3GPP TSG RAN WG1 Meeting #86bis **R1-1608977**
Lisbon, Portugal, 10th – 14th October 2016

Agenda item: 8.1.3.1
Source: AccelerComm
Title: On the maturity of polar decoders, based on a survey of over 150 hardware implementations
Document for: Discussion

I. Introduction

For polar codes, 3GPP RAN1 is considering list decoding with a list size $L \in \{2, 4, 8, 16, 32\}$ and successive cancellation decoding, which can be considered to be a special case of list decoding where $L = 1$. Like Low Density Parity Check (LDPC) codes, the structure of polar decoders scales with the encoded block length N, rather than with the information block length K as in turbo codes. Owing to this, the encoded throughput (measured in Mbps) of a polar (and LDPC) decoder typically remains constant when the coding rate $R = K/N$ is changed. However, since $1/R$ encoded bits must be decoded in order to recover each information bit, the information throughput typically scales down proportionately with the coding rate R, as illustrated in Figure 1. Since the hardware efficiency (measured in Mbps/mm^2 for Application Specific Integrated Circuits (ASICs) or Mbps/kLUT for Field Programmable Gate Arrays (FPGAs)) is given by the ratio of the information throughput to the hardware usage (measured in mm^2 for ASICs or in kLUT for FPGAs), this also scales down proportionately with the coding rate R. Likewise, the energy efficiency (measured in bit/nJ) scales down with coding rate R, since it is given by the ratio of information throughput to the power consumption (measured in mW). Finally, the latency associated with a particular information block length K scales up inversely proportionately with the coding rate R, since it is given by the ratio of the information block length K to the information throughput, as illustrated in Figure 1.

II. Survey of ASIC implementations

A comprehensive survey of 84 polar decoder ASIC implementations from 20 academic publications is provided in [1]. Besides the belief propagation decoders of [2], [3], all of these polar decoders employ list decoding or successive cancellation. The hardware characteristics of these list and successive cancellation decoders are compared in Figure 2. Here, the information throughput, hardware efficiency and energy efficiency have been normalized for the case of a coding rate of $R = 1/2$, as well as for a technology scale of 65 nm.

III. Survey of FPGA implementations

A comprehensive survey of 72 polar decoder FPGA implementations from 11 academic publications is provided in [1]. Besides the SCAN decoder of [8] and the belief propagation decoder of [9], all of these polar decoders employ list decoding or successive cancellation. The hardware characteristics of these list and successive cancellation decoders are compared in Figure 3. Here, the information throughput and hardware efficiency have been normalized for the case of a coding rate of $R = 1/2$, as well as for a technology scale of 40 nm.

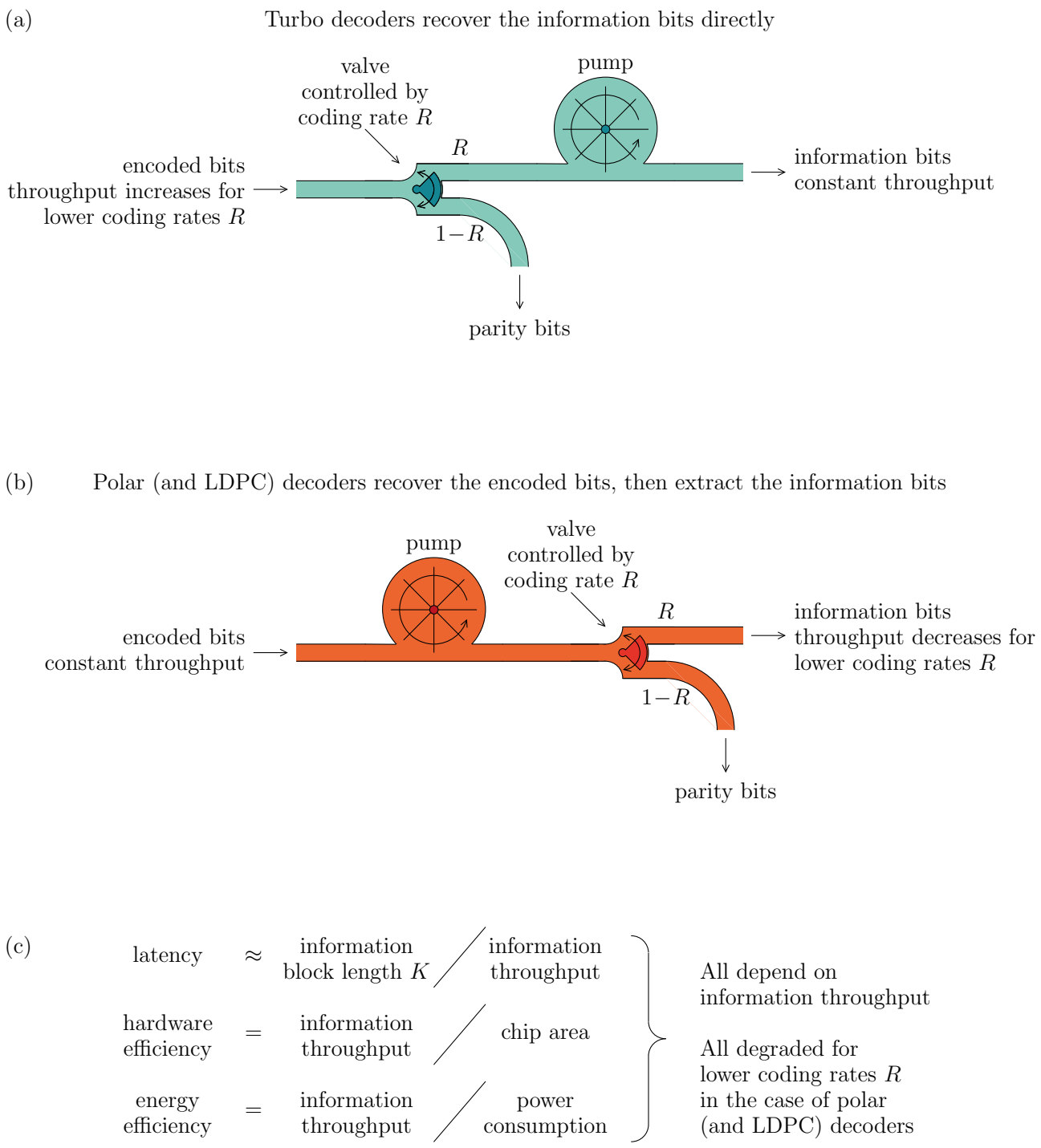
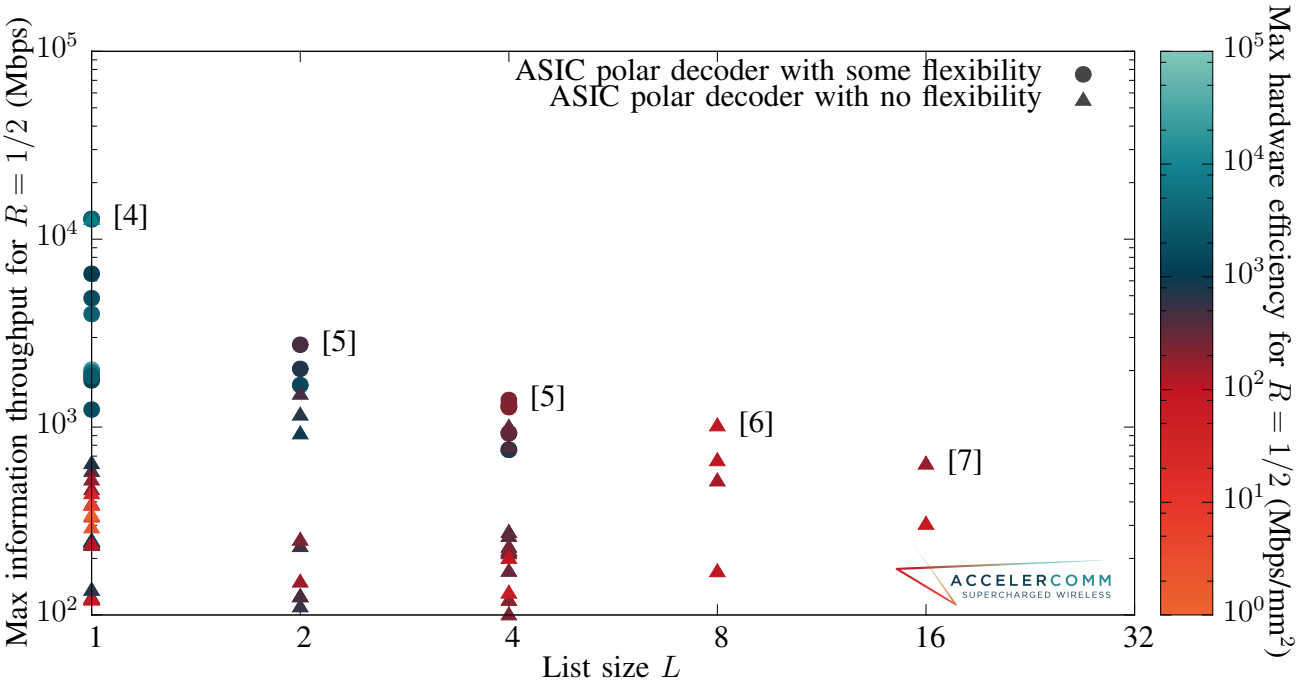
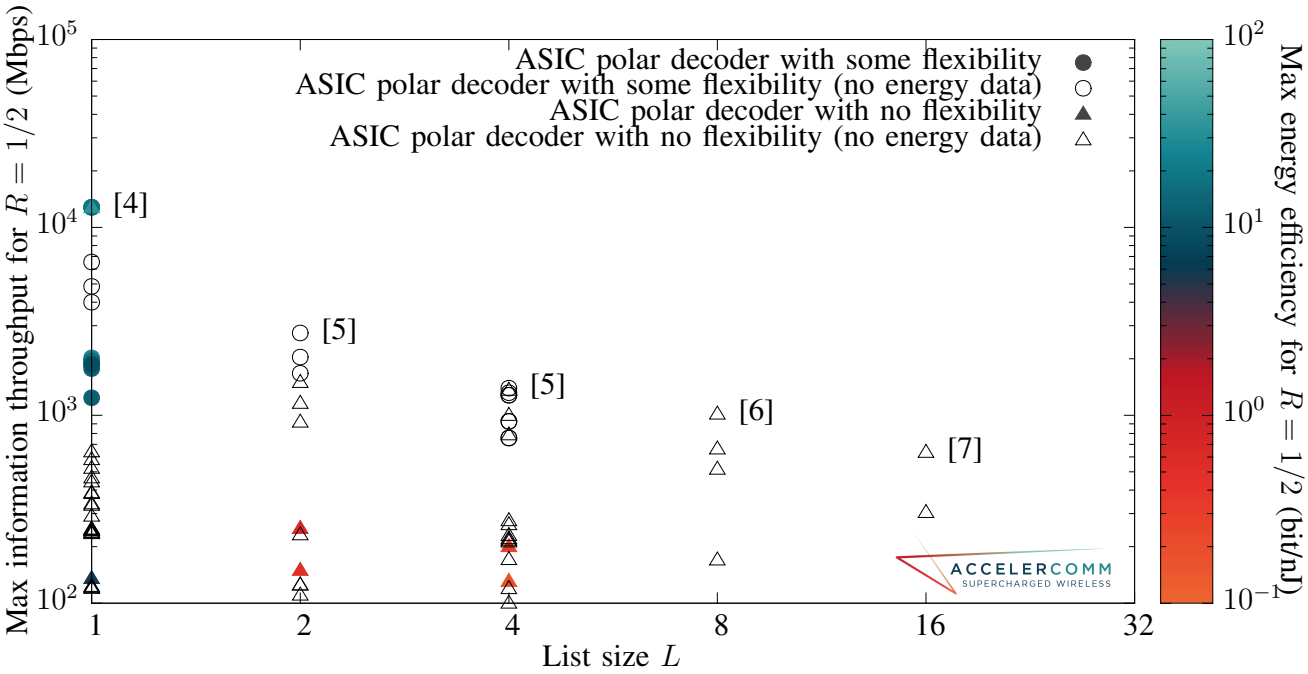


Fig. 1. An analogy using pumps, valves and pipes, to illustrate how the coding rate R of (a) turbo and (b) polar decoders affects their encoded and information throughputs, as well as (c) their latency, hardware efficiency and energy efficiency.

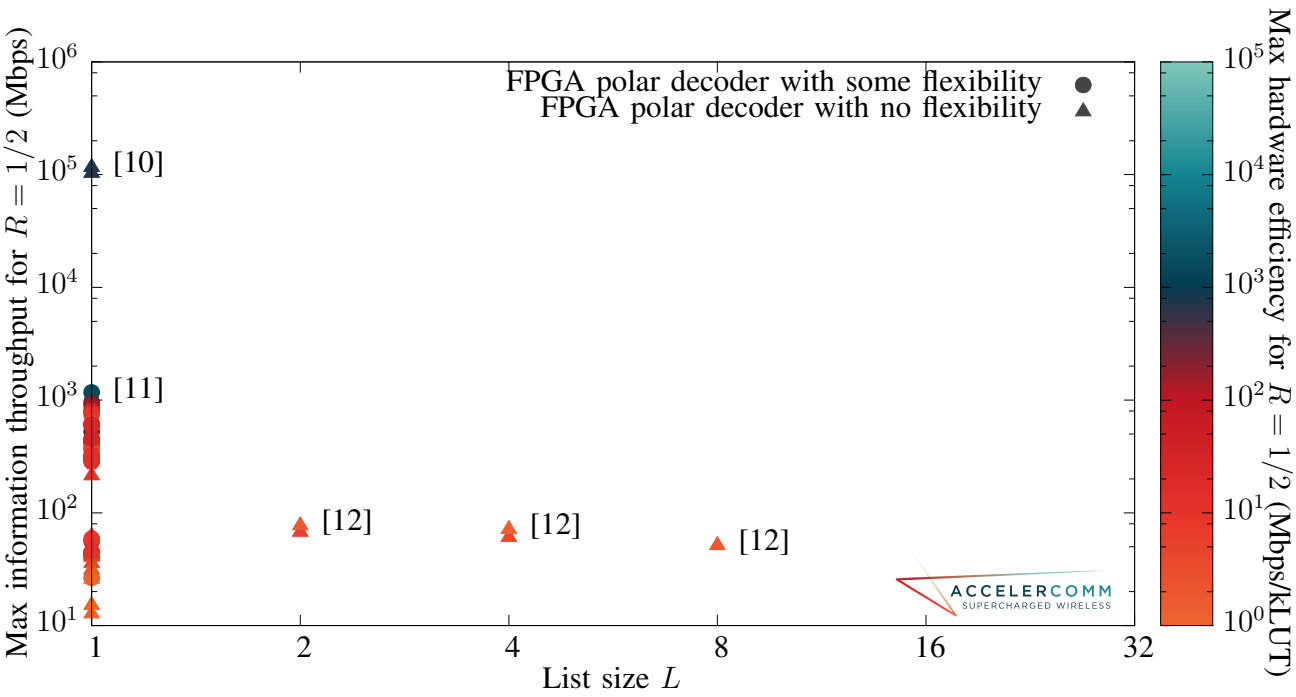


(a)



(b)

Fig. 2. A comparison of list and successive cancellation polar decoder ASIC implementations when scaled to 65 nm and normalized to a coding rate of $R = 1/2$, in terms of list size L, maximum information throughput and (a) maximum hardware efficiency or (b) maximum energy efficiency.



(a)

Fig. 3. A comparison of list and successive cancellation polar decoder FPGA implementations when scaled to 40 nm and normalized to a coding rate of $R = 1/2$, in terms of list size L, maximum information throughput and maximum hardware efficiency.

IV. Maturity of Polar Decoders

As shown in Figure 3, only two papers [4], [10] have demonstrated polar decoders that can achieve information throughputs in excess of the 20 Gbps target for NR. However, the FPGA implementation of [10] achieves information throughputs in excess of 100 Gbps at the cost of being particularly inflexible, since it supports only a single frozen bit pattern, which can be optimized for only a single channel Signal to Noise Ratio (SNR). This design was extended in the follow up paper [4], in order to provide the flexibility to support a small number of different frozen-bit patterns. However, despite achieving only a small amount of flexibility and despite targeting this design for ASIC implementation, the information throughput was significantly reduced to below 20 Gbps for most coding rates R. Furthermore, the designs of [4], [10] employ a list size of $L = 1$, which results in poor Block Error Ratio (BLER) performance compared to turbo and LDPC decoders in wireless channels. Only 11 papers [5]–[7], [12]–[19] have considered list decoders having a list size of $L \in \{2, 4, 8, 16\}$. However, there have been no hardware demonstration of a list decoder having $L = 32$, which is necessary in order to match the BLER performance of turbo and LDPC codes. Furthermore, the highest list size that has been demonstrated for a flexible polar decoder is $L = 4$ in [5]. However, this design supports only coding rate flexibility, since it uses a fixed encoded block length, like nearly all of the flexible polar decoders considered in [1]. Indeed, no polar decoders have been demonstrated that flexibly support a wide range of block lengths K, as well as the full range of coding rates R. Furthermore, in addition to falling short of the BLER performance and flexibility of turbo and LDPC decoders, the polar decoder of [5] achieves worse hardware- and energy-efficiency, as shown in Table I. Here, the polar decoder of [5] is compared with the turbo and LDPC decoders of [20], [21], which are identified as the flexible designs that offer the best throughput, latency, hardware efficiency and energy efficiency in [22].

As shown in Table I, the turbo decoder of [20] offers superior throughput, latency, hardware efficiency and energy efficiency than the polar decoder of [5] at medium and low coding rates R. Furthermore, if the number of supported block lengths and the list size L of this polar decoder were increased such that it offered comparable flexibility and BLER performance as the turbo decoder, then it may be expected

TABLE I
COMPARISON OF THE STATE-OF-THE-ART TURBO, LDPC AND POLAR DECODER ASICS OF [20], [21] AND [5].

Paper	[20]			[21]			[5]		
Year	2014			2013			2016		
Published in	IEEE Trans. Circuits Syst. I			IEEE Trans. Circuits Syst. I			IEEE Trans. VLSI Syst.		
Technology (nm)	90			90			90		
Analysis	Post-layout			Post-layout			Post-synthesis		
Code	Turbo			LDPC			Polar		
Supported standards	LTE			WiMAX, WiFi and G.hn			–		
Flexibility	188 information block lengths and full coding rate flexibility			133 combinations of encoded block length and coding rate			Coding rate flexibility, but fixed encoded block length		
Coding rate R	High 0.95	Medium 0.50	Low 0.33	High 0.83	Medium 0.50	Low –	High 0.90	Medium 0.50	Low 0.33
Information throughput (Mbps)	2274	3028*	3307	857–1957	343–762	–	1662	923**	616**
Latency**** for $K = 1000$ (ns)	440	330*	302	511–1167	1312–2915	–	602	1083**	1625**
Hardware efficiency (Mbps/mm^2)	115	153*	167	154–354	62–138	–	167	93**	62**
Energy efficiency (bit/nJ)	1.57	2.09*	2.28	2.30–5.25***	0.92–2.04***	–	–	–	–

* These characteristics for the medium coding rate have been obtained using linear interpolation between those achieved at the high and low coding rates.
** These characteristics for the medium and low coding rate have been obtained by using the coding rate R to scale those achieved at the high coding rate.
*** The power consumption is stated as 228.36–517.70 mW in [21], but no discussion is provided about how this varies with coding rate. So, the average value of 373.03 mW has been used to calculate these energy efficiencies.
**** Latency is estimated by dividing the information block length $K = 1000$ by the information throughput, since latency is not quantified in [20] or [21]. Note that while none of these decoders support information block lengths of exactly $K = 1000$, these estimates are provided for the sake of illustration.

that the turbo decoder of [20] would offer superior throughput, latency, hardware efficiency and energy efficiency than the polar decoder of [5] at all coding rates R. Indeed, Figures 2 and 3 show that the hardware characteristics of polar decoders degrade linearly as the list size L is increased. The hardware characteristics of the polar decoder of [5] may be expected to degrade further still, if post-layout analysis was employed, rather than only post-synthesis analysis. Indeed, among the 20 academic publications of ASIC implementations considered in [1], all but two of them present only post-synthesis analysis. Furthermore, very few of these publications have quantified the energy efficiency of their polar decoders. This highlights the immaturity of polar decoders.

Observation 1: In contrast to turbo decoders, the information throughput, hardware efficiency and energy efficiency of polar (and LDPC) decoders scale down proportionately with the coding rate, while the latency scales up inversely proportionately.

Observation 2: No polar decoders having the flexibility required for NR have been demonstrated. No polar decoders have been demonstrated that use a list size of above $L = 16$. No polar decoders having any flexibility been demonstrated that use a list size of above $L = 4$.

Observation 3: Turbo decoders offer significantly better BLER performance, flexibility, information throughput, latency, hardware efficiency and energy efficiency than polar decoders, particularly at medium and low coding rates.

Proposal 1: Polar codes should be not be considered further for NR.

REFERENCES

[1] R. G. Maunder, "Survey of ASIC and FPGA implementations of polar decoders," *University of Southampton Dataset*, 2016. [Online]. Available: http://eprints.soton.ac.uk/400401/

[2] B. Yuan and K. K. Parhi, "Early stopping criteria for energy-efficient low-latency belief-propagation polar code decoders," *IEEE Trans. Signal Process.*, vol. 62, no. 24, pp. 6496–6506, Dec 2014.

[3] Y. S. Park, Y. Tao, S. Sun, and Z. Zhang, "A 4.68Gb/s belief propagation polar decoder with bit-splitting register file," in *Proc. Symp. VLSI Circuits*, June 2014.

[4] P. Giard, G. Sarkis, C. Thibeault, and W. J. Gross, "Multi-mode unrolled architectures for polar decoders," *IEEE Trans. Circuits Syst. I*, vol. 63, no. 9, pp. 1443–1453, Sept 2016.

[5] C. Xiong, J. Lin, and Z. Yan, "A multimode area-efficient SCL polar decoder," *IEEE Trans. VLSI Syst.*, 2016.

[6] J. Lin, C. Xiong, and Z. Yan, "A high throughput list decoder architecture for polar codes," *IEEE Trans. VLSI Syst.*, vol. 24, no. 6, pp. 2378–2391, June 2016.

[7] Y. Fan, C. Xia, J. Chen, C. Y. Tsui, J. Jin, H. Shen, and B. Li, "A low-latency list successive-cancellation decoding implementation for polar codes," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 2, pp. 303–317, Feb 2016.

[8] G. Berhault, C. Leroux, C. Jego, and D. Dallet, "Hardware implementation of a soft cancellation decoder for polar codes," in *Proc. Conf. Design Arch. Signal Image Process.*, Sept 2015, pp. 1–8.

[9] A. Pamuk, "An FPGA implementation architecture for decoding of polar codes," in *Proc. Int. Symp. Wireless Commun. Syst.*, Nov 2011, pp. 437–441.

[10] P. Giard, G. Sarkis, C. Thibeault, and W. J. Gross, "237 Gbit/s unrolled hardware polar decoder," *IET Electron. Lett.*, vol. 51, no. 10, pp. 762–763, May 2015.

[11] O. Dizdar and E. Arkan, "A high-throughput energy-efficient implementation of successive cancellation decoder for polar codes using combinational logic," *IEEE Trans. Circuits Syst. I*, vol. 63, no. 3, pp. 436–447, March 2016.

[12] A. Süral and E. Arkan, "An FPGA implementation of successive cancellation list decoding for polar codes," *Bilkent University masters thesis*, February 2016.

[13] A. Balatsoukas-Stimming, A. J. Raymond, W. J. Gross, and A. Burg, "Hardware architecture for list successive cancellation decoding of polar codes," *IEEE Trans. Circuits Syst. II*, vol. 61, no. 8, pp. 609–613, Aug 2014.

[14] 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.

[15] C. Xiong, J. Lin, and Z. Yan, "Symbol-decision successive cancellation list decoder for polar codes," *IEEE Trans. Signal Process.*, vol. 64, no. 3, pp. 675–687, Feb 2016.

[16] Y. Fan, J. Chen, C. Xia, C. y. Tsui, J. Jin, H. Shen, and B. Li, "Low-latency list decoding of polar codes with double thresholding," in *Proc. IEEE Int. Conf. Acoustics, Speech Signal Process.*, April 2015, pp. 1042–1046.

[17] J. Lin, C. Xiong, and Z. Yan, "A reduced latency list decoding algorithm for polar codes," in *Proc. IEEE Workshop Signal Process. Syst.*, Oct 2014, pp. 1–6.

[18] J. Lin and Z. Yan, "An efficient list decoder architecture for polar codes," *IEEE Trans. VLSI Syst.*, vol. 23, no. 11, pp. 2508–2518, Nov 2015.

[19] J. Lin, C. Xiong, and Z. Yan, "A high throughput list decoder architecture for polar codes," *IEEE Trans. VLSI Syst.*, vol. 24, no. 6, pp. 2378–2391, June 2016.

[20] R. Shrestha and R. P. Paily, "High-throughput turbo decoder with parallel architecture for LTE wireless communication standards," *IEEE Trans. Circuits Syst. I*, vol. 61, no. 9, pp. 2699–2710, Sept 2014.

[21] Y. L. Ueng, B. J. Yang, C. J. Yang, H. C. Lee, and J. D. Yang, "An efficient multi-standard LDPC decoder design using hardware-friendly shuffled decoding," *IEEE Trans. Circuits Syst. I*, vol. 60, no. 3, pp. 743–756, March 2013.

[22] AccelerComm, "R1-1608584 Complementary turbo and LDPC codes for NR, motivated by a survey of over 100 ASICs," in *3GPP TSG RAN WG1 #86bis*, Sept. 2016.