# Chapter 1

# Introduction

## 1.1 Historical Overview

Multimedia transmission over time-varying wireless channels presents a number of challenges beyond the existing capabilities of Third-Generation (3G) networks. The prevalent design principles stemming from Shannon's source and channel separation theorem [1] have to be reconsidered. The separation theorem, stating that optimal performance bounds of source and channel decoding can be approached as closely as desired by independently designing source and channel coding, holds only under asymptotic conditions, where both codes may have a length and a complexity tending to infinity, and under the conditions of stationary sources. In practice, the design of the system is heavily constrained in terms of both complexity and delay, and hence source and channel codecs are typically suboptimal [2]. Fortunately, the 'lossy' nature of the audio and video codecs may be concealed with the aid of perceptual masking by exploiting the psycho-acoustic and psycho-visual properties of the human ear and eye. The output of practical source encoders, such as speech codecs [3] and image/video codecs [4], may contain a significant amount of redundancy. Moreover, having a nonzero residual channel error rate may lead to dramatically increased source symbol error rates [2]. In these circumstances, the assumptions of Shannon's separation theorem no longer hold; neither can they be used as a good approximation. One may potentially improve the attainable performance by considering the source and channel code designs jointly, allowing us to exploit both the residual redundancy of the source-coded stream and the redundancy introduced by the channel code, or optimally allocate redundancy (in the best form) within the source and channel coding chain [2, 5–7].

### 1.1.1 Source Coding and Decoding

#### 1.1.1.1 Variable-Length Coding

Source coding constitutes a mapping from a sequence of symbols of an information source to a sequence of codewords (usually binary bits), so that the source symbols can be exactly recovered from the binary bits (lossless source coding) or recovered within some distortion (lossy source coding). Claude Shannon initiated this subject (along with other

subjects in the foundations of information theory) in his ground-breaking paper [1]. The key concept of entropy was formulated here. Shannon then filled the concept of entropy with practical significance by proving his noiseless source coding theorem. Shannon's theorem [1, Theorem 9] states that the entropy rate of a stationary ergodic finite-alphabet Markov source is the optimum rate at which the source can be encoded using 'almost noiseless' fixed-rate block codes or using noiseless variable-rate block codes. To prove the 'fixed-rate half' of the theorem, Shannon constructed fixed-rate block codes using an auxiliary theorem which later became known as the Shannon–McMillan Theorem [8]. In the proof of the 'variable-rate half' of the theorem, Shannon devised a noiseless variable-rate block code (nowadays referred to as a Shannon–Fano code), which encodes a source symbol having a probability of occurrence $p$ into a codeword having a length within one bit of $-\log_2 p$.

The Shannon–Fano coding algorithm is capable of producing fairly efficient variable-length prefix codes based on a set of symbols and their probabilities (estimated or measured), but it is nonetheless suboptimal. As a remedy, Huffman coding [9], named after its inventor, D. A. Huffman, is always capable of producing optimal prefix codes, which approach the lower bound of the expected codeword length under the constraint that each symbol is represented by a codeword formed of an integer number of bits. This constraint, however, is often unnecessary, since the codewords will be packed end-to-end in long sequences. If we consider groups of codes at a time, symbol-by-symbol Huffman coding is only optimal if the probabilities of the symbols are independent and assume a value of $1/2^n$, where $n$ is an integer. In most situations, arithmetic coding [10] is capable of producing a higher overall compression than either Huffman or Shannon–Fano coding, since it is capable of encoding in terms of fractional non-integer numbers of bits, which more closely approximate the actual information content of the symbol. However, arithmetic coding has not superseded Huffman coding, as Huffman coding superseded Shannon–Fano coding, partly because arithmetic coding is more computationally expensive and partly because it is covered by multiple patents. An excellent tutorial on the history of source coding can be found in [11].

In addition to coding efficiency, the robustness of the variable-length codes (VLCs) to transmission errors is another design concern. Takishima *et al.* [12] proposed Reversible Variable-Length Codes (RVLCs) for providing both forward and backward decoding capability so that the correct data can be recovered with as high a probability as possible from the received data when transmission errors have occurred. For example, a decoder can commence by processing the received bit stream in the forward direction, and upon detecting an error proceed immediately to the end of the bit stream and decode in the reverse direction. This technique and a range of other related strategies can be used for significantly reducing the effects of bit errors on the fidelity of the decoded data. Wen and Villasenor [13, 14] proposed a new class of asymmetric RVLCs having the same codeword length distribution as the Golomb–Rice codes [15, 16] as well as exp-Golomb codes [17], and applied them to the video coding framework of the H.263+ [18] and MPEG-4 [19] standards. Later on, in the quest to construct optimal RVLCs, further improved algorithms were proposed in [20–24]. More particularly, Lakovic and Villasenor [22] designed a family of new RVLCs having a free distance of $d_f = 2$, which exhibit a better error-correcting performance than the RVLCs and Huffman codes having a lower free distance. The conditions for the existence of RVLCs and their properties were also studied in [25–27].

For the sake of attaining an even stronger error-correction capability, Buttigieg and Farrell [28, 29] proposed a new class of Variable-Length Error-Correction (VLEC) codes, which generally have higher free distances ($d_f > 2$). It was shown in [29] that the free distance is the most important parameter that determines the decoding performance at high $E_b/N_0$ values. Detailed treatment of various VLEC code properties, their construction, decoding and performance bounds can be found in Buttigieg's PhD thesis [30]. Lamy and Paccaut [31] also investigated a range of low-complexity construction techniques designed for VLEC codes. In fact, both Huffman codes and RVLCs may be viewed as special VLEC codes having a free distance of $d_f < 2$ [32]. Hence a generic construction algorithm may be applied for designing both RVLCs and VLEC codes [32]. For the reader's convenience, we have summarized the major contributions on Variable-Length Coding (VL-coding) in Table 1.1, where they can be seen at a glance.

### 1.1.1.2 *Variable-Length Decoding*

Owing to their prefix property, source messages encoded by Huffman codes as well as RVLCs and VLEC codes may be decoded on a bit-by-bit basis with the aid of their code trees or lookup tables [2, 30, 33]. Given the wide employment of the ITU-T H.26x and ISO/IEC Motion Picture Experts Group (MPEG) series video coding standards, numerous efficient hardware architectures have been proposed [34, 35] for achieving low-power, memory-efficient and high-throughput decoding. These methods aim for low-latency, near-instantaneous decoding and generally assume having an error-free input bit stream. However, owing to their variable-length nature, their decoding is prone to the loss of synchronization when transmission errors occur, and the errors are likely to propagate during the decoding of the subsequent bit stream. In other words, should a decoding error be encountered in the current codeword, a relatively large number of subsequent erroneously decoded codewords are likely to follow, even if in the meantime there were no further errors imposed by the channel. Hence, although these codes may be decoded instantaneously, it is clear that if we were to wait for several codewords before making a decision, a better decoding performance (fewer decoding errors) could be expected. Indeed, the decoding of VLCs in error-prone environments is jointly a segmentation (i.e. recovering source symbol boundaries) and an estimation problem.

Numerous source-decoding algorithms capable of generating a source estimate based on a sequence of received data have been developed [7, 28, 29, 36–46]. In parallel, various trellis representations [29, 37, 38, 40, 41, 44] have been devised for describing the source (coder). In [28, 29] a Maximum Likelihood (ML) sequence estimation algorithm was proposed, which selected the most likely symbol sequence instead of symbol-by-symbol decoding. More specifically, the decoder invoked a modified form of the Viterbi Algorithm (VA) [47] based on the trellis structure of VLEC codes, where each state of the trellis corresponded to the number of bits produced so far and each transition between the states represented the transmission of a legitimate codeword (or source symbol). Maximum *A Posteriori* Probability (MAP) sequence estimation based on the same trellis structure was also studied in [30]. Both algorithms are based on hard decoding, where the inputs of the source decoder are binary zeros as well as ones and the path metric is calculated based on the Hamming distance between the transmitted bits and the received bits. Soft information may be used for improving the attainable decoding performance, resulting in Soft-Input Soft-Output (SISO) decoding algorithms. The term 'soft'

Table 1.1: Major contributions on variable-length coding

| Year | Author(s) | Contribution |
|------|-----------|--------------|
| 1948 | Shannon [1] | Invented the first variable-length coding algorithm, referred to as Shannon–Fano coding, along with the creation of information theory. |
| 1952 | Huffman [9] | Proposed the optimal prefix codes, Huffman codes. |
| 1979 | Rissanen [10] | Proposed arithmetic coding. |
| 1994 | Buttigieg and Farrell [28–30] | Proposed VLEC codes and investigated various VLEC code properties, construction, decoding and performance bounds. |
| 1995 | Takishima *et al.* [12] | Proposed RVLCs for providing both a forward and a backward decoding capability. |
| 1997 | Wen and Villasenor [13, 14] | Proposed a new class of asymmetric RVLCs and applied them to the video coding framework of the H.263+ [18] and MPEG-4 [19] standards. |
| 2001 | Tsai and Wu [20, 21] | Proposed the Maximum Symmetric-Suffix Length (MSSL) metric [20] for designing symmetric RVLCs and the Minimum Repetition Gap (MRG) [21] metric for asymmetric RVLCs. |
| 2002 | Lakovic and Villasenor [22] | Proposed a new family of RVLCs having a free distance of $d_f = 2$. |
| 2003 | Lakovic and Villasenor [23] | Proposed the so-called 'affix index' metric for constructing asymmetric RVLCs. |
|      | Lamy and Paccaut [31] | Proposed low-complexity construction techniques for VLEC codes. |

implies that the decoder is capable of accepting as well as of generating not only binary ('hard') decisions, but also reliability information (i.e. probabilities) concerning the bits.

The explicit knowledge of the total number of transmitted source symbols or transmitted bits can be incorporated into the trellis representations as a termination constraint in order to assist the decoder in resynchronizing at the end of the sequence. Park and Miller [37] designed a symbol-based trellis structure for MAP sequence estimation, where the number of transmitted bits is assumed to be known at the receiver, resulting in the so-called bit-constrained directed graph representation. Each state of the trellis corresponds to a specific source symbol and the number of symbols produced so far. By contrast, when the total number of transmitted symbols is assumed to be known at the receiver, Demir and Sayood [38] proposed a symbol-constrained directed graph representation, where each

state of the trellis corresponds to a specific symbol and the corresponding number of bits that have been produced. Moreover, when both constraints are known at the receiver, Bauer and Hagenauer [40] designed another symbol-trellis representation for facilitating symbol-by-symbol MAP decoding of VLCs, which is a modified version of the Bahl–Cocke–Jelinek–Raviv (BCJR) algorithm [48]. Later they [42] applied their BCJR-type MAP decoding technique to a bit-level trellis representation of VLCs that was originally devised by Balakirsky [41], where the trellis was derived from the corresponding code tree by mapping each tree node to a specific trellis state. The symbol-based trellis has the drawback that for long source sequences the number of trellis states increases quite drastically, while the bit-based trellis has a constant number of states, which only depends on the depth of the VLC tree.

For sources having memory, the inter-symbol correlation is usually modeled by a Markov chain [2], which can be directly represented by a symbol-based trellis structure [36–38, 43]. In order to fully exploit all the available *a priori* source information, such as the inter-symbol correlation or the knowledge of the total number of transmitted bits and symbols, a three-dimensional symbol-based trellis representation was also proposed by Thobaben and Kliewer in [44]. In [45], the same bit trellis as that in [42] was used, but the *A Posteriori* Probability (APP) decoding algorithm was modified for the sake of exploiting the inter-symbol correlation. As a low-complexity solution, sequential decoding [49, 50] was also adopted for VLC decoding. In addition to classic trellis-based representations, Bayesian networks were also employed for assisting the design of VLC decoders [46], where it was shown that the decoding of the Markov source and that of the VLC can be performed separately.

### 1.1.1.3  *Classification of Source-Decoding Algorithms*

Generally, the source-decoding problem may be viewed as a hidden Markov inference problem, i.e. that of estimating the sequence of hidden states of a Markovian source/source-encoder through observations of the output of a memoryless channel. The estimation algorithms may employ different estimation criteria and may carry out the required computations differently. The decision rules can be optimum either with respect to a sequence of symbols or with respect to individual symbols. More explicitly, the source-decoding algorithms can be broadly classified into the following categories [7].

- **MAP/ML Sequence Estimation:** The estimate of the entire source sequence is calculated using the MAP or ML criterion based on all available measurements. When the *a priori* information concerning the source sequence is uniformly distributed, implying that all sequences have the same probability, or simply ignored in order to reduce the computational complexity, the MAP estimate is equivalent to the ML estimate. The algorithms outlined in [28, 29, 36–38, 41, 46] belong to this category.

- **Symbol-by-Symbol MAP Decoding:** This algorithm searches for the estimate that maximizes the *a posteriori* probability for each source symbol rather than for the entire source sequence based on all available measurements. The algorithms characterized in [40, 42–46] fall into this category.

- **Minimum Mean Square Error (MMSE) Decoding:** This algorithm aims at minimizing the expected distortion between the original signal and the reconstructed signal,

where the expected reconstructed signal may be obtained using APP values from the MAP estimation. The algorithm presented in [39] falls into this category.

### 1.1.2   Joint Source–Channel Decoding

Let us consider a classic communication chain, using a source encoder that aims at removing the redundancy from the source signal, followed by a channel encoder that aims to reintroduce the redundancy in the transmitted stream in an efficiently controlled manner in order to cope with transmission errors [51] (Chapter 1, page 2). Traditionally, the source encoder and the channel encoder are designed and implemented separately according to Shannon's source and channel coding separation theorem [1], which holds only under idealized conditions. Therefore Joint Source–Channel Decoding (JSCD) has gained considerable attention as a potentially attractive alternative to the separate decoding of source and channel codes. The key idea of JSCD is to exploit jointly the residual redundancy of the source-coded stream (i.e. exploiting the suboptimality of the source encoder) and the redundancy deliberately introduced by the channel code in order to correct the bit errors and to find the most likely source symbol estimates. This requires modeling of all the dependencies present in the complete source–channel coding chain.

In order to create an exact model of dependencies amenable to optimal estimation, one can construct the amalgam of the source, source encoder and channel encoder models. This amalgamated model of the three elements of the chain has been proposed in [52]. The set of states in the trellis-like joint decoder graph contains state information about the source, the source code and the channel code. The amalgamated decoder may then be used to perform a MAP, ML or MMSE decoding. This amalgamated approach allows for optimal joint decoding; however, its complexity may remain excessive for realistic applications, since the state-space dimension of the amalgamated model explodes in most practical cases. Instead of building the Markov chain of the amalgamated model, one may consider the serial or parallel connection of two Hidden Markov Models (HMMs), one for the source as well as the source encoder and one for the channel encoder, in the spirit of serial and parallel turbo codes. The dimension of the state space required for each of the separate models is then reduced. Furthermore, the adoption of the turbo principle [53] led to the design of iterative estimators operating alternately on each factor of the concatenated model. The estimation performance attained is close to the optimal performance only achievable by the amalgamated model [7].

In [37, 38, 40–46, 54, 55], the serial concatenation of a source and a channel encoder was considered for the decoding of both fixed-length and variable-length codes. *Extrinsic* information was iteratively exchanged between the channel decoder and the source decoder at the receiver. The convergence behavior of iterative source–channel decoding using fixed-length source codes and a serially concatenated structure was studied in [56], and in [45] for VLCs. The gain provided by the iterations is very much dependent on the amount of correlation present at both sides of the interleaver. The variants of the algorithms proposed for JSCD of VLC-encoded sources relate to various trellis representation forms derived for the source encoder, as described in Section 1.1.1.2, as well as to the different underlying assumptions with respect to the knowledge of the length of the symbol or bit sequences. Moreover, a parallel-concatenated source–channel coding and decoding structure designed for VLC-encoded sources is described in [57]. In contrast to a turbo code, the explicit redundancy inherent in one of the component channel codes is replaced by the residual

redundancy left in the source-compressed stream after VLC encoding. It was shown in [57] that the parallel concatenated structure is superior to the serial concatenated structure for the RVLCs considered, and for the Additive White Gaussian Noise (AWGN) channel experiencing low Signal-to-Noise Ratios (SNRs).

Another possible approach is to modify the channel decoder in order to take into account the source statistics and the model describing both the source and the source encoder, resulting in the so-called Source-Controlled Channel Decoding (SCCD) technique [58]. A key idea presented in [58] is to introduce a slight modification of a standard channel-decoding technique in order to take advantage of the source statistics. This idea has been explored [58] first in the case of Fixed-Length Coding (FLC), and has been validated using convolutional codes in the context of transmitting coded speech frames over the Global System of Mobile telecommunication (GSM) [59]. Source-controlled channel decoding has also been combined with both block and convolutional turbo codes, considering FLC for hidden Markov sources [60] or images [61–63]. Additionally, this approach has been extended to JSCD of VLC sources, for example in the source-controlled convolutional decoding of VLCs in [64, 65] and the combined decoding of turbo codes and VLCs in [64–67]. The main contributions on JSCD of VLCs are summarized in Table 1.2.

### 1.1.3 Iterative Decoding and Convergence Analysis

The invention of turbo codes over a decade ago [69] constitutes a milestone in error-correction coding designed both for digital communications and for storage. The parallel concatenated structure, employing interleavers for decorrelating the constituent encoders, and the associated iterative decoding procedure, exchanging only extrinsic information between the constituent decoders, are the key elements that are capable of offering a performance approaching the previously elusive Shannon limit, although they also impose an increasing complexity upon inching closer to the limit, in particular beyond the so-called Shannonian cut-off [70]. Reliable communication for all channel capacity rates slightly in excess of the source entropy rate becomes approachable. The practical success of the iterative turbo decoding algorithm has inspired its adaptation to other code classes, notably serially concatenated codes [71], Turbo Product Codes (TPCs) which may be constituted from Block Turbo Codes (BTCs) [72–74], and Repeat–Accumulate (RA) codes [75], and has rekindled interest [76–78] in Low-Density Parity-Check (LDPC) codes [79], which constitute a historical milestone in iterative decoding.
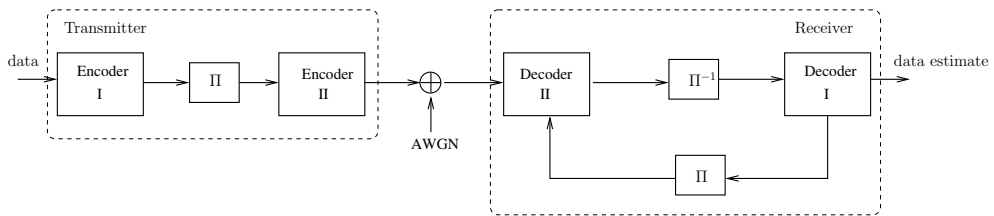
The serially concatenated configuration of components [71] commands particular interest in communication systems, since the 'inner encoder' of such a configuration can be given more general interpretations [53, 80, 81], as seen in Figure 1.1, such as an 'encoder' constituted by a dispersive channel or by the spreading codes used in Code-Division Multiple Access (CDMA). The corresponding iterative decoding algorithm can then be extended into new areas, giving rise to Turbo Equalization (TE) [82], Bit-Interleaved Coded Modulation (BICM) [83, 84], serial concatenated trellis-coded modulation [85, 86], turbo multiuser detection [87, 88], turbo synchronization [89] and coded CDMA [90]. Connection of turbo decoding algorithms to the belief propagation algorithm [91] has also been discussed in [92]. Moreover, factor graphs [93, 94] were employed [95, 96] to provide a unified framework that allows us to understand the connections and commonalities among seemingly different iterative receivers.

Table 1.2: Major contributions on JSCD of VLCs

| Year | Author(s) | Contribution |
|------|-----------|--------------|
| 1991 | Sayood and Borkenhagen [4] | Considered the exploitation of the residual redundancy found in the fixed-length coded quantized-source indices in the design of joint source–channel coders. |
| 1994 | Phamdo and Farvardin [68] | Studied the optimal detection of discrete Markov sources over discrete memoryless channels. |
| 1995 | Hagenauer [58] | Proposed the first SCCD scheme with application in speech transmissions. |
| 1997 | Balakirsky [41] | Proposed a bit-level trellis representation of VLCs and applied it in a joint source–channel VLC sequence estimation scheme. |
| 1998 | Murad and Fuja [52] | Proposed a joint source–channel VLC decoding scheme based on a combined model of the source, the source encoder and the channel encoder. |
|      | Miller and Park [39] | Proposed a sequence-based approximate MMSE decoder for JSCD. |
|      | Demir and Sayood [38] | Investigated a JSCD scheme using MAP sequence estimation based on trellis structure constrained by the number of transmitted bits. |
| 1999 | Park and Miller [37] | Studied a JSCD scheme using both exact and approximate MAP sequence estimation based on a trellis structure constrained by the number of transmitted symbols. |
| 2000 | Bauer and Hagenauer [40] | Proposed a symbol-level trellis representation of VLCs and derived a symbol-by-symbol MAP VLC decoding algorithm. |
|      | Guivarch et al. [64] | Considered SCCD of Huffman codes combined with turbo codes. |
|      | Peng et al. [63] | Investigated a SCCD scheme using turbo codes for image transmission. |
| 2001 | Bauer and Hagenauer [42] | Investigated MAP VLC decoding based on the bit-level trellis proposed in [41]. |
|      | Guyader et al. [46] | Studied JSCD schemes using Bayesian networks. |
|      | Garcia-Frias and Villasenor [60] | Applied SCCD to the decoding of hidden Markov sources. |

Table 1.2: Continued

| Year | Author(s) | Contribution |
|------|-----------|--------------|
| 2002 | Kliewer and Thobaben [43] | Adopted the JSCD scheme of [40] to the decoding of correlated sources. |
| 2003 | Thobaben and Kliewer [44] | Proposed a 3D symbol-based trellis for the sake of fully exploiting the *a priori* source information. |
| 2004 | Jaspar and Vandendorpe [67] | Proposed a SCCD scheme using a hybrid concatenation of three SISO modules. |
| 2005 | Thobaben and Kliewer [45] | Proposed a modified BCJR algorithm [48] incorporating inter-symbol correlation based on the bit-level trellis of [41]. |
|  | Jeanne *et al.* [65] | Studied SCCD schemes using both convolutional codes and turbo codes. |



| Configuration | En-/Decoder I (Outer Code) | En-/Decoder II (Inner Code) |
|---------------|----------------------------|-----------------------------|
| Serial Concatenated Code | FEC En-/Decoder | FEC En-/Decoder |
| LDPC Codes | Check Nodes | Variable Nodes |
| Turbo Equalization | FEC En-/Decoder | Multipath Channel/Equalizer |
| Turbo BICM | FEC En-/Decoder | Mapper/Demapper |
| Turbo Source–Channel Decoding | Source En-/Decoder | FEC En-/Decoder |
| Turbo CDMA | FEC En-/Decoder | Spreading Code/MUD |
| Turbo MIMO | FEC En-/Decoder | Mapper/MIMO Detector |

Figure 1.1: A serial concatenated system with iterative detection/decoding [97], where $\Pi$ denotes the interleaver and $\Pi^{-1}$ represents the de-interleaver. (MIMO: Multiple-Input Multiple-Output; MUD: Multi-User Detection; FEC: Forward Error Correction.)

Typically, Bit-Error Ratio (BER) charts of iterative decoding schemes can be divided into three regions [98]:

- the region of low $E_b/N_0$ values with negligible iterative BER reduction;
- the 'turbo-cliff' region, which is also referred to as the 'waterfall' region, with consistent iterative BER reduction over many iterations;

- the BER floor region often recorded at moderate to high $E_b/N_0$ values, where a rather low BER can be reached after just a few iterations.

Efficient analytical bounding techniques have been found for moderate to high $E_b/N_0$ values in [71, 99–102], where design rules were also derived for the sake of lowering the BER floors. On the other hand, a long-standing open problem was understanding the convergence behavior of iterative decoding algorithms in the turbo-cliff region. A successful approach to this problem was pioneered in [77, 78], which investigated the Probability Density Functions (PDFs) of the information communicated within the iterative decoding algorithms. Based on the evolution of these PDFs during iterative decoding, convergence thresholds have been derived for investigating the iterative decoding of LDPC codes [77] with the aid of the related technique, which is referred to as *density evolution*. Simplified approaches were also proposed by assuming that the PDFs considered above are Gaussian [103], or by observing a single parameter of them such as the SNR measure in [104], a reliability measure in [105] and the mutual information measure in [98, 106]. The convergence prediction capabilities of six different measures were compared in [107], where the mutual information measure used in the EXtrinsic Information Transfer (EXIT) chart [98, 106] analysis was found to have a high accuracy, regardless of the specific iterative decoding schemes considered. Additionally, the cross-entropy [108] was employed by Luo and Sweeney in [109] as a measure used for determining the convergence threshold of turbo decoding algorithms, where neither the knowledge of the source information nor the Gaussian approximation was required, while the relationship between the Minimum Cross-Entropy decoding algorithm and the iterative decoding algorithm was discussed in [110].

EXIT charts have been successfully applied to predict the convergence behavior of various concatenated systems, such as parallel [98, 111] and serially concatenated codes [112], turbo equalization [113], Bit-Interleaved Coded Modulation using Iterative Decoding (BICM-ID) [106, 114] and JSCD [45, 56]. A unified framework derived to apply EXIT chart analysis for various serial concatenation schemes was presented in [115]. EXIT functions were defined with the aid of a universal decoding model in [116] and several properties of the related functions were derived for erasure channels. A specific property allows us to quantify the area under an EXIT function in terms of the conditional entropy. A useful consequence of this result is that the design of capacity-approaching codes reduces to a curve-fitting problem in the context of EXIT charts [116]. For example, EXIT charts have been applied for designing Irregular Convolutional Codes (IrCCs) [80, 81] in order to approach the AWGN channel's capacity, and for designing irregular LDPC codes [117] and Irregular Repeat–Accumulate (IRA) codes [118] for approaching the Multiple-Input Multiple-Output (MIMO) fading channel's capacity.

EXIT charts have then also been extended from 2D to 3D for the sake of analyzing the convergence behavior of multi-stage iterative decoding schemes consisting of more than two components, such as, for example, Parallel Concatenated ('turbo') Codes (PCCs) consisting of three constituent codes [119], three-stage turbo equalization [120] and three-stage iterative decoding and demodulation [121, 122]. A 3D to 2D projection technique was presented in [120–122], which simplifies the convergence analysis of multi-stage systems. Optimizing the activation order of the component decoders was also investigated in [121, 122]. Furthermore, the philosophy of EXIT chart analysis has been extended from

binary decoders to non-binary decoders [123, 124]. The major contributions on iterative decoding and its convergence analysis are summarized in Tables 1.3 and 1.4.

## 1.2 Applications of Irregular Variable-Length Coding

In this monograph we introduce the novel concept of Irregular Variable-Length Codes (IrVLCs) and investigate their applications, characteristics and performance in the context of digital multimedia broadcast telecommunications. More specifically, we consider the employment of IrVLCs to represent or encode multimedia source signals, facilitating their transmission to, and reconstruction by, remote receivers. Explicitly, IrVLCs are defined as encoding schemes that represent particular components or segments of the source signals with different sets of codewords, which comprise various numbers of bits. This may be contrasted with regular Variable-Length Codes (VLCs), which encode all components or segments of the source signals using the same codebook of codewords.

Three particular aspects of IrVLCs designed for digital multimedia telecommunications are investigated in this monograph, namely near-capacity operation, joint source and channel coding, and Unequal Error Protection (UEP).

### 1.2.1 Near-Capacity Operation

In his seminal contribution [133], Shannon considered source signals that contain some redundancy. By definition, the redundant content of a particular source signal may theoretically be accurately predicted, given a perfect knowledge of its unpredictable information content. Provided that unimpaired knowledge of the source signal's information content can be reliably conveyed to the receiver by the channel, the transmitter is allowed to employ source coding in order to remove all the source's redundancy, without jeopardizing the receiver's ability to reconstruct it. When communicating over a perfectly noiseless channel, Shannon [133] showed that the minimum number of bits that are required to reliably convey perfect knowledge of the source signal's information content to the receiver is given by the source's entropy. Note however that the computational complexity and latency imposed by a source codec typically escalates in optimal entropy coding, where all redundancy is removed from the source signal.

Unfortunately, when communicating over noisy channels, which impose uncertainty on the received signal, the reliable transmission of a source signal's information content can never be guaranteed. However, Shannon [133] also showed that if a source signal's information content is conveyed over a noisy channel at a rate (expressed in bits per second) that does not exceed the channel's capacity, then it is theoretically possible to reconstruct it with an infinitesimally low probability of error. This motivates the employment of channel coding that introduces redundancy into the transmitted signal in a specifically designed manner. This redundancy may be exploited in the receiver to mitigate any channel-induced errors within the original non-redundant information content.

As with optimal entropy coding, the computational complexity and latency imposed by a channel codec escalates when approaching the channel's capacity [134]. Indeed, Jacobs and Berlekamp [135] showed that the decoding trees employed by the then state-of-the-art sequential decoding algorithm [136–139] for Convolutional Codes (CCs) [140] have an escalating computational complexity, when operating at an information rate above the so-called cut-off rate of the channel. This cut-off rate has an upper bound, which is equal to

Table 1.3: Major contributions on iterative decoding and convergence analysis (Part 1)

| Year | Author(s) | Contribution |
|------|-----------|--------------|
| 1962 | Galager [79] | Invented LDPC codes. |
| 1966 | Forney [125] | Proposed the first concatenated coding scheme using a convolutional code as the inner code and a Reed–Solomon code as the outer code. |
| 1993 | Berrou *et al.* [69] | Invented turbo codes, which achieved near-Shannon limit performance by employing iterative decoding. |
| 1995 | Douillard *et al.* [82] | Proposed turbo equalization, which was capable of eliminating channel-induced inter-symbol interference by performing channel equalization and channel decoding iteratively. |
|  | Divsalar and Pollara [126] | Extended the turbo principle to multiple parallel concatenated codes. |
| 1996 | Benedetto and Montorsi [99, 127, 128] | Proposed an analytical bounding technique for evaluating the bit error probability of a parallel concatenated coding scheme [99], and also extended the turbo principle to serially concatenated block and convolutional codes [127, 128]. |
|  | Hagenauer *et al.* [72] | Studied the log-likelihood algebra of iterative decoding algorithms, which were generalized for employment in any linear binary systematic block code. |
|  | MacKay and Neal [76] | 'Re-discovered' LDPC codes [79] and proposed a new class of LDPC codes constructed on the basis of sparse random parity check matrices, which was capable of achieving near-Shannon limit performance. |
| 1997 | Benedetto *et al.* [85] | Investigated iterative decoding of an outer convolutional decoder and an inner Trellis-Coded Modulation (TCM) decoder, originally devised by Ungerboeck in 1982 [129]. |
| 1998 | Benedetto *et al.* [71, 130] | Derived design guidelines for serially concatenated codes with the aid of the union-bound technique [71], and also extended the turbo principle to multiple serially concatenated codes [130]. |

Table 1.3: Continued

| Year | Author(s) | Contribution |
|------|-----------|--------------|
| 1998 | Moher [87] | Proposed an iterative multiuser decoder for near-capacity communications. |
| | Reed *et al.* [88] | Investigated iterative multiuser detection designed for Direct-Sequence Code Division Multiple Access (DS-CDMA). |
| | Kschischang and Frey [95] | Studied iterative decoding based on factor graph models. |
| | Divsalar *et al.* [75] | Proposed Repeat–Accumulate (RA) codes, where turbo-like performance was achieved at a low complexity. |
| | Pyndiah [73] | Proposed turbo product codes. |
| | ten Brink *et al.* [84] | Studied iterative decoding between an outer channel decoder and an inner soft-demapper in the context of BICM. |
| | McEliece *et al.* [92] | Discovered the connection between turbo decoding algorithms and Pearl's belief propagation algorithm [91]. |
| 1999 | Wang and Poor [90] | Proposed iterative interference cancelation and channel decoding for coded CDMA schemes. |

the channel's capacity. Throughout the 1960s and 1970s the cut-off rate was regarded as a practical design objective, while approaching the channel capacity was deemed to be an unrealistic design goal. The escalating complexity of CC decoding trees may be avoided by employing the more convenient trellises [141] to represent CCs and by applying the Bahl–Cocke–Jelinek–Raviv (BCJR) [48] algorithm. This forms the basis of turbo codes designed by Berrou *et al.* [142] in 1993, which employ a pair of iteratively decoded CCs and facilitate operation near the channel's capacity without an escalating computational complexity, as described in Section 4.1. The advent of turbo coding led to the 'rediscovery' of Low Density Parity Check (LDPC) codes [143–145], which apply iterative decoding techniques to bipartite graphs [146] in order to facilitate operation near the channel's capacity. Recently, irregular coding techniques [147–149] have enabled 'very-near-capacity' operation, as described in Section 5.5. Indeed, in Chapters 7–9 of this volume we demonstrate a novel application of IrVLCs to near-capacity operation in this manner.

Near-capacity operation is of particular interest in digital multimedia broadcast applications. This is because the channel's capacity reduces with the channel's Signal-to-Noise Ratio (SNR), which depends on the transmit power, on the distance to the remote receiver and on the amount of noise that is imposed upon the signal. Hence, multimedia broadcast schemes having a particular information transmission rate are associated with a minimum threshold

Table 1.4: Major contributions on iterative decoding and convergence analysis (Part 2)

| Year | Author(s) | Contribution |
|---|---|---|
| 1999 | ten Brink [106] | Proposed EXIT charts for analyzing the convergence behavior of iterative decoding and demapping [106]. |
| 2000 | ten Brink [111, 112, 115] | Applied EXIT charts to the convergence analysis of various iterative decoding algorithms, such as iterative decoding of PCCs [111] and Serially Concatenated Codes (SCCs) [112], and to turbo equalization [115]; also proposed a common framework of EXIT chart analysis in [115]. |
| 2001 | Richardson and Urbanke [77] | Proposed the density evolution algorithm for analyzing the convergence behavior of LDPC decoding. |
|  | Richardson *et al.* [78] | Employed the density evolution algorithm for the sake of designing capacity-achieving irregular LDPCs. |
|  | Kschischang *et al.* [93] | Studied the sum–product algorithm with the aid of factor graphs. |
|  | ten Brink [119] | Extended EXIT charts from 2D to 3D for analyzing the convergence behavior of three-stage parallel concatenated codes. |
|  | Grant [123] | Extended EXIT charts to non-binary decoders, where mutual information was calculated based on histogram measurements. |
| 2002 | Tüchler and Hagenauer [80] | Proposed irregular convolutional codes as well as a simplified method for computing mutual information in EXIT chart analysis. |
|  | Tüchler [120] | Applied 3D EXIT charts to the convergence analysis of three-stage serially concatenated turbo equalization schemes. |
|  | Tüchler *et al.* [113, 131] | Proposed turbo MMSE equalization [131] and applied EXIT chart analysis to it [113]. |
| 2003 | Brännström *et al.* [121, 122] | Proposed a 3D-to-2D projection technique for facilitating the EXIT chart analysis of multiple serially concatenated codes. Additionally, a search algorithm was proposed for finding the optimal decoder activation order. |

Table 1.4: Continued

| Year | Author(s) | Contribution |
|------|-----------|--------------|
| 2004 | Ashikhmin *et al.* [116] | Presented a universal model for conducting EXIT chart analysis as well as various properties of EXIT functions, which form the theoretical foundation of EXIT-curve matching techniques. |
| | Land *et al.* [132] | Provided a low-complexity mutual information estimation method, which significantly simplifies the EXIT chart analysis. |
| 2006 | Kliewer *et al.* [124] | Proposed an efficient and low-complexity method of computing non-binary EXIT charts from index-based *a posteriori* probabilities, which may be considered as a generalization of the approach presented in [132]. |

SNR at which the channel's capacity drops below the particular information transmission rate to be maintained. If a multimedia broadcast scheme can be designed to support near-capacity operation, then the threshold SNR required to achieve high-quality reception is reduced.

### 1.2.2   Joint Source and Channel Coding

Shannon's source and channel coding separation theorem [133] states that the removal of undesirable redundancy during source coding and the reintroduction of specifically designed intentional redundancy during channel coding can be performed separately, without jeopardizing our ability to achieve an infinitesimally low probability of transmission errors, while maintaining a near-capacity information transmission rate. However, Shannon's findings are only valid under a number of idealistic assumptions [150], namely that the information is transmitted over an uncorrelated non-dispersive narrowband Additive White Gaussian Noise (AWGN) channel, while potentially imposing an infinite decoding complexity and buffering latency. These assumptions clearly have limited validity for practical finite-delay transmissions over realistic fading wireless channels [151].

Additionally, Shannon assumed that the source is stationary and is losslessly encoded with the aid of entropy-encoded symbols having equal significance and identical error sensitivity. These assumptions have a limited validity in the case of multimedia transmission, since video, image, audio and speech information is typically non-stationary, having characteristics that vary in time and/or space [151, 152]. Furthermore, typically lossy multimedia coding [151, 152] is employed in order to achieve a high level of compression and a concomitant low bandwidth requirement, while exploiting the psycho-visual properties of the human vision or hearing. Finally, the components of the encoded multimedia information typically have varying perceptual significance and a corresponding unequal error sensitivity, since they are often generated using several diverse encoding techniques. For example, video coding typically employs the Discrete Cosine Transform (DCT) [153], Motion Compensation

(MC) [154] and entropy coding [155], as exemplified in the MPEG-1 [156], MPEG-2 [157], MPEG-4 [158], H.261 [159], H.263 [160] and H.264 [161] video codecs.

Hence, the employment of joint source and channel coding [162, 163] is motivated. This may be achieved using diverse methods, which we now briefly discuss.

'Channel-optimized' source coding [163] may be employed to reduce the reconstruction error that results when the channel decoder is unable to correct all transmission errors. Here, the source encoder is designed with special consideration of the transmission errors that are most likely to occur, namely those causing a particular binary codeword to be confused with another similar codeword. In this way, channel-optimized source encoding allocates pairs of similar codewords to represent similar reconstructed source parameter values. For example, this may be applied to scalar quantization [164, 165], where real-valued source samples are represented by one of a number of discrete quantization levels, which are indexed by binary codewords. In this way, the allocation of binary codeword pairs having a low Hamming distance to represent quantization level pairs having a low Euclidean distance was demonstrated in [166, 167]. Similarly, the authors of [168, 169] proposed Channel-Optimized Vector Quantization (COVQ), which employs the channel-optimized indexing of Vector Quantization (VQ) tiles [170]. Note that both scalar and vector quantization are discussed in greater detail in Section 3.3.

Joint source and channel coding can also be beneficially employed if some of the source correlation is not removed during source encoding, resulting in the manifestation of residual redundancy [171] within the resultant source-encoded signal. In the receiver, a model of the source correlation may be employed to exploit the residual redundancy in order to provide an error-correction capability, which may be employed to mitigate any transmission errors that could not be eliminated during channel decoding. The error-correction capability constituted by the source-encoded signal's residual redundancy may be invoked during source decoding in order to provide a Minimum Mean Square Error (MMSE) estimate of the source signal [172–175]. Alternatively, a Maximum *A Posteriori* Probability (MAP) estimate [172, 173, 176] of the source-coded signal may be obtained prior to source decoding by exploiting the aforementioned residual redundancy. Note that MMSE and MAP decoders have the ability to consider the entire source-coded signal sequence at once [175–177], or they may consider the individual symbols of the source-coded signal sequence separately [172–174].

In the same way that residual source redundancy can be exploited in the receiver to provide an error-correction capability, so can redundancy that is intentionally introduced during source encoding. For example, this may be achieved by imposing correlation on the source signal with the aid of Trellis-Coded Quantization (TCQ) [178], as exemplified in [177]. Alternatively, Variable-Length Error Correction (VLEC) coding [179] may be employed to incorporate redundancy within the source-encoded signal. This may be exploited to provide an error correction capability in the receiver [180–183], as detailed in Section 3.3. In Section 4.1 we detail the turbo principle [142], which may be employed for supporting iterative joint source and channel decoding [184–186], where the redundancy introduced by both source and channel coding is alternately exploited so that each provides *a priori* information for the other concerning the source-encoded signal. A novel scheme employing this approach is described in Chapter 6, and hence the detailed literature review of this topic is postponed to Section 6.1.

A further approach to joint source and channel coding that we highlight here employs sophisticated rate allocation in order to jointly optimize the amount of redundancy that is retained or introduced during source encoding and the amount that is introduced during channel encoding. In the scenario where a lossy source codec is employed, an increased degree of lossy compression may be achieved by employing coarser quantization, for example. While this results in additional quantization-induced reconstruction distortion, the associated reduction in the amount of resultant source-coded information facilitates the employment of a lower-rate channel codec, without increasing the overall transmission rate. Since lower-rate channel codecs are associated with higher error-correction capabilities, they are capable of mitigating more channel-induced reconstruction distortion. Hence, rate allocation may be employed to optimize the trade-off between quantization- and channel-induced reconstruction distortion [167, 187, 188].

### 1.2.3 Unequal Error Protection

As mentioned above, this book considers the application of IrVLCs for UEP. In a manner similar to that of [189–191], for example, UEP may be employed to appropriately protect audio-, speech-, image- and video-encoded bit sequences, which are typically generated using diverse encoding techniques and exhibit various error sensitivities. For example, video coding typically employs the DCT and MC, as described above. As noted in [151], typically a higher degree of video reconstruction distortion is imposed by transmission errors that affect the motion vectors of MC than from those inflicted on the DCT-encoded information. Hence, UEP may be employed to protect all MC-related information with a relatively strong error-correction capability, whilst employing a relatively weak error-correction code to protect the DCT-encoded information. This approach may hence be expected to yield a lower degree of video reconstruction distortion than equal protection, as noted in [189–191], for example. In Chapter 7 we demonstrate the novel application of IrVLCs for UEP by employing different sets of VLEC codewords having various error-correction capabilities to appropriately protect particular components of the source signal that have different error sensitivities.

## 1.3 Motivation and Methodology

In the spirit of the remarkable performance of turbo codes, as described in Section 1.1, the turbo principle has been successfully applied to various generalized decoding/detection problems, resulting in diverse turbo transceivers [70]. In the meantime new challenges have arisen: for example, how to choose the parameters of a turbo transceiver consisting of multiple components so that near-capacity performance can be achieved. Owing to iterative decoding, different components of the turbo transceivers are combined and affect each other during each iteration. Hence, in contrast to the design of conventional non-iterative transceivers, the optimization should be considered jointly rather than individually. Moreover, as argued in Section 1.1, EXIT charts have shown their power in analyzing the convergence behavior of iterative receivers. Therefore, in this treatise we mainly rely on this semi-analytical tool as well as on the conventional Monte Carlo simulation method in our attempt to tackle the problem of optimizing multi-stage turbo transceivers and JSCD schemes.

## 1.4 Outline of the Book

Having briefly reviewed the literature of source coding, joint source–channel decoding, concatenated coding and iterative decoding, as well as convergence analysis, let us now outline the organization of the book.

– **Chapter 2: Information Theory Basics**
  In Chapter 2 we commence by laying the information-theoretic foundations for the rest of the book.

– **Chapter 3: Sources and Source Codes**
  This chapter introduces various entropy coding schemes designed for memoryless sources having known alphabets, along with several trellis-based soft-decision decoding methods in contrast to conventional hard-decision decoding. Section 3.2 describes the source models used in this treatise; various source coding schemes designed for memoryless sources are introduced in Section 3.3, such as Huffman codes, Reversible Variable-Length Codes (RVLCs) and Variable-Length Error Correction (VLEC) codes. A generic algorithm contrived for efficiently constructing both RVLCs and VLEC codes having different free distances is also presented. In Section 3.4.1 two different trellis representations of source codes are introduced, together with their conventional binary tree representation. Based on these trellis representations, the Maximum Likelihood Sequence Estimation (MLSE) and MAP decoding algorithm are derived for soft-decision-aided source decoding in Sections 3.4.2.1 and 3.4.2.2 respectively. Finally, the source Symbol-Error Ratio (SER) performance of these decoding algorithms is presented in Section 3.4.3.

– **Chapter 4: Iterative Source–Channel Decoding**
  Based on the soft VLC decoding algorithms described in Chapter 3, various Iterative Source–Channel Decoding (ISCD) schemes may be formed. The SER performance of these schemes contrived for communicating over both AWGN channels and dispersive channels is quantified and their convergence behavior is analyzed with the aid of EXIT charts. More explicitly, Section 4.1 provides an overview of various code concatenation schemes and the so-called 'turbo principle.' A generic SISO *A Posteriori* Probability (APP) decoding module, which serves as the 'kernel' of iterative decoding schemes, is described in Section 4.2. Additionally, the semi-analytical EXIT chart tool for convergence analysis is introduced in Section 4.2.2. Section 4.3 evaluates the performance of ISCD schemes designed for transmission over AWGN channels, along with the EXIT chart analysis of their convergence behavior. More particularly, the effects of different source codes are considered. Furthermore, the performance of ISCD schemes is investigated for transmission over dispersive channels in Section 4.4, which necessitates the employment of channel equalization. Both scenarios, with and without channel precoding, are considered in Sections 4.4.2 and 4.4.3 respectively. A three-stage iterative channel equalization, channel decoding and source decoding scheme is presented in Section 4.4.4.

– **Chapter 5: Three-Stage Serially Concatenated Turbo Equalization**
  In this chapter a three-stage serially concatenated turbo MMSE equalization scheme is investigated. The different components of the iterative receiver are jointly optimized with the aid of EXIT charts for the sake of achieving a near-capacity performance.

More specifically, Section 5.2 provides a short introduction to SISO MMSE equalization. The conventional two-stage turbo equalization scheme is described in Section 5.3, while Section 5.4 presents a three-stage turbo equalization scheme. Both 3D and 2D EXIT chart analysis are carried out in Section 5.4.2, where the channel code is optimized for obtaining the lowest possible $E_b/N_0$ convergence threshold and the activation order of the component decoders is optimized for achieving the fastest convergence, while maintaining a low decoding complexity. Various methods of visualizing the three-stage iterative decoding process are described in Section 5.4.4. The effects of different interleaver block lengths on the attainable system performance are discussed in Section 5.4.5. Section 5.5 demonstrates the technique of EXIT chart matching and the design of IRregular Convolutional Codes (IRCCs), which results in a capacity-approaching three-stage turbo equalization scheme. In addition to unity-rate intermediate codes, the employment of non-unity-rate intermediate codes in three-stage schemes is investigated in Section 5.6.

– **Chapter 6: Irregular Variable-Length Codes for Joint Source and Channel Coding** In Chapter 6 we demonstrate the application of IrVLCs for the joint source and channel coding of video information. The proposed scheme employs the serial concatenation and iterative decoding of a video codec with a channel codec. Our novel video codec represents the video information using Variable-Dimension Vector Quantization (VDVQ) tiles. The VDVQ tiles employed are represented using the corresponding RVLC codewords selected from the VDVQ/RVLC codebook. However, the legitimate use of the VDVQ tiles and their corresponding RVLC codewords is limited by a number of code constraints, which ensure that the VDVQ tiles employed perfectly tessellate, among other desirable design objectives. As a result, different subsets of the RVLC codewords are available at different points during the encoding of the video information, and the proposed approach adopts an IrVLC philosophy.

In the video codec of Chapter 6, the VDVQ/RVLC-induced code constraints are uniquely and unambiguously described by a novel VDVQ/RVLC trellis structure, which resembles the symbol-based VLEC trellis of [181, 192]. Hence, the employment of the VDVQ/RVLC trellis structure allows the consideration of all legitimate transmission frame permutations. This fact is exploited in the video encoder in order to perform novel MMSE VDVQ/RVLC encoding, using a novel variant of the Viterbi algorithm [193].

Additionally, the employment of the VDVQ/RVLC trellis structure during video decoding guarantees the recovery of legitimate – although not necessarily error-free – video information. As a result, the video decoder never has to discard video information. This is unlike conventional video decoders, where a single transmission error may render an entire transmission frame invalid. Furthermore, a novel version of the BCJR algorithm [48] is employed during APP SISO VDVQ/RVLC decoding in order to facilitate the iterative exchange of soft information with the serially concatenated channel decoder, and in order to perform the soft MMSE reconstruction of the video sequence. Finally, since the VDVQ/RVLC trellis structure describes the complete set of VDVQ/RVLC-induced code constraints, all of the associated redundancy is beneficially exploited with the aid of the modified BCJR algorithm.

Owing to its aforementioned benefits and its employment of a joint source and channel coding philosophy, the video transmission scheme of Chapter 6 is shown to outperform the corresponding benchmarks that employ a separate source and channel coding philosophy. Our findings were originally published in [194, 195].

– **Chapter 7: Irregular Variable-Length Codes for EXIT-Chart Matching**
  In Chapter 7, we investigate the application of IrVLCs for Unequal Error Protection (UEP). Here, a number of component VLC codebooks having different error-correction capabilities are employed to encode various fractions of the source symbol frame. When these fractions have different error sensitivities, this approach may be expected to yield a higher reconstruction quality than does equal protection, as noted in [189–191], for example.

  Chapter 7 also investigates the application of IrVLCs to near-capacity operation. Here, a number of component VLC codebooks having different inverted EXIT functions are employed to encode various fractions of the source symbol frame. We show that the inverted IrVLC EXIT function may be obtained as a weighted average of the inverted component VLC EXIT functions. Additionally, the EXIT-chart matching algorithm of [149] is employed to shape the inverted IrVLC EXIT function to match the EXIT function of a serially concatenated inner channel code and to create a narrow but still open EXIT chart tunnel. In this way, iterative decoding convergence to an infinitesimally low probability of error is facilitated at near-capacity SNRs.

  Furthermore, in Chapter 7 the UEP and near-capacity operation of the described scheme is assessed using novel plots that characterize the computational complexity of iterative decoding. More specifically, the average number of Add, Compare and Select (ACS) operations required to reconstruct each source symbol with a high quality is plotted against the channel SNR. These plots are employed to compare the novel IrVLC-based scheme with a suitably designed IrCC and regular VLC-based benchmarks, quantifying the advantages of the IrVLCs. Furthermore, these plots demonstrate that the complexity associated with the bit-based VLEC trellis is significantly lower than that of the corresponding symbol-based trellis. Our findings were originally published in [196, 197] and we proposed attractive near-capacity IrVLC schemes in [198–203].

– **Chapter 8: Genetic Algorithm-Aided Design of Irregular Variable-Length Coding Components**
  In Chapter 8 we introduce a novel Real-Valued Free Distance Metric (RV-FDM) as an alternative to the Integer-Valued Free Distance (IV-FD) for the characterization of the error-correction capability that is associated with VLEC codebooks. Unlike the IV-FD lower bound, the RV-FDM assumes values from the real-valued domain, hence allowing the comparison of the error-correction capability of two VLEC codebooks having equal IV-FD lower bounds. Furthermore, we show that a VLEC codebook's RV-FDM affects the number of inflection points that appear in the corresponding inverted EXIT function. This complements the property [204] that the area below an inverted VLEC EXIT function equals the corresponding coding rate, as well as the property that a free distance of at least two yields an inverted VLEC EXIT function that reaches the top right-hand corner of the EXIT chart.

These properties are exploited by a novel Genetic Algorithm (GA) in order to design beneficial VLEC codebooks having arbitrary inverted EXIT function shapes. We will demonstrate that this is in contrast to the methods of [205–207], which are incapable of designing codebooks having specific EXIT function shapes without imposing a significant level of 'trial-and-error'-based human interaction. This novel GA is shown to be attractive for the design of IrVLC component codebooks for EXIT-chart matching, since Chapter 8 also demonstrates that our ability to create open EXIT-chart tunnels at near-capacity channel SNRs depends on the availability of a sufficiently diverse suite of component codes having a wide variety of EXIT function shapes.

Finally, a suite of component VLEC codebooks designed by the proposed GA is found to facilitate higher-accuracy EXIT-chart matching than a benchmark suite designed using the state-of-the-art method of [207]. Our novel RV-FDM and GA were originally published in [200, 201].

– **Chapter 9: Joint EXIT-Chart Matching of Irregular Variable-Length Coding and Irregular Unity-Rate Coding**
In Chapter 9, we propose a novel modification of the EXIT-chart matching algorithm introduced in [149] that additionally seeks a reduced APP SISO decoding complexity by considering the complexities associated with each of the component codes. Another novel modification introduced in this chapter facilitates the EXIT-chart matching of irregular codes that employ a suite of component codes having the same coding rate.

Additionally, Chapter 9 demonstrates the joint EXIT-chart matching of two serially concatenated irregular codecs, namely an outer IrVLC and an inner Irregular Unity-Rate Code (IrURC). This is achieved by iteratively matching the inverted outer EXIT function to the inner EXIT function and vice versa. By employing an irregular inner code in addition to an irregular outer code, we can afford a higher degree of design freedom than the proposals of [149], which employ a regular inner code. Hence, the proposed approach is shown to facilitate even nearer-capacity operation, which is comparable to that of Irregular Low-Density Parity Checks (IrLDPCs) and irregular turbo codes. Our findings were originally published in [202, 203] and we additionally demonstrated the joint EXIT-chart matching of serially concatenated irregular codecs in [208].

– **Chapter 10: Iteratively Decoded VLC Space–Time Coded Modulation**
In this chapter an Iteratively Decoded Variable-Length Space–Time Coded Modulation (VL-STCM-ID) scheme capable of simultaneously providing both coding and iteration gain as well as multiplexing and diversity gain is investigated. Non-binary unity-rate precoders are employed for assisting the iterative decoding of the VL-STCM-ID scheme. The discrete-valued source symbols are first encoded into variable-length codewords that are spread to both the spatial and the temporal domains. Then the variable-length codewords are interleaved and fed to the precoded modulator. More explicitly, the proposed VL-STCM-ID arrangement is a jointly designed iteratively decoded scheme combining source coding, channel coding and modulation as well as spatial diversity/multiplexing. We demonstrate that, as expected, the higher the source correlation, the higher the achievable performance gain of the scheme becomes. Furthermore, the performance of the VL-STCM-ID scheme is about 14 dB better than

that of the Fixed-Length STCM (FL-STCM) benchmark at a source symbol error ratio of $10^{-4}$.

- **Chapter 11: Iterative Detection of Three-Stage Concatenated IrVLC FFH-MFSK**
  Serially concatenated and iteratively decoded Irregular Variable-Length Coding (IrVLC) combined with precoded Fast Frequency Hopping (FFH) $M$-ary Frequency Shift Keying (MFSK) is considered. The proposed joint source and channel coding scheme is capable of operating at low Signal-to-Noise Ratio (SNR) in Rayleigh fading channels contaminated by Partial Band Noise Jamming (PBNJ). By virtue of its inherent time and frequency diversity, the FFH-MFSK scheme is capable of mitigating the effects of Partial Band Noise Jamming (PBNJ). The IrVLC scheme comprises a number of component variable-length coding codebooks employing different coding rates for encoding particular fractions of the input source symbol stream. These fractions may be chosen with the aid of EXtrinsic Information Transfer (EXIT) charts in order to shape the inverted EXIT curve of the IrVLC codec so that it can be matched with the EXIT curve of the inner decoder. We demonstrate that, using the proposed scheme, an infinitesimally low bit error ratio may be achieved at low SNR values. Furthermore, the system employing a unity-rate precoder yields a substantial SNR gain over that dispensing with the precoder, and the IrVLC-based scheme yields a further gain over the identical-rate single-class VLC-based benchmark.

- **Chapter 12: Conclusions and Future Research**
  The major findings of our work are summarized in this chapter along with our suggestions for future research.

## 1.5   Novel Contributions of the Book

Whilst it may appear somewhat unusual to outline the novel contribution of a book in a style similar to that found in research papers, this research monograph does indeed collate a large body of original contributions to the current state of the art. This justifies the inclusion of the following list.

- A novel and high-efficiency algorithm is proposed for constructing both RVLCs and VLEC codes [32]. For RVLCs the algorithm results in codes of higher efficiency and/or shorter maximum codeword length than the algorithms previously disseminated in the literature. For VLEC codes it significantly reduces the construction complexity imposed, in comparison with the original algorithm of [29, 30].
- Soft VLC decoding based on both symbol-level and bit-level trellis representation is investigated. The performance of different decoding criteria such as MLSE and MAP is evaluated. In particular, the effect of different free distances of the source codes is quantified.
- A novel iterative channel equalization, channel decoding and source decoding scheme is proposed [209, 210]. It is shown that without using channel coding the redundancy in the source code alone is capable of eliminating the channel-induced Inter-Symbol Interference (ISI), provided that the channel equalization and source decoding are performed iteratively at the receiver. When employing channel coding, a three-stage joint source–channel detection scheme is formed, which is capable of outperforming the separate source–channel detection scheme by 2 dB in terms of the $E_b/N_0$ value

required for achieving the same BER. Furthermore, with the aid of EXIT charts, the systems' convergence thresholds and the design of channel precoders are optimized.

- A novel three-stage serially concatenated MMSE turbo equalization scheme is proposed [211]. The performance of conventional two-stage turbo equalization schemes is generally limited by a relatively high BER 'shoulder' [212]. However, when employing a unity-rate recursive convolutional code as the intermediate component, the three-stage turbo equalization scheme is capable of achieving a steep BER turbo 'cliff,' beyond which an infinitesimally low BER is obtained.

- The concept of EXIT modules is proposed, which significantly simplifies the EXIT chart analysis of multiple concatenated systems, casting them from the 3D domain to the 2D domain. As a result, the outer constituent code of the three-stage turbo equalization scheme is optimized for achieving the lowest possible convergence threshold. The activation order of the component decoders is optimized for obtaining the convergence at the lowest possible decoding complexity using the lowest total number of iterations or decoder activations.

- A near-channel-capacity three-stage turbo equalization scheme is proposed [213]. The proposed scheme is capable of achieving a performance within 0.2 dB of the dispersive channel's capacity with the aid of EXIT-chart matching and IRCCs. This scheme is further generalized to MIMO turbo detection [214].

- The employment of non-unity-rate intermediate codes in a three-stage turbo equalization scheme is also investigated. The maximum achievable information rates of such schemes are determined and the serial concatenation of the outer constituent code and the intermediate constituent code is optimized for achieving the lowest possible decoding convergence threshold.

- As an application example, a novel vector quantized RVLC-TCM scheme is proposed for the iterative joint source and channel decoding of video information, and its RVLC symbol-based trellis structure is investigated. Furthermore, a novel version of the BCJR algorithm is designed for the APP SISO VDVQ/RVLC decoding of the video information.

- IrVLC schemes are designed for near-capacity operation and their complexity versus channel SNR characteristics are recorded. The corresponding relationships are also parameterized by the achievable reconstruction quality.

- The error-correction capability of VLECs having the same Integer-Valued Free Distance (IV-FD) is characterized and the relationship between a VLEC's IV-FD and the shape of its inverted EXIT function is established.

- A novel Genetic Algorithm (GA) is devised for designing VLECs having specific EXIT functions.

- A suite of VLECs that are suitable for a wide range of IrVLC applications are found.

- A novel EXIT-chart matching algorithm is devised for facilitating the employment of component codes having the same coding rate.

- This EXIT-chart matching algorithm is further developed to additionally seek a reduced APP SISO-decoding computational complexity.

- A joint EXIT-chart matching algorithm is developed for designing schemes employing a serial concatenation of an irregular outer and inner codec. This results in IrVLC–IrURC schemes designed for near-capacity joint source and channel coding.

- A number of VLC-based system design examples are provided, including a three-stage concatenated VLC-aided space–time scheme.
- As a second VLC-based design example, an IrVLC-coded and unity-rate precoder-assisted noncoherently detected Fast Frequency Hopping $M$-ary Frequency Shift Keying (FFH-MFSK) system communicating in an uncorrelated Rayleigh fading channel contaminated by PBNJ is conceived and investigated. This system may find application in low-complexity noncoherent detection-aided cooperative systems.
- We investigate the serial concatenation of the FFH-MFSK demodulator, unity-rate decoder and IrVLC outer decoder with the aid of EXIT charts, in order to attain a good performance even at low SNR values.
- We contrast the two-stage iterative detection scheme, exchanging extrinsic information between the inner rate-1 decoder and the IrVLC outer decoder, with the three-stage extrinsic information exchange among the demodulator, inner decoder and outer decoder.