Parameter Estimation of K Distribution Based on Second-Kind Statistics

Zengguo Sun
School of Electronic and Information Engineering
Xi'an Jiaotong University
Xi'an, Shaanxi, China
duffer2000@163.com

Chongzhao Han
School of Electronic and Information Engineering
Xi'an Jiaotong University
Xi'an, Shaanxi, China
czhan@mail.xjtu.edu.cn

Xianghui Yuan
School of Electronic and Information Engineering
Xi'an Jiaotong University
Xi'an, Shaanxi, China
yuanxianghui@gmail.com

Abstract - The parameters of K distribution are estimated in this paper, and the log-cumulant estimator is proposed based on second-kind statistics. The performance of the log-cumulant estimator is tested on the Monte Carlo simulations. Parameter estimation results demonstrate that the log-cumulant estimator leads to high estimation accuracy for the small values of the shape parameter.

Keywords: K distribution, parameter estimation, second-kind statistics, log-cumulant estimator.

1 Introduction

The K distribution has been used extensively to represent both sea clutter and land clutter, and it contains the classical Rayleigh distribution as a special case [1-4]. With two parameters (shape parameter and scale parameter), the K distribution can fit the experimental data better than the one-parameter distributions such as Rayleigh [2]. For example, the Rayleigh distribution usually applies when the radar resolution cell is large so that it contains many scatterers, with no one scatterer dominant. However, when the resolution cell size and the grazing angle are small, the K distribution can describe the heavy tails of the clutter more precisely compared to the Rayleigh [2]. In order to use the K distribution in practical applications, its parameters should be estimated accurately. In this paper, the log-cumulant estimator is proposed for the K distribution based on second-kind statistics, which rely on the Mellin transform [5, 6]. We test the performance of the log-cumulant estimator on Monte Carlo simulations. Parameter estimation results from Monte Carlo simulation demonstrate that, for the small values of the shape parameter, which correspond to the severely impulsive samples, the log-cumulant estimator yields high estimation accuracy.

This paper is organized as follows. The K distribution and its simulation method are introduced in Section 2. In order to estimate the parameters of K distribution, the log-cumulant estimator is proposed based on second-kind statistics in Section 3, and its performance is tested on the Monte Carlo simulations in Section 4. Lastly, this paper is concluded in Section 5.

2 K distribution and its simulation

The K distribution has the following probability density function (pdf) [3]

$$f(x) = \frac{4b^{(\nu+1)/2}x^{\nu}}{\Gamma(\nu)} K_{\nu-1}\left(2x\sqrt{b}\right), \tag{1}$$

where ν ($\nu > 0$) is the shape parameter which determines the impulsiveness of the distribution, b ($b > 0$) is the scale parameter, $\Gamma(\cdot)$ is the Gamma function, and $K_{\nu-1}(\cdot)$ is the modified Bessel function of the second kind of order $\nu - 1$. Denoting X as a K-distributed random variable with parameters ν and b, we define a new random variable as

$$Y = \sqrt{b}X. \tag{2}$$

From equation (1), the pdf of Y can be written as

$$f_Y(x) = \frac{4x^{\nu}}{\Gamma(\nu)} K_{\nu-1}(2x). \tag{3}$$

Obviously, the random variable Y follows the K distribution with the shape parameter ν and the unit scale parameter ($b = 1$). In other words, the K-distributed random variable X can be normalized by (2). For various values of ν, the pdf of K distribution is plotted in Fig. 1. Obviously, the value of ν controls the shape of the pdf. The smaller the shape parameter ν is, the spikier the pdf is. It should be noted that the K distribution reduces to the classical Rayleigh distribution when the shape parameter ν tends to infinity [2].

The K distribution is usually described as a compound model that consists of two components: the fast varying speckle component and the slowly varying radar cross section (RCS) component [1, 2]. If we combine these two components, the observation can be expressed as the following product model

$$I = n \cdot \sigma \,, \tag{4}$$

where I is the observed intensity or amplitude, n is the speckle component, and σ is the RCS component. Denoting f_I, f_n and f_σ as the pdfs of the observation, the speckle component, and the RCS component, respectively, we have

$$\begin{aligned} f_I(I) &= \int_0^{+\infty} f_{I|\sigma}(I|\sigma) f_\sigma(\sigma)\, d\sigma \\ &= \int_0^{+\infty} \frac{1}{\sigma} f_n\left(\frac{I}{\sigma}\right) f_\sigma(\sigma)\, d\sigma \end{aligned} . \tag{5}$$

If the speckle component n is modeled with the Rayleigh distribution with the pdf

$$f_n(n) = 2ne^{-n^2} \,, \tag{6}$$

and the RCS component σ is modeled with the Nakagami distribution (Gamma distribution in amplitude format) with the pdf

$$f_\sigma(\sigma) = \frac{2b^\nu \sigma^{2\nu-1} e^{-b\sigma^2}}{\Gamma(\nu)} \,, \tag{7}$$

then the pdf of the observation can be obtained

$$f_I(I) = \frac{4b^{(\nu+1)/2} I^\nu}{\Gamma(\nu)} K_{\nu-1}\left(2I\sqrt{b}\right). \tag{8}$$

Obviously, the equation (8) is just the K distribution whose pdf is shown in (1). Therefore, the K distribution is a composite model, which assumes a Rayleigh-distributed rapidly fluctuating speckle component modulated by a slowly fluctuating Gamma-distributed RCS component [2]. This property can be used for the simulation of K distribution. For various values of the shape parameter ν, the K distributed samples are simulated in Fig. 2. Apparently, the smaller the value of ν is, the more impulsive the K distributed samples are. Since the K distributed samples can be simulated, we can use the Monte Carlo simulation to evaluate the performance of the parameter estimator.

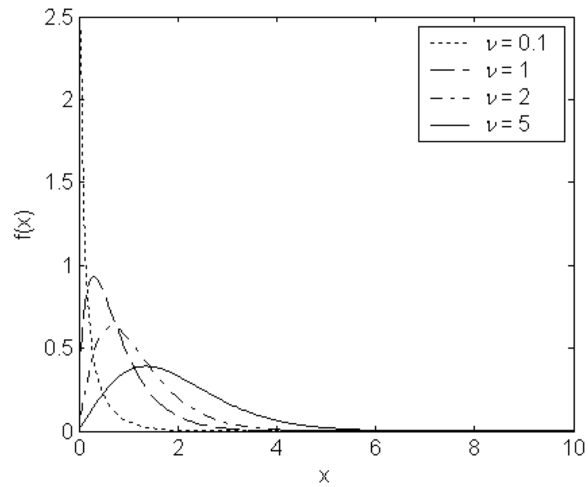
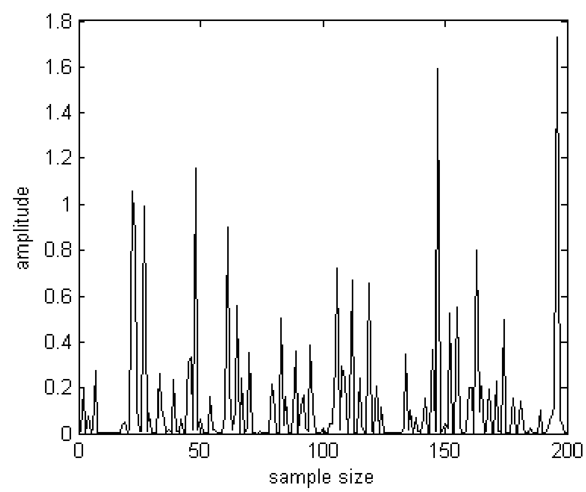
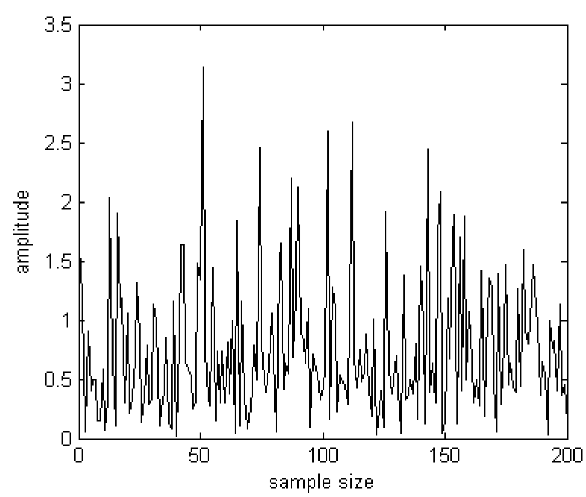


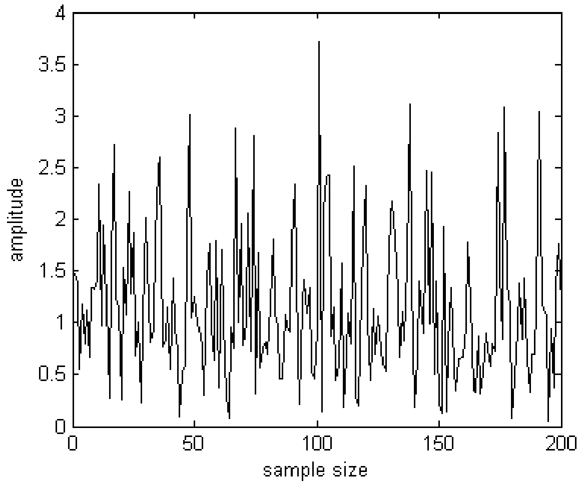
Fig. 1 Pdfs of K distribution ($b = 1$)



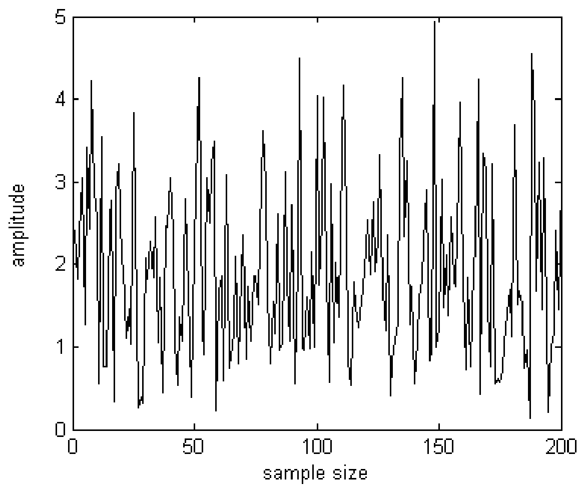
(a) $\nu = 0.1$



(b) $\nu = 1$



(c) $\nu = 2$



(d) $\nu = 5$

Fig. 2 K distributed samples ($b = 1$ and the number of samples is 200)

3 Log-cumulant estimator

In [5], the log-cumulant estimator was used for parameter estimation of the α - stable positive distributions due to its explicit expressions. The log-cumulant estimator is based on second-kind statistics, which rely on the Mellin transform. Denoting g as a function defined over $[0,+\infty)$, its Mellin transform is defined as

$$\mathrm{M}\left[g(x)\right](s) = \int_0^{+\infty} x^{s-1} g(x)\,dx\,, \tag{9}$$

where s is the complex variable of the transform. Specifically, for a pdf f defined in $[0,+\infty)$, analogous to the case of common statistics based on Fourier transform, the second-kind statistic functions are defined as follows [5, 6]:

- Second-kind first characteristic function

$$\Phi(s) = \int_0^{+\infty} x^{s-1} f(x)\,dx \tag{10}$$

- Second-kind second characteristic function

$$\Psi(s) = \log\left(\Phi(s)\right) \tag{11}$$

- r th order second-kind moment (log-moment)

$$\tilde{m}_r = \left.\frac{d^r \Phi(s)}{ds^r}\right|_{s=1} \tag{12}$$

- r th order second-kind cumulant (log-cumulant)

$$\tilde{k}_r = \left.\frac{d^r \Psi(s)}{ds^r}\right|_{s=1} \tag{13}$$

The first two log-cumulants $\tilde{k}_1$ and $\tilde{k}_2$ can be estimated empirically from N samples y_i as follows [6]:

$$\hat{\tilde{k}}_1 = \frac{1}{N}\sum_{i=1}^{N}\left[\log(y_i)\right], \tag{14}$$

$$\hat{\tilde{k}}_2 = \frac{1}{N}\sum_{i=1}^{N}\left[\left(\log(y_i) - \hat{\tilde{k}}_1\right)^2\right]. \tag{15}$$

By substituting (1) into (10) and after some manipulation, the second-kind first characteristic function of K distribution is given by

$$\Phi(s) = \frac{b^{(1-s)/2}\Gamma\left((s+1)/2\right)\Gamma\left(\nu + (s-1)/2\right)}{\Gamma(\nu)}\,. \tag{16}$$

Then, by substituting (16) into (11), the second-kind second characteristic function of K distribution is given by

$$\begin{aligned}\Psi(s) = &\frac{(1-s)\log(b)}{2} + \log\left(\Gamma\left(\frac{s+1}{2}\right)\right) \\ &+ \log\left(\Gamma\left(\nu + \frac{s-1}{2}\right)\right) - \log\left(\Gamma(\nu)\right)\end{aligned}\,. \tag{17}$$

Lastly, substituting (17) into (13), the log-cumulant estimator (i.e., the first and second orders log-cumulants) for the K distribution is obtained as follows:

$$\tilde{k}_1 = -\frac{\log(b)}{2} - \frac{C_e}{2} + \frac{\psi(\nu)}{2}, \tag{18}$$

$$\tilde{k}_2 = \frac{\pi^2}{24} + \frac{\psi(1,\nu)}{4}. \tag{19}$$

Here, C_e is the Euler's constant, $\psi(\cdot)$ is the Digamma function, and $\psi(1,\cdot)$ is the Trigamma function, i.e., the first-order derivative of the Digamma function. By replacing the actual log-cumulants with the sample log-cumulants (equations (14) and (15)), the shape parameter ν can be estimated from (19) using some numerical optimization techniques such as bisection [7], and then the scale parameter b can be obtained readily from (18).

4 Monte Carlo simulation

The log-cumulant estimator for K distribution was tested on Monte Carlo simulations for various true values of the shape parameter ν. The K distributed samples were simulated independently by using the method introduced in Section 2, and the number of samples is 10000. For each true parameter ν, the Monte Carlo simulation experiment was repeated 100 times independently, and then the average and standard deviation of the estimates were computed. The results are shown in Table 1 with standard deviations in parentheses. Fig. 3 shows the performance of the log-cumulant estimator as a function of the true values of ν. Obviously, the log-cumulant estimator leads to high estimation accuracy for the small values of ν (e.g., $\nu = 0.1$). However, the fluctuation of the estimates is apparent as the true value of ν becomes larger.

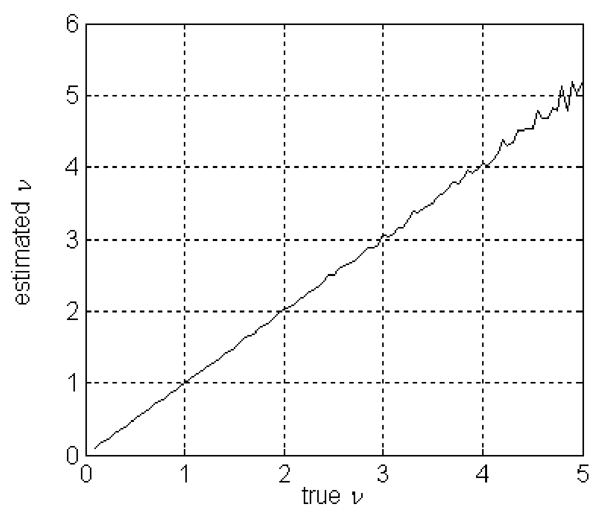
The performance of the log-cumulant estimator can be explained by analyzing the Trigamma function $\psi(1,\nu)$ in (19). We define the right-hand side of (19) as the following function

$$y = \frac{\pi^2}{24} + \frac{\psi(1,\nu)}{4}, \tag{20}$$

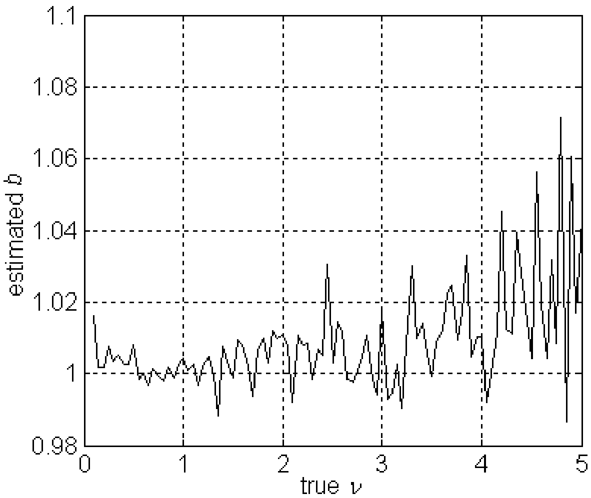
and this function is plotted in Fig. 4. Obviously, when the values of ν are small, the function y decreases rapidly as ν increases. In other words, large values of y easily determine the accurate values of ν. When estimating the parameter ν from (20), y is replaced by the sample log-cumulant $\hat{\tilde{k}}_2$ that is calculated from (15). For a given value of ν, we simulate 10000 K-distributed samples, calculate the sample log-cumulant $\hat{\tilde{k}}_2$, and estimate the parameter ν from (20). The results are shown in Table 2 for various true values of ν. Apparently, for the small values of ν (e.g., $\nu = 0.1$), the K pdf is spiky and the corresponding samples are severely impulsive. Therefore, the sample log-cumulant $\hat{\tilde{k}}_2$ is large, which leads to the accurate estimate of ν. If the larger values are chosen for the ν (e.g., $\nu = 5$), the impulsiveness of the K distributed samples becomes weaker, resulting in the smaller values of $\hat{\tilde{k}}_2$ and subsequently the inaccurate estimate of ν. In a word, due to the appearance of the Trigamma function, the performance of the log-cumulant estimator relies on the shape parameter ν. For the small values of ν, the log-cumulant estimator can achieve high performance.

Table 1 Monte Carlo simulation of log-cumulant estimator (true $b = 1$)

True Value	$\hat{\nu}$	$\hat{b}$
$\nu = 0.1$	0.0999 (0.0015)	1.0020 (0.1048)
$\nu = 0.2$	0.2003 (0.0027)	1.0029 (0.0538)
$\nu = 0.5$	0.5002 (0.0090)	1.0041 (0.0349)
$\nu = 1$	0.9995 (0.0224)	1.0017 (0.0336)
$\nu = 2$	2.0032 (0.1102)	1.0020 (0.0651)
$\nu = 5$	5.0866 (0.7474)	1.0203 (0.1621)

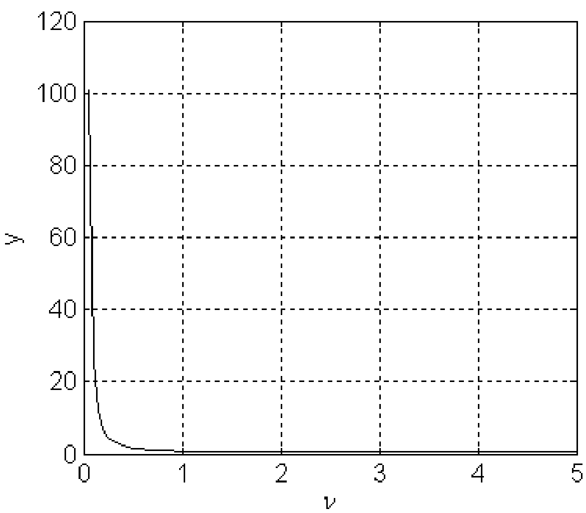


(a) $\hat{\nu}$ varying with true ν

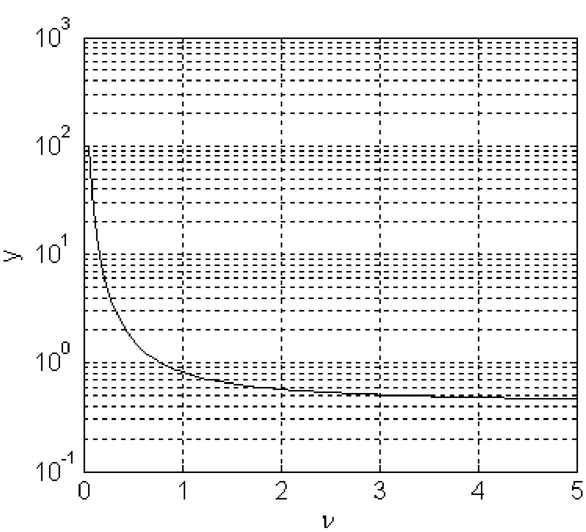


(b) $\hat{b}$ varying with true ν

Fig. 3 Performance of log-cumulant estimator as a function of true ν (true $b=1$)



(a) Linear scale



(b) Logarithmic scale

Fig. 4 Function $y=\pi^2/24+\psi(1,\nu)/4$ versus ν

Table 2 Sample log-cumulants $\hat{\tilde{k}}_2$ and estimated values of ν (true $b=1$)

True Value	$\hat{\tilde{k}}_2$	$\hat{\nu}$
$\nu=0.1$	25.6568	0.1002
$\nu=0.2$	6.9887	0.1998
$\nu=0.5$	1.6484	0.4992
$\nu=1$	0.8191	1.0057
$\nu=2$	0.5727	1.9977
$\nu=5$	0.4602	5.5910

5 Conclusion

In this paper, the log-cumulant estimator based on the second-kind statistics is proposed to estimate the parameters of K distribution, and its performance is tested on Monte Carlo simulations. For the small values of shape parameter, which correspond to the severe impulsiveness of the K distributed samples, the log-cumulant estimator leads to high estimation accuracy due to the appearance of the Trigamma function. In this work, we test the log-cumulant estimator on the simulated data and we emphasize the experimental demonstration. In the future research, we will test the log-cumulant estimator on the real radar data and images, and we should further evaluate the log-cumulant estimator theoretically. In addition, we will compare the log-cumulant estimator with other estimators such as the methods of moments.

Acknowledgments

This work is supported by the State Key Development Program for Basic Research of China (Grant No. 2007CB311006), the National High Technology Research and Development Program of China (Grant No. 2006AA01Z126), the National Natural Science Foundation of China (Grant No. 60602026), and the Specialized Research Fund for the Doctoral Program of Higher Education of China (Grant No. 20070698002).

References

[1] C. Oliver, and S. Quegan, *Understanding Synthetic Aperture Radar Images*, Artech House, Boston, 1998.

[2] M. I. Skolnik, *Introduction to Radar Systems, Third Edition*, McGraw-Hill, New York, 2001.

[3] K. D. Ward, R. J. A. Tough, and S. Watts, *Sea Clutter: Scattering, the K Distribution and Radar Performance*, The Institution of Engineering Technology, London, UK, 2006.

[4] J. B. Billingsley, A. Farina, F. Gini, M. V. Greco, and L. Verrazzani, *Statistical analyses of measured radar ground clutter data*, IEEE Trans. Aerospace and Electronic Systems, vol. 35, no. 2, pp. 579-593, Apr. 1999.

[5] J. M. Nicolas, *Alpha-stable positive distributions: a new approach based on second kind statistics*, Proceedings of the European Signal Processing Conference, Toulouse, France, Sept. 2002, pp. 197-200.

[6] C. Tison, J. M. Nicolas, F. Tupin, and H. Maitre, *A new statistical model for Markovian classification of urban areas in high-resolution SAR images*, IEEE Trans. Geoscience and Remote Sensing, vol. 42, no. 10, pp. 2046-2057, Oct. 2004.

[7] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, Cambridge University Press, Cambridge, UK, 1994.