The Performance of Polar Codes for Multi-Level Flash Memories

Yue Li[1,2], Hakim Alhussien[4], Erich F. Haratsch[4], and Anxiao (Andrew) Jiang[1,3]

[1]Department of Computer Science and Engineering, Texas A&M University, College Station, TX 77843, USA

[2]Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125, USA

[3]Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843, USA

[4]Flash Components Division, LSI Corporation, San Jose, CA 95131, USA.

{*yli, ajiang*}*@cse.tamu.edu*

I. Introduction

To achieve higher storage density, NAND flash geometries keep shrinking while more quantization levels are put into each floating gate transistor. As a result, data are more prone to errors due to process variation and noise. The urgency for improving the reliability of flash memories calls for continuous search for optimal channel coding schemes. Polar codes proposed by Arıkan [1] is the first class of capacity-achieving codes with explicit constructions. Their attractive properties make them a potential candidate for optimal channel coding schemes. However, the practical performance of polar codes in flash memories is still unknown, and applying polar codes to flash channels presents many important challenges. For instance, polar codes require the code length to be an integer power of two which does not fit in flash pages of different sizes; to conduct rigorous experimental analysis, the decoding performance of polar codes need to be compared with that of other ECCs on the same random input and output data sets from flash characterization platforms, and such testing data are not assumed to be the codewords of any ECC; moreover, the construction of polar codes uses the channel statistics, and one concern is that new polar codes need to be frequently constructed for optimized performance when flash memory endures and the channel gradually degrades, which is prohibitively expensive in practice. Motivated by these challenges, we report part of the efforts towards realizing polar decoders for flash channels.

To make polar codewords fit different page sizes of flash memories, length-adapted codes are needed. Punctured polar codes have been studied recently [3] [4]. Puncturing has low implementation complexity, but degrades decoding performance due to the additional erasures introduced. This work explores an alternative approach for length-adapted polar codes through shortening. We propose the schemes for shortening both non-systematic [1] and systematic polar codes [2]. Shortening obtains a shorter codeword by assigning selected codeword symbols of the longer codeword to predetermined values made known both to the encoder and the decoder. The selected symbols are removed before transmission and inserted back before decoding. As the symbols inserted are correct, shortening does not introduce additional noise.

Rate-compatible polar codes can be implemented by adjusting the size of frozen sets without constructions of new codes [3]. We show that this property guarantees the feasibility of a practical adaptive polar decoding framework for flash channels. The decoder adaptively switches to use lower code rates as flash memory endures, and the code of each rate is used for a continuous range of program/erase cycles (PECs). We prove that repeatedly polar code construction is not necessary for such adaptive decoders. With the codes constructed for practical flash channels, we observed the order preservation of subchannel reliability, and our extensive experiments further demonstrate that the decoding performance by using one code closely approaches the optimized performance by constructing codes for different channel parameters.

II. Polar Codes in Flash Memories

We propose the schemes for shortening non-systematic and systematic polar codes for flash memories with different page sizes. The performance of shortened polar codes is further evaluated using data obtained from a flash characterization platform. The shortened polar codes are defined below:

Definition 1. *An (N, K, K')-shortened polar code (SPC) is a polar code of length $N - K'$ obtained from an (N, K)-polar code with block length $N = 2^m$ and information bit length K by assigning K' predetermined input symbols to known values before encoding, and removing K' predetermined codeword symbols after encoding.*

Let us define the notations used later in this section. Consider an (N, K) binary polar code with $N = 2^m$. Let the non-frozen set of the code be $\mathcal{A} \triangleq \{a_1, a_2, \cdots, a_K\} \subseteq \{1, 2, \cdots, N\}$, and let the frozen set $\bar{\mathcal{A}} \triangleq \{b_1, b_2, \cdots, b_{N-K}\}$ be the complement. We also assume that $a_1 < a_2 < \cdots < a_K$ and $b_1 < b_2 < \cdots < b_{N-K}$. Denote the input bits to the encoder by $\boldsymbol{u} \triangleq (u_1, u_2, \cdots, u_N) = (\boldsymbol{u}_{\mathcal{A}}, \boldsymbol{u}_{\bar{\mathcal{A}}})$ to represent $\boldsymbol{u}$, where $\boldsymbol{u}_{\mathcal{A}} \triangleq (u_i : i \in \mathcal{A})$ contains the message bits and $\boldsymbol{u}_{\bar{\mathcal{A}}} \triangleq (u_i : i \in \bar{\mathcal{A}})$ contains the frozen bits. The codeword $\boldsymbol{x} \triangleq (x_1, x_2, \cdots, x_N)$ computed by encoding is written to cells.

Shortened Non-systematic Polar Codes We first study the shortening of non-systematic polar codes (NSPCs) whose encoding of an (N, K)-NSPC follows the linear transformation $\boldsymbol{x} \triangleq \boldsymbol{u}\mathbf{G}$.

Theorem 2. *$(u_{N-K'+1}, u_{N-K'+2}, \cdots, u_N)$ are all 0s if and only if $(x_{N-K'+1}, x_{N-K'+2}, \cdots, x_N)$ are all 0s.*

The above theorem suggests we obtain an (N, K, K')-SPC from an (N, K)-NSPC by setting the last K' input bits to 0s, then removing the last K' codeword symbols after encoding. Among the K' input bits, there are K'' non-frozen bits and $K' - K''$ frozen bits where $K'' = |\{i | i \in \mathcal{A} \text{ and } N - K' + 1 \leqslant i \leqslant N\}|$. An (N, K, K')-SPC obtained through the encoding above has rate $\frac{K-K''}{N-K'} \in [\frac{K-K'}{N-K'}, \frac{K}{N}]$.

Shortened Systematic Polar Codes In practice, systematic codes are preferred because of its lower latency for reading information bits. Systematic polar codes (SYPCs) have been proposed recently by Arıkan [2]:

Definition 3. *[2] Let the sets $\mathcal{B} \subseteq \{1, 2, \cdots, N\}$, and $\bar{\mathcal{B}}$ be the complement. Therefore, $\boldsymbol{u} = (\boldsymbol{x}_{\mathcal{B}}, \boldsymbol{x}_{\bar{\mathcal{B}}})$. Let $\mathbf{G}_{\mathcal{AB}}$ be a submatrix of $\mathbf{G}$ such that for each element $G_{i,j}$, the indices $i \in \mathcal{A}$, $j \in \mathcal{B}$. For any given non-systematic polar encoder with parameter $(\mathcal{A}, \boldsymbol{u}_{\bar{\mathcal{A}}})$, a systematic polar* encoder $(\mathcal{B}, \boldsymbol{u}_{\bar{\mathcal{A}}})$ *exists if there is a one-to-one mapping from* $\boldsymbol{u}_{\mathcal{A}}$ *to* $\boldsymbol{x}_{\mathcal{B}}$ *following* $\boldsymbol{x}_{\mathcal{B}} = \boldsymbol{u}_{\mathcal{A}}\mathbf{G}_{\mathcal{AB}} + \boldsymbol{u}_{\bar{\mathcal{A}}}\mathbf{G}_{\bar{\mathcal{A}}\mathcal{B}}$, $\boldsymbol{x}_{\bar{\mathcal{B}}} = \boldsymbol{u}_{\mathcal{A}}\mathbf{G}_{\mathcal{A}\bar{\mathcal{B}}} + \boldsymbol{u}_{\bar{\mathcal{A}}}\mathbf{G}_{\bar{\mathcal{A}}\bar{\mathcal{B}}}$.

A systematic polar encoder defined above exists if $\mathcal{B} = \mathcal{A}$ [2]. To shorten SYPCs relies on the following theorem:

Theorem 4. *Let* $\mathcal{B} = \mathcal{A}$, *and let* $\boldsymbol{u}_{\bar{\mathcal{A}}}$ *be all* 0*s. There is a one-to-one mapping between* $(u_{a_{K-K'+1}}, u_{a_{K-K'+2}}, \cdots, u_{a_K})$ *and* $(x_{a_{K-K'+1}}, x_{a_{K-K'+2}}, \cdots, x_{a_K})$.

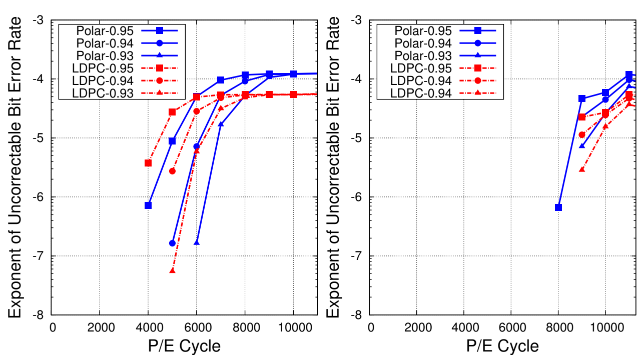
The theorem states that it is feasible to obtain an (N,K,K')-SPC from an (N,K)-SYPC by letting frozen bits be all 0s, and setting the last K' bits of $\boldsymbol{u}_{\mathcal{A}}$ to predetermined values before encoding. The last K' bits of $\boldsymbol{x}_{\mathcal{A}}$ are removed after encoding. An (N,K,K')-SYPC obtained through the encoding above has rate $\frac{K-K'}{N-K'}$.

Performance Evaluations We evaluated the decoding performance of shortened polar codes with the data from the characterizations of MLC flash chips using 2Y-nm technology from some vendor. The characterization process at each PEC sequentially program each page in a block with random input bits, reads the stored (and possibly noisy) data, and erase the block for the next write. As data written to the block are not ECC codewords, coset coding technique is needed to view such random sequences as the codewords of the ECC being evaluated. Fortunately, this is always feasible for polar codes:

Theorem 5. *Given an* (N,K,K')*-SPC from an* (N,K)*-NSPC with frozen set* $\bar{\mathcal{A}}$, *there is a unique* $\boldsymbol{u}'_{\mathcal{A}} \in \{0,1\}^{K-K'}$ *and a unique* $\boldsymbol{u}'_{\bar{\mathcal{A}}} \in \{0,1\}^{N-K-K'+K''}$, *such that*

$$(\boldsymbol{x}', \underbrace{0,\cdots,0}_{K'}) = ((\boldsymbol{u}'_{\mathcal{A}}, \underbrace{0,\cdots,0}_{K''})_{\mathcal{A}}, (\boldsymbol{u}'_{\bar{\mathcal{A}}}, \underbrace{0,\cdots,0}_{K'-K''})_{\bar{\mathcal{A}}}) \cdot \mathbf{G}.$$

Figure 1 shows the average uncorrectable bit error rates (UBERs) of shortened polar codes at different PECs with both hard sensing (Figure 1(a)) and soft sensing (Figure 1(b)). As there are four kinds of pages upper even, upper odd, lower even, and lower odd pages in a block, and each kind of pages have different raw bit error rates, we let each kind of pages use a different polar code. List decoding [5] is used with list size 32. We compare with LDPC codes using min-sum



(a) Hard Decoding (b) Soft Decoding

Fig. 1. The performance of polar codes and LDPC codes at different PECs.

decoding of equivalent rates. Three rates (0.93, 0.94 and 0.95) of interest to flash memories are used. We assume each page stores 8 length-7943 polar codewords shortened from a length-2^{13} polar code constructed for the channel parameters measured at the current PEC. The PECs when decoding failures first occur are of special interest to flash memories. The results suggest both LDPC and polar codes have similar performance in flash memories, and soft sensing significantly improves the endurance of MLCs.

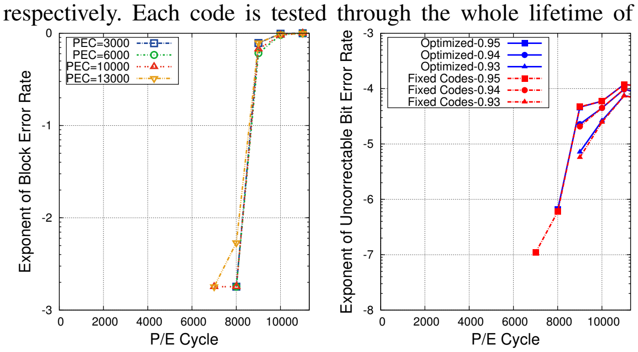
III. Adaptive Polar Decoding for Flash Memories

The channels of flash memories gradually degrade as PEC grows. Specifically, let the flash channel $W(\alpha)$ be parameterized by PEC $\alpha \in \mathbb{N}$, $W(\alpha')$ is degraded respect to $W(\alpha)$ for any α, α' such that $\alpha \leqslant \alpha'$. To keep error rates low, adaptive decoder is used in practice. Let $R_1 > \cdots > R_{k-1}$ be $k-1$ code rates of some channel code C, and let $\alpha_1 < \cdots < \alpha_k$ be k selected PECs. For $i \in \{1,2,\cdots,k-1\}$, an adaptive decoder of C changes the rate of C to R_i at α_i, and uses rate R_i consistently for any $\alpha \in [\alpha_i, \alpha_{i+1})$. We show that polar codes is an excellent candidate for effective adaptive decoding in flash memories in the sense that the construction of new codes is not necessary throughout the lifetime of flash chips, and changing code rate only requires freezing additional input bits. Let σ_W be the polarization order of the subchannels of the code for channel W. If for any PECs α, α' such that $\alpha \leqslant \alpha'$, $\sigma_{W(\alpha)} = \sigma_{W(\alpha')}$ holds (the condition is observed for flash channels in our experiments), we have the next two theorems:

Theorem 6. *For* $i \in \{1,2,\cdots,k-1\}$, *when the decoder changes the code rate* R_i *previously used at* $\alpha_{i+1}-1$ *to* R_{i+1} *at* α_{i+1}, *it only needs to further make the input bits in* $F_{w_{\alpha_{i+1}}} - F_{w_{\alpha_{i+1}-1}}$ *frozen where* $F_{w_{\alpha_{i+1}}}$ *and* $F_{w_{\alpha_{i+1}-1}}$ *are the frozen sets of the two codes.*

Theorem 7. *For* $i \in \{1,2,\cdots,k-1\}$, *given any two PECs* $\alpha, \alpha' \in [\alpha_i, \alpha_{i+1})$, *with the same code rate* R_i *the polar codes for* $W(\alpha)$ *and* $W(\alpha')$ *are equivalent.*

Theorem 7 states that no construction of new code is necessary for the PECs covered by the same code rate. Figure 2(a) shows the block error rates of four polar codes of rate-0.94 for the upper-odd pages constructed at PECs 3×10^3, 6×10^3, 10^4, and 1.3×10^4, respectively. Each code is tested through the whole lifetime of the



(a) Average BERs of upper-odd pages (b) Average UBERs over all pages

Fig. 2. The performance of polar codes constructed at fixed PECs throughout the lifetime of the flash chips. Soft sensing is used.

flash chips. The results suggest the codes yield very similar decoding performance due to the polarization order preservation. Figure 2(b) compares the average UBERs of the codes constructed at 6000 PECs with the optimized performance yield by codes constructed at different PECs. The performance of the scheme without construction of new code closely approaches the optimized performance.

References

[1] E. Arıkan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Information Theory.*, vol. 55, no. 7, pp. 3051–3073, July 2009.

[2] ——, "Systematic polar coding," *IEEE Communications Letters*, vol. 15, no. 8, pp. 860–862, 2011.

[3] A. Eslami and H. Pishro-Nik, "A practical approach to polar codes," in *Proc. ISIT*, 2011, pp. 16–20.

[4] D.-M. Shin, S.-C. Lim, and K. Yang, "Design of length-compatible polar codes based on the reduction of polarizing matrices," *IEEE Transactions on Communications*, vol. 61, no. 7, pp. 2593–2599, 2013.

[5] I. Tal and A. Vardy, "List decoding of polar codes," in *Proc. ISIT*, 2011, pp. 1–5.