Kernelized-Likelihood Ratio Tests for Binary Phase-Shift Keying Signal Detection

Ahmadreza Salehi, Amir Zaimbashi, *Member, IEEE* and Mikko Valkama, *Senior Member, IEEE*

***Abstract*—In this paper, kernelized-likelihood ratio tests (LRTs) for binary phase-shift keying (BPSK) signal detection based on the polynomial kernel function are proposed. Specifically, we kernelize the conventional LRT of BPSK signal detection using the so-called kernel trick, such that the inner product of the conventional LRT is replaced with proper polynomial kernel functions allowing for richer feature space to be deployed in the detection. We also derive computationally efficient recursive implementation structures for the proposed methods, resulting overall in six new detectors. With respect to the noise variance uncertainty (NVU), the proposed detectors can be divided into two general classes, namely i) constant false alarm rate (CFAR) and ii) semi-CFAR (S-CFAR) methods. To facilitate efficient operation under NVU, we also propose a new threshold-setting strategy to adjust the level of the proposed S-CFAR detectors. Additionally, we address the well-known energy detector (ED) under NVU and devise a new fixed-level ED formulation while also obtaining closed-form expressions for its false alarm and detection probabilities. Our extensive simulation results show that the proposed S-CFAR detectors outperform the state-of-the-art BPSK signal detectors with 2.4 dB signal-to-noise ratio (SNR) gain under practical worst-case NVU assumptions, while the performance gain is approximately 5.7 dB without NVU. In the case of the proposed CFAR detectors, the corresponding improvement in detection performance is approximately 1.8 dB.**

***Index Terms*—Detection theory, kernel theory, signal detection, BPSK signal, CFAR and semi-CFAR detector, Likelihood ratio test, noise variance uncertainty.**

I. Introduction

Kernel theory provides a powerful new tool for mathematicians, scientists, and engineers for nonlinear signal processing. More precisely, kernel-based approaches can provide efficient solutions in real-world nonlinear decision making problems, facilitating detection and signal processing algorithms in which the nonlinear characteristics of the input data are exploited through the use of nonlinear kernel functions [1]. Kernel functions are similarity measures that implicitly exploit the structural information of the input data, through appropriate mapping to a higher-dimensional space, known as the feature space [2]. Kernel methods have various applications in, for example, machine learning [3]-[4], pattern recognition [5]-[6], computer vision [7], dictionary learning [8] and fault detection [9].

In general, detection theory is an essential ingredient in the design and optimization of digital receivers of communication systems [10]. In particular, in detection theory, we encounter three different problems, namely the so-called signal activity detection (SAD), data demodulation/detection and classification (or pattern recognition). In all these problems, we need to decide among two or more possible hypotheses based on an observed data set [11]-[12]. Thus, the central problem addressed in the detection theory is to form a function of the observed data set, namely test statistic, and then make a decision based on the obtained test statistic and an appropriate detection threshold. Among the different decision making problems, the SAD is the simplest detection problem, where we aim to decide whether a specific signal is present or whether the received signal contains only noise. Such statistical problem is referred to as the binary hypothesis testing problem since we wish to use the observed data as efficiently as possible in deciding between two possible hypotheses. In general, there are various detection problems in digital communication systems, radar systems, sonar systems, and image processing systems, which can usually be cast as SAD problems [13]-[14].

In the context of signal activity detection problems, there are few works that exploit the benefit of the kernel theory to discriminate between two possible hypotheses, aiming to provide more efficient decision-making methods [9], [15]-[16]. For example, an improved fault detection algorithm to enhance monitoring abilities of nonlinear biological processes was proposed in [9], relying on generalized likelihood ratio test (GLRT) and kernel-based principal component analysis (KPCA). It was shown that the kernel-based GLRT improved the

A. R. Salehi and A. Zaimbashi are with the Optical and RF Communication Systems Laboratory, Department of Electrical Engineering, Shahid Bahonar University of Kerman, Kerman, Iran (e-mail: a.zaimbashi@uk.ac.ir)

M. Valkama is with the Department of Electrical Engineering, Tampere University, Tampere, Finland (e-mail: mikko.valkama@tuni.fi).

fault detection capability, however, the detection performance was not investigated in terms of the detection theory metrics, particularly detection probability curves and false alarm behavior. A kernel-based matched subspace detector (K-MSD) for subpixel target detection in hyperspectral imagery was, in turn, proposed in [15]. It was shown that the K-MSD obtained superior detection performance compared to the conventional MSD. In [16], the KMSD algorithm was adopted for spectrum sensing tasks in a cognitive radio system to improve the detection performance of the conventional MSD in low signal-to-noise ratio (SNR) regimes. Both [15] and [16] consider only real-valued data, where Gaussian radial basis kernels (GRBF) and real-valued polynomial kernels are exploited. In the context of cooperative spectrum sensing, a kernelized Fisher Discriminant Analysis (FDA) has been proposed in [2] for the decision fusion task. For non-cooperative cases, authors of [17] have introduced a kernelized feature template matching (KFTM) based spectrum sensing method. The main drawback of the KFTM spectrum sensing method is that this algorithm needs a-priori knowledge about the principal eigenvector of the primary user signal. In [18], the kernelized energy detection (KED) approach in Gaussian and non-Gaussian noise scenarios was developed for non-cooperative spectrum sensing purposes. This method can be applied only when the receiver is equipped with multiple antennas – otherwise, the test statistic of the devised KED cannot be computed.

In this paper, we capitalize and combine the idea of kernel theory and detection theory in order to develop new SAD algorithms for binary-phase-shift-keying (BPSK) signal detection. To this end, we specifically consider the SAD problem for BPSK signals introduced in [19] and [20], where the authors proposed two SAD algorithms based on the likelihood ratio test (LRT) principle operating in the original input space. In this paper, two general classes of kernelized-likelihood ratio tests for BPSK signal detection based on the polynomial kernel function are proposed and formulated. In the first class, we propose a semi-constant false alarm rate (S-CFAR) detector, while in the second class, some CFAR methods with respect to noise variance uncertainty (NVU) are proposed. Additionally, we derive computationally efficient implementation structures, resulting finally in six new detectors. For the proposed S-CFAR methods, we also propose a new threshold-setting strategy to efficiently determine the threshold level under given worst-case assumptions regarding the level of NVU. Our extensive simulation results show that the proposed detectors significantly outperform the state-of-the-art BPSK signal detectors. This significant improvement in detection performance is due to the richer feature space allowed for by the kernel-based approach, where the simple similarity measure of inner product in the conventional LRT is replaced with appropriate polynomial kernel functions.

Inspired by the kernel theory and detection theory, the main contributions of this article can be summarized as follows:

- New SAD algorithms for binary-phase-shift-keying (BPSK) signal detection are proposed. Specifically, two general classes of kernelized-likelihood ratio tests for BPSK signal detection based on the polynomial kernel function are derived and formulated, where we replace the inner products of tests statistics with that of nonlinear mapped data in the feature space or that of their polynomial kernel to achieve improved detection performance.
- The principle of invariance is exploited to examine the potential CFAR behavior of the proposed kernel-based detectors against noise variance uncertainty (NVU). It is shown that two of the proposed detectors have CFAR property against the NVU, while the other proposed detectors are of semi-CFAR (S-CFAR) nature.
- Some of the proposed detectors build, in their basic form, on exhaustive search, hence we also derive and report new computationally efficient recursive implementation structures. This results overall in six alternative detectors, four of which possess CFAR property while the two are of S-CFAR nature.
- For the S-CFAR methods and the well-known energy detector (ED), we propose a new threshold-setting strategy under NVU, such that the false alarm probability is limited to a given upper bound under the assumed worst-case NVU. In such scenarios, we obtain closed-form expressions for the false alarm and detection probabilities of the modified ED, referred to as the fixed-level ED (FL-ED) in the continuation. This assists in devising new analytical threshold-setting strategies and to quantify the effects of NVU on the detection performance of the proposed FL-ED detector.
- Our extensive simulation results show that the proposed S-CFAR detectors provide mutually comparable detection performance, while outperforming the state-of-the-art BPSK signal detectors by 2.4 dB SNR gain under practical worst-case assumptions of NVU, and by 5.7 dB under zero NVU. In case of the proposed CFAR detectors, the corresponding improvement in detection performance is about 1.8 dB. We also show that the proposed SC-CFAR

methods significantly outperform the proposed FL-ED detector.

- Assuming that each element of the vectors in the considered test statistics is referred to as feature, we will show that the features are all the monomials of degree 1 in the classical BPSK signal detection algorithms while being monomials of degree up to d when transforming the original data into the feature space equivalently through an appropriate polynomial kernel functions of order d. The above significant improvements in detection performance are shown to be stemming from the exploitation of this new richer space, the so-called feature space. Thus, it is shown that in the original LRT detectors, only the inner product of all monomials with degree one is computed, while in case of the proposed detectors, the inner product of all monomials up to degree d are computed.

The rest of this paper is organized as follows: in Section II, the SAD problem of BPSK signals is shortly formulated as a composite hypothesis testing problem, and the LRT-based algorithm of this problem is presented, for reference and comparison purposes, based on [19] and [20]. In Section III, firstly, some preliminaries about kernel theory are shortly introduced. Then, we propose and formulate two general classes of kernelized SAD methods. In Section IV, we investigate the CFAR property of the proposed detectors and introduce a new threshold-setting strategy to tailor the detection thresholds of the proposed S-CFAR detectors. Additionally, we formulate the new energy-based detector under NVU and obtain its closed-form expressions for the false alarm and detection probabilities. Computationally efficient recursive structures of the proposed kernel-based detectors are described in Section V. Extensive simulations results are provided and analyzed in section VI, while Section VII provides the conclusions of the work.

Notation: Throughout the paper, scalars are denoted by non-boldface type, vectors by boldface lowercase letters, and matrices by boldface uppercase letters. The superscripts T and H denote matrix or vector transpose and hermitian, respectively, while $\mathbf{I}$ denotes the identity matrix. We use the notation $\mathcal{CN}(\boldsymbol{\mu}, \mathbf{C})$ to denote the circularly symmetric complex Gaussian distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\mathbf{C}$. The modulus of x is denoted by $|x|$. The inner product of two complex vectors $\mathbf{a}$ and $\mathbf{b}$ is defined as $\mathbf{a}^H\mathbf{b}$ and is denoted by $\langle \mathbf{a}, \mathbf{b} \rangle$.

II. Problem Formulation

We consider the SAD problem introduced in [19] and [20] to decide if a BPSK signal is present (i.e., H_1 hypothesis) or not (i.e., H_0 hypothesis). It is further assumed that there is a set of N discrete-time received samples available, expressed as a vector $\mathbf{y} = [y_1, ..., y_N]^T$. The SAD problem can thus be expressed as a composite hypothesis testing problem, written as [19]

$$\begin{cases} H_0 : \mathbf{y} = \mathbf{w} \\ H_1 : \mathbf{y} = \alpha \mathbf{s} + \mathbf{w}. \end{cases} \tag{1}$$

where

- The noise vector $\mathbf{w} \in \mathcal{C}^N$ is a circularly symmetric complex Gaussian (CSCG) random vector with zero mean and a covariance matrix $\sigma^2\mathbf{I}$ with unknown σ^2, i.e., $\mathbf{w} \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I})$.
- The scalar $\alpha = |\alpha|e^{j\phi}$ is an unknown complex gain of the received BPSK signal, if present, accounting for the channel effects.
- The elements of the vector $\mathbf{s} \in \mathcal{C}^N$ are contained to the set of BPSK alphabets $\mathcal{B} = \{-1, 1\}$, i.e, $\mathbf{s} \in \mathcal{B}^N$ or $s_n \in \mathcal{B}$ where s_n is the n-th entry of the vector $\mathbf{s}$.
- Similar to [19] and [20], we assume that there is no intersymbol interference (ISI) and that there are no inaccuracies in time or carrier synchronization.

For the above SAD problem, the LRT (or GLRT) has been derived [19] and can be expressed as

$$T_1(\mathbf{y}) = \frac{\max_{\mathbf{s}\in\mathcal{B}^N} \left\{ |\langle \mathbf{s}, \mathbf{y} \rangle|^2 \right\}}{\langle \mathbf{y}, \mathbf{y} \rangle} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_1 \tag{2}$$

where $\langle \mathbf{u}, \mathbf{v} \rangle$ denotes the inner product of vectors $\mathbf{u}$ and $\mathbf{v}$, defined as $\mathbf{u}^H\mathbf{v}$. Additionally, the LRT can be expressed in an alternative form as

$$\Lambda_1(\mathbf{y}) = \max_{\mathbf{s}\in\mathcal{B}^N} \left\{ |\langle \mathbf{s}, \mathbf{g} \rangle|^2 \right\} \underset{H_0}{\overset{H_1}{\gtrless}} \zeta_1 \tag{3}$$

where $\mathbf{g} = \frac{\mathbf{y}}{\|\mathbf{y}\|}$, η_1 and ζ_1 are, respectively, the detection threshold of the LRT detectors presented in (2) and (3) to obtain a desired false alarm probability denoted by p_{fa}.

For comparison purposes, we also describe and consider the energy detector (ED), expressed as

$$l_1(\mathbf{y}) = \frac{||\mathbf{y}||^2}{\sigma_n^2} = \frac{\langle \mathbf{y}, \mathbf{y} \rangle}{\sigma_n^2} \underset{H_0}{\overset{H_1}{\gtrless}} \varsigma_1(\sigma_n^2), \tag{4}$$

where $\varsigma_1(\sigma_n^2)$ is the detection threshold. It is generally well-known in the literature that the detection performance of the ED is significantly degraded in the presence of the NVU. In Section IV, we will thus propose a modified ED under the NVU challenge.

III. Kernelized LRT-Based Detectors

A. Kernel Theory Basics

In kernel-based methods, in general, the original input data belonging to the data space $\mathcal{O} \subseteq C^N$ is mapped to a new space $\mathcal{F} \subseteq C^L$ with $L >> N$, known as the feature space [1]. In general, this mapping can be represented by a nonlinear mapping function φ, i.e.,

$$\varphi : \mathcal{O} \to \mathcal{F}, \quad \mathbf{y} \to \varphi(\mathbf{y}), \quad \mathbf{s} \to \varphi(\mathbf{s}), \quad \mathbf{g} \to \varphi(\mathbf{g}) \tag{5}$$

where $\mathbf{y}$ and $\mathbf{s}$ are input vectors in $\mathcal{O}$ which are mapped into a potentially much higher-dimensional feature space $\mathcal{F}$. Then, a given similarity measure (i.e., the inner product) in the original space can be replaced with that of the feature space. To this end, we build on the idea that the original space is limited in resolution, and expressive power and thus the input space $\mathcal{O}$ does not necessarily provide best support for simple inner product type of similarity measure [1]. To circumvent this, the original data space is mapped into the feature space, aiming to obtain a richer space, while the similarity between the two mapped vectors $\varphi(\mathbf{s})$ and $\varphi(\mathbf{y})$ can be measured using the corresponding inner product $\langle \varphi(\mathbf{s}), \varphi(\mathbf{y}) \rangle$. Based on this, the LRT detectors of (2) and (3) can now be re-expressed as

$$T_2(\mathbf{y}) = \frac{\max_{\mathbf{s}\in\mathcal{B}^N} \left\{ |\langle \varphi(\mathbf{s}), \varphi(\mathbf{y}) \rangle|^2 \right\}}{\langle \varphi(\mathbf{y}), \varphi(\mathbf{y}) \rangle} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_2 \tag{6}$$

and

$$\Lambda_2(\mathbf{y}) = \max_{\mathbf{s}\in\mathcal{B}^N} \left\{ |\langle \varphi(\mathbf{s}), \varphi(\mathbf{g}) \rangle|^2 \right\} \underset{H_0}{\overset{H_1}{\gtrless}} \zeta_2 \tag{7}$$

where η_2 and ζ_2 denote the new detection thresholds to achieve a given false alarm probability. It should be noted that the relationships between the feature space thresholds η_2 and ζ_2 and the original thresholds η_1 and ζ_1 are not necessarily easy to obtain even when we know the nonlinear mapping function φ.

In general, due to the high dimensionality of the feature space $\mathcal{F}$, implementing the algorithm directly in the feature space is not necessarily straight-forward or computationally efficient. For example, if φ is a polynomial of order d, then each entry of mapped data $\varphi(\mathbf{y})$ is a monomial of order d. It can be easily shown that the number of different monomials is equal to $\frac{(N+d-1)!}{d!(N-1)!}$, which is a large value if either N or d is large. In addition, it should be noted that computing the similarity measure $\langle \varphi(\mathbf{s}), \varphi(\mathbf{y}) \rangle$ is possible only if the feature mapping φ is explicit. In the kernel-based methods, the mapping function can be defined implicitly. Specifically, we need to define a function, called a kernel, such that [1]

$$\mathcal{K}(\mathbf{s}, \mathbf{x}) = \langle \varphi(\mathbf{s}), \varphi(\mathbf{x}) \rangle \tag{8}$$

where $\mathbf{s} \in \mathcal{O}$ and $\mathbf{x} \in \mathcal{O}$. Thus, the kernel function computes the similarity using the kernel function $\mathcal{K}$ in the input space $\mathcal{O}$ instead of the feature space $\mathcal{F}$. In such cases, we do not require to know the explicit form of the feature map φ, but instead, it is implicitly defined through the corresponding kernel function. This approach is generally known as the kernel trick in the kernel theory. To illustrate this in a concrete manner, we consider the polynomial kernel function with order d and bias parameter c, defined as [1]

$$\mathcal{K}(\mathbf{s}, \mathbf{x}) = (\langle \mathbf{s}, \mathbf{x} \rangle + c)^d \tag{9}$$

To show (8), we consider a simplified example, in which we utilize second-order real polynomial (SORP) kernel, i.e., $d = 2$, and assume that $N = 2$ and $c = 0$. Additionally, we consider two real vectors $\mathbf{s}$ and $\mathbf{x}$, represented as $[s_1, s_2]^T$ and $[x_1, x_2]^T$, respectively. In such case, the SORP kernel can be expanded as

$$\begin{aligned} \mathcal{K}(\mathbf{s}, \mathbf{x}) &= (s_1 x_1 + s_2 x_2)^2 \\ &= \langle \varphi(\mathbf{s}), \varphi(\mathbf{x}) \rangle = \varphi(\mathbf{s})^T \varphi(\mathbf{x}) \end{aligned} \tag{10}$$

where

$$\varphi(\mathbf{s}) = [s_1^2, s_2^2, \sqrt{2} s_1 s_2]^T \tag{11}$$

$$\varphi(\mathbf{x}) = [x_1^2, x_2^2, \sqrt{2} x_1 x_2]^T \tag{12}$$

This clearly shows that the direct computation of the inner product in the feature space and that obtained from the SORP kernel are exactly the same. Thus, using kernels, we never need the coordinates of the data in the feature space since the detection algorithms (6) and (7) only require the inner products between data in the feature space. Additionally, it is useful to note that the use of kernels makes it possible to map the data implicitly into a feature space, i.e., we do not need to represent the feature vectors explicitly. Thus, the number of operations required to compute the inner product by evaluating the kernel function is not necessarily proportional to the number of features.

Remark 1: Let us consider (3) and (7), as examples. Let us call the two vectors $\varphi(\mathbf{s})$ and $\varphi(\mathbf{g})$ as feature vectors. Now, each element of vectors $\mathbf{s}$, $\mathbf{g}$, $\varphi(\mathbf{s})$ and $\varphi(\mathbf{g})$ can be considered as feature in the original (input) or feature (transformed) spaces. In (3), the features are the monomials of degree 1. In contrast, in (7), there are $\frac{(N+d-1)!}{d!(N-1)!}$ distinct features, containing monomials up to and including degree d when transforming the original data (i.e., $\mathbf{s}$ and $\mathbf{g}$) into the feature space (i.e.,

$\varphi(\mathbf{s})$ and $\varphi(\mathbf{g})$) through appropriate polynomial kernel functions of order d. This can be seen in (10)-(12) for the special case of $d = 2$. Thus, it is shown that with the conventional detector in (3), only the inner product of all monomials with degree one is computed, while in the case of the kernelized detector in (7), the inner product of all monomials up to degree d is computed, allowing for richer feature space to be deployed in the detection. Based on this, it can be argued that the kernelized detectors can offer performance improvements compared to the original detectors.

Remark 2: In general, there are two general classes of the real-valued kernel functions, namely i) projective kernel functions such as monomial, polynomial, exponential and sigmoid, and ii) radial kernel functions including Gaussian, Laplacian, multi-quadratic and inverse multi-quadratic [1], [2]. The most common ones are the real polynomial $\mathcal{K}(\mathbf{s},\mathbf{g}) = (\mathbf{s}_1^T\mathbf{g} + c)^d$ and the real Gaussian radial basis function (GRBF) $\mathcal{K}(\mathbf{s},\mathbf{g}) = \exp\left(-||\mathbf{s}-\mathbf{g}||^2/2\sigma_g^2\right)$, where the single parameter $\sigma_g > 0$ is called the width of Gaussian RBF. In our case, however, we cannot utilize any of the radial kernel functions due to the maximization over $\mathbf{s}$ in the considered detectors, which would essentially result in zero detection probability. Among projective kernel functions, the polynomial kernel function is the most common kernel with feasible computational complexity. Besides, the polynomial kernel can be used in complex domains, defined as $\mathcal{K}(\mathbf{s},\mathbf{g}) = (\mathbf{s}^H\mathbf{g} + c)^d$ with $d > 0$, where the monomial kernel is its special case for $c = 0$. We will show that in our detector applications, the polynomial kernel has wider applicability as compared to the monomial one.

B. Kernel-based Detectors

Now, building on the kernel trick, the kernel-based detectors corresponding to (6) and (7) can be represented as

$$T_3(\mathbf{y}) = \frac{\max_{\mathbf{s}\in\mathcal{B}^N}\left\{|\mathcal{K}_1(\mathbf{s},\mathbf{y})|^2\right\}}{\mathcal{K}_2(\mathbf{y},\mathbf{y})} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_3 \tag{13}$$

and

$$\Lambda_3(\mathbf{y}) = \max_{\mathbf{s}\in\mathcal{B}^N}\left\{|\mathcal{K}(\mathbf{s},\mathbf{g})|^2\right\} \underset{H_0}{\overset{H_1}{\gtrless}} \zeta_3 \tag{14}$$

where we consider generally different kernel functions $\mathcal{K}_1$ and $\mathcal{K}_2$ in the numerator and denominator of (13), respectively. In the kernel theory, it is known that different kernel functions can produce nonlinearities in very different ways while can also facilitate different sets of parameters to be tuned or optimized. In general, determining the best kernel functions and optimizing their parameters are open problems, and are also application dependent.

In contrast to existing works on SAD problems with kernels (e.g., [9], [15]), which utilize the Gaussian kernel function, we cannot use this well-known kernel function due to the maximization over $\mathbf{s}$ in the considered detectors, which would essentially result into zero detection probability. Instead, in the following, we adopt polynomial kernels and seek to tune or optimize their parameters to achieve maximum detection probability. To this end, by considering (13) with different polynomial kernels in the numerator and denominator, the polynomial kernelized-LRT detector can now be expressed as

$$T_4(\mathbf{y}) = \frac{\max_{\mathbf{s}\in\mathcal{B}^N}\left\{\left|(\langle\mathbf{s},\mathbf{y}\rangle + c_1)^{d_1}\right|^2\right\}}{(\langle\mathbf{y},\mathbf{y}\rangle + c_2)^{d_2}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_4 \tag{15}$$

where η_4 is the decision threshold that can be chosen such that a desired false alarm probability, denoted as p_{fa}, is obtained. In the above expression, c_1 and c_2 are the bias parameters of the considered polynomial kernels in the numerator and denominator, respectively, while d_1 and d_2 denote the polynomial orders. It is generally noted that the threshold η_4 is a function of N, p_{fa}, d_1, d_2, c_1, and c_2.

Similarly, by considering (14), the polynomial kernelized-LRT can be expressed as

$$\Lambda_4(\mathbf{y}) = \max_{\mathbf{s}\in\mathcal{B}^N}\left\{\left|(\langle\mathbf{s},\mathbf{g}\rangle + \kappa)^{\upsilon}\right|^2\right\} \underset{H_0}{\overset{H_1}{\gtrless}} \zeta_4 \tag{16}$$

where υ denotes the polynomial order and ζ_4 is the corresponding detection threshold. This detector can be re-expressed equivalently as

$$\max_{\mathbf{s}\in\mathcal{B}^N}\left\{\left|\langle\mathbf{s},\mathbf{g}\rangle + \kappa\right|^2\right\} \underset{H_0}{\overset{H_1}{\gtrless}} \zeta_4^{\frac{1}{\upsilon}} \tag{17}$$

This means that the proposed detector in (17) does not depend on the polynomial degree υ. This shows the wider applicability of the polynomial kernel compared to the monomial one.

It can be shown that the proposed detectors in (16) and (17) have the same performance as the original LRT detector introduced in (3). Thus, to introduce a new detector, we express (16) using an alternative form as

$$\Lambda_5(\mathbf{y}) = \max_{\mathbf{s}\in\mathcal{B}^N}\left\{\left|\sum_{k=0}^{\upsilon}\binom{\upsilon}{k}\langle\mathbf{s},\mathbf{g}\rangle^k\kappa^{\upsilon-k}\right|^2\right\} \underset{H_0}{\overset{H_1}{\gtrless}} \zeta_5 \tag{18}$$

Here, we have assumed that ν takes integer values and we have used the binomial expansion of $(\langle \mathbf{s}, \mathbf{g} \rangle + \kappa)^{\upsilon}$ to get (18). Now, performance improvement can be obtained through an appropriate approximation that exploits selected terms of (18). Specifically, by considering the even values of k, we can obtain a detector of the form

$$\Lambda_6(\mathbf{g}) = \max_{\mathbf{s} \in \mathcal{B}^N} \left\{ |\mathcal{L}(\mathbf{g}; \upsilon)|^2 \right\} \underset{H_0}{\overset{H_1}{\gtrless}} \zeta_6 \tag{19}$$

where

$$\mathcal{L}(\mathbf{g}; \upsilon) = \sum_{k=0:even}^{\upsilon} \binom{\upsilon}{k} \langle \mathbf{s}, \mathbf{g} \rangle^k \kappa^{\upsilon - k} \tag{20}$$

In practical example cases $\upsilon = 2$ and $\upsilon = 4$, we get

$$\mathcal{L}(\mathbf{y}; \upsilon) = \begin{cases} \langle \mathbf{s}, \mathbf{g} \rangle^2 + \kappa^2, & \upsilon = 2, \\ \langle \mathbf{s}, \mathbf{g} \rangle^4 + 6 \langle \mathbf{s}, \mathbf{g} \rangle^2 \kappa^2 + \kappa^4, & \upsilon = 4. \end{cases} \tag{21}$$

In order to maximize the detection performance, one can optimize the values of κ. This is discussed and pursued further in Section IV along the numerical results.

The kernelized ED can, in turn, be represented as

$$l_2(\mathbf{y}) = \left(\left\langle \frac{\mathbf{y}}{\sigma_n}, \frac{\mathbf{y}}{\sigma_n} \right\rangle + c_{\mathrm{ED}} \right)^{d_{\mathrm{ED}}} \underset{H_0}{\overset{H_1}{\gtrless}} \varsigma_2(\sigma_n^2). \tag{22}$$

where d_{ED} and c_{ED} denote the order and bias parameters of the corresponding polynomial kernel, and ς_2 is the corresponding detection threshold. This can be equivalently written as

$$l_3(\mathbf{y}) = \left\langle \frac{\mathbf{y}}{\sigma_n}, \frac{\mathbf{y}}{\sigma_n} \right\rangle \underset{H_0}{\overset{H_1}{\gtrless}} \varsigma_3(\sigma_n^2), \tag{23}$$

where $\varsigma_3 = \varsigma_2^{\frac{1}{d_{\mathrm{ED}}}}(\sigma_n^2) - c_{\mathrm{ED}}$. By comparing (4) and (23), we find that the kernelized ED and the original ED have identical performance. This shows that it is not necessarily always possible to improve the detector performance through the kernelization approach.

IV. Threshold-setting Strategy

In general, a detector is said to possess the CFAR property against noise variance uncertainty (NVU), if its false alarm probability is independent of the NVU or the detector statistic is invariant to scale transformation. To investigate the former we need to obtain a closed-form expression for the false alarm probability, while in the latter case it can be examined through the use of the invariance theory. In particular, the CFAR property implies that the detection threshold can be adjusted to attain a desired false alarm probability without the knowledge of the NVU. Before proceeding further, let us introduce two new concepts, i) the level of a test and ii) the size of a test. Assume that empirical false alarm and desired false alarm probabilities are denoted by P_{fa} and p_{fa}, respectively. Now, we say that the devised detector has false alarm probability (P_{fa}) of size p_{fa} provided that $P_{fa} = p_{fa}$. In contrast, it has a false alarm probability (P_{fa}) of level p_{fa} when $P_{fa} \leq p_{fa}$. The first detector is called size-p_{fa} detector (i.e., CFAR detector), while the later one is level-p_{fa} detector [21], [22], [23].

To facilitate efficient operation of the ED under the NVU, we next obtain a closed-form expression for the false alarm probability of ED. This enables us to introduce a new threshold-setting strategy to adjust the threshold of ED under the NVU. To do this, let us define

$$l_1^{(a)}(\mathbf{y}) = \frac{l_1(\mathbf{y})}{\rho} = \frac{\langle \mathbf{y}, \mathbf{y} \rangle}{\sigma_a^2} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\varsigma_1(\sigma_n^2)}{\rho}. \tag{24}$$

where ρ is called the NVU value, defined as

$$\rho = \frac{\sigma_a^2}{\sigma_n^2} > 0 \tag{25}$$

In (25), σ_a^2 represents the actual noise variance, while σ_n^2 is the value of noise variance by which the detection threshold of the ED detector is defined (i.e., the nominal value).

If we assume that $\mathbf{y} \sim \mathcal{CN}(\mathbf{0}, \sigma_a^2 \mathbf{I})$, then it is easy to show that $l_1^{(a)}(\mathbf{y})$ follows a complex central Chi-square distribution with N degrees of freedom under null hypothesis, i.e., $l_1^{(a)}(\mathbf{y}) | H_0 \sim \mathcal{X}_N^2$. Thus, the false alarm probability of the ED can be evaluated as

$$P_{fa}^{(\mathrm{ED})} = Pr\left(l_1^{(a)} \middle| H_0 > \tfrac{\varsigma_1(\sigma_n^2)}{\rho} \right) = Q_{\mathcal{X}_N^2}\left(\tfrac{\varsigma_1(\sigma_n^2)}{\rho}\right), \tag{26}$$

where $Pr(.)$ is probability operator, and $Q_{\mathcal{X}_N^2}(.)$ denotes the right tail of the probability density function (pdf) of $\mathcal{X}_N^2$. Since the function $Q_{\mathcal{X}_N^2}(.)$ is a decreasing function of its argument, the false alarm probability of ED increases by increasing ρ. This is not desirable for signal detection problems. To achieve $P_{fa}^{(\mathrm{ED})} \leq p_{fa}$ in the presence of the NVU, we should thus adjust the ED threshold according to the maximum value of ρ, say ρ_{max}, such that

$$\varsigma_1(\sigma_n^2) = \rho_{max} Q_{\mathcal{X}_N^2}^{-1}(p_{fa}), \tag{27}$$

where $Q_{\mathcal{X}_N^2}^{-1}(.)$ is the inverse of $Q_{\mathcal{X}_N^2}(.)$. The above condition is referred to as the worst case condition of the NVU to ensure that $P_{fa}^{(\mathrm{ED})}$ is always below a p_{fa} level.

As such, the ED with fixed level p_{fa} can be implemented as

$$l_1(\mathbf{y}) = \frac{\langle \mathbf{y}, \mathbf{y} \rangle}{\sigma_n^2} \underset{H_0}{\overset{H_1}{\gtrless}} \rho_{max} Q^{-1}_{\mathcal{X}^2_N}(p_{fa}). \tag{28}$$

This new modified ED is called the fixed-level ED, abbreviated to as FL-ED in the following.

In a similar manner, we can show that $l_1^{(a)}$ follows a complex central Chi-square distribution with N degrees of freedom and with non-centrality parameter $\delta = \frac{|\alpha|^2}{\sigma_a^2}||\mathbf{s}||^2$ under H_1 hypothesis, i.e., $l_1^{(a)}|H_1 \sim \mathcal{X}^2_N(\delta)$. Hence, the detection probability of ED can be computed as

$$\begin{aligned} P_d^{(\mathrm{ED})} \quad (\delta) &= Pr\left(l_1^{(a)} \Big| H_1 > \tfrac{\varsigma_1(\sigma_n^2)}{\rho}\right) \\ &= Q_{\mathcal{X}^2_N(\delta)}\left(\tfrac{\varsigma_1(\sigma_n^2)}{\rho}\right), \end{aligned} \tag{29}$$

where $Q_{\mathcal{X}^2_N(\delta)}(.)$ denotes the right tail of the pdf of $\mathcal{X}^2_N(\delta)$. In the case of $P_{fa}^{(\mathrm{ED})} \leq p_{fa}$, as discussed above, the detection probability can be computed as

$$P_d^{(\mathrm{ED})}(\delta) = Q_{\mathcal{X}^2_N(\delta)}\left(\beta_D Q^{-1}_{\mathcal{X}^2_N}(p_{fa})\right), \tag{30}$$

where

$$\beta_D = \frac{\rho_{max}}{\rho}, \tag{31}$$

is called induced NVU (INUV) factor in the detection probability. From (30), it is seen that the detection probability of ED is a decreasing function of INUV.

For the proposed kernelized detectors, we exploit the principle of invariance to investigate their potential CFAR behavior against NVU since it is not easy to obtain closed-form expressions for their false alarm probabilities. In general, the goal in the invariance theory is to find transformations leaving the considered hypothesis testing problem invariant [24], [25]. In our case, we say a detector with test statistic $\lambda_i(\mathbf{y})$ has CFAR property against the NVU provided that it is invariant under the scale transformation, i.e., $\lambda_i(\sqrt{\rho}\mathbf{y}) = \lambda_i(\mathbf{y})$. This is pursued next.

In the kernelized detector of (15), $T_4(\sqrt{\rho}\mathbf{y}) = T_4(\mathbf{y})$ implies that $c_1 = c_2 = 0$ and $d_1 = d_2$. As such, it is easy to show that the kernelized detector (15) and that of (2) are equivalent, resulting in the same performance. Then, in order to exploit the advantage of kernel theory, we rewrite the decision rule of (15) to obtain

$$T_5(\mathbf{y}) = \frac{\max_{\mathbf{s}\in\mathcal{B}^N}\left\{\left|(\langle \mathbf{s}, \mathbf{y}\rangle + c)^d\right|^2\right\}}{\langle \mathbf{y}, \mathbf{y}\rangle} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_5(\sigma_n^2) \tag{32}$$

Applying the scale transformation to (32) results in $T_5(\sqrt{\rho}\mathbf{y}) = \rho^{d-1}T_5(\mathbf{y})$ when $c = 0$. For $d \neq 1$, this means that this new detector does not possess the CFAR property against the NVU when transforming the receiver noise samples from $\mathbf{y}|H_0 \sim \mathcal{CN}(\mathbf{0}, \sigma_n^2\mathbf{I})$ to $\sqrt{\rho}\mathbf{y}|H_0 \sim \mathcal{CN}(\mathbf{0}, \sigma_a^2\mathbf{I})$. As a result, the detection threshold of test statistic $T_5(\mathbf{y})$ depends on the noise variance, denoted by $\eta_5(\sigma_n^2)$. Under NVU of ρ, we can reformulate (32), expressed as

$$T_6(\mathbf{y}) = \frac{\max_{\mathbf{s}\in\mathcal{B}^N}\left\{\left|\langle \mathbf{s}, \mathbf{y}\rangle^d\right|^2\right\}}{\langle \mathbf{y}, \mathbf{y}\rangle} \underset{H_0}{\overset{H_1}{\gtrless}} \rho^{d-1}\eta_5(\sigma_n^2) \tag{33}$$

In general, in practical situations, the value of ρ is not known a prior; however, we can consider a maximum value for it to set the new detection threshold $\rho^{d-1}\eta_5(\sigma_n^2)$. Additionally, under UVN, we should limit the false alarm probability to an upper bound by using a threshold- setting under the worst-case condition. By doing so, we obtain

$$T_7(\mathbf{y}) = \frac{\max_{\mathbf{s}\in\mathcal{B}^N}\left\{\left|\langle \mathbf{s}, \mathbf{y}\rangle^d\right|^2\right\}}{\langle \mathbf{y}, \mathbf{y}\rangle} \underset{H_0}{\overset{H_1}{\gtrless}} \rho_{max}^{d-1}\eta_5(\sigma_n^2) \tag{34}$$

In general, it should be noted that the new detector does not have the CFAR property. However, since the denominator of (34) represents a noise variance estimate under the hypothesis H_0, we refer to it as a semi-CFAR (S-CFAR) detector. In the sequel, this detector is called polynomial kernel and S-CFAR LRT-based detector, abbreviated as PK-SC-LRT detector. It should be noted that the proposed PK-SC-LRT has CFAR property when $d = 1$, which is equivalent in performance with the classical LRT-based one. This means that the previously proposed LRT detector can be considered as a special case of the proposed one.

In contrast to (34), the proposed test statistics $\Lambda_i(\mathbf{y})$ for $i = 2, ..., 6$ have CFAR property against the noise variance. Thus, they are invariant to any scale transformation, i.e.,

$$\Lambda_i(\sqrt{\rho}\mathbf{y}) = \Lambda_i(\mathbf{y}) \tag{35}$$

for $i = 2, ..., 6$. In the following, the detector in (17) is called the polynomial kernel and CFAR LRT-based detector, namely PK-C-LRT. Similarly, the proposed detector presented in (19) is referred to as the approximated polynomial kernel and CFAR LRT-based detector, namely APK-C-LRT detector.

V. Simplifying the Proposed Kernel- Based LRT Detectors

The proposed detectors in the previous subsection basically need an exhaustive search due to the maximiza-

tion over $\mathbf{s}$. Thus, they are computationally expensive and thus impractical in real communication systems, especially for larger values of N. Thus, to assist practical implementations, we seek next simplified implementation forms of the proposed kernelized detectors. To do this, we firstly introduce in detail the simplified implementation of the detector in (3), while then apply similar approach with the selected actual detectors. Let us denote the estimate of the nth entry of vector $\mathbf{s}$ as $\hat{s_n}$. In [19], firstly, the entries of $\mathbf{s} = [s_1, ..., s_N]^T$ are approximately estimated, expressed as

$$\hat{s}_n = \text{argmax}_{s_n} |r_{n-1} + s_n^* g_n| \tag{36}$$

where $r_n \triangleq r_{n-1} + \hat{s}_n^* g_n$ for $n = 2, ..., N$, and $r_1 = \hat{s}_1^* g_1$ with $\hat{s}_1$ being chosen arbitrarily from $\mathcal{B} = \{-1, 1\}$. This has been obtained based on the fact that the inner product $\langle \mathbf{s}, \mathbf{g} \rangle$ in (3) can be obtained recursively as

$$\langle \mathbf{s}, \mathbf{g} \rangle = \langle \mathbf{s}_{(-N)}, \mathbf{g}_{(-N)} \rangle + s_N^* g_N \tag{37}$$

with $\mathbf{s}_{(-N)}$ and $\mathbf{g}_{(-N)}$ denoting the subvectors that are formed from vectors $\mathbf{s}$ and $\mathbf{g}$ by removing the N-th entries, respectively. Then, the detector in (3) can be approximately implemented as

$$\Lambda_1(\mathbf{y}) \approx |r_N|^2 \underset{H_0}{\overset{H_1}{\gtrless}} \zeta_1' \tag{38}$$

This, in turn, allows us to implement the LRT statistic with computational complexity (CC) of the order of $\mathcal{O}(N|\mathcal{B}|)$ instead of an exhaustive search of the order of $\mathcal{O}(|\mathcal{B}|^{N-1})$ associated to the original LRT detector. This simplified LRT is referred to as SLRT [19]. It can be seen that the block estimation of vector $\mathbf{s}$ in the LRT detector is replaced with a corresponding symbol-by-symbol estimation in the SLRT detector. However, it can be shown that the detection performance of the SLRT is comparable with that of the LRT detector with much lower complexity.

Through the above approach, we can pursue simplified kernel-based detectors. Based on (37) and following similar steps as in the SLRT derivation, the proposed PK-SC-LRT, PK-C-LRT, and APK-C-LRT detectors can be simplified to obtain new detectors called PK-SC-SLRT, PK-C-SLRT and APK-C-SLRT, where the resulting processing steps are summarized in Algorithms 1-3, respectively. The proposed kernel-based detectors are listed in Table I, where their properties and kernel parameters are also summarized for readers' convenience. It is easy to check that the proposed kernelized detectors have the same computational complexity orders as compared with their original counterparts, while the original detectors can basically be considered as special cases of the proposed kernelized ones. For instance, the proposed PK-C-SLRT detector only needs more additions than the C-SLRT detector. Similarly, the proposed PK-SC-SLRT detector needs at most $d-1$ more multiplications and one more addition, when compared with its original counterpart (i.e., the SC-SLRT detector).

Algorithm 1 Proposed PK-SC-SLRT detector, simplified version of (34).

Input: σ_n^2, ρ_{max}, $\eta_5(\sigma_n^2)$, $\mathbf{s}$, $\mathbf{y}$, c, d.
Set $\eta_7' = \rho_{max}^{d-1} \eta_5(\sigma^2)$.
Select $\hat{s_1} \in \mathcal{B}$ **arbitrarily and set** $r_1 = \hat{s}_1^* y_1$,
for $n = 2 : 1 : N$; **do**
  Estimate $\hat{s}_n$ **using**
$$\hat{s}_n = \textbf{argmax}_{s_n} |r_{n-1} + s_n^* y_n|$$
$$r_n \triangleq r_{n-1} + \hat{s}_n^* y_n$$
end for
Compute the PK-SC-SLRT statistic, given by
$$T_7'(\mathbf{y}) = \frac{|(r_N + c)^d|^2}{\langle \mathbf{y}, \mathbf{y} \rangle}$$
Compare $T_7(\mathbf{y})$ **with the detection threshold** η_7',
i.e., $T_7'(\mathbf{y}) \underset{H_0}{\overset{H_1}{\gtrless}} \eta_7'$.

Algorithm 2 Proposed PK-C-SLRT detector, simplified version of (17).

Input: ζ_4', $\mathbf{s}$, $\mathbf{y}$, κ.
Construct $\mathbf{g} = \frac{\mathbf{y}}{\|\mathbf{y}\|}$, where $\mathbf{g} = [g_1, ..., g_N]^T$.
Select $\hat{s_1} \in \mathcal{B}$ **arbitrarily and set** $t_1 = \hat{s}_1^* g_1$,
for $n = 2 : 1 : N$; **do**
  Estimate $\hat{s}_n$ **using**
$$\hat{s}_n = \textbf{argmax}_{s_n} |t_{n-1} + s_n^* g_n|$$
$$t_n \triangleq t_{n-1} + \hat{s}_n^* g_n$$
end for
Compute the PK-C-SLRT statistic, given by
$$\Lambda_4'(\mathbf{g}) = |t_N + \kappa|^2$$
Compare $\Lambda_4'(\mathbf{g})$ **with the detection threshold** ζ_4',
i.e., $\Lambda_4'(\mathbf{g}) \underset{H_0}{\overset{H_1}{\gtrless}} \zeta_4'$.

VI. Simulation results

In this section, numerical simulation results are provided to examine the detection performance of the proposed PK-SC-LRT, PK-SC-SLRT, PK-C-LRT, PK-C-SLRT, APK-C-LRT, APK-C-SLRT detectors and compare them to their classical counterparts (i.e., LRT and SLRT). To do this, unless otherwise stated, the number of received samples is set 6 (i.e., $N = 6$) and the desired false alarm probability of $p_{fa} = 10^{-3}$ is considered.

TABLE I
PROPOSED KERNEL-BASED DETECTORS AND THEIR PROPERTIES.

Proposed Detectors	Properties	Test statistic	Kernel parameters
PK-SC-LRT	Semi-CFAR against NVU	$T_7(\mathbf{y})$	$\mathcal{S}=\{d\}$
PK-C-LRT	CFAR against NVU	$\Lambda_4(\mathbf{y})$ or $\Lambda_4(\mathbf{y})^{\frac{1}{4}}$ of (17)	$\mathcal{S}=\{\kappa\}$
APK-C-LRT	CFAR against NVU	$\Lambda_6(\mathbf{y})$	$\mathcal{S}=\{\kappa\}$
PK-SC-SLRT	Semi-CFAR against NVU, simplified version of PK-SC-LRT	$T_7'(\mathbf{y})$	$\mathcal{S}=\{c,d\}$
PK-C-SLRT	CFAR again st NVU, simplified version of PK-C-LRT	$\Lambda_4'(\mathbf{g})$	$\mathcal{S}=\{\kappa\}$
APK-C-SLRT	CFAR against NVU, simplified version of APK-C-LRT	$\Lambda_6'(\mathbf{g};\upsilon)$	$\mathcal{S}=\{\kappa\}$

Algorithm 3 Proposed APK-C-SLRT detector, simplified version of (19).

1: **Input:** υ, $\zeta_6'(\upsilon)$, $\mathbf{s}$, $\mathbf{y}$, κ.
2: **Construct** $\mathbf{g}=\frac{\mathbf{y}}{\|\mathbf{y}\|}$, where $\mathbf{g}=[g_1,...,g_N]^T$.
3: **Select** $\hat{s_1}\in\mathcal{B}$ **arbitrarily and set** $p_1=\hat{s}_1^* g_1$,
4: **for** $n=2:1:N$; **do**
5: **Estimate** $\hat{s}_n$ **using**

$$\hat{s}_n=\textbf{argmax}_{s_n}|p_{n-1}+s_n^* g_n|$$
$$p_n\triangleq p_{n-1}+\hat{s}_n^* g_n$$

6: **end for**
7: **Compute the APK-C-SLRT statistic, given by**

$$\Lambda_6'(\mathbf{g};\upsilon)=\begin{cases}|p_1+\kappa^2|^2, & \upsilon=2,\\ |p_1^4+6p_1^2\kappa^2+\kappa^4|^2, & \upsilon=4.\end{cases}$$

8: **Compare** $\Lambda_6'(\mathbf{g})$ **with the detection threshold** $\zeta_6'(\upsilon)$, **i.e.,** $\Lambda_6'(\mathbf{g};\upsilon)\underset{H_0}{\overset{H_1}{\gtrless}}\zeta_6'(\upsilon)$.

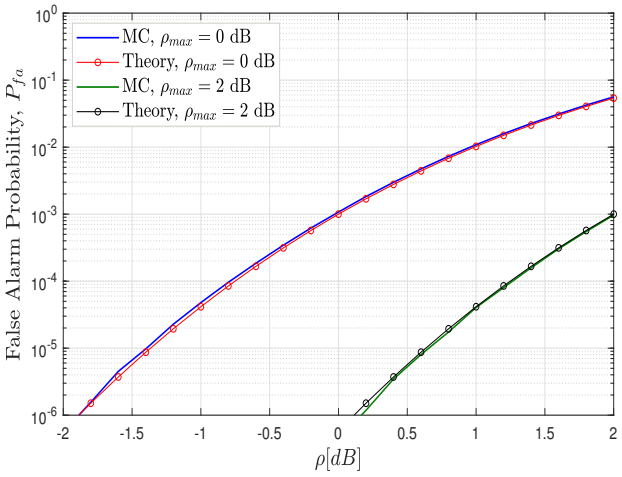

Fig. 1. Probability of false alarm versus noise variance uncertainty ratio ρ for the FL-ED detector when the desired false alarm probability is set equal to $p_{fa}=10^{-3}$.

In the following, for simplicity, we set $\sigma_n^2=1$. Thus $\rho=\sigma_a^2$ and we generate noise samples according to $\mathbf{w}\sim\mathcal{CN}(\mathbf{0},\sigma_a^2\mathbf{I})$. The empirical false alarm probability (i.e., P_{fa}) and detection probability (i.e., P_d) of the above detectors are determined by 10^8 and 10^4 Monte-Carlo (MC) simulation runs, respectively.

To evaluate the detection performance of the proposed detectors, we define input signal-to-noise ratio (SNR) as $\text{SNR}=\frac{|\alpha|^2}{\sigma_a^2}$, where $\text{SNR}[\text{dB}]=10\log_{10}(\text{SNR})$, while the channel gain α is scaled to achieve the desired SNR over all MC simulations. In our case, the elements of the vector $\mathbf{s}$ are taken from the BPSK alphabet (i.e., $\mathcal{B}=\{-1,1\}$), with different random realizations for different MC runs.

A. Verification of FL-ED Analytical Results

In Fig. 1, the probability of false alarm as a function of ρ is plotted for different values of ρ_{max}. The lines with markers "$\circ$" denote the results using the theoretical expression in (26), while the lines without markers denote the MC results. It can be seen that the theoretical results are consistent with the MC ones. Additionally, we observe that the FL-ED detector attains the desired level of 10^{-3} when its threshold is set according to the worse case condition of $\rho_{max}=2$ dB, i.e., we achieve $P_{fa}\leq 10^{-3}$ for $\rho\leq 2$ dB, while it suffers from excessive false alarm probability with $\rho_{max}=0$ dB for $\rho>0$ dB.

Fig. 2 presents the detection probability as a function of SNR for different values of β_{D}. The symbols "$\circ$" denote the results computed with the theoretical expression in (30), while the lines without markers represent again the MC results. We can observe that the theoretical results match the MC ones very well. Additionally, we observe that the detection performance of the proposed FL-ED is degraded as β_{D} increases.

B. False Alarm Regulation Evaluations of Kernelized Detectors

In Fig. 3, the false alarm regulation of the ordinary and kernelized detectors as a function of ρ is examined for two cases of $\rho_{max}=0$ dB and $\rho_{max}=2$ dB. It is seen that the LRT, SLRT, PK-C-LRT and PK-C-SLRT detectors have CFAR property, while the pro-

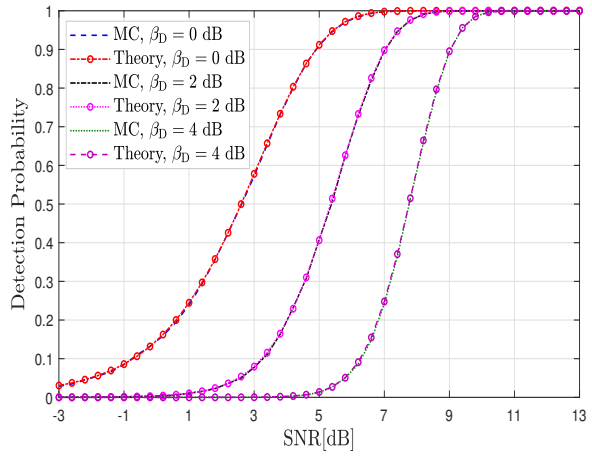

Fig. 2. Detection probability of the proposed FL-ED as a function of SNR for different values of β_D when $\rho_{max} = 2$ dB.

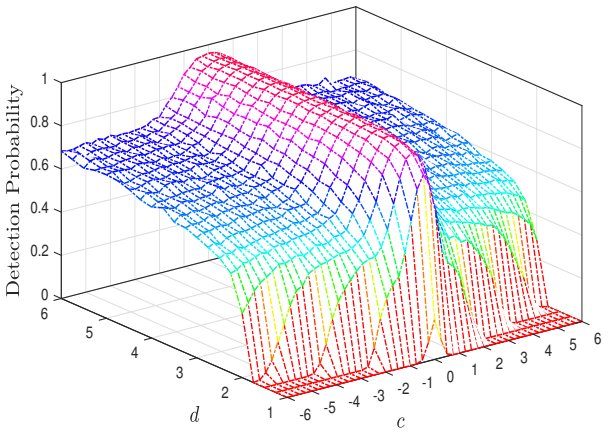

Fig. 4. Detection probability as a function of polynomial kernel's parameters of the decision rule of (32) for $p_{fa} = 10^{-3}$, $\beta_D = 0$ dB with $\rho_{max} = 2$ dB and SNR $= 5$ dB.

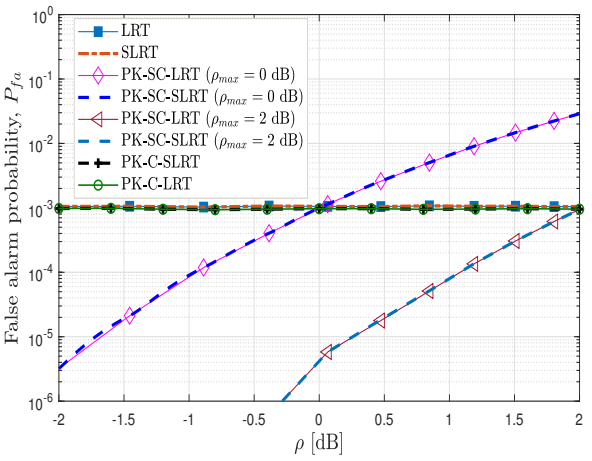

Fig. 3. Probability of false alarm versus noise variance uncertainty ratio ρ for the proposed S-CFAR detectors when the desired false alarm probability is set equal to $p_{fa} = 10^{-3}$.

posed PK-SC-LRT and PK-SC-SLRT detectors suffer from excessive false alarm probabilities in the case of $\rho_{max} = 0$ dB. As can be seen form Fig. 3, the proposed PK-SC-LRT and PK-SC-SLRT detectors are not CFAR detectors, i.e., $P_{fa} \neq 10^{-3}$ when their thresholds are adjusted according to $\rho_{max} = 0$ dB, while they have false alarm probabilities of level p_{fa} when we set their thresholds by $\rho_{max} = 2$ dB, i.e., $P_{fa} \leq 10^{-3}$. In practical situations, it is required for S-CFAR detectors, such as the proposed PK-SC-LRT and PK-SC-SLRT detectors, to guarantee that the obtained false alarm probability is always below a required level p_{fa}. In the following, we consider a worse case condition for NVU, specifically $\rho_{max} = 2$ dB, and set the thresholds accordingly to achieve $P_{fa} \leq p_{fa}$.

C. Detection Performance Evaluations of Kernelized Detectors

For the proposed kernelized detectors, firstly, we need to select suitable parameters for the exploited polynomial kernel functions. The best way of doing this would be to maximize the detection probability under the constraint that the empirical false alarm probability is less than the desired false alarm probability p_{fa}, i.e.,

$$\max_{P_{fa}(\mathcal{S}) \leq p_{fa}} P_d(\mathcal{S}) \tag{39}$$

where $\mathcal{S}$ contains the kernel parameters of different proposed detectors. For instance, we have $\mathcal{S} = \{c, d\}$, and $\mathcal{S} = \{\kappa\}$ for the PK-SC-LRT (SLRT) and APK-C-LRT(SLRT), respectively. Solving such non-linear optimization problem[1] is far from trivial, especially since the closed-form expressions for false alarm probability and detection probability are generally unknown. Thus, in this work, we consider and adopt MC simulation based approach to obtain proper parameters of the polynomial kernel functions, where we fix the false alarm probability and then seek to adjust the parameters of the polynomial kernel functions to maximize the corresponding detection probability.

For the finalized PK-SC-LRT detector with kernel parameters $\mathcal{S} = \{c, d\}$, we follow the above procedure and plot the detection probability of the proposed decision rule in (32) as a function of c and d while setting $p_{fa} = 10^{-3}$ and for SNR $= 5$ dB. The corresponding results are depicted in Fig. 4 for the case of $\beta_D = 0$ dB

[1]Generally, the P_d and P_{fa} are nonlinear functions with respect to kernel parameters; thus we encounter a nonlinear optimization problem.

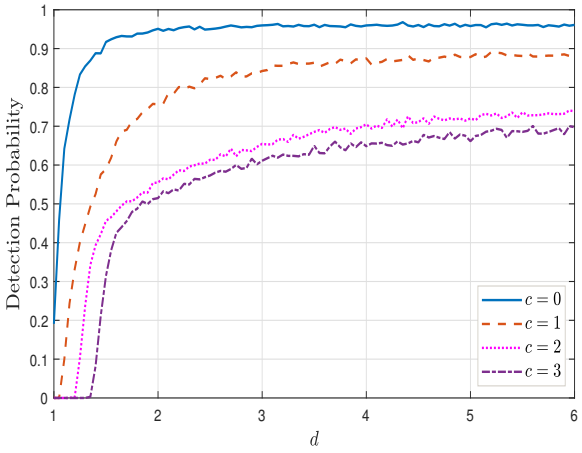

Fig. 5. Detection probability as a function of d for different values of c when the decision rule of (32) is used for $p_{fa} = 10^{-3}$, $\beta_D = 0$ dB with $\rho_{max} = 2$ dB and SNR $= 5$ dB.

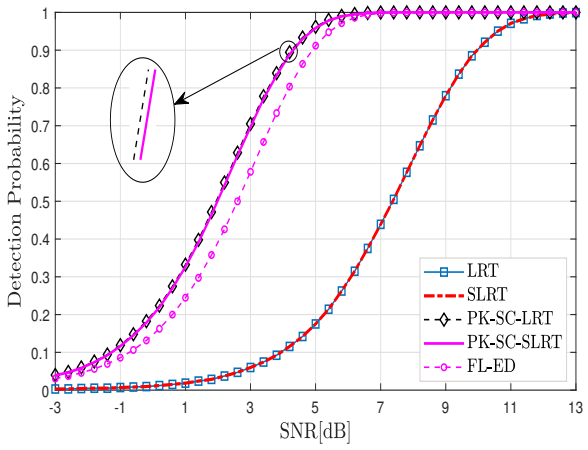

Fig. 6. Detection probability as a function of SNR for the proposed S-CFAR detectors when we set $d = 5$, $c = 0$, $p_{fa} = 10^{-3}$ and $\beta_D = 0$ dB with $\rho_{max} = 2$ dB. $P_{fa} = 10^{-3}$ is achieved with the considered detectors.

with $\rho_{max} = 2$ dB. In such a case, it can be observed from Fig. 4 that the values of d greater than 2 and $c = 0$ result in the best detection performance. This can also be concluded from Fig. 5, where the detection probability is plotted versus d for different values of c. Similar results can be obtained for the PK-SC-SLRT detector. In the following, we choose $c = 0$ and $d = 3$ to compare the detection performance of the proposed PK-SC-LRT and PK-SC-SLRT with that of the classical LRT and SLRT detectors proposed in [19].

In Fig. 6, the detection probability of the semi-CFAR detectors is compared when $\beta_D = 0$ dB with $\rho_{max} = 2$ dB, which corresponds to the case of $P_{fa} = p_{fa} = 10^{-3}$. In this case without any NVU, it is seen that the proposed detectors significantly outperform their classical counterparts, while all the proposed detectors also perform better than FL-ED. Here, the proposed PK-SC-SLRT detector has the same performance as that of PK-SC-LRT with lower computational complexity. This trend can be seen between the previously proposed LRT and SLRT detectors. We see that the proposed PK-SC-SLRT detector requires about 4.3dB to attain $p_d = 0.9$ while it is 10 dB in the SLRT detector, resulting in an SNR gain of about 5.7 dB.

In Fig. 7, the detection performance of the S-CFAR detectors are compared in the case of $\beta_D = 2$ dB with $\rho_{max} = 2$ dB (i.e., $\rho = 0$ dB)[2]. Comparing the detection performances of the proposed detectors in Fig. 6 with that in Fig. 7 indicates that the detection performances of the proposed detectors have been degraded; however, they have significantly better performances than the LRT, SLRT and FL-ED detectors. This degradation in performance is the price paid to attain detectors with level p_{fa}.

In the worse case condition of $\beta_D = 4$ dB with $\rho_{max} = 2$ dB (i.e., $\rho = -2$ dB), we obtain the results depicted in Fig. 8. In this case, the proposed S-CFAR detectors have the same performance with an SNR gain of about 2.4 dB compared with their counterparts. The proposed detectors can also be observed to significantly outperform the FL-ED. Interestingly, however, the FL-ED outperforms the classical LRT and SLRT detectors, as seen through Figs. 6- 8.

For the proposed CFAR detectors the above simulation results are repeated. In Fig. 9, we are interested in finding κ in the proposed PK-C-LRT and APK-C-LRT detectors which maximize their detection probabilities. It is seen that the value of $\kappa = 0$ leads to the same performances of the proposed PK-C-LRT detector with that of the conventional LRT detector, while the values of κ greater than 3 results in the best performance for the APK-C-LRT detector for different values of υ. In the following, we use $\kappa = 3$ to evaluate the performance of the proposed APK-C-LRT detector.

In Fig. 10, the detection performance of the four proposed CFAR detectors are compared with that of the LRT and SLRT detectors. It can be observed that the proposed APK-C-LRT and APK-C-SLRT schemes significantly outperform the proposed PK-C-LRT and PK-C-SLRT detectors, respectively, while the two later

[2]As mentioned before, under noise variance uncertainty, we should limit false alarm probability to an upper bound by using a threshold-setting under the worst-case condition. Therefore, detection performance should be examined with this type of threshold adjusting to satisfy $P_{fa} \leq p_{fa}$.

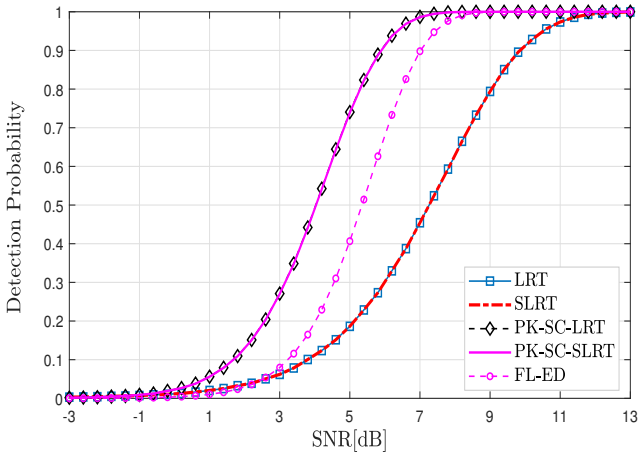

Fig. 7. Detection probability as a function of SNR for the proposed S-CFAR detectors when we set $d = 5$, $c = 0$, $p_{fa} = 10^{-3}$ and $\beta_D = 2$ dB with $\rho_{max} = 2$ dB (i.e., $\rho = 0$ dB). We achieve $P_{fa} < 10^{-3}$ for the proposed S-CFAR detectors while we have $P_{fa} = 10^{-3}$ for the LRT and SLRT detectors.

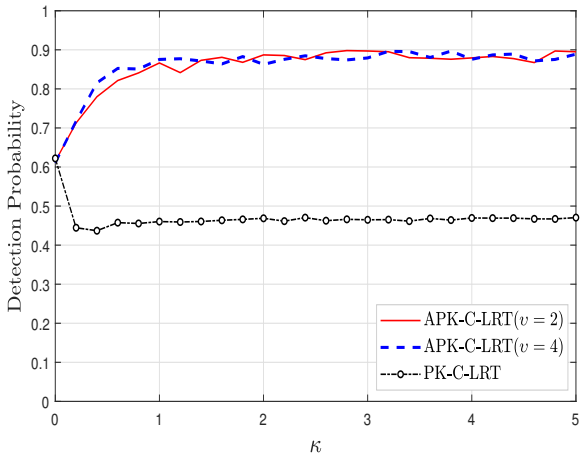

Fig. 9. Detection probability as a function of κ for the proposed CFAR detectors for different values of υ when we set $p_{fa} = 10^{-3}$ and SNR $= 8.2$ dB.

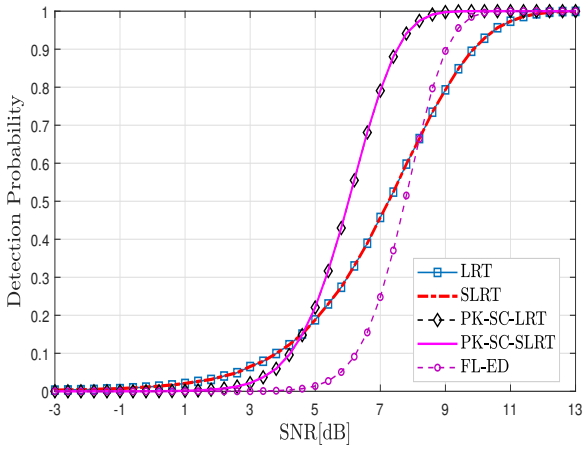

Fig. 8. Detection probability as a function of SNR for the proposed S-CFAR detectors when we set $d = 5$, $c = 0$, $p_{fa} = 10^{-3}$ and $\beta_D = 4$ dB with $\rho_{max} = 2$ dB (i.e., $\rho = -2$ dB). We achieve $P_{fa} < 10^{-3}$ for the proposed S-CFAR detectors while we have $P_{fa} = 10^{-3}$ for the LRT and SLRT detectors.

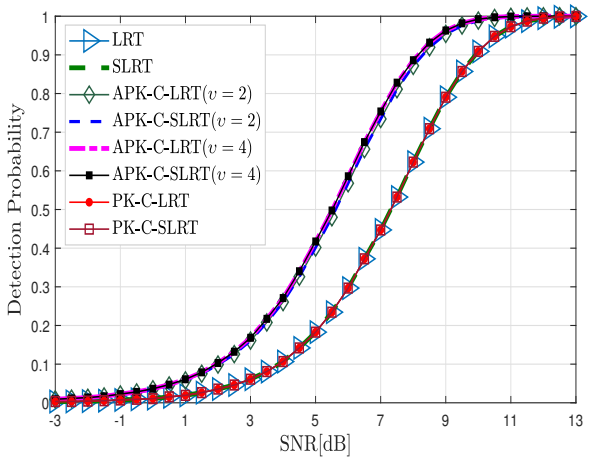

Fig. 10. Detection probability as a function of SNR for the proposed CFAR detectors when we set $\kappa = 3$ and $p_{fa} = 10^{-3}$.

detectors are very close in performance to the LRT and SLRT detectors, respectively.

Finally, the SNR gain as a function of ρ for the proposed PK-SC-SLRT and APK-C-SLRT as compared to the SLRT detector is plotted in Fig. 11, where we achieve detection probability of 0.9 and $P_{fa} \leq 10^{-3}$. Fig. 11 shows how SNR gain varies with ρ for the proposed PK-SC-SLRT, where the detection threshold of the proposed PK-SC-SLRT is adjusted according to $\rho_{max} = 2$ dB. For the worse case condition, the SNR gain of the PK-SC-SLRT detector is about 2.4 dB and it increases with ρ, where the case of $\rho = 2$ dB corresponds to no NVU, resulting in the SNR gain of 5.7 dB. As can be seen from Fig. 11, for the proposed APK-C-SLRT, we obtain a fixed SNR gain of 1.8 dB, regardless of the values of ρ due to its CFAR property. Interestingly, the performance of the proposed S-CFAR methods is significantly better than those of the proposed CFAR detectors even in the considered worse case condition.

It is finally noted that the above improvements in the detection performance through the proposed kernel-based detectors are obtained due to exploiting the richer features of the transformed data (i.e., monomial features with degree up to d or ν) as compared to that of the original data represented by monomials with degree one. This is addressed in further details in Remark 1. Thus, while the original LRT detectors are efficient in finding linear relations, the kernel-based detectors

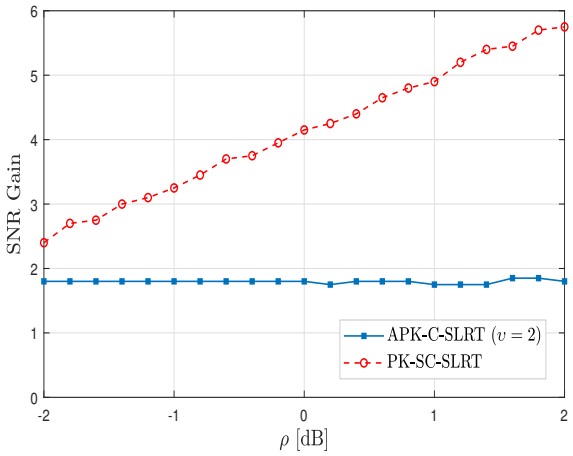

Fig. 11. Signal-to-noise ratio gain of the proposed PK-SC-SLRT detector (with $d = 5$ and $c = 0$) and APK-C-SLRT detector (with $\kappa = 3$ and $\upsilon = 2$) as compared to the SLRT detector to achieve $P_{fa} \leq 10^{-3}$ and detection probability of 0.9.

provide generally a richer space to exploit both linear and nonlinear relations of the received data.

VII. Conclusion

In this article, two general classes of kernelized LRT detector for BPSK signal detection were developed and presented. The first class consists of two new detectors, shown to possess the semi-CFAR property against the noise variance uncertainty, while the proposed detectors of the second class where shown to be of CFAR nature. In deriving the detectors, the inner product of the classical LRT test was replaced with selected polynomial kernel functions, which enables us to exploit some richer features of the transformed space, subsequently improving the detection performance of the proposed detectors. In each detector class, also simplified implementation structures and processing methods with reduced computational complexity but comparable detection performance were described. Although the proposed detectors in the first class are not CFAR detectors with respect to the noise variance uncertainty, we have introduced a new threshold-setting strategy to obtain selected new detectors with a fixed level of p_{fa}. In this case, in the presence of the noise variance uncertainty, we assumed that the maximum variation of the noise variance uncertainty ratio is given or fixed to be able to achieve an empirical false alarm rate below a predetermined level. While building on the polynomial kernel functions in our problem, we optimized the corresponding parameters of the functions to achieve a maximum detection probability under a fixed level of false alarm probability. Our simulation results illustrate the superior performance of the proposed BPSK signal detectors over classical LRT-based detectors. Additionally, the performance of the proposed semi-CFAR methods was shown to be significantly better than those of the proposed CFAR detectors.

Finally, in the presence of NVU, we devised a new fixed-level energy detector (FL-ED), while also derived the corresponding closed-form expressions for both the false alarm probability and the detection probability. It was shown through numerical results that the proposed FL-ED also outperforms the classical LRT and SLRT detectors.

References

[1] B. Schölkopf, A. J. Smola, F. Bach *et al.*, *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT press, 2002.

[2] G. Ding, Q. Wu, Y.-D. Yao, J. Wang, and Y. Chen, "Kernel-based learning for statistical signal processing in cognitive radio networks: Theoretical foundations, example applications, and future directions," *IEEE Signal Processing Magazine*, vol. 30, no. 4, pp. 126–136, June 2013.

[3] D. Marin, M. Tang, I. B. Ayed, and Y. Boykov, "Kernel clustering: density biases and solutions," *IEEE transactions on pattern analysis and machine intelligence*, vol. 41, no. 1, pp. 136–147, December 2017.

[4] M. Kafai and K. Eshghi, "Croification: accurate kernel classification with the efficiency of sparse linear svm," *IEEE transactions on pattern analysis and machine intelligence*, vol. 41, no. 1, pp. 34–48, December 2017.

[5] D. You, O. C. Hamsici, and A. M. Martinez, "Kernel optimization in discriminant analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 3, pp. 631–638, September 2010.

[6] V. Garro and A. Giachetti, "Scale space graph representation and kernel matching for non rigid and textured 3d shape retrieval," *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 6, pp. 1258–1271, September 2015.

[7] H. Wang, D. Mirota, and G. D. Hager, "A generalized kernel consensus-based robust estimator," *IEEE transactions on pattern analysis and machine intelligence*, vol. 32, no. 1, pp. 178–184, July 2009.

[8] A. Golts and M. Elad, "Linearized kernel dictionary learning," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 4, pp. 726–739, April 2016.

[9] M. Mansouri, R. Baklouti, M. F. Harkat, M. Nounou, H. Nounou, and A. B. Hamida, "Kernel generalized likelihood ratio test for fault detection of biological systems," *IEEE transactions on nanobioscience*, vol. 17, no. 4, pp. 498–506, October 2018.

[10] S. Kay, "Statistical decision theory i, and" deterministic signals," *Fundamentals of Statistical Signal Processing, Detection Theory*, 1998.

[11] A. Zaimbashi, "Forward m-ary hypothesis testing based detection approach for passive radar," *IEEE Transactions on Signal Processing*, vol. 65, no. 10, pp. 2659–2671, May 2017.

[12] ——, "Multiband fm-based passive bistatic radar: target range resolution improvement," *IET Radar, Sonar & Navigation*, vol. 10, no. 1, pp. 174–185, January 2016.

[13] Z. Solatzadeh and A. Zaimbashi, "Accelerating target detection in passive radar sensors: Delay-doppler-acceleration estimation," *IEEE Sensors Journal*, vol. 18, no. 13, pp. 5445–5454, May 2018.

[14] J. M. Anderson, "A generalized likelihood ratio test for detecting land mines using multispectral images," *IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 3, pp. 547–551, July 2008.

[15] H. Kwon and N. M. Nasrabadi, "Kernel matched subspace detectors for hyperspectral target detection," *IEEE transactions on pattern analysis and machine intelligence*, vol. 28, no. 2, pp. 178–194, December 2005.

[16] L. Li, S. Hou, and A. L. Anderson, "Kernelized generalized likelihood ratio test for spectrum sensing in cognitive radio," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 8, pp. 6761–6773, April 2018.

[17] S. Hou and R. C. Qiu, "Kernel feature template matching for spectrum sensing," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 5, pp. 2258–2271, November 2014.

[18] A. Margoosian, J. Abouei, and K. N. Plataniotis, "An accurate kernelized energy detection in gaussian and non-gaussian/impulsive noises," *IEEE Transactions on Signal Processing*, vol. 63, no. 21, pp. 5621–5636, July 2015.

[19] A. A. Tadaion, M. Derakhtian, S. Gazor, M. M. Nayebi, and M. R. Aref, "Signal activity detection of phase-shift keying signals," *IEEE Transactions on Communications*, vol. 54, no. 8, pp. 1439–1445, August 2006.

[20] S. Gazor, M. Derakhtian, and A. A. Tadaion, "Computationally efficient maximum likelihood sequence estimation and activity detection for m-psk signals in unknown flat fading channels," *IEEE Signal Processing Letters*, vol. 17, no. 10, pp. 871–874, August 2010.

[21] F. Jalali and A. Zaimbashi, "Cognitive radio spectrum sensing under joint tx/rx i/q imbalance and uncalibrated receiver," *IEEE Systems Journal*, vol. 14, no. 1, pp. 105–112, March 2020.

[22] A. Mehrabian and A. Zaimbashi, "Spectrum sensing in simo cognitive radios under primary user transmitter iq imbalance," *IEEE Systems Journal*, vol. 13, no. 2, pp. 1210–1218, August 2019.

[23] A. Zaimbashi, "Multistatic passive radar sensing algorithms with calibrated receivers," *IEEE Sensors Journal*, vol. 20, pp. 7878 – 7885, March 2020.

[24] A. Zaimbashi and M. Valkama, "Impropriety-based multi-antenna spectrum sensing with i/q imbalanced radios," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 9, pp. 8693–8706, July 2019.

[25] A. Zaimbashi, "Invariant subspace detector in distributed multiple-input multiple output radar: geometry gain helps improving moving target detection," *IET Radar, Sonar Navigation*, vol. 10, no. 5, pp. 923–934, June 2016.