RANDOM PROJECTIONS IN INDIRECT QUANTIZATION PROBLEMS

Yuefeng Cao, Wenyi Zhang

University of Science and Technology of China
Department of Electronic Engineering and Information Science
Hefei, China

ABSTRACT

In the indirect quantization tasks, the quantize-then-estimate (QE) scheme is adopted when the encoders are task-independent or distributed. In some cases, however, the dimension of the observation substantially exceeds that of the state, making the direct quantization of observations highly inefficient. In this work, we propose a random-projection-then-quantization (RPQ) scheme. By randomly projecting high-dimensional observations into a lower-dimensional representation, quantization is applied only to the reduced dimensions rather than all original dimensions. This approach reduces the required bit rate, thereby alleviating communication overhead. The rate-distortion properties of this scheme is analyzed. In the low-rate regime, this scheme outperforms the conventional QE scheme.

Index Terms— Composite source, indirect quantization problem, random projection, universal quantization, rate-distortion theory

1. INTRODUCTION

The problem of semantic communication has long attracted attention [1, 2, 3, 4]. Recent studies incorporate semantic information into the design of communication systems [5, 6, 7, 8]. These works are typically task-oriented rather than aimed at directly compressing or transmitting signals, with objectives closely tied to semantic content. A common model for studying semantic communication is the composite source model [9, 10, 11, 12], which consists of an intrinsic state and an extrinsic observation. The intrinsic state represents the semantic information, while the encoder has access only to the observation.

This problem falls into the category of indirect source coding [13, 14, 15, 16]. In [13, 14], the optimality of an estimate-then-compress (EC) scheme was established. For quantization, the corresponding approach is the estimate-then-quantize (EQ) scheme [17], where the state is first estimated from the observations and then quantized.

Nevertheless, in many practical applications, encoders are universal and task-independent; that is, they are unaware of the semantic relation between the observation and the state, which makes the EQ scheme infeasible. A natural alternative is the quantize-then-estimate (QE) scheme. In [18, 19], the rate–distortion risk of the compress-then-estimate (CE) scheme was analyzed, and a method for designing estimators independent of specific compression techniques was proposed.

In some cases, while the exact relationship between the state and the observation is unknown, it is known that the observation has a much higher dimension than the state. In such highly overdetermined systems, directly quantizing the observation wastes rate. In this paper, we propose a random-projection–then–quantization (RPQ) scheme: the high-dimensional observation is first projected into a lower-dimensional representation and then quantized. Finally, a task-specific estimator reconstructs the state from the quantized low-dimensional representation.

As related works, random projection has been used in the signal processing of practical tasks, such as computer vision [20, 21] and speech recognition [22]. These works empirically demonstrate the potential of random projection, for example, in reducing the amount of transmitted data or improving computational efficiency while maintaining classification accuracy, but they do not investigate it from a rate–distortion perspective. In this work, we model the problem with the composite source model and analyze the RPQ scheme from a rate–distortion perspective.

Our main contributions are as follows:

- We propose the RPQ scheme and characterize its rate–distortion performance for a given projection matrix in the high-rate regime.
- We establish the relationship between the expected estimation distortion and the dimension K of the low-dimensional vector for a common class of source models, and analyze the trade-off between estimation and quantization distortion.
- Through numerical results, we compare the rate–distortion performance of the RPQ scheme with that of the QE and EQ schemes for different values of K. The results show that, under low-rate conditions, random projection can improve rate–distortion performance.

The remainder of the paper is organized as follows:

In Section 2, we describe the source model adopted in this work and formulate the distortion measure and quantizer rate. Section 3 presents the RPQ scheme and analyzes its distortion, including the rate–distortion performance in the high-rate regime for a given dimensionality-reduction matrix Φ. In Section 4, we establish the relationship between the expected estimation distortion and the dimension K of the low-dimensional representation, and analyze the trade-off between estimation and quantization distortions. Section 5 compares the rate–distortion performance of the RPQ scheme under different values of K with that of the EQ and QE schemes via numerical simulations. Finally, Section 6 concludes the paper.

2. PROBLEM FORMULATION

A composite source consists of an intrinsic state S and an extrinsic observation X, where each state value induces a distinct sub-source and distribution of X. In encoding, the encoder observes only X [11, 12], and the state must be inferred from it. In this paper, we focus exclusively on indirect quantization tasks, requiring only the reconstruction of S.

In this paper, we focus on composite sources where $S \in \mathbb{R}^m$ and $X \in \mathbb{R}^n$ are zero-mean and jointly Gaussian, related by $X = HS + W$, with H an $n \times m$ matrix and $W \sim \mathcal{N}(\mathbf{0}, \Sigma_W)$ independent of S. Denote the joint covariance matrix of (S, X) as

$$\begin{bmatrix} \Sigma_S & \Sigma_{SX} \\ \Sigma_{SX}^T & \Sigma_X \end{bmatrix}. \tag{1}$$

We have $\Sigma_{SX} = \Sigma_S H^T$ and $\Sigma_X = H\Sigma_S H^T + \Sigma_W$.

The optimality of the estimate-then-compress (EC) scheme under the mean-squared error (MSE) criterion was proved in [13, 14]. In the EC scheme, the state S is first estimated as $\tilde{S}$ from the observation X, yielding $\tilde{S}$, which is then compressed into $\hat{S}$. The procedure of this scheme is illustrated in Figure 1.

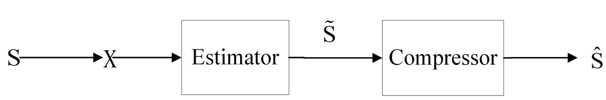


Fig. 1. The procedure of the EC scheme

When the encoder is task-independent, the EC scheme cannot be implemented. In such cases, the compress-then-estimate (CE) scheme can be applied. Here, X is first compressed into $\hat{X}$, and a task-specific estimator then reproduces $\hat{S}$ from $\hat{X}$. The procedure of this scheme is illustrated in Figure 2.

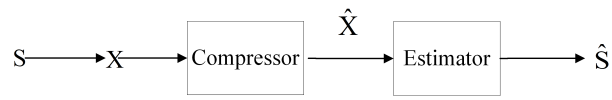


Fig. 2. The procedure of the CE scheme

In this paper, quantization is adopted as the compression method, and the MSE distortion is used as the distortion metric. Denoting the state distortion as d_S, we have:

$$d_S(s, \hat{s}) = \|s - \hat{s}\|_2^2. \tag{2}$$

When designing the quantizer, our objective is to minimize $\mathbb{E}\{d_S(s, \hat{s})\}$. Since S is unobserved by the quantizer, for convenience, this indirect optimization objective can be equivalently expressed in the direct form [11] as:

$$\hat{d}_S(x, \hat{s}) = \mathbb{E}\{d_S(S, \hat{s}) \mid X = x\}. \tag{3}$$

Denote the quantizer output as V. In some schemes, the quantizer produces $\hat{X}$, while in others it directly outputs $\hat{S}$. When using a vector quantizer, the quantizer rate is

$$R = I(X; V) = H(V) - H(V \mid X). \tag{4}$$

Since quantization is a form of deterministic source coding, we have $H(V \mid X) = 0$. Thus, $R = H(V)$. For a scalar quantizer, the rate is $R = \sum_{i=1}^{K} H(V_i)$, where K is the dimension of V.

3. THE RANDOM-PROJECTION-THEN-QUANTIZATION SCHEME

In practical applications, quantizers are often universal and task-independent, and the conventional QE scheme is used. In some cases, however, although the specific source model is unknown, it is known that there is a large dimensionality gap between S and X, making the system highly overdetermined. As a result, directly quantizing X can cause significant performance loss. To save bits, a compressed-sensing-inspired [23] approach can be employed, where a K-dimensional ($K < n$) random projection of X, denoted by Y, is quantized and used to reconstruct S. Since the relationship between S and X is unknown to the quantizer, a random matrix $\Phi \in \mathbb{R}^{K \times n}$ with i.i.d. Gaussian entries is used to project X into a lower-dimensional space. The procedure of the random-projection-then-quantization (RPQ) scheme is illustrated in Figure 3. The Markov chain: $S \leftrightarrow X \leftrightarrow Y \leftrightarrow \hat{Y} \leftrightarrow \hat{S}$ is established.

Evidently, S and Y are also jointly Gaussian. Therefore, the optimal estimator is given by [12]

$$\mathbb{E}\{S \mid Y = y\} = \Sigma_{SY}\Sigma_Y^{-1}y, \tag{5}$$

where Σ_{SY} and Σ_Y denote the cross-covariance matrix between S and Y, and the covariance matrix of Y, respectively.

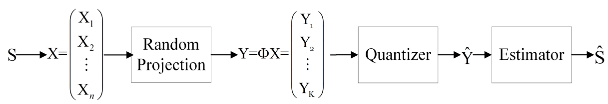


Fig. 3. The procedure of the RPQ scheme

It follows that $\Sigma_Y = \Phi\Sigma_X\Phi^T$ and $\Sigma_{SY} = \Sigma_S H^T\Phi^T$. Let the quantization output of Y be denoted as $\hat{Y}$. The state is then reconstructed as $\hat{S} = \Sigma_{SY}\Sigma_Y^{-1}\hat{Y}$. For convenience, we denote $B = \Sigma_{SY}\Sigma_Y^{-1}$.

Consider the expected distortion between S and $\hat{S}$ for a given Φ.

$$\begin{aligned}\mathbb{E}\{\|S-\hat{S}\|_2^2\} &= \mathbb{E}\{\|S-BY+BY-B\hat{Y}\|_2^2\} \\ &= \mathbb{E}\{\|S-BY\|_2^2\} + 2\mathbb{E}\{\|BY-B\hat{Y}\|_2^2\} \\ &\quad + 2\mathbb{E}\{(S-BY)^T(BY-B\hat{Y})\}.\end{aligned} \tag{6}$$

Since the quantization is independent of the state S, $(S - BY)^T$ and $BY - B\hat{Y}$ are independent. Therefore, the third term in (6) is zero. We have

$$\begin{aligned}\mathbb{E}\{\|S-\hat{S}\|_2^2\} &= \mathbb{E}\{\|S-BY+BY-B\hat{Y}\|_2^2\} \\ &= \mathbb{E}\{\|S-BY\|_2^2\} + 2\mathbb{E}\{\|BY-B\hat{Y}\|_2^2\}.\end{aligned} \tag{7}$$

Since S and Y are jointly Gaussian, a linear relationship exists between them, i.e., $S = BY + Z$, where Z is independent of Y and $Z \sim \mathcal{N}(0, \Sigma_S - \Sigma_{SY}\Sigma_Y^{-1}\Sigma_{SY})$ [12]. Thus, we have

$$\begin{aligned}\mathbb{E}\{\|S-BY\|_2^2\} &= \mathbb{E}\{\|Z\|_2^2\} \\ &= \mathrm{tr}[\Sigma_S - \Sigma_{SY}\Sigma_Y^{-1}\Sigma_{SY}].\end{aligned} \tag{8}$$

Denote this estimation distortion as $D_0(\Phi)$. Then, the total expected distortion of the state S can be expressed as

$$\mathbb{E}\{\|S-\hat{S}\|_2^2\} = D_0(\Phi) + \mathbb{E}\{\|BY - B\hat{Y}\|_2^2\}. \tag{9}$$

Theorem 1. *Under high-resolution conditions, using the universal quantization [24], the distortion-rate performance of the RPQ scheme with projection matrix Φ and scalar universal quantization is approximately*

$$D \approx D_0(\Phi) + \frac{\|B\|_F^2}{12}2^{\frac{2}{K}h(Y)}2^{-\frac{2}{K}R}, \tag{10}$$

where $\|\cdot\|_F$ denotes the Frobenius norm and $h(\cdot)$ denotes the differential entropy.

When the dimensions of Y are quantized separately, the distortion-rate performance of the scheme is:

$$D \approx D_0(\Phi) + \frac{\|B\|_F^2}{12}2^{\frac{2}{K}\sum_{i=1}^{K}h(Y)}2^{-\frac{2}{K}R}, \tag{11}$$

Proof. See Appendix A. □

4. CASE STUDY

In this section, we focus on a specific source model and analyze the rate–distortion performance of the RPQ scheme by taking the expectation over the random projection matrix Φ for a given K. Based on this analysis, we further study the impact of K on the rate–distortion performance of the RPQ scheme. Throughout this section, we assume that the entries of Φ are i.i.d and follow $\mathcal{N}(0, \frac{1}{K})$.

Consider a composite source where $S \sim \mathcal{N}(0, 1)$ and $X = \mathbf{1}_n S + W$, with $\mathbf{1}_n$ the length-n all-ones vector. The observation noise $W = \sigma_W^2 I_n$ is independent of S, where I_n represents the $n \times n$ identity matrix.

From (9), we have

$$\begin{aligned}&\mathbb{E}_\Phi\{D \mid K\} \\ &= \mathbb{E}_\Phi\{D_0(\Phi) \mid K\} + \mathbb{E}_\Phi\{\mathbb{E}\{\|BY - B\hat{Y}\|_2^2\} \mid K\},\end{aligned} \tag{12}$$

where $\mathbb{E}_\Phi\{\cdot \mid K\}$ denotes expectation over Φ with projection dimension K.

Equation (12) shows that the total expected distortion consists of two components: the expected estimation distortion $\mathbb{E}\{D_0(\Phi) \mid K\}$ and the expected quantization distortion.

Theorem 2. *In the composite source described above, when X is projected into a K-dimensional vector Y using a Gaussian random matrix Φ, the expected estimation distortion is*

$$\mathbb{E}\{D_0(\Phi) \mid K\} = \mathbb{E}\{\frac{\sigma_W^2}{\sigma_W^2 + nl}\}, l \sim Beta(\frac{K}{2}, \frac{n-K}{2}), \tag{13}$$

where $Beta(\cdot,\cdot)$ denotes the Beta distribution.

Proof. See Appendix B. □

Remark 1. *Theorem 2 shows that as K increases, the pdf of l shifts toward larger values near $l = 1$, making l more likely to be close to 1. Consequently, the expected estimation distortion decreases with increasing K.*

A special case arises when $K = n$:

$$\begin{aligned}\mathbb{E}\{D_0(\Phi) \mid K\} &= \mathrm{tr}[\Sigma_S - \Sigma_{SY}\Sigma_Y^{-1}\Sigma_{SY}^T] \\ &= \mathrm{tr}[\Sigma_S - \Sigma_{SX}\Phi^T(\Phi\Sigma_X\Phi^T)^{-1}\Phi\Sigma_{SX}^T].\end{aligned} \tag{14}$$

Since Φ is invertible with probability 1 [25],

$$\mathbb{E}\{D_0(\Phi) \mid K\} = \mathrm{tr}[\Sigma_S - \Sigma_{SX}\Sigma_X^{-1}\Sigma_{SX}^T]. \tag{15}$$

Hence, when $K = n$, the expected estimation distortion equals that of directly estimating S from X.

Theorem 1 presents the distortion–rate performance of the RPQ scheme under the high-rate condition. In this regime, (10) and (11) are dominated by $D_0(\Phi)$. Intuitively, since Φ is full rank with probability 1, a larger K preserves more information about X, reducing the estimation distortion $D_0(\Phi)$.

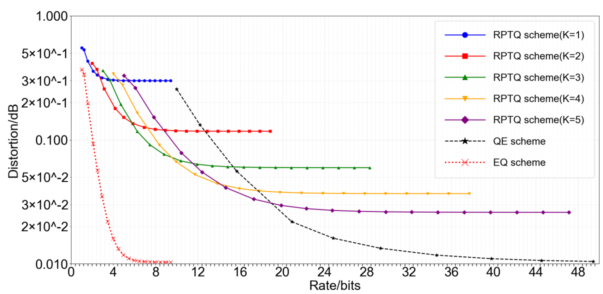
Theorem 2 formally proves this for the composite source considered here. Therefore, in the high-rate regime, dimensionality reduction cannot reduce the total distortion.

Nevertheless, when the high-rate condition is not satisfied, the second term in (6) cannot be neglected. Increasing K reduces the rate allocated to each dimension, enlarging the quantization intervals in scalar quantizers and thereby increasing the quantization distortion. Consequently, varying K introduces a trade-off between estimation and quantization distortions.

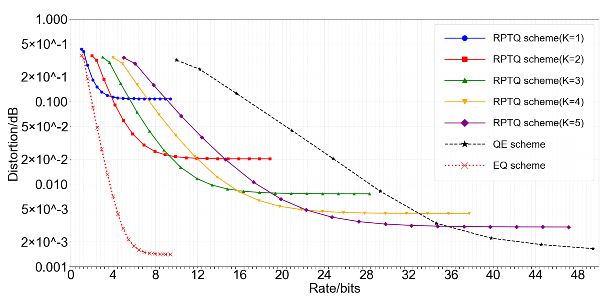
5. NUMERICAL RESULTS

In this section, we present the performance of the RPQ scheme for different values of K and compare it with the QE and EQ schemes. Numerical simulations are conducted on two instances of the source models described in Section 4, with $n = 10$ in both cases and σ_W^2 set to 0.1 and 0.01, respectively.

Scalar uniform quantization is used in all schemes, with K ranging from 1 to 5. For each K, 10,000 instances of Φ are randomly generated, and for each Φ, 50,000 pairs of (S, X) are simulated.



(a)The distortion-rate performances of different schemes when $\sigma_W^2 = 0.1$.



(b)The distortion-rate performances of different schemes when $\sigma_W^2 = 0.01$.

Fig. 4. The distortion-rate performances of different schemes.

Figure 4 compares the performance of the RPQ scheme for different K with the QE and EQ schemes. It can be observed that the EQ scheme outperforms the others, indicating that knowledge of the semantic relationship between the observation and the state improves the quantizer's rate–distortion performance. However, when the quantizer is task-independent, the EQ scheme cannot be implemented. At rates above a certain threshold, the QE scheme outperforms the RPQ scheme, as retaining more information about X reduces estimation distortion, while further increasing each quantizer's precision yields only a limited reduction. In the low-rate regime, however, the contribution of quantization distortion becomes more significant. Thus, by employing random projection, although the estimation distortion increases, the total distortion is substantially reduced due to the higher bit rate allocated to each scalar quantizer dimension. This trend is evident in Figure 4, where the optimal K increases with the rate, highlighting the trade-off between the two distortions.

It can also be observed that for larger σ_W^2, the threshold rate mentioned above decreases. This is because, in such cases, the observation X contains significant noise, and retaining more dimensions effectively improves the signal-to-noise ratio, compensating for the observation noise. In contrast, when σ_W^2 is small, the estimation noise is already low, and the benefit of increasing K becomes negligible. Therefore, moderate dimensionality reduction can still enhance the rate–distortion performance at relatively high rates.

6. CONCLUSION

We have proposed the RPQ scheme to mitigate the rate inefficiency of the conventional QE scheme caused by system overdetermination in certain indirect quantization problems, and we further analyzed its rate–distortion performance. In future work, we plan to study the relationship between estimation distortion and the dimension K in more general source models. Another promising direction is to investigate how, by introducing a feedback system, both the retained dimension K and the projection matrix can be adaptively adjusted when the exact relationship between the state and observation is unknown.

Appendix A

Denote the components of Y and $\hat{Y}$ as $Y = Y_1, Y_2, \dots, Y_K$ and $\hat{Y} = \hat{Y}_1, \hat{Y}_2, \dots, \hat{Y}_K$. So

$$\hat{Y} = Q_1(Y + \mathbf{Z}) - \mathbf{Z}, \tag{16}$$

where $Q_1(\cdot)$ represents the uniform scalar quantization, $\mathbf{Z} = (Z, Z, \cdots)$, $Z \sim U(-\sqrt{3\epsilon_{QE}}, \sqrt{3\epsilon_{QE}}]$ and ϵ_{QE} represents the quantization distortion of each dimension, i.e., $\mathbb{E}\{(Q_1(Y_i + Z) - Z - Y_i)^2\} = \epsilon_{QE}$, where $i = 1, 2, \dots, K$. Hence we have:

$$\mathbb{E}\{\|BY - B\hat{Y}\|_2^2\} = \mathbb{E}\{\|BQ_1(Y + \mathbf{Z}) - BY - B\mathbf{Z}\|_2^2\}. \tag{17}$$

Denote $Q_1(Y+\mathbf{Z}) = q_1, q_2, \cdots, q_K$. Denote the element in the i-th row and j-th column of B as b_{ij}. Define $d_i = (\sum_{j=1}^{K} b_{ij}(q_j - Y_j - Z))^2$. Therefore,

$$\mathbb{E}\{\|BY - B\hat{Y}\|_2^2\} = \sum_{i=1}^{m} \mathbb{E}\{d_i\}. \tag{18}$$

For each $\mathbb{E}\{d_i\}$,

$$\mathbb{E}\{d_i\} = \mathbb{E}_Y\{\mathbb{E}_{Z|Y}\{d_i \mid Y\}\}. \tag{19}$$

First, we fix Y and solve the expectation w.r.t. Z. Denote this expectation as $\mathbb{E}_Z\{d_i\}$. So

$$\mathbb{E}_Z\{d_i\} = \mathbb{E}_Z\{(\sum_{j=1}^{K} b_{ij}(q_j - Y_j - Z))^2 \mid Y\} \tag{20}$$

For a single squared term, we have

$$\mathbb{E}\{b_{ij}^2(q_j - Y_j - Z)^2 \mid Y\} = b_{ij}^2 \epsilon_{QE}. \tag{21}$$

Consider a single cross term ($j \neq k$), we have

$$\begin{aligned} &\mathbb{E}_Z\{b_{ij}b_{ik}(q_j - Y_j - Z)(q_k - Y_k - Z) \mid Y\} \\ &= b_{ij}b_{ik}\mathbb{E}_Z\{(q_j - Y_j - Z)(q_k - Y_k - Z) \mid Y\}. \end{aligned}$$

Denote $\delta_j = Y_j - q_{min}(Y_j)$ and $q_{min}(Y_j)$ is defined as the greatest quantization value less than or equal to Y_j. Denote $q_{j,min} = q_{min}(Y_j)$. Without loss generality, assume $\delta_j \leq \delta_k$ for simplicity.

$$\begin{aligned} &\mathbb{E}_Z\{(q_j - Y_j - Z)(q_k - Y_k - Z) \mid Y\} \\ &= \frac{1}{2\sqrt{3\epsilon_{QE}}}(\int_{-\sqrt{3\epsilon_{QE}}}^{\sqrt{3\epsilon_{QE}}-\delta_k} (-z-\delta_j)(-z-\delta_k)dz \\ &\quad + \int_{\sqrt{3\epsilon_{QE}}-\delta_k}^{\sqrt{3\epsilon_{QE}}-\delta_j} (-z-\delta_j)(-z-\delta_k+2\sqrt{3\epsilon_{QE}})dz \\ &\quad + \int_{\sqrt{3\epsilon_{QE}}-\delta_j}^{\sqrt{3\epsilon_{QE}}} (-z-\delta_j+2\sqrt{3\epsilon_{QE}})(-z-\delta_k+2\sqrt{3\epsilon_{QE}})dz) \\ &= \frac{1}{2\sqrt{3\epsilon_{QE}}}(\int_{-\sqrt{3\epsilon_{QE}}}^{\sqrt{3\epsilon_{QE}}-\delta_k} (z+\delta_j)(z+\delta_k)dz \\ &\quad + \int_{\sqrt{3\epsilon_{QE}}-\delta_k}^{\sqrt{3\epsilon_{QE}}-\delta_j} (z+\delta_j)(z+\delta_k)dz \\ &\quad - 2\sqrt{3\epsilon_{QE}}\int_{\sqrt{3\epsilon_{QE}}-\delta_k}^{\sqrt{3\epsilon_{QE}}-\delta_j} z+\delta_j dz \\ &\quad + \int_{\sqrt{3\epsilon_{QE}}-\delta_j}^{\sqrt{3\epsilon_{QE}}} (z+\delta_j-2\sqrt{3\epsilon_{QE}})(z+\delta_k-2\sqrt{3\epsilon_{QE}})dz) \\ &= \frac{1}{2\sqrt{3\epsilon_{QE}}}(\int_{-\sqrt{3\epsilon_{QE}}}^{\sqrt{3\epsilon_{QE}}-\delta_j} (z+\delta_j)(z+\delta_k)dz \\ &\quad + \int_{\sqrt{3\epsilon_{QE}}-\delta_j}^{\sqrt{3\epsilon_{QE}}} (z+\delta_j-2\sqrt{3\epsilon_{QE}})(z+\delta_k-2\sqrt{3\epsilon_{QE}})dz) \\ &\quad - \int_{\sqrt{3\epsilon_{QE}}-\delta_k}^{\sqrt{3\epsilon_{QE}}-\delta_j} (z+\delta_j)dz. \end{aligned} \tag{22}$$

Denote the two terms of (22) as T_1 and T_2, respectively. Since $\int_{\sqrt{3\epsilon_{QE}}-\delta_j}^{\sqrt{3\epsilon_{QE}}} (z+\delta_j-2\sqrt{3\epsilon_{QE}})(z+\delta_k-2\sqrt{3\epsilon_{QE}})dz = \int_{-\sqrt{3\epsilon_{QE}}-\delta_j}^{-\sqrt{3\epsilon_{QE}}} (z+\delta_j)(z+\delta_k)dz$,

$$\begin{aligned} T_1 &= \frac{1}{2\sqrt{3\epsilon_{QE}}}\int_{-\sqrt{3\epsilon_{QE}}-\delta_j}^{\sqrt{3\epsilon_{QE}}-\delta_j} (z+\delta_j)(z+\delta_k)dz \\ &= \epsilon_{QE}. \end{aligned} \tag{23}$$

$$\begin{aligned} T_2 &= \int_{\sqrt{3\epsilon_{QE}}-\delta_k}^{\sqrt{3\epsilon_{QE}}-\delta_j} (z+\delta_j)dz \\ &= (\delta_k-\delta_j)\sqrt{3\epsilon_{QE}} - \frac{1}{2}(\delta_k-\delta_j)^2. \end{aligned} \tag{24}$$

Therefore,

$$\begin{aligned} &\mathbb{E}_Z\{(q_j - Y_j - Z)(q_k - Y_k - Z) \mid Y\} \\ &= \epsilon_{QE} + \frac{1}{2}(\delta_k-\delta_j)^2 - |\delta_k-\delta_j|\sqrt{3\epsilon_{QE}}. \end{aligned} \tag{25}$$

In the high-resolution cases, the conditional distribution within each quantization lattice is approximately uniform. Denote $\delta = \delta_1, \delta_2, \ldots, \delta_K$, $q_{j,min} = q_{min}(Y_j), j = [1:K]$. Let the quantization step size along each dimension be denoted by Δ_{QE}. Then $f(\delta) \approx \frac{1}{\Delta_{QE}^K}$. $\forall j,k \in [1:K]$, we have $f(\delta_j,\delta_k) \approx \frac{1}{\Delta_{QE}^2}$. Hence, we have

$$\mathbb{E}\{(\delta_k-\delta_j)^2\} \approx 2\epsilon_{QE} \tag{26}$$

$$\mathbb{E}\{|\delta_k-\delta_j|\} \approx \frac{2}{3}\sqrt{3\epsilon_{QE}}. \tag{27}$$

Therefore

$$\mathbb{E}\{(q_j - Y_j - Z)(q_k - Y_k - Z) \mid Y\} \approx 0. \tag{28}$$

So

$$\mathbb{E}\{d_i\} = \sum_{j=1}^{K} b_{ij}^2 \epsilon_{QE}, \tag{29}$$

and

$$\mathbb{E}\{\|BY - B\hat{Y}\|_2^2\} = \|B\|_F^2 \epsilon_{QE}. \tag{30}$$

Denote the rate of the scheme with projection matrix Φ and scalar uniform quantization as $R(\Phi, Q_1)$. We have

$$R(\Phi, Q_1) = h(Y) - K\log\Delta_{QE}. \tag{31}$$

Denote the total distortion of the scheme by D. From (9), we have

$$D_0(\Phi) + \|B\|_F^2 \epsilon_{QE} = D, \tag{32}$$

so

$$\epsilon_{QE} = \frac{D - D_0(\Phi)}{\|B\|_F^2}. \tag{33}$$

Combined with $\epsilon_{QE} = \frac{\Delta_{QE}^2}{12}$, we have

$$R(\Phi, Q_1) = h(Y) - \frac{K}{2} \log \frac{12(D - D_0(\Phi))}{\|B\|_F^2}. \tag{34}$$

Therefore, the distortion-rate performance of the RPQ scheme under high rate is:

$$D \approx D_0(\Phi) + \frac{\|B\|_F^2}{12} 2^{\frac{2}{K} h(Y)} 2^{-\frac{2}{K} R}. \tag{35}$$

When the K dimensions of Y are quantized and encoded separately, (31) becomes:

$$\begin{aligned} R(\Phi, Q_1) &= \sum_{i=1}^{K} (h(Y_i) - \log \Delta_{QE}) \\ &= \sum_{i=1}^{K} h(Y_i) - K \log \Delta_{QE}. \end{aligned} \tag{36}$$

Hence the distortion-rate performance of the RPQ scheme in this case is:

$$D \approx D_0(\Phi) + \frac{\|B\|_F^2}{12} 2^{\frac{2}{K} \sum_{i=1}^{K} h(Y)} 2^{-\frac{2}{K} R}. \tag{37}$$

Appendix B

Since S and Y are jointly Gaussian, the relationship between the two random variables can be written as $S = BY + Z$. Z is a Gaussian noise independent of Y and follows $\mathcal{N}(0, \Sigma_S - \Sigma_{SY}\Sigma_Y^{-1}\Sigma_{SY}^T)$, where Σ_S, Σ_Y and Σ_{SY} represents the covariance matrices of S, Y and their cross-covariance matrix, respectively. So we have

$$D_0(\Phi) = \mathrm{tr}[\Sigma_S - \Sigma_{SY}\Sigma_Y^{-1}\Sigma_{SY}^T]. \tag{38}$$

Specifically in the source described above, we have $\Sigma_S = 1$, $\Sigma_{SY} = \Sigma_{SX}\Phi^T$ and $\Sigma_Y = \Phi\Sigma_X\Phi^T$. Evidently, we have $\Sigma_{SX} = \mathbf{1}_n^T$ and $\Sigma_X = \mathbf{1}_n\mathbf{1}_n^T + \sigma_W^2 I_n$ in this source. Therefore, $\Sigma_{SY} = \mathbf{1}_n^T\Phi^T$. So we have

$$D_0(\Phi) = 1 - \mathbf{1}_n^T\Phi^T(\Phi\Sigma_X\Phi^T)^{-1}\Phi\mathbf{1}_n. \tag{39}$$

Denote $M = \Phi^T(\Phi\Sigma_X\Phi^T)^{-1}\Phi$. For Σ_Y, we have

$$\Sigma_Y = \Phi\mathbf{1}_n\mathbf{1}_n^T\Phi^T + \sigma_W^2\Phi\Phi^T. \tag{40}$$

Denote the column vectors of Φ as $\phi_1, \phi_2, \ldots, \phi_n$. Denote $\phi = \Phi\mathbf{1}_n = \sum_{i=1}^n \phi_i$. So

$$\Sigma_Y = \phi\phi^T + \sigma_W^2\Phi\Phi^T. \tag{41}$$

From Sherman-Morrison formula, we have

$$\begin{aligned} \Sigma_Y^{-1} &= (\sigma_W^2\Phi\Phi^T + \Phi\mathbf{1}_n\mathbf{1}_n^T\Phi^T)^{-1} \\ &= \frac{1}{\sigma_W^2}A^{-1} - \frac{A^{-1}\phi\phi^T A^{-1}}{\sigma_W^4 + \sigma_W^2\phi^T A^{-1}\phi}, \end{aligned} \tag{42}$$

where $A = \Phi\Phi^T$. So

$$M = \Phi^T\left(\frac{1}{\sigma_W^2}A^{-1} - \frac{A^{-1}\phi\phi^T A^{-1}}{\sigma_W^4 + \sigma_W^2\phi^T A^{-1}\phi}\right)\Phi. \tag{43}$$

Further, we have

$$\begin{aligned} \mathbf{1}_n^T M\mathbf{1}_n &= \frac{1}{\sigma_W^2}s^T A^{-1}s - \frac{s^T A^{-1}ss^T A^{-1}s}{\sigma_W^4 + \sigma_W^2 s^T A^{-1}s} \\ &= \frac{s^T A^{-1}s}{\sigma_W^2 + s^T A^{-1}s}. \end{aligned} \tag{44}$$

Denote $t = s^T A^{-1}s$, we have

$$\begin{aligned} \mathrm{tr}[\Sigma_S - \Sigma_{SY}\Sigma_Y^{-1}\Sigma_{SY}^T] &= 1 - \mathbf{1}_n^T M\mathbf{1}_n \\ &= 1 - \frac{t}{\sigma_W^2 + t} \\ &= \frac{\sigma_W^2}{\sigma_W^2 + t}. \end{aligned} \tag{45}$$

Denote $H = \Phi^T A^{-1}\Phi$, we have $t = \mathbf{1}_n^T H\mathbf{1}_n$. Since H is symmetric and idempotent, it follows that

$$\begin{aligned} t &= \mathbf{1}_n^T H^T H\mathbf{1}_n \\ &= \|H\mathbf{1}_n\|_2^2. \end{aligned} \tag{46}$$

For an arbitrary non-zero vector v, consider $\frac{\|Hv\|_2^2}{\|v\|_2^2}$. Without loss of generality, assume that $\|v\|_2^2 = 1$. Denote $e_1 = (1, 0, \ldots, 0)^T$. Then v can be expressed as $v = Ue_1$, where U is an orthogonal matrix. So we have

$$\begin{aligned} \|Hv\|_2^2 &= \|HUe_1\|_2^2 \\ &= e_1^T U^T HUe_1 \\ &= e_1^T(\Phi U)^T((\Phi U)(\Phi U)^T)^{-1}(\Phi U)e_1. \end{aligned} \tag{47}$$

Denote the row vectors of Φ as $\Phi_1^T, \Phi_2^T, \ldots, \Phi_K^T$, i.e.,

$$\Phi = \begin{pmatrix} \Phi_1^T \\ \Phi_2^T \\ \vdots \\ \Phi_K^T \end{pmatrix}. \tag{48}$$

So

$$\Phi U = \begin{pmatrix} \Phi_1^T U \\ \Phi_2^T U \\ \vdots \\ \Phi_K^T U \end{pmatrix}. \tag{49}$$

For the i-th row vector of Φ, $\Phi_i \sim \mathcal{N}(\mathbf{0}, \frac{1}{K} I_n)$. So

$$\begin{aligned} \mathbb{E}\{U^T \Phi_i \Phi^T U\} &= \frac{1}{K} U^T U \\ &= \frac{1}{K} I_n. \end{aligned} \tag{50}$$

Hence ΦU has the same distribution as Φ. Combined with (47), it follows that $\|He_1\|_2^2$ and $\|Hv\|_2^2$ are identically distributed. Therefore, $\frac{t}{n} = \frac{\|H\mathbf{1}_n\|_2^2}{\|\mathbf{1}_n\|_2^2}$ and $\frac{\|He_1\|_2^2}{\|e_1\|_2^2} = \|He_1\|_2^2$ are identically distributed. So for simplicity and without loss of generality, we study the distribution of $\|He_1\|_2^2$ in the following derivation.

$$\begin{aligned} \|He_1\|_2^2 &= e_1^T H^T H e_1 \\ &= e_1 \Phi^T (\Phi\Phi^T)^{-1} \Phi e_1. \end{aligned} \tag{51}$$

Define a random matrix $\Psi = \sqrt{K}\Phi$. Then Ψ has i.i.d. entries following $\mathcal{N}(0,1)$, and $\Psi^T(\Psi\Psi^T)^{-1}\Psi = \Phi^T(\Phi\Phi^T)^{-1}\Phi = H$. So

$$e_1^T H e_1 = e_1^T \Psi^T (\Psi\Psi^T)^{-1} \Psi e_1. \tag{52}$$

Define $a = \Psi e_1$, which is simply the first column of Ψ. Hence

$$e_1^T H e_1 = a^T (\Psi\Psi^T)^{-1} a. \tag{53}$$

Partition Ψ as

$$\Psi = \begin{pmatrix} a & \vdots & W \end{pmatrix}, \tag{54}$$

where the size of W is $K \times (n-1)$. $\Psi\Psi^T = aa^T + WW^T$. Denote $V = WW^T$. From Sherman-Morrison formula, we have

$$(aa^T + V)^{-1} = V^{-1} - \frac{V^{-1} a a^T V^{-1}}{1 + a^T V^{-1} a}. \tag{55}$$

$$\begin{aligned} a^T (aa^T + V)^{-1} a &= a^T V^{-1} a - \frac{(a^T V^{-1} a)^2}{1 + a^T V^{-1} a} \\ &= \frac{a^T V^{-1} a}{1 + a^T V^{-1} a}. \end{aligned} \tag{56}$$

$V \sim W_K(I_K, n-1)$ (W_K denotes the Wishart distribution) and $a \sim \mathcal{N}(\mathbf{0}, I_K)$. So $(n-1)a^T V^{-1} a \sim T^2(K, n-1)$ and $\frac{n-K}{K} a^T V^{-1} a \sim F(K, n-K)$. Since $\frac{n-K}{K} a^T V^{-1} a \sim F(K, n-K)$ follows an $F(K, n-K)$ distribution, it can be expressed as the ratio of two independent chi-square random variables normalized by their respective degrees of freedom, i.e.,

$$\frac{n-K}{K} a^T V^{-1} a = \frac{\frac{U_1}{K}}{\frac{U_2}{n-K}} = \frac{n-K}{K} \frac{U_1}{U_2}, \tag{57}$$

where $U_1 \sim \chi^2(K)$, $U_2 \sim \chi^2(n-K)$, and U_1 and U_2 are independent. After simplification, we obtain:

$$a^T V^{-1} a = \frac{U_1}{U_2}. \tag{58}$$

Hence

$$\|He_1\|_2^2 = \frac{U_1}{U_1 + U_2}. \tag{59}$$

Since $t = \|H\mathbf{1}_n\|_2^2 = n\|He_1\|_2^2$, letting $\|He_1\|_2^2 = l$, with probability density function $f(l)$, we have $t = nl$. Thus

$$\begin{aligned} \mathbb{E}\{\mathrm{tr}[\Sigma_S - \Sigma_{SY}\Sigma_Y^{-1}\Sigma_{SY}^T]\} &= \mathbb{E}\{\frac{\sigma_W^2}{\sigma_W^2 + t}\} \\ &= \int_0^1 f(l) \frac{\sigma_W^2}{nl + \sigma_W^2} dl. \end{aligned} \tag{60}$$

Denote the joint probability density function of U_1 and U_2 by $f_{U_1,U_2}(\cdot)$. We have

$$f_{U_1,U_2}(u_1,u_2) = \frac{1}{2^{\frac{K}{2}}\Gamma(\frac{K}{2})} u_1^{\frac{K}{2}-1} e^{-\frac{u_1}{2}} \frac{1}{2^{\frac{n-K}{2}}\Gamma(\frac{n-K}{2})} u_2^{\frac{n-K}{2}-1} e^{-\frac{u_2}{2}} \tag{61}$$

We have the following transformation:

$$l = \frac{u_1}{u_1 + u_2} \tag{62}$$

and

$$w = u_1 + u_2. \tag{63}$$

The inverse transformation is:

$$u_1 = lw \tag{64}$$

and

$$u_2 = (1-l)w. \tag{65}$$

The Jacobian determinant is

$$J = \begin{vmatrix} w & l \\ -w & 1-l \end{vmatrix} = w. \tag{66}$$

Hence

$$f_{L,W}(l,w) = f_{U_1,U_2}(u_1,u_2)\,|J| = f_{U_1,U_2}(lw,(1-l)w)w. \tag{67}$$

Combined with (61), we have

$$f_{L,W}(l,w) = \frac{1}{\Gamma(\frac{K}{2}))\Gamma(\frac{n-K}{2})2^{\frac{n}{2}}} l^{\frac{K}{2}-1}(1-l)^{\frac{n-K}{2}-1} w^{\frac{n}{2}-1} e^{-\frac{w}{2}}. \tag{68}$$

The marginal pdf of l is obtained by integrating the joint pdf with respect to w:

$$\begin{aligned} f_L(l) &= \int_0^{+\infty} f_{L,W}(l,w) dw \\ &= \frac{l^{\frac{K}{2}-1}(1-l)^{\frac{n-K}{2}-1}}{\Gamma(\frac{K}{2})\Gamma(\frac{n-K}{2})2^{\frac{n}{2}}} \int_0^{+\infty} w^{\frac{n}{2}-1} e^{-\frac{w}{2}} dw \\ &= \frac{l^{\frac{K}{2}-1}(1-l)^{\frac{n-K}{2}-1}}{\Gamma(\frac{K}{2})\Gamma(\frac{n-K}{2})2^{\frac{n}{2}}} 2^{\frac{n}{2}} \Gamma(\frac{n}{2}) \\ &= \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{K}{2})\Gamma(\frac{n-K}{2})} l^{\frac{K}{2}-1}(1-l)^{\frac{n-K}{2}-1}. \end{aligned} \tag{69}$$

Hence

$$\mathbb{E}\{D_0(\Phi) \mid K\} = \mathbb{E}\{\frac{\sigma_W^2}{\sigma_W^2 + nl}\}, l \sim Beta(\frac{K}{2}, \frac{n-K}{2}). \tag{70}$$

7. REFERENCES

[1] Yehoshua Bar-Hillel and Rudolf Carnap, "Semantic information," *The British Journal for the Philosophy of Science*, vol. 4, no. 14, pp. 147–157, 1953.

[2] Luciano Floridi, "Outline of a theory of strongly semantic information," *Minds and Machines*, pp. 197–221, 2004.

[3] Jie Bao, Prithwish Basu, Mike Dean, Craig Partridge, Ananthram Swami, Will Leland, and James A. Hendler, "Towards a theory of semantic communication," in *2011 IEEE Network Science Workshop*, 2011, pp. 110–117.

[4] Brendan Juba, *Universal Semantic Communication*, Springer, 2011.

[5] Petar Popovski, Osvaldo Simeone, Federico Boccardi, Deniz Gunduz, and Onur Sahin, "Semantic-effectiveness filtering and control for post-5g wireless connectivity," 2019.

[6] Marios Kountouris and Nikolaos Pappas, "Semantics-empowered communication for networked intelligent systems," *IEEE Communications Magazine*, vol. 59, no. 6, pp. 96–102, 2021.

[7] Hyowoon Seo, Yoonseong Kang, Mehdi Bennis, and Wan Choi, "Bayesian inverse contextual reasoning for heterogeneous semantics- native communication," *IEEE Transactions on Communications*, vol. 72, no. 2, pp. 830–844, 2024.

[8] Xunze Liu, Yifei Sun, Zhaorui Wang, Lizhao You, Haoyuan Pan, Fangxin Wang, and Shuguang Cui, "User centric semantic communications," 2024.

[9] T. Berger, *Rate Distortion Theory*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1971.

[10] R. Fontana, "Universal codes for a class of composite sources (corresp.)," *IEEE Transactions on Information Theory*, vol. 26, no. 4, pp. 480–482, 1980.

[11] Jiakun Liu, Wenyi Zhang, and H. Vincent Poor, "A rate-distortion framework for characterizing semantic information," in *2021 IEEE International Symposium on Information Theory (ISIT)*, 2021, pp. 2894–2899.

[12] Jiakun Liu, Shuo Shao, Wenyi Zhang, and H. Vincent Poor, "An indirect rate-distortion characterization for semantic sources: General model and the case of gaussian observation," *IEEE Transactions on Communications*, vol. 70, no. 9, pp. 5946–5959, 2022.

[13] J. Wolf and J. Ziv, "Transmission of noisy information to a noisy receiver with minimum distortion," *IEEE Transactions on Information Theory*, vol. 16, no. 4, pp. 406–411, 1970.

[14] R. Dobrushin and B. Tsybakov, "Information transmission with additional noise," *IRE Transactions on Information Theory*, vol. 8, no. 5, pp. 293–304, 1962.

[15] H. Witsenhausen, "Indirect rate distortion problems," *IEEE Transactions on Information Theory*, vol. 26, no. 5, pp. 518–521, 1980.

[16] Photios A. Stavrou, Yanina Shkel, and Marios Kountouris, "Indirect rate distortion functions with f-separable distortion criterion," in *2023 IEEE International Symposium on Information Theory (ISIT)*, 2023, pp. 1050–1055.

[17] Nir Shlezinger, Yonina C. Eldar, and Miguel R. D. Rodrigues, "Hardware-limited task-based quantization," *IEEE Transactions on Signal Processing*, vol. 67, no. 20, pp. 5223–5238, 2019.

[18] Alon Kipnis, Stefano Rini, and Andrea J. Goldsmith, "Multiterminal compress-and-estimate source coding," in *2016 IEEE International Symposium on Information Theory (ISIT)*, 2016, pp. 540–544.

[19] Alon Kipnis, Stefano Rini, and Andrea J. Goldsmith, "The rate-distortion risk in estimation from compressed data," *IEEE Transactions on Information Theory*, vol. 67, no. 5, pp. 2910–2924, 2021.

[20] Vildana Sulić, Janez Perš, Matej Kristan, and Stanislav Kovačič, "Dimensionality reduction for distributed vision systems using random projection," in *2010 20th International Conference on Pattern Recognition*, 2010, pp. 380–383.

[21] Kaihua Zhang, Lei Zhang, and Ming-Hsuan Yang, "Fast compressive tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 10, pp. 2002–2015, 2014.

[22] Tetsuya Takiguchi, Jeff Bilmes, Mariko Yoshii, and Yasuo Ariki, "Evaluation of random-projection-based feature combination on speech recognition," in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2010, pp. 2150–2153.

[23] D.L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[24] J. Ziv, "On universal quantization," *IEEE Transactions on Information Theory*, vol. 31, no. 3, pp. 344–347, 1985.

[25] Greg W Anderson, Alice Guionnet, and Ofer Zeitouni, *An introduction to random matrices*, Number 118. Cambridge university press, 2010.