Spatially coupled turbo-like codes: a new trade-off between waterfall and error floor

Downloaded from: https://research.chalmers.se, 2019-02-27 21:14 UTC

Citation for the original published paper (version of record):
Spatially coupled turbo-like codes: a new trade-off between waterfall and error floor
IEEE Transactions on Communicationshttp://dx.doi.org/10.1109/TCOMM.2019.2896584

N.B. When citing this work, cite the original published paper.
Spatially coupled turbo-like codes: a new trade-off between waterfall and error floor

Saeedeh Moloudi, Michael Lentmaier, Senior Member, IEEE,
and Alexandre Graell i Amat, Senior Member, IEEE

Abstract—Spatially coupled turbo-like codes (SC-TCs) have
been shown to have excellent decoding thresholds due to the
threshold saturation effect. Furthermore, even for moderate
block lengths, simulation results demonstrate very good bit error
rate performance (BER) in the waterfall region. In this paper,
we discuss the effect of spatial coupling on the performance
of TCs in the finite block-length regime. We investigate the
effect of coupling on the error floor performance of SC-TCs
by establishing conditions under which spatial coupling either
preserves or improves the minimum distance of TCs. This allows
us to investigate the error floor performance of SC-TCs by
performing a weight enumerator function (WEF) analysis of the
corresponding uncoupled ensembles. Our results demonstrate
that spatial coupling changes the design trade-off between the
waterfall and error floor performance. Instead of optimizing the
BP threshold of uncoupled TCs, which in turn leads to a higher
error floor, we can take advantage of the threshold saturation
property of SC-TCs. Choosing strong ensembles, characterized
by good MAP thresholds and low error floors, the corresponding
SC-TCs are then able to simultaneously approach capacity and
achieve very low error floor.

Index Terms—Bound on minimum distance, expurgated
bounds, spatially coupled turbo-like codes, union bound, weight
enumerator analysis

I. INTRODUCTION

Turbo codes [1] and low-density parity check (LDPC)
codes [2] are adopted in many communication standards
[3]–[6] because they can practically approach the Shannon
limit. Recently, it has been proved that LDPC convolutional
codes [7], [8] —also known as spatially coupled LDPC (SC-
LDPC) codes—exhibit the remarkable threshold saturation
phenomenon [9]–[12], i.e., for an SC-LDPC ensemble, the
belief propagation (BP) decoder can achieve the threshold
of the optimal maximum-a-posteriori (MAP) decoder of the
underlying uncoupled ensemble. It then follows the remarkable
property that regular SC-LDPC codes achieve capacity with
BP decoding as their variable node degrees tend to infinity.
For finite block lengths, however, larger variable node degrees
increase the number of short cycles in the factor graph [13],
which negatively impacts the performance of a BP decoder.

In this work, we take another approach and consider some
generalized LDPC codes, for which the factor nodes represent
stronger component codes while the variable nodes have
degree one and two only. In particular, we consider a class
of codes defined by factor graphs with convolutional code
trellis constraints that we refer to as turbo-like codes (TCs).
Spatially coupled TCs (SC-TCs) were introduced in [14]–[16],
and it was proved that threshold saturation also occurs for
this class of codes. A density evolution analysis shows that,
by having stronger component codes, SC-TCs can achieve
excellent decoding thresholds with low variable node degrees.

The aim of this paper, which is an extension of [17], is
to investigate the performance of these codes in the finite
block-length regime. We consider the same TC ensembles
as those in [14]–[16], namely parallel concatenated codes
(PCCs) [1], serially concatenated codes (SCCs) [18], [19],
braided convolutional codes (BCCs) [20], [21], and hybrid
concatenated codes (HCCs) [22], [23]. As the first step of
our investigation, using the decoding thresholds on the binary
erasure channel (BEC) obtained in [15], [16] and the method
described in [24], [25], we predict the decoding thresholds
over the additive white Gaussian noise (AWGN) channel.
Using these thresholds together with the provided simulation
results, we discuss the effect of spatial coupling on the
performance of TCs in the waterfall region over the AWGN
channel. Then, we investigate the effect of coupling on the
error floor performance of TCs. We prove that any given
code from one of the uncoupled TC ensembles can be used
to construct a corresponding time-invariant SC-TC such that
for any codeword of a given Hamming weight there exists a
codeword in the corresponding uncoupled code with lower or
equal weight. Thus, the minimum Hamming distance of the
SC-TC is at least as large as that of the original uncoupled
code. The proof relies on some conditions on the permutations
used in the coupled code, which can serve as a guideline
for unwrapping codes from the TC ensembles. Based on
this minimum distance property and results from computer
simulations we conjecture that the error floor performance
of a TC is not degraded by spatial coupling if the derived
conditions are fulfilled, which is not always the case if the
permutations are chosen arbitrarily. A further advantage of
this connection between the minimum distance of TCs and
SC-TCs is that it does not require the computation of finite
length weight enumerator functions (WEFs) of the coupled
ensembles, which is computationally infeasible due to the large
number of different edge types in the factor graph. Instead, we
can perform a WEF analysis for the uncoupled TC ensembles.
The compact graph representation [15] of the PCC ensemble is depicted in Fig. 1(b). Each of the sequences $u$, $v^U$, and $v^L$ is represented by a black circle, referred to as variable node. The trellises corresponding to the component encoders are shown by squares, called factor nodes, and they are labeled by the length of the trellises. The sequences $u$ and $v^U$ are connected to the upper trellis $T^U$. Likewise, a reordered copy of $u$ and $v^L$ are connected to the lower trellis $T^L$. In order to emphasize that a reordered copy of $u$ is connected to $T^L$, the corresponding permutation is represented by a line that crosses the edge which connects $u$ to $T^L$.

Fig. 1(c) shows the compact graph representation of the spatially coupled PCC (SC-PCC) ensemble with coupling memory $m = 1$ at time $t$. Consider a collection of PCC ensembles at time slots $t = 1, \ldots, L$, where $L$ is the coupling length. The SC ensemble can be obtained by dividing the information sequence at time $t$, $u_t$, and its reordered copy, $\tilde{u}_t$, into two subsequences, denoted by $u_{t,j}$ and $\tilde{u}_{t,j}$, $j = 0, 1$, respectively. Then these subsequences are spread over time $t$ and $t + 1$. The input sequence to the upper encoder at time $t$ is the sequence $(u_{t,0}, u_{t-1,1})$, reordered by permutation $\Pi_t^U$.

Likewise, the input sequence to the lower encoder at time $t$ is the sequence $(\tilde{u}_{t,0}, \tilde{u}_{t-1,1})$, reordered by permutation $\Pi_t^L$. The information bits at time slots $t \leq 0$ are initialized by zero and the information bits at $t = L$ are chosen in such a way that $u_{L,1} = 0$ and $\tilde{u}_{L,1} = 0$ (i.e., we consider the termination of the coupled chain).

Fig. 2 shows the compact graph representation of the SCC, BCC, and HCC ensembles, and their corresponding spatially coupled ensembles. We refer the interested reader to [15], [16], for a description and more details on these ensembles. In this paper, we restrict ourselves to PCC, SCC and HCC ensembles with identical 4-state component trellises and generator matrix $G = (1,5/7)$, in octal notation. For the BCC ensemble, we consider two identical 4-state component trellises with generator matrix

$$G(D) = \begin{pmatrix} 1 & 0 & 1/7 \\ 0 & 1 & 5/7 \end{pmatrix}. \quad (1)$$

We also restrict ourselves to systematic TCs and SC-TCs with rate $R = 1/3$. Therefore, for the SCC and HCC ensembles, we consider full puncturing of the parity sequences of the outer encoders [15], [16].

III. SPATIAL COUPLING: WATERFALL REGION PERFORMANCE

A. Asymptotic Performance

Using the decoding threshold of an ensemble computed for the BEC, it is possible to predict its decoding threshold over the AWGN channel [24], [25]. This allows us to use the decoding thresholds of the TC and SC-TCs from [15], [16] to predict the corresponding thresholds over the AWGN channel. The results are shown in Table I. In the table, we report the uncoupled BP thresholds, $E_b/N_0$ [dB]BP, MAP thresholds, $E_b/N_0$ [dB]MAP, and the coupled BP thresholds for coupling

\[ \text{The multiplexer is represented by a rectangular in the compact graph representation.} \]
PCCs, and spatially coupled SCC (SC-SCCs) with R. Finite Block-Length Performance BP thresholds. ensembles, and these ensembles can achieve close-to-capacity This means that threshold saturation occurs for the SC-TC ensemble. The MAP thresholds of the underlying uncoupled ensembles. improve to BP thresholds of the considered SC-TC ensembles improve to results confirm that for large enough coupling memory, the BP thresholds of the considered SC-TC ensembles improve to the MAP thresholds of the underlying uncoupled ensembles. This means that threshold saturation occurs for the SC-TC ensembles, and these ensembles can achieve close-to-capacity BP thresholds.

**B. Finite Block-Length Performance**

Fig. 3 shows BER simulation results for PCCs, SCCs, SC-PCCs, and spatially coupled SCC (SC-SCCs) with R = 1/3 and input block length K = 1024 and K = 4096. The corresponding BP and MAP threshold predictions from Table I are also marked in the figure. For the coupled ensembles, we consider a coupling length L = 100 and a sliding window decoder with window size W = 4 [21].

The decoding latency is W·K. It is well known that the PCC ensemble yields better performance than the SCC ensemble in the waterfall region [18]; however, the SCC ensemble has a much lower error floor than the PCC ensemble. By applying spatial coupling, the performance of the PCC and SCC ensembles improves significantly for both input block lengths. This improvement is more substantial for the SCC ensemble than for the PCC ensemble. For instance, the performance of the SCC ensemble with K = 1024 at BER = 10^{-5} improves more than 1 dB with coupling. The coupling gains are in agreement with the decoding thresholds in Table I. As it can be seen, the gap between the BP and MAP threshold of the SCC ensemble is larger than that of the PCC ensemble, hence the expected gain from coupling is bigger for the SCC ensemble.

In Fig. 3, the uncoupled ensemble with K = 4096 and the coupled ensemble with K = 1024 have equal latency, i.e., both ensembles have a decoding latency of 4096 bits. For this latency, the SC-SCC ensemble performs better than the SCC ensemble. However, in the case of PCCs, for a latency of 4096 bits, the uncoupled ensemble performs slightly better than the corresponding coupled ensemble. Interestingly, for equal latency, the SC-SCC ensemble outperforms the SC-PCC ensemble in the waterfall region. Thus, the SC-SCC ensemble yields better performance in both the waterfall and error floor regions.

In the following section, we investigate the impact of spatial coupling on the error floor performance of TCs in terms of their minimum distance.
Then, for any codeword \( v \) with permutation \( \Pi \), which contains termination as a special case. The result local constraints for \( \Pi \) (Fig. 1(c)). For the SC-PCC, we assume a length-
 generalized to higher coupling memories.

Unfortunately, deriving the WEF for SC-TCs is cumbersome. It is possible to unwrap the PCC to form an SC-PCC, 
by proper choice of the permutation matrices, spatial coupling word weights of SC-TCs and TCs. In particular, we prove that, 

A similar connection between LDPC and SC-LDPC codes is proved in [29]–[31]. Here, we restrict ourselves to 
SC-TCs with coupling memory \( m = 1 \), but the proof can be generalized to higher coupling memories.

**Theorem 1:** Consider an uncoupled PCC, \( \tilde{C} \), (see Fig. 1(b)) with permutation \( \Pi^U \) and parity-check matrices \( H_U \) and \( H_L \) corresponding to the upper and lower component encoders. It is possible to unwrap the PCC to form an SC-PCC, \( C \) (Fig. 1(c)). For the SC-PCC, we assume a length-\( L \) coupled chain with termination or tailbiting, and time-invariant permutations. Let us denote the permutations by \( \Pi^U = \Pi^U \), \( \Pi^t = \Pi^t \), and \( \Pi_t = \Pi \), and assume that they satisfy

\[
\Pi^U = (\Pi^U)^{-1} \cdot \Pi \cdot \Pi^L. \tag{2}
\]

Then, for any codeword \( v \in C \), \( v = (v_1, \ldots, v_L) \), \( v_t = (u_t, v_t^U, v_t^L) \), with Hamming weight \( w_H(v) \), there exists a codeword \( \tilde{v} \in \tilde{C} \) such that

\[
w_H(\tilde{v}) \leq w_H(v) .
\]

**Proof:** We prove this theorem for tailbiting of the coupled chain, which contains termination as a special case. The result is thus valid for both cases. Any codeword \( v \in C \) satisfies the local constraints for \( t = 1, \ldots, L \). Therefore, at time \( t \),

\[
(u_{t,0}^U, u_{t-1,1}^U) \cdot (u_t^U, v_t^U) \cdot H_U^T = 0, \tag{3}
\]

\[
(u_{t,0}^L, u_{t-1,1}^L) \cdot (u_t^L, v_t^L) \cdot H_L^T = 0, \tag{4}
\]

where \( u_t' = u_t \cdot \Pi \). The constraints are linear and time-invariant. Thus, for \( t = 1, \ldots, L \), any superposition of the vectors \( (u_{t-1,1}^U, u_{t,2}^U) \cdot v_t^U \) and \( (u_{t-1,1}^L, u_{t,2}^L) \cdot v_t^L \) satisfies (3) and (4), respectively. In particular, consider

\[
\begin{align*}
\sum_{t=1}^{L} ( (u_{t,0}^U, u_{t-1,1}^U) \cdot \Pi^U \cdot v_t^U ) \\
= \left( \sum_{t=1}^{L} (u_{t,0}^U, u_{t-1,1}^U) \cdot \Pi^U \right) \sum_{t=1}^{L} v_t^U \\
= \left( \sum_{t=1}^{L} u_t \cdot \Pi^U \right) \sum_{t=1}^{L} v_t^U ,
\end{align*}
\]

and

\[
\begin{align*}
\sum_{t=1}^{L} ( (u_{t,0}^L, u_{t-1,1}^L) \cdot \Pi^L \cdot v_t^L ) \\
= \left( \sum_{t=1}^{L} (u_{t,0}^L, u_{t-1,1}^L) \cdot \Pi^L \right) \sum_{t=1}^{L} v_t^L \\
= \left( \sum_{t=1}^{L} u_t \cdot \Pi^L \right) \sum_{t=1}^{L} v_t^L .
\end{align*}
\]

Let

\[
\tilde{u} = \sum_{t=1}^{L} u_t \cdot \Pi^U, \quad \tilde{v}^U = \sum_{t=1}^{L} v_t^U, \quad \tilde{v}^L = \sum_{t=1}^{L} v_t^L .
\]

Then, the vectors obtained from (5) and (6) can be rewritten as \( (\tilde{u}, \tilde{v}^U) \) and \( (\tilde{u} \cdot \Pi^U, \tilde{v}^L) \), respectively.

The vectors from (5) and (6) satisfy (3) and (4), respectively. Thus,

\[
(\tilde{u} \cdot \Pi^U, \tilde{v}^L) \cdot H_U^T = 0, \tag{7}
\]

\[
(\tilde{u} \cdot \Pi^U, \tilde{v}^L) \cdot H_L^T = 0. \tag{8}
\]

Therefore, \( \tilde{v} = (\tilde{u}, \tilde{v}^U, \tilde{v}^L) \) is a codeword of the uncoupled ensemble.

If all nonzero elements of \( v_t \), \( t = 1, \ldots, L \), occur at different positions, then \( w_H(\tilde{v}) = w_H(v) \). Otherwise, the overlap of the nonzero elements reduces the weight of \( \tilde{v} \) and \( w_H(\tilde{v}) < w_H(v) \).

This theorem can be extended to the other TC ensembles.

**Theorem 2:** Consider an uncoupled SCC (BCC/HCC), \( \tilde{C} \), (Fig. 2). It is possible to unwrap the SCC (BCC/HCC) to form an SC-SCC (BCC/SCC, \( \tilde{C} \) (Fig. 2). For the coupled code, we assume a length-\( L \) coupled chain with termination or tailbiting, and time-invariant permutations which satisfy certain conditions. Then, for any codeword \( v \in C \), \( v = (v_1, \ldots, v_t, \ldots, v_L) \), \( v_t = (u_t, v_t^U, v_t^L) \), there exists a codeword \( \tilde{v} \in \tilde{C} \) such that

\[
w_H(\tilde{v}) \leq w_H(v) .
\]

**Proof:** See Appendix.
Corollary 1: The minimum distance of an SC-TC ensemble $\mathcal{C}$ is larger than or equal to the minimum distance of the underlying uncoupled TC ensemble $\tilde{\mathcal{C}}$,

$$d_{\text{min}}(\mathcal{C}) \geq d_{\text{min}}(\tilde{\mathcal{C}}).$$

By the above theorems, we establish conditions on the permutations under which SC-TCs have equal or better minimum distance than their corresponding TCs. These conditions can be considered as guidelines for selecting proper permutations for SC-TCs.

It should be noted that an exact analysis of the error floor performance of SC-TCs would require results on higher order terms of the weight spectrum and their multiplicities, whose efficient computation or estimation is much more complicated and still an open problem. However, based on our analytical results on the minimum distance as well as results from computer simulations we conjecture that a WEF analysis of uncoupled TC ensembles can be used to approximate the error floor of SC-TCs constructed from such TCs.

V. WEIGHT ENUMERATOR ANALYSIS

In this section, we consider uncoupled TC ensembles and describe how to derive upper bounds on the error rate performance and bounds on the minimum distance of these ensembles based on their WEFs [18], [26]. Then, we compare these bounds for different classes of TCs. For that, we first derive the average input-parity WEF (IP-WEF) of the component encoders. In particular, we describe the steps for a rate-$2/3$ recursive systematic convolutional encoder. A similar method can be used to derive the IP-WEF of any convolutional encoder with arbitrary rate $R$. Then, we use the obtained IP-WEFs to compute the average IP-WEFs of the TC ensembles.

A. Input-Parity Weight Enumerator

Let $A(I_1, I_2, P)$ denote the IP-WEF of a rate-$2/3$ recursive systematic convolutional encoder,

$$A(I_1, I_2, P) = \sum_{i_1} \sum_{i_2} \sum_{p} A_{i_1, i_2, p} I_1^{i_1} I_2^{i_2} P^p,$$

where the coefficient $A_{i_1, i_2, p}$ denotes the number of codewords with weight $i_1$, $i_2$, and $p$ for the first input, the second input, and the parity sequence, respectively.

$A(I_1, I_2, P)$ can be computed as follows. For a trellis with $s$ states, transitions within a trellis section can be described by an $s \times s$ matrix $M$. The element of $M$ in the $r$th row and the $c$th column, $[M]_{r,c}$, corresponds to the trellis branch from the $r$th state to the $c$th state. More precisely, $[M]_{r,c}$ is a monomial $I_1^{i_1} I_2^{i_2} P^p$, where $i_1$, $i_2$, and $p$ are the weights corresponding to the first, second, and third outputs of the transition from the $r$th state to the $c$th state. For a trellis with $N$ sections, the overall transition matrix is $M^N$. Considering that the trellis is initialized and terminated to the all-zero state, the IP-WEF is given by the element $[M^N]_{1,1}$.

Example 1: Assume a terminated, rate-2/3 convolutional encoder with three trellis sections and generator matrix in (1).

The transition matrix can be written as

$$M(I_1, I_2, P) = \begin{pmatrix}
1 & I_2 P & I_1 I_2 & I_1 P \\
I_1 & I_1 I_2 P & I_2 P & I_1 P \\
I_2 P & 1 & I_1 P & I_1 I_2 \\
I_1 I_2 P & 1 & I_1 P & I_2 \\
\end{pmatrix},$$

and the IP-WEF becomes

$$A(I_1, I_2, P) = [M^3]_{1,1} = 1 + I_1 I_2 P^2 + 2 I_1 I_2 P + I_1 I_2 P^3 + 2 I_1 I_2^2 P + I_1 I_2^3 P^2 + 3 I_1 I_2^2 P^2 + I_1 I_2^3 P + I_1 I_2^3 P^2.$$

For a rate-1/2 convolutional encoder, we can obtain the transition matrix $M$ in a similar way. Then, the IP-WEF of the encoder is given by $[M^N]_{1,1}$ and can be written as

$$A(I, P) = \sum_i \sum_p A_{i,p} I^i P^p,$$

where $A_{i,p}$ is the number of codewords of input weight $i$ and parity weight $p$.

Consider the PCC ensemble shown in Fig. 2(b). Let $A^{T U}(I, P)$ and $A^{T L}(I, P)$ denote the IP-WEFs corresponding to the upper and lower component encoder, respectively. The overall IP-WEF depends on the IP-WEF of the component encoders and the permutation used. Averaging over all possible permutations, the coefficients of the average IP-WEF of the PCC ensemble, $A^{PCC}_{i,p}$, can be obtained as [26]

$$A^{PCC}_{i,p} = \frac{\sum_{p_1} A^{T U}_{i,p_1} A^{T L}_{i+p_1-p_1}}{\binom{N}{1}}.$$  

(9)

For the SCC ensemble shown in Fig. 2(b), we denote the IP-WEFs of the outer and inner encoder by $A^{T O}(I, P)$ and $A^{T I}(I, P)$, respectively. Similar to PCCs, the average IP-WEF of the SCC ensemble, $A^{SCC}_{i,p}$, can be computed by averaging over all possible permutations [18]. The coefficients $A^{SCC}_{i,p}$ can be written as

$$A^{SCC}_{i,p} = \sum_{p_1, p_2} \frac{A^{T O}_{i,p_1} A^{T I}_{i+p_1-p_2}}{\left(\begin{array}{c}N \\ i+p_1 \end{array}\right)}.$$  

(10)

We denote the IP-WEFs corresponding to the upper and lower component encoders of the BCC ensemble (Fig. 2(c)) by $A^{T U}(I, P)$ and $A^{T L}(I, P)$, respectively. The coefficients of the average IP-WEF, $A^{BCC}_{i,p}$, can be computed as

$$A^{BCC}_{i,p} = \sum_{p_1} \frac{A^{T U}_{i,p_1-p_2-p_1} A^{T L}_{i+p_1-p_2}}{\left(\begin{array}{c}N \\ i+p_1 \end{array}\right)}.$$  

(11)

To compute the average IP-WEF of the HCC ensemble, $A^{HCC}_{i,p}$, it is possible to combine the methods that we used for PCCs and SCCs. First, the average IP-WEF of the parallel component is computed. Then, $A^{HCC}_{i,p}$ can be obtained by substituting $A^{T O}(I, P)$ in (10) by the computed average IP-WEF of the parallel component [22],

$$A^{HCC}_{i,p} = \sum_{p_1} \sum_{p_2} \frac{A^{T U}_{i,p_1} A^{T L}_{i+p_2} A^{T I}_{i+p_1-p_2-p_1}}{\left(\begin{array}{c}N \\ i+p_1+p_2 \end{array}\right)}.$$  

(12)
It is worth mentioning that by the use of the compact graph representation, TCs can be seen as a class of protograph-based generalized LDPC (GLDPC) codes. Therefore, equivalently, it is possible to compute the average IP-WEF of TCs by the method developed for GLDPC codes in [27], [28].

B. Bounds on the Error Probability

Consider transmission of codewords of a rate-$R$ TC ensemble over the AWGN channel. For a maximum likelihood (ML) decoder, the BER is upper bounded by

$$P_b \leq \sum_{i=1}^{N} \frac{N(1/R-1)}{N} \bar{A}_{i,p} Q \left( \sqrt{2(i+p)R \frac{E_b}{N_0}} \right).$$

Likewise, the frame error rate (FER) is upper bounded by

$$P_F \leq \sum_{i=1}^{N} \frac{N(1/R-1)}{N} \bar{A}_{i,p} Q \left( \sqrt{2(i+p)R \frac{E_b}{N_0}} \right),$$

where

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} \exp \left( -u^2/2 \right) du,$$

and $E_b/N_0$ is the signal-to-noise ratio.

Fig. 4 shows the bounds on the BER performance of the different classes of TCs for $R = 1/3$ and $K = 512$. The bounds are truncated at weight $w = 320$, which is larger than the corresponding Gilbert-Varshamov limit. The HCC ensemble has the lowest error floor, while the BCC and PCC ensembles have the highest error floors. Surprisingly, the error floor of the BCC ensemble is not only high but also has the worst slope among all TC ensembles. On the other hand, the excellent MAP thresholds of the BCC ensemble suggest a good performance for this ensemble under MAP decoding. The contradiction between the excellent MAP threshold of the BCC ensemble and its poor bound suggests that the performance is dominated by few bad permutations.

To verify this, we simulated the BCCs for two scenarios: first, a randomly selected but fixed set of permutations; second randomly chosen permutations for each simulated block. The results are shown in Fig. 5, together with the corresponding bounds. The figure shows that the bounds are in agreement with the simulation results for uniformly random permutations. However, it indicates a significant improvement in FER for the fixed set of permutations. For example, at $E_b/N_0 = 2.5$ dB, the FER improves from $9.5 \cdot 10^{-5}$ to $6.8 \cdot 10^{-7}$. This significant improvement caused by fixing the permutations, supports that the high floor of the BCC ensemble is caused by the poor performance of a small fraction of codes. Thus, in the next section, we compute expurgated union bounds.

C. Bound on the Minimum Distance and Expurgated Bounds

Consider a TC ensemble consisting of $\Omega$ codes in total. The value $\Omega$ follows from the different possible combinations of permutations and depends on the type of the ensemble. Assume that all codes in the ensemble are selected with equal probability. Then, the number of codewords with weight $w$ over all possible codes in the ensemble is $\Omega \overline{A}_w$, where $\overline{A}_w$ is the average WEF of the ensemble. Therefore, given an integer value $\hat{d}$, the total number of codewords with weight $w < \hat{d}$ can be computed as

$$\Omega_{w<\hat{d}} = \sum_{w=1}^{\hat{d}-1} \overline{A}_w.$$

By considering that these codewords are spread over different possible codes, we can obtain an upper bound on the number of codes with minimum distance $d_{\min} \geq \hat{d}$,

$$\Omega_{w\geq\hat{d}} \leq \Omega - \sum_{w=1}^{\hat{d}-1} \overline{A}_w.$$
Let $\alpha$ denote the fraction of codes with $d_{\text{min}} \geq \bar{d}$. Then, $\alpha$ is upperbounded by

$$\alpha < 1 - \sum_{w=1}^{\bar{d}-1} \tilde{A}_w.$$  \hfill (15)

For a given $\alpha$ and $\tilde{A}_w$, an analytical bound on the minimum distance of an ensemble can be obtained by computing the largest $\bar{d}$ which satisfies (15). In fact, this bound shows that a fraction $\alpha$ of all possible codes has minimum distance $d_{\text{min}} \geq \bar{d}$. In Fig. 6, considering different classes of TCs with $R = 1/3$, this bound is computed for $\alpha = 0.5$ and different input block lengths. For comparison, we have computed $\tilde{A}_w$ for unstructured $(J, K)$-regular LDPC code ensembles, as introduced and analyzed by Gallager [2], [32]. The corresponding minimum distance bound is also depicted in the figure. As the results show, the HCC ensemble has the best minimum distance, and the $(2, 3)$ LDPC code ensemble has the worst. As an example, for $K = 300$ the values computed for HCCs, BCCs, SCCs, PCCs, $(4, 6)$, and $(2, 3)$ LDPC code ensemble are $\bar{d} = 129, 99, 37, 10, 119$ and 3, respectively.

For demonstrating the impact of the variable node degree $J$, we have computed this bound for $(J, 2J)$ LDPC code ensembles with $J = 2, \ldots, 7$, in comparison with punctured TCs of $R = 1/2$. The results are shown in Fig. 7. According to this figure, LDPC codes can be competitive in terms of minimum distance, provided that the variable node degree is chosen large enough. For example, to obtain a minimum distance bound almost as good as that of the HCC ensemble, the variable node degree $J = 7$ is needed. Among the TC ensembles, according to Fig. 6 and Fig. 7, the minimum distance grows linearly with the input block length for both the BCC and the HCC ensembles [17], [22]. However, the bound on the minimum distance of the HCC ensemble has a higher slope and grows faster than that of the BCC ensemble.

It is also worth to mention that, comparing the results in Fig. 6 and Fig. 7 with the thresholds in Table I, we can observe that the TC ensembles with good MAP threshold also have good minimum distance.

Consider excluding a fraction $(1 - \alpha)$ of codes with $d_{\text{min}} < \bar{d}$ from a TC ensemble. Then, it is possible to compute the upper bound on the performance of this expurgated ensemble. The average BER of the expurgated ensemble is upperbounded by

$$P_b \leq \frac{1}{\alpha} \sum_{i=1}^{kN} \sum_{w=1}^{nN} \frac{i}{N} \tilde{A}_{i,w} \cdot Q \left( \sqrt{2wR \frac{E_b}{N_0}} \right).$$  \hfill (16)

The bounds for the expurgated TC ensembles are shown in Fig. 8 for $\alpha = 0.5$, which means that half of the codes with $d_{\text{min}} < \bar{d}(\alpha)$ are excluded. For comparison, we also provide the corresponding union bounds in the same figure. It can be seen that for all TC ensembles except the PCC ensemble, the error floor estimated by the expurgated bound is much steeper and lower than that resulting from the union bound. In other words, expurgation improves the performance of the SCC, BCC, and HCC ensembles significantly.

For the BCC and HCC ensembles, the gap between the expurgated bound and the union bound is very large and notable. To investigate the influence of expurgation on the performance of these ensembles, in Fig. 9 we provide the expurgated bound on the BER of the BCC and HCC ensembles for $\alpha = 0.5$ and $\alpha = 0.99$. Note that for $\alpha = 0.99$, the expurgated bounds are computed by excluding only 1% of the possible codes, and these bounds are still significantly lower and steeper than the corresponding union bounds. For the BCC ensemble, the gap between the expurgated bounds for $\alpha = 0.5$ and $\alpha = 0.99$ is much smaller than that of the HCC ensemble. Therefore, for $\alpha = 0.99$, the BCC ensemble has slightly steeper and lower error floor than the HCC ensemble. The fact that changing $\alpha$ has a little impact on the expurgation of the BCC ensemble suggests that only a small fraction of the codes have poor distance properties. This means that for
a BCC with randomly selected but fixed permutations, with high probability the error floor is as steep and low as the corresponding expurgated bound for an ML decoder.

VI. DISCUSSION AND CONCLUSION

We investigated the performance of SC-TC ensembles with finite block length in both waterfall and error floor regions. The two primary results can be summarized as follows. First, the performance improvement of spatial coupling in the waterfall region is not only visible asymptotically, for block lengths tending to infinity, as predicted by the threshold saturation [15]. Our simulation results demonstrate that spatial coupling can also improve the waterfall performance even for moderate block length and fixed latency with window decoding. Second, considering certain conditions, spatial coupling either improves or preserves the minimum distance of TCs. Therefore, we can construct from any given good TC an SC-TC with equal or better minimum distance. Using this fact, instead of performing the cumbersome WEF analysis for the coupled ensemble, we derived the WEF for uncoupled ensembles. Then, based on the WEF, we computed bounds on BER performance and the minimum distance of TCs. As the coupled ensembles have equal or larger minimum distance than the uncoupled ensembles, we have used the computed bounds for TCs to approximate the error floor of SC-TCs. The results from the WEF analysis of TCs demonstrate very low error floors for SCC, BCC, and HCC ensembles. Moreover, for the BCC and HCC ensembles, the minimum distance grows linearly with block length.

Comparing the thresholds of SC-TC ensembles (see Table I) and the results from the WEF analysis, we observe that the ensembles with better MAP thresholds also have larger minimum distance and lower error floor. However, so far, only PCCs have been widely used in various standards—such as UMTS and LTE—because of their good BP thresholds and good performance in the waterfall region. Other TC ensembles have received much less attention for commercial use, although they have better MAP threshold and distance properties than PCCs. Our results confirm that the BP thresholds of these ensembles can be significantly improved by applying coupling. Also, regarding the finite length regime, while their error floor stays at very low error probabilities, their waterfall performance gets much closer to capacity. This brings us to the conclusion that by coupling a given ensemble with close to capacity MAP threshold and low error floor, such as SCCs, BCCs, and HCCs, the resulting ensemble is very promising and can perform close-to-capacity, yet achieving low error floor, with a low complexity iterative decoder.

Finally, we should remark that the considered bounds estimate the error floor of an ML decoder. To obtain bounds on the performance of the BP decoder, more investigations on the corresponding absorbing sets [33] and pseudo-codewords [34] need to be done.

APPENDIX A

PROOF OF Theorem 2

We prove the theorem for the general case of tailbiting.

A. Serially Concatenated Codes

Consider the SCC and SC-SCC ensembles in Fig. 2(a) and (b), and assume that

\[ \Pi^{(1)} = \Pi^{(1)} \cdot \Pi^{(2)}. \]

Any codeword \( \tilde{v} \in \mathcal{C} \) satisfies the local constraints for \( t = 1, \ldots, L \). Therefore, at time \( t \),

\[ (u_t, v_t^0) \cdot H_t^T = 0, \]

\[ ((\tilde{v}_{t,0}^0, \tilde{v}_{t-1,1}^0) \cdot \Pi^{(2)} v_t^1) \cdot H_t^T = 0, \]

where \( \tilde{v}_t^i = (u_t, v_t^i) \cdot \Pi^{(1)} \). The constraints are linear and time-invariant. Therefore, any superposition of the vectors \( (u_t, v_t^0) \)
and \((\tilde{v}_{t-1,1}^{(2)}, \tilde{v}_{t}^{(1)} \cdot \Pi^{(2)} v_t^{(1)})\) from different time slots \(t = 1, \ldots, L\), satisfy (17) and (18), respectively. In particular, we can consider
\[
\sum_{t=1}^{L} (u_t \cdot v_t^{O}) = \left(\sum_{t=1}^{L} u_t \cdot \sum_{t=1}^{L} v_t^{O}\right),
\]
and
\[
\sum_{t=1}^{L} \left((\tilde{v}_{t,0}^{(1)}, \tilde{v}_{t-1,1}^{(2)} \cdot \Pi^{(2)} v_t^{(2)})\right) = \left(\sum_{t=1}^{L} \tilde{v}_{t}^{(1)} \cdot \Pi^{(2)} \sum_{t=1}^{L} v_t^{(2)}\right) = \left(\sum_{t=1}^{L} u_t \cdot \Pi^{(1)} \cdot \sum_{t=1}^{L} v_t^{(1)}\right).
\]

Let
\[
\tilde{u} = \sum_{t=1}^{L} u_t, \quad \tilde{v}^{O} = \sum_{t=1}^{L} v_t^{O}, \quad \tilde{v}^{L} = \sum_{t=1}^{L} v_t^{L},
\]
and substitute (19) and (20) into (17) and (18), respectively. Then
\[
(\tilde{u} \cdot \tilde{v}^{O}) \cdot H_{U}^{T} = 0,
\]
\[
(\tilde{u} \cdot \tilde{v}^{O}) \cdot (\Pi^{1n} \tilde{v}^{L}) \cdot H_{U}^{T} = 0.
\]

Therefore, \(\tilde{v} = (\tilde{u}, \tilde{v}^{O}, \tilde{v}^{L})\) is a codeword of the uncoupled code. If all nonzero elements of \(v_t\), \(t = 1, \ldots, L\), occur at different positions, then \(w_H(\tilde{v}) = w_H(v)\). Otherwise, the overlap of the non-zero elements reduces the weight of \(\tilde{v}\) and \(w_H(\tilde{v}) < w_H(v)\).

### B. Braided Convolutional Codes

Consider the SCC and SC-SCC ensembles in Fig. 2(c) and (d), and assume that \(\Pi_t = \Pi, \Pi_t^{U} = \Pi^U \text{ and } \Pi_t^{L} = \Pi^L\). A valid code sequence of \(\tilde{C}\) has to satisfy the local constraints
\[
(u_t \cdot v_t^{(1)} \cdot \Pi^{U} v_t^{(1)}) \cdot H_{U}^{T} = 0,
\]
\[
(u_t \cdot \Pi_t \cdot v_t^{(1)} \cdot \Pi^{L} v_t^{(1)}) \cdot H_{L}^{T} = 0.
\]

for all \(t = 1, \ldots, L\), where \(H_{U}\) and \(H_{L}\) are the parity-check matrices that represent the constraints imposed by the trellises of the upper and lower component encoders, respectively. Since these constraints are linear and time-invariant, it follows that any superposition of vectors \(v_t = (u_t, v_t^{U}, v_t^{L})\) from different time instants \(t \in \{1, \ldots, L\}\) will also satisfy (23) and (24). In particular, if we let
\[
\tilde{u} = \sum_{t=1}^{L} u_t, \quad \tilde{v}^{L} = \sum_{t=1}^{L} v_t^{L}, \quad \tilde{v}^{U} = \sum_{t=1}^{L} v_t^{U}
\]
then
\[
(\tilde{u} \cdot \Pi \tilde{v}^{U} \cdot \Pi^L \tilde{v}^{L}) \cdot H_{L}^{T} = 0.
\]

Here we have implicitly made use of the fact that \(v_t = 0\) for \(t < 1\) and \(t > L\). But now it follows from (25) and (26) that \(\tilde{v} = (\tilde{u}, \tilde{v}^{U}, \tilde{v}^{L}) \in \tilde{C}\), i.e., we obtain a codeword of the uncoupled code. If all nonzero symbols within \(v_t\) occur at different positions for \(t = 1, \ldots, L\), then \(w_H(\tilde{v}) = w_H(v)\). If, on the other hand, the support of nonzero symbols overlaps, the weight of \(\tilde{v}\) is reduced accordingly and \(w_H(\tilde{v}) < w_H(v)\).

The same result can be proved for HCCs by combining the proofs for PCCs and SCCs.

### REFERENCES


**Saeedeh Moloudi** received the Master Degree in Wireless Communications from Shiraz University, Iran in 2012, and Ph.D. degree in communication and information theory, from Lund university, Sweden in 2018. Her main research interests include design and analysis of coding systems and graph-based iterative algorithms.

**Michael Lentmaier** is an Associate Professor at the Department of Electrical and Information Technology at Lund University, which he joined in January 2013. His research interests include design and analysis of coding systems, graph-based iterative algorithms and Bayesian methods applied to decoding, detection and estimation in communication systems. He received the Dipl.-Ing. degree in electrical engineering from University of Ulm, Germany in 1998, and the Ph.D. degree in telecommunication theory from Lund University, Sweden in 2003. He then worked as a Post-Doctoral Research Associate at University of Notre Dame, Indiana and at University of Ulm. From 2005 to 2007 he was with the Institute of Communications and Navigation of the German Aerospace Center (DLR) in Oberpfaffenhofen, where he worked on signal processing techniques in satellite navigation receivers. From 2008 to 2012 he was a senior researcher and lecturer at the Vodafone Chair Mobile Communications Systems at TU Dresden, where he was heading the Algorithms and Coding research group. He is a senior member of the IEEE and served as an editor for IEEE Communications Letters (2010-2013), IEEE Transactions on Communications (2014-2017), and IEEE Transactions on Information Theory (since April 2017). He was awarded the Communications Society and Information Theory Society Joint Paper Award (2012) for his paper Iterative Decoding Threshold Analysis for LDPC Convolutional Codes.

**Alexandre Graell i Amat** received the MSc degree in Telecommunications Engineering from the Universitat Politècnica de Catalunya, Barcelona, Catalonia, Spain, in 2001, and the MSc and the PhD degrees in Electrical Engineering from the Politecnico di Torino, Turin, Italy, in 2000 and 2004, respectively. From September 2001 to May 2002, he was a Visiting Scholar at the University of California San Diego, La Jolla, CA. From 2002 to 2003, he held a visiting appointment at the Universitat Pompeu Fabra and at the Telecommunications Technological Center of Catalonia, both in Barcelona. From 2001 to 2004, he held a part-time appointment at STMicroelectronics Data Storage Division, Milan, Italy, as a Consultant on coding for magnetic recording channels. From 2004 to 2005, he was a Visiting Professor at the Universitat Pompeu Fabra, Barcelona, Spain. From 2006 to 2010, he was with the Department of Electronics, Telecom Bretagne (former ENST Bretagne), Brest, France. In 2011, he joined the Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden, where he is currently a Professor. Since 2019 he is also Adjunct Research Scientist at Simula UiB, Bergen, Norway. His research interests are in the field of coding theory and its application to areas including distributed storage and computing, privacy, and optical communications. Prof. Graell i Amat received the Post-Doctoral Juan de la Cierva Fellowship from the Spanish Ministry of Education and Science and the Marie Curie Intra-European Fellowship from the European Commission. He received the IEEE Communications Society 2010 Europe, Middle East, and Africa Region Outstanding Young Researcher Award. He was the General Co-Chair of the 7th International Symposium on Turbo Codes and Iterative Information Processing, Gothenburg, Sweden, 2012. He was an Associate Editor of the IEEE COMMUNICATIONS LETTERS from 2011 to 2013 and the IEEE TRANSACTIONS ON COMMUNICATIONS from 2011 to 2016. He is currently the Editor-at-Large of the IEEE TRANSACTIONS ON COMMUNICATIONS.