AN OVERVIEW OF ROBUST COMPRESSIVE SENSING OF SPARSE SIGNALS IN IMPULSIVE NOISE

Ana B. Ramirez[*]*, Rafael E. Carrillo*[†]*, Gonzalo Arce*[‡]*, Kenneth E. Barner*[‡] *and Brian Sadler*[§]

[*] Universidad Industrial de Santander, Bucaramanga, Colombia
[†] Signal Processing Laboratory (LTS5), Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland
[‡] Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19716
[§] US Army Research Laboratory, Adelphi, MD, USA, 20783

ABSTRACT

While compressive sensing (CS) has traditionally relied on ℓ_2 as an error norm, a broad spectrum of applications has emerged where robust estimators are required. Among those, applications where the sampling process is performed in the presence of impulsive noise, or where the sampling of the high-dimensional sparse signals requires the preservation of a distance different than ℓ_2. This article overviews robust sampling and nonlinear reconstruction strategies for sparse signals based on the Cauchy distribution and the Lorentzian norm for the data fidelity. The derived methods outperform existing compressed sensing techniques in impulsive environments, thus offering a robust framework for CS.

Index Terms— Compressed sensing, sampling methods, robust signal reconstruction, nonlinear estimation, impulsive noise

1. INTRODUCTION

The theory of compressive sensing (CS) introduces a signal acquisition and reconstruction framework that goes beyond the traditional Nyquist sampling paradigm [1]. Most CS reconstruction algorithms use the ℓ_2 norm as the metric for the residual error. However, it is well-known that least squares based estimators are highly sensitive to outliers present in the measurement vector, leading to a poor performance when the noise does not follow a Gaussian assumption but, instead, is better characterized by heavier-than-Gaussian tailed distributions [2–5]. If the compressive sampling process has infinite, or even very large variance, the reconstructed signal is far from the desired original signal.

Recent works have begun to address the reconstruction of sparse signals whose measurements were acquired such that some metric, different than ℓ_2, is preserved, or when the signals were acquired in the presence of impulsive noise. Popilka *et al.,* propose a reconstruction strategy based on the sparsity of the measurement error pattern to estimate first the error, and then estimate the true signal, in an iterative process [6]. A similar approach is followed in [7] and [8]. These works assume a sparse error and estimate both signal and error at the same stage using a modified ℓ_1 minimization criterion. A drawback of these approaches is that the reconstruction relies on the error sparsity to first estimate the error, but if the sparsity condition is not met, the performance of the algorithm degrades. Sparse models have also been used to address the robust regression problem in [9–11].

Approaches based on robust statistics have also been proposed. Carrillo *et al.,* propose reconstruction approaches based on the Lorentzian norm as data fidelity term [12, 13]. In addition, Ramirez *et al.,* develop an iterative algorithm to solve a Lorentzian ℓ_0-regularized cost function using iterative weighted myriad filters [14]. A similar approach is used in [15] by solving an ℓ_0-regularized least absolute deviation regression problem yielding an iterative weighted median algorithm. Greedy approaches have also been proposed based on M-estimators [16, 17].

Notably, there exists a broad spectrum of applications where practice has shown non-Gaussian, heavy-tailed processes emerge. Examples of such applications are: wireless communications, teletraffic, hydrology, geology, atmospheric noise, economics and image and video processing (see [3] and the references therein). Thus, the motivation is clear for developing robust CS techniques that address these challenging environments. This article covers robust sampling (Section 3) and reconstruction strategies (Section 4) for sparse signals in such environments. We approach the problem from a statistical point of view and review nonlinear methods based on robust statistics [2], specifically methods derived from the Cauchy distribution, that overcome the limitations of traditional linear signal processing methods in the presence of impulsive noise [3].

2. THEORETICAL BACKGROUND

2.1. Compressive Sensing

Let $\mathbf{x} \in \mathbb{R}^n$ be a signal that is either s-sparse or compressible in some representation basis Ψ such that $\mathbf{x} = \Psi\alpha$, where $\alpha \in \mathbb{R}^n$ is the vector of coefficients. Let Φ 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 $\mathbf{x}$ is measured by $\mathbf{y} = \Phi\mathbf{x}$. Setting $\Xi = \Phi\Psi$, the measurement vector becomes $\mathbf{y} = \Xi\alpha$. In the following we assume, without loss of generality, that $\Psi = \mathsf{I}$, the canonical basis for $\mathbb{R}^n$, thus $\mathbf{x} = \alpha$. It has been shown that a convex program (Basis Pursuit) can recover the original signal, $\mathbf{x}$, from $\mathbf{y}$ if the sensing matrices obey the restricted isometry property (RIP) [18]. The restricted isometry constant is defined as follows.

Definition 1 *The s-restricted isometry constant of Φ, δ_s, is defined as the smallest positive quantity such that*

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

holds for all $\mathbf{v} \in \Omega_s$, where $\Omega_s = \{\mathbf{v} \in \mathbb{R}^n \mid \|\mathbf{v}\|_0 \le s\}$. A matrix Φ is said to satisfy the RIP of order s if $\delta_s \in (0, 1)$.

Basically, the RIP dictates that every set of columns of Φ with cardinality smaller than s approximately behaves like an orthonormal system. It has been shown that random matrices with Gaussian or sub-Gaussian entries meet the RIP with high probability provided that $m = O(s \log(n))$ [19].

In a realistic scenario, the measurements are corrupted by noise and can be modeled as $\mathbf{y} = \Phi\mathbf{x} + \mathbf{z}$, where $\mathbf{z}$ is zero-mean additive white noise. Basis Pursuit Denoising (BPD) relaxes the requirement that the reconstructed signal exactly explain the measurements [1], solving the convex problem

$$\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{x}\|_1 \text{ subject to } \|\mathbf{y} - \Phi\mathbf{x}\|_2 \le \epsilon, \tag{1}$$

for some small $\epsilon > 0$. Candès shows in [18] that if $\|\mathbf{z}\|_2 \le \epsilon$ and $\delta_{2s} < \sqrt{2} - 1$, then the reconstructed signal, $\hat{\mathbf{x}}$, is guaranteed to obey $\|\mathbf{x} - \hat{\mathbf{x}}\|_2 \le C\epsilon$, where the constant C depends on δ_{2s}. Unconstrained variations of (1) are also used in the literature, known as ℓ_1-regularized least squares [20]. Other sparse reconstruction approaches include greedy algorithms that iteratively construct approximations. Orthogonal matching pursuit (OMP) [21], and its relatives [22], and iterative hard thresholding (IHT) [23] are examples of this class.

2.2. Lorentzian Norm

The Lorentzian norm for a vector $\mathbf{u} \in \mathbb{R}^m$ is defined as:

$$\|\mathbf{u}\|_{LL_2,\gamma} = \sum_{i=1}^{m} \log\{1 + \gamma^{-2} u_i^2\}, \quad \gamma > 0. \tag{2}$$

The Lorentzian norm (or LL_2 norm) is not a norm in the strict sense, since it does not meet the positive homogeneity and subadditivity properties. However, it defines a robust metric that does not heavily penalize large deviations, with the robustness depending on the scale parameter γ, thus making it an appropriate metric for impulsive environments (optimal under the Cauchy model) [3, 4]. The Lorentzian norm is everywhere continuous and differentiable, and convex near the

origin behaving as an ℓ_2 cost function for small variations. Further justification for the use of the Lorentzian norm is the existence of logarithmic moments for algebraic-tailed distributions, as second moments are infinite or not defined for such distributions and therefore not an appropriate measure of process strength [3].

3. ROBUST SAMPLING USING CAUCHY RANDOM PROJECTIONS

The interest in Cauchy random projections arises due to applications where ℓ_1-distance preservation is preferred in the reduced space. However, since Cauchy random projections have undefined second-order statistics, the RIP condition given by Definition 1 for random matrices is not satisfied, and thus the large suite of methods developed for CS relying on ℓ_2-norm properties can not be used for distance estimation or sparse signal reconstruction [20, 24].

However, Ramirez *et al.*, proposed in [14] a distance-preserving condition between the geometric mean of the projections in the dimensionality reduced space, and the ℓ_1 norm of the original high-dimensional space satisfied by Cauchy random matrices. More precisely, the definition of the distance preservation condition for Cauchy random matrices is as follows.

Definition 2 *A matrix $\mathsf{R} \in \mathbb{R}^{k \times n}$ whose entries $r_{i,j}$ are i.i.d Cauchy random distributed with location parameter 0 and scale parameter 1 (i.e., $r_{i,j} \sim C(0, 1)$), satisfies the following inequalities*

$$(1 - \eta)\|\mathbf{b}\|_1 \le \|\mathsf{R}\mathbf{b}\|_{gm} \le (1 + \eta)\|\mathbf{b}\|_1$$

with high probability for all s-sparse vectors $\mathbf{b} \in \mathbb{R}^n$, where $\|\mathbf{b}\|_1 = \sum_{i=1}^{n} |b_i|$, and $\|\mathbf{z}\|_{gm} = (\prod_{i=1}^{k} |z_i|)^{1/k}$; provided that $k > C_1(\delta)[\log(2/\epsilon) + s \log(n/s)]$, where $\epsilon > 0$, $C_1(\delta) \sim \frac{1}{\delta^2}$ with $\delta > 0$, and $\eta > \frac{\varepsilon + \delta}{1 - \varepsilon}$ with $\varepsilon > 0$.

3.1. Lorentzian Based Coordinate-Descent Algorithm

If the sampling matrix R obeys the distance preservation condition given in Definition 2, then every $s-$sparse signal has a unique image under R, and every sketch $\mathbf{c} = \mathsf{R}\mathbf{b}$ is different for each s-sparse $\mathbf{b}$. Therefore, given the sketch $\mathbf{c}$, a sparse vector $\mathbf{b}$ can be recovered by solving

$$\min_{\mathbf{b} \in \mathbb{R}^n} \|\mathbf{b}\|_0 \text{ subject to } \mathbf{c} = \mathsf{R}\mathbf{b}. \tag{3}$$

The proposed formulation in (3) can be modified to allow a relaxation in the data fidelity term, and in an unconstrained form. That is,

$$\min_{\mathbf{b} \in \mathbb{R}^n} \|\mathbf{c} - \mathsf{R}\mathbf{b}\|_{LL_2,\gamma} + \lambda\|\mathbf{b}\|_0, \tag{4}$$

where $\|\cdot\|_{LL_2,\gamma}$ is the Lorentzian norm, and λ is a regularization parameter. The $\|\cdot\|_0$-norm encourages sparsity in

the solution and the Lorentzian norm is chosen in robust linear regression problems because it has optimality properties for Cauchy distributed samples [3, 4]. The algorithm used to solve the problem in (4) is based on a coordinate-descent method, where at each iteration all the entries of the sparse vector are held constant except one, which is allowed to vary, and is estimated by a weighted myriad operator [14].

3.2. Computer Simulations and Performance

The following results validate the use of Cauchy projections as robust compressive measurements and the ℓ_0-regularized Lorentzian norm method, denoted as $(\ell_0 - \ell_{LL})$, for signal recovery in the presence of impulsive noise environments. In particular, we use alpha-stable distributions $(0 < \alpha < 2)$ to model impulsive noise. The performance of the $(\ell_0 - \ell_{LL})$ method is compared with those achieved using the ℓ_1-regularized least squares $(\ell_1 - \ell_s)$ algorithm [20], the iterative Weighted Median regression *(Wmedian)* algorithm [15], the OMP algorithm *(OMP)* [21], and the ℓ_1-regularized ℓ_1 minimization algorithm $(\ell_1 - \ell_1)$ [25] .

In order to test the robustness of the proposed algorithm to impulsive noise, an 8-sparse signal of length 400 is generated with non-zero entries randomly placed having amplitudes in the interval $(-5, 5)$. The projected signal with 100 $(25\%N)$ samples is then contaminated with noise obeying a statistical model of a "zero-centered" symmetric alpha-stable distribution with $\alpha = 1.2$. Fig. 1 depicts the curves of the normalized mean square error NMSE (dB) vs Geometric-SNR (dB) from the different reconstruction algorithms. Each point in the curves is obtained by averaging 100 independent realizations of the Cauchy random matrix and the noise.

Note in Fig. 1 that the $(\ell_0 - \ell_{LL})$ method outperforms other algorithms. Furthermore, the approach has approximately 3dB of improvement for different G-SNR values over the *(Wmedian)* regression algorithm, and approximately 7dB of improvement for different G-SNR values over the $\ell_1 - \ell_1$ algorithm.

4. ROBUST RECONSTRUCTION METHODS IN THE PRESENCE OF IMPULSIVE NOISE

In this section we review robust reconstruction methods for sparse signals when the measurements are (possibly) corrupted by impulsive noise. The proposed approaches use the Lorentzian norm as the data fidelity metric. The measurement model is $\mathbf{y} = \Phi\mathbf{x}_0 + \mathbf{z}$, where $\mathbf{z}$ is zero-centered additive white noise.

4.1. Lorentzian Based Basis Pursuit

Using the strong theoretical guarantees of ℓ_1 minimization for sparse recovery of underdetermined systems of equations,
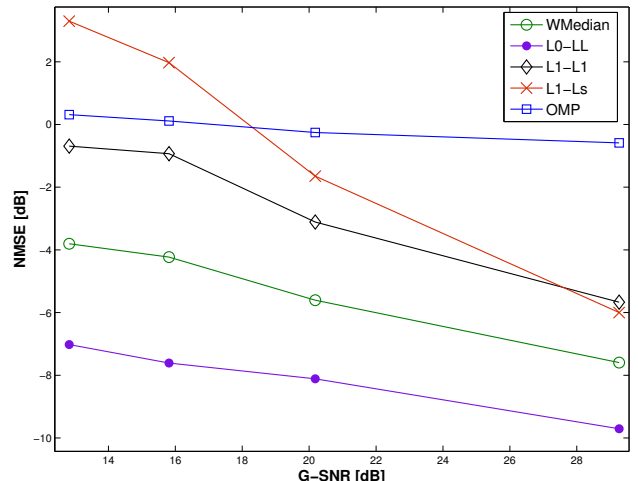

Fig. 1. Normalized MSE reconstruction of sparse signal $(n = 400, s = 8)$ from $(k = 100)$ Cauchy random measurements with α-stable noise $(\alpha = 1.2)$ for various G-SNR values.

Carrillo *et al.,* propose in [12] the following nonconvex constrained optimization problem to estimate a sparse signal from the noisy measurements $\mathbf{y}$:

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{x}\|_1 \text{ subject to } \|\mathbf{y} - \Phi\mathbf{x}\|_{LL_2,\gamma} \leq \rho. \quad (5)$$

The following theorem, proposed in [12], presents an upper bound for the reconstruction error of the proposed estimator in (5).

Theorem 1 *Let Φ be a sensing matrix such that $\delta_{2s} < \sqrt{2} - 1$. Then for any signal $\mathbf{x}_0$ such that $|supp(\mathbf{x}_0)| \leq s$, and observation noise $\mathbf{z}$ with $\|\mathbf{z}\|_{LL_2,\gamma} \leq \rho$, the solution to (5), $\mathbf{x}^*$, obeys the following bound*

$$\|\mathbf{x}_0 - \mathbf{x}^*\|_2 \leq C_s\gamma\sqrt{m(e^\rho - 1)}, \quad (6)$$

where the constant C_s depends only on δ_{2s}.

Theorem 1 shows that the solution to (5) is a sparse signal with an ℓ_2 error that is dependent on logarithmic moments. Note that the dependence on the noise logarithmic moment, rather than its second order moment, makes the formulation in (5) robust and stable to algebraic-tailed and impulsive corrupted samples. The problem in (5) is referred to as Lorentzian BP (LBP). The parameter γ controls the robustness of the employed norm and ρ the radius of the the LL_2 ball thus defining the feasible set. Strategies to estimate both γ and ρ are detailed in [12].

4.2. Lorentzian Based Iterative Hard Thresholding

Even though Lorentzian BP provides a robust CS framework in heavy-tailed environments, numerical algorithms to solve

the proposed optimization problem are slow and complex since the problem in (5) is nonconvex [12]. Therefore, Carrillo and Barner propose in [13] a Lorentzian based iterative hard thresolding (IHT) algorithm. In order to estimate $\mathbf{x}_0$ from $\mathbf{y}$ the following optimization problem is proposed:

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{y} - \Phi\mathbf{x}\|_{LL_2,\gamma} \text{ subject to } \|\mathbf{x}\|_0 \leq s. \quad (7)$$

However, the problem in (7) is non-convex and combinatorial, therefore we derive a suboptimal strategy to estimate $\mathbf{x}_0$ based on the gradient projection algorithm [13]. The proposed strategy is formulated as follows. Let $\mathbf{x}^{(t)}$ denote the solution at iteration time t and set $\mathbf{x}^{(0)}$ to the zero vector. At each iteration t the algorithm computes

$$\mathbf{x}^{(t+1)} = H_s\left(\mathbf{x}^{(t)} + \mu_t\Phi^T W_t(\mathbf{y} - \Phi\mathbf{x}^{(t)})\right) \quad (8)$$

where $H_s(\mathbf{a})$ is the non-linear operator that sets all but the largest (in magnitude) s elements of $\mathbf{a}$ to zero, μ_t is a step size and W_t is an $m \times m$ diagonal matrix defined next. Denote by ϕ_i the i-th row of Φ. Each element on the diagonal is defined as $W_t(i,i) = \gamma^2 \left(\gamma^2 + (y_i - \phi_i^T\mathbf{x}^{(t)})^2\right)^{-1}$, for $i = 1, \ldots, m$.

The algorithm defined by the update in (8) is coined Lorentzian iterative hard thresholding (LIHT). Note that $W_t(i,i) \leq 1$, thus the weights diminish the effect of gross errors by assigning a small weight (close to zero) for large deviations compared to γ, and a weight near one for deviations close to zero. In fact, if W_t is the identity matrix, the algorithm reduces to the ℓ_2 based IHT. The derived algorithm is a fast and simple method that only requires the application of Φ and Φ^T at each iteration.

Although the algorithm is not guaranteed to converge to a global minima of (7), it can be shown that LIHT converges to a local minima [13]. In the following, we show that LIHT has theoretical stability guarantees similar to those of ℓ_2 based IHT. For simplicity of the analysis we set $\mu_t = 1$.

Theorem 2 *Let $\mathbf{x}_0 \in \mathbb{R}^n$. Define $S = supp(\mathbf{x}_0)$, $|S| \leq s$. Suppose $\Phi \in \mathbb{R}^{m \times n}$ meets the RIP of order $3s$ and $\|\Phi\| \leq 1$, where $\|\cdot\|$ denotes the spectral norm. Assume $\mathbf{x}^{(0)} = 0$. Then if $\|\mathbf{z}\|_{LL_2,\gamma} \leq \tau$ and $\delta_{3s} < 1/\sqrt{32}$ the reconstruction error of the LIHT algorithm at iteration t is bounded by*

$$\|\mathbf{x}_0 - \mathbf{x}^{(t)}\|_2 \leq \alpha^t\|\mathbf{x}_0\|_2 + \beta\gamma\sqrt{m(e^\tau - 1)}, \quad (9)$$

where $\alpha = \sqrt{8}\delta_{3s}$ and $\beta = \sqrt{1 + \delta_{2s}}(1 - \alpha^t)(1 - \alpha)^{-1}$.

The results in Theorem 2 can be easily extended to compressible signals using Lemma 6.1 in [22]. Strategies to estimate γ from $\mathbf{y}$ and to adapt the step size μ_t are described in detail in [13].

4.3. Computer Simulations

Numerical experiments that illustrate the effectiveness of the Lorentzian based algorithms in impulsive environments are
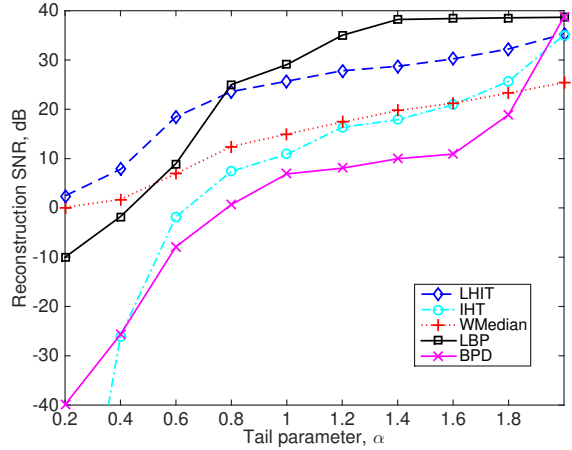
Fig. 2. Reconstruction SNR as a function of the of the tail parameter, α. $\sigma = 0.1$, $s = 8$, $n = 1024$ and $m = 128$.

conducted. 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 $m = 128$. One hundred repetitions of each experiment are averaged and the reconstruction SNR is used as the performance measure.

Weighted median regression *(WMedian)* [15], IHT [23] and BPD [24] are used as benchmarks. Alpha-stable measurement noise is used with the tail parameter, α, varied from 0.2 to 2, *i.e.*, very impulsive compared to the Gaussian case, when α is away from 2. The scale parameter is set as $\sigma = 0.1$, resulting in an SNR of 20.42 dB when $\alpha = 2$. The results are shown in Fig. 2. For small values of α, all methods perform poorly, with LIHT yielding the most acceptable results. Beyond $\alpha = 0.6$, LBP and LIHT produce faithful reconstructions with a SNR greater than 20 dB, outperforming the results of all the other methods. Notice that when $\alpha = 2$ (Gaussian case) the performance of LBP and LIHT is comparable with that of BPD and IHT, which are ℓ_2 based. Of note is that LBP takes 26.31s on average to perform a reconstruction while LIHT takes 0.043s.

5. CONCLUSIONS

This work overviews a set of robust sparse reconstruction algorithms from random measurements, based on the Lorentzian norm as fitting term, having either ℓ_0 or ℓ_1 norms as regularization terms. Numerical results show that the Lorentzian based methods outperform ℓ_2 based methods in impulsive environments, while having similar performance in light-tailed environments.

REFERENCES

[1] E. J. Candès and M. B Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[2] P. J. Huber, *Robust Statistics*, John Wiley & Sons, Inc., 1981.

[3] G. R. Arce, *Nonlinear Signal Processing: A Statistical Approach*, John Wiley & Sons, Inc., 2005.

[4] R. E. Carrillo, T. C. Aysal, and K. E. Barner, "A generalized Cauchy distribution framework for problems requiring robust behavior," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, no. Article ID 312989, pp. 19 pages, 2010.

[5] A. Swami and B. M. Sadler, "On some detection and estimation problems in heavy-tailed noise," *Signal processing, ACM*, vol. 82, no. 12, pp. 1829–1846, Dec. 2002.

[6] B. Popilka, S. Setzer, and G. Steidl, "Signal recovery from incomplete measurements in the presence of outliers," *Inverse Problems and Imaging*, vol. 1, no. 4, pp. 661–672, Nov. 2007.

[7] J. Laska, M. Davenport, and R. G. Baraniuk, "Exact signal recovery from sparsely corrupted measurements through the pursuit of justice," in *Proceedings, IEEE Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, Nov. 2009.

[8] C. Studer, P. Kuppinger, G. Pope, and H. Bolcskei, "Recovery of sparsely corrupted signals," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 3115–3130, 2012.

[9] Y. Jin and B. D. Rao, "Algorithms for robust linear regression by exploiting the connection to sparse signal recovery," in *Proceedings, IEEE ICASSP*, Dallas, TX, Mar. 2010.

[10] G. Mateos and G. B. Giannakis, "Robust nonparametric regression via sparsity control with application to load curve data cleansing," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 1571–1584, June 2012.

[11] K. Mitra, A. Veeraraghavan, and R. Chellappa, "Analysis of sparse regularization based robust regression approaches," *IEEE Transactions on Signal Processing*, vol. 61, no. 5, pp. 1249–1257, May 2013.

[12] R. E. Carrillo, K. E. Barner, and T. C. Aysal, "Robust sampling and reconstruction methods for sparse signals in the presence of impulsive noise," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 392–408, Apr. 2010.

[13] R. E. Carrillo and K. E. Barner, "Lorentzian iterative hard thresholding: Robust compressed sensing with prior information," *IEEE Transactions on Signal Processing*, vol. 61, no. 19, pp. 4822–4833, Oct. 2013.

[14] A. B. Ramirez, G. R. Arce, D. Otero, J. Paredes, and B.M. Sadler, "Reconstruction of sparse signals from l_1 dimensionality-reduced cauchy random-projections," *IEEE Transactions on Signal Processing*, vol. 60, no. 11, pp. 5725–5737, 2012.

[15] J. Paredes and G. R. Arce, "Compressive sensing signal reconstruction by weighted median regression estimates," *IEEE Transactions on Signal Processing*, vol. 59, no. 6, pp. 2585–2601, 2011.

[16] S. A. Razavi, E. Ollila, and V. Koivunen, "Robust greedy algorithms for compressed sensing," in *Proceedings, EUSIPCO*, Bucharest, Rumania, Aug. 2012.

[17] G. Papageorgiou, P. Bouboulis, S. Theodoridis, and K. Themelis, "Robust linear regression analysis: The greedy way," in *Proceedings, EUSIPCO*, Lisbon, Portugal, Sept. 2014.

[18] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *Compte Rendus de l'Academie des Sciences, Paris, Series I*, pp. 589–593, 2008.

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

[20] K. Koh, S. Kim, and S. Boyd, "l1_ls : A matlab solver for large-scale l1-regularized least squares problems," Tech. Report. Stanford University, March 2007.

[21] J. A. 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.

[22] D. Needell and J. A. Tropp, "Cosamp: Iterative signal recovery from incomplete and inaccurate samples," *Applied Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, Apr. 2008.

[23] T. Blumensath and M. E. Davies, "Normalized iterative hard thresholding: guaranteed stability and performance," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 298–309, Apr. 2010.

[24] E. Candès and J. Romberg, "L1magic: Recovery of sparse signals via convex programming," Tech. Report. California Institute of Technology, Oct. 2005.

[25] J. Yang and Y. Zhang, "Alternating direction algorithms for l1-problems in compressive sensing," *SIAM Journal on Scientific Computing*, vol. 33, no. 1, pp. 250–278, 2011.