Error-Pattern-Correcting Cyclic Codes Tailored to a Prescribed Set of Error Cluster Patterns

Jihoon Park, *Member, IEEE*, and Jaekyun Moon, *Fellow, IEEE*

***Abstract*—A new class of cyclic codes is discussed which is highly tailored to a prescribed set of dominant error cluster patterns. The cyclic code construction is based on a generator polynomial that produces a distinct syndrome set for each error pattern in the target set. By tailoring the generator polynomial specifically to the set of dominant error patterns, the code becomes highly effective in handling single and multiple occurrences of dominant error patterns at a very high code rate. A list decoding strategy based on a set of test word-error events is developed for the proposed codes, which efficiently utilizes both the algebraic information from the captured syndrome and the reliability measures provided by the local correlators matched to the dominant error patterns. By forcing a decoder to correct a single-pattern event for each test input word, multiple decoders running in parallel on the list of test words can effectively correct multiple error-pattern occurrences within the channel detector output word.**

***Index Terms*—Cyclic code, dominant error pattern, generator polynomial, list decoding, local error-pattern correlator, reliability measure, syndrome set, test word-error events.**

I. Introduction

INTERSYMBOL interference (ISI) arises in a wide variety of communications channels. High density data storage and band-limited wireline channels are good examples of communications channels wherein ISI is a major source of transmission errors. In many such channels, the characteristics of ISI are well understood and are known *a priori* at the receiver. Input-constrained ISI channels are of specific interest in this paper, as efficient coding methods for such channels are still largely an open area.

This paper is concerned with developing computationally efficient error correction coding methods for input-constrained ISI channels, with emphasis on high-density data storage applications. The structure of ISI typically gives rise to the dominance of a relatively small number of distinct error cluster patterns, regardless of the channel signal-to-noise ratio (SNR). One effective way to design codes for input-constrained ISI channels is to target such dominant error cluster patterns inherent at the channel detector output. Some error patterns in the target set may have large weights while some have low weights. Some patterns are long, while others can be very short. As such, traditional error correction methods, which target either the low weights of the error events or the maximum length of the error events, do not perform well, given a limited code redundancy. In contrast to the traditional approaches, our focus here is on a relatively small set of most likely error cluster patterns that are observed at the output of the optimal channel detector. This set can easily be found using distance analysis or through empirical observation based on simulation or laboratory measurement. This work is in some sense an effort to match the error correction code (ECC) to the given channel characteristics, as the dominance of a given set of error patterns at the channel detector output arises due to the nature of the channel.

Well-known earlier works on coding for input-constrained ISI channels can be found in [1]–[3]. Trellis-coding techniques for binary partial-response channels with particular channel null characteristics were studied in [1], where the Hamming distance of the code was used as the code design criterion after the channel ISI was essentially nullified by precoding. The work of [2] is also based on inverting the channel in some sense and relying on the Hamming distance property of the code to provide a coding gain. In contrast, the matched spectral null (MSN) codes of [3] exploit the channel memory, rather than nullifying it, in order to enhance distance properties at the channel output. The MSN codes are probably the earliest example of distance-enhancing constrained codes. There also have been other constrained codes that are tailored to the given ISI channel to enhance Euclidean distance at the channel output. Well-known examples of distance-enhancing constrained codes are the maximum transition run (MTR) code [4]–[6] and the forbidden-list code [7], [8] that prevent certain error-prone input sequences from entering the channel, based on the *a priori* knowledge of the dominant error patterns. Constrained-coding approaches along this direction are highly effective in the presence of certain signal-dependent noise and nonlinear distortions [9], [10], but in applications where the optimal code rate is high [11], as in high density disk drive systems, the significant rate loss associated with these types of codes has limited their usage.

In contrast, the proposed approach is concerned with error correction codes that do not directly impose a constraint on input sequences. We in fact aim at coding methods that can yield significant error correction powers at very high code rates, including those approaching 1. We do note that both the MSN code of [3] and our code pursue the same types of error patterns in many channels of interest (like the all-ones error pattern polynomials discussed in this paper). The major philosophical difference, however, is that the MSN code can be viewed more as an effort to suppress the problematic error patterns before

Manuscript received July 05, 2007; revised November 20, 2008. Current version published March 18, 2009. The material in this paper was presented in part at The Magnetic Recording Conference (TMRC), Minneapolis, MN, May 2007.

The authors are with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: jhpark@umn.edu; moon@umn.edu).

Communicated by G. Seroussi, Associate Editor for Coding Theory.

Color versions of Figs. 1–13 in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2009.2013019

they occur, whereas our code can be seen as an effort to correct the errors as they occur. Also, unlike the trellis-based codes in [1]–[3], our focus is on designing linear block codes whose algebraic properties are exploited in achieving favorable performance/complexity tradeoffs.

There also have been some attempts to address an observed set of dominant error cluster patterns. In fact, in the magnetic recording industry, parity-bit-based error detection, followed by post-Viterbi error correction using error pattern correlators, has been widely used [12], [13]. Single parity coding is particularly effective in detecting error patterns with odd weights, which are indeed prevalent error events in traditional longitudinal recording. A concatenation of the single-parity code with an existing block code has also been investigated to correct error patterns that are common in longitudinal recording [14]. As the industry shifts to the perpendicular form of recording, however, error detection or correction based on single parity coding has become ineffective, as many significant error patterns now have even weights. A graph-based construction of parity check matrices for detecting dominant error events of general form has been discussed in [15]. In [16], burst error detection codes and cyclic redundancy check (CRC) codes were generalized as error-event detection cyclic codes, to detect single and multiple occurrences of dominant error patterns with the least possible redundancy.

The current authors have also discussed linear block codes based on systematically constructed parity check matrices [17] and CRC codes [18] that can detect a prescribed set of dominant error patterns. However, error correction has depended solely on a reliability measure, such as the soft metric of [19] or the output of correlators matched to likely error events at the Viterbi detector output [12], [13]. In short, while the dominance of a relatively few distinct error patterns in data storage channels has long been recognized and error detection codes have been devised, there have not been systematic efforts to develop specialized error correction codes that can effectively correct these known, dominant error patterns.

In this paper, we propose error-pattern-correcting cyclic codes tailored to any given set of L dominant error cluster patterns. In practice, these patterns should be the ones that make up a very large percentage of all observed occurrences of error patterns. While this paper builds upon our previous works in [20]–[22], it presents a substantially enhanced method to handle multiple error-pattern occurrences by devising a list decoding strategy particularly suited to the presence of dominant error patterns. This paper also presents a theorem that is useful in designing the base code for generating distinct syndrome sets, which is the first step in the overall code design process.

We focus on cyclic codes that have substantial algebraic properties, so that implementation complexity is not a major issue in realizing the idea. Algebraic properties also usually lead to interesting mathematical solutions to the encoder/decoder design problem. Given a set of L targeted, dominant error patterns, the main objective is to effectively correct any single occurrences of the target error patterns as well as their multiple error occurrences.

First, we start with designing a low-degree generator polynomial $g(x)$ that can produce distinct, non-overlapping syndrome sets for the L target error patterns. No two error patterns within the list map to the same syndrome set, and the single occurrence of any target error pattern can be completely identified. Two construction methodologies for such generator polynomials are presented: the search procedure and a length-based construction method. By tailoring the generator polynomial specifically to the set of target error polynomials, the code becomes highly effective in handling frequently observed error patterns. Second, the generator polynomial $g(x)$ is multiplied by a primitive polynomial $p'(x)$ that is not already a factor of $g(x)$, such that the resulting code has a very high rate with the ability to uniquely identify the L dominant patterns preserved. With this $g(x)p'(x)$ as the new generator, we show that the vast majority of multiple error occurrences can easily be distinguished from single occurrences, based on the captured syndrome.

For the single occurrence of any target error pattern within the codeword length, the captured syndrome either uniquely points to the precise error position or suggests a small number of possible positions. Thus, the code itself can algebraically offer complete error correction capability in some cases where the precise position is available; in other cases, the final decision on the actual error position is made, among the few possible positions, based on a certain reliability measure provided by the correlator matched to the corresponding target error pattern. This error pattern correlator can also incorporate the *a priori* probability information that may be available from an outer decoder.

An efficient multiple-error-pattern correction method is also developed which utilizes a set of test word-error events. The idea is to force the decoder to focus on correction of only the single-pattern errors, while decoding is done in parallel with multiple decoders running on many test word-error events simultaneously. The test decoder input words are generated based on the outputs of the local correlators matched to the L dominant error patterns. The test input words are constructed so that with high probability at least one of them is within the correctable range of the transmitted codeword.

An analytical bound on the bit error rate (BER) of the proposed error-pattern-correcting codes is analyzed. The probabilities of multiple occurrences of the target error cluster patterns are also estimated. Both simulated BERs and the analytical bounds are compared for a representative interference-dominant channel, namely, a high density perpendicular magnetic recording channel, in order to demonstrate the viability of the proposed coding scheme. Sector error rates (SERs), assuming concatenation with an outer Reed–Solomon (RS) code, are also given.

II. Construction of High-Rate Error-Pattern-Correcting Cyclic Codes

A. Cyclic Codes Tailored to a Set of Target Error Patterns

Let $e_i(x)$'s, $i = 1, \ldots, L$, be the targeted, dominant error patterns[1] in the form of binary polynomials over *GF*(2). In a cyclic code constructed by a generator polynomial $g(x)$, the syndrome $s_i(x)$ for $e_i(x)$ is given by

$$s_i(x) = e_i(x) + g(x)Q_i(x) \tag{1}$$

[1]In this paper, "error pattern" and "error polynomial" will be used interchangeably.

where $Q_i(x)$ is the quotient when $e_i(x)$ is divided by $g(x)$ over GF(2). The syndrome set $\mathbf{S}_i$ for $e_i(x)$ is generated as the original captured syndrome $s_i(x)$ feeds through the feedback shift register, whose connection weights are specified by $g(x)$, i.e.

$$\mathbf{S}_i = \left\{ s_i(x), s_i^{(1)}(x), s_i^{(2)}(x), \ldots, s_i^{(p_i-1)}(x) \right\} \tag{2}$$

where $s_i^{(\rho)}(x) \triangleq x^\rho s_i(x) = x^\rho e_i(x)$ modulo $g(x)$ for a non-negative integer ρ, and p_i is the period of the syndrome set $\mathbf{S}_i$, such that $s_i^{(\rho)}(x) = s_i^{(\rho+p_i)}(x)$ and p_i is less than or equal to the order p of $g(x)$. Here the order p is the smallest possible integer such that $(x^p + 1)$ is a multiple of $g(x)$. Any syndrome element $s_i^{(\rho)}(x)$, $\rho = 0, \ldots, p_i - 1$ in $\mathbf{S}_i$ indicates the occurrence of the target error polynomial $e_i(x)$.

While the code is a cyclic code, the generator polynomial design is based on neither distance metric nor the maximum length of the burst errors; rather, the code design objective here is to construct a generator polynomial in such a way that the syndrome sets for the L target error polynomials are all distinct. In this sense, our approach to designing ECC is markedly different from the traditional approach of guaranteeing correction of up to t random errors or length-t burst errors within the received codeword.

Generator Polynomial Search Procedure:

Our starting point for constructing a suitable generator polynomial is to establish a sufficient condition for having distinct sets of syndrome patterns, based on the fact that all error polynomials can be factored into only a few irreducible polynomials in most cases.

Let $c_i(x)$, $i = 1, \ldots, L$, be the greatest common divisor (GCD) of $e_i(x)$ and a generator polynomial $g(x)$, i.e., $e_i(x) = c_i(x)e_i'(x)$ and $g(x) = c_i(x)g_i'(x)$, where $\mathrm{GCD}[e_i'(x), g_i'(x)] = 1$. The syndrome $s_i(x)$ for each error polynomial $e_i(x)$ is then given by

$$s_i(x) = c_i(x)[e_i'(x) + g_i'(x)Q_i(x)] \tag{3}$$

where $Q_i(x)$ is the quotient that arises when $g_i'(x)$ divides $e_i'(x)$. Any syndrome element $s_i^{(\rho)}(x)$, $\rho = 0, \ldots, p_i - 1$, in the corresponding syndrome set $\mathbf{S}_i$ contains $c_i(x)$ as a factor.

Property 1: If $c_i(x)$'s are all different, then the syndrome sets $\mathbf{S}_i$'s are distinct among different $e_i(x)$'s.

Proof: Given two different GCDs $c_i(x)$ and $c_j(x)$ for error polynomials $e_i(x)$ and $e_j(x)$, the syndrome elements in the corresponding syndrome sets $\mathbf{S}_i$ and $\mathbf{S}_j$ are respectively given by

$$\begin{aligned} s_i^{(\mu)}(x) &= c_i(x)[x^\mu e_i'(x) + g_i'(x)Q_i(x)] \\ s_j^{(\rho)}(x) &= c_j(x)[x^\rho e_j'(x) + g_j'(x)Q_j(x)] \end{aligned} \tag{4}$$

for any non-negative integers μ and ρ. Assume that $\mathbf{S}_i$ and $\mathbf{S}_j$ are the same. Then, $s_j^{(\rho)}(x)$ (or $s_i^{(\mu)}(x)$) must include $c_i(x)$ (or $c_j(x)$) as a factor, i.e.,

$$[x^\rho e_j'(x) + g_j'(x)Q_j(x)] = a(x)c_i(x)$$

where $a(x)$ is a binary polynomial. Since $g(x) = c_i(x)g_i'(x) = c_j(x)g_j'(x) = c_j(x)[c_i(x)g_j''(x)]$, we can rewrite it as

$$x^\rho e_j'(x) = c_i(x)[a(x) + g_j''(x)Q_j(x)].$$

However, the equality cannot hold for any value of ρ or any polynomial $a(x)$, because $c_i(x) \neq \mathrm{GCD}[e_j(x), g(x)]$, and accordingly, $x^\rho e_j'(x)$ does not have a factor of $c_i(x)$. This is a contradiction to the assumption. Therefore, $\mathbf{S}_i \neq \mathbf{S}_j$. $\square$

Based on this property, we can easily construct a generator polynomial that satisfies the sufficient condition, giving different GCDs for all $e_i(x)$'s, by collecting all irreducible factors of the target error polynomials in the following way:

$$g(x) = p_1^{d_1}(x)p_2^{d_2}(x)\cdots p_K^{d_K}(x) \tag{5}$$

where $p_k(x)$'s, $k = 1, 2, \ldots, K$, are the irreducible factors making up all $e_i(x)$'s, and d_k is the maximum power with which $p_k(x)$ appears in any $e_i(x)$, $i = 1, 2, \ldots, L$.

While having different GCDs for all $e_i(x)$'s guarantees distinct syndrome sets, the converse is not necessarily true. Suppose that $\mathbf{S}_i$ and $\mathbf{S}_j$ are different, i.e., $s_i^{(\mu)}(x) \neq s_j^{(\rho)}(x)$ for any μ and ρ. By using (4), we have

$$c_i(x)[x^\mu e_i'(x) + g_i'(x)Q_i(x)] \neq c_j(x)[x^\rho e_j'(x) + g_j'(x)Q_j(x)].$$

Since $e_i(x) = c_i(x)e_i'(x)$, $e_j(x) = c_j(x)e_j'(x)$ and $g(x) = c_i(x)g_i'(x) = c_j(x)g_j'(x)$, the above equation can be written as

$$[x^\mu e_i(x) + x^\rho e_j(x)] \neq g(x)[Q_i(x) + Q_j(x)].$$

Thus, if $\mathbf{S}_i \neq \mathbf{S}_j$, then $x^\mu e_i(x) + x^\rho e_j(x)$ is not a multiple of $g(x)$ for any μ and ρ, irrespective of $c_i(x) = c_j(x)$ or $c_i(x) \neq c_j(x)$. Also, as long as $x^\mu e_i(x) + x^\rho e_j(x)$ is not a multiple of $g(x)$, $\mathbf{S}_i$ and $\mathbf{S}_j$ are always different. In fact, we make a note of the following obvious property.

Property 2: Given two error polynomials $e_i(x)$ and $e_j(x)$, the corresponding syndrome sets $\mathbf{S}_i$ and $\mathbf{S}_j$ are distinct if and only if $x^\mu e_i(x) + x^\rho e_j(x)$ is not a multiple of $g(x)$ for any non-negative integers μ and ρ.

While $g(x)$ in (5) results in a fairly large number of parity bits leading to a large rate loss and may not be directly useful in high-rate code design, it provides a hint at the proper use of $p_k(x)$'s in constructing a more efficient $g(x)$ structure that can do the job. We retain the basic form of $g(x)$ as in (5), but lower the powers associated with each $p_k(x)$, while monitoring for the sufficient condition of *Property 1* as well as *Property 2*. We basically search for a non-negative integer set $\{\gamma_1, \gamma_2, \ldots, \gamma_K\}$ leading to the minimum degree polynomial of the form

$$g(x) = p_1^{\gamma_1}(x)p_2^{\gamma_2}(x)\cdots p_K^{\gamma_K}(x) \tag{6}$$

that would yield L distinct syndrome sets for the L target error polynomials. The search procedure is summarized as follows.

i) Set up a degree equation as

$$d = m_1\gamma_1 + m_2\gamma_2 + \cdots m_K\gamma_K$$

where m_k, $k = 1, 2, \ldots, K$, is the degree of $p_k(x)$, and start with an integer $d = \{\min d_i \mid (2^{d_i} - 1) \geq L\}$.

ii) Find all possible combinations of non-negative integers $\{\gamma_1, \gamma_2, \ldots, \gamma_K\}$ to yield the value of d, and construct $g(x)$ for each combination as

$$g(x) = p_1^{\gamma_1}(x)p_2^{\gamma_2}(x)\cdots p_K^{\gamma_K}(x).$$

TABLE I
DOMINANT ERROR PATTERNS/POLYNOMIALS FOR A CHANNEL RESPONSE OF $5 + 6D - D^3$ UNDER 100% AWGN

Dominant error pattern	Length	Dominant error polynomial
±[2]	1	1
±[2,-2, 2]	3	$(1+x+x^2)$
±[2,-2]	2	$(1+x)$
±[2,-2, 2,-2]	4	$(1+x)^3$
±[2,-2, 2,-2, 2]	5	$(1+x+x^2+x^3+x^4)$
±[2,-2, 2,-2, 2,-2]	6	$(1+x)(1+x+x^2)^2$
±[2,-2, 2,-2, 2,-2, 2]	7	$(1+x+x^3)(1+x^2+x^3)$
±[2,-2, 2,-2, 2,-2, 2,-2]	8	$(1+x)^7$
±[2,-2, 2,-2, 2,-2, 2,-2, 2]	9	$(1+x+x^2)(1+x^3+x^6)$
±[2,-2, 2,-2, 2,-2, 2,-2, 2,-2]	10	$(1+x)(1+x+x^2+x^3+x^4)^2$

iii) Given targeted, dominant error patterns $e_i(x)$'s, $i = 1, \ldots, L$, compute

$$\{c_1(x), c_2(x), \ldots, c_L(x)\}$$

where $c_i(x)$ is the GCD of $e_i(x)$ and $g(x)$.

iv) If $c_i(x)$'s are all distinct, we are all set; we now have $g(x)$. Else if $e_i(x) + x^\rho e_j(x), i, j \in \{1, 2, \ldots, L\}$ is not a multiple of $g(x)$ for $c_i(x) = c_j(x)$, for any ρ, we are done. Otherwise, $d := d + 1$ and go to ii).

• *Example* 1: Table I lists 10 dominant error patterns (and their polynomial representations) that arise in perpendicular magnetic disk drives at the output of the Viterbi detector, in descending order of the occurrence frequencies. Here, the dominant error patterns make up 99.434% of the observed 220, 207 error patterns at a BER of 2.399×10^{-3}. The channel response corresponds to $5 + 6D - D^3$, which is considered to be a good model for high density perpendicular recording [23]; it has a large dip around the band edge, giving rise to essentially the same set of error patterns as the $(1 + D)^n$ partial response channel extensively studied in [3]. A similar list of frequent error patterns are also given in [24] and [25] for somewhat different models for perpendicular recording. While longer error patterns of the same form yield the same energy in the channel output error sequence, the *a priori* probability of the input sequence that can allow this particular form of error pattern decreases exponentially as a function of the pattern length. Here, our choice of the maximum target error pattern length is driven largely by the complexity and rate requirements of the code.

It is seen in Table I that there are six different irreducible polynomial factors: $p_1(x) = (1 + x)$, $p_2(x) = (1 + x + x^2)$, $p_3(x) = (1+x+x^2+x^3+x^4)$, $p_4(x) = (1+x+x^3)$, $p_5(x) = (1 + x^2 + x^3)$ and $p_6(x) = (1 + x^3 + x^6)$. Collecting each of these irreducible polynomials along with its maximum power as it appears in any error polynomial, we construct $g(x)$ that guarantees distinct GCDs for the listed 10 target error patterns using (5)

$$g(x) = p_1^7(x)\, p_2^2(x)\, p_3^2(x)\, p_4(x)\, p_5(x)\, p_6(x).$$

The resulting degree of $g(x)$ is 31.

Now, going through the proposed search procedure, we quickly find a fairly low degree $g(x)$ in the general form of (6) that produces distinct syndrome sets for the 10 target error polynomials. As a result, two degree-8 generator polynomials are obtained as

$$\begin{aligned} g(x) &= p_1^2(x)\, p_2^1(x)\, p_3^1(x)\, p_4^0(x)\, p_5^0(x)\, p_6^0(x) \\ &= 1 + x^3 + x^5 + x^8 \end{aligned}$$

with $\{\gamma_1 = 2, \gamma_2 = \gamma_3 = 1, \gamma_4 = \gamma_5 = \gamma_6 = 0\}$, and

$$\begin{aligned} g(x) &= p_1^2(x)\, p_2^0(x)\, p_3^0(x)\, p_4^0(x)\, p_5^0(x)\, p_6^1(x) \\ &= 1 + x^2 + x^3 + x^5 + x^6 + x^8 \end{aligned}$$

with $\{\gamma_1 = 2, \gamma_2 = \gamma_3 = \gamma_4 = \gamma_5 = 0, \gamma_6 = 1\}$.

We note that there exist no lower-degree generator polynomials that satisfy the requirement.

Length-based Generator Polynomial Construction: When the coefficients of the error polynomials corresponding to the targeted error patterns are all equal to 1, a nonsearch-based approach is possible in constructing a desired $g(x)$. The approach is based on the observation of the lengths of the target error polynomials. Assume the coefficients of every one of the L target error polynomials $e_i(x)$'s are all one. Then, the following theorem provides a necessary and sufficient condition for finding generator polynomials that produce distinct syndrome sets for the given L all-one's target error polynomials.

Theorem 1: Let $g(x) = g_1(x)g_2(x)$, where $g_1(x)$ is either $(1+x^{p_1})$ or its factor having order p_1, while $g_2(x)$ is either $(1+x+x^2+\cdots+x^{p_2-1})$ or its factor having order p_2. Then, given the all-one's error polynomials $e_i(x)$ of length l_i and $e_j(x)$ of length l_j, $l_i > l_j > 0$, the corresponding syndrome sets $\mathbf{S}_i$ and $\mathbf{S}_j$ are distinct, if and only if

$$\{p_1, p_2\} \neq \left\{ \frac{l_i \pm l_j}{2m}, \frac{l_i \mp l_j}{2n} \right\}$$

for any positive integers m and n.

Proof: [*Necessary condition*] Consider the following polynomials $e_i(x)$ and $e'_j(x)$: $e_i(x) = a(x) + x^{mp_1}b(x)$ and $e'_j(x) = a(x) + b(x)$, where $a(x)$ of length mp_1 and $b(x)$ of length np_2 are all-ones polynomials. Then, $e_i(x)$ and $e'_j(x)$ are given by

$$\begin{aligned} e_i(x) &= (1 + \cdots + x^{mp_1-1} + x^{mp_1} + \cdots + x^{mp_1+np_2-1}) \\ e'_j(x) &= x^{\min[mp_1,\, np_2]}(1 + \cdots + x^{|mp_1-np_2|-1}) \\ &\triangleq x^{\min[mp_1,\, np_2]} e_j(x) \end{aligned}$$

where $e_j(x) \triangleq (1 + \cdots + x^{|mp_1-np_2|-1})$ and $e_j(x) = 0$ if $mp_1 = np_2$. This $e_i(x) + e'_j(x)$ can also be written as

$$\begin{aligned} e_i(x) + e'_j(x) &= [a(x) + x^{mp_1}b(x)] + [a(x) + b(x)] \\ &= b(x)(1 + x^{mp_1}) \end{aligned}$$

over GF(2). Since $g_1(x)$ is a factor of $(1+x^{mp_1})$ for any $m \geq 1$, $e_i(x) + e'_j(x)$ should be a multiple of $g_1(x)$. Also, $b(x)(1 + x) = (1 + x^{np_2})$ is a multiple of $(1 + x^{p_2}) = (1 + x)(1 + x + \cdots + x^{p_2-1})$, and $g_2(x)$ is a factor of $(1 + x + \cdots + x^{p_2-1})$. Consequently, $b(x)$ should also be a multiple of $g_2(x)$. Therefore, $e_i(x) + e'_j(x)$ is a multiple of $g(x)$, i.e.,

$$\begin{aligned} e_i(x) + e'_j(x) &= b(x)(1 + x^{mp_1}) \\ &= e_i(x) + x^{\min[mp_1,\, np_2]} e_j(x) \\ &= g_1(x)g_2(x)Q(x) = g(x)Q(x). \end{aligned}$$

This indicates that $g(x)$ produces the same syndrome set for $e_i(x)$ of length (mp_1+np_2) and $e_j(x)$ of length $|mp_1-np_2|$. Accordingly, If $p_1=\frac{l_i\pm l_j}{2m}$ and $p_2=\frac{l_i\mp l_j}{2n}$, then syndrome sets $\mathbf{S}_i$ and $\mathbf{S}_j$ for $e_i(x)$ of length $l_i\ (=(mp_1+np_2))$ and $e_j(x)$ of length $l_j\ (=|mp_1-np_2|)$ are the same. Therefore, if $\mathbf{S}_i\neq\mathbf{S}_j$, then $\{p_1,p_2\}\neq\{\frac{l_i\pm l_j}{2m},\frac{l_i\mp l_j}{2n}\}$.

[*Sufficient condition*] Suppose that $e_i(x)+x^\rho e_j(x)$ is a multiple of $g(x)=g_1(x)g_2(x)$ for a non-negative integer ρ, i.e., $\mathbf{S}_i=\mathbf{S}_j$. Then, $e_i(x)+x^\rho e_j(x)$ must include $g_1(x)$ and $g_2(x)$ as factors, i.e., $e_i(x)+x^\rho e_j(x)=g_1(x)g_2(x)Q(x)$. Then, this $e_i(x)+x^\rho e_j(x)$ can be written over GF(2) as

$$\begin{aligned}
e_i(x)&+x^\rho e_j(x)=g_1(x)g_2(x)Q(x)\\
&=(1+x+\cdots+x^{l_i-1})+x^\rho(1+x+\cdots+x^{l_j-1})\\
&=[(1+\cdots+x^{\rho+l_j-1})+x^{\rho+l_j}(1+\cdots+x^{l_i-\rho-l_j-1})]\\
&\quad+[(1+\cdots+x^{\rho+l_j-1})+(1+\cdots+x^{\rho-1})]\\
&=x^{\rho+l_j}(1+\cdots+x^{l_i-\rho-l_j-1})+(1+\cdots+x^{\rho-1}). \qquad (7)
\end{aligned}$$

Recall that $g_1(x)$ is either $(1+x^{p_1})$ or its factor, and $g_2(x)$ is either $(1+x+x^2+\cdots+x^{p_2-1})$ or its factor. Thus, for $e_i(x)+x^\rho e_j(x)$ to be multiples of both $g_1(x)$ and $g_2(x)$, the power $(l_i-\rho-l_j)$ needs to be identical to ρ in (7), such that $e_i(x)+x^\rho e_j(x)$ becomes

$$\begin{aligned}
e_i(x)&+x^\rho e_j(x)\\
&=(x^{\rho+l_j}+1)(1+\cdots+x^{\rho-1})\\
&=\left(x^{\frac{l_i+l_j}{2}}+1\right)\left(1+x+\cdots+x^{\frac{l_i-l_j}{2}-1}\right)
\end{aligned}$$

where $\rho=\frac{l_i-l_j}{2}$. Thus, if $\{\frac{l_i+l_j}{2},\frac{l_i-l_j}{2}\}=\{mp_1,np_2\}$ or $\{np_2,mp_1\}$ for any positive integers m and n, $e_i(x)+x^\rho e_j(x)$ is always a multiple of $g(x)=g_1(x)g_2(x)$. The lengths l_i and l_j of $e_i(x)$ and $e_j(x)$ are then given by

$$l_i=mp_1+np_2 \quad \text{and} \quad l_j=|mp_1-np_2|$$

and the order pair $\{p_1,p_2\}$ becomes $\{\frac{l_i\pm l_j}{2m},\frac{l_i\mp l_j}{2n}\}$. Therefore, if $\{p_1,p_2\}\neq\{\frac{l_i\pm l_j}{2m},\frac{l_i\mp l_j}{2n}\}$, then $\mathbf{S}_i\neq\mathbf{S}_j$. □

Theorem 1 can be used to quickly rule out the order pairs that cannot give rise to distinct syndrome sets. We shall call such order pairs "unsuitable" order pairs in the sense that we will have to avoid the corresponding $g_1(x)$'s and $g_2(x)$'s in constructing $g(x)$. Assuming L different lengths, we need to produce $\binom{L}{2}$ order pairs corresponding to the same number of length pairs. A quick glance at Theorem 1, however, reveals that we just need to examine those length pairs for which $(l_i\pm l_j)$ is even, since, otherwise $\frac{(l_i\pm l_j)}{2m}$ cannot be an integer for any positive integer m. It means that we can just examine the length pairs l_i and l_j such that they are either both even or both odd. Thus, given the L all-one's target error polynomials, it is not necessary to exhaust all $\binom{L}{2}$ length pairs. Rather, a set of even lengths and a set of odd lengths can be separately considered in identifying the unsuitable order pairs. Moreover, once the unsuitable order pair $\{\frac{(l_i+l_j)}{2},\frac{(l_i-l_j)}{2}\}$, $l_i>l_j$, is obtained, then all the factors in the form of $\{\frac{(l_i+l_j)}{2m},\frac{(l_i-l_j)}{2n}\}$ are automatically determined as unsuitable order pairs. The acceptable order pairs can be simply obtained after eliminating all unsuitable order pairs. The remaining order pairs point to the desired $g_1(x)$ and $g_2(x)$, and thus to $g(x)$ itself. The construction steps are as follows.

i) Divide l_i's into two sets: a set of even lengths $\{l_{e_1},l_{e_2},\ldots\}$ and a set of odd lengths $\{l_{o_1},l_{o_2},\ldots\}$.

ii) For any length pair in each set, compute the following unsuitable order pair:

$$\left\{\frac{(l_i+l_j)}{2},\frac{(l_i-l_j)}{2}\right\}$$

where both l_i and l_j are elements of either $\{l_{e_1},l_{e_2},\ldots\}$ or $\{l_{o_1},l_{o_2},\ldots\}$, and $l_i>l_j$.

iii) Set up acceptable order pairs $\{\hat{p}_1,\hat{p}_2\}$'s. Here both $\hat{p}_1$ and $\hat{p}_2$ should be neither unsuitable order pairs obtained in ii) nor their factors.

iv) For each $\{\hat{p}_1,\hat{p}_2\}$, construct a generator polynomial as

$$g(x)=g_1(x)g_2(x)$$

where $g_1(x)$ is either $(1+x^{\hat{p}_1})$ or its factor, and $g_2(x)$ is either $(1+x+\cdots+x^{\hat{p}_2-1})$ or its irreducible factor.

• *Example* 2: For the 10 target error patterns shown in Table I, there are five even lengths $l_e=\{2,4,6,8,10\}$ and five odd lengths $l_o=\{1,3,5,7,9\}$. Table II lists 10 unsuitable order pairs corresponding to all possible length pairs in each of the sets l_e and l_o. For an even length pair [6, 2], the corresponding unsuitable order pair is $\{\frac{6+2}{2},\frac{6-2}{2}\}=\{4,2\}$ and its factors $\{1,1\}$, $\{1,2\}$, $\{2,2\}$ and $\{4,1\}$ are immediately ruled out as unsuitable order pairs. In the end, it is easily found that order pairs $\{1,10\}$, $\{2,9\}$, $\{3,8\}$ and $\{5,6\}$ do not belong to the set of the unsuitable order pairs and their factors. Considering the order pairs $\{1,10\}$, $\{2,9\}$, $\{3,8\}$ and $\{5,6\}$ as acceptable order pairs $\{\hat{p}_1,\hat{p}_2\}$'s, we can obtain the following generator polynomials: $g(x)=(1+x)(1+x+x^2+x^3+x^4+x^5+x^6+x^7+x^8+x^9)$ from $\{\hat{p}_1,\hat{p}_2\}=\{1,10\}$, $g(x)=(1+x^2)(1+x^3+x^6)$ from $\{\hat{p}_1,\hat{p}_2\}=\{2,9\}$, $g(x)=(1+x^3)(1+x+x^2+x^3+x^4+x^5+x^6+x^7)$ from $\{\hat{p}_1,\hat{p}_2\}=\{3,8\}$, and $g(x)=(1+x^2)(1+x+x^2)(1+x+x^2+x^3+x^4)$ from $\{\hat{p}_1,\hat{p}_2\}=\{6,5\}$.

TABLE II
UNSUITABLE ORDER PAIRS COMPUTED FROM ALL POSSIBLE EVEN/ODD-NUMBERED LENGTH PAIRS

$[l_{e_1},l_{e_2}]$	Unsuitable order pairs	$[l_{o_1},l_{o_2}]$	Unsuitable order pairs
[4, 2]	$\{3,1\}$	[3, 1]	$\{2,1\}$
[6, 2]	$\{4,2\}$	[5, 1]	$\{3,2\}$
[8, 2]	$\{5,3\}$	[7, 1]	$\{4,3\}$
[10,2]	$\{6,4\}$	[9, 1]	$\{5,4\}$
[6, 4]	$\{5,1\}$	[5, 3]	$\{4,1\}$
[8, 4]	$\{6,2\}$	[7, 3]	$\{5,2\}$
[10,4]	$\{7,3\}$	[9, 3]	$\{6,3\}$
[8, 6]	$\{7,1\}$	[7, 5]	$\{6,1\}$
[10,6]	$\{8,2\}$	[9, 5]	$\{7,2\}$
[10,8]	$\{9,1\}$	[9, 7]	$\{8,1\}$

It turns out that among all possible generator polynomials, including the four generator polynomials given above, the lowest-degree generator polynomials are $g(x)=(1+x^2)(1+x^3+x^6)$ and $g(x)=(1+x^2)(1+x+x^2)(1+x+x^2+x^3+x^4)$. We note that these degree-8 generator polynomials have also been obtained through the search procedure. The orders of the generator polynomials $g(x)=1+x^2+x^3+x^5+x^6+x^8$ and

TABLE III
SYNDROME SETS PRODUCED BY $g(x) = 1 + x^3 + x^5 + x^8$ FOR THE TARGET ERROR PATTERNS OF LENGTHS UP TO 10

Target error polynomial	Syndrome set in decimals												Period P
	$s^{(0)}$	$s^{(1)}$	$s^{(2)}$	$s^{(3)}$	$s^{(4)}$	$s^{(5)}$	$\cdots$	$s^{(9)}$	$\cdots$	$s^{(14)}$	$\cdots$	$s^{(29)}$	$(s^{(\rho)} = s^{(\rho+P)})$
1	128	64	32	16	8	4	$\cdots$	74	$\cdots$	206	$\cdots$	41	30
$(1+x+x^2)$	224	112	56	28	14	7	$\cdots$	233					10
$(1+x)$	192	96	48	24	12	6	$\cdots$	111	$\cdots$	169			15
$(1+x)^3$	240	120	60	30	15	147	$\cdots$	170	$\cdots$	201			15
$(1+x+x^2+x^3+x^4)$	248	124	62	31	155	217							6
$(1+x)(1+x+x^2)^2$	252	126	63	139	209								5
$(1+x+x^3)(1+x^2+x^3)$	254	127	171	193	244	122	$\cdots$	182	$\cdots$	50	$\cdots$	213	30
$(1+x)^7$	255	235	225	228	114	57	$\cdots$	17	$\cdots$	215			15
$(1+x+x^2)(1+x^3+x^6)$	107	161	196	98	49	140	$\cdots$	214					10
$(1+x)(1+x+x^2+x^3+x^4)^2$	33	132	66										3

$g(x) = 1 + x^3 + x^5 + x^8$ are 18 and 30, respectively. The corresponding codes are $(18, 10)$ and $(30, 22)$ cyclic codes. Here, the $(30, 22)$ cyclic code based on $g(x) = 1 + x^3 + x^5 + x^8$ is chosen for further development.

Table III shows syndrome sets in decimal numbers for all the target error patterns, and the periods of the syndrome sets. A syndrome element $s^{(\rho)}$ in Table III is the syndrome for a target error polynomial starting at position ρ. The period P is the smallest integer such that $s^{(\rho)} = s^{(\rho+P)}$. Since the syndrome set for each target error pattern in the list eventually repeats itself as the feedback shift register content keeps shifting, only one syndrome $s^{(0)}$ in each syndrome set needs to be stored for recognizing the captured error pattern.

There still exist two issues at this point: one is that the rate of the code is only $\frac{22}{30} \simeq 0.73$, which is not high enough for high density disk drive applications, and secondly, the periods of many syndrome sets are smaller than the codeword length, meaning that while a computed syndrome $s^{(0)}$ points to a single occurrence of a dominant error pattern in the target list, the error position cannot be identified completely by the captured syndrome. Rather, the syndrome now points to a number of possible error starting positions within the codeword. The second issue can be resolved to a large extent using a channel reliability measure to determine the most likely position of the given error pattern. The first issue can be handled by a code extension method, as discussed next.

B. Extending the Code to Increase the Rate

Given a degree-r $g(x)$ of order p that is tailored to target error polynomials $e_i(x)$'s, $i \in \{1, \ldots, L\}$, consider the following polynomial:

$$g'(x) = g(x)p'(x) \tag{8}$$

where $p'(x)$ is a degree-m primitive polynomial which is not a factor of any $e_i(x)$'s, and the order of $p'(x)$ should be greater than p. Then, the order p' of $g'(x)$ is simply obtained as

$$p' = \mathrm{LCM}[p, 2^m - 1] \tag{9}$$

because p' is the smallest integer, such that both $g(x)$ and $p'(x)$ are factors of $(x^{p'} + 1)$. When $g'(x)$ is taken as the generator polynomial, a $(p', p' - r - m)$ error-pattern-correcting code is constructed, and the code rate R' becomes

$$R' = \frac{p' - r - m}{p'} = \frac{\mathrm{LCM}[p, 2^m - 1] - r - m}{\mathrm{LCM}[p, 2^m - 1]} \tag{10}$$

where $R' \to 1$ as $p' \gg r + m$.

A unique mapping between the syndrome sets and the target error patterns is preserved as with $g(x)$, but now that $g(x)$ is multiplied by $p'(x)$, the periods of new syndrome sets are increased by a factor equal to $(2^m - 1)$.

It is instructive to compare this code with the t-burst-error-correcting Fire code [26], based on the following generator polynomial:

$$g_f(x) = (x^{2t-1} + 1)p_f(x) \tag{11}$$

where $p_f(x)$ is a degree-m primitive polynomial, with m set greater than or equal to t. Given t, the least redundancy is obtained in the case of $m = t$. The order p_f of $g_f(x)$ is the LCM of $(2t - 1)$ and $(2^m - 1)$. As a result, a $(p_f, \, p_f - 2t + 1 - m)$ Fire code can be constructed, and the Fire code can algebraically correct any single burst-error of length t or less [26].

A difference between the two cyclic codes is in the construction methodologies of generator polynomials $g'(x)$ and $g_f(x)$. While the factor $(x^{2t-1} + 1)$ in $g_f(x)$ is determined based only on the maximum length t among target error bursts, in our approach, the fairly low-degree $g(x)$ in $g'(x)$ is tailored to single-error-pattern correction of the L target error polynomials.

• *Example 3:* For $g(x) = 1 + x^3 + x^5 + x^8$ constructed in Examples 1 and 2, any primitive polynomial which is not a factor of the given 10 target error polynomials can be used as $p'(x)$, e.g., $(1+x^2+x^5)$, $(1+x+x^6)$ and $(1+x^3+x^7)$. Here, we choose a degree-6 primitive polynomial $(1 + x + x^6)$ as our $p'(x)$. Accordingly, we obtain

$$\begin{aligned} g'(x) = g(x)p'(x) &= (1 + x^3 + x^5 + x^8)(1 + x + x^6) \\ &= 1 + x + x^3 + x^4 + x^5 + x^8 + x^{11} + x^{14} \end{aligned}$$

where the order of $g'(x)$ is 630, i.e., $\mathrm{LCM}[30, 2^6 - 1]$. Consequently, a $(630, 616)$ code, the rate of which is approximately 0.98, is constructed.

With $g'(x) = (1 + x^3 + x^5 + x^8)(1 + x + x^6)$, Table IV shows the extended syndrome set periods P''s for the given 10 target error polynomials. The period P' of each syndrome set is obtained by the LCM of the period P and the order $(2^6 - 1)$ of $p'(x)$.

It is seen that four target error pattern polynomials (making up 64.924% of the overall error patterns) 1, $(1 + x + x^2)$, $(1 + x + x^3)(1 + x^2 + x^3)$ and $(1 + x + x^2)(1 + x^3 + x^6)$ can be algebraically corrected without any miscorrection. This is because the periods of the corresponding syndrome sets are equal to the codeword length, and a syndrome element in each

TABLE IV
THE EXTENDED PERIODS P''S OF SYNDROME SETS PRODUCED BY $g'(x) = (1 + x^3 + x^5 + x^8)(1 + x + x^6)$

Target error polynomial	Period P	Period P'
1	30	630
$(1+x+x^2)$	10	630
$(1+x)$	15	315
$(1+x)^3$	15	315
$(1+x+x^2+x^3+x^4)$	6	126
$(1+x)(1+x+x^2)^2$	5	315
$(1+x+x^3)(1+x^2+x^3)$	30	630
$(1+x)^7$	15	315
$(1+x+x^2)(1+x^3+x^6)$	10	630
$(1+x)(1+x+x^2+x^3+x^4)^2$	3	63

syndrome set indicates the exact error starting position. While the syndrome sets for the remaining six target error polynomials do not have periods equal to the codeword length, there are only a few possible error starting positions, and the extended periods of the syndrome sets are always greater than the lengths of the target error polynomials.

Note that the number of possible positions can further be reduced by a simple bit-polarity check of the detected codeword based on the already determined error pattern, e.g., given an error pattern $\pm[2, -2, 2, -2, 2, -2, 2, -2, 2, -2]$, the detected binary codeword at each starting position must be either $[1, 0, 1, 0, 1, 0, 1, 0, 1, 0]$ or $[0, 1, 0, 1, 0, 1, 0, 1, 0, 1]$.

C. Detection Capability for Multiple Error Occurrences

The proposed (n, k)[2] error-pattern-correcting cyclic code has been constructed such that

$$e_i(x) + x^\rho e_j(x) \neq g'(x)Q(x)$$

for any $0 \leq \rho \leq n-1$ and $i \neq j \in \{1, \ldots, L\}$. This property naturally guarantees distinct syndrome sets for L target error polynomials $e_i(x)$'s, i.e.,

$$s_i(x) \neq s_j^{(\rho)}(x) \tag{12}$$

for any ρ. The number of such distinct error-pattern pairs is $L(L-1)n$, because given $e_i(x)$ among L target error polynomials, there can be $(L-1)$ possible error polynomials for $e_j(x)$, and the number of possible starting positions for $e_j(x)$ is n.

However, it is possible that double error-pattern events in the form of $e_i(x) + x^\rho e_i(x) = e_i(x)(1 + x^\rho)$, $0 < \rho \leq n-1$, may not be detected, if ρ is a multiple of P_i' which is the extended period of syndrome set for $e_i(x)$. Among $L(n-1)$ such error events, there can be $\sum_{i=1}^{L} \lfloor \frac{n-1}{P_i'} \rfloor$ undetectable double error-pattern events. If $P_i' = n$, then $e_i(x)(1+x^\rho)$ is always detected for any $0 < \rho \leq n-1$, because ρ cannot be a multiple of $P_i' = n$. Thus, given a double-error-pattern event, the conditional probability of undetectable event $P_{u|2}$ is

$$P_{u|2} = \frac{\sum_{i=1}^{L} \left\lfloor \frac{n-1}{P_i'} \right\rfloor}{L(L-1)n + L(n-1)} = \frac{\sum_{i=1}^{L} \left\lfloor \frac{n-1}{P_i'} \right\rfloor}{nL^2 - L}. \tag{13}$$

[2]In this and the following sections, we reserve the symbols n and k to respectively denote the codeword length and input word length of the $(p', p' - r - m)$ error-pattern-correcting cyclic code.

Since the extended periods become closer to the codeword length n of the proposed code, the undetectable probability for a double-error-pattern event is small.

Now, we consider a triple-error-pattern event $e(x) = e_i(x) + x^\rho e_j(x) + x^\nu e_k(x)$, where $i, j, k \in \{1, \ldots, L\}$ and $0 \leq \rho \neq \nu \leq (n-1)$. The corresponding syndrome $s(x)$ is given by $s(x) = s_i(x) + s_j^{(\rho)}(x) + s_k^{(\nu)}(x)$. It has been shown that different GCDs $c_i(x)$'s between $e_i(x)$'s and $g(x)$ ensure distinct syndrome sets $\mathbf{S}_i$'s. Based on this property, if $c_i(x) \neq c_j(x) \neq c_k(x)$, then this $e(x)$ is always detected, because $s(x) = s_i(x) + s_j^{(\rho)}(x) + s_k^{(\nu)}(x) \neq 0$ or

$$s_i(x) + s_j^{(\rho)}(x) \neq s_k^{(\nu)}(x). \tag{14}$$

Thus, it is possible to have an undetectable triple-error-pattern event $e(x)$, only if at least two GCDs are the same, i.e., $s_i(x) + s_j^{(\rho)}(x) = s_k^{(\nu)}(x)$. The number of such cases can be found through an exhaustive search.

Here, instead of doing the exhaustive search, we investigate an upper bound for the probability of such cases. The number of syndrome combinations $s_i(x) + s_j^{(\rho)}(x)$ is $nL^2 - L$, as shown in the denominator of (13), but there can be at most $\sum_k \frac{n}{P_k'}$ cases, where $c_k(x) = c_i(x)$ (or $c_j(x)$), such that $s_i(x) + s_j^{(\rho)}(x) = s_k^{(\nu)}(x)$, $k \in \{1, \ldots, L\}$. This is because given $s_k^{(\nu)}(x)$, there are $\frac{n}{P_k'} - 1$ additional positions to produce $s_k^{(\nu)}(x)$, i.e., the number of possible ν's is $\frac{n}{P_k'}$. Accordingly, the conditional probability of an undetectable triple-error-pattern event $P_{u|3}$ can be upper-bounded by

$$P_{u|3} \leq \frac{\sum_k \frac{n}{P_k'}}{nL^2 - L} \tag{15}$$

where $k \in \{1, \ldots, L\}$ and $c_k(x) = c_i(x)$ (or $c_j(x)$). Again, the conditional probability $P_{u|3}$ turns out to be small.

For other multiple occurrences, e.g., quadruple and quintuple error patterns, similar approaches can be applied to obtain the probability of an undetected error. In sum, the proposed coding has a powerful detection capability for M-target-error-pattern events, $M \geq 2$, as well as any single error patterns. We also note that a large percentage of multiple-error-pattern events can be discriminated from a single error pattern, based on the captured syndrome. Since only L syndrome sets are assigned for identifying L single error patterns, multiple error occurrences can be recognized whenever the captured syndrome is not among the L syndrome sets.

III. List Decoding Based on a Set of Test Word-Error Events

A. List Decoder for Correcting Multiple Error Occurrences

While a significant portion of multiple-error-pattern events can be recognized by the captured syndrome, information on the individual error patterns contained in the multiple-error-pattern event as well as their starting positions are not clearly provided by algebraic information. The number of syndrome patterns that need to be stored can also get very large, resulting in increased implementation complexity. For efficiently correcting multiple-error-pattern events, a list decoding strategy based on a set of test word-error events (or simply test words) is developed. Each test word consists of most probable single target error patterns that are obtained based on "pattern-level" reliability measures. Here, the reliability computation is done in a parallel fashion by local error-pattern correlators, each of which is matched to a targeted, dominant error pattern. Then, a list of candidate codewords is constructed by running the simple single-pattern-correcting decoder in parallel for each test decoder input word. Finally, the most probable codeword is determined among the list of candidate codewords.

Suppose a received word polynomial $\hat{c}(x)$ at the channel detector output is contaminated by M target error patterns, which is expressed over GF(2) as

$$\hat{c}(x) = c(x) + \sum_{k=1}^{M} x^{\rho_k} e_{i_k}(x) \tag{16}$$

where $c(x)$ is the actual transmitted codeword polynomial, ρ_k is any non-negative integer, $e_{i_k}(x) \in \{e_1(x), \ldots, e_L(x)\}$, and $M \geq 2$. The corresponding syndrome $\hat{s}(x)$ is

$$\hat{s}(x) = \sum_{k=1}^{M} s_{i_k}^{(\rho_k)}(x). \tag{17}$$

Suppose that a test word-error, among a given set of test word-errors, is given by $\sum_{k=1}^{M-1} x^{\rho_k} e_{i_k}(x)$. Adding this to $\hat{c}(x)$, the corresponding test word $\tilde{c}(x)$ is obtained

$$\begin{aligned}\tilde{c}(x) &= \hat{c}(x) + \sum_{k=1}^{M-1} x^{\rho_k} e_{i_k}(x)\\ &= c(x) + x^{\rho_M} e_{i_M}(x)\end{aligned} \tag{18}$$

and the syndrome $\tilde{s}(x)$ corresponding to $\tilde{c}(x)$ is

$$\tilde{s}(x) = s_{i_M}^{(\rho_M)}(x). \tag{19}$$

With this syndrome $\tilde{s}(x)$, both the error pattern $e_{i_M}(x)$ and its possible error starting positions are simply identified. The most likely error starting position ρ_M can be estimated based on the reliability measures of the given possible error positions. The M-target-error-pattern event can then be corrected, utilizing the single-pattern correction capability. It is obvious that if a test word-error exactly corresponds to the actual M-error-pattern event $\sum_{k=1}^{M} x^{\rho_k} e_{i_k}(x)$, then the error event is also corrected.

Conceptually, if at least one test word is within the correctable range of the transmitted codeword, then an M-error-pattern event can be corrected. In other words, the single-pattern correction capability of the proposed code can be substantially extended based on a set of test words.

It is instructive to compare the present list decoding scheme to the generalized minimum distance decoding (GMD) [27] and the Chase decoding [28]. They all share the common property of correction based on running relatively simple decoders on a set of test input words that reflect some sort of perturbation of the original decoder input. However, while the test words in the GMD and Chase decoding schemes are constructed based on "bit-level" reliability measures of the original decoder input, the test words utilized in the proposed list decoder are based on the reliability measures of the local patterns in the same input word. In fact, we note that the proposed list decoder is remarkably effective when errors tend to occur in specific patterns. Another difference is that in our case, each decoder, while relying largely on syndrome-computation, is designed to utilize soft information from the channel detector as well.

Now, we describe the test word-error construction procedure. First, the most probable error starting positions for each $e_i(x)$, $i = 1, \ldots, L$, are identified, based on the corresponding local error-pattern correlator, which determines the likelihood of a given error pattern in the local region.[3] A threshold test on the output of the local error pattern correlator can determine how likely the pattern is in the given local regions, as described in Section III-C. If the local correlator output at a certain error starting position for $e_i(x)$ is greater than or equal to the threshold, then this error position and the error pattern are deemed probable; otherwise, they are discarded. It is possible for some $e_i(x)$'s that none of their possible error starting positions are considered probable.

By collecting all probable error starting positions for all $e_i(x)$'s, we obtain a set of probable single target error patterns associated with their starting positions

$$\mathbf{v} = \{x^{\rho_1} e_{i_1}(x), x^{\rho_2} e_{i_2}(x), \ldots, x^{\rho_\mu} e_{i_\mu}(x)\} \tag{20}$$

where $e_{i_1}(x), \ldots, e_{i_\mu}(x) \in \{e_1(x), \ldots, e_L(x)\}$, $\rho_1, \ldots, \rho_\mu \in \{0, \ldots, n-1\}$, and μ represents the total number of local patterns that have passed the threshold check. We note that the selected error patterns $e_{i_1}(x), \ldots, e_{i_\mu}(x)$ are not necessarily all distinct. Assume that the set $\mathbf{v}$ is ordered according to the probability estimate of the local patterns.

The next step is to construct a set of test word-errors, based on the μ potential single error patterns in $\mathbf{v}$. Let $\mathbf{w}_k$, $k = 1, 2, \ldots, \mu$, be a set of $\binom{\mu}{k}$ test words, each of which consists of k most probable single error patterns. Here, $\binom{\mu}{1}$ test words in $\mathbf{w}_1$, e.g., $x^{\rho_1} e_{i_1}(x)$, are generated to correct either single or double error patterns, and $\binom{\mu}{2}$ test words in $\mathbf{w}_2$, e.g., $x^{\rho_1} e_{i_1}(x) + x^{\rho_2} e_{i_2}(x)$, are for correcting either double or triple error patterns. Basically, $\mathbf{w}_{M-1}$ is formed for the purpose of correcting either $(M-1)$ or M target error patterns within the channel detector output word. Thus, for correction up to M error patterns, a set of $\sum_{k=1}^{M-1} \binom{\mu}{k}$ test word-errors is constructed as

$$\mathbf{w} = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{M-1}\}. \tag{21}$$

The overall complexity of the list decoder depends obviously on the parameter μ. The complexity can be traded with performance by increasing the threshold levels for the error patterns, which would decrease μ. For a given μ, complexity can be reduced further by taking only a more probable subset of each $\mathbf{w}_k$, $k = 1, 2, \ldots, M-1$.

It is also possible that a small reliable set of test word-errors can be constructed based on the probability estimates of the

[3]The local error-pattern correlator will be described in detail in the next subsection. In this paper, the correlator output is utilized as a reliability measure (a log-probability-ratio) of $x^\rho e_i(x)$, where ρ is any possible error starting position.

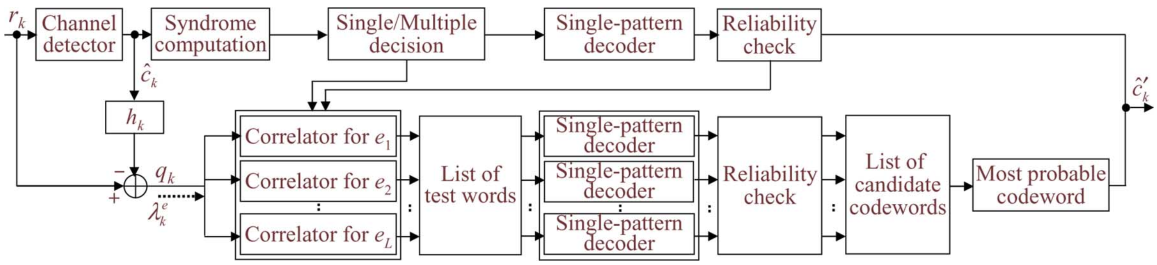

Fig. 1. The list decoder based on a set of test word-error events.

potential word errors. Let $P_i^{(\rho)}$ be a probability measure[4] of the local error pattern $x^\rho e_i(x)$. Then, given the μ patterns $x^{\rho_1}e_{i_1}(x), x^{\rho_2}e_{i_2}(x), \ldots, x^{\rho_\mu}e_{i_\mu}(x)$ in $\mathbf{v}$, we have a probability-measure set

$$\left\{P_{i_1}^{(\rho_1)}, P_{i_2}^{(\rho_2)}, \ldots, P_{i_\mu}^{(\rho_\mu)}\right\}. \tag{22}$$

The probability of a test word-error that contains k local error patterns can then be estimated by multiplying the corresponding k local probability measures, assuming the constituent single error patterns are independent of one another. Given some initial selection of the test word-errors and their corresponding probability estimates, we sort the probability estimates in descending order. Then, by choosing the first ω test word errors, a set of ω test words is obtained.

Fig. 1 shows the block diagram of the overall decoder. The captured syndrome is first examined to determine whether the channel detector output word contains no error, a single error pattern or multiple error patterns. If the captured syndrome indicates the presence of a single dominant pattern, then correction is attempted. If the correction yields the zero syndrome, a reliability check is further applied by testing whether the corresponding error-pattern correlator output gives rise to a high enough output value. If so, the corrected word is finally released as the decision. If not, it is assumed that the original channel output contains multiple error patterns. If the channel output is deemed to contain multiple error patterns, due either to the failure in the final threshold reliability check at the single error-pattern correcting decoder output or to the early indication based on the initial captured syndrome, the parallel local pattern-correlators get activated and start to run on the channel error signal constructed by subtracting the estimated noisy-free channel output from the actual channel output signal (see Fig. 1). The correlator outputs allow construction of the test words that represent perturbations of the original channel detector output word. The perturbed test words are basically constructed by flipping local patterns that are deemed potentially corrupted by dominant error patterns.

A single-pattern-correcting decoder acting on each test word attempts to correct a single dominant error pattern, and if successful, the resulting syndrome becomes zero. If the syndrome becomes zero after correction, then again a threshold-based reliability test is done further to ensure the correction is reliable enough. If the zero syndrome is produced and the subsequent reliability check (specific to the corrected pattern) is satisfied, then the decoder output becomes a candidate codeword; otherwise, the test word is simply discarded. It is possible that a certain test word itself yields the zero syndrome without correction, if it matches the actual codeword. In such a case, the reliability check is not required because the local patterns in that test word are guaranteed to be highly reliable. The final decision is to choose the most probable codeword among potentially many candidate codewords.

Let $\hat{\mathbf{c}}'_\zeta$ and $R(\hat{\mathbf{c}}'_\zeta)$, $\zeta = 1, \ldots, \Omega$, be a candidate codeword and its reliability measure, respectively. The reliability measure $R(\hat{\mathbf{c}}'_\zeta)$ of $\hat{\mathbf{c}}'_\zeta$ is obtained by summing the correlator-based local reliability measures of the constituent single error patterns in $\hat{\mathbf{c}}'_\zeta$. Assuming that $\hat{\mathbf{c}}'_\zeta$ potentially represents an M-error-pattern event $\sum_{k=1}^{M} x^{\rho_k}e_{i_k}(x)$, the reliability measure $R(\hat{\mathbf{c}}'_\zeta)$ is given by

$$R(\hat{\mathbf{c}}'_\zeta) = \sum_{k=1}^{M} C\left(e_{i_k}^{(\rho_k)}\right) \tag{23}$$

where $C(e_{i_k}^{(\rho_k)})$ represents the local error-pattern correlator output for $x^{\rho_k}e_{i_k}(x)$, as given in (35).

Finally, given the reliability measures (or probability measures) of the Ω candidate codewords, the most probable codeword $\hat{\mathbf{c}}'$ is determined by

$$\hat{\mathbf{c}}' = \arg\max_{\hat{\mathbf{c}}'_\zeta} R(\hat{\mathbf{c}}'_\zeta) \tag{24}$$

where $\zeta \in \{1, \ldots, \Omega\}$. We note that if there is no candidate codeword, then the decoder simply releases the channel detector output word.

B. Computation of Correlator-Based Reliability Measures

The reliability measures are computed by the local error-pattern correlators matched to L target error patterns, as shown in Fig. 2. Let r_k be the channel detector input sequence, given by $r_k = c_k * h_k + n_k$, where c_k is the bipolar version of the transmitted codeword sequence, h_k is the ISI channel response of length l_h, and n_k is the zero-mean additive white Gaussian noise (AWGN) sequence with variance σ^2. Also let q_k be the difference between r_k and the convolution of the received codeword sequence $\hat{c}_k$ and h_k, i.e., $q_k = r_k - (\hat{c}_k * h_k) = (c_k - \hat{c}_k) * h_k + n_k$. We assume that a length-l_i target error pattern $e_i(x)$, $i \in \{1, \ldots, L\}$, occurs at positions from $k = \rho_i$ to

[4]The probability measure is based on the local *a posteriori* probability for a single error pattern, given in (27). The local *a posteriori* probability can be provided by the local error-pattern correlator.

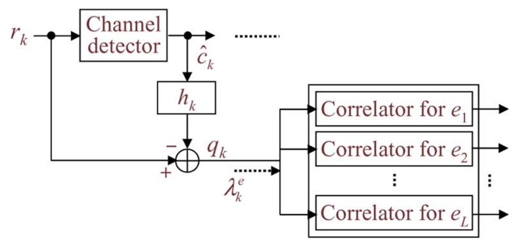

Fig. 2. Local error-pattern correlators matched to L target error patterns.

$k=\rho_i+l_i-1$. Denote this local error pattern as $[\mathbf{e}^{(i)}]_{\rho_i}^{\rho_i+l_i-1}$. Then, q_k can be written as

$$\begin{aligned} q_k &= (c_k-\hat{c}_k)*h_k+n_k \\ &= [\mathbf{e}^{(i)}]_{\rho_i}^{\rho_i+l_i-1}*h_k+n_k \\ &= [\boldsymbol{\epsilon}^{(i)}]_{\rho_i}^{\rho_i+l_i^h}+n_k \end{aligned} \tag{25}$$

where $[\boldsymbol{\epsilon}^{(i)}]_{\rho_i}^{\rho_i+l_i^h} \triangleq [\mathbf{e}^{(i)}]_{\rho_i}^{\rho_i+l_i-1}*h_k$, and $l_i^h \triangleq l_i+l_h-2$.

The probability measure for each possible error starting position ρ of $e_i(x)$, $\rho\in\{0,\ldots,\rho_i,\ldots,n-1\}$, can be computed by the following "local" *a posteriori* probability:

$$\begin{aligned} P_i^{(\rho)} &\triangleq P\left(\left[\boldsymbol{\epsilon}^{(i)}\right]_\rho^{\rho+l_i^h}\Big|[\mathbf{r}]_\rho^{\rho+l_i^h},[\hat{\mathbf{c}}]_\rho^{\rho+l_i^h}\right) \\ &= P\left(\left[\epsilon^{(i)}\right]_\rho^{\rho+l_i^h}\Big|[\mathbf{q}]_\rho^{\rho+l_i^h}\right) \end{aligned} \tag{26}$$

where $[\mathbf{q}]_\rho^{\rho+l_i^h} = [\boldsymbol{\epsilon}^{(i)}+\mathbf{n}]_\rho^{\rho+l_i^h}$. Using Bayes' Theorem, the local *a posteriori* probability can be rewritten as

$$\begin{aligned} &P\left(\left[\boldsymbol{\epsilon}^{(i)}\right]_\rho^{\rho+l_i^h}\Big|[\mathbf{q}]_\rho^{\rho+l_i^h}\right) \\ &\quad= \frac{P\left([\mathbf{q}]_\rho^{\rho+l_i^h}\Big|\left[\boldsymbol{\epsilon}^{(i)}\right]_\rho^{\rho+l_i^h}\right)P\left(\left[\boldsymbol{\epsilon}^{(i)}\right]_\rho^{\rho+l_i^h}\right)}{P\left([\mathbf{q}]_\rho^{\rho+l_i^h}\right)}. \end{aligned} \tag{27}$$

Here, $P([\mathbf{q}]_\rho^{\rho+l_i^h})$ can be approximated by

$$\begin{aligned} P\left([\mathbf{q}]_\rho^{\rho+l_i^h}\right) &\simeq P\left([\mathbf{q}]_\rho^{\rho+l_i^h}\Big|\left[\boldsymbol{\epsilon}^{(i)}\right]_\rho^{\rho+l_i^h}\right)P\left(\left[\boldsymbol{\epsilon}^{(i)}\right]_\rho^{\rho+l_i^h}\right) \\ &\quad+P\left([\mathbf{q}]_\rho^{\rho+l_i^h}\Big|\left[\tilde{\boldsymbol{\epsilon}}^{(i)}\right]_\rho^{\rho+l_i^h}\right)P\left(\left[\tilde{\boldsymbol{\epsilon}}^{(i)}\right]_\rho^{\rho+l_i^h}\right) \end{aligned} \tag{28}$$

where $[\tilde{\boldsymbol{\epsilon}}^{(i)}]_\rho^{\rho+l_i^h}$ is the most probable competing error pattern and, by definition, this is the all-zero pattern, i.e., the most probable competing local pattern is that associated with the channel detector output sequence (maximum likelihood word) itself.

Since q_k is a zero-mean statistically independent Gaussian random variable with variance σ^2, the likelihood functions $P([\mathbf{q}]_\rho^{\rho+l_i^h}|[\boldsymbol{\epsilon}^{(i)}]_\rho^{\rho+l_i^h})$ and $P([\mathbf{q}]_\rho^{\rho+l_i^h}|[\tilde{\boldsymbol{\epsilon}}^{(i)}]_\rho^{\rho+l_i^h})$ are given by

$$\begin{aligned} P\left([\mathbf{q}]_\rho^{\rho+l_i^h}\Big|\left[\boldsymbol{\epsilon}^{(i)}\right]_\rho^{\rho+l_i^h}\right) &= \frac{1}{\sqrt{2\pi\sigma^2}}\,e^{-\frac{1}{2\sigma^2}\sum_{k=\rho}^{\rho+l_i^h}(q_k-\epsilon_k^{(i)})^2} \\ P\left([\mathbf{q}]_\rho^{\rho+l_i^h}\Big|\left[\tilde{\boldsymbol{\epsilon}}^{(i)}\right]_\rho^{\rho+l_i^h}\right) &= \frac{1}{\sqrt{2\pi\sigma^2}}\,e^{-\frac{1}{2\sigma^2}\sum_{k=\rho}^{\rho+l_i^h}q_k^2} \end{aligned} \tag{29}$$

and the probabilities $P([\boldsymbol{\epsilon}^{(i)}]_\rho^{\rho+l_i^h})$ and $P([\tilde{\boldsymbol{\epsilon}}^{(i)}]_\rho^{\rho+l_i^h})$ are

$$\begin{aligned} P\left(\left[\boldsymbol{\epsilon}^{(i)}\right]_\rho^{\rho+l_i^h}\right) &= P\left(\left[\mathbf{e}^{(i)}\right]_\rho^{\rho+l_i-1}\right) \\ &= \prod_{k=0,\,e_k^{(i)}\neq 0}^{l_i-1}P(c_{\rho+k}=-\hat{c}_{\rho+k}) \\ P\left(\left[\tilde{\boldsymbol{\epsilon}}^{(i)}\right]_\rho^{\rho+l_i^h}\right) &= \prod_{k=0,\,e_k^{(i)}\neq 0}^{l_i-1}P\left(c_{\rho+k}=\hat{c}_{\rho+k}\right). \end{aligned} \tag{30}$$

Note that the length-l_i error pattern sequence $[\mathbf{e}^{(i)}]_\rho^{\rho+l_i-1}$ does not necessarily have l_i erroneous bits in general.

For an equally likely c_k, i.e., $P(c_k=\pm1)=\frac{1}{2}$, we have $P([\boldsymbol{\epsilon}^{(i)}]_\rho^{\rho+l_i^h})=P([\tilde{\boldsymbol{\epsilon}}^{(i)}]_\rho^{\rho+l_i^h})$. However, if the interleaved extrinsic information λ_k^e of an outer detector/decoder is available and serves as the *a priori* probability estimate, i.e.

$$\begin{aligned} P(c_k=+1) &= \frac{e^{\lambda_k^e}}{1+e^{\lambda_k^e}} \quad \text{and} \\ P(c_k=-1) &= \frac{1}{1+e^{\lambda_k^e}} \end{aligned} \tag{31}$$

for $\lambda_k^e=\log\frac{P(c_k=+1)}{P(c_k=-1)}$, then $P([\boldsymbol{\epsilon}^{(i)}]_\rho^{\rho+l_i^h})\neq P([\tilde{\boldsymbol{\epsilon}}^{(i)}]_\rho^{\rho+l_i^h})$ in general.

By substituting (29) and (30) into (27), the local *a posteriori* probability $P([\boldsymbol{\epsilon}^{(i)}]_\rho^{\rho+l_i^h}|[\mathbf{q}]_\rho^{\rho+l_i^h})$ can be expressed as

$$\begin{aligned} &P\left(\left[\boldsymbol{\epsilon}^{(i)}\right]_\rho^{\rho+l_i^h}\Big|[\mathbf{q}]_\rho^{\rho+l_i^h}\right) \\ &\quad= \frac{1}{1+\exp\left[-\frac{1}{2\sigma^2}D_i^{(\rho)}\right]\cdot A_i^{(\rho)}} \\ &\quad= \frac{1}{1+\exp\left[-\left\{\frac{1}{2\sigma^2}D_i^{(\rho)}-\log A_i^{(\rho)}\right\}\right]} \end{aligned} \tag{32}$$

where $D_i^{(\rho)}$ and $A_i^{(\rho)}$ are defined by

$$D_i^{(\rho)} \triangleq \sum_{k=\rho}^{\rho+l_i^h}\left[q_k^2-\left(q_k-\epsilon_k^{(i)}\right)^2\right] \tag{33}$$

$$A_i^{(\rho)} \triangleq \frac{P\left(\left[\tilde{\boldsymbol{\epsilon}}^{(i)}\right]_\rho^{\rho+l_i^h}\right)}{P\left(\left[\boldsymbol{\epsilon}^{(i)}\right]_\rho^{\rho+l_i^h}\right)}. \tag{34}$$

Equation (33) represents the "local" error-pattern correlator output in the sense that it essentially describes the correlator operation between q_k and the channel output version of the dominant error pattern $e_i(x)$ within the local region $[\rho,\rho+l_i^h]$. Equation (32) immediately reveals that the log-probability-ratio representation of $P_i^{(\rho)}=P([\boldsymbol{\epsilon}^{(i)}]_\rho^{\rho+l_i^h}|[\mathbf{q}]_\rho^{\rho+l_i^h})$ is given by

$$C\left(e_i^{(\rho)}\right) \triangleq \frac{1}{2\sigma^2}D_i^{(\rho)}-\log A_i^{(\rho)}. \tag{35}$$

C. Threshold-Based Reliability Check

When $[\mathbf{q}]_\rho^{\rho+l_i^h} = [\boldsymbol{\epsilon}^{(i)} + \mathbf{n}]_\rho^{\rho+l_i^h}$, the above correlator yields

$$C\left(e_i^{(\rho)}\right) = \frac{1}{2\sigma^2}\sum_{k=\rho}^{\rho+l_i^h}\left[\left(\epsilon_k^{(i)}\right)^2 + 2\epsilon_k^{(i)} n_k\right] - \log A_i^{(\rho)}$$
$$= \frac{1}{2\sigma^2}\left[E^{(i)} + \sum_{k=\rho}^{\rho+l_i^h} 2\epsilon_k^{(i)} n_k\right] - \log A_i^{(\rho)} \quad (36)$$

where $E^{(i)} \triangleq \|[\boldsymbol{\epsilon}^{(i)}]_\rho^{\rho+l_i^h}|^2$. The maximum likelihood sequence detection (MLSD) channel detector output produces the error pattern $[\mathbf{e}^{(i)}]_\rho^{\rho+l_i-1}$, because this pattern is considered to be the closest to the channel detector input sequence in terms of the Euclidean metric, i.e.

$$\sum_{k=\rho}^{\rho+l_i^h}\left(\epsilon_k^{(i)} + n_k\right)^2 \le \sum_{k=\rho}^{\rho+l_i^h} n_k^2. \quad (37)$$

By rearranging (37), we see $E^{(i)} + \sum_{k=\rho}^{\rho+l_i^h} 2\epsilon_k^{(i)} n_k \le 0$, and consequently, $C(e_i^{(\rho)}) \le -\log A_i^{(\rho)}$.

If the channel detector determines the correct codeword at positions from ρ to $\rho + l_i - 1$, i.e., $[\mathbf{q}]_\rho^{\rho+l_i^h} = [\mathbf{n}]_\rho^{\rho+l_i^h}$, then the correlator gives

$$C\left(e_i^{(\rho)}\right) = \frac{1}{2\sigma^2}\sum_{k=\rho}^{\rho+l_i^h}\left[-\left(\epsilon_k^{(i)}\right)^2 + 2\epsilon_k^{(i)} n_k\right] - \log A_i^{(\rho)}$$
$$= \frac{1}{2\sigma^2}\left[-E^{(i)} + \sum_{k=\rho}^{\rho+l_i^h} 2\epsilon_k^{(i)} n_k\right] - \log A_i^{(\rho)}. \quad (38)$$

The corresponding Euclidean metric relationship is

$$\sum_{k=\rho}^{\rho+l_i^h}\left(\epsilon_k^{(i)} + n_k\right)^2 \ge \sum_{k=\rho}^{\rho+l_i^h} n_k^2. \quad (39)$$

This implies that $-E^{(i)} + \sum_{k=\rho}^{\rho+l_i^h} 2\epsilon_k^{(i)} n_k \ge -2E^{(i)}$ and $C(e_i^{(\rho)}) \ge -\frac{E^{(i)}}{\sigma^2} - \log A_i^{(\rho)}$.
Hence, the correlator-based reliability measure is bounded by

$$C\left(e_i^{(\rho)}\right) \le -\log A_i^{(\rho)} \quad \text{for } [\mathbf{q}]_\rho^{\rho+l_i^h} = \left[\boldsymbol{\epsilon}^{(i)} + \mathbf{n}\right]_\rho^{\rho+l_i^h}$$
$$C\left(e_i^{(\rho)}\right) \ge -\frac{E^{(i)}}{\sigma^2} - \log A_i^{(\rho)} \quad \text{for } [\mathbf{q}]_\rho^{\rho+l_i^h} = [\mathbf{n}]_\rho^{\rho+l_i^h}. \quad (40)$$

For equally-likely c_k's, i.e., $\log A_i^{(\rho)} = 0$, we consider $x^\rho e_i(x)$, $i \in \{1,\ldots,L\}$ and $\rho \in \{0,\ldots,n-1\}$, as a probable error pattern if the corresponding reliability measure $C(e_i^{(\rho)})$ satisfies the following threshold-based reliability check:

$$C\left(e_i^{(\rho)}\right) \ge -\frac{E^{(i)}}{2\sigma^2}. \quad (41)$$

IV. Performance Evaluation

A. Error Rate Analysis for the MLSD Based on the Local Error-Pattern Correlator

The reason that errors tend to occur in specific burst or cluster patterns in ISI channels is because of the small energies associated with these pattern at the channel output. The BER analysis for the MLSD, primarily depending on the Euclidean distance between any two distinguishable signal sequences at the channel detector output, is well known [29]–[32]. In starting an error analysis for our proposed coding scheme, here we set out to quickly revisit the error analysis of the MLSD from the viewpoint of the "local error-pattern correlator."

Given a target error pattern $[\mathbf{e}^{(i)}]_\rho^{\rho+l_i-1}$, its occurrence probability can be written as

$$P\left(\left[\mathbf{e}^{(i)}\right]_\rho^{\rho+l_i-1}, \left[\mathbf{c}^{(i)}\right]_\rho^{\rho+l_i-1}\right)$$
$$= P\left(\left[\mathbf{e}^{(i)}\right]_\rho^{\rho+l_i-1} \middle| [\mathbf{c}^{(i)}]_\rho^{\rho+l_i-1}\right) P\left(\left[\mathbf{c}^{(i)}\right]_\rho^{\rho+l_i-1}\right). \quad (42)$$

Here, $[\mathbf{c}^{(i)}]_\rho^{\rho+l_i-1}$ corresponds to the actual codeword symbols in the local region: $[c_\rho,\ldots,c_{\rho+l_i-1}]$. Note that $[\mathbf{e}^{(i)}]_\rho^{\rho+l_i-1} = [\mathbf{c} - \hat{\mathbf{c}}]_\rho^{\rho+l_i-1}$. As an example, if a length-3 target error pattern $\pm[2,-2,2]$ occurs at positions from ρ to $\rho+2$, i.e., $[\mathbf{e}^{(3)}]_\rho^{\rho+2} = \pm[2,-2,2]$, then the transmitted codeword $[\mathbf{c}]_\rho^{\rho+2}$ must be either $[1,-1,1]$ or $[-1,1,-1]$ in the bipolar representation.

The substitution of $[\mathbf{e}^{(i)}]_\rho^{\rho+l_i-1}$ into (25) yields $[\mathbf{q}]_\rho^{\rho+l_i^h} = [\boldsymbol{\epsilon}^{(i)} + \mathbf{n}]_\rho^{\rho+l_i^h}$. Then, given $[\mathbf{q}]_\rho^{\rho+l_i^h}$, the conditional error-pattern probability $P([\mathbf{e}^{(i)}]_\rho^{\rho+l_i-1}|[\mathbf{c}^{(i)}]_\rho^{\rho+l_i-1})$ is obtained as

$$P\left(\left[\mathbf{e}^{(i)}\right]_\rho^{\rho+l_i-1} \middle| \left[\mathbf{c}^{(i)}\right]_\rho^{\rho+l_i-1}\right)$$
$$= P\left(\sum_{k=\rho}^{\rho+l_i^h} n_k^2 > \sum_{k=\rho}^{\rho+l_i^h} q_k^2\right)$$
$$= Q\left(\frac{\sqrt{E^{(i)}}}{2\sigma}\right) \quad (43)$$

where $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-\frac{t^2}{2}}dt$ for $x \ge 0$. When the energy $E^{(i)}$ is relatively small, the corresponding error pattern would occur more frequently at the channel detector output.

Since $[\mathbf{e}^{(i)}]_\rho^{\rho+l_i-1}$ has two types of sign information (starting with either the positive or the negative sign), i.e., $(-1)^s \times [2,-2,2]$, $s \in \{0,1\}$, the probability $P([\mathbf{c}^{(i)}]_\rho^{\rho+l_i-1})$ can be computed as

$$P\left([\mathbf{c}^{(i)}]_\rho^{\rho+l_i-1}\right) = \sum_{s=0}^{1}\prod_{\substack{k=0,\\ e_k^{(i)}\neq 0}}^{l_i-1} P\left(c_{\rho+k} = (-1)^s \cdot \frac{e_k^{(i)}}{2}\right). \quad (44)$$

When equally likely c_k's are assumed, the probability $P([\mathbf{c}^{(i)}]_\rho^{\rho+l_i-1})$ can be given, irrespective of the error starting position ρ, by

$$P\left(\left[\mathbf{c}^{(i)}\right]_\rho^{\rho+l_i-1}\right) = P\left(\mathbf{c}^{(i)}\right) = \frac{2}{2^{l_i-z_i}} = \frac{1}{2^{l_i-z_i-1}} \quad (45)$$

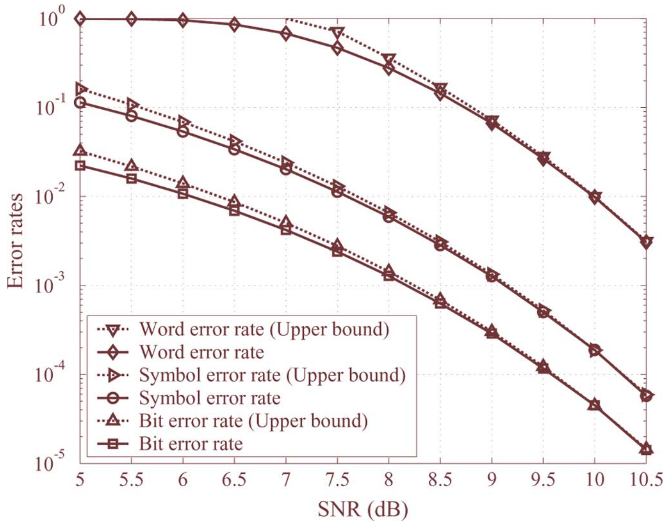

Fig. 3. Comparison of the upper bounds and simulation results for various error rates at the Viterbi detector output.

where z_i is the number of zeros in $[\mathbf{e}^{(i)}]_0^{l_i-1}$. We note that if an updated *a priori* probability for each codeword bit is available, corresponding probability $P([\mathbf{c}^{(i)}]_\rho^{\rho+l_i-1})$ depends on the starting position ρ.

When we consider only the L targeted, dominant error patterns, the BER P_b can then be bounded, using the union bound, as

$$\begin{aligned} P_b &\leq \sum_{i=1}^{L} (l_i - z_i) \cdot Q\left(\frac{\sqrt{E^{(i)}}}{2\sigma}\right) \cdot P\left(\mathbf{c}^{(i)}\right) \\ &= \sum_{i=1}^{L} \hat{P}_{b,i} \triangleq \hat{P}_b \end{aligned} \quad (46)$$

where $\hat{P}_{b,i} \triangleq (l_i - z_i) \cdot Q(\frac{\sqrt{E^{(i)}}}{2\sigma}) \cdot P(\mathbf{c}^{(i)})$.

In the above BER upper bound, a dominant error pattern $e_i(x)$ can start to occur anywhere within an n-bit codeword as well as at its boundary, i.e., $x^\rho e_i(x)$, $\rho \in \{0, \ldots, n-1\}$. Thus, the upper bound for the word error rate P_w is simply

$$P_w \leq \sum_{i=1}^{L} n \cdot Q\left(\frac{\sqrt{E^{(i)}}}{2\sigma}\right) \cdot P\left(\mathbf{c}^{(i)}\right) \triangleq \hat{P}_w. \quad (47)$$

It is possible to have $\hat{P}_w > 1$ for a low SNR. In that case, we just set $\hat{P}_w$ to 1.

Similarly, we can estimate the symbol error rate P_s. Let $l_{\max}$ be the maximum length among the lengths l_i's of the given L target error patterns $e_i(x)$'s. As in obtaining the upper bound of the word error rate, if the starting position ρ for $x^\rho e_i(x)$ is between 0 and $l_{\max} - 1$, then the corresponding symbol is in error. Hence, the upper bound for the symbol error rate P_s can also be expressed as

$$P_s \leq \sum_{i=1}^{L} l_{\max} \cdot Q\left(\frac{\sqrt{E^{(i)}}}{2\sigma}\right) \cdot P\left(\mathbf{c}^{(i)}\right) \triangleq \hat{P}_s. \quad (48)$$

Assuming a 630-bit word and a 10-bit symbol, Fig. 3 compares the upper bounds of the word error rate, the symbol error rate and the BER, i.e., $\hat{P}_w$, $\hat{P}_s$ and $\hat{P}_b$, as well as their simulated error rates P_w, P_s and P_b, for the channel response of $5 + 6D - D^3$. Ten dominant error patterns shown in Table I are used for the computation. The SNR is defined as the ratio of the energy of the channel response to the noise spectral density. We observe that the upper bounds naturally become tighter as the SNR increases.

B. *Error Rate Analysis for the Proposed List Decoder*

While the decoding error rates for conventional t-error-correcting codes are usually obtained by summing the probabilities of having τ errors, for $\tau > t$, within the received word, the error rate performance of the proposed M-error-pattern-correcting list decoder is computed by subtracting the correction probabilities for m-error-pattern events, where $1 \leq m \leq M$, from the overall error probability. Accordingly, we need to compute the occurrence probability of an m-error-pattern event, denoted by P_m.

For estimating P_m, we begin with the word error rate P_w. Whenever m error patterns occur within the channel detector output word, the word is obviously in error. Given the joint probability between the m-pattern error event and the word error event, $P_{m,w}$, we write

$$P_w = \sum_{m \geq 1} P_{m,w} = \sum_{m \geq 1} P_{m|w} P_w \quad (49)$$

where $P_{m|w}$ is the conditional probability of an m-error-pattern event in a given erroneous detector output word, and $\sum_{m \geq 1} P_{m|w} = 1$. In fact, the conditional probability $P_{m|w}$ is related to the symbol error rate P_s, where one symbol consists of $l_{\max}$ bits. While a long single error pattern often yields two symbol errors, as the single-pattern error occurs at a symbol boundary, We also note that sometimes two single-bit errors occur within one symbol. Overall, our empirical observation indicates the assumption that an m-error-pattern event yields m symbol errors gives numerically accurate results.

Now, let n_s be the number of symbols contained in the n-bit word, i.e., $n_s = \lceil \frac{n}{l_{\max}} \rceil$. Then, the conditional probability $P_{m|w}$ can be estimated, using the binomial distribution, as

$$P_{m|w} \simeq \frac{1}{m_0} \binom{n_s}{m} \hat{P}_s^m (1 - \hat{P}_s)^{n_s - m} \triangleq \hat{P}_{m|w} \quad (50)$$

where $m_0 = 1 - (1 - \hat{P}_s)^{n_s} = 1 - \binom{n_s}{0} \hat{P}_s^0 (1 - \hat{P}_s)^{n_s - 0}$, such that $\sum_{m \geq 1} \hat{P}_{m|w} = 1$. From $P_{m|w} \simeq \hat{P}_{m|w}$ in (50) and $P_w \leq \hat{P}_w$ in (47), the probability of an m-error-pattern event P_m is then bounded by

$$\begin{aligned} P_m &\leq \hat{P}_{m|w} \hat{P}_w \\ &= \frac{1}{m_0} \binom{n_s}{m} \hat{P}_s^m (1 - \hat{P}_s)^{n_s - m} \hat{P}_w \triangleq \hat{P}_m. \end{aligned} \quad (51)$$

Fig. 4 compares $\hat{P}_m$ with the captured statistics P_m through computer simulations for 630-bit words at various SNRs. It is seen that the difference between the two probabilities is indeed relatively small at reasonably high SNRs. In particular, the analytical probability estimate matches the captured statistics closely at the 9 dB SNR.

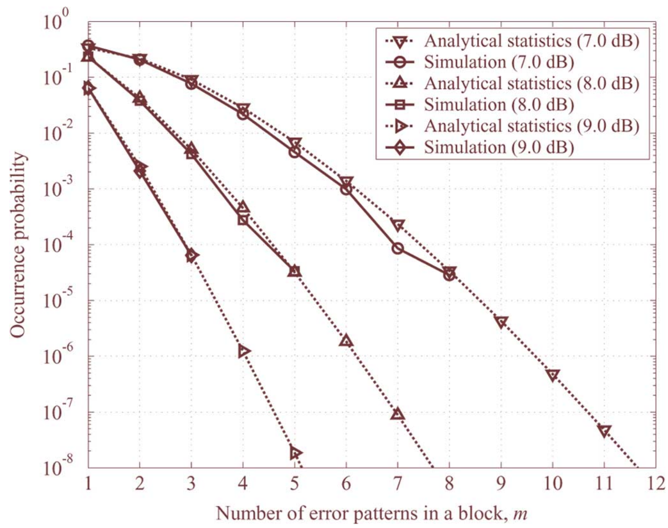

Fig. 4. Comparison of the analytical statistics and captured statistics for multiple error patterns within the channel detector output word.

Using $\sum_{m=1}^{n_s} \hat{P}_{m|w} = 1$, the BER estimate $\hat{P}_b$ in (46) can be written as

$$\hat{P}_b = \sum_{m=1}^{n_s} \hat{P}_{m|w} \hat{P}_b. \tag{52}$$

Here, $\hat{P}_{m|w}\hat{P}_b$ can be viewed as the fraction of the bit error probability due to an m-error-pattern event. Given $\hat{P}_{m|w}\hat{P}_b$, we let $\hat{P}_{c|m}$ be the conditional probability of correct decision for the m-error-pattern event. We also let $\hat{P}_{e|m}$ denote the conditional probability of erroneous decision corresponding to either releasing the detector output word or miscorrection. Then, $\hat{P}_b$ in (52) can be rewritten as

$$\hat{P}_b = \sum_{m=1}^{n_s} [\hat{P}_{c|m} + \hat{P}_{e|m}] \hat{P}_{m|w} \hat{P}_b \tag{53}$$

where $\hat{P}_{c|m} + \hat{P}_{e|m} = 1$.

We now compute $\hat{P}_{c|m}$ over an m-target-error-pattern event. Let us first consider $\hat{P}_{c|1,i}$, which is the correction probability for a single error pattern $x^\rho e_i(x)$. Given $e_i(x)$ and only a few possible error starting positions identified by the captured syndrome, correction is made by determining the maximum likelihood error starting position, based on the reliability measure. As long as the reliability measure at the actual error position satisfies the reliability check given in (41), the single error pattern can be corrected. Then, the conditional probability $\hat{P}_{c|1,i}$ is

$$\begin{aligned}\hat{P}_{c|1,i} &= P\left(C\left(e_i^{(\rho)}\right) \geq -\frac{E^{(i)}}{2\sigma^2} \,|x^\rho e_i(x) \right) \\ &= 1 - Q\left(\frac{2E^{(i)} - 2\sigma^2 \log A_i^{(\rho)}}{\sqrt{4E^{(i)}\sigma^2}} \right).\end{aligned} \tag{54}$$

With equally-likely codewords,

$$\hat{P}_{c|1,i} = 1 - Q\left(\frac{\sqrt{E^{(i)}}}{\sigma} \right). \tag{55}$$

Thus, given L target error patterns $e_i(x)$'s, $i = 1, \ldots, L$, the conditional probability of correct decision for a single-target-error-pattern event, $\hat{P}_{c|1}$ can be written as

$$\hat{P}_{c|1} = \sum_{i=1}^{L} \hat{P}_{c|1,i}. \tag{56}$$

In general, we consider an m-error-pattern event

$$e(x) = \sum_{k=1}^{m} x^{\rho_k} e_{i_k}(x).$$

According to the decoding process described toward the end of Section III-A, a correct decision is made if and only if all of the m local error patterns pass the threshold reliability test based on the matching local correlators. Consequently, the conditional probability of correct decision for this m-pattern event is given by

$$\hat{P}_{c|m} = \sum_{i_1=1}^{L} \cdots \sum_{i_m=1}^{L} \left[\prod_{k=1}^{m} \hat{P}_{c|1,i_k} \right] \tag{57}$$

where

$$\hat{P}_{c|1,i_k} = 1 - Q\left(\frac{\sqrt{E^{(i_k)}}}{\sigma} \right)$$

for equally-likely codewords. Since there are L target error patterns and the local patterns in the m-error-pattern event are not necessarily distinct, a total of L^m error events is considered in (57).

Therefore, when a list decoder aims at correcting up to M target error patterns within the channel detector output word, the resulting BER $\hat{P}_b^M$ can be expressed as

$$\begin{aligned}\hat{P}_b^M &= \hat{P}_b - \sum_{m=1}^{M} \hat{P}_{c|m} \hat{P}_{m|w} \hat{P}_b \\ &= \hat{P}_b \left[1 - \sum_{m=1}^{M} \hat{P}_{c|m} \hat{P}_{m|w} \right].\end{aligned} \tag{58}$$

Fig. 5 illustrates analytical BERs of the seven list decoders based on the $(630, 616)$ error-pattern-correcting cyclic codes, correcting up to M error patterns, where $M = 1, 2, \ldots, 7$, as well as the simulated BERs of the uncoded Viterbi detector. For the BER performance comparison, the variance of the AWGN is increased by a factor of R^{-1}, where R is a code rate of $\frac{616}{630}$. We observe that the list decoders certainly outperform the uncoded Viterbi detector. The return is also quickly diminishing as M increases.

C. Bit Error Rate Simulation Results

For 350 000 codewords transmitted on the channel with response $5 + 6D - D^3$, Table V illustrates the statistics for the number of error patterns within the Viterbi detector output word. In this section, we aim at correcting one to four error patterns. We present the BER simulation results of four list decoders based on the (630, 616) error-pattern-correcting code, as well as their theoretical performance limits $\hat{P}_b^M$'s, where $M = 1, 2, 3, 4$.

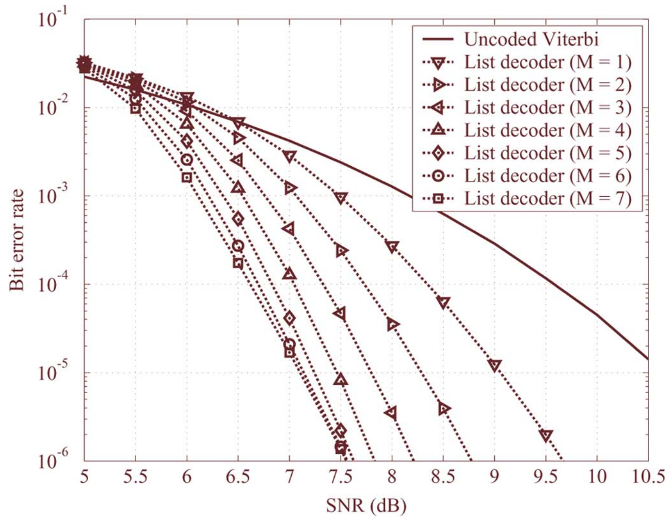

Fig. 5. Analytical BERs of the $(630, 616)$ code based list decoders correcting up to M target error patterns.

TABLE V
NUMBER OF ERROR PATTERNS WITHIN THE 630-BIT CODEWORDS AND CORRESPONDING WORD ERROR RATES AT THE VITERBI DETECTOR OUTPUT ($\mathrm{SNR} = 7.5$ dB, $\mathrm{BER} = 2.399 \times 10^{-3}$)

Number of error patterns within the codeword	Number of occurrences	Word error rate
Single error pattern	118,098	3.3742×10^{-1}
Double error patterns	37,097	1.0599×10^{-1}
Triple error patterns	7,507	2.1449×10^{-2}
Quadruple error patterns	1,136	3.2457×10^{-3}
5 error patterns	152	4.3429×10^{-4}
6 error patterns	15	4.2857×10^{-5}
Overall	164,005	4.6859×10^{-1}

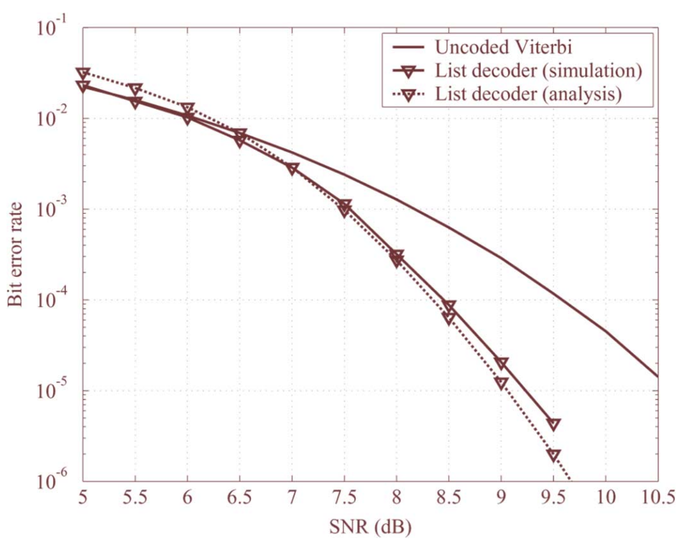

Fig. 6. Simulated BERs versus analytical BERs for the single-error-pattern-correcting list decoder.

1) Single-Error-Pattern Correction $(M = 1)$*:* Fig. 6 compares the simulated BERs of the single-error-pattern-correcting scheme and the analytical BER performance $\hat{P}_b^1$. It is seen that the simulated BER performance is lower-bounded by $\hat{P}_b^1$, as the actual errors may come from error patterns outside the targeted dominant pattern list. Here, there is no need for constructing a set of test words, since any single target error pattern is completely identified by the captured syndrome. As shown in Table IV, four target error patterns $\pm[2]$, $\pm[2,-2,2]$, $\pm[2,-2,2,-2,2,-2,2]$ and $\pm[2,-2,2,-2,2,-2,2,-2,2]$ can be algebraically corrected without any miscorrection because the captured syndrome precisely points to the actual error position. For the remaining six target error patterns, there are only a few possible error starting positions, and the maximum likelihood error position is determined based on the reliability measures of the possible error positions. Besides possible miscorrection for the targeted error patterns, error patterns outside the target set also produce errors, which make the performance sensitive at relatively high SNRs. The SNR gain of the single-pattern-correcting scheme, over the uncoded Viterbi detector, is seen to be around 1.14 dB at BER $= 10^{-4}$.

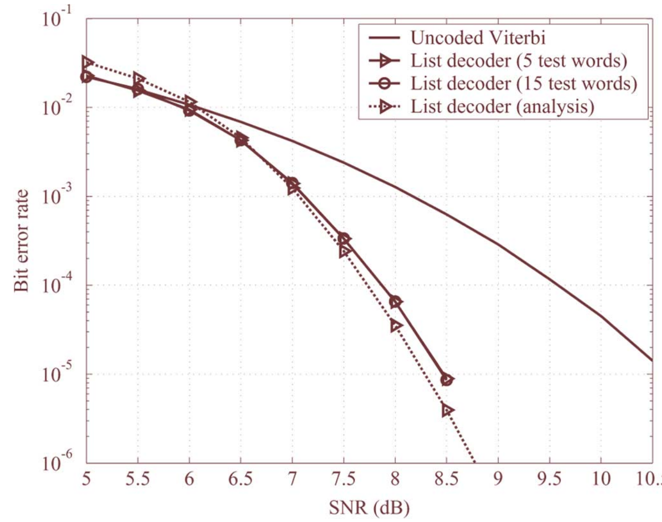

Fig. 7. Simulated BERs versus analytical BERs for the double-error-pattern-correcting list decoder.

2) Double-Error-Pattern Correction $(M = 2)$*:* For correcting double-error-pattern events, test words are constructed by inserting single dominant patterns into the detector output word. For comparison, we consider two sets of test words: one consisting of only five test words, and the other of 15 test words. The simulated BERs of the two double-pattern-correcting list decoders are shown in Fig. 7. The analytical lower bound is also shown. While the list decoder based on 15 test words is expected to perform better than the 5-test-word based list decoder, we observe from the simulation results that the two decoders perform very similarly: the SNR gains are respectively seen to be 1.7 and 1.71 dB for the five- and 15-test-word based list decoders, at $\mathrm{BER} = 10^{-4}$. The results indicate that the probabilities of capturing at least one test word within the correctable range are almost identical for the two sets of test words. However, when a smaller set of test words, e.g., two or three test words, was used, then the performance gain was visibly reduced. Accordingly, 5 test words (5 most probable single error patterns) represent a good choice in this case.

3) Triple-Error-Pattern Correction $(M = 3)$*:* Again, we consider various sets of test words. Here, in addition to five single-pattern test words, we construct $\binom{\mu_2}{2}$ test words based

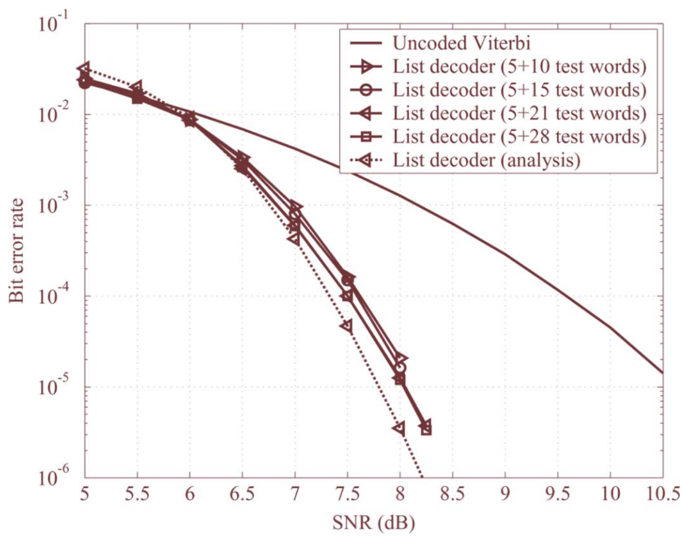

Fig. 8. Simulated BERs versus analytical BERs for the triple-error-pattern-correcting list decoder.

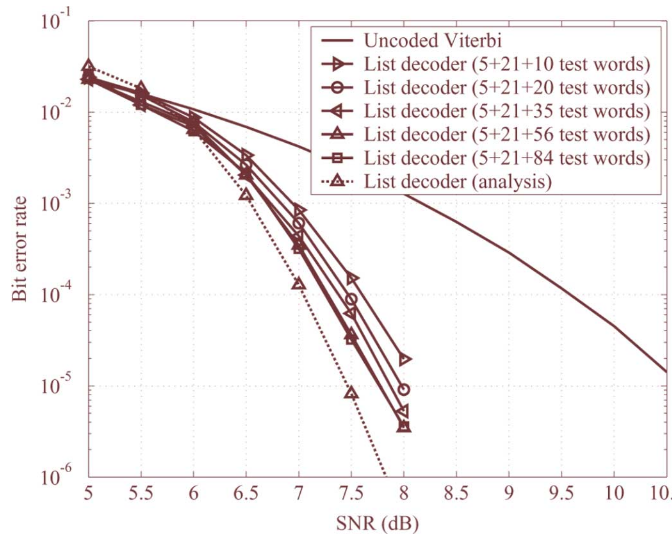

Fig. 9. Simulated BERs versus analytical BERs for the quadruple-error-pattern-correcting list decoder.

on inserting two patterns in the detector output, where μ_2 represents the number of probable local patterns considered in creating the test words involving double-pattern insertion. In general, μ_k can be made smaller than μ in an effort to control complexity.

Here we consider four cases: $\binom{5}{2}=10, \binom{6}{2}=15, \binom{7}{2}=21$, and $\binom{8}{2}=28$ test words, corresponding to $\mu_2 = 5, 6$, 7 and 8, respectively. Since 5 single-pattern test words are already applied for the double-error-pattern correction, the total numbers of test words are correspondingly $15(=5+10), 20(=5+15), 26(=5+21)$, and $33(=5+28)$ for the four list decoders. It is shown in Fig. 8 that the 26- and 33-test-word based list decoders perform similarly, and yield better performance gains compared to the 15- and 20-test-word based list decoders, by approximately 0.1–0.2 dB. With 26 test words, the list decoder produces an SNR gain of 2.13 dB at $\mathrm{BER}=10^{-4}$. The performance gap with $\hat{P}_b^3$ is due to both the possibility of a nondominant error event as well as the limited number of test words used.

4) Quadruple-Error-Pattern Correction $(M=4)$*:* Now, we compare five sets of $\binom{\mu_3}{3}$ test words for quadruple-error-pattern correction: $\binom{5}{3}=10, \binom{6}{3}=20, \binom{7}{3}=35, \binom{8}{3}=56$, and $\binom{9}{3}=84$. By taking 5 single-pattern test words and 21 double-pattern test words for double/triple-error-pattern corrections as well, Fig. 9 compares the simulated BERs of the five quadruple-pattern-correcting list decoders, based on $36(=5+21+10), 46(=5+21+20), 61(=5+21+35), 82(=5+21+56)$, and $110(=5+21+84)$ test words, respectively. It is observed that 82 and 110-test-word-based list decoders give better error rate performances than the other three decoders: a coding gain of 2.34 dB is achieved at $\mathrm{BER}=10^{-4}$.

By summing up the simulation results in Figs. 6 to 9, the BERs of the M-error-pattern-correcting list decoders, as well as the corresponding lower bounds (dotted lines) can be computed, as shown in Fig. 10, where $M=1,2,3,4$. As the number of test words increases, the list decoder apparently performs better, but the relative performance gain is gradually decreased, e.g., 0.57 dB between single-pattern correction and double-pattern correction, 0.42 dB between double- and triple-pattern corrections, and 0.21 dB between triple and quadruple-pattern corrections, at $\mathrm{BER}=10^{-4}$. We also observe that the lower bound becomes increasingly loose, as M increases.

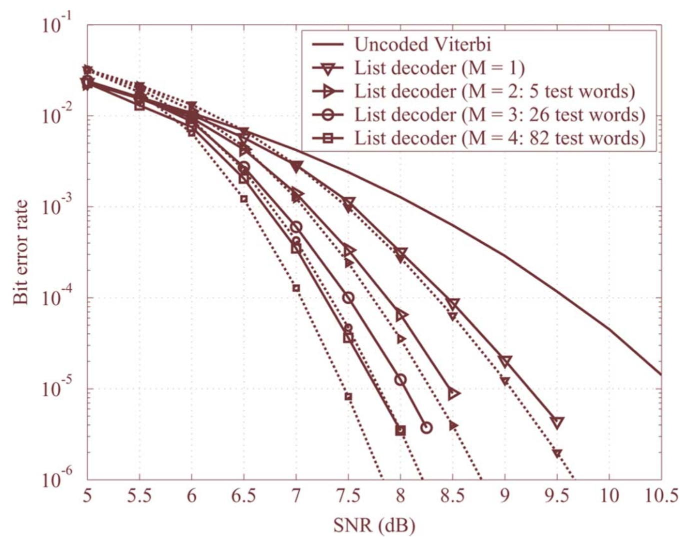

Fig. 10. Comparison of simulated BERs and their lower bounds for the 1-, 2-, 3-, and 4-error-pattern-correcting list decoders.

For the transmitted 350 000 codewords, Table VI illustrates the statistics for the revised number of error patterns within the codewords, at the output of the quadruple-pattern-correcting list decoder based on 82 test words. Comparing to the statistics given in Table V, the number of single error patterns is remarkably reduced $(118\,098 \rightarrow 214)$, and a significant portion of double to quadruple error patterns is also reduced by around 96% $(45\,740 \rightarrow 1\,191)$. Furthermore, it is interesting to observe that the number of more-than-quadruple error patterns, which are not targeted for correction, is also slightly decreased: $167 \rightarrow 145$. We note that the list decoder still attempts to correct the non-targeted multiple error patterns, as long as

TABLE VI
NUMBER OF ERROR PATTERNS WITHIN THE (630, 616) CODEWORDS AND CORRESPONDING WORD ERROR RATES AT THE QUADRUPLE-PATTERN-CORRECTING LIST DECODER OUTPUT ($\mathrm{SNR} = 7.5$ dB, $\mathrm{BER} = 3.63 \times 10^{-5}$)

Number of error patterns within the codeword	Number of occurrences	Word error rate
Single error pattern	214	6.1143×10^{-4}
Double error patterns	693	1.9800×10^{-3}
Triple error patterns	337	9.6286×10^{-4}
Quadruple error patterns	161	4.6000×10^{-4}
5 error patterns	134	3.8286×10^{-4}
6 error patterns	9	2.5714×10^{-5}
7 error patterns	2	5.7143×10^{-6}
Overall	1,550	4.4286×10^{-3}

both the syndrome condition and reliability check are satisfied. As a result, the number of the multiple error patterns can be reduced, if some portions of multiple error patterns are corrected; otherwise, even more multiple error patterns may be produced. As seen in Table VI, the list decoder yields two 7-error-pattern events which did not occur at the Viterbi detector output. Nevertheless, we note that the proposed list decoder does not generate severely long burst errors.

D. Sector Error Rate Performance With Outer RS Code

As another performance comparison, we compute the sector error rate (SER) based on the captured symbol error statistics, assuming a concatenation of an outer t-symbol-correcting RS code. The RS code considered here is a shortened $(410 + 2t, 410, t)$ code based on 10-bit symbols without interleaving [33]. The RS codeword length is matched to the sector length of 4096 bits (or 512 bytes) plus the parity bits. For a given t, comparison is between the Viterbi detector followed by an RS code, and the Viterbi detector followed by the proposed code as the inner code and then the outer RS code. While the SERs in the high error rate region are obtained from direct error counts, the SERs in the low error rate region, e.g., $\mathrm{SER} < 10^{-4}$, are computed using a quasi-analytic simulation method.

One way of doing this calculation is based on the multinomial distribution [13], [34], [35] for the probabilities of τ symbol errors in a $(410 + 2t)$-symbol RS codeword, where $0 \le \tau \le (410 + 2t)$. In the multinomial model, the symbol error statistics are obtained from captured error-pattern events within the outer RS codeword, which are assumed to be independent of each other. However, error events at the inner decoder output are not completely independent, but nonstationary within the channel detector output word. Consequently, the multinomial model does not guarantee rigorous SERs for the proposed list decoder. For the non-stationary symbol error statistics, the block multinomial model has been suggested in [33]. For u symbol errors occurring in an inner codeword consisting of U symbols, $u = 0, \ldots, U$, let p_u be its occurrence probability. The symbol error statistics p_u's are computed from the captured symbol errors within the U-symbol block through computer simulations.

It has been observed that the tail of p_u's asymptotically corresponds to the Gaussian distribution, i.e., $p_u \simeq e^{-(au-b)^2}$, where a and b are any positive real numbers. That is, $-\sqrt{-\log(p_u)}$ for the tail of p_u's can be approximated as a line. Hence, based on the captured symbol error statistics p_u's, $u = 1, \ldots, c < U$, we can estimate $p_{c+1}, \ldots, p_U$, using a Gaussian tail linear extrapolation of the captured statistics. For the Viterbi detector and the quadruple-pattern-correcting list decoder, Table VII illustrates the numbers of occurrences for u-symbol errors through the transmitted 350 000 codeword blocks. It is seen that the number of one to four symbol errors within the blocks is considerably reduced. Also, the number of five to six symbol errors is reduced after the quadruple-error-pattern correction. But, we observe that the number of more than 6 symbol errors is slightly increased, and two 9-symbol errors newly occur at the output of the list decoder. We note that although longer symbol errors are caused by miscorrection of the list decoder, the probability of the miscorrection is very small. Fig. 11 shows the captured p_u's (solid lines), as well as the p_u's (dotted lines) using the Gaussian tail linear extrapolation of the captured statistics given in Table VII. Here, we plot p_u's for $u = 1, \ldots, 15$. While symbol probability tail extrapolation is necessary in this type of analysis, the observed trend in the symbol error statistics indicates that the tail distribution closely follows that of the system without the inner code. We again note that the proposed list decoder does not produce any serious error propagation.

With the captured and extrapolated symbol error statistics p_u's, the probability density function is given by

$$f_{X_0,\ldots,X_U}(x_0, \ldots, x_U) = \frac{B!}{x_0! x_1! \cdots x_U!} \cdot p_0^{x_0} p_1^{x_1} \cdots p_U^{x_U} \tag{59}$$

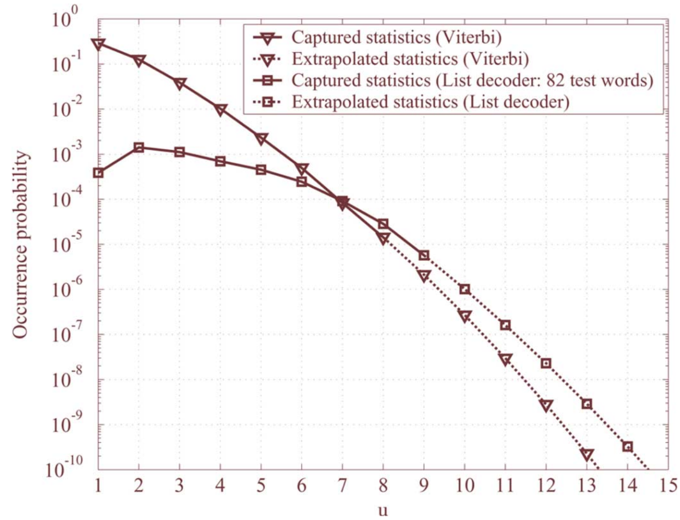

Fig. 11. Captured (solid lines) and extrapolated (dotted lines) symbol error statistics.

TABLE VII
SYMBOL ERROR STATISTICS FOR THE VITERBI DETECTOR AND THE 4-PATTERN-CORRECTING LIST DECODER ($\mathrm{SNR} = 7.5$ dB)

u	Number of occurrences (Viterbi detector)	Number of occurrences (4-pattern-correcting list decoder)
1	101,499	135
2	44,281	492
3	13,597	391
4	3,603	244
5	817	158
6	174	86
7	29	32
8	5	10
9	0	2
Overall	164,005	1,550

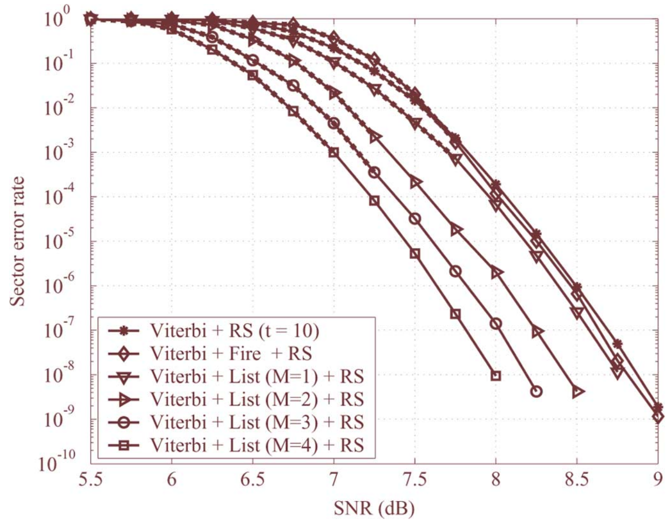

Fig. 12. Sector error rate comparison ($t = 10$).

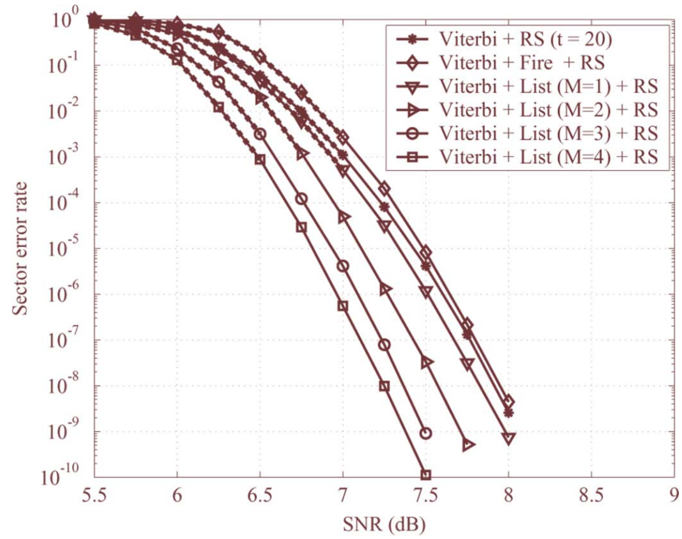

Fig. 13. Sector error rate comparison ($t = 20$).

where B is the number of inner codewords in one sector, x_u is the possible number of occurrences of a u-symbol error in B block codewords, i.e., $B = \sum_{u=0}^{U} x_u$, and p_0 is the probability of no symbol error, i.e., $p_0 = 1 - \sum_{u=1}^{U} p_u$. Then, the block multinomial-based SER is computed as

$$\mathrm{SER} = 1 - \sum_{x_0} \cdots \sum_{x_U} f_{X_0,\ldots X_U}(x_0, \ldots, x_U). \tag{60}$$

Here, the sum is over all combinations of $x_0, \ldots, x_U$ such that $\sum_{u=1}^{U} u \cdot x_u \leq t$.

Fig. 12 compares the block multinomial-based SERs of the $(430, 410, 10)$ RS codes with and without the proposed inner code. The dotted lines shown in the high SER region indicate the SERs based on the direct counts. It is seen that the systems with inner 1-, 2-, 3-, and 4-error-pattern correcting codes achieve SNR gains of 0.13, 0.44, 0.69, and 0.89 dBs at $\mathrm{SER} = 10^{-8}$, respectively, relative to the RS code without an inner code.

Given the 10 dominant error patterns of length-10 or less, a $(19\,437, 19\,408)$ Fire code based on a generator polynomial $g_f(x) = (1 + x^{19})(1 + x^3 + x^{10})$ is also considered for comparison. But, we observed that a (693, 676) Fire code generated by $g_f(x) = (1 + x^{11})(1 + x + x^6)$, which can algebraically correct any single burst errors of length-6 or less, performs better than the $(19\,437, 19\,408)$ Fire code, because of a smaller number of multiple burst errors within the 693-bit word. In addition, the codeword length-693 is close to the codeword length-630 of the proposed code, so that the probabilities of multiple error patterns would be comparable. We use the $(693, 676)$ Fire code as another inner code for comparison purposes. The decoding is based on the error trapping decoder [36]. The corresponding SERs for the inner $(693, 676)$ Fire code are also shown in Fig. 12. Despite the algebraic error correction capability for any targeted and nontargeted single error patterns of length-6 or less, the system with the inner Fire code gives an SNR gain of only 0.055 dB, at $\mathrm{SER} = 10^{-8}$, due to error propagation caused by miscorrection.

Fig. 13 shows similar SER curves for the $(450, 410, 20)$ RS codes with and without the proposed inner code. The $(693, 676)$ Fire code is again considered as an inner code. The SNR gains at $\mathrm{SER} = 10^{-8}$ are respectively seen to be 0.07, 0.34, 0.56, and 0.69 dB with the proposed inner codes, while the system with the inner Fire code performs even worse than the system without an inner code.

V. Generating Soft Output

Generating bit-level soft decisions at the decoder output becomes necessary when the code is to be employed in an iterative system [38]. In our case, a list of candidate codewords that come out of the parallel single-pattern decoders can be used to derive soft decisions.

Once the probability or reliability measure is obtained for each candidate codeword $\hat{\mathbf{c}}'_\zeta$, $\zeta = 1, \ldots, \Omega$, in the list, the conversion of word-level reliability into the bit-level reliability measure of a given bit position can be done using the well-established method of grouping the candidate codewords into two groups, according to the binary value of the hard decision bit in that bit position, and then performing group-wise summing of the word-level probabilities [37].

The novelty in our approach, however, lies in the way it generates the word-level reliability or probability measures. Instead of relying on the usual approach of using the Euclidean distance of the candidate codeword to the observation word, we evaluate the local error pattern probabilities and then construct the codeword probability based on the potential constituent error pattern probabilities. Namely, the *a posteriori* probability of each candidate codeword can be approximated as the product of the local probability measures of its constituent single error patterns, with respect to the channel detector output word. For example, suppose $\hat{\mathbf{c}}'_\zeta$ is a candidate codeword with a potential M-error-pattern event $\sum_{k=1}^{M} x^{\rho_k} e_{i_k}(x)$, with respect to the detector output word $\hat{\mathbf{c}}$. Then, the *a posteriori* probability of this particular codeword can be estimated as

$$P(\hat{\mathbf{c}}'_\zeta | \hat{\mathbf{c}}, \mathbf{r}) = \prod_{k=1}^{M} P_{i_k}^{(\rho_k)} \tag{61}$$

where $\mathbf{r}$ denotes the channel observation at the detector input and, again, $P_{i_k}^{(\rho_k)}$ is the local probability measure of $x^{\rho_k} e_{i_k}(x)$. The application of the proposed code as a building block of an iterative turbo equalization system has been considered in [38] through the use of a soft-input and soft-output decoder.

VI. Conclusion

A new class of high-rate error correcting codes has been developed that provides an efficient correction capability for a specific set of dominant error patterns that are inherent in ISI channels. The proposed code design criterion is new in that the code does not go after low-weight error events or long error events, as in traditional error correction methods. The proposed codes are cyclic block codes, and both the encoder and the decoder are designed under the explicit assumption of an aggressive use of soft information.

The code design is based on a two-step approach. First, a relatively low-rate code is designed by constructing a generator polynomial that yields distinct syndrome sets for all targeted error patterns. Secondly, a new primitive polynomial factor is introduced to form a higher-degree generator polynomial that leads to a considerably lager codeword length while retaining the original error-pattern to syndrome set relationship. The result is a high rate code that is very effective in correcting the single occurrence of any dominant error-pattern.

To handle multiple-pattern occurrences, a new list decoding strategy has been introduced by constructing test decoder input words based on the estimated probabilities of local patterns in the channel detector output word. Both analysis and simulation results indicate that, using the proposed codes, error probabilities can be reduced substantially focusing only on a small dominant error pattern set and thus using only a small coding overhead. Generation of soft bit decisions at the decoder output was also discussed for potential use as a building block of an iterative coding system.

References

[1] J. K. Wolf and G. Ungerboeck, "Trellis coding for partial-response channels," *IEEE Trans. Commun.*, vol. 34, no. 8, pp. 765–773, Aug. 1986.

[2] A. R. Calderbank, C. Heegard, and T. A. Lee, "Binary convolutional codes with application to magnetic recording," *IEEE Trans. Inf. Theory*, vol. 32, no. 6, pp. 797–815, Nov. 1986.

[3] R. Karabed and P. H. Siegel, "Matched spectral-null codes for partial-response channels," *IEEE Trans. Inf. Theory*, vol. 37, no. 3, pp. 818–855, May 1991.

[4] J. Moon and B. Brickner, "Maximum transition run codes for data storage systems," *IEEE Trans. Magn.*, vol. 32, no. 5, pp. 3992–3994, Sep. 1996.

[5] B. Brickner and J. Moon, "Design of a rate 6/7 maximum transition run code," *IEEE Trans. Magn.*, vol. 33, no. 5, pp. 2749–2751, Sep. 1997.

[6] W. G. Bliss, S. She, and L. C. Sundell, "The performance of generalized maximum transition run trellis codes," *IEEE Trans. Magn.*, vol. 34, no. 1, pp. 85–90, Jan. 1998.

[7] R. Karabed and P. H. Siegel, "Coding for higher order partial response channels," *Proc. SPIE*, vol. 2605, pp. 115–126, Oct. 1995.

[8] R. Karabed, P. H. Siegel, and E. Soljanin, "Constrained coding for binary channels with high intersymbol interference," *IEEE Trans. Inf. Theory*, vol. 45, no. 6, pp. 1777–1797, Sept. 1999.

[9] B. Brickner and J. Moon, "Combatting partial erasure and transition jitter in magnetic recording," *IEEE Trans. Magn.*, vol. 36, no. 6, pp. 532–536, Mar. 2000.

[10] R. D. Cideciyan, E. Eleftheriou, B. H. Marcus, and D. S. Modha, "Maximum transition run codes for generalized partial response channels," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 4, pp. 619–634, Apr. 2001.

[11] W. E. Ryan, F. Wang, R. Wood, and Y. Li, "Optimal code rates for the Lorentzian channel: Shannon codes and LDPC codes," *IEEE Trans. Magn.*, vol. 40, no. 6, pp. 3559–3565, Nov. 2004.

[12] T. Conway, "A new target response with parity coding for high density magnetic recording channels," *IEEE Trans. Magn.*, vol. 34, no. 4, pp. 2382–2386, Jul. 1998.

[13] R. D. Cideciyan, J. D. Coker, E. Eleftheriou, and R. L. Galbraith, "Noise predictive maximum likelihood detection combined with parity-based post-processing," *IEEE Trans. Magn.*, vol. 37, no. 2, pp. 714–720, Mar. 2001.

[14] P. Chaichanavong and P. H. Siegel, "Tensor-product parity code for magnetic recording," *IEEE Trans. Magn.*, vol. 42, no. 2, pp. 350–352, Feb. 2006.

[15] B. Vasic, "A graph based construction of high-rate soft-decodable codes for partial response channels," in *Proc. IEEE ICC*, Jun. 2001, pp. 2716–2720.

[16] A. Patapoutian, B. Shen, and P. A. McEwen, "Event error control codes and their applications," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2595–2603, Sep. 2001.

[17] J. Moon and J. Park, "Detection of prescribed error events: Application to perpendicular recording," in *Proc. IEEE ICC*, May 2005, vol. 3, pp. 2057–2062.

[18] J. Moon, J. Park, and J. Lee, "Cyclic redundancy check code based high-rate error-detection code for perpendicular recording," *IEEE Trans. Magn.*, vol. 42, no. 5, pp. 1626–1628, May 2006.

[19] H. Sawaguchi, M. Izumita, and S. Mita, "Soft-output post-processing detection for PRML channels in the presence of data-dependent media noise," in *Proc. IEEE GLOBECOM*, Dec. 2003, pp. 3913–3919.

[20] J. Park and J. Moon, "High-rate error correction codes targeting dominant error patterns," *IEEE Trans. Magn.*, vol. 42, no. 10, pp. 2573–2575, Oct. 2006.

[21] J. Park and J. Moon, "Cyclic codes tailored to a known set of error patterns," in *Proc. IEEE GLOBECOM, GEN03-4*, Nov. 2006, pp. 1–5.

[22] J. Park and J. Moon, "A new class of error-pattern-correcting codes capable of handling multiple error occurrences," *IEEE Trans. Magn.*, vol. 43, no. 6, pp. 2268–2270, Jun. 2007.

[23] M. Madden, M. Öberg, Z. Wu, and R. He, "Read channel for perpendicular magnetic recording," *IEEE Trans. Magn.*, vol. 40, no. 1, pp. 241–246, Jan. 2004.

[24] R. D. Cideciyan, E. Eleftheriou, and T. Mittelholzer, "Perpendicular and longitudinal recording: A signal-processing and coding perspective," *IEEE Trans. Magn.*, vol. 38, no. 4, pp. 1698–1704, Jul. 2002.

[25] P. Kovintavewat, I. Ozgunes, E. Kurtas, J. R. Barry, and S. W. McLaughlin, "Generalized partial-response targets for perpendicular recording with jitter noise," *IEEE Trans. Magn.*, vol. 38, no. 5, pp. 2340–2342, Sep. 2002.

[26] P. Fire, A Class of Multiple-Error-Correcting Binary Codes for Non-Independent Errors Sylvania Electronic Defense Lab., Mountain View, Sylvania Rep., RSL-E-2, Mar. 1959.

[27] G. D. Forney, Jr., "Generalzed-minimum-distance decoding," *IEEE Trans. Inf. Theory*, vol. IT-12, pp. 125–131, Apr. 1966.

[28] D. Chase, "A class of algorithms for decoding block codes with channel measurement information," *IEEE Trans. Inf. Theory*, vol. IT-18, pp. 170–182, Jan. 1972.

[29] G. D. Forney, Jr., "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Trans. Inf. Theory*, vol. IT-18, pp. 363–378, May 1972.

[30] J. Moon and L. R. Carley, "Performance comparison of detection methods in magnetic recording," *IEEE Trans. Magn.*, vol. 26, no. 6, pp. 3155–3172, Nov. 1990.

[31] S. A. Altekar, M. Berggren, B. E. Moision, P. H. Siegel, and J. K. Wolf, "Error-event characterization on partial-response channels," *IEEE Trans. Inf. Theory*, vol. 37, no. 3, pp. 818–855, May 1991.

[32] J. Caroselli and J. K. Wolf, "Error event characterization of partial response systems in magnetic recording systems with medium noise," in *Proc. IEEE GLOBECOM*, Nov. 1998, pp. 2724–2729.

[33] Z. A. Keirn, V. Y. Krachkovsky, E. F. Haratsch, and H. Burger, "Use of redundant bits for magnetic recording: Single-parity codes and Reed-Solomon error-correcting code," *IEEE Trans. Magn.*, vol. 40, no. 1, pp. 225–230, Jan. 2004.

[34] W. Feng, A. Vityaev, G. Burd, and N. Nazari, "On the performance of parity codes in magnetic recording systems," in *Proc. IEEE GLOBECOM*, Nov. 2000, pp. 1877–1881.

[35] W. G. Bliss, R. Karabed, K. Zhang, C. C. Varanasi, and J. Ashley, "Sector error rate estimation of concatenated codes in magnetic recording," in *Proc. IEEE ICC*, June 2001, pp. 2726–2730.

[36] L. Rudolph and M. E. Mitchell, "Implementation of decoders for cyclic codes," *IEEE Trans. Inf. Theory*, vol. IT-10, pp. 259–260, July 1964.

[37] R. M. Pyndiah, "Near-optimum decoding of product codes: Block turbo codes," *IEEE Trans. Commun.*, vol. 46, pp. 1003–1010, Aug. 1998.

[38] H. Alhussien, J. Park, and J. Moon, "Iterative decoding based on error pattern corrections," *IEEE Trans. Magn.*, vol. 44, no. 1, pp. 181–186, Jan. 2008.

Jihoon Park (S'05–M'07) received the B.S. and M.S. degrees in electrical engineering from Hanyang University, Seoul, Korea, in 1994 and 1996, respectively, and the Ph.D. degree in electrical and computer engineering from the University of Minnesota, Minneapolis, MN, in 2007.

From 1996 to 2001, he worked at Samsung Advanced Institute of Technology, Yongin, Korea, as a system engineer for high-speed wireless communication systems. Since 2007, he has been with Hitachi Global Storage Technologies (HGST), San Jose, CA, where he is a research staff member in the HGST San Jose Research Center. His research interests include signal processing, error control coding, and recording system modeling for data storage.

Jaekyun Moon (S'89–M'90–SM'97–F'05) received the B.S.E.E. degree from the State University of New York at Stony Brook and the M.S. and Ph.D. degrees in the Electrical and Computer Engineering Department at Carnegie Mellon University, Pittsburgh, PA.

Since 1990, he has been with the faculty of the Department of Electrical and Computer Engineering at the University of Minnesota, Twin Cities. His research interests are in the area of channel characterization, signal processing and coding for data storage and digital communication.

Dr. Moon received the 1994–1996 McKnight Land-Grant Professorship from the University of Minnesota. He also received the IBM Faculty Development Awards as well as the IBM Partnership Awards. He was awarded the National Storage Industry Consortium (NSIC) Technical Achievement Award for the invention of the maximum transition run (MTR) code, a widely-used error-control/modulation code in commercial storage systems. He served as Program Chair for the 1997 IEEE Magnetic Recording Conference. He is also Past Chair of the Signal Processing for Storage Technical Committee of the IEEE Communications Society. In 2001, he co-founded Bermai, Inc., a fabless semiconductor start-up, and served as founding President and CTO. He served as a Guest Editor for the 2001 IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS issue on Signal Processing for High Density Recording. He also served as an Editor for IEEE TRANSACTIONS ON MAGNETICS in the area of signal processing and coding for 2001-2006. He worked as Chief Technology Officer at Link-A-Media Devices Corporation in 2008.