Graph-Based Solution for Two Scalar Quantization Problems in Network Systems

GRAPH-BASED SOLUTION FOR TWO SCALAR QUANTIZATION PROBLEMS IN NETWORK SYSTEMS

BY

QIXUE ZHENG, B.Sc.

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 (2018) McMaster University

(Electrical & Computer Engineering) Hamilton, Ontario, Canada

TITLE:	Graph-Based Solution for Two Scalar Quantization Problems in Network Systems
AUTHOR:	Qixue Zheng B.Eng., (Automation Engineering) University of Electronic Science and Technology of China, Chengdu, China
SUPERVISOR:	Dr. Sorina Dumitrescu
NUMBER OF PAGES:	xv, 88

To my family and friends

Abstract

This thesis addresses the optimal scalar quantizer design for two problems, i.e. the two-stage Wyner-Ziv coding problem and the multiple description coding problem for finite-alphabet sources. The optimization problems are formulated as the minimization of a weighted sum of distortions and rates. The proposed solutions are globally optimal when the cells in each partition are contiguous. The solution algorithms are both based on solving the single-source or the all-pairs minimum-weight path (MWP) problems in certain weighted directed acyclic graphs (WDAG). When the conventional dynamic programming technique is used to solve the underlying MWP problems the time complexity achieved is $O(N^3)$ for both problems, where N is the size of the source alphabet.

We first present the optimal design of a two-stage Wyner-Ziv scalar quantizer with forwardly or reversely degraded side information (SI) for finite-alphabet sources and SI. We assume that binning is performed optimally and address the design of the quantizer partitions. A solution based on dynamic programming is proposed with $O(N^3)$ time complexity. Further, a so-called *partial Monge property* is additionally introduced and a faster solution algorithm exploiting this property is proposed. Experimental results assess the practical performance of the proposed scheme.

Then we present the optimal design of an improved modified multiple-description scalar

quantizer (MMDSQ). The improvement is achieved by optimizing all the scalar quantizers. The optimization is based on solving the single-source MWP problem in a coupled quantizer graph and the all-pairs MWP problem in a WDAG. Another variant design with the same optimization but enhanced with a better decoding process is also presented to decrease the gap to theoretical bounds. Both designs for the second problem have close or even better performances than the literature as shown in experiments.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor, Dr. Sorina Dumitrescu, for her patience, encouragement and guidance throughout two years. She is always dedicated, efficient and highly disciplined. She always present rigorous, elegant and intelligent works in academic. It is a great honor to be her student.

I would also like to thank my other committee members, Dr. Jiankang Zhang and Dr. Jun Chen, for taking time to read my thesis and examine my work. Their valuable comments and critiques improve the quality of this thesis.

My sincere appreciation also goes to all my friends in Hamilton, who helped me a lot in my life.

Last but not least, I would like to thank my parents and my brother for their unconditional love and support. A special thanks goes to my boyfriend, Zheng Liu, for his love, company and support.

List of Notation

$\mathcal{X}$	The source alphabet
$\bar{\mathcal{X}}$	The extended source alphabet with $x_0 = -\infty$
d	The distortion function $\mathbb{R} \times \mathbb{R} \to [0, \infty)$
Q	A scalar quantizer
$R(Q)$	The rate of the quantizer Q
$D(Q)$	The distortion of the quantizer Q
$H(A)$	The entropy of the variable A
V	The vertex set $\{0, ..., N\}$
E	The edge set $\{(u, v) \in V \mid 0 \leq u < v \leq N\}$
G	The directed acyclic graph with the vertex set V and the edge set E
$G(\omega)$	G with the weight function ω
$\mathbb{G}$	The coupled quantizer directed acyclic graph
$\mathbb{G}(w)$	The coupled quantizer graph with the weight function w

List of abbreviations

2DSQ	Two-Description Scalar Quantizer
ECMDSQ	Entropy-Constrained MDSQ
DAG	Directed Acyclic Graph
DSC	Distributed Source Coding
F-WZ	Two-stage Wyner-Ziv Coding problem with Forwardly Degraded SI
F-WZSQ	Proposed Scheme Based on Scalar Quantization for the F-WZ problem
HB	Heegard-Berger
MDSQ	Multiple Description Scalar Quantizer
MMDSQ	Modified MDSQ
MWP	Minimum-Weight Path
RD	Rate-Distortion
R-WZ	Two-stage Wyner-Ziv Coding Problem with Reversely Degraded SI
R-WZSQ	Proposed Scheme Based on Scalar Quantization for the R-WZ problem

SI	Side Information
SR	Successive Refinement
SRSQ	Successively Refinable Scalar Quantizer
WDAG	Weighted Directed Acyclic Graph
WZ	Wyner-Ziv

Contents

List of Figures

Chapter 1

Introduction

This thesis addresses the scalar quantizer design for the two-stage Wyner-Ziv (WZ) coding problem and for the multiple description coding (MDC) problem. Both problems are relevant to communication in networks. Due to the instability of the transmission channels, some parts of the message may be lost when transmitting from one user to another. If the received message is not complete, the receiver is not able to understand the message. That can be evidenced by our experience in real life, such as when we surf the Internet and the required image does not show up, or the online video can not be loaded. One way to tackle this problem is using multiple channels to send the message in order to increase the probability of the successful transmission. If we can generate different descriptions for one message and transmit them over different channels, then the user has a larger chance to receive the message with acceptable quality. This is called the MDC problem and will be discussed in Section 1.2.

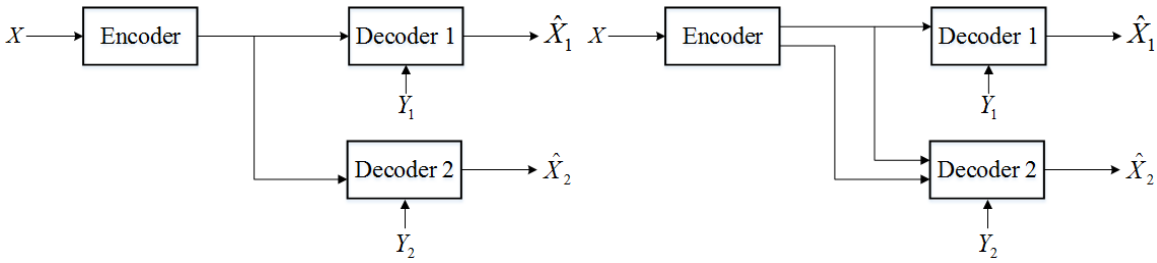
Another way to adapt to the varying quality of channels is to send a coarser description of the message over the channel when it is in a bad quality, then additionally send a refinement of the coarse description when the channel is in a good quality. This way the user can

always recover the source with a tolerable quality. If the receiver already has some side information (SI) about the message, the transmission can be further accelerated by using the correlation between the message and SI, even if the transmitter does not have access to the SI. This is the WZ coding problem and will be presented in Section 1.1.

1.1 Two-stage Wyner-Ziv Coding Problem

Distributed source coding (DSC) refers to the compression of correlated, but isolated sources, which are jointly decoded. The interest in DSC is motivated by applications in sensor networks and video coding. One case of DSC is Wyner-Ziv (WZ) coding, which represents lossy source coding with side information (SI) available only at the decoder[40]. The single-letter characterization of the achievable rate-distortion (RD) region for the WZ problem was derived by Wyner and Ziv in [40].

Motivated by situations where the SI may be present or absent at the decoder, Heegard and Berger [12] and Kaspi [14] studied the scenario where the encoder transmits messages to two decoders, only one of which has SI (depicted in Fig. 1.1(a) with Y_1 a constant). They provided the single-letter characterization of the RD region and explicit expressions for the quadratic Gaussian case and the binary Hamming case. Heegard and Berger generalized the problem to the case of more than two decoders, each with its own SI. We will refer to this general problem as the *Heegard-Berger (HB) problem*. In contrast, we use the term *traditional HB problem* for the two-decoders case where only one has SI. Fig. 1.1(a) illustrates the two-decoder HB problem, where X is the source, while Y_κ is the SI at decoder κ, for $\kappa = 1, 2$. The characterization of the RD region for the HB problem is known when the SI is stochastically degraded [12]. Further contributions to the theoretical study of the HB problem were made in [31] and [32].



(a) Lossy source coding with SI at two decoders. (b) Two-stage WZ coding.

Figure 1.1: System Module.

The problem was further extended to the successive refinement (SR) setting. Fig. 1.1(b) depicts the SR scenario with two stages. Steinberg and Merhav [27] considered multi-stage coding with stochastically degraded SI, i.e., where the decoder receiving higher rate has stronger SI. The authors of [27] characterized the RD region for the two-stage SR problem with degraded SI, i.e., when the Markov chain $X \leftrightarrow Y_2 \leftrightarrow Y_1$ holds. The characterization of the RD region for a general number of stages and degraded SI was given by Tian and Diggavi in [28].

Note that the two decoders in the two-stage scheme of Fig. 1.1(b) could be regarded as two states of a single server-user network, where we expect a coarser reconstruction when SI is weaker and a finer reconstruction when SI is stronger. This interpretation motivates the assumption of degraded SI as in [27]. On the other hand, the two decoders can be regarded as two different users in a multi-user network, where we expect faster decoding for the user with stronger SI. This point of view led Tian and Diggavi [29] to investigate the two-stage coding scenario where the first decoder has stronger SI, i.e., the Markov chain $X \leftrightarrow Y_1 \leftrightarrow Y_2$ holds. They termed this problem *SI-scalable coding*. Tian and Diggavi [29] provided inner and outer bounds to the RD region for general discrete memoryless sources. Furthermore, they derived the complete RD region for multi-stage source coding for quadratic Gaussian source with multiple jointly Gaussian SI.

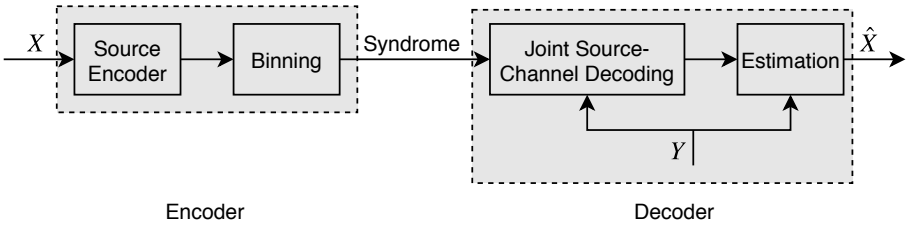


Figure 1.2: Block diagram of a one-stage generic WZ coder.

The research on the theoretical aspects of source coding with varying SI at the decoders was paralleled by the investigation of practical coding schemes. While the information-theoretical results are non-constructive[1] they inspire the practical constructions. The theoretical coding schemes for problems with SI only at the decoder(s) use quantization and binning as building blocks. The diagram of one-encoder one-decoder WZ coding, cited from [41], is shown in Fig. 1.2. For the practical implementation of binning, cosets of powerful linear channel codes are generally used, while for the quantization part various scalar or vector quantizers are employed, including lattice and trellis-based quantizers [21, 24, 41].

Practical schemes for the multiple-decoder WZ problem were proposed in [2, 8, 17, 22, 25, 35, 36, 42]. Cheng and Xiong [2] considered the case when SI is the same at all decoders. Their scheme is based on uniform nested scalar quantizers in conjunction with low density parity check (LDPC) codes for binning. Similar approaches are used in [8, 17, 35, 36, 42] to implement WZ schemes with degraded or identical SI, targeting applications in robust video coding. Ramanan and Walsh [22] proposed a coding scheme for the traditional HB problem using successively refinable trellis coded quantization and LDPC-based codes for binning. Very recently, Shi et. al. [25] have introduced a construction for the traditional HB problem for binary and Gaussian sources based on nested polar codes,

[1] Such results are based on random-coding arguments and show that schemes achieving the claimed performance exist, but do not explain how to construct the corresponding codebooks.

respectively nested polar lattices.

As seen from the above discussion most of existing pracatical schemes for the multiple-decoder WZ problem use uniform quantizer partitions. It is natural to think that better performance can be achieved by employing optimized quantizer partitions. Such an approach was taken by Rebollo-Monedero *et al.* [23] and Muresan and Effros [18, 19] who addressed the design of scalar quantizers for the single-encoder single-decoder WZ problem, under the assumption that the binning is performed optimally achieving the Slepian-Wolf (SW) rate[26]. Both works formulate the problem as the minimization of a weighted sum of the distortion and rate. The algorithm of [23] is an iterative algorithm in the spirit of Max-Lloyd's algorithm, which guarantees only a locally optimal solution in general. The approach of Muresan and Effros [18] is to model the problem as a minimum-weight path (MWP) problem in a certain weighted directed acyclic graph (WDAG). This approach ensures globally optimal solution for the case of finite-alphabet sources, subject to the constraint that the quantizer cells are contiguous[2]. The authors of [18, 19] also proposed globally optimal design algorithms for successively refinable scalar quantizers (SRSQ) (also termed multiresolution scalar quantizers) without SI at the decoders and for multiple description scalar quantizers (MDSQ), subject to the same constraints as above. They addressed both the fixed-rate and entropy-constrained cases. Additionally, Muresan and Effros pointed out that their designs could be easily extended to the case of SRSQ and MDSQ with SI at the decoders. It is worth emphasizing that an algorithm for the entropy-constrained SRSQ design similar to the one of [19] was developed independently by Dumitrescu and Wu in [4]. Additionally, faster globally optimal design algorithms for fixed-rate SRSQ were developed by Dumitrescu and Wu in [38, 5] and for fixed-rate MDSQ

[2] A cell is said to be contiguous if it equals the intersection between the source alphabet and an interval of the real line.

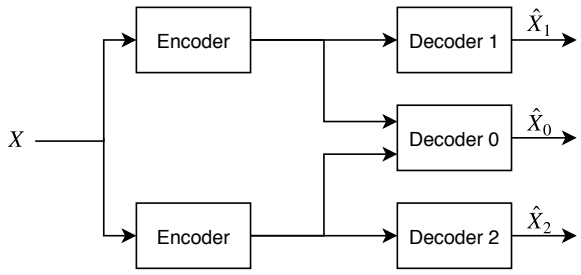


Figure 1.3: Multiple description coding with two channels and three receivers.

in [6, 7], also for finite-alphabet sources under the constraint of cell contiguity. The key technique in the latter works was to prove that the components of the cost function satisfy the so-called *Monge property* [1], which was further exploited to accelerate the design procedure.

1.2 Multiple Description Coding Problem

To guarantee robust communication, two channels are used to transmit the message from the encoder to the decoder. Transmission failure may occur in either one of both channels. When only one of both channels has a successful transmission, a coarse reconstruction is built based on the message from that channel. When the messages from both channels have been successfully received, a finer reconstruction is generated at the user. This scenario can be represented as the module in Fig. 1.3, where the source X is encoded to two messages sent over two channels. Each of the two side decoders only receive the message from one of the channels, while the central decoder receives messages from both channels. Ozarow [20] gave the optimal RD region for a memoryless unit-variance Gaussian source with mean-squared distortion measure. For general independent identically distributed sources and general distortion measurements, an achievable RD region has been found by Gamal

and Cover [11].

For practical implementation, the multiple description scalar quantizer (MDSQ) design of [33] and the entropy-constrained MDSQ (ECMDSQ) design of [34] are both based on a choice of an index assignment and optimization of the central partition. The index assignment step is used to generate the side partitions based on the central partition. The design algorithms of [33], [34] only guarantee local optimality for fixed index assignment. Another design for the multiple-description problem is proposed in [30]. It has a two-stage encoding structure based on scalar quantizers, and is termed modified MDSQ (MMDSQ). The MMDSQ consists of two uniform quantizers with staggered bins at the first stage, forming the side partitions, and another uniform quantizer inside each joint bin at the second stage, forming the central partition. The MMDSQ only finds contiguous intervals at both central and side quantizers, which avoids complex index assignments leading to simpler implementation. Although the side decoders have no access to the partition at the second stage, MMDSQ has the same asymptotic performance as ECMDSQ at high rates. Liu and Zhu [15] further enhanced the MMDSQ scheme by using the refinement message at the second stage to improve the distortions at the side decoders. Their enhancement reduced the asymptotic gap to the theoretical bounds from 3.07 dB to 2.486 dB at high rates. Another two-stage system based on dithered lattice quantizers can be found in [10].

1.3 Original Contribution

Our contributions are presented in two parts, where the contribution of the two-stage WZ problem is described in the first subsection and that of the MDC problem is presented in the Subsection 1.3.2.

1.3.1 Two-Stage Wyner-Ziv Coding Problem

In this thesis we address the design of coding schemes based on scalar quantization for the two-stage WZ coding problem with either forwardly degraded SI, i.e., when $X \leftrightarrow Y_2 \leftrightarrow Y_1$ holds, or reversely degraded SI, i.e., when $X \leftrightarrow Y_1 \leftrightarrow Y_2$ holds. We address the case when the source and the SI have finite alphabets. We use the acronyms F-WZ (respectively, R-WZ) for the two-stage WZ coding problem with forwardly degraded SI (respectively, reversely degraded SI). Additionally, we utilize the notation F-WZSQ and R-WZSQ for the proposed schemes based on scalar quantization for the F-WZ and R-WZ problems, respectively. Our approach is to separate the quantization and the binning parts and, like [23] and [18], to assume that binning and/or nested binning are performed optimally achieving the theoretical limits and focus on the optimal design of the scalar encoder partitions.

The proposed schemes are inspired by the random coding-based schemes used to prove the achievability of the RD regions proposed in [27] and [29], respectively. Thus, the encoder of the F-WZSQ scheme consists of two nested partitions (a coarse and a fine partition), while the encoder of the R-WZSQ scheme is composed of a coarse partition and two independent refinements, one for each decoder. In each case the optimization problem is formulated as the minimization of a weighted sum of the distortions and rates. The proposed solution algorithms are delivered in two stages. First we show how the problem can be decomposed into solving the all-pairs MWP problem in two WDAGs for R-WZSQ, respectively in one WDAG for F-WZSQ, followed by solving the MWP problem in another WDAG. For this we closely follow the approach developed in [19, 4] for entropy-constrained SRSQ design (without SI at the decoder), which also involves optimizing nested partitions. The main difference versus [19, 4] resides in the expression of

the cost function which has to account for the presence of the SI at the decoder. Another difference is manifested in the R-WZSQ case and stems from the fact that the coarse partition has two refinements, not just one as in SRSQ. If conventional algorithms are further used to solve the aforementioned MWP problems then the time complexity of the solution amounts to $O(N^3)$, where N denotes the size of the alphabet of the source X. This claim holds under the assumption that the sizes of the alphabets of Y_1 and Y_2 are $O(N)$. Note that the aforementioned solution algorithm for each problem is globally optimal under the assumption that the cells in each partition are contiguous.

In the following stage of our exposition we introduce the *partial Monge property* and show how the solution developed in the first stage can be accelerated when this property holds. The Monge property was shown to hold in several design problems for systems based on fixed-rate scalar quantizers and was leveraged to achieve significant complexity reduction in comparison with conventional algorithms [39, 38, 5, 6, 7]. It is important to highlight that the aforementioned works which exploit the Monge property require the property to hold for all graph edges of the WDAGs in the problem modeling. Unfortunately, this requirement is not satisfied in the entropy-constrained case, as is ours. However, we have observed empirically that the Monge property is fulfilled for a certain structured subset of the edges of the aforementioned WDAG. We refer to this as the partial Monge property and prove that when it holds it still can be utilized to expedite the solution.

To summarize, our contribution lies in the following aspects.

- We extend the approach of [19, 4] for the design of entropy-constrained SRSQ to obtain globally optimal solutions for the design of F-WZSQ and R-WZSQ schemes for finite alphabet sources and SI, under the assumption that binning is performed optimally using SW coding. The algorithms run in $O(N^3)$ time, where N is the size

of the source alphabet, while the sizes of the alphabets of the SI Y_1 and Y_2 are also $O(N)$. The claim of global optimality holds for the class of F-WZSQs/R-WZSQs with contiguous quantizer cells. This is the first work to address the optimization of the scalar quantizers for the two-stage WZ problem, up to our knowledge.

- We introduce the partial Monge property in a complete WDAG[3] and show how this can be exploited to speed up the dynamic programming solution algorithm for the all-pairs MWP problem.

- We prove that if the partial Monge property holds in the underlying WDAGs then the time complexity of the F-WZSQ and R-WZSQ design algorithms can be significantly reduced.

- We show empirically, using a discretized Gaussian source with discretized Gaussian SI, that the partial Monge property holds in many situations of interest, thus allowing for the fast F-WZSQ/R-WZSQ design algorithm to be employed.

1.3.2 Multiple Description Coding Problem

In the second part of the thesis, we address the design of an improved MMDSQ scheme. The proposed scheme improves Tian and Hemami's work [30] by replacing the uniform partitions with the optimal ones at the two stages. The problem is formulated as a minimization of a weighted sum of rates and distortions. The optimization is resolved by finding the all-pairs MWP in a WDAG followed by solving the single-source MWP problem in a *coupled quantizer graph*.

In the spirit of [15], we also propose the design of improved MMDSQ with enhanced

[3] A WDAG is called complete if any two nodes are connected by an edge.

decoders, which has the same optimization process as the improved MMDSQ but the side distortions are further improved by changing the encoding manner of the messages to be transmitted and the decoding rule at the side decoders.

The advantages of our improved MMDSQ design over existing works are listed as follows.

- Our design uses optimal partitions at both stages such that we can achieve a smaller gap to the theoretical bound than MMDSQ, especially at low rates. Additionally, we can achieve more trade-offs than MMDSQ as shown by our experimental results.
- Our scheme is more flexible than MDSQ in that different trade-offs are obtained just by varying the weights in the objective function without loss in performance.
- Our design is capable of handling both the symmetric and asymmetric cases because there is no constraint on the side partitions, while the MMDSQ only considers the side quantizers with staggered partitions.
- Finally, our design of improved MMDSQ with enhanced decoders has obtained the best results among all the considered designs, i.e., ECMDSQ, MMDSQ, enhanced-MMDSQ [15] and our improved MMDSQ.

1.4 Organization and Related Publications

This thesis consists of five chapters. This chapter presents the introduction and literature review of the two problems, and points out our contributions. Chapter 2 reviews the background knowledge related to our problems including of the graph models for the conventional scalar quantizer and the two-description scalar quantizer with convex cells, and

the theoretical results of WZ coding. In Chapter 3, we investigate the optimal design of a two-stage WZ scalar quantizer in two cases, namely, where the source and the SI form a forward or reversed Markov chain. As a byproduct, we also give the optimal solution for the tradition HB problem. In Chapter 4, the proposed MMDSQ [30] design is demonstrated. Chapter 5 concludes the thesis and presents some directions for future works. Appendix A presents the proofs of the two statements established in Chapter 3 relevant to the Monge property.

The content of this thesis is also contained in three papers. The work in Chapter 3 was first presented in the conference paper [44] in part then extended and summarized in the journal paper [45] that has been submitted to the *IEEE Transactions on Communications*. The content in Chapter 4 is included in [43], which is still in preparation.

Chapter 2

Background Knowledge

This section presents the preliminary, definition and notations. We first review the MWP problem in a WDAG, then describe the graph model for the conventional scalar quantizer design problem in Section 2.2. The graph-based solution for entropy-constrained two-description scalar quantizer design problem is presented next, which is closely related to the fixed-rate design proposed by Dumitrescu and Wu in [6]. Subsection 2.4 presents the theoretical RD region for both one and two-stage WZ coding problem for the Gaussian source and jointly Gaussian SI. Finally, Subsection 2.5 concludes this chapter.

2.1 General Notations

Those notations are used throughout this thesis. Let $d : \mathbb{R} \times \mathbb{R} \rightarrow [0, \infty)$ denotes the distortion function. We will assume that function $d(\cdot)$ is monotone, i.e., for any real x, y_1 and y_2, if $x \leq y_1 < y_2$ or $x \geq y_1 > y_2$, then $d(x, y_1) \leq d(x, y_2)$. Note that the majority of distortion measures of signal quantization used in practice fall into this category. Let the alphabet of the source X be $\mathcal{X} = \{x_1, \cdots, x_N\} \subset \mathbb{R}$, where the elements are labeled in

increasing order. Denote $x_0 = -\infty$ and $\bar{\mathcal{X}} = \mathcal{X} \cup \{x_0\}$. Let $\hat{\mathcal{X}}$ be the reconstruction alphabet of the source X. When the distortion measure is the squared difference we consider $\hat{\mathcal{X}} = \mathbb{R}$. Otherwise, we take a finite set as $\hat{\mathcal{X}}$ with $|\hat{\mathcal{X}}| = O(N)$, where $|S|$ denotes the cardinality of the set S. Further, we say that a set $S \subseteq \mathcal{X}$ is contiguous or convex if there exists $x_u, x_v \in \bar{\mathcal{X}}$ with $u < v$ such that $S = (x_u, x_v]$, where $(x_u, x_v] \triangleq \{x \in \mathcal{X} | x_u < x \leq x_u\}$. For any integer $n \geq 2$, an ascending n-sequence is an n-tuple $\mathbf{r} = (r_0, r_1, \cdots, r_{n-1})$, where $r_0 < r_1 < \cdots < r_{n-1}$ and $r_i \in \bar{\mathcal{X}}$, for $0 \leq i \leq n-1$.

2.2 Graph Model for the Scalar Quantizer Design Problem

The encoder of a scalar quantizer is given by a set of thresholds partitioning the source to disjoint cells. For a finite-alphabet source, the partition can be seen as a path in a WDAG and the entropy-constrained scalar quantizer design problem can be cast as a MWP problem in this WDAG. This section first reviews the MWP problem in a WDAG, then presents a brief discussion of the graph model for the entropy-constraint scalar quantizer.

2.2.1 MWP in A WDAG

A DAG (short for directed acyclic graph) consists of a set of vertices (or nodes) V and a set of directed edges E. In this work we consider $V = \{0, ..., N\}$ and $E = \{(u, v) \in V^2 | 0 \leq u < v \leq N\}$ where N is the size of the source alphabet. We denote by G this DAG. Note that G is a "complete" DAG, meaning that any two nodes are connected by an edge. If we assign a real value, called "weight", to each edge, the graph becomes a WDAG (short for weighted DAG). Let $G(\omega)$ denote the WDAG obtained from the DAG G with

the weight function $\omega : E \to \mathbb{R}$. A path in the WDAG is a sequence of connected edges. Alternatively, a path can be regarded as a sequence of nodes, where any two consecutive nodes are connected by an edge. The weight of the path is the sum of the weights of its edges. The MWP problem in the WDAG is the problem of finding the path of minimum weight from the source node to the final node, where one node is designated as the source and another as the final node. The solution to this problem essentially finds the MWP from the source node to any other node in the graph, i.e., it solves what is referred to as the single-source MWP problem. Let u be the source node. For each $u \leq n \leq N$, let $\hat{W}_u(n)$ denote the weight of the MWP from node u to node n in the WDAG $G(\omega)$. Thus, $\hat{W}_u(u) = 0$ and the following recurrence relation holds

$$\hat{W}_u(n) \triangleq \min_{u \leq m < n} (\hat{W}_u(m) + \omega(m, n)), \tag{2.1}$$

for all $u < n \leq N$. Thus, the single-source MWP problem can be solved using dynamic programming based on (2.1) in $O(N^2)$ time when all edge weights are given. A related problem is the all-pairs MWP problem, which refers to finding the MWP between any pair of nodes of the WDAG. The latter problem can be solved in $O(N^3)$ time, when all edge weights are known, simply by solving the single-source MWP problem N times, each time a different node being the source.

2.2.2 Mapping of The Scalar Quantizer Design Problem to The MWP Problem in A WDAG

A scalar quantizer, denoted by Q, consists of an encoder and a decoder. This work only considers the scalar quantizers with contiguous cells. The encoder is a set of thresholds

$\mathbf{r} = (r_0, r_1, .., r_M)$ dividing the whole source alphabet into M disjoint cells such that i-th cell is $C_i = (r_{i-1}, r_i]$. The decoder reconstructs the source based on the codebook for each cell. Let the average distortion and entropy for a quantizer Q be denoted by $D(Q)$, respectively, $R(Q)$. Any quantizer Q achieving a point on the lower hull of the theoretical RD region is optimal in the sense that there is no other point that can achieve a smaller distortion with the same or lower entropy. As it is well-known, Lagrangian multiplier can help to find the points on the lower convex hull of a region. Consequently, finding an optimal scalar quantizer is equivalent to minimizing the function:

$$\mathcal{O}(Q) = D(Q) + \lambda R(Q), \tag{2.2}$$

for any fixed $\lambda > 0$. A major observation is that $D(Q)$ and $R(Q)$ are additive over code-cells. Let $\hat{x}(C)$ and $p(C) = \sum_{x \in C} p(x)$ denote the codeword, respectively, the probability for a cell $C \subseteq \mathcal{X}$. We can write the entropy and the expected distortion as

$$R(Q) = \sum_{i=1}^{M} h(C_i), \qquad D(Q) = \sum_{i=1}^{M} d(C_i),$$

where $h(C)$ and $d(C)$ for any cell $C \subseteq \mathcal{X}$ are defined as

$$h(C) = -p(C) \log p(C), \qquad d(C) = \sum_{x \in C} p(x) d(x, \hat{x}(C)). \tag{2.3}$$

The mapping between an optimal scalar quantizer and a MWP in a WDAG is performed as following. First, a DAG G can be constructed with vertex set $V = \{0, 1, ..., N\}$ representing the index of $N + 1$ symbols in the set $\bar{\mathcal{X}}$. Any edge (u, v) with $u < v$ in the edge set E corresponds to a cell $(x_u, x_v]$. Then a path from the source node 0 to the final node N has

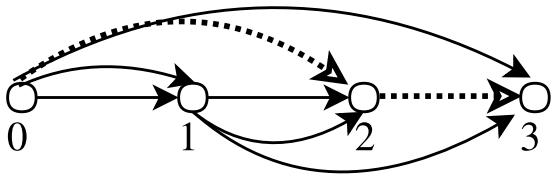


Figure 2.1: The DAG for a source alphabet with $N = 3$. All possible edges are shown. One possible path from node 0 to node 3 appears in dashed arcs, which correspond to the partition $\{\{x_1, x_2\}, \{x_3\}\}$.

a one-to-one correspondence to a partition $\mathbf{r}$ where the i-th edge of the path corresponds to the i-th cell of the partition. An example is illustrated in Fig. 2.1 for $N = 3$. Now define the weight function ω for each edge (u, v) as $\omega(u, v) = d((x_u, x_v]) + \lambda h((x_u, x_v])$. Clearly, the weight of a path equals the cost of the function in (2.2). Therefore, finding the optimal entropy-constrained scalar quantizer is equivalent to solving the single-source MWP problem in the WDAG $G(\omega)$.

In the WDAG $G(\omega)$, the $N + 1$ nodes are already in topological order, finding the shortest path takes $O(|V| + |E|)$ time. The solution algorithm based on (2.1) is shown in Algorithm 1. In conclusion, finding the optimal scalar quantizer for a source with alphabet of size N takes $O(N^2)$ time relying on solving the single-source MWP problem in a WDAG.

This technique of mapping the scalar quantizer problem to a graph problem can also be extended to other lossy source coding scenarios, like multiple resolution coding, multiple description coding and Wyner-Ziv coding as will shown in the next two chapters.

Algorithm 1: Solve the single source MWP problem in $G(\omega)$

```
Let s be an array of size N + 1, which will hold the cost of the MWP from node 0.
  Initialize s[0] = 0, all other s[v] = ∞.
Let t be an array of size N + 1 with all the elements initialized to 0, which will
  hold the last visited node of the MWP from node 0 to all the other nodes.
for v = 1 to N do
    for u = 0 to v − 1 do
        Let w be the weight of the edge from u to v.
        Find the MWP to v:
        if s[v] > s[u] + w then
            s[v] ← s[u] + w;
            t[v] ← u.
```

2.3 Graph Model for the Two-Description Scalar Quantizer Design Problem with Convex Cells

This section introduces the way of using a graph model to solve the two-description scalar quantizer (2DSQ) design problem under the entropy constraint. We restrict the 2DSQ to have convex cells at side quantizers. We will use the *coupled quantizer graph* model proposed in [6], where it is employed to find the optimal fixed-rate 2DSQ with convex cells. In this section, we will present the entropy-constrained 2DSQ. As a result, the coupled quantizer graph used here is different from the one in [6] in that the weight for an edge in our case contains not only the distortion term but also the entropy term. An algorithm is further proposed based on dynamic programming. The time complexity is $O(N^3)$ for a source with alphabet size of N.

A 2DSQ consists of three quantizers $\mathbf{Q} = (Q_1, Q_2, Q_{12})$ where Q_1 and Q_2 are two side quantizers, Q_0 is the central quantizer obtained by intersecting two side partitions. An illustration of the partitions for three quantizers is shown in Fig. 2.2, where the thresholds

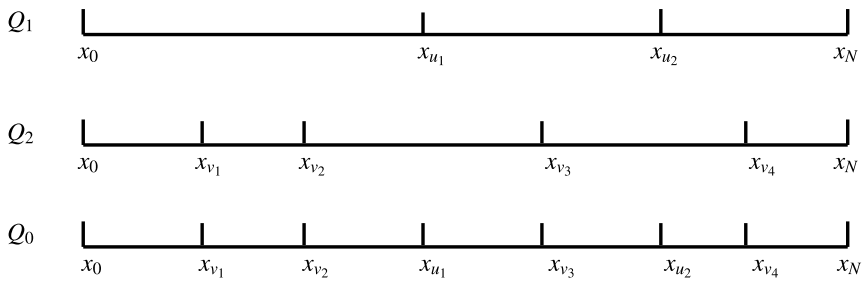


Figure 2.2: Example of partitions of two-description scalar quantizer. Q_1 and Q_2 are two side quantizers, Q_0 is the central quantizer obtained by intersecting two side partitions.

of Q_1 are specified by the variable x_{u_i}, those of Q_2 by x_{v_j}, those of Q_0 by all the x_{u_i} and x_{v_j}, for $i = 1, 2$ and $j = \{1, ..., 4\}$.

Let the distortions at two side decoders denoted by $D_1(\mathbf{Q}), D_2(\mathbf{Q})$ and the distortion at the central decoder by $D_0(\mathbf{Q})$. Let $R_1(\mathbf{Q})$ and $R_2(\mathbf{Q})$ denote the rate of each description. i.e. the rate of Q_1 and Q_2. Then the RD performance of the 2DSQ can be characterized by the quintuple $(R_1(\mathbf{Q}), R_2(\mathbf{Q}), D_0(\mathbf{Q}), D_1(\mathbf{Q}), D_2(\mathbf{Q}))$. An optimal entropy-constrained 2DSQ minimizes a weighted sum of the distortions and rates as follows.

$$\mathcal{O}(\mathbf{Q}) = \sum_{i=0}^{2} \mu_i D_i(\mathbf{Q}) + \sum_{i=1}^{2} \lambda_i R_i(\mathbf{Q}), \tag{2.4}$$

for any $\mu_0, \mu_1, \mu_2, \lambda_1, \lambda_2 > 0$. Any quintuple on the lower convex hull of the theoretical region can be found by minimizing function (2.4) for some positive weights.

Each side quantizer is a partition for the $N + 1$ symbols in $\bar{\mathcal{X}}$. Two side quantizers together determine the central quantizer. Let u denote the index of a threshold in Q_1 and v denote the index of a threshold in Q_2. Then a *coupled quantizer* DAG, denoted by $\mathbb{G}$ can be constructed with the vertex set $\mathbb{V} = \{uv | 0 \leq u, v \leq N\}$ and the edge set $\mathbb{V}$. The edge set $\mathbb{E}$ consists of two types of edges where the edge from uv to $u'v$, denoted by $(uv, u'v)$,

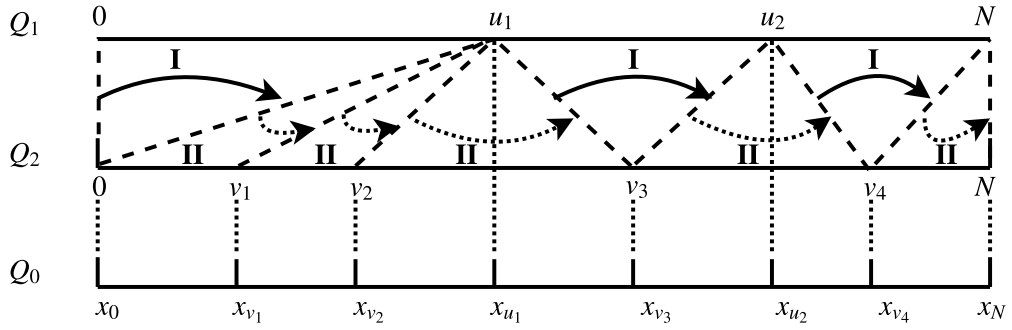


Figure 2.3: A path from 00 to NN in the coupled quantizer graph for the quantizers in Fig. 2.2. Dashed lines connecting a pair of thresholds of two partitions represent a node in the graph. Two types of edges are labeled in **I** and **II**. Each edge generates one side cell and no more than one central cells in Q_0.

with $u < u'$ and $u \leq v$ is called type I edge, while the edge from uv to uv', denoted by (uv, uv'), with $v < u$ and $v < v'$ is called type II edge. Only these two types of edges exist in $\mathbb{V}$. In other words, for a node uv, if $u \leq v$ then only type I edge exist, otherwise only type II edge exists. Note that the coupled quantizer DAG $\mathbb{G}$ here is not a complete graph in that not any two nodes are connected by an edge. The weight for each type I edge is defined as

$$w(uv, u'v) = \mu_1 d((x_u, x_{u'}]) + \lambda_1 h((x_u, x_{u'}]) + \mu_0 d((x_u, x_{\min(v,u')}]).$$

where $d(C)$ and $h(C)$ are the expected distortion and the entropy for cell $C \subseteq \mathcal{X}$ as defined in (2.3). The weight for type II edge is defined as

$$w(uv, uv') = \mu_2 d((x_v, x_{v'}]) + \lambda_2 h((x_v, x_{v'}]) + \mu_0 d(x_v, x_{\min(u,v')}].$$

Now the coupled quantizer graph can be represented by $\mathbb{G}(w)$. The number of type I edges equals the number of cells in Q_1, respectively, the number of type II edges equals the

Algorithm 2: Solve the single-source MWP problem in $\mathbb{G}(w)$

```
for u = 0 to N do
    for v = 0 to N do
        Find the best type I edge to current node uv: for k = 0 to min(u, v) do
            Let w_1 be the weight of the edge from kv to uv.
            Find the MWP to uv:
            if s[uv] > s[kv] + w_1 then
                s[uv] ← s[kv] + w_1;
                t[uv] ← kv.
        Find the best type II edge to current node uv: for k = 0 to min(u, v) do
            Let w_2 be the weight of the edge from uk to uv.
            Find the MWP to uv:
            if s[uv] > s[uk] + w_2 then
                s[uv] ← s[uk] + w_2;
                t[uv] ← uk.
```

number of cells in Q_2. This way a 2DSQ can be mapped to a unique path form the source node 00 to the final node NN in $\mathbb{G}(w)$. Additionally, the weight of a path from 00 to NN equals the cost of the function in (2.4). Consequently, finding the optimal 2DSQ is equivalent to solve the single-source MWP problem in the coupled quantizer graph $\mathbb{G}(w)$. Fig. 2.3, adapted form [6], depicts an example of the coupled quantizer graph $\mathbb{G}(w)$ constructed based on the quantizers in Fig. 2.2. A unique path from the source node to the final node, consisting of both type I and type II edges, corresponds to the partitions in Fig. 2.2. Each edge in $\mathbb{G}(w)$ generates no more than one central cells in Q_0.

The algorithm to solve the single-source MWP problem in $\mathbb{G}(w)$ is shown in Algorithm 2, where s is an array of size $(N+1)^2$, which will hold the weight of the MWP from node 00 to any other nodes. Initialize $s[00] = 0$, all other $s[uv] = \infty$. t is an array of size $(N+1)^2$ with all the elements initialized to 00, which will hold the last visited node of the MWP from node 00 to all the other nodes. The search of the MWP for all nodes proceeds

in the lexicographical order, then for a given node uv, the MWPs from the source node to all the nodes prior to uv have been found. Therefore, the searches in Line $6-8$ and Line 12-14 require $O(N)$ time. Since there are $O(N^2)$ nodes, the total time complexity becomes $O(N^3)$.

2.4 Theoretical Perspective of Wyner-Ziv coding

The one-stage WZ coding with multiple decoders was first addressed by Heegard and Berger [12]. They presented the least rate to satisfy both distortion requirements in the two-decoder WZ coding system in Fig. 1.1(a). They also extended RD results to more than two decoders with degraded SI, i.e. $X \leftrightarrow Y_m \leftrightarrow ... \leftrightarrow Y_1$. We use R_{HB} to denote their results. Assume the source is a standard Gaussian $X \sim \mathcal{N}(0, \sigma_x^2)$ and the side informations is $Y = X + Z$, where $Z \sim \mathcal{N}(0, \sigma_z^2)$ and Z is independent from each other and the source. The theoretical RD region for the two-decoder case is defined as follows.

$$R_{HB}(D_1, D_2) = \begin{cases} \frac{1}{2}\log_2(\frac{\sigma_x^2\sigma_z^2}{D_1(D_2+\sigma_z^2)}), & \text{if } D_1 \leq \frac{D_2\sigma_z^2}{D_2+\sigma_z^2}, D_2 \leq \sigma_x^2, \\ \frac{1}{2}\log_2(\frac{\sigma_x^2}{D_2}), & \text{if } D_1 > \frac{D_2\sigma_z^2}{D_2+\sigma_z^2}, D_2 \leq \sigma_x^2, \\ \frac{1}{2}\log_2(\frac{\sigma_x^2\sigma_z^2}{D_1(\sigma_x^2+\sigma_z^2)}), & \text{if } D_1 \leq \frac{D_2\sigma_z^2}{D_2+\sigma_z^2}, D_2 > \sigma_x^2, \\ 0, & \text{if } D_1 > \frac{D_2\sigma_z^2}{D_2+\sigma_z^2}, D_2 > \sigma_x^2. \end{cases}$$

For the two-stage WZ coding problem, we only consider the cases when either $X \leftrightarrow Y_2 \leftrightarrow Y_1$ or $X \leftrightarrow Y_1 \leftrightarrow Y_2$ holds. The theoretical RD region is known only for the Gaussian source with jointly Gaussian SI in both cases. The former case with forwardly degraded SI is relatively easier to handle for the reason that the syndrome encoded based on the worse SI Y_1 is decodable for the stronger SI Y_2, then the task focuses on refining

the cell partitions. The explicit RD region for jointly Gaussian source given by Tian and Diggavi [28] is presented as follows.

Assume the source is a standard Gaussian $X \sim \mathcal{N}(0, \sigma_x^2)$ and two side informations are $Y_1 = X + N_1 + N_2$, $Y_2 = X + N_2$ where $N_1 \sim \mathcal{N}(0, \sigma_1^2)$, $N_2 \sim \mathcal{N}(0, \sigma_2^2)$ and N_1, N_2 are independent from each other and the source. The achievable RD region is characterized as:

$$R_1 \geq \frac{1}{2} \log \frac{\sigma_x^2(\sigma_1^2 + \sigma_2^2)}{D_1(\sigma_x^2 + \sigma_1^2 + \sigma_2^2)},$$
$$R \geq \frac{1}{2} \log \frac{\sigma_x^2 \sigma_1^2 \sigma_2^2}{D_2(\sigma\_x^2 + \sigma_1^2 + \sigma_2^2)((1-\gamma)^2 D_1 + \gamma \sigma_1^2)},$$

where $\gamma = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$, $0 < D_1 \leq \frac{\sigma_x^2(\sigma_1^2+\sigma_2^2)}{(\sigma_x^2+\sigma_1^2+\sigma_2^2)}$, $0 < D_2 \leq \frac{\sigma_x^2+\sigma_2^2}{(\sigma_x^2+\sigma_2^2)}$ and $D_1 \geq \frac{\gamma\sigma_1^2 D_2}{\gamma\sigma_1^2 - (1-\gamma)^2 D_2}$.

By contrast, the theoretical region of the latter case with reversely degraded SI is derived based on a special binning strategy, i.e. nested binning [29]. This binning structure is depicted in Fig. 2.4. The general idea of nested binning is that if a common partition is used at both decoders, then the binning proceeds in two steps. The cells are first distributed into coarse bins based on the stronger SI Y_1 such that decoder with the stronger SI can identify a cell without ambiguity based on the bin index and Y_1. However, the decoder with the weaker SI may detect multiple codecells with the coarse bin index. As a result, at the second step, each coarse bin is further divided into several finer bins depending on Y_2 such that the decoder with the worse SI can identify a cell with both coarse and fine bin indexes such that both binning can be performed without rate loss.

The achievable RD region with $X \leftrightarrow Y_1 \leftrightarrow Y_2$ is symmetric with the former case. Since the decoder with the weaker SI receives more bit-stream, a smaller distortion at the second decoder is achievable. Thus when $X \leftrightarrow Y_1 \leftrightarrow Y_2$ holds, the theoretical RD region

Figure 2.4: An illustration of the nested binning structure when the Markov chain $X \leftrightarrow Y_1 \leftrightarrow Y_2$ holds. Each coarse bin based on stronger SI Y_1 is further divided into multiple finer bins based on weaker SI Y_2.

is as follows.

$$R = R_{X|Y_2}(D_2) = \frac{1}{2}\log\frac{\sigma_x^2(\sigma_1^2+\sigma_2^2)}{D_2(\sigma_x^2+\sigma_1^2+\sigma_2^2)},$$
$$\text{if } D_1 \le D_1^*, D_2 \le D_2^*, D_1 \ge \frac{\gamma\cdot D_2\cdot\sigma_1^2}{\gamma\cdot\sigma_1^2+D_2\cdot(1-\gamma)^2},$$
$$R = R_{HB}(D_2, D_1) = \frac{1}{2}\log\frac{\sigma_x^2\sigma_1^2\sigma_2^2}{D_1(\sigma_x^2+\sigma_1^2+\sigma_2^2)((1-\gamma)^2D_2+\gamma\sigma_1^2)},$$
$$\text{if } D_1 \le D_1^*, D_2 \le D_2^*, D_1 \le \frac{\gamma\cdot D_2\cdot\sigma_1^2}{\gamma\cdot\sigma_1^2+D_2\cdot(1-\gamma)^2},$$

where $D_1^* = \frac{\sigma_x^2\sigma_1^2}{\sigma_x^2+\sigma_1^2}$, $D_2^* = \frac{\sigma_x^2(\sigma_1^2+\sigma_2^2)}{\sigma_x^2+\sigma_1^2+\sigma_2^2}$, and $\gamma = \frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}$.

2.5 Conclusion

To conclude this chapter, we first reviewed the MWP problem in a WDAG, then presented how to find the optimal scalar quantizer by solving the single-source minimum-weight path problem in the weighted directed acyclic graph, which is the basic and essential technique that will be used in the solutions in following two chapters. Next, we use the graph model to solve the entropy-constrained 2DSQ design problem with convex cells. The theoretical

bounds in Section 2.4 provides the RD guidance for the numerical performance of our solution in Chapter 3. In next chapter, we will present our solution for the two-stage Wyner-Ziv coding problem when the SI is either forwardly or reversely degraded.

Chapter 3

Optimal Design of A Two-stage Wyner-Ziv Scalar Quantizer

3.1 Overview

This chapter provides our solution to the first problem two-stage Wyner-Ziv coding with forwardly/reversely degraded SI, i.e. when $X \leftrightarrow Y_2 \leftrightarrow Y_1$ or $X \leftrightarrow Y_1 \leftrightarrow Y_2$ holds. Both cases will be formulated as minimizing a weighted sum of the rates and distortions. The solutions for both cases are termed the F-WZSQ scheme and the R-WZSQ scheme, respectively. Further, our optimization is based on solving the single-source or all-pairs MWP problems in several WDAGs. Our schemes can guarantee global optimality for finite-alphabet source and SI. The time complexity is $O(N^3)$ if the cardinalities for the source and the SI alphabets are $O(N)$. Although the Monge property does not hold all the time, we exploit empirical conditions to supply the optimization with the partial Monge property, which efficiently reduces the complexity to a inter-point between $O(N^3)$ and

$O(N^2 \log N)$. The effectiveness of both schemes and the partial Monge property are examined by considerable experiments.

This chapter is organized as follows. The following section introduces the notations and the problem formulation. Section 3.3 presents the proposed dynamic programming solution based on the graph model for the optimal R-WZSQ/F-WZSQ design problems. Section 3.4 introduces the partial Monge property and shows how this can be exploited to further reduce the time complexity. Details about the proposed technique, which relies on a modification of an algorithm of Hirschberg and Larmore [13], are given in Section 3.5. Section 3.6 presents extensive experimental results and comparison with the theoretical bounds for a Gaussian source with jointly Gaussian SI. Additionally, the satisfaction of the partial Monge property is empirically investigated. Finally, Section 3.7 concludes this chapter.

3.2 Notations and Problem Formulation

This section starts by presenting notations. Subsection 3.2.2 introduces the R-WZSQ architecture and formulates the problem of optimal R-WZSQ design. The following subsection formulates the problem of optimal F-WZSQ design.

3.2.1 Notations

Let $\mathcal{Y}_1$ and $\mathcal{Y}_2$ denote the alphabets of the SI Y_1 and Y_2, respectively. For discrete random variables A and B, $H(A)$ denotes the entropy of A and $H(A|B)$ denotes the conditional entropy of A given B. For any positive integer k, let $\mathcal{I}_k \triangleq \{1, 2, ..., k\}$. Recall that for any integer $n \geq 2$, an ascending n-sequence is an n-tuple $\mathbf{r} = (r_0, r_1, \cdots, r_{n-1})$, with $r_i \in \bar{\mathcal{X}}$,

for $0 \leq i \leq n-1$, where $r_0 < r_1 < \cdots < r_{n-1}$. For any $x_u, x_v \in \bar{\mathcal{X}}$ with $u < v$, let $\mathcal{T}_{x_u,x_v}$ denote the set of all ascending n-sequences such that $r_0 = x_u$ and $r_{n-1} = x_v$, for all $n \geq 2$.

3.2.2 Optimal R-WZSQ Design Problem

The configuration of the proposed R-WZSQ scheme is inspired by Tian and Diggavi's work [29]. The R-WZSQ is specified by the encoding functions $f_t, t \in \{0,1,2\}$, and decoding functions $g_t, t \in \{1,2\}$, where

$$\begin{aligned} &f_0 : \mathcal{X} \to \mathcal{I}_{M_0}, \quad f_1 : \mathcal{X} \to \mathcal{I}_{M_1}, \quad f_2 : \mathcal{X} \to \mathcal{I}_{M_2}, \\ &g_1 : \mathcal{I}_{M_0} \times \mathcal{I}_{M_1} \times \mathcal{Y}_1 \to \hat{\mathcal{X}}, \quad g_2 : \mathcal{I}_{M_0} \times \mathcal{I}_{M_2} \times \mathcal{Y}_2 \to \hat{\mathcal{X}}, \end{aligned} \tag{3.1}$$

and M_0, M_1, M_2 are positive integers and f_0, f_1, f_2 are surjective. Function f_0 generates the coarse partition, while f_1 and f_2 separately refine the partition f_0. The pair (f_0, f_t) together with g_t forms the quantizer Q_t, for $t = 1, 2$. We will denote by i, respectively j and k, the indexes output by encoder f_0, f_1 and f_2, respectively. We use the notation C_i, $1 \leq i \leq M_0$, for the cells in the coarse partition, i.e., $C_i \triangleq f_0^{-1}(i)$. As shown in Figure 3.1, each C_i is further divided into $M_{1,i}$ and $M_{2,i}$ non-empty sub-cells by the encoding functions f_1 and f_2, respectively, for some $M_{1,i}$, $0 < M_{1,i} \leq M_1$, and some $M_{2,i}$, $0 < M_{2,i} \leq M_1$. Let $C_{ij} \triangleq \{x \in \mathbb{R} | f_0(x) = i \text{ and } f_1(x) = j\}$ and $C'_{ik} \triangleq \{x \in \mathbb{R} | f_0(x) = i \text{ and } f_2(x) = k\}$, for $i \in \mathcal{I}_{M_0}, j \in \mathcal{I}_{M_{1,i}}, k \in \mathcal{I}_{M_{2,i}}$.

We will assume that the cells in each partition, i.e., cells C_i, C_{ij} and C'_{ik} are contiguous. It follows that there is a unique ascending (M_0+1)-sequence $\mathbf{r} \in \mathcal{T}_{x_0,x_N}$ such that $C_i = (r_{i-1}, r_i]$ for $1 \leq i \leq M_0$. Thus, the partition generated by f_0 is completely specified by the sequence $\mathbf{r}$. Likewise, for each $1 \leq i \leq M_0$, the partition of C_i into cells C_{ij} is specified by the ascending $(M_{1,i}+1)$-sequence $\mathbf{s}_i \triangleq (s_{i,0}, s_{i,1}, \cdots, s_{i,M_{1,i}}) \in \mathcal{T}_{r_{i-1},r_i}$, satisfying $C_{ij} =$

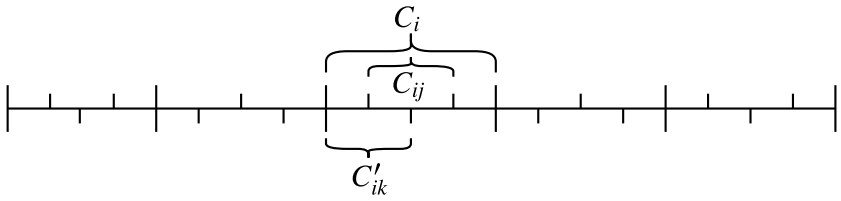


Figure 3.1: Illustration of the three partitions f_0, f_1 and f_2.

$(s_{i,j-1}, s_{i,j}]$ for $1 \leq j \leq M_{1,i}$. Similarly, for each $1 \leq i \leq M_0$, the partition of C_i into cells C'_{ik} is specified by the ascending $(M_{2,i}+1)$-sequence $\mathbf{t}_i \triangleq (t_{i,0}, t_{i,1}, \cdots, t_{i,M_{2,i}}) \in \mathcal{T}_{r_{i-1},r_i}$, where $C'_{ik} = (t_{i,k-1}, t_{i,k}]$ for $1 \leq k \leq M_{2,i}$. Further, let us denote by $\bar{\mathbf{s}}$ the M_0-tuple $(\mathbf{s}_1, \cdots, \mathbf{s}_{M_0})$, and by $\bar{\mathbf{t}}$ the M_0-tuple $(\mathbf{t}_1, \cdots, \mathbf{t}_{M_0})$.

Let I, J and K denote the random variables representing the outputs of f_0, f_1 and f_2, respectively. Decoder g_1 uses I, J and Y_1 to reconstruct the source, while decoder g_2 uses I, K and Y_2 for the source reconstruction. We will assume that the reconstruction at each decoder is optimal, i.e., it minimizes the distortion. Then the decoding functions are defined as follows

$$g_1(i, j, y_1) \triangleq \hat{x}_1(C_{ij}|y_1), \quad g_2(i, k, y_2) \triangleq \hat{x}_2(C'_{ik}|y_2),$$

for $1 \leq i \leq M_0$, $1 \leq j \leq M_{1,i}$, $1 \leq k \leq M_{2,i}$, $y_1 \in \mathcal{Y}_1$ and $y_2 \in \mathcal{Y}_2$, where $\hat{x}_\kappa(C|y_\kappa)$ is defined for any set $C \subseteq \mathcal{X}$ and any $y_\kappa \in \mathcal{Y}_\kappa$, $\kappa \in \{1, 2\}$ as

$$\hat{x}_\kappa(C|y_\kappa) \triangleq \arg\min_{\hat{x} \in \hat{\mathcal{X}}} \mathbb{E}[d(X, \hat{x})|X \in C, Y_\kappa = y_\kappa].$$

Since the decoders are determined given the encoders, it follows that the coding scheme is fully specified by the triple of encoding functions (f_0, f_1, f_2), which we denote by $\mathbf{f}$.

The total message to be transmitted to the two decoders can be split into four parts

$\mathcal{M}_{0,1}$, $\mathcal{M}_1$, $\mathcal{M}_{0,2}$, and $\mathcal{M}_2$. Message $\mathcal{M}_{0,1}$ represents the information needed by decoder 1 to recover the index I with the help of the SI Y_1, while $\mathcal{M}_1$ is the additional information needed at decoder 1 to recover index J based on I and Y_1. Further, $\mathcal{M}_{0,2}$ denotes the message needed at decoder 2 in order to recover the index I with the help of $\mathcal{M}_{0,1}$ and the SI Y_2. Finally, $\mathcal{M}_2$ is the information needed at decoder 2 to recover the index K based on the index I and Y_2. The aforementioned messages are obtained by using SW (short for Slepian-Wolf) coding. Specifically, we assume that the SW coding is performed on blocks of length approaching ∞, so that the limits in the SW theorem [26] are achieved. Thus, the rates of the messages $\mathcal{M}_{0,1}$, $\mathcal{M}_1$, $\mathcal{M}_{0,2}$, and $\mathcal{M}_2$ are $H(I|Y_1)$, $H(J|I,Y_1)$, $H(I|Y_2)-H(I|Y_1)$ and $H(K|I,Y_2)$, respectively. Note that, since the Markov chain $X \leftrightarrow Y_1 \leftrightarrow Y_2$ holds, the aforementioned rates for $\mathcal{M}_{0,1}$ and $\mathcal{M}_{0,2}$ can be achieved by using nested binning, where $\mathcal{M}_{0,1}$ is the index of the coarse bin, while $\mathcal{M}_{0,2}$ is the index of the fine bin inside the coarse bin [29].

Let us denote by $R_1(\mathbf{f})$ the rate for the portion of the message needed by decoder 1 and by $R_2(\mathbf{f})$ the rate for the message portion that only decoder 2 will use. In other words, $R_1(\mathbf{f})$ is the rate for $\mathcal{M}_{0,1}$ and $\mathcal{M}_1$, while $R_2(\mathbf{f})$ is the rate for $\mathcal{M}_{0,2}$ and $\mathcal{M}_2$. Additionally, let $R(\mathbf{f}) \triangleq R_1(\mathbf{f}) + R_2(\mathbf{f})$. Finally, for $\kappa = 1, 2$, let $D_\kappa(\mathbf{f})$ denote the distortion at decoder κ.

We conclude that the RD performance of an R-WZSQ can be characterized by the quadruple $\big(R_1(\mathbf{f}), R(\mathbf{f}), D_1(\mathbf{f}), D_2(\mathbf{f})\big)$. The optimum such quadruple is not unique, rather any such quadruple (we will call them RD quadruples) situated on the lower boundary of the convex hull of the set of all possible RD quadruples is optimal in some sense. Any RD quadruple on the lower convex hull can be obtained by minimizing a weighted sum of the distortions and rates with positive weights [16]. Clearly, if the weights are normalized so

that the weights of the distortion terms add up to 1, the result of the minimization remains the same. Therefore, we will consider as our cost function the following

$$\mathcal{O}(\mathbf{r}, \bar{\mathbf{s}}, \bar{\mathbf{t}}) \triangleq \rho D_1(\mathbf{f}) + (1-\rho) D_2(\mathbf{f}) + \lambda_1 R_1(\mathbf{f}) + \lambda_2 R(\mathbf{f}), \tag{3.2}$$

for some $0 < \rho < 1$ and $\lambda_1, \lambda_2 > 0$. Further, we formulate the problem of optimal R-WZSQ design as follows

$$\min_{\mathbf{r}, \bar{\mathbf{s}}, \bar{\mathbf{t}}} \quad \mathcal{O}(\mathbf{r}, \bar{\mathbf{s}}, \bar{\mathbf{t}}). \tag{3.3}$$

Note that the weights $\rho, 1-\rho, \lambda_1, \lambda_2$ in (3.2) could be interpreted as the priorities that code designers place on the minimization of $D_1(\mathbf{f}), D_2(\mathbf{f}), R_1(\mathbf{f}), R(\mathbf{f})$, respectively. We emphasize that the approach of formulating the optimal design problem as the problem of minimizing a weighted sum of distortions and rates was also adopted in [3, 9, 18, 23].

3.2.3 Optimal F-WZSQ Design Problem

In the case of F-WZSQ the encoders generate only two partitions, a coarse partition to be used at decoder 1 and a fine partition, to be used at decoder 2. Thus, the difference versus the coding scheme in (3.1) is that the encoding function f_1 disappears, or equivalently, $M_1 = 1$. Additionally, out of the four parts constituting the total message to be transmitted to the decoders, only two remain, namely $\mathcal{M}_{0,1}$ and $\mathcal{M}_2$. Message $\mathcal{M}_{0,1}$ is needed at decoder 1 in order to recover index I based on the SI Y_1. Thus it can be transmitted at a rate equal to $H(I|Y_1)$. Since SI Y_2 is stronger than Y_1, the second decoder is able to recover I as well from $\mathcal{M}_{0,1}$. Additionally, the second decoder uses $\mathcal{M}_2$ to recover the refinement index K based on I and Y_2. Therefore, the rate for $\mathcal{M}_2$ equals $H(K|I, Y_2)$. In other words,

$R_1(\mathbf{f}) = H(I|Y_1)$, while $R_2(\mathbf{f}) = H(K|I, Y_2)$.

The cost function is also defined as in (3.2), but is only a function of $\mathbf{r}$ and $\bar{\mathbf{t}}$, i.e.,

$$\mathcal{O}(\mathbf{r}, \bar{\mathbf{t}}) \triangleq \rho D_1(\mathbf{f}) + (1-\rho) D_2(\mathbf{f}) + (\lambda_1 + \lambda_2) R_1(\mathbf{f}) + \lambda_2 R_2(\mathbf{f}). \tag{3.4}$$

The optimization problem is formulated as

$$\min_{\mathbf{r}, \bar{\mathbf{t}}} \quad \mathcal{O}(\mathbf{r}, \bar{\mathbf{t}}). \tag{3.5}$$

For the traditional HB problem with degraded SI, it was shown in [12] that the optimal scheme consists of a two-layer code such that the bad user can decode using the first layer and the good user can decode with both layers, i.e., the scheme for the F-WZ problem. The situations when either Y_1 or Y_2 is stronger are symmetric. When the Markov chain $X \leftrightarrow Y_2 \leftrightarrow Y_1$ holds the SQ-based scheme for the HB problem consists of the same encoding and decoding functions as for F-WZSQ. The distortions $D_\kappa(\mathbf{f})$ and rates $R(\mathbf{f})$, $R_\kappa(\mathbf{f})$, $\kappa = 1, 2$ are defined in the same way. The only difference versus F-WZSQ is that the two messages $\mathcal{M}_{0,1}$ and $\mathcal{M}_2$ are not split between two stages. Thus, the cost function is as in (3.4), but with $\lambda_1 = 0$.

3.3 Solution Algorithm

In this section we present the proposed solution algorithms based on the MWP model. We first describe the solution to the optimal R-WZSQ design problem in subsection 3.3.1. The following subsection presents the preprocessing step whose aim is to make possible the computation of each edge weight in constant time. Then we present the solution to the

optimal F-WZSQ design problem in subsection 3.3.3.

3.3.1 Solution to the Optimal R-WZSQ Design Problem

For $C \subseteq \mathcal{X}$, $y_\kappa \in \mathcal{Y}_\kappa$, $\kappa = 1, 2$, denote

$$\begin{aligned} P_\kappa(C, y_\kappa) &\triangleq \mathbb{P}[X \in C, Y_\kappa = y_\kappa], \\ v_{1,\kappa}(C) &\triangleq \sum_{y_\kappa \in \mathcal{Y}_\kappa} P_\kappa(C, y_\kappa) \mathbb{E}[d(X, \hat{x}_\kappa(C|y_\kappa))|X \in C, Y_\kappa = y_\kappa], \\ v_{2,\kappa}(C) &\triangleq -\sum_{y_\kappa \in \mathcal{Y}_\kappa} P_\kappa(C, y_\kappa) \log_2(P_\kappa(C, y_\kappa)). \end{aligned}$$

Since $D_1(\mathbf{f}) = E[d(X, g_1(I, J, Y_1))]$ and $D_2(\mathbf{f}) = E[d(X, g_2(I, K, Y_2))]$ we obtain

$$D_1(\mathbf{f}) = \sum_{i=1}^{M_0} \sum_{j=1}^{M_{1,i}} v_{1,1}(C_{ij}), \quad D_2(\mathbf{f}) = \sum_{i=1}^{M_0} \sum_{k=1}^{M_{2,i}} v_{1,2}(C'_{ik}). \tag{3.6}$$

The rates $R_1(\mathbf{f})$ and $R_2(\mathbf{f})$ can be written as follows

$$\begin{aligned} R_1(\mathbf{f}) &= H(I|Y_1) + H(J|I, Y_1) = H(I, J|Y_1) \qquad (3.7) \\ &= H(J, I, Y_1) - H(Y_1) = \sum_{i=1}^{M_0} \sum_{j=1}^{M_{1,i}} v_{2,1}(C_{ij}) - H(Y_1), \\ R_2(\mathbf{f}) &= H(I|Y_2) - H(I|Y_1) + H(K|I, Y_2) = H(I, K|Y_2) - H(I|Y_1) \qquad (3.8) \\ &= H(K, I, Y_2) - H(Y_2) - H(I, Y_1) + H(Y_1) \\ &= \sum_{i=1}^{M_0} \sum_{k=1}^{M_{2,i}} v_{2,2}(C'_{ik}) - H(Y_2) - \sum_{i=1}^{M_0} v_{2,1}(C_i) + H(Y_1). \end{aligned}$$

By plugging (3.6)-(3.8) in (3.2) we obtain

$$\mathcal{O}(\mathbf{r},\bar{\mathbf{s}},\bar{\mathbf{t}}) = \rho\sum_{i=1}^{M_0}\sum_{j=1}^{M_{1,i}} v_{1,1}(C_{ij}) + (1-\rho)\sum_{i=1}^{M_0}\sum_{k=1}^{M_{2,i}} v_{1,2}(C'_{ik}) + (\lambda_1+\lambda_2)\sum_{i=1}^{M_0}\sum_{j=1}^{M_{1,i}} v_{2,1}(C_{ij}) - (\lambda_1+\lambda_2)H(Y_1) + \lambda_2\left(\sum_{i=1}^{M_0}\sum_{k=1}^{M_{2,i}} v_{2,2}(C'_{ik}) - \sum_{i=1}^{M_0} v_{2,1}(C_i)\right) + \lambda_2(H(Y_1)-H(Y_2)).$$

Since the quantity $-\lambda_1 H(Y_1) - \lambda_2 H(Y_2)$ is a constant, it can be subtracted from the objective function $\mathcal{O}(\mathbf{r},\bar{\mathbf{s}},\bar{\mathbf{t}})$. After doing so and rearranging the terms the new cost becomes

$$\mathcal{O}'(\mathbf{r},\bar{\mathbf{s}},\bar{\mathbf{t}}) = \sum_{i=1}^{M_0}\Bigg(-\lambda_2 v_{2,1}(C_i) + \underbrace{\sum_{j=1}^{M_{1,i}}\Big(\rho v_{1,1}(C_{ij}) + (\lambda_1+\lambda_2)v_{2,1}(C_{ij})\Big)}_{w_1(C_i,\mathbf{s}_i)} + \underbrace{\sum_{k=1}^{M_{2,i}}\Big((1-\rho)v_{1,2}(C'_{ik}) + \lambda_2 v_{2,2}(C'_{ik})\Big)}_{w_2(C_i,\mathbf{t}_i)}\Bigg). \tag{3.9}$$

We notice from (3.9) that, if C_i is fixed, then the partition $\mathbf{s}_i$ of C_i can be optimized by minimizing the subcost $w_1(C_i,\mathbf{s}_i)$. Likewise, the partition $\mathbf{t}_i$ can be optimized by minimizing $w_2(C_i,\mathbf{t}_i)$. Therefore, for each $x_u, x_v \in \bar{\mathcal{X}}$, with $u < v$, let $\mathbf{s}^*(x_u,x_v)$ and $\mathbf{t}^*(x_u,x_v)$ denote the corresponding optimal partitions of C_i if $C_i = (x_u, x_v]$, i.e.,

$$\mathbf{s}^*(x_u,x_v) \triangleq \arg\min_{\mathbf{s}\in\mathcal{T}_{x_u,x_v}} w_1((x_u,x_v],\mathbf{s}), \tag{3.10}$$

$$\mathbf{t}^*(x_u,x_v) \triangleq \arg\min_{\mathbf{t}\in\mathcal{T}_{x_u,x_v}} w_2((x_u,x_v],\mathbf{t}). \tag{3.11}$$

where $\mathcal{T}_{x_u,x_v}$ was defined in Subsection 3.2.1. Further, for each $x_u, x_v \in \bar{\mathcal{X}}$, with $u < v$, denote

$$w_0(u,v) \triangleq -\lambda_2 v_{2,1}((x_u, x_v]) + w_1((x_u, x_v], \mathbf{s}^*(x_u, x_v)) + w_2((x_u, x_v], \mathbf{t}^*(x_u, x_v)). \quad (3.12)$$

It follows that, if the optimal partitions $\mathbf{s}^*(r_{i-1}, r_i)$ and $\mathbf{t}^*(r_{i-1}, r_i)$ are known for each possible pair (r_{i-1}, r_i) (i.e., for each possible coarse cell C_i), then problem (3.3) reduces to solving the following

$$\min_{M_0, \mathbf{r}} \quad \bar{\mathcal{O}}(\mathbf{r}) \triangleq \sum_{i=1}^{M_0} w_0(a_{i-1}, a_i). \quad (3.13)$$

where $a_i \in V$ such that $x_{a_i} = r_i$ for $0 \leq i \leq M_0$.

The above discussion suggests the following strategy to solve problem (3.3).

- 1) Determine $\mathbf{s}^*(x_u, x_v)$ for all pairs x_u, x_v of elements in $\bar{\mathcal{X}}$, with $u < v$.
- 2) Determine $\mathbf{t}^*(x_u, x_v)$ for all pairs x_u, x_v of elements in $\bar{\mathcal{X}}$, with $u < v$.
- 3) Solve problem (3.13).

Next we will discuss how to solve the problem at each step. The key idea is to model each component problem as an MWP in a WDAG based on the DAG G. Note that any contiguous cell $(x_m, x_n]$ can be associated to the edge (m, n) in the DAG G. Then any partition of some cell $(x_u, x_v]$ into contiguous cells can be regarded as a path in G between the vertices u and v. Furthermore, the cost of the partition can be written as the sum of the costs of the individual cells. Thus, if we define the weight of an edge as the cost of the associated cell, then the cost of the partition becomes equal to the cost of the associated path.

Specifically, consider the partition $\mathbf{s} = (s_0, \cdots, s_M)$ of $(x_u, x_v]$ into M cells, for some

$M > 0$, i.e., $\mathbf{s} \in \mathcal{T}_{x_u,x_v}$. For each $j, 0 \leq j \leq M$, let $q_j \in V$ such that $s_j = x_{q_j}$. Then the sequence $\mathbf{q} = (q_0, \cdots, q_M)$ is an M-edge path from node u to node v in G. For each $i, 1 \leq i \leq M$, the ith edge on this path, namely (q_{i-1}, q_i), corresponds to the ith cell in the partition, namely $(s_{i-1}, s_i]$. Consider now the weight function ω_1 defined as follows.

$$\omega_1(m,n) \triangleq \rho v_{1,1}((x_m, x_n]) + (\lambda_1 + \lambda_2) v_{2,1}((x_m, x_n]). \tag{3.14}$$

Then the cost of the partition $\mathbf{s}$ is equal to the weight of the associated path $\mathbf{q}$ in the WDAG $G(\omega_1)$, i.e., $w_1((x_u, x_v], \mathbf{s}) = \sum_{j=1}^{M} \omega_1(q_{j-1}, q_j)$. Clearly, the aforementioned correspondence between contiguous-cell partitions of $(x_u, x_v]$ and paths from u to v in $G(\omega_1)$ is one-to-one. Therefore, solving problem (3.10), i.e., finding the optimal partition $\mathbf{s}^*(x_u, x_v)$, is equivalent to finding the MWP between the nodes u and v in $G(\omega_1)$. Since in Step 1 we need to find $\mathbf{s}^*(x_u, x_v)$ for all pairs $(u, v) \in E$, it follows that the problem of Step 1 is equivalent to the all-pairs MWP problem in $G(\omega_1)$, which can be solved in $O(N^3)$ time if each edge weight can be evaluated in $O(1)$ time.

Similarly, the problem at Step 2 is equivalent to the all-pairs MWP problem in $G(\omega_2)$, where

$$\omega_2(m,n) \triangleq (1-\rho) v_{1,2}((x_m, x_n]) + \lambda_2 v_{2,2}((x_m, x_n]), \tag{3.15}$$

for each $(m, n) \in E$. Thus, the problem at Step 2 problem can also be solved in $O(N^3)$ time if each edge weight can be evaluated in $O(1)$ time.

Finally, problem (3.13) can be modeled as the MWP problem in the WDAG $G(w_0)$, where the source node is 0, the final node is N and the weighting function w_0 is defined in (3.12). After having solved the problems at Steps 1 and 2 each weight $w_0(u, v)$ can be

determined in constant time and problem (3.13) can be solved in $O(N^2)$ operations.

3.3.2 Preprocessing Step

To make sure that each quantity $\omega_1(m,n)$ and $\omega_2(m,n)$ can be computed in constant time, we include a preprocessing step which evaluates and stores all values $v_{1,\kappa}((x_m,x_n])$ and $v_{2,\kappa}((x_m,x_n])$ for $\kappa=1,2$ and all (m,n) with $0\leq m<n\leq N$. In order to compute the values $v_{2,\kappa}((x_m,x_n])$, we first evaluate for each $\kappa=1,2$ and $y_\kappa\in\mathcal{Y}_\kappa$, the cumulative probabilities $CumP(y_\kappa,n)\triangleq\mathbb{P}[X\in(x_0,x_n],Y_\kappa=y_\kappa]$. This process requires $O(N(|\mathcal{Y}_1|+|\mathcal{Y}_2|))$ time. Then each $v_{2,\kappa}((x_m,x_n])$ is calculated by first computing $P_\kappa((x_m,x_n],y_\kappa)=CumP(y_\kappa,n)-CumP(y_\kappa,m)$ and then performing the summation over y_κ. It follows that the computation of all values $v_{2,\kappa}((x_m,x_n])$ for $\kappa=1,2$ and $(x_m,x_n)\in\bar{\mathcal{X}}\times\bar{\mathcal{X}}$, takes $O(N^2(|\mathcal{Y}_1|+|\mathcal{Y}_2|))$ time. The amount of memory needed to store all these values is clearly $O(N^2)$.

To explain how the quantities $v_{1,\kappa}((x_m,x_n])$ are evaluated, first denote for each (m,n) as above, each $\kappa=1,2$, and each $y_\kappa\in\mathcal{Y}_\kappa$,

$$\gamma_\kappa(m,n,y_\kappa)\triangleq P_\kappa(C,y_\kappa)\mathbb{E}[d(X,\hat{x}_\kappa(C|y_\kappa))|X\in C,Y_\kappa=y_\kappa],$$

where $C=(x_m,x_n]$. Then the following holds

$$v_{1,\kappa}((x_m,x_n])=\sum_{y_\kappa\in\mathcal{Y}_\kappa}\gamma_\kappa(m,n,y_\kappa).$$

Now let us consider the case when the distortion measure is not the squared distance. Recall that in this case $\hat{\mathcal{X}}$ is finite. As shown in [39], since the distortion measure is monotone, for fixed κ and y_κ, all values $\gamma_\kappa(m,n,y_\kappa)$ can be computed in $O(|\bar{\mathcal{X}}||\bar{\mathcal{X}}\cup\hat{\mathcal{X}}|)=O(N^2)$

operations. A simpler technique with the same time complexity was proposed in [5]. It follows that all values $v_{1,\kappa}((x_m, x_n])$ for $\kappa = 1, 2$ and $(x_m, x_n) \in \bar{\mathcal{X}} \times \bar{\mathcal{X}}$, can be evaluated with $O(N^2(|\mathcal{Y}_1| + |\mathcal{Y}_2|))$ time complexity.

When the distortion measure is the squared distance we have $\hat{\mathcal{X}} = \mathbb{R}$. Then the following relations hold

$$\hat{x}_\kappa(C|y_\kappa) = \mathbb{E}[X|X \in C, Y_\kappa = y_\kappa],$$

$$\mathbb{E}[d(X, \hat{x}_\kappa(C|y_\kappa)|X \in C, Y_\kappa = y_\kappa] = \mathbb{E}[X^2|X \in C, Y_\kappa = y_\kappa] - \left(\hat{x}_\kappa(C|y_\kappa)\right)^2.$$

We first compute and store the cumulative first and second moments $Cum_i(y_\kappa, n) \triangleq \sum_{x \le x_n} x^i p_{XY_\kappa}(x, y_\kappa)$ for $i = 1, 2$. Their computation takes $O(N(|\mathcal{Y}_1| + |\mathcal{Y}_2|))$ time. Based on these values, each $\gamma_\kappa(m, n, y_\kappa)$ can be computed in constant time. Thus, the evaluation of all $v_{1,\kappa}((x_m, x_n])$ takes $O(N^2(|\mathcal{Y}_1| + |\mathcal{Y}_2|))$ operations.

To summarize, the time complexity of the preprocessing step amounts to $O(N^2(|\mathcal{Y}_1| + |\mathcal{Y}_2|)) = O(N^3)$ according to our assumption that $|\mathcal{Y}_1| + |\mathcal{Y}_2| = O(N)$. Thus, the inclusion of this step does not change the asymptotical time complexity of $O(N^3)$ for the solution algorithm.

3.3.3 F-WZSQ Design Algorithm

In the F-WZSQ case $D_2(\mathbf{f})$ remains as in (3.6), while $D_1(\mathbf{f})$ becomes

$$D_1(\mathbf{f}) = \sum_{i=1}^{M_0} v_{1,1}(C_i).$$

Additionally, we have

$$R_1(\mathbf{f}) = H(I|Y_1) = H(I, Y_1) - H(Y_1) = \sum_{i=1}^{M_0} v_{2,1}(C_i) - H(Y_1),$$
$$R_2(\mathbf{f}) = H(K|I, Y_2) = H(I, K, Y_2) - H(I, Y_2) = \sum_{i=1}^{M_0}\sum_{k=1}^{M_{2,i}} v_{2,2}(C'_{ik}) - \sum_{i=1}^{M_0} v_{2,2}(C_i).$$

The cost function $\mathcal{O}(\mathbf{r}, \bar{\mathbf{t}})$ is given by

$$\begin{aligned}\mathcal{O}(\mathbf{r}, \bar{\mathbf{t}}) = \rho\sum_{i=1}^{M_0} v_{1,1}(C_i) + (1-\rho)\sum_{i=1}^{M_0}\sum_{k=1}^{M_{2,i}} v_{1,2}(C'_{ik}) + (\lambda_1+\lambda_2)\sum_{i=1}^{M_0} v_{2,1}(C_i) \\ - (\lambda_1+\lambda_2)H(Y_1) + \lambda_2\left(\sum_{i=1}^{M_0}\sum_{k=1}^{M_{2,i}} v_{2,2}(C'_{ik}) - \sum_{i=1}^{M_0} v_{2,2}(C_i)\right).\end{aligned}$$

After removing the constant term $-(\lambda_1+\lambda_2)H(Y_1)$ and rearranging the remaining terms, the cost becomes

$$\begin{aligned}\mathcal{O}'(\mathbf{r}, \bar{\mathbf{t}}) = \sum_{i=1}^{M_0}\Bigg(\rho v_{1,1}(C_i) + (\lambda_1+\lambda_2)v_{2,1}(C_i) - \lambda_2 v_{2,2}(C_i) \\ + \underbrace{\sum_{k=1}^{M_{2,i}}\Big((1-\rho)v_{1,2}(C'_{ik}) + \lambda_2 v_{2,2}(C'_{ik})\Big)}_{w_2(C_i, \mathbf{t}_i)}\Bigg).\end{aligned}$$

Notice that the quantity $w_2(C_i, \mathbf{t}_i)$ is the same as for R-WZSQ. Thus, the optimal partition $\mathbf{t}_i$, for a given C_i, can be found as in the previous section. Thus, problem (3.5) reduces to solving (3.13) with $w_0(u, v)$, for $u < v$ defined as follows

$$\begin{aligned}w_0(u, v) \triangleq \rho v_{1,1}((x_u, x_u]) + (\lambda_1+\lambda_2)v_{2,1}((x_u, x_v]) - \lambda_2 v_{2,2}((x_u, x_v]) \\ + w_2((x_u, x_v], \mathbf{t}^*(x_u, x_v]).\end{aligned} \tag{3.16}$$

In conclusion, problem (3.5) can be solved using the following two steps.

- 1) Determine $\mathbf{t}^*(x_u, x_v)$ for all pairs x_u, x_v of elements in $\mathcal{X}$, with $u < v$.
- 2) Solve problem (3.13) with the definition of w_0 given in (3.16).

The problem at Step 1 is equivalent to the all-pairs MWP problem in $G(\omega_2)$, while the problem at Step 2 is equivalent to the MWP problem in $G(w_0)$. Thus, using conventional algorithms for the aforementioned MWP problems, the time complexity of the solution becomes $O(N^3)$.

3.4 Time Complexity Reduction Using the Partial Monge Property

Notice that the most computationally demanding parts in the solutions to the optimal R-WZSQ and F-WZSQ design problems is solving the all-pairs MWP problem in $G(\omega_1)$ and $G(\omega_2)$, requiring $O(N^3)$ operations. In this section we introduce the partial Monge property and propose a method for reducing the time complexity for this task when the weighting functions ω_1 and ω_2 satisfy it.

If the weight function ω satisfies the Monge property [1] the dynamic programming solution to the single-source MWP problem in $G(\omega)$ can be accelerated by a factor of $N/\log N$[13] or of N [37], thus leading to the acceleration by the same factor of the all-pairs MWP solution algorithm. The general idea behind this complexity reduction is the following. The dynamic programming single-source MWP problem algorithm needs to examine each graph edge in order to determine if that edge is part of an optimal path or not. If all edge weights satisfy the Monge property, after examining a single edge a conclusion

can be drawn about a higher number of edges. Thus, the set of edges which need to be examined is significantly decreased.

Unfortunately, the Monge condition is not fulfilled by our weight functions ω_1 and ω_2. However, we have observed empirically that the Monge property may hold for a structured subset E' of edges. We prove that this partial satisfaction of the Monge property can still be exploited to decrease the running time of the all-pairs MWP algorithm. The basic idea is to exploit the partial Monge property to reduce the number of edges from E' which have to be examined. This idea is used in conjunction with a simple test to determine another set E'' of edges which cannot be in any optimal path and thus need not be examined either. Note that determining each of the sets E' and E'' requires a scan through the whole set of edges E, i.e., $O(N^2)$ operations. Thus, this technique cannot expedite the single-source MWP solution algorithm. However, as we will show shortly, it can effectively speed up the algorithm for the all-pairs MWP problem.

Let V and E be defined as in Subsection 2.2.1.

Definition 1. *We say that the real-valued weight function $\omega : E \to \mathbb{R}$ satisfies the Monge property [1][1] if for all $0 \leq m \leq m' < n \leq n' \leq N$ the following holds*

$$\omega(m, n) + \omega(m', n') \leq \omega(m, n') + \omega(m', n).$$

As pointed out in [1] the Monge property can be extended to weight functions taking values in $\mathbb{R} \cup \{\infty\}$. In this case the addition operation and the order $\leq$ are extended to $\mathbb{R} \cup \{\infty\}$ in a natural way by requiring that $a + \infty = \infty$ for all $a \in \mathbb{R} \cup \{\infty\}$, and $a < \infty$ for all $a \in \mathbb{R}$.

[1]This property has received various denominations in the literature. For instance, the authors of [13], work which we rely upon in this section, refer to this as the concavity property. We prefer to use here the term "Monge property", which has been more widely adopted in the newer literature [1].

Further, for any real-valued weight function $\omega : E \to \mathbb{R}$ denote

$$\Delta_\omega(m,n) \triangleq \omega(m,n+1) + \omega(m+1,n) - \omega(m,n) - \omega(m+1,n+1),$$

for all $0 \le m < n-1 \le N-2$.

Definition 2. *For any real-valued weight function $\omega : E \to \mathbb{R}$, let $(T_1(\omega), T_2(\omega))$ denote the pair of integers (T_1, T_2), $2 \le T_1 \le T_2 \le N$, with maximum $T_2 - T_1$ for which the following holds*

$$\Delta_\omega(m,n) \ge 0 \textit{ for all } 0 \le m,n \le N-1, T_1 \le n-m \le T_2. \tag{3.17}$$

If more such pairs exist, the one with the smallest T_1 is chosen.

Definition 3. *We say that the real-valued weight function $\omega : E \to \mathbb{R}$ satisfies the partial Monge property if $T_1(\omega) < T_2(\omega)$.*

Remark 1. *It is easy to see that the pair $(T_1(\omega), T_2(\omega))$ can be determined in one pass through the edge set E in $O(N^2)$ time.*

Definition 4. *For any a real-valued weight function $\omega : E \to \mathbb{R}$, let $T_3(\omega)$ be the smallest positive integer T_3, smaller than N, satisfying*

$$\omega(m,n) \ge \omega\left(m, \left\lfloor \frac{m+n}{2} \right\rfloor\right) + \omega\left(\left\lceil \frac{m+n}{2} \right\rceil, n\right) \tag{3.18}$$

for all $0 \le m < n \le N, n-m \ge T_3$. If such an integer does not exist we set $T_3(\omega) = N$.

Notice that (3.18) implies that the edge (m,n) can be replaced in any path by other two edges, without increasing the weight of the path. Therefore, we can safely remove all edges

(m, n) with $n - m \geq T_3(\omega)$ when calculating the all-pairs MWPs in $G(\omega)$. Note that the value $T_3(\omega)$ can also be determined in one scan through the edge set E in $O(N^2)$ time.

Consider the single-source MWP problem in $G(\omega)$ with node 0 as the source node. Recall that, for each $0 \leq n \leq N$, $\hat{W}_0(n)$ denotes the weight of the MWP from node 0 to node n in the WDAG $G(\omega)$. Further define $E' \triangleq \{(m, n) \in E | T_1(\omega) - 1 \leq n - m \leq T_2(\omega) + 1\}$ and $E'' \triangleq \{(m, n) \in E | n - m \geq T_3(\omega)\}$. Relation (2.1) and the discussion below equation (3.18) imply that

$$\hat{W}_0(n) = \min(\hat{W}'(n), \hat{W}''(n)), \tag{3.19}$$

where

$$\hat{W}'(n) \triangleq \min_{(m,n)\in E'} (\hat{W}_0(m) + \omega(m, n)), \tag{3.20}$$

$$\hat{W}''(n) \triangleq \min_{(m,n)\in E\setminus(E'\cup E'')} (\hat{W}_0(m) + \omega(m, n)). \tag{3.21}$$

Consider now the weight function $\omega' : E \to \mathbb{R} \cup \{\infty\}$, where $\omega'(m, n) = \omega(m, n)$ if $(m, n) \in E$, and $\omega'(m, n) = \infty$ otherwise. The following result, which is proved in the appendix, is essential for our development.

Proposition 1. *The weight function ω' satisfies the Monge property.*

Further, note that equation (3.20) implies that

$$\hat{W}'(n) \triangleq \min_{0\leq m<n} (\hat{W}_0(m) + \omega'(m, n)). \tag{3.22}$$

We will achieve the complexity reduction by exploiting the Monge property of ω' to expedite the computations in (3.22). For this we will use a modification of the basic algorithm

of Hirschberg and Larmore [13] for solving the single-source MWP problem in a WDAG with Monge weights. More specifically, the algorithm of [13] determines all values $F(n)$, for $1 \leq n \leq N$, where

$$F(n) \triangleq \min_{0 \leq m < n} (F(m) + \omega'(m, n)), \tag{3.23}$$

where $F(0) = 0$ and the weights $\omega'(m, n)$, which are given, satisfy the Monge property. Consider now the upper triangular matrix $\mathcal{G}$, with elements $g(m, n)$, $0 \leq m < n \leq N$, defined as

$$g(m, n) \triangleq F(m) + \omega'(m, n). \tag{3.24}$$

Then the problem of solving (3.23) for all n can be regarded as the problem of finding the minimum element on each column in the upper triangular matrix $\mathcal{G}$, i.e., finding, for $1 \leq n \leq N$,

$$F(n) = \min_{0 \leq m < n} g(m, n). \tag{3.25}$$

The fact that the weights $\omega'(m, n)$ satisfy the Monge property implies that the values $g(m, n)$ also satisfy this property, fact which is straightforward to verify. The authors of [13] exploit the Monge property of the function g to reduce the time complexity from $O(N^2)$ to $O(N \log N)$. Their Basic Algorithm iterates over m from 1 to $N - 1$. For each m, at the end of the $(m - 1)$th iteration the value of $F(m)$ is computed. The algorithm is based on comparing elements of the matrix. Note that, while the weights ω' are all available at the beginning, the matrix elements are not. Specifically, an element $g(m, n)$ can be accessed only after the $(m - 1)$th iteration, i.e., after $F(m)$ was computed. We will refer to the Basic Algorithm of [13] as algorithm $\mathcal{A}$.

Now consider a modification of problem (3.23) as follows

$$F(n) \triangleq \min_{0 \leq m < n} (L(m) + \omega'(m, n)), \tag{3.26}$$

where $L(1) = 0$ and $L(m)$ is computed based on $F(m)$, for each $1 \leq m \leq n-1$ according to a specified procedure. Further, modify the definition of $g(m, n)$ in (3.24) as follows

$$g(m, n) \triangleq L(m) + \omega'(m, n), \tag{3.27}$$

for $0 \leq m < n \leq N$. Then problem (3.26) remains equivalent to problem (3.25) of finding all column minima in the modified matrix $\mathcal{G}$. Relation (3.27) implies that the elements $g(m, n)$ of the modified upper triangular matrix $\mathcal{G}$ still satisfy the Monge property Then problem (3.25) can be solved by using algorithm $\mathcal{A}$ enhanced with a procedure which evaluates $L(m)$ based on $F(m)$, immediately after the latter was computed. We will refer to this algorithm as $\mathcal{EA}$ (short for Enhanced $\mathcal{A}$). Clearly, the running time of $\mathcal{EA}$ will be equal to the running time of $\mathcal{A}$ augmented by the time needed to evaluate $L(m)$ from $F(m)$, for all m.

To solve problem (3.22) for all n we will use algorithm $\mathcal{EA}$ with $\hat{W}_0(m)$ in place of $L(m)$ and $\hat{W}'(n)$ in place of $F(n)$. The enhancement procedure computes each $\hat{W}_0(n)$ based on $\hat{W}'(n)$ using the computations in (3.21) and (3.19). The running time to solve the minimization in (3.21) for given n is $O(T(\omega))$ operations and doing so for all n requires $O(T(\omega)N)$ operations, where $T(\omega) \triangleq T_1(\omega) - 2 + \max(0, T_3(\omega) - T_2(\omega) - 2)$. Thus, by employing the enhanced algorithm to solve the single-source MWP problem in $G(\omega)$ leads to a time complexity of $O(N(T(\omega) + \log N))$. Further, by using $\mathcal{EA}$ repeatedly to solve the all-pairs MWP problem in $G(\omega)$ the time complexity achieved is $O(N^2(T(\omega) +$

$\log N)$). We point out that for this technique to be applied we first must determine the values $T_1(\omega), T_2(\omega), T_3(\omega)$. This process takes $O(N^2)$ operations and thus it does not contribute to increasing the overall asymptotical time complexity.

It is important to point out that in [13] it is assumed that the weights $\omega'(m,n)$ are real-valued. This implies that all values $g(m,n)$ are finite, while in our case some of them are ∞. For this reason we need to perform some slight adjustments to algorithm $\mathcal{A}$. These are explained in detail in the following section, where we also show that they do not impact the algorithm correctness.

Let us discuss now the impact in terms of running time of using the above development to solve the all-pairs MWP problem in our WDAG of interest, namely $G(\omega_1)$ and $G(\omega_2)$. According to (3.14) and (3.15) the weight function ω_1 and ω_2 comply to the following general form

$$\omega(m,n) = \mu v_{1,\kappa}((x_m, x_n]) + \lambda v_{2,\kappa}((x_m, x_n]), \tag{3.28}$$

for some positive μ and λ. For simplicity we use the notation T_1, T_2, T_3, T instead of $T_1(\omega), T_2(\omega)$, $T_3(\omega), T(\omega)$, respectively, in the rest of the paper. Notice that the values T_1, T_2 and T_3 depend on the joint probability distribution of X and Y_κ, denoted by p_{XY_κ}, and on the ratio λ/μ. In our experiments, where we used discretized Gaussian sources with discretized Gaussian SI, we found that there exist two thresholds $\tau_1(p_{XY_\kappa}) \leq \tau_2(p_{XY_\kappa})$ such that when $\lambda/\mu < \tau_1(p_{XY_\kappa})$, we have $T_3 \leq T_2$, while for $\lambda/\mu > \tau_2(p_{XY_\kappa})$ we have $T_3 = N$. Thus, when $\lambda/\mu < \tau_1(p_{XY_\kappa})$ the running time of $\mathcal{EA}$ is $O(N(T_1 + \log N))$. We have observed empirically that T_1 could be lower than $N/10$ when $\lambda/\mu < \tau_1(p_{XY_\kappa})$, which leads to the conclusion that applying $\mathcal{EA}$ may lead to significant savings in running time. On the other hand, when $\lambda/\mu > \tau_2(p_{XY_\kappa})$ we have $T > N/2$ thus, the proposed complexity reduction is not sufficient to reduce the asymptotical time complexity.

We have observed in our experiments that in the F-WZSQ case, the condition $\lambda/\mu < \tau_1(p_{XY_\kappa})$ holds in many cases of interest. Thus, in such cases, by using $\mathcal{EA}$ the running time to solve the F-WZSQ design problem (excluding the preprocessing stage) decreases to $O(N^2(T_1 + \log N))$.

In the R-WZSQ case the condition $\lambda/\mu < \tau_1(p_{XY_\kappa})$ is also satisfied in at least one of the two graphs $G(\omega_1)$ and $G(\omega_2)$ in most cases of interest, but rarely in both of them. However, even if the solution to the all-pairs MWP problem is accelerated in only one of the two WDAGs, this contributes signficantly to the reduction of the actual running time, even if the asymptotical value still remains $O(N^3)$. More specifically, the constant hidden in the big-O notation is reduced in half.

3.5 Algorithm $\mathcal{EA}$

This section presents algorithm $\mathcal{EA}$ in detail. The following notations will be used

$$g(m,n) \triangleq \hat{W}_0(m) + \omega'(m,n),$$
$$g_2(m,n) \triangleq \hat{W}_0(m) + \omega(m,n),$$

for $0 \leq m < n \leq N$. Further, denote $\mathcal{S} \triangleq \{k | 0 \leq k < T_1 - 1 \text{ or } T_2 + 1 < k < T_3\}$. For each $1 \leq n \leq N$, let $bestleft(n)$ denote the value of m achieving the minimum in (3.22) (which also achieves the minimum in (3.20)), and let $bestleft_2(n)$ be the value of m achieving the minimum in (3.21). Further, let $bestleft_0(n)$ denote the node before n in the optimum path from 0 to n in $G(\omega)$. In virtue of (3.19) $bestleft_0(n)$ is the best of $bestleft(n)$ and $bestleft_2(n)$.

The pseudocode of algorithm $\mathcal{EA}$ is presented on the next page. The algorithm exploits

the fact that the function g satisfies the Monge property, fact which follows easily based on Proposition 1. Algorithm $\mathcal{EA}$ uses a deque (i.e., a double-ended queue) $\mathcal{D}$. At all times $\mathcal{D}$ will contain a sequence of increasing integers in the range between 0 and $N-1$. The element at the front, which is the smallest in the deque, will be denoted f, and the next element $f2$. The element at the rear, which is the largest, will be denote r, and the previous element $r2$. Note that $f2$ and $r2$ are defined only when the deque has at least two elements. The update operations allowed on $\mathcal{D}$ are $RemoveFront$, which deletes f, $RemoveRear$, which deletes r and $InsertAtRear(m)$, which appends m at the rear. The access of f, $f2$, r and $r2$ is also allowed on $\mathcal{D}$.

Algorithm $\mathcal{EA}$: Solution to the single source MWP problem in $G(\omega)$.

begin
  $\hat{W}_0(0) \leftarrow 0$, $\mathcal{D} \leftarrow \{0\}$
  for $m = 1$ **to** $N-1$ **do**
    $\hat{W}'(m) \leftarrow g(f,m)$, $bestleft(m) \leftarrow f$
    $\hat{W}''(m) \leftarrow \min\limits_{k, m-k \in \mathcal{S}} g_2(k,m)$
    $bestleft_2(m) \leftarrow \arg\min\limits_{k, m-k \in \mathcal{S}} g_2(k,m)$
    $\hat{W}_0(m) \leftarrow \min(\hat{W}'(m), \hat{W}''(m))$; Compute $bestleft_0(m)$
    while $|\mathcal{D}| > 1$ *and* $g(f2, m+1) \leq g(f, m+1)$ **do**
      $RemoveFront$
    while $|\mathcal{D}| > 1$ *and* $Bridge(r2, r, m)$ **do**
      $RemoveRear$
    $InsertAtRear(m)$
  $\hat{W}'(N) \leftarrow g(f,N)$, $bestleft(N) \leftarrow f$
  $\hat{W}''(N) \leftarrow \min\limits_{k, N-k \in \mathcal{S}} g_2(k,N)$
  $bestleft_2(N) \leftarrow \arg\min\limits_{k, N-k \in \mathcal{S}} g_2(k,N)$
  $\hat{W}_0(N) \leftarrow \min(\hat{W}'(N), \hat{W}''(N))$; Compute $bestleft_0(N)$

The deque contains all current candidates for $bestleft(m)$, for all m which are yet to

be considered. The algorithm uses the procedure $Bridge(r2, r, m)$, where $r2 < r < m$, which returns true if and only if $g(r, k) \geq \min(g(r2, k), g(m, k))$, for all $m < k \leq N$.

We point out that in the Basic Algorithm of [13] operation $InsertAtRear(m)$ is performed only if $g(m, N) < g(r, N)$. However, a careful examination of the proof of correctness given in [13] reveals that the algorithm is still correct if that condition is removed.

$Bridge(a, b, c)$

```
1  begin
2      max ← min(N, b + T_2 + 1);
3      if c=max then return true;
4      min ← max(c + 1, b + T_1 − 1);
5      low ← min; high ← max;
6      if g(a, max) ≤ g(b, max) then return true;
7      while high − low ≥ 2 do
8          mid ← ⌊(low + high)/2⌋;
9          if g(a, mid) ≤ g(b, mid) then
10             low ← mid
11         else
12             high ← mid
13     if g(c, high) ≤ g(b, high) then
14         return true
15     else
16         return false
```

The fact that function g satisfies the Monge property implies that the following property holds. Its proof is deferred to the appendix.

The Forward Property (FP): Let $0 \leq a < b < c < d \leq N$.

FP1) If $g(b, c) < g(a, c)$ then $g(b, d) \leq g(a, d)$.

FP2) If $g(b, c) < g(a, c)$ and $g(b, d) \neq \infty$ then $g(b, d) < g(a, d)$.

Note that in the case when the weights have only finite values (as in [13]) a stronger

variant of FP holds, where the inequality $g(b,d) \leq g(a,d)$ in FP1 is always strict. The proof of correctness of algorithm $\mathcal{A}$ given in [13] relies on the strong FP. However, a careful examination of their proof leads to the conclusion that only the weaker FP1 and FP2 are sufficient. Specifically, FP is invoked in the proof in four places and in each of them FP1 is actually used. The Monge condition (referred to as the concavity condition in [13]) is also invoked at the end of the proof, where actually FP2 suffices.

The subroutine $Bridge(a,b,c)$ proposed in [13] relies on the stronger FP and uses a binary search over the set of integers from c to N to determine whether some $k, c < k \leq N$ exists such that $g(b,k) < \min(g(a,k), g(c,k))$. Specifically, the procedure finds the smallest such value if it exists. Clearly, for such a k we have $g(b,k) \neq \infty$. Thus, it is safe to restrict the search range to the range for which $g(b,k) \neq \infty$, i.e., from $\max(c+1, b+T_1-1)$ to $\min(N, b+T_2+1)$. Then the stronger FP holds for this range and no further adjustment is needed. The pseudocode of the subroutine $Bridge(a,b,c)$ is shown in the previous page.

3.6 Experimental Results

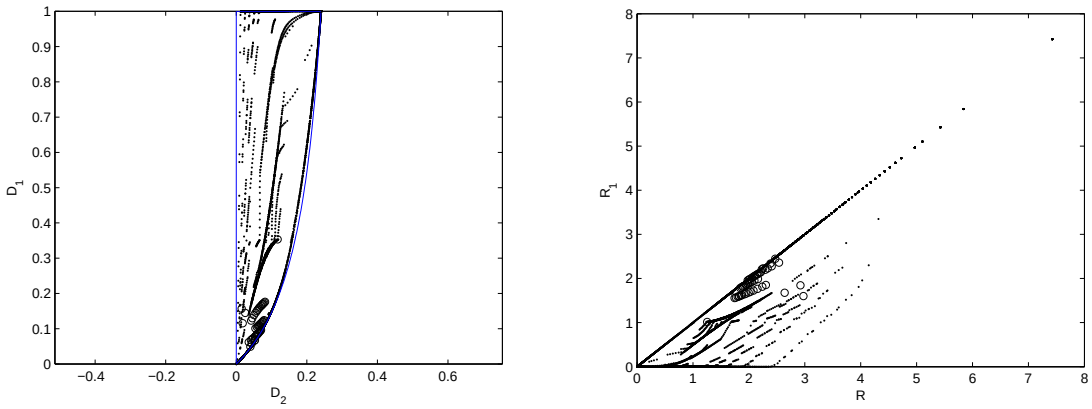
This section assesses the practical performance of the proposed design algorithms for the two scenarios considered in this work. In our experiments the source X is obtained by discretizing a continuous Gaussian variable $\tilde{X}$ with mean 0 and variance 1. Specifically, $N = 1000$ and the source alphabet $\mathcal{X}$ is formed of the centroids of the intervals $(-\infty, -6), (6, \infty)$ and of the sets obtained by partitioning $(-6, 6)$ into 998 equal-size intervals. The distortion measure is the squared distance and $\hat{\mathcal{X}} = \mathbb{R}$. For $\kappa = 1, 2$, the SI Y_κ is obtained by discretizing the random variable $\tilde{X} + Z_\kappa$, where Z_κ is Gaussian and independent of $\tilde{X}$. Specifically, for $\kappa = 1, 2$, the alphabet $\mathcal{Y}_\kappa$ consists of 300 values, which

are the centroids of the intervals $(-\infty, -6), (6, \infty)$ and of the sets obtained by partitioning $(-6, 6)$ into 298 equal-size intervals. More details about each Z_κ will be given when discussing each scenario.

Since we will compare our results with the theoretical bounds for the continuous Gaussian source, we will evaluate the performance of our schemes for the continuous case, i.e., when the source is $\tilde{X}$ and the SI is $\tilde{X} + Z_\kappa$. Note that there is no difference between the continuous case and the discretized version in terms of rate, but only in terms of distortion. Namely, when evaluating the distortion $\tilde{D}_\kappa(\mathbf{f})$ for the continuous case we need to account for the distortion due to the discretization as well. Throughout this section we use the notations D_κ, R and R_κ instead of $\tilde{D}_\kappa(\mathbf{f})$, $R(\mathbf{f})$ and $R_\kappa(\mathbf{f})$, respectively, for $\kappa = 1, 2$. We first present the results for the F-WZSQ problem in subsection 3.6.2. We continue with the experimental results for the R-WZSQ scenario in subsection 3.6.3. We end the section with a discussion of our empirical observations regarding the satisfaction of the partial Monge property and its impact on the running time in subsection 3.6.4.

3.6.1 Discussion of Traditional HB Problem Results

In this subsection we present experimental results for the traditional HB problem, i.e., when the SI at the first decoder Y_1 is a constant. We have $Z_2 = N_2$, where $N_2 \sim \mathcal{N}(0, 1/\sqrt{10})$, and N_2 is independent of $\tilde{X}$. Recall that for the HB problem we have $\lambda_1 = 0$ in the cost function (3.4). We solve the optimization problem (3.5) for $\rho \in \{0.01, 0.02, 0.03, 0.04, 0.05, 0.1, 0.11, 0.12,\ 0.13, 0.14, 0.2, 0.3,\ 0.4, 0.5, 0.6\}$. The values of λ_2 range from 0.0001 to 0.2 in increments of 0.0005.



(a) (b)

(c)

Figure 3.2: Traditional HB problem where SI may be absent. (a) shows the achieved distortion region. Theoretical distortion outline is marked in blue. (b) shows corresponding achieved rate region. (c) shows the rate difference of achieved rate pair $(R_1(D_1), R_1 + R_2(D_1, D_2))$ to the theoretical rate pair $(R_1^*(D_1)), (R_1 + R_2)^*(D_1, D_2)$. Circles in all three figures marker the points with a gap higher than 0.265 in R_1.

In Fig. 3.2(a) we plot the distortion pairs (D_1, D_2) achieved in our experiments, while in Fig. 3.2(b) we show the rate pairs $(R_1, R_1 + R_2)$. We point out that even though in the HB problem only the total rate is of interest we find instructive to analyze how the total rate is split between the two virtual stages of the encoder.

We emphasize that the convex closure of the set of distortion pairs achieved in our experiments closely matches the theoretical achievable region specified in [12]. The upper boundary of the distortion region corresponds to the case when D_1 is maximum and, hence $R_1 = 0$. The right-side boundary contains the points achieved when $R_2 = 0$. On the other hand, the leftmost boundary is approached when the total rate $R_1 + R_2$ is high enough. We point out that in our experiments when $\rho \geq 0.15$ the distortion pairs are on the right boundary, i.e., $R_2 = 0$.

Fig. 3.2(c) plots the difference between the practical rate pairs $(R_1, R_1 + R_2)$ and the theoretical lower bounds in rate for the corresponding distortion pair (D_1, D_2), i.e., $(R_{WZ}(D_1), R_{HB}(D_1, D_2))$ [12]. Note that $R_{WZ}(\cdot)$ and $R_{HB}(\cdot)$ denote the rate-distortion function in the WZ and HB scenarios, respectively. We see that the gap in the total rate is within 0.275 bits/sample. This result is very encouraging since it shows that the gap is very close to that predicted by the high rate quantization theory between scalar quantization and infinite dimension vector quantization, namely of 0.25 bits/sample.

On the other hand we see that the gap in the first-stage rate can be higher than 0.25 bits, reaching up to 0.42 bits. The rate pairs and distortion pairs for which this happens are marked in circles in 3.2(a) and Fig. 3.2(b). We notice that this points have a higher concentration in the region where D_1 is relatively small, which agrees with the intuition that it is more difficult to achieve small distortions. This result indicates that when D_1 is small, beside the expected rate loss of 0.25 bits attributed to the use of SQ versus infinite

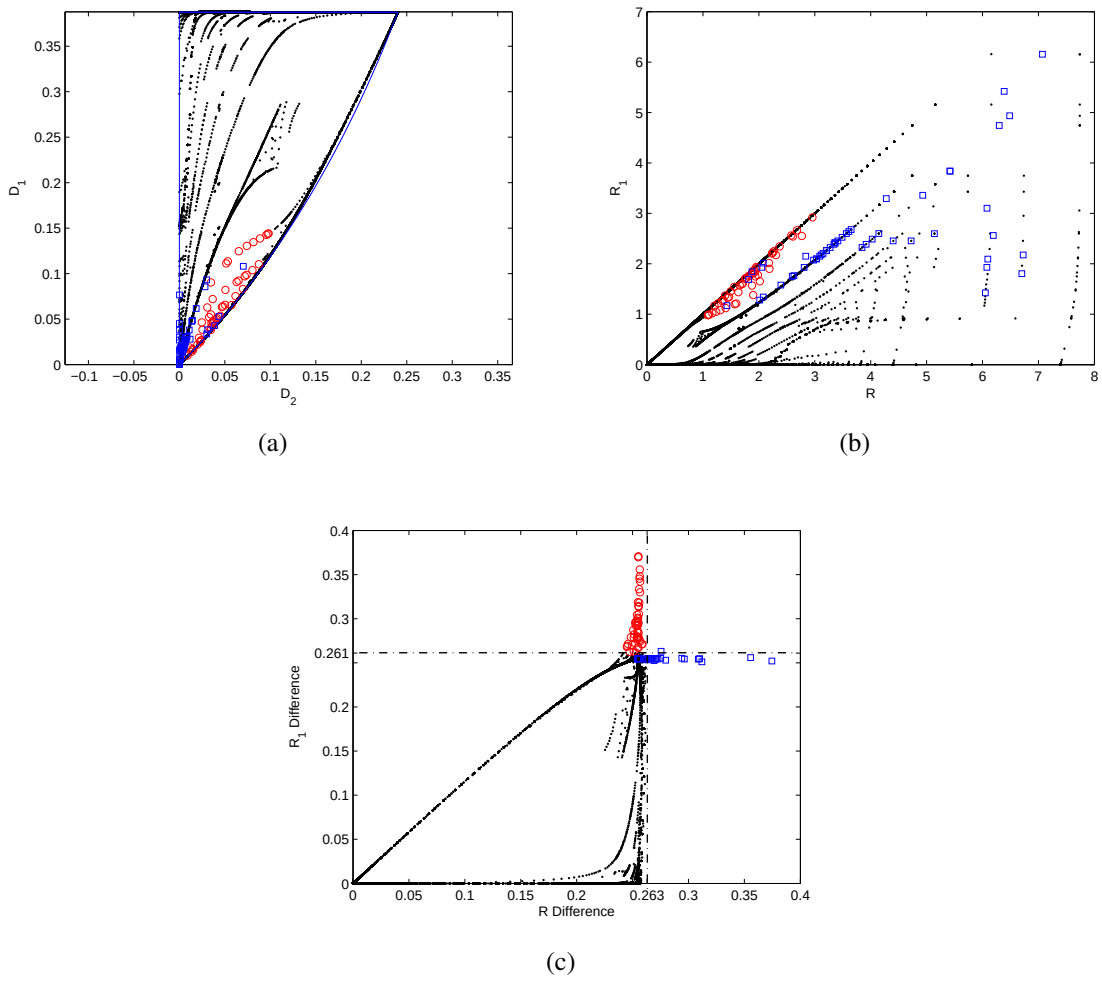
dimension VQ, there is some extra loss at the first stage virtual encoder. It is important to point out that this extra loss at the first virtual stage is completely canceled out after the second virtual stage, so that the difference in total rate versus the theoretical limit remains 0.25 bits, as shown in Fig. 3.2(c).

The parameters ρ, λ_2 have a strong correlation with achieved quadruples. Since the first decoder does not have a SI, but the second decoder has a strong SI, it's rewarding to have a big R_1 reducing D_1 for some fixed rate sum to minimize weighted distortion. In our experiments we have obtained $R_2 > 0$ only when $\rho < 0.15$. otherwise $R_1 =$ rate sum. For a fixed ρ, as the decreasing of λ_2, R_1 is increasing in a slope approximately parallel to the upper boundary. The smaller ρ, the further away from the upper boundary achieved points will be. For a smaller ρ, it is also possible to have $R_2 = 0$ when λ_2 is relatively small.

3.6.2 Discussion of F-WZSQ Results

In the F-WZSQ case we have $Z_1 = N_1 + N_2$ and $Z_2 = N_2$, where $N_1 \sim \mathcal{N}(0, \frac{1}{\sqrt{10}}), N_2 \sim \mathcal{N}(0, \frac{1}{\sqrt{10}})$, and N_1 and N_2 are independent of each other and of $\tilde{X}$. The value of ρ used in our experiments are $0.05, 0.1, 0.102, 0.105, 0.11, 0.12, 0.13, 0.15, 0.2, 0.3, 0.5, 0.8, 0.95$. The values of λ_1 are in the range of $(10^{-5}, 0.9)$, and values of λ_2 are in the range of $(10^{-5}, 0.3)$.

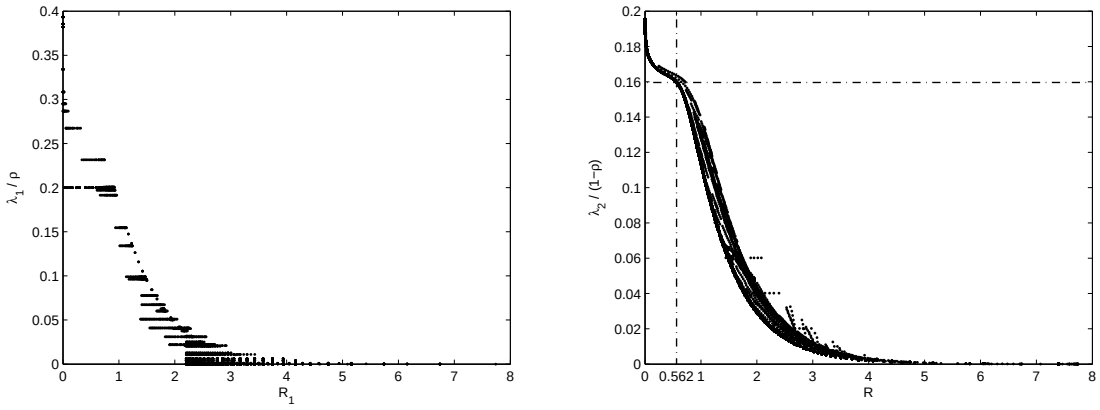
The distortion pairs (D_1, D_2) and the rate pairs (R_1, R) are plotted in Fig. 3.3(a) and Fig. 3.3(b), respectively. Fig. 3.3(a) also shows the boundary of the theoretically achievable distortion region given in [28] (in blue). Fig. 3.3(c) plots the difference between the practical rate pairs (R_1, R) and the pair of theoretical lower bounds $(R_{WZ}(D_1), R_{HB}(D_1, D_2))$ [28]. We see that in most of our experiments the gap in both R_1 and R is within 0.263. The corresponding points are marked using black dots in all three subfigures of Fig. 3.3.



(a) (b)

(c)

Figure 3.3: F-WZSQ results. (a) Practical and theoretical (blue line) distortion region. (b) Practical rate region. (c) Difference between R_1, respectively R, and the corresponding theoretical rate bounds for all the distortion pairs in (a). Circle markers are for the cases when the gap in R_1 is higher than 0.261, square markers are for the cases when the gap in R is higher than 0.263.



(a) (b)

Figure 3.4: (a) Plot of $\frac{\lambda_1}{\rho}$ versus R_1 when $R \geq 2.2$; (b) Plot of $\frac{\lambda_2}{1-\rho}$ versus R when $R_2 > 0.001$.

The points which do not fit in the aforementioned category (termed "extra loss points") exhibit an additional loss in either R_1 (points marked using circles) or in R (points marked using squares). The cases with extra loss in R_1 appear for relatively small D_1 and R_2. The cases with extra loss in R are mostly occurring when D_2 is very small, hence R is large. Note that the rate gap between scalar quantization and infinite dimension vector quantization, predicted by the high rate quantization theory for the single encoder-decoder pair problem is 0.254 bits/sample [46]. The existence of additional rate loss on top of these 0.254 bits can be attributed to the additional tension introduced in the optimization problem because of the need to meet the quality requirements at two decoders instead of one, while preserving rate constraints at two encoders as opposed to one.

It is instructive to analyze how the choice of the parameters ρ, λ_1 and λ_2 influences the algorithm outcome. In our experiments we have obtained $R_1 > 0$ only when $\frac{\lambda_1}{\rho} < 0.4$, while $R_2 > 0$ was obtained only when $\frac{\lambda_2}{1-\rho} < 0.9$. We point out that for $\frac{\lambda_2}{1-\rho} > 0.2$, R_2 is very small, namely $R_2 \leq 0.001$. Our results show a strong correlation between R_1 and

the value of $\frac{\lambda_1}{\rho}$ when R is higher than 2.2 bits, and between R and $\frac{\lambda_2}{1-\rho}$ when $R_2 > 0.001$. Specifically, Fig. 3.4(a), where we plot the value of $\frac{\lambda_1}{\rho}$ versus R_1, for the cases when $R \geq 2.2$, shows that R_1 tends to increase with the decrease of $\frac{\lambda_1}{\rho}$. Further, Fig. 3.4(b), containing the plot of $\frac{\lambda_2}{1-\rho}$ versus R when $R_2 > 0.001$, shows that R increases as $\frac{\lambda_2}{1-\rho}$ becomes smaller. Additionally, notice that we have $R_2 > 0.001$ and $R_1 + R_2 \geq 0.57$ only if $\frac{\lambda_2}{1-\rho} \leq 0.16$. This observation will be useful in the last subsection where we discuss the satisfaction of the partial Monge property.

3.6.3 Discussion of R-WZSQ Results

In the R-WZSQ case we have $Z_1 = N_1$ and $Z_2 = N_1 + N_2$, where $N_1 \sim \mathcal{N}(0, \frac{1}{\sqrt{10}}), N_2 \sim \mathcal{N}(0, \frac{1}{\sqrt{10}})$, and N_1 and N_2 are independent of each other and of $\tilde{X}$. The values of ρ used in our experiments are $0.1, 0.12, 0.15, 0.2, 0.85, 0.9, 0.95, 0.96, 0.97$. The values of λ_1 range between 0.01 and 0.1. The values of λ_2 range between 10^{-5} and 0.4.

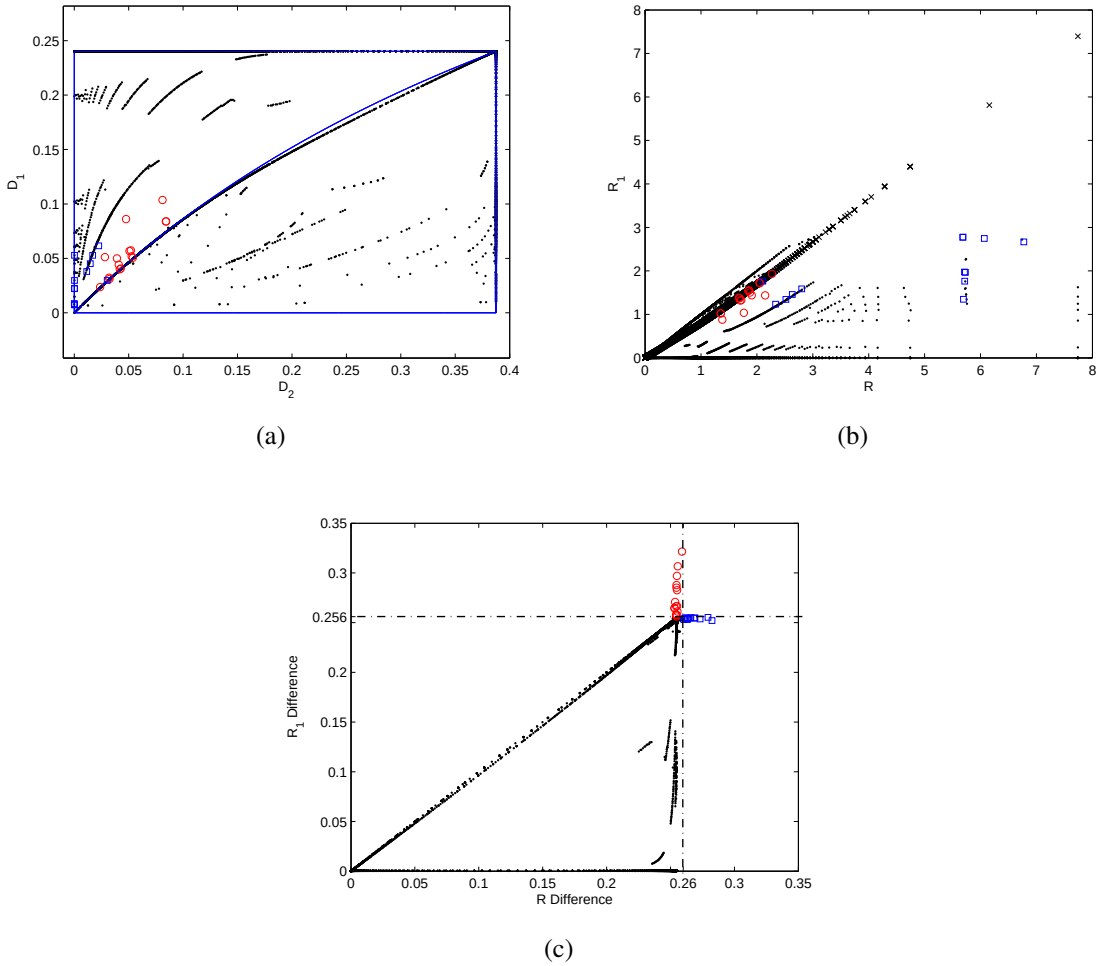
Tian and Diggavi [29] showed that the achievable RD region they proposed for the RDSI-WZ problem is exact in the quadratic Gaussian case with jointly Gaussian SI. Moreover, they showed that any rate pair on the lower boundary of the rate region for given (D_1, D_2) can be achieved with only two codebooks, a coarse codebook to be used at one decoder, and a finer codebook, to be used at the other decoder. Which decoder recovers the finer codebook depends on the particular distortion pair (D_1, D_2). Our experimental results confirm this property since all the time at most one of the two quantizers Q_1 and Q_2 has a more refined partition than the coarse partition f_0. Fig. 3.5(a) and Fig. 3.5(b) plot the achieved distortion pairs and rate pairs, respectively. Unlike the case with forwardly degraded SI the theoretical distortion region is a filled rectangle. Its boundaries are shown in blue in Fig. 3.5(a). The curve connecting the bottom left corner with the top

right corner corresponds to the case when both decoders use the coarse partition, i.e. there is no-refinement in either of Q_1 and Q_2. We will refer to this case as the "no refinement" case. Note that the blue curve contains the theoretical no-refinement distortion pairs, while the black points situated close to this curve represent the practical no-refinement pairs. The no-refinement rate pairs are marked using crosses in Fig. 3.5(b).

The no-refinement distortion curve separates the distortion region into two sub-regions: lower and upper. The upper distortion sub-region represents the case when only quantizer Q_2 has a refined partition. The corresponding rate pairs appear below the no-refinement curve in Fig. 3.5(b). The lower distortion sub-region contains the distortion pairs achieved when only Q_1 has a refinement. The corresponding rate pairs are above the no-refinement curve in Fig. 3.5(b). Notice that the rate sub-region for the latter case is much smaller than the other sub-region. This is because for a fixed sum rate, once R_2 is big enough, Q_2 gets refinements.

Fig. 3.5(c) plots the difference between the practical rate pairs (R_1, R) and the pair of theoretical lower bounds [29]. We observe that in most of our experiments the gap in both R_1 and rate-sum is within 0.26. The remaining points exhibit an extra loss either only in R_1 (points marked with circles) or only in rate-sum (points marked with squares). Similarly, to the case with FDSI-WZ, the cases with extra loss in R_1 appear for relatively small D_1 and R_2. The cases with extra loss in R are mostly occurring when D_2 is very small, hence R is large.

In our experiments we found that when Q_1 has a refinement we have $\frac{\lambda_1+\lambda_2}{\rho} < 0.255$ and $\frac{\lambda_2}{1-\rho} >= 0.9$. On the other hand, when Q_2 has a refinement we have $\frac{\lambda_1}{\rho} < 0.84$ and $\frac{\lambda_2}{1-\rho} < 0.44$, with $\frac{\lambda_2}{1-\rho} > 0.26$ only when $R < 0.44$.



(a) (b)

(c)

Figure 3.5: R-WZSQ results. (a) Practical and theoretical (blue line) distortion region. (b) Practical rate region. (c) Difference between R_1, respectively R, and the corresponding theoretical rate bounds for all the distortion pairs in (a). Circle markers are for the cases when the gap in R_1 is higher than 0.256, square markers are for the cases when the gap in R is higher than 0.26.

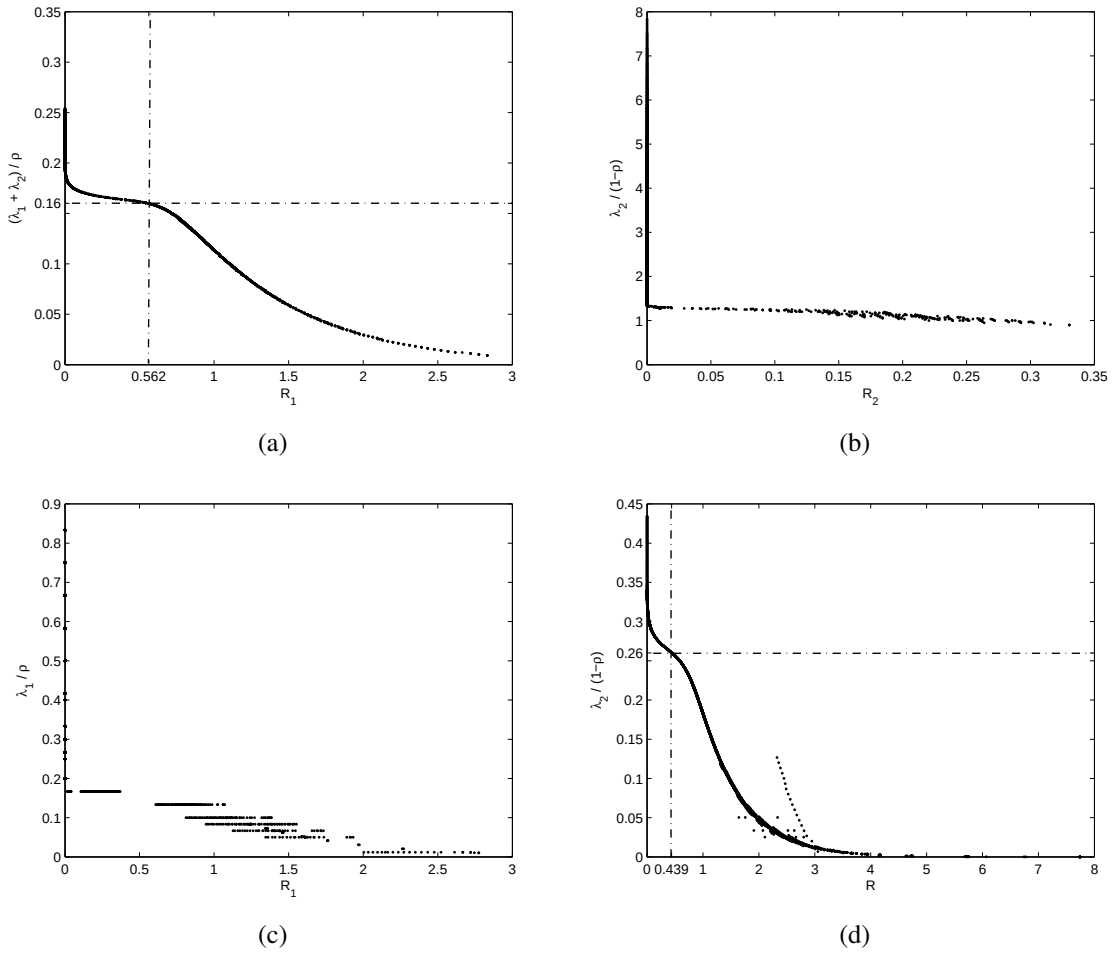


Figure 3.6: (a) Relation between $\frac{\lambda_1+\lambda_2}{\rho}$ and R_1 when Q_1 has a refinement. (b) Relation between $\frac{\lambda_2}{1-\rho}$ and R_2 when Q_1 has a refinement. (c) Relation between $\frac{\lambda_1}{\rho}$ and R_1 when Q_2 has a refinement. (d) Relation between $\frac{\lambda_2}{1-\rho}$ and R when Q_2 has a refinement. The points which deviate significantly from the main curve correspond to very small refinement in Q_2.

When Q_1 has a refinement we found that R_1 increases as $\frac{\lambda_1+\lambda_2}{\rho}$ decreases, as seen in Fig. 3.6(a), while R_2 increases as $\frac{\lambda_2}{1-\rho}$ decreases, as seen in Fig. 3.6(b). When Q_2 has a refinement we notice from Fig. 3.6(c) that R_1 tends to increase as $\frac{\lambda_1}{\rho}$ decreases. Additionally, the sum-rate R generally increases as $\frac{\lambda_2}{1-\rho}$ decreases, except when the refinement in Q_2 is very small as seen in Fig. 3.6(d).

3.6.4 Fulfillment of the Partial Monge Property

In this subsection we first evaluate T_1, T_2 and T_3 for the graph $G(\omega)$ with the edge weight function given in (3.28). We consider the SI Y_κ obtained by discretizing $\tilde{X} + Z$, where Z is Gaussian and independent of $\tilde{X}$. We consider three cases with the following variances $\sigma_Z^2 = \frac{1}{\sqrt{10}}, \frac{2}{\sqrt{10}}$ and 0.8.

In Table 3.1 we show the values of T_1, T_2 and T_3 for the aforementioned cases of SI for several values of λ/μ ranging from 0.05 to 0.5. We observe that as the ratio λ/μ increases T_1 and T_3 are nondecreasing, while T_2 is nonincreasing at a very slow rate. Another interesting observation is that, for fixed λ/μ, T_1 and T_2 are nondecreasing as the SI becomes weaker (T_1 changing at a very slow rate), while T_3 is nonincreasing.

For the strongest SI, we have $\tau_1(p_{XY_\kappa}) \approx 0.16$ and $\tau_2(p_{XY_\kappa}) \approx 0.1635$. For the second strongest SI we have $\tau_1(p_{XY_\kappa}) \approx \tau_2(p_{XY_\kappa}) \approx 0.26$.

Recall that for the F-WZSQ design the all-pairs MWP problem has to be solved only in $G(\omega_2)$. The edge weights are given in (3.15), which corresponds to equation (3.28) with $\kappa = 2$, $\lambda = \lambda_2$ and $\mu = 1 - \rho$. Recall that SI Y_2 used in our experiments for F-WZSQ design has $\sigma_Z^2 = \frac{1}{\sqrt{10}}$, hence it is the strongest among the three cases considered in this subsection. Thus, when $\frac{\lambda_2}{1-\rho} < 0.16$ a significant complexity reduction can be achieved. As seen in Fig. 3.4(b) all of the cases corresponding to a sum-rate larger than 0.57 and

$R_2 > 0.001$ are obtained when this condition holds.

The R-WZSQ design algorithm has to solve the all-pairs MWP problem in both $G(\omega_1)$ and $G(\omega_2)$. For $G(\omega_1)$ we have $\kappa = 1$, $\lambda = \lambda_1 + \lambda_2$ and $\mu = \rho$. The SI Y_1 used in the experiments for R-WZSQ is the strongest among the three considered in this subsection. Thus, a significant complexity reduction can be achieved when $\frac{\lambda_1+\lambda_2}{\rho} < 0.16$. On the other hand, for $G(\omega_2)$ we have $\kappa = 2$, $\lambda = \lambda_2$ and $\mu = 1 - \rho$. The SI Y_2 has $\sigma_Z^2 = \frac{2}{\sqrt{10}}$. A considerable complexity reduction can be obtained when $\frac{\lambda_2}{1-\rho} < 0.26$. As seen from Fig. 3.6, in order to achieve $R_1 > 0.57$ or $R > 0.44$ at least one of conditions $\frac{\lambda_1+\lambda_2}{\rho} < 0.16$ and $\frac{\lambda_2}{1-\rho} < 0.26$ must hold. In such a case, the all-pairs MWP problem in at least one of the two WDAGs will run considerably faster. However, cases when both conditions $\frac{\lambda_1+\lambda_2}{\rho} < 0.16$ and $\frac{\lambda_2}{1-\rho} < 0.26$ are satisfied are more rare. Thus, the asymptotical time complexity will be reduced only in a smaller number of cases, however, in many cases the constant hidden in the big-Oh notation will be reduced in half, effectively decreasing the practical running time.

Table 3.1: T_1, T_2, T_3 Experimental Data

	$\sigma_Z^2 = \frac{1}{\sqrt{10}}$			$\sigma_Z^2 = \frac{2}{\sqrt{10}}$			$\sigma_Z^2 = 0.8$		
λ/μ	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3
0.05	34	363	98	34	496	96	34	531	95
0.1	50	363	166	51	489	156	51	526	155
0.16	66	363	285	67	482	221	68	510	218
0.1635	66	363	424	68	482	225	69	510	222
0.2	74	363	1000	77	476	262	78	510	257
0.26	86	358	1000	90	473	448	92	508	314
0.3	93	358	1000	98	473	1000	100	508	1000
0.4	107	358	1000	116	471	1000	119	506	1000
0.5	120	357	1000	131	465	1000	135	496	1000

3.7 Conclusion

In this work, we address the design of a two-stage Wyner-Ziv scalar quantizer with forwardly or reversely degraded side information (SI) for finite-alphabet sources and SI. We assume that the binning is performed perfectly so that the theoretical limits are achieved and focus on the optimization of the quantizer partitions. The optimization problem aims to minimize a weighted sum of distortions and rates. The proposed solution is based on solving the single-source or the all-pairs minimum-weight path problem in some weighted directed acyclic graphs. By employing dynamic programming, which is the conventional solution for the underlying MWP problem, the time complexity achieved is $O(N^3)$, where N denotes the size of the source alphabet. Further, we introduce a so-called partial Monge property and propose a technique to exploit it in order to expedite the solution algorithm. The proposed solution algorithm is globally optimal when the quantizer cells are contiguous. Experimental results using a discretized Gaussian source with discretized Gaussian SI assess the practical performance of the proposed scheme and show that the partial Monge property holds in many situations of interest.

In next chapter, we will present how to utilize graphs to solve the multiple description coding problem.

Chapter 4

Improved Two-Stage Multiple Description Scalar Quantizer

4.1 Overview

In this chapter, we present an improved MMDSQ design with optimized encoder partitions. Recall that the MMDSQ [30] has two scalar quantizers with staggered thresholds at the first stage as side quantizers. Then each joint cell formed by intersecting two side cells is further divided into a fixed number of finer cells forming the central quantizer at the second stage. The MMDSQ of [30] is attractive in comparison with the ECMDSQ of [34] because of its simplicity. However, its performance is comparable with that of [34] only at high rates. The proposed improved MMDSQ has performance very close to ECMDSQ at low rates as well.

In our improved MMDSQ scheme, the optimization is achieved by solving the all-pairs MWP problem in a WDAG $G(\omega)$ followed by solving another single-source MWP problem in a coupled quantizer graph $\mathbb{G}(w)$. We also discuss a variant of the improved

MMDSQ with enhanced decoders, like in [15], where the side decoders use the refinement information at the second stage. The aim of the variant design is to decrease the gap to the theoretical bound.

This chapter is organized as follows. Section 4.2 introduces the notations and the problem formulation. The structure of our improved MMDSQ is presented. The problem is formulated as a minimization of weighted sum of all distortions and rates. We present our improved MMDSQ design in Section 4.3 and the variant design with enhanced decoders in the following section. Both designs have the same optimization process, but the enhanced design has a different decoding rule. Some numerical results are exhibited in Section 4.5 for a Gaussian source in the symmetric case. Finally, Section 4.6 concludes this chapter.

4.2 Problem Formulation

First we review the MMDSQ design proposed in [30]. The MMDSQ system operates in two stages: at the first stage, two staggered uniform quantizers with bin size Δ generate two side partitions and produce a joint uniform quantizer with bin size $\Delta/2$, i.e., the intersection of two side partitions; at the second stage, each bin in the joint quantizer is further divided into a fixed number of uniform finer bins forming the central partition as shown in Fig. 4.1. The messages from the first stage are entropy encoded into each description, and the refinement from the second stage is entropy encoded and evenly split between two descriptions. The two side decoders recover the source using the information at the first stage based on each side partition, whereas the central decoder uses the information at both stages to reconstruct the source. Notice that the MMDSQ can be regarded as consisting of four quantizers $\mathbf{Q} = (Q_1, Q_2, Q_{12}, Q_0)$, where Q_1 and Q_2 are two side quantizers, Q_{12} is the joint quantizer, and Q_0 is the central quantizer as shown in Fig. 4.1.

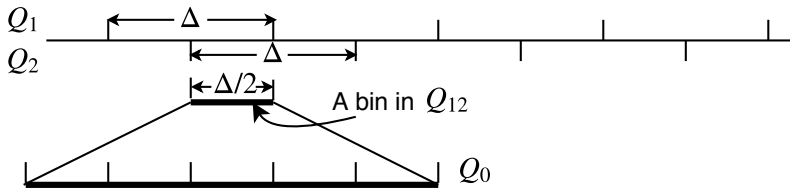


Figure 4.1: Structure of MMDSQ.

The improvement we propose resides in optimizing the encoder partitions rather than considering uniform partitions for Q_1, Q_2 and Q_0. One difference comparing to the design in [6] resides in the addition of the central quantizer Q_0 which is actually a refinement of Q_{12}.

Let C_i, C_j denote a cell in Q_1, respectively, Q_2, let C_{ij} denote the intersection of cell C_i and C_j, and C_{ijk} denote a cell in Q_0 obtained by partitioning C_{ij} for $1 \leq i \leq M_1, 1 \leq j \leq M_2, 1 \leq k \leq M_{0,ij}$, as shown in Fig. 4.2. Denote the number of cells in quantizer Q_i by M_i for $i = 0, 1, 2$, and let $M_0 = \max_{ij} M_{0,ij}$, where $M_{0,ij}$ is the number of finer bins within each C_{ij}. Let $D_l(\mathbf{Q})$ and $R_l(\mathbf{Q})$ represent the distortion, respectively, the rate of Q_l for $l = \{0, 1, 2\}$. The RD region of our improved MMDSQ can be characterized by the tuple $(R_1(\mathbf{Q}), R_2(\mathbf{Q}), R_0(\mathbf{Q}), D_0(\mathbf{Q}), D_1(\mathbf{Q}), D_2(\mathbf{Q}))$.

Let an ascending sequence $\mathbf{s} = (s_0, s_1, ..., s_{M_1})$ denote the thresholds of Q_1 such that $C_i = (s_{i-1}, s_i]$, and another ascending sequence $\mathbf{t} = (t_0, t_1, ..., t_{M_2})$ denote the thresholds of Q_2 such that $C_j = (t_{j-1}, t_j]$, for $1 \leq i \leq M_1, 1 \leq j \leq M_2$. Let $\mathbf{r}_{ij} = (r_{ij,0}, r_{ij,1}, ..., r_{ij,M_{0,ij}})$ denote the finer partition of cell C_{ij}, for $1 \leq i \leq M_1, 1 \leq j \leq M_2$. Note that if C_{ij} is empty, $\mathbf{r}_{ij}$ is empty. Further denote $\bar{\mathbf{r}} = (\mathbf{r}_{11}, \mathbf{r}_{12}, ..., \mathbf{r}_{M_1 M_2})$, which

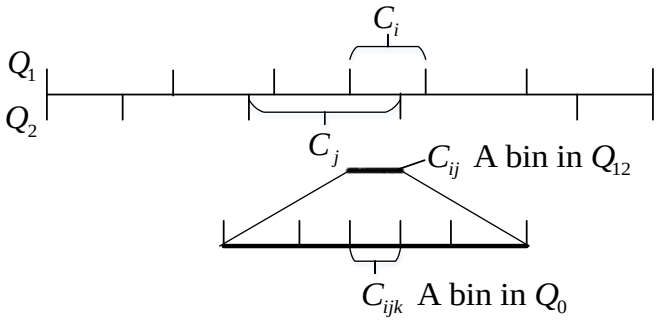


Figure 4.2: Structure of the improved MMDSQ.

represents the full partition at the central quantizer. Define

$$\mathcal{L}(\mathbf{s},\mathbf{t},\bar{\mathbf{r}}) \triangleq (1-\mu_1-\mu_2)D_0(\mathbf{Q}) + \mu_1 D_1(\mathbf{Q}) + \mu_2 D_2(\mathbf{Q}) \\ + \lambda_0 R(\mathbf{Q}) + \lambda_1 R_1(\mathbf{Q}) + \lambda_2 R_2(\mathbf{Q}) \tag{4.1}$$

for some $\mu_1, \mu_2, \lambda_0, \lambda_1, \lambda_2 > 0$ and $\mu_1 + \mu_2 < 1$, where $R(\mathbf{Q}) \triangleq R_0(\mathbf{Q}) + R_1(\mathbf{Q}) + R_2(\mathbf{Q})$. Then the problem of finding the optimal improved-MMDSQ is to minimize the weighted sum of distortions and rates over $\mathbf{s}, \mathbf{t}, \bar{\mathbf{r}}$ as follows,

$$\min_{\mathbf{s},\mathbf{t},\bar{\mathbf{r}}} \mathcal{L}(\mathbf{s},\mathbf{t},\bar{\mathbf{r}}). \tag{4.2}$$

Any tuple $\big(R_1(\mathbf{Q}), R_2(\mathbf{Q}), R(\mathbf{Q}), D_0(\mathbf{Q}), D_1(\mathbf{Q}), D_2(\mathbf{Q})\big)$ on the convex hull of the set of all possible tuples can be found by solving problem (4.2) for some $\mu_1, \mu_2, \lambda_0, \lambda_1, \lambda_2 > 0$ and $\mu_1 + \mu_2 < 1$.

4.3 Improved MMDSQ Design

For any given interval $(a, b]$, the best decoder has to choose the reconstruction value $\hat{x}(a, b]$ minimizing the distortion for cell $(a, b]$, i.e.,

$$\hat{x}(a, b] = \arg \min_{\hat{x}_i \in \hat{\mathcal{X}}} \mathbb{E}[d(X, \hat{x}_i) | X \in (a, b]].$$

The distortions can be represented as

$$D_1(Q) = \sum_{i=1}^{M_1} d(C_i), \quad D_2(Q) = \sum_{j=1}^{M_2} d(C_j), \quad D_0(Q) = \sum_{ij \in M_1 \times M_2} \sum_{k=1}^{M_{0,ij}} d(C_{ijk}). \tag{4.3}$$

where $d(C)$, for any $C \subseteq \mathcal{X}$, is defined as

$$d(C) \triangleq \mathbb{E}[d(X, \hat{x}(C)) | X \in C] \tag{4.4}$$

Denote by I, J, K the random variable taking as value i, j and k, respectively. Using entropy coding, the rates can be written as:

$$R_1(Q) = H(I) = \sum_{i=1}^{M_1} h(C_i), \quad R_2(Q) = H(J) = \sum_{j=1}^{M_2} h(C_j), \tag{4.5}$$

$$R_0(Q) = H(IJK) - H(IJ) = \sum_{ij \in M_1 \times M_2} \Big(\sum_{k=1}^{M_{0,ij}} h(C_{ijk}) - h(C_{ij}) \Big),$$

where for any cell $C \subseteq \mathcal{X}$, $P(C)$ and $h(C)$ are defined as

$$P(C) \triangleq P(x \in C), \quad h(C) \triangleq -P(C) \log_2 P(C). \tag{4.6}$$

By plugging (4.3) and (4.5) into (4.1), we have the complete objective function as

$$\begin{aligned}\mathcal{L}(\mathbf{s},\mathbf{t},\bar{\mathbf{r}}) \triangleq & \sum_{i=1}^{M_1}\Big(\mu_1 d(C_i)+(\lambda_0+\lambda_1)h(C_i)\Big) \\ & +\sum_{j=1}^{M_2}\Big(\mu_2 d(C_j)+(\lambda_0+\lambda_2)h(C_j)\Big) \\ & +\sum_{ij\in M_1\times M_2}\Big(-\lambda_0 h(C_{ij})+\underbrace{\sum_{k=1}^{M_{0,ij}}(1-\mu_1-\mu_2)d(C_{ijk})+\lambda_0 h(C_{ijk})}_{\pi(C_{ij},\mathbf{r}_{ij})}\Big).\end{aligned} \tag{4.7}$$

For fixed C_i, C_j, the intersection set C_{ij} is fixed, then the partition $\mathbf{r}_{ij}$ of C_{ij} can be optimized by minimizing the sub-cost $\pi(C_{ij},\mathbf{r}_{ij})$. Notice that the sub-cost $\pi(C_{ij},\mathbf{r}_{ij})$ depends on the cell C_{ij}, not on the index ij. Therefore, for any given cell $(x_u,x_v]$ with $u<v$, let $\mathbf{r}^*_{x_u,x_v}$ denote the optimal partition of cell $(x_u,x_v]$ as

$$\mathbf{r}^*_{x_u,x_v} \triangleq \arg\min_{\mathbf{r}_{x_u,x_v}\in\mathcal{T}_{x_u,x_v}} \pi\big((x_u,x_v],\mathbf{r}_{x_u,x_v}\big). \tag{4.8}$$

where $\mathcal{T}_{x_u,x_v}$ is defined in Section 3.2.1 as the set of all ascending n-sequence $\mathbf{r}$ with $r_0=x_u$, $r_{n-1}=x_v$ for all $n\geq 2$. The optimization problem (4.8) can be formulated as an all-pairs MWP problem in the WDAG $G(\omega)$ based on the DAG G, while the weight of each edge (m,n) is

$$\omega(m,n) \triangleq (1-\mu_1-\mu_2)d((x_m,x_n])+\lambda_0 h((x_m,x_n]). \tag{4.9}$$

Any partition $\mathbf{r}=(r_0,r_1,...,r_M)$ of $(x_u,x_v]$ for any integer $M>1$, corresponds to a path from source node u to final node v. This correspondence is one-to-one. The weight of the path equals the cost in (4.8), i.e., $\sum_{i=1}^{M}\omega(r_{i-1},r_i)=\pi((x_u,x_v],\mathbf{r}_{x_u,x_v})$. It follows

that finding the optimal $\mathbf{r}^*_{x_u,x_v}$ is equivalent to finding the MWP from node u to node v in WDAG $G(\omega)$. Recall that $\hat{W}_u(v)$ denotes the weight of the MWP from u to v. We have

$$\hat{W}_u(v) = \min_{u<v'<v} \{\hat{W}_u(v') + \omega(v', v)\}, \tag{4.10}$$

where $\omega(v', v)$ is defined in (4.9). Solving the MWP problem for all possible (u, v) takes $O(N^3)$ times if the weight for each edge in $G(\omega)$ can be accessed in constant time.

Since $\pi\big((x_u, x_v], \mathbf{r}^*_{x_u,x_v}\big)$ can be pre-calculated for any cell $(x_u, x_v]$, the optimization problem (4.2) reduces to solving the minimization over all sequences $\mathbf{s}$ and $\mathbf{t}$. Notice that the cost function (4.7) becomes a function of only $\mathbf{s}$ and $\mathbf{t}$. To optimize Q_1, Q_2, Q_{12} simultaneously, we will convert the problem to the MWP problem in a coupled quantizer graph.

Consider the coupled quantizer graph $\mathbb{G}(w)$ based on the DAG $\mathbb{G}$, defined in Chapter 2.3, with the weight function w for two types of edges defined as follows,

$$\begin{aligned} w(uv, u'v) = {} & \mu_1 d(u, u'] + (\lambda_0 + \lambda_1) h(u, u'] \\ & - \lambda_0 h(u, \min(u', v)] + \pi((u, \min(u', v)], \mathbf{r}^*_{u,\min(u',v)}), \end{aligned} \tag{4.11}$$

$$\begin{aligned} w(uv, uv') = {} & \mu_2 d(v, v'] + (\lambda_0 + \lambda_2) h(v, v'] \\ & - \lambda_0 h(v, \min(v', u)] + \pi((v, \min(v', u)], \mathbf{r}^*_{v,\min(v',u)}). \end{aligned} \tag{4.12}$$

Note that the coupled quantizer graph is used to solve the fixed-rate 2DSQ problem in [6], while we use it to find the optimal partitions in order to improve the MMDSQ design. The differences reside in that the weight function is different from the weight functions of [6], which further results in different time complexities.

The MWP in $\mathbb{G}(w)$ from 00 to NN corresponds to the optimal partitions for quantizers Q_1, Q_2, Q_{12} minimizing (4.7). Note that the way to map (Q_1, Q_2, Q_{12}) to the WDAG $\mathbb{G}(w)$ and the algorithm to find the single-source MWP have been presented in Section 2.3. Therefore, the MWP in $\mathbb{G}(w)$ from the source node to the final node corresponds to the sequence $\mathbf{s}, \mathbf{t}, \bar{\mathbf{r}}$ minimizing (4.7).

To make sure the weight for each edge in graph $G(\omega)$ and $\mathbb{G}(w)$ can be accessed in constant time, a pre-processing procedure is needed like the one in Section 3.3.2. Assuming the distortion measure takes mean square error, the cumulative moments $\sum_{i=1}^{c} x_i^j p(x_i)$ for $j = 0, 1, 2$ is computed and stored for all $1 \leq c \leq N$. Then for any possible cell $(x_c, x_{c'}]$ with $c' > c$, relevant cumulative moments can be computed using $\sum_{i=1}^{c'} x_i^j p(x_i) - \sum_{i=1}^{c} x_i^j p(x_i)$ for all $j = 0, 1, 2$ such that $d(C)$ (4.4) and $h(C)$ (4.6) can be computed in constant time for any cell $C \subseteq \mathcal{X}$. To summarize, there are two steps in the optimization process:

- Solve the all-pairs MWP problem in WDAG $G(\omega)$ using the recursion in (4.10) with the weight function defined in (4.9). Store the minimum cost and last cutting point for each pair (u, v) with $0 \leq u < v \leq N$. The time complexity in this step is $O(N^3)$.

- Find the MWP from node 00 to NN in WDAG $\mathbb{G}(w)$ using the algorithm in Algorithm 2 with the weight defined in (4.11) and (4.12). This step requires $O(N^3)$ time.

The total time complexity is $O(N^3)$. Back-tracking is used when reconstructing the whole paths in graph $G(\omega)$ and $\mathbb{G}(w)$.

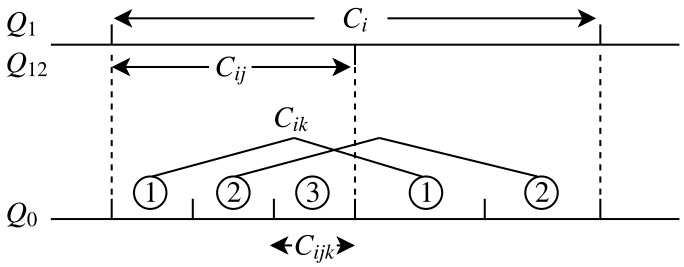


Figure 4.3: The improved MMDSQ with enhanced decoders.

4.4 Improved MMDSQ Design with Enhanced Decoders

In this section, we present a variant of the improved MMDSQ, which uses the refinement from the second stage to improve the reconstruction at the side decoders. This idea is inspired by Liu and Zhu [15]. Instead of applying first entropy coding conditioned on the joint partition for the second stage followed by splitting the bitstream evenly between the two descriptions, we first split the refinement to both side partitions then use entropy coding conditioned on the corresponding side partitions. This way finer reconstruction values can be computed at the side decoders.

Fig. 4.3 depicts an example of the cell partitions where C_i in Q_1 contains two joint cells. Each joint cell C_{ij} is further divided into three or two finer cells C_{ijk} by Q_0. In the improved MMDSQ design, the side decoders are only able to decode the coarse cell C_i or C_j because the refinement information K is encoded conditioned on the joint cell C_{ij}. In the variant design, we encode all the refinement conditioned on the first partition C_i and sent all the refinement to the first decoder one half of the time. During the other half of time, we encode all the refinement conditioned on the second partition C_j and send all of them to the second decoder. This way the side decoder receiving the refinement is able to identify finer cells C_{ik} (C_{jk}) as shown in Fig. 4.3, which are unions of the cells C_{ijk} with

the same index ik (jk). Then the average distortion at each side decoder equals the mean of the distortions with and without refinement.

Although the modification in the entropy coding of the refinement message will increase the total rate, this rate loss will be compensated by the improvement in the side distortions. Note that this procedure is applied to the improved MMDSQ method optimized as in the previous section. Thus, we can narrow the gap to the theoretical bound without changing the optimization process.

Let the modified rate of the first and the second description be denoted by $R_1'(\mathbf{Q})$ respectively, $R_2'(\mathbf{Q})$. Let $\bar{R}_1(\mathbf{Q})$ and $\bar{R}_2(\mathbf{Q})$ represent the average rate of each description. Denote the improved distortion at two decoders by $D_1'(\mathbf{Q})$, respectively, $D_2'(\mathbf{Q})$. Let $\bar{D}_1(\mathbf{Q})$ and $\bar{D}_2(\mathbf{Q})$ denote the average distortion at each side decoder. Note that R_0 is not needed in the variant design, since all the refinement has been encoded with the side partitions. Then the rates can be written as follows,

$$
\begin{aligned}
R_1'(\mathbf{Q}) &= H(IK) = \sum_{i=1}^{M_1} \sum_{k=1}^{M_{0,i}} h(C_{ik}), \\
R_2'(\mathbf{Q}) &= H(JK) = \sum_{j=1}^{M_2} \sum_{k=1}^{M_{0,j}} h(C_{jk}), \\
\bar{R}_1(\mathbf{Q}) &= \big(R_1(\mathbf{Q}) + R_1'(\mathbf{Q})\big)/2, \\
\bar{R}_2(\mathbf{Q}) &= \big(R_2(\mathbf{Q}) + R_2'(\mathbf{Q})\big)/2,
\end{aligned}
\tag{4.13}
$$

where the $R_1(\mathbf{Q})$ and $R_2(\mathbf{Q})$ are defined in (4.5), and $h(C)$ any $C \subseteq \mathcal{X}$ is defined in (4.6). $M_{0,i}$ is the number of cell C_{ik} within each cell C_i, and $M_{0,j}$ is the number of cell C_{jk} within each cell C_j.

Correspondingly, the distortions can be written as

$$
\begin{aligned}
D_1'(\mathbf{Q}) &= \sum_{i=1}^{M_l} \sum_{k=1}^{M_{0,i}} d(C_{ik}), \\
D_2'(\mathbf{Q}) &= \sum_{j=1}^{M_l} \sum_{k=1}^{M_{0,j}} d(C_{jk}), \\
\bar{D}_1(\mathbf{Q}) &= \big(D_1(\mathbf{Q}) + D_1'(\mathbf{Q})\big)/2, \\
\bar{D}_2(\mathbf{Q}) &= \big(D_2(\mathbf{Q}) + D_2'(\mathbf{Q})\big)/2,
\end{aligned}
\tag{4.14}
$$

where $d(C)$ for any $C \subseteq \mathcal{X}$ is defined in (4.4).

4.5 Experimental Result

In this section, we assess the performance of the improved MMDSQ in the symmetric case where both side quantizers have the same weights, i.e., $\mu_1 = \mu_2$, $\lambda_1 = \lambda_2$ in (4.1). In our experiments the source X is obtained in the same way as in Section 3.6, i.e. by discretizing a continuous Gaussian variable $\tilde{X}$ with mean 0 and variance 1. The size of source alphabet $\mathcal{X}$ is 1000. The distortion measure is the squared distance and $\hat{\mathcal{X}} = \mathbb{R}$. We compared our two methods with the results of ECMDSQ [34], MMDSQ [30] and enhanced MMDSQ [15].

Each triple of parameters $(\mu_1, \lambda_0, \lambda_1)$ corresponds to one RD tuple $(R_1, R_2, R, D_0, D_1, D_2)$. The results shown in Fig. 4.4 and Fig. 4.5 have the average rate per description equals $1, 2$ and 3 bits, respectively. It can be observed that our method achieves more points than the MMDSQ especially at low rates, specifically, only two trade-off points can be obtained by MMDSQ at 1 or 2 bits/description. In addition, our method always achieves

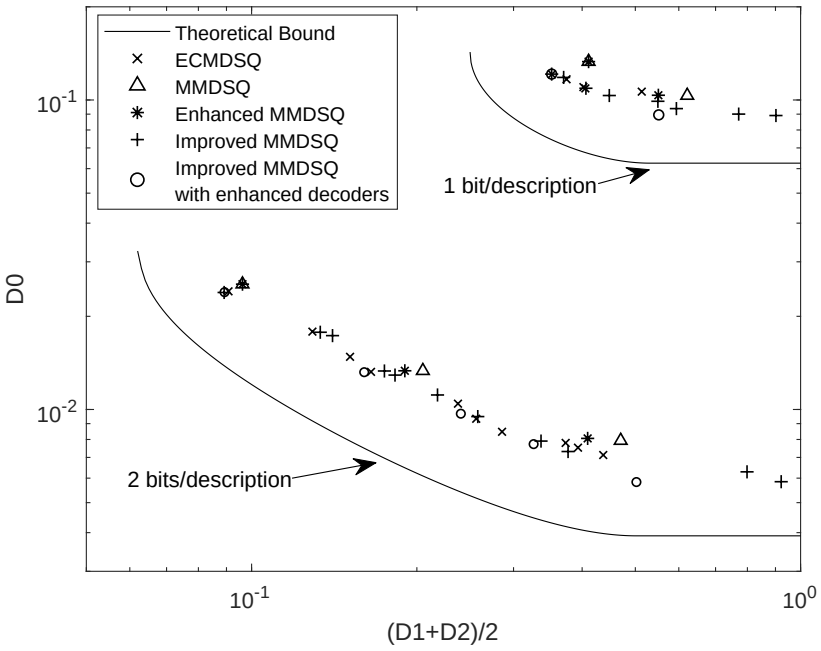


Figure 4.4: Performance at rate 1 bit/description (top) and 2 bits/description (bottom). The index assignment matrix size in ECMDSQ is $\sqrt{M} = 4$ and 8, respectively.

a performance no worse than MMDSQ. At low rates, i.e., 1 or 2 bits/description, the improved MMDSQ achieves significantly lower central distortion D_0 for the same average side distortion $\frac{(D_1+D_2)}{2}$. At high rate, i.e., 3 bits/description, the difference is smaller but still noticeable.

Comparing to ECMDSQ, the improved MMDSQ shows a competitive performance. The points achieved by our improved MMDSQ situate closely to the points achieved by ECMDSQ at all rates. Similar to ECMDSQ, the central distortion D_0 achieved by improved MMDSQ gradually decrease from the biggest to the smallest value at all rates. ECMDSQ has a slight advantage in the case with small D_1 and big D_0 because our side decoders do not use the message from the second stage, while in ECMDSQ the side decoders use all the available messages. However, our improved MMDSQ has a tendency to achieve better

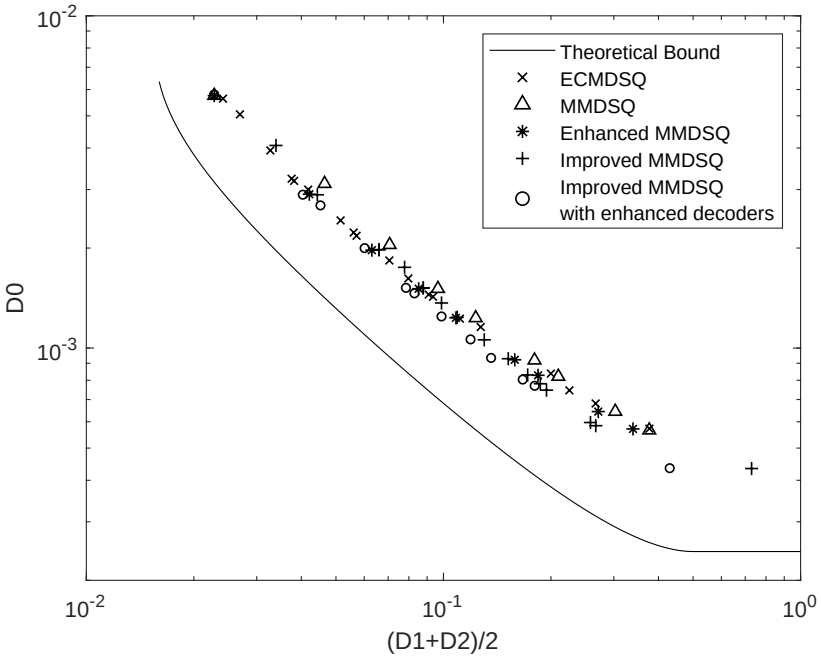


Figure 4.5: Performance at 3 bits/description. ECMDSQ has the index assignment matrix size $\sqrt{M} = 16$.

results when D_0 is really small and D_1 is relatively big. This may be attributed to the index assignment procedure used in ECMDSQ, which restricts the achievable final partitions. Finding the optimal index assignment for ECMDSQ is still a open problem.

The variant of our method always achieves the best performance among all the five methods, although the difference is small at low rates as shown in Fig. 4.4 and 4.5. At high rate, it is clear to see all the circles representing our variant design situate below all the other points when D_0 is small. However, the number of achieved tradeoff points is limited when the rate is low. Therefore, our improved MMDSQ with enhanced decoders is more suitable to be used at high rates where the effect of rate loss is negligible compared to the effect of improvement in the side distortions.

4.6 Conclusion

This chapter proposes the improved MMDSQ design. The improvement resides in using optimized encoder partitions. The optimization problem is formulated as a minimization of a weighted sum of the rates and distortions aiming at those RD tuples on the lower convex hull of the achievable RD region. The improved MMDSQ can guarantee global optimality for finite-alphabet source when all involved scalar quantizers have contiguous cells. The optimization is based on solving the MWP problem in a WDAG and the single-source MWP problem in a coupled quantizer graph. The continuum tradeoffs are achieved by varying the weights in the objective function. We also present a variant of the improved MMDSQ with enhanced decoders, which uses the refinement information at the second stage to improve the distortion at side decoders. This variant design effectively narrows the gap to the theoretical bound at high rates as shown in the experimental results.

Chapter 5

Conclusion and Future Work

This work proposes graph-based solutions for two scalar quantization problems in network systems, namely, two-stage Wyner-Ziv coding problem and multiple description coding problem. Both problems are formulated as the minimization of rates and distortions such that all points lying on the lower convex hull of the theoretical rate-distortion region can be found.

The first design is for the two-stage Wyner-Ziv problem with forwardly/reversely degraded side information (SI). We assume that the binning is performed on blocks of infinite length such that the limit in Slepian-Wolf theorem can be achieved. We present two kinds of encoding structures based on scalar quantization for each SI degrading scenario and focus on finding the optimal partitions. The proposed solution is based on solving the single-source and all-pairs minimum-weight path (MWP) problem in some weighted directed acyclic graphs (WDAG). The time complexity achieved by conventional dynamic programming is $O(N^3)$, where N is the size of the source alphabet. Furthermore, a partial Monge property is proposed to expedite the solution of the all-pairs MWP problem in a WDAG. A thorough algorithm exploiting this property is presented. Through our extensive

experimental results, it is clear to see that the partial Monge property holds in many situations of interest and that our coding scheme can achieve expected performance compared to the rate-distortion (RD) bound. As a byproduct, we also obtain a solution for the traditional Heegard-Berger problem. The proposed design guarantees global optimality when the quantizer cells are contiguous.

The other design for multiple description coding problem is an improved scheme for the modified multiple description scalar quantizer (MMDSQ) [30], termed improved MMDSQ. The improvement consists in that we replace the uniform partitions by the optimal ones which minimize a weighted sum of the rates and distortions. The solution algorithm is based on solving the all-pairs MWP problem in a WDAG and a single-source MWP problem in a coupled quantizer graph. We also propose a variant of the improved MMDSQ with enhanced decoders, which uses the refinement message at the second stage to improve distortions at side decoders. The improved MMDSQ achieves a performance close to ECMDSQ at all rates. Meanwhile, it has a significantly better performance than MMDSQ at low rates, while at high rates both designs achieve similar results. The variant design obtains the smallest gap to the theoretical bounds among all the other considered designs at all rates.

Although the work shows good performances of the graph-based solutions for both problems, there are still some aspects that can be further explored.

- An interesting direction for future work is to investigate theoretically if the partial Monge property holds for general sources and SI or to derive sufficient conditions under which it is satisfied.

- It is also interesting to consider extending the WZ encoding scheme to more than two stages.

- Our variant of the improved MMDSQ design just simply modifies the decoding rule after the partitions are obtained. It is interesting to incorporate the modified distortions in the optimization to generate the optimal quantizers under the modified decoding rule.

Appendix A

Proof of Proposition 1. We have to show that the following holds for all $1 \leq m \leq m' < n \leq n' \leq N$,

$$\omega'(m,n) + \omega'(m',n') \leq \omega'(m,n') + \omega'(m',n). \tag{A.1}$$

If $n - m' < T_1 - 1$ then $\omega'(m',n) = \infty$, while if $n' - m > T_2 + 1$ then $\omega'(m,n') = \infty$. In either case the right hand side of (A.1) equals ∞, thus the relation is satisfied. It remains to consider the case when $n - m' \geq T_1 - 1$ and $n' - m \leq T_2 + 1$. In this case all quantities in (A.1) are real values. Note that if $m = m'$ or $n = n'$ the relation is trivially satisfied. Therefore, let us assume that $m < m'$ and $n < n'$. For any k such that $m \leq k < m'$, denote

$$\Delta(k,n,n') \triangleq \omega'(k,n') + \omega'(k+1,n) - \omega'(k,n) - \omega'(k+1,n').$$

The quantities appearing on the right hand side of the above equation are all real values, therefore, the expression is well defined. Further, we have

$$\Delta(k,n,n') = \omega(k,n') + \omega(k+1,n) - \omega(k,n) - \omega(k+1,n') = \sum_{j=n}^{n'-1} \Delta_\omega(k,j).$$

For $m \leq k \leq m' - 1$ and $n \leq j \leq n' - 1$, we have

$$T_1 \leq n - m' + 1 \leq j - k \leq n' - m - 1 \leq T_2.$$

Then we have $\Delta_\omega(k, j) \geq 0$ in virtue of (3.17). It follows that $\Delta(k, n, n') \geq 0$ and further, that

$$\omega'(m, n') + \omega'(m', n) - \omega'(m, n) - \omega'(m, n') = \sum_{k=m}^{m'-1} \Delta(k, n, n') \geq 0.$$

This observation completes the proof. □

Proof of FP. When $g(a, d) = \infty$ the claim holds trivially. Let us assume now that $g(a, d) \neq \infty$. Since the inequality $g(b, c) < g(a, c)$ is strict, we have $g(b, c) \neq \infty$. The Monge property

$$g(a, c) + g(b, d) \leq g(b, c) + g(a, d) \tag{A.2}$$

further implies that $g(a, c) \neq \infty$ and $g(b, d) \neq \infty$. Then (A.2) is equivalent to

$$g(a, c) - g(b, c) \leq g(a, d) - g(b, d).$$

The expression on the left hand side is strictly positive according to the hypothesis, thus $g(a, d) - g(b, d) > 0$, proving the claim. □

Bibliography

[1] Burkard, R. E., Klinz, B., and Rudolf, R. (1996). Perspectives of Monge properties in optimization. *Discrete Applied Mathematics*, **70**(2), 95–161.

[2] Cheng, S. and Xiong, Z. (2005). Successive refinement for the Wyner-Ziv problem and layered code design. *IEEE Transactions on Signal Processing*, **53**(8), 3269–3281.

[3] Chou, P. A., Lookabaugh, T., and Gray, R. M. (1989). Entropy-constrained vector quantization. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **37**(1), 31–42.

[4] Dumitrescu, S. and Wu, X. (2002). Optimal multiresolution quantization for scalable multimedia coding. In *Proc. IEEE Inf. Theory Workshop (ITW)*, pages 139–142, Bangalore, India.

[5] Dumitrescu, S. and Wu, X. (2004). Algorithms for optimal multi-resolution quantization. *J. Algorithms*, **50**(1), 1–22.

[6] Dumitrescu, S. and Wu, X. (2005). Optimal two-description scalar quantizer design. *Algorithmica*, **41**(4), 269–287.

[7] Dumitrescu, S. and Wu, X. (2007). Lagrangian optimization of two-description scalar quantizers. **53**(11), 3990–4012.

[8] Fan, X., Au, O. C., Cheung, N. M., Chen, Y., and Zhou, J. (2010). Successive refinement based Wyner–Ziv video compression. *Signal Processing: Image Communication*, **25**(1), 47–63.

[9] Fleming, M., Zhao, Q., and Effros, M. (2004). Network vector quantization. *IEEE Transactions on information Theory*, **50**(8), 1584–1604.

[10] Frank-Dayan, Y. and Zamir, R. (2002). Dithered lattice-based quantizers for multiple descriptions. *IEEE Transactions on Information Theory*, **48**(1), 192–204.

[11] Gamal, A. E. and Cover, T. (1982). Achievable rates for multiple descriptions. *IEEE Transactions on Information Theory*, **28**(6), 851–857.

[12] Heegard, C. and Berger, T. (1985). Rate distortion when side information may be absent. *IEEE Transactions on information Theory*, **31**(6), 727–734.

[13] Hirschberg, D. S. and Larmore, L. L. (1987). The least weight subsequence problem. *SIAM J. Computing*, **16**(4), 628–638.

[14] Kaspi, A. H. (1994). Rate-distortion function when side-information may be present at the decoder. *IEEE Transactions on information Theory*, **40**(6), 2031–2034.

[15] Liu, M. and Zhu, C. (2009). Enhancing two-stage multiple description scalar quantization. *IEEE Signal Processing Letters*, **16**(4), 253–256.

[16] Luenberger, D. G. and Ye, Y. (2008). *Linear and nonlinear programming*. Springer, New York, third edition.

[17] Majumdar, A. and Ramchandran, K. (2004). Video multicast over lossy channels

based on distributed source coding. In *Proc. IEEE Int. Conf. Image Process. (ICIP)*, pages 3093–3096, Singapore.

[18] Muresan, D. and Effros, M. (2002). Quantization as histogram segmentation: globally optimal scalar quantizer design in network systems. In *Proc. Data Compress. Conf. (DCC)*, pages 302–311, Snowbird, UT.

[19] Muresan, D. and Effros, M. (2008). Quantization as histogram segmentation: optimal scalar quantizer design in network systems. *IEEE Transactions on information Theory*, **54**(1), 344–366.

[20] Ozarow, L. (1980). On a source-coding problem with two channels and three receivers. *The Bell System Technical Journal*, **59**(10), 1909–1921.

[21] Pradhan, S. S. and Ramchandran, K. (2003). Distributed source coding using syndromes (DISCUS): Design and construction. *IEEE Transactions on information Theory*, **49**(3), 626–643.

[22] Ramanan, S. and Walsh, J. M. (2011). Practical codes for lossy compression when side information may be absent. In *Proc. IEEE Int. Conf. Acoust., Speech and Signal Process. (ICASSP)*, pages 3048–3051, Prague, Czech Republic.

[23] Rebollo-Monedero, D., Zhang, R., and Girod, B. (2003). Design of optimal quantizers for distributed source coding. In *Proc. Data Compress. Conf. (DCC)*, pages 13–22, Snowbird, UT.

[24] Shamai, S., Verdú, S., and Zamir, R. (1998). Systematic lossy source/channel coding. *IEEE Transactions on information Theory*, **44**(2), 564–579.

[25] Shi, J., Liu, L., Gündüz, D., and Ling, C. (2017). Polar codes and polar lattices for the Heegard-Berger problem. *arXiv preprint arXiv:1702.01042*.

[26] Slepian, D. and Wolf, J. (1973). Noiseless coding of correlated information sources. *IEEE Transactions on information Theory*, **19**(4), 471–480.

[27] Steinberg, Y. and Merhav, N. (2004). On successive refinement for the Wyner-Ziv problem. *IEEE Transactions on information Theory*, **50**(8), 1636–1654.

[28] Tian, C. and Diggavi, S. N. (2007). On multistage successive refinement for Wyner–Ziv source coding with degraded side informations. *IEEE Transactions on information Theory*, **53**(8), 2946–2960.

[29] Tian, C. and Diggavi, S. N. (2008). Side-information scalable source coding. *IEEE Transactions on information Theory*, **54**(12), 5591–5608.

[30] Tian, C. and Hemami, S. S. (2005). A new class of multiple description scalar quantizer and its application to image coding. *IEEE Signal Processing Letters*, **12**(4), 329–332.

[31] Timo, R., Chan, T., and Grant, A. (2011). Rate distortion with side-information at many decoders. *IEEE Transactions on information Theory*, **57**(8), 5240–5257.

[32] Unal, S. and Wagner, A. B. (2017). An LP upper bound for rate distortion with variable side information. In *Proc. IEEE Data Compress. Conf. (DCC)*, pages 370–379, Snowbird, UT.

[33] Vaishampayan, V. A. (1993). Design of multiple description scalar quantizers. *IEEE Transactions on Information Theory*, **39**(3), 821–834.

[34] Vaishampayan, V. A. and Domaszewicz, J. (1994). Design of entropy-constrained multiple-description scalar quantizers. *IEEE Transactions on Information Theory*, **40**(1), 245–250.

[35] Wang, H. and Ortega, A. (2004). Scalable predictive coding by nested quantization with layered side information. In *Proc. IEEE Int. Conf. Image Process. (ICIP)*, pages 1755–1758, Singapore.

[36] Wang, J., Majumdar, A., and Ramchandran, K. (2005). On enhancing mpeg video broadcast over wireless networks with an auxiliary broadcast channel. In *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Process. (ICASSP)*, pages 1101–1104, Philadelphia, PA.

[37] Wilber, R. E. (1988). The concave least-weight subsequence problem revisited. *J. Algorithms*, **9**(3), 418–425.

[38] Wu, X. and Dumitrescu, S. (2002). On optimal multi-resolution scalar quantization. In *Proc. Data Compress. Conf. (DCC)*, pages 322–331, Snowbird, UT.

[39] Wu, X. and Zhang, K. (1993). Quantizer monotonicities and globally optimal scalar quantizer design. *IEEE Transactions on information Theory*, **39**(3), 1049–1053.

[40] Wyner, A. and Ziv, J. (1976). The rate-distortion function for source coding with side information at the decoder. *IEEE Transactions on information Theory*, **22**(1), 1–10.

[41] Xiong, Z., Liveris, A. D., and Cheng, S. (2004). Distributed source coding for sensor networks. *IEEE Signal Process. Mag.*, **21**(5), 80–94.

[42] Xu, Q. and Xiong, Z. (2006). Layered Wyner–Ziv video coding. *IEEE Transactions on information Theory*, **15**(12), 3791–3803.

[43] Zheng, Q. and Dumitrescu, S. (2018a). Improved two-stage multiple description scalar quantizer. in preparation.

[44] Zheng, Q. and Dumitrescu, S. (2018b). Optimal design of a two-stage wyner-ziv scalar quantizer with degraded side information. presented at 29th Biennial Symposium on Communications (BSC), Toronto, Canada.

[45] Zheng, Q. and Dumitrescu, S. (2018c). Optimal design of a two-stage wyner-ziv scalar quantizer with forwardly/reversely degraded side information. submitted.

[46] Ziv, J. (1985). On universal quantization. *IEEE Transactions on information Theory*, **31**(3), 344–347.