Binary Block Codes for Noisy Channels with Unknown Offset

Weber, Jos H.; Bu, Renfei; Cai, Kui; Schouhamer Immink, Kees A.

DOI
10.1109/TCOMM.2020.2986200

Publication date
2020

Document Version
Accepted author manuscript

Published in
IEEE Transactions on Communications

Citation (APA)
Weber, J. H., Bu, R., Cai, K., & Schouhamer Immink, K. A. (2020). Binary Block Codes for Noisy Channels with Unknown Offset. *IEEE Transactions on Communications*, *68*(7), 3975-3983. Article 9058711. https://doi.org/10.1109/TCOMM.2020.2986200

Binary Block Codes for Noisy Channels with Unknown Offset

Jos H. Weber, *Senior Member, IEEE,* Renfei Bu, Kui Cai, *Senior Member, IEEE,* and Kees A. Schouhamer Immink, *Fellow, IEEE*

Abstract

Decoders minimizing the Euclidean distance between the received word and the candidate codewords are known to be optimal for channels suffering from Gaussian noise. However, when the stored or transmitted signals are also corrupted by an unknown offset, other decoders may perform better. In particular, applying the Euclidean distance on normalized words makes the decoding result independent of the offset. The use of this distance measure calls for alternative code design criteria in order to get good performance in the presence of both noise and offset. In this context, various adapted versions of classical binary block codes are proposed, such as (i) cosets of linear codes, (ii) (unions of) constant weight codes, and (iii) unordered codes. It is shown that considerable performance improvements can be achieved, particularly when the offset is large compared to the noise.

Jos H. Weber and Renfei Bu are with Delft University of Technology, Delft, The Netherlands, E-mail: j.h.weber@tudelft.nl, r.bu@tudelft.nl. Kui Cai is with Singapore University of Technology and Design (SUTD), Singapore, E-mail: cai_kui@sutd.edu.sg. Kees A. Schouhamer Immink is with Turing Machines Inc., Rotterdam, The Netherlands, E-mail: immink@turing-machines.com.

Most of this research was done while the first author was visiting SUTD in Oct.-Nov. 2018. This work is supported by Singapore Ministry of Education Academic Research Fund Tier 2 MOE2016-T2-2-054

Index Terms

Binary block codes, decoding criteria, noise, offset, performance evaluation.

I. Introduction

Besides the omnipresent noise, an unknown offset is another nuisance in many communication and storage systems. While noise may vary from symbol to symbol, it is often assumed that the offset is constant within a block of symbols. For example, charge leakage from memory cells may cause such an offset of the stored signal values [12]. While Euclidean distance based decoders are known to be optimal if the transmitted or stored signal is only disturbed by Gaussian noise, they may perform badly if there is offset as well. On the other hand, decoders based on the Pearson correlation coefficient are completely immune to offset mismatch, at the expense of a higher noise sensitivity [8].

Various methods to deal with offset mismatch have been proposed. One way is the use of fixed predetermined pilot symbols, from which the offset can be estimated. After subtraction of the offset from the received sequence, the decoder can deal with the noise as usual. However, the pilot symbols lead to a redundancy increase, of course. An alternative method is dynamic threshold detection [7], in which the information is encoded using a conventional error-correcting code and the actual offset is estimated based on the disturbed received symbol sequence, that is re-scaled accordingly and further processed using the Chase algorithm.

In the methodology considered in this paper, no offset estimation is required. Hence, in contrast to prior methods, no extra redundancy and/or operations to deal with offset cancellation are needed. Actually, offset immunity is guaranteed by considering normalized codewords in the decoding process, rather than the codewords themselves. The price to pay for this virtue is a worse noise resistance, since the normalization brings the codewords closer to each other in Euclidean space. Also, the decoding complexity is typically high, which makes the method infeasible for long codes.

With respect to the design of codes that work well in combination with decoders that are immune to offset mismatch, the emphasis has been on constructing a set of codewords $\mathcal{S} \subseteq \{0, 1, \ldots, q-1\}^n$, $q \geq 2$, with the following property [14]. If a vector $(u_1, u_2, \ldots, u_n)$ is in $\mathcal{S}$, then any vector of the type $(u_1 + c, u_2 + c, \ldots, u_n + c)$, with $c \in \mathbb{R}$, $c \neq 0$, is not in $\mathcal{S}$. This indeed avoids codeword ambiguity for such decoders, but the error rate may still be too high due to the noise, since the codewords can be quite close to each other in $\mathbb{R}^n$.

In this paper, we focus on the binary case, i.e., $q = 2$. We design codes that work well with offset-resistant decoders, even if there is considerable noise. One approach is based on classical linear block codes. However, rather than using these codes as such, we consider carefully chosen shifts of these codes, i.e., cosets. Another approach is based on constant weight codes [11]. These are known to be intrinsically resistant to offset mismatch. In particular, we investigate unions of such codes. Finally, we revisit the concept of unordered codes [2], that turns out to be a promising alternative.

The rest of this paper is organized as follows. In Section II, we present the channel model and further preliminaries. Next, we analyze the distance measure under consideration for the binary case in Section III. Based on this analysis, we propose appropriate codes in Sections IV-VI, followed by a performance evaluation in Section VII. Finally, the paper is concluded in Section VIII.

II. Preliminaries

We consider the binary case, in the sense that we have two real signal levels, l_0 and l_1, and that we use codes over GF(2). By appropriate scaling and shifting operations, we assume without loss of generality that $l_0 = 0$ and $l_1 = 1$. Given the context, these zeroes and ones will be considered as elements of either $\mathbb{R}$, e.g., if we use them as signal values, or GF(2), e.g., if we perform algebraic operations on codewords. Furthermore, let '+' denote the real addition and let '$\oplus$' denote the XOR addition.

We assume a channel such that

$$\mathbf{r} = \mathbf{x} + \boldsymbol{\nu} + b\mathbf{1}, \tag{1}$$

where

- $\mathbf{x} = (x_1, \ldots, x_n)$ is the transmitted codeword taken from a code $\mathcal{S} \subseteq \{0,1\}^n \subset \mathbb{R}^n$,
- $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_n) \in \mathbb{R}^n$ is the noise vector, where the ν_i are independently normally (Gaussian) distributed with mean 0 and standard deviation σ,
- b is a real number representing the unknown channel offset,
- $\mathbf{1}$ is the real all-one vector $(1, \ldots, 1)$ of length n, and
- $\mathbf{r} \in \mathbb{R}^n$ is the received vector.

Note that we assume that the noise may vary from symbol to symbol, while the offset is fixed within a block of codeword symbols. The offset value may vary from codeword to codeword though [8]. This precludes the usage of regular offset control estimation based on previously retrieved codewords.

A general decoding technique upon receipt of the vector $\mathbf{r}$ is to choose as the decoder output a codeword optimizing some criterion. In the case of Gaussian noise without offset mismatch, it is well known that minimizing the Euclidean distance between the received vector and the candidate codewords achieves maximum likelihood decoding. The squared Euclidean distance between $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ is defined by

$$\delta(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^{n} (u_i - v_i)^2. \tag{2}$$

Upon receipt of a vector $\mathbf{r}$, a Euclidean decoder outputs

$$\arg\min_{\hat{\mathbf{x}} \in \mathcal{S}} \delta(\mathbf{r}, \hat{\mathbf{x}}). \tag{3}$$

When there is offset mismatch besides the Gaussian noise, then a good alternative, inspired by the well-known Pearson correlation coefficient, is to apply the squared Euclidean distance principle on vectors which are normalized by subtracting their average value from each coordinate [8]. This leads to the distance

$$\delta^*(\mathbf{u},\mathbf{v}) = \delta(\mathbf{u} - \bar{\mathbf{u}}\mathbf{1}, \mathbf{v} - \bar{\mathbf{v}}\mathbf{1}), \tag{4}$$

where $\bar{\mathbf{w}} = \frac{1}{n}\sum_{i=1}^{n} w_i$. Note that this is not a metric in the strict mathematical sense, since, for example, $\delta^*(\mathbf{u},\mathbf{v}) = 0$ may hold for vectors $\mathbf{u}$ and $\mathbf{v} \neq \mathbf{u}$. Actually, an interpretation of (4) is that the vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ are mapped to vectors in the hyperplane $\{\mathbf{y} \in \mathbb{R}^n : \bar{\mathbf{y}} = 0\}$ by orthogonal projection, i.e., in the direction $\mathbf{1}$, and that then the squared Euclidean distance between these projections is calculated. As a consequence, codeword pairs $(\mathbf{u},\mathbf{v})$ such that $\mathbf{u} = \mathbf{v} + c\mathbf{1}$, $c \in \mathbb{R}$, $c \neq 0$, should be avoided, since these cannot be distinguished from each other.

Upon receipt of a vector $\mathbf{r}$, a decoder using measure (4) outputs

$$\arg\min_{\hat{\mathbf{x}} \in \mathcal{S}} \delta^*(\mathbf{r}, \hat{\mathbf{x}}). \tag{5}$$

This criterion is known to be immune to offset mismatch, in the sense that the decoding result is independent of the value of b. However, it is more sensitive to noise than (3), due to the projection as just described, which brings codewords closer together. It has also been shown in [8] that rather than minimizing $\delta^*(\mathbf{r}, \hat{\mathbf{x}})$ among all candidate codewords $\hat{\mathbf{x}} \in \mathcal{S}$, we may as well minimize $\delta(\mathbf{r}, \hat{\mathbf{x}} - \bar{\hat{\mathbf{x}}}\mathbf{1})$, called the modified Pearson distance in [8], since it leads to the same result.

The word error rate (WER) of a code $\mathcal{S}$ when there is no offset mismatch, i.e., $b = 0$, can be upper bounded by using a union bound type of argument. If (3) is used as the decoding criterion, then it is well known [8] that

$$\begin{aligned} \mathrm{WER} &\leq \frac{1}{|\mathcal{S}|}\sum_{\mathbf{u}\in\mathcal{S}}\sum_{\mathbf{v}\in\mathcal{S},\mathbf{v}\neq\mathbf{u}} Q\left(\frac{\sqrt{\delta(\mathbf{u},\mathbf{v})}}{2\sigma}\right) \\ &= \sum_{\alpha\in\mathbb{R}} N_\alpha Q\left(\frac{\sqrt{\alpha}}{2\sigma}\right), \end{aligned} \tag{6}$$

where

$$Q(z) = \frac{1}{\sqrt{2\pi}}\int_z^\infty e^{-u^2/2}du \tag{7}$$

and

$$N_\alpha = \frac{1}{|\mathcal{S}|}\sum_{\mathbf{u}\in\mathcal{S}} |\{\mathbf{v}\in\mathcal{S} : \mathbf{v}\neq\mathbf{u}\wedge\delta(\mathbf{u},\mathbf{v})=\alpha\}|. \tag{8}$$

If (5) is used, we denote the word error rate as WER^*. Then it has been shown in [8] that

$$\begin{aligned} \mathrm{WER}^* &\leq \frac{1}{|\mathcal{S}|}\sum_{\mathbf{u}\in\mathcal{S}}\sum_{\mathbf{v}\in\mathcal{S},\mathbf{v}\neq\mathbf{u}} Q\left(\frac{\sqrt{\delta^*(\mathbf{u},\mathbf{v})}}{2\sigma}\right) \\ &= \sum_{\alpha\in\mathbb{R}} N^*_\alpha Q\left(\frac{\sqrt{\alpha}}{2\sigma}\right), \end{aligned} \tag{9}$$

where

$$N^*_\alpha = \frac{1}{|\mathcal{S}|}\sum_{\mathbf{u}\in\mathcal{S}} |\{\mathbf{v}\in\mathcal{S} : \mathbf{v}\neq\mathbf{u}\wedge\delta^*(\mathbf{u},\mathbf{v})=\alpha\}|. \tag{10}$$

Define

$$\delta_{\min} = \min_{\mathbf{u},\mathbf{v}\in\mathcal{S},\mathbf{u}\neq\mathbf{v}} \delta(\mathbf{u},\mathbf{v}) \tag{11}$$

and

$$\delta^*_{\min} = \min_{\mathbf{u},\mathbf{v}\in\mathcal{S},\mathbf{u}\neq\mathbf{v}} \delta^*(\mathbf{u},\mathbf{v}). \tag{12}$$

For small values of the noise standard deviation σ, we have

$$\mathrm{WER} \approx N_{\delta_{\min}} \times Q\left(\frac{\sqrt{\delta_{\min}}}{2\sigma}\right), \tag{13}$$

if (3) is used, and

$$\mathrm{WER}^* \approx N^*_{\delta^*_{\min}} \times Q\left(\frac{\sqrt{\delta^*_{\min}}}{2\sigma}\right), \tag{14}$$

if (5) is used.

In this paper we focus on the design of binary codes avoiding codeword pairs $(\mathbf{u},\mathbf{v})$ with small $\delta^*(\mathbf{u},\mathbf{v})$ values, since this has a positive impact on the word error rate, which is apparent from the stated expressions.

III. Analysis of $\delta^*(\mathbf{u},\mathbf{v})$ for Binary Vectors

The weight of a binary vector $\mathbf{u}$ is defined by

$$w(\mathbf{u}) = |\{i : u_i = 1\}|. \tag{15}$$

Further, define

$$N(\mathbf{u},\mathbf{v}) = |\{i : u_i = 0 \wedge v_i = 1\}| \tag{16}$$

for any two binary vectors $\mathbf{u}$ and $\mathbf{v}$ of length n. Hence, the Hamming distance between such vectors can be expressed as

$$d(\mathbf{u},\mathbf{v}) = N(\mathbf{v},\mathbf{u}) + N(\mathbf{u},\mathbf{v}). \tag{17}$$

Note that for binary vectors of length n the squared Euclidean distance equals the Hamming distance, i.e.,

$$\delta(\mathbf{u},\mathbf{v}) = d(\mathbf{u},\mathbf{v}), \tag{18}$$

while we have the following result for $\delta^*(\mathbf{u},\mathbf{v})$.

Theorem 1: For any binary vectors $\mathbf{u}$ and $\mathbf{v}$ of length n, it holds that

$$\delta^*(\mathbf{u},\mathbf{v}) = d(\mathbf{u},\mathbf{v}) - \frac{(N(\mathbf{v},\mathbf{u}) - N(\mathbf{u},\mathbf{v}))^2}{n}. \tag{19}$$

Proof. Let $A = |\{i : u_i = 1 \wedge v_i = 1\}|$, $B = N(\mathbf{u}, \mathbf{v})$, and $C = N(\mathbf{v}, \mathbf{u})$. Then $\bar{\mathbf{u}} = (A+C)/n$ and $\bar{\mathbf{v}} = (A+B)/n$, and thus

$$\begin{aligned}
\delta^*(\mathbf{u}, \mathbf{v}) &= \delta(\mathbf{u} - \bar{\mathbf{u}}\mathbf{1}, \mathbf{v} - \bar{\mathbf{v}}\mathbf{1}) \\
&= \sum_{i=1}^{n} \left(u_i - v_i - \frac{(A+C)-(A+B)}{n} \right)^2 \\
&= \sum_{i=1}^{n} \left((u_i - v_i)^2 - \left(\frac{C-B}{n} \right)^2 \right) \\
&= B + C - \frac{(C-B)^2}{n},
\end{aligned}$$

which corresponds to the stated result. ■

For convenience, we define

$$m(\mathbf{u}, \mathbf{v}) = \min\{N(\mathbf{v}, \mathbf{u}), N(\mathbf{u}, \mathbf{v})\}. \tag{20}$$

Corollary 1: For binary vectors $\mathbf{u}$ and $\mathbf{v}$ of length n with $d(\mathbf{u}, \mathbf{v}) = d$, $1 \leq d \leq n$, and $m(\mathbf{u}, \mathbf{v}) = m$, $0 \leq m \leq \lfloor d/2 \rfloor$, we have

$$\delta^*(\mathbf{u}, \mathbf{v}) = d - \frac{(d-2m)^2}{n}. \tag{21}$$

Proof. This is an immediate consequence of Theorem 1 by noting that $|N(\mathbf{v}, \mathbf{u}) - N(\mathbf{u}, \mathbf{v})| = d(\mathbf{u}, \mathbf{v}) - 2m(\mathbf{u}, \mathbf{v})$. ■

This result is illustrated in Table I.

Corollary 2: When considering $\delta^*(\mathbf{u}, \mathbf{v})$ as a function of d and m as given in (21), it shows the following behavior.

- For fixed d, $\delta^*(\mathbf{u}, \mathbf{v})$ is strictly increasing in m, with minimum $d - d^2/n$ at $m = 0$ and maximum d (if d even) or $d - 1/n$ (if d odd) at $m = \lfloor d/2 \rfloor$.
- For fixed $m > \lfloor n/4 \rfloor$, $\delta^*(\mathbf{u}, \mathbf{v})$ is strictly increasing in d, with minimum $2m$ at $d = 2m$ and maximum $4m(n-m)/n$ at $d = n$.

TABLE I

$\delta^*(\mathbf{u},\mathbf{v})$ FOR GIVEN VALUES $d(\mathbf{u},\mathbf{v})=d$ AND $m(\mathbf{u},\mathbf{v})=m$

	$m=0$	$m=1$	$m=2$	$\cdots$	$m=\lfloor \frac{n}{2} \rfloor$
$d=1$	$1-\frac{1}{n}$	$\times$	$\times$	$\cdots$	$\times$
$d=2$	$2-\frac{4}{n}$	2	$\times$	$\cdots$	$\times$
$d=3$	$3-\frac{9}{n}$	$3-\frac{1}{n}$	$\times$	$\cdots$	$\times$
$d=4$	$4-\frac{16}{n}$	$4-\frac{4}{n}$	4	$\cdots$	$\times$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$d=n-2$	$2-\frac{4}{n}$	$6-\frac{16}{n}$	$10-\frac{36}{n}$	$\cdots$	$\times$
$d=n-1$	$1-\frac{1}{n}$	$5-\frac{9}{n}$	$9-\frac{25}{n}$	$\cdots$	$\times$ (n even) $n-1$ (n odd)
$d=n$	0	$4-\frac{4}{n}$	$8-\frac{16}{n}$	$\cdots$	n (n even) $n-\frac{1}{n}$ (n odd)

- For fixed $m \leq \lfloor n/4 \rfloor$, $\delta^*(\mathbf{u},\mathbf{v})$ is strictly increasing from $d = \max\{2m, 1\}$ to $d = 2m + \lfloor n/2 \rfloor$, and then strictly decreasing from $d = 2m + \lceil n/2 \rceil$ to $d = n$. The absolute maximum when n is even is $n/4 + 2m$ (reached at $d = 2m + n/2$). The absolute maximum when n is odd is $(n^2 - 1)/(4n) + 2m$ (reached at $d = 2m + (n-1)/2$ and $d = 2m + (n+1)/2$). The absolute minimum is 0 (reached at $d = n$) if $m = 0$ and $2m$ (reached at $d = 2m$) if $m > 0$.

Proof. These results follow by analyzing (21) using basic calculus tools. ■

Observe that the lowest values of $\delta^*(\mathbf{u},\mathbf{v})$ appear when m is small and d/n is either close to 0 or close to 1. See also Table II. In particular, note that in the design of binary codes without codeword pairs $(\mathbf{u},\mathbf{v})$ with small $\delta^*(\mathbf{u},\mathbf{v})$ values, codeword pairs with large Hamming distances and small $m(\mathbf{u},\mathbf{v})$ values should be avoided. This is a big contrast with classical code design and will be further explored in the next sections.

IV. Coset Codes

A binary block code $\mathcal{S}$ of length n is a subset of $\{0,1\}^n$. A linear binary block code of length n, dimension k, and minimum Hamming distance $d_{\min}$, is denoted as an $[n, k, d_{\min}]$ code. In classical code design, the emphasis was on achieving high code rates, avoiding vector pairs with small Hamming distances, and allowing simple encoding and decoding procedures. Here, we have an additional challenge, as just discussed at the end of previous section. A first priority, when decoding according to (5), is that $\delta^*_{\min} > 0$. Hence, the main focus in literature so far, see, e.g., [14], [15], has been on avoiding codeword pairs $(\mathbf{u},\mathbf{v})$ with $\delta^*(\mathbf{u},\mathbf{v}) = 0$. For the binary case, this leads to the code $\{0,1\}^n \setminus \{\mathbf{1}\}$ of size $2^n - 1$. It has an extremely high code rate, very close to 1, but it suffers from the fact that the number of codewords is not a power of two, which makes information encoding cumbersome. Furthermore, we conclude from the previous section that $\delta^*_{\min} = 1 - 1/n$ for this code, which may be too low to offer

TABLE II

SMALLEST POSSIBLE VALUES OF $\delta^*(\mathbf{u},\mathbf{v})$ IN INCREASING ORDER ($d(\mathbf{u},\mathbf{v})=d$, $m(\mathbf{u},\mathbf{v})=m$)

$\delta^*(\mathbf{u},\mathbf{v})$	Remark
0	For $d=n$, $m=0$, $n\geq 1$
$1-\frac{1}{n}$	For $d=1$ or $d=n-1$, $m=0$, $n\geq 2$
$2-\frac{4}{n}$	For $d=2$ or $d=n-2$, $m=0$, $n\geq 4$
$3-\frac{9}{n}$	For $d=3$ or $d=n-3$, $m=0$, $6\leq n\leq 9$
2	For $d=2$, $m=1$, $n\geq 2$
2	For $d=4$, $m=0$, $n=8$
...	...

sufficient resistance to the noise. Therefore, it is of interest to investigate possibilities of increasing δ^*_{min} and/or enabling easy implementation by sacrificing some rate. The important result presented in the next theorem will be used in order to do so. Let $\mathcal{S}_{\boldsymbol{\alpha}}$ denote the coset of $\mathcal{S}$ obtained by adding the fixed binary vector $\boldsymbol{\alpha}$ of length n to all codewords of $\mathcal{S}$:

$$\mathcal{S}_{\boldsymbol{\alpha}}=\{\boldsymbol{\alpha}\oplus\mathbf{c}|\mathbf{c}\in\mathcal{S}\}. \tag{22}$$

Note that $\mathcal{S}_{\boldsymbol{\alpha}}$ is a code that has the same length, cardinality, rate, and minimum Hamming distance as $\mathcal{S}$, while information can also be uniquely mapped to its codewords by using an encoding procedure for $\mathcal{S}$ followed by a simple shift operation.

Theorem 2: Let $\mathcal{S}$ be a binary $[n, k, d_{\min}]$ code with $d_{\min} \geq 2$, which contains the all-one vector, i.e., $\mathbf{1} \in \mathcal{S}$. Then, for any binary vector $\boldsymbol{\alpha}$ of length n with weight $\lfloor \frac{d_{\min}}{2} \rfloor$, $\lceil \frac{d_{\min}}{2} \rceil$, $n - \lfloor \frac{d_{\min}}{2} \rfloor$, or $n - \lceil \frac{d_{\min}}{2} \rceil$, it holds that

$$\delta^*(\mathbf{u}, \mathbf{v}) \geq d_{\min} \left(1 - \frac{d_{\min}}{n}\right) \tag{23}$$

for all $\mathbf{u}, \mathbf{v} \in \mathcal{S}_{\boldsymbol{\alpha}}$, $\mathbf{u} \neq \mathbf{v}$.

Proof. Since $\mathcal{S}$ contains the all-zero and all-one words, it does not contain codewords $\mathbf{c}$ with $1 \leq w(\mathbf{c}) \leq d_{\min} - 1$ or $n - d_{\min} + 1 \leq w(\mathbf{c}) \leq n - 1$. Hence, due to the linearity, the Hamming distance between any two different codewords in $\mathcal{S}$ is either (i) in between $d_{\min}$ and $n - d_{\min}$ (both inclusive) or (ii) equal to n. Since a shift operation on a code is invariant with respect to Hamming distance, the same holds for any two different codewords $\mathbf{u}$ and $\mathbf{v}$ in $\mathcal{S}_{\boldsymbol{\alpha}}$. We will show that the stated result holds for both cases (i) and (ii).

If $d_{\min} \leq d(\mathbf{u}, \mathbf{v}) \leq n - d_{\min}$, then

$$\begin{aligned} \delta^*(\mathbf{u}, \mathbf{v}) &= d(\mathbf{u}, \mathbf{v}) - \frac{(N(\mathbf{v}, \mathbf{u}) - N(\mathbf{u}, \mathbf{v}))^2}{n} \\ &\geq d(\mathbf{u}, \mathbf{v}) - \frac{(d(\mathbf{u}, \mathbf{v}))^2}{n} \\ &\geq d_{\min} - \frac{(d_{\min})^2}{n} = d_{\min} \left(1 - \frac{d_{\min}}{n}\right), \end{aligned}$$

where the first equality follows from Theorem 1, the first inequality from (17), and the second inequality from the distance restrictions in this case, while taking into account that the parabola $x - x^2/n$ with domain $[d_{\min}, n - d_{\min}]$ obtains its minimum value at the boundaries of this interval.

If $d(\mathbf{u},\mathbf{v})=n$, then

$$\begin{aligned}
\delta^*(\mathbf{u},\mathbf{v}) &= d(\mathbf{u},\mathbf{v})-\frac{(N(\mathbf{v},\mathbf{u})-N(\mathbf{u},\mathbf{v}))^2}{n}\\
&= n-\frac{(n-2m(\mathbf{u},\mathbf{v}))^2}{n}\\
&\geq n-\frac{(n-2\lfloor\frac{d_{\min}}{2}\rfloor)^2}{n}\\
&= 4\left(\left\lfloor\frac{d_{\min}}{2}\right\rfloor-\frac{(\lfloor\frac{d_{\min}}{2}\rfloor)^2}{n}\right)\\
&\geq 2d_{\min}-2-\frac{(d_{\min})^2}{n}\\
&\geq d_{\min}-\frac{(d_{\min})^2}{n}=d_{\min}\left(1-\frac{d_{\min}}{n}\right),
\end{aligned}$$

where the first equality follows from Theorem 1, the second equality from the fact that $\mathbf{u}=\mathbf{1}\oplus\mathbf{v}$ in this case, the first inequality from the fact that $m(\mathbf{u},\mathbf{v})=\min\{w(\mathbf{u}),n-w(\mathbf{u})\}\geq\lfloor\frac{d_{\min}}{2}\rfloor$ due to the specific weight of $\boldsymbol{\alpha}$, and the last inequality from the fact that $d_{\min}\geq 2$. ■

Note that for any binary linear block code $\mathcal{S}$ containing the all-one vector $\delta^*_{\min}=0$ since $\delta^*(\mathbf{0},\mathbf{1})=0$. Theorem 2 shows that $\delta^*_{\min}$ significantly increases by using a well-chosen coset of $\mathcal{S}$ rather than $\mathcal{S}$ itself. Actually, when n is large compared to $d_{\min}$, it follows from (23) that $\delta^*_{\min}$ is close to $\delta_{\min}=d_{\min}$, and thus that the noise performance of the decoder using (5) is close to the noise performance of the decoder using (3), while the former has the advantage of being immune to offset mismatch, in contrast to the latter. Since many classical binary linear block codes do contain the all-one vector, we can try to exploit Theorem 2 in order to design codes which are immune to offset mismatch while having a good noise performance as well. This will be further explored in the following subsections.

A. *Cosets of the Repetition Code*

In the $[n, 1, n]$ repetition code the single information bit is repeated $n-1$ more times. Hence the code has only two codewords, the all-zero and the all-one word. Therefore, $\delta^*_{\min} = 0$ for this code. However, by taking a coset, this can be increased to (almost) the Hamming distance, as shown next.

Theorem 3: Let $\mathcal{S}$ be the binary $[n, 1, n]$ code. Then, for any binary vector $\boldsymbol{\alpha}$ of length n with weight $\lfloor \frac{n}{2} \rfloor$ or $\lceil \frac{n}{2} \rceil$, it holds that

$$\delta^*_{\min} = \begin{cases} n & \text{if } n \text{ is even,} \\ n - 1/n & \text{if } n \text{ is odd,} \end{cases} \tag{24}$$

for $\mathcal{S}_{\boldsymbol{\alpha}}$.

Proof. Since $\mathcal{S}_{\boldsymbol{\alpha}} = \{\boldsymbol{\alpha}, \mathbf{1} \oplus \boldsymbol{\alpha}\}$, it has $\delta^*_{\min} = \delta^*(\boldsymbol{\alpha}, \mathbf{1} \oplus \boldsymbol{\alpha})$, which gives the stated result by applying Theorem 1, while observing that $d(\boldsymbol{\alpha}, \mathbf{1} \oplus \boldsymbol{\alpha}) = n$ and that $|N(\boldsymbol{\alpha}, \mathbf{1} \oplus \boldsymbol{\alpha}) - N(\mathbf{1} \oplus \boldsymbol{\alpha}, \boldsymbol{\alpha})|$ equals $|n/2 - n/2| = 0$ if n is even and $|(n+1)/2 - (n-1)/2| = 1$ if n is odd. ■

B. *Codes with a Single Parity Bit*

Another simple way to provide protection against errors is to use a single parity bit. A codeword then consists of $n - 1$ information bits followed by one parity bit. The parity bit can be chosen in such a way that the number of ones in each codeword is even, in which case the code is indicated as $\mathcal{S}_{\text{even}}$, or by making the number of ones odd, in which case the code is indicated as $\mathcal{S}_{\text{odd}}$. Both $\mathcal{S}_{\text{even}}$ and $\mathcal{S}_{\text{odd}}$ have length n, redundancy 1, code rate $1-1/n$, and minimum Hamming distance 2. Note that $\mathcal{S}_{\text{odd}}$ can be considered to be a coset of the linear $[n, n-1, 2]$ code $\mathcal{S}_{\text{even}}$, i.e., $\mathcal{S}_{\text{odd}} = (\mathcal{S}_{\text{even}})_{\boldsymbol{\alpha}}$ with $\boldsymbol{\alpha}$ being a vector of length n and weight 1.

The use of these codes to deal with noise and offset issues was already briefly discussed in [9], where hybrid Pearson and Euclidean detection was considered. By substituting the value zero for the weighing parameter γ in [9, Eq. (35)] (and then squaring because of a different notation), it appears that a $\delta^*_{\min}$ of $2-4/n$ can be obtained by using a single parity bit. However, this result only holds for even values of n, $n \geq 4$, as shown in the next theorem.

Theorem 4: For binary codes using a single parity bit, the $\delta^*_{\min}$ values are as stated in Table III.

Proof. Let $\mathbf{u}_i$, with $0 \leq i \leq n$, denote the binary vector of length n starting with i ones followed by $n-i$ zeroes. Note that $m(\mathbf{u}_i, \mathbf{u}_j) = 0$ and $d(\mathbf{u}_i, \mathbf{u}_j) = |i-j|$ for all i and j.

If n is odd, then $\mathcal{S}_{\text{even}}$ does not contain $\mathbf{u}_n = \mathbf{1}$, but it does contain both $\mathbf{u}_0 = \mathbf{0}$ and $\mathbf{u}_{n-1}$. Hence it follows from Table II that $\delta^*_{\min} = 1 - 1/n$ for $\mathcal{S}_{\text{even}}$. The same conclusion holds for $\mathcal{S}_{\text{odd}}$ if n is odd, since it does not contain $\mathbf{u}_0 = \mathbf{0}$, but it does contain both $\mathbf{u}_n = \mathbf{1}$ and $\mathbf{u}_1$.

If n is even, then $\mathcal{S}_{\text{even}}$ does contain both $\mathbf{u}_0 = \mathbf{0}$ and $\mathbf{u}_n = \mathbf{1}$, and thus $\delta^*_{\min} = \delta^*(\mathbf{0}, \mathbf{1}) = 0$. Further, Theorem 3 gives that $\mathcal{S}_{\text{odd}} = (\mathcal{S}_{\text{even}})_{\mathbf{u}_1}$ has $\delta^*_{\min} = 2$ if $n = 2$, while Theorem 2 gives that it has $\delta^*_{\min} \geq 2 - 4/n$ in the case of even $n \geq 4$. Equality in this last case follows by observing that $\mathbf{u}_1, \mathbf{u}_{n-1} \in \mathcal{S}_{\text{odd}}$ and $\delta^*(\mathbf{u}_1, \mathbf{u}_{n-1}) = 2 - 4/n$. ■

Hence, from the $\delta^*_{\min}$ perspective, there is a significant difference between $\mathcal{S}_{\text{even}}$ and $\mathcal{S}_{\text{odd}}$ in case n is even.

C. Cosets of (Shortened) Hamming Codes

Next, we consider the important family of $[2^s-1, 2^s-1-s, 3]$ Hamming codes $\mathcal{H}_s$ [10], with $s \geq 3$. The $s \times (2^s - 1)$ parity-check matrix H_s of $\mathcal{H}_s$ consists of all possible columns of length s except the all-zero column. Since these codes contain both the all-zero and the all-one words, they have $\delta^*_{\min} = 0$, but it follows from Theorem 2 that there exist cosets

TABLE III

$\delta^*_{\min}$ FOR SINGLE PARITY CODES

	$n=2$	n even, $n \geq 4$	n odd
$\mathcal{S}_{\text{even}}$	0	0	$1-\frac{1}{n}$
$\mathcal{S}_{\text{odd}}$	2	$2-\frac{4}{n}$	$1-\frac{1}{n}$

with $\delta^*_{\min} \geq 3-9/(2^s-1)$. Equality holds, since each coset can be shown to contain words achieving this value.

The lengths of such Hamming codes are rather restrictive, but actually any length n with $s+1 \leq n \leq 2^s-2$ can be achieved, while maintaining redundancy s and Hamming distance (at least) 3, by applying an appropriate shortening procedure on $\mathcal{H}_s$. Specifically, this can be done by removing h columns from H_s such that the rank of the new parity-check matrix remains s, which leads to a $[2^s-1-h, 2^s-1-s-h, 3]$ code for any $1 \leq h \leq 2^s-s-2$. Next, we will investigate which $\delta^*_{\min}$ values can be achieved when shortening.

If $3 \leq h \leq 2^{s-1}+2$, $s \geq 4$, then we can choose the columns to be removed from H_s in such a way that their XOR sum is the all-zero column, which implies that the resulting shortened code contains the all-one word. This enables the construction of a coset with $\delta^*_{\min} \geq 3-9/(2^s-1-h)$ according to Theorem 2.

If $1 \leq h \leq 2$, then the shortened code does not contain the all-one word, no matter how we choose the removed column(s). These two cases will be discussed next.

If $h=1$, $s \geq 3$, the Hamming distances d between different codewords appearing in the shortened code of length $n = 2^s - 2$ all satisfy $3 \leq d \leq n - 2$. From Table I we can thus conclude that $\delta^*_{\min} \geq 2 - 4/n$. This lower bound is achieved between the all-zero word and any of the codewords of weight $n - 2$. There are $n/2$ such codewords. Each of them has its zeroes in two positions for which the corresponding columns in the parity-check matrix have their XOR sum equal to the removed column in the shortening process. Taking a coset rather than the code itself may reduce the average number of nearest neighbours, but it will not increase $\delta^*_{\min}$.

If $h = 2$, $s \geq 3$, the Hamming distances d between different codewords appearing in the shortened code of length $n = 2^s - 3$ all satisfy $3 \leq d \leq n - 1$. From Table I we can thus conclude that $\delta^*_{\min} \geq 1 - 1/n$. This lower bound is only achieved for the all-zero word and the single codeword of weight $n - 1$. Taking a coset by shifting the code over a vector $\boldsymbol{\alpha}$ of weight one, where the single one in $\boldsymbol{\alpha}$ is not in the position where the codeword of weight $n - 1$ has its single zero, does increase $\delta^*_{\min}$ to $2 - 4/n$. Other choices of $\boldsymbol{\alpha}$ do not lead to higher values of $\delta^*_{\min}$.

Hence, for $h = 1$ or $h = 2$, neither the shortened codes nor their cosets achieve $\delta^*_{\min} = 3 - 9/n$. However, in case we would like to use a code for which the length n is of the format $2^z - 2$ or $2^z - 3$, $z \geq 3$, it is still possible to have $\delta^*_{\min} \geq 3 - 9/n$, by shortening $\mathcal{H}_{z+1}$ by $h = 2^z + 1$ or $h = 2^z + 2$ positions and taking appropriate cosets as indicated before. Note that the resulting codes have redundancy $z + 1$ rather than z, which is the price to be paid for the higher $\delta^*_{\min}$ value.

To summarize the results discussed in this subsection, we provide in Table IV an overview of parameters which can be achieved for appropriately chosen cosets of (shortened) Hamming codes.

TABLE IV

REDUNDANCY AND LOWER BOUNDS ON $\delta^*_{\min}$ OF COSETS OF (SHORTENED) HAMMING CODES OF ANY LENGTH $n \geq 5$

length n	redundancy	$\delta^*_{\min}$
$n = 2^z - 3, 2^z - 2$ (with $z \geq 3$)	$\lceil \log_2(n+1) \rceil$	$2 - \frac{4}{n}$
$n = 2^z - 3, 2^z - 2$ (with $z \geq 3$)	$\lceil \log_2(n+1) \rceil + 1$	$3 - \frac{9}{n}$
$n \geq 7, n \neq 2^z - 3, 2^z - 2 \ \forall z \geq 4$	$\lceil \log_2(n+1) \rceil$	$3 - \frac{9}{n}$

D. Other Coset Codes

The approach of the previous subsections can be applied to any binary block code to obtain further trade-offs between code rate and $\delta^*_{\min}$ values. Descriptions of various celebrated classes of codes, such as BCH and Reed-Muller codes, can be found in text books like [11] and [10]. For the many codes containing the all-one vector, Theorem 2 is a key tool.

V. UNIONS OF CONSTANT WEIGHT CODES

It is well known that constant weight codes are intrinsically resistant to offset mismatch. Here, we will show this once more in the context of our framework, for completeness. Furthermore, we will propose a method of taking the union of several constant weight codes in order to obtain codes with a low redundancy and high $\delta^*_{\min}$.

A binary constant weight code, indicated as $\mathcal{C}(n, M, d, w)$, is a set of M binary vectors of length n, weight w, and mutual Hamming distance at least d, where $0 \leq w \leq n$, $2 \leq d \leq n$, and d is even. For example, the set of all words of length n and weight w is a $\mathcal{C}(n, \binom{n}{w}, 2, w)$

code.

Theorem 5: For any constant weight code with minimum Hamming distance $d_{\min}$, it holds that $\delta^*_{\min} = d_{\min}$.

Proof. For any two vectors $\mathbf{u}$ and $\mathbf{v}$ of the same weight and length, it holds that $N(\mathbf{u},\mathbf{v}) = N(\mathbf{v},\mathbf{u})$ and thus, according to Theorem 1, $\delta^*(\mathbf{u},\mathbf{v}) = d(\mathbf{u},\mathbf{v})$, which implies the statement. ■

Rather than just taking one constant weight code, we may also consider taking the union of several constant weight codes. Based on the findings of Section III, we have the following result.

Theorem 6: For any code which consists of the union of t constant weight codes of the same length n and Hamming distance d, i.e., $\cup_{i=1}^{t}\mathcal{C}(n, M_i, d, w_i)$, such that $n \geq (d+1)^2$, $0 \leq w_1 < w_2 < \cdots < w_t \leq n$, $w_{j+1} - w_j \geq d+1$ for all $j = 1, 2, \ldots, t-1$, and $w_t - w_1 \leq n - d - 1$, it holds that $\delta^*_{\min} \geq d$.

Proof. For any two different codewords $\mathbf{u}$ and $\mathbf{v}$ from the same constant weight subcode, $\delta^*(\mathbf{u},\mathbf{v}) = d(\mathbf{u},\mathbf{v}) \geq d$.

For any two codewords $\mathbf{u}$ and $\mathbf{v}$ from different constant weight subcodes, we have that

$$\begin{aligned}
\delta^*(\mathbf{u},\mathbf{v}) &= d(\mathbf{u},\mathbf{v}) - \frac{(N(\mathbf{v},\mathbf{u}) - N(\mathbf{u},\mathbf{v}))^2}{n} \\
&\geq |w(\mathbf{u}) - w(\mathbf{v})| - \frac{|w(\mathbf{u}) - w(\mathbf{v})|^2}{n} \\
&\geq d + 1 - \frac{(d+1)^2}{n} \geq d + 1 - 1 = d,
\end{aligned}$$

where the first equality follows from Theorem 1, the first inequality from the fact that $d(\mathbf{u},\mathbf{v}) \geq |w(\mathbf{u}) - w(\mathbf{v})| = |N(\mathbf{v},\mathbf{u}) - N(\mathbf{u},\mathbf{v})|$, the second inequality from the weight restrictions as stated in the theorem, while taking into account that the parabola $x - x^2/n$ with domain $[d+1, n-d-1]$ obtains its minimum value at the boundaries of this interval, and the third inequality from the fact that $n \geq (d+1)^2$.

In conclusion, $\delta^*_{\min}$ is at least equal to d. ■

Codes constructed as (the union of) constant weight codes typically possess less algebraic structure than the coset codes of the previous section, but they may have favourable redundancy and distance properties. For example, we consider the constant weight code of length 7 containing all the 35 words of weight 3. The code's Hamming distance is 2 and because of Theorem 5 also $\delta^*_{\min} = 2$. By selecting 32 out of the 35 words we obtain a code which can protect messages of 5 bits, thus the code has redundancy $7 - 5 = 2$. Note that the coset of the $[7, 4, 3]$ Hamming code presented in the previous section has redundancy $7 - 4 = 3$ and $\delta^*_{\min} = 12/7 < 2$, so the constant weight code is better in both aspects. However, it is not systematic in the sense that information bits can be separated from the check bits. Further, note that for the cosets of longer Hamming codes $\delta^*_{\min} = 3 - 9/n > 2$, so those have better noise resistance than a constant weight code of the same length and Hamming distance 2.

A. Codes with Redundancy 2 and $\delta^*_{\min} = 2$

The codes presented in Subsection IV-B have redundancy 1 and $\delta^*_{\min}$ as indicated in Table III, for any length $n \geq 2$. Here, we will present for any length $n \geq 3$, (unions of) constant weight codes with redundancy 2 and $\delta^*_{\min} = 2$.

When $3 \leq n \leq 8$, it holds that $\binom{n}{\lfloor n/2 \rfloor} \geq 2^{n-2}$. Hence, by selecting 2^{n-2} codewords of $\mathcal{C}(n, \binom{n}{\lfloor n/2 \rfloor}, 2, \lfloor n/2 \rfloor)$, we obtain a code of length n with redundancy 2 and $\delta^*_{\min} = 2$.

When $n \geq 9$, then we have the following (almost) systematic construction. Let $\mathbf{m}$ be any message vector of length $n - 2$. We append to $\mathbf{m}$ a vector $\mathbf{n}$ of length 2 to form a codeword

$\mathbf{c} = (\mathbf{m}, \mathbf{n})$ of length n, where

$$\mathbf{n} = \begin{cases} 11 & \text{if } w(\mathbf{m}) \equiv n-4 \pmod 3, \\ 10 & \text{if } w(\mathbf{m}) \equiv n-3 \pmod 3, \\ 00 & \text{if } w(\mathbf{m}) \equiv n-2 \pmod 3. \end{cases} \tag{25}$$

The only exception to this rule is that if $w(\mathbf{m}) = 0$, i.e., $\mathbf{m} = \mathbf{0}$, and $n \equiv 2 \pmod 3$, then we set $\mathbf{c} = 1100\ldots01$. Hence, $\mathbf{m}$ can always be retrieved from $\mathbf{c}$ by omitting the last two bits, except when these bits are equal to 01, in which case $\mathbf{m} = \mathbf{0}$. Note that this code is a collection of constant weight codes, of length n and Hamming distance 2 each, where all the weights appearing are equal to $n-2$ modulo 3 and at least equal to 1 and at most equal to $n-2$. Hence, all the weight and length requirements from Theorem 6 are satisfied, and thus this theorem gives that the code has $\delta^*_{\min} \geq 2$, where equality holds since codeword pairs meeting this bound are readily identified.

B. Codes with $\delta^*_{\min} > 2$

In order to apply Theorems 5 and 6 to obtain codes with $\delta^*_{\min} > 2$, there is a need for constant weight codes with Hamming distance larger than 2. An introduction on such codes is given in [11, Chapter 17], with extensive tables of (bounds on) the code sizes provided in [11, Appendix A]. More recent tables are available via [5].

As an example, we note from [11, App. A, Fig. 3] that there exists a code with length 12 and Hamming distance 4, in which each of the 132 codewords has weight 6. Hence, by selecting 128 of these words, we obtain a code with length 12, redundancy $12-\log_2 128 = 5$, and $\delta^*_{\min} = 4$.

VI. Unordered Codes

In this section we do not present new constructions, but we revisit classes of codes that have been designed for other purposes, but also turn out to be useful in the context considered here.

We say that a transmitted or stored binary codeword suffers from unidirectional errors if all the errors are either of the $0 \to 1$ type or of the $1 \to 0$ type [2]. A necessary and sufficient condition for a code to be capable of detecting any number of unidirectional errors is that the code is unordered, i.e., $m(\mathbf{u}, \mathbf{v}) = \min\{N(\mathbf{v}, \mathbf{u}), N(\mathbf{u}, \mathbf{v})\} \geq 1$ for all codewords $\mathbf{u}$ and $\mathbf{v} \neq \mathbf{u}$. Berger codes [1] are unordered codes which are constructed by taking information words of length k and then appending a tail of length $\lceil \log_2(k+1) \rceil$ which represents the binary representation of the number of zeroes in the information word.

The concept of unordered codes has been extended to t-EC AUED (t error correcting and all unidirectional error detecting) codes [4], [3]. A necessary and sufficient condition for a code to have this property is that

$$m(\mathbf{u}, \mathbf{v}) \geq t + 1 \tag{26}$$

for all codewords $\mathbf{u}$ and $\mathbf{v} \neq \mathbf{u}$. Note that unordered codes appear as a special case by setting $t = 0$. An excellent collection of papers on codes dealing with unidirectional errors has been composed by Blaum [2]. Typically, a t-EC AUED is constructed by taking a classical linear block code with Hamming distance $2t + 1$, which guarantees the correction of up to t errors, and then adding extra bits to the codewords to obtain the detection capability of all unidirectional errors [3].

It follows from Corollary 2 and (26) that any t-EC AUED code $\mathcal{S}$ has

$$\delta^*_{\min} \geq \min_{\mathbf{u}, \mathbf{v} \in \mathcal{S}, \mathbf{v} \neq \mathbf{u}} 2m(\mathbf{u}, \mathbf{v}) \geq 2t + 2. \tag{27}$$

Hence, the vast literature on t-EC AUED codes can be used to find codes with a desired $\delta^*_{\min}$ value. Most of these codes have the virtue of being systematic. For example, it follows from [6] that there exists a systematic 1-EC AUED code with length 9 and redundancy 6. From (27), with $t = 1$, it follows that this code has $\delta^*_{\min} \geq 4$. It should be mentioned that there does exist a code with length 9, $\delta^*_{\min} = 4$ and redundancy 5, obtained by taking a subset of size 16 of the 18 codewords of the constant weight code of length 9, weight 4, and Hamming distance 4 [11, App. A, Fig. 3], but this code is not systematic.

VII. Performance Evaluation

As already argued, the decoder in (3) is optimal with respect to dealing with Gaussian noise, but not capable of handling substantial offset mismatch. On the other hand, the decoder in (5) is completely immune to offset mismatch, at the expense of a higher noise sensitivity. As usual, applying coding techniques will improve the error performance, at the expense of an increased redundancy. The codes proposed in the previous sections have been designed for channels suffering from both noise and offset. Their suitability for such channels is based on their $\delta^*_{\min}$ values. Substitution in (14) leads to an approximation of the WER at high SNR. However, more in-depth research is required to check their actual performance in case of low or moderate SNR.

As an example case, we investigate the performance of (a coset of) the $[15, 11, 3]$ Hamming code $\mathcal{H}_4$, as presented in Subsection IV-C, in various scenarios. Simulation results are shown in Figures 1-4. For both the Hamming code $\mathcal{H}_4$ itself and the coset $\mathcal{H}_{4,\boldsymbol{\alpha}}$ obtained by shifting $\mathcal{H}_4$ over a vector $\boldsymbol{\alpha}$ of length 15 and weight 1, we show the WER values for the Euclidean-based decoder (ED) from (3) and the Pearson-based decoder (PD) from (5). In the figures the abbreviation SNR stands for signal-to-noise ratio, which we define as $-20 \log_{10} \sigma$ dB.

In Figure 1, the offset is very small, $b = 0.02$. In such a case ED performs best. Further,

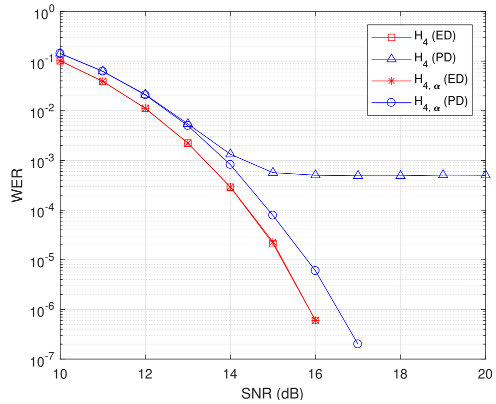

Fig. 1. Word error rate versus signal-to-noise ratio simulation results, in the case of channel offset $b = 0.02$, for the code $\mathcal{H}_4$ and its coset $\mathcal{H}_{4,\boldsymbol{\alpha}}$, each in combination with ED (3) and PD (5).

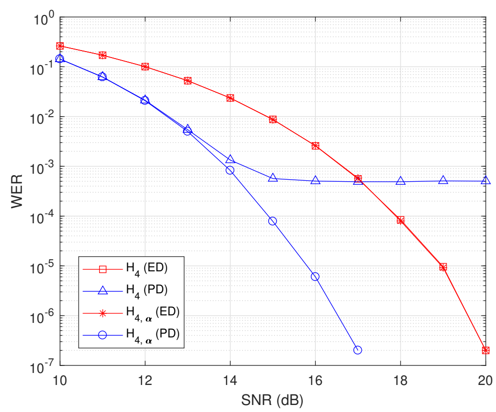

Fig. 2. Word error rate versus signal-to-noise ratio simulation results, in the case of channel offset $b = 0.2$, for the code $\mathcal{H}_4$ and its coset $\mathcal{H}_{4,\boldsymbol{\alpha}}$, each in combination with ED (3) and PD (5).

observe that ED has almost the same performance for $\mathcal{H}_4$ and its coset, but that for PD the coset performs considerably better than the code itself. In fact, since $\mathcal{H}_4$ contains both the all-zero word $\mathbf{0}$ and the all-one word $\mathbf{1}$ having $\delta^*(\mathbf{0}, \mathbf{1}) = 0$, and since $\delta^*(\mathbf{r}, \mathbf{0}) = \delta^*(\mathbf{r}, \mathbf{1})$ for any received vector $\mathbf{r}$, the WER of $\mathcal{H}_4$ approaches for high SNR the value of (14), i.e., it

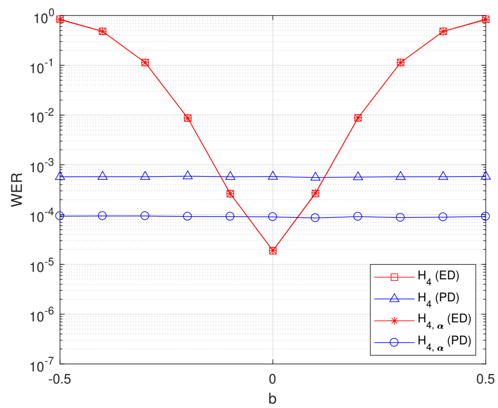

Fig. 3. Word error rate versus channel offset b simulation results, in the case of an SNR of 15 dB, for the code $\mathcal{H}_4$ and its coset $\mathcal{H}_{4,\boldsymbol{\alpha}}$, each in combination with ED (3) and PD (5).

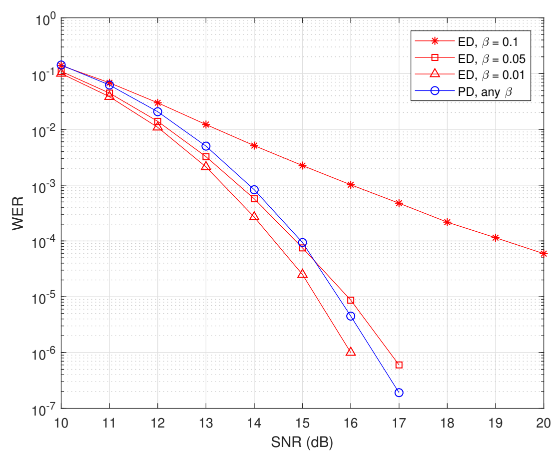

Fig. 4. Word error rate versus signal-to-noise ratio simulation results, in the case of Gaussian distributed channel offset with mean 0 and standard deviation β, for the coset $\mathcal{H}_{4,\boldsymbol{\alpha}}$ of $\mathcal{H}_4$ in combination with ED (3) and PD (5).

has an error floor at $N_0^* \times Q(0) = (2/|\mathcal{H}_4|) \times (1/2) = 1/2048 = 5 \times 10^{-4}$. If the offset value is increased, as done in Figure 2, $b = 0.2$, we observe that the PD performance does not change, as expected since PD is immune to offset mismatch, but that the ED performance is now worse than the performance of the PD with the coset. In Figure 3, the SNR is fixed at 15 dB and the WER is given as a function of the channel offset b. Indeed, we observe that for small $|b|$ ED is the best, but that for larger $|b|$ PD in combination with the coset is superior. Finally, remember that the offset value b may vary from codeword to codeword. In Figure 4, we assume that the offset is i.i.d. Gaussian with mean 0 and standard deviation β. Results are shown for the coset $\mathcal{H}_{4,\boldsymbol{\alpha}}$. The PD performance is of course independent of β, but the ED performance rapidly gets worse with growing values of β.

VIII. Concluding Remarks

In this paper, we have proposed adaptations of classical binary block codes to make them work well with a decoding criterion that guarantees immunity to channel offset mismatch. This immunity generally comes at the price of a high noise sensitivity, but it has been shown that appropriate code design can considerably mitigate this negative effect.

A major concern, however, is the fact that the evaluation of criterion (5) in an exhaustive way is infeasible for large codes. Another issue is that in our analysis, we focused on $\delta^*_{\min}$, but ignored the number of nearest neighbours. Though the former is indeed of utmost importance with respect to the WER performance, the latter could play an important role too. For example, in [7], an extended Hamming code of length 72 and dimension 64 is used, without the shift to a coset as proposed here. Since the code contains both the all-zero and all-one words, it has $\delta^*_{\min} = 0$, but this is only achieved between these two codewords, occurring with a negligible probability of $2^{-64} \approx 5 \times 10^{-20}$ each. Hence, using the code itself rather than a coset is no problem in the case of a large size.

In conclusion, we think that for the codes proposed in this paper, the most promising opportunities, from the application perspective, are for relatively small codes. For example, such codes could be used as inner codes in a concatenated coding scheme, where for the inner decoding (5) is used, while for the outer decoding a fast traditional hard-decision

(Reed-Solomon) decoder is applied.

Besides the additive disturbances, i.e., noise and offset, as in our model $\mathbf{r} = \mathbf{x} + \boldsymbol{\nu} + b\mathbf{1}$, channels may suffer from multiplicative effects as well. This could lead to the more extensive channel model $\mathbf{r} = a(\mathbf{x}+\boldsymbol{\nu})+b\mathbf{1}$, where $a > 0$ is called the gain [8]. Like the offset b, the gain a is assumed to be constant for a transmitted codeword, but it may vary from one codeword to the next. A decoding criterion, that is immune to both gain and offset mismatch, has been proposed in [8], and some basic properties for the binary case have been presented in [13]. An interesting topic for future research is the design of suitable codes for this scenario as well.

REFERENCES

[1] J. M. Berger, "A Note on Error Detection Codes for Asymmetric Channel," *Information and Control*, vol. 4, pp. 68-73, March 1961.

[2] M. Blaum, *Codes for Detecting and Correcting Unidirectional Errors*, IEEE Computer Society Press, Los Alamitos, CA, 1993.

[3] M. Blaum and H. van Tilborg, "On t-Error Correcting/All Unidirectional Error Detecting Codes," *IEEE Transactions on Computers,* vol. C-38, pp. 1493-1501, Nov. 1989.

[4] B. Bose and T. R. N. Rao, "Theory of Unidirectional Error Correcting/Detecting Codes," *IEEE Transactions on Computers,* vol. C-31, pp. 521-530, June 1982.

[5] A. E. Brouwer, "Bounds for Binary Constant Weight Codes," `https://www.win.tue.nl/~aeb/codes/Andw.html`.

[6] J. Bruck and M. Blaum, "New Techniques for Constructing EC/AUED Codes," *IEEE Transactions on Computers,* vol. C-41, pp. 1318-1324, Oct. 1992.

[7] K. A. S. Immink, K. Cai, and J. H. Weber, "Dynamic Threshold Detection Based on Pearson Distance Detection," *IEEE Transactions on Communications*, vol. 66, no. 7, pp. 2958-2965, July 2018.

[8] K. A. S. Immink and J. H. Weber, "Minimum Pearson Distance Detection for Multi-Level Channels with Gain and/or Offset Mismatch," *IEEE Transactions on Information Theory,* vol. 60, pp. 5966-5974, Oct. 2014.

[9] K. A. S. Immink and J. H. Weber, "Hybrid Minimum Pearson and Euclidean Distance Detection," *IEEE Transactions on Communications*, vol. 63, no. 9, pp. 3290-3298, Sept. 2015.

[10] S. Lin and D. J. Costello, Jr., *Error Control Coding (Second Edition)*, Pearson Prentice Hall, 2004.

[11] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*, North-Holland Publishing Company, 1977.

[12] F. Sala, K. A. S. Immink, and L. Dolecek, “Error Control Schemes for Modern Flash Memories: Solutions for Flash Deficiencies,” *IEEE Consumer Electronics Magazine,* vol. 4 (1), pp. 66-73, Jan. 2015.

[13] J. H. Weber and K. A. S. Immink, “Properties of Binary Pearson Codes,” in *Proc. 2018 International Symposium on Information Theory and its Applications (ISITA)*, Singapore, pp. 637-641, Oct. 28-31, 2018.

[14] J. H. Weber, K. A. S. Immink, and S. Blackburn, “Pearson Codes,” *IEEE Transactions on Information Theory,* vol. IT-62, no. 1, pp. 131-135, Jan. 2016.

[15] J. H. Weber, T. G. Swart, and K. A. S. Immink, “Simple Systematic Pearson Coding,” in *Proceedings of the 2016 IEEE International Symposium on Information Theory*, Barcelona, Spain, pp. 385-389, July 10-15, 2016.