IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

IMT Atlantique / Département MEE
Laboratoire Lab-STICC / équipe CODES
Technopole Brest Iroise – CS 83818 – 29238 Brest – France

Project: PhD Thesis

Topic: Novel algorithms for ultra-high throughput trellis-based decoding

Domain: Channel coding, digital communications

Duration: 36 months **Start:** from 01.09.2021

Co-directors: Charbel Abdel Nour and Catherine Douillard **Supervisor:** Stefan Weithoffer

Contact: firstname.lastname@imt-atlantique.fr

A. Context

The information society we live in today relies on forward-error-correction (FEC) techniques for point-to-point communications. In the wireless domain, FEC based on Turbo codes enabled the mobile internet and ubiquitous video streaming with third and fourth generation (3G/4G) wireless communication systems. Indeed, for standards like UMTS, which only required a throughput of 1Mb/s in its 1999 release, to LTE, which required 100Mb/s in the 2008 release, Turbo codes with their excellent error correcting performance for the targeted error rates, built-in code rate flexibility and ease of encoding were a perfect match. Turbo codes are also used in current and future releases of LTE-A Pro (4G+), which requires a throughput of several Gb/s and support for almost 200 different frame sizes and a wide range of code rates. But even before iterative decoding with soft-in-soft-out decoding algorithms, trellis based decoding of convolutional codes was successfully employed in a wide range of applications.

However, the widespread use of Turbo and Convolutional codes and the low level of targeted rates led to only a few decoding algorithms being considered for practical use cases. For convolutional codes, the Viterbi algorithm and its variants (Soft-Viterbi, Bi-soft Viterbi, etc) are considered [1, 2, 3]. For Turbo Codes, the BCJR/MAP algorithm and its variants [4, 5, 6] in the logarithmic domain represent the standard algorithms used in software and hardware implementations alike. This textbook knowledge has not been significantly challenged in almost 30 years [7]. At the same time both the Viterbi algorithm and the MAP algorithm suffer from significant limitations through the inherently serial recursive calculations along the code trellis. These impose a lower boundary on the computational complexity of the algorithms and also limit the achievable parallelism for high throughput and low latency implementations in both hardware and software [8, 9, 10]. On another note, the iterative decoding principle introduced with Turbo codes played a significant role in the re-discovery of Low Density Parity Check (LDPC) codes [11]. Their advent coincided with the need for higher throughput. It seemed then natural to resort to LDPC codes for high throughput applications taking into account the ease of achieving high degrees of parallelism for their corresponding decoders. As a consequence to the absence of major novelty for a long period of time and the availability of an alternative solution (LDPC codes) for high throughput applications a common belief was set that Turbo codes no longer hold any mystery for the experts of the field and that there is no room for overcoming their above-listed drawbacks.

Recently, the non-binary (NB) counterparts of Turbo codes over Galois Fields GF(q) have been shown to have considerable potential to outperform the binary case [12]. Moreover, they can be mapped directly to the corresponding modulation symbols of higher order modulation schemes [13]. Unfortunately again, the strong dependency upon the MAP algorithm for decoding has carried over from binary Turbo codes and its complexity in the non-binary domain has been a limiting/prohibitive factor for the design of competitive hardware architectures, in particular for higher Galois field orders.

It is important to note that compared to the other major and recent FEC codes, Turbo codes still offer significant advantages in terms of error correcting performance, especially for short frame sizes and severe channel conditions, and the ease of support for frame size and code rate flexibilities. These advantages play a significant role in improving the spectral efficiency of communication systems constrained by available spectral resources on the one hand and the need for the support for massive access on the other hand.

The stringent requirements of the emerging 5G Ultra-Reliable Low-Latency Communication (URLLC) scenario, for example severely challenge decoding architectures for all major code families. Indeed, one of the outcomes of the recently finished EU project EPIC [14] targeting ultra-high throughput coding, where we were actively involved, clearly states that there is no silver bullet. Today, there is no code family that is able to achieve jointly ultra-high throughput, low latency, reasonable chip area and power consumption,

excellent error correction capabilities while providing the level of support for frame size and code rate flexibilities required for practical applications. Unfortunately, this lack of a leading FEC technology creates a void, exploited during the adoption in standards to base the choice on political/patent and influence related non-technical arguments. An example is the 5G New Radio (NR), where LDPC codes were chosen over Turbo Codes. In addition, Polar codes replaced Convolutional codes for the FEC over the control channel in the 5G NR. However, despite this drawback and for swift deployment purposes, the 3GPP standardization committee submitted its 5G solution to meet the IMT2020 requirements based jointly on LTE evolution and on the first 5G NR specifications [15]. The underlying idea is the upgrade towards LTE Evolution-based 5G solution in the near future while preparing the more costly 5G NR deployment in a few years. Being the FEC solution of LTE evolution, the need for high-throughput Turbo decoders that allow for rate and frame size flexibility becomes a short-term need.

To set Turbo codes, invented in Brest 30 years ago by Claude Berrou [16]**, again on the path of being THE leading technology, their decoding algorithms need to be re-invented. This should be possible helped by the advances in the field related to the other major code families and by nowadays highly improved processing capabilities. This also represents a motivating task from the scientific and the political aspects targeting jointly scientific leadership and French/European sovereignty on major technologies.**

B. Positioning and recent works/results

In recent works, we started to tackle this challenging and disruptive task. Indeed in [17], we introduced a new trellis-based decoding algorithm, the Local-SOVA algorithm. It employs a forward and backward recursion like the conventional max-Log-MAP algorithm, but features a new way of computing the extrinsic information. This new way is based upon a new path-merging operation which leads to a hierarchical structure and allows for a reduction in the computational complexity for high radix orders. Decoding with a higher radix is an essential technique for increasing the decoder throughput, since several consecutive trellis steps are processed in parallel. This leads to a reduction in decoding latency and state metric memory, which in turn, increases the decoder throughput and area efficiency for conventional Turbo decoders (PMAP, XMAP). Therefore, a high radix decoding is highly desirable. However, the increased complexity and long critical path for higher radix orders in the past made radix orders ≥8 inefficient and Turbo decoders typically use radix-2 or radix-4 computations. In [17], a computational complexity reduction of around **30%** w.r.t. addition and comparison operations for Local SOVA is shown for radix orders 4 and 8 on the example of the (8-state) LTE code. Similar ideas applied to the decoding of high rate LTE Turbo codes on their dual-trellis yielded a significant reduction in the number of required lookup tables [18].

In parallel for NB-Turbo Codes, we proposed in [19] significant modifications to the max-Log-MAP (MLM) algorithm to reduce the decoding complexity. Our proposed complexity reductions were inspired by the bubble check algorithm, which had been previously applied to the decoding of NB-LDPC codes. To calculate the recursion metrics and extrinsic information, this algorithm resorts to a subset of the $n_m < q$ highly reliable transitions in the code trellis. These n_m transitions are processed using a 2D-bubble sorter for the add compare select (ACS) operations of the state metric and extrinsic computation. A complexity reduction of up to 50%-75% in terms of ACS operations was achieved at the cost of <0.3 dB in terms of frame error rate (FER) performance when compared to the MLM. Due to the increase in number of states for the component NB- convolutional codes, the application Local SOVA remains to be done and is expected to achieve yet greater levels of complexity reduction compared to the binary case.

These different types of efficient algorithmic simplifications (Local SOVA, bubble check, etc) that we proposed during the EPIC project represent only a short subset of the ideas that were identified with high technical potential and as serious candidates for patent filings. These ideas are planned as future works within this PhD topic. Note also that the listed findings have a wide reach well beyond Turbo codes. In fact, they do directly apply to all techniques that use trellis-based decoding with soft output at the receiver side such as complex detectors (for multiple antenna systems, for interference cancellation, for multi-user detection, etc), some demappers (for continuous phase modulation for example) and the widely used convolutional codes. For these cases, the application of Local-SOVA is expected to reach even more impressive complexity reductions mainly due to the generally larger number of considered trellis states compared to the 8-state trellis of the LTE Turbo code.

Motivated by these large complexity reductions, we recently proposed hardware architectures for the computational units of the Local-SOVA algorithm in [20]. They apply additional complexity reductions to the Local-SOVA algorithm leading to efficient hardware architectures for the Add-Compare-Select-Units (ACSU) and Soft-Output-Units (SOU). These works gave, for the first time, implementation results for the Local-SOVA computational units at a 22nm technology. The results not only confirm the projected complexity reductions from [17] which were based on a mere counting of arithmetic operations, but exceed the complexity reduction projections. Indeed, up to 75% area savings for the SOU and approximately 45% for the overall comparison with the respective MLM units were observed. Building on these initial results on the hardware implementation of new decoding algorithms for Turbo decoding, the ANR funded JCJC project "TurboLEAP", which is set to start 01.03.2021 and led by Stefan Weithoffer, will investigate the design of ultra-high throughput hardware architectures for the decoding of Turbo codes, focusing mainly on hardware design aspects. This project is set to provide the necessary expertise in hardware design to demonstrate the practical feasibility of the proposals.

These previous examples therefore illustrate the potential of revisiting the way how we understand trellis based decoding. There is an immense potential for significant complexity reductions if the decoding algorithm can be more closely tailored to the application and even completely new low complexity decoding algorithms are possible based on the knowledge and insight already recently gained.

C. Topic, environment and outcomes

The PhD candidate is expected to play a significant role in the original approach applied traditionally by the members of the Mathematical and Electrical Engineering (MEE) department of IMT Atlantique consisting of a close interaction between Silicon and Algorithm design. Contrary to classical procedures, in this approach, algorithms are thought and derived from the start to consider at equal footing the hardware implementation and performance related constraints. Indeed, the candidate is expected to explore the full potential of Local SOVA in addition to all the identified high potential alternative simplified decoding algorithms during the EPIC project. Under the supervision of well established researchers in the field of Turbo coding/decoding and hardware implementations, mainly Prof. Catherine Douillard and associate professors Charbel Abdel Nour and Stefan Weithoffer, he will work in close collaboration with the PhD candidate planned to be recruited in the context of the JCJC project on hardware design. Together, both candidates shall achieve the required level of synergy between the simplified design of decoding algorithms and their impact on actual hardware implementations. This synergy is of primary importance for achieving ambitious results that are expected to be demonstrated by the proof-of concept platform planned to be developed within the JCJC project.

D. Financing

36 Months (Internal / ARED)

E. Candidate profile

The candidate should hold a Master degree or an engineering degree in digital communications. The following qualifications are beneficial for the completion of the project:

- Experience with one or more of the following languages: C, C++, Matlab, Python
- Advanced lectures on channel coding and communication systems.

F. References

[1] A. Viterbi, «Error bounds for convolutional codes and an asymptotically optimum decoding algorithm,» *IEEE Transactions on Information Theory,* vol. 13, n° 12, pp. 260-269, 1967.

[2] Q. Huang, Q. Xiao, L. Quan, Z. Wang et S. Wang, «Trimming Soft-Input Soft-Output Viterbi Algorithms,» *IEEE Transactions on Communications,* vol. 64, n° 17, pp. 2952-2960, 2016.

[3] O. Joeressen, M. Vaupel et H. Meyr, «Soft-output Viterbi decoding: VLSI implementation issues,» chez *IEEE 43rd Vehicular Technology Conference*, 1993.

[4] L. Bahl, J. Cocke, F. Jelinek et J. Raviv, «Optimal decoding of linear codes for minimizing symbol error rate (Corresp.),» *IEEE Transactions on Information Theory,* vol. 20, n° 12, pp. 284-287, 1974.

[5] T. Ilnseher, F. Kienle, C. Weis et N. Wehn, «A 2.15GBit/s turbo code decoder for LTE advanced base station applications,» chez *2012 7th International Symposium on Turbo Codes and Iterative Information Processing (ISTC)*, 2012.

[6] S. Weithoffer, F. Pohl et N. Wehn, «On the applicability of trellis compression to Turbo-Code decoder hardware architectures,» chez *2016 9th International Symposium on Turbo Codes and Iterative Information Processing (ISTC)*, 2016.

[7] S. Weithoffer, C. A. Nour, N. Wehn, C. Douillard et C. Berrou, «25 Years of Turbo Codes: From Mb/s to beyond 100 Gb/s,» chez *2018 IEEE 10th International Symposium on Turbo Codes & Iterative Information Processing (ISTC)*, 2018.

[8] B. L. Gal et C. Jego, «Low-latency and high-throughput software turbo decoders on multi-core architectures,» *Annales Des Télécommunications,* vol. 75, n° 11, pp. 27-42, 2020.

[9] C. Roth, S. Belfanti, C. Benkeser et Q. Huang, «Efficient Parallel Turbo-Decoding for High-Throughput Wireless Systems,» *IEEE Transactions on Circuits and Systems,* vol. 61, n° 16, pp. 1824-1835, 2014.

[10] X. Chen, Y. Chen, Y. Li, Y. Huang et X. Zeng, «A 691 Mbps 1.392mm 2 configurable radix-16 turbo decoder ASIC for 3GPP-LTE and WiMAX systems in 65nm CMOS,» chez *2013 IEEE Asian Solid-State Circuits Conference (A-SSCC)*, 2013.

[11] D. J. C. MacKay et R. M. Neal, «Near Shannon limit performance of low density parity check codes,» *Electronics Letters,* vol. 33, n° 16, pp. 457-458, 1996.

[12] R. Klaimi, C. A. Nour, C. Douillard et J. Farah, «Design of Low-Complexity Convolutional Codes over GF(q),» chez *2018 IEEE Global Communications Conference (GLOBECOM)*, 2018.

[13] R. Klaimi, C. A. Nour, C. Douillard et J. Farah, «Union Bound Evaluation for Non-Binary Turbo Coded Modulations,» *IEEE Communications Letters,* vol. 24, n° 16, pp. 1178-1182, 2020.

[14] EPIC, «Enabling practical wireless tb/s communications with next generation channel coding (EPIC),» [En ligne]. Available: https://epic-h2020.eu/. [Accès le 27 01 2021].

[15] A. G. W. Yong, «Draft LTI on 3GPP 5G initial description template,» 2017.

[16] C. Berrou, A. Glavieux et P. Thitimajshima, «Near Shannon limit error-correcting coding and decoding: Turbo-codes. 1,» chez *Proceedings of ICC '93 - IEEE International Conference on Communications*, 1993.

[17] V. H. S. Le, C. A. Nour, E. Boutillon et C. Douillard, «Revisiting the Max-Log-Map Algorithm With SOVA Update Rules: New Simplifications for High-Radix SISO Decoders,» *IEEE Transactions on Communications,* vol. 68, n° %14, pp. 1991-2004, 2020.

[18] V. H. S. Le, C. A. Nour, C. Douillard et E. Boutillon, «A Low-Complexity Dual Trellis Decoding Algorithm for High-Rate Convolutional Codes,» chez *2020 IEEE Wireless Communications and Networking Conference (WCNC)*, 2020.

[19] R. Klaimi, C. A. Nour, C. Douillard et J. Farah, «Low-complexity decoders for non-binary turbo codes,» chez *2018 IEEE 10th International Symposium on Turbo Codes & Iterative Information Processing (ISTC)*, 2018.

[20] S. Weithoffer, R. Klaimi, C. A. Nour, N. Wehn et C. Douillard, «Low-complexity Computational Units for the Local-SOVA Decoding Algorithm,» chez *2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications*, 2020.