l_1-norm Based GWLP for Robust Frequency Estimation

Yuan Chen[1] , Liangtao Duan[1], Weize Sun[2, *] and Jingxin Xu[3]

Abstract: In this work, we address the frequency estimation problem of a complex single-tone embedded in the heavy-tailed noise. With the use of the linear prediction (LP) property and ℓ_1-norm minimization, a robust frequency estimator is developed. Since the proposed method employs the weighted ℓ_1-norm on the LP errors, it can be regarded as an extension of the ℓ_p-generalized weighted linear predictor. Computer simulations are conducted in the environment of α-stable noise, indicating the superiority of the proposed algorithm, in terms of its robust to outliers and nearly optimal estimation performance.

Keywords: Robust frequency estimation, linear prediction, impulsive noise, weighted ℓ_1-norm minimization.

1 Introduction

Frequency estimation from a finite number of complex sinusoids is an of importance research topic and attract much attention, in many applications such as array processing, digital communications, and biomedical engineering [Chan, So and Huang (2015); Qian, Huang and So (2016); Kay (1993)]. Among numerous techniques developed in the literature, the maximum likelihood (ML) and nonlinear least squares (NLS) [Stoica (2005)] are most representative ones. Nevertheless, since the frequency is estimated in terms of excessive iterations on a nonlinear cost function, both ML and NLS suffer from a high computational complexity. To decrease such high complexity requirements, [So, Chan and Ho (2005); So and Chan (2006)] suggest generalized weighted linear predictor (GWLP), whose main idea is reformulating the nonlinear problem as a linear one by employing the linear prediction (LP) property of the complex tone, while the frequency is updated iteratively according to the weighted least squares (WLS) technique. Although it provides the optimum estimation performance for Gaussian noise, it cannot work properly to α-stable noise, which occurs in many practical scenarios [Nikias and Shao (1995); Laguna-Sanchez and Lopez-Guerrero (2014); Zoubir and Koivunen (2012)]. This is because the LS approach in GWLP is sensitive to outliers of impulsive noise.

To resist outliers, the generalized version of the GWLP, referred to as the ℓ_p-GWLP [Chen, Yang and Huang (2018)], is devised by replacing the WLS with the weighted ℓ_p-norm $(1 < p < 2)$ [Shao and Nikias (1993)]. The iteratively reweighted least squares (IRLS)

[1] University of Science & Technology Beijing, Beijing, 100083, China.
[2] ShenZhen University, Shenzhen, 518060, China.
[3] Department of Housing and Public Works, Queensland Government, Brisbane, Australia.
[*] Corresponding Author: Weize Sun. Email: Proton198601@hotmail.com.

[Byrd and Payne (1979)] is then utilized in GWLP to obtain the estimate of the frequency. However, this method cannot be applied in the case of $p = 1$, due to the divergence of the IRLS method. In this work, we address the problem of extending ℓ_p-GWLP with $1 < p < 2$ to $p = 1$. To guarantee the convergence of IRLS algorithm, we redefine the reweighting matrix by adding a small constant in each diagonal elements [Wu (2018)].

The rest of this paper is organized as follows. The development of the ℓ_1-GWLP is presented in Section 2, where the symmetric α-stable (SαS) distribution is taken as an illustration. In Section 3, computer simulations are carried out to demonstrate the effectiveness of the proposed scheme. Finally, conclusions are drawn in Section 4.

2 Proposed method

Without loss of generality, we start with the signal model:

$$\mathbf{y} = \mathbf{s} + \mathbf{q}, \tag{1}$$

where $\mathbf{y} = [y_1 \; y_2 \; \cdots \; y_N]^T \in \mathbb{C}^N$ is the observed data vector with ^{T} being the transpose operator, $\mathbf{s} = [s_1 \; s_2 \; \cdots \; s_N]^T \in \mathbb{C}^N$ is the signal, and $\mathbf{q} = [q_1 \; q_2 \; \cdots \; q_N]^T \in \mathbb{C}^N$ denotes the noise vector whose elements are independent identically distributed (IID) complex isotropic SαS random variable, with zero location parameter and covariation matrix $\gamma \mathbf{I}_N$. Here α and γ denote the shape parameter and dispersion of the noise, respectively.

In this paper, the single complex tone is considered, whose n-th element has the form of

$$s_n = A \exp(j(\omega n + \theta)), \tag{2}$$

where $A \in (0, +\infty)$, $\omega \in (-\pi, \pi)$ and $\theta \in [0, 2\pi)$ are the unknown amplitude, frequency and phase, respectively. The task is estimating the frequency ω from observations $\mathbf{y}$. Then the LP property of s_n is expressed as

$$s_n = \rho s_{n-1}, \tag{3}$$

where $\rho = \exp(j\omega)$. Let $\mathbf{x}_1 = [y_2 \; y_3 \; \cdots \; y_N]^T$ and $\mathbf{x}_2 = [y_1 \; y_2 \; \cdots \; y_{N-1}]^T$. The LP error vector between $\mathbf{x}_1$ and $\mathbf{x}_2$ is

$$\mathbf{e} = \mathbf{x}_1 - \rho \mathbf{x}_2, \tag{4}$$

where $\mathbf{e} = [e_1 \; e_2 \; \cdots \; e_{N-1}]^T \in \mathbb{C}^{N-1}$. As it is discussed in Shao et al. [Shao and Nikias (1993)], the LP error vector $\mathbf{e}$ follows the same distribution with the noise vector $\mathbf{q}$ which is SαS distribution with covariation matrix $\mathbf{D}$. According to the statistical property of SαS distribution, the (m, n) entry of $\mathbf{D}$, denoted by $[\mathbf{D}]_{m,n}$, can be expressed as

$$[\mathbf{D}]_{m,n} = \frac{E\{e_m |e_n|^{-1} e_n^*\}}{E\{|e_n|\}} \gamma_{e_n}, \tag{5}$$

where γ_{e_n} denotes the dispersion of residual e_n. Employing the fact that $e_n = q_{n+1} - \rho q_n$ and $e_{n-1} = q_n - \rho q_{n-1}$ as well as [Tsakalides (1995)], $\mathbf{D}$ has the form of

$$\mathbf{D} = C(\gamma, \rho) \mathbf{W}^{-1} \tag{6}$$

where $^{-1}$ denotes the matrix inverse operator, $C(\gamma, \rho) = \frac{(1+\rho^2)^{1-\frac{1}{\alpha}} \Gamma(\frac{1}{2})}{\sqrt{\pi}\Gamma(1-\frac{1}{\alpha})} \gamma^{1-\frac{1}{\alpha}}$ and

$$\mathbf{W} = \begin{bmatrix} 1+|\rho|^{\alpha} & -\rho & 0 & \cdots & 0 \\ -\rho^*|\rho|^{\alpha-2} & 1+|\rho|^{\alpha} & -\rho & \cdots & 0 \\ 0 & -\rho^*|\rho|^{\alpha-2} & 1+|\rho|^{\alpha} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1+|\rho|^{\alpha} \end{bmatrix}^{-1}. \tag{7}$$

With the use of LP property in (4), our task can be converted from a nonlinear problem of estimating ω into the linear one of estimating ρ.

According to (6), e_n is dependent on e_{n-1}, since covariation matrix is not diagonal. Therefore, to ensure the optimum estimation performance, the dependence on $\mathbf{e}$ should be removed by whitening transform [Kessy, Lewin and Strimmer (2015)], which is

$$\mathbf{e}_w = \mathbf{W}\mathbf{e}, \tag{8}$$

where $\mathbf{e}_w = [e_{w_1}\ e_{w_2}\ \cdots\ e_{w_{N-1}}]^T$ is the whitened noise. It is worth to point out that after whitening process in (8), the new residual $\mathbf{e}_w$ still follows the IID $S\alpha S$ distribution because of the linear combination in the pre-process [Shao and Nikias (1993)] .

Employing the ℓ_1-norm on (8), the estimate of, denoted by $\hat{\rho}$, can be obtained by minimizing the cost function $J(\rho)$:

$$\hat{\rho} = \arg\min_{\rho} J(\rho), \tag{9}$$

where $J(\rho) = \sum_{n=1}^{N-1} |e_{w_n}| = \sum_{n=1}^{N-1} e_{w_n}^* \frac{1}{|e_{w_n}|} e_{w_n}$. Based on the definition of e_{w_n} in (8), $J(\rho)$ in the vector form can be expressed as

$$\begin{aligned} J(\rho) &= \mathbf{e}_w^H \mathrm{diag}\left\{ |e_{w_1}|^{p-2}\ |e_{w_2}|^{-1}\ \cdots\ |e_{w_{N-1}}|^{-1} \right\} \mathbf{e}_w \\ &= (\mathbf{W}\mathbf{e})^H \mathbf{O} (\mathbf{W}\mathbf{e}), \end{aligned} \tag{10}$$

where ^{H} denotes the Hermitian operator and

$$\mathbf{O} = \left\{ \frac{1}{|e_{w_1}|}\ \frac{1}{|e_{w_2}|}\ \cdots\ \frac{1}{|e_{w_{N-1}}|} \right\} \tag{11}$$

is a diagonal matrix. From (11), we can see that once the whitened residual $|e_{w_n}|$ in $\mathbf{O}$ is zero, $\frac{1}{|e_{w_1}|}$goes to infinity, resulting in the divergence of the IRLS method. Therefore, to this problem, we define $|e_{w_n}|$ with

$$|e_{w_n}| \approx \sqrt{e_{w_n}^2 + \epsilon}, \tag{12}$$

where ϵ is chosen as a small positive constant.

Substituting (4) into (10) yields

$$J(\rho) = (\mathbf{x}_1 - \rho\mathbf{x}_2)^H \mathbf{Q} (\mathbf{x}_1 - \rho\mathbf{x}_2), \tag{13}$$

where

$$\mathbf{Q} = \mathbf{W}^H \mathbf{P} \mathbf{W}, \tag{14}$$

With

$$\mathbf{P} = \left\{ \frac{1}{\sqrt{e_{w_1}^2 + \epsilon}} \; \frac{1}{\sqrt{e_{w_2}^2 + \epsilon}} \; \cdots \; \frac{1}{\sqrt{e_{w_{N-1}}^2 + \epsilon}} \right\} \tag{15}$$

Then the $(k+1)$-th iteration estimate of ρ, denoted by $\hat{\rho}^{(k+1)}$, can be obtained by minimizing (9) using the IRLS [Merle and Spath (1974)],

$$\hat{\rho}^{(k+1)} = \frac{\mathbf{x}_1^H \mathbf{Q}^{(k)} \mathbf{x}_2}{\mathbf{x}_1^H \mathbf{Q}^{(k)} \mathbf{x}_1}, \tag{16}$$

where

$$\mathbf{Q}^{(k)} = \left(\mathbf{W}^{(k)}\right)^H \mathbf{P}^{(k)} \mathbf{W}^{(k)} \tag{17}$$

With

$$\mathbf{P}^{(k)} = \begin{bmatrix} \sqrt{\left(e_{w_1}^{(k)}\right)^2 + \epsilon} & 0 & \cdots & 0 \\ 0 & \sqrt{\left(e_{w_2}^{(k)}\right)^2 + \epsilon} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\left(e_{w_{N-1}}^{(k)}\right)^2 + \epsilon} \end{bmatrix} \tag{18}$$

$$\mathbf{W}^{(k)} = \begin{bmatrix} 1 + |\rho^{(k)}|^\alpha & -\rho^{(k)} & \cdots & 0 \\ -\left(\rho^{(k)}\right)^* |\rho^{(k)}|^{\alpha-2} & 1 + |\rho^{(k)}|^\alpha & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 + |\rho^{(k)}|^\alpha \end{bmatrix}, \tag{19}$$

$$\mathbf{e}_{w_n}^{(k)} = \left[\left(\mathbf{W}^{(k)}\right)^{-\frac{1}{p}} \mathbf{e}^{(k)} \right]_n \tag{20}$$

$$\mathbf{e}^{(k)} = \mathbf{x}_1 - \rho^{(k)} \mathbf{x}_2. \tag{21}$$

The steps of the proposed algorithm is summarized in Tab. 1.

Table 1: Summary of proposed algorithm

(i)	Find the initial estimation $\hat{\rho}$ as 1;
(ii)	Compute $\mathbf{Q}^{(k)}$ using (17)-(21);
(iii)	Update $\hat{\rho}^{(k)}$ using (16);
(iv)	Repeat Steps (ii)--(iii) until the relative error $\frac{\lvert\hat{\rho}^{(k)} - \hat{\rho}^{(k-1)}\rvert}{\lvert\hat{\rho}^{(k-1)}\rvert} < \epsilon$ with ϵ being the tolerance;
(v)	Obtain the frequency estimate, denoted by $\hat{\omega}$, using $\omega = \angle\hat{\rho}$, where $\angle\cdot$ is the angle in $[-\pi, \pi]$

3 Simulation results

Computer simulations have been conducted to assess the estimation performance of the proposed scheme. The mean square error (MSE) and bias of $\hat{\omega}$ are employed as the performance metrics, defining as $E\left\{(\omega-\hat{\omega})^2\right\}$ and $\omega-E\left\{\hat{\omega}\right\}$, respectively. The Cramer-Rao lower bound (CRLB) [Kozick and Sadler (2016)] for $\hat{\omega}$ is included as the benchmark while the comparison with the GWLP, and ℓ_p-GWLP are also provided with $p=1.4$ [Chen and So (2016)]. The proposed method, GWLP and ℓ_p-GWLP employ the same stopping criterion when the tolerance $\varepsilon=10^{-6}$ is reached. The signal is generated according to (2) with $A=1$, $\omega=1.25$ and $\theta=0.5$. It is worth pointing out that since the second-order power of the $S\alpha S$ model diverges, the geometric SNR (GSNR) [Gonzalez and Gonzalez (2006)] is employed to produce different values of γ. The GSNR is defined as.

$$GSNR=\frac{A^2}{\sigma C_g^{\frac{1}{\alpha}-1}} \tag{22}$$

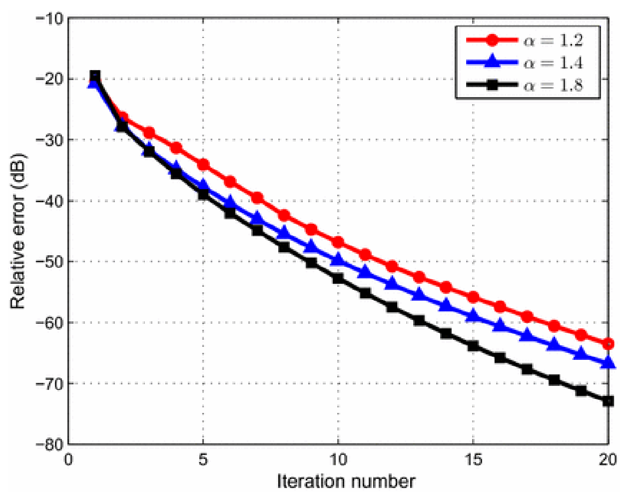

Figure 1: Relative error *vs.* iteration number at GSNR=15 dB

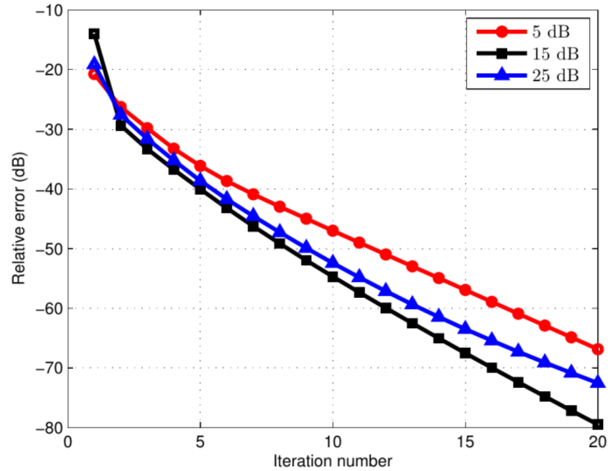

Figure 2: Relative error *vs.* iteration number at α =1

where $C_g \approx 1.78$ and $\sigma = \gamma^{\frac{1}{\alpha}}$ All our results are averages of 5000 Monte Carlo trials with a data length of N=50.

First of all, the convergence of the proposed algorithm and the other two methods are investigated. Fig. 1 shows the relative error *vs.* iteration number for different values of α at GSNR=15 dB, while those in Figs. 2 and 3 study the environments of different GSNR and different p with $\alpha = 1$, respectively. It is indicated in Figs. 1, 2 and 3 that the proposed method always converges and it has a faster convergence rate, when GSNR, α and p becomes larger.

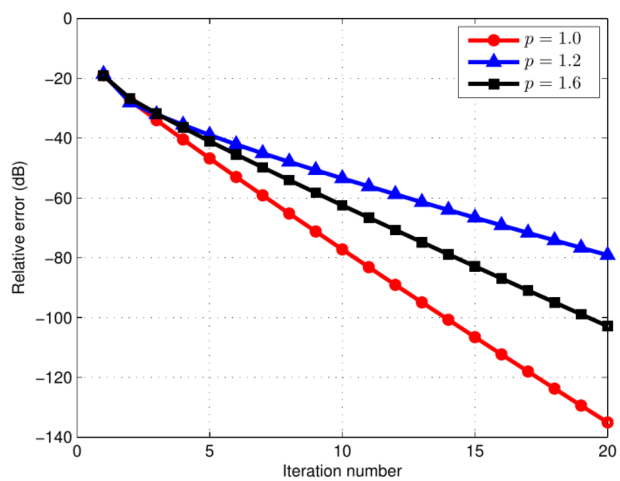

Figure 3: Relative error *vs.* iterations at α =1 and GSNR=15 dB

Second, we study the frequency estimation performance for different GSNR conditions with $\alpha = 1$. Fig. 4 plots the MSE of $\hat{\omega}$ *vs.* GSNR. It is seen that in the presence of $S\alpha S$ noise, the ℓ_1-GWLP significantly outperforms the GWLP and ℓ_p-GWLP methods due to the smaller gap between the MSE of $\hat{\omega}$ and CRLB. In the case of the lower GSNR, say, GSNR<10 dB, the proposed method performs better than the other two schemes. In Fig. 5, the corresponding frequency biases are shown. It can be observed that the biases of the ℓ_1-GWLP, GWLP and ℓ_p-GWLP were negligible at sufficiently high GSNRs, i.e., GSNR>15 dB, indicating that the unbiasedness of all three frequency estimation methods.

Finally, we study the performance for different α and data length.

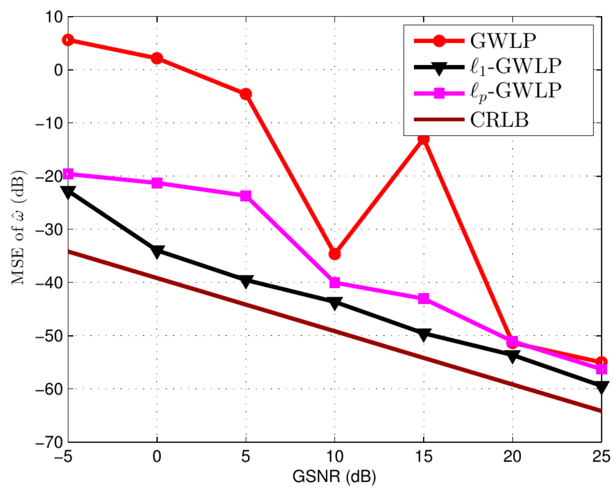

Figure 4: MSE of estimated frequency *vs.* GSNR at α =1

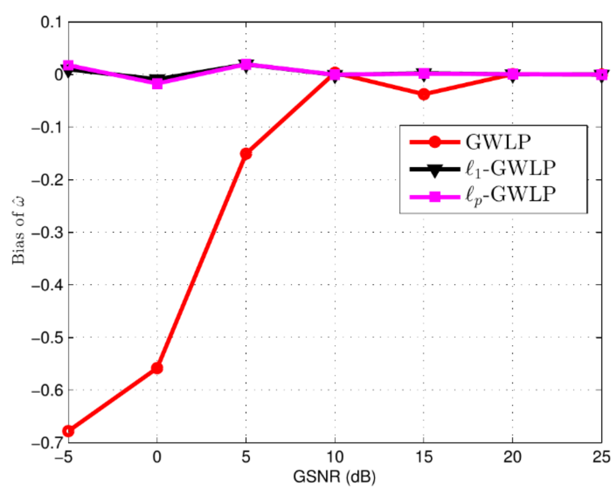

Figure 5: Bias of frequency *vs.* GSNR

The MSEs of frequency *vs.* α are plotted in Fig. 6, where the parameters are same with the previous test with GNSR=15 dB. It is observed that the ℓ_1-GWLP performs best among the GWLP and ℓ_p-GWLP, indicating the robustness of our algorithm. Fig. 7 shows the MSE of $\hat{\omega}$ *vs*. data length at GSNR=15 dB. It is seen that the ℓ_1-GWLP is again superior to the other two algorithms. Furthermore, the GWLP method is not stable even in large data length, verifying its sensitivity to outliers. Note that the corresponding bias of frequency is not included here because they were similar to those in the second test.

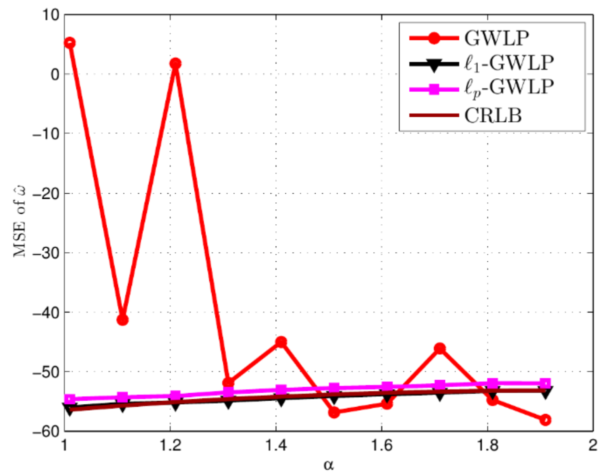

Figure 6: MSE of estimated frequency *vs.* α at GSNR=15 dB

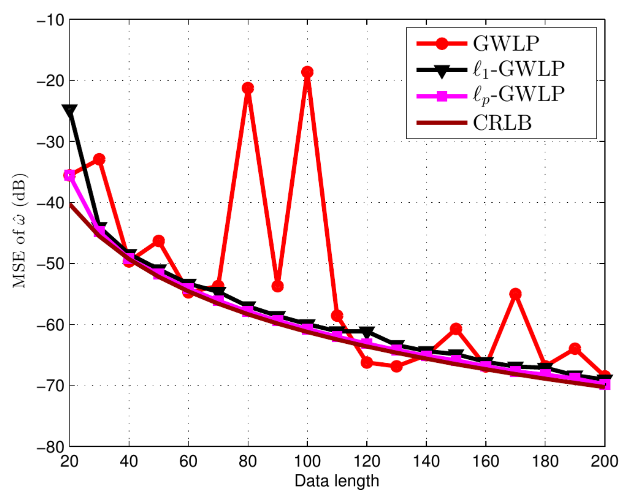

Figure 7: MSE of estimated frequency *vs.* N at GSNR=15 dB

In summary, in the scenarios of different α, the ℓ_1-GWLP is robust and nearly optimal. At the lower GSNR or smaller data length, the proposed method still performs better than all other estimators. Moreover, simulation results indicate that our algorithm is not very sensitive to the estimated value of α .

4 Conclusion

In this paper, with the use of the ℓ_1-norm minimization on the whitened linear prediction errors, a robust and nearly optimal frequency estimator, namely, ℓ_1-GWLP, is devised, for the single complex sinusoid in the presence of the $S\alpha S$ noise. Simulation results indicate that the ℓ_1-GWLP can resist outliers and is superior to the original GWLP, and ℓ_p-GWLP estimators. It is worth to point out that although the single complex tone is taken as an illustration in this paper, our method can also be extended to the multiple complex-valued and real-valued scenarios.

References

Byrd, R.; Payne, D. (1979): Convergence of the IRLS Algorithm for Robust Regression. Technical Report 313. John Hopkins University, Baltimore, Maryland, United States.

Chen, C.; So, H. O. (2016): Variance analysis of unbiased least ℓ_p-norm *estimator* in non-Gaussian noise. *Signal Processing*, vol. 122, pp. 190-203.

Chan, F. K. W.; So, H. C.; Huang, L. T. (2015): Parameter estimation and identifiability in bistatic multiple-input multipleoutput radar. *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, no. 3, pp. 2047-2056.

Chen, Y.; Yang, X. L.; Huang, L. T. (2018): Robust frequency estimation in symmetric alpha-stable noise. *Computers, Materials & Continua*, vol. 54, no. 2, pp.103-136.

Gonzalez, J. G.; Paredes, J. L. (2006): Zero-order statistics: a mathematical framework for the processing and characterization of very impulsive signals. *IEEE Transactions on Signal Processing*, vol. 54, no. 10, pp. 3839-3851.

Kay, S. M. (1993): *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, NJ: Englewood Cliffs.

Kessy, A.; Lewin, A.; Strimmer, K. (2015): Optimal whitening and decorrelation. *Statistics*, pp. 1-5.

Kozick, R. J.; Sadler, B. M. (2016): Maximum-Likelihood array processing in non-Gaussian noise with Gaussian mixtures. *IEEE Transactions on Signal Processing*, vol. 48, no. 12, pp. 3520-3535.

Laguna-Sanchez, G.; Lopez-Guerrero, M. (2014): An experimental study of the effect of human activity on the alpha-stable characteristics of the power-line noise. pp. 6-11*Proceedings of the 18th IEEE International Symposium on Power Line Communications and Its Applications*, Glasgow, United Kingdom.

Merle, G.; Spath, H. (1974): Computational experiences with discrete ℓ_p-Approximation. *Computing*, vol. 12, no. 4, pp. 315-321.

Nikias, C. L.; Shao, M. (1995): *Signal Processing with Alpha-Stable Distribution and Applications*. John Wiley Sons Inc., New York.

Qian, C.; Huang, L.; So, H. C. (2016): Enhanced PUMA for direction-of-arrival estimation and its performance analysis. *IEEE Transactions on Signal Processing*, vol. 63, no. 20, pp. 5358-5368.

Stoica, P.; Moses, R. (2005): *Spectral Analysis of Signals*, Prentice-Hall, NJ: Upper Saddle River.

So, H. C.; Chan, F. K. W.; Ho, K. C. (2005): Linear prediction approach for efficient frequency estimation of multiple realsinusoids: algorithms and analyses. *IEEE Transactions on Signal Processing*, vol. 53, no. 7, pp. 2290-2305.

So, H. C.; Chan, F. K. W. (2006): A generalized weighted linear predictor frequency estimation approach for a complex sinusoid. *IEEE Transactions on Signal Processing*, vol. 54, no. 4, pp. 1304-1315.

Shao, M.; Nikias, C. L. (1993): Signal processing with fractional lower order moments: stable processes and their applications. *IEEE Transactions on Signal Processing*, vol. 81, pp. 986-1010.

Tsakalides, P. (1995): *Array Signal Processing with Alpha Stable Distributions*. *Ph.D.* dissertation, University of South California, Los Angeles, CA.

Wu, C. R. (2018): Time optimization of multiple knowledge transfers in the big data environment. *Computers, Materials & Continua*, vol. 54, no. 3, pp. 269-285.

Zoubir, A. M.; Koivunen, V. (2012): Robust estimation in signal processing: a tutorial-style treatment of fundamental concepts. *Circuits, Systems Signal Processing*, vol. 37, no. 10, pp. 4637-4650.