Code Design for Short Blocks: A Survey

Gianluigi Liva, Lorenzo Gaudio, Tudor Ninacs and Thomas Jerkovits

Abstract—The design of block codes for short information blocks (e.g., a thousand or less information bits) is an open research problem which is gaining relevance thanks to emerging applications in wireless communication networks. In this work, we review some of the most recent code constructions targeting the short block regime, and we compare then with both finite-length performance bounds and classical error correction coding schemes. We will see how it is possible to effectively approach the theoretical bounds, with different performance vs. decoding complexity trade-offs.

I. INTRODUCTION

DURING the past sixty years, a formidable effort has been channeled in the research of capacity-approaching error correcting codes [1]. Initially the attention was directed to short and medium-length linear block codes [2] (with some notable exceptions, see e.g. [3], [4]), mainly for complexity reasons. As the idea of code concatenation [5] got established in the coding theorists community [6], the design of long channel codes became a viable solution to approach channel capacity. The effort resulted in a number of practical code constructions allowing reliable transmission at fractions of decibels from the Shannon limit [7]–[16] with low-complexity (sub-optimum) decoding.

The interest in short and medium-block length codes (i.e., codes with dimension \(k\) in the range of 50 to 1000 bits) has been rising again recently, mainly due to emergent applications requiring the transmission of short data units. Examples of such applications are machine-type communications, smart metering networks, remote command links and messaging services (see e.g. [17]–[20]).

When the design of short iteratively-decodable codes is attempted, it turns out that some classical code construction tools which have been developed for turbo-like codes tend to fail in providing codes with acceptable performance. This is the case, for instance, of density evolution [21] and extrinsic information transfer (EXIT) charts [22], which are well-established techniques to design powerful long low-density parity-check (LDPC) and turbo codes. The issue is due to the asymptotic (in the block length) nature of density evolution and EXIT analysis which fail to properly model the iterative decoder in the short block length regime. However, competitive LDPC and turbo code designs for moderate-length and small blocks have been proposed, mostly based on heuristic construction techniques [23]–[44]. While iterative codes retain a large appeal due their low decoding complexity, more sophisticated decoding algorithms [45]–[49] are feasible for short blocks leading to solutions that are performance-wise competitive (if not superior) with respect to iterative decoding of short turbo and LDPC codes.\(^1\)

II. A CASE STUDY

In this section, we provide an exemplary comparison of short codes. We focus on the case study of codes with block length and code dimension \(n = 128\) and \(k = 64\) bits, respectively, which are the parameters of the shortest code recently standardized by Consultative Committee for Space Data Systems (CCSDS) [52] for satellite telecommand links [53]. The performance of the schemes is measures in terms of codeword error rate (CER) versus signal-to-noise ratio (SNR) over the binary-input additive white Gaussian noise (bi-AWGN) channel, with SNR given by the \(E_b/N_0\) ratio (here, \(E_b\) is the energy per information bit and \(N_0\) the single-sided noise power spectral density). Besides, we discuss other metrics such as the capability to detect errors and (although not exhaustively) the complexity of decoding. For this block size, we defined a list of viable candidate solutions comprising

- i. Short binary LDPC and turbo codes, and their non-binary counterparts;
- ii. The \((128, 64)\) extended Bose-Chaudhuri-Hocquenghem (BCH) code (with minimum distance 22), under ordered statistics decoding (OSD);
- iii. Two tail-biting convolutional codes with memory \(m = 8\) and \(m = 11\);
- iv. A polar code under successive cancellation (SC) decoding and under CRC-aided list decoding.

The performance of the codes is compared in Figure 1 with three finite-length performance benchmarks, i.e., the 1959 Shannon’s sphere packing bound (SPB)\(^2\) [57] (—), Gallager’s random coding bound (RCB) [58] for the bi-AWGN channel (-----), and the normal approximation of [59] (------).\(^3\) As reference, the performance of the \((128, 64)\) binary protograph-based [27], [63] LDPC code from the CCSDS telecommand standard [53] is provided too (----). The CCSDS LDPC code performs somehow poorly in terms of coding gain. The code is outperformed at moderate error rates (CER \(\approx 10^{-4}\)) even by a standard regular \((3, 6)\) LDPC code (-----). The CCSDS LDPC code is also outperformed by an accumulate-repeat-3-accumulate (AR3A) LDPC code [64] (----- ---) an by an accumulate-repeat-jagged-accumulate (ARJA) LDPC code (----- ----).

\(^1\)Further approaches deserving a particular attention for short and moderate-length codes are, among others, those in [50], [51].

\(^2\)Additionally to Shannon’s 1959 SPB, one may consider the comparison with bounds relying on error exponents following the 1967 SPB [54]–[56].

\(^3\)Excellent surveys on performance bounds in the finite block length regime are given in [59]–[61]. A useful library of routines for the calculation of the benchmarks is available at https://sites.google.com/site/durisi/software [62].
At low error rates (e.g. CER \( \approx 10^{-6} \)) the CCSDS LDPC code is likely to attain lower error rates than the above-introduced LDPC code competitors thanks to its remarkable distance properties [27]. The four binary LDPC codes introduced so far perform relatively poorly with respect to the benchmarks (roughly 1 dB away from the RCB at CER \( \approx 10^{-4} \)). Despite its uninspiring performance, we shall see in the Section II.B that the CCSDS LDPC code design is particularly suited for application to satellite telecommand links.

The performance of a turbo code introduced in [67] based on 16-states component recursive convolutional codes is also provided (- - -). The turbo code shows superior performance with respect to binary LDPC codes, down to low error rates. The code attains a CER \( \approx 10^{-4} \) at almost 0.4 dB from the RCB. The code performance diverges remarkably from the RCB at lower error rates, due to the relatively low code minimum distance.\(^5\)

Results for both non-binary turbo (- - -) and LDPC (- - -) codes are included in Figure 1. Both codes have been constructed over a finite field of order 256. The turbo code is based on memory-1 time-variant recursive convolutional codes [42]. The choice of memory-1 component codes enables the use of the fast Fourier transform (FFT) to reduce complexity of their forward-backward decoding algorithm [34]. The non-binary LDPC is based on an ultra-sparse parity-check matrix [35]. Details on the code structure are provided in [53], [68], [69]. Both codes attain visible gains with respect to their binary counterparts, performing on the RCB (and with 0.7 dB from the normal approximation reference) down to low error rates (no floors down to CER \( \approx 10^{-9} \) were observed in [68]).

For the block length considered in this comparison, a viable alternative to the use of codes with iterative decoders is provided by OSD. Contrary to iterative decoding, OSD [45] does not require any particular code structure, and hence can be applied to any (linear) block code. In Figure 1, the performance of a (128, 64) extended BCH code with minimum distance 22 is displayed. The variant of OSD used for the simulation is the one based on the identification of the most reliable basis. Test error patterns up to a maximum weight of 4 have been used, resulting in a decoder list of \( \approx 6.8 \times 10^5 \) codewords. The BCH code performance is close to the normal approximation benchmark, gaining \( \approx 0.6 \) dB over non-binary turbo and LDPC codes at CER \( \approx 10^{-4} \). The same decoding algorithm has been applied to the binary image of the non-binary LDPC code (- - -). Interestingly, a non-binary LDPC CER is almost indistinguishable from the one of the BCH code, highlighting the sub-optimality of iterative decoding. The error probability of a polar code under SC decoding (- - -) is included. A more appropriate comparison, able to fully exploit the potential of polar codes under list decoding, is a concatenation of an inner polar code with an outer high-rate error detection code as proposed in [49]. The error probability of the concatenation using a CRC-7 as an outer code is shown (- - -). The polar code has parameters (128, 71), while the outer CRC code has generator polynomial \( g(x) = x^7 + x^3 + 1 \), leading to a code with dimension 64. A list size of 32 has been used in the simulation. The code outperforms all the competitors relying on iterative decoding algorithms, down to a CER \( \approx 10^{-6} \), where the code performance curve intersects the one of the non-binary turbo and LDPC codes. It is anyhow expected that, by changing the design target for the polar code (resulting in a different set of frozen bits), a different trade-off between low and high SNR performance can be achieved.

Finally, the CER of three tail-biting convolutional codes has been included [70]. The first code (- - -) is based on a memory-8 encoder with generator polynomials (in octal form) given by [515 677]. The second code (- - -) is based on a memory-11 encoder with generator polynomials [5537 6131]. The wrap-around Viterbi algorithm (WAVA) algorithm has been used for decoding [71]. The memory-11 convolutional code reaches the performance of the BCH and LDPC codes under OSD. The memory-8 code loses 1 dB at CER \( \approx 10^{-6} \), but still outperforms binary LDPC and turbo codes over the whole simulation range. The third code (- - -) is based on a memory-14 encoder [72] with generator polynomials (in octal form) given by [75063 56711]. The code outperforms all other codes in Figure 1 (at the expense of a high decoding complexity due to the large number of states in the code trellis).

A. The Elephant in the Room: Complexity

In the comparison presented at the beginning of this Section, an important aspect has been (purposely) overlooked: the cost of decoding. The codes that perform close to the SPB rely on relatively complex decoding algorithms. An exhaustive decoding complexity comparison would require a lengthy and rigorous analysis. Moreover, aspects that are not directly measurable in terms of algorithmic complexity (such as, for example, the probability vs. log-likelihood ratio domain form of the decoding algorithms) but still have large impact in hardware implementation can be difficult compared. We provide next only a few qualitative remarks on complexity aspects for the decoding algorithms employed in the simulations.

Remark 1 (Binary vs. non-binary iterative decoding). Binary iterative decoding for LDPC and turbo codes can be efficiently performed in the logarithmic domain, with obvious benefits for finite precision (hardware) implementations. The belief propagation algorithm for the non-binary LDPC and turbo codes presented in this manuscript is performed in the probability domain to allow for FFT-based decoding at the check nodes [73]. Thanks to the FFT, complexity of iterative decoding is proportional to \( q \log_2 q \) (being \( q \) the field order), whereas the conventional iterative decoding complexity would scale with \( q^2 \). From an algorithmic complexity viewpoint, it has been estimated that the FFT-based decoding of the (128, 64) non-binary LDPC code is \( \approx 64 \) times larger than...
the one of (iterative decoding of) the CCSDS LDPC code [69]. Complexity reductions for non-binary LDPC codes can be obtained by applying sub-optimum check node update rules, with various trade-offs between coding gain and decoding complexity [74].

**Remark 2 (OSD vs. non-binary iterative decoding).** For the code parameters adopted in this comparison, OSD and non-binary belief propagation decoding over a finite field of order 256 have similar decoding complexities, as documented in [75]. However, the decoding complexity of OSD scales less favorably with the block length than that of (non-binary) belief propagation (which is linear in the block length for a fixed iteration count and ensemble degree distributions pair). Hence, for larger block lengths OSD may be considered impractical. Efficient OSD variants have been introduced during the past decade, which may extend the range of interest for OSD algorithms (see e.g. [46]).

### B. Error Detection

Some of the algorithms used to decode the codes in Figure 1 are complete, i.e., the decoder output is always a codeword. Incomplete algorithms, such as belief propagation for LDPC codes, may output an *erasure*, i.e., the iterative decoder may converge to a decision that is not a (valid) codeword. Hence, while for complete decoders all error events are *undetected*, incomplete ones provide the additional capability of discarding some decoder outputs when decoding does not succeed. In some applications, it is of paramount importance to deliver very low undetected error rates. This is the case, for instance,
of telecommand systems, where wrong command sequences may be harmful. The CCDS LDPC code has been designed with this objective in mind, trading part of the coding gain for a strong error detection capability [76]. Complete decoders, such as those based on OSD and Viterbi decoding, may be used in such critical applications by adding an error detection mechanism. One possibility would be to include an outer error detection code. Nevertheless, in the short block length regime the introduced overhead might be unacceptable. In this context, a more appealing solution is provided by a post-decoding threshold test as proposed in [77]. Denote by \( \mathbf{y} = (y_1, y_2, \ldots, y_n) \), with \( x + n \), the bi-AWGN channel output for a given transmitted codeword \( x \) (\( n \) is the noise contribution here). We refer to the conditional distribution of \( y \) given \( x \) as \( p(y|x) \). We further denote the maximum likelihood (ML) decoder decision as

\[
x_{\text{ML}} := \arg \max_{x \in \mathcal{C}} p(y|x).
\]

In [77] the metric

\[
\Lambda_{\text{ML}}(y) := \frac{\sum_{x \in \mathcal{C}, x \neq x_{\text{ML}}} p(y|x)}{\sum_{x \in \mathcal{C}} p(y|x)}
\]  

(1)

was proposed and it was proved that the rule for discarding the decoder decision given by the threshold test

\[
\Lambda_{\text{ML}}(y) < \exp(\alpha T)
\]

is optimal in the sense of minimizing the undetected error probability for a given (overall) error probability. The metric (1) is in general complex to compute (with some notable exceptions, see e.g. [78], [79]) due to evaluation of the denominator of (1) (which requires a sum over all possible codewords) and to the need of the ML decision \( x_{\text{ML}} \). In the case of OSD (and of list decoders in general) an approximation of the metric (1) can be easily obtained by summing the conditional distribution \( p(y|x) \) over the codewords present in the list, only. The resulting metric would then be given by

\[
\Lambda_{\text{OSD}}(y) := \frac{\sum_{x \in \mathcal{L}, x \neq x_{\text{OSD}}} p(y|x)}{\sum_{x \in \mathcal{C}} p(y|x)}
\]  

(2)

with

\[
x_{\text{OSD}} := \arg \max_{x \in \mathcal{L}} p(y|x).
\]

being \( \mathcal{L} \) the list produced by the OSD algorithm. While the performance of the test based on the metric (1) has been extensively studied (see e.g. [77], [80]) the authors are not aware of any attempt at analyzing the performance of the metric (2).

III. CONCLUSIONS

An overview of the recent efforts in the design and analysis of efficient error correcting codes for the short block length regime has been provided. A case study tailored to \( (128, 64) \) binary linear block codes has been used to discuss some of the trade-offs between coding gain and decoding complexity for some of the best known code/decoding schemes. The comparison, though incomplete, highlights some promising directions for the design of short and moderate-size block codes.

ACKNOWLEDGMENTS

The authors would like to thank Igal Sason for his valuable comments which helped to improve the review of finite-length performance benchmarks.

REFERENCES

[1] C. Shannon, “A mathematical theory of communication,” Bell System Tech. J., vol. 27, pp. 379–423, 623–656, Jul./Oct. 1948.
[2] E. R. Berlekamp, Key papers in the development of coding theory. IEEE Press, 1974.
[3] P. Elias, “Error-free coding,” Transactions of the IRE, vol. 4, no. 4, pp. 29–37, September 1954.
[4] R. G. Gallager, Low-Density Parity-Check Codes. Cambridge, MA: M.I.T. Press, 1963.
[5] G. D. Forney, Jr., Concatenated Codes. Cambridge, MA: USA: M.I.T. Press, 1966.
[6] D. J. Costello, Jr. and G. D. Forney, Jr., “Channel coding: The road to channel capacity,” Proceedings of the IEEE, vol. 95, no. 6, pp. 1150–1177, June 2007.
[7] C. Berrou, A. Glavieux, and P. Thitimajshima, “Near Shannon limit error-correcting coding and decoding: Turbo-codes,” in Proc. IEEE Int. Conf. Commun. (ICC), Geneva, Switzerland, May 1993.
[8] D. J. C. MacKay, “Good error-correcting codes based on very sparse matrices;” IEEE Trans. Inf. Theory, vol. 45, no. 2, pp. 399–431, Mar 1999.
[9] T. Richardson, M. Shokrollahi, and R. Urbanke, “Design of capacity-approaching irregular low-density parity-check codes,” IEEE Trans. Inf. Theory, vol. 47, no. 2, pp. 619–637, Feb. 2001.
[10] T. Richardson and R. 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.
[11] M. Luby, M. Mitzenmacher, M. A. Shokrollahi, and D. A. Spielman, “Improved Low-Density Parity-Check Codes Using Irregular Graphs,” IEEE Trans. Inf. Theory, vol. 47, no. 2, pp. 585–598, Feb. 2001.
[12] 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.
[13] H. D. Pfister and I. Sason, “Accumulate-repeat-accumulate codes: Capacity-achieving ensembles of systematic codes for the erasure channel with bounded complexity;” IEEE Trans. Inf. Theory, vol. 53, no. 6, pp. 2088–2115, June 2007.
[14] 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, July 2009.
[15] M. Lentmaier, A. Sridharan, D. Costello, Jr., and K. Zigangirov, “Iterative decoding threshold analysis for LDPC convolutional codes;” IEEE Trans. Inf. Theory, vol. 56, no. 10, pp. 5274–5289, Oct. 2010.
[16] S. Kudekar, T. Richardson, and R. 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.
[17] T. de Cola, E. Paolini, G. Liva, and G. Calzolari, “Reliability options for data communications in the future deep-space missions;” Proc. IEEE, vol. 99, no. 11, pp. 2056–2074, Nov. 2011.
[18] F. Boccardi, R. W. Heath, 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.
[19] E. Paolini, C. Stefanovic, G. Liva, and P. Popovski, “Coded random access: applying codes on graphs to design random access protocols;” IEEE Commun. Mag., vol. 53, no. 6, pp. 144–150, June 2015.
[20] G. Durisi, T. Koch, and P. Popovski, “Towards massive, ultra-reliable, and low-latency wireless communications with short packets;” Proc. IEEE, 2016, to appear.
[21] T. Richardson and R. Urbanke, Modern coding theory. Cambridge University Press, 2008.
[22] T. ten Brink, “Convergence behavior of iteratively decoded parallel concatenated codes;” IEEE Trans. Commun., vol. 49, no. 10, pp. 1727–1737, Oct. 2001.
[76] S. Dolinar, K. Andrews, F. Pollara, and D. Divsalar, “Bounded angle iterative decoding of LDPC codes,” in Proc. 2008 IEEE Milcom, Nov. 2008, pp. 1–6.

[77] G. D. Forney, Jr., “Exponential error bounds for erasure, list, and decision feedback schemes,” IEEE Trans. Inf. Theory, vol. 14, no. 2, pp. 206–220, Mar. 1968.

[78] E. Hof, I. Sason, and S. Shamai, “On optimal erasure and list decoding schemes of convolutional codes,” in Proc. Tenth Int. Symp. Commun.

[79] A. R. Williamson, M. J. Marshall, and R. D. Wesel, “Reliability-output decoding of tail-biting convolutional codes,” IEEE Trans. Commun., vol. 62, no. 6, pp. 1768–1778, June 2014.

[80] E. Hof, I. Sason, and S. Shamai, “Performance bounds for erasure, list, and decision feedback schemes with linear block codes,” IEEE Trans. Inf. Theory, vol. 56, no. 8, pp. 3754–3778, Aug 2010.