# Polarimetric Detection in Compound Gaussian Clutter With Kronecker Structured Covariance Matrix

Yikai Wang, *Student Member, IEEE*, Wei Xia, Zishu He, Hongbin Li, *Senior Member, IEEE*, and Athina P. Petropulu, *Fellow, IEEE*

***Abstract*—In this paper, we consider polarimetric adaptive detection in compound Gaussian clutter whose covariance matrix (CM) has a Kronecker structure. We derive the Cramér–Rao bound of a Kronecker structured CM estimate and analyze the constant false alarm rate property of the adaptive subspace matched filter detector that uses Kronecker structured estimates. We provide a general expression for the average signal-to-clutter ratio loss (SCRL) as a function of the mean square error of the covariance estimate. The aforementioned expression is helpful in determining how many samples are required in order to achieve a desired average SCRL level in practical scenarios. Based on that expression, we show that the required sample size can be effectively reduced by exploiting the Kronecker structure of the clutter CM. We also derive the asymptotic detection performance of the adaptive subspace matched filter. The analysis of SCRL and detection performance can be extended to more general scenarios, especially when the maximum-likelihood estimate of the structured CM involves solving fixed point equations. Numerical simulations validate the merits of the proposed methods.**

***Index Terms*—Polarimetric detection, compound Gaussian clutter, Kronecker structure, constant false alarm rate, Cramér–Rao bound, SCR loss, asymptotic performance.**

## I. Introduction

IT IS well known that polarization diversity [1] is a useful tool for improving the performance of radar detection [2], [4], [9], [10], estimation [3], and tracking [5]. Unlike conventional radar which employs only one polarization in both transmitter and receiver, polarimetric radar can operate in different polarization channels, and thus can acquire complete polarimetric information of the target and the environment [6]. In radar adaptive detection, the unknown clutter covariance matrix (CM) is estimated based on sufficient amount of the training data [7]. For polarimetric radar, the received signals are required to be processed in the polarization-space-time joint domain [12], [13], [39], and thus CM estimation in that context requires more training samples as compared to conventional radar. However, the structure of the CM in this scenario can be exploited to reduce the amount of training data needed [17], [24], [26], [52].

In polarimetric array radar, the CM is expressed as the Kronecker product of the polarization covariance matrix $\mathbf{R}_p$ and the spatial covariance matrix $\mathbf{R}_s$ [9], [11]–[13], where $\mathbf{R}_p$ is determined by the clutter terrain, and $\mathbf{R}_s$ is related to the array configuration (see [14] and also Sec. II). Exploiting such structure can significantly reduce the number of unknown parameters, and can enable high estimation accuracy with fewer training data. Maximum likelihood (ML) estimation of a Kronecker structured CM in the presence of Gaussian clutter is discussed in [19], [20], while ML estimation in the presence of compound Gaussian clutter is discussed in [21]. Further, two fast non-iterative asymptotically efficient CM estimates are developed in [20], and CM estimation for rank deficient submatrices $\mathbf{R}_p$ and $\mathbf{R}_s$ is discussed in [22].

The performance of adaptive detection depends on the signal-to-clutter ratio loss (SCRL) of the minimum variance distortioness response (MVDR) filter and the training size (convergence rate) [7], [11], [23]. The SCRL is defined as the ratio between the signal-to-clutter ratio (SCR) of the MVDR filter using the true CM and that using a CM estimate, and quantifies how close the performance of an adaptive detector using a CM estimate is to that using the true CM. It is well known that, to achieve a 3 dB average SCRL, the sample covariance matrix (SCM) [7] requires approximately $2N$ training samples, where $N$ is the signal dimension. However, the required training size can be significantly reduced by exploiting structural information on the clutter, which may be available due to array geometry [24], [26], [52], or known clutter properties [27]. For example, for the persymmetric structured CM, 3 dB average SCRL can be achieved with $N$ training samples [24]. The average SCRL for the Hermitian (SCM) and persymmetric structured CM can be computed based on the probability distribution of the SCRL. For determining the training size required, one needs to compute the statistics of SCRL as function of the CM estimate. However, for the compound Gaussian case,

Manuscript received January 5, 2017; revised May 10, 2017; accepted June 6, 2017. Date of publication June 16, 2017; date of current version June 28, 2017. The associate editor coordinating the review of this manuscript and approving it for publication was Kostas D. Berberidis. This work was supported in part by China Scholarship Council and in part by the Natural Science Foundation of China under Grants 61671137, 61671139, and 61401062. *(Corresponding author: Yikai Wang.)*

Y. Wang was with the Rutgers, The State University of New Jersey, New Brunswick, NJ 08854 USA. He is now with the School of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: nbwyk.123@163.com).

W. Xia and Z. He are with the School of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: wx@uestc.edu.cn; zshe@uestc.edu.cn).

H. Li is with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030 USA (e-mail: Hongbin.Li@stevens.edu).

A. P. Petropulu is with the Department of Electrical and Computer Engineering, Rutgers, The State University of New Jersey, New Brunswick, NJ 08854 USA (e-mail: athinap@rci.rutgers.edu).

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

Digital Object Identifier 10.1109/TSP.2017.2716912

or for other structured covariance matrices, such as Toeplitz and Kronecker structures, the CM estimates are obtained by solving fixed point equations [21]. Thus, it is rather difficult, if not impossible, to obtain accurate statistical properties for the SCRL.

In this paper, we consider a Kronecker structured CM, which arises in the context of polarimetric array radar. We derive a compact form of the Cramér-Rao bound (CRB) of a CM estimate with Kronecker structure in compound Gaussian scenarios, and provide a general expression of the average SCRL as function of the mean-square error (MSE) of the CM estimate. Based on the invariance principle [28], we establish the constant false alarm rate (CFAR) property and the approximate average SCRL of the adaptive detector which uses a Kronecker structured ML estimate. This expression can be extended to analyze the performance of adaptive detectors using other structured CMs, such as the ML estimate of the CM, obtained through solving fixed point equations. We also provide the average SCRL for a CM with group symmetric structures [30], commonly employed in radar CFAR detection [31], such as persymmetric, circulant matrices [52]. Finally, we derive asymptotic expressions for the probability of false alarm and probability of detection. The asymptotic results are found to be accurate in the small sample regime, which is attributed to the superior convergence rate of the estimates exploiting the Kronecker structure in the CM.

Our main contributions are the following:

1) We provide a general expression of the average SCRL as function of the MSE of the CM estimate (see Theorem 1). This is helpful in determining how many secondary samples are required for achieving a desired level of SCRL, especially when the CM estimate is obtained as the solution of fixed point equations. Examples of popular structured CMs in radar detection are provided (see Table I). This result can be regarded as the generalization of the well known RMB rule [7] from the unstructured case to the structured case.
2) We show the merits of exploiting the Kronecker structure in the CM for both radar detection and estimation. We provide a compact expression of the CRB for Kronecker structured CM estimation in compound Gaussian clutter (see Proposition 2). We also provide in closed form the required training size in order to achieve a desired average SCRL (see Corollary 1).
3) Based on the invariance principle, we analyze the CFAR property of detectors using Kronecker structured CM estimates (see Proposition 3), and determine the asymptotic detection performance of the adaptive subspace matched filter that uses a Kronecker ML estimate (see Proposition 4).

The rest of this paper is organized as follows. We describe the signal model of the polarimetric detection problem in Sec. II. We briefly introduce some background on the adaptive subspace matched detector in Sec. III. We present our results on Kronecker structured CM estimation and detection performance analysis in Sec. IV and Sec. V respectively. Finally, we provide numerical simulations to validate our theoretical results in Sec. VI, and draw conclusions in Sec. VII.

*Notation:* Throughout this paper, $\mathbb{R}$ and $\mathbb{C}$ denote real and complex domains, respectively. Bold-faced upper and lower letters denote matrices and vectors respectively, while light-faced lower letters denote scalar quantities. $N$-dimensional vector $\mathbf{e}_i \in \mathbb{R}^{N\times 1}$ denotes the $i$th column of the $N$-dimensional identity matrix $\mathbf{I}_N$. Superscripts $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^*$, $\{\cdot\}^\dagger$ denote the matrix transpose, Hermitian transpose, conjugate and pseudo-inverse operators, respectively. $\mathbf{A}^{-T}$ denotes $(\mathbf{A}^{-1})^T$. $[\mathbf{Q}]_{ij}$ denotes the $(i,j)$th entry of the matrix $\mathbf{Q}$, and $\det\{\cdot\}$ and $\mathrm{tr}\{\cdot\}$ denote the matrix determinant and trace operators, respectively. $\mathbf{A} \succeq \mathbf{B}$ denotes that $\mathbf{A}-\mathbf{B}$ is positive semi-definite. $\Re\{\cdot\}$ and $\Im\{\cdot\}$ denote the real and imaginary parts of a complex number or matrix, respectively. $\mathrm{vec}\{\cdot\}$ denotes the vectorization operator, which stacks the columns of a matrix into a vector, $\mathrm{vech}\{\cdot\}$ and $\underline{\mathrm{vech}}\{\cdot\}$ stack the elements below the main diagonal columnwise, with "vech" including the main diagonal, and "$\underline{\mathrm{vech}}$" not including the main diagonal [37]. "$\otimes$" and "$\|\cdot\|$" denote the Kronecker product and the Frobenius norm, respectively. "$\propto$" means proportional to. $\mathbb{E}\{\cdot\}$ and $\mathrm{cov}\{\cdot\}$ denote mathematical expectation and covariance, respectively.

## II. Signal Model and Problem Formulation

We now consider a polarimetric array with $N_s$ electrical-magnetic vector-sensors, with each sensor measuring the electromagnetic wave in $N_p$ polarization channels [2], [12], [13]. Depending on the polarimetric radar scheme, $N_p$ may take the values 1, 2, 3, 4 [6]. The received signals in the cell under test are down-converted to baseband or to an intermediate frequency in all the polarization channels at each sensor, they are processed by the corresponding matched filters, and they are sampled and stacked into a $N$-dimensional vector $\mathbf{y} \in \mathbb{C}^{N\times 1}$ where $N = N_p N_s$. Let $\mathbf{y}_k \in \mathbb{C}^{N\times 1}, k = 1, \cdots, K$ be $K$ independent identical distributed (i.i.d) signal-free training data (secondary data), obtained from adjacent range cells. The binary hypothesis testing problem can be formulated as

$$\begin{cases} \mathcal{H}_0: & \mathbf{y} = \mathbf{n} \\ & \mathbf{y}_k = \mathbf{n}_k, \quad k = 1, 2, \cdots, K \\ \mathcal{H}_1: & \mathbf{y} = \mathbf{A}\boldsymbol{\alpha} + \mathbf{n} \\ & \mathbf{y}_k = \mathbf{n}_k, \quad k = 1, 2, \cdots, K, \end{cases} \tag{1}$$

where $\mathbf{A} = \mathbf{I}_{N_p} \otimes \mathbf{a} \in \mathbb{C}^{N\times N_p}$ is the system response of the polarimetric array, $\mathbf{a} \in \mathbb{C}^{N_s\times 1}$ is the spatial steering vector, $\boldsymbol{\alpha} \in \mathbb{C}^{N_p\times 1}$ denotes the unknown deterministic polarimetric scattering vector of the target, and $\mathbf{n}$ and $\mathbf{n}_k$ denote the corresponding clutter returns. Based on this, we have that $\mathbf{A}\boldsymbol{\alpha} = \boldsymbol{\alpha} \otimes \mathbf{a}$. The clutter is assumed to be compound Gaussian distributed, and can be expressed in terms of a real positive texture component $\tau$ and a speckle component $\mathbf{c}$ as

$$\mathbf{n} = \sqrt{\tau}\mathbf{c}. \tag{2}$$

The texture can be deterministic or random. Here, we assume that $\tau$ obeys the inverse Gamma (IG) distribution with shape parameter $\upsilon$ [38]. However, the main result corresponding to other textures can be obtained similarly. The probability density

function (PDF) of the IG distribution is

$$p_\tau(\tau) \triangleq p_\tau(\tau; v) = \frac{v^v}{\Gamma(v)\tau^{v+1}} \exp\left(-\frac{v}{\tau}\right), \quad \tau > 0, \tag{3}$$

where $\Gamma(\cdot)$ is the Gamma function. In this paper, we assume that the shape parameter is known. When the shape parameter is unknown, it can be estimated along the lines of [32]. The speckle component is assumed to be circular complex Gaussian distributed with zero mean and covariance matrix $\mathbf{R}$, i.e., $\mathbf{c} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R})$. The $K$ i.i.d training samples have a similar texture $\tau_k$ and speckle $\mathbf{c}_k$ representation, i.e., $\mathbf{n}_k = \sqrt{\tau_k}\mathbf{c}_k$.

In the polarimetric array context, the clutter signal can be regarded as a sum of $N_c$ clutter patches in the same range cell (similar to the well known space-time adaptive processing model [40]), i.e.,

$$\mathbf{c} = \sum_{i=1}^{N_c} \boldsymbol{\alpha}_i^c \otimes \mathbf{a}_i^c \in \mathbb{C}^{N\times 1}, \tag{4}$$

In the above equation, $\mathbf{a}_i^c \in \mathbb{C}^{N_s\times 1}$ and $\boldsymbol{\alpha}_i^c \in \mathbb{C}^{N_p\times 1}$ are the spatial steering vector and the polarization scattering vector of the $i$th clutter patch, respectively. $\boldsymbol{\alpha}_i^c$ is complex Gaussian distributed with zero mean, and for monostatic full polarization radar with $N_p = 3$, it has the covariance matrix equal to [14]

$$\mathbb{E}\left\{\boldsymbol{\alpha}_i^c(\boldsymbol{\alpha}_i^c)^H\right\} = \sigma^2(i)\begin{bmatrix} 1 & 0 & \epsilon\sqrt{\gamma} \\ 0 & \delta & 0 \\ \epsilon^*\sqrt{\gamma} & 0 & \gamma \end{bmatrix}, \tag{5}$$

where $\sigma^2(i)$ is the clutter power of the $i$th clutter patch, and is determined by the transmitter's power, the beamforming process and the clutter scattering; $\delta$, $\epsilon$ and $\gamma$ represent correlation parameters normalized by $\sigma^2(i)$, and are determined by the terrain of the illuminated area. Typical values of $\delta$, $\epsilon$ and $\gamma$ of various types of terrain clutter may be found in [14]. Here we assume that all the clutter patches are associated with the same terrain and i.i.d, and thus the polarization and spatial covariance matrices are respectively defined as

$$\mathbf{R}_p \triangleq \begin{bmatrix} 1 & 0 & \epsilon\sqrt{\gamma} \\ 0 & \delta & 0 \\ \epsilon^*\sqrt{\gamma} & 0 & \gamma \end{bmatrix}, \tag{6}$$

and

$$\mathbf{R}_s \triangleq \sum_{i=1}^{N_c} \sigma^2(i)\mathbf{a}_i^c(\mathbf{a}_i^c)^H. \tag{7}$$

Based on the above, the clutter CM has a Kronecker structure [9], [11]–[13], i.e.,

$$\mathbf{R} = \mathbb{E}\{\mathbf{c}\mathbf{c}^H\} = \mathbf{R}_p \otimes \mathbf{R}_s. \tag{8}$$

For the sake of simplicity, we assume that both $\mathbf{R}_p$ and $\mathbf{R}_s$ are invertible matrices. We also consider the noise-free signal model, which corresponds to high clutter-to-noise (CNR) scenarios. In a practical scenario, the Kronecker structure is valid within some error. The effect of structure model error is discussed in Sec. VI.

In addition to polarimetric radar, examples of scenarios in which such structured CM arises include channel modeling for multiple-input multiple-output (MIMO) communications [15], [16], MIMO radar [17], and modeling of spatio-temporal noise of MEG/EEG data [18]. Therefore, the results that will be presented in the following can be applied in those applications as well.

## III. Preliminary of Adaptive Detection

In this section, we employ a two-step adaptive subspace matched filter (ASMF) detector [8], [48]. First, we derive the generalized likelihood ratio test (GLRT) detector [23] by assuming that the CM is known, and then obtain a suitable estimate of the CM based on secondary data; this estimate is substituted into the derived detector in place of the exact CM.

With the assumption that the CM is known *a priori*, the detector is given by the GLRT method, i.e.,

$$\frac{\max_{\boldsymbol{\alpha}} \int p(\mathbf{y}\,|\mathcal{H}_1, \boldsymbol{\alpha}\,, \tau)p_\tau(\tau)d\tau}{\int p(\mathbf{y}\,|\mathcal{H}_0\,, \tau)p_\tau(\tau)d\tau} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \xi, \tag{9}$$

where $\xi$ is a threshold, and for the conditional PDF of $\mathbf{y}$ under $\mathcal{H}_q, q = 0, 1$, it holds that

$$p(\mathbf{y}|\mathcal{H}_q, \boldsymbol{\alpha}, \tau) \propto \frac{1}{\det(\tau\mathbf{R})} \cdot \exp\left\{-\frac{(\mathbf{y} - q\mathbf{A}\boldsymbol{\alpha})^H\mathbf{R}^{-1}(\mathbf{y} - q\mathbf{A}\boldsymbol{\alpha})}{\tau}\right\}. \tag{10}$$

Using the Gamma integral, i.e.,

$$\int \frac{1}{\tau^N} \exp\left\{-\frac{A}{\tau}\right\} d\tau = \frac{\Gamma(N-1)}{A^{N-1}}, \tag{11}$$

and based on (3) and (10), (9) can be rewritten as

$$\frac{\left[v + (\mathbf{y} - \mathbf{A}\hat{\boldsymbol{\alpha}})^H\mathbf{R}^{-1}(\mathbf{y} - \mathbf{A}\hat{\boldsymbol{\alpha}})\right]^{-v-N}}{\left[v + \mathbf{y}^H\mathbf{R}^{-1}\mathbf{y}\right]^{-v-N}} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta \tag{12}$$

where $\hat{\boldsymbol{\alpha}} = (\mathbf{A}^H\mathbf{R}^{-1}\mathbf{A})^{-1}\mathbf{A}^H\mathbf{R}^{-1}\mathbf{y}$ is the ML estimate of $\boldsymbol{\alpha}$ [48]. The detector of (12) can be rewritten as

$$T(\mathbf{y}, \mathbf{R}, \mathbf{A}, v) = \frac{\mathbf{y}^H\mathbf{R}^{-1}\mathbf{A}(\mathbf{A}^H\mathbf{R}^{-1}\mathbf{A})^{-1}\mathbf{A}^H\mathbf{R}^{-1}\mathbf{y}}{\mathbf{y}^H\mathbf{R}^{-1}\mathbf{y} + v} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta_1. \tag{13}$$

On replacing the covariance matrix $\mathbf{R}$ in (13) by an estimate $\hat{\mathbf{R}}$, obtained from the secondary data, we get the ASMF detector

$$T(\mathbf{y}, \hat{\mathbf{R}}, \mathbf{A}, v) = \frac{\mathbf{y}^H\hat{\mathbf{R}}^{-1}\mathbf{A}(\mathbf{A}^H\hat{\mathbf{R}}^{-1}\mathbf{A})^{-1}\mathbf{A}^H\hat{\mathbf{R}}^{-1}\mathbf{y}}{\mathbf{y}^H\hat{\mathbf{R}}^{-1}\mathbf{y} + v} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \xi. \tag{14}$$

Two popular CM estimates are the maximum likelihood estimate (MLE) and the normalized sample covariance matrix (NSCM)

estimate, given respectively as [36]

$$\hat{\mathbf{R}}_{\text{MLE}} = \frac{v+N}{K}\sum_{k=1}^{K}\frac{\mathbf{y}_k\mathbf{y}_k^H}{v+\mathbf{y}_k^H\hat{\mathbf{R}}_{\text{MLE}}^{-1}\mathbf{y}_k}, \tag{15}$$

$$\hat{\mathbf{R}}_{\text{NSCM}} = \frac{N}{K}\sum_{k=1}^{K}\frac{\mathbf{y}_k\mathbf{y}_k^H}{\mathbf{y}_k^H\mathbf{y}_k}. \tag{16}$$

MLE may be obtained recursively; convergence and uniqueness of MLE are proven in [47]. The NSCM estimate is given in closed form, but it is a biased estimate [45]. However, detectors utilizing these two estimators do not work well with limited sample size [33] (See also Sec. VI).

## IV. Estimation of Kronecker Structured CM

### A. Kronecker Structured CM Estimators

In this section, we consider CM estimation for a Kronecker structured CM. In particular, we propose two estimators, namely, the Kronecker-structured MLE (KMLE) and the Kronecker-structured NSCM (KNSCM). Estimates of $\mathbf{R}_p$ and $\mathbf{R}_s$ are obtained via the proposed estimators, and subsequently, the estimate of $\mathbf{R}$ is obtained according to (8).

*1) Kronecker MLE:* The PDF of the training samples $\boldsymbol{Y} = [\mathbf{y}_1, \cdots, \mathbf{y}_K] \in \mathbb{C}^{N\times K}$ conditioned on the textures $\tau_k$s is given by

$$p(\boldsymbol{Y}|\tau_1,\cdots,\tau_K) \propto \frac{1}{\prod_{k=1}^{K}\tau_k \det(\mathbf{R})} \cdot \exp\left\{-\sum_{k=1}^{K}\frac{\mathbf{y}_k{}^H\mathbf{R}^{-1}\mathbf{y}_k}{\tau_k}\right\}. \tag{17}$$

After integration over all the $\tau_k$s, and by using (11), the marginal PDF of $\boldsymbol{Y}$ becomes

$$p(\boldsymbol{Y};\mathbf{R}) \propto \frac{1}{\det(\mathbf{R})^K}\prod_{k=1}^{K}\left(v+\mathbf{y}_k^H\mathbf{R}^{-1}\mathbf{y}_k\right)^{(-v-N)}, \tag{18}$$

where $\mathbf{R} = \mathbf{R}_p \otimes \mathbf{R}_s$.

In order to obtain KMLE, we need to maximize the likelihood function given by (18) with respect to $\mathbf{R}_p$ and $\mathbf{R}_s$. For the case of compound Gaussian clutter with deterministic texture ($v = 0$), KMLE is obtained via the majorization minimization (MM) method [21]. The MM method can be easily extended to random textures. For example, for the compound Gaussian case with IG texture, the ML estimates of $\mathbf{R}_s$ and $\mathbf{R}_p$ can be obtained by solving the fixed point equations

$$\hat{\mathbf{R}}_s = \frac{v+N}{KN_p}\sum_{k=1}^{K}\frac{\mathbf{Y}_k\hat{\mathbf{R}}_p^{-T}\mathbf{Y}_k^H}{v+\mathbf{y}_k^H(\hat{\mathbf{R}}_p^{-1}\otimes\hat{\mathbf{R}}_s^{-1})\mathbf{y}_k}, \tag{19}$$

$$\hat{\mathbf{R}}_p = \frac{v+N}{KN_s}\sum_{k=1}^{K}\frac{\mathbf{Y}_k^T\hat{\mathbf{R}}_s^{-T}\mathbf{Y}_k^*}{v+\mathbf{y}_k^H(\hat{\mathbf{R}}_p^{-1}\otimes\hat{\mathbf{R}}_s^{-1})\mathbf{y}_k}, \tag{20}$$

where $\mathbf{Y}_k = [\mathbf{y}_{k1},\cdots,\mathbf{y}_{kN_p}] \in \mathbb{C}^{N_s\times N_p}$, and $\mathbf{y}_{ki} \in \mathbb{C}^{N_s\times 1}$, $i = 1,\cdots,N_p$ is the received data in the $i$th polarization channel of the $k$th training sample $\mathbf{y}_k = \left[\mathbf{y}_{k1}^T\ \mathbf{y}_{k2}^T\ \cdots\ \mathbf{y}_{kN_p}^T\right]^T$. Finally, KMLE is given by

$$\hat{\mathbf{R}}_{\text{KMLE}} = \hat{\mathbf{R}}_p \otimes \hat{\mathbf{R}}_s. \tag{21}$$

Note that for $v = 0$, KMLE corresponds to Tyler's estimator with deterministic texture [21]. In order to avoid the scale ambiguity resulting from the unknown deterministic texture, normalization should be applied such that $\text{tr}\{\hat{\mathbf{R}}_{\text{KMLE}}\} = N$ [21]. For the random texture ($v > 0$), this normalization could also be applied to KMLE, however, it is not necessary to do so, because the shape parameter already provides information on the amplitude of the clutter returns. The normalization can speed up convergence because it brings the estimate at each iteration closer to the solution of the fixed point equations, especially in cases of bad initialization. For $v \to \infty$, KMLE corresponds to the Flip-Flop algorithm [20] for the Gaussian case. Further, based on (19) and (20), in order to guarantee that KMLE is non-singular, $K$ should satisfy $K \geq \max\{\frac{N_p}{N_s}, \frac{N_s}{N_p}\}$.

It is well known that the ML estimate is asymptotically unbiased and efficient for many common cases[1] [51]. Also, the ML estimate is invariant, which may lead to the CFAR property of the ASMF detector used along with the ML estimate (See Sec. V.A and [31]). Here we give some preliminaries of the invariance principle, which is central in detection analysis.

*Definition 1 (Model Invariance [28]):* Let $\mathcal{P} = \{\boldsymbol{P}_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \boldsymbol{\Omega}\}$ be a family of distributions determined by the parameter space $\boldsymbol{\Omega}$ and the random variable $\boldsymbol{Y} \sim \boldsymbol{P}_{\boldsymbol{\theta}}$. Define the transformation group [29] $\mathcal{G}$ of one-to-one functions $\{g\}$ on $\boldsymbol{Y}$. If $g\boldsymbol{Y} \sim \boldsymbol{P}_{\boldsymbol{\theta}'} \in \mathcal{P}$, then we say $\mathcal{P}$ is invariant under $\mathcal{G}$. This induces a group $\mathcal{F}$ on the parameter space $\boldsymbol{\Omega}$, i.e., $f \in \mathcal{F} : \boldsymbol{\Omega} \mapsto \boldsymbol{\Omega}$ such that $f\boldsymbol{\theta} = \boldsymbol{\theta}'$.

*Definition 2 (Invariant Estimator [28]):* Let $\mathcal{P}$ be invariant under the group $\mathcal{G}$, and $\boldsymbol{Y} \sim \boldsymbol{P}_{\boldsymbol{\theta}}$. $\hat{\boldsymbol{\theta}}(\boldsymbol{Y})$ is an invariant estimator of $\boldsymbol{\theta}$ under the transformation group $\mathcal{G}$, if $\hat{\boldsymbol{\theta}}(g\boldsymbol{Y}) = f\hat{\boldsymbol{\theta}}(\boldsymbol{Y})$ for all $\boldsymbol{\theta} \in \boldsymbol{\Omega}$, i.e., the estimate of $f\boldsymbol{\theta}$ is $f\hat{\boldsymbol{\theta}}(\boldsymbol{Y})$.

*Proposition 1:* Let $\mathcal{P} = \{\boldsymbol{P}_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \boldsymbol{\Omega}\}$ be invariant under $\mathcal{G}$. For $g \in \mathcal{G}$ and $\boldsymbol{Y} \sim \boldsymbol{P}_{\boldsymbol{\theta}}$, if the Jacobian of the transformation $d(g\boldsymbol{Y}) = \det\{\boldsymbol{\mathcal{J}}\}d\boldsymbol{Y}$ is independent of $\boldsymbol{\theta}$, where $\det\{\boldsymbol{\mathcal{J}}\}$ is the Jacobian determinant, then the ML estimate of $\boldsymbol{\theta}$ is an invariant estimator under $\mathcal{G}$.

*Proof:* See Appendix A. ■

It is easy to verify that the PDF given by (18) is invariant under the transformation group $\mathcal{G}_\kappa$, where the group $\mathcal{G}_k$ and its induced $\mathcal{F}_k$ are respectively given by

$$\mathcal{G}_\kappa = \{g|g\boldsymbol{Y} = \mathbf{G}_\kappa\boldsymbol{Y}\}, \mathcal{F}_\kappa = \{f|f\mathbf{R} = \mathbf{G}_\kappa\mathbf{R}\mathbf{G}_\kappa^H\} \tag{22}$$

where $\mathbf{G}_\kappa = \mathbf{G}_p \otimes \mathbf{G}_s$ with $\mathbf{G}_p \in \mathbb{C}^{N_p\times N_p}$ and $\mathbf{G}_s \in \mathbb{C}^{N_s\times N_s}$ being nonsingular matrices, and the corresponding parameter space being

$$\boldsymbol{\Omega}_\kappa = \{\mathbf{R}|\mathbf{R} = \mathbf{R}_p \otimes \mathbf{R}_s, \mathbf{R}_p \in \mathbb{H}^+_{N_p\times N_p}, \mathbf{R}_s \in \mathbb{H}^+_{N_s\times N_s}\}, \tag{23}$$

where $\mathbb{H}^+_{N\times N}$ is the set of $N$ by $N$ positive definite Hermitian matrices.

[1]The ML estimate may be biased, and not asymptotically efficient in some cases [34], e.g., when the unknown parameters are on the boundary of a continuous parameter space, or the parameter space is finite or discrete.

Then, note that the corresponding Jacobian determinant $\det\{\mathcal{J}\} = \det\{\mathbf{G}_\kappa \mathbf{G}_\kappa^H\}^K$ of $\mathcal{G}_\kappa$ is independent of $\mathbf{R}$. Then, according to Proposition 1, KMLE is an invariant estimator under $\mathcal{G}_\kappa$.

*Remark 1:* The PDF given by (18) is also invariant under the general linear transformation group [29] $\mathcal{G}_\upsilon$, with

$$\mathcal{G}_\upsilon = \{g | g\mathbf{Y} = \mathbf{G}_\upsilon \mathbf{Y}\}, \mathcal{F}_\upsilon = \{f | f\mathbf{R} = \mathbf{G}_\upsilon \mathbf{R} \mathbf{G}_\upsilon^H\} \quad (24)$$

where $\mathbf{G}_\upsilon \in \mathbb{C}^{N\times N}$ is a nonsingular matrix and

$$\mathbf{\Omega}_\upsilon = \{\mathbf{R} | \mathbf{R} \in \mathbb{H}_{N\times N}^+\}. \quad (25)$$

Therefore, MLE given by (15) is an invariant estimator under $\mathcal{G}_\upsilon$ for $\mathbf{\Omega}_\upsilon$. It is obvious that $\mathcal{G}_\kappa$ is a subgroup [29] of $\mathcal{G}_\upsilon$ and $\mathbf{\Omega}_\kappa$ is a subset of $\mathbf{\Omega}_\upsilon$. According to Proposition 1, MLE is also an invariant estimator under $\mathcal{G}_\kappa$ for $\mathbf{\Omega}_\kappa$. In contrast, KMLE is not invariant under the group $\mathcal{G}_\upsilon$.

There are also some other subgroups of the general linear group $\mathcal{G}_\upsilon$ comprised by a set of structural matrices [30], such as persymmetric matrices [33] and circulant matrices [52]. This structural information is commonly exploited in radar CFAR detection to improve the detection performance [31]. Also, not all structured CMs can form a subgroup of $\mathcal{G}_\upsilon$. For example, Toeplitz structured CMs cannot form a subgroup because the product of two Toeplitz matrices and the inversion of a Toeplitz matrix is not Toeplitz.

*2) Kronecker NSCM Estimator:* Matrix inversions and iterations are required to obtain KMLE, which leads to a relatively high computational load. Inspired by the NSCM estimator given in (16), we now consider an alternative closed-form estimator.

It holds that

$$\begin{aligned}\mathbb{E}\{\mathbf{Y}_k \mathbf{Y}_k^H | \tau_k\} &= \tau_k \mathrm{tr}(\mathbf{R}_p)\mathbf{R}_s, \\ \mathbb{E}\{\mathbf{Y}_k^T \mathbf{Y}_k^* | \tau_k\} &= \tau_k \mathrm{tr}(\mathbf{R}_s)\mathbf{R}_p. \end{aligned} \quad (26)$$

Eq. (26) implies that $\mathbf{R}_s$ (or $\mathbf{R}_p$) may be respectively estimated by $\mathbf{Y}_k \mathbf{Y}_k^H$ (or $\mathbf{Y}_k^T \mathbf{Y}_k^*$) up to a real-valued scalar, e.g., $\tau_k \mathrm{tr}(\mathbf{R}_p)$ (or $\tau_k \mathrm{tr}(\mathbf{R}_s)$). Note that there is no loss of correlation among the entries due to the scalar, which means that the ratio between off-diagonal and diagonal entries is not changed. Hence, we may impose a condition on the trace such that $\mathrm{tr}(\hat{\mathbf{R}}_{\mathrm{NSCM}}) = N$ and the scale parameter can be absorbed in the clutter texture. Thus, the estimates of the sub-matrices of KNSCM are respectively given by

$$\begin{aligned}\tilde{\mathbf{R}}_s &= \frac{N_s}{K} \sum_{k=1}^K \frac{\mathbf{Y}_k \mathbf{Y}_k^H}{\hat{\beta}_k}, \\ \tilde{\mathbf{R}}_p &= \frac{N_p}{K} \sum_{k=1}^K \frac{\mathbf{Y}_k^T \mathbf{Y}_k^*}{\hat{\beta}_k}, \end{aligned} \quad (27)$$

where $\hat{\beta}_k = \mathrm{tr}(\mathbf{Y}_k \mathbf{Y}_k^H)$ is introduced to normalize $\tilde{\mathbf{R}}_s$ and $\tilde{\mathbf{R}}_p$. Subsequently, the KNSCM estimate of the CM is given by

$$\hat{\mathbf{R}}_{\mathrm{KNSCM}} = \tilde{\mathbf{R}}_p \otimes \tilde{\mathbf{R}}_s. \quad (28)$$

According to (27) and (28), KNSCM can be implemented easily since there is no matrix inversion calculation. However, KNSCM is biased due to its trace constraint, and is not invariant under group $\mathcal{G}_\kappa$. Thus more false alarms may occur with the KNSCM-ASMF detector.

### B. The Cramér-Rao Bound

With the Kronecker structure considered, the parameter space $\mathbf{\Omega}_\upsilon$ for CM estimation is reduced to $\mathbf{\Omega}_\kappa$. Naturally, the estimation accuracy can be improved where there are fewer unknown parameters to be estimated.

To be specific, for the unstructured covariance matrix $\mathbf{R} \in \mathbf{\Omega}_\upsilon$, the unknown parameter vector is given by

$$\boldsymbol{\theta}_\upsilon = [\Re\{\mathrm{vech}\{\mathbf{R}\}\}^T, \Im\{\underline{\mathrm{vech}}\{\mathbf{R}\}\}^T]^T \in \mathbb{R}^{N^2\times 1}. \quad (29)$$

Similarly, for $\mathbf{R} = \mathbf{R}_p \otimes \mathbf{R}_s \in \mathbf{\Omega}_\kappa$, the unknown parameter vector is given by

$$\boldsymbol{\theta}_\kappa \triangleq [\boldsymbol{\theta}_p^T \quad \boldsymbol{\theta}_s^T]^T \in \mathbb{R}^{(N_p^2+N_s^2)\times 1} \quad (30)$$

where $\boldsymbol{\theta}_p$ and $\boldsymbol{\theta}_s$ contain respectively $N_p^2$ and $N_s^2$ real components of the Hermitian sub-matrices $\mathbf{R}_p$ and $\mathbf{R}_s$.

As is well known, the MSE of an unbiased estimate of the unknown parameter vector $\boldsymbol{\theta} \in \mathbb{R}^{r\times 1}$ is lower bounded by the CRB, which is a popular metric of estimation accuracy [41]. The CRB is given by the inverse of the Fisher information matrix (FIM) $\boldsymbol{\mathcal{I}}$. In the compound Gaussian case, the FIM of the unknown elements in $\mathbf{R}$ (see eq. (19b) in [37]) can be rewritten as

$$\boldsymbol{\mathcal{I}} = K\mathbf{H}^H \mathbf{\Sigma} \mathbf{H}, \quad (31)$$

where

$$\mathbf{H} = \frac{\partial \mathrm{vec}\{\mathbf{R}\}}{\partial \boldsymbol{\theta}^T} \in \mathbb{C}^{N^2\times r}, \quad (32)$$

$r = N^2$ corresponds to $\boldsymbol{\theta}_\upsilon$, $r = N_p^2 + N_s^2$ corresponds to $\boldsymbol{\theta}_\kappa$ and

$$\mathbf{\Sigma} = v_1(\mathbf{R}^{-T} \otimes \mathbf{R}^{-1}) - v_2 \mathrm{vec}(\mathbf{R}^{-1})\mathrm{vec}(\mathbf{R}^{-1})^H, \quad (33)$$

where $v_1$ and $v_2$ are related to the texture distribution [37]. For the inverse Gamma texture, it holds that

$$v_1 \triangleq \frac{v+N}{v+N+1}, \quad v_2 \triangleq \frac{1}{v+N+1}. \quad (34)$$

According to the results in [20], [41], the MSE of an unbiased CM estimate is lower bounded by the CRB matrix, i.e.,

$$\boldsymbol{\mathcal{E}} \succeq \mathbf{H}\boldsymbol{\mathcal{I}}^\dagger \mathbf{H}^H, \quad (35)$$

where

$$\boldsymbol{\mathcal{E}} \triangleq \mathrm{cov}\left\{\mathrm{vec}\{\hat{\mathbf{R}}\}\right\} = \mathbb{E}\left\{\delta\mathbf{R}\delta\mathbf{R}^H\right\} \quad (36)$$

is the MSE of the estimate $\hat{\mathbf{R}}$ with $\delta\mathbf{R} \triangleq \mathrm{vec}\{\hat{\mathbf{R}} - \mathbf{R}\}$.

*Proposition 2:* The CRB matrices for $\mathbf{\Omega}_\kappa$ and $\mathbf{\Omega}_\upsilon$ are respectively given by

$$\mathbf{\Xi} = \frac{c_1}{K}(\mathbf{R}^T \otimes \mathbf{R})^{\frac{1}{2}} \mathbf{P}_\kappa (\mathbf{R}^T \otimes \mathbf{R})^{\frac{1}{2}} + \frac{c_2}{K}\mathrm{vec}(\mathbf{R})\mathrm{vec}(\mathbf{R})^H, \quad (37)$$

$$\mathbf{\Lambda} = \frac{c_1}{K}(\mathbf{R}^T \otimes \mathbf{R}) + \frac{c_2}{K}\mathrm{vec}(\mathbf{R})\mathrm{vec}(\mathbf{R})^H \quad (38)$$

where

$$c_1 \triangleq \frac{1}{v_1}, \quad c_2 \triangleq \frac{v_2}{v_1(v_1 - v_2 N)}, \tag{39}$$

$\mathbf{P}_\kappa = [(\mathbf{R}^T \otimes \mathbf{R})^{-\frac{1}{2}} \mathbf{H}_\kappa][(\mathbf{R}^T \otimes \mathbf{R})^{-\frac{1}{2}} \mathbf{H}_\kappa]^\dagger$, and $\mathbf{H}_\kappa = \frac{\partial \mathrm{vec}\{\mathbf{R}\}}{\partial \boldsymbol{\theta}_\kappa^T}$.

*Proof:* See Appendix B. ■

The improvement on the estimation accuracy, due to exploiting the Kronecker structure, can be evaluated by the ratio of the traces of the CRB matrices of $\boldsymbol{\Omega}_\kappa$ and $\boldsymbol{\Omega}_\upsilon$, i.e.,

$$\mathcal{L} \triangleq \frac{\mathrm{tr}\{\boldsymbol{\Lambda}\}}{\mathrm{tr}\{\boldsymbol{\Xi}\}}. \tag{40}$$

Since the CRBs are inversely proportional to the training size $K$, the above metric also determines the training size required by KMLE relative to MLE, in order to achieve the same estimation accuracy.

As shown in Appendix B, it holds that $\boldsymbol{\Sigma}^{-1} = K\boldsymbol{\Lambda}$. According to (31) and (35), the trace of the CRB matrix of $\boldsymbol{\Omega}_\kappa$ may be rewritten as

$$\mathrm{tr}\{\boldsymbol{\Xi}\} = \mathrm{tr}\{\mathbf{P}\boldsymbol{\Lambda}\} \tag{41}$$

where $\mathbf{P}$ is the projection matrix of the column space of $\boldsymbol{\Sigma}^{\frac{1}{2}}\mathbf{H}$. Note that the rank of $\mathbf{P}$ is $N_s^2 + N_p^2 - 1$ [20]. Using Lemma 1 of [42], we have $\mathcal{L} > 1$, which means that the estimation accuracy of the CM improves when the Kronecker structure of the CM is exploited.

*Remark 2:* In this paper, the estimation error is evaluated based on the CRB and the MSE defined through the square of the Frobenius norm. Although this is a common choice in estimation problems, one needs to note that the parameter space for CM estimation is a manifold of Hermitian positive definite matrices, rather than a vector space. According to the intrinsic estimation theory [43], the intrinsic CRB and the natural metric on this manifold which involves the generalized eigenvalues of the CM and its estimate, would be more appropriate in this context [43, Theorem 4]. However, applying the natural metric does not modify the hierarchy between the estimators as compared to the flat metric (this is also mentioned in [44]). Also, the Frobenius norm provides more intuitive and concise results for theoretic analysis, and works well in our simulations.

## V. Detection Performance Analysis

### A. CFAR Analysis

The constant false alarm ratio (CFAR) property guarantees that the detection threshold is independent of the clutter parameters, and thus is critical for radar detection. It is obvious that the proposed ASMF detector is not CFAR with respect to the texture component, because it is related to the shape parameter $v$. If we need a CFAR detector with respect to the texture, we can employ the adaptive subspace detector (ASD) given in [48] and Tyler's estimator [21], which is a special case of (14) with $v = 0$. In the following, we examine the CFAR property with respect to the Kronecker structured CM for the ASMF detector using $\hat{\mathbf{R}}_{\mathrm{KMLE}}$ and $\hat{\mathbf{R}}_{\mathrm{KNSCM}}$, namely, KMLE-ASMF and KNSCM-ASMF.

*Proposition 3:* The KMLE-ASMF detector is CFAR with respect to the Kronecker structured CM estimate, while the KNSCM-ASMF detector is not.

*Proof:* See Appendix C. ■

The KNSCM-ASMF detector is approximately CFAR when the true CM is close to the identity matrix, because NSCM estimators are unbiased when the CM is the identity matrix along the lines of [45], [46] (see also Sec. VI).

### B. Average SCR Loss

The SCR loss (SCRL) of an adaptive detector is defined as the ratio of the SCR of the MVDR filter using a CM estimate, $\hat{\mathbf{R}}$, and that of the optimal MVDR filter using the true CM, $\mathbf{R}$ [7], [24], i.e.,

$$\rho \triangleq \rho(\hat{\mathbf{R}}) = \frac{|\mathbf{t}^H \hat{\mathbf{R}}^{-1}\mathbf{t}|^2}{(\mathbf{t}^H \mathbf{R}^{-1}\mathbf{t})(\mathbf{t}^H \hat{\mathbf{R}}^{-1}\mathbf{R}\hat{\mathbf{R}}^{-1}\mathbf{t})}, \tag{42}$$

where $\mathbf{t} \in \mathbb{C}^{N\times 1}$ is the steering vector of the target and equals $\mathbf{t} = \mathbf{A}\boldsymbol{\alpha} = \boldsymbol{\alpha} \otimes \mathbf{a}$ for the polarimetric array model in (1). Note that SCRL is scale-invariant with respect to $\mathbf{R}$ and $\hat{\mathbf{R}}$, so the SCRL is only related to the texture by $\hat{\mathbf{R}}$.

We develop a unified method to quantify the relation between the SCRL and the sample support by approximating the average SCRL as a function of the MSE of the CM estimate.

*Theorem 1:* Given a covariance matrix estimate $\hat{\mathbf{R}}$ with sufficiently small error, the average SCRL of the adaptive-subspace-matched-filter detector may be approximately expressed as

$$\bar{\rho} \triangleq \mathbb{E}\{\rho(\hat{\mathbf{R}})\} \approx 1 - \mathrm{tr}\{\boldsymbol{\Delta}\bar{\boldsymbol{\mathcal{E}}}\} \tag{43}$$

where $\boldsymbol{\Delta} \triangleq \left(\mathbf{I}_N - \mathbf{e}_1\mathbf{e}_1^T\right)^T \otimes \mathbf{e}_1\mathbf{e}_1^T$,

$$\bar{\boldsymbol{\mathcal{E}}} = \left((\mathbf{R}^{-\frac{1}{2}}\mathbf{U})^T \otimes \mathbf{U}^H\mathbf{R}^{-\frac{1}{2}}\right)\boldsymbol{\mathcal{E}}\left((\mathbf{R}^{-\frac{1}{2}}\mathbf{U})^T \otimes \mathbf{U}^H\mathbf{R}^{-\frac{1}{2}}\right)^H,$$

$\boldsymbol{\mathcal{E}}$ is the MSE of the estimate $\hat{\mathbf{R}}$ given in (36), and the unitary matrix $\mathbf{U}$ satisfies that $\mathbf{U}^H\mathbf{R}^{-\frac{1}{2}}\mathbf{t} = (\mathbf{t}^H\mathbf{R}^{-1}\mathbf{t})^{\frac{1}{2}}\mathbf{e}_1$.

*Proof:* See Appendix D. ■

According to Theorem 1, the average SCRL involves the sum of the corresponding $N-1$ diagonal entries of $\bar{\boldsymbol{\mathcal{E}}}$, which can be regarded as the whitened MSE of the CM estimate $\hat{\mathbf{R}}$. Specifically, we just need to calculate the whitened MSE corresponding to the entries of the first column or the first row of $\mathbf{R}$ except for the $(1,1)$th entry. Theorem 1 is suitable when the asymptotic MSE of a CM estimate is available. Fortunately, the asymptotic MSEs of several popular CM estimates can be easily obtained, such as the ML estimate and the generalized least square estimate [20].

*Example 1:* In the Gaussian scenario, the sample covariance matrix is given by $\mathbf{S} = \frac{1}{K}\sum_{k=1}^K \mathbf{y}_k\mathbf{y}_k^H$, the MSE of which is $\boldsymbol{\mathcal{E}} = \mathrm{cov}\{\mathrm{vec}\{\mathbf{S}\}\} = \frac{1}{K}(\mathbf{R}^T \otimes \mathbf{R})$. Thus, in view of Theorem 1, we have $\bar{\rho} \approx 1 - \frac{N-1}{K}$ for the SCM. With $\bar{\rho} = 0.5$, the number of training samples is $K \approx 2N - 2$, which is very close to the result $K = 2N - 3$, reported in [7].

*Example 2:* In the Gaussian scenario, the complex-valued persymmetric structured covariance matrix $\mathbf{R} \in \mathbb{C}^{N\times N}$ [24] can be equivalently transformed into a real-valued symmetric counterpart $\mathbf{R}_{\mathrm{ps}} = \mathbf{T}\mathbf{R}\mathbf{T}^H \in \mathbb{R}^{N\times N}$, where $\mathbf{T} \in \mathbb{C}^{N\times N}$ is

given in [25, eq. (13)]. The real-valued CM estimator is given by $\hat{\mathbf{R}}_{\rm ps} = \Re\{\mathbf{TST}^H\}$, with $\mathbf{S} = \frac{1}{K}\sum_{k=1}^{K}\mathbf{y}_k\mathbf{y}_k^H$, and $K\hat{\mathbf{R}}_{\rm ps}$ is a real Wishart matrix with $2K$ degrees of freedom and the parameter matrix equal to $\frac{1}{2}\mathbf{R}_{\rm ps}$ [25]. The asymptotic MSE of $\hat{\mathbf{R}}_{\rm ps}$ equals $\boldsymbol{\mathcal{E}} = \mathrm{cov}\{\mathrm{vec}\{\hat{\mathbf{R}}_{\rm ps}\}\} = \frac{1}{2K}(\mathbf{R}_{\rm ps}^T \otimes \mathbf{R}_{\rm ps})(\mathbf{I}_{N^2} + \mathbf{K}_N)$ [50], where $\mathbf{K}_N = \sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{e}_i\mathbf{e}_j^T \otimes \mathbf{e}_j\mathbf{e}_i^T$ is the commutation matrix. Note that $\mathrm{tr}\{\boldsymbol{\Delta}\mathbf{K}_N\} = 0$ and $\mathbf{K}_N(\mathbf{A}\otimes\mathbf{B}) = (\mathbf{B}\otimes\mathbf{A})\mathbf{K}_N$. According to Theorem 1, the average SCRL of the detector with the persymmetric structured CM is given by $\bar{\rho}_{\rm ps} \approx 1 - \frac{N-1}{2K}$. With $\bar{\rho}_{\rm ps} = 0.5$, the amount of training samples is $K \approx N-1$. Thus, detectors using the persymmetric structure can reduce by one half the number of training samples as compared to the unstructured counterpart.

Further, if the estimate $\hat{\mathbf{R}}$ is an invariant estimate under a subgroup of $\mathcal{G}_v$, and $\mathbf{U}^H\mathbf{R}^{-\frac{1}{2}}$ belongs to the subgroup, we have that

$$\bar{\boldsymbol{\mathcal{E}}}\left(\mathbf{R}, \hat{\mathbf{R}}(\boldsymbol{Y})\right) = \bar{\boldsymbol{\mathcal{E}}}\left(f\mathbf{R}, \hat{\mathbf{R}}(g\boldsymbol{Y})\right) = \boldsymbol{\mathcal{E}}(\mathbf{I}_N), \tag{44}$$

where $\boldsymbol{\mathcal{E}}(\mathbf{I}_N)$ is the MSE of the estimate when $\mathbf{R} = \mathbf{I}_N$. According to (44), the average SCRL is an invariant loss function [28] under the group, which means that the average SCRL is independent of the CM and the steering vector.

To be specific, for KMLE, let $\mathbf{G}_\kappa = \mathbf{U}^H\mathbf{R}^{-\frac{1}{2}}$ with the unitary matrix $\mathbf{U}$ given by (70). Because $\mathbf{G}_\kappa \in \mathcal{G}_\kappa$ and $\mathbf{G}_\kappa\hat{\mathbf{R}}(\boldsymbol{Y})\mathbf{G}_\kappa^H = \hat{\mathbf{R}}(\mathbf{G}_\kappa\boldsymbol{Y})$ is KMLE of the identity matrix according to Definition 2, the average SCRL of KMLE is independent of the steering vector and the CM.

*Corollary 1:* In compound Gaussian clutter, for a sufficiently large $K$, the average SCRL for the adaptive-subspace-matched-filter detector using the KMLE $\hat{\mathbf{R}}_{\rm KMLE}$ is given by

$$\mathbb{E}\{\rho(\hat{\mathbf{R}}_{\rm KMLE})\} \approx 1 - \frac{c_1 M}{K} \tag{45}$$

where $c_1$ is given in (39) and

$$M = \frac{N_s - 1}{N_p} + \frac{N_p - 1}{N_s}. \tag{46}$$

*Proof:* See Appendix E. ■

According to (45), if $N_p = 1$, we have $M = N_s - 1$ and $\hat{\mathbf{R}}_{\rm KMLE}$ is equivalent to $\hat{\mathbf{R}}_{\rm MLE}$. On the other hand, for large $N_p$ and $N_s$, we have $c_1 \approx 1$ and $M \approx \frac{N_s}{N_p} + \frac{N_p}{N_s} \geq 2$. This implies that the smallest number of training samples for an average SCRL of 3 dB is only $K = 2c_1M \approx 4$. Intuitively, this is because the polarimetric and spatial channels provide extra training samples for the estimation of $\mathbf{R}_s$ and $\mathbf{R}_p$.

*Remark 3:* Corollary 1 can be extended to invariant estimates of other group symmetric covariance matrices. To be specific, the constant $c_1$ and $M$ are pertinent to the texture and the structural information, respectively. Thus, the average SCRLs of various scenarios with different texture modeling and group symmetric structures of the CM can be readily derived by simply substituting the corresponding $c_1$ for the particular texture [37], [45], and the corresponding $M$ for the structured CM. The corresponding $M$ can be obtained by calculating the CRB of the structured CM estimate at the identity matrix along the lines of Corollary 1. In Table I , we list results of several popular group invariant structures [30] without a proof, namely, Hermitian, persymmetric, circulant[2] and the quaternion Q-proper [30], [53] ($\mathbf{R}_p = \mathbf{I}_{N_p}$ and $N_p = 2$) CMs. Note that the improvement factors in the denominators in Table I (e.g., the improvement factor corresponding to the perymmetric structure equals 2) are related to the intrinsic degrees of the freedom of the corresponding groups defined in [30]. Based on Corollary 1 and Table I, we can further conjecture that the average SCRL for the Kronecker structured CM with group symmetric structured submatrices, e.g., $\mathbf{R}_p$ and $\mathbf{R}_s$ are endowed with the persymmetric structure [35]. Let $\delta_p$ and $\delta_s$ be the improvement factors for the corresponding group symmetric structures of $\mathbf{R}_p$ and $\mathbf{R}_s$ respectively. The corresponding $M$ for the Kronecker structured CM with group symmetric structured submatrices may be given by

$$M = \frac{N_s - 1}{\delta_s N_p} + \frac{N_p - 1}{\delta_p N_s}. \tag{47}$$

TABLE I
THE AVERAGE SCRLS FOR DIFFERENT STRUCTURES

| Hermitian | persymmetric | circulant | Q-proper |
|---|---|---|---|
| $1 - \frac{c_1(N-1)}{K}$ | $1 - \frac{c_1(N-1)}{2K}$ | $1 - \frac{c_1(N-1)}{KN}$ | $1 - \frac{c_1(N-1)}{2K}$ |

### C. Asymptotic False Alarm and Detection Probabilities

In this section, we assess the asymptotic performance of the KMLE-ASMF detector along the lines of SCRL analysis. The detection statistic of the KMLE-ASMF detector can be approximated by its first-order Taylor expansion at the true CM, i.e.,

$$T_\kappa \triangleq T(\hat{\mathbf{R}}_{\rm KMLE}) \approx T(\mathbf{R}) + \frac{\partial T(\mathbf{R})}{\partial \boldsymbol{\theta}_\kappa^T}(\boldsymbol{\theta}_\kappa - \hat{\boldsymbol{\theta}}_\kappa), \tag{48}$$

where we use the simplified notation $T(\mathbf{R}) \triangleq T(\mathbf{y}, \mathbf{R}, \mathbf{A}, v)$, and $\hat{\boldsymbol{\theta}}_\kappa$ is the real-valued components of the KMLE similar to $\boldsymbol{\theta}_\kappa$ in (30).

With the chain rule of the derivative, we have that

$$\frac{\partial T(\mathbf{R})}{\partial \boldsymbol{\theta}_\kappa^T} = \frac{\partial T(\mathbf{R})}{\partial \mathrm{vec}^T\{\mathbf{R}\}}\frac{\partial \mathrm{vec}\{\mathbf{R}\}}{\partial \boldsymbol{\theta}_\kappa^T} = \frac{\partial T(\mathbf{R})}{\partial \mathrm{vec}^T\{\mathbf{R}\}}\mathbf{H}_\kappa. \tag{49}$$

Note that $\frac{\partial T(\mathbf{R})}{\partial \boldsymbol{\theta}_\kappa^T}$ is real-valued, because $T(\mathbf{R})$ and $\boldsymbol{\theta}_k$ are real. According to (48) and (49), $T_\kappa$ conditioned on $T(\mathbf{R})$ is asymptotically Gaussian distributed, i.e.,

$$T_\kappa \sim \mathcal{N}(T(\mathbf{R}), \sigma^2_{\rm KMLE}), \tag{50}$$

where $\sigma^2_{\rm KMLE}$ is given by

$$\sigma^2_{\rm KMLE} = \frac{\partial T(\mathbf{R})}{\partial \mathrm{vec}^T(\mathbf{R})}\boldsymbol{\Xi}\frac{\partial T(\mathbf{R})}{\partial \mathrm{vec}(\mathbf{R}^T)}. \tag{51}$$

With the whitening technique (see (73) in Appendix C), we have that

$$T_\kappa = \frac{\mathbf{q}^H\tilde{\mathbf{M}}^{-1}\boldsymbol{\mathcal{Q}}(\boldsymbol{\mathcal{Q}}^H\tilde{\mathbf{M}}^{-1}\boldsymbol{\mathcal{Q}})^{-1}\boldsymbol{\mathcal{Q}}^H\tilde{\mathbf{M}}^{-1}\mathbf{q}}{\mathbf{q}^H\tilde{\mathbf{M}}^{-1}\mathbf{q} + v}, \tag{52}$$

[2]Circulant matrices can be diagonalized by the discrete Fourier transform (DFT) matrix [30], [52], and thus we only need to calculate the CRB of a diagonal matrix at $\mathbf{R} = \mathbf{I}_N$.

where $\tilde{\mathbf{M}}$ and $\boldsymbol{\mathcal{Q}}$ are given by (72) and (74) respectively, and $\mathbf{q} = \sqrt{\tau}\mathbf{z}$. Under hypothesis $\mathcal{H}_q$, conditioned on $\tau$, $\mathbf{z}|\tau \sim \mathcal{CN}(\boldsymbol{\mu}_q, \mathbf{I}_N)$, $q = 0, 1$, where

$$\boldsymbol{\mu}_q = q\tau^{-\frac{1}{2}}(\boldsymbol{\alpha}^H \mathbf{A}^H \mathbf{R}^{-1}\mathbf{A}\boldsymbol{\alpha})^{\frac{1}{2}}\mathbf{e}_1. \tag{53}$$

Here we assume that the texture component $\tau$ at the cell under test is independent of its corresponding speckle component $\mathbf{c}$.

In light of (50), we have that

$$T_\kappa \sim \mathcal{N}(T(\mathbf{q}, \mathbf{I}_N, \boldsymbol{\mathcal{Q}}, v), \sigma^2) \tag{54}$$

where the variance $\sigma^2 = \mathbf{d}^H \boldsymbol{\Xi}(\mathbf{I}_N)\mathbf{d}$, $\boldsymbol{\Xi}(\mathbf{I}_N)$ is the CRB given by (37) when $\mathbf{R} = \mathbf{I}_N$ and

$$\begin{aligned}\mathbf{d} &= \frac{\partial T(\mathbf{q}, \tilde{\mathbf{M}}, \boldsymbol{\mathcal{Q}}, v)}{\partial \mathrm{vec}(\tilde{\mathbf{M}}^T)}\Big|_{\tilde{\mathbf{M}}=\mathbf{I}_N} \\ &= \frac{\mathrm{vec}\{-\boldsymbol{\mathcal{Q}}\mathbf{q}\mathbf{q}^H - \mathbf{q}\mathbf{q}^H\boldsymbol{\mathcal{Q}}^H + \boldsymbol{\mathcal{Q}}\mathbf{q}\mathbf{q}^H\boldsymbol{\mathcal{Q}}^H\}}{\mathbf{q}^H\mathbf{q} + v} \\ &\quad + \frac{|\boldsymbol{\mathcal{Q}}\mathbf{q}|^2 \mathrm{vec}\{\mathbf{q}\mathbf{q}^H\}}{(\mathbf{q}^H\mathbf{q} + v)^2}.\end{aligned} \tag{55}$$

The derivation of (55) is given in Appendix F.

Based on (54), the asymptotic expressions of the probability of false alarm $P_{fa}$ and the probability of detection $P_d$ of the KMLE-ASMF detector are given in the following proposition.

*Proposition 4:* The asymptotic expressions of false alarm probability $P_{fa}$ and detection probability $P_d$ of the KMLE-ASMF detector are respectively given by

$$\begin{aligned}P_{fa} &= \boldsymbol{P}\{T_\kappa \geq \xi | \mathcal{H}_0\} \\ &= \mathbb{E}_{|\mathcal{H}_0, \mathbf{q}}\left\{1 - \Phi\left(\frac{\xi - T(\mathbf{q}, \mathbf{I}_N, \boldsymbol{\mathcal{Q}}, v)}{\sigma}\right)\right\},\end{aligned} \tag{56}$$

$$\begin{aligned}P_d &= \boldsymbol{P}\{T_\kappa \geq \xi | \mathcal{H}_1\} \\ &= \mathbb{E}_{|\mathcal{H}_1, \mathbf{q}}\left\{1 - \Phi\left(\frac{\xi - T(\mathbf{q}, \mathbf{I}_N, \boldsymbol{\mathcal{Q}}, v)}{\sigma}\right)\right\},\end{aligned} \tag{57}$$

where $\mathbb{E}_{|\mathcal{H}_q, \mathbf{q}}\{\cdot\}$ denotes expectation under hypothesis $\mathcal{H}_q$, $\Phi(\cdot)$ is the cumulative distribution function of the normalized Gaussian distribution, and the variance equals $\sigma^2 = \mathbf{d}^H \boldsymbol{\Xi}(\mathbf{I}_N)\mathbf{d}$.

Closed-form expressions of the expectations appearing in Proposition 4 are hardly tractable. Instead, we can use the Monte-Carlo sampling method to evaluate the integrals in (56) and (57). In light of the theoretical results on SCRL by exploiting the Kronecker structure in Corollary 1, the asymptotic detection performance may be achieved in the small sample regime.

## VI. Simulation Results

In this section, we compare the estimation accuracy of the proposed covariance matrix estimates, namely KMLE, and KNSCM against MLE and NSCM. We also compare KMLE and MLE against KNSCM and NSCM after they have been normalized in the same way as KNSCM and NSCM, such that $\mathrm{tr}(\hat{\mathbf{R}}) = N$; the normalized estimates will be respectively denoted by NKMLE and NMLE. It is assumed that $[\mathbf{R}_p]_{ij} = \epsilon_p^{|i-j|}$, $[\mathbf{R}_s]_{ij} = \epsilon_s^{|i-j|}$, $\epsilon = \epsilon_p = \epsilon_s$, for the sake of simplicity. We use

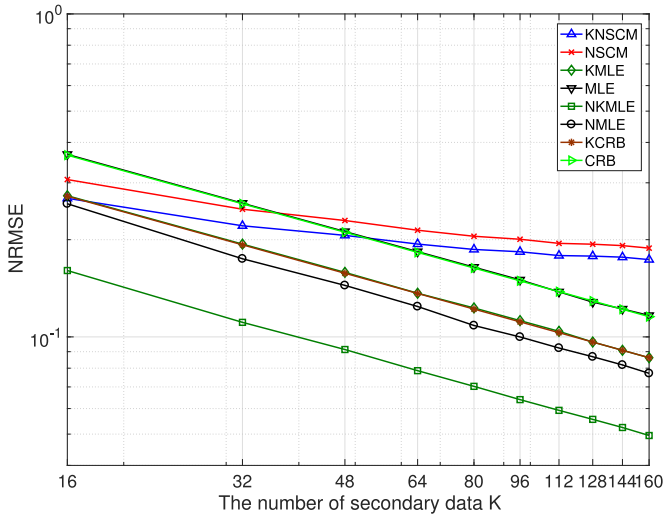


Fig. 1. Estimation accuracy comparison with $N_p = 2$, $N_s = 8$, $\epsilon = 0.9$.

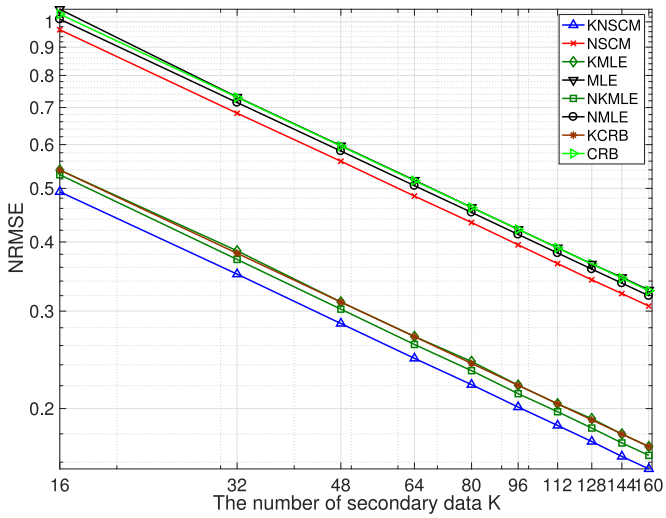


Fig. 2. Estimation accuracy comparison with $N_p = 2$, $N_s = 8$, $\epsilon = 0$.

the normalized root mean-square error (NRMSE) as the performance metric, which is defined as

$$\mathrm{NRMSE} \triangleq \sqrt{\frac{\mathbb{E}\left\{\left\|\hat{\mathbf{R}} - \mathbf{R}\right\|^2\right\}}{\|\mathbf{R}\|^2}}. \tag{58}$$

As illustrated in Fig. 1 and Fig. 2, by taking into account the Kronecker structure, both Kronecker structured estimates outperform their counterparts. Also, CM estimation under the trace constraint corresponds to the constrained CRB with equality constraints [54], which is lower than the CRB derived in Proposition 2. Thus, the trace constraint may provide extra information for CM estimation. Due to this reason, KNSCM and NSCM may achieve better estimation performance in the small sample regime as compared to KMLE and MLE. For large $\epsilon$ ($\epsilon = 0.9$), NKMLE and NMLE achieve lower NRMSEs than the corresponding CRBs, and also lower than KNSCM and NSCM estimates respectively (see Fig. 1). Also, for large $\epsilon$, the

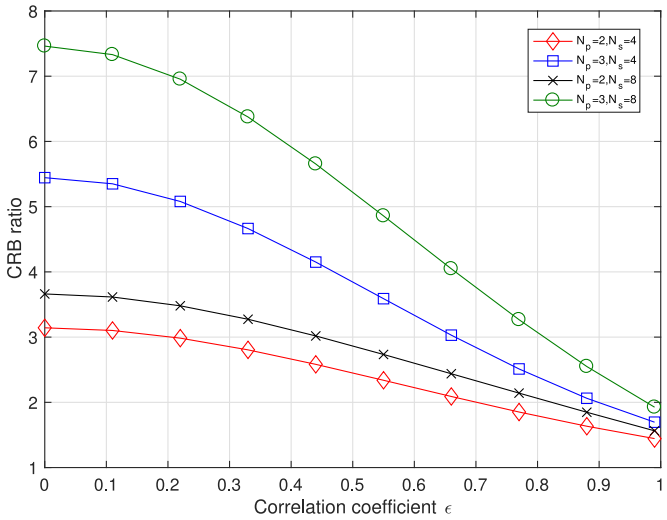


Fig. 3. The CRB ratio $\mathcal{L}$ as a function of the correlation coefficient $\epsilon$.

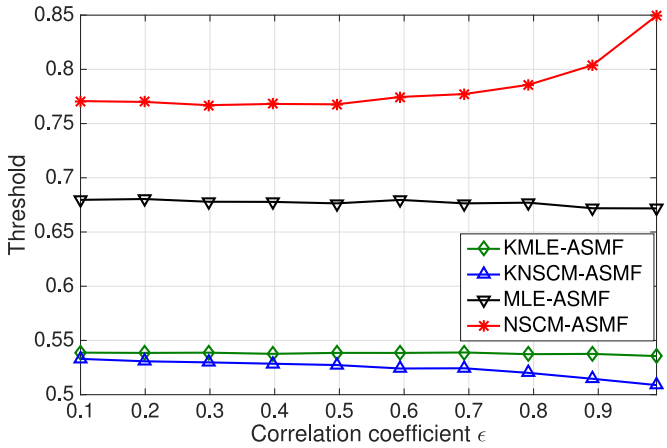


Fig. 4. Detection thresholds as a function of the correlation coefficient $\epsilon$, where $N_p = 2$, $N_s = 4$.

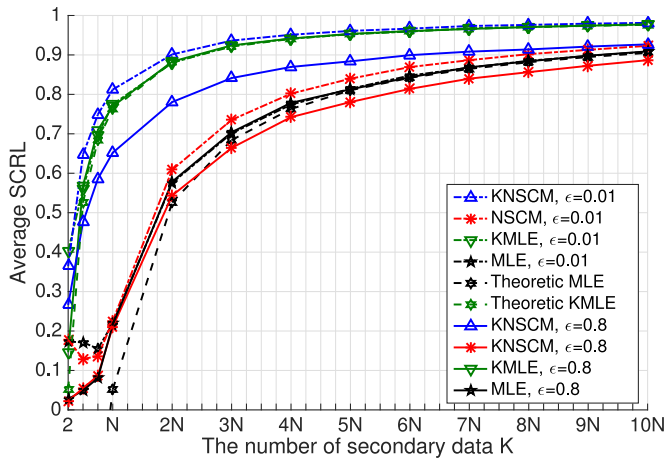


Fig. 5. Average SCRL as a function of the number of secondary data with $N_s = 4$, $v = 4$, $N_p = 2$.

NRMSEs of both KNSCM and NSCM decrease slowly in the large sample regime. For small $\epsilon$ ($\epsilon = 0$), KNSCM and NSCM can achieve better performance than NKMLE and NMLE (see Fig. 2). This is because NSCM is unbiased only when $\epsilon = 0$, and is a special case of the theoretical fixed point (TFP) estimate, which is the benchmark of Tyler's estimator [45], [46]. Similar results can be obtained for random texture cases and for KNSCM.

Next, we use the CRB ratio, $\mathcal{L}$, defined in (40) to evaluate the impact on the number of required training samples of exploiting the Kronecker structure for CM estimation, for different number of antennas, $N_s$, or polarimetric channels, $N_p$. As can be seen in Fig. 3, $\mathcal{L}$ is generally larger than one. This means that the Kronecker-structured estimator can accordingly achieve the same estimation accuracy as its unstructured counterpart with fewer secondary samples. As the correlation coefficient $\epsilon$ decreases, both the condition number of $\mathbf{R}$ and the sample size required by the Kronecker-structured estimator decrease. It can be observed that exploiting the Kronecker structure, the estimation accuracy can be further improved with either more antennas or more polarimetric channels.

To verify the CFAR property of the aforementioned detectors, we consider the scenario of a polarimetric uniform linear array [2], [12], [39] with $N_p = 2, N_s = 4$. The thresholds of the detectors with different correlation coefficients $\epsilon$ are compared in Fig. 4. It is observed that the thresholds of the NSCM and KNSCM detectors vary with different values of $\epsilon$; albeit the thresholds for both NSCM and KNSCM detectors level off in the small correlation coefficient regime; this is because both NSCM estimators are nearly unbiased when the true CM approaches the identity matrix with decreasing $\epsilon$, according to [45]. On the other hand, both the MLE and KMLE detectors are endowed with the CFAR property. Thus, both the MLE-based and KMLE-based detectors are more robust to time varying clutter.

In Figs. 5 and 6, we evaluate the average SCRLs of the considered detectors, namely, KMLE-ASMF, KNSCM-ASMF, MLE-ASMF, and NSCM-ASMF, as function of the number of secondary samples for $N_p = 2$ and $N_p = 3$ polarimetric channels, respectively. Note that we use the pseudo-inverse $\hat{\mathbf{R}}^{\dagger}$ in place of $\hat{\mathbf{R}}^{-1}$ for MLE and NSCM when $K < N$. The theoretical results (see (45)) are also shown on the figures. It is observed that the numerical results are accurately predicted by the theoretical expression in the large secondary sample regime. Remarkably, for KMLE, the prediction is effective even in the small sample regime ($K = 4$).

The average SCRLs of KMLE and MLE remain constant as the correlation coefficient $\epsilon$ varies, while the average SCRLs of NSCM and KNSCM do not. It is also observed that the average SCRLs of KNSCM and NSCM is higher than their MLE counterparts, when $\epsilon = 0.01$. This result is consistent with the results of Fig. 2. With $N_s = 4$, according to Table I, in order to achieve the same average SCRL, the MLE needs $\frac{N-1}{M} = 4$ times more training samples than KMLE for $N_p = 2$, and $\frac{N-1}{M} \approx 7.333$ times more samples for $N_p = 3$. This implies that exploiting the Kronecker structure can significantly reduce the amount of the required secondary data, especially with more polarimetric channels. It can be observed in Figs. 5 and 6 that for $\bar{\rho} = 0.5$, or 3 dB average SCRL, the detectors which exploit

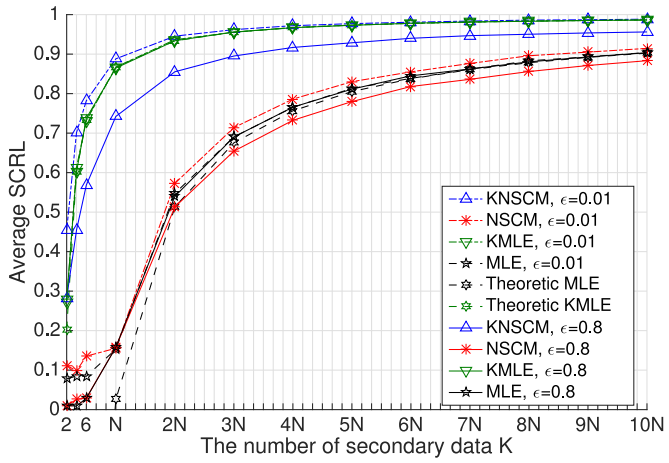


Fig. 6. Average SCRL as a function of the number of secondary data with $N_s = 4$, $v = 4$, $N_p = 3$.

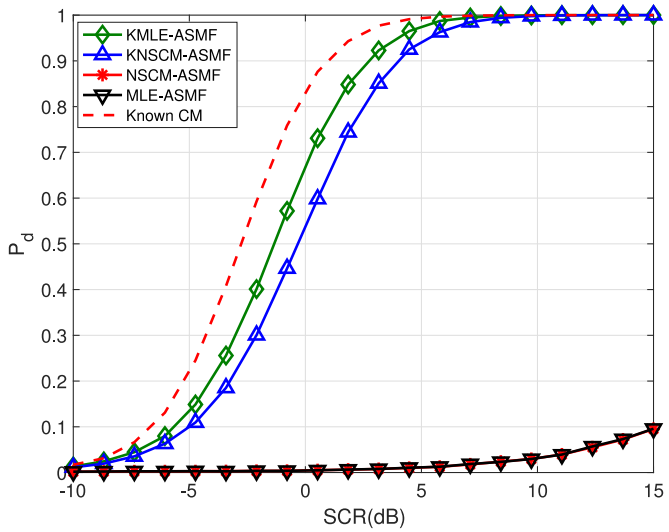


Fig. 7. Probability of detection as a function of SCR with $N_p = 2$, $N_s = 8$, $v = 4$, $\epsilon = 0.9$, $P_{fa} = 0.001$, $K = N$.

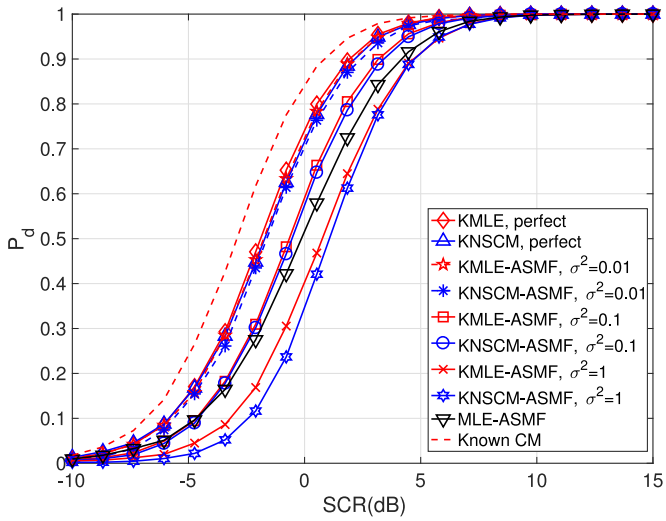


Fig. 8. Probability of detection as a function of SCR with $N_p = 2$, $N_s = 8$, $v = 4$, $\epsilon = 0.9$, $P_{fa} = 0.001$, $K = 2N$.

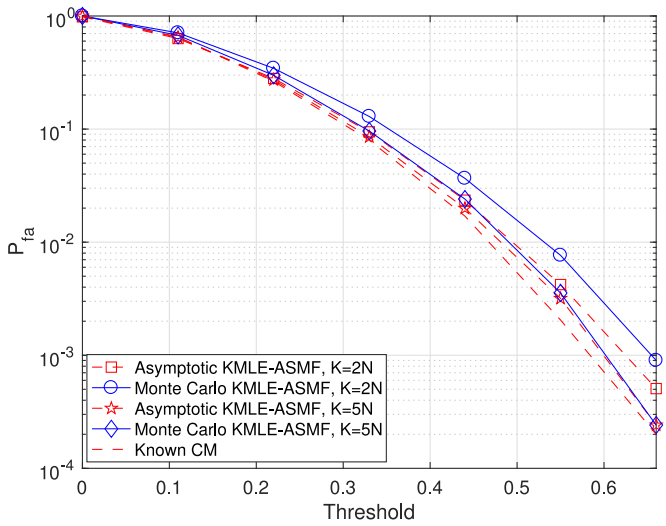


Fig. 9. Probability of false alarm as a function of the detection threshold with $v = 4$, $N_p = 2$, $N_s = 4$, $\epsilon = 0.9$.

the prior Kronecker structure information are able to work well even with only 4 secondary samples.

In Figs. 7 and 8, we show the detection performance of the ASMF detectors using KMLE, MLE, KNSCM and NSCM estimates as a function of SCR. Here, the SCR is defined as

$$\text{SCR(dB)} = 10\log_{10}\frac{v-1}{v}\frac{\boldsymbol{\alpha}^H\mathbf{A}^H\mathbf{R}^{-1}\mathbf{A}\boldsymbol{\alpha}}{N}. \qquad (59)$$

It is observed in Fig. 7 that the detectors which exploit the Kronecker structure can achieve considerably higher probability of detection, with a small amount of secondary data ($K = N$), while the performance of their unstructured counterparts is poor under the same conditions.

We also consider the case in which the Kronecker structure of the CM is imperfect, i.e., $\mathbf{R} = \mathbf{R}_p \otimes \mathbf{R}_s + \sigma^2\mathbf{R}_0$, where $\mathbf{R}_0 \in \mathbb{C}^{N\times N}$ is a random Hermitian positive definite matrix sampled from the complex Wishart distribution with $N$ degrees of freedom and the parameter matrix $\mathbf{I}_N$ models the imperfection, with $\sigma^2$ controlling the amount of imperfection. In our experiment, we normalize the CM such that $\text{tr}\{\mathbf{R}\} = N$, and maintain the same SCR as (59) for different levels of $\sigma^2$ by adjusting the power of the steering vector of the target. The results are shown in Fig. 8, where one can see that while for large $\sigma^2$ ($\sigma^2 = 1$), the detection performance of the KMLE-ASMF and KNSCM-ASMF detectors is worse than that of the MLE-ASMF detector, for small $\sigma^2$, exploiting the Kronecker structure of the CM can achieve better performance. Determining whether the CM has a Kronecker structure is discussed in [19], [20]. Once the Kronecker structure exists even within a certain degree of error, one can exploit it to achieve performance gains. This is because the likelihood function involving the Kronecker structured CM in the hypothesis test [19], [20] is higher than its unstructured counterpart.

Fig. 9 and Fig. 10 illustrate the asymptotic $P_{fa}$ and $P_d$ of the KMLE-ASMF detector when $K = 2N$ and $K = 5N$, respectively. It is observed that the asymptotic results are fairly

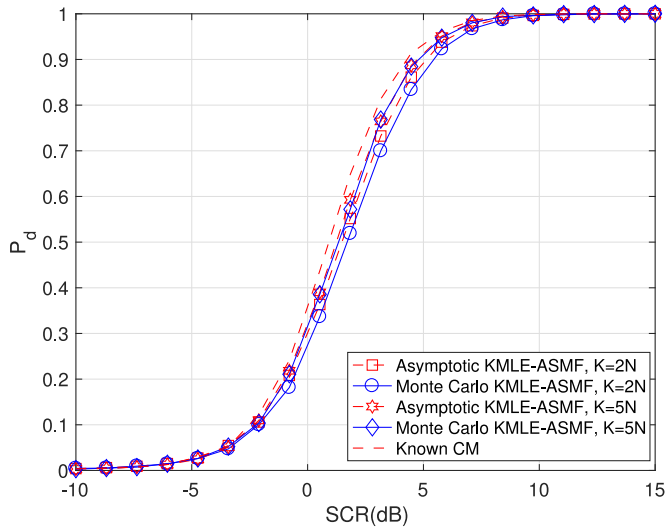


Fig. 10. Probability of detection as a function of SCR with $P_{fa} = 0.001$, $v = 4, N_p = 2, N_s = 4, \epsilon = 0.9$.

accurate and can provide a good suggestion for threshold setting and performance prediction. It is also observed that due to the improvement on sample support, the asymptotic performance is effective for $K = 5N$.

## VII. Conclusions

We have considered the problem of exploiting the *a priori* knowledge of a Kronecker structure of the CM for polarimetric detection problem in compound Gaussian clutter. We have shown via theory and simulations that the Kronecker-structured estimates and detectors can achieve significantly improved performance with limited number of secondary samples. We have analyzed the CFAR property for the proposed detectors using different CM estimates and derived a general closed-form representation of the average SCRL that can be used to determine the required training size for adaptive detection. It has been demonstrated that the adaptive detector using the Kronecker MLE can achieve asymptotically optimal performance with a relatively small number of secondary data. This result is helpful in threshold selection and performance prediction of the adaptive detector. The proposed analysis can be readily extended to adaptive detection in more general scenarios, where the ML estimate of the structured CM involves solving fixed-point equations.

## Appendix

### A. Proof of Proposition 1

*Proof:* Based on the Jacobian transformation and model invariance, the PDF of $\boldsymbol{Y}' = g\boldsymbol{Y}$ satisfies

$$p(\boldsymbol{Y}'; f\boldsymbol{\theta}) = p(g^{-1}\boldsymbol{Y}'; \boldsymbol{\theta})\det\{\boldsymbol{\mathcal{J}}\}^{-1}. \tag{60}$$

The proof of Proposition 1 is obtained as follows:

$$\begin{aligned}\hat{\boldsymbol{\theta}}_{\mathrm{ML}}(\boldsymbol{Y}) &= \arg\max_{\boldsymbol{\theta}} p(\boldsymbol{Y};\boldsymbol{\theta}) = \arg\max_{\boldsymbol{\theta}} p(g\boldsymbol{Y}; f\boldsymbol{\theta})\det\{\boldsymbol{\mathcal{J}}\} \\ &= f^{-1}\arg\max_{f\boldsymbol{\theta}} p(g\boldsymbol{Y}; f\boldsymbol{\theta}) = f^{-1}\hat{\boldsymbol{\theta}}_{\mathrm{ML}}(g\boldsymbol{Y}).\end{aligned} \tag{61}$$

In (61), the first equality is based on the definition of the ML estimate. The second one follows from (60). The third one follows because $\det\{\boldsymbol{\mathcal{J}}\}$ is independent of $\boldsymbol{\theta}$ and $f^{-1}$ exists. ■

### B. Proof of Proposition 2

*Proof:* Define $\mathbf{H}_1 \triangleq (\mathbf{R}^T \otimes \mathbf{R})^{-\frac{1}{2}}\mathbf{H}$, and $\mathbf{i}_N \triangleq \mathrm{vec}\{\mathbf{I}_N\}$. According to the fact that $\mathrm{vec}\{\mathbf{R}^{-1}\} = (\mathbf{R}^T \otimes \mathbf{R})^{-\frac{1}{2}}\mathbf{i}_N$, and the matrix inversion lemma [49],

$$\left(\mathbf{Q} - \mathbf{q}\mathbf{q}^H\right)^{\dagger} = \mathbf{Q}^{\dagger} + \frac{\mathbf{Q}^{\dagger}\mathbf{q}\mathbf{q}^H\mathbf{Q}^{\dagger}}{1 - \mathbf{q}^H\mathbf{Q}^{\dagger}\mathbf{q}} \tag{62}$$

where the semi-positive definite matrix $\mathbf{Q}$ and the vector $\mathbf{q}$ satisfy $\mathbf{q}^H\mathbf{Q}^{\dagger}\mathbf{q} < 1$, the CRB matrix given in (35) may be rewritten as

$$\begin{aligned}K \cdot \mathbf{H}\boldsymbol{\mathcal{I}}^{\dagger}\mathbf{H}^H &= \mathbf{H}(\mathbf{H}^H\boldsymbol{\Sigma}\mathbf{H})^{\dagger}\mathbf{H}^H \\ &= (\mathbf{R}^T \otimes \mathbf{R})^{\frac{1}{2}}\left(\frac{1}{v_1}\mathbf{P}_1 + \frac{v_2\mathbf{P}_1\mathbf{i}_N\mathbf{i}_N^H\mathbf{P}_1}{v_1^2 - v_1v_2\mathbf{i}_N^H\mathbf{P}_1\mathbf{i}_N}\right)(\mathbf{R}^T \otimes \mathbf{R})^{\frac{1}{2}}\end{aligned} \tag{63}$$

where $\mathbf{P}_1 \triangleq \mathbf{H}_1(\mathbf{H}_1^H\mathbf{H}_1)^{\dagger}\mathbf{H}_1^H = \mathbf{H}_1\mathbf{H}_1^{\dagger}$ is the projection matrix of $\mathbf{H}_1$.

To be specific, it is easy to verify that $\mathbf{H}_v = \frac{\partial \mathrm{vec}\{\mathbf{R}\}}{\partial \boldsymbol{\theta}_v^T} \in \mathbb{C}^{N^2 \times N^2}$ corresponding to the unknown parameter set $\boldsymbol{\theta}_v$ is nonsingular, and thus the corresponding $\mathbf{H}_1$ is nonsingular. Based on this, $\mathbf{P}_1 = \mathbf{I}_N$. According to (63), we have (38).

For the Kronecker structured CM, note that

$$\mathbf{H}_1 = (\mathbf{R}^T \otimes \mathbf{R})^{-\frac{1}{2}}\left[\frac{\partial \mathrm{vec}\{\mathbf{R}\}}{\partial \boldsymbol{\theta}_p^T}, \frac{\partial \mathrm{vec}\{\mathbf{R}\}}{\partial \boldsymbol{\theta}_s^T}\right]. \tag{64}$$

Due to (64), the $i$th column ($i = 1, \cdots, N_p^2$) of $\mathbf{H}_1$ can be expressed by

$$(\mathbf{R}^T \otimes \mathbf{R})^{-\frac{1}{2}}\frac{\partial \mathrm{vec}\{\mathbf{R}\}}{\partial [\boldsymbol{\theta}_p]_i} = \mathrm{vec}\left\{\left(\mathbf{R}_p^{-\frac{1}{2}}\frac{\partial \mathbf{R}_p}{\partial [\boldsymbol{\theta}_p]_i}\mathbf{R}_p^{-\frac{1}{2}}\right) \otimes \mathbf{I}_{N_s}\right\}. \tag{65}$$

Also, $\mathbf{i}_N$ can be rewritten as

$$\mathbf{i}_N = \mathrm{vec}\{\mathbf{I}_{N_p} \otimes \mathbf{I}_{N_s}\}. \tag{66}$$

Note that $\frac{\partial \mathrm{vec}\{\mathbf{R}_p\}}{\partial \boldsymbol{\theta}_p^T}$ is nonsingular, because $\frac{\partial \mathrm{vec}\{\mathbf{R}_p\}}{\partial \boldsymbol{\theta}_p^T}$ corresponds to the unknown parameters of the Hermitian matrix $\mathbf{R}_p$. Therefore, $(\mathbf{R}_p^T \otimes \mathbf{R}_p)^{-\frac{1}{2}}\frac{\partial \mathrm{vec}\{\mathbf{R}_p\}}{\partial \boldsymbol{\theta}_p^T}$ is also nonsingular. As a result, $\mathbf{i}_{N_p}$ can be a linear combination of the columns of $(\mathbf{R}_p^T \otimes \mathbf{R}_p)^{-\frac{1}{2}}\frac{\partial \mathrm{vec}\{\mathbf{R}_p\}}{\partial \boldsymbol{\theta}_p^T}$, which means $\mathbf{I}_{N_p}$ can be expressed as a linear combination of the matrices $\mathbf{R}_p^{-\frac{1}{2}}\frac{\partial \mathbf{R}_p}{\partial [\boldsymbol{\theta}_p]_i}\mathbf{R}_p^{-\frac{1}{2}}, i = 1, \cdots, N_p^2$. Based on (65) and the fact that $a\mathbf{A}_1 \otimes \mathbf{B} + b\mathbf{A}_2 \otimes \mathbf{B} = (a\mathbf{A}_1 + b\mathbf{A}_2) \otimes \mathbf{B}$, $\mathbf{i}_N$ can be expressed linearly by the columns of $\mathbf{H}_1$, and thus we have

$$\mathbf{P}_1\mathbf{i}_N = \mathbf{i}_N. \tag{67}$$

According to (63) and (67), we have (37). ■

### C. Proof of Proposition 3

*Proof:* We first validate the CFAR property of the KMLE-ASMF detector. That means the detection statistic is independent of $\mathbf{R}_p$ and $\mathbf{R}_s$.

Define $\boldsymbol{\mathcal{A}} \triangleq \mathbf{R}^{-\frac{1}{2}}\mathbf{A}$, $\boldsymbol{\mathcal{A}}_1 \triangleq \boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{A}}^H\boldsymbol{\mathcal{A}})^{-\frac{1}{2}} = \mathbf{I}_{N_p} \otimes (\mathbf{a}^H\mathbf{R}_s^{-1}\mathbf{a})^{-\frac{1}{2}}\mathbf{R}_s^{-\frac{1}{2}}\mathbf{a}$ and

$$\hat{\mathbf{M}} \triangleq \mathbf{R}^{-\frac{1}{2}}\hat{\mathbf{R}}_{\text{KMLE}}\mathbf{R}^{-\frac{1}{2}} = \hat{\mathbf{M}}_p \otimes \hat{\mathbf{M}}_s \tag{68}$$

where $\hat{\mathbf{M}}_p = \mathbf{R}_p^{-\frac{1}{2}}\hat{\mathbf{R}}_p\mathbf{R}_p^{-\frac{1}{2}}$, and $\hat{\mathbf{M}}_s = \mathbf{R}_s^{-\frac{1}{2}}\hat{\mathbf{R}}_s\mathbf{R}_s^{-\frac{1}{2}}$.

The detection statistics in (14) can be expressed as

$$T(\mathbf{y}, \hat{\mathbf{R}}_{\text{KMLE}}, \mathbf{A}, v) = T(\mathbf{w}, \hat{\mathbf{M}}, \boldsymbol{\mathcal{A}}_1, v) \tag{69}$$

where $\mathbf{w}|_\tau = (\mathbf{R})^{-\frac{1}{2}}\mathbf{y} \sim \mathcal{CN}(\mathbf{0}, \tau\mathbf{I}_N)$ under hypothesis $\mathcal{H}_0$. Note that the textures $\tau$ and $\tau_k$ are independent of the corresponding speckles.

We further define an $N \times N$ unitary matrix

$$\mathbf{U} \triangleq \mathbf{U}_p \otimes \mathbf{U}_s, \tag{70}$$

such that the first columns of the unitary matrices $\mathbf{U}_p$ and $\mathbf{U}_s$ are $(\boldsymbol{\alpha}^H\mathbf{R}_p^{-1}\boldsymbol{\alpha})^{-\frac{1}{2}}\mathbf{R}_p^{-\frac{1}{2}}\boldsymbol{\alpha}$ and $(\mathbf{a}^H\mathbf{R}_s^{-1}\mathbf{a})^{-\frac{1}{2}}\mathbf{R}_s^{-\frac{1}{2}}\mathbf{a}$ respectively. Eq. (69) may be rewritten as

$$T(\mathbf{y}, \hat{\mathbf{R}}_{\text{KMLE}}, \mathbf{A}, v) = T(\mathbf{q}, \tilde{\mathbf{M}}, \boldsymbol{\mathcal{A}}_2, v) \tag{71}$$

where $\boldsymbol{\mathcal{A}}_2 \triangleq \mathbf{U}^H\boldsymbol{\mathcal{A}}_1 = \mathbf{U}_p^H \otimes \ \mathbf{e}_1,\ \mathbf{e}_1 \in \mathbb{R}^{N_s \times 1},\ \mathbf{q} \triangleq \mathbf{U}^H\mathbf{w}$ with $\mathbf{q}|\tau \sim \mathcal{CN}(\mathbf{0}, \tau\mathbf{I}_N)$, and

$$\tilde{\mathbf{M}} \triangleq \mathbf{U}^H\hat{\mathbf{M}}\mathbf{U} = (\mathbf{U}_p^H\hat{\mathbf{M}}_p\mathbf{U}_p) \otimes (\mathbf{U}_s^H\hat{\mathbf{M}}_s\mathbf{U}_s) \tag{72}$$

Thus, according to the invariance property of KMLE, $\tilde{\mathbf{M}}$ is KMLE of the identity matrix $\mathbf{I}_N$. Thus, $\tilde{\mathbf{M}}$ is independent of $\mathbf{R}$.

Further, with the fact that $\mathbf{U}_p^H \otimes \mathbf{e}_1 = (\mathbf{I}_{N_p} \otimes \mathbf{e}_1)\mathbf{U}_p^H$, it is easy to obtain that

$$T(\mathbf{q}, \tilde{\mathbf{M}}, \boldsymbol{\mathcal{A}}_2, v) = T(\mathbf{q}, \tilde{\mathbf{M}}, \boldsymbol{\mathcal{Q}}, v) \tag{73}$$

where

$$\boldsymbol{\mathcal{Q}} = \mathbf{I}_{N_p} \otimes \mathbf{e}_1. \tag{74}$$

Now, all the components of the detection statistic of KMLE-ASMF are independent of $\mathbf{R}_p$ and $\mathbf{R}_s$. Therefore, the CFAR property with respect to the Kronecker structured CM $\mathbf{R}$ can be guaranteed for the KMLE-ASMF detector.

The second part of Proposition 3 is trivial. CFAR with respect to the Kronecker structured CM of KNSCM-ASMF cannot be guaranteed, because KNSCM is not an invariant estimator under group $\mathcal{G}_\kappa$. Note that $\mathbf{Y}_k = \sqrt{\tau_k}\mathbf{R}_s^{\frac{1}{2}}\mathbf{Z}_k\mathbf{R}_p^{\frac{T}{2}}$ where $\text{vec}\{\mathbf{Z}_k\} = (\tau_k\mathbf{R}_p \otimes \mathbf{R}_s)^{-\frac{1}{2}}\mathbf{y}_k \sim \mathcal{CN}(0, \mathbf{I}_N)$ conditioned on $\tau_k$. It is readily verified that both $\mathbf{R}_s^{-\frac{1}{2}}\mathbf{Y}_k\mathbf{Y}_k^H\mathbf{R}_s^{-\frac{1}{2}} = \tau_k\mathbf{Z}_k\mathbf{R}_p^T\mathbf{Z}_k^H$ and $\mathbf{R}_p^{-\frac{1}{2}}\mathbf{Y}_k^T\mathbf{Y}_k^*\mathbf{R}_p^{-\frac{1}{2}} = \tau_k\mathbf{Z}_k^T\mathbf{R}_s^T\mathbf{Z}_k^*$ are related to $\mathbf{R}_p$ and $\mathbf{R}_s$. ■

### D. Proof of Theorem 1

*Proof:* Let us express the SCRL in (42) compactly as

$$\rho(\hat{\mathbf{R}}) \triangleq \frac{w}{u}, \tag{75}$$

with $w \triangleq |\mathbf{t}^H\hat{\mathbf{R}}^{-1}\mathbf{t}|^2$, and $u \triangleq (\mathbf{t}^H\mathbf{R}^{-1}\mathbf{t})(\mathbf{t}^H\hat{\mathbf{R}}^{-1}\mathbf{R}\hat{\mathbf{R}}^{-1}\mathbf{t})$. For large $K$, the Taylor expansion of $\rho(\hat{\mathbf{R}})$ up to the second-order with respect to $\hat{\mathbf{R}}$ at $\hat{\mathbf{R}} = \mathbf{R}$ is given by

$$\rho(\hat{\mathbf{R}}) \approx 1 + \dot{\boldsymbol{\rho}}^T\delta\boldsymbol{\theta} + \frac{1}{2}\delta\boldsymbol{\theta}^T\ddot{\boldsymbol{\rho}}\delta\boldsymbol{\theta}, \tag{76}$$

where $\delta\boldsymbol{\theta} = \hat{\boldsymbol{\theta}} - \boldsymbol{\theta}$, $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{\theta}$ are respectively the $r$-dimensional unknown real parameters in $\hat{\mathbf{R}}$ and $\mathbf{R}$ same as defined in (32) and

$$\dot{\boldsymbol{\rho}} \triangleq \frac{\partial\rho}{\partial\hat{\boldsymbol{\theta}}}\bigg|_{\hat{\mathbf{R}}=\mathbf{R}}, \ddot{\boldsymbol{\rho}} \triangleq \frac{\partial^2\rho}{\partial\hat{\boldsymbol{\theta}}\partial\hat{\boldsymbol{\theta}}^T}\bigg|_{\hat{\mathbf{R}}=\mathbf{R}} \tag{77}$$

are the first and second derivatives of the SCRL $\rho$ evaluated at $\hat{\mathbf{R}} = \mathbf{R}$, respectively.

Then, we can calculate the entries of the first and second derivatives of $u$ and $w$. Following the differentiation rule of a matrix inverse $d\mathbf{M}^{-1} = -\mathbf{M}^{-1}d\mathbf{M}\mathbf{M}^{-1}$, we have that

$$\frac{\partial w}{\partial[\hat{\boldsymbol{\theta}}]_i} = -2\mathbf{t}^H\hat{\mathbf{R}}^{-1}\mathbf{t}\mathbf{t}^H\hat{\mathbf{R}}^{-1}\dot{\mathbf{R}}_i\hat{\mathbf{R}}^{-1}\mathbf{t}, \tag{78}$$

$$\frac{\partial u}{\partial[\hat{\boldsymbol{\theta}}]_i} = -2\mathbf{t}^H\mathbf{R}^{-1}\mathbf{t}\mathbf{t}^H\hat{\mathbf{R}}^{-1}\dot{\mathbf{R}}_i\hat{\mathbf{R}}^{-1}\mathbf{R}\hat{\mathbf{R}}^{-1}\mathbf{t}, \tag{79}$$

$$\begin{aligned}\frac{\partial^2 w}{\partial[\hat{\boldsymbol{\theta}}]_i\partial[\hat{\boldsymbol{\theta}}]_j} &= 2\mathbf{t}^H\hat{\mathbf{R}}^{-1}\dot{\mathbf{R}}_i\hat{\mathbf{R}}^{-1}\mathbf{t}\mathbf{t}^H\hat{\mathbf{R}}^{-1}\dot{\mathbf{R}}_j\hat{\mathbf{R}}^{-1}\mathbf{t} \\ &\quad + 2\mathbf{t}^H\hat{\mathbf{R}}^{-1}\mathbf{t}\mathbf{t}^H\hat{\mathbf{R}}^{-1}\dot{\mathbf{R}}_i\hat{\mathbf{R}}^{-1}\dot{\mathbf{R}}_j\hat{\mathbf{R}}^{-1}\mathbf{t} \\ &\quad + 2\mathbf{t}^H\hat{\mathbf{R}}^{-1}\mathbf{t}\mathbf{t}^H\hat{\mathbf{R}}^{-1}\dot{\mathbf{R}}_j\hat{\mathbf{R}}^{-1}\dot{\mathbf{R}}_i\hat{\mathbf{R}}^{-1}\mathbf{t} \\ &\quad - 2\mathbf{t}^H\hat{\mathbf{R}}^{-1}\mathbf{t}\mathbf{t}^H\hat{\mathbf{R}}^{-1}\ddot{\mathbf{R}}_{ij}\hat{\mathbf{R}}^{-1}\mathbf{t},\end{aligned} \tag{80}$$

$$\begin{aligned}\frac{\partial^2 u}{\partial[\hat{\boldsymbol{\theta}}]_i\partial[\hat{\boldsymbol{\theta}}]_j} &= \mathbf{t}^H\mathbf{R}^{-1}\mathbf{t} \\ &\quad \cdot\Big\{2\mathbf{t}^H\hat{\mathbf{R}}^{-1}\dot{\mathbf{R}}_j\hat{\mathbf{R}}^{-1}\dot{\mathbf{R}}_i\hat{\mathbf{R}}^{-1}\mathbf{R}\hat{\mathbf{R}}^{-1}\mathbf{t} \\ &\quad + 2\mathbf{t}^H\hat{\mathbf{R}}^{-1}\dot{\mathbf{R}}_i\hat{\mathbf{R}}^{-1}\dot{\mathbf{R}}_j\hat{\mathbf{R}}^{-1}\mathbf{R}\hat{\mathbf{R}}^{-1}\mathbf{t} \\ &\quad + 2\mathbf{t}^H\hat{\mathbf{R}}^{-1}\dot{\mathbf{R}}_i\hat{\mathbf{R}}^{-1}\mathbf{R}\hat{\mathbf{R}}^{-1}\dot{\mathbf{R}}_j\hat{\mathbf{R}}^{-1}\mathbf{t} \\ &\quad - 2\mathbf{t}^H\hat{\mathbf{R}}^{-1}\ddot{\mathbf{R}}_{ij}\hat{\mathbf{R}}^{-1}\mathbf{R}\hat{\mathbf{R}}^{-1}\mathbf{t}\Big\},\end{aligned} \tag{81}$$

where $\dot{\mathbf{R}}_i \triangleq \frac{\partial\hat{\mathbf{R}}}{\partial[\hat{\boldsymbol{\theta}}]_i}$, and $\ddot{\mathbf{R}}_{ij} \triangleq \frac{\partial^2\hat{\mathbf{R}}}{\partial[\hat{\boldsymbol{\theta}}]_i\partial[\hat{\boldsymbol{\theta}}]_j}$.

By the quotient rule,

$$\dot{\rho}_i = \frac{u\dot{w}_i - w\dot{u}_i}{u^2}, \tag{82}$$

$$\ddot{\rho}_{ij} = \frac{u^2\ddot{w}_{ij} - uw\ddot{u}_{ij} - u\dot{u}_i\dot{w}_j + u\dot{w}_i\dot{u}_j - 2u\dot{u}_j\dot{w}_i + 2w\dot{u}_j\dot{u}_i}{u^3}, \tag{83}$$

where $\dot{\rho}_i \triangleq \frac{\partial\rho(\hat{\mathbf{R}})}{\partial[\hat{\boldsymbol{\theta}}]_i}$ and $\ddot{\rho}_{ij} \triangleq \frac{\partial^2\rho(\hat{\mathbf{R}})}{\partial[\hat{\boldsymbol{\theta}}]_i\partial[\hat{\boldsymbol{\theta}}]_j}$ denote the first and second derivative of the SCRL $\rho$, respectively, and the other parameters in (83) are similarly defined.

According to (78) and (79), for the first-order derivatives evaluated at $\hat{\mathbf{R}} = \mathbf{R}$, it holds that

$$\dot{u}_i\big|_{\hat{\mathbf{R}}=\mathbf{R}} = \dot{w}_i\big|_{\hat{\mathbf{R}}=\mathbf{R}} = -2\mathbf{t}^H\mathbf{R}^{-1}\mathbf{t}\mathbf{t}^H\mathbf{R}^{-1}\dot{\mathbf{R}}_i\mathbf{R}^{-1}\mathbf{t}, \tag{84}$$

and $u = w = |\mathbf{t}^H \mathbf{R}^{-1} \mathbf{t}|^2$. With (82) and (84), we have

$$\dot{\boldsymbol{\rho}} = \mathbf{0}, \tag{85}$$

and the $(i,j)$th entry of the second derivative (77) can be rewritten as

$$\begin{aligned} \ddot{\rho}_{ij}|_{\hat{\mathbf{R}}=\mathbf{R}} &= \frac{u\ddot{w}_{ij} - w\ddot{u}_{ij}}{u^2}|_{\hat{\mathbf{R}}=\mathbf{R}} \\ &= \frac{2\mathbf{t}^H \mathbf{R}^{-1} \dot{\mathbf{R}}_i \mathbf{R}^{-1} \mathbf{t}\mathbf{t}^H \mathbf{R}^{-1} \dot{\mathbf{R}}_j \mathbf{R}^{-1} \mathbf{t}}{|\mathbf{t}^H \mathbf{R}^{-1} \mathbf{t}|^2} \\ &\quad - \frac{2\mathbf{t}^H \mathbf{R}^{-1} \dot{\mathbf{R}}_i \mathbf{R}^{-1} \dot{\mathbf{R}}_j \mathbf{R}^{-1} \mathbf{t}}{|\mathbf{t}^H \mathbf{R}^{-1} \mathbf{t}|}. \end{aligned} \tag{86}$$

Making use of the facts that $\mathrm{tr}\{\mathbf{AB}\} = \mathrm{tr}\{\mathbf{BA}\}$ and $\mathrm{tr}\{\mathbf{ABCD}\} = \mathrm{vec}^T\{\mathbf{B}^T\} \left(\mathbf{C}^T \otimes \mathbf{A}\right) \mathrm{vec}\{\mathbf{D}\}$, (86) can be rewritten as

$$\ddot{\rho}_{ij}|_{\hat{\mathbf{R}}=\mathbf{R}} = 2\mathrm{vec}^T\{\dot{\mathbf{R}}_i^T\} \left((\boldsymbol{\Upsilon} - \mathbf{R}^{-1})^T \otimes \boldsymbol{\Upsilon}\right) \mathrm{vec}\{\dot{\mathbf{R}}_j\}, \tag{87}$$

where $\boldsymbol{\Upsilon} \triangleq \frac{\mathbf{R}^{-1}\mathbf{t}\mathbf{t}^H\mathbf{R}^{-1}}{\mathbf{t}^H\mathbf{R}^{-1}\mathbf{t}}$. Using the definition of $\mathbf{H}$ given by (32) in (87), we have that

$$\ddot{\boldsymbol{\rho}} = 2\mathbf{H}^H \left[(\boldsymbol{\Upsilon} - \mathbf{R}^{-1})^T \otimes \boldsymbol{\Upsilon}\right] \mathbf{H}. \tag{88}$$

Also, since $\mathrm{vec}\{\mathbf{R}\} = \mathbf{H}\boldsymbol{\theta}$ (we assume that $\mathbf{R}$ is a linear function of $\boldsymbol{\theta}$ and thus $\mathbf{H}$ is a constant matrix), the MSE defined in (36) can be expressed as the $\boldsymbol{\mathcal{E}} = \mathbf{H}\mathbb{E}\{\delta\boldsymbol{\theta}\delta\boldsymbol{\theta}^T\}\mathbf{H}^H$, so the average SCRL can be expressed as

$$\mathbb{E}\{\rho(\hat{\mathbf{R}})\} \approx 1 + \mathrm{tr}\left\{\left[(\boldsymbol{\Upsilon} - \mathbf{R}^{-1})^T \otimes \boldsymbol{\Upsilon}\right] \boldsymbol{\mathcal{E}}\right\}. \tag{89}$$

Then define a unitary matrix $\mathbf{U}$ such that $\mathbf{U}^H \mathbf{R}^{-\frac{1}{2}} \mathbf{t} = (\mathbf{t}^H \mathbf{R}^{-1} \mathbf{t})^{\frac{1}{2}} \mathbf{e}_1$, and

$$\boldsymbol{\Delta} \triangleq \left(\mathbf{I}_N - \mathbf{e}_1 \mathbf{e}_1^T\right)^T \otimes \left(\mathbf{e}_1 \mathbf{e}_1^T\right). \tag{90}$$

Note that $\mathbf{R}^{-1} = \mathbf{R}^{-\frac{1}{2}} \mathbf{U}\mathbf{U}^H \mathbf{R}^{-\frac{1}{2}}$, and that $\boldsymbol{\Upsilon}$ can be rewritten as

$$\boldsymbol{\Upsilon} = \mathbf{R}^{-\frac{1}{2}} \mathbf{U} \left(\mathbf{e}_1 \mathbf{e}_1^T\right) \mathbf{U}^H \mathbf{R}^{-\frac{1}{2}}. \tag{91}$$

Thus, we have

$$\mathrm{tr}\left\{\left[(\boldsymbol{\Upsilon} - \mathbf{R}^{-1})^T \otimes \boldsymbol{\Upsilon}\right] \boldsymbol{\mathcal{E}}\right\} = -\mathrm{tr}\left\{\boldsymbol{\Delta}\bar{\boldsymbol{\mathcal{E}}}\right\} \tag{92}$$

where

$$\bar{\boldsymbol{\mathcal{E}}} = \left((\mathbf{R}^{-\frac{1}{2}}\mathbf{U})^T \otimes \mathbf{U}^H \mathbf{R}^{-\frac{1}{2}}\right) \boldsymbol{\mathcal{E}} \left((\mathbf{R}^{-\frac{1}{2}}\mathbf{U})^T \otimes \mathbf{U}^H \mathbf{R}^{-\frac{1}{2}}\right)^H. \tag{93}$$

The theorem follows by substituting (92) into (89). ■

### *E. Proof of Corollary 1*

*Proof:* This proof is followed by substituting the CRB matrix, given in (37), into Theorem 1. According to (44), we just need to calculate the average SCRL when $\mathbf{R} = \mathbf{I}_N$. Notice that $\mathrm{tr}\{\boldsymbol{\Delta}\mathrm{vec}\{\mathbf{I}_N\}\mathrm{vec}^H\{\mathbf{I}_N\}\} = 0$, the average SCRL can be rewritten as

$$\mathbb{E}\{\rho(\hat{\mathbf{R}}_{\mathrm{KMLE}})\} \approx 1 - \frac{c_1 \mathrm{tr}\{\boldsymbol{\Delta}\tilde{\mathbf{P}}_\kappa\}}{K}. \tag{94}$$

Note that $\tilde{\mathbf{P}}_\kappa$ is a projection matrix of the column subspace of the matrix $\tilde{\mathbf{H}}_\kappa$, where $\tilde{\mathbf{H}}_\kappa = \mathbf{H}_\kappa|_{\mathbf{R}=\mathbf{I}_N}$. First, we make a linear transformation on the columns of $\tilde{\mathbf{H}}_\kappa$, i.e., $\mathbf{G} = \tilde{\mathbf{H}}_\kappa \mathbf{J}$, where $\mathbf{J}$ is a nonsingular matrix such that

$$\begin{aligned} \mathbf{G} = &\left[\mathbf{E}_{p11}, \cdots, \mathbf{E}_{pN_pN_p}, \mathbf{E}_{s11}, \cdots, \mathbf{E}_{sN_sN_s}\right] \\ &\in \mathbb{R}^{(N_p^2+N_s^2)\times N^2}, \end{aligned} \tag{95}$$

where $\mathbf{E}_{pij} = \mathrm{vec}\{\mathbf{e}_i \mathbf{e}_j^T \otimes \mathbf{I}_{N_s}\} \in \mathbb{R}^{N^2\times 1}$ and $\mathbf{E}_{skl} = \mathrm{vec}\{\mathbf{I}_{N_p} \otimes \mathbf{e}_k \mathbf{e}_l^T\} \in \mathbb{R}^{N^2\times 1}, i,j = 1,\cdots,N_p, k,l = 1,\cdots,N_s$. By this transformation, we convert the unknown real-valued parameter $\boldsymbol{\theta}_\kappa \in \mathbb{R}^{(N_p^2+N_s^2)\times 1}$ into the $N_p^2$ complex-valued entries $[\mathbf{R}_p]_{ij}$ and $N_s^2$ complex entries $[\mathbf{R}_s]_{ij}$ of the Hermitian matrices $\mathbf{R}_p$ and $\mathbf{R}_s$ [51], [55]. Note that such column transformation will not change the projection matrix, so it holds that

$$\tilde{\mathbf{P}}_\kappa = \tilde{\mathbf{H}}_\kappa (\tilde{\mathbf{H}}_\kappa^H \tilde{\mathbf{H}}_\kappa)^\dagger \tilde{\mathbf{H}}_\kappa^H = \mathbf{G}(\mathbf{G}^H\mathbf{G})^\dagger \mathbf{G}^H. \tag{96}$$

Due to (95), we have

$$\mathbf{G}^H\mathbf{G} = \begin{bmatrix} N_s\mathbf{I}_{N_p^2} & \mathbf{i}_p\mathbf{i}_s^T \\ \mathbf{i}_s\mathbf{i}_p^T & N_p\mathbf{I}_{N_s^2} \end{bmatrix} \tag{97}$$

where $\mathbf{i}_p \triangleq \mathrm{vec}\{\mathbf{I}_{N_p}\}$ and $\mathbf{i}_s \triangleq \mathrm{vec}\{\mathbf{I}_{N_s}\}$.

It can be verified that the generalized inverse of $\mathbf{G}^H\mathbf{G}$ is

$$\begin{aligned} &(\mathbf{G}^H\mathbf{G})^\dagger \\ &= \begin{bmatrix} \frac{1}{N_s}\mathbf{I}_{N_p^2} - \frac{N_p+2N_s}{N_s(N_s+N_p)^2}\mathbf{i}_p\mathbf{i}_p^T & -\frac{\mathbf{i}_p\mathbf{i}_s^T}{(N_p+N_s)^2} \\ -\frac{\mathbf{i}_s\mathbf{i}_p^T}{(N_p+N_s)^2} & \frac{1}{N_p}\mathbf{I}_{N_s^2} - \frac{N_s+2N_p}{N_p(N_p+N_s)^2}\mathbf{i}_s\mathbf{i}_s^T \end{bmatrix}. \end{aligned} \tag{98}$$

According to (95), the entries of $\mathbf{G}$ are "0"s or "1"s, so the entries of $\tilde{\mathbf{P}}_\kappa$ can be obtained by the sum of the corresponding entries of $(\mathbf{G}^H\mathbf{G})^\dagger$ given by (98). Note that the $n$th entry of $\mathrm{vec}\{\mathbf{R}\}$ has the form of $[\mathrm{vec}\{\mathbf{R}\}]_n = [\mathbf{R}_p]_{ij}[\mathbf{R}_s]_{kl}$, with $n = (i-1)N_s + (j-1)NN_s + k + (l-1)N$. According to (95) and (98), the diagonal entries of $\tilde{\mathbf{P}}_\kappa$ are given by

$$\left[\tilde{\mathbf{P}}_\kappa\right]_{nn} = \begin{cases} \dfrac{N_p+N_s-1}{N} & i=j, k=l \\ \dfrac{1}{N_p} & i=j, k\neq l \\ \dfrac{1}{N_s} & i\neq j, k=l \\ 0 & i\neq j, k\neq l \end{cases}. \tag{99}$$

According to (94) and (99), Corollary 1 follows. ■

### *F. Derivation of (55)*

*Proof:* Following the differentiation rules of the matrix inversion $d\mathbf{M}^{-1} = -\mathbf{M}^{-1}d\mathbf{M}\mathbf{M}^{-1}$ and $d\mathbf{y}^H\mathbf{M}^{-1}\mathbf{y} =$

$\text{vec}^H\{-\mathbf{M}^{-1}\mathbf{y}\mathbf{y}^H\mathbf{M}^{-1}\}d\text{vec}(\mathbf{M})$, we have

$$\begin{aligned}&\frac{\partial \mathbf{q}^H\tilde{\mathbf{M}}^{-1}\boldsymbol{\mathcal{Q}}(\boldsymbol{\mathcal{Q}}^H\tilde{\mathbf{M}}^{-1}\boldsymbol{\mathcal{Q}})^{-1}\boldsymbol{\mathcal{Q}}^H\tilde{\mathbf{M}}^{-1}\mathbf{q}}{\partial \text{vec}(\tilde{\mathbf{M}}^*)}\\ &= \text{vec}\{-\tilde{\mathbf{M}}^{-1}\boldsymbol{\mathcal{Q}}(\boldsymbol{\mathcal{Q}}^H\tilde{\mathbf{M}}^{-1}\boldsymbol{\mathcal{Q}})^{-1}\boldsymbol{\mathcal{Q}}^H\tilde{\mathbf{M}}^{-1}\mathbf{q}\mathbf{q}^H\tilde{\mathbf{M}}^{-1}\\ &\quad-\tilde{\mathbf{M}}^{-1}\mathbf{q}\mathbf{q}^H\tilde{\mathbf{M}}^{-1}\boldsymbol{\mathcal{Q}}(\boldsymbol{\mathcal{Q}}^H\tilde{\mathbf{M}}^{-1}\boldsymbol{\mathcal{Q}})^{-1}\boldsymbol{\mathcal{Q}}^H\tilde{\mathbf{M}}^{-1}\\ &\quad+\tilde{\mathbf{M}}^{-1}\boldsymbol{\mathcal{Q}}(\boldsymbol{\mathcal{Q}}^H\tilde{\mathbf{M}}^{-1}\boldsymbol{\mathcal{Q}})^{-1}\boldsymbol{\mathcal{Q}}^H\tilde{\mathbf{M}}^{-1}\mathbf{q}\\ &\quad\cdot\mathbf{q}^H\tilde{\mathbf{M}}^{-1}\boldsymbol{\mathcal{Q}}(\boldsymbol{\mathcal{Q}}^H\tilde{\mathbf{M}}^{-1}\boldsymbol{\mathcal{Q}})^{-1}\boldsymbol{\mathcal{Q}}^H\tilde{\mathbf{M}}^{-1}\}. \qquad (100)\end{aligned}$$

Therefore, using (100), we have

$$\begin{aligned}&\frac{\partial \mathbf{q}^H\tilde{\mathbf{M}}^{-1}\boldsymbol{\mathcal{Q}}(\boldsymbol{\mathcal{Q}}^H\tilde{\mathbf{M}}^{-1}\boldsymbol{\mathcal{Q}})^{-1}\boldsymbol{\mathcal{Q}}^H\tilde{\mathbf{M}}^{-1}\mathbf{q}}{\partial \text{vec}(\tilde{\mathbf{M}}^*)}\Big|_{\tilde{\mathbf{M}}=\mathbf{I}}\\ &= \text{vec}\{-\boldsymbol{\mathcal{Q}}\mathbf{q}\mathbf{q}^H-\mathbf{q}\mathbf{q}^H\boldsymbol{\mathcal{Q}}^H+\boldsymbol{\mathcal{Q}}\mathbf{q}\mathbf{q}^H\boldsymbol{\mathcal{Q}}^H\} \qquad (101)\end{aligned}$$

and

$$\frac{\partial \mathbf{q}^H\tilde{\mathbf{M}}^{-1}\mathbf{q}}{\partial \text{vec}(\tilde{\mathbf{M}}^*)}\Big|_{\tilde{\mathbf{M}}=\mathbf{I}} = \text{vec}\{-\mathbf{q}\mathbf{q}^H\}. \qquad (102)$$

Then in light of (82), (101) and (102), we have (55). ■

## REFERENCES


[1] D. Giuli, "Polarization diversity in radars," *Proc. IEEE*, vol. 74, no. 2, pp. 245–269, Feb. 1986.

[2] M. Hurtado and A. Nehorai, "Polarimetric detection of targets in heavy inhomogeneous clutter," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1349–1361, Apr. 2008.

[3] A. Nehorai and P. Eytan, "Vector-sensor array processing for electromagnetic source localization," *IEEE Trans. Signal Process.*, vol. 42, no. 2, pp. 376–398, Feb. 1994.

[4] C. Hao, S. Gazor, X. Ma, S. Yan, C. Hou, and D. Orlando, "Polarimetric detection and range estimation of a point-like target," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 52, no. 2, pp. 603–616, Apr. 2016.

[5] M. Hurtado, T. Zhao, and A. Nehorai, "Adaptive polarized waveform design for target tracking based on sequential Bayesian inference," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1120–1133, Mar. 2008.

[6] M. Hurtado, J. Xiao, and A. Nehorai, "Target estimation, detection, and tracking," *IEEE Signal Process. Mag.*, vol. 26, no. 1, pp. 42–52, Jan. 2009.

[7] I. Reed, J. Mallett, and L. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-10, no. 6, pp. 853–863, Nov. 1974.

[8] J. Liu, Z. Zhang, Y. Yang, and H. Liu, "A CFAR adaptive subspace detector for first-order or second-order Gaussian signals based on a single observation," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5126–5140, Nov. 2011.

[9] J. Liu, Z. Zhang, and Y. Yang, "Optimal waveform design for generalized likelihood ratio and adaptive matched filter detectors using a diversely polarized antenna," *Signal Process.*, vol. 92, no. 4, pp. 1126–1131, Apr. 2012.

[10] J. Liu, Z. Zhang, and Y. Yang, "Performance enhancement of subspace detection with a diversely polarized antenna," *IEEE Signal Process. Lett.*, vol. 19, no. 1, pp. 4–7, Jan. 2012.

[11] J. Liu, W. Liu, B. Chen, H. Liu, H. Li, and C. Hao, "Modified Rao test for multichannel adaptive signal detection," *IEEE Trans. Signal Process.*, vol. 64, no. 3, pp. 714–725, Feb. 2016.

[12] G. Cui, L. Kong, X. Yang, and J. Yang, "Distributed target detection with polarimetric MIMO radar in compound-Gaussian clutter," *Digit. Signal Process.*, vol. 22, no. 3, pp. 430–438, May 2012.

[13] G. Alfano, A. De Maio, and E. Conte, "Polarization diversity detection of distributed targets in compound-Gaussian Clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 40, no. 2, pp. 755–765, May 2004.

[14] L. Novak, M. Burl, and W. Irving, "Optimal polarimetric processing for enhanced target detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, no. 1, pp. 234–244, Jan. 1993.

[15] J. Kermoal, L. Schumacher, K. Pedersen, P. Mogensen and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 6, pp. 1211–1226, Aug. 2002.

[16] K. Yu, M. Bengtsson, B. Ottersten, D. McNamara, P. Karlsson, and M. Beach, "Modeling of wideband MIMO radio channels based on NLoS indoor measurements," *IEEE Trans. Veh. Technol.*, vol. 53, no. 8, pp. 655–665, May 2004.

[17] S. Zhou, H. Liu, B. Liu, and K. Yin, "Adaptive MIMO radar target parameter estimation with Kronecker-product structured interference covariance matrix," *Signal Process.*, vol. 92, no. 5, pp. 1177–1188, Jul. 2012.

[18] J. Munck, H. Huizenga, L. Waldorp, and R. Heethaar, "Estimating stationary dipoles from MEG/EEG data contaminated with spatially and temporally correlated background noise," *IEEE Trans. Signal Process.*, vol. 50, no. 7, pp. 1565–1572, Jul. 2002.

[19] N. Lu and D. Zimmerman, "The likelihood ratio test for a separable covariance matrix," *Statist. Probab. Lett.*, vol. 73, no. 5, pp. 449–457, May 2005.

[20] K. Werner, M. Jansson, and P. Stoica, "On estimation of covariance matrices with Kronecker product structure," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 478–491, Feb. 2008.

[21] Y. Sun, P. Vabu, and D. Palomar, "Robust estimation of structured covariance matrix for heavy-tailed elliptical distributions," *IEEE Trans. Signal Process.*, vol. 64, no. 14, pp. 3576–3590, Jul. 2015.

[22] M. Castaneda and J. Nossek, "Estimation of rank deficient covariance matrices with Kronecker Structure," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2014, pp. 394–398.

[23] E. Kelly, "An adaptive detection algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-22, no. 2, pp. 115–127, Mar. 1986.

[24] J. Liu, W. Liu, H. Liu, B. Chen, X. Xia, and F. Dai, "Average SINR calculation of a persymmetric sample matrix inversion beamformer," *IEEE Trans. Signal Process.*, vol. 64, no. 8, pp. 2135–2145, Apr. 2016.

[25] F. Pascal, P. Forster, J. Ovarlez, and F. Pascal, "Persymmetric adaptive radar detectors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, no. 4, pp. 2376–2390, Oct. 2011.

[26] A. De Maio, D. Orlando, C. Hao, and G. Foglia, "Adaptive detection of point-like targets in spectrally symmetric interference," *IEEE Trans. Signal Process.*, vol. 64, no. 12, pp. 3207–3220, Mar. 2016.

[27] C. Hao, D. Orlando, G. Foglia, and G. Giunta, "Knowledge-based adaptive detection: Joint exploitation of clutter and system symmetry properties," *IEEE Signal Process. Lett.*, vol. 23, no. 10, pp. 1489–1493, Oct. 2016.

[28] E. Lehmann and J. Romano, *Testing Statistical Hypotheses (Springer Texts in Statistics)*, 3rd ed. New York, NY, USA: Springer, 2006.

[29] C. Curtis and I. Reiner, Representation theory of finite groups and associative algebras. Providence, RI, USA: Amer. Math. Soc., 1962.

[30] I. Soloveychik, D. Trushin, and A. Wiesel, "Group symmetric robust covariance estimation," *IEEE Trans. Signal Process.*, vol. 64, no. 1, pp. 244–257, Jan. 2016.

[31] A. De Maio, D. Orlando, I, Soloveychik, and A. Wiesel, "Invariance theory for adaptive detection in interference with group symmetric covariance matrix," *IEEE Trans. Signal Process.*, vol. 64, no. 23, pp. 6299–6312, Dec. 2016.

[32] A. Balleri, A. Nehorai, and J. Wang, "Maximum likelihood estimation for compound-Gaussian clutter with inverse gamma texture," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 2, pp. 775–780, Apr. 2007.

[33] Y. Gao, G. Liao, S. Zhu, and D. Yang, "A persymmetric GLRT for adaptive detection in compound-Gaussian clutter with random texture," *IEEE Signal Process. Lett.*, vol. 20, no. 6, pp. 615–618, Jun. 2013.

[34] E. Lehmann and G. Casella, *Theory of Point Estimation* (Springer Texts in Statistics), 2nd ed. New York, NY, USA: Springer, 1998.

[35] M. Jansson, P. Wirfalt, K. Werner, and B. Ottersten, "ML estimation of covariance matrices with Kronecker and persymmetric structure," in *Proc. IEEE 13th Digit. Signal Process. Workshop/5th IEEE Signal Process. Educ. Workshop*, 2009, pp. 298–301.

[36] F. Gini and M. Greco, "Covariance matrix estimation for CFAR detection in correlated heavy tailed clutter," *Signal Process.*, vol. 82, no. 12, pp. 1847–1859, Dec. 2002.

[37] J. Wang, A. Dogandzic, and A. Nehorai, "Maximum likelihood estimation of compound-Gaussian clutter and target parameters," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 3884–3898, Jan. 2006.

[38] K. Sangston, F. Gini, and M. Greco, "Coherent radar target detection in heavy-tailed compound-Gaussian clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 1, pp. 64–77, Jan. 2012.

[39] H. Park, J. Li, and H. Wang, "Polarization-space-time domain generalized likelihood ratio detection of radar targets," *Signal Process.*, vol. 41, no. 2, pp. 153–164, Jan. 1995.

[40] J. Ward, "Space-time adaptive processing for airborne radar," *Lincoln Lab., Mass. Inst. Technol.*, Lincoln, MA, USA, Tech. Rep. ESC-TR-94-109, Dec. 1994.
[41] P. Stoica and T. Marzetta, "Parameter estimation problems with singular information matrices," *IEEE Trans. Signal Process.*, vol. 49, no. 1, pp. 87–90, Jan. 2001.
[42] D. Romero and R. López-Valcarce, "Nearly-optimal compression matrices for signal power estimation," in *Proc. IEEE 15th Int. Workshop Signal Process. Adv. Wireless Commun.*, 2014, pp. 434–438.
[43] S. T. Smith, "Covariance, subspace, and intrinsic Cramér–Rao bounds," *IEEE Trans. Signal Process.*, vol. 53, no. 5, pp. 1610–1630, May 2005.
[44] O. Besson, S. Bidon, and J. Tourneret, "Bounds for estimation of covariance matrices from heterogeneous samples," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3357–3362, Jul. 2008.
[45] F. Pascal, P. Forster, J. Ovarlez, and P. Larzabal, "Performance analysis of covariance matrix estimates in impulsive noise," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2206–2217, Jun. 2008.
[46] S. Bausson, F. Pascal, P. Forster, J. Ovarlez, and P. Larzabal, "First-and second-order moments of the normalized sample covariance matrix of spherically invariant random vectors," *IEEE Signal Process. Lett.*, vol. 14, no. 6, pp. 425–428, Jun. 2007.
[47] A. Wiesel, "Geodesic convexity and covariance estimation," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6182–6189, Dec. 2012.
[48] S. Kraut, L. Scharf, and R. Butler, "The adaptive coherence estimator: A uniformly most powerful invariant adaptive detection statistic," *IEEE Trans. Signal Process.*, vol. 53, no. 2, pp. 427–438, Feb. 2005.
[49] P. Brandt and M. Pedersen, "The matrix cookbook," Tech. Univ. Denmark, Lyngby, Denmark, Nov. 2008.
[50] M. Bilodeau and D. Brenner, *Theory of Multivariate Statistics.* New York, NY, USA: Springer, 1999.
[51] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory.* Englewood Cliffs, NJ, USA: PTR Prentice-Hall, 1993.
[52] E. Conte, A. De Maio, and G. Ricci, "Adaptive CFAR detection in compound-Gaussian clutter with circulant covariance matrix," *IEEE Signal Process. Lett.*, vol. 7, no. 3, pp. 63–65, Mar. 2000.
[53] Y. Wang, W. Xia, Z. He, and Y. Sun, "Polarimetric detection for vector-sensor processing in quaternion proper Gaussian noises," *Multidimensional Syst. Signal Process.*, vol. 27, no. 2, pp. 597–618, Apr. 2016.
[54] J. Gorman and A. Hero, "Lower bounds for parametric estimation with constraints," *IEEE Trans. Inf. Theory*, vol. 36, no. 6, pp. 1285–1301, Nov. 1990.
[55] E. Ollila, V. Koivunen, and J. Eriksson, "On the Cramér–Rao bound for the constrained and unconstrained complex parameter," in *Proc. 5th IEEE Sensor Array Multichannel Signal Process. Workshop*, 2008, pp. 414–418.

**Yikai Wang** (S'16) received the B.S. degree in 2011 from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, where he is currently working toward the Ph.D. degree in signal and information processing. From September 2015 to March 2017, he was a Visiting Student with the Department of Electrical and Computer Engineering, Rutgers, The State University of New Jersey. His current research interests include machine learning, radar signal processing, and adaptive signal detection algorithm.

**Wei Xia** received the B.S. degree in communication engineering and the M.S. and Ph.D. degrees in electronic engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2002, 2005, and 2008, respectively. Since 2009, he has been with the School of Electronic Engineering, UESTC, where he is currently an Associate Professor. From March 2015 to March 2016, he was on a 12-month sabbatical leave as a Visiting Scholar at the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, New Jersey, USA. His general research interests include statistical signal processing, adaptive signal processing, radar signal processing, and advanced implementation techniques of signal processing algorithm.

**Zishu He** received the B.S., M.S., and Ph.D. degrees in signal and information processing from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 1984, 1988, and 2000 respectively. He is currently a Professor of signal and information processing in the School of Electronic Engineering, UESTC. His current research interests include array signal processing, digital beam forming, the theory on MIMO communication and MIMO radar, adaptive signal processing, and interference cancellation.

**Hongbin Li** (M'99–SM'08) received the B.S. and M.S. degrees from the University of Electronic Science and Technology of China, Chengdu, China, in 1991 and 1994, respectively, and the Ph.D. degree from the University of Florida, Gainesville, FL, USA, in 1999, all in electrical engineering.

From July 1996 to May 1999, he was a Research Assistant in the Department of Electrical and Computer Engineering, University of Florida. Since July 1999, he has been with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ, USA, where he became a Professor in 2010. He was a Summer Visiting Faculty Member at the Air Force Research Laboratory in the summers of 2003, 2004, and 2009. His general research interests include statistical signal processing, wireless communications, and radars.

Dr. Li received the IEEE Jack Neubauer Memorial Award in 2013 from the IEEE Vehicular Technology Society, the Outstanding Paper Award from the IEEE AFICON Conference in 2011, the Harvey N. Davis Teaching Award in 2003, and the Jess H. Davis Memorial Award for excellence in research in 2001 from the Stevens Institute of Technology, and the Sigma Xi Graduate Research Award from the University of Florida in 1999. He is a member of the IEEE SPS Signal Processing Theory and Methods Technical Committee (TC) and the IEEE SPS Sensor Array and Multichannel TC, an Associate Editor for *Signal Processing* (Elsevier), IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE SIGNAL PROCESSING LETTERS, and IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, as well as a Guest Editor for IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING and *EURASIP Journal on Applied Signal Processing*. He has been involved in various conference organization activities, including serving as a General Co-Chair for the 7th IEEE Sensor Array and Multichannel Signal Processing Workshop, Hoboken, NJ, USA, June 17–20, 2012. He is a member of Tau Beta Pi and Phi Kappa Phi.

**Athina P. Petropulu** (M'87–SM'01–F'08) received the undergraduate degree from the National Technical University of Athens, Athens, Greece, and the M.Sc. and Ph.D. degrees from the Northeastern University, Boston, MA, USA, all in electrical and computer engineering.

Since 2010, she has been a Professor in the Department of Electrical and Computer Engineering, Rutgers, The State University of New Jersey, New Brunswick, NJ, USA, where she was also the Chair of the department during 2010–2016. Before that she was a faculty at Drexel University. She held Visiting Scholar appointments at SUPELEC, Universite Paris Sud, Princeton University, and the University of Southern California. Her research interests include the area of statistical signal processing, wireless communications, signal processing in networking, physical layer security, and radar signal processing. Her research has been funded by various government industry sponsors including the National Science Foundation, the Office of Naval research, the US Army, the National Institute of Health, the Whitaker Foundation, Lockheed Martin.

Dr. Petropulu received the 1995 Presidential Faculty Fellow Award given by NSF and the White House. She was the Editor-in-Chief of the IEEE TRANSACTIONS ON SIGNAL PROCESSING (2009–2011), IEEE Signal Processing Society Vice President—Conferences (2006–2008), and member-at-large of the IEEE Signal Processing Board of Governors. She was the General Chair of the 2005 International Conference on Acoustics Speech and Signal Processing, Philadelphia, PA, USA. In 2005, she received the IEEE SIGNAL PROCESSING MAGAZINE Best Paper Award, and in 2012 the IEEE Signal Processing Society Meritorious Service Award for "exemplary service in technical leadership capacities." She was an IEEE Distinguished Lecturer for the Signal Processing Society for 2017–2018.