Low-Complexity Detection for Balanced Codes in AWGN Channels with Offset

Antonino Favano, Luca Barletta, *Member, IEEE*, Marco Sforzin, Paolo Amato, and Marco Ferrari

***Abstract*—Low-complexity detection schemes are studied for additive white Gaussian noise channels with an unknown and unbounded offset constant over each memory read. Detectors based on the Pearson distance are analyzed, and a new lower bound on the word error rate of Modified Pearson (MP) detection is derived. Three novel detectors are proposed: the Simplified Pearson (SP), the Ultra-Simplified Pearson (USP), and the Adjusted-Threshold (AT) detectors. The USP and AT detectors are designed to be robust against destructive readings. The proposed schemes are particularly suited for memory systems employing ramp-reading architectures. The analysis demonstrates that the proposed detectors achieve competitive error-rate performance with significantly reduced complexity compared to MP detection.**

***Index Terms*—AWGN Channel with Offset, Low-Complexity Detection, Pearson Distance, Ramp Reading, Destructive Reading**

I. Introduction

Emerging memories (EM) are becoming an attractive alternative to well-established technologies such as DRAM and flash memories. Research on EMs focuses on developing solutions that overcome the limitations of conventional memories. Several EM technologies, including PCM, CeRAM, FeRAM, ReRAM, and 3D XPoint, have been proposed [2]–[6]. In addition, EMs are promising enablers for new paradigms that require real-time processing and/or the storage of massive amounts of information [2].

Data stored in a memory device is subject to various impairments arising from the physical properties of the storage medium, whether optical, magnetic, or electronic. Such impairments may be induced by factors such as temperature, humidity, or charge magnitude [7], [8]. Therefore, data retrieved from accessing a memory device typically exhibits deviations in both mean and variance relative to the nominal stored values. To mitigate these impairments, several coding schemes have been proposed [9]–[14]. However, many of these schemes incur prohibitive computational complexity for high-speed applications [15].

Whereas mature technologies such as DRAM and NAND employ well-optimized and highly reliable sensing strategies, EM systems are still in an early stage of development. Consequently, the design of fast and reliable sensing strategies is critical to enable the practical deployment of EMs in real-world applications.

By *memory read* we denote any memory access that retrieves information stored within a given set of memory cells. A widely used model for noisy memory reads is the additive white Gaussian noise (AWGN) channel with an unknown random offset [8]. Depending on the technology, the offset may remain constant or vary throughout each memory read [16]–[19]. This work focuses on the former case. For an unknown, unbounded offset that is constant during each memory read, the modified Pearson (MP) distance detector proposed in [15] achieves maximum-likelihood (ML) decoding [20]. However, as the block length increases, the complexity of Pearson-type detectors becomes a critical issue, particularly when combined with error-correcting codes, which add further computational overhead [17]. As a result, both Pearson and MP detection may be impractical for real-world implementations.

An example of a low-complexity detector for AWGN channels with unknown offset is the self-referenced reading (SRR) detector, introduced in [21]. Its efficiency is achieved through constrained coding techniques that restrict the Hamming weight spectrum of the stored codewords [22], combined with reading architectures capable of efficiently tracking memory impairments. A key limitation of SRR detection, however, is that its error-rate performance can be optimized only for a specific target signal-to-noise ratio (SNR).

A. Contributions

The first contribution of this work is an error-rate analysis of the MP detector for uncoded input. We prove that, as the block length increases, the performance of the MP detector over an AWGN channel with offset converges to that of Euclidean distanced based detection over a standard AWGN channel.

The main contribution is the design and analysis of reliable, low-complexity alternatives to Pearson-based methods and to the SRR detector [21]. The proposed solutions exploit constrained coding techniques and efficient reading architectures, achieving error-rate performance comparable to that of MP detection. In addition, we develop specialized strategies for memory systems affected by destructive reading [2]. While optimized for ramp-reading architectures, the proposed detectors can also be adapted to other reading schemes.

The first scheme, the Simplified Pearson (SP) detector, reduces the complexity of MP detection by approximating the Pearson distance metric. We then introduce the Ultra-Simplified Pearson (USP) detector, which further simplifies SP

Part of this work was presented at the 2024 IEEE International Symposium on Information Theory [1].

Antonino Favano and Luca Barletta are with the Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, 20133 Milan, Italy (e-mail: antonino.favano@polimi.it; luca.barletta@polimi.it).

Marco Sforzin and Paolo Amato are with Micron Technology Inc., 20871 Vimercate, Italy (e-mail: msforzin@micron.com; pamato@micron.com).

Marco Ferrari is with the Istituto di Elettronica e di Ingegneria dell'Informazione e delle Telecomunicazioni, Consiglio Nazionale delle Ricerche, 20133 Milan, Italy (e-mail: marcopietro.ferrari@cnr.it).

and requires only a single memory read, making it especially suitable for destructive-reading scenarios. Finally, we propose the Adjusted-Threshold (AT) detector, which is likewise robust to destructive reading and provides superior error-rate performance compared to USP.

For all of the proposed detectors, we present numerical evaluations of error-rate performance and, for SP and AT, derive analytical approximations of the error probability. Comparisons with the optimal MP detector show that the proposed schemes, despite their low complexity, closely match the performance of MP detection and consistently outperform the SRR detector [21] across all SNR regimes.

This work focuses exclusively on efficient detection techniques. However, since the proposed detectors do not require perfectly balanced constrained codes, they can also be combined with classical systematic channel coding, as shown in [23].

B. Paper Outline

The remainder of this paper is organized as follows. Section II introduces the AWGN channel model with offset, the characteristics of the targeted memory technology, and the constrained encoding strategy under consideration. Section III reviews the MP distance detector proposed in [15]. Section IV presents an error-rate analysis of MP detection. Section V introduces the SP detector as a low-complexity alternative to MP detection and evaluates its performance. The USP detector, derived as a further simplification of SP, is also discussed in this section. Section VI proposes the AT detector, another low-complexity scheme. Finally, Section VII concludes the paper.

C. Notation

Unless otherwise stated, lowercase letters denote deterministic scalars (x), bold lowercase letters denote deterministic vectors ($\mathbf{x}$), and uppercase letters denote random variables (X, $\mathbf{X}$). Given a vector $\mathbf{x}$ of length K, x_i denotes its ith component and $x_{i:K}$ its ith order statistic (i.e., the ith smallest value). The set of real numbers is denoted by $\mathbb{R}$, and the finite field of order 2 by $\mathbb{F}_2$.

The Binomial distribution with n trials and success probability p is denoted by $\mathcal{B}(n,p)$, and the Gaussian distribution with mean μ and variance σ^2 by $\mathcal{N}(\mu,\sigma^2)$. The probability density function (pdf) and cumulative distribution function (cdf) of the standard normal distribution $\mathcal{N}(0,1)$ are denoted by $\phi(x)$ and $\Phi(x)$, respectively, and the Q-function is defined as $Q(x) \triangleq 1-\Phi(x)$. Finally, $\mathbb{P}(\mathcal{A})$ denotes the probability of event $\mathcal{A}$, $\mathbb{P}_X$ and $\mathbb{E}[X]$, respectively, the distribution and the expected value of X, and $*$ the convolution operator.

II. System Model

We model memory reads as the output of a binary-input AWGN channel with random offset. The input-output relationship is:

$$\mathbf{Y} = a(\mathbf{X}+\mathbf{Z}) + b\mathbf{1}, \tag{1}$$

where $\mathbf{Y} \in \mathbb{R}^K$ is the noisy read vector of length K, $a > 0$ is an unknown channel gain, and $\mathbf{X} = m(\mathbf{C})$ is the modulated sequence derived from the stored length-K binary codeword $\mathbf{C} \in C$. The mapping $m(\mathbf{C})$ converts $0 \mapsto 1$ and $1 \mapsto -1$. The unknown offset $b \in \mathbb{R}$ is assumed constant during each read, and $\mathbf{1}$ denotes the all-ones vector. Finally, $\mathbf{Z} \in \mathbb{R}^K$ is an independent and identically distributed (iid) Gaussian vector with entries $Z_i \sim \mathcal{N}(0,\sigma^2)$, where σ^2 is the noise variance.

Throughout this work, we set $a = 1$ without loss of generality, as any other positive gain can be incorporated into the SNR.

A. Memory Features and Hamming Weight Profile

While the proposed detection schemes can be applied to various memory types, we focus on ramp-reading architectures, which, when combined with constrained coding techniques, enable computationally efficient implementations suitable for EM systems [21].

A *ramp reading architecture* applies a controlled, monotonically increasing signal (e.g., a voltage ramp) to a set of memory cells. During the ramp, each cell output is continuously monitored to identify when it crosses a critical threshold. The threshold-crossing times are proportional to the stored values, inherently providing an ordering among the read cells. For instance, given (1), ramp reading would provide an output vector consisting of the ordered entries of $\mathbf{Y}$.

The ordered reads produced by ramp architectures enable efficient implementation of Pearson-based detectors. We also develop solutions for memory technologies characterized by destructive readings, i.e., reading procedures that inherently alter or erase the stored information. Memory systems with destructive readings require detection to be performed in a single access. For example, FeRAM is affected by this limitation [24].

Under ramp sensing, detector performance depends on the Hamming weight profile of stored data [21]. Uncoded binary sequences do not guarantee a controlled Hamming weight spectrum; therefore, constrained coding can be used to define low-complexity detectors. In this work, the performance of the MP detector is analyzed assuming uncoded input. In contrast, for the newly proposed detectors, we consider a class of constrained codes called $(w_{\min}, w_{\max})$-balanced codes, i.e., codes whose codewords have Hamming weight W constrained to the interval $[w_{\min}, w_{\max}]$, where $w_{\min} \le K/2 \le w_{\max}$.

Definition 1 ($(w_{\min}, w_{\max})$-Balanced Codes)**.** Let $\mathbf{B}$ be an uncoded binary sequence of length K_u, and $\mathbf{C}$ its encoded sequence of length $K > K_u$. Let W be the Hamming weight of $\mathbf{C}$. We define the following family of distributions $\mathbb{P}_W$

$$\mathcal{S}(w_{\min}, w_{\max}) \triangleq \{\mathbb{P}_W : \operatorname{supp}(W) = \{w_{\min}, \dots, w_{\max}\}\}, \tag{2}$$

with $0 < w_{\min} \le K/2 \le w_{\max} < K$. A $(w_{\min}, w_{\max})$-balanced code maps $\mathbf{B}$ to $\mathbf{C}$ such that its Hamming weight $W \sim \mathbb{P}_W \in \mathcal{S}(w_{\min}, w_{\max})$, with support size

$$\Delta_W = |\operatorname{supp}(W)| = w_{\max} - w_{\min} + 1. \tag{3}$$

Notice that Δ_W provides a measure of the degree of *balance* of the code, with smaller values of Δ_W corresponding to a *closer* match between the numbers of 0s and 1s in each $\mathbf{C}$ in the codebook, i.e., stronger *balance*.

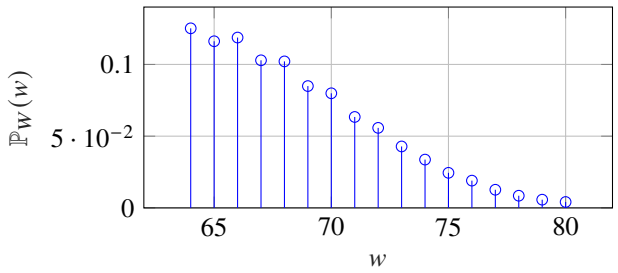

Figure 1. Hamming Weight distribution for $K = 132$ and $W \sim \mathrm{SSF}(64, 80)$.

Among the class of balanced codes, we focus on the Sub-Sequence Flipping (SSF) code from [21] due to its efficiency. For a binary vector $\mathbf{B}$ of length $K_u = 128$, the SSF encoder partitions $\mathbf{B}$ into four consecutive 32-bit subsequences. Based on four additional parity *flip* bits, each subsequence is either left unchanged or *flipped*, i.e., bitwise inverted. The *flip* bits are selected to optimize the overall balance between the numbers of 0s and 1s in the resulting codeword. For $K_u = 128$, the SSF code yields a $(64, 80)$-balanced code with codewords $\mathbf{C}$ of length $K = 132$. Let $\mathbf{B}$ be drawn uniformly at random from $\mathbb{F}_2^{K_u}$. After SSF encoding, the resulting Hamming weight W of $\mathbf{C}$ follows the distribution $W \sim \mathrm{SSF}(64, 80)$, with $\mathbb{P}_W$ shown in Fig. 1.

III. Previous Works

We introduce common detectors for the AWGN channel with an unknown, constant offset, focusing on the Pearson distance detector [25] and the MP detector [15].

A. Pearson Distance Detector

For two sequences $\mathbf{x}$ and $\mathbf{y}$, the Pearson distance is defined as

$$\delta_{\mathrm{P}}(\mathbf{x}, \mathbf{y}) = 1 - \rho_{\mathbf{x},\mathbf{y}}, \tag{4}$$

where $\rho_{\mathbf{x},\mathbf{y}} = \left(\sum_{i=1}^{K}(x_i - \bar{\mathbf{x}})(y_i - \bar{\mathbf{y}})\right)/\left(\sigma_{\mathbf{x}}\sigma_{\mathbf{y}}\right)$ is the Pearson correlation coefficient. Here, $\bar{\mathbf{x}} = \frac{1}{K}\sum_{i=1}^{K} x_i$, $\sigma_{\mathbf{x}}^2 = \sum_{i=1}^{K}(x_i - \bar{\mathbf{x}})^2$, with analogous definitions for $\mathbf{y}$.

The estimate of the stored sequence is obtained as

$$\hat{\mathbf{c}}_{\mathrm{P}}(\mathbf{Y}) = \arg\min_{\mathbf{c}\in C} \delta_{\mathrm{P}}(\mathbf{Y}, m(\mathbf{c})). \tag{5}$$

The codebook must exclude the all-0 and all-1 sequences. Indeed, suppose $\mathbf{y}$ in $\delta_{\mathrm{P}}(\mathbf{x}, \mathbf{y})$ is either the modulated all-0 or all-1 word. Then, $\rho_{\mathbf{x},\mathbf{y}}$ is not defined, as $\sigma_{\mathbf{y}} = 0$ in both cases.

B. Modified Pearson Distance Detector

The MP distance [8, Eq.(17)] improves robustness against AWGN:

$$\delta_{\mathrm{MP}}(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{K}(x_i - y_i + \bar{\mathbf{y}})^2. \tag{6}$$

Given a stored codeword $\mathbf{C}$ and its noisy observation $\mathbf{Y}$, the MP detector estimates $\mathbf{C}$ as

$$\hat{\mathbf{c}}_{\mathrm{MP}}(\mathbf{Y}) = \arg\min_{\mathbf{c}\in C} \delta_{\mathrm{MP}}(\mathbf{Y}, m(\mathbf{c})). \tag{7}$$

To guarantee unambiguous detection, it suffices for the codebook to exclude either the all-0 or all-1 sequence. Suppose again that $\mathbf{y}$ in $\delta_{\mathrm{MP}}(\mathbf{x}, \mathbf{y})$ corresponds to either the modulated all-0 or all-1 word. In both cases, we have $\delta_{\mathrm{MP}}(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{K} x_i^2$ and, therefore, the two cases cannot be distinguished, nor can a reliable estimate $\hat{\mathbf{c}}_{\mathrm{MP}}(\mathbf{Y})$ be provided. To avoid this ambiguity, we exclude the all-1 sequence and define the codebook $C = \mathbb{F}_2^K \setminus \{\mathbf{1}\}$. We assume equiprobable codewords, $\mathbb{P}_{\mathbf{C}}(\mathbf{c}) = 1/(2^K - 1)$, $\forall \mathbf{c} \in C$. It has been shown in [26] that (7) corresponds to the ML detection criterion for the considered scenario ($a = 1$, $b \in \mathbb{R}$).

C. Implementation of the Modified Pearson Distance Detector

An efficient implementation of MP detection, based on a cascade of Hamming weight estimation and Slepian detection [27], is presented in [15]. The Hamming weight is estimated as

$$\hat{W}_{\mathrm{MP}}(\mathbf{Y}) = \arg\min_{k\in[0,K-1]} \delta_k(\mathbf{Y}), \tag{8}$$

with $\delta_0(\mathbf{Y}) = 0$ and, for $k = 1, \ldots, K-1$,

$$\delta_k(\mathbf{Y}) = \sum_{i=1}^{k}\left(Y_{i:K} - \bar{\mathbf{Y}} + \frac{K+1-2i}{K}\right) \tag{9}$$

$$= \delta_{k-1}(\mathbf{Y}) + Y_{k:K} - \bar{\mathbf{Y}} + \frac{K+1-2k}{K}. \tag{10}$$

Denoting the difference between consecutive terms as

$$D_k(\mathbf{Y}) = \delta_k(\mathbf{Y}) - \delta_{k-1}(\mathbf{Y}) = Y_{k:K} - \bar{\mathbf{Y}} + \frac{K+1-2k}{K}, \tag{11}$$

the Slepian sorting algorithm assigns the estimated codewords as

$$\hat{C}_{\mathrm{MP},i} = \begin{cases} 1, & i \in \left\{j : Y_j \le Y_{\hat{W}_{\mathrm{MP}}:K}\right\}, \\ 0, & i \in \left\{j : Y_j > Y_{\hat{W}_{\mathrm{MP}}:K}\right\}. \end{cases} \tag{12}$$

Thus, the first $\hat{W}_{\mathrm{MP}}$ memory reads are detected as 1s, and the remaining reads as 0s.

Thanks to the ordered reads provided by ramp sensing, the recursive computation in (10) can be implemented particularly efficiently.

IV. Performance Analysis of the Modified Pearson Detector

We now evaluate the error-rate performance of the MP detector for the AWGN channel with an unbounded offset.

A. Word Error Rate

Let E_{MP} and $\mathbb{P}_{\epsilon,\mathrm{MP}}$ denote the number of symbol errors and the word error rate (WER), respectively, under MP detection.

Theorem 1. *The WER of MP detection over an AWGN channel with random and unbounded offset is lower-bounded by*

$$\mathbb{P}_{\epsilon,\mathrm{MP}} \ge \underline{\mathbb{P}}_{\epsilon,\mathrm{MP}} = \frac{1 - \left(1 - Q\left(\frac{1}{\sigma}\right)\right)^K - 2^{-K}}{1 - 2^{-K}}. \tag{13}$$

Proof. The MP detector cannot output the all-1 codeword since the codebook is $C = \mathbb{F}_2^K \setminus \{\mathbf{1}\}$.

Consider input codewords drawn uniformly from $C' = \mathbb{F}_2^K$, i.e., with $\mathbb{P}_{\mathbf{C}'}(\mathbf{c}) = 2^{-K}$ for all $\mathbf{c} \in C'$. Then:

$$\begin{aligned}
&\mathbb{P}_{\epsilon,\mathrm{MP}} \\
&= \mathbb{P}(\hat{\mathbf{c}}_{\mathrm{MP}}(\mathbf{Y}) \neq \mathbf{C}) && (14)\\
&= \mathbb{P}(\hat{\mathbf{c}}_{\mathrm{MP}}(\mathbf{Y}') \neq \mathbf{C}' | \mathbf{C}' \neq \mathbf{1}) && (15)\\
&= \frac{\mathbb{P}(\hat{\mathbf{c}}_{\mathrm{MP}}(\mathbf{Y}') \neq \mathbf{C}',\ \mathbf{C}' \neq \mathbf{1})}{1 - \mathbb{P}_{\mathbf{C}'}(\mathbf{1})} && (16)\\
&= \frac{\mathbb{P}(\hat{\mathbf{c}}_{\mathrm{MP}}(\mathbf{Y}') \neq \mathbf{C}') - \mathbb{P}(\hat{\mathbf{c}}_{\mathrm{MP}}(\mathbf{Y}') \neq \mathbf{1} | \mathbf{C}' = \mathbf{1})\mathbb{P}_{\mathbf{C}'}(\mathbf{1})}{1 - \mathbb{P}_{\mathbf{C}'}(\mathbf{1})} && (17)\\
&= \frac{\mathbb{P}(\hat{\mathbf{c}}_{\mathrm{MP}}(\mathbf{Y}') \neq \mathbf{C}') - \mathbb{P}_{\mathbf{C}'}(\mathbf{1})}{1 - \mathbb{P}_{\mathbf{C}'}(\mathbf{1})} && (18)\\
&\geq \frac{\mathbb{P}(\hat{\mathbf{c}}_{\mathrm{ML}}(\mathbf{Y}') \neq \mathbf{C}') - \mathbb{P}_{\mathbf{C}'}(\mathbf{1})}{1 - \mathbb{P}_{\mathbf{C}'}(\mathbf{1})} && (19)\\
&\geq \frac{\mathbb{P}(\hat{\mathbf{c}}_{\mathrm{ML}}(\mathbf{Y}' - b\mathbf{1}) \neq \mathbf{C}') - \mathbb{P}_{\mathbf{C}'}(\mathbf{1})}{1 - \mathbb{P}_{\mathbf{C}'}(\mathbf{1})} && (20)\\
&= \frac{1 - \left(1 - Q\left(\frac{1}{\sigma}\right)\right)^K - 2^{-K}}{1 - 2^{-K}}, && (21)
\end{aligned}$$

where $\mathbf{Y}' = m(\mathbf{C}') + b + \mathbf{Z}$. Step (18) follows because the MP detector never chooses the codeword $\mathbf{1}$, hence $\mathbb{P}(\hat{\mathbf{c}}_{\mathrm{MP}}(\mathbf{Y}') \neq \mathbf{1} | \mathbf{C}' = \mathbf{1}) = 1$. Step (19) arises from the definition of an ML detector; step (20) corresponds to providing b as side information; and (21) holds because $\hat{\mathbf{c}}_{\mathrm{ML}}$ reduces to the Euclidean detector for an AWGN channel without offset, with noise variance σ^2. □

An upper bound[1] on $\mathbb{P}_{\epsilon,\mathrm{MP}}$, for $\sigma \ll 1$, is derived in [8]:

$$\mathbb{P}_{\epsilon,\mathrm{MP}} < \overline{\mathbb{P}}_{\epsilon,\mathrm{MP}} = K \cdot Q\left(\frac{1}{\sigma}\sqrt{1 - \frac{1}{K}}\right). \tag{22}$$

Let the WER of a Euclidean detector over an AWGN channel without offset be

$$\mathbb{P}_{\epsilon,\mathrm{AWGN}(\sigma^2)} = 1 - (1 - Q(1/\sigma))^K. \tag{23}$$

Theorem 2. *The bounds on $\mathbb{P}_{\epsilon,\mathrm{MP}}$ also bound $\mathbb{P}_{\epsilon,\mathrm{AWGN}(\sigma^2)}$:*

$$\underline{\mathbb{P}}_{\epsilon,\mathrm{MP}} \leq \mathbb{P}_{\epsilon,\mathrm{AWGN}(\sigma^2)} \leq \overline{\mathbb{P}}_{\epsilon,\mathrm{MP}}. \tag{24}$$

Proof. The difference between $\mathbb{P}_{\epsilon,\mathrm{AWGN}(\sigma^2)}$ and $\underline{\mathbb{P}}_{\epsilon,\mathrm{MP}}$ in (13) is

$$\mathbb{P}_{\epsilon,\mathrm{AWGN}(\sigma^2)} - \underline{\mathbb{P}}_{\epsilon,\mathrm{MP}} = \left(1 - Q\left(\frac{1}{\sigma}\right)\right)^K \cdot \left(\frac{2^{-K}}{1 - 2^{-K}}\right). \tag{25}$$

Both terms on the right-hand side lie in $[0, 1]$, hence the lower bound holds. As for the upper bound, we have

$$\mathbb{P}_{\epsilon,\mathrm{AWGN}(\sigma^2)} \leq 1 - \left(1 - K \cdot Q\left(\frac{1}{\sigma}\right)\right) \leq \overline{\mathbb{P}}_{\epsilon,\mathrm{MP}}, \tag{26}$$

where the first upper bound holds by Bernoulli's inequality and the second bound by monotonicity of the Q-function, with $\sqrt{1 - 1/K} \leq 1$ for $K \geq 1$. □

In Fig. 2, we compare $\mathbb{P}_{\epsilon,\mathrm{AWGN}(\sigma^2)}$ and the WER bounds of the MP detector to numerical evaluations of $\mathbb{P}_{\epsilon,\mathrm{MP}}$, for uncoded data with $K = 16, 64, 128$, $b = 0.3$, and a Binomial Hamming weight profile $\mathcal{B}(K, 1/2)$[2]. For all K, the lower bound $\underline{\mathbb{P}}_{\epsilon,\mathrm{MP}}$ is tighter at low SNR, while for small K, the upper bound is tighter at high SNR. As K increases, the lower bound closely matches numerical evaluations of $\mathbb{P}_{\epsilon,\mathrm{MP}}$ across all SNR values.

[1] In [8, Eq. (28)], the correct scaling term for the Q-function argument is $1/\sigma$ instead of $1/(2\sigma)$.

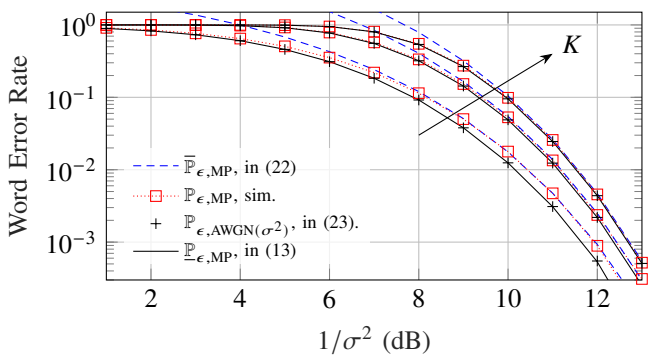

Figure 2. WER curves of bounds and Monte Carlo evaluation under MP detection with offset $b = 0.3$ and of the AWGN channel, without offset, under Euclidean detection vs SNR, for uncoded data, $K = 16, 64, 128$, and Binomial Hamming weight distribution $W \sim \mathcal{B}(K, 1/2)$.

Thanks to Theorem 2, we have that

$$|\mathbb{P}_{\epsilon,\mathrm{MP}} - \mathbb{P}_{\epsilon,\mathrm{AWGN}(\sigma^2)}| \leq \overline{\mathbb{P}}_{\epsilon,\mathrm{MP}} - \underline{\mathbb{P}}_{\epsilon,\mathrm{MP}} = G(K, \sigma). \tag{27}$$

We study this gap in the asymptotic regime where $K \to \infty$ under the constraint:

$$K \cdot Q(1/\sigma) = c, \quad c \in (0, 1/2). \tag{28}$$

In other words, we fix a constant c that deterministically ties K and σ by (28). This models systems operating at a fixed, low error probability as K increases.

Theorem 3. *In the asymptotic regime* (28)*, for sufficiently large K, the gap $G(K, \sigma)$ in* (27) *is upper-bounded by:*

$$G(K, \sigma) < U(K, \sigma) - L(K, c) \tag{29}$$

where $L(K, c) = \frac{1 - (1 - c/K)^K - 2^{-K}}{1 - 2^{-K}}$, and

$$U(K, \sigma) = \frac{c}{\sqrt{1 - 1/K}} \cdot \frac{1}{1 - \frac{1}{2\ln K - 2\ln(2\ln K)}} \cdot \left(\frac{K}{2c}\right)^{1/K}. \tag{30}$$

Moreover,

$$\lim_{K \to \infty} [U(K, \sigma) - L(K, c)] = c - (1 - e^{-c}), \tag{31}$$

which vanishes as $c \to 0$.

Proof: See Appendix A. ■

Figure 2 confirms numerically that, for large K, e.g., $K = 128$, $\mathbb{P}_{\epsilon,\mathrm{MP}} \approx \mathbb{P}_{\epsilon,\mathrm{AWGN}(\sigma^2)}$ at any SNR. Therefore, as K increases, the MP detector's error rate tends to that of a Euclidean detector for an AWGN channel without offset.

[2] We slightly abuse notation, since the all-1 word does not belong to C. To obtain the correct distribution, the probability mass at Hamming weight K can be set to zero, and the probabilities of the remaining weights adjusted accordingly.

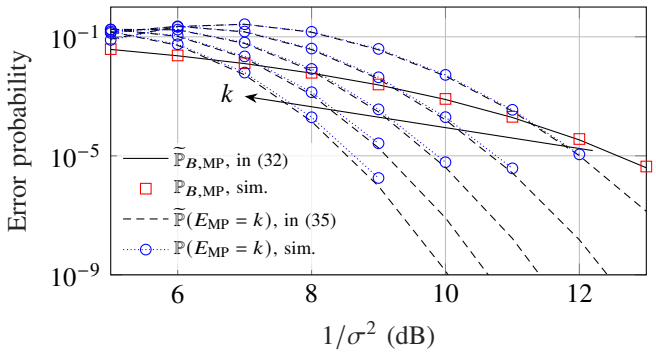

Figure 3. Error probability curves under MP detection vs SNR, for $K = 132$, $b = 0.3$, $W \sim \mathrm{SSF}(64, 80)$, and for $k = 2, \ldots, 6$.

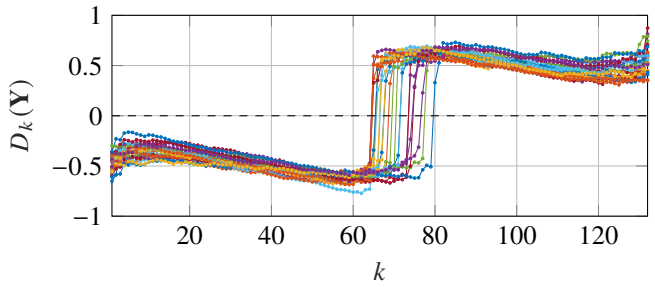

Figure 4. Examples of $D_k(\mathbf{Y})$'s, for 25 realizations of $\mathbf{Y}$ at SNR $1/\sigma^2 = 13$ dB, with $b = 0.3$ and $W \sim \mathrm{SSF}(64, 80)$.

B. Bit Error Rate

We now evaluate the bit error rate (BER) of the MP detector for uncoded input sequences. As K and SNR increase, we expect it to behave similarly to a Euclidean detector for an AWGN channel without offset.

Proposition 1. *The BER of the MP detector, $\mathbb{P}_{B,\mathrm{MP}}$, can be approximated as*

$$\mathbb{P}_{B,\mathrm{MP}} \approx \widetilde{\mathbb{P}}_{B,\mathrm{MP}} = Q\left(\frac{1}{\sigma}\right). \tag{32}$$

Proof. Let E_{MP} be the number of errors produced by the MP detector. From Theorem 3, for sufficiently large K:

$$\mathbb{P}(E_{\mathrm{MP}} = k) \approx \widetilde{\mathbb{P}}(E_{\mathrm{MP}} = k) \tag{33}$$

$$= \mathbb{E}\left[\widetilde{\mathbb{P}}(E_{\mathrm{MP}} = k \mid W)\right] \tag{34}$$

$$= \binom{K}{k} Q(1/\sigma)^k (1 - Q(1/\sigma))^{K-k}. \tag{35}$$

Therefore, the BER can be approximated by

$$\mathbb{P}_{B,\mathrm{MP}} \approx \widetilde{\mathbb{P}}_{B,\mathrm{MP}} = \frac{\sum_{k=1}^{K} k \cdot \widetilde{\mathbb{P}}(E_{\mathrm{MP}} = k)}{K} \tag{36}$$

$$= \frac{\mathbb{E}\left[\mathcal{B}\left(K, Q\left(\frac{1}{\sigma}\right)\right)\right]}{K} = Q\left(\frac{1}{\sigma}\right), \tag{37}$$

which proves (32). □

Remark 1. Consider any $(w_{\min}, w_{\max})$-balanced input code C with $w_{\min} \neq w_{\max}$. Then, at high SNR and for sufficiently large K, the error-rates of the MP detector are again approximately equal to those of the Euclidean detector for an AWGN channel without offset and with input codebook restricted to C. Moreover, if the Hamming weight distribution of $\mathbf{C} \in C$ is not highly concentrated at $w_{\min}$ and/or $w_{\max}$, (32) in Proposition 1 remains a reliable BER approximation.

To clarify Remark 1, notice that a $(w_{\min}, w_{\max})$-balanced code C, with $w_{\min} \neq w_{\max}$ necessarily has minimum Hamming distance 1. Let N_1 denote the average number of neighbors at minimum distance. Any codeword $\mathbf{C} \in C$ with Hamming weight $w_{\min} < w < w_{\max}$ has K nearest neighbors at distance one. In contrast, since codewords of weights $w_{\min} - 1$ or $w_{\max} + 1$ do not belong to C, codewords of weight $w_{\min}$ or $w_{\max}$ have fewer than K nearest neighbors. Furthermore, following [8], an upper bound on the WER for $\sigma \ll 1$ is given by $\mathbb{P}_{\epsilon,\mathrm{MP}} \leq N_1 \cdot Q(\frac{1}{\sigma}\sqrt{1 - 1/K})$. At high SNR, errors due to decisions toward minimum-distance neighbors dominate. Consequently, as $\sigma \to 0$, we obtain $\mathbb{P}_{\epsilon,\mathrm{MP}} \approx N_1 \cdot Q(\frac{1}{\sigma}\sqrt{1 - 1/K})$. For sufficiently large K, the WER under MP detection asymptotically tends to that of a Euclidean detector over an AWGN channel without offset and with input codebook restricted to C, i.e., $N_1 \cdot Q(\frac{1}{\sigma}\sqrt{1 - 1/K}) \approx N_1 \cdot Q(1/\sigma)$. Since single-bit errors dominate, the BER can be approximated as $\mathbb{P}_{B,\mathrm{MP}} \approx \mathbb{P}_{\epsilon,\mathrm{MP}}/K \approx N_1/K \cdot Q(1/\sigma)$. Finally, given that the Hamming weight distribution of $\mathbf{C} \in C$ is not highly concentrated at $w_{\min}$ and/or $w_{\max}$, we have $N_1 \approx K$ and (32) again provides a reliable approximation of the BER under MP detection.

Remark 1 applies to the SSF code, since its average number of minimum distance neighbors yields $N_1/K \approx 0.94$. Figure 3, shows the BER and error probability curves for $K = 132$ and $b = 0.3$, using the balanced code Hamming weight distribution $W \sim \mathrm{SSF}(64, 80)$. The black line is the analytical BER approximation, while the red + markers show Monte Carlo simulation results. The black dashed lines and blue dotted lines with circular markers show, respectively, the analytical approximation $\widetilde{\mathbb{P}}(E_{\mathrm{MP}} = k)$ and the numerical evaluation $\mathbb{P}(E_{\mathrm{MP}} = k)$ for $k = 2, \ldots, 6$. The close agreement between analytical and simulated results confirms the reliability of both Remark 1 and Proposition 1.

We consider the MP detector as a baseline for evaluating the performance of the proposed detectors.

V. Simplified Pearson Detector

The MP detector requires the computation of the sample average $\bar{\mathbf{Y}}$ over all reads and the evaluation of K metrics via (10). Although these metrics can be computed recursively, their evaluation still incurs non-negligible complexity. At high SNR, the function $k \mapsto \delta_k(\mathbf{Y})$ becomes sharp, meaning that the difference $D_k(\mathbf{Y})$ in (11) exhibits a sudden sign change around the true weight. Figure 4 shows examples of $D_k(\mathbf{Y})$ for 25 realizations of $\mathbf{Y}$ at SNR $1/\sigma^2 = 13$ dB, with $K = 132$, $b = 0.3$, and $W \sim \mathrm{SSF}(64, 80)$.

Definition 2 (Simplified Pearson Detector)**.** Consider a $(w_{\min}, w_{\max})$-balanced code with weight distribution and width $\Delta_W \leq \frac{K-1}{2}$. The SP detector estimates the Hamming weight as

the smallest $k \in \{w_{\min}+1, \ldots, w_{\max}+1\}$ such that $D_k(\mathbf{Y}) > 0$, i.e.,

$$\hat{W}_{\mathrm{SP}}(\mathbf{Y}) = \min\{k : D_k(\mathbf{Y}) > 0\} - 1. \tag{38}$$

$$= \min\left\{k : \left(Y_{k:K} - \bar{\mathbf{Y}} + \frac{K+1-2k}{K}\right) > 0\right\} - 1, \tag{39}$$

where (39) follows from (11). If $D_k(\mathbf{Y}) \le 0$ for all k's, the detector outputs $\hat{W}_{\mathrm{SP}}(\mathbf{Y}) = w_{\max}$. The codeword $\hat{\mathbf{C}}_{\mathrm{SP}}$ is then recovered using Slepian's sorting algorithm:

$$\hat{C}_{\mathrm{SP},i} = \begin{cases} 1, & i \in \{j : Y_j \le Y_{\hat{W}_{\mathrm{SP}}:K}\}, \\ 0, & i \in \{j : Y_j > Y_{\hat{W}_{\mathrm{SP}}:K}\}. \end{cases} \tag{40}$$

Remark 2. Balanced codes with weight distribution $\mathbb{P}_W \in \mathcal{S}(w_{\min}, w_{\max})$ and $\Delta_W \le \frac{K-1}{2}$ are not only useful for reducing complexity but also necessary for good high-SNR behavior. As $\sigma \to 0$, the metrics $D_k(\mathbf{y})$ depend deterministically on the true Hamming weight w of $\mathbf{y}$:

$$\lim_{\sigma\to 0} D_k(\mathbf{y}) = \begin{cases} -1 + \frac{1+2(w-k)}{K}, & k \le w, \\ +1 + \frac{1+2(w-k)}{K}, & k > w. \end{cases} \tag{41}$$

If the range of k is not restricted, $D_k(\mathbf{y})$ may still be positive for $k \neq w$, even as $\sigma \to 0$. For example, if $w = K/2+1$, then $\lim_{\sigma\to 0} D_k(\mathbf{y}) = \mp 1 + 1 + (2-2k)/K$, which is positive for $k = 1, w+1, \ldots, K$. To ensure reliable detection, we require $\lim_{\sigma\to 0} D_k(\mathbf{y}) \le 0$ for $k \le w$. Since (41) is proportional to $w-k$, the worst-case offset is $\Delta_W = w_{\max} - w_{\min} + 1$, leading to

$$\lim_{\sigma\to 0} D_k(\mathbf{y}) \le -1 + \frac{1+2\Delta_W}{K}, \quad k \le w. \tag{42}$$

Thus, good high-SNR detection behavior requires

$$\Delta_W \le \frac{K-1}{2}. \tag{43}$$

The complexity of computing $\delta_k(\mathbf{Y})$ and $D_k(\mathbf{Y})$ is similar. The benefit of SP detection emerges when combined with ramp sensing and balanced codes. After computing $\bar{\mathbf{Y}}$, the MP detector must still evaluate all K metrics to estimate $\hat{W}_{\mathrm{MP}}(\mathbf{Y})$. In contrast, the SP detector requires at most Δ_W metrics to obtain $\hat{W}_{\mathrm{SP}}(\mathbf{Y})$. One could reduce MP complexity by setting $\delta_{w_{\min}-1} = 0$ and restricting the evaluation to $\delta_{w_{\min}}, \ldots, \delta_{w_{\max}}$. However, thanks to ordered reads, the SP detector can stop as soon as $D_k(\mathbf{Y}) > 0$, typically after only $W - w_{\min} \le \Delta_W$ evaluations. This early stopping gives SP detection a practical edge.

A. Error-Rate Analysis of the Hamming Weight Estimation

We now analyze the performance of the SP detector by characterizing the distribution of the Hamming weight estimate. We begin with two auxiliary lemmas.

Lemma 1. *Consider two classes of independent memory reads* $\mathbf{U}^{(1)}$ *and* $\mathbf{U}^{(2)}$*, such that*

$$U_k^{(i)} \sim \mathcal{N}\left(\mu_i, \sigma^2\right), \qquad k = 1, \ldots, n_i, \tag{44}$$

for $i = 1, 2$*. Let* $\mathbf{U}$ *denote the order-statistic vector obtained by merging the samples* $\mathbf{U}^{(1)}$ *and* $\mathbf{U}^{(2)}$ *and sorting them in ascending order. Then, the cdf of the* j*th order statistic* U_j *is*

$$F_{U_j}(u) = \sum_{k=j}^{n_1+n_2} \left\{\mathcal{B}\left(n_1, \Phi^{(1)}(u)\right) * \mathcal{B}\left(n_2, \Phi^{(2)}(u)\right)\right\}(k), \tag{45}$$

where $\Phi^{(i)}$ *is the cdf of* $\mathcal{N}(\mu_i, \sigma^2)$*.*

Proof. Let V_i be the number of samples in class i that fall below a threshold u. Then $V_i \sim \mathcal{B}\left(n_i, \Phi^{(i)}\right)$, with

$$\Phi^{(i)}(u) = \mathbb{P}(U_k^{(i)} \le u) = \Phi\left(\frac{u-\mu_i}{\sigma}\right), \quad i = 1, 2. \tag{46}$$

Since V_1 and V_2 are independent, we have

$$\mathbb{P}(U_j \le u) = \sum_{k=j}^{n_1+n_2} \mathbb{P}(U_k \le u, U_{k+1} > u) \tag{47}$$

$$= \sum_{k=j}^{n_1+n_2} \mathbb{P}(V_1 + V_2 = k) \tag{48}$$

$$= \sum_{k=j}^{n_1+n_2} \left\{\mathcal{B}\left(n_1, \Phi^{(1)}(u)\right) * \mathcal{B}\left(n_2, \Phi^{(2)}(u)\right)\right\}(k). \tag{49}$$

□

Lemma 2. *Let a codeword* $\mathbf{C}$ *of Hamming weight* w *be stored and let* $\mathbf{Y}$ *denote its noisy read. The distribution of the Hamming weight estimate* $\hat{W}_{SP}(\mathbf{Y})$ *provided by the SP detector is given by*

$$\mathbb{P}(\hat{W}_{SP} = k) = \sum_{w=0}^{K-1} \mathbb{P}(W = w)\mathbb{P}(\hat{W}_{SP} = k | W = w), \tag{50}$$

where $\mathbb{P}(\hat{W}_{SP} = k | W = w) = F_{\hat{W}_{SP};w}(k+1) - F_{\hat{W}_{SP};w}(k)$ *and*

$$F_{\hat{W}_{SP};w}(k) \triangleq \mathbb{P}(\hat{W}_{SP} < k) \tag{51}$$

$$\approx 1 - \mathbb{E}\left[F_{U_k}\left(b + \bar{\mathbf{Z}} - \frac{1-2(k-w)}{K}\right)\right], \tag{52}$$

with F_{U_k} *defined in Lemma 1 for the order statistic vector* $\mathbf{U}$ *of the mixed populations*

$$\mathbf{Y}^{(-1)} = (Y_i)_{i \in \{j \in \{1,\ldots,K\}:\ X_j = -1\}}, \tag{53a}$$

$$\mathbf{Y}^{(+1)} = (Y_i)_{i \in \{j \in \{1,\ldots,K\}:\ X_j = +1\}}. \tag{53b}$$

Proof. The codeword average is $\bar{\mathbf{X}} = (K-2w)/K$, so that

$$\bar{\mathbf{Y}} = \frac{K-2w}{K} + b + \bar{\mathbf{Z}}, \tag{54}$$

with $\bar{\mathbf{Z}} = \frac{1}{K}\sum_{k=1}^{K} Z_k$ and $Z_k \sim \mathcal{N}(0, \sigma^2)$, $\forall k$. Hence $\bar{\mathbf{Z}} \sim \mathcal{N}(0, \sigma^2/K)$. Using (39), the SP estimate can be equivalently expressed as

$$\hat{W}_{\mathrm{SP}}(\mathbf{Y}) = \min\left\{k : Y_{k:K} - b - \bar{\mathbf{Z}} + \frac{1+2(w-k)}{K} > 0\right\} - 1. \tag{55}$$

Define the estimation error $\epsilon_w = \hat{W}_{\mathrm{SP}}(\mathbf{Y}) - w$. Then $\epsilon_w = \hat{k} - 1 - w$ with $\hat{k}$ satisfying the condition in (55). Its complementary cdf is approximately

$$\mathbb{P}(\epsilon_w > \epsilon) \approx \mathbb{P}\left(Y_{(w+\epsilon+1):K} - b - \bar{\mathbf{Z}} \le \frac{1+2\epsilon}{K}\right), \tag{56}$$

with $\epsilon \in \mathbb{Z} \cap [-K, K]$, under the assumption that if the condition in (55) is not satisfied for $\tilde{k}$, then it is not satisfied for any $k < \tilde{k}$ (monotonic crossing assumption).

Assuming independence between $Y_{(w+\epsilon+1):K}$ and $\bar{\mathbf{Z}}$, and applying Lemma 1, we obtain

$$\mathbb{P}(\epsilon_w > \epsilon) \approx \mathbb{P}\left(Y_{(w+\epsilon+1):K} \leq b + \bar{\mathbf{Z}} + \frac{1+2\epsilon}{K}\right) \tag{57}$$

$$= \mathbb{E}\left[F_{U_{w+\epsilon+1}}\left(b + \bar{\mathbf{Z}} + \frac{1+2\epsilon}{K}\right)\right]. \tag{58}$$

Hence, the error probability mass function is

$$P_{\epsilon_w}(\epsilon) = \mathbb{P}(\epsilon_w > \epsilon - 1) - \mathbb{P}(\epsilon_w > \epsilon). \tag{59}$$

For a general Hamming weight distribution $\mathbb{P}(W = w)$, the error probability follows as

$$\mathbb{P}(\hat{W}_{\mathrm{SP}} - w = \epsilon) = \sum_{w=0}^{K-1} P_{\epsilon_w}(\epsilon)\mathbb{P}(W = w). \tag{60}$$

Note that the distribution of $\hat{W}_{\mathrm{SP}}$ under the monotonic crossing assumption can be approximated as

$$F_{\hat{W}_{\mathrm{SP}};w}(k) = \mathbb{P}(\hat{W}_{\mathrm{SP}} < k) \approx \mathbb{P}(D_k \geq 0), \tag{61}$$

because $D_k \geq 0$ implies that $\hat{W}_{\mathrm{SP}}$ is surely set to a value less than k. Notice that the converse is also true: if $\hat{W}_{\mathrm{SP}} = t < k$, then $D_{t+1} \geq 0$, and hence $D_k \geq 0$. Therefore, the two events $\hat{W}_{\mathrm{SP}} < k$ and $D_k \geq 0$ are equivalent under the monotonic crossing assumption. Then, by (54) we have

$$F_{\hat{W}_{\mathrm{SP}};w}(k) \approx \mathbb{P}(D_k \geq 0) \tag{62}$$

$$= \mathbb{P}\left(Y_{k:K} - b - \bar{\mathbf{Z}} + \frac{1-2(k-w)}{K} \geq 0\right) \tag{63}$$

$$= 1 - \mathbb{P}\left(Y_{k:K} \leq b + \bar{\mathbf{Z}} - \frac{1-2(k-w)}{K}\right) \tag{64}$$

$$\approx 1 - \mathbb{E}\left[F_{U_k}\left(b + \bar{\mathbf{Z}} - \frac{1-2(k-w)}{K}\right)\right], \tag{65}$$

where the last step follows from (58). Finally, by noticing that $\mathbb{P}(\hat{W}_{\mathrm{SP}} = k|W = w) = F_{\hat{W}_{\mathrm{SP}};w}(k+1) - F_{\hat{W}_{\mathrm{SP}};w}(k)$, we get (50). □

B. Bit Error Rate

We now evaluate the BER of the SP detector.

Proposition 2. *Let E_{SP} denote the number of bit errors at the output of the SP detector. An approximation of the BER is*

$$\widetilde{\mathbb{P}}_{B,\mathrm{SP}} \triangleq \frac{\sum_{k=1}^{K} k \cdot \widetilde{\mathbb{P}}(E_{\mathrm{SP}} = k)}{K}, \tag{66}$$

where

$$\widetilde{\mathbb{P}}(E_{\mathrm{SP}} = k) = \mathbb{E}\Big[\widetilde{\mathbb{P}}(E_{SP} = k \mid W)\Big]. \tag{67}$$

Consider a trial with 3 possible mutually exclusive outcomes, where p_i is the probability of the ith outcome and $\sum_{i=1}^{3} p_i = 1$. We define the Trinomial distribution as

$$\mathcal{T}\begin{pmatrix} n_1 & n_2 & n_3 \\ p_1 & p_2 & p_3 \end{pmatrix} = \frac{(n_1+n_2+n_3)!}{n_1!n_2!n_3!} p_1^{n_1} p_2^{n_2} p_3^{n_3}, \tag{68}$$

where n_i is the number of times we observe the ith outcome, over $\sum_{i=1}^{3} n_i$ trials.

Then, we can define $\widetilde{\mathbb{P}}(E_{SP} = k \mid W = w)$ as

$$\widetilde{\mathbb{P}}(E_{SP} = k \mid W = w) = \sum_{i=0}^{k} \mathcal{T}\begin{pmatrix} w-i & 0 & i \\ p_{A_1} & 1-p_{A_1}-p_{A_2} & p_{A_2} \end{pmatrix} \cdot \mathcal{T}\begin{pmatrix} k-i & 0 & K-w-(k-i) \\ p_{B_1} & 1-p_{B_1}-p_{B_2} & p_{B_2} \end{pmatrix}, \tag{69}$$

with

$$p_{A_1} = \Phi\Big(\tfrac{K-1+2(k-2i)}{K\sigma_n}\Big), \quad p_{A_2} = 1 - \Phi\Big(\tfrac{K+1+2(k-2i)}{K\sigma_n}\Big), \tag{70}$$

$$p_{B_1} = \Phi\Big(\tfrac{-K-1+2(k-2i)}{K\sigma_n}\Big), \quad p_{B_2} = 1 - \Phi\Big(\tfrac{-K+1+2(k-2i)}{K\sigma_n}\Big), \tag{71}$$

and $\sigma_n = \sigma\sqrt{\frac{K-1}{K}}$.

Proof. The event $\{\hat{W}_{\mathrm{SP}} = k\}$ implies $\{D_k < 0 < D_{k+1}\}$. Assuming a single zero crossing of the sequence $\{D_k\}$, we have $\mathbb{P}(\hat{W}_{\mathrm{SP}} = k) = \mathbb{P}(D_k < 0 < D_{k+1})$. The event $\{D_k < 0 < D_{k+1}\}$ can be expressed as

$$Y_{k:K} < b + \bar{\mathbf{Z}} + \frac{2(k-W)-1}{K} < Y_{(k+1):K} - \frac{2}{K}. \tag{72}$$

This inequality defines an upper threshold for $Y_{k:K}$ and a lower threshold for $Y_{(k+1):K}$:

$$Y_{k:K} < t_{-1} \triangleq b + \bar{\mathbf{Z}} + \frac{2(k-W)-1}{K}, \tag{73}$$

$$Y_{(k+1):K} > t_{+1} \triangleq b + \bar{\mathbf{Z}} + \frac{2(k-W)+1}{K}. \tag{74}$$

Consider the classes (-1) and $(+1)$ defined in (53) and introduce the intervals $T_{-1} = (-\infty, t_{-1})$, $T_0 = [t_{-1}, t_{+1}]$, and $T_{+1} = (t_{+1}, \infty)$. For each class, the probability of observing n_1 elements in T_{-1}, n_2 elements in T_0, and n_3 elements in T_{+1} follows the Trinomial distribution in (68), where p_1, p_2, and p_3 are the probabilities that any given Y_i falls into T_{-1}, T_0, or T_{+1}, respectively. The two trinomials, corresponding to classes (-1) and $(+1)$, are independent. We focus on the case $n_2 = 0$ for both classes, i.e., no samples fall in T_0. Bit errors occur when elements of class (-1) fall in T_{+1} and elements of class $(+1)$ in T_{-1}. Observing k errors corresponds to i elements of class (-1) in T_{+1} and $k-i$ elements of class $(+1)$ in T_{-1}, for $i = 0, \ldots, k$. Summing over all such cases yields (69). □

Remark 3. When $\lim_{K\to\infty} k/K = 0$, the trinomial distributions reduce to binomials. In this regime,

$$\lim_{K\to\infty} \widetilde{\mathbb{P}}(E_{\mathrm{SP}} = k \mid W = w)$$

$$= \{\mathcal{B}(w, 1-\Phi(1/\sigma)) * \mathcal{B}(K-w, \Phi(-1/\sigma))\}(k) \tag{75}$$

$$= \{\mathcal{B}(w, 1-\Phi(1/\sigma)) * \mathcal{B}(K-w, 1-\Phi(1/\sigma))\}(k) \tag{76}$$

$$= \mathcal{B}(K, 1-\Phi(1/\sigma))(k) \tag{77}$$

$$= \mathbb{P}(E_{\mathrm{AWGN}(\sigma^2)} = k), \tag{78}$$

where $E_{\mathrm{AWGN}(\sigma^2)}$ denotes the number of symbol errors under Euclidean detection on an AWGN channel without offset and with noise variance σ^2, i.e., $E_{\mathrm{AWGN}(\sigma^2)} \sim \mathcal{B}\left(K, Q\left(\frac{1}{\sigma}\right)\right)$. Equivalently, as T_0 disappears, the error distributions of both

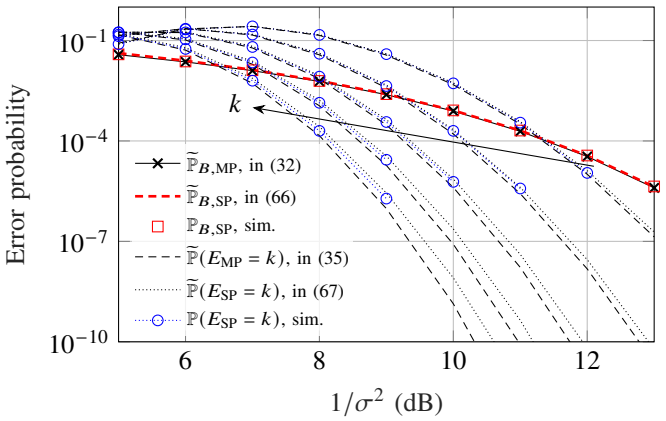

Figure 5. Error probability curves under SP detection vs SNR, for $K = 132$, $b = 0.3$, $W \sim \mathrm{SSF}(64, 80)$, and $k = 2, \ldots, 6$.

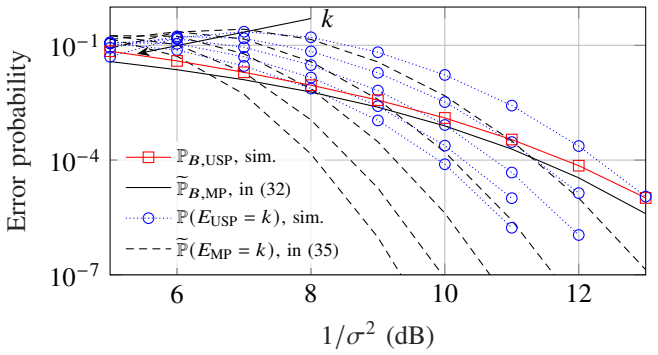

Figure 6. Error probability curves under USP detection vs SNR, for $K = 132$, $b = 0.3$, $W \sim \mathrm{SSF}(64, 80)$, and for $k = 2, \ldots, 6$.

classes reduce to identical binomials, and the total number of errors follows a single binomial law. Therefore, as K grows, the error-rate performance of the SP detector converges to that of the MP detector.

Numerical Results: Figure 5 shows the BER performance of the SP detector. The black solid line with crosses denotes the approximated BER of the MP detector. Red markers and the red dashed line correspond to the simulated and analytical approximated BER curves of the SP detector, respectively. Black dashed, black dotted, and blue dotted curves with circular markers depict, respectively, the probabilities of k errors obtained from the MP approximation, the SP approximation, and SP simulation, for $k = 2, \ldots, 6$. As already shown in Fig. 3, the analytical approximations $\mathbb{P}_{B,\mathrm{MP}}$ and $\mathbb{P}(E_{\mathrm{MP}} = k)$ match the MP simulations closely; to avoid clutter, Fig. 5 reports only $\widetilde{\mathbb{P}}_{B,\mathrm{MP}}$ and $\widetilde{\mathbb{P}}(E_{\mathrm{MP}} = k)$.

The results confirm that the SP detector achieves performance nearly indistinguishable from that of the MP detector, while the analytical approximations for the SP error probabilities are in excellent agreement with Monte Carlo simulations. Thus, the SP detector provides a lower-complexity alternative with negligible performance loss.

C. Ultra-Simplified Pearson Detector

We now introduce a further simplification of the SP detector, designed to be resistant to destructive reads. Consider again the metric difference $D_k(\mathbf{Y})$ in (11). Notice that the $D_k(\mathbf{Y})$'s depend on the arithmetic mean $\bar{\mathbf{Y}}$; therefore, the SP detector requires a first memory read to compute $\bar{\mathbf{Y}}$, followed by a second read to evaluate the $D_k(\mathbf{Y})$'s. As a result, the SP detector is not suitable for memory technologies affected by destructive reads.

Define the metric jump difference as

$$J_k(\mathbf{Y}) \triangleq D_k(\mathbf{Y}) - D_{k-1}(\mathbf{Y}) \tag{79}$$

$$= (\delta_k(\mathbf{Y}) - \delta_{k-1}(\mathbf{Y})) - (\delta_{k-1}(\mathbf{Y}) - \delta_{k-2}(\mathbf{Y})) \tag{80}$$

$$= Y_{k:K} - Y_{k-1:K} - \frac{2}{K}, \qquad k = 2, \ldots, K. \tag{81}$$

The key observation is that the $J_k(\mathbf{Y})$'s are independent of $\bar{\mathbf{Y}}$. If the estimation of the Hamming weight relies only on the $J_k(\mathbf{Y})$'s, we do not need to compute $\bar{\mathbf{Y}}$ and can therefore further simplify detection.

Definition 3 (Ultra-Simplified Pearson Detector)**.** Let us consider any $(w_{\min}, w_{\max})$-balanced code. The USP detector estimates the Hamming weight as

$$\hat{W}_{\mathrm{USP}}(\mathbf{Y}) = \underset{k \in \mathrm{supp}(W)}{\arg\max}\ J_{k+1}(\mathbf{Y}) \tag{82}$$

$$= \underset{k \in \mathrm{supp}(W)}{\arg\max}\ Y_{(k+1):K} - Y_{k:K}. \tag{83}$$

The estimated binary codeword $\hat{\mathbf{C}}_{\mathrm{USP}}$ is then obtained through Slepian's sorting algorithm:

$$\hat{C}_{\mathrm{USP},i} = \begin{cases} 1, & i \in \{j : Y_j \le Y_{\hat{W}_{\mathrm{USP}}:K}\}, \\ 0, & i \in \{j : Y_j > Y_{\hat{W}_{\mathrm{USP}}:K}\}. \end{cases} \tag{84}$$

Remark 4. USP detection can, in principle, be applied to any Hamming weight distribution, even when $\mathrm{supp}(W) = \{0, \ldots, K-1\}$. In practice, at intermediate SNR values, the hypotheses corresponding to the extremes of the weight distribution must be excluded, since in those positions the measurement $J_{k+1}(\mathbf{Y})$ is particularly noisy.

Compared to the evaluation of the MP and SP metrics, i.e., $\delta_k(\mathbf{Y})$ and $D_k(\mathbf{Y})$'s, the computational complexity is further reduced: by (83), each $J_k(\mathbf{Y})$ requires only the difference of two consecutive ordered samples. For USP detection, a theoretical error-rate analysis is significantly more involved than for the MP and SP detectors. Therefore, we provide only numerical results.

While USP detection has the advantage of being independent of $\bar{\mathbf{Y}}$ and thus suitable for destructive reads, its error-rate performance is not optimal. Figure 6 compares the error-rate of the MP and USP detectors. Specifically, we evaluate via Monte Carlo simulation the BER of the USP detector, $\mathbb{P}_{B,\mathrm{USP}}$, and the probability of k errors, $\mathbb{P}(E_{\mathrm{USP}} = k)$. The BER curves are shown for $K = 132$, $b = 0.3$, and $W \sim \mathrm{SSF}(64, 80)$. The results show that, although the USP detector does not achieve optimal performance, its error probability is still independent of the channel offset b.

VI. Adjusted-Threshold Detector

We now introduce another detection method, namely the AT detector. Similarly to the SRR detector proposed in [21],

the AT detector aims to produce and adjust an estimate of the offset b while reading the memory. As with the USP detector, the main advantage is its low complexity: it does not require the evaluation of $\bar{\mathbf{Y}}$ nor a second read. Therefore, it is suitable for memory technologies affected by destructive reads. In addition, the AT detector provides improved error-rate performance compared to both the SRR and USP detectors.

The AT detector estimates an optimal threshold to separate 0s from 1s, i.e., it estimates the offset b.

Definition 4 (Adjusted-Threshold Detector)**.** Let us consider any $(w_{\min}, w_{\max})$-balanced code. For any $\mathbf{C}$, we know that $X_{i:K} = -1$ for all $i = 1, \ldots, w_{\min}$. Using this, we estimate the threshold as

$$\hat{b} = \frac{1}{m} \sum_{i=\ell}^{\ell+m-1} (Y_{i:K} + 1) = 1 + \frac{1}{m} \sum_{i=\ell}^{\ell+m-1} Y_{i:K}, \tag{85}$$

where $1 \le m \le w_{\min}$ is the number of elements in the sum, and ℓ is chosen such that $1 \le \ell \le \ell + m - 1 \le w_{\min}$. Once $\hat{b}$ is computed, the C_i's are estimated as

$$\hat{C}_{\mathrm{AT},i} = \begin{cases} 1, & i \in \{j : Y_j < \hat{b} \vee j \le w_{\min}\}, \\ 0, & i \in \{j : Y_j \ge \hat{b} \wedge j > w_{\min}\}. \end{cases} \tag{86}$$

Remark 5. Thanks to the symmetry of the AWGN, for a codeword of weight w, the $(w/2)$th order statistic of $\mathbf{Y}$ is expected to belong to the least noisy samples of the population $\mathbf{Y}^{(-1)}$ defined in (53a). Therefore, for a given m, a suitable choice of ℓ to optimize the accuracy of $\hat{b}$ is $\ell = \lfloor w_{\min}/2 - m/2 \rfloor + 1$.

The AT detector can estimate $\hat{\mathbf{C}}_{\mathrm{AT}}$ in a single read, since $\hat{b}$ can be computed before $Y_{w_{\min}+1:K}$ is sensed. For sufficiently small values of m, the AT detector has even lower complexity than all other considered detectors. Indeed, for a given m, the evaluation of $\hat{b}$ requires only m additions and one division. Furthermore, in our experiments, even very small values of m (e.g., $m = 2$) yield error-rate performance comparable to that of the MP and SP detectors.

A. Error-Rate Analysis

We now analyze the performance of the AT detector. Given (85), if ℓ is chosen according to Remark 5, we expect a reliable estimate of $\hat{b}$ with reduced correlation to the noisier samples, i.e., the largest and smallest elements of $\mathbf{Y}^{(-1)}$, defined in (53a). In this case, it is reasonable to approximate $\hat{b}$ as independent of the errors. Moreover, the classification of 1s and 0s depends solely on the estimated threshold $\hat{b}$. As a result, the channel can be modeled as an AWGN channel with additional interference caused by the estimation error $b - \hat{b}$.

Lemma 3. *For a given choice of the parameter ℓ and m of the AT detector, the estimation error of the offset $b - \hat{b}$ is distributed as*

$$f_{b-\hat{b}}(t) = m \cdot f_{T_{\ell,m}}(-mt), \tag{87}$$

where

$$\begin{aligned} f_{T_{\ell,m}}(t) = \alpha \int_{-\infty}^{t/m} dz_1 \int_{\frac{t-z_1}{m-1}}^{t-(m-1)z_1} dz_3 \phi_\sigma(z_1) \left[\Phi_\sigma(z_1)\right]^{\ell-1} \\ \cdot \phi_\sigma(z_3)(1-\Phi_\sigma(z_3))^{w-\ell-m+1} \\ \cdot \int_{-\infty}^{\infty} e^{-i(t-z_1-z_3)x} \beta(x, z_1, z_3)^{m-2} dx, \end{aligned} \tag{88}$$

with $\alpha = \frac{w!}{2\pi(\ell-1)!(w-\ell-m+1)!(m-2)!}$ *and*

$$\beta(x, z_1, z_3) = \frac{1}{2} e^{-\frac{x^2\sigma^2}{2}} \left[erf\left(\frac{z_3 - ix\sigma^2}{\sqrt{2\sigma^2}}\right) - erf\left(\frac{z_1 - ix\sigma^2}{\sqrt{2\sigma^2}}\right) \right]. \tag{89}$$

Proof. We start with the *simplest* case where $m = 1$, i.e.,

$$\hat{b} = 1 + Y_{\ell:K} \approx 1 + Y^{(-1)}_{\ell:w} = b + Z_{\ell:w} \tag{90}$$

where the approximation assumes that the samples $\mathbf{Y}^{(+1)}$ in (53b) do not affect $Y_{\ell:K}$. The pdf of $Z_{\ell:w}$ is given by

$$f_{Z_{\ell:w}}(z) = w \binom{w-1}{\ell-1} \phi_\sigma(z) \left(\Phi_\sigma(z)\right)^{\ell-1} (1 - \Phi_\sigma(z))^{w-\ell} \tag{91}$$

where $\phi_\sigma(z) = \frac{1}{\sigma}\phi(z/\sigma)$ and $\Phi_\sigma(z) = \Phi(z/\sigma)$. The pdf of $b - \hat{b}$ is then

$$f_{b-\hat{b}}(t) = f_{Z_{\ell:w}}(-t). \tag{92}$$

For $m = 2$, we have

$$\hat{b} = 1 + \frac{Y_{\ell:K} + Y_{\ell+1:K}}{2} \tag{93}$$

$$\approx 1 + \frac{Y^{(-1)}_{\ell:w} + Y^{(-1)}_{\ell+1:w}}{2} = b + \frac{Z_{\ell:w} + Z_{\ell+1:w}}{2}. \tag{94}$$

The pdf of $M_2 \triangleq \frac{Z_{\ell:w} + Z_{\ell+1:w}}{2}$ is

$$f_{M_2}(t) = 2 \int_{-\infty}^{t} f_{Z_{\ell:w}, Z_{\ell+1:w}}(x, 2t - x) dx \tag{95}$$

$$\begin{aligned} = 2 \int_{-\infty}^{t} \frac{w!}{(\ell-1)!(w-\ell-1)!} \left[\Phi_\sigma(x)\right]^{\ell-1} \\ \cdot \left[1 - \Phi_\sigma(2t-x)\right]^{w-\ell-1} \cdot \phi_\sigma(x)\phi_\sigma(2t-x) dx. \end{aligned} \tag{96}$$

Thus, the pdf of $b - \hat{b}$ is

$$f_{b-\hat{b}}(t) = f_{M_2}(-t) \tag{97}$$

Finally, we generalize the approach to any m. Specifically,

$$\hat{b} = 1 + \frac{1}{m} \sum_{i=\ell}^{\ell+m-1} Y_{i:K} \tag{98}$$

$$\approx 1 + \frac{1}{m} \sum_{i=\ell}^{\ell+m-1} Y^{(-1)}_{i:w} = b + \frac{1}{m} \sum_{i=\ell}^{\ell+m-1} Z_{i:w}. \tag{99}$$

Next, consider

$$T_{\ell,m} = \sum_{i=\ell}^{\ell+m-1} Z_{i:w}, \tag{100}$$

which is the key to evaluating

$$f_{b-\hat{b}}(t) = m \cdot f_{T_{\ell,m}}(-mt). \tag{101}$$

To generalize the analysis for arbitrary m and ℓ, we follow the approach in [28]. Consider (100), the joint vector

$$\widetilde{\mathbf{Z}} = \left(Z_{\ell:w}, \sum_{j=\ell+1}^{\ell+m-2} Z_{j:w}, Z_{\ell+m-1:w} \right), \tag{102}$$

and

$$\check{Z} = \sum_{j=\ell+1}^{\ell+m-2} Z_{j:w} | Z_{\ell:w}, Z_{\ell+m-1:w}. \tag{103}$$

For $m > 2$, we obtain

$$f_{T_{\ell,m}}(t) = \int_{-\infty}^{t/m} \int_{\frac{t-z_1}{m-1}}^{t-(m-1)z_1} f_{\widetilde{\mathbf{Z}}}(z_1, t - z_1 - z_3, z_3) dz_3 dz_1 \tag{104}$$

$$= \int_{-\infty}^{t/m} \int_{\frac{t-z_1}{m-1}}^{t-(m-1)z_1} f_{Z_{\ell:w}, Z_{\ell+m-1:w}}(z_1, z_3) \cdot f_{\check{Z}}(t - z_1 - z_3 | z_1, z_3) dz_3 dz_1 \tag{105}$$

$$= \mathbb{E}\Big[f_{\check{Z}}(t - Z_{\ell:w} - Z_{\ell+m-1:w} | Z_{\ell:w}, Z_{\ell+m-1:w}) \Big]. \tag{106}$$

The conditional distribution is

$$f_{Z_{\ell+1:w} \cdots Z_{\ell+m-2:w} | Z_{\ell:w}, Z_{\ell+m-1:w}} = (m-2)! \prod_{j=\ell+1}^{\ell+m-2} \frac{\phi_\sigma(z_j)}{\Phi_\sigma(Z_{\ell+m-1:w}) - \Phi_\sigma(Z_{\ell:w})}, \tag{107}$$

for $Z_{\ell:w} < z_{\ell+1} < \cdots < z_{\ell+m-2} < Z_{\ell+m-1:w}$.

In other words, the set of conditional variates is equivalent to a set of unconditional order statistics of cardinality $m-2$, where the marginal pdfs are normalized over the interval $(Z_{\ell:w}, Z_{\ell+m-1:w})$. This shows that

$$f_{Z_{\ell+1:w} \cdots Z_{\ell+m-2:w} | Z_{\ell:w}, Z_{\ell+m-1:w}} = f_{\widetilde{Z}_{1:m-2} \cdots \widetilde{Z}_{m-2:m-2}},$$

where the $\widetilde{Z}_{j:m-2}$'s have the pdf

$$f_{\widetilde{Z}_{j:m-2}}(z_j) = \frac{\phi_\sigma(z_j)}{\Phi_\sigma(Z_{\ell+m-1:w}) - \Phi_\sigma(Z_{\ell:w})}, \tag{108}$$

for $z_j \in (Z_{\ell:w}, Z_{\ell+m-1:w})$. Therefore,

$$f_{T_{\ell,m}}(t) = \mathbb{E}\Big[f_{\check{Z}}(t - Z_{\ell:w} - Z_{\ell+m-1:w} | Z_{\ell:w}, Z_{\ell+m-1:w}) \Big] \tag{109}$$

$$= \mathbb{E}\Big[f_{\sum_{j=1}^{m-2} \widetilde{Z}_{j:m-2}}(t - Z_{\ell:w} - Z_{\ell+m-1:w}) \Big] \tag{110}$$

$$= \mathbb{E}\Big[f_{\sum_{j=1}^{m-2} \widetilde{Z}_j}(t - Z_{\ell:w} - Z_{\ell+m-1:w}) \Big]. \tag{111}$$

Since the $\widetilde{Z}_j$'s are iid, the pdf of $\sum_{j=1}^{m-2} \widetilde{Z}_j$ can be computed either via the $(m-2)$-fold convolution of $f_{\widetilde{Z}_1}$ or by using the characteristic function:

$$f_{\sum_{j=1}^{m-2} \widetilde{Z}_j}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \left[\theta_{\widetilde{Z}_1}(x) \right]^{m-2} dx, \tag{112}$$

where

$$\theta_{\widetilde{Z}_1}(x) = \int_{Z_{\ell:w}}^{Z_{\ell+m-1:w}} e^{ix\widetilde{z}} \frac{\phi_\sigma(\widetilde{z})}{\Phi_\sigma(Z_{\ell+m-1:w}) - \Phi_\sigma(Z_{\ell:w})} d\widetilde{z}. \tag{113}$$

Finally, we define

$$\beta(x, z_1, z_3) = \int_{z_1}^{z_3} e^{ix\widetilde{z}} \phi_\sigma(\widetilde{z}) d\widetilde{z} \tag{114}$$

$$= \frac{1}{2} e^{-\frac{x^2\sigma^2}{2}} \left[\operatorname{erf}\left(\frac{z_3 - ix\sigma^2}{\sqrt{2\sigma^2}} \right) - \operatorname{erf}\left(\frac{z_1 - ix\sigma^2}{\sqrt{2\sigma^2}} \right) \right]. \tag{115}$$

This concludes the proof of (88). □

Finally, we compute the bit error-rate statistics at the output of the hard decision.

Proposition 3. *Let E_{AT} be the random number of bit errors at the output of the AT detector. The BER under AT detection can be approximated as*

$$\widetilde{\mathbb{P}}_{B,\mathrm{AT}} = \frac{\sum_{k=1}^{K} k \cdot \widetilde{\mathbb{P}}(E_{\mathrm{AT}} = k)}{K}, \tag{116}$$

where

$$\widetilde{\mathbb{P}}(E_{AT} = k) = \mathbb{E}\Big[\widetilde{\mathbb{P}}(E_{AT} = k \mid W) \Big]. \tag{117}$$

and

$$\widetilde{\mathbb{P}}(E_{AT} = k \mid W = w) = \int_{-\infty}^{\infty} f_{b-\hat{b}}(t) \sum_{i=0}^{k} \binom{w}{i} (p_{-1})^i (1-p_{-1})^{w-i} \cdot \binom{K-w}{k-i} (p_1)^{k-i} (1-p_1)^{K-w-k+i} dt, \tag{118}$$

with $f_{b-\hat{b}}(t)$ from Lemma 3 and

$$p_{-1}(t) = 1 - \Phi_\sigma(1+t), \quad p_1(t) = \Phi_\sigma(-1+t). \tag{119}$$

Proof. By (87) in Lemma 3 and assuming independence between $\hat{b}$ and the bit errors, we obtain (118) directly. The sum in (118) accounts for the probability of observing k bit errors, across the two classes defined in (53), as a function of the offset estimate error t. The outer integral then averages this probability over the distribution of $b - \hat{b}$. Finally, (117) is obtained by averaging (118) over the distribution of W, and the BER under AT detection follows by plugging (117) into (116). □

Notice that, the complete formulations of (116) and (117) are not in closed form and require the evaluation of multiple integrals. Nonetheless, at high SNR, reliable Monte Carlo simulations become computationally prohibitive, while (116) and (117) become viable solutions to estimate the error-rate of the AT detector.

B. Numerical Results

We now demonstrate the agreement between the error-rate analysis and the Monte Carlo simulations. In Fig. 7, we compare the error-rate performance of the MP and AT detectors, for $K = 132$, $b = 0.3$, and $W \sim \mathrm{SSF}(64, 80)$. For the AT detector, we set $\ell = 32$ and $m = 2$. The approximations derived in the analysis are accurate across all SNR values. Despite its low complexity, the AT detector achieves error-rate performance close to that of the optimal

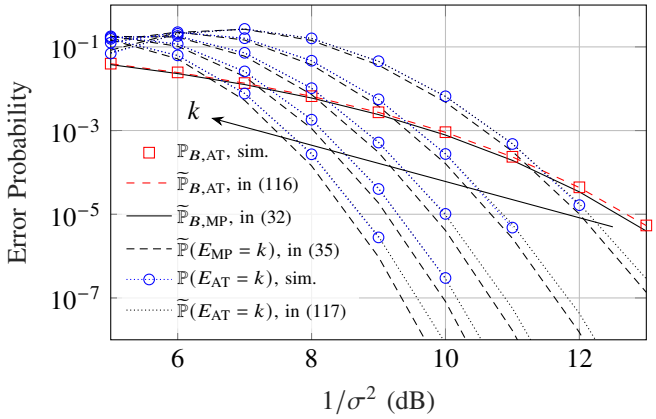

Figure 7. Error probability curves under AT detection with $m = 2$ and $\ell = 32$ vs SNR, for $K = 132$, $b = 0.3$, $W \sim \mathrm{SSF}(64, 80)$, and for $k = 2, \ldots, 6$.

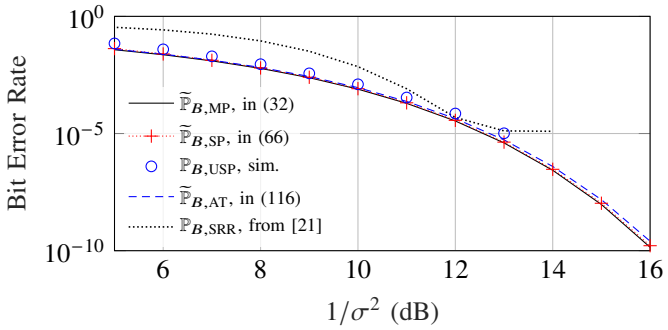

Figure 8. Bit error rate curves vs SNR for all investigated detectors, $K = 132$, $b = 0.3$, and $W \sim \mathrm{SSF}(64, 80)$.

Pearson Detector. Moreover, for $W \sim \mathrm{SSF}(64, 80)$ and a suitable choice of ℓ, our numerical simulations consistently show that, at sufficiently high SNR, the BER is independent of m. Finally, Fig. 8 compares the BER performance of the AT detector with all the other considered detection techniques, including the SRR detector from [21], for $K = 132$, $b = 0.3$, and $W \sim \mathrm{SSF}(64, 80)$. While the AT and SRR detectors have similar complexity, the SRR detector relies on precomputed parameters that depend on the target SNR, Hamming weight distribution, and the slope used in ramp sensing. Consequently, the SRR can only be optimized for a specific SNR. For example, to evaluate $\mathbb{P}_{\boldsymbol{B},\mathrm{SRR}}$ in Fig. 8, the target SNR was set to $1/\sigma^2 = 12$ (dB). As a result, the SRR detector's error-rate performance quickly deviates from the target, making it less flexible or requiring constant slope and SNR tracking to maintain optimal performance. By contrast, the AT detector provides reliable performance across all SNR levels and only requires knowledge of $w_{\min}$.

VII. Conclusion

For an additive white Gaussian noise (AWGN) channel with an unknown and unbounded offset, we derived a tight lower bound on the word error rate under modified Pearson (MP) detection. We also showed that the MP detector achieves error-rate performance closely matching that of a Euclidean detector for an AWGN channel without offset. We then proposed several new low-complexity detection schemes. First, the simplified Pearson (SP) detector is obtained from a simplification of the MP detection. Second, to address the limitations caused by destructive readings, we further specialized the SP detector and defined the ultra-simplified Pearson (USP) detector. Finally, we introduced the adjusted-threshold (AT) detector, another solution robust to destructive reads. We showed that both the SP and AT detectors, despite their lower computational complexity, retain error-rate performance close to that of the optimal MP detector. Overall, the proposed detectors provide practical and efficient building blocks for enabling reliable operation in emerging memory technologies.

Appendix A
Proof of Theorem 3

Under the constraint (28), the lower bound (21) becomes a function of only K and c:

$$L(K, c) = \frac{1 - (1 - c/K)^K - 2^{-K}}{1 - 2^{-K}}, \tag{120}$$

while the upper bound (22) can be rewritten as follows:

$$U(K, \sigma) = c \cdot \frac{Q\left(\frac{1}{\sigma}\sqrt{1 - \frac{1}{K}}\right)}{Q(1/\sigma)}. \tag{121}$$

Let $x = 1/\sigma$. We use the well-known Mill's ratio bounds:

$$\left(\frac{1}{z} - \frac{1}{z^3}\right)\frac{e^{-z^2/2}}{\sqrt{2\pi}} < Q(z) < \frac{1}{z}\frac{e^{-z^2/2}}{\sqrt{2\pi}}, \quad z > 0. \tag{122}$$

Let $z_1 = x\sqrt{1 - 1/K}$ and $z_2 = x$. Applying (122) to (121) gives:

$$\begin{aligned}\frac{Q(z_1)}{Q(z_2)} &< \frac{\frac{1}{z_1\sqrt{2\pi}}e^{-z_1^2/2}}{\left(\frac{1}{z_2} - \frac{1}{z_2^3}\right)\frac{e^{-z_2^2/2}}{\sqrt{2\pi}}} = \frac{z_2}{z_1}\frac{1}{1 - 1/z_2^2}e^{(z_2^2 - z_1^2)/2} \\ &= \frac{1}{\sqrt{1 - 1/K}} \cdot \frac{1}{1 - 1/x^2} \cdot \exp\left(\frac{x^2}{2K}\right).\end{aligned} \tag{123}$$

Let $x_K = Q^{-1}(c/K)$. The constraint implies $x = x_K$. This gives a non-elementary upper bound on U:

$$U(K, \sigma) < c \cdot \frac{1}{\sqrt{1 - 1/K}} \cdot \frac{1}{1 - 1/x_K^2} \cdot \exp\left(\frac{x_K^2}{2K}\right). \tag{124}$$

Next, let us find elementary bounds for x_K^2. Suppose $c/K < 1/2$ (so $x_K > 0$). Then, we have $c/K = Q(x_K) < \frac{1}{2}e^{-x_K^2/2}$. Rearranging the inequality yields:

$$\frac{2c}{K} < e^{-x_K^2/2} \implies \ln\left(\frac{2c}{K}\right) < -\frac{x_K^2}{2} \implies x_K^2 < 2\ln\left(\frac{K}{2c}\right).$$

As for the asymptotic behavior of x_K^2, we can write

$$\limsup_{K\to\infty} \frac{x_K^2}{2\ln K} \le \lim_{K\to\infty}\left(1 - \frac{\ln(2c)}{\ln K}\right) = 1. \tag{125}$$

For the lower bound, we use (122) with $z = x_K$:

$$\frac{c}{K} = Q(x_K) > \left(\frac{1}{x_K} - \frac{1}{x_K^3}\right)\frac{e^{-x_K^2/2}}{\sqrt{2\pi}}.$$

Taking logarithms, dividing by $\ln K$, and rearranging gives:

$$\frac{x_K^2}{2\ln K} > 1 - \frac{\ln(x_K)}{\ln K} + \frac{\ln(1-1/x_K^2)}{\ln K} - \frac{\ln(c\sqrt{2\pi})}{\ln K}.$$

As $K \to \infty$, $x_K \to \infty$. From (125), we know x_K^2 grows at most as fast as $2\ln K$, so $\ln(x_K)$ grows at most as fast as $\frac{1}{2}\ln(2\ln K)$. Therefore, $\lim_{K\to\infty} \frac{\ln(x_K)}{\ln K} = 0$. The other two fractional terms also tend to 0. Then, we have:

$$\liminf_{K\to\infty} \frac{x_K^2}{2\ln K} \geq 1.$$

Since the limit superior is ≤ 1 and the limit inferior is ≥ 1, the limit must exist and be equal to 1.

As a byproduct, note that a concrete lower bound for x_K^2 for sufficiently large K is $x_K^2 > 2\ln K - 2\ln(2\ln K)$. Combining the bounds on x_K^2 with (124) yields result (30) for sufficiently large K.

The final upper bound on the gap G for sufficiently large K is obtained by combining our results:

$$G(K,\sigma) \leq U(K,\sigma) - L(K,c). \tag{126}$$

To verify that this bound is asymptotically tight, we evaluate the limit of $U(K,\sigma)$ as $K \to \infty$:

- $\lim_{K\to\infty} \frac{c}{\sqrt{1-1/K}} = c.$
- $\lim_{K\to\infty} \frac{1}{1-\frac{1}{2\ln K - 2\ln(2\ln K)}} = \frac{1}{1-0} = 1.$
- $\lim_{K\to\infty} \left(\frac{K}{2c}\right)^{1/K} = \lim_{K\to\infty} \exp\left(\frac{\ln(K/2c)}{K}\right) = e^0 = 1.$

Thus, $\lim_{K\to\infty} U(K,\sigma) = c$ and:

$$\lim_{K\to\infty} U(K,\sigma) - L(K,c) = c - \lim_{K\to\infty} L(K,c) = c - (1-e^{-c}). \tag{127}$$

REFERENCES

[1] A. Favano, L. Barletta, M. Sforzin, P. Amato, and M. Ferrari, "Low-complexity Pearson-based detection for AWGN channels with offset," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Athens, Greece, Jul. 2024, pp. 1592–1597.

[2] R. Gastaldi and G. Campardo, *In Search of the Next Memory: Inside the Circuitry from the Oldest to the Emerging Non-Volatile Memories*. Springer, 2017.

[3] D. J. Wouters, R. Waser, and M. Wuttig, "Phase-change and redox-based resistive switching memories," *Proc. of the IEEE*, vol. 103, no. 8, pp. 1274–1288, 2015.

[4] G. Muller, N. Nagel, C.-U. Pinnow, and T. Rohr, "Emerging non-volatile memory technologies," in *European Solid-State Circuits Conf. (ESSCIRC)*. IEEE, 2003, pp. 37–44.

[5] Y. Zhou and S. Ramanathan, "Mott memory and neuromorphic devices," *Proc. of the IEEE*, vol. 103, no. 8, pp. 1289–1310, 2015.

[6] F. T. Hady, A. Foong, B. Veal, and D. Williams, "Platform storage performance with 3D XPoint technology," *Proc. of the IEEE*, vol. 105, no. 9, pp. 1822–1833, 2017.

[7] K. A. S. Immink, "A survey of codes for optical disk recording," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 4, pp. 756–764, 2001.

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

[9] A. Jiang, R. Mateescu, M. Schwartz, and J. Bruck, "Rank modulation for flash memories," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2659–2673, 2009.

[10] K. A. S. Immink and J. H. Weber, "Very efficient balanced codes," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 2, pp. 188–192, 2010.

[11] F. Sala, R. Gabrys, and L. Dolecek, "Dynamic threshold schemes for multi-level non-volatile memories," *IEEE Trans. Commun.*, vol. 61, no. 7, pp. 2624–2634, 2013.

[12] H. Zhou, M. Schwartz, A. A. Jiang, and J. Bruck, "Systematic error-correcting codes for rank modulation," *IEEE Trans. Inf. Theory*, vol. 61, no. 1, pp. 17–32, 2014.

[13] B. Peleato, R. Agarwal, J. M. Cioffi, M. Qin, and P. H. Siegel, "Adaptive read thresholds for NAND flash," *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3069–3081, 2015.

[14] K. A. S. Immink and K. Cai, "Composition check codes," *IEEE Trans. Inf. Theory*, vol. 64, no. 1, pp. 249–256, 2017.

[15] K. A. S. Immink, K. Cai, and J. H. Weber, "Dynamic threshold detection based on Pearson distance detection," *IEEE Trans. Commun.*, vol. 66, no. 7, pp. 2958–2965, 2018.

[16] K. A. S. Immink and J. H. Weber, "Hybrid minimum Pearson and Euclidean distance detection," *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3290–3298, 2015.

[17] J. H. Weber, R. Bu, K. Cai, and K. A. S. Immink, "Binary block codes for noisy channels with unknown offset," *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 3975–3983, 2020.

[18] R. Bu, J. H. Weber, and K. A. S. Immink, "Maximum likelihood decoding for channels with Gaussian noise and signal dependent offset," *IEEE Trans. Commun.*, vol. 69, no. 1, pp. 85–93, 2020.

[19] K. A. S. Immink and J. H. Weber, "A new detection method for noisy channels with time-varying offset," *IEEE Trans. Inf. Theory*, vol. 68, no. 5, pp. 3108–3114, 2022.

[20] 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, 2018.

[21] M. Sforzin, G. Mirichigni, A. Orlando, and P. Amato, "Self-referenced read methodology for EMs," in *IEEE Intern. Memory Workshop (IMW)*. IEEE, 2018, pp. 1–4.

[22] K. A. S. Immink, "Innovation in constrained codes," *IEEE Commun. Mag.*, vol. 60, no. 10, pp. 20–24, 2022.

[23] M. Ferrari, A. Favano, M. Sforzin, P. Amato, and L. Barletta, "Do we really need balanced error correcting codes?" in *SOFTCOM 2025*, Split, Croatia, Sep. 2025, pp. 1–5.

[24] J. S. Meena, S. M. Sze, U. Chand, and T.-Y. Tseng, "Overview of emerging nonvolatile memory technologies," *Nanosc. Res. Lett.*, vol. 9, no. 1, p. 526, 2014.

[25] J. H. Weber, K. A. S. Immink, and S. R. Blackburn, "Pearson codes," *IEEE Trans. Inf. Theory*, vol. 62, no. 1, pp. 131–135, 2015.

[26] R. Bu and J. H. Weber, "Maximum likelihood decoding for multi-level cell memories with scaling and offset mismatch," in *IEEE Intern. Conf. Commun. (ICC)*. IEEE, 2019, pp. 1–6.

[27] D. Slepian, "Permutation modulation," *Proc. of the IEEE*, vol. 53, no. 3, pp. 228–236, 1965.

[28] S. S. Nam, M.-S. Alouini, and H.-C. Yang, "An MGF-based unified framework to determine the joint statistics of partial sums of ordered random variables," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5655–5672, 2010.