Performance of HARQ-Assisted OFDM Systems Contaminated by Impulsive Noise: Finite-Length LDPC Code Analysis

Tong Bai, Student Member, IEEE, Chao Xu, Member, IEEE, Rong Zhang, Senior Member, IEEE, Anas F. Al Rawi, Member, IEEE, and Lajos Hanzo, Fellow, IEEE

Abstract—Impulsive noise (IN) constitutes a limiting factor in power-line communication systems, especially in those relying on orthogonal frequency division multiplexing (OFDM). If a single time-domain (TD) sample is contaminated by an impulse, all subcarriers become contaminated after the spreading of the discrete Fourier transform (DFT)-based demodulation. In order to alleviate this problem, hybrid automatic repeat-and-request (HARQ) is often invoked. Moreover, the powerful low-density parity-check (LDPC) codes have been increasingly employed in a variety of current and next-generation communication standards. Against this background, in this work, we explicitly characterize the performance of LDPC-coded HARQ-assisted OFDM systems in the face of IN, where the performance metrics of the outage probability (OP) and the average number of retransmissions as well as the effective throughput are analyzed. First of all, we conceive a new algorithm for evaluating the OP in a realistic finite-length LDPC regime, by adapting both the so-called density evolution technique and the waterfall signal-to-noise ratio analysis method. Following this, both the average number of retransmission attempts and the effective throughput are investigated based on our analysis of the OP. The accuracy of the proposed analysis is confirmed by the simulation results, which also effectively quantify the impact of IN on HARQ-assisted OFDM systems in a finite-length LDPC regime.

Index Terms—HARQ, OFDM, impulsive noise, LDPC codes, density evolution, waterfall performance analysis.

I. INTRODUCTION

Impulsive noise (IN) constitutes a deleterious factor in many communication systems, for example, in digital subscriber lines [1], power line communications (PLC) [2], wireless communications [3] and acoustic underwater communications [4]. The IN typically originates from electromagnetic and electronic devices or aquatic animals’ activities. It affects the data transmission in the form of random bursts, each of which typically spans a short duration and has an extremely high power. The detrimental effects of IN become particularly damaging in orthogonal frequency division multiplexing (OFDM) systems, because even if a single time-domain (TD) sample is corrupted by an impulse, all subcarriers of the OFDM symbol become contaminated owing to the spreading effect of the discrete Fourier transform (DFT)-based demodulator at the receiver.

In order to alleviate the deleterious effects of IN, a range of mitigation techniques have been proposed for employment both at transmitter and receiver [2], [5]–[12]. Explicitly, both channel coding [5] and interleaving [6] as well as automatic repeat-and-request (ARQ) [7] have been invoked at the transmitter. By contrast, the mitigation techniques used at the receiver can be generally classified into parametric and non-parametric approaches [13]. As for the parametric methods, IN is assumed to obey a specific statistical model and its parameters can be readily estimated during the training stage. By adapting the mitigation techniques to the noise’s statistics, we may eliminate the IN by employing the techniques of nonlinear preprocessing [8], adaptive filtering [9], symbol detection [10] and iterative decoding [11]. By contrast, the noise statistics are unknown in the context of non-parametric approaches. Hence the receiver has to deal with the IN blindly, by typically invoking erasure decoding [12] or compressed-sensing aided mitigation [2].

More particularly, characterized by its high robustness to sudden perturbations and low-complexity implementation, hybrid ARQ (HARQ) has been widely invoked in diverse communication systems [14]–[16], in cooperation with channel coding. HARQ can be generally classified into three types, namely the Type-I, Type-II and Type-III schemes. Specific to the Type-I HARQ scheme [17], if a packet is correctly received, a positive acknowledgement (ACK) flag is sent back to the transmitter and then the next packet is transmitted. By contrast, if a corrupted one is received, a negative ACK (NACK) is fed back to the transmitter and then the packet is retransmitted, until the transmitter receives an ACK or the affordable number of retransmissions reaches its maximum limit. As for the Type-II HARQ [18] scheme, its operation is the same as that of the Type-I HARQ arrangement, except for the fact that the corrupted received previous copies are saved for joint demodulation in combination with the newly received copy. As for the Type-III HARQ scheme, additional parity is transmitted for each transmission attempt. The benefit of this design philosophy is that the sophisticated HARQ mentioned is only activated, when it is required, because the channel coding mechanism was overwhelmed by a large noise impulse. In the literature, HARQ has been widely investigated in non-static channels from the perspective of information theory [19]–[21] or code design [22]–[24]. However, the analysis...
of HARQ-assisted systems in IN environments is still in its infancy.

As a benefit of their near-capacity performance and low-complexity parallel decoding structure, low-density parity check (LDPC) codes [25] constitute a strong channel-coding candidate for diverse communication systems. In order to analyze the performance of LDPC codes, Richardson et al. proposed the density evolution (DE) concept [26], which is an analytical tool conceived for characterizing the performance of LDPC codes by evaluating the probability density function (PDF) of the log-likelihood ratios (LLRs) during the iterations between the variable nodes (VN) and check nodes (CN). To further extend the performance analysis of LDPC codes, Yazdani et al. conceived the waterfall performance analysis concept [27]. Unfortunately this approach cannot be directly utilized in our system, because it assumes that the noise variance is constant.

Against the background, we analyze the performance of LDPC-coded HARQ-assisted OFDM systems contaminated by IN in terms of their outage probability (OP), average number of retransmission attempts and effective throughput. The main contributions are summarized as follows.

- We conceive an LDPC-coded HARQ-assisted OFDM system for communications in IN environments. Specifically, we design a soft demodulator, which is intrinsically amalgamated with our Type-II HARQ scheme for operation in hostile IN environments.
- We then propose a methodology for analyzing the OP, the average number of transmissions and the effective throughput of our system, in a realistic finite-length codeword regime. Specifically, as for the OP analysis, we modify the DE concept for determining the LDPC codes’s waterfall performance signal-to-noise ratio (SNR) and then extend our solution to the finite-length LDPC regime. In order to characterize the error floor of our system, we propose a new algorithm for evaluating the OP performance.
- We verify the accuracy of the proposed analysis technique through extensive simulation results, which further quantify the impact of IN on our LDPC-coded HARQ-assisted OFDM system.

The paper is organized as follows. The system model and noise model are described in Section II. Then the OP is analyzed mathematically and verified experimentally in Section III. In Section IV, both the average number of retransmissions and the effective throughput are quantified. Finally, we present conclusions in Section V.

II. SYSTEM DESCRIPTION

As discussed in Section I, the Type-I and Type-II HARQ schemes have a wide range of applications in diverse systems [28]–[30], hence we focus our attention on their analysis in this paper. In this section, we describe the general LDPC-coded OFDM system, the noise model and the pair of considered retransmission schemes considered.

A. LDPC-Coded OFDM Systems

1) Transmitter: As illustrated in Fig. 1, we consider an LDPC-coded OFDM system, where an LDPC codeword spans over $L_B$ OFDM symbols, whereas an OFDM symbol consists of $M$ subcarriers. More explicitly, as shown in Fig. 2, a bit sequence denoted by $a$ and having the elements $a \in \{0, 1\}$ is encoded by an LDPC encoder, leading to the sequence of $c$. The interleaved bits denoted by $d$ are mapped to a particular constellation which has the alphabet $\mathcal{X}$ in the symbol mapper, resulting in samples denoted by $X = [X_1, X_2, \ldots, X_{L_B}]$. Each element is processed by an OFDM modulator and we denote the $l_B$th OFDM symbol obtained in the TD by $x_{l_B}$, where we have $l_B \in [1, L_B]$. For simplicity, we omit the process of adding a cyclic prefix. The signals are processed with the aid of channel inversion based frequency-domain (FD) equalization to compensate for the dispersive channel impulse response and then passed through the channel. The received signal is contaminated by both the TD background and impulsive noise. Let us denote the superposition of TD background and impulsive noise by $n_{l_B}$. Then, the received signals is expressed as:

$$y_{l_B} = \sqrt{\rho}x_{l_B} + n_{l_B},$$

where $\rho$ is the power of received signals, while $y_{l_B}, x_{l_B}, n_{l_B}$ are vectors having the dimension of $M$.

2) Noise Model: We consider both the background and impulsive noise. The background noise is modeled by an additive white Gaussian noise (AWGN) process [13]. As for the IN, it generally consists of short pulses, whose amplitude, duration and arrival time are all random variables. The amplitude is mainly modeled by either Gaussian [31] or Middleton’s Class A [32] or alternatively Weibull PDFs [33]. The main models routinely used for the impulse duration are either the partitioned Markov Chain (PMC) [34] or the log-normal PDFs [33]. The common distribution of the arrival time is typically modeled by PMC [34], Poisson [35] and gated Bernoulli-Gaussian (GBG) PDFs [36].

Bearing both the analytical tractability and empirical characteristics in mind, in this treatise the bursty GBG model is invoked for representing the noise process. To elaborate, the GBG model is widely invoked in the literature [36], [37], where the IN is modeled as a sequence of independent, Bernoulli distributed pulses, each of whose width exactly covers the duration of a single TD OFDM sample. However, the disadvantage of the GBG model is its inability to reflect the bursty nature of IN. To overcome this shortcoming, here the IN is modeled using the bursty GBG model of [38]. Specifically, the IN process consists of a sequence of independent bursts, each of which is constituted by $M$ pulses. In this case, any IN burst may contaminate an OFDM symbol. Here let us denote the background noise component without IN by $w_{l_B}$.
and the noise component containing both the background and IN by \( i_{B,t} \). Then the total received noise is expressed as \( n_{B,t} = (1-b)w_{B,t} + b i_{B,t} \), where \( w_{B,t} \) and \( i_{B,t} \) are complex-valued Gaussian vectors of the dimension of \( M \), while the binary flag of \( b \in \{0,1\} \) is used to specify the presence or absence of IN. More particularly, \( p_i = \Pr(b = 1) \) is the probability of IN occurrence. Let us furthermore denote the variance of \( w_{B,t} \) and \( i_{B,t} \) by \( \sigma_w^2 \) and \( \sigma_i^2 \), respectively, whilst representing the ratio between \( \sigma_i^2 \) and \( \sigma_w^2 \) by \( \kappa = \sigma_i^2/\sigma_w^2 \).

3) Receiver: Equipped with the DFT-based OFDM demodulator at the receiver, both the received signal and the noise are transformed to the frequency domain. Then, the resultant signal is formulated as:

\[
Y_{B,t} = \sqrt{\rho}X_{B,t} + N_{B,t},
\]

where \( Y_{B,t}, X_{B,t} \) and \( N_{B,t} \) are the complex-valued vectors with the dimension of \( M \). Let us denote the variance of the noise in an OFDM symbol by \( \sigma_N^2 \). Since the DFT is a linear operation, we have \( \sigma_N^2 = (1-b)\sigma_w^2 + b\sigma_i^2 \). Here we assume that the binary flag \( b \), which indicates the presence or absence of IN, can be perfectly estimated at the receiver side and \( \sigma_N^2 \) is then fed to the symbol detector.

\[ B. \ Type-I \ HARQ-Assisted \ LDPC-Coded \ Scheme \]

In the Type-I HARQ-assisted scheme, we assume that each packet consists of a single LDPC codeword. If an LDPC codeword is correctly received, an ACK flag is fed back to the transmitter, which triggers the transmission of the next packet. However, if the LDPC codeword in the packet is unsuccessfully decoded, the received packet is discarded and a NACK flag is sent back, which triggers the retransmission of the original packet. This retransmission process continues until either the packet is correctly received or the maximum number of transmissions is reached. If the LDPC codeword still remains corrupted, the packet is discarded and then an outage event is reported. This loss of the packet can be solved by ARQ in the upper radio link layer. Let us denote the maximum number of transmission rounds by \( T \). Then the received copy of the \( l_{B,t} \)th OFDM symbol is expressed as:

\[
Y_{l_{B,t}} = \sqrt{\rho}X_{l_{B,t}} + N_{l_{B,t}},
\]

where we have \( t \in [1, T] \) and \( l_{B,t} = [1, L_B] \).

\[ C. \ Type-II \ HARQ-Assisted \ LDPC-Coded \ Scheme \]

Again, we assume that each packet consists of a single LDPC codeword. In contrast to the Type-I HARQ-assisted scheme, the Type-II HARQ-aided arrangement saves the unsuccessfully decoded OFDM symbols and combines them with the newly received copy for joint detection and decoding. Similarly, if the packet is still detected with errors when the maximum number of transmissions is reached, an outage is reported. Note that the retransmissions take place on the basis of LDPC codewords, while the received symbol combining is carried out on the basis of OFDM symbols. Let us denote the noise variance of the \( l_{B,t} \)th OFDM symbol at the \( t \)th ARQ transmission attempt by \( \sigma_i^2 \). The copies of an OFDM symbol are then multiplied by the weights of \( \alpha_{l_{B,t}} = \sqrt{\rho}/\sigma_i^2 \) and summed together based on the MRC principle as [39]:

\[
Y_{l_{B,t}} = \sum_t \alpha_{l_{B,t}} Y_{l_{B,t}} = \sum_t \sqrt{\rho} \frac{1}{\sigma_{i,t}^2} \left( \sqrt{\rho} X_{l_{B,t}} + N_{l_{B,t}} \right) = \sum_t \frac{\rho}{\sigma_{i,t}^2} X_{l_{B,t}} + \sum_t \frac{\sqrt{\rho}}{\sigma_{i,t}^2} N_{l_{B,t}}.
\]  

The noise variance associated with the combined \( l_{B,t} \)th OFDM symbol at the output of the MRC detector is expressed as:

\[
\sigma_i^2 = \mathbb{E} \left[ \left( \sum_t \frac{\sqrt{\rho}}{\sigma_{i,t}^2} N_{l_{B,t}} \right)^2 \right] = \sum_t \frac{\rho}{\sigma_{i,t}^2}.
\]

It can be readily seen from (4) and (5) that the SNR of the \( l_{B,t} \)th OFDM symbol at the \( t \)th ARQ transmission attempt by \( \sigma_i^2 \). The copies of an OFDM symbol are then multiplied by the weights of \( \alpha_{l_{B,t}} = \sqrt{\rho}/\sigma_i^2 \) and summed together based on the MRC principle as [39]:

\[
p(Y_{l_{B,t}}|X_{l_{B,t}}) = \frac{1}{\pi \sum_t \sigma_{l_{B,t}}^2} \times \exp \left( - \frac{Y_{l_{B,t}} - (\sum_t \frac{\sqrt{\rho}}{\sigma_{l_{B,t}}^2}) X_{l_{B,t}}}{\sum_t \frac{1}{\sigma_{l_{B,t}}^2}} \right).
\]  

Then the a posteriori LLRs gleaned from the MRC detector are calculated as:

\[
L_p(d_k) = \ln \frac{\sum_{X_{l_{B,t}} \in \mathcal{X}_{d_k=1}} p(Y_{l_{B,t}}|X_{l_{B,t}}) p(X_{l_{B,t}})}{\sum_{X_{l_{B,t}} \in \mathcal{X}_{d_k=0}} p(Y_{l_{B,t}}|X_{l_{B,t}}) p(X_{l_{B,t}})}.
\]

where \( \mathcal{X}_{d_k=1} \) and \( \mathcal{X}_{d_k=0} \) denote the BPSK subsets, when the specific bit \( d_k \) is fixed to 1 and 0, respectively.
III. Outage Probability Analysis

In order to characterize the reliability of our HARQ-assisted LDPC-coded systems in IN environments, in this section we analyze their OP, namely the probability that a packet is unsuccessfully decoded within the maximum affordable number of transmissions. Specifically, we briefly review the concept of DE and waterfall-SNR analysis. Following this, we appropriately adapt DE and extend the waterfall-SNR analysis concept for IN environments. Furthermore, we propose a specific algorithm for determining the OP versus the SNR for the pair of HARQ transmission schemes considered. Finally, our analytical and experimental results are presented.

A. Block Error Rate of Finite-Length LDPC Codes

The development of DE [26], [40] and of waterfall-SNR analysis [27] facilitates the block error rate (BLER) analysis of finite-length LDPC codes. More explicitly, DE can be used for determining the SNR threshold to be satisfied for achieving correct decoding, by evaluating the PDF of the LLRs during the decoding iterations between VNs and CNs. Based on the result of DE, the waterfall-SNR analysis is capable of providing a BLER approximation for finite-length LDPC codes. In the rest of this subsection, we will briefly introduce the original DE and waterfall-SNR philosophy in the absence of IN and then present numerical results for characterizing the associated accuracy.

1) Density Evolution: In AWGN channels associated with a noise variance of \( \sigma^2 \), the LLRs output by a soft detector typically obey the classic Gaussian distribution having the mean of \( 2/\sigma^2 \) and the variance of \( 4/\sigma^2 \). Hence, the initial PDF of the LLRs forwarded to the VN of the LDPC decoder can be expressed as:

\[
f^{(0)}_v = \mathcal{N}\left(\frac{2}{\sigma^2}, \frac{4}{\sigma^2}\right),
\]

(7)

where \( f^{(0)}_v \) denotes the PDF of the input LLRs forwarded to the VNs of the LDPC decoder, when the iteration index is \( l = 0 \). Let us denote the PDF of the LLRs forwarded to a CN of the LDPC decoder at the \( l \)-th iteration by \( f^{(l)}_v \). The density evolution observed at the CN and VN is iteratively updated, respectively, as follows [41]:

\[
f^{(l)}_v = \Lambda^{-1}\left[\sum_{i=1}^{d_v-1} \mu_i \left(\Lambda \left[f^{(l-1)}_v\right] \otimes (i-1)^{\otimes i-1}\right)\right],
\]

(8)

\[
f^{(l)}_v = f^{(0)}_v \otimes \sum_{i=1}^{d_v-1} \lambda_i \left(f^{(l)}_v\right) \otimes (i-1)^{\otimes i-1},
\]

(9)

where \( d_v \) and \( d_e \) denote the number of neighbors of a VN and a CN, respectively. Furthermore, \( \lambda_i \) and \( \mu_i \) are the fractions of edges belonging to the degree-\( i \) VNs and CNs, respectively, while \( \Lambda[\cdot] \) and \( \Lambda^{-1}[\cdot] \) represent the changes of density due to the transformations of \( g(\cdot) \) and \( g^{-1}(\cdot) \), respectively, where \( g(\cdot) = [\text{sign}(\cdot), \ln\coth((\cdot)/2)] \). Moreover, \( \otimes \) represents convolution. Assuming that the all-zero codeword (\( x = +1 \)) is transmitted, the error probability after \( l \) iterations denoted by \( P^{(l)}_b \) can be expressed as [26]:

\[
P^{(l)}_b = \int_{-\infty}^{0} f^{(l)}_v(x)dx.
\]

(10)

With the aid of (7), (8), (9) and (10), we may search all possible values of \( \sigma^2 \) for obtaining the threshold value, when \( P^{(l)}_b \) converges to zero as the number of iterations tends to infinity. Let us denote the threshold value by \( \sigma^2_{\text{th}} \), yielding:

\[
\sigma^2_{\text{th}} = \sup \left\{ \sigma^2 : \lim_{l \to \infty} P^{(l)}_b (\sigma^2) = 0 \right\}.
\]

(11)

Here the threshold \( \sigma^2_{\text{th}} \) obtained from (11) represents the waterfall-SNR value for infinite-length LDPC codes. However, the waterfall-SNR is higher for the finite-length LDPC codes. Based on the result of DE, let us now discuss the BLER analysis of finite-length LDPC codes.

2) Waterfall Performance Analysis: In general, for an LDPC codeword having the length of \( L \) and the threshold noise variance of \( \sigma^2_{\text{th}} \), the average number of correctable bit errors can be approximated by \( E_b = L \times Q(1/\sigma_{\text{th}}) \) for a binary input AWGN channel [27], where \( Q(\cdot) \) is the Gaussian Q-function. In this case, the BLER can be obtained by comparing \( E_{\text{th}} \) to the PDF of the actual number of bit errors, denoted by \( E_{\text{obs}} \). More explicitly, given a binary-input AWGN channel having a noise variance of \( \sigma^2 \), its BER can be calculated as \( P_b = Q(1/\sigma) \), while the PDF of \( E_{\text{obs}} \) associated with an LDPC codeword length of \( L \) can be formulated in a binomial form as:

\[
f_{E_{\text{obs}}} = \left(\frac{L}{E_{\text{obs}}} \right) P_b^{E_{\text{obs}}}(1 - P_b)^{L - E_{\text{obs}}},
\]

(12)

Given a specific value of \( L \) of our interest, the binomial form of (12) can be accurately approximated by the Gaussian distribution, i.e. we have \( N[L P_b, L P_b(1 - P_b)] \). Accordingly, the BLER denoted by \( P_B \) associated with \( L \) is readily expressed as:

\[
P_B(L) = Q\left(\frac{E_{\text{th}} - L P_b}{\sqrt{L P_b (1 - P_b)}}\right).
\]

(13)

3) Numerical Results: Fig. 3 presents the BLER of finite-length LDPC codes in the context of BPSK-modulated systems communicating over AWGN channels. Specifically, the 1/2-rate regular LDPC codes configured by Table I is used as an example. Three observations can be inferred from the figure. Firstly, the finite-length performance analysis provides a very close match to the simulations. Specifically, the SNR gap between the theoretical analysis and simulation results is within 0.1 dB when \( L = 2048 \), and it becomes even more accurate when \( L \) reaches 6144 or higher values. Second, the finite-length performance analysis usually provides an underestimated BLER prediction for LDPC codes, because of the

<table>
<thead>
<tr>
<th>Description</th>
<th>Value or algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of neighbors of a VN</td>
<td>( d_v = 3 )</td>
</tr>
<tr>
<td>Number of neighbors of a CN</td>
<td>( d_e = 6 )</td>
</tr>
<tr>
<td>Number of decoding iterations</td>
<td>50</td>
</tr>
<tr>
<td>Decoding algorithm</td>
<td>Sum-product algorithm</td>
</tr>
</tbody>
</table>
difference between the accurate binomial form of (12) and the approximate Gaussian distribution invoked in (13). Thirdly, as expected, the BLER performance is improved upon increasing the codeword length. This also demonstrates the importance of finite-length LDPC code analysis.

B. Outage Probability Analysis under IN

The occurrence of IN changes the distribution of the LLR input to the LDPC decoder, which eminently influences the BLER of the LDPC codes. In this subsection, we modify the DE and the waterfall-SNR analysis, in order to facilitate the analysis in an IN environment. In order to present the OP caused by IN, we propose a specific algorithm for calculating the outage of LDPC-coded OFDM systems and then extend it to the Type-I and Type-II HARQ schemes.

1) BLER of Coded OFDM Systems Inflicting IN: Since the channel-induced erroneous transmissions have to be symmetric for accurate DE [40], we have to prove that the channel noise is indeed symmetric in our system, which is seen as follows:

$$P(y|x = 1) = \frac{1 - p_i}{\sqrt{2\pi\sigma_w^2}} \exp \left( -\frac{(y - 1)^2}{2\sigma_w^2} \right) + \frac{p_i}{\sqrt{2\pi\sigma_i^2}} \exp \left( -\frac{(y - 1)^2}{2\sigma_i^2} \right)$$

$$= \frac{1 - p_i}{\sqrt{2\pi\sigma_w^2}} \exp \left( -\frac{(-y + 1)^2}{2\sigma_w^2} \right) + \frac{p_i}{\sqrt{2\pi\sigma_i^2}} \exp \left( -\frac{(-y + 1)^2}{2\sigma_i^2} \right)$$

$$= P(-y|x = -1).$$  \hspace{1cm} (14)

Since an LDPC codeword conveys $L_B$ OFDM symbols, the IN corrupting an LDPC codeword may contaminate several OFDM symbols. The distribution of the initial LLRs depends on the number of impulsive-noise-infested OFDM symbols in an LDPC codeword. More explicitly, if the number of contaminated OFDM symbols is low in a long LDPC codeword, then the average initial LLRs become high and hence the LDPC codeword has a high probability of being correctly decoded.

Let us denote the number of error-infested OFDM symbols in an LDPC codeword by $L_1$, where we have $L_1 \in [0, L_B]$. Then, the initial PDF of the LLRs gleaned from the detector and characterized in (7) can be reformulated as a superposition, which is a function of $L_1$, formulated as:

$$f_v^{(o)}(L_1) = (L_B - L_1)N\left( \frac{2}{\sigma_w^2}, \frac{4}{\sigma_w^2} \right) + L_1N\left( \frac{2}{\sigma_i^2}, \frac{4}{\sigma_i^2} \right).$$  \hspace{1cm} (15)

The PDF of the LLRs is iteratively updated with the aid of (8) and (9). Then, the noise threshold of the channel below which error-free decoding becomes possible can be obtained with the aid of (10) and $\sigma_w^2 = \sup \{ \sigma_w^2 : \lim_{l \to \infty} P_b^o(\sigma_w^2) = 0 \}$. To elaborate a little further, we use the background noise to represent the channel’s noise threshold, while we represent the relationship between the background and IN using $\kappa = \sigma_i^2/\sigma_w^2$.

Given the knowledge of the above noise threshold, we now extend the waterfall-SNR analysis to our IN environment. As observed in Fig. 3, the gap between two curves becomes smaller upon increasing the LDPC codeword length. Therefore, we use the relationship between $E_b$ and $E_{obs}$ in the OFDM symbols not infested by IN to approximate the overall BLER within an LDPC codeword containing $L_B$ OFDM symbols, because the number of uncontaminated OFDM symbols is usually higher than the number of the symbols corrupted.

For an LDPC codeword spanning $L_B$ OFDM symbols having $M$ subcarriers, when $L_1$ symbols are corrupted by IN, the PDF of $E_{obs}$ can be expressed in a similar form to (12) upon replacing $L$ by $M(L_B - L_1)$, yielding:

$$f_{E_{obs}}(L_1) = \left( M(L_B - L_1) \right)(P_b) E_{obs} (1 - P_b)^{L-B-E_{obs}},$$  \hspace{1cm} (16)

which can be approximated by the Gaussian distribution and then the BLER associated with $L_1$ can be formulated as:

$$P_B(L_1) \approx Q\left( \frac{E_b - M(L_B - L_1)P_b}{\sqrt{M(L_B - L_1)P_b(1 - P_b)}} \right),$$  \hspace{1cm} (17)

where we have $E_{ih} = M(L_B - L_1)Q(1/\sigma_{ah})$.

In order to combine the analytical results associated with different values of $L_1$, we propose a convenient algorithm. Before explicitly outlining the algorithm, we have to make an observation. As detailed in [1], bit errors floors exist in IN environments, which are determined by the noise impulse’s occurrence frequency. This in turn determines the number of imperfectly detected OFDM symbols per LDPC codeword, hence directly linking them to the OP. In other words, the OP is discretized into the following values $P_B(0), P_B(1), \ldots, P_B(L_B)$. More explicitly, the probability that $L_1$ OFDM symbols are contaminated within an LDPC codeword constituted by a total of $L_B$ OFDM symbols, is given by the binomial expression of $p(L_1) = \binom{L_B}{L_1}p_i^{L_1}(1 - p_i)^{L_B - L_1}$, where $p_i$ represents the occurrence probability of IN. Let us use $P_{Bf}(L_1)$ to denote the specific BLER at which an outage floor occurs, which is determined by the SNR value. If the SNR encountered is lower than that required for correctly decoding an LDPC codeword conveying $L_1$ error-infested OFDM symbols, an
outage event is declared. The corresponding OP of $P_{\text{OF}}(L_I)$ can be expressed as:

$$P_{\text{OF}}(L_I) = \begin{cases} 1, & \text{if } L_I = 0, \\ 1 - \sum_{l=0}^{L_I-1} p(L), & \text{if } 1 \leq L_I \leq L_B. \end{cases}$$ (18)

Let us now briefly summarize the BLEP calculation in the face of IN.

- Firstly, the channel threshold is obtained with the aid of (15), (8), (9), (10) and (11).
- Secondly, we calculate $P_B(L_I)$ upon increasing the SNR with the aid of (17).
- Thirdly, we compare $P_B(L_I)$ to $P_{\text{OF}}(L_I + 1)$. If $P_B(L_I) < P_{\text{OF}}(L_I + 1)$, we set $P_B = P_{\text{OF}}(L_I + 1)$ until the SNR value reaches its threshold value required for correctly decoding the LDPC codewords having $L_I + 1$ error-infested OFDM symbols.
- Fourthly, we set $L_I = L_I + 1$ and then go to the second step.

This calculation can be formulated using the pseudo-code shown in Algorithm 1.

**Algorithm 1** $P_{\text{out}}$ calculation in IN environment

**Require:** $L_B$, $M$, $k$, $p_i$, SNR$_{\text{start}}$, SNR$_{\text{end}}$

1: $p = \text{zeros}(1, L_B + 1)$
2: $P_{\text{OF}} = \text{zeros}(1, L_B + 2)$
3: for $L_I = 0 : L_B$ do
4: Obtain $\sigma^2_{B,I}$ by (15), (8), (9), (10) and (11);
5: $p(L_I) = (\frac{L_I}{L_B})^k \cdot (1 - p_i)^{L_B - L_I}$;
6: if $L_I = 0$ then
7: $P_{\text{OF}}(L_I) = 1$;
8: else
9: $P_{\text{OF}}(L_I) = 1 - \sum_{l=0}^{L_I-1} p(L)$;
10: end if
11: end for
12: $L_I = 0$;
13: for SNR = SNR$_{\text{start}}$ : SNR$_{\text{end}}$ do
14: Obtain $P_B(L_I)$ through (17);
15: if $P_B(L_I) > P_{\text{OF}}(L_I)$ then
16: $P_B(SNR) = P_{\text{OF}}(L_I)$;
17: else if $P_B(L_I) < P_{\text{OF}}(L_I + 1)$ then
18: $P_B(SNR) = P_{\text{OF}}(L_I + 1)$;
19: $L_I = L_I + 1$;
20: else
21: $P_B(SNR) = P_B(L_I)$;
22: end if
23: end for

2) Type-I HARQ-Assisted LDPC-Coded OFDM Scheme:
Similar to the above LDPC coded systems, its HARQ-assisted counterpart also exhibits OP floors, which cannot be avoided even upon employing an infinite number of transmission attempts. To elaborate, when the first transmission attempt of a packet is unsuccessful due to the occurrence of IN, the following attempts may also encounter IN occurrences, (the probability of such events is dependent on $p_i$), which results in an outage. Let us denote by $P_{\text{OF},I}(L_I)$ the specific BLER at which an outage floor occurs after unsuccessfully decoding an LDPC codeword having $L_I$ error-infested OFDM symbols in Type-I HARQ schemes. Then, given a maximum number of transmission attempts $T$, $P_{\text{OF},I}(L_I)$ can be expressed as:

$$P_{\text{OF},I}(L_I) = \begin{cases} 1, & \text{if } L_I = 0 \\ (1 - \sum_{l=0}^{L_I-1} p(L))^T, & \text{if } 1 \leq L_I \leq L_B. \end{cases}$$ (19)

Assuming that the IN occurrence events of the consecutive transmissions are independent of each other, the OP for the Type-I HARQ-assisted scheme denoted is given by:

$$P_{\text{out},I}(L_I) = (P_{\text{B}}(L_I))^T,$$ (20)

where $P_{\text{B}}(L_I)$ can be calculated with the aid of DE and the finite-length analysis of (17). Then, the analytical OP can be calculated using Algorithm 1 upon replacing $P_B$ and $P_{\text{OF}}$ by $P_{\text{out},I}$ in (19) and by $P_{\text{OF},I}$ in (20), respectively.

3) Type-II HARQ-Assisted LDPC-Coded OFDM Scheme:
In Type-II HARQ-assisted schemes, all the copies of a packet are combined for joint demodulation and decoding, so that the resultant SNR of an OFDM symbol becomes the sum of the SNRs of the corresponding OFDM symbols in the copies received. The resultant PDF of the LLRs of the combined packet may vary, as a function of the number of OFDM symbols contaminated by IN in the copies. For example, given an LDPC codeword spanning $L_B$ OFDM symbols, there are $T^{L_B+1}$ legitimate possibilities for the LLRs’ PDF observed following the detection of the packets combined after $T$ transmissions. Note however that some of the $T^{L_B+1}$ possibilities yield the same LLR distribution. Let us hence denote the total number of different LLR distributions by $J$.

The OP of the scheme is then evaluated step by step as follows:

- Firstly, we categorize all the combinations according to the resultant LLRs’ PDF denoted by $f_j$, $j = \{1, \cdots, J\}$, and then obtain the noise thresholds to be satisfied $\sigma^2_j$, $j = \{1, \cdots, J\}$, for correct LDPC decoding under the different LLR PDFs, with the aid of the DE technique.
- Secondly, we sort the required noise thresholds $\sigma^2_j$ to create $\hat{\sigma}^2_j$ so that we arrive at $\hat{\sigma}^{j+1}_j < \hat{\sigma}^j$, where $j \in \{1, J - 1\}$ and then calculate the corresponding occurrence probabilities denoted by $P(\hat{\sigma}^j)$, where $j = \{1, \cdots, J\}$.
- Thirdly, the components denoted by $P_{\text{out},II}(\hat{\sigma}^j)$ can be obtained in a form similar to (17).
- Fourthly, the outage floor denoted by $P_{\text{OF},II}(j)$ is obtained as $P_{\text{OF},II}(j) = 1 - \sum_{k=0}^{J-1} P(\hat{\sigma}^k)$.
- Finally, the OP floors can also be calculated by Algorithm 1 upon replacing $P_{\text{OF}}$ and $P_B$ by $P_{\text{OF},II}$ and $P_{\text{out},II}$, respectively.

C. Numerical Results

In this subsection, we compare the OP of our Type-I and Type-II HARQ-assisted OFDM systems protected by 1/2-rate regular LDPC codes configured in Table I and operating in IN environments, using our proposed analytical formulas and our simulation results. For simplicity, the LDPC codeword length is set to 4096, which spans over $L_B = 2$ OFDM
symbols. Each OFDM symbol has 2048 subcarriers. BPSK modulation is used. Note that $E_s/N_0$ refers to the signal-to-background noise ratio experienced by each symbol and it is not normalized to the total number of transmission attempts.

The performance attained under various conditions is portrayed in Fig. 4, Fig. 5 and Fig. 6. Here we summarize our general observations. Firstly, the analytical curves are confirmed by the simulation results. Secondly, as anticipated, the OP curves obey a stair-case shape. Thirdly, the OP curves of Type-II HARQ-assisted systems exhibit more steps than those of the Type-I HARQ-assisted systems, because the received packets have a larger number of legitimate combinations due to the packet combining operation in the Type-II HARQ-assisted systems. For example, when $L_B = 2$, the components are categorized into three types, i.e. $I = 0$, $I = 1$, and $I = 2$, for the Type-I HARQ-assisted scheme, while we have 10 possible components for the Type-II HARQ-assisted scheme. In the following, we will elaborate on the specific system configuration for each figure and discuss the numerical results individually.

Fig. 4 shows the performance of Type-I and Type-II HARQ systems for IN occurrence probabilities of $p_i = 0.01, 0.05$ and 0.1. The maximum number of transmissions is set to 3. For the IN, we have $\kappa = 20$ dB. Our observations are as follows. Firstly, a lower $p_i$ implies that the OP floors occur at higher $E_s/N_0$ values, which is because a lower $p_i$ implies that a lower number of OFDM symbols per LDPC codeword is likely to be infested by IN and hence its outage floor occurs at a lower $P_{out}$. In our Type-I HARQ-assisted scheme associated with $T = 3$, for example, the probability of an LDPC code being corrupted by IN is $7.88 \times 10^{-6}$ and $6.86 \times 10^{-3}$ for $p_i = 0.01$ and $p_i = 0.1$, respectively. Secondly, as expected, our Type-II HARQ-assisted systems outperforms the Type-I HARQ-assisted systems. More explicitly, the former requires 4.77 dB lower $E_s/N_0$ to achieve $P_{out} = 10^{-6}$, compared to the Type-I HARQ, because the Type-II HARQ is capable of combining up to 3 packets for joint decoding.

Fig. 5 shows the performance of our Type-I and Type-II HARQ systems for impulsive-to-background noise ratios of $\kappa = 10, 20$ and 30 dB. The maximum number of transmissions is set to 3. For the IN, we have $p_i = 0.05$. It can be inferred from the figures that we need a higher $E_s/N_0$ for overcoming the OP floors upon increasing $\kappa$. This is because we need a higher $E_s/N_0$ for correctly decoding the received LDPC codeword at the same number of error-infested OFDM symbols, when $\kappa$ is higher.

Fig. 6 shows the performance of Type-I and Type-II HARQ systems for the maximum number of transmissions given by $T = 1, 2$ and 3. As for the IN, we set $p_i = 0.05$ and $\kappa = 20$ dB. It can be directly inferred from the figures that we need a lower $E_s/N_0$ to attain a specific OP upon increasing $T$. For the case of the Type-I HARQ, this is because the IN bursts are more likely to be avoided upon using more transmission attempts. Having a higher $T$ reduces the probability that an LDPC codeword remains infested by IN and hence the OP floors occur at lower OP. Again, in Type-II HARQ, several transmissions are combined for joint decoding, hence the overall received $E_s/N_0$ becomes higher, which results in a higher probability of correct decoding and hence the OP is reduced.

IV. THE NUMBER OF PACKET TRANSMISSION ATTEMPTS AND GOODPUT

Some applications, such as, lip-synchronized video streaming, are sensitive to delays. The average number of ARQ transmissions is one of the factors affecting the delay. In a related context, goodput is defined as the number of successfully delivered bits per unit time, which reflects the transmission efficiency of the system. In this section, we characterize the average number of packet transmissions and the attainable goodput of the HARQ-assisted schemes, both analytically and by our numerical results.

A. Mathematical Analysis

1) Number of Packet Transmission Attempts: Let us denote the average number of packet transmissions by $E[t]$, which is expressed as [42]:

$$E[t] = 1 + \sum_{t=1}^{T-1} P_{out,1/II}'t,$$

where $P_{out,1/II}'$ represents the OP of Type-I and Type-II HARQ after $t$ transmissions, respectively.

2) Goodput: Upon denoting the goodput by $\hat{R}$, we calculated it as the long-term average ratio of the number of successfully delivered bits over the total number of bits required [43], [44]. The initial transmission rate for a packet is denoted by $R_{I,1}$ and $R_{II,1}$ for the Type-I and Type-II HARQ schemes, respectively. When $(t - 1)$ extra transmission attempts are used, the transmission rate becomes $R_{I,II,t} = R_{II,1}/t$. In this case, the long-term average transmission rate becomes $R_{I,II} = R_{I,II,1}/E[t]$ and the goodput is expressed as:

$$\hat{R}_{I,1/II} = \frac{R_{I,II,1}(1 - P_{out,1/II}'t)}{E[t]},$$

where $P_{out,1/II}'$ is the OP of the Type-I and Type-II HARQ, when the number of the transmission attempts is $T$.

B. Numerical Results

The average number of transmission attempts and the goodput are compared between the Type-I and Type-II HARQ-assisted systems, coded by 1/2-rate $(3, 6)$ regular LDPC codes in our IN environment. Similar to the setting in Section III, an LDPC codeword length is set as 4096, which spans over $L_B = 2$ OFDM symbols. Each OFDM symbol has 2048 subcarriers. BPSK modulation is used and $T = 3$ is set for the maximum number of transmissions. As for the IN, we set $\kappa = 20$ dB and $p_i = 0.05$.

Fig. 7a depicts our comparison between the Type-I and Type-II HARQ-assisted systems in terms of the average number of transmissions. Several observations can be made. Firstly, as expected, the Type-II HARQ outperforms the Type-I HARQ, again because the Type-II HARQ combines the consecutively received copies so that the resultant SNR becomes higher than that of any of the single packets. The other reason...
The theoretical results are calculated according to Algorithm 1. The code is configured in Table I and the modulation scheme is BPSK. The maximum number of transmission times $T$ is set as 3. For the IN, $\kappa = 20$ dB. The theoretical results are calculated according to Algorithm 1.

Fig. 4: OP performance comparison of Type-I and Type-II HARQ systems under various values of IN occurrence probability, $p_i$. The code is 1/2-rate regular LDPC code configured in Table I and the modulation scheme is BPSK. The maximum number of transmission times $T$ is set as 3. For the IN, $\kappa = 20$ dB. The theoretical results are calculated according to Algorithm 1.

is that even if each transmission attempt suffers from IN, the index of the error-infested symbol may not be the same, hence the combined codeword may become correctly decoded. As for the range $1 \leq T \leq 2$, the two types of HARQ have almost identical performance. This is because when the first trial is erroneously decoded, the second transmission will be triggered regardless on which type of HARQ is used. Secondly, the average number of transmission attempts is consistent with the corresponding OP performance.

Fig. 7b presents our comparison between the Type-I and Type-II HARQ-assisted systems in terms of their goodput. Our observations are as follows. Firstly, the Type-II HARQ outperforms the Type-I HARQ in terms of its goodput, especially when $\hat{R} < 0.25$. Secondly, the goodput has an approximately linear relationship with $E_s/N_0$ in Fig. 7a, because our region of interest in terms of $P_{\text{out}}$ in (22) is close to 0.

V. CONCLUSIONS

The promising paradigm of industrial Internet of Things imposes a stringent reliability requirement on communication systems. The tolerable OP in smart grids, for example, has to be less than $10^{-6}$ [45]. Table II summarizes the required $E_s/N_0$ of both the Type-I and Type II HARQ schemes for achieving the OP of $10^{-6}$ in various setting of $\kappa$. It can be seen that the required $E_s/N_0$ increases upon increasing $\kappa$. Moreover, Table II also confirms that the Type-II HARQ outperforms the Type-I HARQ for all the scenarios considered.

In a nutshell, we analyzed HARQ-assisted OFDM systems contaminated by IN in a realistic finite-length LDPC regime, both with the aid of our modified density evolution and waterfall-SNR analysis. The performance was characterized in terms of the OP, the average number of transmissions and the effective throughput. The simulation results confirm the
Type-II HARQ.

HARQ-assisted schemes in hostile IN environments.

accuracy of our analysis and quantify the efficiency of our HARQ-assisted schemes in hostile IN environments.

<table>
<thead>
<tr>
<th>HARQ scheme</th>
<th>$\kappa = 10$ dB</th>
<th>$\kappa = 20$ dB</th>
<th>$\kappa = 30$ dB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type I</td>
<td>3.2 dB</td>
<td>11.7 dB</td>
<td>21.5 dB</td>
</tr>
<tr>
<td>Type II</td>
<td>-0.5 dB</td>
<td>7.0 dB</td>
<td>18.8 dB</td>
</tr>
</tbody>
</table>

Fig. 6: OP performance comparison of Type-I and Type-II HARQ systems under various values of maximum number transmission times, $T$. The code is 1/2-rate regular LDPC code configured in Table I and the modulation scheme is BPSK. For the IN, $p_i = 0.05$ and $\kappa = 20$ dB. The theoretical results are calculated according to Algorithm 1.

Fig. 7: The average number of transmissions and goodput comparison between Type-I and Type-II HARQ systems. The code is 1/2-rate regular LDPC code configured in Table I and the modulation scheme is BPSK. For the IN, $p_i = 0.05$ and $\kappa = 20$ dB. The theoretical results are calculated according to (21) and (22), for average number of transmissions and goodput, respectively.

REFERENCES


IEEE DRAFT


I. Recommendation G.993.1, “Very high speed digital subscriber line (VDSL) transceivers,”

I. R. G.9960, “Unified high-speed wireline-based home networking Architectures – system architecture and physical layer specification.”


