Bit-error Resilient Index Assignment for Multiple Description Scalar Quantizers

Sorina Dumitrescu, *Senior Member, IEEE*, and Yinghan Wan

***Abstract*—This work addresses the problem of increasing the robustness to bit errors for two description scalar quantizers. Our approach is to start with an m-diagonal index assignment and further apply a permutation to the indexes of each description to increase the minimum Hamming distance d_{min} of the set of valid index pairs. In particular, we show how to construct linear permutation pairs achieving $d_{min} \geq 3$, and establish a lower bound in terms of the description rate R, for the highest value of m for which such permutations exist.**

For the case when one description is known to be correct we propose a new performance measure, denoted by $d_{side,min}$. This represents the minimum Hamming distance of the set of indexes of one description, when the index of the other description is fixed. We prove the close connection between the problem of robust permutations design under the new criterion and the antibandwidth problem in a certain graph derived from a hypercube. Leveraging this connection we settle the problem of existence of permutations achieving $d_{side,min} \geq 2$, respectively $d_{min} \geq 2$, and show their construction. Further, we develop a technique for constructing linear permutation pairs achieving $d_{side,min} \geq h$ based on linear $(R, \lceil \log_2 m \rceil)$ channel codes of minimum Hamming distance $h+1$. Additionally, tight bounds in terms of R, on the maximum achievable value of $d_{side,min}$ are derived for $m = 2, 3, 4$.

***Index Terms*—Multiple descriptions, bit-error resilient index assignment, minimum Hamming distance, antibandwidth.**

I. Introduction

A (symmetric) multiple description (MD) code produces a number of descriptions of a signal such that the quality of the reconstruction increases with the number of decoded descriptions. Practical MD coding has been extensively studied over the past two decades due to its applications in modern communications systems [1] – [25].

While MD schemes traditionally target robustness against description loss, a natural question is whether the redundancy that is intentionally built into the system can also be used to combat other channel impairments, such as bit errors. Indeed, the ability of combating bit errors (with or without additional channel coding) has been attested and exploited via joint source-channel decoding [26] – [32]. On the other hand, the design of MD codes to strengthen this ability has received only little attention [33], [34]. The authors of [33] consider the case of a two description scalar quantizer (2DSQ) and identify as a measure for index assignment (IA) robustness to bit errors, the minimum Hamming distance of the set of valid pairs of two description indexes. Their procedure for robust IA design is split into two steps: (1) given a value d, find a set of index pairs of minimum Hamming distance d; (2) assign index pairs from the set determined at the first stage to central partition cells. The second problem is solved heuristically via a genetic algorithm, while for the first problem no systematic solution for the general case is presented other than exhaustive search, which is practical only for small values of R. An efficient solution for the first problem is proposed only for the case of 1-bit redundancy, achieving $d = 2$. In [34] the authors consider a multiple description vector quantizer with a general number of descriptions over space-time orthogonal block coded slow Rayleigh fading channels. The problem of optimal IA design is formulated by modeling the concatenation of the IA, modulators, space-time encoders, multiple-antenna channel, space-time soft decoder, linear combiner and the maximum a posteriori probability detector as an equivalent discrete memoryless channel, and it is solved by using the binary switching heuristic algorithm.

This work addresses the problem of increasing the bit-error resilience of the IA in the case of 2DSQ, without decreasing its resilience to description loss. Our approach is to start from an initial IA, which is known to be good for the conventional 2DSQ problem, such as the IA's proposed in [1], [14], [20], [23], and apply a permutation to the indexes of each description. Such a technique does not change the performance of the 2DSQ in the conventional sense, i.e., when the descriptions are not corrupted by bit errors, but it has the potential of increasing the bit-error resilience at the central decoder.

For the scenario when both descriptions may carry bit errors we use the minimum Hamming distance of the set of valid index pairs, denoted by d_{min}, as a performance measure, following [33]. Specifically, we will show how to construct linear permutation pairs achieving $d_{min} \geq 3$ and establish a lower bound on the maximum m, for given R, for which such permutations exist. Note that R represents the rate of each description, while m denotes the number of diagonals occupied by the valid index pairs in the initial IA matrix. Clearly, by using such permutations the central decoder is able to correct any one bit error pattern in the pair of received indexes.

Another interesting scenario is when one description is known to be correct. This could happen in the transmission over wireless channels of fluctuating quality. Assume that

The first author is with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada. The second author was with the same department at the time the paper was written. Emails: sorina@mail.ece.mcmaster.ca, katrinawyh@gmail.com.

This paper was presented in part at the Canadian Workshop on Information Theory, Toronto, June 2013, and at the International Symposium on Information Theory, Istanbul, July 2013.

the sequence of indexes corresponding to each description is applied a systematic channel code whose decoder either corrects all the errors or declares a decoding failure. Consider now the event when the quality of the first channel is such that all the errors can be corrected, while the second channel is so bad that the error correcting code is ineffective. In such a case for each pair of received indexes, the decoder knows for sure that the first index is correct, while for the second index there is the possibility of bit errors.

For the aforementioned scenario when one description is known to be correct we propose a better suited performance criterion, termed the *side minimum Hamming distance* of the IA, and denoted by $d_{side,min}$. This notion is defined as the minimum Hamming distance of the set of valid indexes of the description which may carry errors, when the index in the correct description is fixed. Interestingly, we show that the problem of designing permutations achieving $d_{side,min} \geq d$ is closely connected to the antibandwidth problem in a modified R-dimensional hypercube, denoted by $Q_R(d)$, where edges connect any two vertices at Hamming distance at most $d-1$. Specifically, such a permutation exists if and only if the antibandwidth of $Q_R(d)$ is larger or equal to m. The antibandwidth problem is very difficult in general, but its solution is known for the hypercube, i.e. for $Q_R(2)$, [35], [36]. Leveraging the results of [35], [36] we establish the highest value of m, in terms of R (and thus, the lowest redundancy of the IA), for which permutations achieving $d_{side,min} \geq 2$ exist, and show their construction. By proving, additionally, that a permutation pair achieves $d_{min} \geq 2$ if and only if $d_{side,min} \geq 2$, we also settle the problem of existence of permutation pairs achieving $d_{min} \geq 2$.

For the case when $d > 2$ the antibandwidth problem in $Q_R(d)$ has not been studied. Therefore, in an attempt to simplify the problem, we turn our attention to the design of linear permutations that increase $d_{side,min}$. We show how to construct linear permutations achieving $d_{side,min} \geq h$ based on linear $(R, \lceil \log_2 m \rceil)$ channel codes with minimum Hamming distance $h+1$. This result is one of the main contributions of this work. We further exploit this technique to construct permutations achieving $d_{side,min} \geq 3$ based on shortened Hamming codes, and establish a lower bound in terms of R, on the largest value of m for which such permutations exist. Furthermore, we determine tight bounds in terms of R, on the highest value of $d_{side,min}$ achievable via general permutations (not necessarily linear), for $m = 2, 3, 4$.

Parts of this work were first presented in the conference papers [37], [38]. In comparison with [37], [38], the present work includes experimental results and provides more detailed proofs and literature review, as well as new theoretical results, namely Propositions 4, 5 and 7. Note that Propositions 4 and 5 establish tight bounds in terms of R, on the largest value of $d_{side,min}$ achievable via general permutations, for $m = 2, 3, 4$. Further, Proposition 7 proposes the construction of linear permutations achieving $d_{min} \geq 3$ for $m = 2$ and any $R \geq 6$, thus improving upon the result of Proposition 6 specialized to $m = 2$.

The paper is structured as follows. Section II presents most of the notations used in this work, the definition of a 2DSQ and its operation for on/off channels. In Section III d_{min} and $d_{side,min}$ are considered as criteria for robustness of the IA and the problem of designing permutations that increase the robustness of m-diagonal IA's is formulated. Some preliminary results are proven. The following section establishes the connection with the antibandwitdh problem and derives results establishing the existence and construction of permutations achieving d_{min} and $d_{side,min}$ larger than or equal to 2. Section V presents a technique for the construction of linear permutations achieving $d_{side,min} \geq h$ based on linear $(R, \lceil \log_2 m \rceil)$ channel codes of minimum Hamming distance $h+1$. This technique is further exploited to construct permutations achieving $d_{side,min} \geq 3$ and to establish lower bounds for the maximum m for which such permutations exist. Additionally, for the case when $m = 2, 3, 4$, tight bounds in terms of R, for the maximum $d_{side,min}$ achievable using general permutations, are derived. Section VI addresses the scenario when both descriptions may carry bit errors and proposes the construction of linear permutations with $d_{min} \geq 3$. Experimental results validating the robustness of the proposed IA's are presented in section VII. Finally, section VIII concludes the paper.

II. Definitions and Notations

Let $\mathbb{F}_2$ denote the binary field with elements 0 and 1. We use lower case letters in bold to denote bit sequences. If the length of the bit sequence $\mathbf{u}$ is k then we write $\mathbf{u} = (u_1, u_2, \cdots, u_k)$. Note that $\mathbf{u}$ is a row vector in the vector space $\mathbb{F}_2^k$. Additionally, denote $\mathbf{u}_s^t = (u_s, u_{s+1}, \cdots u_t)$, for any $1 \leq s \leq t \leq k$. For all integers i and $k > 0$ such that $0 \leq i \leq 2^k - 1$, let $\boldsymbol{\beta}_k(i)$ denote the k-bit representation of i starting with the most significant bit and ending with the least significant bit. When $k = R$ we will omit the subscript, i.e., we will use $\boldsymbol{\beta}(i)$ instead of $\boldsymbol{\beta}_R(i)$. Conversely, for any bit sequence $\mathbf{b} \in \mathbb{F}_2^k$, $k > 0$, $\eta(\mathbf{b})$ denotes the corresponding integer in natural binary representation, i. e., $\eta(\mathbf{b}) \triangleq \sum_{s=1}^k b_s 2^{k-s}$. For any positive integer k, $\mathbf{0}_k$, respectively $\mathbf{1}_k$, denotes the all zero, respectively all one, vector in $\mathbb{F}_2^k$. The subscript k will be omitted when it is understood from the context.

We will use upper case letters in bold to denote matrices with elements in $\mathbb{F}_2$. The set of all n-by-k matrices with elements in $\mathbb{F}_2$ is denoted by $\mathcal{M}_{n\times k}$. $\mathbf{I}_k$ denotes the k-by-k identity matrix, $\mathbf{0}_{n\times k}$ denotes the n-by-k zero matrix, and $\mathbf{D}_k$ denotes the matrix in $\mathcal{M}_{k\times k}$ with elements $D(i,j) = 1$ if and only if $i = j$ or $i = j - 1$, while the remaining elements are 0. For any matrix $\mathbf{A}$, $\mathbf{A}^T$ denotes its transpose, and $det(\mathbf{A})$ denotes its determinant.

The addition of integers is denoted by "+", while the addition in the binary field $\mathbb{F}_2$ is denoted by "$\oplus$". The component wise addition of vectors in $\mathbb{F}_2^k$ and of matrices in $\mathcal{M}_{n\times k}$ are also denoted by "$\oplus$". For any two sets $\mathcal{A}, \mathcal{B} \subseteq \mathbb{F}_2^k$, we define their sum $\mathcal{A} \oplus \mathcal{B} \triangleq \{u \oplus v : u \in \mathcal{A}, v \in \mathcal{B}\}$ (**note**: this should not be confused with the notion of direct sum).

For any binary vector $\mathbf{u} \in \mathbb{F}_2^k$, its Hamming weight $H(\mathbf{u})$ is defined as the number of components equal to 1. The Hamming distance between two binary vectors $\mathbf{u}$ and $\mathbf{v}$ of the same dimension is $d(\mathbf{u}, \mathbf{v}) \triangleq H(\mathbf{u} \oplus \mathbf{v})$. For any set

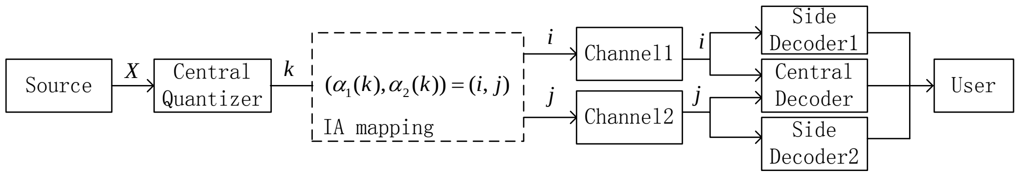

Fig. 1. Block diagram of a 2DSQ system for on/off channels.

$\mathcal{A} \subseteq \mathbb{F}_2^k$, we define $H_{min}(\mathcal{A}) \triangleq \min_{\mathbf{u} \in \mathcal{A}} H(\mathbf{u})$. Additionally, if $|\mathcal{A}| \geq 2$, we denote by $d_{min}(\mathcal{A})$ the minimum Hamming distance between any two different elements of $\mathcal{A}$.

A (symmetric) two description scalar quantizer (2DSQ) (Fig. 1) operates as follows. The source sample x is encoded first by a so-called central quantizer q to an index $k \in \{0, \cdots, N-1\}$. Every index k is further mapped to a pair of indexes (i, j) via the IA mapping $\alpha : \{0, \cdots, N-1\} \to \mathcal{I}_R \times \mathcal{I}_R$, where R is a positive integer, and $\mathcal{I}_R \triangleq \{0, 1, 2, \cdots, 2^R - 1\}$. We will denote by $Im(\alpha)$ the set of assigned index pairs, i.e., $Im(\alpha) = \{\alpha(k) : 0 \leq k \leq N-1\}$ and by α_s the s-th component of α, $s = 1, 2$, i.e., $\alpha(k) = (\alpha_1(k), \alpha_2(k))$ for any k. The index $\alpha_s(k)$ represents the s-th description. The R-bit sequence $\boldsymbol{\beta}(\alpha_s(k))$ is sent over the s-th channel, for $s = 1, 2$.

In a conventional 2DSQ it is assumed that each channel either transmits correctly or breaks down. Therefore, at the receiver end there are three non-trivial decoders: the central decoder $g_0 : Im(\alpha) \to \hat{\mathcal{X}}_0$, for the case when both i and j are received, where (i, j) is the transmitted pair, and the side decoders g_1, respectively g_2, when only i, respectively j, is received. Note that $g_s : \mathcal{I}_R \to \hat{\mathcal{X}}_s$, $s = 1, 2$, and $\hat{\mathcal{X}}_t \subset \mathbb{R}$, $t = 0, 1, 2$.

The performance at each decoder is measured by the expected distortion between the source and the reconstruction. The distortion measure that we assume in this work is the squared distance.

In a 2DSQ system there is a trade-off between the quality of reconstruction at the central decoder and the reconstruction at the side decoders. Therefore, the distortions at all three decoders cannot be minimized simultaneously. To account for this, the performance measure for the 2DSQ can be formulated as a weighted sum of the distortion at the central decoder D_0 and of the average distortion at the side decoders $(D_1 + D_2)/2$. Vaishampayan proved in [1] that the optimal 2DSQ must contain convex cells (i.e., intervals) in the central partition. On the other hand, the cells of the side encoder partitions are not necessarily convex. Additionally, note that the number N of cells in the central partition, is the main parameter that controls the trade-off between the quality at the central and side decoders.

Further, notice that when the central quantizer partition is fixed (and hence N is fixed), the IA α determines the performance at the side decoders. Therefore, the IA is a crucial component of the 2DSQ system. Vaishampayan introduced in [1] the notion of *spread* as a measure of the quality of the IA. For $t = 1, 2$ and $\ell \in \mathcal{I}_R$, the spread of index ℓ in description t is the number of central quantizer cells between the infimum and the supremum of the cell with index ℓ of side encoder t. Further, the spread of the IA is defined as the maximum value of the spreads of all cells in both descriptions. Using a high resolution argument Vaishampayan [1] showed that the problem of optimizing the IA is connected to the problem of minimizing the spread.

Vaishampayan proposed in [1] good IA's and provided a lower bound on the minimum spread for a restricted class of IA's. Note that the IA can be illustrated using a $2^R \times 2^R$ matrix (or table) where only positions corresponding to valid index pairs (i, j) are filled and the remaining positions are empty. Specifically, the position (i, j) contains k if $\alpha(k) = (i, j)$, and is empty if (i, j) is not in $Im(\alpha)$. The author of [1] advocates the selection of valid index pairs that fill only elements on the main diagonal and on the diagonals closest to the main diagonal. He further proposes several classes of IA's for the case when an odd number of diagonals are filled.

The problem of minimizing the spread of an IA was also addressed in [39] and [40]. In [39] the problem is solved for the no redundancy case, i.e., when $N = 2^{2R}$. In [40] tighter bounds on the optimum spread are proved.

Another case where the problem of optimal IA is settled is when the side encoders cells are enforced to be convex sets (i.e., intervals). Such a 2DSQ is termed *convex* 2DSQ. In this case N cannot be larger than $2^{R+1} - 1$. The optimal convex 2DSQ design was studied in [11], [16]–[18], [20]. In [11] it was shown that the staggered IA is optimal for the convex 2DSQ and the result was generalized to more than two descriptions in [20]. Note that the staggered IA fills the main diagonal and the diagonal above it and satisfies $\alpha(2\ell) = (\ell, \ell)$ and $\alpha(2\ell + 1) = (\ell, \ell + 1)$.

Another direction in the study of IA design is in the context of multiple description lattice vector quantizers (MDLVQ). In this case the central quantizer has a lattice codebook, termed the central lattice. The IA assigns to each central lattice point a pair of points from a sublattice. These sublattice points represent the two descriptions. The problem of optimal IA design for MDLVQ with two descriptions was addressed in [4] and for general number of descriptions in [13], [14], [19]. A variation of an MD lattice scalar quantizer is considered in [23], namely by using translated lattices as the side codebooks. Inspired by [14] and taking advantage of the specifics of their setup, the authors of [23] propose a simple technique for the optimal IA design for certain values of N. For the case of two descriptions their IA's fill an even number of diagonals in the IA matrix, thus complementing the IA's proposed in [1].

III. Problem Formulation and Preliminary Results

In this work we assume that the channels over which the two description are transmitted may carry bit errors. A minimum

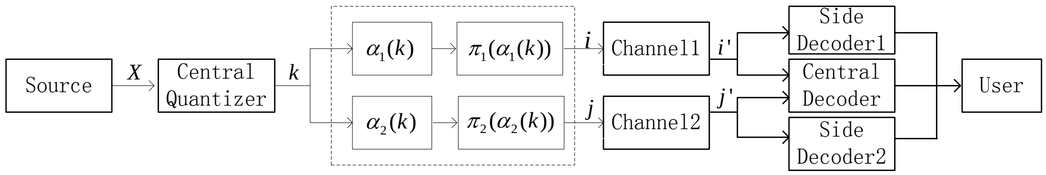

Fig. 2. Block diagram of a 2DSQ system with permutations applied to indexes of each description.

Hamming distance decoder is used as the central decoder. The details of the decoding process are provided in Appendix A. Additionally, it is shown in the same appendix that the minimum Hamming distance decoder is a good approximation for the optimum decoder when the bit errors are independent and the bit error rate is sufficiently small.

We are concerned with the design of index assignments that are robust to bit errors. Therefore, let us first clarify our measure of robustness. Let us denote

$$d_{min}(\alpha) \triangleq d_{min}(\{(\boldsymbol{\beta}(i), \boldsymbol{\beta}(j)) : (i,j) \in Im(\alpha)\}).$$

Note that in the case when both channels may carry bit errors, the central decoder can correct all error patterns with at most $\left\lceil \frac{d_{\min}(\alpha)}{2} \right\rceil - 1$ bit errors. Therefore, we will use $d_{min}(\alpha)$ as a measure of robustness to bit errors of the IA α.

On the other hand, for the situation when one description is known to be correct we propose a better performance measure. Assume that description 1 (i.e., index i) is known to be correct. Then the decoder only needs to look into the set of second description indexes j such that $(i,j) \in Im(\alpha)$. Then a relevant performance measure is what we refer to as the *side 2 minimum Hamming distance* denoted by $d_{2,\min}(\alpha)$ and defined as

$$d_{2,\min}(\alpha) \triangleq \min_{i \in \mathcal{I}_R} d_{\min}(\{\boldsymbol{\beta}(j) : (i,j) \in Im(\alpha)\}).$$

Clearly, the minimum Hamming distance decoder is able to correct any error pattern with at most $\left\lceil \frac{d_{2,\min}(\alpha)}{2} \right\rceil - 1$ bit errors. The side 1 minimum Hamming distance $d_{1,\min}(\alpha)$ is defined similarly, i.e.,

$$d_{1,\min}(\alpha) \triangleq \min_{j \in \mathcal{I}_R} d_{\min}(\{\boldsymbol{\beta}(i) : (i,j) \in Im(\alpha)\}).$$

Further, define

$$d_{side,\min}(\alpha) \triangleq \min(d_{1,\min}(\alpha), d_{2,\min}(\alpha)).$$

It is clear that for any IA α, we have $d_{side,\min}(\alpha) \geq d_{\min}(\alpha)$. We will see shortly (in Remark 1) that there are IA's of interest for which the inequality is strict.

In order to increase the robustness when the channels may introduce bit errors, while maintaining the performance of the 2DSQ for the on/off channels, we start with an IA known to be good for the conventional 2DSQ and apply an index permutation to the index output by each description.

Let $\pi_s : \mathcal{I}_R \to \mathcal{I}_R$ be the permutation applied to indexes of description s, $s = 1, 2$. We will use the notation $\boldsymbol{\pi}$ for the permutation pair (π_1, π_2). Thus, a new IA, denoted by $\boldsymbol{\pi} \circ \alpha$, is generated, with $(\boldsymbol{\pi} \circ \alpha)(k) = (\pi_1(\alpha_1(k)), \pi_2(\alpha_2(k)))$, for any $k \in \mathcal{I}_R$. Fig. 2 shows the 2DSQ system obtained after applying the permutation pair $\boldsymbol{\pi}$.

We will consider initial IA's with the assigned pairs filling the main diagonal and the closest $m-1$ diagonals in the IA matrix, as advocated in [1], [11], [20], [23]. We will refer to such IA's as m-diagonal IA's, formally defined as follows.

Definition 1. *Let $R \geq 2$ and $2 \leq m \leq 2^R$. An IA $\alpha : \{0, 1, \cdots, N-1\} \to \mathcal{I}_R \times \mathcal{I}_R$ is called m-diagonal if $Im(\alpha)$ is the set of all pairs $(a, a+\tau)$ satisfying $0 \leq a \leq 2^R - 1$ and $\max(-a, -m+1+\lfloor \frac{m}{2} \rfloor) \leq \tau \leq \min(\lfloor \frac{m}{2} \rfloor, 2^R - 1 - a)$.*

It can be easily verified that for any m-diagonal IA α, one has $d_{min}(\alpha) = d_{1,min}(\alpha) = d_{2,min}(\alpha) = 1$.

As shown in [41], the expected distortion D of a noisy channel quantizer satisfying the centroid condition can be written as the sum of two components, one due to the quantization of the source, and the other due to the channel noise. Thus, assuming that g_0 satisfies the centroid condition (15) (in Appendix A), the expected distortion at the central decoder can be decomposed as

$$D_0 = D_{0,S} + D_{0,C},$$

where $D_{0,S}$ is the distortion due to quantization, given by (14), and $D_{0,C}$ is the channel distortion given by (16). It is easy to see that by applying the permutation pair $\boldsymbol{\pi}$ the source distortion $D_{0,S}$ of (14) is not affected. On the other hand, the channel distortion $D_{0,C}$ can be decreased with a careful of choice of $\boldsymbol{\pi}$, as we will show in this work. To achieve this goal we will construct permutation pairs that increase d_{min} and $d_{side,min}$ beyond 1. We will illustrate this possibility with an example shortly. For this we need to present first a result which shows that in order to ensure that $d_{min}(\boldsymbol{\pi} \circ \alpha) \geq 2$ it is sufficient to have $d_{side,min}(\boldsymbol{\pi} \circ \alpha) \geq 2$, a relation which is easier to verify.

Proposition 1. *Let $\boldsymbol{\pi} = (\pi_1, \pi_2)$ be a pair of permutations of the set $\mathcal{I}_R$ and $\alpha : \{0, 1, \cdots, N-1\} \to \mathcal{I}_R \times \mathcal{I}_R$. Then $d_{min}(\boldsymbol{\pi} \circ \alpha) \geq 2$ if and only if $d_{side,min}(\boldsymbol{\pi} \circ \alpha) \geq 2$.*

Proof: The implication "$\Rightarrow$" is obvious. Assume now that $d_{side,min}(\boldsymbol{\pi} \circ \alpha) \geq 2$ and let $(i_1, j_1) \neq (i_2, j_2)$ be two index pairs in $Im(\alpha)$. We have to prove that

$$d\big(\boldsymbol{\beta}(\pi_1(i_1)), \boldsymbol{\beta}(\pi_2(i_2))\big) + d\big(\boldsymbol{\beta}(\pi_1(j_1)), \boldsymbol{\beta}(\pi_2(j_2))\big) \geq 2. \quad (1)$$

If $i_1 \neq i_2$ and $j_1 \neq j_2$ then $\pi_1(i_1) \neq \pi_1(i_2)$ and $\pi_2(j_1) \neq \pi_2(j_2)$. It follows that $d(\boldsymbol{\beta}(\pi_1(i_1)), \boldsymbol{\beta}(\pi_1(i_2))) \geq 1$ and $d(\boldsymbol{\beta}(\pi_2(j_1)), \boldsymbol{\beta}(\pi_2(j_2))) \geq 1$, which implies (1). If $i_1 = i_2$ then necessarily $j_1 \neq j_2$ and $d_{2,min}(\boldsymbol{\pi} \circ \alpha) \geq 2$ implies that $d(\boldsymbol{\beta}(\pi_2(j_1)), \boldsymbol{\beta}(\pi_2(j_2))) \geq 2$. Finally, if $j_1 = j_2$ then the inequality $d_{1,min}(\boldsymbol{\pi} \circ \alpha) \geq 2$ leads to $d(\boldsymbol{\beta}(\pi_1(i_1)), \boldsymbol{\beta}(\pi_1(i_2))) \geq 2$, further implying (1). ∎

TABLE I
EXAMPLE OF PERMUTATION π THAT INCREASES THE ROBUSTNESS OF A 2-DIAGONAL IA FOR $R = 4$.

i	$\pi(i)$	$\boldsymbol{\beta}(\pi(i))$
0	0	0000
1	15	1111
2	1	0001
3	14	1110
4	3	0011
5	12	1100
6	2	0010
7	13	1101
8	6	0110
9	9	1001
10	7	0111
11	8	1000
12	5	0101
13	10	1010
14	4	0100
15	11	1011

Further, to facilitate the computation of $d_{side,min}(\boldsymbol{\pi} \circ \alpha)$ consider the following definition.

Definition 2. *For any integers $R \geq 2$, $2 \leq m \leq 2^R$, and any one-to-one mapping $\pi : \mathcal{I}_R \to \mathcal{I}_R$, define*

$$\mu(\pi, m) \triangleq \min_{\substack{j_1, j_2 \in \mathcal{I}_R \\ 1 \leq |j_1 - j_2| \leq m-1}} d(\boldsymbol{\beta}(\pi(j_1)), \boldsymbol{\beta}(\pi(j_2))).$$

It can be easily seen that for any permutation pair $\boldsymbol{\pi} = (\pi_1, \pi_2)$ and any m-diagonal IA α, one has $d_{s,\min}(\boldsymbol{\pi} \circ \alpha) = \mu(\pi_s, m)$, for $s = 1, 2$. Therefore,

$$d_{side,\min}(\boldsymbol{\pi} \circ \alpha) = \min\{\mu(\pi_1, m), \mu(\pi_2, m)\}.$$

It follows that in order to find a permutation pair $\boldsymbol{\pi}$ with $d_{side,\min}(\boldsymbol{\pi} \circ \alpha) = d$ for some d, it is sufficient to find a permutation π with $\mu(\pi, m) = d$ and then let $\boldsymbol{\pi} = (\pi, \pi)$.

Example 1. *Let $R = 4$, $m = 2$ and α be a 2-diagonal IA with $Im(\alpha) = \{(i, i) : i \in \mathcal{I}_4\} \cup \{(i, i+1) : i \in \mathcal{I}_4 \setminus \{15\}\}$. Consider now the permutation π of the set $\mathcal{I}_4$ defined in Table I. Let $\pi_1 = \pi_2 = \pi$. By inspecting Table I we notice that $d(\boldsymbol{\beta}(\pi(i)), \boldsymbol{\beta}(\pi(i+1)))$ is either 3 or 4. Therefore, $\mu(\pi, 2) = 3$, leading to $d_{side,min}(\boldsymbol{\pi} \circ \alpha) = 3$. Proposition 1 further implies that $d_{min}(\boldsymbol{\pi} \circ \alpha) \geq 2$. Since $(0, 0), (2, 2) \in Im(\alpha)$ and $d(\boldsymbol{\beta}(\pi(0)), \boldsymbol{\beta}(\pi(2))) + d(\boldsymbol{\beta}(\pi(0)), \boldsymbol{\beta}(\pi(2))) = 2$, it further follows that $d_{min}(\boldsymbol{\pi} \circ \alpha) = 2$.*

Remark 1. *Notice that Example 1 also illustrates the fact that there are permutations where $d_{side,min}(\boldsymbol{\pi} \circ \alpha) > d_{min}(\boldsymbol{\pi} \circ \alpha)$, thus supporting the claim that $d_{side,min}$ is a better performance measure for bit-error resilience than d_{min} for the scenario when one channel is known to be error-free.*

Before ending this section we would like to point out that the technique of applying an appropriate permutation to the indexes output by a quantizer in order to increase the error resilience over a noisy channel, has been extensively investigated. Notice that such a permutation determines the assignment of quantizer outputs to binary indexes transmitted over the channel. Numerous studies were dedicated to the design and analysis of robust index assignments, among which are [42] – [53]. A major distinction between the aforementioned problem and ours is that the minimum Hamming distance criterion is not appropriate for the former problem since no permutation can increase the Hamming distance of the set of codewords.

IV. CONNECTION BETWEEN SIDE MINIMUM HAMMING DISTANCE AND THE ANTIBANDWIDTH PROBLEM

In this section, we show that the problem of designing permutations achieving high side minimum Hamming distance for the m-diagonal initial IA, is closely related to the antibandwidth problem in a certain graph derived from a hypercube. To this end we first introduce the antibandwidth problem. Then we describe the relation between the antibandwidth and robust permutations. Finally, we draw a conclusion about the existence of permutations achieving $d_{side,\min} \geq 2$, and present their construction based on known results on the antibandwidth of a hypercube.

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be an undirected graph, where $\mathcal{V}$ denotes the set of vertices and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$, denotes the set of edges. Let the size of $\mathcal{V}$ be n, i.e., $n = |\mathcal{V}|$. A labeling or numbering of the vertices of $\mathcal{G}$ is a one-to-one mapping ω: $\mathcal{V} \to \{0, 1, \cdots, n-1\}$. The antibandwidth of the graph $\mathcal{G}$, denoted by $ab(\mathcal{G})$, is defined as

$$ab(\mathcal{G}) = \max_{\omega} \min_{(u,v) \in \mathcal{E}} |\omega(v) - \omega(u)|. \quad (2)$$

The antibandwidth problem was studied in [54] – [60]. Finding the antibandwidth of a graph is in general a very difficult problem. It was proved that for general graphs the problem of determining if the antibandwidth is larger than some given value is NP-complete [54]. The exact value of the antibandwidth and the achieving labeling are known only for certain special cases such as the hypercube [35], [36].

To establish the connection between our problem and the antibandwidth problem we need first the following definition.

Definition 3. *For any positive integer ℓ let Q_ℓ denote the ℓ-dimensional hypercube, i.e., the undirected graph with vertex set $\{0, 1\}^\ell$, where the edges connect any two vertices at Hamming distance 1. Further, for any integer d, $2 \leq d \leq \ell$, define $Q_\ell(d)$ as the undirected graph with the same vertex set as Q_ℓ and the edge set consisting of all pairs $(\mathbf{u}, \mathbf{v})$ such that $\mathbf{u} \neq \mathbf{v}$ and $d(\mathbf{u}, \mathbf{v}) \leq d - 1$. (Notice that $Q_\ell(2) = Q_\ell$.)*

Theorem 1. *For integers $R \geq 2$, and m, $2 \leq m \leq 2^R$, the following relation holds*

$$\max\{\mu(\pi, m) : \pi \in \mathcal{I}_R\} = \max\{d : ab(Q_R(d)) \geq m\}. \quad (3)$$

Proof: Relation $ab(Q_R(d)) \geq m$ is equivalent to the fact that there exists a numbering $\omega : \{0, 1\}^R \to \mathcal{I}_R$ such that

$$|\omega(\mathbf{v}) - \omega(\mathbf{u})| \geq m \text{ holds for all } \mathbf{u}, \mathbf{v} \in \{0, 1\}^R \text{ with } \mathbf{u} \neq \mathbf{v} \text{ and } d(\mathbf{u}, \mathbf{v}) \leq d - 1. \quad (4)$$

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TIT.2015.2413780

Statement (4) is further equivalent to the following

$$d(\mathbf{u},\mathbf{v}) > d-1 \text{ holds for all } \mathbf{u},\mathbf{v} \in \{0,1\}^R \\ \text{with } \mathbf{u} \neq \mathbf{v} \text{ and } |\omega(\mathbf{v})-\omega(\mathbf{u})| < m. \quad (5)$$

Let $\pi = \eta \circ \omega^{-1}$. Thus, π is a permutation of $\mathcal{I}_R$ and $j = \omega(\boldsymbol{\beta}(\pi(j)))$ for any $j \in \mathcal{I}_R$. Then statement (5) is equivalent to

$$d(\boldsymbol{\beta}(\pi(j_1)), \boldsymbol{\beta}(\pi(j_2))) \geq d \text{ holds for all } j_1, j_2 \in \mathcal{I}_R \\ \text{with } j_1 \neq j_2 \text{ and } |j_1 - j_2| \leq m-1,$$

which is further equivalent to $\mu(\pi, m) \geq d$. Since the relation $\pi = \eta \circ \omega^{-1}$ determines a one-to-one correspondence between labellings ω of $Q_R(d)$ and permutations π of the set $\mathcal{I}_R$, equality (3) easily follows. ■

Theorem 1 implies that if we knew the value of $ab(Q_R(d))$ for any d, then we could find the largest value of $\mu(\pi, m)$ for given m and R, via (3). Unfortunately, the value of $ab(Q_R(d))$ has not been studied for $d > 2$. On the other hand, the antibandwidth of the hypercube (i.e., when $d = 2$) is known. In particular, the antibandwidth achieving permutation for Q_R was found in [35], while the value of the antibandwidth was determined in [36]. In view of Theorem 1 the aforementioned results can be exploited to determine when permutations achieving $\mu(\pi, m) \geq 2$ do exist and to construct them. In order to describe the permutation achieving the hypercube antibandwidth we need the following definition.

Definition 4. *For $R \geq 1$, the Hales order $\leq_H$ on the vertices of Q_R is defined by $\mathbf{u} \leq_H \mathbf{v}$, if (1) $H(\mathbf{u}) < H(\mathbf{v})$, or (2) $H(\mathbf{u}) = H(\mathbf{v})$ and $\mathbf{u}$ is greater than or equal to $\mathbf{v}$ in lexicographic order relative to the right to left order of the coordinates.*

Example 2. *The vertices of Q_4 listed in Hales order are:* 0000, 0001, 0010, 0100, 1000, 0011, 0101, 1001, 0110, 1010, 1100, 0111, 1011, 1101, 1110, 1111.

The following lemma is based on the results of [35], [36].

Lemma 1. *For $R \geq 2$ the following holds*

$$ab(Q_R) = 2^{R-1} - \sum_{k=0}^{R-2} \binom{k}{\lfloor \frac{k}{2} \rfloor}. \quad (6)$$

Furthermore, the permutation achieving the antibandwidth of Q_R is obtained by numbering the vertices of Q_R with even Hamming weight first then those with odd Hamming weight, in Hales order.

Theorem 1, Lemma 1 and Proposition 1 lead to the following result.

Proposition 2. *Let $R \geq 2$ and $2 \leq m \leq 2^R$. Then there are permutation pairs $\boldsymbol{\pi}$ of the set $\mathcal{I}_R$ such that $d_{min}(\boldsymbol{\pi} \circ \alpha) \geq 2$, respectively $d_{side,min}(\boldsymbol{\pi} \circ \alpha) \geq 2$, for any m-diagonal initial IA α, if and only if*

$$m \leq 2^{R-1} - \sum_{k=0}^{R-2} \binom{k}{\lfloor \frac{k}{2} \rfloor}.$$

Additionally, a permutation pair achieving the above equalities is $\boldsymbol{\pi} = (\pi, \pi)$, where $\pi = \eta \circ \omega_R^{-1}$ and ω_R is the labeling of Q_R described in Lemma 1.

TABLE II
THE LARGEST VALUE OF m FOR WHICH PERMUTATIONS π WITH $\mu(\pi, m) \geq 2$ EXIST, FOR $2 \leq R \leq 10$.

R	2	3	4	5	6	7	8	9	10
$m(R)$	1	2	4	9	19	41	82	175	361

Table II lists the largest value of m (denoted by $m(R)$) for which permutations π with $\mu(\pi, m) \geq 2$ exist, for $2 \leq R \leq 10$.

Example 3. *The permutation π indicated in Proposition 2 for $R = 4$, is specified by the following list of values, which represents the sequence of $\pi(x)$ for $x \in \{0, \cdots, 15\}$, in increasing order of x: $0, 3, 5, 9, 6, 10, 12, 15, 1, 2, 4, 8, 7, 11, 13, 14$. According to Proposition 2, for any index assignment α with at most 4 diagonals, one has $d_{side,min}((\pi, \pi) \circ \alpha) \geq 2$ and $d_{min}((\pi, \pi) \circ \alpha) \geq 2$.*

V. LINEAR PERMUTATIONS THAT INCREASE THE SIDE MINIMUM HAMMING DISTANCE

As shown in the previous section, given an m-diagonal initial IA, the problem of determining whether there is a permutation achieving $d_{side,min} \geq d$ is equivalent to determining whether the antibandwidth of $Q_R(d)$ is larger than or equal to m. Unfortunately, there are no results available for the antibandwidth of $Q_R(d)$ with $d \geq 3$. Additionally, the problem of determining if the antibandwidth of a general graph is larger than some given value is NP-complete [54]. These facts motivate our attempt to simplify the problem by confining our attention to the set of linear permutations.

In this section we first show that the value of $\mu(\pi, m)$ for a linear permutation π can be characterized in terms of the minimum Hamming distance of a certain linear channel code and the minimum Hamming weight of a restricted set of binary sequences. Based on this result we further present a simple construction of linear permutations achieving $\mu(\pi, m) \geq h$ based on linear $(R, \lceil \log_2 m \rceil)$ channel codes of minimum Hamming distance $h+1$. Using further this approach in conjunction with shortened Hamming codes, we determine a lower bound in terms of R, on the largest m for which linear permutations with $\mu(\pi, m) \geq 3$ exist. Finally, tight bounds on the maximum value of $\mu(\pi, m)$, in terms of R, are derived for $m = 2, 3, 4$.

Definition 5. *A one-to-one mapping $\pi : \mathcal{I}_R \to \mathcal{I}_R$ is called a linear permutation if there is a full rank matrix $\mathbf{G}_\pi \in \mathcal{M}_{R\times R}$ such that $\boldsymbol{\beta}(\pi(j)) = \boldsymbol{\beta}(j)\mathbf{G}_\pi$, for any $j \in \mathcal{I}_R$.*

Definition 6. *For any integers $R \geq 1$, k, $1 \leq k \leq R$, and any matrix $\mathbf{A} \in \mathcal{M}_{k\times R}$ with $rank(\mathbf{A}) = k$, define*

$$\mathcal{C}(\mathbf{A}) \triangleq \{\mathbf{bA} : \mathbf{b} \in \mathbb{F}_2^k\}$$
$$\mathcal{S}(\mathbf{A}) \triangleq \{\mathbf{bA} : \mathbf{b} = (\mathbf{0}_{k-t}, \mathbf{1}_t), 1 \leq t \leq k\}.$$

Additionally, for any integer m, $2 \le m \le 2^k$, define

$$\mathcal{V}(m, \mathbf{A}) \triangleq \{\mathbf{bA} : \mathbf{b} \in \mathbb{F}_2^k, \eta(\mathbf{b}) \ge 2^k - m + 1\}.$$

Notice that $\mathcal{C}(\mathbf{A})$ is the linear code generated by matrix $\mathbf{A}$. We will use the shorter notation $d_{min}(\mathbf{A})$ for $d_{min}(\mathcal{C}(\mathbf{A}))$. The following result provides a simpler way for determining the value of $\mu(\pi, m)$ for a linear permutation π. Proposition 2 implies that for any $m > 2^{R-1}$ one has $\mu(\pi, m) = 1$ for all π. Therefore, from now on we will only be concerned with $m \le 2^{R-1}$.

Theorem 2. *Consider integers $R \ge 2$ and $m, 2 \le m \le 2^{R-1}$. Let $\mathbf{G} \in \mathcal{M}_{R\times R}$ be a full rank matrix. Let $k = \lceil \log_2 m \rceil$ and denote by $\mathbf{A}$, respectively $\mathbf{B}$, the submatrix of $\mathbf{G}$ formed out of the first $R - k$ rows, respectively the last k rows. Then the linear permutation π defined by $\mathbf{G}$ has the following property*

$$\mu(\pi, m) = \min\{d_{min}(\mathbf{B}), H_{min}(\mathcal{S}(\mathbf{A}) \oplus \mathcal{V}(m, \mathbf{B}))\}. \quad (7)$$

The proof of this theorem is deferred to Appendix B. The following proposition leverages the above theorem to develop a simple, yet elegant technique to construct robust permutations based on linear channel codes with good Hamming distance properties.

Proposition 3. *Consider integers $R \ge 2$ and $m, 2 \le m \le 2^{R-1}$. Let $k = \lceil \log_2 m \rceil$ and $(\mathbf{I}_k \quad \mathbf{P}_{k\times(R-k)})$ be the generator matrix of an (R, k) linear channel code with minimum Hamming distance h. Then the permutation π generated by the following matrix*

$$\mathbf{G} = \begin{pmatrix} \mathbf{0}_{(R-k)\times k} & \mathbf{D}_{R-k} \\ \mathbf{I}_k & \mathbf{P}_{k\times(R-k)} \end{pmatrix}, \quad (8)$$

satisfies the relations

$$h - 1 \le \mu(\pi, m) \le h.$$

Proof: The second inequality follows from Theorem 2. In order to prove the first inequality, first let $\mathbf{A}$ denote the submatrix of $\mathbf{G}$ formed of the first $R-k$ rows of $\mathbf{G}$ and let $\mathbf{B}$ denote the submatrix formed by the remaining rows. Clearly, any vector in $\mathcal{S}(\mathbf{A})$ has Hamming weight 1, thus any vector in $\mathcal{S}(\mathbf{A}) \oplus \mathcal{V}(m, \mathbf{B})$ has Hamming weight at least $d_{min}(\mathbf{B}) - 1$. Now the conclusion follows by applying Theorem 2. ∎

Remark 2. *The above result together with Theorem 2 imply that the largest value of $\mu(\pi, m)$ achievable with a linear permutation π of the set $\mathcal{I}_R$ is either equal to or one unit less than the largest minimum Hamming distance of a linear $(R, \lceil \log_2 m \rceil)$ channel code.*

Using the technique developed in Proposition 3 linear permutations π achieving $\mu(\pi, m) \ge 2$, respectively $\mu(\pi, m) \ge 3$, can be constructed based on shortened Hamming codes [61], [62] of minimum Hamming distance at least 3, respectively 4. The following result establishes implicit lower bounds on the largest values of m in terms of R, for which such constructions are possible.

Corollary 1. 1) *Let $R \ge 3$ and $m \ge 2$ be integers. If $\lceil \log_2 m \rceil \le R - \log_2(R + 1)$ then there are linear permutations $\pi \in \mathcal{I}_R$ achieving $\mu(\pi, m) \ge 2$.*

2) *Let $R \ge 4$ and $m \ge 2$ be integers. If $\lceil \log_2 m \rceil \le R - 1 - \log_2 R$ then then there are linear permutations $\pi \in \mathcal{I}_R$ achieving $\mu(\pi, m) \ge 3$.*

Proof: Let $k = \lceil \log_2 m \rceil$. It is known that the minimum Hamming distance of a linear channel block code equals the smallest number of columns of the parity check matrix that add up to the zero vector [62]. The parity check matrix of a shortened Hamming code has all the columns non-zero and distinct. Therefore, if $R \ge 3$, the minimum Hamming distance of a shortened Hamming code is at least 3. Notice that as long as $R \ge 3$ and $R \le 2^{R-k} - 1$ such a code exists. Finally, note that the above conditions are met when $\lceil \log_2 m \rceil \le R - \log_2(R + 1)$, thus proving the first claim.

In view of Proposition 3, in order to prove the second claim it is sufficient to show that when m satisfies the specified condition there exists an (R, k) shortened Hamming code of minimum Hamming distance at least 4. The columns of the parity check matrix of such a code have odd Hamming weight and are distinct. Such a code exists if $R \ge 4$ and $R \le 2^{R-k-1}$, conditions that are satisfied when $\lceil \log_2 m \rceil \le R - 1 - \log_2 R$. Thus, the conclusion follows. ∎

By exploiting Proposition 3 the highest value of $\mu(\pi, m)$ achievable with linear permutations $\pi \in \mathcal{I}_R$, denoted by $\theta_{lin}(R, m)$, can be determined exactly for certain values of m. The following two results establish the value of $\theta_{lin}(R, m)$ for $m = 2, 3, 4$. Additionally, they also determine the exact value or tight bounds for the largest $\mu(\pi, m)$ over all permutations $\pi \in \mathcal{I}_R$, denoted by $\theta(R, m)$, for $m = 2, 3, 4$. The proofs of the following two propositions are deferred to Appendix B.

Proposition 4. *For any $R \ge 2$ one has $\theta(R, 2) = \theta_{lin}(R, 2) = R - 1$.*

Proposition 5. *The following statements hold.*

1) *For any $n \ge 1$, $R = 3n + 1$ and $m = 3, 4$, one has*

$$\theta_{lin}(R, m) = \theta(R, m) = \left\lfloor \frac{2R}{3} \right\rfloor. \quad (9)$$

2) *For any $n \ge 1$, $R = 3n$ or $R = 3n + 2$ and $m = 3, 4$, one has*

$$\left\lfloor \frac{2R}{3} \right\rfloor - 1 = \theta_{lin}(R, m) \le \theta(R, m) \le \left\lfloor \frac{2R}{3} \right\rfloor. \quad (10)$$

In view of the connection between $\mu(\pi, m)$ and the antibandwidth problem highlighted by Theorem 1, the results obtained in this section can be used to determine bounds on the antibandwidth of the graph $Q_R(d)$ for certain values of $d \ge 3$. Since the antibandwith for such graphs is not known, establishing such bounds represents an interesting result, which is stated in the following corollary.

Corollary 2. 1) *For $R \ge 4$, one has $ab(Q_R(3)) \ge 2^{R-1-\lceil \log_2 R \rceil}$.*

2) *For $R \ge 2$, one has $ab(Q_R(R - 1)) \ge 2$ and $ab(Q_R(R)) = 1$.*

3) *For $R \ge 3$, one has $ab(Q_R(\lfloor \frac{2R}{3} \rfloor + 1)) \le 2$ and $ab(Q_R(\lfloor \frac{2(R-1)}{3} \rfloor)) \ge 4$.*

Proof: Notice first that Theorem 1 and the fact that $\theta(R, m) \ge \theta_{lin}(R, m)$ imply that if $\theta_{lin}(R, m) \ge d$ then

$ab(Q_R(d)) \geq m$. Then the first claim follows from the second claim of Proposition 1. The second claim follows immediately from Proposition 4. To prove the third claim we use Proposition 5. The fact that $\theta(R,3) \leq \lfloor \frac{2R}{3} \rfloor$ implies that $ab(Q_R(\lfloor \frac{2R}{3} \rfloor + 1)) \leq 2$. Further, according to Proposition 5, we have $\theta(R,4) \geq \lfloor \frac{2(R-1)}{3} \rfloor$. Therefore, we obtain that $ab(Q_R(\lfloor \frac{2(R-1)}{3} \rfloor)) \geq 4$. ■

VI. Permutations for Achieving a Minimum Hamming Distance of 3

In this section we address the scenario when both descriptions may be corrupted by errors. Specifically, we construct linear permutations achieving $d_{min} \geq 3$ and establish an implicit bound in terms of R, for the maximum m for which such permutations exist. The proofs are deferred to Appendix C.

Proposition 6. *Let R and $m \geq 2$ be positive integers such that $R \geq 5 + \lceil \log_2 m \rceil + \log_2(\lceil \log_2 m \rceil + 2)$. Let $k = \lceil \log_2 m \rceil$ and consider the generator matrices*

$$\mathbf{G}_1 = \begin{pmatrix} \mathbf{0}_{(R-k)\times k} & \mathbf{D}_{R-k} \\ \mathbf{I}_k & \mathbf{P}_1 \end{pmatrix},$$
$$\mathbf{G}_2 = \begin{pmatrix} \mathbf{0}_{(R-k-1)\times(k+1)} & \mathbf{D}_{R-k-1} \\ \mathbf{I}_{k+1} & \mathbf{P}_2 \end{pmatrix}, \quad (11)$$

where $\mathbf{P}_1 \in \mathcal{M}_{k\times(R-k)}$, $\mathbf{P}_2 \in \mathcal{M}_{(k+1)\times(R-k-1)}$, and the following properties are satisfied.

- **C1)** *Any two rows of $\mathbf{P}_1$ are different and the Hamming weight of any row of $\mathbf{P}_1$ is an odd number larger than or equal to 3.*
- **C2)** *If the Hamming weight of a row of $\mathbf{P}_1$ is 3 then the last component of that row is 0.*
- **C3)** *Every row of $\mathbf{P}_1$ starts with $(1,0,1)$.*
- **C4)** *Any two rows of $\mathbf{P}_2$ are different and the Hamming weight of any row of $\mathbf{P}_2$ is an odd number larger than or equal to 3.*
- **C5)** *Every row of $\mathbf{P}_2$ starts with $(0,1,0)$.*

Then the pair $\boldsymbol{\pi} = (\pi_1, \pi_2)$ of linear permutations, where π_i is generated by $\mathbf{G}_i$, $i = 1,2$, satisfies $d_{min}(\boldsymbol{\pi} \circ \alpha) \geq 3$, for any m-diagonal IA α.

Remark 3. *Proposition 6 provides an upper bound for the smallest value of R, for given m, for which there are linear permutation pairs satisfying $d_{min}(\boldsymbol{\pi} \circ \alpha) \geq 3$, for any m-diagonal IA α. For $m = 2$ this upper bound equals 8. The following result shows that for $m = 2$ this upper bound can be lowered to 6.*

Proposition 7. *Let $R \geq 6$ and consider the matrices*

$$\mathbf{G}_1 = \begin{pmatrix} \mathbf{0}_{(R-1)\times 1} & \mathbf{D}_{R-1} \\ 1 & \mathbf{1}_{R-1} \end{pmatrix},$$
$$\mathbf{G}_2 = \begin{pmatrix} \mathbf{0}_{(R-2)\times(2)} & \mathbf{D}_{R-2} \\ 01 & 1\mathbf{0}_{R-6}110 \\ 11 & 0\mathbf{0}_{R-6}011 \end{pmatrix}. \quad (12)$$

Then the pair $\boldsymbol{\pi} = (\pi_1, \pi_2)$ of linear permutations, where π_i is generated by $\mathbf{G}_i$, $i = 1,2$, satisfies $d_{min}(\boldsymbol{\pi} \circ \alpha) \geq 3$, for any 2-diagonal IA α.

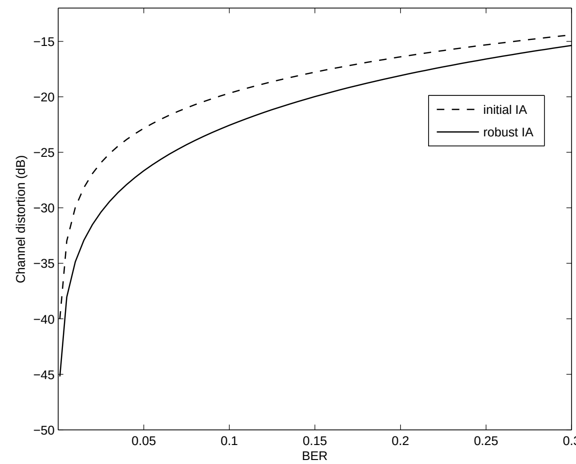

Fig. 3. Performance of a 2-diagonal IA for $R = 3$, before and after applying the proposed permutation, when description 1 is correct.

VII. Experimental Results

The purpose of this section is to assess in practice the performance of the proposed robust IA's in comparison with the initial IA's. Our tests are performed on a zero mean, unit variance, memoryless Gaussian source. In each case, the 2DSQ is optimized by using Vaishampayan's algorithm [1]. We will consider two cases: 1) channel 1 transmits correctly and channel 2 is a binary symmetric channel with bit error rate (BER) ϵ; 2) both channels are independent binary symmetric channels with the same BER ϵ. In both cases the performance of each IA is measured by the channel distortion at the central decoder in dB, i.e., by $10\log_{10} D_{0,C}$, where $D_{0,C}$ is defined in (16) in Appendix A. The BER's considered are in the interval $(0.001, 0.3)$. The decoder used in our experiments is the minimum Hamming distance decoder given by (18) for case 1), respectively (17) for case 2).

Fig. 3 and 4 plot the performance of a robust IA compared to the initial IA for case 1), i.e., when description 1 is known to be correct. For both figures the permutation used to increase the robustness of the IA is based on the construction proposed in Proposition 3. For Fig. 3 the initial IA is 2-diagonal with $R = 3$. The robust IA has $d_{side,min} = 2$ and is obtained by applying the permutation generated by the following matrix

$$\begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

For Fig. 4, the initial IA is 3-diagonal with $R = 6$. The robust IA has $d_{side,min} = 3$ and is obtained by applying the permutation generated by the matrix

$$\begin{pmatrix} 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}.$$

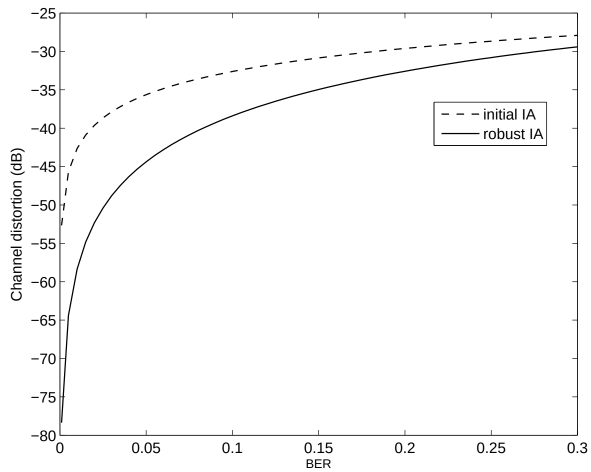

Fig. 4. Performance of a 3-diagonal IA for $R = 6$, before and after applying the proposed permutation, when description 1 is correct.

TABLE III
VALUE OF $D_{0,C}(d)$ FOR THE INITIAL IA ($D_{0,C,init}(d)$) AND FOR THE ROBUST IA ($D_{0,C,robust}(d)$) FOR $0 \leq d \leq 6$.

d	0	1	2	3	4	5	6
$D_{0,C,init}(d)$	0	5	59	311	963	1964	2801
$D_{0,C,robust}(d)$	0	0	143	824	1895	3039	3593

As it can be seen from Figs. 3 and 4 the IA obtained after applying the permutation significantly outperforms the initial IA for all values of the BER in the range $(0.001, 0.3)$. Thus, these results validate the effectiveness of the proposed method for increasing the bit-error resilience for the scenario when one description is known to be correct.

Fig. 5 plots the performance of a proposed robust IA for the case when both descriptions may carry errors. The initial IA is 2-diagonal and the permutations applied to indexes of the two descriptions are constructed as in Proposition 7. After applying the permutations the value of d_{min} becomes 3. As Fig. 5 shows the robust IA outperforms the initial IA only when BER is very small, i.e., $\epsilon < 0.05$ approximately. In an attempt to understand why the IA after permutation becomes worse than the initial IA when $\epsilon > 0.05$ we compute for both IA's the quantities

$$D_{0,C}(d) \triangleq \sum_{i' \in \mathcal{I}_R} \sum_{j' \in \mathcal{I}_R} \sum_{(i,j) \in \mathcal{H}_d(i',j')} P(i,j)(g_0(i,j) - \bar{g}_0(i',j'))^2,$$

where $\mathcal{H}_d(i',j')$ is the set of pairs $(i,j) \in Im(\alpha)$ such that $(\boldsymbol{\beta}(i), \boldsymbol{\beta}(j))$ is at Hamming distance d from $(\boldsymbol{\beta}(i'), \boldsymbol{\beta}(j'))$, for $0 \leq d \leq 2R$. Notice that $D_{0,C}(d)$ represents the distortion incurred by error patterns of Hamming weight d and

$$D_{0,C} = \sum_{d=1}^{2R} \epsilon^d (1-\epsilon)^{2R-d} D_{0,C}(d). \quad (13)$$

Table III lists the values of $D_{0,C}(d)$ for the initial IA ($D_{0,C,init}(d)$) and for the robust IA ($D_{0,C,robust}(d)$) for $0 \leq d \leq 6$. Notice that $D_{0,C,robust}(1) = 0$, which means that all 1-bit errors are corrected with the robust IA. This was expected because $d_{min} = 3$ in this case. The fact that $D_{0,C,robust}(1) < D_{0,C,init}(1)$ is the reason for the superiority of the robust IA when ϵ is sufficiently small, because in such a case the term corresponding to $D_{0,C}(1)$ dominates in (13). On the other hand, notice that $D_{0,C,robust}(d)$ is larger than $D_{0,C,init}(d)$ for all $2 \leq d \leq 6$, and the gap between the two quantities is much larger than $D_{0,C,init}(1) - D_{0,C,robust}(1)$. This is the reason for the better performance of the initial IA for ϵ above 0.05. This observation motivates as a future research direction the design of permutations that beside ensuring $d_{min} = 3$ also achieve sufficiently small values for $D_{0,C}(d)$ for $d > 1$.

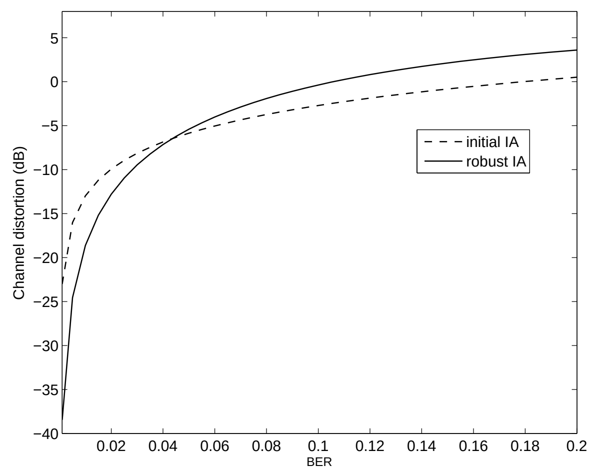

Fig. 5. Performance of a 2-diagonal IA with $R = 6$, before and after applying the permutations, when both descriptions may contain errors.

VIII. CONCLUSION

This work addresses the problem of increasing the bit-error resilience of index assignments (IA) for two description scalar quantizers, without compromising their robustness to description loss. To this end we start from an initial m-diagonal IA and apply a permutation to the indexes of each description. As a performance criterion for the design of such permutations we use the minimum Hamming distance d_{min} of the set of valid index pairs. We present the construction of linear permutations achieving $d_{min} \geq 3$, concurrently determining a lower bound for the largest m for which such permutations exist, for fixed description rate R.

Another scenario we direct our attention to is the case when one description is known to be correct. For such a case we introduce a more appropriate performance criterion. This is the minimum Hamming distance of the set of indexes of one description when the index of the other description is fixed, denoted by $d_{side,min}$. Interestingly, we prove that the problem of designing permutations to increase $d_{side,min}$ is closely related to the antibandwidth problem in a certain graph derived from a hypercube. We further exploit this connection and known results on the hypercube antibandwidth to determine the highest value of m, in terms of R, for which permutations achieving $d_{side,min} \geq 2$ exist, and present their construction. It turns out that this result also settles the existence of permutation pairs achieving $d_{min} \geq 2$ since the latter relation is equivalent to $d_{side,min} \geq 2$.

Finally, we investigate the construction of linear permutations that increase $d_{side,min}$. We develop a simple technique to construct linear permutations satisfying $d_{side,min} \geq h$ based on linear $(R, \lceil \log_2 m \rceil)$ channel codes with minimum Hamming distance $h+1$. By exploiting this result we further establish a lower bound in terms of R, for the highest m for which there are permutations achieving $d_{side,min} \geq 3$. Additionally, tight bounds on the maximum value of $d_{side,min}$ achievable via general permutations are further derived for $m = 2, 3, 4$.

Future work directions include considering more than two descriptions and/or a vector central quantizer instead of a scalar quantizer.

Appendix A
Decoder for 2DSQ with Bit Errors

In this appendix we discuss the central decoder of a 2DSQ for the case when the descriptions may be corrupted by errors. For this let us review first the central distortion $D_{0,S}$ in the conventional case without bit errors,

$$D_{0,S} = \sum_{k=0}^{N-1} \int_{A_k^0} (x - g_0(\alpha(k)))^2 f_X(x)\, dx, \tag{14}$$

where $A_k^0 = \{x | q(x) = k\}$ and $f_X(x)$ denotes the probability density function of the source, which is assume to be i.i.d. It is well-known that the decoding mapping g_0 which minimizes $D_{0,S}$, is the mapping defined by the centroid condition [1], i.e.

$$g_0(\alpha(k)) = \mu(A_k^0), \tag{15}$$

where

$$\mu(A) \triangleq \frac{\int_A x f_X(x)\, dx}{\int_A f_X(x)\, dx}.$$

In the case when one or both channels may incur bit errors, the received index pair (i', j') is not necessarily in $Im(\alpha)$. Therefore, the decoding mapping has to be defined for all pairs $(i', j') \in \mathcal{I}_R \times \mathcal{I}_R$. Let $\bar{g}_0 : \mathcal{I}_R \times \mathcal{I}_R \to \mathbb{R}$ denote the decoding mapping. Let the channel distortion $D_{0,C}$ be defined as

$$D_{0,C} = \sum_{(i,j) \in Im(\alpha)} \sum_{i' \in \mathcal{I}_R} \sum_{j' \in \mathcal{I}_R} P(i,j) P_e(i', j' | i, j) \times (g_0(i,j) - \bar{g}_0(i', j'))^2, \tag{16}$$

where $P(i,j)$ is the probability that (i,j) is output by the 2DSQ and $P_e(i', j'|i,j)$ is the conditional probability that the index pair (i', j') is received conditioned on (i,j) being sent. Note that $P(i,j) = \int_{A_k^0} f_X(x)\, dx$, where $\alpha(k) = (i,j)$. We will assume throughout this work that all bit errors are i.i.d. with bit error rate (BER) ϵ. Then

$$P_e(i', j'|i,j) = \epsilon^d (1-\epsilon)^{2R-d},$$

where $d = d(\boldsymbol{\beta}(i), \boldsymbol{\beta}(i')) + d(\boldsymbol{\beta}(j), \boldsymbol{\beta}(j'))$.

As argued in Section III the expected distortion D_0 at the central decoder can be decomposed as

$$D_0 = D_{0,S} + D_{0,C}.$$

Notice that $D_{0,S}$ is not affected by the mapping $\bar{g}_0$. Thus, the optimal decoder which minimizes the expected distortion D_0 has to minimize $D_{0,C}$, leading to

$$\bar{g}_{0,opt}(i', j') = \arg\min_{y \in \mathbb{R}} \sum_{(i,j) \in Im(\alpha)} P(i,j) P_e(i', j'|i,j) \times (g_0(i,j) - y)^2,$$

and, further, to

$$\bar{g}_{0,opt}(i', j') = \frac{\sum_{(i,j) \in Im(\alpha)} P(i,j) P_e(i', j'|i,j) g_0(i,j)}{\sum_{(i,j) \in Im(\alpha)} P(i,j) P_e(i', j'|i,j)},$$

for every $(i', j') \in \mathcal{I}_R \times \mathcal{I}_R$. Clearly, the optimal decoder $\bar{g}_{0,opt}$ needs knowledge of the BER ϵ. Therefore, this decoder may lead to performance degradation if ϵ is not estimated correctly. To avoid this we propose a different decoder which does not need knowledge of ϵ and is close to the optimum when ϵ is small. In order to introduce the proposed decoder we need the following notation. For each received index pair $(i', j') \in \mathcal{I}_R \times \mathcal{I}_R$, let $\mathcal{H}(i', j')$ denote the set of pairs $(i,j) \in Im(\alpha)$ such that $(\boldsymbol{\beta}(i), \boldsymbol{\beta}(j))$ is situated at the smallest Hamming distance from $(\boldsymbol{\beta}(i'), \boldsymbol{\beta}(j'))$. Then define $\bar{g}_0(i', j')$ as

$$\bar{g}_0(i', j') = \frac{\sum_{(i,j) \in \mathcal{H}(i',j')} g_0(i,j) P(i,j)}{\sum_{(i,j) \in \mathcal{H}(i',j')} P(i,j)}. \tag{17}$$

The proposed decoder is in essence a minimum Hamming distance decoder. Precisely, the decoder looks in $Im(\alpha)$ for the index pairs (i,j) whose binary representation is closest in Hamming distance to that of the received pair. If only one such pair (i,j) is found, then its reconstruction is used. On the other hand, if more such pairs are found then the reconstruction is computed as the mean of the reconstructions corresponding to all such pairs. It is easy to see that for every $(i', j') \in \mathcal{I}_R \times \mathcal{I}_R$ the following relation holds

$$\lim_{\epsilon \to 0} \bar{g}_{0,opt}(i', j') = \bar{g}_0(i', j').$$

In the above discussion we have assumed that both descriptions may be affected by errors. Another scenario of interest is when one description is known to be correct, and only the other description may be corrupted by errors. Let us assume that only the second channel may incur errors and let (i, j') denote the received index pair. Then index i acts as side information, indicating that the transmitted index j must be an element of the set $\mathcal{J}(i) \triangleq \{j : (i,j) \in Im(\alpha)\}$. Let $\bar{g}_{0,2} : \mathcal{I}_R \times \mathcal{I}_R \to \mathbb{R}$ denote the decoding mapping in this scenario. Then $D_{0,C}$ becomes

$$D_{0,C} = \sum_{i \in \mathcal{I}_R} \sum_{j \in \mathcal{J}(i)} \sum_{j' \in \mathcal{I}_R} P(i,j) P_e(j'|j) (g_0(i,j) - \bar{g}_{0,2}(i,j'))^2,$$

where $P_e(j'|j) = \epsilon^d (1-\epsilon)^{R-d}$, for $d = d(\boldsymbol{\beta}(j), \boldsymbol{\beta}(j'))$. Then the decoder that minimizes $D_{0,C}$ is

$$\bar{g}_{0,2,opt}(i,j') = \frac{\sum\limits_{j\in\mathcal{J}(i)} P(i,j)P_e(j'|j)g_0(i,j)}{\sum\limits_{j\in\mathcal{J}(i)} P(i,j)P_e(j'|j)}.$$

The above decoder also needs knowledge of ϵ and may lead to loss in performance if ϵ is not estimated correctly. Therefore, we propose the following decoder, which is close to the optimum for small ϵ.

$$\bar{g}_{0,2}(i,j') = \frac{\sum\limits_{j\in\mathcal{H}_i(j')} g_0(i,j)P(i,j)}{\sum\limits_{j\in\mathcal{H}_i(j')} P(i,j)}, \tag{18}$$

where $\mathcal{H}_i(j')$ denotes the set of indexes $j \in \mathcal{J}(i)$ such that $\boldsymbol{\beta}(j)$ is situated at the smallest Hamming distance from $\boldsymbol{\beta}(j')$. The proposed decoder can also be regarded as a minimum Hamming distance decoder since, if there is only one index j such that $\boldsymbol{\beta}(j)$ is closest in Hamming distance to $\boldsymbol{\beta}(j')$, its corresponding reconstruction is used for decoding. If more such j's are found then the reconstruction is computed as the mean of the reconstructions corresponding to all such j's. It can be easily seen that for every $(i,j') \in \mathcal{I}_R \times \mathcal{I}_R$ the following relation holds

$$\lim_{\epsilon\to 0} \bar{g}_{0,2,opt}(i,j') = \bar{g}_{0,2}(i,j').$$

Appendix B
Proof of Results in Section V

In this appendix we present the proofs of Theorem 2 and Propositions 4 and 5.

Proof of Theorem 2. We will first show that inequality "$\geq$" holds between the two sides of (7). For this we have to show that for any $j_1, j_2 \in \mathcal{I}_R$ with $0 < j_2 - j_1 \leq m-1$ the following inequality is satisfied

$$d(\boldsymbol{\beta}(\pi(j_1)), \boldsymbol{\beta}(\pi(j_2))) \geq \min\{d_{min}(\mathbf{B}), H_{min}(\mathcal{S}(\mathbf{A}) \oplus \mathcal{V}(m,\mathbf{B}))\}. \tag{19}$$

Before proceeding to the formal proof we first present a brief description of the proof idea. Consider partitioning the set $\mathcal{I}_R$ into 2^{R-k} subsets $\mathcal{J}_c \triangleq \{n \in \mathbb{N} : c \times 2^k \leq n < (c+1) \times 2^k\}$, for $0 \leq c \leq 2^{R-k} - 1$. Notice that all elements in a subset have the same bits in the first $R-k$ positions of their binary representation. Therefore, the mapping $\boldsymbol{\beta} \circ \pi$ maps each subset $\mathcal{J}_c$ onto a coset of the channel code generated by matrix $\mathbf{B}$. Each coset is obtained by shifting the channel code by some vector in $\mathbb{F}_2^R$, therefore the minimum Hamming distance between any binary vectors in each coset equals $d_{min}(\mathbf{B})$. It follows that, if integers j_1 and j_2 fall in the same subset $\mathcal{J}_c$, relation (19) is satisfied. The remaining case is when j_1 is in some $\mathcal{J}_c$ and j_2 is in $\mathcal{J}_{c+1}$. Then the vector $\boldsymbol{\beta}(\pi(j_2))$ can be obtained from $\boldsymbol{\beta}(\pi(j_1))$ by adding the shift vector between the cosets corresponding to $\mathcal{J}_c$ and $\mathcal{J}_{c+1}$, and a vector in the channel code. Finally, it turns out that the former vector is an element of the set $\mathcal{S}(\mathbf{A})$, while the latter vector is in $\mathcal{V}(m,\mathbf{B})$. Thus, relation (19) is satisfied.

Let us now proceed to the formal proof. First notice that the linearity of the permutation implies that

$$d(\boldsymbol{\beta}(\pi(j_1)), \boldsymbol{\beta}(\pi(j_2))) = H((\boldsymbol{\beta}(j_1) \oplus \boldsymbol{\beta}(j_2))\mathbf{G}). \tag{20}$$

Let us fix arbitrary j_1 and j_2 as above and consider the unique integers c and e such that $j_1 = c \times 2^k + e$, $0 \leq c \leq 2^{R-k} - 1$ and $0 \leq e \leq 2^k - 1$ (thus, $j_1 \in \mathcal{J}_c$). Let $\tau \triangleq j_2 - j_1$. Thus, $j_2 = j_1 + \tau$ and $\tau \leq m-1$. Then $\boldsymbol{\beta}(j_1) = (\boldsymbol{\beta}_{R-k}(c), \boldsymbol{\beta}_k(e))$ and one of the following two cases is possible: $e+1 \leq e+\tau \leq 2^k - 1$ (i.e., $j_2 \in \mathcal{J}_c$) or $2^k \leq e+\tau \leq 2^{k+1}-1$ (i.e., $j_2 \in \mathcal{J}_{c+1}$).

Case 1: $e+1 \leq e+\tau \leq 2^k-1$. Then $\boldsymbol{\beta}(j_2) = (\boldsymbol{\beta}_{R-k}(c), \boldsymbol{\beta}_k(e+\tau))$, leading to $\boldsymbol{\beta}(j_1) \oplus \boldsymbol{\beta}(j_2) = (\mathbf{0}_{R-k}, \boldsymbol{\beta}_k(e) \oplus \boldsymbol{\beta}_k(e+\tau))$, and $\boldsymbol{\beta}_k(e) \oplus \boldsymbol{\beta}_k(e+\tau) \neq \mathbf{0}_k$. This implies that

$$H((\boldsymbol{\beta}(j_1) \oplus \boldsymbol{\beta}(j_2))\mathbf{G}) = H((\boldsymbol{\beta}_k(e) \oplus \boldsymbol{\beta}_k(e+\tau))\mathbf{B}) \geq d_{min}(\mathbf{B}). \tag{21}$$

Case 2: $2^k \leq e+\tau \leq 2^{k+1}-1$. Then one has $\boldsymbol{\beta}(j_2) = (\boldsymbol{\beta}_{R-k}(c+1), \boldsymbol{\beta}_k(e+\tau-2^k))$. Additionally, relation $\boldsymbol{\beta}_{R-k}(c+1) = \boldsymbol{\beta}_{R-k}(c) \oplus (\mathbf{0}_{s-1}, \mathbf{1}_{R-k-s+1})$ holds, where s denotes the position of the rightmost 0 in $\boldsymbol{\beta}_{R-k}(c)$ (thus, $1 \leq s \leq R-k$). Note that $\boldsymbol{\beta}_{R-k}(c)$ must contain at least one 0 since $c < 2^{R-k} - 1$. It follows that

$$(\boldsymbol{\beta}(j_1) \oplus \boldsymbol{\beta}(j_2))\mathbf{G} = (\mathbf{0}_{s-1}, \mathbf{1}_{R-k-s+1})\mathbf{A} \oplus \mathbf{b}\mathbf{B}, \tag{22}$$

where $\mathbf{b} = \boldsymbol{\beta}_k(e) \oplus \boldsymbol{\beta}_k(e+\tau-2^k)$. Clearly, $(\mathbf{0}_{s-1}, \mathbf{1}_{R-k-s+1})\mathbf{A} \in \mathcal{S}(\mathbf{A})$. Additionally, using the fact that $\eta(\mathbf{b}_1 \oplus \mathbf{b}_2) \leq \eta(\mathbf{b}_1) + \eta(\mathbf{b}_2)$ for any $\mathbf{b}_1, \mathbf{b}_2 \in \mathbb{F}_2^k$, it follows that $e = \eta(\mathbf{b} \oplus \boldsymbol{\beta}_k(e+\tau-2^k)) \leq \eta(\mathbf{b}) + (e+\tau-2^k)$. This implies that $\eta(\mathbf{b}) \geq 2^k - \tau \geq 2^k - m + 1$, i.e. that $\mathbf{b}\mathbf{B} \in \mathcal{V}(m,\mathbf{B})$. Thus, (22) leads to

$$H((\boldsymbol{\beta}(j_1) \oplus \boldsymbol{\beta}(j_2))\mathbf{G}) \geq H_{min}(\mathcal{S}(\mathbf{A}) \oplus \mathcal{V}(m,\mathbf{B})). \tag{23}$$

Relations (21) and (23) imply (19).

Next we will show that

$$\mu(\pi, m) \leq d_{min}(\mathbf{B}). \tag{24}$$

For this it is sufficient to prove that for any $\mathbf{b} = (b_1, \cdots, b_k) \in \mathbb{F}_2^k \setminus \{\mathbf{0}_k\}$ there are $j_1 \in \{0, \cdots, 2^R - 2\}$ and $\tau \in \{1, 2, \cdots, \min(m-1, 2^R - j_1 - 1)\}$ such that $\boldsymbol{\beta}(j_1) \oplus \boldsymbol{\beta}(j_1+\tau) = (\mathbf{0}_{R-k}, \mathbf{b})$. Assume first that $b_1 = 0$. Then $1 \leq \eta(\mathbf{b}) \leq 2^{k-1} - 1 \leq m-1$. Thus, we may take $\tau = \eta(\mathbf{b})$. Let $j_1 = c \times 2^k$ for some c such that $0 \leq c \leq 2^{R-k} - 1$. Then $j_1 + \tau = c \times 2^k + \eta(\mathbf{b})$ and it is clear that the requirements on j_1 and τ are satisfied.

Now assume that $b_1 = 1$. Then $2^k - 1 \geq \eta(\mathbf{b}) \geq 2^{k-1}$, which implies that $1 \leq 2^k - \eta(\mathbf{b}) \leq 2^{k-1} \leq m-1$. Thus, we may take $\tau = 2^k - \eta(\mathbf{b})$ and $j_1 = c \times 2^k + \eta(\mathbf{b}) - 2^{k-1}$ for some c such that $0 \leq c \leq 2^{R-k} - 1$. Then $j_1 + \tau = c \times 2^k + 2^{k-1}$ and it follows that the desired requirements are satisfied. Thus, the proof of (24) is completed.

To complete the proof of the theorem it remains to show that

$$\mu(\pi, m) \leq H_{min}(\mathcal{S}(\mathbf{A}) \oplus \mathcal{V}(m,\mathbf{B})). \tag{25}$$

For this it is sufficient to prove that for any s, $1 \leq s \leq R-k$, and any $\mathbf{b} \in \mathbb{F}_2^k$ satisfying $\eta(\mathbf{b}) \geq 2^k - m + 1$, there are integers $j_1 \in \{0, \cdots, 2^R - 2\}$ and

$\tau \in \{1, 2, \cdots, \min(m-1, 2^R - j_1 - 1)\}$ such that $\boldsymbol{\beta}(j_1) \oplus \boldsymbol{\beta}(j_1+\tau) = (\mathbf{0}_{s-1}, \mathbf{1}_{R-k-s+1}, \mathbf{b})$. Let us fix such an s and $\mathbf{b}$. The fact that $\eta(\mathbf{b}) \geq 2^k - m + 1$ implies that $2^k - \eta(\mathbf{b}) \leq m-1$. Thus, we may choose $\tau = 2^k - \eta(\mathbf{b})$ and $j_1 = c \times 2^k + \eta(\mathbf{b})$, where $c = 2^{R-k-s} - 1$. Then $j_1 + \tau = (c+1) \times 2^k = 2^{R-s} \leq 2^{R-1}$ and $\boldsymbol{\beta}(j_1) \oplus \boldsymbol{\beta}(j_1+\tau) = (\mathbf{0}_{s-1}, \mathbf{1}_{R-k-s+1}, \mathbf{b})$. It follows that $(\boldsymbol{\beta}(j_1) \oplus \boldsymbol{\beta}(j_1+\tau))\mathbf{G} = (\mathbf{0}_{s-1}, \mathbf{1}_{R-k-s+1})\mathbf{A} \oplus \mathbf{bB}$ holds, fact which implies (25). Finally, relations (19), (24) and (25) lead to the conclusion that (7) holds, thus completing the proof. ∎

In the proofs of Propositions 4 and 5 we will use the following definition. For any integers, $R \geq 2$ and $k, 1 \leq k \leq R$, let $\delta_{lin}(R, k)$ denote the largest minimum Hamming distance of linear (R, k) channel code.

Proof of Proposition 4. Notice that any vector in $\mathbb{F}_2^R$ is at distance R from exactly one other vector. Therefore, for any permutation π of the set $\mathbb{F}_2^R$ one has $\min\{d(\boldsymbol{\beta}(\pi(0)), \boldsymbol{\beta}(\pi(1))), d(\boldsymbol{\beta}(\pi(1)), \boldsymbol{\beta}(\pi(2)))\} \leq R-1$, which implies that $\theta(R, 2) \leq R-1$. On the other hand, according to Proposition 3 one has $\theta_{lin}(R, 2) \geq \delta_{lin}(R, 1) - 1 = R-1$. Using further the inequality $\theta_{lin}(R, 2) \leq \theta(R, 2)$, the conclusion of the proposition follows. ∎

In order to prove Proposition 5 we need the following lemma.

Lemma 2. *Let $R \geq 2$. Then $\delta_{lin}(R, 2) = \lfloor \frac{2R}{3} \rfloor$.*

Proof: An $(R, 2)$ linear block code is a set $\mathcal{C} = \{\mathbf{0}_k, \mathbf{u}, \mathbf{v}, \mathbf{u} \oplus \mathbf{v}\}$, where $\mathbf{u}$ and $\mathbf{v}$ are two different non-zero vectors in $\mathbb{F}_2^R$. Let d_1 denote the number of indexes i, $1 \leq i \leq R$, such that $u_i = 1$ and $v_i = 0$. Let d_2 denote the number of indexes i, $1 \leq i \leq R$, such that $u_i = v_i = 1$ and let d_3 denote the number of indexes i, $1 \leq i \leq R$, such that $u_i = 0$ and $v_i = 1$. We may assume that $d_1 + d_2 + d_3 = R$ since if the equality does not hold we may always construct an $(R, 2)$ linear block code with at least the same minimum Hamming distance where the equality holds. Then one has $H(\mathbf{u}) = R - d_3$, $H(\mathbf{v}) = R - d_1$ and $H(\mathbf{u} \oplus \mathbf{v}) = R - d_2$. It follows that $d_{min}(\mathcal{C}) = R - \max\{d_1, d_2, d_3\}$. We conclude that the linear code $\mathcal{C}$ maximizing $d_{min}(\mathcal{C})$ can be found by solving the following optimization problem.

$$
\begin{aligned}
&\min_{d_1, d_2, d_3} \max\{d_1, d_2, d_3\} \\
&\text{subject to } 0 \leq d_1, d_2, d_3 \leq R-1 \\
&\qquad\qquad d_1 + d_2 + d_3 = R \\
&\qquad\qquad d_1, d_2, d_3 \text{ are integers.}
\end{aligned}
$$

Because the problem is symmetric in the three variables we may assume without loss of generality that $d_1 \leq d_2 \leq d_3$. Then the cost function equals d_3. Additionally, we have $R = d_1 + d_2 + d_3 \leq 3d_3$, which implies that $d_3 \geq \lceil \frac{R}{3} \rceil$. It follows that if there is a feasible solution satisfying $d_3 = \lceil \frac{R}{3} \rceil$ then this is the optimal solution. Such a feasible solution must obey the relations

$$0 \leq d_1 \leq d_2 \leq d_3 = \left\lceil \frac{R}{3} \right\rceil, \quad d_1 + d_2 = R - \left\lceil \frac{R}{3} \right\rceil. \tag{26}$$

When $R = 3n$ for some integer $n \geq 1$, the only solution to (26) is $d_1 = d_2 = d_3 = n$. When $R = 3n+2$ for some integer $n \geq 0$, then the only solution to (26) is $d_1 = n$ and $d_2 = d_3 = n+1$. Finally, when $R = 3n+1$ for some integer $n \geq 1$, then the system of relations (26) admits two solutions, namely $(d_1, d_2, d_3) = (n, n, n+1)$ and $(d_1, d_2, d_3) = (n-1, n+1, n+1)$. With these observations, the proof is completed. ∎

Proof of Proposition 5. We will first prove that for any $R \geq 2$ and $m = 3, 4$, one has

$$\left\lfloor \frac{2R}{3} \right\rfloor - 1 \leq \theta_{lin}(R, m) \leq \theta(R, m) \leq \left\lfloor \frac{2R}{3} \right\rfloor. \tag{27}$$

Let π be a permutation of $\mathbb{F}_2^R$. Then $\mu(\pi, 3) \leq \min\{H(\boldsymbol{\beta}(\pi(0)) \oplus \boldsymbol{\beta}(\pi(1))), H(\boldsymbol{\beta}(\pi(1)) \oplus \boldsymbol{\beta}(\pi(2)), H(\boldsymbol{\beta}(\pi(0)) \oplus \boldsymbol{\beta}(\pi(2)))\}$. Since the set $\mathcal{C} = \{\mathbf{0}_R, \boldsymbol{\beta}(\pi(0)) \oplus \boldsymbol{\beta}(\pi(1)), \boldsymbol{\beta}(\pi(1)) \oplus \boldsymbol{\beta}(\pi(2)), \boldsymbol{\beta}(\pi(0)) \oplus \boldsymbol{\beta}(\pi(2))\}$ forms an $(R, 2)$ linear block code it follows that $\mu(\pi, 3) \leq \delta_{lin}(R, 2)$. Using the fact that $\mu(\pi, 4) \leq \mu(\pi, 3)$ and Lemma 2, the last inequality in (27) follows. Corroborating, further, with the fact that $\theta_{lin}(R, m) \leq \theta(R, m)$ and with Proposition 3 the proof is completed.

Let us prove now claim 1). Let $R = 3n+1$ for some $n \geq 1$. Based on (27) and the fact that $\theta_{lin}(R, 4) \leq \theta_{lin}(R, 3)$, in order to prove (9) it is sufficient to construct a linear permutation π of the set $\mathcal{I}_R$, achieving $\mu(\pi, 4) = \lfloor \frac{2R}{3} \rfloor$. For this consider the permutation matrix $\mathbf{G} = \begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix}$, where $\mathbf{A} = (\mathbf{0}_{(R-2)\times 2}, \mathbf{D}_{R-2} \oplus \mathbf{E})$, $\mathbf{E} \in \mathcal{M}_{(R-2)\times(R-2)}$ is the matrix whose only non-zero element is situated on row $R-3$ and column $R-2$, and $\mathbf{B} = \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{pmatrix}$, with $\mathbf{u}_1 = (1, 0, \mathbf{1}_{2n-1}, \mathbf{0}_n)$, $\mathbf{u}_2 = (0, 1, \mathbf{0}_{n-1}, \mathbf{1}_{2n-1}, 0)$. Notice that $det(\mathbf{G}) = det \begin{pmatrix} \mathbf{B} \\ \mathbf{A} \end{pmatrix}$. Since the matrix on the righthand-side is upper triangular with all diagonal elements equal to 1, its determinant is 1. Therefore, one has $rank(\mathbf{G}) = R$.

According to Theorem 2, in order to prove that $\mu(\pi, 4) = \lfloor \frac{2R}{3} \rfloor$ it is sufficient to show that

$$\min\{d_{min}(\mathbf{B}), H_{min}(\mathcal{S}(\mathbf{A}) \oplus \mathcal{V}(4, \mathbf{B}))\} = 2n.$$

The above relation is is obvious when $n = 1$. Let us consider now the case $n > 1$. The fact that $d_{min}(\mathbf{B}) = 2n$ is immediate. Next we will show that $H_{min}(\mathcal{S}(\mathbf{A}) \oplus \mathcal{V}(4, \mathbf{B})) = 2n$. Notice that $\mathcal{V}(4, \mathbf{B}) = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_1 \oplus \mathbf{u}_2\}$. Further notice that $\mathcal{S}(\mathbf{A}) = \{\mathbf{v}_s : 1 \leq s \leq R-2\}$, where

$$
\begin{aligned}
&\mathbf{v}_s = (\mathbf{0}_{s+1}, 1, \mathbf{0}_{3n-s-2}, 1) \text{ when } 1 \leq s \leq 3n-2, \\
&\mathbf{v}_{3n-1} = (\mathbf{0}_{3n-2}, 1).
\end{aligned}
$$

Let $\mathbf{u} \in \mathcal{V}(4, \mathbf{B})$ and $\mathbf{w} \in \mathcal{S}(\mathbf{A})$. Since $u_R = 0$, $w_R = 1$ and $H(\mathbf{w}_1^{R-1}) \leq 1$, it follows that $H(\mathbf{w} \oplus \mathbf{u}) \geq H(\mathbf{u}) \geq 2n$. Thus, the claim is proved.

Let us prove claim 2). Assume that $R = 3n$ or $R = 3n+2$, for some $n \geq 1$. According to relation (27), Lemma 2 and the fact that $\theta_{lin}(R, 4) \leq \theta_{lin}(R, 3)$, it is sufficient to show that for any linear permutation π of $\mathbb{F}_2^R$ one has $\mu(\pi, 3) \neq \delta_{lin}(R, 2)$. Let $\mathbf{G} = \begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix}$ be the permutation matrix of π, where $\mathbf{A} \in \mathcal{M}_{(R-2)\times R}$ and $\mathbf{B} \in \mathcal{M}_{2\times R}$ with $\mathbf{B} = \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}$. Notice that the set $\mathcal{S}(\mathbf{A}) \cup \{\mathbf{u}, \mathbf{v}\}$ forms a basis of the vector space $\mathbb{F}_2^R$.

According to Theorem 2, the equality $\mu(\pi, 3) = \delta_{lin}(R, 2)$ would hold if and only if the following relations were satisfied: $d_{min}(\mathbf{B}) = \delta_{lin}(R, 2)$ and

$$H(\mathbf{w} \oplus \mathbf{u}) \geq \delta_{lin}(R, 2),$$
$$H(\mathbf{w} \oplus \mathbf{u} \oplus \mathbf{v}) \geq \delta_{lin}(R, 2), \quad (28)$$

for any $\mathbf{w} \in \mathcal{S}(\mathbf{A})$. Assume now that $d_{min}(\mathbf{B}) = \delta_{lin}(R, 2)$ holds. We will show that it is impossible to find $R-2$ vectors $\mathbf{w}$ that together with $\mathbf{v}$ form a linearly independent set, and satisfy relations (28). For this let d_1 denote the number of indexes i, $1 \leq i \leq R$, such that $u_i = 1$ and $v_i = 0$. Let d_2 denote the number of indexes i, $1 \leq i \leq R$, such that $u_i = v_i = 1$ and let d_3 denote the number of indexes i, $1 \leq i \leq R$, such that $u_i = 0$ and $v_i = 1$. The fact that $d_{min}(\mathbf{B}) = \delta_{lin}(R, 2)$ implies that $d_1 + d_2 + d_3 = R$ since otherwise all binary vectors in $\mathbf{B}$ would have the value 0 in some position i_0, which would imply that $d_{min}(\mathbf{B}) \leq \delta_{lin}(R-1, 2)$. Lemma 2 implies that $\delta_{lin}(R-1, 2) < \delta_{lin}(R, 2)$, thus leading to a contradiction.

Further, let δ_1 denote the number of indexes i, $1 \leq i \leq R$, such that $u_i = 1$, $v_i = 0$ and $w_i = 1$. Let δ_2 denote the number of indexes i, $1 \leq i \leq R$, such that $u_i = v_i = w_i = 1$ and let δ_3 denote the number of indexes i, $1 \leq i \leq R$, such that $u_i = 0$, $v_i = w_i = 1$. Then relations (28) are equivalent to

$$d_1 - \delta_1 + d_2 - \delta_2 + \delta_3 \geq \delta_{lin}(R, 2),$$
$$d_1 - \delta_1 + d_3 - \delta_3 + \delta_2 \geq \delta_{lin}(R, 2). \quad (29)$$

Next we will differentiate between the cases $R = 3n$ and $R = 3n + 2$.

Case 1. $R = 3n$. Then $\delta_{lin}(R, 2) = 2n$ and according to the proof of Lemma 2 one must have $d_1 = d_2 = d_3 = n$. Substituting in (29) and rearranging one obtains $\delta_3 \geq \delta_1 + \delta_2$ and $\delta_2 \geq \delta_1 + \delta_3$, which lead to $\delta_1 = 0$ and $\delta_2 = \delta_3$. Let $\mathcal{P}$ denote the set of indexes i, $1 \leq i \leq R$, such that $u_i = 1$ and $v_i = 0$. Then condition $\delta_1 = 0$ implies that $w_i = 0$ for all $i \in \mathcal{P}$. The set of vectors $\mathbf{w}$ satisfying the latter condition forms a vector space $\mathcal{T}$ of dimension $R - |\mathcal{P}| = 2n$ and $\mathbf{v}$ is an element of this space as well. Therefore, the maximum number of vectors from $\mathcal{T} \setminus \{\mathbf{v}\}$ that together with $\mathbf{v}$ form a linearly independent set, is $2n - 1$. When $n > 1$ one has $2n-1 < 3n-2 = R-2$ proving the claim. Consider now the case $n = 1$. Assume without loss of generality (w.l.o.g.) that $\mathbf{u} = (1, 1, 0)$ and $\mathbf{v} = (0, 1, 1)$. Then the additional condition $\delta_2 = \delta_3$ implies that the only non-zero vector $\mathbf{w}$ satisfying (28) is $\mathbf{w} = (0, 1, 1) = \mathbf{v}$. Thus, the linear independence condition is violated. With this observation the proof for **Case 1** is completed.

Case 2. $R = 3n+2$. Then $\delta_{lin}(R, 2) = 2n+1$ and according to the proof of Lemma 2 only the following three possibilities exist: $(d_1, d_2, d_3) = (n, n+1, n+1)$, $(d_1, d_2, d_3) = (n+1, n, n+1)$ and $(d_1, d_2, d_3) = (n+1, n+1, n)$. The latter two cases can be treated similarly. Therefore, we will address only the first two. Assume that $(d_1, d_2, d_3) = (n, n+1, n+1)$. Substituting in (29) and rearranging one obtains $\delta_3 \geq \delta_1 + \delta_2$ and $\delta_2 \geq \delta_1 + \delta_3$, leading to $\delta_1 = 0$ and $\delta_2 = \delta_3$. By a similar argument as in **Case 1** it follows that the maximum number of vectors satisfying $\delta_1 = 0$, which together with $\mathbf{v}$ form a linearly independent set, is $2n + 1$, value which is larger or equal to $R - 2$ if and only if $n = 1$. Let now $n = 1$ and assume w.r.g. that $\mathbf{u} = (1, 1, 1, 0, 0)$ and $\mathbf{v} = (0, 1, 1, 1, 1)$. Then the only nonzero vectors satisfying the conditions $\delta_1 = 0$ and $\delta_2 = \delta_3$, are $(0, 1, 0, 1, 0)$, $(0, 0, 1, 0, 1)$, $(0, 1, 0, 0, 1)$ $(0, 0, 1, 1, 0)$, $(0, 1, 1, 1, 1)$. Notice that the fifth vector is $\mathbf{v}$, while the sum of the first two vectors and the sum of the following two vectors are each equal to $\mathbf{v}$. It follows that it is not possible to choose $R - 2 = 3$ of these vectors to form a linearly independent set together with $\mathbf{v}$.

Assume now that $(d_1, d_2, d_3) = (n + 1, n, n + 1)$. Substituting in (29) and rearranging one obtains $\delta_3 \geq \delta_1 + \delta_2$ and $1 + \delta_2 \geq \delta_1 + \delta_3$, which further implies that $\delta_1 = 0$. The maximum number of vectors satisfying $\delta_1 = 0$, which together with $\mathbf{v}$ form a linearly independent set, is $2n$, number which is smaller than $R-2$ for any $n \geq 1$. With this observation the proof of **Case 2** and of the proposition are completed. ■

APPENDIX C
PROOF OF RESULTS IN SECTION VI

In this appendix we present the proofs of Propositions 6 and 7.

Proof of Proposition 6. Property **C1** implies that $d_{min}(\mathbf{B}_1) \geq 4$ [61], where $\mathbf{B}_1 = (\mathbf{I}_k, \mathbf{P}_1)$. To see why this holds let $\mathbf{H}_1 = (\mathbf{I}_{R-k}, \mathbf{P}_1^T)$ be a parity check matrix of the linear code generated by $\mathbf{B}_1$. It is known that $d_{min}(\mathbf{B}_1)$ equals the smallest number of columns of $\mathbf{H}_1$ that equal the zero vector. Property **C1** implies that this number cannot be smaller than 4. Further, notice that the total possible number of vectors in $\mathbb{F}_2^{R-k}$ satisfying properties **C1-C3** is $2^{R-k-4} - 1$. Therefore, such a matrix $\mathbf{P}_1$ exists if and only if $R - k \geq 5$ and $k \leq 2^{R-k-4} - 1$, conditions which are satisfied by the hypothesis.

Property **C4** implies that $d_{min}(\mathbf{B}_2) \geq 4$ [61], where $\mathbf{B}_2 = (\mathbf{I}_{k+1}\ \mathbf{P}_2)$. Notice that the total possible number of vectors in $\mathbb{F}_2^{R-k}$ satisfying properties **C4** and **C5** is $2^{R-k-5} - 1$. Therefore, such a matrix $\mathbf{P}_2$ exists if and only if $R - k \geq 6$ and $k + 1 \leq 2^{R-k-5} - 1$, conditions which are satisfied by the hypothesis.

Consider now an m-diagonal IA α, and two distinct pairs of valid indexes $(a, a + \tau)$ and $(a', a' + \tau')$ in $Im(\alpha)$. Then we have to prove that

$$H((\boldsymbol{\beta}(a) \oplus \boldsymbol{\beta}(a'))\mathbf{G}_1) + H((\boldsymbol{\beta}(a + \tau) \oplus \boldsymbol{\beta}(a' + \tau'))\mathbf{G}_2) \geq 3.$$

Using Theorem 2 it follows immediately that $d_{1,min}(\boldsymbol{\pi} \circ \alpha) \geq 3$ and $d_{2,min}(\boldsymbol{\pi} \circ \alpha) \geq 3$, Therefore, it remains to prove that $H((\boldsymbol{\beta}(a) \oplus \boldsymbol{\beta}(a'))\mathbf{G}_1)$ and $H((\boldsymbol{\beta}(a + \tau) \oplus \boldsymbol{\beta}(a' + \tau'))\mathbf{G}_2)$ cannot be equal to 1 simultaneously.

For this let us first determine the set $\mathcal{W} \triangleq \{\mathbf{w} \in \mathbb{F}_2^R : H(\mathbf{w}\mathbf{G}_1) = 1\}$. Clearly, $\mathcal{W} = \{\mathbf{w}_t, 1 \leq t \leq R\}$, where $\mathbf{w}_t$ is the unique vector satisfying $\mathbf{w}_t\mathbf{G}_1 = (\mathbf{0}_{t-1}, 1, \mathbf{0}_{R-t})$, $1 \leq t \leq R$. It follows that

$$\mathbf{w}_t = (\mathbf{0}_{t-k-1}, \mathbf{1}_{R-t+1}, \mathbf{0}_k),\ \text{for}\ k + 1 \leq t \leq R. \quad (30)$$

Additionally, one has

$$\mathbf{w}_t = (\mathbf{0}_{R-k+t-1}, 1, \mathbf{0}_{k-t}) \oplus_{i=1}^{h} \mathbf{w}_{t_i+k}, \text{ for } 1 \le t \le k,$$

where h denotes the Hamming weight of the t-th row of matrix $\mathbf{P}_1$, and $t_1, t_2, \cdots, t_h$ are the positions of the non-zero components of the t-th row of matrix $\mathbf{P}_1$, with $1 \le t_1 < t_2 < \cdots < t_h \le R-k$. Property **C1** implies that h is odd, while **C3** leads to $t_1 = 1$ and $t_2 = 3$. Therefore, $\mathbf{w}_t$ becomes

$$\mathbf{w}_t = (11, \mathbf{0}_{t_3-3}, \mathbf{1}_{R-k-t_3+1}, \mathbf{0}_{t-1}, 1, \mathbf{0}_{k-t}) \text{ for } h = 3, \quad (31)$$

$$\mathbf{w}_t = (11, \mathbf{0}_{t_3-3}, \mathbf{1}_{t_4-t_3}, \cdots, \mathbf{0}_{t_h-t_{h-1}}, \mathbf{1}_{R-k-t_h+1}, \mathbf{0}_{t-1}, 1, \mathbf{0}_{k-t}) \text{ for } h \ge 5. \quad (32)$$

To proceed with the proof let $\mathbf{u} \triangleq \boldsymbol{\beta}(a) \oplus \boldsymbol{\beta}(a')$ and $\mathbf{x} \triangleq \mathbf{u} \oplus \boldsymbol{\beta}(a+\tau) \oplus \boldsymbol{\beta}(a'+\tau')$. Next we will determine the form of vector $\mathbf{x}$. For this we will first analyze the form of $\boldsymbol{\beta}(a) \oplus \boldsymbol{\beta}(a+\tau)$. Let $a = c \times 2^k + e$, where $0 \le c \le 2^{R-k}-1$, $0 \le e \le 2^k - 1$. Then $\boldsymbol{\beta}(a) = (\boldsymbol{\beta}_{R-k}(c), \boldsymbol{\beta}_k(e))$ and for $a+\tau$ one of the following three situations is possible: **S1)** $0 \le e+\tau \le 2^k - 1$; **S2)** $-2^k \le e+\tau \le -1$; **S3)** $2^k \le e+\tau \le 2^{k+1}-1$.

If **S1** holds then $\boldsymbol{\beta}(a+\tau) = (\boldsymbol{\beta}_{R-k}(c), \boldsymbol{\beta}_k(e+\tau))$, leading to $\boldsymbol{\beta}(a) \oplus \boldsymbol{\beta}(a+\tau) = (\mathbf{0}_{R-k}, \boldsymbol{\beta}_k(e) \oplus \boldsymbol{\beta}_k(e+\tau))$. In the case **S2** one has $\boldsymbol{\beta}(a+\tau) = (\boldsymbol{\beta}_{R-k}(c-1), \boldsymbol{\beta}_k(e+\tau+2^k))$. Let s denote the position of the rightmost 1 in $\boldsymbol{\beta}_{R-k}(c)$. Then $1 \le s \le R-k$ and $\boldsymbol{\beta}_{R-k}(c-1) = \boldsymbol{\beta}_{R-k}(c) \oplus (\mathbf{0}_{s-1}, \mathbf{1}_{R-k-s+1})$. Finally, in the case **S3** one has $\boldsymbol{\beta}(a+\tau) = (\boldsymbol{\beta}_{R-k}(c+1), \boldsymbol{\beta}_k(e+\tau-2^k))$. Then relation $\boldsymbol{\beta}_{R-k}(c+1) = \boldsymbol{\beta}_{R-k}(c) \oplus (\mathbf{0}_{s-1}, \mathbf{1}_{R-k-s+1})$ holds, where s denotes the position of the rightmost 0 in $\boldsymbol{\beta}_{R-k}(c)$ (thus, $1 \le s \le R-k$).

Further, let $a' = c' \times 2^k + e'$ with $0 \le c' \le 2^{R-k} - 1$, $0 \le e' \le 2^k - 1$. For $i = 1, 2, 3$, we say that $\mathbf{S}i$ holds for $a' + \tau'$ when the relations corresponding to $\mathbf{S}i$ are valid when e and τ are replaced by e' and τ', respectively. Combining now all possible cases for $a+\tau$ and for $a'+\tau'$ we obtain the following possible cases for $\mathbf{x}$.

- **X1)** When **S1** holds for both $a+\tau$ and $a'+\tau'$, then one has $\mathbf{x}_1^{R-k} = \mathbf{0}_{R-k}$.
- **X2)** When a) **S1** holds for $a+\tau$ and either **S2** or **S3** holds for $a'+\tau'$, or b) **S1** holds for $a'+\tau'$ and either **S2** or **S3** holds for $a+\tau$, then $\mathbf{x}_1^{R-k} = (\mathbf{0}_{s-1}, \mathbf{1}_{R-k-s+1})$ for some $1 \le s \le R-k$.
- **X3)** When either **S2** or **S3** holds for $a+\tau$ and either **S2** or **S3** holds for $a'+\tau'$ then $\mathbf{x}_1^{R-k} = \mathbf{0}_{R-k}$ or $\mathbf{x}_1^{R-k} = (\mathbf{0}_{s-1}, \mathbf{1}_{t-s}, \mathbf{0}_{R-k-t+1})$ for some $1 \le s < t \le R-k$.

To complete the proof we have to show that if $\mathbf{u} \in \mathcal{W}$ then $H(\mathbf{y}\mathbf{G}_2) > 1$, where $\mathbf{y} \triangleq \mathbf{u} \oplus \mathbf{x}$. From the form of matrix $\mathbf{G}_2$ it follows that

$$\mathbf{y}\mathbf{G}_2 = \mathbf{y}_1^{R-k-1}\mathbf{A}_2 \oplus \mathbf{y}_{R-k}^{R}\mathbf{B}_2. \quad (33)$$

Additionally, the fact that the number of 1's in the first $k+1$ positions of $\mathbf{y}\mathbf{G}_2$ equals $H(\mathbf{y}_{R-k}^{R})$ implies that

$$H(\mathbf{y}\mathbf{G}_2) \ge H(\mathbf{y}_{R-k}^{R}). \quad (34)$$

Further, relation (33) together with the fact that $H(\mathbf{v}_1 \oplus \mathbf{v}_2) \ge |H(\mathbf{v}_1) - H(\mathbf{v}_2)|$, for any $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{F}_2^R$, lead to

$$H(\mathbf{y}\mathbf{G}_2) \ge |H(\mathbf{y}_1^{R-k-1}\mathbf{A}_2) - H(\mathbf{y}_{R-k}^{R}\mathbf{B}_2)|. \quad (35)$$

Assume now that $\mathbf{u} \in \mathcal{W}$. We need to consider two cases.

Case A1) $\mathbf{u} = \mathbf{w}_{k+t}$ for $1 \le t \le R-k$. Then relation (30) implies that $\mathbf{u} = (\mathbf{0}_{t-1}, \mathbf{1}_{R-k-t+1}, \mathbf{0}_k)$. Notice that $\mathbf{u}_{R-k+1}^{R} = \mathbf{0}_k$, which implies that $e = e'$. Next we need to consider three subcases.

Subcase A1.1) $\mathbf{x}_1^{R-k} = \mathbf{0}$. Then $\mathbf{y}_1^{R-k} = (\mathbf{0}_{t-1}, \mathbf{1}_{R-k-t+1})$. This implies that $y_{R-k} = 1$, yielding $\mathbf{y}_{R-k}^{R}\mathbf{B}_2 \ge d_{min}(\mathbf{B}_2) \ge 4$. Further, notice that when $t = R-k$ then $\mathbf{y}_1^{R-k-1}\mathbf{A}_2 = \mathbf{0}_R$, while for $t < R-k$ one has $\mathbf{y}_1^{R-k-1}\mathbf{A}_2 = (\mathbf{0}_{k+t}, 1, \mathbf{0}_{R-1-k-t})$. We conclude that $H(\mathbf{y}_1^{R-k-1}\mathbf{A}_2) \le 1$. Applying now (35) one obtains that $H(\mathbf{y}\mathbf{G}_2) \ge 3$.

Subcase A1.2) $\mathbf{x}_1^{R-k} = (\mathbf{0}_{s-1}, \mathbf{1}_{R-k-s+1})$ for some $1 \le s \le R-k$. We will first show that

$$H(\mathbf{y}_1^{R-k-1}\mathbf{A}_2) \le 2. \quad (36)$$

If $s = t$ then $\mathbf{y}_1^{R-k} = \mathbf{0}_{R-k}$, thus (36) holds. Consider now the case $s \ne t$. Assume without loss of generality that $s < t$. Then one has $\mathbf{y}_1^{R-k-1} = (\mathbf{0}_{s-1}, \mathbf{1}_{t-s}, \mathbf{0}_{R-k-t})$ and (36) holds.

Next we will show that $\mathbf{y}_{R-k+1}^{R} \ne \mathbf{0}_k$. Recall that $e = e'$ and that **X2** holds. Assume that **S1** holds for $a+\tau$ and **S2** holds for $a'+\tau'$. Then $\mathbf{y}_{R-k+1}^{R} = \boldsymbol{\beta}_k(e) \oplus \boldsymbol{\beta}_k(e'+\tau'+2^k)$. Since $e'+\tau'+2^k > e' = e$ it follows that $\boldsymbol{\beta}_k(e) \oplus \boldsymbol{\beta}_k(e'+\tau'+2^k) \ne \mathbf{0}_k$ proving our claim. Assume now that **S1** holds for $a+\tau$ and **S3** holds for $a'+\tau'$. Then $\mathbf{y}_{R-k+1}^{R} = \boldsymbol{\beta}_k(e) \oplus \boldsymbol{\beta}_k(e'+\tau'-2^k) \ne \mathbf{0}_k$ since $e'+\tau'-2^k < e' = e$. The remaining cases of **X2** lead similarly to the desired conclusion.

The fact that $\mathbf{y}_{R-k+1}^{R} \ne \mathbf{0}_k$ implies that $H(\mathbf{y}_{R-k}^{R}\mathbf{B}_2) \ge d_{min}(\mathbf{B}_2)$. Corroborating with (36) and (35) one obtains that $H(\mathbf{y}\mathbf{G}_2) \ge 2$, thus completing the proof.

Subcase A1.3) $\mathbf{x}_1^{R-k} = (\mathbf{0}_{s-1}, \mathbf{1}_{q-s}, \mathbf{0}_{R-k-q+1})$ for some $1 \le s < q \le R-k$. If $t = s$ then $\mathbf{y}_1^{R-k} = (\mathbf{0}_{q-1}, \mathbf{1}_{R-k-q+1})$, while when $t = q$ we have $\mathbf{y}_1^{R-k} = (\mathbf{0}_{s-1}, \mathbf{1}_{R-k-s+1})$. Both these cases can be treated like **A1.1**.

Consider now the case $t < s$ (the discussions for $s < t < q$ and for $q < t$ are similar). Then $\mathbf{y}_1^{R-k} = (\mathbf{0}_{t-1}, \mathbf{1}_{s-t}, \mathbf{0}_{q-s}, \mathbf{1}_{R-k-q+1})$. Notice that $y_{R-k} = 1$. Therefore, $H(\mathbf{y}_{R-k}^{R}\mathbf{B}_2) \ge d_{min}(\mathbf{B}_2) \ge 4$. If $q = R-k$ then $H(\mathbf{y}_1^{R-k-1}\mathbf{A}_2) = 2$ and the claim follows via (35).

Now consider the case when $q < R-k$. We will first show that $\mathbf{y}_{R-k+1}^{R} \ne \mathbf{0}_k$. Notice that if **S2** held for both $a+\tau$ and $a'+\tau'$, the fact that $s, q < R-k$ would imply that both $\boldsymbol{\beta}_{R-k}(c)$ and $\boldsymbol{\beta}_{R-k}(c')$ end with a 0, thus contradicting the fact that $u_{R-k} = 1$. We conclude that **S2** cannot hold for both $a+\tau$ and $a'+\tau'$. It can be similarly shown that **S3** cannot hold for both $a+\tau$ and $a'+\tau'$. Therefore, since we are in the case **X3** we may assume without loss of generality that **S2** holds for $a+\tau$ and **S3** holds for $a'+\tau'$. Then $\mathbf{y}_{R-k+1}^{R} = \boldsymbol{\beta}_k(e+\tau+2^k) \oplus \boldsymbol{\beta}_k(e'+\tau'-2^k)$. Since $e'+\tau'-2^k < e' = e < e+\tau+2^k$ it follows that $\mathbf{y}_{R-k+1}^{R} \ne \mathbf{0}_k$. Using further the fact that $y_{R-k} = 1$ it follows that $H(\mathbf{y}_{R-k}^{R}) \ge 2$ and the claim follows via (34).

Case A2) $\mathbf{u} = \mathbf{w}_t$ for $1 \le t \le k$. Then from (32) and (31) it follows that $\mathbf{u}_1^3 = (1, 1, 0)$ and $u_{R-k} = 1$. Examining all possibilities for $\mathbf{x}$ we conclude that $\mathbf{x}_1^3 \in \mathbb{F}_2^3 \setminus \{(1, 0, 1)\}$, which implies that $\mathbf{y}_1^3 = \mathbf{u}_1^3 \oplus \mathbf{x}_1^3 \in \mathbb{F}_2^3 \setminus \{(0, 1, 1)\}$. Denote now $\mathbf{z} \triangleq \mathbf{y}_1^{R-k-1}\mathbf{A}_2$. Then $\mathbf{z}_{k+2}^{k+4} \in \mathbb{F}_2^3 \setminus \{(0, 1, 0)\}$. Assume now that $H(\mathbf{y}_{R-k}^{R}) = 1$. Then property **C5** and (33) imply that

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TIT.2015.2413780

$\mathbf{y}\mathbf{G}_2$ has the value 1 in at least one of the positions $k+2, k+3$ or $k+4$. Additionally, the vector $\mathbf{y}\mathbf{G}_2$ has one of the first $k+1$ components equal to 1, yielding $H(\mathbf{y}\mathbf{G}_2) \geq 2$.

By (34), it remains to discuss only the case when $\mathbf{y}_{R-k}^R = \mathbf{0}_k$. Then $\mathbf{y}\mathbf{G}_2 = \mathbf{z}$ and the non-trivial cases are when $H(\mathbf{z}_{k+2}^{k+4}) \leq 1$, i.e. when $\mathbf{x}_1^3 \in \{(1,1,0),(0,1,0),(1,0,0),(1,1,1)\}$. Further, notice that we have $y_{R-k} = 0$ and $u_{R-k} = 1$, implying that $x_{R-k} = 1$, which rules out case **X3**, therefore $\mathbf{x}_1^3 \notin \{(1,1,0),(0,1,0),(1,0,0)\}$. Now consider $\mathbf{x}_1^3 = (1,1,1)$. Then **X2** must hold with $s = 1$. Assume now that (31) holds. Then we have $\mathbf{y}_1^{R-k-1} = (0,0,\mathbf{1}_{t_3-3},\mathbf{0}_{R-k-t_3})$. Note that by property **C2**, relation $t_3 < R-k$ is valid implying that $H(\mathbf{z}) \geq 2$. The same conclusion follows when (32) holds, thus completing the proof of the theorem. ■

Proof of Proposition 7. Let $\mathcal{W} = \{\mathbf{w}_q, 1 \leq q \leq R\}$, where $\mathbf{w}_q$ is the unique vector satisfying $\mathbf{w}_q\mathbf{G}_1 = (\mathbf{0}_{q-1},1,\mathbf{0}_{R-q})$, $1 \leq q \leq R$. Then

$$\mathbf{w}_q = (\mathbf{0}_{q-2},\mathbf{1}_{R-q+1},0) \text{ for } 2 \leq q \leq R. \tag{37}$$

and

$$\mathbf{w}_1 = \begin{cases} ((\mathbf{10})_{\frac{R-1}{2}},1) \text{ if } R \text{ is odd} \\ ((\mathbf{10})_{\frac{R-2}{2}},11) \text{ if } R \text{ is even} \end{cases}. \tag{38}$$

Further, let $\mathcal{W}' = \{\mathbf{w}'_q, 1 \leq q \leq R\}$, where $\mathbf{w}'_q$ is the unique vector satisfying $\mathbf{w}_q\mathbf{G}_2 = (\mathbf{0}_{q-1},1,\mathbf{0}_{R-q})$, $1 \leq q \leq R$. Then

$$\mathbf{w}'_q = (\mathbf{0}_{q-3},\mathbf{1}_{R-q+1},00) \text{ for } 3 \leq q \leq R. \tag{39}$$

and

$$\mathbf{w}'_1 = (\mathbf{1}_{R-5},00111) \text{ and } \mathbf{w}'_2 = (\mathbf{1}_{R-5},01110). \tag{40}$$

Consider a 2-diagonal IA α satisfying Definition 1, and two distinct pairs of valid indexes $(a, a+\tau)$ and $(a', a'+\tau')$ in $Im(\alpha)$. Clearly, we have $\tau, \tau' \in \{0,1\}$. Further, denote $\mathbf{u} \triangleq \boldsymbol{\beta}(a) \oplus \boldsymbol{\beta}(a')$ and $\mathbf{x} \triangleq \mathbf{u} \oplus \boldsymbol{\beta}(a+\tau) \oplus \boldsymbol{\beta}(a'+\tau')$. Since Theorem 2 implies that $d_{1,min}(\boldsymbol{\pi} \circ \alpha) \geq 3$ and $d_{2,min}(\boldsymbol{\pi} \circ \alpha) \geq 3$, it only remains to prove that $H(\mathbf{u})$ and $H(\mathbf{u} \oplus \mathbf{x})$ cannot be equal to 1 simultaneously.

To determine the form of $\boldsymbol{\beta}(a+\tau)$ and $\boldsymbol{\beta}(a'+\tau')$ we can apply the analysis from the previous proof specialized to the case when $k = 1$. Notice that only cases **S1** and **S3** may hold for $a+\tau$ and they become: **S1)** $0 \leq e+\tau \leq 1$; **S3)** $2 \leq e+\tau \leq 3$. Clearly, the latter case holds if and only if $e = \tau = 1$. It follows further that $\mathbf{x}$ falls in one of the three cases **X1-3** described in the previous proof, specialized to $k = 1$. Notice that in case **X3** one has $x_R = 0$. Additionally, in case **X2**, relations $u_R = 0$ and $x_R = 0$ cannot hold simultaneously. This is because $u_R = 0$ implies that $e = e'$. Since **S1** must hold either for $a+\tau$, or for $a'+\tau'$, while **S3** must hold for the other, it follows that $e = e' = 1$ and $\tau \neq \tau'$. Finally, this implies that $x_R = 1$. We conclude that one of the following possible cases holds for $\mathbf{x}$:

Y1) $\mathbf{x}_1^{R-1} = \mathbf{0}_{R-1}$.

Y2) $\mathbf{x}_1^{R-1} = (\mathbf{0}_{s-1},\mathbf{1}_{R-s})$, for some $1 \leq s \leq R-1$, and $x_R = 0$ only if $u_R = 1$.

Y3) $\mathbf{x} = (\mathbf{0}_{s-1},\mathbf{1}_{t-s},\mathbf{0}_{R-t+1})$, where $1 \leq s < t \leq R-1$ and s is the position of the rightmost 0 in $\boldsymbol{\beta}(a)$ and t is the position of the rightmost 0 in $\boldsymbol{\beta}(a')$ or viceversa (the case $s = t$ is covered by **Y1).**

To complete the proof it is sufficient to show that if $\mathbf{u} \in \mathcal{W}$ then $\mathbf{y}$ is not in $\mathcal{W}'$, where $\mathbf{y} \triangleq \mathbf{u} \oplus \mathbf{x}$.

Assume that $\mathbf{u} = \mathbf{w}_q$ for some $2 \leq q \leq R$. If **Y1** holds then $\mathbf{y}_1^{R-1} = (\mathbf{0}_{q-2},\mathbf{1}_{R-q+1})$. Thus, it is clear that $\mathbf{y}$ is not in $\mathcal{W}'$. If **Y2** holds then $\mathbf{y}_{R-1}^R = 01$, which again implies that $\mathbf{y}$ is not in $\mathcal{W}'$. Assume now that **Y3** holds. Then relation (37) implies that $\mathbf{u}_{R-1}^R = 10$, which further leads to the conclusion that $t = R-1$. This implies that $\mathbf{x} = (\mathbf{0}_{s-1},\mathbf{1}_{R-1-s},00)$, where $1 \leq s < R-1$. It follows that $\mathbf{y}_{R-1}^R = 10$, which implies that $\mathbf{y} \neq \mathbf{w}'_\ell$ for $3 \leq \ell \leq R$ or $\ell = 1$. It remains to show that $\mathbf{y} \neq \mathbf{w}'_2$. Note that if $\mathbf{y} = \mathbf{w}'_2$, then the first bit of $\mathbf{y}$ must be 1, which happens only if a) $s = 1$ and $q > 2$ or if b) $s > 1$ and $q = 2$. In both cases $\mathbf{y} = (\mathbf{1}_\ell,\mathbf{0}_{R-\ell-2},10)$, where $\ell = |q-1-s|$. Then clearly, $\mathbf{y} \neq \mathbf{w}'_2$.

It remains to consider $\mathbf{u} = \mathbf{w}_1$. Because $u_R = 1$ it follows that one of a, a' is even and the other is odd, therefore **Y3** may not hold. Then, when R is odd, one has $\mathbf{y}_1^{R-1} = ((\mathbf{10})_\ell,(\mathbf{01})_{\frac{R-1}{2}-\ell})$ for $0 \leq \ell \leq \frac{R-1}{2}$, or $\mathbf{y}_1^{R-1} = ((\mathbf{10})_\ell,1,(\mathbf{10})_{\frac{R-1}{2}-\ell-1},1)$ for $0 \leq \ell \leq \frac{R-1}{2}-1$. When R is even, one has $\mathbf{y}_1^{R-1} = ((\mathbf{10})_\ell,0,(\mathbf{10})_{\frac{R}{2}-1-\ell})$ or $\mathbf{y}_1^{R-1} = ((\mathbf{10})_\ell,1,(\mathbf{10})_{\frac{R}{2}-1-\ell})$ for $0 \leq \ell \leq \frac{R}{2}-1$. It is clear now that $\mathbf{y}$ is not in $\mathcal{W}'$. ■

References

[1] V. A. Vaishampayan, "Design of multiple description scalar quantizers," *IEEE Trans. Inform. Theory*, vol. 39, no. 3, pp. 821–834, May 1993.

[2] V. A. Vaishampayan and J. Domaszewicz, "Design of entropy-constrained multiple description scalar quantizers", *IEEE Trans. Inform. Theory*, vol. 40, no. 1, pp. 245–250, Jan. 1994.

[3] Y. Wang, M. T. Orchard, V. A. Vaishampayan, and A. R. Reibman, "Multiple description coding using pairwise correlating transforms," *IEEE Trans. Image Proc.*, vol. 10, no. 3, pp. 351–366, Mar. 2001.

[4] V. A. Vaishampayan, N. Sloane, and S. Servetto, "Multiple description vector quantization with lattice codebooks: design and analysis," *IEEE Trans. Inform. Theory*, vol. 47, no. 5, pp. 1718–1734, Jul. 2001.

[5] V. K. Goyal, "Multiple description coding: compression meets the network", *IEEE Signal Proc. Magazine*, vol. 18, no.5, pp. 74–93, Sept. 2001.

[6] V. K Goyal and J. Kovacevic, "Generalized multiple description coding with correlating transforms," *IEEE Trans. Inform. Theory*, vol. 47, no. 6, pp. 2199–2224, Sept. 2001.

[7] V. K Goyal, J. A. Kelner, and J. Kovacevic, "Multiple description vector quantization with a coarse lattice", *IEEE Trans. Inform. Theory*, vol. 48, no. 3, pp. 781–7884, Mar. 2002.

[8] R. Venkataramani, G. Kramer, and V. K. Goyal, "Multiple description coding with many channels", *IEEE Trans. Inform. Theory*, vol. 49, no. 9, pp. 2106–2114, Sept. 2003.

[9] S. S. Pradhan, R. Puri, and K. Ramchandran, "n-channel symmetric multiple descriptions - Part I: (n,k) source-channel erasure codes", *IEEE Trans. Inform. Theory*, vol. 50, no. 1, pp. 47–61, Jan. 2004.

[10] C. Tian and S. S. Hemami, "Universal multiple description scalar quantization: analysis and design," *IEEE Trans. Inform. Theory*, vol. 50, no. 9, pp. 2089–2102, Sept. 2004.

[11] S. Dumitrescu and X. Wu, "Optimal two-description scalar quantizer design," *Algorithmica*, vol. 41, no. 4, pp. 269–287, Feb. 2005.

[12] S. S. Pradhan, R. Puri, and K. Ramchandran, "n-channel symmetric multiple descriptions - Part II: An achievable rate-distortion region", *IEEE Trans. Inform. Theory*, vol. 51, no. 4, pp. 1377–1392, Apr. 2005.

[13] J. Ostergaard, J. Jensen, and R. Heusdens, "n-channel entropy-constrained multiple description lattice vector quantization", *IEEE Trans. Inform. Theory*, vol. 52, no. 5, pp. 1956–1973, May 2006.

[14] X. Huang and X. Wu, "Optimal index assignment for multiple description lattice vector quantization," *Proc. IEEE DCC*, pp. 272–281, Snowbird, UT, Mar. 2006.

[15] J. Chen, C. Tian, T. Berger, and S. S. Hemami, "Multiple description quantization via Gram-Schmidt orthogonalization," *IEEE Trans. Inform. Theory*, vol. 52, no. 12, pp. 5197–5217, Dec. 2006.

[16] S. Dumitrescu and X. Wu, "Lagrangian optimization of two-description scalar quantizers," *IEEE Trans. Inform. Theory*, vol. 53, no. 11, pp. 3990–4012, Nov. 2007.

[17] D. Muresan and M. Effros, "Quantization as histogram segmentation: optimal scalar quantizer design in network systems," *IEEE Trans. Inform. Theory*, vol. 54, no. 1, pp. 344–366, Jan. 2008.

[18] A. Gyorgy, T. Linder, and G. Lugosi, "Tracking the best quantizer," *IEEE Trans. Inform. Theory*, vol. 54, no. 4 pp. 16049–1625, Apr. 2008.

[19] M. Liu and C. Zhu, "M-description lattice vector quantization: index assignment and analysis", *IEEE Trans. Signal Proc.*, vol. 57, no. 6, pp. 2258–2274, Jun. 2009.

[20] S. Dumitrescu and X. Wu, "On properties of locally optimal multiple description scalar quantizers with convex cells," *IEEE Trans. Inform. Theory*, vol. 55, no. 12, pp. 5591–5606, Dec. 2009.

[21] J. Ostergaard, R. Heusdens, and J. Jensen, "n-Channel asymmetric entropy-constrained multiple-description lattice vector quantization", *IEEE Trans. Inform. Theory*, vol. 56, no. 12, pp. 6354–6375, Dec. 2010.

[22] J. Chen, Y. Zhang, and S. Dumitrescu, "Gaussian multiple description coding with low-density generator matrix codes," *IEEE Trans. Commun.*, vol. 60, no. 3, pp. 676–687, Mar. 2012.

[23] G. Zhang, J. Klejsa, W. B. Kleijn, "Optimal index assignment for multiple description scalar quantization with translated lattice codebooks," *IEEE Trans. Signal Proc.*, vol. 60, no. 8, pp. 4444–4451, Aug. 2012.

[24] Y. Zhang, S. Dumitrescu, J. Chen, and Z. Sun, " LDGM-based multiple description coding for finite alphabet sources," *IEEE Trans. Commun.*, vol. 60, no. 12, pp. 3671–3682, Dec. 2012.

[25] Z. Gao and S. Dumitrescu, "Flexible symmetric multiple description lattice scalar quantizer with $L \geq 3$ descriptions", *IEEE Trans. Commun.*, vol. 62, no. 12, pp. 4281–4292, Dec. 2014.

[26] M. Srinivasan, "Iterative decoding of multiple descriptions," *Proc. IEEE DCC*, pp. 463–472, Snowbird, UT, Mar. 1999.

[27] J. Barros, J. Hagenauer, and N. Gortz, "Turbo cross decoding of multiple descriptions," *IEEE ICC*, 2002.

[28] T. Guionnet and C. Guillemot, "Soft decoding of multiple descriptions," *IEEE ICME*, Lausanne, 2002.

[29] J. N. Laneman, E. Martinian, G. W. Wornell, and J. G. Apostolopoulos, "Source-channel diversity for parallel channels", *IEEE Trans. Inform. Theory*, vol. 51, no. 10, pp. 3518–3539, Oct. 2005.

[30] I. Bahceci, Y. Altunbasak, and T. M. Duman, "A turbo-coded multiple-description system for multiple antennas," *IEEE Trans. Communic.*, vol. 54, pp. 187–191, Feb. 2006.

[31] X. Wu, X. Wang, and Z. Wang, "Resource-scalable joint source-channel MAP and MMSE estimation of multiple descriptions", *IEEE Trans. Signal Proc.*, vol. 57, no. 1, pp. 279–288, Jan. 2009.

[32] K.-Y. Yen, C.-F. Wu, and W.-W. Chang "Iterative decoding of convolutionally encoded multiple descriptions," *Proc. IEEE DCC*, pp. 470, Snowbird, UT, Mar. 2009.

[33] R. Ma and F. Labeau, "Error-resilient multiple description coding," *IEEE Trans. Signal Proc.*, vol. 56, no. 8, pp. 3996–4007, Aug. 2008.

[34] Y. Zhou and W.-Y. Chan, "Multiple description quantizer design for multiple-antenna systems with MAP detection," *IEEE Trans. Commun.*, vol. 58, no. 1, pp. 136–145, Jan. 2010.

[35] L. Harper, "Optimal numberings and isoperimetric problems on graphs," *J. Combinatorial Theory*, vol. 1, no. 3, pp. 385–393, 1966.

[36] X. Wang, X. Wu, and S. Dumitrescu, "On explicit formulas for bandwidth and antibandwidth of hypercubes," *Discrete Applied Mathematics*, vol. 157, no. 8, pp. 1947–1952, Apr. 2009.

[37] Y. Wan and S. Dumitrescu, "Index Assignment Capable of Detecting One Bit Errors for Multiple Description Scalar Quantizers", *Canadian Workshop Inform. Theory*, Toronto, June 2013.

[38] S. Dumitrescu and Y. Wan, "Bit-error Resilient Index Assignment for Multiple Description Scalar Quantizers", *IEEE ISIT*, Istanbul, Turkey, Jul. 2013.

[39] T. Y. Berger-Wolf and E. M. Reingold, "Index Assignment for Multichannel Communication Under Failure," *IEEE Trans. Inform. Theory*, vol. 48, no. 10, pp. 2656–2668, Oct. 2002.

[40] J. Balogh and J. A. Csirik, "Index assignment for two-channel quantization," *IEEE Trans. Inform. Theory*, vol. 50, no. 11, pp. 2737–2751, Nov. 2004.

[41] R. E. Totty and G. C. Clark, "Reconstruction error in waveform transmission," *IEEE Trans. Inform. Theory*, vol. 13, pp. 336–338, Apr. 1967.

[42] T. R. Crimmins, H. M. Horwitz, C. J. Palermo, and R. V. Palermo, "Minimization of mean-square error for data transmitted via group codes," *IEEE Trans. Inform. Theory*, vol. IT-15, no. 1, pp. 72–78, Jan. 1969.

[43] N. Rydbeck and C.-E. W. Sundberg, "Analysis of digital errors in nonlinear PCM systems", *IEEE Trans. Commun.*, vol. 24, no. 1, pp. 59–65, Jan. 1976.

[44] J. R. B. De Marca and N. S. Jayant, "An algorithm for assigning binary indices to the codevectors of a multi-dimensional quantizer," *Proc. IEEE ICC*, Seattle, WA, Jun. 1987.

[45] K. Zeger and A. Gersho, "Pseudo-Gray coding", *IEEE Trans. Commun.*, vol. 38, no. 12, pp. 2147–2158, Dec. 1990.

[46] D.-M. Chiang and L. C. Potter, "Minmax non-redundant channel coding for vector quantization," *Proc. IEEE ICASSP*, Minneapolis, MN, vol. V, pp. 617–620, Apr. 1993.

[47] H.-S. Wu and J. Barba, "Index allocation in vector quantization for noisy channels," *Electron. Lett.*, vol. 29, no. 15, pp. 1317–1319, Jul. 1993.

[48] S. W. McLaughlin, D. L. Neuhoff, and J. J. Ashley, "Optimal binary index assignments for a class of equiprobable scalar and vector quantizers", *IEEE Trans. Inform. Theory*, vol. 41, no. 6, pp. 2031–2037, Nov. 1995.

[49] P. Knagenhjelm and E. Agrell, "The Hadamard transform - a tool for index assignment," *IEEE Trans. Inform. Theory*, vol. 42, no. 4, pp. 1139–1151, Jul. 1996.

[50] R. Hagen and P. Hedelin, "Robust vector quantization by a linear mapping of a block code," *IEEE Trans. Inform. Theory*, vol. 45, no. 1, pp. 200–218, Jan. 1999.

[51] M. Skoglund, "On channel-constrained vector quantization and index assignment for discrete memoryless channels," *IEEE Trans. Inform. Theory*, vol. 45, no. 7, pp. 2615–2622, Nov. 1999.

[52] B. Farber and K. Zeger, "Quantizers with uniform encoders and channel optimized decoders," *IEEE Trans. Inform. Theory*, vol. 50, no. 1, pp. 62–77, Jan. 2004.

[53] B. Farber and K. Zeger, "Quantizers with uniform decoders and channel-optimized encoders," *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 640–661, Feb. 2006.

[54] J. Y-T. Leung, O. Vornberger and J. D. Witthoff "On some variants of the bandwidth minimization problem," *SIAM J. Comput.*, vol. 13, no. 3, pp. 650–667, Jul. 1984.

[55] Z. Miller and D. Pritikin, "On the separation number of a graph," *Networks* vol. 19, no. 6, pp. 651–666, Oct. 1989.

[56] A. Raspaud, H. Schroder, O. Sykora, L. Torok, and I. Vrto, "Antibandwidth and cyclic antibandwidth of meshes and hypercubes," *Discrete Mathematics*, vol. 309, no. 11, pp. 3541–3552, 2009.

[57] L. Torok and I. Vrto, "Antibandwidth of d-Dimensional Meshes," *Intern. Workshop Combinatorial Algorithms*, *LNCS*, vol. 5874, pp. 471–477, Jun. 28-Jul. 2, 2009.

[58] R. Bansal and K. Srivastava, "Memetic algorithm for the antibandwidth maximization problem," *J. Heuristics*, vol. 17, no. 1, pp 39–60, Feb. 2011.

[59] M. Lozano, A. Duarte, F. Gortazar, and R. Marti, "Variable neighborhood search with ejection chains for the antibandwidth problem", *J. Heuristics*, vol. 18, no. 6, pp 919–938, Dec. 2012.

[60] S. Dobrev, R. Kralovic, D. Pardubska, L. Torok, and I. Vrto, "Antibandwidth and cyclic antibandwidth of Hamming graphs," *Discrete Applied Mathematics*, vol. 161, no. 10–11, pp. 1402–1408, Jul. 2013.

[61] R. W. Hamming, "Error detection and error correcting codes," *The Bell Syst. Tech. J.*, vol. 29, no. 2, pp. 147–160, Apr. 1950.

[62] S. Lin and D. J. Costello Jr., *Error Control Coding - Fundamentals and Applications*, 2nd Ed., Pearson Prentice Hall, Pearson Education, Upper Saddle River, NJ, 2004.

PLACE PHOTO HERE

Sorina Dumitrescu (M’05-SM’13) received the B.Sc. and Ph.D. degrees in mathematics from the University of Bucharest, Romania, in 1990 and 1997, respectively. From 2000 to 2002 she was a Postdoctoral Fellow in the Department of Computer Science at the University of Western Ontario, London, Canada. Since 2002 she has been with the Department of Electrical and Computer Engineering at McMaster University, Hamilton, Canada, where she held Postdoctoral, Research Associate, and Assistant Professor positions, and where she is currently an Associate Professor. Her current research interests include multimedia coding and communications, network-aware data compression, multiple description codes, joint source-channel coding, signal quantization. Her earlier research interests were in formal languages and automata theory. Dr. Dumitrescu held an NSERC University Faculty Award during 2007-2012.

PLACE PHOTO HERE

Yinghan Wan received her B.Eng. degree in automation from University of Science and Technology of China, Hefei, China, in 2009 and her M.A.Sc. degree in electrical and computer engineering from McMaster University, Hamilton, Canada, in 2011. She is currently working at Genband Canada on software design. Her research interests include signal quantization.