Title	Covering Codes as Near-Optimal Quantizers for Distributed Hypothesis Testing Against Independence
Author(s)	Khaledian, Fatemeh; Asvadi, Reza; Dupraz, Elsa; Matsumoto, Tad
Citation	2024 IEEE Information Theory Workshop (ITW): 67-72
Issue Date	2024-11-24
Type	Conference Paper
Text version	author
URL	http://hdl.handle.net/10119/19412
Rights	This is the author's version of the work. Copyright (C) 2024 IEEE. 2024 IEEE Information Theory Workshop (ITW), Shenzhen, China, pp. 67-72. DOI: https://doi.org/10.1109/ITW61385.2024.10806995. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
Description	2024 IEEE Information Theory Workshop (ITW), Shenzhen, China, November 24-28, 2024

Covering Codes as Near-Optimal Quantizers for Distributed Hypothesis Testing Against Independence

Fatemeh Khaledian*, Reza Asvadi*, Elsa Dupraz†, and Tad Matsumoto†‡

*Faculty of Electrical Engineering, Shahid Beheshti University, Tehran, Iran

†IMT Atlantique, CNRS UMR 6285, Lab-STICC, Brest, France

‡JAIST and University of Oulu (Emeritus)

***Abstract*—We explore the problem of distributed Hypothesis Testing (DHT) against independence, focusing specifically on Binary Symmetric Sources (BSS). Our investigation aims to characterize the optimal quantizer among binary linear codes, with the objective of identifying optimal error probabilities under the Neyman-Pearson (NP) criterion for short code-length regime. We define optimality as the direct minimization of analytical expressions of error probabilities using an alternating optimization (AO) algorithm. Additionally, we provide lower and upper bounds on error probabilities, leading to the derivation of error exponents applicable to large code-length regime. Numerical results are presented to demonstrate that, with the proposed algorithm, binary linear codes with an optimal covering radius perform near-optimally for the independence test in DHT.**

I. Introduction

In collaborative decision-making scenarios, multiple agents or sensors send their local observations to a decision center that aims to infer a global state. Distributed Hypothesis Testing (DHT) refers to a particular case where the decision center's objective is to decide between two hypotheses, $\mathcal{H}_0$ and $\mathcal{H}_1$, leveraging coded version of a source X and a side information Y. The decision process is characterized by Type-I and Type-II error probabilities, represented by α_n and β_n, respectively. Information-theoretic DHT analyses often investigate the achievable error exponent of Type-II error probability, while satisfying a certain constraint ϵ on Type-I error probability.

Berger [1] introduced the DHT problem, which has sparked significant interest. Subsequently Ahlswede and Csiszar [2], and later Han [3], considered a two-node system represented in Fig. 1, where the first node collects data and communicates with a second node acting as the decision center over a noiseless, rate-limited link. In [2], an achievable optimal error exponent of β_n is provided for a special case of "testing against independence", which is of great interest in scenarios where detecting absence of presence of statistical dependence between sources may influence further analysis or decoding. Several other variations of DHT problems have been explored in [4]–[10].

However, practical coding perspective of DHT has received much less attention compared to the information-theoretic aspects of the problem. While certain studies, such as those by Haim and Kochman in [11], have integrated linear codes into their analysis of the error exponent, the optimal properties of such codes remain uncertain. Regarding the design of practical quantizers for DHT, an iterative alternating optimization (AO) algorithm [12] is designed, considering a criterion of distributional distance (Bhattacharyya distance), to find optimal quantizers in a multiple sensor setup, as discussed in [13]. Furthermore, [14] studied the optimal scalar quantization scheme for the distributed independence test in the case of Gaussian sources. Additionally, in [15], memoryless quantization methods are employed due to their memory-efficiency and low-latency characteristics for use in a multiple-node setup.

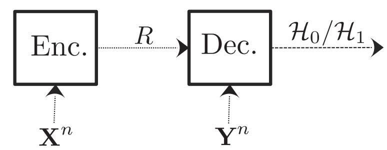


Fig. 1: Two-node system for distributed hypothesis testing problem.

Despite the efforts in optimizing local quantizers in DHT problems, the scenario of testing against independence has garnered comparatively less attention. This paper presents two main contributions. First, it explores the utilization of linear block codes as the local quantizer component, by relying on exact analytical expressions for Type-I and Type-II error probabilities provided in [16]. An iterative AO algorithm is proposed to identify the optimal characteristics of the binary local quantizer by optimizing the *coset leader spectrum* of linear block codes while also optimizing a decision rule under Neyman-Pearson (NP) criterion. Then, it derives error exponents for Type-I and Type-II error probabilities using binary linear codes as the local quantizer component. In addition to the exact error exponent derivations, to grasp a general tendency of Type-I and Type-II error exponents, upper and lower bounds of Type-I and Type-II error probabilities are derived.

II. Problem Statement

A. DHT against independence for Binary Symmetric Sources

Consider the system depicted in Fig. 1, where the first and second nodes observe random vectors $\mathbf{X}^n$ and $\mathbf{Y}^n$ of length n, respectively. As a Binary Symmetric Source (BSS) case, consider the scenario X and Y take values in a binary alphabet, and $P_X(x) = P_Y(y) \sim \text{Bern}(\frac{1}{2})$. Although the marginal distributions of X and Y are identical, Y may represent a noisy version of X. Denote $Y \triangleq X \oplus W$, where $\oplus$ is the summation over the binary field. Consider the following two hypotheses,

$$\mathcal{H}_0 : W \sim \text{Bern}\,(p_0), \tag{1}$$

$$\mathcal{H}_1 : W \sim \text{Bern}\,(p_1), \tag{2}$$

where Bern(p) denotes the Bernoulli distribution with the parameter p, and $0 \leq p \leq 1$. We also assume throughout the paper $0 \leq p_0 < \frac{1}{2}$. Obviously $\mathcal{H}_1$ indicates that X and Y are independent when $p_1 = \frac{1}{2}$. As illustrated in Fig. 1, the first node transmits a coded version of $\mathbf{X}^n$ under a communication rate R over a noiseless channel. The second node makes a decision based on the coded version of $\mathbf{X}^n$ and the side information vector $\mathbf{Y}^n$. This decision may result in two types of errors, with their probabilities defined as follows

$$\alpha_n = Pr(\mathcal{H}_1|\mathcal{H}_0), \tag{3}$$

$$\beta_n = Pr(\mathcal{H}_0|\mathcal{H}_1). \tag{4}$$

In these expressions, α_n is Type-I error probability and β_n is Type-II error probability. In [2], it is demonstrated that for testing against independence, information-theoretic quantizer-based schemes are optimal in minimizing Type-II error probability β_n under the constraint $\alpha_n \leq \epsilon$.

B. Linear block codes and useful properties

Consider a binary linear code $\mathcal{C}$ defined by a $k \times n$ generator matrix $\mathbf{G}$ with a rate of $R = \frac{k}{n}$ as the binary quantizer component [17]. According to the standard array concept [18], the minimum Hamming weight vector $d_H(\cdot)$ in each *coset* $\mathcal{C}_\mathbf{s}$ associated to the syndrome $\mathbf{s}$ is referred to as the *coset leader* defined as

$$L(\mathcal{C}_\mathbf{s}) \triangleq \arg \min_{\mathbf{z} \in \mathcal{C}_\mathbf{s}} d_H(\mathbf{z}),$$

and all coset leaders are denoted by $\mathcal{L} = \{L(\mathcal{C}_\mathbf{s}) : \text{for all possible } \mathbf{s}\}$.

Definition 1 (*covering radius* [19]): The covering radius of a code $\mathcal{C} \subseteq \{0,1\}^n$ is the smallest integer such that every vector $\mathbf{x} \in \{0,1\}^n$ is covered by a radius ρ Hamming ball centered at a point $\mathbf{c} \in \mathcal{C}$, i.e.,

$$\rho(\mathcal{C}) = \max_{\mathbf{x} \in \{0,1\}^n} \min_{\mathbf{c} \in \mathcal{C}} d_H(\mathbf{x}, \mathbf{c}).$$

Definition 2 (*coset leader spectrum*): Let $\mathcal{C}$ be a binary linear code. Its *coset leader spectrum* $\mathbf{N} = (N_0, N_1, ..., N_\rho)$ is a vector of length $\rho + 1$, where

$$N_i = |\{L \in \mathcal{L} : d_H(L) = i\}|,$$

i.e., the i-th component of $\mathbf{N}$ is the number of coset leaders of weight i in $\mathcal{L}$.

We represent a binary linear code associated with the generator matrix $\mathbf{G}$ and the covering radius ρ as $[n,k]_\rho$ throughout the paper. For an $[n,k]_\rho$, It is shown that $\sum_{i=0}^{\rho} \binom{n}{i} \geq 2^{(n-k)}$ called sphere-covering bound [19], and $\rho \leq n - k$ called Singleton bound [19]. These two inequalities provide general lower and upper bounds for the covering radius ρ, respectively.

III. Analysis of Optimum Quantizer

A. The quantizer structure under Neyman-Pearson criterion

The optimal binary quantizer for minimizing the average distortion criterion is known as Minimum Distance (MD) quantizer [20]. An MD quantizer implemented by a binary linear block code operates so that for a source vector $\mathbf{x}^n$, the quantized vector $\mathbf{u}_q^k$ is given by

$$\mathbf{u}_q^k = \arg \min_{\mathbf{u}^k \in \{0,1\}^k} d_H\left(\mathbf{x}^n, \mathbf{x}_q^n\right),$$

where $\mathbf{x}_q^n = \mathbf{u}_q^k \mathbf{G}$. In the design of the rule to decide between hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ various methods are available. Given that the probability distribution of the noise W is known under both $\mathcal{H}_0$ and $\mathcal{H}_1$, we use the NP lemma [21] which minimizes β_n under a certain constraint on α_n. By considering the BSS model as defined in Section II-A, the NP lemma provides following criterion [11]

$$\sum_{j=1}^{n} (x_{q,j} \oplus y_j) \lessgtr \gamma_t,$$

where $\gamma_t \in \mathbb{N}$ is an integer threshold chosen to minimize β_n while satisfying the constraint $\alpha_n \leq \epsilon$. Note that the symbol $<$ indicates a decision in favor of H_0, while $>$ indicates a decision in favor of H_1.

We next consider exact analytical expressions for Type-I and Type-II error probabilities provided in [16] for the MD quantizer. Specifically in the context of testing against independence, the expressions of [16] can be simplified as follows

$$\alpha_n = \frac{1}{N} \sum_{\gamma=\gamma_t+1}^{n} \sum_{i=0}^{\rho} \sum_{u=0}^{\min(\gamma,i)} \Gamma_{\gamma,i,u} p_0^j \left(1 - p_0\right)^{n-j} N_i, \tag{5}$$

$$\beta_n = \frac{1}{N} \left(\frac{1}{2}\right)^n \sum_{\gamma=0}^{\gamma_t} \sum_{i=0}^{\rho} \sum_{u=0}^{\min(\gamma,i)} \Gamma_{\gamma,i,u} N_i, \tag{6}$$

where $\Gamma_{\gamma,i,u} = \binom{i}{u}\binom{n-i}{\gamma-u}$ and $j = i + \gamma - 2u$. The total number of the *coset leaders* is denoted by $N = \sum_{i=0}^{\rho} N_i$. Additionally, for the weight of i, $\Gamma_{\gamma,i,u}$ represents the number of all possible vectors of weight γ with u 1's in common with $\mathbf{x}^n$.

B. Short code-length regime

In [16], the performance of the quantizer was evaluated for specific linear codes, notably BCH codes, with specific coset leader spectrums. Here, alternatively, our objective is to formulate an optimization problem *only under* generic constraints: Minimize β_n while $\alpha_n \leq \epsilon$. Consider $\mathbf{N} = (N_0, \ldots, N_\rho)$ as the *coset leader spectrum* of an hypothetical binary linear code satisfying the condition $N_0+\ldots+N_\rho=N$. In the following, (5) and (6) can be rewritten as

$$\alpha_n=\frac{1}{N}\sum_{i=0}^{\rho}W_{\alpha_i}N_i, \tag{7}$$

$$\beta_n=\frac{1}{N}\sum_{i=0}^{\rho}W_{\beta_i}N_i, \tag{8}$$

where

$$W_{\alpha_i}=\sum_{\gamma=\gamma_t+1}^{n}\sum_{u=0}^{\min(\gamma,i)}\Gamma_{\gamma,i,u}p_0^{j_i}\left(1-p_0\right)^{(n-j_i)}, \tag{9}$$

$$W_{\beta_i}=\left(\frac{1}{2}\right)^n\sum_{\gamma=0}^{\gamma_t}\sum_{u=0}^{\min(\gamma,i)}\Gamma_{\gamma,i,u}. \tag{10}$$

Lemma 1: With the blocklength n and an integer threshold $0\leq\gamma_t\leq n$, Type-II error probability can be expressed as:

$$\beta_n=\left(\frac{1}{2}\right)^n\sum_{\gamma=0}^{\gamma_t}\binom{n}{\gamma}. \tag{11}$$

Proof: According to (10), we have

$$W_{\beta_i}=\left(\frac{1}{2}\right)^n\sum_{\gamma=0}^{\gamma_t}\sum_{u=0}^{\min(\gamma,i)}\binom{i}{u}\binom{n-i}{\gamma-u}. \tag{12}$$

Consider the summation term with respect to u in (12), we separate it into the following two cases:
i) if $\gamma<i$, by utilizing the Vandermonde's identity [22], we have

$$\sum_{u=0}^{\gamma}\binom{i}{u}\binom{n-i}{\gamma-u}=\binom{n}{\gamma}.$$

ii) if $\gamma>i$, by using the binomial formula, we have

$$\begin{aligned}(1+x)^n&=\sum_{\gamma=0}^{n}\binom{n}{\gamma}x^\gamma,\\ (1+x)^i(1+x)^{n-i}&=\sum_{u=0}^{i}\binom{i}{u}x^u\sum_{u'=0}^{n-i}\binom{n-i}{u'}x^{u'}\sum_{\gamma=0}^{n}\delta_{\gamma,u+u'}\\ &=\sum_{\gamma=0}^{n}x^\gamma\sum_{u=0}^{i}\binom{i}{u}\sum_{u'=0}^{n-i}\binom{n-i}{u'}\delta_{u',\gamma-u}\\ &=\sum_{\gamma=0}^{n}x^\gamma\sum_{u=0}^{i}\binom{i}{u}\binom{n-i}{\gamma-u}\Big|_{0\leq\gamma-u\leq n-i},\end{aligned}$$

where $\delta_{(\cdot,\cdot)}$ denotes the Kronecker delta function. Then, the following equality is derived:

$$\sum_{\substack{u=0\\ \gamma-(n-i)\leq u\leq\gamma}}^{i}\binom{i}{u}\binom{n-i}{\gamma-u}=\binom{n}{\gamma}, \tag{13}$$

given that $\gamma-(n-i)=\gamma-n<0$ when $i=0$ is the minimum possible value for the summation index, the summation in (13) is properly bounded. Therefore, we can express (8) as,

$$\begin{aligned}\beta_n&=\frac{1}{N}\left(\frac{1}{2}\right)^n\sum_{i=0}^{\rho}\sum_{\gamma=0}^{\gamma_t}\binom{n}{\gamma}N_i\\ &=\left(\frac{1}{2}\right)^n\sum_{\gamma=0}^{\gamma_t}\binom{n}{\gamma}.\end{aligned}$$

■

According to Lemma 1, minimizing β_n is equivalent to minimizing the decision threshold γ_t. First of all, suppose our goal is to satisfy the constraint $\alpha_n\leq\epsilon$ for a possible minimum threshold γ_t. In this case, we can formulate a minimization problem as follows

$$\min_{\{\mathbf{N}_i\}}\ e=\sum_{i=1}^{\rho}W_{\alpha_i}N_i, \tag{14}$$

subject to,

$$\begin{cases}0\leq N_i\leq\binom{n}{i}\ ;\ 1\leq i\leq\rho & \text{(14a)}\\ \sum_{i=1}^{\rho}N_i=N-1, & \text{(14b)}\end{cases}$$

where W_{α_i} is defined as in (9). In this formulation, (14a) comes from the finite number of vectors with a length of n and a weight of i, additionally (14b) stems from the fact $N_0=1$[1].

By solving the minimization problem outlined in (14), (14a) and (14b), aimed at optimizing the *coset leader spectrum* $\mathbf{N}^*$ of an hypothetical binary linear code $[n,k]_{\rho^*}$, we can meet the constraint $\alpha_n\leq\epsilon$. Notably, the problem described in (14), along with constraints (14a) and (14b), falls into the category of integer linear programming (ILP) problems.

Corollary 1: Given an $[n,k]_{\rho^*}$, there exists a polynomial-time algorithm that finds an optimal solution for the ILP problem described in (14), with the constraints (14a) and (14b). The solution is a vector with components, at most, $6\left(n-k\right)^3\Phi$.[2]

The *proof* is directly derived from [23, Theorem 16.2] and [23, Corollary 17.1c], according to the Singleton bound. ■

To determine the optimal decision threshold γ_t^* while ensuring that the constraint $\alpha_n\leq\epsilon$ is satisfied, we propose an AO algorithm, outlined in Algorithm 1.

C. Large code-length regime

Following the discussions in the previous Section for the local quantizers built from linear block codes first, we then investigate characteristics of the optimal quantizer for a large code-length regime.

Proposition 1: Let the DHT problem defined as in Sections II and III, Type-I and Type-II error probabilities are bounded with the exponential forms with the following exponents

$$E_0=D_b\left(\frac{\gamma_t}{n}||p_0\right), \tag{15}$$

$$E_1=1-H_b\left(\frac{\gamma_t}{n}\right), \tag{16}$$

with $p_0\leq\frac{\gamma_t}{n}\leq\frac{1}{2}$. Here, $D_b\left(p||q\right)$ represents the binary Kullback-Leibler divergence between the probability pair

[1]Particularly, N_0 represents the all-zero codeword of the binary linear code considered, hence $N_0=1$.

[2]Φ represents the facet complexity of a rational polyhedron associated with the ILP problem, defined by $P=\{x\mid Ax\leq b\}$, where it is equal to the maximum row size of the matrix $[A\ \ b]$ [23].

Algorithm 1: Coset Leader Spectrum Optimization

Procedure Integer Linear Programming (n, k, p_0, ϵ)
Set $\gamma_t = 0$;
Define the ILP problem in (14), (14a) and (14b);
Solve the ILP for $\gamma_t = 0$;
Compute α_n from (7);
if $\alpha_n \le \epsilon$ **then**
  return $\gamma_t^* = 0$, $\mathbf{N}^*$;
end if
else
  $\gamma_t = \gamma_t + 1$;
  while $\alpha_n > \epsilon$ **do**
    Update $\mathbf{W}_\alpha$ and initialize $\mathbf{N}_i$;
    Solve the ILP;
    Update α_n;
    $\gamma_t = \gamma_t + 1$;
  end while
end if
return γ_t^*, $\mathbf{N}^*$;
end Procedure

(p, q), where $(p, q) \in [0, 1]$, and $H_b(\cdot)$ denotes the binary entropy function.

Proof: We first derive an upper bound for Type-I error probability. According to Vandermond's identity [22] $\binom{i}{u}\binom{n-i}{\gamma-u} \le \binom{n}{\gamma}$ and using (5) we find

$$\begin{aligned}
\alpha_n &\le \frac{1}{N} \sum_{\gamma=\gamma_t+1}^{n} \sum_{i=0}^{\rho} \sum_{u=0}^{\min(\gamma,i)} \binom{n}{\gamma} p_0^j (1-p_0)^{n-j} N_i \\
&= \frac{1}{N} \sum_{\gamma=\gamma_t+1}^{n} \binom{n}{\gamma} p_0^\gamma (1-p_0)^{n-\gamma} \\
&\times \sum_{i=0}^{\rho} \left(\frac{p_0}{1-p_0}\right)^i N_i \sum_{u=0}^{\min(\gamma,i)} \left(\frac{1-p_0}{p_0}\right)^{2u} \\
&\overset{(a)}{=} \frac{1}{N} \sum_{\gamma=\gamma_t+1}^{n} \binom{n}{\gamma} p_0^\gamma (1-p_0)^{n-\gamma} \\
&\underbrace{\times \sum_{i=0}^{\rho} \left(\frac{p_0}{1-p_0}\right)^i \left[\frac{\left(\frac{1-p_0}{p_0}\right)^{2\min(\gamma,i)} - 1}{\left(\frac{1-p_0}{p_0}\right)^2 - 1}\right]}_{B} N_i,
\end{aligned}$$

where (a) originates from substituting the inner summation with the sum of a geometric sequence with a common ratio $\left(\frac{1-p_0}{p_0}\right)^2 = \kappa$. On the other hand, let

$$\Delta \triangleq \sum_{i=0}^{\rho} \left(\frac{1}{\kappa}\right)^{\frac{i}{2}} \left[\kappa^i - 1\right] N_i,$$

hence we have $B \overset{(b)}{\le} \Delta$ where (b) arises from $1 \le \kappa$ and the property of the minimum function $\min(\gamma, i) \le i$. With the upper-bounding expression for B

$$\alpha_n \overset{(c)}{\le} \frac{\Delta}{N} 2^{-nD_b\left(\frac{1+\gamma_t}{n}||p_0\right)},$$

where (c) is derived from [24, Lemma 4.7.2] under the conditions $np_0 - 1 < \gamma_t < n - 1$. Using Stirling's bounds [25] as $\sqrt{2\pi} n^{\left(n+\frac{1}{2}\right)} e^{-n} \le n! \le n^{\left(n+\frac{1}{2}\right)} e^{1-n}$, we can write

$$\begin{aligned}
\alpha_n &= \frac{1}{N} \sum_{\gamma=\gamma_t+1}^{n} \binom{n}{\gamma} p_0^\gamma (1-p_0)^{n-\gamma} \sum_{i=0}^{\rho} \left(\frac{p_0}{1-p_0}\right)^i N_i \\
&\times \sum_{u=0}^{\min(\gamma,i)} \left(\frac{1-p_0}{p_0}\right)^{2u} \frac{\binom{i}{u}\binom{n-i}{\gamma-u}}{\binom{n}{\gamma}} \\
&\overset{(d)}{\ge} \frac{4\pi^2 e^{-3}}{N} \sum_{\gamma=\gamma_t+1}^{n} \binom{n}{\gamma} p_0^\gamma (1-p_0)^{n-\gamma} \sum_{i=0}^{\rho} \left(\frac{p_0}{1-p_0}\right)^i N_i \\
&\times \sum_{u=0}^{\min(\gamma,i)} \left(\frac{1-p_0}{p_0}\right)^{2u} \sqrt{\frac{\gamma}{n.i}} 2^{-nH_b\left(\frac{\gamma}{n}\right)} \\
&\overset{(e)}{\ge} \frac{\Delta'}{N} \left[8(\gamma_t+1)\left(1 - \frac{\gamma_t+1}{n}\right)\right]^{-\frac{1}{2}} 2^{-nD_b\left(\frac{\gamma_t+1}{n}\|p_0\right)},
\end{aligned}$$

where $\Delta' = 4\pi^2 e^{-3} \sum_{i=0}^{\rho} \left(\frac{1}{i}\right)^{\frac{1}{2}} \left(\frac{p_0}{1-p_0}\right)^{i+2} N_i$. (d) comes from Jensen's inequality [26] for $H_b(\cdot)$ and the conditions $\gamma \le n$, $u \le i$. Additionally (e) arises the lower bound in [24, Lemma 4.7.2]. According to Lemma 1 and by employing a similar mathematical manipulation, we establish lower and upper bounds for Type-II error probability as follows

$$\left[8\gamma_t\left(1 - \frac{\gamma_t}{n}\right)\right]^{-\frac{1}{2}} 2^{-nD_b\left(1-\frac{\gamma_t}{n}||\frac{1}{2}\right)} \le \beta_n \le 2^{-nD_b\left(1-\frac{\gamma_t}{n}||\frac{1}{2}\right)},$$

such that for sufficiently n

$$\begin{aligned}
E_0 &= -\left\{\frac{1}{n}\log\alpha_n\right\} && ; \quad n \to \text{large} \\
&= D_b\left(\frac{\gamma_t}{n}||p_0\right) && ; \quad p_0 \le \frac{\gamma_t}{n} \le 1
\end{aligned}$$

and

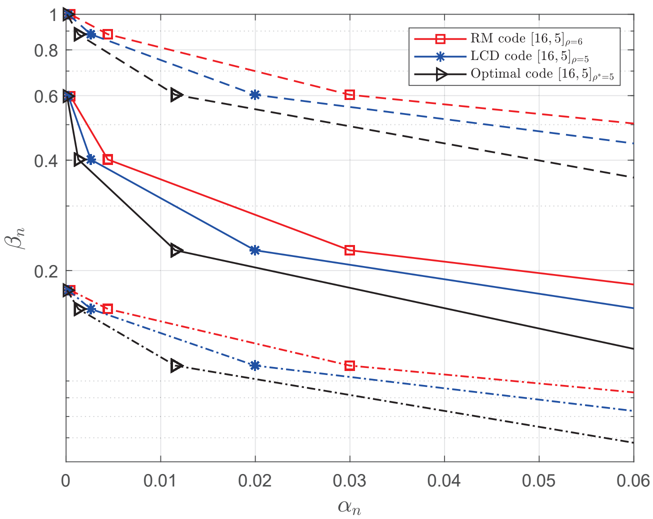
$$\begin{aligned}
E_1 &= -\left\{\frac{1}{n}\log\beta_n\right\} && ; \quad n \to \text{large} \\
&= D_b\left(1 - \frac{\gamma_t}{n}||\frac{1}{2}\right) \\
&\overset{(f)}{=} 1 - H_b\left(\frac{\gamma_t}{n}\right) && ; \quad 0 \le \frac{\gamma_t}{n} \le \frac{1}{2}
\end{aligned}$$

where (f) arises from the symmetry property of the binary entropy function $H_b(\cdot)$, when $\frac{\gamma_t}{n} \le \frac{1}{2}$.[1] ■

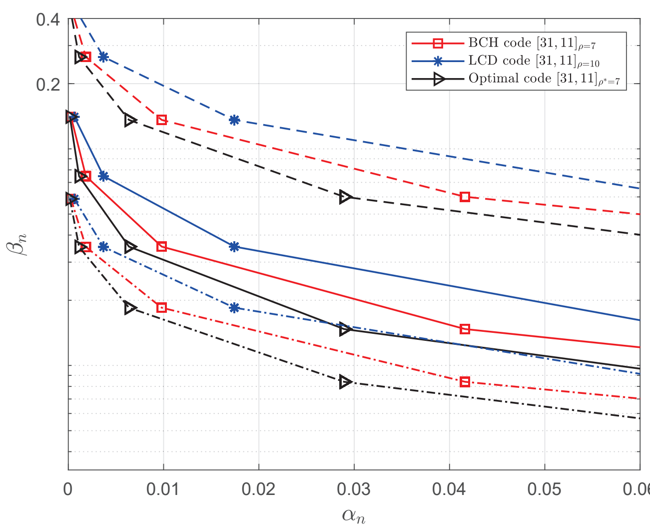
Corollary 2: Considering the DHT problem defined in Sections II and utilizing an $[n, k]_\rho$ as the local quantizer with a sufficiently large blocklength n, the *only* effective parameter influencing Type-I and Type-II error probabilities is the *normalized decision threshold* $\frac{\gamma_t}{n}$, when $p_0 \le \frac{\gamma_t}{n} \le \frac{1}{2}$.

proof: The proof directly follows from Proposition 1. ■

[1] a detailed proof is available at http://arxiv.org/abs/2410.15839.



(a) ROC curve for Reed-Muller (RM) code $[16, 5]_{\rho=6}$, linear complementary dual (LCD) code $[16, 5]_{\rho=5}$, and the optimization result code $[16, 5]_{\rho^*=5}$.



(b) ROC curve for BCH code $[31, 11]_{\rho=7}$, linear complementary dual (LCD) code $[31, 11]_{\rho=10}$, and the optimization result code $[31, 11]_{\rho^*=7}$.

Fig. 2: ROC curves: exact values (solid lines), lower bounds (dash-dotted lines), and upper bounds (dashed lines).

IV. Numerical Results

We set $\rho = n - k$ to test Algorithm 1, according to the Singleton bound. For certain values of n, k provided in Table I, the optimal solution precisely matches the *coset leader spectrum* of existing linear codes[1]. It should be noticed that, however, in some other cases, the solutions do *not* correspond, to our best knowledge, to the *coset leader spectrum* of any known linear block code, indicating that optimal DHT quantizers should not necessarily be based on existing linear codes. We investigate Type-II error probability, given the constraint $\epsilon \leq 0.06$ on Type-I error probability and $p_0 = 0.05$. Fig. 2a and Fig. 2b

[1]Remarkably, these codes exhibit an optimal covering radius.

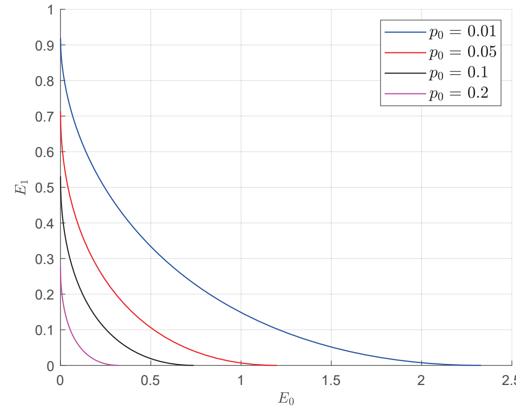


Fig. 3: The tradeoff between Type-I and Type-II in terms of E_0 and E_1.

TABLE I. COSET LEADER SPECTRUM OBTAINED AS SOLUTIONS, IN THE INDICATED CASES, OF ALGORITHM 1

Optimal linear block code	n	k	$\mathbf{N}^*$
Hamming	7	4	$(1, 7)$
Reed-Muller	8	4	$(1, 8, 7)$
Hamming	15	11	$(1, 15)$
Golay	23	12	$(1, 23, 253, 1771)$
Extended Golay	24	12	$(1, 24, 276, 2024, 1771)$
Hamming	31	26	$(1, 31)$

provide Type-I versus Type-II error probabilities, referred to Reciever Operating Characteristic (ROC) curve, where, the lower and upper bounds are shown using Proposition 1; Fig. 2a considers the Reed-Muller code $[16, 5]_{\rho=6}$ as well as the linear complementary dual (LCD) code $[16, 5]_{\rho=5}$ compared to hypothetical linear code $[16, 5]_{\rho^*=5}$ obtained by Algorithm 1. Similarly, Fig. 2b presents the results for the case of the BCH code $[31, 11]_{\rho=7}$ and the LCD code $[31, 11]_{\rho=10}$ compared to the hypothetical linear code $[31, 11]_{\rho^*=7}$ derived from Algorithm 1.The numerical results demonstrate that as the covering radius ρ of $[n, k]_\rho$ decreases, we approach $[n, k]_{\rho^*}$ with ρ^* satisfying the equality in the Sphere-covering bound [27, Table I]. It is noteworthy that the gap between the lower and upper bounds for $[n, k]_{\rho^*}$ and $[n, k]_\rho$ with equal ρ is attributed to the difference in the decision thresholds from the AO algorithm and the thresholds chosen according to the NP criterion, respectively. Moreover, it is observed that as the code length n increases from $n = 16$ in Fig. 2a to $n = 31$ in Fig. 2b, the gap between the lower and upper bounds and the exact Type-II error probability decreases.

In Fig. 3, the results are generated utilizing Proposition 1 for $p_0 \leq \frac{\gamma_t}{n} \leq \frac{1}{2}$. The figure indicates that the smaller p_0 value, i.e., the smaller value of the lower bound for $\frac{\gamma_t}{n}$, the larger exponents for Type-I and Type-II error probabilities.

Acknowledgement

This work has received a French government support granted to the Cominlabs excellence laboratory and managed by the National Research Agency in the Investing for the Future program under reference ANR-10-LABX-07-01.

References

[1] T. Berger, "Decentralized estimation and decision theory," in *IEEE Seven Springs Workshop on Inf. Theory, Mt. Kisco, NY*, 1979.

[2] R. Ahlswede and I. Csiszár, "Hypothesis testing with communication constraints," *IEEE Trans. Inf. Theory*, vol. 32, no. 4, pp. 533–542, 1986.

[3] T. Han, "Hypothesis testing with multiterminal data compression," *IEEE Trans. Inf. Theory*, vol. 33, no. 6, pp. 759–772, 1987.

[4] M. Wigger and R. Timo, "Testing against independence with multiple decision centers," in *2016 Int. Conf. Signal Process. and Commun. (SPCOM)*, Bangaro, Inadia, 2016, pp. 1–5.

[5] M. S. Rahman and A. B. Wagner, "On the optimality of binning for distributed hypothesis testing," *IEEE Trans. Inf. Theory*, vol. 58, no. 10, pp. 6282–6303, 2012.

[6] X. Xu and S.-L. Huang, "On distributed hypothesis testing with constant-bit communication constraints," in *2021 IEEE Inf. Theory Workshop (ITW)*, Kanazawa, Japan, 2021, pp. 1–6.

[7] S. Salehkalaibar and M. Wigger, "Distributed hypothesis testing with variable-length coding," *IEEE J. Select. Areas in Inf. Theory*, vol. 1, no. 3, pp. 681–694, 2020.

[8] S. Espinosa, J. F. Silva, and P. Piantanida, "New results on testing against independence with rate-limited constraints," in *2019 IEEE Global Conf. Signal and Inf. Process. (GlobalSIP)*, Ottawa, ON, Canada, 2019, pp. 1–5.

[9] S. Salehkalaibar and V. Y. Tan, "Distributed sequential hypothesis testing with zero-rate compression," in *2021 IEEE Inf. Theory Workshop (ITW)*, Kanazawa, Japan, 2021, pp. 1–5.

[10] S. Salehkalaibar and M. Wigger, "Distributed hypothesis testing based on unequal-error protection codes," *IEEE Trans. Inf. Theory*, vol. 66, no. 7, pp. 4150–4182, 2020.

[11] E. Haim and Y. Kochman, "Binary distributed hypothesis testing via Körner-Marton coding," in *2016 IEEE Inf. Theory Workshop (ITW)*, Cambridge, UK, pp. 146–150.

[12] D. Bertsekas and J. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Prentice Hall, Englewood Cliff, NJ, 1989.

[13] M. Longo, T. D. Lookabaugh, and R. M. Gray, "Quantization for decentralized hypothesis testing under communication constraints," *IEEE Tran. Inf. Theory*, vol. 36, no. 2, pp. 241–255, 1990.

[14] M. Chen, W. Liu, B. Chen, and J. Matyjas, "Quantization for distributed testing of independence," in *2010 13th Int. Conf. Inf. Fusion*, Edinburgh, UK, 2010, pp. 1–5.

[15] Y. Inan, M. Kayaalp, A. H. Sayed, and E. Telatar, "A fundamental limit of distributed hypothesis testing under memoryless quantization," in *Proc. IEEE Int. Con. Commun. (ICC)*, Seoul, Korea, Republic of, May 2022, pp. 4824–4829.

[16] E. Dupraz, I. Salihou Adamou, R. Asvadi, and T. Matsumoto, "Practical short-length coding schemes for binary distributed hypothesis testing," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, 2024.

[17] T. Richardson and R. Urbanke, *Modern Coding Theory*. Cambridge Univ. Press, 2008.

[18] S. Lin and D. Costello, *Error Control Coding: Fundamentals and Applications*, ser. Pearson education. Pearson-Prentice Hall, 2004.

[19] G. Cohen, I. Honkala, S. Litsyn, and A. Lobstein, *Covering Codes*. Elsevier, 1997.

[20] R. Gray, "Vector quantization," *IEEE ASSP Mag.*, vol. 1, no. 2, pp. 4–29, April 1984.

[21] E. L. Lehmann, J. P. Romano, and G. Casella, *Testing Statistical Hypotheses*. New York: Springer-Verlag, 2005, vol. 3.

[22] L. Comtet, *Advanced Combinatorics: The art of finite and infinite expansions*. Springer Science & Business Media, 2012.

[23] A. Schrijver, *Theory of Linear and Integer Programming*. John Wiley & Sons, 1998.

[24] R. B. Ash, *Information Theory*. Courier Corporation, 2012.

[25] A. Prügel-Bennett, *The probability companion for engineering and computer science*. Cambridge University Press, 2020.

[26] M. Thomas and A. T. Joy, *Elements of information theory*. Wiley-Interscience, 2006.

[27] R. Graham and N. Sloane, "On the covering radius of codes," *IEEE Trans. Inf. Theory*, vol. 31, no. 3, pp. 385–401, 1985.