



Article

Passive MIMO Radar Detection with Unknown Colored Gaussian Noise

Yongjun Liu ^{1,*}, Guisheng Liao ¹, Haichuan Li ¹, Shengqi Zhu ¹, Yachao Li ¹ and Yingzeng Yin ²

¹ National Laboratory of Radar Signal Processing, Xidian University, Xi'an 710071, China; liaogs@xidian.edu.cn (G.L.); hclixd@stu.xidian.edu.cn (H.L.); sqzhu@xidian.edu.cn (S.Z.); ycli@mail.xidian.edu.cn (Y.L.)

² National Key Laboratory of Antennas and Microwave Technology, Xidian University, Xi'an 710071, China; yyzeng@mail.xidian.edu.cn

* Correspondence: yjliu@xidian.edu.cn

Abstract: The target detection of the passive multiple-input multiple-output (MIMO) radar that is comprised of multiple illuminators of opportunity and multiple receivers is investigated in this paper. In the passive MIMO radar, the transmitted signals of illuminators of opportunity are totally unknown, and the received signals are contaminated by the colored Gaussian noise with an unknown covariance matrix. The generalized likelihood ratio test (GLRT) is explored for the passive MIMO radar when the channel coefficients are also unknown, and the closed-form GLRT is derived. Compared with the GLRT with unknown transmitted signals and channel coefficients but a known covariance matrix, the proposed method is applicable for a more practical case when the covariance matrix of colored noise is unknown, although it has higher computational complexity. Moreover, the proposed GLRT can achieve similar performance as the GLRT with the known covariance matrix when the number of training samples is large enough. Finally, the effectiveness of the proposed GLRT is verified by several numerical examples.



Citation: Liu, Y.; Liao, G.; Li, H.; Zhu, S.; Li, Y.; Yin, Y. Passive MIMO Radar Detection with Unknown Colored Gaussian Noise. *Remote Sens.* **2021**, *13*, 2708. <https://doi.org/10.3390/rs13142708>

Academic Editor: Andrzej Stateczny

Received: 18 May 2021

Accepted: 6 July 2021

Published: 9 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: radar detection; passive radar; colored Gaussian noise; generalized likelihood ratio test; multiple-input multiple-output

1. Introduction

In the field of radar, passive radar has been an important research area [1–28]. In passive radar, the transmitters are usually the noncooperative illuminators of opportunity. These illuminators can be digital audio broadcast stations [26], digital television stations [27], or commercial cellular phone base stations [28]. Compared with active radar [29–39], passive radar requires less infrastructure, and is also covert and low-cost, since passive radar does not require to be equipped with its own transmitters. Due to this, passive radar has been attracted great attention.

Compared with active radar that knows the transmitted signal, passive radar does not know the transmitted signal, which complicates target detection for passive radar. For passive radar detection, according to the property of the transmitted signal of the illuminator of opportunity, the unknown transmitted signal is usually modeled as either a stochastic model or a deterministic model. For the stochastic model, the transmitted signal is usually modeled as a stochastic process, while it is deterministic but unknown for the deterministic model. For the deterministic model, the traditional method that utilizes the correlation between the direct-path (transmitter to receiver) and target-path (transmitter to target to receiver) signals is proposed in [1]. However, the performance of the method in [1] will be significantly degraded when the signal-to-noise ratio (SNR) of the received direct-path signal is low. To improve the detection performance of passive radar, the generalized likelihood ratio test (GLRT) [40] is proposed in [2], in which the unknown parameters are replaced with the corresponding maximum likelihood estimates (MLEs).

Moreover, for a bistatic passive polarimetric radar network, the GLRT is explored in [4]. For the stochastic model, the GLRT for a multistatic passive radar is investigated in [5], where the unknown transmitted signal of the illuminator of opportunity is treated as a stochastic process.

In general, the transmit power of the employed noncooperative illuminators of opportunity in passive radar is low, which limits the performance of target detection. Nevertheless, the detection performance of passive radar can be improved by utilizing multiple illuminators and receivers [7], named passive multiple-input multiple-output (MIMO) radar in [23]. In [8], a linear fusion method is proposed for the passive MIMO radar. In this method, for each receiver and transmitter pair of the passive MIMO radar, the cross-correlation between the received target-path and direct-path signals is computed first, and then the weighted combination of those cross correlations forms the final test. The weights are optimized according to a revised deflection coefficient criterion. For passive MIMO radar comprised of multiple illuminators, the GLRT is investigated in [9]. To obtain the GLRT requires obtaining the unknown parameters estimations. In [41], based on the expectation-maximization principle, the authors propose an iterative estimator to estimate the target delay and Doppler for the bistatic radar, where the unknown signal transmitted from the noncooperative illuminators of opportunity is modeled as an autoregressive process. Furthermore, in [42] the authors provide a survey of parameter estimation and target detection with limited data for the phased-array radar, distributed MIMO radar, and passive radar when the parametric auto-regressive model is used for target detection. To derive a closed-form detector for the passive MIMO radar, the Rao test is derived in [10]. To reduce the computational complexity, the authors devise a detector for passive/active MIMO radar in [11]. In [12], the signal transmitted by the noncooperative illuminator of opportunity is assumed to be a discrete-time, zero-mean, second-order cyclostationary signal. The most locally powerful, invariant-test-inspired detector is proposed for the MIMO passive bistatic radar system, which consists of a transmitter, a reference array, and a surveillance array. In [13], the signal sequence of the sources of opportunity is treated as Gaussian random vectors with zero mean and unknown covariance matrices, and the noises in each surveillance channel and reference channel are assumed to be either independent with identical variances or arbitrarily correlated. Under those assumptions, the GLRT is investigated. In [16], both the Bayesian test and GLRT are derived for passive MIMO radar when the noise variance is either known or unknown. In [17], the signal format of the communication signal is exploited to improve the detection performance of the passive MIMO radar when the variance of white noise is known/unknown. Moreover, in [18], the GLRT for the passive MIMO radar is derived when the signal format of the communication signal is known but the covariance matrix of the colored noise is unknown. In [19], it is assumed that the direct-path signals are not available in the passive MIMO radar, and the preamble information of the transmitted signals of noncooperative illuminators of opportunity is known. Moreover, in [19], the noise is assumed to be Gaussian white noise, and the GLRTs are, respectively, investigated when the white noise variance is known or unknown. In practice, the signal format and preamble information of transmitted signals may not be known, and the direct-path signals are usually available. In this case, the GLRT for the passive MIMO radar network as depicted in Figure 1 is explored in [23]. In [23], the target-path and direct-path signals are totally unknown, and the covariance matrix of noise is known or the noise is already whitened. Whitening the noise is non-trivial and would require some sort of training samples. However, in practice, these whitened data are unlikely to be present.

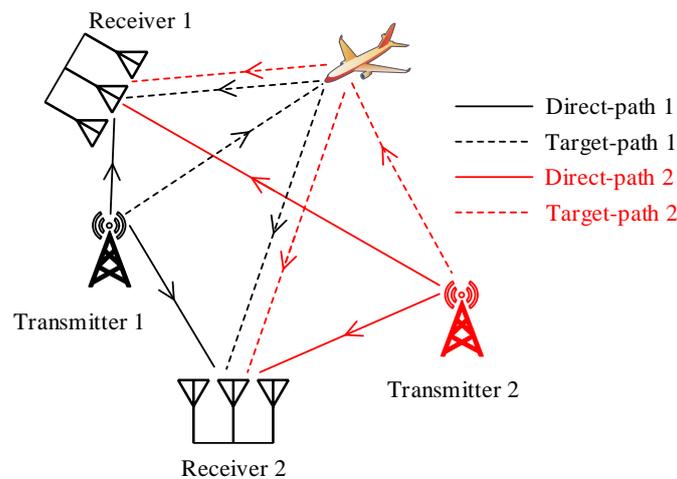


Figure 1. The notional passive MIMO radar consists of two multichannel receiver and two transmitters.

In this paper, we consider a more practical case where the covariance matrix of colored noise in a passive MIMO radar network is unknown, i.e., the colored noise is un-whitened, and the GLRT will be developed. The developed GLRT can simultaneously estimate the statistics of the colored noise while performing the passive radar target detection. This is a very practical problem: passive radar detection in un-whitened colored noise, for which GLRT-based algorithms have not yet been published to date. These algorithms are highly nontrivial and do not follow easily from the work in [23]. Moreover, the developed GLRT approaches the previous GLRT performance with the known covariance matrix when the number of training samples used to estimate the covariance matrix is increased.

The major technical contributions of this paper are summarized as follows. (1) We proposed a target detection method for the passive MIMO radar when the transmitted signal, channel coefficients, and covariance matrix of colored Gaussian noise are unknown. (2) A closed-form GLRT is derived in this paper.

The remainder of this paper is organized as follows. In Section 2, the signal model and assumption are developed. In Section 3, the GLRT for the passive MIMO radar with unknown colored noise is derived. Numerical results are presented in Section 4. Finally, conclusions are drawn in Section 5.

2. Signal Model and Assumption

In this section, the assumptions of a passive MIMO radar network will be made and the signal model will be constructed.

In this paper, we assume that there are N_r receivers and N_t transmitters in the passive MIMO radar network. In each receiver, there is an antenna array which can separate the direct-path and target-path signals by spatial filtering [23]. Suppose that the signals transmitted by different noncooperative illuminators of opportunity or transmitters occupy different frequency bands (In a cellular system, this is usually true for the signals transmitted by neighboring base stations in order to minimize the mutual interference.) [23]. Hence, the received signals due to the transmission from different transmitter can be separated by frequency domain. Moreover, suppose that in the received signals, the clutter has been suppressed by clutter cancellation techniques [20,43], which is the same assumption made in [23].

For each hypothesized range-Doppler cell under test, passive MIMO radar requires to determine whether there is a target or not. For each range-Doppler cell under test, the Doppler shift and range are known, which can be used to compensate the Doppler shift and time delay of the received signal.

Following the compensation in [23], after delay-Doppler compensation, the target-path signal (surveillance channel signal) vector of N time samples at the j th receiver due

to the transmission from the i th transmitter is represented as $\mathbf{y}_{i,j}^s = [y_{i,j}^s(0), y_{i,j}^s(1), \dots, y_{i,j}^s(N-1)]^T$, and the direct-path signal (reference channel signal) vector of N time samples at the j th receiver due to the transmission from the i th transmitter is represented as $\mathbf{y}_{i,j}^r = [y_{i,j}^r(0), y_{i,j}^r(1), \dots, y_{i,j}^r(N-1)]^T$, for $i = 0, 1, \dots, N_t - 1$ and $j = 0, 1, \dots, N_r - 1$, as per the model in [23], where $(\cdot)^T$ denotes the transpose operator.

Similar to [29], we suppose that the training samples, i.e., secondary data, are available, and they have the same probability distribution as the colored Gaussian noise in the reference and surveillance channels (the training samples can be obtained by monitoring the frequency bands, which do not have a communication signal in them). Moreover, suppose that for each receiver and transmitter pair we can obtain K training samples, and the k th training samples obtained in the j th receiver and in the frequency band occupied by the i th transmitter is $\mathbf{y}_{i,j,k}^n = [y_{i,j,k}^n(0), y_{i,j,k}^n(1), \dots, y_{i,j,k}^n(N-1)]^T$, for $i = 0, 1, \dots, N_t - 1$, $j = 0, 1, \dots, N_r - 1$, and $k = 0, 1, \dots, K - 1$.

For the target detection in passive MIMO radar, under hypothesis H_1 , i.e., target present, the compensated signals and secondary data can be represented as

$$\begin{aligned} H_1: \mathbf{y}_{i,j}^s &= \mu_{i,j}^s \mathbf{s}_i + \mathbf{n}_{i,j}^s \\ \mathbf{y}_{i,j}^r &= \mu_{i,j}^r \mathbf{s}_i + \mathbf{n}_{i,j}^r \\ \mathbf{y}_{i,j,k}^n &= \mathbf{n}_{i,j,k}^n. \end{aligned} \quad (1)$$

Under hypothesis H_0 , i.e., target absent, the compensated signals and secondary data can be described as

$$\begin{aligned} H_0: \mathbf{y}_{i,j}^s &= \mathbf{n}_{i,j}^s \\ \mathbf{y}_{i,j}^r &= \mu_{i,j}^r \mathbf{s}_i + \mathbf{n}_{i,j}^r \\ \mathbf{y}_{i,j,k}^n &= \mathbf{n}_{i,j,k}^n. \end{aligned} \quad (2)$$

In (1) and (2), $\mu_{i,j}^r$ and $\mu_{i,j}^s$ respectively denote the complex reference and surveillance channel coefficients from the i th transmitter to the j th receiver, $\mathbf{s}_i = [s_i(0), s_i(1), \dots, s_i(N-1)]^T$ is an $N \times 1$ vector collecting the unknown transmitted signal samples of the i th transmitter, the $N \times 1$ vectors $\mathbf{n}_{i,j}^s = [n_{i,j}^s(0), n_{i,j}^s(1), \dots, n_{i,j}^s(N-1)]^T$ and $\mathbf{n}_{i,j}^r = [n_{i,j}^r(0), n_{i,j}^r(1), \dots, n_{i,j}^r(N-1)]^T$ denote the surveillance and reference channel noise at the j th receiver focused on the transmission from the i th transmitter, respectively, and the $N \times 1$ vector $\mathbf{n}_{i,j,k}^n = [n_{i,j,k}^n(0), n_{i,j,k}^n(1), \dots, n_{i,j,k}^n(N-1)]^T$ denotes the k th training samples obtained at the j th receiver focused on the transmission from the i th transmitter.

Suppose that the colored noise vectors $\mathbf{n}_{i,j}^s$, $\mathbf{n}_{i,j}^r$, and $\mathbf{n}_{i,j,k}^n$ for $i = 0, 1, \dots, N_t - 1$, $j = 0, 1, \dots, N_r - 1$, $k = 0, 1, \dots, K - 1$ are all independent identically distributed, and each of them is the Gaussian random vector with zero mean and covariance matrix $\boldsymbol{\Sigma} \in \mathbb{C}^{N \times N}$, where $\boldsymbol{\Sigma}$ is a positive definite matrix, which is assumed to be unknown in this paper.

3. GLRT for Passive MIMO Radar

In this section, the probability distribution of the received signal is formulated and the GLRT for passive MIMO radar is investigated.

3.1. Probability Distribution of Received Signal

Following the assumption in Section II, under hypothesis H_1 , the joint conditional probability density function (PDF) of the received target-path signal, direct-path signal, and secondary data in (1) is

$$p^1(\mathbf{y}^r, \mathbf{y}^s, \mathbf{y}^n | \boldsymbol{\mu}^r, \boldsymbol{\mu}^s, \mathbf{s}, \boldsymbol{\Sigma}) = \frac{1}{\pi^{Nc_0} |\boldsymbol{\Sigma}|^{c_0}} \exp \left\{ - \sum_{i=0}^{N_t-1} \sum_{j=0}^{N_r-1} \left[\sum_{k=0}^{K-1} (\mathbf{y}_{i,j,k}^n)^H \boldsymbol{\Sigma}^{-1} \mathbf{y}_{i,j,k}^n + (\mathbf{y}_{i,j}^s - \mu_{i,j}^s \mathbf{s}_i)^H \boldsymbol{\Sigma}^{-1} (\mathbf{y}_{i,j}^s - \mu_{i,j}^s \mathbf{s}_i) + (\mathbf{y}_{i,j}^r - \mu_{i,j}^r \mathbf{s}_i)^H \boldsymbol{\Sigma}^{-1} (\mathbf{y}_{i,j}^r - \mu_{i,j}^r \mathbf{s}_i) \right] \right\}, \quad (3)$$

where $c_0 = N_t N_r (K + 2)$, $|\cdot|$ represents the determinant of a matrix, $(\cdot)^H$ denotes Hermitian conjugate operator, $\mathbf{y}^r = \left[(\mathbf{y}_0^r)^T, (\mathbf{y}_1^r)^T, \dots, (\mathbf{y}_{N_t-1}^r)^T \right]^T$ represents the received signal vector that collects all the reference channel samples with $\mathbf{y}_i^r = [\mathbf{y}_{i,0}^r, \mathbf{y}_{i,1}^r, \dots, \mathbf{y}_{i,N_r-1}^r]^T$, $\mathbf{y}^s = \left[(\mathbf{y}_0^s)^T, (\mathbf{y}_1^s)^T, \dots, (\mathbf{y}_{N_t-1}^s)^T \right]^T$ denotes the received signal vector that collects all the surveillance channel samples with $\mathbf{y}_i^s = [\mathbf{y}_{i,0}^s, \mathbf{y}_{i,1}^s, \dots, \mathbf{y}_{i,N_r-1}^s]^T$, $\mathbf{y}^n = \left[(\mathbf{y}_0^n)^T, (\mathbf{y}_1^n)^T, \dots, (\mathbf{y}_{N_t-1}^n)^T \right]^T$ represents the received signal vector that collects all the training samples with $\mathbf{y}_i^n = \left[(\mathbf{y}_{i,0}^n)^T, (\mathbf{y}_{i,1}^n)^T, \dots, (\mathbf{y}_{i,N_r-1}^n)^T \right]^T$ and $\mathbf{y}_{i,j}^n = [\mathbf{y}_{i,j,0}^n, \mathbf{y}_{i,j,1}^n, \dots, \mathbf{y}_{i,j,K-1}^n]^T$, $\boldsymbol{\mu}^r = \left[(\boldsymbol{\mu}_0^r)^T, (\boldsymbol{\mu}_1^r)^T, \dots, (\boldsymbol{\mu}_{N_t-1}^r)^T \right]^T$ denotes the vector that collects all the reference channel coefficients with $\boldsymbol{\mu}_i^r = [\mu_{i,0}^r, \mu_{i,1}^r, \dots, \mu_{i,N_r-1}^r]^T$, $\boldsymbol{\mu}^s = \left[(\boldsymbol{\mu}_0^s)^T, (\boldsymbol{\mu}_1^s)^T, \dots, (\boldsymbol{\mu}_{N_t-1}^s)^T \right]^T$ represents the vector that collects all the surveillance channel coefficients with $\boldsymbol{\mu}_i^s = [\mu_{i,0}^s, \mu_{i,1}^s, \dots, \mu_{i,N_r-1}^s]^T$, and $\mathbf{s} = \left[(\mathbf{s}_0)^T, (\mathbf{s}_1)^T, \dots, (\mathbf{s}_{N_t-1})^T \right]^T$ collects the \mathbf{s}_i for all i .

Under hypothesis H_0 , the joint conditional PDF of the received target-path and direct-path signals, and secondary data in (2) is

$$p^0(\mathbf{y}^r, \mathbf{y}^s, \mathbf{y}^n | \boldsymbol{\mu}^r, \mathbf{s}, \boldsymbol{\Sigma}) = \frac{1}{\pi^{Nc_0} |\boldsymbol{\Sigma}|^{c_0}} \exp \left\{ - \sum_{i=0}^{N_t-1} \sum_{j=0}^{N_r-1} \left[\sum_{k=0}^{K-1} (\mathbf{y}_{i,j,k}^n)^H \boldsymbol{\Sigma}^{-1} \mathbf{y}_{i,j,k}^n + (\mathbf{y}_{i,j}^r - \mu_{i,j}^r \mathbf{s}_i)^H \boldsymbol{\Sigma}^{-1} (\mathbf{y}_{i,j}^r - \mu_{i,j}^r \mathbf{s}_i) + (\mathbf{y}_{i,j}^s)^H \boldsymbol{\Sigma}^{-1} \mathbf{y}_{i,j}^s \right] \right\}. \quad (4)$$

3.2. GLRT Derivation

According to the Neyman–Pearson Lemma [44,45], we can obtain that the likelihood ratio test has the largest detection probability when the probability of false alarm (P_{fa}) is fixed. Nevertheless, for the problem considered in this paper, the likelihood ratio test can not be obtained because the PDF of the received signals depends on the unknown channel coefficients, transmitted signals, and covariance matrix of colored noise. To solve this problem, a typical method is to use the maximum likelihood estimates of unknown parameters as the true values. This method is called GLRT. The GLRT for our problem will be investigated in this subsection.

For the passive MIMO radar, the GLRT can be written as

$$\frac{\max_{\boldsymbol{\mu}^r, \boldsymbol{\mu}^s, \mathbf{s}, \boldsymbol{\Sigma}} p^1(\mathbf{y}^r, \mathbf{y}^s, \mathbf{y}^n | \boldsymbol{\mu}^r, \boldsymbol{\mu}^s, \mathbf{s}, \boldsymbol{\Sigma})}{\max_{\boldsymbol{\mu}^r, \mathbf{s}, \boldsymbol{\Sigma}} p^0(\mathbf{y}^r, \mathbf{y}^s, \mathbf{y}^n | \boldsymbol{\mu}^r, \mathbf{s}, \boldsymbol{\Sigma})} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\geq}} \gamma, \quad (5)$$

where γ is the detection threshold which is usually determined according to a desired P_{fa} in practice.

The GLRT in (5) can also be written as

$$\max_{\mu^r, \mu^s, \mathbf{s}, \Sigma} l^1(\mathbf{y}^r, \mathbf{y}^s, \mathbf{y}^n | \mu^r, \mu^s, \mathbf{s}, \Sigma) - \max_{\mu^r, \mathbf{s}, \Sigma} l^0(\mathbf{y}^r, \mathbf{y}^s, \mathbf{y}^n | \mu^r, \mathbf{s}, \Sigma) \stackrel{\mathcal{H}_1}{\geq} \gamma^0, \tag{6}$$

where $\gamma^0 = \ln \gamma$, $l^0(\mathbf{y}^r, \mathbf{y}^s, \mathbf{y}^n | \mu^r, \mathbf{s}, \Sigma) = \ln p^0(\mathbf{y}^r, \mathbf{y}^s, \mathbf{y}^n | \mu^r, \mathbf{s}, \Sigma)$ and $l^1(\mathbf{y}^r, \mathbf{y}^s, \mathbf{y}^n | \mu^r, \mu^s, \mathbf{s}, \Sigma) = \ln p^1(\mathbf{y}^r, \mathbf{y}^s, \mathbf{y}^n | \mu^r, \mu^s, \mathbf{s}, \Sigma)$ denote the log-likelihood functions under H_0 and H_1 , respectively.

From (3), the log-likelihood function under H_1 is

$$l^1(\mathbf{y}^r, \mathbf{y}^s, \mathbf{y}^n | \mu^r, \mu^s, \mathbf{s}, \Sigma) = c_1 - c_0 \ln |\Sigma| - \sum_{i=0}^{N_t-1} \sum_{j=0}^{N_r-1} \left[\sum_{k=0}^{K-1} (\mathbf{y}_{i,j,k}^n)^H \Sigma^{-1} \mathbf{y}_{i,j,k}^n + (\mathbf{y}_{i,j}^s - \mu_{i,j}^s \mathbf{s}_i)^H \Sigma^{-1} (\mathbf{y}_{i,j}^s - \mu_{i,j}^s \mathbf{s}_i) + (\mathbf{y}_{i,j}^r - \mu_{i,j}^r \mathbf{s}_i)^H \Sigma^{-1} (\mathbf{y}_{i,j}^r - \mu_{i,j}^r \mathbf{s}_i) \right], \tag{7}$$

where $c_1 = -NN_tN_r(K + 2) \ln \pi$.

Similarly, from (4), the log-likelihood function under H_0 is

$$l^0(\mathbf{y}^r, \mathbf{y}^s, \mathbf{y}^n | \mu^r, \mathbf{s}, \Sigma) = c_1 - c_0 \ln |\Sigma| - \sum_{i=0}^{N_t-1} \sum_{j=0}^{N_r-1} \left[\sum_{k=0}^{K-1} (\mathbf{y}_{i,j,k}^n)^H \Sigma^{-1} \mathbf{y}_{i,j,k}^n + (\mathbf{y}_{i,j}^r - \mu_{i,j}^r \mathbf{s}_i)^H \Sigma^{-1} (\mathbf{y}_{i,j}^r - \mu_{i,j}^r \mathbf{s}_i) + (\mathbf{y}_{i,j}^s)^H \Sigma^{-1} \mathbf{y}_{i,j}^s \right]. \tag{8}$$

In order to obtain the optimal solution to the problem in (6), the optimal solution to the optimization problem

$$\max_{\mu^r, \mu^s, \mathbf{s}, \Sigma} l^1(\mathbf{y}^r, \mathbf{y}^s, \mathbf{y}^n | \mu^r, \mu^s, \mathbf{s}, \Sigma) \tag{9}$$

is required to be obtained.

To obtain the optimal solution to the optimization problem in (9), let the derivative of (7) with respect to the covariance matrix Σ be the zero matrix, and then we can achieve

$$\hat{\Sigma} = \frac{1}{c_0} \sum_{i=0}^{N_t-1} \sum_{j=0}^{N_r-1} \left[(\mathbf{y}_{i,j}^r - \mu_{i,j}^r \mathbf{s}_i) (\mathbf{y}_{i,j}^r - \mu_{i,j}^r \mathbf{s}_i)^H + (\mathbf{y}_{i,j}^s - \mu_{i,j}^s \mathbf{s}_i) (\mathbf{y}_{i,j}^s - \mu_{i,j}^s \mathbf{s}_i)^H + \sum_{k=0}^{K-1} \mathbf{y}_{i,j,k}^n (\mathbf{y}_{i,j,k}^n)^H \right], \tag{10}$$

where $\hat{\Sigma}$ satisfies (9) for any $\mathbf{s}_i, \mu_{ij}^r, \mu_{ij}^s, i = 0, 1, \dots, N_t - 1, j = 0, 1, \dots, N_r - 1$.

According to Theorem 3.1.4 in [46], $\hat{\Sigma}$ will be a positive definite matrix with probability 1 if and only if

$$N \leq N_t N_r (K + 2). \tag{11}$$

Suppose that the inequation in (11) is satisfied. Replacing Σ in (7) with $\hat{\Sigma}$ in (10) and utilizing the formula $\mathbf{x}^H \mathbf{A} \mathbf{x} = \text{tr}(\mathbf{A} \mathbf{x} \mathbf{x}^H)$, (7) can be rewritten as

$$\begin{aligned} l^1(\mathbf{y}^r, \mathbf{y}^s, \mathbf{y}^n | \boldsymbol{\mu}^r, \boldsymbol{\mu}^s, \mathbf{s}, \hat{\Sigma}) &= c_1 - c_0 \ln |\hat{\Sigma}| - \sum_{i=0}^{N_t-1} \sum_{j=0}^{N_r-1} \sum_{k=0}^{K-1} \left[\left(\mathbf{y}_{i,j,k}^n \right)^H \hat{\Sigma}^{-1} \mathbf{y}_{i,j,k}^n \right. \\ &\quad \left. + \left(\mathbf{y}_{i,j}^r - \mu_{i,j}^r \mathbf{s}_i \right)^H \hat{\Sigma}^{-1} \left(\mathbf{y}_{i,j}^r - \mu_{i,j}^r \mathbf{s}_i \right) + \left(\mathbf{y}_{i,j}^s - \mu_{i,j}^s \mathbf{s}_i \right)^H \hat{\Sigma}^{-1} \left(\mathbf{y}_{i,j}^s - \mu_{i,j}^s \mathbf{s}_i \right) \right] \\ &= c_1 - c_0 \ln |\hat{\Sigma}| - \text{tr} \left(\hat{\Sigma}^{-1} \sum_{i=0}^{N_t-1} \sum_{j=0}^{N_r-1} \sum_{k=0}^{K-1} \left[\mathbf{y}_{i,j,k}^n \left(\mathbf{y}_{i,j,k}^n \right)^H \right. \right. \\ &\quad \left. \left. + \left(\mathbf{y}_{i,j}^r - \mu_{i,j}^r \mathbf{s}_i \right) \left(\mathbf{y}_{i,j}^r - \mu_{i,j}^r \mathbf{s}_i \right)^H + \left(\mathbf{y}_{i,j}^s - \mu_{i,j}^s \mathbf{s}_i \right) \left(\mathbf{y}_{i,j}^s - \mu_{i,j}^s \mathbf{s}_i \right)^H \right] \right) \end{aligned} \quad (12)$$

$$= c_1 - c_0 \ln |\hat{\Sigma}| - \text{tr}(N_t N_r (K+2) \mathbf{I}_N) \quad (13)$$

$$= c_1 + c_2 - c_0 \ln |\hat{\Sigma}|, \quad (14)$$

where in going from (12) to (13) we use (10), $c_2 = -NN_t N_r (K+2)$, and \mathbf{I}_N represents an $N \times N$ identity matrix.

Let $\mathbf{Y}_{sr} = [\mathbf{Y}_s, \mathbf{Y}_r]$ denote the $N \times 2N_r N_t$ matrix that collects all the received surveillance and reference channel signals, where $\mathbf{Y}_s = [\mathbf{Y}_{0'}^s, \mathbf{Y}_{1'}^s, \dots, \mathbf{Y}_{N_t-1'}^s]$, $\mathbf{Y}_r = [\mathbf{Y}_{0'}^r, \mathbf{Y}_{1'}^r, \dots, \mathbf{Y}_{N_t-1'}^r]$, $\mathbf{Y}_i^s = [\mathbf{y}_{i,0'}^s, \mathbf{y}_{i,1'}^s, \dots, \mathbf{y}_{i,N_r-1'}^s]$, and $\mathbf{Y}_i^r = [\mathbf{y}_{i,0'}^r, \mathbf{y}_{i,1'}^r, \dots, \mathbf{y}_{i,N_r-1'}^r]$, for $i = 0, 1, \dots, N_t - 1$. Let $\mathbf{S} = [\mathbf{s}_0, \mathbf{s}_1, \dots, \mathbf{s}_{N_t-1}]$ denote an $N \times N_t$ matrix that collects all the unknown transmit signals. Let $\mathbf{U}_{sr} = [\mathbf{U}_s, \mathbf{U}_r]$ represent an $N_t \times 2N_r N_r$ matrix which collects all the surveillance and reference channel coefficients, where $\mathbf{U}_s = \text{blkdiag}\{(\boldsymbol{\mu}_0^s)^T, (\boldsymbol{\mu}_1^s)^T, \dots, (\boldsymbol{\mu}_{N_t-1}^s)^T\}$, $\mathbf{U}_r = \text{blkdiag}\{(\boldsymbol{\mu}_0^r)^T, (\boldsymbol{\mu}_1^r)^T, \dots, (\boldsymbol{\mu}_{N_t-1}^r)^T\}$, and $\text{blkdiag}\{\mathbf{A}_0, \mathbf{A}_1, \dots, \mathbf{A}_m\}$ denote the block diagonal matrix with the main diagonal elements being the matrix blocks.

Then (10) can also be rewritten as

$$\hat{\Sigma} = \frac{1}{c_0} \left[(\mathbf{Y}_{sr} - \mathbf{S} \mathbf{U}_{sr}) (\mathbf{Y}_{sr} - \mathbf{S} \mathbf{U}_{sr})^H + \mathbf{R}_n \right], \quad (15)$$

where

$$\mathbf{R}_n = \sum_{i=0}^{N_t-1} \sum_{j=0}^{N_r-1} \sum_{k=0}^{K-1} \mathbf{y}_{i,j,k}^n \left(\mathbf{y}_{i,j,k}^n \right)^H. \quad (16)$$

Using (15), the optimization problem in (9) can be rewritten as

$$\max_{\boldsymbol{\mu}^r, \boldsymbol{\mu}^s, \mathbf{s}} l^1(\mathbf{y}^r, \mathbf{y}^s, \mathbf{y}^n | \boldsymbol{\mu}^r, \boldsymbol{\mu}^s, \mathbf{s}, \hat{\Sigma}) = \max_{\mathbf{S}, \mathbf{U}_{sr}} l^1(\mathbf{y}^r, \mathbf{y}^s, \mathbf{y}^n | \mathbf{S}, \mathbf{U}_{sr}, \hat{\Sigma}) = \max_{\mathbf{S}, \mathbf{U}_{sr}} c_1 + c_2 - c_0 \ln |\hat{\Sigma}|. \quad (17)$$

We can see that the log-likelihood function in (17) can be viewed as either a function of $\boldsymbol{\mu}^r, \boldsymbol{\mu}^s, \mathbf{s}$, and $\hat{\Sigma}$, or a function of $\mathbf{S}, \mathbf{U}_{sr}$, and $\hat{\Sigma}$.

Ignoring the constant terms in (17), the optimization problem in (17) can be rewritten as

$$\min_{\mathbf{S}, \mathbf{U}_{sr}} \ln |\hat{\Sigma}|. \quad (18)$$

Note that

$$|\hat{\Sigma}| = \frac{1}{c_0^N} \left| (\mathbf{Y}_{sr} - \mathbf{S} \mathbf{U}_{sr}) (\mathbf{Y}_{sr} - \mathbf{S} \mathbf{U}_{sr})^H + \mathbf{R}_n \right|. \quad (19)$$

Using Theorem 3.1.4 in [46], we can achieve that the matrix

$$\mathbf{Q} = \mathbf{R}_n + (\mathbf{Y}_{sr} - \mathbf{S} \mathbf{U}_{sr}) (\mathbf{Y}_{sr} - \mathbf{S} \mathbf{U}_{sr})^H = c_0 \hat{\Sigma} \quad (20)$$

which is a positive definite matrix with probability one, if

$$N \leq N_t N_r (K + 2). \quad (21)$$

If the inequation in (21) satisfies, then $\hat{\Sigma}$ in (19) is a positive definite matrix with probability one. In order to minimize $\ln|\hat{\Sigma}|$ with respect to \mathbf{S} for any \mathbf{U}_{sr} as in (18), let the derivative of $\ln|\hat{\Sigma}|$ with respect to \mathbf{S}^* be the zero matrix, i.e.,

$$\frac{\partial \ln|\hat{\Sigma}|}{\partial \mathbf{S}^*} = \text{tr} \left\{ \mathbf{Q}^{-1} \frac{\partial \mathbf{Q}}{\partial \mathbf{S}^*} \right\} = -\mathbf{Q}^{-1} \mathbf{Y}_{sr} \mathbf{U}_{sr}^H + \mathbf{Q}^{-1} \mathbf{S} \mathbf{U}_{sr} \mathbf{U}_{sr}^H = \mathbf{Q}^{-1} (\mathbf{S} \mathbf{U}_{sr} \mathbf{U}_{sr}^H - \mathbf{Y}_{sr} \mathbf{U}_{sr}^H) = \mathbf{0}, \quad (22)$$

where $(\cdot)^*$ represents the conjugate operator.

Since the matrix \mathbf{Q}^{-1} is a nonsingular matrix, from (22) we can obtain the \mathbf{S} that satisfies (18) for any \mathbf{U}_{sr} as

$$\hat{\mathbf{S}} = \mathbf{Y}_{sr} \mathbf{U}_{sr}^H (\mathbf{U}_{sr} \mathbf{U}_{sr}^H)^\dagger, \quad (23)$$

where $(\mathbf{A})^\dagger$ denotes the Moore Penrose inverse of \mathbf{A} .

Using (23), the optimization problem in (18) can be rewritten as

$$\min_{\mathbf{U}_{sr}} \left| \mathbf{R}_n + (\mathbf{Y}_{sr} - \mathbf{Y}_{sr} \tilde{\mathbf{U}}_{sr}) (\mathbf{Y}_{sr} - \mathbf{Y}_{sr} \tilde{\mathbf{U}}_{sr})^H \right|, \quad (24)$$

where $\tilde{\mathbf{U}}_{sr} = \mathbf{U}_{sr}^H (\mathbf{U}_{sr} \mathbf{U}_{sr}^H)^\dagger \mathbf{U}_{sr}$.

Let F be the object function in (24). We can achieve

$$\begin{aligned} F &= \left| \mathbf{R}_n + (\mathbf{Y}_{sr} - \mathbf{Y}_{sr} \tilde{\mathbf{U}}_{sr}) (\mathbf{Y}_{sr} - \mathbf{Y}_{sr} \tilde{\mathbf{U}}_{sr})^H \right| = \left| \mathbf{R}_n + \mathbf{Y}_{sr} (\mathbf{I}_{2N_r N_t} - \tilde{\mathbf{U}}_{sr}) (\mathbf{I}_{2N_r N_t} - \tilde{\mathbf{U}}_{sr})^H \mathbf{Y}_{sr}^H \right| \\ &= \left| \mathbf{R}_n + \mathbf{Y}_{sr} \mathbf{P}_{\mathbf{U}_{sr}^H}^\perp \mathbf{P}_{\mathbf{U}_{sr}^H}^\perp \mathbf{Y}_{sr}^H \right|, \end{aligned} \quad (25)$$

where

$$\mathbf{P}_{\mathbf{U}_{sr}^H}^\perp = \mathbf{I}_{2N_r N_t} - \tilde{\mathbf{U}}_{sr} = \mathbf{I}_{2N_r N_t} - \mathbf{U}_{sr}^H (\mathbf{U}_{sr} \mathbf{U}_{sr}^H)^\dagger \mathbf{U}_{sr} \quad (26)$$

is the orthogonal projection matrix of \mathbf{U}_{sr}^H [47].

Note that $\mathbf{P}_{\mathbf{U}_{sr}^H}^\perp = \mathbf{P}_{\mathbf{U}_{sr}^H}^{\perp H}$, and $\mathbf{P}_{\mathbf{U}_{sr}^H}^\perp \mathbf{P}_{\mathbf{U}_{sr}^H}^\perp = \mathbf{P}_{\mathbf{U}_{sr}^H}^\perp$. Hence, (25) can be simplified as

$$F = \left| \mathbf{R}_n + \mathbf{Y}_{sr} \mathbf{P}_{\mathbf{U}_{sr}^H}^\perp \mathbf{P}_{\mathbf{U}_{sr}^H}^\perp \mathbf{Y}_{sr}^H \right| = \left| \mathbf{R}_n + \mathbf{Y}_{sr} \mathbf{P}_{\mathbf{U}_{sr}^H}^\perp \mathbf{Y}_{sr}^H \right|. \quad (27)$$

Since $\mathbf{P}_{\mathbf{U}_{sr}^H}^\perp$ is orthogonal projection matrix, the eigenvalues of $\mathbf{P}_{\mathbf{U}_{sr}^H}^\perp$ are either 1 or 0. Using Theorem 21.5.7 in [47], we can obtain that $\mathbf{P}_{\mathbf{U}_{sr}^H}^\perp$ is orthogonally diagonalizable, i.e.,

$$\mathbf{P}_{\mathbf{U}_{sr}^H}^\perp = \mathbf{V} \mathbf{D} \mathbf{V}^H = [\mathbf{V}_m, \mathbf{V}_{2N_r N_t - m}] \begin{bmatrix} \mathbf{I}_m & \mathbf{0}_{m \times (2N_r N_t - m)} \\ \mathbf{0}_{(2N_r N_t - m) \times m} & \mathbf{0}_{2N_r N_t - m} \end{bmatrix} \begin{bmatrix} \mathbf{V}_m^H \\ \mathbf{V}_{2N_r N_t - m}^H \end{bmatrix} = \mathbf{V}_m \mathbf{V}_m^H, \quad (28)$$

where $\mathbf{D} = \text{diag}\{\mathbf{I}_m, \mathbf{0}_{2N_r N_t - m}\}$ with $\mathbf{0}_n$ denotes an $n \times n$ zero matrix, \mathbf{V} is an orthogonal matrix, i.e., $\mathbf{V} \mathbf{V}^H = \mathbf{V}^H \mathbf{V} = \mathbf{I}_{2N_r N_t}$, \mathbf{V}_m collects the first m column vectors of \mathbf{V} , $\mathbf{V}_{2N_r N_t - m}$ collects the last $2N_r N_t - m$ column vectors of \mathbf{V} , $\mathbf{V}_m^H \mathbf{V}_m = \mathbf{I}_m$, $\mathbf{V}_{2N_r N_t - m}^H \mathbf{V}_{2N_r N_t - m} = \mathbf{0}_{(2N_r N_t - m) \times (2N_r N_t - m)}$ with $\mathbf{0}_{n \times m}$ denotes an $n \times m$ zero matrix, and $m = \rho(\mathbf{P}_{\mathbf{U}_{sr}^H}^\perp)$ with $\rho(\cdot)$ represents the rank of a matrix.

Plugging (28) into (27) yields

$$F = \left| \mathbf{R}_n + \mathbf{Y}_{sr} \mathbf{V}_m \mathbf{V}_m^H \mathbf{Y}_{sr}^H \right|. \quad (29)$$

According to Theorem 3.1.4 in [46], if

$$N_t N_r K \geq N, \quad (30)$$

the matrix \mathbf{R}_n will be positively definite with probability one. We assume that (30) is satisfied. By using $|\mathbf{A}_{n \times n} \mathbf{B}_{n \times n}| = |\mathbf{A}_{n \times n}| |\mathbf{B}_{n \times n}|$ and $|\mathbf{I}_m + \mathbf{A}_{m \times n} \mathbf{B}_{n \times m}| = |\mathbf{I}_n + \mathbf{B}_{n \times m} \mathbf{A}_{m \times n}|$, (29) can be rewritten as

$$F = |\mathbf{R}_n| \left| \mathbf{I}_N + \mathbf{R}_n^{-1} \mathbf{Y}_{sr} \mathbf{V}_m \mathbf{V}_m^H \mathbf{Y}_{sr}^H \right| = |\mathbf{R}_n| \left| \mathbf{I}_r + \mathbf{V}_m^H \mathbf{Y}_{sr}^H \mathbf{R}_n^{-1} \mathbf{Y}_{sr} \mathbf{V}_m \right|. \quad (31)$$

Since $\mathbf{V}_m^H \mathbf{V}_m = \mathbf{I}_m$, (31) can be simplified as

$$\begin{aligned} F &= |\mathbf{R}_n| \left| \mathbf{V}_m^H \mathbf{I}_{2N_r N_t} \mathbf{V}_m + \mathbf{V}_m^H \mathbf{Y}_{sr}^H \mathbf{R}_n^{-1} \mathbf{Y}_{sr} \mathbf{V}_m \right| \\ &= |\mathbf{R}_n| \left| \mathbf{V}_m^H \left(\mathbf{I}_{2N_r N_t} + \mathbf{Y}_{sr}^H \mathbf{R}_n^{-1} \mathbf{Y}_{sr} \right) \mathbf{V}_m \right| \\ &= |\mathbf{R}_n| \left| \mathbf{V}_m^H \mathbf{\Psi} \mathbf{V}_m \right|, \end{aligned} \quad (32)$$

where

$$\mathbf{\Psi} = \mathbf{I}_{2N_r N_t} + \mathbf{Y}_{sr}^H \mathbf{R}_n^{-1} \mathbf{Y}_{sr} \quad (33)$$

is a $2N_r N_t \times 2N_r N_t$ positive definite matrix.

From (32), the optimization problem in (24) can be rewritten as the following optimization problem

$$\begin{aligned} \min_{\mathbf{V}_m} \quad & F = |\mathbf{R}_n| \left| \mathbf{V}_m^H \mathbf{\Psi} \mathbf{V}_m \right| \\ \text{subject to} \quad & \mathbf{V}_m^H \mathbf{V}_m = \mathbf{I}_m. \end{aligned} \quad (34)$$

To solve the optimization problem in (34), we introduce Theorem 1 (the proof is given in Appendix A).

Theorem 1. Let \mathbf{A} be an $n \times n$ positive Hermitian matrix, and $\mathbf{X}_m = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m]$ be an $n \times m$ matrix with $m \leq n$. The solution to

$$\begin{aligned} \arg \min_{\mathbf{X}_m} \quad & |\mathbf{X}_m^H \mathbf{A} \mathbf{X}_m| \\ \text{subject to} \quad & \mathbf{X}_m^H \mathbf{X}_m = \mathbf{I}_m \end{aligned} \quad (35)$$

is $\hat{\mathbf{X}}_m = [\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \dots, \hat{\mathbf{x}}_m]$, which collects the m eigenvectors corresponding to the m smallest eigenvalues of \mathbf{A} , i.e., $\hat{\mathbf{x}}_i \in \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ and $\hat{\mathbf{x}}_i \neq \hat{\mathbf{x}}_j$, for $i \neq j$, where $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ are the m eigenvectors corresponding to the m smallest eigenvalues of \mathbf{A} . Moreover, the optimum value of the objective function in (35) is the product of m smallest eigenvalues of \mathbf{A} .

According to Theorem 1, the optimal solution to (34) is $\hat{\mathbf{V}}_m = [\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \dots, \boldsymbol{\psi}_m]$ and the optimal value is $\hat{F}(m) = |\mathbf{R}_n| \prod_{i=1}^m \lambda_i$, where λ_i , for $i = 1, 2, \dots, 2N_r N_t$, is the eigenvalue of $\mathbf{\Psi}$, $\boldsymbol{\psi}_i$ is the corresponding eigenvector, and $1 \leq \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_{2N_r N_t}$. Note that the optimal value $\hat{F}(m)$ depends on the rank $m = \rho(\mathbf{P}_{\mathbf{U}_{sr}^H}^\perp)$. Obviously, the smaller the m , the smaller the optimal value $\hat{F}(m)$ will be.

Since $\mathbf{P}_{\mathbf{U}_{sr}^H} = \tilde{\mathbf{U}}_{sr} = \mathbf{U}_{sr}^H (\mathbf{U}_{sr} \mathbf{U}_{sr}^H)^\dagger \mathbf{U}_{sr}$ is the projection matrix of \mathbf{U}_{sr}^H and the rank of \mathbf{U}_{sr}^H satisfies $\rho(\mathbf{U}_{sr}^H) \leq N_t$, the rank of $\mathbf{P}_{\mathbf{U}_{sr}^H}$ is $\rho(\mathbf{P}_{\mathbf{U}_{sr}^H}) \leq N_t$. Since $\mathbf{P}_{\mathbf{U}_{sr}^H}^\perp = \mathbf{I}_{2N_r N_t} - \mathbf{P}_{\mathbf{U}_{sr}^H}$, the rank of $\mathbf{P}_{\mathbf{U}_{sr}^H}^\perp$ satisfies

$$m = \rho(\mathbf{P}_{\mathbf{U}_{sr}^H}^\perp) = 2N_r N_t - \rho(\mathbf{P}_{\mathbf{U}_{sr}^H}) \geq 2N_r N_t - N_t. \quad (36)$$

Using (36) and Theorem 1, we can obtain that the optimal solution to the optimization problem in (34) is

$$\hat{\mathbf{V}}_{2N_r N_t - N_t} = \mathbf{\Psi}_{2N_r N_t - N_t} = [\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \dots, \boldsymbol{\psi}_{2N_r N_t - N_t}] \tag{37}$$

and the corresponding optimal value is

$$\hat{F}(2N_r N_t - N_t) = |\mathbf{R}_n| \prod_{i=1}^{2N_r N_t - N_t} \lambda_i. \tag{38}$$

The optimal solution $\hat{\mathbf{U}}_{sr}$ to the optimization problem in (24) is (the proof is shown in Appendix B)

$$\hat{\mathbf{U}}_{sr} = \mathbf{\Gamma}^H \mathbf{\Omega} \mathbf{\Psi}_{N_t}^H, \tag{39}$$

where $\mathbf{\Psi}_{N_t} = [\boldsymbol{\psi}_{2N_r N_t - N_t + 1}, \boldsymbol{\psi}_{2N_r N_t - N_t + 2}, \dots, \boldsymbol{\psi}_{2N_r N_t}]$, $\mathbf{\Omega}$ is any $N_t \times N_t$ invertible diagonal matrix, and $\mathbf{\Gamma}$ is any $N_t \times N_t$ orthogonal matrix, i.e., $\mathbf{\Gamma}^H \mathbf{\Gamma} = \mathbf{\Gamma} \mathbf{\Gamma}^H = \mathbf{I}_{N_t}$.

Using (9), (17), and (38), we can obtain

$$\begin{aligned} & \max_{\boldsymbol{\mu}^r, \boldsymbol{\mu}^s, \mathbf{s}, \boldsymbol{\Sigma}} l^1(\mathbf{y}^r, \mathbf{y}^s, \mathbf{y}^n | \boldsymbol{\mu}^r, \boldsymbol{\mu}^s, \mathbf{s}, \boldsymbol{\Sigma}) = \max_{\boldsymbol{\mu}^r, \boldsymbol{\mu}^s, \mathbf{s}} l^1(\mathbf{y}^r, \mathbf{y}^s, \mathbf{y}^n | \boldsymbol{\mu}^r, \boldsymbol{\mu}^s, \mathbf{s}, \hat{\boldsymbol{\Sigma}}) \\ & = \max_{\mathbf{S}, \mathbf{U}_{sr}} l^1(\mathbf{y}^r, \mathbf{y}^s, \mathbf{y}^n | \mathbf{S}, \mathbf{U}_{sr}, \hat{\boldsymbol{\Sigma}}) = \max_{\mathbf{S}, \mathbf{U}_{sr}} c_1 + c_2 - c_0 \ln |\hat{\boldsymbol{\Sigma}}| \\ & = c_1 + c_2 - c_0 \ln \frac{\hat{F}(2N_r N_t - N_t)}{c_0^N} = c_1 + c_2 + c_3 - c_0 \ln \left(|\mathbf{R}_n| \prod_{i=1}^{2N_r N_t - N_t} \lambda_i \right), \end{aligned} \tag{40}$$

where $c_3 = NN_t N_r (K + 2) \ln [N_t N_r (K + 2)]$.

Similar to the previous processing, under the hypothesis H_0 , the following optimization problem is required to be solved

$$\max_{\boldsymbol{\mu}^r, \mathbf{s}, \boldsymbol{\Sigma}} l^0(\mathbf{y}^r, \mathbf{y}^s, \mathbf{y}^n | \boldsymbol{\mu}^r, \mathbf{s}, \boldsymbol{\Sigma}). \tag{41}$$

Using the previous estimation method, we can achieve that

$$\begin{aligned} \hat{\boldsymbol{\Sigma}} &= \frac{1}{c_0} \sum_{i=0}^{N_t-1} \sum_{j=0}^{N_r-1} \left[(\mathbf{y}_{i,j}^r - \boldsymbol{\mu}_{i,j}^r \mathbf{s}_i) (\mathbf{y}_{i,j}^r - \boldsymbol{\mu}_{i,j}^r \mathbf{s}_i)^H + \mathbf{y}_{i,j}^s (\mathbf{y}_{i,j}^s)^H + \sum_{k=0}^{K-1} \mathbf{y}_{i,j,k}^n (\mathbf{y}_{i,j,k}^n)^H \right] \\ &= \frac{1}{c_0} [(\mathbf{Y}_r - \mathbf{S} \mathbf{U}_r) (\mathbf{Y}_r - \mathbf{S} \mathbf{U}_r)^H + \mathbf{R}_{sn}], \end{aligned} \tag{42}$$

is the optimal solution to the optimization problem in (41) for any $\mathbf{s}_i, \boldsymbol{\mu}_{i,j}^r, i = 0, 1, \dots, N_t - 1, j = 0, 1, \dots, N_r - 1$, where

$$\mathbf{R}_{sn} = \sum_{i=0}^{N_t-1} \sum_{j=0}^{N_r-1} \mathbf{y}_{i,j}^s (\mathbf{y}_{i,j}^s)^H + \sum_{i=0}^{N_t-1} \sum_{j=0}^{N_r-1} \sum_{k=0}^{K-1} \mathbf{y}_{i,j,k}^n (\mathbf{y}_{i,j,k}^n)^H. \tag{43}$$

Substituting (42) into (8) with $\boldsymbol{\Sigma} = \hat{\boldsymbol{\Sigma}}$ produces

$$l^0(\mathbf{y}^r, \mathbf{y}^s, \mathbf{y}^n | \boldsymbol{\mu}^r, \mathbf{s}, \hat{\boldsymbol{\Sigma}}) = c_1 + c_2 - c_0 \ln |\hat{\boldsymbol{\Sigma}}|. \tag{44}$$

From (44), we can achieve that the optimization problem in (41) can be reformulated as

$$\max_{\boldsymbol{\mu}^r, \mathbf{s}} l^0(\mathbf{y}^r, \mathbf{y}^s, \mathbf{y}^n | \boldsymbol{\mu}^r, \mathbf{s}, \hat{\boldsymbol{\Sigma}}) = \max_{\mathbf{S}, \mathbf{U}_r} l^0(\mathbf{y}^r, \mathbf{y}^s, \mathbf{y}^n | \mathbf{S}, \mathbf{U}_r, \hat{\boldsymbol{\Sigma}}) = \max_{\mathbf{S}, \mathbf{U}_r} c_1 + c_2 - c_0 \ln |\hat{\boldsymbol{\Sigma}}|. \tag{45}$$

In (45), we can see that the log-likelihood function $l^0(\cdot)$ can be viewed as either a function of μ^r, \mathbf{s} , and $\hat{\Sigma}$, or a function of \mathbf{S}, \mathbf{U}_r , and $\hat{\Sigma}$.

Ignoring the constant terms in (45), the optimization problem in (45) can be rewritten as

$$\min_{\mathbf{S}, \mathbf{U}_r} \ln|\hat{\Sigma}|. \tag{46}$$

Using Theorem 3.1.4 in [46], we can achieve that the matrix $\hat{\Sigma}$ is a positive definite matrix with probability one, if

$$N \leq N_t N_r (K + 2). \tag{47}$$

Assume that (47) is true. In order to minimize $\ln|\hat{\Sigma}|$ with respect to \mathbf{S} for any \mathbf{U}_r as in (46), let the derivative of $\ln|\hat{\Sigma}|$ with respect to \mathbf{S} be the zero matrix, and we can achieve that

$$\hat{\mathbf{S}} = \mathbf{Y}_r \mathbf{U}_r^H (\mathbf{U}_r \mathbf{U}_r^H)^\dagger. \tag{48}$$

Using (48), the optimization problem in (46) can be rewritten as

$$\min_{\mathbf{U}_r} \left| (\mathbf{Y}_r - \mathbf{Y}_r \tilde{\mathbf{U}}_r) (\mathbf{Y}_r - \mathbf{Y}_r \tilde{\mathbf{U}}_r)^H + \mathbf{R}_{sn} \right|, \tag{49}$$

where $\tilde{\mathbf{U}}_r = \mathbf{U}_r^H (\mathbf{U}_r \mathbf{U}_r^H)^\dagger \mathbf{U}_r$.

Similar to previous processing, we can achieve that the optimal solution $\hat{\mathbf{U}}_r$ to the problem in (49) is

$$\hat{\mathbf{U}}_r = \mathbf{\Gamma}'^H \mathbf{\Omega}' \mathbf{\Psi}'_{N_t}^H, \tag{50}$$

where $\mathbf{\Omega}'$ is any $N_t \times N_t$ invertible diagonal matrix, $\mathbf{\Gamma}'$ is any $N_t \times N_t$ orthogonal matrix, and $\mathbf{\Psi}'_{N_t} = [\psi'_{N_r N_t - N_t + 1}, \psi'_{N_r N_t - N_t + 2}, \dots, \psi'_{N_r N_t}]$, in which ψ'_i , for $i = 1, 2, \dots, N_r N_t$, is the eigenvector corresponding to the eigenvalue λ'_i of

$$\mathbf{\Psi}' = \mathbf{I}_{N_r N_t} + \mathbf{Y}_r^H \mathbf{R}_{sn}^{-1} \mathbf{Y}_r \tag{51}$$

that is an $N_r N_t \times N_r N_t$ positive definite matrix, and $1 \leq \lambda'_1 \leq \lambda'_2 \leq \dots \leq \lambda'_{N_r N_t}$.

The optimal value of (49) is

$$\hat{F}'(N_r N_t - N_t) = |\mathbf{R}_{sn}| \prod_{i=1}^{N_r N_t - N_t} \lambda'_i. \tag{52}$$

Using (41), (45), and (52), we can obtain

$$\begin{aligned} & \max_{\mu^r, \mathbf{s}, \Sigma} l^0(\mathbf{y}^r, \mathbf{y}^s, \mathbf{y}^n | \mu^r, \mathbf{s}, \Sigma) = \max_{\mu^r, \mathbf{s}} l^0(\mathbf{y}^r, \mathbf{y}^s, \mathbf{y}^n | \mu^r, \mathbf{s}, \hat{\Sigma}) = \max_{\mathbf{S}, \mathbf{U}_r} l^0(\mathbf{y}^r, \mathbf{y}^s, \mathbf{y}^n | \mathbf{S}, \mathbf{U}_r, \hat{\Sigma}) \\ & = \max_{\mathbf{S}, \mathbf{U}_r} c_1 + c_2 - c_0 \ln|\hat{\Sigma}| = c_1 + c_2 - c_0 \ln \frac{\hat{F}'(N_r N_t - N_t)}{c_0^N} \\ & = c_1 + c_2 + c_3 - c_0 \ln \left(|\mathbf{R}_{sn}| \prod_{i=1}^{N_r N_t - N_t} \lambda'_i \right). \end{aligned} \tag{53}$$

Substituting (40) and (53) into (6) yields

$$c_0 \ln \frac{|\mathbf{R}_{sn}| \prod_{i=1}^{N_r N_t - N_t} \lambda'_i \mathcal{H}_1}{|\mathbf{R}_n| \prod_{i=1}^{2N_r N_t - N_t} \lambda_i \mathcal{H}_0} \geq \gamma^0, \tag{54}$$

i.e., the GLRT is equivalent to

$$\frac{|\mathbf{R}_{sn}| \prod_{i=1}^{N_r N_t - N_t} \lambda_i'}{\lambda_i} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \tilde{\gamma}, \quad (55)$$

where $\tilde{\gamma} = \exp\{\gamma^0/c_0\}$.

3.3. Performance Analysis and Discussion

The computational complexity of the proposed GLRT in (55) is approximately $\mathcal{O}(3N^3 + 3N^2 N_r N_t + 5N N_r^2 N_t^2 + N N_r N_t + 9N_r^3 N_t^3 + 2N_r N_t K)$. The computational complexity of the GLRT in [23] is approximately $\mathcal{O}(N N_r^2 N_t + 9N_r^3 N_t^3)$. Obviously, the computational complexity of the proposed method is higher than the GLRT in [23], since the proposed method does not know the covariance matrix of colored noise, which complicates the target detection and increases the computational complexity.

In practice, the proposed GLRT will determine whether there is a target or not for each range and Doppler cell. Hence, for multiple targets that are in different range or Doppler cells, the proposed GLRT is applicable. That is to say, the proposed GLRT work for multi-target detection.

4. Simulation

In this section, several simulations results are provided to show the performance of the proposed GLRT. In the following simulations, when the channel coefficients, transmitted signal, and covariance matrix of colored noise are unknown, the proposed GLRT is denoted by 'UK CovM'. When the channel coefficients and transmitted signals are unknown whereas the covariance matrix of colored noise is known, the GLRT in [23] is denoted by 'K CovM'. In the simulations, the results are obtained by performing 10^5 Monte Carlo experiments. The number of transmitted signal samples, receivers, and transmitters is $N = 10$, $N_r = 2$, and $N_t = 2$, respectively. Furthermore, the transmitted signal of the i th transmitter is $\mathbf{s}_i = \exp\{j\boldsymbol{\theta}_i\}$, where $\boldsymbol{\theta}_i \in \mathbb{R}^{N \times 1}$ is a random phase vector with each independent component uniformly distributed on $[0, 2\pi]$ (same as [23]). Suppose that all the independent noise vectors in $\{\mathbf{n}_{i,j}^r, \mathbf{n}_{i,j}^s, \text{ and } \mathbf{n}_{i,j,k}^n\}$ for $i = 0, 1, \dots, N_t - 1, j = 0, 1, \dots, N_r - 1, k = 0, 1, \dots, K - 1$ obey the same complex Gaussian distribution with zero mean and covariance matrix $\boldsymbol{\Sigma}_t$, where the m th row and n th column element of $\boldsymbol{\Sigma}_t$ is m/n , for $m \leq n$.

The SNR of the received target-path signal is defined as

$$\text{TNR} = \frac{1}{N N_r N_t \sigma_1^2} \sum_{i=0}^{N_t-1} \sum_{j=0}^{N_r-1} \|\mu_{ij}^s \mathbf{s}_i\|^2, \quad (56)$$

where σ_1^2 represents the colored noise power in the surveillance channel.

The SNR of the received direct-path signal is defined as

$$\text{DNR} = \frac{1}{N N_r N_t \sigma_2^2} \sum_{i=0}^{N_t-1} \sum_{j=0}^{N_r-1} \|\mu_{ij}^r \mathbf{s}_i\|^2, \quad (57)$$

where σ_2^2 represents the colored noise power in the reference channel.

In the simulation, μ_{ij}^s and μ_{ij}^r are randomly drawn from the Gaussian distribution with zero mean and unit variance. Then, they are scaled to achieve the desired TNR and DNR according to (56) and (57), respectively.

4.1. Variation of P_d with P_{fa}

The dependence of the probability of detection (P_d) on P_{fa} is illustrated in Figure 2. In Figure 2a, the TNR is -15 dB and the DNR is -30 dB. We choose the number of training samples to be $K = 3, K = 9$, and $K = 15$, respectively. In Figure 2b, the TNRs are set

as -18 dB, -15 dB, and -12 dB, respectively. The DNR is -30 dB and $K = 15$ training samples are used.

Figure 2a shows that the performance of the proposed GLRT is close to that of the GLRT in [23] when more than nine training samples are employed for the illustrated cases. This is due to that the estimate of unknown covariance matrix of colored noise becomes more and more accurate with the increase of training samples. As depicted in Figure 2b, the performance of the proposed GLRT and that of the GLRT in [23] are improved when the TNR is increased, since more accurate estimates of unknown parameters can be obtained with higher TNR.

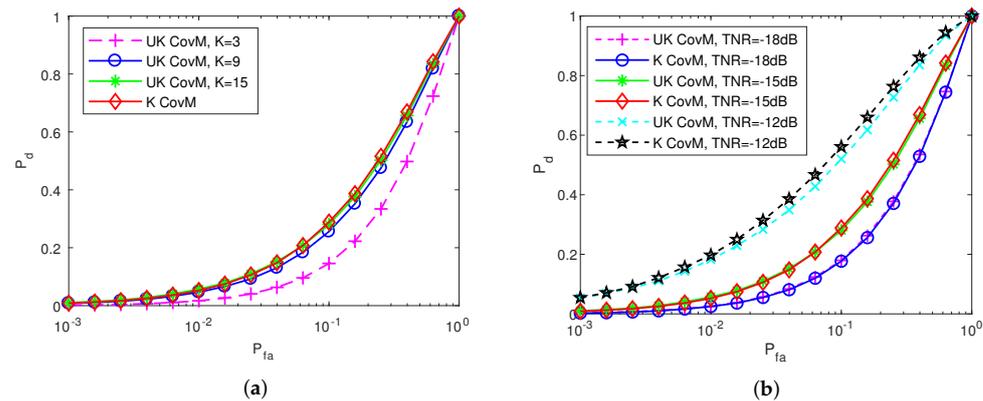


Figure 2. P_d vs. P_{fa} . (a) TNR = -15 dB. (b) $K = 15$.

Under different DNRs, the dependence of P_d on P_{fa} is illustrated in Figure 3. In Figure 3, the TNR is chosen to be -15 dB, and there are $K = 15$ training samples. The DNRs are -30 dB, -10 dB, and 10 dB, respectively. When the DNR is improved, both the performance of the proposed GLRT and that of the GLRT in [23] are improved, since more accurate estimates of unknown parameters are obtained with high DNR. Compared with the proposed GLRT, the GLRT in [23] can achieve greater performance gain when the DNR is increased, since the proposed GLRT needs to estimate more unknown parameters.

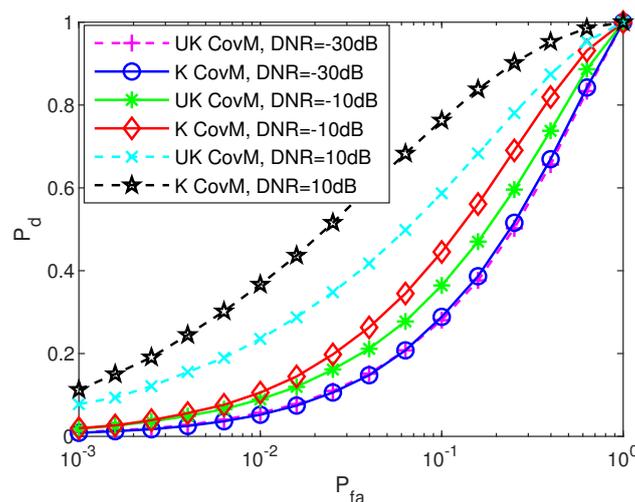


Figure 3. P_d vs. P_{fa} (TNR = -15 dB, $K = 15$).

4.2. Variation of P_d with TNR

The dependence of the P_d on TNR is illustrated in Figure 4. In Figure 4a, we consider that there are $K = 6$, $K = 9$, $K = 15$, and $K = 21$ training samples, respectively. The P_{fa} is chosen to be 10^{-3} , and the DNR is set as -30 dB. In Figure 4b, the DNR is -30 dB, and we consider that there are $K = 18$ training samples. The P_{fa} s are 10^{-3} , 10^{-2} , and 10^{-1} , respectively.

As shown in Figure 4, the P_d of both the GLRT in [23] and the proposed GLRT is increased with the increase of TNR, because the estimates of the unknown parameters become more accurate when TNR is increased. Moreover, as the number of training samples increases, the proposed GLRT gradually approaches the GLRT in [23]. Increasing the P_{fa} , the detection threshold will be reduced, and hence the P_d will be improved.

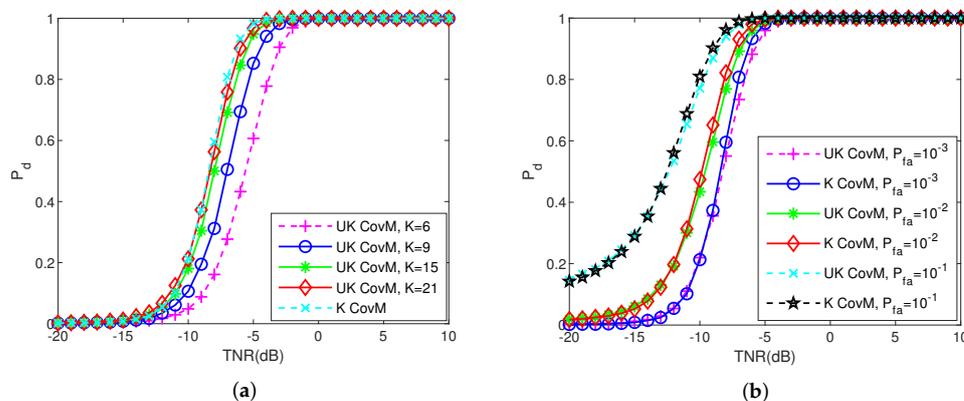


Figure 4. The variation of P_d with TNR. (a) DNR = -30 dB, $P_{fa} = 10^{-3}$. (b) DNR = -30 dB, $K = 18$.

In Figure 5a, the variation of the P_d with TNR is illustrated with different DNRs. In Figure 5a, the DNRs are considered to be -30 dB, -10 dB, and 10 dB, respectively, and there are $K = 21$ training samples. The P_{fa} is set as 10^{-3} .

In Figure 5b, the variation of the P_d with TNR for different numbers of receivers and transmitters is shown. In Figure 5b, there are $K = 20$ training samples, the DNR is -30 dB, and the P_{fa} is chosen to be 10^{-3} . There are $(N_t = 1, N_r = 2)$, $(N_t = 2, N_r = 2)$, $(N_t = 2, N_r = 3)$, and $(N_t = 3, N_r = 3)$ transmitters and receivers, respectively.

In Figure 5a, as expected, with an increase in DNR, the performance of both the GLRT in [23] and the proposed GLRT is improved, because the estimate accuracy of unknown parameters is improved when the DNR is improved. In Figure 5b, as expected, the P_d is improved when either the number of transmitters or the number of receivers is improved. Obviously, with more transmitters and receivers, we can obtain more observations that will contribute to obtaining more accurate estimates of unknown parameters. Hence, we can achieve higher P_d with more transmitters and receivers.

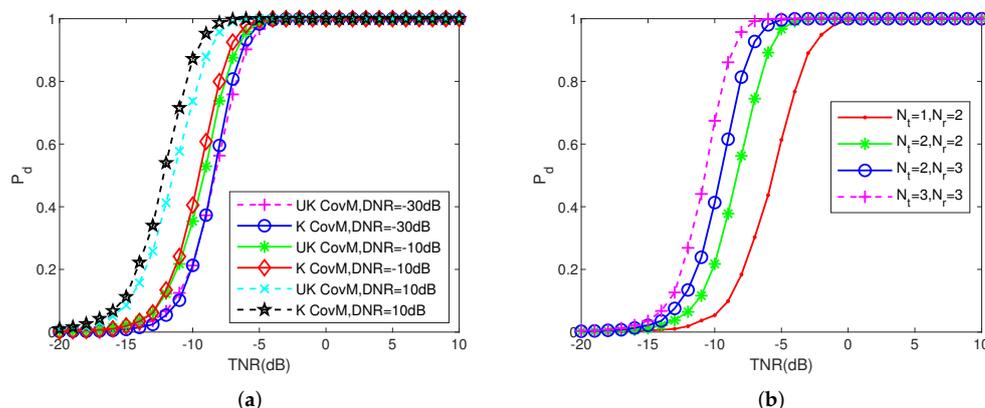


Figure 5. The variation of P_d with TNR. (a) $P_{fa} = 10^{-3}$, $K = 21$. (b) $P_{fa} = 10^{-3}$, $K = 20$, DNR = -30 dB.

4.3. P_d Loss

The performance of the GLRT in [23] is better than that of the proposed GLRT, since the covariance matrix is unknown for the proposed GLRT. To measure the performance degradation, the following P_d loss is utilized, which is defined as

$$L_{Pd} = P_{d_k} - P_{d_{UK}}, \quad (58)$$

where $P_{d_{UK}}$ denotes the P_d of the proposed GLRT, and P_{d_k} denotes the P_d of the GLRT in [23].

In Figure 6, the P_d loss of the proposed method is shown. In Figure 6a, the P_{fa} is 10^{-3} , the DNR is -30 dB, and the TNRs are, respectively, set as -14 dB, -12 dB, -10 dB, and -8 dB. In Figure 6b, the DNR is considered to be -30 dB, the TNR is chosen to be -12 dB, and the P_{fa} s are 10^{-3} , 10^{-2} , and 10^{-1} , respectively.

As expected, the P_d loss decreases as the number of training samples increases. This is due to that we can achieve more accurate estimates of the unknown covariance matrix of colored noise with the increase of the number of training samples. As shown in Figure 6, when there are more than 25 training samples, the P_d loss will be close to 0. For low TNR, the P_d loss is smaller as shown in Figure 6a. For low TNR, both the GLRT in [23] and the proposed GLRT have small P_d , which will result in a small P_d loss. Moreover, the proposed GLRT needs to estimate more unknown parameters than the GLRT in [23]. Using the same training samples, the GLRT in [23] will achieve more performance improvement when the TNR is increased. Hence, the P_d loss will be increased when the TNR is increased. In Figure 6b, for large P_{fa} , the P_d loss is larger. Under large P_{fa} , both the P_d of the GLRT in [23] and that of the proposed GLRT are large. For large P_{fa} , to achieve the same P_d loss, it needs more training samples for the proposed GLRT.

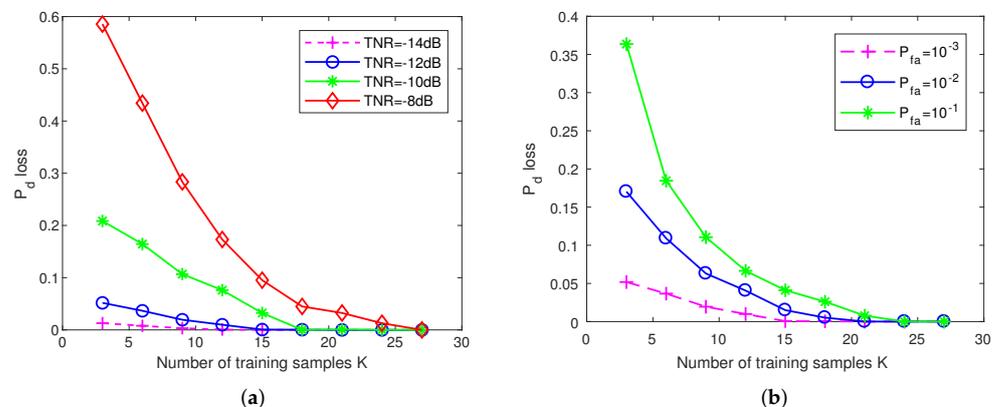


Figure 6. The P_d loss of the proposed method. (a) DNR = -30 dB, $P_{fa} = 10^{-3}$. (b) DNR = -30 dB, TNR = -12 dB.

In Figure 7, the dependence of the P_d loss on the number of training samples is illustrated. In Figure 7, the TNR is -12 dB, the P_{fa} is 10^{-3} , and the DNRs are -30 dB, -10 dB, and 10 dB, respectively. As the number of training samples increases, the P_d loss is decreased. Moreover, the proposed GLRT will need more training samples to obtain the same P_d loss when the DNR is improved.

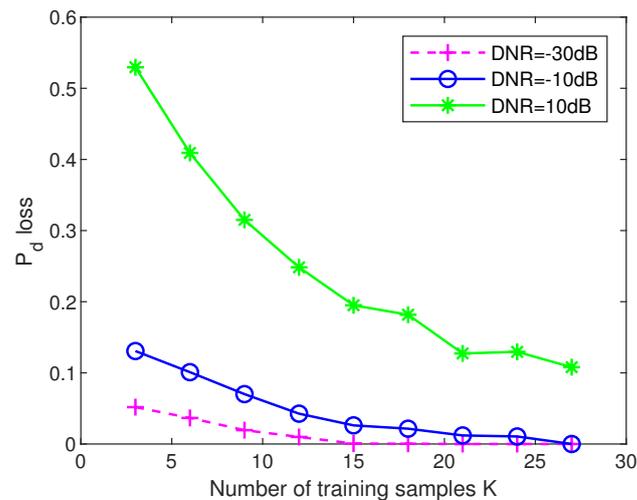


Figure 7. The P_d loss of the proposed method (TNR = -12 dB, $P_{fa} = 10^{-3}$).

5. Conclusions

In this paper, target detection of a passive MIMO radar is studied when the transmitted signals, complex channel coefficients, and covariance matrix of colored Gaussian noise are unknown. A GLRT for passive MIMO radar is derived by utilizing the training samples. Simulation results show that the proposed GLRT can almost achieve the same performance of a GLRT with a known covariance matrix if there are sufficient training samples. In this paper, we assume that the noise and clutter are independent and identically distributed, while they might be heterogeneous in practice. In our future work, target detection for passive MIMO radar in the heterogeneous environment will be investigated. Moreover, the intelligent methods has become a hot topic for target detection nowadays, such as machine-learning-based approaches. In our future work, we will also investigate the intelligent method for target detection in the passive MIMO radar network.

Author Contributions: Conceptualization, Y.L. (Yongjun Liu) and G.L.; methodology, Y.L. (Yongjun Liu); software, Y.L. (Yongjun Liu) and H.L.; validation, Y.L. (Yongjun Liu), G.L., and S.Z.; formal analysis, Y.L. (Yongjun Liu) and G.L.; investigation, Y.L. (Yongjun Liu) and S.Z.; resources, Y.L. (Yongjun Liu), Y.Y., and Y.L. (Yachao Li); data curation, Y.L. (Yongjun Liu) and H.L.; writing—original draft preparation, Y.L. (Yongjun Liu) and H.L.; writing—review and editing, Y.L. (Yongjun Liu); visualization, Y.L. (Yongjun Liu) and S.Z.; supervision, Y.L. (Yachao Li), G.L., and Y.Y.; project administration, Y.L. (Yongjun Liu), G.L., and Y.L. (Yachao Li); funding acquisition, Y.L. (Yongjun Liu), G.L., and Y.L. (Yachao Li). All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China under Grant No. 62001352, by the National Key R & D Program of China under Grant No. 2018YFB2202500, by the Foundation for Innovative Research Groups of the National Natural Science Foundation of China under Grant No. 61621005, and by the National Natural Science Foundation of China under Grant No. 61671352.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: The authors would like to thank the anonymous reviewers for their valuable and useful comments and suggestions that helped improve the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Proof of Theorem 1

Since \mathbf{A} is a positive Hermitian matrix, according to Theorem 21.5.7 in [47], \mathbf{A} is orthogonally diagonalizable, i.e., $\mathbf{A} = \mathbf{VDV}^H$, where $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n]$ is an orthogonal matrix, $\mathbf{D} = \text{diag}\{d_1, d_2, \dots, d_n\}$ is a diagonal matrix, and $0 < d_1 \leq d_2 \leq \dots \leq d_n$.

For any $n \times m$ matrix \mathbf{X}_m that satisfies the constraint in (35), we can obtain that the $m \times m$ matrix $\mathbf{G}(\mathbf{X}_m) = \mathbf{X}_m^H \mathbf{A} \mathbf{X}_m$ is also a positive Hermitian matrix, since \mathbf{A} is a positive Hermitian matrix. Thus, according to Theorem 21.5.7 in [47], $\mathbf{G}(\mathbf{X}_m)$ is orthogonally diagonalizable, i.e., $\mathbf{G}(\mathbf{X}_m) = \mathbf{W}\Phi\mathbf{W}^H$, where $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m]$ is an $m \times m$ orthogonal matrix, $\Phi = \text{diag}\{\phi_1, \phi_2, \dots, \phi_m\}$ is a diagonal matrix, and $0 < \phi_1 \leq \phi_2 \leq \dots \leq \phi_m$.

Let $G(\mathbf{X}_m) = |\mathbf{X}_m^H \mathbf{A} \mathbf{X}_m| = |\mathbf{G}(\mathbf{X}_m)|$ represent the objective function in (35). Since \mathbf{W} is an orthogonal matrix, we can obtain

$$G(\mathbf{X}_m) = |\mathbf{X}_m^H \mathbf{A} \mathbf{X}_m| = |\mathbf{W}^H \mathbf{X}_m^H \mathbf{A} \mathbf{X}_m \mathbf{W}| = |\mathbf{Z}_m^H \mathbf{A} \mathbf{Z}_m|, \quad (\text{A1})$$

where $\mathbf{Z}_m = \mathbf{X}_m \mathbf{W}$ with the i th column of \mathbf{Z}_m being $\mathbf{z}_i = \mathbf{X}_m \mathbf{w}_i$, for $i = 1, 2, \dots, m$.

Since $\mathbf{G}(\mathbf{X}_m) = \mathbf{W}\Phi\mathbf{W}^H$, we can obtain

$$G(\mathbf{X}_m) = |\Phi| = \prod_{i=1}^m \phi_i \quad (\text{A2})$$

$$= |\mathbf{Z}_m^H \mathbf{A} \mathbf{Z}_m| = G(\mathbf{Z}_m), \quad (\text{A3})$$

where from (A2) to (A3) we use (A1).

Suppose that $\prod_{i=1}^m \phi_i < \prod_{i=1}^m d_i$, i.e., suppose that there is a solution \mathbf{Z}_m that satisfies $G(\mathbf{Z}_m) = \prod_{i=1}^m \phi_i < \prod_{i=1}^m d_i$. Since ϕ_i and d_i , for $i = 1, 2, \dots, m$, are positive and non-decreasing with the increase of i , there must be some i' that satisfies

$$\phi_{i'} < d_{i'}. \quad (\text{A4})$$

Note that $\mathbf{Z}_m^H \mathbf{Z}_m = \mathbf{I}_m$. Hence, the first i' columns of \mathbf{Z}_m , denoted as $\mathbf{Z}_{i'}$, form the basis for an i' -dimensional subspace $\mathbb{S}^{i'}$ of n -dimensional complex space \mathbb{C}^n . Any $n \times 1$ vector \mathbf{x} in the subspace $\mathbb{S}^{i'}$ can be represented as

$$\mathbf{x} = \sum_{i=1}^{i'} \alpha_i \mathbf{z}_i, \quad (\text{A5})$$

where α_i , for $i = 1, 2, \dots, i'$, are the coefficients of \mathbf{x} in the i' -dimensional subspace $\mathbb{S}^{i'}$.

Hence, we can obtain

$$\max_{\mathbf{x} \in \mathbb{S}^{i'}, \mathbf{x} \neq \mathbf{0}} \frac{\mathbf{x}^H \mathbf{A} \mathbf{x}}{\mathbf{x}^H \mathbf{x}} = \max_{\alpha_i, i=1,2,\dots,i'} \frac{\sum_{i=1}^{i'} |\alpha_i|^2 \phi_i}{\sum_{i=1}^{i'} |\alpha_i|^2} = \phi_{i'}. \quad (\text{A6})$$

Using Theorem 4.2.6 in [48], we can obtain that

$$\min_{\dim(\mathbb{S})=i'} \max_{\mathbf{x} \in \mathbb{S}, \mathbf{x} \neq \mathbf{0}} \frac{\mathbf{x}^H \mathbf{A} \mathbf{x}}{\mathbf{x}^H \mathbf{x}} = d_{i'} \leq \phi_{i'}, \quad (\text{A7})$$

where $\dim(\mathbb{S})$ represents the dimension of the subspace \mathbb{S} . Obviously, the obtained solution in (A7) contradicts the assumption in (A4). This completes the proof. Theorem 1 can also be proved by majorization theory [49,50].

Appendix B. Maximum Likelihood Estimate $\hat{\mathbf{U}}$

The optimal solution to (34) is $\hat{\mathbf{V}}_{2N_r N_t - N_t} = \mathbf{\Psi}_{2N_r N_t - N_t} = [\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \dots, \boldsymbol{\psi}_{2N_r N_t - N_t}]$. Hence, we can obtain

$$\mathbf{P}_{\hat{\mathbf{U}}_{sr}^H}^\perp = \mathbf{I}_{2N_r N_t} - \hat{\mathbf{U}}_{sr}^H (\hat{\mathbf{U}}_{sr} \hat{\mathbf{U}}_{sr}^H)^\dagger \hat{\mathbf{U}}_{sr} = \mathbf{I}_{2N_r N_t} - \mathbf{P}_{\hat{\mathbf{U}}_{sr}^H} = \mathbf{\Psi}_{2N_r N_t - N_t} \mathbf{\Psi}_{2N_r N_t - N_t}^H. \quad (\text{A8})$$

Furthermore, we can obtain

$$\mathbf{P}_{\hat{\mathbf{U}}_{sr}^H} = \hat{\mathbf{U}}_{sr}^H (\hat{\mathbf{U}}_{sr} \hat{\mathbf{U}}_{sr}^H)^\dagger \hat{\mathbf{U}}_{sr} = \mathbf{I}_{2N_r N_t} - \mathbf{P}_{\hat{\mathbf{U}}_{sr}^H}^\perp = \mathbf{I}_{2N_r N_t} - \mathbf{\Psi}_{2N_r N_t - N_t} \mathbf{\Psi}_{2N_r N_t - N_t}^H. \quad (\text{A9})$$

Note that $\mathbf{\Psi} = [\mathbf{\Psi}_{2N_r N_t - N_t}, \mathbf{\Psi}_{N_t}]$ and $\mathbf{\Psi} \mathbf{\Psi}^H = \mathbf{I}_{2N_r N_t}$, where $\mathbf{\Psi}_{N_t} = [\boldsymbol{\psi}_{2N_r N_t - N_t + 1}, \boldsymbol{\psi}_{2N_r N_t - N_t + 2}, \dots, \boldsymbol{\psi}_{2N_r N_t}]$ is a $2N_r N_t \times N_t$ full-column rank matrix. Hence, we can obtain that

$$\begin{aligned} \mathbf{P}_{\hat{\mathbf{U}}_{sr}^H} &= \mathbf{\Psi} \mathbf{\Psi}^H - \mathbf{\Psi}_{2N_r N_t - N_t} \mathbf{\Psi}_{2N_r N_t - N_t}^H \\ &= [\mathbf{\Psi}_{2N_r N_t - N_t} \quad \mathbf{\Psi}_{N_t}] \begin{bmatrix} \mathbf{I}_{2N_r N_t - N_t} & \mathbf{0}_{(2N_r N_t - N_t) \times N_t} \\ \mathbf{0}_{N_t \times (2N_r N_t - N_t)} & \mathbf{I}_{N_t} \end{bmatrix} \begin{bmatrix} \mathbf{\Psi}_{2N_r N_t - N_t}^H \\ \mathbf{\Psi}_{N_t}^H \end{bmatrix} \\ &\quad - [\mathbf{\Psi}_{2N_r N_t - N_t} \quad \mathbf{\Psi}_{N_t}] \begin{bmatrix} \mathbf{I}_{2N_r N_t - N_t} & \mathbf{0}_{(2N_r N_t - N_t) \times N_t} \\ \mathbf{0}_{N_t \times (2N_r N_t - N_t)} & \mathbf{0}_{N_t} \end{bmatrix} \begin{bmatrix} \mathbf{\Psi}_{2N_r N_t - N_t}^H \\ \mathbf{\Psi}_{N_t}^H \end{bmatrix} \\ &= [\mathbf{\Psi}_{2N_r N_t - N_t} \quad \mathbf{\Psi}_{N_t}] \begin{bmatrix} \mathbf{0}_{2N_r N_t - N_t} & \mathbf{0}_{(2N_r N_t - N_t) \times N_t} \\ \mathbf{0}_{N_t \times (2N_r N_t - N_t)} & \mathbf{I}_{N_t} \end{bmatrix} \begin{bmatrix} \mathbf{\Psi}_{2N_r N_t - N_t}^H \\ \mathbf{\Psi}_{N_t}^H \end{bmatrix} = \mathbf{\Psi}_{N_t} \mathbf{\Psi}_{N_t}^H. \end{aligned} \quad (\text{A10})$$

Now we show that $\hat{\mathbf{U}}_{sr}^H = \mathbf{\Psi}_{N_t} \boldsymbol{\Omega} \boldsymbol{\Gamma}$, where $\boldsymbol{\Omega}$ is any $N_t \times N_t$ invertible diagonal matrix, and $\boldsymbol{\Gamma}$ is any $N_t \times N_t$ orthogonal matrix, i.e., $\boldsymbol{\Gamma}^H \boldsymbol{\Gamma} = \boldsymbol{\Gamma} \boldsymbol{\Gamma}^H = \mathbf{I}_{N_t}$. We can verify that

$$\mathbf{P}_{\hat{\mathbf{U}}_{sr}^H} = \hat{\mathbf{U}}_{sr}^H (\hat{\mathbf{U}}_{sr} \hat{\mathbf{U}}_{sr}^H)^\dagger \hat{\mathbf{U}}_{sr} = \mathbf{\Psi}_{N_t} \boldsymbol{\Omega} \boldsymbol{\Gamma} (\boldsymbol{\Gamma}^H \boldsymbol{\Omega} \mathbf{\Psi}_{N_t}^H \mathbf{\Psi}_{N_t} \boldsymbol{\Omega} \boldsymbol{\Gamma})^\dagger \boldsymbol{\Gamma}^H \boldsymbol{\Omega} \mathbf{\Psi}_{N_t}^H. \quad (\text{A11})$$

Note that $\mathbf{\Psi}_{N_t}^H \mathbf{\Psi}_{N_t} = \mathbf{I}_{N_t}$ and $(\boldsymbol{\Gamma}^H \boldsymbol{\Omega} \mathbf{\Psi}_{N_t}^H \mathbf{\Psi}_{N_t} \boldsymbol{\Omega} \boldsymbol{\Gamma})^\dagger = \boldsymbol{\Gamma}^H \boldsymbol{\Omega}^{-2} \boldsymbol{\Gamma}$. Hence, (A11) can be simplified as

$$\mathbf{P}_{\hat{\mathbf{U}}_{sr}^H} = \mathbf{\Psi}_{N_t} \boldsymbol{\Omega} \boldsymbol{\Gamma} \boldsymbol{\Gamma}^H \boldsymbol{\Omega}^{-2} \boldsymbol{\Gamma} \boldsymbol{\Gamma}^H \boldsymbol{\Omega} \mathbf{\Psi}_{N_t}^H = \mathbf{\Psi}_{N_t} \mathbf{\Psi}_{N_t}^H, \quad (\text{A12})$$

which is the same result in (A10).

Hence, the maximum likelihood estimate $\hat{\mathbf{U}}_{sr}$ is

$$\hat{\mathbf{U}}_{sr} = \boldsymbol{\Gamma}^H \boldsymbol{\Omega} \mathbf{\Psi}_{N_t}^H. \quad (\text{A13})$$

References

1. Liu, J.; Li, H.; Himed, B. On the performance of the cross-correlation detector for passive radar applications. *Signal Process.* **2015**, *113*, 32–37. [CrossRef]
2. Cui, G.; Liu, J.; Li, H.; Himed, B. Signal detection with noisy reference for passive sensing. *Signal Process.* **2015**, *108*, 389–399. [CrossRef]
3. Subedi, S.; Zhang, Y.D.; Amin, M.G.; Himed, B. Group sparsity based multi-target tracking in passive multi-static radar systems using Doppler-only measurements. *IEEE Trans. Signal Process.* **2016**, *64*, 3619–3634. [CrossRef]
4. Prateek, G.; Hurtado, M.; Nehorai, A. Target detection using weather radars and electromagnetic vector sensors. *Signal Process.* **2017**, *137*, 387–397. [CrossRef]
5. Zhang, X.; Li, H.; Himed, B. Multistatic detection for passive radar with direct-path interference. *IEEE Trans. Aerosp. Electron. Syst.* **2017**, *53*, 915–925. [CrossRef]
6. Zhang, Y.D.; Amin, M.G.; Himed, B. Structure-aware sparse reconstruction and applications to passive multistatic radar. *IEEE Aerosp. Electron. Syst. Mag.* **2017**, *32*, 68–78. [CrossRef]

7. He, Q.; Blum, R.S. The significant gains from optimally processed multiple signals of opportunity and multiple receive stations in passive radar. *IEEE Signal Process. Lett.* **2014**, *21*, 180–184. [[CrossRef](#)]
8. Zhao, H.Y.; Liu, J.; Zhang, Z.J.; Liu, H.; Zhou, S. Linear fusion for target detection in passive multistatic radar. *Signal Process.* **2017**, *130*, 175–182. [[CrossRef](#)]
9. Chalise, B.K.; Himed, B. GLRT detector in single frequency multi-static passive radar systems. *Signal Process.* **2018**, *142*, 504–512. [[CrossRef](#)]
10. Fazlollahpoor, M.; Derakhtian, M.; Khorshidi, S. Rao Detector for Passive MIMO Radar With Direct-Path Interference. *IEEE Trans. Aerosp. Electron. Syst.* **2020**, *56*, 2999–3009. [[CrossRef](#)]
11. Li, Y.; He, Q.; Blum, R.S. Limited-Complexity Receiver Design for Passive/Active MIMO Radar Detection. *IEEE Trans. Signal Process.* **2019**, *67*, 3258–3271. [[CrossRef](#)]
12. Horstmann, S.; Ramírez, D.; Schreier, P.J. Two-Channel Passive Detection of Cyclostationary Signals. *IEEE Trans. Signal Process.* **2020**, *68*, 2340–2355. [[CrossRef](#)]
13. Santamaria, I.; Scharf, L.L.; Via, J.; Wang, H.; Wang, Y. Passive detection of correlated subspace signals in two MIMO channels. *IEEE Trans. Signal Process.* **2017**, *65*, 5266–5280. [[CrossRef](#)]
14. Santamaria, I.; Via, J.; Scharf, L.L.; Wang, Y. A GLRT approach for detecting correlated signals in white noise in two MIMO channels. In Proceedings of the 2017 25th European Signal Processing Conference (EUSIPCO), Kos, Greece, 28 August–2 September 2017; IEEE: Hoboken, NJ, USA, 2017; pp. 1395–1399.
15. Santamaria, I.; Scharf, L.L.; Cochran, D.; Via, J. Passive detection of rank-one signals with a multiantenna reference channel. In Proceedings of the 2016 24th European Signal Processing Conference (EUSIPCO), Budapest, Hungary, 29 August–2 September 2016; IEEE: Hoboken, NJ, USA, 2016; pp. 140–144.
16. Howard, S.D.; Sirianunpiboon, S. Passive radar detection using multiple transmitters. In Proceedings of the 2013 Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 3–6 November 2013; IEEE: Hoboken, NJ, USA, 2013; pp. 945–948.
17. Karthik, A.K.; Blum, R.S. Improved Detection Performance for Passive Radars Exploiting Known Communication Signal Form. *IEEE Signal Process. Lett.* **2018**, *25*, 1625–1629. [[CrossRef](#)]
18. Liu, Y.; Blum, R.S.; Liao, G.; Zhu, S. Passive MIMO radar detection exploiting known format of the communication signal observed in colored noise with unknown covariance matrix. *Signal Process.* **2020**, *174*, 107611. [[CrossRef](#)]
19. Liu, Y.; Liao, G.; Xu, J.; Yang, Z.; Yin, Y. Improving detection performance of passive MIMO radar by exploiting the preamble information of communications signal. *IEEE Syst. J.* **2020**, 1–12. [[CrossRef](#)]
20. Colone, F.; O’hagan, D.; Lombardo, P.; Baker, C. A multistage processing algorithm for disturbance removal and target detection in passive bistatic radar. *IEEE Trans. Aerosp. Electron. Syst.* **2009**, *45*, 698–722. [[CrossRef](#)]
21. Zaimbashi, A.; Derakhtian, M.; Sheikhi, A. GLRT-based CFAR detection in passive bistatic radar. *IEEE Trans. Aerosp. Electron. Syst.* **2013**, *49*, 134–159. [[CrossRef](#)]
22. Masjedi, M.; Moddares-Hashemi, M.; Sadri, S. Theoretical approach for target detection and interference cancellation in passive radars. *IET Radar Sonar Navigat.* **2013**, *7*, 205–216. [[CrossRef](#)]
23. Hack, D.E.; Patton, L.K.; Himed, B.; Saville, M.A. Detection in passive MIMO radar networks. *IEEE Trans. Signal Process.* **2014**, *62*, 2999–3012. [[CrossRef](#)]
24. Hack, D.E.; Patton, L.K.; Himed, B.; Saville, M.A. Centralized passive MIMO radar detection without direct-path reference signals. *IEEE Trans. Signal Process.* **2014**, *62*, 3013–3023.
25. Wang, Y.; Scharf, L.L.; Santamaria, I.; Wang, H. Canonical correlations for target detection in a passive radar network. In Proceedings of the 2016 50th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 6–9 November 2016; IEEE: Hoboken, NJ, USA, 2016; pp. 1159–1163.
26. Daun, M.; Nickel, U.; Koch, W. Tracking in multistatic passive radar systems using DAB/DVB-T illumination. *Signal Process.* **2012**, *92*, 1365–1386. [[CrossRef](#)]
27. Wang, H.; Wang, J.; Zhong, L. Mismatched filter for analogue TV-based passive bistatic radar. *IET Radar Sonar Navigat.* **2011**, *5*, 573–581. [[CrossRef](#)]
28. Sun, H.; Tan, D.K.; Lu, Y.; Lesturgie, M. Applications of passive surveillance radar system using cell phone base station illuminators. *IEEE Aerosp. Electron. Syst. Mag.* **2010**, *25*, 10–18. [[CrossRef](#)]
29. Kelly, E.J. An adaptive detection algorithm. *IEEE Trans. Aerosp. Electron. Syst.* **1986**, *2*, 115–127. [[CrossRef](#)]
30. Xiao, Y.; Cui, G.; Yi, W.; Kong, L.; Yang, J. Adaptive detection and estimation for an unknown occurring interval signal in correlated Gaussian noise. *Signal Process.* **2015**, *108*, 440–450. [[CrossRef](#)]
31. He, Q.; Lehmann, N.H.; Blum, R.S.; Haimovich, A.M. MIMO radar moving target detection in homogeneous clutter. *IEEE Trans. Aerosp. Electron. Syst.* **2010**, *46*, 1290–1301. [[CrossRef](#)]
32. Li, H.; Wang, Z.; Liu, J.; Himed, B. Moving target detection in distributed MIMO radar on moving platforms. *IEEE J. Sel. Top. Signal Process.* **2015**, *9*, 1524–1535. [[CrossRef](#)]
33. Liu, W.; Wang, Y.; Liu, J.; Xie, W.; Chen, H.; Gu, W. Adaptive detection without training data in colocated MIMO radar. *IEEE Trans. Aerosp. Electron. Syst.* **2015**, *51*, 2469–2479. [[CrossRef](#)]
34. Liu, J.; Zhou, S.; Liu, W.; Zheng, J.; Liu, H.; Li, J. Tunable Adaptive Detection in Colocated MIMO Radar. *IEEE Trans. Signal Process.* **2017**, *66*, 1080–1092. [[CrossRef](#)]

35. Wang, Z.; Li, H.; Himed, B. A sparsity based GLRT for moving target detection in distributed MIMO radar on moving platforms. In Proceedings of the 2015 49th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 8–11 November 2015; IEEE: Hoboken, NJ, USA, 2015; pp. 90–94.
36. Dogandzic, A.; Nehorai, A. Generalized multivariate analysis of variance-A unified framework for Signal Processing in correlated noise. *IEEE Signal Process. Mag.* **2003**, *20*, 39–54. [[CrossRef](#)]
37. Robey, F.C.; Fuhrmann, D.R.; Kelly, E.J.; Nitzberg, R. A CFAR adaptive matched filter detector. *IEEE Trans. Aerosp. Electron. Syst.* **1992**, *28*, 208–216. [[CrossRef](#)]
38. Monticciolo, P.; Kelly, E.; Porakis, J. A noncoherent adaptive detection technique. *IEEE Trans. Aerosp. Electron. Syst.* **1992**, *28*, 115–124. [[CrossRef](#)]
39. Tang, B.; Liu, J.; Huang, Z.; Wang, G.; Fan, F. Adaptive Target Detection in Gaussian Clutter Edges. *IEEE Trans. Aerosp. Electron. Syst.* **2020**, *56*, 1662–1673. [[CrossRef](#)]
40. Steven, M.K. *Fundamentals of Statistical Signal Processing*; PTR Prentice-Hall: Englewood Cliffs, NJ, USA, 1993.
41. Wang, F.; Li, H.; Zhang, X.; Himed, B. Signal Parameter Estimation for Passive Bistatic Radar With Waveform Correlation Exploitation. *IEEE Trans. Aerosp. Electron. Syst.* **2018**, *54*, 1135–1150. [[CrossRef](#)]
42. Wang, F.; Wang, P.; Zhang, X.; Li, H.; Himed, B. An Overview of Parametric Modeling and Methods for Radar Target Detection With Limited Data. *IEEE Access* **2021**, *9*, 60459–60469. [[CrossRef](#)]
43. Palmer, J.E.; Searle, S.J. Evaluation of adaptive filter algorithms for clutter cancellation in passive bistatic radar. In Proceedings of the 2012 IEEE Radar Conference, Atlanta, GA, USA, 7–11 May 2012; IEEE: Hoboken, NJ, USA, 2012; pp. 0493–0498.
44. Scharf, L.L.; Demeure, C. *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*; Addison-Wesley: Reading, MA, USA, 1991; Volume 63.
45. Kay, S.M. *Fundamentals of Statistical Signal Processing. Detection Theory, Volume II*; Prentice Hall: Upper Saddle River, NJ, USA, 1998.
46. Muirhead, R.J. *Aspects of Multivariate Statistical Theory*; John Wiley & Sons: Hoboken, NJ, USA, 1982; Volume 197.
47. Harville, D.A. *Matrix Algebra from a Statistician's Perspective*; Springer: Berlin/Heidelberg, Germany, 1997; Volume 1.
48. Horn, R.A.; Johnson, C.R. *Matrix Analysis*; Cambridge University Press: Cambridge, UK, 2013.
49. Tang, B.; Tang, J.; Peng, Y. MIMO radar waveform design in colored noise based on information theory. *IEEE Trans. Signal Process.* **2010**, *58*, 4684–4697. [[CrossRef](#)]
50. Tang, B.; Li, J. Spectrally Constrained MIMO Radar Waveform Design Based on Mutual Information. *IEEE Trans. Signal Process.* **2019**, *67*, 115–127. [[CrossRef](#)]