Quantization for Compressed Sensing Reconstruction

John Z. Sun, Vivek K. Goyal

Quantization for Compressed Sensing Reconstruction

John Z. Sun and Vivek K Goyal

Massachusetts Institute of Technology, Cambridge, MA 02139 USA
johnsun@mit.edu, vgoyal@mit.edu

Abstract:

Quantization is an important but often ignored consideration in discussions about compressed sensing. This paper studies the design of quantizers for random measurements of sparse signals that are optimal with respect to mean-squared error of the lasso reconstruction. We utilize recent results in high-resolution functional scalar quantization and homotopy continuation to approximate the optimal quantizer. Experimental results compare this quantizer to other practical designs and show a noticeable improvement in the operational distortion-rate performance.

1. Introduction

In practical systems where information is stored or transmitted, data must be discretized using a quantization scheme. The design of the optimal quantizer for a given stochastic source has been well studied and is surveyed in [6]. Here, optimal means the quantizer minimizes the error as measured by some distortion metric. In this paper, we explore optimal quantization for an emerging nonadaptive compression paradigm called compressed sensing (CS) [1, 4]. Several authors have studied the asymptotic reconstruction performance of quantized random measurements assuming a mean-squared error (MSE) distortion metric [3, 5]. Other previous work presented modifications to existing reconstruction algorithms to mitigate distortion resulting from standard quantizers [3, 7] or modified quantization that can be viewed as the binning of quantizer output indexes [10].

Our contribution is to reduce distortion due to quantization through design of the quantizer itself. The key observation is simply that the random measurements are used as arguments in a *nonlinear* reconstruction function. Thus, minimizing the MSE of the measurements is not equivalent to minimizing the MSE of the reconstruction. We use the theory for high-resolution distributed functional scalar quantization (DFSQ) recently developed in [9] to design optimal quantizers for random measurements. To obtain concrete results, we choose a particular reconstruction function (lasso [11]) and distributions for the source data and sensing matrix. However, the general principle of obtaining improvements through the use of DFSQ theory holds more generally, and we address the conditions that must be satisfied for sensing and reconstruction. Also, rather than develop results for fixed and variable rate in parallel, we present only fixed rate. To concentrate on the central ideas, we choose signal and sensing models that obviate discussion of quantizer overload.

2. Background

In our notation, a random vector is always lowercase and in bold. A subscript then indicates an element of the vector. Also, an unbolded vector y corresponds to a realization of the random vector $\mathbf{y}$.

2.1 Distributed functional scalar quantization

In standard fixed-rate scalar quantization [6], one is asked to design a quantizer Q that operates separably over its components and minimizes MSE between a probabilistic source vector $\mathbf{y} \in \mathbb{R}^M$ and its quantized representation $\hat{\mathbf{y}} = Q(\mathbf{y})$. The resulting optimization is

$$\min_Q E\left[\|\mathbf{y} - Q(\mathbf{y})\|^2\right],$$

subject to the constraint that the maximum number of codewords or quantization levels for each $\mathbf{y_i}$ is less than 2^{R_i}. We can use high-resolution theory to find the quantizer point density of the optimal quantizer.

In DFSQ [9], the goal is to create a quantizer that minimizes distortion for some scalar function $g(\mathbf{y})$ of the source vector $\mathbf{y}$ rather than the vector itself. Hence, the optimization is now

$$\min_Q E\left[|g(\mathbf{y}) - g(Q(\mathbf{y}))|^2\right]$$

such that the maximum number of codewords or quantization levels representing each $\mathbf{y_i}$ is less than 2^{R_i}. To apply the following model, we need $g(\cdot)$ and $f_{\mathbf{y}}(\cdot)$ to satisfy certain conditions:

C1. $g(\mathbf{y})$ is smooth and monotonic for each $\mathbf{y_i}$.

C2. The partial derivative $g_i(y) = \partial g(y)/\partial y_i$ is defined and bounded for each i.

C3. The joint pdf of the source variables $f_{\mathbf{y}}(y)$ is smooth and supported in a compact subset of $\mathbb{R}^M$.

For valid $g(\cdot)$ and $f_{\mathbf{y}}(\cdot)$ pairs, we define a set of functions

$$\gamma_i(t) = \left(E\left[|g_i(\mathbf{y})|^2 \mid \mathbf{y_i} = t\right]\right)^{1/2}. \quad (1)$$

We call $\gamma_i(t)$ the *sensitivity* of $g(\mathbf{y})$ with respect to the source variable $\mathbf{y_i}$. The optimal point density is then

$$\lambda_i(t) = C\left(\gamma_i^2(t) f_{\mathbf{y_i}}(t)\right)^{1/3}, \quad (2)$$

for some normalization constant C, which leads to a total operational distortion-rate

$$D(\{R_i\}) = \sum_i 2^{-2R_i} E\left[\frac{\gamma_i^2(\mathbf{y_i})}{12\lambda_i^2(\mathbf{y_i})}\right]. \quad (3)$$

The sensitivity $\gamma_i(t)$ serves to reshape the quantizer, giving better resolution to regions of $\mathbf{y_i}$ that have more impact on $g(\mathbf{y})$, thereby reducing MSE.

Similar results for variable-rate quantizers are also presented in [9]. However, we will only consider the fixed-rate case in this paper. The theory of DFSQ can be extended to a vector of functions, where $\mathbf{x_j} = g^{(j)}(\mathbf{y})$ for $1 \le j \le N$. Since the cost function is additive in its components, we can show that the overall sensitivity for each component $\mathbf{y_i}$ is

$$\gamma_i(t) = \frac{1}{N}\sum_{j=1}^{N} \gamma_i^{(j)}(t), \quad (4)$$

where $\gamma_i^{(j)}(t)$ is the sensitivity of the function $g^{(j)}(\mathbf{y})$ with respect to $\mathbf{y_i}$.

2.2 Compressed Sensing

CS refers to estimation of a signal at a resolution higher than the number of data samples, taking advantage of sparsity or compressibility of the signal and randomization in the measurement process [1, 4]. We will consider the following formulation. The input signal $x \in \mathbb{R}^N$ is K-sparse in some orthonormal basis Ψ, meaning the transformed signal $u = \Psi^{-1}x \in \mathbb{R}^N$ contains only K nonzero elements. Consider a length-M measurement vector $y = \Phi x$, where $\Phi \in \mathbb{R}^{M\times N}$ with $K < M < N$ is a realization of $\mathbf{\Phi}$. The major innovation in CS (for the case of sparse u considered here) is that recovery of x from y via some computationally-tractable reconstruction method can be guaranteed asymptotically almost surely.

Many reconstruction methods have been proposed including a linear program called basis pursuit [2] and greedy algorithms like orthogonal matching pursuit (OMP) [12]. In this paper, we focus on a convex optimization called lasso [11], which takes the form

$$\hat{x} = \arg\min_x \left(\|y - \Phi x\|_2^2 + \mu\|\Psi^{-1}x\|_1\right). \quad (5)$$

As one sample result, lasso leads to perfect sparsity pattern recovery with high probability if $M \sim 2K\log(N-K) + K$ under certain conditions on $\mathbf{\Phi}$, μ, and the scaling of the smallest entry of u [13]. Unlike in [5], our concern in this paper is not how the scaling of M affects performance, but rather how the accuracy of the lasso computation (5) is affected by quantization of y.

A method for understanding the set of solutions to (5) is the homotopy continuation (HC) method [8]. HC considers the regularization parameter μ at an extreme point (e.g., very large μ so the reconstruction is all zero) and slowly varies μ so that all sparsities and the resulting reconstructions are obtained. It is shown that there are N values of μ where the lasso solution changes sparsity, or equivalently $N+1$ intervals over which the sparsity does not change. For μ in the interior of one of these intervals, the reconstruction is determined uniquely by the solution of a linear system of equations involving a submatrix of Φ. In particular, for a specific choice μ^* and observed random measurements y,

$$2\Phi_{J_{\mu^*}}^T \Phi_{J_{\mu^*}} \hat{x} + \mu^* v = 2\Phi_{J_{\mu^*}}^T y, \quad (6)$$

where $v = \mathrm{sgn}(\hat{x})$ and $\Phi_{J_{\mu^*}}$ is the submatrix of Φ with columns corresponding to the nonzero elements $J_{\mu^*} \subset \{1, 2, \ldots, N\}$ of $\hat{x}$.

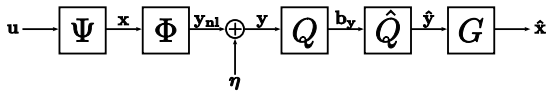

Figure 1: A compressed sensing model with quantization of measurement vector $\mathbf{y}$. The vector $\mathbf{y_{nl}}$ denotes the noiseless random measurements.

3. Problem Model

Figure 1 presents a CS model with quantization. Assume without loss of generality that $\Psi = I_N$ and hence the (random) signal $\mathbf{x} = \mathbf{u}$ is K-sparse. Also assume a random matrix $\mathbf{\Phi}$ is used to take measurements, and additive Gaussian noise perturbs the resulting signal, meaning the continuous-valued measurement vector is $\mathbf{y} = \mathbf{\Phi x} + \boldsymbol{\eta}$. The sampler wants to transmit the measurements with total rate R and encodes $\mathbf{y}$ into a transmittable bitstream $\mathbf{b_y}$ using encoder Q. Next, a decoder $\hat{Q}$ produces a quantized signal $\hat{\mathbf{y}}$ from $\mathbf{b_y}$. Finally, a reconstruction algorithm G outputs an estimate $\hat{\mathbf{x}}$. The function G is a black box that may represent lasso, OMP or another CS reconstruction algorithm.

We now present a probabilistic model for the input source and sensing matrix. It is chosen to guarantee finite support on both the input and measurement vectors, and prevent overload errors for quantizers with small R. However, we emphasize that the following theory is general, and other choices for $\mathbf{x}$ and $\mathbf{\Phi}$ are possible for large enough R.

Assume the K-sparse vector $\mathbf{x}$ has random sparsity $\mathbf{J}$ chosen uniformly from all possibilities, and each nonzero component $\mathbf{x_i}$ is distributed iid $\mathcal{U}(-1,1)$. Also assume the additive noise vector $\boldsymbol{\eta}$ is distributed iid Gaussian with zero mean and variance σ^2. Finally, let $\mathbf{\Phi}$ correspond to random projections such that each column $\boldsymbol{\phi_j} \in \mathbb{R}^M$ has unit energy ($\|\boldsymbol{\phi_j}\|^2 = 1$). The columns of $\mathbf{\Phi}$ thus form a set of N random vectors chosen uniformly on the unit $(M-1)$-hypersphere. Since $\mathbf{y} = \Phi\mathbf{x}$,

$$\mathbf{y_i} = \sum_{j=1}^{N} \mathbf{\Phi_{ij}x_j} = \sum_{j\in\mathbf{J}} \underbrace{\mathbf{\Phi_{ij}x_j}}_{\mathbf{z_{ij}}}.$$

The distribution of each $\mathbf{z_{ij}}$ is found using derived distributions. The resulting pdfs can be shown to be iid $f_{\mathbf{z}}(z)$, where $\mathbf{z}$ is a scalar random variable that is identical in distribution to each $\mathbf{z_{ij}}$. The distribution of $\mathbf{y_i}$ is then the $K-1$ convolution cascade of $f_{\mathbf{z}}(z)$ with itself. Thus, $f_{\mathbf{y}}(y)$ is smooth and supported for $\{|\mathbf{y_i}| \le K\}$, satisfying

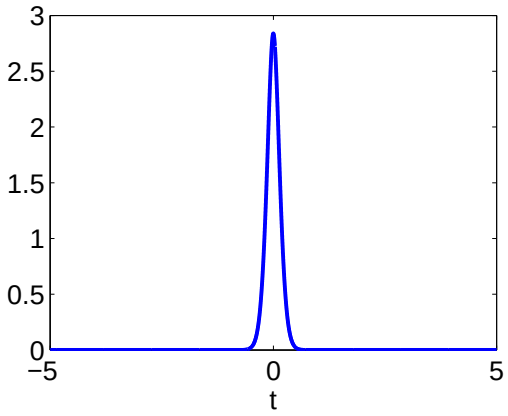

Figure 2: Distribution $f_{\mathbf{y_i}}(t)$ for $(K, M, N) = (5, 71, 100)$. The support of $\mathbf{y_i}$ is the range $[-K, K]$, where K is the sparsity of the input signal. However, the probability is only non-negligible for small $\mathbf{y_i}$.

condition C3 for DFSQ. Figure 2 illustrates the distribution of $\mathbf{y_i}$ for a particular case.

The reconstruction algorithm G is a function of the measurement vector $\mathbf{y}$ and sampling matrix $\mathbf{\Phi}$. We will show that if $G(\mathbf{y}, \mathbf{\Phi})$ is lasso with a proper relaxation variable μ, then conditions C1 and C2 are met. Using HC, we see $G(\mathbf{y}, \mathbf{\Phi})$ is a piecewise smooth function that is also piecewise monotonic with every $\mathbf{y_i}$ for a *fixed* μ. Moreover, for every μ the reconstruction is an affine function of the measurements through (6), so the partial derivative with respect to any element y_i is piecewise defined and smooth (constant in this case). Conditions C1 and C2 are therefore satisfied.

4. Optimal Quantizer Design

We now pose the optimal fixed-rate quantizer design as a DFSQ problem. For a given noise variance σ^2, choose an appropriate μ^* to form the best reconstruction $\hat{\mathbf{x}}$ from the unquantized random measurements $\mathbf{y}$. We produce M quantizers to transmit the elements of $\mathbf{y}$ such that the decoded message $\hat{\mathbf{y}}$ will minimize the distortion between $\tilde{\mathbf{x}} = G(\mathbf{y}, \mathbf{\Phi})$ and $\hat{\mathbf{x}} = G(\hat{\mathbf{y}}, \mathbf{\Phi})$ for a total rate R. Note G can be visualized as a set of N scalar functions $\hat{\mathbf{x}}_\mathbf{j} = G^{(j)}(\hat{\mathbf{y}}, \mathbf{\Phi})$ that are identical in distribution due to symmetry in the randomness of $\mathbf{\Phi}$. Since the sparse input signal is assumed to have uniformly distributed sparsity and $\mathbf{\Phi}$ distributes energy uniformly to all measurements $\mathbf{y_i}$ in expectation, we argue by symmetry that each measurement is allotted the same number of bits and that every measurement's quantizer is the same. Moroever, since the functions representing the reconstruction are identical, we argue using (4) that the overall sensitivity $\gamma_{\text{cs}}(\cdot)$ is the same as the sensitivity of any $G^{(j)}(\hat{\mathbf{y}}, \mathbf{\Phi})$. Computing (2) yields the point density $\lambda_{\text{cs}}(\cdot)$.

This is when the homotopy continuation method becomes extremely useful. For a given realization of $\mathbf{\Phi}$ and $\boldsymbol{\eta}$, we can use HC to determine how many elements in the reconstruction are nonzero for μ^*, denoted J_{μ^*}. Equation (6) is then used to find $\partial G^{(j)}(y, \Phi)/\partial y_i$, which is needed to compute $\gamma_{\text{cs}}(\cdot)$. To simplify our notation, let $A = \Phi_{J_{\mu^*}}$. The resulting differentials can be expressed as

$$\frac{\partial G^{(j)}(y, \Phi)}{\partial y_i} = \left[\left(A^T A\right)^{-1} A^T \right]_{ji}. \tag{7}$$

We now present the sensitivity through the following theorem:

Theorem 1 *Let the noise variance be σ^2 and choose an appropriate μ^*. Define $\mathbf{y}_{\setminus \mathbf{i}}$ to be all the elements of a vector $\mathbf{y}$ except $\mathbf{y_i}$. The sensitivity of each element $\mathbf{y_i}$, which is denoted $\gamma_i^{(j)}(t)$, can be written as*

$$\left(E_{\mathbf{\Phi}, \mathbf{y}_{\setminus \mathbf{i}}} \left[\frac{f_{\mathbf{y_i}|\Phi}(t|\Phi)}{f_{\mathbf{y_i}}(t)} \left[\left(A^T A\right)^{-1} A^T \right]_{ji} \mid \mathbf{y_i} = t \right] \right)^{\frac{1}{2}},$$

where A is the submatrix of Φ as described in HC for μ^* and some observation $\mathbf{y}$. Moreover, for any Φ and its corresponding J, $f_{\mathbf{y_i}|\mathbf{\Phi}}(t|\Phi)$ is the convolution cascade of $\{\mathbf{z_j} \sim \mathcal{U}(-\Phi_{ij}, \Phi_{ij})\}$ for $j \in J$. By symmetry arguments, $\gamma_{\text{cs}}(t) = \gamma_i^{(j)}(t)$ for any i and j.

This expectation is difficult to calculate but can be approached through L Monte Carlo trials on $\mathbf{\Phi}$, $\boldsymbol{\eta}$, and $\mathbf{x}$. For each trial, we can compute the partial derivative using (7). We denote the Monte Carlo approximation to that function to be $\gamma_{\text{cs}}^{(L)}(\cdot)$. Its form is

$$\gamma_{\text{cs}}^{(L)}(t) = \frac{1}{L} \sum_{\ell=1}^{L} \left(\frac{f_{\mathbf{y_i}|\mathbf{\Phi}}(t|\Phi_\ell)}{f_{\mathbf{y_i}}(t)} \left[\left(A_\ell^T A_\ell\right)^{-1} A_\ell^T \right]_{ji}^2 \right)^{\frac{1}{2}}, \tag{8}$$

with i and j arbitrarily chosen. By the weak law of large numbers, the empirical mean of L realizations of the random parameters should approach the true expectation for L large.

We now substitute (8) into (2) to find the Monte Carlo approximation to the optimal quantizer for compressed sensing. It becomes

$$\lambda_{\text{cs}}^{(L)}(t) = C \left(\gamma_{\text{cs}}^{(L)}(t) f_{\mathbf{y_i}}(t) \right)^{1/3}, \tag{9}$$

for some normalization constant C. Again by the weak law of large numbers, $\lambda_{\text{cs}}^{(L)}(t) \xrightarrow{p} \lambda_{\text{cs}}(t)$ for L large.

5. Experimental Results

We compare the CS-optimized quantizer, called the "sensitive" quantizer, to a uniform quantizer and "ordinary" quantizer $\lambda_{\text{ord}}(t)$ which is optimized for the distribution of $\mathbf{y}$. This means the ordinary quantizer would be best if we want to minimize distortion between $\mathbf{y}$ and $\hat{\mathbf{y}}$, and hence has a flat sensitivity curve over the support of $\mathbf{y}$. The sensitive quantizer $\lambda_{\text{cs}}(t)$ is found using (9) and the uniform quantizer $\lambda_{\text{uni}}(t) = c$, where c is a normalization constant.

Using 1000 Monte Carlo trials, we estimate $\gamma_{\text{cs}}(t)$. The resulting point density functions for the three quantizers are illustrated in Figure 3.

Experimental results are performed on a Matlab testbench. Practical quantizers are designed by extracting codewords

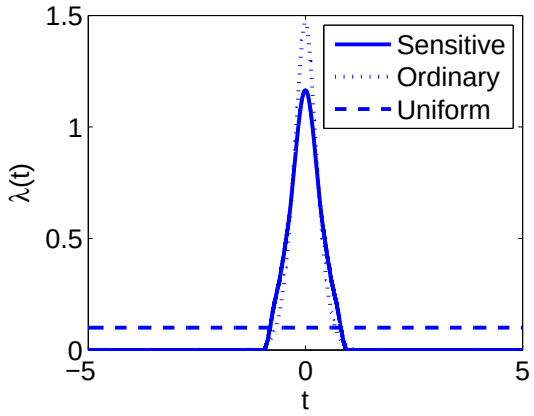

Figure 3: Estimated point density functions $\lambda_{\mathrm{cs}}(t)$, $\lambda_{\mathrm{ord}}(t)$, and $\lambda_{\mathrm{uni}}(t)$ for $(K, M, N) = (5, 71, 100)$.

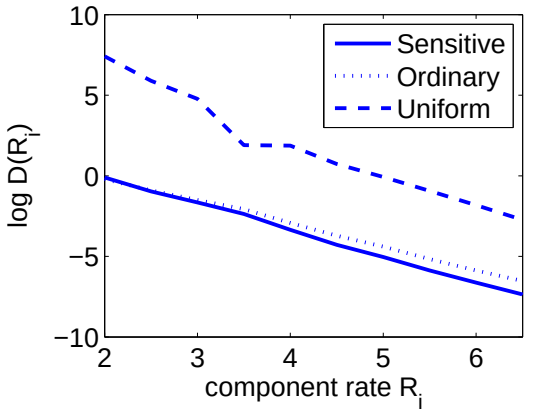

Figure 4: Results for distortion-rate for the three quantizers with $\mu = 0.01$ and $\sigma^2 = 0.3$. We see that the sensitive quantizer has the least distortion.

from the cdf of the normalized point densities. In the approximation, the ith codeword is the point t such that

$$\int_{-\infty}^{t} \lambda_{\mathrm{cs}}(t')dt' = \frac{i - 1/2}{2^{R_i}},$$

where R_i is the rate for each measurement. The partition points are then chosen to be the midpoints between codewords.

We compare the sensitive quantizer to uniform and ordinary quantizers using the parameters $\mu = 0.1$ and $\sigma^2 = 0.3$. Results are shown in Figure 4.

We find the sensitive quantizer performs best in experimental trials for this combination of μ and σ^2 at sufficiently high rates. This makes sense because $\lambda_{\mathrm{cs}}(t)$ is a high-resolution approximation and should not necessarily perform well at very low rates.

6. Conclusion

We present a high-resolution approximation to an optimal quantizer for the storage or transmission of random measurements in a compressed sensing system with lasso reconstruction. Using DFSQ and HC, we find a sensitivity function $\gamma_{\mathrm{cs}}(\cdot)$ that determines the optimal point density function $\lambda_{\mathrm{cs}}(\cdot)$ of such a quantizer. Experimental results show that the operational distortion-rate is best when using this so called "sensitive" quantizer.

We conclude that proper quantization in compressed sensing is not simply a function of the distribution of the random measurements themselves (using either a high-resolution approximation or practical algorithms like Lloyd-Max). Rather, quantization adds a non-constant effect, called functional sensitivity [9], on the distortion between the the lasso reconstructions of the random measurements and its quantized version.

A significant amount of work can still be done in this area. Parallel developments could be made for variable-rate quantizers. Also, this theory can be extended to other probabilistic signal and sensing models, and CS reconstruction methods.

References:

[1] E. J. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inform. Theory*, 52(2):489–509, 2006.

[2] S. Chen, D. L. Donoho, and M. A. Saunders. Atomic decomposition by basis pursuit. *SIAM J. Sci. Comp.*, 20(1):33–61, 1999.

[3] W. Dai, H. Vinh Pham, and O. Milenkovic. Quantized compressive sensing. arXiv:0901.0749v2 [cs.IT]., 2009.

[4] D. L. Donoho. Compressed sensing. *IEEE Trans. Inform. Theory*, 52(4):1289–1306, 2006.

[5] V. K. Goyal, A. K. Fletcher, and S. Rangan. Compressive sampling and lossy compression. *IEEE Sig. Process. Mag.*, 25(2):48–56, 2008.

[6] R. M. Gray and D. L. Neuhoff. Quantization. *IEEE Trans. Inform. Theory*, 44(6):2325–2383, 1998.

[7] L. Jacques, D. K. Hammond, and M. J. Fadili. Dequantized compressed sensing with non-Gaussian constraints. arXiv:0902.2367v2 [math.OC]., 2009.

[8] D. M. Malioutov, M. Cetin, and A. S. Willsky. Homotopy continuation for sparse signal representation. In *Proc. IEEE ICASSP*, pp. 733–736, 2006.

[9] V. Misra, V. K. Goyal, and L. R. Varshney. Distributed functional scalar quantization: High-resolution analysis and extensions. arXiv:0811.3617v1 [cs.IT]., 2008.

[10] R. J. Pai. Nonadaptive lossy encoding of sparse signals. Master's thesis, Massachusetts Inst. of Tech., Cambridge, MA, 2006.

[11] R. Tibshirani. Regression shrinkage and selection via the lasso. *J. Royal Stat. Soc., Ser. B*, 58(1):267–288, 1996.

[12] J. A. Tropp. Greed is good: Algorithmic results for sparse reconstruction. *IEEE Trans. Inform. Theory*, 50(10):2231–2242, 2004.

[13] M. J. Wainwright. Sharp thresholds for high-dimensional and noisy recovery of sparsity. *Department of Statistics, UC Berkley, Tech. Rep 709*, 2006.