## **ROBUST SAMPLING AND RECONSTRUCTION METHODS FOR COMPRESSED SENSING**

Rafael E. Carrillo<sup>\*</sup>, Kenneth E. Barner<sup>\*</sup> and Tuncer C. Aysal<sup>†</sup>

\*University of Delaware, †Cornell University

## ABSTRACT

Recent results in compressed sensing show that a sparse or compressible signal can be reconstructed from a few incoherent measurements. Compressive sensing systems are not immune to noise, which is always present in practical acquisition systems. In this paper we propose robust methods for sampling and reconstructing sparse signals in the presence of impulsive noise. Analysis of the proposed methods demonstrates their robustness under heavy-tailed models. Simulations show that the proposed methods outperform existing compressed sensing techniques in impulsive environments, while having similar performance in light-tailed environments.

*Index Terms*— Sampling methods, signal reconstruction, nonlinear estimation, impulse noise

## 1. INTRODUCTION

Compressed sensing (CS) is a recently introduced novel framework that goes against the traditional data acquisition paradigm. CS demonstrates that a sparse, or compressible, signal can be acquired using a low rate acquisition process that projects the signal onto a small set of vectors incoherent with the sparsity basis [1]. Since noise is always present in real data acquisition systems, a range of different algorithms have been developed that enable approximate reconstruction of sparse signals from noisy compressive measurements [2, 3, 4].

Noise contributions to the overall system can be separated into two models: *observation* noise and *sampling* noise. Observation noise is any perturbation introduced to the underlying signal *prior* to the sampling process, *e.g.*, noisy channel effects or salt and pepper noise in images. The (additive) model of the signal in this case is:  $x = x_0 + w$ , where  $x_0 \in \mathbb{R}^n$  is the original signal and w is the additive noise.

Sampling noise, in contrast, introduces perturbations to the measurements in conjunction with the sampling process, *i.e.*,  $y = y_0(x) + z$  where  $y_0(x) \in \mathbb{R}^m$ , m < n, is the vector of samples, or measurements of x, and z is the corrupting noise, *e.g.*, quantization noise or sensor noise. If we consider  $\Phi$  as the linear measurement process, then the overall noise contribution is  $r = z + \Phi w$ . When w and z are both Gaussian, r is also Gaussian and both problems can be simultaneously addressed by Gaussian–derived reconstruction techniques [3]. In contrast to the typical Gaussian assumption, heavy– tailed processes exhibit very large, or infinite, variance. Existing reconstruction algorithms operating on such processes yield estimates far from the desired original signal. Similarly, linear projections are severely degraded by the presence of large amplitude outliers, *e.g.*, spreading impulses throughout the measurements and precluding fair reconstructions even when robust reconstruction techniques are employed.

This paper develops robust methods for sampling and reconstructing sparse signals in the presence of impulsive noise. We approach the problem from a statistical point of view using robust methods derived from the Generalized Cauchy distribution (GCD). Robust sampling operators, based on the weighted myriad estimators, are proposed. Also introduced are geometric reconstruction algorithm based on  $L_1$  minimization with a Lorentzian norm feasible set constraint. Presented analysis shows the proposed methods yields dramatic performance improvements in heavy-tailed environments and comparable performance in light-tailed environments.

## 2. BACKGROUND AND MOTIVATION

#### 2.1. Compressed Sensing Review

Let  $x \in \mathbb{R}^n$  be a signal that is either *s*-sparse or compressible in some orthogonal basis  $\Psi$ . The signal is *s*-sparse if at most *s* of its coefficients are nonzero, where  $s \ll n$ . The signal is compressible if its ordered set of coefficients decays rapidly and *x* is well-approximated by the first *s* coefficients. In the following we assume, without loss of generality, that  $\Psi = I$ .

Let  $\Phi$  be an  $m \times n$  sensing matrix, m < n, with rows that form a set of vectors incoherent with the sparsity basis [1]. The signal x is measured by  $y = \Phi x$ . It has been shown that a linear program (Basis Pursuit) can recover the original signal, x, from y if the number of measurements is of the order of  $O(s \log(n))$  and the sensing matrices obey the restricted isometry property [1]. The restricted isometry constant definition is as follows.

**Definition 1** For every integer  $1 \leq S \leq m$ , define  $\delta_S$ , the *S*-restricted isometry constant of  $\Phi$ , as the smallest positive quantity such that

$$(1 - \delta_S) \|v\|_2^2 \le \|\Phi v\|_2^2 \le (1 + \delta_S) \|v\|_2^2 \tag{1}$$

This work was supported in part by NSF under grant 0728904.

# for all subsets T of cardinality at most S and vectors v supported on T.

In a realistic scenario, the measurements are noise corrupted and can be modeled as  $y = \Phi x + z$ , where z is zero-mean additive white noise. Basis Pursuit Denoising (BPD) relaxes the requirement that the reconstructed signal exactly explain the measurements [2], solving the optimization problem

$$\min_{x \in \mathbb{R}^n} \|x\|_1 \quad \text{s. t.} \quad \|y - \Phi x\|_2 \le \epsilon, \tag{2}$$

for some small  $\epsilon > 0$ . In [5] it is shown that if  $||z||_2 \le \epsilon$  and  $\delta_{2s} < \sqrt{2} - 1$ , then the reconstructed signal,  $\hat{x}$ , is guaranteed to obey  $||x - \hat{x}||_2 \le C\epsilon$ , where the constant *C* depends on  $\delta_{2s}$ . Other sparse solution determination approaches include greedy algorithms that iteratively construct approximations, Matching pursuit and Orthogonal Matching Pursuit (OMP)[4] being examples. OMP, for instance, solves a LS regression at each iteration, thereby inducing a denoising effect.

Both aforementioned approaches rely on Gaussian statistics, or finite variance, assumptions. In impulsive or heavy– tailed observation noise cases, however, the variance can be very large, or infinite, thus yielding estimates with significant deviations. Impulsive cases, for example, yield large amplitude outlier components spread across all measurements by the (linear) sampling process, causing a breakdown in the recovery process. Even in impulsive sampling cases, where relatively few measurements are corrupted, the gross errors breakdown deriving premises, consequently yielding aberrant reconstructions.

## 2.2. GCD Based Robust Estimation

We propose the use of robust statistics to derive CS methods suitable for impulsive environments. The M–GC estimator is derived in [6] as a Generalized Cauchy density ML estimator, or M–estimator for the cost function  $\rho(x) = \log\{\sigma^p + |x|^p\}$ . This estimator is extendable to a weighted filter structure admitting real–valued weights, thereby providing a robust correlation estimation between two vectors or sequences.

**Definition 2** Let  $x = [x_1, ..., x_n]$  be a vector of observations and  $h = [h_1, ..., h_n]$  a vector of real valued weights. The weighted M–GC estimate is defined as

$$\hat{\theta} = \arg\min_{\theta} \left[ \sum_{i=1}^{N} \log\{\sigma^p + |h_i| |sgn(h_i)x_i - \theta|^p\} \right].$$
(3)

The p = 2 case defines the standard Cauchy distribution optimal weighted *myriad* filter [7]

$$\hat{\theta} = myriad(\sigma; |h_i| \circ sgn(h_i)x_i)|_{i=1}^n.$$
(4)

The M–GC cost function defines a pseudo–norm for  $\mathbb{R}^m$ , and therefore a metric for the space.

**Definition 3** For  $u \in \mathbb{R}^m$ , the  $LL_p$  norm of u is defined as

$$||u||_{LL_{p},\sigma} = \sum_{i=1}^{m} \log\{1 + \sigma^{-p} |u_{i}|^{p}\}, \ \sigma > 0.$$
 (5)

The  $LL_p$  norm doesn't over penalize large deviations, and is therefore a robust metric appropriate for impulsive environments. Importantly, the p = 2 case defines the well known Lorentzian norm with property  $||u||_{LL_2,\sigma} \leq \sigma^{-2} ||u||_2^2$ . In general, the  $LL_p$  norms can be used to define robust regressors for statistical estimation [6].

## 3. ROBUST SAMPLING FUNCTIONS

Of interest here is the design of a robust information operator  $I : \mathbb{R}^n \to \mathbb{R}^m$  that samples *m* pieces of information from *x*, *i.e.*, an operator immune to outlier corruption. We propose the nonlinear weighted myriad filter as such a sampling operator. As myriad filter estimates are asymptotically light–tailed distributed, and the operator converge to a linear filter (sampling process) as a limiting case [7], standard Gaussian/linear sampling–derived reconstruction algorithms can be applied to the (robust) measurements.

**Definition 4** Let  $\Phi \in \mathbb{R}^{m \times n}$  be a random measurement matrix and  $\phi_{ij}$  its ij-th entry. The myriad projections are defined as

$$f_K(\phi_i, x) = a_i \cdot myriad(K; |\phi_{ij}| \circ sgn(\phi_{ij})x_j)|_{j=1}^n \quad (6)$$

where  $a_i = \sum_{j=1}^{n} |\phi_{ij}|$  is a scaling factor introduced to preserve the amplitude of the measurements.

K is known as the linearity parameter and plays an important role in the filtering or sampling process, especially in regards to outlier rejection and asymptotic (linear) behavior [7].

**Property 1** In the limit as  $K \to \infty$ , the weighted myriad measurement reduces to a linear projection on to  $\phi_i$ , i.e.,

$$\lim_{K \to \infty} f_K(\phi_i, x) = \sum_{j=1}^n \phi_{ij} x_j.$$
(7)

The parameter K controls a tradeoff between robust performance and linear fidelity. That is, large K values lead to highly linear projections, but at the cost of less resilience to outliers. Conversely, small K values yield robust operation, but a measurement process that, especially as  $K \rightarrow 0$ , is highly nonlinear. Experimental results show that good performance is achieved in both Gaussian and impulsive environments by a linearity parameter, which is fundamentally related to the scale of the process, set to a value in the neighborhood of the process Median Absolute Deviation (MAD). An observation to make is that when the signal is sparse in the canonical basis and sparse–like impulsive noise is added directly, the signal and noise become undistinguishable, unless the noise has significantly larger amplitude.

#### 4. ROBUST RECONSTRUCTION ALGORITHMS

Consider next the case of sampling noise. The objective in this case is the design of robust reconstruction algorithms,  $A: \mathbb{R}^m \to \mathbb{R}^n$ , that that produce fair signal reconstructions from a small set of measurements. Let  $x_o \in \mathbb{R}^n$  be an *s*sparse signal and  $\Phi \in \mathbb{R}^{m \times n}$  a measurement matrix. The measurement model is  $y = \Phi x_o + z$ , where *z* is white additive noise with impulsive behavior. To estimate  $x_0$  from *y* we propose the following non-linear optimization problem:

$$\min_{x \in \mathbb{D}^n} \|x\|_1 \text{ s. t. } \|y - \Phi x\|_{LL_2, \gamma} \le \epsilon.$$
(8)

The following result presents an upper bound for the reconstruction error of the proposed estimator.

**Lemma 1** Assume  $\delta_{2s} < \sqrt{2} - 1$ . Then for any signal  $x_o$  such that  $|T_0| \leq s$ , where  $T_0 = supp(x_o)$ , and observation noise z with  $||z||_{LL_2,\gamma} \leq \epsilon$ , the solution to (8), denoted as  $x^*$ , obeys the following bound

$$\|x^* - x_o\|_2 \le C_s \cdot \gamma \cdot \sqrt{m(e^{2\epsilon} - 1)},\tag{9}$$

where the constant  $C_s$  depends only on  $\delta_{2s}$ .

*Proof outline:* Proof of lemma 1 follows directly from establishing a relationship between the  $LL_2$  norm and the  $L_2$  norm and the results of Theorem 1.2 in [5].

Notably,  $\gamma$  controls the robustness of the employed norm and  $\epsilon$  the radius of the feasibility set  $LL_2$  ball. Assuming a standard Cauchy model for the noise, with scale parameter  $\sigma$ ,  $\mathbb{E}||z||_{LL_2,\gamma} = m\mathbb{E}\log\{1+\gamma^{-2}z_i^2\} = 2m\log(1+\gamma^{-1}\sigma)$ . We use this value for  $\epsilon$  and set  $\gamma$  as MAD(y).

Debiasing is achieved through robust regression on a subset of indexes of  $\hat{x}$  using the Lorentzian norm. The subset is defined as  $I = \{i : |\hat{x}_i| > \alpha\}, \alpha = \lambda \max_i |\hat{x}_i|$ , where  $0 < \lambda < 1$ . Thus  $\tilde{x}_I \in \mathbb{R}^d$  is defined as

$$\tilde{x}_I = \arg\min_{x \in \mathbb{R}^d} \|y - \Phi_I x\|_{LL_2,\sigma}$$
(10)

where d = |I|. The final reconstruction, after the regression  $(\tilde{x})$  is defined as  $\tilde{x}_I$  for indexes in the subset I and zero outside I. This algorithm is referred to as Lorentzian BP.

## 5. EXPERIMENTAL RESULTS

Numerical experiments illustrate the effectiveness of myriad measurements and Lorentzian BP as robust techniques for CS. All experiments utilize synthetic *s*-sparse signals in a Hadamard basis, with s = 8 and n = 1024. The nonzero coefficients have equal amplitude, equiprobable sign, randomly chosen position, and average power fixed to 0.78. Gaussian sensing matrices are employed with m = 128. Two hundred repetitions of each experiment are averaged and reconstruction SNR is used as the performance measure.

### 5.1. Observation Noise Results

Consider first experiments validating the use of myriad projections as robust sampling functions. We start with an example of a single impulse added to the original signal, thus showing the outlier rejection property of myriad measurements. The amplitude of the impulse is  $10^6$  and the reconstruction is performed using OMP for both linear and myriad projections. The reconstruction SNR is -81.61dB for the linear projections and 34.30dB for myriad projections (K = MAD(x) = 6.4). Results are summarized in Fig. 1 (Left).

Next, we compare myriad measurements, with varying linearity parameter K, to linear measurements in the noiseless case. BP and OMP reconstructions are used for both myriad and linear measurements. Results are summarized in Fig. 1 (Middle). Notably, myriad measurements yield fair reconstructs in the noiseless case (as  $K \to \infty$ ). Having established the applicability of myriad measurements in even noiseless cases, we now address the more demanding heavytailed environments. Specifically, Fig. 1 (Right) presents results for  $\alpha$ -stable noise corruption, with different values of  $\alpha$ and fixed scale parameter  $\gamma = 0.1$ . As  $K \to 0$ , performance degrades because myriad projections tend to a selection type estimator. As K increases, linearity and performance increase until maximum performance is achieved. Beyond this point, performance decreases due to loss of robustness. Note that the optimal K is largely independent of  $\alpha$ , with  $\alpha = 2$  representing the Gaussian special case in which performance is independent of K beyond the fixed minimum point.

#### 5.2. Sampling Noise Results

Lorentzian BP is derived from Cauchy statistics. We therefore present experiments comparing Lorentzian BP with BPD and OMP under Cauchy sampling noise. The Cauchy scale parameter,  $\sigma$ , is varied from  $10^{-3}$  to 10, resulting in a variation of the geometrical SNR [7] from 28.9321 dB to -11.0679 dB. As expected, Lorentzian BP outperforms both BPD and OMP since it is optimal for Cauchy statistics, Fig. 2 (Left). As perhaps a more realistic scenario, consider a mixed noise environment, e.g., p-Gaussian. We set the Gaussian component variance to  $\sigma^2 = 10^{-2}$ , resulting in an SNR of 18.9321 dB when p = 0. The amplitude of the outliers is set as  $\delta = 10^3$ and p is varied from  $10^{-3}$  to 0.5, Fig. 2 (Middle). The results demonstrate that Lorentzian BP outperforms BPD and OMP. Moreover, the Lorentzian BP results are stable over a range of contamination factors p, including contaminations up to 5% of the measurements, making Lorentzian BP a desirable method when measurements are lost or erased.

The final experiment explores the behavior of Lorentzian BP in very impulsive environments. We compare again against BPD and OMP, this time with  $\alpha$ -Stable sampling noise. The scale parameter of the noise is set as  $\sigma = 0.1$  for all cases and the tail parameter,  $\alpha$ , is varied from 0.2 to 2, *i.e.*, very impulsive to the Gaussian case, Fig. 2 (Right). For small values of



Fig. 1. Comparison of linear and myriad projections. L: Single outlier, M: Noiseless case, R:  $\alpha$ -S environment.



Fig. 2. Comparison of Lorentzian BP in impulsive environments. L: Cauchy, M: Contaminated Gaussian, R:  $\alpha$ -Stable.

 $\alpha$ , all methods perform poorly, with Lorentzian BP yielding the most acceptable results. Beyond  $\alpha = 0.8$ , Lorentzian BP produces faithful reconstructions with a SNR greater than 20 dB, and often 30 dB greater than BPD and OMP results. It is of notice that when  $\alpha = 2$  (Gaussian case) the performance of Lorentzian BP is comparable with that of BPD and OMP.

#### 6. CONCLUSIONS

This paper presents robust sampling and reconstruction methods for sparse signals in impulsive environments. Myriad projections are proposed as sampling operators to address problems with impulsive observation noise. Properties of the proposed sampling function are analyzed, and it is noted that reconstruction performance depends on a linearity parameter, K, which can be adapted to the signal and noise environment. Importantly, myriad projections can be used with standard Gaussian-derived reconstruction algorithms. To address the problem of heavy-tailed sampling noise, Lorentzian basis pursuit is proposed. A reconstruction bound is derived that depends on the noise strength and a tunable parameter of the Lorentzian norm. Methods to estimate the adjustable parameters in the sampling functions and reconstruction algorithms are proposed, although computation of their optimal values remains an open question. Thus Myriad projections and Lorentzian BP offer a robust framework for CS in heavytailed environments, with performance comparable to existing methods in less demanding light-tailed environments.

#### 7. REFERENCES

- [1] E. J. Candes, "Compressive sampling," in *Proceedings*, *Int. Congress of Mathematics*, Madrid, Spain, Aug. 2006.
- [2] E. J. Candes, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on pure and applied mathematics*, vol. 59, no. 8, pp. 1207–1223, Aug. 2006.
- [3] J. Haupt and R. Nowak, "Signal reconstruction from noisy random projections," *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 4036–4048, Sept. 2006.
- [4] J. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.
- [5] E. J. Candes, "The restricted isometry property and its implications for compressed sensing," *Compte Rendus de l'Academie des Sciences, Paris, Series I*, pp. 589–593, 2008.
- [6] R.E. Carrillo, T. C. Aysal, and K. E. Barner, "Generalized cauchy distribution based robust estimation," in *Proceed*ings, IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, Las Vegas, Nevada, Apr. 2008.
- [7] G. R. Arce, Nonlinear Signal Processing: A Statistical Approach, John Wiley & Sons, Inc., 2005.