Neurocomputing 223 (2017) 103–106

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Brief papers

Mean-square analysis of the gradient projection sparse recovery algorithm based on non-uniform norm

F.Y. Wu, F. Tong*

Key laboratory of underwater acoustic communication and marine information technology of the Minister of Education, Xiamen University, Xiamen, China

ARTICLE INFO

Communicated by Huaping Liu

Keywords:
Non-uniform norm
Sparse recovery
Compressed sensing
Gradient projection

ABSTRACT

With the previously proposed non-uniform norm called l_N-norm, which consists of a sequence of l_1-norm or l_0-norm elements according to relative magnitude, a novel l_N-norm sparse recovery algorithm can be derived by projecting the gradient descent solution to the reconstruction feasible set. In order to gain analytical insights into the performance of this algorithm, in this letter we analyze the steady state mean square performance of the gradient projection l_N-norm sparse recovery algorithm in terms of different sparsity, as well as additive noise. Numerical simulations are provided to verify the theoretical results.

1. Introduction

Obtaining the sparse recovery solution under the framework of compressed sensing (CS) has gained more and more attention in recent years. The l_1-norm reconstruction algorithms, such as the interior-point method for large-scale l_1-regularized least squares (l1-ls) method [2] and gradient projection for sparse reconstruction [3], are usually adopted by searching for a solution with minimum l_1-norm. In comparison, because direct search of minimum l_0-norm will generally lead to NP hard problem [4–7], various approximation methods are also investigated to solve the difficulties caused by l_0-norm [4–6] such as the smoothing l_0–norm (SL0) [8] and l_0–norm zero-point attracting projection (l_0–ZAP) [9]. Another type of popular approaches is greedy method such as matching pursuit (MP) and orthogonal matching pursuit (OMP) [4,5], by which the approximation is generated via an iteratively process to search the column vectors that most closely resemble the required. Based on block based least square and minimum mean-square-error cost function, iteratively reweighted least squares (IRLS) minimization technique is used for iterative sparse recovery [6]. Moreover, in [10] the Frobenius norm and the l_1-norm of the Euclidean norm are used to design an iterative optimization algorithm for the structured sparse coding model.

In [11,12], a new non-uniform norm called l_N-norm, which consisted of a sequence of l_0 or l_1 norm elements according to relative magnitude, is proposed to exploit the sparseness while providing adaptability to different sparsity of sources. It is pointed out in [11] that imposing l_N-norm constraint on the Least Mean Square (LMS) iteration yields enhanced convergence rate as well as better tolerance upon different sparsity in system identification. The sparsity exploitation performance of the l_N-norm LMS algorithm has been compared with that of other existing sparsity-aware algorithms in [14,15]. In [16], the concept of non-uniform norm is further combined with variable step-size to derive the p norm variable step-size LMS algorithm, which claims to outperform the classic l_N-norm LMS.

The concept of l_N-norm can be also introduced into the sparse signal reconstruction at the presence of source signals associated with different sparsity. Similar to previous gradient projection type approaches [3], the iterative optimization solution of the proposed l_N-norm sparse reconstruction can be derived by directly minimizing the l_N-norm cost function via the steepest descent method, and then affine projecting the solution to the feasible set. However, there is a lack of analytical steady state performance in terms of l_N-norm sparsity recovery. In this letter, the steady state mean-square performance of the l_N-norm gradient projection sparse recovery algorithm is theoretically performed. Finally numerical simulation results are provided to verify the analysis.

2. Derivation of the non-uniform norm CS Algorithm

The problem to obtain l_1-norm or l_0-norm sparse solution can be respectively expressed as:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{Ax} = \mathbf{y} \tag{1}$$

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{subject to} \quad \mathbf{Ax} = \mathbf{y} \tag{2}$$

where $\|\mathbf{x}\|_1 = \sum_{i=1}^{n} |x(i)|$ and $\|\mathbf{x}\|_0 = \#\{i|x(i) \neq 0, \quad i = 1, \ldots, n\}$.

In [11,12], the concept of p-norm like is introduced and defined as:

* Corresponding author.
E-mail address: ftong@xmu.edu.cn (F. Tong).

http://dx.doi.org/10.1016/j.neucom.2016.10.032
Received 1 September 2015; Received in revised form 14 September 2016; Accepted 25 October 2016
Available online 28 October 2016

$$\|\mathbf{x}\|_p^p = \sum_{i=1}^{n} |x(i)|^p, \quad 0 \le p \le 1. \tag{3}$$

Furthermore, a non-uniform norm, denoted as $\|\mathbf{x}\|_N$ in this study, is defined in [12]. It is noticeable that $\|\mathbf{x}\|_N$ utilizes different value of p for each entry of $\mathbf{x}$, i.e. $\|\mathbf{x}\|_N = \sum_{i=1}^{n} |x(i)|^{p_i}, \quad 0 \le p_i \le 1$. Moreover, by classifying the entries of $x(i)$ into 'large' and 'small' ones according to relative magnitude [12], the non-uniform norm actually consists of a sequence of l_1-norm or l_0-norm elements, with p_i taking 0 for 'large' entries and taking 1 for 'small' entries respectively.

As the definition indicates, the non-uniform norm is a quasi-norm. Similarly, the l_N-norm sparse solution can be expressed as:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_N \ \text{ subject to } \ \mathbf{Ax} = \mathbf{y} \tag{4}$$

(4) can be further relaxed as the following problem:

$$\mathbf{x}_{opt} = \arg \min_{\mathbf{x}} (\|\mathbf{y} - \mathbf{Ax}\| + \lambda \|\mathbf{x}\|_N) \tag{5}$$

where $\|\mathbf{y} - \mathbf{Ax}\|$ means Eulidean norm and $\mathbf{x}_{opt}$ is the l_N-norm sparse optimization solution, $\lambda > 0$ is a factor to balance the contribution of the l_N-norm sparsity criterion. Different from the classic Eulidean norm, the adjustability of p_i parameter enables certain adaptability to the sparsity by implicitly tuning between the l_0 norm and l_1 norm [12], as taking $p_i = 0$ or $p_i = 1$ to large or small element will lead to different mixing pattern of l_1-norm or l_0-norm element.

As a penalty term in the modified LMS cost function, the iterative optimization of the non-uniform norm is achieved by gradient descent in [12]. Inspired by similar gradient projection method [3], the l_N norm sparse recovery can be obtained by iteratively projecting the l_N-norm gradient descent sparse solution to the feasible set of reconstruction. The proposed sparse recovery algorithm is described using pseudo-codes as Table 1.

In Table 1, $\mathbf{A}^{+} = \mathbf{A}^T(\mathbf{AA}^T)^{-1}$, μ is the step size and $\mathbf{x}_j = [x(1)_j, \ldots, x(k)_j, \ldots, x(n)_j]^T$ at j-th iteration. sgn() denotes sign function defined as:

$$\text{sgn}(x) = \begin{cases} 0, & x = 0 \\ \frac{x}{|x|}, & x \ne 0 \end{cases} \tag{9}$$

The iteration of the l_N norm CS method can be divided as two parallel processes: the first one is finding l_N norm gradient descent sparest solution with (7), the second one is using affine projection via Laplacian method [8,9] to obtain l_N norm CS solution from the reconstruction feasible set by (8).

3. Steady state performance analysis of the l_N norm CS algorithm

Without loss of generality, in this letter the entries of mixing matrix $\mathbf{A}$ are independently sampled from a normal distribution with mean zero and variance of $\frac{1}{m}$, which ensure each column vector of $\mathbf{A}$ are normalized.

We define the misalignment vector as $\mathbf{h}_j = \mathbf{x}_j - \mathbf{x}_O$, where $\mathbf{x}_O$ is the optimum gradient descent solution, and the actual possible sampling result $\mathbf{y}_1$ related to the optimum gradient descent solution can be rewritten as:

$$\mathbf{y}_1 = \mathbf{Ax}_O + \mathbf{w} \tag{10}$$

where $\mathbf{w}$ can be regarded as deviation noise signal between the actual possible sampling result $\mathbf{y}_1$ and $\mathbf{Ax}_O$. As the gradient descent iteration is used to find the l_N sparsest solution $\mathbf{x}$ of

$$\mathbf{y} = \mathbf{Ax} \tag{11}$$

Thus the reconstruction error can be expressed as:

$$\begin{aligned} \mathbf{e} &= \mathbf{y}_1 - \mathbf{y} \\ &= \mathbf{Ax}_O + \mathbf{w} - \mathbf{Ax} \\ &= \mathbf{w} - \mathbf{A}(\mathbf{x} - \mathbf{x}_O) \end{aligned} \tag{12}$$

Theorem 1. Suppose $\mathbf{x}_O$ is the original signal, and $\mathbf{x}$ is the optimum solution of l_N norm CS method, the final mean square derivation (MSD) in steady state is:

$$D(\infty) = E\{tr[\mathbf{P}_\infty]\} \le \frac{n^2}{2mn - m^2}\left[2\frac{n-m}{n}\gamma\cdot\mu\cdot SR + \mu^2 + \left(\frac{m}{n}\right)^2 \sigma_{\mathbf{w}}^2\right] \tag{13}$$

where $\gamma = \max(|\mathbf{x}_j - \mathbf{x}_O|)$, $\sigma_{\mathbf{w}}^2$ denotes the variance of the noise signal $\mathbf{w}$. The sparse ratio (SR) is defined as the ratio of the number of nonzero elements to the number of the whole original source signal $\mathbf{x}_O$.

Proof: In (6), the value of $f_{j-1}(k)$ just can be 0 or 1, then in (7), we denote:

$$g(x_{j-1}(k)) = \frac{\mu\cdot\text{sgn}(x_{j-1}(k))\cdot f_{j-1}(k)}{1 + \varepsilon|x_{j-1}(k)|}, \qquad 1 \le k \le n; \tag{14}$$

Obviously, the value of $g(x_{j-1}(k))$ just can be 0 or $\pm \frac{\mu}{1+\varepsilon|x_{j-1}(k)|}$, denote the vector $\mathbf{g}_{\,j-1}$ as:

$$\mathbf{g}_{\,j-1} = [g(x_{j-1}(1)), \ldots g(x_{j-1}(k))\ldots g(x_{j-1}(n))]^T, \qquad 1 \le k \le n; \tag{15}$$

We can get $|\mathbf{g}_{\,j-1}| < \mu\cdot\mathbf{1}$, where $\mathbf{1}$ has the same size of $\mathbf{x}$, then (7) can be rewritten as:

$$\mathbf{x}_j = \mathbf{x}_{j-1} + \mathbf{g} \tag{16}$$

Combination of (8) and (10) leads to:

$$\mathbf{x}_{j+1} = \mathbf{x}_j + \mathbf{A}^{+}(\mathbf{w} - \mathbf{A}(\mathbf{x}_j - \mathbf{x}_O)) \tag{17}$$

Using the definition of the misalignment vector $\mathbf{h}_j = \mathbf{x}_j - \mathbf{x}_O$, thus we have:

$$\mathbf{h}_{j+1} = \mathbf{h}_j + \mathbf{A}^{+}(\mathbf{w} - \mathbf{Ah}_j) + \mathbf{g} = (\mathbf{I} - \mathbf{A}^{+}\mathbf{A})\mathbf{h}_j + \mathbf{A}^{+}\mathbf{w} + \mathbf{g} \tag{18}$$

By post-multiplying both sides of (18) with their respective transposes, we have:

$$\begin{aligned} \mathbf{h}_{j+1}\mathbf{h}_{j+1}^T = \ & (\mathbf{I} - \mathbf{A}^{+}\mathbf{A})\mathbf{h}_j\mathbf{h}_j^T(\mathbf{I} - \mathbf{A}^{+}\mathbf{A})^T + (\mathbf{I} - \mathbf{A}^{+}\mathbf{A})\mathbf{h}_j\mathbf{w}^T(\mathbf{A}^{+})^T \\ & + (\mathbf{I} - \mathbf{A}^{+}\mathbf{A})\mathbf{h}_j\mathbf{g}^T + \mathbf{A}^{+}\mathbf{w}\mathbf{h}_j^T(\mathbf{I} - \mathbf{A}^{+}\mathbf{A})^T + \mathbf{A}^{+}\mathbf{w}\,\mathbf{w}^T(\mathbf{A}^{+})^T + \mathbf{A}^{+}\mathbf{w}\mathbf{g}^T \\ & + \mathbf{g}\mathbf{h}_j^T(\mathbf{I} - \mathbf{A}^{+}\mathbf{A})^T + \mathbf{g}\mathbf{A}^{+}\mathbf{w} + \mathbf{g}\mathbf{g}^T \end{aligned} \tag{19}$$

As the noise $\mathbf{w}$ is independent with $\mathbf{x}$ and $\mathbf{A}$. Let $\mathbf{P}_j = E[\mathbf{h}_j\mathbf{h}_j^T]$, $\mathbf{R} = E[\mathbf{A}^{+}\mathbf{A}]$, taking expectations on both sides of (19) will lead to:

$$\begin{aligned} \mathbf{P}_{j+1} = \ & (\mathbf{I} - \mathbf{R})\mathbf{P}_j(\mathbf{I} - \mathbf{R})^T + 2(\mathbf{I} - \mathbf{R})E[\mathbf{h}_j\mathbf{g}^T] \\ & + E[\mathbf{g}\mathbf{g}^T] + \sigma_{\mathbf{w}}^2\mathbf{A}^{+}(\mathbf{A}^{+})^T \end{aligned} \tag{20}$$

where $\sigma_{\mathbf{w}}^2$ denotes the variance of the noise $\mathbf{w}$. According to the central limit theorem [13], the following equation holds approximately

$$(\mathbf{AA}^T)^{-1} \approx \frac{m}{n}\mathbf{I} \tag{21}$$

Table 1
Pseudo-Codes of The Proposed Algorithm.

1:	Initialize $\mathbf{x}_0 = \mathbf{A}^{+}\mathbf{y}$, $\mu = \frac{1}{\max(|\mathbf{x}|)}$; choose μ_{th}, α, ε, d, J.
2:	Judge whether stop condition is satisfied $\mu > \mu_{th}$. If satisfied, send $\mathbf{x}_j$ back to $\mathbf{x}$ and exit; otherwise turn to Step 3.
3:	*for* $j = 1:J$ $f_{j-1}(k) = \frac{\text{sgn}(|x_{j-1}(k)| - \alpha\cdot\mu) + 1}{2}$, $1 \le k \le n$; (6) $x_j(k) = x_{j-1}(k) + \frac{\mu\cdot\text{sgn}(x_{j-1}(k))\cdot f_{j-1}(k)}{1+\varepsilon|x_{j-1}(k)|}$; (7) $\mathbf{x}_{j+1} = \mathbf{x}_j + \mathbf{A}^{+}(\mathbf{y} - \mathbf{Ax}_j)$; (8) *end for*
4:	Step size updates: $\mu = d\mu$;

where $\mathbf{A}^T\mathbf{A} \approx \mathbf{I}$ since the entries of mixing matrix $\mathbf{A}$ are independently sampled from a normal distribution with mean zero and variance of $\frac{1}{m}$. Then, we have:

$$\mathbf{A}^{+}(\mathbf{A}^{+})^T = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}(\mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1})^T \approx \frac{m}{n}\mathbf{A}^T\left(\mathbf{A}^T\frac{m}{n}\right)^T = \left(\frac{m}{n}\right)^2\mathbf{I} \tag{22}$$

and

$$\mathbf{R} = E[\mathbf{A}^{+}\mathbf{A}] = E[\mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{A}] \approx \frac{m}{n}\mathbf{I} \tag{23}$$

Substituting (22) and (23) into (20), then one can get

$$\begin{aligned}\mathbf{P}_{j+1} = &(1-\tfrac{m}{n})^2\mathbf{P}_j + 2(1-\tfrac{m}{n})E[\mathbf{h}_j\mathbf{g}^T]\\ &+E[\mathbf{g}\mathbf{g}^T] + \sigma_{\mathbf{w}}^2(\tfrac{m}{n})^2\mathbf{I}\end{aligned} \tag{24}$$

where $(1-\frac{m}{n})^2 < 1$ means that the iteration process is convergent.

Considering that the value of $g(x_{j-1}(k))$ just can be 0 or $\pm\frac{\mu}{1+\varepsilon|x_{j-1}(k)|}$, and the number of nonzero elements in vector $\mathbf{g}_{j-1}$ is proportional to SR. Take trace of the expectation of $\mathbf{h}_j\mathbf{g}^T$, one can get

$$E\{tr[|E[\mathbf{h}_j\mathbf{g}^T]|]\} \le E\{tr[E|(\mathbf{x}_j-\mathbf{x}_O)\mathbf{g}^T|]\} \le \gamma\cdot\mu\cdot SR \tag{25}$$

where $\gamma = \max(|\mathbf{x}_j - \mathbf{x}_O|)$. Then

$$E\{tr[\mathbf{P}_{j+1}]\} \le \left(1-\frac{m}{n}\right)^2 E\{tr[\mathbf{P}_j]\} + 2\left(1-\frac{m}{n}\right)\gamma\cdot\mu\cdot SR + \mu^2 + \left(\frac{m}{n}\right)^2\sigma_{\mathbf{w}}^2 \tag{26}$$

Thus the final steady state mean square deviation is obtained as:

$$D(\infty) = E\{tr[\mathbf{P}_\infty]\} \le \frac{n^2}{2mn-m^2}\left[2\frac{n-m}{n}\gamma\cdot\mu\cdot SR + \mu^2 + \left(\frac{m}{n}\right)^2\sigma_{\mathbf{w}}^2\right] \tag{27}$$

4. Numerical simulation

In this section, numerical simulations are performed to verify the theoretical analysis of mean-square performance. The MSD of the sparse reconstruction is defined as:

$$MSD = \|\mathbf{x} - \mathbf{x}_O\|_2^2 \tag{28}$$

where $\mathbf{x}_O$ is ideal sparse signal size of $n \times 1$. In order to formulate the CS problem, the matrix and vectors are generated according to (10), and the nonzero value of $\mathbf{x}_O$ is generated according to normal distribution process as following:

$$x_i \sim SR\cdot N(0, \sigma_u^2) + (1-SR)\cdot N(0, \sigma_w^2), \qquad \forall\, 0 \le i \le n \tag{29}$$

where σ_u^2 and σ_w^2 are the variance of the zero-mean white Gaussian random distributed sources in active mode and inactive mode respectively. In the sources σ_w^2 actually models the noise at the form of small value"inactive" sources. The"inactive" sources can be regarded as the noise signal $\mathbf{w}$. Then the SNR of the source signal is defined as:

$$SNR = 10\cdot\log_{10}\frac{SR\cdot\sigma_u^2}{(1-SR)\cdot\sigma_w^2} \tag{30}$$

According to the discussion above, for each column of the $m \times n$ dimension matrix $\mathbf{A}$, vector elements are generated according to normal distribution process, with mean zero and variance of $\frac{1}{m}$. The algorithm parameters are set as: $\alpha = 4$, $d = 0.8$, $\varepsilon = 4$, $J = 3$, $\mu_{th} = 0.001$. If there is no other explanation, the parameters in the proposed algorithm are set as described above. All the simulations were repeated 100 times with the same conditions.

First, the steady-state MSD with respect to different noise level is compared between numerical simulation and theoretical analysis, with $SR = 0.1, n = 300, m = 100$. Both the analytical and simulated steady-state MSDs are shown in Fig. 1, It is observed that the simulation results are generally consistent with our analytical MSD performance especially when the SNR of source signal is more than 27 dB. For the aim of performance comparison, simulation results of the sl0 norm sparse recovery algorithm [8] are also provided in Fig. 1, which indicating that the l_N norm CS method outperforms the sl0 algorithm. Furthermore, iterative convergences of the l_N norm CS method associated with SNR of 25 dB, 27 dB, and 30 dB are provided in Fig. 2, which verified the effectiveness of the theoretical MSD analysis.

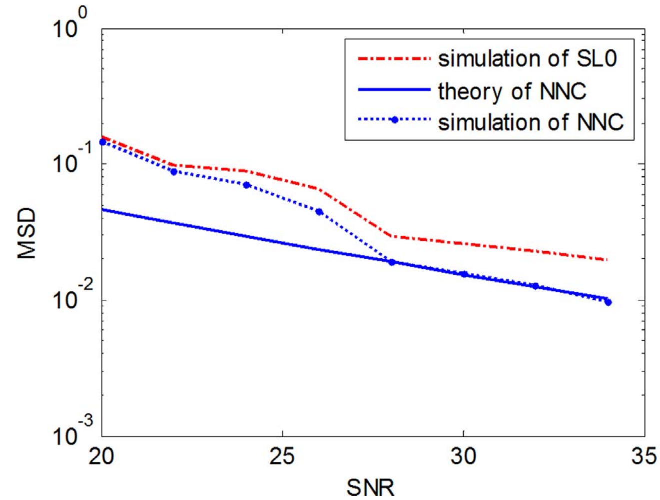


Fig. 1. Analytical and simulation results under different SNR.

Next, the Theorem 1 is investigated with numerical simulations in terms of different sparsity with the parameter of $SR = 0.01 + 0.02(i-1)$, $i = 1, \ldots, 5$ and the SNR of 30 dB. To guarantee the sparse signal $\mathbf{x}_O$ and the measurement matrix $\mathbf{A}$ meet the condition of RIP [1–6], the rows of $\mathbf{A}$ are set to be big enough, as m=130, 140, and 150 respectively. The corresponding steady-state MSDs are depicted in Fig. 3. One can observe that MSDs of simulated results are consistent with the analytical results. It is also evident from Fig. 3 that, with increasing of the SR, both the analytical and the numerical simulation MSDs exhibit a rising trend, due to the contribution of the first term in (13).

5. Conclusion

Extending from the application of the non-uniform norm for sparsity exploitation in the framework of LMS algorithm, a novel l_N-norm sparse recovery algorithm is derived in this letter, by projecting the gradient descent l_N norm solution to the reconstruction feasible set. Meanwhile, the steady state MSD performance of the proposed l_N

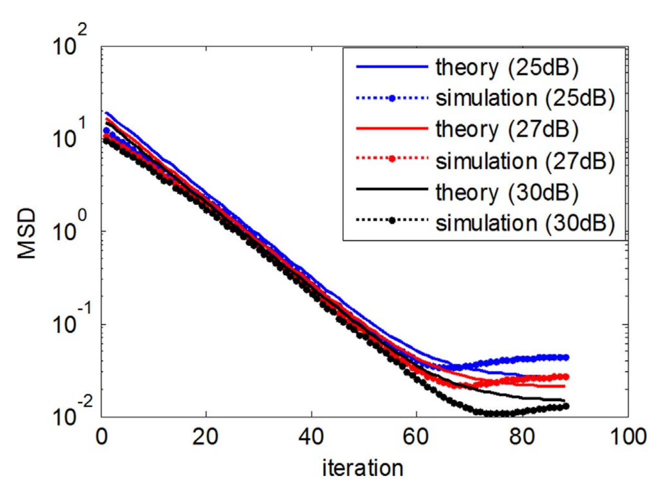


Fig. 2. The iteration process of theoretical and simulation MSD.

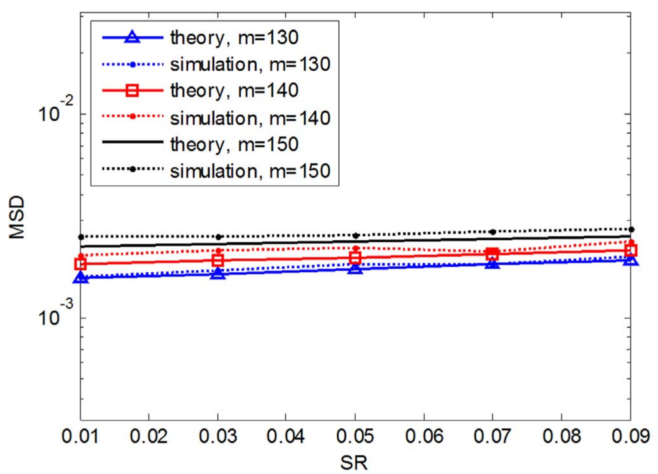


Fig. 3. The MSD performance with respect to different sparsity.

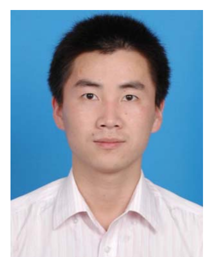
norm compressed sensing signal recovery algorithm is theoretically investigated in terms of different sparsity as well as different additive noise. The performance analysis is verified by numerical simulations.

Acknowledgment

The authors are grateful for the funding by Grants from the National Natural Science Foundation of China (Project no. 11274259 and Project no. 11574258) in support of the present research. The authors also would like to thank the anonymous reviewers for improving the quality of this letter.

References

[1] I. Takigawa, M. Kudo, J. Toyama, performance analysis of minimum l1-norm solutions for underdetermined source separation, IEEE Trans. Signal Process. 52 (3) (2004) 582–591 (Mar.).
[2] S.J. Kim, K. Koh, M. Lustig, S. Boyd, D. Gorinevsky, An interior point method for large-scale l1-regularized least squares, IEEE J. Sel. Top. Signal Process. 1 (4) (2007) 606–617 (Dec.).
[3] M. Figueiredo, R. Nowak, S. Wright, Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems, IEEE J. Sel. Top. Signal Process. 1 (4) (2007) 586–597 (Dec.).
[4] J.A. Tropp, Greed is good: algorithmic results for sparse approximation, IEEE Trans. Inform. Theory 50 (10) (2004) 2231–2242 (Oct.).
[5] J.A. Tropp, A.C. Gilbert, Signal recovery from random measurements via orthogonal matching pursuit, IEEE Trans. Inform. Theory 53 (12) (2007) 4655–4666 (Dec.).
[6] I. Daubechies, R. DeVore, M. Fornasier, C.S. Gunturk, Iteratively reweighted least squares minimization for sparse recovery, Commun. Pure Appl. Maths. 63 (1) (2010) 1–38 (Jan.).
[7] Y. Gu, J. Jin, S. Mei, l0 norm constraint LMS algorithm for sparse system identification, IEEE Signal Process. Lett. 16 (9) (2009) 774–777 (Sep.).
[8] M. Hosein, B.Z. Massoud, J. Christian, A fast approach for overcomplete sparse decomposition based on smoothed l0 Norm, IEEE Trans. Signal Process. 57 (1) (2009) 289–301 (Jan.).
[9] J. Jin, Y. Gu, S. Mei, A stochastic gradient approach on compressive sensing signal reconstruction based on adaptive filtering framework, IEEE J. Sel. Top. Signal Process. 4 (2) (2010) 409–420 (Apr.).
[10] H. Liu, Y. Liu, F. Sun, Robust exemplar extraction using structured sparse coding, IEEE Trans. Neural Netw. Learn. Syst. 26 (8) (2015) 1816–1821 (Aug.).
[11] F.Y. Wu, F. Tong, Gradient optimization p-norm-like constraint LMS algorithm for sparse system estimation, Signal Process. 93 (4) (2013) 961–967 (Apr.).
[12] F.Y. Wu, F. Tong, Non-uniform norm constraint LMS algorithm for sparse system identification, IEEE Commun. Lett. 17 (2) (2013) 385–388 (Feb.).
[13] O. Kallenberg, Foundation of Modern Probability, 2nd ed., Springer, New York, 2002.
[14] C. Wang, Y. Zhang, Y. Wei, N. Li, A new l0-LMS algorithm with adaptive zero attractor, IEEE Commun. Lett. 19 (12) (2015) 2150–2153 (Dec.).
[15] H. Zayyani, M. Korki, F. Marvasti, A distributed 1-bit compressed sensing algorithm robust to impulse noise, IEEE Commun. Lett. 20 (6) (2016) 1132–1135.
[16] M.L. Aliyu, M.A. Alkassim, M.S. Salman, A p-norm variable step-size LMS algorithm for sparse system identification, Signal Image Video Process. 9 (7) (2015) 1559–1565.



Feiyun Wu received his B.S., M.S and Ph.D degrees from Nanchang University, Nanchang, Sun Yat-Sen University, Guangzhou and Xiamen University, Xiamen, China, 2006, 2010, 2016, respectively. From 2006-2008, he worked as a college lecturer in Electronic Engineering Department of Nanchang University Gongqing College. During 2013-2015, he was a visiting Ph.D student in University of Delaware, DE, US, Now he is an assistant professor in Northwestern Polytechnical University, Xi'an, China. His research interests include adaptive signal processing for underwater acoustics communications, compressed sensing, ICA and wavelet algorithms and their applications.

Feng Tong received his PhD in underwater acoustics at Xiamen University, China in 2000. From 2000–2002, he worked as a post-doctoral fellow in the Department of Radio Engineering, Southeast University, China. Since 2003, he has been a research associate at the Department of Manufacturing Engineering and Engineering Management, City University of Hong Kong for one and a half year. From Dec 2009 to Dec 2010 He was a visiting scholar in Department of Computer Science and Engineering, University of California San Diego, USA. Currently he is a professor with the Department of Applied Marine Physics and Engineering, Xiamen University, China. His research interests focus on underwater acoustic communication and acoustic signal processing. He is a member of IEEE, ASC (Acoustical Society of China) and CSIS (China Ship Instrument Society). He serves on the editorial board for the Journal of Marine Science and Application.