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

Stian Normann Anfinsen, *Student 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 multi-look polarimetric radar data. The domain of the transform integral is the cone of complex positive definite matrices, which allows for transformation of the covariance matrix distributions used to model multilook polarimetric radar data. 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*—Radar polarimetry, synthetic aperture radar, Mellin transform, matrix-variate statistics, parameter estimation, method-of-log-cumulants, doubly stochastic product model**

I. Introduction

POLARIMETRIC radar has become an important remote sensing instrument due to its ability to discriminate between different scattering mechanisms. It can therefore characterise physical properties of the target that cannot be determined from single polarisation (mono-pol) radar measurements. To fully utilise the polarimetric information captured, it is necessary to analyse the complex correlations between all polarimetric channels, incorporating all intensity and phase information. This requires relatively complicated data models, that together with the speckle phenomenon, inherent to all types of coherent imaging, make analysis of multiple polarization radar data a challenging task.

It was noted already by Epstein [1] 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 [2], [3] utilised 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 [2], [3], which Nicolas applied to a number of doubly stochastic distributions, as well as the positive alpha-stable distribution [4], and members of the generalised gamma distribution (GγD) family (e.g., the Weibull and log-normal distribution) [5]. The same method has earlier been applied to GγDs [6], 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 modelling [7], [8], [9], speckle filtering [10], [11], [12], classification [13], segmentation [14], [15], change detection [16], [17], [18], interferometric coherence estimation [19], and image compression [20].

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 [21], 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 organised as follows. In Section II

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

we describe the data delivered by mono-pol and polarimetric radars, together with the probability density functions (PDFs) commonly used to model the data. In Section III 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 for convenience:

NOMENCLATURE

CF	characteristic function
CGF	cumulant generating function
FT	Fourier transform
GγD	generalised 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

II. RADAR DATA MODELS

A full-polarimetric imaging radar separately transmits orthogonally polarised microwave pulses, and measures orthogonal polarisations 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 polarisation, and assuming no atmospheric disturbance (i.e. zero 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 polarisations x and y. The superscript indicates transmitted/incident ($\downarrow$) or received/backscattered ($\uparrow$) wave. SAR systems normally use linear polarisations (horizontal and vertical), while using circular polarisations (left and right) is another choice. The scattering coefficients S_{ij} are subscripted with the associated receive and transmit polarisation, 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 [22], since the vector can be linearly transformed to emphasise physical interpretations of the elements (i.e., Pauli basis), or the dimension can be reduced to $d = 3$ by assuming reciprocity of the target (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 polarisation SAR ($d = 2$).

Radar images are affected by an interference phenomenon which is a characteristic of all coherent imaging systems. The noise-like 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 characterised 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 $\mathbf{s} \sim \mathcal{N}_d^{\mathbb{C}}(\mathbf{0}, \boldsymbol{\Sigma})$, where $\mathbf{0}$ is a column vector of zeros and $\boldsymbol{\Sigma} = \mathrm{E}\{\mathbf{s}\mathbf{s}^H\}$ is the covariance matrix of $\mathbf{s}$. $\mathrm{E}\{\cdot\}$ denotes the expectation value. The PDF of $\mathbf{s}$ is

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

where $|\cdot|$ is the determinant operator. It follows that if $L \geq d$ and the $\mathbf{s}_\ell$ are independent, then the scaled sample covariance matrix, defined as $\mathbf{Z} = L\mathbf{C}$, follows the nonsingular complex Wishart distribution [23]:

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

where $\mathrm{etr}(\cdot) = \exp(\mathrm{tr}(\cdot))$ is the exponential trace operator and $\boldsymbol{\Sigma} = \mathrm{E}\{\mathbf{C}\} = \mathrm{E}\{\mathbf{Z}\}/L$. We write this as $\mathbf{Z} \sim \mathcal{W}_d^{\mathbb{C}}(L,\boldsymbol{\Sigma})$. The normalisation constant $\Gamma_d(L)$ is the multivariate gamma function of the complex case, 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 $\mathbf{Z} = L\mathbf{C}$ [21]. The PDF of $\mathbf{C}$ becomes

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

In the one-dimensional 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 $C \sim \gamma(L,\sigma)$.

For the Gaussian model, we denote the scaled covariance matrix $\mathbf{Z}$ by $\mathbf{W}$, and $\mathbf{C}$ by $\tilde{\mathbf{W}}$. We also refer to the PDF in (7) as $p_{\mathbf{W}}(\mathbf{W})$, to emphasise that it is a complex Wishart distribution.

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 modelling 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 [26], [27], has been shown to be both mathematically tractable and successful for modelling and prediction purposes in coherent imaging. In the multilook polarimetric version, which is extensively reviewed in [25], it decomposes $\mathbf{Z}$ as a product of two independent stochastic variables,

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

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

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	$\delta(t-1)$	$\mathcal{W}_d^{\mathbb{C}}(\mathbf{\Sigma}, L)$	$\frac{L^{Ld}}{\Gamma_d(L)} \frac{\lvert\mathbf{C}\rvert^{L-d}}{\lvert\mathbf{\Sigma}\rvert^{L}} \operatorname{etr}(-L\mathbf{\Sigma}^{-1}\mathbf{C})$	[23]
$\bar{\gamma}(\alpha)$	$\frac{\alpha^\alpha}{\Gamma(\alpha)} t^{\alpha-1} \exp(-\alpha t)$	$\mathcal{K}_d(\mathbf{\Sigma}, L, \alpha)$	$\frac{2\lvert\mathbf{C}\rvert^{L-d}(L\alpha)^{\frac{\alpha+Ld}{2}}}{\lvert\mathbf{\Sigma}\rvert^{L}\Gamma_d(L)\Gamma(\alpha)} \left(\operatorname{tr}(\mathbf{\Sigma}^{-1}\mathbf{C})\right)^{\frac{\alpha-Ld}{2}} K_{\alpha-Ld}\left(2\sqrt{L\alpha \operatorname{tr}(\mathbf{\Sigma}^{-1}\mathbf{C})}\right)$	[24]
$\bar{\gamma}^{-1}(\lambda)$	$\frac{(\lambda-1)^\lambda}{\Gamma(\lambda)} \frac{1}{t^{\lambda+1}} \exp\left(-\frac{\lambda-1}{t}\right)$	$G_d^0(\mathbf{\Sigma}, L, \lambda)$	$\frac{L^{Ld}\lvert\mathbf{C}\rvert^{L-d}}{\Gamma_d(L)\lvert\mathbf{\Sigma}\rvert^{L}} \frac{\Gamma(Ld+\lambda)(\lambda-1)^\lambda}{\Gamma(\lambda)} \left(L\operatorname{tr}(\mathbf{\Sigma}^{-1}\mathbf{C}) + \lambda - 1\right)^{-\lambda-Ld}$	[25]
$\bar{\mathcal{F}}(\xi, \zeta)$	$\frac{\Gamma(\xi+\zeta)}{\Gamma(\xi)\Gamma(\zeta)} \frac{\xi}{\zeta-1} \frac{\left(\frac{\xi}{\zeta-1}t\right)^{\xi-1}}{\left(\frac{\xi}{\zeta-1}t+1\right)^{\xi+\zeta}}$	$\mathcal{U}_d(\mathbf{\Sigma}, L, \xi, \zeta)$	$\frac{L^{Ld}\lvert\mathbf{C}\rvert^{L-d}}{\Gamma_d(L)\lvert\mathbf{\Sigma}\rvert^{L}} \frac{\Gamma(\xi+\zeta)}{\Gamma(\xi)\Gamma(\zeta)} \left(\frac{\xi}{\zeta-1}\right) \Gamma(Ld+\zeta)$ $\times U\left(Ld+\zeta, Ld-\xi+1, L\operatorname{tr}(\mathbf{\Sigma}^{-1}\mathbf{C})\xi/(\zeta-1)\right)$	[9]

with individual distributions. The positive, scalar and unit mean RV T generates texture, assuming that its contribution is independent of polarisation 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.

In [25], the family of generalised 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 normalised 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 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

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

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

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

[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.

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 [28]. Then

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

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 [29]:

$$\begin{aligned} c_\nu\{X\} &= m_\nu\{X\} \\ &\quad - \sum_{i=1}^{\nu-1}\binom{\nu-1}{i-1}c_i\{X\}m_{\nu-i}\{X\} \end{aligned} \tag{20}$$

and reversely through

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

where $B_\nu(\cdot)$ is the νth complete Bell polynomial [30]. All relations for classical univariate statistics are summarised in the diagram of Figure 1.

Moments and cumulants can be generalised 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 [31], [32]

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

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

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

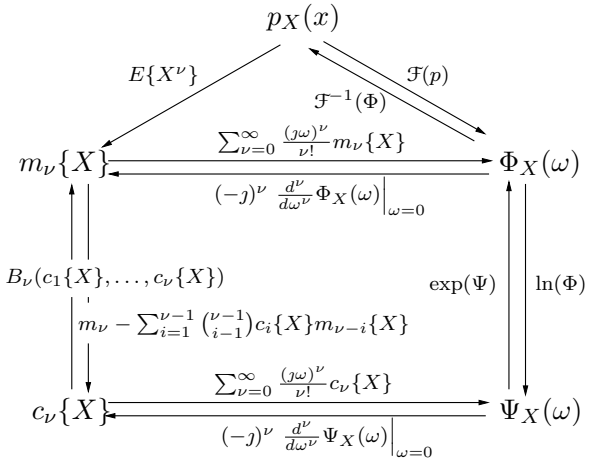

Fig. 1. Relations in univariate classical statistics.

where $\mathrm{tr}(\cdot)$ is the trace operator and $\Re\mathrm{e}\{\cdot\}$ 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 $\mathrm{tr}(\boldsymbol{\Omega}^H\mathbf{Z})$ [31].

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 [33], 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 [1], [34]

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

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

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

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

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

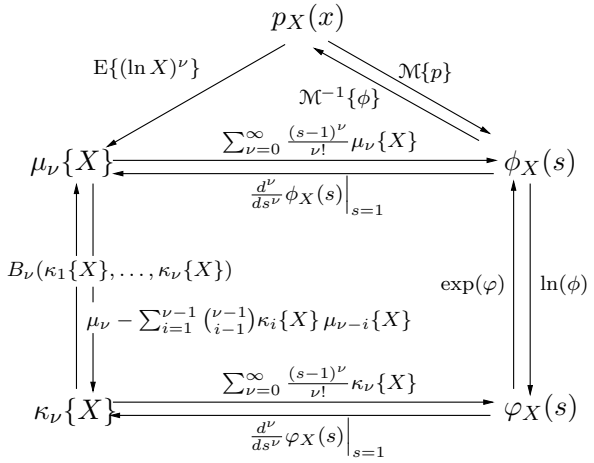

Fig. 2. Relations in univariate Mellin kind statistics.

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 [2] 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

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

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

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

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

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

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

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

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. Figure 2 summarises all relations for the univariate MKS.

Nicolas derived the MKS for the gamma, the inverse gamma, and the Fisher-Snedecor distribution [2], [3], 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 generalised transform (that he named the M-transform) for matrix-valued functions in [35], and referred to it in [36] as a generalised

MT. The complex version relevant to our study was presented in [21].

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

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

with transform variable $s \in \mathbb{C}$, whenever the integral exists.

It is duly noted in [35] 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. This problem is not associated with the multivariate MT defined in [37] as

$$\begin{aligned}&\mathcal{M}\{f(\mathbf{z})\}(s_1, \ldots, s_d)\\ &\quad = \int_0^\infty \cdots \int_0^\infty \prod_{i=1}^{d} z_i^{s_i - 1} f(\mathbf{z})\, dz_1 \cdots\, dz_d\end{aligned} \tag{35}$$

with $\mathbf{z} = [z_1, \ldots, z_d]^T \in \mathbb{C}^d$ and f defined on $\mathbb{C}^d$, which can in principle be extended to the matrix-variate case. Nevertheless, we shall refer to (34) as the matrix-variate MT.

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

$$\phi_{\mathbf{Z}}(s) = \mathrm{E}\{|\mathbf{Z}|^{s-d}\} = \mathcal{M}\{p_{\mathbf{Z}}(\mathbf{Z})\}(s) \tag{36}$$

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

$$\mu_\nu\{\mathbf{Z}\} = \left.\frac{d^\nu}{ds^\nu} \phi_{\mathbf{Z}}(s)\right|_{s=d}. \tag{37}$$

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

$$\begin{aligned}\phi_{\mathbf{Z}}(s) &= \int_{\Omega_+} e^{(s-d)\ln|\mathbf{Z}|} p_{\mathbf{Z}}(\mathbf{Z})\, d\mathbf{Z}\\ &= \sum_{\nu=0}^{\infty} \frac{(s-d)^\nu}{\nu!} \mu_\nu\{\mathbf{Z}\}\end{aligned} \tag{38}$$

in terms of the $\mu_\nu\{\mathbf{Z}\}$. The derivation of (38) reveals that

$$\begin{aligned}\mu_\nu\{\mathbf{Z}\} &= \mathrm{E}\{(\ln|\mathbf{Z}|)^\nu\}\\ &= \int_{\Omega_+} (\ln|\mathbf{Z}|)^\nu p_{\mathbf{Z}}(\mathbf{Z})\, d\mathbf{Z}\end{aligned} \tag{39}$$

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

$$\varphi_{\mathbf{Z}}(s) = \ln \phi_{\mathbf{Z}}(s)\,. \tag{40}$$

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

$$\kappa_\nu\{\mathbf{Z}\} = \left.\frac{d^\nu}{ds^\nu} \varphi_{\mathbf{Z}}(s)\right|_{s=d}. \tag{41}$$

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

$$\varphi_{\mathbf{Z}}(s) = \ln \phi_{\mathbf{Z}}(s) = \sum_{\nu=0}^{\infty} \frac{(s-d)^\nu}{\nu!} \kappa_\nu\{\mathbf{Z}\} \tag{42}$$

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 summarised in Figure 3. 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).

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

$$\begin{aligned}\phi_{\mathbf{W}}(s) &= \mathcal{M}\{p_{\mathbf{W}}(\mathbf{W})\}(s) \\ &= \frac{\Gamma_d(L+s-d)}{\Gamma_d(L)} |\mathbf{\Sigma}|^{s-d} .\end{aligned} \tag{43}$$

The MLCs are found to be

$$\kappa_1\{\mathbf{W}\} = \psi_d^{(0)}(L) + \ln|\mathbf{\Sigma}| \tag{44}$$

$$\kappa_\nu\{\mathbf{W}\} = \psi_d^{(\nu-1)}(L) , \quad \nu > 1 \tag{45}$$

where we introduce the νth-order *multivariate polygamma function* as

$$\psi_d^{(\nu)}(L) = \sum_{i=0}^{d-1} \psi^{(\nu)}(L-i) . \tag{46}$$

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

$$\psi^{(\nu)}(L) = \frac{d^{\nu+1} \ln\Gamma(L)}{dL^{\nu+1}} , \quad \nu \geq 0 . \tag{47}$$

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

$$\kappa_1\{\tilde{\mathbf{W}}\} = \psi_d^{(0)}(L) + \ln|\mathbf{\Sigma}| - d\ln L \tag{48}$$

$$\kappa_\nu\{\tilde{\mathbf{W}}\} = \psi_d^{(\nu-1)}(L) , \quad \nu > 1 . \tag{49}$$

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 utilising 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.

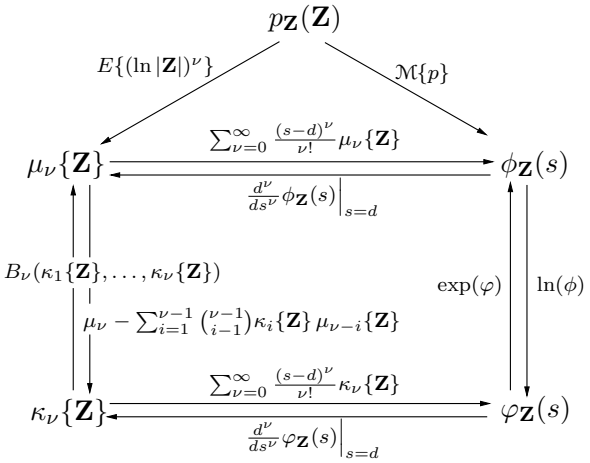

Fig. 3. Relations in matrix-variate Mellin kind statistics.

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,

$$X = U + V \tag{50}$$

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

$$\Phi_X(\omega) = \Phi_U(\omega) \cdot \Phi_V(\omega) \tag{51}$$

$$\Psi_X(\omega) = \Psi_U(\omega) + \Psi_V(\omega) \tag{52}$$

$$c_\nu\{X\} = c_\nu\{U\} + c_\nu\{V\} . \tag{53}$$

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

$$\begin{aligned}p_X(x) &= (p_U * p_V)(x) \\ &= \int_{-\infty}^{+\infty} p_U(u) p_V(x-u)\, du\end{aligned} \tag{54}$$

where $*$ denotes the convolution operator, which corresponds to a multiplication in the FT domain, as seen in (51).

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

$$X = U \cdot V, \tag{55}$$

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

$$\phi_X(s) = \phi_U(s) \cdot \phi_V(s) \tag{56}$$
$$\varphi_X(s) = \varphi_U(s) + \varphi_V(s) \tag{57}$$
$$\kappa_\nu\{X\} = \kappa_\nu\{U\} + \kappa_\nu\{V\}\,. \tag{58}$$

The usual interpretation [2] 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

$$\begin{aligned} p_X(x) &= \int_0^\infty p_{X|V}(x|v)\, p_V(v)\, dv \\ &= \int_0^\infty p_U\Big(\frac{x}{v}\Big)\, p_V(v)\, \frac{dv}{v} \end{aligned} \tag{59}$$

which is known as the Mellin convolution. The operation is denoted $p_X(x) = (p_U \mathbin{\hat{\star}} p_V)(x)$. The product in (56) 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

$$\begin{aligned} &(f \mathbin{\hat{\star}} g)(\mathbf{U}) \\ &\quad= \int_{\Omega_+} |\mathbf{V}|^{-d} f(\mathbf{V}^{-\frac{1}{2}}\mathbf{U}\mathbf{V}^{-\frac{1}{2}}) g(\mathbf{V})\, d\mathbf{V} \\ &\quad= \int_{\Omega_+} |\mathbf{V}|^{-d} g(\mathbf{V}^{-\frac{1}{2}}\mathbf{U}\mathbf{V}^{-\frac{1}{2}}) f(\mathbf{V})\, d\mathbf{V} \end{aligned} \tag{60}$$

which is an associative and commutative operation.

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

$$\begin{aligned} &\mathcal{M}\{(f \mathbin{\hat{\star}} g)(\mathbf{U})\}(s) \\ &\qquad= \mathcal{M}\{f(\mathbf{U})\}(s) \cdot \mathcal{M}\{g(\mathbf{U})\}(s)\,. \end{aligned} \tag{61}$$

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 $\mathbf{U} = \mathbf{V}^{-\frac{1}{2}}\mathbf{X}\mathbf{V}^{-\frac{1}{2}}$ and $d\mathbf{U} = d\mathbf{X}/|\mathbf{V}|^d$ [21]. This yields

$$\begin{aligned} &\mathcal{M}\{f(\mathbf{U})\}(s) \cdot \mathcal{M}\{g(\mathbf{V})\}(s) \\ &= \int_{\Omega_+} (|\mathbf{X}|/|\mathbf{V}|)^{s-d} f(\mathbf{V}^{-\frac{1}{2}}\mathbf{X}\mathbf{V}^{-\frac{1}{2}}) |\mathbf{V}|^{-d}\, d\mathbf{X} \\ &\qquad \times \int_{\Omega_+} |\mathbf{V}|^{s-d} g(\mathbf{V})\, d\mathbf{V} \\ &= \int_{\Omega_+} |\mathbf{X}|^{s-d} \Big[\int_{\Omega_+} |\mathbf{V}|^{-d} f(\mathbf{V}^{-\frac{1}{2}}\mathbf{X}\mathbf{V}^{-\frac{1}{2}}) g(\mathbf{V})\, d\mathbf{V} \Big] d\mathbf{X} \\ &= \mathcal{M}\{(f \mathbin{\hat{\star}} g)(\mathbf{X})\}(s) \end{aligned} \tag{62}$$

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 [2], [3].

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

$$\begin{aligned} &(f \mathbin{\hat{\otimes}} g)(\mathbf{U}) \\ &\qquad= \int_{\Omega_+} |\mathbf{V}|^{d} f(\mathbf{V}^{\frac{1}{2}}\mathbf{U}\mathbf{V}^{\frac{1}{2}}) g(\mathbf{V})\, d\mathbf{V}\,. \end{aligned} \tag{63}$$

This operation is neither associative nor commutative.

Theorem 2 (Type I matrix-variate Mellin correlation): Under the assumptions presented in Definition 6, then

$$\begin{aligned} &\mathcal{M}\{(f \mathbin{\hat{\otimes}} g)(\mathbf{U})\}(s) \\ &\qquad= \mathcal{M}\{f(\mathbf{U})\}(s) \cdot \mathcal{M}\{g(U)\}(2d - s)\,. \end{aligned} \tag{64}$$

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

$$\begin{aligned}&(f \mathbin{\hat{\circledast}} g)(\mathbf{U}) \\ &\quad = \int_{\Omega_+} |\mathbf{V}|^{-d} f(\mathbf{V}^{\frac{1}{2}}\mathbf{U}\mathbf{V}^{\frac{1}{2}}) g(\mathbf{V})\, d\mathbf{V}\,.\end{aligned} \tag{65}$$

This is neither an associative nor commutative operation.

Theorem 3 (Type II matrix-variate Mellin correlation): Under the assumptions presented in Definition 6, then

$$\begin{aligned}&\mathcal{M}\{(f \mathbin{\hat{\circledast}} g)(\mathbf{U})\}(s) \\ &\quad = \mathcal{M}\{f(\mathbf{U})\}(s) \cdot \mathcal{M}\{g(-\mathbf{U})\}(s)\,.\end{aligned} \tag{66}$$

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 [21, Th. 6.2].

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

$$\mathbf{X} = \mathbf{U}\,\mathbf{V} \tag{67}$$

where $\mathbf{X}$, $\mathbf{U}$, and $\mathbf{V}$ are complex and positive definite Hermitian matrices. With the same approach as in (59) we establish that the PDF of $\mathbf{X}$ is

$$\begin{aligned}&p_{\mathbf{X}}(\mathbf{X}) \\ &\quad = \int_{\Omega_+} p_{\mathbf{X}|\mathbf{V}}(\mathbf{X}|\mathbf{V}) p_{\mathbf{V}}(\mathbf{V})\, d\mathbf{V} \\ &\quad = \int_{\Omega_+} |\mathbf{V}|^{-d} p_{\mathbf{U}}(\mathbf{V}^{-\frac{1}{2}}\mathbf{X}\mathbf{V}^{-\frac{1}{2}}) p_{\mathbf{V}}(\mathbf{V})\, d\mathbf{V}\,.\end{aligned} \tag{68}$$

This is exactly $(p_{\mathbf{U}} \mathbin{\hat{\star}} p_{\mathbf{V}})(\mathbf{X})$, as should be expected from (62), 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 Definition 3 and 5. It follows from Theorem 1 that

$$\phi_{\mathbf{X}}(s) = \phi_{\mathbf{U}}(s) \cdot \phi_{\mathbf{V}}(s) \tag{69}$$

$$\varphi_{\mathbf{X}}(s) = \varphi_{\mathbf{U}}(s) + \varphi_{\mathbf{V}}(s) \tag{70}$$

$$\kappa_\nu\{\mathbf{X}\} = \kappa_\nu\{\mathbf{U}\} + \kappa_\nu\{\mathbf{V}\}\,. \tag{71}$$

in the matrix-variate case.

Example (Multilook polarimetric product model): We now return to the multilook polarimetric product model for the radar data covariance matrix $\mathbf{Z} = L\mathbf{C}$. The model can be written as: $\mathbf{Z} = \mathbf{T}\mathbf{W}$, 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 entry, namely T. Without entering the stringent argument in terms of differential calculus, we state that an integral $\int_{\Omega_+} f(|\mathbf{T}|)\, d\mathbf{T}$ can be replaced with $\int_{\mathbb{R}^+} f(T^d)\, dT$. We thus have

$$\begin{aligned}\phi_{\mathbf{T}}(s) &= \int_{\Omega_+} |\mathbf{T}|^{s-d} p_{\mathbf{T}}(\mathbf{T})\, d\mathbf{T} \\ &= \int_0^\infty t^{d(s-d)} p_T(t)\, dt \\ &= \sum_{\nu=0}^{\infty} \frac{[d(s-d)]^\nu}{\nu!} \mu_\nu\{T\} \\ &= \phi_T(d(s-d)+1)\,.\end{aligned} \tag{72}$$

This implies that

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

Faá di Bruno's formula is 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 $\mathbf{Z}$ as

$$\begin{aligned}\kappa_\nu\{\mathbf{Z}\} &= \mu_\nu\{\mathbf{Z}\} \\ &\quad - \sum_{i=1}^{\nu-1} \binom{\nu-1}{i-1} \kappa_i\{\mathbf{Z}\}\mu_{\nu-i}\{\mathbf{Z}\}\,.\end{aligned} \tag{75}$$

The first MLCs are expressed as

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

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

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

We use $\mathbf{Z} = \mathbf{T}\mathbf{W}$ and (71) to prove that

$$\kappa_\nu\{\mathbf{Z}\} = \kappa_\nu\{\mathbf{W}\} + d^\nu \kappa_\nu\{T\} \tag{79}$$

but, since the observable of multilook polarimetric radar is $\mathbf{C} = \mathbf{Z}/L = T\tilde{\mathbf{W}}$, we are more interested in the $\kappa_\nu\{\mathbf{C}\}$. With the $\kappa_\nu\{\mathbf{W}\}$ expanded, the MLCs of $\mathbf{C}$ evaluate under the product model to

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

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

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

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

given a sample of N covariance matrices: $\mathscr{C} = \{\mathbf{C}_i\}_{i=1}^N$. 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 visualisation tool, which efficiently explains some uses of MKS and provides intuition about the MoMLC. It is a straight-forward extension of the log-cumulant diagrams used by Nicolas [2], [3] 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 $\nu_1, \cdots, \nu_q$ 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

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

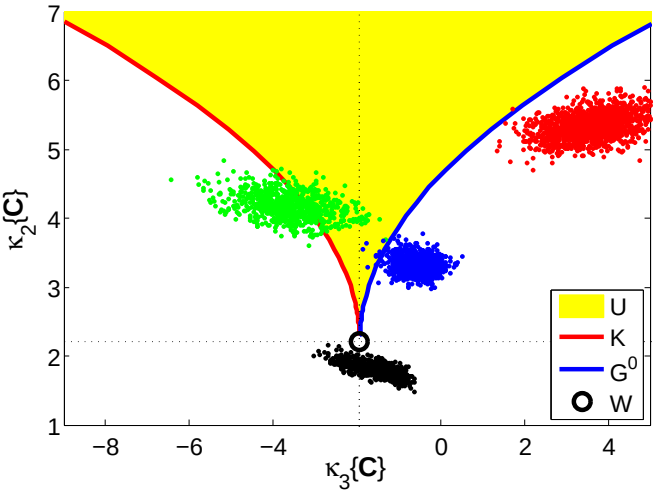

Fig. 4. Matrix log-cumulant diagram showing the manifolds of theoretical MLCs for the complex Wishart, $\mathcal{K}$, $\mathcal{G}^0$ and $\mathcal{U}$ distribution, as well as a collection of sample MLCs representing forest (green), ocean (blue), urban area (red) and a wheat crop (black).

where we have changed the notation of the theoretical νth-order MLC from $\kappa_\nu\{\mathbf{C}\}$ to $\kappa_\nu(\boldsymbol{\vartheta})$ to emphasise 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

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

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

Like Nicolas [2], [3], 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 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 Figure 4, 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 parametrised by one parameter, thus they are represented by a curve (red and blue, respectively), which

is a one-dimensional manifold. The $\mathcal{U}$ distribution has two texture parameters and is represented by a surface (yellow area), which is a two-dimensional 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 centred at the point which represents the Wishart distribution. The impact of nonzero texture on the MLCs is measured relative to these axes.

Given a sample $\mathscr{C} = \{\mathbf{C}_i\}_{i=1}^N$ of covariance matrices, we can compute the sample MLCs of order $\nu_1, \cdots, \nu_q$ and plot them as a point in MLC space. This has been done in Figure 4 for one forest sample (shown as green points) and one wheat crop sample (black) 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 (red) 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 $\mathscr{C}$.

MoMLC parameter estimation can now be visualised 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 is assigned as an estimate. An estimator based on a single νth-order MLC relies on a projection in the direction normal to the νth-order coordinate. This is illustrated in Figure 5, displaying estimators for the $\mathcal{K}$ distribution texture parameter α based on the point $(\langle\kappa_3(\mathscr{C})\rangle, \langle\kappa_2(\mathscr{C})\rangle)$ 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-order and third-order MLC equation, respectively. The dashed arrows visualise 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 $\mathscr{C}$ is utilised more

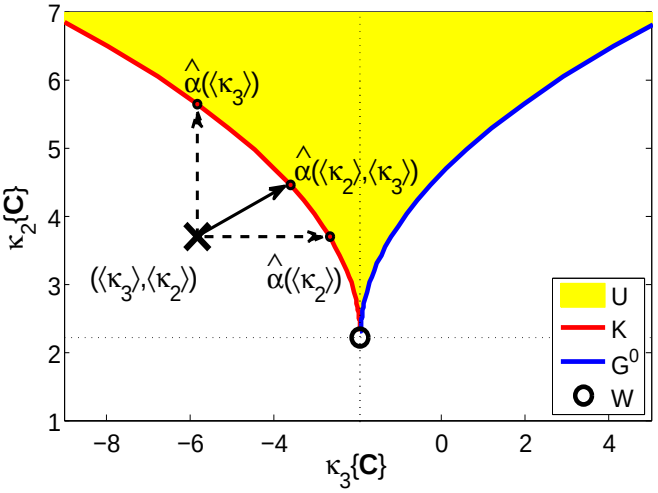

Fig. 5. MLC space interpretation of three estimators of the $\mathcal{K}$ distribution texture parameter α. The first estimator is based on $\langle\kappa_2\{\mathbf{C}\}\rangle$, the other is based on $\langle\kappa_3\{\mathbf{C}\}\rangle$, and the third is based on both.

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 Figure 5. The sample MLC point is projected onto the point on the $\mathcal{K}$ distribution curve that minimises 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 Figure 4 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 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 $\mathbf{\Sigma}$, defined as the sample mean of $\mathscr{C}$.

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 $\tau = \mathrm{tr}\left(\hat{\mathbf{\Sigma}}^{-1}\mathbf{C}\right)$. The derivation is shown in Appendix B. Another approach taken by Freitas and Frery et al. [25], [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

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

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 [2], [3]. 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

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

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

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

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

Figure 6 and 7 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

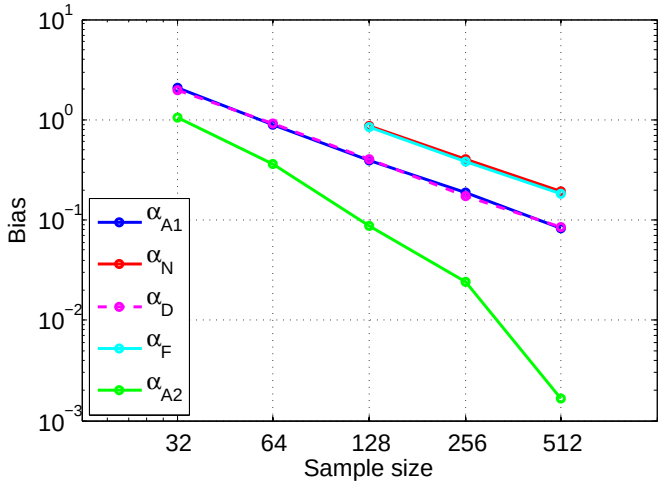

Fig. 6. Bias of estimators for the $\mathcal{K}$ distribution texture parameter α as function of sample size N.

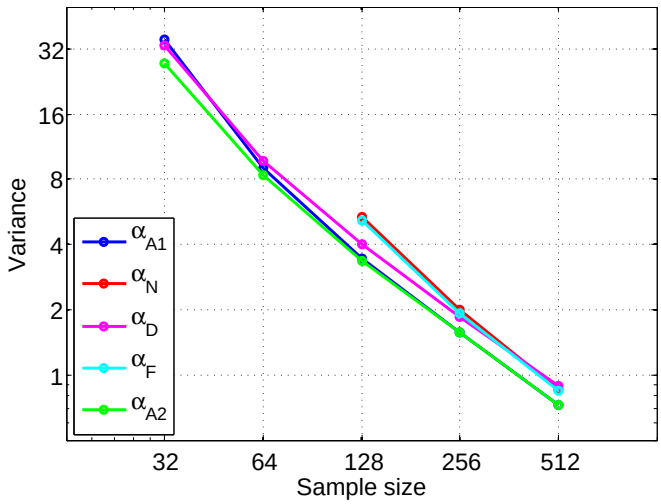

Fig. 7. Variance of estimators for the $\mathcal{K}$ distribution texture parameter α as function of sample size N.

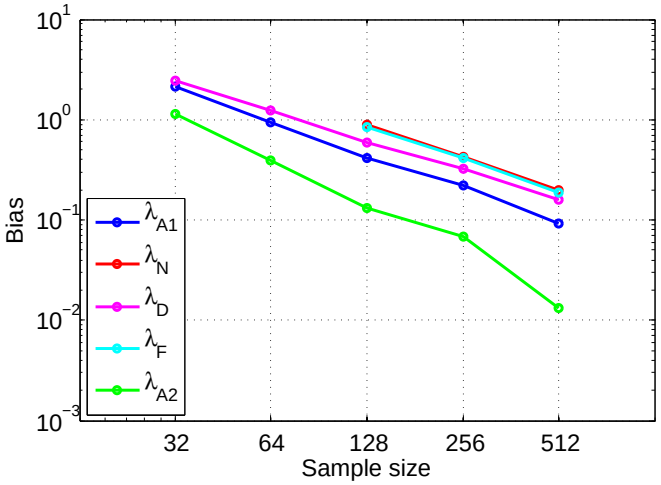

Fig. 8. Bias of estimators for the $\mathcal{G}^0$ distribution texture parameter λ.

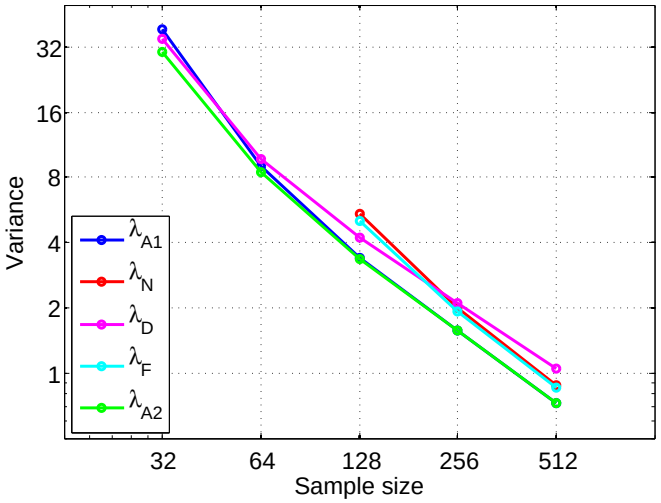

Fig. 9. Variance of estimators for the $\mathcal{G}^0$ distribution texture parameter λ.

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.

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

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

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

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

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 parametrised by $L = 10$ and $\lambda = 10$. The bias and variance results in Figure 8 and 9 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.

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

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

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

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

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

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

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. Conclusions

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

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

Accordingly, the Mellin kind CGF is

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

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 utilised the well-known relation

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

and the multivariate polygamma function introduced in (46). 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 $\varphi_{\mathbf{W}}^{(\nu)}(s) = \frac{d^\nu}{ds^\nu}\varphi_{\mathbf{W}}(s)$. 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

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

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

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,

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

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

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

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

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

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 (48) and (49).

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

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

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, $\{\tau_i = \mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{C}_i)\}_{i=1}^N$.

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.

ACKNOWLEDGEMENT

The authors wish to thank Anthony Paul Doulgeris at the University of Tromsø and Gabriele Moser at the University of Genoa for their valuable contribution through discussions of the topic and comments on the manuscript.

REFERENCES

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

[2] 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, in French.

[3] ——, "Application de la transformée de Mellin: Étude des lois statistiques de l'imagerie cohérente," Ecole Nationale Supérieure des Télécommunications, Paris, France, Tech. Rep. 2006D010, 2006, in French.

[4] R. D. Pierce, "Application of the positive alpha-stable distribution," in *Proc. IEEE Signal Proc. Workshop on Higher-Order Stat., SPW-HOS '97*, Banff, Canada, Jul. 1997, pp. 420–424.

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

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

[7] 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.

[8] ——, "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.

[9] 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.

[10] J.-M. Nicolas, "A Fisher-MAP filter for SAR image processing," in *Proc. IEEE Int. Geosc. Remote Sens. Symp., IGARSS'03*, vol. 3, Toulouse, France, Jul. 2003, pp. 1996–1998.

[11] A. Achim, E. E. Kuruoğlu, 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.

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

[13] 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.

[14] 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 Int. Geosci. Remote Sens. Symp., IGARSS'07*, vol. 1, Barcelona, Spain, Jul. 2007, pp. 3891–3894.

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

[16] F. Bujor, E. Trouvé, 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.

[17] 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.

[18] ——, "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.

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

[20] C. Valade and J.-M. Nicolas, "Homomorphic wavelet transform and new subband statistics models for SAR image compression," in *Proc. IEEE Int. Geosc. Remote Sens. Symp., IGARSS'04*, vol. 1, Anchorage, USA, Sep. 2004, pp. 285–288.

[21] A. M. Mathai, *Jacobians of Matrix Transformations and Functions of Matrix Arguments*. New York, USA: World Scientific, 1997, ch. 6, Def. 6.3.

[22] J.-S. Lee and E. Pottier, *Polarimetric Radar Imaging: From Basics to Applications*, ser. Optical Science and Engineering. Boca Raton, USA: CRC Press, 2009, no. 143.

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

[24] 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 Int. Geosc. Remote Sens. Symp., IGARSS'94*, vol. 4, Pasadena, USA, Aug. 1994, pp. 2179–2181.

[25] 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.

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

[27] 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.

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

[29] W. P. Johnson, "The curious history of Faà di Bruno's formula," *Am. Math. Monthly*, vol. 109, no. 3, pp. 217–234, Mar. 2002.

[30] L. Comtet, *Advanced Combinatorics: The Art of Finite and Infinite Expansions*. Dordrecht, The Netherlands: Reidel Publishing Company, 1974, ch. 3.3, pp. 133–136.

[31] H. H. Andersen, M. Højbjerre, D. Sørensen, and P. S. Eriksen, *Linear and Graphical Models for the Multivariate Complex Normal Distribution*, ser. Lecture Notes in Statistics. New York, USA: Springer, 1995.

[32] C.-Y. Chi, C.-C. Feng, C.-H. Chen, and C.-Y. Chen, *Blind Equalization and System Identification*. London: Springer, 2006, ch. 3, pp. 103–104.

[33] T. Kollo and D. von Rosen, *Advanced Multivariate Statistics with Matrices*. Dordrecht, The Netherlands: Springer, 2005.

[34] J. Bertrand, P. Bertrand, and J.-P. Ovarlez, "The Mellin transform," in *The Transform and Applications Handbook*, 2nd ed., A. D. Poularikas, Ed. Boca Raton, US: CRC Press, 2000, ch. 11.

[35] A. M. Mathai, "Some results on functions of matrix argument," *Math. Nachr.*, vol. 84, no. 1, pp. 171–177, 1978.
[36] ——, "Distribution of the canonical correlation matrix," *Ann. Inst. Statist. Math.*, vol. 33, part A, pp. 35–43, 1981.
[37] B. Mathur and S. Krishna, "On multivariate fractional integration operators," *Indian J. Pure Appl. Math.*, vol. 8, pp. 1078–1082, 1977.
[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 on Science and Applications of SAR Polarimetry and Polarimetric Interferometry (POLinSAR '09)*, vol. ESA SP-668, Frascati, Italy, 8 pp., Jan. 2009.
[40] S. N. Anfinsen, "Statistical analysis of multilook polarimetric radar images with the Mellin transform," Ph.D. dissertation, University of Tromsø, Tromsø, Norway, in prep, 2010.
[41] D. Casasent and D. Psaltis, "Deformation invariant, space-invariant optical pattern recognition," in *Progress in Optics*, E. Wolf, Ed. New York, USA: North-Holland, 1978, vol. XVI, ch. V.
[42] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*. Boca Raton, USA: 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 Trans. Geosci. Remote Sens.*, 2010, submitted, Available: http://www.phys.uit.no/~stiann/sna-gof-submitted.pdf.
[44] S. N. Anfinsen, A. P. Doulgeris, and T. Eltoft, "Estimation of the equivalent number of looks in polarimetric SAR imagery," in *Proc. IEEE Int. Geosci. Remote Sens. Symp., IGARSS'08*, vol. 4, Boston, USA, Jul. 2008, 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 '91*, vol. 2. Honolulu, USA: IEEE, Oct. 1991, pp. 1217–1222.
[48] D. Blacknell, "Comparison of parameter estimator for K-distribution," *IEE Proc. Radar, Sonar, Navig.*, vol. 141, no. 1, pp. 45–52, Feb. 1994.
[49] O. Harant, L. Bombrun, M. Gay, R. Fallourd, E. Trouvé, and F. Tupin, "Segmentation and classification of polarimetric SAR data based on the KummerU distribution," in *Proc. 4th Int. Workshop on Science and Applications of SAR Polarimetry and Polarimetric Interferometry (POLinSAR '09)*, vol. ESA SP-668, Frascati, Italy, 6 pp., Apr. 2009.