Covariance Matrix Estimation of Texture Correlated Compound-Gaussian Vectors for Adaptive Radar Detection

ÇAĞATAY CANDAN
Middle East Technical University (METU), Ankara, Turkey

FRÉDÉRIC PASCAL, Senior Member, IEEE
Universitè Paris-Saclay, Gif-sur-Yvette, France

In this article, covariance matrix estimation of compound-Gaussian vectors with texture-correlation (spatial correlation for the adaptive radar detectors) is examined. The texture parameters are treated as hidden random parameters whose statistical description is given by a Markov chain. States of the chain represent the value of texture coefficient and the transition probabilities establish the correlation in the texture sequence. An expectation–maximization (EM) method-based covariance matrix estimation solution is given for both noiseless and noisy snapshots. An extension to the practically important case of persymmetric covariance matrices is developed and possible extensions to other structured covariance matrices are described. The numerical results indicate that the benefit of utilizing spatial correlation in the covariance matrix estimation can be significant especially when the total number of snapshots in the secondary data is small. From applications viewpoint, the suggested model is well suited for the adaptive target detection in sea clutter, where some spatial correlation between range cells has been experimentally observed. The performance improvements of the suggested approach for small number of snapshots can be particularly important in this application area.

Manuscript received 23 February 2022; revised 23 June 2022 and 29 October 2022; accepted 4 November 2022. Date of publication 10 November 2022; date of current version 9 June 2023.

DOI. No. 10.1109/TAES.2022.3221385

Refereeing of this contribution was handled by F. Gini.

Authors' addresses: Çağatay Candan is with the Department of Electrical and Electronics Engineering, Middle East Technical University (METU), TR06800 Ankara, Turkey, E-mail: (ccandan@metu.edu.tr). Frédéric Pascal is with CNRS, CentralSupélec, Laboratoire des signaux et systémes (UMR8506), Université Paris-Saclay, 91190 Gif-sur-Yvette, France, E-mail: (frederic.pascal@l2s.centralesupelec.fr). *(Corresponding author: Çağatay Candan.)*

I. INTRODUCTION

The problem of accurate and reliable (outlier robust) covariance matrix estimation from a limited number of snapshots is critically important in many applications. The problem is typically treated for jointly Gaussian multivariates (Gaussian vectors); but in some applications, such as the clutter modeling, the total power of a snapshot vector can fluctuate significantly from snapshot-to-snapshot leading to deviations from the standard model. For such applications the outlier robust estimators from statistics literature, such as the Tyler's estimator [1], normalized sample covariance matrix [2], and Huber's M-estimator [3], [4], have been studied in depth to include some application specific needs in the design, such as sample deficiency, low-rank constraint, known matrix structure, knowledge-aided operation [5], [6], [7], [8], [9], [10]. The common assumption in these studies is the statistical independence of snapshot vectors. In this study, we examine the case of correlated snapshots for the covariance matrix estimation and present a versatile solution for noisy/noiseless snapshots. The case of noisy correlated snapshots is of practical importance for airborne/seaborne adaptive radar detectors which is the main motivation for this study.

Adaptive radar detectors aim to directly or indirectly estimate the interference (clutter plus noise) covariance matrix from a set of snapshots called secondary dataset [11], [12]. More specifically, the Kelly's test [13] and adaptive coherence estimator (ACE) [14] indirectly estimate the unknown interference covariance matrix in the generalized likelihood ratio test (GLRT) framework. These two tests are similar but have important differences in their assumptions. Kelly's test assumes that the covariance matrix of the primary data, which is also called cell-under test (CUT), is identical to the covariance matrix of the secondary data cells (homogeneous environment assumption). ACE assumes that covariances matrices of secondary data cells are identical; but differ from the CUT covariance matrix by an unknown scaling factor (partially homogeneous environment assumption), [11, Sec. 2.3]. Hence, ACE assumes that the primary cell can be at different average power level compared with the secondary data cells. In literature, the compound-Gaussian models are suggested to take into account such power fluctuations [15]. For example, the clutter snapshot vector $\mathbf{x}_k$ can be modeled as a compound-Gaussian vector $\mathbf{x}_k = \sqrt{\tau_k}\mathbf{u}_k$, where τ_k (texture component) is a random variable independent of the Gaussian vector $\mathbf{u}_k$ (speckle component). With the compound-Gaussian model, the texture parameter τ_k in Kelly's test becomes a global constant for both CUT and secondary data cells. ACE also assumes a fixed value for the secondary data texture parameter; but this value is independent of the texture parameter of CUT. Note that both methods assume a constant texture value for secondary data cells. This assumption, i.e., the constant texture parameter assumption, is valid when the shape parameter ν of K-distributed clutter is arbitrarily large (as $\nu \to \infty$, the K-distribution approaches Rayleigh distribution with constant

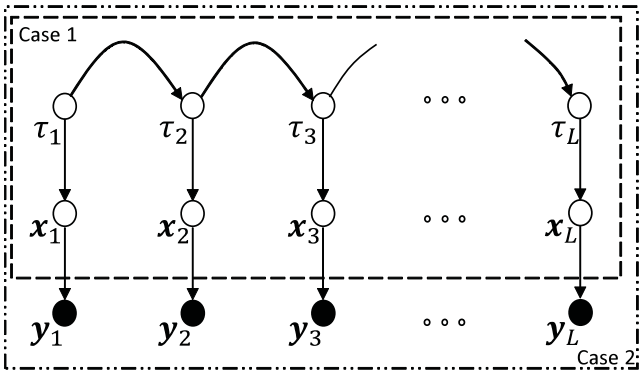


Fig. 1. Hidden Markov model for Case 1: Noise-free snapshots (inner rectangle with $\mathbf{x}_k$'s as observations) and Case 2: Noisy snapshots (outer rectangle with $\mathbf{y}_k$'s as observations) studied in Section III. In both models, τ_k variables are hidden variables to be estimated from available observations.

power [16]) or when the texture is fully correlated (spatial correlation) across the data cells. It is known that homogenous or partially homogenous environment assumption is not satisfied, or satisfied for only a limited number of cells in a spiky clutter environment [17, Figs. 2 and 3], [16, Sec. 4.3.2]. Hence, the assumptions of Kelly's and ACE detectors are in mismatch with the experimentally observed statistical description of sea clutter. This mismatch is considered to be at least partially responsible for the deviation of the experimental adaptive detector performance from the theoretical predictions [18], [19].

This article suggests a novel compound-Gaussian model with texture-correlation. The texture-correlation in this study is induced through a Markov chain. From an adaptive detection perspective, the states of the chain can be considered to reflect the power ratio between CUT and a secondary data cell. This ratio is taken as 1 and a deterministic (nonrandom) unknown for Kelly's and ACE tests, respectively. Here, we allow this ratio to be a random variable and study the covariance matrix estimation for this setup. The suggested Markov model is also in agreement with some empirical observations on sea clutter, where a sequence of high clutter power cells are observed to be followed by a sequence of low powered cells, as shown in [16, Fig. 2.21] and also with the experimentally observed spatial correlation of sea clutter [17, Figs. 2 and 3].

More specifically, [16, Fig. 2.21] shows the sea-clutter measurements and their spatial correlation for different operating conditions in relation with the grazing angle, range resolution, sea state, etc. It can be observed that there exists a significant correlation between the clutter powers for more than ten range cells in some scenarios. It has also been observed in [20, Fig. 2] that for the range resolutions of 10 m or less, there exists some spatial correlation between the neighboring cells in the field data recorded by IPIX radar of McMaster University (Canada). In the same study, the periodicity of clutter power, due to the periodic behavior of sea surface and wave motion, is also observed at high range resolutions, say at 3 m. The adaptive target detection methods in literature typically ignore the spatial correlation, with the notable exception of [21], and run the risk of underperforming due to its mismatch. In [21], the effect of spatial correlation on the covariance matrix estimation from secondary data and its effect on the performance of ACE detector is studied. Results indicate that there is a significant performance improvement when spatial correlation is taken into account. In another related work, the clutter subspace is estimated via the expectation–maximization formulation for the case of nonhomogenous but uncorrelated texture parameters [22]. Different from this work, we consider the spatially correlated clutter and also take into account the effect of white noise (thermal noise) in the covariance matrix estimation.

The main contribution of this article is to build a hidden Markov model (HMM)-based scheme to take into account both amplitude/power fluctuations and spatial correlation in the secondary data. To this aim, we study the compound-Gaussian model for the texture-correlated snapshots and develop a maximum likelihood (ML) covariance matrix estimator, via the EM method. The suggested method is easily applicable for the structured covariance matrix estimation. To illustrate that, we also examine the important case of persymmetric covariance matrix estimation.

Notation: Small and capital bold letters, as in $\mathbf{x}$ and $\mathbf{M}$, refer to vectors and matrices, respectively. The probability mass function of discrete random variables, which is the texture variable in this study, is denoted with $p(\tau = t)$. Here τ is the random variable and t is an element of sample space $\mathcal{K}, t \in \mathcal{K} = \{\kappa_1, \ldots, \kappa_S\}$. The probability density functions of continuous random variables are denoted with $f(\mathbf{x})$. Here $\mathbf{x}$ denotes a random vector with the joint density $f(\mathbf{x})$. The density function for mixed random variables are also denoted with $f(\mathbf{x}, \tau)$, where the Dirac delta functions are utilized to represent the finite sample space $\mathcal{K} = \{\kappa_1, \ldots, \kappa_S\}$ of the discrete random variable τ. The conditional densities are shown with $f(\mathbf{x}_k|\tau_k = \kappa_s)$. The parameters of the densities are shown as in $f(\mathbf{x}_k|\tau_k; \mathbf{M})$, that is the nonrandom parameters are separated by the semicolon which is the matrix parameter $\mathbf{M}$ in this example. The operators $E\{\cdot\}$ and $E\{\cdot|A\}$ denote the expectation and conditional expectation given the event A, respectively.

II. HMM-BASED TEXTURE-CORRELATED COMPOUND-GAUSSIAN MODEL

1) Compound-Gaussian Signal Model: The snapshot vector $\mathbf{x}_k$ is formed by the product of a scalar variable $\sqrt{\tau_k}$ (texture amplitude) and $N \times 1$ dimensional vector $\mathbf{c}_k$ (speckle)

$$\mathbf{x}_k = \sqrt{\tau_k}\mathbf{c}_k, \quad k = \{1, \ldots, L\}. \tag{1}$$

The speckle component $\mathbf{c}_k$ is assumed to be independent and identically distributed (iid) with circularly symmetric complex Gaussian random vector with zero mean and covariance matrix $\mathbf{M}$, $\mathbf{c}_k \sim CN(\mathbf{0}, \mathbf{M})$.

The texture variables τ_k are discrete-valued random variables with the sample space $\mathcal{K} = \{\kappa_1, \ldots, \kappa_S\}$. We assume that the texture variables (the parameter in relation with the

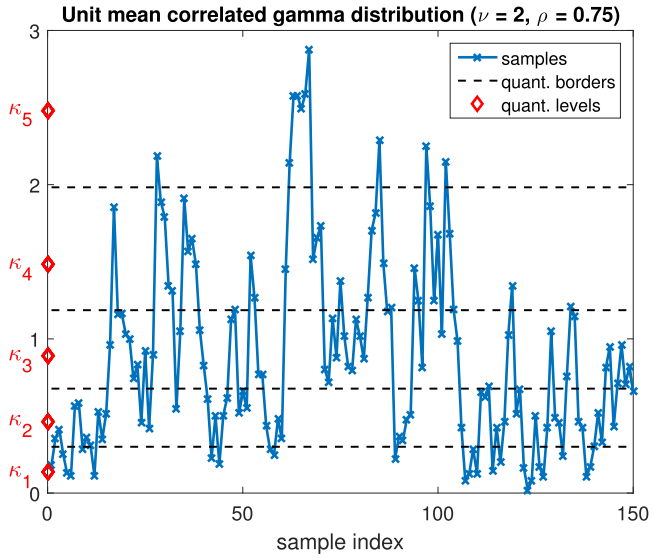


Fig. 2. Samples of unit mean correlated gamma distribution with parameters $\nu = 2$ $\rho = 0.75$. Five-level quantization values $\{\kappa_1, \ldots, \kappa_5\}$ and the quantization borders are also shown.

clutter power in radar applications) are mapped to a total S levels, say $\{0, 10, 20, 40\}$ dB, indicating the severity of clutter in a range cell.[1] We assume that the clutter severity level persists for a number of range cells, say a strong clutter cell is more likely to be followed by another strong clutter cell; as observed in some sea-clutter measurements. The main goal is to make use of this dependency in the estimation of clutter covariance matrix and also in adaptive target detection.

The statistical dependency of the texture variables τ are modeled by a Markov structure shown in Fig. 1 and the joint density of the texture variables can be written as follows:

$$\begin{aligned} &p(\tau_1 = t_1, \tau_2 = t_2, \ldots, \tau_L = t_L) \\ &\quad = p(\tau_1 = t_1) \prod_{l=2}^{L} p(\tau_l = t_l | \tau_{l-1} = t_{l-1}) \end{aligned} \tag{2}$$

where $t_l \in \mathcal{K} = \{\kappa_1, \ldots, \kappa_S\}$, $l = \{1, \ldots, L\}$. The state-transition probabilities of the Markov chain is denoted with $\pi_{ij} = P(\tau_k = \kappa_i \,|\, \tau_{k-1} = \kappa_j)$. It is easy to see from (1) that the snapshot vector $\mathbf{x}_k$ given $\tau_k = \kappa_s \in \mathcal{K}$ has the distribution of $\mathbf{x}_k \mid \tau_k = \kappa_s \sim CN(\mathbf{0}, \kappa_s \mathbf{M})$.

Our goal is the estimation of the covariance matrix $\mathbf{M}$ given the observations $\mathbf{x}_l$, $l = \{1, \ldots, L\}$ in the absence of information on the texture variables, i.e., when the texture variables are hidden. With marginalization operation over the texture variables, the joint density of observation vectors can be written as follows:

$$\begin{aligned} f(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_L) = &\sum_{t_1 \in \mathcal{K}, \ldots, t_L \in \mathcal{K}} \prod_{l=1}^{L} f(\mathbf{x}_l | \tau_l = t_l) \\ &p(\tau_1 = t_1, \ \ldots \ , \tau_L = t_L). \end{aligned} \tag{3}$$

Note that, the argument of the sum on the right-hand side of (3) is the joint density $f(x_1, \ \ldots \ , x_L, \ \tau_1, \ \ldots \ , \tau_L)$. We give an EM-based solution to the problem by treating the texture variables as hidden variables in this study.

2) Noisy Snapshot Case: In many applications, the snapshot vector $\mathbf{x}_k$ is not directly observed, but its noisy version $\mathbf{y}_k$ is available

$$\mathbf{y}_k = \mathbf{x}_k + \mathbf{n}_\mathbf{k}, \quad k = \{1, \ldots, L\}. \tag{4}$$

Here the noise vector $\mathbf{n}_\mathbf{k}$ is assumed to be independent of all other random variables and is Gaussian distributed, $\mathbf{n}_k \sim CN(\mathbf{0}, \sigma_n^2 \mathbf{I})$. The covariance matrix estimation problem for this setup is the estimation of $\mathbf{M}$ given the noisy snapshots $\mathbf{y}_k$, $k = \{1, \ldots, L\}$. The graphical model for this case is illustrated as Case 2 in Fig. 1.

In the context of adaptive target detectors, the first cell, that is $\mathbf{y}_1$, can be considered as CUT with the texture parameter τ_1. The other cells correspond to the secondary data cells with texture parameters τ_k. The texture variables are correlated according to the density given in (2). The central goal is to estimate the CUT clutter covariance matrix from the secondary data cells $\{\mathbf{y}_2, \ldots, \mathbf{y}_L\}$. In Section IV, we suggest a method for the construction of the HMM structure required for the described system model.

3) Identifiability Problem: Since the snapshot $\mathbf{x}_k = \sqrt{\tau_k}\mathbf{c}_k$ is the product of random variable τ_k and random vector $\mathbf{c}_k$, the factors of the product can be identified apart from a constant scaling factor. Stated differently, the substitution of τ_k with $\alpha\tau_k$, that is $\tau_k \leftarrow \alpha\tau_k$, and $\mathbf{c}_k \leftarrow \alpha^{-1}\mathbf{c}_k$ yields an identical product for $\tau_k\mathbf{c}_k$. Hence, the unknown covariance matrix $\mathbf{M}$ can only be identified apart from a constant. (Note that due to the discrete-valued τ_k definition in this study, the scaling factor α is limited to the ratio of two outcomes in the sample space $\mathcal{K} = \{\kappa_1, \ldots, \kappa_S\}$ of this variable.) To uniquely identify $\mathbf{M}$, a constraint on $\mathbf{M}$, such as $\mathrm{tr}\{\mathbf{M}\} = N$, is required. Here, without any loss of generality, we assume that the texture variable is scaled such that $E\{\tau_k\} = 1$.

III. COVARIANCE MATRIX ESTIMATION FOR TEXTURE CORRELATED SNAPSHOTS

We present an EM method-based solution for the covariance matrix estimation for noise-free and noisy snapshots in this section.

A. Case 1: Noise-Free Snapshots

We assume L snapshots $\mathbf{x}_k = \sqrt{\tau_k}\mathbf{c}_k$, $k = \{1, 2, \ldots, L\}$ are given for the estimation of $\mathbf{M}$. The texture variable vector $\boldsymbol{\tau} = [\tau_1 \ \ldots \ \tau_L]$ is the hidden (unobserved) vector variable of the problem and

[1] The conventional texture variable is a continuous-valued random variable, typically representing the clutter power from a range cell in radar signal processing applications. With the introduced discrete-valued texture variable, we consider that the clutter power is mapped (quantized) to the values representing low, medium, and high power cases with some granularity. Fig. 2 illustrates this operation for the gamma distributed correlated texture variable for a five-level quantization. The introduced discretization/quantization of clutter power is necessitated by the need of the posterior distribution calculation of hidden variables (τ) in the graphical model shown in Fig. 1. More details about the HMM model and its construction is given in Section IV.

$N \times L$ dimensional $\mathbf{X} = [\mathbf{x}_1 \ \dots \ \mathbf{x}_L]$ is the matrix of observation vectors. For the EM formulation, we specify the complete dataset as $\mathbf{X}_{\text{comp}} = \{\mathbf{X}, \boldsymbol{\tau}\}$. The log-likelihood of complete dataset, $\Lambda(\mathbf{X}_{\text{comp}}) = \log(f(\mathbf{X}_{\text{comp}})) = \log(f(x_1, \ \dots \ , x_L, \ \tau_1, \ \dots \ , \ \tau_L))$ can be expressed as

$$\Lambda(\mathbf{X}_{\text{comp}}) \stackrel{c}{=} -\sum_{l=1}^{L} \log(|\tau_l \mathbf{M}|) + \frac{\mathbf{x}_l^H \mathbf{M}^{-1} \mathbf{x}_l}{\tau_l}. \tag{5}$$

Here $\stackrel{c}{=}$ shows to presence of additional terms on the right-hand side of the equation that do not affect the subsequent optimization steps for the estimation of $\mathbf{M}$ matrix.

The expectation-step (E-step) of the EM-algorithm requires the expectation calculation of $\Lambda(\mathbf{X}_{\text{comp}})$ over the posterior distribution of hidden variables $\boldsymbol{\tau}$ given the observation $\mathbf{X}$. E-step can be explicitly written as

$$\begin{aligned} J(\mathbf{M}, \mathbf{M}^{(k)}) &= E\left\{\Lambda(\mathbf{X}_{\text{comp}}) | \mathbf{X}; \mathbf{M}^{(k)}\right\} \\ &\stackrel{c}{=} -L\log(|\mathbf{M}|) - \sum_{l=1}^{L} E_{\tau_l|\mathbf{X};\mathbf{M}^{(k)}}\left\{\frac{1}{\tau_l}\right\} \mathbf{x}_l^H \mathbf{M}^{-1} \mathbf{x}_l \end{aligned} \tag{6}$$

where $\mathbf{M}^{(k)}$ is the estimate for the unknown covariance matrix at the kth iteration.

The maximization-step (M-step) requires the optimization of $J(\mathbf{M}, \mathbf{M}^{(k)})$ over the deterministic variable $\mathbf{M}$. By differentiating (6), we can get the maximizer as

$$\begin{aligned} \mathbf{M}^{(k+1)} &= \frac{1}{L} \sum_{l=1}^{L} E_{\tau_l|\mathbf{X};\mathbf{M}^{(k)}}\left\{\frac{1}{\tau_l}\right\} \mathbf{x}_l \mathbf{x}_l^H \\ &= \frac{1}{L} \sum_{l=1}^{L} \left(\sum_{s=1}^{S} \frac{p\left(\tau_l = \kappa_s | \mathbf{X}; \mathbf{M}^{(k)}\right)}{\kappa_s} \right) \mathbf{x}_l \mathbf{x}_l^H \end{aligned} \tag{7}$$

where $\mathbf{M}^{(k+1)}$ is the updated covariance matrix estimate. EM-algorithm is run by inserting the updated variable $\mathbf{M}^{(k+1)}$ into E-step, repeating the process until convergence. EM-algorithm is an ascent algorithm that increases the likelihood of the parameter estimate at every update. The algorithm is guaranteed to convergence; but, possibly to a local maxima [23].

It should be noticed that the implementation of E-step requires the marginal posterior distribution of hidden variables τ_l given the observations $\mathbf{X}$. For the problem of interest, this is the marginal posterior distribution of each texture variable given the noise-free snapshot vectors, that is $p(\tau_l = \kappa_s | \mathbf{X}; \mathbf{M}^{(k)})$, which is given as follows.

1) Posterior Probability Calculation via $\alpha\beta$ Recursion: The $\alpha\beta$ recursion is one of the methods for the marginal posterior probability calculation for the HMM structures with the graphical model in Fig. 1 [24].

The joint density of $\{\tau_l, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_l\}$ is denoted as $\alpha_l^{(k)}(\kappa_s) = p(\tau_l = \kappa_s, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_l; \mathbf{M}^{(k)})$, $l = \{1, \dots, L\}$. In this definition, the superscript (k) denotes the dependence of the probability expression to the covariance matrix estimate at the (k)th iteration, i.e., $\mathbf{M}^{(k)}$. The α-recursion is initialized with $\alpha_1^{(k)}(\kappa_s) = p(\tau_1 = \kappa_s) f(\mathbf{x}_1 | \tau_1 = \kappa_s; \mathbf{M}^{(k)})$ and recursively evaluated for $l \geq 2$ via

$$\alpha_l^{(k)}(\kappa_s) = f\left(\mathbf{x}_l \mid \tau_l = \kappa_s; \mathbf{M}^{(k)}\right) \sum_{s'=1}^{S} \pi_{ss'} \alpha_{l-1}^{(k)}(\kappa_{s'}). \tag{8}$$

Here $\pi_{ss'}$ is the state-transition probability, $\pi_{ss'} = p(\tau_l = \kappa_s | \tau_{l-1} = \kappa_{s'})$ and $f(\mathbf{x}_l \mid \tau_l = \kappa_s; \mathbf{M}^{(k)})$ is the likelihood value of $\mathbf{x}_l | \tau_l = \kappa_s \sim CN(\mathbf{0}, \kappa_s \mathbf{M}^{(k)})$.

The function $\beta_l^{(k)}(\kappa_s) = f(\mathbf{x}_{l+1}, \dots, \mathbf{x}_L \mid \tau_l = \kappa_s; \mathbf{M}^{(k)})$ denotes a likelihood value for the lth texture variable for $l = \{1, \dots, L-1\}$. β-recursion is initialized with $\beta_L^{(k)}(\kappa_s) = 1 \ \forall \kappa_s \in \mathcal{K}$ and β-sequence recursively evaluated for decreasing l from L to 2

$$\beta_{l-1}^{(k)}(\kappa_s) = \sum_{s'=1}^{S} f\left(\mathbf{x}_l \mid \tau_l = \kappa_{s'}; \mathbf{M}^{(k)}\right) \pi_{s's} \beta_l^{(k)}(\kappa_{s'}). \tag{9}$$

Finally, the posterior probability that we are after, $p(\tau_l = \kappa_s \mid \mathbf{X}; \mathbf{M}^{(k)})$, is the normalized version of the product of α and β sequences obtained from $\alpha\beta$ recursion, [24]

$$p\left(\tau_l = \kappa_s \mid \mathbf{X}; \mathbf{M}^{(k)}\right) \propto \alpha_l^{(k)}(\kappa_s) \beta_l^{(k)}(\kappa_s). \tag{10}$$

The suggested covariance matrix estimation algorithm is summarized in Table I. A ready-to-use MATLAB implementation, including the persymmetric case, is provided in [25].

We would like to note that the suggested algorithm does not assume any structure on $\mathbf{M}$ matrix. Yet, any additional knowledge on the structure of $\mathbf{M}$ matrix can be easily taken

---

Algorithm 1: Proposed Method for Texture-Correlated Noise-free Snapshots, [25].

Input : $\mathbf{X} = [\mathbf{x}_1 \, \mathbf{x}_2 \ \dots \ \mathbf{x}_L]$, maxiter: maximum number of iterations

Output : $\mathbf{M}_n$ (normalized covariance matrix), η (power normalization factor)

Requires: HMM initial state probabilities $p(\tau_1 = \kappa_s)$, HMM state transition probabilities $(\pi_{ss'})$, the conditional distribution of the observation given the hidden state, $(f(\mathbf{x}_l | \tau_l))$.

1: Initial Conditions: $k = 0$, $\mathbf{M}^{(0)} = \mathbf{I}$

2: Marginal posterior density calculation:
$\alpha_1^{(k)}(\kappa_s) = p(\tau_1 = \kappa_s) f\left(\mathbf{x}_1 | \tau_1 = \kappa_s; \mathbf{M}^{(k)}\right)$
for $l = 2 : L$,
$\alpha_l^{(k)}(\kappa_s) = f(\mathbf{x}_l \mid \tau_l = \kappa_s; \mathbf{M}^{(k)}) \sum_{s'=0}^{S} \pi_{ss'} \alpha_{l-1}^{(k)}(\kappa_{s'})$
end
$\beta_L^{(k)}(\kappa_s) = 1$
for $l = L : -1 : 2$,
$\beta_{l-1}^{(k)}(\kappa_s) = \sum_{s'=1}^{S} f(\mathbf{x}_l | \tau_l = \kappa_{s'}; \mathbf{M}^{(k)}) \pi_{s's} \beta_l^{(k)}(\kappa_{s'})$
end
$p(\tau_l = \kappa_s \mid \mathbf{X}; \mathbf{M}^{(k)}) \propto \alpha_l^{(k)}(\kappa_s) \beta_l^{(k)}(\kappa_s)$

3: Covariance matrix update:
$\mathbf{M}^{(k+1)} = \frac{1}{L} \sum_{l=1}^{L} \left(\sum_{s=1}^{S} p\left(\tau_l = \kappa_s | \mathbf{X}; \mathbf{M}^{(k)}\right) \frac{1}{\kappa_s} \right) \mathbf{x}_l \mathbf{x}_l^H$

4: $k \leftarrow k + 1$

5: **if** $k <$ maxiter, goto Step-2

6: **Return** $\mathbf{M}_n = \mathbf{M}^{(k)} \times N / \text{tr}\{\mathbf{M}^{(k)}\}$, $\eta = \text{tr}\{\mathbf{M}^{(k)}\}/N$

TABLE I
Suggested 3/5/7 Level Quantization Values, in Decibels, for Unit Mean Gamma Distributed Texture Parameter With Shape Parameter ν

	3 Level			5 Level					7 Level						
	κ_{-1}	κ_0	κ_1	κ_{-2}	κ_{-1}	κ_0	κ_1	κ_2	κ_{-3}	κ_{-2}	κ_{-1}	κ_0	κ_1	κ_2	κ_3
$\nu = 1$	-17.05	-1.76	3.24	-19.8	-5.09	-0.8	2.26	5.15	-21.68	-7.24	-3.14	-0.41	1.81	3.87	6.11
$\nu = 1.25$	-11.77	-1.53	2.97	-14.37	-4.55	-0.76	2.03	4.71	-16.16	-6.5	-2.9	-0.43	1.6	3.5	5.61
$\nu = 1.50$	**-9.17**	**-1.27**	**2.81**	-11.51	-4.04	-0.66	1.9	4.4	-13.12	-5.81	-2.63	-0.39	1.48	3.25	5.24
$\nu = 1.75$	-7.7	-1.08	2.68	-9.83	-3.64	-0.58	1.8	4.15	-11.29	-5.27	-2.41	-0.36	1.38	3.05	4.94
$\nu = 2$	-6.74	-0.94	2.57	-8.69	-3.32	-0.51	1.71	3.95	-10.03	-4.84	-2.23	-0.32	1.31	2.89	4.69
$\nu = 4$	**-3.82**	**-0.45**	**2.02**	-5.1	-2.12	-0.26	1.32	3.01	-5.96	-3.15	-1.48	-0.18	0.99	2.17	3.56
$\nu = 6$	-2.88	-0.3	1.73	-3.89	-1.66	-0.18	1.12	2.54	-4.56	-2.48	-1.17	-0.12	0.84	1.82	3.01
$\nu = 8$	-2.39	-0.22	1.54	-3.25	-1.4	-0.13	0.99	2.25	-3.81	-2.1	-0.99	-0.09	0.74	1.61	2.67
$\nu = 10$	**-2.08**	**-0.18**	**1.4**	-2.83	-1.23	-0.11	0.9	2.05	-3.33	-1.85	-0.88	-0.07	0.67	1.46	2.42
$\nu = 100$	-0.57	-0.02	0.5	-0.79	-0.35	-0.01	0.32	0.71	-0.93	-0.54	-0.25	-0.01	0.23	0.5	0.84

into account in the M-step of the algorithm. As an example, when $\mathbf{M}$ matrix is restricted to the class of persymmetric matrices; the M-step of the suggested scheme should be restricted to the search space of persymmetric matrices. In the literature, there are several studies describing such restrictions and these studies can be integrated easily with the proposed method. As an example, with the persymmetric matrix constraint, the required modification is simply the replacement of M-step in (7) with

$$\mathbf{M}_{\text{persym}}^{(k+1)} = \frac{1}{2L}\sum_{l=1}^{L}\left(\sum_{s=1}^{S}\frac{p\left(\tau_l=\kappa_s|\mathbf{X};\mathbf{M}^{(k)}\right)}{\kappa_s}\right)\left(\mathbf{x}_l\mathbf{x}_l^H+\mathbf{x}_l^R\left(\mathbf{x}_l^R\right)^H\right). \tag{11}$$

Here $\mathbf{x}_k^R$ denotes the reversed (flipped up-down) and conjugated vector $\mathbf{x}_k$ (see the Appendix for more details).

As another example, the set of feasible covariance matrices can be restricted to $\mathcal{S}=\{\mathbf{R}:\mathbf{R}=\mathbf{M}+\sigma^2\mathbf{I},\mathbf{M}\succeq 0,\text{rank}(\mathbf{M})\le r\}$, as in [26] for the goal of low-rank covariance matrix estimation. The algorithm in [26] can also be utilized in the set up of this article by replacing the sample covariance matrix calculation step (given as $\mathbf{X}\mathbf{X}^H$ in [26]) with (7) and running the low-rank covariance matrix estimation method in [26]) after this replacement.

B. Case 2: Noisy Snapshots

In this section, we extend the algorithm developed for noise-free snapshots to the noisy snapshots. The compound-Gaussian snapshots $\mathbf{x}_k$ are assumed to be observed according to the following model:

$$\mathbf{y}_k=\mathbf{x}_k+\mathbf{n}_k=\sqrt{\gamma_k}\mathbf{c}_k+\mathbf{n}_k,\quad k=\{1,\ldots,L\}. \tag{12}$$

Here, the noise vector $\mathbf{n}_k$ is assumed to be zero-mean white Gaussian noise, $\mathbf{n}_k\sim CN(\mathbf{0},\sigma_n^2\mathbf{I})$. The problem statement is the estimation of covariance matrix $\mathbf{M}$ given the observation matrix $\mathbf{Y}=[\mathbf{y}_1,\ \mathbf{y}_2,\ \ldots,\mathbf{y}_L]$, as illustrated in Case 2 of Fig. 1.

To adapt the earlier solution to this problem, we include the noise-free snapshots $\mathbf{x}_k$ among the hidden variables of the problem. Hence, the complete dataset for this problem becomes $\mathbf{Y}_{\text{comp}}=\{\boldsymbol{\tau},\mathbf{X},\mathbf{Y}\}$ with $\boldsymbol{\tau}=[\tau_1,\ \ldots,\ \tau_L]$ and $\mathbf{X}=[\mathbf{x}_1,\ \ldots,\ \mathbf{x}_L]$. The joint density function of the complete dataset can be factorized as follows:

$$\begin{aligned} f(\mathbf{Y}_{\text{comp}}) &= f(\tau_{1:L},\mathbf{x}_{1:L},\mathbf{y}_{1:L};\mathbf{M}) \\ &= f(\tau_1)f(\mathbf{x}_1|\tau_1;\mathbf{M})f(\mathbf{y}_1|\mathbf{x}_1) \\ &\quad \prod_{l=2}^{L}p(\tau_l|\tau_{l-1})f(\mathbf{x}_l|\tau_l;\mathbf{M})f(\mathbf{y}_l|\mathbf{x}_l). \end{aligned} \tag{13}$$

Note that the conditional densities $f(\mathbf{x}_l|\tau_l;\mathbf{M})$ are the only factors having dependency on $\mathbf{M}$, where $\mathbf{x}_l\mid\tau_l=\kappa_s\sim CN(\mathbf{0},\kappa_s\mathbf{M})$. Due to noisy observations, we have an additional conditional density for the observation vector which is $\mathbf{y}_l\mid\mathbf{x}_l\sim CN(\mathbf{x}_l,\sigma_n^2\mathbf{I})$.

1) Expectation Step: The conditional expectation of complete data log-likelihood $J(\mathbf{M},\mathbf{M}^{(k)})=E_{\mathbf{X},\boldsymbol{\tau}|\mathbf{Y},\theta^{\text{old}}}\{\log f(\mathbf{Y}_{\text{comp}})\}$ is

$$\begin{aligned} J\left(\mathbf{M},\mathbf{M}^{(k)}\right) &\overset{c}{=} \sum_{l=1}^{L}E_{\mathbf{x}_l,\tau_l|\mathbf{Y}}\{\log f(\mathbf{x}_l|\tau_l;\mathbf{M})\} \\ &\overset{c}{=} -\sum_{l=1}^{L}E_{\mathbf{x}_l,\tau_l|\mathbf{Y}}\left\{\log|\tau_l\mathbf{M}|+\frac{\mathbf{x}_l^H\mathbf{M}^{-1}\mathbf{x}_l}{\tau_l}\right\} \\ &\overset{c}{=} -L\log|\mathbf{M}| \\ &\quad -\sum_{l=1}^{L}E_{\tau_l|\mathbf{Y}}\left\{E_{\mathbf{x}_l|\tau_l,\mathbf{Y}}\left\{\frac{\mathbf{x}_l^H\mathbf{M}^{-1}\mathbf{x}_l}{\tau_l}\right\}\right\}. \end{aligned} \tag{14}$$

The inner expectation in (14), $E_{\mathbf{x}_l|\tau_l,\mathbf{Y}}\{\mathbf{x}_l^H\mathbf{M}^{-1}\mathbf{x}_l/\tau_l\}$, can be evaluated by first rewriting the argument of the expectation as $\text{tr}(\mathbf{M}^{-1}\mathbf{x}_l\mathbf{x}_l^H)/\tau_l$ and interchanging trace-expectation operations and utilizing $E_{\mathbf{x}_l|\tau_l,\mathbf{Y}}\{\mathbf{x}_l\mathbf{x}_l^H\}=\overline{\mathbf{x}}_{l,\tau_l}\overline{\mathbf{x}}_{l,\tau_l}^H+\mathbf{K}_{\tau_l}$. This expectation easily follows from the conditional density of $\mathbf{x}_l$ given $\{\tau_l,\mathbf{Y}\}$.

It can be easily verified from the graphical model in Fig. 1 that $\mathbf{x}_l$ is independent of $\mathbf{x}_k$, $k\neq l$ conditioned on τ_l and $\mathbf{y}_l$. Hence, we only need to process the observation $\mathbf{y}_l=\mathbf{x}_l+\mathbf{n}_l$ for the estimation of $\mathbf{x}_l$ when the outcomes for τ_l and $\mathbf{y}_l$ are provided, i.e., fixed, via the conditioning operation. The *a priori* distribution of $\mathbf{x}_l$ given the texture parameter τ_l and the kth iteration estimate of $\mathbf{M}$ matrix is $\mathbf{x}_l|\tau_l\sim CN(\mathbf{0},\tau_l\mathbf{M}^{(k)})$. The noise vector is independent of the signal vector and has the distribution $\mathbf{n}_l\sim CN(\mathbf{0},\sigma_n^2\mathbf{I})$. Under these conditions, the posterior distribution of $\mathbf{x}_l$ given $\mathbf{y}_l$ and τ_l, that is $\mathbf{x}_l\mid\{\mathbf{y}_l,\tau_l\}$, is also jointly Gaussian

distributed [23, Sec. 4.3] with $\mathbf{x}_l \mid \{\mathbf{y}_l, \tau_l\} \sim CN(\overline{\mathbf{x}}_{l,\tau_l}, \mathbf{K}_{\tau_l})$, where

$$\overline{\mathbf{x}}_{l,\tau_l} = \tau_l \mathbf{M}^{(k)} \left(\tau_l \mathbf{M}^{(k)} + \sigma_n^2 \mathbf{I}\right)^{-1} \mathbf{y}_l$$
$$\mathbf{K}_{\tau_l} = \left(\frac{1}{\sigma_n^2}\mathbf{I} + \frac{1}{\tau_l}\left(\mathbf{M}^{(k)}\right)^{-1}\right)^{-1}. \quad (15)$$

Once the inner expectation result is inserted in (14), we have

$$\begin{aligned} J(\mathbf{M}, \mathbf{M}^{(k)}) &\stackrel{c}{=} -L \log|\mathbf{M}| \\ &\quad - \sum_{l=1}^{L} E_{\tau_l|\mathbf{Y}} \left\{ \frac{\overline{\mathbf{x}}_{l,\tau_l}^H \mathbf{M}^{-1} \overline{\mathbf{x}}_{l,\tau_l} + \mathrm{tr}\left(\mathbf{M}^{-1}\mathbf{K}_{\tau_l}\right)}{\tau_l} \right\} \\ &\stackrel{c}{=} -L \log|\mathbf{M}| - \sum_{l=1}^{L}\sum_{s=1}^{S} \frac{p(\tau_l = \kappa_s|\mathbf{Y})}{\kappa_s} \overline{\mathbf{x}}_{l,\kappa_s}^H \mathbf{M}^{-1} \overline{\mathbf{x}}_{l,\kappa_s} \\ &\quad - L \sum_{s=1}^{S} \frac{\bar{p}(\tau = \kappa_s|\mathbf{Y})}{\kappa_s} \mathrm{tr}\left(\mathbf{M}^{-1}\mathbf{K}_{\kappa_s}\right). \end{aligned} \quad (16)$$

Here

$$\bar{p}(\tau = \kappa_s|\mathbf{Y}) = \frac{1}{L}\sum_{l=1}^{L} p(\tau_l = \kappa_s|\mathbf{Y}) \quad (17)$$

denotes the average posterior of texture variable over all snapshots. The marginal posterior probability calculation is similar to the noiseless case. The main difference in marginal probability calculation is the replacement of earlier likelihood function $f(\mathbf{x}_l|\tau_l)$ with $f(\mathbf{y}_l|\tau_l)$. Details are given in Algorithm 2.

2) Maximization Step: The maximization step involves the maximization of $J(\mathbf{M}, \mathbf{M}^{(k)})$ given by (16) with respect to $\mathbf{M}$. By differentiation, we get the extrema as

$$\begin{aligned} \mathbf{M}^{(k+1)} &= \frac{1}{L}\sum_{l=1}^{L}\sum_{s=1}^{S} \frac{p(\tau_l = \kappa_s|\mathbf{Y})}{\kappa_s} \overline{\mathbf{x}}_{l,\kappa_s} \overline{\mathbf{x}}_{l,\kappa_s}^H \\ &\quad + \sum_{s=1}^{S} \frac{\bar{p}(\tau = \kappa_s|\mathbf{Y})}{\kappa_s} \mathbf{K}_{\kappa_s}. \end{aligned} \quad (18)$$

Note that, the covariance matrix estimation step for the noiseless and noisy snapshots, given in (7) and (18), respectively, are quite similar. In the noisy case, the mse-optimal estimate for the snapshot vector, that is $\overline{\mathbf{x}}_{l,\kappa_s}$ is utilized instead of $\mathbf{x}_l$. Also, the estimation error on $\mathbf{x}_l$ is taken into account with the inclusion of the error covariance matrix $\mathbf{K}_{\kappa_s}$ in (18).

The maximization step generating the covariance matrix estimate in (18) does not assume any matrix structure. As in the noise-free case, the maximization step can be easily updated to include the structural constraints. For example, with the persymmetry constraint, the update equation becomes

$$\begin{aligned} \mathbf{M}_{\mathrm{persym}}^{(k+1)} &= \frac{1}{2L}\sum_{l=1}^{L}\sum_{s=1}^{S} \frac{p(\tau_l = \kappa_s|\mathbf{Y})}{\kappa_s} \left(\overline{\mathbf{x}}_{l,\kappa_s}\overline{\mathbf{x}}_{l,\kappa_s}^H + \overline{\mathbf{x}}_{l,\kappa_s}^R \left(\overline{\mathbf{x}}_{l,\kappa_s}^R\right)^H\right) \\ &\quad + \sum_{s=1}^{S} \frac{\bar{p}(\tau = \kappa_s|\mathbf{Y})}{\kappa_s} \mathbf{K}_{\kappa_s}. \end{aligned} \quad (19)$$

Algorithm 2: Proposed Method for Texture-Correlated Noisy Snapshots.

Input : $\mathbf{Y} = [\mathbf{y}_1\, \mathbf{y}_2\, \ldots\, \mathbf{y}_L]$, maxiter: maximum number of iterations
Output : $\mathbf{M}_n$ (normalized covariance matrix), η (power normalization factor)
Requires: HMM initial state probabilities $p(\tau_1 = \kappa_s)$, HMM state transition probabilities $(\pi_{ss'})$, the conditional distribution of the observation given the hidden state, $(f(\mathbf{y}_l|\tau_l))$.

1: Initial Conditions: $k = 0$, $\mathbf{M}^{(0)} = \mathbf{I}$
2: Marginal posterior density calculation:
$\alpha_1^{(k)}(\tau_s) = p(\tau_1 = \kappa_s) f(\mathbf{y}_1|\tau_1 = \kappa_s; \mathbf{M}^{(k)})$
for $l = 2 : L$,
$\alpha_l^{(k)}(\kappa_s) = f(\mathbf{y}_l|\tau_l = \kappa_s; \mathbf{M}^{(k)}) \sum_{s'=0}^{S} \pi_{ss'} \alpha_{l-1}^{(k)}(\kappa_{s'})$;
end
$\beta_L^{(k)}(\kappa_s) = 1$
for $l = L : -1 : 2$,
$\beta_{l-1}^{(k)}(\kappa_s) = \sum_{s'=1}^{S} f(\mathbf{y}_l|\tau_l = \kappa_{s'}; \mathbf{M}^{(k)}) \pi_{s's} \beta_l^{(k)}(\kappa_{s'})$;
end
$p(\tau_l = \kappa_s \mid \mathbf{Y}; \mathbf{M}^{(k)}) \propto \alpha_l^{(k)}(\kappa_s) \beta_l^{(k)}(\kappa_s)$
Snapshots:
for $s = 1 : S$,
for $l = 1 : L$,
$\overline{\mathbf{x}}_{l,\kappa_s} = \kappa_s \mathbf{M}^{(k)} \left(\kappa_s \mathbf{M}^{(k)} + \sigma_n^2 \mathbf{I}\right)^{-1} \mathbf{y}_l$
end;
$\mathbf{K}_{\kappa_s} = \left(\frac{1}{\sigma_n^2}\mathbf{I} + \frac{1}{\kappa_s}\left(\mathbf{M}^{(k)}\right)^{-1}\right)^{-1}$
end;
3: Covariance matrix update:
$\mathbf{M}^{(k+1)} = \frac{1}{L}\sum_{l=1}^{L}\sum_{s=1}^{S} \frac{p(\tau_l = \kappa_s|\mathbf{Y})}{\kappa_s} \overline{\mathbf{x}}_{l,\kappa_s} \overline{\mathbf{x}}_{l,\kappa_s}^H + \sum_{s=1}^{S} \frac{\bar{p}(\tau = \kappa_s|\mathbf{Y})}{\kappa_s} \mathbf{K}_{\kappa_s}$
4: $k \leftarrow k + 1$
5: **if** $k <$ maxiter, goto Step-2
6: **Return** $\mathbf{M}_n = \mathbf{M}^{(k)} \times N/\mathrm{tr}\{\mathbf{M}^{(k)}\}$, $\eta = \mathrm{tr}\{\mathbf{M}^{(k)}\}/N$

Note that $\mathbf{M}_{\mathrm{persym}}^{(k+1)}$ in (19) is persymmetric provided that $\mathbf{K}_{\kappa_s}$ is persymmetric. The suggested EM algorithm initialization $\mathbf{M}^{(0)} = \mathbf{I}$ given in Algorithm 2 guarantees that.

IV. CONSTRUCTION OF HIDDEN MARKOV MODEL

In almost all applications, the texture parameter τ of the compound model is a continuous random variable. For example, the frequently used K-distributed sea-clutter model utilizes the gamma distribution as the texture distribution [16]. The gamma distribution has two parameters, namely, shape and scale parameters. In sea-clutter modeling, the shape parameter characterizes the spikiness of the clutter. A small shape parameter ν, say $\nu = 0.5$, corresponds to a highly spiky clutter and as ν increases, the spikiness is reduced. The scale parameter of the gamma distribution is related with its spread, that is when a gamma variate is multiplied/scaled by α, the resulting random variable is also a gamma variate with the same shape parameter; but its scale parameter is multiplied/scaled by α. In this study, we consider texture variables having unit mean value.

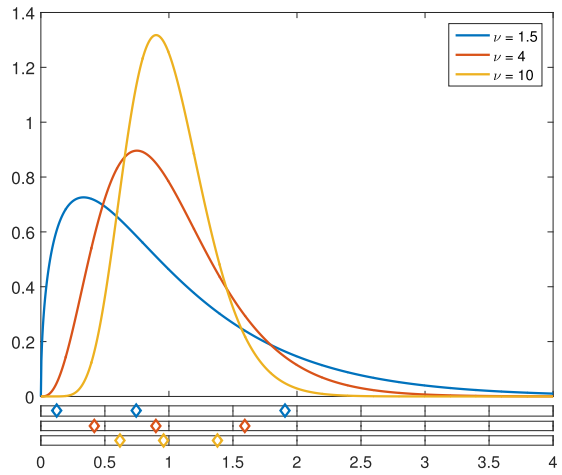


Fig. 3. Unit mean gamma distributions with shape parameters $\nu = \{1.5, 4, 10\}$ and three-level quantization values for each distribution (shown with the matching color diamond-shaped marker).

(When texture is gamma distributed, this corresponds to setting the scale parameter as $1/\nu$.) Working with a unit mean texture variable is not a limitation of the study as previously mentioned in the identifiability discussions given in Section II.

The conventional continuous-valued texture random variable can be utilized in the suggested HMM-based covariance matrix estimation setup provided that 1) the texture random variable can be accurately quantized taking into account the needs of the covariance matrix estimation application; 2) the state transition probabilities of the chain can be properly assigned. The solution of these problems are studied under the topic of HMM training in the literature [27]. Here, we deviate from the standard HMM training practice for the first problem (quantization of continuous valued texture random variable) which is mainly due to the application specific needs of the covariance matrix estimation problem.

1) Quantization of Continuous-Valued Texture Random Variable: Fig. 3 shows the density of unit mean gamma distributions with different shape parameters. In the same figure, a set of three-level quantization values for each distribution is also given. (The quantization values are indicated with diamond markers whose color match the color of the continuous density in the figure.) The choice for the number of quantization levels and their values is the first problem of HMM training [27]. For the specific application of interest, we suggest to utilize a modification of classical Lloyd-Max quantizer [28], [29].

The classical Lloyd-Max quantizer is the mse-optimal quantizer that minimizes the quantization error, $E_x\{(x - Q(x))^2\}$ [28], [29]. Here $Q(\cdot)$ refers to the quantization operation. The main application of the mse-optimal quantizer, with the error definition specialized to the quantization error $x - Q(x)$, is the source coding. Different from the source coding problem, the continuous valued texture variable τ is not to be reconstructed in this application. On the contrary, τ is a nuisance parameter of the problem whose value is not of any interest. An examination of Algorithm 1 should reveal that it is not the random variable τ, but its reciprocal $1/\tau$ affects the steps of estimation algorithm. More specifically, in the third step of Algorithm 1, the conditional mean of $1/\tau$, which is explicitly shown as $E_{\tau_l|\mathbf{X};\mathbf{M}^{(k)}}\{1/\tau_l\}$ in (6), is evaluated. Given this, we suggest to set the quantizer such that the reciprocal of τ is to be mse-optimal represented by minimizing $E_\tau\{(1/\tau - Q(\tau))^2\}$. A straightforward extension of the conventional Lloyd-Max quantizer to this problem leads to

$$\begin{aligned} \text{Step 1}: \quad & p_{q+1} = \frac{1}{2}(\kappa_q + \kappa_{q+1}) \quad q = \{0, \ldots, S-2\} \\ \text{Step 2}: \quad & \frac{1}{\kappa_q} = \frac{\int_{p_q}^{p_{q+1}} \frac{1}{\tau} f_\tau(\tau) d\tau}{\int_{p_q}^{p_{q+1}} f_\tau(\tau) d\tau} \quad q = \{0, \ldots, S-1\} \end{aligned} \tag{20}$$

where $p_0 \triangleq 0$ and $p_S \triangleq \infty$. In (20), κ_q, $q = \{0, \ldots, S-1\}$ indicate the quantization values and p_{q+1} indicates the separation boundary between κ_q and κ_{q+1}. Step 1 of (20) is identical to the classical Lloyd-Max quantizer, while the conditional mean of $1/\tau$, instead of τ, is calculated in Step 2. Similar to the conventional Lloyd-Max algorithm, these steps are run sequentially starting from an inital set of κ_q values, until convergence.

The quantization values shown in Fig. 3 correspond to the application of the modified Lloyd-Max quantizer for three quantization levels. The location of diamond markers on x-axis denotes the quantization values $\{\kappa_0, \kappa_1, \kappa_2\}$ for each distribution. It can be observed that the quantization values are more spread for small valued shape parameters and as the shape parameter increases, the quantization levels concentrate around unity. Table I gives the numerical values for the quantization for different cases. (The cases shown in Fig. 3 are marked with bold case.) The numerical values reflect the skewness of the distribution toward 0 for small ν. As ν increases, the distribution becomes almost symmetric around 1; it is not exactly symmetric, since the distribution is a one-sided.

The suggested method in (20) for setting quantization values is applicable to any texture distribution with density $f_\tau(\tau)$. If the density is not available in closed form, the conditional mean operation in Step 2 of (20) can be evaluated by the conditional mean of the empirical data. Hence, a distribution fitting to the empirical data is not required when setting the required quantization values for the HMM structure.

2) Assignment of State Transition Probabilities: For the solution of the second problem of HMM training, we follow the standard practice given in [27]. We assign the state transition probabilities according to the ratio of expected number of transitions from state-i to state-j and the expected number of times state-i is visited. Hence, following the conventional method, the assignment of state-transition probabilities reduces to counting the number of transitions between states.

3) Comments: For illustration purposes, we generate a sequence of correlated unit mean gamma random variables with shape parameter $\nu = 2$ and scale parameter $1/2$

according to the linear transformation method described in [30]. The covariance function between the kth and lth element of the sequence is set as $\rho^{|k-l|}$ with $\rho = 0.75$. Fig. 2 shows a sample run containing a total of 150 samples. The quantization values κ_q for five-level quantization and the quantization borders are also shown in the figure. An examination of this figure reveals that pairs of consecutive quantized samples exhibit a tendency to stay in the same or neighboring state. This is perhaps most clearly seen for the samples quantized to κ_5 value in the figure. The set of next-states visited when the current-state is κ_5 is exclusively limited κ_4 to and κ_5 in this Monte Carlo run. In fact, a large number of Monte Carlo runs indicate that this is the case 90% of the time. The main of goal of this article is to make use of this knowledge, that is the observed persistence of states, in the covariance matrix estimation. Furthermore, the suggested HMM-based model easily enables to empirical observations, such as "the state κ_5 occurs rarely; but when it occurs, it persists for a number samples" into the covariance matrix estimation setting.

V. NUMERICAL EXPERIMENTS

We present performance comparisons for the suggested method with two numerical experiments. In the first experiment, the covariance matrix of an AR(1) process is estimated from noiseless and noisy snapshots. The comparison metric is the Frobenius norm[2] $||A||_F = \sqrt{\text{tr}(\mathbf{A}^T\mathbf{A})}$. In the second experiment, we explore the significance of the proposed method for the adaptive target detection application.

In both experiments, we present comparisons with the Tyler's estimator [1] and the clairvoyant estimator. The clairvoyant estimator is assumed to have access to the noiseless snapshots and the true value of the texture parameter for each snapshot and can be expressed for general and persymmetric matrices as follows:

$$\widehat{\mathbf{M}}^{\text{clair}} = \frac{1}{L}\sum_{l=1}^{L}\frac{1}{\tau_l}\mathbf{x}_l\mathbf{x}_l^H$$

$$\widehat{\mathbf{M}}^{\text{clair}}_{\text{persym}} = \frac{1}{2L}\sum_{l=1}^{L}\frac{1}{r_l}\left(\mathbf{x}_l\mathbf{x}_l^H + \mathbf{x}_l^R\left(\mathbf{x}_l^R\right)^H\right). \quad (21)$$

Note that the results for the clairvoyant estimator should be interpreted as a performance bound for other estimators. Tyler's estimator and its persymmetric version are described in the Appendix of this study.

[2]We have also made comparisons with the metric proposed by Förstner, $d(\mathbf{M}, \widehat{\mathbf{M}}) = ||\ln(\mathbf{M}^{-1/2}\widehat{\mathbf{M}}\mathbf{M}^{-1/2})||_F = \sqrt{\sum_i \ln^2(\lambda_i)}$ (λ_i are the generalized eigenvalues of $\mathbf{M}$ and $\widehat{\mathbf{M}}$) [31]. Förstner's metric is designed for the positive definite matrix comparisons and it is affine invariant, i.e., invariant to a measurement unit change, say from meters to kilometers. We only report the results for more commonly used Frobenius norm; since the results for both metrics are similar.

A. Experiment 1: Estimation of AR(1) Autocorrelation Matrix

In this case, the unknown covariance matrix $\mathbf{M}$ is taken as a Toeplitz matrix with the first row entries $[1\ \rho\ \rho^2 \dots\ \rho^{N-1}]$. This matrix corresponds the covariance matrix of a stationary AR(1) random process having the autocorrelation sequence $r[k] = \rho^{|k|}$. The texture parameter of this experiment is assumed to be quantized to five-levels $\tau_k \in \{1/100, 1/10, 1, 10, 100\}$. In this experiment, the probability of preserving the same state is 0.9, $\pi_{ii} = 0.9$; the transitions to other states are equally probable, i.e. $\pi_{ij} = 0.025, i \neq j$. Hence, a state is maintained, i.e., no state transitions, on the average of $1/(1-\pi_{ii}) = 10$ consecutive observations.

Fig. 4(a) shows the results for $N = 10$ dimensional covariance matrix estimation for the noiseless case. Plots with dashed lines in all figures show the results when the persymmetry constraint is taken into account. [For the stationary AR(1) process, the unknown covariance matrix is Toeplitz; hence, satisfies the persymmetry condition, as explained in the Appendix.] It can be observed from Fig. 4(a) that the inclusion of persymmetry constraint brings significant gains. The suggested method can be noted to closely track the clairvoyant estimator at almost all L values. When compared with the Tyler's estimator, the suggested scheme achieves the performance of the Tyler's estimator with 5–10 fewer number of snapshots. This result can be significant in applications where the number of snapshots is scarce.

The results of the same experiment in the presence of noise is given in Fig. 4(b). Here, the snapshots are assumed to be contaminated with white-Gaussian noise with variance 1/SNR. Note that in this experiment, the signal power is normalized to 1, that is $r[k]|_{k=0} = 1$. Fig. 4(b) shows that the performance of suggested method approaches to the performance of the clairvoyant estimator with increasing SNR.

B. Experiment 2: Adaptive Radar Detection Application

We study the SNR loss due to imperfect covariance matrix estimation within the context of adaptive target detection. We consider an X-band (wavelength of $\lambda = 3$ cm) pulse Doppler radar system operating at a constant pulse repetition frequency (PRF) of 3 kHz, transmitting $N = 10$ pulses. The system operates under the influence of sea clutter with mean velocity $\mu_c = 3$ m/s and spread $\sigma_c = 1$ m/s. The clutter-to-noise ratio (CNR) is 10 dB, noise variance is $\sigma_n^2 = 1$.

The clutter power spectral density is assumed to be in Gaussian form leading to the clutter autocorrelation of $r_c[k] = \text{CNR}\, e^{j2\pi f_d k}\rho^{k^2}$ with $f_d = 2\mu_c/(\lambda\,\text{PRF}) = 1/15$ and $\rho = \exp(-8\pi^2\sigma_c^2/(\lambda\,\text{PRF})^2) \approx 0.990$. The kth row, lth column entry of the clutter covariance matrix $\mathbf{M}_0$ is then $r_c[k-l]$. As in earlier experiment, the texture parameter is quantized to five levels, $\kappa_s \in \{-20, -10, 0, 10, 20\}$ dB; the state transition probabilities are set as $\pi_{ii} = 0.9$ and

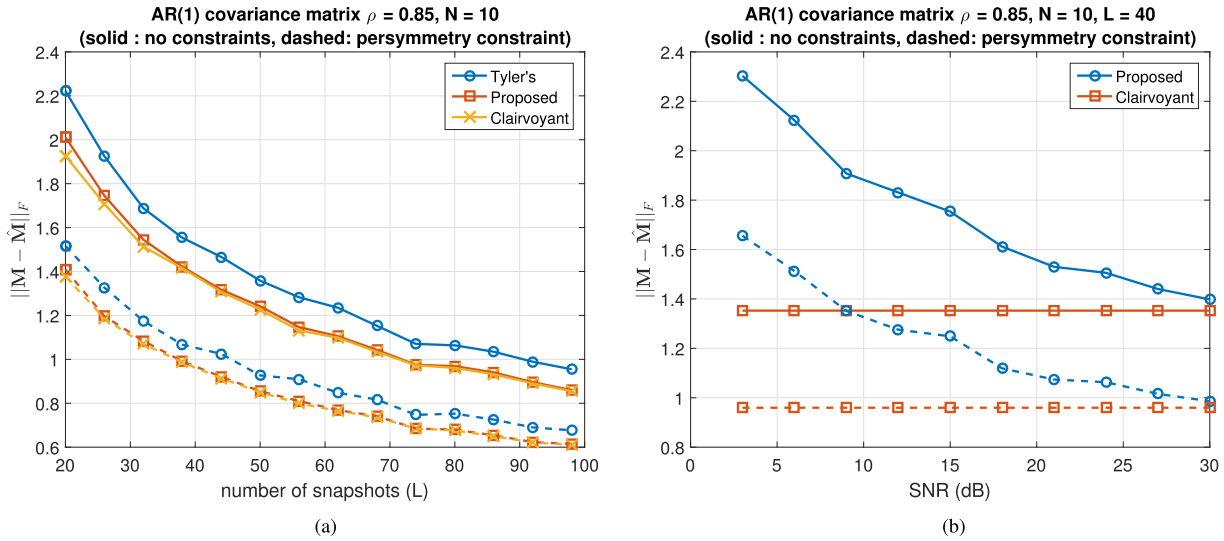


Fig. 4. Experiment 1: AR(1) covariance matrix estimation. (a) Noiseless snapshots. (b) Noisy snapshots.

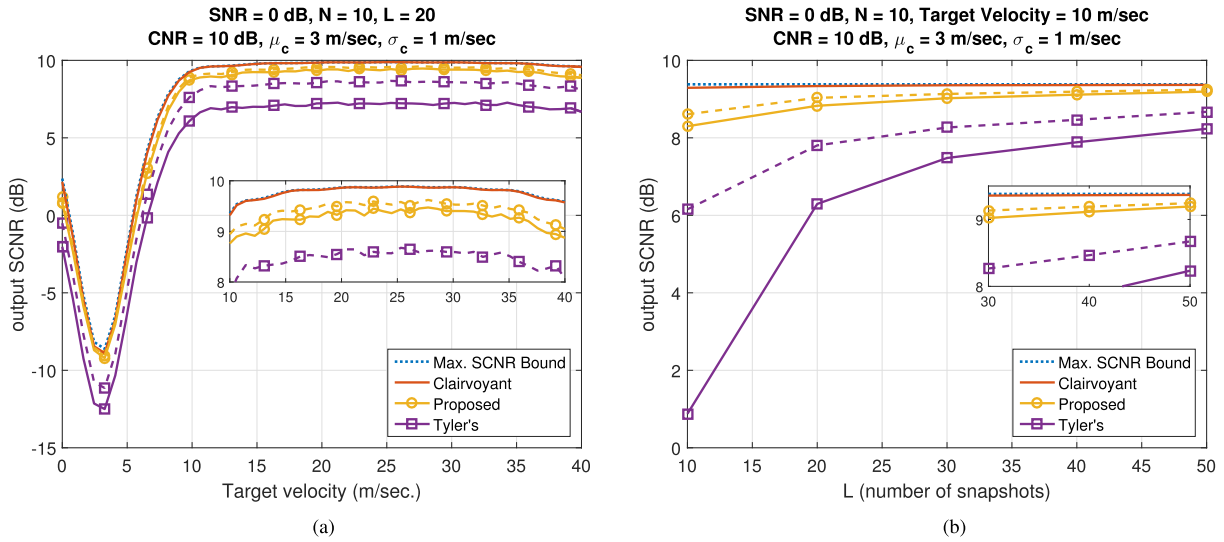


Fig. 5. Experiment 2: Average output SCNR comparison. Dashed and solid lines correspond to estimators with and without persymmetry constraint, respectively. (a) SCNR versus Target Velocity. (b) SCNR versus number of snapshots.

$\pi_{ij} = 0.025\ i \neq j$. These numerical values are also roughly inline with the measurements in [17, Figs. 2 and 3].

In this experiment, we present a comparison for the average output signal-to-clutter-and-noise ratio (SCNR), $\text{SCNR}_{\text{out}}(\omega) = |\mathbf{w}^H \mathbf{s}_\omega|^2 / \mathbf{w}^H (\mathbf{M} + \sigma_n^2 \mathbf{I}) \mathbf{w}$, for two-stage adaptive detector (AMF) with the linear combination weight $\mathbf{w} = (\widehat{\mathbf{M}} + \sigma_n^2 \mathbf{I})^{-1} \mathbf{s}_\omega$ [11]. Here $\mathbf{s}_\omega = \sqrt{\text{SNR}}[1\ e^{j\omega}\ \ldots\ e^{j(N-1)\omega}]^T$ is the Doppler steering vector of the target with velocity v_t and $\omega = 4\pi v_t / (\lambda\, \text{PRF})$ is the phase progression due to the Doppler effect. In this comparison, we assume that $\widehat{\mathbf{M}}$ is generated from an observation of L snapshots via different methods and compare the achievable output SCNR for each method. We also present the maximum SCNR bound as a comparison metric. This bound is achieved when true $\mathbf{M}$ (instead of its estimate) is utilized in the weight vector calculation, $\text{SCNR}^{\max} = \mathbf{s}_\omega^H (\mathbf{M} + \sigma_n^2 \mathbf{I})^{-1} \mathbf{s}_\omega$.

Fig. 5(a) shows average SCNR attained with $L = 20$ snapshots at various target velocities. The proposed detector operates 0.3 dB (with persymmetry constraint) and 0.5 dB (without persymmetry constraint) below the performance bounds (clairvoyant detector and SCNR bound) at target velocity of 10 m/s. The detector with Tyler's estimate has a loss around 1.5 dB (with persymmetry constraint) and 3.1 dB (without persymmetry constraint) and the SNR loss is not recoverable even at high target velocities. Note that, in this experiment the Tyler's estimator does not estimate the clutter covariance matrix $\mathbf{M}$; but the interference matrix (clutter plus noise covariance matrix), that is $\mathbf{M} + \sigma_n^2 \mathbf{I}$. Fig. 5(b) presents the results when target velocity is fixed to 10 m/s and secondary data size (L) is varied. The SNR loss further increases for small number of snapshots.

1) K-distributed Clutter Experiment: This experiment examines the case of model mismatch which is the case

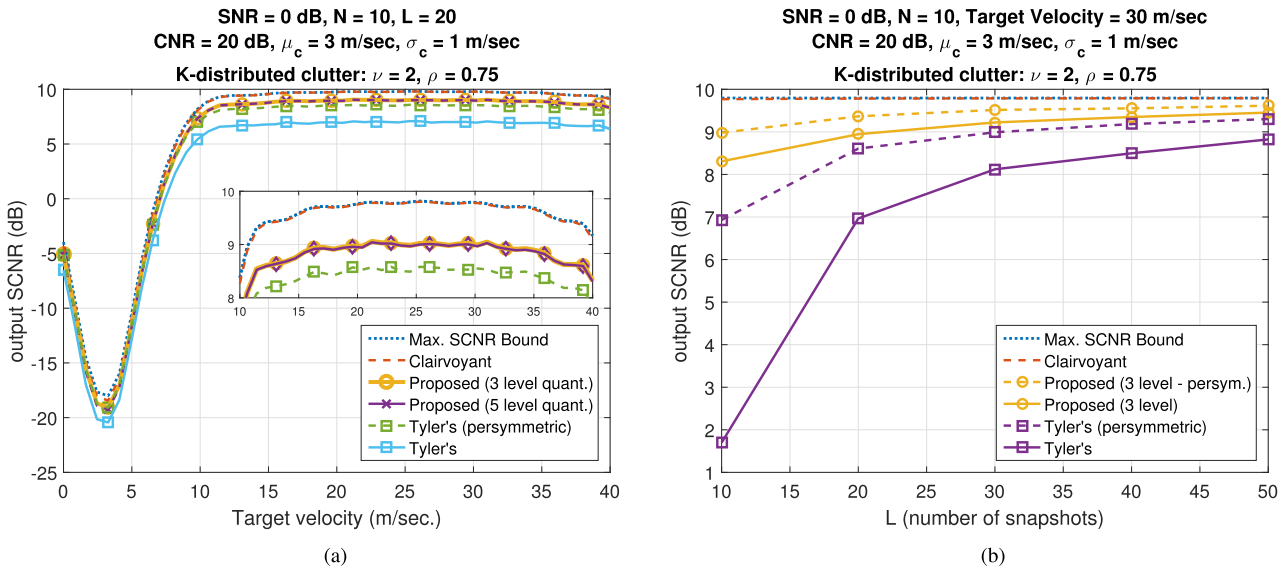


Fig. 6. Experiment 2 (K-distributed clutter). Average output SCNR comparison. Dashed and solid lines correspond to estimators with and without persymmetry constraint, respectively. (a) SCNR versus Target Velocity. (b) SCNR versus number of snapshots.

in general for seaborne/airborne adaptive target detectors. The clutter is taken as K-distributed. That is, the texture parameter is gamma distributed. The shape parameter of the distribution is set as $\nu = 2$. The texture correlation is $r_\tau[k] = \rho^{|k|}$ with $\rho = 0.75$. Fig. 2 shows a realization of gamma variable sequence generated with these parameters.

We follow the procedure given in Section IV to construct the Markov model and use both three and five level quantization values given in Table I for $\nu = 2$ in this comparison. Fig. 6 shows the results for $N = 10$ and $L = 20$. The other experiment parameters are given in the figure title.

We observe from Fig. 6(a) that the performance of the proposed method for the quantization levels of three and five are almost identical. The performance plots for the proposed method in Fig. 6(a) do not assume the persymmetry condition. The effect of persymmetry is examined in Fig. 6(b). From Fig. 6(a), we see that the SNCR loss for the velocity of 30 m/s (high velocity region) when compared with SCNR bound or clairvoyant estimator is 0.8 dB for the proposed estimators (without persymmetry); while it is 1.2 and 2.7 dB for the Tyler's estimator with and without persymmetry constraint, respectively.

Fig. 6(b) gives the results for a fixed velocity of 30 m/s for different number of snapshots. In this figure, only three-level quantization results are shown for figure clarity; since five-level quantization results are almost identical. In this figure, the results for the proposed method with and without persymmetry constraint are given. We see that for $L = 20$ case, the gap to the SCNR bound further reduces to 0.4 dB (with persymmetry constraint) from 0.8 dB (without persymmetry constraint). Hence the proposed method (with persymmetry constraint) brings an SNR improvement of 0.8 dB over the Tyler's method (with persymmetry constraint). Or stated differently, the Tyler's estimator requires $L = 50$ snapshots to operate at the output SCNR level of the proposed method with $L = 20$ snapshots. Yet, we would like to remind that in many adaptive radar detection applications, the number of snapshots can be limited by other factors, such as clutter discretes, other targets, etc. Hence, utilizing fewest possible snapshots can be mandated by other considerations. It can be seen from Fig. 6(b) that for $L = 10$, the gap between proposed and the Tyler's method is almost 2 dB.

VI. CONCLUSION

In this article, we have examined the effect of texture correlation in the covariance matrix estimation problem for compound-Gaussian vectors. To the best of our knowledge, such a study is not available in literature, with the notable exception of [21], in spite of experimental observations of clutter power spatial correlation especially in seaborne data. Numerical results have shown that the additional information on the spatial correlation can yield critical performance improvements both in covariance matrix estimation and adaptive target detection applications for small number of snapshots. These findings are especially important for adaptive radar detection applications, where the number of snapshots can be naturally very limited due to the constraints of the radar system and environment.

The study can be extended toward two different venues: (1) The first one is the HMM-based signal model given in this study can be further examined under the general title of the Markov modulated processes, as in [32]. The second venue is the reformulation of the problem with a suitable conjugate prior for the texture variables enabling closed form posterior density solutions other than the one given with the HMM structure. Both venues can yield interesting results that may be worth some further research effort.

Finally, a ready-to-use MATLAB implementation of the algorithm in Table I, including the persymmetric case, is provided in [25].

APPENDIX

A. Persymmetric Matrices, Tyler's Covariance Matrix Estimator

1) Persymmetric Covariance Matrices: A matrix is called persymmetric if it has symmetry across its antidiagonal. Stated differently, matrices that satisfy $\mathbf{A} = \mathbf{J}\mathbf{A}^T\mathbf{J}$ with the exchange matrix $\mathbf{J}$ are called persymmetric. Exchange matrix is a permutation matrix which reverses the order of entries in a vector. ($\mathbf{J} = [\mathbf{e}_N\, \mathbf{e}_{N-1}\, \ldots\, \mathbf{e}_1]$ and $\mathbf{e}_k$ is the kth column of $N \times N$ dimensional identity matrix.) Toeplitz matrices having constant values on the antidiagonals, $\mathbf{T}_{k,k+i} = c_i$, are also persymmetric.

Covariance matrices are, by definition, Hermitian-symmetric. Covariance matrices for stationary processes are also Toeplitz, hence they are both Hermitian-symmetric ($\mathbf{R} = \mathbf{R}^H$) and persymmetric ($\mathbf{R} = \mathbf{J}\mathbf{R}^T\mathbf{J}$). Combining two symmetry relations, we have $\mathbf{R}^H = \mathbf{J}\mathbf{R}^T\mathbf{J}$ which is the definition for centrohermitian matrices. Centrohermitian matrices are conjugate-symmetric across its center entry. As an example, a 3×3 centrohermitian matrix $\mathbf{C}$ has elements which are Hermitian symmetric about $\mathbf{C}_{2,2}$ entry, the central entry. Note that the centrosymmetry condition $\mathbf{R}^H = \mathbf{J}\mathbf{R}^T\mathbf{J}$ can also be written as $\mathbf{R} = \mathbf{J}\mathbf{R}^*\mathbf{J}$, where $\mathbf{R}^*$ is the complex-conjugate of the matrix $\mathbf{R}$.

2) Tyler's Estimator: We give brief information on the Tyler's covariance matrix estimator and its persymmetric version. The Tyler's estimator can be derived as the ML estimator for the compound-Gaussian vectors with unknown deterministic texture parameters (nuisance parameters for the covariance matrix estimation problem).

The snapshots $\mathbf{x}_k = \sqrt{\tau}_k \mathbf{c}_k$ $k = \{1, \ldots, L\}$ (with deterministic texture parameters τ_k and iid Gaussian distributed speckle component $\mathbf{c}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{M})$) are jointly Gaussian distributed $\mathbf{x}_k \sim \mathcal{CN}(\mathbf{0}, \tau_k\mathbf{M})$ with the log-likelihood function

$$\Lambda(\mathbf{M}, \tau_1, \ldots, \tau_L) \stackrel{c}{=} -L\log(|\mathbf{M}|) - \sum_{k=1}^{L}\left(N\log(\tau_k) + \frac{1}{\tau_k}\mathbf{x}_k^H\mathbf{M}^{-1}\mathbf{x}_k\right). \tag{22}$$

Maximizing the log-likelihood expression over τ_k yields $\tau_k^{(\text{opt})} = \mathbf{x}_k^H\mathbf{M}^{-1}\mathbf{x}_k/N$. After the substitution of the optimized values into (22), we get the compressed log-likelihood expression

$$\Lambda\left(\mathbf{M}, \tau_1^{(\text{opt})}, \ldots, \tau_L^{(\text{opt})}\right) \stackrel{c}{=} -L\log(|\mathbf{M}|) - N\sum_{k=1}^{L}\log\left(\mathbf{x}_k^H\mathbf{M}^{-1}\mathbf{x}_k\right).$$

Optimizing the compressed log-likelihood over $\mathbf{M}$, we get the fixed-point relation defining the Tyler's estimator

$$\mathbf{M}_{\text{Tyler}} = \frac{N}{L}\sum_{k=1}^{L}\frac{\mathbf{x}_k\mathbf{x}_k^H}{\mathbf{x}_k^H\mathbf{M}^{-1}\mathbf{x}_k}. \tag{23}$$

For the case of persymmetric matrices, we have the identity of $\mathbf{x}_k^H\mathbf{M}^{-1}\mathbf{x}_k = (\mathbf{x}_k^R)^H\mathbf{M}^{-1}\mathbf{x}_k^R$, where $\mathbf{x}_k^R = \mathbf{J}\mathbf{x}_k^*$ denotes the reversed (flipped up-down) and conjugated vector $\mathbf{x}_k$. Using this identity, the log-likelihood expression in (22), can be rewritten for persymmetric matrices as

$$\Lambda(\mathbf{M}, \tau_1, \ldots, \tau_L) \stackrel{c}{=} -L\log(|\mathbf{M}|) - \sum_{k=1}^{L}\left(N\log(\tau_k) + \frac{1}{2\tau_k}\left(\mathbf{x}_k^H\mathbf{M}^{-1}\mathbf{x}_k + \left(\mathbf{x}_k^R\right)^H\mathbf{M}^{-1}\mathbf{x}_k^R\right)\right). \tag{24}$$

Starting from (24) and repeating the same steps, we get the persymmetric version of the Tyler's estimator

$$\mathbf{M}_{\text{persym}} = \frac{N}{L}\sum_{k=1}^{L}\frac{\mathbf{x}_k\mathbf{x}_k^H + \mathbf{x}_k^R\left(\mathbf{x}_k^R\right)^H}{\mathbf{x}_k^H\mathbf{M}_{\text{persym}}^{-1}\mathbf{x}_k + \left(\mathbf{x}_k^R\right)^H\mathbf{M}_{\text{persym}}^{-1}\mathbf{x}_k^R}. \tag{25}$$

For more information, see [33], [34], and [9].

REFERENCES

[1] D. E. Tyler, "A distribution-free m-estimator of multivariate scatter," *Ann. Statist.*, vol. 15, no. 1, pp. 234–251, 1987.

[2] 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, 2002.

[3] P. J. Huber, "Robust estimation of a location parameter," *Ann. Math. Statist.*, vol. 35, no. 1, pp. 73–101, 1964. [Online]. Available: http://www.jstor.org/stable/2238020

[4] R. A. Maronna, "Robust m-estimators of multivariate location and scatter," *Ann. Statist.*, vol. 4, no. 1, pp. 51–67, 1976. [Online]. Available: http://www.jstor.org/stable/2957994

[5] Y. I. Abramovich, B. A. Johnson, and N. K. Spencer, "Sample-deficient adaptive detection: Adaptive scalar thresholding versus CFAR detector performance," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 46, no. 1, pp. 32–46, Jan. 2010.

[6] Y. Sun, A. Breloy, P. Babu, D. P. Palomar, F. Pascal, and G. Ginolhac, "Low-complexity algorithms for low rank clutter parameters estimation in radar systems," *IEEE Trans. Signal Process.*, vol. 64, no. 8, pp. 1986–1998, Apr. 2016.

[7] A. Breloy, G. Ginolhac, F. Pascal, and P. Forster, "Robust covariance matrix estimation in heterogeneous low rank context," *IEEE Trans. Signal Process.*, vol. 64, no. 22, pp. 5794–5806, Nov. 2016.

[8] I. Soloveychik and A. Wiesel, "Tyler's covariance matrix estimator in elliptical models with convex structure," *IEEE Trans. Signal Process.*, vol. 62, no. 20, pp. 5251–5259, Oct. 2014.

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

[10] F. Bandiera, O. Besson, and G. Ricci, "Knowledge-aided covariance matrix estimation and adaptive detection in compound-Gaussian noise," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5391–5396, Oct. 2010.

[11] F. Bandiera, D. Orlando, and G. Ricci, *Advanced Radar Detection Schemes Under Mismatched Signal Models*. San Rafael, CA, USA: Morgan & Claypool, 2009.

[12] A. D. Maio and M. S. Greco, *Modern Radar Detection Theory*. Rijeka, Croatia: Scitech Publishing, 2015.

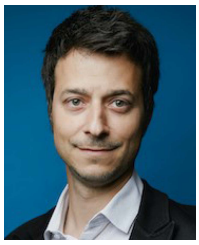
[13] E. J. Kelly, "An adaptive detection algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-22, no. 2, pp. 115–127, Mar. 1986.

[14] E. Conte, M. Lops, and G. Ricci, "Asymptotically optimum radar detection in compound-Gaussian clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, no. 2, pp. 617–625, Apr. 1995.
[15] J.-P. Ovarlez, F. Pascal, and P. Forster, "Covariance matrix estimation in SIRV and elliptical processes and their application in radar detection," in *Modern Radar Detection Theory*, A. D. Maio and M. S. Greco, Eds. Rijeka, Croatia: Scitech Publishing, 2015, ch. 8, pp. 295–327.
[16] K. Ward, R. Tough, and S. Watts, *Sea Clutter: Scattering, the K Distribution and Radar Performance*. London, U.K.: Institution Eng. Technol., 2013.
[17] S. Watts and K. D. Ward, "Spatial correlation in K-distributed sea clutter," *IEE Proc. F., Commun., Radar Signal Process*., vol. 134, no. 6, pp. 526–532, Oct. 1987.
[18] F. Gini, M. V. Greco, M. Diani, and L. Verrazzani, "Performance analysis of two adaptive radar detectors against non-Gaussian real sea clutter data," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, no. 4, pp. 1429–1439, Oct. 2000.
[19] A. D. Maio, G. Foglia, E. Conte, and A. Farina, "CFAR behavior of adaptive detectors: An experimental analysis," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 41, no. 1, pp. 233–251, Jan. 2005.
[20] M. Greco, F. Gini, and M. Rangaswamy, "Statistical analysis of measured polarimetric clutter data at different range resolutions," *IEE Proc.-Radar, Sonar Navigation*, vol. 153, pp. 473–481, Dec. 2006.
[21] R. Couillet, M. S. Greco, J.-P. Ovarlez, and F. Pascal, "RMT for whitening space correlation and applications to radar detection," in *Proc. IEEE 6th Int. Workshop Comput. Adv. Multi-Sensor Adaptive Process.*, 2015, pp. 149–152.
[22] R. S. Raghavan, "Statistical interpretation of a data adaptive clutter subspace estimation algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 2, pp. 1370–1384, Apr. 2012.
[23] B. C. Levy, *Principles of Signal Detection and Parameter Estimation*, 1st ed. New York, NY, USA: Springer, 2008.
[24] D. Barber and A. T. Cemgil, "Graphical models for time-series," *IEEE Signal Process. Mag.*, vol. 27, no. 6, pp. 18–28, Nov. 2010.
[25] C. Candan and F. Pascal, "Covariance matrix estimation of texture correlated compound-Gaussian vectors for adaptive radar detection," *MATLAB Code Files*, 2022. [Online]. Available: https://doi.org/10.24433/CO.6015307.v1
[26] A. P. Shikhaliev, L. C. Potter, and Y. Chi, "Low-rank structured covariance matrix estimation," *IEEE Signal Process. Lett.*, vol. 26, no. 5, pp. 700–704, May 2019.
[27] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, Feb. 1989.
[28] S. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 129–137, Mar. 1982.
[29] J. Max, "Quantizing for minimum distortion," *IRE Trans. Inf. Theory*, vol. 6, no. 1, pp. 7–12, 1960.
[30] Y. Dong, L. Rosenberg, and G. Weinberg, "Generating correlated gamma sequences for sea-clutter simulation," Defence Science and Technology Organisation, Canberr, ACT, Australia, Tech. Rep. DSTO–TR-2688, 2012.
[31] W. Förstner and B. Moonen, "A metric for covariance matrices," in *Geodesy-The Challenge of the 3rd Millennium*, Dr E. W. Grafarend, Dr F. W. Krumm, Dr V. S. Schwarze, F. Krumm, and V. S. Schwarze eds. Berlin, Heidelberg: Springer, 2003, pp. 113–128.
[32] Y. Ephraim and W. J. J. Roberts, "An EM algorithm for Markov modulated Markov processes," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 463–470, Feb. 2009.
[33] M. Jansson and P. Stoica, "Forward-only and forward-backward sample covariances—A comparative study," *Elsevier Signal Process.*, vol. 77, no. 3, pp. 235–245, 1999.
[34] E. Conte and A. De Maio, "Exploiting persymmetry for CFAR detection in compound-Gaussian clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 2, pp. 719–724, Apr. 2003.

Çağatay Candan has received the B.S. degree from Middle East Technical University, Ankara, Turkey, in 1996, and the M.S. degree from Bilkent University, Ankara, Turkey, in 1998, both in electrical and electronics engineering, and Ph.D. degree in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2004.

He is currently a Professor with the Department of Electrical and Electronics Engineering, Middle East Technical University, Ankara, Turkey. His research interests include statistical signal processing and its applications.

Dr. Candan was the recipient of the Research Encouragement Award from TUBITAK in 2015. He has been serving as an Associate Editor for the IEEE COMMUNICATIONS LETTERS and *IEEE Signal Processing Magazine*, since 2021.



Frédéric Pascal (Senior Member, IEEE) received the M.Sc. degree in 2003, the Ph.D. degree in 2006, and the HDR degree in 2012.

He is a Full Professor of AI with L2S lab, CentraleSupélec, University Paris-Saclay, Gif-sur-Yvette, France. Between March 2008 and December 2011 (resp. Jan. 2012 and Dec. 2013), he was an Assistant Professor (resp. Associate Professor) with SONDRA, (joint French-Singaporean lab.) CentraleSupélec. Between August 2013 and 2014, he was a Visiting Associate Professor with the ECE department, National University of Singapore, Singapore. Between January 2017 and December 2018, he was the head of the "Signals and Statistics" group of L2S. Since December 2019, he has been the Coordinator of activities in artificial intelligence with CentraleSupélec, and also the Chairholder of the Givaudan Chair on Data Sciences. He has authored/coauthored more than one hundred papers in the top journals and conferences in signal processing, image processing, and statistics. His research interests include estimation, detection and classification for statistical signal processing and applications in radar and image processing.

Dr. Pascal served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, from 2015 to 2018, for the EURASIP *Journal on Advances in Signal Processing*, from 2015 to 2021, and for Elsevier *Signal Processing*, from 2018 to 2021. Since September 2017, he is in the Executive Committee of the DATAIA institute as the Program Coordinator, and he was appointed as Co-Director of the institute, in April 2021. From 2015 to 2017, Frederic Pascal was the Chair of the EURASIP SAT in Theoretical and Methodological Trends in Signal Processing (TMTSP) and he is a member of the IEEE Signal Processing Society SAM technical committee, from January 2015 to present. Between 2019 and 2021, he is the Vice Chair of the Data Science Initiative of the IEEE Signal Processing Society.