Optimal Two-description Scalar Quantizer Design

Sorina Dumitrescu
Department of Electrical and Computer Engineering
McMaster University
Hamilton, ON, Canada
sorina@mail.ece.mcmaster.ca

Xiaolin Wu *
Department of Computer and Information Science
Polytechnic University
Brooklyn, New York, NY 11201
xwu@poly.edu

Abstract

Multiple description quantization is a signal compression technique for robust networked multimedia communication. In this paper we consider the problem of optimally quantizing a random variable into two descriptions, while each description being produced by a side quantizer of convex codecells. The optimization objective is to minimize the expected distortion given the probabilities of receiving either and both descriptions. The problem is formulated as one of shortest path in a weighted directed acyclic graph with constraints on the number and types of edges. An $O(K_1K_2N^3)$ time algorithm for designing the optimal two-description quantizer is presented, where N is the cardinality of the source alphabet, and K_1, K_2 are the number of codewords of the two quantizers respectively. This complexity is reduced to $O(K_1K_2N^2)$ by exploiting the Monge property of the objective function. Furthermore, if $K_1 = K_2 = K$ and the two descriptions are transmitted through two channels of the same statistics, then the optimal two-description quantizer design problem can be solved in $O(KN^2)$ time.

Key words: *Quantization, multiple description signal representation, multimedia communications, Monge property, matrix search, k-link shortest path.*

*Supported by NSF Grant CCR-0208678.

1 Introduction

Quantization is a common technique of compressing multimedia signals such as images, video and audio [5]. Data compression is achieved by quantizing signal samples from a representation of higher resolution to lower resolution so that fewer bits suffice to code each sample of the quantized signal. Optimal quantization falls into the class of resource allocation problems in operational research. The central issue is how to describe a random variable X or a random vector $\mathbf{X}$ to the maximum precision (or minimum distortion) possible using a given number of bits. The problem is called optimal scalar or vector quantizer design depending on whether the input is a random variable or a random vector. This paper is restricted to the treatment of scalar quantization.

In conventional single description quantization, a quantized signal is coded and transmitted in a single bit stream through a communication channel. If the channel fails then the reconstruction of the signal will be necessarily interrupted or abandoned at the receiver. The modern IP communication networks, however, offer multiple routes between any two nodes. This design of distributed communication can be utilized to improve the error resilience of conventional quantizers. Multiple description or networked quantization is such a technique [4, 8, 9, 10] of robust data communications.

A multiple description quantizer consists of two or more so-called side quantizers, each of which separately generates a "partial" or coarse description of the input signal (called side description). Each side description is self contained in a sense that it can be decoded autonomously to reconstruct the signal at certain fidelity without the knowledge of other description(s). At the same time the decoders of multiple side quantizers can collaborate to generate a refined joint description of the input signal if more or all side descriptions are received. The fidelity of the joint description increases in general in the number of received descriptions. Figure 1 presents the effect of passing a signal through a two-description quantizer.

Multiple description quantizers are particularly useful in multimedia streaming over lossy packet-switched networks like the internet. They offer a graceful degradation in quality of service in adverse network conditions such as congestions and local outage, as opposed to outright stoppage of network service in a single description of the multimedia content. Other applications are communications using antenna diversity or distributed data storage systems. In this paper we focus on two-description

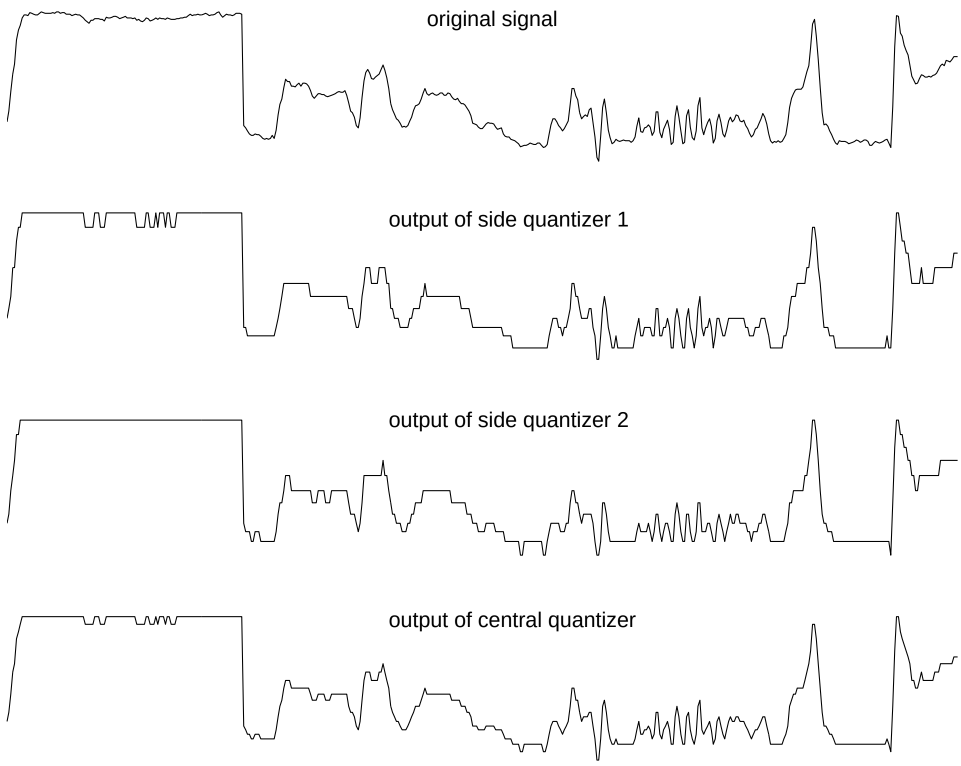

Figure 1: A signal and the corresponding side and joint descriptions as the outputs of a two-description scalar quantizer.

quantizers (2DQ), i.e., the case of two side descriptions. We will also restrict the side quantizers of the 2DQ's to have convex codecells.

This paper is structured as follows. In the next section we present the necessary definitions and notations, and formulate the problem of optimal two-description scalar quantizer (2DQ) design for the class of 2DQ's with convex codecells. In Section 3 we present solutions for this problem. Subsection 3.1 describes how the problem can be modelled as a weighted shortest path problem with constraints on the number and type of edges, and then presents an $O(K_1K_2N^3)$ time algorithm to solve the problem, where N is the size of the source alphabet, and K_1, K_2 are the number of code words of the two side quantizers, respectively. In subsection 3.2 we prove that the time complexity of the solution can be reduced to $O(K_1K_2N^2)$. For this we break the problem into multiple instances of matrix search, and show that each matrix is totally monotone based on the Monge property satisfied by the cost function. As a result, the fast algorithm SMAWK introduced in [1] can be applied to solve each instance of matrix search. This leads to the claimed reduction in time complexity. Section 4 addresses the problem of optimal design of symmetric 2DQ. The symmetry means that the two side quantizers operate at the same rate ($K_1 = K_2 = K$) and have the same probability of contributing to the reconstruction of the input signal. First it is proved in Subsection 4.1 that there is an optimal symmetric 2DQ with the thresholds of the two side quantizers interleaved, a property which can be naturally used to simplify the design algorithm and decrease the time complexity to $O(KN^2)$. This more efficient algorithm is still based on multiple applications of the fast matrix search technique SMAWK of [1]. Subsection 4.2 shows that the optimal symmetric 2DQ design can be done by a new simpler algorithm. The new algorithm avoids the procedure of recursive matrix reduction as in SMAWK, while achieving the same time complexity. This simplified solution hinges again on the Monge property of the cost function.

2 Definitions, Notations, Problem Formulation

Let X be a random variable over a finite alphabet $\mathcal{A} = \{x_1, x_2, \cdots, x_N\} \subset \mathcal{R}$, $x_i < x_{i+1}$, $1 \leq i \leq N-1$. Let the probability mass function (pmf) of X be $p_i = p(X = x_i)$, $1 \leq i \leq N$. A convex subset of the alphabet $\mathcal{A}$ is any set $c(a, b] = \{x_i | a < i \leq b\}$ for some integers a, b, $0 \leq a \leq b \leq N$.

A quantizer can be defined in the most general way as an arbitrary partition of the alphabet $\mathcal{A}$. However, in this paper we are concerned only with quantizers whose partition consists of convex subsets of the alphabet $\mathcal{A}$, and we will impose this condition throughout the paper. Consequently, a quantizer Q is defined as a partition of the alphabet $\mathcal{A}$ into convex subsets $c(q_j, q_{j+1}]$ (called codecells), $0 \leq j \leq K-1$, where $0 = q_0 < q_1 < \cdots q_{K-1} < q_K = N$, for some K, $K < N$. Note that the intervals of integers $(q_j, q_{j+1}]$, $0 \leq j \leq K-1$, form a partition of $(0, N]$, which we denote by $(q_0, q_1, \cdots, q_{K-1}, q_K)$. The quantizer Q can be identified by the above partition of the interval of integers $(0, N]$. Hence, we write

$$Q = (q_0, q_1, \cdots, q_{K-1}, q_K). \tag{1}$$

We use interchangeably the terms quantizer and partition. The values q_j are called thresholds of the quantizer (or of the partition). The codecell $c(q_j, q_{j+1}]$ is simply denoted by $(q_j, q_{j+1}]$.

To measure the quantizer performance we consider a distortion function $d(x, y)$, $d : \mathcal{R} \times \mathcal{R} \rightarrow [0, \infty)$, which is monotone, i.e., it observes the following relation

$$d(x, y_1) \leq d(x, y_2), \text{ for all real values } x, y_1, y_2 \text{ such that } x \leq y_1 < y_2 \text{ or } x \geq y_1 > y_2, \tag{2}$$

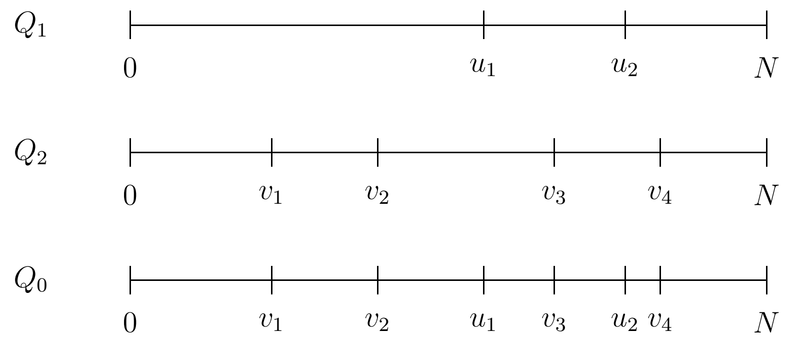

Figure 2: Example of 2DQ. Q_1 and Q_2 are the side quantizers. The central quantizer Q_0 is obtained by intersecting the two side partitions.

which is satisfied by all distortion functions used in practice.

The reproduction codewords assigned to individual codecells are chosen from another finite alphabet $\mathcal{B} = \{y_1, y_2, \cdots, y_M\} \subset \mathcal{R}$, such that $y_i < y_{i+1}$, $1 \leq i \leq M-1$, and $\mathcal{A} \subseteq \mathcal{B}$. Namely, for any codecell $(a, b]$, the reproduction codeword $\mu(a, b]$ is the one which minimizes the distortion, i.e.

$$\mu(a, b] = \arg\min_{y_j \in \mathcal{B}} \sum_{i=a+1}^{b} d(x_i, y_j) p_i. \tag{3}$$

The distortion of the codecell $(a, b]$ is thus

$$D(a, b] = \sum_{i=a+1}^{b} d(x_i, \mu(a, b]) p_i = \min_{y_j \in \mathcal{B}} \sum_{i=a+1}^{b} d(x_i, y_j) p_i. \tag{4}$$

By convention, $D(a, a] = 0$. The distortion of the quantizer Q is the sum of distortions of its codecells, i.e.,

$$D(Q) = \sum_{j=0}^{K-1} D(q_j, q_{j+1}]. \tag{5}$$

A two-description quantizer (2DQ) $\mathbf{Q} = (Q_1, Q_2, Q_0)$ is composed of three quantizers Q_1, Q_2, Q_0. Q_1 and Q_2 are called side quantizers, and Q_0, the central quantizer is obtained by intersecting the two side partitions. Hence, if K_i is the number of codecells of the side quantizer Q_i, $i = 1, 2$, then the central quantizer has at most $K_1 + K_2 - 1$ codecells. Fig. 2 illustrates an example of 2DQ.

Traditionally, the goal of multiple description quantizer design is to minimize the distortion of the central description (the one obtained when both side descriptions are available), while meeting given upper bounds on the side distortions [10]. We formulate the problem as one of minimizing an expected distortion of the 2DQ. Let ω_i denote the probability that only the output of side quantizer Q_i, $i = 1, 2$, is available to reconstruct the signal. Let further ω_0 be the probability that the outputs of both side quantizers are available. In the latter case, the two side descriptions can be refined to the description of Q_0. Finally, the probability of no description received is $1 - \omega_1 - \omega_2 - \omega_0$. Thus, the expected distortion of the 2DQ can be expressed as

$$\bar{D}(\mathbf{Q}) = (1 - \omega_1 - \omega_2 - \omega_0)\sigma^2 + \sum_{i=0}^{2} \omega_i D(Q_i), \tag{6}$$

where σ^2 is $D(0, N]$ (or the variance of X in the case when the distortion function is the squared distance). The problem of optimal 2DQ design can hence be stated as follows.

Problem 1. Construct a 2DQ $\mathbf{Q} = (Q_1, Q_2, Q_0)$, where each side quantizer Q_i has K_i codecells, $i = 1, 2$, such that the expected distortion $\bar{D}(\mathbf{Q})$ is minimized.

Note that in our definition of 2DQ the codecells of the side quantizers are convex. This convexity restriction does not preclude optimality in single description scalar quantization [5] but it may in two-description quantization [10]. However, optimal 2DQ can have convex codecells if the weights ω_1 and ω_2 of the side distortions are much larger than the weight ω_0 of the central distortion. In this case, the emphasis in the minimization of (6) is on lowering the side distortions. Since low side distortions can be achieved with convex codecells, there is no loss of optimality under the convexity constraint when the ratios ω_1/ω_0, ω_2/ω_0 become large enough. To avoid tedious terminology, in the sequel we will simply use 2DQ to refer to 2DQ's with convex codecells.

The problem of optimal scalar 2DQ design was studied by others [6, 7, 9, 10]. The authors of [10] seek to minimize the distortion of the central quantizer while meeting some constraints on the side distortions. The codecells of the side quantizers are not restricted to be convex. However, once an index assignment is chosen, the design algorithm is confined to that assignment. The constrained optimization problem is transformed into a Lagrangian form. The Lagrangian minimization is done by an iterative gradient descent algorithm.

Closer to our formulation is the treatment of [6, 7]. Muresan and Effros [6] addressed the case of convex codecells in the side quantizers, too. They introduced a weighted directed acyclic graph WDAG, called *partial RD graph*, and mapped the problem of optimal entropy-constrained 2DQ design to one of the shortest path in the partial RD graph. The cost to be minimized is a weighted sum of distortions and rates of the side and central quantizers. Since the graph has $O(N^2)$ vertices and $O(N^3)$ edges, the solution can be obtained in $O(N^3)$ time, assuming that the proper Lagrangian multiplier can be found in a constant number of iterations. In [7] Muresan and Effros also treated the case of fixed-rate 2DQ (Problem 1). They showed that the problem was again equivalent to a shortest path problem, but in a more complex WDAG, with $O(K_1K_2N^2)$ vertices and $O(K_1K_2N^3)$ edges. Consequently, the solution can be computed in $O(K_1K_2N^3)$ time.

3 Optimal Scalar 2DQ Design

In this section we first describe the graph model and sketch a simple $O(K_1K_2N^3)$ time algorithm to solve the problem. Then we proceed to reduce the time complexity to $O(K_1K_2N^2)$ by exploiting a monotonicity of the cost function.

3.1 Solution Based on The Graph Model

We map Problem 1 to a constrained shortest path problem in a WDAG simpler than the one of [7]. This WDAG has $O(N^2)$ vertices and $O(N^3)$ edges. Actually, this WDAG has the same structure as the partial RD graph of [6], i.e., the same vertices and edges, only the weights of the edges are different, according to our cost function (they are only weighted sum of distortions, not of distortions and entropies). The constraint on the number of codecells (i.e., on the data rate of the quantizer) becomes a constraint on the number and type of edges, as explained in what follows. Let us construct a WDAG $G = (V, E)$. Each member (a node) of the vertex set V is an ordered pair of integers u and v, $0 \leq u, v \leq N$. Such a node, labelled by uv (the order of u and v matters), corresponds to a pair of thresholds of the two side partitions: u is a threshold of Q_1 and v a threshold of Q_2. The edges of G are of two types. The edges of type I are those from the node uv to the node $u'v$ for all $0 \leq u \leq v \leq N$ and $u < u' \leq N$. The edges of type II are those from the node uv to the node uv' for all $0 \leq v < u \leq N$ and $v < v' \leq N$. Note that the outgoing edges of some vertex uv are all either of type I or of type II, depending on whether $u \leq v$ or $u > v$. The source of the graph is the node 00 and the final node is NN. By a path in the graph G we will understand in the sequel any path from the source to the final node, unless otherwise specified.

Any 2DQ is mapped to a distinct path of G in the following way. Let $Q_1 = (u_0, u_1, \cdots, u_{K_1})$ and $Q_2 = (v_0, v_1, \cdots, v_{K_2})$ be the side quantizers of the 2DQ $\mathbf{Q}$. The path is incrementally constructed by starting from the source 00 (i.e., u_0v_0) and adding one edge at a time. Let the current vertex be

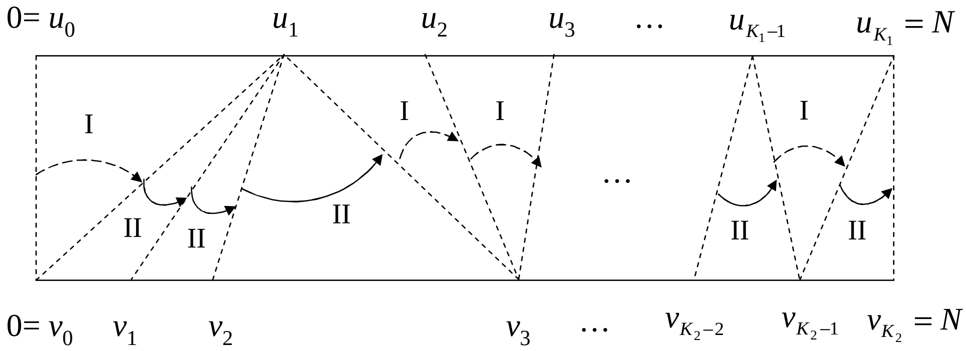

Figure 3: A path in the WDAG G corresponding to a 2DQ, in which dashed lines connecting thresholds of the two side partitions represent the graph nodes; arrowed arcs represent edges whose types are labelled.

u_iv_j (i.e., the path has been constructed up to u_iv_j) for some $i, j, 0 \le i \le K_1$,$0 \le j \le K_2$. Then the next edge to be added is $(u_iv_j, u_{i+1}v_j)$ if $u_i \le v_j$, or it is (u_iv_j, u_iv_{j+1}) if $u_i > v_j$. Note that only one of these two situations is possible. Thus, the path constructed this way is unique. Note that for any threshold u_i (resp. v_j) of the first (resp. second) side quantizer, there is a vertex u_iv (resp. uv_j) visited by the path. Also note that this path has K_1 edges of type I and K_2 edges of type II. Figure 3 illustrates a path in the graph G corresponding to a 2DQ.

Conversely, each path of G generates by this construction a distinct 2DQ in the following way. The thresholds of the first side partition are the elements of the set $S_1 = \{u|uv$ is a vertex in the path$\}$. If the path has K_1 edges of type I, then the above set has $K_1 + 1$ elements, which means that the first side quantizer has K_1 codecells. On the other side, the thresholds of the second side partition are the elements of the set $S_2 = \{v|uv$ is a vertex in the path$\}$. If the path has K_2 edges of type II, then the second side partition has $K_2 + 1$ thresholds or K_2 codecells.

Note that an edge of type I, $(uv, u'v)$, appears in a path corresponding to a 2DQ if and only if $(u, u']$ is a codecell of the first side partition, v is a threshold of the second side partition, and moreover, v is the smallest threshold of Q_2 that is larger or equal to u (the last condition can be proved by contradiction). Hence $(u, \min(v, u')]$ is a codecell of the central partition. Based on these observations, the weight assigned to the edge $(uv, u'v)$ is the contribution to the expected distortion of these two codecells:

$$w(uv, u'v) = \omega_1 D(u, u'] + \omega_0 D(u, \min(v, u')]. \tag{7}$$

Symmetrically, an edge of type II, (uv, uv') appears in a path corresponding to a 2DQ if and only if $(v, v']$ is a codecell of the side quantizer Q_2 and u is the smallest threshold of Q_1 that is strictly larger than v. This implies that $(v, \min(v', u)]$ is a codecell of the central quantizer. Therefore, the weight assigned to this edge is

$$w(uv, uv') = \omega_2 D(v, v'] + \omega_0 D(v, \min(v', u)]. \tag{8}$$

We call the WDAG constructed above *coupled quantizer graph.* As we have seen the coupled quantizer graph associates any 2DQ $\mathbf{Q}$ with a path from the source node 00 to the final node NN. It is easy to verify that the total weight of this path equals $\bar{D}(\mathbf{Q}) - (1 - \omega_1 - \omega_2 - \omega_0)\sigma^2$. Moreover, the number of codecells of the side quantizer Q_i is K_i for $i = 1, 2$, if and only if the corresponding path consists of K_1 edges of type I and K_2 edges of type II. Consequently, Problem 1 can be equivalently formulated as follows.

Problem 2. Find a minimum-weight path among all paths in the coupled quantizer graph G with K_1 edges of type I and K_2 edges of type II.

A solution to Problem 2, without using any properties of the cost function, is to find recursively, for each node uv and each pair of integers s, t, $0 \le s \le K_1$, $0 \le t \le K_2$, a path from the source 00 to uv, with s edges of type I and t edges of type II, whose weight is minimal. Denote by $W_{s,t}(uv)$ the weight of this path. Also denote by $W^I_{s,t}(uv)$ ($W^{II}_{s,t}(uv)$, respectively) the minimal path weight among all paths ending at uv, with s edges of type I and t edges of type II, which have the last edge of type I (type II, respectively). Then the recursions to solve the problem are:

$$\begin{aligned} W_{s,t}(uv) &= \min\{W^I_{s,t}(uv), W^{II}_{s,t}(uv)\}, && (9)\\ W^I_{s,t}(uv) &= \min_{0 \le u'' < u, u" \le v} \{W_{s-1,t}(u''v) + w(u''v, uv)\}, && (10)\\ W^{II}_{s,t}(uv) &= \min_{0 \le v'' < u, v'' < v} \{W_{s,t-1}(uv'') + w(uv'', uv)\}. && (11) \end{aligned}$$

The above computations can be organized by a dynamic programming process such that all pairs s and t (corresponding to the number of edges of the two types) are processed in lexicographical order. Given s and t, the value $W_{s,t}(u, v)$ is found for all nodes uv of the graph. For a given node uv, $W_{s,t}(uv)$ is computed by applying recursions (10), (11) and then (9). This can be done since the values $W_{s-1,t}(u''v)$ and $W_{s,t-1}(uv'')$ have already been computed (the pairs $s-1, t$, and $s, t-1$, precede s, t in the lexicographical order).

The search in each of equations (10) and (11) requires $O(N)$ time, while the search in (9) takes constant time. Hence computing $W_{s,t}(uv)$ for a given pair s, t and a given node uv takes $O(N)$ time. Since the graph G has $O(N^2)$ nodes, and there are $O(K_1K_2)$ pairs of s and t, the total time complexity becomes $O(K_1K_2N^3)$. The time complexity achieved is the same as in [7], but we can show that it can be further reduced.

3.2 Fast Design Algorithm Aided by Monge Property

In this section we reduce the time complexity of optimal 2DQ design thanks to a property of the underlying cost function, called total monotonicity [1].

Let us first reexamine recursion (10). For fixed s, t and v, consider the $(v+1) \times N$ matrix $M_{s,t,v}$ with rows indexed by u'', $0 \le u'' \le v$, columns indexed by u, $1 \le u \le N$, and with elements $M_{s,t,v}(u'', u)$ defined by

$$M_{s,t,v}(u'', u) = \begin{cases} W_{s-1,t}(u''v) + w(u''v, uv) & \text{if } u'' < u \\ \infty & \text{otherwise.} \end{cases} \tag{12}$$

Then relation (10) becomes

$$W^I_{s,t}(uv) = \min_{0 \le u'' \le v} M_{s,t,v}(u'', u). \tag{13}$$

In other words, the minimization problem of (10) for fixed s, t, v and each $u, 1 \le u \le N$, is to find the minimum element of every column u of the matrix $M_{s,t,v}$ (or all column minima of the matrix $M_{s,t,v}$).

Exhaustive search can complete this task in $O(N^2)$ time. But the time complexity can be drastically reduced if the matrix is totally monotone [1]. An $m \times n$ rectangular matrix with elements $M(i, j)$, $a \le i \le a + m - 1, b \le j \le b + n - 1$, is said to be totally monotone (with respect to column minima) if and only if for all $a \le i_1 < i_2 \le a + m - 1$ and $b \le j_1 < j_2 \le b + n - 1$, the following implication holds:

$$M(i_1, j_1) \ge M(i_2, j_1) \Rightarrow M(i_1, j_2) \ge M(i_2, j_2). \tag{14}$$

Aggarwal *et al.* [1] showed that if an $m \times n$ rectangular matrix is totally monotone, then all its column minima can be computed in $O(m+n)$ time, comparing to the $O(mn)$ time of the exhaustive search.

The matrix $M_{s,t,v}$ does not actually satisfy the property of total monotonicity, but it can be split into two sub-matrices $A_{s,t,v}$ and $B_{s,t,v}$ each of which satisfies it. The sub-matrix $A_{s,t,v}$ contains the columns 1 through v of the matrix $M_{s,t,v}$, and the sub-matrix $B_{s,t,v}$ contains the columns $v+1$ through N. Finding all column minima of $M_{s,t,v}$ is equivalent to finding all column minima of the two sub-matrices $A_{s,t,v}$ and $B_{s,t,v}$.

A sufficient condition for the total monotonicity of a matrix is the Monge condition [2]. Let f be a real valued function defined on a set $\mathcal{Z}$ of ordered pairs of integers. It is said that f satisfies the Monge condition if and only if for all integers $i_1 < i_2, j_1 < j_2$ such that $(i_1, j_1), (i_2, j_2), (i_1, j_2), (i_2, j_1) \in \mathcal{Z}$ the following inequality holds:

$$f(i_1, j_1) + f(i_2, j_2) \leq f(i_1, j_2) + f(i_2, j_1). \tag{15}$$

The above property of a function can be extended to matrices as well by considering a matrix as a function of two integer variables.

Proposition 1. Both sub-matrices $A_{s,t,v}$ and $B_{s,t,v}$ are totally monotone.

Proof. We will show that both sub-matrices $A_{s,t,v}$ and $B_{s,t,v}$ satisfy the Monge condition, which implies that they are totally monotone [2]. For the sub-matrix $A_{s,t,v}$ the Monge condition is

$$M_{s,t,v}(u_1'', u_1) + M_{s,t,v}(u_2'', u_2) \leq M_{s,t,v}(u_1'', u_2) + M_{s,t,v}(u_2'', u_1), \tag{16}$$

for all integers u_1'', u_2'', u_1, u_2 with $0 \leq u_1'' < u_2'' \leq v$ and $1 \leq u_1 < u_2 \leq v$. The above inequality is trivially satisfied when $u_2'' \geq u_1$ because $M_{s,t,v}(u_2'', u_1) = \infty$ by (12), hence the right hand side becomes ∞. Consequently, it remains to prove it for the case $u_2'' < u_1$. In this case, by (12) and by (7), relation (16) becomes

$$\begin{aligned} &W_{s-1,t}(u_1''v) + (\omega_1 + \omega_0)D(u_1'', u_1] + W_{s-1,t}(u_2''v) + (\omega_1 + \omega_0)D(u_2'', u_2] \leq \\ &\quad W_{s-1,t}(u_1''v) + (\omega_1 + \omega_0)D(u_1'', u_2] + W_{s-1,t}(u_2''v) + (\omega_1 + \omega_0)D(u_2'', u_1]. \end{aligned} \tag{17}$$

After subtracting $W_{s-1,t}(u_1''v)$ and $W_{s-1,t}(u_2''v)$ from both sides, and dividing by $\omega_1 + \omega_0$ when $\omega_1 + \omega_0 \neq 0$ (when $\omega_1 + \omega_0 = 0$ relation (17) is trivially true), (17) becomes

$$D(u_1'', u_1] + D(u_2'', u_2] \leq D(u_1'', u_2] + D(u_2'', u_1]. \tag{18}$$

Since the function $D(a, b]$ satisfies the Monge condition for all monotone distortion functions $d(\cdot,\cdot)$, as proved by Wu and Zhang [11], it follows that (18) is valid for all $0 \le u_1'' < u_2'' < u_1 < u_2 \le v$. Thus, the proof that the Monge condition is satisfied by the sub-matrix $A_{s,t,v}$ is completed.

In the case of the sub-matrix $B_{s,t,v}$, the Monge condition is equivalent to

$$M_{s,t,v}(u_1'', u_1) + M_{s,t,v}(u_2'', u_2) \le M_{s,t,v}(u_1'', u_2) + M_{s,t,v}(u_2'', u_1), \tag{19}$$

for all integers u_1'', u_2'', u_1, u_2 with $0 \le u_1'' < u_2'' \le v$ and $v + 1 \le u_1 < u_2 \le N$. By (12) and by (7), (19) becomes

$$\begin{aligned} &W_{s-1,t}(u_1''v) + \omega_1 D(u_1'', u_1] + \omega_0 D(u_1'', v] + W_{s-1,t}(u_2''v) + \omega_1 D(u_2'', u_2] + \omega_0 D(u_2'', v] \le \\ &\quad W_{s-1,t}(u_1''v) + \omega_1 D(u_1'', u_2] + \omega_0 D(u_1'', v] + W_{s-1,t}(u_2''v) + \omega_1 D(u_2'', u_1] + \omega_0 D(u_2'', v]. \end{aligned} \tag{20}$$

The inequality is trivially satisfied when $\omega_1 = 0$. When $\omega_1 \neq 0$ the inequality is equivalent to

$$D(u_1'', u_1] + D(u_2'', u_2] \le D(u_1'', u_2] + D(u_2'', u_1]. \tag{21}$$

Since the function $D(a, b]$ satisfies the Monge condition it follows that (21) holds for all integers u_1'', u_2'', u_1, u_2 with $0 \le u_1'' < u_2'' \le v$ and $v+1 \le u_1 < u_2 \le N$. Consequently, the proof of the Monge condition for the sub-matrix $B_{s,t,v}$ is completed, too. $\square$

As a consequence of Proposition 1, the algorithm proposed by Aggarwal *et al.* [1] can be applied to compute all column minima of the sub-matrices $A_{s,t,v}$ and $B_{s,t,v}$ in $O(N)$ time. In other words, given s, t, and v, the values $W^I_{s,t}(uv)$ over all u, $1 \le u \le N$, can be found in $O(N)$ time. Therefore, the evaluation of $W^I_{s,t}(uv)$ over all nodes uv when s and t are given takes $O(N^2)$ time.

Given s and t, the values of $W^{II}_{s,t}(uv)$ over all nodes uv can also be computed in $O(N^2)$ time. To see this we consider the $u \times N$ rectangular matrix $M'_{s,t,u}$ for each triple s, t, u, with rows indexed by $v'', 0 \le v'' \le u - 1$, columns indexed by $v, 1 \le v \le N$, and with elements $M'_{s,t,u}(v'', v)$ defined by

$$M'_{s,t,u}(v'', v) = \begin{cases} W_{s,t-1}(uv'') + w(uv'', uv) & \text{if } v'' < v \\ \infty & \text{otherwise.} \end{cases} \tag{22}$$

Then for fixed u, computing $W^{II}_{s,t}(uv)$ over all v, $1 \le v \le N$, requires finding all column minima for the matrix $M'_{s,t,u}$. This matrix also can be split into two sub-matrices, one containing the columns 1 through u, and the other containing the columns $u + 1$ through N. Using the same idea as in the proof of Proposition 1, it can be shown that these two sub-matrices are totally monotone, hence the complexity claim.

Finally, since there are $O(K_1K_2)$ pairs of s and t the overall time complexity of the proposed optimal 2DQ design algorithm is $O(K_1K_2N^2)$.

4 Symmetric 2DQ

An interesting case is when the two side quantizers operate at the same rate ($K_1 = K_2 = K$), and their distortions are equally weighted in the total expected distortion, i.e., $\omega_1 = \omega_2 = \omega$. We call such a two-description quantizer *symmetric* 2DQ. This situation arises in practice. For instance, in diversity communication systems, the outputs of the two quantizers are transmitted through two independent channels that operate at the same data rate $\log_2 K$ and have the same probability q of successful transmission. The probability that the receiver gets only one of the two descriptions is $q(1-q)$, hence $\omega_1 = \omega_2 = \omega = q(1-q)$. The probability that both descriptions arrive successfully is $\omega_0 = q^2$. In conclusion, the expected distortion of the 2DQ becomes

$$\bar{D}(\mathbf{Q}) = (1-q)^2\sigma^2 + q(1-q)(D(Q_1) + D(Q_2)) + q^2 D(Q_0). \tag{23}$$

4.1 Property of Optimal Symmetric 2DQ

We next show that in order to find an optimal symmetric 2DQ we may restrict our attention to a special class of highly structured 2DQ's. This will allow for the further reduction of the time complexity $O(K^2N^2)$ of the design algorithm by a factor of K.

Proposition 2. There is an optimal symmetric 2DQ with interleaved thresholds of the side quantizers:

$$u_0 \leq v_0 \leq u_1 \leq v_1 \leq u_2 \leq v_2 \leq \ \leq u_{K-1} \leq v_{K-1} \leq u_K \leq v_K, \tag{24}$$

where $Q_1 = (u_0, u_1, u_2, \cdots, u_K)$ and $Q_2 = (v_0, v_1, v_2, \cdots, v_K)$ are the side partitions.

Proof. There are $2K+1$ inequalities in the sequence above. The proof is through induction on i that an optimal 2DQ satisfies the first i inequalities. For $i = 1$ the statement is obvious because $u_0 = 0 = v_0$.

The next is to prove the inductive step $i \to i+1$. The nontrivial situations are $2 \leq i \leq 2K-2$.

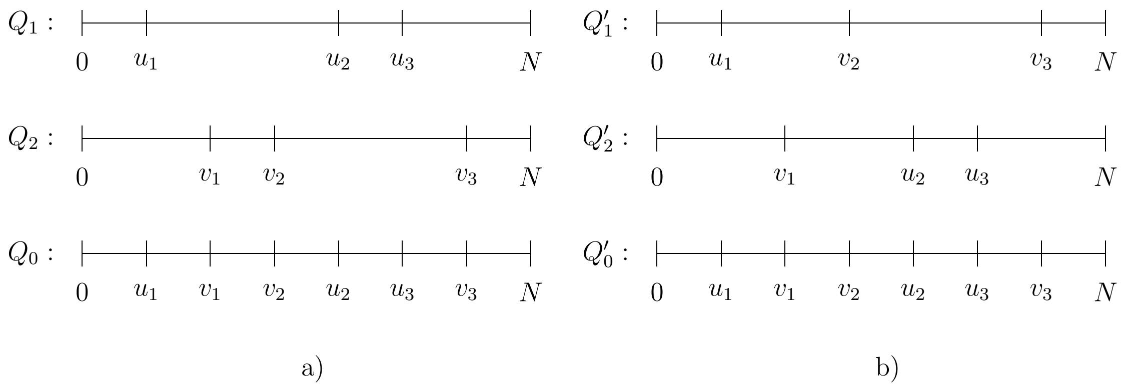

Figure 4: Interchange and modification of codecells in Case 1.

Let $\mathbf{Q}$ be an optimal 2DQ for which the first i inequalities hold. We proceed for even and odd i separately.

Case 1. Consider an even index $i = 2j$ for some $1 \leq j \leq K-1$. The $(i-1)^{th}$ and the i^{th} inequalities are, respectively, $u_{j-1} \leq v_{j-1}$, $v_{j-1} \leq u_j$. If the $(i+1)^{th}$ inequality in the sequence of (24) does not hold, then $u_j > v_j$. Also, $u_{j-1} \leq v_{j-1}$ and $v_{j-1} < v_j$ imply that $u_{j-1} \leq v_j$. We construct a new 2DQ $\mathbf{Q'} = (Q_1', Q_2', Q_0')$, by interchanging the last $K-j$ codecells between the side quantizers, and modifying the j-th codecell accordingly. This is effected by interchanging the last $K-j+1$ thresholds between the side quantizers (i.e., u_k is interchanged with v_k, for all $k, j \leq k \leq K$). In other words, Q_1' has the codecells $(u_0, u_1], \cdots (u_{j-2}, u_{j-1}], (u_{j-1}, v_j], (v_j, v_{j+1}], \cdots (v_{K-1}, v_K]$, and Q_2' has the codecells $(v_0, v_1], \cdots (v_{j-2}, v_{j-1}], (v_{j-1}, u_j], (u_j, u_{j+1}], \cdots (u_{K-1}, u_K]$. This construction is illustrated in Figure 4, for $K = 4$ and $j = 2$. Note that the central partitions of the two 2DQ's remain the same. The codecells which have been interchanged between the two side partitions do not affect the expected distortion because the side quantizers have the same weighting in the expected distortion. Hence any difference in expected distortions is only due to the modification in the j^{th} codecell of each side partition:

$$\bar{D}(\mathbf{Q'}) - \bar{D}(\mathbf{Q}) = \omega(D(u_{j-1}, v_j] + D(v_{j-1}, u_j] - D(u_{j-1}, u_j] - D(v_{j-1}, v_j]). \tag{25}$$

Since the function $D(a, b]$ satisfies the Monge condition [11, 3], the relations $u_{j-1} \leq v_{j-1} < v_j < u_j$ imply that

$$D(u_{j-1}, v_j] + D(v_{j-1}, u_j] \leq D(u_{j-1}, u_j] + D(v_{j-1}, v_j]. \tag{26}$$

Consequently, $\bar{D}(\mathbf{Q'}) \leq \bar{D}(\mathbf{Q})$. Hence, the 2DQ $\mathbf{Q'}$ is optimal, too. But the new 2DQ satisfies the

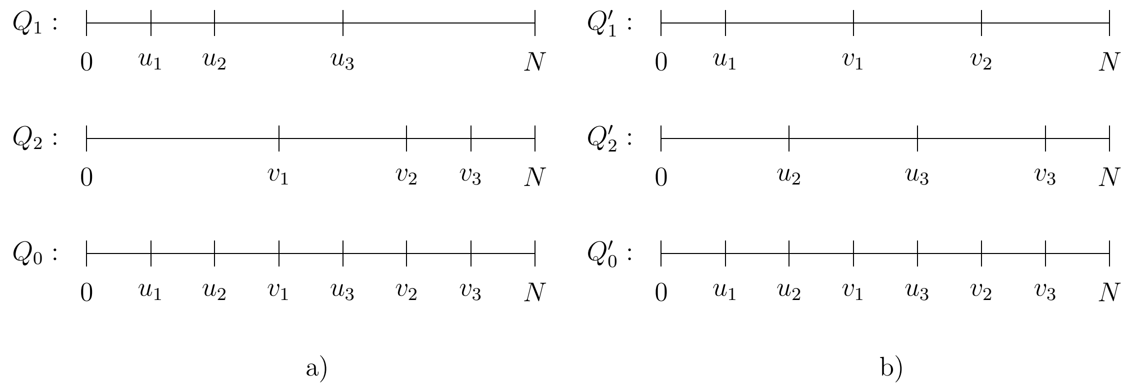

Figure 5: Interchange and modification of codecells in Case 2.

first $i+1$ inequalities. Indeed, the first $i-1$ inequalities are identical to the previous ones and the i^{th} and $(i+1)^{th}$ are equivalent, respectively, to $v_{j-1} \leq v_j$ and $v_j \leq u_j$. Hence the proof of Case 1 is completed.

Case 2. Consider an odd index $i = 2j+1$ for some $1 \leq j \leq K-2$. The $(i-1)^{th}$ and the i^{th} inequalities in the sequence (24) are, respectively, $v_{j-1} \leq u_j$, $u_j \leq v_j$. If the next inequality in the sequence does not hold, then $v_j > u_{j+1}$. Let l be the smallest positive integer ($l \geq 1$) such that $v_{j+l} \leq u_{j+l+1}$. Such an integer always exists because $v_{j+(K-j-1)} < N = u_{j+(K-j-1)+1}$. It follows that $v_{j+l-1} > u_{j+l}$. We construct a new 2DQ $\mathbf{Q}' = (Q'_1, Q'_2, Q'_0)$ by interchanging some thresholds between the side partitions. Namely, the thresholds $u_{j+1}, \cdots, u_{j+l}$, are interchanged with $v_j, \cdots, v_{j+l-1}$, respectively. Consequently, Q'_1 has the codecells $(u_0, u_1], \cdots, (u_{j-1}, u_j], (u_j, v_j]$, $(v_j, v_{j+1}], \cdots, (v_{j+l-2}, v_{j+l-1}]$, $(v_{j+l-1}, u_{j+l+1}], (u_{j+l+1}], u_{j+l+2}], \cdots (u_{K-1}, u_K]$. Quantizer Q'_2 has the codecells $(v_0, v_1], \cdots, (v_{j-2}, v_{j-1}], (v_{j-1}, u_{j+1}]$, $(u_{j+1}, u_{j+2}], \cdots$, $(u_{j+l-1}, u_{j+l}]$, $(u_{j+l}, v_{j+l}]$, $(v_{j+l}, v_{j+l+1}]$, $\cdots (v_{K-1}, v_K]$. This construction is illustrated in Figure 5, for $K = 4$, $j = 1$ and $l = 2$. Although some codecells have been interchanged between the side partitions, they do not affect the expected distortion. Also note that the central partition Q'_0 is identical to Q_0. In order to evaluate the difference in expected distortion between the two 2DQ's we only have to examine the change in the $(j+1)^{th}$ and the $(j+l+1)^{th}$ codecells of the first side quantizer, and the j^{th} and the $(j+l)^{th}$ codecells of the second side quantizer; namely

$$\begin{aligned}\bar{D}(\mathbf{Q}') - \bar{D}(\mathbf{Q}) = \omega(D(u_j, v_j] + D(v_{j-1}, u_{j+1}] - D(u_j, u_{j+1}] - D(v_{j-1}, v_j]) + \\ \omega(D(u_{j+l}, v_{j+l}] + D(v_{j+l-1}, u_{j+l+1}] - D(u_{j+l}, u_{j+l+1}] - D(v_{j+l-1}, v_{j+l}]).\end{aligned} \tag{27}$$

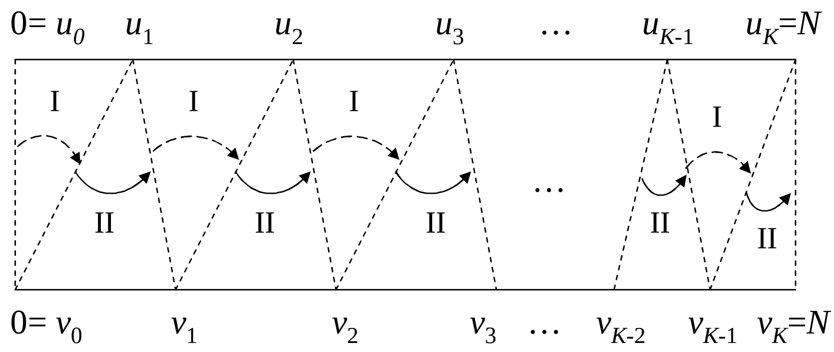

Figure 6: Alternating path in the coupled quantizer graph.

Since the function $D(a, b]$ satisfies the Monge condition, and since $v_{j-1} \leq u_j < u_{j+1} < v_j$, and $u_{j+l} < v_{j+l-1} < v_{j+l} \leq u_{j+l+1}$, it follows that $\bar{D}(\mathbf{Q}') - \bar{D}(\mathbf{Q}) \leq 0$. Hence the 2DQ $\mathbf{Q}'$ is optimal, too. Also in the new 2DQ the first $i+1$ inequalities in the sequence (24) are satisfied, proving Case 2 as well. □

A symmetric 2DQ with interleaved thresholds of the side quantizers corresponds to a path in the coupled quantizer graph of edges of alternating types I and II, which we call alternating path. Specifically, an alternating path (Figure 6) is of the form:

$$00, u_1 0, u_1 v_1, u_2 v_1, u_2 v_2, \cdots, u_{K-1} v_{K-1}, N v_{K-1}, NN. \tag{28}$$

Proposition 2 implies that in order to construct an optimal symmetric 2DQ with K codecells in each side quantizer, it is sufficient to find a $2K$-link minimum-weight alternating path in G.

This problem can be solved by using a similar idea as the one described in Section 3. The main difference is that instead of considering all $O(K^2)$ pairs of s and t (the number of edges of type I and type II, respectively), only $O(K)$ combinations (s with $s-1$, and s with s) need to be processed. This is because for any alternating path ending at some node uv, that has s edges of type I, the number of edges of type II is either s (if $u \leq v$) or $s-1$ (if $u > v$). Consequently, the number of subproblems to be solved in the dynamic programming process is reduced. For each such combination the problem is converted into $O(N)$ problems of matrix search in a totally monotone matrix of dimension $O(N) \times O(N)$. Using the algorithm of [1] to solve these matrix search problems, we arrive at the total time of $O(KN^2)$. Note that the complexity of optimal symmetric 2DQ design is asymptotically as low as that of globally optimal single-description scalar quantizer design [11].

4.2 Simplified Algorithm for the Symmetric Case

Although the use of Proposition 2 in conjunction with the fast matrix search technique of [1] can reduce the complexity of optimal symmetric 2DQ design to $O(KN^2)$, the algorithm of [1] has a quite complex structure, requiring a recursive reduction of the underlying matrix (see the original paper for detail). In this section we propose instead a much simpler nonrecursive algorithm whose implementation is more suitable for engineering practice. Furthermore, due to its unrecursive nature, our algorithm is faster in practice even though its asymptotical complexity is also $O(KN^2)$.

We show first that the coupled quantizer graph can be simplified in the symmetric case. Indeed, Proposition 2 reveals that the edges of the minimum-weight path in the coupled quantizer graph for symmetric 2DQ are highly structured. Specifically, the only edges of type I to be possibly included in the shortest path are those from node uv to node $u'v$ such that $0 \le u \le v \le u' \le N$ and $u < u'$ (hence, those with $u < u' < v$ are unnecessary and can be removed); the only edges of type II to be included in the desired path are those from uv to uv' such that $0 \le v < u \le v' \le N$ (likewise, those with $v < v' < u$ can be removed). Removing the unused edges results in a graph in which any path from the source to the final node is necessarily an alternating path. Moreover, since the side descriptions are symmetrical (namely, their distortions are equally weighted in the total expected distortion), it is no longer necessary to distinguish between the edge types.

Based on the above observations, we further simplify the coupled quantizer graph G to $\underline{G} = (\underline{V}, \underline{E})$. The set of vertices $\underline{V}$ consists of all ordered pairs of integers a and b (denoted by ab) with $0 \le a \le b \le N$. Such a pair ab corresponds to a pair of consecutive thresholds of the central partition, hence a represents a threshold of one side partition and b represents a threshold of the other side partition, but it does not matter which of Q_1 and Q_2. The source node is 00 and the final node is NN. The edges of $\underline{G}$ connect any node ab to any node bc with $a < c$. Such an edge corresponds to the case when a, b and c are three consecutive thresholds of the central partition; or equivalently, a and c are two consecutive thresholds of a side partition and b is a threshold of the other side partition. The weight assigned to such an edge is

$$\underline{w}(ab, bc) = \omega D(a, c] + \omega_0 D(a, b]. \tag{29}$$

There is a one-to-one correspondence between the symmetric 2DQ's with the property (24) of interleaved side partitions' thresholds, and the paths in the WDAG $\underset{\sim}{\mathrm{G}}$, from 00 to NN, with exactly $2K$ edges. This correspondence associates to a 2DQ as in Proposition 2, the path:

$$00, 0u_1, u_1v_1, v_1u_2, u_2v_2, v_2u_3, u_3v_3 \cdots, u_{K-1}v_{K-1}, v_{K-1}N, NN. \tag{30}$$

Moreover, the weight of this path equals $\bar{D}(\mathbf{Q}) - (1 - \omega_1 - \omega_2 - \omega_0)\sigma^2$. Consequently, the problem of optimal 2DQ design in the symmetric case is equivalent to the minimum-weight $2K$-link path problem for the WDAG $\underset{\sim}{\mathrm{G}}$.

For each i, $1 \le i \le 2K$, and each node ab, denote by $\underset{\sim}{\mathrm{W}}_i(ab)$ the weight of the minimum-weight i-edge path from the source to the node ab. $\underset{\sim}{\mathrm{W}}_0(ab) = 0$ by convention; and for $i = 1$, $\underset{\sim}{\mathrm{W}}_i(ab)$ is defined only for the nodes ab with $a = 0$. Clearly, the following recursion holds:

$$\underset{\sim}{\mathrm{W}}_i(ab) = \min_{0 \le \xi \le a, \xi < b} \{\underset{\sim}{\mathrm{W}}_{i-1}(\xi a) + \underset{\sim}{\mathrm{w}}(\xi a, ab)\}. \tag{31}$$

We denote by $\xi_i(a, b)$ the optimal threshold ξ where the minimum in the right hand side of (31) is realized (in the case of multiple solutions, the largest is picked).

The following algorithm can be used to solve the problem. For each $2 \le i \le 2K$ compute $\underset{\sim}{\mathrm{W}}_i(ab)$ for all nodes ab using (31). The search in (31) requires $O(N)$ time. Multiplying by the number of nodes ($O(N^2)$) and the number of different i values, the time complexity amounts to $O(KN^3)$. This value does not yet support our claim. We need another property of the graph $\underset{\sim}{\mathrm{G}}$ to reduce the time complexity.

Proposition 3. For all $1 \le i \le 2K$, and $a \le a'$, $b \le b'$, $a \le b$, $a' \le b'$, (if $i = 1$ then $a = a' = 0$), the following inequality holds

$$\xi_i(a, b) \le \xi_i(a', b'). \tag{32}$$

We defer the proof of Proposition 3 to Appendix in order not to interrupt the presentation flow. The following corollary is a direct consequence of Proposition 3 and relation (31).

Corollary. For all $0 \le a < b \le N$, and $2 \le i \le 2K$, the following relation holds:

$$\underset{\sim}{\mathrm{W}}_i(ab) = \min_{\xi_i(a,b-1) \le \xi \le \xi_i(a+1,b), \xi \le a} \{\underset{\sim}{\mathrm{W}}_{i-1}(\xi a) + \underset{\sim}{\mathrm{w}}(\xi a, ab)\}. \tag{33}$$

Now we are ready to present a new simple algorithm to solve the optimal symmetric 2DQ design problem. Treat the values of $\underset{\sim}{\mathrm{W}}_i(ab)$ for a given i, and all $a, b, 0 \le a \le b \le N$, as the elements of an upper triangular matrix $\underline{W}_i$. The computations are organized such that the entries of the matrix are filled column by column from left to right, and inside each column from the bottom to the top. When proceeding to the position (a, b), $a < b$, the entries of the matrix $\underline{W}_i$, situated left to below (a, b) are already known, hence, $\xi_i(a, b-1)$ and $\xi_i(a+1, b)$ are also known. Consequently, the relation (33) can be used to compute $\underline{W}_i(ab)$. The elements situated on the main diagonal are evaluated by applying (31). The main diagonal needs $O(N^2)$ time since each element takes $O(N)$ time. But any of the other N diagonals above the main diagonal needs only $O(N)$ time. We call the set of entries $\underline{W}_i(ab)$ with $b = a+j$, $0 \le a \le N-j$, the j^{th} superdiagonal. Each element on the j^{th} superdiagonal is computed in $O(\xi_i(a+1, a+j) - \xi_i(a, a+j-1))$ time. Hence the total time spent on the j^{th} superdiagonal is $O(\sum_{a=0}^{N-j}(\xi_i(a+1, a+j) - \xi_i(a, a+j-1))) = O(\xi_i(N-j+1, N) - \xi_i(0, j-1)) = O(N)$. Over N superdiagonals, the time requirement amounts to $O(N^2)$. Multiplying by the number of instances $2K$, we obtain the total time complexity of $O(KN^2)$.

When evaluating the time complexity of the algorithms proposed in this work, we have consistently assumed that each value $D(a, b]$ can be accessed in constant time. For monotone distortion functions $d(\cdot, \cdot)$ all values $D(a, b]$, $0 \le a \le b \le N$, can be precomputed in $O(MN)$ time, where M is the size of the alphabet $\mathcal{B}$, as shown in [11, 3]. Since usually $M = O(N)$, this preprocessing step takes $O(N^2)$ time, not affecting the time complexity of the proposed algorithms.

5 Conclusion

We developed an $O(K_1K_2N^2)$ time algorithm for designing optimal two-description scalar quantizers of convex codecells for the very large family of monotone distortion functions, where K_1 and K_2 are the number of code words of the two side quantizers, and N is the size of symbol alphabet. This complexity can be reduced to $O(KN^2)$ if both side quantizers have the same channel statistics and $K_1 = K_2 = K$.

Acknowledgement

We thank anonymous reviewers for their professionalism and many valuable comments and suggestions that greatly improved the manuscript in its contents and clarity.

References

[1] A. Aggarval, M. Klave, S. Moran, P. Shor, and R. Wilber, " Geometric applications of a matrix-searching algorithm", *Algorithmica*, 2(1987), pp.195-208.

[2] A. Apostolico and Z. Galil(eds.), *Pattern Matching Algorithms*, New York 1997.

[3] S. Dumitrescu, X. Wu, "Algorithms for Optimal Multi-resolution Quantization", *Journal of Algorithms*, vol. 50, pp. 1-22, January 2004.

[4] M. Fleming, Q. Zhao and M. Effros, "Network vector quantization", to appear in *IEEE Trans. Inform. Th.*.

[5] R. M. Gray, D. L. Neuhoff, "Quantization", *IEEE Trans. Inform. Th.*, vol. 44, no. 6, pp. 2325-2383, Oct. 1998.

[6] D. Muresan and M. Effros, "Quantization as histogram segmentation: globally optimal scalar quantizer design in network systems", *Proc. Data Compression Conf.'2002*, pp. 302-311, April 2002.

[7] D. Muresan and M. Effros, "Quantization as histogram segmentation: globally optimal scalar quantizer design in network systems", in preparation.

[8] D. Rebollo-Monedero, R. Zhang and B. Girod, "Design of optimal quantizers for distributed source coding", *Proc. Data Compression Conf.'2003*, pp. 13-22, March 2003.

[9] C. Tian, S. S. Hemami, "Universal multiple description scalar quantization: analysis and design", *Proc. Data Compression Conf.'2003*, pp. 183-192, March 2003.

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

[11] X. Wu and K. Zhang, "Quantizer monotonicities and globally optimal scalar quantizer design", *IEEE Trans. Inform. Theory*, vol. 39, pp. 1049-1053, May 1993.

Appendix

In this appendix we present the Proof of Proposition 3. For this we need first to prove two lemmas. Lemma 1 describes some situations in which two edges can be replaced by other two edges such that the total weight decreases. These situations are illustrated in Figure 7. An edge $(\mu a, ab)$ is represented by the union of the two segment lines connecting the points μ,a and b. μ and b correspond to thresholds of one side partition, hence they are represented on the same horizontal, and a corresponds to a threshold of the other side partition hence it is drawn on another horizontal.

Lemma 1. Let μa, ab, $\mu' a'$ and $a'b'$ be nodes in the WDAG $\underline{\mathrm{G}}$, such that $\mu \leq \mu'$, $a \leq a'$, $b \leq b'$, $\mu < b$ and $\mu' < b'$. Then the following assertions hold:

i) if $\mu' \leq a$ and $\mu' < b$, then

$$\underline{\mathrm{w}}(\mu' a, ab) + \underline{\mathrm{w}}(\mu a', a'b') \geq \underline{\mathrm{w}}(\mu a, ab) + \underline{\mathrm{w}}(\mu' a', a'b'); \tag{34}$$

ii) if $a' \leq b$ and $\mu' < b$, then

$$\underline{\mathrm{w}}(\mu a, ab') + \underline{\mathrm{w}}(\mu' a', a'b) \geq \underline{\mathrm{w}}(\mu a, ab) + \underline{\mathrm{w}}(\mu' a', a'b'); \tag{35}$$

iii) if $\mu' \leq a$ and $a' \leq b$, then

$$\underline{\mathrm{w}}(\mu' a, ab') + \underline{\mathrm{w}}(\mu a', a'b) \geq \underline{\mathrm{w}}(\mu a, ab) + \underline{\mathrm{w}}(\mu' a', a'b'). \tag{36}$$

Proof. i) By replacing the weights of the edges according to (29), relation (34) becomes equivalent to

$$\omega D(\mu', b] + \omega_0 D(\mu', a] + \omega D(\mu, b'] + \omega_0 D(\mu, a'] \geq \omega D(\mu, b] + \omega_0 D(\mu, a] + \omega D(\mu', b'] + \omega_0 D(\mu', a']. \tag{37}$$

A sufficient condition for the above inequality to hold is that the following two relations to be valid:

$$D(\mu, b] + D(\mu', b'] \leq D(\mu, b'] + D(\mu', b] \tag{38}$$

$$D(\mu, a] + D(\mu', a'] \leq D(\mu, a'] + D(\mu', a]. \tag{39}$$

The inequalities (38) and (39) are indeed valid because the function $D(\cdot, \cdot]$ satisfies the Monge condition [11]. Hence the proof of point i) is completed. For proving the conclusions of points

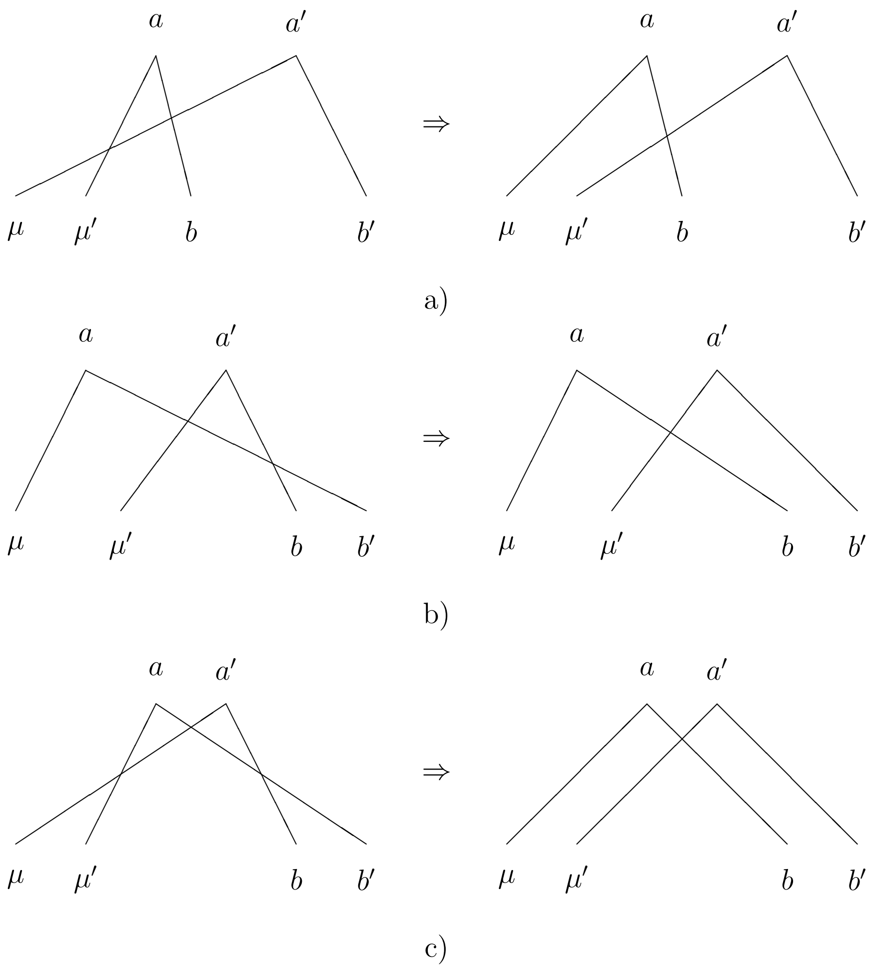

Figure 7: Illustration of Lemma 1. By replacing the two edges to the left by the two edges to the right, the total weight decreases. a) Lemma 1 i); b) Lemma 1 ii); c) Lemma 1 iii).

ii) and iii), again the weights of the edges are replaced in the relations (35) and (36). Thus, (35) becomes equivalent to (38), and (36) becomes equivalent to (39). As we have seen, both (38) and (39) are true. □

Lemma 2. For all $i, 1 \leq i \leq 2K$, $\underline{\mathrm{W}}_i(ab)$ (as a function of two integer variables: a and b) satisfies the Monge condition.

Proof. We prove this result by induction on i. For $i = 1$, the conclusion trivially follows because a and a' can only take the value 0. Assume now that the conclusion of the lemma is true for some fixed value of i and let us prove it for $i+1$. Consequently, we have to show that

$$\underline{\mathrm{W}}_{i+1}(ab) + \underline{\mathrm{W}}_{i+1}(a'b') \leq \underline{\mathrm{W}}_{i+1}(ab') + \underline{\mathrm{W}}_{i+1}(a'b) \tag{40}$$

for any a, a', b, b' such that $a \le a' \le b \le b'$. Denote by ξ, the value $\xi_{i+1}(a, b']$ and by ξ', the value $\xi_{i+1}(a', b]$. Hence, $\xi \le a$, and $\xi < b'$, $\xi' \le a'$, and $\xi' < b$. Applying the definition of $\xi_{i+1}(a, b']$ and $\xi_{i+1}(a', b]$, we obtain that

$$\underline{\mathrm{W}}_{i+1}(ab') = \underline{\mathrm{W}}_i(\xi a) + \underline{\mathrm{w}}(\xi a, ab'), \tag{41}$$

$$\underline{\mathrm{W}}_{i+1}(a'b) = \underline{\mathrm{W}}_i(\xi' a') + \underline{\mathrm{w}}(\xi' a', a'b). \tag{42}$$

Farther we need to distinguish between the cases $\xi \ge \xi'$ and $\xi < \xi'$.

Case $\xi \ge \xi'$. Since $\xi \le a \le a'$, it follows that $\xi' \le a$ and $\xi \le a'$. The definitions of $\underline{\mathrm{W}}_{i+1}(ab)$ and $\underline{\mathrm{W}}_{i+1}(a'b')$ imply, respectively, that

$$\underline{\mathrm{W}}_{i+1}(ab) \le \underline{\mathrm{W}}_i(\xi' a) + \underline{\mathrm{w}}(\xi' a, ab), \tag{43}$$

$$\underline{\mathrm{W}}_{i+1}(a'b') \le \underline{\mathrm{W}}_i(\xi a') + \underline{\mathrm{w}}(\xi a', a'b'). \tag{44}$$

According to the inductive hypothesis, $\underline{\mathrm{W}}_i(\cdot, \cdot)$ satisfies the Monge condition. Hence, since $\xi' \le \xi \le a \le a'$, it follows that

$$\underline{\mathrm{W}}_i(\xi' a) + \underline{\mathrm{W}}_i(\xi a') \le \underline{\mathrm{W}}_i(\xi a) + \underline{\mathrm{W}}_i(\xi' a'). \tag{45}$$

Also Lemma 1, iii) can be applied for $\mu = \xi'$ and $\mu' = \xi$. Combining these two results, we obtain that the sum of the righthand sides of inequalities (43) and (44) is smaller or equal to the sum of the righthand sides of equalities (41) and (42). This implies that (40) is satisfied.

Case $\xi < \xi'$. Since $\xi' < b \le b'$, it follows that $\xi < b$ and $\xi' < b'$. We also have $\xi \le a$ and $\xi' \le a'$. Consequently, the following inequalities hold:

$$\underline{\mathrm{W}}_{i+1}(ab) \le \underline{\mathrm{W}}_i(\xi a) + \underline{\mathrm{w}}(\xi a, ab), \tag{46}$$

$$\underline{\mathrm{W}}_{i+1}(a'b') \le \underline{\mathrm{W}}_i(\xi' a') + \underline{\mathrm{w}}(\xi' a', a'b'). \tag{47}$$

Applying Lemma 1, ii) for $\mu = \xi$ and $\mu' = \xi'$, we obtain that the sum of the righthand sides of inequalities (46) and (47) is smaller or equal to the sum of the righthand sides of equalities (41) and (42). This implies that (40) is satisfied.$\square$

Proposition 3. For all $1 \le i \le 2K$, and $a \le a'$, $b \le b'$, $a \le b$, $a' \le b'$, (if $i = 1$ then $a = a' = 0$), the following inequality holds

$$\xi_i(a, b) \le \xi_i(a', b'). \tag{48}$$

Proof. Assume that

$$\xi_i(a, b) > \xi_i(a', b'). \tag{49}$$

We will show that this assumption leads to a contradiction. Let $\mu = \xi_i(a', b')$ and $\mu' = \xi_i(a, b)$. Then

$$\underline{\mathrm{W}}_i(ab) = \underline{\mathrm{W}}_{i-1}(\mu' a) + \underline{\mathrm{w}}(\mu' a, ab), \tag{50}$$

$$\underline{\mathrm{W}}_i(a'b') = \underline{\mathrm{W}}_{i-1}(\mu a') + \underline{\mathrm{w}}(\mu a', a'b'). \tag{51}$$

Note that $\mu \leq a'$ and $\mu < b'$. Also $\mu' \leq a$ and $\mu' < b$. Using the inequality $\mu < \mu'$, we obtain that $\mu \leq a$ and $\mu < b$. Furthermore, $\mu' \leq a'$ and $\mu' < b'$. These imply that

$$\underline{\mathrm{W}}_i(ab) \leq \underline{\mathrm{W}}_{i-1}(\mu a) + \underline{\mathrm{w}}(\mu a, ab), \tag{52}$$

$$\underline{\mathrm{W}}_i(a'b') < \underline{\mathrm{W}}_{i-1}(\mu' a') + \underline{\mathrm{w}}(\mu' a', a'b'). \tag{53}$$

The inequality in (53) is strict by the definition of of $\xi_i(a', b')$ in conjunction with of $\xi_i(a', b') < \mu'$ (49). Since the function $\underline{\mathrm{W}}_{i-1}(\cdot, \cdot)$ satisfies the Monge condition, it further follows that

$$\underline{\mathrm{W}}_{i-1}(\mu a) + \underline{\mathrm{W}}_{i-1}(\mu' a') \leq \underline{\mathrm{W}}_{i-1}(\mu' a) + \underline{\mathrm{W}}_{i-1}(\mu a'). \tag{54}$$

Moreover, Lemma 1, i) can be applied, and combining these two observations, yields that the sum of the righthand sides of inequalities (52) and (53), which we denote by $\mathcal{S}_1$, is smaller or equal than the sum of the righthand sides of equalities (50) and (51), which we denote by $\mathcal{S}_2$. Relations (50), (51), (52) and (53) further imply that

$$\underline{\mathrm{W}}_i(ab) + \underline{\mathrm{W}}_i(a'b') < \mathcal{S}_1 \leq \mathcal{S}_2 \leq \underline{\mathrm{W}}_i(ab) + \underline{\mathrm{W}}_i(a'b'), \tag{55}$$

which is a contradiction. □