ON THE CHARACTERIZATION, GENERATION, AND EFFICIENT ESTIMATION OF THE COMPLEX MULTIVARIATE GGD

Rami Mowakeaa, Zois Boukouvalas, Tülay Adalı

University of Maryland, Baltimore County
Baltimore, Maryland 21250
Email: {ramo1,zb1,adali}@umbc.edu

Charles Cavalcante

Federal University of Ceará, Fortaleza-CE
Email: charles@gtel.ufc.br

ABSTRACT

The complex multivariate generalized Gaussian distribution (CMGGD) is a flexible parametrized distribution suitable for a variety of applications. Previous work in this area is either limited to the univariate case or, in the multivariate case, restricts the complex vectors, unjustifiably, to be circular. In both cases, algorithms for parameter estimation also suffer from convergence or accuracy limitations over the complete range of their parameters. In this work, we develop the probability density function (PDF) for CMGGD that properly describes noncircular complex data. We then develop a fixed-point algorithm for the estimation of parameters of the CMGGD that is both rapid in its convergence and accurate for the complete shape parameter range. We quantify performance against other algorithms while varying noncircularity, shape parameter and data dimensionality and demonstrate robustness and gains in performance, especially for noncircular data.

1. INTRODUCTION

Generalized Gaussian distribution (GGD), though simple in form, provides a sufficiently flexible model for many applications since it can approximate both super- and sub-Gaussian distributions through a single shape parameter β. Its multivariate formulation [1] defines marginals that are all GGD in shape. The complex version of multivariate GDD (CMGGD) extends the use of this attractive model to applications where data are best represented as complex valued. Recent efforts to generate, characterize and estimate parameters of CMGGD are a testament to the growing interest by the statistical signal processing community [2–4]. Such models are useful, for example, in the joint processing of functional magnetic resonance imaging (fMRI) data [5, 6]. In addition, flexible complex independent component analysis (ICA) algorithms are needed for processing fMRI data in its native complex domain [7]. The purpose of CMGGD is to combine joint modeling of datasets with a flexible parameterizable distribution in the complex domain.

In [4], a probability density function (PDF) is derived for CMGGD, but the variables are assumed to be circular, a mathematically convenient, but often unreasonable simplification for most applications [8–11]. Also, the fixed point algorithm derived is shown to converge only over a subset of the parameter space. In [3], a PDF is derived that accounts for noncircularity but for the univariate case alone. The Newton-Raphson algorithm introduced here suffers in performance as the shape parameter diverges from that of a Gaussian.

In this paper, we construct a PDF for CMGGD that incorporates noncircularity by developing the relation between CMGGD and MGGD. This allows the extension of the Riemannian averaged fixed point (RA-FP) estimator for MGGD [12] to a new complex RA-FP (CRA-FP) estimator for the CMGGD augmented covariance matrix. To the best of our knowledge, no existing PDF and parameter estimator take into account the noncircularity of a multivariate CMGGD model while guaranteeing convergence for all values of β. We present simulation results to demonstrate the desirable performance of the new estimator.

The rest of the paper is organized as follows: in Section 2, we review the mathematical foundation and notation used throughout the paper and establish the connection between CMGGD and MGGD resulting in the CMGGD PDF. Section 3 develops CRA-FP and Section 4 describes the numerical experiments performed. Finally, Section 5 discusses the results of the paper.

2. MATHEMATICAL DEVELOPMENT

2.1. Background and Notation

The PDF of a complex random vector $\mathbf{z} = \mathbf{z}_r + j\mathbf{z}_i \in \mathbb{C}^p$, where $\mathbf{z}_r, \mathbf{z}_i \in \mathbb{R}^p$, if it exists, is defined as the joint PDF of its real and imaginary vector components. For a point $\mathbf{z} \in \mathbb{C}^p$, we recall the isomorphism between $\mathbb{C}^p$ and $\mathbb{R}^{2p}$, in which an equivalent representation results from cascading the real and imaginary components to form a *real composite* vector $\mathbf{z}_\mathbb{R} = \left[\mathbf{z}_r^T, \mathbf{z}_i^T\right]^T \in \mathbb{R}^{2p}$, where the superscript $(.)^T$ indicates the transpose. A thorough presentation of equivalent real and complex forms can be found in [8, 11]. Another equivalent, and quite useful, form is the *complex augmented* one $\underline{\mathbf{z}} = [\mathbf{z}^T, \mathbf{z}^H]^T$, where the superscript $(.)^H$ indicates the Hermitian, or conjugate transpose. The redundancy apparent in this form may appear unnecessary, but in fact it allows the use of Wirtinger calculus to derive real-valued functions of complex variables with simplicity while preserving the intuition of the complex form [13]. Unless otherwise noted, we assume with no loss of generality that the mean of each random vector is $\mathbf{0}$. For a real composite random vector, second-order statistics are given by the covariance matrix

$$\mathbf{C}_{z_\mathbb{R} z_\mathbb{R}} = \mathbb{E}\{\mathbf{z}_\mathbb{R}\mathbf{z}_\mathbb{R}^T\} = \begin{bmatrix} \mathbf{C}_{z_r z_r} & \mathbf{C}_{z_r z_i} \\ \mathbf{C}_{z_r z_i}^T & \mathbf{C}_{z_i z_i} \end{bmatrix},$$

where $\mathbf{C}_{z_r z_r}$, $\mathbf{C}_{z_r z_i}$, and $\mathbf{C}_{z_i z_i}$ are the covariance matrices of the real-real, real-imaginary, and imaginary-imaginary real-valued component vectors of $\mathbf{z}$ respectively. For complex augmented random vectors, second-order statistics are given by the augmented covariance matrix

$$\underline{\mathbf{C}}_{zz} = \mathbb{E}\{\underline{\mathbf{z}}\,\underline{\mathbf{z}}^H\} = \begin{bmatrix} \mathbf{C}_{zz} & \tilde{\mathbf{C}}_{zz} \\ \tilde{\mathbf{C}}_{zz}^* & \mathbf{C}_{zz}^* \end{bmatrix},$$

where $\mathbf{C}_{zz}$ is the covariance matrix:

$$\mathbf{C}_{zz} = \mathbb{E}\{\mathbf{z}\mathbf{z}^H\} = \mathbf{C}_{z_r z_r} + \mathbf{C}_{z_i z_i} + j\left(\mathbf{C}_{z_r z_i}^T - \mathbf{C}_{z_r z_i}\right),$$

and $\tilde{\mathbf{C}}_{zz}$ is the complimentary covariance matrix:

$$\tilde{\mathbf{C}}_{zz} = \mathbb{E}\{\mathbf{z}\mathbf{z}^T\} = \mathbf{C}_{z_r z_r} - \mathbf{C}_{z_i z_i} + j\left(\mathbf{C}_{z_r z_i}^T + \mathbf{C}_{z_r z_i}\right).$$

We note that the sub-blocks of the real composite covariance matrix *fully* describe the second-order complex statistics in the augmented covariance matrix.

To relate these two representations, we denote the set of real symmetric positive definite matrices of size $2p \times 2p$ by $\mathbb{P}_{\mathbb{R}}^{2p}$ and the set of complex augmented Hermitian positive definite matrices of size $2p \times 2p$ by $\mathbb{P}_{\mathbb{C}}^{2p}$. Let the continuous linear function $f : \mathbb{P}_{\mathbb{R}}^{2p} \rightarrow \mathbb{P}_{\mathbb{C}}^{2p}$ such that

$$f\left(\mathbf{\Sigma}_{\mathbb{R}}\right) = \underline{\mathbf{\Sigma}} = \frac{1}{2}\mathbf{T}_p \mathbf{\Sigma}_{\mathbb{R}} \mathbf{T}_p^H$$

where $T_p = \begin{bmatrix} \mathbf{I}_p & j\mathbf{I}_p \\ \mathbf{I}_p & -j\mathbf{I}_p \end{bmatrix}$. It is straightforward to show that f is one-to-one and onto, and therefore f is a bijection with inverse given by

$$f^{-1}\left(\underline{\mathbf{\Sigma}}\right) = \frac{1}{2}\mathbf{T}_p^H \underline{\mathbf{\Sigma}} \mathbf{T}_p.$$

For a univariate complex random variable z with variance $\sigma^2 = \mathbb{E}\{|z|^2\}$ and complimentary covariance $\tilde{\sigma}^2 = \mathbb{E}\{z^2\}$, we define the noncircularity coefficient

$$\rho = \left|\frac{\mathbb{E}\{z^2\}}{\mathbb{E}\{|z|^2\}}\right| \in [0, 1],$$

as a measure of second-order noncircularity. When $\rho = 0$, a complex random variable is uncorrelated with its complex conjugate and is called second-order circular, or proper. When a random variable is maximally correlated with its complex conjugate, $\rho = 1$ and it is called rectilinear. Circularity for a complex random variable z, a stronger assumption than propriety, implies that z and $e^{j\alpha}z, \alpha \in \mathbb{R}$ have the same PDF. For a complete treatment of noncircularity and various extensions for random vectors, see [14].

2.2. CMGGD Development

The CMGGD PDF, assuming circular complex random variables only, is given by [4]

$$f_{\mathbf{z}}(\mathbf{z}) = \frac{\beta\Gamma(p)b^{-p/\beta}}{\pi^p\Gamma\left(p/\beta\right)}\left|\mathbf{\Sigma}_{zz}\right|^{-1} e^{-\frac{1}{b}\left(\mathbf{z}^H\mathbf{\Sigma}_{zz}^{-1}\mathbf{z}\right)^{\beta}},$$

where $\mathbf{z} \in \mathbb{C}^p$ is a complex multivariate generalized Gaussian random vector, b is the scale parameter, β is the shape parameter, and $\mathbf{\Sigma}_{zz} \in \mathbb{C}^{p\times p}$ is the complex scatter matrix that only accounts for circular data. Generating samples from this distribution is performed using the stochastic representation theorem [4]

$$\mathbf{z} =_d \mathcal{R}\mathbf{A}\mathbf{u}, \tag{1}$$

where $=_d$ indicates equality in distribution, $\mathcal{R} = \mathcal{G}^{1/2\beta}$ is a real, non-negative variate with $\mathcal{G} \sim \text{Gamma}\left(\frac{p}{\beta}, b\right)$, $\mathbf{A}$ is a matrix such that $\mathbf{A}\mathbf{A}^H = \mathbf{\Sigma}_{zz}$, and $\mathbf{u}$ is a random vector, independent of $\mathcal{R}$, uniformly distributed on the surface of the p-dimensional complex unit sphere defined as

$$\mathbb{C}S^p \triangleq \{\mathbf{u} \in \mathbb{C}^p : \|\mathbf{u}\| = 1\}.$$

We note that the scatter matrix $\mathbf{\Sigma}_{zz}$ is a scaled version of the covariance matrix $\mathbf{C}_{zz} = b\mathbf{\Sigma}_{zz}$ and therefore, the complimentary covariance matrix is neither taken into account in the PDF nor in the generation process.

To develop a form that accounts for full noncircularity, we write the real-valued MGGD PDF for dimension $2p$ as

$$f(\mathbf{y}) = \frac{\beta\Gamma\left(p\right)b^{-p/\beta}}{\pi^p\Gamma\left(p/\beta\right)}\left|\mathbf{\Sigma}\right|^{-1/2} e^{-\frac{1}{b}\left(\mathbf{y}^T\mathbf{\Sigma}\mathbf{y}\right)^{\beta}}, \tag{2}$$

where $\mathbf{y} \in \mathbb{R}^{2p}$ is a real input vector, $\mathbf{\Sigma} \in \mathbb{R}^{2p\times 2p}$ is the real symmetric positive definite scatter matrix, and p, b, and β are defined as before [1, 15]. Generation of MGGD can be performed according to (1) where $\mathcal{R}$ and $\mathbf{A}$ are generated as in the complex case, but $\mathbf{u}$ is distributed uniformly on the surface of the real $2p$-dimensional unit sphere defined as

$$\mathbb{R}S^{2p} \triangleq \{\mathbf{u} \in \mathbb{R}^{2p} : \|\mathbf{u}\| = 1\}.$$

However, the two definitions of $\mathbf{u}$ coincide when the complex vector $\mathbf{u}$ takes the real composite form $\mathbf{u}_{\mathbb{R}} = \left[\mathbf{u}_r^T, \mathbf{u}_i^T\right]^T$, i.e.,

$$\mathbf{u} \in \mathbb{C}S^p \iff \mathbf{u}_{\mathbb{R}} \in \mathbb{R}S^{2p}.$$

This leads to the following result: The generation of a real multivariate generalized Gaussian vector in $\mathbb{R}^{2p}$ is equivalent to the generation of a complex multivariate generalized Gaussian vector in $\mathbb{C}^p$. The advantage of this connection is that the scatter matrix of the MGGD vector in $\mathbb{R}^{2p}$, along with β, fully capture the statistical information, including noncircularity, of the equivalent CMGGD vector in $\mathbb{C}^p$.

Using identities in [3] that relate the complex augmented and real composite forms, the CMGGD PDF that takes into account full noncircularity can be written as:

$$f_{\mathbf{z}}(\underline{\mathbf{z}}) = \frac{2\beta\Gamma(p)b^{-p/\beta}}{\pi^p\Gamma\left(p/\beta\right)}\left|\underline{\mathbf{\Sigma}}_{zz}\right|^{-1/2} e^{-\frac{1}{b}\left(\underline{\mathbf{z}}^H\underline{\mathbf{\Sigma}}_{zz}^{-1}\underline{\mathbf{z}}\right)^{\beta}}. \tag{3}$$

We choose $b = \left[\frac{2^{1/\beta}\Gamma\left(\frac{p+1}{\beta}\right)}{\Gamma\left(\frac{p}{\beta}\right)}\right]^{\beta}$ as in [1] to obtain the variance normalized distribution with $\underline{\mathbf{\Sigma}}_{zz} = \underline{\mathbf{C}}_{zz}$ This assertion is maintained through the rest of the paper and we do not distinguish between the scatter matrix $\underline{\mathbf{\Sigma}}_{zz}$ and the covariance matrix $\underline{\mathbf{C}}_{zz}$ going forward.

3. PARAMETER ESTIMATION

Estimation of the shape parameter β can be performed in a straightforward manner using a technique based on Newton-Raphson as in [12], so we focus on the estimation of parameters of the scatter matrix $\underline{\mathbf{\Sigma}}_{zz}$. To that end, we develop the complex Riemannian averaged fixed point algorithm (CRA-FP) using the fixed point algorithm developed in [12] that implements successive averages of fixed point iterates by using Riemannian geometry on $\mathbb{P}_{\mathbb{R}}^{2p}$. The idea for the update rule is to transform the current iterate, $\underline{\mathbf{\Sigma}}_{zz}^{[n]}$, from $\mathbb{P}_{\mathbb{C}}^{2p}$ to $\mathbb{P}_{\mathbb{R}}^{2p}$ using the bijection f defined in Section 2, and then apply RA-FP repeatedly until convergence. Pseudo-code for CRA-FP is given in Algorithm 1. Next, we prove the convergence of CRA-FP in the following proposition.

Proposition 1 *Let $\underline{\mathbf{\Sigma}}_{zz}$ be a unique local maximum of the likelihood function of (3) in a neighborhood U of the space of augmented complex matrices $\mathbb{P}_{\mathbb{C}}^{2p}$. Then, from an initial point $\underline{\mathbf{\Sigma}}_{zz}^{[0]} \in U$, the sequence of iterates generated by the CRA-FP algorithm converges to $\underline{\mathbf{\Sigma}}_{zz}$.*

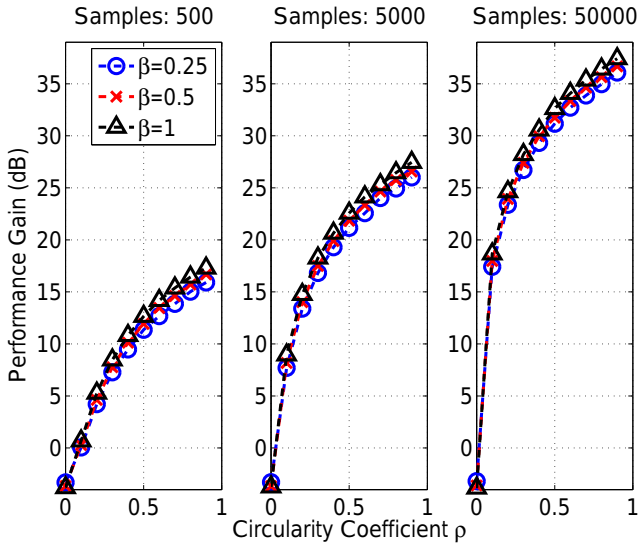

Fig. 1. Average performance gain of utilizing full statistical information for different sample sizes with fixed data dimensionality ($p = 3$). Each point is the average of 1000 independent runs.

Algorithm 1 CRA-FP Algorithm

1: **Input**: $\mathbf{X} \in \mathbb{C}^{p \times T}, \beta$
2: Initialize $\underline{\mathbf{\Sigma}}_{zz}^{[0]} = \mathbf{I}_{2p}$ or by the method of moments (MOM). The same may be done for β if not given
3: Find $\mathbf{\Sigma}_{\mathbb{R}}^{[0]} = f^{-1}\left(\underline{\mathbf{\Sigma}}_{zz}^{[0]}\right) \in \mathbb{P}_{\mathbb{R}}^{2p}$
4: Apply the RA-FP algorithm to produce $\left\{\mathbf{\Sigma}_{\mathbb{R}}^{[n]}\right\}$, the sequence of convergent iterates corresponding to $\left\{\underline{\mathbf{\Sigma}}_{zz}^{[n]}\right\}$
5: **Output**: $\underline{\mathbf{\Sigma}}_{zz}^{[K]} = f\left(\mathbf{\Sigma}_{\mathbb{R}}^{[K]}\right) \in \mathbb{P}_{\mathbb{C}}^{2p}$, the image of the final iterate of RA-FP when a specific tolerance or maximum iteration is met

Proof: Since $\mathbb{P}_{\mathbb{R}}^{2p}$ is compact and f is continuous, $\mathbb{P}_{\mathbb{C}}^{2p}$ is compact. We assume with no loss of generality that U is closed. Then, U and $f^{-1}(U)$ are also compact. Therefore, any sequence of points convergent to $\underline{\mathbf{\Sigma}}_{zz}$ in U has a corresponding sequence of points convergent to $f^{-1}\left(\underline{\mathbf{\Sigma}}_{zz}\right)$ in $f^{-1}(U)$ and vice versa. Hence, $\mathbf{\Sigma}_{\mathbb{R}} = f^{-1}\left(\underline{\mathbf{\Sigma}}_{zz}\right)$ is a unique local maximum of the equivalent likelihood function of (2) in $f^{-1}\left(U\right)$. Following [16, 17], $\mathbf{\Sigma}_{\mathbb{R}}$ is a fixed point with respect to RA-FP applied to (2). Using an initial point $\mathbf{\Sigma}_{\mathbb{R}}^{[0]} = f^{-1}\left(\underline{\mathbf{\Sigma}}_{zz}^{[0]}\right)$, the iterates generated by RA-FP converge to $\mathbf{\Sigma}_{\mathbb{R}}$.

More formally, for any $\epsilon_1 > 0$, there exists $M_1 \in \mathbb{N}$ such that for any $n > M_1$, $\|\mathbf{\Sigma}_{\mathbb{R}}^{[n]} - \mathbf{\Sigma}_{\mathbb{R}}\|_F < \epsilon_1$ where we have chosen the Frobenius norm with no loss in generality. Then, the sequence of iterates generated by CRA-FP is:

$$\left\{\underline{\mathbf{\Sigma}}_{zz}^{[n]}\right\} = \left\{\frac{1}{2}\mathbf{T}_p\mathbf{\Sigma}_{\mathbb{R}}^{[n]}\mathbf{T}_p^H\right\} = \frac{1}{2}\mathbf{T}_p\left\{\mathbf{\Sigma}_{\mathbb{R}}^{[n]}\right\}\mathbf{T}_p^H.$$

To prove convergence:

$$\left\|\underline{\mathbf{\Sigma}}_{zz}^{[n]} - \underline{\mathbf{\Sigma}}_{zz}\right\|_F = \left\|\frac{1}{2}\mathbf{T}_p\mathbf{\Sigma}_{\mathbb{R}}^{[n]}\mathbf{T}_p^H - \frac{1}{2}\mathbf{T}_p\mathbf{\Sigma}_{\mathbb{R}}\mathbf{T}_p^H\right\|_F =$$
$$\left\|\frac{1}{2}\mathbf{T}_p\left(\mathbf{\Sigma}_{\mathbb{R}}^{[n]} - \mathbf{\Sigma}_{\mathbb{R}}\right)\mathbf{T}_p^H\right\|_F \leq \frac{1}{2}\left\|\mathbf{T}_p\right\|_F\left\|\mathbf{\Sigma}_{\mathbb{R}}^{[n]} - \mathbf{\Sigma}_{\mathbb{R}}\right\|_F\left\|\mathbf{T}_p^H\right\|_F$$
$$= 2p^2\left\|\mathbf{\Sigma}_{\mathbb{R}}^{[n]} - \mathbf{\Sigma}_{\mathbb{R}}\right\|_F = \epsilon_2.$$

Finally, for any $\epsilon_2 > 0$, let $\epsilon_1 = \frac{1}{2p^2}\epsilon_2$. Then, there exists $M_2 \in \mathbb{N}$ such that $\left\|\underline{\mathbf{\Sigma}}_{zz}^{[n]} - \underline{\mathbf{\Sigma}}_{zz}\right\|_F < \epsilon_2$ for $n > M_2$. ■

The guaranteed convergence of the algorithm, in addition to its lack of adhoc parameters, make CRA-FP a fast, accurate, and robust parameter estimator for CMGGD data over the entire range of shape parameter β.

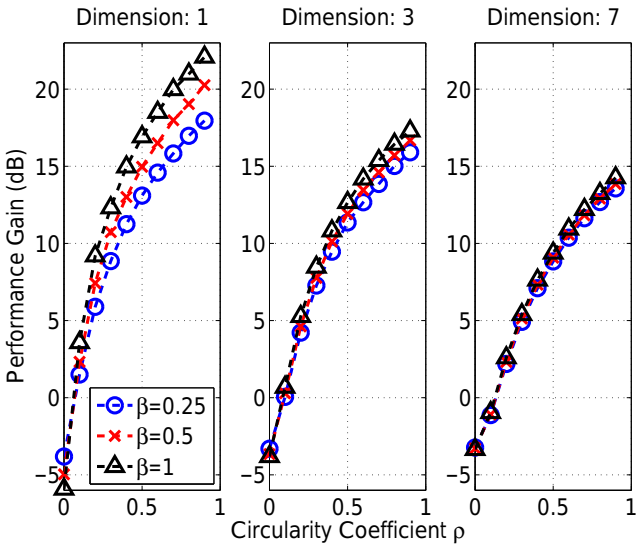

Fig. 2. Average performance gain of utilizing full statistical information for different data dimensionality with a fixed sample size ($N = 500$). Each point is the average of 1000 independent runs.

4. EXPERIMENTAL RESULTS

To demonstrate the performance of CRA-FP, we perform a series of experiments. In order to generate CMGGD data, we utilize an MGGD data generator with real dimensionality of $2p$, where p is the dimension of the complex vector as before. To simplify the experiments, we first consider the noncircularity coefficients $\rho_{kj} = \left|\frac{\mathbb{E}\{z_k z_j\}}{\mathbb{E}\{z_k z_j^*\}}\right|$,where z_k and z_j are the kth and jth entries of the complex random vector $\mathbf{z}$. Then, an additional assumption is imposed on the generated data both to set off-diagonal terms of $\mathbf{C}_{zz}$ and $\tilde{\mathbf{C}}_{zz}$ to 0 as well as to equate the diagonal terms to generate augmented covariance matrices of the form

$$\underline{\mathbf{C}}_{zz} = \begin{bmatrix} \mathbf{I} & j\rho\mathbf{I} \\ -j\rho\mathbf{I} & \mathbf{I} \end{bmatrix} \in \mathbb{C}^{2p \times 2p},$$

where ρ is the desired circularity coefficient.

In the first experiment, we generate CMGGD samples with dimension $p = 3$ with values of noncircularity $\rho = 0.1k,\ k \in \{0, 1, \ldots, 9\}$. For $\beta \in \{0.25, 0.5, 1\}$, we estimate the augmented covariance matrix using CRA-FP and the fixed point algorithm in [4]. Note that we restrict the range of β to the Gaussian and super-Gaussian range since the algorithm in [4] is only guaranteed to converge in this range. Then, we compute the Frobenius norm error between each of the estimates and the true augmented covariance matrix. The performance gain is determined as the ratio of the error of the competing method to that of CRA-FP. We repeat each experiment 1000 times and average the results. The plot in Fig. 1 shows the performance gain measured in dB when using CRA-FP over the CMGGD estimator given in [4] for a number of samples $N \in \{500, 5000, 50000\}$. In this figure, we see that with fixed dimensionality, the performance gain increases with an increasing number of samples. Also, there is a small performance loss for low values of noncircularity. Both can be explained by the increased model complexity of CRA-FP ($3p$ real parameters) over the CMGGD estimator in [4] ($2p$ real parameters). In other words, the noncircularity at which the curve crosses the 0 dB gain level becomes closer to 0 as the sample size increases, warranting increasing model complexity in all but the near-circular case.

In the second experiment, we fix the sample size to $N = 500$ and vary the data dimensionality $p \in \{1, 3, 7\}$. Figure 2 plots the result of this experiment with each point averaged over 1000 independent runs. We note in this case that as the data dimensionality increases, performance decreases as the distribution becomes poorly represented by the fixed sample size. For the same reason, the noncircularity at which the curve crosses the 0 dB gain level becomes closer to 0 as the dimensionality decreases, increasing the advantage for the additional model complexity.

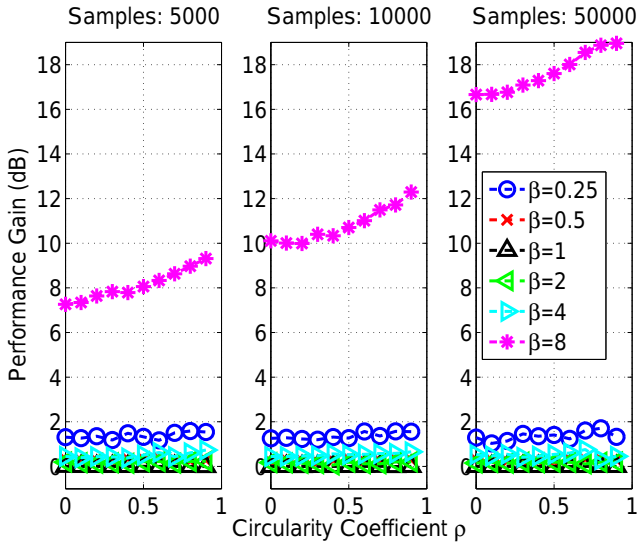

Fig. 3. Average gain in performance of utilizing the fixed point CRA-FP algorithm with dimensionality $p = 1$ against the Newton-derived CGGD algorithm. Each point is the average of 1000 independent runs.

For the third experiment, we generate univariate CGGD data ($p = 1$) with values of noncircularity $\rho = 0.1k,\ k \in \{0, 1, \ldots, 9\}$. For $\beta \in \{0.25, 0.5, 1, 2, 4, 8\}$, we estimate the augmented covariance matrix using CRA-FP and the Newton-Raphson approach in [3]. Then, as in the first experiment, we compute the Frobenius norm error between each of the estimates and the ideal augmented covariance matrix and compute the performance gain. We repeat each experiment 1000 times and average the results. The plot in Fig. 3 shows the gain in performance measured in dB when using CRA-FP over the CGGD algorithm for a number of samples $N \in \{5000, 10000, 50000\}$. We note that the CGGD algorithm suffers in performance with respect to CRA-FP at extreme values of the shape parameter β. Also, CRA-FP outperforms CGGD for every value of β except the Gaussian case ($\beta = 1$) for which they are equal.

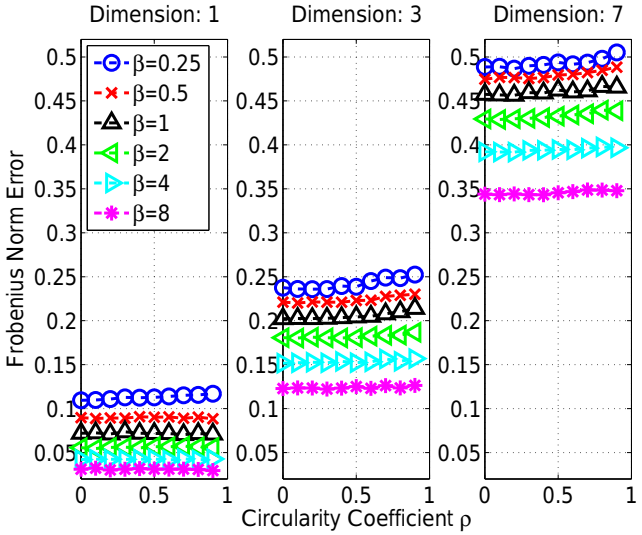

Fig. 4. Average absolute Frobenius norm error utilizing full statistical information for different data dimensionality with a fixed number of samples ($N = 500$). Each point is the average of 1000 independent runs.

Finally, we show the absolute Frobenius norm error of CRA-FP from the first experiment ($N = 500,\ p \in \{1, 3, 7\}$). Only the top p rows of the augmented covariance matrix are used for the error calculation. Figure 4 shows this plot where each point is the average of 1000 independent runs. As expected, the error increases as the dimensionality increases when the sample size is held constant.

5. DISCUSSION

We provide a characterization of CMGGD that takes into account full noncircularity. We develop a new fixed point algorithm, CRA-FP that avoids deficiencies in statistical modeling, convergence limitations, and/or inaccuracy associated with previous approaches. We also provide mathematical proof for the convergence of CRA-FP. Our numerical simulations show that the new algorithm clearly outperforms others in all but the near-circular case. This is related to the model selection issue as discussed in [18], since when noncircularity is low, keeping the model simple—in this case circular—provides advantages in performance. One possible application of these techniques is complex-valued blind source separation (BSS) using independent component analysis (ICA) and independent vector analysis (IVA), for example for the analysis of functional magnetic resonance imaging data using ICA or IVA [19].

References

[1] E. Gómez, M. Gomez-Vilegas, and J. Marin, "A multivariate generalization of the power exponential family of distributions," *Communications in Statistics-Theory and Methods*, vol. 27, no. 3, pp. 589–600, 1998.

[2] M. Novey, T. Adalı, and A. Roy, "Circularity and Gaussianity detection using the complex generalized Gaussian distribution," *IEEE Signal Processing Letters*, vol. 16, no. 11, pp. 993–996, Nov. 2009.

[3] ——, "A complex generalized Gaussian distribution—characterization, generation, and estimation," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1427–1433, March 2010.

[4] E. Ollila, D. E. Tyler, V. Koivunen, and H. V. Poor, "Complex elliptically symmetric distributions: Survey, new results and applications," *Signal Processing, IEEE Transactions on*, vol. 60, no. 11, pp. 5597–5625, 2012.

[5] J. Laney, K. P. Westlake, S. Ma, E. Woytowicz, V. D. Calhoun, and T. Adalı, "Capturing subject variability in fMRI data: A graph-theoretical analysis of GICA vs. IVA," *Journal of neuroscience methods*, vol. 247, pp. 32–40, 2015.

[6] J.-H. Lee, T.-W. Lee, F. A. Jolesz, and S.-S. Yoo, "Independent vector analysis (IVA): multivariate approach for fMRI group study," *Neuroimage*, vol. 40, no. 1, pp. 86–109, 2008.

[7] W. Du, Y. Levin-Schwartz, G.-S. Fu, S. Ma, V. D. Calhoun, and T. Adalı, "The role of diversity in complex ICA algorithms for fMRI analysis," *Journal of Neuroscience Methods*, vol. 264, pp. 129–135, 2016.

[8] T. Adalı, P. Schreier, and L. Scharf, "Complex-valued signal processing: The proper way to deal with impropriety," *IEEE Transactions on Signal Processing*, vol. 59, no. 11, pp. 5101–5125, Nov. 2011.

[9] T. Adalı and P. Schreier, "Optimization and estimation of complex-valued signals: Theory and applications in filtering and blind source separation," *IEEE Signal Processing Magazine*, vol. 31, no. 5, pp. 112–128, Sept. 2014.

[10] P. J. Schreier and L. L. Scharf, *Statistical signal processing of complex-valued data: the theory of improper and noncircular signals*. Cambridge University Press, 2010.

[11] T. Adalı and H. Li, "Complex-valued adaptive signal processing," in *Adaptive Signal Processing*. John Wiley & Sons, Inc., 2010, pp. 1–85.

[12] Z. Boukouvalas, S. Said, L. Bombrun, Y. Berthoumieu, and T. Adalı, "A new Riemannian averaged fixed-point algorithm for MGGD parameter estimation," *IEEE Signal Processing Letters*, vol. 22, no. 12, pp. 2314–2318, Dec. 2015.

[13] P. Bouboulis, "Wirtinger's calculus in general Hilbert spaces," *CoRR*, vol. abs/1005.5170, 2010.

[14] B. Picinbono, "On circularity," *IEEE Transactions on Signal Processing*, vol. 42, no. 12, pp. 3473–3482, Dec. 1994.

[15] Z. Boukouvalas, G.-S. Fu, and T. Adalı, "An efficient multivariate generalized Gaussian distribution estimator: Application to IVA," in *49th Annual Conference on Information Sciences and Systems (CISS), 2015*, March 2015, pp. 1–4.

[16] S. Sra and R. Hosseini, "Geometric optimisation on positive definite matrices for elliptically contoured distributions," in *Advances in Neural Information Processing Systems 26*, 2013, pp. 2562–2570.

[17] F. Pascal, L. Bombrun, J.-Y. Tourneret, and Y. Berthoumieu, "Parameter estimation for multivariate generalized Gaussian distributions," *IEEE Transactions on Signal Processing*, vol. 61, no. 23, pp. 5960–5971, Dec. 2013.

[18] X. L. Li, T. Adalı, and M. Anderson, "Noncircular principal component analysis and its application to model selection," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4516–4528, Oct. 2011.

[19] T. Adalı, M. Anderson, and G.-S. Fu, "Diversity in independent component and vector analyses: Identifiability, algorithms, and applications in medical imaging," *IEEE Signal Processing Magazine*, vol. 31, no. 3, pp. 18–33, May 2014.