A NEW ALPHA AND GAMMA BASED MIXTURE APPROXIMATION FOR HEAVY-TAILED RAYLEIGH DISTRIBUTION

Mohammadreza Hassannejad Bibalan, Hamidreza Amindavar

Department of Electrical Engineering, Amirkabir University of Technology, Tehran, Iran.
mr.hassannezhad@aut.ac.ir , hamidami@aut.ac.ir

ABSTRACT

In this paper, a novel bi-parameter mixture approximation for modeling the statistical properties of *heavy-tailed* Rayleigh distribution is developed. Heavy-tailed Rayleigh distribution is basically the amplitude PDF of bi-variate isotropic α-stable distribution which appears in the envelope distribution of impulsive interferences and SAR return signals. Closed-form expression for PDF is a serious consideration in signal processing with α-stable distribution. In this case, mixture approximation by means of Rayleigh and a generalization of Cauchy distribution is a simple but comprehensive approach. Several models try to give a model for the mixture ratio based on α. We propose two new methods for mixture ratio estimation based on both α and γ. Numerical results verify that the proposed mixture approximation is closer to the theoretical and empirical heavy-tailed Rayleigh distribution.

***Index Terms*—** mixture approximation, heavy-tailed Rayleigh distribution, symmetric α-stable, characteristic exponent, dispersion.

1. INTRODUCTION

Stable distributions emerge innately in the study of heavy-tailed distributions[1]. A generalization of the central limit theorem says that the sum of a number of i.i.d random processes with thick tails and infinite variances have a stable distribution as the number of summands increases . The precise modeling of the properties of SAR data plays a critical role in statistical signal processing [2]. Stable law is a well-validated and more conventional model which can also illustrate impulsive and skewed behavior [3]. A comprehensive but simple model with less imposed burden for *SαS* distributions is interpreted in [4]. In [2] a generalization of the Rayleigh model for SAR images which inherently is a Rayleigh-mixture model is introduced. Furthermore, it proposes the negative order moments as a parameter estimation technique with more versatility in SAR signal processing compared to conventional Rayleigh, Weibull, and k-distribution. Recently, a closed-form mixture approximation is introduced for the envelope distribution of the complex α-stable SAR texture and speckle components [5]. Previous models [6] consider only the role of α in mixture ratio estimation, whereas our two new models are based on α and γ.

The paper is organized as follows. In Section 2, we provide some necessary preliminaries on α-stable processes and heavy-tailed Rayleigh distribution and mixture approximation used to describe its properties. In Section 3, we present our mixture approximation, which includes two novel parameter estimation methods based on α and γ.In Section 4, the performance of our proposed approximation is evaluated and compared with the performance of the existing methods. Finally, we conclude in Section 5.

2. HEAVY-TAILED RAYLEIGH MODEL

2.1. α-Stable Distribution

α-stable PDF is procured by taking the inverse Fourier transform of its characteristic function [7]; however, it does not provide an analytic formula. α-stable distribution, $S(\alpha, \gamma, \beta, \mu)$, is completely determined by four parameters.α is the *characteristic exponent* and it determines the shape of the distribution, ($0 < \alpha \leq 2$), γ is the *dispersion* of the distribution and plays a similar role to the variance of the Gaussian distribution to specify the expansion of the distribution around its *location*, ($\gamma > 0$), β is the index of skewness ($-1 \leq \beta \leq 1$) and μ is the *location* parameter, ($\mu \epsilon \mathbb{R}$). The case of $\beta = 0$ corresponds to the *Symmetric α-Stable* (*SαS*) distribution. In this paper a special case of zero-mean *SαS* distribution is considered, which is defined in terms of its characteristic function, as the following,

$$\varphi_{\alpha,\gamma}(\boldsymbol{\omega}) = \exp(-\gamma|\boldsymbol{\omega}|^{\alpha}). \tag{1}$$

2.2. Heavy-Tailed Rayleigh Distribution

Below is called in the literature as the heavy-tailed Rayleigh distribution [8].

$$f_{\alpha,\gamma}(r) = r\int_0^{\infty} \omega \exp(-\gamma\omega^{\alpha}) J_0(r\omega) d\omega. \tag{2}$$

This new generalized form of the Rayleigh distribution can illustrate impulsive behaviour and has thicker tails rather than the classical Rayleigh distribution. Therefore, the noise-free SAR return signal amplitude PDF can be modeled by heavy-tailed Rayleigh distribution. Analytical formula [8] only exists for two special value of α. The case $\alpha = 2$, as the following,

$$f_{2,\gamma}(r) = \frac{r}{2\gamma}\exp(-\frac{r^2}{4\gamma}), \quad (3)$$

which is basically the classical Rayleigh distribution. The other case is $\alpha = 1$,

$$f_{1,\gamma}(r) = \frac{\gamma r}{(r^2+\gamma^2)^{\frac{3}{2}}}, \quad (4)$$

which is corresponds to the Cauchy distribution.

2.3. Mixture Approximation Based on α

Since the $S\alpha S$ random variable representation as a scale mixture of the Gaussian r.v. is traditional, it suggests that a mixture model may be useful. Therefore, its PDF can be approximated by a finite Gaussian Mixture Model (GMM), but an accurate Gaussian mixture approximation demands several terms [9]. In [5], a mixture approximation for the amplitude PDF of isotropic α-stable distribution is given as the following,

$$f^{app}_{\alpha,\gamma}(r) = (1-\varepsilon_\alpha)f_{1,\gamma}(r) + \varepsilon_\alpha f_{2,\gamma}(r). \quad (5)$$

ε_α, is the mixture ratio, a *concave* function of characteristic exponent, $\varepsilon_\alpha : [1, 2] \to [0, 1]$. This mixture model has several appealing properties: it requires only 2 parameters and it can capture the algebraic tail as well as the mode; it is tractable and conventional Expectation Maximization (EM) algorithms can be used with simple parameterizations. If the number of samples is small, the conditional ML estimate of γ can be obtained via polynomial rooting. Some more accurate estimate of ε_α are provided as follows [4],[7].

1. McCulloch equation:

$$\varepsilon_\alpha = \alpha - 1, \quad (6)$$

2. LM-1 method:

$$\varepsilon_\alpha = 2\Big(\frac{\alpha-1}{\alpha}\Big), \quad (7)$$

3. LM-2 method:

$$\varepsilon_\alpha = \frac{4}{3}\Big(\frac{\alpha^2-1}{\alpha^2}\Big), \quad (8)$$

4. FLOM method:

$$\varepsilon_\alpha = \frac{\Gamma(1-\frac{q}{\alpha}) - \Gamma(1-\frac{q}{2})}{\Gamma(1-q) - \Gamma(1-\frac{q}{2})}, -2 < q < \alpha, \quad (9)$$

where, q denotes the q^{th}-order moment. These four methods are only based on the α and they connive the effect of γ in mixture ratio estimation problem, which is the disadvantage of these methods. Also, these methods are based on the moment and log-moment of $S\alpha S$ distribution, whereas the heavy-tailed Rayleigh distribution is the amplitude of complex $S\alpha S$ distribution.

3. PROPOSED MIXTURE APPROXIMATION BASED ON α , γ

In this section, we propose a new mixture approximation for modelling the PDF in (2), which its mixture ratio is a function of α and γ (i.e. $\varepsilon_{\alpha,\gamma}$). Its PDF is given by,

$$f^{app}_{\alpha,\gamma}(r) = (1-\varepsilon_{\alpha,\gamma})f_{1,\gamma}(r) + \varepsilon_{\alpha,\gamma} f_{2,\gamma}(r). \quad (10)$$

It can be seen that the for $\alpha = 1$, $\varepsilon_{\alpha,\gamma} = 0$ and for $\alpha = 2$, $\varepsilon_{\alpha,\gamma}$ yields to 1. Notice that, the density given in (10) can be expressed equivalently in Fourier domain as characteristic function. We will confine our attention to zero-mean bi-variate isotropic α-stable distribution. Our entire goal is $f^{app}_{\alpha,\gamma}(r) = f_{\alpha,\gamma}(r)$. Using (1),(2) and (10) we have,

$$\exp(-\gamma|\boldsymbol{\omega}|^\alpha) = (1-\varepsilon_{\alpha,\gamma})\exp(-\gamma|\boldsymbol{\omega}|) + \varepsilon_{\alpha,\gamma}\exp(-\gamma|\boldsymbol{\omega}|^2), \quad (11)$$

where,

$$\boldsymbol{\omega} = (\omega_1, \omega_2) = \omega_1 + j\omega_2. \quad (12)$$

3.1. PDF in the Origin Method

Two different approaches has been utilized to estimating the mixture ratio. The first basis comes from the value of PDF in the origin, which must be the same on both sides of (10) and we named it *PDF in the Origin (PO)* method. Equivalently we can write,

$$f^{app}_{\alpha,\gamma}(0) = (1-\varepsilon_{\alpha,\gamma})f_{1,\gamma}(0) + \varepsilon_{\alpha,\gamma} f_{2,\gamma}(0). \quad (13)$$

The value of a function at 0 is just the integral of its Fourier transform in frequency domain. Hence,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp[-\gamma(\omega_1^2+\omega_2^2)^{\alpha/2}]d\omega_1 d\omega_2 =$$
$$(1-\varepsilon_{\alpha,\gamma})\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp[-\gamma(\omega_1^2+\omega_2^2)^{\frac{1}{2}}]d\omega_1 d\omega_2 +$$
$$\varepsilon_{\alpha,\gamma}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp[-\gamma(\omega_1^2+\omega_2^2)]d\omega_1 d\omega_2. \quad (14)$$

If we make the substitution $\omega_1 = \rho\cos(\theta), \omega_2 = \rho\sin(\theta)$, and integrating over θ, we obtain,

$$2\pi\int_0^\infty \rho\exp(-\gamma\rho^\alpha)d\rho = 2\pi(1-\varepsilon_{\alpha,\gamma})\int_0^\infty \rho\exp(-\gamma\rho)d\rho$$
$$+ 2\pi\varepsilon_{\alpha,\gamma}\int_0^\infty \rho\exp(-\gamma\rho^2)d\rho \quad (15)$$

These integrals can be simply calculated. So,

$$\frac{\Gamma(\frac{2}{\alpha})}{\alpha\gamma^{\frac{2}{\alpha}}} = \frac{1-\varepsilon_{\alpha,\gamma}}{\gamma^2} + \frac{\varepsilon_{\alpha,\gamma}}{2\gamma}, \tag{16}$$

where, $\Gamma(\cdot)$ represents the Gamma function. Eventually, the mixture ratio is obtained as the following,

$$\varepsilon_{\alpha,\gamma} = \frac{2[\gamma^2\Gamma(\frac{2}{\alpha}) - \alpha\gamma^{\frac{2}{\alpha}}]}{\alpha\gamma^{\frac{2}{\alpha}}[\gamma-2]}, \gamma \neq 2. \tag{17}$$

3.2. PDF Energy Method

The second method considers the equality of probability density function energy on both sides of (10). We also named it as *PDF-Energy (PE)* method. According to the Parseval's theorem, the total energy stored in r domain, is equal to the total energy stored in ω domain. Therefore, one can be written,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \|\varphi_\alpha(\boldsymbol{\omega})\|^2 d\omega_1 d\omega_2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \|(1-\varepsilon_{\alpha,\gamma})\varphi_1(\boldsymbol{\omega}) + \varepsilon_{\alpha,\gamma}\varphi_2(\boldsymbol{\omega})\|^2 d\omega_1 d\omega_2. \tag{18}$$

The right side of the equation contains three terms. For simplicity, this equation can best be described as,

$$A = B(1-\varepsilon_{\alpha,\gamma})^2 + C\varepsilon_{\alpha,\gamma}^2 + 2D\varepsilon_{\alpha,\gamma}(1-\varepsilon_{\alpha,\gamma}) \tag{19}$$

where, A,B,C and D are some coefficients related to characteristic function, as the following,

$$\begin{aligned}
A &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp[-2\gamma(\omega_1^2+\omega_2^2)^{\alpha/2}] d\omega_1 d\omega_2 \\
&= 2\pi\int_0^{\infty} \rho\exp(-2\gamma\rho^\alpha)d\rho = \frac{2\pi\Gamma(\frac{2}{\alpha})}{\alpha(2\gamma)^{\frac{2}{\alpha}}}, \\
B &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp[-2\gamma(\omega_1^2+\omega_2^2)^{\frac{1}{2}}] d\omega_1 d\omega_2 \\
&= 2\pi\int_0^{\infty} \rho\exp(-2\gamma\rho)d\rho = \frac{\pi}{2\gamma^2}, \\
C &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp[-2\gamma(\omega_1^2+\omega_2^2)] d\omega_1 d\omega_2 \\
&= 2\pi\int_0^{\infty} \rho\exp(-2\gamma\rho^2)d\rho = \frac{\pi}{2\gamma}, \\
D &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp[-\gamma(\omega_1^2+\omega_2^2+(\omega_1^2+\omega_2^2)^{\frac{1}{2}})] d\omega_1 d\omega_2 \\
&= 2\pi\int_0^{\infty} \rho\exp[-\gamma(\rho^2+\rho)]d\rho \\
&= \pi\int_0^{\infty} \Big\{\exp(-\gamma y) - \exp[-\gamma(y^2+y)]\Big\} dy \\
&= \frac{\pi}{\gamma} - \frac{\pi}{2}\sqrt{\frac{\pi}{\gamma}}\exp(\frac{\gamma}{4})[1-\Phi(\frac{\sqrt{\gamma}}{2})],
\end{aligned} \tag{20}$$

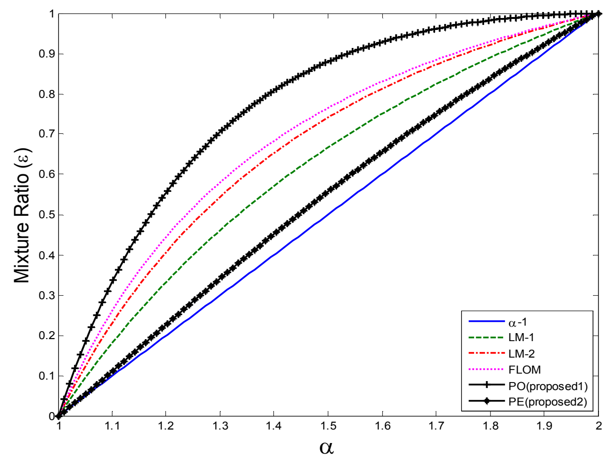

Fig. 1. Mixture ratio versus $\alpha \in [1,2]$, $\gamma = 1.5$, $q = 0.5$.

where $\Phi(\cdot)$ is the *error* function. Substituting in (19), we have a quadratic equation which is a polynomial equation of degree two.

$$(B+C-2D)\varepsilon_{\alpha,\gamma}^2 - 2(B-D)\varepsilon_{\alpha,\gamma} + (B-A) = 0. \tag{21}$$

We need to analyze separately an extra step to descent roots from two to one. The only acceptable root is,

$$\varepsilon_{\alpha,\gamma} = \frac{(B-D)+\sqrt{D^2+A(B+C-2D)-BC}}{B+C-2D} \tag{22}$$

Note that, ε_α in (6),(7),(8) is a concave function of α, whereas ε_α in (9) may be a concave or convex function based on the values of α, q. Furthermore, $\varepsilon_{\alpha,\gamma}$ in (17),(22) is also a concave or convex function based on the values of α, γ. In Fig.1 the behaviour of mixture ratio versus the different values of α is demonstrated for previous and two new models.

4. SIMULATION RESULTS

In this section, we present simulation results obtained by our proposed methods (PO,PE) based on the mixture approximation of heavy-tailed Rayleigh distribution with those obtained using $\alpha - 1$, LM-1, LM-2 and FLOM.

4.1. KS-distance

If the data come from a distribution function which has a closed form, then it is quite natural to estimate the unknown parameters by fitting the empirical PDF points to the theoretical distribution function or its approximation and using the criteria of KS-distance, as the following,

$$D_{KS} = \sup_r \|F_{\alpha,\gamma}^{emp}(r) - F_{\alpha,\gamma}^{app}(r)\|, \tag{23}$$

where, sup is the supremum ,$F_{\alpha,\gamma}(\cdot)$ and $F_{\alpha,\gamma}^{emp}(\cdot)$ denote the cumulative distribution function (CDF) and the empirical

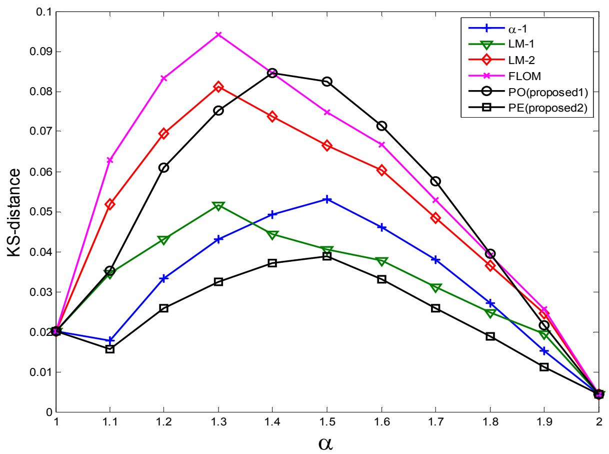

Fig. 2. KS-distance versus $\alpha \in [1, 2]$, $\gamma = 2.5$, $q = 0.5$.

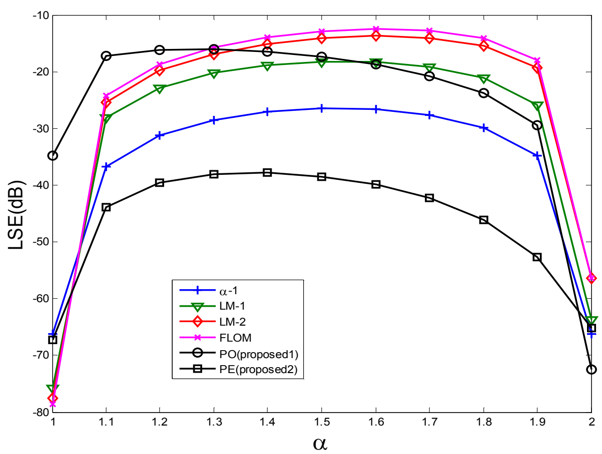

Fig. 3. LSE versus $\alpha \in [1, 2]$, $\gamma = 2.5$, $q = 0.5$.

CDF respectively. We generate $N = 20,000$ i.i.d samples of the $S\alpha S$ random variables with $\gamma = 2.5$ for different values of $\alpha \in [1, 2]$. Next, we evaluate the empirical CDF with mixture approximation and calculate respective KS-distances. The simulation is repeated 100 times and the best order of moment for FLOM method was $q = 0.5$. Fig.2 shows that the PE method has the minimum distance.

4.2. Estimation of α

Now consider the estimation of α when γ is known. In this case the Least Squared estimate of α can be obtained by,

$$\hat{\alpha}_{LS} = \arg\min_{\alpha} |f_{\alpha,\gamma}(r) - f^{app}_{\alpha,\gamma}(r)|^2. \tag{24}$$

The parameters used in the simulation are: $\gamma = 2.5$, $\alpha \in [1, 2]$ and $q = 0.5$. For visually assessing the quality of our proposed method we show the obtained results in Fig.3. Compared with the other methods, the parameter estimated by PE method has the minimum LS error except at $\alpha = 1, 2$.

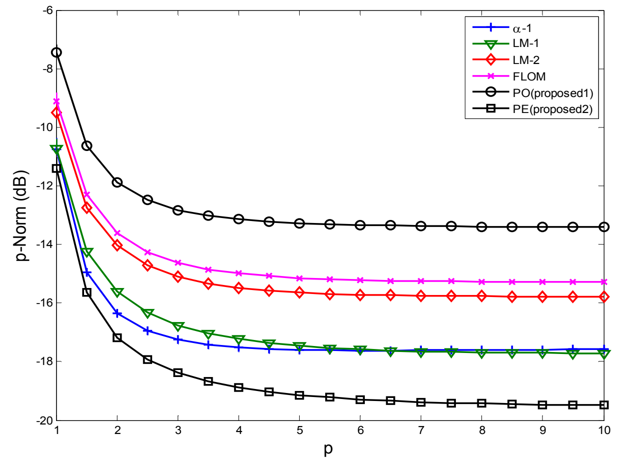

Fig. 4. p-Norm versus different values of p. $\alpha = 1.5$, $\gamma = 1.5$, $q = 0.5$.

4.3. p-Norm Error of Approximated PDF

As a representative comparison, the accuracy of our proposed methods is quantified with p-Norm error between $f_{\alpha,\gamma}(r)$ and $f^{app}_{\alpha,\gamma}(r)$. p-Norm is given by,

$$||f_{\alpha,\gamma}(r) - f^{app}_{\alpha,\gamma}(r)||_p = \left(\int_0^{\infty} \left| f_{\alpha,\gamma}(r) - f^{app}_{\alpha,\gamma}(r) \right|^p dr \right)^{\frac{1}{p}}. \tag{25}$$

We apply the proposed parametric methods for mixture ratio estimation with $\gamma = 2.5$, $q = 0.5$ and $\alpha = 1.5$. In Fig.4 the results of simulation for six different method are demonstrated for $1 \le p \le 10$. In this case, the PE method has the best and the PO method has the worst performance.

5. CONCLUSION

The problem of mixture approximation for heavy-tailed Rayleigh distribution was considered. This paper assumed two new models for mixture ratio estimation based on the α and γ. Comparing the performance of all the estimators, it was observed that the PE method for the mixture ratio estimation had better numerical results regarding to the another methods.

6. REFERENCES

[1] C.L. Nikias and M. Shao, *Signal processing with α-stable distributions and applications*, Wiley, 1995.

[2] G. Moser, J. Zerubia, and S.B. Serpico, "SAR amplitude probability density function estimation based on a generalized Gaussian model," *Image Processing, IEEE Transactions on*, vol. 15, no. 6, pp. 1429–1442, June 2006.

[3] W.J. Zeng, H.C. So, and A.M. Zoubir, "An lp-norm minimization approach to time delay estimation in impulsive noise," *Digital Signal Processing*, vol. 23, no. 4, pp. 1247 – 1254, 2013.

[4] X.T. Li, J. Sun, L.W. Jin, and M. Liu, "Bi-parameter CGM model for approximation of α-stable PDF," *Electronics Letters*, vol. 44, no. 18, pp. 1096–1097, August 2008.

[5] X. Li, S. Wang, L. Fan, and X. Lei, "Mixture approximation to the amplitude statistics of isotropic α-stable clutter," *Signal Processing*, vol. 99, pp. 86–91, January 2014.

[6] P.R. Hill, A.M. Achim, D.R. Bull, and M.E. Al-Mualla, "Dual-tree complex wavelet coefficient magnitude modelling using the bivariate Cauchy-Rayleigh distribution for image denoising," *Signal Processing*, vol. 105, no. 0, pp. 464 – 472, 2014.

[7] Xinya Ma and C.L. Nikias, "Parameter estimation and blind channel identification in impulsive signal environments," *Signal Processing, IEEE Transactions on*, vol. 43, no. 12, pp. 2884–2897, Dec 1995.

[8] E.E. Kuruoglu and J. Zerubia, "Modeling SAR images with a generalization of the Rayleigh distribution," *Image Processing, IEEE Transactions on*, vol. 13, pp. 527–533, April 2004.

[9] A Swami, "Non-Gaussian mixture models for detection and estimation in heavy-tailed noise," in *Acoustics, Speech, and Signal Processing, 2000. ICASSP '00. Proceedings. 2000 IEEE International Conference on*, 2000, vol. 6, pp. 3802–3805.