

Trends in Channel Coding for 6G

This article provides an overview of the evolution and future directions of error-correction coding (channel coding) in wireless communication systems, emphasizing its critical role from 2G to the upcoming 6G era.

By SISI MIAO^{ID}, Student Member IEEE, CLAUS KESTEL^{ID}, Graduate Student Member IEEE, LUCAS JOHANNSEN^{ID}, MARVIN GEISELHART^{ID}, Graduate Student Member IEEE, LAURENT SCHMÄLEN^{ID}, Fellow IEEE, ALEXIOS BALATSOUKAS-STIMMING^{ID}, Member IEEE, GIANLUIGI LIVA^{ID}, Senior Member IEEE, NORBERT WEHN, AND STEPHAN TEN BRINK^{ID}, Fellow IEEE

ABSTRACT | Error correction coding (i.e., channel coding) is a key ingredient of any digital communications system. In mobile wireless communications, channel codes have evolved from simple convolutional codes in Global System for Mobile Communications (GSM) (2G), parallel concatenated (turbo) codes in Universal Mobile Telecommunications Service (UMTS) (3G), and long-term evolution (LTE) (4G), to carefully designed multirate/multilength low-density parity-check (LDPC) codes in 5G, combined with polar codes for short messages on the synchronization channel. Based on this rich history, and by accounting for the technological advances in very large-scale integration, this article will outline some recent trends in channel coding as they may be applied in 6G systems, ranging from novel approaches for short blocklengths such as automorphism ensemble decoding, via ideas of coding for multiple access, to concepts for unified coding schemes that may simplify encoding/decoding hardware at competitive error-correcting performance.

KEYWORDS | 6G; channel coding; wireless communications.

I. INTRODUCTION

The paradigm shift caused by the introduction of turbo codes [32] and the rediscovery of low-density parity-check (LDPC) codes [85], [155], [229] had a profound impact on the development of wireless communication systems. Since the last decade of the past century, the revolutionary principles at the foundations of turbo and LDPC codes provided a practical means to approach channel capacity [217], yielding tremendous gains in terms of energy and spectral efficiency in wireless channels. Turbo and LDPC codes currently dominate the landscape of channel codes in modern radio communication standards. Turbo codes play a prominent role in the 3rd and 4th generations of the 3rd Generation Partnership Project (3GPP) cellular standard [1], [3], [4], as well as in satellite/space communications [45], [72]. LDPC codes replaced turbo codes in the 3GPP 5G new radio (NR) standard [5] and have been adopted in space communications [45], video broadcasting systems [71], [73], and the wireless local area network (WLAN) standard [105]. Polar codes, introduced more recently [21], joined LDPC codes in the family of channel codes that lay at the core of 5G NR [5], providing support for the control channel of the standard.

The ground-breaking progress made in coding theory during the past three decades is inextricably intertwined with the enormous progress in microelectronics driven by Moore's law [106]. This progress combined with efficient decoder architectures turned the promises of coding theory into hardware implementations that populate all modern wireless communication devices. In fact, a deeper understanding of the implementation challenges has been fundamental in guiding coding theorists in the exploration of turbo, LDPC, and polar code structures that can

Manuscript received 2 January 2024; revised 10 May 2024; accepted 4 June 2024. Date of publication 26 June 2024; date of current version 30 October 2024. This work was supported in part by German Federal Ministry of Education and Research (BMBF) under Grant 16KISK004, Grant 16KISK010, Grant 16KISK019 (Open6GHz), and Grant 16KISK022 (6G-RIC); and in part by Dutch Research Council (NWO) through the Open Technology Programme (OTP) BIT-FREE Project. (Corresponding author: Stephan ten Brink.)

Sisi Miao and **Laurent Schmalen** are with the Communications Engineering Lab, Karlsruhe Institute of Technology (KIT), 76187 Karlsruhe, Germany.

Claus Kestel and **Norbert Wehn** are with the Division of Microelectronic Systems Design, University of Kaiserslautern-Landau, 67663 Kaiserslautern, Germany.

Lucas Johannsen is with the Hochschule Koblenz, Koblenz Karthause, 56075 Koblenz, Germany.

Marvin Geiselhart and **Stephan ten Brink** are with the Institute of Telecommunications, University of Stuttgart, 70569 Stuttgart, Germany (e-mail: tenbrink@inue.uni-stuttgart.de).

Alexios Balatsoukas-Stimming is with the Signal Processing for Communications Lab, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands.

Gianluigi Liva is with the Institute of Communications and Navigation of the German Aerospace Center (DLR), 82234 Wessling, Germany.

Digital Object Identifier 10.1109/JPROC.2024.3416050

0018-9219 © 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See <https://www.ieee.org/publications/rights/index.html> for more information.

approach the ultimate performance limits, enabling at the same time the use of efficient hardware architectures. In the following, we review some of the milestones in coding theory, with a time horizon that begins from the early 1990s (what we call the “mobile Internet era”). Connections between the new developments and their adoption in wireless communications will be established, paying particular attention to engineering aspects that led to practical, efficient implementations.

On the Notation: In the following, we will consider (n, k) binary linear block codes, where n is the blocklength (i.e., the length of codewords) in bits and k is the code dimension. The code rate will be denoted by $R = k/n$. We will also use $O(\cdot)$ for Landau’s big- O notation.

II. CHANNEL CODING FOR THE MOBILE INTERNET ERA

The history of coding theory is the history of a fascinating intellectual endeavor that started in the 1940s, around the time when Shannon [217] laid the foundations of information theory. The search for channel codes capable of achieving the limits set in [217] involved important contributions from several domains, well beyond coding theory, including discrete mathematics, theoretical computer science, and physics, among others. For a wide-spectrum historical perspective on the achievements in channel coding from the outset, we refer the interested reader to the extensive surveys in [52], [53], and [88] and the detailed historical notes provided at the end of the chapters in [200]. In the following, we will focus on the most recent developments, starting from the turbo code revolution of the early 1990s. In particular, attention will be paid to techniques that are relevant to wireless communications standards. Therefore, recent developments in the rich and elegant field of algebraic coding theory will be mostly excluded from this survey. Fig. 1 may serve the reader as a timeline of the major advances discussed in the following.

A. Iterative Decoding Revolution

The dawn of what is often referred to as *modern coding theory* coincides with two major developments that occurred during the last decade of the past century. The introduction of turbo codes [32] surprised the channel coding community, providing a first practical means for constructing error correction codes capable of working close to the Shannon limit with moderate-to-low decoding complexity. Shortly after that, LDPC codes, invented in the early 1960s [85] and mostly forgotten in the following 30 years, were rediscovered and shown to compete in performance with turbo codes [155], [229].

The period that immediately followed the introduction of turbo codes and the rediscovery of LDPC codes was characterized by widespread efforts to explain the outstanding performance achieved by the two code classes. Research focused on three main avenues: the development of a unifying theory of iterative decoding algorithms, the introduction of new techniques to study the behavior of

iterative decoders, and the analysis of turbo and LDPC codes from a distance spectrum perspective.

LDPC codes admit a simple graphical representation in terms of Tanner graphs [238], where each codeword bit is associated with a *variable node*, each parity-check equation is associated with a *check node*, and edges connect variable and check nodes according to the constraints defined by the LDPC codes’ parity-check matrix. Due to the low density of their parity-check matrices, the Tanner graphs of LDPC codes are *sparse* bipartite graphs. Tanner graphs can be used to cast the probabilistic iterative decoding algorithm of LDPC codes, originally described in [85], in terms of the more general sum-product algorithm (SPA). Including state variables, factor graphs [127], [258], [259] can be used to describe the trellis of block and convolutional codes, extending the theory of Tanner graphs to provide a unifying framework for the description of turbo and LDPC codes. An almost equivalent formalism was independently proposed in [13]. By instantiating the SPA as a message passing (MP) algorithm over one particular factor graph representation associated with the trellis diagram of a code, one recovers the Bahl–Cocke–Jelinek–Raviv (BCJR) algorithm (also known as *forward–backward* algorithm) [25]. The factor graph of a turbo code is obtained by connecting the factor graphs of the component convolutional codes according to the interleaver specification, and iterative decoding of turbo codes follows by applying the SPA to the overall graph. Due to their generality, Tanner and factor graphs can be used to define a class of codes (the class of *codes on graphs*) that is much wider than the ones of turbo and LDPC codes. The correspondence of the SPA and the belief propagation (BP) algorithm [177], designed for statistical inference in Bayesian networks, was established in [159]. When applied to a factor graph that does not contain cycles, SPA was shown to be optimal in providing exact a posteriori probabilities (APPs) [80], [258], [259]. However, Tanner graph descriptions of good codes must contain cycles, or, conversely, codes whose Tanner graph is cycle-free feature pathologically small minimum distances [74]. The SPA and related (heuristic) simplifications were used to decode product codes [93], [187] and more general classes of codes on graphs, referred to as *generalized* LDPC codes [138].

B. Analysis of Iterative Decoders: EXIT Charts and Density Evolution

In classical coding theory, the analysis of the performance of a code is intimately related to the analysis of its weight enumerator (distance spectrum). This kind of analysis presumes the use of maximum likelihood (ML) or bounded-distance decoding algorithms and uses the weight enumerator in conjunction with union bounds to obtain upper bounds on the bit/block error probability [204]. The analysis of turbo and LDPC codes under BP decoding required the development of a completely different perspective. Taking advantage of the *local* nature

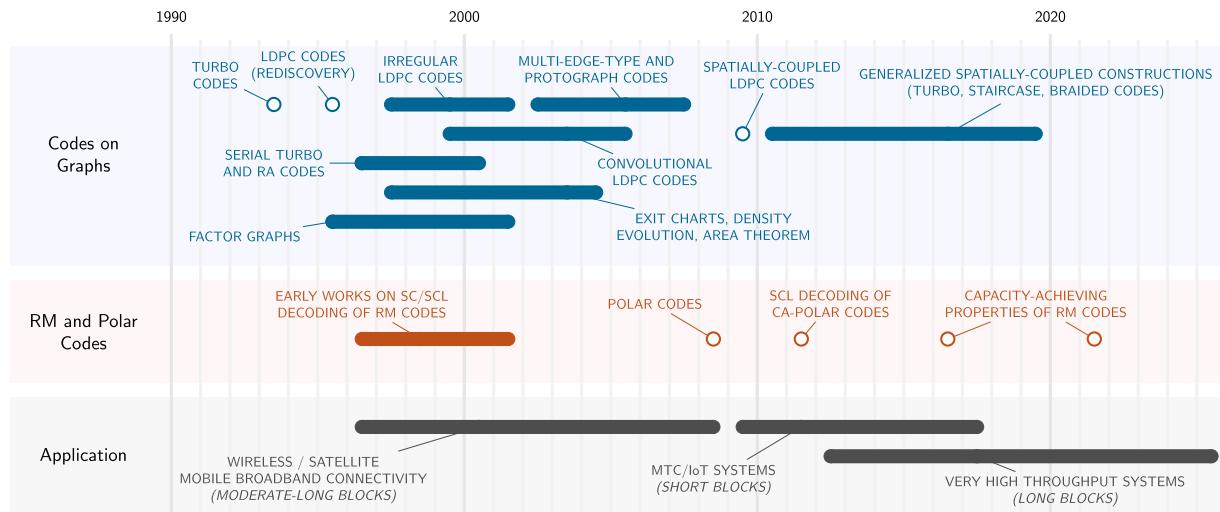


Fig. 1. Timeline of the development of coding theory, starting from the iterative decoding revolution in the early 90s. The emphasis is on the broad class of codes on graphs and polar codes. Applications driving the development of coding techniques, from a requirement point of view, are included in the lower part.

of the BP algorithm, approaches that treated the messages (exchanged between the nodes of the factor graph) as random variables emerged [62], [69], [152], [153], [154], [196], [241], [242], [243], extending the analysis of the binary MP algorithms (*Gallager-A* and *Gallager-B* algorithms) originally introduced in [85]. All these approaches track the evolution of the message distributions to determine the conditions under which the iterative decoder converges to a vanishingly small bit error probability. Among the techniques, extrinsic information transfer (EXIT) analysis played a prominent role thanks to a number of properties [24] that have been central in establishing fundamental results in coding theory. Introduced in [242] as a means to visualize the convergence process in the iterative decoding of turbo codes, EXIT charts [241], [242] provide an approximation of the density evolution analysis [196] over general communication channels. The approximation lies in describing the distribution of the messages (extrinsic estimates) exchanged between the two BCJR decoders of a parallel turbo code through a single parameter, namely, the mutual information between a message and the corresponding codeword bit. Over the additive white Gaussian noise (AWGN) channel, the analysis assumes that the messages follow a distribution, conditioned on the associated bit value, that is Gaussian. On the contrary, the density evolution analysis of [196] does not enforce any specific message distribution, and it tracks the evolution of the actual message density (e.g., by finely quantizing the domain of the density). Despite its simpler form, the EXIT analysis provides extremely accurate estimates of the iterative decoding threshold of turbo code ensembles over several communication channels of practical interest. Furthermore, the EXIT analysis yields exact results over the binary erasure channel (BEC); assuming a uniform distribution at the encoder input, over the BEC the SPA, applied to the factor graph of a binary linear code, results in messages that carry either an *erasure* (full uncertainty) or the correct bit value. Messages follow

a Bernoulli distribution, which can be fully characterized by the message erasure probability or, equivalently, by its complement to 1, i.e., by mutual information. Over the BEC, a fundamental property of EXIT charts arises; the area underlined by the EXIT function of an (n, k) binary linear block code equals $1 - R$ [24]. This result is called the *area theorem* and has become a handy tool for the design of concatenated codes and the optimization of their iterative decoders.

In particular, density evolution and EXIT analysis provided useful tools to optimize turbo and LDPC codes under BP decoding. For the specific case of LDPC codes, the iterative decoding threshold obtained by density evolution analysis was used as a figure of merit to optimize the degrees of the nodes in the code graph. It was observed that LDPC codes with irregular degree distributions can achieve decoding thresholds close to Shannon's limit [49], [154], [196]. In contrast, regular LDPC codes are bound away from the limit under BP decoding. Remarkably, LDPC code degree distribution sequences achieving the capacity of the BEC were designed in [175]. A drawback of the result in [175] is the growth of the average node degrees as the decoding thresholds approach the BEC capacity, resulting in a decoding complexity that increases as the threshold improves. It was realized that this limitation is inherent to LDPC codes whose Tanner graph does not possess state variable nodes [205]. By allowing state variables in the code graph, degree distributions capable of achieving the BEC capacity with bounded complexity per information bit were proposed in [179].

C. Understanding Error Floors: Distance Spectra, Stopping Sets, Trapping Sets, and Pseudo-Codewords

Both density evolution and EXIT analysis provide a characterization of the performance of turbo and LDPC code ensembles, in the limit of large blocklengths. The asymptotic nature of this kind of analysis does not capture

finite-length effects that may be especially visible in the low bit/block error probability regime. By observing that, at low error probabilities, the performance of a turbo/LDPC code is often driven by its minimum distance, the iterative decoder threshold analysis is often complemented by an analysis of the distance properties of the codes. Following the ensemble-based perspective already adopted for the analysis of the distance properties of regular LDPC code ensembles in [85], the *uniform interleaver* approach was used in [30] to derive the average weight enumerator of turbo code ensembles. It soon became clear that *parallel turbo codes* (i.e., turbo codes obtained by parallel concatenation of two recursive convolutional encoders, separated by an interleaver, as originally conceived in [32]) possess small minimum distances. However, the use of large interleavers drastically reduces the number of low-weight codewords, a phenomenon referred to as *spectral thinning* [178]. A fundamental limitation of parallel turbo code ensembles was established later in [38]; the minimum distance of parallel turbo codes can grow only as $O(\log n)$. Although this limitation did not affect the performance of turbo codes at moderate-to-high bit error rates (BERs) (e.g., above 10^{-5}), it caused the emergence of *error floors* at lower error rates [201]. The error floor phenomenon prevented the adoption of parallel turbo codes in applications that demand very low error rates. Serial turbo codes [29] were introduced to circumvent this issue, trading some tenths of a decibel in coding gain for drastically lower error floors. A general treatment of the distance properties of various classes of turbo-like code ensembles was introduced in [63], [113]. A similar analysis was developed for irregular LDPC code ensembles [40], [60], [146], extending the results already available for regular codes in [85]. Using asymptotic enumeration techniques, it was possible to identify conditions on the degree distributions that allow discrimination between *good* LDPC code ensembles, where the minimum distance grows linearly in n and *bad* LDPC code ensembles, where the minimum distance grows sub-linearly in the blocklength, often as $O(\log n)$. Some LDPC codes were shown to suffer from error floors that were not caused by low-weight codewords, but rather by specific subgraph configurations (involving a small number of variable nodes in the Tanner graph) that were able to trap the BP decoder [156]. On the BEC, such configurations are associated with *stopping sets* [59], and their enumeration allows for a tight estimation of the average ensemble error probability [59], [174]. On other communication channels, various subgraph configurations that produce decoding failures were introduced and analyzed, including *trapping sets* [48], [165], [195], *pseudo-codewords* [84], [123], and *absorbing sets* [16], [28].

D. Early Works on Successive Cancellation Decoding of Reed–Muller codes and on Convolutional LDPC Codes

In the years following the discovery of turbo codes, most of the attention of the coding theory community was

dedicated to the understanding of iterative decoders and in developing the full potential of turbo and LDPC code constructions. Other contributions, whose potential was not immediately recognized, set the ground for important future discoveries. Dorsch's decoding algorithm [65] was revisited in [83], introducing a detailed analysis of its decoding complexity. The algorithm, currently referred to as ordered statistics decoding (OSD), enables near-ML decoding of short binary linear block codes with feasible complexity. In [103], [104], and [233], recursive successive cancellation (SC) and successive cancellation list (SCL) decoding of Reed–Muller (RM) codes [169], [191] were introduced and analyzed, anticipating some of the design elements laying at the foundation of polar codes. A convolutional variant of LDPC block codes was presented in [111]. Together with polar codes, convolutional LDPC codes, and the related (more general) class of spatially coupled LDPC (SC-LDPC) codes, would later be recognized to be simple code constructions capable of achieving the capacity of binary-input memoryless symmetric channels.

E. Engineering Turbo and LDPC Codes for Wireless and Satellite Communication Systems

As the theory that explains the outstanding performance of turbo and LDPC codes was blooming, engineering efforts started to unfold with the objective of delivering the gains observed on paper to real-world communications systems. It was soon understood that while the analysis of turbo codes benefits from assuming *random* interleavers, the implementation complexity of turbo decoders would benefit from *structured* interleavers. Judicious interleaver designs, capable of greatly improving the error floor performance of turbo codes based on random interleavers, were devised and analyzed in terms of implementation complexity [55], [56], [86], [235], [236]. Turbo code decoding is inherent serial and the interleaver plays a crucial role in the parallelization of the decoding process. Therefore, conflict-free interleavers that enable efficient parallelization were a breakthrough for higher throughput. Nonbinary turbo codes that improved on the error floor of the original turbo codes were subsequently introduced [66]. Modifications to the schedule used to exchange messages between component code decoders, with the aim of introducing a level of parallelism, were developed [266], [268]. The capacity-approaching performance of turbo codes, as well as the progress made in their implementation, allowed for early adoption by major wireless communications standards. Focusing on cellular systems, turbo codes were adopted by 3GPP for the 3rd generation of the standard (Universal Mobile Telecommunications Service (UMTS) [4]) as well as by the competing standardization effort carried out by 3rd Generation Partnership Project 2 (3GPP2) with the CDMA2000 standard [1]. The two standards employed two different classes of turbo codes. While the UMTS standard used turbo codes based on two rate-1/2 eight-state recursive convolutional encoders (resulting in a rate 1/3 turbo

code), the CDMA2000 system followed the adoption of two rate-1/3 eight-state recursive convolutional encoders, generating a rate-1/5 turbo code. In both cases, higher code rates can be achieved by puncturing. The UMTS turbo code allows the encoding of information blocks ranging from 40 to 5114 bits, with a fine granularity in the definition of the input block size. On the contrary, the CDMA2000 turbo code allows for a limited set of input block sizes, ranging from 378 to 20 730 bits.

As for turbo codes, also for LDPC codes, it was quickly realized that the implementation of efficient encoders and decoders entailed the use of Tanner graphs with specific structures. On the encoder side, *unstructured* LDPC code constructions exhibit an encoding complexity that grows as $O(n^2)$. By exploiting the sparse parity-check matrix, it was shown in [197] that encoding could be simplified, limiting the number of parity bits whose calculation required quadratic complexity. Bridging the theory of (serial) turbo codes with the theory of LDPC codes, repeat-accumulate (RA) and irregular repeat-accumulate (IRA) codes were introduced in [63] and [112]. Based on the concatenation of an low-density generator matrix (LDGM) code with an inner binary accumulator, RA and IRA codes are classes of LDPC codes whose encoding has linear complexity. Thanks to this property, which is shared by other accumulator-based LDPC constructions [8], [180], IRA codes and their variations became the default choice for the design of LDPC codes with efficient encoders. The Tanner graph of accumulator-based codes is partially structured; the variable and check nodes associated with the accumulator form a chain that can be visualized as a double-diagonal in the code parity-check matrix. Introducing further structure in the code Tanner graph is of paramount importance from the decoder perspective, too. In particular, constructions of parity-check matrices formed by arrays of cyclic submatrices were developed [82], [124], [252], enabling a large reduction in the complexity in the implementation of the message exchange between the variable and check nodes in partially parallel BP decoder architectures [158]. The resulting LDPC codes possess a quasi-cyclic structure, yielding additional algorithmic benefits on the encoder implementation [142]. In particular, if the submatrices are chosen to be cyclic permutation matrices or zero matrices, efficient message schedules can be used to reduce the complexity of BP decoding [101], [218], [268] and trading parallelism, i.e., throughput versus chip area and power. By combining designs based on arrays of cyclic matrices with accumulator-based constructions, LDPC codes amenable to low-complexity encoder and decoder implementations were devised [8], [239], [269]. The introduction of multi-edge type (MET) [198] and protograph-based [245] LDPC code ensembles provided elegant frameworks for analyzing and designing hardware-friendly LDPC codes [8], [11], [61], [148]. Among protograph-based LDPC codes, a family of accumulate-repeat-accumulate (ARA) codes (modified by the introduction of a jagged accumulator)

were standardized by NASA-JPL for spacecraft telemetry [11], [18], [45], jointly with a high-rate quasi-cyclic LDPC code (whose design was based on finite geometries [127]). Protograph-based LDPC codes were part of the 3GPP2 ultra mobile broadband (UMB) standardization [2], paving the way to their adoption in the 5G NR standard [5], [199]. The 5G NR LDPC codes are defined by two protographs (or, equivalently, by their *base matrix* representation). Similar to the space telemetry protograph LDPC codes [45], they employ state variable nodes as well as a considerable fraction of degree-1 variable nodes, following a construction that is reminiscent of one of Raptor codes [47], [225].

F. Achieving Capacity: Spatially Coupled LDPC Codes and Polar Codes

While turbo and LDPC code technologies were reaching maturity, two major breakthroughs were unfolding. In [230], terminated convolutional LDPC codes were analyzed from a density evolution perspective over the BEC. Iterative decoding thresholds extremely close to the capacity of the BEC were shown to be achievable by regular convolutional LDPC, with gaps that diminished for larger variable and check node degrees, in contrast with what happens for regular block LDPC code ensembles. At about the same time, following the introduction of the area theorem [24], a deeper understanding of the performance of LDPC code ensembles over erasure channels was developed, allowing the computation of decoding thresholds under optimum maximum a posteriori (MAP) decoding [160], [161]. The striking numerical coincidence between the MAP decoding threshold of block LDPC code ensembles and the iterative decoding threshold of their convolutional counterparts was recognized a few years later [129], [137]. The phenomenon, referred to as *threshold saturation*, was rigorously analyzed in [129] and [264]. Various classes of LDPC codes (including convolutional LDPC codes) were shown to benefit from threshold saturation [166], [248]. SC-LDPC was the general term that was widely adopted to define this special type of LDPC codes. The threshold saturation phenomenon was shown to take place on general binary-input memoryless symmetric channels [131], [132]. A fundamental consequence of this result is the capability of regular SC-LDPC code ensembles to *universally* achieve the capacity of general binary-input memoryless symmetric channels, as the variable/check node degrees increase, without the need to customize the ensemble construction to a specific channel. Threshold saturation in spatially coupled turbo-like codes was analyzed in [167] and [188]. More details on SC-LDPC codes will be given in Section III-B.

Contemporary with the discovery of the threshold saturation phenomenon, the invention of polar codes [21] showed that channel capacity can be achieved by a *deterministic* code construction, without resorting to ensemble-based arguments. More importantly, polar codes reached

the result with $O(n \log n)$ encoding and decoding complexity. The result was achieved by an information-theoretic code construction, where two copies of a binary-input memoryless symmetric channel are turned into two *synthesized* channels: a downgraded channel (i.e., a channel with capacity smaller than the one of the original channel) and an upgraded channel (a channel with capacity larger than the one of the original channel). The sum of the synthesized channels' capacities preserved the sum of the original channels' capacities. By applying the principle in a recursive manner, it was shown that the synthesized channels polarize; for large blocklengths, all but a vanishing small fraction of synthesized channels have a capacity that is either 1, yielding ideal transmission conditions, or 0, i.e., the synthesized channel is useless. Channel capacity is then achieved by encoding information on the ideal synthesized channels, by *freezing* the input of useless channels to 0 and by employing a simple multistage decoding principle (referred to as SC decoding) at the decoder end. The elegance and simplicity of the approach made polar codes an immediate candidate for future communication systems. However, two drawbacks affected the construction. First, the definition of the synthesized channels hosting the frozen bits depends on the communication channel. Therefore, polar codes do not possess the universal capacity achieving property of SC-LDPC code ensembles. Second, convergence to channel capacity under SC decoding takes place very slowly [23], [168]; polar codes exhibit a competitive performance under SC decoding only for impractically large blocklengths. While the first issue has limited practical impact (the choice of the frozen bits is relatively robust, even when performed over a mismatched channel), the second issue was a main roadblock on the adoption of polar codes in real systems. By engineering the polar code construction, it was shown in [237] that SCL decoding of polar codes, aided by the concatenation of an inner polar code with an outer high-rate block code (often a cyclic redundancy check (CRC)), allowed to tightly approach finite-length performance bounds [183] in the short blocklength regime, overthrowing nonbinary LDPC codes [57], [64], [151], [184] as best low-complexity scheme for short blocks (the results were later matched by concatenated CRC-convolutional codes [262]). A similar observation was made in [173]. In the construction of [173] and [237], the outer high-rate block code is used to filter the list produced by the inner SCL polar decoder, similar to the list decoding schemes for turbo and convolutional codes [171], [172]. By employing complexity- and latency-reduction techniques such as a log-likelihood ratio (LLR)-based formulation [26], fast decoding based on constituent codes [193], [203], [224], adaptive list sizes [140], and combining successive cancellation flip (SCF) decoding [12] with SCL decoding [224], [265], SCL decoding became feasible in practice. Thanks to their excellent performance at short blocks, CRC-concatenated polar codes were included in the 5G NR standard to protect the control channel [5], [34].

From a theoretical viewpoint, the invention of polar codes triggered a renewed interest in recursive code constructions. In [130], a long-lasting conjecture was settled by showing that RM codes achieve the capacity of the BEC under MAP decoding. The result was extended to a wider class of codes, including extended primitive narrow-sense Bose–Chaudhuri–Hocquengham (BCH) codes. Remarkably, a key ingredient of the proof was again the area theorem [24]. The capacity achieving properties of RM codes under bitwise MAP decoding were recently extended to binary-input memoryless symmetric channels [192].

III. NEW DEVELOPMENTS IN CODING THEORY AND COMMUNICATIONS

In this section, we highlight a few recent advancements in the general field of communications and coding theory, which we believe have the potential to impact 6G. First, we highlight novel advances in decoding algorithms for short codes in Section III-A. Short codes will be important as machine-type communication (MTC) will become a major driver of 6G traffic. Afterward, in Section III-B, we highlight spatially coupled codes, which are codes that offer unprecedented performance for streaming applications or very large blocklengths. 6G will not only encompass radio access with high-speed data transfer (with applications such as extended reality, 8K, beyond video streaming, and so on) but also high-throughput applications such as fronthaul and backhaul transmission. Therefore, in contrast to previous mobile communication systems, long codes with good performance and a simple decoding algorithm will also be needed. As 6G will also require low-cost and low-energy transceivers operating at high speeds, we investigate hybrid decoding in Section III-C. Hybrid decoding offers high decoding gains at very low decoder complexity for high-throughput applications. In Section III-D, we show how coding theory can be used to improve massive, uncoordinated multiple access, which can become an important operation mode in 6G with MTC. Finally, Section III-E showcases probabilistic constellation shaping (PCS), which can be potentially used in 6G to further increase the performance toward lower signal-to-noise ratios (SNRs) by providing a so-called *shaping gain* and to add another layer of rate adaptivity.

A. Practical Near-ML Decoding of Short Codes

As the volume and variety of mission-critical MTC grow, ultrareliable low-latency communication (URLLC) modes will become increasingly important [157]. A straightforward way to ensure low latency on the physical layer is to use short packets. This in turn implies the use of error correction coding with short blocklengths, which makes it challenging to achieve good error-correcting performance. On the other hand, the restriction to short blocklengths makes near-maximum-likelihood decoding algorithms more practically feasible even if they have high asymptotic complexity. An overview of the performance of

Table 1 Properties of Decoding Algorithms for Short Error-Correcting Codes

Algorithm	Code Universality	Rate Regime	Parallelizability
OSD	Yes	Low-rate	Low
GRAND	Yes	High-rate	High
LP Decoding	Yes	Any	Medium
PA Decoding	No	Low-rate	High

various error-correcting codes and decoding algorithms in the short blocklength regime was given in [54]. Tail-biting convolutional codes (TBCCs) and polar codes were shown to perform particularly well in the very short blocklength regime. We note that the performance of short polar-like codes can be further enhanced by using more advanced decoding algorithms [181] or polarization-adjusted convolutional codes [22]. In this section, we give a brief overview of some additional and more recent developments in this area, some of which have already been covered in [246]. In particular, we discuss OSD, guessing random additive noise decoding (GRAND), linear programming (LP) decoding, and projection-aggregation (PA) decoding of RM codes. Table 1 gives an overview of various properties of the aforementioned algorithms that are pertinent to low-latency and/or high-reliability MTC.

1) *Ordered Statistics Decoding*: OSD [83] is a soft-decision decoder that can achieve near-ML performance. It is a universal decoding algorithm in the sense that it can decode any linear code and it is often used as a benchmark for other decoding algorithms. The main idea behind OSD is to form the most reliable basis (MRB) of a received noisy codeword using the most reliable bits and to perform bit-flips on the MRB to generate multiple candidate codewords through re-encoding. The most likely candidate is then declared as the decoded codeword. While the candidate generation can be highly parallelized, it has very high computational complexity when near-ML decoding is required because a large number of bit-flips need to be explored. Lower complexity versions of OSD have been proposed, such as box-and-match OSD [250], but these tend to be more serial in nature. Moreover, the MRB calculation requires Gaussian elimination, which has high complexity and can also not be parallelized well.

For these reasons, the study of efficient hardware implementations of OSD, which is a necessary condition for its use in any communications system, has received very little attention to date, with the notable exceptions of the early works of [214] and [215] and the more recent work of [143]. Nevertheless, OSD may still be an interesting decoding algorithm for applications that require very high reliability but have more relaxed latency and bandwidth constraints, such as massive sensing networks.

2) *Guessing Random Additive Noise Decoding*: The main idea behind GRAND is to guess the noise that resulted in the received noisy signal and to undo its effect to recover the transmitted codeword. In its simplest form for a binary code over the binary symmetric channel (BSC), GRAND

generates binary test patterns \mathbf{e} that are added to the received noisy signal \mathbf{y} until $(\mathbf{y} \oplus \mathbf{e})\mathbf{H}^\top = 0$, i.e., until $\mathbf{y} \oplus \mathbf{e}$ is a valid codeword. This approach is similar to Chase decoding, the main difference being that in Chase decoding, and after guessing the noise patterns, the list is hard decision decoded, while in GRAND, we only perform a membership check. GRAND was extended to employ partial soft information in [68] and full soft information in [227]. Similar to OSD, GRAND is a universal decoding algorithm that can decode any linear code, although it is most practical in the high-rate regime and with short codes.

The test pattern generation for soft-information GRAND has high complexity as a new test pattern order needs to be calculated for each received noisy codeword. Ordered reliability bits guessing random additive noise decoding (ORBGRAND) [67] solves this problem by first sorting the received symbols in increasing order of reliability and then generating test patterns based on the logistic weight of each pattern, which can be computed once offline.

The first hardware implementation for GRAND was described in [6] and was shown to achieve a throughput of up to 64 Gb/s for a code with $n = 128$ due to the high parallelizability of GRAND. Due to its lower complexity and deterministic test pattern generation, ORBGRAND has recently received significant attention from a hardware implementation perspective, with some examples being [7], [51], and [261]. Finally, an ORBGRAND ASIC manufactured in 40-nm technology was presented in [194].

3) *Linear Programming Decoding*: Maximum-likelihood decoding of binary linear codes can be formulated as a binary linear program [76]. In particular, given an LLR vector γ and the $(n-k) \times n$ parity-check matrix \mathbf{H} of a code, ML decoding is equivalent to solving

$$\begin{aligned} & \min_{\mathbf{x}} \gamma^\top \mathbf{x} \\ & \text{s.t. } \mathbf{x}\mathbf{H}^\top = \mathbf{0} \\ & \quad \mathbf{x} \in \{0, 1\}^n. \end{aligned} \quad (1)$$

LP decoding has the so-called *ML certificate property*, which means that if the algorithm outputs a codeword, it is guaranteed to be the ML codeword [76].

In practice, solving (1) is too complex and relaxed versions of this problem are solved instead. A particularly interesting relaxation is obtained by using the alternating direction method of multipliers (ADMM) [27], which results in an MP decoding algorithm with manageable complexity and good error-correcting performance, especially in the high SNR regime where it does not suffer from error floors.

ADMM-based decoding has also received some hardware implementation attention. Initial implementation attempts were presented in [58] and [256]. More recent works have focused on the Euclidean projection onto

the parity polytope step of ADMM decoding, which is particularly challenging from a hardware implementation perspective [89], [244], [257]. The remaining challenges for the implementation of ADMM-based decoders are outlined in [126] and [257, Sec. VI].

4) Projection-Aggregation Decoding of RM Codes: The aforementioned decoding algorithms are universal in the sense that they can decode any linear code. However, when restricting ourselves to codes with a specific structure, it is usually possible to derive more efficient decoding algorithms. One example is the recently introduced class of PA-based decoding algorithms for RM codes that were first proposed in [263] in the form of the recursive PA (RPA) decoder that achieves near-ML decoding performance. These decoding algorithms exploit the recursive structure of RM codes by decoding several RM subcodes that are obtained through subspace projections. The results of these partial decodings are then aggregated into a final decoding result. Unfortunately, the computational complexity of the RPA decoder scales exponentially with the RM code order r , meaning that it is mostly interesting for low-order (i.e., low-rate) RM codes.

Several works have attempted to reduce the complexity of RPA decoding. For example, Hashemipour-Nazari et al. [98] showed that decoding iterations in recursive levels are not very beneficial and can be removed. In a similar vein, Lian et al. [144] proposed to remove recursive levels completely by performing projections on higher order subspaces. In a different direction, Fathollahi et al. [75] proposed to use multiple parallel RPA decoders, each with a significantly reduced number of projections, thus lowering the overall complexity while largely maintaining the error-correcting performance. The work of [141] has lowered the average computational complexity of RPA by taking advantage of syndrome-based early stopping techniques along with a scheduling scheme, while Hashemipour-Nazari et al. [99] used aggressive projection pruning.

Recent hardware implementation results for the simplified algorithm proposed in [98] show an area efficiency improvement of almost an order of magnitude for a similar decoding latency compared to a state-of-the-art SCL decoder for polar codes [100] in the short blocklength and low-rate regime. Unfortunately, the energy efficiency is still worse than that of the SCL decoder, but including some of the other aforementioned simplifications in the hardware implementation is expected to further improve the results.

B. Spatially Coupled Codes

6G will offer new functionalities expanding beyond the capabilities of the current mobile communication system [110]. These include, but are not limited to, high-speed point-to-point communications for fast data transfer, point-to-point communications in the terahertz, and visible light regime to provide high-speed access to users or to interconnect base stations [96, Fig. 1]. These

applications will likely operate at data rates that are multiple orders above the typical data rates in conventional mobile communication systems. Hence, we need codes that are tailored to high-throughput applications. Coding for high-throughput applications has become extremely important in the past few years in optical communications [92], [139]. The drive for longer transmission distances and throughputs in the range of 100 Gbit/s–1.2 Tbit/s have led to coding schemes that offer high net coding gains (NCGs) and high throughput with low-complexity, ultra-parallelized decoders. These can form a blueprint for novel high-data-rate applications in 6G.

In contrast to typical mobile networks, optical communication networks are typically circuit switched and the data stream is not packetized but rather continuous. Also, when operating at high data rates, latency due to packetization and processing is typically not significant and we can use codes that have either a large blocklength or even operate in a continuous mode, similar to traditional nonterminated convolutional codes.

In this section, we introduce a class of codes that enable both operating modes (large blocklengths or continuous) and offer high NCGs with moderate decoding complexity. This class of codes is the class of *spatially coupled* codes. Besides polar codes, the concept of *spatial coupling* was the second coding breakthrough in the past decade. Spatially coupled codes originate from LDPC convolutional codes [111], a generalization of LDPC codes with a superimposed convolutional structure. A similar idea was already pursued in the late 1950s and early 1960s with the concept of recurrent codes (see [260] and references therein). In fact, it was only realized in the late 2000s that the asymptotic decoding performance of a certain class of SC-LDPC codes approaches the performance of optimum decoding [135], [136], [137] with simple suboptimum BP decoding. By spatial coupling, we can construct very powerful codes that can be decoded with a simple windowed decoder [109].

Spatially coupled codes have in various forms tailored to high-throughput applications. One example is the staircase code of rate 239/255 presented in [226] targeted for low-complexity optical communications with hard decision decoding (HDD). The staircase code can be classified as spatially coupled BCH product code (PC) with a natural windowed decoder implementation. Spatially coupled LDPC codes are a natural extension of the ubiquitous LDPC codes and have been successfully employed in long-haul optical communications, where they are able to achieve record NCGs of about 12 dB at residual BERs of 10^{-15} and a rate of 4/5, proven by an field-programmable gate array (FPGA)-based emulation [206], [211].

In order to describe SC-LDPC codes, we first take on an encoder perspective. An LDPC encoder is a block encoder, i.e., it encodes a block of information bits u_t into a codeword x_t . This is shown in Fig. 2(a). The encoder of an SC-LDPC code can be interpreted as an encoder of

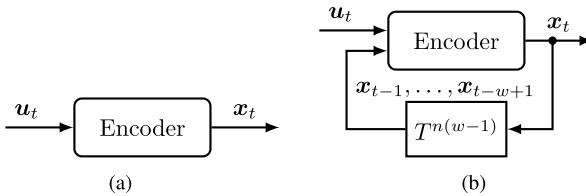


Fig. 2. Comparison of (a) conventional block LDPC encoder and (b) spatially coupled LDPC encoder.

a recurrent code [260] and is shown in Fig. 2(b). The recurrent encoder generates batches x_t of a codeword. The batch x_t at time t does not only depend on u_t but also on $w - 1$ previous batches $x_{t-1}, \dots, x_{t-w+1}$. The parameter w is called the *coupling width* of the code. The term *spatial coupling* highlights the dependency of x_t on previous batches along the dimension t . Each batch is hence also called a *spatial position*. The case with $w = 2$ is particularly interesting as it can lead to very simple and effective decoders and is called *unit-memory coupling* [210].

At the beginning of the transmission (for $t = 1$), we assume that the contents of the feedback memory are zero, i.e., the previous batches at undefined times $t < 1$ are $x_0 = x_{-1} = \dots = 0$. After starting the transmission, we could in principle run the encoding procedure without interrupting it. This is sometimes called *left-termination* or *unterminated* operation and has shown to be an effective coding method in high-throughput applications [212]. Frequently however, the encoding is terminated after the transmission of L batches, where L denotes the *replication factor* or *spatial length* of the code. Termination means that $x_t = 0$ for $t > L$, provided that the input $u_t = 0$ for $t > L$. Termination yields a block code again, where the length n of the block code amounts to $L\tilde{n}$ bits, with \tilde{n} being the size of one batch x_t . Termination is realized by slightly reducing the number of information bits to $\tilde{k}' < \tilde{k}$ in the final $w - 1$ batches, i.e., u_{L-w+1}, \dots, u_L . The encoder then generates $\tilde{k} - \tilde{k}'$ extra parity bits that are required to enforce the termination condition. These extra parity bits lead to a *rate loss* (i.e., reduce the effective number of information bits) and possible difficulties when implementing a framing structure. However, note that the effect of the rate loss becomes negligible if L is increased. Termination is often used to fit multiple batches into a larger superframe, a networking container, or to prevent error propagation.

SC-LDPC codes can also be described by a parity-check matrix with a banded structure [92]. For example, we provide a parity-check matrix for a time-invariant, terminated SC-LDPC codes with $L = 5$ and $w = 2$

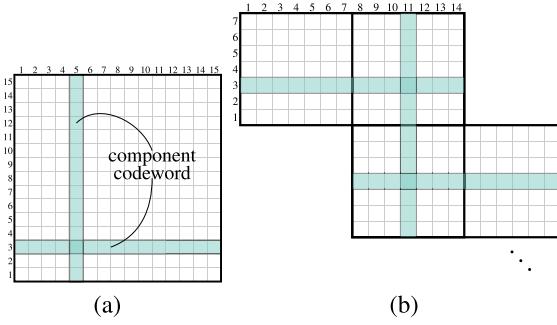
$$\mathbf{H} = \begin{pmatrix} \mathbf{H}_0 & & & & \\ \mathbf{H}_1 & \mathbf{H}_0 & & & \\ & \mathbf{H}_1 & \mathbf{H}_0 & & \\ & & \mathbf{H}_1 & \mathbf{H}_0 & \\ & & & \mathbf{H}_1 & \mathbf{H}_0 \\ & & & & \mathbf{H}_1 \end{pmatrix}$$

with $\dim \mathbf{H}_i = \tilde{m} \times \tilde{n}$ and $\dim \mathbf{H} = (L + w - 1)\tilde{m} \times L\tilde{n}$. Each block column of \mathbf{H} corresponds to a spatial position and hence one batch of size \tilde{n} . The encoding procedure can also be described using \mathbf{H} . If we decompose $\mathbf{H}_0 = (\mathbf{H}_{0,\text{inf.}} \ \mathbf{H}_{0,p})$, where $\mathbf{H}_{0,p}$ is nonsingular, we have the condition $\mathbf{H}_0 \mathbf{x}_1^T = 0$. If we assume *systematic encoding* with $\mathbf{x}_t = (u_t \ p_t)$, where p_t are the parity bits of batch t , the condition for the first batch becomes $\mathbf{H}_{0,\text{inf.}} \mathbf{u}_1^T + \mathbf{H}_{0,p} \mathbf{p}_1^T = 0$ and we can express $\mathbf{p}_1^T = \mathbf{H}_{0,p}^{-1} \mathbf{H}_{0,\text{inf.}} \mathbf{u}_1^T$, leading to encoding of the first batch. Similarly, the parity bits of the second (and subsequent) batch can be encoded as $\mathbf{p}_t = \mathbf{H}_{0,p}^{-1} (\mathbf{H}_1 \mathbf{x}_{t-1}^T + \mathbf{H}_{0,\text{inf.}} \mathbf{u}_t^T)$, akin to the illustration in Fig. 2. Special care needs to be taken for the final termination [111], [240], i.e., for computing $\mathbf{x}_{L-w+2}, \dots, \mathbf{x}_L$.

Spatially coupled LDPC codes require novel approaches for code optimization and code design. An initial design can be obtained from a conventional LDPC code, and however, we need to pay attention to multiple details. The coupling adds another dimension that we can exploit to design codes with fast convergence, requiring only a limited number of iterations for convergence [19]. Optimizing the distribution of the coupling can lead to codes that have optimized convergence speed [208], [209]. Furthermore, we can make use of the modulation format or other properties of the transmission system to trigger the decoding process more effectively [42], [43], [94]. While spatially coupled LDPC codes have a very efficient windowed decoder, the decoder can intrinsically lead to a phenomenon called *decoder stalling*, which leads to long bursts of errors. These bursts can be mitigated using an optimized decoder scheduling [122]. On the other hand, spatially coupled LDPC codes have built-in burst correction capabilities, which can be exploited during code design [20].

C. Recent Advances in Hybrid Decoding of Product-Like Codes

With increasing data rates, 6G will likely offer modes for high-throughput communications, e.g., in fronthaul, backhaul, for 8K video streaming, extended reality, and more. It is important to reduce the complexity of the receiver integrated circuit [46] in order to increase both the energy efficiency of the base station (and hence the operating expenses for network providers) and of the user equipment, which is typically constrained by the battery capacity. Channel decoding, and in particular soft decision decoding (SDD), is typically responsible for a large part of the receiver's power consumption [133]. HDD usually leads to extremely power-efficient receivers at the expense of a significantly reduced coding gain. In recent years, hybrid decoding emerged as a means to extend HDD by soft information in order to improve the coding gains without significantly decreasing the power efficiency of the receiver implementation. As of today, hybrid decoding only gives satisfactory results for a specific class of codes that are used in high-throughput applications. In this section,

**Fig. 3.** Diagrams of (a) product codes and (b) staircase codes.

we introduce these product codes (and derived schemes) and describe hybrid decoding for this class of codes. The development of low-complexity hybrid decoding algorithms for general coding schemes is an open research problem.

PCs [70] and their generalized versions, referred to as generalized product codes (GPCs), are competitive forward error correction (FEC) schemes for high-throughput communication systems. They are often used in such systems, as extremely efficient, highly parallelizable decoders for HDD exist. Fig. 3(a) illustrates a PC codeword as a 2-D array, where each row and column is protected by an (n, k, t) component code \mathcal{C} . Typically, this component code is a Reed–Solomon (RS) code or a BCH code, resulting in a PC rate of $((k/n))^2$. Employing a spatially coupled structure, PCs can be further generalized to, for example, staircase codes [226], zipper codes [234], and braided codes [77]. Fig. 3(b) depicts the structure of a staircase code, which comprises an infinite sequence of blocks of size $(n/2) \times (n/2)$, achieving a rate of $2(k/n) - 1$. Each row and column is a codeword of a component code. Similar to PCs, every two neighboring blocks are protected by a component code \mathcal{C} . As staircase codes are not terminated, a sliding window decoder is employed for decoding.

Initially, GPCs were employed for HDD assuming the BSC as channel model. Known as iterative bounded-distance decoding (iBDD), the decoding process involves alternatively decoding the rows and columns of GPCs using a component decoder D_C for BCH or RS codes. The decoder D_C is an algebraic bounded-distance decoder able to decode up to t errors. iBDD is a low-complexity decoder, requiring only hard MP. In an iterative MP decoder operating at high throughputs (such as LDPC decoders or iBDD), the internal decoder data flow is a good indicator of the decoder power consumption. In an MP decoder, messages are passed between the computing nodes of a decoder (typically one node per bit), which requires space and power on the decoder. Soft messages are typically quantized and require q bit/message, whereas hard messages only require a single bit/message. iBDD can further reduce the number of binary messages by only exchanging messages for bits that undergo changes, which reduces the internal decoder data flow by two orders of magnitude compared to BP decoders for LDPC codes. We refer the interested reader to [226] for details. Therefore,

energy-efficient decoding of codes with large blocklengths can be implemented.

Later on, to approach the capacity of channels such as the binary-input AWGN (BI-AWGN) channel, SDD was introduced. The Chase–Phydiah turbo product decoding (TPD) [187] algorithm, a famous example of SDD, yields an additional decoding gain of 1–2 dB compared to iBDD. However, apart from the high computational complexity, TPD also requires soft MP during iterative decoding, making adapting TPD for high-speed communication systems challenging.

In this section, we review recent results of hybrid decoding for GPCs. Hybrid decoding aims at exploiting soft information to aid hard decision decoders, achieving performance similar to SDD using only binary or ternary low-complexity MP. One promising approach to enhance HDD is employing an enhanced component code decoder. For example, error-and-erasure decoders (EaEDs) are explored in [190], [228], and [270], generalized minimal distance decoder (GMDD) [81] is studied in [220] and [223], and a list-based component code decoder was used in [162].

A major issue with this class of codes are *mis corrections*, where the component code decoder yields a valid yet incorrect codeword, which is then undetectable by the algebraic decoder, hindered the performance improvement achieved by enhanced component decoders, as observed in [190]. Mis corrections occur frequently as the component codes are often very simple high-rate algebraic codes with only limited error-correcting capabilities (e.g., two or three-error-correcting BCH codes). To address this issue, several heuristic strategies were proposed.

For example, iBDD with scaled reliability (iBDD-SR) introduced in [221] combines the scaled soft-channel output LLR with the component decoder decision to increase the reliability of the component decoder. Later works extended this principle [145], [219], [222], [223], with iBDD with combined reliability (iBDD-CR) using an optimized combining rule derived via density evolution [222]. In particular, the BEE-PC and BEE-SCC decoders proposed in [219] combine EaED and iBDD-CR, yielding good performance with relatively high decoding complexity.

In [134], soft-aided bit marking (SABM) was proposed, which marks a set of highly reliable bits based on the channel LLR. During component code decoding, the decision of the component code decoder D_C is only accepted if it does not flip a highly reliable bit. Compared to iBDD-SR, SABM requires less storage overhead. The shared drawback of iBDD-SR and SABM is that the channel LLRs are never updated during the decoding and become stale in the later decoding iterations.

In [95], anchor decoding (AD) was introduced, which does not use any soft-channel information but utilizes the special property of the iBDD decoding of GPCs. A set of anchor bits are identified as the bits that do not change during iterative decoding. AD is based on the heuristic that a bit, on which both row and column decoder agree,

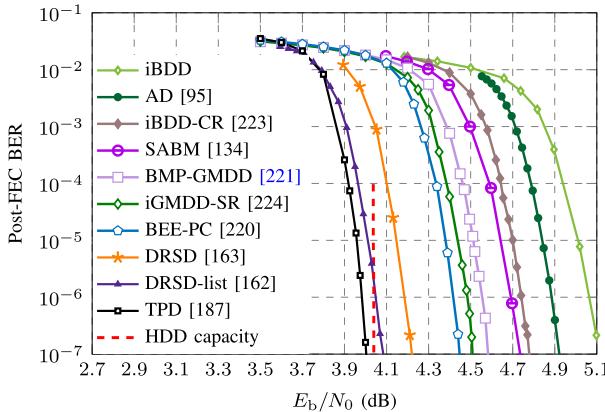


Fig. 4. Performance comparison of different decoders for PC with a rate of 0.87. The component code used is based on [255, 239, $t = 2$] BCH code with either 1-bit extension or its even-weight subcode.

is likely to be error-free. Then, the anchor bits are used to prevent miscorrections with a similar principle as in SABM, i.e., the output of the component code decoder D_C should not be accepted if it conflicts with any of the anchor bits.

The dynamic reliability score decoder (DRSD) [163], [164] combines the ideas of SABM and AD where each bit is assigned a reliability score indicating its reliability. The scores are initialized using channel LLRs and increased if a bit stays constant during decoding and decreased if a bit is flipped by the row and column decoders. Compared to SABM and AD, DRSD requires more storage space but delivers a significantly larger decoding gain compared with the former two.

Figs. 4 and 5 demonstrate the decoding performance of various algorithms for a PC and a staircase code, respectively. The component code of the PC is constructed with a [255, 239, $t = 2$] BCH code with 1-bit extension or its even-weight subcode. The reason for extending or taking the subcode is to increase the minimal distance by 1 with negligible rate loss, leading to a reduced miscorrection rate and improved iterative decoding performance. In addition, for EaED, this is beneficial as it allows the correction of one additional erasure. The staircase code is constructed with [255, 231, $t = 3$] BCH code with 1-bit shortening as staircase codes require component codes with even blocklength. For both codes, compared to iBDD, soft-aided decoding algorithms provide a significant decoding gain with similarly low complexity.

D. Coding for Massive Uncoordinated Multiple Access

During the past few years, coding theory has emerged as a fundamental tool in the domain of massive *uncoordinated* multiple access communications. As we shall see next, channel codes play a double role in this context. First, they provide an essential mechanism to protect data packets from noise and multiuser interference. Second, by casting the uncoordinated multiple access problem from a code-theoretic point of view, new classes of energy-efficient uncoordinated multiple access protocols have been

introduced. The interest in uncoordinated (i.e., *grant-free*) access protocols is widely acknowledged [35], and it is mostly motivated by the rapid growth of the Internet of Things (IoT) and, more generally, of large-scale wireless sensor networks—which is giving rise to new challenges for the medium access control layer of future cellular systems.

Uncoordinated access protocols allow to remove the overhead entailed by the typical handshake used by terminals to obtain an allocation of radio resources. This is especially important in IoT systems, where the activity of the nodes is typically sporadic and unpredictable. Owing to the traffic characteristics and the large population of terminals, resorting to a grant-free approach seems to be the only viable solution. In uncoordinated access protocols, user terminals transmit their data without negotiating with the base station an allocation of resources. This results in transmissions that may randomly collide over the selected radio resources. The amount of mutual interference generated by such transmissions may prevent the successful reception of the data. However, if the traffic load is low enough, the probability of success can still meet the (typically mild) requirements of the IoT network. While the 5G NR standard includes a first example of a grant-free access protocol, its design does not meet the efficiency required by a large-scale IoT network [121]. Historically considered in the domain of medium access control protocols [9], [10], the problem of reliable communication with uncoordinated terminals through a shared medium can largely benefit from a coding-theoretic approach. In [147], a first class of schemes built on the theory of LDPC codes was introduced. The multiple access channel (MAC) technique of [147], devised as a generalization of an earlier contention resolution diversity slotted Aloha (CRDSA) protocol [44], assumes the transmissions to be aligned to a slot structure. In particular, the slots are organized in MAC frames, composed of M slots each. When a user attempts transmission over a MAC frame, it selects a set of slots over which copies of the data packet are sent.

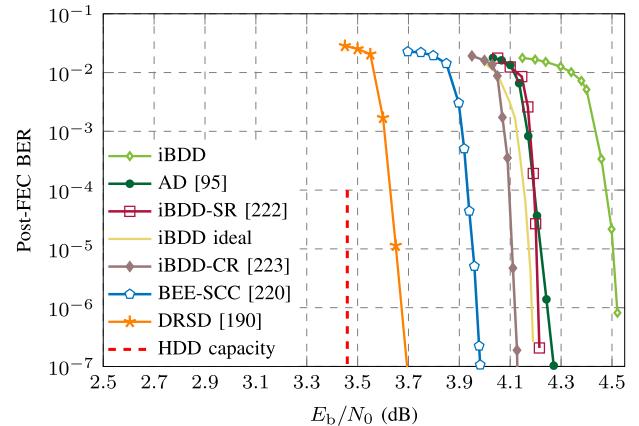
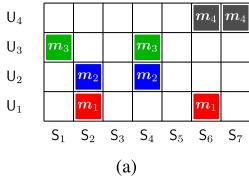
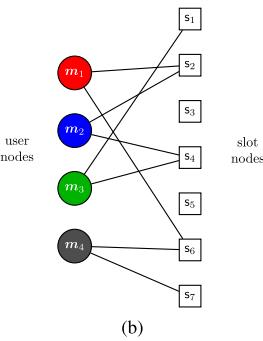


Fig. 5. Performance comparison of different decoders for a staircase code with a rate of 0.811. The component code used is the [255, 231, $t = 3$] BCH code with 1-bit shortening.

User 1: $x_1 \rightarrow \text{hash}(m_1) = \{2, 6\}$
User 2: $x_2 \rightarrow \text{hash}(m_2) = \{2, 4\}$
User 3: $x_3 \rightarrow \text{hash}(m_3) = \{1, 4\}$
User 4: $x_4 \rightarrow \text{hash}(m_4) = \{6, 7\}$



(a)



(b)

Fig. 6. Transmission according to CRDSA/IRSA. (a) Example of transmission pattern over an MAC frame composed of seven slots. (b) Corresponding bipartite graph representation.

The choice of slots is made by hashing the packet to be transmitted. An example of transmission is depicted in Fig. 6(a). Here, four users access a frame composed of seven slots, and each user transmits two copies of its packet. On the decoder side, packets received in *singleton* slots (i.e., slots where only one transmission is present) are decoded. By hashing the decode packets, the slot indexes where their copies have been sent are retrieved. Interference cancellation is then performed to remove the interference generated by the copies, eventually revealing new singleton slots. The process is iterated until no singleton slot (hence, no decodable packet) can be found. The described successive interference cancellation (SIC) algorithm admits a simple representation on a bipartite graph, where a user node is associated with every active user, a slot node with every slot in the MAC frame, and an edge connects a user node with a slot node if the corresponding user transmits a packet copy in the associated slot. Continuing with the example, the bipartite graph for the MAC frame of Fig. 6(a) is shown in Fig. 6(b). It is easy to recognize that the iterative SIC algorithm corresponds to an edge peeling process over the bipartite graph. This observation allowed to bridge the analysis of the SIC algorithm with the analysis of peeling decoders of LDPC codes [147]. Density evolution techniques were used to improve the performance of the protocol; by allowing the number of copies generated by users to follow a carefully designed degree distribution, an efficiency approaching 1 packet per slot was achieved [147], [170], while the classic Aloha slot protocol saturates at $1/e$ packets per slot [10]. In contrast to CRDSA, where the number of packet copies is fixed, irregular repetition rates are used by the construction of [147]. For this reason, the protocol of [147] was named irregular repetition slotted Aloha (IRSA). Further generalizations of the schemes of CRDSA and IRSA have been introduced and analyzed, including the following:

- frameless Aloha [231], [232], which exploits a feedback link to modify the distribution of the repetitions, as the SIC proceeds, to improve the success probability;

- Coded slotted Aloha (CSA) [176], where the packet repetition entailed by CRDSA/IRSA is replaced by the use of short erasure codes, allowing for a fine control of the tradeoff between energy consumption and throughput;
- a spatially coupled version of CSA [202], with remarkable gains in throughput;
- slot-asynchronous variants of CRDSA [50].

The physical layer of CRDSA/IRSA was optimized, in combination with the use of multiple antennas at the base station, in [251]. Thanks to tangible gains over classical random access protocols (such as Aloha and slotted Aloha), CRDSA/IRSA have been included in the DVB-RCS2 standard for interactive satellite services [72].

The remarkable performance of the class of protocols stemming from CRDSA comes with two important open points. First, the use of packet repetition, where the energy used to transmit the different packet copies is not used in the decoding process, implies an energy cost that needs to be carefully analyzed. Second, this class of protocols builds upon the Aloha protocol; channel coding is applied as an additional layer that jointly encodes data units, which are subsequently transmitted by means of Aloha. Therefore, there is no guarantee that the achieved performance is close to the best that can be achieved with the given energy/bandwidth constraints. This second point was exacerbated by the fact that, not until recently [182], an information-theoretic framework to provide limits on the performance achievable by grant-free access schemes was missing. The foundations for establishing theoretical limits on (massive) uncoordinated MAC protocols were laid in [182]. The key observation was that, when a large number of users access the communication medium without coordination, assigning user-specific resources to enable user separation is highly inefficient. Here, “resources” can be interpreted in very general terms as different codebooks assigned to different users. Note in fact that, from an information-theoretic perspective, time/frequency division multiple access can be considered as a (highly structured) code. Hence, in contrast to coordinated MAC schemes, in uncoordinated protocols, all users employ the same codebook. To grasp the generality of this observation, let us consider the simple example of a slotted Aloha protocol, operated over an MAC frame of M slots. Here, the codebook is structured in a way that every codeword is composed of $M - 1$ null slots and a single slot that contains the transmitted packet; any active user in the frame can originate any such codeword. In the simple frame-and-symbol synchronous setting where the channel introduces AWGN, the model can be cast as

$$\mathbf{y} = \sum_{w=1}^K \mathbf{x}^{(i_w)} + \mathbf{z} \quad (2)$$

where $\mathbf{y} = (y_1, y_2, \dots, y_n)$ is the sequence of n real channel observations, $\mathbf{x}^{(i_w)}$ is the codeword transmitted by the w th

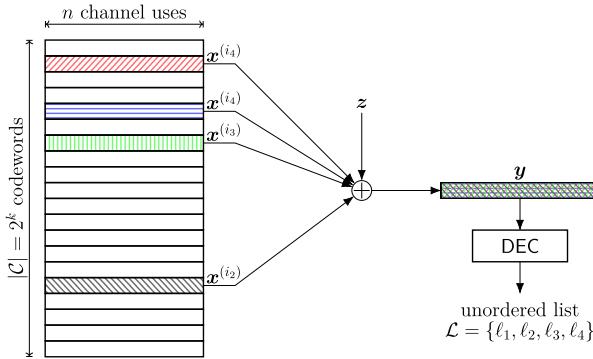


Fig. 7. Uncoordinated multiple access model from [182].

active user, K is (random) number of active users, and $z = (z_1, z_2, \dots, z_n)$ is the sequence of additive Gaussian noise terms. The setting is depicted in Fig. 7. By enforcing $E[|X|^2] = 1$ and assuming $Z \sim \mathcal{N}(0, \sigma^2)$, the SNR is defined as $\text{SNR} = 1/2\sigma^2$. The codewords belong to a code \mathcal{C} with blocklength n , where the number of codewords is $|\mathcal{C}| = 2^k$, i.e., each user can transmit k information bits. In the idealized case where the receiver is able to obtain an accurate estimate of the number of active users, the decoder outputs an unordered list \mathcal{L} of indexes pointing to K codewords. The per-user probability of error is then given by

$$P_e = \frac{1}{K} \sum_{w=1}^K P(i_w \notin \mathcal{L}). \quad (3)$$

In [182], an achievable region (P_e, K, SNR) for given code parameters (n, k) was established, providing a benchmark that can be used to evaluate the performance of grant-free access schemes. From a coding theory viewpoint, a second major contribution of [182] was to cast the decoding problem in a compressive sensing (CS) framework, providing a direction to design coding and decoding schemes capable of approaching the achievability bounds. The sensing matrix here is obtained by stacking all codewords in an array of size $n \times |\mathcal{C}|$. It is easy to observe that even when users transmit moderate amounts of information (e.g., k is in the order of a few tens of bits), the size of the sensing matrix is tremendously large to enable the application of any CS algorithm. Inspired by these insights, several schemes relying on CS sensing techniques were proposed, applying a divide-and-conquer approach to break the curse of dimensionality. In [14], a coded CS scheme was proposed, which relies on splitting a message to be transmitted in subblocks, where each subblock is encoded separately. By doing so, the decoding of the superposition of the messages is performed on a subblock basis, reducing the size of the sensing matrix employed by the CS decoder. The subblocks associated with each user message are then stitched together by means of an outer tree code. Variations on this theme were proposed in [15] and [78]. In particular, Fengler et al. [78] replace the inner code (sensing matrix) of [14] with a sparse regression

code [254]. An alternative line of work [186], [247], [253], which is closely related to the existing two-step random access procedure of the 5G NR standard, employs a CS phase to detect the preambles selected by users from a common dictionary to signal their activity. Each preamble identifies the resource that each user adopts in the subsequent transmission. The approach realizes a two-phase mechanism, where the second phase can be seen as a special instance of a nonorthogonal multiple access protocol. For example, in [186], the preambles define a sparse interleaving pattern that is used in subsequent transmissions, which are encoded with an LDPC code. At the decoder end, after identifying the transmitted preambles, a joint Tanner graph describing the superposition of the permuted LDPC code words is constructed and used to decode the overlapping transmissions. In [17] and [125], it was shown that a slotted Aloha scheme supported by a strong error correction code (LDPC or polar) can perform close to theoretical limits over block fading channels. The suitability of short polar codes to support transmission in uncoordinated MACs is confirmed by the more elaborate construction of [185], where a short polar code is used in a spread-spectrum scheme.

The brief summary provided above on the remarkable developments in the design of advanced grant-free MAC protocols, witnessed during the past few years, shows that a wide variety of coding solutions are available to attack the problem. The level of maturity of the techniques is still relatively low, as most of the studies have focused on rather idealized settings, with limited insights on the robustness of the schemes to the harsh propagation conditions that may be encountered in cellular wireless systems. Hence, there is still considerable space to improve on some of the already proposed protocols and to engineer new ones, taking into account realistic system constraints. From a coding theory point of view, research challenges related to the complexity versus network scaling tradeoff have been identified in [149]. For a more detailed account of the problem, we refer the interested reader to [150].

E. Constellation Shaping

We have seen that powerful code constructions exist and that we can approach the Shannon limit with modern coding schemes relatively easily. Modern coding hence enables us to increase the power efficiency of communication systems at the expense of redundancy (parity bits). To transmit the parity bits, we either need to operate at a higher rate and hence larger bandwidth or to use higher order modulation formats. Higher order modulation formats, such as quadrature phase shift keying (QPSK), 16-quadrature amplitude modulation (QAM), and 64-QAM, are used in today's mobile networks. The combination of coding and modulation is called coded modulation (CM).

CM itself has a long history with trellis-coded modulation (TCM) and multilevel coding (MLC) being first

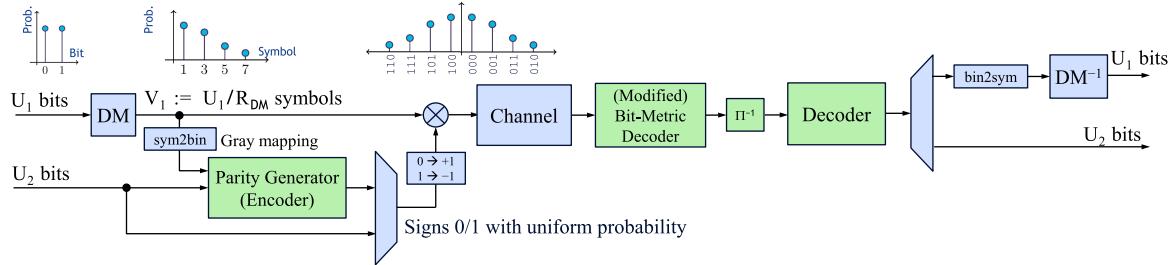


Fig. 8. Block diagram of probabilistic amplitude shaping (PAS) transmitter and receiver of a single quadrature of the PAS system. Channel coding-related blocks are highlighted in green. Image adapted from [207].

popular schemes [107], [249], [255]. Both schemes suffer from the drawback that flexibility in terms of adaptive coding rates and adaptive modulation formats is difficult to achieve. Hence, these schemes are often employed in point-to-point applications with a fixed code and modulation format, e.g., in the V.32 modem standard for communication over analog phone channels [108].

Mobile networks need to operate over a wide variety of rapidly changing channel conditions, covering a wide range of SNRs, modulation formats ranging from binary phase shift keying (BPSK) to 256-QAM with code rates ranging from $R = 1/5$ to $R = 8/9$ in 5G-NR (to meet peak decoder throughputs of 5 Gbit/s) and operating with information blocklengths ranging from $k = 100$ to $k = 8000$ [199]. Combining such a hyperflexible coding scheme with traditional CM schemes, such as TCM and MLC, is difficult, which is why often a pragmatic approach to CM is chosen, where the encoder output is interleaved and fed to the modulator. Such an approach is also called bit-interleaved coded modulation (BICM) [41]. Interestingly, such a pragmatic approach yields nearly optimal performance with square QAM constellations and Gray bit labeling, provided that the SNR (and hence rate R) is sufficiently high.

In high-throughput optical communications, rate adaptivity only became necessary in the 2010s, when systems needed to incorporate multiple modulation formats and modes to cover various applications, ranging from data-center interconnects to submarine optical cables. Due to stringent requirements on residual (post-decoding) error rate, decoder implementation, and parallelization, hyperflexible codes, such as those used in 5G-NR [199], could not be used. Interestingly, in optical communications, the rate adaptivity problem was solved by a CM technique called probabilistic amplitude shaping (PAS) [37].

PAS solves the rate adaptivity problem by introducing a new device, the distribution matcher (DM) [216]. The DM generates a sequence of nonuniformly distributed amplitudes (belonging to a set $\{1, 3, 5, 7, \dots\}$) from a sequence of uniformly distributed bits, where the distribution of the amplitudes can be configured. PAS then uses a fixed code of fixed rate (for which a highly optimized decoder architecture exists) to generate sign values (± 1) with which the amplitudes are multiplied. Two such values

are then used to generate complex modulation symbols. A block diagram of PAS is shown in Fig. 8. PAS has shown to enable both high-rate adaptivity and transmission reach gains in optical communications [39], [207] (see also [36] for a tutorial-like introduction in the context of optical communications). The receiver of PAS is simple, with a soft demodulator (also called *bit-metric decoder*) generating the channel decoder input followed by a decoding step and inverse DM. The only variable part of PAS is the DM.

As a side effect, the use of PAS entails a *shaping gain* if an optimized modulation constellation, resembling a Gaussian channel input, is used [79]. PAS is an implementation of *probabilistic constellation shaping*, which mimics a Gaussian channel input by adapting the probability of each modulation symbol [128]. The benefits of PAS compared to conventional modulation formats are shown in Fig. 9, where we use a rate $R = 4/5$ SC-LDPC code together with either PAS or conventional QAM modulation [207] leading to different effective number of bits per QAM symbol (bpQs). The code has coupling width $w = 2$ and $n = 24\,000$ bit/batch and is described in detail in [90] and [91]. For decoding, we use a windowed decoder [109] with a window of $W_D = 15$ batches and carry out ten iterations per window step. In PAS, the effective number of bpQs is configured via the rate R_{RM} of the DM. We can see an extra coding gain of ≈ 1 dB. Constellation shaping promises extra gains of up to 1.53 dB [79]. The success of PAS in the optical communications community led to a renewed interest in constellation shaping and also *geometric* constellation shaping (see [102]) has been considered, where the modulation alphabet is optimized.

Although PAS and constellation shaping in general offer immediate coding gains, it has not yet been adopted in mobile communication networks. So far, the extra coding gains of approximately 1 dB and the extra benefit in terms of rate adaptivity have not been sufficient to warrant a major overhaul of the physical layer architecture. Also, in contrast to optical communications, which mostly operate as single-carrier or few-carrier systems, mobile networks ubiquitously employ multicarrier modulation formats such as orthogonal frequency-division multiplexing (OFDM). In OFDM, flexibility to varying channel conditions can be achieved by changing the rate and modulation format individually per subcarrier. In the future, and in particular in 6G applications where we need to transmit

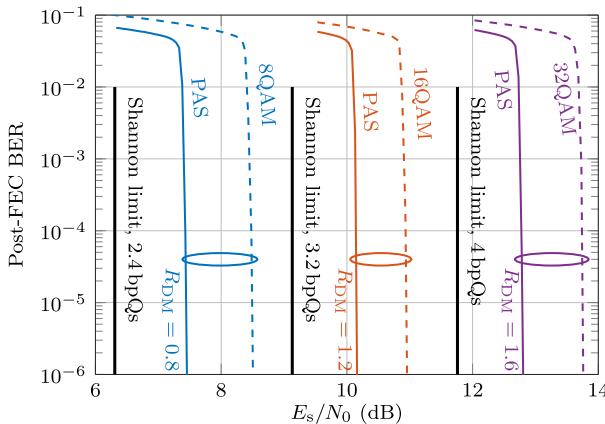


Fig. 9. Simulation results of a spatially coupled LDPC code with PAS and comparison with legacy, unshaped QAM modulation formats. We also show the respective Shannon limits for the fixed number of bpQs.

large amounts of data with high data rates, PAS may become a viable option to ensure flexibility and rate adaptivity, if the latter cannot be guaranteed by the coding layer itself.

IV. IMPLEMENTATION: A TECHNOLOGY PERSPECTIVE

A. Trends in Very Large-Scale Integration Technology

Already in 1968, Berlekamp [31] stated: “From a practical standpoint, the essential limitation of all coding and decoding schemes proposed to date has not been Shannon’s capacity but the complexity of the decoder.” This statement is more valid than ever. For several decades, progress in microelectronics driven by Moore’s law was an enabler for the implementation of complex decoding algorithms, large leaps in throughput, lower latency, lower power, and low area. Microelectronics is still following Moore’s law, but slowing down for the following reasons, to name but a few of them.

- 1) Design costs in advanced technology nodes are skyrocketing, i.e., a state-of-the-art system-on-chip design in 5-nm technology costs about U.S. \$500 million.
- 2) A minimum cost per usable gate was reached at the 28-nm technology node, after which the cost increased with every new technology generation.
- 3) Analog circuits and SRAMs can no more profit from further scaling.
- 4) With further scaling, the transistor density is increasing faster than the power reduction of a single transistor. Thus, the power density is continuously increasing.
- 5) The delay and power consumption of interconnect is rising due to the increasing resistance.

Thus, forward error correction is no longer just a matter of spectral efficiency or bit/frame error rate. When it comes to implementation, channel coding requires a cross-layer approach covering information theory, algorithms, efficient

architectures, and implementation in semiconductor technology. Previously mentioned examples of the interrelation of code design and efficient implementation in the past are conflict-free interleaver for long-term evolution (LTE) turbo code decoding that enable highly parallel architectures or quasi-cyclic structures of LDPC codes that enable low complexity encoding and efficient partially parallel decoder architectures.

B. Implementation Challenges

From an implementation perspective, there is a fundamental tradeoff between information theory and efficient decoder architectures [118]. Information theory is based on randomness/irregularity (interleaver in turbo codes, Tanner graph in LDPC codes, and position of frozen bits in polar codes) and complex decoding algorithms. Efficient decoding architectures on the other hand require low complexity parallel algorithms with large locality. For example, if we are targeting a throughput of 1 Tbit/s with a decoder frequency of 1 GHz, we have to decode 1000 bits in a single clock cycle. Hence, we always have to trade off communication key performance indicators (KPIs), i.e., coding gain, versus implementation KPIs [120]. Implementation KPIs are silicon area, information throughput, latency, power and derived KPIs area efficiency (bit/s/mm²), energy efficiency (pJ/bit), and power density (W/mm²)

$$\text{Area Efficiency} = \frac{\text{Throughput}}{\text{Area}} \quad (4)$$

$$\text{Energy Efficiency} = \frac{\text{Power}}{\text{Throughput}} \quad (5)$$

$$\text{Power Density} = \frac{\text{Power}}{\text{Area}}. \quad (6)$$

If the power consumption of the channel decoder is fixed, e.g., due to thermal design power/power density limits, increasing the throughput is equivalent to improving the energy efficiency by the same factor. Let us assume that the power budget for the channel decoder is 1 W and the throughput is 1 Tbit/s. Then, the necessary energy efficiency is 1 pJ/bit. For comparison, a multiply-accumulate operation in single precision floating-point format needs 1.2 pJ in a 14-nm technology. If the area efficiency is fixed, e.g., due to silicon cost limits, minimizing the power density is equivalent to minimizing the energy efficiency. Thus, minimizing the energy efficiency is a key KPI.

The development of coding schemes that enable high throughput, low latency, and energy-efficient decoder architectures is a multiobjective optimization problem since we have conflicting requirements. For example, a large coding gain demands large block sizes, a large number of iterations in the case of iterative decoding algorithms, and a low code rate. However, energy-efficient, high-throughput, and low-latency decoders require a high code rate, a low number of iterations, and small block sizes. Depending on the application, some objectives may be considered more important than others. Since 6G

has to support a large spectrum of application scenarios with divergent KPI requirements, a deep understanding of the various tradeoffs in the multiobjective design space is mandatory [267]. Another challenge that results from the large application spectrum is the required flexibility. The ultimate goal is a code-agnostic decoding algorithm that can support all different codes, and, from an implementation perspective, a highly regular multidecoder architecture that is composed of a set of low-complexity decoders that are loosely coupled. Flexibility in coding gain, block sizes, and so on should be adjusted by the number of activated cores. Such an architecture enables a reduction in the power and power density if not all cores are activated. For example, a round-robin scheduling can activate different cores to distribute heat. Furthermore, the yield and the reliability can be increased since defective decoders, due to manufacturing defects or aging effects, can be switched off. In [213], an LDPC decoder was presented that pursues this architectural approach. The decoder consists of 16 standard min-sum decoders. Typically, only one decoder has to be activated. However, if the decoding process fails, up to 16 decoders, dependent on the required coding gain, are activated in parallel using a saturated min-sum decoding algorithm in which, e.g., 4 bits with the lowest reliability are enumerated resulting in 16 parallel decoding processes. A gain of 0.7 dB and a power density reduction by one order of magnitude were demonstrated. Ensemble decoding is another promising approach that falls into this architectural category.

C. Ensemble Decoder Architectures

In ensemble decoding, several, preferably low-complexity, decoders work in parallel with minimal information exchange. From a coding perspective, this combination of decoders can either improve the error correction performance or decode different parts of the code word on different cores to achieve blocklength and rate flexibility. The high locality and the low complexity of the individual decoders yield savings in power consumption. As mentioned above, an additional, SNR-dependent disabling of cores and round-robin scheduling can be applied to better distribute heat dissipation across the chip. The efficiency of this approach was demonstrated for automorphism ensemble decoding (AED) implementations in [117]. The discovery of suitable automorphisms of polar codes [87] enables several low complexity SC decoders to work in parallel on permuted representations of the LLRs (Fig. 10). After decoding, the permutations are reversed and the best result is chosen according to a path metric. Compared to state-of-the-art SCL decoders used for 5G, this AED implementation achieves up to 0.5-dB gain in frame error rate (FER) at 10^{-5} , $\times 8.9$ improvement in area efficiency, and $\div 4.6$ improvement in energy efficiency. In addition, the latency of this architecture is constant and independent of the number of cores, i.e., the error correction performance. AED is therefore well suited for 6G URLLC applications.

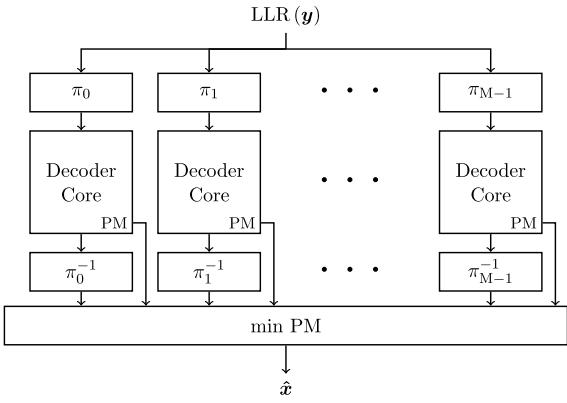


Fig. 10. Automorphism ensemble decoder architecture.

D. Cross Layer Approach/SCL Decoder Evolution

As already stated, a cross-layer approach is very important to achieve efficient implementations. To illustrate the potential of such an approach, we refer to Fig. 11. This figure presents different optimization steps of high-throughput polar SCL- L decoders for polar codes with a code length of 128, code rate $R \approx (1/2)$, and list size $L \in \{8, 4\}$ providing similar error correction performance. All decoders target a similar throughput of 64 Gbit/s and were implemented in a 28-nm technology using the same EDA design flow. The detailed implementation results are presented in Table 2.

The starting point is a state-of-the-art fully unrolled and pipelined CRC aided (CA)-fast simplified successive cancellation list (FSSCL)-8 decoder architecture with full node-level parallelism [119], as shown in Fig. 11(a). Due to numerous generated candidate paths, large area portions are occupied by rate-1 nodes and the corresponding sorters, which are highlighted in red. In an SCL decoder with list size L , according to [97], a rate-1 node with node size N_v generates $L \cdot 2^{\min(L, N_v)}$ from L input paths by applying Chase-like decoding.

Based on this observation, a first optimization step focused on the rate-1 nodes. The candidate paths exhibit a partial order (PO), which can be used in a parallel candidate generation algorithm to eliminate unneeded paths while preserving the error correction performance. A further candidate reduction can be carried out by a double-threshold scheme. This extended threshold partial order (ExPOS)-based rate-1 node candidate generation has a small impact on the error correction but further reduces implementation costs [115] yielding a gain of $\times 2.6$ in area efficiency and $\div 2.6$ in energy efficiency. The resulting decoder is shown in Fig. 11(b).

The red area in this figure marks the single parity check (SPC) nodes [116] with related sorters. A similar optimization as described above is also applied to these nodes. The resulting decoder is shown in Fig. 11(c) and provides further gains of $\times 1.3$ in area efficiency and $\div 1.5$ in energy efficiency.

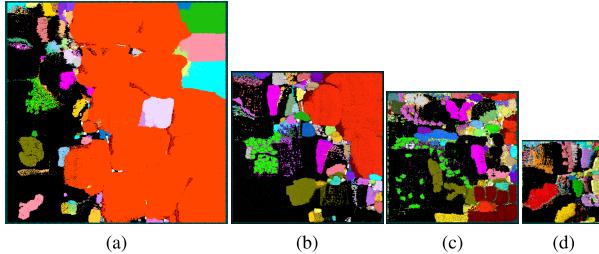


Fig. 11. Layouts of unrolled and pipelined SCL decoder evolution for $N = 128$ in a 28-nm technology with equal area scaling and similar FER performance: (a) CA-Fast-SSCL-8 decoder [119]. (b) ExPOS-rate-1 node optimization [115]. (c) ExPOS-SPC node optimization [116]. (d) Fast-SSCAL-4 decoder [114]. Computational kernels are represented by different colors, and delay line memory is colored in black.

The aforementioned optimizations focused on the nodes only. In the next optimization step, SCL decoding is combined with automorphisms, resulting in successive cancellation automorphism list (SCAL) decoding [114]. The SCL decoder is initialized with permutations of the received data corresponding to selected automorphisms, which compete against each other inside the decoder. This combination improves the error correction capability or a similar error correction compared to SCL and CA-SCL decoding can be achieved with a reduced list size and, thus, a reduced decoder complexity. Fig. 11(d) shows a SCAL-4 decoder of a $\mathcal{P}(128, 60)$ polar code, which is comparable to the aforementioned CA-SCL-8 decoders in terms of code dimension and error correction performance (see also Table 2). Compared to the decoder in Fig. 11(c), a $\times 2.5$ higher area efficiency and $\div 2.2$ better energy efficiency are observed.

V. THOUGHTS ON GUIDELINES FOR 6G CODING

As we have seen in the previous chapters, the last couple of years brought significant advances in code design and decoding algorithms. It is crucial to recognize that coding gain alone is not the sole determinant of system performance. The introduced hardware implementation KPIs emerge as a critical factor, influencing the feasibility and efficiency of deploying advanced coding schemes. Moreover, an important requirement since the very beginning of mobile communications systems is the flexibility in packet

Table 2 ASIC Implementation Results of SCL and SCAL Decoders for Polar Codes With $N = 128$ in a 28-nm Technology

Decoder shown in Fig.	11a	11b	11c	11d
Payload / (bit)	64	64	64	60
E_b/N_0 / (dB) @ FER 10^{-5}	4.8	4.8	4.8	4.3
Coded Throughput / (Gbit/s)	53	64	64	64
Area / (mm^2)	3.15	1.47	1.11	0.44
Area Eff. / (Gbit/s/ mm^2)	16.8	43.6	57.8	145.1
Energy Eff. / (pJ/bit)	62.5	23.6	16.0	7.4

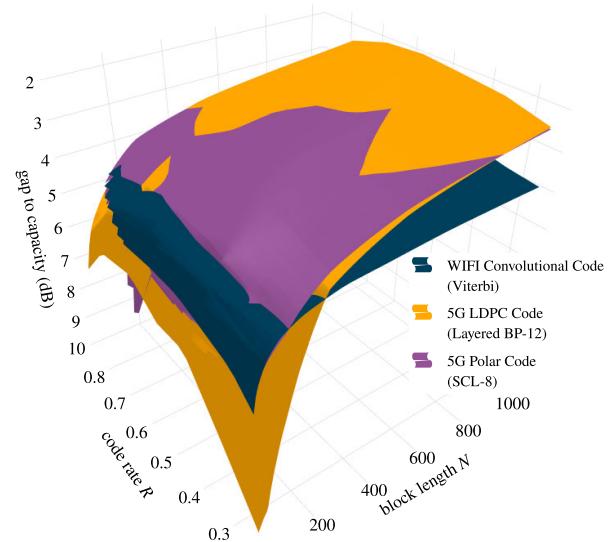


Fig. 12. Required SNR to achieve a BER of 10^{-4} for various blocklengths, code rates, and code/decoder combinations. The combination that performs best is on top [33].

size and coding rate. A scheme fit for 6G must clearly offer such flexibility and, hence, narrow down possible code candidates. The used channel codes and their state-of-the-art decoders from previous standards (such as 5G) serve as a baseline for both performance and flexibility. Fig. 12 illustrates that none of the major previous coding schemes can offer good performance over the whole spectrum of rates and lengths. It shows the required E_b/N_0 to reach a target BER of 10^{-4} for convolutional codes (from Wi-Fi IEEE 802.11, using 64-state Viterbi decoding), 5G polar codes (under SCL-8), and 5G LDPC codes (using layered BP with 12 iterations). We simulated each code/decoder combination for each point a dense grid of parameters R and N and interpolated in between. The surface visible on top is the code/decoder combination that works best in each case. Clearly, the availability of the Viterbi algorithm as a low-complexity ML decoder makes convolutional codes dominate in the short blocklengths (up to $N \approx 128$). Their coding gain, however, does not improve with the blocklength, and hence, they are quickly surpassed by polar codes. Finally, for lengths more larger than $N \approx 512$, the 5G LDPC codes take the lead.

The chart of Fig. 12 suggests the noble objective of a “unified” coding scheme for 6G, that is, to converge to a single family of codes that can be parameterized to different length and rate combinations while operating close to the respective capacity limit. Moreover, the flexibility for decoder implementation, trading off coding gain versus hardware complexity, would constitute a welcome feature.

VI. CONCLUSION

This survey article tried to shed some light on the evolution of channel coding over the past many years, focusing on what we refer to as the “mobile Internet era,” with 6G being its most recent off-spring. Since the advent of

modern channel codes and their corresponding iterative decoding strategies in the early 1990s, the theoretical insights and technological advances have been tremendous. It is no longer just about getting closer to capacity, but flexibility in codeword length and rate adaptivity

has become of prime importance. Together with devising elegant decoder architectures that satisfy the various efficiency metrics, we have come closer to finding a “unified coding” scheme, but more research is required—we are not there yet. ■

REFERENCES

- [1] *Physical Layer Standard for CDMA2000 Spread-Spectrum Systems, Release C*, Standard C.S0002-C Version 1.0, 3rd Gener. Partnership Project 2 (3GPP2), May 2002.
- [2] *Physical Layer for Ultra Mobile Broadband (UMB) Air Interface Specification*, Standard C.S0084-001-0 v3.0, 3rd Gener. Partnership Project 2 (3GPP2), Aug. 2008.
- [3] *Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Multiplexing and Channel Coding (Release 9)*, Standard TS 36.212, 3rd Gener. Partnership Project (3GPP), 2008.
- [4] *Universal Mobile Telecommunications System (UMTS): Multiplexing and Channel Coding (FDD)*, Standard TS 125.212 V3.4.0, 3rd Gener. Partnership Project (3GPP), Sep. 2020.
- [5] *Tech. Specification Group Radio Access Network—NR—Multiplexing Channel Coding*, Standard TS 38.212 V16.5.0, 3rd Gener. Partnership Project (3GPP), Mar. 2021.
- [6] S. M. Abbas, T. Tonnellier, F. Ercan, and W. J. Gross, “High-throughput VLSI architecture for GRAND,” in *Proc. IEEE Workshop Signal Process. Syst. (SiPS)*, Oct. 2020, pp. 1–6.
- [7] S. M. Abbas, T. Tonnellier, F. Ercan, M. Jalaleddine, and W. J. Gross, “High-throughput and energy-efficient VLSI architecture for ordered reliability bits GRAND,” *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 30, no. 6, pp. 681–693, Jun. 2022.
- [8] A. Abbasfar, D. Divsalar, and K. Yao, “Accumulate-repeat-accumulate codes,” *IEEE Trans. Commun.*, vol. 55, no. 4, pp. 692–702, Apr. 2007.
- [9] N. Abramson, “The ALOHA system—Another alternative for computer communications,” in *Proc. Fall Joint Comput. Conf.*, vol. 37, 1970, pp. 281–285.
- [10] N. Abramson, “The throughput of packet broadcasting channels,” *IEEE Trans. Commun.*, vol. COM-25, no. 1, pp. 117–128, Jan. 1977.
- [11] S. Abu-Surra, D. Divsalar, and W. E. Ryan, “Enumerators for protograph-based ensembles of LDPC and generalized LDPC codes,” *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 858–886, Feb. 2011.
- [12] O. Afisisidis, A. Balatsoukas-Stimming, and A. Burg, “A low-complexity improved successive cancellation decoder for polar codes,” in *Proc. 48th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2014, pp. 2116–2120.
- [13] S. M. Aji and R. J. McEliece, “The generalized distributive law,” *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 325–343, Mar. 2000.
- [14] V. K. Amalladinne, J.-F. Chamberland, and K. R. Narayanan, “A coded compressed sensing scheme for unsourced multiple access,” *IEEE Trans. Inf. Theory*, vol. 66, no. 10, pp. 6509–6533, Oct. 2020.
- [15] V. K. Amalladinne, A. K. Pradhan, C. Rush, J. Chamberland, and K. R. Narayanan, “Unsourced random access with coded compressed sensing: Integrating AMP and belief propagation,” *IEEE Trans. Inf. Theory*, vol. 68, no. 4, pp. 2384–2409, Apr. 2022.
- [16] B. Amiri, C.-W. Lin, and L. Dolecek, “Asymptotic distribution of absorbing sets and fully absorbing sets for regular sparse code ensembles,” *IEEE Trans. Commun.*, vol. 61, no. 2, pp. 455–464, Feb. 2013.
- [17] K. Andreev, P. Rybin, and A. Frolov, “Coded compressed sensing with list recoverable codes for the unsourced random access,” *IEEE Trans. Commun.*, vol. 70, no. 12, pp. 7886–7898, Dec. 2022.
- [18] K. S. Andrews, D. Divsalar, S. Dolinar, J. Hamkins, C. R. Jones, and F. Pollara, “The development of turbo and LDPC codes for deep-space applications,” *Proc. IEEE*, vol. 95, no. 11, pp. 2142–2156, Nov. 2007.
- [19] V. Aref, L. Schmalen, and S. ten Brink, “On the convergence speed of spatially coupled LDPC ensembles,” in *Proc. 51st Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Oct. 2013, pp. 342–349.
- [20] V. Aref, N. Rengaswamy, and L. Schmalen, “Finite-length analysis of spatially-coupled regular LDPC ensembles on burst-erasure channels,” *IEEE Trans. Inf. Theory*, vol. 64, no. 5, pp. 3431–3449, May 2018.
- [21] E. Arikan, “Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels,” *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3051–3073, Jul. 2009.
- [22] E. Arikan, “From sequential decoding to channel polarization and back again,” 2019, *arXiv:1908.09594*.
- [23] E. Arikan and E. Telatar, “On the rate of channel polarization,” in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2009, pp. 1493–1495.
- [24] A. Ashikhmin, G. Kramer, and S. ten Brink, “Extrinsic information transfer functions: Model and erasure channel properties,” *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2657–2673, Nov. 2004.
- [25] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, “Optimal decoding of linear codes for minimizing symbol error rate,” *IEEE Trans. Inf. Theory*, vol. IT-20, no. 2, pp. 284–287, Mar. 1974.
- [26] A. Balatsoukas-Stimming, M. B. Parizi, and A. Burg, “LLR-based successive cancellation list decoding of polar codes,” *IEEE Trans. Signal Process.*, vol. 63, no. 19, pp. 5165–5179, Oct. 2015.
- [27] S. Barman, X. Liu, S. C. Draper, and B. Recht, “Decomposition methods for large scale LP decoding,” *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 7870–7886, Dec. 2013.
- [28] E. B. Yacoub, “Trapping and absorbing set enumerators for irregular generalized low-density parity-check code ensembles,” *IEEE Trans. Inf. Theory*, vol. 69, no. 6, pp. 3637–3662, Jun. 2023.
- [29] S. Benedetto and G. Montorsi, “Serial concatenation of block and convolutional codes,” *Electron. Lett.*, vol. 32, no. 10, p. 887, 1996.
- [30] S. Benedetto and G. Montorsi, “Unveiling turbo codes: Some results on parallel concatenated coding schemes,” *IEEE Trans. Inf. Theory*, vol. 42, no. 2, pp. 409–428, Mar. 1996.
- [31] E. Berlekamp, *Algebraic Coding Theory*. New York, NY, USA: McGraw-Hill, 1961.
- [32] C. Berrou, A. Glavieux, and P. Thitimajshima, “Near Shannon limit error-correcting coding and decoding: Turbo-codes. 1,” in *Proc. IEEE Int. Conf. Commun.*, Geneva, Switzerland, May 1993, pp. 1064–1070.
- [33] P. Bezner and M. Geiselhart. (2023). *CoComBA: A Code Comparison and Benchmarking Assistant*. [Online]. Available: <https://cocomba.inue.uni-stuttgart.de/>
- [34] V. Bioglio, C. Condo, and I. Land, “Design of polar codes in 5G new radio,” *IEEE Commun. Surveys Tuts.*, vol. 23, no. 1, pp. 29–40, 1st Quart., 2021.
- [35] F. Boccardi, R. W. Heath Jr., A. Lozano, T. L. Marzetta, and P. Popovski, “Five disruptive technology directions for 5G,” *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.
- [36] G. Böcherer, P. Schulte, and F. Steiner, “Probabilistic shaping and forward error correction for fiber-optic communication systems,” *J. Lightw. Technol.*, vol. 37, no. 2, pp. 230–244, Jan. 15, 2019.
- [37] G. Böcherer, P. Schulte, and F. Steiner, “Bandwidth efficient and rate-matched low-density parity-check coded modulation,” *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 4651–4665, Dec. 2015.
- [38] M. Breiling, “A logarithmic upper bound on the minimum distance of turbo codes,” *IEEE Trans. Inf. Theory*, vol. 50, no. 8, pp. 1692–1710, Aug. 2004.
- [39] F. Buchali, F. Steiner, G. Bocherer, L. Schmalen, P. Schulte, and W. Idler, “Rate adaptation and reach increase by probabilistically shaped 64-QAM: An experimental demonstration,” *J. Lightw. Technol.*, vol. 34, no. 7, pp. 1599–1609, Apr. 1, 2016.
- [40] D. Burshtein and G. Miller, “Asymptotic enumeration methods for analyzing LDPC codes,” *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1115–1131, Jun. 2004.
- [41] G. Caire, G. Taricco, and E. Biglieri, “Bit-interleaved coded modulation,” *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 927–946, May 1998.
- [42] S. Cammerer, V. Aref, L. Schmalen, and S. ten Brink, “Triggering wave-like convergence of tail-biting spatially coupled LDPC codes,” in *Proc. Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2016, pp. 93–98.
- [43] S. Cammerer, L. Schmalen, V. Aref, and S. ten Brink, “Wave-like decoding of tail-biting spatially coupled LDPC codes through iterative demapping,” in *Proc. 9th Int. Symp. Turbo Codes Iterative Inf. Process. (ISTC)*, Sep. 2016, pp. 121–125.
- [44] E. Casini, R. De Gaudenzi, and O. del Rio Herrero, “Contention resolution diversity slotted Aloha (CRDSA): An enhanced random access scheme for satellite access packet networks,” *IEEE Trans. Wireless Commun.*, vol. 6, no. 4, pp. 1408–1419, Apr. 2007.
- [45] *TM Synchronization and Channel Coding*, CCSDS Standard 131.O.B.2, Blue Book, Aug. 2011.
- [46] M. Chafii, L. Bariah, S. Muhaidat, and M. Debbah, “Twelve scientific challenges for 6G: Rethinking the foundations of communications theory,” *IEEE Commun. Surveys Tuts.*, vol. 25, no. 2, pp. 868–904, 2nd Quart., 2023.
- [47] T.-Y. Chen, K. Vakilian, D. Divsalar, and R. D. Wesel, “Protograph-based raptor-like LDPC codes,” *IEEE Trans. Commun.*, vol. 63, no. 5, pp. 1522–1532, May 2015.
- [48] S. K. Chilappagari, D. V. Nguyen, B. Vasic, and M. W. Marcellin, “On trapping set and guaranteed error correction capability of LDPC codes and GLDPC codes,” *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1600–1611, Apr. 2010.
- [49] S.-Y. Chung, G. D. Forney, T. J. Richardson, and R. Urbanke, “On the design of low-density parity-check codes within 0.0045 dB of the Shannon limit,” *IEEE Commun. Lett.*, vol. 5, no. 2, pp. 58–60, Feb. 2001.
- [50] F. Clazzer, C. Kissling, and M. Marchese, “Enhancing contention resolution ALOHA using combining techniques,” *IEEE Trans. Commun.*, vol. 66, no. 6, pp. 2576–2587, Jun. 2018.
- [51] C. Condo, “A fixed latency ORBGRAND decoder architecture with LUT-aided error-pattern scheduling,” *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 69, no. 5, pp. 2203–2211,

- May 2022.
- [52] D. J. Costello, J. Hagenauer, H. Imai, and S. B. Wicker, "Applications of error-control coding," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2531–2560, Oct. 1998.
- [53] D. J. Costello, Jr. and G. D. Forney, "Channel coding: The road to channel capacity," *Proc. IEEE*, vol. 95, no. 6, pp. 1150–1177, Jun. 2007.
- [54] M. C. Coşkun et al., "Efficient error-correcting codes in the short blocklength regime," *Phys. Commun.*, vol. 34, pp. 66–79, Jun. 2019.
- [55] S. Crozier and P. Guinand, "High-performance low-memory interleaver banks for turbo-codes," in *Proc. IEEE Veh. Technol. Conf. Fall*, Atlantic City, NJ, USA, Oct. 2001, pp. 2394–2398.
- [56] S. Crozier and P. Guinand, "Distance upper bounds and true minimum distance results for turbo codes designed with DRP interleavers," in *Proc. Int. Symp. Turbo Codes Rel. Topics*, Brest, France, Sep. 2003.
- [57] M. C. Davey and D. J. C. MacKay, "Low density parity check codes over GF(q)," in *Proc. Inf. Theory Workshop*, Killarney, Ireland, Jun. 1998, pp. 70–71, doi: [10.1109/ITW1998.706440](https://doi.org/10.1109/ITW1998.706440).
- [58] I. Debbabi, B. Le Gal, N. Khouja, F. Tili, and C. Jego, "Analysis of ADMM-LP algorithm for LDPC decoding, a first step to hardware implementation," in *Proc. IEEE Int. Conf. Electron., Circuits, Syst. (ICECS)*, Dec. 2015, pp. 356–359.
- [59] C. Di, D. Proietti, I. E. Telatar, T. J. Richardson, and R. L. Urbanke, "Finite-length analysis of low-density parity-check codes on the binary erasure channel," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1570–1579, Jun. 2002.
- [60] C. Di, T. J. Richardson, and R. L. Urbanke, "Weight distribution of low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 4839–4855, Nov. 2006.
- [61] D. Divsalar, S. Dolinar, C. Jones, and K. Andrews, "Capacity-approaching protograph codes," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 6, pp. 876–888, Aug. 2009.
- [62] D. Divsalar, S. Dolinar, and F. Pollara, "Iterative turbo decoder analysis based on density evolution," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 5, pp. 891–907, May 2001.
- [63] D. Divsalar, H. Jin, and R. J. McEliece, "Coding theorems for turbo-like codes," in *Proc. Annu. Allerton Conf. Commun. Control. Comput.*, Monticello, IL, USA, Sep. 1998, pp. 201–210.
- [64] L. Dolecek, D. Divsalar, Y. Sun, and B. Amiri, "Non-binary protograph-based LDPC codes: Enumerators, analysis, and designs," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 3913–3941, Jul. 2014.
- [65] B. Dorsch, "A decoding algorithm for binary block codes and J-ary output channels (corresp.)," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 3, pp. 391–394, May 1974.
- [66] C. Douillard and C. Berrou, "Turbo codes with rate-m/(m+1) constituent convolutional codes," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 1630–1638, Oct. 2005.
- [67] K. R. Duffy, W. An, and M. Médard, "Ordered reliability bits guessing random additive noise decoding," *IEEE Trans. Signal Process.*, vol. 70, pp. 4528–4542, 2022.
- [68] K. R. Duffy, M. Médard, and W. An, "Guessing random additive noise decoding with symbol reliability information (SRGRAND)," *IEEE Trans. Commun.*, vol. 70, no. 1, pp. 3–18, Jan. 2022.
- [69] H. El Gamal and A. R. Hammons, "Analyzing the turbo decoder using the Gaussian approximation," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 671–686, Feb. 2001.
- [70] P. Elias, "Coding for noisy channels," *IRE Conv. Rec.*, vol. 3, no. 4, pp. 37–46, 1955.
- [71] *Frame Structure, Channel Coding and Modulation for a Second Generation Digital Terrestrial Television Broadcasting System (DVB-T2)*, ETSI Standard EN 302 755 v1.1.2, 2009.
- [72] *Interaction Channel for Satellite Distribution Systems (DVB-RCS)*, ETSI Standard EN 301 790 v1.5.1, 2009.
- [73] *Second Generation Framing Structure, Channel Coding and Modulation Systems for Broadcasting, Interactive Services, News Gathering and Other Broadband Satellite Applications (DVB-S2)*, ETSI Standard EN 302 307 v1.2.1, 2009.
- [74] T. Etzion, A. Trachtenberg, and A. Vardy, "Which codes have cycle-free Tanner graphs?" *IEEE Trans. Inf. Theory*, vol. 45, no. 6, pp. 2173–2181, Sep. 1999.
- [75] D. Fathollahi, N. Farsad, S. A. Hashemi, and M. Mondelli, "Sparse multi-decoder recursive projection aggregation for Reed–Muller codes," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Melbourne, VIC, Australia, Jul. 2021, pp. 1082–1087, doi: [10.1109/ISIT45174.2021.9517887](https://doi.org/10.1109/ISIT45174.2021.9517887).
- [76] J. Feldman, M. J. Wainwright, and D. R. Karger, "Using linear programming to decode binary linear codes," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 954–972, Mar. 2005.
- [77] A. J. Felstrom, D. Truhachev, M. Lentmaier, and K. S. Zigangirov, "Braided block codes," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2640–2658, Jun. 2009.
- [78] A. Fengler, P. Jung, and G. Caire, "SPARCs for unsourced random access," *IEEE Trans. Inf. Theory*, vol. 67, no. 10, pp. 6894–6915, Oct. 2021.
- [79] G. Forney, R. Gallager, G. Lang, F. Longstaff, and S. Qureshi, "Efficient modulation for band-limited channels," *IEEE J. Sel. Areas Commun.*, vol. SAC-2, no. 5, pp. 632–647, Sep. 1984.
- [80] G. D. Forney, "Codes on graphs: Normal realizations," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 520–548, Feb. 2001.
- [81] G. D. Forney Jr., "Generalized minimum distance decoding," *IEEE Trans. Inf. Theory*, vol. IT-12, no. 2, pp. 125–131, Apr. 1966.
- [82] M. P. C. Fossorier, "Quasi-cyclic low-density parity-check codes from circulant permutation matrices," *IEEE Trans. Inf. Theory*, vol. 50, no. 8, pp. 1788–1793, Aug. 2004.
- [83] M. P. C. Fossorier and S. Lin, "Soft-decision decoding of linear block codes based on ordered statistics," *IEEE Trans. Inf. Theory*, vol. 41, no. 5, pp. 1379–1396, Sep. 1995.
- [84] B. J. Frey, R. Koetter, and A. Vardy, "Signal-space characterization of iterative decoding," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 766–781, Feb. 2001.
- [85] R. Gallager, "Low-density parity-check codes," Ph.D. dissertation, Massachusetts Inst. Technol., Cambridge, MA, USA, 1963.
- [86] R. Garello, G. Montorsi, S. Benedetto, and G. Cancellieri, "Interleaver properties and their applications to the trellis complexity analysis of turbo codes," *IEEE Trans. Commun.*, vol. 49, no. 5, pp. 793–807, May 2001.
- [87] M. Geiselhart, A. Elkelesh, M. Ebada, S. Cammerer, and S. ten Brink, "On the automorphism group of polar codes," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2021, pp. 1230–1235.
- [88] M. Geiselhart, F. Krieg, J. Clausius, D. Tandler, and S. ten Brink, "6G: A welcome chance to unify channel coding?" *IEEE BITS Inf. Theory Mag.*, vol. 3, no. 1, pp. 67–80, Mar. 2023.
- [89] F. Gensheimer, T. Dietz, K. Kraft, S. Ruzika, and N. Wehn, "A reduced-complexity projection algorithm for ADMM-based LP decoding," *IEEE Trans. Inf. Theory*, vol. 66, no. 8, pp. 4819–4833, Aug. 2020.
- [90] A. Ghazisaeidi et al., "Advanced C+L-band transoceanic transmission systems based on probabilistically shaped PDM-64QAM," *J. Lightw. Technol.*, vol. 35, no. 7, pp. 1291–1299, Jan. 23, 2017.
- [91] A. Ghazisaeidi et al., "Transoceanic transmission systems using adaptive multirate FECs," *J. Lightw. Technol.*, vol. 33, no. 7, pp. 1479–1487, Feb. 5, 2015.
- [92] A. Graell i Amat and L. Schmalen, "Forward error correction for optical transponders," in *Springer Handbook of Optical Networks* (Springer Handbooks), B. Mukherjee, I. Tomkos, M. Tornatore, P. Winzer, and Y. Zhao, Eds. Cham, Switzerland: Springer, 2020, doi: [10.1007/978-3-030-16250-4_7](https://doi.org/10.1007/978-3-030-16250-4_7).
- [93] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Trans. Inf. Theory*, vol. 42, no. 2, pp. 429–445, Mar. 1996.
- [94] C. Häger, A. Graell i Amat, F. Bränström, A. Alvarado, and E. Agrell, "Terminated and tailbiting spatially coupled codes with optimized bit mappings for spectrally efficient fiber-optical systems," *J. Lightw. Technol.*, vol. 33, no. 7, pp. 1275–1285, Apr. 1, 2015.
- [95] C. Häger and H. D. Pfister, "Approaching miscorrection-free performance of product codes with anchor decoding," *IEEE Trans. Commun.*, vol. 66, no. 7, pp. 2797–2808, Jul. 2018.
- [96] T. Harter et al., "Generalized Kramers–Kronig receiver for coherent terahertz communications," *Nature Photon.*, vol. 14, no. 10, pp. 601–606, Oct. 2020.
- [97] S. A. Hashemi, C. Condo, and W. J. Gross, "Fast simplified successive-cancellation list decoding of polar codes," in *Proc. IEEE Wireless Commun. Netw. Conf. Workshops (WCNCW)*, Mar. 2017, pp. 1–6.
- [98] M. Hashemipour-Nazari, K. Goossens, and A. Balatsoukas-Stimming, "Hardware implementation of iterative projection-aggregation decoding of Reed–Muller codes," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2021, pp. 8293–8297.
- [99] M. Hashemipour-Nazari, K. Goossens, and A. Balatsoukas-Stimming, "Multi-factor pruning for recursive projection-aggregation decoding of RM codes," in *Proc. IEEE Workshop Signal Process. Syst. (SiPS)*, Nov. 2022, pp. 1–6.
- [100] M. Hashemipour-Nazari, Y. Ren, K. Goossens, and A. Balatsoukas-Stimming, "Pipelined architecture for soft-decision iterative projection aggregation decoding for RM codes," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 70, no. 12, pp. 5468–5481, Dec. 2023.
- [101] D. E. Hocevar, "A reduced complexity decoder architecture via layered decoding of LDPC codes," in *Proc. IEEE Workshop Signal Process. Syst. (SiPS)*, Austin, TX, USA, Oct. 2004, pp. 107–112.
- [102] C. Häger, A. Graell i Amat, A. Alvarado, and E. Agrell, "Design of APSK constellations for coherent optical channels with nonlinear phase noise," *IEEE Trans. Commun.*, vol. 61, no. 8, pp. 3362–3373, Aug. 2013.
- [103] I. Dumer and K. Shabunov, "Recursive decoding of Reed–Muller codes," in *Proc. IEEE Int. Symp. Inf. Theory*, Oct. 1999, pp. 61–69.
- [104] I. Dumer and K. Shabunov, "Recursive list decoding of Reed–Muller codes," in *Information, Coding and Mathematics*, M. Blaum, P. Farrell, and H. van Tilborg, Eds. Boston, MA, USA: Kluwer, 2002, pp. 279–298.
- [105] *Wireless LAN Medium Access Control and Physical Layer specifications: Enhancements for Higher Throughput*, IEEE Standard P802.11n D1.0, 2006.
- [106] IEEE Spectrum. (2015). *50 Years of Moore's Law*. [Online]. Available: <https://spectrum.ieee.org/special-reports/50-years-of-moores-law/>
- [107] 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, May 1977.
- [108] *A Family of 2-Wire, Duplex Modems Operating at Data Signalling Rates of up to 9600 Bit/s for Use on the General Switched Telephone Network and on Leased Telephone-type Circuits*, Standard V.32, Int. Telecommun. Union (ITU), 1988.
- [109] A. R. Iyengar, M. Papaleo, P. H. Siegel, J. K. Wolf, A. Vanelli-Coralli, and G. E. Corazza, "Window decoding of protograph-based LDPC convolutional codes over erasure channels," *IEEE Trans. Inf. Theory*, vol. 58, no. 4, pp. 2303–2320, Apr. 2012.
- [110] W. Jiang, B. Han, M. A. Habibi, and H. D. Schotten, "The road towards 6G: A comprehensive survey," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 334–366, 2021.
- [111] A. J. Felstrom and K. S. Zigangirov, "Time-varying

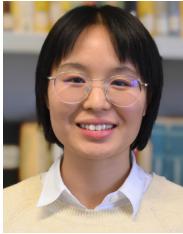
- periodic convolutional codes with low-density parity-check matrix," *IEEE Trans. Inf. Theory*, vol. 45, no. 6, pp. 2181–2191, Sep. 1999.
- [112] H. Jin, A. Khandekar, and R. McEliece, "Irregular repeat-accumulate codes," in *Proc. Int. Symp. Turbo Codes Rel. Topics*, Brest, France, Sep. 2000.
- [113] H. Jin and R. J. McEliece, "Coding theorems for turbo code ensembles," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1451–1461, Jun. 2002.
- [114] L. Johannsen, C. Kestel, M. Geiselhart, T. Vogt, S. ten Brink, and N. Wehn, "Successive cancellation automorphism list decoding of polar codes," in *Proc. 12th Int. Symp. Topics Coding (ISTC)*, Sep. 2023, pp. 1–5.
- [115] L. Johannsen, C. Kestel, O. Griebel, T. Vogt, and N. Wehn, "Partial order-based decoding of rate-1 nodes in fast simplified successive-cancellation list decoders for polar codes," *Electronics*, vol. 11, no. 4, p. 560, Feb. 2022.
- [116] L. Johannsen, C. Kestel, T. Vogt, and N. Wehn, "Parallel single parity check nodes for high-throughput fast-SSCL polar code decoders," in *Proc. IEEE Symp. Future Telecommun. Technol. (SOFTT)*, Johor Bahru, Malaysia, Nov. 2022, pp. 28–34.
- [117] C. Kestel, M. Geiselhart, L. Johannsen, S. ten Brink, and N. Wehn, "Automorphism ensemble polar code decoders for 6G URLLC," in *Proc. 26th Int. ITG Workshop Smart Antennas 13th Conf. Syst., Commun., Coding*, Braunschweig, Germany, Feb. 2023, pp. 1–6.
- [118] C. Kestel, M. Herrmann, and N. When, "When channel coding hits the implementation wall," in *Proc. IEEE 10th Int. Symp. Turbo Codes Iterative Inf. Process. (ISTC)*, Dec. 2018, pp. 1–6.
- [119] C. Kestel et al., "A 506 Gbit/s polar successive cancellation list decoder with CRC," in *Proc. IEEE 31st Annu. Int. Symp. Pers., Indoor Mobile Radio Commun.*, London, U.K., Aug. 2020, pp. 1–7.
- [120] E. Kiehn, N. Wehn, and H. Meyer, "On complexity, energy- and implementation-efficiency of channel decoders," *IEEE Trans. Commun.*, vol. 59, no. 12, pp. 3301–3310, Dec. 2011.
- [121] J. Kim, G. Lee, S. Kim, T. Taleb, S. Choi, and S. Bahk, "Two-step random access for 5G system: Latest trends and challenges," *IEEE Netw.*, vol. 35, no. 1, pp. 273–279, Jan. 2021.
- [122] K. Kläber, S. Cammerer, L. Schmalen, and S. ten Brink, "Avoiding burst-like error patterns in windowed decoding of spatially coupled LDPC codes," in *Proc. IEEE 10th Int. Symp. Turbo Codes Iterative Inf. Process. (ISTC)*, Dec. 2018, pp. 1–5.
- [123] R. Koetter, W.-C.-W. Li, P. O. Vontobel, and J. L. Walker, "Characterizations of pseudo-codewords of (low-density) parity-check codes," *Adv. Math.*, vol. 213, no. 1, pp. 205–229, Aug. 2007.
- [124] Y. Kou, S. Lin, and M. P. C. Fossorier, "Low-density parity-check codes based on finite geometries: A rediscovery and new results," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 2711–2736, Nov. 2001.
- [125] S. S. Kowshik, K. Andreev, A. Frolov, and Y. Polyanskiy, "Energy efficient coded random access for the wireless uplink," *IEEE Trans. Commun.*, vol. 68, no. 8, pp. 4694–4708, Aug. 2020.
- [126] K. Kraft and N. Wehn, "ADMM-based ML decoding: From theory to practice," in *Proc. ICASSP - IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2021, pp. 8278–8282.
- [127] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.
- [128] F. R. Kschischang and S. Pasupathy, "Optimal nonuniform signaling for Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 913–929, May 1993.
- [129] S. Kudekar, T. J. Richardson, and R. L. Urbanke, "Threshold saturation via spatial coupling: Why convolutional LDPC ensembles perform so well over the BEC," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 803–834, Feb. 2011.
- [130] S. Kudekar, S. Kumar, M. Mondelli, H. D. Pfister, E. Sasoglu, and R. L. Urbanke, "Reed-Muller codes achieve capacity on erasure channels," *IEEE Trans. Inf. Theory*, vol. 63, no. 7, pp. 4298–4316, Jul. 2017.
- [131] S. Kudekar, T. Richardson, and R. L. Urbanke, "Spatially coupled ensembles universally achieve capacity under belief propagation," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 7761–7813, Dec. 2013.
- [132] S. Kumar, A. J. Young, N. Macris, and H. D. Pfister, "Threshold saturation for spatially coupled LDPC and LDGM codes on BMS channels," *IEEE Trans. Inf. Theory*, vol. 60, no. 12, pp. 7389–7415, Dec. 2014.
- [133] N. Magotra and J. Larimer, "Energy efficient digital signal processing," in *Proc. 53rd IEEE Int. Midwest Symp. Circuits Syst.*, Aug. 2010, pp. 1053–1056.
- [134] Y. Lei et al., "Improved decoding of staircase codes: The soft-aided bit-marking (SABM) algorithm," *IEEE Trans. Commun.*, vol. 67, no. 12, pp. 8220–8232, Dec. 2019.
- [135] M. Lentmaier, G. P. Fettweis, K. Sh. Zigangirov, and D. J. Costello, "Approaching capacity with asymptotically regular LDPC codes," in *Proc. Inf. Theory Appl. Workshop*, La Jolla, CA, USA, Feb. 2009, pp. 173–177.
- [136] M. Lentmaier, D. G. M. Mitchell, G. Fettweis, and D. J. Costello, "Asymptotically good LDPC convolutional codes with AWGN channel thresholds close to the Shannon limit," in *Proc. 6th Int. Symp. Turbo Codes Iterative Inf. Process.*, Brest, France, Sep. 2010, pp. 324–328.
- [137] M. Lentmaier, A. Sridharan, D. J. Costello, and K. S. Zigangirov, "Iterative decoding threshold analysis for LDPC convolutional codes," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 5274–5289, Oct. 2010.
- [138] M. Lentmaier and K. S. Zigangirov, "On generalized low-density parity-check codes based on Hamming component codes," *IEEE Commun. Lett.*, vol. 3, no. 8, pp. 248–250, Aug. 1999.
- [139] A. Leven and L. Schmalen, "Status and recent advances on forward error correction technologies for lightwave systems," *IEEE J. Lightw. Technol.*, vol. 32, no. 16, pp. 2735–2750, Aug. 15, 2014.
- [140] B. Li, H. Shen, and D. Tse, "An adaptive successive cancellation list decoder for polar codes with cyclic redundancy check," *IEEE Commun. Lett.*, vol. 16, no. 12, pp. 2044–2047, Nov. 2012.
- [141] J. Li, S. M. Abbas, T. Tonnellier, and W. J. Gross, "Reduced complexity RPA decoder for Reed-Muller codes," in *Proc. 11th Int. Symp. Topics Coding (ISTC)*, Montreal, QC, Canada, 2021, pp. 1–5, doi: 10.1109/ISTC49272.2021.9594060.
- [142] Z. Li, L. Chen, L. Zeng, S. Lin, and W. Fong, "Efficient encoding of quasi-cyclic low-density parity codes," *IEEE Trans. Commun.*, vol. 54, no. 1, pp. 71–81, Jan. 2006.
- [143] Z. Li, L. Zhang, Y. Shen, A. Burg, X. You, and C. Zhang, "List ordered statistics decoders for polar codes," in *Proc. 56th Asilomar Conf. Signals, Syst., Comput.*, Oct. 2022, pp. 628–633.
- [144] M. Lian, C. Häger, and H. D. Pfister, "Decoding Reed-Muller codes using redundant code constraints," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Los Angeles, CA, USA, Jun. 2020, pp. 42–47, doi: 10.1109/ISIT44484.2020.9174087.
- [145] G. Liga, A. Sheikh, and A. Alvarado, "A novel soft-aided bit-marking decoder for product codes," in *Proc. 45th Eur. Conf. Opt. Commun. (ECOC)*, Sep. 2019, pp. 1–4.
- [146] S. Litsyn and V. Shevelev, "Distance distributions in ensembles of irregular low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 49, no. 12, pp. 3140–3159, Dec. 2003.
- [147] G. Liva, "Graph-based analysis and optimization of contention resolution diversity slotted Aloha," *IEEE Trans. Commun.*, vol. 59, no. 2, pp. 477–487, Feb. 2011.
- [148] G. Liva and M. Chiani, "Photograph LDPC codes design based on EXIT analysis," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Washington, DC, USA, Nov. 2007, pp. 3250–3254.
- [149] G. Liva and Y. Polyanskiy, "On coding techniques for unsourced multiple-access," in *Proc. 55th Asilomar Conf. Signals, Syst., Comput.*, Oct. 2021, pp. 1507–1514.
- [150] G. Liva and Y. Polyanskiy, "Unsourced multiple access: A coding paradigm for massive random access," *Proc. IEEE*, to be published.
- [151] G. Liva, E. Paolini, B. Matuz, S. Scalise, and M. Chiani, "Short turbo codes over high order fields," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2201–2211, Jun. 2013.
- [152] M. Luby, M. Mitzenmacher, and A. Shokrollahi, "Analysis of random processes via and-or tree evaluation," in *Proc. 9th Annu. ACM-SIAM Symp. Discrete Algs.* San Francisco, CA, USA, 1998, pp. 364–373.
- [153] M. G. Luby, M. Mitzenmacher, M. A. Shokrollahi, D. A. Spielman, and V. Stemann, "Practical loss-resilient codes," in *Proc. 29th Annu. ACM Symp. Theory Comput.*, El Paso, TX, USA, 1997, pp. 150–159.
- [154] M. Luby et al., "Improved low-density parity-check codes using irregular graphs," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 585–598, Feb. 2001.
- [155] D. MacKay and R. Neal, "Good codes based on very sparse matrices," in *Proc. 5th IMA Int. Conf. Cryptography Coding*, in Lecture Notes in Computer Science, vol. 1025, C. Boyd, Ed. Berlin, Germany: Springer, Dec. 1995, pp. 100–111.
- [156] D. J. C. MacKay and M. S. Postol, "Weaknesses of Margulis and Ramanujan–Margulis low-density parity-check codes," *Electron. Notes Theor. Comput. Sci.*, vol. 74, pp. 97–104, Oct. 2003.
- [157] N. H. Mahmood, I. Atzeni, E. A. Jorswieck, and O. L. A. López, "Ultra-reliable low-latency communications: Foundations, enablers, system design, and evolution towards 6G," *Found. Trends Commun. Inf. Theory*, vol. 20, nos. 5–6, pp. 512–747, 2023.
- [158] M. M. Mansour and N. R. Shanbhag, "High-throughput LDPC decoders," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 11, no. 6, pp. 976–996, Dec. 2003.
- [159] R. McEliece, D. MacKay, and J.-F. Cheng, "Turbo decoding as an instance of Pearl's 'belief propagation' algorithm," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 2, pp. 140–152, Feb. 1998.
- [160] C. Measson, A. Montanari, and R. Urbanke, "An upper-bound for the ML threshold of iterative coding systems over the BEC," in *Proc. Ann. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, USA, Sep. 2003.
- [161] C. Measson, A. Montanari, and R. Urbanke, "Maxwell construction: The hidden bridge between iterative and maximum a posteriori decoding," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5277–5307, Dec. 2008.
- [162] S. Miao, L. Rapp, and L. Schmalen, "Improved soft-aided decoding of product codes with adaptive performance-complexity trade-off," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Sep. 2022, pp. 1–4.
- [163] S. Miao, L. Rapp, and L. Schmalen, "Improved soft-aided decoding of product codes with dynamic reliability scores," *J. Lightw. Technol.*, vol. 40, no. 22, pp. 7279–7288, Aug. 26, 2022.
- [164] S. Miao, L. Rapp, and L. Schmalen, "Improved soft-aided error- and erasure decoding of product codes with dynamic reliability scores," in *Proc. Opt. Fiber Commun. Conf. Exhib. (OFC)*, Mar. 2022, pp. 1–3.
- [165] O. Milenkovic, E. Soljanin, and P. Whiting, "Asymptotic spectra of trapping sets in regular and irregular LDPC code ensembles," *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 39–55, Jan. 2007.
- [166] D. G. M. Mitchell, M. Lentmaier, and D. J. Costello, "Spatially coupled LDPC codes constructed from photographs," *IEEE Trans. Inf. Theory*, vol. 61, no. 9, pp. 4866–4889,

- Sep. 2015.
- [167] S. Moloudi, M. Lentmaier, and A. Graell i Amat, "Spatially coupled turbo-like codes," *IEEE Trans. Inf. Theory*, vol. 63, no. 10, pp. 6199–6215, Oct. 2017.
- [168] M. Mondelli, S. H. Hassani, and R. L. Urbanke, "Unified scaling of polar codes: Error exponent, scaling exponent, moderate deviations, and error floors," *IEEE Trans. Inf. Theory*, vol. 62, no. 12, pp. 6698–6712, Dec. 2016.
- [169] D. E. Muller, "Application of Boolean algebra to switching circuit design and to error detection," *Trans. IRE Prof. Group Electron. Comput.*, vol. 3, no. 3, pp. 6–12, 1954.
- [170] K. R. Narayanan and H. D. Pfister, "Iterative collision resolution for slotted ALOHA: An optimal uncoordinated transmission policy," in *Proc. 7th Int. Symp. Turbo Codes Iterative Inf. Process. (ISTC)*, Gothenburg, Sweden, Aug. 2012, pp. 136–139.
- [171] K. R. Narayanan and G. L. Stuber, "List decoding of turbo codes," *IEEE Trans. Commun.*, vol. 46, no. 6, pp. 754–762, Jun. 1998.
- [172] C. Nill and C.-E.-W. Sundberg, "List and soft symbol output Viterbi algorithms: Extensions and comparisons," *IEEE Trans. Commun.*, vol. 43, no. 2/3/4, pp. 277–287, Feb. 1995.
- [173] K. Niu and K. Chen, "CRC-aided decoding of polar codes," *IEEE Commun. Lett.*, vol. 16, no. 10, pp. 1668–1671, Oct. 2012.
- [174] A. Orlitsky, K. Viswanathan, and J. Zhang, "Stopping set distribution of LDPC code ensembles," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 929–953, Mar. 2005.
- [175] P. Oswald and A. Shokrollahi, "Capacity-achieving sequences for the erasure channel," *IEEE Trans. Inf. Theory*, vol. 48, no. 12, pp. 3017–3028, Dec. 2002.
- [176] E. Paolini, G. Liva, and M. Chiani, "Coded slotted ALOHA: A graph-based method for uncoordinated multiple access," *IEEE Trans. Inf. Theory*, vol. 61, no. 12, pp. 6815–6832, Dec. 2015.
- [177] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1988.
- [178] L. C. Perez, J. Seghers, and D. J. Costello, "A distance spectrum interpretation of turbo codes," *IEEE Trans. Inf. Theory*, vol. 42, no. 6, pp. 1698–1709, Nov. 1996.
- [179] H. D. Pfister, I. Sason, and R. Urbanke, "Capacity-achieving ensembles for the binary erasure channel with bounded complexity," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2352–2379, Jul. 2005.
- [180] H. D. Pfister and P. H. Siegel, "The serial concatenation of rate-1 codes through uniform random interleavers," *IEEE Trans. Inf. Theory*, vol. 49, no. 6, pp. 1425–1438, Jun. 2003.
- [181] J. Piao, K. Niu, J. Dai, and C. Dong, "Approaching the normal approximation of the finite blocklength capacity within 0.025 dB by short polar codes," *IEEE Wireless Commun. Lett.*, vol. 9, no. 7, pp. 1089–1092, Jul. 2020.
- [182] Y. Polyanskiy, "A perspective on massive random-access," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2017, pp. 2523–2527.
- [183] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.
- [184] C. Poulliat, M. Fossorier, and D. Declercq, "Design of regular $(2, d_c)$ -LDPC codes over GF(q) using their binary images," *IEEE Trans. Commun.*, vol. 56, no. 10, pp. 1626–1635, Oct. 2008.
- [185] A. K. Pradhan, V. K. Amalladinne, K. R. Narayanan, and J.-F. Chamberland, "Polar coding and random spreading for unsourced multiple access," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6.
- [186] A. K. Pradhan, V. K. Amalladinne, A. Vem, K. R. Narayanan, and J.-F. Chamberland, "Sparse IDMA: A joint graph-based coding scheme for unsourced random access," *IEEE Trans. Commun.*, vol. 70, no. 11, pp. 7124–7133, Nov. 2022.
- [187] R. M. Pyndiah, "Near-optimum decoding of product codes: Block turbo codes," *IEEE Trans. Commun.*, vol. 46, no. 8, pp. 1003–1010, Aug. 1998.
- [188] M. Qiu, X. Wu, J. Yuan, and A. Graell i Amat, "Generalized spatially-coupled parallel concatenated codes with partial repetition," *IEEE Trans. Commun.*, vol. 70, no. 9, pp. 5771–5787, Sep. 2022.
- [189] L. Rapp, S. Miao, and L. Schmalen, "Optimized soft-aided decoding of OFEC and staircase codes," 2024, arXiv:2404.19532.
- [190] L. Rapp and L. Schmalen, "Error-and-erasure decoding of product and staircase codes," *IEEE Trans. Commun.*, vol. 70, no. 1, pp. 32–44, Jan. 2022.
- [191] I. Reed, "A class of multiple-error-correcting codes and the decoding scheme," *Trans. IRE Prof. Group Inf. Theory*, vol. 4, no. 4, pp. 38–49, Sep. 1954.
- [192] G. Reeves and H. D. Pfister, "Reed–Muller codes on BMS channels achieve vanishing bit-error probability for all rates below capacity," *IEEE Trans. Inf. Theory*, vol. 70, no. 2, pp. 920–949, Feb. 2024, doi: 10.1109/TIT.2023.3286452.
- [193] Y. Ren, A. T. Kristensen, Y. Shen, A. Balatsoukas-Stimming, C. Zhang, and A. Burg, "A sequence repetition node-based successive cancellation list decoder for 5G polar codes: Algorithm and implementation," *IEEE Trans. Signal Process.*, vol. 70, pp. 5592–5607, 2022.
- [194] A. Riaz et al., "A sub-0.8 pJ/b 16.3 Gbps/mm² universal soft-detection decoder using ORBGRAND in 40 nm CMOS," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2023, pp. 432–434.
- [195] T. Richardson, "Error floors of LDPC codes," in *Proc. Ann. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, USA, Sep. 2003.
- [196] T. J. Richardson and R. L. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 599–618, Feb. 2001.
- [197] T. J. Richardson and R. L. Urbanke, "Efficient encoding of low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 638–656, Feb. 2001.
- [198] T. Richardson and R. Urbanke, "Multi-edge type LDPC codes," presented at the Workshop Honoring Prof. R. McEliece 60th Birthday, Pasadena, CA, USA, May 2002.
- [199] T. Richardson and S. Kudekar, "Design of low-density parity check codes for 5G new radio," *IEEE Commun. Mag.*, vol. 56, no. 3, pp. 28–34, Mar. 2018.
- [200] T. Richardson and R. Urbanke, *Modern Coding Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2008.
- [201] P. Robertson, "Illuminating the structure of code and decoder of parallel concatenated recursive systematic (turbo) codes," in *Proc. IEEE Global Telecommun. Conf.*, vol. 3, San Francisco, CA, USA, Jun. 1994, pp. 1298–1303.
- [202] E. Sandgren, A. Graell i Amat, and F. Brännström, "On frame asynchronous coded slotting ALOHA: Asymptotic, finite length, and delay analysis," *IEEE Trans. Commun.*, vol. 65, no. 2, pp. 691–704, Feb. 2017.
- [203] G. Sarkis, P. Giard, A. Vardy, C. Thibeault, and W. J. Gross, "Fast list decoders for polar codes," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 2, pp. 318–328, Feb. 2016.
- [204] I. Sason and S. Shamai (Shitz), "Performance analysis of linear codes under maximum-likelihood decoding: A tutorial," *Found. Trends Commun. Inf. Theory*, vol. 3, pp. 1–222, Jul. 2006.
- [205] I. Sason and R. Urbanke, "Parity-check density versus performance of binary linear block codes over memoryless symmetric channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1611–1635, Jul. 2003.
- [206] L. Schmalen, V. Aref, J. Cho, D. Suikat, D. Rösener, and A. Leven, "Spatially coupled soft-decision error correction for future lightwave systems," *J. Lightw. Technol.*, vol. 33, no. 5, pp. 1109–1116, Mar. 12, 2015.
- [207] L. Schmalen, "Probabilistic constellation shaping: Challenges and opportunities for forward error correction," in *Proc. Opt. Fiber Commun. Conf. Expo. (OFC)*, San Diego, CA, USA, Mar. 2018, pp. 1–3.
- [208] L. Schmalen and V. Aref, "Spatially coupled LDPC codes with non-uniform coupling for improved decoding speed," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Visby, Sweden, Aug. 2019, pp. 1–5.
- [209] L. Schmalen, V. Aref, and F. Jardel, "Non-uniformly coupled LDPC codes: Better thresholds, smaller rate-loss, and less complexity," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Aachen, Germany, Jun. 2017, pp. 376–380.
- [210] L. Schmalen, D. Suikat, V. Aref, and D. Rösener, "On the design of capacity-approaching unit-memory spatially coupled LDPC codes for optical communications," in *Proc. 42nd Eur. Conf. Opt. Commun.*, Sep. 2016, pp. 1–3.
- [211] L. Schmalen, D. Suikat, D. Rösener, V. Aref, A. Leven, and S. ten Brink, "Spatially coupled codes and optical fiber communications: An ideal match?" in *Proc. IEEE 16th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2015, pp. 460–464.
- [212] L. Schmalen, D. Suikat, D. Rösener, and A. Leven, "Evaluation of left-terminated spatially coupled LDPC codes for optical communications," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Sep. 2014, pp. 1–3.
- [213] S. Scholl, P. Schläfer, and N. Wehn, "Saturated min-sum decoding: An 'afterburner' for LDPC decoder hardware," in *Proc. Design, Autom. Test Eur. Conf. Exhib. (DATE)*, Mar. 2016, pp. 1219–1224.
- [214] S. Scholl, C. Stumm, and N. Wehn, "Hardware implementations of Gaussian elimination over GF(2) for channel decoding algorithms," in *Proc. Africon*, Sep. 2013, pp. 1–5.
- [215] S. Scholl and N. Wehn, "Advanced hardware architecture for soft decoding Reed–Solomon codes," in *Proc. 8th Int. Symp. Turbo Codes Iterative Inf. Process. (ISTC)*, Aug. 2014, pp. 22–26.
- [216] P. Schulte and G. Böcherer, "Constant composition distribution matching," *IEEE Trans. Inf. Theory*, vol. 62, no. 1, pp. 430–434, Jan. 2016.
- [217] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, Jul. 1948.
- [218] E. Sharon, S. Litsyn, and J. Goldberger, "Efficient serial message-passing schedules for LDPC decoding," *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 4076–4091, Nov. 2007.
- [219] A. Sheikh, A. Graell i Amat, and A. Alvarado, "Novel high-throughput decoding algorithms for product and staircase codes based on error-and-erasure decoding," *J. Lightw. Technol.*, vol. 39, no. 15, pp. 4909–4922, Mar. 7, 2021.
- [220] A. Sheikh, A. Graell i Amat, and G. Liva, "Binary message passing decoding of product codes based on generalized minimum distance decoding," in *Proc. 53rd Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2019, pp. 1–5.
- [221] A. Sheikh, A. Graell i Amat, and G. Liva, "Binary message passing decoding of product-like codes," *IEEE Trans. Commun.*, vol. 67, no. 12, pp. 8167–8178, Dec. 2019.
- [222] A. Sheikh, A. Graell i Amat, G. Liva, and A. Alvarado, "Refined reliability combining for binary message passing decoding of product codes," *J. Lightw. Technol.*, vol. 39, no. 15, pp. 4958–4973, May 18, 2021.
- [223] A. Sheikh, A. Graell i Amat, G. Liva, C. Hager, and H. D. Pfister, "On low-complexity decoding of product codes for high-throughput fiber-optic systems," in *Proc. IEEE 10th Int. Symp. Turbo Codes Iterative Inf. Process. (ISTC)*, Dec. 2018, pp. 1–5.
- [224] Y. Shen, A. Balatsoukas-Stimming, X. You, C. Zhang, and A. P. Burg, "Dynamic SCL decoder with path-flipping for 5G polar codes," *IEEE Wireless Commun. Lett.*, vol. 11, no. 2,

- pp. 391–395, Feb. 2022.
- [225] A. Shokrollahi, "Raptor codes," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2551–2567, Jun. 2006.
- [226] B. P. Smith, A. Farhood, A. Hunt, F. R. Kschischang, and J. Lodge, "Staircase codes: FEC for 100 Gb/s OTN," *J. Lightw. Technol.*, vol. 30, no. 1, pp. 110–117, Jan. 1, 2012.
- [227] A. Solomon, K. R. Duffy, and M. Médard, "Soft maximum likelihood decoding using GRAND," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6.
- [228] D. K. Soma, A. K. Pradhan, and K. Narayanan, "Errors and erasures decoding of product codes for optical transport networks," *IEEE Commun. Lett.*, vol. 25, no. 8, pp. 2482–2486, Aug. 2021.
- [229] D. A. Spielman, "Linear-time encodable and decodable error-correcting codes," *IEEE Trans. Inf. Theory*, vol. 42, no. 6, pp. 1723–1731, Jun. 1996.
- [230] A. Sridharan, M. Lentmaier, D. J. Costello Jr., and K. S. Zigangirov, "Convergence analysis of a class of LDPC convolutional codes for the erasure channel," in *Proc. Ann. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, USA, Sep. 2004.
- [231] C. Stefanović, P. Popovski, and D. Vukobratović, "Frameless ALOHA protocol for wireless networks," *IEEE Commun. Lett.*, vol. 16, no. 12, pp. 2087–2090, Dec. 2012.
- [232] C. Stefanović and P. Popovski, "ALOHA random access that operates as a rateless code," *IEEE Trans. Commun.*, vol. 61, no. 11, pp. 4653–4662, Nov. 2013.
- [233] N. Stolte, "Rekursive Codes mit der Plotkin-Konstruktion und ihre Decodierung," Ph.D. dissertation, Fac. Elect. Eng. Inf. Technol., Technische Universität Darmstadt, Darmstadt, Germany, 2002.
- [234] A. Y. Sukmadji, U. Martínez-Pérez, and F. R. Kschischang, "Zipper codes," *J. Lightw. Technol.*, vol. 40, no. 19, pp. 6397–6407, Jul. 27, 2022.
- [235] J. Sun and O. Y. Takeshita, "Interleavers for turbo codes using permutation polynomials over integer rings," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 101–119, Jan. 2005.
- [236] O. Y. Takeshita and D. J. Costello, "New deterministic interleaver designs for turbo codes," *IEEE Trans. Inf. Theory*, vol. 46, no. 6, pp. 1988–2006, Sep. 2000.
- [237] I. Tal and A. Vardy, "List decoding of polar codes," *IEEE Trans. Inf. Theory*, vol. 61, no. 5, pp. 2213–2226, May 2015.
- [238] R. Tanner, "A recursive approach to low complexity codes," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 5, pp. 533–547, Sep. 1981.
- [239] R. Tanner, "On quasi-cyclic repeat-accumulate codes," in *Proc. Ann. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, USA, Sep. 1999, pp. 249–259.
- [240] M. B. S. Tavares, "On low-density parity-check convolutional codes: Constructions, analysis and VLSI implementation," Ph.D. dissertation, Fac.
- Elect. Eng. Inf. Technol., TU Dresden, Dresden, Germany, 2010.
- [241] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1727–1737, Oct. 2001.
- [242] S. ten Brink, "Convergence of iterative decoding," *Electron. Lett.*, vol. 35, no. 10, pp. 806–808, May 1999.
- [243] S. ten Brink, G. Kramer, and A. Ashikhmin, "Design of low-density parity-check codes for modulation and detection," *IEEE Trans. Commun.*, vol. 52, no. 4, pp. 670–678, Apr. 2004.
- [244] H. B. Thameur, B. L. Gal, N. Khouja, F. Tili, and C. Jego, "Hardware design of Euclidean projection modules for ADMM LDPC decoding," in *Proc. 25th IEEE Int. Conf. Electron., Circuits Syst. (ICECS)*, Dec. 2018, pp. 73–76.
- [245] J. Thorpe, "Low-density parity-check (LDPC) codes constructed from graphs," NASA JPL, Pasadena, CA, USA, IPN Progress Tech. Rep. 42-154, Aug. 2003.
- [246] T. Tonnellier, M. Hashemipour, N. Doan, W. J. Gross, and A. Balatsoukas-Stimming, "Towards practical near-maximum-likelihood decoding of error-correcting codes: An overview," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2021, pp. 8283–8287.
- [247] D. Truhachev, M. Bashir, A. Karami, and E. Nassaji, "Low-complexity coding and spreading for unsourced random access," *IEEE Commun. Lett.*, vol. 25, no. 3, pp. 774–778, Mar. 2021.
- [248] D. Truhachev, D. Mitchell, M. Lentmaier, D. Costello, and A. Karami, "Code design based on connecting spatially coupled graph chains," *IEEE Trans. Inf. Theory*, vol. 65, no. 9, pp. 5604–5617, Sep. 2019.
- [249] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 1, pp. 55–67, Jan. 1982.
- [250] A. Valembois and M. Fossorier, "Box and match techniques applied to soft-decision decoding," *IEEE Trans. Inf. Theory*, vol. 50, no. 5, pp. 796–810, May 2004.
- [251] L. Valentini, M. Chiani, and E. Paolini, "Massive grant-free access with massive MIMO and spatially coupled replicas," *IEEE Trans. Commun.*, vol. 70, no. 11, pp. 7337–7350, Nov. 2022.
- [252] B. Vasic and O. Milenkovic, "Combinatorial constructions of low-density parity-check codes for iterative decoding," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1156–1176, Jun. 2004.
- [253] A. Vem, K. R. Narayanan, J. Chamberland, and J. Cheng, "A user-independent successive interference cancellation based coding scheme for the unsourced random access Gaussian channel," *IEEE Trans. Commun.*, vol. 67, no. 12, pp. 8258–8272, Dec. 2019.
- [254] R. Venkataraman, S. Tatikonda, and A. Barron, "Sparse regression codes," *Found. Trends Commun. Inf. Theory*, vol. 15, nos. 1–2, pp. 1–195, 2019.
- [255] U. Wachsmann, R. F. H. Fischer, and J. B. Huber, "Multilevel codes: Theoretical concepts and practical design rules," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1361–1391, Jul. 1999.
- [256] M. Wasson and S. C. Draper, "Hardware based projection onto the parity polytope and probability simplex," in *Proc. 49th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2015, pp. 1015–1020.
- [257] M. Wasson, M. Milicevic, S. C. Draper, and G. Gulak, "Hardware-based linear program decoding with the alternating direction method of multipliers," *IEEE Trans. Signal Process.*, vol. 67, no. 19, pp. 4976–4991, Oct. 2019.
- [258] N. Wiberg, "Codes and decoding on general graphs," Ph.D. dissertation, Dept. Elect. Eng., Linköping Univ., Linköping, Sweden, 1996.
- [259] N. Wiberg, H. Loeliger, and R. Kotter, "Codes and iterative decoding on general graphs," *Eur. Trans. Telecommun.*, vol. 6, no. 5, pp. 513–525, Sep. 1995.
- [260] A. Wyner and R. Ash, "Analysis of recurrent codes," *IEEE Trans. Inf. Theory*, vol. IT-9, no. 3, pp. 143–156, Jul. 1963.
- [261] J. Xiao, Y. Zhou, S. Song, and Z. Wang, "A low-latency and area-efficient ORBGRAND decoder for polar codes," in *Proc. 4th Inf. Commun. Technol. Conf. (ICTC)*, May 2023, pp. 10–15.
- [262] H. Yang, E. Liang, M. Pan, and R. D. Wesel, "CRC-aided list decoding of convolutional codes in the short blocklength regime," *IEEE Trans. Inf. Theory*, vol. 68, no. 6, pp. 3744–3766, Jun. 2022.
- [263] M. Ye and E. Abbe, "Recursive projection-aggregation decoding of Reed-Muller codes," *IEEE Trans. Inf. Theory*, vol. 66, no. 8, pp. 4948–4965, Aug. 2020.
- [264] A. Yedla, Y. Jian, P. S. Nguyen, and H. D. Pfister, "A simple proof of Maxwell saturation for coupled scalar recursions," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 6943–6965, Nov. 2014.
- [265] Y. Yongrun, P. Zhiwen, L. Nan, and Y. Xiaohu, "Successive cancellation list bit-flip decoder for polar codes," in *Proc. 10th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Oct. 2018, pp. 1–6.
- [266] S. Yoon and Y. Bar-Ness, "A parallel MAP algorithm for low latency turbo decoding," *IEEE Commun. Lett.*, vol. 6, no. 7, pp. 288–290, Jul. 2002.
- [267] H. Zhang and W. Tong, "Channel coding for 6G extreme connectivity-requirements, capabilities and fundamental tradeoffs," *IEEE BITs Inf. Theory Mag.*, vol. 3, no. 1, pp. 54–66, Mar. 2023.
- [268] J. Zhang and M. P. C. Fossorier, "Shuffled iterative decoding," *IEEE Trans. Commun.*, vol. 53, no. 2, pp. 209–213, Feb. 2005.
- [269] Y. Zhang and W. Ryan, "Structured eIRA codes: Performance analysis and construction," *IEEE Trans. Commun.*, vol. 55, no. 5, pp. 837–844, May 2007.
- [270] M. Zhu, M. Jiang, C. Zhao, and L. Hu, "Error-and-erasures decoding of product-like codes with BCH and SPC components," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC Workshops)*, Aug. 2023.

ABOUT THE AUTHORS

Sisi Miao (Student Member, IEEE) received the bachelor's degree in communication engineering from the Ocean University of China, Qingdao, China, in 2018, and the master's degree in INFOTECH from the University of Stuttgart, Stuttgart, Germany, in 2021. She is currently working toward the Ph.D. degree at the Communications Engineering Laboratory, Karlsruhe Institute of Technology, Karlsruhe, Germany, with a focus on classical and quantum error control coding.



Claus Kestel (Graduate Student Member, IEEE) received the Diploma degree in computer engineering from the University of Mannheim, Mannheim, Germany, in 2006. He is currently working toward the Ph.D. degree at the Channel-Coding Group, Chair for Microelectronic Systems Design, University of Kaiserslautern-Landau.

From 2008 to 2013, he was with genua GmbH, Munich, Germany, where he was engaged in the development of network security products as a Software Developer. His current research interests focus on polar codes and their efficient implementations in hardware.



Lucas Johannsen received the B.Eng. degree in mechatronics and the M.Eng. degree in system engineering from Koblenz University of Applied Sciences, Koblenz, Germany, in 2014 and 2016, respectively. He is currently working toward the Ph.D. degree at the Channel-Coding Group, Chair for Microelectronic Systems Design, University of Kaiserslautern-Landau.

During his bachelor's and master's studies, he worked as an Intern at thyssenkrupp Packaging Steel GmbH, Andernach, Germany. Since 2016, he has been a Scientific Assistant at Koblenz University of Applied Sciences. His main research topic is the efficient hardware implementation of polar code decoders.



Marvin Geiselhart (Graduate Student Member, IEEE) received the B.Sc. and M.Sc. degrees (Hons.) in electrical engineering and information technology from the University of Stuttgart, Stuttgart, Germany, in 2017 and 2019, respectively, where he is currently working toward the Ph.D. degree.

His main research topic is channel coding, particularly polar coding and algebraic coding for low-latency applications.



Laurent Schmalen (Fellow, IEEE) received the Dipl.-Ing. degree in electrical engineering and information technology and the Dr.-Ing. degree from RWTH Aachen University, Aachen, Germany, in 2005 and 2011, respectively.

From 2011 to 2019, he was with Alcatel-Lucent Bell Labs and Nokia Bell Labs, Stuttgart, Germany. From 2014 to 2019, he was also a Guest Lecturer with the University of Stuttgart, Stuttgart, Germany. Since 2019, he has been a Professor at Karlsruhe Institute of Technology, Karlsruhe, Germany, where he co-heads the Communications Engineering Laboratory. His research interests include channel coding, modulation formats, and optical communications.

Dr. Schmalen was a recipient and a co-recipient of several awards, including the E-Plus Award for his Ph.D. thesis, the 2013 Best Student Paper Award from the IEEE Signal Processing Systems Workshop, and the 2016 and 2018 *Journal of Lightwave Technology* Best Paper Awards. He is an Associate Editor of IEEE TRANSACTIONS ON COMMUNICATIONS.



Alexios Balatsoukas-Stimming (Member, IEEE) received the Ph.D. degree in computer and communications sciences from the École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, in 2016.

He was a Postdoctoral Researcher at European Laboratory for Particle Physics (CERN) and the Telecommunications Circuits Laboratory, EPFL, from 2017 to 2019. He has been a Visiting Postdoctoral Researcher with Cornell University, Ithaca, NY, USA; and the University of California at Irvine, Irvine, CA, USA. He is currently an Assistant Professor with the Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands. His research interests include very large-scale integration (VLSI) circuits for communications, error correction coding theory and practice, and applications of approximate computing and machine learning to signal processing for communications.



Gianluigi Liva (Senior Member, IEEE) received the M.S. and Ph.D. degrees in electrical engineering from the University of Bologna, Bologna, Italy, in 2002 and 2006, respectively.

Since 2006, he has been with the Institute of Communications and Navigation, German Aerospace Center (DLR), Wessling, Germany, where he leads the Information Transmission Group. His research interests include multiple access and error control coding.



Dr. Liva has been serving as an Associate Editor in coding and information theory for IEEE TRANSACTIONS ON COMMUNICATIONS since 2020.

Norbert Wehn holds the Chair for Microelectronic System Design, Department of Electrical Engineering and Information Technology, University Kaiserslautern-Landau (RPTU). He has more than 450 publications in various fields of microelectronic system design and holds 21 patents. His research interests are very large-scale integration (VLSI) architectures for mobile communication; forward error correction techniques; low-power techniques; advanced system-on-chip (SoC) and memory architectures; and reliability issues in SoC, the Internet of Things (IoT), and hardware accelerators for machine learning.



Stephan ten Brink (Fellow, IEEE) has been the Head of the Institute of Telecommunications, University of Stuttgart, Stuttgart, Germany, since July 2013. Before that, he spent 14 years in industry, with Bell Laboratories (mobile telephony), Holmdel, NJ, USA; and Realtek Semiconductor Corporation (wireless LAN systems), Irvine, CA, USA.

Dr. ten Brink was a recipient and a co-recipient of several awards, including the Vodafone Innovation Award and the IEEE Stephen O. Rice Paper Prize.

