DISTRIBUTED QUANTIZATION FOR COMPRESSED SENSING
Amirpasha Shirazinia, Saikat Chatterjee, Mikael Skoglund

Communication Theory Department, ACCESS Linnaeus Centre, KTH Royal Institute of Technology

ABSTRACT
We study distributed coding of compressed sensing (CS) measurements using vector quantizer (VQ). We develop a distributed framework for realizing optimized quantizer that enables encoding CS measurements of correlated sparse sources followed by joint decoding at a fusion center. The optimality of VQ encoder-decoder pairs is addressed by minimizing the sum of mean-square errors between the sparse sources and their reconstruction vectors at the fusion center. We derive a lower-bound on the end-to-end performance of the studied distributed system, and propose a practical encoder-decoder design through an iterative algorithm.

Index Terms—Compressed sensing, distributed source coding, vector quantization, correlation, mean square error.

1. INTRODUCTION
With the rapid growth of sensor networks and their popularity to accomplish autonomous tasks such as sensing and computation, distributed coding of correlated sources has attracted much attention. In short, distributed coding considers separately encoding of correlated sources (from different terminals), and decoding the coded symbols jointly at a fusion center (FC). Distributed lossless and lossy source coding are well-developed fields and various theoretical results on this topic have been established [1–7]. The characteristics of sensor networks motivate the development of new techniques and algorithms which need to be aware of the systems’ limited resources, computational complexity and power consumption. In this regard, compressed sensing (CS) [8] can be considered as an emerging tool for signal compression and acquisition that significantly reduces costs due to sampling and reconstruction, leading to lower power consumption and computations. By exploiting signal’s sparsity in a known domain, CS allows the signal to be retrieved from relatively few measurements through a non-linear procedure.

The problem of source coding (through quantization) for CS measurements in a point-to-point setup has gained significant attention recently, and a wide range of interesting problems regarding the design and analysis of CS reconstruction and quantization algorithms have been formulated, see, e.g., [9–17, 17–22]. Considering communication channel errors, in [23], we have studied joint source-rithms have been formulated, see, e.g., [9–17, 17–22]. Considering the design and analysis of CS reconstruction and quantization algorithms recently, and a wide range of interesting problems regarding measurements in a point-to-point setup has gained significant attention in [23], we have studied joint source-coding schemes for realizing optimized quantizer that enables encoding CS measurements followed by joint decoding at the FC. To the best of our knowledge, the distributed design and analysis of VQ for CS measurements—which can be of great importance in sensor networks—have not been addressed yet.

In this work, we consider, without loss of generality, two linear CS systems where they obtain few number of measurements from two correlated sparse vectors. The low-dimensional (possibly noisy) measurements are encoded using VQ, and then transmitted to a FC for joint reconstruction of correlated sparse sources. As a performance criterion in this distributed setup, we are interested in the sum of mean square error (MSE) distortions between the sparse source vectors and their reconstruction vectors at the FC. By applying the end-to-end MSE criterion, our contributions are as following: (1) Establishing (necessary) conditions for optimality of VQ encoder-decoder pairs. (2) Deriving analytical expressions for the minimum mean square error (MMSE) estimators of correlated sparse sources from noisy CS measurement vectors; the MMSE estimation is an integral part of optimum VQ scheme in distributed CS setup. (3) Providing a theoretical lower-bound on the MSE performance. (4) Developing a practical VQ encoder-decoder design algorithm through an iterative algorithm.

2. DISTRIBUTED MMSE ESTIMATION FOR CS
In this section, we introduce the CS system setup, and study the MMSE estimation of correlated sparse sources from noisy CS measurements in a distributed framework. We consider an exact K-sparse (in a known basis) vector \( \Theta \in \mathbb{R}^N \) comprised of \( K \) non-zero coefficients (\( K < N \)). We define the random support set of the vector \( \Theta \) as \( \Theta = [\Theta_1, \ldots, \Theta_N]^T \) as \( S \ni \{n \in \{1, 2, \ldots, N\} : \Theta_n \neq 0 \} \) with \( |S| = \|\Theta\|_0 = K \). Here, \( \| \cdot \|_0 \) and \( \| \cdot \|_2 \) denote the \( l_0 \) norm and cardinality, respectively. The non-zero coefficients of \( \Theta \) are drawn according to \( \Theta_n \sim \mathcal{N}(0, \sigma_n^2) \), \( n \in S \). We assume a distributed CS system with two correlated sources \( X_1, X_2 \in \mathbb{R}^N \) in which the correlation is established according to

\[
X_l = \Theta + Z_l, \quad l \in \{1, 2\},
\]

where \( Z_l \sim [Z_{1l}, \ldots, Z_{Nl}]^T \in \mathbb{R}^N \) is an exact K-sparse vector with a common support set \( \hat{S} \) whose non-zero components are drawn as \( Z_{nl} \in \mathcal{N}(0, \sigma_n^2) \), \( n \in \hat{S} \); thus \( \|Z_l\|_0 = K, \ l \in \{1, 2\} \). We assume that \( Z_1 \) and \( Z_2 \) are uncorrelated with each other and with the common signal \( \Theta \). This joint sparsity model (also known as JSM-2) for distributed CS was first introduced in [27].

Here, we stress that although the extension to arbitrary number of correlated sources is straightforward, the assumption of a distributed system with two sources simplifies the presentation through-out the paper. We assume, without loss of generality, that \( \sigma_{Z_l}^2 = \sigma_{Z_1}^2 = \sigma_{Z_2}^2 = \sigma_Z^2 + \sigma_s^2 = 1 \). To measure the amount of correlation between sources, we define the correlation ratio as

\[
\rho \triangleq \frac{\sigma_Z^2}{\sigma_s^2}.
\]
Hence, $\sigma_{x_1}^2 = \frac{1}{\rho_1}$ and $\sigma_{x_2}^2 = \frac{1}{\rho_2}$. Also, note that $\rho \to \infty$ implies that the sources are highly correlated, whereas $\rho \to 0$ means that they are highly uncorrelated. The remote sparse sources $X_1$ and $X_2$ are measured by CS-based sensors, leading to measurement vectors $Y_1 \in \mathbb{R}^{M_1}$ and $Y_2 \in \mathbb{R}^{M_2}$ described as

$$Y_1 = \Phi_1 X_1 + W_1, \quad l \in \{1, 2\}, \quad \|X_1\|_0 = K,$$

where $\Phi_1 \in \mathbb{R}^{M_1 \times N}$ is a fixed sensing matrix of the $l^{th}$ sensor, and $W_1 \in \mathbb{R}^{M_1}$ is an additive measurement noise vector independent of other sources. Without loss of generality, we assume that $M_1 = M_2 \equiv M$, and in the spirit of CS, we consider that $K < M < N$.

In order to reconstruct an unknown sparse source from noisy under-sampled measurement vectors in a distributed fashion, we define CS reconstruction distortion as

$$D_{cs} = \frac{1}{2K} \sum_{l=1}^{2} \mathbb{E}[(\|X_l - \tilde{X}_l\|_2^2)],$$

where $\tilde{X}_l \in \mathbb{R}^N (l \in \{1, 2\})$ is an estimation vector of the sparse source $X_l$ from noisy CS measurements $Y_1$ and $Y_2$. Here, $\| \cdot \|_2$ denotes $\ell_2$ norm. In this work, we focus on the Bayesian minimum mean square error (MMSE) framework for distributed reconstruction since we are interested in the MSE performance of random sparse vectors with known priors. Later in Section 4, it will be shown that the MMSE estimator is central in developing distributed quantizer design procedure which is the main focus of this paper. The following proposition provides an analytical expression for the MMSE estimator. We omit the proof (given in [28]) due to space limitation.

**Proposition 1** Consider the following assumptions: (i) the $K$-sparse sources $X_1, X_2$ are correlated based on the model (1) with ratio $\rho$. (ii) The $K$ elements of the support set are uniformly drawn from all $\binom{N}{K}$ possibilities. (iii) The measurement noise vector is distributed as $W_l \sim \mathcal{N}(0, \sigma_N^2 I_{M_1})$, $l \in \{1, 2\}$, which is uncorrelated with the CS measurements and sources. Then, the MMSE estimator of $X_l$ given the noisy CS measurement vector $y = [y_1^T, y_2^T]^T$ that minimizes $D_{cs}$ in (4), is obtained by $\tilde{x}_l^*(y) \triangleq \mathbb{E}[X_l | y]$ which has the following closed form expression

$$\tilde{x}_l^*(y) \triangleq [\tilde{x}_1^*(y) \tilde{x}_2^*(y)]^T = \sum_{S \subset \Omega} \beta_S \tilde{x}_l^*(y, S) \sum_{S \subset \Omega} \beta_S,$$

where $\tilde{x}_l^*(y, S) \triangleq \mathbb{E}[x_l | y, S]$ and within its support

$$\tilde{x}_l^*(y, S) = \left[ \begin{array}{cccc} I_K & I_K & 0_K & 0_K \\ 0_K & I_K & 0_K & 0_K \\ C & D^{-1} \end{array} \right],$$

and otherwise zeros. Further,

$$\beta_S = e^{\frac{1}{2}(\rho^2 (N^2 s^2 - F^2 \Phi_1^T \Phi_2 \Phi_1) - 2FN^2 - \sum_{l=1}^{2} \phi_i M_{x_l} \Phi_2 S)}.$$
\[
D_l(y_1, i_1) \triangleq \mathbb{E} \| X_1 - D_l(l_1, l_2)\|_2^2 | y_1, i_1 \|_2^2 | y_1 + \sum_{i_2=0}^{q_l-1} P(i_2 | y_1) \left( \| D_l(l_1, i_2) \|_2^2 - 2\mathbb{E}[X_l^T | y_1, i_2 | D_l(l_1, i_2)] \right)
\]
\[
\mathbb{E}[\| X_1 \|_2^2 | y_1] + \sum_{i_2=0}^{q_l-1} \int_{y_2 \in \mathbb{R}^{q_l}} \left( \| D_l(l_1, i_2) \|_2^2 - 2\mathbb{E}[X_l^T | y_1, y_2 | D_l(l_1, i_2)] \right) p(y_2 | y_1) dy_2
\]
\[
i_1^* = \arg \min_{i_1 \in \mathcal{I}} \left\{ \sum_{i_2=0}^{q_l-1} \int_{y_2 \in \mathbb{R}^{q_l}} \left[ \| D_l(l_1, i_2) \|_2^2 - 2X_l^T(y_1, y_2)D_l(l_1, i_2) \right] p(y_2 | y_1) dy_2 \right\},
\]

where \( p(y_1) \) is the \( M \)-fold probability density function (pdf) of the measurement vector \( Y_1 \). Since \( p(y_1) \) is a non-negative value, in order to optimize the mapping \( \mathbb{E}_1 \) in the sense of minimizing \( D_l \), it suffices to minimize the expression inside the braces in (8). Thus, the optimized encoding index \( i_1^* \) is obtained by
\[
i_1^* = \arg \min_{i_1 \in \mathcal{I}} \{ D_l(y_1, i_1) + D_2(y_1, i_1) \}.
\]

Now, \( D_l(y_1, i_1) \) can be rewritten as (10) on top of the page, where (a) follows by expanding the conditional expectation and the fact that \( X_1 \) and \( D_l(l_1, l_2) \) are independent conditioned on \( y_1, i_1, i_2 \). Further, (b) follows from marginalization of the expression inside the brackets in (a) over \( i_2 \) and \( y_2 \). In a similar way, \( D_2(y_1, i_1) \) can be parameterized similar to (10) with the only difference that \( X_1 \) and \( D_2(i_1, i_2) \) are replaced with \( X_2 \) and \( D_2(i_1, i_2) \), respectively. Following (9) and (10), the MSE-minimizing encoding index \( i_1 \) is given by (11), where \( X_l^+(y_1, y_2) \|_2^2 = [X_l^+(y_1, y_2) X_l^+(y_1, y_2)]^T \) and \( D(l_1, i_2) \|_2^2 = [D_l(l_1, i_2) D_l(l_1, i_2)]^T \). Here, the codevectors \( D(l_1, i_2), (l \in \{1, 2\}) \), are given, and the vector \( X_l^+(y_1, y_2) \) denotes the MMSE estimators derived in Proposition 1. It should be mentioned that although the observation at terminal 2, \( y_2 \), appears in the formulation of the optimized encoder at terminal 1, it is finally integrated out.

Assuming all encoders are fixed, it can be shown that the MSE-minimizing decoder is given by
\[
D_l(l_1, i_2) = \mathbb{E}[X_l | i_1, i_2] = \int_{y_1 \in \mathcal{I}} \int_{y_2 \in \mathbb{R}^{q_l}} X_l^+(y_1, y_2) p(y_1, y_2) dy_1 dy_2
\]
\[
\int_{y_2 \in \mathcal{I}} \int_{y_2 \in \mathbb{R}^{q_l}} p(y_1, y_2) dy_1 dy_2,
\]

where the second equality can be shown by marginalizing of the conditional expectation over \( Y_1 \) and \( Y_2 \) and using the Bayes’ rule.

When there is no correlation between sources (\( \rho \to 0 \)), then it can be shown that the optimized encoder (11) and the optimized decoder in (12) boil down to the optimized encoder and decoder in the point-to-point source coding of CS measurements, cf. [23].

We emphasize that we do not assume any sparse structure on the reconstructed vectors at the receiving-ends in contrast to conventional \( l_1 \)-norm reconstruction methods. The reason is due to the fact that we are interested in the final reconstruction MSE, where considering any kind of sparse structure might degrade the performance.

Now, we analyze the end-to-end MSE for our studied distributed system. Recall the MMSE estimation of the correlated sources \( \hat{X}_l^+ \|_2^2 = \mathbb{E}[X_l | Y_1, Y_2], l \in \{1, 2\} \), then we rewrite the end-to-end MSE as
\[
D \|_{(a)} \left[ \frac{1}{2K} \sum_{l=1}^{2} \mathbb{E}[\| X_l - \hat{X}_l^+ \|_2^2] + \frac{1}{2K} \sum_{l=1}^{2} \mathbb{E}[\| X_l^+ - \hat{X}_l^+ \|_2^2] \right] = D_{\text{cor}} + D_t,
\]

where (a) can be shown by the definition of the MMSE estimator \( \hat{X}_l^+ \) and by using the Markov property \( X_l \to (l_1, l_2) \to \hat{X}_l, l \in \{1, 2\} \).

Interestingly, (13) implies that, without loss of optimality, the end-to-end MSE, denoted by \( D \), can be summed up as CS reconstruction MSE (of the MMSE estimator), denoted by \( D_{\text{cor}} \), and quantization MSE, denoted by \( D_t \). This property together with information theoretic results in [5] can be used to develop a lower-bound on \( D \) provided by the following theorem. We omit the proof (given in [28]) due to lack of space.

**Theorem 2** Consider the assumptions given in Proposition 1. Let the total quantization rate be \( R = R_1 + R_2 \) bits/vector where \( R_l \) is the assigned quantization rate at terminal \( l \in \{1, 2\} \), then the asymptotic (in quantization rate) end-to-end MSE (8) is lower-bounded as
\[
D \geq \max \left\{ D_{\text{cor}}(b), D_{\text{cor}} \right\},
\]

where \( D_{\text{cor}}(b) = \frac{1}{2} \left( 1 + \frac{\rho^2}{(1 + \rho)^2} \right)^2 - \frac{2(n-\log_2(\frac{\rho^2}{(1 + \rho)^2}))}{(1 + \rho)^2} \). And \( D_{\text{cor}} \) is calculated by (4) corresponding to the distortion due to MMSE estimation of correlated sparse sources \( X_1 \) and \( X_2 \) from noisy CS measurements derived in (5) of Proposition 1.

When CS measurements are noisy, i.e. \( \sigma_{x_1}^2, \sigma_{x_2}^2 \neq 0 \), it can be verified as the quantization rate \( R \) increases, the end-to-end MSE saturates to \( D_{\text{cor}} \) since \( D_{\text{cor}}(b) \) decays exponentially, but \( D_{\text{cor}} \) becomes constant by quantization rate. The source correlation ratio \( \rho \) also plays an important role on the level of the lower-bound (14). By taking the first derivative of \( D_{\text{cor}}(b) \) in (15) with respect to \( \rho \), it can be verified that the derivative is always negative which means as the source correlation \( \rho \) increases, the lower-bound decreases.

### 5. Practical Design Algorithm

The necessary optimal conditions in (11) (and its equivalence \( i_2^* \)) and (12) can be combined in an alternate-iterate training algorithm in order to design optimized encoder-decoder pairs for distributed quantization of CS measurements which converges to locally optimal encoder regions and decoder codevectors. A natural order to optimize the mappings is: 1) the first encoder, 2) the first decoder, 3) the second encoder and 4) the second decoder.

To implement the algorithm, we make some modifications. The integral in (11) (for calculating the optimized encoder) cannot be solved in closed form in general. Hence, we modify the integral, and compute it numerically. The integral in (11) can be approximated as
\[
\| D(l_1, i_2) \|_2^2 = \int \mathbb{E}[X | y] P(i_2 | y_1) - 2P(i_2 | y_1) \mathbb{E}[X | y_1, i_2, D(l_1, i_2)]
\]

where we have only approximated \( y_1 \) by its scalar-quantized representation, denoted by \( \tilde{y}_1 \), using \( r_0 \)-bit nearest-neighbor coding (using, e.g., LBG algorithm [30]). Hence, \( y_1 \) is discretized, and
measurements improve the performance which is expected since the theory. From Figure 1(a), we observe that increasing number of CS measurements. Further, the vectors $y_1$ and $y_2$ are pre-quantized using $r_y = 3$ bits per measurement entry. We vary the correlation ratio from very low $\rho = 10^{-2}$ to very high values $\rho = 10^2$, and compare the simulation results with the lower-bound derived in (14) of Theorem 2. The results are shown in Figure 1(b) for measurement rates $\alpha = \frac{2}{3}, \frac{1}{3}, \frac{1}{2}$. As would be expected, at a fixed quantization rate $R$ and correlation ratio $\rho$, increasing $\rho$ improves the performance since the sources are reconstructed from larger sets of information. Hence, the end-to-end MSE decreases, and the curves approach the lower-bound. The correlation is a useful factor to reduce the end-to-end distortion which is also reflected from the lower-bound. This behavior can be interpreted as follows. When the sources are fully correlated, $\rho \rightarrow \infty$, the two sources can be viewed as a single source, and the FC is able to reconstruct the source jointly from two sets of received quantized indexes. Hence, the performance is maximized. On the other hand, when the sources are uncorrelated, $\rho \rightarrow 0$, the FC reconstructs the sources from two sets of independent quantized indexes. Thus, there is no gain in joint decoding at the FC.

Now, we investigate how the performance varies by quantization rate. We use the simulation parameter set $(N = 8, K = 2, \alpha = 5/8, \text{SNMR}_1 = \text{SNMR}_2 = 10)$ dB. In Figure 1(c), we illustrate the end-to-end MSE of the proposed design method as a function of total quantization rate $R = R_1 + R_2$ (with $R_1 = R_2$) for two values of correlation ratios: $\rho = 1$ (low-correlated sources) and $\rho = 100$ (high-correlated sources). The simulation curves are compared with the lower-bound in (14) corresponding to $\rho = 1, 100$. From Figure 1(c), we observe that the performance improves by increasing quantization rate. Moreover, increasing correlation between sources reduces the MSE as observed from other experiments too. We also note that as quantization rate increases, all the curves converge to their respective MSE floors, specified by $D_{cs}$ that can be seen from the lower-bounds in Figure 1(c).

7. CONCLUSIONS

We studied the design and analysis of the distributed vector quantization of CS measurements. We derived necessary conditions for optimality of encoder-decoder pairs by minimizing end-to-end MSE. We analyzed the MSE and proved that it is the sum of CS reconstruction MSE (of MMSE estimator) and quantization MSE. This result helped us to derive a lower-bound on the end-to-end MSE. Simulations revealed that correlation between sources, besides compression resources e.g. measurement and quantization rates, is an effective factor on CS reconstruction MSE as well as quantization MSE.
8. REFERENCES


