# Correspondence

## A Complex Generalized Gaussian Distribution—Characterization, Generation, and Estimation

Mike Novey, Tülay Adalı, *Fellow, IEEE*, and Anindya Roy

***Abstract*—The generalized Gaussian distribution (GGD) provides a flexible and suitable tool for data modeling and simulation, however the characterization of the complex-valued GGD, in particular generation of samples from a complex GGD have not been well defined in the literature. In this correspondence, we provide a thorough presentation of the complex-valued GGD by: i) constructing the probability density function (pdf); ii) defining a procedure for generating random numbers from the complex-valued GGD; and iii) implementing a maximum likelihood estimation (MLE) procedure for the shape and covariance parameters in the complex domain. We quantify the performance of the MLE with simulations and actual radar data.**

***Index Terms*—Complex-valued signal processing, generalized Gaussian distribution (GGD), maximum likelihood estimation (MLE).**

### I. Introduction

The generalized Gaussian distribution (GGD) has found wide use in modeling various physical phenomena in the signal processing community. For example, the GGD has been used to model synthetic aperture radar [1] and echocardiogram [2] images; features for face recognition [3]; load demand in power systems [4]; and subband signals in images [5]. The use of the GGD has also found utility in independent component analysis (ICA) where it is used as a flexible class of source density models (see e.g., [6]–[8]).

The GGD family of densities [9] is obtained by generalizing the Gaussian density to provide a variable rate of decay and is given by

$$p_X(x;\sigma,c) = \frac{c}{2\sigma\Gamma(1/c)} e^{-(\frac{|x|}{\sigma})^c}$$

where $\Gamma(\cdot)$ is the Gamma function, $\sigma$ is the standard deviation, and $c$ is the shape parameter. What makes the GGD appropriate in so many applications is its flexible parametric form which adapts to a large family of symmetric distributions, from super-Gaussian to sub-Gaussian including specific densities such as Laplacian $(c = 1)$ and Gaussian $(c = 2)$. Although the GGD has found wide use, most applications employ the univariate version. A bivariate GGD is introduced in [10] and used in modeling a video coding scheme in [11] along with a maximum likelihood estimate (MLE) for the shape parameter based on a chi-square test, however, the covariance matrix estimate presented is not an MLE. Complex-valued GGD models have been described much less frequently and assume that the signal is circular, i.e., invariant to rotation, as in [1], [6]—both papers use an MLE for estimating the shape parameter.

Manuscript received April 01, 2009; accepted October 02, 2009. First published November 06, 2009; current version published February 10, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Alfred Hanssen.

The authors are with the University of Maryland Baltimore County, Catonsville, MD 21228 USA (e–mail: mnovey1@umbc.edu; adali@umbc.edu; anindya@umbc.edu).

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

Digital Object Identifier 10.1109/TSP.2009.2036049

In this correspondence, we extend the results for the complex normal distribution defined in [12] and [13] to the GGD, by providing a fully complex distribution denoted as CGGD, given in (7). As in the univariate case, the CGGD adapts to a large family of bivariate symmetric distributions, from super-Gaussian to sub-Gaussian including specific densities such as Laplacian and Gaussian distributions. Since the CGGD is also a member of the elliptically symmetric distributions, the normalized kurtosis values of the real and imaginary parts of a complex random variable are a scaled version of the complex kurtosis where the scale factor is nonnegative and is a function of noncircularity as shown in [14]. Since the kurtosis of the complex Gaussian is zero, as in the real-valued case, positive normalized kurtosis values imply a super-Gaussian distribution, i.e., a sharper peak with heavier tails, and negative normalized kurtosis values imply sub-Gaussian distributions.

Recently, use of the full second-order statistics of complex random variables, namely the information in the commonly used covariance as well as the pseudocovariance matrices [12], [15], have proven useful in signal processing. The second-order statistics are used to classify the variable as second-order circular or second-order noncircular, i.e., a complex-valued random variable is second-order circular if the pseudocovariance matrix is zero. Due to the recent interest in incorporating the circular/noncircular properties of complex-valued signals, the CGGD is not restricted to the circular case but is also parameterized by the second moment thus providing a means of varying the noncircularity of the distribution. To enhance the usefulness of the CGGD, we also provide a method for generating samples from a CGGD as well as an MLE for its shape and covariance parameters. We provide simulations to quantify the MLE's performance and then test on actual radar data.

### II. Complex Preliminaries

A complex variable $z$ is defined in terms of two real variables $z_R$ and $z_I$ as $z = z_R + j z_R$ where $j = \sqrt{-1}$ and alternately as the bivariate vector $\mathbf{z}_b = [z_R, z_I]^T$. It is also convenient to work with the augmented vector defined in [13] and [16] as $\mathbf{z}_a = [z, z^*]^T$ where $\mathbf{z}_a = \begin{bmatrix} 1 & j \\ 1 & -j \end{bmatrix} \mathbf{z}_b$.

Similarly, a complex random variable is defined as $Z = Z_R + jZ_I$ along with the bivariate $\mathbf{Z}_b = [Z_R, Z_I]^T$ and augmented $\mathbf{Z}_a = [Z, Z^*]^T$ vector forms. Assuming that $E\{Z\} = 0$, the bivariate covariance matrix is, thus

$$\mathbf{C}_b = E\left\{\mathbf{Z}_b \mathbf{Z}_b^T\right\} = \begin{bmatrix} \sigma_R^2 & \rho \\ \rho & \sigma_I^2 \end{bmatrix}$$

and the augmented covariance matrix is

$$\begin{aligned} \mathbf{C}_a &= E\left\{\mathbf{Z}_a \mathbf{Z}_a^H\right\} \\ &= \begin{bmatrix} \sigma_R^2 + \sigma_I^2 & \left(\sigma_R^2 - \sigma_I^2\right) + j2\rho \\ \left(\sigma_R^2 - \sigma_I^2\right) - j2\rho & \sigma_R^2 + \sigma_I^2 \end{bmatrix} \end{aligned}$$

where $\sigma^2$ is the variance and $\rho$ is the correlation $E\{Z_R Z_I\}$. It is clear that the augmented covariance matrix will have real-valued diagonal elements $\mathbf{C}_{a(0,0)} = \mathbf{C}_{a(1,1)}$ and complex-valued off-diagonal elements $\mathbf{C}_{a(1,0)} = \mathbf{C}^*_{a(0,1)}$. For $Z$ to be second-order circular, $\mathbf{C}_{a(1,0)} = \mathbf{C}^*_{a(0,1)} = 0$, i.e., the variance of $Z_R$ and $Z_I$ are the same and $Z_R$ and $Z_I$ are uncorrelated. A measure of second-order noncircularity [17], [18] is $|E\{Z^2\}|/E\{ZZ^*\}$ with bounds $0 \leq |E\{Z^2\}|/E\{ZZ^*\} \leq 1$ and $|E\{Z^2\}| = 0$ indicates circular data. A stronger definition of circularity, *strict circularity*, is based on the probability density function

of the complex random variable such that for any $\alpha$, the pdf of $Z$ and $e^{j\alpha}Z$ are the same [15].

We note the following identities from [13] and [19] between the bivariate and augmented covariances:

$$\begin{aligned} \mathbf{z}_b^T \mathbf{C}_b^{-1} \mathbf{z}_b &= \mathbf{z}_a^H \mathbf{C}_a^{-1} \mathbf{z}_a \\ 2\sqrt{|\mathbf{C}_b|} &= \sqrt{|\mathbf{C}_a|} \\ 2\,\mathrm{eigenvals}(\mathbf{C}_b) &= \mathrm{eigenvals}(\mathbf{C}_a) \end{aligned} \tag{1}$$

where $|\cdot|$ is the determinant and $\mathrm{eigenvals}(\cdot)$ are the eigenvalues.

## III. CGGD Construction

### A. Pdf Derivation

In this section, we derive the CGGD by first constructing the bivariate pdf and then extending this distribution to the complex case with a general augmented covariance matrix, i.e., for noncircular data. We note that the bivariate GGD is defined in [11], however the construction shown here provides insight into how the random variables are generated in Section III-B.

We begin by defining a complex random variable $Z = Re^{j\Theta}$ that is a function of two random variables $R$ and $\Theta$. The magnitude $R$, is a modified Gamma variate defined as $R = G^{1/q}$ where $G \sim \mathrm{Gamma}(2/q, 1)$ is a gamma-distributed random variable with shape parameter $2/q$ and unit scale and $\Theta \sim \mathcal{U}(0, 2\pi)$ has a uniform distribution.

Before finding the pdf of $Z$, we first find the pdf of $R$ starting with the univariate gamma distribution defined as

$$p_G(x, 2/q) = \frac{x^{(2/q-1)}}{\Gamma(2/q)} e^{-x}$$

where $\Gamma$ is the gamma function. The Gamma random variable is then raised to the $(1/q)$th power resulting in the pdf of $R$ given by

$$p_R(r) = \frac{qr}{\Gamma(2/q)} e^{-r^q} \tag{2}$$

where we used the transform of a random variable to a power, i.e., if $R = X^{1/q}$ then $p_R(r) = qr^{q-1} p_X(r^q)$.

The complex variable $Z = Re^{j\Theta}$ can be rewritten in the bivariate case as $z_R = r\cos(\theta)$ and $z_I = r\sin(\theta)$ with inverses $r = |z| = \sqrt{z_R^2 + z_I^2}$ and $\theta = \mathrm{atan}(z_I/z_R)$. The joint distribution of $\mathbf{Z}_b = [Z_R, Z_I]^T$ is found through the density transform as

$$p_{\mathbf{Z}_b}(\mathbf{z}_b) = \frac{1}{|J|} p_{(R,\Theta)} \left( \sqrt{z_R^2 + z_I^2}, \mathrm{atan}(z_I/z_R) \right)$$

where $p_{(R,\Theta)}$ is the joint distribution of $R$ and $\Theta$ and the determinant of the Jacobian is

$$|J| = \left| \begin{bmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{bmatrix} \right| = r.$$

Noting that $p_{(R,\Theta)}(r, \theta) = p_R(|z|)(1/2\pi)$ due to the independence of $R$ and $\Theta$ and $p_\Theta(\theta) = (1/2\pi)$, our joint distribution becomes

$$p_{\mathbf{Z}_b}(\mathbf{z}_b) = \frac{q}{2\pi\Gamma(2/q)} e^{-(z_R^2+z_I^2)^{q/2}} = \frac{c}{\pi\Gamma(1/c)} e^{-(z_R^2+z_I^2)^c} \tag{3}$$

where we substituted $c = q/2$ in the last line so the pdf is Gaussian when $c = 1$. The expression in (3) is a GGD that is circular due to the invariance to $\theta$ with the variance a function of $c$. For the variable to have unit variance (normalized), we first solve for the second moment, $E\{Z_R^2\} = E\{Z_I^2\}$, with the integral

$$\begin{aligned} E\left\{Z_R^2\right\} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} z_R^2 \frac{c}{\pi\Gamma(1/c)} e^{-(z_R^2+Z_I^2)^c}\,dxdy \\ &= \int_0^{\infty}\int_0^{2\pi} r^2\cos^2(\theta) \frac{c}{\pi\Gamma(1/c)} e^{-(r^2)^c}\; r\,d\theta dr \\ &= \frac{c}{\Gamma(1/c)}\left[\frac{\Gamma(2/c)}{2c}\right] = \frac{\Gamma(2/c)}{2\Gamma(1/c)} \end{aligned} \tag{4}$$

which results in a normalizing term $\eta(c) = (\Gamma(2/c)/2\Gamma(1/c))$ and the second line follows from a rectangular to polar coordinate substitution—we can similarly show $E\{Z_R Z_I\} = 0$ which is expected since the random variable $Z$ is circular. We use this result to normalize the variance of $Z_R$ and $Z_I$ to unity through the linear transform $\mathbf{W}_b = \mathbf{N}\mathbf{Z}_b$ where $\mathbf{N} = \sqrt{1/\eta(c)}\mathbf{I}$, i.e., dividing by the standard deviation. Applying this transform to (3), we obtain

$$p_{W_b}(\mathbf{w}_b) = \frac{1}{|\mathbf{N}|} p_{\mathbf{Z}_b}(\mathbf{N}^{-1}\mathbf{w}_b) = \frac{\beta(c)}{2} e^{-[\eta(c)(\mathbf{w}_b^T\mathbf{w}_b)]^c} \tag{5}$$

where $\beta(c) = (c\Gamma(2/c)/\pi\Gamma(1/c)^2)$ and $E\{\mathbf{W}_b\mathbf{W}_b^T\} = \mathbf{I}$ for $c > 0$. The identities in (1) allow us to rewrite (5) in the complex-augmented form as

$$p_{W_a}(\mathbf{w}_a) = \frac{\beta(c)}{2} e^{-[\eta(c)(\mathbf{w}_a^H\mathbf{w}_a)]^c} \tag{6}$$

where $p_{W_a}(\mathbf{w}_a) = p_{W_b}(\mathbf{w}_b)$. Equation (6) is primarily notational since pdfs are defined with respect to real variables, however, this form allows us to work with probabilistic descriptions directly in the complex domain as described in [20]. Our goal, however, is to have a form with a general augmented covariance matrix. We tailor the covariance matrix by applying a linear transform $\mathbf{T}_a$ to the normalized data through $\mathbf{V}_a = \mathbf{T}_a\mathbf{W}_a$. The covariance matrix is now given by

$$\mathbf{C}_a = E\left\{\mathbf{V}_a\mathbf{V}_a^H\right\} = \mathbf{T}_a E\left\{\mathbf{W}_a\mathbf{W}_a^H\right\}\mathbf{T}_a^H = \mathbf{T}_a\mathbf{I}_a\mathbf{T}_a^H = \mathbf{T}_a\mathbf{T}_a^H.$$

Due to the unique form of the augmented covariance matrix, the diagonal terms are real valued and the off-diagonal terms are conjugates, $\mathbf{T}_a^H = \mathbf{T}_a$. Now given any arbitrary augmented covariance, we find the transform $\mathbf{T}_a$ using the matrix square root, i.e., $\mathbf{T}_a = \sqrt{\mathbf{C}_a}$. The matrix square root can be found using the eigenvalue decomposition of $\mathbf{C}_a$, such that $\mathbf{C}_a = \mathbf{V}^H\mathbf{\Lambda}\mathbf{V}$, where $\mathbf{V}$ is the matrix of eigenvectors and $\mathbf{\Lambda}$ is the diagonal matrix of real-valued eigenvalues due to the Hermitian symmetric properties of the covariance matrix. It is easy to show that $\sqrt{\mathbf{C}_a} = \mathbf{V}^H\sqrt{\mathbf{\Lambda}}\mathbf{V}$ and also $|\mathbf{T}_a| = \sqrt{|\mathbf{C}_a|}$.

Applying the transform $\mathbf{T}_a$ to the pdf (6), we obtain

$$\begin{aligned} p_{\mathbf{V}_a}(\mathbf{v}_a) &= \frac{1}{|\mathbf{T}_a|} p_{W_a}\left(\mathbf{T}_a^{-1}\mathbf{v}_a\right) \\ &= \frac{1}{2|\mathbf{T}_a|}\beta(c) e^{-[\eta(c)(\mathbf{v}_a^H\mathbf{T}_a^{-H}\mathbf{T}_a^{-1}\mathbf{v}_a)]^c} \\ &= \frac{1}{\sqrt{|\mathbf{C}_a|}}\beta(c) e^{-[\eta(c)(\mathbf{v}_a^H\mathbf{C}_a^{-1}\mathbf{v}_a)]^c} \end{aligned} \tag{7}$$

which defines the general CGGD distribution parameterized by the shape $c$ and augmented covariance matrix $\mathbf{C}_a$.

### B. CGGD Generation

To generate CGG distributed samples with pdf (7) using Matlab (www.mathworks.com), we use the same procedure for constructing

the pdf as outlined in Section III-A, i.e., first generate the bivariate normalized random variable, then substitute the augmented form, and lastly apply a transform to yield the desired covariance. Given the shape parameter $c$, where $c = 1$ is Gaussian, and augmented covariance $\mathbf{C}_a$, the following procedure generates $N$ independent complex variables:

1) Generate $n = 1, \ldots, N$ complex samples:

$$z(n) = \text{gamrnd}(1/c, 1)^{1/(2c)} e^{(j2\pi \text{ rand})};$$

2) Normalize the complex variance:

$$w = z/\sqrt{\eta_c(c)} \text{ where } \eta_c(c) = \frac{\Gamma(2/c)}{\Gamma(1/c)};$$

3) Form augmented vector:

$$\mathbf{w}_a = [w, \text{ conj}(w)]^T;$$

4) Calculate transform from $\mathbf{C}_a$ using matrix square root:

$$\mathbf{T}_a = \text{sqrtm}(\mathbf{C}_a);$$

5) Apply transform:

$$\mathbf{v}_a = \mathbf{T}_a \mathbf{w}_a$$

where gamrnd, sqrtm, conj, and rand are Matlab functions.

### C. MLE Estimator for the Covariance $\mathbf{C}_b$ and Shape Parameter $c$

Our approach for estimating the shape parameter and covariance matrix is to use a maximum likelihood approach. Since our parameters are both real valued and complex valued, we choose to work in the real domain for our MLE, i.e., the bivariate vector form. Our starting point is the log of the pdf (7) with $N$ independent and identically distributed samples which results in the likelihood function

$$L(\mathbf{v}_b; \boldsymbol{\phi}) = N\ln(\beta(c)) - \frac{N}{2}\ln(|\mathbf{C}_b|) - \eta^c(c) \sum_{t=1}^{N} \left(\mathbf{v}_b^T(t)\mathbf{C}_b^{-1}\mathbf{v}_b(t)\right)^c \tag{8}$$

where our parameter vector is $\boldsymbol{\phi} = [\sigma_R^2, \sigma_I^2, \rho, c]^T$ and the bivariate vectors are substituted for the augmented vectors. Setting the derivative of (8) to zero does not yield a closed form solution to the parameter vector, and hence a numerical solution is warranted. Our method is the three step procedure:

1) The initial covariance matrix is estimated using the sample covariance $\hat{\mathbf{C}}_b^0 = (1/N)\sum_t \mathbf{v}_b(t)\mathbf{v}_b^T(t)$;
2) The initial shape parameter $\hat{c}^0$ is estimated using a moment estimator as suggested in [5], [21];
3) A Newton–Raphson iteration is used to find the final estimated values $\hat{\boldsymbol{\phi}}$.

We show in the simulations section that this three step procedure provides fast convergence, typically in five steps with an accuracy of $10^{-5}$, over a wide range of parameter values.

In step two, we implement a method of moments estimator prior to the Newton–Raphson iteration to aid in convergence. The moment used in the estimator is the scale-invariant fourth moment term defined as

$$\kappa(v) = \frac{E\left\{Z_R^4\right\}}{E\left\{Z_R^2\right\}^2} + \frac{E\left\{Z_I^4\right\}}{E\left\{Z_I^2\right\}^2}.$$

Using a procedure similar to the one given in (4), we find

$$\kappa(v) = \frac{3\Gamma(1/c)\Gamma(3/c)}{[\Gamma(2/c)]^2}. \tag{9}$$

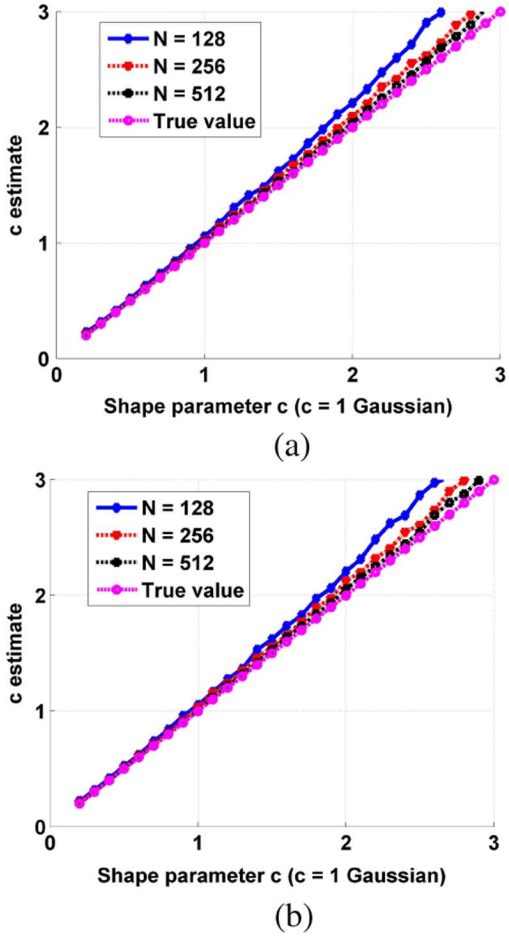


Fig. 1. Shape parameter estimate versus true shape parameter with circular and noncircular data with sample sizes of 128, 256, and 512. (a) Circular. (b) Noncircular $|E\{Z^2\}| = 0.9$.

Our moment estimator then solves for the root of

$$f(c) = \frac{E\left\{Z_R^4\right\}}{E\left\{Z_R^2\right\}^2} + \frac{E\left\{Z_I^4\right\}}{E\left\{Z_I^2\right\}^2} - \frac{3\Gamma(1/c)\Gamma(3/c)}{\Gamma(2/c)^2}$$

resulting in the estimator

$$\hat{c} = \arg\min_c |f(c)| \tag{10}$$

over the domain $c \in [0.1, 4]$ for the simulations.

The moment estimator given in (10) and the sample covariance are not MLEs, however they provide an accurate initial value to the Newton–Raphson iteration defined as

$$\boldsymbol{\phi}^n = \boldsymbol{\phi}^{n-1} - \mathbf{H}^{-1}\nabla$$

where $\nabla = (\partial L/\partial \boldsymbol{\phi})$ is the gradient and $\mathbf{H} = (\partial^2 L/\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}^T)$ is the Hessian matrix evaluated at $\boldsymbol{\phi}^{n-1}$. Both $\nabla$ and $\mathbf{H}$ are derived in the Appendix. The MLE coded in Matlab can be found at http://mlsp.umbc.edu/resources.

## IV. MLE Performance

Simulations, using data generated with the procedure in Section III-B, are used to quantify the performance of the MLE method outlined in Section III-C; the results are the average of

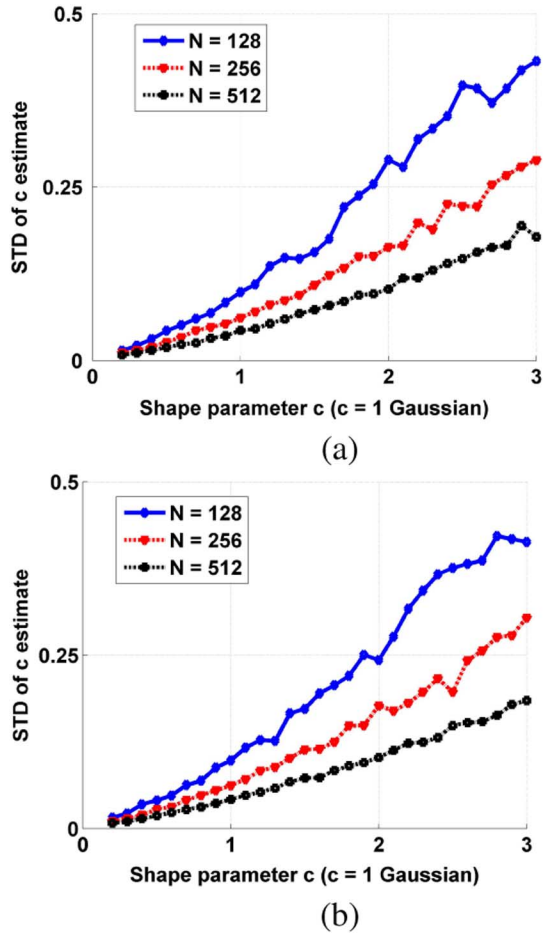


Fig. 2. Shape parameter estimator's standard deviation versus shape parameter with circular and noncircular data. (a) Circular. (b) Noncircular $|E\{Z^2\}| = 0.9$.

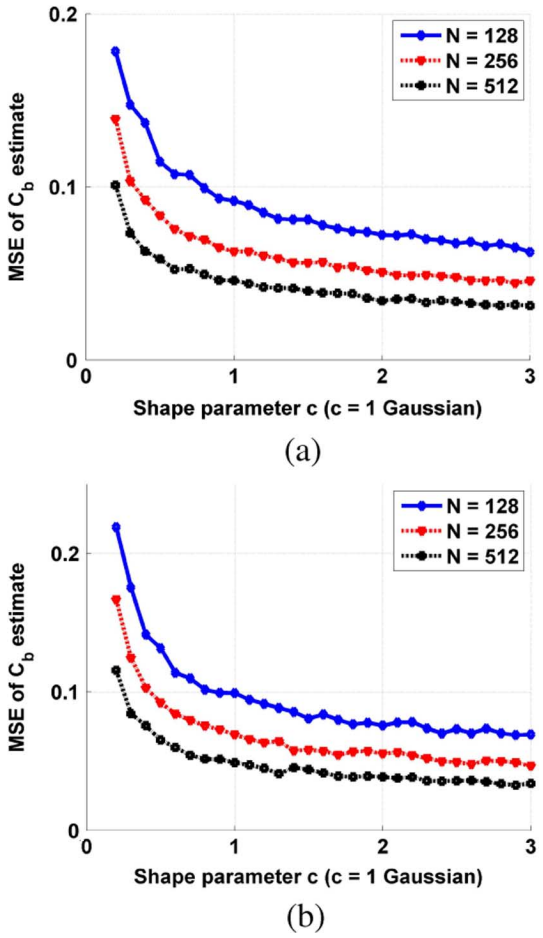


Fig. 3. MSE between $\mathbf{C}_b$ estimate and the true covariance versus shape parameter with circular and noncircular data. (a) Circular. (b) Noncircular $|E\{Z^2\}| = 0.9$.

500 runs. We then test the MLE on actual sea clutter which is a good source of complex-valued data with a nonstationary distribution.

The results of the shape parameter estimator are shown in Figs. 1 and 2. In Fig. 1, we plot the shape parameter estimate versus the true shape parameter with sample sizes of 128, 256, and 512 with circular and noncircular data—the noncircular data has $|E\{Z^2\}| = 0.9$. The results show how well the MLE tracks the true value with only a slight positive bias when $c > 2$ and $N = 128$. Also note that the performance does not degrade with this high value of noncircularity. Fig. 2 depicts the standard deviation of the shape parameter estimate with the same data as the previous figure. As indicated, the standard deviation increases linearly with the shape parameter and is the same for both circular and noncircular data. Fig. 3 depicts the performance of the $\mathbf{C}_b$ estimate by plotting the mean square error (MSE) between the estimate and the true covariance matrix using circular and noncircular data. What the figure shows is that the MLEs performance increases with both sample size and shape parameter with near identical performance using both circular and noncircular data. Fig. 4 compares the performance of the MLE and the sample covariance estimator with $N = 256$ by depicting the MSE of both estimators versus shape parameter. What we glean from the figure is that the MLE shows better performance then the sample covariance estimator as expected, however both estimators perform the same when $c = 1$ since the sample covariance is the MLE for the Gaussian case. Fig. 5 depicts the number of steps for the Newton–Raphson iteration to converge versus shape parameter. As seen in the figure, the MLE converges in about five iterations on average.

Next, we test the MLE on complex-valued sea-clutter data with a small target collected with the McMaster University IPIX radar off the coast of Canada, http://soma.crl.mcmaster.ca/ipix/, see [22] for more details. The data that we are using is from file 19 with radar parameters: X-band, 30 m range resolution, horizontal polarization, and pulse repetition time of $10^{-3}$ seconds. The data is collected in blocks of 256 time samples with adjacent blocks overlapping by 128 samples. Each block is then transformed to the frequency domain which is then tested with the MLE. We test two range gates, one with a small target in clutter and one with clutter only. Fig. 6 depicts the shape parameter estimate for each block with and without a target. What we glean from the figure is that when a target is present the distribution becomes more super-Gaussian as seen by the smaller shape parameter values. This is expected since a target in the frequency domain is a line component causing a heavier tail. The clutter-only data is closer to a Gaussian distribution but still shows areas of low $c$ values, around block 150 for example. This demonstrates the non-stationary nature of sea clutter due to the wind and wave interactions with the sea and the ability of our MLE to follow these fluctuations. This demonstrates the utility of modeling sea clutter with the CGGD.

## V. Conclusion

We introduced a complex-valued generalized Gaussian probability distribution along with a procedure to generate samples from the CGGD. Also presented is a maximum likelihood estimator for the

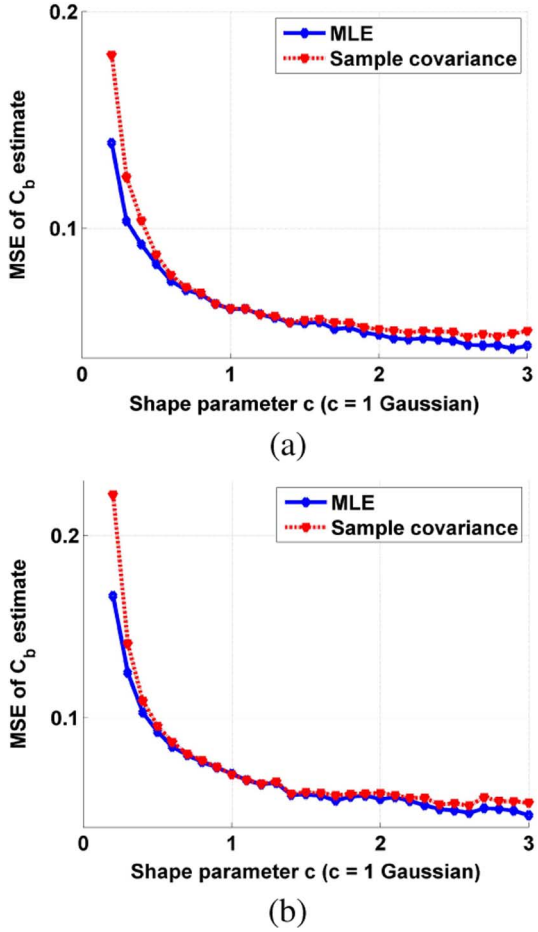


Fig. 4. Mean square error of $\mathbf{C}_b$ estimate using MLE and sample covariance estimator versus shape parameter with $N = 256$. (a) Circular. (b) Noncircular $|E\{Z^2\}| = 0.9$.

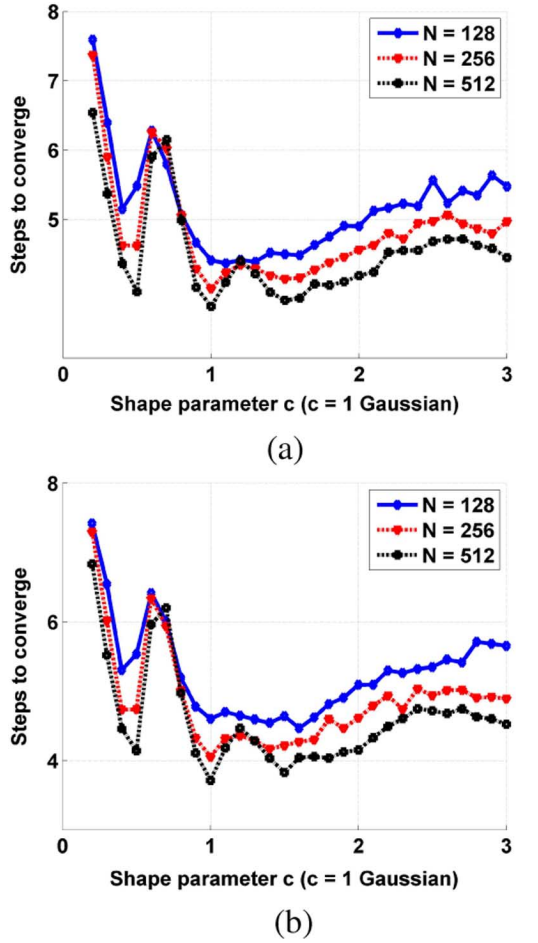


Fig. 5. Number of steps for MLE to converge versus shape parameter with circular and noncircular data. (a) Circular. (b) Noncircular $|E\{Z^2\}| = 0.9$.

shape parameter and covariance matrix using a Newton–Raphson iteration. We show empirically the performance of the MLE on circular and noncircular simulated data and then on complex-valued radar data.

## APPENDIX

### A. Derivation of Gradient of Likelihood Function

We begin with the gradient of (8) with respect to the parameter vector $\boldsymbol{\phi}$ given by

$$\frac{\partial L}{\partial \sigma_R^2} = \frac{-N\sigma_I^2}{2|\mathbf{C}_b|} - \eta^c(c)c\sum_{t=1}^{N}(m(t))^{c-1}m_x(t),$$

$$\frac{\partial L}{\partial \sigma_I^2} = \frac{-N\sigma_R^2}{2|\mathbf{C}_b|} - \eta^c(c)c\sum_{t=1}^{N}(m(t))^{c-1}m_y(t),$$

$$\frac{\partial L}{\partial \rho} = \frac{N\rho}{|\mathbf{C}_b|} - \eta^c(c)c\sum_{t=1}^{N}(m(t))^{c-1}m_\rho(t), \text{ and}$$

$$\frac{\partial L}{\partial c} = \frac{N\beta'(c)}{\beta(c)} - \eta^c(c)\ln(\eta(c))\eta'(c)\sum_{t=1}^{N}m^c(t) - \eta^c(c)\sum_{t=1}^{N}m^c(t)\ln(m(t))$$

where $m(t) = \mathbf{v}_b^T(t)\mathbf{C}_b^{-1}\mathbf{v}_b(t)$, $m_x(t) = (\partial m(t))/(\partial\sigma_R^2) = (y^2 - \sigma_I^2 m(t))/(|\mathbf{C}_b|)$, $m_y(t) = (\partial m(t))/(\partial\sigma_I^2) = (x^2 - \sigma_R^2 m(t))/(|\mathbf{C}_b|)$, $m_\rho(t) = (\partial m(t))/(\partial\rho) = (2\rho m(t) - 2xy)/(|\mathbf{C}_b|)$, $\beta'(c) = (\partial\beta(c))/(\partial c) = (\beta(c))/(c) + (2\beta(c))/(c^2)(\psi(1/c) - \psi(2/c))$, $\eta'(c) = (\partial\eta(c))/(\partial c) = (\eta(c))/(c^2)(\Psi(1/c) - 2\Psi(2/c))$, and $\Psi(\cdot)$ is the digamma function.

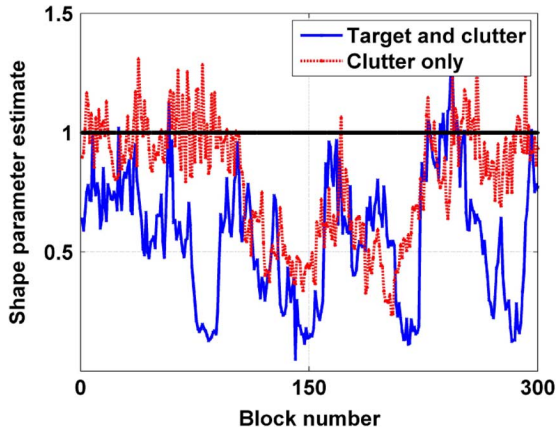


Fig. 6. Shape parameter estimate using the frequency domain sea clutter data with and without a target versus block number.

### B. Derivation of Hessian of Likelihood Function

The terms of the Hessian are the second and cross derivatives of (8) with respect to the parameter vector $\boldsymbol{\phi}$ resulting in

$$\frac{\partial^2 L}{\partial\left(\sigma_R^2\right)^2} = \frac{N\left(\sigma_I^2\right)^2}{2|\mathbf{C}_b|^2} - \eta^c(c)c\sum_{t=1}^{N}\left[(c-1)(m(t))^{c-2}m_x^2(t) + (m(t))^{c-1}m_{xx}(t)\right]$$

$$\frac{\partial^2 L}{\partial (\sigma_I^2)^2} = \frac{N\left(\sigma_R^2\right)^2}{2|\mathbf{C}_b|^2} - \eta^c(c)c\sum_{t=1}^{N}\left[(c-1)(m(t))^{c-2}m_y^2(t) + (m(t))^{c-1}m_{yy}(t)\right]$$

$$\frac{\partial^2 L}{\partial \rho^2} = N\frac{|\mathbf{C}_b| + 2\rho^2}{|\mathbf{C}_b|^2} - \eta^c(c)c\sum_{t=1}^{N}\left[(c-1)(m(t))^{c-2}m_\rho^2(t) + (m(t))^{c-1}m_{\rho\rho}(t)\right]$$

$$\frac{\partial^2 L}{\partial \sigma_R^2 \partial \sigma_I^2} = -N\frac{|\mathbf{C}_b| - \sigma_R^2\sigma_I^2}{2|\mathbf{C}_b|^2} - \eta^c(c)c\sum_{t=1}^{N}\left[(c-1)(m(t))^{c-2} \times m_x(t)m_y(t) + (m(t))^{c-1}m_{xy}(t)\right]$$

$$\frac{\partial^2 L}{\partial \sigma_R^2 \partial \rho} = \frac{-N\sigma_I^2\rho}{|\mathbf{C}_b|^2} - \eta^c(c)c\sum_{t=1}^{N}\left[(c-1)(m(t))^{c-2}m_x(t)m_\rho(t) + (m(t))^{c-1}m_{x\rho}(t)\right]$$

$$\frac{\partial^2 L}{\partial \sigma_I^2 \partial \rho} = \frac{-N\sigma_R^2\rho}{|\mathbf{C}_b|^2} - \eta^c(c)c\sum_{t=1}^{N}\left[(c-1)(m(t))^{c-2}m_y(t)m_\rho(t) + (m(t))^{c-1}m_{y\rho}(t)\right]$$

$$\frac{\partial^2 L}{\partial \sigma_R^2 \partial c} = -\eta^c(c)\left[c\left(\ln(\eta) + \frac{c\eta'(c)}{\eta}\right) + 1\right]\sum_{t=1}^{N}(m(t))^{c-1}m_x(t) - \eta^c(c)c\sum_{t=1}^{N}(m(t))^{c-1}\ln(m(t))m_x(t)$$

$$\frac{\partial^2 L}{\partial \sigma_I^2 \partial c} = -\eta^c(c)\left[c\left(\ln(\eta) + \frac{c\eta'(c)}{\eta}\right) + 1\right]\sum_{t=1}^{N}(m(t))^{c-1}m_y(t) - \eta^c(c)c\sum_{t=1}^{N}(m(t))^{c-1}\ln(m(t))m_y(t)$$

$$\frac{\partial^2 L}{\partial \rho \partial c} = -\eta^c(c)\left[c\left(\ln(\eta) + \frac{c\eta'(c)}{\eta}\right) + 1\right]\sum_{t=1}^{N}(m(t))^{c-1}m_\rho(t) - \eta^c(c)c\sum_{t=1}^{N}(m(t))^{c-1}\ln(m(t))m_\rho(t)$$

and

$$\frac{\partial^2 L}{\partial c^2} = N\left[\frac{-(\beta'(c))^2}{(\beta(c))^2} + \frac{\beta''(c)}{\beta(c)}\right] - \left[\eta^c(c)\ln^2(\eta(c)) - \eta'(c)(\eta(c))^{c-1} - \eta''(c)\eta^c(c)\ln(\eta)\right] \times \sum_{t=1}^{N}m^c(t) + \eta^c(c)\ln(\eta)\eta'(c)\sum_{t=1}^{N}m^c(t)\ln(m(t)) - \eta^c(c)\left[\ln(\eta(c)\sum_{t=1}^{N}m^c(t)\ln(m(t)) + \sum_{t=1}^{N}m^c(t)\ln^2(m(t))\right]$$

where

$$m_{xx}(t) = \frac{\partial m_x(t)}{\partial \sigma_R^2} = \frac{-y^2\sigma_I^2 - \sigma_I^2 m_x(t)|\mathbf{C}_b| + \left(\sigma_I^2\right)^2 m(t)}{|\mathbf{C}_b|^2}$$

$$m_{yy}(t) = \frac{\partial m_y(t)}{\partial \sigma_I^2} = \frac{-x^2\sigma_R^2 - \sigma_R^2 m_y(t)|\mathbf{C}_b| + \left(\sigma_R^2\right)^2 m(t)}{|\mathbf{C}_b|^2}$$

$$m_{\rho\rho}(t) = \frac{\partial m_\rho(t)}{\partial \rho} = \frac{2}{|\mathbf{C}_b|^2}[|\mathbf{C}_b|(m(t) + \rho m_\rho(t)) + 2\rho^2 m(t) - 2xy\rho]$$

$$m_{xy}(t) = \frac{\partial m_x(t)}{\partial \sigma_I^2} = \frac{-y^2\sigma_R^2 - |\mathbf{C}_b|(m(t) + \sigma_I^2 m_y(t)) + \sigma_R^2\sigma_I^2 m(t)}{|\mathbf{C}_b|^2}$$

$$m_{x\rho}(t) = \frac{\partial m_x(t)}{\partial \rho} = \frac{2y^2\rho - \sigma_I^2(m_\rho(t)|\mathbf{C}_b| + 2\rho m(t))}{|\mathbf{C}_b|^2}$$

$$m_{y\rho}(t) = \frac{\partial m_y(t)}{\partial \rho} = \frac{2x^2\rho - \sigma_R^2(m_\rho(t)|\mathbf{C}_b| + 2\rho m(t))}{|\mathbf{C}_b|^2}$$

$$\beta''(c) = \frac{\partial \beta'(c)}{\partial c} = \frac{\beta'(c)c - \beta(c)}{c^2} + \frac{2}{c^4}[\Psi(1/c)(\beta'(c)c^2 - 2c\beta(c)) - \beta(c)\Psi(1,1/c)] - \frac{2}{c^4}[\Psi(2/c)(\beta'(c)c^2 - 2c\beta(c)) - 4\beta(c)\Psi(1,2/c)]$$

and

$$\eta''(c) = \frac{1}{c^4}[(\eta'(c)c^2 - 2c\eta(c))(\Psi(1/c) - 2\Psi(2/c)) - \eta(c)(\Psi(1,1/c) - 2\Psi(1,2/c))].$$

## REFERENCES


[1] M. S. Davis, P. Bidigare, and D. Chang, "Statistical modeling and ML parameter estimation of complex SAR imagery," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Nov. 2007, pp. 500–502.

[2] O. Bernard, J. D'Hooge, and D. Fribouler, "Statistical modeling of the radio-frequency signal in echocardiographic images based on generalized Gaussian distribution," in *Proc. 3rd IEEE Int. Symp. Biomed. Imag.: Nano to Macro*, Arlington, VA, Apr. 2006, pp. 153–156.

[3] D. Gonzalez-Jimenez, F. Perez-Gonzalez, P. Comesana-Alfaro, L. Perez-Freire, and J. L. Alba-Castro, "Modeling Gabor coefficients via generalized Gaussian distributions for face recognition," in *Proc. IEEE Int. Conf. Image Process.*, San Antonio, TX, Oct. 2007, vol. 4, pp. 485–488.

[4] M. O. Mohamed Mahmoud, M. Jaidane-Saidane, J. Souissi, and N. Hizaoui, "Modeling of the load duration curve using the asymmetric generalized Gaussian distribution: Case of the Tunisian power system," in *Proc. IEEE Power and Energy Soc. Gen. Meet.—Convers. Deliv. Elect. Energy in the 21st Century*, Pittsburgh, PA, Jul. 2008, pp. 1–7.

[5] M. Do and M. Vetterli, "Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance," *IEEE Trans. Image Process.*, vol. 11, no. 2, pp. 146–158, Feb. 2002.

[6] M. Novey and T. Adalı, "Adaptable nonlinearity for complex maximization of nongaussianity and a fixed-point algorithm," in *Proc. IEEE Workshop on Machine Learn. Signal Process. (MLSP)*, Maynooth, Ireland, Sep. 2006, pp. 79–84.

[7] J. Cao and N. Murata, "A stable and robust ICA algorithm based on t-distribution and generalized Gaussian distribution models," in *Proc. IEEE Signal Process. Soc. Workshop on Neural Netw. Signal Process.*, Madison, WI, Aug. 1999, pp. 283–292.

[8] Z. Koldovsk, Y. Tichavsk, and E. Oja, "Efficient variant of algorithm fastICA for independent component analysis attaining the Cramer-Rao lower bound," *IEEE Trans. Neural Netw.*, vol. 17, no. 5, pp. 1265–1277, 2006.

[9] N. Johnson and S. Kotz, *Distributions in Statistics—Continuous Univariate Distributions-2*. Boston, MA: Houghton Mifflin , 1970.

[10] T. Taguchi, "On a generalization of Gaussian distribution," *Ann. Inst. Statist. Math.*, vol. 30, pp. 211–242, 1978.

[11] M. Coban and R. Mersereau, "Adaptive subband video coding using bivariate generalized Gaussian distribution model," in *Proc. ICASSP*, Atlanta, GA, May 1996, vol. 4, pp. 1990–1993.

[12] B. Picinbono, "Second-order complex random vectors and normal distributions," *IEEE Trans. Signal Process.*, vol. 44, no. 10, pp. 2637–2640, Oct. 1996.

[13] A. van den Bos, "The multivariate complex normal distribution—A generalization," *IEEE Trans. Inf. Theory*, vol. 41, no. 2, pp. 537–539, Mar. 1995.
[14] E. Ollila and V. Koivunen, "Adjusting the generalized likelihood ratio test of circularity robust to non-normality," in *Proc. IEEE Int. Workshop Signal Process.*, Perugia, Italy, Jun. 2009.
[15] B. Picinbono, "On circularity," *IEEE Trans. Signal Process.*, vol. 42, pp. 3473–3482, Dec. 1994.
[16] P. J. Schreier and L. L. Scharf, "Second-order analysis of improper complex random vectors and processes," *IEEE Trans. Signal Process.*, vol. 51, pp. 714–725, Mar. 2003.
[17] J. Eriksson and V. Koivunen, "Complex random vectors and ICA models: Identifiability, uniqueness, and separability," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 1017–1029, Mar. 2006.
[18] P. Schreier, "The degree of impropriety (noncircularity) of complex random vectors," in *Proc. ICASSP*, Las Vegas, NV, Apr. 2008, pp. 3909–3912.
[19] A. van den Bos, "Complex gradient and Hessian," *Inst. Elect. Eng. Proc. Image Signal Process.*, vol. 141, pp. 380–382, Dec. 1994.
[20] T. Adalı, H. Li, M. Novey, and J.-F. Cardoso, "Complex ICA using nonlinear functions," *IEEE Trans. Signal Process.*, vol. 56, pp. 4536–4544, Sep. 2008.
[21] M. Varanassi and B. Aazhang, "Parametric generalized Gaussian density estimation," *J. Acoust. Soc. Amer.*, vol. 85, pp. 1404–1415, Oct. 1989.
[22] S. Haykin, *Adaptive Radar Signal Processing*. Hoboken, NJ: Wiley, 2007.

# Passive Range-Difference Estimation in a Dispersive Medium

Y. T. Chan, Thomas K. C. Lo, and H. C. So

***Abstract*—Passive source localization can be achieved by utilizing the range-difference (RD) measurements between the source and several spatially separated sensors. When the signal propagation medium is nondispersive, the RD is typically obtained from time-difference-of-arrival (TDOA) between sensor outputs. However, for a dispersive medium which corresponds to in-solid localization, the TDOA information is not reliable. In this correspondence, two RD estimation algorithms for dispersive medium are derived and analyzed, assuming that the dispersion curve is available. Computer simulations are included to contrast the estimation performance of the two methods with the Cramér-Rao lower bound (CRLB) for different conditions.**

***Index Terms*—Dispersive medium, range-difference (RD) estimation, source localization, time-difference-of-arrival (TDOA) estimation.**

## I. Introduction

A conventional approach for passive source localization is to utilize the range-difference (RD) measurements between spatially separated sensors in an array. From the RD information, a set of hyperbolic equations is constructed and the source position can then be calculated with the use of the known sensor positions [1]–[4]. The minimum number of sensors required for two-dimensional positioning is three, and four for the three-dimensional case. When the signal propagation speed is a constant, the RD estimates are obtained by multiplying the velocity with the time-difference-of-arrival (TDOA) measurements, that is, the differences in arrival times between pairs of sensor outputs which receive the signal. There has been extensive research [5]–[8] on TDOA or RD estimation in nondispersive scenarios which have applications in global positioning systems [9], speaker tracking [10] and mobile phone location [11]. When the medium is dispersive, the propagation velocity is a function of frequency and the RDs cannot be obtained directly using the TDOA estimates. This is because different frequency components have different propagation speeds. Important application examples for positioning in a dispersive medium include localization of earthquake epicenters and underground explosions [12], microseismic events in mines [13], and tactile interaction in novel tangible human-computer interfaces [14], [15]. This correspondence is on the development of accurate RD estimation algorithms with known frequency-dependent velocity, denoted by $v(\omega)$ where $\omega$ is the radian frequency. Note that if $v(\omega)$ is not available, its experimental determination consists of placing sensors at a known separation distance $d$, measuring their phase-difference $\phi(\omega)$ of an input variable-frequency sinusoid, and the velocity is then computed as $v(\omega) = \omega d/\phi(\omega)$, based on different frequencies [16], [17].

There is not much literature on TDOA or RD estimation in a dispersive medium. Assuming that the TDOA is a polynomial in $\omega$, [5] finds the TDOA as the value that maximizes a weighted sum of cosine functions. The argument of the cosine is the difference between the measured phase-difference $\phi(\omega)$, and the theoretical, which is $\mathrm{TDOA} \times \omega$. The sum is over all $\omega$ within the signal bandwidth. There is some difficulty with obtaining the optimal weights in [5], and [6] provides an improvement which is applicable when the spectral matrix satisfies an autoregressive model. As our goal is localization with a known $v(\omega)$ [16], [17], the best way is to estimate the RDs directly. The development below gives two RD estimators. The first is nonlinear which corresponds to the maximum likelihood method but requires iteration. The second is linear and thus there is a unique global solution, although its performance is slightly suboptimal.

The rest of the correspondence is organized as follows. In Section II, two RD estimation methods for a pair of sensor outputs are derived and analyzed. Both algorithms operate in the frequency domain. It is proved that their RD variances can attain the optimality benchmark, namely, the Cramér-Rao lower bound (CRLB). Extension to more than two sensors is also studied. Numerical examples are provided in Section III to evaluate the RD estimation performance in the dispersive medium. Finally, conclusions are drawn in Section IV.

## II. Algorithm Development

Consider an array of $L$ sensors at known positions $(x_i, y_i)$, $i = 1, 2, \ldots, L$, receiving signal from a passive source at unknown position $(x, y)$ in a dispersive medium. The signal propagation velocity is a function of frequency, denoted by $v(\omega)$, and is assumed known. Let the distance between the source and the $i$th receiver be $R_i = \sqrt{(x_i - x)^2 + (y_i - y)^2}$. The RD between the $i$th and $j$th receivers, denoted by $\Delta_{ij}$, is then:

$$\begin{aligned}\Delta_{ij} &= R_i - R_j \\ &= \sqrt{(x_i - x)^2 + (y_i - y)^2} - \sqrt{(x_j - x)^2 + (y_j - y)^2}. \quad (1)\end{aligned}$$

We see that the source position can be solved from the hyperbolic equations constructed from $\{\Delta_{ij}\}$.

Manuscript received March 17, 2009; accepted September 21, 2009. First published November 06, 2009; current version published February 10, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Chong-Meng Samson See. This work was supported by a grant from CityU (Project No. 7002364).

Y. T. Chan is with the Department of Electrical and Computer Engineering, Royal Military College of Canada, Kingston, Ontario, Canada K7K 7B4.

T. K. C. Lo and H. C. So are with the Department of Electronic Engineering, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong.

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

Digital Object Identifier 10.1109/TSP.2009.2035987