Slepian–Wolf Coded Nested Lattice Quantization for Wyner–Ziv Coding: High-Rate Performance Analysis and Code Design

Zhixin Liu, Samuel Cheng, *Member, IEEE*, Angelos D. Liveris, and Zixiang Xiong, *Senior Member, IEEE*

***Abstract*—Nested lattice quantization provides a practical scheme for Wyner–Ziv coding. This paper examines the high-rate performance of nested lattice quantizers and gives the theoretical performance for general continuous sources. In the quadratic Gaussian case, as the rate increases, we observe an increasing gap between the performance of finite-dimensional nested lattice quantizers and the Wyner–Ziv distortion-rate function. We argue that this is because the boundary gain decreases as the rate of the nested lattice quantizers increases. To increase the boundary gain and ultimately boost the overall performance, a new practical Wyner–Ziv coding scheme called Slepian–Wolf coded nested lattice quantization (SWC-NQ) is proposed, where Slepian–Wolf coding is applied to the quantization indices of the source for the purpose of compression with side information at the decoder. Theoretical analysis shows that for the quadratic Gaussian case and at high rate, SWC-NQ performs the same as conventional entropy-coded lattice quantization with the side information available at both the encoder and the decoder. Furthermore, a nonlinear minimum mean-square error (MSE) estimator is introduced at the decoder, which is theoretically proven to degenerate to the linear minimum MSE estimator at high rate and experimentally shown to outperform the linear estimator at low rate. Practical designs of one- and two-dimensional nested lattice quantizers together with multilevel low-density parity-check (LDPC) codes for Slepian–Wolf coding give performance close to the theoretical limits of SWC-NQ.**

***Index Terms*—Lattice quantization and lattice channel code, low-density parity-check (LDPC) codes, nested lattices, Slepian–Wolf coding, syndrome-based compression, Wyner—Ziv coding.**

I. Introduction

WYNER–ZIV coding [37], or lossy source coding with side information at the decoder, is one of the main problems considered in network information theory [9, Ch. 14]. It generalizes lossless source coding with side information at the decoder—a special case of Slepian–Wolf coding [33].[1] The rate–distortion (RD) function of Wyner–Ziv coding is known for both discrete and continuous alphabet cases of the source and the side information with a general distortion metric in [37], [38]. It is derived by using a technique called "binning" that divides the set of *jointly typical sequences* [9] into bins which are as *far apart* (in terms of the correlation statistics) as possible. The binning scheme for Wyner–Ziv coding can be applied to other related problems (e.g., Gelfand–Pinsker coding [17] and its special case of dirty-paper coding [8]) based on the duality between source coding and channel coding with side information [2], [27]. The theoretical analyses in [37], [38] are based on *random binning* which, due to its lack of structure, does not indicate how practical code design should be done.

In their information-theoretical work, Zamir *et al.* [42] outlined a structured *algebraic binning* scheme based on a pair of *nested* linear/lattice codes for Wyner–Ziv coding of binary symmetric/quadratic Gaussian sources, where the fine code in the nested pair plays the role of source coding while the coarse code does channel coding. The quadratic Gaussian case corresponds to when the correlation between the source X and the side information Y can be modeled by an additive white Gaussian noise (AWGN) channel as $X = Y + Z, Z \sim N(0, \sigma_Z^2)$, with mean-square error (MSE) distortion and arbitrarily distributed Y. Note that Wyner–Ziv coding in general suffers a rate loss when compared to coding with side information available both at the encoder and the decoder. The quadratic Gaussian case is special because there is no rate loss with Wyner–Ziv coding in this case.[2] Furthermore, it is shown in [42] that the Wyner–Ziv RD function in this special case is asymptotically achievable using nested lattices, under the assumption that the lattices are ideally sphere-packed as the lattice dimensions go to infinity. However, high-dimensional lattice codes are difficult to implement in practice. Thus, structured binning via nested lattice codes only facilitates high-dimensional asymptotic analysis [42].

In this paper, we analyze the performance of *finite-dimensional* nested lattice quantizers for continuous sources under the high-rate assumption. Here the high-rate assumption is consistent with the one in classic quantization theory [18], meaning that the source X is uniformly distributed inside the fine lattice

Manuscript received August 23,2005; revised July 5, 2006. This work was supported by the National Science Foundation under Grant CCF-0430720. The material in this paper was presented in part at the Data Compression Conference, Snowbird, UT, March 2004.

Z. Liu and Z. Xiong are with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843 USA (e-mail: liuzx@ece.tamu.edu;zx@ece.tamu.edu).

S. Cheng was with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843 USA. He is now with the School of Electrical and Computer Engineering, University of Oklahoma–Tulsa, Tulsa, OK 74135 USA (e-mail: phsamuel@gmail.com).

A. D. Liveris was with the Department of Electrical and Computer Engineering, Texas A&M University. He is now with Microwave Networks Inc., Stafford, TX 77477 USA (e-mail: alive@ieee.org).

Communicated by S. A. Savari, Associate Editor for Source Coding.

Color versions of Figs. 1, 3–7, 9, and 11 are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2006.881833

[1]Throughout the paper, Slepian–Wolf coding means near-lossless source coding with side information at the decoder.

[2]It was only shown in [37] that Wyner–Ziv coding of X suffers no rate loss when X and Y are zero mean and jointly Gaussian with MSE distortion. Pradhan *et al.* [27] recently extended this no rate loss result to the more general quadratic Gaussian case.

cell of the quantizer. The distortion-rate (DR) performance is analyzed for both the general and the quadratic Gaussian cases. For general continuous sources, the distortion under a specific rate consists of two parts: one from source coding and another from channel coding. For the quadratic Gaussian case, a tight lower bound of the DR function is given, showing an increasing gap from the Wyner–Ziv limit, as the rate increases. Based on our analysis, we argue that this increasing gap is due to the decreasing boundary gain as the rate increases. Thus, a practical approach to boosting the overall performance is to increase the boundary gain with a second stage of binning, which groups the support region of the fine lattice into k cosets. This way, the volume of the support region decreases by a factor of k while the decoding error probability stays the same. According to the definition in [11], the boundary gain increases without changing the dimensionality of the lattices. Since various possible boundary gains are realizable using the second stage of binning, there is only maximally 1.53-dB granular gain left unexploited by the quantizer. Thus, using Slepian–Wolf coding, for second-stage binning allows us to show the theoretical performance limits at high rate.

Following this logic, we introduce a new framework for Wyner–Ziv coding of continuous independent and identically distributed (i.i.d.) sources based on Slepian–Wolf coded nested quantization (SWC-NQ). Slepian–Wolf coding [33] here refers to near-lossless source coding with side information at the decoder. Practical syndrome-based schemes for Slepian–Wolf coding using channel codes have been studied in [1], [16], [20], [21], [26]. The role of Slepian–Wolf coding in SWC-NQ is to exploit the correlation between the quantized source and the side information for further compression and by making the overall channel code stronger. SWC-NQ generalizes the classic source coding approach of entropy-coded quantization in the sense that the quantizer performs quite well alone and can exhibit further rate savings by employing a powerful Slepian–Wolf code. Moreover, it connects network information theory with the rich areas of practical lattice source code (e.g., [7]) and channel code (e.g., low-density parity-check (LDPC) codes [14], [23]) designs, making it feasible to devise codes that can approach the Wyner–Ziv DR function.

For the quadratic Gaussian case, we establish the high-rate performance of SWC-NQ with ideal Slepian-Wolf coding, assuming there is no channel decoding error in the latter. We show that SWC-NQ with finite-dimensional nested lattice quantizer at high rate achieves the same performance of classic entropy-coded lattice quantization as if the side information were also available at the encoder. For example, with ideal Slepian–Wolf coding, one-/two-dimensional SWC-NQ performs 1.53/1.36 dB away from the Wyner–Ziv DR function for quadratic Gaussian sources at high rate.

We also implement one- and two-dimensional nested lattice quantizers in the rate range of 1.0–6.0 bits per sample (b/s), for the case when Y is also Gaussian (hence X and Y are jointly Gaussian), which is a special case of the quadratic Gaussian scenario. Our experiments using nested lattice quantizers together with irregular LDPC codes for Slepian–Wolf coding give performance close to the corresponding limit at high rate. Our work thus shows that SWC-NQ provides an efficient scheme for practical Wyner–Ziv coding with low-dimensional lattice quantizers at high rate.

Although our theoretical analysis assumes high rate, when a *nonlinear* minimum MSE estimator is applied at the decoder, our simulated DR performance of SWC-NQ at low rate matches that of classic entropy-coded quantization at low rate when the side information is also available at the encoder. At high rate, the nonlinear estimator degenerates to the linear one used in our high-rate performance analysis.

We note that nonlinear estimation at the decoder can yield significant gains only for low rate and for high rate it cannot help noticeably. This is confirmed by the agreement of the high-rate analysis results in this paper, which assume that the linear estimation is used, with the high-rate simulation results, for which the nonlinear estimation method is always used.

The contributions of this paper are as follows.

1) High-rate performance analysis of nested lattice quantization for all dimensions, indicating for any finite dimension an increasing gap in distortion from the Wyner–Ziv DR function as the rate increases.
2) Introduction of the SWC-NQ framework for Wyner–Ziv coding. The DR performance of SWC-NQ for the quadratic Gaussian case is presented, showing agreement with the performance of entropy-coded lattice quantization in classic source coding.
3) Our proof that the performance loss of Wyner–Ziv coding of quadratic Gaussian sources with nested lattice quantization at a fixed high rate is independent of the source correlation, with or without Slepian–Wolf coding.
4) A nonlinear minimum MSE estimator at the decoder of the nested lattice quantizer, which improves the quantizer performance at low rate.
5) Practical designs of one-dimensional scalar and two-dimensional hexagonal nested lattice quantizers and multilevel irregular LDPC codes for Slepian–Wolf coding, confirming our high-rate performance analyses for both nested lattice quantization and SWC-NQ.

The rest of this paper is organized as follows. Section II gives the background on Slepian–Wolf coding and Wyner–Ziv coding. Section III defines lattices and nested lattices. Section IV is devoted to the high-rate performance analysis of nested lattice quantization. Section V introduces SWC-NQ and analyzes its performance. Section VI proposes an optimal quantizer decoder, discusses practical nested lattice quantizer design and multilevel LDPC codes for Slepian–Wolf coding, and presents simulation results. Section VII concludes the paper.

A. Related Works

As mentioned earlier, Zamir *et al.* [42] studied the high-dimension asymptotics of nested lattice quantization for Wyner–Ziv coding. Practical approaches to Wyner–Ziv coding have recently been investigated in [26], [36], [31], [24], [12], [28], [5], [22], [40]. For example, in DISCUS [26], two source codes (scalar quantization and trellis-coded quantizer (TCQ)) and two channel codes (scalar coset code and trellis-based coset code [35]) are used in source–channel coding for the Wyner–Ziv problem, resulting in four combinations. One of

them (scalar quantization with scalar coset code) is nested scalar quantization and another one (TCQ with trellis-based coset code, also suggested in [36]) can effectively be considered as nested TCQ.

A recent work [29] starts with nonuniform quantization with index reuse and Slepian–Wolf coding and shows the same high-rate theoretical performance as ours when the quantizer becomes an almost uniform one without index reuse. This agrees with our finding in Section V that at high rate, the nested quantizer asymptotically becomes a non-nested regular one so that strong channel coding is guaranteed.

Servetto [31] explored explicit nested lattice constructions based on similar sublattices [6]. But we point out that results presented in this paper contradict those in [31], [32] in three aspects: 1) Whereas our analysis in Section IV shows that finite-dimensional nested lattice quantization performs increasingly worse than the Wyner–Ziv limit as the rate increases,[3] [31, Fig. 3] seems to indicate that the performance of a low-dimensional nested lattice quantizer is a constant gap (in decibels) away from the Wyner–Ziv limit in the 2.0–7.0 b/s rate range. 2) Our simulation results with two-dimensional nested lattice quantization shown in Fig. 11(b) are much worse than those in [31, Fig. 4]. 3) In [32], the author attempts to apply nested lattice quantization to dense sensor networks with limited resource (e.g., fixed rate), where high correlation among sensor outputs is achieved by increasing the number of sensors. In contrast, we show in Section IV that, at fixed rate and dimensionality, the gap between the performance of a nested lattice quantizer and the Wyner–Ziv limit is independent of the source correlation at high rate—hence, there is no high-correlation asymptotics in Wyner–Ziv coding at high rate.

II. Theoretical Backgrounds

In this section, we give the background on Slepian–Wolf coding and Wyner–Ziv coding.

A. Slepian–Wolf Coding

Slepian–Wolf coding is concerned with near-lossless source coding with side information at the decoder. For lossless compression of a pair of correlated, discrete random variables X and Y, a rate of $R_X + R_Y = H(X,Y)$ is sufficient if they are encoded jointly [9]. However, Slepian and Wolf [33] showed that the rate $R_X + R_Y = H(X,Y)$ is almost sufficient even for separate encoding (with joint decoding) of X and Y. Specifically, the Slepian–Wolf theorem says that the achievable region for coding X and Y is given by

$$R_X \geq H(X \,|\, Y),\ R_Y \geq H(Y \,|\, X),\ R_X + R_Y \geq H(X,Y). \tag{1}$$

This result shows that there is no loss of coding efficiency with separate encoding when compared to joint encoding as long as joint decoding is performed. When the side information (e.g., Y) is perfectly available at the decoder, then the aim of Slepian–Wolf coding is to compress X to the rate limit $H(X \,|\, Y)$.

B. Wyner–Ziv Coding

Wyner–Ziv coding [37], [38] deals with the problem of RD with side information at the decoder. It asks the question of how many bits are needed to encode X under the constraint that $E\{d(X,\hat{X})\} \leq D$, assuming the side information Y is available at the decoder but not at the encoder. This problem generalizes the setup of [33] in that coding of X is lossy with respect to a fidelity criterion rather than lossless. For both discrete and continuous alphabets of $\mathcal{X}$ and general distortion metrics $d(\cdot)$, Wyner and Ziv [37] gave the RD function $R_{\mathrm{WZ}}(D)$ for this problem as $R_{\mathrm{WZ}}(D) = \inf I(X; A \,|\, Y)$, where the infimum is taken over all auxiliary random variables A such that $Y \rightarrow X \rightarrow A$ is a Markov chain and there exists a function $\hat{X} = \hat{X}(A,Y)$ satisfying $E\{d(X,\hat{X})\} \leq D$. According to [37]

$$R_{\mathrm{WZ}}(D) \geq R_{X \,|\, Y}(D) = \inf_{\{\hat{X} \in \mathcal{X}: E\{d(X,\hat{X})\} \leq D\}} I(X; \hat{X} \,|\, Y)$$

where $R_{X \,|\, Y}(D)$ is the classic RD function of coding X with Y available at the encoder (and the decoder). This means that, compared to coding of X when the side information Y is also available at the encoder, there is in general a rate loss with Wyner–Ziv coding. Zamir quantified this loss in [41], showing a < 0.22 b/s loss for binary sources with Hamming distance and a < 0.5 b/s loss for continuous sources with MSE distortion.

When D is very small and the source is discrete-valued, the Wyner–Ziv problem degenerates to the Slepian–Wolf problem with

$$R_{\mathrm{WZ}}(D) = R_{X \,|\, Y}(D) = H(X \,|\, Y).$$

Another interesting setup is the quadratic Gaussian case with the source model being $X = Y + Z$ and $Z \sim N(0, \sigma_Z^2)$, then

$$R_{\mathrm{WZ}}(D) = R_{X \,|\, Y}(D) = \frac{1}{2} \log^+ \left[\frac{\sigma_Z^2}{D} \right]$$

where $\log^+ x = \max\{\log x, 0\}$, i.e., there is no rate loss in this case. Note that Y is arbitrarily distributed [27]. When Y is also Gaussian (then X and Y are jointly Gaussian memoryless sources), let the covariance matrix of (X_i, Y_i) be

$$\Lambda = \begin{bmatrix} \sigma_X^2 & \rho \sigma_X \sigma_Y \\ \rho \sigma_X \sigma_Y & \sigma_Y^2 \end{bmatrix}$$

[3] The intuitive explanation for this increasing performance gap from the Wyner–Ziv limit at high rate is given by a reviewer of [19], [31] as follows: Under the quadratic Gaussian correlation model $X = Y + Z$, with $N \sim N(0, \sigma_Z^2)$, the noise Z should in high probability be contained in the coarse lattice cell in order for the nested lattice coding scheme to function well. Otherwise, the probability of decoding error will be too high, and will dominate the total distortion, in particular, at high rate. Since a cubic cell does not match the spherical shape of a multidimensional Gaussian, we must take a large margin in the scaling of the coarse lattice to make the decoding error probability small enough. Moreover, the smaller the total distortion needs to be, the higher this margin must be.

Alternatively, using the concept of "cell overloading" [18], we can say that the Voronoi cell of a low-dimensional (or any finite-dimensional) lattice quantizer does not match the spherical shape of a multi-dimensional Gaussian, i.e., we always have cell overloading with low-dimensional lattice quantizers. Relatively large cell size (or low rate) are needed to avoid cell overloading. But as the rate increases, the cell size decreases; this leads to more and more severe cell overloading and explains the increasing performance gap from the Wyner-Ziv limit.

with $|\rho| < 1$ for all n, then

$$R_{\mathrm{WZ}}(D) = R_{X|Y}(D) = \frac{1}{2}\log^{+}\left[\frac{\sigma_X^2(1-\rho^2)}{D}\right].$$

This case is of special interest in practice because many image and video sources can be modeled as jointly Gaussian and Wyner–Ziv coding suffers no rate loss. For the sake of simplicity, we consider this specific case in our code designs.

III. Lattices and Nested Lattices

A. Lattices

For a set of n independent basis vectors $\{\boldsymbol{m}_1, \boldsymbol{m}_2, \ldots, \boldsymbol{m}_n\}$, an unbounded n-dimensional lattice Λ is defined by

$$\Lambda = \{\boldsymbol{l} = M\boldsymbol{i} : \boldsymbol{i} \in \mathbb{Z}^n\} \tag{2}$$

and its generator matrix $M = [\boldsymbol{m}_1 \boldsymbol{m}_2 \cdots \boldsymbol{m}_n]$. The nearest neighbor quantizer $Q_\Lambda(\cdot)$ associated with Λ is given by

$$Q_\Lambda(\mathbf{x}) = \arg\min_{\boldsymbol{l}\in\Lambda} |\mathbf{x} - \boldsymbol{l}|. \tag{3}$$

The basic Voronoi cell of Λ, which specifies the shape of the nearest-neighbor decoding region, is

$$\mathcal{V} = \{\mathbf{x} : Q_\Lambda(\mathbf{x}) = \mathbf{0}\}. \tag{4}$$

Associated with the Voronoi cell $\mathcal{V}$ are several important quantities: the cell volume V, the second moment σ^2 and the normalized second moment $G(\Lambda)$, defined by

$$V = \int_{\mathcal{V}} d\mathbf{x}, \quad \sigma^2 = \frac{1}{nV}\int_{\mathcal{V}} |\mathbf{x}|^2\, d\mathbf{x}, \quad \text{and} \quad G(\Lambda) = \frac{\sigma^2}{V^{\frac{2}{n}}} \tag{5}$$

respectively. The minimum of $G(\Lambda)$ over all lattices in $\mathbb{R}^n$ is denoted as G_n. By [7], $G_n \geq \frac{1}{2\pi e}, \forall n$, and $\lim_{n\to\infty} G_n = \frac{1}{2\pi e}$.

B. Nested Lattices

A pair of n-dimensional lattices (Λ_1, Λ_2) with corresponding generator matrices M_1 and M_2 is *nested* if there exists an $n \times n$ integer matrix P such that $M_2 = M_1 \times P$ and $|\det P| > 1$. In this case, $\frac{V_2}{V_1}$ is called the *nesting ratio*, and Λ_1 and Λ_2 are called the *fine* and *coarse* lattices, respectively.

For a pair of nested lattices (Λ_1, Λ_2), the points in the set $\Lambda_1/\Lambda_2 \triangleq \{\Lambda_1 \bigcap \mathcal{V}_2\}$ are called the *coset leaders* of Λ_2 relative to Λ_1, where $\mathcal{V}_2$ is the basic Voronoi cell of Λ_2. For each $\mathbf{v} \in \Lambda_1/\Lambda_2$ the set of shifted lattice points $C(\mathbf{v}) \triangleq \{\mathbf{v} + \boldsymbol{l}, \forall \boldsymbol{l} \in \Lambda_2\}$ is called a *coset* of Λ_2 relative to Λ_1. The jth point of $C(\mathbf{v})$ is denoted as $\boldsymbol{c}_j(\mathbf{v})$. Then

$$C(\mathbf{0}) = \{\boldsymbol{c}_j(\mathbf{0}), \quad \forall j \in \mathbb{Z}\} = \Lambda_2 \tag{6}$$

and

$$\bigcup_{\mathbf{v}\in\Lambda_1/\Lambda_2} C(\mathbf{v}) = \Lambda_1. \tag{7}$$

Since

$$\boldsymbol{c}_j(\mathbf{v}) \in \Lambda_1, \qquad \forall j \in \mathbb{Z} \tag{8}$$

we further define

$$R_j(\mathbf{v}) = \{\mathbf{x} : Q_{\Lambda_1}(\mathbf{x}) = \boldsymbol{c}_j(\mathbf{v})\}$$

as the Voronoi region associated with $\boldsymbol{c}_j(\mathbf{v})$ in Λ_1, and $R(\mathbf{v}) = \bigcup_{j=-\infty}^{\infty} R_j(\mathbf{v})$, then

$$\bigcup_{j=-\infty}^{\infty} \bigcup_{\mathbf{v}\in\Lambda_1/\Lambda_2} R_j(\mathbf{v}) = \bigcup_{\mathbf{v}\in\Lambda_1/\Lambda_2} R(\mathbf{v}) = \mathbb{R}^n. \tag{9}$$

An examples of $\mathbf{v}, C(\mathbf{v})$, and $R(\mathbf{v})$ for $n = 2$ are shown in Fig. 1.

IV. Nested Lattice Quantization

Throughout this paper, we use the correlation model of $X = Y + Z$, where X is the source to be coded, Y is the side information, and Z is the noise. Y and Z are independent. In this section, we discuss the performance of nested lattice quantization for general sources where Y and Z are arbitrarily distributed continuous sources with zero mean, and for the quadratic Gaussian case with $Z \sim N(0, \sigma_Z^2)$. For both cases, MSE is used as the distortion measure.

Zamir *et al.*'s nested lattice quantization scheme [42] works as follows: Let the pseudorandom vector U (the dither), known to both the quantizer encoder and the decoder, be uniformly distributed over the basic Voronoi cell $\mathcal{V}_1$ of the fine lattice Λ_1. For a given target average distortion D, denote $\alpha = \sqrt{1 - \frac{D}{\sigma_Z^2}}$ as the estimation coefficient. Given the n-dimensional realizations of the source, the side information and the dither as $\mathbf{x}, \mathbf{y}$, and $\mathbf{u}$, respectively, then according to [42], the nested quantizer encoder quantizes $\alpha\mathbf{x} + \mathbf{u}$ to the nearest point $\mathbf{x}_{Q_{\Lambda_1}} = Q_{\Lambda_1}(\alpha\mathbf{x} + \mathbf{u})$ in Λ_1, computes $\mathbf{s} = \mathbf{x}_{Q_{\Lambda_1}} - Q_{\Lambda_2}(\mathbf{x}_{Q_{\Lambda_1}})$ which is the coset shift of $\mathbf{x}_{Q_{\Lambda_1}}$ with respect to Λ_2, and transmits the index corresponding to this coset shift.

The nested quantizer decoder receives $\mathbf{s}$, forms $\mathbf{t} = \mathbf{s} - \mathbf{u} - \alpha\mathbf{y}$, and reconstructs $\mathbf{x}$ as $\hat{\mathbf{x}} = \mathbf{y} + \alpha(\mathbf{t} - Q_{\Lambda_2}(\mathbf{t}))$ using linear combination and dithering in estimation.

It is shown in [42] that the Wyner–Ziv DR function $D_{\mathrm{WZ}}(R) = \sigma_Z^2 2^{-2R}$ is achievable with infinite-dimensional nested lattice quantization for quadratic Gaussian case.

In this paper, we analyze the high-rate performance of finite-dimensional nested lattice quantizers. Our analysis is based on the high-resolution assumption, which means that D is small compared to σ_Z^2. Consequently, V_1 is small enough so that the conditional pdf of X given Y, $f(\mathbf{x}|\mathbf{y})$, is approximately constant over each Voronoi cell of Λ_1. Under the high-rate assumption, $\alpha = 1$. In addition, dithering is not needed in our high-rate analysis. With $\alpha = 1$ and $\mathbf{u} = \mathbf{0}$, the encoder/decoder described above simplifies to the following.

- The encoder quantizes $\mathbf{x}$ to $\mathbf{x}_{Q_{\Lambda_1}} = Q_{\Lambda_1}(\mathbf{x})$, computes $\mathbf{s} = \mathbf{x}_{Q_{\Lambda_1}} - Q_{\Lambda_2}(\mathbf{x}_{Q_{\Lambda_1}})$, and transmits an index corresponding to the coset leader $\mathbf{s}$.
- Upon receiving $\mathbf{s}$, the decoder forms $\mathbf{t} = \mathbf{s} - \mathbf{y}$ and reconstructs $\mathbf{x}$ as $\hat{\mathbf{x}} = \mathbf{y} + \mathbf{t} - Q_{\Lambda_2}(\mathbf{t}) = \mathbf{s} + Q_{\Lambda_2}(\mathbf{y} - \mathbf{s})$.

In the performance analysis, we limit ourselves to this simplified nested lattice quantization scheme for high rate, which is shown in Fig. 2 and was also used in [31].

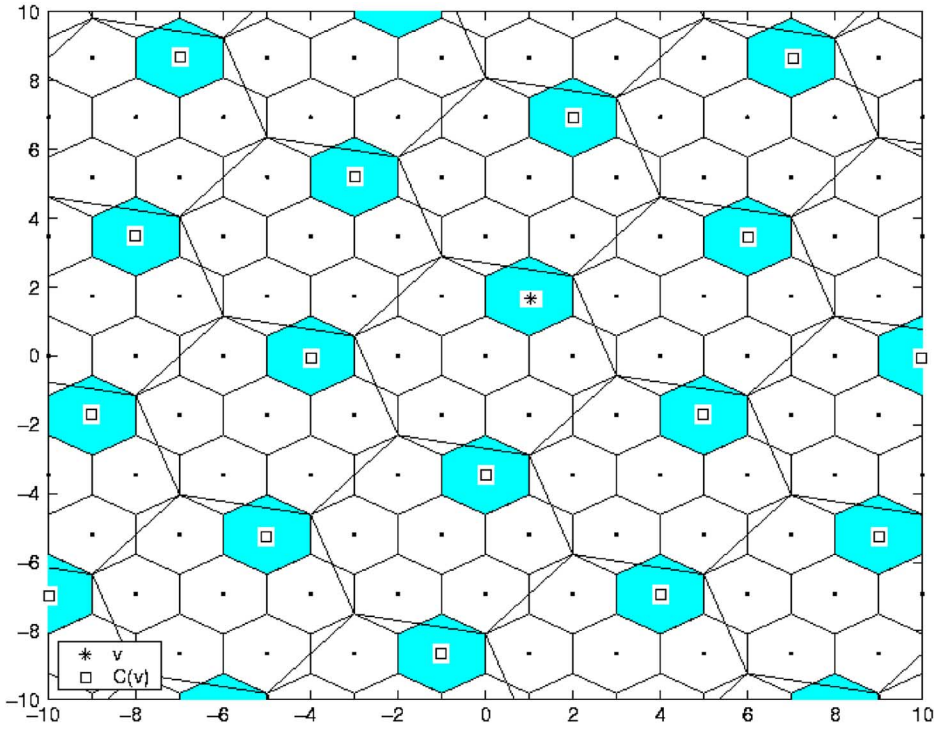


Fig. 1. An example of $\mathbf{v}, C(\mathbf{v})$, and $R(\mathbf{v})$ for $n = 2$, where the shaded regions correspond to $R(\mathbf{v})$.

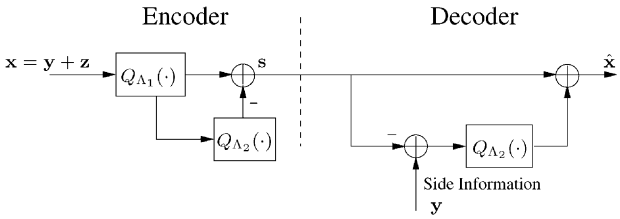


Fig. 2. The simplified nested lattice quantizer for Wyner–Ziv coding.

A. High-Rate Performance for General Sources With Arbitrary Distribution

Theorem 4.1: If a pair of n-dimensional nested lattices (Λ_1, Λ_2) with nesting ratio $N = \frac{V_2}{V_1}$ is used for nested lattice quantization, the distortion per dimension in Wyner–Ziv coding of X (with decoder side information Y) at high rate is

$$D_n = G(\Lambda_1)V_1^{\frac{2}{n}} + \frac{1}{n}E_{\mathbf{Z}}\left[\|Q_{\Lambda_2}(\mathbf{Z})\|^2\right]. \tag{10}$$

Proof: Since

$$\mathbb{R}^n = \bigcup_{j=-\infty}^{\infty} \bigcup_{\mathbf{v}\in\Lambda_1/\Lambda_2} R_j(\mathbf{v}) \tag{11}$$

the average distortion for a given realization of the side information $\mathbf{y}$ is

$$\begin{aligned}
D(\mathbf{y}) &= \int_{\mathbb{R}^n} f(\mathbf{x}\,|\,\mathbf{y})\|\mathbf{x}-\hat{\mathbf{x}}\|^2\,d\mathbf{x} \\
&= \sum_{\mathbf{v}\in\Lambda_1/\Lambda_2}\sum_{j=-\infty}^{\infty}\int_{\mathbf{x}\in R_j(\mathbf{v})} f(\mathbf{x}\,|\,\mathbf{y}) \\
&\quad\times \|\mathbf{x}-\boldsymbol{c}_j(\mathbf{v})+\boldsymbol{c}_j(\mathbf{v})-\hat{\mathbf{x}}\|^2\,d\mathbf{x} \\
&= \sum_{\mathbf{v}\in\Lambda_1/\Lambda_2}\sum_{j=-\infty}^{\infty}\int_{\mathbf{x}\in R_j(\mathbf{v})} f(\mathbf{x}\,|\,\mathbf{y})[\|\mathbf{x}-\boldsymbol{c}_j(\mathbf{v})\|^2 \\
&\quad+\|\boldsymbol{c}_j(\mathbf{v})-\hat{\mathbf{x}}\|^2+2\langle\mathbf{x}-\boldsymbol{c}_j(\mathbf{v}), \\
&\quad\boldsymbol{c}_j(\mathbf{v})-\hat{\mathbf{x}}\rangle]\,d\mathbf{x} \\
&\overset{(a)}{=} \sum_{\mathbf{v}\in\Lambda_1/\Lambda_2}\sum_{j=-\infty}^{\infty}\Bigg[f(\boldsymbol{c}_j(\mathbf{v})\,|\,\mathbf{y}) \\
&\quad\times\int_{\mathbf{x}\in R_j(\mathbf{v})}\|\mathbf{x}-\boldsymbol{c}_j(\mathbf{v})\|^2\,d\mathbf{x} \\
&\quad+\int_{\mathbf{x}\in R_j(\mathbf{v})} f(\mathbf{x}\,|\,\mathbf{y})\|\boldsymbol{c}_j(\mathbf{v})-\hat{\mathbf{x}}\|^2\,d\mathbf{x}\Bigg] \\
&\overset{(b)}{=} \sum_{\mathbf{v}\in\Lambda_1/\Lambda_2}\sum_{j=-\infty}^{\infty}\Bigg[f(\boldsymbol{c}_j(\mathbf{v})\,|\,\mathbf{y})nG(\Lambda_1)V_1^{1+\frac{2}{n}} \\
&\quad+\int_{\mathbf{x}\in R_j(\mathbf{v})} f(\mathbf{x}\,|\,\mathbf{y})\|Q_{\Lambda_2}(\boldsymbol{c}_j(\mathbf{v})) \\
&\quad-Q_{\Lambda_2}(\mathbf{y}-\boldsymbol{c}_j(\mathbf{v})+Q_{\Lambda_2}(\boldsymbol{c}_j(\mathbf{v})))\|^2\,d\mathbf{x}\Bigg] \\
&\overset{(c)}{=} nG(\Lambda_1)V_1^{\frac{2}{n}} \\
&\quad+\sum_{j=-\infty}^{\infty}\sum_{\mathbf{v}\in\Lambda_1/\Lambda_2}\int_{\mathbf{x}\in R_j(\mathbf{v})} f(\mathbf{x}\,|\,\mathbf{y})\|Q_{\Lambda_2}(\mathbf{x}-\mathbf{y})\|^2\,d\mathbf{x} \\
&= nG(\Lambda_1)V_1^{\frac{2}{n}} \\
&\quad+\int_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}\,|\,\mathbf{y})\|Q_{\Lambda_2}(\mathbf{x}-\mathbf{y})\|^2\,d\mathbf{x}
\end{aligned} \tag{12}$$

where (a) comes from the high rate assumption and

$$\int_{\mathbf{x}\in R_j(\mathbf{v})} < \mathbf{x}-\boldsymbol{c}_j(\mathbf{v}), \boldsymbol{c}_j(\mathbf{v})-\hat{\mathbf{x}} > \, d\mathbf{x} = 0$$

which is due to the fact that $\mathbf{x}-\boldsymbol{c}_j(\mathbf{v})$ is odd spherical symmetric for $\mathbf{x}\in R_j(\mathbf{v})$ and both $\boldsymbol{c}_j(\mathbf{v})$ and $\hat{\mathbf{x}}$ are fixed for $\mathbf{x}\in R_j(\mathbf{v})$ with given $\mathbf{v}$ and $\mathbf{y}$, (b) is due to $\boldsymbol{c}_j(\mathbf{v}) = Q_{\Lambda_1}(\mathbf{x})$ for $\mathbf{x}\in R_j(\mathbf{v})$ and $\hat{\mathbf{x}} = \boldsymbol{c}_j(\mathbf{v}) - Q_{\Lambda_2}(\boldsymbol{c}_j(\mathbf{v})) + Q_{\Lambda_2}(\mathbf{y}-\boldsymbol{c}_j(\mathbf{v}) + Q_{\Lambda_2}(\boldsymbol{c}_j(\mathbf{v})))$, and (c) is because

$$\begin{aligned}
&\sum_{\mathbf{v}\in\Lambda_1/\Lambda_2}\sum_{j=-\infty}^{\infty} f(\boldsymbol{c}_j(\mathbf{v})\,|\,\mathbf{y})V_1 \\
&\quad = \sum_{\mathbf{v}\in\Lambda_1/\Lambda_2}\sum_{j=-\infty}^{\infty}\int_{\mathbf{x}\in R_j(\mathbf{v})} f(\mathbf{x}\,|\,\mathbf{y})\,d\mathbf{x} \\
&\quad = \int_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}\,|\,\mathbf{y})\,d\mathbf{x} = 1
\end{aligned} \tag{13}$$

and $Q_{\Lambda_2}(\mathbf{a}+Q_{\Lambda_2}(\mathbf{b})) = Q_{\Lambda_2}(\mathbf{a})+Q_{\Lambda_2}(\mathbf{b}), \forall \mathbf{a},\mathbf{b}\in\mathbb{R}^n$, which leads to

$$\begin{aligned}
&\int_{\mathbf{x}\in R_j(\mathbf{v})} f(\mathbf{x}\,|\,\mathbf{y})\,\|Q_{\Lambda_2}(\boldsymbol{c}_j(\mathbf{v})) \\
&\quad - Q_{\Lambda_2}(\mathbf{y}-\boldsymbol{c}_j(\mathbf{v}) + Q_{\Lambda_2}(\boldsymbol{c}_j(\mathbf{v})))\|^2\,d\mathbf{x} \\
&\qquad = \int_{\mathbf{x}\in R_j(\mathbf{v})} f(\mathbf{x}\,|\,\mathbf{y})\,\|Q_{\Lambda_2}(\boldsymbol{c}_j(\mathbf{v})) \\
&\quad - Q_{\Lambda_2}(\mathbf{y}-\boldsymbol{c}_j(\mathbf{v})) - Q_{\Lambda_2}(\boldsymbol{c}_j(\mathbf{v}))\|^2\,d\mathbf{x} \\
&\qquad = \int_{\mathbf{x}\in R_j(\mathbf{v})} f(\mathbf{x}\,|\,\mathbf{y})\,\|Q_{\Lambda_2}(\mathbf{y}-\boldsymbol{c}_j(\mathbf{v}))\|^2\,d\mathbf{x} \\
&\qquad = \int_{\mathbf{x}\in R_j(\mathbf{v})} f(\mathbf{x}\,|\,\mathbf{y})\,\|Q_{\Lambda_2}(\mathbf{x}-\mathbf{y})\|^2\,d\mathbf{x}.
\end{aligned} \tag{14}$$

Therefore, the average distortion per dimension over all realizations of $\mathbf{y}$ is

$$\begin{aligned}
D_n &= \frac{1}{n}E_Y[D(\mathbf{y})] \\
&= G(\Lambda_1)V_1^{\frac{2}{n}} + \frac{1}{n}\int_{\mathbf{x}}\int_{\mathbf{y}} f(\mathbf{x},\mathbf{y})\,\|Q_{\Lambda_2}(\mathbf{x}-\mathbf{y})\|^2\,d\mathbf{x}d\mathbf{y} \\
&= G(\Lambda_1)V_1^{\frac{2}{n}} + \frac{1}{n}\int_{\mathbf{y}} f(\mathbf{y})\int_{\mathbf{z}} f(\mathbf{z})\,\|Q_{\Lambda_2}(\mathbf{z})\|^2\,d\mathbf{z}d\mathbf{y} \\
&= G(\Lambda_1)V_1^{\frac{2}{n}} + \frac{1}{n}E_{\mathbf{Z}}\left[\|Q_{\Lambda_2}(\mathbf{Z})\|^2\right].
\end{aligned} \tag{15}$$

□

Remarks:

- For a fixed pair of the nested lattices (Λ_1,Λ_2), D_n only depends on Z, i.e., the correlation between X and Y. It is independent of the marginal distribution of X (or Y).
- The first term $D_S \triangleq G(\Lambda_1)V_1^{\frac{2}{n}}$ in D_n is due to lattice quantization, which is determined by the geometric structure and V_1. It is the same as the MSE for classic lattice quantizers [11]. The second term $D_C \triangleq \frac{1}{n}E_{\mathbf{Z}}[\|Q_{\Lambda_2}(\mathbf{Z})\|^2]$ is the loss due to nesting (or lattice channel coding). It depends on V_2 and the distribution of Z, and is characterized by the error probability of the lattice channel code.

B. The Quadratic Gaussian Case When $n\to\infty$

Corollary 4.1: In the quadratic Gaussian case

$$\lim_{n\to\infty} D_n = D_{\mathrm{WZ}} = \sigma_Z^2 2^{-2R}. \tag{16}$$

Proof: Since the nested lattice quantizer is a fixed-rate quantizer with rate $R = \frac{1}{n}\log(\frac{V_2}{V_1})$ per dimension, then (10) can be rewritten as

$$D_n = G(\Lambda_1)V_2^{\frac{2}{n}}2^{-2R} + \frac{1}{n}E_{\mathbf{Z}}\left[\|Q_{\Lambda_2}(\mathbf{Z})\|^2\right]. \tag{17}$$

For the quadratic Gaussian case, according to [42, eq. (3.14)]

$$\lim_{n\to\infty}\frac{1}{n}\log V_2 = \frac{1}{2}\log\left(2\pi e\sigma_Z^2\right) \tag{18}$$

if we assume Λ_2 is a *good* AWGN channel σ_Z^2-code [7], meaning $\mathcal{V}_2$ approximates a Euclidean ball of radius $\sqrt{n}\sigma_Z$. Then

$$\lim_{n\to\infty} G(\Lambda_1)V_2^{\frac{2}{n}}2^{-2R} = \frac{1}{2\pi e}2\pi e\sigma_Z^2 2^{-2R} = D_{\mathrm{WZ}}(R). \tag{19}$$

At the same time, according to [42, (3.12)], for any $\varepsilon>0$ and sufficiently large n, $P_r\{\mathbf{Z}\notin\mathcal{V}_2\} < \varepsilon$ for the *good* Λ_2, hence

$$\lim_{n\to\infty}\frac{1}{n}E_{\mathbf{Z}}\left[\|Q_{\Lambda_2}(\mathbf{Z})\|^2\right] = 0. \tag{20}$$

Consequently

$$\lim_{n\to\infty} D_n = D_{\mathrm{WZ}} = \sigma_Z^2 2^{-2R} \tag{21}$$

for the quadratic Gaussian case. □

The limit (16) we obtain under the high rate assumption is consistent with results in [42], which assert that nested lattice quantization can achieve the Wyner–Ziv limit asymptotically as the dimensionality n goes to infinity for all rates.

C. A Lower Bound on the DR Performance With Finite n in the Quadratic Gaussian Case

The source coding loss $D_S = G(\Lambda_1)V_1^{\frac{2}{n}}$ in (10) is in an explicit form, while the channel coding loss

$$D_C = \frac{1}{n}E_{\mathbf{Z}}[\|Q_{\Lambda_2}(\mathbf{Z})\|^2]$$

not so clear. In the quadratic Gaussian case with $Z\sim(0,\sigma_Z^2)$, we obtain from Theorem 4.1 a lower bound on the high-rate DR performance of finite-dimensional nested lattice quantizers.

Corollary 4.2: For $X = Y+Z, Z\sim N(0,\sigma_Z^2)$, the operational DR function $D_n(R)$ of Wyner–Ziv coding of X (with

decoder side information Y) using n-dimensional nested lattice quantizers is lower-bounded at high rate by

$$\bar{D}_n(R) \triangleq \min_{V_2>0}\{D_S(R)+\bar{D}_C\}, \qquad \text{for } n>1 \tag{22}$$

with

$$D_S(R) = G(\Lambda_1)V_2^{\frac{2}{n}}2^{-2R} \tag{23}$$

and

$$\begin{aligned}\bar{D}_C = &\frac{n-1}{n\Gamma\left(\frac{n+1}{2}\right)2^{\frac{n}{2}}\pi^{\frac{1}{2}}}\\ &\times\sum_{\boldsymbol{l}\in\Lambda_2} l^2\int_{l-r}^{l+r}\int_0^{\cos^{-1}\left(\frac{l^2+u^2-r^2}{2lu}\right)}\sin^{n-2}\theta\, d\theta\\ &\times\frac{u^{n-1}}{\sigma_Z^n}\exp\left(-\frac{u^2}{2\sigma_Z^2}\right)du\end{aligned} \tag{24}$$

where $l \triangleq \|\boldsymbol{l}\|, r=\pi^{-\frac{1}{2}}(\Delta\Gamma(\frac{n}{2}+1)V_2)^{\frac{1}{n}}$ is the packing radius [7, p. 6] of Λ_2, and Δ its density [7, Table 1.2]. When $n=1$, the exact best possible high-rate DR performance is

$$\begin{aligned}D_1(R) = \min_{V_2>0}\Bigg\{&G(\Lambda_1)V_2^2 2^{-2R}\\ &+\frac{1}{2}V_2^2\sum_{i=0}^{\infty}(2i+1)\text{erfc}\left(\frac{V_2}{\sqrt{2}\sigma_Z}\left(i+\frac{1}{2}\right)\right)\Bigg\}.\end{aligned} \tag{25}$$

Proof:

i) Rate computation: The nested lattice quantizer is a fixed rate quantizer with $R=\frac{1}{n}\log_2(\frac{V_2}{V_1})$.

ii) Distortion computation: The source coding loss is

$$G(\Lambda_1)V_1^{\frac{2}{n}} = G(\Lambda_1)V_2^{\frac{2}{n}}2^{-2R} = D_S(R).$$

In the quadratic Gaussian case, the channel coding loss D_C in the second term of (10) can be evaluated as

$$\begin{aligned}D_C &= \frac{1}{n}E_{\mathbf{z}}\left[\|Q_{\Lambda_2}(\mathbf{Z})\|^2\right]\\ &= \frac{1}{n}\sum_{\boldsymbol{l}\in\Lambda_2}\int_{\mathbf{z}\in\mathcal{V}_2(\boldsymbol{l})}\|Q_{\Lambda_2}(\mathbf{Z})\|^2 f(\mathbf{z})\,d\mathbf{z}\\ &= \frac{1}{n}\sum_{\boldsymbol{l}\in\Lambda_2}\|\boldsymbol{l}\|^2\int_{\mathbf{z}\in\mathcal{V}_2(\boldsymbol{l})} f(\mathbf{z})\,d\mathbf{z}\\ &= \frac{1}{n}\sum_{\boldsymbol{l}\in\Lambda_2} l^2 P_r(\mathbf{z}\in\mathcal{V}_2(\boldsymbol{l}))\end{aligned} \tag{26}$$

where $\mathcal{V}_2(\boldsymbol{l})$ is the Voronoi cell associated with the lattice point $\boldsymbol{l}\in\Lambda_2$.

For the one-dimensional ($n=1$) case, $P_r(\mathbf{z}\in\mathcal{V}_2(\boldsymbol{l}))$ can be expressed in terms of the erfc function, then $E_{\mathbf{Z}}[\|Q_{\Lambda_2}(\mathbf{Z})\|^2]$ becomes [10]

$$\begin{aligned}&\frac{1}{n}E_{\mathbf{Z}}[\|Q_{\Lambda_2}(\mathbf{Z})\|^2]\\ &\quad= \frac{1}{2}V_2^2\sum_{i=0}^{\infty}(2i+1)\text{erfc}\left(\frac{V_2}{\sqrt{2}\sigma_Z}\left(i+\frac{1}{2}\right)\right).\end{aligned} \tag{27}$$

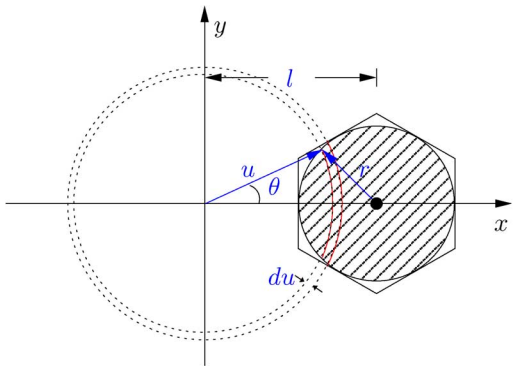


Fig. 3. Geometry used in evaluating $P_r(\mathbf{z}\in\mathcal{V}_2(\boldsymbol{l}))$ for the two-dimensional case.

We hence have

$$\begin{aligned}D_1(R) = \min_{V_2>0}\Bigg\{&G(\Lambda_1)V_2^2 2^{-2R}\\ &+\frac{1}{2}V_2^2\sum_{i=0}^{\infty}(2i+1)\text{erfc}\left(\frac{V_2}{\sqrt{2}\sigma_Z}\left(i+\frac{1}{2}\right)\right)\Bigg\}.\end{aligned} \tag{28}$$

For the case when $n>1$, since the distribution of $\mathbf{z}$ is spherically symmetric, without loss of generality, we can assume that $\boldsymbol{l}\in\Lambda_2$ lies on an axis (e.g., the horizontal axis, as shown in Fig. 3 for $n=2$) and use polar coordinate systems when computing $P_r(\mathbf{z}\in\mathcal{V}_2(\boldsymbol{l}))$. Let $\mathcal{S}$ be the packing sphere [7, p. 6] of the Voronoi region $\mathcal{V}_2(\boldsymbol{l})$. The volume of $\mathcal{S}$ is

$$V_{\mathcal{S}} = \frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}+1\right)}r^n = \Delta V_2. \tag{29}$$

When $n=2$, from the (u,θ) polar coordinate system in Fig. 3, we have

$$\begin{aligned}P_r(\mathbf{z}\in\mathcal{V}_2(\boldsymbol{l})) &\geq P_r(\mathbf{z}\in\mathcal{S})\\ &= \int_{l-r}^{l+r}\int_0^{2\cos^{-1}\left(\frac{l^2+u^2-r^2}{2lu}\right)} u\frac{1}{2\pi\sigma_Z^2}\\ &\quad\times\exp\left(-\frac{u^2}{2\sigma_Z^2}\right)d\theta du\\ &= \frac{1}{\pi}\int_{l-r}^{l+r}\cos^{-1}\left(\frac{l^2+u^2-r^2}{2lu}\right)\frac{u}{\sigma_Z^2}\\ &\quad\times\exp\left(-\frac{u^2}{2\sigma_Z^2}\right)du.\end{aligned} \tag{30}$$

Similarly, when $n=3$, from the polar coordinate system (u,θ,θ_1), we have

$$\begin{aligned}P_r(\mathbf{z}\in\mathcal{V}_2(\boldsymbol{l})) &\geq P_r(\mathbf{z}\in\mathcal{S})\\ &= \int_{l-r}^{l+r}\int_0^{\cos^{-1}\left(\frac{l^2+u^2-r^2}{2lu}\right)}\sin\theta\, d\theta\\ &\quad\times\int_0^{2\pi}d\theta_1 u^2\frac{1}{(2\pi)^{\frac{3}{2}}\sigma_Z^3}\exp\left(-\frac{u^2}{2\sigma_Z^2}\right)du\\ &= \frac{1}{\sqrt{2\pi}}\int_{l-r}^{l+r}\int_0^{\cos^{-1}\left(\frac{l^2+u^2-r^2}{2lu}\right)}\\ &\quad\times\sin\theta d\theta\frac{u^2}{\sigma_Z^3}\exp\left(-\frac{u^2}{2\sigma_Z^2}\right)du.\end{aligned} \tag{31}$$

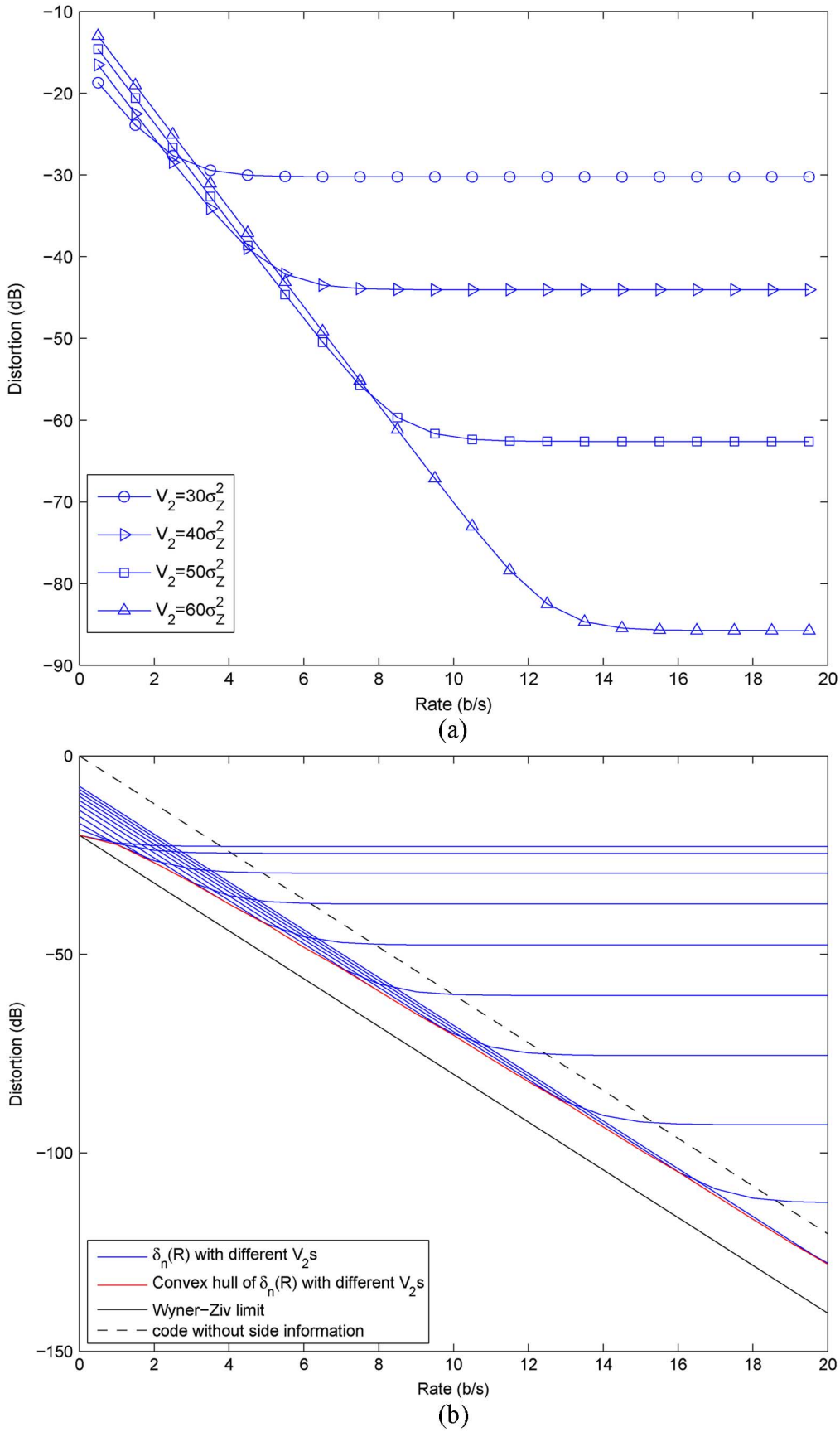


Fig. 4 (a) Distortions with different V_2's in the two-dimensional case. (b) $\bar{D}_n(R)$ is the convex hull of distortions for different V_2's.

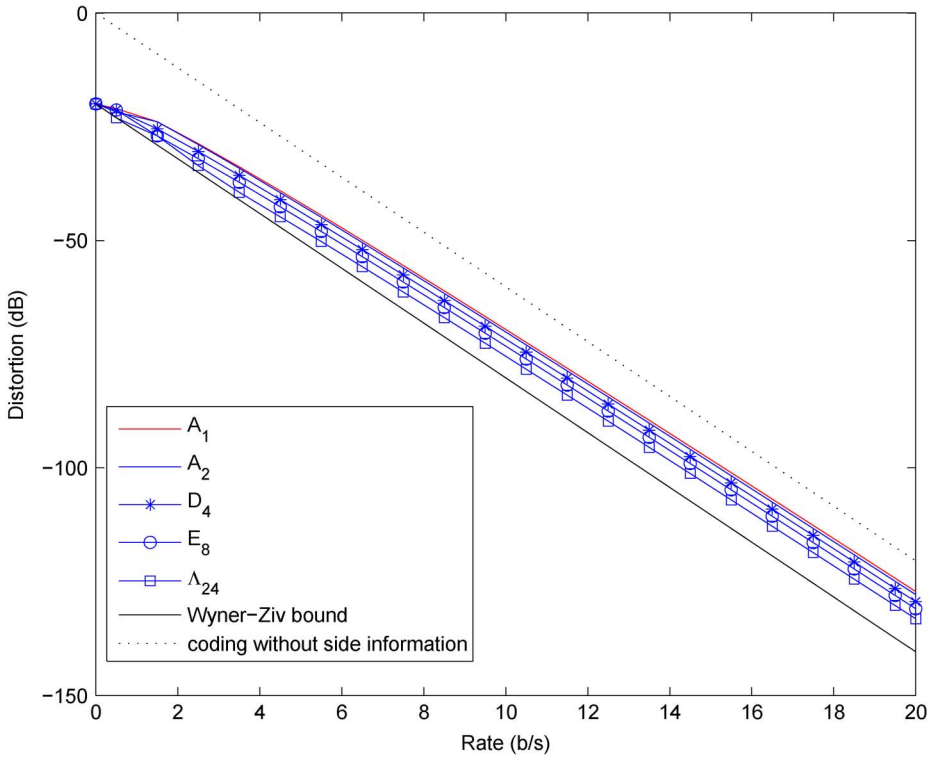


Fig. 5. The lower bound $\bar{D}(R)$ of $D_n(R)$ for different dimensions, with $\sigma_Z^2 = 0.01$, in the quadratic Gaussian case.

For $n > 3$, we generalize (31) by using the polar coordinate system $(u, \theta, \theta_1, \ldots, \theta_{n-2})$ and its Jacobian determinant [3, p. 904] to get

$$\begin{aligned} &P_r(\mathbf{z} \in \mathcal{V}_2(\boldsymbol{l})) \\ &\quad \geq P_r(\mathbf{z} \in \mathcal{S}) \\ &\quad = \int_{l-r}^{l+r} \int_0^{\cos^{-1}\left(\frac{l^2+u^2-r^2}{2lu}\right)} \sin^{n-2}\theta d\theta \int_0^{\pi} \sin^{n-3}\theta_1 d\theta_1 \\ &\qquad \cdots \int_0^{\pi} \sin\theta_{n-3} d\theta_{n-3} \\ &\qquad \times \int_0^{2\pi} d\theta_{n-2} \frac{u^{n-1}}{(2\pi)^{\frac{n}{2}}\sigma_Z^n} \exp\left(-\frac{u^2}{2\sigma_Z^2}\right) du \\ &\quad = \frac{n-1}{\Gamma\left(\frac{n+1}{2}\right) 2^{\frac{n}{2}} \pi^{\frac{1}{2}}} \int_{l-r}^{l+r} \int_0^{\cos^{-1}\left(\frac{l^2+u^2-r^2}{2lu}\right)} \\ &\qquad \times \sin^{n-2}\theta d\theta \frac{u^{n-1}}{\sigma_Z^n} \exp\left(-\frac{u^2}{2\sigma_Z^2}\right) du \end{aligned} \tag{32}$$

because

$$\int_0^{\pi} \sin^{n-3}\theta_1 d\theta_1 \cdots \int_0^{\pi} \sin\theta_{n-3} d\theta_{n-3} \int_0^{2\pi} d\theta_{n-2} = \frac{(n-1)\pi^{\frac{n-1}{2}}}{\Gamma(\frac{n+1}{2})}.$$

Combining (30)–(32) with (26), we obtain the lower bound $\bar{D}_C$ in (24) for D_C when $n > 1$. Hence we have (22). □

Fig. 4(a) shows distortions with different V_2's using nested A_2 lattices in two dimensions with $\sigma_Z^2 = 0.01$. The lower bound $\bar{D}_2(R)$ is the lower convex hull of all operational DR points with different V_2, as shown in Fig. 4(b). We observe from Fig. 4(b) that the gap (in decibels) from $\bar{D}_n(R)$ to $D_{\mathrm{WZ}}(R)$ keeps increasing as the rate increases. This is due to the fact that the source coding loss

$$D_S = G(\Lambda_1) V_1^{\frac{2}{n}} = \frac{1}{12 g(\Lambda_1)} \left(\frac{V_2}{N}\right)^{\frac{2}{n}} = \frac{1}{12 g(\Lambda_1)} V_2^{\frac{2}{n}} 2^{-2R}$$

is bounded away from 2^{-2R} with increasing V_2, where $g(\Lambda_1) \triangleq \frac{1}{12G(\Lambda_1)}$ is the granular gain [11] of lattice Λ_1, and $R = \frac{1}{n}\log N$, with N being the nesting ratio. Fig. 5 plots $\bar{D}_n(R)$ for $n = 1, 2, 4, 8$ and 24 with $\sigma_Z^2 = 0.01$. We see that for fixed but finite n, the gap (in decibels) between $\bar{D}_n(R)$ and the Wyner–Ziv DR function $D_{\mathrm{WZ}}(R) = \sigma_Z^2 2^{-2R}$ is an increasing function of R, and that for fixed R, it is a decreasing function of n.

From (10), (26), (30), and (32), we see that the gap between $D_n(R)$ and our lower bound $\bar{D}_n(R)$ for $n > 1$ is

$$\frac{1}{n} \sum_{\boldsymbol{l} \in \Lambda_2} l^2 [P_r(\mathbf{z} \in \mathcal{V}_2(\boldsymbol{l})) - P_r(\mathbf{z} \in \mathcal{S})]. \tag{33}$$

Since $\mathcal{V}_2(\boldsymbol{l})$ will be more and more spherical-like as n increases, our lower bound is asymptotically tight as n goes to infinity. When $n = 2$, as we shall see from Fig. 11(b) in Section VI-C, this gap is 0.6 dB at high rate. The following corollary asserts that this gap is independent of σ_Z^2 (or the correlation between X and Y).

D. Performance Under Varying Source Correlation in the Quadratic Gaussian Case

With the source model $X = Y + Z$ in the quadratic Gaussian case, we have $\sigma_Z^2 = \sigma_X^2(1-\rho^2) = (\sigma_Y^2+\sigma_Z^2)(1-\rho^2)$. If σ_Y^2 is fixed, then the correlation coefficient ρ varies with σ_Z^2.

Corollary 4.3: For fixed rate R and dimensionality n, the lower bound $\bar{D}_n(R)$ of $D_n(R)$ in (22) remains a constant gap (in decibels) from the Wyner–Ziv limit $D_{\mathrm{WZ}}(R)$ for all $\sigma_Z^2 > 0$.

Proof: For any point $\boldsymbol{l} \in \Lambda_2$, according to (2), $\boldsymbol{l} = M_2\mathbf{i}$ for some $\boldsymbol{i} \in \mathbb{Z}^n$, then $l^2 = \|\boldsymbol{l}\|^2 = \boldsymbol{i}^T M_2^T M_2 \boldsymbol{i}$ and $V_2 = [\det(M_2^T M_2)]^{\frac{1}{2}}$ [7, p. 4]. Consider a *similar* lattice [6] of Λ_2, denoted as Λ_2', with packing radius $r' = 1$ and generator matrix M_2', then Λ_2 is a scaled version of Λ_2' with scaling factor r, which is also the packing radius of Λ_2 defined as $r = \pi^{-\frac{1}{2}}(\Delta\Gamma(\frac{n}{2}+1)V_2)^{\frac{1}{n}}$ in Corollary 4.2. Thus, $M_2 = rM_2'$ and $\boldsymbol{l} = rM_2'\boldsymbol{i}$.

Let $\beta = \sqrt{\boldsymbol{i}^T M_2'^T M_2' \boldsymbol{i}}$, $\tilde{u} = \frac{u}{\sigma_Z}$, and $\tilde{r} = \frac{r}{\sigma_Z}$, then $l = \beta r = \beta\sigma_Z\tilde{r}$. Starting from (23) and (24), we get

$$\begin{aligned} D_S(R) &= G(\Lambda_1)V_2^{\frac{2}{n}}2^{-2R} \\ &= G(\Lambda_1)\pi(\Delta\Gamma(\frac{n}{2}+1))^{-\frac{2}{n}}\sigma_Z^2\tilde{r}^2 2^{-2R} \end{aligned} \tag{34}$$

and

$$\begin{aligned} \bar{D}_C &= \frac{n-1}{n\Gamma(\frac{n+1}{2})2^{\frac{n}{2}}\pi^{\frac{1}{2}}} \sum_{\boldsymbol{l}\in\Lambda_2} l^2 \int_{l-r}^{l+r} \int_0^{\cos^{-1}(\frac{l^2+u^2-r^2}{2lu})} \\ &\quad \times \sin^{n-2}\theta d\theta \frac{u^{n-1}}{\sigma_Z^n} \exp\left(-\frac{u^2}{2\sigma_Z^2}\right) du \\ &= \frac{(n-1)\sigma_Z^2\tilde{r}^2}{n\Gamma\left(\frac{n+1}{2}\right)2^{\frac{n}{2}}\pi^{\frac{1}{2}}} \sum_{\boldsymbol{i}\in\mathbb{Z}^n;\beta=\sqrt{\boldsymbol{i}^T M_2'^T M_2'\boldsymbol{i}}} \beta^2 \int_{(\beta-1)\tilde{r}}^{(\beta+1)\tilde{r}} \\ &\quad \times \int_0^{\cos^{-1}(\frac{(\beta^2-1)\tilde{r}^2+\tilde{u}^2}{2\beta\tilde{r}\tilde{u}})} \sin^{n-2}\theta\, d\theta \tilde{u}^{n-1} \\ &\quad \times \exp\left(-\frac{\tilde{u}^2}{2}\right) d\tilde{u}. \end{aligned} \tag{34}$$

Thus, the sum of $D_S(R)$ and $\bar{D}_C$ can be written as

$$D_S(R) + \bar{D}_C = \sigma_Z^2 \Psi_{\Lambda_1,\Lambda_2}(\tilde{r}, R) \tag{35}$$

with the function $\Psi_{\Lambda_1,\Lambda_2}(\tilde{r}, R)$ depending on σ_Z only through $\tilde{r} = \frac{r}{\sigma_Z}$. Therefore, for fixed R, both the optimal $\tilde{r}$, denoted as $\tilde{r}^*$, which minimizes $\Psi_{\Lambda_1,\Lambda_2}(\tilde{r}, R)$, and the resulting minimal $\Psi_{\Lambda_1,\Lambda_2}(\tilde{r}^*, R)$ are independent of σ_Z. Then the lower bound of distortion is

$$\begin{aligned} \bar{D}_n(R) &= \min_{V_2>0}(D_S(R) + \bar{D}_C) \\ &= \min_{\tilde{r}>0} \sigma_Z^2 \Psi_{\Lambda_1,\Lambda_2}(\tilde{r}, R) = \sigma_Z^2 \Psi_{\Lambda_1,\Lambda_2}(\tilde{r}^*, R) \end{aligned} \tag{36}$$

which is proportional to σ_Z^2. Since the Wyner–Ziv limit is $D_{\mathrm{WZ}}(R) = \sigma_Z^2 2^{-2R}$, then the gap (in decibels) between $\bar{D}_n(R)$ and $D_{\mathrm{WZ}}(R)$ for fixed R and n is

$$10\log_{10}\frac{\bar{D}_n(R)}{D_{\mathrm{WZ}}(R)} = 10\log_{10}\frac{\Psi_{\Lambda_1,\Lambda_2}(\tilde{r}^*, R)}{2^{-2R}} \tag{37}$$

which is a constant that is independent of σ_Z^2 (or the correlation between X and Y). □

In addition to the above result about the constant gap (in decibesl) between the lower bound $\bar{D}_n(R)$ and $D_{\mathrm{WZ}}(R)$ for varying source correlation, we have a similar result when $\bar{D}_n(R)$ is replaced by $D_n(R)$.

Corollary 4.4: For $X = Y + Z, Z \sim N(0, \sigma_Z^2)$, the operational DR function

$$D_n(R) = \min_{V_2>0}\left\{G(\Lambda_1)V_2^{\frac{2}{n}}2^{-2R} + \frac{1}{n}E_{\mathbf{Z}}[\|Q_{\Lambda_2}(\mathbf{Z})\|^2]\right\} \tag{38}$$

of Wyner–Ziv coding of X (with decoder side information Y) using n-dimensional nested lattice quantizers at high but fixed rate R remains a constant gap (in decibels) from the Wyner–Ziv limit $D_{\mathrm{WZ}}(R)$ for all $\sigma_Z^2 > 0$.

Proof: It suffices to show that $D_n(R)$ is proportional to σ_Z^2. Let $\tilde{\mathbf{z}} = \frac{\mathbf{z}}{\sigma_Z}$ and define lattice $\tilde{\Lambda}_1$ and $\tilde{\Lambda}_2$ with generator matrix $\tilde{M}_1 = \frac{1}{\sigma_Z}M_1$ and $\tilde{M}_2 = \frac{1}{\sigma_Z}M_2$, respectively, then for any lattice point $\tilde{\boldsymbol{l}} \in \tilde{\Lambda}_2$, $\tilde{\boldsymbol{l}} = \frac{1}{\sigma_Z}\boldsymbol{l}$, where $\boldsymbol{l}$ is the corresponding lattice point of Λ_2. In addition, $\tilde{l} = \|\tilde{\boldsymbol{l}}\| = \frac{1}{\sigma_Z}l$, and the volume of the Voronoi cell of $\tilde{\Lambda}_2$ is $\tilde{V}_2 = \frac{1}{\sigma_Z^n}V_2$. Since the normalized second moment remains unchanged with respect to lattice scaling and rotation, we have

$$\begin{aligned} D_n(R) &= \min_{V_2>0}\left\{G(\Lambda_1)V_2^{\frac{2}{n}}2^{-2R} + \frac{1}{n}E_{\mathbf{Z}}\left[\|Q_{\Lambda_2}(\mathbf{Z})\|^2\right]\right\} \\ &= \min_{V_2>0}\left\{G(\Lambda_1)V_2^{\frac{2}{n}}2^{-2R} + \frac{1}{n}\sum_{\boldsymbol{l}\in\Lambda_2} l^2 \right. \\ &\quad \left. \times \int_{\mathbf{z}\in\mathcal{V}_2(\boldsymbol{l})} \frac{1}{(2\pi\sigma_Z^2)^{\frac{n}{2}}}\exp\left(-\frac{\|\mathbf{z}\|^2}{2\sigma_Z^2}\right) d\mathbf{z}\right\} \\ &= \sigma_Z^2 \min_{\tilde{V}_2>0}\left\{G(\tilde{\Lambda}_1)\tilde{V}_2^{\frac{2}{n}}2^{-2R} + \frac{1}{n}\sum_{\tilde{\boldsymbol{l}}\in\tilde{\Lambda}_2}\tilde{l}^2 \right. \\ &\quad \left. \times \int_{\tilde{\mathbf{z}}\in\tilde{\mathcal{V}}_2(\tilde{\boldsymbol{l}})} \frac{1}{(2\pi)^{\frac{n}{2}}}\exp\left(-\frac{\|\tilde{\mathbf{z}}\|^2}{2}\right) d\tilde{\mathbf{z}}\right\} \end{aligned} \tag{39}$$

which is proportional to σ_Z^2. □

V. Slepian–Wolf Coded Nested Lattice Quantization (SWC-NQ)

A. Motivation of SWC-NQ

Recall from Theorem 4.1 that the distortion per dimension of the nested lattice quantizer is $D_n = D_S + D_C$, where $D_S = G(\Lambda_1)V_1^{\frac{2}{n}}$ is the source coding loss, characterized by the granular gain $g(\Lambda_1)$ of Λ_1 and the boundary gain $b(\Lambda_2)$ of Λ_2, whereas the channel coding loss $D_C \triangleq \frac{1}{n}E_{\mathbf{Z}}[\|Q_{\Lambda_2}(\mathbf{Z})\|^2]$ is characterized by the error probability of lattice channel decoding. Suppose the coarse lattice Λ_2 with Voronoi region $\mathcal{V}_2$ in the n-dimensional space has the same overload probability as a cubic support region of side E centered at the origin, then,

according to [11], $b(\Lambda_2)$ is defined as the ratio of the normalized volume E^2 of the cubic support region to the normalized volume $V_2^{\frac{2}{n}}$. That is,

$$b(\Lambda_2) = \frac{E^2}{V_2^{\frac{2}{n}}}. \tag{40}$$

Then

$$D_S = G(\Lambda_1)V_1^{\frac{2}{n}} = \frac{1}{12g(\Lambda_1)}V_2^{\frac{2}{n}}N^{-\frac{2}{n}} = \frac{1}{12g(\Lambda_1)}\frac{E^2}{b(\Lambda_2)}N^{-\frac{2}{n}} \tag{41}$$

where N is the nesting ratio. If N is fixed, D_S will decrease with increasing $b(\Lambda_2)$; but D_C remains unaffected because the overload probability stays unchanged.

To increase the boundary gain $b(\Lambda_2)$, a second-stage of binning can be applied to the quantization indices. The essence of binning is using a channel code to partition the support region into cosets. Assume ideal channel code is employed to partition the support region $\mathcal{V}_2$ into k cosets without decoding errors and denote the set consisting of the coset leaders as $\mathcal{L}$, then $|\mathcal{L}| = k$ and the support region for the quantization indices (or the nested quantizer), which is also the set of the quantization cells associated with $\mathcal{L}$, has volume $\frac{V_2}{k}$. Thus, the *effective* volume of the support region decreases by a factor of k after the second stage of binning, and therefore the boundary gain $b(\Lambda_2)$ increases by a factor of $k^{\frac{2}{n}}$.

We thus propose a framework for Wyner–Ziv coding of i.i.d. sources based on SWC-NQ, which follows NQ by Slepian–Wolf coding to perform second-stage binning. Despite the fact that there is almost no correlation among the nested quantization indices that identify the coset leaders $\mathbf{s} = \mathbf{x}_{Q_{\Lambda_1}} - Q_{\Lambda_2}(\mathbf{x}_{Q_{\Lambda_1}}) \in \Lambda_1/\Lambda_2$ of nested lattice pair (Λ_1, Λ_2), there still remains correlation between $\mathbf{s}$ and the side information $\mathbf{y}$, especially at high rate. Write $\mathbf{W} = Q_{\Lambda_1}(\mathbf{X})$ and $\mathbf{S} = \mathbf{W} - Q_{\Lambda_2}(\mathbf{W})$. Ideal Slepian–Wolf coding can be used to compress $\mathbf{S}$ to the rate of $R = \frac{1}{n}H(\mathbf{S}\,|\,\mathbf{Y})$ per dimension. In practice, state-of-the-art channel codes, such as LDPC codes, can be used to approach the Slepian–Wolf limit $\frac{1}{n}H(\mathbf{S}\,|\,\mathbf{Y})$ [20]. The role of Slepian–Wolf coding in SWC-NQ is thus to exploit the correlation between $\mathbf{S}$ and $\mathbf{Y}$ for further compression.

B. High-Rate Performance for the Quadratic Gaussian Case

Lemma 5.1: For the quadratic Gaussian case, a lower bound for the high-rate performance of SWC-NQ with a pair of nested lattices (Λ_1, Λ_2) is given as

$$D_n(R) \geq G(\Lambda_1)2^{\frac{2}{n}h'(\mathbf{X},\Lambda_2)}\sigma_Z^2 2^{-2R} + \bar{D}_C \tag{42}$$

where

$$h'(\mathbf{X}, \Lambda_2) \triangleq -\int_{\mathbf{x}\in\mathbb{R}^n} \bar{f}(\mathbf{x}) \log_2 \left[\sum_{j=-\infty}^{\infty} \bar{f}(\mathbf{x} + \frac{\boldsymbol{c}_j(\mathbf{0})}{\sigma_Z}) \right] d\mathbf{x} \tag{43}$$

$\bar{f}(\cdot)$ is the probability density function (pdf) of an n-dimensional i.i.d. Gaussian source with mean $\mathbf{0}$ and covariance matrix $\mathbf{I}_{n\times n}$, $\boldsymbol{c}_j(\mathbf{0})$ is defined in Section III-B as lattice points of Λ_2, and $\bar{D}_C$ is the same as given in (24).

Proof: See Appendix A. □

The lower bounds of $D_n(R)$ for $n = 1$ and $n = 2$ with different V_2 are plotted in Fig. 6. It indicates that the *best* performance of a Slepian–Wolf coded nested scalar quantizer remains a constant gap (in decibels) from the Wyner–Ziv limit at high rate. Here the *best* means that the minimal achievable distortion D over all possible V_2 for a given rate R. Before rigorously stating our main result in a theorem, we give the following lemma.

Lemma 5.2: For nested lattice quantization with $\mathbf{W} = Q_{\Lambda_1}(\mathbf{X})$ and $\mathbf{S} = \mathbf{W} - Q_{\Lambda_2}(\mathbf{W}), H(\mathbf{S}\,|\,\mathbf{Y}) = H(\mathbf{W}\,|\,\mathbf{Y})$ at high rate.

Proof: See Appendix B. □

Theorem 5.2: For the quadratic Gaussian case, the optimal DR performance of SWC-NQ using n-dimensional nested lattices at high rate is

$$D_n^*(R) \triangleq \min_{V_2} D_n(R) = 2\pi e G(\Lambda_1)\sigma_Z^2 2^{-2R}. \tag{44}$$

Proof: 1) By Lemma 5.2

$$\begin{aligned} nR &= H(\mathbf{S}\,|\,\mathbf{Y}) = H(\mathbf{W}\,|\,\mathbf{Y}) \\ &= H\left(Q_{\Lambda_1}(\mathbf{X})\,|\,\mathbf{Y}\right) \\ &= h(\mathbf{X}\,|\,\mathbf{Y}) - \log V_1 \\ &= \frac{n}{2}\log\left(2\pi e\sigma_Z^2\right) - \log V_1 \end{aligned} \tag{45}$$

and $D_n(R) = D_S + D_C = D_S = G(\Lambda_1)V_1^{\frac{2}{n}}$ since $D_C = 0$ under ideal Slepian–Wolf coding. Combine R and D_n through V_1 and we get the DR function as

$$\lim_{V_2\to\infty} D_n(R) = 2\pi e G(\Lambda_1)\sigma_Z^2 2^{-2R}. \tag{46}$$

Since

$$D_n^*(R) = \min_{V_2} D_n(R) \leq \lim_{V_2\to\infty} D_n(R)$$

we have

$$D_n^*(R) \leq 2\pi e G(\Lambda_1)\sigma_Z^2 2^{-2R}. \tag{47}$$

2) Denote $\mathbf{w} \triangleq Q_{\Lambda_1}(\mathbf{x})$, and

$$\mathcal{S}_1 \triangleq \{(\mathbf{X}, \hat{\mathbf{X}}) : \frac{1}{n}E[d(\mathbf{X}, \hat{\mathbf{X}})] \leq D_n\}.$$

The rate of Wyner–Ziv coding with respect to D_n is [37]

$$\begin{aligned} nR^*(D_n) &= \min_{p(\mathbf{s}),p(\hat{\mathbf{x}}\,|\,\mathbf{s},\mathbf{y}),(\mathbf{X},\hat{\mathbf{X}})\in\mathcal{S}_1} I(\mathbf{X};\mathbf{S}\,|\,\mathbf{Y}) \\ &\overset{(a)}{=} \min_{p(\mathbf{s}),p(\hat{\mathbf{x}}\,|\,\mathbf{s},\mathbf{y}),(\mathbf{X},\hat{\mathbf{X}})\in\mathcal{S}_1} H(\mathbf{S}\,|\,\mathbf{Y}) \\ &\overset{(b)}{=} \min_{p(\mathbf{s}),p(\hat{\mathbf{x}}\,|\,\mathbf{s},\mathbf{y}),(\mathbf{X},\hat{\mathbf{X}})\in\mathcal{S}_1} H(\mathbf{W}\,|\,\mathbf{Y}). \end{aligned} \tag{48}$$

where (a) comes from $H(\mathbf{S}\,|\,\mathbf{X},\mathbf{Y}) = 0$ and (b) is due to Lemma 5.2.

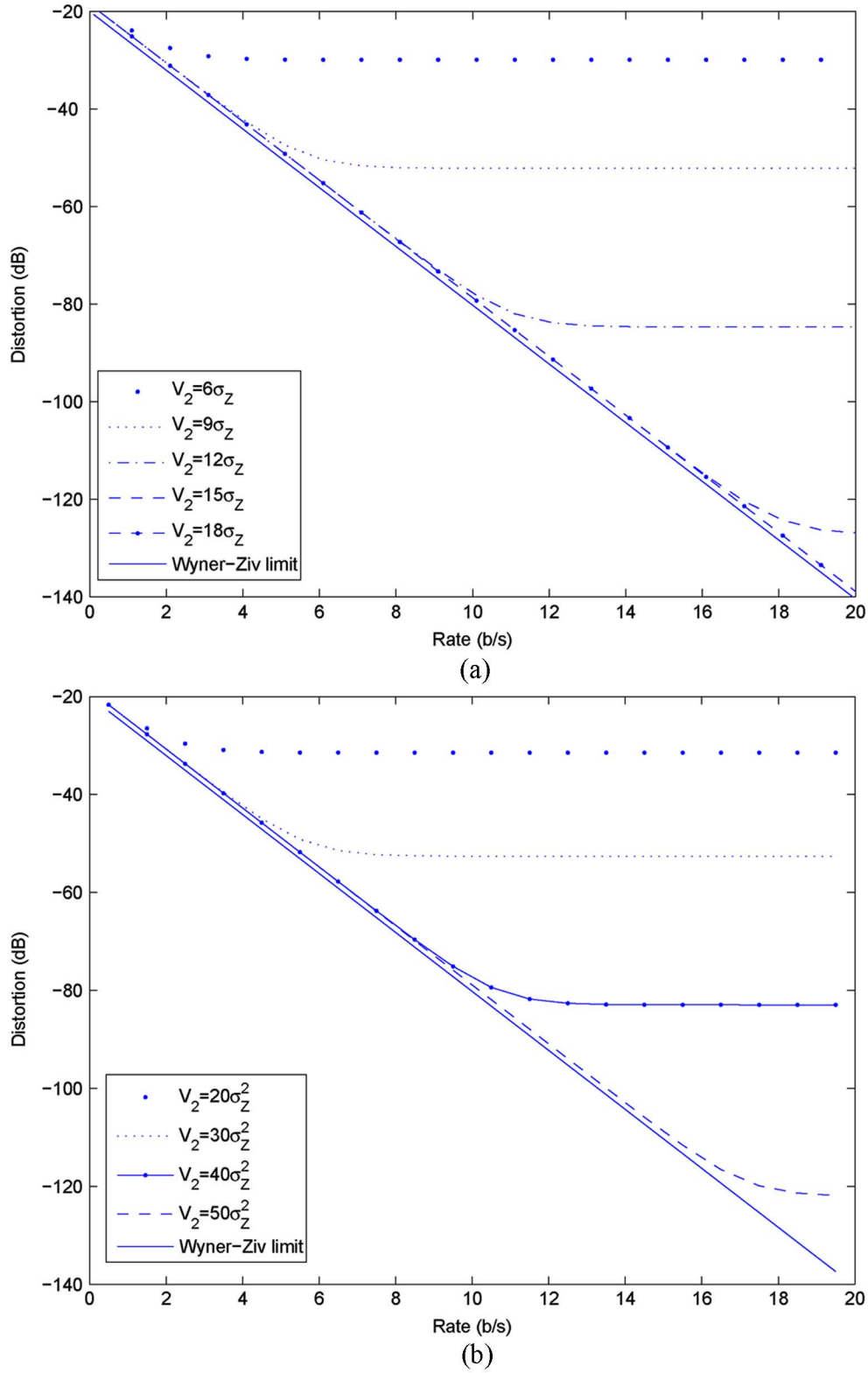


Fig. 6. Lower bounds of (a) $D_1(R)$ and (b) $D_2(R)$ with different V_2's.

Define $\mathcal{S}_2 \triangleq \{(\mathbf{X},\mathbf{W}) : \frac{1}{n}E[d(\mathbf{X},\mathbf{W})] \le D_n\}$. From Theorem 4.1

$$\begin{aligned}\frac{1}{n}E[d(\mathbf{X},\hat{\mathbf{X}})] &= G(\Lambda_1)V_1^{\frac{2}{n}} + \frac{1}{n}E_{\mathbf{Z}}\left[\|Q_{\Lambda_2}(\mathbf{Z})\|^2\right] \\ &= \frac{1}{n}E[d(\mathbf{X},\mathbf{W})] + \frac{1}{n}E_{\mathbf{Z}}\left[\|Q_{\Lambda_2}(\mathbf{Z})\|^2\right] \\ &\ge \frac{1}{n}E[d(\mathbf{X},\mathbf{W})] \end{aligned} \tag{49}$$

then for $\forall(\mathbf{X},\hat{\mathbf{X}}) \in \mathcal{S}_1$

$$D_n \ge \frac{1}{n}E[d(\mathbf{X},\hat{\mathbf{X}})] \ge \frac{1}{n}E[d(\mathbf{X},\mathbf{W})], \tag{50}$$

which means $(\mathbf{X}, \hat{\mathbf{X}}) \in \mathcal{S}_2$. Hence, $\mathcal{S}_1 \subseteq \mathcal{S}_2$, and

$$nR^*(D_n) = \min_{p(\mathbf{s}),p(\hat{\mathbf{x}}\,|\,\mathbf{s},\mathbf{y}),(\mathbf{X},\hat{\mathbf{X}})\in\mathcal{S}_1} H(\mathbf{W}\,|\,\mathbf{Y}) \geq \min_{p(\mathbf{s}),p(\hat{\mathbf{x}}\,|\,\mathbf{s},\mathbf{y}),(\mathbf{X},\hat{\mathbf{X}})\in\mathcal{S}_2} H(\mathbf{W}\,|\,\mathbf{Y}). \quad (51)$$

Since

$$H(\mathbf{W}\,|\,\mathbf{Y}) = \frac{n}{2}\log(2\pi e\sigma_Z^2) - \log V_1$$

and

$$\frac{1}{n}E[d(\mathbf{X},\mathbf{W})] = G(\Lambda_1)V_1^{\frac{2}{n}}$$

(51) becomes

$$\begin{aligned} nR^*(D_n) &\geq \min_{p(\mathbf{s}),p(\hat{\mathbf{x}}\,|\,\mathbf{s},\mathbf{y}),(\mathbf{X},\hat{\mathbf{X}})\in\mathcal{S}_2} \frac{n}{2}\log\left(2\pi e\sigma_Z^2\right) - \log V_1 \\ &= \min_{\frac{1}{n}E[d(\mathbf{X},\mathbf{W})]\leq D_n} \times \left\{\frac{n}{2}\log\left(2\pi e\sigma_Z^2\right) - \frac{n}{2}\log\frac{\frac{1}{n}E[d(\mathbf{X},\mathbf{W})]}{G(\Lambda_1)}\right\} \\ &= \min_{\frac{1}{n}E[d(\mathbf{X},\mathbf{W})]\leq D_n} \frac{n}{2}\log\frac{2\pi eG(\Lambda_1)\sigma_Z^2}{\frac{1}{n}E[d(\mathbf{X},\mathbf{W})]} \\ &= \frac{n}{2}\log\frac{2\pi eG(\Lambda_1)\sigma_Z^2}{D_n}. \end{aligned} \quad (52)$$

We thus have

$$D_n^*(R) \geq 2\pi eG(\Lambda_1)\sigma_Z^2 2^{-2R}. \quad (53)$$

Combining (47) and (53), we conclude that, at high rate, the *best* DR performance of the quadratic Gaussian SWC-NQ using n-dimensional lattices is

$$D_n^*(R) = 2\pi eG(\Lambda_1)\sigma_Z^2 2^{-2R}. \quad (54)$$

□

Thus, at high rate and for the quadratic Gaussian case, SWC-NQ performs the same as classic entropy-coded lattice quantization with the side information available at both the encoder and decoder. Specifically, the DR functions with one-dimensional (scalar) lattice and two-dimensional (hexagonal) nested lattices are 1.53 and 1.36 dB away from the Wyner–Ziv DR function, respectively.

We found that for finite rate R and small n (e.g., $n = 1$ and 2), the optimal V_2, denoted as V_2^*, that minimizes the distortion $D_n(R)$ is also finite. Fig. 7(a) and (b) plots the optimal V_2^* as functions of R for $n = 1$ and $n = 2$ (scaled by σ_Z and σ_Z^2, respectively). We see that as R goes to infinity, V_2^* also goes to infinity. In addition, we observe from Fig. 6 that for fixed R and n, $D_n(R)$ stays roughly unchanged for $V_2 > V_2^*$.

Remarks:

- Since SWC-NQ has its root in variable-rate quantization, it is not surprising to see (54) as an elegant generalization of the result in entropy-coded quantization from classic source coding to Wyner–Ziv coding.
- SWC-NQ relies on conditional entropy coding (or Slepian–Wolf coding implemented via channel coding) to achieve rate savings after nested lattice quantization. That fact that (54) holds means two things: 1) Just like entropy coding can achieve all the boundary gain in classic source coding of Gaussian sources [18], Slepian–Wolf coding in SWC-NQ can realize all the remaining boundary gain left unexploited by the coarse lattice channel code of the nested lattice quantizer. 2) Ideal Slepian–Wolf coding also renders the channel code loss D_C in $D_n = D_S + D_C$ to zero. Thus, with ideal Slepian-Wolf coding, the only remaining loss in D_n is the granular loss portion of D_S, which is maximally 1.53 dB [11].

Similar to Corollary 4.4, we give (without proof) the following corollary of Theorem 5.2 for high-rate SWC-NQ.

Corollary 5.5: For fixed dimensionality n, the optimal performance of high-rate SWC-NQ in the quadratic Gaussian case is $10\log_{10} 2\pi eG(\Lambda_1)$ decibels from the Wyner–Ziv limit $D_{\mathrm{WZ}}(R)$ for all $\sigma_Z^2 > 0$ (or any correlation between X and Y).

The above result is stronger than that in Corollary 4.3 because the $10\log_{10} 2\pi eG(\Lambda_1)$ decibel gap is also independent of the rate R. We thus conclude that the high-rate performance loss of Wyner–Ziv coding of quadratic Gaussian sources with nested lattice quantization is independent of the source correlation, regardless of whether Slepian–Wolf coding is used or not. There are only high-dimension asymptotics with nested lattice quantization (studied in [42]) and high-rate asymptotics with SWC-NQ (presented in Theorem 5.2).

VI. Code Design and Simulation Results

A. Design of Nested Lattice Quantizer

The nested lattice quantizer design problem involves optimizing the nesting in the quantizer encoder (under a fixed nesting ratio N) and devising the minimum MSE estimator at the decoder.

1) Optimizing the Nesting Scheme of Lattices: Lattices with the densest packing (i.e., the best channel code) and the thinnest covering (i.e., the best source code) are introduced in [7]. For example, the hexagon lattice A_2 is the best lattice source code and the best lattice channel code at $n = 2$. To optimize the nesting scheme, Λ_2 should be *clean* and geometrically similar to Λ_1 [6], [31], where the former means that the Voronoi cell boundaries for Λ_2 do not intersect Λ_1. We follow the scheme suggested by Conway *et al.* in [6] to search for clean and similar nested lattice pairs. Fig. 8 illustrates the nesting lattice pair for A_2 with nesting ratio $N = 31$.

2) The Optimal Decoder: The optimal decoder for the nested lattice quantizer is the one that minimizes the MSE between X and the reconstructed $\hat{X}$.

Theorem 6.3: The minimum MSE decoder for the nested lattice quantizer is

$$\hat{\mathbf{x}} = E[\mathbf{X}\,|\,\mathbf{s},\mathbf{y}] = \frac{\int_{R(\mathbf{s})} \mathbf{x} f(\mathbf{x}\,|\,\mathbf{y})\,d\mathbf{x}}{\int_{R(\mathbf{s})} f(\mathbf{x}\,|\,\mathbf{y})\,d\mathbf{x}} \quad (55)$$

where $\mathbf{s} = \mathbf{x}_{Q_{\Lambda_1}} - Q_{\Lambda_2}(\mathbf{x}_{Q_{\Lambda_1}}) \in \Lambda_1/\Lambda_2$ is the received coset leader for $\mathbf{x}$ at the decoder.

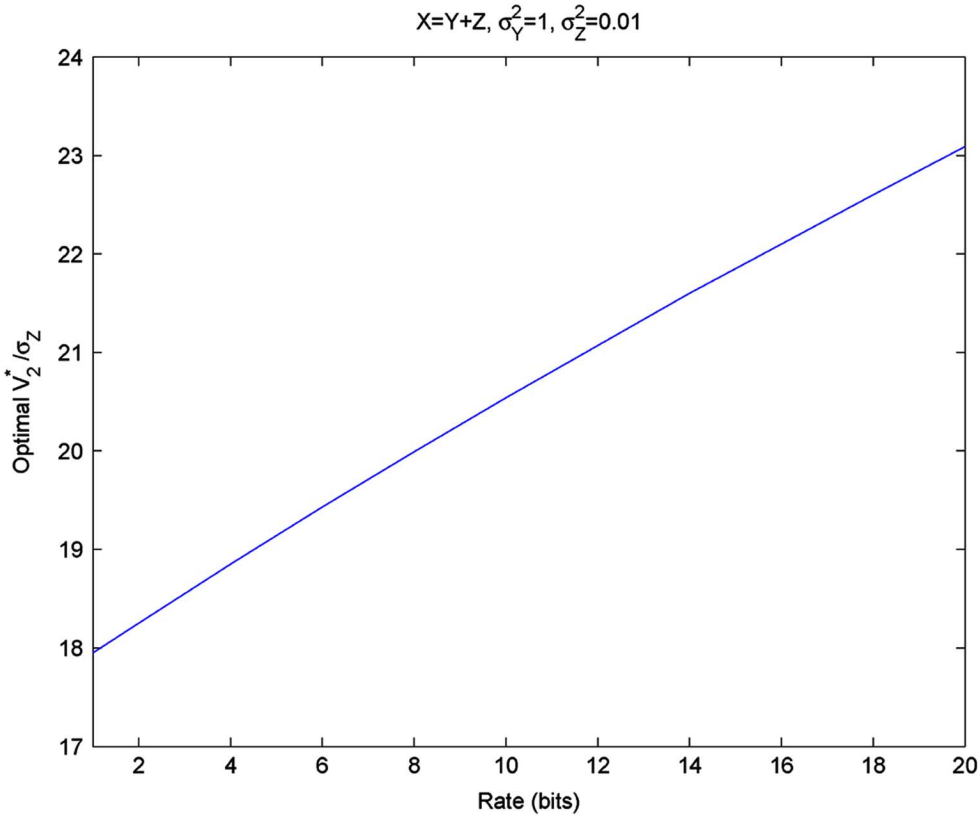


Fig. 7. (a) $\frac{V_2^*}{\sigma_Z}$ versus R for $n = 1$ and (b) $\frac{V_2^*}{\sigma_Z^2}$ versus R for $n = 2$.

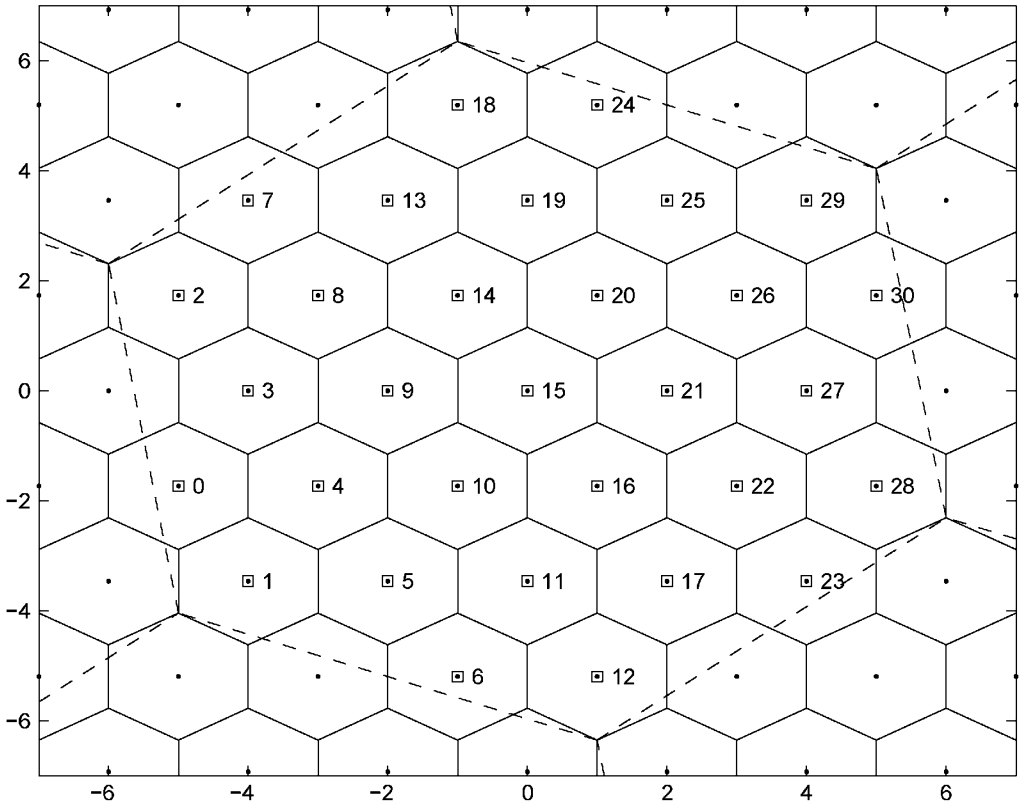


Fig. 8. A clean and geometrically similar nested hexagonal lattice pair with nesting ratio $N = 31$.

Proof:

$$\begin{aligned}\hat{\mathbf{x}} &= E[\mathbf{X}\,|\,\mathbf{s},\mathbf{y}] = \int_{\mathbb{R}^n} \mathbf{x} f(\mathbf{x}\,|\,\mathbf{s},\mathbf{y})\,d\mathbf{x} \\ &= \int_{\mathbb{R}^n} \mathbf{x}\frac{f(\mathbf{x},\mathbf{s}\,|\,\mathbf{y})}{p(\mathbf{s}\,|\,\mathbf{y})}\,d\mathbf{x} \\ &= \frac{\int_{\mathbb{R}^n} \mathbf{x} f(\mathbf{x},\mathbf{s}\,|\,\mathbf{y})\,d\mathbf{x}}{p(\mathbf{s}\,|\,\mathbf{y})} \\ &= \frac{\int_{\mathbb{R}^n} \mathbf{x} f(\mathbf{x}\,|\,\mathbf{y}) p(\mathbf{s}\,|\,\mathbf{x},\mathbf{y})\,d\mathbf{x}}{\int_{R(\mathbf{s})} f(\mathbf{x}\,|\,\mathbf{y})\,d\mathbf{x}}. \end{aligned} \tag{56}$$

Since $\mathbf{S},\mathbf{X}$ and $\mathbf{Y}$ form a Markov chain $\mathbf{Y} \longleftrightarrow \mathbf{X} \longleftrightarrow \mathbf{S}$

$$p(\mathbf{s}\,|\,\mathbf{x},\mathbf{y}) = p(\mathbf{s}\,|\,\mathbf{x}) = \begin{cases} 0, & \text{if } \mathbf{x} \notin R(\mathbf{s}) \\ 1, & \text{if } \mathbf{x} \in R(\mathbf{s}). \end{cases} \tag{57}$$

Thus,

$$\hat{\mathbf{x}} = \frac{\int_{\mathbb{R}^n} \mathbf{x} f(\mathbf{x}\,|\,\mathbf{y}) p(\mathbf{s}\,|\,\mathbf{x})\,d\mathbf{x}}{\int_{R(\mathbf{s})} f(\mathbf{x}\,|\,\mathbf{y})\,d\mathbf{x}} = \frac{\int_{R(\mathbf{s})} \mathbf{x} f(\mathbf{x}\,|\,\mathbf{y})\,d\mathbf{x}}{\int_{R(\mathbf{s})} f(\mathbf{x}\,|\,\mathbf{y})\,d\mathbf{x}}. \tag{58}$$

Note that this optimal quantizer decoder (58), which is in the form of a nonlinear estimator, is consistent with the centroid condition in classic minimum MSE lattice (or vector) quantizer design [18].

As an example, in the quadratic Gaussian case with $X = Y + Z, Y \sim N(0,\sigma_Y^2)$ and $Z \sim N(0,\sigma_Z^2)$, when $n = 1$

$$f(x\,|\,y) = \frac{1}{\sqrt{2\pi\sigma_Z^2}} e^{-\frac{(x-y)^2}{2\sigma_Z^2}}$$

the optimal decoder for a nested scalar quantizer can be expressed as

$$\hat{x} = \frac{\sum_{j=-\infty}^{\infty} \int_{sq_1+jq_2}^{(s+1)q_1+jq_2} x e^{-\frac{(x-y)^2}{2\sigma_Z^2}}\,dx}{\sum_{j=-\infty}^{\infty} \int_{sq_1+jq_2}^{(s+1)q_1+jq_2} e^{-\frac{(x-y)^2}{2\sigma_Z^2}}\,dx} \tag{59}$$

where q_1 and q_2 are the step sizes of the nested scalar quantizers Λ_1 and Λ_2 with $q_2 = Nq_1$ and N being the nesting ratio.

Nonlinear estimation at the decoder plays an important role at low rate. Fig. 9 shows the improvement gained at low rate by using the nonlinear minimum MSE estimator of (58) versus the linear estimator $\mathbf{s} + Q_{\Lambda_2}(\mathbf{y} - \mathbf{s})$ valid for high rate (see Section IV) for $n = 2$ with $\sigma_Y^2 = 1$ and $\sigma_Z^2 = 0.01$.

Corollary 6.6: The nonlinear minimum MSE estimator of (58) degenerates to the linear one $\hat{\mathbf{x}} = \mathbf{s} + Q_{\Lambda_2}(\mathbf{y} - \mathbf{s})$ at high rate.

Proof: At high rate

$$\begin{aligned}\hat{\mathbf{x}} &= \frac{\int_{R(\mathbf{s})} \mathbf{x} f(\mathbf{x}\,|\,\mathbf{y})\,d\mathbf{x}}{\int_{R(\mathbf{s})} f(\mathbf{x}\,|\,\mathbf{y})\,d\mathbf{x}} \\ &= \frac{\sum_{j=-\infty}^{\infty} \int_{R_j(\mathbf{s})} \mathbf{x} f(\mathbf{x}\,|\,\mathbf{y})\,d\mathbf{x}}{\sum_{j=-\infty}^{\infty} \int_{R_j(\mathbf{s})} f(\mathbf{x}\,|\,\mathbf{y})\,d\mathbf{x}} \\ &\overset{(a)}{=} \frac{\sum_{j=-\infty}^{\infty} c_j(\mathbf{s}) \int_{R_j(\mathbf{s})} f(\mathbf{x}\,|\,\mathbf{y})\,d\mathbf{x}}{\sum_{j=-\infty}^{\infty} \int_{R_j(\mathbf{s})} f(\mathbf{x}\,|\,\mathbf{y})\,d\mathbf{x}} \\ &\overset{(b)}{=} \frac{[\mathbf{s} + Q_{\Lambda_2}(\mathbf{y}-\mathbf{s})] \int_{\mathcal{V}_1(\mathbf{s}+Q_{\Lambda_2}(\mathbf{y}-\mathbf{s}))} f(\mathbf{x}\,|\,\mathbf{y})\,d\mathbf{x}}{\int_{\mathcal{V}_1(\mathbf{s}+Q_{\Lambda_2}(\mathbf{y}-\mathbf{s}))} f(\mathbf{x}\,|\,\mathbf{y})\,d\mathbf{x}} \\ &= \mathbf{s} + Q_{\Lambda_2}(\mathbf{y}-\mathbf{s}) \end{aligned} \tag{60}$$

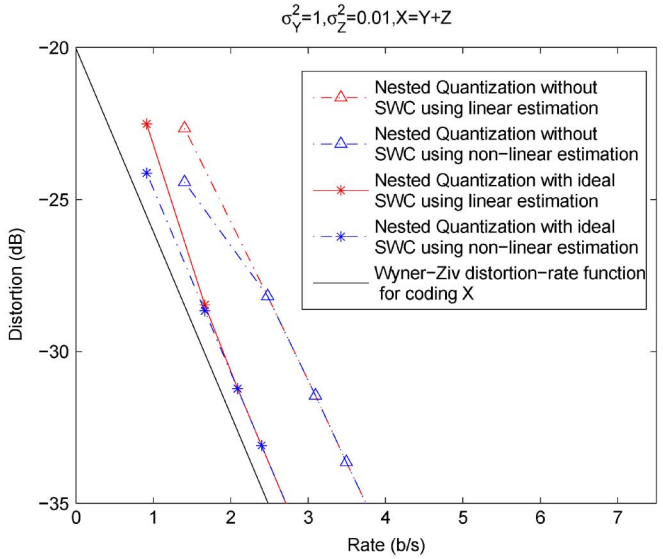


Fig. 9. Improvement gained at low rate by using the nonlinear minimum MSE estimator versus the linear estimator for $n = 2, \sigma_Y^2 = 1$ and $\sigma_Z^2 = 0.01$.

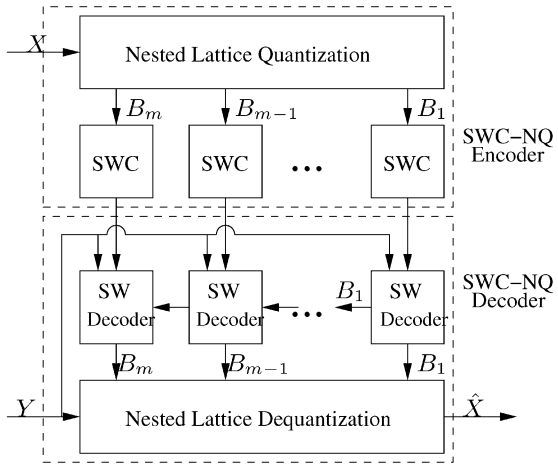


Fig. 10. SWC-NQ with multilevel Slepian–Wolf coding.

where (a) and (b) are due to the high rate assumption and

$$\mathcal{V}_1(\mathbf{s}+Q_{\Lambda_2}(\mathbf{y}-\mathbf{s})) = \{\mathbf{x} : Q_{\Lambda_1}(\mathbf{x}) = \mathbf{s}+Q_{\Lambda_2}(\mathbf{y}-\mathbf{s})\}. \ \square$$

Recall that the linear estimator of (60) is the one we use for high-rate performance analysis in Section IV. But nonlinear estimation is employed for all rate in our simulations. Since (58) involves integration over a disconnected region consisting of many isolated Voronoi cells, we use the Monte Carlo method to compute this integration in our simulations.

B. Practical Slepian–Wolf Code Design

We compress $\mathbf{S}$ using multilevel Slepian–Wolf coding with $\mathbf{Y}$ as the decoder side information. Denote $J(0 \le J \le N-1)$ as the index of $\mathbf{S}$ and write J as $B_m B_{m-1} \cdots B_2 B_1$ in its binary representation, where B_m is the most significant bit of J, and B_1 the least significant bit. At first B_1 is compressed using the first Slepian–Wolf code to rate $R_1 = H(B_1 \,|\, \mathbf{Y})$, then B_2 is compressed with the second Slepian–Wolf code to rate $R_2 = H(B_2 \,|\, \mathbf{Y}, B_1)$, and so on. Finally, B_m is compressed with the mth Slepian–Wolf code to rate $R_m = H(B_m \,|\, \mathbf{Y}, B_1, \ldots, B_{m-1})$. By the chain rule

$$R_1 + R_2 + \cdots + R_m = H(J \,|\, \mathbf{Y}) = H(\mathbf{S} \,|\, \mathbf{Y}).$$

The rate per dimension is

$$R = \frac{1}{n}(R_1 + R_2 + \cdots + R_m) = \frac{1}{n} H(\mathbf{S} \,|\, \mathbf{Y}).$$

By splitting $\mathbf{S}$ into multiple bit planes, well-studied binary channel codes can be used to implement Slepian–Wolf coding of each of them. The idea is to treat $B_i (1 \le i \le m)$ as an input into some "hypothetical" channel with output $(\mathbf{Y}, B_1, B_2, \ldots, B_{i-1})$. At the encoder, the syndrome of the designed channel code for a sequence of realizations of B_i is computed and passed to the decoder. Therefore, the compression rate is one minus the channel code rate and thus the ideal code rate is $1 - R_i = 1 - H(B_i \,|\, \mathbf{Y}, B_1, \ldots, B_{i-1})$. Decoding the ith bit plane B_i is similar to conventional channel decoding except one thing: instead of decoding into a channel codeword, the decoder estimates the sequence of realizations of B_i using the received syndrome in conjunction with $(\mathbf{Y}, B_1, B_2, \ldots, B_{i-1})$. Fig. 10 depicts SWC-NQ with multilevel Slepian–Wolf coding.

In practice, LDPC codes have been used to implement the Slepian–Wolf codes for their near-capacity performance [22]. To design the Slepian-Wolf code for the ith bit plane B_i, it is essential to estimate the desired channel code rate, i.e., $1 - H(B_i \,|\, \mathbf{Y}, B_1, \ldots, B_i - 1)$, via gathering the statistics for the hypothetical channel with input B_i and output $(\mathbf{Y}, B_1, \ldots, B_i - 1)$. In the following, we illustrate this Slepian–Wolf code design process, starting from estimating $\Pr(B_1 = b_1, \ldots, B_i = b_i \,|\, \mathbf{Y} = \mathbf{y})$, for all $\mathbf{y}$ and $1 \le i \le m$.

Because B_i is a deterministic function of $\mathbf{X}$, with a slight abuse of notation, we define $b_i(\cdot)$ as this function. In other words, $B_i = b_i(\mathbf{X})$. Then we have

$$\begin{aligned}\Pr(B_1 = b_1, \ldots, B_i = b_i \,|\, \mathbf{Y} = \mathbf{y}) &= \int_{b_1(\mathbf{x})=b_1, \ldots, b_i(\mathbf{x})=b_i} f(\mathbf{x} \,|\, \mathbf{y})\, d\mathbf{x} \\ &= \int_{b_1(\mathbf{z}+\mathbf{y})=b_1, \ldots, b_i(\mathbf{z}+\mathbf{y})=b_i} f(\mathbf{z})\, d\mathbf{z}. \end{aligned} \tag{61}$$

In the one-dimensional case (with $n = 1$), the integration interval $\{z \,|\, b_1(z+y) = b_1, \ldots, b_i(z+y) = b_i\}$ in (61) corresponds to a union of infinite number of intervals. However, since $f(z)$ decays exponentially from the origin, we can approximate the integral, thus, $\Pr(B_1 = b_1, \ldots, B_i = b_i \,|\, \mathbf{Y} = \mathbf{y})$, accurately as a linear combination of a few Gaussian tail probabilities.

In higher dimensional cases (with $n \ge 2$), however, $\Pr(B_1 = b_1, \ldots, B_i = b_i \,|\, \mathbf{Y} = \mathbf{y})$ cannot be obtained analytically. For example, the integration region $\{\mathbf{z} \,|\, b_1(\mathbf{z}+\mathbf{y}) = b_1, \ldots, b_i(\mathbf{z}+\mathbf{y}) = b_i\}$ in the two-dimensional case corresponds to a union of hexagons and hence no simple analytical solution can be found. Although numerical integration may be used to evaluate (61), we use Monte Carlo simulations because they are more flexible.

Specifically, we first quantize the real axis into K intervals and partition all possible values of $\mathbf{Y}$ into K^n regions $\mathbb{Y}_j$ for $j = 1, \ldots, K^n$. Denote $\mathbb{Y}(\mathbf{y})$ as the region that contains $\mathbf{y}$ and define $1(\cdot)$ as an indicator function that equals to one if its argument is true and zero otherwise. We then draw L independent samples of $\mathbf{Y}' \sim N(\mathbf{0}, \sigma_Y^2 \mathbf{I}_n)$ and $\mathbf{Z}' \sim N(\mathbf{0}, \sigma_Z^2 \mathbf{I}_n)$, where $\mathbf{I}_n$ is the $n \times n$ identity matrix, and let the lth samples be $\bar{\mathbf{y}}_l$ and $\bar{\mathbf{z}}_l$, respectively. Then $\bar{\mathbf{x}}_l = \bar{\mathbf{y}}_l + \bar{\mathbf{z}}_l, 1 \le l \le L$ will be statistically equivalent to samples of $\mathbf{X}$. Finally, by simple number counting, we approximate the probability $\Pr(B_1 = b_1, \ldots, B_i = b_i \,|\, \mathbf{Y} = \mathbf{y})$ in (61) as shown in

$$\begin{aligned}\Pr(B_1 = b_1, \ldots, B_i = b_i \,|\, \mathbf{Y} = \mathbf{y}) &\approx \Pr(B_1 = b_1, \ldots, B_i = b_i \,|\, \mathbf{Y} \in \mathbb{Y}(\mathbf{y})) \\ &\approx \frac{\sum_{l=1}^{L} 1(\bar{\mathbf{y}}_l \in \mathbb{Y}(\mathbf{y}), b_1(\bar{\mathbf{x}}_l) = b_1, \ldots, b_i(\bar{\mathbf{x}}_l) = b_i)}{\sum_{l=1}^{L} 1(\bar{\mathbf{y}}_l \in \mathbb{Y}(\mathbf{y}))}.\end{aligned} \tag{62}$$

1) Desired Channel Code Rate Computation: Armed with (62), we estimate the LDPC code rate for the ith bit plane B_i as shown in the first expression at the bottom of the page, where

$$\Pr(B_1 = b_1, \ldots, B_i = 1 \,|\, \mathbf{Y} \in \mathbb{Y}_j)$$

and

$$\Pr(B_1 = b_1, \ldots, B_{i-1} = b_{i-1} \,|\, \mathbf{Y} \in \mathbb{Y}_j)$$

are obtained directly from (62)

$$\mathcal{H}(p) = p \log_2 \frac{1}{p} + (1-p) \log_2 \frac{1}{1-p}$$

and

$$\Pr(\mathbf{Y} \in \mathbb{Y}_j, B_1 = b_1, \ldots, B_{i-1} = b_{i-1})$$

estimated by using similar Monte Carlo simulations as in (62) with

$$\begin{aligned}&\Pr(\mathbf{Y} \in \mathbb{Y}_j, B_1 = b_1, \ldots, B_{i-1} = b_{i-1}) \\ &= \frac{1}{L} \sum_{l=1}^{L} 1(\bar{\mathbf{y}}_l \in \mathbb{Y}_j, b_1(\bar{\mathbf{x}}_l) = b_1, \ldots, b_{i-1}(\bar{\mathbf{x}}_l) = b_{i-1}).\end{aligned} \tag{63}$$

2) Channel Estimation: Recall that at the ith bit plane, Slepian–Wolf decoding can be viewed as channel decoding over a hypothetical channel with input B_i and output $(\mathbf{Y}, B_1, B_2, \ldots, B_{i-1})$. Assuming that $\mathbf{Y} = \mathbf{y}, B_1 = b_1, \ldots, B_{i-1} = b_{i-1}$, the channel statistics, which is needed for decoding B_i, is entirely captured by the log-likelihood ratio (LLR) shown as the second array of equations at the bottom of the page, where $\mathrm{L}_{\mathrm{ap}} = \log \frac{P_r(B_i=1)}{P_r(B_i=0)}$ is the *a priori* LLR, which can be estimated as

$$\log \frac{\sum_{l=1}^{L} 1(B_i(\bar{\mathbf{x}}_l) = 1)}{\sum_{l=1}^{L} 1(B_i(\bar{\mathbf{x}}_l) = 0)}.$$

In practice, L_{ch} is used to design the LDPC code degree profiles with the help of density evolution [4], [30]. It also served as an initial estimate during decoding. L_{ch} is updated after each decoding iteration. After the final iteration, it is added to the *a priori* LLR L_{ap} for the estimation of B_i. Finally, the estimate $\hat{b}_i$ will be one if the sum is positive and zero otherwise.

While we code the bit planes from the least significant to the most significant levels (i.e., in a bottom-up fashion), the reverse (i.e., top-down) order is also feasible. Theoretically, there is no performance difference between these two Slepian–Wolf coding schemes since the chain rule will result in the same joint conditional entropy $H(\mathbf{S} \,|\, \mathbf{Y})$. Coding from top down is needed for successive refinement in practical Wyner–Ziv coding [4]. However, coding from bottom up in practice can give slightly better performance at high rate. For example, as shown in

$$\begin{aligned}&1 - H(B_i \,|\, \mathbf{Y}, B_1, \ldots B_{i-1}) \\ &\quad = 1 - \sum_{j=1}^{K^n} \sum_{\substack{b_1 \in \{0,1\} \\ \cdots \\ b_{i-1} \in \{0,1\}}} \Pr(\mathbf{Y} \in \mathbb{Y}_j, B_1 = b_1, \ldots, B_{i-1} = b_{i-1}) \mathcal{H}(\Pr(B_i = 1 \,|\, \mathbf{Y} \in \mathbb{Y}_j, B_1 = b_1, \ldots, B_{i-1} = b_{i-1})) \\ &\quad = 1 - \sum_{j=1}^{K^n} \sum_{\substack{b_1 \in \{0,1\} \\ \cdots \\ b_{i-1} \in \{0,1\}}} \Pr(\mathbf{Y} \in \mathbb{Y}_j, B_1 = b_1, \ldots, B_{i-1} = b_{i-1}) \mathcal{H}\left(\frac{\Pr(B_1 = b_1, \ldots, B_i = 1 \,|\, \mathbf{Y} \in \mathbb{Y}_j)}{\Pr(B_1 = b_1, \ldots, B_{i-1} = b_{i-1} \,|\, \mathbf{Y} \in \mathbb{Y}_j)}\right)\end{aligned}$$

$$\begin{aligned}\mathrm{L}_{\mathrm{ch}} &= \log \frac{P_r(\mathbf{Y} = \mathbf{y}, B_1 = b_1, \ldots, B_{i-1} = b_{i-1} \,|\, B_i = 1)}{P_r(\mathbf{Y} = \mathbf{y}, B_1 = b_1, \ldots, B_{i-1} = b_{i-1} \,|\, B_i = 0)} \\ &= \log \frac{P_r(\mathbf{Y} = \mathbf{y}, B_1 = b_1, \ldots, B_{i-1} = b_{i-1}, B_i = 1)}{P_r(\mathbf{Y} = \mathbf{y}, B_1 = b_1, \ldots, B_{i-1} = b_{i-1}, B_i = 0)} - \log \frac{P_r(B_i = 1)}{P_r(B_i = 0)} \\ &= \log \frac{P_r(B_1 = b_1, \ldots, B_{i-1} = b_{i-1}, B_i = 1 \,|\, \mathbf{Y} = \mathbf{y})}{P_r(B_1 = b_1, \ldots, B_{i-1} = b_{i-1}, B_i = 0 \,|\, \mathbf{Y} = \mathbf{y})} - \log \frac{P_r(B_i = 1)}{P_r(B_i = 0)} \\ &\triangleq \log \frac{P_r(B_1 = b_1, \ldots, B_{i-1} = b_{i-1}, B_i = 1 \,|\, \mathbf{Y} = \mathbf{y})}{P_r(B_1 = b_1, \ldots, B_{i-1} = b_{i-1}, B_i = 0 \,|\, \mathbf{Y} = \mathbf{y})} - \mathrm{L}_{\mathrm{ap}}\end{aligned}$$

TABLE I
THE CONDITIONAL ENTROPY, LDPC CODE RATE, AND THE CORRESPONDING DEGREE DISTRIBUTION POLYNOMIALS $\lambda(x)$ AND $\rho(x)$ FOR EACH BIT PLANE OF ONE-DIMENSIONAL SLEPIAN–WOLF CODED NESTED LATTICE QUANTIZATION WITH NESTING RATIO $N = 64$. $H(S \mid Y) = 4.9618231$ B/S; $R = 5.009$ B/S

Bit plane	Conditional entropy	LDPC code rate	Degree polynomials	
			$\lambda(x)$	$\rho(x)$
6	0.1316364	0.850	$0.114444x + 0.179976x^2 + 0.147790x^6 + 0.080683x^7 + 0.012620x^{16} + 0.020889x^{18} + 0.132564x^{20} + 0.033668x^{21} + 0.001529x^{37} + 0.060635x^{41} + 0.019736x^{68} + 0.033480x^{71} + 0.123705x^{78} + 0.038282x^{86}$	x^{40}
5	0.8304147	0.141	$0.345952x + 0.183430x^2 + 0.000084x^3 + 0.009634x^4 + 0.185891x^5 + 0.000104x^6 + 0.115577x^{14} + 0.021200x^{15} + 0.010848x^{16} + 0.108563x^{46} + 0.015865x^{48} + 0.002726x^{60} + 0.000124x^{99}$	$0.8x^3 + 0.2x^4$
4	0.9997720	0	N/A	N/A
1-3	1.0000000	0	N/A	N/A

TABLE II
THE CONDITIONAL ENTROPY, LDPC CODE RATE, AND THE CORRESPONDING DEGREE DISTRIBUTION POLYNOMIALS $\lambda(x)$ AND $\rho(x)$ FOR EACH BIT PLANE OF TWO-DIMENSIONAL SLEPIAN–WOLF CODED NESTED LATTICE QUANTIZATION WITH NESTING RATIO (A) $N = 7$; $\frac{1}{2}H(\mathbf{S} \mid \mathbf{Y}) = 0.91278$ B/S; $R = 0.98$ B/S AND (B) $N = 31$; $\frac{1}{2}H(\mathbf{S} \mid \mathbf{Y}) = 1.66376325$ B/S; $R = 1.749$ B/S

Bit plane	Conditional entropy	LDPC code rate	Degree polynomials	
			$\lambda(x)$	$\rho(x)$
3	0.4307767	0.494	$0.154196x + 0.607571x^2 + 0.191598x^8 + 0.046635x^9$	$0.5x^5 + 0.5x^6$
2	0.6187541	0.342	$0.188401x + 0.432910x^2 + 0.107433x^{11} + 0.271255x^{12}$	$0.3x^4 + 0.7x^5$
1	0.7760298	0.204	$0.142139x + 0.250041x^2 + 0.027853x^7 + 0.116382x^8 + 0.463585x^{39}$	$0.1x^5 + 0.9x^6$

(a)

Bit plane	Conditional entropy	LDPC code rate	Degree polynomials	
			$\lambda(x)$	$\rho(x)$
5	0.6043050	0.349	$0.193289x + 0.450023x^2 + 0.356688x^{11}$	$0.4x^4 + 0.6x^5$
4	0.4695057	0.502	$0.147216x + 0.603324x^2 + 0.122048x^8 + 0.127413x^9$	$0.3x^5 + 0.7x^6$
3	0.7903277	0.178	$0.149076x + 0.247592x^2 + 0.039490x^7 + 0.108175x^8 + 0.455668x^{39}$	$0.4x^5 + 0.6x^6$
2	0.7097249	0.255	$0.229576x + 0.416782x^2 + 0.142361x^{11} + 0.211281x^{12}$	$0.2x^3 + 0.8x^4$
1	0.7536632	0.218	$0.255287x + 0.418496x^2 + 0.300630x^{11} + 0.025587x^{12}$	$0.6x^3 + 0.4x^4$

(b)

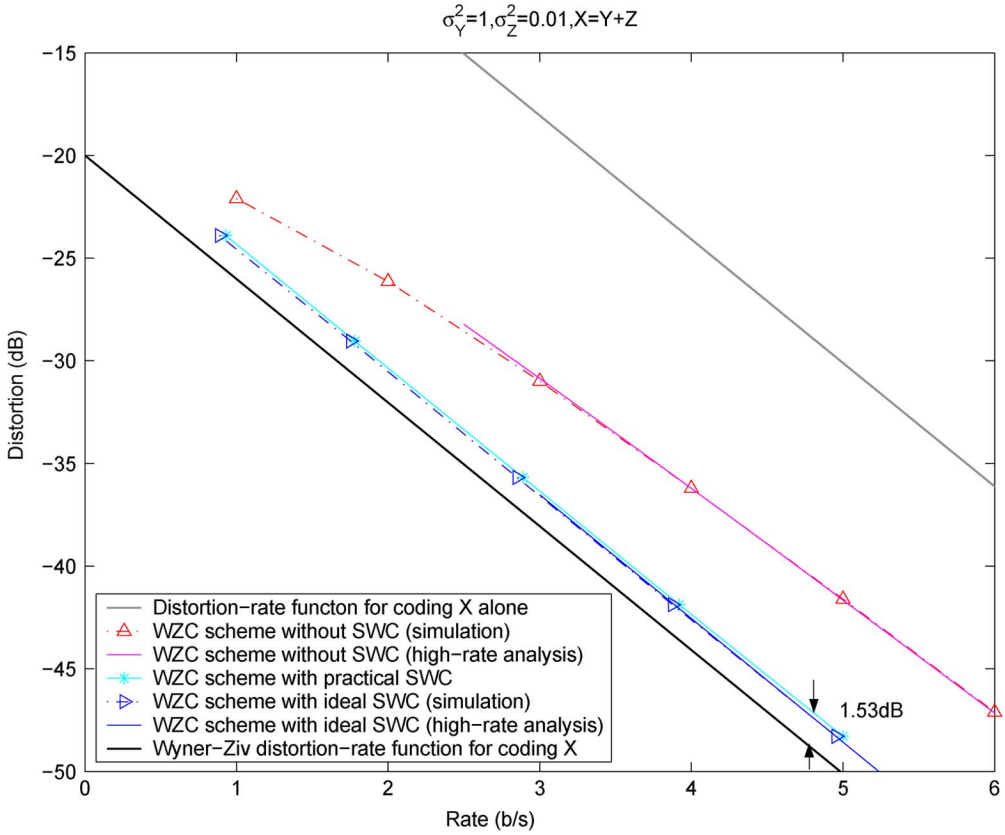
Table I, the conditional entropies of several lower bit planes in the one-dimensional case are almost equal to one and hence the corresponding bit planes are not compressed in practice. Therefore, Slepian–Wolf coding is only employed to compress the two most significant bit planes. Since Slepian–Wolf coding is near- lossless, the fewer the bit planes we code, the smaller the extra distortion introduced in Slepian–Wolf decoding. In the case of coding the bit planes from top down, more bit planes have their conditional entropies quite smaller than one; Slepian–Wolf coding of these bit planes result in higher extra distortion (see [4] for details).

C. Simulation Results

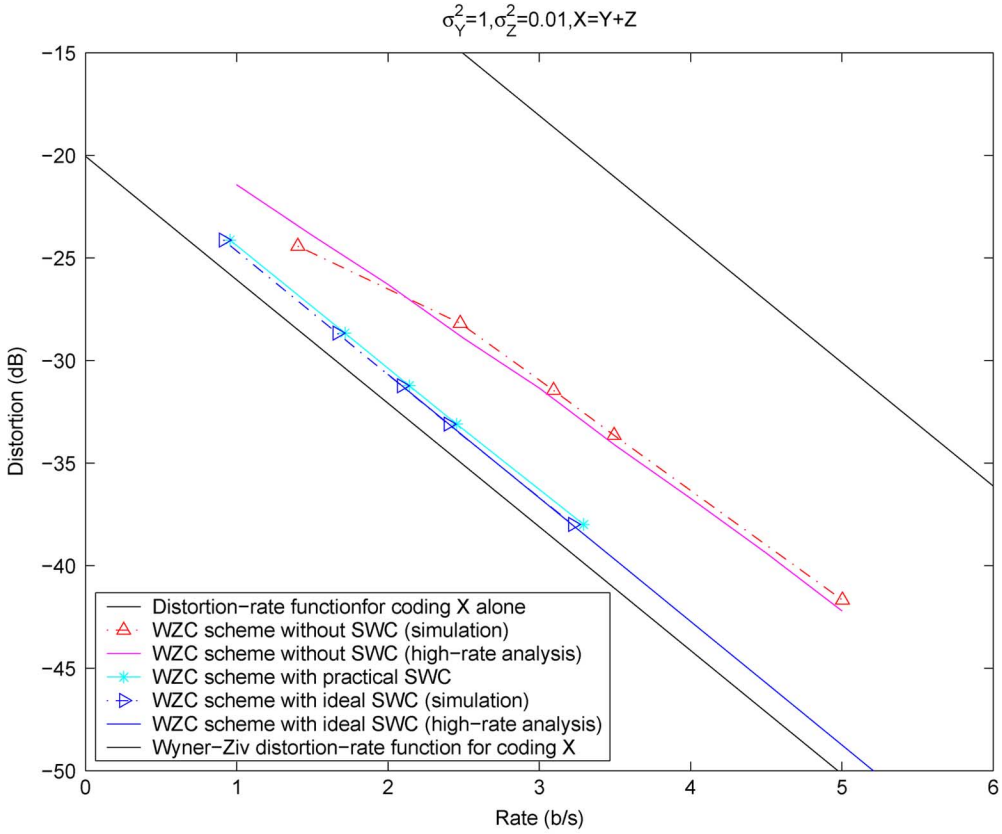
Assuming the source X and the decoder side information Y are related by $X = Y + Z$, with $Y \sim N(0, 1.0)$ and $Z \sim N(0, 0.01)$, extensive simulations have been carried out to evaluate nested lattice quantization in the one- and two-dimensional cases and our proposed SWC-NQ scheme for Wyner–Ziv coding of X. The bit-error rate of our practical Slepian–Wolf decoder is less than 10^{-6} and the errors are accounted for in our reported MSE distortion.

In the one-dimensional case, Fig. 11(a) shows results with nested lattice quantization alone and with SWC-NQ. Nested lattice/scalar quantization alone exhibits a 3.95–9.60-dB performance gap from $D_{\mathrm{WZ}}(R)$ for R =1.0–6.0 b/s, which agrees with the exact high-rate performance given in (29) in Corollary 4.2. With SWC-NQ, we observe that the gap between our simulation results with ideal Slepian–Wolf coding (with rate computed as $H(\mathbf{S} \mid \mathbf{Y})$) and $D_{\mathrm{WZ}}(R)$ is indeed 1.53 dB at high rate; using practical Slepian–Wolf codes based on irregular LDPC codes of length 10^6 bits (with profiles from Table I), this gap is 1.66–1.80 dB for $R = 0.93$–5.0 b/s.

For two-dimensional nested lattice quantization, we use the A_2 hexagonal lattices. Table II lists the conditional entropy and practical LDPC code rate for each bit plane when the nesting ratio $N = 7$ and 31, along with the profiles of the practical LDPC codes. Fig. 11(b) gives results with nested lattice quantization alone and with SWC-NQ. With two-dimensional



(a)



(b)

Fig. 11. Results based on (a) one-dimensional nested lattice quantization with and without SWC and (b) two-dimensional nested A_2 lattice quantization with and without SWC.

nested lattice quantization alone, our simulated performance is 4.06–8.48 dB worse than $D_{\mathrm{WZ}}(R)$ for $R = 1.40$–5.0 b/s. At high rate, our simulated points are 0.6 dB above the lower bound given in (26) in Corollary 4.2. Since the sphere-based lower bound becomes tighter and tighter as n increases (and exact as n goes to infinity), we conclude that 0.6 dB (observed at $n = 2$) is the maximum gap between the lower bound in Corollary 4.2 and $D_{\mathrm{WZ}}(R)$ at high rate.

With SWC-NQ, we also see from Fig. 11(b) that the gap between our results with ideal Slepian–Wolf coding (assuming $R = \frac{1}{2}H(\mathbf{S}\,|\,\mathbf{Y})$ bits per second) and $D_{\mathrm{WZ}}(R)$ is 1.36 dB at high rate. With practical Slepian–Wolf coding based on irregular LDPC codes (of length 10^6 bits), this gap is 1.67–1.84 dB for $R = 0.95$–3.29 b/s.

VII. Conclusion

In this paper, we have analyzed the high-rate performance of nested lattice quantization for Wyner–Ziv coding, showing an increasing gap from the theoretical limit as the rate increases. The reason for the increase of the gap is mainly because the boundary loss is an increasing function of the rate. To compensate for the boundary loss, an SWC-NQ framework has been proposed for Wyner–Ziv coding, where Slepian–Wolf coding plays the role of second-stage binning to save rate after nested lattice quantization. Assuming ideal Slepian–Wolf coding, SWC-NQ is shown to perform a constant gap (in decibels) away from the Wyner–Ziv DR function at high rate. This result mirrors that from entropy-coded quantization in classic source coding. A nonlinear minimum MSE estimator at the decoder is introduced and used in simulations that degenerates to the linear estimator we use in our high-rate performance analysis. Simulations with one- and two-dimensional nested lattice quantization and SWC-NQ (with ideal Slepian-Wolf coding) for quadratic Gaussian sources show agreement with our high-rate analytical results. Using irregular LDPC codes for practical Slepian–Wolf coding in SWC-NQ exhibits a roughly constant gap from the Wyner–Ziv limit for a wide range of rates.

We have also proved that the performance loss of Wyner–Ziv coding of quadratic Gaussian sources with nested lattice quantization at a fixed high rate is independent of the source correlation, with or without Slepian–Wolf coding.

Appendix A

Proof of Lemma 5.1 for the lower bound on the DR performance SWC-NQ in the quadratic Gaussian case.

1) Rate Computation: The rate for SWC-NQ is

$$R = \frac{1}{n}H(\mathbf{S}\,|\,\mathbf{Y}). \tag{64}$$

Since at high rate

$$\begin{aligned} p(\mathbf{s}\,|\,\mathbf{y}) &= \sum_{j=-\infty}^{\infty}\int_{\mathbf{x}\in R_j(\mathbf{s})} f_{\mathbf{X}\,|\,\mathbf{Y}}(\mathbf{x})\,d\mathbf{x} \\ &= \sum_{j=-\infty}^{\infty}\int_{\mathbf{x}\in R_0(\mathbf{s})} f_{\mathbf{Z}}(\mathbf{x}+\boldsymbol{c}_j(\mathbf{0}))\,d\mathbf{x} \\ &\overset{(a)}{=} \sum_{j=-\infty}^{\infty} f_{\mathbf{Z}}(\mathbf{s}+\boldsymbol{c}_j(\mathbf{0}))V_1 \\ &= g(\mathbf{s})V_1 \end{aligned} \tag{65}$$

where (a) is due to the high rate assumption and

$$g(\mathbf{x}) \triangleq \sum_{j=-\infty}^{\infty} f_{\mathbf{Z}}(\mathbf{x}+\boldsymbol{c}_j(\mathbf{0})).$$

Then the achievable rate of SWC-NQ is

$$\begin{aligned} nR = H(\mathbf{S}\,|\,\mathbf{Y}) &= -\sum_{\mathbf{s}\in\Lambda_1/\Lambda_2} p(\mathbf{s}\,|\,\mathbf{y})\log_2[p(\mathbf{s}\,|\,\mathbf{y})] \\ &\overset{(b)}{=} -\sum_{\mathbf{s}\in\Lambda_1/\Lambda_2} p(\mathbf{s}\,|\,\mathbf{y})\log_2[g(\mathbf{s})V_1] \\ &= -\sum_{\mathbf{s}\in\Lambda_1/\Lambda_2}\sum_{j=-\infty}^{\infty}\int_{\mathbf{x}\in R_0(\mathbf{s})} f_{\mathbf{Z}}(\mathbf{x}+\boldsymbol{c}_j(\mathbf{0}))\,d\mathbf{x}\log_2 g(\mathbf{s}) \\ &\quad -\log_2 V_1 \\ &\overset{(c)}{=} -\sum_{j=-\infty}^{\infty}\sum_{\mathbf{s}\in\Lambda_1/\Lambda_2}\int_{\mathbf{x}\in R_0(\mathbf{s})} f_{\mathbf{Z}}(\mathbf{x}+\boldsymbol{c}_j(\mathbf{0}))\log_2 g(\mathbf{x})\,d\mathbf{x} \\ &\quad -\log_2 V_1 \\ &\overset{(d)}{=} -\int_{\mathbf{x}\in\mathbb{R}^n} f_{\mathbf{Z}}(\mathbf{x})\log_2 g(\mathbf{x})\,d\mathbf{x} - \log_2 V_1 \\ &= -\int_{\mathbf{x}\in\mathbb{R}^n} f_{\mathbf{Z}}(\mathbf{x})\log_2\left[\sum_{j=-\infty}^{\infty} f_{\mathbf{Z}}(\mathbf{x}+\boldsymbol{c}_j(\mathbf{0}))\right]d\mathbf{x} \\ &\quad -\log_2 V_1 \end{aligned}$$

where (b) and (c) use the high rate assumption and (d) is due to the periodic property of $g(\cdot)$, i.e., $g(\mathbf{x}-\boldsymbol{l}) = g(\mathbf{x}), \forall \boldsymbol{l}\in\Lambda_2$. Thus. the achievable rate of SWC-NQ is

$$nR = H(\mathbf{S}\,|\,\mathbf{Y}) = h'(\mathbf{X},\Lambda_2) + \log_2\sigma_Z^n - \log_2 V_1. \tag{66}$$

2) Distortion Computation: From Theorem 4.1, the average distortion of nested lattice quantization over all realizations of $(\mathbf{X},\mathbf{Y})$ is $D_n = G(\Lambda_1)V_1^{\frac{2}{n}} + \frac{1}{n}E_{\mathbf{Z}}[\|Q_{\Lambda_2}(\mathbf{Z})\|^2]$, which can be lower-bounded as

$$D_n \geq G(\Lambda_1)V_1^{\frac{2}{n}} + \bar{D}_C \tag{67}$$

according to (26) and (32)

Because Slepian–Wolf coding is near-lossless, the distortion of SWC-NQ is also D_n. Combining D_n and R through V_1 in (66) and (67), we obtain the DR performance of SWC-NQ with a pair of n-dimensional nested lattices (Λ_1,Λ_2) as

$$D_n(R) \geq G(\Lambda_1)2^{\frac{2}{n}h'(\mathbf{X},\Lambda_2)}\sigma_Z^2 2^{-2R} + \bar{D}_C. \tag{68}$$

Appendix B

Proof of Lemma 5.2: This proof closely follows [37, p. 3, remark 3)] with slight modifications. Let $\delta \triangleq \min_{\mathbf{w}\neq\hat{\mathbf{x}}} d(\mathbf{w},\hat{\mathbf{x}}) > 0$.

Here δ is actually the minimum distance between points in Λ_2. Thus, if $(\mathbf{X}, \hat{\mathbf{X}}) \in \mathcal{S}_1$

$$\begin{aligned}\lambda &\triangleq P_r\{\mathbf{W} \neq \hat{\mathbf{X}}\} \leq \frac{1}{\delta} E[d(\mathbf{W}, \hat{\mathbf{X}})] \\ &\overset{(a)}{\leq} \frac{1}{\delta}(E[d(\mathbf{W}, \mathbf{X})] + E[d(\mathbf{X}, \hat{\mathbf{X}})]) \end{aligned} \tag{69}$$

where (a) is due to the triangle inequality. From Theorem 4.1, $D = \frac{1}{n}E[d(\mathbf{X}, \hat{\mathbf{X}})] = D_S + D_C$, where $D_S = \frac{1}{n}E[d(\mathbf{W}, \mathbf{X})]$ is the source coding loss and D_C is the channel coding loss, then

$$\lambda \leq \frac{2nD}{\delta}. \tag{70}$$

Now since $\hat{\mathbf{X}}$ is a function of $\mathbf{S}$ and $\mathbf{Y}$, Fano's inequality [9], [15] implies that

$$\begin{aligned} H(\mathbf{W} \,|\, \mathbf{S}, \mathbf{Y}) &\leq -\lambda \log \lambda - (1-\lambda)\log(1-\lambda) + \lambda \log(|\mathcal{W}|) \\ &\triangleq \varepsilon(\lambda) \end{aligned} \tag{71}$$

then

$$\begin{aligned} H(\mathbf{S} \,|\, \mathbf{Y}) &\geq I(\mathbf{W}; \mathbf{S} \,|\, \mathbf{Y}) \\ &= H(\mathbf{W} \,|\, \mathbf{Y}) - H(\mathbf{W} \,|\, \mathbf{S}, \mathbf{Y}) \\ &\geq H(\mathbf{W} \,|\, \mathbf{Y}) - \varepsilon\left(\frac{2nD}{\delta}\right). \end{aligned} \tag{72}$$

Meanwhile

$$\begin{aligned} H(\mathbf{W} \,|\, \mathbf{Y}) &\geq I(\mathbf{W}; \mathbf{S} \,|\, \mathbf{Y}) \\ &= H(\mathbf{S} \,|\, \mathbf{Y}) - H(\mathbf{S} \,|\, \mathbf{W}, \mathbf{Y}) \\ &= H(\mathbf{S} \,|\, \mathbf{Y}). \end{aligned} \tag{73}$$

At high rate, $D \to 0$ and $\varepsilon(\frac{2nD}{\delta}) \to 0$, thus $H(\mathbf{S} \,|\, \mathbf{Y}) = H(\mathbf{W} \,|\, \mathbf{Y})$.

ACKNOWLEDGMENT

The authors gratefully thank the anonymous reviewers and Prof. Tie Liu for their constructive comments.

REFERENCES

[1] A. Aaron and B. Girod, "Compression with side information using turbo codes," in *Proc. Data Compression Conf. (DCC'02)*, Snowbird, UT, Mar. 2002, pp. 252–261.

[2] R. Barron, B. Chen, and G. Wornell, "The duality between information embedding and source coding with side information and some applications," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1159–1180, May 2003.

[3] J. Bronski1 and R. McLaughlin, "Rigorous estimates of the tails of the probability distribution function for the random linear shear model," *J. Statist. Phys.*, vol. 98, pp. 897–915, Feb. 2000.

[4] 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.

[5] J. Chou, S. Pradhan, and K. Ramchandran, "Turbo and trellis-based constructions for source coding with side information," in *Proc. Data Compression Conf. (DCC'03)*, Snowbird, UT, Mar. 2003, pp. 33–42.

[6] J. Conway, E. Rains, and N. Sloane, "On the existence of similar sublattices," *Canad. J. Math.*, vol. 51, pp. 1300–1306, 1999.

[7] J. Conway and J. Sloane, *Sphere Packings. Lattices and Groups*. New York: Springer-Verlag, 1998.

[8] M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 3, pp. 439–441, May 1983.

[9] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[10] L. Dalton, Analysis of 1-D Nested Lattice Quantization and Slepian-Wolf Coding For Wyner-Ziv Coding of i.i.d. Sources Texas A&M Univ., College Stataion, Tech. Rep., May 2002.

[11] M. Eyuboglu and D. Forney, Jr., "Lattice and trellis quantization with lattice-and trellis-bounded codebooks—high-rate theory for memoryless sources," *IEEE Trans. Inf. Theory*, vol. 39, no. 1, pp. 46–59, Jan. 1993.

[12] H. Feng, Q. Zhao, and M. Effros, "Network source coding using entropy constrained dithered quantization," in *Proc. data Comression Conf. (DCC'03)*, Snowbird, UT, Mar. 2003, pp. 427–436.

[13] D. Forney Jr., "Coset codes-Part II: Binary lattices and related codes," *IEEE Trans. Inf. Theory*, vol. 34, no. 5, pp. 1152–1187, Sep. 1988.

[14] R. Gallager, *Low Density Parity Check Codes*. Cambridge, MA: MIT Press, 1963.

[15] ——, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[16] J. Garcia-Frias and Y. Zhao, "Compression of correlated binary sources using turbo codes," *IEEE Commun. Lett.*, vol. 5, no. 10, pp. 417–419, Oct. 2001.

[17] S. Gel'fand and M. Pinsker, "Coding for channel with random parameters," *Probl. Contr. and Inf. Theory*, vol. 9, pp. 19–31, 1980.

[18] R. Gray and D. Neuhoff, "Quantization," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2325–2383, Oct. 1998.

[19] Z. Liu, S. Cheng, A. Liveris, and Z. Xiong, "Slepian-Wolf coded nested quantization (SWC-NQ) for Wyner-Ziv coding: Performance analysis and code design," in *Proc. Data Compression Conf. (DCC'04)*, Snowbird, UT, Mar. 2004, pp. 322–331.

[20] A. Liveris, Z. Xiong, and C. Georghiades, "Compression of binary sources with side information at the decoder using LDPC codes," *IEEE Commun. Lett.*, vol. 6, no. 10, pp. 440–442, Oct. 2002.

[21] A. Liveris, Z. Xiong, and C. Georghiades, "Distributed compression of binary sources using conventional parallel and serial concatenated convolutional codes," in *Proc. Data Compression Conf. (DCC'03)*, Snowbird, UT, Mar. 2003, pp. 193–202.

[22] A. Liveris, Z. Xiong, and C. Georghiades, "Nested convolutional/turbo codes for the binary Wyner-Ziv problem," in *Proc. Int. Conf. Image Processiong (ICIP'03)*, Barcelona, Spain, Sep. 2003, pp. I-601–I-604.

[23] D. MacKay, "Good error-correcting codes based on very sparse matrices," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 399–431, Mar. 1999.

[24] P. Mitran and J. Bajcsy, "Coding for the Wyner-Ziv problem with turbo-like codes," in *Proc. Int. Symp. Information Theory*, Lausanne, Switzerland, Jun./Jul. 2002, p. 91.

[25] J. Pearl, *Probability Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Francisco, CA: Morgan Kaufmann, 1988.

[26] 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.

[27] S. Pradhan, J. Chou, and K. Ramchandran, "Duality between source coding and channel coding and its extension to the side information case," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1181–1203, May 2003.

[28] D. Rebollo-Monedero, R. Zhang, and B. Girod, "Design of optimal quantizers for distributed source coding," in *Proc. Data Compression Conf. (DCC'03)*, Snowbird, UT, Mar. 2003, pp. 13–22.

[29] D. Rebollo-Monedero, A. Aaron, and B. Girod, "Transforms for high-rate distributed source coding," in *Proc. 37th Asilomar Conf.* , Pacific Grove, CA, Nov. 2003, pp. 850–854.

[30] T. Richardson, M. Shokrollahi, and R. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 619–637, Feb. 2001.

[31] S. Servetto, "Lattice quantization with side information," in *Proc. Data Compression Conf. (DCC'00)*, Snowbird, UT, Mar. 2000, pp. 510–519.

[32] S. Servetto, "On the feasibility of large-scale wireless sensor networks," in *Proc. 40th Allerton Conf. Communication, Control, and Computing*, Monticello, IL, Oct. 2002.

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

[34] V. Tarokh, A. Vardy, and K. Zeger, "Universal bound on the performance of lattice codes," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 670–681, Mar. 1999.

[35] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 1, pp. 55–67, Jan. 1982.

[36] X. Wang and M. Orchard, "Design of trellis codes for source coding with side information at the decoder," in *Proc. Data Compression Conference (DCC'01)*, Snowbird, UT, Mar. 2001, pp. 361–370.

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

[38] A. Wyner, "The rate-distortion function for source coding with side information at the decoder-II: General sources," *Inf. Contr.*, vol. 38, pp. 60–80, 1978.

[39] ——, "Recent results in the Shannon theory," *IEEE Trans. Inf. Theory*, vol. IT-20, pp. 2–10, Jan. 1974.

[40] Y. Yang, S. Cheng, Z. Xiong, and W. Zhao, "Wyner-Ziv coding based on TCQ and LDPC codes," in *Proc. 37th Asilomar Conf.*, Pacific Grove, CA, Nov. 2003, pp. 825–829.

[41] R. Zamir, "The rate loss in the Wyner-Ziv problem," *IEEE Trans. Inf. Theory*, vol. 42, no. 6, pp. 2073–2084, Nov. 1996.

[42] R. Zamir, S. Shamai (Shitz), and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1250–1276, Jun. 2002.