Robust Sparse Recovery in Impulsive Noise via Continuous Mixed Norm

Amirhossein Javaheri, Hadi Zayyani, *Member, IEEE*, Mario A. T. Figueiredo, *Fellow, IEEE*, and Farrokh Marvasti, *Life Senior Member, IEEE*

***Abstract*—This letter investigates the problem of sparse signal recovery in the presence of additive impulsive noise. The heavy-tailed impulsive noise is well modeled with stable distributions. Since there is no explicit formula for the probability density function of $S\alpha S$ distribution, alternative approximations are used, such as, generalized Gaussian distribution, which imposes ℓ_p-norm fidelity on the residual error. In this letter, we exploit a continuous mixed norm (CMN) for robust sparse recovery instead of ℓ_p-norm. We show that in blind conditions, i.e., in the case where the parameters of the noise distribution are unknown, incorporating CMN can lead to near-optimal recovery. We apply alternating direction method of multipliers for solving the problem induced by utilizing CMN for robust sparse recovery. In this approach, CMN is replaced with a surrogate function and the majorization–minimization technique is incorporated to solve the problem. Simulation results confirm the efficiency of the proposed method compared to some recent algorithms for robust sparse recovery in impulsive noise.**

***Index Terms*—Continuous mixed norm (CMN), impulsive noise, majorization–minimization (MM), robust sparse recovery, symmetric α-Stable ($S\alpha S$) distribution.**

I. Introduction

In a Compressive Sensing (CS) problem, the objective is to reconstruct a sparse signal from lower-dimensional linear measurements. This has found many applications in various areas of signal processing within the last decade [1], [2].

Sometimes, there exists heavy-tailed impulsive noise in the measurements, which strongly degrades the performance of CS reconstruction. For this purpose, there is a class of impulsive noise robust sparse recovery algorithms introduced within different applications in image [3] and speech [4] signal processing, and even wireless sensor networks [5]. There are, in addition, noise-robust regression algorithms proposed for applications in beam-forming [6] and direction of arrival (DOA) estimation [7]. In [8], the impulsive noise is treated as a sparse vector, whereas a joint sparse recovery method is proposed in [9] to recover the original signal. There are several iterative soft (e.g., [10]) and hard-thresholding algorithms, some of which are based on Lorentzian-norm as the fidelity criterion [11]. In recent approaches, ℓ_p-norm fidelity is incorporated for sparse recovery robust to impulsive noise. For example, robust greedy pursuit algorithms based on ℓ_p-correlation in ℓ_p-space are devised in [12]. Moreover, an ℓ_p-ℓ_1 minimization approach named Lp-ADM is proposed in [13], in which ℓ_p-norm is used for error penalization and ℓ_1-norm is employed for sparsity. The ℓ_p-norm is also used via ADMM for robust matrix completion [14].

Since the closed form PDF of the $S\alpha S$ distribution does not exist, other PDFs are employed for the approximation of this heavy-tailed distribution. One alternative [13], is a zero-mean generalized Gaussian distribution (GGD) with shape parameter $p\,(0 < p < 2)$. The exact choice of the shape parameter requires the parameter α to be known. A more general modeling with unknown noise parameters is proposed in [15], but a set of training noise samples are required in that approach. In this letter, we examine the blind recovery problem where the original signal and the parameters of $S\alpha S$ noise distribution are both unknown and we have no training data.

II. Problem Formulation

The problem explored in this letter is to recover a sparse signal given its linear random measurements corrupted with impulsive noise. Suppose $\mathbf{x} \in \mathbb{R}^n$ is the original sparse signal and $\mathbf{A} \in \mathbb{R}^{m\times n}$ is the measurement matrix. The linear measurements $\mathbf{Ax}$, are then, added with impulsive noise modeled as symmetric α-stable i.i.d. components, i.e., each component is a random variable, denoted by N, which follows a $S\alpha S$ distribution ($\beta = 0$) with scale parameter γ and location parameter $\delta = 0$. We assume the PDF of this distribution is modeled with GGD as follows:

$$N \sim \mathcal{S}(\alpha, 0, \gamma, 0), \quad f_N(n) = \frac{\alpha}{2\sigma_n \Gamma\left(\frac{1}{\alpha}\right)} \exp\left(-\frac{|n|^\alpha}{\sigma_n^\alpha}\right) \qquad (1)$$

where σ_n is a constant (proportional to γ) and Γ denotes the gamma function.

The objective is to recover $\mathbf{x}$, given the noisy measurements $\mathbf{y} = \mathbf{Ax} + \mathbf{n} \in \mathbb{R}^m$. In a blind problem, the parameters of the $S\alpha S$ distribution modeling the impulsive noise are unknown. If we treat these unknown parameters, denoted by Θ, as unobserved latent variables, we can apply EM algorithm to find the original sparse signal. The EM algorithm

Manuscript received April 27, 2018; revised May 31, 2018; accepted June 1, 2018. Date of publication June 11, 2018; date of current version June 25, 2018. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Marco Felipe Duarte. *(Corresponding author: Hadi Zayyani.)*

A. Javaheri and F. Marvasti are with the Department of Electrical Engineering, Sharif University of Technology, Tehran 11155-11365, Iran (e-mail: javaheri_amirhossein@ee.sharif.edu; marvasti@sharif.edu).

H. Zayyani is with the Department of Electrical and Computer Engineering, Qom University of Technology (QUT), Qom 12345, Iran (e-mail: zayyani@qut.ac.ir).

M. A. T. Figueiredo is with Instituto de Telecomunicações and Instituto Superior Técnico, Universidade de Lisboa, Lisboa 1649-004, Portugal (e-mail: mario.figueiredo@tecnico.ulisboa.pt).

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2018.2846479

[16] is an iterative method comprising of expectation (E) and maximization (M) steps, i.e., in each iteration, we have to maximize $Q(\mathbf{x}|\mathbf{x}^{(t)}) = \mathbb{E}_{\Theta|\mathbf{y},\mathbf{x}^{(t)}}[\log f(\mathbf{y},\Theta|\mathbf{x}) + \log f(\mathbf{x})]$ with respect to $\mathbf{x}$. For $\mathbf{x}^{(t)}$ and $\mathbf{x}^{(t+1)}$ obtained from the EM algorithm, it can be proven that $f(\mathbf{y}|\mathbf{x}^{(t+1)})f(\mathbf{x}^{(t+1)}) > f(\mathbf{y}|\mathbf{x}^{(t)})f(\mathbf{x}^{(t)})$. This means that the EM, iteratively maximizes the posterior PDF $f(\mathbf{x}|\mathbf{y})$. Now, using Bayes' rule, we have $\mathbb{E}_{\Theta|\mathbf{y},\mathbf{x}^{(t)}}[\log f(\mathbf{y},\Theta|\mathbf{x})] = \mathbb{E}_{\Theta|\mathbf{y},\mathbf{x}^{(t)}}[\log f(\mathbf{y}|\Theta,\mathbf{x})] + \mathbb{E}_{\Theta|\mathbf{y},\mathbf{x}^{(t)}}[\log f(\Theta|\mathbf{x})] = E_1 + E_2$. Assume that the unknown noise parameters (Θ) are independent from $\mathbf{x}$. Therefore, $f(\Theta|\mathbf{x}) = f(\Theta)$ and the term E_2 may be discarded during the maximization. Also, suppose σ_n is given and α is the only unknown parameter of GGD. Then, $\Theta = \alpha$ and the probability density $f(\mathbf{y}|\Theta,\mathbf{x})$ is equal to $f(\mathbf{y}|\alpha,\mathbf{x})$. Hence, using (1) for i.i.d. components of noise signal, we may write $f(\mathbf{y}|\alpha,\mathbf{x}) = \prod_{i=1}^{m} f_N(n_i|\alpha) = \frac{\alpha^m}{(2\sigma_n\Gamma(\frac{1}{\alpha}))^m}\exp\left(-\frac{||\mathbf{n}||_\alpha^\alpha}{\sigma_n^\alpha}\right)$, where $\mathbf{n} = \mathbf{y} - \mathbf{A}\mathbf{x}$. Therefore, the EM optimization problem is finally reduced to

$$\mathbf{x}^{(t+1)} = \underset{\mathbf{x}}{\operatorname{argmax}} -\int_\alpha f(\alpha|\mathbf{y},\mathbf{x}^{(t)})\frac{||\mathbf{y}-\mathbf{A}\mathbf{x}||_\alpha^\alpha}{\sigma_n^\alpha}\mathrm{d}\alpha + \log f(\mathbf{x}) \tag{2}$$

where the latter results from discarding the terms in $Q(\mathbf{x}|\mathbf{x}^{(t)})$, independent from $\mathbf{x}$.

III. Proposed Algorithm

If we ignore $\mathbf{y}$ and $\mathbf{x}^{(t)}$ while integrating over α, we can replace $f(\alpha|\mathbf{y},\mathbf{x}^{(t)})$ with a general function denoted by $\lambda(\alpha)$. Hence, problem (2) can be restated as a general optimization problem as follows:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmax}} -\int_\alpha \lambda(\alpha)\frac{||\mathbf{y}-\mathbf{A}\mathbf{x}||_\alpha^\alpha}{\sigma_n^\alpha}\mathrm{d}\alpha + \log f(\mathbf{x}) \tag{3}$$

Now, suppose $\lambda(\alpha)$ has a support within the range $\alpha \in [p_s, p_f]$. If we also assume $\mathbf{x}$ has a Laplacian prior, it implies $\log f(\mathbf{x}) = -\mu||\mathbf{x}||_1$. Thus, problem (3) may be restated as

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}}\ C(\mathbf{x}) = \ell\left(\frac{\mathbf{y}-\mathbf{A}\mathbf{x}}{\sigma_n}\right) + \mu||\mathbf{x}||_1 \tag{4}$$

where $\ell(\mathbf{v}) = \mathrm{CMN}(\mathbf{v}) = \int_{p_s}^{p_f}\lambda(p)||\mathbf{v}||_p^p\mathrm{d}p$ is what we call the continuous mixed norm (CMN) [17] of vector $\mathbf{v}$. In fact, we have shown that problem (2), under the assumptions discussed in the beginning of this section, is equivalent to a sparse recovery problem in which, CMN is incorporated as fidelity criterion. Now, inserting the auxiliary variable $\mathbf{z} = \frac{1}{\sigma_n}(\mathbf{A}\mathbf{x} - \mathbf{y})$ into problem (4) and using the augmented Lagrangian method (ALM), this problem is transformed into

$$\min_{\mathbf{x},\mathbf{z}} C^{\mathcal{L}}(\mathbf{x},\mathbf{z},\boldsymbol{\eta}) = \ell(\mathbf{z}) + \mu||\mathbf{x}||_1 + \boldsymbol{\eta}^T\left(\frac{\mathbf{A}\mathbf{x}-\mathbf{y}}{\sigma_n} - \mathbf{z}\right) + \frac{\sigma}{2}\left\|\frac{\mathbf{A}\mathbf{x}-\mathbf{y}}{\sigma_n} - \mathbf{z}\right\|^2 \tag{5}$$

where σ is a positive constant. Applying ADMM [18] to solve this problem, we obtain the following alternating steps.

A. $\mathbf{z}$ Update Step

The optimization problem associating with this step is:

$$\mathbf{z}^{(k+1)} = \underset{\mathbf{z}}{\operatorname{argmin}}\ C^{\mathcal{L}}(\mathbf{x}^{(k)},\mathbf{z},\boldsymbol{\eta}^{(k)}) \tag{6}$$

The solution to (6) is found by applying the majorization minimization (MM) technique [19], [20]. To this aim, we need to find a surrogate function for $\ell(\mathbf{z})$ at each iteration.

Lemma 1: For any $0 < p \le q$ and any real z_i', the term $s(z_i,z_i') = |z_i|^q\left(\frac{p}{q}|z_i'|^{p-q}\right) + \left(1-\frac{p}{q}\right)|z_i'|^p$ is a surrogate function, with respect to z_i, for $|z_i|^p$.

Proof: The proof consists in investigating $s(z_i,z_i') - |z_i|^p \ge 0$, using first and second order derivatives with respect to z_i. ■

Assume $q \ge p_f$. Using lemma 1, we may write:

$$\begin{aligned}\ell(\mathbf{z}) &= \sum_i \int_{p_s}^{p_f}\lambda(p)|z_i|^p\mathrm{d}p \\ &\le \sum_i \int_{p_s}^{p_f}\lambda(p)\left(|z_i|^q\left(\frac{p}{q}|z_i'|^{p-q}\right) + \left(1-\frac{p}{q}\right)|z_i'|^p\right)\mathrm{d}p \\ &= \sum_i |z_i|^q\phi_{p_s,p_f}^{(q)}(z_i') + \sum_i \psi_{p_s,p_f}^{(q)}(z_i') = \ell_S(\mathbf{z},\mathbf{z}')\end{aligned} \tag{7}$$

Hence, $\ell_S(\mathbf{z},\mathbf{z}')$ is a surrogate function for $\ell(\mathbf{z})$ ($\ell_S(\mathbf{z},\mathbf{z}) = \ell(\mathbf{z})$ and $\ell_S(\mathbf{z},\mathbf{z}') > \ell(\mathbf{z})$ for any $\mathbf{z} \neq \mathbf{z}'$). Although, there is no closed-form relation for the above integral with $\lambda(\alpha) = f(\alpha|\mathbf{y},\mathbf{x}^{(t)})$, but assuming uniform distribution, we obtain

$$\phi_{p_s,p_f}^{(q)}(z_i') = \frac{|z_i'|^{p_f}(p_f\log|z_i'| - 1) - |z_i'|^{p_s}(p_s\log|z_i'| - 1)}{(p_f - p_s)q|z_i'|^q\log^2|z_i'|} \tag{8}$$

Now let $\mathbf{z}' = \mathbf{z}^{(k)}$; using the MM, it suffices to solve the following optimization problem at each iteration:

$$\begin{aligned}\mathbf{z}^{(k+1)} &= \underset{\mathbf{z}}{\operatorname{argmin}}\ C_S^{\mathcal{L}}(\mathbf{x}^{(k)},\mathbf{z},\boldsymbol{\eta}^k) \\ &= \underset{\mathbf{z}}{\operatorname{argmin}}\ \ell_S(\mathbf{z},\mathbf{z}^{(k)}) + \frac{\sigma}{2}\left\|\frac{\mathbf{A}\mathbf{x}^{(k)}-\mathbf{y}}{\sigma_n} - \mathbf{z} + \frac{\boldsymbol{\eta}^{(k)}}{\sigma}\right\|^2\end{aligned} \tag{9}$$

Hence, depending on the value of q, we have to deal with the following problems:

1) $q = 1$ and $p_s < p_f \le 1$: In this case we have

$$\ell_S(\mathbf{z},\mathbf{z}^{(k)}) = \sum_i |z_i|\phi_{p_s,p_f}^{(1)}(z_i^{(k)}) + \sum_i \psi_{p_s,p_f}^{(1)}(z_i^{(k)})$$

Now, substituting $\phi_{p_s,p_f}^{(1)}(z_i^{(k)})$ from (8) and discarding the term $\sum_i \psi_{p_s,p_f}^{(1)}(z_i^{(k)})$, problem (9) is transformed into an ℓ_1 minimization problem, where the solution is obtained via the soft-thresholding operator $\mathcal{S}_{\mathbf{T}}$ [21] as

$$\mathbf{z}^{(k+1)} = \mathcal{S}_{\mathbf{T}^{(k)}}\left(\frac{\mathbf{A}\mathbf{x}^{(k)}-\mathbf{y}}{\sigma_n} + \frac{\boldsymbol{\eta}^{(k)}}{\sigma}\right) \tag{10}$$

and $\mathbf{T}^{(k)} = \frac{1}{\sigma}[\phi_{p_s,p_f}^{(1)}(z_1^{(k)}),\cdots,\phi_{p_s,p_f}^{(1)}(z_m^{(k)})]^T$.

2) $q = 2$ and $p_s < p_f \le 2$: The choice of $q = 2$ results in a quadratic formulation for $\ell_S(\mathbf{z},\mathbf{z}^{(k)})$, i.e., substituting $\phi_{p_s,p_f}^{(2)}(z_i^{(k)})$ from (8), we have $\ell_S(\mathbf{z},\mathbf{z}^{(k)}) =$

$\mathbf{z}^T \mathbf{W}(\mathbf{z}^{(k)})\mathbf{z}$, where $\mathbf{W}(\mathbf{z}^{(k)}) = \mathrm{diag}(\boldsymbol{\phi}^{(2)}_{p_s,p_f}(\mathbf{z}^{(k)}))$. Hence, problem (9) will have a solution obtained by

$$\mathbf{z}^{(k+1)} = \left(\mathbf{I} + \frac{2}{\sigma}\mathbf{W}(\mathbf{z}^{(k)})\right)^{-1} \left(\frac{\mathbf{A}\mathbf{x}^{(k)} - \mathbf{y}}{\sigma_n} + \frac{\boldsymbol{\eta}^{(k)}}{\sigma}\right) \tag{11}$$

B. $\mathbf{x}$ Update Step

This step involves the following optimization problem:

$$\mathbf{x}^{(k+1)} = \underset{\mathbf{x}}{\mathrm{argmin}}\, C_S^{\mathcal{L}}(\mathbf{x}, \mathbf{z}^{(k+1)}, \boldsymbol{\eta}^{(k)})$$
$$= \underset{\mathbf{x}}{\mathrm{argmin}}\, \frac{\sigma}{2}\left\|\frac{\mathbf{A}\mathbf{x} - \mathbf{y}}{\sigma_n} - \mathbf{z}^{(k+1)} + \frac{\boldsymbol{\eta}^{(k)}}{\sigma}\right\|^2 + \mu||\mathbf{x}||_1 \tag{12}$$

solving this ℓ_1-problem using MM with $\lambda_0 > ||\mathbf{A}||^2/\sigma_n^2$ and $\tilde{\mathbf{x}}^{(k)} = \mathbf{x}^{(k)} - \frac{1}{\lambda_0 \sigma_n}\mathbf{A}^T\left(\frac{\mathbf{A}\mathbf{x}^{(k)} - \mathbf{y}}{\sigma_n} - \mathbf{z}^{(k+1)} + \frac{\boldsymbol{\eta}^{(k)}}{\sigma}\right)$, yields [21]

$$\mathbf{x}^{(k+1)} = \mathcal{S}_{\frac{\mu}{\sigma\lambda_0}}\left(\tilde{\mathbf{x}}^{(k)}\right) \tag{13}$$

C. Multiplier Update Step

The update formula for $\boldsymbol{\eta}$ is given in Algorithm 1.

D. Remark

The weight function $\phi^{(q)}_{p_s,p_f}(z_i')$ given in (8) is undefined when $|z_i'| = 0$. To deal with this problem, a common alternative [22] is to consider a small regularizing parameter ϵ in the definition of CMN, i.e., we let $\mathrm{CMN}(\mathbf{z}) = \int_{p_s}^{p_f} \lambda(p) \sum_i (|z_i| + \epsilon)^p \mathrm{d}p$. The derivation for the ϵ-regularized problem, with approximation, is similar to what is obtained in $\mathbf{z}$-update step in section III, except that we let $|z_i^{(k)}| \leftarrow |z_i^{(k)}| + \epsilon$ in the computation of $\phi^{(q)}_{p_s,p_f}(z_i^{(k)})$. In addition, to increase the rate of convergence of the algorithm, we apply a continuation method on the regularizing parameter μ (as proposed in [23]) with the minimum threshold $\mu_{\min}$. Hence, the modified steps of the proposed algorithm, named CMN-ADM, are given in Algorithm 1.

E. Convergence

Although, the convergence of MM for ℓ_1-problems [20], [23] and ADMM for ℓ_p-problems [13] have been studied in the literature, the proof of convergence for the proposed algorithm is not easy, since it applies MM (instead of exact solution) for mixed ℓ_p-norm minimization within each step of an ADMM algorithm. We can inspire from [24] and [25] for ϵ-regularized ℓ_p-minimization in the single analysis of the $\mathbf{z}$ update step, but the extension to the CMN-ADM algorithm in general, remains an open problem. What we can say is that, according to simulations, it seems that the algorithm is convergent for $p_s, p_f > 1$, i.e., when the cost function in (5) is convex.

IV. Simulation Results

In this section, we conduct experiments to compare the reconstruction quality of our method with some state-of-the-art algorithms for robust sparse recovery. In particular, we use Lp-ADM, YALL1 [26], BP-SEP with ADMM [27] and Huber-FISTA [28].[1] Based on different choices for p_s, p_f and q, we obtain different versions of the proposed algorithm. In the following experiments, we run 3 versions with $(p_s = 0, p_f = 1, q = 1)$, $(p_s = 0, p_f = 1, q = 2)$ and $(p_s = 0, p_f = 2, q = 2)$. We choose $\sigma_n = 1$ and the stopping criterion is set to primal or dual error tolerance of $1e-5$ with 100 maximum iterations. The experiments and results are categorized into the following sub-sections:

Algorithm 1: Proposed Robust CS Algorithm (CNM-ADM).

Set $0 \le p_s < p_f \le q \le 2$, $\mu, \sigma, \lambda_0, \zeta > 0$, $\mu_{\min}, \epsilon \ll 1$.
Initialize $\boldsymbol{\eta}^{(0)} = \mathbf{0}$, $\mathbf{x}^{(0)} = \mathbf{0}$, $\mathbf{z}^{(0)} = -\mathbf{y}$, $k = 0$.
repeat
Update $\mathbf{z}^{(k+1)}$ using (10) or (11) (with $|z_i^{(k)}| \leftarrow |z_i^{(k)}| + \epsilon$)
Update $\mathbf{x}^{(k+1)}$ using (13)
Update $\boldsymbol{\eta}^{(k+1)} = \boldsymbol{\eta}^{(k)} + \sigma\left(\frac{\mathbf{A}\mathbf{x}^{(k+1)} - \mathbf{y}}{\sigma_n} - \mathbf{z}^{(k+1)}\right)$
Update $\mu \leftarrow \max\{\zeta\mu, \mu_{\min}\}$ and $k \leftarrow k+1$
until A stopping criterion is reached

A. Average Preference Ratio

In this scenario, a set of N_t k-sparse vectors of size $n \times 1$ are randomly generated where $N_t = 60$, $n = 128$ and $k = 7$. For each vector $\mathbf{x}_i$, the elements of the support set are chosen uniformly at random and the values are generated according to normal distribution. Each sample vector is compressed via a random Gaussian measurement matrix $\mathbf{A}_i$ of dimensions $m \times n$. We choose $m = 50$. The measurement vector $\mathbf{y}_i = \mathbf{A}_i\mathbf{x}_i$ is then added with $S\alpha S$ noise with distribution parameters α and γ. These values are chosen from $\alpha \in \{0.5, 1, 1.5\}$ and $\gamma \in \{1e-4, 1e-3, 1e-2, 1e-1\}$. The noisy signal is then given to the robust sparse recovery algorithms to obtain an estimate $\hat{\mathbf{x}}_i$. Finally we calculate SNR performance as $\mathrm{SNR} = 20\log_{10}\left(\frac{||\mathbf{x}_i||}{||\hat{\mathbf{x}}_i - \mathbf{x}_i||}\right)$ and average the results over random turns (random $\mathbf{x}_i$ and $\mathbf{A}_i$). For this part, we only compare our results with Lp-ADM algorithm. We choose $\sigma = 1$, $\mu_{\min} = 5e-1$, $\zeta = 0.95$, $\epsilon = 1e-2$ and $\lambda_0 = 2$. We have used the source-code for Lp-ADM with default parameters. Since the problem is blind, we do not know the best choice of p (ℓ_p norm) in Lp-ADM. Thus, to compare the performance of the algorithms, we plot the diagram of average SNR versus p. We then calculate the *preference ratio*, defined as the portion (in %) of the region $p \in (0, 2)$ where the SNR curves corresponding to different versions of our algorithm (these curves are constant lines versus p) lie above the Lp-ADM curve. These regions for each curve are highlighted in the preference diagram depicted in Fig. 1. In fact, this experiment demonstrates the probability that our algorithms outperform Lp-ADM when the problem is blind and p is chosen uniformly at random.

Furthermore, in a blind problem, the regularization path for specifying μ as proposed in [13] is not applicable. In this case, we choose $\mu = \xi||\mathbf{A}^T\mathbf{x}||_\infty$ (as proposed in [23]) with $\xi = 0.1$ for all algorithms. Table I shows the preference ratios for different values of $S\alpha S$ noise parameters. According to this table, the

[1] The source-codes for these algorithms are available at https://github.com/FWen/Lp-Robust-CS.git

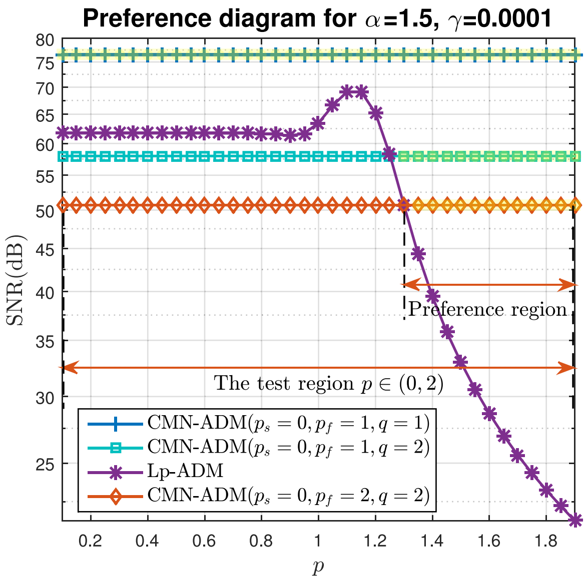


Fig. 1. The preference diagram of SNR versus p (GGD approximation shape parameter) depicting the preference region for proposed robust CS recovery methods (experiment IV-A). The parameters of $S\alpha S$ noise are $\alpha = 1.5$, $\gamma = 1e-4$.

TABLE I
PREFERENCE RATIO (IN %) OF DIFFERENT VERSIONS OF THE PROPOSED ALGORITHM, AS DEFINED IN SECTION IV-A, COMPARED TO LP-ADM

	$\gamma = 1e-4$	$\gamma = 1e-3$	$\gamma = 1e-2$	$\gamma = 1e-1$
$\alpha = 0.5$	100 100 0	100 100 0	100 100 0	100 0 0
$\alpha = 1$	87 100 100	63 87 100	100 100 100	100 100 0
$\alpha = 1.5$	100 100 100	100 100 100	57 78 100	100 100 100

The values are reported in three-tuples corresponding to $(p_s = 0, p_f = 1, q = 1)$, $(p_s = 0, p_f = 1, q = 2)$, and $(p_s = 0, p_f = 2, q = 2)$ versions, respectively.

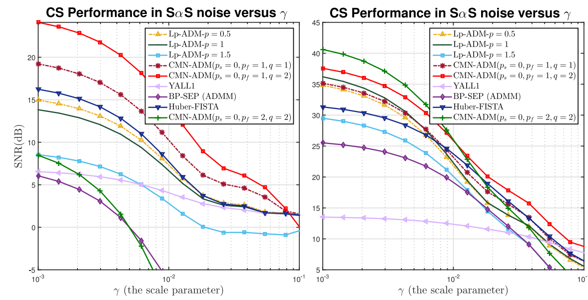
version of the proposed algorithm with $(p_s = 0, p_f = 1, q = 1)$ has best performance in blind settings. Even the worst case corresponding to the version $(p_s = 0, p_f = 2, q = 2)$ which fails at $\alpha = 0.5$, outperforms Lp-ADM in most cases, especially for $\alpha > 0.5$ and $\gamma < 1e-1$.

B. The Influence of Noise Power

In this part, we examine the effect of noise power on the performance of robust sparse recovery algorithms. The settings and the parameters (including μ) for the algorithms are similar to the previous scenario except that α is fixed and the SNR performance is depicted versus the parameter γ, which in some sense specifies the additive noise power. We run Lp-ADM with shape parameter $p \in \{0.5, 1, 1.5\}$ and the results are compared with those of our methods, as well as YALL1, BP-SEP (ADMM) and Huber-FISTA. Fig. 2(a) shows the SNR performance of robust CS algorithms in $S\alpha S$ noise versus the scale parameter γ. In Fig. 2(a) we have chosen $\alpha = 0.5$ and Fig. 2(b) depicts the results for $\alpha = 1$. As shown in these figures, the proposed algorithms clearly outperform competing algorithms specifically for $\alpha = 0.5$.

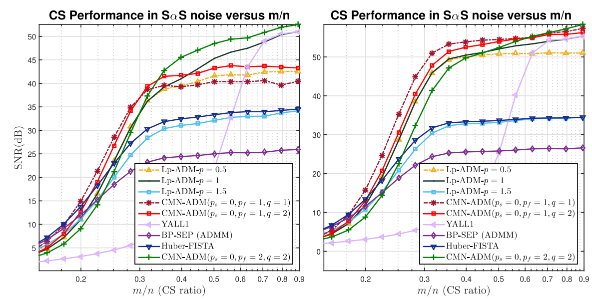
C. The Effect of CS Factor

This part demonstrates the performance of the proposed algorithms in terms of the CS factor, i.e., the ratio m/n where m equals the number of measurements and n denotes the size of the sparse signal. Similar to previous sections (using the same parameters), we generate random 8-sparse vectors of size $n = 128$, but m varies from $\lfloor 0.1n \rfloor$ to $\lfloor 0.9n \rfloor$. The sparse signals are then corrupted with $S\alpha S$ impulsive noise. We consider two cases where in the first, we let $\alpha = 1$ and $\gamma = 1e-3$ and in the second, $\alpha = 1.5$ and $\gamma = 1e-3$ are chosen. The SNR performance of the sparse recovery algorithms is finally depicted versus the ratio m/n. Fig. 3(a) and Fig. 3(b) show the results for the first and the second settings, respectively. For $\alpha = 1$ the proposed algorithm with $p_s = 0, p_f = 2, q = 2$ has the best performance while for $\alpha = 1.5$ the version with $p_s = 0, p_f = 1, q = 1$ yields more robust recovery.



(a) (b)

Fig. 2. SNR performance of robust sparse recovery algorithms versus γ (the unknown scale parameter of $S\alpha S$ impulsive noise) (experiment IV-B). (a) $S\alpha S$ noise with $\alpha = 0.5$ (b) $S\alpha S$ noise with $\alpha = 1$.



(a) (b)

Fig. 3. SNR performance of robust sparse recovery algorithms versus CS factor m/n (experiment IV-C). (a) $S\alpha S$ noise, $\alpha = 1$, $\gamma = 1e-3$ (b) $S\alpha S$ noise, $\alpha = 1.5$, $\gamma = 1e-3$.

V. CONCLUSION

In this letter, we explored the problem of blind sparse signal recovery in the presence of impulsive noise. We modeled the noise signal with symmetric α-stable i.i.d. components and we used GGD to approximate the PDF of the noise distribution. In blind conditions, the parameter α of the $S\alpha S$ noise model is unknown; treating this parameter as latent variable, we applied an EM algorithm to obtain an iterative approach to near optimal recovery of the original signal where in each iteration, an M-step optimization problem is solved. Under some assumptions, this problem was shown to be equivalent to a CS problem in which, a continuous mixed norm is used as fidelity criterion. We employed ADMM with MM technique to iteratively solve the corresponding problem. The performance of the proposed algorithm in blind CS recovery in impulsive noise, was finally examined via simulation results.

REFERENCES

[1] Z. Zhang, Y. Xu, J. Yang, X. Li, and D. Zhang, "A survey of sparse representation: Algorithms and applications," *IEEE Access*, vol. 3, pp. 490–530, 2015.

[2] F. Marvasti *et al.*, "A unified approach to sparse signal processing," *EURASIP J. Adv. Signal Process.*, vol. 44, 2012.

[3] H. Hosseini, F. Hessar, and F. Marvasti, "Real-time impulse noise suppression from images using an efficient weighted-average filtering," *IEEE Signal Process. Lett.*, vol. 22, no. 8, pp. 1050–1054, Aug. 2014.

[4] C. K. W. Wong, F. Marvasti, and W. G. Chambers, "Implementation of recovery of speech with impulsive noise on a DSP chip," *Electron. Lett.*, vol. 31, no. 17, pp. 1412–1413, 1995.

[5] H. Zayyani, M. Korki, and F. Marvasti, "A distributed 1-bit compressed sensing algorithm robust to impulsive noise," *IEEE Commun. Lett.*, vol. 20, no. 6, pp. 1089–7798, 2016.

[6] X. Jiang, W.-J. Zeng, H. C. So, A. M. Zoubir, and T. Kirubarajan, "Beamforming via nonconvex linear regression, *IEEE Trans. Signal Process.*, vol. 64, no. 7, pp. 1714–1728, Apr. 2016.

[7] W.-J. Zeng, H. C. So, and L. Huang, "ℓ_p-MUSIC: Robust direction-of-arrival estimator for impulsive noise environments, *IEEE Trans. Signal Process.*, vol. 61, no. 17, pp. 4296–4308, Sep. 2013.

[8] C. Studer, P. Kuppinger, G. Pope, and H. Bolcskey, "Recovery of sparsely corrupted signals," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 3115–3130, May 2012.

[9] J. Shang, Z. Wang, and Q. Huang, "A robust algorithm for joint sparse recovery in presence of impulsive noise," *IEEE Signal Process. Lett.*, vol. 22, no. 8, pp. 1166–1170, Aug. 2015.

[10] S. Zahedpou, S. Feizi, A. Amini, M. Ferdosizadeh, and Farokh Marvasti, "Impulsive noise cancellation based on soft decision and recursion," *IEEE Trans. Instrum. Meas.*, vol. 58, no. 8, pp. 2780–2790, Aug. 2009.

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

[12] W. J. Zeng, H. C. So, and X. Jiang, "Outlier-robust greedy pursuit algorithms in ℓ_p-space for sparse approximation," *IEEE Trans. Signal Process.*, vol. 64, no. 1, pp. 60–75, Jan. 2016.

[13] F. Wen, P. Liu, Y. Liu, R. C. Qiu, and W. Yu, "Robust sparse recovery in impulsive noise via ℓ_p-ℓ_1 optimization," *IEEE Trans. Signal Process.*, vol. 65, no. 1, pp. 105–118, Jan. 2017.

[14] W.-J. Zeng and H. C. So, "Outlier-robust matrix completion via ℓ_p-minimization, *IEEE Trans. Signal Process.*, vol. 66, no. 5, pp. 1125–1140, Mar. 2018.

[15] R. Hu, Y. Fu, Z. Chen, Y. Xiang, and R. Rong, "Robust sparse signal recovery in the presence of the $S\alpha S$ noise," *IEEE Signal Process. Lett.*, vol. 23, no. 11, pp. 1687–1691, Nov. 2016.

[16] R. J. A. Little and D. B. Rubin, *Statistical Analysis With Missing Data* (Wiley Series in Probability and Statistics). New York, NY, USA: Wiley, 1987.

[17] H. Zayyani, "Continuous mixed p-norm adaptive algorithm for system identification," *IEEE Signal Process. Lett.*, vol. 21, no. 9, pp. 1108–1110, Sep. 2014.

[18] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.

[19] K. Lange, *MM Optimization Algorithms*. Philadelphia, PA, USA: SIAM, 2016.

[20] M. A. T. Figueiredo, J. M. Bioucas-Diaz, and R. D. Nowak, "Majorization-minimization algorithms for wavelet-based image restoration," *IEEE Trans. Image Process.*, vol. 16, no. 12, pp. 2980–2991, Dec. 2007.

[21] I. Duabechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Commun. Pure Appl. Math.*, vol. 57, no. 11, pp. 1413–1457, 2004.

[22] R. Chartrand and W. Yinl, "Iteratively reweighted algorithms for compressive sensing," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Las Vegas, NV, USA, 2008, pp. 3869–3872.

[23] S. J. Wright, R. D. Nowak, and M. A. T. Figueiredo, "Sparse reconstruction by separable approximation," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2479–2493, Jul. 2009.

[24] B. D. Rao and K. Kreutz-Delgado, "An affine scaling methodology for best basis selection," *IEEE Trans. Signal Process.*, vol. 47, no. 1, pp. 187–200, Jan. 1999.

[25] S. Foucart and M. J. Lai, "Sparsest solutions of underdetermined linear systems via l_q-minimization for $0 < q \leq 1$," *Appl. Comput. Harmon. Anal.*, vol. 26, no. 3, pp. 395–407, 2009.

[26] J. F. Yang and Y. Zhang, "Alternating direction algorithms for ℓ_1-problems in compressive sensing," *SIAM J. Sci. Comput.*, vol. 33, pp. 250–278, 2011.

[27] C. Studer and R. G. Baraniuk, "Stable restoration and separation of approximately sparse signals," *Appl. Comput. Harmon. Anal.*, vol. 37, no. 1, pp. 12–35, 2014.

[28] D. S. Pham and S. Venkatesh, "Efficient algorithms for robust recovery of images from compressed data," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 4724–4737, Dec. 2013.