THE VALUE OF REDUNDANT MEASUREMENT IN COMPRESSED SENSING

Victoria Kostina,[p] *Marco F. Duarte,*[d] *Sina Jafarpour,*[p] *Robert Calderbank* [p,d]

[p] Princeton University, Princeton, NJ 08544
[d] Duke University, Durham, NC 27707

ABSTRACT

The aim of compressed sensing is to recover attributes of sparse signals using very few measurements. Given an overall bit budget for quantization, this paper demonstrates that there is value to redundant measurement. The measurement matrices considered here are required to have the property that signal recovery is still possible even after dropping certain subsets of D measurements. It introduces the concept of a measurement matrix that is weakly democratic in the sense that the amount of information about the signal carried by each of the designated D-subsets is the same. Examples of deterministic measurement matrices that are weakly democratic are constructed by exponentiating codewords from the binary second order Reed Muller code. The value in rejecting D measurements that are on average larger, is to be able to provide a finer grid for vector quantization of the remaining measurements, even after discounting the original budget by the bits used to identify the reject set. Simulation results demonstrate that redundancy improves recovery SNR, sometimes by a wide margin. Optimum performance occurs when a significant fraction of measurements are rejected.

Index Terms— Compressed sensing, quantization, democracy, saturation

1. INTRODUCTION

Democracy is the principle that the individual bits in a coarsely quantized representation of a signal are all given *equal weight* in the signal approximation. It is a mathematical property characteristic of certain sigma-delta converters [1] but not of standard binary or decimal expansions. The principle was introduced by Calderbank and Daubechies [2] who proved that democratic representations cannot achieve the same accuracy as optimal nondemocratic schemes.

In compressed sensing (CS), a signal $\alpha \in \mathbb{R}^C$ is sampled via the linear measurements $y = \Phi\alpha$, where Φ is an $N \times C$ measurement matrix, and $y \in \mathbb{R}^N$ is the vector of samples acquired. Laska et al. [3] formalized the concept of democratic measurement in the CS framework. The principle that each of the N measurements be given equal weight is expressed in the requirement that any $N - D$ measurements should be sufficient to *robustly recover* the sparse input signal. One of the virtues of democracy is that it avoids the situation where removal of some measurements results in high distortion whereas removal of others has negligible effect. Laska et al. used the Restricted Isometry Property (RIP) introduced by Candes and Tao [4] to prove that random Gaussian measurement matrices satisfy this strong notion of democracy.

Laska et al. [3] also investigated how to quantize CS measurements. Their baseline is conventional Shannon-Nyquist uniform sampling where we would scale down the analog signal amplitude (and therefore increase the quantization error) to avoid the gross saturation errors that occur when the signal amplitude exceeds the dynamic range of the quantizer. Their proposed recovery algorithm simply rejects all measurements that fall above the saturation level G of the quantizer. The optimal saturation level is found through experiment to be markedly different from zero. In this case, the reduced distortion on the measurements that are retained outweighs the loss in fidelity from rejecting measurements that saturate.

The analysis given by Laska et al. [3] can be improved in several ways. First, we note that that the columns of a random measurement matrix may be viewed as points randomly drawn an N-dimensional Euclidean sphere. Thus, encoding CS measurements via scalar quantization is suboptimal for the resulting Gaussian distribution. Additionally, since the bitrate per measurement remains constant through the rejection process, the total number of bits used to encode the measurements is dependent on the rejection rate D/N. Finally, there is no accounting for the number of bits needed to encode the indices of the rejected measurements, which is required to perform signal recovery. Our goal in this paper is to address these issues and demonstrate significant further improvements in CS recovery enabled by democratic measurement matrices.

Our proposed problem setup is as follows. We are given a bit budget B and an initial set of N measurements, with the choice to reject D measurements from this set. Our bit budget is divided between (i) a set of bits to encode the set of indices

This work was supported in part by NSF under grant DMS-0914892, by ONR under grant N00014-08-1-1110, and by AFOSR under grants FA9550-09-1-0643 and FA9550-09-1-0551. V. Kostina was supported in part by the Natural Sciences and Engineering Research Council of Canada. M. F. Duarte was supported in part by NSF Supplemental Funding DMS-0439872 to UCLA-IPAM, P.I. R. Caflisch. Email: vkostina@princeton.edu, marco.duarte@duke.edu, sina2jp@princeton.edu, calderbk@princeton.edu.

for the rejected measurements, and (ii) the remaining bits that encode the values of the preserved measurements. Similarly to [3], our objective in rejecting measurements is to be able to place a finer mesh over the measurements that are retained, so that recovery from such quantized measurements can be more accurate. However, as we allow more diversity among the sets of measurements that can be rejected, we are shifting the role of more bits from encoding measurement values to encoding rejected indices, making the quantization mesh coarser. This tradeoff suggests that we weaken the concept of democracy by allowing only specific sets of measurements to be available for rejection.

Random measurement matrices in CS play a role similar to that of random coding in Shannon theory, in that both provide worst case guarantees in the context of an adversarial signal/error model. Random matrices are easy to construct, and can be proven to satisfy RIP with high probability [5]. However, storing the entries of a random matrix may require significant space and there is also no algorithm for efficiently verifying whether a measurement matrix satisfies RIP. Calderbank, Howard and Jafarpour [6] considered deterministic matrices and provided easily verifiable criteria that guarantee the measurement matrix acts like a near isometry on the overwhelming majority of sparse signals. Probability still plays a critical role, but it enters the signal model rather than the construction of the measurement matrix.

While our weaker notion of democracy applies to both random and deterministic matrices, we focus in this paper on deterministic constructions for measurement matrices that are provably democratic. The columns of these matrices may be viewed as points that are uniformly distributed over the Euclidean sphere. Similarly to the random matrix case, the distribution of the resulting CS measurements is essentially Gaussian and scalar quantization is suboptimal. We therefore apply vector quantization to the measurements preserved to further improve the performance of measurement quantization and CS signal recovery. We demonstrate experimentally that measurement rejection achieve significant improvements on CS recovery performance by rejecting structured subsets of measurements from democratic matrices. Furthermore, we show that the additional considerations taken in this paper make the optimal rejection levels higher than those achieved by the framework of [3].

This paper is organized as follows. Section 2 describes our proposed structured rejection algorithm. Section 3 gives examples of deterministic measurement matrices that are weakly democratic. Section 4 presents simulation results showing that our approach outperforms the conventional one in typical scenarios. Section 5 concludes the paper.

2. STRUCTURED REJECTION ALGORITHM

Identifying the entries of y to be rejected in order to achieve the optimum performance is a complex problem in itself. The implementation of the reject-and-quantize idea in the context of vector quantization meets several challenges.

First, we want to reduce the variance of the surviving measurements so that they are easier to quantize. One might choose to reject the largest entries, or the smallest entries, or some of the very large and some of the very small entries. Second, rejection changes the number of bits that are available per measurement. We wish to not only reject entries that are outside the dynamic range of the quantizer, but also to redistribute the bit budget among the surviving measurements. Third, the overhead entailed in encoding the indices for the set of rejected measurements might negate the gain in bit rate due to the bit budget redistribution.

Suppose first that we fix the value of D and reject (suboptimally) the D largest magnitude entries of y. The vector $\tilde{y}$ containing the remaining $N-D$ entries will have, on the average, smaller norm compared to that of a vector with $N-D$ entries drawn at random from y. This allows us to reduce sizes of the individual quantization cells, in turn decreasing the quantization error. The overhead size to encode the set of rejected measurements is $\log_2 \binom{N}{D} \sim D \log_2 \frac{N}{D}$ bits.

To mitigate the overhead, instead of allowing all measurement subsets of size D to be rejected, we fix a collection of rejectable subsets of D measurements. The richer the collection, the more freedom we have as to which measurements to reject, but also the more overhead to describe the rejected subset. To explore this tradeoff, we consider several different collections of rejectable subsets of size D, where we assume that $N = 2^m$ for some positive integer m.

First, we define the collection Ω_s for any integer $1 \leq s \leq m-1$ containing sets of $D = 2^{m-s}$ measurements as follows. Consider the row numbers from 1 to 2^m as the members of a finite field $\mathbb{F}_2^m$. The subsets in Ω_s correspond to all translates of all subspaces of $\mathbb{F}_2^m$ of codimension s. Since there are a total of $\frac{\prod_{k=1}^{s}[2^m-2^{k-1}]}{\prod_{k=1}^{s}[2^s-2^{k-1}]}$ subspaces of codimension s and a total of 2^s translates, it is easy to check that less than $(m-s+2)s$ bits are required to encode a subset from Ω_s.

Next, we define the collection $\Omega_s^{\perp} \subset \Omega_s$ that contains subsets of $\mathbb{F}_2^m$ corresponding to translates of a the subspace spanned by s elements from the generator set $\{2^0, 2^1, 2^2, \ldots, 2^m\}$. Since there are a total of $\binom{m}{s}$ subsets of size s and a total of 2^s translates, number of bits required to index subsets from $\Omega_s^{\perp}$ is $\log_2 2^s \binom{m}{s} \approx s\left(1 + \log_2 \frac{m}{s}\right)$.

The structured rejection algorithm is summarized below.

- Fix a collection Ω of subsets with D elements.
- Reject the measurements for the subset $\omega \in \Omega$ that features the largest norm.
- Quantize the remaining measurements $\tilde{y}$ using a vector quantizer trained accordingly.
- Pass the quantized representation of $\tilde{y}$ and rejected subset ω to the CS recovery algorithm.

- Remove the rows of the measurement matrix indexed by ω and recover α using the quantized version of $\tilde{y}$.

3. DEMOCRATIC MEASUREMENT MATRICES

Given m odd and $0 \leq r \leq (m-1)/2$, the Delsarte-Goethals set $DG(m,r)$ is a vector space of $2^{m(r+2)}$ binary symmetric matrices of size $m \times m$ with the property that the modulo 2 sum of two distinct matrices has rank at least $m-2r$. The matrices in $DG(m,r)$ can be parameterized by elements $a_0, a_1, \ldots, a_r$ from the finite field $\mathbb{F}_2^m$ by setting

$$xP(a_0,\ldots,a_r)y^\top = \mathrm{Tr}\left[a_0xy + \sum_{t=1}^{r} a_t(x^{2^t+1} + xy2^t + 1)\right],$$

where Tr denotes the trace from $\mathbb{F}_2^m$ to $\mathbb{F}_2$, and multiplication in the right side is modulo an irreducible polynomial used to generate the finite field $\mathbb{F}_2^m$. The set $DG(m,(m-1)/2)$ is the vector space of all binary symmetric matrices.

Let $N = 2^m$ and $C = 2^{m(r+2)}$. The Delsarte Goethals frame ($DGF(m,r)$ for short) [6] is an $N \times C$ complex matrix where the rows are indexed by binary m-tuples x, and the columns are indexed by pairs (P,b) where $P \in DG(m,r)$ is a binary symmetric matrix and b is a binary m-tuple. The entry in row $x \in \mathbb{F}_2^m$ and column (P,b) is given by

$$\varphi_{P,b}(x) = \imath^{xPx^\top + 2bx^\top}. \tag{1}$$

Theorem 1. *The $DGF(m,(m-1)/2)$ is weakly democratic; two sets of rows indexed by affine subspaces of the same dimension are equivalent.*

Proof. All arithmetic in the exponent of (1) takes place in $\mathbb{Z}_4$, the ring of integers modulo 4, and that for all binary vectors $w \in \mathbb{F}_2^m$, $(x+2w)P(x+2w)^\top = xPx^\top$. Since

$$\begin{aligned}(x \oplus y)P(x \oplus y)^\top &= (x+y)P(x+y)^\top \\ &= xPx^\top + yPy^\top + 2yPx^\top,\end{aligned}$$

it follows that interchanging rows of the DGF indexed by x and $x \oplus y$ can be realized by interchanging columns (P,b) and $(P, b+yP)$ and multiplying by the phase factor $yPy^\top$. Next, let A be a nonsingular linear transformation. We have

$$(xA)P(xA)^\top = x(APA^\top)x^\top = xP_Ax^\top + 2d_{Q_A}x^\top,$$

where P_A, Q_A are binary symmetric matrices such that

$$APA^\top = P_A + 2Q_A \pmod 4,$$

and d_{Q_A} is the diagonal of Q_A. It follows that the permutation of DGF rows $x \to xA$ can be realized as the permutation of columns $(P,b) \to (P_A, d_{Q_A} \oplus b \oplus bA^\top)$. Thus, the measurements $x \in \Omega$ are equivalent to the measurements in the set $x' \in \Omega A \oplus y$ for a signal f' obtained from a modulation and permutation of the original signal f. These two operations preserve sparsity, and the corrections can be implemented during signal recovery. □

It can also be shown that the $DGF(m,0)$ is weakly democratic in that two sets of rows indexed by affine subspaces of codimension 1 are equivalent.

4. EXPERIMENTS

We now evaluate the CS recovery performance of the structured rejection algorithm described in Section 2 via a suite of simulations and compare it to that of the approach in [3].

Our measurement matrix Φ is the real matrix with $N = 2^{m+1}$ rows and $C = 2^{2m}$ columns obtained from the $DGF(m,0)$ via Gray mapping, i.e., doubling the number of rows and storing the real part of the original matrix in the upper half and the imaginary part in the lower half.

To perform vector quantization, we used the scalar vector quantizer (SVQ) [7] trained as follows. For each collection Ω of rejectable subsets of D measurements considered, we generated a sequence of 200 $k-$sparse $\alpha \in \mathbb{R}^C$ with standard Gaussian nonzero entries. For each α we calculated N measurements as $y = \Phi\alpha$ and rejected the entries of y indexed by the elements of the largest norm subset from to Ω. A maximum of 10 iterations of the two-step training algorithm from [7] were performed to fit the quantizer to the distribution of surviving measurements. We also performed uniform scalar quantization (USQ) as a baseline [3].

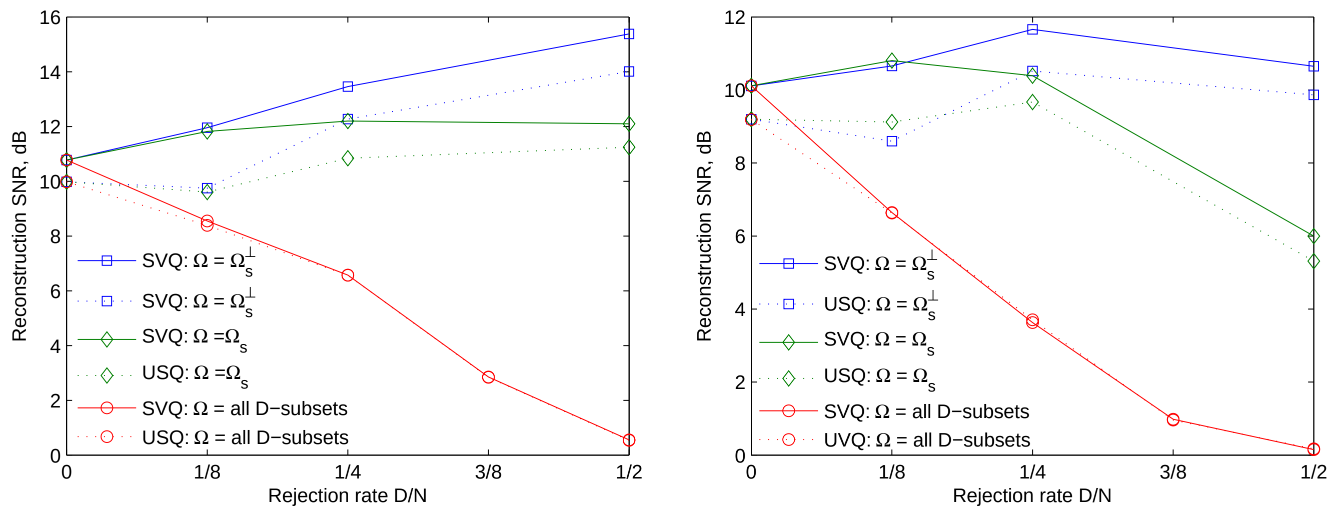
To evaluate the performance of the new approach, 100 measurement vectors y were generated and quantized in the same manner. Basis pursuit denoising [8] was then used to obtain the estimate of the signal $\hat{\alpha}$ using the SPGL1 toolbox [9]. As in [3], the recovery SNR served as the performance measure:

$$SNR = 20\log_{10}\frac{\|\alpha\|_2}{\|\alpha - \hat{\alpha}\|_2}$$

Figure 1 plots the recovery SNR versus the rejection rate D/N of the structured ($\Omega = \Omega_s$ or $\Omega_s^\perp$) and unstructured ($\Omega =$ all subsets of size D) rejection approaches. The bitrate per surviving measurement after taking into account the rejected index coding overhead is plotted in Fig. 2. The rejected index coding overhead of the unstructured rejection is so large that the resulting bitrate per measurement is smaller than that of the conventional approach (i.e. at $D = 0$). As a result, the unstructured rejection approach fails to match the performance of the conventional approach. Structured rejection allows us to boost the bit rate as D increases; the gain due to finer quantization outweighs the performance degradation due to the loss of measurements, as shown in Fig. 1. Once the number of rejected measurements D becomes large, the tradeoff is broken and the improved quantization accuracy cannot balance the degraded recovery performance due to undermeasurement. The figure shows that the tradeoff's optimal rejection rate D/N is smaller as the measurement ratio N/C decreases. The results also show that the smaller collection $\Omega_s^\perp$ outperforms Ω_s, suggesting that having a smaller collection Ω suffices in this case.

5. CONCLUSION

There is value to redundant measurement in compressed sensing. With a fixed total bit budget, rejecting D measurements allows to provide a finer quantization mesh on the measurements that remain. The result is an improvement in recovery performance: as our simulation results demonstrate, given N measurements, it is better to reject up to half of them and quantize the rest finely rather than keep them all but quantize them coarsely. To realize these gains, weak democracy of the measurement matrix is required, meaning that



(a) (b)

Figure 1: Reconstruction SNR versus fraction of rejected measurements D/N. We use a $DGF(7,0)$ measurement matrix on $K = 20$-sparse signals α with $\mathcal{N}(0,1)$ nonzero entries. The dimensionality of the signal space is (a) $C = 2N$ (a) and (b) $C = 5N$ (C random columns of the measurement matrix are used). We use SVQ with block length 32. We set a total bit budget of $B = 2N$ bits. SVQ consistently outperforms USQ and reaches optimal performance for rejection rates of $1/2$ and $1/4$, respectively.

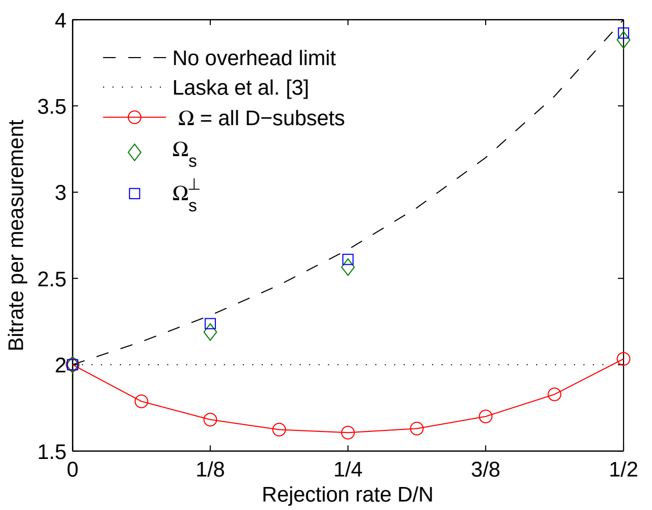


Figure 2: Bitrate per surviving measurement versus D/N for the total bit budget $B = 2N$ bits. $DGF(7,0)$.

recovery must be robust to dropping certain subsets of D measurements. The Delsarte-Goethals frame is an example of a deterministic measurement matrix that is weakly democratic. Future work will focus on characterizing the optimal points on the rejection-structure tradeoff to find suitable collections and sizes of rejection supports Ω.

6. REFERENCES

[1] C.S. Güntürk, *Harmonic analysis of two problems in signal quantization and compression*, Ph.D. thesis, Princeton, 2001.

[2] R Calderbank and I. Daubechies, "The pros and cons of democracy," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1721–1725, 2002.

[3] J. N. Laska, P. T. Boufounos, M. A. Davenport, and R. G. Baraniuk, "Democracy in action: Quantization, saturation, and compressive sensing," *Preprint*, 2009.

[4] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inform. Theory*, vol. 51, pp. 4203–4215, Dec. 2005.

[5] R. G. Baraniuk, M. A. Davenport, R. A. DeVore, and M. B. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approx.*, vol. 28, no. 3, pp. 253–263, Dec. 2008.

[6] R. Calderbank, S. Howard, and S. Jafarpour, "Construction of a large class of deterministic matrices that satisfy a statistical isometry property," *IEEE J. Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 358–374, 2010.

[7] R. Laroia and N. Farvardin, "A structured fixed-rate vector quantizer derived from a variable-length scalar quantizer: Part I - Memoryless sources," *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 851, 1993.

[8] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Review*, vol. 43, no. 1, pp. 129–159, 2001.

[9] E. van den Berg and M. P. Friedlander, "Probing the Pareto frontier for basis pursuit solutions," *SIAM J. on Sci. Comp.*, vol. 31, no. 2, pp. 890–912, Nov. 2008.