

Optimal Design of a Two-stage Wyner-Ziv Scalar Quantizer with Forwardly/Reversely Degraded Side Information

Qixue Zheng and Sorina Dumitrescu, *Senior Member, IEEE*

Abstract—This work addresses 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 the binning is performed optimally and address the design of the quantizer partitions. The optimization problem is formulated as the minimization of a weighted sum of distortions and rates. The proposed solution is globally optimal when the cells in each partition are contiguous. The solution algorithm is based on solving the single-source or the all-pairs minimum-weight path (MWP) problem in certain weighted directed acyclic graphs. When the conventional dynamic programming technique is used to solve the underlying MWP problems, the time complexity achieved is $O(N^3)$, where N is the size of the source alphabet. 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.

Index Terms—Wyner-Ziv coding, degraded side information, scalar quantization, globally optimal algorithm, minimum-weight path problem, Monge property.

I. INTRODUCTION

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 [1]. The single-letter characterization of the achievable rate-distortion (RD) region for the WZ problem was derived by Wyner and Ziv in [1]. Heegard and Berger [2] and Kaspi [3] studied the scenario where the encoder transmits messages to two decoders, only one of which has SI, referred to as the HB problem. They provided the single-letter characterization of the RD region. Additionally, Heegard and Berger generalized the problem to the case of more than two decoders, each with its own SI, and provided the characterization of the RD region when the SI is stochastically degraded [2].

The problem was further extended to the successive refinement (SR) setting [4]–[6]. Figure 1 depicts the SR scenario with two stages and SI at the two decoders, where X is the source, while Y_κ is the SI at decoder κ , for $\kappa = 1, 2$. Steinberg and Merhav [4] considered multi-stage coding with stochastically degraded SI, i.e., where the decoder receiving

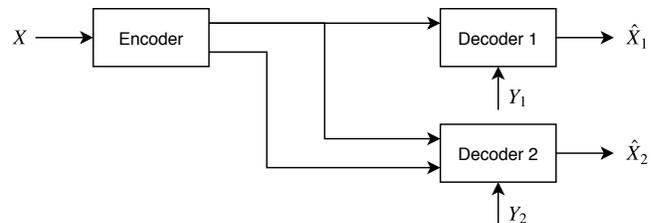


Figure 1. Two-stage Wyner-Ziv coding.

higher rate has stronger SI. They 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 general number of stages and degraded SI was given by Tian and Diggavi in [5]. In [6] Tian and Diggavi investigated 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*. Further, they provided inner and outer bounds to the RD region for general discrete memoryless sources and derived the complete RD region for the multi-stage case for the quadratic Gaussian source with jointly Gaussian SI.

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-constructive¹, 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. 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 [7]–[9].

Practical schemes for the multiple-decoder WZ problem were proposed in [10]–[17]. Cheng and Xiong [10] considered the case when the 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 [11]–[15] to implement WZ schemes with degraded or identical SI, targeting applications in robust video coding. Further, Ramanan and Walsh [16] proposed a coding scheme for the HB problem using successively

¹Such 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.

The authors are with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada (Emails: zhengq15@mcmaster.ca; sorina@mail.ece.mcmaster.ca). This work was supported in part by an NSERC Discovery Grant.

refinable trellis coded quantization and LDPC-based codes for binning. Very recently, Shi et. al. [17] have introduced a construction for the HB problem for binary and Gaussian sources based on nested polar codes, respectively nested polar lattices.

As seen from the above discussion, most of the existing practical schemes for the multiple-decoder WZ problem use uniform quantizer partitions. Better performance is expected to be achieved by employing optimized quantizer partitions. Such an approach was taken by Rebollo-Monedero *et al.* [18] and by Muresan and Effros [19], [20], 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 rate [21]. Both works formulate the problem as the minimization of a weighted sum of the distortion and rate. The algorithm of [18] 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 [19], [20] 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 contiguous². The authors of [19], [20] 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 [20] was developed independently by Dumitrescu and Wu in [22]. Additionally, faster globally optimal design algorithms for fixed-rate SRSQ were developed by Dumitrescu and Wu in [23], [24] and for fixed-rate MDSQ in [25], [26], 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* [27], which was further exploited to accelerate the design procedure.

In this work, 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 [18] and [20],

to assume that the 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 derived in [4] and [6], respectively. Thus, the encoder of the F-WZSQ consists of two nested partitions (a coarse and a fine partition), while the encoder of the R-WZSQ 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 [20], [22] for entropy-constrained SRSQ design (without SI at the decoder), which also involves optimizing nested partitions. The main difference versus [20], [22] resides in the expression of the cost function, which has to account for the presence of the SI at the decoders. 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 (based on dynamic programming) 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 optimal design problems for systems based on fixed-rate scalar quantizers and was leveraged to achieve significant complexity reduction in comparison with conventional algorithms [23]–[26], [28]. 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 set of edges of the aforementioned WDAGs. We refer to this as the partial Monge property and prove that, when it is satisfied, it still can be utilized to expedite the solution.

To summarize, our contribution lies in the following aspects.

- We extend the approach of [20], [22] 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 the binning is performed optimally achieving the Slepian-Wolf limits. This is the first work to address the optimization of the scalar quantizers for the two-stage WZ problem, up to our knowledge.

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

- We introduce the partial Monge property in a complete WDAG³ 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 is fulfilled in many situations of interest, thus allowing for the fast F-WZSQ/R-WZSQ design algorithm to be employed.

We would like to mention that our conference paper [29] addresses only the optimal F-WZSQ design. It includes only a succinct presentation of the fast solution enabled by the partial Monge property, without covering the algorithm details and the proof of correctness. Additionally, work [29] does not contain the empirical verification of the fulfillment of the partial Monge property.

This paper is organized as follows. The following section introduces the notations and the problem formulation. Section III presents the proposed dynamic programming solution based on the MWP model for the optimal R-WZSQ/F-WZSQ design problems. Section IV 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 [30], are given in Section V. Section VI presents extensive experimental results and comparison with the theoretical bounds for a Gaussian source with Gaussian SI. Additionally, the satisfaction of the partial Monge property is empirically investigated. Finally, Section VII concludes the paper.

II. NOTATIONS AND PROBLEM FORMULATION

This section starts by presenting general notations. Subsection II-B 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.

A. Notations

Let $d : \mathbb{R} \times \mathbb{R} \rightarrow [0, \infty)$ denote the distortion function. We will assume that d 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 $\mathcal{X} = \{x_1, \dots, x_N\} \subset \mathbb{R}$, where the elements are labeled in increasing order. Denote $x_0 = -\infty$ and $\tilde{\mathcal{X}} = \mathcal{X} \cup \{x_0\}$. Let \mathcal{Y}_1 and \mathcal{Y}_2 denote the alphabets of the SI Y_1 and Y_2 , respectively. 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 if there exist $x_u, x_v \in \mathcal{X}$ with $u < v$ such that

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

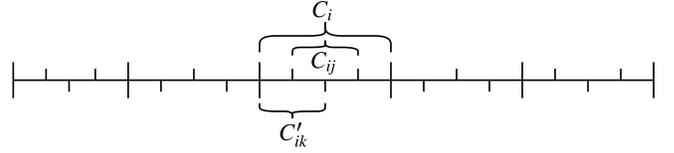


Figure 2. Illustration of the three partitions f_0, f_1 and f_2 .

$S = (x_u, x_v]$, where $(x_u, x_v] \triangleq \{x \in \mathcal{X} | x_u < x \leq x_v\}$. 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 $\mathcal{J}_k \triangleq \{1, 2, \dots, k\}$. For any integer $n \geq 2$, an ascending n -sequence is an n -tuple $\mathbf{r} = (r_0, r_1, \dots, r_{n-1})$, where $r_0 < r_1 < \dots < r_{n-1}$ and $r_i \in \tilde{\mathcal{X}}$, for $0 \leq i \leq n-1$. For any $x_u, x_v \in \tilde{\mathcal{X}}$ 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 $n \geq 2$.

B. Optimal R-WZSQ Design Problem

The configuration of the proposed R-WZSQ scheme is inspired by Tian and Diggavi's work [6]. 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

$$\begin{aligned} f_0 : \mathcal{X} &\rightarrow \mathcal{J}_{M_0}, \quad f_1 : \mathcal{X} \rightarrow \mathcal{J}_{M_1}, \quad f_2 : \mathcal{X} \rightarrow \mathcal{J}_{M_2}, \\ g_1 : \mathcal{J}_{M_0} \times \mathcal{J}_{M_1} \times \mathcal{Y}_1 &\rightarrow \hat{\mathcal{X}}, \quad g_2 : \mathcal{J}_{M_0} \times \mathcal{J}_{M_2} \times \mathcal{Y}_2 \rightarrow \hat{\mathcal{X}}, \end{aligned} \quad (1)$$

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 encoders 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 2, 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 $M_{2,i}, 0 < M_{2,i} \leq M_2$. 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 $i \in \mathcal{J}_{M_0}, j \in \mathcal{J}_{M_{1,i}}, k \in \mathcal{J}_{M_{2,i}}$.

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 $\mathbf{s}_i \triangleq (s_{i,0}, s_{i,1}, \dots, s_{i,M_{1,i}}) \in \mathcal{T}_{r_{i-1}, r_i}$ satisfying $C_{ij} = (s_{i,j-1}, s_{i,j}]$, 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 $\mathbf{t}_i \triangleq (t_{i,0}, t_{i,1}, \dots, t_{i,M_{2,i}}) \in \mathcal{T}_{r_{i-1}, r_i}$, 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 $(\mathbf{s}_1, \dots, \mathbf{s}_{M_0})$, and by $\bar{\mathbf{t}}$ the M_0 -tuple $(\mathbf{t}_1, \dots, \mathbf{t}_{M_0})$.

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 . We assume that when coding the aforementioned message, the binning is performed on blocks of length approaching ∞ , so that the limits in the Slepian-Wolf Theorem [21] 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 [6].

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 rate-distortion performance of an R-WZSQ can be characterized by the quadruple $(R_1(\mathbf{f}), R(\mathbf{f}), D_1(\mathbf{f}), D_2(\mathbf{f}))$. 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. Note that any RD quadruple on the lower convex hull can be obtained by minimizing a weighted sum of the distortions and rates with positive weights [31]. 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

$$\mathcal{O}(\mathbf{r}, \bar{\mathbf{s}}, \bar{\mathbf{t}}) \triangleq \rho D_1(\mathbf{f}) + (1 - \rho) D_2(\mathbf{f}) + \lambda_1 R_1(\mathbf{f}) + \lambda_2 R(\mathbf{f}), \quad (2)$$

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

$$\min_{\mathbf{r}, \bar{\mathbf{s}}, \bar{\mathbf{t}}} \mathcal{O}(\mathbf{r}, \bar{\mathbf{s}}, \bar{\mathbf{t}}). \quad (3)$$

Note that the weights $\rho, 1 - \rho, \lambda_1, \lambda_2$ in (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 [18], [19], [32], [33].

C. 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 (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 (2), but is only a function of \mathbf{r} and $\bar{\mathbf{t}}$, i.e.,

$$\mathcal{O}(\mathbf{r}, \bar{\mathbf{t}}) \triangleq \rho D_1(\mathbf{f}) + (1 - \rho) D_2(\mathbf{f}) + (\lambda_1 + \lambda_2) R_1(\mathbf{f}) + \lambda_2 R_2(\mathbf{f}). \quad (4)$$

The problem of optimal F-WZSQ design is formulated as

$$\min_{\mathbf{r}, \bar{\mathbf{t}}} \mathcal{O}(\mathbf{r}, \bar{\mathbf{t}}). \quad (5)$$

III. DYNAMIC PROGRAMMING SOLUTION BASED ON THE MWP PARADIGM

In this section, we present the proposed solution algorithms based on the MWP model. We first review the MWP problem in a complete WDAG and its dynamic programming solution in subsection III-A. Next we describe the solution to the optimal R-WZSQ design problem in subsection III-B. The following subsection presents the preprocessing step, whose aim is to make possible the computation of each edge weight in constant time. Finally, the solution to the optimal F-WZSQ design problem is discussed in subsection III-D.

A. 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 $V = \{0, \dots, N\}$ and $E = \{(u, v) \in V^2 | 0 \leq u < v \leq N\}$. 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 $\omega : E \rightarrow \mathbb{R}$. 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

3) Solve problem (14).

Next we will discuss how to solve the problem at each step. The key idea is to model each component problem as an MWP problem in a WDAG obtained from 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 $\mathbf{s} = (s_0, \dots, s_M)$ of $(x_u, x_v]$ into M cells, for some $M > 0$, i.e., $\mathbf{s} \in \mathcal{J}_{x_u, x_v}$. For each $j, 0 \leq j \leq M$, let $q_j \in V$ such that $s_j = x_{q_j}$. Then the sequence $\mathbf{q} = (q_0, \dots, q_M)$ is an M -edge path from node u to node v in G . For each $i, 1 \leq i \leq M$, the i th edge on this path, namely (q_{i-1}, q_i) , corresponds to the i th cell in the partition, namely $(s_{i-1}, s_i]$. Consider now the weight function ω_1 defined as follows

$$\omega_1(m, n) \triangleq \rho v_{1,1}((x_m, x_n]) + (\lambda_1 + \lambda_2) v_{2,1}((x_m, x_n]). \quad (15)$$

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 (11), 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 at 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

$$\omega_2(m, n) \triangleq (1 - \rho) v_{1,2}((x_m, x_n]) + \lambda_2 v_{2,2}((x_m, x_n]), \quad (16)$$

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

Finally, problem (14) can be modelled as the MWP problem in the WDAG $G(\omega_0)$, where the source node is 0, the final node is N and $\omega_0(u, v) \triangleq w_0(x_u, x_v)$, for w_0 defined in (13). After having solved the problems at Steps 1 and 2, each weight $\omega_0(u, v)$ can be determined in constant time and, thus, problem (14) can be solved in $O(N^2)$ operations.

C. 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 $\kappa = 1, 2$, $0 \leq m < n \leq N$. 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 $(x_m, x_n) \in \tilde{\mathcal{X}} \times \tilde{\mathcal{X}}$, 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 one has

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

Let us first 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 [28], since the distortion measure is monotone, for fixed κ and y_κ , all values $\gamma_\kappa(m, n, y_\kappa)$ can be computed in $O(|\hat{\mathcal{X}}||\tilde{\mathcal{X}} \cup \hat{\mathcal{X}}|) = O(N^2)$ operations. A simpler technique with the same time complexity was proposed in [24]. It follows that all values $v_{1,\kappa}((x_m, x_n])$ for $\kappa = 1, 2$ and $(x_m, x_n) \in \tilde{\mathcal{X}} \times \tilde{\mathcal{X}}$, 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

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

We first compute and store the cumulative first and second moments, for $i = 1, 2$, $Cum_i(y_\kappa, n) \triangleq \sum_{x \leq x_n} x^i p_{XY_\kappa}(x, y_\kappa)$. 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 $|\mathcal{Y}_1| + |\mathcal{Y}_2| = O(N)$. Corroborating with the discussion in the previous subsection, we conclude that the asymptotical time complexity of the solution algorithm for the optimal R-WZSQ problem is $O(N^3)$ when conventional methods are used to solve the underlying MWP problems.

D. F-WZSQ Design Algorithm

In the F-WZSQ case, $D_2(\mathbf{f})$ remains as in (7), while $D_1(\mathbf{f})$ becomes

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

Additionally, we have

$$\begin{aligned} R_1(\mathbf{f}) &= H(I|Y_1) = H(I, Y_1) - H(Y_1) \\ &= \sum_{i=1}^{M_0} v_{2,1}(C_i) - H(Y_1), \end{aligned}$$

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

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

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

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 (5) reduces to solving (14) with $w_0(x_u, x_v)$, for $u < v$, defined as follows

$$\begin{aligned} w_0(x_u, x_v) &\triangleq \rho v_{1,1}((x_u, x_u]) + (\lambda_1 + \lambda_2)v_{2,1}((x_u, x_v]) \\ &\quad - \lambda_2 v_{2,2}((x_u, x_v]) + w_2((x_u, x_v], \mathbf{t}^*(x_u, x_v)). \end{aligned} \quad (17)$$

In conclusion, problem (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 $\tilde{\mathcal{X}}$ with $u < v$.
- 2) Solve problem (14) with the definition of w_0 given in (17).

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 $G(\omega_0)$, where $\omega_0(u, v) \triangleq w_0(x_u, x_v)$. Thus, using the conventional dynamic programming algorithms for the aforementioned MWP problems, the time complexity of the solution becomes $O(N^3)$.

IV. TIME COMPLEXITY REDUCTION USING THE PARTIAL MONGE PROPERTY

Clearly, the most computationally demanding parts in the solutions to the optimal R-WZSQ and F-WZSQ design problems are 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 of this task when the weight functions ω_1 and ω_2 satisfy it.

If the weight function ω satisfies the Monge property [27], the dynamic programming solution to the single-source MWP problem in $G(\omega)$ can be accelerated by a factor of $N/\log N$ [30] or of N [34], thus leading to the acceleration by the same factor of the all-pairs MWP algorithm. The general idea behind this complexity reduction is the following. The dynamic programming single-source MWP 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 larger number of edges. Thus, the set of edges which need to be further investigated 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. In this section 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 in E' which are examined. This idea is used in conjunction with a simple test to determine another set E'' of edges which cannot appear in any optimal path and thus need not be checked 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 III-A.

Definition 1: We say that the real-valued weight function $\omega : E \rightarrow \mathbb{R}$ satisfies the Monge property [27]⁴ 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 [27], 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 that $a < \infty$, for all $a \in \mathbb{R}$.

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

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

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

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

$$\Delta_\omega(m, n) \geq 0, \text{ for all } 0 \leq m, n \leq N - 1, T_1 \leq n - m \leq T_2. \quad (18)$$

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

⁴This property has received various denominations in the literature. For instance, the authors of [30], 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 [27].

Definition 3: We say that the real-valued weight function $\omega : E \rightarrow \mathbb{R}$ 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 real-valued weight function $\omega : E \rightarrow \mathbb{R}$, let $T_3(\omega)$ be the smallest positive integer T_3 , smaller than N , satisfying

$$\omega(m, n) \geq \omega\left(m, \left\lfloor \frac{m+n}{2} \right\rfloor\right) + \omega\left(\left\lceil \frac{m+n}{2} \right\rceil, n\right), \quad (19)$$

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

Notice that (19) implies that the edge (m, n) can be replaced in any path by two other 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 MWP 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 $E' \triangleq \{(m, n) \in E \mid T_1(\omega) - 1 \leq n - m \leq T_2(\omega) + 1\}$ and $E'' \triangleq \{(m, n) \in E \mid n - m \geq T_3(\omega)\}$. Relation (6) and the discussion below equation (19) imply that

$$\hat{W}_0(n) = \min(\hat{W}'(n), \hat{W}''(n)), \quad (20)$$

where

$$\hat{W}'(n) \triangleq \min_{(m,n) \in E'} (\hat{W}_0(m) + \omega(m, n)), \quad (21)$$

$$\hat{W}''(n) \triangleq \min_{(m,n) \in E \setminus (E' \cup E'')} (\hat{W}_0(m) + \omega(m, n)). \quad (22)$$

Consider now the weight function $\omega' : E \rightarrow \mathbb{R} \cup \{\infty\}$, 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 (21) implies that

$$\hat{W}'(n) \triangleq \min_{0 \leq m < n} (\hat{W}_0(m) + \omega'(m, n)). \quad (23)$$

We will achieve the complexity reduction by exploiting the Monge property of ω' to expedite the computations in (23). For this, we will use a modification of the Basic Algorithm of Hirschberg and Larmore [30] for solving the single-source MWP problem in a WDAG with Monge weights. More specifically, the algorithm of [30] determines all values $F(n)$, for $1 \leq n \leq N$, where

$$F(n) \triangleq \min_{0 \leq m < n} (F(m) + \omega'(m, n)), \quad (24)$$

$F(0) = 0$, 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

$$g(m, n) \triangleq F(m) + \omega'(m, n). \quad (25)$$

Then the problem of solving (24) 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 $1 \leq n \leq N$,

$$F(n) = \min_{0 \leq m < n} g(m, n). \quad (26)$$

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 [30] 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 [30] as algorithm \mathcal{A} .

Now consider a modification of problem (24) as follows

$$F(n) \triangleq \min_{0 \leq m < n} (L(m) + \omega'(m, n)), \quad (27)$$

where $L(0) = 0$ and $L(m)$ is computed based on $F(m)$, for each $1 \leq m \leq n - 1$, according to a specified procedure. Further, let us modify the definition of $g(m, n)$ in (25) as follows

$$g(m, n) \triangleq L(m) + \omega'(m, n), \quad (28)$$

for $0 \leq m < n \leq N$. Then problem (27) remains equivalent to problem (26) of finding all column minima in the modified matrix \mathcal{G} . Relation (28) implies that the elements $g(m, n)$ of the modified upper triangular matrix \mathcal{G} still satisfy the Monge property. Then problem (27) can be solved by using algorithm \mathcal{A} enhanced with a procedure which evaluates $L(m)$ based on $F(m)$, immediately after the latter is computed. We will refer to this algorithm as \mathcal{EA} (short for Enhanced \mathcal{A}). Clearly, the running time of \mathcal{EA} is 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 (23) 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 (22) and (20). The running time to solve the minimization in (22) for a 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 $G(\omega)$, we obtain 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 $O(N^2(T(\omega) + \log N))^5$. We point out that, for this technique to be applicable, 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.

⁵Recall that this evaluation of the time complexity holds when each edge weight can be computed in constant time. Thus, it does not account for the preprocessing step described in subsection III-C.

It is important to point out that in [30] 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 WDAGs of interest, namely $G(\omega_1)$ and $G(\omega_2)$. According to (15) and (16), the weight functions ω_1 and ω_2 comply to the following general form

$$\omega(m, n) = \mu v_{1,\kappa}((x_m, x_n]) + \lambda v_{2,\kappa}((x_m, x_n]), \quad (29)$$

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 $\mathcal{E}\mathcal{A}$ 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 $\mathcal{E}\mathcal{A}$ 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 decrease 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 $\mathcal{E}\mathcal{A}$, the running time to solve the F-WZSQ design problem, excluding the preprocessing stage, decreases to $O(N^2(T_1 + \log N))$. The total time complexity, which accounts for the preprocessing stage as well, is then $O(N^2(T_1 + \log N + |Y_1| + |Y_2|))$. Note that this value is $o(N^3)$ when⁶ $T_1 + |Y_1| + |Y_2| = o(N)$.

In the R-WZSQ case, the condition $\lambda/\mu < \tau_1(p_{XY_\kappa})$ is also satisfied in at least one of the two WDAGs $G(\omega_1)$ 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 significantly 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 when, additionally, the preprocessing stage takes only $o(N^3)$ operations (i.e., when $|Y_1| + |Y_2| = o(N)$).

V. ALGORITHM $\mathcal{E}\mathcal{A}$

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

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

⁶It is said that $f(N) = o(g(N))$ if $\lim_{N \rightarrow \infty} \frac{f(N)}{g(N)} = 0$.

Algorithm $\mathcal{E}\mathcal{A}$: 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
     $\hat{W}'(m) \leftarrow g(f, m), \text{bestleft}(m) \leftarrow f;$ 
     $\hat{W}''(m) \leftarrow \min_{k, m-k \in \mathcal{S}} g_2(k, m)$ 
     $\text{bestleft}_2(m) \leftarrow \arg \min_{k, m-k \in \mathcal{S}} g_2(k, m)$ 
     $\hat{W}_0(m) \leftarrow \min(\hat{W}'(m), \hat{W}''(m));$  Compute
     $\text{bestleft}_0(m)$ 
    while  $|\mathcal{D}| > 1$  and  $g(f_2, m + 1) \leq g(f, m + 1)$  do
         $\perp$  RemoveFront
    while  $|\mathcal{D}| > 1$  and Bridge( $r_2, r, m$ ) do
         $\perp$  RemoveRear
         $\perp$  InsertAtRear( $m$ )
 $\hat{W}'(N) \leftarrow g(f, N), \text{bestleft}(N) \leftarrow f;$ 
 $\hat{W}''(N) \leftarrow \min_{k, N-k \in \mathcal{S}} g_2(k, N)$ 
 $\text{bestleft}_2(N) \leftarrow \arg \min_{k, N-k \in \mathcal{S}} g_2(k, N)$ 
 $\hat{W}_0(N) \leftarrow \min(\hat{W}'(N), \hat{W}''(N));$  Compute  $\text{bestleft}_0(N)$ 

```

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 $\text{bestleft}(n)$ denote the value of m achieving the minimum in (23) (which also achieves the minimum in (21)) and let $\text{bestleft}_2(n)$ be the value of m achieving the minimum in (22). Further, let $\text{bestleft}_0(n)$ denote the node before n in the optimum path from 0 to n in $G(\omega)$. In virtue of (20), $\text{bestleft}_0(n)$ is the best of $\text{bestleft}(n)$ and $\text{bestleft}_2(n)$.

The pseudocode of algorithm $\mathcal{E}\mathcal{A}$ is presented at the top of this column. The algorithm exploits the fact that the function g satisfies the Monge property, fact which follows easily based on Proposition 1. Algorithm $\mathcal{E}\mathcal{A}$ 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 f_2 . The element at the rear, which is the largest, will be denoted r , and the previous element r_2 . Note that f_2 and r_2 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 , f_2 , r , and r_2 is also allowed on \mathcal{D} .

The deque contains all current candidates for $\text{bestleft}(m)$, for all m which are yet to be considered. The algorithm uses the procedure *Bridge*(r_2, r, m), where $r_2 < r < m$, which returns true if and only if $g(r, k) \geq \min(g(r_2, k), g(m, k))$, for all $m < k \leq N$.

We point out that in the Basic Algorithm of [30] operation *InsertAtRear*(m) is performed only if $g(m, N) < g(r, N)$. However, a careful examination of the proof of correctness given in [30] reveals that the algorithm is still correct if that condition is removed.

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

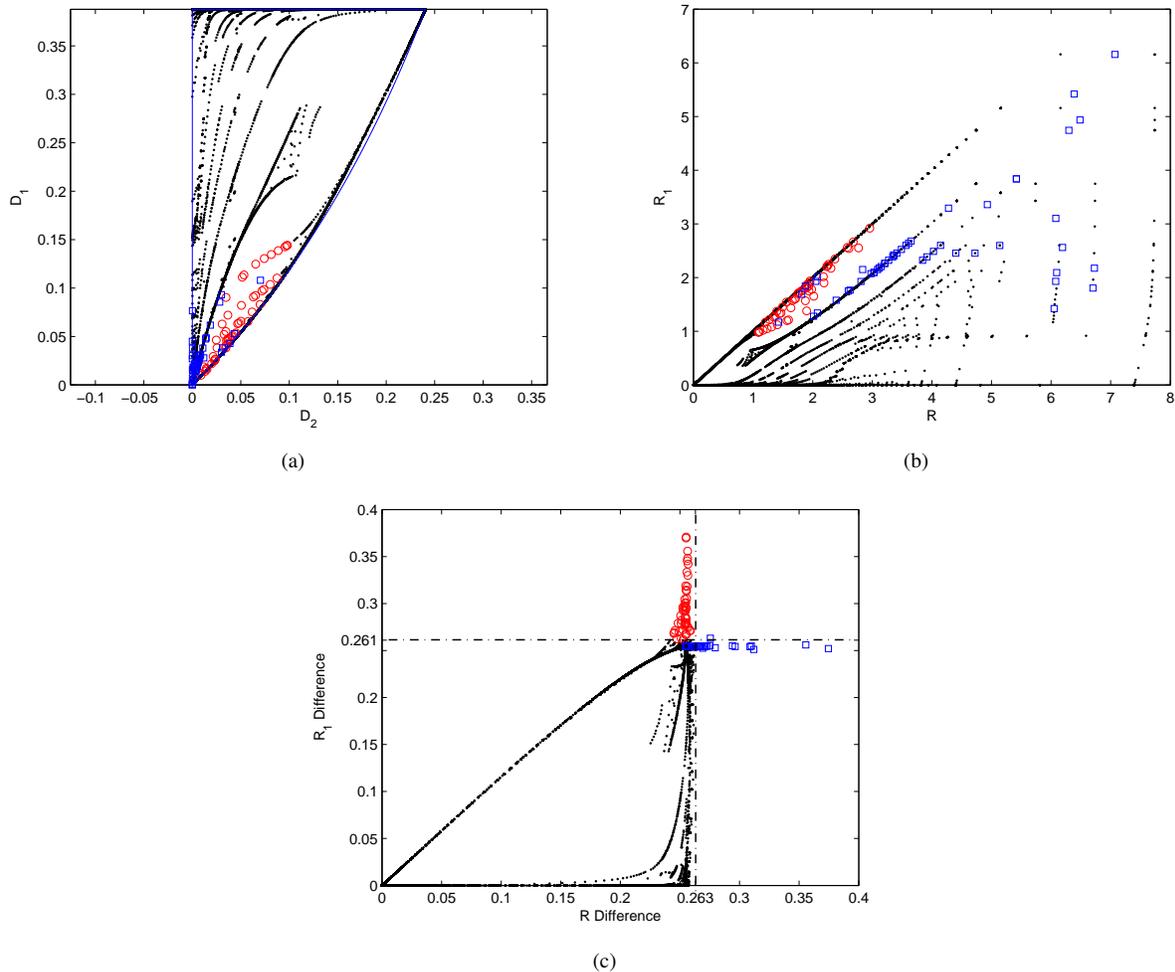


Figure 3. F-WZSQ results. (a) Practical and theoretical (blue line) distortion region. (b) Practical rate region. (c) Difference between R_1 , respectively R , and the corresponding theoretical rate bounds for all the distortion pairs in (a). Circle markers are for the cases when the gap in R_1 is higher than 0.261, square markers are for the cases when the gap in R is higher than 0.263.

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 [30]) 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 [30] 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 [30]) is also invoked at the end of the proof, where actually FP2 suffices.

The subroutine $Bridge(a, b, c)$ proposed in [30] 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 as follows.

$Bridge(a, b, c)$

begin

```

    max ← min(N, b + T2 + 1);
    if c=max then return true;
    min ← max(c + 1, b + T1 - 1);
    low ← min; high ← max;
    if g(a, max) ≤ g(b, max) then return true;
    while high - low ≥ 2 do
        mid ← ⌊(low + high)/2⌋;
        if g(a, mid) ≤ g(b, mid) then
            low ← mid
        else
            high ← mid
    if g(c, high) ≤ g(b, high) then
        return true
    else
        return false

```

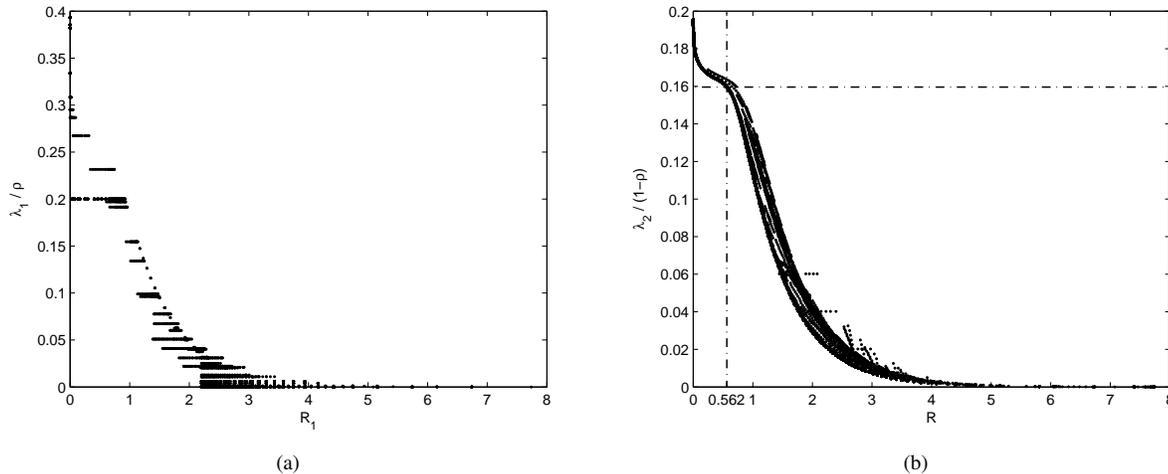


Figure 4. (a) Plot of $\frac{\lambda_1}{\rho}$ versus R_1 when $R_2 \geq 2.2$. (b) Plot of $\frac{\lambda_2}{1-\rho}$ versus R when $R_2 > 0.001$.

VI. 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 \tilde{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 $\tilde{\mathcal{X}} = \mathbb{R}$. For $\kappa = 1, 2$, the SI Y_κ is obtained by discretizing the random variable $\tilde{X} + Z_\kappa$, where Z_κ is 0-mean Gaussian and independent of \tilde{X} . Specifically, for $\kappa = 1, 2$, the alphabet \mathcal{Y}_κ 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 VI-A. We continue with the experimental results for the R-WZSQ scenario in subsection VI-B. 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 VI-C.

A. 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 values 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 the 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 Figure 3(a) and Figure 3(b), respectively. Figure 3(a) also shows the boundary of the theoretically achievable distortion region given in [5] (in blue). Additionally, Figure 3(c) plots the difference between the practical rate pairs (R_1, R) and the pair of theoretical lower bounds $(R_{WZ}(D_1), R_{HB}(D_1, D_2))$ [5], where $R_{WZ}(\cdot)$ denotes the RD function in the single encoder-decoder pair WZ scenario, while $R_{HB}(\cdot)$ denotes the RD function for the general HB problem where each decoder has its own SI. 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 Figure 3. 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 red circles) or in R (points marked using blue 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 dimensional vector quantization predicted by the high rate quantization theory for the single encoder-decoder pair problem is 0.254 bits/sample [35]. 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

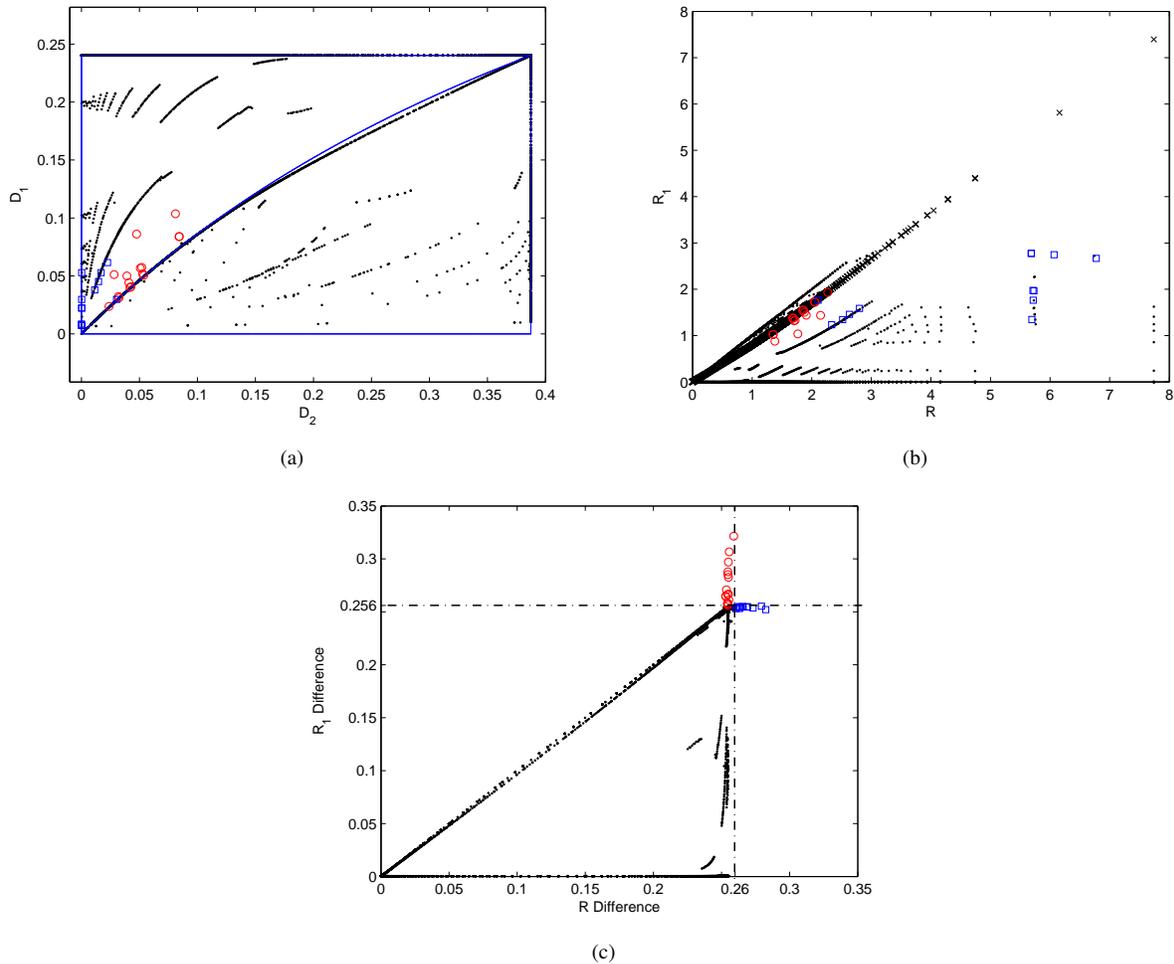


Figure 5. R-WZSQ results. (a) Practical and theoretical (blue line) distortion region. (b) Practical rate region. (c) Difference between R_1 , respectively R , and the corresponding theoretical rate bounds for all the distortion pairs in (a). Circle markers are for the cases when the gap in R_1 is higher than 0.256, square markers are for the cases when the gap in R is higher than 0.26.

$\frac{\lambda_1}{\rho}$ when R is higher than 2.2 bits, and between R and $\frac{\lambda_2}{1-\rho}$ when $R_2 > 0.001$. Specifically, Figure 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, Figure 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.

B. 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 [6] showed that the achievable RD region they proposed for the R-WZ problem is exact in the quadratic

Gaussian case with jointly Gaussian SI. Moreover, they proved 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 fine codebook, to be used at the other decoder. Which decoder recovers the fine 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 . Figure 5(a) and Figure 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 Figure 5(a). The curve connecting the bottom left corner with the top right corner corresponds to the case when both decoders use only 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 Figure 5(b).

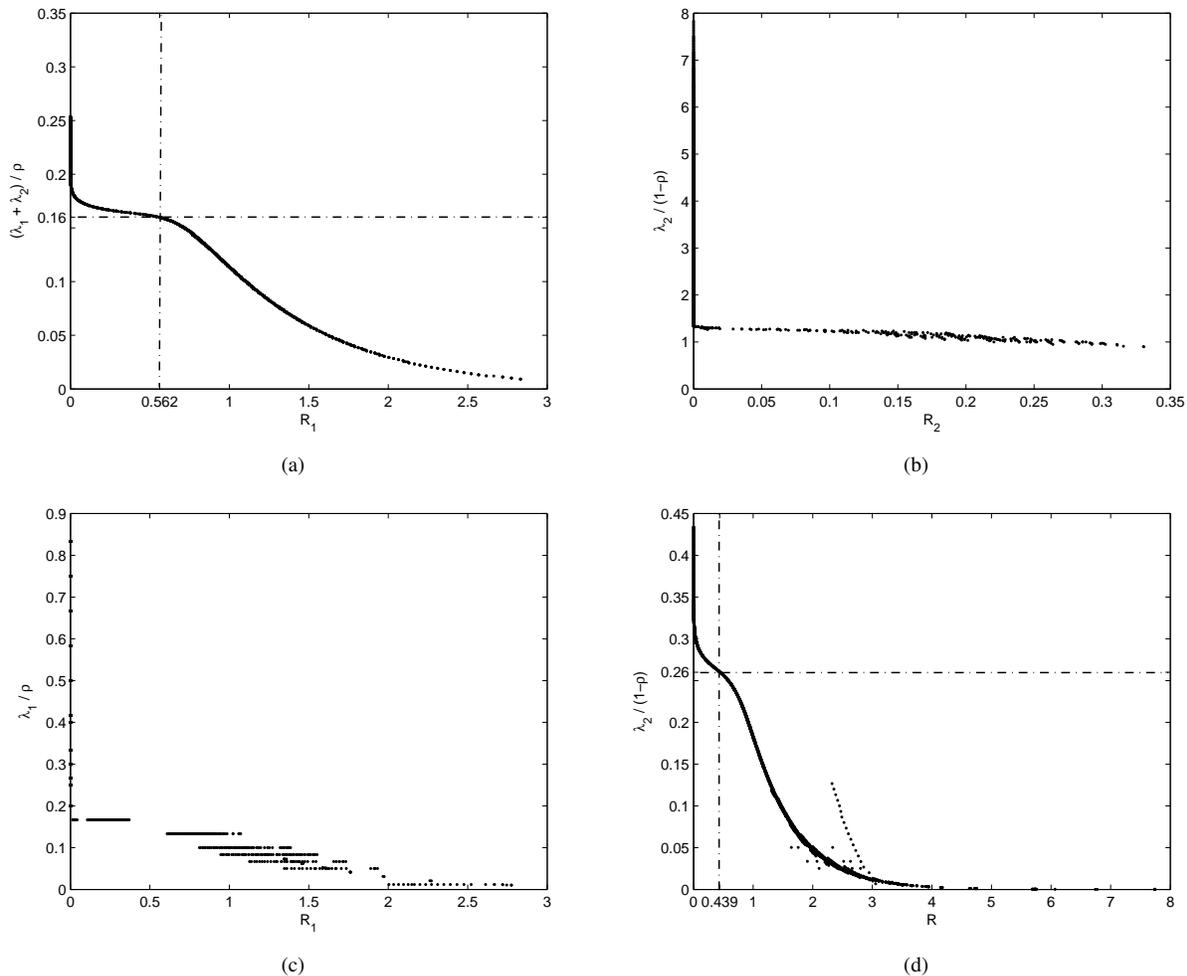


Figure 6. (a) Relation between $\frac{\lambda_1 + \lambda_2}{\rho}$ and R_1 when Q_1 has a refinement. (b) Relation between $\frac{\lambda_2}{1 - \rho}$ and R_2 when Q_1 has a refinement. (c) Relation between $\frac{\lambda_1}{\rho}$ and R_1 when Q_2 has a refinement. (d) Relation between $\frac{\lambda_2}{1 - \rho}$ 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 .

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 Figure 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 Figure 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 a refinement.

Figure 5(c) plots the difference between the practical rate pairs (R_1, R) and the pair of theoretical lower bounds [6]. We observe that in most of our experiments the gap in both R_1 and R is within 0.26. The remaining points exhibit an extra loss either only in R_1 (points marked with red circles) or only in R (points marked with blue squares). Similarly to F-WZSQ, 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 refine-

ment, we have $\frac{\lambda_1 + \lambda_2}{\rho} < 0.255$ and $\frac{\lambda_2}{1 - \rho} \geq 0.9$. 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 $R < 0.44$. When Q_1 has a refinement, we found that R_1 increases as $\frac{\lambda_1 + \lambda_2}{\rho}$ decreases, as seen in Figure 6(a), while R_2 increases as $\frac{\lambda_2}{1 - \rho}$ decreases, as seen in Figure 6(b). When Q_2 has a refinement, we notice from Figure 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 Figure 6(d).

C. 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 weight function ω given in (29). We consider the SI Y_κ obtained by discretizing $\tilde{X} + Z$, where Z is 0-mean Gaussian and independent of \tilde{X} . We will discuss three cases with the variances $\sigma_Z^2 = \frac{1}{\sqrt{10}}, \frac{2}{\sqrt{10}}$ and 0.8.

In Table I 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 (16), which corresponds to equation (29) with $\kappa = 2$, $\lambda = \lambda_2$ and $\mu = 1 - \rho$. Recall that the 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 Figure 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 Figure 6, in order to achieve $R_1 > 0.57$ or $R > 0.44$, at least one of the 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-O notation will be reduced in half, effectively decreasing the practical running time.

VII. 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 (MWP) problem in some weighted directed acyclic graphs. By employing dynamic programming, which is the conventional solution for the underlying MWP problems, 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. We point out that the proposed solution 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 schemes and show that the partial Monge property holds in many situations

of interest. 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 this property is satisfied.

APPENDIX

Proof of Proposition 1: We have to show that the following holds, for all $0 \leq m \leq m' < n \leq n' \leq N$,

$$\omega'(m, n) + \omega'(m', n') \leq \omega'(m, n') + \omega'(m', n). \quad (30)$$

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 (30) 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 (30) 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

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

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 $\Delta_\omega(k, j) \geq 0$ in virtue of (18). It follows that $\Delta(k, n, n') \geq 0$ and, further, that

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

This observation completes the proof. \blacksquare

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

$$g(a, c) + g(b, d) \leq g(b, c) + g(a, d) \quad (31)$$

further implies that $g(a, c) \neq \infty$ and $g(b, d) \neq \infty$. Then (31) 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. \blacksquare

ACKNOWLEDGMENT

The authors would like to thank the Associate Editor and the anonymous reviewers for their valuable comments and suggestions, which helped improve the quality and presentation of the work.

Table I
 T_1, T_2, T_3 EXPERIMENTAL DATA

λ/μ	$\sigma_Z^2 = \frac{1}{\sqrt{10}}$			$\sigma_Z^2 = \frac{2}{\sqrt{10}}$			$\sigma_Z^2 = 0.8$		
	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

REFERENCES

[1] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inf. Theory*, vol. 22, no. 1, pp. 1–10, Jan. 1976.

[2] C. Heegard and T. Berger, "Rate distortion when side information may be absent," *IEEE Trans. Inf. Theory*, vol. 31, no. 6, pp. 727–734, Nov. 1985.

[3] A. H. Kaspi, "Rate-distortion function when side-information may be present at the decoder," *IEEE Trans. Inf. Theory*, vol. 40, no. 6, pp. 2031–2034, Nov. 1994.

[4] Y. Steinberg and N. Merhav, "On successive refinement for the Wyner-Ziv problem," *IEEE Trans. Inf. Theory*, vol. 50, no. 8, pp. 1636–1654, Aug. 2004.

[5] C. Tian and S. N. Diggavi, "On multistage successive refinement for Wyner-Ziv source coding with degraded side informations," *IEEE Trans. Inf. Theory*, vol. 53, no. 8, pp. 2946–2960, July 2007.

[6] C. Tian and S. N. Diggavi, "Side-information scalable source coding," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5591–5608, Dec. 2008.

[7] S. Shamaï, S. Verdú, and R. Zamir, "Systematic lossy source/channel coding," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 564–579, Mar. 1998.

[8] S. S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (DISCUS): design and construction," *IEEE Trans. Inf. Theory*, vol. 49, no. 3, pp. 626–643, Mar. 2003.

[9] Z. Xiong, A. D. Liveris, and S. Cheng, "Distributed source coding for sensor networks," *IEEE Signal Process. Mag.*, vol. 21, no. 5, pp. 80–94, Aug. 2004.

[10] S. Cheng and Z. Xiong, "Successive refinement for the Wyner-Ziv problem and layered code design," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3269–3281, Aug. 2005.

[11] A. Majumdar and K. Ramchandran, "Video multicast over lossy channels based on distributed source coding," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Singapore, Oct. 2004, pp. 3093–3096.

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

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

[14] Q. Xu and Z. Xiong, "Layered Wyner-Ziv video coding," *IEEE Trans. Inf. Theory*, vol. 15, no. 12, pp. 3791–3803, Dec. 2006.

[15] X. Fan, O. C. Au, N. M. Cheung, Y. Chen, and J. Zhou, "Successive refinement based Wyner-Ziv video compression," *Signal Process.: Image Commun.*, vol. 25, no. 1, pp. 47–63, Jan. 2010.

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

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

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

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

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

[21] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. 19, no. 4, pp. 471–480, July 1973.

[22] S. Dumitrescu and X. Wu, "Optimal multiresolution quantization for scalable multimedia coding," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Bangalore, India, Oct. 2002, pp. 139–142.

[23] X. Wu and S. Dumitrescu, "On optimal multi-resolution scalar quantization," in *Proc. Data Compress. Conf. (DCC)*, Snowbird, UT, Mar. 2002, pp. 322–331.

[24] S. Dumitrescu and X. Wu, "Algorithms for optimal multi-resolution quantization," *J. Algorithms*, vol. 50, no. 1, pp. 1–22, Jan. 2004.

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

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

[27] R. E. Burkard, B. Klinz, and R. Rudolf, "Perspectives of Monge properties in optimization," *Discrete Applied Mathematics*, vol. 70, no. 2, pp. 95–161, Sept. 1996.

[28] X. Wu and K. Zhang, "Quantizer monotonicities and globally optimal scalar quantizer design," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 1049–1053, May 1993.

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

[30] D. S. Hirschberg and L. L. Larmore, "The least weight subsequence problem," *SIAM J. Computing*, vol. 16, no. 4, pp. 628–638, Aug. 1987.

[31] D. G. Luenberger and Y. Ye, *Linear and nonlinear programming*, 3rd ed. New York: Springer, 2008.

[32] M. Fleming, Q. Zhao, and M. Effros, "Network vector quantization," *IEEE Trans. Inf. Theory*, vol. 50, no. 8, pp. 1584–1604, Aug. 2004.

[33] P. A. Chou, T. Lookabaugh, and R. M. Gray, "Entropy-constrained vector quantization," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 1, pp. 31–42, Jan. 1989.

[34] R. E. Wilber, "The concave least-weight subsequence problem revisited," *J. Algorithms*, vol. 9, no. 3, pp. 418–425, Sept. 1988.

[35] J. Ziv, "On universal quantization," *IEEE Trans. Inf. Theory*, vol. 31, no. 3, pp. 344–347, May 1985.



Qixue Zheng (M'05-SM'13) received her B.Eng degree in automation engineering from University of Electronic Science and Technology of China, Chengdu, China, in 2016, and her M.A.Sc. in electrical and computer engineering from McMaster University, Ontario, Canada, in 2018. Her research interests include multiple description coding, quantization and multimedia design.

PLACE
PHOTO
HERE

Sorina Dumitrescu (M'05-SM'13) received the B.Sc. and Ph.D. degrees in mathematics from the University of Bucharest, Romania, in 1990 and 1997, respectively. From 2000 to 2002 she was a Postdoctoral Fellow in the Department of Computer Science at the University of Western Ontario, London, Canada. Since 2002 she has been with the Department of Electrical and Computer Engineering at McMaster University, Hamilton, Canada, where she held a Postdoctoral and a Research Associate position, and where she is currently an Associate

Professor. Her current research interests include multimedia coding and communications, network-aware data compression, joint source-channel coding, signal quantization. Her earlier research interests were in formal languages and automata theory. She was a recipient of the NSERC University Faculty Award during 2007-2012.