Maximum Likelihood Decoding for Channels with Uniform Noise and Signal Dependent Offset

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

DOI
10.1109/ISIT44484.2020.9174270

Publication date
2020

Document Version
Final published version

Published in
2020 IEEE International Symposium on Information Theory (ISIT)

Maximum Likelihood Decoding for Channels with Uniform Noise and Signal Dependent Offset

Renfei Bu
Delft University of Technology
Delft, The Netherlands
r.bu@tudelft.nl

Jos H. Weber
Delft University of Technology
Delft, The Netherlands
j.h.weber@tudelft.nl

Kees A. Schouhamer Immink
Turing Machines Inc.
Rotterdam, The Netherlands
immink@turing-machines.com

***Abstract*—Maximum likelihood (ML) decision criteria have been developed for channels suffering from signal independent offset mismatch. Here, such criteria are considered for signal dependent offset, which means that the value of the offset may differ for distinct signal levels rather than being the same for all levels. An ML decision criterion is derived, assuming uniform distributions for both the noise and the offset. In particular, for the proposed ML decoder, bounds are determined on the standard deviations of the noise and the offset which lead to a word error rate equal to zero. Simulation results are presented confirming the findings.**

***Index Terms*—Maximum likelihood decoding, offset mismatch, signal dependent offset.**

I. Introduction

In data communication and storage systems, noise is usually an important issue, but other physical factors may hamper the reliability of the transmitted or stored data as well. For example, in flash memories, the number of electrons of a flash cell decreases with time and some cells become defective over time [1]. The amount of charge (electrons) leakage, which can be modeled as offset mismatch, depends on various physical parameters, such as the device's temperature, the magnitude of the charge, and the time elapsed between writing and reading the data [2]. In digital optical recording, fingerprints and scratches on the surface of discs result in offset variations of retrieved signals [3]. For direct conversion receivers, the local oscillator is the main source of dc-offset [4].

Consider transmitting a codeword $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ from a codebook $\mathcal{S} \subseteq \mathcal{Q}^n$, where

$$\mathcal{Q} = \{0, 1, \ldots, q-1\}$$

is a q-ary alphabet. In [5], the authors study a channel model where the received sequence $\mathbf{r} = (r_1, r_2, \ldots, r_n)$ is

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

i.e., the transmitted symbols of a codeword are corrupted by noise, represented by the sequence $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ of independent noise samples v_i, and by offset, represented by $b\mathbf{1}$, where $\mathbf{1}$ is the all-one vector of length n. Note that the offset value b is independent of the signal levels in this model. In this paper, however, we study a model in which it is assumed that the offset value depends on the signal level. With every signal level $j \in \mathcal{Q}$ we associate an offset value b_j. The transmitted symbols of a codeword are corrupted by noise, that varies from symbol to symbol, and by offset, that is equal to b_j for each transmitted symbol equal to j. The symbols of the received vector are thus

$$r_i = x_i + v_i + b_{x_i}, \tag{2}$$

where v_i is a noise sample and b_{x_i} represents the signal dependent offset. We assume that the values of the offset, b_j, $j = 0, \ldots, q-1$, may vary from codeword to codeword, but that for each j it is fixed within a codeword of length n. Like the noise values, the offset values are unknown to both the receiver and the transmitter. The channel model in vector form yields

$$\mathbf{r} = \mathbf{x} + \mathbf{v} + \mathbf{b}_{\mathbf{x}}, \tag{3}$$

where $\mathbf{b}_{\mathbf{x}} = (b_{x_1}, b_{x_2}, \ldots, b_{x_n})$. For example,

$$\mathbf{b}_{\mathbf{x}} = (b_0, b_1, b_0, b_1)$$

when the binary word $\mathbf{x} = (0, 1, 0, 1)$ is transmitted.

Much of the current literature pays particular attention to the critical role played by offset [5]–[8], where the offset is the same for all signal levels. In [5], Immink and Weber proposed Pearson distance-based decoders that are immune to gain and/or offset mismatch, at the expense of higher noise sensitivity. Blackburn [6] investigated a maximum likelihood (ML) criterion for channels with Gaussian noise and unknown gain and offset mismatch. In a subsequent study, ML decision criteria were derived for Gaussian noise channels when assuming various distributions for the offset in the absence of gain mismatch [7]. Another relevant reference to this study is the work in [8], where an ML decoding criterion is investigated for channels with bounded noise and offset. This includes an example case where both the noise and the offset are uniformly distributed.

A common feature of these prior studies is the assumption that the offset is independent of signal levels, i.e, it has a fixed value b for all symbols within a transmitted codeword. As stated, this paper is concerned with a different situation, where the offset is a signal dependent parameter, i.e., it equals b_j for signal level j. The model (3) is appropriate in a number of scenarios. For example, the binary input user data is stored as the two resistance states of a spin-torque transfer magnetic random access memory (STT-MRAM) cell [9]. A signal dependent offset model is reasonable when process variation causes an asymmetric distribution of both the low and high resistance states. The model is also appropriate for the multilevel-cell storage capacity and retention of multilevel-cell phase-change memory, which is adversely affected by resistance states dependent drift and noise [10]. Moreover, degradation of the data reliability can be modelled as signal dependent offset model, for the situation that with the increase of temperature, the low signal level hardly changes, while the high signal level decreases, leading to a drift of the high signal level to the low signal level [11].

In this paper, we present an ML decoding criterion for channels with uniform noise and signal dependent offset mismatch. The uniform distribution and the Gaussian distribution are two classical probability distributions of stochastic processes, that are often used to model real-world noise and offset phenomena. Which of the two is the most appropriate depends on the situation under consideration. On the one hand, the Gaussian model is a very natural one, but the fact that it is unbounded may make it less suitable in some cases [12], [13]. In flash memories, for instance, the impact of parameters such as charge leakage on the retrieved data value should not be arbitrarily large. Consequently, not taking into account the bounded nature of stochastic variations may lead to impracticable model-based inferences. Therefore, the uniform model is adopted in this paper. To the best of our knowledge, this is the first paper investigating channels with uniform noise and signal dependent offset. The case of Gaussian noise and signal dependent offset is investigated in [14].

A consequence of having bounded noise and offset is that a zero word error rate (WER) is achievable under certain constraints. A major result of this paper is that we provide sufficient conditions on the standard deviations of the noise and offset in combination with the code properties to guarantee the zero WER.

The outline of the rest of this paper is as follows. We first review the classical Euclidean and Pearson distance-based decoding criteria in Section II. In Section III, we present an ML decoding criterion for channels with uniform noise and signal dependent offset. In Section IV, a zero WER is shown to be achievable if the standard deviations of the noise and offset satisfy certain conditions. Finally, simulation results are given for our ML decoder in comparison with Euclidean and Pearson distance-based decoding criteria in Section V. Conclusion and discussion in Section VI terminate the paper.

II. Preliminaries

This section describes two well-known decoding criteria. The first one is the Euclidean distance-based (ED) decoding criterion, which outputs

$$\mathbf{x}_o = \arg\min_{\hat{\mathbf{x}} \in \mathcal{S}} \delta_E(\mathbf{r}, \hat{\mathbf{x}}),$$

for a received vector $\mathbf{r}$, where

$$\delta_E(\mathbf{r}, \hat{\mathbf{x}}) = \sum_{i=1}^{n} (r_i - \hat{x}_i)^2. \quad (4)$$

It is known to be optimal if the signal is only disturbed by Gaussian noise, but may perform badly if there is offset.

A distance measure [5], inspired by the well-known Pearson correlation coefficient, is proposed in situations which require resistance towards unknown offset mismatch b. Note that here the offset b is the same for all signal levels. For any vector $\mathbf{u} \in \mathbb{R}^n$, let

$$\bar{\mathbf{u}} = \frac{1}{n} \sum_{i=1}^{n} u_i$$

denote the average symbol value. The modified Pearson distance (PD) is defined by

$$\delta_P(\mathbf{r}, \hat{\mathbf{x}}) = \sum_{i=1}^{n} (r_i - \hat{x}_i + \bar{\hat{\mathbf{x}}})^2. \quad (5)$$

This is actually applying the squared Euclidean distance on codewords which are normalized by subtracting their vector average value from each coordinate. A PD decoder chooses a codeword minimizing this distance, that is,

$$\mathbf{x}_o = \arg\min_{\hat{\mathbf{x}} \in \mathcal{S}} \delta_P(\mathbf{r}, \hat{\mathbf{x}}).$$

This criterion is immune to signal independent offset mismatch, thus it is the ML choice for channels $\mathbf{r} = \mathbf{x} + b\mathbf{1}$. Note that decoders based on distance (5) cannot distinguish between the vectors $\mathbf{x}$ and $\mathbf{x} + c\mathbf{1}$, $c \in \mathbb{R}$. The use of the Pearson distance demands that the set of codewords satisfies certain special properties. Such sets are called Pearson codes [15].

III. Maximum Likelihood Decoding

In this section, we present an ML decoding criterion for channels suffering from uniform noise and signal dependent offset. We start by specifying the noise and offset distributions and introducing some further notation, and then present the main result.

For the noise vector $\mathbf{v} = (v_1, \ldots, v_n)$, we assume that the v_i are independently uniformly distributed with mean 0 and variance σ^2. Hence, the probability density function of each v_i, $i = 1, 2, \ldots, n$, is

$$\upsilon(v_i) = \begin{cases} \dfrac{1}{2\sqrt{3}\sigma}, & -\sqrt{3}\sigma < v_i < \sqrt{3}\sigma, \\ 0, & \text{otherwise}, \end{cases} \quad (6)$$

leading to a probability density function $\chi(\mathbf{v}) = \prod_{i=1}^{n} \upsilon(v_i)$ for $\mathbf{v}$.

We assume that the b_j are independently uniformly distributed with mean 0 and standard deviations β_j. The probability density function of each b_j, $j = 0, 1, \ldots, q-1$. is

$$\zeta(b_j) = \begin{cases} \dfrac{1}{2\sqrt{3}\beta_j}, & -\sqrt{3}\beta_j < b_j < \sqrt{3}\beta_j, \\ 0, & \text{otherwise}. \end{cases} \quad (7)$$

For $j = 0, 1, \ldots, q-1$, let $\mathbf{x}^{(j)}$ denote the index set indicating the positions in $\mathbf{x}$ for which the symbol value equals j. For example, in case $q = 2$,

$$\mathbf{x}^{(0)} = \{1, 3\} \text{ and } \mathbf{x}^{(1)} = \{2, 4\}$$

when $\mathbf{x} = (0,1,0,1)$.

For the received vector $\mathbf{r}$ and a candidate codeword $\hat{\mathbf{x}} \in \mathcal{S}$, we define

$$u_j(\mathbf{r},\hat{\mathbf{x}}) = \min\left(\{r_i - \hat{x}_i + \sqrt{3}\sigma \,\middle|\, i \in \hat{\mathbf{x}}^{(j)}\} \cup \{\sqrt{3}\beta_j\}\right),$$
$$l_j(\mathbf{r},\hat{\mathbf{x}}) = \max\left(\{r_i - \hat{x}_i - \sqrt{3}\sigma \,\middle|\, i \in \hat{\mathbf{x}}^{(j)}\} \cup \{-\sqrt{3}\beta_j\}\right),$$

for $j = 0,1,\ldots,q-1$. These parameters correspond to upper and lower bounds on the possible values of b_j for $\hat{\mathbf{x}}$ when $\mathbf{r}$ is received. Further, let

$$m_j(\mathbf{r},\hat{\mathbf{x}}) = \max\{u_j(\mathbf{r},\hat{\mathbf{x}}), l_j(\mathbf{r},\hat{\mathbf{x}})\},$$

and

$$I_j(\mathbf{r},\hat{\mathbf{x}}) = \max\{u_j(\mathbf{r},\hat{\mathbf{x}}) - l_j(\mathbf{r},\hat{\mathbf{x}}), 0\}$$

for $j = 0,1,\ldots,q-1$. Next, we present a decoding criterion and show that it is ML for the channel under consideration.

Theorem 1: If the noise and the offsets in (3) have probability density functions as (6) and (7), respectively, then ML decoding is achieved by maximizing

$$\prod_{j=0}^{q-1} I_j(\mathbf{r},\hat{\mathbf{x}}) \tag{8}$$

over all codewords $\hat{\mathbf{x}} \in \mathcal{S}$.

Proof: If a vector $\mathbf{r}$ is received, ML decoding must determine a codeword $\hat{\mathbf{x}} \in \mathcal{S}$ maximizing $\mathrm{P}(\mathbf{r}\,|\hat{\mathbf{x}})$, that is, the probability that $\mathbf{r}$ is received, given $\hat{\mathbf{x}}$ is sent. This satisfies

$$\begin{aligned}
\mathrm{P}(\mathbf{r}\,|\hat{\mathbf{x}}) &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \chi(\mathbf{r}-\hat{\mathbf{x}}-b_{\hat{\mathbf{x}}}) \prod_{j=0}^{q-1}(\zeta(b_j)db_j) \\
&= \frac{1}{(2\sqrt{3}\sigma)^n}\prod_{j=0}^{q-1}\left(\int_{l_j(\mathbf{r},\hat{\mathbf{x}})}^{m_j(\mathbf{r},\hat{\mathbf{x}})} \frac{1}{2\sqrt{3}\beta_j}db_j\right) \\
&= \frac{\prod_{j=0}^{q-1} I_j(\mathbf{r},\hat{\mathbf{x}})}{(2\sqrt{3}\sigma)^n(2\sqrt{3})^q\prod_{j=0}^{q-1}\beta_j}.
\end{aligned} \tag{9}$$

The first equality is due to the channel model (3). The second equality follows from the probability density functions (6) and (7). The third equality follows from the observation that

$$\int_{l_j(\mathbf{r},\hat{\mathbf{x}})}^{m_j(\mathbf{r},\hat{\mathbf{x}})} db_j = m_j(\mathbf{r},\hat{\mathbf{x}}) - l_j(\mathbf{r},\hat{\mathbf{x}}) = I_j(\mathbf{r},\hat{\mathbf{x}})$$

for all j. Since the denominator in (9) is a constant term for all candidate codewords, we can ignore it during the maximization process, which gives (8). ■

The bounded nature of the noise and the offset has interesting consequences with respect to the WER of the ML decoder, as will be further explored in the next section.

IV. Zero Word Error Rate Analysis

Most interestingly, for the ML decoder based on (8), a word error rate (WER) of zero is achieved if the standard deviations of the noise and the offset satisfy certain conditions. This is shown in the next theorem.

Theorem 2: If the noise and the offsets in (3) have probability density functions as (6) and (7), respectively, with

$$\sigma \le \min_{\substack{\mathbf{s},\mathbf{c}\in\mathcal{S} \\ \mathbf{s}\neq\mathbf{c}}}\left(\frac{\max\limits_{i\in\{1,\ldots,n\}}\{|c_i - s_i| - \sqrt{3}(\beta_{c_i}+\beta_{s_i})\}}{2\sqrt{3}}\right) \tag{10}$$

or

$$\sigma \le \min_{\substack{\mathbf{s},\mathbf{c}\in\mathcal{S} \\ \mathbf{s}\neq\mathbf{c}}}\left(\frac{\max\limits_{j\in\mathcal{Q}}\left\{\max\limits_{i,k\in\mathbf{c}^{(j)}}\{s_k - s_i - \sqrt{3}(\beta_{s_k}+\beta_{s_i})\}\right\}}{4\sqrt{3}}\right), \tag{11}$$

then the ML decoder achieves a WER equal to zero.

Proof: Assume that $\mathbf{x} \in \mathcal{S}$ is sent and $\mathbf{r} = \mathbf{x} + \mathbf{v} + b_{\mathbf{x}}$ is received. We will show that if (10) or (11) holds, then $\mathrm{P}(\mathbf{r}\,|\hat{\mathbf{x}}) = 0$ for all codewords $\hat{\mathbf{x}} \neq \mathbf{x}$. First of all, note that

$$\begin{aligned}
&u_j(\mathbf{r},\hat{\mathbf{x}}) - l_j(\mathbf{r},\hat{\mathbf{x}}) \\
&= \min\left(\{r_i - \hat{x}_i + \sqrt{3}\sigma \,\middle|\, i \in \hat{\mathbf{x}}^{(j)}\} \cup \{\sqrt{3}\beta_j\}\right) \\
&\quad - \max\left(\{r_i - \hat{x}_i - \sqrt{3}\sigma \,\middle|\, i \in \hat{\mathbf{x}}^{(j)}\} \cup \{-\sqrt{3}\beta_j\}\right) \\
&= \min\left(\{r_i - \hat{x}_i + \sqrt{3}\sigma \,\middle|\, i \in \hat{\mathbf{x}}^{(j)}\} \cup \{\sqrt{3}\beta_j\}\right) \\
&\quad + \min\left(\{-(r_i - \hat{x}_i) + \sqrt{3}\sigma \,\middle|\, i \in \hat{\mathbf{x}}^{(j)}\} \cup \{\sqrt{3}\beta_j\}\right) \\
&= \min\Bigg(\{2\sqrt{3}\beta_j\} \cup \{\min_{i\in\hat{\mathbf{x}}^{(j)}}\{-|r_i - \hat{x}_i| + \sqrt{3}\sigma + \sqrt{3}\beta_j\}\} \\
&\quad \cup \{\min_{i,k\in\hat{\mathbf{x}}^{(j)}}\{(r_i - \hat{x}_i) - (r_k - \hat{x}_k) + 2\sqrt{3}\sigma\}\}\Bigg), \\
&= \min\Bigg(\{2\sqrt{3}\beta_j\} \cup \{\min_{i\in\hat{\mathbf{x}}^{(j)}}\{-|r_i - \hat{x}_i| + \sqrt{3}\sigma + \sqrt{3}\beta_j\}\} \\
&\quad \cup \{\min_{i,k\in\hat{\mathbf{x}}^{(j)}}\{r_i - r_k + 2\sqrt{3}\sigma\}\}\Bigg),
\end{aligned} \tag{12}$$

for $j = 0,1,\ldots q-1$.

Next, we show that if (10) or (11) holds, (12) will be negative for some j whenever $\hat{\mathbf{x}} \neq \mathbf{x}$. Note that the final expression in (12) contains a union of three terms, where the first term is always positive since β_j is positive. We show that the second term is negative for some j if (10) holds and that the third term is negative for some j if (11) holds.

For each $\hat{\mathbf{x}} \in \mathcal{S}, \hat{\mathbf{x}} \neq \mathbf{x}$, let j_0 be a symbol from $\mathcal{Q}$ such that a position i_0 in $\hat{\mathbf{x}}^{(j_0)} \subseteq \{1,\ldots,n\}$ maximizes the expression $|x_i - \hat{x}_i| - \sqrt{3}(\beta_{x_i} + \beta_{\hat{x}_i})$. That is,

$$i_0 = \arg\max_{i\in\{1,\ldots,n\}}\left\{|x_i - \hat{x}_i| - \sqrt{3}(\beta_{x_i} + \beta_{\hat{x}_i})\right\},$$
$$j_0 = \hat{x}_{i_0}.$$

Note that this j_0 is not necessarily the same for each $\hat{\mathbf{x}}$. If (10) holds, then we have

$$\begin{aligned}
&\min_{i\in\hat{\mathbf{x}}^{(j_0)}}\{-|r_i-\hat{x}_i|+\sqrt{3}\sigma+\sqrt{3}\beta_j\}\\
&=\min_{i\in\hat{\mathbf{x}}^{(j_0)}}\{-|r_i-\hat{x}_i|-\sqrt{3}\sigma-\sqrt{3}\beta_{x_i}+\sqrt{3}(2\sigma+\beta_{x_i}+\beta_j)\}\\
&<\min_{i\in\hat{\mathbf{x}}^{(j_0)}}\{-|r_i-\hat{x}_i|-|v_i+b_{x_i}|+\sqrt{3}(2\sigma+\beta_{x_i}+\beta_j)\}\\
&=\min_{i\in\hat{\mathbf{x}}^{(j_0)}}\{-|r_i-\hat{x}_i|-|r_i-x_i|+\sqrt{3}(2\sigma+\beta_{x_i}+\beta_j)\}\\
&\leq\min_{i\in\hat{\mathbf{x}}^{(j_0)}}\{-|x_i-\hat{x}_i|+\sqrt{3}(2\sigma+\beta_{x_i}+\beta_j)\}\\
&=-\max_{i\in\hat{\mathbf{x}}^{(j_0)}}\{|x_i-\hat{x}_i|-\sqrt{3}(\beta_{x_i}+\beta_{\hat{x}_i})\}+2\sqrt{3}\sigma\\
&=-\max_{i\in\{1,\ldots,n\}}\{|x_i-\hat{x}_i|-\sqrt{3}(\beta_{x_i}+\beta_{\hat{x}_i})\}+2\sqrt{3}\sigma\\
&\leq 0
\end{aligned}$$

The first inequality follows from the fact that $|v_i+b_{x_i}|\leq|v_i|+|b_{x_i}|<\sqrt{3}\sigma+\sqrt{3}\beta_{x_i}$, the second inequality from the triangular inequality, and the last inequality from (10). Thus the second term of (12) is negative for some j whenever $\hat{\mathbf{x}}\neq\mathbf{x}$ if (10) holds.

Similarly, for each $\hat{\mathbf{x}}\in\mathcal{S},\hat{\mathbf{x}}\neq\mathbf{x}$, let j_1 be a symbol from $\mathcal{Q}$ such that

$$j_1=\arg\max_{j\in\mathcal{Q}}\left\{\max_{i,k\in\hat{\mathbf{x}}^{(j)}}\{x_k-x_i-\sqrt{3}(\beta_{x_k}+\beta_{x_i})\}\right\}.$$

If (11) holds, then we have

$$\begin{aligned}
&\min_{i,k\in\hat{\mathbf{x}}^{(j_1)}}\{r_i-r_k+2\sqrt{3}\sigma\}\\
&<\min_{i,k\in\hat{\mathbf{x}}^{(j_1)}}\{r_i-r_k-(v_i-v_k)\}+4\sqrt{3}\sigma\\
&=\min_{i,k\in\hat{\mathbf{x}}^{(j_1)}}\{x_i-x_k+b_{x_i}-b_{x_k}\}+4\sqrt{3}\sigma\\
&<\min_{i,k\in\hat{\mathbf{x}}^{(j_1)}}\{x_i-x_k+\sqrt{3}(\beta_{x_i}+\beta_{x_k})\}+4\sqrt{3}\sigma\\
&=-\max_{i,k\in\hat{\mathbf{x}}^{(j_1)}}\{x_k-x_i-\sqrt{3}(\beta_{x_i}+\beta_{x_k})\}+4\sqrt{3}\sigma\\
&=-\max_{j\in\mathcal{Q}}\left\{\max_{i,k\in\hat{\mathbf{x}}^{(j)}}\{x_k-x_i-\sqrt{3}(\beta_{x_i}+\beta_{x_k})\}\right\}+4\sqrt{3}\sigma\\
&\leq 0.
\end{aligned}$$

The first inequality follows because $v_i-v_k\leq|v_i|+|v_k|<2\sqrt{3}\sigma$, the second inequality follows because $b_{x_i}-b_{x_k}\leq|b_{x_i}|+|b_{x_k}|<\sqrt{3}(\beta_{x_i}+\beta_{x_k})$, and the last one from (11).

In conclusion, we have for any codeword $\hat{\mathbf{x}}\neq\mathbf{x}$ that $I_j(\mathbf{r},\hat{\mathbf{x}})=0$ for some j if (10) or (11) holds. Hence,

$$\prod_{j=0}^{q-1}I_j(\mathbf{r},\hat{\mathbf{x}})=\mathrm{P}(\mathbf{r}\,|\hat{\mathbf{x}})=0$$

for all codewords $\hat{\mathbf{x}}\neq\mathbf{x}$, while

$$\prod_{j=0}^{q-1}I_j(\mathbf{r},\mathbf{x})>0.$$

Hence, the transmitted codeword is always the outcome of the decoding procedure maximizing (8), and thus ML decoding achieves a WER equal to zero. ■

Next, we give a sufficient condition to achieve zero WER for the ML decoder in the binary case, i.e., $q=2$.

Corollary 3: If the noise and the offsets have probability density functions as (6) and (7), respectively, with

$$2\sigma+\beta_0+\beta_1\leq\frac{1}{\sqrt{3}},\tag{13}$$

then the ML decoder achieves a WER equal to zero for a binary codebook.

Proof: In the binary case, the expression $|c_i-s_i|-\sqrt{3}(\beta_{c_i}+\beta_{s_i})$ in (10) has one of three values depending on c_i and s_i, i.e.,

$$|c_i-s_i|-\sqrt{3}(\beta_{c_i}+\beta_{s_i})=\begin{cases}-2\sqrt{3}\beta_0, & \text{if } (c_i,s_i)=(0,0);\\ -2\sqrt{3}\beta_1, & \text{if } (c_i,s_i)=(1,1);\\ 1-\sqrt{3}(\beta_0+\beta_1), & \text{otherwise.}\end{cases}\tag{14}$$

Since β_0 and β_1 are both positive, $-2\sqrt{3}\beta_0$ and $-2\sqrt{3}\beta_1$ are both negative, and thus if (13) holds then it immediately follows from the fact that σ is positive as well that

$$1-\sqrt{3}(\beta_0+\beta_1)>0.$$

For any codewords $\mathbf{s}$ and $\mathbf{c}\neq\mathbf{s}$, there exists at least one position, k, such that $c_k\neq s_k$, and then we have $|c_k-s_k|=1$ and $\beta_{c_k}+\beta_{s_k}=\beta_0+\beta_1$. In conclusion, if (13) holds, maximizing (14) over $i\in\{1,\ldots,n\}$ outputs $1-\sqrt{3}(\beta_0+\beta_1)$ as its maximum value for any codewords $\mathbf{s}$ and $\mathbf{c}\neq\mathbf{s}$. Hence, according to Theorem 2, the ML decoder achieves a WER equal to zero when (13) holds. ■

Theorem 2 has important implications for developing zero WER codes for channels suffering from uniform noise and signal dependent offset. Code design for these channels is beyond the scope of this paper, but in the next section we provide a performance analysis for a simple example code to show the advantage of the ML decoding technique in comparison to the ED and PD decoders as presented in Section II. Also, it will be illustrated that a zero WER indeed appears in cases that the standard deviations of the noise and the offsets are sufficiently small.

V. Performance Evaluation

Simulated WER results are shown in Fig. 1 for the binary codebook

$$\mathcal{S}^*=\{(0,0,0),(1,1,0),(1,0,1),(0,1,1)\}$$

of length 3 and size 4, in combination with ED, PD, and ML decoders. This simple codebook is used to demonstrate some important WER characteristics. The standard deviations of the signal dependent offsets are set to $\beta_0=0.2$ and $\beta_1=0.15$. We observe that the performance of each of the three decoders declines with increasing values of σ. The PD decoder has the worst performance among these three. The performance of

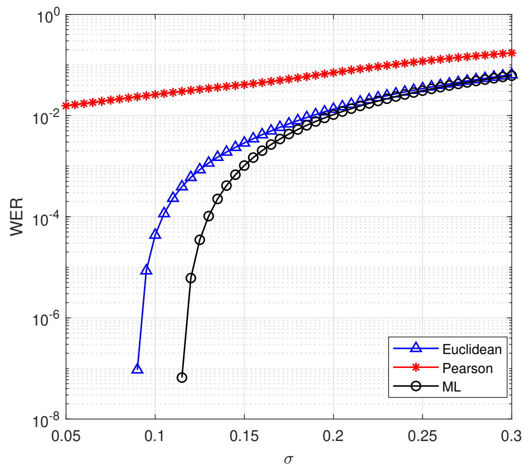

Fig. 1. WER versus the standard deviation σ of the uniform noise for $\mathcal{S}^*$ in combination with ED, PD, and ML decoders, when the standard deviations of the uniform offsets are $\beta_0 = 0.2$, $\beta_1 = 0.15$.

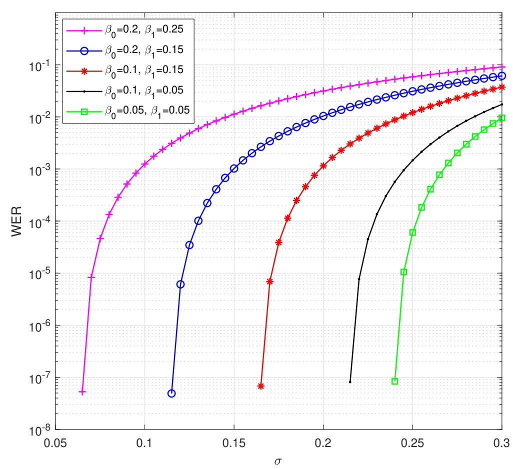

Fig. 2. WER versus the standard deviation σ of the uniform noise for $\mathcal{S}^*$ in combination with ML decoding, with different values of the uniform offset standard deviations β_0, β_1.

the proposed ML decoder is better than for the ED and PD decoders.

Simulation results for $\mathcal{S}^*$ with various values of β_0 and β_1 are shown in Fig. 2. The results from Fig. 2 confirm Corollary 3. Zero WER of the ML decoder is indeed achieved if $2\sigma + \beta_0 + \beta_1 \leq 1/\sqrt{3} \approx 0.58$. We can also observe this in Fig. 1, where for the ML decoder zero WER is achieved when the value of σ is less than $(1/\sqrt{3} - 0.20 - 0.15)/2 \approx 0.11$.

VI. Conclusion and Discussion

In this paper, we have considered channels with not only noise, but also another important channel impairment, *offset*. We have investigated the situation that the noise and the offset are uniformly distributed, where, most importantly, the offset is assumed to be signal dependent. An ML decoding criterion is derived for such channels. This theoretical work can serve as the basis for the design of advanced channel coding schemes for the offset mismatch channel. We have also shown that the ML decoder can achieve a zero WER when the standard deviations of the noise and the offset are small enough. More broadly, research is also needed to determine an ML decoding criterion for the case that the signal dependent offsets have some correlation. Perhaps the concept of copulas [16] could be used in this case. Further investigations on how codebooks can be generated satisfying the conditions as stated in Theorem 2 given σ, β_0, and β_1 will be of interest as well.

References

[1] D. Ajwani, I. Malinger, U. Meyer, and S. Toledo, "Characterizing the performance of flash memory storage devices and its impact on algorithm design," in *Proc. Int. Workshop on Experimental and Efficient Algorithms*. Berlin, Heidelberg: Springer, May 2008, pp. 208–219.

[2] 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, no. 1, pp. 66–73, Jan. 2015.

[3] G. Bouwhuis, J. Braat, A. Huijser, J. Pasman, G. van Rosmalen, and K. A. S. Immink, *Principles of Optical Disc Systems*. Bristol and Boston: Adam Hilger, 1985.

[4] R. Svitek and S. Raman, "DC offsets in direct-conversion receivers: characterization and implications," *IEEE Microwave Mag.*, vol. 6, no. 3, pp. 76–86, Sep. 2005.

[5] K. A. S. Immink and J. H. Weber, "Minimum Pearson distance detection for multilevel channels with gain and/or offset mismatch," *IEEE Trans. Inform. Theory*, vol. 60, no. 10, pp. 5966–5974, Oct. 2014.

[6] S. R. Blackburn, "Maximum likelihood decoding for multilevel channels with gain and offset mismatch," *IEEE Trans. Inform. Theory*, vol. 62, no. 3, pp. 1144–1149, Mar. 2016.

[7] J. H. Weber and K. A. S. Immink, "Maximum likelihood decoding for Gaussian noise channels with gain or offset mismatch," *IEEE Commun. Lett.*, vol. 22, no. 6, pp. 1128–1131, Jun. 2018.

[8] R. Bu and J. H. Weber, "Decoding criteria and zero wer analysis for channels with bounded noise and offset," in *IEEE International Conference on Communications*, May 2019, pp. 1–6.

[9] K. Cai and K. A. S. Immink, "Cascaded channel model, analysis, and hybrid decoding for spin-torque transfer magnetic random access memory," *IEEE Trans. Magn.*, vol. 53, no. 11, pp. 1–11, Nov 2017.

[10] H. Pozidis, N. Papandreou, A. Sebastian, T. Mittelholzer, M. BrightSky, C. Lam, and E. Eleftheriou, "A framework for reliability assessment in multilevel phase-change memory," in *IEEE International Memory Workshop*, May 2012, pp. 143–146.

[11] X. Kou, J. Schmalhorst, A. Thomas, and G. Reiss, "Temperature dependence of the resistance of magnetic tunnel junctions with MgO barrier," *Applied physics letters*, vol. 88, no. 21, p. 212115, May 2006.

[12] M. Grigoriu, *Applied Non-Gaussian Processes: Examples, Theory, Simulation, Linear Random Vibration, and MATLAB Solutions*. Englewood Cliffs, NJ: PTR Prentice Hall, 1995.

[13] A. D'Onofrio, *Bounded Noises in Physics, Biology, and Engineering*. New York, NY: Springer, 2013.

[14] R. Bu, J. H. Weber, and K. A. S. Immink, "Maximum likelihood decoding for channels with Gaussian noise and signal dependent offset," manuscript submitted to IEEE Transactions on Communications.

[15] J. H. Weber, K. A. S. Immink, and S. R. Blackburn, "Pearson codes," *IEEE Trans. Inform. Theory*, vol. 62, no. 1, pp. 131–135, Jan. 2016.

[16] R. B. Nelsen, *An Introduction to Copulas*. Springer Science & Business Media, 2007.