This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Application of the Matrix-Variate Mellin Transform to Analysis of Polarimetric Radar Images

Stian Normann Anfinsen, *Member, IEEE*, and Torbjørn Eltoft, *Member, IEEE*

***Abstract*—In this paper, we propose to use a matrix-variate Mellin transform in the statistical analysis of multilook polarimetric radar data. The domain of the transform integral is the cone of complex positive definite matrices, which allows for transformation of the distributions used to model the polarimetric covariance and coherency matrix. Based on the matrix-variate Mellin transform, an alternative characteristic function is defined, from which we can retrieve a new kind of matrix log-moments and log-cumulants. It is demonstrated that the matrix log-cumulants are of great value to analysis of polarimetric radar data, and that they can be used to derive estimators for the distribution parameters with low bias and variance.**

***Index Terms*—Doubly stochastic product model, matrix-variate statistics, Mellin transform (MT), method of log-cumulants, parameter estimation, radar polarimetry, synthetic aperture radar (SAR).**

NOMENCLATURE

CF	Characteristic function.
CGF	Cumulant generating function.
FT	Fourier transform.
GγD	Generalized gamma distribution.
MLC	Matrix log-cumulant.
MLM	Matrix log-moment.
MoLC	Method of log-cumulants.
MoMLC	Method of matrix log-cumulants.
MT	Mellin transform.
MKS	Mellin kind statistics.
PDF	Probability density function.
RV	Random variable.
SAR	Synthetic aperture radar.

I. INTRODUCTION

POLARIMETRIC radar has become an important remote sensing instrument due to its ability to discriminate between different scattering mechanisms. It can therefore characterize physical properties of the target that cannot be determined from single polarization (mono-pol) radar measurements. To fully utilize the polarimetric information captured, it is necessary to consider the complex correlations between all polarimetric channels, thereby including all intensity and phase information. This is taken care of in the doubly stochastic polarimetric product model [1], [2] which describes speckle and texture as independent sources of data variability and provides good model fit to the data. However, the product model produces difficult mathematical expressions for the probability density function (PDF) and the statistical moments which complicates its application to image analysis. We here present an analytic framework that alleviates this problem.

It was noted already by Epstein [3] that the Mellin transform (MT) is a natural analytical tool to use when studying the distribution of products and quotients of independent random variables (RV). Nicolas [4], [5] utilized this fact in the analysis of compounded distributions used to model synthetic aperture radar (SAR) data. He introduced a new theoretical framework by replacing the Fouriér transform (FT) with the MT in the definition of the characteristic function (CF) and cumulant generating function (CGF). This framework was originally coined second kind statistics but we shall refer to it as Mellin kind statistics (MKS). From the resulting Mellin kind CF and CGF, one can retrieve the statistics known as log-moments and log-cumulants.

The most important development under this framework is the method of log-cumulants (MoLC) for parameter estimation [4], [5] which Nicolas applied to a number of doubly stochastic distributions as well as the positive alpha-stable distribution [6] and members of the generalized gamma distribution (GγD) family (e.g., the Weibull and log-normal distribution) [7]. The same method has earlier been applied to GγDs [8], though without relating it to the MT. The list of recently proposed SAR image analysis and image processing algorithms that employ the MoLC covers diverse applications such as statistical modeling [9]–[11], speckle filtering [12]–[15], classification [16], segmentation [17], [18], change detection [19]–[21], interferometric coherence estimation [22], and image compression [23].

Being aware of the impact that MKS has had on mono-pol SAR image analysis, we here extend the theory to the matrix-variate case which describes multilook polarimetric radar data. This is done by introducing a matrix-variate version of the MT [24] which is used to define a Mellin kind CF and CGF of random matrices. We then show how matrix log-moments and matrix log-cumulants can be obtained from the Mellin kind matrix CF and CGF, respectively. For all the theoretical derivations, we highlight the analogy with the univariate case developed by Nicolas and also the classical theory where the FT is used instead of the MT.

The paper is organized as follows. In Section II, we describe the data delivered by mono-pol and polarimetric radars together with the PDFs commonly used to model the data. In Section III,

Manuscript received May 1, 2010; revised November 5, 2010; accepted November 20, 2010.

The authors are with the Department of Physics and Technology, University of Tromsø, 9037 Tromsø, Norway (e-mail: stian.normann.anfinsen@uit.no; torbjorn.eltoft@uit.no).

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/TGRS.2010.2103945

0196-2892/$26.00 © 2011 IEEE

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

we review the classical definition of moments and cumulants and give an overview of univariate MKS before presenting the extension to complex matrix-variate MKS including new definitions and derivations of key properties. The application to parameter estimation for multilook polarimetric radar data distributions using the method of matrix log-cumulants (MoMLC) is presented in Section IV, accompanied by numerical simulations that document the improvement of estimator precision and accuracy. We give our conclusions in Section V.

Our convention for notation is that scalar values are denoted as lower or upper case standard weight characters, vectors are lower case boldface characters, and matrices are upper case boldface characters. Except for scalar random variables, we do not distinguish between random variables and instances of random variables, as such can be ascertained through context. A list of acronyms is provided in the next page for convenience.

II. Radar Data Models

A full-polarimetric imaging radar separately transmits orthogonally polarized microwave pulses and measures orthogonal polarizations of the received signal. For each pixel, the measurements result in a matrix of scattering coefficients. These are complex-valued dimensionless numbers that describe the transformation of the transmitted electromagnetic field to the received electromagnetic field for all combinations of transmit and receive polarization, and assuming no atmospheric disturbance, such as Faraday rotation.

The transformation can be expressed as

$$\begin{bmatrix} \mathcal{E}_x^{\uparrow} \\ \mathcal{E}_y^{\uparrow} \end{bmatrix} = \frac{e^{\jmath kr}}{r} \begin{bmatrix} S_{xx} & S_{xy} \\ S_{yx} & S_{yy} \end{bmatrix} \begin{bmatrix} \mathcal{E}_x^{\downarrow} \\ \mathcal{E}_y^{\downarrow} \end{bmatrix} \tag{1}$$

where k denotes wavenumber and r is the distance between radar and target. The subscript of the electromagnetic field components $\mathcal{E}_i^{\downarrow}$ and $\mathcal{E}_i^{\uparrow}$, $i \in \{x, y\}$ refers to one of the orthogonal polarizations x and y. The superscript indicates transmitted/incident ($\downarrow$) or received/backscattered ($\uparrow$) wave. SAR systems normally use linear polarizations (horizontal and vertical) while circular polarizations (left and right) are another choice. The scattering coefficients S_{ij} are subscripted with the associated receive and transmit polarization, in that order. Together, they form the scattering matrix, denoted $\mathbf{S} = [S_{ij}] \in \mathbb{C}^{2\times 2}$.

The scattering vector may be defined as

$$\mathbf{s} = [S_{xx}\ S_{xy}\ S_{yx}\ S_{yy}]^T = \mathrm{vec}(\mathbf{S}^T) \in \mathbb{C}^d \tag{2}$$

where $(\cdot)^T$ means transposition, $\mathrm{vec}(\cdot)$ denotes vectorisation by column stacking, and $d = \dim(\mathbf{s}) = 4$ is the vector dimension. Other definitions are possible [25] since the vector can be linearly transformed to emphasize physical interpretations of the elements (i.e., Pauli basis), or the dimension can be reduced to $d = 3$ by assuming target reciprocity (i.e., $S_{xy} = S_{yx}$). A reduced version also results when only a subset of $\mathbf{S}$ is measured, such as for mono-pol SAR ($d = 1$) and dual-polarization SAR ($d = 2$).

Radar images are affected by an interference phenomenon which is a characteristic of all coherent imaging systems. The noiselike effect of interference, known as speckle, can be mitigated by a processing step called multilooking. Multiple measurements are obtained by splitting the Doppler bandwidth into a number of subbands, each giving rise to a separate image referred to as a look. All looks are averaged in the power domain to produce multilook data.

The matrix $\mathbf{S}$ and the vector $\mathbf{s}$ are single-look complex-format representations of polarimetric radar data. The L looks extracted in the multilooking process may be represented by the set $\{\mathbf{s}_\ell\}_{\ell=1}^L$ or $\{S_\ell\}_{\ell=1}^L$ in the mono-pol case. The data formats obtained in the multilook intensity domain are

$$C = \frac{1}{L}\sum_{\ell=1}^{L} S_\ell S_\ell^*, \quad d = 1 \tag{3}$$

where $(\cdot)^*$ denotes complex conjugation, and

$$\mathbf{C} = \frac{1}{L}\sum_{\ell=1}^{L} \mathbf{s}_\ell \mathbf{s}_\ell^H, \quad d > 1 \tag{4}$$

where $(\cdot)^H$ denotes the Hermitian (conjugate transposition) operator. We refer to $C \in \mathbb{R}^+$ as the multilook intensity and $\mathbf{C} \in \Omega_+ \subset \mathbb{C}^{d\times d}$ as the multilook polarimetric covariance matrix. Note that C is a real positive RV, whereas $\mathbf{C}$ is a random matrix defined on the cone[1] Ω_+ of positive definite complex Hermitian matrices

$$\Omega_+ = \{\mathbf{X} : \mathbf{X} \succ \mathbf{0},\ \mathbf{X} = \mathbf{X}^H \in \mathbb{C}^{d\times d}\} \tag{5}$$

where $\mathbf{X} \succ \mathbf{0}$ means that $\mathbf{X}$ is positive definite.

A. Gaussian Model

It is commonly assumed that the scattering vector elements are jointly circular complex Gaussian. This is strictly justified only for homogeneous regions of the image characterized by fully developed speckle and no texture. The notion of texture is here defined as spatial variation in the backscatter that is due to target variability, i.e., fluctuations in the radar cross section. The Gaussian model only encompasses variability due to the stochastic interference pattern, that is, speckle.

Assume for the moment that $\mathbf{s}$ is zero mean and circular complex multivariate Gaussian, denoted $\mathbf{s} \sim \mathcal{N}_d^{\mathbb{C}}(\mathbf{0}, \mathbf{\Sigma})$, where $\mathbf{0}$ is a column vector of zeros and $\mathbf{\Sigma} = \mathrm{E}\{\mathbf{s}\mathbf{s}^H\}$ is the covariance matrix of $\mathbf{s}$. $\mathrm{E}\{\cdot\}$ denotes the expectation value. The PDF of $\mathbf{s}$ is

$$p_{\mathbf{s}}(\mathbf{s}; \mathbf{\Sigma}) = \frac{1}{\pi^d |\mathbf{\Sigma}|} \exp(-\mathbf{s}^H \mathbf{\Sigma}^{-1} \mathbf{s}) \tag{6}$$

where $|\cdot|$ is the determinant operator. It follows that if $L \geq d$ and the $\mathbf{s}_\ell$ are independent, then the unnormalized sample covariance matrix, defined as $\mathbf{Z} = L\mathbf{C}$, follows the non-singular complex Wishart distribution [26]:

$$p_{\mathbf{Z}}(\mathbf{Z}; L, \mathbf{\Sigma}) = \frac{|\mathbf{Z}|^{L-d}}{\Gamma_d(L)|\mathbf{\Sigma}|^L} \mathrm{etr}(-\mathbf{\Sigma}^{-1}\mathbf{Z}) \tag{7}$$

[1] A cone is defined as a subset of a vector space that is closed under multiplication by positive scalars.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

where $\mathrm{etr}(\cdot) = \exp(\mathrm{tr}(\cdot))$ is the exponential trace operator and $\mathbf{\Sigma} = \mathrm{E}\{\mathbf{C}\} = \mathrm{E}\{\mathbf{Z}\}/L$. We write this as $\mathbf{Z} \sim \mathcal{W}_d^{\mathbb{C}}(L, \mathbf{\Sigma})$. The normalization constant $\Gamma_d(L)$ is the multivariate gamma function of the complex kind, defined as

$$\begin{aligned}\Gamma_d(L) &= \int_{\Omega_+} |\mathbf{Z}|^{L-d} \mathrm{etr}(-\mathbf{Z})\, d\mathbf{Z} \\ &= \pi^{d(d-1)/2} \prod_{i=0}^{d-1} \Gamma(L-i) \end{aligned} \tag{8}$$

where $\Gamma(L)$ is the standard Euler gamma function. We further have $p_{\mathbf{C}}(\mathbf{C}) = p_{\mathbf{Z}}(L\mathbf{C})|J_{\mathbf{Z}\to\mathbf{C}}|$, where $|J_{\mathbf{Z}\to\mathbf{C}}| = L^{d^2}$ is the Jacobian determinant of the transformation $\mathbf{Z} = L\mathbf{C}$ [24]. The PDF of $\mathbf{C}$ becomes

$$p_{\mathbf{C}}(\mathbf{C}; L, \mathbf{\Sigma}) = \frac{L^{Ld}}{\Gamma_d(L)} \frac{|\mathbf{C}|^{L-d}}{|\mathbf{\Sigma}|^L} \mathrm{etr}(-L\mathbf{\Sigma}^{-1}\mathbf{C}). \tag{9}$$

In the 1-D case, the complex Wishart distribution reduces to the gamma distribution with PDF

$$p_C(c; L, \sigma) = \frac{L^L}{\Gamma(L)} \frac{c^{L-1}}{\sigma^L} \exp\left(-\frac{Lc}{\sigma}\right) \tag{10}$$

where $\sigma = \mathrm{E}\{C\}$. This is denoted $C \sim \gamma(L, \sigma)$.

In the following, we denote the unnormalized covariance matrix $\mathbf{Z}$ by $\mathbf{W}$ when it follows a complex Wishart distribution and the corresponding $\mathbf{C}$ by $\tilde{\mathbf{W}}$. We also refer to the PDF in (7) as $p_{\mathbf{W}}(\mathbf{W})$ to emphasize the underlying model.

B. Product Model

As described above, the randomness of a radar image measurement is commonly attributed to two unrelated factors, namely, speckle and texture. The latter represents the natural spatial variation of the radar cross section which is generally not perfectly homogeneous for pixels that are thematically mapped as one class. While the Gaussian model only accounts for speckle, several statistical models exist that also incorporate texture, either by assuming statistics that imply a non-Gaussian scattering vector, or explicitly, by modeling texture as a separate RV. The latter case leads to a doubly stochastic model with a so-called compounded distribution.

The well-known product model, described e.g., in [1], [2], [28], has been shown to be both mathematically tractable and successful for modeling and prediction purposes in coherent imaging. In the multilook polarimetric version, which is extensively reviewed in [2], it decomposes $\mathbf{Z}$ as a product of two independent stochastic variables

$$\mathbf{Z} = T\mathbf{W} \tag{11}$$

with individual distributions. The positive, scalar, and unit mean RV T generates texture, assuming that its contribution is independent of polarization and common for all channels. The matrix variable $\mathbf{W} \sim \mathcal{W}_d^{\mathbb{C}}(L, \mathbf{\Sigma})$ models speckle. The PDF of $\mathbf{Z}$ depends on the PDF of the multilook texture RV T. Fig. 1 illustrates the relationships between the covariance matrix representations introduced thus far.

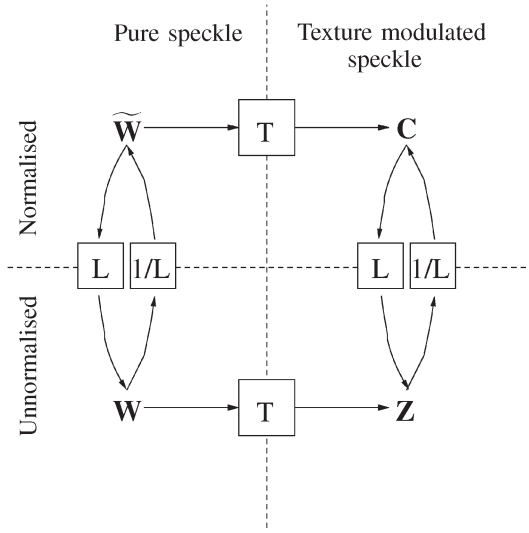

Fig. 1. Relations between the matrix representations introduced in the product model: $\mathbf{Z} = T\mathbf{W}$; $\mathbf{C} = T\tilde{\mathbf{W}}$ and by the normalizations: $\mathbf{Z} = L\mathbf{C}$; $\mathbf{W} = L\tilde{W}$.

In [2], the family of generalized inverse Gaussian distributions is proposed as a model for T. Table I lists the gamma ($\bar{\gamma}$), inverse gamma ($\bar{\gamma}^{-1}$), and Fisher–Snedecor ($\bar{\mathcal{F}}$) distribution as possible choices of $p_T(t)$, giving both notation and expression of their PDF. We remark that the distributions have been normalized to unit mean, indicated by the overbar in the given symbol, which fixes and obsoletes one parameter of the unconstrained distribution. The table also presents the resulting distributions for $\mathbf{C}$, calculated from

$$\begin{aligned} p_{\mathbf{C}}(\mathbf{C}) &= \int_0^\infty p_{\mathbf{Z}|T}(L\mathbf{C}|t) p_T(t) |J_{\mathbf{Z}\to\mathbf{C}}|\, dt \\ &= |J_{\mathbf{Z}\to\mathbf{C}}| \int_0^\infty p_{\mathbf{W}}(tL\mathbf{C}) p_T(t)\, dt. \end{aligned} \tag{12}$$

These covariance matrix distributions are the matrix-variate $\mathcal{K}$ distribution, the matrix-variate $\mathcal{G}^0$ distribution, and the $\mathcal{U}$-distribution. The $\mathcal{K}$ distribution and the $\mathcal{U}$ distribution are named, respectively, after the second kind modified Bessel function of order ν, denoted $K_\nu(\cdot)$, and the second kind confluent hypergeometric Kummer function, denoted $U(\cdot,\cdot,\cdot)$. The complex Wishart distribution is also presented and interpreted as a special case of the product model with a constant texture parameter whose PDF is written in terms of a Dirac delta function, $\delta(t)$.

We remark that the PDF of a product of random variables is known as a Mellin convolution of the factor densities. Thus, (12) can be viewed as a Mellin convolution, as we shall return to later.

III. Mellin Kind Statistics

A. Classical Statistics

A scalar statistical moment captures certain characteristics of a statistical distribution by projecting its PDF onto a scalar. In

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE I
TEXTURE AND COVARIANCE MATRIX DISTRIBUTIONS UNDER THE DOUBLY STOCHASTIC PRODUCT MODEL

$p_T(t)$ of texture variable T		$p_{\mathbf{C}}(\mathbf{C})$ of covariance matrix $\mathbf{C}$		Ref.
Constant	$\delta(t-1)$	$\mathcal{W}_d^{\mathbb{C}}(\mathbf{\Sigma}, L)$	$\frac{L^{Ld}}{\Gamma_d(L)} \frac{\lvert\mathbf{C}\rvert^{L-d}}{\lvert\mathbf{\Sigma}\rvert^{L}} \operatorname{etr}(-L\mathbf{\Sigma}^{-1}\mathbf{C})$	[26]
$\bar{\gamma}(\alpha)$	$\frac{\alpha^\alpha}{\Gamma(\alpha)} t^{\alpha-1} \exp(-\alpha t)$	$\mathcal{K}_d(\mathbf{\Sigma}, L, \alpha)$	$\frac{2\lvert\mathbf{C}\rvert^{L-d}(L\alpha)^{\frac{\alpha+Ld}{2}}}{\lvert\mathbf{\Sigma}\rvert^{L}\Gamma_d(L)\Gamma(\alpha)} \left(\operatorname{tr}(\mathbf{\Sigma}^{-1}\mathbf{C})\right)^{\frac{\alpha-Ld}{2}} K_{\alpha-Ld}\left(2\sqrt{L\alpha \operatorname{tr}(\mathbf{\Sigma}^{-1}\mathbf{C})}\right)$	[27]
$\bar{\gamma}^{-1}(\lambda)$	$\frac{(\lambda-1)^\lambda}{\Gamma(\lambda)} \frac{1}{t^{\lambda+1}} \exp\left(-\frac{\lambda-1}{t}\right)$	$G_d^0(\mathbf{\Sigma}, L, \lambda)$	$\frac{L^{Ld}\lvert\mathbf{C}\rvert^{L-d}}{\Gamma_d(L)\lvert\mathbf{\Sigma}\rvert^{L}} \frac{\Gamma(Ld+\lambda)(\lambda-1)^\lambda}{\Gamma(\lambda)} \left(L\operatorname{tr}(\mathbf{\Sigma}^{-1}\mathbf{C}) + \lambda - 1\right)^{-\lambda-Ld}$	[2]
$\bar{\mathcal{F}}(\xi, \zeta)$	$\frac{\Gamma(\xi+\zeta)}{\Gamma(\xi)\Gamma(\zeta)} \frac{\xi}{\zeta-1} \frac{\left(\frac{\xi}{\zeta-1}t\right)^{\xi-1}}{\left(\frac{\xi}{\zeta-1}t+1\right)^{\xi+\zeta}}$	$\mathcal{U}_d(\mathbf{\Sigma}, L, \xi, \zeta)$	$\frac{L^{Ld}\lvert\mathbf{C}\rvert^{L-d}}{\Gamma_d(L)\lvert\mathbf{\Sigma}\rvert^{L}} \frac{\Gamma(\xi+\zeta)}{\Gamma(\xi)\Gamma(\zeta)} \left(\frac{\xi}{\zeta-1}\right)^{Ld} \Gamma(Ld+\zeta)$ $\times U\left(Ld+\zeta, Ld-\xi+1, L\operatorname{tr}(\mathbf{\Sigma}^{-1}\mathbf{C})\xi/(\zeta-1)\right)$	[11]

the univariate case, the νth-order raw moment of a real RV X is defined as

$$m_\nu\{X\} = \mathrm{E}\{X^\nu\} = \int_{-\infty}^{+\infty} x^\nu p_X(x)\, dx \tag{13}$$

and the νth-order central moment as

$$\begin{aligned}\bar{m}_\nu\{X\} &= \mathrm{E}\left\{(X-m_1)^\nu\right\} \\ &= \int_{-\infty}^{+\infty} (x-m_1)^\nu p_X(x)\, dx\end{aligned} \tag{14}$$

where $m_1 = \mathrm{E}\{X\}$. In both cases, we assume that the integrals exist. Let $\mathcal{F}\{\cdot\}(\omega)$ and $\mathcal{F}^{-1}\{\cdot\}(x)$ be the forward and inverse FT, respectively. The classical CF is defined as $\Phi_X(\omega) = \mathrm{E}\{e^{\jmath\omega X}\} = \mathcal{F}\{p_X(x)\}(\omega)$, with ω a real number and $\jmath$ the imaginary unit.[2] The CF always exists. When all moments exist, the CF can be written as

$$\begin{aligned}\Phi_X(\omega) &= \int_{-\infty}^{+\infty} e^{\jmath\omega x} p_X(x)\, dx \\ &= \sum_{\nu=0}^{\infty} \frac{(\jmath\omega)^\nu}{\nu!} m_\nu\{X\}\end{aligned} \tag{15}$$

using the Maclaurin series expansion of the exponential function. The νth-order moment can in this case be retrieved from

$$m_\nu\{X\} = (-\jmath)^\nu \frac{d^\nu}{d\omega^\nu} \Phi_X(\omega)\bigg|_{\omega=0}. \tag{16}$$

A statistical distribution is uniquely specified by its CF if all of its moments are finite and the power series expansion for its CF converges absolutely near the origin [29]. Then,

$$\begin{aligned}p_X(x) &= \mathcal{F}^{-1}\left\{\Phi_X(\omega)\right\}(x) \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-\jmath\omega x} \Phi_X(\omega)\, d\omega.\end{aligned} \tag{17}$$

[2]The exponential in the Fouriér transform may be defined with or without a negative sign in the exponent. We have chosen the latter version.

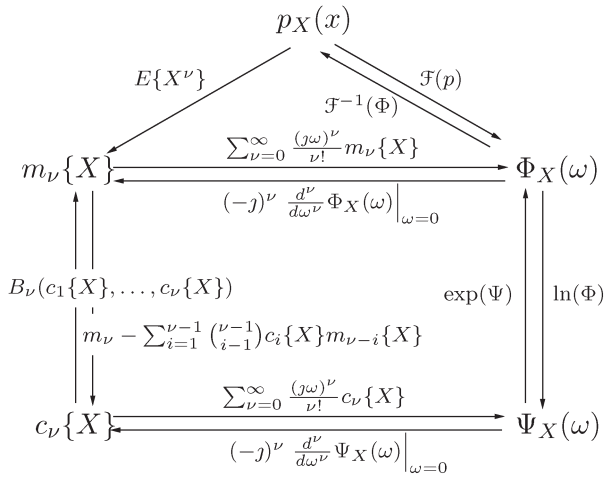

Fig. 2. Relations in univariate classical statistics.

The CGF of X is defined as $\Psi_X(\omega) = \ln \Phi_X(\omega)$. When the moments $m_\nu\{X\}$ exist, so do the cumulants $c_\nu\{X\}$ that are coefficients of the power series expansion

$$\Psi_X(\omega) = \sum_{\nu=0}^{\infty} \frac{(\jmath\omega)^\nu}{\nu!} c_\nu\{X\}. \tag{18}$$

The cumulants can be retrieved from the CGF as

$$c_\nu\{X\} = (-\jmath)^\nu \frac{d^\nu}{d\omega^\nu} \Psi_X(\omega)\bigg|_{\omega=0} \tag{19}$$

by analogy with the moments. Moments and cumulants are related by a combinatorial version of Faá di Bruno's formula [30]

$$c_\nu\{X\} = m_\nu\{X\} - \sum_{i=1}^{\nu-1} \binom{\nu-1}{i-1} c_i\{X\} m_{\nu-i}\{X\} \tag{20}$$

and reversely through

$$m_\nu\{X\} = B_\nu\left(c_1\{X\}, \ldots, c_\nu\{X\}\right) \tag{21}$$

where $B_\nu(\cdot)$ is the νth complete Bell polynomial [31]. All relations for classical univariate statistics are summarized in the diagram of Fig. 2.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Moments and cumulants can be generalized to random vectors: $\mathbf{x} \sim p_{\mathbf{x}}(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^n$, random matrices: $\mathbf{X} \sim p_{\mathbf{X}}(\mathbf{X})$, $\mathbf{X} \in \mathbb{R}^{m\times n}$, and the corresponding complex cases: $\mathbf{z} \sim p_{\mathbf{z}}(\mathbf{z})$, $\mathbf{z} \in \mathbb{C}^n$ and $\mathbf{Z} \sim p_{\mathbf{Z}}(\mathbf{Z})$, $\mathbf{Z} \in \mathbb{C}^{m\times n}$. The CF of a complex random vector $\mathbf{z}$ is defined as [32], [33]

$$\Phi_{\mathbf{z}}(\boldsymbol{\omega}) = \mathrm{E}\left\{e^{\jmath\mathcal{R}e\{\boldsymbol{\omega}^H\mathbf{z}\}}\right\}, \quad \boldsymbol{\omega} \in \mathbb{C}^n \tag{22}$$

and the CF of a complex random matrix $\mathbf{Z}$ as

$$\Phi_{\mathbf{Z}}(\boldsymbol{\Omega}) = \mathrm{E}\left\{e^{\jmath\mathcal{R}e\{\mathrm{tr}(\boldsymbol{\Omega}^H\mathbf{Z})\}}\right\}, \quad \boldsymbol{\Omega} \in \mathbb{C}^{m\times n} \tag{23}$$

where $\mathrm{tr}(\cdot)$ is the trace operator and $\mathcal{R}e\{\cdot\}$ extracts the real part of a complex expression, while the vector $\boldsymbol{\omega}$ and matrix $\boldsymbol{\Omega}$ are transform variables. We note that the CFs are defined in terms of the standard complex vector inner product $\boldsymbol{\omega}^H\mathbf{z}$ and complex matrix inner product $\mathrm{tr}(\boldsymbol{\Omega}^H\mathbf{Z})$ [32].

The νth-order moments of $\mathbf{z}$ and $\mathbf{Z}$ are retrieved from

$$m_\nu\{\mathbf{z}\} = (-\jmath)^\nu \frac{\partial^\nu}{\partial\boldsymbol{\omega}^\nu}\Phi_{\mathbf{z}}(\boldsymbol{\omega})\bigg|_{\boldsymbol{\omega}=\mathbf{0}} \tag{24}$$

$$m_\nu\{\mathbf{Z}\} = (-\jmath)^\nu \frac{\partial^\nu}{\partial\boldsymbol{\Omega}^\nu}\Phi_{\mathbf{Z}}(\boldsymbol{\Omega})\bigg|_{\boldsymbol{\Omega}=\mathbf{0}} \tag{25}$$

where $\partial^\nu/\partial\boldsymbol{\omega}^\nu$ and $\partial^\nu/\partial\boldsymbol{\Omega}^\nu$ is multi-index notation for sequential νth-order partial differentiation with respect to all elements in $\boldsymbol{\omega}$ and $\boldsymbol{\Omega}$, respectively. Non-scalar moments and cumulants of real random vectors and matrices are defined in [34] and the theory can be extended to the complex case, but this is outside the scope of our work.

B. Univariate Mellin Kind Statistics

The MT of the real-valued function $f(x)$ defined on $\mathbb{R}^+$ is given as [3], [35]

$$F(s) = \mathcal{M}\left\{f(x)\right\}(s) = \int_0^\infty x^{s-1}f(x)\,dx \tag{26}$$

where the transform variable $s \in \mathbb{C}$. Under certain restrictions on $f(x)$, $F(s)$ is analytic in a strip parallel to the imaginary axis. The inverse MT is

$$\begin{aligned} f(x) &= \mathcal{M}^{-1}\left\{F(s)\right\}(x) \\ &= \frac{1}{2\pi\jmath}\int_{a-\jmath\cdot\infty}^{a+\jmath\cdot\infty} x^{-s}F(s)\,ds. \end{aligned} \tag{27}$$

The integration limits denote a line integral along any line $s = a \in \mathbb{R}$ parallel to the imaginary axis, which must lie within the analytic strip of $F(s)$.

Nicolas [4] proposed to replace the FT with the MT in the definition of the CF for the RV X, thus defining the Mellin kind CF as

$$\phi_X(s) = \mathrm{E}\left\{X^{s-1}\right\} = \mathcal{M}\left\{p_X(x)\right\}(s). \tag{28}$$

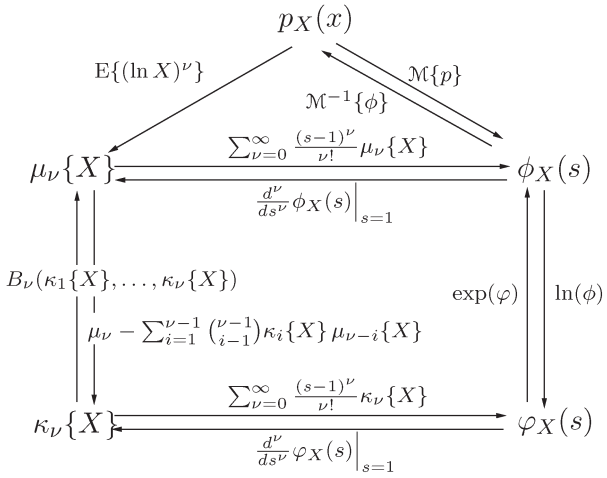

Fig. 3. Relations in univariate Mellin kind statistics.

The domain of the MT restricts this definition to positive RVs $X \in \mathbb{R}^+$. By expanding $\phi_X(s)$ as in the classical case, we obtain

$$\begin{aligned} \phi_X(s) &= \int_0^\infty e^{(\ln x)(s-1)}p_X(x)\,dx \\ &= \sum_{\nu=0}^{\infty}\frac{(s-1)^\nu}{\nu!}\mu_\nu\{X\} \end{aligned} \tag{29}$$

with the νth-order Mellin kind moment defined as

$$\begin{aligned} \mu_\nu\{X\} &= \mathrm{E}\left\{(\ln X)^\nu\right\} \\ &= \int_0^\infty (\ln x)^\nu p_X(x)\,dx \end{aligned} \tag{30}$$

provided the integral exists. The derivation of (29) reveals that $\phi_X(s)$ is a power series expansion of the terms $\mu_\nu\{X\}$, appropriately termed log-moments, when they exist. When pursuing the analogy with the classical case, it is found that

$$\mu_\nu\{X\} = \frac{d^\nu}{ds^\nu}\phi_X(s)\bigg|_{s=1}. \tag{31}$$

The Mellin kind CGF is further defined as $\varphi_X(s) = \ln\phi_X(s)$, from which the Mellin kind cumulants, also known as log-cumulants, can be retrieved as

$$\kappa_\nu\{X\} = \frac{d^\nu}{ds^\nu}\varphi_X(s)\bigg|_{s=1} \tag{32}$$

given that the corresponding log-moment exists. When all log-cumulants exist, the Mellin kind CGF can be expanded as

$$\varphi_X(s) = \sum_{\nu=0}^{\infty}\frac{(s-1)^\nu}{\nu!}\kappa_\nu\{X\}. \tag{33}$$

The relation between the Mellin kind CF and CGF is the same as in the classical case, hence so is the relation between the log-moments and log-cumulants. Fig. 3 summarizes all relations for the univariate MKS.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE II
MELLIN KIND STATISTICS OF UNIVARIATE DISTRIBUTIONS FOR REAL POSITIVE TEXTURE VARIABLES

$p_T(t)$	Characteristic function $\phi_T(s)$	Log-cumulants $\kappa_\nu(T)$
$\bar{\gamma}(\alpha)$	$\frac{\Gamma(\alpha+s-1)}{\alpha^{s-1}\Gamma(\alpha)}$	$\kappa_1 = \psi^{(0)}(\alpha) - \ln(\alpha)$ $\kappa_{\nu>1} = \psi^{(\nu-1)}(\alpha)$
$\bar{\gamma}^{-1}(\lambda)$	$\left(\frac{\lambda-1}{\lambda}\right)^{s-1} \frac{\Gamma(\lambda+1-s)}{\lambda^{1-s}\Gamma(\lambda)}$	$\kappa_1 = \ln(\lambda-1) - \psi^{(0)}(\lambda)$ $\kappa_{\nu>1} = (-1)^\nu \psi^{(\nu-1)}(\lambda)$
$\bar{\mathcal{F}}(\xi,\zeta)$	$\left(\frac{\zeta-1}{\zeta}\right)^{s-1} \frac{\Gamma(\xi+s-1)}{\xi^{s-1}\Gamma(\xi)} \frac{\Gamma(\zeta+1-s)}{\zeta^{1-s}\Gamma(\zeta)}$	$\kappa_1 = \psi^{(0)}(\xi) - \psi^{(0)}(\zeta) + \ln\left(\frac{\zeta-1}{\xi}\right)$ $\kappa_{\nu>1} = \psi^{(\nu-1)}(\xi) + (-1)^\nu \psi^{(\nu-1)}(\zeta)$

Nicolas derived the MKS for the gamma, the inverse gamma, and the Fisher–Snedecor distribution [4], [5], among others. These results are listed in Table II for the unit mean version of these distributions.

C. Complex Matrix-Variate Mellin Kind Statistics

Mathai proposed a generalized transform (that he named the M-transform) for matrix-valued functions in [36] and referred to it in [37] as a generalized MT. The complex version relevant to this paper was presented in [24].

Definition 1 (Complex Matrix-Variate MT): Let $f(\mathbf{Z})$ be a real-valued scalar function defined on a cone of $d \times d$ Hermitian matrices that are either positive definite, negative definite, or null, and let f be symmetric in the sense $f(\mathbf{ZV}) = f(\mathbf{VZ})$ where $\mathbf{V}$ and $\mathbf{Z}$ belong to the same cone. The complex matrix-variate MT is then defined as

$$\mathcal{M}\left\{f(\mathbf{Z})\right\}(s) = \int_{\Omega_+} |\mathbf{Z}|^{s-d} f(\mathbf{Z})\, d\mathbf{Z} \tag{34}$$

with transform variable $s \in \mathbb{C}$ whenever the integral exists. It is duly noted in [36] that since $\mathcal{M}\{f(\mathbf{Z})\}(s)$ is a function of the complex scalar transform variable s, whereas $f(\mathbf{Z})$ is defined on a matrix space, the transform is not unique.

The symmetry requirement in the definition of the matrix-variate MT restricts in theory the range of PDFs it can be applied to. In practice, however, it does not pose any problems for the compound Wishart-type distributions used for multilook polarimetric radar data. In these functions (See Table I), the matrix variable $\mathbf{Z}$ occurs inside determinant and trace operators that are symmetric themselves, hence the overall PDFs are also symmetric in the required sense. We may therefore use the transform to define MKS for the complex matrix-variate case.

Definition 2 (Mellin Kind Matrix-Variate CF): The Mellin kind CF of the complex random matrix $\mathbf{Z}$ is defined as

$$\phi_{\mathbf{Z}}(s) = \mathrm{E}\left\{|\mathbf{Z}|^{s-d}\right\} = \mathcal{M}\left\{p_{\mathbf{Z}}(\mathbf{Z})\right\}(s) \tag{35}$$

when $\mathbf{Z}$ and $p_{\mathbf{Z}}(\mathbf{Z})$ satisfy all requirements of the complex matrix-variate MT.

Definition 3 (Mellin Kind Matrix Moments): The νth-order Mellin kind matrix moment of $\mathbf{Z}$ is defined as

$$\mu_\nu\{\mathbf{Z}\} = \left.\frac{d^\nu}{ds^\nu}\phi_{\mathbf{Z}}(s)\right|_{s=d}. \tag{36}$$

If all Mellin kind matrix moments exist, the Mellin kind CF can be written as the power series expansion

$$\begin{aligned}\phi_{\mathbf{Z}}(s) &= \int_{\Omega_+} e^{(s-d)\ln|\mathbf{Z}|} p_{\mathbf{Z}}(\mathbf{Z})\, d\mathbf{Z} \\ &= \sum_{\nu=0}^{\infty} \frac{(s-d)^\nu}{\nu!} \mu_\nu\{\mathbf{Z}\}\end{aligned} \tag{37}$$

in terms of the $\mu_\nu\{\mathbf{Z}\}$. The derivation of (37) reveals that

$$\begin{aligned}\mu_\nu\{\mathbf{Z}\} &= \mathrm{E}\left\{(\ln|\mathbf{Z}|)^\nu\right\} \\ &= \int_{\Omega_+} (\ln|\mathbf{Z}|)^\nu\, p_{\mathbf{Z}}(\mathbf{Z})\, d\mathbf{Z}\end{aligned} \tag{38}$$

when the integral converges, which justifies the denotation of $\mu_\nu\{\mathbf{Z}\}$ as a matrix log-moment (MLM).

Definition 4 (Mellin Kind Matrix-Variate CGF): The Mellin kind CGF of the complex random matrix $\mathbf{Z}$ is defined as

$$\varphi_{\mathbf{Z}}(s) = \ln \phi_{\mathbf{Z}}(s). \tag{39}$$

Definition 5 (Mellin Kind Matrix Cumulants): The νth-order Mellin kind matrix cumulant of $\mathbf{Z}$ is defined as

$$\kappa_\nu\{\mathbf{Z}\} = \left.\frac{d^\nu}{ds^\nu}\varphi_{\mathbf{Z}}(s)\right|_{s=d}. \tag{40}$$

When all Mellin kind matrix moments exist, so do the Mellin kind matrix cumulants, and the Mellin kind CGF can be expanded as

$$\varphi_{\mathbf{Z}}(s) = \ln \phi_{\mathbf{Z}}(s) = \sum_{\nu=0}^{\infty} \frac{(s-d)^\nu}{\nu!} \kappa_\nu\{\mathbf{Z}\} \tag{41}$$

in terms of the $\kappa_\nu\{\mathbf{Z}\}$, that are also called matrix log-cumulants (MLCs).

As we see, there is a complete analogy with the MKS derived in Section III-B for the univariate case, as summarized in Fig. 4. The Mellin kind matrix-variate CF and CGF are related by the same logarithmic transformation as in the univariate case. Thus, the conversion between MLMs and MLCs is also given in terms of Faá di Bruno's formula and the complete Bell polynomial by analogy with (20) and (21).

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

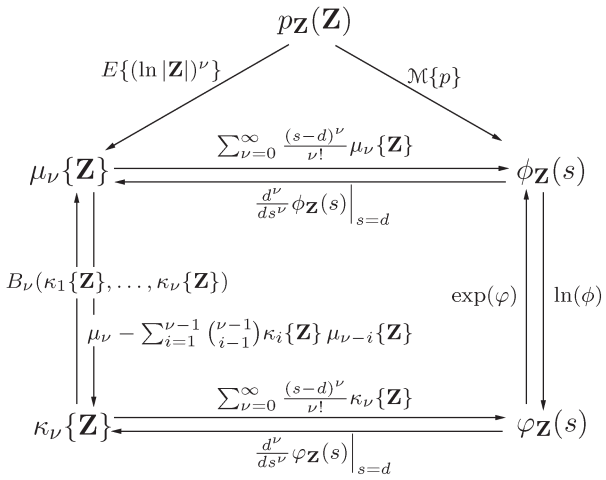

Fig. 4. Relations in matrix-variate Mellin kind statistics.

Example (Complex Wishart Distribution): The Mellin kind CF of a complex Wishart matrix $\mathbf{W}\sim\mathcal{W}_d^{\mathbb{C}}(L,\boldsymbol{\Sigma})$ is derived in Appendix A as

$$\begin{aligned}\phi_{\mathbf{W}}(s)&=\mathcal{M}\left\{p_{\mathbf{W}}(\mathbf{W})\right\}(s)\\&=\frac{\Gamma_d(L+s-d)}{\Gamma_d(L)}|\boldsymbol{\Sigma}|^{s-d}.\end{aligned}\tag{42}$$

The MLCs are found to be

$$\kappa_1\{\mathbf{W}\}=\psi_d^{(0)}(L)+\ln|\boldsymbol{\Sigma}|\tag{43}$$

$$\kappa_\nu\{\mathbf{W}\}=\psi_d^{(\nu-1)}(L),\quad\nu>1\tag{44}$$

where we introduce the νth-order multivariate polygamma function as

$$\psi_d^{(\nu)}(L)=\sum_{i=0}^{d-1}\psi^{(\nu)}(L-i).\tag{45}$$

This is a convenient extension of the ordinary polygamma functions, defined as the logarithmic derivatives of the gamma function

$$\psi^{(\nu)}(L)=\frac{d^{\nu+1}\ln\Gamma(L)}{dL^{\nu+1}}\quad\nu\geq0.\tag{46}$$

Let $\tilde{\mathbf{W}}=\mathbf{W}/L$ be the normalized Wishart matrix whose PDF is given in (9). The MLCs of $\tilde{\mathbf{W}}$ are derived as

$$\kappa_1\{\tilde{\mathbf{W}}\}=\psi_d^{(0)}(L)+\ln|\boldsymbol{\Sigma}|-d\ln L\tag{47a}$$

$$\kappa_\nu\{\tilde{\mathbf{W}}\}=\psi_d^{(\nu-1)}(L),\quad\nu>1.\tag{47b}$$

The MLMs of $\mathbf{W}$ and $\tilde{\mathbf{W}}$ can be found by inserting the MLCs into the equivalent formula of (21).

Log-statistics of $\mathbf{W}$ and $\tilde{\mathbf{W}}$ were first derived in [38] without utilizing the Mellin transform and not for a general order ν. They were also used in [39] but interpreted as log-moments and log-cumulants of the positive scalar RV $|\mathbf{W}|$ rather than of the matrix $\mathbf{W}$. A detailed derivation is given in Appendix A.

D. Some Properties of the Matrix-Variate Product Model

We shall now look at some fundamental properties of the MT which makes it a natural replacement of the FT when working with a multiplicative signal model and extend this exposition to the complex matrix-variate case.

1) Univariate Additive Model: Let X, U, and V be real scalar RVs whose moments and cumulants all exist, and assume that U and V are statistically independent. For the additive stochastic signal model

$$X=U+V\tag{48}$$

we find that the CF, the CGF, and the cumulants of X, as defined in the classical case using the FT, can be written as

$$\Phi_X(\omega)=\Phi_U(\omega)\cdot\Phi_V(\omega)\tag{49}$$

$$\Psi_X(\omega)=\Psi_U(\omega)+\Psi_V(\omega)\tag{50}$$

$$c_\nu\{X\}=c_\nu\{U\}+c_\nu\{V\}.\tag{51}$$

These relations also hold when the signal model is generalized to the multivariate, matrix-variate, and complex case. The PDF of X is given by the convolution

$$\begin{aligned}p_X(x)&=(p_U*p_V)(x)\\&=\int_{-\infty}^{+\infty}p_U(u)p_V(x-u)du\end{aligned}\tag{52}$$

where $*$ denotes the convolution operator, which corresponds to a multiplication in the FT domain, as seen in (49).

2) Univariate Product Model: Now consider the stochastic product model

$$X=U\cdot V\tag{53}$$

with the additional constraints that X, U, and $V\in\mathbb{R}^+$. Observe that the Mellin kind CF, CGF, and cumulants defined with the MT take the form

$$\phi_X(s)=\phi_U(s)\cdot\phi_V(s)\tag{54}$$

$$\varphi_X(s)=\varphi_U(s)+\varphi_V(s)\tag{55}$$

$$\kappa_\nu\{X\}=\kappa_\nu\{U\}+\kappa_\nu\{V\}.\tag{56}$$

The usual interpretation [4] is to view the MT as a Laplace transform computed on a logarithmic scale. The logarithmic transformation translates the product model into an additive one, which explains why the Mellin kind CF inherits the multiplicative property of the classical CF, whereas the CGF and the cumulants take over the additive property from their classical counterparts.

The PDF of X can be found from

$$\begin{aligned}p_X(x)&=\int_0^\infty p_{X|V}(x|v)p_V(v)dv\\&=\int_0^\infty p_U\left(\frac{x}{v}\right)p_V(v)\frac{dv}{v}\end{aligned}\tag{57}$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

which is known as the Mellin convolution. The operation is denoted $p_X(x) = (p_U \hat{\star} p_V)(x)$. The product in (54) is the MT domain equivalent.

3) Matrix-Variate Product Model: Before we are ready to consider the matrix-variate product model, we shall establish a matrix-variate Mellin convolution theorem using the matrix-variate MT of Definition 1. We also find it natural to include some closely related correlation theorems. We start by defining the matrix-variate Mellin convolution.

Definition 6 (Matrix-Variate Mellin Convolution): Let $f(\mathbf{U})$ and $g(\mathbf{U})$ be two functions defined on the cone of positive definite (or negative definite) complex Hermitian matrices. Further assume that $\mathbf{U}$ and $\mathbf{V}$ both belong to the domain of f and g and that the functions are symmetric in the sense that $f(\mathbf{UV}) = f(\mathbf{VU})$. We define the matrix-variate Mellin convolution of f and g as

$$\begin{aligned}(f \hat{\star} g)(\mathbf{U}) &= \int_{\Omega_+} |\mathbf{V}|^{-d} f\left(\mathbf{V}^{-\frac{1}{2}}\mathbf{U}\mathbf{V}^{-\frac{1}{2}}\right) g(\mathbf{V}) d\mathbf{V} \\ &= \int_{\Omega_+} |\mathbf{V}|^{-d} g\left(\mathbf{V}^{-\frac{1}{2}}\mathbf{U}\mathbf{V}^{-\frac{1}{2}}\right) f(\mathbf{V}) d\mathbf{V} \end{aligned} \tag{58}$$

which is an associative and commutative operation.

Theorem 1 (Matrix-Variate Mellin Convolution): Under the assumptions presented in Definition 6, then

$$\mathcal{M}\{(f \hat{\star} g)(\mathbf{U})\}(s) = \mathcal{M}\{f(\mathbf{U})\}(s) \cdot \mathcal{M}\{g(\mathbf{U})\}(s). \tag{59}$$

Proof: Introduce the substitution $\mathbf{X} = \mathbf{UV}$ and note that $\mathbf{X}$ must belong to the same matrix space as $\mathbf{U}$ and $\mathbf{V}$. Furthermore, we have $\mathbf{U} = \mathbf{V}^{-(1/2)}\mathbf{X}\mathbf{V}^{-(1/2)}$ and $d\mathbf{U} = d\mathbf{X}/|\mathbf{V}|^d$ [24]. This yields

$$\begin{aligned}&\mathcal{M}\{f(\mathbf{U})\}(s) \cdot \mathcal{M}\{g(\mathbf{V})\}(s) \\ &= \int_{\Omega_+} (|\mathbf{X}|/|\mathbf{V}|)^{s-d} f\left(\mathbf{V}^{-\frac{1}{2}}\mathbf{X}\mathbf{V}^{-\frac{1}{2}}\right) |\mathbf{V}|^{-d} d\mathbf{X} \\ &\quad \times \int_{\Omega_+} |\mathbf{V}|^{s-d} g(\mathbf{V}) d\mathbf{V} \\ &= \int_{\Omega_+} |\mathbf{X}|^{s-d} \left[\int_{\Omega_+} |\mathbf{V}|^{-d} f\left(\mathbf{V}^{-\frac{1}{2}}\mathbf{X}\mathbf{V}^{-\frac{1}{2}}\right) g(\mathbf{V}) d\mathbf{V}\right] d\mathbf{X} \\ &= \mathcal{M}\{(f \hat{\star} g)(\mathbf{X})\}(s) \end{aligned} \tag{60}$$

where in the last transition, we identify the term in the square brackets as the Mellin convolution. ■

We have shown that the MT provides a convolution theorem for the product model like the FT does for the additive model, and that this extends to matrix-variate theory. By further analogy with the Fourier transform, the Mellin transform also has a correlation theorem. The following operation reduces in the univariate case to the Mellin correlation, as Nicolas defines it in [4], [5].

Definition 7 (Type I Matrix-Variate Mellin Correlation): Under the assumptions presented in Definition 6, we define the type I matrix-variate Mellin correlation of f and g as

$$(f \hat{\otimes} g)(\mathbf{U}) = \int_{\Omega_+} |\mathbf{V}|^d f\left(\mathbf{V}^{\frac{1}{2}}\mathbf{U}\mathbf{V}^{\frac{1}{2}}\right) g(\mathbf{V}) d\mathbf{V}. \tag{61}$$

This operation is neither associative nor commutative.

Theorem 2 (Type I Matrix-Variate Mellin Correlation): Under the assumptions presented in Definition 6, then

$$\mathcal{M}\left\{(f \hat{\otimes} g)(\mathbf{U})\right\}(s) = \mathcal{M}\left\{f(\mathbf{U})\right\}(s) \cdot \mathcal{M}\left\{g(U)\right\}(2d - s). \tag{62}$$

The proof is given in [40].

We also present an alternative definition. It reduces in the univariate case to a relation often referred to as a Mellin correlation (see e.g., [41]).

Definition 8 (Type II Matrix-Variate Mellin Correlation): Under the assumptions presented in Definition 6, we define the type II matrix-variate Mellin correlation of f and g as

$$(f \hat{\circledast} g)(\mathbf{U}) = \int_{\Omega_+} |\mathbf{V}|^{-d} f\left(\mathbf{V}^{\frac{1}{2}}\mathbf{U}\mathbf{V}^{\frac{1}{2}}\right) g(\mathbf{V}) d\mathbf{V}. \tag{63}$$

This is neither an associative nor commutative operation.

Theorem 3 (Type II Matrix-Variate Mellin Correlation): Under the assumptions presented in Definition 6, then

$$\mathcal{M}\left\{(f \hat{\circledast} g)(\mathbf{U})\right\}(s) = \mathcal{M}\left\{f(\mathbf{U})\right\}(s) \cdot \mathcal{M}\left\{g(-\mathbf{U})\right\}(s). \tag{64}$$

The proof follows in the same manner as for the Theorem 1 and is therefore omitted. A more general theorem which reduced to both Theorems 2 and 3 was presented in [24, Th. 6.2].

We will now explain the relevance to a matrix-variate product model expressed by the (ordinary) matrix product

$$\mathbf{X} = \mathbf{UV} \tag{65}$$

where $\mathbf{X}$, $\mathbf{U}$, and $\mathbf{V}$ are complex and positive definite Hermitian matrices. With the same approach as in (57), we establish that the PDF of $\mathbf{X}$ is

$$\begin{aligned} p_{\mathbf{X}}(\mathbf{X}) &= \int_{\Omega_+} p_{\mathbf{X}|\mathbf{V}}(\mathbf{X}|\mathbf{V}) p_{\mathbf{V}}(\mathbf{V}) d\mathbf{V} \\ &= \int_{\Omega_+} |\mathbf{V}|^{-d} p_{\mathbf{U}}\left(\mathbf{V}^{-\frac{1}{2}}\mathbf{X}\mathbf{V}^{-\frac{1}{2}}\right) p_{\mathbf{V}}(\mathbf{V}) d\mathbf{V}. \end{aligned} \tag{66}$$

This is exactly $(p_{\mathbf{U}} \hat{\star} p_{\mathbf{V}})(\mathbf{X})$ as should be expected from (60), which justifies the definition of the matrix-variate Mellin convolution. Assume that all MLMs and MLCs of $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{X}$ exist, as given in Definitions 3 and 5. It follows from Theorem 1 that

$$\phi_{\mathbf{X}}(s) = \phi_{\mathbf{U}}(s) \cdot \phi_{\mathbf{V}}(s) \tag{67}$$

$$\varphi_{\mathbf{X}}(s) = \varphi_{\mathbf{U}}(s) + \varphi_{\mathbf{V}}(s) \tag{68}$$

$$\kappa_\nu\{\mathbf{X}\} = \kappa_\nu\{\mathbf{U}\} + \kappa_\nu\{\mathbf{V}\}. \tag{69}$$

in the matrix-variate case.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Example (Multilook Polarimetric Product Model): We now return to the multilook polarimetric product model for the observable covariance matrix $\mathbf{C}$. The model can be written as $\mathbf{C} = \mathbf{T}\tilde{\mathbf{W}}$ with $\mathbf{T} = T\,\mathbf{I}_d$, where T is the texture RV and $\mathbf{I}_d$ is the $d \times d$ identity matrix. We note that the matrix $\mathbf{T}$ contains only one functionally independent element, namely, T. Without entering the stringent argument in terms of differential calculus, we state that

$$\begin{aligned}\phi_{\mathbf{T}}(s) &= \int_{\Omega_+} |\mathbf{T}|^{s-d} p_{\mathbf{T}}(\mathbf{T}) d\mathbf{T} \\ &= \int_0^\infty t^{d(s-d)} p_T(t) dt \end{aligned} \tag{70}$$

where we have used $|\mathbf{T}| = T^d$. It follows that

$$\begin{aligned}\phi_{\mathbf{T}}(s) &= \sum_{\nu=0}^{\infty} \frac{[d(s-d)]^\nu}{\nu!} \mu_\nu\{T\} \\ &= \phi_T\left(d(s-d)+1\right) \end{aligned} \tag{71}$$

which leads to

$$\phi_{\mathbf{C}}(s) = \phi_T\left(d(s-d)+1\right) \cdot \phi_{\tilde{\mathbf{W}}}(s) \tag{72}$$

by virtue of (67).

Equation (70) further implies

$$\mu_\nu\{\mathbf{T}\} = \left.\frac{d^\nu}{ds^\nu}\phi_{\mathbf{T}}(s)\right|_{s=d} = d^\nu \mu_\nu\{T\}. \tag{73}$$

Faá di Bruno's formula is then used to prove

$$\kappa_\nu\{\mathbf{T}\} = d^\nu \kappa_\nu\{T\} \tag{74}$$

and the matrix-variate version of the formula yields the MLCs of $\mathbf{C}$ as

$$\kappa_\nu\{\mathbf{C}\} = \mu_\nu\{\mathbf{C}\} - \sum_{i=1}^{\nu-1} \binom{\nu-1}{i-1} \kappa_i\{\mathbf{C}\}\mu_{\nu-i}\{\mathbf{C}\}. \tag{75}$$

The first MLCs are expressed as

$$\kappa_1\{\mathbf{C}\} = \mu_1\{\mathbf{C}\} \tag{76}$$

$$\kappa_2\{\mathbf{C}\} = \mu_2\{\mathbf{C}\} - \mu_1^2\{\mathbf{C}\} \tag{77}$$

$$\kappa_3\{\mathbf{C}\} = \mu_3\{\mathbf{C}\} - 3\mu_1\{\mathbf{C}\}\mu_2\{\mathbf{C}\} + 2\mu_1^3\{\mathbf{C}\}. \tag{78}$$

We use $\mathbf{C} = \mathbf{T}\tilde{\mathbf{W}}$ and (69) to prove that

$$\kappa_\nu\{\mathbf{C}\} = \kappa_\nu\{\tilde{\mathbf{W}}\} + d^\nu \kappa_\nu\{T\}. \tag{79}$$

After inserting the $\kappa_\nu\{\tilde{\mathbf{W}}\}$ from (47), the MLCs of $\mathbf{C}$ evaluate under the product model to

$$\kappa_1\{\mathbf{C}\} = \psi_d^{(0)}(L) + \ln|\boldsymbol{\Sigma}| - d\left(\ln L - \kappa_1\{T\}\right) \tag{80}$$

$$\kappa_\nu\{\mathbf{C}\} = \psi_d^{(\nu-1)}(L) + d^\nu \kappa_\nu\{T\}, \quad \nu > 1 \tag{81}$$

for a general texture variable with unspecified distribution. To obtain the MLCs of a specific distribution, the texture variable log-cumulants $\kappa_\nu\{T\}$, such as those listed in Table II, must be inserted.

We finally note that sample MLMs, denoted $\langle \mu_\nu\{\mathbf{C}\}\rangle$, are calculated with the simple sample mean estimator

$$\langle \mu_\nu\{\mathbf{C}\}\rangle = \frac{1}{N}\sum_{i=1}^{N} \left(\ln|\mathbf{C}_i|\right)^\nu \tag{82}$$

given a sample of N covariance matrices $\mathcal{C} = \{\mathbf{C}_i\}_{i=1}^N$. Sample MLCs $\langle \kappa_\nu\{\mathbf{C}\}\rangle$ are obtained from (75) by combining sample MLMs instead of theoretical MLMs.

IV. Applications

In this section, we discuss application of matrix-variate MKS. Before presenting specific examples of MoMLC algorithms for parameter estimation and demonstrating their effectiveness, we introduce the MLC diagram. The MLC diagram is not an application in its own right but serves as a visualization tool which efficiently explains some uses of MKS and provides intuition about the MoMLC. It is a straightforward extension of the log-cumulant diagrams used by Nicolas [4], [5] for univariate MKS.

A. Matrix Log-Cumulant Diagrams

The MLC diagram generally displays a q-dimensional space where each dimension represents one particular MLC with unique order ν. Let $\nu_1, \ldots, \nu_q$ be the orders of the chosen MLCs. In this MLC space, we plot: i) The manifolds spanned by the theoretical MLCs that can be attained under given models; and ii) points that represent the empirical sample MLCs computed from data samples.

Define $\boldsymbol{\vartheta}$ as the vector that contains all texture parameters of a certain distribution model. Thus, $\boldsymbol{\vartheta}_{\mathcal{K}} = [\alpha]$, $\boldsymbol{\vartheta}_{\mathcal{G}^0} = [\lambda]$ and $\boldsymbol{\vartheta}_{\mathcal{U}} = [\xi, \zeta]^T$ are the respective texture parameter vectors of the $\mathcal{K}$, $\mathcal{G}^0$ and $\mathcal{U}$ distribution. Assume that the parameters L and $\boldsymbol{\Sigma}$ are fixed, such that the theoretical MLCs only vary through $\boldsymbol{\vartheta}$. The MLC space manifold spanned by a general model is denoted

$$\mathcal{M}(\boldsymbol{\vartheta}) = \left\{\left(\kappa_{\nu_1}(\boldsymbol{\vartheta}), \kappa_{\nu_2}(\boldsymbol{\vartheta}), \ldots, \kappa_{\nu_q}(\boldsymbol{\vartheta})\right)\right\} \tag{83}$$

where we have changed the notation of the theoretical νth-order MLC from $\kappa_\nu\{\mathbf{C}\}$ to $\kappa_\nu(\boldsymbol{\vartheta})$ to emphasize that the points that constitute the manifold are functions of $\boldsymbol{\vartheta}$. The dimension of the manifold $\mathcal{M}(\boldsymbol{\vartheta})$ is the same as the dimension of the vector $\boldsymbol{\vartheta}$, that is, the number of texture parameters. We next define the vector of sample MLCs as

$$\langle \boldsymbol{\kappa}(\mathcal{C})\rangle = \left[\langle \kappa_{\nu_1}(\mathcal{C})\rangle, \ldots, \langle \kappa_{\nu_q}(\mathcal{C})\rangle\right] \tag{84}$$

and note that the sample MLCs have also been given a new notation $\langle \kappa_\nu(\mathcal{C})\rangle$ to stress that they are computed from the data sample $\mathcal{C}$.

Like Nicolas [4], [5], we concentrate on diagrams that plot the third-order log-cumulant against the second-order log-cumulant. We have shown that under the polarimetric product model, MLCs of order two and higher are independent of the scale matrix $\boldsymbol{\Sigma}$. Assuming that the equivalent number of looks L is a global constant for the data set, this diagram

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

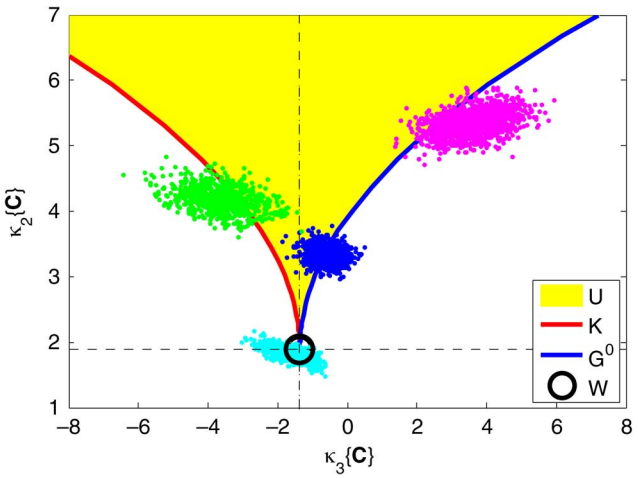

Fig. 5. Matrix log-cumulant diagram showing the manifolds of theoretical MLCs for the complex Wishart, $\mathcal{K}$, $\mathcal{G}^0$ and $\mathcal{U}$ distribution, as well as a collection of sample MLCs representing forest (green), ocean (blue), urban area (magenta), and a wheat crop (cyan).

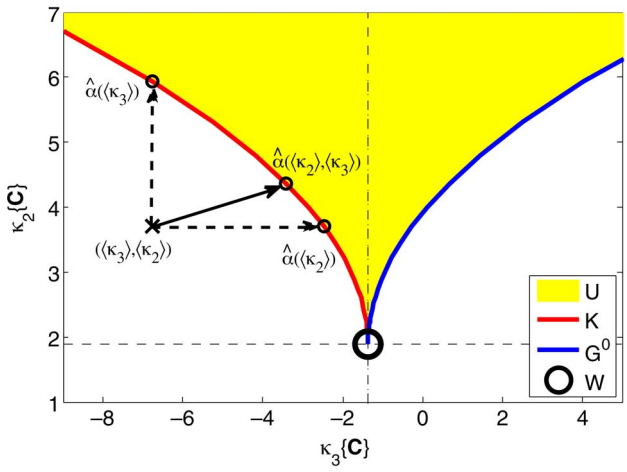

Fig. 6. MLC space interpretation of three estimators of the $\mathcal{K}$ distribution texture parameter α. The first estimator is based on $\langle\kappa_2\{\mathbf{C}\}\rangle$, the other is based on $\langle\kappa_3\{\mathbf{C}\}\rangle$, and the third is based on both.

is of particular interest since it displays the solitary impact of the texture parameters upon the models. Therefore, it also provides valuable insight about how the texture parameters can be estimated.

As seen in Fig. 5, the manifolds of our selected distribution models have different dimensions. The Wishart distribution has no texture parameters and is therefore represented by a point (black circle), which can be viewed as a zero-dimensional manifold. The texture of the $\mathcal{K}$ and $\mathcal{G}^0$ distribution is parameterized by one parameter, thus they are represented by a curve (red and blue, respectively) which is a 1-D manifold. The $\mathcal{U}$ distribution has two texture parameters and is represented by a surface (yellow area), which is a 2-D manifold. This is valid also when plotting higher dimensional tuples of MLCs, that is, for MLC spaces with dimension higher than two. The dashed coordinates are centered at the point which represents the Wishart distribution, more specifically at $[\kappa_3, \kappa_2] = [\psi_d^{(2)}(L), \psi_d^{(1)}(L)]$. The impact of nonzero texture on the MLCs is measured relative to these axes.

Given a sample $\mathcal{C} = \{\mathbf{C}_i\}_{i=1}^N$ of covariance matrices, we can compute the sample MLCs of order $\nu_1, \ldots, \nu_q$ and plot them as a point in MLC space. This has been done in Fig. 5 for one forest sample (shown as green points) and one wheat crop sample (cyan) taken from a polarimetric NASA/JPL AIRSAR C-band image of Flevoland, The Netherlands. We have also plotted sample MLCs computed from an ocean sample (blue) and an urban area sample (magenta) extracted from an image of San Francisco, United States, captured by the same sensor. Both images are from 1989. Multiple points have been obtained for each class by bootstrap sampling [42] of $\mathcal{C}$.

MoMLC parameter estimation can now be visualized as a projection of the sample MLCs onto the manifolds representing the models. The manifolds are functions of the texture parameters, and the parameter values at the projection point are assigned as an estimate. An estimator based on a single νth-order MLC relies on a projection in the direction normal to the corresponding coordinate. This is illustrated in Fig. 6, displaying estimators for the $\mathcal{K}$ distribution texture parameter α based on the point $(\langle\kappa_3(\mathcal{C})\rangle, \langle\kappa_2(\mathcal{C})\rangle)$ in MLC space, which is shown as the black symbol "$\times$". The estimators denoted $\hat{\alpha}(\langle\kappa_2\rangle)$ and $\hat{\alpha}(\langle\kappa_3\rangle)$ are based on the second- and third-order MLC equation, respectively. The dashed arrows visualize their projection of the sample MLC point onto the red curve representing the $\mathcal{K}$ distribution. Note that an estimator requires at least as many sample MLCs as the number of texture parameters to be estimated. For instance, the $\mathcal{K}$ and $\mathcal{G}_0$ distribution require one, while the $\mathcal{U}$ distribution requires two sample MLCs. The estimators $\hat{\alpha}(\langle\kappa_2\rangle)$ and $\hat{\alpha}(\langle\kappa_3\rangle)$ use exactly the required number.

As indicated, it is possible to design an estimator which is based on more sample MLCs than there are texture parameters, which implies that more information about $\boldsymbol{\vartheta}$ is extracted and $\mathcal{C}$ is utilized more efficiently. One way to combine the information contained in multiple sample MLCs is to derive a squared Mahalanobis distance (d_M^2) between the sample MLC point and the points on the model manifold. The estimation problem then reduces to a minimisation of the distance measure with respect to $\boldsymbol{\vartheta}$. To find an expression for d_M^2, we must derive the (approximate) mean values and covariance matrix of the sample MLCs.

The minimum of d_M^2 defines a new projection of the sample MLC point onto the model manifold, as illustrated by the solid arrow in Fig. 6. The sample MLC point is projected onto the point on the $\mathcal{K}$ distribution curve that minimizes d_M^2, and the associated value of α defines the estimate $\hat{\alpha}(\langle\kappa_2\rangle, \langle\kappa_3\rangle)$. We can see this estimate as a weighted mean of $\hat{\alpha}(\langle\kappa_2\rangle)$ and $\hat{\alpha}(\langle\kappa_3\rangle)$. The information content of an individual sample MLC is proportional to its precision (i.e., inverse variance) and determines its contribution to the overall estimate. The shape of the sample MLC clusters in Fig. 5 shows that the sample variance increases with MLC order, as expected.

A detailed derivation of the Mahalanobis distance is given in [43], where we also discuss the coupling of the estimation problem and the problem of measuring goodness-of-fit (GoF) for distribution models. The geometrical interpretation of distances in MLC space in terms of model fit is intuitive. We also find it much easier to observe deviations between data and model

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

in MLC space than by comparing data histograms with model densities, which is the alternative normally resorted to in the literature. This point is highlighted in [39]. In [43] we derive the sample distribution of d_M^2, such that formal GoF testing can be performed, which conforms with visual inspection of model fit in the MLC diagram.

B. Parameter Estimation

In this section, we discuss MKS-based estimation algorithms for parameters of the distributions presented in Table I.

1) Equivalent Number of Looks: The equivalent number of looks L can be estimated from the first-order MLC equation of the complex Wishart distribution. This yields the maximum likelihood solution proposed in [38], [44]

$$\psi_d^{(0)}(\hat{L}) - d\ln\hat{L} = \langle\kappa_1\{\mathbf{C}\}\rangle - \ln|\mathbf{\Sigma}| \tag{85}$$

where we must insert the first-order sample MLC and an estimate of $\mathbf{\Sigma}$ before solving for L by numerical methods. We use the maximum likelihood estimate of $\mathbf{\Sigma}$, which is the sample mean of $\mathcal{C}$ in the complex Wishart case.

In principle, we can also solve for L from the MLC equations of the product model in (11) and avoid the Wishart constraint. However, these MLC equations contain texture parameters already from the first order, and all unknown parameters must therefore be estimated jointly from a system of equations. Higher order MLCs can also be used to improve the estimator in (85). None of these approaches have been attempted in practice.

In the following, we shall assume that an estimate of L has been provided and treat it as a known constant.

2) Matrix-Variate $\mathcal{K}$ Distribution: Under this distribution, the texture parameter α is related to the second-order MLC through

$$\kappa_2\{\mathbf{C}\} = d^2\psi^{(1)}(\alpha) + \psi_d^{(1)}(L) \tag{86}$$

and the estimate $\hat{\alpha}_{A_1}$ is obtained by solving

$$\psi^{(1)}(\hat{\alpha}_{A_1}) = \frac{\langle\kappa_2\{\mathbf{C}\}\rangle - \psi_d^{(1)}(L)}{d^2}. \tag{87}$$

Alternative estimators are proposed in Frery *et al.* [45] and Doulgeris *et al.* [46], where the former is just a mono-pol version of the latter. Doulgeris' estimator is

$$\hat{\alpha}_D = \frac{d(Ld+1)}{L\widehat{\mathrm{Var}}\{\tau\} - d} \tag{88}$$

where $\tau = \mathrm{tr}(\hat{\mathbf{\Sigma}}^{-1}\mathbf{C})$. The derivation is shown in Appendix B. Another approach taken by Freitas and Frery *et al.* [2], [45] is to derive estimators from fractional moments of the mono-pol intensity C. By combining the half- and quarter-order moments, they found that

$$\frac{\Gamma^2\left(\hat{\alpha}_F+\frac{1}{4}\right)}{\Gamma(\hat{\alpha}_F)\Gamma\left(\hat{\alpha}_F+\frac{1}{2}\right)}\frac{\Gamma^2\left(L+\frac{1}{4}\right)}{\Gamma(L)\Gamma\left(L+\frac{1}{2}\right)} - \frac{\left\langle C^{\frac{1}{4}}\right\rangle^2}{\left\langle C^{\frac{1}{2}}\right\rangle} = 0 \tag{89}$$

which can be solved for $\hat{\alpha}_F$. This method provides one estimate per polarimetric channel. The final estimate is an average of the mono-pol estimates.

Averaging over mono-pol estimates can also be carried out for the mono-pol version of $\hat{\alpha}_{A_1}$ (i.e., with $d=1$), which is the estimator derived by Nicolas from univariate MKS [4], [5]. We denote this estimator as $\hat{\alpha}_N$ and include it in the comparison in order to quantify the gain of using the full polarimetric information contained in $\mathbf{C}$ with respect to the information contained in intensity channels only. On a historic note, we remark that the mono-pol MKS-based estimator of Nicolas was proposed earlier by Kreithen and Hogan [47] and Blacknell [48], although without relating it to Mellin transform theory.

The final estimator we present is the one we have proposed in [43] based on multiple MLCs, as discussed in Section IV-A. It is defined as

$$\hat{\alpha}_{A_2} = \arg\left\{\min_{\alpha}\left\{d_M^2\right\}\right\} \tag{90}$$

where the squared Mahalanobis distance

$$d_M^2 = \left(\langle\boldsymbol{\kappa}\rangle - \boldsymbol{\kappa}\right)^T\mathbf{K}^{-1}\left(\langle\boldsymbol{\kappa}\rangle - \boldsymbol{\kappa}\right) \tag{91}$$

contains the sample MLC vector $\langle\boldsymbol{\kappa}\rangle = [\langle\kappa_2\rangle, \langle\kappa_3\rangle]^T$, its mean vector $\boldsymbol{\kappa} = \mathrm{E}\{\langle\boldsymbol{\kappa}\rangle\} = [\kappa_2, \kappa_3]^T$, and the covariance matrix

$$\begin{aligned}\mathbf{K} &= \mathrm{Cov}\left\{\langle\boldsymbol{\kappa}\rangle\right\} \\ &= \begin{bmatrix}\kappa_4 + 2\kappa_2^2 & \kappa_5 + 6\kappa_2\kappa_3 \\ \kappa_5 + 6\kappa_2\kappa_3 & \kappa_6 + 9\kappa_2\kappa_4 + 9\kappa_3^2 + 6\kappa_2^3\end{bmatrix}.\end{aligned} \tag{92}$$

The sample MLCs are fixed after a data sample is collected, and the minimization is performed by varying $\boldsymbol{\kappa}$ and $\mathbf{K}$ that both depend on α through the theoretical MLCs.

Figs. 7 and 8 show bias and variance of all estimators, obtained from Monte Carlo simulations with $L=10$ and $\alpha=10$. They clearly show that the estimators based on the full polarimetric covariance matrix ($\hat{\alpha}_{A_1}$, $\hat{\alpha}_D$ and $\hat{\alpha}_{A2}$) outperform those based on intensities only ($\hat{\alpha}_N$ and $\hat{\alpha}_F$), both in terms of bias and variance. From the latter group, $\hat{\alpha}_F$ ranks slightly better than $\hat{\alpha}_N$. For the truly polarimetric estimators, we see that $\hat{\alpha}_{A_2}$ has the superior bias properties, while the bias of $\hat{\alpha}_{A_1}$ and $\hat{\alpha}_D$ is very similar. Estimator $\hat{\alpha}_{A_2}$ is best also when it comes to variance but is approached by $\hat{\alpha}_{A_1}$ as the sample size increases.

The behavior observed in Figs. 7 and 8 is representative not only for the chosen parameter values but for the parameter space as a whole. This also applies to the simulations reported in the next sections for other distributions.

3) Matrix-Variate $\mathcal{G}^0$ Distribution: For this distribution with texture parameter λ, the second-order MLC is

$$\kappa_2\{\mathbf{C}\} = d^2\psi^{(1)}(\lambda) + \psi_d^{(1)}(L) \tag{93}$$

which leads to an estimator $\hat{\lambda}_{A_1}$ by solving

$$\psi^{(1)}(\hat{\lambda}_{A_1}) = \frac{\langle\kappa_2\{\mathbf{C}\}\rangle - \psi_d^{(1)}(L)}{d^2} \tag{94}$$

that is identical to $\hat{\alpha}_{A_1}$. The method of Doulgeris, derived in Appendix B, yields

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

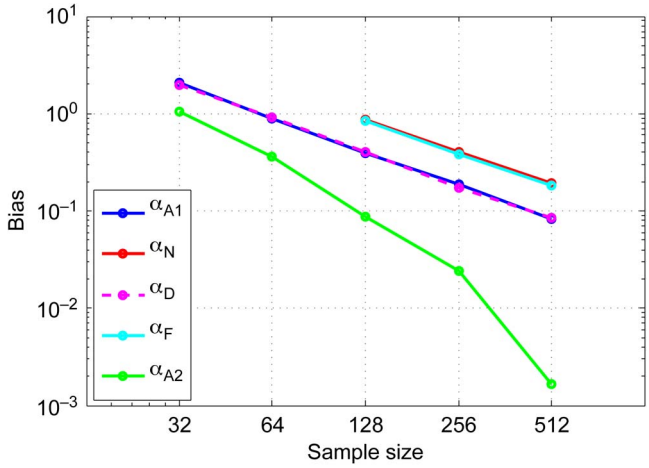

Fig. 7. Bias of estimators for the $\mathcal{K}$ distribution texture parameter α as function of sample size N.

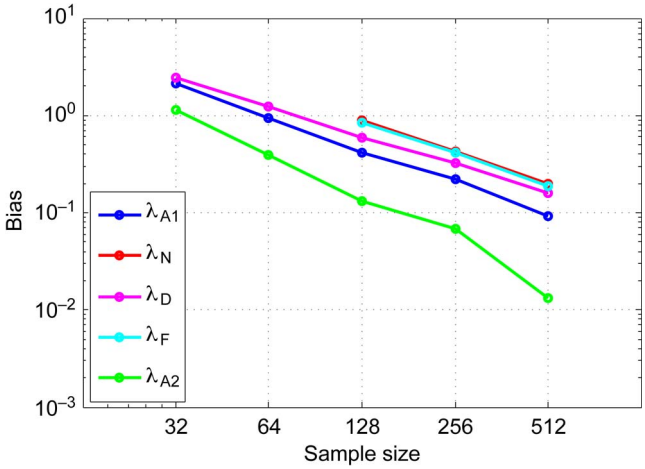

Fig. 9. Bias of estimators for the $\mathcal{G}^0$ distribution texture parameter λ.

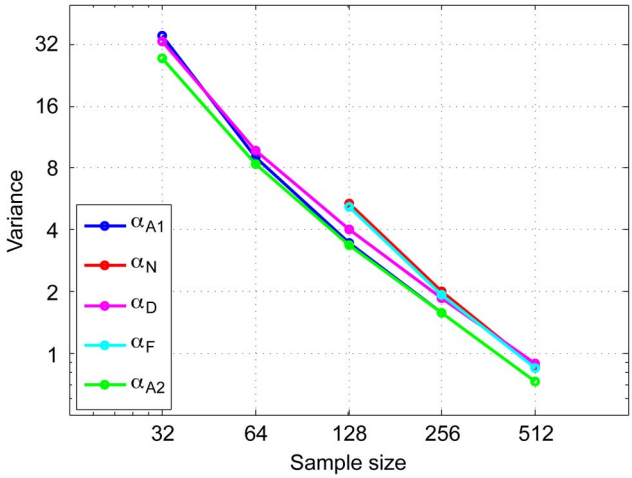

Fig. 8. Variance of estimators for the $\mathcal{K}$ distribution texture parameter α as function of sample size N.

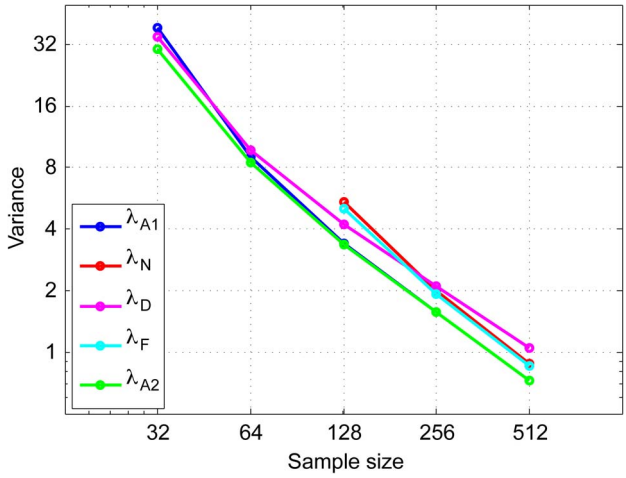

Fig. 10. Variance of estimators for the $\mathcal{G}^0$ distribution texture parameter λ.

$$\hat{\lambda}_D = \frac{2L\widehat{\mathrm{Var}}\{\tau\} + d(Ld-1)}{L\widehat{\mathrm{Var}}\{\tau\} - d} \tag{95}$$

while the fractional moment estimator is an average of the mono-pol estimates defined as the solution of

$$\frac{\Gamma^2\left(\hat{\lambda}_F - \frac{1}{4}\right)}{\Gamma(\hat{\lambda}_F)\Gamma\left(\hat{\lambda}_F - \frac{1}{2}\right)} \frac{\Gamma^2\left(L + \frac{1}{4}\right)}{\Gamma(L)\Gamma\left(L + \frac{1}{2}\right)} - \frac{\left\langle C^{\frac{1}{4}}\right\rangle^2}{\left\langle C^{\frac{1}{2}}\right\rangle} = 0. \tag{96}$$

In addition, an estimator $\hat{\lambda}_N$ is obtained by averaging the mono-pol estimates produced by $\hat{\lambda}_{A_1}$ for $d = 1$, while $\hat{\lambda}_{A_2}$ is defined in the same way as $\hat{\alpha}_{A_2}$, as given by (90).

We have performed Monte Carlo simulations for the estimators of λ with $\mathcal{G}^0$ distributed data parameterized by $L = 10$ and $\lambda = 10$. The bias and variance results in Figs. 9 and 10 are very similar to those reported for the estimators of α. The main difference is that $\hat{\lambda}_D$, which does not use MKS, is superseded in terms of variance by the MKS estimators based on intensities only, $\hat{\lambda}_N$ and $\hat{\lambda}_F$, for $N > 200$. The preferred estimator is $\hat{\lambda}_{A_2}$, due to its superior bias and variance. Estimator $\hat{\lambda}_{A_1}$ has a comparably low variance for $N > 100$, which makes it a good alternative, due to a slightly lower complexity. Note that the variance of estimators $\hat{\lambda}_{A_1}$ and $\hat{\lambda}_{A_2}$ are largely coincident in the upper range of sample sizes and can therefore be difficult to distinguish in Fig. 10.

4) $\mathcal{U}$ Distribution: This distribution has two texture parameters, ξ and ζ. The estimation procedure therefore requires two MLC equations

$$\kappa_2\{\mathbf{C}\} = d^2\left(\psi^{(1)}(\xi) + \psi^{(1)}(\zeta)\right) + \psi_d^{(1)}(L) \tag{97}$$

$$\kappa_3\{\mathbf{C}\} = d^3\left(\psi^{(2)}(\xi) - \psi^{(2)}(\zeta)\right) + \psi_d^{(2)}(L) \tag{98}$$

from which we can jointly determine the estimates $\hat{\xi}_{A_1}$ and $\hat{\zeta}_{A_1}$ by solving the equation system

$$\psi^{(1)}(\hat{\xi}_{A_1}) + \psi^{(1)}(\hat{\zeta}_{A_1}) = \frac{\langle\kappa_2\{\mathbf{C}\}\rangle - \psi_d^{(1)}(L)}{d^2} \tag{99}$$

$$\psi^{(2)}(\hat{\xi}_{A_1}) - \psi^{(2)}(\hat{\zeta}_{A_1}) = \frac{\langle\kappa_3\{\mathbf{C}\}\rangle - \psi_d^{(2)}(L)}{d^3}. \tag{100}$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

The alternative estimators are $\hat{\xi}_N$ and $\hat{\zeta}_N$, that is, the averaged mono-pol estimates obtained from $\hat{\xi}_{A_1}$ and $\hat{\zeta}_{A_1}$ with $d=1$. These have been implemented by the authors of [49].

The results for the $\mathcal{U}$ distribution estimators are similar to those reported for the $\mathcal{K}$ and $\mathcal{G}^0$ distributions, and are therefore omitted.

V. Conclusion

We have used a matrix-variate Mellin transform previously introduced by Mathai to extend the framework that we call Mellin kind statistics from the univariate to the matrix-variate case describing multilook polarimetric radar data. We have further defined the Mellin kind characteristic function and cumulant generating function for the matrix-variate case, and used them to define matrix log-moments and matrix log-cumulants. We have then proven the matrix-variate Mellin convolution theorem and used it to develop expressions for Mellin kind statistics of the multilook polarimetric product model. Specific expressions for important distributions, such as the matrix-variate $\mathcal{K}$ distribution, $\mathcal{G}^0$ distribution, and $\mathcal{U}$ distribution, have been given.

Mellin kind moments and cumulants are computed on a logarithmic scale, and the impact of speckle and texture therefore can be separated in the matrix-variate log-cumulant domain, which provides a valuable analysis tool for the doubly stochastic product model. Simulations have demonstrated the superior bias and variance properties possessed by estimators derived with the method of matrix log-cumulants. We have also used matrix log-cumulant space as a visualization tool to provide intuition about estimation algorithms and model assessment that uses Mellin kind statistics. The mathematical tractability and the simplicity of the obtained expressions show, together with the excellent estimator properties documented, that the matrix-variate Mellin transform is a natural tool for analysis of multilook polarimetric radar data.

Appendix A
Mellin Kind Statistics for the Complex Wishart Distribution

Let $\mathbf{W} \sim \mathcal{W}_d^{\mathbb{C}}(L, \boldsymbol{\Sigma})$ follow the complex Wishart distribution given in (7). The matrix-valued Mellin transform of $p_{\mathbf{W}}(\mathbf{W}; L, \boldsymbol{\Sigma})$, and hence the Mellin kind CF of the random matrix $\mathbf{W}$ is then

$$\begin{aligned}\phi_{\mathbf{W}}(s) &= \mathcal{M}\left\{p_{\mathbf{W}}(\mathbf{W})\right\}(s) \\ &= \int_{\Omega_+} |\mathbf{W}|^{s-d} p_{\mathbf{W}}(\mathbf{W}) d\mathbf{W} \\ &= \frac{\Gamma_d(L+s-d)}{\Gamma_d(L)} \frac{|\boldsymbol{\Sigma}|^{L+s-d}}{|\boldsymbol{\Sigma}|^L} \\ &\quad \times \int_{\Omega_+} p_{\mathbf{W}}(\mathbf{W}; L+s-d, \boldsymbol{\Sigma}) d\mathbf{W} \\ &= \frac{\Gamma_d(L+s-d)}{\Gamma_d(L)} |\boldsymbol{\Sigma}|^{s-d}. \end{aligned} \tag{101}$$

Accordingly, the Mellin kind CGF is

$$\begin{aligned}\varphi_{\mathbf{W}}(s) &= \ln \phi_{\mathbf{W}}(s) \\ &= \ln \Gamma_d(L+s-d) - \ln \Gamma_d(L) + (s-d) \ln |\boldsymbol{\Sigma}|. \end{aligned} \tag{102}$$

We shall use the result

$$\begin{aligned}\frac{d}{dL}\Gamma_d(L) &= \frac{d}{dL}\left(\pi^{d(d-1)/2} \prod_{i=0}^{d-1} \Gamma(L-i)\right) \\ &= \pi^{d(d-1)/2} \sum_{i=0}^{d-1} \left(\frac{d}{dL}\Gamma(L-i) \prod_{\substack{j=0 \\ j\neq i}}^{d-1} \Gamma(L-j)\right) \\ &= \pi^{d(d-1)/2} \prod_{j=0}^{d-1} \Gamma(L-j) \sum_{i=0}^{d-1} \psi^{(0)}(L-i) \\ &= \Gamma_d(L) \psi_d^{(0)}(L) \end{aligned} \tag{103}$$

which is obtained by straightforward application of the product rule of differentiation. We have also utilized the well-known relation

$$\frac{d}{dL}\Gamma(L) = \Gamma(L)\psi^{(0)}(L) \tag{104}$$

and the multivariate polygamma function introduced in (45). Remark that (103) is a multivariate version of (104). We also need the result

$$\frac{d}{dL}\psi_d^{(\nu)}(L) = \psi_d^{(\nu+1)}(L) \tag{105}$$

whose proof is trivial.

Equations (103) and (105) are used to deduce the derivatives of $\varphi_{\mathbf{W}}(s)$, denoted as $\varphi_{\mathbf{W}}^{(\nu)}(s) = (d^\nu/ds^\nu)\varphi_{\mathbf{W}}(s)$. The MLCs can then be written as $\kappa_\nu\{\mathbf{W}\} = \varphi_{\mathbf{W}}^{(\nu)}(d)$. By repeated differentiation of (102) and induction, we find that

$$\kappa_1\{\mathbf{W}\} = \psi_d^{(0)}(L) + \ln|\boldsymbol{\Sigma}| \tag{106}$$

$$\kappa_\nu\{\mathbf{W}\} = \psi_d^{(\nu-1)}(L), \quad \nu > 1. \tag{107}$$

Let $\mathbf{X}$ be a $d \times d$ complex positive definite matrix and $\mathbf{A}$ an equal-size real constant matrix. The scaling property of the matrix-variate MT

$$\mathcal{M}\left\{f(\mathbf{AX})\right\}(s) = |\mathbf{A}|^{-s} \mathcal{M}\left\{f(\mathbf{X})\right\}(s) \tag{108}$$

is easily verified by evaluating the integral with a simple substitution of variables. For $\mathbf{A} = a\mathbf{I}_d$ with a real and positive scalar constant a, we get $\mathcal{M}\{f(a\mathbf{X})\}(s) = a^{-ds}\mathcal{M}\{f(\mathbf{X})\}(s)$. This is used to show that

$$\begin{aligned}\mathcal{M}\left\{p_{\tilde{\mathbf{W}}}(\tilde{\mathbf{W}})\right\}(s) &= L^{d^2}\mathcal{M}\left\{p_{\mathbf{W}}(L\tilde{\mathbf{W}})\right\}(s) \\ &= L^{-d(s-d)}\mathcal{M}\left\{p_{\mathbf{W}}(\mathbf{W})\right\}(s). \end{aligned} \tag{109}$$

Recall the definition of $\tilde{\mathbf{W}} = \mathbf{W}/L$, which gives

$$\begin{aligned}\phi_{\tilde{\mathbf{W}}}(s) &= L^{-d(s-d)}\phi_{\mathbf{W}}(s) \\ &= \frac{\Gamma_d(L+s-d)}{\Gamma_d(L)} \left(\frac{|\boldsymbol{\Sigma}|}{L^d}\right)^{s-d}. \end{aligned} \tag{110}$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

This is used to show that

$$\varphi_{\tilde{\mathbf{W}}}(s) = \varphi_{\mathbf{W}}(s) - (s-d)d\ln L \tag{111}$$

and the MLCs of $\tilde{\mathbf{W}}$ follow immediately as given in (47).

Appendix B
Doulgeris' Parameter Estimators

Doulgeris *et al.* [46] derived their estimator for the texture parameter of the matrix-variate $\mathcal{K}$ distribution from moments of the Hotelling–Lawley trace. This is an important test statistic in multivariate statistics, defined as

$$\tau = \mathrm{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{C}). \tag{112}$$

It is easily shown that $\mathrm{E}\{\tau\} = d$, and the variance of τ is

$$\mathrm{Var}\{\tau\} = \mathrm{E}\{T^2\}\left(d^2 + \frac{d}{L}\right) - d^2. \tag{113}$$

Given a choice of the texture RV T, we may solve for the texture parameter to obtain an estimator. For instance, with $T \sim \bar{\gamma}(\alpha)$ we have $\mathrm{E}\{T^2\} = (\alpha+1)/\alpha$, which yields the following estimator for the $\mathcal{K}$ distribution parameter α:

$$\hat{\alpha}_D = \frac{d(Ld+1)}{L\widehat{\mathrm{Var}}\{\tau\} - d}. \tag{114}$$

This estimator was given in [46]. The variance of τ is estimated with a standard variance estimator from a population of Hotelling–Lawley traces, $\{\tau_i = \mathrm{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{C}_i)\}_{i=1}^N$.

When $T \sim \bar{\gamma}^{-1}(\lambda)$, we have $\mathrm{E}\{T^2\} = (\lambda-1)/(\lambda-2)$ which is used to derive

$$\hat{\lambda}_D = \frac{2L\widehat{\mathrm{Var}}\{\tau\} + d(Ld-1)}{L\widehat{\mathrm{Var}}\{\tau\} - d} \tag{115}$$

for the $\mathcal{G}^0$ distribution parameter λ. The method is not pursued for the $\mathcal{U}$ distribution, since it would require derivation of higher moments of τ to solve for both texture parameters.

Acknowledgment

The authors would like to thank A. Doulgeris at the University of Tromsø and G. Moser at the University of Genoa for their valuable contribution through discussions of the topic and comments on the manuscript. We also acknowledge the anonymous reviewers for their suggestions that have helped improve the paper.

References

[1] R. Touzi, W. M. Boerner, J.-S. Lee, and E. Lüneburg, "A review of polarimetry in the context of synthetic aperture radar: Concepts and information extraction," *Can. J. Remote Sens.*, vol. 30, no. 3, pp. 380–407, 2004.

[2] C. C. Freitas, A. C. Frery, and A. H. Correia, "The polarimetric G distribution for SAR data analysis," *Environmetrics*, vol. 16, no. 1, pp. 13–31, Feb. 2005.

[3] B. Epstein, "Some applications of the Mellin transform in statistics," *Ann. Math. Stat.*, vol. 19, no. 3, pp. 370–379, Sep. 1948.

[4] J.-M. Nicolas, "Introduction aux statistique de deuxième espèce: Application des logs-moments et des logs-cumulants à l'analyse des lois d'images radar," *Traitement du Signal*, vol. 19, no. 3, pp. 139–167, 2002, English translation in [40].

[5] J.-M. Nicolas, "Application de la transformèe de Mellin: Étude des lois statistiques de l'imagerie cohèrente," Ecole Nationale Supèrieure des Tèlècommunications, Paris, France, Tech. Rep. 2006D010, 2006.

[6] R. D. Pierce, "Application of the positive alpha-stable distribution," in *Proc. IEEE SPW-HOS*, Banff, Canada, Jul. 1997, pp. 420–424.

[7] E. W. Stacy, "A generalization of the gamma distribution," *Ann. Math. Stat.*, vol. 33, no. 3, pp. 1187–1192, Sep. 1962.

[8] E. W. Stacy and G. A. Mihram, "Parameter estimation for a generalized gamma distribution," *Technometrics*, vol. 7, no. 3, pp. 349–358, Aug. 1965.

[9] G. Moser, J. Zerubia, and S. B. Serpico, "Dictionary-based stochastic expectation-maximization for SAR amplitude probability density function estimation," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 1, pp. 188–200, Jan. 2006.

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

[11] L. Bombrun and J.-M. Beaulieu, "Fisher distribution for texture modeling of polarimetric SAR data," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 3, pp. 512–516, Jul. 2008.

[12] J.-M. Nicolas, "A Fisher-MAP filter for SAR image processing," in *Proc. IEEE IGARSS*, Toulouse, France, Jul. 21–25, 2003, vol. 3, pp. 1996–1998.

[13] A. Achim, E. E. Kuruolu, and J. Zerubia, "SAR image filtering based on the heavy-tailed Rayleigh model," *IEEE Trans. Image Process.*, vol. 15, no. 9, pp. 2686–2693, Sep. 2006.

[14] G. Chen and X. Liu, "Wavelet-based SAR image despeckling using Cauchy pdf modeling," in *Proc. IEEE Radar Conf.*, Rome, Italy, May 26–30, 2008, pp. 1–5.

[15] J. J. Ranjani and S. J. Thiruvengadam, "Dual-tree complex wavelet transform based SAR despeckling using interscale dependence," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 6, pp. 2723–2731, Jun. 2010.

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

[17] D. Benboudjema, F. Tupin, W. Pieczynski, M. Sigelle, and J.-M. Nicolas, "Unsupervised segmentation of SAR images using triplet Markov fields and Fisher noise distributions," in *Proc. IEEE IGARSS*, Barcelona, Spain, Jul. 23–27, 2007, vol. 1, pp. 3891–3894.

[18] F. Galland, J.-M. Nicolas, H. Sportouche, M. Roche, F. Tupin, and P. Rèfrègier, "Unsupervised synthetic aperture radar image segmentation using Fisher distributions," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 8, pp. 2966–2972, Aug. 2009.

[19] F. Bujor, E. Trouvè, L. Valet, J.-M. Nicolas, and J.-P. Rudant, "Application of log-cumulants to the detection of spatiotemporal discontinuities in multitemporal SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 10, pp. 2073–2084, Oct. 2004.

[20] G. Moser and S. B. Serpico, "Generalized minimum-error thresholding for unsupervised change detection from SAR amplitude imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 10, pp. 2972–2982, Oct. 2006.

[21] G. Moser and S. B. Serpico, "Unsupervised change detection from multichannel SAR data by Markovian data fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 7, pp. 2114–2128, Jul. 2009.

[22] R. Abdelfattah and J.-M. Nicolas, "Interferometric SAR coherence magnitude estimation using second kind statistics," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, pt. 2, no. 7, pp. 1942–1953, Jul. 2006.

[23] C. Valade and J.-M. Nicolas, "Homomorphic wavelet transform and new subband statistics models for SAR image compression," in *Proc. IEEE IGARSS*, Anchorage, AK, Sep. 20–24, 2004, vol. 1, pp. 285–288.

[24] A. M. Mathai, *Jacobians of Matrix Transformations and Functions of Matrix Arguments*. New York: World Scientific, 1997, ch. 6.

[25] J.-S. Lee and E. Pottier, *Polarimetric Radar Imaging: From Basics to Applications*. Boca Raton, FL: CRC Press, 2009.

[26] N. R. Goodman, "Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction)," *Ann. Math. Stat.*, vol. 34, no. 1, pp. 152–177, Mar. 1963.

[27] J.-S. Lee, D. L. Schuler, R. H. Lang, and K. J. Ranson, "K-distribution for multi-look processed polarimetric SAR imagery," in *Proc. IEEE IGARSS*, Pasadena, CA, Aug. 1994, vol. 4, pp. 2179–2181.

[28] C. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar Images,* 2nd ed. Raleigh, NC: SciTech Publishing, 2004.

[29] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes,* 4th ed. New York: McGraw-Hill, 2002.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

[30] W. P. Johnson, "The curious history of Faà di Bruno's formula," *Amer. Math. Monthly*, vol. 109, no. 3, pp. 217–234, Mar. 2002.
[31] L. Comtet, *Advanced Combinatorics: The Art of Finite and Infinite Expansions*. Dordrecht, The Netherlands: Reidel, 1974, ch. 3.3, pp. 133–136.
[32] H. H. Andersen, M. Højbjerre, D. Sørensen, and P. S. Eriksen, *Linear and Graphical Models for the Multivariate Complex Normal Distribution*. New York: Springer-Verlag, 1995.
[33] C.-Y. Chi, C.-C. Feng, C.-H. Chen, and C.-Y. Chen, *Blind Equalization and System Identification*. London, U.K.: Springer-Verlag, 2006, ch. 3, pp. 103–104.
[34] T. Kollo and D. von Rosen, *Advanced Multivariate Statistics With Matrices*. Dordrecht, The Netherlands: Springer-Verlag, 2005.
[35] J. Bertrand, P. Bertrand, and J.-P. Ovarlez, "The Mellin transform," in *The Transform and Applications Handbook*, A. D. Poularikas, Ed., 2nd ed. Boca Raton, FL: CRC Press, 2000, ch. 11.
[36] A. M. Mathai, "Some results on functions of matrix argument," *Math. Nachr.*, vol. 84, no. 1, pp. 171–177, 1978.
[37] A. M. Mathai, "Distribution of the canonical correlation matrix," *Ann. Inst. Stat. Math.*, vol. 33, pt. A, no. 1, pp. 35–43, Dec. 1981.
[38] S. N. Anfinsen, A. P. Doulgeris, and T. Eltoft, "Estimation of the equivalent number of looks in polarimetric synthetic aperture radar imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 11, pp. 3795–3809, Nov. 2009.
[39] S. N. Anfinsen, T. Eltoft, and A. P. Doulgeris, "A relaxed Wishart model for polarimetric SAR data," in *Proc. 4th Int. Workshop Sci. Appl. POLinSAR*, Frascati, Italy, Apr. 2009, vol. ESA SP-668, 8 pp.
[40] S. N. Anfinsen, "Statistical analysis of multilook polarimetric radar images with the Mellin transform," Ph.D. dissertation, University of Tromsø, Tromsø, Norway, May, 2010.
[41] D. Casasent and D. Psaltis, "Deformation invariant, space-invariant optical pattern recognition," in *Progress in Optics*, E. Wolf, Ed. New York: North-Holland, 1978, ch. V.
[42] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*. Boca Raton, FL: Chapman & Hall, 1993.
[43] S. N. Anfinsen and T. Eltoft, "Goodness-of-fit tests for multilook polarimetric radar data based on the Mellin transform," *IEEE Geosci. Remote Sens. Lett.*, vol. 49, no. 8, Aug. 2011, 18 pp., to be published.
[44] S. N. Anfinsen, A. P. Doulgeris, and T. Eltoft, "Estimation of the equivalent number of looks in polarimetric SAR imagery," in *Proc. IEEE IGARSS*, Boston, MA, Jul. 2008, vol. 4, pp. 487–490.
[45] A. C. Frery, A. H. Correia, and C. C. Freitas, "Classifying multifrequency fully polarimetric imagery with multiple sources of statistical evidence and contextual information," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3098–3109, Oct. 2007.
[46] A. P. Doulgeris, S. N. Anfinsen, and T. Eltoft, "Classification with a non-Gaussian model for PolSAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 10, pp. 2999–3009, Oct. 2008.
[47] D. E. Kreithen and G. G. Hogan, "Statistical analysis of Ka-band sea clutter," in *Proc. IEEE OCEANS*, Honolulu, HI, Oct. 1991, vol. 2, pp. 1217–122.
[48] D. Blacknell, "Comparison of parameter estimators for K-distribution," *Proc. Inst. Elect. Eng.—Radar, Sonar, Navig.*, vol. 141, no. 1, pp. 45–52, Feb. 1994.
[49] O. Harant, L. Bombrun, M. Gay, R. Fallourd, E. Trouvè, and F. Tupin, "Segmentation and classification of polarimetric SAR data based on the KummerU distribution," in *Proc. 4th Int. Workshop Sci. Appl. POLinSAR*, Frascati, Italy, Apr. 2009, vol. ESA SP-668, 6 pp.

Stian Normann Anfinsen (S'06–M'10) received the Cand.Mag. (B.Sc.), Cand.Scient. (M.Sc.), and Ph.D. degrees in physics from the University of Tromsø, Tromsø, Norway, in 1997, 2000, and 2010, respectively, and the M.Sc. degree in communications, control and digital signal processing from the University of Strathclyde, Glasgow, U.K., in 1998.

He was with satellite ground station system provider Kongsberg Spacetec in Tromsø from 2001 to 2005, before he moved to the University of Tromsø, Department of Physics and Technology, where he is currently a Postdoctoral Research Fellow and also affiliated with the Barents Remote Sensing School. His current research focuses on statistical modeling, parameter estimation, classification, change detection, and target detection in polarimetric radar images.

Torbjørn Eltoft (A'92–M'98) received the Cand.Real. (M.Sc.) and Dr.Scient. degrees from the University of Tromsø, Tromsø, Norway, in 1981 and 1984, respectively.

Since 1988, he has been with the University of Tromsø, where he is currently a Professor in electrical engineering in the Department of Physics and Technology. He is also Director of the Barents Remote Sensing School, a Ph.D. school at the department, and an Adjunct Professor at the Northern Research Institute, Tromsø. He is an Associate Editor for *Pattern Recognition*. His current research interests include multidimensional signal and image analysis with application in remote sensing, statistical models, neural networks, and machine learning.

Dr. Eltoft is the recipient of the year 2000 Outstanding Paper Award in Neural Networks from the IEEE Neural Networks Council and an Honorable Mention in the year 2003 Pattern Recognition Journal Best Paper Award.