One-Bit Target Detection in Colocated MIMO Radar With Colored Background Noise

Yu-Hang Xiao, *Member, IEEE*, David Ramírez, *Senior Member, IEEE*, Lei Huang, *Senior Member, IEEE*, Xiao Peng Li, *Member, IEEE*, and Hing Cheung So, *Fellow, IEEE*

***Abstract*—One-bit sampling has emerged as a promising technique in multiple-input multiple-output (MIMO) radar systems due to its ability to significantly reduce data volume, hardware complexity, and power consumption. Nevertheless, current detection methods have not adequately addressed the impact of colored noise, which is frequently encountered in real scenarios. In this paper, we present a novel detection method that accounts for colored noise in MIMO radar systems. Specifically, we derive Rao's test by computing the derivative of the likelihood function with respect to the target reflectivity parameter and the Fisher information matrix, resulting in a detector that takes the form of a weighted matched filter. To ensure constant false alarm rate (CFAR), we also consider noise covariance uncertainty and examine its effect on the probability of false alarm. The detection probability is also studied analytically. Simulation results demonstrate that the proposed detector provides considerable performance gains in the presence of colored noise.**

***Index Terms*—One-bit analog-to-digital converter (ADC), multiple-input multiple-output (MIMO) radar, Rao's test, target detection.**

Received 26 April 2024; revised 20 September 2024; accepted 16 October 2024. Date of publication 4 November 2024; date of current version 25 November 2024. The work of Yu-Hang Xiao was supported by the National Natural Science Foundation of China under Grant 62201359. The work of David Ramírez was supported in part by the MICIU/AEI/10.13039/501100011033/FEDER, UE, under Grant PID2021-123182OB-I00 (EPiCENTER) and Grant CPP2022-009537 (PreCARE), in part by the Office of Naval Research (ONR) Global under Grant N62909-23-1-2002, in part by the Comunidad de Madrid under Grant IND2023/TIC-27508 (IRIS), and in part by the Spanish Ministry of Economic Affairs and Digital Transformation and the European Union-NextGenerationEU through the UNICO 5G I+D SORUS project. The work of Lei Huang was supported in part by the National Science Fund for Distinguished Young Scholars under Grant 61925108, in part by the Key Project of International Cooperation and Exchanges of the National Natural Science Foundation of China under Grant 62220106009, in part by the project of Shenzhen Peacock Plan Teams under Grant KQTD20210811090051046 and Shenzhen University 2035 Program for Excellent Research. The associate editor coordinating the review of this article and approving it for publication was Prof. Wei Yi. *(Corresponding author: Lei Huang.)*

Yu-Hang Xiao, Lei Huang and Xiao Peng Li are with the State Key Laboratory of Radio Frequency Heterogeneous Integration, Shenzhen University, Shenzhen 518060, China (e-mail: yuhangxiao@szu.edu.cn; lhuang@szu.edu.cn; x.p.li@szu.edu.cn).

David Ramírez is with the Department of Signal Theory and Communications, Universidad Carlos III de Madrid, 28911 Madrid, Spain, and also with Gregorio Marañón Health Research Institute, 28007 Madrid, Spain (e-mail: david.ramirez@uc3m.es).

Hing Cheung So is with the Department of Electrical Engineering, City University of Hong Kong, 999077 Hong Kong, China (e-mail: hcso@ee.cityu.edu.hk).

Digital Object Identifier 10.1109/TSP.2024.3484582

I. Introduction

APPLICATIONS on small platforms, such as drones, are increasingly driving the development of colocated multiple-input multiple-output (MIMO) radar systems. These platforms operate under stringent resource constraints, making high-precision, high-speed sampling both costly and technically challenging. Additionally, the large data volume generated by traditional sampling methods pose significant challenges for storage, transmission, and processing. One-bit quantization addresses these issues by reducing data volume, lowering power consumption, and simplifying hardware requirements, enabling effective radar sensing without compromising real-time capabilities. As a result, one-bit radar has become an attractive and necessary choice for modern radar systems, leading to significant advancements in radar processing and imaging in recent years [1], [2], [3], [4], [5], [6].

Although one-bit sampling results in information loss, recent studies have demonstrated that this can be effectively mitigated through advanced signal processing techniques. In some instances, these methods can even enhance the overall system performance, for example, through higher sampling rates [7]. Moreover, one-bit radar has proven capable of performing all the functions of traditional high-bit radar, including direction-of-arrival (DOA) estimation [8], [9], [10], [11], range and Doppler estimation [12], [13], [14], [15], detection [16], [17], tracking [18], and imaging [19], [20]. Consequently, one-bit radar is emerging as an important development direction in the radar field, with profound implications for the design and application of future radar systems, particularly in the context of efficient and accurate target detection.

In practical radar signal processing, colored noise is an unavoidable factor that introduces significant challenges compared to white noise. Unlike white noise, which is uncorrelated and typically easier to handle in detection algorithms, colored noise exhibits correlation across time or space, complicating the detection process. The presence of colored noise requires more sophisticated methods, as the noise covariance matrix is no longer diagonal, which impacts the formulation and solution of detection problems. Ignoring the noise correlation can lead to suboptimal detection performance, as standard algorithms designed for white noise may fail to accurately model the noise characteristics, resulting in increased probability of false alarm or missed detection. Despite these challenges, addressing

colored noise is crucial for enhancing the applicability and robustness of radar detection systems in real-world scenarios.

However, most current research on one-bit radar operates under the assumption of white noise, which is often unrealistic in practice [24], [25], [26], [27], [28], [29], [30], [31], [32], [33]. This discrepancy arises because, in the presence of colored noise, the likelihood function is described by central/non-central orthant probabilities, which lack closed-form expressions. Although these probabilities can be evaluated numerically using fast algorithms, they pose challenges for detection tasks since most detection criteria, such as the generalized likelihood ratio test (GLRT), are based on likelihood functions. Concretely, the orthant probabilities make it exceedingly difficult to determine the maximum likelihood estimate (MLE) of the target reflectivity parameter, thereby hindering the formulation of the GLRT. In this paper, we demonstrate that the derivative of the orthant probability with respect to the mean value can be expressed as an orthant probability of a lower dimension, which can also be numerically evaluated. Using this property, we construct a Rao's test [34], [35], [36] to address this problem, circumventing the need to compute the MLE.

It is important to note that Rao's test does not need the MLE only in scenarios where the likelihood under the null is simple, which implies no unknown parameters [37]. Consequently, our detector's development is based on the assumption that the noise covariance matrix is known. However, in practical applications, this matrix is often estimated from noise-only samples using, for instance, the algorithms described in [38], [39], [40], [41]. Hence, our subsequent analysis focuses on the impact of estimation errors on the detector's performance. Initially, we explore how an increased false alarm rate can result from noise covariance mismatch. Specifically, we examine the detector's distribution when the actual noise covariance matrix is assumed to be at a specific mismatched value. These results are then averaged, considering the prior distribution of estimation errors as provided in [38], using an improved Monte Carlo method. Employing this approach enables a more nuanced understanding of how estimation errors, characterized by the prior distribution, might influence the performance of the detector.

Continuing with our analysis, we examine the detector's detection performance and simplify the results to a non-central χ^2 distribution. This simplification provides a clearer view of how noise covariance mismatch affects the detection process, particularly by decreasing the non-centrality parameter. This reduction can be interpreted as a decrease in the "distance" between the null and non-null distributions, which provides valuable insights into the impact of covariance mismatch on the overall detection performance.

Finally, our theoretical findings are validated via computer simulations. It is shown that under realistic settings, the performance degradation due to noise covariance mismatch is almost negligible, especially when the number of noise-only samples is sufficient. This is a common scenario in one-bit processing, as the sampling rate can be significantly increased due to the simple structure of one-bit ADCs. Additionally, the simulation results confirm that the theoretical analysis accurately reflects the impact of noise mismatch on the null distribution, which allows us to adjust the threshold to maintain a constant false alarm rate (CFAR).

The contributions of this paper are summarized as follows:

1) **Development of Rao's Test for One-Bit Target Detection in Colored Noise**: This paper extends our previous work on a white noise detector [7]. To the best of our knowledge, this is the first study in the field of one-bit radar processing that takes into account colored noise. This advancement marks a significant step forward in enhancing the applicability and accuracy of one-bit radar systems.
2) **Accurate Characterization of Null and Non-Null Distributions**: We derive accurate expressions for the null and non-null distributions of the detector, which are crucial for accurately predicting the false alarm and detection probabilities. More importantly, they provide a deeper understanding of the detector's behavior, aiding in a more informed and nuanced approach to radar detection challenges.
3) **Analysis of Noise Covariance Matrix Mismatch Impact**: We conduct an in-depth study of how discrepancies in the noise covariance matrix influence detection performance, including its impact on the null distribution. Our analysis reveals that the noise covariance matrix mismatch leads to an increased false alarm rate, necessitating adaptive threshold adjustments to preserve the CFAR property. Additionally, by approximating the non-null distribution by a non-central χ^2 distribution, we show that performance degradation can be quantified as a decrease in the non-centrality parameter. This insight provides a clear and direct understanding of how noise covariance matrix mismatch affects system performance.

The remainder of this paper is organized as follows: Section II presents the signal model for one-bit detection in colocated MIMO radar under colored noise conditions. Section III details the derivation of a detector based on Rao's test. The analysis of its null and non-null distributions is conducted in Sections IV and V, respectively, which also study the effect of noise mismatch. Section VI provides simulation results to corroborate the theoretical calculations. The paper concludes with a summary of the main findings.

Notation

Throughout this paper, we use boldface uppercase letters for matrices and boldface lowercase letters for column vectors, while lowercase letters denote scalar quantities. The notation $\mathbf{A} \in \mathbb{R}^{p\times q}$ $(\mathbb{C}^{p\times q})$ indicates that $\mathbf{A}$ is a $p \times q$ real (complex) matrix. The (i, j)th entry of $\mathbf{A}$ is denoted by $\mathbf{A}(i, j)$, and $\mathbf{a}(i)$ refers to the ith entry of the vector $\mathbf{a}$. The trace of $\mathbf{A}$ is represented as $\mathrm{tr}(\mathbf{A})$. The function $\mathrm{Diag}(\mathbf{A})$ retrieves the diagonal matrix of $\mathbf{A}$, and $\mathrm{diag}(\mathbf{a})$ produces a diagonal matrix with the elements of $\mathbf{a}$. The superscripts $(\cdot)^{-1}$, $(\cdot)^T$, and $(\cdot)^H$ represent the matrix inverse, transpose, and Hermitian transpose operators, respectively. The operators $\mathbb{E}[a]$ and $\mathbb{V}[a]$ denote the expected value and variance of a, respectively, while $\mathbb{C}[a, b]$ is

the covariance between a and b. The symbol $\sim$ means "distributed as". The terms χ_f^2 and $\chi_f^2(\delta^2)$ refer, respectively, to the central and non-central Chi-squared distributions, where f is the number of degrees of freedom (DOFs), and δ^2 is the non-centrality parameter. Finally, the operators $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ extract the real and imaginary parts of their arguments, $\imath$ denotes the imaginary unit, and $\mathrm{sign}(\cdot)$ indicates the sign of its argument.

II. Signal Model

We begin by examining a colocated MIMO radar setup, encompassing p transmit and m receive antennas. The transmit array emits a probing signal of length n, $\mathbf{S} \in \mathbb{C}^{p\times n}$, which is then reflected by a far-field point source. The received signal at the input of the ADCs can be represented as:

$$\mathbf{X} = \left[\mathbf{x}_1, \cdots, \mathbf{x}_n\right] = \beta \mathbf{a}_r(\phi)\mathbf{a}_t^T(\phi)\mathbf{S} + \mathbf{N}, \tag{1}$$

where β is the target reflectivity, ϕ is the angel of interest, $\mathbf{a}_t(\phi) \in \mathbb{C}^{p\times 1}$ and $\mathbf{a}_r(\phi) \in \mathbb{C}^{m\times 1}$ are the transmit and receive steering vectors, respectively. The term $\mathbf{N} \in \mathbb{C}^{m\times n}$ denotes additive noise that is composed of n independent and identically distributed Gaussian vectors with zero mean and covariance matrix $\mathbf{\Sigma_N}$.

After one-bit quantization, the signal becomes

$$\mathbf{Y} = \left[\mathbf{y}_1, \cdots, \mathbf{y}_n\right] = \mathcal{Q}(\mathbf{X}) = \mathrm{sign}(\mathrm{Re}(\mathbf{X})) + \imath \mathrm{sign}(\mathrm{Im}(\mathbf{X})), \tag{2}$$

where $\mathcal{Q}(\cdot)$ is the complex-valued quantization function.

Our objective is to detect the presence or absence of a target by analyzing the quantized observations $\mathbf{Y}$. Hypothesis $\mathcal{H}_1$ means that the target is present, while hypothesis $\mathcal{H}_0$ implies target absence. For the one bit measurements, the problem of target detection boils down to

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{Y} = \mathcal{Q}\left(\mathbf{N}\right), \\ \mathcal{H}_1 &: \mathbf{Y} = \mathcal{Q}\left(\beta\mathbf{W} + \mathbf{N}\right). \end{aligned} \tag{3}$$

That is, we test $\mathcal{H}_0 : \beta = 0$ vs. $\mathcal{H}_1 : \beta \neq 0$. To simplify the derivation, we have included all known parameters into a single term, denoted as $\mathbf{W} = \mathbf{a}_r(\phi)\mathbf{a}_t^T(\phi)\mathbf{S}$.

Unlike the scenario discussed in [7], this paper delves into a more intricate yet commonly encountered situation where the noise is colored, implying that the covariance matrix $\mathbf{\Sigma_N}$ deviates from being a diagonal matrix. Under such circumstances, the likelihood function is given by the central/non-central orthant probability, which is markedly more complex than the Q function required in the case of white noise [7]. Consequently, there is a need to devise a new detector to address the case of colored noise, which we develop under the following assumptions, commonly found in the MIMO radar target detection literature [24], [25], [26], [27], [28], [29], [30], [31], [32], [33]:

1) **The target reflectivity parameter β remains invariant throughout the observation period.** This assumption is based on the stability of target characteristics within short observation windows, a widely accepted premise in signal processing, especially when target movements are slow or the observation time is brief.
2) **The columns of $\mathbf{N}$ are independently and identically distributed (i.i.d.) with a complex circular Gaussian distribution $\mathcal{CN}(\mathbf{0}, \mathbf{\Sigma_N})$.** This implies that the noise is temporally white but spatially correlated, which is a common scenario in array radar detection, as discussed in [24], [25], [26], [27], [28], [29], [30], [31], [32], [33].
3) **The covariance matrix $\mathbf{\Sigma_N}$ has been estimated using training data and is known to the receiver.** In practice, this involves collecting noise-only samples and applying parameter estimation techniques, which are widely adopted in modern signal processing [24], [25], [26], [27], [28], [29], [30]. Specifically, in the one-bit setting, methods for estimating the covariance matrix have been developed in [38], [39], [40], [41], making the estimation of the noise covariance matrix feasible. It is also important to note that, in traditional ∞-bit systems, adaptive methods can bypass the need for training data by incorporating the estimation process within the detector's formulation [31], [32], [33]. However, this approach is technically prohibitive in the one-bit case, as it involves the computation of orthant probabilities, which do not have closed-form expressions. Finally, we account for the estimation error of the covariance matrix in our performance analysis, ensuring that our model realistically captures the potential degradation of imperfect estimation on detection performance.

III. Detector Design

Since the non-central orthant probability lacks a closed-form expression, standard criteria like the GLRT cannot be applied. Instead, we resort to numerical methods for constructing the detector. In this section, we demonstrate that the derivative of the orthant probability can be represented as a lower-dimensional orthant probability, which is computationally tractable. This insight leads us to formulate Rao's test as a weighted sum of squared derivatives.

We begin by stacking the real and imaginary parts of the received signal $\mathbf{x}_i$ as $\underline{\mathbf{x}}_i = \left[\mathrm{Re}(\mathbf{x}_i)^T, \mathrm{Im}(\mathbf{x}_i)^T\right]^T$ and the quantized signal $\mathbf{y}_i$ as $\underline{\mathbf{y}}_i = \left[\mathrm{Re}(\mathbf{y}_i)^T, \mathrm{Im}(\mathbf{y}_i)^T\right]^T$. Given the circular nature of the noise, the covariance matrix of $\underline{\mathbf{x}}_i$ is[1]

$$\mathbf{\Sigma}_{\underline{\mathbf{x}}_i} = \frac{1}{2}\begin{bmatrix} \mathrm{Re}(\mathbf{\Sigma_N}) & -\mathrm{Im}(\mathbf{\Sigma_N}) \\ \mathrm{Im}(\mathbf{\Sigma_N}) & \mathrm{Re}(\mathbf{\Sigma_N}) \end{bmatrix}. \tag{4}$$

Defining $\mathbf{w}_i = \mathbf{u}_i + \imath\mathbf{v}_i$, where $\mathbf{w}_i$ is the ith column of $\mathbf{W}$, and $\beta = a + \imath b$, the mean of $\underline{\mathbf{x}}_i$ is

$$\mathbb{E}[\underline{\mathbf{x}}_i] = \boldsymbol{\upsilon}_i = \begin{bmatrix} a\mathbf{u}_i - b\mathbf{v}_i \\ a\mathbf{v}_i + b\mathbf{u}_i \end{bmatrix}. \tag{5}$$

Additionally, the orthant probability is defined as

$$P(\boldsymbol{\mu}, \mathbf{\Sigma}) = \int_0^\infty \cdots \int_0^\infty \phi_k(\mathbf{x}; \boldsymbol{\mu}, \mathbf{\Sigma}) \mathrm{d}x_1 \ldots \mathrm{d}x_k, \tag{6}$$

[1]In the following, we will omit the subindex i, since $\mathbf{\Sigma}_{\underline{\mathbf{x}}_i}$ does not very with i.

where $\phi_k(\mathbf{x};\boldsymbol{\mu},\boldsymbol{\Sigma})$ is the probability density function (PDF) of a k-dimensional Gaussian distribution $\mathcal{N}(\boldsymbol{\mu},\boldsymbol{\Sigma})$:

$$\phi_k(\mathbf{x};\boldsymbol{\mu},\boldsymbol{\Sigma})=\frac{1}{(2\pi)^{\frac{k}{2}}\,|\boldsymbol{\Sigma}|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right). \tag{7}$$

Based on these ingredients, and similar to [42], we can write the likelihood of the ith observation as

$$\mathcal{L}(\underline{\mathbf{y}}_i;\boldsymbol{\theta})=\Pr\{\underline{\mathbf{y}}_i;\boldsymbol{\theta}\}=P\left(\boldsymbol{\mu}_i,\boldsymbol{\Omega}_i\right), \tag{8}$$

where $\boldsymbol{\theta}=[a,b]^T$ encapsulates the real and imaginary components of the unknown reflectivity $\beta=a+\imath b$. The means are $\boldsymbol{\mu}_i=\mathbf{Z}_i\boldsymbol{\nu}_i$, with $\mathbf{Z}_i=\mathrm{diag}(\underline{\mathbf{y}}_i)$ and $\boldsymbol{\nu}_i=\mathbf{D}^{-\frac{1}{2}}\boldsymbol{\upsilon}_i$, where $\mathbf{D}=\mathrm{Diag}(\boldsymbol{\Sigma}_{\underline{\mathbf{x}}})$. The matrix $\mathbf{C}=\mathbf{D}^{-\frac{1}{2}}\boldsymbol{\Sigma}_{\underline{\mathbf{x}}}\mathbf{D}^{-\frac{1}{2}}$ is the coherence matrix [45], while $\boldsymbol{\Omega}_i=\mathbf{Z}_i\mathbf{C}\mathbf{Z}_i$ account for sample-specific information. Notably, $\boldsymbol{\mu}_i$ depends on $\boldsymbol{\theta}$ via $\boldsymbol{\upsilon}_i$, whereas $\boldsymbol{\Omega}_i$ does not depend on $\boldsymbol{\theta}$. Considering the independence between samples, the log-likelihood function under $\mathcal{H}_1$ is obtained as:

$$\mathcal{L}(\mathbf{Y};\boldsymbol{\theta})=\sum_{i=1}^{n}\mathcal{L}(\underline{\mathbf{y}}_i;\boldsymbol{\theta})=\sum_{i=1}^{n}\log P(\boldsymbol{\mu}_i,\boldsymbol{\Omega}_i), \tag{9}$$

and by noting that under $\mathcal{H}_0$, $\boldsymbol{\theta}=\boldsymbol{\theta}_0=[0,0]^T$, the log-likelihood function becomes

$$\mathcal{L}(\mathbf{Y};\boldsymbol{\theta}_0)=\sum_{i=1}^{n}\log P(\mathbf{0}_{2m},\boldsymbol{\Omega}_i), \tag{10}$$

which only depends on the known covariance matrices.

Since there are no unknown parameters under $\mathcal{H}_0$, we propose to use Rao's test to address this problem, given by

$$T_{\mathrm{R}}=\left(\frac{\partial\mathcal{L}(\mathbf{Y};\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right)^T\mathbf{F}^{-1}(\boldsymbol{\theta}_0)\left(\frac{\partial\mathcal{L}(\mathbf{Y};\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right), \tag{11}$$

where $\mathbf{F}(\boldsymbol{\theta})$ is the Fisher information matrix (FIM):

$$\mathbf{F}(\boldsymbol{\theta})=\mathbb{E}\left[\frac{\partial\mathcal{L}(\mathbf{Y};\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\frac{\partial\mathcal{L}(\mathbf{Y};\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^T}\right]. \tag{12}$$

Denoting $P_i=P(\boldsymbol{\mu}_i,\boldsymbol{\Omega}_i)$, the derivatives are

$$\frac{\partial\mathcal{L}(\mathbf{Y};\boldsymbol{\theta})}{\partial a}=\sum_{i=1}^{n}\frac{1}{P_i}\frac{\partial P_i}{\partial a}=\sum_{i=1}^{n}\sum_{j=1}^{2m}\frac{1}{P_i}\frac{\partial P_i}{\partial\boldsymbol{\mu}_i(j)}\frac{\partial\boldsymbol{\mu}_i(j)}{\partial a} \tag{13a}$$

$$\frac{\partial\mathcal{L}(\mathbf{Y};\boldsymbol{\theta})}{\partial b}=\sum_{i=1}^{n}\frac{1}{P_i}\frac{\partial P_i}{\partial b}=\sum_{i=1}^{n}\sum_{j=1}^{2m}\frac{1}{P_i}\frac{\partial P_i}{\partial\boldsymbol{\mu}_i(j)}\frac{\partial\boldsymbol{\mu}_i(j)}{\partial b}, \tag{13b}$$

where the right-most derivatives are

$$\frac{\partial\boldsymbol{\mu}_i}{\partial a}=\mathbf{Z}_i\mathbf{D}^{-\frac{1}{2}}\begin{bmatrix}\mathbf{u}_i\\ \mathbf{v}_i\end{bmatrix}=\mathbf{Z}_i\mathbf{a}_i \tag{14a}$$

$$\frac{\partial\boldsymbol{\mu}_i}{\partial b}=\mathbf{Z}_i\mathbf{D}^{-\frac{1}{2}}\begin{bmatrix}-\mathbf{v}_i\\ \mathbf{u}_i\end{bmatrix}=\mathbf{Z}_i\mathbf{b}_i. \tag{14b}$$

To compute the derivative of the orthant probability with respect to the mean values, we introduce the following lemma.

Lemma 1: For an orthant probability $P(\boldsymbol{\mu},\boldsymbol{\Sigma})$, the derivative with respect to the jth element of $\boldsymbol{\mu}$ is

$$\frac{\partial P(\boldsymbol{\mu},\boldsymbol{\Sigma})}{\partial\boldsymbol{\mu}(j)}=\frac{1}{\sqrt{2\pi\boldsymbol{\Sigma}(j,j)}}P(\boldsymbol{\omega}(\boldsymbol{\mu},j),\mathbf{R}(\boldsymbol{\Sigma},j)), \tag{15}$$

where $\boldsymbol{\omega}(\boldsymbol{\mu},j)$ denotes the reduced vector after removing the jth element of $\boldsymbol{\mu}$ and

$$\mathbf{R}(\boldsymbol{\Sigma},j)=\boldsymbol{\Theta}(\boldsymbol{\Sigma},j)-\frac{1}{\boldsymbol{\Sigma}(j,j)}\mathbf{r}_j\mathbf{r}_j^T, \tag{16}$$

where $\boldsymbol{\Theta}(\boldsymbol{\Sigma},j)$ denotes the reduced matrix after removing the jth row and column of $\boldsymbol{\Sigma}$, $\boldsymbol{\Sigma}(i,j)$ is the (i,j)th element of $\boldsymbol{\Sigma}$ and $\mathbf{r}_j$ is the reduced vector after removing the jth element of the jth column of $\boldsymbol{\Sigma}$.

Proof: See Appendix A. ■

This lemma has demonstrated that the derivative of a k-dimensional orthant probability can be computed from a $(k-1)$-dimensional orthant probability. Consequently, it enables us to express:

$$\frac{\partial P_i}{\partial\boldsymbol{\mu}_i(j)}\bigg|_{\boldsymbol{\theta}=\mathbf{0}}=\frac{1}{\sqrt{2\pi}}P(\mathbf{0}_{2m-1},\mathbf{R}(\boldsymbol{\Omega}_i,j)). \tag{17}$$

Subsequently, the derivatives in (13) become:

$$\frac{\partial\mathcal{L}(\mathbf{Y};\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\bigg|_{\boldsymbol{\theta}=\mathbf{0}}=\begin{bmatrix}\mathrm{tr}(\mathbf{A}^T\mathbf{L}\mathbf{P}^{-1})\\ \mathrm{tr}(\mathbf{B}^T\mathbf{L}\mathbf{P}^{-1})\end{bmatrix}, \tag{18}$$

where the $2m\times n$ matrices $\mathbf{A}$ and $\mathbf{B}$ are defined as:

$$\mathbf{A}=[\mathbf{a}_1,\cdots,\mathbf{a}_n],\qquad \mathbf{B}=[\mathbf{b}_1,\cdots,\mathbf{b}_n], \tag{19}$$

and $\mathbf{P}$ is the $n\times n$ diagonal matrix:

$$\mathbf{P}=\mathrm{diag}(P(\mathbf{0}_{2m},\boldsymbol{\Omega}_1),\cdots,P(\mathbf{0}_{2m},\boldsymbol{\Omega}_n)). \tag{20}$$

The $2m\times n$ matrix $\mathbf{L}$ is defined as:

$$\mathbf{L}=\left[\mathbf{Z}_1\frac{\partial P_1}{\partial\boldsymbol{\mu}_1}\bigg|_{\boldsymbol{\theta}=\mathbf{0}},\cdots,\mathbf{Z}_n\frac{\partial P_n}{\partial\boldsymbol{\mu}_n}\bigg|_{\boldsymbol{\theta}=\mathbf{0}}\right], \tag{21}$$

with each element given by:

$$\frac{\partial P_i}{\partial\boldsymbol{\mu}_i}\bigg|_{\boldsymbol{\theta}=\mathbf{0}}=\frac{1}{\sqrt{2\pi}}[P(\mathbf{0}_{2m-1},\mathbf{R}(\boldsymbol{\Omega}_i,1)),\cdots,P(\mathbf{0}_{2m-1},\mathbf{R}(\boldsymbol{\Omega}_i,2m))]^T. \tag{22}$$

In addition, Appendix B shows that the FIM is

$$\mathbf{F}(\boldsymbol{\theta}_0)=\upsilon^2\mathbf{I}_2, \tag{23}$$

where

$$\upsilon^2=\mathrm{tr}\left(\boldsymbol{\Delta}_1^T\boldsymbol{\Delta}_1\mathbf{O}^{-1}\right)=\mathrm{tr}\left(\boldsymbol{\Delta}_2^T\boldsymbol{\Delta}_2\mathbf{O}^{-1}\right), \tag{24}$$

with $\mathbf{O}$ defined in (99), and $\boldsymbol{\Delta}_1$ and $\boldsymbol{\Delta}_2$ defined in (100).

As a result, Rao's test is formulated as:

$$T_{\mathrm{R}}=\frac{1}{\upsilon^2}\left[\mathrm{tr}^2(\mathbf{A}^T\mathbf{L}\mathbf{P}^{-1})+\mathrm{tr}^2(\mathbf{B}^T\mathbf{L}\mathbf{P}^{-1})\right]\underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}}\gamma. \tag{25}$$

Combining (25) with (21) and (19), it becomes clear that the detector can be interpreted as a weighted matched filter, where by the weights are calculated from the partial derivatives in

matrix $\mathbf{L}$ and the orthant probabilities in matrix $\mathbf{P}$, which are determined by the elements of the noise covariance matrix.

Remark 1: To build the detector, it is necessary to compute 2^{2m} orthant probabilities of dimension $2m$ to construct $\mathbf{O}$ and $\mathbf{P}$. We demonstrate in Appendix B that these probabilities can be grouped into sets of four with identical values, reducing the required calculations to 2^{2m-2}. Regarding the derivatives, our computations involve $2mn$ orthant probabilities of dimension $2m-1$, which may be identical for repeated observations. Taking into account the previously discussed symmetry, the computational load for the derivatives is minimized to $\min(m2^{2m-1}, 2m\tilde{n})$, assuming there are $\tilde{n}\,(\leq n)$ distinct observations. Additionally, these probabilities need to be calculated only once for a specific noise covariance matrix and do not require updates unless the matrix changes. Several efficient algorithms for evaluating orthant probabilities, as documented in [43], [44], significantly reduce the computational burden in detector construction, enhancing the feasibility of practical, real-world applications.

IV. Null Distribution

In this section, we delve into the null distribution of the proposed detector T_{R}. Initially, we consider the scenario with perfectly known noise covariance matrix. Subsequently, we analyze the effect on the false alarm rate when the noise covariance matrix is known up to some estimation error. Our study employs an improved Monte Carlo approach, which begins with a distribution for a specific error matrix. We then compute the average effect over the prior distribution of this error matrix via Monte Carlo integration, leading to the null distribution for the scenario of imperfect noise covariance matrix estimation.

A. Distribution With Known Noise Covariance

Let us start by rewriting the test statistic as

$$T_{\mathrm{R}} = w_1^2 + w_2^2, \tag{26}$$

where

$$w_1 = \frac{1}{\upsilon}\mathrm{tr}(\mathbf{A}^T\mathbf{L}\mathbf{P}^{-1}), \qquad w_2 = \frac{1}{\upsilon}\mathrm{tr}(\mathbf{B}^T\mathbf{L}\mathbf{P}^{-1}). \tag{27}$$

Since the samples are i.i.d. under $\mathcal{H}_0$, it follows straightforwardly from the central limit theorem that w_1 and w_2 follow asymptotically ($n \to \infty$) a 2-dimensional joint Gaussian distribution. In addition, recalling that

$$\mathbf{w} = [w_1, w_2]^T = \frac{1}{\upsilon}\left[\frac{\partial \mathcal{L}(\mathbf{Y};\boldsymbol{\theta})}{\partial a}, \frac{\partial \mathcal{L}(\mathbf{Y};\boldsymbol{\theta})}{\partial b}\right]^T, \tag{28}$$

and taking (23) into account, it is easy to show that

$$\mathbb{E}[\mathbf{w}] = \mathbf{0}_2, \qquad \mathbb{E}[\mathbf{w}\mathbf{w}^T] = \mathbf{I}_2, \tag{29}$$

which allows us to conclude that T_{R} is asymptotically ($n \to \infty$) Chi-square distributed with 2 degrees of freedom (DoFs),

$$T_{\mathrm{R}} \sim \chi_2^2. \tag{30}$$

That is, T_{R} is exponentially distributed with parameter $1/2$, the probability of false alarm becomes

$$P_{\mathrm{fa}}(\gamma) = \Pr\{T_{\mathrm{R}} > \gamma\} = \exp(-\gamma/2), \tag{31}$$

and the detection threshold can be obtained as

$$\gamma = -2\log(P_{\mathrm{fa}}). \tag{32}$$

B. Distribution With Estimated Noise Covariance

In the preceding derivations, the noise covariance matrix was assumed to be perfectly known to the receiver. However, in real problems, it must be estimated beforehand using, for instance, the algorithms in [38], [39], [40]. Accordingly, we need to account for the mismatch between the true and estimated noise covariance matrices and assess its impact. In this subsection, the scenario where the estimated covariance matrix, $\boldsymbol{\Sigma}_{\underline{\mathbf{x}}}$, and the true covariance matrix, $\boldsymbol{\Sigma}'_{\underline{\mathbf{x}}}$, differ is examined, aiming to study the changes of the null distribution. We proceed with two assumptions: 1) the true covariance matrix $\boldsymbol{\Sigma}'_{\underline{\mathbf{x}}}$ is close to the estimate $\boldsymbol{\Sigma}_{\underline{\mathbf{x}}}$, a condition that can be ensured by a sufficient amount of (training) samples, and 2) the prior distribution of $\boldsymbol{\Sigma}'_{\underline{\mathbf{x}}}$ given $\boldsymbol{\Sigma}_{\underline{\mathbf{x}}}$ is known, which was obtained in [38].

We start by analyzing the joint distribution of w_1 and w_2 for the estimated covariance matrix, which is given by the following theorem.

Theorem 1: Under noise covariance mismatch, the distribution of $\mathbf{w} = [w_1, w_2]^T$ can be asymptotically ($n \to \infty$) approximated by the real Gaussian distribution with zero mean and covariance matrix

$$\boldsymbol{\Sigma}_w = \mathbb{E}[\mathbf{w}\mathbf{w}^T] = \frac{\upsilon_1^2}{\upsilon^2}\mathbf{I}_2, \tag{33}$$

where

$$\upsilon_1^2 = \mathrm{tr}\left(\boldsymbol{\Delta}_1^T\boldsymbol{\Delta}_1\mathbf{G}\right) = \mathrm{tr}\left(\boldsymbol{\Delta}_2^T\boldsymbol{\Delta}_2\mathbf{G}\right), \tag{34}$$

with

$$\mathbf{G} = \mathrm{diag}\left(\frac{O'_1}{O_1^2}, \cdots, \frac{O'_\kappa}{O_\kappa^2}\right), \tag{35}$$

and $O'_1, \cdots, O'_\kappa$ are defined in (102).

Proof: See Appendix C. ∎

Using Theorem 1, we have

$$\frac{\upsilon^2}{\upsilon_1^2}T_{\mathrm{R}} \sim \chi_2^2. \tag{36}$$

Therefore, the probability of false alarm can be rewritten as

$$P_{\mathrm{fa}}(\gamma) = \exp\left(-\frac{\upsilon^2}{2\upsilon_1^2}\gamma\right), \tag{37}$$

and the detection threshold is

$$\gamma = -\frac{2\upsilon_1^2}{\upsilon^2}\log(P_{\mathrm{fa}}). \tag{38}$$

C. Average Over Prior Distribution of Estimation Error

In this section, we proceed to average the null distribution over the prior distribution of the estimation error. Such a prior distribution can be derived using analytical methods detailed in the literature, such as [38]. Assuming a known prior PDF for $\boldsymbol{\Sigma}'_{\underline{\mathbf{x}}}$, denoted by $f(\boldsymbol{\Sigma}'_{\underline{\mathbf{x}}})$, the probability of false alarm can be computed as

$$\begin{aligned} P_{\mathrm{fa}}(\gamma) &= \int \Pr\{T_{\mathrm{R}} > \gamma | \boldsymbol{\Sigma}'_{\underline{\mathbf{x}}}\} f(\boldsymbol{\Sigma}'_{\underline{\mathbf{x}}}) \mathrm{d}\boldsymbol{\Sigma}'_{\underline{\mathbf{x}}} \\ &= \int \exp\left(-\frac{\upsilon^2}{2\upsilon_1^2}\gamma\right) f(\boldsymbol{\Sigma}'_{\underline{\mathbf{x}}}) \mathrm{d}\boldsymbol{\Sigma}'_{\underline{\mathbf{x}}}. \end{aligned} \tag{39}$$

Given that evaluating υ_1^2 for each $\boldsymbol{\Sigma}'_{\underline{\mathbf{x}}}$ involves computing 2^{2m-2} orthant probabilities, which do not have a closed-form, computing this integral directly is very challenging. To circumvent this, we present an improved Monte Carlo method to approximate (39). Concretely, we generate a sequence of K covariance matrices from the prior distribution $f(\boldsymbol{\Sigma}'_{\underline{\mathbf{x}}})$ and approximate the false alarm probability by averaging the outcomes as

$$P_{\mathrm{fa}}(\gamma) \approx \frac{1}{K}\sum_{i=1}^{K} \exp\left(-\frac{\upsilon^2}{2\upsilon_{1,i}^2}\gamma\right), \tag{40}$$

where $\upsilon_{1,i}^2 = \mathrm{tr}(\boldsymbol{\Delta}_1^T \boldsymbol{\Delta}_1 \mathbf{G}_i) = \mathrm{tr}(\boldsymbol{\Delta}_2^T \boldsymbol{\Delta}_2 \mathbf{G}_i)$, with

$$\mathbf{G}_i = \mathrm{diag}\left(\frac{O'_{1,i}}{O_1^2}, \cdots, \frac{O'_{\kappa,i}}{O_\kappa^2}\right), \tag{41}$$

and $O'_{j,i} = P(\mathbf{0}_{2m}, \mathbf{C}'_{j,i})$, with $\mathbf{C}'_{j,i} = \boldsymbol{\Gamma}_j \mathbf{C}'_i \boldsymbol{\Gamma}_j$, where $\mathbf{C}'_i$ is the coherence matrix of the ith covariance matrix sample, and $\boldsymbol{\Gamma}_j$ is defined in Appendix B.

Achieving an accurate approximation requires a large K, thereby increasing computational complexity since computing $\upsilon_{1,i}^2$, for $i = 1, \ldots, K$, necessitates the calculation of 2^{2m-2} orthant probabilities. However, this process can be optimized by using a Taylor's expansion around $\mathbf{C}_j = \boldsymbol{\Gamma}_j \mathbf{C} \boldsymbol{\Gamma}_j$, where $\mathbf{C}$ is the coherence matrix of the estimated noise covariance matrix, $\boldsymbol{\Sigma}_{\underline{\mathbf{x}}}$, to approximate the orthant probabilities required for $\upsilon_{1,i}^2$, corresponding to each $\boldsymbol{\Sigma}'_{\underline{\mathbf{x}}_i}$. This approximation can be formulated as:

$$P(\mathbf{0}_{2m}, \mathbf{C}'_{j,i}) \approx P(\mathbf{0}_{2m}, \mathbf{C}_j) + \left.\frac{\partial P(\mathbf{0}_{2m}, \mathbf{C})}{\partial \mathbf{c}}\right|_{\mathbf{c}=\mathbf{c}_j} (\mathbf{c}'_{j,i} - \mathbf{c}_j), \tag{42}$$

where

$$\mathbf{c} = \mathrm{vech}(\mathbf{C}) = [\mathbf{C}(1,2), \cdots, \mathbf{C}(2m-1, 2m)]^T \tag{43}$$

is the vectorization of the upper triangular part of $\mathbf{C}$, excluding the main diagonal, which stacks the free parameters of $\mathbf{C}$ in a vector. Moreover, $\mathbf{c}_j = \mathrm{vech}(\mathbf{C}_j)$ and $\mathbf{c}'_{j,i} = \mathrm{vech}(\mathbf{C}'_{j,i})$. The partial derivative in this expression can be efficiently computed using the following theorem.

Theorem 2: The derivative of the orthant probability $P(\mathbf{0}_k, \mathbf{C})$ with respect to the correlation coefficient $\mathbf{C}(r,s)$, $r < s$, is

$$\frac{\partial P(\mathbf{0}_{2m}, \mathbf{C})}{\partial \mathbf{C}(r,s)} = \frac{P(\mathbf{0}_{k-2}, \bar{\mathbf{C}})}{2\pi\sqrt{1 - |\mathbf{C}(r,s)|^2}}, \tag{44}$$

where $\bar{\mathbf{C}} = [\boldsymbol{\Theta}(\boldsymbol{\Theta}(\mathbf{C}^{-1}, r), s-1)]^{-1}$.

Proof: See Appendix D. ∎

By employing this approach, we can avoid computing the orthant probability for each $\boldsymbol{\Sigma}'_{\underline{\mathbf{x}}_i}$. Instead, we only need to compute the derivative and then generate a large number of $\boldsymbol{\Sigma}'_{\underline{\mathbf{x}}_i}$ samples to obtain a reliable approximation to the null distribution. Let us denote:

$$\upsilon_{1,i}^2 = \upsilon^2 + \upsilon_{\triangle,i}^2, \tag{45}$$

where

$$\upsilon_{\triangle,i}^2 = \mathrm{tr}\left(\boldsymbol{\Delta}_1^T \boldsymbol{\Delta}_1 (\mathbf{G}_i - \mathbf{O}^{-1})\right) = \mathrm{tr}\left(\boldsymbol{\Delta}_2^T \boldsymbol{\Delta}_2 (\mathbf{G}_i - \mathbf{O}^{-1})\right), \tag{46}$$

and

$$\mathbf{G}_i - \mathbf{O}^{-1} = \mathrm{diag}\left(\frac{\Delta O_{1,i}}{O_1^2}, \cdots, \frac{\Delta O_{\kappa,i}}{O_\kappa^2}\right). \tag{47}$$

In this expression, $\Delta O_{j,i} = O'_{j,i} - O_j$ can be approximated using (42) as

$$\Delta O_{j,i} = \left.\frac{\partial P(\mathbf{0}_{2m}, \mathbf{C})}{\partial \mathbf{c}}\right|_{\mathbf{c}=\mathbf{c}_j} (\mathbf{c}'_{j,i} - \mathbf{c}_j). \tag{48}$$

Thus, orthant probabilities are computed only once for the derivatives of each $\mathbf{C}_j$ and not for every realization of $\boldsymbol{\Sigma}'_{\underline{\mathbf{x}}_i}$. Subsequently, substituting (45) into (37) allows the average probability of false alarm to be estimated as:

$$P_{\mathrm{fa}}(\gamma) \approx \frac{1}{K}\sum_{i=1}^{K} \exp\left(-\frac{\upsilon^2}{2\left(\upsilon^2 + \upsilon_{\triangle_i}^2\right)}\gamma\right). \tag{49}$$

Complexity Comparison: Previously, as described in (40) and (42), evaluating $K2^{2m-2}$ orthant probabilities is required to construct (40). However, by applying Taylor's expansion, the number of orthant probabilities needed is reduced to a fixed value of $(m^2 - m)2^{2m-1}$. This fixed number no longer scales with the number of samples drawn from the prior distribution of covariance matrix estimation errors. Consequently, we can achieve high approximation accuracy with a large number of samples. Additionally, this sampling can be directly applied to the shrinkage factor in (49), eliminating the need to compute the distribution for each sample. This approach significantly accelerates the distribution computation, making the process much more efficient.

V. Non-Null Distribution

In this section, the non-null distribution of the proposed detector is examined. First, a generalized non-central χ^2 distribution is introduced through the analysis of the joint distribution of w_1 and w_2. Afterwards, a simplified representation is derived for the low signal-to-noise ratio (SNR) scenario using a Taylor's expansion, yielding a standard non-central χ^2 distribution.

A. Fundamental Result

Let us start by computing the mean and covariance matrix of $\mathbf{w}$ under $\mathcal{H}_1$, which are presented in the following theorem.

Theorem 3: Under $\mathcal{H}_1$, the mean and covariance matrix of $\mathbf{w}$ are

$$\mathbf{u}_w = \frac{1}{\upsilon}\begin{bmatrix}\mathrm{tr}(\mathbf{E}_1\mathbf{Q}^T)\\ \mathrm{tr}(\mathbf{E}_2\mathbf{Q}^T)\end{bmatrix}, \qquad \mathbf{\Sigma}_w = \begin{bmatrix}\sigma_1^2 & \sigma_{12}\\ \sigma_{12} & \sigma_2^2\end{bmatrix}, \tag{50}$$

where $\mathbf{E}_l = \mathbf{\Delta}_l\mathbf{O}^{-1}, l = 1, 2,$ and $\mathbf{Q}(i,j) = P(\mathbf{\Gamma}_j\boldsymbol{\nu}_i, \mathbf{\Gamma}_j\mathbf{C}\mathbf{\Gamma}_j), i = 1, \ldots, n, j = 1, \ldots, \kappa.$ The proof of this result, along with the definition of the elements of the covariance matrix, is given in Appendix E.

Having obtained the mean and covariance matrix of $\mathbf{w}$ under $\mathcal{H}_1$, we now define

$$\mathbf{\Sigma}_w = \mathbf{P}^T\mathbf{\Lambda}\mathbf{P}, \qquad \mathbf{m} = \mathbf{P}\mathbf{\Sigma}_w^{-\frac{1}{2}}\mathbf{u}_w. \tag{51}$$

Then, as in [7], the detector can be rewritten as

$$T_{\mathrm{R}} = \lambda_1(\nu_1 + m_1)^2 + \lambda_2(\nu_2 + m_2)^2, \tag{52}$$

where $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2)$, $\mathbf{m} = [m_1, m_2]^T$, and ν_1, ν_2 are mutually independent standard Gaussian random variables. Thus, the detection probability of T_{R} is given by a general non-central χ^2 distribution:

$$T_{\mathrm{R}} = \sum_{l=1}^{2}\lambda_l\chi_1^2(m_l^2), \tag{53}$$

which can be numerically evaluated [46].

It is easily proved that in the case of a mismatched noise covariance matrix, the result is adjusted by substituting $\mathbf{Q}$ with $\mathbf{Q}'$, which is defined as follows:

$$\mathbf{Q}'(i,j) = P(\mathbf{\Gamma}_j\boldsymbol{\nu}_i', \mathbf{\Gamma}_j\mathbf{C}'\mathbf{\Gamma}_j), \tag{54}$$

Here, $\mathbf{C}'$ represents the coherence matrix corresponding to the mismatched covariance matrix $\mathbf{\Sigma}_{\underline{\mathbf{x}}}'$ and $\boldsymbol{\nu}_i' = \mathrm{diag}(\mathbf{\Sigma}_{\underline{\mathbf{x}}}')^{-\frac{1}{2}}\boldsymbol{\upsilon}_i$ characterizes the changes in the mean due to the covariance mismatch.

B. Low-SNR Approximation

Despite its near-exact nature, the aforementioned approximation requires the computation of the orthant probability for all n samples. Moreover, it does not allow for an insightful comparison between matched and mismatched cases. Thus, in this subsection, a simplified approximation is presented for the low-SNR regime via a Taylor's approximation, which avoids additional orthant probability computations and allows for the aforementioned comparison. The derived result is presented in the following theorem.

Theorem 4: In the low-SNR regime, where $|\beta| = \mathcal{O}(n^{-\frac{1}{2}})$, the mean and covariance matrix of vector $\mathbf{w}$ for the matched noise covariance case are:

$$\mathbf{u}_w = \upsilon\begin{bmatrix}a\\ b\end{bmatrix} + \mathcal{O}(n^{-\frac{1}{2}}), \qquad \mathbf{\Sigma}_w = \mathbf{I}_2 + \mathcal{O}(n^{-\frac{1}{2}}). \tag{55}$$

Furthermore, for the mismatched noise covariance case, they become

$$\mathbf{u}_w = \frac{1}{\upsilon}\begin{bmatrix}a\varsigma_{11}+b\varsigma_{12}\\ a\varsigma_{21}+b\varsigma_{22}\end{bmatrix} + \mathcal{O}(n^{-\frac{1}{2}}), \quad \mathbf{\Sigma}_{\mathbf{w}} = \frac{\upsilon_1^2}{\upsilon^2}\mathbf{I}_2 + \mathcal{O}(n^{-\frac{1}{2}}), \tag{56}$$

where $\varsigma_{ij} = \mathrm{tr}\left(\mathbf{\Delta}_i^T\mathbf{\Delta}_j'\mathbf{O}^{-1}\right), i, j = 1, 2.$

Proof: See Appendix F. ∎

Taking this result into account, it is straightforward to show for the matched case that

$$T_{\mathrm{R}} \sim \chi_2^2(\delta^2), \tag{57}$$

where $\delta^2 = \upsilon^2|\beta|^2$. For the mismatched noise covariance matrix, the approximation yields

$$\frac{\upsilon^2}{\upsilon_1^2}T_{\mathrm{R}} \sim \chi_2^2(\delta'^2), \tag{58}$$

where $\delta'^2 = (a^2\varsigma_1^4 + b^2\varsigma_2^4)/\upsilon_1^2$.

Combined with (30) and (36), we can see that the detection power is determined by the non-centrality parameters δ^2 and δ'^2. A direct comparison shows that the noise covariance mismatch has caused the detection performance to decrease from δ^2 to δ'^2.

Remark 2: It is worth mentioning that the above results can also be applied to scenarios where the signal matrix $\mathbf{W}$ is mismatched due to factors such as steering vector inaccuracies or mismatches in the signal waveform, which are common in practical scenarios. Our analysis, which considers inaccuracies in the diagonal elements of the noise covariance matrix, can be equivalently modeled as errors in $\mathbf{W}$ and thus is applicable to address such mismatches as well. This extension provides additional insight into the effect of signal parameter mismatch on detection performance, allowing for potential optimization of the waveform to either improve robustness or enhance mismatch discrimination capabilities, depending on the desired application.

C. Performance Loss

Building on the results above, we can now compare the detection performance of a one-bit system with that of an ∞-bit system. In the latter, the detector is known as the generalized matched filter, whose performance has been analyzed in [37, (Section 4.4.1)] for the real-valued case. Extending this analysis to the complex circular Gaussian noise case, the distributions under the two hypotheses are:

$$T_\infty = |\mathrm{tr}(\mathbf{W}^H\mathbf{\Sigma}_{\mathbf{N}}^{-1}\mathbf{X})|^2 \sim \begin{cases}\chi_2^2, & \text{under } \mathcal{H}_0,\\ \chi_2^2(\delta_\infty^2), & \text{under } \mathcal{H}_1,\end{cases} \tag{59}$$

where $\delta_\infty^2 = |\beta|^2\mathrm{tr}^2(\mathbf{W}'\mathbf{\Sigma}_{\mathbf{N}}^{-1}\mathbf{W})$. Consequently, the performance loss can be measured by the reduction in the non-centrality parameter, quantified as:

$$\frac{\delta_\infty^2}{\delta^2} = \frac{\mathrm{tr}^2(\mathbf{W}'\mathbf{\Sigma}_{\mathbf{N}}^{-1}\mathbf{W})}{\upsilon^2}. \tag{60}$$

It is well known that in the case of white noise, the performance loss can be easily calculated as $\pi/2 \approx 1.96$ dB [7]. However, in the presence of colored noise, the calculation of υ^2 becomes significantly more complex and does not have a straightforward form, making it difficult to directly compare with the numerator. Therefore, we choose to study this relationship numerically to quantify the performance as a function of noise correlation, which will be presented in the numerical simulations.

VI. Numerical Results

In this section, we conduct numerical simulations to validate our theoretical findings. Initially, we evaluate the accuracy of the derived theoretical null distribution in both matched and mismatched scenarios. Subsequently, we examine the accuracy of the derived theoretical non-null distributions in both scenarios, along with their low-SNR approximations. Finally, we compare the detection performance of the proposed detector, using the receiver operating characteristics (ROC) curve, with the existing one-bit white noise detector proposed in [7].

We examine a colocated MIMO radar system equipped with a uniform linear array, with inter-element spacing set to half the wavelength. The system parameters are configured as $m = p = 4$, $n = 2000$, and the direction-of-arrival (DOA) θ is set to 30°. Similar to [47], [48], we employ an orthogonal linear frequency modulation (LFM) waveform for transmission, directed at angle θ:

$$\mathbf{S}(k,l) = \frac{\exp\left\{\frac{\imath}{n}[2\pi(l-1) + \pi(l-1)^2 + (k-1)\sin(\theta)]\right\}}{\sqrt{p}}, \tag{61}$$

where $k = 1, \ldots, p$ and $l = 1, \ldots, n$.

The noise covariance matrix is constructed as

$$\mathbf{\Sigma_N} = \frac{\alpha \mathbf{H}\mathbf{H}^H + \mathbf{I}_m}{\mathrm{tr}(\alpha \mathbf{H}\mathbf{H}^H + \mathbf{I}_m)/m}. \tag{62}$$

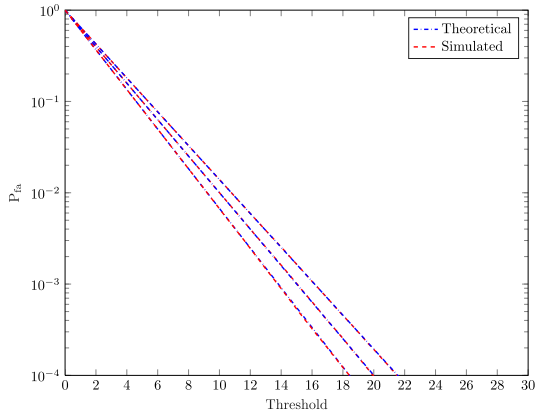
where $\mathbf{H} \in \mathbb{C}^{m \times m}$ has i.i.d. elements drawn from a standard complex-valued Gaussian distribution, and α is a scaling factor used to modulate the correlation coefficients. In this simulation, we set $\alpha = 1$. In the scenario where noise covariance matrix estimation error is present, the true noise covariance matrix is expressed as

$$\mathbf{\Sigma'_N} = \mathbf{\Sigma_N} + \Delta\mathbf{\Sigma_N}. \tag{63}$$

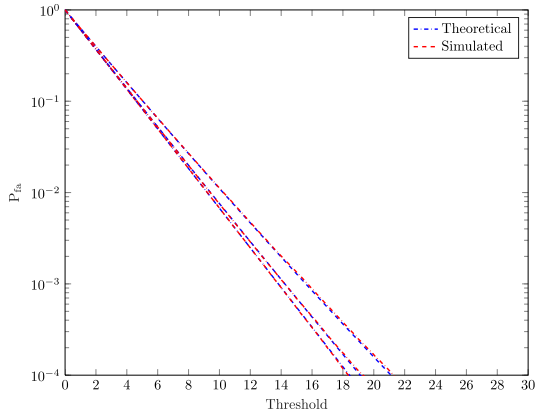
where $\mathbf{\Sigma_N}$ is the one used in the formulation of the detector. The perturbation matrix $\Delta\mathbf{\Sigma_N}$ is generated as a Hermitian matrix, with its lower triangular elements drawn from a normal distribution $\mathcal{N}(\mathbf{0}_{m^2}, \rho^2\mathbf{I}_{m^2})$, where ρ^2 acts as a scaling factor quantifying the level of error in the covariance matrix, which, in practice, is proportional to the inverse of the length of the training sequence. For instance, for a training sequence of length 1000, it is reasonable to assume $\rho^2 < 0.001$, according to the analysis in [38]. To ensure that the generated covariance matrix $\mathbf{\Sigma'_N}$ remains positive definite, we exclude rare instances where $\mathbf{\Sigma'_N}$ is not positive definite from the simulation.

A target with a constant reflectivity parameter is assumed. Specifically, β is an unknown value whose phase is uniformly distributed on the interval $[0, 2\pi)$ and its amplitude is set according to the SNR, defined as:

$$\mathrm{SNR} = 10\log_{10}\left(\frac{p|\beta|^2}{\mathrm{tr}(\mathbf{\Sigma_N})}\right). \tag{64}$$



(a) Fixed noise covariance matrix

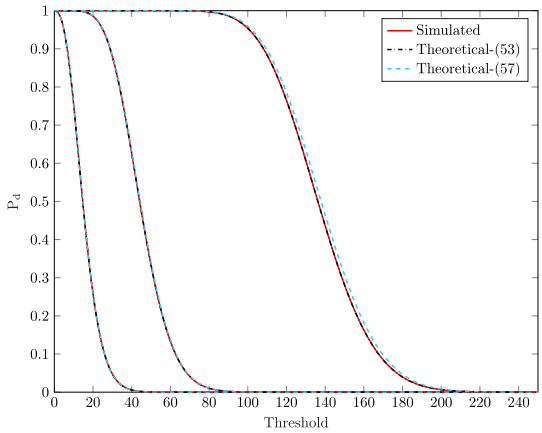


(b) Random noise covariance matrix

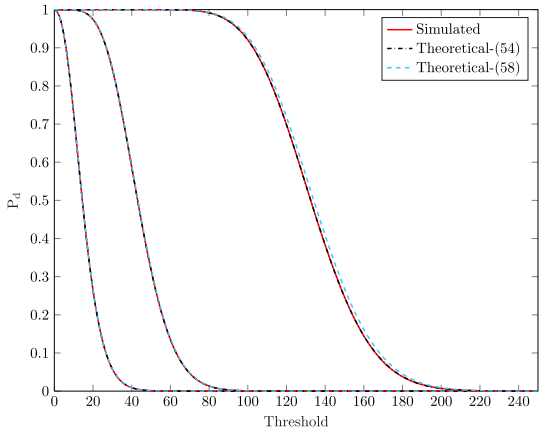
Fig. 1. Probability of false alarm versus threshold with $m = p = 4$, $n = 2000$, and $\rho = [0, 0.1, 0.2]$.

A. Null Distribution

Initially, we examine a scenario where the true covariance matrix is assumed to be a specific value, essentially considering one realization of (63). Fig. 1(a) illustrates the probability of false alarms for varying threshold in an experiment characterized by mismatch levels $\rho = [0, 0.05, 0.1]$ (curves are ordered from left to right). Note that we chose relatively large ρ values in this figure to make the curves more distinguishable, whereas in realistic conditions, smaller values (e.g., $\rho \leq 0.03$) are more typical. For the no-mismatch case ($\rho = 0$), we compare the false alarm probability derived from Monte Carlo simulations with the theoretical one in (31). As depicted in the leftmost curve of the figure, there is an almost perfect alignment between the theoretical predictions and simulation results. This congruence holds true even in the presence of mismatched noise covariance matrices ($\rho \neq 0$), demonstrating the accuracy of (37).



(a) Matched case



(b) Mismatched case

Fig. 2. Probability of detection versus threshold in an experiment with $m = p = 4$, $n = 2000$, and SNR $= [-30, -25, -20]$ dBs.

Subsequently, we evaluate our method's precision for computing the average probability of false alarms. In this approach, $\mathbf{\Sigma_N}$ is not fixed but generated from $K = 1000$ samples according to (63), and the proposed approximation in (49) is used. Using the same parameters as in the previous experiment, Fig. 1(b) clearly validates the theoretical approximation's precision in (49). Moreover, the advantage of (49) lies in its computational efficiency with respect to (40). This efficiency stems from the rapid evaluation of the orthant probability for each sample. In contrast, (40) requires computing orthant probabilities for every sample of the noise covariance matrix mismatch, a process that can be markedly time-intensive. As a result, it has been omitted from the simulation.

B. Non-Null Distribution

We now turn our attention to the accuracy of the derived non-null distributions. For the matched scenario, we compare the probability of detection obtained through Monte Carlo simulations with that predicted by (53) and (57). It should be noted that, in practical applications, the detection probability usually ranges in the vicinity of one. Consequently, the figures use a linear scale for the detection probability, in contrast to the logarithmic scale used for the false alarm rate. Fig. 2(a) shows these values for an experiment with SNR $= [-30, -25, -20]$ dBs (from left to right), which confirms the accuracy of the theoretical results. Fig. 2(b) considers a mismatched scenario, with $\rho = 0.1$, and the angle used in the waveform parameter $\mathbf{W}$ is set as $\phi' = 35°$. Correspondingly, (54) and (58) are used as theoretical counterparts for comparison.. Again, this analysis reveals a high degree of agreement between our theoretical findings and the corresponding simulations.

Additionally, Table I measures the approximation errors for (53) and (54), as well as for the low-SNR approximations, given by (57) and (58). The approximation errors are computed using the Cramér-von Mises goodness-of-fit criterion [49]:

$$\epsilon = \frac{1}{Q}\sum_{i=1}^{K}\left|F(c_i) - \hat{F}(c_i)\right|^2, \tag{65}$$

where $F(c_i)$ is the empirical cumulative distribution function (CDF), $\hat{F}(c_i)$ is the proposed approximation, and $Q = 2000$ is the number of points sampled from the CDFs. Although the low-SNR approximation achieves worse accuracy, it is notably more straightforward to compute, as discussed in Section V-B, offering a practical advantage.

C. Detection Performance

In this section, we evaluate the detection performance of the proposed detector and compare it with the white noise one-bit detector derived in [7]. We consider two key scenarios for this evaluation. In the first scenario, the noise covariance matrix is perfectly known to the receiver, and in the second, there is a mismatch in the noise covariance matrix due to estimation errors. For the mismatched scenario, we consider the model described in (63) with $\rho = [0.01, 0.02, 0.03]$. These values are intentionally selected to simulate moderate and significant noise mismatches, mirroring conditions frequently encountered in real-world applications. Our simulations involve a total of $n = 2000$ samples with the SNR fixed at -31 dBs.

As illustrated in Fig. 3, which shows the ROC curves, the proposed detector shows enhanced detection performance in comparison to the white noise detector across all the considered levels of ρ. This finding highlights the critical role of considering the colored nature of the noise. When $\rho = 0.03$, a modest decline in detection performance is observed relative to scenarios with known noise covariance matrix. However, this decline is quantitatively less pronounced than typically expected for such degree of noise mismatch. This effect is attributed to the concurrent rightward shifts in both the null and non-null distributions caused by noise mismatches, as discussed in previous sections, which mitigates the extent of performance degradation. These results imply that at lower mismatch levels ($\rho = [0.01, 0.02]$), the impact on detection performance is relatively negligible.

Furthermore, according to previous studies on one-bit covariance matrix estimation [38], estimation errors are often even lower than $\rho = 0.02$, thereby supporting the reliability of the proposed detection method in practical scenarios where noise covariance matrix estimation is necessary. Nonetheless,

TABLE I
ERRORS OF DIFFERENT APPROXIMATION METHODS AT DIFFERENT SNRS

	Matched				Mismatched		
SNR	-30 dB	-25 dB	-20 dB		-30 dB	-25 dB	-20 dB
Eq. (53)	1.29×10^{-8}	7.79×10^{-8}	2.28×10^{-8}	Eq. (54)	8.09×10^{-9}	9.41×10^{-9}	5.40×10^{-8}
Eq. (57)	2.66×10^{-8}	3.25×10^{-6}	1.38×10^{-4}	Eq. (58)	4.88×10^{-8}	2.12×10^{-6}	1.16×10^{-4}

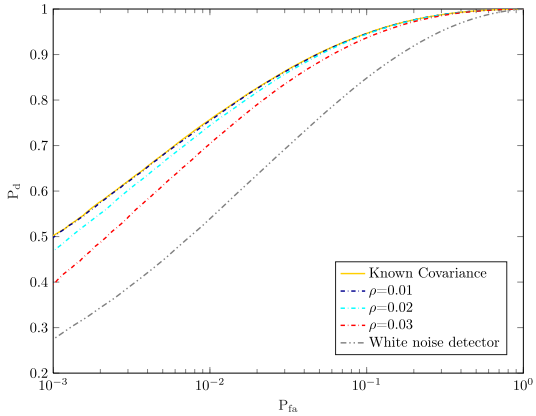


Fig. 3. Detection performance under mismatched noise covariance.

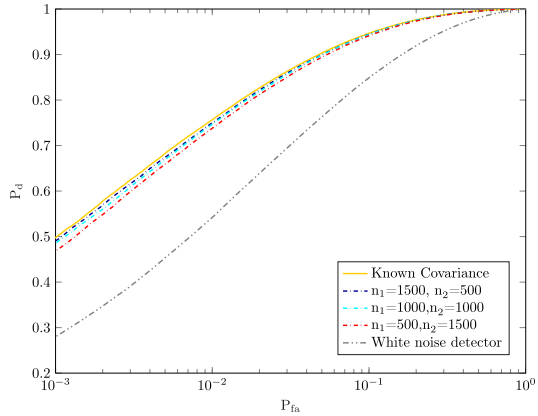


Fig. 4. Impact of training sequence length on detection performance.

the statistical analysis of noise mismatch impact is still crucial to maintain the CFAR property.

We also explore a more realistic scenario where the noise covariance matrix is estimated using noise-only samples. For this purpose, the observation interval is divided into two windows. The first window collects n_1 noise-only samples, which are used for the estimation of the noise covariance matrix using the algorithm presented in [38]. The signal transmission, of length n_2, then occurs in the subsequent phase. Since, contrary to the proposed detector, the white noise detector does not use the noise covariance matrix, we exclude the training phase for this detector, dedicating the entire interval to signal transmission. To maintain a fair comparison, the total transmitted power of the signal is kept constant.

In the experiments with the proposed detector, we evaluate three configurations for the training sequence: $[n_1, n_2] = [500, 1500]$, $[1000, 1000]$, and $[1500, 500]$. These configurations are chosen to assess the impact of varying lengths of the training sequence on the detection performance. As shown in Fig. 4, the proposed detector consistently outperforms the white noise detector. Furthermore, the performance curve for all $[n_1, n_2]$ settings closely mirrors that of the scenario with perfectly known noise covariance. This observation confirms that the current estimation approach yields sufficiently precise covariance matrix information, thereby ensuring the robust performance of our detection method. Hence, the effectiveness of the proposed detector in real-world scenarios, where noise covariance estimation is crucial, is underscored by this result.

D. Performance Loss

In this subsection, we conduct a numerical illustration of the performance loss computed by (60). We consider noise with

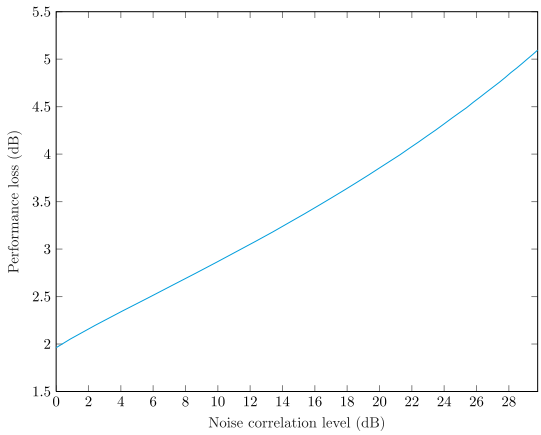


Fig. 5. Performance loss versus noise correlation level.

different correlation levels controlled by α through (62), where α linearly grows from 0 to 1 and only one $\mathbf{H}$ is generated for all trials. In Fig. 5, we plot the performance loss (in dB) as a function of the noise correlation level. To quantify the noise correlation level, we use the determinant of the noise covariance matrix, a well-established measure of correlation in covariance matrices, i.e., NL (dBs) $= -10 \cdot \log_{10}(|\mathbf{\Sigma_N}|)$. The results, depicted in Fig. 5, show that starting from 1.96 dB in the case of uncorrelated noise, the performance loss increases gradually with increasing noise correlation. This finding suggests that in the presence of colored noise, the performance degradation of one-bit systems does increase, but not drastically so.

In addition, while one-bit sampling experiences performance loss under the same system settings, it significantly reduces the total data volume. For scenarios where data volume is

a constraint—such as in small platform radars that need to transmit raw data to a ground station for processing—one-bit sampling can actually achieve a performance gain by allowing for a larger number of samples under the same data volume constraint. For instance, compared with 16-bit systems, one-bit sampling offers approximately a $10 \cdot \log_{10}(16) \approx 12$dB performance gain, which helps offset the previously mentioned performance loss, making it competitive in such scenarios.

VII. Conclusion

In this study, we derived a novel Rao's test for one-bit target detection in MIMO radar systems operating in colored noise environments, generalizing our prior work [7]. The detector is designed as a weighted matched filter, with weights derived from orthant probabilities tied to noise covariance matrix elements. This approach shows enhanced robustness and significant performance gains in colored noise scenarios compared to the white noise detector [7]. Through comprehensive theoretical analysis, we obtained closed-form approximations for both the null and non-null distributions, enabling accurate calculations of false alarm and detection probabilities. We also assessed the impact of noise covariance matrix mismatch, highlighting how it increases the false alarm probability and providing the necessary adjustments to maintain the CFAR property. The analysis of the non-null distribution revealed that performance degradation due to covariance mismatch can be quantified by a decrease in the non-centrality parameter of a chi-squared distribution. Simulation results confirmed the effectiveness and practical applicability of the proposed detector in realistic radar detection scenarios.

Appendix A
Proof of Lemma 1

We consider μ_1 as an example, as any other derivative can be obtained by a simple permutation of the components of the vector. First, we define the PDF of a zero-mean Gaussian

$$f(x_1, \cdots, x_k) = \phi_k(\mathbf{x}; \mathbf{0}_k, \boldsymbol{\Sigma}), \tag{66}$$

and rewrite (6) as

$$P(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \int_{-\infty}^{\mu_1} \int_{-\infty}^{\mu_2} \cdots \int_{-\infty}^{\mu_k} f(x_1, \cdots, x_k) dx_1 \cdots dx_k. \tag{67}$$

Taking into account the definition of the partial derivative of P with respect to μ_1, given by

$$\frac{\partial P}{\partial \mu_1} = \lim_{\Delta\mu_1 \to 0} \frac{P(\boldsymbol{\mu} + \Delta\mu_1 \mathbf{e}_1, \boldsymbol{\Sigma}) - P(\boldsymbol{\mu}, \boldsymbol{\Sigma})}{\Delta\mu_1}, \tag{68}$$

where $\mathbf{e}_1 = [1, 0, \cdots, 0]^T$, we get

$$\begin{aligned} \frac{\partial P}{\partial \mu_1} &= \lim_{\Delta\mu_1 \to 0} \frac{\int_{\mu_1}^{\mu_1+\Delta\mu_1} \int_{-\infty}^{\mu_2} \cdots \int_{-\infty}^{\mu_k} f(x_1, \cdots, x_k) dx_1 \cdots dx_k}{\Delta\mu_1} \\ &= \int_{-\infty}^{\mu_2} \cdots \int_{-\infty}^{\mu_k} f(\mu_1, x_2, \cdots, x_k) dx_2 \cdots dx_k. \end{aligned} \tag{69}$$

Now, using Bayes's theorem to decompose the joint PDF, the derivative becomes

$$\begin{aligned} \frac{\partial P}{\partial \mu_1} &= f_{x_1}(\mu_1) \int_{-\infty}^{\mu_2} \cdots \int_{-\infty}^{\mu_k} f(x_2, \cdots, x_k | x_1 = \mu_1) dx_2 \cdots dx_k \\ &= f_{x_1}(\mu_1) P(\boldsymbol{\omega}(\boldsymbol{\mu}, 1), \mathbf{R}(\boldsymbol{\Sigma}, 1)), \end{aligned} \tag{70}$$

where $\boldsymbol{\omega}(\boldsymbol{\mu}, j)$ and $\mathbf{R}(\boldsymbol{\Sigma}, j)$ are defined in the lemma. The proof follows from

$$f_{x_1}(\mu_1) = \frac{1}{\sqrt{2\pi \boldsymbol{\Sigma}(1,1)}}. \tag{71}$$

Appendix B
Proof of (23)

Firstly, we arrange all possibilities for $\mathbf{y}$ in ascending order of their binary forms:

$$\boldsymbol{\tau}_1 = [-1, \cdots, -1, -1]^T \tag{72a}$$

$$\boldsymbol{\tau}_2 = [-1, \cdots, -1, +1]^T \tag{72b}$$

$$\vdots$$

$$\boldsymbol{\tau}_\kappa = [+1, \cdots, +1, +1]^T \tag{72c}$$

where $\kappa = 2^{2m}$. We define $\boldsymbol{\Gamma}_j = \mathrm{diag}(\boldsymbol{\tau}_j)$, and set

$$O_j = \Pr\{\underline{\mathbf{y}} = \boldsymbol{\tau}_j | \mathcal{H}_0\} = P(\mathbf{0}, \boldsymbol{\Gamma}_j \mathbf{C} \boldsymbol{\Gamma}_j) = P(\mathbf{0}, \mathbf{C}_j), \tag{73}$$

and

$$\mathbf{d}_j = \boldsymbol{\Gamma}_j \left. \frac{\partial P(\boldsymbol{\mu}, \mathbf{C}_j)}{\partial \boldsymbol{\mu}} \right|_{\boldsymbol{\mu} = \mathbf{0}_{2m}}, \tag{74}$$

where the partial derivative of $P(\boldsymbol{\mu}, \mathbf{C}_j)$ is defined analogously to (22). We first consider the first order statistic:

$$\mathbb{E}\left[\left. \frac{\partial \mathcal{L}(\underline{\mathbf{y}}_i; \boldsymbol{\theta})}{\partial a} \right|_{\boldsymbol{\theta} = \boldsymbol{\theta}_0} \right] = \sum_{j=1}^{\kappa} O_j \frac{\mathbf{a}_i^T \mathbf{d}_j}{O_j} = \mathbf{a}_i^T \bar{\mathbf{d}}, \tag{75}$$

where $\bar{\mathbf{d}} = \sum_{j=1}^{\kappa} \mathbf{d}_j$.

We proceed to prove that $\bar{\mathbf{d}} = 0$. Defining j^* as the index such that $\boldsymbol{\Gamma}_{j^*} = -\boldsymbol{\Gamma}_j$, and given $\boldsymbol{\Gamma}_j \mathbf{C} \boldsymbol{\Gamma}_j = \boldsymbol{\Gamma}_{j^*} \mathbf{C} \boldsymbol{\Gamma}_{j^*}$, it is straightforward to prove that $\mathbf{d}_j = -\mathbf{d}_{j^*}$. Since the map between j and j^* is unique, we can conclude that $\bar{\mathbf{d}} = \mathbf{0}$, resulting in:

$$\mathbb{E}\left[\left. \frac{\partial \mathcal{L}(\underline{\mathbf{y}}_i; \boldsymbol{\theta})}{\partial a} \right|_{\boldsymbol{\theta} = \boldsymbol{\theta}_0} \right] = 0. \tag{76}$$

Similarly, the expectation of the derivative of the log-likelihood function with respect to b is 0.

Now, we study the second order statistics. Using (13) and (14), we obtain

$$\mathbb{E}\left[\left. \left(\frac{\partial \mathcal{L}(\underline{\mathbf{y}}_i; \boldsymbol{\theta})}{\partial a} \right)^2 \right|_{\boldsymbol{\theta} = \boldsymbol{\theta}_0} \right] = \sum_{j=1}^{\kappa} O_j \frac{(\mathbf{a}_i^T \mathbf{d}_j)^2}{O_j^2} = \sum_{j=1}^{\kappa} \frac{(\mathbf{a}_i^T \mathbf{d}_j)^2}{O_j}. \tag{77}$$

Taking into account that observations are independent and (76),

this leads to

$$\mathbb{E}\left[\left(\frac{\partial\mathcal{L}(\mathbf{Y};\boldsymbol{\theta})}{\partial a}\right)^2\Bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right]=\sum_{i=1}^{n}\sum_{j=1}^{\kappa}\frac{(\mathbf{a}_i^T\mathbf{d}_j)^2}{O_j}. \tag{78}$$

Similarly, we have

$$\mathbb{E}\left[\left(\frac{\partial\mathcal{L}(\mathbf{Y};\boldsymbol{\theta})}{\partial b}\right)^2\Bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right]=\sum_{i=1}^{n}\sum_{j=1}^{\kappa}\frac{(\mathbf{b}_i^T\mathbf{d}_j)^2}{O_j}, \tag{79}$$

$$\mathbb{E}\left[\frac{\partial\mathcal{L}(\mathbf{Y};\boldsymbol{\theta})}{\partial a}\frac{\partial\mathcal{L}(\mathbf{Y};\boldsymbol{\theta})}{\partial b}\Bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right]=\sum_{i=1}^{n}\sum_{j=1}^{\kappa}\frac{\mathbf{a}_i^T\mathbf{d}_j\mathbf{b}_i^T\mathbf{d}_j}{O_j}. \tag{80}$$

We now explore the "symmetric" properties of the orthant probabilities $\{O_1,\cdots,O_\kappa\}$ and the derivative vectors $\{\mathbf{d}_1,\cdots,\mathbf{d}_\kappa\}$. Define the matrices:

$$\mathbf{T}_1=\begin{bmatrix}\mathbf{0}_m & \mathbf{I}_m\\ -\mathbf{I}_m & \mathbf{0}_m\end{bmatrix},\mathbf{T}_2=\begin{bmatrix}\mathbf{I}_m & \mathbf{0}_m\\ \mathbf{0}_m & -\mathbf{I}_m\end{bmatrix},\mathbf{T}_3=\begin{bmatrix}\mathbf{0}_m & \mathbf{I}_m\\ \mathbf{I}_m & \mathbf{0}_m\end{bmatrix}, \tag{81}$$

and let j_k be the integer such that

$$\boldsymbol{\tau}_{j_k}=\mathbf{T}_1^k\boldsymbol{\tau}_j,\quad k=0,1,2,3. \tag{82}$$

Given the circular nature of the noise, the coherence matrix remains invariant under the transformations:

$$\mathbf{C}=\mathbf{T}_1^T\mathbf{C}\mathbf{T}_1=\mathbf{T}_3\mathbf{T}_2\mathbf{C}\mathbf{T}_2\mathbf{T}_3. \tag{83}$$

Now, we study the relationships between the orthant probabilities corresponding to $\underline{\mathbf{y}}=\boldsymbol{\tau}_j$, $\underline{\mathbf{y}}=\mathbf{T}_2\boldsymbol{\tau}_j$, and $\underline{\mathbf{y}}=\mathbf{T}_3\boldsymbol{\tau}_j$. It is easy to prove that

$$\Pr\{\underline{\mathbf{y}}=\mathbf{T}_2\boldsymbol{\tau}_j;\boldsymbol{\theta}=\boldsymbol{\theta}_0\}=P(\mathbf{0},\mathbf{T}_2\mathbf{C}_j\mathbf{T}_2). \tag{84}$$

The transformation $\underline{\mathbf{y}}=\mathbf{T}_3\boldsymbol{\tau}_j$ corresponds to a reordering of the elements in $\underline{\mathbf{y}}$. Correspondingly, we can reorder the elements of the coherence matrix, which yields:

$$\Pr\{\underline{\mathbf{y}}=\mathbf{T}_3\boldsymbol{\tau}_j;\boldsymbol{\theta}=\boldsymbol{\theta}_0\}=P(\mathbf{0},\mathbf{T}_3\mathbf{C}_j\mathbf{T}_3). \tag{85}$$

As a result,

$$\begin{aligned}O_{j_1}&=\Pr\{\underline{\mathbf{y}}=\mathbf{T}_1\boldsymbol{\tau}_j;\boldsymbol{\theta}=\boldsymbol{\theta}_0\}=\Pr\{\underline{\mathbf{y}}=\mathbf{T}_2\mathbf{T}_3\boldsymbol{\tau}_j;\boldsymbol{\theta}=\boldsymbol{\theta}_0\}\\&=P(\mathbf{0},\mathbf{T}_3\mathbf{T}_2\mathbf{C}_j\mathbf{T}_2\mathbf{T}_3)=P(\mathbf{0},\mathbf{C}_j)=O_j.\end{aligned} \tag{86}$$

For the derivatives $\mathbf{d}_j$ in (74), we use the fact that

$$\mathbf{C}_{j_1}=\boldsymbol{\Gamma}_{j_1}\mathbf{C}\boldsymbol{\Gamma}_{j_1}=\mathbf{T}_3\boldsymbol{\Gamma}_j\mathbf{C}\boldsymbol{\Gamma}_j\mathbf{T}_3=\mathbf{T}_3\mathbf{C}_j\mathbf{T}_3, \tag{87}$$

which correspond to a reordering of the elements in $\mathbf{C}_j$. Accordingly, we can reorder the derivative vector in the same manner, leading to

$$\frac{\partial P(\boldsymbol{\mu},\mathbf{C}_{j_1})}{\partial\boldsymbol{\mu}}\Bigg|_{\boldsymbol{\mu}=\mathbf{0}_{2m}}=\mathbf{T}_3\left.\frac{\partial P(\boldsymbol{\mu},\mathbf{C}_j)}{\partial\boldsymbol{\mu}}\right|_{\boldsymbol{\mu}=\mathbf{0}_{2m}}. \tag{88}$$

In addition, since $\boldsymbol{\Gamma}_{j_1}=\mathbf{T}_1\boldsymbol{\Gamma}_j\mathbf{T}_3$, by combining (74) and (88), we obtain

$$\mathbf{d}_{j_1}=\mathbf{T}_1\mathbf{d}_j. \tag{89}$$

Furthermore, it follows from (14) that

$$\mathbf{b}_i=-\mathbf{T}_1\mathbf{a}_i. \tag{90}$$

Using (86), (89), (90) and the fact that $\mathbf{T}_1^T\mathbf{T}_1=\mathbf{I}_{2m}$, it is easy to prove that:

$$\frac{(\mathbf{a}_i^T\mathbf{d}_j)^2}{O_j}=\frac{(\mathbf{b}_i^T\mathbf{d}_{j_1})^2}{O_{j_1}}. \tag{91}$$

Moreover, since $\boldsymbol{\tau}_j=\mathbf{T}_1^4\boldsymbol{\tau}_j$, we can divide the set $\{1,\cdots,\kappa\}$ into $\kappa/4$ subsets, each with elements $\{j_{k_0},j_{k_1},j_{k_2},j_{k_3}\}$. This division allows us to simplify the summations in (78) and (79) as the sum of $\kappa/4$ individual summations. For each subset, the corresponding summation is

$$\sum_{k=0}^{3}\frac{(\mathbf{a}_i^T\mathbf{d}_{j_k})^2}{O_{j_k}}=\sum_{k=0}^{3}\frac{(\mathbf{b}_i^T\mathbf{d}_{j_k})^2}{O_{j_k}}. \tag{92}$$

Summarizing these subsets yields

$$\sum_{j=1}^{\kappa}\frac{(\mathbf{a}_i^T\mathbf{d}_j)^2}{O_j}=\sum_{j=1}^{\kappa}\frac{(\mathbf{b}_i^T\mathbf{d}_j)^2}{O_j}. \tag{93}$$

For the covariance term in (80), using the symmetries, we can obtain the following relation:

$$\frac{\mathbf{a}_i^T\mathbf{d}_j\mathbf{b}_i^T\mathbf{d}_j}{O_j}=-\frac{\mathbf{a}_i^T\mathbf{d}_{j_1}\mathbf{b}_i^T\mathbf{d}_{j_1}}{O_{j_1}}, \tag{94}$$

which yields

$$\sum_{k=0}^{3}\frac{\mathbf{a}_i^T\mathbf{d}_{j_k}\mathbf{b}_i^T\mathbf{d}_{j_k}}{O_{j_k}}=0. \tag{95}$$

Consequently, (80) becomes

$$\sum_{j=1}^{\kappa}\frac{\mathbf{a}_i^T\mathbf{d}_j\mathbf{b}_i^T\mathbf{d}_j}{O_j}=0, \tag{96}$$

and the FIM simplifies to

$$\mathbf{F}(\boldsymbol{\theta}_0)=\upsilon^2\mathbf{I}_2, \tag{97}$$

where

$$\upsilon^2=\sum_{i=1}^{n}\sum_{j=1}^{\kappa}\frac{(\mathbf{a}_i^T\mathbf{d}_j)^2}{O_j}=\sum_{i=1}^{n}\sum_{j=1}^{\kappa}\frac{(\mathbf{b}_i^T\mathbf{d}_j)^2}{O_j}. \tag{98}$$

Finally, by defining

$$\mathbf{O}=\operatorname{diag}\left(O_1,\cdots,O_\kappa\right), \tag{99}$$

and $\boldsymbol{\Delta}_1$, $\boldsymbol{\Delta}_2\in\mathbb{R}^{n\times\kappa}$ with elements

$$\boldsymbol{\Delta}_1(i,j)=\mathbf{a}_i^T\mathbf{d}_j,\qquad \boldsymbol{\Delta}_2(i,j)=\mathbf{b}_i^T\mathbf{d}_j, \tag{100}$$

υ^2 can be expressed as

$$\upsilon^2=\operatorname{tr}\left(\boldsymbol{\Delta}_1^T\boldsymbol{\Delta}_1\mathbf{O}^{-1}\right)=\operatorname{tr}\left(\boldsymbol{\Delta}_2^T\boldsymbol{\Delta}_2\mathbf{O}^{-1}\right). \tag{101}$$

This completes the proof of (23).

APPENDIX C
PROOF OF THEOREM 1

It is straightforward to obtain the asymptotic distribution of $\mathbf{w}$ and its mean, so this appendix computes the covariance matrix. Thus, we begin by defining

$$O'_j = P(\mathbf{0}, \mathbf{\Gamma}_j \mathbf{C}' \mathbf{\Gamma}_j), \tag{102}$$

where $\mathbf{C}'$ is the coherence matrix of $\mathbf{\Sigma}'_{\underline{\mathbf{x}}}$. Following the same argument as in (77) in Appendix B, we have

$$\mathbb{E}\left[\left(\frac{\partial \mathcal{L}(\underline{\mathbf{y}}_i;\boldsymbol{\theta})}{\partial a}\right)^2\right] = \sum_{j=1}^{\kappa} O'_j \frac{(\mathbf{a}_i^T \mathbf{d}_j)^2}{O_j^2} = \sum_{j=1}^{\kappa} (\mathbf{a}_i^T \mathbf{d}_j)^2 \frac{O'_j}{O_j^2}. \tag{103}$$

Likewise, we compute

$$\mathbb{E}\left[\left(\frac{\partial \mathcal{L}(\mathbf{Y};\boldsymbol{\theta})}{\partial a}\right)^2\right] = \sum_{i=1}^{n}\sum_{j=1}^{\kappa} O'_j \frac{(\mathbf{a}_i^T \mathbf{d}_j)^2}{O_j^2}, \tag{104a}$$

$$\mathbb{E}\left[\left(\frac{\partial \mathcal{L}(\mathbf{Y};\boldsymbol{\theta})}{\partial b}\right)^2\right] = \sum_{i=1}^{n}\sum_{j=1}^{\kappa} O'_j \frac{(\mathbf{b}_i^T \mathbf{d}_j)^2}{O_j^2}, \tag{104b}$$

$$\mathbb{E}\left[\frac{\partial \mathcal{L}(\mathbf{Y};\boldsymbol{\theta})}{\partial a}\frac{\partial \mathcal{L}(\mathbf{Y};\boldsymbol{\theta})}{\partial b}\right] = \sum_{i=1}^{n}\sum_{j=1}^{\kappa} O'_j \frac{\mathbf{a}_i^T \mathbf{d}_j \mathbf{b}_i^T \mathbf{d}_j}{O_j^2}. \tag{104c}$$

Once again, by employing the circularity property, it can be shown that

$$O'_j = O'_{j_k}, \quad k = 0, 1, 2, 3. \tag{105}$$

Combining this property with (94) and (104c), we have

$$\mathbb{E}\left[\frac{\partial \mathcal{L}(\mathbf{Y};\boldsymbol{\theta})}{\partial a}\frac{\partial \mathcal{L}(\mathbf{Y};\boldsymbol{\theta})}{\partial b}\right] = 0. \tag{106}$$

Analogously, we obtain

$$\mathbb{E}\left[\left(\frac{\partial \mathcal{L}(\mathbf{Y};\boldsymbol{\theta})}{\partial a}\right)^2\right] = \mathbb{E}\left[\left(\frac{\partial \mathcal{L}(\mathbf{Y};\boldsymbol{\theta})}{\partial b}\right)^2\right] = \upsilon_1^2, \tag{107}$$

where υ_1^2, i.e., the variance in the mismatched case, is defined in (34). Consequently, the covariance matrix of $\mathbf{w} = [w_1, w_2]^T$ is

$$\mathbf{\Sigma}_w = \frac{\upsilon_1^2}{\upsilon^2}\mathbf{I}_2, \tag{108}$$

which completes the proof of Theorem 1.

APPENDIX D
PROOF OF THEOREM 2

The proof of Theorem 2 is based on Price's Theorem [50], which is summarized in the following lemma. For clarity, we present its simplified version.

Lemma 2: Let $\mathbf{x} = [x_1, \cdots, x_k]^T$ be a k-dimensional vector following a zero-mean Gaussian distribution with unit variances and coherence matrix $\mathbf{C}$. Consider k nonlinear functions $g_i(x), i = 1, 2, \ldots, k$, and define the kth order correlation coefficient of the outputs as

$$R = \mathbb{E}\left[\prod_{i=1}^{k} g_i(x_i)\right]. \tag{109}$$

Then, the partial derivative of R with respect to the elements of the coherence matrix $\mathbf{C}$ is given by

$$\frac{\partial R}{\partial \mathbf{C}(r,s)} = \mathbb{E}\left[g'_r(x_r)\, g'_s(x_s) \prod_{\substack{i=1 \\ i\neq r,s}}^{k} g_i(x_i)\right], \tag{110}$$

where $g'_i(x_i)$ is the derivative of $g_i(x_i)$ with respect to x_i and $r < s$.

To proceed, we define

$$g_i(x) = \begin{cases} 1, & x \geq 0, \\ 0, & x < 0, \end{cases} \tag{111}$$

for $i = 1, \ldots, k$, which yields

$$\mathbb{E}\left[\prod_{i=1}^{k} g_i(x_i)\right] = \Pr\{x_1 > 0, \cdots, x_k > 0\} = P(\mathbf{0}_k, \mathbf{C}). \tag{112}$$

In addition, we have

$$g'_r(x_r) g'_s(x_s) = \delta(x_r)\delta(x_s), \tag{113}$$

where $\delta(\cdot)$ is the Dirac delta function. Therefore, Lemma 2 allows us to write the derivative of the orthant probability as

$$\begin{aligned}\frac{\partial P(\mathbf{0}_k, \mathbf{C})}{\partial \mathbf{C}(r,s)} &= \mathbb{E}\left[\delta(x_r)\delta(x_s) \prod_{\substack{i=1 \\ i\neq r,s}}^{k} g_i(x_i)\right] \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \delta(x_r)\delta(x_s) \prod_{\substack{i=1 \\ i\neq r,s}}^{k} g_i(x_i) f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x}. \end{aligned} \tag{114}$$

Denoting $\bar{\mathbf{x}}$ as the remaining vector after removing x_r and x_s from $\mathbf{x}$, we can write $f_{\mathbf{x}}(\mathbf{x}) = f_{\bar{\mathbf{x}}}(\bar{\mathbf{x}}|x_r, x_s) f_{x_r,x_s}(x_r, x_s)$, which yields

$$\begin{aligned}\frac{\partial P(\mathbf{0}_k, \mathbf{C})}{\partial \mathbf{C}(r,s)} &= f_{x_r,x_s}(0,0) \\ &\times \int_0^{\infty} \cdots \int_0^{\infty} f_{\bar{\mathbf{x}}}(\bar{\mathbf{x}}|0,0) d\bar{\mathbf{x}}. \end{aligned} \tag{115}$$

According to the conditional distribution of a multivariate Gaussian distribution [51], $\bar{\mathbf{x}}$ conditioned on $x_r = x_s = 0$ is a zero mean Gaussian vector with covariance matrix

$$\bar{\mathbf{C}} = [\mathbf{\Theta}(\mathbf{\Theta}(\mathbf{C}^{-1}, r), s-1)]^{-1}. \tag{116}$$

Therefore, we obtain

$$\int_0^{\infty} \cdots \int_0^{\infty} f_{\bar{\mathbf{x}}}(\bar{\mathbf{x}}|0,0) d\bar{\mathbf{x}} = P(\mathbf{0}_{k-2}, \bar{\mathbf{C}}), \tag{117}$$

and taking into account that

$$f_{x_r,x_s}(0,0)=\frac{1}{2\pi\sqrt{1-|\mathbf{C}(r,s)|^2}}, \tag{118}$$

the derivative becomes

$$\frac{\partial P(\mathbf{0}_k,\mathbf{C})}{\partial \mathbf{C}(r,s)}=\frac{P(\mathbf{0}_{k-2},\bar{\mathbf{C}})}{2\pi\sqrt{1-|\mathbf{C}(r,s)|^2}}, \tag{119}$$

which completes the proof of Theorem 2.

Appendix E
Proof of Theorem 3

The definition of the expectation allows us to write

$$\mathbb{E}\left[\frac{\partial\mathcal{L}(\underline{\mathbf{y}}_i;\boldsymbol{\theta})}{\partial a}\right]=\sum_{j=1}^{\kappa}\mathbf{Q}(i,j)\frac{\mathbf{a}_i^T\mathbf{d}_j}{O_j}, \tag{120}$$

and taking into account the independence of the observations, we get

$$\begin{aligned}\boldsymbol{\mu}_w(1)&=\frac{1}{\upsilon}\mathbb{E}\left[\frac{\partial\mathcal{L}(\mathbf{Y};\boldsymbol{\theta})}{\partial a}\right]=\frac{1}{\upsilon}\sum_{i=1}^{n}\sum_{j=1}^{\kappa}\mathbf{Q}(i,j)\frac{\mathbf{a}_i^T\mathbf{d}_j}{O_j}\\&=\frac{1}{\upsilon}\mathrm{tr}(\mathbf{E}_1\mathbf{Q}^T).\end{aligned} \tag{121}$$

Similarly, we have

$$\boldsymbol{\mu}_w(2)=\frac{1}{\upsilon}\mathbb{E}\left[\frac{\partial\mathcal{L}(\mathbf{Y};\boldsymbol{\theta})}{\partial b}\right]=\frac{1}{\upsilon}\mathrm{tr}(\mathbf{E}_2\mathbf{Q}^T). \tag{122}$$

Furthermore, the variance of w_1 is

$$\sigma_1^2=\frac{1}{\upsilon^2}\mathbb{V}\left[\frac{\partial\mathcal{L}(\mathbf{Y};\boldsymbol{\theta})\}}{\partial a}\right]=\frac{1}{\upsilon^2}\sum_{i=1}^{n}\mathbb{V}\left[\frac{\partial\mathcal{L}(\underline{\mathbf{y}}_i;\boldsymbol{\theta})}{\partial a}\right], \tag{123}$$

where each of the n variances on the right-hand-side of the above expression can be obtained as

$$\begin{aligned}\mathbb{V}\left[\frac{\partial\mathcal{L}(\underline{\mathbf{y}}_i;\boldsymbol{\theta})}{\partial a}\right]&=\sum_{j=1}^{\kappa}\mathbf{Q}(i,j)\frac{(\mathbf{a}_i^T\mathbf{d}_j)^2}{O_j^2}\\&\quad-\left(\sum_{j=1}^{\kappa}\mathbf{Q}(i,j)\frac{\mathbf{a}_i^T\mathbf{d}_j}{O_j}\right)^2.\end{aligned} \tag{124}$$

Plugging (124) into (123) yields

$$\begin{aligned}\sigma_1^2&=\frac{1}{\upsilon^2}\sum_{i=1}^{n}\sum_{j=1}^{\kappa}\mathbf{Q}(i,j)\frac{(\mathbf{a}_i^T\mathbf{d}_j)^2}{O_j^2}\\&\quad-\frac{1}{\upsilon^2}\sum_{i=1}^{n}\left(\sum_{j=1}^{\kappa}\mathbf{Q}(i,j)\frac{\mathbf{a}_i^T\mathbf{d}_j}{O_j}\right)^2.\end{aligned} \tag{125}$$

Similarly, we have

$$\begin{aligned}\sigma_2^2&=\frac{1}{\upsilon^2}\sum_{i=1}^{n}\sum_{j=1}^{\kappa}\mathbf{Q}(i,j)\frac{(\mathbf{b}_i^T\mathbf{d}_j)^2}{O_j^2}\\&\quad-\frac{1}{\upsilon^2}\sum_{i=1}^{n}\left(\sum_{j=1}^{\kappa}\mathbf{Q}(i,j)\frac{\mathbf{b}_i^T\mathbf{d}_j}{O_j}\right)^2,\end{aligned} \tag{126}$$

and

$$\begin{aligned}\sigma_{12}&=\frac{1}{\upsilon^2}\sum_{i=1}^{n}\sum_{j=1}^{\kappa}\mathbf{Q}(i,j)\frac{\mathbf{a}_i^T\mathbf{d}_j\mathbf{b}_i^T\mathbf{d}_j}{O_j^2}\\&\quad-\frac{1}{\upsilon^2}\sum_{i=1}^{n}\left(\sum_{j=1}^{\kappa}\mathbf{Q}(i,j)\frac{\mathbf{a}_i^T\mathbf{d}_j}{O_j}\right)\times\left(\sum_{j=1}^{\kappa}\mathbf{Q}(i,j)\frac{\mathbf{b}_i^T\mathbf{d}_j}{O_j}\right).\end{aligned} \tag{127}$$

Appendix F
Proof of (55) and (56)

We begin with the case of matched noise covariance. Given that β is assumed to be of order $\mathcal{O}(n^{-\frac{1}{2}})$, it can be expressed as $\beta=\frac{1}{\sqrt{n}}(a_0+\imath b_0)$. Under such circumstances, a first-order Taylor's expansion of $\mathbf{Q}(i,j)$ at $[a,b]^T=[0,0]^T$ yields

$$\begin{aligned}\mathbf{Q}(i,j)=O_j+\frac{1}{\sqrt{n}}\left[a_0\left.\frac{\partial\mathbf{Q}(i,j)}{\partial a}\right|_{a=0}+b_0\left.\frac{\partial\mathbf{Q}(i,j)}{\partial b}\right|_{b=0}\right]\\+\mathcal{O}(n^{-1}).\end{aligned} \tag{128}$$

Taking into account (17), the derivatives are:

$$\left.\frac{\partial\mathbf{Q}(i,j)}{\partial a}\right|_{a=0}=\mathbf{a}_i^T\mathbf{d}_j=\boldsymbol{\Delta}_1(i,j), \tag{129a}$$

$$\left.\frac{\partial\mathbf{Q}(i,j)}{\partial b}\right|_{b=0}=\mathbf{b}_i^T\mathbf{d}_j=\boldsymbol{\Delta}_2(i,j), \tag{129b}$$

which yields

$$\mathbf{Q}=\mathbf{1}_n\mathbf{o}^T+\frac{1}{\sqrt{n}}(a_0\boldsymbol{\Delta}_1+b_0\boldsymbol{\Delta}_2)+\mathcal{O}(n^{-1}). \tag{130}$$

where $\mathbf{o}=[O_1,\cdots,O_\kappa]^T$. Since $\mathbf{E}_1=\boldsymbol{\Delta}_1\mathbf{O}^{-1}$, it can be shown that

$$\begin{aligned}\mathbf{u}_w(1)&=\frac{1}{\upsilon}\mathrm{tr}(\mathbf{E}_1\mathbf{Q}^T)\\&=\frac{\mathbf{1}_n^T\boldsymbol{\Delta}_1\mathbf{1}_\kappa}{\upsilon}+\frac{1}{\sqrt{n}\upsilon}\left[a_0\mathrm{tr}\left(\boldsymbol{\Delta}_1^T\boldsymbol{\Delta}_1\mathbf{O}^{-1}\right)\right.\\&\quad\left.+b_0\mathrm{tr}\left(\boldsymbol{\Delta}_2^T\boldsymbol{\Delta}_1\mathbf{O}^{-1}\right)\right]+\mathcal{O}(n^{-\frac{1}{2}}).\end{aligned} \tag{131}$$

Using (100), we have

$$\mathbf{1}_n^T\boldsymbol{\Delta}_1\mathbf{1}_\kappa=\sum_{i=1}^{n}\sum_{j=1}^{\kappa}\mathbf{a}_i^T\mathbf{d}_j=\left(\sum_{i=1}^{n}\mathbf{a}_i^T\right)\left(\sum_{j=1}^{\kappa}\mathbf{d}_j\right), \tag{132}$$

and recalling that

$$\sum_{j=1}^{\kappa}\mathbf{d}_j=\bar{\mathbf{d}}=0, \tag{133}$$

we get $\mathbf{1}_n^T\boldsymbol{\Delta}_1\mathbf{1}_\kappa=0$. Moreover, as shown in Appendix B, $\mathrm{tr}(\boldsymbol{\Delta}_1^T\boldsymbol{\Delta}_2\mathbf{O}^{-1})=0$ and $\mathrm{tr}(\boldsymbol{\Delta}_1^T\boldsymbol{\Delta}_1\mathbf{O}^{-1})=\mathrm{tr}(\boldsymbol{\Delta}_2^T\boldsymbol{\Delta}_2\mathbf{O}^{-1})=\upsilon^2$. Thus, we have

$$\mathbf{u}_w(1)=\frac{\upsilon a_0}{\sqrt{n}}+\mathcal{O}(n^{-\frac{1}{2}}), \tag{134}$$

and similarly,

$$\mathbf{u}_w(2)=\frac{\upsilon b_0}{\sqrt{n}}+\mathcal{O}(n^{-\frac{1}{2}}). \tag{135}$$

For the covariance matrix, using (125), the variance σ_1^2 is

$$\sigma_1^2 = \frac{1}{\upsilon^2}\mathbb{E}\left[\left(\frac{\partial \mathcal{L}(\mathbf{Y};\boldsymbol{\theta})}{\partial a}\right)^2\right] - [\mathbf{u}_w(1)]^2 = 1 + \frac{1}{\sqrt{n}\upsilon^2}\sum_{i=1}^{n}\sum_{j=1}^{\kappa}\frac{a_0(\mathbf{a}_i^T\mathbf{d}_j)^3}{O_j^2} + \mathcal{O}(n^{-1}). \tag{136}$$

Similarly, σ_2^2 becomes

$$\sigma_2^2 = 1 + \frac{1}{\sqrt{n}\upsilon^2}\sum_{i=1}^{n}\sum_{j=1}^{\kappa}\frac{b_0(\mathbf{b}_i^T\mathbf{d}_j)^3}{O_j^2} + \mathcal{O}(n^{-1}), \tag{137}$$

and the covariance σ_{12} is

$$\sigma_{12} = \frac{1}{\sqrt{n}\upsilon^2}\sum_{i=1}^{n}\sum_{j=1}^{\kappa}\frac{\mathbf{a}_i^T\mathbf{d}_j\mathbf{b}_i^T\mathbf{d}_j}{O_j^2}(a_0\mathbf{a}_i^T\mathbf{d}_j + b_0\mathbf{b}_i^T\mathbf{d}_j) + \mathcal{O}(n^{-1}). \tag{138}$$

Given that υ^2 is of order n, it follows that

$$\boldsymbol{\Sigma}_w = \mathbf{I}_2 + \mathcal{O}(n^{-\frac{1}{2}}). \tag{139}$$

For the mismatched case where the true noise covariance matrix is $\boldsymbol{\Sigma}'_{\underline{x}}$, we define:

$$\mathbf{a}'_i = \mathrm{Diag}(\boldsymbol{\Sigma}'_{\underline{x}})^{-\frac{1}{2}}\begin{bmatrix}\mathbf{u}_i\\ \mathbf{v}_i\end{bmatrix}, \tag{140a}$$

$$\mathbf{b}'_i = \mathrm{Diag}(\boldsymbol{\Sigma}'_{\underline{x}})^{-\frac{1}{2}}\begin{bmatrix}-\mathbf{v}_i\\ \mathbf{u}_i\end{bmatrix}, \tag{140b}$$

$$\mathbf{d}'_j = \boldsymbol{\Gamma}_j\left.\frac{\partial P(\boldsymbol{\mu}, \boldsymbol{\Gamma}_j\mathbf{C}'\boldsymbol{\Gamma}_j)}{\partial \boldsymbol{\mu}}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}. \tag{140c}$$

Mirroring the arguments in the previous subsection, we obtain:

$$\left.\frac{\partial \mathbf{Q}'(i,j)}{\partial a}\right|_{a=0} = \mathbf{a}_i'^T\mathbf{d}'_j = \boldsymbol{\Delta}'_1(i,j), \tag{141a}$$

$$\left.\frac{\partial \mathbf{Q}'(i,j)}{\partial b}\right|_{b=0} = \mathbf{b}_i'^T\mathbf{d}'_j = \boldsymbol{\Delta}'_2(i,j). \tag{141b}$$

In addition, taking into account that $\mathrm{tr}(\boldsymbol{\Delta}_1^T\boldsymbol{\Delta}'_2\mathbf{O}^{-1})$ and $\mathrm{tr}(\boldsymbol{\Delta}_2^T\boldsymbol{\Delta}'_1\mathbf{O}^{-1})$ are no longer 0 due to the parameter mismatch, the expressions for $\mathbf{u}_w(1)$ and $\mathbf{u}_w(2)$ are:

$$\mathbf{u}_w(1) = \frac{a_0\mathrm{tr}(\boldsymbol{\Delta}_1^T\boldsymbol{\Delta}'_1\mathbf{O}^{-1}) + b_0\mathrm{tr}(\boldsymbol{\Delta}_1^T\boldsymbol{\Delta}'_2\mathbf{O}^{-1})}{\sqrt{n}\upsilon} \tag{142a}$$

$$\mathbf{u}_w(2) = \frac{a_0\mathrm{tr}(\boldsymbol{\Delta}_2^T\boldsymbol{\Delta}'_1\mathbf{O}^{-1}) + b_0\mathrm{tr}(\boldsymbol{\Delta}_2^T\boldsymbol{\Delta}'_2\mathbf{O}^{-1})}{\sqrt{n}\upsilon}. \tag{142b}$$

As proved in Appendix C, the covariance matrix under $\mathcal{H}_0$ for the mismatched case is $\frac{\upsilon_1^2}{\upsilon^2}\mathbf{I}_2$. Following similar arguments, we can conclude that the covariance matrix under low SNR is:

$$\boldsymbol{\Sigma}_w = \frac{\upsilon_1^2}{\upsilon^2}\mathbf{I}_2 + \mathcal{O}(n^{-\frac{1}{2}}). \tag{143}$$

This completes the proof.

REFERENCES

[1] J. Ren and J. Li, "One-bit digital radar," in *Proc. 51st Asilomar Conf. Signals Syst. Comput.*, 2017, pp. 1142–1146.

[2] B. Jin, J. Zhu, Q. Wu, Y. Zhang, and Z. Xu, "One-bit LFMCW radar: Spectrum analysis and target detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 4, pp. 2732–2750, Aug. 2020.

[3] A. Ameri, A. Bose, J. Li, and M. Soltanalian, "One-bit radar processing with time-varying sampling thresholds," *IEEE Trans. Signal Process.*, vol. 67, no. 20, pp. 5297–5308, Oct. 2019.

[4] F. Foroozmehr, M. Modarres-Hashemi, and M. M. Naghsh, "Transmit code and receive filter design for PMCW radars in the presence of one-bit ADC." *IEEE Trans. Aerosp. Electron. Syst.*, vol. 58, no. 4, pp. 3078–3089, Jan. 2022.

[5] L. Ni, D. Zhang, Y. Sun, Y., N. Liu, J. Liang, and Q. Wan, "Detection and localization of one-bit signal in multiple distributed subarray systems." *IEEE Trans. Signal Process.*, vol. 71, pp. 2776–2791, Aug. 2023.

[6] A. Ameri and M. Soltanalian, "One-bit radar processing for moving target detection," in *Proc. IEEE Radar Conf. (RadarConf)*, 2019, pp. 1–6.

[7] Y.-H. Xiao, D. Ramírez, P. J. Schreier, C. Qian, and L. Huang, "One-bit target detection in collocated MIMO radar and performance degradation analysis," *IEEE Trans. Veh. Technol.*, vol. 71, no. 9, pp. 9363–9374, Sep. 2022.

[8] O. Bar-Shalom and A. J. Weiss, "DOA estimation using one-bit quantized measurements," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 38, no. 3, pp. 868–884, Jul. 2002.

[9] L. Feng, L. Huang, Q. Li, Z.-Q. He, and M. Chen, "An off-grid iterative reweighted approach to one-bit direction of arrival estimation." *IEEE Trans. Veh. Technol.*, vol. 72, no. 6, pp. 8134–8139, Jan. 2023.

[10] K. Yu, Y. D. Zhang, M. Bao, Y. Hu, and Z. Wang, "DOA estimation from one-bit compressed array data via joint sparse representation," *IEEE Signal Process. Lett.*, vol. 23, no. 8, pp. 1279–1283, Sep. 2016.

[11] C. L Liu and P. P. Vaidyanathan, "One-bit sparse array DOA estimation," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, New Orleans, LA, USA, Mar. 2017, pp. 3126–3130

[12] X. Shang, H. Zhu, and J. Li, "Range-Doppler imaging via one-bit PMCW radar," in *Proc. IEEE 11th Sensor Array Multichannel Signal Process. Workshop (SAM)*, Hangzhou, China, 2020, pp. 1–5.

[13] X. Shang, J. Li, and P. Stoica, "Weighted SPICE algorithms for range-Doppler imaging using one-bit automotive radar," *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 4, pp. 1041–1054, Jun. 2021.

[14] F. Xi, Y. Xiang, Z. Zhang, S. Chen, and A. Nehorai, "Joint angle and Doppler frequency estimation for MIMO radar with one-bit sampling: A maximum likelihood-based method," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 6, pp. 4734–4748, Dec. 2020.

[15] F. Xi, Y. Xiang, S. Chen, and A. Nehorai, "Gridless parameter estimation for one-bit MIMO radar with time-varying thresholds," *IEEE Trans. Signal Process.*, vol. 68, pp. 1048–1063, 2020.

[16] S. Yang, W. Yi, A. Jakobsson, Y. Wang, and H. Xiao, "Weak signal detection with low-bit quantization in colocated MIMO radar," *IEEE Trans. Signal Process.*, vol. 71, pp. 447–460, 2023.

[17] S. Yang, Y. Lai, A. Jakobsson, and W. Yi, "Hybrid quantized signal detection with a bandwidth-constrained distributed radar system," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 6, pp. 7835–7850, Jul. 2023.

[18] M. Stein, A. Kurzl, A. Mezghani, and J. A. Nossek, "Asymptotic parameter tracking performance with measurement data of 1-bit resolution," *IEEE Trans. Signal Process.*, vol. 63, no. 22, pp. 6086–6095, Nov. 2015.

[19] B. Zhao, L. Huang, and W. Bao, "One-bit SAR imaging based on single-frequency thresholds," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 7017–7032, Sep. 2019.

[20] S. Ge, D. Feng, S. Song, J. Wang, and X. Huang, "Sparse logistic regression-based one-bit SAR imaging," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, pp. 1–15, 2023.

[21] A. Aubry, A. De Maio, L. Lan, and M. Rosamilia, "Adaptive radar detection and bearing estimation in the presence of unknown mutual coupling," *IEEE Trans. Signal Process.*, vol. 71, pp. 1248–1262, 2023.

[22] T. Wang, D. Xu, C. Hao, P. Addabbo, and D. Orlando, "Clutter edges detection algorithms for structured clutter covariance matrices," *IEEE Trans. Process. Lett.*, vol. 29, pp. 642–646, 2022.

[23] S. Han, L. Yan, Y. Zhang, P. Addabbo, C. Hao, and D. Orlando, "Adaptive radar detection and classification algorithms for multiple coherent signals," *IEEE Trans. Signal Process.*, vol. 69, pp. 560–572, 2021.

[24] W. Liu, J. Liu, L. Huang, D. Zou, and Y. Wang, "Rao tests for distributed target detection in interference and noise," *Signal Process.*, vol. 117, pp. 333–342, Dec. 2015.
[25] W. Liu, W. Xie, J. Liu, and Y. Wang, "Adaptive double subspace signal detection in Gaussian background–Part I: Homogeneous environments," *IEEE Trans. Signal Process.*, vol. 62, no. 9, pp. 2345–2357, May 2014.
[26] A. De Maio, "Rao test for adaptive detection in Gaussian interference with unknown covariance matrix," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3577–3584, Jul. 2007.
[27] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 1, pp. 208–216, Jan. 1992.
[28] E. Conte, A. De Maio, and G. Ricci, "GLRT-based adaptive detection algorithms for range-spread targets," *IEEE Trans. Signal Process.*, vol. 49, no. 7, pp. 1336–1348, Jul. 2001.
[29] L. Lan et al., "GLRT-based adaptive target detection in FDA-MIMO radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 57, no. 1, pp. 597–613, Feb. 2021.
[30] E. J. Kelly, "An adaptive detection algorithm," *IEEE Trans. Aerosp. Electron. Syst.,* vol. AES-22, no. 1, pp. 115–127, Mar. 1986.
[31] B. Tang and P. Stoica, "MIMO multifunction RF systems: Detection performance and waveform design," *IEEE Trans. Signal Process.*, vol. 70, pp. 4381–4394, 2022.
[32] W. Liu, Y. Wang, J. Liu, W. Xie, H. Chen, and W. Gu, "Adaptive detection without training data in colocated MIMO radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 51, no. 3, pp. 2469–2479, Jul. 2015.
[33] J. Liu, S. Zhou, W. Liu, J. Zheng, H. Liu, and J. Li, "Tunable adaptive detection in colocated MIMO radar," *IEEE Trans. Signal Process.*, vol. 66, no. 4, pp. 1080–1092, Feb. 2018.
[34] O. Besson, "Rao, Wald, and Gradient tests for adaptive detection of Swerling I targets," *IEEE Trans. Signal Process.*, vol. 71, pp. 3043–3052, 2023.
[35] A. De Maio, S. M. Kay, and A. Farina, "On the invariance, coincidence, and statistical equivalence of the GLRT, Rao test, and Wald test," *IEEE Trans. Signal Process.*, vol. 58, no. 4, pp. 1967–1979, Apr. 2010.
[36] A. P. Shikhaliev and B. Himed, "GLR, Rao, and Wald tests for distributed parametric detection in subspace interference," *IEEE Trans. Signal Process.*, vol. 71, pp. 388–400, 2023.
[37] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Englewood Cliffs, NJ, USA: Prentice Hall, 1998.
[38] Y.-H. Xiao, L. Huang, D. Ramírez, C. Qian, and H. C. So, "Covariance matrix recovery from one-bit data with non-zero quantization thresholds: Algorithm and performance analysis," *IEEE Trans. Signal Process.*, vol. 71, pp. 4060–4076, 2023.
[39] A. Eamaz, F. Yeganegi, and M. Soltanalian, "Covariance recovery for one-bit sampled non-stationary signals with time-varying sampling thresholds," *IEEE Trans. Signal Process.*, vol. 70, pp. 5222–5236, 2022.
[40] C.-L. Liu and Z.-M. Lin, "One-bit autocorrelation estimation with nonzero thresholds," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Toronto, Canada, Jun. 2021, pp. 4520–4524.
[41] S. Dirksen, J. Maly, and H. Rauhut, "Covariance estimation under one-bit quantization," *Ann. Statist.*, vol. 50, no. 6, pp. 3538–3562, Dec. 2022.
[42] P.-W. Wu, L. Huang, D. Ramírez, Y.-H. Xiao, and H. C. So, "One-bit spectrum sensing for cognitive radio," *IEEE Trans. Signal Process.*, vol. 72, pp. 549–564, 2024.
[43] T. Koyama and A. Takemura, "Calculation of orthant probabilities by the holonomic gradient method," *Jpn. J. Ind. Appl. Math.*, vol. 32, no. 1, pp. 187–204, 2015.
[44] T. Miwa, A. Hayter, and S. Kuriki, "The evaluation of general non-centred orthant probabilities." *J. R. Statist. Soc. B*, vol. 65, no.1, pp. 223–234, 2003.
[45] D. Ramírez, I. Santamaría, and L. Scharf, *Coherence: In Signal Processing and Machine Learning.* Berlin, Germany: Springer Nature, 2023.
[46] J. P. Imhof, "Computing the distribution of quadratic forms in normal variables," *Biometrika*, vol. 48, nos. 3–4, pp. 419–426, 1961.
[47] G. Cui, H. Li, and M. Rangaswamy, "MIMO radar waveform design with constant modulus and similarity constraints," *IEEE Trans. Signal Process.*, vol. 62, no. 2, pp. 343–353, Jan. 2014.
[48] O. Aldayel, V. Monga, and M. Rangaswamy, "Successive QCQP refinement for MIMO radar waveform design under practical constraints," *IEEE Trans. Signal Process.*, vol. 64, no. 14, pp. 3760–3774, Jul. 2016.
[49] T. W. Anderson, "On the distribution of the two-sample Cramér-von Mises criterion," *Ann. Math. Statist.*, vol. 33, no. 3, pp. 1148–1159, Sep. 1962.
[50] R. Price, "A useful theorem for nonlinear devices having Gaussian inputs," *IRE Trans. Inf. Theory*, vol. 4, no. 2, pp. 69–72, Jun. 1958.
[51] M. L. Eaton, *Multivariate Statistics: A Vector Space Approach*. Hoboken, NJ, USA: Wiley, 1983.

Yu-Hang Xiao (Member, IEEE) was born in Anhui, China, in 1992. He received the B.E. degree in microelectronics from Harbin Engineering University, China, in 2012, and master's and Ph.D. degrees in information and communication engineering from Harbin Institute of Technology (HIT), China, in 2014 and 2018, respectively. From 2015 to 2017, he was a visiting Ph.D. student with the Department of Electrical and Computer Engineering, McMaster University, Canada. From 2019 to 2021, he was a Postdoctoral Researcher with the Signal and System Theory Group, Paderborn University, Germany. Since 2021, he has been an Assistant Professor with the College of Electronics and Information Engineering, Shenzhen University. His research interests include statistical signal processing, hypotheses testing, and one-bit signal processing.

David Ramírez (Senior Member, IEEE) received the Telecommunication Engineer degree and the Ph.D. degree in electrical engineering from the Universidad de Cantabria, Spain, in 2006 and 2011, respectively. From 2006 to 2011, he was with the Communications Engineering Department, University of Cantabria, Spain. In 2011, he joined as a Research Associate the University of Paderborn, Germany, and later on, he became an Assistant Professor (Akademischer Rat). Currently, he is an Associate Professor with the Universidad Carlos III de Madrid. He has been a Visiting Researcher with the University of Newcastle, Australia and University College London. His research interests include signal processing for wireless communications, statistical signal processing, change-point management, and signal processing over graphs. He has been involved in several national and international research projects on these topics. He was the recipient of the 2012 IEEE Signal Processing Society Young Author Best Paper Award and the 2013 Extraordinary Ph.D. Award of the University of Cantabria. Moreover, he served as an Associate Editor for IEEE TRANSACTIONS ON SIGNAL PROCESSING and the Publications Chair of the 2018. IEEE Workshop on Statistical Signal Processing. He is a member of the IEEE Technical Committee on Signal Processing Theory and Methods and EURASIP Technical Area Committee on Theoretical and Methodological Trends in Signal Processing.

Lei Huang (Senior Member, IEEE) received the B.Sc. and Ph.D. degrees in electronic engineering from Xidian University, Xian, China, in 2000 and 2005, respectively. Currently, he is with the College of Electronics and Information Engineering, Shenzhen University, as a Distinguished Professor, and established the Shenzhen Key Laboratory of Advanced Navigation Technology (ANT) as the Founding Director. Currently, he is the Executive Dean of the College of Electronics and Information Engineering, Shenzhen University. His research interests include spectral estimation, array signal processing, statistical signal processing, and their applications in radar, navigation and wireless communications. In these areas, he has published 100 IEEE journal papers, and undertaken 20 national and provincial key projects, such as the Key Project of the National Natural Science Foundation of China (NSFC) and Joint Project of NSFC-RGC (Hong Kong). He was the winner of the Distinguished Young Scientists of NSFC, Program of the New Century Excellent Talents in University of Ministry of Education (MOE) of China, and Leading Talent of Special Support Program of Guangdong Province. He won the first prize of Technological Invention Award of the Chinese Institute of Electronics (CIE), and the second prize of Natural Science Award in University of MOE of China. He is currently severing as a Senior Area Editor of IEEE TRANSACTIONS ON SIGNAL PROCESSING, and was an Associate Editor of IEEE TRANSACTIONS ON SIGNAL PROCESSING (2015–2019). He was also on the editorial boards of Elsevier-*Digital Signal Processing* (2012–2019) and has been on the editorial boards of IET SIGNAL PROCESSING (2017–present), and an Elected Member of Sensor Array and Multichannel (SAM) Technical Committee of the IEEE Signal Processing Society (2016–present). He was elected an IET Fellow in 2018.

Xiao Peng Li (Member, IEEE) received the B.Eng. degree as an outstanding graduate in electronic science and technology from Yanshan University, Qinhuangdao, China, in 2015, and the M.Sc. (Hons.) degree in electronic information engineering and the Ph.D. degree in electrical engineering from the City University of Hong Kong, Hong Kong SAR, China, in 2018 and 2022, respectively. He was a Research Assistant with the Department of Information Engineering, Shenzhen University, Shenzhen, China from 2018 to 2019, and a Postdoctoral Fellow with the Department of Electrical Engineering, City University of Hong Kong from 2022 to 2023. Currently, he is an Assistant Professor with the College of Electronics and Information Engineering, Shenzhen University. His research interests include robust signal processing, sparse recovery, matrix processing, tensor processing, optimization methods, machine learning, and their applications in various areas of engineering, including target estimation, image recovery, video restoration, hyperspectral unmixing, and stock market analysis.

Hing Cheung So (Fellow, IEEE) was born in Hong Kong. He received the B.Eng. degree from the City University of Hong Kong, in 1990, and the Ph.D. degree from The Chinese University of Hong Kong, in 1995, both in electronic engineering. From 1990 to 1991, he was an Electronic Engineer with the Research and Development Division, Everex Systems Engineering Ltd., Hong Kong. From 1995 to 1996, he was a Postdoctoral Fellow with The Chinese University of Hong Kong. From 1996 to 1999, he was a Research Assistant Professor with the Department of Electronic Engineering, City University of Hong Kong, where he is currently a Professor. His research interests include detection and estimation, fast and adaptive algorithms, multidimensional harmonic retrieval, robust signal processing, source localization, and sparse approximation. He has been on the editorial boards of IEEE SIGNAL PROCESSING MAGAZINE (2014–2017), IEEE TRANSACTIONS ON SIGNAL PROCESSING (2010–2014), *Signal Processing* (2010–), and *Digital Signal Processing* (2011–). He was also the Lead Guest Editor for IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, special issue on "Advances in Time/Frequency Modulated Array Signal Processing" in 2017. In addition, he was an Elected Member in Signal Processing Theory and Methods Technical Committee (2011–2016) of the IEEE Signal Processing Society where he was the Chair in the awards subcommittee (2015–2016).