

vector \mathbf{y} rather than the vector itself. Hence, the optimization is now

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

such that the maximum number of 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 (piecewise) smooth and monotonic for each \mathbf{y}_i .

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

C3. The joint pdf of the source variables $f_{\mathbf{y}}(\mathbf{y})$ 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

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

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

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

for some normalization constant C . This leads to a total operational distortion-rate

$$D(\{R_i\}) \approx \sum_{i=1}^M 2^{-2R_i} E \left[\frac{\gamma_i^2(\mathbf{y}_i)}{12\lambda_i^2(\mathbf{y}_i)} \right]. \quad (4)$$

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. The theory of DFSQ can be extended to a vector of functions, where $\hat{\mathbf{x}}_j = g^{(j)}(\mathbf{y})$ for $1 \leq j \leq N$. Since the cost function is additive in its components, we can show that the overall sensitivity for each component \mathbf{y}_i is

$$\gamma_i(t) = \frac{1}{N} \sum_{j=1}^N \gamma_i^{(j)}(t), \quad (5)$$

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

Similar results for variable-rate quantizers are also presented in [10]. However, we will only consider the fixed-rate case in this paper.

B. 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 [2], [3], [12]. 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 $\mathbf{u} = \Psi^{-1}x \in \mathbb{R}^N$ contains only K nonzero elements. Consider a length- M measurement vector $\mathbf{y} = \Phi\mathbf{x}$, where $\Phi \in \mathbb{R}^{M \times N}$ with $K < M < N$ is a realization of Φ . The major innovation in CS (for the case of sparse u considered here) is that recovery of x from \mathbf{y} via some computationally-tractable reconstruction method can be guaranteed asymptotically almost surely.

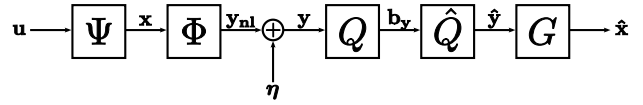


Fig. 1: A compressed sensing model with quantization of noisy measurements \mathbf{y} . The vector \mathbf{y}_{n1} denotes the noiseless random measurements.

Many reconstruction methods have been proposed including a linear program called basis pursuit [13] and greedy algorithms like orthogonal matching pursuit (OMP) [14], [15]. In this paper, we focus on a convex optimization called lasso [11], which takes the form

$$\hat{x} = \arg \min_x (\|\mathbf{y} - \Phi\mathbf{x}\|_2^2 + \mu\|\Psi^{-1}\mathbf{x}\|_1). \quad (6)$$

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 Φ , μ , and the scaling of the smallest entry of u [16]. 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 (6) is affected by quantization of y .

A method for understanding the set of solutions to (6) is the homotopy continuation (HC) method [17]. 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 \mathbf{y} ,

$$2\Phi_{J_{\mu^*}}^T \Phi_{J_{\mu^*}} \hat{x} + \mu^* v = 2\Phi_{J_{\mu^*}}^T \mathbf{y}, \quad (7)$$

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

III. 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 Φ is used to take measurements, and additive Gaussian noise perturbs the resulting signal, meaning the continuous-valued measurement vector is $\mathbf{y} = \Phi\mathbf{x} + \boldsymbol{\eta}$. 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 .

Assume the K -sparse vector \mathbf{x} has random sparsity \mathbf{J} chosen uniformly from all possibilities, and each nonzero component x_i is distributed iid $\mathcal{U}(-1, 1)$. This corresponds to the least-informative prior for bounded and sparse random vectors. Also assume the additive noise vector $\boldsymbol{\eta}$ is distributed iid Gaussian with zero mean and variance σ^2 . Finally, let Φ correspond to random projections such that each column $\phi_j \in \mathbb{R}^M$ has unit energy ($\|\phi_j\|^2 = 1$). The columns of Φ thus form a set of N random vectors chosen uniformly on the unit $(M-1)$ -hypersphere. The cumulative distribution function (cdf) of each matrix element Φ_{ij} is described in the following lemma:

Lemma 1. Assume $\phi_j \in \mathbb{R}^M$ is a random vector uniformly chosen on a unit $(M-1)$ -hypersphere for $M \geq 2$. Then the cdf of each element Φ_{ij} is

$$F_{\Phi_{ij}}(v, M) = \begin{cases} 1 - T(v, M), & 0 \leq v \leq 1; \\ T(-v, M), & -1 \leq v < 0; \\ 0, & \text{otherwise,} \end{cases}$$

where

$$T(v, M) = \frac{\Gamma(\frac{M}{2})}{\sqrt{\pi} \Gamma(\frac{M-1}{2})} \int_0^{\arccos(v)} (\sin \theta)^{M-2} d\theta$$

and $\Gamma(\cdot)$ is the Gamma function.

We find the pdf of Φ_{ij} by differentiating the cdf or use a tractable computational approximation. Since $\mathbf{y} = \Phi \mathbf{x}$,

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

The distribution of each z_{ij} is found using derived distributions. The resulting pdfs can be shown to be iid $f_z(z)$, where \mathbf{z} is a scalar random variable that is identical in distribution to each z_{ij} . The distribution of \mathbf{y}_i is then the $K-1$ convolution cascade of $f_z(z)$ with itself. Thus, $f_y(y)$ is smooth and supported for $\{|y_i| \leq K\}$, satisfying condition C3 for DFSQ. Figure 2 illustrates the distribution of \mathbf{y}_i for a particular choice of parameters.

The reconstruction algorithm G is a function of the measurement vector \mathbf{y} and sampling matrix Φ . We will show that if $G(\mathbf{y}, \Phi)$ is lasso with a proper relaxation variable μ , then conditions C1 and C2 are met. Using HC, we see $G(\mathbf{y}, \Phi)$ 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 (7), 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.

IV. 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 $\tilde{\mathbf{x}}$ from

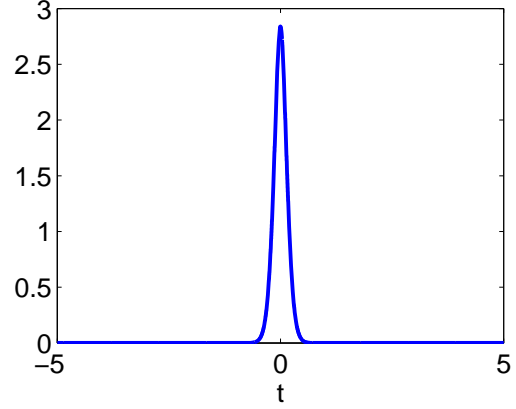


Fig. 2: Distribution $f_{y_i}(t)$ for $(K, M, N) = (5, 71, 100)$. The support of \mathbf{y}_i is the range $[-K, K]$, where K is the sparsity of the input signal. However, the probability is only non-negligible for small \mathbf{y}_i .

the unquantized random measurements \mathbf{y} . We produce M quantizers to transmit the elements of \mathbf{y} such that the quantized measurements $\hat{\mathbf{y}}$ will minimize the distortion between $\tilde{\mathbf{x}} = G(\mathbf{y}, \Phi)$ and $\hat{\mathbf{x}} = G(\hat{\mathbf{y}}, \Phi)$ for a total rate R . Note G can be visualized as a set of N scalar functions $\tilde{\mathbf{x}}_j = G^{(j)}(\hat{\mathbf{y}}, \Phi)$ that are identical in distribution due to the randomness in Φ . Since the sparse input signal is assumed to have uniformly distributed sparsity and Φ distributes energy equally 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. Moreover, again by symmetry in Φ , the functions representing the reconstruction are identical in expectation and we argue using (5) that the overall sensitivity $\gamma_{cs}(\cdot)$ is the same as the sensitivity of any $G^{(j)}(\hat{\mathbf{y}}, \Phi)$. Computing (3) yields the point density $\lambda_{cs}(\cdot)$.

This is when the homotopy continuation method becomes useful. For a given realization of Φ and $\boldsymbol{\eta}$, we can use HC to determine how many elements in the reconstruction are nonzero for μ^* , denoted J_{μ^*} . Equation (7) is then used to find $\partial G^{(j)}(\mathbf{y}, \Phi) / \partial y_i$, which is needed to compute $\gamma_{cs}(\cdot)$. The resulting differentials can be defined as

$$G_i^{(j)}(\mathbf{y}, \Phi) = \frac{\partial G^{(j)}(\mathbf{y}, \Phi)}{\partial y_i} \quad (8)$$

$$= \left[\left(\Phi_{J_{\mu^*}}^T \Phi_{J_{\mu^*}} \right)^{-1} \Phi_{J_{\mu^*}}^T \right]_{ji}. \quad (9)$$

We now present the sensitivity through the following theorem.

Theorem 1. Let the noise variance be σ^2 and choose an appropriate μ^* . Define $\mathbf{y}_{\setminus i}$ 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

$$\left(E_{\Phi} \left[\frac{f_{\mathbf{y}_i | \Phi}(t | \Phi)}{f_{\mathbf{y}_i}(t)} E_{\mathbf{y}_{\setminus i}} \left[\left| G_i^{(j)}(\mathbf{y}, \Phi) \right|^2 \mid \mathbf{y}_i = t, \Phi \right] \right] \right)^{\frac{1}{2}}.$$

For any Φ and its corresponding J found through HC, $f_{\mathbf{y}_i|\Phi}(t|\Phi)$ is the convolution cascade of $\{\mathbf{z}_{ij} \sim \mathcal{U}(-\Phi_{ij}, \Phi_{ij})\}$ for $j \in J$. By symmetry arguments, $\gamma_{cs}(t) = \gamma_i^{(j)}(t)$ for any i and j .

Proof: By symmetry arguments, we can consider any j for the partial derivative in the sensitivity equation without loss of generality. Noting (8), we define

$$\Gamma_i^{(j)}(t, \Phi) = E_{\mathbf{y}_i} \left[\left| G_i^{(j)}(\mathbf{y}, \Phi) \right|^2 \mid \mathbf{y}_i = t \right],$$

and then modify (2) in the following steps:

$$\begin{aligned} \gamma_i^{(j)}(t) &= \left(E \left[\left| G_i^{(j)}(\mathbf{y}, \Phi) \right|^2 \mid \mathbf{y}_i = t \right] \right)^{\frac{1}{2}} \\ &= \left(E_{\Phi} \left[\Gamma_i^{(j)}(t, \Phi) \mid \mathbf{y}_i = t \right] \right)^{\frac{1}{2}} \\ &= \left(\int f_{\Phi|\mathbf{y}_i}(\Phi|t) \Gamma_i^{(j)}(t, \Phi) d\Phi \right)^{\frac{1}{2}} \\ &= \left(E_{\Phi} \left[\frac{f_{\mathbf{y}_i|\Phi}(t|\Phi)}{f_{\mathbf{y}_i}(t)} \Gamma_i^{(j)}(t, \Phi) \right] \right)^{\frac{1}{2}}. \end{aligned}$$

Plugging in (9) will give us the final form of the theorem. Given a realization Φ , $\mathbf{y}_i = \sum_{j \in J} \Phi_{ij} \mathbf{x}_j = \sum_{j \in J} \mathbf{z}_{ij}$, meaning $\mathbf{z}_{ij} \sim \mathcal{U}(-\Phi_{ij}, \Phi_{ij})$. The conditional probability $f_{\mathbf{y}_i|\Phi}(y|\Phi)$ can be found by taking the $K-1$ convolution chain of the set of density functions representing the K nonzero \mathbf{z}_{ij} 's. ■

The expectation in Theorem 1 is difficult to calculate but can be approached through L Monte Carlo trials on Φ , η , and \mathbf{x} . For each trial, we can compute the partial derivative using (9). We denote the Monte Carlo approximation to that function to be $\gamma_{cs}^{(L)}(\cdot)$. Its form is

$$\gamma_{cs}^{(L)}(t) = \frac{1}{L} \sum_{\ell=1}^L \left(\frac{f_{\mathbf{y}_i|\Phi}(t|\Phi_{\ell})}{f_{\mathbf{y}_i}(t)} \left[G_i^{(j)}(y_{\ell}, \Phi_{\ell}) \right]^2 \right)^{\frac{1}{2}}, \quad (10)$$

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 (10) into (3) to find the Monte Carlo approximation to the optimal quantizer for compressed sensing. It becomes

$$\lambda_{cs}^{(L)}(t) = C \left(\gamma_{cs}^{(L)}(t) f_{\mathbf{y}_i}(t) \right)^{1/3}, \quad (11)$$

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

V. EXPERIMENTAL RESULTS

We compare the CS-optimized quantizer, called the ‘‘sensitive’’ quantizer, to a uniform quantizer and ‘‘ordinary’’ quantizer which is optimized for the distribution of \mathbf{y} through (1). 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

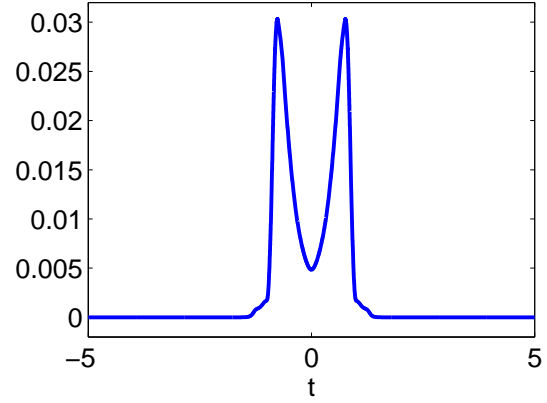


Fig. 3: Estimated sensitivity $\gamma_{cs}(t)$ via Monte Carlo trials and importance sampling for $(K, M, N) = (5, 71, 100)$.

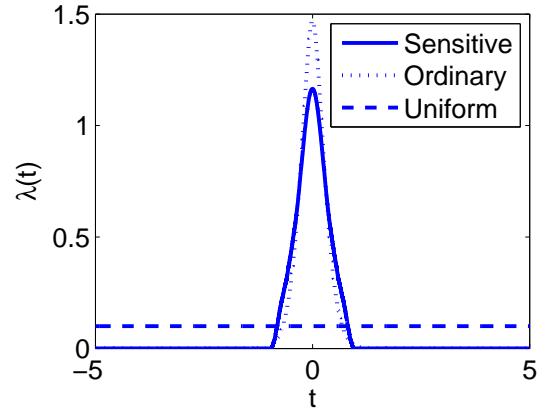


Fig. 4: Estimated point density functions $\lambda_{cs}(t)$, $\lambda_{ord}(t)$, and $\lambda_{uni}(t)$ for $(K, M, N) = (5, 71, 100)$.

$\lambda_{cs}(t)$ is found using (11) and the uniform quantizer $\lambda_{uni}(t)$ is constant and normalized to integrate to 1.

If we restrict ourselves to fixed-rate scalar quantizers, the high-resolution approximation for quantization distortion (4) can be used. The distortion for an arbitrary quantizer $\lambda_q(t)$ with rate R is

$$\begin{aligned} D(R) &\approx 2^{-2R} E \left[\frac{\gamma_{cs}^2(\mathbf{y}_i)}{12\lambda_q^2(\mathbf{y}_i)} \right] \\ &= 2^{-2R} \int \frac{\gamma_{cs}^2(t) f_{\mathbf{y}_i}(t)}{12\lambda_q^2(t)} dt. \end{aligned} \quad (12)$$

Using 1000 Monte Carlo trials, we estimate $\gamma_{cs}(t)$ in Figure 3. Note that the estimate is found through importance sampling since there is low probability of getting samples for large \mathbf{y}_i in Monte Carlo simulations. The sensitivity is symmetric and has peaks away from zero because of the structure in (9). The resulting point density functions for the three quantizers are illustrated in Figure 4.

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

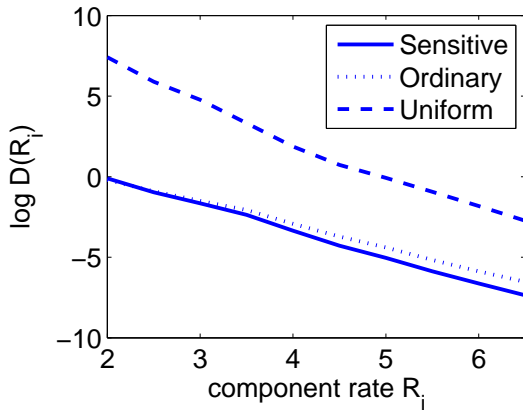


Fig. 5: Results for distortion-rate for the three quantizers with $\sigma^2 = 0.3$ and $\mu = .01$. We see the sensitive quantizer has the least distortion.

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

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

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 $\sigma^2 = 0.3$ and $\mu = 0.1$. Results are shown in Figure 5.

We find the sensitive quantizer performs best in experimental trials for this combination of σ^2 and μ at sufficiently high rates. This makes sense because $\lambda_{cs}(t)$ is a high-resolution approximation and should not necessarily perform well at very low rates. Numerical comparisons between experimental data and the estimated quantization distortion in (12) are similar.

VI. FINAL THOUGHTS AND FUTURE WORK

We present a high-resolution approximation to an optimal quantizer for the storage or transmission of random measurements in a compressed sensing system. We integrate ideas from functional scalar quantization and the homotopy continuation view of lasso to find a sensitivity function $\gamma_{cs}(\cdot)$ that determines the optimal point density function $\lambda_{cs}(\cdot)$ of such a quantizer. Experimental results show that the operational distortion-rate is best when using this so called “sensitive” quantizer.

Our main finding is that proper quantization in compressed sensing is not simply a function of the distribution of random measurements (using either high-resolution approximation or practical algorithms like Lloyd-Max). Rather, quantization adds a non-constant effect, called functional sensitivity, on

the distortion between the lasso reconstructions of random measurements and its quantized version. In the case of lasso reconstruction, the homotopy continuation method allows us to find the sensitivity analytically or through Monte Carlo simulations.

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 that satisfy DFSQ conditions.

ACKNOWLEDGMENT

The authors thank Lav Varshney and Vinith Misra for help in extending the DFSQ theory. We also acknowledge Joel Goodman, Keith Forsythe and Andrew Bolstad for their input.

REFERENCES

- [1] R. M. Gray and D. L. Neuhoff, “Quantization,” *IEEE Trans. Inform. Theory*, vol. 44, pp. 2325–2383, Oct. 1998.
- [2] E. J. Candès, J. Romberg, and T. Tao, “Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information,” *IEEE Trans. Inform. Theory*, vol. 52, pp. 489–509, Feb. 2006.
- [3] D. L. Donoho, “Compressed sensing,” *IEEE Trans. Inform. Theory*, vol. 52, pp. 1289–1306, Apr. 2006.
- [4] E. J. Candès and J. Romberg, “Encoding the ℓ_p ball from limited measurements,” in *Proc. IEEE Data Compression Conf.*, (Snowbird, UT), pp. 33–42, Mar. 2006.
- [5] V. K. Goyal, A. K. Fletcher, and S. Rangan, “Compressive sampling and lossy compression,” *IEEE Sig. Process. Mag.*, vol. 25, pp. 48–56, Mar. 2008.
- [6] W. Dai, H. V. Pham, and O. Milenkovic, “Quantized compressive sensing,” arXiv:0901.0749v2 [cs.IT], Mar. 2009.
- [7] P. T. Boufounos and R. G. Baraniuk, “1-bit compressive sensing,” in *Proc. Conf. on Inform. Sci. & Sys.*, (Princeton, NJ), pp. 16–21, Mar. 2008.
- [8] L. Jacques, D. K. Hammond, and M. J. Fadili, “Dequantized compressed sensing with non-Gaussian constraints,” arXiv:0902.2367v2 [math.OC], Feb. 2009.
- [9] R. J. Pai, “Nonadaptive lossy encoding of sparse signals,” Master’s thesis, Massachusetts Inst. of Tech., Cambridge, MA, Aug. 2006.
- [10] V. Misra, V. K. Goyal, and L. R. Varshney, “Distributed functional scalar quantization: High-resolution analysis and extensions,” arXiv:0811.3617v1 [cs.IT], Nov. 2008.
- [11] R. Tibshirani, “Regression shrinkage and selection via the lasso,” *J. Royal Stat. Soc., Ser. B*, vol. 58, no. 1, pp. 267–288, 1996.
- [12] E. J. Candès and T. Tao, “Near-optimal signal recovery from random projections: Universal encoding strategies?,” *IEEE Trans. Inform. Theory*, vol. 52, pp. 5406–5425, Dec. 2006.
- [13] S. Chen, D. L. Donoho, and M. A. Saunders, “Atomic decomposition by basis pursuit,” *SIAM J. Sci. Comp.*, vol. 20, no. 1, pp. 33–61, 1999.
- [14] S. G. Mallat and Z. Zhang, “Matching pursuits with time-frequency dictionaries,” *IEEE Trans. Signal Process.*, vol. 41, pp. 3397–3415, Dec. 1993.
- [15] J. A. Tropp, “Greed is good: Algorithmic results for sparse reconstruction,” *IEEE Trans. Inform. Theory*, vol. 50, pp. 2231–2242, Oct. 2004.
- [16] M. J. Wainwright, “Sharp thresholds for high-dimensional and noisy recovery of sparsity,” *Department of Statistics, UC Berkeley, Tech. Rep 709*, 2006.
- [17] D. M. Malioutov, M. Cetin, and A. S. Willsky, “Homotopy continuation for sparse signal representation,” in *Proc. IEEE Acoustics, Speech and Sig. Proc. Conf. (Philadelphia, PA)*, pp. 733–736, Mar. 2006.