

Limited-Complexity Receiver Design for Passive/Active MIMO Radar Detection

Yang Li, *Student Member, IEEE*, Qian He, *Senior Member, IEEE*, and Rick S. Blum, *Fellow, IEEE*

Abstract—In this paper we develop efficient methods for devising lower complexity receivers that can achieve performance close to the full complexity receivers for passive/active multiple-input-multiple-output (MIMO) radar. The method employed eliminates some parts of the test statistic to lower either hardware or software complexity. For the case of spatially uncorrelated reflection coefficients and spatially white clutter-plus-noise, the test statistic requires the computation of a set of matched filters, each matched to a signal from a different transmitter. In this case our method is equivalent to selecting a specific set of transmitters to provide optimum performance. In the more general case of correlated clutter-plus-noise and reflection coefficients, then the test statistic requires the computation of a larger set of matched filters. These matched filters correlate the clutter-plus-noise free signal received at one receive antenna due to the signal transmitted from some transmit antenna and the signal received at another receive antenna. In the more general case our algorithm picks the best of these matched filters to implement when the total number of these matched filters one can implement is limited.

Index Terms—MIMO radar, matched filter, transmitter selection, target detection.

I. INTRODUCTION

The performance of multiple-input-multiple-output (MIMO) radar systems has been widely investigated since 2004 [1]–[11]. MIMO radar can benefit from additional spatial separated antennas since it can observe a target from different directions [12]. Passive radar has also attracted attention over the past few years [13]–[15] due to the advantages of low cost, low probability of intercept, etc. In passive radar, existing illuminators of opportunity be employed to save the cost and energy on transmission.

Passive MIMO radar [16] employs multiple existing illuminators and multiple receivers. In recent work, the performance of passive MIMO radar systems has been investigated intensively. Target detection using the generalized likelihood ratio test has been studied with [16] or without [17] the consideration of direct-path signals. The modified Cramér-Rao lower bounds for target parameter estimation are derived in [18], [19].

The work of Y. Li and Q. He was supported by the National Nature Science Foundation of China under Grants No. 61571091 and 61371184, and Huo Yingdong Education Foundation under Grant No. 161097. The work of R. S. Blum was supported by the National Science Foundation under Grant No. ECCS-1405579.

Y. Li and Q. He are with University of Electronic Science and Technology of China, Chengdu, Sichuan 611731 China (emails: yangliuestc@163.com, qianhe@uestc.edu.cn). Rick S. Blum is with Lehigh University, Bethlehem, PA 18015 USA (email: rblum@eecs.lehigh.edu).

In [18], universal mobile telecommunications system (UMTS) signals are used to estimate target delay and Doppler, and in [19], the authors studied target velocity and location estimation from L-band digital aeronautical communication system signals. The work in [20] investigated the ambiguity function for UMTS-based passive radar under both coherent and non-coherent processing. In [21], target tracking is studied for passive MIMO radar.

Passive and active MIMO radar implementations require large hardware and software complexity when a large number of transmitters are present, so that the lower complexity approaches studied here, like transmitter selection approaches, are of considerable interest. Our problem is somewhat similar to antenna selection. In [22], [23], the antenna selection strategies for minimizing the average error probability in a communication system have been investigated when maximum likelihood or zero forcing receiver is used. In [24], the authors consider the optimal antenna subset selection in a communication system with space-time coding in flat fading channels based on exact or statistical channel knowledge. The work in [25] investigated the optimal antenna selection for maximizing the channel capacity in a communication system under the assumption that only the long-term channel statistics are known. In [26], a geometry-based sensor selection method is investigated for Kalman filtering. In [27], statistical information is used to select transmitters in a MIMO radar for improving target detection performance by increasing the total amount of average incoming energy. In [28]–[30], antenna selection in distributed MIMO radar for target localization by minimizing the trace of Cramér-Rao bound is studied. The optimal antenna selection and placement based on Fisher information matrix for estimating target location is investigated in [31]. In [32], the antenna selection for minimizing the volume of an η -confidence ellipsoid of estimation error is presented.

For MIMO radar systems, the hardware and software complexity depends heavily on the number of matched filters (MFs) employed. For many practical scenarios, when the number of transmitters is large and the clutter-plus-noise is non-ideal, the number of candidate matched MFs and the associated hardware (e.g., adder and multiplier) required is typically large. It is necessary to control the complexity and cost and to simultaneously achieve the best possible performance. The proposed method can be used in active MIMO

radar systems. If the location of the illuminators of opportunity are known, the statistical properties of the target reflection coefficients and clutter-plus-noise are learned or estimated, and the direct-path signals transmitted from different illuminators can be well estimated, we can also use the approach in passive MIMO radar. We derive the log-likelihood ratio (LLR) function for a general case accounting for possibly correlated target reflection coefficients and clutter-plus-noise, and show that a limited-complexity receiver can be achieved by matched filter (MF) selection.

For the spatially uncorrelated reflection coefficients and spatially white clutter-plus-noise, we show that the MF selection is equivalent to selecting a subset of transmitters to provide optimum performance. We first consider the case where the signals transmitted by the transmitters are mutually orthogonal and maintain orthogonality for different delays and prove that at each receiver selecting the MFs corresponding to the largest signal-to-clutter-plus-noise ratios (SCNRs) leads to the best detection performance. Then, we consider the case where the transmitted signals may not be mutually orthogonal. For a special case where each receiver selects one MF, a closed-form solution is presented. When each receiver selects an arbitrary number of MFs, a closed-form solution is not available to optimize detection probability. In these cases, suboptimal criteria such as Kullback-Leibler (KL) distance and divergence [12] can be employed. In this paper, we use KL distance as a low complexity approximate measure for detection performance considering that it is typically used in the Neyman-Pearson settings [33]. A greedy algorithm adopted from [34]–[36] is considered in this paper. Numerical experiments justify our approach.

The rest of the paper is organized as follows. The signal model for target detection is presented in Section II, where the likelihood ratio test is derived. The lower-complexity receiver design method is introduced in Section III. Transmitter selection is studied in Section IV. Numerical examples are presented in Section V. Finally, Section VI concludes the paper.

II. SIGNAL MODEL FOR TARGET DETECTION

Consider a MIMO radar system with M transmitters and N receivers located at known positions $(x_{t,m}, y_{t,m})$, $m = 1, 2, \dots, M$ and $(x_{r,n}, y_{r,n})$, $n = 1, 2, \dots, N$, respectively. The signal from the m -th transmitter is assumed known or perfectly estimated¹ and written as $\sqrt{E_m} s_m(t)$, $m = 1, 2, \dots, M$, where $\int_{\mathcal{T}} |s_m(t)|^2 dt = 1$, E_m is the transmitted energy, and $\mathcal{T}$ is the observation interval. Consider a possible static target in the cell-

under-test located at (x, y) . The received signal at the n -th single antenna receiver can be written as

$$r_n(t) = \sum_{m=1}^M \frac{\beta_{mn} \sqrt{E_m}}{R_{t,m} R_{r,n}} s_m(t - \tau_{mn}) + w_n(t), \quad t \in \mathcal{T} \quad (1)$$

in which $w_n(t)$ represents the clutter-plus-noise, assumed to be temporally white zero-mean complex Gaussian such that $\mathbb{E}\{w_i(t)w_j^*(u)\} = N_{ij}\delta(t-u)$, where N_{ij} is the (i, j) -th element of a positive definite Hermitian matrix $\mathbf{N}$, and $\mathbb{E}\{\cdot\}$ denotes expectation. The reflection coefficient β_{mn} is complex Gaussian distributed with zero-mean and variance σ_{mn}^2 while being independent of the clutter-plus-noise components. The term $R_{t,m}$ is the distance between the m -th transmitter and the target, $R_{r,n}$ is the distance between the n -th receiver and the target, and τ_{mn} is the time delay between the m -th transmitter and the n -th receiver. They satisfy $R_{t,m} = \sqrt{(x_{t,m} - x)^2 + (y_{t,m} - y)^2}$, $R_{r,n} = \sqrt{(x_{r,n} - x)^2 + (y_{r,n} - y)^2}$, and $\tau_{mn} = (R_{t,m} + R_{r,n})/c$, where c is the speed of light. Define

$$\boldsymbol{\xi} = [\boldsymbol{\xi}_1^T, \dots, \boldsymbol{\xi}_N^T]^T, \quad (2)$$

where $\boldsymbol{\xi}_n = [\xi_{1n}, \dots, \xi_{Mn}]^T$, and $\xi_{mn} = \beta_{mn} \sqrt{E_m} / (R_{t,m} R_{r,n})$. Define the covariance matrix of $\boldsymbol{\xi}$ as $\mathbb{E}\{\boldsymbol{\xi}\boldsymbol{\xi}^H\} = \boldsymbol{\Lambda}$, which is a positive definite Hermitian matrix. Note that for correlated target reflection coefficients, $\boldsymbol{\Lambda}$ is non-diagonal, while for spatially uncorrelated target reflection coefficients, $\boldsymbol{\Lambda}$ is diagonal. From (1), the target detection problem can be formulated as

$$\mathcal{H}_0 : r_n(t) = w_n(t), \quad t \in \mathcal{T}, \quad (3)$$

$$\mathcal{H}_1 : r_n(t) = \sum_{m=1}^M \xi_{mn} s_m(t - \tau_{mn}) + w_n(t), \quad t \in \mathcal{T}. \quad (4)$$

Assuming the MN received waveforms $s_m(t - \tau_{mn})$ ($m = 1, \dots, M$ and $n = 1, \dots, N$) are linearly independent (extensions are possible), the signals can be represented by their components in terms of a set of basis functions. The log-likelihood ratio (LLR) that considers these components as observations is given by (see Appendix A for complete derivation)

$$\mathcal{L} = C + \mathbf{x}^H ((\mathbf{N} \otimes \boldsymbol{\Xi})^{-1} - (\mathbf{N} \otimes \boldsymbol{\Xi} + \boldsymbol{\Psi} \boldsymbol{\Lambda} \boldsymbol{\Psi}^H)^{-1}) \mathbf{x}, \quad (5)$$

and the test statistic is given by

$$T = \mathbf{x}^H ((\mathbf{N} \otimes \boldsymbol{\Xi})^{-1} - (\mathbf{N} \otimes \boldsymbol{\Xi} + \boldsymbol{\Psi} \boldsymbol{\Lambda} \boldsymbol{\Psi}^H)^{-1}) \mathbf{x}. \quad (6)$$

where $\otimes$ denotes Kronecker product, the constant $C = \ln(\text{Det}(\mathbf{N} \otimes \boldsymbol{\Xi})) - \ln(\text{Det}(\mathbf{N} \otimes \boldsymbol{\Xi} + \boldsymbol{\Psi} \boldsymbol{\Lambda} \boldsymbol{\Psi}^H))$, in which $\text{Det}(\cdot)$ denotes determinant, $\boldsymbol{\Xi}$ is an $MN \times MN$ matrix with the $((n_1 - 1)M + m_1, (n_2 - 1)M + m_2)$ -th element $(n_1, n_2 = 1, \dots, N$ and $m_1, m_2 = 1, \dots, M)$ given by

$$\begin{aligned} & \Xi_{(n_1-1)M+m_1, (n_2-1)M+m_2} \\ &= \int_{\mathcal{T}} s_{m_1}^*(t - \tau_{m_1 n_1}) s_{m_2}(t - \tau_{m_2 n_2}) dt, \end{aligned} \quad (7)$$

¹The case with imperfectly estimated waveforms is considered in numerical experiments.

and $\Psi = \text{Diag}\{\Psi_1, \dots, \Psi_N\}$, where Ψ_n is an $MN \times M$ matrix whose m -th column is the $((n-1)M + m)$ -th column of Ξ . In (6), the $MN^2 \times 1$ complex Gaussian vector

$$\mathbf{x} = [\mathbf{x}_1^T, \dots, \mathbf{x}_N^T]^T \quad (8)$$

collects all the MF outputs, where $\mathbf{x}_n = [x_{11n}, \dots, x_{MNn}]^T$ is an $MN \times 1$ vector with the $((n'-1)M + m)$ -th element ($m = 1, \dots, M$ and $n' = 1, \dots, N$) given by $x_{mn'n} = \int_{\mathcal{T}} s_m^*(t - \tau_{mn'}) r_n(t) dt$.

III. LIMITED-COMPLEXITY RECEIVER DESIGN

From (6), we see that the test statistic and hence the detection performance is dependent on the received signals only via the MF output vector $\mathbf{x}$. The size of $\mathbf{x}$ determines the complexity of the associated hardware or software. We propose to select a subset of the vector $\mathbf{x}$ for subsequent processing to reduce complexity. Before proceeding, define a selection vector

$$\mathbf{a} \triangleq [\mathbf{a}_1^T, \dots, \mathbf{a}_N^T]^T, \quad (9)$$

where $\mathbf{a}_n = [a_{11n}, \dots, a_{MNn}]^T$, in which $a_{mn'n} \in \{1, 0\}$ indicating whether or not the signal associated with the (m, n') -th transmitter to receiver path is processed at the n -th receiver. Define a selection matrix

$$\mathbf{J}(\mathbf{a}) \triangleq \text{Diag}\{\mathbf{J}_1(\mathbf{a}_1), \dots, \mathbf{J}_N(\mathbf{a}_N)\} \quad (10)$$

where $\text{Diag}\{\cdot\}$ denotes a block diagonal matrix and

$$\mathbf{J}_n(\mathbf{a}_n) \triangleq \text{diag}_r\{\mathbf{a}_n\}, \quad (11)$$

in which $\text{diag}_r\{\cdot\}$ represents a diagonal matrix with the argument on its diagonal, but with the all-zero rows removed [37], [38]. The size of $\mathbf{J}_n(\mathbf{a}_n)$ is $u_n \times MN$, where

$$u_n = \|\mathbf{a}_n\|_0 \quad (12)$$

is the number of paths to be processed at receiver n and $\|\cdot\|_0$ denotes the ℓ_0 -norm operator. Assume that at least one path is processed at each receiver, such that $u_n \geq 1$ for all n . For a given selection, the MFs corresponding to the zero elements in $\mathbf{a}$ are no longer needed and the associated hardware or software can be saved. Accordingly, the MF output vector is reduced from $\mathbf{x}$ to $\mathbf{J}(\mathbf{a})\mathbf{x}$. Then the test statistic is changed from (6) to

$$T_s = (\mathbf{J}(\mathbf{a})\mathbf{x})^H \left((\mathbf{J}(\mathbf{a})\Sigma_0\mathbf{J}^T(\mathbf{a}))^{-1} - (\mathbf{J}(\mathbf{a})\Sigma_1\mathbf{J}^T(\mathbf{a}))^{-1} \right) \mathbf{J}(\mathbf{a})\mathbf{x}, \quad (13)$$

where

$$\Sigma_0 = \mathbb{E}\{\mathbf{x}\mathbf{x}^H | \mathcal{H}_0\} = \mathbf{N} \otimes \Xi \quad (14)$$

and

$$\Sigma_1 = \mathbb{E}\{\mathbf{x}\mathbf{x}^H | \mathcal{H}_1\} = \mathbf{N} \otimes \Xi + \Psi\Lambda\Psi^H. \quad (15)$$

For the special case, where signals associated with all paths are processed, (13) is equivalent to the test statis-

tic in (6). From (13), we see that $\sum_{n=1}^N u_n$ MFs, $(1 + \sum_{n=1}^N u_n) \sum_{n=1}^N u_n$ multipliers, and $(1 + \sum_{n=1}^N u_n)(-1 + \sum_{n=1}^N u_n)$ adders are required.

The test statistic T_s is compared to a threshold γ , such that a decision for $\mathcal{H}_1$ is made if $T_s > \gamma$ and $\mathcal{H}_0$ otherwise. Suppose, to limit the cost, receiver n can at most process signals associated with A_n ($A_n \leq MN$) paths, namely $u_n \leq A_n$. If the Neyman-Pearson criterion is employed and $\gamma(P_{FA}, \mathbf{a})$ is chosen to provide a false alarm probability P_{FA} when a particular $\mathbf{a}$ is employed, the optimal selection can be obtained by solving the following optimization problem

$$\mathbf{P}_1 \begin{cases} \max_{\mathbf{a} \in \{0,1\}^{MN^2}} & Pr(T_s > \gamma(P_{FA}, \mathbf{a}) | \mathcal{H}_1) \\ \text{s.t.} & 1 \leq u_n \leq A_n, n = 1, \dots, N \end{cases} \quad (16a)$$

The solution of $\mathbf{P}_1$ provides guidance to system designers on how to maximize detection performance with limited budget for a general case where the target reflection coefficients and the clutter-plus-noise components can be correlated. For simplification and to get insight, next we concentrate on the case of spatially uncorrelated coefficients and spatially white clutter-plus-noise.

IV. TRANSMITTER SELECTION FOR PASSIVE/ACTIVE MIMO RADAR

In this section, we consider the case under the assumption of spatially uncorrelated target reflection coefficients and spatially white clutter-plus-noise components. In this case, the previously discussed limited-complexity receiver is equivalent to the transmitter selection problem for passive/active radar, as shown in the sequel.

For spatially uncorrelated target reflection coefficients, the covariance matrix Λ of ξ in (2) can be written as

$$\Lambda = \text{Diag}\{\Lambda_1, \dots, \Lambda_N\}, \quad (17)$$

where

$$\Lambda_n = \text{diag}\left\{\frac{\sigma_{1n}^2 E_1}{(R_{t,1} R_{r,n})^2}, \dots, \frac{\sigma_{Mn}^2 E_M}{(R_{t,M} R_{r,n})^2}\right\}. \quad (18)$$

in which $\text{diag}\{\cdot\}$ denotes a diagonal matrix. For spatially white clutter-plus-noise, the covariance matrix defined after (1) can be reduced to

$$\mathbf{N} = \text{diag}\{N_{11}, N_{22}, \dots, N_{NN}\} \triangleq N_0 \mathbf{I}_N, \quad (19)$$

where $N_{ii} = N_0$, $i = 1, \dots, N$ is the power spectral density (PSD) of $w_i(t)$ and $\mathbf{I}_N$ is a $N \times N$ identity matrix.

Substituting (17), (18) and (19) into (6), the test statistic becomes²

$$T = \frac{1}{N_0} \sum_{n=1}^N \mathbf{x}_n^H (\Xi_n + N_0 \Lambda_n^{-1})^{-1} \mathbf{x}_n, \quad (20)$$

²Even if the MN received waveforms $s_m(t - \tau_{mn})$ are not linear independent, the test statistic in (20) is also make sense because the matrix $\Xi_n + N_0 \Lambda_n^{-1}$ in (20) is always invertible.

where Ξ_n is an $M \times M$ matrix with the (i, j) -th element given by $\Xi_{n,(i,j)} = \int_{\mathcal{T}} s_i^*(t - \tau_{in}) s_j(t - \tau_{jn}) dt$ and

$$\mathbf{x}_n = [x_{1n}, \dots, x_{Mn}]^T \quad (21)$$

is now reduced to $M \times 1$ dimensional, in which

$$x_{mn} = \int_{\mathcal{T}} s_m^*(t - \tau_{mn}) r_n(t) dt. \quad (22)$$

Therefore, in this case the MF output vector in (8) is reduced to $MN \times 1$ dimensional, given by

$$\mathbf{x} = [\mathbf{x}_1^T, \dots, \mathbf{x}_N^T]^T = [x_{11}, x_{21}, \dots, x_{MN}]^T. \quad (23)$$

It is seen from (21) that at the n -th receiver, the MF output vector $\mathbf{x}_n$ has M elements, each corresponding to the contribution due to a given transmitter. Thus, the selection of a subset of $\mathbf{x} = [\mathbf{x}_1^T, \dots, \mathbf{x}_N^T]^T$ implies the selection of transmitters at each of the receivers. Accordingly, we redefine

$$\mathbf{a}_n = [a_{1n}, \dots, a_{Mn}]^T \quad (24)$$

and the overall selection vector in (9) is now reduced to an $MN \times 1$ transmitter selection vector,

$$\mathbf{a} = [\mathbf{a}_1^T, \dots, \mathbf{a}_N^T]^T = [a_{11}, a_{21}, \dots, a_{MN}]^T, \quad (25)$$

where $a_{mn} \in \{0, 1\}$ indicates whether or not transmitter m is selected for processing at receiver n .

The test statistic after transmitter selection can be written as

$$T_s = \sum_{n=1}^N (\mathbf{J}_n(\mathbf{a}_n) \mathbf{x}_n)^H \left((\mathbf{J}_n(\mathbf{a}_n) \Sigma_{0,n} \mathbf{J}_n^T(\mathbf{a}_n))^{-1} - (\mathbf{J}_n(\mathbf{a}_n) \Sigma_{1,n} \mathbf{J}_n^T(\mathbf{a}_n))^{-1} \right) \mathbf{J}_n(\mathbf{a}_n) \mathbf{x}_n, \quad (26)$$

where the selection matrix

$$\mathbf{J}_n(\mathbf{a}_n) = \text{diag}_r \{a_{1n}, \dots, a_{Mn}\} \quad (27)$$

is $u_n \times M$ dimensional,

$$\Sigma_{0,n} = \mathbb{E}\{\mathbf{x}_n \mathbf{x}_n^H | \mathcal{H}_0\} = N_0 \Xi_n, \quad (28)$$

and

$$\Sigma_{1,n} = \mathbb{E}\{\mathbf{x}_n \mathbf{x}_n^H | \mathcal{H}_1\} = \Xi_n \mathbf{\Lambda}_n \Xi_n^H + N_0 \Xi_n. \quad (29)$$

Next, we first assume the waveforms are orthogonal and provide a closed-form solution to the optimization problem $\mathbf{P}_1$. Then, for non-orthogonal waveforms, sub-optimal solutions based on KL distance are presented.

A. Orthogonal Waveforms

Assume the transmitted signals are mutually orthogonal and maintain orthogonality for any delay τ of interest, namely,

$$\int_{\mathcal{T}} s_{m_1}(t) s_{m_2}^*(t - \tau) dt = 0, \text{ for } m_1 \neq m_2, \forall \tau. \quad (30)$$

Thus, $\Xi_n = \mathbf{I}_M$ and the test statistic in (26) can be written as

$$T_s = \frac{1}{N_0} \sum_{n=1}^N (\mathbf{J}_n(\mathbf{a}_n) \mathbf{x}_n)^H \left[\mathbf{J}_n(\mathbf{a}_n) (N_0 \mathbf{\Lambda}_n^{-1} + \mathbf{I}_M)^{-1} \mathbf{J}_n^T(\mathbf{a}_n) \right] \mathbf{J}_n(\mathbf{a}_n) \mathbf{x}_n, \quad (31)$$

$$= \sum_{n=1}^N \sum_{m=1}^M \frac{E_m \sigma_{mn}^2 a_{mn}}{N_0 (E_m \sigma_{mn}^2 + N_0 (R_{t,m} R_{r,n})^2)} |x_{mn}|^2. \quad (32)$$

$$= \sum_{n=1}^N \sum_{m=1}^M \frac{E_m \sigma_{mn}^2 a_{mn}}{N_0 (E_m \sigma_{mn}^2 a_{mn} + N_0 (R_{t,m} R_{r,n})^2)} |x_{mn}|^2, \quad (33)$$

where the last equality is obtained considering that the a_{mn} in the numerator of (32) equals either 0 or 1, hence adding an a_{mn} in the denominator as per (33) does not change the result. Define the signal-to-clutter-plus-noise ratio (SCNR) of the (m, n) -th path as $\eta_{mn} = E_m \sigma_{mn}^2 / N_0 (R_{t,m} R_{r,n})^2$.

Then, (33) can be rewritten as a function of the SCNRs as follows

$$T_s = \sum_{n=1}^N \sum_{m=1}^M \zeta_{mn}, \quad (34)$$

where $\zeta_{mn} = \frac{\rho_{mn}}{N_0(\rho_{mn}+1)} |x_{mn}|^2$, and $\rho_{mn} = \eta_{mn} a_{mn}$ is non-negative. For later use, define $\boldsymbol{\rho}_n = [\rho_{1n}, \dots, \rho_{Mn}]^T$ the transmitter selection vector weighted by SCNRs, $\boldsymbol{\rho} = [\boldsymbol{\rho}_1^T, \dots, \boldsymbol{\rho}_N^T]^T$, and $\boldsymbol{\eta} = [\eta_{11}, \eta_{21}, \dots, \eta_{MN}]^T$.

Denote the cumulative distribution function (cdf) of ζ_{mn} under $\mathcal{H}_i$ ($i = 0$ or 1) by $F_{\zeta_{mn}|\mathcal{H}_i}(z, \rho_{mn}) = \Pr(\zeta_{mn} \leq z, \mathcal{H}_i)$. It is easy to see that when $\rho_{mn} = 0$,

$$F_{\zeta_{mn}|\mathcal{H}_i}(z, \rho_{mn} = 0) = \begin{cases} 1 & z \geq 0 \\ 0 & z < 0 \end{cases} \quad (35)$$

Recall from (22) that $x_{mn} \sim \mathcal{CN}(0, N_0)$ under $\mathcal{H}_0$ and $x_{mn} \sim \mathcal{CN}(0, N_0(1 + \eta_{mn}))$ under $\mathcal{H}_1$. Thus, when $\rho_{mn} > 0$, it can be shown that $\zeta_{mn} \sim \frac{\rho_{mn}}{2(\rho_{mn}+1)} \chi_2^2$ under $\mathcal{H}_0$ and $\zeta_{mn} \sim \frac{\rho_{mn}}{2} \chi_2^2$ under $\mathcal{H}_1$, so the corresponding $F_{\zeta_{mn}|\mathcal{H}_i}(z, \rho_{mn})$ can be obtained based on the cumulative distribution function (cdf) of the Chi-squared χ_2^2 distribution.

Hence, the cdf of the test statistic in (34) under hypothesis $\mathcal{H}_i$ can be calculated by [39]

$$F_{T_s|\mathcal{H}_i}(z, \boldsymbol{\rho}) = \Pr(T_s \leq z | \mathcal{H}_i) = F_{\zeta_{11}|\mathcal{H}_i}(z, \rho_{11}) * F_{\zeta_{21}|\mathcal{H}_i}(z, \rho_{21}) * \dots * F_{\zeta_{MN}|\mathcal{H}_i}(z, \rho_{MN}) \quad (36)$$

$$= \int_{\mathbb{R}^{MN-1}} F_{\zeta_{11}|\mathcal{H}_i}(z - z_{21} - \dots - z_{MN}, \rho_{11}) dF_{\zeta_{21}|\mathcal{H}_i}(z_{21}, \rho_{21}) \dots dF_{\zeta_{MN}|\mathcal{H}_i}(z_{MN}, \rho_{MN}) \quad (37)$$

where $*$ denotes the convolution operator³, and $\mathbb{R}^{MN-1}$ means the $(MN - 1)$ -dimensional real space since there are $(MN - 1)$ convolution operators in (36).

Lemma 1. Denote by $\rho_{(1)}, \rho_{(2)}, \dots, \rho_{(MN)}$ the decreasing sequence of nonnegative $\rho_{11}, \rho_{21}, \dots, \rho_{MN}$ and define $\bar{\boldsymbol{\rho}} = [\rho_{(1)}, \rho_{(2)}, \dots, \rho_{(MN)}]^T$. Let $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ be two feasible solutions for $\mathbf{P}_1$, and correspondingly $\boldsymbol{\rho}_{\boldsymbol{\alpha}} = \boldsymbol{\alpha} \odot \boldsymbol{\eta}$ and $\boldsymbol{\rho}_{\boldsymbol{\beta}} = \boldsymbol{\beta} \odot \boldsymbol{\eta}$, where $\odot$ denotes Hadamard product. If $\bar{\boldsymbol{\rho}}_{\boldsymbol{\alpha}} \succeq \bar{\boldsymbol{\rho}}_{\boldsymbol{\beta}}$, where ' $\succeq$ '

³The convolution between two cdfs $F_1(z)$ and $F_2(z)$ is defined as $F_1(z) * F_2(z) = \int_{-\infty}^{\infty} F_1(z - y) dF_2(y)$.

means element-wise no less than⁴, then, for the hypothesis testing problem characterized by the test statistic in (34), the detection probability under Neyman-Pearson criterion satisfies

$$P_D(\rho_\alpha) \geq P_D(\rho_\beta) \quad (38)$$

where $P_D(\rho) = \Pr(T_s > \gamma | \mathcal{H}_1)$ and γ is determined by the required level of false alarm probability P_{FA} and the weighted transmitter selection vector ρ .

Proof. It can be proved (see Appendix B) that $P_D(\rho)$ is a strictly monotone increasing function with respect to any element ρ_{mn} ($m = 1, \dots, M, n = 1, \dots, N$) of ρ when all the other elements are fixed.

From (36), according to the commutative law of convolution,

$$\begin{aligned} F_{T_s|\mathcal{H}_i}(z, \rho_\alpha) &= F_{\zeta_{11}|\mathcal{H}_i}(z, \rho_{\alpha 11}) * \dots * F_{\zeta_{MN}|\mathcal{H}_i}(z, \rho_{\alpha MN}) \\ &= F_{\zeta_{11}|\mathcal{H}_i}(z, \rho_{\alpha(1)}) * \dots * F_{\zeta_{MN}|\mathcal{H}_i}(z, \rho_{\alpha(MN)}) \\ &= F_{T_s|\mathcal{H}_i}(z, \overline{\rho_\alpha}), \end{aligned} \quad (39)$$

where $i = 0$ or 1 . Then

$$P_{FA} = 1 - F_{T_s|\mathcal{H}_0}(\gamma, \rho_\alpha) = 1 - F_{T_s|\mathcal{H}_0}(\gamma, \overline{\rho_\alpha}). \quad (40)$$

From (40), $\gamma(P_{FA}, \rho_\alpha) = \gamma(P_{FA}, \overline{\rho_\alpha})$, and

$$\begin{aligned} P_D(\rho_\alpha) &= 1 - F_{T_s|\mathcal{H}_1}(\gamma(P_{FA}, \rho_\alpha), \rho_\alpha) \\ &= 1 - F_{T_s|\mathcal{H}_1}(\gamma(P_{FA}, \overline{\rho_\alpha}), \overline{\rho_\alpha}) = P_D(\overline{\rho_\alpha}). \end{aligned} \quad (41)$$

Similarly,

$$P_D(\rho_\beta) = P_D(\overline{\rho_\beta}). \quad (42)$$

Because $\rho_{\alpha(k)} \geq \rho_{\beta(k)}$, for all k , $k = 1, \dots, MN$, then, based on the monotonicity of $P_D(\rho)$, it is clear that

$$P_D(\overline{\rho_\alpha}) \geq P_D(\overline{\rho_\beta}). \quad (43)$$

Then from (41), (42) and (43),

$$P_D(\rho_\alpha) = P_D(\overline{\rho_\alpha}) \geq P_D(\overline{\rho_\beta}) = P_D(\rho_\beta) \quad (44)$$

which completes the proof. $\square$

From Lemma 1, the following conclusion follows.

Theorem 1. Under the assumption of uncorrelated target reflection coefficients, spatially white clutter-plus noise, and orthogonal transmitted waveforms, the optimal solution of $\mathbf{P}_1$ can be obtained by selecting the transmitters corresponding to the A_n largest⁵ SCNRs at receiver n , $n = 1, \dots, N$, i.e.,

$$a_{mn}^* = \begin{cases} 1, & m \in \mathcal{S}_n \\ 0, & \text{else,} \end{cases} \quad (45)$$

where $\mathcal{S}_n$ denotes the index set of m associated with the A_n largest η_{mn} in $\mathcal{I}_n = \{\eta_{1n}, \dots, \eta_{Mn}\}$.

⁴When $\overline{\rho_\alpha} \succeq \overline{\rho_\beta}$, the k -th element $\rho_{\alpha(k)}$ of $\overline{\rho_\alpha}$ is no less than the k -th element $\rho_{\beta(k)}$ of $\overline{\rho_\beta}$ for all k , $k = 1, \dots, MN$.

⁵If two transmitters, say the m_1 -th and m_2 -th transmitter ($m_1 < m_2$), lead to the same SCNR $\eta_{m_1 n} = \eta_{m_2 n}$, then the one with lower index is selected first such that we select the m_1 -th transmitter before the m_2 -th transmitter.

B. General Waveforms

This section generalizes the discussion to possibly non-orthogonal transmitted waveforms. First, we consider a special case where each receiver selects one transmitter, a closed-form solution is presented. Further, when each receiver selects an arbitrary number of transmitters, KL distance as a suboptimal criterion is employed.

1) *General Waveforms with $A_n=1$:* First, we consider the case where $A_n = 1$ for $n = 1, \dots, N$. For each n , there is only one non-zero element in the set $\{a_{1n}, \dots, a_{Mn}\}$, leading to $u_n = 1$ and the $\mathbf{J}_n(\mathbf{a}_n)$ in (27) becomes an $1 \times M$ dimensional row vector. Thus, the terms in (26) can be simplified to

$$\mathbf{J}_n(\mathbf{a}_n)\mathbf{x}_n = \sum_{m=1}^M a_{mn}x_{mn} \quad (46)$$

and from (24) and (27), recalling that a_{mn} equals either 0 or 1, we have

$$\begin{aligned} &\left[\mathbf{J}_n(\mathbf{a}_n)\boldsymbol{\Sigma}_{0,n}\mathbf{J}_n^T(\mathbf{a}_n) \right]^{-1} - \left[\mathbf{J}_n(\mathbf{a}_n)\boldsymbol{\Sigma}_{1,n}\mathbf{J}_n^T(\mathbf{a}_n) \right]^{-1} \\ &= \left(\sum_{m=1}^M \Sigma_{0,n,(m,m)} a_{mn} \right)^{-1} - \left(\sum_{m=1}^M \Sigma_{1,n,(m,m)} a_{mn} \right)^{-1} \\ &= \sum_{m=1}^M \left(\frac{1}{\Sigma_{0,n,(m,m)}} - \frac{1}{\Sigma_{1,n,(m,m)}} \right) a_{mn} \end{aligned} \quad (47)$$

where $\Sigma_{i,n,(m,m)}$ ($i = 0$ or 1) denotes the (m,m) -th element of the matrix $\boldsymbol{\Sigma}_{i,n}$. From (18), (28) and (29),

$$\begin{aligned} \Sigma_{0,n,(m,m)} &= N_0 \Xi_{n,(mm)} \\ &= N_0 \int_{\mathcal{T}} s_m^*(t - \tau_{mn}) s_m(t - \tau_{mn}) dt \\ &= N_0 \end{aligned} \quad (48)$$

and

$$\begin{aligned} \Sigma_{1,n,(m,m)} &= N_0 \Xi_{n,(m,m)} + \sum_{i=1}^M \sum_{j=1}^M \Xi_{n,(m,i)} \Lambda_{n,(i,j)} \Xi_{n,(j,m)} \\ &= N_0 + \sum_{i=1}^M \frac{E_i \sigma_{in}^2 |\Xi_{n,(i,m)}|^2}{(R_{t,i} R_{r,n})^2}. \end{aligned} \quad (49)$$

where $\Xi_{n,(i,j)}$ and $\Lambda_{n,(i,j)}$ denote the (i,j) -th element of $\boldsymbol{\Xi}_n$ and $\boldsymbol{\Lambda}_n$, respectively. Substituting (46)-(49) into (26), we obtain

$$\begin{aligned} T_s &= \sum_{n=1}^N \left(\sum_{m=1}^M x_{mn}^* a_{mn} \sum_{m=1}^M \frac{\varrho_{mn} a_{mn}}{N_0(1 + \varrho_{mn})} \sum_{m=1}^M x_{mn} a_{mn} \right) \\ &= \sum_{n=1}^N \sum_{m=1}^M \frac{\varrho_{mn} a_{mn}}{N_0(1 + \varrho_{mn})} |x_{mn}|^2 \\ &= \sum_{n=1}^N \sum_{m=1}^M \frac{\varrho_{mn} a_{mn}}{N_0(1 + \varrho_{mn} a_{mn})} |x_{mn}|^2, \end{aligned} \quad (50)$$

where $\varrho_{mn} = \sum_{i=1}^M \frac{E_i \sigma_{in}^2 |\Xi_{n,(i,m)}|^2}{N_0 (R_{t,i} R_{r,n})^2}$. From (22), we see that for the case of general waveform, $x_{mn} \sim \mathcal{CN}(0, N_0)$ under $\mathcal{H}_0$ and $x_{mn} \sim \mathcal{CN}(0, (1 + \varrho_{mn})N_0)$ under $\mathcal{H}_1$. The terms x_{mn} are mutually independent for different n .

Recall that $A_n = 1$, so the correlation between different MF outputs at receiver n can be ignored. Note that the only difference between (34) and (50) is that all the terms η_{mn} associated with (34) are changed to ϱ_{mn} . Hence, by changing the term η_{mn} in Theorem 1 to ϱ_{mn} , the optimal solution for the case $A_1 = A_2 = \dots, A_N = 1$ is given by

$$a_{mn}^* = \begin{cases} 1, & m \in \mathcal{S}_n \\ 0, & \text{else,} \end{cases} \quad (51)$$

where $\mathcal{S}_n = \arg \max_m \{\varrho_{mn}\}$, $n = 1, \dots, N$.

2) *General Waveform with general A_n* : Now we study a more general case where $A_n > 1$ for certain n . Since for this case, analytically evaluating detection probability is intractable, it is difficult to solve $\mathbf{P}_1$. An alternative way is to change the objective function from the detection probability to some other heuristic criteria [12], [37]. Based on Stein's lemma, for a fixed value P_{FA} [37], [38],

$$\ln(P_M(\mathbf{a})) \xrightarrow{a.s.} -D_{KL}(\mathbf{a}), \text{ for } P_M(\mathbf{a}) \rightarrow 0 \quad (52)$$

where $\xrightarrow{a.s.}$ means converges almost surely. $P_M(\mathbf{a}) = 1 - Pr(T_s > \gamma(P_{FA}, \mathbf{a}) | \mathcal{H}_1)$ denotes the probability of miss and

$$D_{KL}(\mathbf{a}) = \int \ln \left(\frac{p_{\mathbf{y}|\mathcal{H}_1}(\mathbf{y})}{p_{\mathbf{y}|\mathcal{H}_0}(\mathbf{y})} \right) p_{\mathbf{y}|\mathcal{H}_1}(\mathbf{y}) d\mathbf{y} \quad (53)$$

denotes the KL distance between $p_{\mathbf{y}|\mathcal{H}_1}(\mathbf{y})$ and $p_{\mathbf{y}|\mathcal{H}_0}(\mathbf{y})$, in which

$$\mathbf{y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_N^T]^T \quad (54)$$

and $\mathbf{y}_n = \mathbf{J}_n(\mathbf{a}_n)\mathbf{x}_n$. So the KL-distance is a reasonable metric to characterise the detection performance in the Neyman-Pearson settings. Next, we use the KL-distance as the alternate to facilitate the evaluation and optimization.

It is obvious that $\mathbf{y}_n$, $n = 1, \dots, N$ are mutually independent for different n under both $\mathcal{H}_0$ and $\mathcal{H}_1$. Therefore

$$p_{\mathbf{y}|\mathcal{H}_i}(\mathbf{y}) = \prod_{n=1}^N p_{\mathbf{y}_n|\mathcal{H}_i}(\mathbf{y}_n) \quad (55)$$

where $p_{\mathbf{y}|\mathcal{H}_i}(\mathbf{y})$ and $p_{\mathbf{y}_n|\mathcal{H}_i}(\mathbf{y}_n)$ are the probability density functions (pdfs) of $\mathbf{y}$ and $\mathbf{y}_n$ ($n = 1, \dots, N$) under $\mathcal{H}_i$ ($i = 0, 1$), respectively. Thus, the KL distance between $p_{\mathbf{y}|\mathcal{H}_1}(\mathbf{y})$ and $p_{\mathbf{y}|\mathcal{H}_0}(\mathbf{y})$ in (53) can be written as

$$\begin{aligned} D^{KL}(\mathbf{a}) &= \int \left(\sum_{n=1}^N \ln \left(\frac{p_{\mathbf{y}_n|\mathcal{H}_1}(\mathbf{y}_n)}{p_{\mathbf{y}_n|\mathcal{H}_0}(\mathbf{y}_n)} \right) \right) \prod_{n=1}^N p_{\mathbf{y}_n|\mathcal{H}_1}(\mathbf{y}_n) d\mathbf{y} \\ &= \sum_{n=1}^N \left(\int \ln \left(\frac{p_{\mathbf{y}_n|\mathcal{H}_1}(\mathbf{y}_n)}{p_{\mathbf{y}_n|\mathcal{H}_0}(\mathbf{y}_n)} \right) p_{\mathbf{y}_n|\mathcal{H}_1}(\mathbf{y}_n) d\mathbf{y}_n \right. \\ &\quad \times \left. \prod_{k \neq n} \int p_{\mathbf{y}_k|\mathcal{H}_1}(\mathbf{y}_k) d\mathbf{y}_k \right) \\ &= \sum_{n=1}^N \int \ln \left(\frac{p_{\mathbf{y}_n|\mathcal{H}_1}(\mathbf{y}_n)}{p_{\mathbf{y}_n|\mathcal{H}_0}(\mathbf{y}_n)} \right) p_{\mathbf{y}_n|\mathcal{H}_1}(\mathbf{y}_n) d\mathbf{y}_n \\ &= \sum_{n=1}^N D_n^{KL}(\mathbf{a}_n), \end{aligned} \quad (56)$$

where

$$D_n^{KL}(\mathbf{a}_n) = \int \ln \frac{p_{\mathbf{y}_n|\mathcal{H}_1}(\mathbf{y}_n)}{p_{\mathbf{y}_n|\mathcal{H}_0}(\mathbf{y}_n)} p_{\mathbf{y}_n|\mathcal{H}_1}(\mathbf{y}_n) d\mathbf{y}_n \quad (57)$$

is the KL distance between the two distributions $p_{\mathbf{y}_n|\mathcal{H}_1}(\mathbf{y}_n)$ and $p_{\mathbf{y}_n|\mathcal{H}_0}(\mathbf{y}_n)$. The distributions of $\mathbf{y}_n$, $n = 1, 2, \dots, N$, under the two hypotheses are

$$\mathbf{y}_n | \mathcal{H}_0 \sim \mathcal{CN}(\mathbf{0}, \mathbf{J}(\mathbf{a}_n) \boldsymbol{\Sigma}_{0,n} \mathbf{J}^T(\mathbf{a}_n)) \quad (58)$$

and

$$\mathbf{y}_n | \mathcal{H}_1 \sim \mathcal{CN}(\mathbf{0}, \mathbf{J}(\mathbf{a}_n) \boldsymbol{\Sigma}_{1,n} \mathbf{J}^T(\mathbf{a}_n)), \quad (59)$$

$n = 1, 2, \dots, N$, where $\boldsymbol{\Sigma}_{0,n}$ and $\boldsymbol{\Sigma}_{1,n}$ are given in (28) and (29). Inserting (58) and (59) into (57), we obtain [40]

$$\begin{aligned} D_n^{KL}(\mathbf{a}_n) &= \text{Tr} \left(\mathbf{J}_n(\mathbf{a}_n) \boldsymbol{\Sigma}_{1,n} \mathbf{J}_n^T(\mathbf{a}_n) (\mathbf{J}_n(\mathbf{a}_n) \boldsymbol{\Sigma}_{0,n} \mathbf{J}_n^T(\mathbf{a}_n))^{-1} \right) \\ &\quad - \ln \frac{\text{Det}(\mathbf{J}_n(\mathbf{a}_n) \boldsymbol{\Sigma}_{1,n} \mathbf{J}_n^T(\mathbf{a}_n))}{\text{Det}(\mathbf{J}_n(\mathbf{a}_n) \boldsymbol{\Sigma}_{0,n} \mathbf{J}_n^T(\mathbf{a}_n))} - u_n \end{aligned} \quad (60)$$

where $\text{Tr}(\cdot)$ denotes trace.

The KL-optimal transmitter selection can be obtained by solving the optimization problem

$$\mathbf{P}_2 \begin{cases} \max_{\mathbf{a} \in \{0,1\}^{MN}} D^{KL}(\mathbf{a}) \\ \text{s.t.} \quad 1 \leq u_n \leq A_n, n = 1, 2, \dots, N \end{cases} \quad (61)$$

From (56) and (61), we see that solving $\mathbf{P}_2$ is equivalent to solving the following N optimization problems. For $n = 1, \dots, N$,

$$\max_{\mathbf{a}_n \in \{0,1\}^M} D_n^{KL}(\mathbf{a}_n) \quad \text{s.t.} \quad 1 \leq u_n \leq A_n. \quad (62)$$

Considering that, from (11),

$$\mathbf{J}_n^T(\mathbf{a}_n) \mathbf{J}_n(\mathbf{a}_n) = \text{diag}\{a_{1n}, \dots, a_{Mn}\}, \quad (63)$$

we change the optimization variable in (62) from $\mathbf{a}_n$ to $\mathbf{J}_n$, where $\mathbf{J}_n$ has one unit element per row and all the other elements are zeros such that $\mathbf{J}_n \mathbf{J}_n^T = \mathbf{I}_{u_n}$. Then, an alternative optimization problem is obtained

$$\max_{\mathbf{J}_n} D_n^{KL}(\mathbf{J}_n) \quad (64a)$$

$$\text{s.t.} \quad \mathbf{J}_n \mathbf{J}_n^T = \mathbf{I}_{u_n} \quad (64b)$$

$$\mathbf{J}_n \in \{0, 1\}^{u_n \times M} \quad (64c)$$

$$1 \leq u_n \leq A_n, \quad (64d)$$

where $n = 1, \dots, N$.

Since solving (64) is NP hard [40], we employ a greedy-based method [29], [34]–[36] to find a suboptimal solution. In the greedy-based method, the rows of the selection matrix $\mathbf{J}_n$ are determined one by one. Initially, we consider selecting just a single row of $\mathbf{J}_n$ and use (64a) to compute

$$\tilde{j}_1 = \arg \max_{j_1 \in \Omega} D_n^{KL}(\mathbf{e}_{j_1}), \quad (65)$$

where $\Omega = \{1, 2, \dots, M\}$ and $\mathbf{e}_{j_1}$ is the j_1 -th row of the identity matrix $\mathbf{I}_M$. Once $\tilde{j}_1$ is obtained, we set $\mathbf{J}_n = \mathbf{e}_{\tilde{j}_1}$ and $\Omega = \Omega \setminus \tilde{j}_1$. Then, we move on to select the second

row of $\mathbf{J}_n$ and use (64a) again to compute

$$\tilde{j}_2 = \arg \max_{j_2 \in \Omega} D_n^{KL}([\mathbf{J}_n^T, \mathbf{e}_{j_2}^T]^T). \quad (66)$$

Once $\tilde{j}_2$ is obtained, we set $\mathbf{J}_n = [\mathbf{J}_n^T, \tilde{\mathbf{e}}_{\tilde{j}_2}^T]^T$ and $\Omega = \Omega \setminus \tilde{j}_2$. Next, we select the third row of $\mathbf{J}_n$ to get $\tilde{j}_3$, so on and so forth, up to the A_n -th step. The greedy-based algorithm is summarized in Algorithm 1.

Algorithm 1 Greedy-Based Algorithm

Input: $\Sigma_{0,n}$, $\Sigma_{1,n}$ and A_n , $n = 1, \dots, N$.

For $n = 1 : N$ **do**

Set $\Omega = \{1, 2, \dots, M\}$ and use (64a) to compute

$$\tilde{j}_1 = \arg \max_{j_1 \in \Omega} D_n^{KL}(\mathbf{e}_{j_1}).$$

Set $\mathbf{J}_n = \mathbf{e}_{\tilde{j}_1}$ and $\Omega = \Omega \setminus \tilde{j}_1$.

For $i = 2 : A_n$ **do**

Use (64a) to compute

$$\tilde{j}_i = \arg \max_{j_i \in \Omega} D_n^{KL}([\mathbf{J}_n^T, \mathbf{e}_{j_i}^T]^T).$$

Set $\mathbf{J}_n = [\mathbf{J}_n^T, \mathbf{e}_{\tilde{j}_i}^T]^T$ and $\Omega = \Omega \setminus \tilde{j}_i$.

End for

Set $\tilde{\mathbf{a}}_{mn} = [\mathbf{J}_n^T \mathbf{J}_n]_{m,m}$ and $\tilde{\mathbf{a}}_n = [\tilde{a}_{1n}, \dots, \tilde{a}_{Mn}]^T$.

End for

Output: $\tilde{\mathbf{a}} = [\tilde{\mathbf{a}}_1^T, \dots, \tilde{\mathbf{a}}_N^T]^T$, which is the resulting selection vector.

Complexity: Assume A_n ($n = 1, \dots, N$) are fixed constants. For Algorithm 1, at the i -th iteration, we need to compute D_n^{KL} as per (64a) $(M - i)$ times, where the complexity for computing (64a) is on the order of $O(i^3)$, so the complexity at the i -th step is on the order of $O(i^3(M - i)) = O(M)$. Thus, the total complexity of using Algorithm 1 is on the order of $O(MN)$. For comparison, suppose we let $\|\mathbf{a}_n\| = A_n$ and use exhaustive search, then we need to compute (64a) $\binom{M}{A_n} = O(M^{A_n})$ times and the complexity for computing (64a) is on the order of $O(A_n^3)$. So the total complexity of using the brute force method is $O(\sum_{i=1}^N M^{A_n} A_n^3) = O(\sum_{i=1}^N M^{A_n})$.

V. NUMERICAL EXPERIMENTS

In this section, numerical examples are presented to illustrate our findings. We set the transmitted energy $E_m = 10^6$ for all m , $m = 1, \dots, M$. The transmitted waveforms are $s_m(t) = \frac{1}{\sqrt{T}} \exp(j2\pi f_m t)$, $0 < t < T$, where $T = 1\text{ms}$ and f_m is the frequency of the m -th transmitter. Define $\mathbf{f} = [f_1, f_2, \dots, f_M]$ as the frequency vector. The target is located at $(x, y) = (0, 0)$ km. The variance of the reflection coefficient associated with the (m, n) -th path is set as $\sigma_{mn}^2 = 10$ for all m and n , and the PSD of the clutter-plus-noise is set as $N_{ii} = 10^{-6}$ for all i , $i = 1, \dots, N$. All curves are obtained using 10^4 Monte Carlo simulations.

A. Optimal Selection for Orthogonal Waveforms

Assume $M = 3$ transmitters located at $(x_{t,1}, y_{t,1}) = (0, 1)$ km, $(x_{t,2}, y_{t,2}) = (0, 2)$ km, and $(x_{t,3}, y_{t,3}) = (0, 3)$ km,

TABLE I: Detection probability for different selections

Selection Combination	SCNR (dB)	P_D
$\{< 1, 1 >, < 1, 2 >\}$	$\{10, 10\}$	0.8798
$\{< 1, 1 >, < 2, 2 >\}$	$\{10, 3.98\}$	0.7286
$\{< 1, 1 >, < 3, 2 >\}$	$\{10, 0.46\}$	0.6868
$\{< 2, 1 >, < 1, 2 >\}$	$\{3.98, 10\}$	0.7422
$\{< 2, 1 >, < 2, 2 >\}$	$\{3.98, 3.98\}$	0.4354
$\{< 2, 1 >, < 3, 2 >\}$	$\{3.98, 0.46\}$	0.3196
$\{< 3, 1 >, < 1, 2 >\}$	$\{0.46, 10\}$	0.6785
$\{< 3, 1 >, < 2, 2 >\}$	$\{0.46, 3.98\}$	0.3117
$\{< 3, 1 >, < 3, 2 >\}$	$\{0.46, 0.46\}$	0.1880

respectively. The $N = 2$ receivers are located at $(x_{r,1}, y_{r,1}) = (-1, 0)$ km and $(x_{r,1}, y_{r,2}) = (1, 0)$ km. Suppose the number of transmitters that can be selected at the receivers are $A_1 = A_2 = 1$. The target reflection coefficients and the clutter-plus-noise are spatially white. Let the frequency vector be $\mathbf{f} = [\frac{10}{T}, \frac{20}{T}, \frac{30}{T}]$, which ensures that the waveforms are approximately orthogonal. The false alarm probability is 10^{-2} . Table I shows the detection performance of all selection schemes. Denote by $< m, n >$ the m -th transmitter being selected at the n -th receiver. We can see that higher detection probability can be achieved when the subset of the selected transmitters have larger SCNRs. For example, the corresponding SCNRs of the selection $\{< 2, 1 >, < 1, 2 >\}$ are $\{3.98, 10\}$ dB, which is larger than the the corresponding SCNRs of the selection $\{< 3, 1 >, < 3, 2 >\}$, which are $\{0.46, 0.46\}$ dB. The resulting detection probabilities are 0.7422 and 0.1880 respectively. Clearly, the former selection with higher SCNRs has larger detection probability. We can see that optimal selection is $\{< 1, 1 >, < 1, 2 >\}$, and the corresponding SCNRs are the largest, which agrees with Theorem 1.

B. Optimal Selection for Non-orthogonal Waveforms

The previous results assumed orthogonal signals. Next, we remove this assumption to investigate the general case. Assume there are $M = 4$ transmitters located at $(x_{t,1}, y_{t,1}) = (0, 1.5)$ km, $(x_{t,2}, y_{t,2}) = (0, -1.5)$ km, $(x_{t,3}, y_{t,3}) = (0, 3)$ km and $(x_{t,4}, y_{t,4}) = (0, -3)$ km, respectively. The $N = 2$ receivers located at $(x_{r,1}, y_{r,1}) = (-1, 0)$ km and $(x_{r,1}, y_{r,2}) = (1, 0)$ km. We consider three scenarios. The frequency vector $\mathbf{f}$ for each scenario is $[\frac{10}{T}, \frac{10}{T}, \frac{30}{T}, \frac{30}{T}]$ (Sc. 1), $[\frac{10}{T}, \frac{10}{T}, \frac{10}{T}, \frac{10}{T}]$ (Sc. 2) and $[\frac{10}{T}, \frac{20}{T}, \frac{30}{T}, \frac{40}{T}]$ (Sc. 3). In Fig. 1, the selection combinations are $\{< 1, 1 >, < 1, 2 >\}$ (Sel. 1), $\{< 1, 1 >, < 2, 1 >, < 1, 2 >, < 2, 2 >\}$ (Sel. 2), $\{< 1, 1 >, < 3, 1 >, < 1, 2 >, < 3, 2 >\}$ (Sel. 3) and all transmitters being selected (Sel. 4). The other parameters are the same as those in the previous example. In scenario 1, we can see that the corresponding SCNRs of selection 2, $\{6.48, 6.48, 6.48, 6.48\}$ dB are larger than the corresponding SCNRs of selection 3, $\{6.48, 0.46, 6.48, 0.46\}$ dB, but the later selection has better detection performance, so the selection which has the best performance in terms of the ROCs may not correspond to choosing the largest

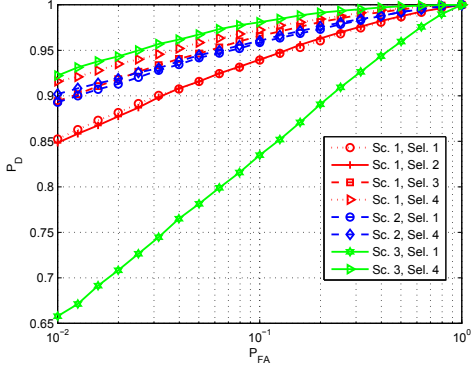


Fig. 1: ROC curves under different scenarios and selections.

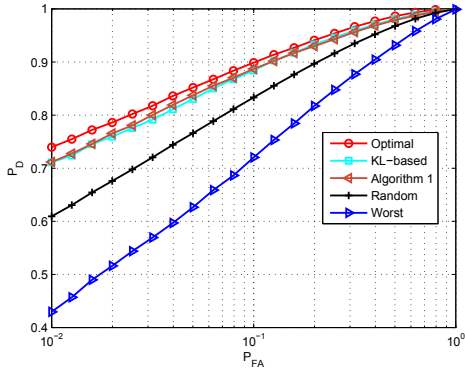


Fig. 2: The ROC curves (averaged over a hundred different realizations of transmitter replacement) of the optimal selection, KL-based selection, Algorithm 1-based selection, random selection and worst selection.

SCNRs when the transmitted waveforms are not orthogonal. In scenario 2, where all the transmitted waveforms are the same, we can see the detection performance of selection 1 is almost the same with selection 4 (which selects all the transmitters), which indicates that selecting only some of the transmitters as opposed to all of them may lose very little performance. In scenario 3, the waveforms are approximately orthogonal. We can see that detection performance in scenario 3 is best among the three scenarios if we select all transmitters while the detection performance of the optimal selection in scenario 3 performs poorly if the number of transmitter that can be selected is $A_1 = A_2 = 1$.

C. KL-Based Selection for Non-orthogonal Waveforms

Now we consider the case where $M = 8$ transmitters are randomly and uniformly located in a ring with inner radius 2 km and outer radius 4 km. There are two receivers located at $(x_{r,1}, y_{r,1}) = (-1, 0)$ km and $(x_{r,2}, y_{r,2}) = (1, 0)$ km. The frequency vector is $\mathbf{f} = [\frac{1}{2T}, \frac{2}{2T}, \dots, \frac{4}{T}]$. The number of transmitters that can be selected at each receiver is $A_1 = A_2 = 2$. The other parameters are the same as those in Fig. 1. In Fig. 2, the ROC curve averaged

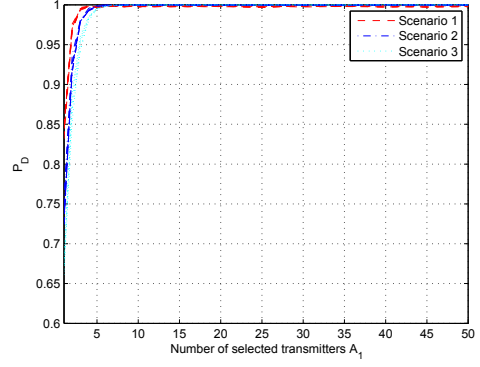


Fig. 3: Detection probabilities of the Algorithm 1-based selection versus the number of selected transmitters.

over a hundred different realizations of transmitter replacement are plotted for the optimal selection (obtained via exhaustive search), KL-based selection, Algorithm 1-based selection, random selection and the worst selection (obtained via exhaustive search). We can see that, for any fixed P_{FA} , the optimal and worst selection lead to the largest and smallest P_D , as expected. The P_D obtained using random selection is larger than the worst case. The P_D for the KL-based selection is better than the random selection. The P_D for the Algorithm 1-based selection is almost on top of that for the KL-based selection. We find that the ROC curves obtained using both the KL-based and Algorithm 1-based methods are close enough to the optimal selection, which implies the effectiveness of these proposed methods.

D. Large Saving of Complexity

Consider the case where the reflection coefficients are uncorrelated and the transmitted waveforms are non-orthogonal. Assume there are $M = 50$ transmitters randomly located in a ring with inner radius 4 km and outer radius 6 km. A single receiver is located at $(x_{r,1}, y_{r,1}) = (0.5, 0)$ km. We consider three scenarios, the frequency vector $\mathbf{f}$ for which are $[\frac{3}{50T}, \frac{6}{50T}, \dots, \frac{3}{T}]$ (Scenario 1), $[\frac{6}{50T}, \frac{12}{50T}, \dots, \frac{6}{T}]$ (Scenario 2) and $[\frac{9}{50T}, \frac{18}{50T}, \dots, \frac{9}{T}]$ (Scenario 3), respectively. Due to large number of transmitters, the optimal selection based on exhaustive search becomes infeasible. Thus, we employ Algorithm 1 in this example to select transmitters. Letting $P_{FA} = 10^{-2}$, for each scenario, the resulting P_D for five different (random) transmitter placements are plotted versus the allowed number A_1 of selected transmitters, as shown in Fig. 3. It is seen that for all the tested scenarios and transmitter placements, selecting $A_1 = 6$ transmitters leads to almost the same P_D as selecting all transmitters. Clearly, using the proposed method, a large saving of complexity can be gained accordingly.

E. Extension

1) *Correlated Reflections and Clutter-Plus-Noise*: In Fig. 4, the parameters are the same with the case in Fig.

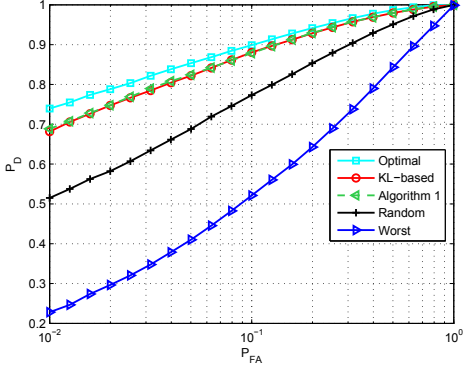


Fig. 4: The ROC curves (averaged over a hundred different realizations of transmitter replacement) of the optimal selection, KL-based selection, Algorithm 1-based selection, random selection and worst selection under correlated reflections and clutter-plus-noise.

2 except for the clutter-plus-noise and the reflection coefficients are all not spatially white. The correlation of clutter-plus-noise are set as $N_{ij}=10^{-6}$ for $i = j$ and 10^{-7} for $i \neq j$, $i, j = 1, 2$, and the correlation of reflection coefficients are set as $\mathbb{E}\{\beta_{mn}\beta_{m'n'}^*\} = 10$ for $m = m'$, $n = n'$ and 1 otherwise. In Fig. 4, the ROC curve averaged over a hundred different realizations of transmitter replacement are plotted for the optimal selection, KL-based selection, Algorithm 1-based selection, random selection and the worst selection. We can see that, for any fixed P_{FA} , the optimal and worst selection lead to the largest and smallest P_D , as expected. The P_D obtained using random selection is larger than the worst case. The P_D for the KL-based selection is better than the random selection. The P_D for the Algorithm 1-based selection is almost on top of that for the KL-based selection. We find that the ROC curves obtained using both the KL-based and Algorithm 1-based methods are close enough to the optimal selection, which implies these proposed methods are again effective in finding good approaches without exhaustive search.

2) *Imperfectly Estimated Waveforms*: For passive radar, the waveforms of the transmitted signals must be estimated. Now consider the case where the transmitted signals are not estimated perfectly. Denote by $\hat{s}_m(t)$ the estimated waveform for the m -th ($m = 1, \dots, M$) transmitted signal. Assume

$$\hat{s}_m(t) = s_m(t) + e_m(t), 0 < t < T_m \quad (67)$$

where $e_m(t)$ is the measurement error. Assume $e_m(t)$ is a zero-mean complex Gaussian random process with autocorrelation function

$$\mathbb{E}\{e_m(t_1)e_m^*(t_2)\} = \begin{cases} N_{e_m} & t_1 = t_2 \\ 0 & t_1 \neq t_2 \end{cases}. \quad (68)$$

Define the signal-to-error ratio (SER) for the m -th

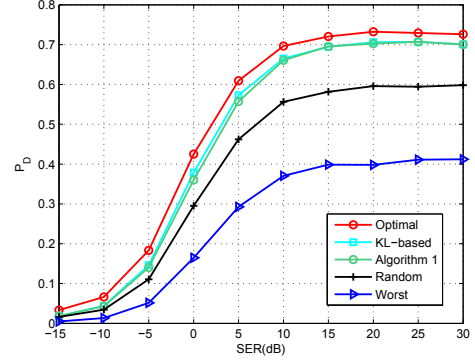


Fig. 5: Average detection probability for the optimal selection, KL-based selection, Algorithm 1-based selection, random selection and worst selection versus SER.

waveform as

$$\text{SER}_m = \frac{\int_0^T |s_m(t)|^2 dt}{N_{e_m} T}. \quad (69)$$

We assume $\text{SER}_1 = \text{SER}_2 = \dots = \text{SER}_M$. In Fig. 5, the P_{FA} is fixed to 10^{-2} , and the other parameters (including the SCNR) are the same as those in Fig. 2, except for the inaccurate waveform estimation. Thus, if the waveforms were perfectly estimated, the P_D should take values on the left hand side of Fig. 2. We plot the average detection probability for the optimal selection, KL-based selection, Algorithm 1-based selection, random selection and worst selection versus SERs. From this figure, we see that the detection probability obtained by both the KL-based and Algorithm 1-based methods are close enough to the optimal selection even in very low SER, which implies these proposed methods are also effectiveness even if the transmitted waveforms are estimated. For large enough SER, it is seen that the P_D approaches the corresponding values in Fig. 2 for the perfect estimated waveforms case, as expected.

VI. CONCLUSION

We studied the limited-complexity receiver design for passive/active MIMO radar, considering that usually only a limited number of MFs can be implemented at each receiver due to cost considerations. For a general case with possibly nonorthogonal signals and possibly correlated clutter-plus-noise and reflection coefficients, we investigated the target detection performance under the Neyman-Pearson criterion and formulated an optimization problem to maximize the detection performance for selecting a set of matched filters. For the case of spatially white clutter-plus-noise and uncorrelated reflection coefficients, we showed that the problem is equivalent to selecting a subset of transmitters. When the signals transmitted by different transmitters are mutually orthogonal, we proved that the selection leading to larger SCNRs can achieve larger detection probability. Thus, at each receiver maximum detection probability

can be achieved if we select the transmitters corresponding to the largest SCNRs. When the transmitted waveforms are nonorthogonal, we replaced the original objective function with the KL distance and showed that the KL-based transmitter selection method can lead to a detection performance that is close to the optimal selection. Further, optimally selecting a few transmitters can lead to almost the same performance as selecting all transmitters. Simulations show that this method can also be used in the more general case of correlated clutter-plus-noise and reflection coefficients and in the case where the waveform is estimated.

APPENDIX A

DERIVATION OF LOG-LIKELIHOOD RATIO IN (5)

Define $\mathcal{S} \triangleq \{s_m(t - \tau_{mn}), t \in \mathcal{T} | m = 1, \dots, M, n = 1, \dots, N\}$. Since the elements in $\mathcal{S}$ are linear independent, the dimension of $\text{span}(\mathcal{S})$ is $K = MN$, where $\text{span}(\mathcal{S})$ is the set of all linear combinations of the elements of $\mathcal{S}$. Let $\{\phi_1(t), \dots, \phi_K(t)\}$ be an orthonormal basis of $\text{span}(\mathcal{S})$. We extend the set $\{\phi_1(t), \dots, \phi_K(t)\}$ to a complete orthonormal basis, which is $\{\phi_1(t), \dots, \phi_K(t), \phi_{K+1}(t), \dots\}$. Thus,

$$r_n(t) = \sum_{k=1}^{\infty} \langle r_n(t), \phi_k(t) \rangle \phi_k(t) = \sum_{k=1}^{\infty} r_{kn} \phi_k(t) \quad (70)$$

where

$$r_{kn} = \langle r_n(t), \phi_k(t) \rangle \triangleq \int_{\mathcal{T}} r_n(t) \phi_k^*(t) dt \quad (71)$$

is the inner product between $r_n(t)$ and $\phi_k(t)$.

Under $\mathcal{H}_0$, r_{kn} can be written as $r_{kn}|\mathcal{H}_0 = \langle w_n(t), \phi_k(t) \rangle \triangleq w_{kn}$, where w_{kn} is complex Gaussian distributed [41] with zero-mean and variance 1, and

$$\begin{aligned} \mathbb{E}\{r_{kn} r_{k'n'}^* | \mathcal{H}_0\} &= \mathbb{E}\{w_{kn} w_{k'n'}^* | \mathcal{H}_0\} \\ &= \mathbb{E}\left\{ \int_{\mathcal{T}} w_n(t) \phi_k^*(t) dt \int_{\mathcal{T}} w_{n'}^*(t') \phi_{k'}(t') dt' \right\} \\ &= \int_{\mathcal{T}} \int_{\mathcal{T}} \mathbb{E}\{w_n(t) w_{n'}^*(t')\} \phi_k^*(t) \phi_{k'}(t') dt dt' \\ &= \int_{\mathcal{T}} \int_{\mathcal{T}} N_{nn'} \delta(t - t') \phi_k^*(t) \phi_{k'}(t') dt dt' \\ &= N_{nn'} \delta_{kk'}, \end{aligned} \quad (72)$$

where k and k' are positive integers, $n, n' = 1, \dots, N$ and

$$\delta_{kk'} = \begin{cases} 1 & k = k' \\ 0 & k \neq k'. \end{cases} \quad (73)$$

Under $\mathcal{H}_1$, r_{kn} can be written as

$$\begin{aligned} r_{kn}|\mathcal{H}_1 &= \langle \sum_{m=1}^M \xi_{mn} s_m(t - \tau_{mn}) + w_n(t), \phi_k(t) \rangle \\ &= \langle \sum_{m=1}^M \xi_{mn} s_m(t - \tau_{mn}), \phi_k(t) \rangle + w_{kn}. \end{aligned} \quad (74)$$

Since $\{\phi_1(t), \dots, \phi_K(t)\}$ is a orthonormal basis of $\text{span}(\mathcal{S})$, in (74), $s_m(t - \tau_{mn})$ can be written as

$$\begin{aligned} s_m(t - \tau_{mn}) &= \sum_{k=1}^K \langle s_m(t - \tau_{mn}), \phi_k(t) \rangle \phi_k(t) \\ &= \sum_{k=1}^K s_{kmn} \phi_k(t) \end{aligned} \quad (75)$$

where $s_{kmn} = \langle s_m(t - \tau_{mn}), \phi_k(t) \rangle$. Hence, the term $\langle \sum_{m=1}^M \xi_{mn} s_m(t - \tau_{mn}), \phi_k(t) \rangle$ in (74) can be written as

$$\begin{aligned} &\langle \sum_{m=1}^M \xi_{mn} s_m(t - \tau_{mn}), \phi_k(t) \rangle \\ &= \langle \sum_{m=1}^M \xi_{mn} \left(\sum_{k'=1}^K s_{k'mn} \phi_{k'}(t) \right), \phi_k(t) \rangle \end{aligned} \quad (76)$$

$$= \begin{cases} \sum_{m=1}^M \xi_{mn} s_{kmn} & k \leq K \\ 0 & k > K. \end{cases} \quad (77)$$

From (77), the equation (74) can be written as

$$r_{kn}|\mathcal{H}_1 = \begin{cases} \mathbf{u}_{kn}^T \boldsymbol{\xi}_n + w_{kn} & k \leq K \\ w_{kn} & k > K \end{cases} \quad (78)$$

where $\mathbf{u}_{kn} = [s_{k1n}, \dots, s_{kMn}]^T$. It is clear $\mathbf{u}_{kn}^T \boldsymbol{\xi}_n \sim \mathcal{CN}(0, \mathbf{u}_{kn}^T \boldsymbol{\Lambda}_{nn} \mathbf{u}_{kn})$, where $\boldsymbol{\Lambda}_{nn} = \mathbb{E}\{\boldsymbol{\xi}_n \boldsymbol{\xi}_n^H\}$. From (72) and (78), we see that

$$\begin{aligned} &\mathbb{E}\{r_{kn} r_{k'n'}^* | \mathcal{H}_1\} \\ &= \begin{cases} N_{nn'} \delta_{kk'} & k > K \text{ or } k' > K. \\ \mathbf{u}_{kn}^T \boldsymbol{\Lambda}_{nn'} \mathbf{u}_{k'n'}^* + N_{nn'} \delta_{kk'} & k, k' \leq K \end{cases} \end{aligned} \quad (79)$$

where $\boldsymbol{\Lambda}_{nn'} = \mathbb{E}\{\boldsymbol{\xi}_n \boldsymbol{\xi}_{n'}^H\}$. Define

$$\mathbf{r}_k \triangleq [r_{k1}, r_{k2}, \dots, r_{kN}]^T \quad (80)$$

and

$$\mathbf{R}_k \triangleq [\mathbf{r}_1^T, \dots, \mathbf{r}_K^T]. \quad (81)$$

From (72) and (79), we see that r_{kn} and $r_{k'n'}$ ($k \neq k'$) are mutually independent for $k > K$ or $k' > K$. Thus, $\mathbf{r}_k$ and $\mathbf{r}_{k'}$ ($k \neq k'$) are mutually independent for $k > K$ or $k' > K$. Thus, the likelihood ratio is [41]

$$\begin{aligned} \mathcal{L} &= \lim_{i \rightarrow \infty} \frac{p(\mathbf{R}_i | \mathcal{H}_1)}{p(\mathbf{R}_i | \mathcal{H}_0)} \\ &= \frac{p(\mathbf{R}_K | \mathcal{H}_1)}{p(\mathbf{R}_K | \mathcal{H}_0)} \lim_{i \rightarrow \infty} \prod_{i=K+1}^{\infty} \frac{p(\mathbf{r}_i | \mathcal{H}_1)}{p(\mathbf{r}_i | \mathcal{H}_0)} = \frac{p(\mathbf{R}_K | \mathcal{H}_1)}{p(\mathbf{R}_K | \mathcal{H}_0)}. \end{aligned} \quad (82)$$

Let

$$\mathbf{v}_n = [r_{n1}, \dots, r_{nN}]^T. \quad (83)$$

It is clear $\mathbf{v}_n^T$ is the n -th row of $\mathbf{R}_K$. Define

$$\boldsymbol{\Upsilon}_{n_1, n_2, i} \triangleq \mathbb{E}\{\mathbf{v}_{n_1} \mathbf{v}_{n_2}^H | \mathcal{H}_i\}, \quad (84)$$

where $i = 0, 1$. From (72), the (k_1, k_2) -th ($k_1, k_2 = 1, \dots, K$) element of $\boldsymbol{\Upsilon}_{n_1, n_2, 0}$ is

$$\boldsymbol{\Upsilon}_{n_1, n_2, 0, (k_1, k_2)} = N_{n_1 n_2} \delta_{k_1 k_2} \quad (85)$$

Thus,

$$\boldsymbol{\Upsilon}_{n_1, n_2, 0} = N_{n_1 n_2} \mathbf{I}_K. \quad (86)$$

From (79), the (k_1, k_2) -th element of $\boldsymbol{\Upsilon}_{n_1, n_2, 1}$ is

$$\boldsymbol{\Upsilon}_{n_1, n_2, 1, (k_1, k_2)} = \mathbf{u}_{k_1}^T \boldsymbol{\Lambda}_{n_1 n_2} \mathbf{u}_{k_2}^* + N_{n_1 n_2} \delta_{k_1 k_2}. \quad (87)$$

Thus,

$$\Upsilon_{n_1, n_2, 1} = N_{n_1 n_2} \mathbf{I}_K + \mathbf{U}_{n_1} \mathbf{\Lambda}_{n_1 n_2} \mathbf{U}_{n_2}^H, \quad (88)$$

where $\mathbf{U}_n = [\mathbf{u}_{1n}, \dots, \mathbf{u}_{Kn}]^T$. Let $\mathbf{v} = [\mathbf{v}_1^T, \dots, \mathbf{v}_N^T]^T$. From (86) and (88), we obtain

$$\mathbb{E}\{\mathbf{v}\mathbf{v}^H | \mathcal{H}_0\} = \mathbf{N} \otimes \mathbf{I}_K \quad (89)$$

and

$$\mathbb{E}\{\mathbf{v}\mathbf{v}^H | \mathcal{H}_1\} = \mathbf{N} \otimes \mathbf{I}_K + \mathbf{\Phi} \mathbf{\Lambda} \mathbf{\Phi}^H, \quad (90)$$

where $\mathbf{\Phi} = \text{Diag}(\mathbf{U}_1, \dots, \mathbf{U}_N)$ and $\mathbf{\Lambda} = \mathbb{E}\{\xi\xi^H\}$. So the log-likelihood ratio is

$$\begin{aligned} \mathcal{L} &= \ln \frac{p(\mathbf{R}_K | \mathcal{H}_1)}{p(\mathbf{R}_K | \mathcal{H}_0)} = \ln \frac{p(\mathbf{v} | \mathcal{H}_1)}{p(\mathbf{v} | \mathcal{H}_0)} \\ &= C + \mathbf{v}^H \left((\mathbf{N} \otimes \mathbf{I}_K)^{-1} - (\mathbf{N} \otimes \mathbf{I}_K + \mathbf{\Phi} \mathbf{\Lambda} \mathbf{\Phi}^H)^{-1} \right) \mathbf{v}, \end{aligned} \quad (91)$$

where

$$C = \ln \text{Det}(\mathbf{N} \otimes \mathbf{I}_K) - \ln \text{Det}(\mathbf{N} \otimes \mathbf{I}_K + \mathbf{\Phi} \mathbf{\Lambda} \mathbf{\Phi}^H) \quad (92)$$

is a constant. Define

$$\mathbf{s} \triangleq [s_1(t - \tau_{11}), \dots, s_M(t - \tau_{MN})]^T \quad (93)$$

and

$$\phi \triangleq [\phi_1(t), \dots, \phi_K(t)]^T. \quad (94)$$

From (75), we obtain

$$\mathbf{s} = \mathbf{S} \phi. \quad (95)$$

where the $\mathbf{S}$ a $K \times K$ matrix whose k -row is

$$\mathbf{S}_k = [s_{k11}, \dots, s_{kMN}]. \quad (96)$$

Let $\mathbf{x}_n$ be a $K \times 1$ vector with the $(n'(M-1) + m)$ -th ($m = 1, \dots, M$, $n' = 1, \dots, N$) element

$$x_{mn'n} = \langle r_n(t), s_m(t - \tau_{mn'}) \rangle. \quad (97)$$

Define

$$\begin{aligned} &\langle r_n(t), \mathbf{s} \rangle \\ &\triangleq [\langle r_n(t), s_1(t - \tau_{11}) \rangle, \dots, \langle r_n(t), s_M(t - \tau_{MN}) \rangle]^T. \end{aligned} \quad (98)$$

From (83), (95), (97) and (98), we obtain

$$\mathbf{x}_n = \langle r_n(t), \mathbf{s} \rangle = \mathbf{S}^* \langle r_n(t), \phi \rangle = \mathbf{S}^* \mathbf{v}_n. \quad (99)$$

Since the elements of $\mathcal{S}$ are linear independent, they can also be a basis of $\mathcal{S}$. Thus, the matrix $\mathbf{S}^*$ is invertible such that

$$\mathbf{v}_n = (\mathbf{S}^*)^{-1} \mathbf{x}_n. \quad (100)$$

Let $\mathbf{x} = [\mathbf{x}_1^T, \dots, \mathbf{x}_N^T]^T$. From (100), we obtain

$$\mathbf{v} = (\mathbf{I}_N \otimes (\mathbf{S}^*)^{-1}) \mathbf{x} \quad (101)$$

Plugging (101) into (91), we obtain

$$\begin{aligned} T &= C + ((\mathbf{I}_N \otimes (\mathbf{S}^*)^{-1}) \mathbf{x})^H \left((\mathbf{N} \otimes \mathbf{I}_K)^{-1} \right. \\ &\quad \left. - (\mathbf{N} \otimes \mathbf{I}_K + \mathbf{\Phi} \mathbf{\Lambda} \mathbf{\Phi}^H)^{-1} \right) (\mathbf{I}_N \otimes (\mathbf{S}^*)^{-1}) \mathbf{x} \\ &= C + \mathbf{x}^H (\mathbf{I}_N \otimes \mathbf{S}^T)^{-1} \left((\mathbf{N} \otimes \mathbf{I}_K)^{-1} \right. \\ &\quad \left. - (\mathbf{N} \otimes \mathbf{I}_K + \mathbf{\Phi} \mathbf{\Lambda} \mathbf{\Phi}^H)^{-1} \right) (\mathbf{I}_N \otimes \mathbf{S}^*)^{-1} \mathbf{x} \\ &= C + \mathbf{x}^H \left((\mathbf{N} \otimes \mathbf{S}^* \mathbf{S}^T)^{-1} \right. \\ &\quad \left. - (\mathbf{N} \otimes \mathbf{S}^* \mathbf{S}^T + (\mathbf{I}_N \otimes \mathbf{S}^*) \mathbf{\Phi} \mathbf{\Lambda} \mathbf{\Phi}^H (\mathbf{I}_N \otimes \mathbf{S}^T))^{-1} \right) \mathbf{x} \\ &= C + \mathbf{x}^H \left((\mathbf{N} \otimes \mathbf{\Xi})^{-1} - (\mathbf{N} \otimes \mathbf{\Xi} + \mathbf{\Psi} \mathbf{\Lambda} \mathbf{\Psi}^H)^{-1} \right) \mathbf{x}, \end{aligned} \quad (102)$$

where

$$\mathbf{\Xi} = \mathbf{S}^* \mathbf{S}^T \quad (103)$$

and

$$\begin{aligned} \mathbf{\Psi} &= (\mathbf{I}_N \otimes \mathbf{S}^*) \mathbf{\Phi} = \text{Diag}(\mathbf{S}^*, \dots, \mathbf{S}^*) \text{Diag}(\mathbf{U}_1, \dots, \mathbf{U}_N) \\ &= \text{Diag}(\mathbf{\Psi}_1, \dots, \mathbf{\Psi}_N), \end{aligned} \quad (104)$$

in which

$$\mathbf{\Psi}_n = \mathbf{S}^* \mathbf{U}_n \quad (105)$$

is an $MN \times M$ matrix. The $((n_1 - 1)M + m_1, (n_2 - 1)M + m_2)$ -th element ($n_1, n_2 = 1, \dots, N$ and $m_1, m_2 = 1, \dots, M$) of $\mathbf{\Xi}$ is

$$\begin{aligned} &\Xi_{(n_1-1)M+m_1, (n_2-1)M+m_2} \\ &= \mathbf{S}_{((n_1-1)M+m_1)}^* \mathbf{S}_{((n_2-1)M+m_2)}^T \\ &= \sum_{k=1}^K s_{km_1 n_1}^* s_{km_2 n_2} \\ &= \int_T s_{m_2}(t - \tau_{m_2 n_2}) s_{m_1}^*(t - \tau_{m_1 n_1}) dt. \end{aligned} \quad (106)$$

The $((n_1 - 1)M + m_1, m)$ -th elements of $\mathbf{\Psi}_n$ is

$$\begin{aligned} \Psi_{n, ((n_1-1)M+m_1, m)} &= \sum_{k=1}^K s_{km_1 n_1}^* s_{kmn} \\ &= \int_T s_m(t - \tau_{mn}) s_{m_1}^*(t - \tau_{m_1 n_1}) dt. \end{aligned} \quad (107)$$

$$(108)$$

From (103) and (104), the constant C in (92) can also be written as

$$\begin{aligned} C &= \left(\ln \text{Det}(\mathbf{N} \otimes \mathbf{I}_K) + \ln \text{Det}(\mathbf{I}_N \otimes \mathbf{\Xi}) \right) \\ &\quad - \ln \text{Det}(\mathbf{I}_N \otimes \mathbf{\Xi}) - \ln \text{Det}(\mathbf{N} \otimes \mathbf{I}_K + \mathbf{\Phi} \mathbf{\Lambda} \mathbf{\Phi}^H) \\ &= \ln \text{Det}(\mathbf{N} \otimes \mathbf{\Xi}) - \ln \text{Det}(\mathbf{N} \otimes \mathbf{\Xi} + \mathbf{\Psi} \mathbf{\Lambda} \mathbf{\Psi}^H). \end{aligned} \quad (109)$$

APPENDIX B

MONOTONICITY OF P_d WITH RESPECT TO ρ_{mn} FOR LEMMA 1

Define $\Omega_\rho = \{\rho_{11}, \rho_{21}, \dots, \rho_{MN}\}$. We consider two cases. When the elements in $\Omega_\rho \setminus \rho_{mn}$ are all zero, we will prove P_D is strictly monotone increasing on $(0, +\infty)$ with respect to ρ_{mn} . When the elements in $\Omega_\rho \setminus \rho_{mn}$ are not all zero, we will prove P_D is strictly monotone increasing on $[0, +\infty)$ with respect to ρ_{mn} .

In the first case, the elements in $\Omega_\rho \setminus \rho_{mn}$ are all zero. From (34), we obtain $T_s = \zeta_{mn}$. Since $\zeta_{mn}|\mathcal{H}_0 \sim \frac{\rho_{mn}}{2(\rho_{mn}+1)}\chi_2^2$ and $\zeta_{mn}|\mathcal{H}_1 \sim \frac{\rho_{mn}}{2}\chi_2^2$, we obtain

$$P_{FA} = \exp\left(-\frac{1+\rho_{mn}}{\rho_{mn}}\gamma\right) \quad (110)$$

$$P_D = \exp\left(-\frac{1}{\rho_{mn}}\gamma\right). \quad (111)$$

From (110), we obtain

$$\gamma = -\frac{\rho_{mn}}{1+\rho_{mn}} \ln P_{FA}. \quad (112)$$

Inserting (112) into (111), we obtain

$$P_D = \exp\left(\frac{\ln P_{FA}}{1+\rho_{mn}}\right) \quad (113)$$

Note that $0 < P_{FA} < 1$, $\ln P_{FA} < 0$, so P_D is strictly monotone increasing on $(0, +\infty)$ with respect to ρ_{mn} .

In the second case, the elements in $\Omega_\rho \setminus \rho_{mn}$ are not all zero, we denote the indices of these nonzero elements by $1_+, \dots, P_+$, where P the number of nonzero elements in $\Omega_\rho \setminus \rho_{mn}$. Since selecting one more transmitter would not lead to worse detection performance,

$$P_D(\dots, \rho_{mn}, \dots) \geq P_D(\dots, \rho_{mn} = 0, \dots), \forall \rho_{mn} > 0. \quad (114)$$

Thus, if

$$\frac{\partial P_D}{\partial \rho_{mn}} > 0, \forall \rho_{mn} > 0, \quad (115)$$

then P_D is strictly monotone increasing on $[0, +\infty)$ with respect to ρ_{mn} .

Proof of (115): Define

$$P_{FA} = 1 - F_{T_s|\mathcal{H}_0}(\gamma, \rho) \triangleq \varphi_0(\gamma, \rho), \quad (116)$$

and

$$P_D = 1 - F_{T_s|\mathcal{H}_1}(\gamma, \rho) \triangleq \varphi_1(\gamma, \rho). \quad (117)$$

(116) determines a explicit function $\gamma = \gamma(P_{FA}, \rho)$. Since P_{FA} and all the elements in $\Omega \setminus \rho_{mn}$ are fixed, γ is only determined by ρ_{mn} . For a simpler notation, we use $\gamma(\rho_{mn})$ instead of $\gamma(P_{FA}, \rho)$. Similarly, we use $\varphi_0(\gamma, \rho_{mn})$, $\varphi_1(\gamma, \rho_{mn})$ and $P_D(\rho_{mn})$ instead of $\varphi_0(\gamma, \rho)$, $\varphi_1(\gamma, \rho)$ and $P_D(\rho)$, respectively.

According to the theory of calculus,

$$\frac{\partial P_D}{\partial \rho_{mn}} = \frac{\partial \varphi_1}{\partial \gamma} \frac{\partial \gamma}{\partial \rho_{mn}} + \frac{\partial \varphi_1}{\partial \rho_{mn}} \quad (118)$$

$$\frac{\partial \gamma}{\partial \rho_{mn}} = -\frac{\frac{\partial \varphi_0}{\partial \rho_{mn}}}{\frac{\partial \varphi_0}{\partial \gamma}}. \quad (119)$$

From (35), (37), (116) and (117), we can get

$$\varphi_i(\gamma, \rho_{mn}) = \int_{\gamma}^{\infty} f_i(z) * f_{\zeta_{mn}|\mathcal{H}_i}(z, \rho_{mn}) dz, \quad (120)$$

where $i = 0, 1$,

$$f_i(z) = f_{\zeta_{1+}|\mathcal{H}_i}(z, \rho_{1+}) * \dots * f_{\zeta_{P+}|\mathcal{H}_i}(z, \rho_{P+}). \quad (121)$$

In addition, $f_{\zeta_{mn}|\mathcal{H}_i}(z, \rho_{mn})$ and $f_{\zeta_{p+}|\mathcal{H}_i}(z, \rho_{p+})$, $p =$

$1, \dots, P$, are the probability density functions (pdfs)⁶ of ζ_{mn} and ζ_{p+} under $\mathcal{H}_i$ ($i = 0, 1$), respectively. Since $\zeta_{mn}|\mathcal{H}_0 \sim \frac{\rho_{mn}}{2(\rho_{mn}+1)}\chi_2^2$ and $\zeta_{mn}|\mathcal{H}_1 \sim \frac{\rho_{mn}}{2}\chi_2^2$,

$$\begin{aligned} f_{\zeta_{mn}|\mathcal{H}_0}(z, \rho_{mn}) &= \frac{\rho_{mn}+1}{\rho_{mn}} \exp\left(-\frac{\rho_{mn}+1}{\rho_{mn}}z\right)u(z) \\ f_{\zeta_{mn}|\mathcal{H}_1}(z, \rho_{mn}) &= \frac{1}{\rho_{mn}} \exp\left(-\frac{1}{\rho_{mn}}z\right)u(z). \end{aligned} \quad (122)$$

where $u(z)$ is a step function. Similarly,

$$\begin{aligned} f_{\zeta_{p+}|\mathcal{H}_0}(z, \rho_{p+}) &= \frac{\rho_{p+}+1}{\rho_{p+}} \exp\left(-\frac{\rho_{p+}+1}{\rho_{p+}}z\right)u(z) \\ f_{\zeta_{p+}|\mathcal{H}_1}(z, \rho_{p+}) &= \frac{1}{\rho_{p+}} \exp\left(-\frac{1}{\rho_{p+}}z\right)u(z). \end{aligned} \quad (123)$$

Thus,

$$\frac{\partial \varphi_i}{\partial \gamma} = -f_i(\gamma) * f_{\zeta_{mn}|\mathcal{H}_i}(\gamma, \rho_{mn}) \quad (124)$$

$$\frac{\partial \varphi_i}{\partial \rho_{mn}} = \int_{\gamma}^{\infty} f_i(z) * \frac{\partial f_{\zeta_{mn}|\mathcal{H}_i}(z, \rho_{mn})}{\partial \rho_{mn}} dz \quad (125)$$

From (124), $\frac{\partial \varphi_i}{\partial \gamma} < 0$, $i = 0, 1$. Thus, from (118) and (119), (115) is equivalent to

$$\frac{\partial \varphi_1}{\partial \gamma} \frac{\partial \varphi_0}{\partial \rho_{mn}} - \frac{\partial \varphi_0}{\partial \gamma} \frac{\partial \varphi_1}{\partial \rho_{mn}} > 0, \quad \forall \rho_{mn} > 0. \quad (126)$$

Plugging (124) and (125) into (126), we obtain

$$\begin{aligned} &\int_{\gamma}^{\infty} f_1(z) * \frac{\partial f_{\zeta_{mn}|\mathcal{H}_1}(z, \rho_{mn})}{\partial \rho_{mn}} dz \\ &\quad \times (f_0(z) * f_{\zeta_{mn}|\mathcal{H}_0}(z, \rho_{mn})) \\ &- \int_{\gamma}^{\infty} f_0(z) * \frac{\partial f_{\zeta_{mn}|\mathcal{H}_0}(z, \rho_{mn})}{\partial \rho_{mn}} dz \\ &\quad \times (f_1(z) * f_{\zeta_{mn}|\mathcal{H}_1}(z, \rho_{mn})) > 0. \end{aligned} \quad (127)$$

In (127),

$$\begin{aligned} &\int_{\gamma}^{\infty} f_1(z) * \frac{\partial f_{\zeta_{mn}|\mathcal{H}_1}(z, \rho_{mn})}{\partial \rho_{mn}} dz \\ &= f_1(\gamma) * \frac{\partial f_{\zeta_{mn}|\mathcal{H}_1}(\gamma, \rho_{mn})}{\partial \rho_{mn}} * u(-\gamma) \\ &= f_1(\gamma) * \left(\frac{-\gamma}{\rho_{mn}^2} \exp\left(-\frac{1}{\rho_{mn}}\gamma\right) \right) \\ &= f_1(\gamma) * f_{\zeta_{mn}|\mathcal{H}_1}(\gamma, \rho_{mn}) * f_{\zeta_{mn}|\mathcal{H}_1}(\gamma, \rho_{mn}). \end{aligned} \quad (128)$$

Similarly,

$$\begin{aligned} &\int_{\gamma}^{\infty} f_0(z) * \frac{\partial f_{\zeta_{mn}|\mathcal{H}_0}(z, \rho_{mn})}{\partial \rho_{mn}} dz \\ &= \frac{1}{(\rho_{mn}+1)^2} f_0(\gamma) * f_{\zeta_{mn}|\mathcal{H}_0}(\gamma, \rho_{mn}) * f_{\zeta_{mn}|\mathcal{H}_0}(\gamma, \rho_{mn}) \end{aligned} \quad (129)$$

Thus, (127) is equivalent to

$$\begin{aligned} &f_1(\gamma) * f_{\zeta_{mn}|\mathcal{H}_1}(\gamma, \rho_{mn}) * f_{\zeta_{mn}|\mathcal{H}_1}(\gamma, \rho_{mn}) \\ &\quad \times (f_0(\gamma) * f_{\zeta_{mn}|\mathcal{H}_0}(\gamma, \rho_{mn})) \\ &- \frac{1}{(\rho_{mn}+1)^2} f_0(\gamma) * f_{\zeta_{mn}|\mathcal{H}_0}(\gamma, \rho_{mn}) * f_{\zeta_{mn}|\mathcal{H}_0}(\gamma, \rho_{mn}) \\ &\quad \times (f_1(\gamma) * f_{\zeta_{mn}|\mathcal{H}_1}(\gamma, \rho_{mn})) > 0. \end{aligned} \quad (130)$$

⁶Different from (36), the convolution between two pdfs $g_1(z)$ and $g_2(z)$ is defined as $g_1(z) * g_2(z) = \int_{-\infty}^{\infty} g_1(z-y)g_2(y)dy$.

From (121), (122) and (123), the inequality (130) is equivalent to

$$\begin{aligned} & \int_{\mathfrak{D}} \frac{1}{\rho_{mn}^2 \prod_{p=1}^P \rho_{p+}} \exp\left(-\frac{1}{\rho_{mn}}(\gamma - z_{P+1} - \dots z_1)\right) \\ & \times \exp\left(-\frac{1}{\rho_{1+}} z_1\right) \dots \exp\left(-\frac{1}{\rho_{P+}} z_P\right) \exp\left(-\frac{1}{\rho_{mn}} z_{P+1}\right) \\ & \times \left(\left(f_0(\gamma) * f_{\zeta_{mn}|\mathcal{H}_0}(\gamma, \rho_{mn}) \right) \right. \\ & \left. - \left(f_1(\gamma) * f_{\zeta_{mn}|\mathcal{H}_1}(\gamma, \rho_{mn}) \right) \exp(-\gamma) \prod_{p=1}^P (1 + \rho_{p+}) \right) dz > 0. \end{aligned} \quad (131)$$

where $dz = dz_1 \dots dz_{P+1}$ and $\mathfrak{D} = \{z_1, z_2, \dots, z_{P+1} | z_1 \geq 0, \dots, z_{P+1} \geq 0, z_1 + z_2 + \dots + z_{P+1} \leq \gamma\}$. If we can prove

$$\begin{aligned} & \left(f_0(\gamma) * f_{\zeta_{mn}|\mathcal{H}_0}(\gamma, \rho_{mn}) \right) \\ & - \left(f_1(\gamma) * f_{\zeta_{mn}|\mathcal{H}_1}(\gamma, \rho_{mn}) \right) \exp(-\gamma) \prod_{p=1}^P (1 + \rho_{p+}) > 0, \end{aligned} \quad (132)$$

then (131) can hold. (132) is equivalent to

$$\begin{aligned} & \int_{\mathfrak{D}'} \frac{\prod_{p=1}^P (1 + \rho_{p+})}{\prod_{p=1}^P (\rho_{p+})} \exp(-\gamma) \exp\left(-\frac{1}{\rho_{mn}}(\gamma - z'_P - \dots z'_1)\right) \\ & \times \exp\left(-\frac{1}{\rho_{1+}} z'_1\right) \dots \exp\left(-\frac{1}{\rho_{P+}} z'_P\right) dz' > 0 \end{aligned} \quad (133)$$

where $dz' = dz'_1 \dots dz'_P$, $\mathfrak{D}' = \{z'_1, \dots, z'_P | z'_1 \geq 0, \dots, z'_P \geq 0, z'_1 + \dots + z'_P \leq \gamma\}$. Since (133) is obviously true, so are (132), this completes the proof of (115).

REFERENCES

- [1] E. Fishler, A. Haimovich, R. Blum, D. Chizhik, L. Cimini, and R. Valenzuela, "MIMO radar: an idea whose time has come," in *Proceedings of the 2004 IEEE Radar Conference (IEEE Cat. No. 04CH37509)*, April 2004, pp. 71–78.
- [2] A. M. Haimovich, R. S. Blum, and L. J. Cimini, "MIMO radar with widely separated antennas," *IEEE Signal Processing Magazine*, vol. 25, no. 1, pp. 116–129, 2008.
- [3] J. Li and P. Stoica, "MIMO radar with colocated antennas," *IEEE Signal Processing Magazine*, vol. 24, no. 5, pp. 106–114, Sept. 2007.
- [4] E. Fishler, A. Haimovich, R. Blum, R. Cimini, D. Chizhik, and R. Valenzuela, "Performance of MIMO radar systems: advantages of angular diversity," in *Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers, 2004.*, vol. 1, Nov. 2004, pp. 305–309.
- [5] P. Wang, H. Li, and B. Himed, "A parametric moving target detector for distributed MIMO radar in non-homogeneous environment," *IEEE Transactions on Signal Processing*, vol. 61, no. 9, pp. 2282–2294, May 2013.
- [6] M. Rossi, A. M. Haimovich, and Y. C. Eldar, "Spatial compressive sensing for MIMO radar," *IEEE Transactions on Signal Processing*, vol. 62, no. 2, pp. 419–430, Jan. 2014.
- [7] Y. Yu, A. P. Petropulu, and H. V. Poor, "MIMO radar using compressive sampling," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 1, pp. 146–163, Feb. 2010.
- [8] T. Aittomaki and V. Koivunen, "Performance of MIMO radar with angular diversity under swerling scattering models," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 1, pp. 101–114, Feb. 2010.
- [9] J. Liu, S. Zhou, W. Liu, J. Zheng, H. Liu, and J. Li, "Tunable adaptive detection in colocated MIMO radar," *IEEE Transactions on Signal Processing*, vol. 66, no. 4, pp. 1080–1092, Feb. 2018.
- [10] J. Liu, H. Li, and B. Himed, "Persymmetric adaptive target detection with distributed MIMO radar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, no. 1, pp. 372–382, Jan. 2015.
- [11] Z. Cheng, Z. He, Z. Wang, and Y. Lu, "Detection performance analysis for distributed mimo radar," *Journal of Radars*, vol. 6, no. 1, pp. 81–89, 2017.
- [12] E. Fishler, A. Haimovich, R. S. Blum, L. J. Cimini, D. Chizhik, and R. A. Valenzuela, "Spatial diversity in radar models and detection performance," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 823–838, March 2006.
- [13] K. Pölönen and V. Koivunen, "Detection of DVB-T2 control symbols in passive radar systems," in *2012 IEEE 7th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, June 2012, pp. 309–312.
- [14] J. Liu, H. Li, and B. Himed, "Two target detection algorithms for passive multistatic radar," *IEEE Transactions on Signal Processing*, vol. 62, no. 22, pp. 5930–5939, Nov. 2014.
- [15] Q. Wu, Y. D. Zhang, M. G. Amin, and B. Himed, "Space-time adaptive processing and motion parameter estimation in multistatic passive radar using sparse Bayesian learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 2, pp. 944–957, Feb. 2016.
- [16] D. E. Hack, L. K. Patton, B. Himed, and M. A. Saville, "Detection in passive MIMO radar networks," *IEEE Transactions on Signal Processing*, vol. 62, no. 11, pp. 2999–3012, 2014.
- [17] —, "Centralized passive MIMO radar detection without direct-path reference signals," *IEEE Transactions on Signal Processing*, vol. 62, no. 11, pp. 3013–3023, 2014.
- [18] S. Gogineni, M. Rangaswamy, B. D. Rigling, and A. Nehorai, "Cramer-Rao bounds for UMTS-based passive multistatic radar," *IEEE Transactions on Signal Processing*, vol. 62, no. 1, pp. 95–106, 2014.
- [19] A. Filip and D. Shutin, "Cramér-Rao bounds for L-band digital aeronautical communication system type 1 based passive multiple-input multiple-output radar," *IET Radar, Sonar and Navigation*, vol. 10, no. 2, pp. 348–358, 2015.
- [20] S. Gogineni, M. Rangaswamy, B. D. Rigling, and A. Nehorai, "Ambiguity function analysis for UMTS-based passive multistatic radar," *IEEE Transactions on Signal Processing*, vol. 62, no. 11, pp. 2945–2957, 2014.
- [21] J. Yi, X. Wan, H. Leung, and F. Cheng, "MIMO passive radar tracking under a single frequency network," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 8, pp. 1661–1671, 2015.
- [22] R. W. Heath Jr and A. Paulraj, "Antenna selection for spatial multiplexing systems based on minimum error rate," in *Communications, 2001. ICC 2001. IEEE International Conference on*, vol. 7, IEEE, 2001, pp. 2276–2280.
- [23] D. Gore, R. Heath, and A. Paulraj, "Statistical antenna selection for spatial multiplexing systems," in *Communications, 2002. ICC 2002. IEEE International Conference on*, vol. 1, IEEE, 2002, pp. 450–454.
- [24] D. Gore, A. J. Paulraj et al., "MIMO antenna subset selection with space-time coding," *Signal Processing, IEEE Transactions on*, vol. 50, no. 10, pp. 2580–2588, 2002.
- [25] L. Dai, S. Sfar, and K. Letaief, "Optimal antenna selection based on capacity maximization for MIMO systems in correlated channels," *IEEE Transactions on Communications*, vol. 54, no. 3, pp. 563–573, 2006.
- [26] L. M. Kaplan, "Global node selection for localization in a distributed sensor network," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, no. 1, pp. 113–135, 2006.
- [27] J. Akhtar, "Antenna selection for MIMO radar transmitters," in *Radar Conference (RADAR), 2011 IEEE*, 2011, pp. 018–021.
- [28] H. Godrich, A. P. Petropulu, and H. V. Poor, "Sensor selection in distributed multiple-radar architectures for localization: A knapsack problem formulation," *IEEE Transactions on Signal Processing*, vol. 60, no. 1, pp. 247–260, 2012.
- [29] C. Shi, F. Wang, M. Sellathurai, and J. Zhou, "Transmitter subset selection in FM-based passive radar networks for joint target parameter estimation," *IEEE Sensors Journal*, vol. 16, no. 15, pp. 6043–6052, 2016.
- [30] B. Ma, H. Chen, B. Sun, and H. Xiao, "A joint scheme of antenna selection and power allocation for localization in MIMO radar sensor networks," in *Signal Processing (ICSP), 2014 12th International Conference on*, IEEE, 2014, pp. 2226–2229.

- [31] B. Sun, H. Chen, D. Yang, and X. Li, "Antenna selection and placement analysis of MIMO radar networks for target localization," *International Journal of Distributed Sensor Networks*, vol. 2014, 2014.
- [32] S. Joshi and S. Boyd, "Sensor selection via convex optimization," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 451–462, Feb 2009.
- [33] S. Eguchi and J. Copas, "Interpreting Kullback–Leibler divergence with the Neyman–Pearson lemma," *Journal of Multivariate Analysis*, vol. 97, no. 9, pp. 2034–2040, 2006.
- [34] M. Skolnik, G. Nemhauser, and J. Sherman, "Dynamic programming applied to unequally spaced arrays," *IEEE Transactions on Antennas and Propagation*, vol. 12, no. 1, pp. 35–43, 1964.
- [35] Y. I. Abramovich, A. Y. Gorokhov, N. K. Spencer, and P. Turcaj, "Adaptive switched antenna control: from radar to MIMO applications," in *Sensor Array and Multichannel Signal Processing Workshop Proceedings, 2004.* IEEE, 2004, pp. 494–498.
- [36] A. Gorokhov, D. A. Gore, and A. J. Paulraj, "Receive antenna selection for MIMO spatial multiplexing: theory and algorithms," *IEEE Transactions on signal processing*, vol. 51, no. 11, pp. 2796–2807, 2003.
- [37] S. P. Chepuri and G. Leus, "Sparse sensing for distributed Gaussian detection," in *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on.* IEEE, 2015, pp. 2394–2398.
- [38] —, "Sparse sensing for distributed detection," *IEEE Transactions on Signal Processing*, vol. 64, no. 6, pp. 1446–1460, March 2016.
- [39] A. N. Shiryaev, *Probability.* Springer-Verlag, New York, 1996.
- [40] D. Bajović, B. Sinopoli, and J. Xavier, "Sensor selection for event detection in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4938–4953, 2011.
- [41] H. L. Van Trees, *Detection, estimation, and modulation theory.* John Wiley & Sons, 2004.