A Novel Concatenated Coding Scheme: RS-SC-LDPC Codes

Jie Qiu, Li Chen*, Senior Member, IEEE, and Shiqiu Liu, Member, IEEE

Abstract—A novel concatenated coding scheme is introduced, where the Reed-Solomon (RS) code is concatenated with the spatially coupled low-density parity-check (SC-LDPC) code, namely the RS-SC-LDPC code. We first propose the locally systematic encoding of the inner code and its termination when considering the concatenated code has a finite length. A decoding scheme is then proposed, in which the belief propagation (BP) based sliding window decoding (SWD) decodes the inner code with the assistance of the outer Berlekamp-Massey (BM) decoding feedback. This research shows the RS-SC-LDPC codes can achieve a high decoding performance without yielding an error floor until the bit error rate (BER) of $10^{-8}$. They also significantly outperform the incumbent RS-convolutional concatenated (RS-CC) codes, as well as the BCH-SC-LDPC codes.

Index Terms—Concatenated code, RS code, SC-LDPC code.

I. INTRODUCTION

Based on the work of Thorpe [1], spatially coupled low-density parity-check (SC-LDPC) codes [2] can be constructed by first coupling and then lifting [3] a chain of block protographs, yielding the Tanner graph [4] and hence defining the parity-check matrix of the code. Due to its convolutional structure, the belief propagation (BP) based sliding window decoding (SWD) [5] can be utilized to yield a low message recovery latency. There exist approaches to improve the code’s performance, such as eliminating the dominant trapping sets [6] and designing the SC-LDPC codes with a large girth [7], both of which can lead to a lower error floor.

The SC-LDPC code is suitable for data stream transmission but its SWD often yields burst errors. In order to eliminate them, it is desirable to concatenate the SC-LDPC code with an outer code. Reed-Solomon (RS) codes are maximum distance separable codes that have a strong burst error correction capability. They can be employed to eliminate the remaining errors yielded by the SWD. Therefore, they were first considered to be concatenated with SC-LDPC codes in [8]. The concatenated code will have a greater minimum Hamming distance and hence yield a better asymptotic performance where the error floor appears. Moreover, the SWD may suffer from error propagation, that is, a failed decoding window will affect the following ones. Hence, the SWD-Berlekamp-Massey (SWD-BM) algorithm with feedback was also proposed in [8] to decode the concatenated code in order to attenuate the error propagation. In order to decode the RS codes based on the SWD outcome, the SC-LDPC code should hold the locally systematic encoding property. While for a finite length code, termination tail is also needed for the inner code [9]. Therefore, this correspondence presents a more comprehensive construction as well as performance of the code. In particular, we have presented the locally systematic encoding of the inner code, in which its termination has been considered. This research also looks into the choice of outer code in maximizing the SWD-BM performance. Our simulation results show the concatenated codes do not yield an error floor until the bit error rate (BER) of $10^{-8}$. They can also outperform the incumbent RS-convolutional concatenated (RS-CC) code [10] and the existing BCH-SC-LDPC code [11].

II. CONSTRUCTION OF RS-SC-LDPC CODES

A. The RS Codes

Let $\mathbb{F}_{2^p}$ denote the finite field of size $2^p$, where $p$ is a positive integer. An $(n, k)$ RS code defined over $\mathbb{F}_{2^p}$ has length $n = 2^p - 1$ and dimension $k$. Given a message vector $\mathbf{m} = (m_0, m_1, \ldots, m_{n-1}) \in \mathbb{F}_{2^p}^n$, its polynomial presentation is $m(x) = m_0 + m_1 x + \cdots + m_{k-1} x^{k-1}$. Generator polynomial

The codeword $c(x) = x^{n-k} m(x) + x^{n-k} m(x) \mod g(x)$

$\quad = c_0 + c_1 x + \cdots + c_{n-1} x^{n-1}$. (1)

B. The SC-LDPC Codes

The protograph-based SC-LDPC codes can be constructed by first coupling a chain of block protographs with a coupling width of $\omega$. A block protograph is a bipartite graph with $n_c$ check nodes (CNs) and $n_v$ variable nodes (VNs). It can be defined by a base matrix $B$ of size $n_c \times n_v$, where $B = \sum_{i=0}^{L-1} B_i$ and the submatrices $B_i$ represent the edge connections from the VNs at time $t$ to the CNs at time $t + i$. Consequently, the coupled protographs can be defined.

A finite length coupled protograph is obtained by coupling $L$ block protographs, where $L$ is called the coupling length. This coupled protograph can be described by the following base matrix of size $(L + \omega)n_c \times Ln_v$.

$$B_{[0,L-1]} = \begin{bmatrix}
B_0 & \cdots & \\
\vdots & \ddots & \\
B_{L-1} & \cdots & B_0 \\
\end{bmatrix}.$$ (2)
The designed code rate is \( R_{SC}^{(L)} = 1 - \frac{(L+\omega)n_e}{L_n} \). Let \( O_\alpha \) and \( I_\alpha \) denote an all-zero matrix and an identity matrix of size \( a \times a \), respectively. \( I_\theta^{(\omega)} \) denotes the shifted identity matrix with each row of \( I_\theta \) cyclically shifted to the left by \( \theta \) positions, where \( \theta \) is called the shifting factor. It has been shown that \( B_{[0..1]} \) can be systematically designed, ensuring a girth of at least six [7]. In this letter, protographs of the SC-LDPC codes are designed using this paradigm. Applying Fossorier’s condition [12], parity-check matrix \( H_{[0..1]} \) of the designed SC-LDPC code can be obtained by the \( M \)-folded graph lifting [3] based on \( B_{[0..1]} \), where each zero entry is replaced by \( O_M \) while others are replaced by \( I_M^{(\omega)} \), and \( M \) is called the lifting factor. Hence, \( H_{[0..1]} \) can have a girth of at least eight and its transpose can be written as

\[
H_{[0..1]}^T = \begin{bmatrix}
H_0^T(0) & \cdots & H_0^T(\omega)

H_1^T(1) & \cdots & H_1^T(\omega+1)

\vdots & \ddots & \vdots

H_{(L-1)}^T(L-1) & \cdots & H_{(L-1)}^T(L+\omega-1)
\end{bmatrix}
\]

Let \( N_c = n_oM \) and \( N_v = n_vM \). \( H_i(t) \) is of size \( N_c \times N_v \), where \( t = 0, 1, \ldots, L+\omega-1 \). Furthermore, we let \( h_1^{(\lambda,\mu)}(t) \in \{0, 1\} \) denote the row-\( \lambda \) column-\( \mu \) entry of \( H_i(t) \).

C. The Concatenation

Let \( \lambda \) denote the number of outer RS codes and \( \gamma \) denote the individual RS codeword index, where \( 0 \leq \gamma \leq \lambda - 1 \). Given \( \lambda \) message vectors \( \omega \), each of them is encoded by an \((n, k)\) RS code as in (1). Then \( \lambda \) RS codewords \( \omega \) are converted into a binary sequence, which is further divided into \( L \) blocks of equal length as \( U_{[0..1]} = (U_{[0..1]}^{(0)} \cdots U_{[0..1]}^{(L-1)}) \), where \( U_{[0..1]} = (U_{1,0}, U_{1,1}, \ldots, U_{1,N_v-N_c+1}) \in F_2^{N_v-N_c} \). The above concatenation requires

\[
\lambda \cdot \gamma \cdot n_p = L(N_c - N_v)
\]

for the finite length concatenated code. In practice, if dimension of the inner code, i.e., \( L(N_c - N_v) \), does not span \( \lambda \) RS codewords, zero padding can be applied. Finally let \( V_{[0..1]} \) denote the concatenated codeword, where \( V_{[0..1]} = (V_{[0..1]}^{(0)} \cdots V_{[0..1]}^{(L-1)}) \in F_2^{L(\lambda,\gamma)} \). The inner encoding should be locally systematic so that an outer decoding can function with the output of the inner SWD. This is addressed in the following section.

III. Encoding of the Inner Code

A. Locally Systematic Encoding

A finite length SC-LDPC code should ensure \( V_{[0..1]} \cdot H_{[0..1]}^T = \Omega \), which implies

\[
V_{[0..1]} \cdot H_0^T(0) + V_{[0..1]} \cdot H_1^T(1) + \cdots + V_{[0..1]} \cdot H_{[0..1]}(L+\omega-1) = \Omega
\]

for \( t = 0, 1, \ldots, L-1 \). Since the encoder is locally systematic, we have \( V_{[0..1]} = [V_{[0..1]}^{(0)} \cdots V_{[0..1]}^{(L-1)}] \), where \( V_{[0..1]}^{(0)} = U_{[0..1]} \) and \( V_{[0..1]}^{(1)} \) is the parity-check portion of length \( N_c \). Provided that \( H_{[0..1]}(t) \) is full rank, (4) can be utilized to generate \( V_{[0..1]} \). For the designed code, \( H_0(t) \) is generated by applying the \( M \)-folded lifting on \( B_0 \) at time instant \( t \). Hence, let \( B_0 = [B_0^{(0)} \ B_0^{(1)}] \) and \( B_0^{(1)} \) is designed as an identity matrix ensuring \( B_0 \) being full rank. Consequently, \( H_0(t) \) is also full rank, where \( H_0(t) = [H_0^{(0)}(t) \ H_0^{(1)}(t)] \) and \( H_1^{(1)}(t) \) are the lifted outcomes of \( B_0^{(0)} \) and \( B_0^{(1)} \), respectively. Since \( H_0^{(1)}(t) \) is the key to generate the check bits \( V_{t,j} \), where \( j = N_c - N_c, N_c - N_c + 1, \ldots, N_c - 1 \), we need to specify the parity-check equations that generate \( V_{t,j} \), including the column indices of the non-zero entries in \( H_0^{(1)}(t) \).

Let \( \rho \) denote the index of the parity-check equation at each time instant, where \( \rho = 0, 1, \ldots, N_c - 1 \). Furthermore, let \( \rho = j - (N_c - N_c) \) denote the position of \( V_{t,j} \) in \( V_{[0..1]}^{(1)} \), where \( \rho = 0, 1, \ldots, N_c - 1 \). Note that \( H_0^{(1)}(t) \) can be obtained by replacing each zero entry in \( B_0^{(1)} \) by \( O_M \) while others by \( I_M^{(\omega)} \). Since \( V_{[M-\rho]} \) is the transpose of \( B_0^{(0)} \), we can determine \( \rho \) by

\[
\rho = \frac{\rho_c}{M} \mod M.
\]

Therefore, the locally systematic encoding follows

\[
V_{t,j} = U_{t,j}.
\]

for \( j = 0, 1, \ldots, N_v - N_c - 1 \). Otherwise, for \( j = N_v - N_c, N_v - N_c + 1, \ldots, N_v - 1 \),

\[
V_{t,j} = \sum_{\rho=0}^{N_c-N_v-1} V_{t,j} \cdot h_0^{(\rho,\mu)}(t) + \sum_{\rho=0}^{N_v-N_c-1} V_{t,j} \cdot h_1^{(\rho,\mu)}(t).
\]

B. Termination of the Inner Code

For the finite length concatenated code, the termination tail can be generated by using the partial syndrome \( P_L \) [9], which will be defined later. The calculation of partial syndrome requires matrix inverse, which is described as follows.

Let \( V_{t,[L,L+\tau-1]} \) denote the termination tail of length \( \tau N_v \), where \( \tau = \omega + 1 \). The concatenated codeword satisfies

\[
V_{[0..1],[L,L+\tau-1]} \cdot H_{[0..1],[L,L+\tau-1]} = \Omega_{[0..1],[L,L+\tau-1]} \cdot \Omega, \tag{8}
\]

where \( \Omega_{[0..1],[L,L+\tau-1]} \) is a zero vector of length \( (L+\tau+\omega)N_c \). The above equation implies

\[
V_{[0..1],[L,L+\tau-1]} \cdot H_{[0..1],[L,L+\tau-1]} = \Omega_{[0..1],[L,L+\tau-1]} \cdot P_L, \tag{9}
\]

\[
V_{[L,L+\tau-1]} \cdot H_{[L,L+\tau-1]} = P_L \mid \Omega_{[0..1],[L,L+\tau-1]}, \tag{10}
\]

where \( P_L \) is the partial syndrome of length \( \omega N_v \). Note that \( H_{[L,L+\tau-1]} \) may have dependent columns, which should be excluded, resulting in a matrix \( F(L) \) of size \( \tau N_v \times \eta \), where \( 0 < \eta \leq (\omega + \tau)N_c \). As a result, equation (10) becomes

\[
V_{[L,L+\tau-1]} \cdot F(L) = \Omega_{[L+\tau-1]}, \tag{11}
\]

where we denote \( \Omega = [\Omega_{[0..1],[L,L+\tau-1]} \mid \Omega_{[0..1],[L,L+\tau-1]}] \) which is of length \( \eta \).

Since \( F(L) \) may not allow right inverse, \( F(L) \) needs to be decomposed into two submatrices, one of which is full rank. The tail bits can be generated by using these two submatrices, respectively. First, we exclude all dependent rows of \( F(L) \), obtaining a full rank square matrix \( F_0(L) \) of size \( \eta \times \eta \). The remaining rows of \( F(L) \) constitute \( F_1(L) \) of size \( (\tau N_v - \eta) \times \eta \). Then, \( V_{[L,L+\tau-1]} \) can be decomposed into \( V_{[L,L+\tau-1]}^{(0)} \) of length \( \eta \) and \( V_{[L,L+\tau-1]}^{(1)} \) of length \( \tau N_v - \eta \). They correspond
Fig. 1. Block diagram of the SWD-BM decoder.

to $F_0(L)$ and $F_1(L)$, respectively. Therefore, equation (11) can be split into

$$V_{r=0}^{(0)}[L, L+\tau-1] \cdot F_0(L) = \Omega_0,$$

$$V_{r=1}^{(1)}[L, L+\tau-1] \cdot F_1(L) = \Omega_1,$$

where $\Omega = \Omega_0 + \Omega_1$. Since $F_1(L)$ is formed with dependent rows, we can generate the corresponding tail $V_{r=1}^{(1)}[L, L+\tau-1]$ randomly, and obtain $\Omega_1$ using (13). $\Omega_0$ can be further obtained by $\Omega_0 = \Omega - \Omega_1$. Since $F_0(L)$ is a full rank square matrix, $V_{r=0}^{(0)}[L, L+\tau-1]$ can be generated by $V_{r=0}^{(0)}[L, L+\tau-1] = \Omega_0 \cdot F_0^{-1}(L)$. Consequently, the termination tail $V_{r=1}^{(1)}[L, L+\tau-1]$ is determined.

IV. THE SWD-BM DECODING

The proposed SWD-BM decoding for the RS-SC-LDPC codes is shown as in Fig. 1. The inner SWD performs log-BP decoding with a window of size $W$ at each decoding instant, aiming to recover the targeted symbols, where $\omega + 1 \leq W \leq L$. For a finite length RS-SC-LDPC code, the entire codeword will be recovered in $L$ window decoding events. After each window decoding, log-likelihood ratios (LLRs) of the targeted symbols of the window will be estimated. Hard decisions are made based on the estimated LLRs. The BM algorithm further decodes the outer RS code based on the SWD outcome, which can correct at most $\left\lfloor \frac{n-k}{2} \right\rfloor$ symbol errors. If the BM decoding succeeds, it will feed back its estimation to the SWD in order to improve the concatenated code’s performance. That says deterministic probabilities (of either one or zero) of those symbols are fed back. Their LLRs would be set to a deterministic scale. Let $L_1 = n_0$ denote the outer codeword length over $\mathbb{F}_2$ and $L_{II} = N_o - N_e$ denote the number of targeted information symbols in each decoding window. We categorize the decoding into three cases.

Case 1: When $L_1 = L_{II}$, each set of targeted symbols is protected by an RS code. If the BM decoding succeeds, LLRs of the targeted symbols will be adjusted to the deterministic scales. Otherwise, their LLRs remain unchanged.

Case 2: When $L_1 < L_{II}$, each set of targeted symbols is protected by several RS codes. Note that $L_1$ may not divide $L_{II}$. In that case, the remaining targeted symbols will be decoded after the next window.

Case 3: When $L_1 > L_{II}$, more than one set of targeted symbols are protected by an RS code. It means that more than one SWD event have been performed before the outer code can be decoded.

It is assumed that the SWD and the BM decoding interleave, where one starts right after another completes. Since the BM algorithm performs finite field operations, its latency is normally marginal compared to the SWD which performs floating point operations and being iterative. However, if the outer code enlarges as in Case 3, the outer decoding latency cannot be ignored. In practice, a timer would be needed to coordinate the SWD and the BM decoding.

V. PERFORMANCE ANALYSIS

This section studies performance of the RS-SC-LDPC codes. Our simulations were performed over the additive white Gaussian noise (AWGN) channel using BPSK. The signal-to-noise ratio (SNR) is defined as $E_b/N_0$, where $E_b$ denotes the transmitted energy per information bit and $N_0$ denotes the noise power spectral density. The parameters of the simulated inner codes are listed in Table I.

Fig. 2 shows performance of RS-SC-LDPC codes which consist of the SC-LDPC-1 code as the inner code and different outer RS codes. Each decoding window performs 100 log-BP iterations. We first consider the choice of the outer code. For the SC-LDPC-1 code, there are 378 targeted information symbols in each decoding window. The RS codes over $\mathbb{F}_{64}$ are chosen as outer codes so that their SWD-BM decoding functions as in Case 1. The (63, 59) RS code can eliminate the error floor, which was otherwise observed for the SC-LDPC-1 code, until the BER of $10^{-8}$. Concatenating the (63, 59) outer RS code yields a greater minimum Hamming distance for the concatenated code, leading to a better asymptotic performance including the error floor. However, Fig. 2 also shows when concatenating with the (63, 61) RS code, error floor starts to appear at the SNR region of 1.7-1.85 dB. We therefore look into the numerical insight of our simulation by measuring ratio of outer decoding failures when concatenating with the (63, 61) and the (63, 59) RS codes, which is presented as in Table I. It can be seen when concatenating with the (63, 59) RS code, the ratio drops by an order of magnitude as the SNR increases, indicating a tendency for the error rate curve in continuing the fall. However, this is not the case for the (63, 61) RS code whose flattened ratios reflect the error floor of the concatenated code. Concatenating with the (31, 29) and
catenated codes have the same length. Adopting the notations between the inner and outer decoders) of 20. The two decoding algorithm [10] with a maximum global iterations the (soft-decision) Viterbi-BM algorithm or the iterative soft convolutional code. The RS-CC code is decoded by either

However, as the codeword length and the window size enlarge, error rate (FER). This is due to the use of stronger outer codes. We further compare the performance between the BCH-SC-LDPC code, demonstrating its potential for streaming applications,” in Proc. IEEE Int. Symp. Inf. Theory (ISIT), Xi’an, China, Oct. 2019, pp. 1–6.

VI. C ONCLUSION

This letter has proposed the RS-SC-LDPC codes, which yield a low message recovery latency and a high decoding performance. The SWD-BM decoding has also been proposed for the code. This research has shown that the concatenated codes can remove the error floor of the SC-LDPC codes. The decoding performance can be maximized by matching the size of targeted symbol set of the SWD and the length of the outer RS code. This concatenated coding scheme can yield a substantial performance improvement over the popular RS-CC code and the BCH-SC-LDPC code, demonstrating its potential of being applied in future communication systems.

REFERENCES