

Folded Successive Cancellation Decoding of Polar Codes

Sinan Kahraman^{*†}, Emanuele Viterbo^{*} and Mehmet E. Çelebi[†]

^{*} Electrical and Computer Systems Eng., Monash University, Melbourne, VIC 3800, Australia

[†] Electronics and Communication Eng., Istanbul Technical University, Istanbul, 34469, Turkey

Email: kahraman@ieee.org, emanuele.viterbo@monash.edu, mecelebi@itu.edu.tr

Abstract—Polar codes are the first explicit class of codes that are provably capacity-achieving under the successive cancellation (SC) decoding. As a suboptimal decoder, SC has quasi-linear complexity $N(1 + \log N)$ in the code length N . In this paper, we propose a new non-binary SC decoder with reduced complexity $\frac{N}{2}(1 + \log \frac{N}{2})$ based on the folding operation, which was first proposed in [11] to implement folded tree maximum-likelihood decoding of polar codes. Simulation results for the additive white Gaussian noise channel show that folded SC decoders can achieve the same error performance of standard SC by suitable selecting the folding of the polar code.

I. INTRODUCTION

The existence of capacity-achieving codes was proved by Shannon's noisy channel coding theorem [1]. Recently, channel polarization was introduced by Arıkan [2], and polar codes were shown to be the first provable class of capacity-achieving codes under the low complexity, suboptimal successive cancellation (SC) decoding method. SC is known as the conventional decoder for polar codes and has quasi-linear complexity $\mathcal{O}(N \log N)$ in the code length, N .

Close to capacity performance of polar codes can be obtained with rather long codes, where the SC decoding implementation complexity can become an issue. Research has developed along two lines: (i) reducing SC complexity implementation (ii) improving the error performance with higher complexity decoders. Along the first line, efficient SC decoder implementations have been studied in [3],[4],[5],[6], based on specific methods to optimize the SC decoder in hardware. In [5], a specific scheduling for butterfly-based SC decoder was presented enabling resource sharing to reduce complexity. A semi-parallel decoder was proposed in [6] as a simple architecture for resource sharing with a small increase in latency. Along the second line, to improve error performance, higher complexity decoders have been investigated such as list decoder in [7], belief propagation (BP) in [9],[8] and the sphere decoding based maximum likelihood (ML) decoders in [10],[11].

We propose a non-binary *folded SC decoder* to reduce the decoding complexity of polar coding. The proposed method is based on the folded decoding tree structure, introduced in [11] to implement an efficient ML decoder of polar codes.

Briefly, polar codes are constructed from the n -fold Kronecker product $\mathbf{F}^{\otimes n}$ of a 2×2 kernel matrix \mathbf{F} . In [11], it was observed that $\mathbf{F}^{\otimes n}$ has the form of a well-known fractal in chaotic phenomena, the Sierpinski triangle, introduced by

W. Sierpinski in 1915, [12]. The triangular structure of $\mathbf{F}^{\otimes n}$ defines a binary decoding tree for polar codes. The folding operation transforms such decoding tree of height N into a nonbinary tree with height $N/2$.

In [11], it was also shown that the fractal structure of $\mathbf{F}^{\otimes n}$ results in N alternative foldings yielding different nonbinary tree structures.

In this paper, we show that, by applying once the folding operation, a conventional SC decoder can be re-designed as a nonbinary half-length SC decoder. In the following we will refer to this as *folded-SC decoder*. Instead of the likelihood ratios used in the $1 + \log N$ steps of the conventional SC decoder, we consider conditional probabilities of bit-pairs (defined by the folding operation) in $1 + \log \frac{N}{2}$ steps of the proposed folded-SC decoder.

The folding operation enables us to construct $\log N$ alternative pairings of bits, which results in folded SC decoders with the same complexity but different error performances for a given code.

Simulation results show that a small subset of the $\log N$ alternative folded-SC decoders provide the same bit error rate (BER) and frame error rate (FER) of the conventional SC decoder.

The organization of this paper is given as follows. The folded-SC decoder is described in Section II. In Section III, we provide simulation results to investigate the performance of the proposed method. In Section IV, we analyze complexity of the proposed method. In Section V, we give conclusions and address some open problems for future work.

II. FOLDED SUCCESSIVE CANCELATION DECODER

A. Polar codes with conventional SC decoder

A polar code $\mathcal{C} : (N, K, \mathcal{F})$ of length N and dimension K is defined by the set of frozen bits \mathcal{F} . We can generate the codewords $\mathbf{x} = (x_0, \dots, x_{N-1})^T$ of \mathcal{C} as

$$\mathbf{x} = \mathbf{G}_n \mathbf{d} \quad (1)$$

where the vector $\mathbf{d} = (d_0, \dots, d_{N-1})^T$ has $N - K$ bits in positions \mathcal{F} frozen to '0' and K bits in positions $\bar{\mathcal{F}}$ used for the information bits. We recall that the triangular matrix $\mathbf{G}_n = \mathbf{F}^{\otimes n}$ is the n -fold iterated Kronecker product of the kernel matrix $\mathbf{F} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

The coded bits in \mathbf{x} are modulated using a BPSK modulation (i.e., '1' \rightarrow +1, '0' \rightarrow -1) into a signal vector $\tilde{\mathbf{x}}$ which is transmitted over the AWGN channel

$$\tilde{\mathbf{y}} = \tilde{\mathbf{x}} + \mathbf{z}, \quad (2)$$

where \mathbf{z} is the additive white Gaussian noise with zero mean and variance σ^2 .

We recall the pseudo code of the conventional successive cancellation decoding algorithm in Table I, [2], [7]. Let $n = \log N$ and $\hat{\mathbf{d}}$ be the estimated bits. The SC decoder estimates bits in the order $\alpha(0), \alpha(1), \dots, \alpha(N-1)$, depending on the SC architecture [2]. Let $\hat{\mathbf{d}}_* = (\dots, \hat{d}_{\alpha(0)}, \dots, \hat{d}_{\alpha(j-1)}, \dots)$ be the vector with partial decisions after the first j steps (the remaining entries are yet to be determined). We assume that $\tilde{\mathbf{y}}$ is the received noisy vector. Then we define the conditional probabilities as $W_{\alpha(i)}^j(\tilde{\mathbf{y}}, \hat{\mathbf{d}}_*|0)$ and $W_{\alpha(i)}^j(\tilde{\mathbf{y}}, \hat{\mathbf{d}}_*|1)$ for the j^{th} step. For the simplicity of notation, we will use $W_{\alpha(i)}^j(\cdot|0)$ and $W_{\alpha(i)}^j(\cdot|1)$, instead.

TABLE I

Algorithm: Conventional successive cancellation decoding

```

1: for all  $i = 0, 1, \dots, N-1$  do
2:   calculate successively  $W_{\alpha(i)}^n(\cdot|0), W_{\alpha(i)}^n(\cdot|1)$ 
3:   if  $d_{\alpha(i)}$  is frozen then
4:     set  $d_{\alpha(i)} = 0$ 
5:   else
6:     if  $W_{\alpha(i)}^n(\cdot|0) < W_{\alpha(i)}^n(\cdot|1)$  then
7:       set  $d_{\alpha(i)} = 1$ 
8:     else
9:       set  $d_{\alpha(i)} = 0$ 
10:    end if
11:  end if
12:  broadcast current decisions to other levels
13: end for
    
```

To describe the computations of W , the unit circuit is shown in Fig.1. Successive computations of the conditional probabilities in the unit circuit from step j^{th} to step $(j+1)^{\text{th}}$ can be given as

$$W_1^{j+1}(\cdot|0) = W_1^j(\cdot|0) \cdot W_2^j(\cdot|0) + W_1^j(\cdot|1) \cdot W_2^j(\cdot|1)$$

$$W_1^{j+1}(\cdot|1) = W_1^j(\cdot|1) \cdot W_2^j(\cdot|0) + W_1^j(\cdot|0) \cdot W_2^j(\cdot|1)$$

$$W_2^{j+1}(\cdot|0) = \begin{cases} W_1^j(\cdot|0) \cdot W_2^j(\cdot|0) & \text{if } \hat{d} = 0 \\ W_1^j(\cdot|1) \cdot W_2^j(\cdot|0) & \text{if } \hat{d} = 1 \end{cases}$$

$$W_2^{j+1}(\cdot|1) = \begin{cases} W_1^j(\cdot|1) \cdot W_2^j(\cdot|1) & \text{if } \hat{d} = 0 \\ W_1^j(\cdot|0) \cdot W_2^j(\cdot|1) & \text{if } \hat{d} = 1 \end{cases}$$

Alternatively, the likelihood ratios $L^j(\cdot) = W^j(\cdot|0)/W^j(\cdot|1)$ can be updated by the following expressions:

$$L_1^{j+1}(\cdot) = \frac{L_1^j(\cdot) \cdot L_2^j(\cdot) + 1}{L_1^j(\cdot) + L_2^j(\cdot)}$$

$$L_2^{j+1}(\cdot) = \begin{cases} L_1^j(\cdot) L_2^j(\cdot) & \text{if } \hat{d} = 0 \\ L_2^j(\cdot) / L_1^j(\cdot) & \text{if } \hat{d} = 1 \end{cases}$$

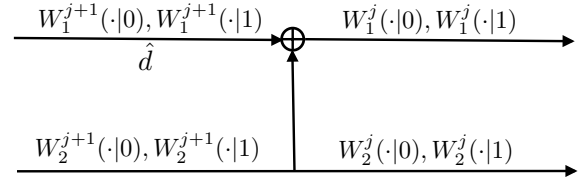


Fig. 1. Unit circuit with conditional probabilities.

B. Description of Folding Operation on SC Decoder

Polar codes are defined by n -fold Kronecker product of a kernel matrix \mathbf{F} as given by $\mathbf{F}^{\otimes n}$. The fractal structure on $\mathbf{F}^{\otimes n}$ was first observed in [11]. The structure has self similarities that are repeated in different scales for any polarization step. This geometry enables the folding operation to construct a nonbinary tree structure equivalent to the binary tree structure of the polar code. Furthermore, it is possible to obtain $\log N$ alternative foldings which result in different nonbinary trees [11].

The folding operation enables us to write the encoding equation $\mathbf{x} = \mathbf{F}^{\otimes n} \mathbf{d}$ as

$$\begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N-1} \end{bmatrix} = \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N/2-1} \\ x_{N/2} \\ x_{N/2+1} \\ \vdots \\ x_{N-1} \end{bmatrix} = \mathbf{F}^{\otimes(n-1)} \begin{bmatrix} d_0 \oplus d_{N/2} \\ d_1 \oplus d_{N/2+1} \\ \vdots \\ d_{N/2-1} \oplus d_{N-1} \\ d_{N/2} \\ d_{N/2+1} \\ \vdots \\ d_{N-1} \end{bmatrix} \quad (3)$$

In general, we can write the set of pairs of bit indices which are added in the first half of (3) as

$$\mathcal{I} = \{\mathcal{I}_\ell\} = \left\{ \left(\frac{N}{2} - \ell, N - \ell \right), \ell = 1, \dots, N/2 \right\} \quad (4)$$

Then, we can construct a nonbinary tree of half height, where the branches are labeled by the bit pairs \mathcal{I}_ℓ at each level for $\ell = 1, \dots, N/2$.

It can be shown that different foldings can be described in terms of permutations of the input bits. More detailed descriptions about folding operation can be found in [11].

Here, we can consider the folding operation for SC decoding. The folded-SC decoding algorithm can be constructed as a non-binary version of SC decoding, using conditional probabilities of bit pairs, i.e., $W^j(\cdot|\varphi)$ for the bit pairs $\varphi = \{00, 01, 10, 11\}$. The folded-SC decoder has only $\frac{n}{2} + 1$ steps. The folding operation defines $\varphi_1 = d_{\alpha(i)} \oplus d_{\alpha(i+\frac{n}{2})}$ and $\varphi_2 = d_{\alpha(i+\frac{n}{2})}$.

Successive computations of the conditional probabilities in the folded unit circuit in Fig. 2 from step j to step $j+1$ are given as:

$$\bar{W}_1^{j+1}(\cdot|00) = W_1^j(\cdot|00) \cdot W_2^j(\cdot|00) + W_1^j(\cdot|01) \cdot W_2^j(\cdot|01) + W_1^j(\cdot|10) \cdot W_2^j(\cdot|10) + W_1^j(\cdot|11) \cdot W_2^j(\cdot|11)$$

$$\bar{W}_1^{j+1}(\cdot|01) = W_1^j(\cdot|00) \cdot W_2^j(\cdot|01) + W_1^j(\cdot|01) \cdot W_2^j(\cdot|00) + W_1^j(\cdot|10) \cdot W_2^j(\cdot|11) + W_1^j(\cdot|11) \cdot W_2^j(\cdot|10)$$

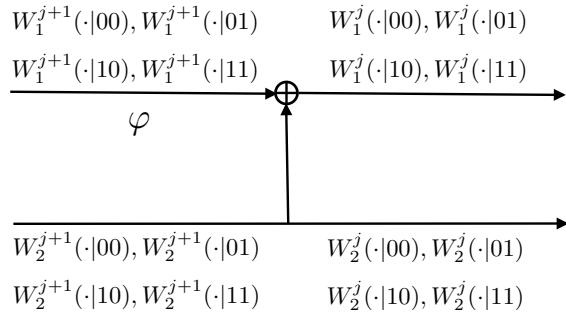


Fig. 2. Folded unit circuit with conditional probabilities.

$$\bar{W}_1^{j+1}(\cdot|10) = W_1^j(\cdot|00) \cdot W_2^j(\cdot|10) + W_1^j(\cdot|01) \cdot W_2^j(\cdot|11) + W_1^j(\cdot|10) \cdot W_2^j(\cdot|00) + W_1^j(\cdot|11) \cdot W_2^j(\cdot|01)$$

$$\bar{W}_1^{j+1}(\cdot|11) = W_1^j(\cdot|00) \cdot W_2^j(\cdot|11) + W_1^j(\cdot|01) \cdot W_2^j(\cdot|10) + W_1^j(\cdot|10) \cdot W_2^j(\cdot|01) + W_1^j(\cdot|11) \cdot W_2^j(\cdot|00)$$

$$\bar{W}_2^{j+1}(\cdot|00) = \begin{cases} W_1^j(\cdot|00) \cdot W_2^j(\cdot|00) & \text{if } \varphi = 00, \\ W_1^j(\cdot|01) \cdot W_2^j(\cdot|00) & \text{if } \varphi = 01, \\ W_1^j(\cdot|10) \cdot W_2^j(\cdot|00) & \text{if } \varphi = 10, \\ W_1^j(\cdot|11) \cdot W_2^j(\cdot|00) & \text{if } \varphi = 11. \end{cases}$$

$$\bar{W}_2^{j+1}(\cdot|01) = \begin{cases} W_1^j(\cdot|01) \cdot W_2^j(\cdot|01) & \text{if } \varphi = 00, \\ W_1^j(\cdot|00) \cdot W_2^j(\cdot|01) & \text{if } \varphi = 01, \\ W_1^j(\cdot|11) \cdot W_2^j(\cdot|01) & \text{if } \varphi = 10, \\ W_1^j(\cdot|10) \cdot W_2^j(\cdot|01) & \text{if } \varphi = 11. \end{cases}$$

$$\bar{W}_2^{j+1}(\cdot|10) = \begin{cases} W_1^j(\cdot|10) \cdot W_2^j(\cdot|10) & \text{if } \varphi = 00, \\ W_1^j(\cdot|11) \cdot W_2^j(\cdot|10) & \text{if } \varphi = 01, \\ W_1^j(\cdot|00) \cdot W_2^j(\cdot|10) & \text{if } \varphi = 10, \\ W_1^j(\cdot|01) \cdot W_2^j(\cdot|10) & \text{if } \varphi = 11. \end{cases}$$

$$\bar{W}_2^{j+1}(\cdot|11) = \begin{cases} W_1^j(\cdot|11) \cdot W_2^j(\cdot|11) & \text{if } \varphi = 00, \\ W_1^j(\cdot|10) \cdot W_2^j(\cdot|11) & \text{if } \varphi = 01, \\ W_1^j(\cdot|01) \cdot W_2^j(\cdot|11) & \text{if } \varphi = 10, \\ W_1^j(\cdot|00) \cdot W_2^j(\cdot|11) & \text{if } \varphi = 11. \end{cases}$$

Note that some of the conditional probabilities will be set to zero when a decision is made at a previous stage (broadcast phase). The corresponding a posteriori probabilities (APP) on a bit pair $(b_1 b_2)$ can then be computed by normalization as:

$$W_1^{j+1}(\cdot|b_1 b_2) = \frac{\bar{W}_1^{j+1}(\cdot|b_1 b_2)}{\sum_{q_1 q_2} \bar{W}_1^{j+1}(\cdot|q_1 q_2)}$$

$$W_2^{j+1}(\cdot|b_1 b_2) = \frac{\bar{W}_2^{j+1}(\cdot|b_1 b_2)}{\sum_{q_1 q_2} \bar{W}_2^{j+1}(\cdot|q_1 q_2)}, \text{ if } \varphi = \varphi_1 \varphi_2$$

Differently from the binary case log-likelihood ratios can not be used and four dimensional vectors of APP need to be stored. In the next section, we investigate the error performance of the folded-SC decoder in AWGN channels.

III. SIMULATION RESULTS

We consider polar codes for the block lengths 128 and 256. Conventional SC and the proposed folded-SC decoders are applied for P(128,120), P(256,247), P(128,64) and P(256,128). High rate codes, P(128,120), P(256,247), were constructed using the standard approach in [9] for $E_b/N_0 = 1$ dB, and the half rate codes, P(128,64) and P(256,128), for $E_b/N_0 = 0$ dB. We compare the performance of all the n possible folding operations for the folded-SC decoder ($n = 7$ for 128 and $n = 8$ for 256). Simulation results of the bit error rates are given in Fig.3-4 for P(128,120) and P(256,247) under SC decoding and the alternative folded-SC decoders. The alternative folding operations result in different pairings of bits: the frozen bits may end up in a pair or with a non frozen bit, and will result in different decoding performance. There are three groups of in alternative folded-SC decoders with the same error performance. These groups for P(128,120) can be given as G_1 :{FSC-F,FSC-G}, G_2 :{FSC-B,FSC-C,FSC-D,FSC-E}, G_3 :{FSC-A}. In the case of P(256,247), groups are G_1 :{FSC-F,FSC-G,FSC-H}, G_2 :{FSC-C,FSC-D,FSC-E}, G_3 :{FSC-A,FSC-B}. The curves for G_1, G_2 and G_3 are dashed, dot-dashed, and dotted, respectively.

It is interesting to notice that G_1 for the case of P(128,120) and P(256,247) yields approximately the same performance of the conventional SC decoder. G_2 is close to the G_1 in terms of the error performances. G_3 performs the worst performance for the case of P(128,120) and P(256,247). Frame error rates of SC and folded-SC decoders are given in Fig.5-6 for P(128,120) and P(256,247).

Bit and frame error rates of P(128,64) and P(256,128) are given in Fig.7-8 and Fig.9-10. There are five performance groups for P(128,64) and six groups for P(256,128) for the alternative folded-SC decoders with the same error performances. The group with the best performance is G_1 :{FSC-F,FSC-G} for P(128,64) and G_1 :{FSC-F,FSC-G,FSC-H} for P(256,128).

Simulation results show that alternative foldings have different error performance for a given polar code. The main reason for such performance difference is related to the positions of the frozen bits. Let us define a *constrained bit pair* as a folded bit pair with one frozen bit and one information bit. Constrained bit pairs degrade the performance of the folded-SC decoding since the frozen information is not fully propagated in the SC decoding algorithm. We notice that performance groups of alternative folded-SC decoders have the same number of constrained bit pairs. In future work we will study if it is possible to construct good polar-like codes for the folded SC decoders by limiting the number of constrained bit pairs.

IV. COMPLEXITY

In this section, we discuss the complexity and latency of the folded-SC decoder. It will be shown that folded SC decoder has significant advantages in these respects. We can show that folded-SC decoder has a lower complexity compared to SC decoder. The estimated complexity of conventional

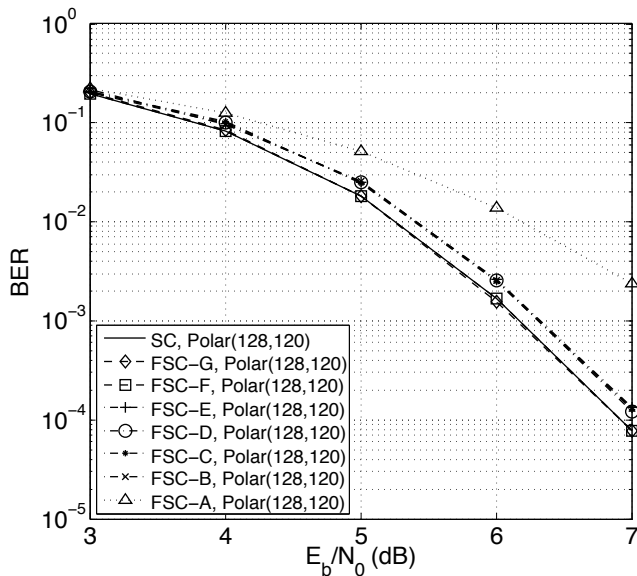


Fig. 3. Bit error rate performances of SC and folded SC for Polar code (128,120).

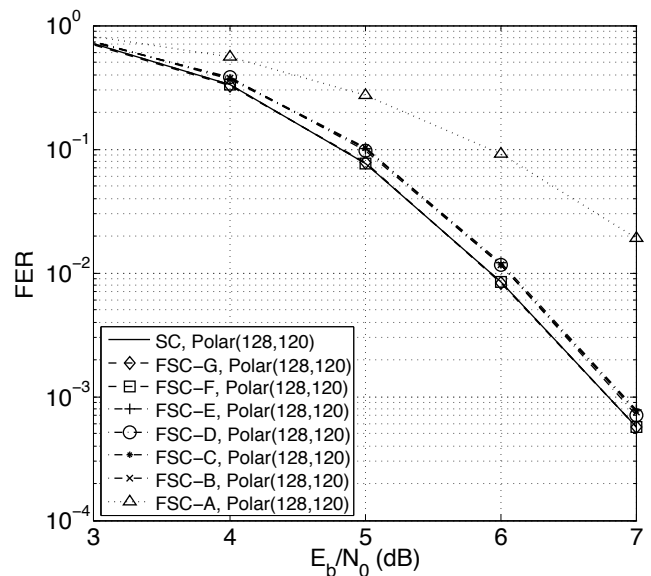


Fig. 5. Frame error rate performances of SC and folded SC for Polar code (128,120).

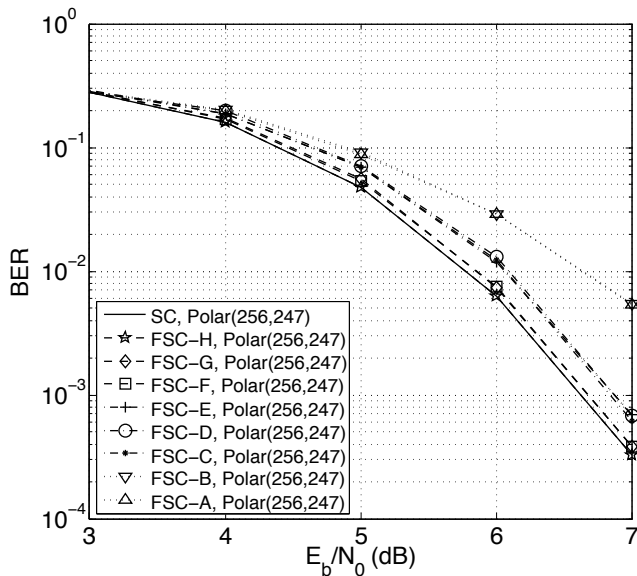


Fig. 4. Bit error rate performances of SC and folded SC for Polar code (256,247).

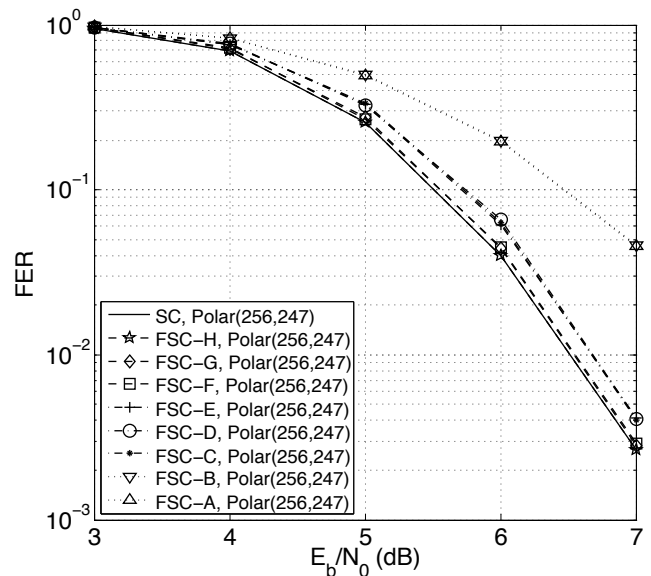


Fig. 6. Frame error rate performances of SC and folded SC for Polar code (256,247).

SC decoding under single processor computational model was defined by [2] as the *time complexity*. Conventional SC decoding has $N(1 + \log N)$ time complexity and $2N - 1$ time complexity if the best-possible parallelization is used. Folded SC decoding can be seen as non-binary of SC decoding with half the number of steps. Hence, the the time complexity is $N/2(1 + \log N/2)$ under the single processor computational model and $N - 1$ under best-possible parallelization. It can be seen that folded SC decoding has a significantly lower complexity than the conventional SC decoder.

ACKNOWLEDGMENT

Sinan Kahraman was supported by The Scientific and Technological Research Council of Turkey (TUBITAK) under grant 1059B141200235. This work was performed during his visit

to the Software Defined Telecommunications (SDT) Lab. at Monash University, Australia.

Emanuele Viterbo's contribution to this paper was made possible by NPRP grant # NPRP5-597-2-241 from the Qatar National Research Fund (a member of Qatar Foundation).

V. CONCLUSIONS

A nonbinary SC decoder was proposed for folded polar codes. By the use of proposed technique, the required complexity reduces to $\frac{N}{2}(1 + \log \frac{N}{2})$ from $N(1 + \log N)$, which is the complexity of the conventional SC decoding. We show that the $(\log N)$ alternative folded-SC decoders with the same complexity have different performance. We show that at least some of the alternative folded-SC decoders can achieve the same error performance of the conventional SC decoders.

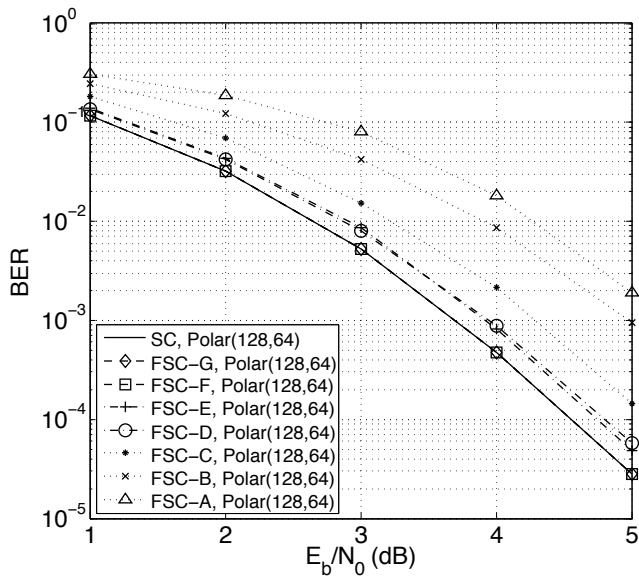


Fig. 7. Bit error rate performances of SC and folded SC for polar code (128,64).

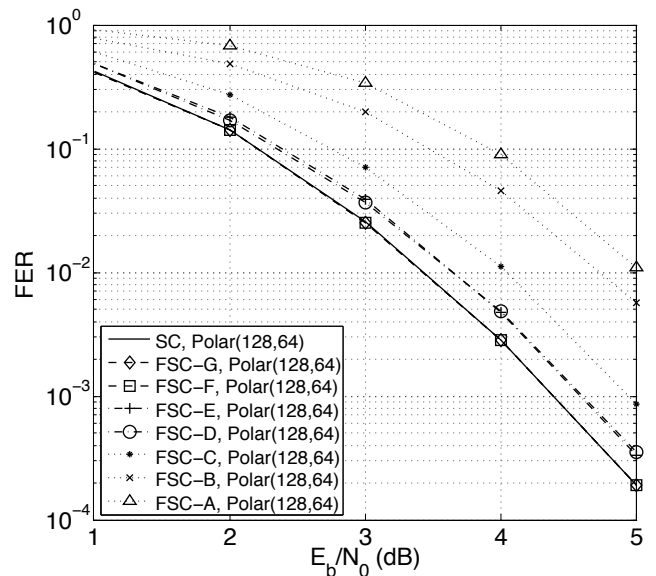


Fig. 9. Frame error rate performances of SC and folded SC for polar code (128,64).

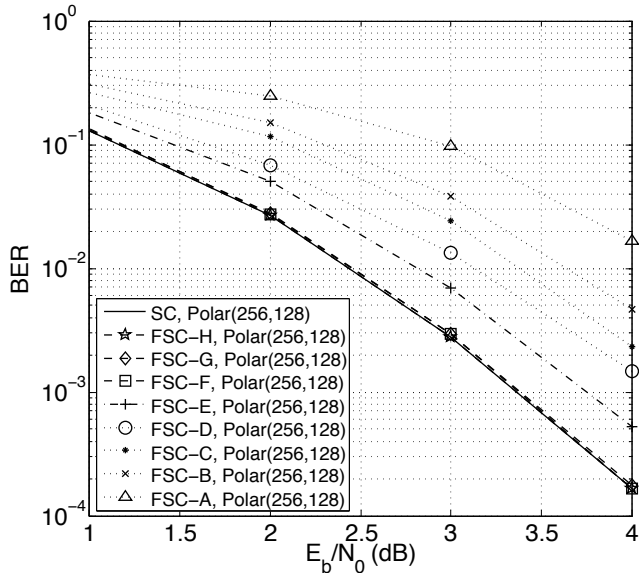


Fig. 8. Bit error rate performances of SC and folded SC for polar code (256,128).

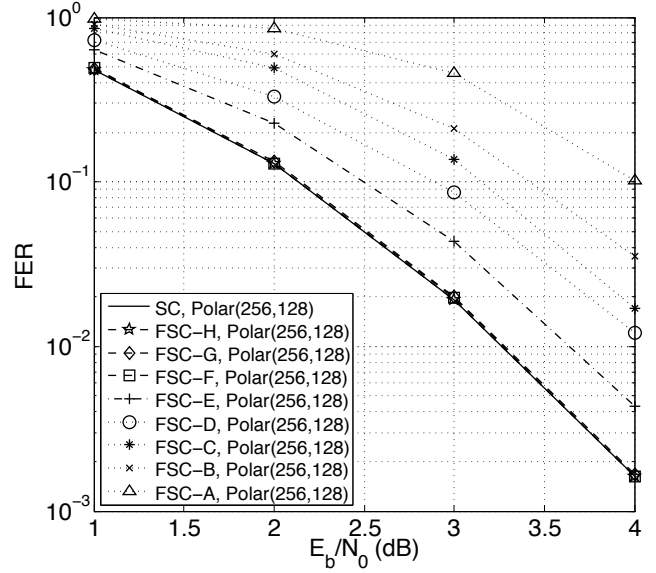


Fig. 10. Frame error rate performances of SC and folded SC for polar code (256,128).

Future work will focus on multiple foldings for SC decoding to achieve further complexity reductions.

REFERENCES

- [1] C.E. Shannon, "A mathematical theory of communication," *Bell System Tech. J.*, 27(2), pp. 379-423, 623-656, 1948.
- [2] E. Arkan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inform. Theory*, vol. 55, no. 7, pp. 3051-3073, 2009.
- [3] A. Pamuk, "An FPGA Implementation Architecture for Decoding of Polar Codes," in *Proc. 2011 IEEE Int. Symp. Wireless Comm. Systems*, (Aachen, Germany), pp. 437-441, 2011.
- [4] P. Trifanov, "Efficient Design and Decoding of Polar Codes," in *IEEE Trans. Commun.*, vol. 60, no. 11, pp. 3221-3227, 2012.
- [5] C. Leroux, I. Tal, A. Vardy, and W. J.Gross "Hardware architectures for successive cancellation decoding of polar codes," in *Proc. 2011 IEEE Conf. Int. Acoust., Speech, Signal Process. (ICASSP)*, pp. 1665-1668, 2011.
- [6] C. Leroux, A. J. Raymond, G. Sarkis, and W. J.Gross "A Semi-Parallel Successive-Cancellation Decoder for Polar Codes," *IEEE Trans. Signal Processing*, vol. 61, no. 2, pp. 289-299, 2013.
- [7] I. Tal, and A. Vardy, "List Decoding of Polar Codes," in *Proc. 2011 IEEE Int. Symp. Inform. Theory*, (St.Petersburg, Russia), pp. 1-5, 2011.
- [8] N. Hussami, S. B. Korada, and R. Urbanke, "Performance of polar codes for channel and source coding," in *Proc. 2009 IEEE Int. Symp. Inform. Theory*, (Seoul, Korea), pp. 1488-1492, 2009.
- [9] E. Arkan, "A performance comparison of polar codes and Reed-Muller codes," in *IEEE Commun. Letter*, vol. 12, no. 6, pp. 447-449, 2008.
- [10] S. Kahraman, and M. E. Çelebi, "Code based efficient maximum-likelihood decoding of short polar codes," in *Proc. 2012 IEEE Int. Symp. Inform. Theory*, (Cambridge, USA), pp. 1967-1971, 2012.
- [11] S. Kahraman, E. Viterbo, and M. E. Çelebi, "Folded Tree Maximum-Likelihood Decoder for Kronecker Product-Based Codes," in *Proc. 2013 Allerton Conference on Communication, Control, and Computing (Allerton)*, (Monticello IL, USA), 2013.
- [12] W. Sierpinski, "Sur une courbe dont tout point est un point de ramification," *C.R. Acad. Paris*, 160, pp. 302-305, 1915.