NATIONAL AND KAPODISTRIAN UNIVERSITY OF ATHENS

SCHOOL OF SCIENCE
DEPARTMENT OF INFORMATICS AND TELECOMMUNICATION

MSc THESIS

Forward-Error Correction for High-Capacity Transmission Systems

Safiya J. Dabwan
7115192100004

Supervisors: **Stylianos Sygletos,** Associate Professor
Dimitris Syvridis, Professor

ATHENS

OCTOBER 2022

MSc THESIS

Forward-Error Correction for High-Capacity Transmission Systems

Safiya J. Dabwan
7115192100004

SUPERVISOR: **Stylianos Sygletos,** Associate Professor
Dimitris Syvridis, Professor

ABSTRACT

This study investigates the interplay between forward error correction FEC and digital back-propagation DBP nonlinearity compensation on a long-haul fibre channel. First, a research-based approach is used to identify the state-of-the-art technologies in FEC for the fibre channel and adapt them to the final design. The design choices includes the usage of trivial bit interleaved coded modulation T-BICM architecture with a concatenated code scheme that uses an iterative soft decoder. The requirement for a concatenated FEC implementation motivated another investigation of a well-performing code combination. The Irregular LDPC and quasi-cyclic QC-LDPC, adopted from DVB-S2 and IEEE 802.11 standards, respectively, were each concatenated with staircase code and compared based on the attained performance. We prove that increasing the fibre transmission distance by a factor of 1/3, from 300km to 400km, while maintaining the same performance and using the same overhead, i.e. 27.5\% is achievable when implementing DBP with 2 steps/span or 3 steps/span, depending on whether the decoding iterations are 10 or 5. This study concludes with favouring LDPC from DVB-S2 over IEEE 802.11's QC-LDPC for long haul fibre channel. The conclusion is made based on the better attained performance for LDPC-DVB, due to its long code lengths, and its support for high coding rates resulting low overhead requirement.

SUBJECT AREA: Forward Error Correction

KEYWORDS: Fibre Channel, LDPC, Digital Back Propagation, BICM

Co-funded by the Erasmus+ Programme of the European Union

Erasmus Mundus Joint Master's Degree
SMART Telecom and Sensing NETworks (SMARTNET) - 2020/2022 intake
Aston University, Triangle, B4 7ET / Birmingham, UK

Email: aipt_smartnet@aston.ac.uk, Website: smartnet.astonphotonics.uk/

ACKNOWLEDGEMENT

This Master Thesis has been accomplished in the framework of the European Funded Project: **SMART Telecom and Sensing Networks (SMARTNET)** - Erasmus+ Programme Key Action 1: Erasmus Mundus Joint Master Degrees – Ref. Number 2017 – 2734/001 – 001, Project number - 586686-EPP- 1-2017-1-UK-EPPKA1-JMD-MOB, coordinated by **Aston University**, and with the participation of **Télécom SudParis, member of IP Paris** and **National and Kapodistrian University of Athens**.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

The dependency on the fibre channel to fulfil the enormous internet traffic demand is a crucial motivation to fully maximise its capacity. These future high-capacity networks can achieve high spectral efficiency, by deploying high multilevel modulation formats. Adapting to a higher modulation scheme requires an increase in the signal quality to maintain the signal-to-noise ratio within the receiver's requirement. Increasing the signal power level is constrained by the nonlinear behaviour of the fibre channel due to the Kerr effect. Digital signal processing strategies, such as Digital Back Propagation, which mainly address the nonlinear behaviour of the fibre channel are attracting the attention of photonics researchers. In parallel, strategies concerned with decreasing the error by performing corrections to the distorted received information, such as forward error correction FEC have also emerged. The implementation of FEC mandates adding processing units that utilize information overhead to correct errors. Integrating coding components adds a complexity cost to the system and translates into increased energy consumption. In addition, the required overhead for the error correction decreases the spectral efficiency. Therefore, careful consideration is taken in the design phase to achieve the least requirement of complexity and overhead. Implementing FEC in nonlinear fibre channels had recently started to catch more attention. The interplay between FEC on the one hand and the other compensation elements existing in the fibre DSP chain on the other is a vital research topic that requires more engagement.

This work investigates the effect of the fibre nonlinearities compensation, using DBP, upon the integration with coded modulation. A literature review is conducted to identify the state-of-art in FEC for fibre channels. Accordingly, the coded-modulation architecture alongside the coding algorithms is chosen. The existence of two coding candidates, irregular LDPC and QC-LDPC, partially adapted from recognised standards, DVB-S2 and IEEE802.11, motivated a sub-research question of which performs better. The comparative study in [5] addressed this question for the wireless channel; however, there is no research that conducted such a study on the nonlinear fibre channel. That motivated the endeavour to investigate which configuration works best for the fibre channel and subsequently use it in the primary model. Numerical simulations were developed in MATLAB for an end-to-end coded modulated system, which integrates DBP compensation. The model is based on the design choices made at the early research stage and specific simulation experiments that were held and reported in this document to define some design metrics that required optimisation for the particular implementation. The simulation's intensity required the usage of the high-performance computing system of the Aston Institute of Photonics Technologies. The end-

to-end coded-modulation system's numerical simulation can be used to perform further investigations on the performance impact of other DSP technologies when integrated with coded modulation; probabilistic and geometric shaping are potential candidates.

Following this chapter, the background chapter presents the literature review and information about the history of FEC, its related important milestones, concepts and debates. Later, Chapter 3 explains the procedure and the appropriate design choices and algorithms used. The following chapter presents the results and provides a thorough analysis. Last but not least, the conclusion chapter highlights the main achievements of the undertaken research and proposes directions for future study.

2. BACKGROUND

To achieve successful transmission of information between two ends, a reliable channel is required. Reliability is an ever-existing issue due to the imperfections of the communication medium. Therefore, various strategies that serve the purpose of assuring the correct reception of the message were developed in the field of information theory and coding from very early. Among them, Forward-Error Correction (FEC) has been one of the most powerful. FEC procedure involves the addition of a small efficient fraction of overhead OH to the original message prior to transmission. At the receiver, this overhead is used to detect if the received message had experienced any erroneous events and correct them, up to a certain extent. From an information-theoretic perspective, the introduction of the overhead redundancy targets to increase the amount of information that is transmitted from the input to the output of the channel, quantified by what is known as mutual information (MI) [1].

Mutual information is an information-theoretic measure that identifies the probabilistic dependency between the channel's output and input. To provide a more precise definition of this term, let us assume a typical scenario of message transmission over a communication channel, depicted in Figure 2.1. A discrete memoryless source (DMS) is assumed at the transmitter generating a continuing flow of discrete and independent symbols. The DMS feeds the channel with its input X, where X is a random variable that accepts any value x from the space of possible values, aka alphabets $\mathscr{X}$. At the other end, the receiver receives the channel output, which is consequently also a random variable, Y. Ideally, those two random variables are to be identical, however, the various impairments (e.g. white Gaussian noise) of the transmission channel prevent this from happening. The more severe the effect of the channel is on the input x, the more variant the output y becomes and the effective information carried in y about x decreases as it gets mixed with channel noise. Mathematically, the amount of information carried out by Y about X, i.e. the Mutual Information (MI) can be defined by the averaged conditional probability logarithm [1] :

$$I(X;Y) = \sum_{x \in \mathscr{X}} \sum_{y \in \mathscr{Y}} P[X = x, Y = y] \log_2 \frac{P[x|y]}{P[x]} \tag{2.1}$$

In his remarkable work published in 1948 [6], Shannon used MI to define R; the maximum attainable quantity of information in a symbol i.e. bits that can be accommodated in a single transmittable symbol while maintaining a reliable communication, which he called capacity. It represented an upper-bound metric of the channel's capability to transmit information. The work of Shannon inspired a

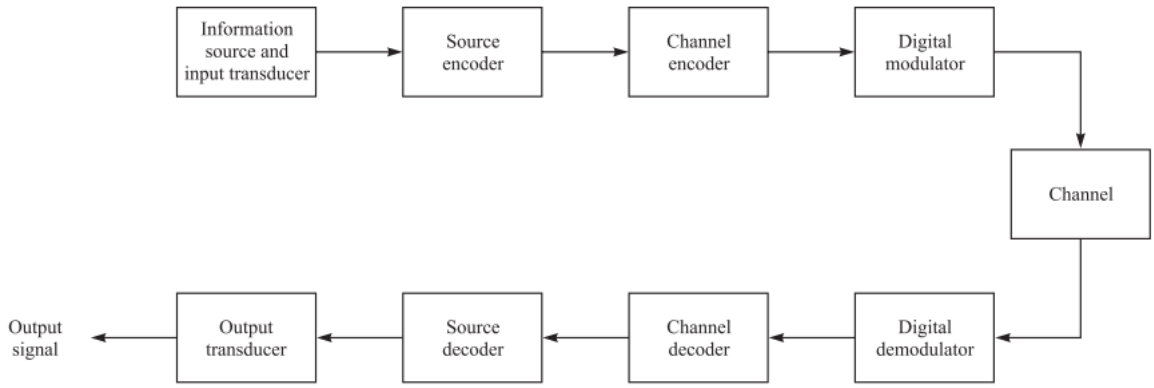


Figure 2.1: Block Diagram of Typical Transmission System (Adapted from [1])

century-long research endeavour in the FEC field. The quest of FEC is concerned with developing codes that are able to enhance the performance to get closer to the channel capacity.

2.1 Forward Error Correction development

Forward Error Correction FEC, is a course of processing the message undergoes at both ends of the communication system to deal with errors. At the transmitting end, the encoder uses the message bits to evaluate the parity bits based on the encoding scheme it follows. At the receiving end, the decoder tries to make the best interpretation out of the received distorted bits to detect the occurrence of errors and correct them.
As Figure 2.2 shows, the encoding process takes place on a limited number of bits, k. Once encoded, the parity bits addition represented increases the size of the output binary sequence to become N, where $N > k$.

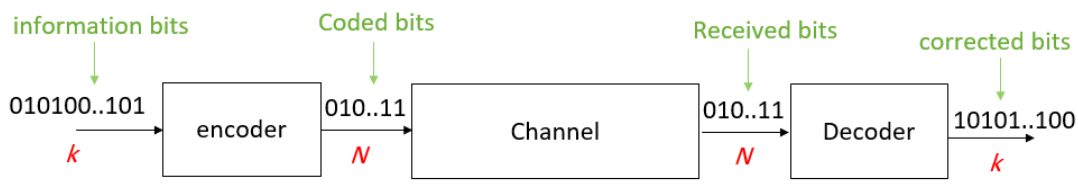


Figure 2.2: Transmission with FEC integration

Encoding and decoding blocks underwent intensive growth over the seven past decades. The development of FEC included new encoding schemes, decoding

algorithms, and coding architectures and setups. There is also an application-dependent line of development. Many FEC schemes were designed to serve certain types of communication channels. FEC in IEEE 802.11 developed for wireless channels is a representative example [7].
A milestone in the development of FEC was the pioneering work of Claude Shannon who introduced the concept of channel coding. His contributions marked the start of the FEC from the dark ages and fueled the progress that followed up. Although the development of the encoder moved rapidly, the decoding part didn't follow the same pace because of the complexity of the corresponding algorithms. The limited computational capabilities of the available hardware at that time played a role in slowing down technological developments. That didn't last long though. After a couple of years, advanced Coded-Modulation CM architectures got into the picture, with the recent stages marking the research endeavours in establishing flexible high-performing FEC architectures. A visualization of the different development stages in the FEC history is shown in 2.3. A more detailed discussion follows in the below subsections.

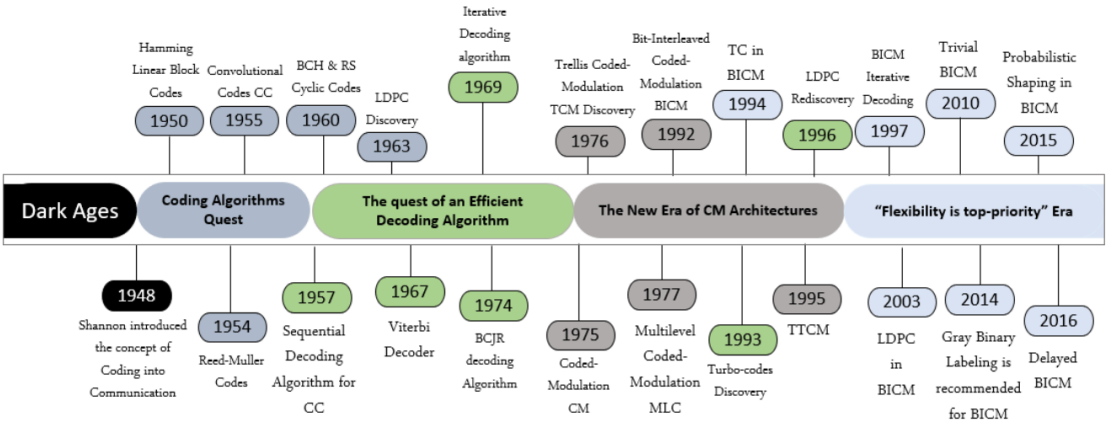


Figure 2.3: FEC Evolution Timeline

2.1.1 Dark ages - before FEC

Earlier to FEC, increasing the rate at which data was transmitted while maintaining reliability was thought to be limited only by signal power, channel bandwidth and modulation level [1]. However, Shannon in his work [6] showed that the statistical behaviour of the information source plays a significant role and based on this characteristic he defined an upper bound of the achievable transmission rate for bandwidth and power-limited AWGN channels. He even went beyond that

stage by suggesting how to manipulate the statistical characteristics of the information source in a way that increases the mutual information. This manipulation was nothing else but channel coding, or FEC. He showed that by applying random coding he could achieve maximum utilization of the channel without reliability loss. Shannon, however, did not manage to demonstrate any practical algorithmic implementation of his coding work. Therefore, the research endeavour of the following decade got directed towards establishing one.

2.1.2 The quest of a coding scheme

During the 50s and 60s, various coding strategies were invented. These differed in the way parity bits were generated. Hence, a natural differentiation in the performance and the implementation complexity occurred.

The coding scheme categories include Linear Block Code LBC, and Convolutional Code CC. The Cyclic Code, which is quite popular nowadays is a subcategory of LBC. The LBC deals with the information bits in finite uniform blocks. It uses a generator matrix formed by linear relationships between the information bits to generate the parity bits [1]. On the other hand, CC uses shift registers to enforce state-dependent relationships, forming a finite-state Markov process [2]. As a result, the output of the convolutional encoder is dependent on the input information bit as well as the state of the registers resulting from the preceding bits [1]. The first invented coding scheme was of LBC nature, developed by Richard Hamming and called after him [8]. Hamming codes are simple to implement but were limited in their error correction ability to a single error. The single error correction limitation did not last long, as Reed-Muller coding scheme was established a couple of years later to tackle this issue [9]. After that, CC saw the light, upon its introduction in [10]. CC gained great acceptance, as it exhibited simpler encoding requirements relative to LBC [2]. Also, the decoding complexity had reduced with the introduction of trellis representation for codes [2].

Later on, another encoding breakthrough emerged; the discovery of the Cyclic Codes. Originally they were introduced as BCH codes [11] and Reed-Solomon RS [12] codes. Eventually, Low-Density Parity Check LDPC Code also emerged in Gallager's master's thesis [13]. However, the LDPC was more computationally demanding than what technology could support at that time. As a result, it was forgotten for 3 decades, before being rediscovered by Dan Spielman in his thesis [14]. That justifies the initial lack of interest in LDPC codes right after their inven-

tion. However, nowadays LDPC is among the best-performing coding schemes [2]. LDPC is the main subject of this thesis, therefore it will be described in more detail in the next chapter.

Another coding scheme that came into the arena was Turbo code (TC). TC was discovered a couple of years before the rediscovery of LDPC. A TC encoder consists of two concatenated encoders of convolutional type in parallel, separated by a pseudorandom interleaver. The input of the encoder is a sequence of the information bits, while the output is a combination of the original sequence, a sequence of parity bits generated from the first convolutional encoder, and another sequence of parity bits generated from the interleaved version of the same sequence of bits. Scientific literature points out that TC is a strong competitor in the performance race among FECs [2].

2.1.3 The pursuit of an efficient decoding algorithm

Decoding reverses the effect of encoding. However, despite its complementary operation, it is characterized by a significantly higher complexity. This has driven researchers to focus on encoding schemes that can be easily decoded. They understood that a well-encoded system should have a minimal-complex decoding process possible. That explains the interference between the encoding and decoding eras in 2.3. Thus, reducing the complexity of the decoding process while maintaining high performance was one of the main targets in FEC research.

Decoding techniques can be classified into either hard or soft, or into optimum and sub-optimum. Hard decoding, is the traditional method for decoding, which corresponds to using binary sharp estimates as inputs to the decoder i.e. 1 and 0. Soft decoding, on the other hand, uses soft values that carry the reliability information as well [2]. Soft decoding is the state of the art in decoding as it achieves better performance, despite its complexity.

Ideally, decoding can be achieved with a Maximum Likelihood ML operation that obtains the most probable codeword in the codebook space. However, this procedure is computationally very extensive, as it includes enumerating all the codebook's codewords to check the closest one to the received distorted codeword. The ML decoder formal mathematical equation is represented as [15]:

$$\underline{\hat{x}}^{ml} \triangleq argmax\{\log p_{Y|X}(y[n]|x[n])\} \tag{2.2}$$

The right side of the equation is the vector form of all the symbols consisting of a single codeword. Y and X are the random variables representing the channel's output and input, respectively. Since Equation 2.2 is characterized by an exponential complexity, other sub-optimal methods representing reasonable approximation have been invented [1]. Those algorithms compromise the performance for reduced complexity. A good example is the bit-wise iterative decoding that repetitively occurs in the literature and the recent FEC implementations. Furthermore, by observing the development timeline in Figure 2.3, we can see that sequential decoding was the first decoding algorithm. It was a tree-search-based algorithm, that became very popular because of its implementation simplicity and its pipelined structure, which was vital for a good performance [2].

In 1967, the remarkable invention of the decoding algorithm known today as Viterbi Algorithm VA emerged [16]. VA is a maximum-likelihood optimal decoding algorithm developed specifically for convolutional codes. Its moderate complexity and high performance, contributed to its acceptance in many applications [1]. Afterwards, in 1969, Berlekamp–Massey algorithm was published in [17] to become for a decade the standard decoding algorithm. It decoded BCH codes and exhibited high computational efficiency. It was developed based on obtaining the shortest linear feedback shift register needed to attain the target sequence [2]. Last but not least, this era concluded with the development of BCJR algorithm [18]. Unlike VA, BCJR is dependent on the Posteriori Probability APP in its estimation. It mandates an iterative implementation to enable the exchange of data, and as a result, it is characterized by a more complex implementation [2]. The complexity issue hindered the ability of this algorithm to replace VA. However, it most importantly contributed to introducing the concept of Soft-Input-Soft-Output SISO. SISO system facilitates the exchange of prior knowledge to future iterations in a Maximum A posteriori MAP decoder. The SISO idea, played a major role in the implementation of iterative turbo decoding [2].

2.1.4 Coded-Modulation CM Architecture

The Coded Modulation era was based on the idea to build efficient systems by integrating coding and modulation altogether. It was the next obvious step after the maturity coding schemes had reached at that time, leveraging also the computational developments of the current hardware. The first CM system was a Trellis Coded Modulation TCM in [19]. It mainly relied on coupling a convolutional encoder with a high-level constellation mapper, working closely together to reduce

the error probability. Multilevel Coded Modulation MLC followed up in the famous paper of Imai et al. [20]. The principal idea of MLC was about offering multistage coding through cascading several independent encoders and decoders, combined with modulation. Although the latter CM schemes could achieve a very good performance, they had the disadvantage of a fixed and complex design, which forced scientists to start striving for more adaptable design solutions.

The target was soon achieved with the Bit Interleaved Coded Modulation BICM system, which was introduced in 1992. BICM systems caught the attention of researchers for decades afterwards and marked the current era of channel encoding.

2.1.5 The era of flexible and adaptive systems

The motivation behind BICM introduction was to enforce adaptable system design by decoupling processing functions, such as modulation and coding. The BICM was introduced by [21]. It normally consists of a binary encoder, permutation, and sub-optimal bit-wise decoding, each of them being an independent modular block. Bit level encoding is facilitated by an interleaver separating the blocks of the encoder and the modulator. Its modular structure allows for amending the configuration of each block without affecting the others, enabling simpler designs with less complexity, while maintaining a reasonably good performance [15]. Later on, BICM underwent many developments, as it was easier to experiment with different variants. For instance, research took place to compare the performance of TTCM (Turbo codes in TCM) and BICM with Turbo codes [22], showing that there are cases where TTCM outperforms. However, the over-performance of TTCM over BICM, [15], does not necessarily mean that TTCM should be preferred over BICM. Firstly because there are still scenarios where BICM maintains performance superiority, e.g when increasing the coding rate of the encoder [23]. Secondly, the inherent flexibility of BICM simplifies the parameter tuning procedure allowing an easier search for best-performing configurations. That's how the researchers' attention was shifted towards developing variants of BICM.

Among the variants of BICM that had seen the light is; BICM with LDPC [24]. BICM originally constitutes a bit-interleaving block, however, [24] demonstrates that the scheme where LDPC is used as an encoder does not require an interleaver. The implicit random nature of LDPC enforces the statistical independence of the bits [24]. Thus, the computational complexity is reduced. Observations

like this paved the way for the introduction of BICM-Trivial or BICMT [?], which is based on the usage of a trivial interleaver, that does no interleaving whatsoever.

Another variant of BICM proved to achieve an enhancement in the performance, by including the state-of-the-art input distribution shaping [25]. The particularly used input distribution shaping technique in [25] is Probabilistic Shaping PS. Essentially, PS observes the input binary stream and tries to correspondingly invent an “alphabetical language” by grouping bits. The number of bits to be grouped is found by a search process for the arrangement that achieves an alphabetic frequency distribution equivalent to the design probabilistic distribution, which is usually Gaussian. The hatch is that the most frequent alphabets are assigned to the low-power signals on the constellation graph. An illustrated explanation of the concept can be found in [26]. To date, implementations targeting achieving high performance by integrating probabilistic shaping with BICM are being developed. Discussions have been also going about the key role of bit-labelling proper selection on the performance of BICM schemes. In [27], Gray code binary labelling was proven to optimise the performance of BICM.

When it comes to decoding, BICM has two possible configurations. BICM with Iterative Decoding BICM-ID, or BICM with Bit-wise decoding BICM-BW. The simplest one is the BICM-BW, where the received bits are taken out of their symbol-based assembly and treated independently. The result is not exploiting all the received information, however, that is an acceptable performance-complexity trade-off [28]. The other decoding approach, BICM-ID, introduced in [29], mainly targeted mitigating the loss of information due to the sub-optimal implementation. It creates an information feedback loop between the demodulator and the decoder to exchange information iteratively. Thus, it exploits some of the bit dependency within a single symbol instead of neglecting it as it happens with the BICM-ID scheme. A new configuration of BICM is the Delayed BICM. It uses memory to restore information from past bits relating the current bit [30].

BICM is the state-of-the-art in FEC. Hence it was chosen as the basic configuration for this study. More details about its structure and components are presented in Chapter 3.

2.2 The debate on performance evaluation for an FEC-based system

The accurate evaluation of the communication channel performance is a vital component in the process of designing a communication system. The consequences of inaccurate metric results in a misguided development process. Normally, channel quality is quantified by the bit error rate BER. The BER is defined by the probability of a received bit to not match the transmitted bit, given C coding scheme, M modulation method, and $P_{Y|X}$ channel, as $Pr(\hat{b} \neq b|C, M, P_{Y|X})$. In Monte Carlo-based numerical simulations, the number of bits required to measure a h BER, needs to be at least 100/h to achieve a less than 10% error in our results [31].

For a high-capacity transmission link, such as a long-haul fibre channel, the typical requirement for the BER after decoding BER_{post} is to be as low as 10^{-15} [32], [33]. Such a low BER imposes the need for at least 10^{17} bits to achieve a 10% error margin. This massive number of bits makes very difficult the practical performance evaluation of such systems. Thus, design simulations and experiments of high-capacity links exclude implementing channel coding and replacing it with the so-called FEC-Threshold examination.

As Figure 2.4 demonstrates, the FEC threshold examination consists of three steps. First, deploying the communication system without including channel coding. Then, measuring a metric β that reflects the performance quality of the transmission. Finally, β is compared against a pre-defined threshold θ, at which if it exceeded it, then the desired BER_{post} e.g. 10^{-15} is guaranteed to be achieved with an arbitrary coding channel.

Figure 2.4: FEC-Threshold Definition

This workaround exempts researchers from including FEC in their system studies and running extensive simulations. However, it also brings up several important research questions. The first question is; what could θ, the required threshold, be? researches such as [34, 35], and [36] addressed the question by perform-

ing the necessary extensive simulations for certain choices of channels, modulation format, and symbol rate, and concluded their research with a ready-to-use BER_{pre} threshold requirement for achieving the desired BER_{post}. Another question is about the appropriate performance metric that can accurately and reliably resemble the performance of the communication system between the FEC components at both ends. Metrics that exhibit dependency on the type of channel, or the modulation scheme need to be identified and disqualified.

A prevalent metric is the BER before decoding, i.e. BER_{pre}. Using BER_{pre} perfectly suited the popular usage of hard decoders and when soft decoders emerged, it was the default metric to be used. However, this notion was critically criticised and proven to be unreliable in the work conducted by A. Alvarado et al. [33]. The authors performed two types of simulation experiments. One where they deployed FEC and measured the performance. The other is where they followed the notion of using FEC-threshold, by measuring several types of metrics β's (refer to Figure 2.4. The surveyed β's include BER_{pre}, as well as information-theoretic values; mutual information MI and general mutual information GMI. The GMI is a statistical measure for the dependency between the channel observation and its input, but contrary to MI, it measures the dependency at a bit level rather than a symbol level (the modulated system is assumed) [1]. The reported results in [33] had proven that using BER_{pre} as a β metric does work only in the cases where FEC schemes with hard decoding are assumed. When soft decoding is adopted, an accurate β would be MI. Furthermore, GMI would be the best β metric in the case of BICM system, as it is based on a bit-level processing [33].

2.3 The state-of-the-art in FEC for fibre channel

According to [37], the main FEC schemes for fibre optical channel are; staircase codes (SC) [38], Quasi-Cyclic LDPC (QC-LDPC) [39], Spatially-Coded LDPC (SC-LDPC) [40], Soft-Decision Turbo Product-Code (SD-TPC). [41]. Each one of them varies in its setup, it has its own advantages and disadvantages, but all manage to reduce the BER to $10^{-}15$ [32]. The FEC schemes listed above differ in various ways. Their coding structure is one of them. Some of them consist of a single encoder and a decoder in the transmitter and the receiver, respectively, whereas others are a concatenation of encoding and decoding elements. A representative example is a transmitter, where we could have two encoders of different types, inner and outer. While at the receiver, there would be a compat-

ible concatenation of two decoders, addressing sequentially the correspondent codeword frame. A concatenated scheme holds higher complexity than a single ex-coder configuration [39], it is a necessity for addressing error-floor phenomena. The error floor occurs in scenarios where the BER cannot drop below a certain level, e.g. see Figure 2.5. Concatenated FEC schemes have been used in various works, [42, 25, 43] to name few.

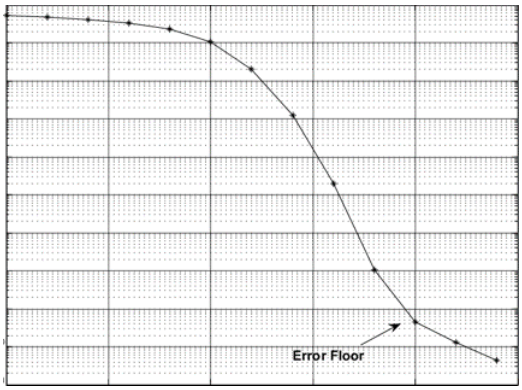


Figure 2.5: The Issue of Error Floor as Low BER is Obtained - BER at the y-axis and Eb/No (dB) at the x-axis (Adapted from [2])

The decoding strategy is another point FEC schemes might differ. There are usually two decoding techniques; either soft decoding SD or hard decoding HD. HD, which is the simplest form of decoding, uses binary values as input for the decoder and performs its operation upon them. The moderate complexity of this technique made it find its way to commercial applications. SD, on the other hand, deals with the soft values of the bits. A bit's soft value consists of its binary value, 0 or 1, and the certainty of the belief about the selected binary value. The number of bits required to represent a soft value is apparently more than the hard one. A single bit of information would be represented by more than one bit, hence, the complexity of SD increases [1]. Next generation FEC of fibre channel utilises SD, due to its enhanced performance advantage [44], and the revolution the hardware sector is experiencing [39] in terms of computational power. Eventually, a differentiation metric that can be used over the various FEC schemes, is the overall complexity. Addressing complexity requires a thorough consideration of different aspects of the scheme, including code construction, type of constraints imposed, codeword length, decoding algorithm, and decoding quantisation [37]. Another differentiation metric that can be used to compare FEC schemes is the overhead OH. The OH is a representation of the ratio of the parity bits counts to the original number of information bits. $OH\%$ can be calculated using the coding rate R

through:

$$OH\% = \frac{1}{R_T} - 1 \tag{2.3}$$

where R_T is the total coding rate of the concatenation if existed. It is the result of multiplying each concatenated scheme's coding rate; $R_T = R_{En_1} R_{En_2}$

A work that benchmarked several proprietary FECs currently available in the market for long-haul optical fibre channels, based on the OH is [3]. Figure 2.6 shows the OH% added by each FEC scheme with the developing company name. The OH% requirement ranged between 15% and 35%, which represents the state-of-the-art acheivement in terms of OH% requirement.

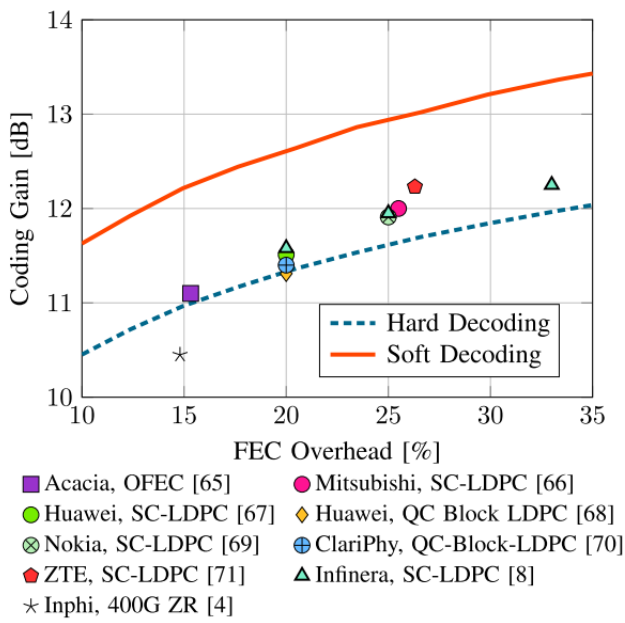


Figure 2.6: The overhead cost of various FEC systems developed for the fibre channel. [Adopted from [3]]

3. PROCEDURE

The goal of this study is to test the performance influence of DBP compensation over fibre channel when integrated with coded-modulation. Therefore, we developed a simulation model of a coded-modulated transmission system, which uses the nonlinear fibre channel as a communication medium. We added DBP compensation to the model, and checked the prior and post performance in terms of the bit error rate BER. This chapter discusses the design choices and their theoretical principles in the following sequence:

- The coded-modulation CM architecture choice.
- The coding used in the CM architecture, from encoding point of view then decoding.
- The methodology at which the information propagates in the channel and how information is extracted from the channel observations i.e. mapping and demapping.
- Interleaving Implementation; a function part of the chosen CM architecture.
- Fibre channel definition and a description of the DBP implementation for nonlinearities compensation.

3.1 Coded-Modulation CM Architecture

There are several potential candidates that could have been used as CM architecture for our model. The chosen one is the state-of-the-art in coding in general and fibre FEC implementations in particular, bit interleaved coded modulation BICM [15]. BICM architecture consists of a concatenation of encoder and mapper which independently functions, and usually are separated by a permuting function that is called interleaver. Figure 3.1 demonstrates the building blocks of a BICM architecture. As previously mentioned in Chapter 2, the advantage of using BICM is its support for a bit-level processing, which makes it a sub-optimal but also affordable and practical implementation. The aforementioned feature, implicitly facilitates another advantage, which is the ability to choose the encoder and mapper specifications independently from each other. This simplifies the implementation further more and allows an adaptive and configurable system that adjust itself as per the channel conditions, which consequently enhances the capacity utilisation. Those key advantages facilitated the increased interest in BICM at which is was adapted in wide range of practical implementations of communication systems, such as DVB-S2 for satellite communication, wireless LANs, DSL, WiMax, and in the evolutionary cellular systems; LTE and 5G-NR [28, 45].

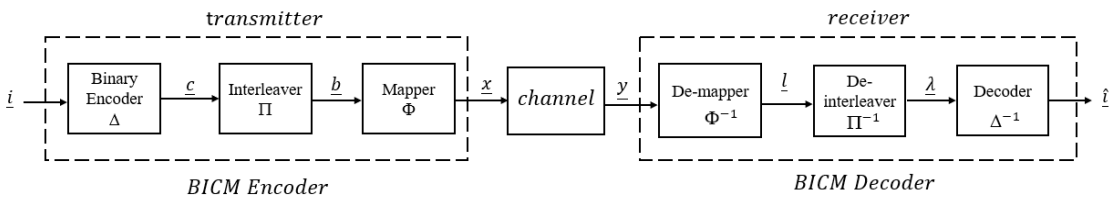


Figure 3.1: Communication Block Diagram Based on BICM Configuration

3.2 The encoding scheme

The BER requirement for a high-capacity channel as fibre channel, is very low; 10^{-15} [32]. Coding schemes that guarantee such low performance levels are mostly concatenations of two encoders and decoders [46]. This is due to the error floor phenomena discussed in Chapter 2. The concatenation of an inner soft code and outer hard code had proven to be a successful workaround this problem. Accordingly, we choose for our implementation of encoders, a pair of an outer En_1 and an inner En_2 encoders. At the the outer encoder, staircase codes based on the published work in [47] is chosen. Staircase code is commonly used in fibre channel's FEC implementation, due to its ability to correct large number of errors and reduce the BER to very low levels, i.e. 10^{-15}. Figure 3.2 demonstrates the described concatenation. The interleaver separating both encoders, assures that decorrelation of the errors that take place at the inner decoding process at the receiving end, to mitigate their effect on the outer decoder [33].

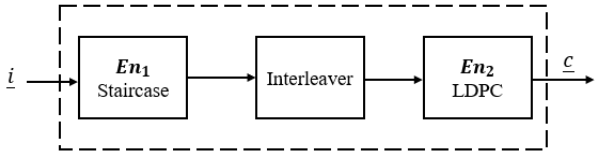


Figure 3.2: The Binary Encoder Concatenation

The encoder performs coding by adding a number of parity bits, P, to a sequence of information bits of length K. The result of adding those two, the parity bits and the information bits, generate the codeword, which hold the value $N = P + K$ as its length. Two essential parameters that are commonly used to characterise any coding scheme is the code-length, N, and the coding rate, R. The coding rate is the ratio of the information bits count to the code length, where higher coding rates correspond to a smaller overhead but weaker error correction capability. The equation that defines overhead as a function of coding rate is presented in Chapter 2.

Table 3.1: Description of the Examined Inner Encoders

	Examined Inner Encoders	
	Encoder-a: DVB code	Encoder-b: QC code
Type	Irregular LDPC	Quasi-Cyclic LDPC
Usage	Digital Video Broadcasting DVB Satellite Communication	IEEE 802.11 WiFi
Codeword Length(s)	64,800	1944 1296 648
Supported Coderates	1/4 1/3 2/5 1/2 3/5 2/3 3/4 4/5 5/6 8/9 9/10	1/2 2/3 3/4 5/6
Standardizing Organization	European Telecommunications Standards Institute - ETSI	IEEE

The chosen staircase code comes in two different coding rates, 0.9412 and 0.75 corresponding to 6.25% and 33.33% of overhead, respectively. Our ability to choose any of them is highly dependent on the attained performance from inner coder. If the inner decoder managed to decrease the BER to the FEC threshold level defined for 6.25% staircase code as 4.7×10^{-3}, then this code is guaranteed to generate a BER as low as 10^{-15}. The other staircase code with the 33.33% OH requirement has a more modest BER FEC threshold of 2.12×10^{-2}. The usage of FEC threshold paradigm to predict the performance post outer decoding, is justified by the hard nature of the staircase decoder. Refer to Chapter 2 for more details.

3.2.1 The inner code choice

Based on the earlier discussion, a good choice of the inner code to a critical enabler for achieving the required performance at the lowest possible OH cost at the staircase code. Consequently, two inner code candidates were examined and assessed comparatively to subsequently choose the one best performing. The

Table 3.2: The Investigated Configurations for each Inner Coder

Configuration	DVB (OH%)	QC (OH%)
Config-1	R = 3/4 (33.33%)	R = 3/4, N = 648, Z = 27 (33.33%)
Config-2	R = 5/6 (20%)	R = 3/4, N = 1944, Z = 81 (33.33%)
Config-3	R = 9/10 (11.11%)	R = 5/6, N = 648, Z = 27 (20%)
Config-4		R = 5/6, N = 1944, Z = 81 (20%)

investigated codes; irregular LDPC adopted from DVB-S2 standard developed for satellite communication [48], and the Wifi protocol's quasi-cyclic QC-LDPC from IEEE 802.11 protocol [49]. Table 3.1, describes the important characteristics of each coding candidate. The DVB code has a very long single codeword length, while the QC code supports three different code-lengths. Both codes have a range of supported coding rates. Although both codes intersect in most supported coding rates, DVB code tends to support higher coding rates.
To investigate the QC and DVB codes in an appropriate comparative form, certain identical coding rates were chosen from each protocol, 5/6 and 3/4. In the case of DVB, 9/10 was also picked in order to examine the system with the highest possible coding rate. While DVB code comes in single code-length, QC supports three different length, two at which were chosen to allow better characterisation of the scheme; 648 and 1944. The details of the studied configurations is shown in Table 3.2. DVB has three configurations which are defined as per the coding rate, while the QC uses four configurations with different coding rate and length. The definition of each code scheme is made through the parity check matrix PCM.

3.2.2 Parity check matrix - a code definer

The parity check matrix PCM, is a $P \times N$ binary matrix that represents all the parity equations by marking the participant bits by 1. Figure 3.3 demonstrates the method at which PCM is generated. They start from the parity check equations c_j which define the linear algebraic relations that gather the bits b_i. In the example shown in Figure 3.3, a total of 3 parity bits exist, which justifies the rows count the PCM, H i.e. the first 3 columns resembling $b_1, b_2, \& b_3$.

Parity check matrix of Irregular LDPC from DVB-S2 standard

Being an LDPC based code, the H-matrix is sparse. The fact that it belongs to irregular LDPC suggests that the columns and rows weights are not consistent (the weight means the number of ones in a row or column). Although it was optimised to be used for satellite communication, it has found applications in various fields, due to its good performance and wide off-the-shelf availability. Authors in [33],

$$c_1 : b_1 \oplus b_4 \oplus b_5 \oplus b_7 = 0$$

$$c_2 : b_2 \oplus b_4 \oplus b_6 \oplus b_7 = 0 \longrightarrow \mathbf{H} = \begin{pmatrix} 1001101 \\ 0101011 \\ 0010111 \end{pmatrix}$$

$$c_3 : b_3 \oplus b_5 \oplus b_6 \oplus b_7 = 0$$

Parity Equations Parity Check Matrix

Figure 3.3: Parity Check Matrix Formation from the Parity Equations. The $\oplus$ operator is modulo 2 addition.[Adopted from [4] and edited]

for example, had used it in their experimental setup for examining FEC on fibre channel. The PCM has 64800 columns as a result of the long codeword length it supports. As it supports various coding rates, each coding rate has its own parity check matrix that can be found in the standard definition appendix [48].

Parity check matrix of Quasi-Cyclic QC LDPC from WiFi protocol

The H-matrix for QC-LDPC is distinguished for its cyclic nature and systematic reformation. The matrix defined in [49] is specified in terms of three pieces of information; the base matrix, the expansion factor Z, and the codeword length. The based matrix is represented by a matrix of numbers ranging from -1 to Z. Those numerals are then replaced by an identity matrix of size Z or a version of it. The version of the identity matrix is dependent on the particular number it is replacing in the base matrix. For example, if it is replacing a 1, then it corresponds to an identity matrix that is circularly shifted to the right by 1, and the logic is the same for all other numbers. The number 0 and -1 might be less intuitive, as 0 corresponds to a straight identity matrix, and -1 corresponds to a matrix of zeros. The typical length for codewords generated by this protocol is relatively small, which results a smaller PCMs.

3.2.3 Encoding Algorithm

Encoding takes place following a certain procedure that starts with the coding scheme's parity check matrix, which are usually used to define coding schemes. PCM can be defined as:

$$H_{(N-K\times N)} = [-P^T | I_{N-K}] \tag{3.1}$$

where I_{N-K} is identity matrix of size $(N-K \times N-K)$. The reformulation of H, performed with Gaussian elimination, results an important element that makes

up the Generator matrix G, P matrix. Consequently, the generator matrix that is required to generate the codewords, is defined based on the obtained P as:

$$G_{(K \times N)} = [I_K \mid P] \tag{3.2}$$

By having the generator matrix, the encoder that generates codewords C becomes a mere matrix multiplication between the information bits sequence, of size K and the generator matrix, as in Equation 3.3.

$$C_{1 \times N} = x_{1 \times K} \, G_{K \times N} \tag{3.3}$$

Equation 3.3 validates the increase of the bits sequence length from K to N upon performing coding, at which the extra parity bits are utilised in decoding.

3.3 The Decoding strategy

The decoder is responsible for utilising the overhead in the coded data to correct the errors the message experiences upon the transmission. The choice of the decoding algorithm is made to mainly address the complexity concern, as decoding is the most expensive element in the coding process. Hence, we favour the sub-optimal algorithm, min-sum which operates on soft input values to enable a state-of-the-art implementation of soft iterative decoding. This decoder collects the information carried by each bit in the codeword concerning each and every message bits in an iterative message passing manner. Properly designed coding schemes performs better as the number of allowed iterations increases, at an expense of increased complexity. The min-sum algorithm is a variant of the Belief Propagation BP algorithm that offers a less complex solution at an acceptable performance loss. Decoding Algorithms are usually explained on a type of graphs that facilitate a better visualisation of the message passing between code bits.

3.3.1 Tanner Graph - A visualisation tool for decoding

There are several ways to present codes other than parity check matrices; one graphical way is Tanner graph i.e. bipartite graph. It is a convenient method for visualising the relationship between bits, and the information exchange between code bits during the decoding process. Figure 3.4 demonstrates several representations of an exemplary dummy code, starting from primitive parity check equations, PCM, and Tanner graph.

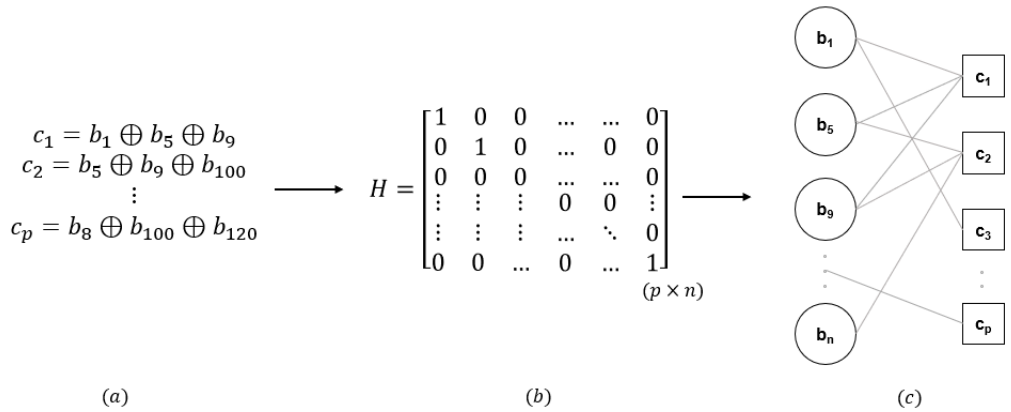


Figure 3.4: Different Representations of a channel Code. (a) Algebraic Parity Equations. (b) Parity Check Matrix. (c) Tanner Graph

In Figure 3.4(c), information relating the parity bit equations is assembled to the right and the code bits to the left. The b_i's are called variable nodes while the c_j's are check nodes. The equations shown in 3.4(a) define the edges between the check nodes and their participant variable nodes. If accurate channel observations were received about b_1, b_5, and b_9 then the substitution in c_1 will results zero. This acts as an evidence for the correct transmission of the participant variable bits. The aim of the soft iterative decoder is to gradually accumulate such beliefs about each message bit and form a decision about its value.
The nodes shown in Figure 3.4(c) are loaded with soft values about the binary bits generated by the mapper from the channel observations. A soft value is a signed number that holds the bit's hard value (whether it is believed to be 1 or 0) in their sign, where 1 corrsponds to positive and 0 corresponds to negative sign. The absolute value represents the certainty level i.e. confidence of the belief. Details concerning the generation of the soft value will be discussed in the context of demapping. These soft values feed the min-sum algorithm

3.3.2 The Min-sum Algorithm

At the algorithm's initial state, every variable node holds its intrinsic soft L-value which it receives from the demapper. The final state, on the other hand, is an accumulation of as many as possible of extrinsic L-values concerning each message bit, which are collected within the permitted decoding iterations number. Starting from the initial state, the dynamics of the algorithm are as follows [4, 50]:

1. Each variable node sends its intrinsic L-value to all the check nodes it is connected to.

2. Check nodes uses the received information and generates a belief concerning each variable bit by performing the check-equations. The processed values are not binary values, thus, modulo-2 addition cannot be simply performed over beliefs. Rather, single parity check SPC decoding is performed, which as defined as following. Suppose the following check node equation in the hard-value space:

$$Z = X_1 \oplus X_2 \oplus X_3 \tag{3.4}$$

where the X's hold binary values and $\oplus$ represents a modulo-2 addition. When transforming Equation ?? to the soft-value space, where the X's hold soft values, l_{X_k}:

$$sgn(l_Z) = sgn(l_{X_1})\ sgn(l_{X_2})\ sgn(l_{X_3}) \tag{3.5}$$

$$|l_Z| = abs(f[f(l_{X_1})\ + f(l_{X_2})\ + f(l_{X_3}])]) \tag{3.6}$$

where $f(x) = log(tanh\frac{|x|}{2})$ The functions shown in Equations 3.5 and 3.6 represent the exact implementation of the original algorithm, BP. Min-sum, however, is an approximated variant that uses an approximation for computing the function in Equation 3.6;

$$|l_Z| = min\,(l_{X_1}, l_{X_2}, l_{X_3}) \tag{3.7}$$

The complete mathematical proof for Equation 3.5 and 3.6 is presented in [51]. In both implementations, When the check node computes the extrinsic L-value for a variable node, b_1 for example in check node c_1, it excludes b_1's intrinsic value. This exclusion assures independence between observations over iterations for the sake of an unbiased propagation of the belief.

3. The check node passes to each variable node extrinsic L-value it measure about them based on the information received from other variable nodes participating in the equations.

4. Variable nodes receive from different check nodes what the other check equation's participants say about its value. The variable node at this stage acts as a repetitive code decoder, as the received information from different check nodes are statistically independent. Therefore, it simply adds the received extrinsic L-values to its intrinsic L-value.

The steps 1-4 represent a single decoding iteration, which can be repeated as per the defined number of allowed iterations I_{max}. The selection of the maximum number of iterations allowed enhances the quality of the belief, but comes at a computational complexity and energy consumption costs. The iterative soft decoding algorithm, min-sum, is dependent on the soft values generated by the demapper based on the received channel's observations.

3.4 Mapping & Demapping

The mapper and demapper perform modulation and demodulation to assure an efficient utilisation of the available channel while maintaining the BER level within the transmission reliability definition. Another aspect they also responsible for, is the bits sequences assignment to constellation's complex signals while maintaining the least error probability.

3.4.1 Mapper

The mapper shown in Figure 3.1 receives interleaved coded binary bits and divides them into symbols i.e. sequences of m long bits. The m value is dependent on the multi-level modulation format choice, which increases the spectral efficiency the higher it gets, but at requirement of higher SNR. Our modulation format choice is the 256-QAM, which offers a high level of spectral efficiency, 8 bits/symbol/Hz. It is also offers an adequately challenging scenario for our coded-modulation system that enables a better characterisation and assessment opportunities [1]. After chunking the bits stream into 8 bits long symbols, the mapper assigns the binary sequences to the constellation's 256 complex signals which have differentiable amplitude and phase characteristics. Labelling, is another responsibility mapper which is mainly concerned with mapping the binary sequences to the complex signals. Many different labelling strategies exist [52], including binary and Gray. For BICM architecture, studies had show that Gray mapping attains the best performance [53, 15], therefore, it we chose it for our implementation. The specifications of the mapper, including the constellation format and labelling, is a shared information with the demapper which reverses the mapper's work.

3.4.2 Demapper

The main task of the demapper is to perform demodulation to recover the message bits from the distorted received channel's observation. There are two methods of demapping, hard demapping and soft demapping. The conventional way, hard demapping, is the simplest and least computationally expensive technique which is a practical option for demapping implementation. It works by defining equidistant thresholds that separates the constellation's symbols and then classifying the received signal according to the area it lays on. An illustration of the hard demapper's thresholds is shown in Figure 3.5 for a 4-QAM constellation. As the figure demonstrates, the four symbols have their own areas (each shaded with a different colour) that are defined by the purple threshold lines that are parallel to the constellation's axes. The simplicity of this technique is not only due tho the easy definition of those thresholds, but also because of the delivery of the estimation which is exactly m-bits long. The hard demapper does not provide any confidence information about its estimate, which means that its perception to a measurement recorded close to the constellation point, is equivalent to another one recorded close to the threshold line. This information omit, is a disadvantage of this technique, and restricts its performance level.

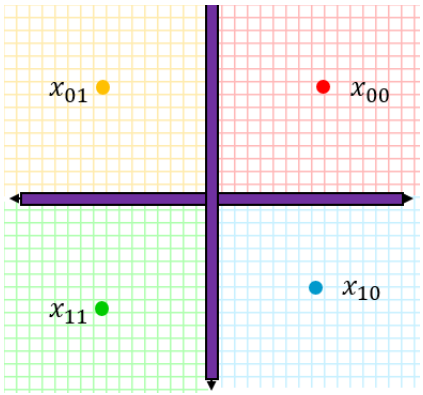


Figure 3.5: Hard demapping threshold's for a 4-QAM constellation

Soft demapper, addresses the shortcomings of the hard demapper by providing information related to the confidence too, at a complexity cost. Soft demapper differs from hard demapper by three main points. First, it does not define the separating lines between symbols using thresholds, rather it depends on a probabilistic approach. Second, it reports its estimation's confidence. Third, it requires at least one extra bit to represent besides the hard value's m-bits the confidence. The topic of quantising the soft values is out of this research scope, but interested readers can refer to [54].

The high performance of soft demapper combined with soft iterative decoding made it be the state-of-the-art in FEC, especially with the receiver's developing computational capabilities [15]. The soft demapping that was used in this work is based on Logarithmic Likelihood Ratio LLR, which works at a bit level. The bits of the constellation's symbol are processed independently each as per its bit position. While a 4-QAM symbol consists of 2 bits, $b_1 b_2$, is assumed, the algorithm [15] is explained in following the steps:

1. To start with b_1's value, the whole constellation's symbols are divided into two groups, Group-0 and Group-1. Group-0 consists of the symbols identify 0 at their first bit position. Group-1 contains the rest of the symbols which, on the contrary, has the first bit position equal to 1. Refer to Figure 3.6(a), to visualise the Gray-mapped 4-QAM constellation being divided as per the first bit position. The upper intensively blue shaded area represents Group-0, while the lower one is Group-1 for the first bit position. The forming threshold based on this is parallel to the horizontal axis.

2. The Euclidean distance ED separating the received signal from each constellation point is calculated and then fed into the assumed channel's Gaussian probability density function PDF, characterised by a zero mean and an arbitrary variance σ. The prior knowledge of σ is not required as it would be cancelled out in a later stage. This step results the probability of the received signal being certain symbol, given the channel observation based on the ED. There are various assumption introduced in this step, including that the channel noise is Gaussian, and a uniform type of distortion is experienced among all the symbols and the bit's positions.

3. The probabilities computed at step 3 are added up for each group following the classification logic in step 2. The result of this, is a summation of the probabilities for b_1 being 1, and another summation of it being 0.

4. The logarithm of the ratio is computed and referred to as, Logarithm-Likelihood-ratio LLR or the L-value using the information collected in step 3.

   The mathematical equation representing LLR estimation for the 2nd bit position in symbol n is [15]:

$$l_{b_1}[n] = log_{10} \frac{\sum_{i \in I_{b_1=0}} \; p_{Y|x}(y|x_i) \;\; P_{X|b_1}(x_i|0)}{\sum_{i \in I_{b_1=1}} \; p_{Y|x}(y|x_i) \;\; P_{X|b_1}(x_i|1)} \tag{3.8}$$

5. Steps from 1-4 are repeated for the second bit position. Figure 3.6(b), shows the division based on the second bit position in the constellation.

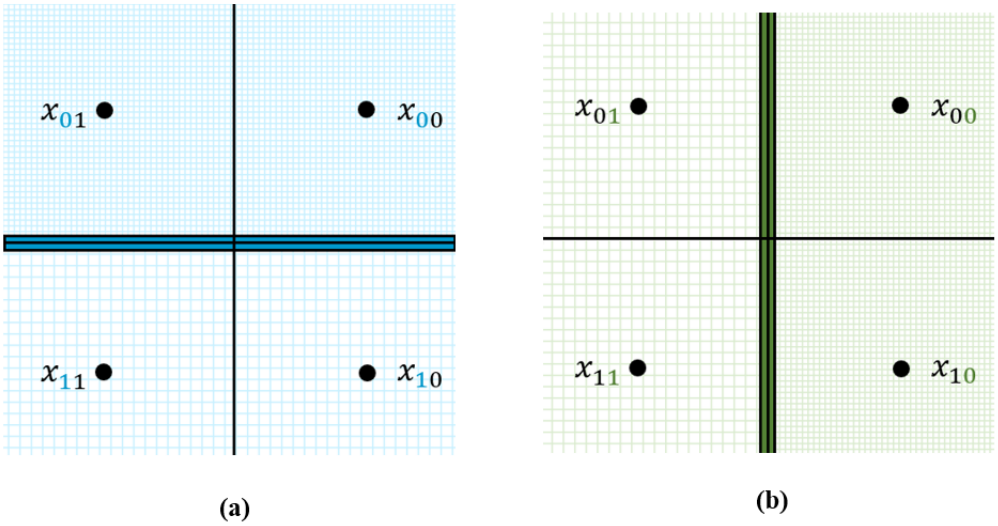


Figure 3.6: Soft demapping threshold's for a Gray labelled 4-QAM constellation based on each bit position

Soft demapping is very dependent on the choice of the labelling scheme, as it directly affects the probability of errors. The thresholds shown in Figure 3.6 are identical to the one generated by the hard demapper shown in Figure 3.5. This is due to the usage of Gray labelling. In case of using another label scheme such as the arbitrary one used in Figure 3.7, the thresholds for each bit position differ.

3.5 Interleaving

Deploying permutation of bits between the encoders concatenation, as well as before mapping, proved to be effective in enhancing the performance [15]. It mainly breaks the implicit statistical dependency between bits belonging to the same symbol or same codeword, which maximizes the efficiency of the decoder that is build upon the assumption of independent bits. There are two architectures of interleavers, S-interleavers (for single), and M-Intereavers (for multiple) [15]. Figure 3.8 shows the architecture of both interleavers. S-interleaver receives p codewords, permutes them, then pass them to the following functional block. M-interleaver on the other hand, uses m parallel interleavers that get parts of several codewords depending on the specifications of the interleaving function. Another parameter that requires optimization in the M-interleaver is number of the parallel interleavers to be used. Due to the simplicity of the S-Interleaver, it was preferred

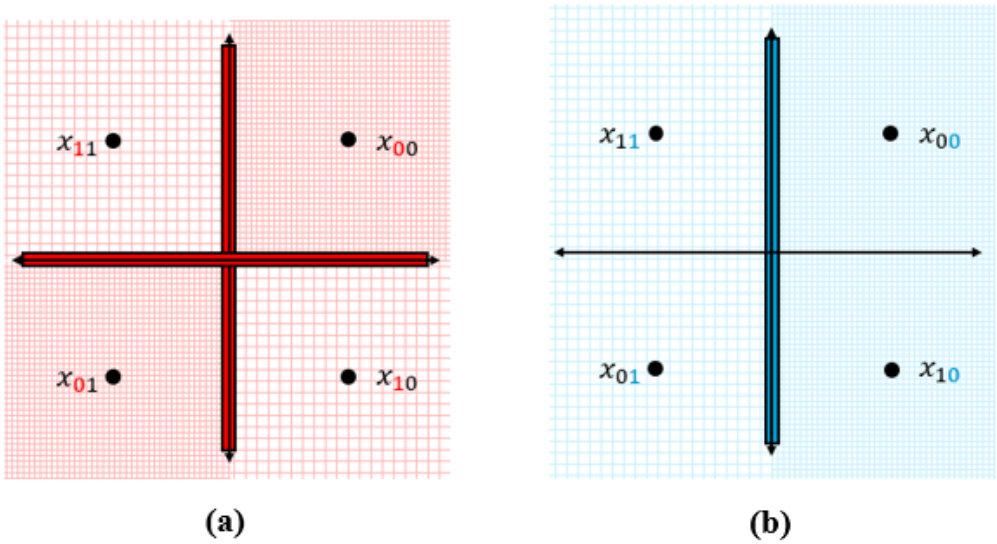


Figure 3.7: Soft demapping threshold's for an arbitrary labelled 4-QAM constellation based on each bit position. (a) Thresholds for first bit position. (b) Thresholds for section bit position.

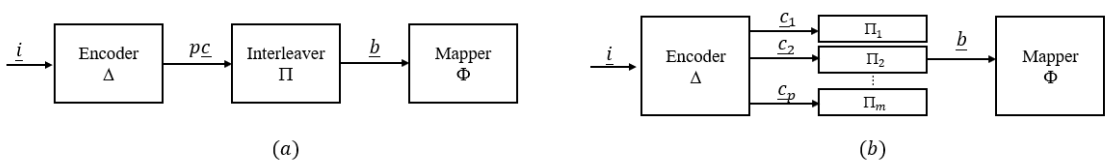


Figure 3.8: Types of Interleavers (a) S-Interleaver, (b) M-Interleaver

over the M-Interleaver. Ideally, interleavers work best with an infinite number of codewords being interleaved. However, due to the impracticality of that, interleaving is optimized to work on the largest possible number of codewords. An experiment was performed to measure the performance enhancement as the number of codewords being permuted increases. The permutation takes place, by buffering the required number of codewords, generating a pseudo-random array of an equivalent size from a specific seed, then reorder the buffered codewords's bits according to those random indices. At the receiving end, undoing the interleaving is performed at the deinterleaver. It has the seed's information that was used to generate the pseudo-random indices' array at the transmitting part, which enables it to recover the correct order of the interleaved bits. Although the interleaver is a major component in the original BICM, its impact on the performance has been questioned and in some work denied [55], and brought forward a new variant called Trivial-BICM. In this work, interleaving's impact would be studied on the

simulated transmission model.
Up to this point, a thorough discussion related to the system and its design elements was conducted. Due to the limited computational capabilities, numerical simulations that measure for a BER as low as 10^{-15} are not possible. Therefore the inner coder is implemented in the simulated model, and further outer coder processing is assumed based on the results reported in [47]. The assumption is valid by assuring the compliance to the FEC-threshold criteria. Therefore, the simulations excluded staircase codes, but it will come back to the picture when the system's total overhead cost is defined. This workaround is a common practice [33] among researchers working on high capacity fibre channel.

3.6 Fibre channel

The FEC subjects of this study were examined on two different states of the Fibre optical Channel, a state includes nonlinearities and another one where they are compensated with DBP.

3.6.1 Nonlinear fibre channel

A recognisable way of representing light propagation in fibre and the material-signal interactions is non-linear Schrodinger equation NLSE (Equation 3.9) [56]. The implemented numerical model simulates signal's propagation in the fibre channel by solving the NLSE using split-step Fourier SSF method [57].

$$\frac{\partial A}{\partial z} + \frac{\partial^2 A}{\partial t^2} + \frac{\alpha}{2} A = i\gamma |A|^2 A \tag{3.9}$$

The simulated fiber channel accepts as an input quadrature amplitude modulated QAM, pulse-shaped signal with a raised cosine filter with a roll-off factor of 0.25. The transmission is tested among a range of launch power, while experiencing chromatic dispersion, attenuation, and nonlinearities due to Kerr effect. Along the propagation channel there inline amplifiers are installed at the end of each span. Those amplifiers boost the signal up, and add an amplification noise to the signal. At the receiver, chromatic dispersion is equalized, the modulated information is recovered by performing matched-filtering. Demodulation takes place, and the transmitted data is received with some error. Table 3.3 lists the specifications related to the fiber channel used in this study's experiments.

Table 3.3: Fiber Link Specifications

Parameter	Value
Baud-rate	60 Gbaud
Span Distance	100 km
Number of Spans	3
Attenuation	0.2 dB/Km
Nonlinearity Coefficient	1.4 $W^{-1}Km^{-1}$
Dispersion Coefficient	17 $psnm^{-1}Km^{-1}$
Amplifier Noise Figure	5 dB

3.6.2 DBP-Compensated Fiber Channel

As for the nonlinear behaviour of the fiber channel being the dominant type of distortion that limits its performance, techniques that compensate this type of impairment thrived. A common technique to do that is Digital Backpropagation DBP technique. DBP is implemented based on SSF method again. However it works in a reversed form in order to compensate the nonlinear effect generated by the fiber channel. The SSF method is an approximation that deals with the nonlinear and linear effect separately [58]. The solution also includes splitting the propagation through a single span to small steps and integrate them to find an approximate solution[58]. While implementing DBP using SSF method, the more steps considered per span, the more accurate the non-linearity approximation, and therefor the mitigation is better. However, that comes at the cost of increased complexity [58]. Accordingly the selection of the number of steps/span in the simulations was conservative. The investigating study that tested coding on DBP-compensated channel used 2, 3, and 4 steps/span only.

4. RESULTS AND ANALYSIS

The results presented in this chapter are obtained assuming the numerical system model and parameters of Chapter 3. We first identify the critical FEC design parameters and perform system-level optimisation targeting a moderate complexity and optimum performance for the lowest achievable OH%. The parameters under consideration include the interleaving size and the maximum number of iterations the iterative decoder permits. In our simulation experiments we have considered two FEC candidates, candidate-1; QC-LDPC from IEEE802.11 protocol, and candidate-2; irregular LDPC from DVB-S2 standard. Based on the obtained results, the optimum design FEC choices are made and used for subsequent system-level studies that investigate the impact of DBP-based non-linearity compensation over the fibre channel. Our goal is to benchmark the performance of those two FEC candidates over the nonlinear fibre channel, to identify their optimum settings and explore operational system margins as a function of the fibre channel length.

4.1 Assessing the interleaving's significance on performance

As discussed in the previous chapters, the basic architecture of BICM involves interleaving between encoders and the mapper. Challenging this design choice, however, was one of the early determinations this research started with. Not only because there is literature that challenges its true importance, [55], but this function comes also at a complexity and power consumption cost. It is always wise to direct the DSP computational power towards the strategies that achieve higher performance gain. For the fibre channel, a critical question that needs to be addressed is the impact of the interleaving size on the overall system performance.

The coded modulation system that has been illustrated in the block diagram of Figure 3.1 is simulated here numerically using MATLAB. The experiment is performed over the fibre channel defined in Section 3.6. The total transmission distance is 300km. The initial simulation experiments were conducted on the transmission channel without assuming non-linearity compensation whatsoever. Nevertheless, we have assumed an ideal compensation of chromatic dispersion effects in the digital domain. Furthermore, we assumed a Nyquist-shaped pulses with a roll-off factor of 0.25, modulated with a 256-QAM scheme and transmitted in a 60 baudrate channel. A high modulation scheme results an 8-bits per symbol which achieves a high spectral efficiency.

As the subject channel has nonlinear characteristics, the simulation experiments

investigated the targeted configurations over a range of launch power levels. Then, measuring for each of them the achieved BER and selecting the minimum. Figure 4.1 demonstrates the behaviour of the performance metric, BER, as a function of the launch power in dBm. The lowest BER is always observed around 0 dBm launch power, whether the system is coded or not.

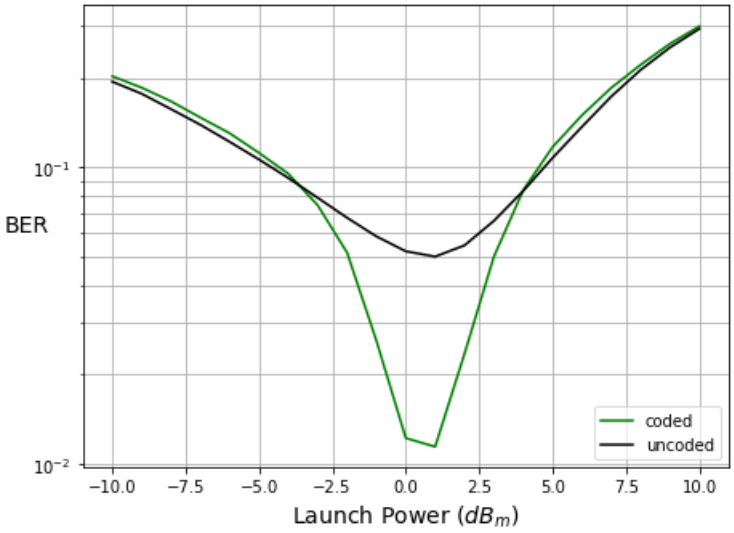


Figure 4.1: Fibre Channel performance measured against launch power in coded and uncoded modes

The configuration above is tested in different scenarios for various interleaver sizes. The random interleaving structure described in Chapter 3 is used. The size of an interleaver indicates the number of codewords that are buffered and permuted together before feeding the mapper. We have examined interleaving sizes of 0, 10, 20, and 30 on the configuration of our two FEC candidates. Both used an iterative decoder, and to achieve a fair comparison a single iteration was allowed. A size-0 interleaver corresponds to no interleaving.

Figure 4.2 demonstrates the BER performance of each FEC candidate, in the various configurations discussed in Chapter 3. For the LDPC DVB-S2 case, we considered three separate coding schemes of different rate. For the QC-LDPC we considered four separate coding schemes, of different coding rate and codeword length. Refer to Table 3.2 for more information about each configuration characteristics. The results show no significant influence of interleaving on the system's BER performance. This applies to uncoded and coded transmissions for both FEC candidates, regardless of the configuration.

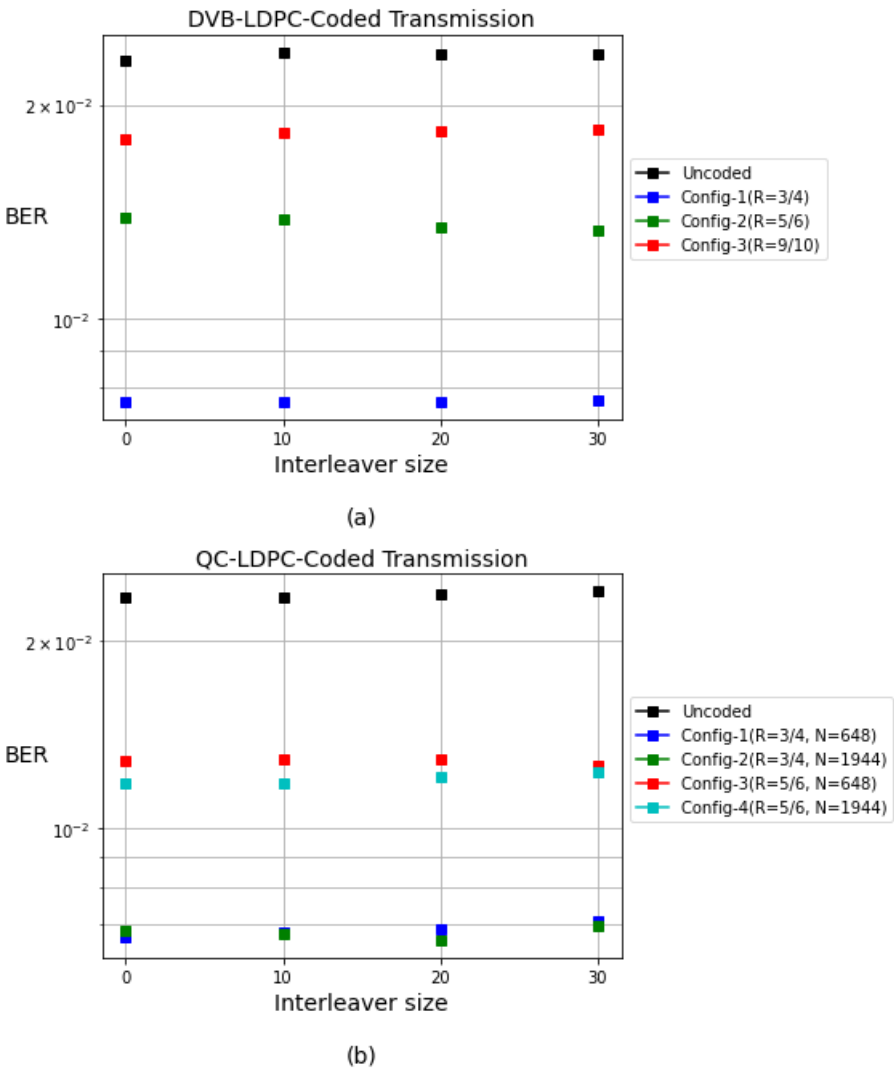


Figure 4.2: The Performance of Coded Systems with Different Interleaving Capacities. (a) FEC Based on DVB-S2 Standard. (b) QC-LDPC from IEEE 802.11 Protocol to the right

4.2 Impact of increasing maximum iteration of the decoder

Following the discussion in Chapter 3, the decoding functionality is performed in a number of iterations. The maximum permissible number I_{max} is a parameter that is subject to optimisation depending on the particular transmission scenario. The higher the allowed I_{max} is, the better the attained performance. However, this enhancement is constrained by the complexity the I_{max} imposes on the system. In addition, increasing the I_{max} should be carefully chosen as per the encoding scheme girth [46]. By aligning I_{max} to the coding girth, cycles are avoided and better performance is attained [1]. Therefore, a set of simulation experiments

were performed to investigate the optimum number of decoding iterations for each FEC candidate. A non-interleaved coded-modulated system was simulated over the aforementioned fibre channel assuming the following set of I_{max}'s values: 1, 5, 10, and 20.

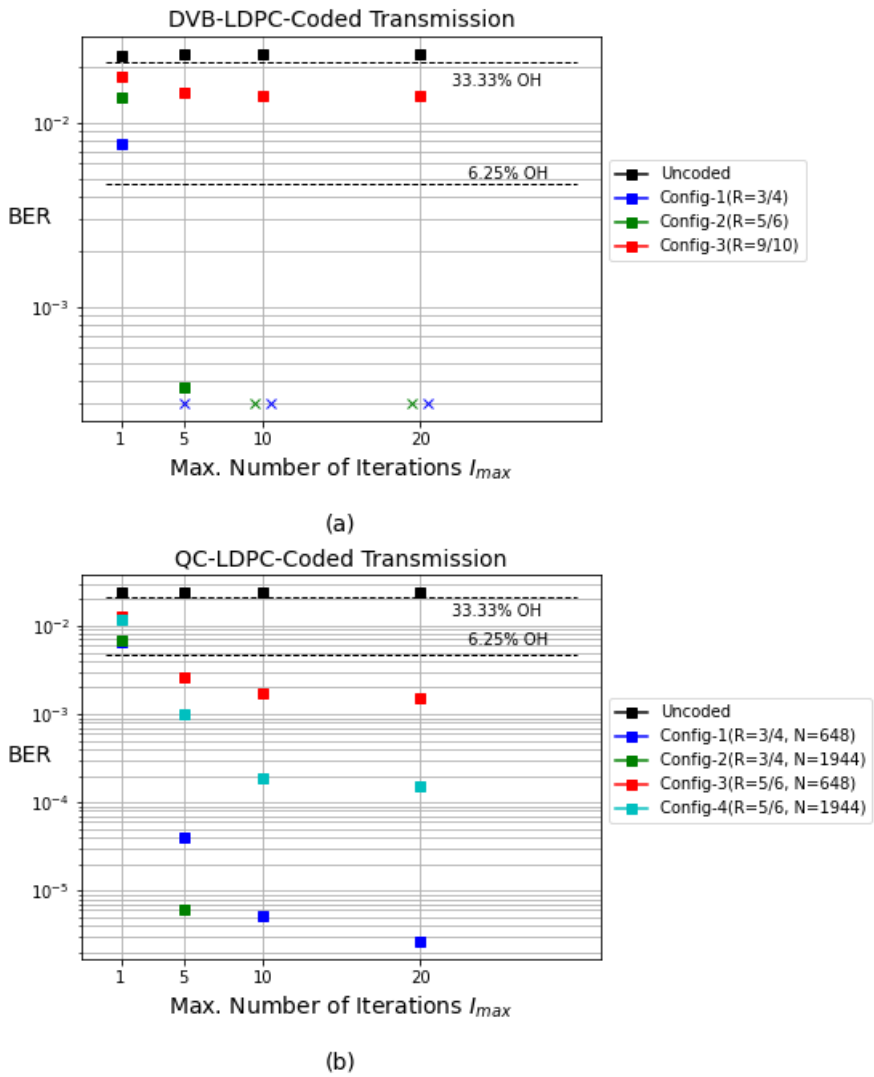


Figure 4.3: The Performance of Coded Systems with Different Decoding's Iterations. (a) FEC Based on DVB-S2 Standard. (b) QC-LDPC from IEEE 802.11. The coding rates {3/4, 5/6, 9/10} corresponds to the OH% {33.33%, 20%, 11.11%} respectively

Due to the nature of the study that is Monte-Carlo based, the simulations enabled to measure only up to 10^{-4} BER with an accuracy of %10. To achieve this, the simulated bit phrase had over 10^6 bits. Despite this large bit number, there were cases in which we did not manage to capture even a single error to characterise

the performance improvement that was achieved. In Figure 4.3 those cases are reported with a cross and signify the BER resolution of our simulator.
The horizontal dotted lines represent the BER performance thresholds required by the outer staircase decoder, for reaching a target BER of (10^{-15}). The upper threshold line is for a staircase code that requires a modest BER to function properly, but with a relatively high requirement of overhead (33.33%). The lower line has a more demanding BER threshold (4.7×10^{-3}), but it can be offered at a much lower overhead cost (6.25%). In both cases, the overhead introduced by the staircase decoder is added to the overall OH% of the FEC concatenated system. Figure 4.3 shows that the transmission system is incapable of achieving the desired BER without coding, as it couldn't pass reaching of coding on the performance. Figure 4.3 clearly demonstrates that increasing the number of iterations to 5, significantly improves the performance for all the configurations in both coding schemes. All the squares are shifted to below both targeted performance thresholds, which means that 5 iterations is a fair choice for the system's number of iterations. Therefore, the following investigations will use 5 iterations. An exception to the latter observation, is the Config-3 in the DVB scheme, which corresponds to the lowest overhead (11.11%). Its low overhead explains why it is tough to influence its performance. In a trial to address that, a higher number of iterations i.e. 10 will also be used in the next phase simulations. This inclusion will allow us to determine the existence of a particular decoder and DBP configuration that can achieve the desired performance, if any.

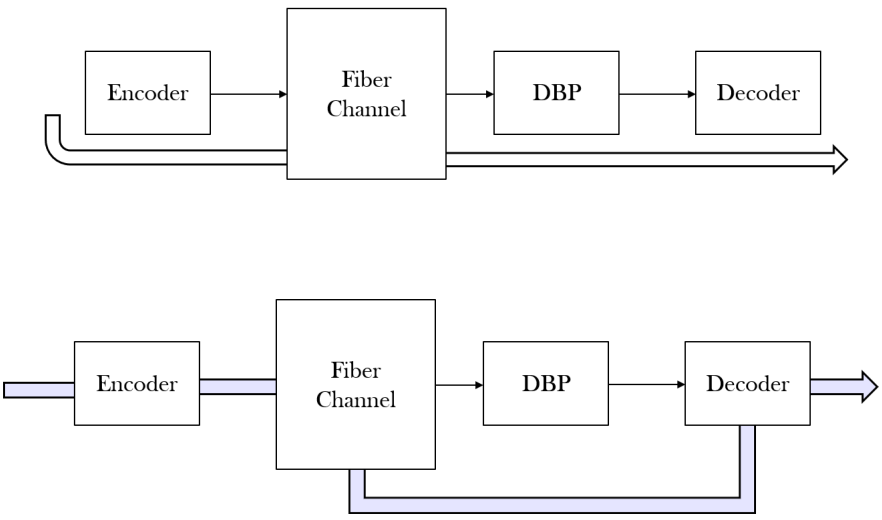
We also observed that without coding, the performance does not qualify to pass any of the two hard-decoder's thresholds. This suggests that coding is quite important for the system in order for it to achieve the required reliability. Furthermore, the various QC-LDPC coding schemes, shown in Figure 4.3(b), differ in the coding rate and the code length. Although the configurations have similar coding rates, they exhibit differentiation in the behaviour due to the code length difference. If we compare the red squares to the cyan, or the blue to the green, we would spot that there is a supremacy of performance for longer codes.

4.3 Combining FEC with DBP-based non-linearity compensation

The purpose of this section is to evaluate the effect of compensating the fibre channel non-linearities using the DBP technique in the presence of coding. The same 300km transmission link defined in the previous sections with 256 QAM

modulation format is considered also here. As explained in Chapter 3, the accuracy of the DBP compensation is critically affected by the number of computational steps per span of the back-propagation link. Although increasing this number brings better performance, this comes at a cost of high complexity. Therefore, a conservative selection of the steps/span parameter is made when implementing DBP for the sake of preserving design practicality. The steps/span under investigation were 2, 3, and 4.

Following the conclusions of Sections 4.1 and 4.2, the Coded-Modulated architecture does not include interleaving, and only performs iterative decoding with 5 or 10 iterations. Based on that, five different scenarios were considered for each subject coding scheme (refer to Figure 4.4). The first is the baseline, which represents transmission with neither coding nor DBP compensation. The second scenario assumed coded transmission with the correspondent FEC, but without any DBP compensation. The third, fourth and fifth scenarios represent coded transmissions, alongside DBP with 2, 3, and 4 steps/span, respectively. All the different scenarios are illustrated in Figure 4.4.



As discussed in Chapter 3, DVB-LDPC includes 3 configurations, while QC-LDPC, four, see Table 3.2 for the details of each configuration. Each one was examined The different FEC configurations were examined for two different maximum decoding iterations numbers, namely 5 and 10.

The numerical simulations took place on an HPC system. Each setup ran over a range of launch powers -10dBm to 10 dBm, and the minimum BER was recorded.

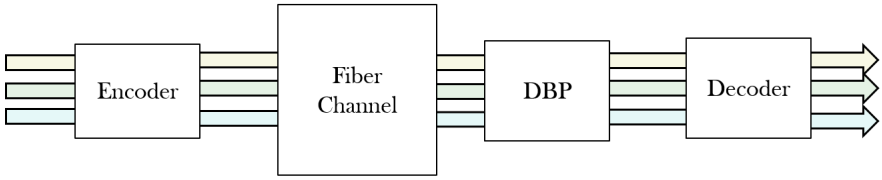


Figure 4.4: Block diagrams of the various simulation scenarios under investigation

The results of the DVB-coded system and QC-LDPC are shown in Figures 4.5 and 4.6, respectively. In those figures the horizontal lines mark the BER performance that needs to be achieved so the outer hard-decoder of the staircase codes is able to function and attain the targeted BER; 10^{-15}. The line labelled with 6.25% represents OH requirement of the staircase code, while the 33.33% the highest one. A lesser OH requirement comes at the cost of a lower BER_{pre} threshold, and vice versa.

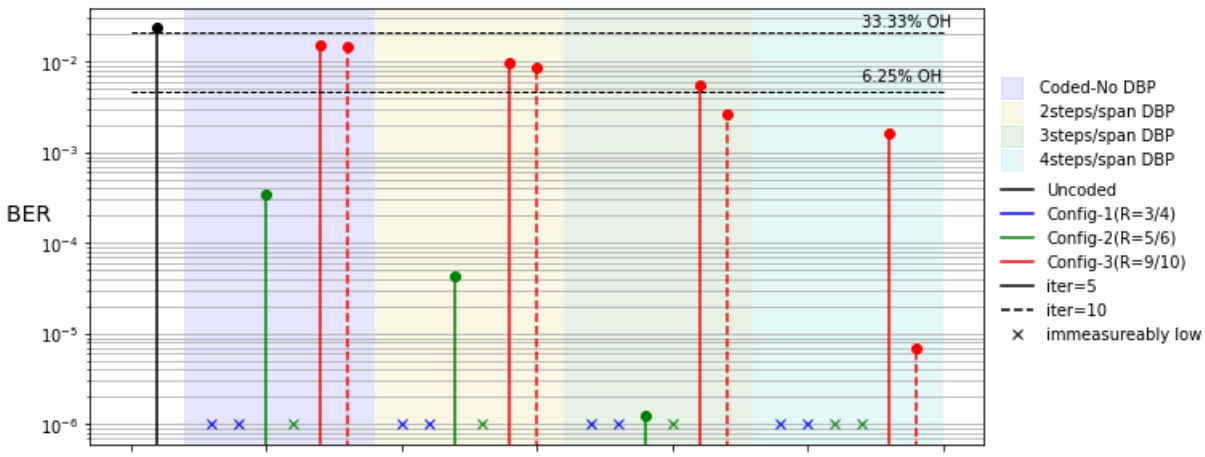


Figure 4.5: Performance of DVB-Coded Systems with and without DBP Compensation on a 300km long fibre channel. The coding rates {3/4, 5/6, 9/10} corresponds to the OH% {33.33%, 20%, 11.11%} respectively

In Figure 4.5 we can see that the uncoded scheme fails to pass the staircase decoder threshold, which means that concatenation is mandatory to attain the desired reliable performance. Coding with DVB succeeds in passing the required threshold without the help of DBP-based compensation (see the area shaded with purple in Figure 4.5). Indeed, we can obtain the desired BER of 10^{-15} with only 5 decoding iterations and with an overall overhead of 27.5%. The increase of iterations from 1 to 5 (when compared with the results of Figure 4.3) reports an

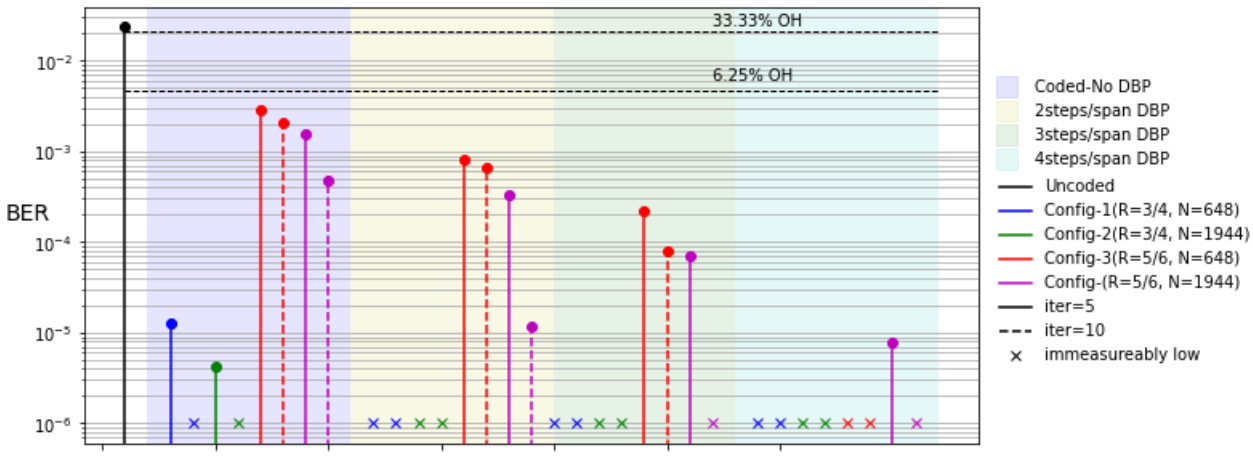


Figure 4.6: Performance of QC-LDPC-Coded Systems with and without DBP Compensation on a 300km long fibre channel. The coding rates, R, {3/4, 5/6} corresponds to the OH% {33.33%, 20%} respectively, and N resembles the code length

enhancement of 20.64% in the OH. On the other hand, integrating a DBP compensation with 3 steps/span and using 10 decoding iterations achieves a 30% enhancement with a resultant OH of 18.06%. This particular OH can be also achieved using less number of iterations i.e. 5, but at the cost of a higher number of steps/span in the DBP (4 steps/span).

Figure 4.6 demonstrates the performance results for a QC-LDPC-based coded system. We notice that there are two main configurations achieving different OHs and none of them requires implementing DBP, due to the relatively low coding rates this scheme is based on. The first configuration achieves 41.66% OH by using a 5 iterations decoder and any coding scheme with 3/4 coding rate (Config-1 or Config-2). The second scores a less OH, 27%, as it is dependent on a coding scheme with higher rates, 5/6, represented by Config-3 and Config-4. The 27% OH coding configuration is favoured over the 41.66%. When comparing the competent results from DVB and QC, we notice that DVB could lower the OH% by 9.44% after fulfilling the required conditions of 3 steps/span and 10 decoding iterations. This is due to the fact that DVB standards can offer very high coding rates such as 9/10.

4.4 Towards achieving a better reach with coding & DBP

The approach used in Section 4.3 is implemented for an increased number of fibre spans. The same experiment is repeated for 4 i.e. a total of 400km. To enable the comparison with the previous 300km case all the results are presented in the same format.

The results shown in Figure 4.7 and 4.8 suggest that there are three different ways to achieve the 10^{-15} BER target. The first is by accommodating a 41.67% OH without the need for DBP with a scheme of 3/4 coding rate and 5 iterations (i.e. Config-1 in the case of DVB and Config-1 or Config-2 in the case of QC). The second way is quite similar, except that it uses a scheme of 5/6 coding rate (i.e. Config-2 in the case of DVB and Config-3 or Config-4 in the case of QC)) resulting in an overall OH of 60%. When comparing the second way to the first, we notice it requires an increased overhead although its inner encoder uses a higher coder rate. That is due to the performance enhancement the lower coding rate (3/4) achieves enabling the utilisation of an outer staircase code with a higher coding rate i.e. 0.9412 or equivalently 6.25% OH. The third way is by incorporating DBP in a system that uses 5/6 coding rate at the inner encoder adding an OH of 27.5%. The DBP should have 2 steps/span if 10 decoding iterations are considered, or a 3 steps/span if 5 were used instead.

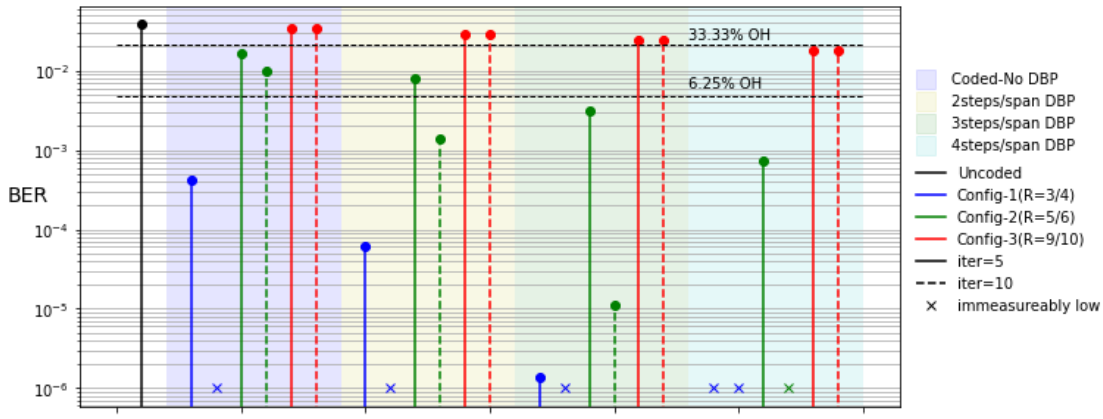


Figure 4.7: Performance of DVB-Coded Systems with and without DBP Compensation on a 400km long fibre channel. The coding rates {3/4, 5/6, 9/10} corresponds to the OH% {33.33%, 20%, 11.11%} respectively and N = 64800

Another way of representing the data by highlighting the total overhead required to implement the various coded scenarios mentioned above may facilitate a bet-

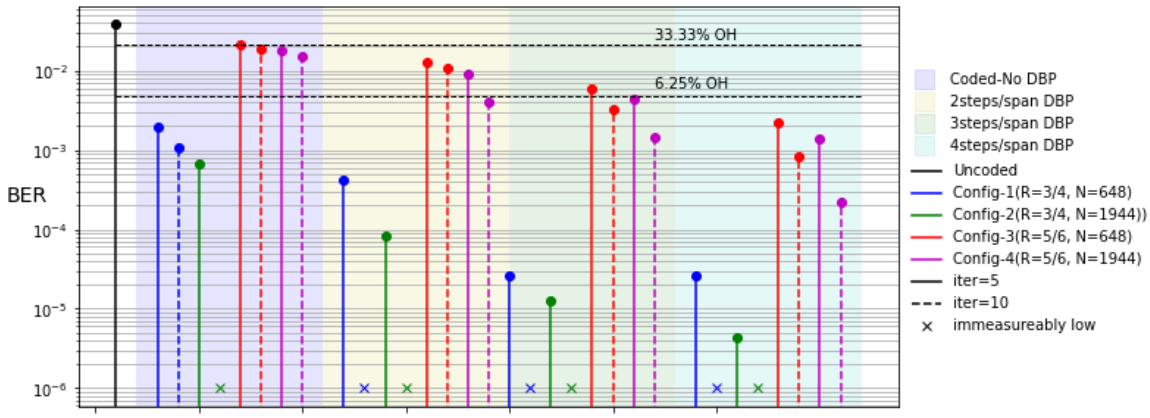


Figure 4.8: Performance of QC-LDPC-Coded Systems with and without DBP Compensation on a 400km long fibre channel. The coding rates {3/4, 5/6} corresponds to the OH% {33.33%, 20%} respectively, and N resembles the code length

ter understanding of the results. Tables 4.1 and 4.2 summarise the configurations which successfully achieved the targeted BER using DVB and QC based FECs, over the investigated channel lengths.

As already mentioned, the system is using a concatenation of an inner code with an HD staircase outer code. The optimum inner code is to be chosen from one of the FEC schemes under study, i.e. LDPC from DVB-S2 or the QC-LDPC. The coding rate of the outer code is dependent on the performance enhancement attained by employing the inner code. There are two possible coding rate configurations for the outer code, R_{outer}, 0.9412 and 0.75, which correspond to an OH of 6.25% and 33.33%, respectively. The DVB-based inner encoder, R_{inner}, supports coding rates of 3/4, 5/6, and 9/10, which correspond to the OH% of 33.33, 20, and 11.11 whilst the QC-LDPC supports only the rates of 3/4 and 5/6. The various configurations shown in Table 4.1 and 4.2, are based on 4 specifications; the steps/span requirement in DBP, the iterations number of decoding I_{max}, the coding rate of the inner coder R_{inner}, and the coding rate of the outer staircase code R_{outer}. For each distance, there are several configurations that could achieve the required reliability level at different costs. The cost is mainly quantified by the overhead amount. Therefore, the configurations are ordered based on the total overhead of the concatenation, OH_T, which is measured using Equation 2.3.

By observing Table 4.1, we notice that the concatenation that is based on an inner code of 5/6 rate results in a better OH than the one that uses a rate of 9/10.

Table 4.1: The configurations which reported a performance achieving the target BER in DVB-based codes

distance (km)	OH_T %	Configuration										
		DBP (steps/span)				I_{max}		R_{inner} (OH%)			R_{outer} (OH%)	
		0	2	3	4	5	10	3/4 (33.33)	5/6 (20)	9/10 (11.11)	0.75 (33.33)	0.9412 (6.25)
300	48.15											
	27.5											
	18.06											
400	48.15											
	41.67											
	27.5											

This might sound counter intuitive, due to the higher OH% the 5/6 code holds relative to 9/10. However, it makes sense if we take into account the significant performance enhancement the 5/6 coder achieves, allowing an outer code of only 6.25% overhead instead of 33.33%. Achieving the desired performance is possible without the use of DBP at an overall overhead cost of 27.5%. A decrease of 8.94% can be achieved if we use DBP with either 3 steps/span or 4 steps/span, and decoder iterations of 10 or 5, respectively. It is also evident that adding the 3 steps/span DBP increases the system reach by 33.33% without affecting the required overhead, which remains 27.5%.

Table 4.2 clearly demonstrates that when implementing QC-codes as an inner code for a 300km long transmission, there is no point in adding DBP to the system. Indeed, the relatively low coding rates this scheme supports, make it sufficient to enhance the performance up to the desired level with FEC without the need for DBP. In the case of 400km long transmission, we can attain the same performance as the 300km long channel at the same overhead cost, 27.5%, by implementing a 3 steps/span DBP.

Overall, both coding schemes achieved the targeted performance successfully. However, When comparing the precise performance enhancement caused by codes having identical coding rates, from the different investigated schemes, we notice that DVB outperforms. The first blue stem in both figures 4.7 and 4.8 is a sound proof. This over-performance is justified by the code length difference

Table 4.2: The configurations which reported a performance achieving the target BER in QC-based codes

distance (km)	OH_T %	Configuration									
		DBP (steps/span)				I_{max}		R_{inner} (OH%)		R_{outer} (OH%)	
		0	2	3	4	5	10	3/4 (33.33)	5/6 (20)	0.75 (33.33)	0.9412 (6.25)
300	27.5										
400	60										
	41.67										
	27.5										

between both schemes, when LDPC codes perform best for long codewords [1]. DVB-S2 codes are of 64800 bits long, while QC-LDPC supports a range of code lengths between 648 and 1944 bits per codeword. The huge gap in the length is due to the channel characteristics those codes were developed for. The investigated QC-LDPC codes are adapted from IEEE 802.11 protocol which is mainly developed for fading wireless channel. Therefore, smaller packets design suits the application which frequently requires re-transmission, thus, long codewords are impractical for such a channel.

5. CONCLUSIONS

This study aimed the development of a forward error correction FEC system for long-haul fibre channels integrated with a DBP-based fibre nonlinearity compensation scheme targeting the stringent BER performance requirement of 10^{-15}. The significance of compensating nonlinearities using DBP compensation in a coded system is investigated. A concatenated FEC system is used in a trivial bit interleaved coded modulation T-BICM architecture. The FEC concatenation consists of an outer encoder, inner encoder, soft inner decoder, and hard outer decoder. The outer coder uses staircase code with two options of coding rates; 0.9412 and 0.75. The outer coder corrects most of the errors at a low overhead cost if the inner coder successfully corrects a small number of errors to elevate the performance to a certain required level. The higher the attained performance enhancement post inner soft decoding, the lower the overhead required for the outer code. The inner encoder is adopted from recognised standards defined for other channel types that are attractive for their off-the-shelf availability. The surveyed inner coding schemes are irregular LDPC adopted from DVB-S2 standard and IEEE 802.11's Quasi-cyclic QC-LDPC. The investigated coding rates are 3/4 and 5/6 in both FECs, besides the addition of 9/10, which is only supported by DVB-S2.
In the scenario of 300km long fibre channel, the least attained overhead, 18.06%, is encountered when using DVB-Staircase concatenation with an individual overhead of 11.11% and 6.25% at the inner and outer coders, respectively. This achievement's prerequisites are either implementing a DBP with 3 steps/span, alongside 10 decoding iterations, or using 4 steps/span with 5 decoding iterations.

A less complex option that does not require DBP, and uses five decoding iterations is the concatenation of inner and outer codes holding overheads of 20% and 6.25%, respectively. The total overhead of this concatenation is 27.5%, which marks a minor difference of 9.44% relative to the DBP-integrated implementation. This study proves that it is possible to achieve a one-third increase in the transmission reach, from 300km to 400km, while using the same overhead requirement if a 2 steps/span DBP with 10 decoding iterations is used. Alternatively, a 3 steps/span DBP with 5 decoding iterations can be adopted. The preferable implementation is a vital point that requires further study by taking also into account the complexity of the iterative decoding along with the DBP complexity. The time frame restriction for this thesis prevented us from diving into addressing this question. However, we acknowledge its importance for and we are considering itwas conducted It is also found that the QC-LDPC-staircase concatenation does

not require DBP to meet the performance requirements of a 300km fibre channel. This is due to the low coding rates QC-LDPC codes support, which make them adequate and robust enough to obtain the required performance enhancement, although at the cost of higher overhead requirements.

The comparative study concludes by favouring DVB-based codes over QC-LDPC. The long codeword length they support makes them perform better than short QC-LDPC codes, which aligns with the expectations built upon the nature of LDPC codes. LDPC codes are known to perform better in very long code-length configurations. The other reason for recommending the LDPC codes adopted from the DVB-S2 protocol is its support for low overhead formats, which aligns with the long-haul fibre channel's FEC requirement.

ABBREVIATIONS

FEC	Forward Error Correction
MI	Mutual Information
DMS	Discrete Memoryless Source
OH	Overhead
IEEE	Institute of Electrical and Electronics Engineers
LDPC	Low Density Parity Check
QC-LDPC	Quasi-Cyclic Low Density Parity Check
SC-LDPC	Spatially Coded Low Density Parity Check
SD-TPC	Soft-Decision Turbo Product-Code
DBP	Digital Back Propagation
SD	Soft Decoding
BER	Bit Error Rate
HD	Hard Decoding
NCG	Net Coding Gain
SNR	Signal to Noise Ratio
ED	Euclidean Distance
LLR	Logarithm-Likelihood-ratio

REFERENCES

[1] J. G. Proakis and M. Salehi, *Digital Communication*, 2021.

[2] G. D. Forney and D. J. Costello, “Channel coding: The road to channel capacity,” *Proceedings of the IEEE*, vol. 95, no. 6, pp. 1150–1177, 2007.

[3] M. Torbatian, D. Lavery, M. Osman, D. Yao, D. Millar, Y. Gao, A. Kakkar, Z. El-Sahn, C. Doggart, A. E. Morra, N. Abughalieh, S. Yang, X. Chen, R. Maher, H. Sun, K. T. Wu, and P. Kandappan, “Performance oriented DSP for flexible long haul coherent transmission,” *Journal of Lightwave Technology*, vol. 40, no. 5, pp. 1256–1272, 2022.

[4] R. A. Stine, “Coding theory,” pp. 261–270, 2009.

[5] N. S. and M. H., “Performance Evaluation and Assessment of LDPC Codec over DVB-S2 and WLAN802.11n Applications,” *International Journal of Computer Applications*, vol. 181, no. 40, pp. 9–13, 2019.

[6] C. E. Shannon, “A Mathematical Theory of Communication,” *Bell System Technical Journal*, vol. 27, no. 4, pp. 623–656, 1948.

[7] *802.11n IEEE Standard*, 2009, no. December.

[8] R. W. Hamming, “Error Detecting and Error Correcting Codes,” *Bell System Technical Journal*, vol. 29, no. 2, pp. 147–160, apr 1950. [Online]. Available: https://ieeexplore.ieee.org/document/6772729

[9] I. Reed, “A class of multiple-error-correcting codes and the decoding scheme,” *Transactions of the IRE Professional Group on Information Theory*, vol. 4, no. 4, pp. 38–49, sep 1954. [Online]. Available: http://ieeexplore.ieee.org/document/1057465/

[10] P. Elias, “Coding for noisy channels,” *IRE Conv. Rec*, pp. 37–46, 1955.

[11] R. Bose and D. Ray-Chaudhuri, “On a class of error correcting binary group codes,” *Information and Control*, vol. 3, no. 1, pp. 68–79, mar 1960. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0019995860902874

[12] I. S. Reed and G. Solomon, “Polynomial Codes Over Certain Finite Fields,” *Journal of the Society for Industrial and Applied Mathematics*, vol. 8, no. 2, pp. 300–304, jun 1960. [Online]. Available: http://epubs.siam.org/doi/10.1137/0108018

[13] R. G. Gallager, “Low-Density Parity-Check Codes.” Ph.D. dissertation, Cambridge, MA: MIT Press, 1963.

[14] M. Sipser and D. Spielman, “Expander codes,” *IEEE Transactions on Information Theory*, vol. 42, no. 6, pp. 1710–1722, 1996. [Online]. Available: http://ieeexplore.ieee.org/document/556667/

[15] A. Alvarado and L. Szczecinski, *Bit interleaved Coded Modulation*, 2015.

[16] A. J. Viterbi, “Error bounds for convolutional codes and an asymptotically optimum decoding algorithm,” *IEEE Trans. Inform. Theory*, vol. IT-13, no. 4, pp. 260–269, 1967.

[17] J. L. Massey, “Shift-register synthesis and BCH decoding,” *IEEE Trans. Inform. Theory*, vol. IT-15, no. 1, pp. 122–127.

[18] L. R. et al. Bahl, “Optimal decoding of linear codes for minimizing symbol error rate,” *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 284–287, 1974.

[19] G. Ungerboeck and I. Csajka, “On improving data-link performance by increasing channel alphabet and introducing sequence decoding,” in *International Symposium on Information Theory (ISIT)*, 1976.

[20] H. Imai and S. Hirakawa, “A new multilevel coding method using error-correcting codes,” *IEEE Trans. Inf. Theory,*, vol. IT-23, no. 3, pp. 371–377, 1977.

[21] E. Zehavi, “8-PSK Trellis Codes for a Rayleigh Channel,” *IEEE TRANSACTIONS ON COMMUNICATION*, vol. 40, no. 5, pp. 873–884, 1992.

[22] D. Divsalar and F. Pollara, “On the design of turbo codes,” *TDA Progress Report 42-12*, pp. 99–121, 1995.

[23] P. R. Wörz and T., “Bandwidth-efficient turbo trellis-coded modulation using punctured component code,” *IEEE J. Sel. Areas Commun*, vol. 16, no. 2, pp. 206–218, 1998.

[24] H. G. Zhang, D. F. Yuan, P. M. Ma, and X. M. Yang, “Low-Density Parity-Check Code(LDPC) Schemes with BICM,” *International Conference on Communication Technology Proceedings, ICCT*, vol. 2, pp. 1148–1151, 2003.

[25] T. Fehenberger, G. Bocherer, A. Alvarado, and N. Hanik, “LDPC coded modulation with probabilistic shaping for optical fiber systems,” *Conference on Optical Fiber Communication, Technical Digest Series*, vol. 2015-June, no. March, pp. 2–5, 2015.

[26] P. Momtahan, “Ah, So That’s How Probabilistic Constellation Shaping Works!” 2020.

[27] A. Alvarado, F. Brannstrom, E. Agrell, and T. Koch, “High-SNR Asymptotics of Mutual Information for Discrete Constellations With Applications to BICM,” *IEEE Transactions on Information Theory*, vol. 60, no. 2, pp. 1061–1076, feb 2014. [Online]. Available: http://ieeexplore.ieee.org/document/6671479/

[28] A. G. I. Fàbregas, A. Martinez, and G. Caire, “Bit-interleaved coded modulation,” *Foundations and Trends in Communications and Information Theory*, vol. 5, no. 1-2, pp. 1–153, 2008.

[29] Xiaodong Li, A. Chindapol, and J. Ritcey, “Bit-interleaved coded modulation with iterative decoding and 8 PSK signaling,” *IEEE Transactions on Communications*, vol. 50, no. 8, pp. 1250–1257, aug 2002. [Online]. Available: http://ieeexplore.ieee.org/document/1025496/

[30] H. Ma, W. K. Leung, X. Yan, K. Law, and M. Fossorier, “Delayed bit interleaved coded modulation,” in *2016 9th International Symposium on Turbo Codes and Iterative Information Processing (ISTC)*. IEEE, sep 2016, pp. 86–90. [Online]. Available: http://ieeexplore.ieee.org/document/7593082/

[31] M.C. Jeruchim, P. Balaban, and K.S. Shanmugan, *Simulation of Communication Systems Second Edition*. New York: Kluwer, 2000.

[32] J. Cho, L. Schmalen, and P. J. Winzer, “Normalized Generalized Mutual Information as a Forward Error Correction Threshold for Probabilistically Shaped QAM,” *European Conference on Optical Communication, ECOC*, vol. 2017-Septe, no. 1, pp. 1–3, sep 2017.

[33] A. Alvarado, E. Agrell, D. Lavery, R. Maher, and P. Bayvel, “Replacing the Soft-Decision FEC Limit Paradigm in the Design of Optical Communication Systems,” *Journal of Lightwave Technology*, vol. 33, no. 20, pp. 4338–4352, 2015.

[34] I.-T. S. Group, “ITU-T Rec. G.975.1 (02/2004) Forward error correction for high bit-rate DWDM submarine systems,” *Standard*, pp. 1–58, 2005. [Online]. Available: papers2://publication/uuid/9A662586-4C7D-4DFB-B457-FE756BF7E680

[35] T. Mizuochi, Y. Miyata, T. Kobayashi, K. Ouchi, K. Kuno, K. Kubo, K. Shimizu, H. Tagami, H. Yoshida, H. Fujita, M. Akita, and K. Motoshima, “Forward error correction based on block turbo code with 3-bit soft decision for 10-Gb/s optical communication systems,” *IEEE Journal on Selected Topics in Quantum Electronics*, vol. 10, no. 2, pp. 376–386, 2004.

[36] I. B. Djordjevic and B. Vasic, “Projective geometry LDPC codes for ultralong-haul WDM high-speed transmission,” *IEEE Photonics Technology Letters*, vol. 15, no. 5, pp. 784–786, 2003.

[37] Z. Xiao, M. Li, and F. Yu, “Performance and power of soft-decision SDFEC for 100G -800G applications,” *Optics InfoBase Conference Papers*, vol. Part F174-, pp. 2020–2022, 2020.

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

[39] D. Chang, F. Yu, Z. Xiao, Y. Li, N. Stojanovic, C. Xie, X. Shi, X. Xu, and Q. Xiong, “FPGA verification of a single QC-LDPC code for 100 Gb/s optical systems without error floor down to BER of 10-15,” *Optics InfoBase Conference Papers*, pp. 15–17, 2011.

[40] L. Schmalen, V. Aref, J. Cho, D. Suikat, D. Rosener, and A. Leven, “Spatially coupled soft-decision error correction for future lightwave systems,” *Journal of Lightwave Technology*, vol. 33, no. 5, pp. 1109–1116, 2015.

[41] R. M. Pyndiah, “Near-optimum decoding of product codes: block turbo codes,” *IEEE Transactions on Communications*, vol. 46, no. 8, pp. 1003–1010, 1998.

[42] A. Alvarado, E. Agrell, D. Lavery, R. Maher, and P. Bayvel, “Replacing the Soft-Decision FEC Limit Paradigm in the Design of Optical Communication Systems,” *Journal of Lightwave Technology*, vol. 33, no. 20, pp. 4338–4352, 2015.

[43] R. Maher, T. Xu, L. Galdino, M. Sato, A. Alvarado, K. Shi, S. J. Savory, B. C. Thomsen, R. I. Killey, and P. Bayvel, “Spectrally Shaped DP-16QAM Super-Channel Transmission with Multi-Channel Digital Back-Propagation,” *Scientific Reports*, vol. 5, no. 1, 2015.

[44] A. Tychopoulos, O. Koufopavlou, and I. Tomkos, “FEC in optical communications - A tutorial overview on the evolution of architectures and the future prospects of outband and inband FEC for optical communications,” *IEEE Circuits and Devices Magazine*, vol. 22, no. 6, pp. 79–86, 2006.

[45] J. H. Bae, A. Abotabl, H. P. Lin, K. B. Song, and J. Lee, “An overview of channel coding for 5G NR cellular communications,” *APSIPA Transactions on Signal and Information Processing*, vol. 8, 2019.

[46] G. Tzimpragos, C. Kachris, I. B. Djordjevic, M. Cvijetic, D. Soudris, and I. Tomkos, “A Survey on FEC Codes for 100 G and beyond Optical Networks,” *IEEE Communications Surveys and Tutorials*, vol. 18, no. 1, pp. 209–221, 2016.

[47] L. M. Zhang and F. R. Kschischang, “Staircase codes with 6% to 33% overhead,” *Journal of Lightwave Technology*, vol. 32, no. 10, pp. 1999–2002, 2014.

[48] European Telecommunications Standards Institute, “EN 302 307-1 - V1.4.1 - Digital Video Broadcasting (DVB); Second generation framing structure, channel coding and modulation systems for Broadcasting, Interactive Services, News Gathering and other broadband satellite applications; Part 1: DVB-S2,” *En 302 307-1 V1.4.1*, vol. 1, pp. 1–80, 2014.

[49] IEEE, “IEEE Std 802.11n,” New York, Tech. Rep., 2009.

[50] A. Thangaraj. Soft-input soft-output iterative message passing decoder for ldpc codes. IIT Madras. [Online]. Available: https://www.youtube.com/watch?v=JoSltDBOLGoab_channel=NPTEL-NOCIITM

[51] V. Neubauer, Andre; Freudenberger, Jurgen; Kuhn, *coding theory-algorithms, architectures, and applications*, 2007.

[52] D. C. Hung, “Comparison of BICM-ID to Turbo code with high-ary modulation in wide band communication systems in the future,” *Proceedings - 2019 19th International Symposium on Communications and Information Technologies, ISCIT 2019*, pp. 249–254, 2019.

[53] C. Stierstorfer and R. F. Fischer, “(Gray) mappings for bit-interleaved coded modulation,” *IEEE Vehicular Technology Conference*, pp. 1703–1707, 2007.

[54] H. Tagami, T. Kobayashi, Y. Miyata, K. Ouchi, K. Sawada, K. Kubo, K. Kuno, H. Yoshida, K. Shimizu, T. Mizuochi, and K. Motoshima, “A 3-bit soft-decision IC for powerful forward error correction in 10-Gb/s optical communication systems,” *IEEE Journal of Solid-State Circuits*, vol. 40, no. 8, pp. 1695–1703, 2005.

[55] A. Alvarado, L. Szczecinski, and E. Agrell, “On the performance of BICM with trivial interleavers in nonfading channels,” *IEEE International Conference on Communications*, vol. 2013, no. 216715, pp. 7–12, 2011.

[56] G. Agrawal, *Applications of Nonlinear Fiber Optics*, San Diego, California, 1993.

[57] J. A. C. Weideman and B. M. Herbst, “Split-Step Methods for the Solution of the Nonlinear Schrödinger Equation,” *SIAM Journal on Numerical Analysis*, vol. 23, no. 3, pp. 485–507, jun 1986. [Online]. Available: http://epubs.siam.org/doi/10.1137/0723033

[58] E. Ip and J. M. Kahn, “Compensation of dispersion and nonlinear impairments using digital backpropagation,” *Journal of Lightwave Technology*, vol. 26, no. 20, pp. 3416–3425, 2008.