Index Assignment for Robust Multiple Description Scalar Quantizer

INDEX ASSIGNMENT FOR ROBUST MULTIPLE DESCRIPTION SCALAR QUANTIZER

BY

YINGHAN WAN, B.Eng.

A THESIS

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

Master of Applied Science (2011)	McMaster University
(Electrical & Computer Engineering)	Hamilton, Ontario, Canada

TITLE:	Index Assignment for Robust Multiple Description Scalar Quantizer
AUTHOR:	Yinghan Wan B.Eng., (Automation) Univ. of Sci. and Tech. of China, Hefei, China
SUPERVISOR:	Dr. Sorina Dumitrescu
NUMBER OF PAGES:	x, 72

Abstract

Conventional multiple description coding (MDC) is a source coding technique which provides resilience against packet loss. On the other hand, the correlation introduced between descriptions can be used to combat bit errors as well. While the latter feature of MDC has been attested and exploited in prior work, only few attempts have been made to design MDC with higher bit error resilience ability.

This thesis makes some progress in the latter direction by addressing the problem of robust (i.e., bit error resilient) index assignment (IA) design for two description scalar quantizers. Our approach is to start from an initial IA which is known to be good for the conventional two description problem, and then apply permutations to indices in each description to increase a minimum Hamming distance-like performance measure.

The criterion of increasing the minimum Hamming distance between valid index pairs (d_{min}), has been considered in prior work, however an efficient IA construction was presented only for the case of $d_{min} = 2$ and low redundancy.

The contribution of this thesis is the following. For the scenario when one description is known to be error free, a new measure for IA robustness is proposed, which is termed minimum side Hamming distance ($d_{side,min}$). This quantity is defined as the minimum Hamming distance between valid indices of one description for fixed index

of the other description. It is further shown that the problem of robust permutations design under the new criterion is closely connected to the anti-bandwidth problem in a certain graph derived from a hypercube. Leveraging this connection, permutations achieving $d_{side,min} = 2$ are proposed for all redundancy levels. Furthermore, for general values of $d_{side,min}$, a simple construction of permutations achieving $d_{side,min}$ is presented, based on channel codes of appropriate block length and rate, and with minimum distance $d_{side,min} + 1$, respectively, $d_{side,min}$, for two types of initial IA (diagonal, respectively, square-based). The application of this result to achieve IA with $d_{side,min} = 3$ is further discussed for a wide range of redundancy levels.

Finally, for the scenario when both descriptions may carry bit errors, simple constructions of permutations achieving $d_{min} = 3$ are proposed for the high redundancy case.

Acknowledgement

I would like to take this opportunity to thank the many people who made this thesis possible.

First and foremost, I would like to express my gratitude to my supervisor Dr. Sorina Dumitrescu, for her patience, enthusiasm and dedication to research. Throughout my thesis writing, she provided encouragement, kind advice, and a lot of good ideas. Especially her rigorous attitude to research and work will be always remembered.

I would also like to thank Dr. J. K. Zhang and Dr. S. Shirani for serving my committee. My gratitude also goes to my fellow colleagues Jiayi, Ying, Heng, Jie, Xiao, Yang and Li, for their providing a fun environment where to learn and grow in these two years. Besides, I wish to give my special thank to Xinxin, Ling and Sijiao for helping me get through difficult times and their emotional support. I am also deeply indebted to my boyfriend Dong, for his love and patience with me.

Last and the most important, I wish to thank my parents for giving birth to me and supporting me spiritually throughout my life. Also thank my aunt and uncle, who liked calling me during the two years to tell me their love and caring for me.

Notations and Abbreviations

$d_H(\cdot,\cdot)$	Hamming distance between two bit sequences
$H_w(\cdot)$	Hamming weight of a bit sequence
$\alpha(\cdot)$	Index assignment mapping function
$\boldsymbol{\pi}$	Index permutation pair
$\oplus$	Modulo 2 addition
$\lceil\cdot\rceil$	Ceiling function
$\lfloor\cdot\rfloor$	Floor function
$ab(\cdot)$	Anti-bandwidth function
$\leq_H$	Hales order
A^T	Transpose of matrix A
MD	Multiple Description
MDC	Multiple Description Coding
MDSQ	Multiple Description Scalar Quantizer
$\mathbf{K-DSQ}$	K-Descriptions MDSQ
MDVQ	Multiple Description Vector Quantizer
IA	Index Assignment
MAP	Maximum A Posteriori
FEC	Forward Error Correction
BER	Bit Error Rate

Contents

List of Figures

Chapter 1

Introduction

Multiple Description Coding (MDC) is to compress a source into multiple descriptions such that each individual description should lead to the signal reconstruction to some acceptable quality, while more descriptions, when collaboratively decoded, improve the quality of the reconstruction, which is commensurate with the number of decoded descriptions. As one technique for MDC, Multiple Description Scalar Quantizer (MDSQ) is to generate descriptions by using scalar quantizers. A mechanism which governs the trade-off between the quality of central and side descriptions in MDSQ is the Index Assignment (IA). In conventional MDC redundancy is introduced in order to combat only description loss. On the other hand, this redundancy can also be exploited to correct other channel impairments, like bit-errors.

While the problem of MDC design to combat only description loss has been extensively studied, the design of robust MDC, which additionally has increased bit error detection/correction ability has received less attention. This thesis makes an advance in the latter direction by addressing the problem of index assignment design for bit error resilient MDSQ.

In this thesis, we will use the attribute “conventional”for any MDC technique which is concerned only with combating description loss, and the attribute “robust ”for any MDC techniques which is additionally concerned with alleviating the effect of bit errors.

This chapter briefly describes applications and four approaches for conventional MDC. Then we present the motivation for robust MDSQ and discuss previous work on this topic. Finally we present the contribution and the organization of this thesis.

1.1 Applications for Conventional MDC

In a packet-switched network, all the transmitted data is allocated into suitable sized packets, and the packets are sent over a shared network that routes each packet independently from all others and assigns transmission resources as needed. Packets may be lost in data networks for a variety of reasons. For instance, packets may be dropped due to network congestion or a packet may be considered lost if it does not reach the destination by some prescribed time limit due to routing delay. On the other hand, if the transmission also incurs bit errors and the channel code used to protect the data packet fails to correct all the errors, the packet is declared lost.

One method to recover from losses is packet retransmission. This technique is well suited to many applications, but the delay in receiving a retransmitted packet may be not affordable when it is much longer than the interval between received packets, especially for real-time services. Therefore, MDC arose as an attractive alternative to combat packet loss, where each packet can be considered as a separate description.

The use of MDC ensures that when some packets are lost, the received packets can still be decoded to a certain reconstruction quality which degrades gracefully as the number of losses increases.

Another application of MDC is in diversity communication systems. In such a system, there are several channels from a single sender to the destination. Each description is transmitted over a separate channel. If the same data were transmitted over all the channels, there would not be any advantage in the case when all channels transmit successfully comparing to the case only one description is successfully received. With MDC, the separate descriptions are different but correlated. Thus MDC guarantees that when only one channel works, a coarse version of the source signal could be reconstructed, while when more channels work the descriptions received refine each other to yield a reconstruction of higher fidelity.

1.2 Techniques for Conventional MDC

For independent identical distributed sources, some popular MDC techniques are (Goyal (2001)): (1) progressive coding with unequal erasure protection; (2) MD quantization; (3) MD correlated transforms; (4) MD coding with frames.

Progressive Coding with Unequal Erasure Protection is constructed based on a progressive or successively refinable code stream, which is divided into consecutive segments of non-decreasing length. Each segment is protected by a Reed-Solomon channel code, all codewords having the same length. Thus, early segments, which also have higher importance, are assigned higher protection levels comparing to segments appearing later in the code stream. The descriptions are formed across channel codewords, in other words, each description contains one symbol from each codeword. Thus, when only k out of the total N descriptions are received, the decoder can recover all segments of at most $N-k$ symbols, based on the channel erasure protection.

These segments form a prefix of the code stream, which is successfully decoded. The rate-distortion optimal design of such MDC systems were addressed by Mohr *et al.* (1999); Puri and Ramchandran (1999); Stankovic *et al.* (2004); Dumitrescu *et al.* (2004).

In **MD Quantization**, MD quantizers are used to produce separate descriptions. The MDSQ optimal design was first addressed by Vaishampayan (1993) for the case with two balanced descriptions, i.e., where the two descriptions have the same rate and achieve the same distortion when decoded individually. He also introduced the notion of IA and proposed IA's with good performance. Optimal MD scalar or vector quantizers design was also addressed by Vaishampayan and Domaszewicz (1994); Vaishampayan and Batllo (1998); Vaishampayan *et al.* (2001); Fleming *et al.* (2004); Dumitrescu and Wu (2005, 2007, 2009); Gyorgy *et al.* (2008); Muresan and Effros (2008). The problem of index assignment was further discussed by Balogh and Csirik (2004) for two descriptions balanced scalar quantizers, by Berger-Wolf and Reingold (2002) for general MDSQ and by Yahampath (1999), and by Gortz and Leelapornchai (2003) for MD vector quantizers.

In **MD Correlated Transforms**, the basic idea is to introduce correlation between a pair of random variables with a linear transform (Goyal and Kovacevic (1998, 2001); Wang *et al.* (1997)). The statistic dependencies between transform coefficients can be useful since the estimation of transform coefficients that are in a lost description is improved.

In **MD Coding with Frames**, the main idea is to left-multiply a source vector with a rectangular matrix in order to produce transform coefficients (Goyal *et al.* (1998, 1999)), which is further quantized and partitioned into descriptions. The source vector can be reconstructed by a least-square problem to find a possible vector so as to minimize the square of the difference between quantized expansion coefficients and the product.

1.3 Motivation and Literature Review

In MDC, separate descriptions should be different enough in order to refine each other. On the other hand, if the separate descriptions are completely uncorrelated, then the best reconstruction at the central decoder (where all descriptions are available) is achieved, but the reconstruction at some side decoders (where only one description is available) is very poor. Thus, to obtain good enough quality at all side decoders, the descriptions have to be correlated. Therefore, redundancy is introduced in the system. In conventional MDC this redundancy is used only for packet loss resilience. But it is natural to think that this redundancy could be useful for robustness against bit errors.

The idea of exploiting the redundancy left in the coded sequence in order to increase bit error resilience was widely used in the literature during the past two decades, in the case of single description coding (Sayood and Borkenhagen (1991); Phamdo and Farvardin (1994); Park and Miller (1999); Bauer and Hagenauer (2000); Hagenauer and Gortz (2003); Wu *et al.* (2004); Wang and Wu (2010)).

Then naturally this idea was extended to MDC. Decoding schemes were developed by Barros *et al.* (2002); Guionnet and Guillemot (2002); Zhou and Chan (2004); Bahceci *et al.* (2006); Wu *et al.* (2009), which exploit the dependency between descriptions to combat bit errors at the central decoder, with or without the aid of channel coding. Barros *et al.* (2002); Guionnet and Guillemot (2002) considered a balanced MDSQ with two descriptions, while Wu *et al.* (2009) used a general MDQ with general number of descriptions as the MD code.

Very recently, the problem of robust MDC design was also addressed by Ma and Labeau (2008) and Zhou and Chan (2010). An MDVQ over space-time orthogonal block coded slow Rayleigh fading channels is considered by Zhou and Chan (2010). The problem of optimal robust 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 posterior probability detector as an equivalent discrete memoryless channel, and it is solved by using a heuristic algorithm, while the robust IA design by Zhou and Chan (2010) depends on the channel parameters, Ma and Labeau (2008) take a different approach. They consider a balanced MDSQ with two descriptions (2-DSQ) and identify as a measure for IA robustness the minimum Hamming distance $h_{\min}$ of the set $\mathcal{C}$ of valid pairs of two descriptions indices (i, j). Their procedure for robust IA design is split into two steps: (1) given a value for $h_{\min}$, find a set $\mathcal{C} \subseteq \{0, \cdots, 2^R - 1\} \times \{0, \cdots, 2^R - 1\}$ of minimum Hamming distance $h_{\min}$ where R is the number of bits used to represent an index of each description; (2) assign indices l of the central partition to pairs (i, j) in $\mathcal{C}$ so as to achieve balanced descriptions with as low side distortions as possible. 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. They propose an efficient solution for the first problem only for the case of 1-bit redundancy, i.e., $2R - \log_2 |\mathcal{C}| = 1$, which achieves $h_{\min} = 2$. Furthermore, Ma and Labeau (2009) introduced a robust multiple description scalar quantizer system with three state Markov chain as testbed, they also combined forward error correction (FEC) techniques and MDC to combat packet loss and bit errors.

1.4 Contribution and Organization of Thesis

This thesis addresses the problem of designing robust IA for balanced 2-DSQ, which does not depend on the channel parameters and thus is not prone to performance degradation in case of mismatch. Our approach is to start from an initial IA which is good for the conventional 2-DSQ problem, and apply a permutation to indices in each description, i.e., permutation π_t for description t, $t = 1, 2$.

Applying the permutation pair $\boldsymbol{\pi} = (\pi_1, \pi_2)$ does not change the performance of the 2-DSQ in the conventional sense, i.e., when the descriptions are not corrupted by bit errors, but has the potential of increasing the bit error resilience at the central decoder. The following scenarios are considered at the central decoder: (1) when one description is correct and the decoder knows which one, referred to as *Scenario II*; (2) when both descriptions may carry errors, referred to as *Scenario III.* The performance measure for error resilience in *Scenario III* considered in this work is the same as in the work of Ma and Labeau (2008), and we refer to it as the minimum Hamming distance of the IA. For *Scenario II*, on the other hand, we identify a better suited performance criterion, termed the minimum side Hamming distance of the IA, which is defined as the minimum Hamming distance of the set of valid indices in one description when the index in the other description is fixed.

As initial IA, we use the diagonal IA (where the valid pairs occupy consecutive diagonals in the IA matrix) and the square IA (where the valid positions in the IA matrix occupy equal size squares along the main diagonal). Then we propose efficient constructions of robust permutation pairs $\boldsymbol{\pi}$ as follows.

For the case of diagonal initial IA, we establish the equivalence between the problem of robust permutation design achieving minimum side Hamming distance $d_{\min}$ and the anti-bandwidth problem in a certain graph derived from a hypercube. For $d_{\min} = 2$, the corresponding graph is exactly the hypercube, for which the anti-bandwidth problem has been solved. Therefore, we apply these results to construct permutations achieving minimum side Hamming distance equal to 2. Furthermore, for both diagonal and square initial IA we present a general and simple construction of permutations with minimum side Hamming distance $d_{\min}$, based on linear channel codes of minimum Hamming distance $d_{\min} + 1$, respectively $d_{\min}$, of appropriate rate and block length. We further discuss the application of this result with shortened Hamming codes of minimum Hamming distance 3 or 4, as well as simple channel codes corresponding to the case of high redundancy. Finally, for *Scenario III*, we propose permutations achieving minimum Hamming distance 3 for both the diagonal and the square IA in the high redundancy case.

The rest of the thesis is structured as follows. Chapter 2 presents the notations and background for MDSQ, discusses decoding schemes for different scenarios, and introduces the notion of robust permutation along with the performance measures of minimum side Hamming distance and minimum Hamming distance, for *Scenario II* and *Scenario III* respectively. Moreover, the problem of robust IA is converted to the problem of finding the permutation $\boldsymbol{\pi}$ which increases the robustness. In Chapter 3, we show the connection between robust IA problem for *Scenario II* and the anti-bandwidth problem. Chapter 4 and Chapter 5 discuss robust IA design based on linear permutations for *Scenario II* and *Scenario III* respectively. We present some experimental results in Chapter 6. Chapter 7 concludes this thesis and outlines some future works.

Chapter 2

Index Assignment for MDSQ

In this chapter, we introduce the background and notations for MDSQ, as well as the index assignment, which influences the quantization performance. Since in conventional MDSQ design, all channels are assumed as on/off or so-called erasure channels, descriptions are received correctly or totally lost. However in practice channels are usually noisy, which leads to bit errors during transmission so that MDSQ for noisy channel is also discussed in this chapter. Finally, we use the last section to discuss IA for robust MDSQ problem formulation.

2.1 Conventional MDSQ

2.1.1 Background and Notations

A K-DSQ consists of a set of K encoders and $2^K - 1$ decoders. Descriptions generated by K encoders are transmitted through K independent channels. The decoders corresponding to the sets which have only one description are called side decoders, and the decoders with set of more than one descriptions are so-called joint decoders. Among the $2^K - 1$ decoders of K-DSQ, there are K side decoders and $2^K - 1 - K$ joint decoders. Note that there is a special case of joint decoders called central decoder, which corresponds to all the K descriptions. This thesis considers only balanced 2-DSQ, where there are 2 descriptions with the same rate and the same quality of the reconstruction.

Let X be a continuous random variable with probability density function $f_X(x)$. The encoder of a 2-DSQ operates as follows. The source sample x is encoded first by a single description scalar quantizer q to an index $k \in \{0, \cdots, N-1\}$. q is called the central quantizer and N is the total number of central cells. Every index k is further mapped to an index pair $\alpha(k) = (\alpha_1(k), \alpha_2(k))$. This mapping is called the index assignment. Each component of the pair is an index of one side description. In all, the encoding scheme for conventional 2-DSQ can be described as $x \xrightarrow{q} q(x) \xrightarrow{\alpha} (\alpha_1(q(x)), \alpha_2(q(x)))$. For convenience, the notation (i, j) is used instead of $(\alpha_1(q(x)), \alpha_2(q(x)))$. Index i takes values in the set $\{0, \cdots, 2^{R_1} - 1\}$, and j takes values in $\{0, \cdots, 2^{R_2} - 1\}$, where R_t is the rate of description t, $t = 1, 2$. Let $b(i), b(j)$ denote the R_1-bit and R_2-bit binary representations of i, j respectively. Furthermore, $b(i)$ is sent over channel 1, and $b(j)$ over channel 2.

In a conventional 2-DSQ system, it is assumed that each channel either transmits correctly or it breaks down. Therefore, at the receiver end there are 2 decoders, one for each non-empty subset of received descriptions: the central decoder g_0 and the side decoders g_1, g_2. The side decoders are two one-to-one mappings: $g_1 : \{0, \cdots, 2^{R_1} - 1\} \rightarrow \hat{\mathcal{X}}_1$, $g_2 : \{0, \cdots, 2^{R_2} - 1\} \rightarrow \hat{\mathcal{X}}_2$, where $\hat{\mathcal{X}}_1, \hat{\mathcal{X}}_2 \in \mathbb{R}$ are two sets of reproduction values called codebooks. Denote $\hat{x}_i^1 = g_1(i)$ and $\hat{x}_j^2 = g_2(j)$. The central decoder is a

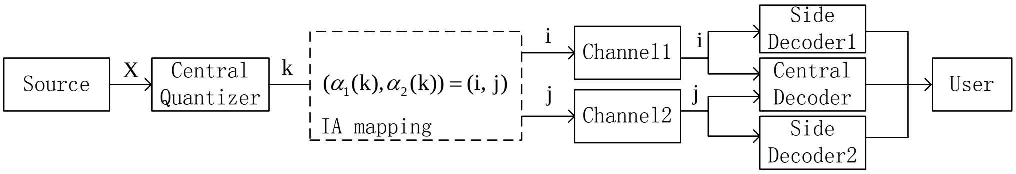


Figure 2.1: 2-DSQ system with two channels and three receivers

mapping $g_0 : \mathcal{C} \to \hat{\mathcal{X}}_0$, where $\mathcal{C}$ is the set of *transmitted* or *valid* pairs of indices (i, j), in other words $\mathcal{C} = \{(\alpha_1(q(x)), \alpha_2(q(x)))|x \in \mathbb{R}\}$. Denote $\hat{x}^0_{(i,j)} = g_0(i, j)$ for $(i, j) \in \mathcal{C}$. Fig. 2.1 illustrates a 2-DSQ system. The performance of the 2-DSQ is measured by the expected distortion between the source and its reconstruction at the receiver side. In this work, we use the squared error i.e., $d(x, y) = (x - y)^2$, as a distortion measure. Without bit-errors, the expected distortion at each decoder is given by

$$D_1 = E(d(X, \hat{\mathcal{X}}_1)) = \sum_i \int_{A^1_i} (x - \hat{x}^1_i)^2 f_X(x)\, dx,$$

where $A^1_i = \{x|\alpha_1(g(x)) = i\}$. The sets A^1_i form the partition of side description 1.

$$D_2 = E(d(X, \hat{\mathcal{X}}_2)) = \sum_j \int_{A^2_j} (x - \hat{x}^2_j)^2 f_X(x)\, dx,$$

where $A^2_j = \{x|\alpha_2(g(x)) = j\}$. The sets A^2_j form the partition of side description 2.

$$D_0 = E(d(X, \hat{\mathcal{X}}_0)) = \sum_{k=0}^{N-1} \int_{A^0_k} (x - \hat{x}^0_{(\alpha_1(k),\alpha_2(k))})^2 f_X(x)\, dx,$$

where $A^0_k = \{x|q(x) = k\}$. The sets A^0_k form the central partition.

Note that given fixed central quantizers, cells and IA, the optimized reconstruction values, i.e., which minimize the three distortions, must satisfy the centroid condition:

$$\hat{x}_i^1 = \frac{\int_{A_i^1} x f_X(x)\, dx}{\int_{A_i^1} f_X(x)\, dx},$$

$$\hat{x}_j^2 = \frac{\int_{A_j^2} x f_X(x)\, dx}{\int_{A_j^2} f_X(x)\, dx},$$

$$\hat{x}^0_{(\alpha_1(k),\alpha_2(k))} = \frac{\int_{A_k^0} x f_X(x)\, dx}{\int_{A_k^0} f_X(x)\, dx}.$$

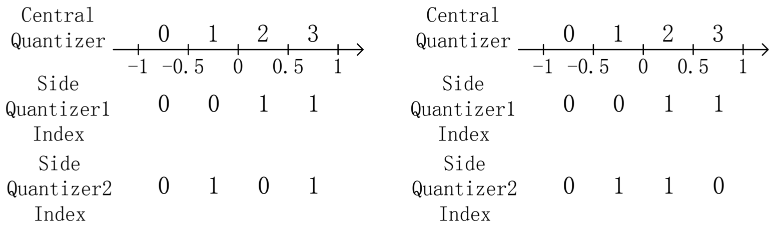
We will assume throughout this work that the centroid condition is satisfied.

In a 2-DSQ system, there is a trade-off between the quality of the side and central descriptions. In other words, the three descriptions cannot be very good simultaneously. A mechanism which controls this trade-off is the IA, which will be discussed in more details shortly. In this work, we are concerned with balanced 2-DSQ, i.e., where $R_1 = R_2 = R$ and $D_1 \approx D_2$.

The optimization problem for a 2-DSQ can be formulated as the problem of minimizing the central distortion subject to constraints on the side distortions. Vaishampayan (1993) attacked this problem by converting it to the minimization of a weighted sum of the three distortions:

$$\min \quad D_0 + \lambda(D_1 + D_2). \tag{2.1}$$

Moreover, Vaishampayan showed that the optimal 2-DSQ which minimizes (2.1) must have convex cells i.e., intervals in the central partition; in other words, the central



(a) (b)

Figure 2.2: Two MDSQ's for a uniform distributed source

partition is characterized by a set of thresholds $t_0 < t_1 < \cdots < t_{N-2}$ and

$$\begin{aligned} A_0^0 &= (-\infty,\ t_0], \\ A_k^0 &= (t_{k-1},\ t_k], \qquad 1 \le k \le N-2, \\ A_{N-1}^0 &= (t_{N-2},\ \infty). \end{aligned}$$

2.1.2 Index Assignment

The design of a 2-DSQ usually consists of two components: the selection of IA and optimizing the structure of the quantizer for the chosen IA. As a measure of the IA quality, Vaishampayan (1993) introduced the notion of *spread* which is defined follows. For every $t = 1, 2$ and $k \in \{0, 1, \cdots, 2^R - 1\}$, $s^{(t)}(k)$ is the number of central cells spanned by index k of the t-th description, i.e., $s^{(t)}(k) = \max \alpha_t^{-1}(k) - \min \alpha_t^{-1}(k) + 1$. The reconstruction accuracy of index k at side decoder t increases, as the spread $s^{(t)}(k)$ decreases. To maintain a uniform reconstruction quality, it is desirable to have the spreads of all indices equal in each description. Moreover, for balanced descriptions, we would like the spreads of all side cells to be approximately equal in both descriptions. Define the maximum spread $S_{\max}(\alpha)$ as the maximum value of the spreads of all cells in both descriptions. Then given a fixed number of cells N in the central partition, a good IA α should minimize $S_{\max}(\alpha)$.

Table 2.1: Reconstruction Level

	$\hat{x}^0_{(0,0)}$	$\hat{x}^0_{(0,1)}$	$\hat{x}^0_{(1,0)}$	$\hat{x}^0_{(1,1)}$	$\hat{x}^1_0$	$\hat{x}^1_1$	$\hat{x}^2_0$	$\hat{x}^2_1$
(a)	-0.75	-0.25	0.25	0.75	-0.5	0.5	-0.25	0.25
(b)	-0.75	-0.25	0.75	0.25	-0.5	0.5	0	0

To illustrate the significance of the spread, let us consider a simple example inspired by Vaishampayan (1993). Fig. 2.2 shows two 2-DSQ's with the same central quantizer, but different index assignments. Notice that for IA in case (a) the maximum spread is $S_{\max}(\alpha) = 3$, while in case (b) it is $S_{\max}(\alpha) = 4$. Let X here be a uniformly distributed random variable over the interval (-1,1), and the rate of each side description be $R = 1$ bit per source sample. The corresponding reconstruction codebook for each 2-DSQ is given in Table 2.1, while the expected distortions of central or side quantizers are given in Table 2.2. Notice that for both index assignments, the central distortion and the side distortion of description 1 are the same, while the side distortion of the second description is smaller in case (a) than in case (b). Thus this example validates the general principle that better performance is ensured with smaller $S_{\max}(\alpha)$.

Popularly, the IA is represented as a matrix (see examples in Fig. 2.3). Each cell inside the matrix containing a number, represents a central partition cell, and the number is its index. Then the indices of the cell's row and column stand for the indices for description 1 and description 2, respectively.

Vaishampayan proposed good IA's for balanced 2-DSQ, where central quantization cells correspond to matrix cells on $m = 2k + 1$ diagonals, the main diagonal and k closest diagonals above and below main diagonal, respectively. Vaishampayan's

Table 2.2: Central and Side Distortion

	D_0	D_1	D_2
(a)	0.0208	0.0833	0.2708
(b)	0.0208	0.0833	0.3333

	0	1	2	3	4	5	6	7
0	0	1						
1	2	3	4					
2		5	6	7				
3			8	9	10			
4				11	12	13		
5					14	15	16	
6						17	18	19
7							20	21

(a) 3-diagonal IA

	0	1	2	3	4	5	6	7
0	0	1						
1	2	3						
2			4	5				
3			6	7				
4					8	9		
5					10	11		
6							12	13
7							14	15

(b) 2-by-2 square IA

Figure 2.3: Two examples of IA's.

designs can be extended to the case of $m = 2k$ by using the main diagonal and k, respectively $k-1$, closest diagonals above, respectively below the main diagonal. We will refer to such IA's using m diagonals, as m-diagonal IA's. Fig. 2.3(a) illustrates a 3-diagonal IA. Another type of IA that we will consider in this work is the 2^l-by-2^l square IA, where the occupied matrix cells form 2^{R-l} 2^l-by-2^l squares placed along the main diagonal. Moreover, the indices of central quantizer cells corresponding to the i-th square , $1 \leq i \leq 2^{R-l}$, are in the range $\{2^{2l}(i-1), 2^{2l}(i-1)+1, \cdots, 2^{2l}i-1\}$. Fig. 2.3(b) shows a 2-by-2 square IA.

For the m-diagonal IA or the 2^l-by-2^l square IA, the parameter m, respectively l, controls the trade-off between the quality of the central and side descriptions. This is because as m, respectively l, increases, the number of central cells increases, thus making the central distortion smaller, while the spread increases as well making the side distortions larger.

2.2 Decoder for 2-DSQ with Bit-Errors

In the case when one or both descriptions channels may incur bit errors, the redundancy between descriptions can be used to improve the reconstructions at the central decoder. From this point of view, it is relevant to distinguish between three scenarios at the central decoder: (*Scenario I*) neither description has errors; (*Scenario II*) only one channel incurs bit errors and the central decoder knows which one; (*Scenario III*) both descriptions may incur bit errors. *Scenario I* corresponds to the conventional 2-DSQ and was discussed previously. Next we address the central decoder in *Scenario II* and *Scenario III*.

2.2.1 Central Decoder in *Scenario II*

As shown in Lemma 1 of the work by Zeger and Gersho (1990), the average distortion D of a noisy channel quantizer with the mean-square distortion function, that satisfies the centroid condition, can be written as the sum of two parts, where one is produced by the quantizer and denoted by D_S, and the other is due to the channel noise and denoted by D_C. Therefore, D can be expressed as $D = D_S + D_C$. Thus, the expected distortion at the central decoder can be decomposed as

$$D_0 = D_{0,s} + D_{0,c},$$

where $D_{0,s}$ is given by

$$D_{0,s} = \sum_{k=0}^{N-1} \int_{A_k^0} (x - \hat{x}^0_{(\alpha_1(k),\alpha_2(k))})^2 f_X(x)dx$$

and $D_{0,c}$ will be discussed shortly. Without loss of generality, we assume description 1 is received correctly while description 2 might be with errors. Let (i, j') denote the index pair received at the central decoder, and let the central decoder be $\bar{g}_0 : \{0, \cdots, 2^R - 1\} \times \{0, \cdots, 2^R - 1\} \to \mathbb{R}$. For each $i \in \{0, \cdots, 2^R - 1\}$, denote $\mathcal{J}(i) = \{j|(i, j) \in \mathcal{C}\}$ as the set of j's such that (i, j)'s are valid index pairs. Then $D_{0,c}$ is

$$\begin{aligned} D_{0,c} &= E\{d(\hat{x}^0_{(\boldsymbol{i},\boldsymbol{j})}, \bar{g}_0(\boldsymbol{i}, \boldsymbol{j'}))\} \\ &= \sum_{i=0}^{2^R-1} \sum_{j\in\mathcal{J}(i)} \sum_{j'=0}^{2^R-1} P(i, j)P_e(j'|j)d(\hat{x}^0_{(i,j)}, \bar{g}_0(i, j')) \end{aligned} \tag{2.2}$$

where $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{j'}$ are random variables, representing the input to channel 1, input to channel 2 and output of channel 2, respectively. Moreover, $P(i, j)$ is the probability that (i, j) is output by the 2-DSQ, $P_e(j'|j)$ is the conditional probability that j' is received conditioned on j being sent. Assume that all bit errors are independent and identically distributed (i.i.d) with bit error rate (BER) ϵ, then $P_e(j'|j)$ can be expressed as $P_e(j'|j) = \epsilon^d(1-\epsilon)^{R-d}$ in which d is the Hamming distance between $b(j)$ and $b(j')$.

Notice that $D_{0,s}$ is not affected by the choice of the decoders. Thus, the optimal decoder which minimizes the expected distortion D_0 should minimize $D_{0,c}$. Notice that

$$D_{0,c} = \sum_{i=0}^{2^R-1} \sum_{j'=0}^{2^R-1} D_c(i, j'),$$

where

$$D_c(i,j') = \sum_{j \in \mathcal{J}(i)} P(i,j) P_e(j'|j) d(\hat{x}^0_{(i,j)}, \bar{g}_0(i,j')).$$

Thus, for given pair i, j', the optimal decoder $\bar{g}_{0,opt}(i,j')$ has to minimize $D_c(i,j')$, in other words

$$\bar{g}_{0,opt}(i,j') = \arg\min_{y \in \mathbb{R}} \sum_{j \in \mathcal{J}(i)} P(i,j) P_e(j'|j) d(\hat{x}^0_{(i,j)}, y),$$

and further

$$\bar{g}_{0,opt}(i,j') = \begin{cases} \hat{x}^0_{(i,j')} & \text{for } j' \in \mathcal{J}(i) \\ \frac{\sum\limits_{j \in \mathcal{J}(i)} P(i,j) P_e(j'|j) g_0(i,j)}{\sum\limits_{j \in \mathcal{J}(i)} P(i,j) P_e(j'|j)} & \text{for } j' \notin \mathcal{J}(i) \end{cases}. \tag{2.3}$$

Notice that the optimal decoder needs knowledge of the BER ϵ. Moreover, this decoder can lead to performance degradation in case of mismatch, i.e., if ϵ is not estimated correctly. Ma and Labeau (2008) presented another decoder, called suboptimal decoder:

$$\bar{g}_{0,subopt}(i,j') = \begin{cases} \hat{x}^0_{(i,j')} & \text{, for } j' \in \mathcal{J}(i) \\ \frac{\sum\limits_{j \in \mathcal{J}(i)} \hat{x}^0_{(i,j)} P(i,j)}{\sum\limits_{j \in \mathcal{J}(i)} P(i,j)} & \text{, for } j' \notin \mathcal{J}(i) \end{cases} \tag{2.4}$$

which does not depend on the channel statistics. Notice that $\bar{g}_{0,subopt}(i,j') = \hat{x}^1_i$ when $j' \notin \mathcal{J}(i)$, thus this decoder actually discards the description with errors.

We propose a new decoder which does not need knowledge of ϵ, and is more accurate than the sub-optimal decoder. For each $i, j' \in \{0, \cdots, 2^R - 1\}$ and $d \in \{0, \cdots, R\}$, define $\mathcal{H}_{i,d}(j') = \{j \in \mathcal{J}(i) | d_H(b(j), b(j')) = d\}$ as the set of j's in $\mathcal{J}(i)$ with Hamming distance to j' equal to d. Notice that $d_H(u,v)$ denotes the Hamming distance between u and v, i.e., the number of bits in which the two bit sequences differ. Let $d_{\min}(i,j') = \min_{\mathcal{H}_{i,d}(j')\neq\phi} d$, be the minimum Hamming distance between the received pairs and sent valid pairs. Furthermore, let $\mathcal{H}_i(j') = \mathcal{H}_{i,d_{\min}(i,j')}(j')$, i.e., the set of valid j's for given i closest in Hamming distance to j'. The proposed decoder is motivated by the assumption that ϵ is very small ($\epsilon \ll 0.5$) and thus $\epsilon^d(1-\epsilon)^{n-d} \gg \epsilon^{d+s}(1-\epsilon)^{n-d-s}$ for $1 \leq s \leq R-d$. Then the dominant terms in $D_c(i,j')$ are those corresponding to $d = d_{\min}(i,j')$:

$$\begin{aligned} D_c(i,j') &= \sum_{j\in\mathcal{J}(i)} P(i,j)P_e(j'|j)d(\hat{x}^0_{(i,j)}, \bar{g}_0(i,j')) \\ &= \sum_{d\geq d_{\min}(i,j')} \sum_{\substack{j\in\mathcal{J}(i) \\ d_H(b(j),b(j'))=d}} P(i,j)d(\hat{x}^0_{(i,j)}, \bar{g}_0(i,j'))\epsilon^d(1-\epsilon)^{R-d} \\ &\approx \sum_{\substack{j\in\mathcal{J}(i) \\ d_H(b(j),b(j'))=d_{\min}(i,j')}} P(i,j)d(\hat{x}^0_{(i,j)}, \bar{g}_0(i,j'))\epsilon^{d_{\min}(i,j')}(1-\epsilon)^{R-d_{\min}(i,j')}. \end{aligned} \tag{2.5}$$

The value of $\bar{g}_0(i,j')$ which minimizes the latter expression is given by

$$\bar{g}_0(i,j') = \frac{\sum_{j\in\mathcal{H}_i(j')} \hat{x}^0_{(i,j)}P(i,j)}{\sum_{j\in\mathcal{H}_i(j')} P(i,j)} \tag{2.6}$$

Note that the proposed decoder is in essence a minimum Hamming distance decoder. Precisely, the decoder looks in $\mathcal{J}(i)$ for the closest index j to j'. If only one such j is found, then $\bar{g}_0(i,j') = g_0(i,j)$, which actually means that j' is decoded as j. If more such j's are found then $\bar{g}_0(i,j')$ is computed as in (2.6).

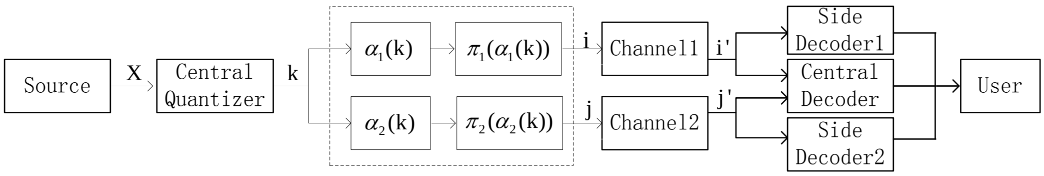


Figure 2.4: MDSQ system with bit-errors for two channels and three receivers

2.2.2 Decoder for ***Scenario III***

For *Scenario III*, we use a similar minimum Hamming distance decoder. Let (i', j') be the received index pair. Define $\mathcal{H}_d(i', j') = \{(i, j) \in \mathcal{C} | d_H(b(i, j), b(i', j')) = d\}$ and $d'_{\min}(i', j') = \min_{\mathcal{H}_d(i',j') \neq \phi} d$. Denote $\mathcal{H}(i', j') = \mathcal{H}_{d'_{\min}(i',j')}(i', j')$, then the proposed decoder is

$$\tilde{g}_0(i', j') = \frac{\sum_{(i,j)\in\mathcal{H}(i',j')} \hat{x}^0_{(i,j)} P(i, j)}{\sum_{(i,j)\in\mathcal{H}(i',j')} P(i, j)}. \tag{2.7}$$

2.3 Problem Formulation

In conventional 2-DSQ design, the 2-DSQ performance is optimized for on/off channels. Our goal is to increase the robustness when the channels may addtionally introduce bit-errors, while maintaining the performance in the bit error-free case. For this, we start from a 2-DSQ optimized in the traditional sense and further apply an index permutation to the index output by each description. The index permutations do not affect the performance in the bit error-free case, but have the potential to increase the bit error resilience. Let $\pi_1 : \{0, 1, \cdots, 2^R - 1\} \to \{0, 1\}^R$ be the permutation applied to indices in description 1 and $\pi_2 : \{0, 1, \cdots, 2^R - 1\} \to \{0, 1\}^R$ be the permutation for description 2. We will use the notation $\boldsymbol{\pi}$ for the permutation pair (π_1, π_2). Thus, a new IA is generated, denoted by $\boldsymbol{\pi} \circ \alpha$, where $\boldsymbol{\pi} \circ \alpha(k) = (\pi_1(\alpha_1(k)), \pi_2(\alpha_2(k)))$ for any $k \in \{0, 1, \cdots, N\}$. Fig. 2.4 shows the 2-DSQ system with index assignment

$\boldsymbol{\pi} \circ \alpha$.

As discussed in the previous section, the distortion of a quantizer satisfying the centroid condition over a noisy channel can be decomposed into two components: quantization distortion D_S and channel distortion D_C. Applying the permutation π_t to indices of description $t\,(t = 1, 2)$ result in a relabeling of the cells of side quantizer t. This relabeling does not change the quantizers distortion $D_{t,s}$, but can change the channel distortion $D_{t,c}$. In this work, we are interested only in reducing $D_{0,c}$, i.e., the channel distortion at the central decoder, and disregard $D_{1,c}$ and $D_{2,c}$.

From the previous sections, it is clear that the error corrective capability of the IA is related to the minimum Hamming distance. In order to formulate in more detail a criteria for robust IA for *Scenario III* and *Scenario II*, let us first introduce some notations.

For any set $\mathcal{A} \subseteq \{0, 1\}^n$ for some n, define the minimum Hamming distance of $\mathcal{A}$ as

$$d_{\min}(\mathcal{A}) = \min_{\substack{u,v \in \mathcal{A} \\ u \neq v}} d_H(u, v) \tag{2.8}$$

For *Scenario III*, the decoder looks in the whole codebook $\mathcal{C}$ for the closest codeword to the received pair. Therefore, the minimum Hamming distance of the whole set $\mathcal{C}$ can be considered a measure for the robustness of the IA. The higher $d_{\min}(\mathcal{C})$, the larger the number of detected and corrected errors. Specifically, the decoder can detect all bit error patterns containing at most $d_{\min}(\mathcal{C}) - 1$ bit errors and can correct all error pattern with at most $\left\lceil \frac{d_{\min}(\mathcal{C})}{2} \right\rceil - 1$, where $\lceil \cdot \rceil$ denotes the ceiling function.

Define the minimum distance of an IA α, denoted by $d_{\min}(\alpha)$, as the minimum

Hamming distance of the set of valid codewords pairs

$$d_{\min}(\alpha) = d_{\min}(\mathcal{C}). \tag{2.9}$$

Then the goal in designing a robust IA $\boldsymbol{\pi} \circ \alpha$ for *Scenario III* is to increase $d_{\min}(\boldsymbol{\pi} \circ \alpha)$.

Notice that the criterion of large $d_{\min}(\alpha)$ for the robustness of IA α, was discussed first by Ma and Labeau (2008). However, they applied this criterion only for *Scenario II*. We introduce next a better suited criterion for robustness of IA in *Scenario II*, which is also easier to satisfy. Let us assume that description 1 is correct, while description 2 may carry bit errors.

Define the minimum side 2 Hamming distance of IA α, denoted by $d_{2,\min}(\alpha)$, as follows

$$d_{2,\min}(\alpha) = \min_{i \in \{0,1,\cdots,2^R-1\}} d_{\min}(\mathcal{J}(i)). \tag{2.10}$$

It can be easily seen that the decoder for *Scenario II* is able to detect all error pattern with at most $d_{2,\min}(\alpha) - 1$ bit errors, and is able to correct any error pattern with at most $\left\lceil \frac{d_{2,\min}(\alpha)}{2} \right\rceil - 1$ bit errors. Therefore, our criteria for robust IA in *Scenario II* is to increase $d_{2,\min}(\boldsymbol{\pi} \circ \alpha)$. It is clear from the definition that for any IA α, we have

$$d_{2,\min}(\alpha) \geq d_{\min}(\alpha), \tag{2.11}$$

and thus

$$d_{2,\min}(\boldsymbol{\pi} \circ \alpha) \geq d_{\min}(\boldsymbol{\pi} \circ \alpha), \tag{2.12}$$

for any $\boldsymbol{\pi}$. However, in many cases it is possible to design permutation pairs $\boldsymbol{\pi}$ such

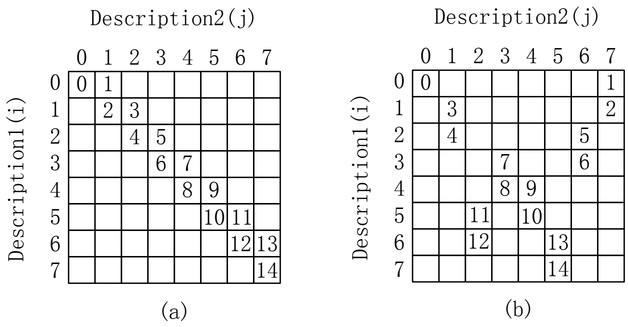


Figure 2.5: Robust IA example. (a) Original IA. (b) IA after permutation

that

$$d_{2,\min}(\boldsymbol{\pi} \circ \alpha) > d_{\min}(\boldsymbol{\pi} \circ \alpha), \tag{2.13}$$

Such an example is presented next. Fig. 2.5 gives two IA matrices with size $2^3 \times$

Table 2.3: Permutation $\boldsymbol{\pi}$

i/j	0	1	2	3	4	5	6	7
$\pi_1(i)$	000	001	010	011	100	101	110	111
$\pi_2(j)$	000	111	001	110	011	100	010	101

2^3, i.e., $R = 3$, and 15 occupied cells, i.e., $N = 15$. Fig. 2.5(b) is the IA matrix after permutation pair $\boldsymbol{\pi}$ applied to the initial IA α shown in Fig. 2.5(a), where the permutations of each side description are given in Table 2.3.

It is easy to see that $d_{2,\min}(\alpha) = d_{\min}(\alpha) = 1$, while $d_{2,\min}(\alpha) = 2 > d_{\min}(\alpha) = 1$. If the pair $(i, j) = (0, 0)$, i.e., [000,000], is transmitted

$$[000, 000] \overset{trans}{\longrightarrow} [000, 001] \overset{decode}{\longrightarrow} [000, 001] \tag{2.14}$$

$$[000, 000] \overset{trans}{\longrightarrow} [000, 001] \overset{decode}{\longrightarrow} [000, 000] \tag{2.15}$$

When there is one-bit error through transmission, [000,000] may become [000,001].

According to the decoding strategy described before, two different decoding results are obtained. As (2.15) shows, in the case of IA with robust permutation, the decoder is able to correct the error, while in the case of (2.14), it is not.

In this work, we will design robust permutation pairs for the case when the initial IA is diagonal or square based. Notice that when α is an m-diagonal IA, $d_{2,\min}(\alpha)$ can be rewritten as

$$d_{2,\min}(\boldsymbol{\pi} \circ \alpha) = \min_{\substack{j_1,j_2 \in \{0,1,\cdots,2^R-1\} \\ 1 \le |j_1-j_2| \le m-1}} d_H(\pi_2(j_1), \pi_2(j_2)), \tag{2.16}$$

while when α is a 2^l-by-2^l square IA, we have

$$d_{2,\min}(\boldsymbol{\pi} \circ \alpha) = \min_{\substack{j_1,j_2 \in \{0,1,\cdots,2^R-1\}, j_1 \neq j_2 \\ b(j_1) \oplus b(j_2) = \underbrace{0\cdots0}_{R-l} u_1 \cdots u_l \\ u_1,\cdots,u_l \in \{0,1\}}} d_H(\pi_2(j_1), \pi_2(j_2)) \tag{2.17}$$

Finally, note that for the case when description 2 is correct and description 1 may contain bit errors, the minimum side 1 Hamming distance, $d_{1,\min}(\alpha)$, can be defined similarly to (2.10):

$$d_{1,\min}(\alpha) = \min_{j \in \{0,1,\cdots,2^R-1\}} d_{\min}(\{i : (i,j) \in \mathcal{C}\}).$$

Clearly, $d_{t,\min}(\boldsymbol{\pi} \circ \alpha)$ depends only on π_t, $t = 1, 2$. Furthermore, due to the symmetry of the diagonal and square based IA, a permutation π_2 achieving $d_{2,\min}(\boldsymbol{\pi} \circ \alpha) = d$ can be used for description 1 as well (i.e., $\pi_1 = \pi_2$) to achieve $d_{1,\min}(\boldsymbol{\pi} \circ \alpha) = d$. Therefore, without loss of generality, we will address only $d_{2,\min}(\boldsymbol{\pi} \circ \alpha)$ in the rest of this thesis.

Chapter 3

Connection with Anti-bandwidth Problem

In this chapter, we show that the problem of designing a robust permutation for the m-diagonal IA for *Scenario II* is closely related to the anti-bandwidth problem in a certain graph derived from a hypercube. Firstly, we introduce the anti-bandwidth problem. Then we describe the relation between anti-bandwidth and robust permutation. Finally we draw a conclusion about the existence of a permutation achieving $d_{2,\min}$ equal to 2, and present its construction based on known results about the anti-bandwidth of a hypercube.

3.1 Definition of Anti-bandwidth

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be an undirected graph, where $\mathcal{V}$ denotes the set of vertices or nodes, and $\mathcal{E}$ denotes the set of edges which connect pairs of vertices. Let the size of $\mathcal{V}$ be n, i.e., $n = |\mathcal{V}|$. A labeling or numbering μ of the vertices of $\mathcal{G}$ is an one to one mapping

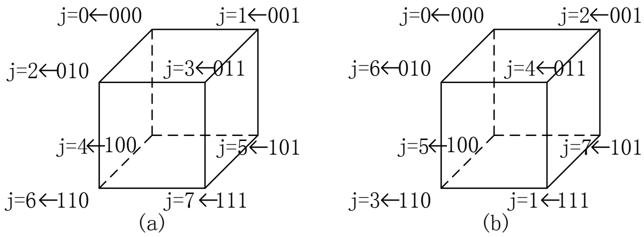


Figure 3.1: Different numbering of a hypercube

μ: $\mathcal{V} \to \{0, 1, \cdots, n-1\}$. In other words, the labeling μ assigns to each vertex v in $\mathcal{V}$ a unique number $\mu(v)$ in $\{0, 1, \cdots, n-1\}$. Note that any numbering μ defines a total order $\leq_\mu$ of the vertices in $\mathcal{V}$ as follows: for any $u, v \in \mathcal{V}$, $u \leq_\mu v$ if and only if $\mu(u) \leq \mu(v)$.

Definition 3.1. *For a positive integer n, define the n-dimensional hypercube Q_n as a graph with 2^n vertices and $2^{n-1}n$ edges, in which every vertex corresponds to an n-bit sequence and any two vertices are connected by an edge if an only if their corresponding bit sequences differ in only 1 bit.*

Then clearly a numbering μ in an n-dimension hypercube is an injective mapping from n-bit sequences to natural values from $\{0, 1, \cdots, 2^n - 1\}$. Fig. 3.1 shows two labelings of a 3-dimensional hypercube.

The bandwidth of a numbering μ is denoted by $bw(\mu)$ and is defined as

$$bw(\mu) = \max_{(v,u) \in \mathcal{E}} |\mu(v) - \mu(u)|. \tag{3.1}$$

The bandwidth of a graph $\mathcal{G}$ is the minimum bandwidth among all the numberings μ

$$bw(\mathcal{G}) = \min_{\mu} \max_{(v,u) \in \mathcal{E}} |\mu(v) - \mu(u)|. \tag{3.2}$$

The term of anti-bandwidth was first introduced by Torok and Vrto (2009). The anti-bandwidth problem can be regarded as the dual of the bandwidth problem. In other words, finding the anti-bandwidth problem is equivalent to labeling vertices of a graph such that the minimum absolute difference of labels of adjacent vertices to be maximized. More precisely, the anti-bandwidth of a labeling μ, is denoted by $ab(\mu)$ and is defined as

$$ab(\mu) = \min_{(v,u)\in\mathcal{E}} |\mu(v) - \mu(u)|. \tag{3.3}$$

The anti-bandwidth of a graph $\mathcal{G}$, denoted $ab(\mathcal{G})$, is further defined as

$$ab(\mathcal{G}) = \max_{\mu} \min_{(v,u)\in\mathcal{E}} |\mu(v) - \mu(u)|. \tag{3.4}$$

It was proved that for general graphs the problem of determining if the anti-bandwidth is larger than some given value is NP-complete (Leung *et al.* (1984)). The value of the anti-bandwidth and the achieving labeling are known only for certain cases.

3.2 Connection between Robust Index Assignment and Anti-bandwidth

Definition 3.2. *For positive integers n and d with $2 \leq d \leq n$, define $Q_n(d)$ as an n-dimensional modified hypercube where all the vertices are n-bit sequences and all edges are any two pairs of vertices (u,v) such that $u \neq v$ and $d_H(u,v) \leq d-1$.*

Next we show that the problem of robust permutation design for *Scenario II* and m-diagonal initial index assignment is closely connected with the anti-bandwidth problem for a certain *modified hypercube*. Note that the anti-bandwidth of $Q_n(d)$ is given by

$$ab(Q_n(d)) = \max_{\mu} \min_{\substack{u \neq v \\ d_H(u,v) \leq d-1}} |\mu(u) - \mu(v)|. \tag{3.5}$$

Theorem 3.1. *Given an m-diagonal IA α with $m \geq 2$ and a positive integer $d \geq 2$, there is a permutation $\pi_2 : \{0, 1, \cdots, 2^R - 1\} \to \{0,1\}^R$ such that $d_{2,\min}(\boldsymbol{\pi} \circ \alpha) \geq d$ if and only if $ab(Q_R(d)) \geq m$.*

Proof. The relation $ab(Q_R(d)) \geq m$ is equivalent to the fact that there exists a numbering $\mu : \{0,1\}^R \to \{0, 1, \cdots, 2^R - 1\}$ such that

$$|\mu(v) - \mu(u)| \geq m \text{ for all } u, v \in \{0,1\}^R \text{ with } u \neq v \text{ and } d_H(u,v) \leq d-1. \tag{3.6}$$

Let $\pi_2 = \mu^{-1}$, i.e., $\pi_2 : \{0, 1, \cdots, 2^R - 1\} \to \{0,1\}^R$ such that $\mu(\pi_2(j)) = j$ for any $j \in \{0, 1, \cdots, 2^R - 1\}$ and $\pi_2(\mu(u)) = u$ for any $u \in \{0,1\}^R$.
Let $v = \pi_2(j_1)$ and $u = \pi_2(j_2)$. Then $|\mu(\pi_2(j_2)) - \mu(\pi_2(j_1))| = |j_2 - j_1|$. Statement (3.6) is equivalent to $|j_2 - j_1| \geq m$ for all $j_1, j_2 \in \{0, 1, \cdots, 2^R - 1\}$ such that $j_1 \neq j_2$ and $d_H(\pi_2(j_1), \pi_2(j_2)) \leq d - 1$, which is further equivalent to the statement: for any $j_1, j_2 \in \{0, 1, \cdots, 2^R - 1\}$ such that $j_1 \neq j_2$ and $|j_2 - j_1| \leq m - 1$ we have $d_H(\pi(j_1), \pi(j_2)) \geq d$. Finally, by (2.16), the above statement is equivalent to
$d_{2,\min}(\boldsymbol{\pi} \circ \alpha) \geq d$. □

If we knew the value of $ab(Q_R(d))$ for any d, then we could find the maximum d for which $ab(Q_R(d)) \geq m$. Unfortunately, the value of $ab(Q_R(d))$ is not known but for $d = 2$, i.e., when $Q_R(d)$ is the R-dimensional hypercube Q_R. The anti-bandwidth achieving permutation for Q_R was found by Harper (1966). In order to describe the permutation we need the following definition.

Definition 3.3. *For a hypercube Q_n with vertex set $\{0,1\}^n$, the Hales order $\leq_H$ on*

Table 3.1: Anti-bandwidth of $Q_R(d)$ with Different R

R	2	3	4	5	6
$ab(Q_R(2))$	1	2	4	9	19

$\mathcal{V}$ is defined by $u \leq_H v$, if (1) $H_w(u) < H_w(v)$, or (2) $H_w(u) = H_w(v)$ and u is greater than or equal to v in lexicographic order relative to the right to left order of the coordinates. Note that $H_w(u)$ is the Hamming weight of vertex u, in other words the number of 1's in u.

For example, if $n = 4$ then the Hales numbering of the 16 4-bit sequences is $0000 < 0001 < 0010 < 0100 < 1000 < 0011 < 0101 < 1001 < 0110 < 1010 < 1100 < 0111 < 1011 < 1101 < 1110 < 1111$.

Lemma 3.1. *(Harper (1966))The permutation achieving the anti-bandwidth of the hypercube is obtained by numbering the vertices of the hypercube with even Hamming weight first then those with odd Hamming weight, in increasing Hales order.*

The value of the hypercube anti-bandwidth was given by Wang *et al.* (2009):

$$ab(Q_R(2)) = 2^{R-1} - \sum_{k=0}^{R-2} \binom{k}{\lfloor \frac{k}{2} \rfloor}. \tag{3.7}$$

Table 3.1 gives $ab(Q_R(2))$ for different values of R. Theorem 3.1, Lemma 3.1 and relation (3.6) lead to the following result.

Proposition 3.1. *Let α be an m-diagonal IA. Then there is a permutation π_2 such that $d_{2,\min}(\boldsymbol{\pi} \circ \alpha) \geq 2$ if and only if*

$$m \leq 2^{R-1} - \sum_{k=0}^{R-2} \binom{k}{\lfloor \frac{k}{2} \rfloor}, \tag{3.8}$$

Furthermore, if the above relation holds, the permutation π_2 *can be constructed by numbering the side 2 indices with increasing Hales order, the even Hamming weight sequences first, followed by the odd Hamming weight sequences.*

Chapter 4

Robust IA Design for *Scenario II* Based on Linear Permutation

As shown in the previous chapter, the problem of finding a robust permutation for the diagonal IA in *Scenario II* is equivalent to the anti-bandwidth problem, which is NP-complete. Therefore polynomial time algorithms to solve the problem are not known. On the other hand, it is desirable to have simple constructions for the permutation π_2, thus we restrict our attention in the rest of the thesis to linear permutations, i.e., permutations which are linear mappings.

In this chapter, we propose simple constructions of robust linear permutations for *Scenario II*, for both cases of diagonal and squared initial IA.

Definition 4.1. *An one-to-one mapping* $\pi_2 : \{0, 1, \cdots, 2^R - 1\} \to \{0, 1\}^R$ *is called linear permutation if there is an* $R \times R$ *matrix* G_{π_2} *with elements in the binary field GF(2) such that*

$$\pi_2(j) = b(j) \cdot G_{\pi_2},$$

where "$\cdot$" denotes matrix multiplication and $b(j)$ denotes the R-dimensional row vector with elements in $GF(2)$, which is the binary representation of j with the first component being the most significant bit and the last component being the least significant bit.

Notice that G_{π_2} is a matrix of maximum rank. Moreover, any $R \times R$ dimensional matrix G of full rank defines a permutation. Therefore, we will refer to any $R \times R$ matrix G of full rank with elements in $GF(2)$ as a permutation matrix.

For convenience, the following notations are useful in the rest of the thesis. For any integer $t \geq 0$, I_t denotes the identity matrix with dimension $t \times t$, 1_t denotes the all 1's t-dimensional row vector and 0_t is the all 0's t-dimensional row vector. For any two row vectors u of dimension k_1 and v of dimension k_2, $[u|v]$ is the $(k_1 + k_2)$-dimensional row vector obtained by concatenating u and v. Moreover, for any matrix A, A^T denotes its transpose.

4.1 m-diagonal Initial IA

In this section, we first present a result which shows how an IA $(\boldsymbol{\pi} \circ \alpha)$ with side 2 minimum Hamming distance larger or equal than some d, can be constructed based on a certain channel code with minimum Hamming distance $d+1$ of appropriate rate and block length. Then we discuss the application of this result to several values of m.

4.1.1 Robust Linear Permutation for Diagonal IA Based on Linear Channel Codes

Theorem 4.1. *For some integer k with $1 \leq k \leq R-1$, let G_2 be the generator matrix of an (R,k) systematic linear block code, i.e., $G_2 = [I_k \ P_{k\times(R-k)}]$. Assume that the minimum Hamming distance of the channel code is $d+1$ with $d \geq 1$. Consider now the matrix G given by*

$$G = \left(\begin{array}{c:ccccccc}
 & 1 & 1 & 0 & 0 & 0 & \cdots & 0 \\
 & 0 & 1 & 1 & 0 & 0 & \cdots & 0 \\
 & 0 & 0 & 1 & 1 & 0 & \cdots & 0 \\
\mathbf{0} & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\
 & 0 & 0 & \cdots & 0 & 1 & 1 & 0 \\
 & 0 & 0 & \cdots & 0 & 0 & 1 & 1 \\
 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 \\
\hdashline
\multicolumn{8}{c}{G_2}
\end{array}\right)
\begin{array}{l}
\left.\vphantom{\begin{array}{c}1\\1\\1\\1\\1\\1\\1\end{array}}\right\} R-k \\
\left.\vphantom{G_2}\right\} k
\end{array}
\qquad \underbrace{\qquad}_{k} \ \underbrace{\qquad\qquad\qquad}_{R-k} \tag{4.1}$$

Then G has full rank and the associated permutation π_2 has the property that

$$d_H(\pi_2(a), \pi_2(a+\tau)) \geq d, \tag{4.2}$$

for any $a \in \{0, 1, \cdots, 2^R - 2\}$ and any $\tau \in \{1, 2, \cdots, \min(2^k - 1, 2^R - a - 1)\}$.

Proof. Let G_1 denote the $(R-k) \times R$ dimensional matrix formed with the first $R-k$ rows in G. Recall that $G_2 = [I_k \ P_{k\times(R-k)}]$. Notice that by switching the positions of G_1 and G_2, an upper triangle matrix is obtained with all elements on the main diagonal equal to 1, which therefore has full rank. Since elementary row operations do not change the row space of a matrix, it follows that G has full rank as well.

Now fix arbitrary $a \in \{0, \cdots, 2^R - 2\}$ and $\tau \in \{1, 2, \cdots, \min(2^k - 1, 2^R - a - 1)\}$. Let $b(a) = [u_1 \cdots u_R]$ and $b(a+\tau) = [v_1 \cdots v_R]$. In general, for a binary sequence $[x_1 \cdots x_n]$ for some integer n, we will use the notation x_m^k for the sub-sequence $[x_m \cdots x_k]$ where $1 \leq m \leq k \leq n$.

By the definition of π_2, we have

$$\pi_2(a) = b(a) \cdot G \text{ and } \pi_2(a+\tau) = b(a+\tau) \cdot G. \tag{4.3}$$

Then

$$\begin{aligned} d_H(\pi_2(a), \pi_2(a+\tau)) &= d_H(b(a) \cdot G, b(a+\tau) \cdot G) \\ &= H_w(b(a) \cdot G \oplus b(a+\tau) \cdot G) \\ &= H_w((b(a) \oplus b(a+\tau)) \cdot G), \end{aligned} \tag{4.4}$$

where $H_w(\cdot)$ denotes the Hamming weight function and $\oplus$ is the modulo 2 component wise addition.

Then it remains to prove that

$$H_w((b(a) \oplus b(a+\tau)) \cdot G) \geq d. \tag{4.5}$$

Now consider the unique non-negative integers c and e such that $a = c \times 2^k + e$, where $c \geq 0$ and $0 \leq e \leq 2^k - 1$. Then u_1^{R-k} is the $(R-k)$-bit representation of c while u_{R-k+1}^R is the k-bit representation of e. Notice that $a + \tau = c \times 2^k + e + \tau$. Next we will consider two cases.

Case 1: $e+\tau \leq 2^k-1$. Then v_1^{R-k} is the $(R-k)$-bit representation of c, thus $v_1^{R-k}=u_1^{R-k}$. On the other hand, v_{R-k+1}^R is the k-bit representation of $e+\tau$. Since $e+\tau \geq e+1$, it follows that v_{R-k+1}^R and u_{R-k+1}^R differ in at least one bit, thus $u_{R-k+1}^R \oplus v_{R-k+1}^R \neq 0_k$. Then

$$\begin{aligned} H_w((b(a)\oplus b(a+\tau))\cdot G) &= H_w([0_{R-k}|u_{R-k+1}^R \oplus v_{R-k+1}^R]\cdot G) \\ &= H_w((u_{R-k+1}^R \oplus v_{R-k+1}^R)\cdot G_2) \\ &\geq d+1, \end{aligned} \tag{4.6}$$

where the last inequality follows from the fact that the minimum Hamming distance of the linear code generated by G_2 is $d+1$. We have used the well-known fact that the minimum Hamming distance of a linear block code equals to the minimum Hamming weight of the non-zero vectors in the row space of the generator matrix.

Case 2: $e+\tau \geq 2^k$. Then $a+\tau=(c+1)2^k+(e+\tau-2^k)$, where $0 \leq e+\tau-2^k \leq e-1$. Clearly v_{R-k+1}^R is the k-bit representation of $e+\tau-2^k$, which differs in at least one bit from u_{R-k+1}^R. Then $v_{R-k+1}^R \oplus u_{R-k+1}^R \neq 0_k$, which implies that

$$H_w((u_{R-k+1}^R \oplus v_{R-k+1}^R)\cdot G_2) \geq d+1. \tag{4.7}$$

As for v_1^{R-k}, it is the $(R-k)$-bit binary representation of $c+1$. Further, let t denote the position of the rightmost 0 in u_1^{R-k}, i.e., $u_t=0$ and $u_{t+1}^{R-k}=1_{R-k-t}$. Then it is easy to see that $v_1^{t-1}=u_1^{t-1}, v_t=1$ and $v_{t+1}^{R-k}=0_{R-k-t}$, which implies that

$u_1^{R-k} \oplus v_1^{R-k} = [0_{t-1}|1_{R-k-t+1}]$. Then we have

$$H_w((u_1^{R-k} \oplus v_1^{R-1})G_1) = H_w([0_{t-1}|1|0_{R-t}]) = 1. \tag{4.8}$$

Finally,

$$(b(a) \oplus b(a+\tau)) \cdot G = (u_1^{R-k} \oplus v_1^{R-k}) \cdot G_1 \oplus (u_{R-k+1}^R \oplus v_{R-k+1}^R) \cdot G_2. \tag{4.9}$$

Moreover, since the Hamming distance is a metric function and thus satisfies the triangle inequality (Lin and Costello (2004)), i.e., $d_H(u, v) \geq |d_H(u, 0_R) - d_H(v, 0_R)|$, it follows that

$$H_w(u \oplus v) \geq |H_w(u) - H_w(v)|, \quad \forall u, v \in \{0, 1\}^R. \tag{4.10}$$

Finally combing (4.7), (4.8), (4.9) and (4.10), relation (4.5) follows. Thus, the proof of Theorem 4.1 is completed. □

The next result follows immediately from Theorem 4.1.

Corollary 4.1. *For π_2 defined in Theorem 4.1, the following inequality holds*

$$d_{2,\min}(\boldsymbol{\pi} \circ \alpha) \geq d, \tag{4.11}$$

for any m-diagonal IA satisfying $m \leq 2^k$.

Thus, Theorem 4.1 can be used to construct robust IA's based on known channel codes with large minimum Hamming distance. Next we discuss the application of Theorem 4.1 for the cases $k = 1$ and $k = 2$, using channel codes with largest minimum

Hamming distance and for more general k, using shortened (R, k) Hamming codes of distance 3 or 4.

4.1.2 Application of Theorem 4.1 for $k = 1$ and $k = 2$

For $k = 1$, i.e., $m = 2$, the $(R, 1)$ linear channel code of maximum minimum Hamming distance is the repetition code. Applying Theorem 4.1 with G_2 generator matrix of repetition code, the permutation matrix is given by

$$G = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 & \cdots & 0 \\ 0 & & \cdots & 0 & 1 & 1 & 0 & 0 \\ 0 & & \cdots & & 0 & 1 & 1 & 0 \\ 0 & & \cdots & & & 0 & 1 & 1 \\ 0 & & & \cdots & & & 0 & 1 \\ 1 & & & \cdots & & & & 1 \end{pmatrix}. \tag{4.12}$$

Since the minimum Hamming distance of the repetition code is R, we have the following result.

Proposition 4.1. *For any initial 2-diagonal IA α, permutation π_2 with the permutation matrix shown as (4.12) achieves $d_{2,\min}(\boldsymbol{\pi} \circ \alpha) = R - 1$.*

For $k = 2$, i.e., $m \in \{2, 3, 4\}$, the $(R, 2)$ linear channel code with the maximum minimum Hamming distance is generated by the following generator matrix

$$G_2 = \Bigg(\overbrace{\begin{matrix} 1 & \cdots & 1 \\ 0 & \cdots & 0 \end{matrix}}^{\lfloor \frac{R+1}{3} \rfloor} \Bigg| \overbrace{\begin{matrix} 1 & \cdots & 1 \\ 1 & \cdots & 1 \end{matrix}}^{R-2\lfloor \frac{R+1}{3} \rfloor} \Bigg| \overbrace{\begin{matrix} 0 & \cdots & 0 \\ 1 & \cdots & 1 \end{matrix}}^{\lfloor \frac{R+1}{3} \rfloor} \Bigg), \tag{4.13}$$

where $\lfloor \cdot \rfloor$ denotes the floor function. Accordingly, the minimum Hamming distance of this channel code is $\lfloor \frac{2R}{3} \rfloor$. Thus the following proposition holds.

Proposition 4.2. *For any initial m-diagonal IA α with $m \in \{2, 3, 4\}$, permutation π_2 with the permutation matrix shown as (4.13) achieves $d_{2,\min}(\boldsymbol{\pi} \circ \alpha) \geq \lfloor \frac{2R}{3} \rfloor - 1$.*

4.1.3 Application of Theorem 4.1 Using Hamming Codes

For any positive integer $t \geq 3$, there exists a Hamming code with the following parameters:

Code length:	$R = 2^t - 1$
Number of information symbols:	$k = 2^t - t - 1$
Minimum Hamming distance among codewords:	$d_{\min} = 3$

The parity check matrix of such a Hamming code can be arranged in the following form:

$$H = \begin{bmatrix} Q_{t \times (2^t - t - 1)} & I_t \end{bmatrix}.$$

It is possible to delete any s columns from the parity check matrix of a Hamming code to obtain a new parity check matrix for a shortened Hamming code with following parameters(Lin and Costello (2004)):

Code length:	$R = 2^t - s - 1$
Number of information symbols:	$k = 2^t - t - s - 1$
Minimum Hamming distance among codewords:	$d_{\min} \geq 3$

The minimum Hamming distance of a shortened Hamming code is at least the same as its original Hamming code. According to the shortened Hamming code, $t = R - k$ which is at least 3, and $s = 2^t - R - 1$ which is no less than 1. Then relation $R - k \geq \log_2(R+2)$ holds and we get the following proposition.

Proposition 4.3. *For any k and $R \geq 1$ such that $R-k \geq 3$ and $R-k \geq \log_2(R+2)$ and any m-diagonal IA α with $m \leq 2^k$, there is a linear permutation π_2 such that*

$$d_{2,\min}(\boldsymbol{\pi} \circ \alpha) \geq 2. \tag{4.14}$$

If we delete from the sub-matrix Q all the columns of even weight, we obtain a modified parity check matrix written as $[Q' \ I_k]$ where Q' consists of $2^{k-1} - k$ columns of odd weight, which means there are no three columns adding to zero. Whereas, for any column in Q' of Hamming weight 3, it is possible to find three columns in I such that their logical summation is zero. Therefore, the minimum Hamming distance of this Hamming code is exactly 4. The shortened Hamming code with this parity check matrix has the properties(Lin and Costello (2004))

Code length:	$R = 2^{t-1}$
Number of information symbols:	$k = 2^{t-1} - t$
Minimum Hamming distance among codewords:	$d_{\min} = 4$

Then we have the following result.

Proposition 4.4. *For any $t \geq 3$, $R = 2^{t-1}$ and $k = 2^{t-1} - t$ and any m-diagonal IA α with $m \leq 2^k$, there is a linear permutation π_2 with*

$$d_{2,\min}(\boldsymbol{\pi} \circ \alpha) \geq 3. \tag{4.15}$$

4.2 2^l-by-2^l Square Initial IA

4.2.1 Robust Linear Permutation for Square IA Based on Linear Block Codes

Theorem 4.2. *For some integer l with $1 \leq l \leq R-1$, let G_2 be the generator matrix of an (R, l) systematic linear block code, i.e., $G_2 = [I_l \ P_{l\times(R-l)}]$. Assume that the minimum Hamming distance of the channel code is d for $d \geq 1$. Consider now the* *matrix G given by*

$$G = \begin{array}{c} \left(\begin{array}{c:ccccccc} & 1 & 1 & 0 & 0 & 0 & \cdots & 0 \\ & 0 & 1 & 1 & 0 & 0 & \cdots & 0 \\ & 0 & 0 & 1 & 1 & 0 & \cdots & 0 \\ \mathbf{0} & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ & 0 & 0 & \cdots & 0 & 1 & 1 & 0 \\ & 0 & 0 & \cdots & 0 & 0 & 1 & 1 \\ & 0 & 0 & \cdots & 0 & 0 & 0 & 1 \\ \hdashline \multicolumn{8}{c}{G_2} \end{array}\right) \begin{array}{l} \left.\vphantom{\begin{array}{c}1\\1\\1\\1\\1\\1\\1\end{array}}\right\} R-l \\ \left.\vphantom{G_2}\right\} l \end{array} \\ \underbrace{\qquad}_{l} \quad \underbrace{\qquad\qquad\qquad\qquad\qquad}_{R-l} \end{array} . \tag{4.16}$$

Then G is full ranked and the associated permutation π_2 satisfies

$$d_H(\pi_2(a_1), \pi_2(a_2)) \geq d, \tag{4.17}$$

for all $a_1, a_2 \in \{0, 1, \cdots, 2^R - 1\}$ such that $a_1 \neq a_2$ and $b(a_1) \oplus b(a_2)$ has the $R - l$ most significant bits equal to 0.

Proof. The fact that G is full ranked follows from the proof of Theorem 4.1. Let $a_1, a_2 \in \{0, 1, \cdots, 2^R - 1\}$ such that $b(a_1) \oplus b(a_2) = [0_{R-l}|u_1 \cdots u_l]$ where $u_t \in \{0, 1\}$ with $t \in \{1, 2, \cdots, l\}$ and $[u_1 \cdots u_l] \neq 0_l$. Then

$$\begin{aligned} d_H(\pi_2(a_1), \pi_2(a_2)) &= d_H(b(a_1) \cdot G, b(a_2) \cdot G) \\ &= H_w((b(a_1) \oplus b(a_2)) \cdot G) = H_w([u_1 \cdots u_l] \cdot G_2) \geq d. \end{aligned}$$

This completes the proof of Theorem 4.2. $\square$

For the cells in a 2^l-by-2^l square, their binary representations of side 2 indices differ in the last l bits for a given side 1 index. In other words, $j_1, j_2 \in \mathcal{J}(i)$ if and only if $b(j_1) \oplus b(j_2) = [0_{R-l}|u_1 \cdots u_l]$ for some $[u_1 \cdots u_l] \in \{0,1\}^l$. Then we obtain the following corollary.

Corollary 4.2. *For π_2 given by Theorem 4.2, the following inequality holds*

$$d_{2,\min}(\boldsymbol{\pi} \circ \alpha) \geq d, \tag{4.18}$$

for any 2^l-by-2^l IA.

4.2.2 Application of Theorem 4.2 for $l = 1$, $l = 2$ and $l = R-1$

Notice that the permutation matrix from Theorem 4.2 is the same as the permutation matrix from Theorem 4.1 when $l = k$. Thus, for $l = 1$ and $l = 2$, the matrix is given by (4.12) and (4.13) respectively.

Proposition 4.5. *For any initial 2-by-2 square IA α, permutation π_2 with the permutation matrix given by (4.12) achieves $d_{2,\min}(\boldsymbol{\pi} \circ \alpha) = R$.*

Proposition 4.6. *For $R \geq 3$ and any initial 2^2-by-2^2 square IA α, permutation π_2 with G_2 in Theorem 4.2 given by (4.13) achieves $d_{2,\min}(\boldsymbol{\pi} \circ \alpha) \geq \left\lfloor \frac{2R}{3} \right\rfloor$.*

For $l = R-1$, the $(R, R-1)$ linear block code which achieves maximum minimum Hamming distance is the single parity check code . A single parity check code is with a single parity check digit and whose generator matrix is given by(Lin and Costello (2004))

$$G_2 = \begin{bmatrix} I_{R-1} & 1_{R-1}^T \end{bmatrix}. \tag{4.19}$$

Accordingly, the parity check matrix of (4.19) is $H = 1_R$. And the minimum Hamming distance of the single parity check code is $d = 2$. Then the following proposition holds.

Proposition 4.7. *For any initial 2^{R-1}-by-2^{R-1} square IA α, consider permutation π_2 with permutation matrix G_{π_2} given by*

$$G_{\pi_2} = \left(\begin{array}{c:c} 0_{R-1} & \\ \hdashline I_{R-1} & \end{array} 1_R^T \right). \tag{4.20}$$

Then the following inequality holds.

$$d_{2,\min}(\boldsymbol{\pi} \circ \alpha) \geq 2. \tag{4.21}$$

4.2.3 Application of Theorem 4.2 with Hamming Code

Using a shortened Hamming codes with following parameters:

Code length:	$R = 2^t - s - 1$
Number of information symbols:	$l = 2^t - t - s - 1$
Minimum Hamming distance among codewords:	$d_{\min} \geq 3$

for any $1 \leq s \leq 2^t - 1$ and $t \geq 3$, we get the following proposition.

Proposition 4.8. *For any $l, R \geq 1$ such that $R - l \geq 3$ and $R - l \geq \log_2(R+1)$ and any 2^l-by-2^l square IA α, there is a linear permutation π_2 such that*

$$d_{2,\min}(\boldsymbol{\pi} \circ \alpha) \geq 3. \tag{4.22}$$

We properly design the parity check matrix of the shortened Hamming code shown in Section 4.1.3 with the following parameters:

Code length:	$R = 2^{t-1}$
Number of information symbols:	$l = 2^{t-1} - t$
Minimum Hamming distance among codewords:	$d = 4$

The next proposition immediately follows.

Proposition 4.9. *For any $t \geq 3$, $R = 2^{t-1}$ and $l = 2^{t-1} - t$ and any 2^l-by-2^l IA α, there is a linear permutation π_2 with*

$$d_{2,\min}(\boldsymbol{\pi} \circ \alpha) \geq 4. \tag{4.23}$$

Chapter 5

Robust IA Design for *Scenario III* Based on Linear Permutations

The index assignment schemes for robust 2-DSQ with either description being error-prone are presented in this chapter. Let $i \in \{0, 1, \cdots, 2^R - 1\}$ and $j \in \{0, 1, \cdots, 2^R - 1\}$ denote the indices of description 1 and description 2, respectively. In the following discussion, we propose the construction of robust linear permutations achieving minimum Hamming distance 3 for 2-diagonal initial IA and 2-by-2 square initial IA. The general design procedure is still under study.

5.1 2-by-2 Square Initial IA

Consider first general 2^l-by-2^l square initial IA α. Let the binary representation of indices i, respectively j, be $b(i) = [i_1 i_2 \cdots i_R]$ respectively $b(j) = [j_1 j_2 \cdots j_R]$. For all the cells in the same square, the codewords $(b(i), b(j))$ differ only in the bits $[i_{R-l+1} \cdots i_R]$ and $[j_{R-l+1} \cdots j_R]$. For example, the cells in the first square in the matrix shown in Fig. 2.3(b) which corresponds to $R = 3$ and $l = 1$, are $[00|0, 00|0]$, $[00|0, 00|1]$, $[00|1, 00|0]$ and $[00|1, 00|1]$, thus the first two bits are the same for all the cells. In general, a cell in a 2^l-by-2^l square IA can be represented by only $R + l$ bits instead of $2R$ bits, where the first $R - l$ bits are the common bits for all cells in the same square, while the remaining $2l$ bits are for telling the two descriptions apart. Therefore, the initial codeword of one cell can be written as $[u|v_1|v_2]$, where $u = [i_1 i_2 \cdots i_{R-l}] = [j_1 j_2 \cdots j_{R-l}]$, $v_1 = [i_{R-l+1} \cdots i_R]$ and $v_2 = [j_{R-l+1} \cdots j_R]$. Consider now applying a pair $\boldsymbol{\pi} = (\pi_1, \pi_2)$ of linear permutations to the initial IA, with

$$G_{\pi_1} = \begin{pmatrix} A_1 \\ \hdashline B_1 \end{pmatrix} \begin{matrix} \}\, R-l \\ \}\, l \end{matrix} \quad , \quad G_{\pi_2} = \begin{pmatrix} A_2 \\ \hdashline B_2 \end{pmatrix} \begin{matrix} \}\, R-l \\ \}\, l \end{matrix} \quad . \tag{5.1}$$

Then $[\pi_1(i)|\pi_2(j)] = [uA_1 \oplus v_1 B_1 | uA_2 \oplus v_2 B_2] = [u|v_1|v_2] \cdot G_{\boldsymbol{\pi}}$ where $G_{\boldsymbol{\pi}}$ is the $(R + l) \times (2R)$-dimensional matrix

$$G_{\boldsymbol{\pi}} = \begin{array}{c} \left(\begin{array}{c:c} A_1 & A_2 \\ \hdashline B_1 & 0_R \\ \hdashline 0_R & B_2 \end{array}\right) \\ \underbrace{\hphantom{A_1}}_{R} \; \underbrace{\hphantom{A_2}}_{R} \end{array} \begin{matrix} \}\, R-l \\ \}\, l \\ \}\, l \end{matrix} \quad . \tag{5.2}$$

Moreover, the set of all codeword pairs $[\pi_1(i)|\pi_2(j)]$ coincides with the row space of $G_{\boldsymbol{\pi}}$. Therefore, $d_{\min}(\boldsymbol{\pi} \circ \alpha)$ coincides with the minimum Hamming weight of all non-zero vectors in the row space of $G_{\boldsymbol{\pi}}$.

Proposition 5.1. *Let $R \geq 5$, α be a 2-by-2 square IA and matrix G given by*

$$G = \begin{array}{c} \left(\begin{array}{c:c} A_1 & A_2 \\ \hdashline B_1 & 0_R \\ \hdashline 0_R & B_2 \end{array}\right) \begin{array}{l} \}\, R-1 \\ \}\, 1 \\ \}\, 1 \end{array} \\ \quad\underbrace{}_{R}\ \underbrace{}_{R} \end{array}, \tag{5.3}$$

where $B_1 = B_2 = [1|0_{R-4}|101]$, $A_1 = [I_{R-1}\ 0_{R-1}^T]$ and A_2 is the $(R-1) \times R$ matrix defined as

$$A_2(k_1, k_2) = \begin{cases} 1, & \text{if } k_1 = k_2 \text{ or } k_1 = k_2 - 1 \\ 0, & \text{otherwise} \end{cases}.$$

Then $G_1 = \begin{pmatrix} A_1 \\ B_1 \end{pmatrix}$ and $G_2 = \begin{pmatrix} A_2 \\ B_2 \end{pmatrix}$ are full rank and the associated permutation pair $\boldsymbol{\pi}$ has the property that

$$d_{\min}(\boldsymbol{\pi} \circ \alpha) \geq 3. \tag{5.4}$$

In order to prove the proposition, we need the following result from Lin and Costello (2004).

Lemma 5.1. *Let C be a linear code with parity check matrix H. The minimum weight (or the minimum distance) of C is equal to the smallest number of columns of H that sum to 0.*

Then we give the proof of Proposition 5.1 as following.

Proof. Matrix G_1 is lower triangle with all elements on the diagonal equal to 1. Thus, clearly it has full rank. For G_2, apply the following row operations. First add the second last and the third last rows to the last row. Then switch the positions of A_2

and B_2. The obtained matrix is upper triangle with all 1's on the main diagonal. Thus, G_2 is full ranked as well.

Now we proceed to proving relation (5.4) It is known that elementary row operations, i.e., interchanging any tow rows or adding one row to another, on a matrix, do not change its row space. Then add the last row to the first row and then add the first row and the last fourth row to the last second row. The obtained G' is in systematic form, i.e., $G' = [I_{R+1}\ P]$, with P given by

$$P = \left(\begin{array}{c:c} P_1 & \\ \hdashline P_2 & P_4 \\ \hdashline P_3 & \end{array}\right),$$

where $P_1 = [1|0_{R-5}|10]$, $P_4 = [1|0_{R-3}|111]^T$, $P_3 = \left(\begin{smallmatrix} 1 & 0_{R-4} & 1 \\ 0_{R-4} & 1 & 0 \end{smallmatrix}\right)$, and P_2 is an $(R-2) \times (R-2)$ matrix defined as

$$P_2(k_1, k_2) = \begin{cases} 1, \text{ if } k_1 = k_2 \text{ or } k_1 = k_2 - 1 \\ 0, \text{ otherwise} \end{cases}.$$

In view of Lemma 5.1, we need only to prove that the smallest number of columns of the parity check matrix H of G' that sum to 0 is at least 3.

It is clear that $H = [P^T\ I_{R-1}]$ derived from the systematic form of G'. Since all columns of H are different, the conclusion follows (only two identical vectors can add up to 0). Therefore, the minimum Hamming distance of $\mathcal{C}$ after applying permutation pair $\boldsymbol{\pi}$ is at least 3, i.e., $d_{\min}(\pi \circ \alpha) \geq 3$. □

Proposition 5.2. *Let matrix G be given by*

$$G = \begin{array}{c} \left(\begin{array}{c:c} A_1' & A_2' \\ \hdashline B_1' & 0_7 \\ \hdashline 0_7 & B_2' \end{array}\right) \begin{array}{l} \}6 \\ \}1 \\ \}1 \end{array} \\ \underbrace{\hphantom{A_1'}}_{7} \ \underbrace{\hphantom{A_2'}}_{7} \end{array}, \tag{5.5}$$

where $A_1' = [I_6 \ 0_6^T]$, $B_1' = [1100101]$, $B_2' = [1100101]$ and then A_2' is the 6×7 dimensional matrix defined as

$$A_2'(k_1, k_2) = \begin{cases} 1, & \text{if } k_1 = k_2 \text{ or } k_1 = k_2 - 1 \text{ or } k_1 = k_2 - 2 \\ 0, & \text{otherwise} \end{cases}$$

and $A_2'(5,1) = 1$. Then G has full rank and the associated permutation pair $\boldsymbol{\pi}$ has the property that

$$d_{\min}(\boldsymbol{\pi} \circ \alpha) \geq 4. \tag{5.6}$$

Proof. For G, apply the following row operations. Add the last row to the first row, and then add the first, second and fifth row to the last second row. Now G is written as its systematic format $G' = [I_8 \ P_{8\times6}]$ so G is full rank.

Correspondingly, the parity check matrix of G can be written as $H = [P_{6\times8}^T \ I_6]$ in which the weights of all columns in $P_{6\times8}^T$ are 3. Therefore, it is impossible to choose less than 3 columns from $P_{6\times8}^T$ sum to 0 and also no columns from I_8 can be summed to 0. In other words, we have to pick columns from both I_8 and $P_{6\times8}^T$ so as to obtain 0 column. Since the minimum weight can be achieved from the logical sum of columns from $P_{6\times8}^T$ is 2, at least 2 other columns need to be picked to offset this weight. That is, the smallest number of columns of parity check matrix summing to 0 is 4. Therefore, the minimum Hamming distance of $\mathcal{C}$ with permutation π given by the proposition above is at least 4, i.e., $d_{\min}(\pi \circ \alpha) \geq 4$. $\square$

5.2 2-Diagonal Initial IA

In this section,a permutation pair $\boldsymbol{\pi}$ is proposed achieving $d_{\min}(\pi \circ \alpha) = 3$, for the case of initial 2-diagonal IA.

Proposition 5.3. *For some integer $R \geq 6$, let two matrices G_{π_1} and G_{π_2} given by*

$$G_{\pi_t} = \left(\begin{array}{c:c} A_0^t & A_1^t \\ \hdashline \multicolumn{2}{c}{A_2^t} \end{array}\right) \tag{5.7}$$

where $t = 1, 2$. $A_0^1 = 0_{R-1}^T$, A_0^2 is a $(R-2) \times 2$ dimensional matrix with all elements zero, the matrix $A_2^1 = 1_R$, $A_2^2 = \left(\begin{smallmatrix} 0 & 1 & 1 & 0_{R-6} & 1 & 1 & 0 \\ 1 & 1 & 0 & 0_{R-6} & 0 & 1 & 1 \end{smallmatrix}\right)_{2\times R}$,

$$A_1^1(k_1, k_2) = \begin{cases} 1 \ \textit{if} \ k_1 = k_2 \ \textit{or} \ k_1 = k_2 - 1 \\ 0 \ \textit{otherwise} \end{cases}$$

with dimension $(R-1) \times (R-1)$ and A_1^2 is the $(R-2) \times (R-2)$ dimensional matrix defined as

$$A_1^2(k_1, k_2) = \begin{cases} 1 \ \textit{if} \ k_1 = k_2 \ \textit{or} \ k_1 = k_2 - 1 \\ 0 \ \textit{otherwise} \end{cases} .$$

Then G_{π_1} and G_{π_2} are both full rank and the associated permutation pairs $\boldsymbol{\pi} = (\pi_1, \pi_2)$

has the property that

$$d_{\min}(\boldsymbol{\pi} \circ \alpha) = 3. \tag{5.8}$$

Proof. G_{π_1} has full rank by Theorem 4.1. For G_{π_2}, first switch the positions of the two rows in A_2^2, and denote the obtained matrix by $A_2'^2$. Then switch the positions of $A_2'^2$ and $(A_0^2\ A_1^2)$. The new matrix is upper triangular with all 1's on the main diagonal. Therefore, G_{π_2} is full rank as well.

Let the binary representations of two arbitrary distinct cells of the 2-diagonal IA α, be $[e_1|e_2]$ and $[\bar{e}_1|\bar{e}_2]$, where $e_1, \bar{e}_1$ are for the first description in each cell, while $e_2, \bar{e}_2$ are for the second description. Thus, the Hamming distance between the codewords of the two cells after permutation is given by

$$d_H = H_w((e_1 \oplus \bar{e}_1) \cdot G_{\pi_1} + (e_2 \oplus \bar{e}_2) \cdot G_{\pi_2}).$$

It is clear that the weight of any row in G_{π_1} or G_{π_2} is greater than or equal to 1, therefore, $H_w((e_1 \oplus \bar{e}_1) \cdot G_{\pi_1}) \geq 1$ if $e_1 \neq \bar{e}_1$ and $H_w((e_2 \oplus \bar{e}_2) \cdot G_{\pi_2}) \geq 1$ if $e_2 \neq \bar{e}_2$. For $e_1 = \bar{e}_1$, the two cells are on the same row of the IA matrix, i.e., $e_2 = \bar{e}_2 + 1$ or $\bar{e}_2 = e_2 + 1$ where we use $''+''$ as the addition of natural numbers, more precisely we use $e_2 + 1$ instead of $b(b^{-1}(e_2) + 1)$. For $e_2 = \bar{e}_2$, the two cells are on the same column, i.e., $e_1 = \bar{e}_1 + 1$ or $\bar{e}_1 = e_1 + 1$. Thus, the Hamming distance between their codewords for the two cases is at least 3 by Theorem 4.1. Therefore, in order to complete the proof of (5.8), we only need to show that

$$\text{if } H_w((e_1 \oplus \bar{e}_1) \cdot G_{\pi_1}) = 1 \text{ then } H_w((e_2 \oplus \bar{e}_2) \cdot G_{\pi_2}) \geq 2$$

or

$$\text{if } H_w((e_2 \oplus \bar{e}_2) \cdot G_{\pi_2}) = 1 \text{ then } H_w((e_1 \oplus \bar{e}_1) \cdot G_{\pi_1}) \geq 2.$$

Notice that for the 2-diagonal IA α, there are two relations between two side indices of a cell, i.e., $e_2 = e_1$ or $e_2 = e_1 + 1$. Hence, there are four possible combinations of two arbitrary cells: (1) $[e_1|e_1]$, $[\bar{e}_1|\bar{e}_1]$; (2) $[e_1|e_1+1]$, $[\bar{e}_1|\bar{e}_1]$; (3) $[e_1|e_1]$, $[\bar{e}_1|\bar{e}_1+1]$; (4) $[e_1|e_1+1]$, $[\bar{e}_1|\bar{e}_1+1]$. Since $d_H = H_w((e_1 \oplus \bar{e}_1) \cdot G_{\pi_1} + (e_2 \oplus \bar{e}_2) \cdot G_{\pi_2})$, the second combination and the third combination are equivalent from the point of view of d_H. Thus, only three cases have to be considered.

Now if remains to show that for each of the above three cases, if $H_w((e_1 \oplus \bar{e}_1) \cdot G_{\pi_1}) = 1$, then $H_w((e_2 \oplus \bar{e}_2) \cdot G_{\pi_2}) \neq 1$. The detailed proof can be found in Appendix.

$\square$

Chapter 6

Experiments and Discussions

The purpose of this chapter is to assess the performance in practice of the proposed permuted IA in comparison with the original index assignment. Our tests are performed on a zero mean, unit variance, memoryless Gaussian source. In each case, the central quantizer of the 2-DSQ is optimized by using Vaishampayan's algorithm (Vaishampayan (1993)). In all cases, we consider transmission over independent channels with the same bit error probability $\epsilon \in (0.001, 0.3)$. The performance of each IA is measured by the distortion portion due to the channel at the central decoder in dB, i.e., $10 \log_{10} D_{0,c}$.

We have performed tests for both *Scenario II* and *Scenario III*, in both cases the separate bit errors are independent, identically distributed. In all figures in this chapter, the curve corresponding to the expected central distortion $D_{0,c}$ calculated with the IA before permutation is labeled *Initial IA*, while the label *Robust IA* is for the curve corresponding to the robust IA obtained after applying the permutation presented in chapter 4 or chapter 5. In both cases, the decoder is the minimum Hamming distance decoder given by (2.6) for *Scenario II*, respectively (2.7) for *Scenario*

III. For Fig. 6.1 and Fig. 6.2, which consider *Scenario II*, the labels *Initial(suboptimal)* and *Robust(suboptimal)* indicate that the sub-optimal decoder proposed by Ma and Labeau (2008) with the same IA's as *Initial IA* and *Robust IA*, respectively. Recall that the sub-optimal decoder discards the incorrect description, when the index pair is not valid (see 2.4)

6.1 Experiments for *Scenario II*

Fig. 6.1, Fig. 6.2 and Fig. 6.3 plot the performance of index assignment for *Scenario II* with and without index permutation designed by the method described in Chapter 4, versus BER. Each figure illustrates one of the following cases respectively: (1) 3-diagonal initial IA with $R = 6$; (2) 2-by-2 square initial IA with $R = 4$; (3) 2-diagonal IA with $R = 3$.

For Fig. 6.1 and Fig. 6.2, minimum side 2 Hamming distance of the codebook is 3, which means that 1 bit error can be corrected. With minimum Hamming distance decoder proposed in chapter 2, the expected distortion with IA after permutation is much less than the one before permutation, specifically the maximum difference is 17.5 dB for both cases and it is achieved when $\epsilon = 0.001$. Naturally, as the BER increases, the gap between the performance of the robust IA and initial IA decreases correspondingly. This is due to the fact that as BER increases, the probability of the error patterns which can be corrected by the robust IA decreases. Regarding to the performance of the sub-optimal decoder, notice that it is worse than minimum Hamming distance decoder for small BER ($\epsilon \leq 0.2$). For Fig. 6.3, the minimum side 2 Hamming distance of robust IA is 2. As it can be seen, the robust IA is better than initial IA. An interesting observation here is that the performance gap between the

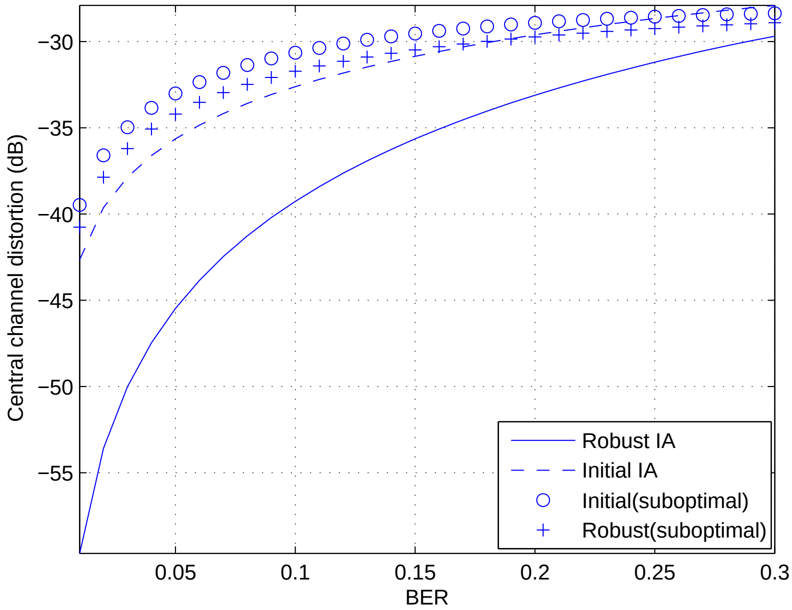


Figure 6.1: 3-diagonal IA with $R = 6$ for *Scenario II*

initial IA and the robust IA does not vary greatly with BER as in the other cases.

In summary, compared with the initial IA, the robust IA based on linear permutation associated with minimum Hamming distance decoder can achieve significant improvements against bit errors, especially when the bit error rate is small.

6.2 Experiments for *Scenario III*

Experimental results for robust 2-by-2 square IA and robust 2-diagonal IA based on linear permutation presented by Chapter 5 are plotted in Fig. 6.4 and Fig. 6.5 respectively. For the purpose of examining the performance of index permutation in details, zooming in figures for both IA schemes are also drawn.

For Fig. 6.4 $R = 5$ and for Fig. 6.5 $R = 6$, the minimum Hamming distance is 3

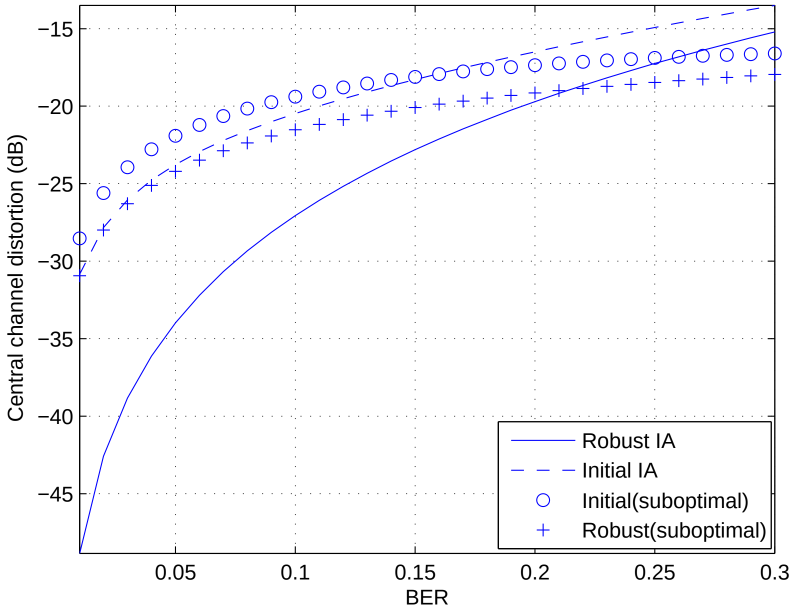


Figure 6.2: 2-by-2 IA with $R = 4$ for *Scenario II*

for robust IA while for initial IA it is 1. As it can be seen, the robust IA outperforms initial IA only when BER is very small, i.e., $\epsilon < 0.05$ roughly. The fact that robust IA is worse than the initial IA as BER increases, which is contrary to the intuition that initial IA should be worse than robust IA (since the former has smaller $d_{\min}$). In an attempt to explain this, notice that $D_{0,c} = \sum_{k=1}^{2R} P(k) \times \epsilon^k \times (1-\epsilon)^{2R-k}$ where $P(k)$ is the portion of the distortion when sent and decoded index pairs differ in k bits. Actually, we have observed that $P(1)$ is higher for initial IA than for robust IA, while $P(k)$ is much higher for robust IA than for the initial IA, in both square and diagonal cases, which means the robust IA introduces higher distortion due to 2 bit errors, and as ϵ increases, this term becomes dominant in $D_{0,c}$.

To summarize, the robust IA with $d_{\min} = 3$ based on linear permutation for initial 2-by-2 square IA with $R = 5$ and initial 2-diagonal IA with $R = 6$ of *Scenario III*

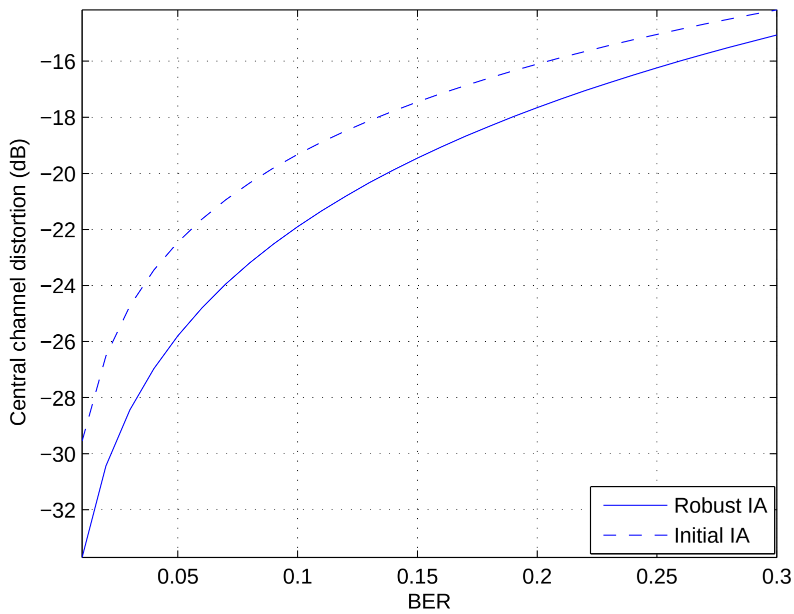
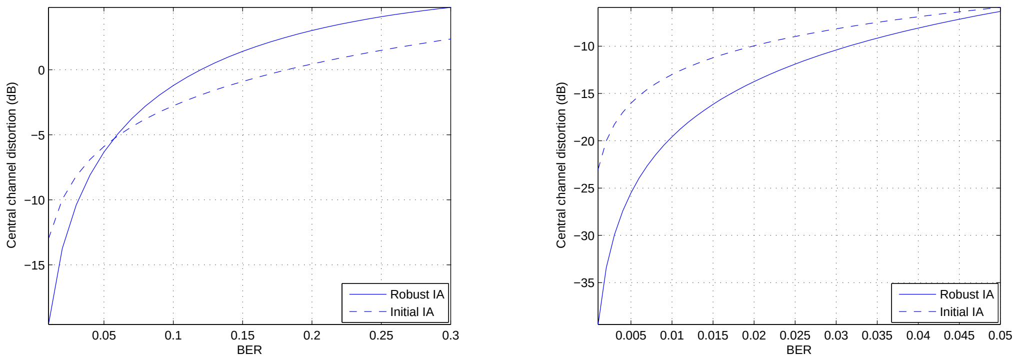


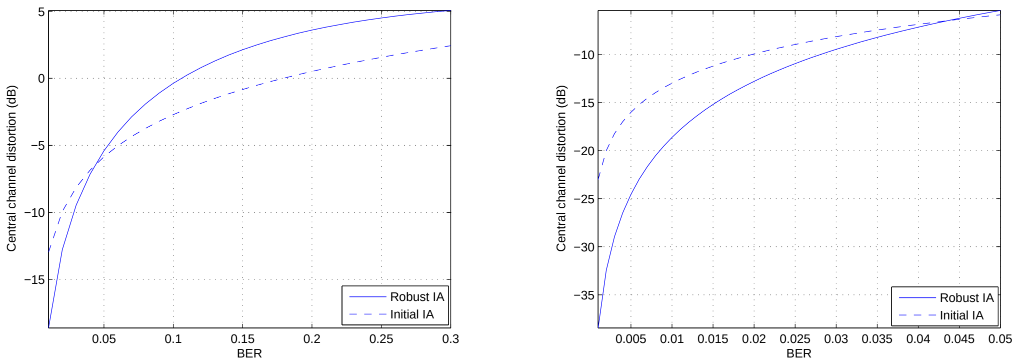
Figure 6.3: 2-diagonal IA with $R = 3$ for *Scenario II*

outperforms the ones before permutation when bit error rate is less than 5%.



(a) (b)

Figure 6.4: 2-by-2 square IA with $R = 5$ for *Scenario III*



(a) (b)

Figure 6.5: 2-diagonal IA with $R = 6$ for *Scenario III*

Chapter 7

Conclusion and Future Work

7.1 Conclusion

For conventional two description scalar quantizer (2-DSQ), some techniques to design index assignment (IA) with good performance, i.e., small distortion, have been presented in prior work. In this thesis, we proposed an approach based on applying an index permutation to an initial IA (diagonal or square based) such as to increase the minimum Hamming distance and thus to increase the robustness of the 2-DSQ's against bit errors.

For the scenario when only one description may incur bit errors and the decoder knows which one, i.e., (*Scenario II*), we introduced a new performance measure for IA robustness, named minimum side Hamming distance $d_{side,\min}$, which is defined as the minimum Hamming distance between valid index pairs for a fixed index of the error-free description. Moreover, we established the connection between the anti-bandwidth problem in a certain graph derived from a hypercube and the problem of designing a robust permutation for the diagonal IA of *Scenario II*. Further, using known results related to the anti-bandwidth of a hypercube, we proposed permutations achieving $d_{side,\min} = 2$. Then we introduced a simple construction for *Scenario II* achieving a general minimum side Hamming distance $d_{side,\min}$ using linear block codes with minimum Hamming distance at least $d_{side,\min} + 1$ and $d_{side,\min}$ for diagonal IA and square IA, respectively.

For the case when both descriptions may incur bit errors, i.e.,(*Scenario III*), we described how to construct permutations achieving minimum Hamming distance 3 for diagonal and square-based initial IA's, in the high redundancy case.

We also performed an experimental study of the proposed index permutation techniques and compared them with the initial index assignment. We observed that the IA after permutation for *Scenario II* works much better than initial IA in terms of central channel distortion, at the same time it maintains the performance of the source distortion. And for *Scenario III*, only when bit error rate (BER) is low, the central channel distortion after permutation is less than the initial IA, while as the BER increases the initial IA shows its advantage.

7.2 Future Work

As we have seen, the proposed robust index permutations for the diagonal and square initial IA show good performance for small BER in both scenarios using minimum Hamming distance decoder. However, some issues still remain unsolved:

1. The assumption of the whole thesis is based on 2 balanced descriptions scalar quantizer, hence a natural question is how to construct robust index permutation for arbitrary number of descriptions and unbalanced case.

2. We have considered only index permutations based on a given initial index assignment and quantizer so far. The problem optimizing the quantizer, the index assignment and the index permutation at the same time is still not solved. In addition, besides m-diagonal IA and 2^l-by-2^l square IA, other IA matrices are also interesting to be considered.

3. For *Scenario III*, current index permutations guarantee that only 1-bit error can be corrected for two initial IA's with specified m and l. However, general constructions to combat larger number of bit errors for general value of m and l are not known, and would be interesting to be investigated.

Appendix A

Here we present the details of the proof of Proposition 5.3.

As explained in Section 5.2, to complete the proof that G_{π_1} and G_{π_2} are two valid permutation matrices achieving minimum Hamming distance $d_{\min} = 3$, it remains to show that $H_w((e_2 \oplus \bar{e}_2) \cdot G_{\pi_2}) = 1$ and $H_w((e_1 \oplus \bar{e}_1) \cdot G_{\pi_1}) = 1$ can not hold simultaneously. For this, define the sets

$$\mathcal{U} = \{u \in \{0,1\}^R | H_w(u \cdot G_{\pi_1}) = 1\}, \tag{A.1}$$

$$\mathcal{V} = \{v \in \{0,1\}^R | H_w(v \cdot G_{\pi_2}) = 1\}. \tag{A.2}$$

By solving the equations $u \cdot G_{\pi_1} = b$ and $v \cdot G_{\pi_2} = b$ for any $b \in \{0,1\}^R$ of Hamming weight 1, we obtain that

$$\mathcal{U} = \{[0_{R-t_1-1}|1_{t_1}|0] | 1 \leq t_1 \leq R-1\} \cup \begin{cases} (10)^k 1, & \text{if } R = 2k+1 \\ (10)^{k-1} 11, & \text{if } R = 2k \end{cases}, \tag{A.3}$$

$$\mathcal{V} = \{[0_{R-t_2-2}|1_{t_2}|00]\} \cup \{1_{R-5}00111\} \cup \{1_{R-5}01110\}, \tag{A.4}$$

where for any n-dimensional row vector e, e^k denote the nk-dimensional row vector obtained by repeating e k times.

Due to the 2-diagonal IA matrix, there are only three cases to be considered: (1) $e_2 = e_1$, $\bar{e}_2 = \bar{e}_1$; (2) $e_2 = e_1 + 1$, $\bar{e}_2 = \bar{e}_1$; (3) $e_2 = e_1 + 1$, $\bar{e}_2 = \bar{e}_1 + 1$. We will discuss the three cases as follows.

Case 1. In this case, we have $e_2 \oplus \bar{e}_2 = e_1 \oplus \bar{e}_1$. According to (A.3) and (A.4), we have $\mathcal{U} \neq \mathcal{V}$, in other words, when $e_1 \oplus \bar{e}_1 \in \mathcal{U}$ and $e_2 \oplus \bar{e}_2 \in \mathcal{V}$ then $e_1 \oplus \bar{e}_1 \neq e_2 \oplus \bar{e}_2$, which contradicts the condition of this case.

Case 2. In this case we have $e_2 \oplus \bar{e}_2 = (e_1 + 1) \oplus \bar{e}_1$. Write $e_1 = [b_1 \cdots b_{R-t-1} | 0 | 1_t]$, $(0 \leq t \leq R-1)$, i.e., $R-t$ is the position of the rightmost 0 in e_1, then $e_2 = e_1 + 1 = [b_1 \cdots b_{R-t-1} | 1 | 0_t]$. Therefore, $e_1 \oplus (e_1 + 1) = [0_{R-t-1} | 1_{t+1}]$, which further implies that

$$
\begin{aligned}
e_2 \oplus \bar{e}_2 =& (e_1 + 1) \oplus \bar{e}_1 \\
=& e_1 \oplus e_1 \oplus (e_1 + 1) \oplus \bar{e}_1 \\
=& e_1 \oplus \bar{e}_1 \oplus [0_{R-t-1} | 1_{t+1}].
\end{aligned}
\tag{A.5}
$$

Based on the above relation, all possible values of $e_2 \oplus \bar{e}_2$ corresponding to $e_1 \oplus \bar{e}_1 \in \mathcal{U}$, are presented in Table A.1. By inspecting the Table A.1, it is clear that when $e_1 \oplus \bar{e}_1 \in \mathcal{U}$, we have $e_2 \oplus \bar{e}_2 \notin \mathcal{V}$, which means that $H_w((e_2 \oplus \bar{e}_2) \cdot G_{\pi_2}) \geq 2$.

Case 3. In this case we have $e_2 \oplus \bar{e}_2 = (e_1 + 1) \oplus (\bar{e}_1 + 1)$. Again, write $e_1 = [b_1 \cdots b_{R-t-1} | 0 | 1_t]$ for some $0 \leq t \leq R-1$ and $\bar{e}_1 = [\bar{b}_1 \cdots \bar{b}_{R-s-1} | 0 | 1_s]$ for some $0 \leq s \leq R-1$. Without loss of generality, assume that $t \geq s$.

When $R = 2k+1$ and $e_1 \oplus \bar{e}_1 = [(10)^k | 1]$, the rightmost bit of $\bar{e}_1$ must be 0, i.e., $s = 0 < t$. Then $e_1 + 1 = e_1 \oplus [0_{R-t-1} | 1_{t+1}]$ and $e_2 + 1 = e_1 \oplus [0_{R-1} | 1]$, which imply

Table A.1: Relation between $e_1 \oplus \bar{e}_1$ and $e_2 \oplus \bar{e}_2$ for $e_2 = e_1 + 1, \bar{e}_2 = \bar{e}_1$

<table>
<tr><td colspan="2"></td><td>$e_1 \oplus \bar{e}_1$</td><td>$e_2 \oplus \bar{e}_2$</td></tr>
<tr><td rowspan="2">case (a)</td><td>$R = 2k+1$</td><td>$(10)^k 1$</td><td>$(10)^s 0 (10)^{k-s}, \forall s, 0 \leq s \leq k$
$(10)^s 1 (10)^{k-s}, \forall s, 1 \leq s \leq k-1$</td></tr>
<tr><td>$R = 2k$</td><td>$(10)^{k-1} 11$</td><td>$(10)^s 0 (10)^{k-s-1} 0, \forall s, 0 \leq s \leq k-1$
$(10)^s 1 (10)^{k-s-1} 0, \forall s, 0 \leq s \leq k-1$</td></tr>
<tr><td colspan="2">case (b)</td><td>$0_{R-t_1-1} | 1_{t_1} | 0$
$\forall t_1, 1 \leq t_1 \leq R-1$</td><td>$1_{R-s} | 1_s, \forall s, 1 \leq s \leq R$
$0_{R-1-s-t} | 1_s | 0_t | 1, \forall s, t \geq 1, s+t \leq R$</td></tr>
</table>

that

$$
\begin{aligned}
e_2 \oplus \bar{e}_2 =& e_1 \oplus \bar{e}_1 \oplus [0_{R-t-1} | 1_t | 0] \\
=& [(10)^k | 1] \oplus [0_{R-t-1} | 1_t | 0],
\end{aligned}
\tag{A.6}
$$

where $1 \leq t \leq R-1$. Clearly, $e_2 \oplus \bar{e}_2 \notin \mathcal{V}$.

Assume now that $R = 2k$ and $e_1 \oplus \bar{e}_1 = [(10)^k | 11]$. Again, the last bit of the bit sequence is 1, which implies that the rightmost bit of $\bar{e}_1$ is 0, i.e., $s = 0 < t$. Thus, $e_2 \oplus \bar{e}_2 = (e_1 + 1) \oplus (\bar{e}_1 + 1)$ has the two rightmost bits 01. Obviously, $e_2 \oplus \bar{e}_2 \notin \mathcal{V}$.

Finally, assume that $e_1 \oplus \bar{e}_1 = [0_{R-t_1-1} | 1_{t_1} | 0]$ for some $1 \leq t_1 \leq R-1$. Because the last bit of $e_1 \oplus \bar{e}_1$ is 0, it follows that either $s = t = 0$ or $s, t > 0$. If $s = t = 0$, then $e_2 \oplus \bar{e}_2 = e_1 \oplus \bar{e}_1 \notin \mathcal{V}$. Let us consider now the case when $s, t > 0$. Further, the fact that $t \geq s$ and that the last two bits of $e_1 \oplus \bar{e}_1$ are 10, implies that $s = 1 < t$. Then

$$\begin{aligned} e_2 \oplus \bar{e}_2 =& e_1 \oplus \bar{e}_1 \oplus [0_{R-t-1}|1_{t-1}|00] \\ =& [0_{R-1-\max\{t_1,t\}}|1_{|t_1-t|}|0_{\min\{t_1,t\}}|10] \\ =& [0_{R-1-\max\{t_1,t\}}|1_{|t_1-t|}|0_{\min\{t_1,t\}}|10], \end{aligned} \tag{A.7}$$

which clearly is not in $\mathcal{V}$.

With this the proof is completed.

Bibliography

Bahceci, I., Altunbasak, Y., and Duman, T. M. (2006). A turbo-coded multiple-description system for multiple antennas. *IEEE Transactions on Communications*, **54**, 187–191.

Balogh, J. and Csirik, J. A. (2004). Index assignment for two-channel quantization. *IEEE Transactions on Information Theory*, **50**, 2737–2751.

Barros, J., Hagenauer, J., and Gortz, N. (2002). Turbo cross decoding of multiple descriptions. In *IEEE International Conference on Communications*.

Bauer, R. and Hagenauer, J. (2000). Symbol-by-symbol map decoding of variable length codes.

Berger-Wolf, T. Y. and Reingold, E. M. (2002). Index assignment for multichannel communication under failure. *IEEE Transaction on Information Theory*, pages 2656–2667.

Dumitrescu, S. and Wu, X. (2005). Optimal two-description scalar quantizer design. *Algorithmica*, **41**, 269–287.

Dumitrescu, S. and Wu, X. (2007). Lagrangian optimization of two-description scalar quantizers. *IEEE Transactions on Information Theory*, **53**, 3990–4012.

Dumitrescu, S. and Wu, X. (2009). On properties of locally optimal multiple description scalar quantizers with convex cells. *IEEE Transactions on Information Theory*, **55**, 5591–5606.

Dumitrescu, S., Wu, X., and Wang, Z. (2004). Globally optimal uneven error-protected packetization of scalable code streams. *IEEE Transactions on Multimedia*, **6**, 230–239.

Fleming, M., Zhao, Q., and Effros, M. (2004). Network vector quantization. *IEEE Transactions on Information Theory*, **50**, 1584–1604.

Gortz, N. and Leelapornchai, P. (2003). Optimization of the index assignments for multiple description vector quantizers. *IEEE Trans. Commun*, **48**, 192–204.

Goyal, V. (2001). Multiple description coding: compression meets the network. *IEEE Signal Processing Magazine*, **18**, 74–93.

Goyal, V., Kovacevic, J., and Vetterli, M. (1998). Multiple description transform coding: Robustness to erasures using tight frame expansions.

Goyal, V. K. and Kovacevic, J. (1998). Optimal multiple description transform coding of gaussian vectors. In *Data Compression Conference*, pages 388–397.

Goyal, V. K. and Kovacevic, J. (2001). Generalized multiple description coding with correlating transforms. *IEEE Transactions on Information Theory*, **47**, 2199–2224.

Goyal, V. K., Kovacevic, J., and Vetterli, M. (1999). Quantized frame expansions as source-channel codes for erasure channels. In *Data Compression Conference*, pages 326–335.

Guionnet, T. and Guillemot, C. (2002). Soft decoding of multiple descriptions.

Gyorgy, A., Linder, T., and Lugosi, G. (2008). Tracking the best quantizer. *IEEE Transactions on Information Theory*, **54**, 1604–1625.

Hagenauer, J. and Gortz, N. (2003). The turbo principle in joint source-channel coding. *IEEE Transactions on Information Theory*.

Harper, L. (1966). Optimal numberings and isoperimetric problems on graphs. *Journal of Combinatorial Theory*.

Leung, J. Y.-T., Vornberger, O., and Witthoff, J. D. (1984). On some variants of the bandwidth minimization problem. *SIAM J. Comput.*, **13**, 650–667.

Lin, S. and Costello, D. J. (2004). *Error Control Coding, Second Edition*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.

Ma, R. and Labeau, F. (2008). Error-resilient multiple description coding. *IEEE Transactions on Signal Processing*, **56**, 3996–4007.

Ma, R. and Labeau, F. (2009). End-to-end performance of robust multiple description scalar quantizer. In *Vehicular Technology Conference*.

Mohr, A. E., Riskin, E. A., and Ladner, R. E. (1999). Graceful degradation over packet erasure channels through forward error correction. pages 92–101.

Muresan, D. and Effros, M. (2008). Quantization as histogram segmentation: Optimal scalar quantizer design in network systems. *IEEE Transactions on Information Theory*, **54**, 344–366.

Park, M. and Miller, D. (1999). Joint source-channel decoding for variable-length encoded data by exact and approximate map sequence estimation. *Acoustics, Speech, and Signal Processing, IEEE International Conference on*, **5**, 2451–2454.

Phamdo, N. and Farvardin, N. (1994). Optimal detection of discrete markov sources over discrete memoryless channels - applications to combined source-channel coding. *IEEE Transactions on Information Theory*, **40**, 186–193.

Puri, R. and Ramchandran, K. (1999). Multiple description source coding through forward error correction codes.

Sayood, K. and Borkenhagen, J. C. (1991). Use of residual redundancy in the design of joint source/channel coders. *IEEE Transactions on Communications*, **39**, 838–846.

Stankovic, V., Hamzaoui, R., and Xiong, Z. (2004). Efficient channel code rate selection algorithms for forward error correction of packetized multimedia bitstreams in varying channels. *IEEE Transactions on Multimedia*, **6**, 240–248.

Torok, L. and Vrto, I. (2009). Antibandwidth of d-dimensional meshes. pages 471–477.

Vaishampayan, V. (1993). Design of multiple description scalar quantizer. *IEEE Transactions on Information Theory*, **39**, 821–834.

Vaishampayan, V. A. and Batllo, J.-C. (1998). Asymptotic analysis of multiple description quantizers. *IEEE Transactions on Information Theory*, **44**, 278–284.

Vaishampayan, V. A. and Domaszewicz, J. (1994). Design of entropy-constrained multiple-description scalar quantizers. *IEEE Transactions on Information Theory*, **40**, 245–250.

Vaishampayan, V. A., Sloane, N. J. A., and Servetto, S. D. (2001). Multiple-description vector quantization with lattice codebooks: Design and analysis. *IEEE Transactions on Information Theory*, **47**, 1718–1734.

Wang, X. and Wu, X. (2010). Index assignment optimization for joint source-channel map decoding. *Trans. Comm.*, **58**, 901–910.

Wang, X., Wu, X., and Dumitrescu, S. (2009). On explicit formulas for bandwidth and antibandwidth of hypercubes. *Discrete Applied Mathematics*, **157**, 1947–1952.

Wang, Y., Orchard, M. T., and Reibman, A. R. (1997). Multiple description image coding for noisy channels by pairing transform coefficients. pages 419–424.

Wu, X., Dumitrescu, S., and Wang, Z. (2004). Monotonicity-based fast algorithms for map estimation of markov sequences over noisy channels. *IEEE Transactions on Information Theory*, **50**, 1539–1544.

Wu, X., Wang, X., and Wang, Z. (2009). Resource-scalable joint source-channel map and mmse estimation of multiple descriptions. *IEEE Transactions on Signal Processing*, **57**, 279–288.

Yahampath, P. (1999). On index assignment and the design of multiple description quantizers. In *International Conference on Acoustics, Speech, and Signal Processing*.

Zeger, K. and Gersho, A. (1990). Pseudo-gray coding. *IEEE Transaction on Communications*, pages 2147–2158.

Zhou, Y. and Chan, W.-Y. (2004). Multiple description quantizer design using a channel optimized quantizer approach.

Zhou, Y. and Chan, W.-Y. (2010). Multiple description quantizer design for multiple-antenna systems with map detection. *IEEE Transactions on Communications*, **58**, 136–145.