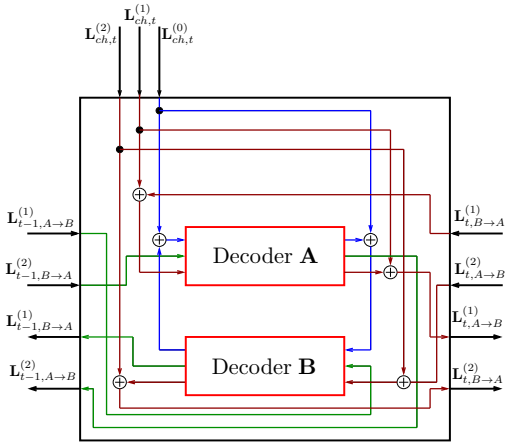

Fig. 5. BCC Decoding Block at time t

current time, namely those which are related to the first inputs of the component decoders. Blue lines show this kind of LLRs.

As the considered channel is the BEC, the LLRs from the channel and from the outputs of the decoders can only be one of the values $+\infty$, 0 and $-\infty$ for 0, erasure and 1 and the only combinations that are possible to happen are as follows:

$$+\infty+\infty=+\infty \quad 0+\infty=+\infty$$
$$-\infty-\infty=-\infty \quad 0-\infty=-\infty$$

IV. Probability of Bit Erasure for Component Convolutional Codes

In order to derive an analytical expression for the probability of erasure of BCCs, the probabilities of erasure of the component codes are required. To catch this goal, we use the method proposed in [8] and apply it to convolutional codes with rate $R=2/3$.

The extrinsic output erasure probabilities are functions of the input erasure probabilities p_0, p_1 and p_2,

$$p_{e,0}=f_0(p_0,p_1,p_2) \tag{1}$$
$$p_{e,1}=f_1(p_0,p_1,p_2)$$
$$p_{e,2}=f_2(p_0,p_1,p_2)$$

For each component code, there is a BCJR decoder and its lth input at trellis section n is denoted by $L^{(l)}_{\text{in},n}$.[1]

Moving over the trellis, forward and backward state metric values are obtained from the following equations[2]:

$$\alpha_n(\sigma)=\max^*_{\sigma'}(\gamma_n(\sigma',\sigma)+\alpha_{n-1}(\sigma')) \tag{2}$$

$$\beta_{n-1}(\sigma')=\max^*_{\sigma}(\gamma_n(\sigma',\sigma)+\beta_n(\sigma)) \tag{3}$$

where σ and σ' denote the states at time n and $n-1$, respectively, and

$$\gamma_n(\sigma',\sigma)=\sum_{l=1}^{3} L^{(l)}_{\text{in},n}\cdot\left(\frac{1}{2}-v^{(l)}_n\right)$$

Finally, the lth output (extrinsic information) can be calculated as

$$L^{(l)}_{\text{out},n}=\max^*_{(\sigma',\sigma):v^{(l)}_n=0}(\alpha_{n-1}(\sigma')+\gamma_n(\sigma',\sigma)+\beta_n(\sigma)) - \max^*_{(\sigma',\sigma):v^{(l)}_n=1}(\alpha_{n-1}(\sigma')+\gamma_n(\sigma',\sigma)+\beta_n(\sigma))$$

We define the metric vectors $\boldsymbol{\mu}_{\alpha,n}$ and $\boldsymbol{\mu}_{\beta,n}$, whose length is equal to the number of the states. $\boldsymbol{\mu}_{\alpha,n}(i)$ is the forward metric of ith state at time n and $\boldsymbol{\mu}_{\beta,n}$ has the backward metric of the ith state as the ith element. Since for both $\boldsymbol{\mu}_{\alpha,n}$ and $\boldsymbol{\mu}_{\beta,n}$ nonzero elements are always equal, we can normalize these entires to 1.

Due to the linearity of the code, we can assume in the analysis that the all-zero codeword has been transmitted. Let $\mathcal{M}_\alpha=\{\boldsymbol{\sigma}^{(1)}_\alpha,\boldsymbol{\sigma}^{(2)}_\alpha,\dots\boldsymbol{\sigma}^{(|\mathcal{M}_\alpha|)}_\alpha\}$ and $\mathcal{M}_\beta$ denote the set of all possible $\boldsymbol{\mu}_{\alpha,n}$ and $\boldsymbol{\mu}_{\beta,n}$, respectively. $\boldsymbol{\mu}_{\alpha,n}$ is one of the elements of $\mathcal{M}_\alpha$.

Example 1: Consider the rate $R=2/3$ convolutional code with generator matrix

$$\mathbf{G}(D)=\begin{pmatrix}1 & 0 & \frac{1}{1+D+D^2}\\ 0 & 1 & \frac{1+D^2}{1+D+D^2}\end{pmatrix}=\begin{pmatrix}1 & 0 & 1/7\\ 0 & 1 & 5/7\end{pmatrix} \tag{4}$$

[1] We use l to denote the symbol index from the perspective of the component decoder, whereas k denotes the index from the perspective of the overall code. Likewise, t is the time index of the code sequence $\mathbf{v}_t$, while n denotes the index within the trellis of a component decoder at a given time instant t.

[2] $\max^*$ denotes the Jacobian logarithm.

For this code, using observer canonical form, there are four states and in this case, $\mathcal{M}_\alpha$ and $\mathcal{M}_\beta$ are equal and have finite number of elements.

$$\mathcal{M}_\alpha = \mathcal{M}_\beta =$$

$$\{(1,0,0,0),(1,1,0,0),(1,0,0,1),(1,0,1,0),(1,1,1,1)\}$$

The sequence $\ldots, \boldsymbol{\mu}_{\alpha,n-1}, \boldsymbol{\mu}_{\alpha,n}, \boldsymbol{\mu}_{\alpha,n+1}, \ldots$ forms a Markov chain with transition matrix $\mathbf{M}_\alpha$, in which $\mathbf{M}_\alpha(j,k)$ is the probability of coming from state $\boldsymbol{\sigma}_\alpha^{(j)}$ to sate $\boldsymbol{\sigma}_\alpha^{(k)}$. This probability depends on the input erasure probabilities $p_l,\ l = 0,1,2$. Using the following formula, we can obtain the steady state distribution of the Markov chain,

$$\pi_\alpha = \mathbf{M}_\alpha \pi_\alpha \ . \tag{5}$$

With the same method, $\mathbf{M}_\beta$ and π_β are obtained.

For the encoder defined in (4) we get

$$\mathbf{M}_\alpha = \begin{bmatrix} (1-p)^2(2p+1) & (1-p)^2 & (1-p)^3 & 0 & 0 \\ p^2(1-p) & 0 & p(1-p)^2 & p^3-2p+1 & (1-p)^2 \\ p^2(1-p) & p(1-p) & p(1-p)^2 & 0 & 0 \\ p^2(1-p) & p(1-p) & p(1-p)^2 & 0 & 0 \\ p^3 & p^2 & p^2(3-2p) & p^2(2-p) & p(2-p) \end{bmatrix}$$

For simpler presentation we have asssumed that p_0, p_1 and p_2 are equal to p, however in general the elements of this matrix are calculated as a function of these three variables.

Define the matrices $\mathbf{T}^{(l)}$

$$T_{i,j}^{(l)} = P\left(L_{\text{out},n}^{(l)} = 0 | \boldsymbol{\mu}_{\alpha,n} = \boldsymbol{\sigma}_\alpha^{(i)}, \boldsymbol{\mu}_{\beta,n+1} = \boldsymbol{\sigma}_\beta^{(j)}\right)$$

The probability of erasure is equal to

$$\begin{aligned} p_e^{(l)} &= P\left(L_{\text{out},n}^{(l)} = 0\right) = \\ &\sum_{i=1}^{|\mathcal{M}_\alpha|} \sum_{j=1}^{|\mathcal{M}_\beta|} P\left(L_{\text{out},n}^{(l)} = 0 | \boldsymbol{\mu}_{\alpha,n} = \boldsymbol{\sigma}_\alpha^{(i)}, \boldsymbol{\mu}_{\beta,n+1} = \boldsymbol{\sigma}_\beta^{(j)}\right) \\ &\cdot P\left(\boldsymbol{\mu}_{\alpha,n} = \boldsymbol{\sigma}_\alpha^{(i)}\right) \cdot P\left(\boldsymbol{\mu}_{\beta,n+1} = \boldsymbol{\sigma}_\beta^{(j)}\right) \\ &= \pi_\alpha \cdot \mathbf{T}^{(l)} \cdot \pi_\beta \ . \end{aligned} \tag{6}$$

Using the above formula, the desired transfer functions of (1) are acquired.

V. Analysis of Iterative Decoding

A. Density Evolution for BCCs

By means of the erasure probability of the component decoders, we are able to calculate the evolution of the erasure probability during the decoding procedure.[3] As the decoder is the same in all iterations we can use the transfer functions obtained in the previous section recursively to obtain the exact decoding probability of erasure after a certain number of iterations.

[3]An assumption in density evolution is that the messages exchanged by the decoders are independent. For turbo codes it has been shown in [9] that this can be achieved by considering a windowed BCJR decoder. A formal proof for BCCs is still an open problem, but we expect that the technique in [9] can be generalized to the ensembles considered here.

For coupled BCCs, the decoding probability of erasure for symbol $l = 0,1,2$ of decoder A at time t after i iterations can be obtained as

$$p_{D_A,0}^{(i,t)} = f_{D_A,0}\left(q_{D_B,0}^{(i-1)}, q_{D_B,1}^{(i-1)}, q_{D_B,2}^{(i-1)}\right) \tag{7}$$

$$p_{D_A,1}^{(i,t)} = f_{D_A,1}\left(q_{D_B,0}^{(i-1)}, q_{D_B,1}^{(i-1)}, q_{D_B,2}^{(i-1)}\right) \tag{8}$$

$$p_{D_A,2}^{(i,t)} = f_{D_A,2}\left(q_{D_B,0}^{(i-1)}, q_{D_B,1}^{(i-1)}, q_{D_B,2}^{(i-1)}\right) \ , \tag{9}$$

where

$$q_{D_B,0}^{(i-1)} = \epsilon \cdot p_{D_B,0}^{(i-1,t)} \tag{10}$$

$$q_{D_B,1}^{(i-1)} = \epsilon \cdot p_{D_B,2}^{(i-1,t-1)} \tag{11}$$

$$q_{D_B,2}^{(i-1)} = \epsilon \cdot p_{D_B,1}^{(i-1,t+1)} \ . \tag{12}$$

Here $f_{D_A,l}$ is the transfer function for the lth symbol of the component decoder D_A and ϵ denotes the erasure probability of the channel. Because of the symmetric design, the update equations for decoder D_B are identical to those of decoder D_A after interchanging D_A and D_B in equations (7)–(12).

The initial LLRs from before the first iteration are assumed to be set to zero, i.e., $p_{D_A,l}^{(i=0,t)} = p_{D_B,l}^{(i=0,t)} = 1$ for $l = 0,1,2$. However, for coupling length L, all messages which come from time $t < 1$ or $t > L$ are assumed to be known, i.e., all probabilities with time index $t < 1$ or $t > L$ are equal to zero. The decoding probability of erasure at time t for blockwise BCCs after i iterations is:

$$p_{e,t} = \epsilon \cdot p_{D_A,0}^{(i,t)} \cdot p_{D_B,0}^{(i,t)}$$

As a special case, an uncoupled BCC can be achieved by tailbiting and the assumption of $L = 1$. The transfer function for the uncoupled case can be achieved from the above mentioned equations for coupled BCCs by omitting the time index t.

B. Results and Discussion

We want to evaluate the largest probability of erasure of the channel ϵ for which the probability of erasure of BP decoding $p_{e,t}$ converges to zero for all t. To obtain such a threshold, $p_{e,t}$ is evaluated as a function of the number of iterations for different values of ϵ (density evolution). We consider two examples of blockwise BCC ensembles of rate $R = 1/3$ with identical component encoders. The first case corresponds to component encoders with $\mathbf{G}(D)$ as defined in (4), i.e., with generators (1/7, 5/7) in octal form. In the second case we consider the generators (1/5, 7/5), i.e., the feedback polynomial is exchanged. The thresholds ϵ_{BP} for the uncoupled case (tailbiting with $L = 1$) and ϵ_{SC} for the coupled (terminated) case are shown in Table I[4]. It can be observed that spatial coupling leads to a significantly better BP decoding threshold. The value $\epsilon_{\text{SC}}^{\text{W}}$ denotes the threshold that can be achieved with a sliding window decoder that starts at time instant $t = 1$. For the first encoder we see that $\epsilon_{\text{SC}}^{\text{W}}$ is worse than ϵ_{SC}, which is due to the fact that the decoder converges better from

[4]It should be noted that the threshold for tailbiting, due to the circular structure, is equal to the threshold for the uncoupled case for any value of L.

TABLE I
THRESHOLDS OF BLOCKWISE BCCS.

$\mathbf{G}(D)$	ϵ_{BP}	ϵ_{MAP}	ϵ_{SC}	$\epsilon_{\text{SC}}^{\text{W}}$
(1/7, 5/7)	0.5541	0.6654	0.6609	0.6554
(1/5, 7/5)	0.5541	0.6654	0.6609	0.6609

the end of the coupled sequence. This suboptimality of the window decoding threshold can be avoided by exchanging the generator polynomials.

Table I shows also an upper bound ϵ_{MAP} on the threshold of an optimal MAP decoder of the uncoupled codes. This upper bound can be obtained according to the area theorem [10] as solution to the following equation:

$$\int_{\epsilon_{\text{MAP}}}^{1} \bar{p}_{\text{extr}}(p)dp = R \ . \tag{13}$$

Here R is the rate of the BCC and

$$\bar{p}_{\text{extr}}(p) = \frac{1}{3}\left(p_{D_A,0}^{(\infty)} \cdot p_{D_B,0}^{(\infty)} + p_{D_A,2}^{(\infty)} \cdot p_{D_B,1}^{(\infty)} + p_{D_A,1}^{(\infty)} \cdot p_{D_B,2}^{(\infty)}\right)$$

denotes the extrinsic probability of erasure of uncoupled BCCs, which is a function of the channel parameter p. To solve this equation we compute the area under the curve $\bar{p}_{\text{extr}}(p)$ for a sufficiently large number of decoding iterations.

We see that the BP threshold of the coupled codes is close to ϵ_{MAP}. However, some gap is still visible for the considered ensemble, which is equivalent to the one introduced in [6]. As shown in [11], this gap to the MAP threshold vanishes if the original ensemble is generalized to larger coupling memories $m_{\text{BCC}} > 1$.

Another interesting observation is that for BCCs the coupling gain, i.e., the gap between ϵ_{SC} and ϵ_{BP} appears to be significantly larger than for conventional turbo codes (i.e., parallel concatenated convolutional codes). Although thresholds for spatially coupled turbo codes have not yet been investigated in the literature, this follows from the gap between their uncoupled BP threshold and MAP threshold, which have been determined in [10]. The large threshold improvement for the coupled case can be justified by looking at the transfer function of the component codes, as illustrated in Fig. 6.

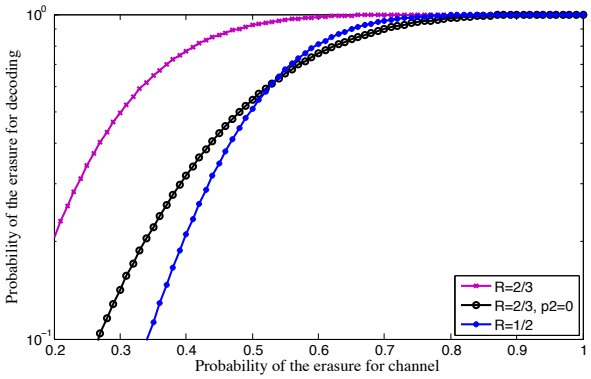

Fig. 6. Probability of erasure of the component convolutional codes.

On one hand, since a rate $R = 2/3$ encoder performs worse than a $R = 1/2$ encoder it is natural that the BP threshold of an uncoupled turbo code is better than the BP threshold of an uncoupled BCC code. However, assuming that the second input of the $R = 2/3$ encoder is known (i.e., $p_2 = 0$) we effectively obtain an equivalent encoder of $R = 1/2$ whose performance is considerably improved (see black curve in Fig. 6). This effect appears for the BCC ensemble at time $t = 1$ and $t = L$, and it propagates further to the other time instants during the iterative decoding procedure, resulting in a threshold improvement.

VI. CONCLUSIONS

In this paper, we derived exact density evolution equations for blockwise BCCs under BP decoding over the BEC. Considering component encoders of memory $m = 2$ we computed BP thresholds for the coupled (terminated) and uncoupled (tailbiting) case and compared them with an upper bound on the MAP threshold. Our threshold analysis confirms the conjecture made in [6] that terminated BCCs can have much better thresholds than their tailbiting counterparts. The threshold of the considered original BCC ensemble is already close to the MAP threshold but for a vanishing gap some generalization to larger coupling memories is required. A major advantage of the BCC construction compared to other turbo-like codes are the superior minimum distance properties in combination with the capacity approaching thresholds.

REFERENCES

[1] A.J. Feltström, D. Truhachev, M. Lentmaier, and K.Sh. Zigangirov, "Braided block codes," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2640–2658, June 2009.

[2] P. Elias, "Error free coding," *IRE Trans. Inform. Theory*, vol. PGIT-4, pp. 29–37, 1954, Also in *Key Papers in Development of Coding Theory*, IEEE press, New York, NY, 1974.

[3] Y.Y. Jian, H.D. Pfister, K.R. Narayanan, R. Rao, and R. Mazahreh, "Iterative hard-decision decoding of braided BCH codes for high-speed optical communication," in *Proc. IEEE Global Telecommunications Conference, 2013. GLOBECOM '13.*, Dec. 2013.

[4] J. Justesen, K.J. Larsen, and L.A. Pedersen, "Error correcting coding for OTN," *IEEE Commun. Magazine*, vol. 48, no. 9, pp. 70–75, 2010.

[5] B.P. Smith, A. Farhood, A. Hunt, F.R. Kschischang, and J. Lodge, "Staircase codes: FEC for 100 Gb/s OTN," *J. Lightwave Technol.*, vol. 30, no. 1, pp. 110–117, 2012.

[6] W. Zhang, M. Lentmaier, K.Sh. Zigangirov, and D.J. Costello, Jr., "Braided convolutional codes: a new class of turbo-like codes," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 316–331, Jan. 2010.

[7] S. Kudekar, T.J. Richardson, and R.L. Urbanke, "Threshold saturation via spatial coupling: Why convolutional LDPC ensembles perform so well over the BEC," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 803 –834, Feb. 2011.

[8] B.M. Kurkoski, P.H. Siegel, and J.K. Wolf, "Exact probability of erasure and a decoding algorithm for convolutional codes on the binary erasure channel," in *Proc. IEEE Global Telecommunications Conference, 2003. GLOBECOM '03.*, Dec. 2003, vol. 3.

[9] M. Lentmaier, D.V. Truhachev, K.Sh. Zigangirov, and D.J. Costello, Jr., "An analysis of the block error probability performance of iterative decoding," *IEEE Trans. Inf. Theory*, vol. 51, no. 11, pp. 3834–3855, Nov. 2005.

[10] Cyril Méasson, *Conservation laws for coding*, Ph.D. thesis, ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE, 2006.

[11] S. Moloudi, M. Lentmaier, and A. Graell i Amat, "Braided convolutional codes - a class of spatially coupled turbo-like codes," in *Proc. International Conference on Signal Processing and Communications (SPCOM)*, Bangalore, India, July 2014.