Minimum Pearson Distance Detection for Multi-Level Channels with Gain and/or Offset Mismatch

Kees Schouhamer Immink
TURING MACHINES Inc., Rotterdam
Jos Weber
Delft University of Technology, Delft

June 3, 2014

Outline

- Introduction

Outline

- Introduction
- Channels with gain and offset mismatch

Outline

- Introduction
- Channels with gain and offset mismatch
- Prior Art: Euclidean distance detection

Outline

- Introduction
- Channels with gain and offset mismatch
- Prior Art: Euclidean distance detection
- Pearson distance detection and constrained coding

Outline

- Introduction
- Channels with gain and offset mismatch
- Prior Art: Euclidean distance detection
- Pearson distance detection and constrained coding
- Channels with offset-only mismatch

Outline

- Introduction
- Channels with gain and offset mismatch
- Prior Art: Euclidean distance detection
- Pearson distance detection and constrained coding
- Channels with offset-only mismatch
- Conclusions

Channels with gain and offset mismatch

Let $\boldsymbol{x} = (x_1, \ldots, x_n)$ be a codeword of n symbols over the alphabet $\mathcal{Q} = \{0, 1, \ldots, q-1\}$ taken from a codebook S.

In AWGN channels, it is assumed that if $\boldsymbol{x}$ is sent that $\boldsymbol{r} = \boldsymbol{x} + \boldsymbol{\nu}$ is received, where $\boldsymbol{\nu}$ is an additive Gaussian, $N(0, \sigma)$, noise term vector.

Here we assume an AWGN offset and gain *mismatch* channel, i.e.,

$$\boldsymbol{r} = a(\boldsymbol{x} + \boldsymbol{\nu}) + b,$$

where $a > 0$ and b are constants, $a > 0$.

Prior Art re gain and offset mismatch

Detection is normally very sensitive to gain and offset mismatch, and a loss of performance may occur.

Prior Art solutions

- Automatic offset and gain control based on the previous words received. (Not possible for Flash memories.)
- Insertion of reference symbols by which actual gain and offset can be estimated.
- Intrinsically resistant codes.

Prior Art: Euclidean Distance Detection

Apply Minimum Euclidean Distance Detection detection

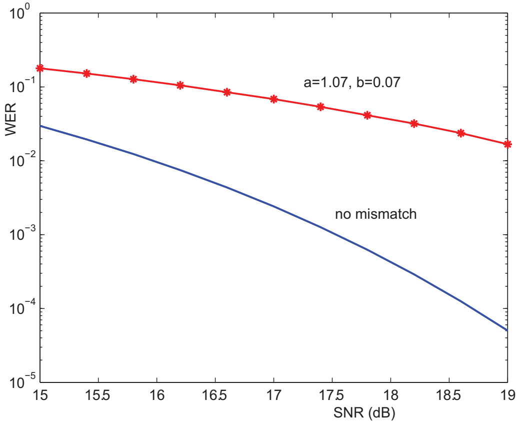
$$\boldsymbol{x}_o = \arg\min_{\hat{\boldsymbol{x}} \in S} d(\hat{\boldsymbol{x}}),$$

where

$$d(\hat{\boldsymbol{x}}) = \sum_{i=1}^{n} (r_i - \hat{x}_i)^2 = \sum_{i=1}^{n} (ax_i + b - \hat{x}_i)^2$$

$$d(\hat{\boldsymbol{x}}) = \sum (ax_i + b)^2 - 2a \sum x_i \hat{x}_i - 2b \sum \hat{x}_i + \sum \hat{x}_i^2$$

Mismatch



Word error rate of matched and mismatched channel ($a = 1.07$, $b = 0.07$), where $n = 8$ and $q = 4$.

All codewords in a *balanced code* must satisfy

$$\sum \hat{x}_i = d_c$$

and

$$\sum \hat{x}_i^2 = d_p$$

then

$$d'(\hat{\boldsymbol{x}}) = -\sum r_i \hat{x}_i,$$

is independent of a and b.

Disadvantage: High redundancy of balanced codes.

New Approach: Minimum Pearson distance detection

Apply Minimum Pearson Distance Detection detection

$$\boldsymbol{x}_o = \arg\min_{\hat{\boldsymbol{x}} \in S} \delta(\hat{\boldsymbol{x}}),$$

where

$$\delta(\hat{\boldsymbol{x}}) = 1 - \rho_{\boldsymbol{r},\hat{\boldsymbol{x}}}, \tag{1}$$

where

$$\rho_{\boldsymbol{r},\hat{\boldsymbol{x}}} = \frac{\sum_{i=1}^{n}(r_i - \overline{r})(\hat{x}_i - \overline{\hat{x}})}{\sigma_r \sigma_{\hat{x}}} \tag{2}$$

is the well-known *(Pearson) correlation coefficient*. We define

$$\overline{\hat{x}} = \frac{1}{n}\sum_{i=1}^{n}\hat{x}_i, \quad \sigma_{\hat{x}}^2 = \sum_{i=1}^{n}(\hat{x}_i - \overline{\hat{x}})^2.$$

Minimum Pearson distance detection II

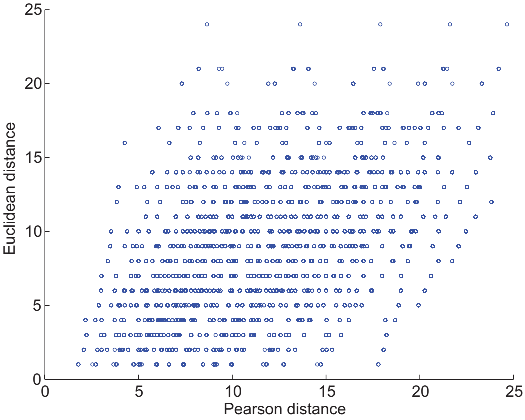
The Pearson distance, $1 - \rho_{\boldsymbol{r},\hat{\boldsymbol{x}}}$, is invariant to changes in translation or scale (up to a sign) in the two vectors $\boldsymbol{r}$ and $\hat{\boldsymbol{x}}$. That is,

$$\rho_{\boldsymbol{r},\hat{\boldsymbol{x}}} = \rho_{c_1 + c_2\boldsymbol{r},\hat{\boldsymbol{x}}},$$

where c_1 and $c_2 \geq 0$ are real constants.

Thus: Minimum Pearson distance detection is immune to offset and gain mismatch! But some problems have to be solved.

Scatter Diagram



Scatter diagram of Euclidean distance versus Pearson distance for $n = 6$ and $q = 3$.

Constrained codes

What is wrong and why do we need coding??

- The Pearson distance is undefined for codewords $\boldsymbol{x}$ with $\sigma_x = 0$.
- A Pearson-distance-based detector cannot distinguish between the words $\boldsymbol{x}$ and $c_1 + c_2\boldsymbol{x}$.

Constrained coding is needed to exclude codewords
a) $\boldsymbol{x} = c$ and
b) $c_1 + c_2\boldsymbol{x}$ if $\boldsymbol{x} \in S$.

A set of T-constrained codewords is denoted by S_T consists of codewords, where T, $0 < T \leq q$, *preferred* symbols **must** appear at least once in a codeword. The size, $|S_T|$, equals

$$|S_T| = \sum_{i=0}^{T} (-1)^i \binom{T}{T-i} (q-i)^n, \;\; n \geq T. \tag{3}$$

Definition: The set S_1 contains codewords where the symbol '0' appears at least once.

The number of constrained codewords is

$$|S_1| = q^n - (q-1)^n, q > 1.$$

Definition: The set S_2 contains codewords where both the symbols '0' and '$q-1$' appear at least once.

The number of constrained codewords is

$$|S_2| = q^n - 2(q-1)^n + (q-2)^n, q > 1.$$

For $q = 2$, we simply have

$$|S_1| = 2^n - 1 \text{ and } |S_2| = 2^n - 2.$$

The all-'0' and all-'$q-1$' words are deleted.

Properties of T-constrained codes

Property A: If $\boldsymbol{x} \in S_1$ then $\boldsymbol{x} + c \notin S_1$ for all non-zero $c \in \mathbb{R}$.
Proof: By definition, $\boldsymbol{x}$ has at least one position, say k, where $x_k = 0$. Hence if $c < 0$, then $x_k + c < 0$, and thus $\boldsymbol{x} + c \notin S_1$. If $c > 0$, then $x_i + c > 0$ for all i, and thus $\boldsymbol{x} + c \notin S_1$ since $\boldsymbol{x} + c$ does not contain the symbol '0'

Properties of T-constrained codes

Property B: If $\boldsymbol{x} \in S_2$ then $c_1 + c_2\boldsymbol{x} \notin S_2$ for all $c_1, c_2 \in \mathbb{R}$ with $(c_1, c_2) \neq (0, 1)$ and $c_2 > 0$.
Proof: Suppose $c_1 + c_2\boldsymbol{x} \in S_2$. Since $c_2 > 0$ and $x_i \geq 0$ for all i, it follows that $c_1 \leq 0$ in order to have at least one '0' in $c_1 + c_2\boldsymbol{x}$. Since $c_1 < 0$ would result in at least one negative value in $c_1 + c_2\boldsymbol{x}$ (in a position where $\boldsymbol{x}$ has a '0'), it follows that $c_1 = 0$ and thus $c_2 \neq 1$. If $0 < c_2 < 1$, then all symbols in $c_1 + c_2\boldsymbol{x} = c_2\boldsymbol{x}$ are smaller then $q - 1$, while if $c_2 > 1$, then at least one value in $c_1 + c_2\boldsymbol{x}$ is larger than $q - 1$ (in a position where $\boldsymbol{x}$ has a '$q - 1$'). In conclusion, $c_1 + c_2\boldsymbol{x} \notin S_2$.

Offset and gain mismatch

The word error rate (WER) is upperbounded by

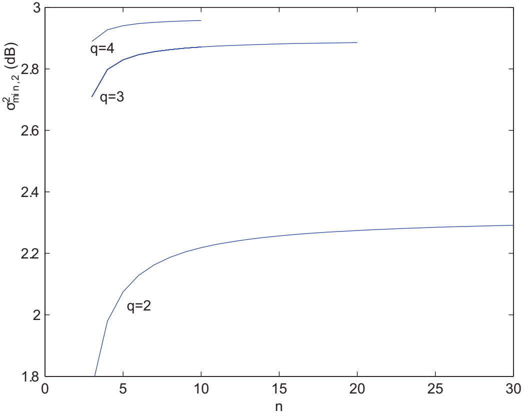
$$\mathrm{WER} < N_1 Q(\frac{d_{\min,2}}{2\sigma}), \quad \sigma << 1,$$

where

$$d^2_{\min,2} = \min_{\substack{\boldsymbol{x},\hat{\boldsymbol{x}} \in S_2 \\ \boldsymbol{x} \neq \hat{\boldsymbol{x}}}} 2\sigma_x^2(1 - \rho_{\boldsymbol{x},\hat{\boldsymbol{x}}}),$$

and N_1 is a constant.

Minimum squared distance, $d^2_{\min,2}$



Detection loss, $-20\log(d_{\min,2})$ (dB), as a function of the codeword length n with $q = 2, 3, 4$ as a parameter. Although lines have been drawn, the curves consist of discrete points.

Special case: $a = 1$

Here we assume an AWGN *offset mismatch* channel, i.e.,

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

where b a constant. We propose the distance measure

$$\delta_1(\boldsymbol{r}, \hat{\boldsymbol{x}}) = \sum_{i=1}^{n} (r_i - \hat{x}_i + \overline{\hat{x}})^2, \tag{4}$$

where, as we may notice, we have deleted $\sigma_{\hat{x}}$, in (1). We first show that $\delta_1(\boldsymbol{r}, \hat{\boldsymbol{x}})$ is, as claimed, independent of the channel's offset b

Analysis of offset mismatch case

$$\begin{aligned}\delta_1(\boldsymbol{r}, \hat{\boldsymbol{x}}) &= \sum_{i=1}^{n} (x_i' + b - \hat{x}_i + \overline{\hat{x}})^2 \\ &= \sum_{i=1}^{n} (x_i' - \hat{x}_i + \overline{\hat{x}})^2 + 2b \sum_{i=1}^{n} (x_i' - \hat{x}_i + \overline{\hat{x}}) + b^2,\end{aligned}$$

where $x_i' = x_i + \nu_i$. By definition, we have

$$\sum_{i=1}^{n} (x_i' - \hat{x}_i + \overline{\hat{x}}) = \sum_{i=1}^{n} x_i',$$

so that

$$\delta_1(\boldsymbol{r}, \hat{\boldsymbol{x}}) \equiv \sum_{i=1}^{n} (x_i' - \hat{x}_i + \overline{\hat{x}})^2. \tag{5}$$

It is immediate that the minimization of $\delta_1(\boldsymbol{r}, \hat{\boldsymbol{x}})$ is *intrinsically resistant* to the channel offset b.

Equivalence

We have

$$\boldsymbol{x}_o = \arg\min_{\hat{\boldsymbol{x}} \in S} \delta_1(\boldsymbol{r}, \hat{\boldsymbol{x}}). \tag{6}$$

A scaled (with a positive factor) or translated version of the metric is said to be *equivalent* with the original version of the metric. Metric equivalence will be denoted by the $\equiv$ sign since the outcome of (6) is unchanged. Thus,

$$\delta_1(\boldsymbol{r}, \hat{\boldsymbol{x}}) \equiv c_1\delta_1(\boldsymbol{r}, \hat{\boldsymbol{x}}) + c_2,$$

where $c_1 > 0$ and c_2 are constants.

The receiver decides that the codeword $\boldsymbol{x}_o$ was sent if (4) attains its least value for $\hat{\boldsymbol{x}} = \boldsymbol{x}_o$, that is

$$\boldsymbol{x}_o = \arg\min_{\hat{\boldsymbol{x}} \in S_1} \delta_1(\boldsymbol{r}, \hat{\boldsymbol{x}}) = \arg\min_{\hat{\boldsymbol{x}} \in S_1} \sum_{i=1}^{n} (r_i - \hat{x}_i + \overline{\hat{x}})^2. \tag{7}$$

We have $r_i = x_i + \nu_i$, so that

$$\delta_1(\boldsymbol{r}, \hat{\boldsymbol{x}}) = \sum_{i=1}^{n} (x_i + \nu_i - \hat{x}_i + \overline{\hat{x}})^2.$$

The analysis is simplified by noticing that

$$\delta_1(\boldsymbol{r}, \hat{\boldsymbol{x}}) \equiv \sum_{i=1}^{n} (x_i + \nu_i - \hat{x}_i + \bar{\hat{x}} - \bar{x})^2.$$

Define $\boldsymbol{e} = \boldsymbol{x} - \hat{\boldsymbol{x}}$ and $\bar{e} = \bar{x} - \bar{\hat{x}}$, then

$$\delta_1(\boldsymbol{r}, \hat{\boldsymbol{x}}) = \sum_{i=1}^{n} (e_i - \bar{e} + \nu_i)^2.$$

The detector errs, i.e. $\boldsymbol{x}_o \neq \boldsymbol{x}$, if there is at least one codeword $\hat{\boldsymbol{x}} \in S_1, \hat{\boldsymbol{x}} \neq \boldsymbol{x}$, such that

$$\delta_1(\boldsymbol{r}, \hat{\boldsymbol{x}}) < \delta_1(\boldsymbol{r}, \boldsymbol{x}),$$

or

$$\sum_{i=1}^{n} (e_i - \bar{e} + \nu_i)^2 < \sum_{i=1}^{n} \nu_i^2 \tag{8}$$

or

$$2\sum_{i=1}^{n}\nu_i(e_i-\overline{e})+\sum_{i=1}^{n}(e_i-\overline{e})^2<0. \tag{9}$$

The left-hand side of (9) is a stochastic variable with distribution $N(\alpha_1,\beta_1\sigma^2)$, where

$$\alpha_1=\sum_{i=1}^{n}(e_i-\overline{e})^2$$

and

$$\beta_1=4\sum_{i=1}^{n}(e_i-\overline{e})^2.$$

We define the square of the distance between the vectors $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$ by

$$d_1^2(\boldsymbol{x},\hat{\boldsymbol{x}})=\frac{4\alpha_1^2}{\beta_1}=\sum_{i=1}^{n}(e_i-\overline{e})^2.$$

The probability that $\delta_1(\boldsymbol{r}, \hat{\boldsymbol{x}}) < \delta_1(\boldsymbol{r}, \boldsymbol{x})$ equals

$$Q\left(\frac{d_1(\boldsymbol{x}, \hat{\boldsymbol{x}})}{2\sigma}\right),$$

where the Q-function is defined by

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{u^2}{2}} du.$$

The word error rate (WER) is upperbounded by

$$\mathrm{WER} < \frac{1}{|S_1|} \sum_{\boldsymbol{x} \in S_1} \sum_{\hat{\boldsymbol{x}} \neq \boldsymbol{x}} Q\left(\frac{d_1(\boldsymbol{x}, \hat{\boldsymbol{x}})}{2\sigma}\right). \quad (10)$$

Define the square of the minimum distance between any possible pair of codewords in S_1 by

$$d^2_{\min,1} = \min_{\substack{\boldsymbol{x}, \hat{\boldsymbol{x}} \in S_1 \\ \boldsymbol{x} \neq \hat{\boldsymbol{x}}}} d_1^2(\boldsymbol{x}, \hat{\boldsymbol{x}}) = \min_{\boldsymbol{e} \neq 0} \sum_{i=1}^{n} (e_i - \bar{e})^2. \quad (11)$$

Analysis V

The WER is overbounded by

$$\mathrm{WER} < N_1 Q(\frac{d_{\min,1}}{2\sigma}), \sigma << 1, \tag{12}$$

where N_1 is the average number of pairs of codewords (neighbors) at minimum distance, $d_{\min,1}$.

Analysis VI

We simply find

$$d_{\min,1} = \sqrt{1 - \frac{1}{n}}, \tag{13}$$

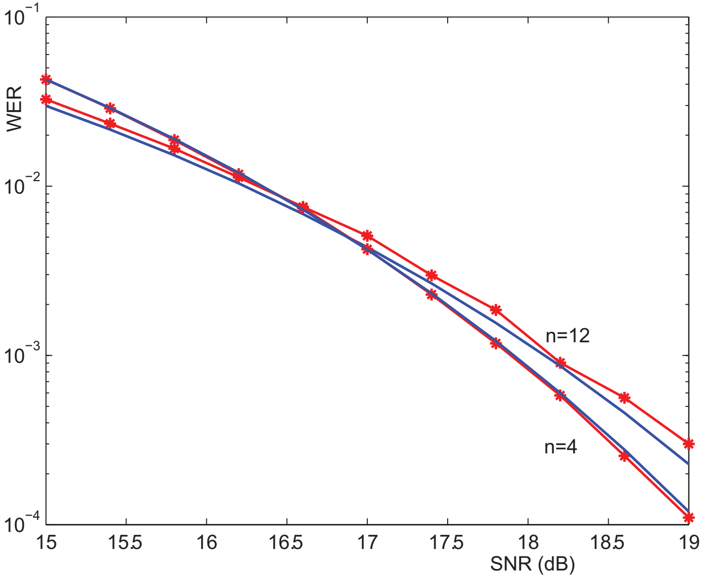
and we obtain an upperbound to the word error rate (WER)

$$\mathrm{WER} < \frac{2n(q-1)}{q} Q\left(\frac{1}{2\sigma}\sqrt{1 - \frac{1}{n}}\right). \tag{14}$$

The figure shows the WER, computed using the above upperbound (14), as a function of the signal-to-noise-ratio (SNR), where the quantity SNR is defined by

$$\mathrm{SNR(dB)} = -20 \log_{10} \sigma.$$

Performance of new system



Open problems

- Is the T-constrained code the only class of codes that can be used with Pearson detection? Are there larger sets?

Open problems

- Is the T-constrained code the only class of codes that can be used with Pearson detection? Are there larger sets?
- Can we improve the noise margin by extra coding?

Open problems

- Is the T-constrained code the only class of codes that can be used with Pearson detection? Are there larger sets?
- Can we improve the noise margin by extra coding?
- How to efficiently encode and decode T-constrained codes?

Open problems

- Is the T-constrained code the only class of codes that can be used with Pearson detection? Are there larger sets?
- Can we improve the noise margin by extra coding?
- How to efficiently encode and decode T-constrained codes?
- Can we combine the Euclidean and Pearson distance metrics?

Conclusions

- Minimum Pearson distance detection is intrinsically resistant against gain and/or offset mismatch.
- The detection is used with T-constrained codes, where in each codeword $T = 2$ the '0' and $q - 1$ symbols appear at least once.
- The redundancy of T-constrained codes is much less than that of balanced codes of the same length.
- We have analyzed the error performance in the presence of additive Gaussian noise.