Error correction in data storage systems using polar codes

Nana Kobina Gerrar$^{1,2}$ | Shengmei Zhao$^3$ | Lingjun Kong$^3$

$^1$ College of Overseas Education, Nanjing University of Posts and Telecommunications, Nanjing, China
$^2$ Department of Electrical/Electronics Engineering, Cape Coast Technical University, Cape Coast, Ghana
$^3$ Institute of Signal Processing and Transmission, Nanjing University of Posts and Telecommunications, Nanjing, China

Correspondence
Nana Kobina Gerrar, College of Overseas Education, Nanjing University of Posts and Telecommunications, Nanjing, China.
Email: nanagerrar@yahoo.com

Funding information
National Natural Science Foundation of China, Grant/Award Numbers: 61475075, 61871234; China NSF for Young Scientist, Grant/Award Number: 61501250

Abstract
This paper investigates error correction in two-dimensional (2-D) intersymbol interference (ISI) channels using polar codes. 2-D channels offer high data storage capacity compared to traditional one-dimensional (1-D) channels but suffer from greater ISI effects. Error correction codes can be used to mitigate the effects of ISI in 2-D data storage systems and improve their performance and reliability. Polar codes achieve the symmetric capacity for the class of binary-input discrete memoryless channels (B-DMCs). A post-processing technique is proposed to improve the performance of polar codes in the 2-D ISI channel. The simplified successive cancellation (SSC) polar decoder is useful for data storage systems owing to its good waterfall region performance, low error floors and good throughput rates and is therefore adopted for the study. Simulation results indicate that the proposed technique has an error correction performance comparable to some existing decoders and has better performance than other decoding techniques.

1 | INTRODUCTION

Recent telecommunication trends and advances in technology, such as internet-of-things (IoT) and cloud storage, present new data storage implications. It is projected that by the year 2025, as much as 163 zettabytes (ZB) of data will be created through various data applications [1]. Data storage technology has witnessed a rapid evolutionary trajectory over the past decades. This progress has led to advances in storage capacity and data transfer rates. Today, we have data storage systems that can hold terabytes (TBs) of data and possess writing/reading speeds of several Megabits/Sec (MB/s). However, there is a need for cheaper, higher capacity and more efficient data storage techniques. This will make it possible to meet the increasing data storage requirements of recent and future data storage applications.

One-dimensional (1-D) channels have traditionally been used in data storage systems; however, they possess limited data storage capacity. Two-dimensional (2-D) storage systems can achieve exponential areal density compared to 1-D tracks. Intersymbol interference (ISI) in 1-D data storage systems is limited and its effects are mitigated by the presence of a sequence detector [2]. However, ISI is a major concern for multi-dimensional data storage techniques [3]. 2-D systems, with multiple dimensions, describe storage conditions with smaller track-pitch and bit-width and so suffer from greater ISI effects [2]. One of the earliest investigations of data storage in 2-D channels can be found in [4]. Here, the authors study the possibility of increasing conventional recording beyond the perceived limit of 1 Tb/in$^2$. Since then, several studies have investigated magnetic recording in 2-D channels [5, 6]. In [5], schemes for signal processing in the two-dimensional magnetic recording (TDMR) channel are investigated. The channel considered exhibits write errors, 2-D interference, additive white Gaussian noise (AWGN) and random jitter. Also, in [6], the feasibility of attaining an areal density of 4 Tb/in$^2$ in data storage systems is addressed by applying a TDMR technique with particular emphasis on mitigating the effects of down-track and across-track location errors.

One way to handle errors in data storage systems is through replication. Here, copies of the data are stored in different locations. The backup information then helps in the recovery of the stored data in the event of data corruption or read
failure. However, replication can lead to performance degradation and it requires extra resources to replicate the data [7]. On the other hand, error correction codes (ECCs) improve the performance and reliability of data storage systems [8], thereby providing avenues for enhancing the advancements in storage capacity and access speeds. The unique characteristics of data storage systems imply that the application of recent advances in ECCs is not straightforward. For example, storage system technology requires high access speeds and therefore requires high throughput rates. Owing to the existence of different data types, there is also a challenge in applying ECCs to data storage systems [9]. Meta-data, for example, may require frequent updating of the recorded information [4]. Many coding options have been applied to improve error performance in data storage systems. Erasure codes have been popular in data storage technology and have largely been implemented by Reed-Solomon (RS) codes [10]. RS coding for distributed storage systems is effective but requires computations that make the encoding and decoding complexity very high for large code lengths [7]. Low-density parity check (LDPC) codes have emerged as an efficient substitute for RS codes. Properly adjusted signal processing and recent coding techniques can mitigate the effect of 2-D ISI as well as other channel impairments such as timing artefacts, jitter, and electronic noise [2]. LDPC codes have good performance in a wide range of practical applications. LDPC codes have been regarded as the coding scheme of choice in ECCs for data storage systems. Works concerning the application of binary LDPC codes to 2-D channels appear in [11–16].

Polar codes [17] are able to achieve the symmetric capacity for the class of binary input-discrete memoryless channels (B-DMCs). Various modifications to the conventional successive cancellation (SC) decoder [18–21] have significantly improved the performance of polar codes at finite code lengths. Desirable characteristics of polar codes such as good waterfall region performance, low error floors, and moderate throughput are important requirements for error correction schemes in data storage systems. Simplified successive cancellation (SSC) decoding of polar codes [19] provides throughput values that are sufficient for data storage [8]. Additionally, polar codes have been applied to 2-D data storage channels [22–24]. The authors in [22] use the mutual information between the input and output of a 2-D detector to optimize polar codes for the ISI data storage system. Bhatia et al. [23], propose the use of polar codes in place of LDPC codes in an interleaved multistage decoding (MSD) scheme. In this configuration, each sub-channel has a unique polar code. Fayyaz and Barry [24] also consider polar codes for the ISI channel, an MSD scheme for polar codes is investigated, and the corresponding extrinsic information transfer (EXIT) curve for the ISI channel is provided.

Noise can be added to the sub-optimal decoder to improve its successful detection of a received signal [25]. Also, in [26], white noise is added to a nonlinear system to improve signal detection. Perturbation has been applied to improve decoding in Polar and other error correction codes. In the presence of un-converged errors, the detection of signals can be aided through stochastic perturbation in nonlinear systems [27]. The authors achieve this by applying noise with varying power levels to parallel BP (belief propagation) decoders and selecting the decoder output with the best characteristics. The authors in [28] study the application of a dynamic perturbation method for the decoding of a polar-CRC cascaded code. Here, codewords are decoded each time by dynamically changing the variance of the perturbation noise. In [29], a scheme is proposed that improves the multi-round robin BP decoding of short LDPC codes through an input perturbation technique. In this technique, perturbation is done iteratively on a few symbols to widen the search space.

We presented a perturbed decoding approach to the decoding of polar codes in [30]. In this study, we present an enhanced perturbed polar decoding method to improve the performance of error correction in the 2-D data storage channel. A post-processing technique is proposed to improve the performance of the simplified successive cancellation (SSC) polar decoder in the 2-D ISI channel. Numerical simulations are carried out to verify the performance of the proposed technique. The simulation results indicate up to 0.75 dB performance gain over existing techniques. The genetic algorithm produces perturbation vectors that aid in the quick convergence of the technique. Two termination criteria ensure reduced complexity of the scheme. Post-processing technique has been applied to channel codes to improve the error correction performance. The proposed technique applies post-processing to improve the performance of polar codes in data storage systems. To the best of our knowledge, this is the first post-processing error correction technique applied to the decoding of polar codes in data storage systems.

The rest of the paper is organized as follows. In Section 2, the system model of the proposed technique is presented. This section entails an overview of emerging data storage systems, the discussion of the 2-D ISI channel and the proposed post-processing polar decoding technique for the 2-D ISI channel. In Section 3, a numerical analysis of the proposed technique is undertaken and the results and implications are discussed. The conclusion appears in Section 4.

2 | SYSTEM MODEL

In this section, emerging data storage systems are introduced, the 2-D ISI channel is presented, an overview of polar codes is given, and the proposed post-processing technique which applies polar codes to improve the error correction performance in the 2-D ISI channel is discussed.

2.1 | Overview of emerging data storage systems

Enhancements in magnetic recording technology have the potential to increase the storage capacity of data storage systems compared to conventional or traditional techniques like perpendicular magnetic recording (PMR). These novel techniques include bit-patterned media recording (BPMR), heat-assisted magnetic recording (HAMR) and two-dimensional magnetic recording (TDMR). For BPMR technology, the storage of each
bit of data is restricted to one magnetic recording medium or grain [4] as opposed to several grains per bit in traditional techniques like PMR. Special patterns are thus created out of the magnetic medium to store the data. These grains, though considerably smaller, are designed to be magnetically steady. In HAMR, the medium to be written onto is briefly heated before the writing process begins. This pre-processing enables the enhancement of the magnetization ability of the recording medium and also permits the use of much smaller grains or domains for the storage of data bits. The write heads in HAMR technology can access a smaller area on the magnetic medium. These new techniques aim at increasing the areal storage density of magnetic storage beyond 1 Tb/in². However, these new methods offer additional challenges for data storage efforts due to smaller track pitch and inter-track interference (ITI) leading to greater ISI effects [31]. Ensuring thermal stability and practical switching fields are some of the challenges for emerging recording technologies [4].

Shingled-writing is an important realization of magnetic recording in 2-D storage channels. In shingled writing, magnetic grains of relatively smaller sizes are used to store information. The write heads are therefore larger when compared to the tracks leading to the head accessing more than one track at a time during the writing process. As tracks are brought relatively closer together, there is an overlap between the current track and previously written tracks [2]. There is therefore an overhead cost in implementing shingled writing which requires re-writing tracks affected by the current operation. However, TDMR technology, unlike BPMR and HAMR, is compatible with current PMR methods. This makes TDMR easier to implement in existing data storage solutions. However, also for TDMR, the hard disk drive (HDD) architecture and software implementations may have to be modified to accommodate the TDMR technology. All these new techniques (TDMR, BPMR, and HAMR) pose additional ISI effects. There have been some suggestions in literature to overcome the challenges high storage capacity techniques face. Some of these suggestions include improved low-complexity read-back signal detection schemes and coding design optimization techniques [31].

2.2 | 2-Dimensional channel

2-D magnetic recording depends on advanced signal processing methods [32]. Figure 1 shows a discrete-time communication system which is a representation of a 2-D ISI channel. The input/write information to the system is \( x(i, j) \), \( h(i, j) \) represents the channel response, and \( e(i, j) \) is additive white Gaussian noise (AWGN) with zero mean and variance \( \sigma^2 \). The channel response is assumed to be symmetric both along-track and cross-track. The output/read information of the system is \( y(i, j) \) [33].

For a signal transmitted using a 2-D channel with an array of \( M \) rows and \( N \) columns, the channel output \( y(i, j) \) can be expressed according to [34] as:

\[
y(i, j) = \sum_{a=1}^{M} \sum_{b=1}^{N} h(a, b) x(i - a, j - b) + e(i, j) \tag{1}
\]

The input data is assumed to be \(-1\) outside the data range and selected with equal probability from the alphabet \( \{+1, -1\} \). If \( L_M \) and \( L_N \) denote the interference lengths in the horizontal and vertical directions, respectively, then, the 2-D channel response matrix [15] of size \( L_M \times L_N \) is expressible as:

\[
H = \begin{bmatrix}
  h_{(1,1)} & h_{(1,2)} & \cdots & h_{(1,L_N)} \\
  h_{(2,1)} & h_{(2,2)} & \cdots & h_{(2,L_N)} \\
  \vdots & \vdots & \ddots & \vdots \\
  h_{(L_M,1)} & h_{(L_M,2)} & \cdots & h_{(L_M,L_N)}
\end{bmatrix} \tag{2}
\]

To reduce the effects of ISI, and thus obtain an accurate estimate of the read-back signal, \( \hat{x}(i, j) \), a 2-D detection technique is required. Detectors are an important part of the 2-D data channel system configuration. Popular detection solutions developed for 1-D systems include Viterbi equalization and decision-feedback equalization (DFE). While 1-D detection systems are less complex as they relate only to the duration of the channel response and do not depend on the size of the input data, detectors for multi-dimensional channels are much more difficult to model due to the increasing complexity with larger data sizes [35]. An important desirable characteristic for a 2-D detector is the ability to deal with high levels of ITI. Additionally, a 2-D detector should be able to work with long code blocks and enable the attainment of performance close to capacity [4]. Detectors should also have the ability to mitigate the effects of ISI. A number of detectors have been suggested for 2-D data channels. These include the Bahl-Cocke-Jelinek-Raviv (BCJR) detector, which is considered to be an optimal 2-D detector but has prohibitive complexity making it unsuitable for most practical applications.

2.3 | Polar codes

An \( (N, K) \) polar code is synthesized with code length \( N \), and information set cardinality \( K \). The rate of the code can be expressed as:

\[
R = \frac{K}{N}. \tag{3}
\]

The polar encoding process requires the specification of a generator matrix \( G_N \). \( G_N \) can be expanded as:

\[
G_N = B_N F^\otimes u \]
where \(B_N\) is a bit-reversal matrix and \(F^\otimes n\) is the \(n\)th Kronecker power of \(F\). \(F_2\) can further be expressed as:

\[
F_2 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}
\]

The polar encoding process is completed according to [17] as:

\[
\mathbf{x} = \mathbf{u}G_N
\]

with \(\mathbf{x} = [x_1, x_2, ..., x_N]\) as the encoded bits and \(\mathbf{u} = [u_1, u_2, ..., u_n]\) representing the information bits. The set of information bits \(\mathbf{u}\) can further be divided into two sets; the free set \(\mathbf{u}_F\) and the frozen bit set \(\mathbf{u}_C\). The polar encoding process produces two categories of bit channels. The first group of channels have the symmetric capacity, \(K(W)\), that approach channel capacity, and the second group has \(L(W)\), that approach zero. Information can then be reliably transmitted over the channels using the first group of channels.

Several decoding algorithms have been developed for polar codes [18–21]. These decoding algorithms explore the improvement of the performance of the conventional polar decoding algorithm; the successive cancellation (SC) decoder. The SC decoder has poor performance at finite code lengths and also has high latency. Belief propagation (BP) [18], and SSC [19] decoding algorithms of polar codes are favored for error correction in data storage systems owing to their desirable characteristics such as good waterfall region, low error floors, and moderate throughput. In this study, we employ the SSC polar decoding technique for the 2-D ISI channel.

The message passing method used in SSC decoding is shown using the binary code tree in Figure 2. In SSC decoding, two categories of leaf nodes are considered; rate-one nodes and rate-zero nodes. If it is determined that the leaf nodes attached to a node are rate-one, decoding is done immediately, thereby simplifying the decoding process and offering improvement in latency and a reduction in the complexity. If \(D_q\) is considered to be the depth of the full binary tree with \(2^n\) leaves and \(n > 0\), a node, \(q\), can be identified with depth \(d_q\), parent \(p_q\), left child node \(g_{left}\), and right child node \(g_{right}\). Soft-information exchanged between node \(q\), its parent, left and right children are represented by \(\alpha_q, \alpha_{p_q}, \alpha_{g_{left}}, \alpha_{g_{right}}\). Also, estimated codewords generated by node \(q\) and its left and right children are represented by \(\beta_q, \beta_{g_{left}}, \beta_{g_{right}}\) respectively, with node \(q\) considered as the local decoder for its constituent code. The message passing process can be described as follows: First, node \(q\) receives soft-information vector, \(\alpha_q\), from its parent node, \(p_q\), and produces a codeword \(\beta_q\) after interaction with its left and right children. It should be noted that, for rate-one nodes, the evaluation of the output of the binary quantizer is done directly as \(\beta_q = h(\alpha_q)\). In the second step, all the bits that are related to node \(q\) as its children or descendants are computed immediately. The estimation of the bits is expressible using [19]

\[
[q_{\min}y_q, ..., q_{\max}y_q] = \beta_q G_{n-d_q}
\]

where \(y_q\) is the set containing the indices of all leaf nodes that are descendants of node \(q\) and \(G_{n-d_q}\) is the generator matrix for node \(q\).

In successive cancellation (SC), channel estimation is carried out by applying a decision directed channel estimation (DDCE) technique [17]. Likelihood ratios (LRs) are used to make decisions about the outcome of the decoding process. The output sequence \(y_i^N\) is received at the decoder input. The SC decoder generates estimates of the transmitted bits \(\{\hat{u}\}\) as \(\hat{u}\) in a sequential order \(\hat{u}_0, ..., \hat{u}_{N-1}\), this is done after considering \(Ai, u_F, Y\), and \(\gamma_i\) where \(A\) is the information set. Bit indexes relating to the frozen or fixed bits are determined directly and assigned their fixed values i.e. \(\hat{u}_0 = 0\). The sequence of bits can be estimated by computing the LR of the \(\hat{u}\) as:

\[
L_{N}^{(i)}(y^N_i, \hat{u}^{-1}) = \frac{L_{N}^{(i)}(y^N_i, \hat{u}^{-1}|0)}{L_{N}^{(i)}(y^N_i, \hat{u}^{-1}|1)}
\]

The SC decoder applies a decision criterion for the determination of the information bits. The criteria can be expressed as:

\[
\hat{u}_i = \begin{cases} 0, & \text{if } L_{N}^{(i)}(y^N_i, \hat{u}^{-1}) \geq 1 \\ 1, & \text{otherwise} \end{cases}
\]

Channel estimation in simplified successive cancellation (SSC) closely follows the technique adopted in [17]. However, for rate-one and rate-zero nodes there is no need for channel estimation as the bits are determined directly.

2.4 Post-processing decoding technique for 2-D ISI channel

Most conventional decoders make provision for one decoder output, per each received signal, during the data transmission process. This situation limits the capacity of the decoder in
the event of a wrongly decoded output. Multiple decoder outputs from a received signal can improve the reliability of the transmission process and avoid the need to re-send information that has been incorrectly decoded. Perturbation of the received signal can make available multiple decoder outputs per each received signal. Perturbation of the received signal can be explained as the addition of independent random noise to improve the performance of the suboptimal decoder [25]. The added noise is independent of the noise the transmitted signal experiences in the channel. The post-processing technique proposed in this paper takes advantage of the beneficial effects of perturbation to improve the performance of the SSC decoder in the 2-D ISI channel. The fundamental principle behind the perturbed decoding technique is illustrated in Figure 3. For each valid polar codeword, \( c = c_1, c_2, \ldots, c_g \), an error correction region is specified. The set of all error correction regions for all the valid codewords is specified as \( g = g_1, g_2, \ldots, g_s \), with \( S \) as the total number of valid codewords. For example, if a received signal \( y \) falls in its designated error correction region, \( g_2 \), the decoder can successfully estimate the transmitted codeword. However, for a received signal that falls outside its designated error correction region, the decoded sequence will be an invalid codeword. The addition of random noise, \( n \), to the received signal, \( y \), can shift it into its designated error correction region resulting in a successful decoding process.

The system model for the post-processing technique for the 2-D ISI channel is illustrated in Figure 4. The input information to be transmitted over the 2-D ISI channel is processed by a concatenated coding process where a cyclic redundancy check (CRC) sequence is generated from it. Then, after polar encoding, the signal goes through the 2-D channel where it encounters ISI and is observed with AWGN. The received/read signal, after 2-D detection, is decoded by a conventional decoding algorithm at the destination (receiver). A CRC verifies the output. In the event of a successful CRC, the decoding process ends. However, in the event of a failed CRC, the proposed post-processing algorithm is activated. The post-processing technique is composed of two parts; a perturbed algorithm and a genetic algorithm.

In the perturbed algorithm, independent random noise is added to the signal received from the 2-D detector, and the perturbed signal is again decoded by the conventional decoder. The process is repeated until a correct/valid codeword is obtained, or the threshold for a maximum number of iterations is achieved. If after the set threshold, a valid codeword is not found, the obtained perturbation noise sequences are refined by a genetic algorithm process. The new perturbation sequences are then added to the received signal, in turns, until a valid codeword is found or the set threshold for the genetic algorithm process is also attained. This is to ensure that any additional perturbation vectors are inherently better than the previous ones to guarantee a faster convergence for the proposed algorithm. If the genetic algorithm process also fails to produce a valid codeword then a decoding error is declared.

For the channel input information, \( u \), a CRC sequence is generated which makes it possible to verify the accuracy of the decoded sequence from the SSC polar decoder at the receiver. A CRC-16-CCITT polynomial is used to generate the CRC sequence. The polar encoding procedure follows (5) to produce encoded vector \( x \). The encoded vector is transmitted through the 2-D ISI channel and observed with AWGN. The output of the 2-D ISI channel goes to the 2-D detector. The Gaussian approximation low-complexity iterative row-column soft decision feedback algorithm (GA-IRCSDFA) detector proposed in [36] is adopted in this study. This detector has lower complexity than the original low-complexity iterative row-column soft decision feedback algorithm (IRCSDFA) detector [37]. The GA-IRCSDFA detector improves the complexity of the original IRCSDFA detector by reducing the strip-width of the row and column detectors to 1. Based on Gaussian approximation (GA), the output of the GA-IRCSDFA detector for either the row or column direction can be modelled according to [36] as:

\[
y(i, j) = \sum_{b=1}^{L_M} b(d_M, b) x(i, j - b + d_N) + e_r(i, j)
\] (9)

with \( d_M \) and \( d_N \) as the non-causality offsets. \( e_r(i, j) \) captures the equivalent noise and can further be expanded as:

\[
e_r(i, j) = \sum_{d_p, d_j=1}^{L_M} \sum_{b=1}^{L_M} b(a, b) x(i - a + d_M, j - b + d_N) + e(i, j)
\] (10)

The output of the 2-D detector serves as an input to the SSC polar decoder. The output of the SSC polar decoder, \( d \), is verified by CRC. In the event of a successful CRC, the decoding process ends. However, if CRC fails, the post-processing SSC (PPSSC) decoder is activated. With the PPSSC decoder, noise is added to the received signal which is the output from the 2-D detector. The perturbed received signal is expressible as:

\[
y_{p1} = y + n_{p1}
\] (11)
Figure 4 2-D ISI post-processing decoder

\( n_p \) indicates the random noise and is AWGN with variance \( \sigma^2 \). The perturbed received signal is then decoded by the SSC decoder and a decoded sequence is generated, \( d_p \). This is an iterative process that continues until the correct codeword is obtained or the threshold for the maximum number of iterations, \( T_1 \), is attained.

The PPSSC decoder has an additional feature to hasten the convergence of the iterative process and obtain the correct codeword. A genetic algorithm procedure enhances the iteration process by producing perturbation vectors that are inherently better than the previous ones. The perturbation noise sequences \( \{ n_{p1}, n_{p2}, \ldots, n_{pT_1} \} \), already generated, are used as the initial population for the genetic algorithm process. These perturbation noise sequences are assessed to determine their suitability as a solution to the decoding problem by applying a cost function. The fitness value of the \( i \)th member of the population is therefore evaluated according to [38] as:

\[
F_c = \sqrt{\frac{\sum_{i \in \mathcal{A}} [d - d_p(i)]^2}{|\mathcal{A}|}} \tag{12}
\]

where \( d \) is the decoded sequence of the received signal, \( d_p \) is the decoded sequence of the perturbed signal and \( \mathcal{A} \) is the information set of the polar code. A perturbed decoded sequence nearer to the originally decoded sequence can be considered closer to the solution with the assumption that the incorrect output information deviates from the correct output information by a small margin [38]. Also, the fitness score for each member of the population can be computed. This score is the ratio of the fitness score of the \( i \)th member of the population and cumulative sum of the fitness score of all the members of the population and can be expressed as:

\[
F_s = \frac{F_c(j)}{\sum_{j=1}^{T_1} F_c(j)} \tag{13}
\]

Using the fitness scores, members of the population can be selected to participate in the genetic operation process. The first step of the genetic operation process is crossover. In crossover, there is an exchange of sub-parts (bit-strings) between selected members of the population (parents) to form new individuals or offspring. For the single-point crossover technique, a random crossover point is established. A uniform random number \( \{a, a \in [0,1]\} \) determines whether the selected members of the population undergo the crossover process. The pre-selected crossover probability, \( p_c \), is compared to \( a \). When the condition \( a < p_c \) is met, then, a crossover point is determined as \( \text{cpoint} = \text{round}(a \times cl) \) where \( cl \) is the chromosome length. The perturbation vectors are encoded as chromosomes in the genetic algorithm process. At the crossover point, the selected members exchange sub-parts to form new individuals or offspring.
However, if \( a > p \), then no crossover occurs for the current pair. The next genetic operation is mutation. Mutation, which enhances genetic diversity, is done by randomly flipping one or more of the bits of the offspring created. A similar criterion described for the crossover operation is used to determine which offspring undergoes the mutation process. Here too, a uniform random number \( a \) is generated \( \{a, a \in [0,1]\} \) for each offspring created. If the condition \( a < p_m \) is satisfied, the offspring is muted. \( p_m \) indicates mutation probability. At the mutation point determined as \( m_{\text{point}} = \text{round}(a \cdot N) \), the bit corresponding to that point is flipped, that is a 1 for a 0 and a 0 for a 1. From the crossover and mutation operations, new members are created and a new population emerges. Each new population corresponds to a new generation that produces decoded sequences with higher likelihoods. The perturbed received signal, as a result of the genetic algorithm process, is expressible as:

\[
y_{p2} = y + n_{p2}
\]

where \( n_{p2} \) represents the random noise from the genetic algorithm process. The new outputs of the SSC polar decoder \( d_{p2} \) are then assessed by CRC. This iterative process continues until the correct codeword is obtained or the set threshold for the maximum number of generations \( T_2 \) is obtained. There is a decoding failure if the genetic algorithm process is also unable to generate a correct codeword. Algorithm 1 shows the post-processing process in detail.

### 2.5 Complexity analysis

The additional computation requirements of the polar decoding process, brought about by the PPSSC decoder are discussed in this section. The PPSSC decoder produces multiple decoded sequences in the event of a failed CRC. Consequently, the implementation of the PPSSC decoder leads to a higher computational complexity compared to using the SSC decoder without any post-processing. The specific analysis of the additional computation requirements of the proposed technique is discussed.

The evaluation is done by assessing the number of additional unit calculations which must be computed when the PPSSC decoder is employed [39]. In the following analysis, the computational cost associated with the CRC check is not considered [40]. For a code of length \( N \), the number of computations required to generate multiple codeword candidates by perturbing the output-read signal assuming the worst-case scenario of attaining the threshold limit is \( C_p T_1 N \), where \( C_p \) represents the unit calculations required and \( T_1 \) is the set threshold. Additionally, for the genetic algorithm process, the number of computations required can be captured by \( C_s T_2 N \) with \( C_s \) as the unit calculations required and, \( T_2 \), the set threshold. Here, \( C_p \) represents the unit calculations required by the genetic algorithm process. Therefore, the overall additional complexity brought about by the PPSSC decoder is \( O(C_p T_1 N + C_s T_2 N) \). The details of the unit calculations required in the estimation of \( C_p \) and \( C_s \) is shown in Table 1. The complexity of the PPSSC decoder is compared with other polar decoders in Table 2. \( I_{\text{max}} \) represents maximum iteration in BP decoding and \( L \) the list size in SCL decoding. From the table, it can be seen that the PPSSC decoder has a greater computational requirement than the other polar decoders. In PPSSC decoding, SSC decoding is performed for all information bits and additional SSC decoding for those bits that are incorrectly decoded according to Algorithm 1. The PPSSC decoder is only activated when the decoded sequence of the conventional SSC decoder fails the CRC. Therefore, the overall additional complexity occasioned by the proposed PPSSC decoder is minimal and its contribution to the overall decoding complexity of the whole scheme is not prohibitive.

### 3 NUMERICAL RESULTS AND ANALYSIS

In this section, numerical results of the performance of the PPSSC, SSC and BP decoders on the 2-D ISI channel are
TABLE 1 Unit calculations

| Perturbation algorithm calculations ($C_p$) | Genetic algorithm calculations ($C_g$) |
|-------------------------------------------|--------------------------------------|
| Multiplication (generation of perturbation noise) | Division, square root (fitness value evaluation) |
| Addition (perturbing the received signal) | Summation, division (fitness score evaluation) |
| Comparison (selection) | Comparison, addition (crossover) |
| Comparison, addition (mutation) | Addition (perturbing the received signal) |

TABLE 2 Complexity comparison of polar decoding schemes

| Technique | Complexity |
|-----------|------------|
| SC        | $O(N \log N)$ |
| BP        | $O(I_{max} N \log N)$ |
| SCL       | $O(L N \log N)$ |
| SSC       | $O(N \log N)$ |
| PPSSC     | $O(C_p N \log N) + O(C_g N \log N)$ |

Presented. A (1024, 512) polar code is used for the simulations at a code rate of 0.5. The channel response matrix of the 2-D ISI channel, $H_1$, is set at

$$H_1 = \begin{bmatrix} 0.0368 & 0.1435 & 0.0368 \\ 0.2299 & 0.8966 & 0.2299 \\ 0.0368 & 0.1435 & 0.0368 \end{bmatrix}$$

(15)

$H_1$ corresponds to a magnetic recording density of 4 Tb/in². The row matrix relates to the along-track direction, and the column relates to the cross-track direction. The BP decoder performs a maximum of 50 iterations. For the PPSSC decoder, the threshold for a maximum number of iterations is set at 10 and the threshold for the maximum number of generations is set at 50. The parameters of the simulation are summarized in Table 3. Figure 5 compares the BER performance of the PPSSC, SSC and BP decoders on the 2-D ISI channel. It can be observed that at a BER of $10^{-3}$, the PPSSC decoder exhibits a performance gain of 0.40 dB when compared to the SSC decoder and a performance gain of 0.18 dB when compared to the BP decoder.

TABLE 3 Parameters of the simulation

| Parameter                  | Value |
|----------------------------|-------|
| Channel                   | 2D ISI |
| Code length                | 1024  |
| Code rate                  | 0.5   |
| Modulation scheme          | BPSK  |
| Number of iterations (BP)  | 50    |
| Number of iterations (PPSSC)| 10    |
| Number of generations (PPSSC)| 50    |

The simulation results of the FER performance of the PPSSC, SSC and BP decoders on the 2-D ISI channel are shown in Figure 6. At a FER of $10^{-3}$, there is a 0.75 dB performance gain for the PPSSC decoder over the SSC decoder and a 0.40 dB gain over the BP decoder. In Figure 7, the BER performance of the PPSSC decoder at different iteration thresholds is presented. It can be observed that by increasing the threshold limit ($T_1$) the performance of the PPSSC decoder improves. At a BER of $10^{-3}$, the performance improves by 0.35 dB when $T_1$ is increased from 10 to 30. The improved performance can be explained by the fact that at a higher threshold the search space is widened making it easier to find a suitable perturbation vector that results in a valid codeword. However, increasing the threshold limit contributes to an increased complexity measure and a greater latency requirement. Additionally, the FER results for $T_1$ values confirm the superiority of the PPSSC decoder at higher threshold values as shown in Figure 8. For example, there is a 0.2 dB performance gain with $T_1 = 30$ compared to when $T_1 = 10$. The BER results for different code rates ($R = 0.5, 0.75$) of the PPSSC decoder is shown in Figure 9. There is a 1.1 dB...
loss in performance at a code rate of 0.75 compared to the performance at 0.5. Figure 10 shows the FER results for different code rates ($R = 0.5, 0.75$) of the PPSSC decoder. The results also show a decline of 1.2 dB when the PPSSC decoder is used at a code rate of 0.75 compared to the performance at 0.5.

4 | CONCLUSION

In this paper, the implementation of polar codes in data storage systems was investigated. A post-processing algorithm was proposed to improve the performance of the SSC polar decoder in the 2-D ISI channel. Perturbation of the received signal produces multiple decoder outputs from which a valid output/read decoded sequence can be obtained. A genetic algorithm improves the likelihoods of the perturbation vectors leading to faster convergence. Simulation results indicate that the proposed decoder improves the performance of conventional decoders for up to 0.75 dB on the 2-D ISI channel with little additional overhead complexity requirement. Increasing the threshold limit for the maximum number of iterations improves the performance of the proposed decoder but leads to higher complexity and latency requirements.

ACKNOWLEDGEMENT
This work was supported in part by the National Natural Science Foundation of China (Grant No. 61475075, 61871234), the project funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions. It is also supported by China NSF for Young Scientist (Grant No. 61501250).

REFERENCES
1. Reinsel, D., et al.: The digitization of the world from edge to core. An IDC White Paper No. 44413318, November 2018.
2. Garani, S., et al.: Signal processing and coding techniques for 2-D magnetic recording: An overview. Proc. IEEE 106(2), 286–318 (2018)
3. Huang, G.: Holographic memory. MIT Technol. Rev. 9 (2005)
4. Wood, R., et al.: The feasibility of magnetic recording at 10 terabits per square inch on conventional media. IEEE Trans. Mag. 45(2), 917–923 (2009)
5. Hwang, E., et al.: Signal processing for near 10 Tbits/in² density in two-dimensional magnetic recording (TDMR). IEEE Trans. Mag. 46(6), 1813–1816 (2010)
6. Hwang, E., et al.: Investigation of two-dimensional magnetic recording (TDMR) with position and timing uncertainty at 4 Tbit/in². 47(12), 4775–4780 (2011)
7. Plank, J., Thomassen, M.: A practical analysis of low-density parity check erasure codes for wide-area storage applications. In: Proceeding International Conference on Dependable Systems and Networks, Florence, Italy, 28 June–1 July, 2004
8. Sarks, G., Gross, J.: Polar codes for data storage applications. In: Proceeding IEEE International Conference on Computing, Networking, and Communications (ICNC), San Diego, CA, USA, 28–31 Jan. 840–844 (2013)
9. Casste, Y.: Coding techniques for data-storage systems. PhD thesis, California Institute of Technology, (2008)
10. Lee, O., et al.: A method for storage node allocation in erasure code based storage systems. In: Proceeding IEEE 3rd International Conference on Collaboration and Internet Computing San Jose, CA, USA, 15–17 Oct. 449–454 (2017)
11. Suzutou, R., et al.: Performance evaluation of TDMR R/W channel with head skew by LDPC coding and iterative decoding system. IEEE Trans. Mag. 1–4, 51(11), (2015)
12. Wang, Y., et al.: Improved BER performance with rotated head array and 2-D detector in two dimensional magnetic recording. IEEE Trans. Mag. 1–6, 52(7), (2016)
13. Shafiee, S., et al.: Novel joint detection and decoding schemes for TDMR channels. IEEE Trans. Mag. 1–6, 51(4), (2015)
14. Pituso, K., et al.: An intertrack interference subtraction scheme for a rate-4/5 modulation code for two dimensional magnetic recording. IEEE Trans. Lett. 7, (2016)
15. Kong, L., et al.: EXIT-chart-based LDPC code design for 2D ISI channels. IEEE Trans. Mag. 49(6), 2823–2826 (2013)
16. Chen, P., et al.: The design of protograph LDPC codes for 2-D magnetic recording channels. IEEE Trans. Mag. 1–4, 51(11), (2015)
17. Ankan, E.: Channel polarization: A method for constructing capacity achieving codes for symmetric binary-input memoryless channels. IEEE Trans. Inf. Theory 55(7), 3051–3079 (2009)
18. Hussami, N., et al.: Performance of polar codes for channel and source coding. In: Proceeding IEEE International Symp. in Information Theory (ISIT), Seoul, South Korea, 28 June–3 July 1488–1492 (2009)
19. Alamdar-Yazdi, A., Kschischang, F.: A simplified successive-cancellation decoder for polar codes. IEEE Comm. Lett. 15(12), 1378–1380 (2011)
20. Tal, I., Vardy, A.: List decoding of polar codes. IEEE Trans. Info. Theory 61(5), 2213–2226 (2015)
21. Niu, K., Chen, K.: Stack decoding of polar codes. IEEE Electronic Lett 48(12), 695–697 (2012)
22. Kong, L., et al.: Polar code design for ultra-high density magnetic recording channels. In: Proceeding Asia-Pacific Recording Conference, Shanghai, China 15–17 Nov. (2018)
23. Bhatia, A., et al.: Polar codes for magnetic recording channels. In: Proceeding IEEE Information Theory Workshop (ITW) Jerusalem, Israel, 26 Apr.-1 May 15 (2015)
24. Fayyaz, U., Barry, J.: Polar code design for intersymbol channels. In: Proceeding Symp on Sel. Areas in Comm. Austin, TX, USA, 2357–2361 (2014)
25. Kay, S.: Can detectability be improved by adding noise?. IEEE Signal Process. 7(1), 8–10 (2000)
26. Sutera, A.: Stochastic perturbation of a pure connective motion. J. Atmos. Sci. 37, 245–249 (1980)
27. Arli, A., Gazi, O: ‘Noise-aided belief propagation list decoding of polar codes. IEEE Comm. Lett. 23(8), 1285–1288 (2019)
28. Xiao, D., Gu, Z.: Dynamic perturbation decoding of polar-CRC cascaded code. In: Proceeding 2020 International Wireless Communications, and Mobobile Computing Limassol, Cyprus 5–19 Jun. (2020)
29. Lee, H., et al.: Multi-round belief propagation decoding with impulsive perturbation for short LDPC codes. IEEE Wireless Comm. Lett. 7(9), 1491–1494 (2020)
30. Gerrar, N., et al.: A CRC-aided perturbed decoding of polar codes. In: Proceeding 14th International Conference on Wireless Communications, Networ and Mob. Comps., Chongqing, China, (2018)
31. Guan, Y., et al.: Coding and signal processing for ultra-high density magnetic recording channels. In: Proceeding International Conference on Comp. Net., and Comm., Honolulu, HI, USA, 194–199 3–6 Feb., (2014)
32. Krishnan, A., et al.: 2-D magnetic recording: Read channel modelling and detection. IEEE Trans. Mag. 45(10), 3830–3836 (2009)
33. Zhao, X.: ‘LMMSSE equalization and DFE in 2-D ISI channels.’; In: Proceeding IEEE 4th International Congress on Image and Signal Processing, 15–17 Oct., (2011)
34. Chen, J., Siegel, P.: On the symmetric information rate of two-dimensional finite-state ISI channels. IEEE Trans. Info. Theory 52(1), 320–323 (2006)
35. Marrow, M., Wolf, J.: Iterative detection of 2-Dimensional ISI channels. In: Proceeding Information Theory Workshop (ITW). Paris, France, 31 March–4 Apr., 1311–134 (2003)
36. Zheng, J., et al.: Low-complexity iterative row-column soft decision feedback algorithm for 2-D intersymbol interference channel detection with Gaussian approximation. IEEE Trans. Mag. 49(8), 4768–4773 (2013)
37. Cheng, T., et al.: Row-column soft-decision feedback algorithm for two-dimensional intersymbol interference. IEEE Signal Process. Lett. 14(7), 433–436 (2007)
38. Jin, W., Fossorier, M.: ‘Reliability-based list decoding with multiple biases.’, IEEE Trans. Info. Theory 53(1), 105–120 (2007)
39. Chen, K.: Improved successive cancellation decoding of polar codes. IEEE Trans. Comm. 61(8), 3100–3107 (2013)
40. Guo, J., et al.: Multi-CRC polar codes and their applications. IEEE Comm. Lett. 20(2), 212–215 (2016)