IET Radar, Sonar & Navigation

Research Article

IET Journals
The Institution of Engineering and Technology

Covariance-free non-homogeneity STAP detector in compound Gaussian clutter based on robust statistics

ISSN 1751-8784
Received on 27th April 2019
Revised 15th July 2019
Accepted on 9th August 2019
E-First on 8th November 2019
doi: 10.1049/iet-rsn.2019.0201
www.ietdl.org

Ahmed A. Abouelfadl[1] ✉, Ioannis Psaromiligkos[1], Benoit Champagne[1]

[1]*Department of Electrical and Computer Engineering, McGill University, Montreal, QC, Canada*

✉ *E-mail: ahmed.abouelfadl@mail.mcgill.ca*

Abstract: Space-time adaptive processing (STAP) detects targets by computing adaptive weight vectors for each cell under test using its covariance matrix, as estimated from surrounding secondary cells. In this study, the non-homogeneity detector (NHD) excludes the anomalous secondary cells that adversely affect the detection performance. The existing robust NHDs require estimating the covariance matrix of each secondary cell, which hinders their implementation in modern radars with large-dimensional range cells. In this study, the authors propose a new low-complexity NHD that is suitable for highly correlated clutter environments with both Gaussian and non-Gaussian heavy-tailed distributions. The proposed detector, which is based on the projection depth function from the field of robust statistics, features a non-parametric and covariance-free test statistic. As a result, its computational complexity is much lower than that of current NHDs, such as the widely used normalised adaptive matched filter (NAMF) detector, especially for large-dimensional range cells. In Monte Carlo simulations with different clutter distributions and radar system configurations, the proposed detector shows comparable performance to that of NAMF. The low complexity and robust performance of the new detector make it particularly attractive for real-time applications.

1 Introduction

Modern radar systems can jointly estimate the range, speed, and direction of a detected target. The range and speed can be measured based on the time difference and Doppler frequency shift between the transmitted signal and the target return received by the radar, respectively. For angle measurement, phased array radars employ antenna arrays and beamforming techniques [1]. More specifically, following radio-frequency down-conversion, the range information is obtained by sampling the received antenna signals within a temporal window in the pulse repetition interval (PRI). To estimate the target Doppler frequency, coherent pulse-Doppler radars transmit a sequence of identical time-shifted pulses that together define the coherent pulse interval (CPI). The use of an antenna array makes it possible to measure phase differences between different spatial receiving channels, which after sampling result in a set of directions [2].

Upon reflection of the transmitted pulses by a target, the radar antenna receives distorted versions of these pulses due to other scatterers, clutter, and noise. A space-time adaptive processing (STAP) detector discretely scans the range dimension and, for each range bin, arranges the data along the angle and Doppler dimensions into a vector, called a range cell. It then linearly combines the spatio-temporal data in each range cell to form the test statistics. To this end, it needs to compute a set of weight vectors corresponding to the different spatio-temporal 'look' directions, which depend on the covariance matrix of the background clutter and noise within the cell under test (CUT), also called the primary cell [2]. However, this covariance matrix is not known in practice and it is commonly estimated from the adjacent range cells, known as the secondary or training cells in this context.

The estimation of the covariance matrix from the secondary cells relies on the assumption that they are homogeneous, i.e. independent and identically distributed (iid). In reality, the homogeneity assumption is hardly met due to the presence of discrete scatterers, in-band interferers, target-dependent jammers [3, 4], or a combination thereof. In this case, the estimated covariance matrix does not represent accurately the background clutter and noise, and hence, the weight vectors computed from this matrix lead to a performance degradation of the STAP detector. To tackle this problem, the non-homogeneity detector (NHD) was introduced to detect the anomalous secondary cells to be censored from covariance matrix estimation [5].

Conceptually, a secondary cell is considered to be homogeneous to its surrounding secondary cells if it shares with them the same covariance matrix up to a scalar. Since the true covariance matrix of a given secondary cell is unknown, the work in [5] used the generalised inner product (GIP) test to examine the similarity between this unknown covariance matrix and the test covariance matrix estimated from the surrounding secondary cells. Later, the normalised adaptive matched filter (NAMF) test was used as an NHD with Gaussian and non-Gaussian clutter models in [6], where the NAMF detector was shown to be the most robust NHD.

Recent research efforts on NHD have focused on improving the performance of the aforementioned classical detectors or reducing their complexity. For instance, Jiang and Wang [7] proposed a soft NHD concept, wherein the covariance matrix of the CUT is calculated using the weighted secondary cells assuming Gaussian distributed clutter. In turn, the calculation of weight for each secondary cell is formulated as a non-linear optimisation problem based on the output of a modified version of the adaptive matched filter (AMF). An iterative approximate maximum-likelihood (ML) approach based on the GIP detector was developed in [8] for estimating the subset of non-homogeneous cells. This approach shows a comparable performance to the iterative original GIP test using the ML covariance estimator for the Gaussian interference [9]. In addition, a large body of research has been devoted to reduce the dimensions of the STAP detection problem using different transformation and rank reduction techniques as in [10], which can also be applied to NHD [11]. However, these partially adaptive detectors generally exhibit inferior performance compared to their fully adaptive counterparts. Based on the GIP detector, other NHD procedures for the special cases of spaceborne or side-looking radars were introduced in [12, 13].

The above referenced covariance-based NHDs share the need to estimate the covariance matrix and its inverse (known as the precision matrix) for each secondary cell, which leads to a high computational cost, especially for non-Gaussian clutter. Some covariance estimators need a priori knowledge about the clutter distribution [14], which is imperfect in most cases, while other estimators need to solve non-convex optimisation problems with high computational complexity [15]. To avoid such difficulties, a

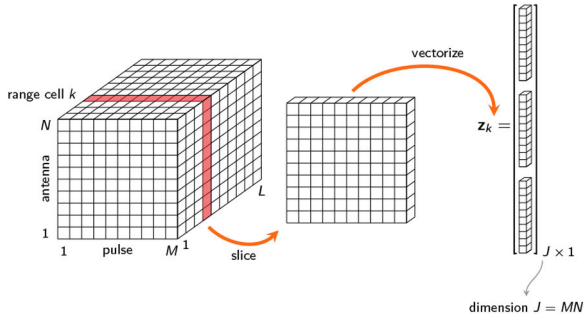

Fig. 1 *Illustration of STAP data 'cube' and formation of range cells*

covariance-free NHD with a comparable performance to the GIP was introduced in [16]. However, it is known that the GIP test is not robust, especially in non-Gaussian clutter scenarios [6].

In this paper, we focus on the problem of detecting the non-homogeneous cells within the secondary cells for correlated clutter with Gaussian as well as non-Gaussian distributions. We introduce a novel covariance-free, non-parametric NHD based on the projection depth (PD) function, a well-known tool in the field of robust statistics [17]. By exploiting the properties of this function, the new detector does not require a priori knowledge about the clutter distribution, and more importantly, evades estimating the covariance or the precision matrices, which reduces the computational burden significantly. Simulation results for different clutter distributions show that the proposed detector can maintain the robust performance of covariance-based, fully-adaptive detectors such as the NAMF while inheriting the simple computational structure of the PD function. Preliminary results of this work were presented in an abridged conference paper [18]. This paper provides new results and extensions including: proofs of key results that are central to the derivation of the new detector, detailed performance analysis including Fisher-consistency under different clutter models, guidelines on the choice of the proposed NHD parameters with complexity analysis, and extended simulation studies for different radar operational conditions and a more challenging benchmark.

The rest of the paper is organised as follows: Section 2 provides the background about STAP and the clutter signal model. The non-homogeneity detection problem is presented in Section 3. In Section 4, the proposed NHD is introduced and its approximate equivalence to the NAMF detector is proven. The comparative evaluation of the proposed and NAMF detectors through Monte Carlo simulation is presented in Section 5. Section 6 concludes the paper.

Notation: Matrices and column vectors are denoted in boldface uppercase and lowercase letters, respectively; $\boldsymbol{I}_K$ indicates a $K \times K$ identity matrix; $(\cdot)^{\mathrm{T}}$, $(\cdot)^{\mathrm{H}}$, $\otimes$, $\|\cdot\|_2$, and $\det(\cdot)$ denote the transpose, the Hermitian transpose, the Kronecker product, the ℓ_2-norm, and determinant operations, respectively. $E(\cdot)$ is the mathematical expectation; $\Re(\cdot)$ and $\Im(\cdot)$ denote the real and imaginary parts of their complex-valued arguments, respectively.

2 Signal model

In this section, we briefly review the STAP concept for pulse-Doppler radars and present a clutter model based on the spherically invariant random process (SIRP) that can represent both Gaussian and non-Gaussian clutters.

2.1 STAP model

Consider a pulsed Doppler radar using a uniform linear array (ULA) of N antenna elements that are spaced $d = \lambda/2$ apart, where λ is the wavelength at the radar's centre frequency. The radar simultaneously transmits from each antenna element a sequence of M coherent pulses with a PRI T, which define the so-called *slow* time domain. The transmitted signal from each antenna element is assumed to be narrowband, i.e. its bandwidth B satisfies $B \ll c/Nd$, where c is the speed of light [2].

Upon reflection by a moving point target with azimuth angle θ_{t} from the boresight of the radar antenna array (planar geometry is assumed), a target return will be received with a Doppler frequency shift f_{d}, normalised with respect to the pulse repetition frequency $1/T$. Let T_{u} be the time delay corresponding to the radar maximum unambiguous range, while the time delay corresponding to the radar range resolution is $1/B$. Hence, the total number of range cells in the so-called *fast* time domain is

$$L = \lfloor T_{\mathrm{u}} B \rfloor \tag{1}$$

where $\lfloor \cdot \rfloor$ denotes the floor function. Therefore, the STAP data can be visualised as an $L \times N \times M$ data 'cube' as shown in Fig. 1. If the target's range corresponds to the kth range cell, its spatio-temporal data is an $N \times M$ matrix that contains the received signal from each antenna element and PRI within the target's range cell.

The M-dimensional temporal and N-dimensional spatial steering vectors of the target are, respectively, given by [2]

$$\boldsymbol{b}(f_{\mathrm{d}}) = \left[1 \quad \mathrm{e}^{\mathrm{j}2\pi f_{\mathrm{d}}} \quad \cdots \quad \mathrm{e}^{\mathrm{j}2\pi(M-1)f_{\mathrm{d}}}\right]^{\mathrm{T}} \tag{2a}$$

$$\boldsymbol{a}(\theta_{\mathrm{t}}) = \left[1 \quad \mathrm{e}^{\mathrm{j}2\pi\frac{d}{\lambda}\sin(\theta_{\mathrm{t}})} \quad \cdots \quad \mathrm{e}^{\mathrm{j}2\pi\frac{(N-1)d}{\lambda}\sin(\theta_{\mathrm{t}})}\right]^{\mathrm{T}} \tag{2b}$$

The compound spatio-temporal steering vector $\boldsymbol{s}(f_{\mathrm{d}}, \theta_{\mathrm{t}}) \in \mathbb{C}^J$, where $J = MN$, is defined as

$$\boldsymbol{s}(f_{\mathrm{d}}, \theta_t) = \frac{\boldsymbol{b}(f_{\mathrm{d}}) \otimes \boldsymbol{a}(\theta_{\mathrm{t}})}{\| \boldsymbol{b}(f_{\mathrm{d}}) \otimes \boldsymbol{a}(\theta_{\mathrm{t}}) \|_2} \tag{3}$$

To simplify the presentation, $\boldsymbol{s}(f_{\mathrm{d}}, \theta_{\mathrm{t}})$ will be denoted as $\boldsymbol{s}$ hereinafter. The baseband signal $\boldsymbol{r}$ received from the target is

$$\boldsymbol{r} = a\boldsymbol{s} \tag{4}$$

where a is an unknown deterministic complex amplitude (i.e. Swerling case 0 [19]).

Depending on whether a target is present or not, the total received signal vector $\boldsymbol{z} \in \mathbb{C}^J$ in a given range cell is expressed as

$$H_1 : \boldsymbol{z} = \boldsymbol{r} + \boldsymbol{c} + \boldsymbol{n} \tag{5a}$$

$$H_0 : \boldsymbol{z} = \boldsymbol{c} + \boldsymbol{n} \tag{5b}$$

where H_0 and H_1 are the null and alternative hypotheses, respectively, $\boldsymbol{r}$ is the target return as in (4), $\boldsymbol{c}$ is the clutter vector and $\boldsymbol{n}$ is the noise vector; $\boldsymbol{n}$ and $\boldsymbol{c}$ are assumed to be statistically independent. The noise vector $\boldsymbol{n}$ is drawn from a complex circular symmetric Gaussian distribution $\mathcal{CN}(\boldsymbol{0}, \varsigma_n^2 \boldsymbol{I}_J)$ with zero mean and covariance matrix $\varsigma_n^2 \boldsymbol{I}_J$, where ς_n^2 is the noise variance. The clutter vector is modelled as [2]

$$\boldsymbol{c} = \sum_{i=0}^{N_c - 1} \kappa_i \boldsymbol{e}_i \tag{6}$$

where N_c is the number of clutter patches, κ_i is the complex amplitude of the ith patch and $\boldsymbol{e}_i \in \mathbb{C}^J$ is the corresponding steering vector, which admits the form of (3).

Let $\boldsymbol{R} = E(\boldsymbol{z}\boldsymbol{z}^H)$ be the covariance matrix of the received signal $\boldsymbol{z}$ in (5). For each CUT, STAP aims at forming the optimal beamforming (or weight) vector in real-time to maximise the received signal-to-interference-plus-noise ratio with respect to $\boldsymbol{s}$. Under the minimum variance distortionless response criterion, the optimal weight vector takes the form [2]

$$\boldsymbol{w} = g\boldsymbol{R}^{-1}\boldsymbol{s} \tag{7}$$

where g is a complex scalar.

For the complex vector $z = z_{\mathrm{R}} + jz_{\mathrm{I}}$, where $z_{\mathrm{R}} = \Re(z)$ and $z_{\mathrm{I}} = \Im(z)$, the covariance matrix is expressed as [20]

$$\boldsymbol{R} = \boldsymbol{R}_{z_{\mathrm{R}}z_{\mathrm{R}}} + \boldsymbol{R}_{z_{\mathrm{I}}z_{\mathrm{I}}} + j\left(\boldsymbol{R}_{z_{\mathrm{R}}z_{\mathrm{I}}}^T - \boldsymbol{R}_{z_{\mathrm{R}}z_{\mathrm{I}}}\right) \tag{8}$$

where $\boldsymbol{R}_{z_{\mathrm{R}}z_{\mathrm{R}}} = E(z_{\mathrm{R}}z_{\mathrm{R}}^T)$, $\boldsymbol{R}_{z_{\mathrm{R}}z_{\mathrm{I}}} = E(z_{\mathrm{R}}z_{\mathrm{I}}^T)$, $\boldsymbol{R}_{z_{\mathrm{I}}z_{\mathrm{I}}} = E(z_{\mathrm{I}}z_{\mathrm{I}}^T)$, and $\boldsymbol{z}$ is a proper complex signal, i.e. $\boldsymbol{R}_{z_{\mathrm{R}}z_{\mathrm{I}}} = -\boldsymbol{R}_{z_{\mathrm{R}}z_{\mathrm{I}}}^T$ and $\boldsymbol{R}_{z_{\mathrm{R}}z_{\mathrm{R}}} = \boldsymbol{R}_{z_{\mathrm{I}}z_{\mathrm{I}}}$, which is common in the radar context. Moreover, it is customary to assume that the in-phase and quadrature components of $\boldsymbol{z}$ are independent, i.e. $\boldsymbol{R}_{z_{\mathrm{R}}z_{\mathrm{I}}} = \boldsymbol{0}$, where $\boldsymbol{0}$ is $J \times J$ zero matrix [21, 22]. Hence

$$\boldsymbol{R} = 2\boldsymbol{R}_{z_{\mathrm{R}}z_{\mathrm{R}}} = 2\boldsymbol{R}_{z_{\mathrm{I}}z_{\mathrm{I}}} \tag{9}$$

In practice, the covariance matrix $\boldsymbol{R}$ is unknown and different techniques are used to estimate it from the adjacent $L-1$ secondary cells, assuming no guard cells. In the case of Gaussian clutter, the ML estimator is the sample covariance matrix (SCM) given below:

$$\hat{\boldsymbol{R}}_{\mathrm{SCM}} = \frac{1}{L-1}\sum_{l=1}^{L-1} z_l z_l^H \tag{10}$$

where z_l denotes the total received signal in the lth secondary cell, and the condition $L-1 \geq 2J$ is needed to ensure robustness. If $\boldsymbol{c}$ follows a non-Gaussian distribution, the SCM is neither a consistent nor robust estimator and other estimators should be used. More details on these estimators will be presented shortly.

2.2 SIRP clutter model

In this paper, we are concerned with coherent processing of the received signal vector $\boldsymbol{z}$ in (5), where both the real and imaginary parts (i.e. in-phase and quadrature components) of each vector entry are considered. In this regard, it is essential to employ a probabilistic model of the clutter vector $\boldsymbol{c}$ in (6) that takes into account the joint statistics of the real and imaginary parts of all its entries. In particular, for proper clutter modelling, both the spatio-temporal correlation properties and probability density function (PDF) of the clutter envelope should comply with experimental data. Under the SIRP model, which meets these requirements [23], the clutter vector is modelled as a product of two independent components, i.e. a zero-mean complex Gaussian vector, known as the speckle component, and a positive random variable, known as the texture component and assumed to vary slowly across range cells. Therefore, the clutter vector in (6) can be represented as

$$\boldsymbol{c} = v\boldsymbol{y}, \tag{11}$$

where $\boldsymbol{y} \in \mathbb{C}^J$ follows a complex Gaussian distribution $\mathscr{CN}(\boldsymbol{0}, \boldsymbol{\Sigma})$ with zero mean and covariance matrix $\boldsymbol{\Sigma}$, and v is a positive random variable. By choosing the proper PDF of the texture component v in the SIRP model (11), denoted as $f_V(v)$ in the sequel, we can obtain different non-Gaussian clutter distributions, also known as the compound Gaussian distributions, while the particular choice $v = 1$ (with probability one) yields the Gaussian clutter model. Moreover, through a suitable choice of the covariance matrix $\boldsymbol{\Sigma}$ of the Gaussian speckle vector $\boldsymbol{y}$, the desired spatio-temporal correlation properties can be fulfilled. The PDF of $\boldsymbol{c}$ can be expressed as [6]

$$f_{\boldsymbol{c}}(\boldsymbol{c}) = (2\pi)^{-J}\det(\boldsymbol{\Sigma})^{-1/2}h_{2J}(\boldsymbol{c}^H\boldsymbol{\Sigma}^{-1}\boldsymbol{c}) \tag{12}$$

where the function $h_{2J}(x)$ is defined as

$$h_{2J}(x) = \int_0^\infty v^{-J}\exp\left(-\frac{x}{v^2}\right)f_V(v)\mathrm{d}v\,. \tag{13}$$

The covariance matrix of the SIRP vector $\boldsymbol{c}$ is given by $\boldsymbol{R}_{\boldsymbol{c}} = E(v^2)\boldsymbol{\Sigma}$.

Other models for compound Gaussian clutter use zero memory non-linear (ZMNL) transformations. These methods apply non-linear transformations on sequences of coherent Gaussian samples that result in the desired marginal PDF of the clutter envelope. However, due to the non-linear transformations, the covariance matrix of the resulting non-Gaussian clutter is related to that of the original Gaussian samples in an intricate manner, which makes it difficult to obtain the desired covariance matrix. Moreover, these methods do not guarantee that the resulting covariance matrix is non-negative definite [23]. On the contrary, the SIRP model in (12) allows controlling both the envelope PDF and the covariance matrix of the generated clutter.

One of the most common clutter distributions is the K-distribution, which provides a good fit to the envelope of the data acquired from different environments. The K-distribution of the clutter envelope is given by [6]

$$f(r) = \frac{2\delta}{\Gamma(\alpha)}\left(\frac{\delta r}{2}\right)^{\alpha}K_{\alpha-1}(\delta r), \tag{14}$$

where $\alpha > 0$ and $\delta > 0$ are the shape and scale parameters, respectively, $\Gamma(\cdot)$ is the Gamma function, and $K_\alpha(\cdot)$ is the modified Bessel function of the second kind of order α. In order to arrive at the K-distribution for the clutter envelope using the SIRP model, the PDF of the texture component $f_V(v)$ should be selected as [6]

$$f_V(v) = \frac{2\delta}{\Gamma(\alpha)2^{\alpha}}(\delta v)^{2\alpha-1}\exp\left(-\delta^2v^2\right). \tag{15}$$

In this case, the second moment of v is given by $E(v^2) = 2\alpha/\delta^2$.

3 Non-homogeneity detector

To calculate the adaptive weight vector $\boldsymbol{w}$ in (7) for a given CUT within the available L range cells, one needs to estimate the covariance matrix of this CUT from the adjacent $L_1 = L-1$ secondary cells z_l, where $l \in \mathscr{L} = \{1, \ldots, L_1\}$, that together form the secondary sample matrix $\boldsymbol{Z} = [z_1, \ldots, z_{L_1}] \in \mathbb{C}^{J\times L_1}$. To censor non-homogeneous secondary cells from the estimation, the NHD decides if a secondary cell, say z_k for $k \in \mathscr{L}$, is non-homogeneous with respect to the remaining $L_2 = L_1 - 1$ secondary cells z_l for $l \in \mathscr{L} - \{k\}$, which together forms a matrix $\boldsymbol{Z}_k$ (obtained from $\boldsymbol{Z}$ by removing the column z_k). The NHD is basically a STAP detector that sequentially processes the L_1 secondary cells with one of them, z_k, temporarily considered as the CUT (also termed secondary CUT), while the remaining secondary cells $\boldsymbol{Z}_k$ are used to estimate the covariance matrix of z_k.

A basic test employs the general inner product (GIP), which is equivalent to the square of the Mahalanobis distance [5], i.e.,

$$\Lambda_{\mathrm{GIP}} = (z_k - \hat{\boldsymbol{\mu}})^H\hat{\boldsymbol{R}}^{-1}(z_k - \hat{\boldsymbol{\mu}}) \underset{H_1}{\overset{H_0}{\gtrless}} \eta_1 \tag{16}$$

where $\hat{\boldsymbol{\mu}} \in \mathbb{C}^{J\times 1}$ is the sample mean of $\boldsymbol{z}_k$, $\hat{\boldsymbol{R}} \in \mathbb{C}^{J\times J}$ is its estimated covariance matrix, and η_1 is a threshold that is determined based on the required probability of false alarm P_{F}. In this test, H_0 is the null hypothesis that $\boldsymbol{z}_k$ is homogeneous with respect to $\boldsymbol{Z}_k$, while H_1 is the alternative hypothesis. However, the GIP test is not robust in non-Gaussian clutter environment as reported in [6], where a more robust detector, namely the normalised adaptive matched filter (NAMF), is proposed as

$$\begin{aligned}\Lambda_{\mathrm{NAMF}} &= \frac{\left|\hat{\boldsymbol{w}}^H \boldsymbol{z}_k\right|^2}{(\hat{\boldsymbol{w}}^H \hat{\boldsymbol{R}} \hat{\boldsymbol{w}})(\boldsymbol{z}_k^H \hat{\boldsymbol{R}}^{-1} \boldsymbol{z}_k)} \\ &= \frac{\left|\boldsymbol{s}^H \hat{\boldsymbol{R}}^{-1} \boldsymbol{z}_k\right|^2}{(\boldsymbol{s}^H \hat{\boldsymbol{R}}^{-1} \boldsymbol{s})(\boldsymbol{z}_k^H \hat{\boldsymbol{R}}^{-1} \boldsymbol{z}_k)} \underset{H_1}{\overset{H_0}{\gtrless}} \eta_2\end{aligned} \tag{17}$$

where $\hat{\boldsymbol{w}} = g\hat{\boldsymbol{R}}^{-1}\boldsymbol{s}$. For this detector, P_F has been derived in [6] assuming Gaussian clutter, but it is not tractable analytically for non-Gaussian SIRP clutter. In the latter case, Monte Carlo simulations are used to set the threshold.

While the SCM in (10) is the ML estimator in the case of Gaussian clutter, the ML estimator of the covariance matrix in the case of compound Gaussian clutter cannot generally be obtained in analytical form. Tyler introduced a generalisation of the ML estimator for elliptical distributions (which include Gaussian along with other distributions) that can be expressed as the solution to the non-linear equation [24]

$$\hat{\boldsymbol{R}} = \frac{J}{L_1} \sum_{\substack{l=1 \\ l\neq k}}^{L_1} \frac{\boldsymbol{z}_l \boldsymbol{z}_l^H}{\boldsymbol{z}_l^H \hat{\boldsymbol{R}}^{-1} \boldsymbol{z}_l} \tag{18}$$

However, besides the difficulties posed by solving (18) due to the high computational cost, it needs a large number of secondary cells L_1 for estimator accuracy [6]. An approximation to the ML estimator for the covariance matrix of SIRP clutter is given by [6]

$$\hat{\boldsymbol{R}}_{\mathrm{SIRP}} = \frac{1}{L_1} \sum_{\substack{l=1 \\ l\neq k}}^{L_1} \zeta_l \boldsymbol{z}_l \boldsymbol{z}_l^H, \tag{19}$$

where

$$\zeta_l = \frac{h_{2J+2}(\boldsymbol{z}_l^H \hat{\boldsymbol{R}}_{\mathrm{SIRP}}^{-1} \boldsymbol{z}_l)}{h_{2J}(\boldsymbol{z}_l^H \hat{\boldsymbol{R}}_{\mathrm{SIRP}}^{-1} \boldsymbol{z}_l)} \tag{20}$$

where the function $h_{2J}(\cdot)$ is defined in (13). The scalar ζ_l cannot be expressed in a closed form, since both sides of (19) contain $\hat{\boldsymbol{R}}_{\mathrm{SIRP}}$, but it can be found by the iterative expectation-maximisation (EM) algorithm [25]. However, the EM algorithm converges slowly, especially for low values of α [6]. Moreover, the estimator in (19) needs a priori knowledge of the clutter distribution. Another approximation to the ML covariance estimator in case of non-Gaussian clutter is the iterative normalised SCM (NSCM) that is obtained through the following recursive formula [21]

$$\hat{\boldsymbol{R}}_{\mathrm{NSCM}}^{(t+1)} = \frac{J}{L_1} \sum_{\substack{l=1 \\ l\neq k}}^{L_1} \frac{\mathfrak{R}(\boldsymbol{z}_l)\mathfrak{R}(\boldsymbol{z}_l^T)}{\mathfrak{R}(\boldsymbol{z}_l^T)(\hat{\boldsymbol{R}}_{\mathrm{NSCM}}^{(t)})^{-1}\mathfrak{R}(\boldsymbol{z}_l)} \tag{21}$$

where t denotes the iteration index. The computation is initialised with the estimator [21]

$$\hat{\boldsymbol{R}}_{\mathrm{NSCM}}^{(0)} = \frac{J}{L_1} \sum_{\substack{l=1 \\ l\neq k}}^{L_1} \frac{\mathfrak{R}(\boldsymbol{z}_l)\mathfrak{R}(\boldsymbol{z}_l^T)}{\mathfrak{R}(\boldsymbol{z}_l^T)\mathfrak{R}(\boldsymbol{z}_l)} \tag{22}$$

Although it is also based on iterative procedures, its rate of convergence is faster than the EM-based algorithm mentioned above for the solution of (19) and (20), and it has been reported to converge after only four iterations [21]. Moreover, the NSCM shows a detection performance that is very close to that of the EM-based estimator [26]. Henceforth, whenever we use $\hat{\boldsymbol{R}}$ we mean $\hat{\boldsymbol{R}}_{\mathrm{NSCM}}$.

4 Proposed NHD

In this section, we first introduce the PD function and use it to provide covariance-free interpretations of the GIP and NAMF test statistics. We then introduce a covariance-free NHD that employs a novel non-parametric (distribution-free) test statistic based on the PD function and extend it to the case of correlated clutter.

4.1 Projection depth function

Let $\boldsymbol{z} \in \mathbb{C}^J$ be a random vector with joint cumulative distribution function (CDF) $F(\boldsymbol{z})$. A depth function is a random scalar $D(\boldsymbol{z}, F) \in [0, 1]$, defined as a function of $\boldsymbol{z}$ and taking into account the features of its distribution F. Ideally, the value of $D(\boldsymbol{z}, F)$ provides an inverse measure of 'distance' from a central point (such as the median or the mean of the distribution F), which can used for the centre-outward ordering of observations of vector $\boldsymbol{z}$ [27]. Based on this ordering, outliers can be detected when their distance from the center is larger than a certain threshold. Hence, the concepts of depth function and outliers are related. Specifically, we can define a measure of outlyingness as the function [28]

$$O(\boldsymbol{z}, F) = \frac{1}{D(\boldsymbol{z}, F)} - 1 \tag{23}$$

Let $\mu(\cdot)$ and $\sigma(\cdot)$ be univariate location and scale measures, respectively. Then, the projection-based outlyingness of $\boldsymbol{z}$ is [28]

$$O(\boldsymbol{z}, F) = \sup_{\boldsymbol{u} \in \mathbb{C}^J,\ \|\boldsymbol{u}\| = 1} \frac{\left|\boldsymbol{u}^H \boldsymbol{z} - \mu(F_{\boldsymbol{u}})\right|}{\sigma(F_{\boldsymbol{u}})} \tag{24}$$

where $F_{\boldsymbol{u}}$ is the CDF of $\boldsymbol{u}^H\boldsymbol{z}$. In practice, the sample version of (24) is found by replacing $F_{\boldsymbol{u}}$ by its empirical version $\hat{F}_{\boldsymbol{u}}$.

The projection-based outlyingness has a higher breakdown value in comparison to other types of outlyingness functions [28], which motivates its use in this work. To understand the concept of the breakdown value, consider the estimation of a scalar parameter θ from n observations $X_n = \{x_1, \ldots, x_n\}$, with $T_\theta(X_n)$ denoting the resulting estimator. Let us assume that out of these observations, m are replaced by arbitrary values (outliers), resulting in the contaminated sample set $X_{n,m}$. The estimator $T_\theta(X_{n,m})$ is calculated for the same parameter θ, but from the contaminated set $X_{n,m}$. The finite sample breakdown value of the estimator $T_\theta(\cdot)$ is the smallest ratio of contamination m/n for which the distance between $T_\theta(X_n)$ and $T_\theta(X_{n,m})$ can become arbitrarily large for certain choices of outliers [29].

The projection-based outlyingness function in (24) is a robust alternative to the Mahalanobis distance and, hence, to the GIP. A non-parametric GIP NHD detector was introduced in [16] for Gaussian clutter using the outlyingness function in (24) that evades the high computational burden of estimating the covariance matrix and its inverse with increasing dimensions of the range cells. We refer to it as the projection depth GIP (PD-GIP). However, the generation of the projection vectors $\boldsymbol{u}$ to approximate the supremum operation requires calculating the median of the secondary cells in $\boldsymbol{Z}_k$ for each $\boldsymbol{z}_k$; besides, the performance of the original GIP detector in case of non-Gaussian clutter environment is not robust [6]. Below, we propose a covariance-free detector based on (24) that is approximately equivalent to the NAMF detector in its robust performance, while at the same time sharing the non-parametric character of (24) and its lower computational complexity.

4.2 Covariance-free reformulation of GIP and NAMF

We begin by stating a proposition about the equivalence of the outlyingness function in (24) to the GIP in (16). This equivalence, which was demonstrated in [30] for the case of real-valued data in image processing applications, is extended here to complex-valued radar observations, as needed to comply with the case of coherent clutter model under consideration in this paper.

Proposition 1: Let $\boldsymbol{Z} = [\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{L_1}] \in \mathbb{C}^{J \times L_1}$ be a secondary sample matrix. For any target steering vector $\boldsymbol{s}$ as in (3) and an arbitrary secondary cell $\boldsymbol{z}_k$, where $k \in \{1, \ldots, L_1\}$, is associated with an estimated covariance matrix $\hat{\boldsymbol{R}}$ and mean vector $\hat{\boldsymbol{\mu}}$, we have

$$\sup_{\|\boldsymbol{u}\|=1} \left(\frac{\left| \boldsymbol{u}^H \boldsymbol{z}_k - \hat{\mu}(\boldsymbol{u}^H \boldsymbol{Z}_k) \right|}{\hat{\sigma}(\boldsymbol{u}^H \boldsymbol{Z}_k)} \right)^2 = (\boldsymbol{z}_k - \hat{\boldsymbol{\mu}})^H \hat{\boldsymbol{R}}^{-1} (\boldsymbol{z}_k - \hat{\boldsymbol{\mu}}) \tag{25}$$

and

$$\sup_{\|\boldsymbol{u}\|=1} \left(\frac{\left| \boldsymbol{u}^H \boldsymbol{s} \right|}{\hat{\sigma}(\boldsymbol{u}^H \boldsymbol{Z}_k)} \right)^2 = \boldsymbol{s}^H \hat{\boldsymbol{R}}^{-1} \boldsymbol{s} \tag{26}$$

where $\hat{\mu}(\boldsymbol{u}^H \boldsymbol{Z}_k)$, $\hat{\sigma}(\boldsymbol{u}^H \boldsymbol{Z}_k)$ are the sample mean and standard deviation (SD) of $\boldsymbol{u}^H \boldsymbol{Z}_k$, respectively, and $\boldsymbol{Z}_k$ denotes the secondary cells after excluding $\boldsymbol{z}_k$.

In (25) and (26), the supremum operation is taken over all unit-norm vectors $\boldsymbol{u} \in \mathbb{C}^J$. The proof of Proposition 1 is given in Appendix 1. We note that the denominator of the test statistic of the NAMF detector in (17) can be expressed using the outlyingness function in (24) as shown in Appendix 1, specifically (48) and (50). However, the numerator of (17), $\boldsymbol{s}^H \hat{\boldsymbol{R}}^{-1} \boldsymbol{z}_k$, cannot be directly expressed in terms of (24). To circumvent this difficulty, we suggest replacing $\boldsymbol{z}_k$ in the numerator of (17) by $(\boldsymbol{s}^H \boldsymbol{z}_k)\boldsymbol{s}$ to obtain

$$\left| \boldsymbol{s}^H \hat{\boldsymbol{R}}^{-1} (\boldsymbol{s}^H \boldsymbol{z}_k) \boldsymbol{s} \right|^2 = \left| \boldsymbol{s}^H \boldsymbol{z}_k \right|^2 \left(\boldsymbol{s}^H \hat{\boldsymbol{R}}^{-1} \boldsymbol{s} \right)^2 \tag{27}$$

The following proposition states that, in case of a dominant target, the expression in (27) is approximately equivalent to $\left| \boldsymbol{s}^H \hat{\boldsymbol{R}}^{-1} \boldsymbol{z}_k \right|^2$, which is the numerator of (17).

Proposition 2: Let $\boldsymbol{s}$, $\boldsymbol{z}_k$, and $\hat{\boldsymbol{R}}$ be as defined in Proposition 1, then $(\boldsymbol{s}^H \boldsymbol{z}_k)\boldsymbol{s}^H \hat{\boldsymbol{R}}^{-1} \boldsymbol{s}$ has the same target's signal component as $\boldsymbol{s}^H \hat{\boldsymbol{R}}^{-1} \boldsymbol{z}_k$.

The proof of this proposition is provided in Appendix 2. The next proposition introduces a modified test statistic, which approximates the original NAMF test statistic in (17) in terms of the projection-based outlyingness in (24).

Proposition 3: Let $\boldsymbol{s}$, $\boldsymbol{z}_k$, $\boldsymbol{Z}_k$, and $\hat{\boldsymbol{R}}$ be as defined in Proposition 1. Then

$$\begin{aligned} \Lambda'_{\mathrm{NAMF}} &\triangleq \frac{\left| \boldsymbol{s}^H \hat{\boldsymbol{R}}^{-1} (\boldsymbol{s}^H \boldsymbol{z}_k) \boldsymbol{s} \right|^2}{(\boldsymbol{s}^H \hat{\boldsymbol{R}}^{-1} \boldsymbol{s})(\boldsymbol{z}_k^H \hat{\boldsymbol{R}}^{-1} \boldsymbol{z}_k)} \\ &= \frac{\left| \boldsymbol{s}^H \boldsymbol{z}_k \right|^2 \sup_{\|\boldsymbol{u}\|=1} \left(\frac{\left| \boldsymbol{u}^H \boldsymbol{s} \right|}{\hat{\sigma}(\boldsymbol{u}^H \boldsymbol{Z}_k)} \right)^2}{\sup_{\|\boldsymbol{u}\|=1} \left(\frac{\left| \boldsymbol{u}^H \boldsymbol{z}_k \right|}{\hat{\sigma}(\boldsymbol{u}^H \boldsymbol{Z}_k)} \right)^2} \end{aligned} \tag{28}$$

The proof of this proposition is provided in Appendix 3. As observed from (28), the test statistic Λ'_{NAMF} is covariance-free. Moreover, besides its approximate equivalence to the NAMF test in (17) as shown in Proposition 2, it inherits the non-parametric characteristic of the projection-based outlyingness.

4.3 Robust, covariance-free, and non-parametric NHD

Although the projection-based outlyingness in (24) does not dictate a specific scale measure, the median absolute deviation (MAD) has been widely used in robust statistics to detect outliers due to its robustness with respect to heavy-tailed distributions and higher breakdown value compared to the SD [31].

For the real-valued random sample data $X_n = [x_1, \ldots, x_n]$ with order statistics $x_{(1)} \leq \cdots \leq x_{(n)}$, the sample median $\mathrm{med}(X_n)$ and sample MAD $\mathrm{mad}(X_n)$ are calculated as [32]

$$\mathrm{med}(X_n) = \begin{cases} x_{((n+1)/2)} & n \text{ is odd} \\ 0.5(x_{(n/2)} + x_{((n/2)+1)}) & n \text{ is even} \end{cases} \tag{29}$$

and

$$\mathrm{mad}(X_n) = \mathrm{med}(|x_i - \mathrm{med}(X_n)|), \quad i = 1, \ldots, n \tag{30}$$

respectively. The population MAD of the random variable X $\mathrm{MAD}(X)$ is related to its population SD $\sigma(X)$ as [33]

$$\mathrm{MAD}(X) = k_f \sigma(X) \tag{31}$$

where k_f is a positive constant to achieve consistency and its value depends on the population CDF of X. For the standard normal distribution, $k_f \simeq 0.6745$ [31]. In the absence of outliers, the sample versions $\mathrm{mad}(X_n)$ and $\hat{\sigma}(X_n)$ are related, approximately, by the same constant k_f even with a sample size as low as 10 [31].

The breakdown value of the MAD is 0.5 [33], which is the best possible breakdown value, compared to a value of 0 for the SD [29]. Therefore, the MAD is more robust than the SD, especially for heavy-tailed clutter distributions as the K-distribution. Since heavy-tailed distributions tend to have many outliers with very high values, the MAD constitutes a better estimate for the scale parameter than the SD and leads to a lower threshold for the same false alarm rate and, consequently, a better detection.

By employing the $\mathrm{mad}(\boldsymbol{u}^H \boldsymbol{Z}_k)$ as a robust scale measure instead of $\hat{\sigma}(\boldsymbol{u}^H \boldsymbol{Z}_k)$ in (28), we obtain

$$\Lambda'_{\mathrm{NAMF}} \simeq \frac{\left| \boldsymbol{s}^H \boldsymbol{z}_k \right|^2 \sup_{\|\boldsymbol{u}\|=1} \left(\frac{k_f \left| \boldsymbol{u}^H \boldsymbol{s} \right|}{\mathrm{mad}(\boldsymbol{u}^H \boldsymbol{Z}_k)} \right)^2}{\sup_{\|\boldsymbol{u}\|=1} \left(\frac{k_f \left| \boldsymbol{u}^H \boldsymbol{z}_k \right|}{\mathrm{mad}(\boldsymbol{u}^H \boldsymbol{Z}_k)} \right)^2} \tag{32}$$

where for the proper complex signal vectors $\boldsymbol{Z}_k$, we have according to (8)

$$\mathrm{mad}(\boldsymbol{u}^H \boldsymbol{Z}_k) = 2\mathrm{mad}(\Re(\boldsymbol{u}^H \boldsymbol{Z}_k)) = 2\mathrm{mad}(\Im(\boldsymbol{u}^H \boldsymbol{Z}_k)) \tag{33}$$

Under the SIRP model considered in this paper, all the CDFs $F_{\boldsymbol{u}}$ of the projections of a given SIRP vector are the same [22]; this means that the value of k_f, that is determined based on $F_{\boldsymbol{u}}$, does not depend on the projection vector $\boldsymbol{u}$. To verify this independence for the considered signal vector, a secondary sample matrix $\boldsymbol{Z} \in \mathbb{C}^{J \times L}$ of uncorrelated clutter vectors is simulated with the dimension of the secondary cells fixed at $J = 20$, while L/J changes from 2 to 10. For each sample size L, a sample of k_f from different 1000 projection vectors $\boldsymbol{u}$ is calculated using (31), but using the sample MAD $\mathrm{mad}(\boldsymbol{u}^H \boldsymbol{Z}_k)$ as defined in (30). The relative SD (RSD) of the k_f sample is

$$\mathrm{RSD}(k_f) = \frac{\hat{\sigma}(k_f)}{\hat{\mu}(k_f)} \tag{34}$$

where $\hat{\sigma}(k_f)$ and $\hat{\mu}(k_f)$ are the sample SD and mean of k_f, respectively, averaged over 10^4 trials. As Fig. 2 shows, the value of k_f exhibits a low variation for both of the considered distributions

at all considered values of L. This also shows that (31) holds also for mad $\left(\boldsymbol{u}^H\boldsymbol{Z}_k\right)$ even for low sample size, as shown in [31].

With the agreement of the presented simulation results with the theoretical analysis in [22, 31], the constant k_f can be taken out of the supremum in (32). Therefore, the proposed test (Λ_{PD}) based on the NAMF and PD outlyingness is

$$\Lambda_{\mathrm{PD-NAMF}} \triangleq \frac{\left|\boldsymbol{s}^H\boldsymbol{z}_k\right|^2 \sup_{\|\boldsymbol{u}\|=1}\left(\frac{\left|\boldsymbol{u}^H\boldsymbol{s}\right|}{\mathrm{mad}(\boldsymbol{u}^H\boldsymbol{Z}_k)}\right)^2}{\sup_{\|\boldsymbol{u}\|=1}\left(\frac{\left|\boldsymbol{u}^H\boldsymbol{z}_k\right|}{\mathrm{mad}(\boldsymbol{u}^H\boldsymbol{Z}_k)}\right)^2} \underset{H_1}{\overset{H_0}{\gtrless}} \eta_3 \tag{35}$$

We call the detector based on the test statistic in (35) as the PD-NAMF. Theoretically, implementing the supremum requires calculating the projections of an infinite number of vectors that cover the unit hypersphere in J-dimensional space. In practice, as shown in [30, 34] for different applications, the supremum can be approximated by taking the maximum magnitude of a finite number (Q) of projections of $\boldsymbol{z}_k$ or $\boldsymbol{s}$ on randomly generated vectors over this hypersphere. As suggested in [35], each of these vectors is obtained by first generating J independent complex Gaussian variates $u_i \sim \mathcal{CN}(0,1)$, $1 \le i \le J$ with zero mean and unit variance to form the vector $\boldsymbol{u} = [u_1, \ldots, u_J]$, and then normalising $\boldsymbol{u}$ with respect to $\|\boldsymbol{u}\|_2$. As the steps above show, the generation method used in this work is totally independent of the steering vector $\boldsymbol{s}$ or the CUT $\boldsymbol{z}_k$, hence it is performed once and the obtained vectors are stored to be used for all range cells and any steering vector $\boldsymbol{s}$. This off-line method of generation is different from that used in [16]. In the latter, the projection vectors were recomputed for each CUT from the secondary cells $\boldsymbol{Z}$. The discussion on the choice of Q is left for Section 5.

4.4 Correlated clutter

The correlation matrix of the clutter signal vector $\boldsymbol{c}$ in (6) is a Kronecker product of the temporal (i.e. between pulses) and spatial (i.e. between antenna elements) covariance matrices $\boldsymbol{\Psi}_{\mathrm{t}}$ and $\boldsymbol{\Psi}_{\mathrm{s}}$, respectively, i.e. [36]

$$\boldsymbol{\Psi} = \boldsymbol{\Psi}_{\mathrm{t}} \otimes \boldsymbol{\Psi}_{\mathrm{s}} \tag{36}$$

The spatial correlation of the clutter depends on the inter-element spacing of the antenna array as in [37, 38]. It can be approximated as

$$\boldsymbol{\Psi}_{\mathrm{s}} = \left[\rho_{\mathrm{s}}^{|i-j|}\right], \quad 1 \le i, j \le N \tag{37}$$

where ρ_{s} is the one-lag spatial correlation coefficient. Based on experimental measurements for different clutter environments, e.g. [39, 40], the temporal covariance matrix of the clutter can be expressed similarly as

$$\boldsymbol{\Psi}_{\mathrm{t}} = \left[\rho_{t}^{|i-j|}\right], \quad 1 \le i, j \le M \tag{38}$$

where ρ_{t} is the one-lag temporal correlation coefficient.

To achieve the constant false alarm rate (CFAR) property, a crucial feature of a robust radar detector is the scale invariance of its test statistic. However, the distribution of the sample version of (24) depends on the scale parameter estimator if the latter is not Fisher consistent, as shown in [28, Remark 3.4]. Since (24) is the main building block of the PD-NAMF in (35), the dependency of the detector's test statistic on the scale measure estimator degrades its detection performance. This is more emphasised considering the rapid variations of the clutter signal statistics in many real radar scenarios. Therefore, we need to delve into the Fisher-consistency of the mad($\boldsymbol{u}^H\boldsymbol{Z}_k$) as a scale measure in the case of correlated clutter considered in this paper.

To define the Fisher consistency, let T_θ be an estimator for the parameter θ such that $T_\theta = f(\hat{G})$, where $\hat{G}$ is the empirical CDF based on independent sample data and f is a continuous function in the space of distribution functions. The estimator T_θ is said to be Fisher-consistent if $f(G) = \theta$, where G is the population CDF from which the sample data is drawn [17]. From (31), we can outline two necessary conditions for the Fisher consistency of the mad($\boldsymbol{u}^H\boldsymbol{Z}_k$): (i) the convergence of the sample SD $\hat{\sigma}$ to its population version, and (ii) the convergence of empirical CDF $\hat{F}_{\boldsymbol{u}}$ to the population CDF $F_{\boldsymbol{u}}$, since k_f depends on $F_{\boldsymbol{u}}$.

First, we begin the discussion with the Fisher consistency of $\hat{\sigma}$ in the case of correlated data. Based on the results of [41], the convergence rate of $\hat{\sigma}$ to its population version σ in the case of correlated data is lower compared to that of uncorrelated data. This slowdown depends on the correlation coefficient ρ. More specifically, for $\rho = 0.75$, the rate of convergence is 3.5 times lower than that for the uncorrelated data [41]. Second, we consider the convergence of the empirical CDF to the population CDF for correlated data, which is addressed in [42, Theorem 1]. Let $\{x_i\}_{i=1}^n$, be random univariate samples that follow a joint normal distribution with the correlation matrix $\boldsymbol{\Phi}$. If $\{x_i\}_{i=1}^n$ are not weakly correlated, then $E[\hat{G} - G]^2$ does not tend to 0 as $n \to \infty$.

Based on the aforementioned discussions, $\hat{\sigma}$ and $\hat{F}_{\boldsymbol{u}}$ calculated from correlated data samples do not converge to their corresponding population versions, even if the data samples follow the Gaussian distribution for which $\hat{\sigma}$ is the ML estimator of the SD. Therefore, given the strong correlation shown by the available experimental data for different clutter environments [39, 40], the mad$\left(\boldsymbol{u}^H\boldsymbol{Z}_k\right)$ is not a Fisher-consistent scale estimator in the considered application. To handle this problem, we propose decorrelating $\boldsymbol{Z}$ before applying (35). The decorrelated secondary cells are

$$\boldsymbol{Z}_{\mathrm{d}} = \hat{\boldsymbol{\Psi}}^{-1/2}\boldsymbol{Z} \tag{39}$$

where $\hat{\boldsymbol{\Psi}}$ is an estimate of the correlation matrix of $\boldsymbol{Z}$. To keep the non-parametric characteristic of the PD-NAMF, we use a non-parametric correlation estimator.

There are two prevalent nonparametric rank correlation coefficients, namely, the Kendall's and Spearman correlation coefficients. Compared to the Spearman coefficient, the Kendall's has a lower bias, shows better accuracy at lower number of samples, and has a lower mean square error (MSE) for heavily correlated data [43]. However, calculating the Spearman coefficient has lower computational complexity than that of the Kendall coefficient. Generally, the Kendall correlation matrix estimator of the E-dimensional vector $\boldsymbol{x}$ calculated from the sample data $\boldsymbol{X} \in \mathbb{R}^{E\times D}$ is given by [44]

$$\hat{\psi}_{jk}^K = \frac{2}{D(D-1)}\sum_{i<i'}\mathrm{sign}(x_{ji} - x_{ji'})\mathrm{sign}(x_{ki} - x_{ki'}) \tag{40}$$

where $\hat{\psi}_{jk}^K$ is the (j,k)th entry of $\hat{\boldsymbol{\Psi}}^K$, $1 \le j,k \le E$, $1 \le i \le D$, and $2 \le i' \le D$. The Spearman correlation matrix estimator calculated from the same sample data is [44]

$$\hat{\psi}_{jk}^S = \frac{\sum_{i=1}^D\left(o(x_{ji}) - \bar{D}\right)\left(o(x_{ki}) - \bar{D}\right)}{\sqrt{\sum_{i=1}^D\left(o(x_{ji}) - \bar{D}\right)^2\sum_{i=1}^{L_1}\left(o(x_{ki}) - \bar{D}\right)^2}} \tag{41}$$

where $\bar{D} = (D+1/2)$ and $o(x_{ki})$ denotes the order of x_{ki} within $x_{k1}, \ldots, x_{kD}$.

The correlation estimators in (40) and (41) cannot be directly applied to the complex-valued secondary cells assumed in this paper. Based on (8), the correlation matrix $\hat{\boldsymbol{\Psi}} \in \mathbb{C}^{MN\times MN}$ is given by

$$\hat{\boldsymbol{\Psi}} = 2\hat{\boldsymbol{\Psi}}_{z_{\mathrm{R}}z_{\mathrm{R}}} = 2\hat{\boldsymbol{\Psi}}_{z_{\mathrm{I}}z_{\mathrm{I}}} \tag{42}$$

where $\hat{\boldsymbol{\Psi}}_{z_R z_R}$ and $\hat{\boldsymbol{\Psi}}_{z_I z_I}$ are the estimated autocorrelation of $\Re(z)$ and $\Im(z)$, respectively. Equation (42) is applied to both Kendall and Spearman correlation matrix estimators $\hat{\boldsymbol{\Psi}}^K$ and $\hat{\boldsymbol{\Psi}}^S$, respectively.

Remark 1: It should be emphasised that the estimation of the correlation matrix $\hat{\boldsymbol{R}}^{-1} = [p_{ij}]$ is not the same as estimating the covariance matrix $\boldsymbol{R}$. The estimation of $\boldsymbol{\Psi}$ can be seen as a step to estimate $\boldsymbol{R}$; a step that should be followed by estimating the SDs of the components of z_k. To illustrate this, the correlation matrix $\boldsymbol{\Psi} = [\psi_{ij}]$ is related to the covariance matrix $\boldsymbol{R} = [r_{ij}]$ as

$$r_{ij} = \sigma_i \sigma_j \psi_{ij}, \quad 1 \le i, j \le MN \tag{43}$$

where σ_i, σ_j are the SDs of the ith and jth components of z_k, respectively. The step of estimating the SDs of z_k's components is cumbersome for non-Gaussian clutter models in addition to the need to calibrate the resulting covariance matrix by solving multiple optimisation problems, as shown in [45]. Under the proposed algorithm, it suffices to estimate $\boldsymbol{\Psi}$; avoiding the complexity of estimating $\boldsymbol{R}$.

Remark 2: In the case of Gaussian distributed data, both the Spearman and Kendall coefficients are related to the linear Pearson correlation coefficient $\hat{\psi}^P_{jk}$ by [44]

$$\hat{\psi}^P_{jk} = \sin\left(\frac{\pi}{2}\hat{\psi}^K_{jk}\right) = 2\sin\left(\frac{\pi}{6}\hat{\psi}^S_{jk}\right). \tag{44}$$

Nonetheless, we do not use the transformed coefficients for two reasons. The first is that they are derived for the Gaussian distributed data, while we do not make any assumptions about the distribution of the received signal vector. The second is that the transformations in (44) do not guarantee the positive semi-definiteness of the estimated matrices [46], in contrast to the estimators in (40) and (41).

The flow of the PD-NAMF with both of Kendall and Spearman decorrelation matrices is shown in Figs. 3 and 4, respectively.

5 Performance assessment

In this section, the performance of the PD-NAMF in (35) is compared to that of the NAMF detector in (17) using Monte Carlo simulations. To justify the robustness of the PD-NAMF, we evaluate its performance with different clutter distributions and signal configurations. Moreover, we study the different choices for the algorithm parameters; specifically, the type of correlation estimator used and the minimum required a number of projections. Finally, we investigate the complexity and the execution time of the PD-NAMF compared to the NAMF detector for different design parameters.

5.1 Simulation parameters

The simulated radar signal has a fixed dimension $J = 16$ and L_1 is either 65 or 33 cells. The NHD is applied on a sequential basis on each secondary cell where $\hat{\boldsymbol{R}}$ and $\text{mad}(\boldsymbol{u}^H \mathbf{Z}_k)$ are estimated from the remaining $L_2 = 64$ or 32 cells. An interfering target is injected in a secondary CUT, representing a non-homogeneous cell, with a normalised Doppler frequency $f_d = 0.3$ and azimuth angle $\theta_t = 35°$.

The Kendall correlation matrix $\hat{\boldsymbol{\Psi}}^K$ is estimated once from all the L_1 secondary cells $\boldsymbol{Z}$ including the secondary CUT z_k, i.e. it is not recalculated for each z_k. However, the Spearman correlation matrix $\hat{\boldsymbol{\Psi}}^S$, due to its lower immunity to outliers, is calculated for each secondary CUT z_k from the remaining secondary cells. If the secondary CUT is included in the calculations of $\hat{\boldsymbol{\Psi}}^S$, a self-nulling effect appears at the output of the detector, especially at low number of secondary cells and/or high interfering target's power.

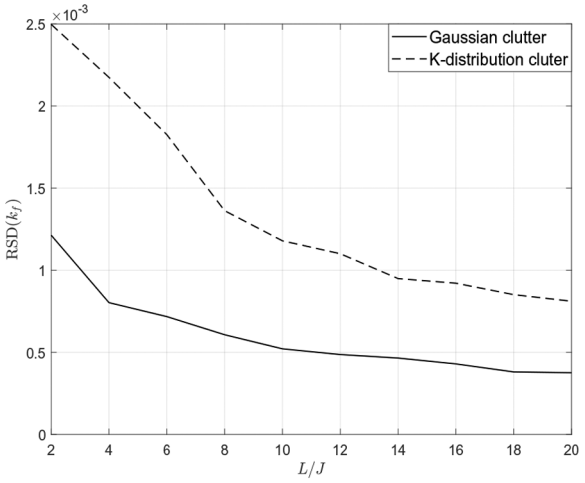

Fig. 2 *RSD of k_f at different number of samples (secondary cells) with $J = 20$*

Input: $\mathbf{Z}, \eta_3$

Calculate $\hat{\boldsymbol{\Psi}}^K$

$\mathbf{Z} \leftarrow (\hat{\boldsymbol{\Psi}}^K)^{-1/2}\mathbf{Z}$

Generate $\mathbf{U} = [\mathbf{u}_1, \cdots, \mathbf{u}_Q]$

for $k = 1$ to L **do**

Perform the proposed test as (35)

end for

Fig. 3 *Algorithm 1a: Using Kendall*

Input: $\mathbf{Z}, \eta_3$

Generate $\mathbf{U} = [\mathbf{u}_1, \cdots, \mathbf{u}_Q]$

for $k = 1$ to L **do**

$\mathbf{Z}_k \leftarrow [\mathbf{z}_l], l \in \mathcal{L} - \{k\}$

Calculate $\hat{\boldsymbol{\Psi}}^S$ from $\mathbf{Z}_k$

$\mathbf{Z}_k \leftarrow (\hat{\boldsymbol{\Psi}}^S)^{-1/2}\mathbf{Z}_k$

Perform the proposed test as (35)

end for

Fig. 4 *Algorithm 1b: Using Spearman*

Regarding the clutter vectors, they are generated as proper complex SIRP vectors with independent quadrature and in-phase components. For the clutter's envelope distribution, we consider two extreme cases: K-distributed clutter with $\alpha = 0.1$, which represents heavy-tailed spiky clutter, and Gaussian clutter. For the K-distributed clutter, δ is allowed to be randomly and independently changed from a range cell to another as indicated by Michels *et al.* [14].

As for the value of δ, Melebari *et al.* [47] give the measured values between (0, 1], which is the range of values considered in most of the relevant work in the literature, where only the shape parameter is considered to have an impact on the detection performance [6]. However, Antipov [48] provides experimental data showing that $1 \le \delta \le 2$. Therefore, we examine the performance of both the NAMF and the PD-NAMF in K-distributed clutter with the foregoing two cases of δ for each range cell: $\delta \sim \mathcal{U}(0, 1]$, as a default case, and $\delta \sim \mathcal{U}[1, 2]$, where $\mathcal{U}$ denotes the uniform distribution. The average clutter-to-noise ratio (CNR) is assumed to be 20 dB. The one-lag spatial and temporal correlation coefficients of the clutter in (37) and (38) are $\rho_s = \rho_t = 0.99$ [38, 40], unless other values are specified.

The projection vectors are generated randomly over the J-dimensional unit hypersphere as defined in Proposition 1. The

default number of projections is $Q = 4J$, however, we consider other values for Q later in this section. For all detection performance simulations, P_F is set to 0.01. The probability of detection P_D is evaluated versus the input signal-to-noise (SNR) of the interfering target using Monte Carlo simulation with 10^5 trials.

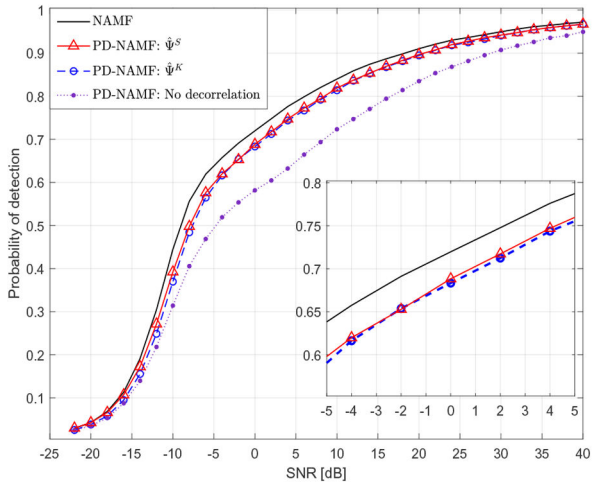

Fig. 5 *Detection performance in K-distributed clutter* ($\alpha = 0.1, \delta \sim \mathcal{U}(0,1], J = 16, L_2 = 64$)

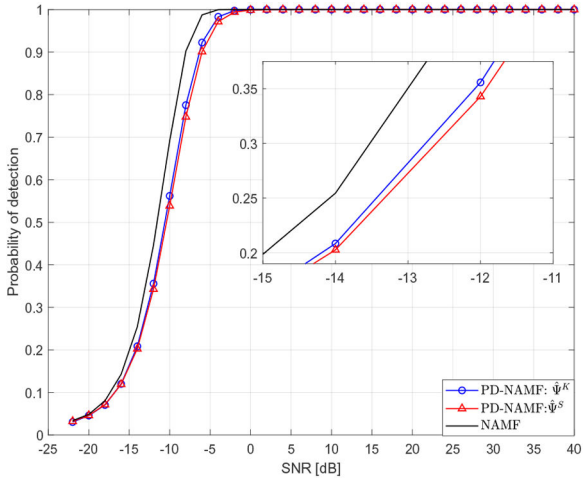

Fig. 6 *Detection performance in Gaussian clutter* ($J = 16$, $L_2 = 64$)

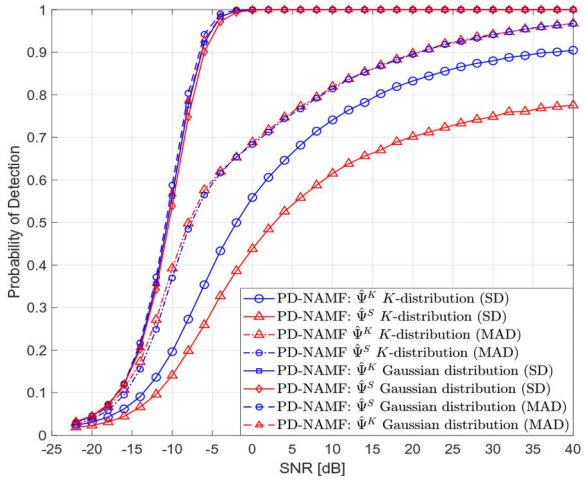

Fig. 7 *Detection performance in Gaussian and K-distributed clutter for PD-NAMF with MAD and SD* ($J = 16$, $L_2 = 64$)

5.2 Results

5.2.1 Low-dimensional case ($J/L_2 = 0.25$)*:* The STAP radar signal is considered low-dimensional when $J \leq 0.5L_2$. Fig. 5 shows the detection performance of the PD-NAMF and the NAMF detectors in K-distributed clutter with $\alpha = 0.1$ and $\delta \sim \mathcal{U}(0,1]$. We can observe that the detection performance of the PD-NAMF is comparable to that of the NAMF detector with a maximum loss in P_D of 0.051 in the case of $\hat{\boldsymbol{\Psi}}^S$. For SNR > -7 dB, it is also observed that using $\hat{\boldsymbol{\Psi}}^K$ provides a relative improvement over $\hat{\boldsymbol{\Psi}}^S$, narrowing the loss in P_D relative to the NAMF detector to 0.036. The difference in the performance between the NAMF detector and the PD-NAMF is getting narrower for SNR values beyond 5 dB and below −6 dB. While the observed slight advantage of the NAMF over the PD-NAMF is not common among all the cases studied in this paper, it also comes with the cost of much higher complexity, as we show shortly. In Fig. 5, we can also point out that the decorrelation is not only dictated by the theoretical need to achieve the Fisher consistency of $\text{mad}(\boldsymbol{u}^H\boldsymbol{Z}_k)$, but it also has a significant effect on the performance of the PD-NAMF.

The performance in Gaussian distributed clutter is shown in Fig. 6. Compared to the NAMF detector, the PD-NAMF has a maximum loss in P_D of 0.051 with both $\hat{\boldsymbol{\Psi}}^S$ and $\hat{\boldsymbol{\Psi}}^K$. Moreover, the performance of the PD-NAMF with $\hat{\boldsymbol{\Psi}}^S$ is similar to, or slightly better than that with $\hat{\boldsymbol{\Psi}}^K$.

To validate our claim of the robustness of the PD-NAMF using the MAD, we replaced MAD by SD in (32) and we evaluated the resulting detection performance in both Gaussian and K-distributed clutters. The results are shown in Fig. 7. As we can observe, the performance of the PD-NAMF using the SD is almost equivalent to that using the MAD in the case of the Gaussian distribution for both Spearman $\hat{\boldsymbol{\Psi}}^S$ and Kendall $\hat{\boldsymbol{\Psi}}^K$ decorrelation matrices. This is attributed to the equivalence of the SD and MAD, up to a constant k_f, in the case of the Gaussian distribution as we indicate in (31), which is based on [31]. In the case of the K-distribution, however, the performance of the PD-NAMF with both $\hat{\boldsymbol{\Psi}}^S$ and $\hat{\boldsymbol{\Psi}}^K$ degrades when SD is used in place of MAD. This is consistent with the theoretical reasoning provided in Section 4.3. We can also notice that when using SD, the PD-NAMF with $\hat{\boldsymbol{\Psi}}^K$ is more robust than the one with $\hat{\boldsymbol{\Psi}}^K$, which is ascribed to the higher robustness of the former in the presence of outliers as shown in [49].

It is important to demonstrate the performance of the original GIP and the PD-GIP detectors compared to the PD-NAMF in simulation. Interestingly, to the best of our knowledge, the detection performance of GIP in correlated compound Gaussian clutter has not been investigated in the open literature. Furthermore, the detection performance of the PD-GIP has not yet been investigated. For a fair comparison between the PD-GIP and the proposed PD-NAMF, we use a modified version of PD-GIP that differs from the one originally proposed in [16] in the following ways: the secondary cells are decorrelated using $\hat{\boldsymbol{\Psi}}^K$, and the projection vectors are generated in the same way as in the proposed PD-NAMF. To make this point clear, we refer to this detector as the 'modified PD-GIP'. Fig. 8 shows the performance of GIP, PD-GIP, modified PD-GIP, and PD-NAMF for both Gaussian and K-distributed clutters. We first note from the figure that the modifications made to the PD-GIP contribute to improve its performance. Furthermore, both the GIP and the modified PD-GIP show a performance degradation in the case of Gaussian clutter of ∼5 and 7 dB compared to the proposed PD-NAMF, respectively. However, this degradation is much greater in the case of K-distributed clutter. These results are consistent with the false alarm results for GIP presented in [6].

The effect of the scale parameter on the detection performance of both detectors can be observed in Fig. 9. The PD-NAMF performs approximately the same with both $\hat{\boldsymbol{\Psi}}^S$ and $\hat{\boldsymbol{\Psi}}^K$. The maximum detection loss by the PD-NAMF relative to NAMF is

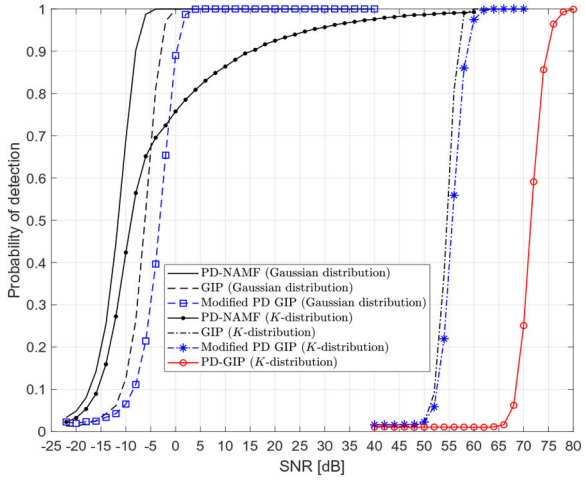

Fig. 8 *Detection performance in Gaussian and K-distributed clutter for GIP, modified PD-GIP, and PD-NAMF* ($J = 16$, $L_2 = 64$)

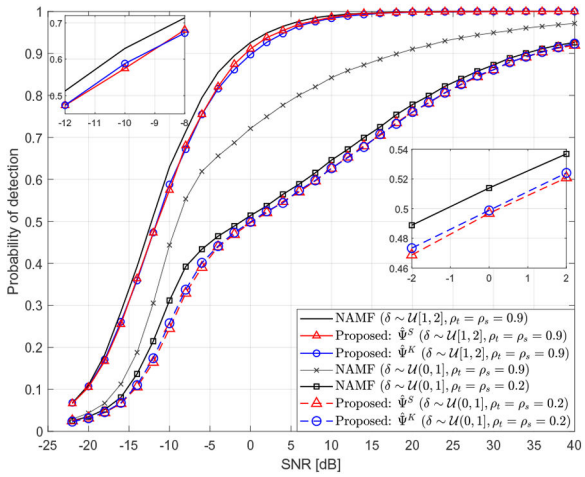

Fig. 9 *Detection performance in K-distributed clutter* ($\alpha = 0.1$, $\delta \sim \mathcal{U}[1,2]$ or $\mathcal{U}(0,1]$, $J = 16$, $L_2 = 64$, $\rho_s = \rho_t = 0.99$ or 0.2)

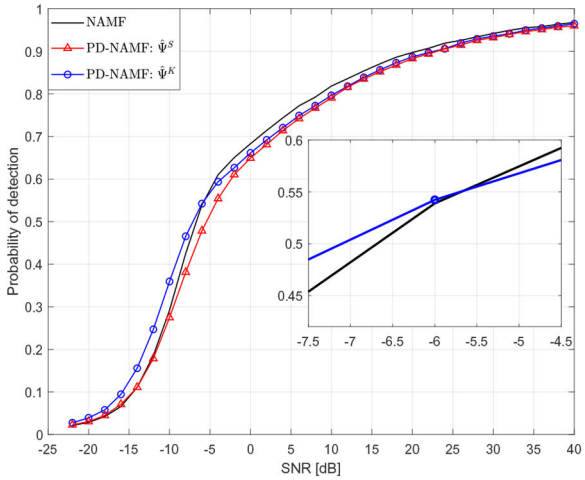

Fig. 10 *Detection performance in K-distributed clutter* ($\alpha = 0.1, \delta \sim \mathcal{U}(0,1], J = 16, L_2 = 32$)

0.034. In general, the detection performance of both the PD-NAMF and NAMF detectors are considerably affected by the change in the scale parameter of the clutter. In addition to the typical choice of 0.99 for both ρ_s and ρ_t, we consider the case of a low value $\rho_s = \rho_t = 0.2$ for these coefficients. The results are also shown in Fig. 9. The lower correlation coefficients of the clutter lead to a degraded detection performance, which coincides with the results in [50, Fig. 8].

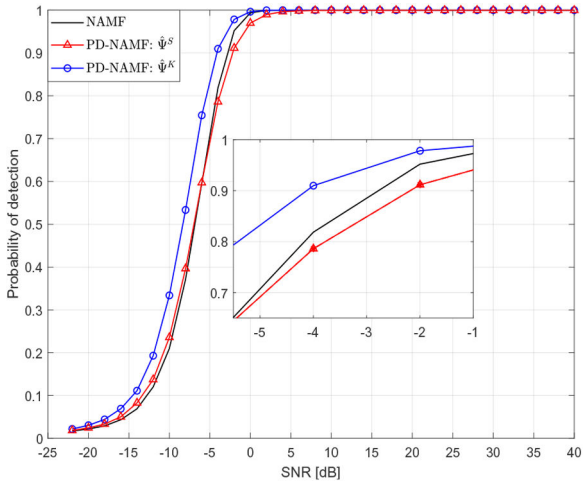

Fig. 11 *Detection performance in Gaussian distributed clutter* ($J = 16$, $L_2 = 32$)

5.2.2 Higher-dimensional case ($J/L_2 = 0.5$)*:* The performance of the PD-NAMF is investigated at a higher-dimensional case, where $L_2 = 32$ and $J = 16$. As we notice in Fig. 10, in the presence of K-distributed clutter, the PD-NAMF with $\hat{\boldsymbol{\Psi}}^K$ provides a relative advantage over the NAMF at SNR<–6 dB, with a maximum increase in P_D of 0.06. Beyond this point, the maximum loss in detection of the PD-NAMF with $\hat{\boldsymbol{\Psi}}^K$ relative to NAMF detector is 0.03. With $\hat{\boldsymbol{\Psi}}^S$ the PD-NAMF shows a maximum loss in P_D of 0.061 relative to the NAMF for $-10\,\text{dB} \leq \text{SNR} \leq 0\,\text{dB}$. It is noteworthy that the overall performance of both detectors is relatively degraded by lowering L_2/J as we observe by comparing the performance of each detector in Fig. 5 with its counterpart in Fig. 10.

As Fig. 11 depicts for the Gaussian clutter, P_D of the PD-NAMF with $\hat{\boldsymbol{\Psi}}^K$ is higher than that of the NAMF with a maximum difference of 0.16. It is noteworthy that this improvement in P_D provided by the PD-NAMF is higher than any loss it shows relative to the NAMF in the previous cases. When the PD-NAMF uses $\hat{\boldsymbol{\Psi}}^S$, it shows a maximum improvement of 0.014 over the NAMF. Beyond the crossover point at SNR = −6 dB, the PD-NAMF with $\hat{\boldsymbol{\Psi}}^S$ shows a comparable detection performance to the NAMF with a maximum loss in P_D of 0.041.

To summarise, using $\hat{\boldsymbol{\Psi}}^K$ with the PD-NAMF improves the detection performance over that of the NAMF detector in case of high-dimensional signals at all SNR values in Gaussian clutter and at lower SNR values for K-distributed clutter. This is explained by the robustness of the Kendall's coefficient in small sample conditions as mentioned before.

5.2.3 Number of projections: Theoretically, the higher the number of random projections Q, the more accurate (25) holds [30]. However, in practice, the used number of projections should be as small as possible for fast computations. Unfortunately, there is no analytical method to determine the minimum number of projections required for (25) to hold at a given approximation level; consequently, simulations are used to determine this value as in [30, 34]. While the simulations in these references are concerned with the convergence of (25), the simulation in this paper is concerned with maximising P_D at a given level of false alarm.

Fig. 12 illustrates the power of a secondary CUT with homogeneous interference (clutter and noise) or an interfering target's signal power at the output of the PD-NAMF with $\hat{\boldsymbol{\Psi}}^K$ and $\hat{\boldsymbol{\Psi}}^S$, as a function of the ratio Q/J, averaged over 10^5 trials. The powers of both interference and target signals of the PD-NAMF, with both correlation estimators, are normalised with respect to those of the NAMF detectors, respectively. The simulated clutter

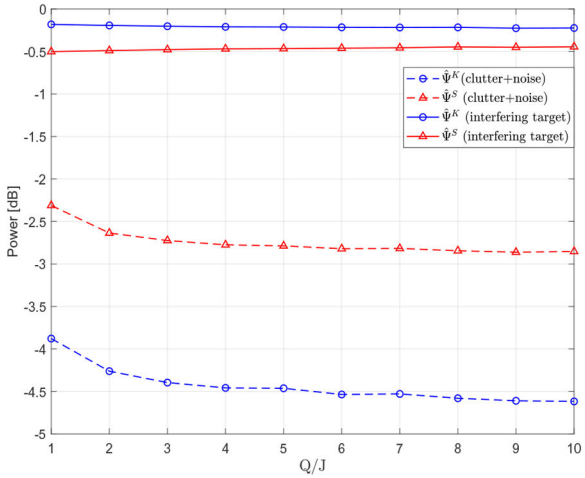

Fig. 12 *Interference and the interfering target powers at the output of the PD-NAMF in K-distributed clutter* ($\alpha = 0.1, \delta \sim \mathcal{U}(0,1], J = 16, L_2 = 64$)

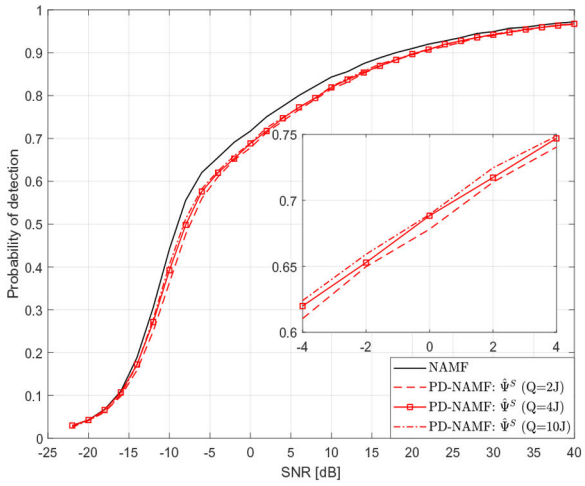

Fig. 13 *Detection performance in K-distributed clutter with different Q values* ($\hat{\mathbf{\Psi}}^S$, $\alpha = 0.1, \delta \sim \mathcal{U}(0,1], J = 16, L_2 = 64$)

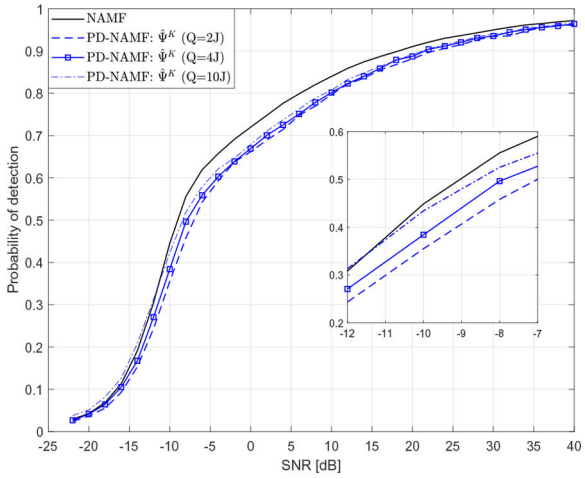

Fig. 14 *Detection performance in K-distributed clutter with different Q values* ($\hat{\mathbf{\Psi}}^K$, $\alpha = 0.1, \delta \sim \mathcal{U}(0,1], J = 16, L_2 = 64$)

envelope follows the K-distribution and the radar signal is low-dimensional ($J = L/4$). As shown in Fig. 12, using $\hat{\mathbf{\Psi}}^K$ results in a lower interference power level at the output of the PD-NAMF than $\hat{\mathbf{\Psi}}^S$, which explains its superior performance relative to the latter. The interference level of both of the correlation estimators decreases as Q increases up to $Q = 4J$, beyond this point the

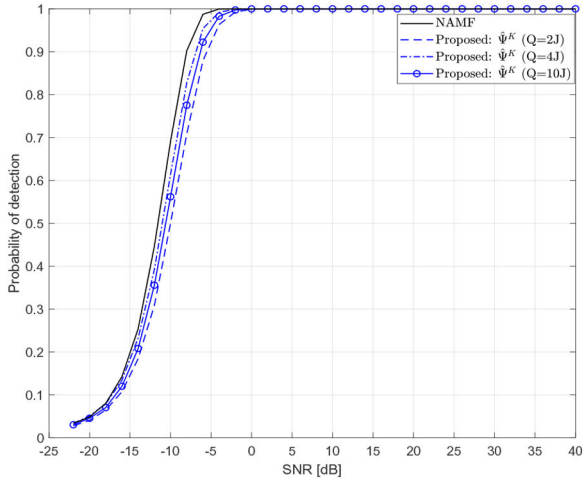

Fig. 15 *Detection performance in Gaussian distributed clutter with different Q* ($\hat{\mathbf{\Psi}}^K$, $J = 16, L_2 = 64$)

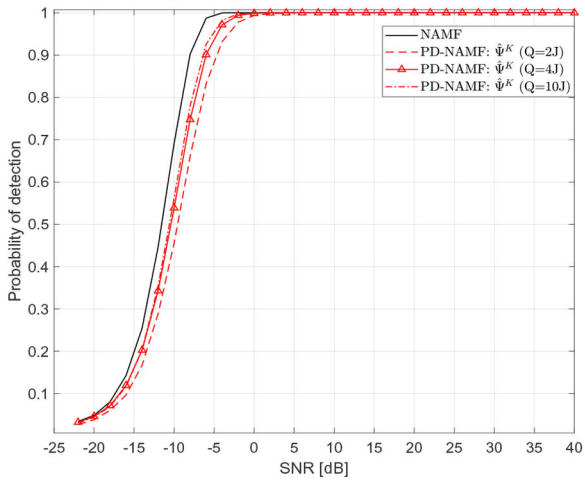

Fig. 16 *Detection performance in Gaussian distributed clutter with different Q* ($\hat{\mathbf{\Psi}}^S$, $J = 16, L_2 = 64$)

interference level is almost constant with $\hat{\mathbf{\Psi}}^S$ while it decreases slightly using $\hat{\mathbf{\Psi}}^K$. In general, it is also obvious that the interference power at the output of the PD-NAMF for $\hat{\mathbf{\Psi}}^K$ and $\hat{\mathbf{\Psi}}^S$ is lower than that at the output of the NAMF detector. This demonstrates the validity of the proof in Appendix 1. Fig. 12 also reveals that Q has a negligible effect on the target signal level at the output of the PD-NAMF for both decorrelators with a relative higher target's signal level for $\hat{\mathbf{\Psi}}^K$ than that of $\hat{\mathbf{\Psi}}^S$.

The final choice of the minimum required Q is based on the detection performance of the PD-NAMF as depicted in Fig. 13, where $\hat{\mathbf{\Psi}}^S$ is used in the presence of K-distributed clutter ($\delta \sim \mathcal{U}(0,1]$) and $J = L_2/4$. We can see that the increase of Q beyond $4J$ has a negligible impact on the detection performance and for most values of SNR there is no difference in the performance. When Q is reduced to $2J$, P_D decreases slightly with a maximum loss of 0.022. The same is shown for $\hat{\mathbf{\Psi}}^K$ in Fig. 14, but with a slight improvement with $Q = 10J$ at lower SNR values even over the NAMF. To investigate the dependence of Q on the clutter distribution, we performed additional simulations for different values of Q, but in the presence of Gaussian clutter. As shown in Figs. 15 and 16, for both $\hat{\mathbf{\Psi}}^K$ and $\hat{\mathbf{\Psi}}^S$, the performances of the PD-NAMF in the Gaussian clutter for different values of Q exhibit the same trend as in the K-distributed clutter shown in Figs. 13 and 14. Therefore, we can conclude that using $Q = 4J$ projections is an appropriate rule of thumb that does not depend on the clutter distribution.

Table 1 Performed operations by the proposed and the NAMF tests

	NAMF		Proposed (Kendall)		Proposed (Spearman)	
	Complexity	Repetitions	Complexity	Repetitions	Complexity	Repetitions
multiplication	$\mathcal{O}(J^2)$	$5L_2$	$\mathcal{O}(J^2L_1)$	$1/L_1$	$\mathcal{O}(J^2L_2)$	1
—	$\mathcal{O}(J)$	L_2	$\mathcal{O}(QJ)$	2	$\mathcal{O}(QJ)$	2
—	$\mathcal{O}(J^3)$	$4L_2+3$	$\mathcal{O}(QJL_2)$	2	$\mathcal{O}(QJL_2)$	2
inverse	$\mathcal{O}(J^3)$	5	$\mathcal{O}(J^3)$	$1/L_1$	$\mathcal{O}(J^3)$	1
sort	—	—	$\mathcal{O}(QL_2\log(L_2))$	4	$\mathcal{O}(QL_2\log(L_2))$	4
correlation	—	—	$\mathcal{O}(J^2L_1^2)$	$1/L_1$	$\mathcal{O}(J^2L_2\log(L_2))$	1

5.2.4 Complexity analysis: The complexity of the PD-NAMF (with both $\hat{\boldsymbol{\Psi}}^K$ and $\hat{\boldsymbol{\Psi}}^S$) compared to the NAMF is analysed in terms of the required arithmetic operations and the run time of each detector. Table 1 summarises the mathematical operations performed by the PD-NAMF compared to the NAMF for each secondary cell. The reported complexities are based on the Gaussian–Jordan elimination, Schoolbook, and merge sort algorithms for matrix inversion, matrix multiplication, and sorting, respectively [51]. By $1/L_1$ we mean that $\hat{\boldsymbol{\Psi}}^K$ is calculated once for all the L_1 secondary cells and not for each cell in contrast to $\hat{\boldsymbol{\Psi}}^S$, which is estimated for each secondary cell from the remaining L_2 cells.

The computation reduction is more obvious in Fig. 17, where the run times of the NAMF detector and the PD-NAMF are computed on the same platform dedicated only for this job. For both versions of the PD-NAMF with $\hat{\boldsymbol{\Psi}}^S$ and $\hat{\boldsymbol{\Psi}}^K$, the figure shows their average run times normalised by the run time of NAMF with $L_2 = 4J$ secondary cells for different J. It is conspicuous that the PD-NAMF, either with $\hat{\boldsymbol{\Psi}}^S$ or $\hat{\boldsymbol{\Psi}}^K$, substantially reduces the NHD run time depending on J. The larger the dimension of the cell J (and consequently L_2), the greater the reduction, which is of a great importance for modern radar systems with large antenna arrays.

Remark 2: The complexity of the PD-NAMF can be reduced further using parallel processing, given the independence of the random projections from each other. Moreover, the use of parallel programming on graphical processing units (GPUs) can reduce the complexity of calculating the median, and consequently the MAD, as in [52]. Furthermore, the calculation of $\hat{\boldsymbol{\Psi}}^K$, the most computationally demanding step, can be parallelised as well [53]. Nevertheless, the parallelisation is only possible partially in the NAMF due to the iterative nature of the robust covariance estimators.

6 Conclusion

In this paper, we introduced a novel covariance-free, non-parametric NHD detector for correlated clutter environments with Gaussian and non-Gaussian distributions. Based on the projection depth function, the proposed PD-NAMF avoids the computationally expensive estimation of the covariance matrix. Interestingly, the larger the dimension of the radar signal vector, the higher the computation reduction the PD-NAMF provides relative to the NAMF. This advantage fosters the application of the PD-NAMF in modern radars with large antenna arrays. Further, this significant complexity reduction is not achieved at the expense of degraded performance. That is, the detection performance of the new detector is shown to be comparable to, and in some cases better than, the full adaptive NAMF detector at different dimensions and clutter distributions. With this robust performance and the considerable reduction in computations, the PD-NAMF is superior to its covariance-based counterparts in the literature for real-time applications and it can be a more efficient replacement of the computationally demanding GIP and NAMF detectors in iterative NHD approaches. The feasible utilisation of parallel processing and GPUs paves the way for more efficient implementations of the PD-NAMF in the future.

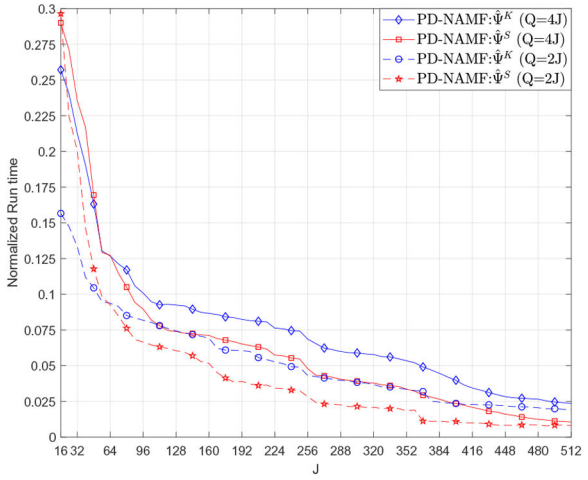

Fig. 17 *Run times of the PD-NAMF normalised by that of NAMF ($L_2 = 4J$)*

7 References

[1] Budge, M., German, S.: '*Basic radar analysis*' (Artech House, Norwood, MA, USA, 2015)

[2] Guerci, J.: '*Space-time adaptive processing for radar*' (Artech House, Norwood, MA, USA, 2014)

[3] Abouelfadl, A.A., Ahmed, F.M., Soliman, M.S.: 'A novel noise-free jamming technique against LFM-PC search radar'. Int. Japan-Egypt Conf. on Electronics, Communication and Computers (JEC-ECC), Cairo, Egypt, 31 May - 2 June 2016, pp. 127–130

[4] Abouelfadl, A.A., Samir, A.M., Ahmed, F.M., *et al.*: 'Performance analysis of LFM pulse compression radar under effect of convolution noise jamming'. National Radio Science Conf. (NRSC), Aswan, Egypt, February 2016, pp. 282–289

[5] Melvin, W.L., Wicks, M.C.: 'Improving practical space-time adaptive radar'. Proc. of the IEEE National Radar Conf., Syracuse, NY, USA, May 1997, pp. 48–53

[6] Rangaswamy, M.: 'Statistical analysis of the nonhomogeneity detector for non-Gaussian interference backgrounds', *IEEE Trans. Signal Process.*, 2005, **53**, pp. 2101–2111

[7] Jiang, L., Wang, T.: 'Robust non-homogeneity detector based on reweighted adaptive power residue', *IET Radar, Sonar Navig.*, 2016, **10**, pp. 1367–1374

[8] Han, S., Maio, A.D., Carotenuto, V., *et al.*: 'Censoring outliers in radar data: an approximate ML approach and its analysis', *IEEE Trans. Aerosp. Electron. Syst.*, 2019, **55**, (2) pp. 534–546

[9] Han, S., Pallotta, L., Carotenuto, V., *et al.*: 'An approximate regularized ML approach to censor outliers in Gaussian radar data', *IEEE Access*, 2019, **7**, pp. 66263–66274

[10] Setlur, P., Rangaswamy, M.: 'A family of random and random type projections for radar STAP'. IEEE Radar Conf. (RadarConf18), Oklahoma City, OK, USA, April 2018, pp. 0856–0861

[11] Li, Z., Liu, H., Zhang, Y., *et al.*: 'Robust nonhomogeneous training samples detection method for space-time adaptive processing radar using sparse-recovery with knowledge-aided', *J. Appl. Remote Sens.*, 2017, **11**, (4), pp. 045013-1–045013-13

[12] Wang, W., Zou, L., Wang, X., *et al.*: 'Deterministic-aided single dataset stap method based on sparse recovery in heterogeneous clutter environments', *EURASIP J. Adv. Signal Process.*, 2018, **2018**, p. 24

[13] Bowen, Z., Zhiguang, D., Lan, X.: 'Secondary non-homogeneity detector for STAP'. IEEE 3rd Int. Conf. on Signal and Image Processing (ICSIP), Shenzhen, China, July 2018, pp. 382–384

[14] Michels, J.H., Rangaswamy, M., Himed, B.: 'Performance of parametric and covariance based STAP tests in compound-Gaussian clutter', *Digit. Signal Process.*, 2002, **12**, (2), pp. 307–328

[15] Fan, J., Liao, Y., Liu, H.: 'An overview of the estimation of large covariance and precision matrices', *Econ. J.*, 2016, **19**, (1), pp. C1–C32

[16] Schoenig, G.N., Picciolo, M.L., Mili, L.: 'Improved detection of strong nonhomogeneities for STAP via projection statistics'. IEEE Int. Radar Conf., Arlington, VA, USA, May 2005, pp. 720–725

[17] Huber, P.J.: '*Robust statistics*' (Wiley, Hoboken, NJ, USA, 2009, 2nd edn.)

[18] Abouelfadl, A.A., Psaromiligkos, I., Champagne, B.: 'A low-complexity nonparametric STAP detector'. IEEE National Aerospace and Electronics Conf. (NAECON), Dayton, OH, USA, July 2018, pp. 592–596

[19] Richards, M., Holm, W., Melvin, W., *et al.*: '*Principles of modern radar: basic principles*' (Institution of Engineering and Technology, Stevenage, UK, 2012)

[20] Adali, T., Schreier, P.J.: 'Optimization and estimation of complex-valued signals: theory and applications in filtering and blind source separation', *IEEE Signal Process. Mag.*, 2014, **31**, pp. 112–128

[21] Conte, E., Maio, A.D., Ricci, G.: 'Recursive estimation of the covariance matrix of a compound-Gaussian process and its application to adaptive CFAR detection', *IEEE Trans. Signal Process.*, 2002, **50**, pp. 1908–1915

[22] Conte, E., Longo, M.: 'Characterisation of radar clutter as a spherically invariant random process', *IEE Proc. F – Commun., Radar Signal Process.*, 1987, **134**, pp. 191–197

[23] Rangaswamy, M., Weiner, D., Ozturk, A.: 'Computer generation of correlated non-Gaussian radar clutter', *IEEE Trans. Aerosp. Electron. Syst.*, 1995, **31**, pp. 106–116

[24] Soloveychik, I., Wiesel, A.: 'Tyler's covariance matrix estimator in elliptical models with convex structure', *IEEE Trans. Signal Process.*, 2014, **62**, pp. 5251–5259

[25] Pulsone, N.B., Raghavan, R.S.: 'Analysis of an adaptive CFAR detector in non-Gaussian interference', *IEEE Trans. Aerosp. Electron. Syst.*, 1999, **35**, (3), pp. 903–916

[26] He, Y., Jian, T., Su, F., *et al.*: 'CFAR assessment of covariance matrix estimators for non-Gaussian clutter', *Sci. China Inf. Sci.*, 2010, **53**, pp. 2343–2351

[27] Zuo, Y., Serfling, R.: 'General notions of statistical depth function', *Ann. Stat.*, 2000, **28**, pp. 461–482

[28] Zuo, Y.: 'Projection-based depth functions and associated medians', *Ann. Stat.*, 2003, **31**, pp. 1460–1490

[29] McKean, J.W., Hettmansperger, T.P.: 'Rank-based analysis of linear models and beyond: a review', in Liu, R.Y., McKean, J.W. (Eds.): '*Robust rank-based and nonparametric methods*' (Springer International Publishing, Cham, Switzerland, 2016), pp. 1–24,

[30] Velasco-Forero, S., Angulo, J.: 'Random projection depth for multivariate mathematical morphology', *IEEE J. Sel. Top. Signal Process.*, 2012, **6**, pp. 753–763

[31] Hayes, K.: 'Finite-sample bias-correction factors for the median absolute deviation', *Commun. Stat. - Simul. Comput.*, 2014, **43**, (10), pp. 2205–2212

[32] Olive, D.J.: 'Sufficient statistics', in '*Statistical theory and inference*' (Springer International Publishing, Cham, Switzerland, 2014), pp. 215–256

[33] Rousseeuw, P.J., Croux, C.: 'Alternatives to the median absolute deviation', *J. Am. Stat. Assoc.*, 1993, **88**, (424), pp. 1273–1283

[34] Cuesta-Albertos, J., Nieto-Reyes, A.: 'The random Tukey depth', *Comput. Stat. Data Anal.*, 2008, **52**, (11), pp. 4979–4988

[35] Muller, M.E.: 'A note on a method for generating points uniformly on N-dimensional spheres', *Commun. ACM*, 1959, **2**, pp. 19–20

[36] Ward, J.: 'Space-time Adaptive Processing for Airborne Radar', Technical report, Massachusetts Institute of Technology, Lincoln Laboratory, 1994

[37] Younis, M., Laux, C., Loinger, A., *et al.*: 'Concept and performance of internal instrument calibration for multi-channel SAR'. 12th European Conf. on Synthetic Aperture Radar (EUSAR), Aachen, Germany, June 2018, pp. 1–5

[38] Attia, E., Steinberg, B.: 'Self-cohering large antenna arrays using the spatial correlation properties of radar clutter', *IEEE Trans. Antennas Propag.*, 1989, **37**, (1), pp. 30–38

[39] Aubry, A., Maio, A.D., Pallotta, L.: 'A geometric approach to covariance matrix estimation and its applications to radar problems', *IEEE Trans. Signal Process.*, 2018, **66**, pp. 907–922

[40] Gini, F., Greco, M.: 'Covariance matrix estimation for CFAR detection in correlated heavy tailed clutter', *Signal Process.*, 2002, **82**, (12), pp. 1847–1859

[41] Sharma, S.C.: 'The effects of correlation among observations on the consistency property of sample variance', *Commun. Stat. - Theory Methods*, 1986, **15**, (4), pp. 1125–1152

[42] Azriel, D., Schwartzman, A.: 'The empirical distribution of a large number of correlated normal variables', *J. Am. Stat. Assoc.*, 2015, **110**, (511), pp. 1217–1228

[43] Xu, W., Hou, Y., Hung, Y. S., *et al.*: 'A comparative analysis of Spearman's rho and Kendall's tau in normal and contaminated normal models', *Signal Process.*, 2013, **93**, (1), pp. 261–276

[44] Loh, P.-L., Tan, X.L.: 'High-dimensional robust precision matrix estimation: cellwise corruption under ε-contamination', *Electron. J. Stat.*, 2018, **12**, (1), pp. 1429–1467

[45] Zhao, T., Liu, H.: 'Calibrated precision matrix estimation for high-dimensional elliptical distributions', *IEEE Trans. Inf. Theory/Prof. Tech. Group Inf. Theory*, 2014, **60**, (12), pp. 7874–7887

[46] Zhao, T., Roeder, K., Liu, H.: 'Positive semidefinite rank-based correlation matrix estimation with application to semiparametric graph estimation', *J. Comput. Graph. Stat.*, 2014, **23**, (4), pp. 895–922

[47] Melebari, A., Mishra, A.K., Gaffar, M.Y.A.: 'Statistical analysis of measured high resolution land clutter at X-band and clutter simulation'. European Radar Conf. (EuRAD), Paris, France, September 2015, pp. 105–108

[48] Antipov, I.: 'Analysis of sea clutter data', DSTO Electronic and Surveillance Research Laboratory Salisbury, Australia, 1998

[49] Croux, C., Dehon, C.: 'Influence functions of the Spearman and Kendall correlation measures', *Stat. Methods. Appt.*, 2010, **19**, pp. 497–515

[50] Himonas, S.D., Barkat, M.: 'Adaptive CFAR detection in partially correlated clutter', *IEE Proc. F - Radar Signal Process.*, 1990, **137**, pp. 387–394

[51] Cormen, T.H., Leiserson, C.E., Rivest, R.L., *et al.*: '*Introduction to algorithms*' (The MIT Press, Cambridge, MA, USA, 2009, 3rd Edn.)

[52] Blanchard, J.D., Opavsky, E., Uysaler, E.: 'Selecting multiple order statistics with a graphics processing unit', *ACM Trans. Parallel Comput.*, 2016, **3**, pp. 1–23

[53] Liu, Y., Pan, T., Green, O., *et al.*: 'Parallelized Kendall's tau coefficient computation via SIMD vectorized sorting on many-integrated-core processors', ArXiv e-prints, p. arXiv:1704.03767, 2017, pp. 1–29

8 Appendix

8.1 Appendix 1: Proof of Proposition 1

Proof: Assuming that the covariance matrix $\hat{\boldsymbol{R}}$ is positive definite, it is invertible and admits of a square-root $\hat{\boldsymbol{R}}^{1/2}$, which is also invertible. Hence, applying the Cauchy–Schwartz inequality

$$\begin{aligned}|\boldsymbol{u}^H\boldsymbol{z}_k|^2 &= \left|\boldsymbol{u}^H\hat{\boldsymbol{R}}^{\frac{1}{2}}\hat{\boldsymbol{R}}^{-\frac{1}{2}}\boldsymbol{z}_k\right|^2 \le \|\,\boldsymbol{u}^H\hat{\boldsymbol{R}}^{\frac{1}{2}}\|^2\ \|\,\hat{\boldsymbol{R}}^{-\frac{1}{2}}\boldsymbol{z}_k\|^2 \\ &\le (\boldsymbol{u}^H\hat{\boldsymbol{R}}\boldsymbol{u})(\boldsymbol{z}_k^H\boldsymbol{R}^{-1}\boldsymbol{z}_k)\end{aligned} \tag{45}$$

Consequently, for any non-zero vector $\boldsymbol{u} \in \mathbb{C}^J$, we have

$$\frac{|\boldsymbol{u}^H\boldsymbol{z}_k|^2}{\boldsymbol{u}^H\hat{\boldsymbol{R}}\boldsymbol{u}} \le \boldsymbol{z}_k^H\hat{\boldsymbol{R}}^{-1}\boldsymbol{z}_k \tag{46}$$

Hence

$$\sup_{\|\,\boldsymbol{u}\,\|=1}\frac{|\boldsymbol{u}^H\boldsymbol{z}_k|^2}{\boldsymbol{u}^H\hat{\boldsymbol{R}}\boldsymbol{u}} = \boldsymbol{z}_k^H\hat{\boldsymbol{R}}^{-1}\boldsymbol{z}_k \tag{47}$$

Among the possible scale measures of the scalar random variable $\boldsymbol{u}^H\boldsymbol{z}_k$, let us consider the sample variance $\hat{\sigma}^2(\boldsymbol{u}^H\boldsymbol{Z}_k) = \boldsymbol{u}^H\hat{\boldsymbol{R}}\boldsymbol{u}$. Then according to (47)

$$\sup_{\|\,\boldsymbol{u}\,\|=1}\left(\frac{|\boldsymbol{u}^H\boldsymbol{z}_k|}{\hat{\sigma}(\boldsymbol{u}^H\boldsymbol{Z}_k)}\right)^2 = \boldsymbol{z}_k^H\hat{\boldsymbol{R}}^{-1}\boldsymbol{z}_k \tag{48}$$

Based on (9), $\hat{\sigma}(\boldsymbol{u}^H\boldsymbol{Z}_k)$ is given by

$$\hat{\sigma}(\boldsymbol{u}^H\boldsymbol{Z}_k) = 2\hat{\sigma}(\Re(\boldsymbol{u}^H\boldsymbol{Z}_k)) = 2\hat{\sigma}(\Im(\boldsymbol{u}^H\boldsymbol{Z}_k)) \tag{49}$$

Repeating the same steps from (45)–(48), but with replacing $\boldsymbol{z}_k$ by $\boldsymbol{s}$, a similar relation can be proven for the steering vector $\boldsymbol{s}$, i.e.

$$\sup_{\|\,\boldsymbol{u}\,\|=1}\left(\frac{|\boldsymbol{u}^H\boldsymbol{s}|}{\hat{\sigma}(\boldsymbol{u}^H\boldsymbol{Z}_k)}\right)^2 = \boldsymbol{s}^H\hat{\boldsymbol{R}}^{-1}\boldsymbol{s} \tag{50}$$

For the centralised vector $\boldsymbol{z}_k - \hat{\mu}$, we can write

$$\sup_{\|\,\boldsymbol{u}\,\|=1}\left(\frac{|\boldsymbol{u}^H\boldsymbol{z}_k - \hat{\mu}|}{\hat{\sigma}(\boldsymbol{u}^H\boldsymbol{Z}_k)}\right)^2 = (\boldsymbol{z}_k - \hat{\mu})^H\hat{\boldsymbol{R}}^{-1}(\boldsymbol{z}_k - \hat{\mu}) \qquad \square \tag{51}$$

8.2 Appendix 2: Proof of proposition 2

Proof: Considering the kth secondary cell $\boldsymbol{z}_k = a\boldsymbol{s} + \boldsymbol{c} + \boldsymbol{n}$, where for convenience we let $\boldsymbol{s} = [s_1, \ldots, s_J]^T$, $\boldsymbol{c} = [c_1, \ldots, c_J]^T$, and $\boldsymbol{n} = [n_1, \ldots, n_J]^T$, then we have

$$\boldsymbol{s}^H \boldsymbol{z}_k \boldsymbol{s}^H \hat{\boldsymbol{R}}^{-1} \boldsymbol{s} = \left(a \sum_{i=1}^{J} s_i s_i^* + \sum_{i=1}^{J} s_i^* (c_i + n_i) \right) \cdot \sum_{i=1}^{J} \left(s_i^* \sum_{j=1}^{J} p_{ij} s_j \right) \tag{52}$$

where $\hat{\boldsymbol{R}}^{-1} = [p_{ij}]$. From (3), $\sum_{i=1}^{J} s_i s_i^* = 1$, hence

$$\boldsymbol{s}^H \boldsymbol{z}_k \boldsymbol{s}^H \hat{\boldsymbol{R}}^{-1} \boldsymbol{s} = a \sum_{i=1}^{J} s_i^* \sum_{j=1}^{J} p_{ij} s_j + \sum_{j=1}^{J} \left(s_j^* \sum_{i=1}^{J} \left(s_i^* \sum_{k=1}^{J} p_{ik} s_k \right) (c_j + n_j) \right) \tag{53}$$

$$\boldsymbol{s}^H \hat{\boldsymbol{R}}^{-1} \boldsymbol{z}_k = a \sum_{i=1}^{J} s_i^* \sum_{j=1}^{J} p_{ij} s_j + \sum_{i=1}^{J} s_i^* \sum_{j=1}^{J} p_{ij} (c_j + n_j) \tag{54}$$

From (53) and (54), we can observe that the two tests have the same target signal component (i.e. the first term in each equation, which is equivalent to $a(\boldsymbol{s}^H \hat{\boldsymbol{R}}^{-1} \boldsymbol{s})$), but they differ in the interference component. There is no analytical way to compare the interference components in the two tests due to the different random weights of each term, hence, we rely on simulation to compare them. Nevertheless, as (53) considers only the interference component in the spatio-temporal direction of the target, its average interference power is anticipated to be lower than that of (54), which is confirmed by simulation in Section 5. □

8.3 Appendix 3: Proof of proposition 3

Proof: Using the term in (27), we obtain the modified NAMF test statistic

$$\Lambda'_{\text{NAMF}} = \frac{\left| \boldsymbol{s}^H \hat{\boldsymbol{R}}^{-1} (\boldsymbol{s}^H \boldsymbol{z}_k) \boldsymbol{s} \right|^2}{(\boldsymbol{s}^H \hat{\boldsymbol{R}}^{-1} \boldsymbol{s})(\boldsymbol{z}_k^H \hat{\boldsymbol{R}}^{-1} \boldsymbol{z}_k)} \tag{55}$$

Based on (50), we have

$$\left| (\boldsymbol{s}^H \boldsymbol{z}_k) \boldsymbol{s}^H \hat{\boldsymbol{R}}^{-1} \boldsymbol{s} \right|^2 = \left| (\boldsymbol{s}^H \boldsymbol{z}_k) \left(\sup_{\| \boldsymbol{u} \| = 1} \frac{|\boldsymbol{u}^H \boldsymbol{s}|}{\hat{\sigma}(\boldsymbol{u}^H \boldsymbol{Z}_k)} \right)^2 \right|^2 \tag{56}$$

By substitution of (48), (50), and (56) into (55) and after simple manipulations, Λ'_{NAMF} reduces to

$$\Lambda'_{\text{NAMF}} = \frac{\left| \boldsymbol{s}^H \boldsymbol{z}_k \right|^2 \sup_{\| \boldsymbol{u} \| = 1} \left(\frac{|\boldsymbol{u}^H \boldsymbol{s}|}{\hat{\sigma}(\boldsymbol{u}^H \boldsymbol{Z}_k)} \right)^2}{\sup_{\| \boldsymbol{u} \| = 1} \left(\frac{|\boldsymbol{u}^H \boldsymbol{z}_k|}{\hat{\sigma}(\boldsymbol{u}^H \boldsymbol{Z}_k)} \right)^2} \quad \square \tag{57}$$