Communications-Inspired Sensing: A Case Study on Waveform Design

Wenshu Zhang, Student Member, IEEE, and Liuqing Yang*, Senior Member, IEEE

Abstract—Information theory, and particularly the mutual information (MI), has provided fundamental guidance for communications research. In Bell’s 1993 paper, the MI was first applied to radar waveform design. Similar to its communications counterpart, the solution comes in a water-filling form. However, the practical meaning of MI in the sensing context remains unclear to date. Recently, Yang and Blum’s 2007 paper shows that under the white noise assumption, the optimum water-filling scheme simultaneously maximizes the MI and minimizes the estimation minimum mean square error (MMSE). Such an equivalence, however, does not hold when the target parameter statistics are not perfectly known as shown in Yang and Blum’s subsequent work. To further the understanding of the practical meaning of MI and to establish a connection between the MI and commonly adopted MSE measures for sensing, this paper takes a fresh look at the target estimation problem. We consider the general colored noise, incorporate the normalized MSE (NMSE), and develop joint robust designs for both the transmitter (waveforms) and the receiver (estimator) under various target and noise uncertainty models. Our results show that: i) the optimum waveform designs resulted from the MI, MMSE and NMSE criteria are all different; and ii) compared to MMSE, the NMSE-based designs share more similarities with the MI-based ones, especially when the target and noise statistics are not perfectly known.

I. INTRODUCTION

In multi-input multi-output (MIMO) communication systems, multiple transmit and/or receive antennas can increase the diversity to combat channel fading for enhanced transmission reliability and increase the degrees of freedom for improved data rate. Partly inspired by these benefits, MIMO sensing has drawn great interests in recent years (see e.g., [1]–[5], [7]–[10]). In such systems, a particularly critical issue is the waveform optimization. Solutions to this problem mainly fall into two categories, the space-time correlation optimization of the transmitted waveforms (see e.g., [2], [15], [16]), and the specific time-domain signal design given the desired space-time correlation properties (see e.g., [7]). In this paper, we focus on the former. Bell’s 1993 paper first used the mutual information (MI) to design radar waveforms for the estimation of an extended target [2]. His MI-based water-filling approach has been extended by several recent works.

In particular, [15] deals with multiple extended targets using a large coherent phased array, and [16] considers the detection of an extended target. Both of them adopt the MI as the optimization criterion.

The MI is an essential measure in the field of communications. However, its role in sensing is not yet clear. In an attempt to link the MI criterion with more direct performance indicators in sensing, and particularly target parameter estimation applications, Yang and Blum studied the extended target estimation problem in a widely separated MIMO radar scenario. In [25], it is shown that the MI and the minimum mean square error (MMSE) criteria lead to the same optimum water-filling strategy, assuming perfectly known target and white noise power spectral densities (PSDs). These waveform designs were then extended in [26] to account for bounded uncertainty in the target PSD. In contrast to [25], [26] shows that the MI and MMSE criteria result in distinct waveform designs.

Though the results in [25] and [26] shed some light on the possible connection between the MI and MMSE measures, they are based on limiting assumptions such as white noise and perfectly known noise PSD. In this paper, we will further these existing works and reveal more intrinsic connections between the MI and MSE measures in a sensing setup.

For comparison convenience, we will consider a MIMO radar setup as in [25] and [26]. However, rather than the widely separated MIMO radar, we employ a mixed MIMO structure with widely separated transmit array elements and closely spaced receive array elements. This is due to the inherent identifiability issue associated with the former setup as will be detailed in Section II.

Our contributions in this paper are three-fold. First, we take into consideration the more general and practical colored Gaussian noise that can emerge in various situations. For example, the received signal may be affected by unwanted interferences including jammers. The noise spectrum might also be shaped by the antenna and RF filters [22]. It turns out that in the presence of colored Gaussian noise, the equivalence between the MI and MMSE design criteria established in [25] does not hold, even when the target and noise PSDs are both perfectly known. Secondly, we introduce the normalized MSE (NMSE) minimizing criterion for radar waveform designs. Not only is it more meaningful for parameter estimation problems, but it also exhibits more similar behaviors with the MI criterion than its MMSE counterpart, especially in robust designs. Last but not least, we provide joint robust designs for both the probing waveforms at the transmitter and the estimator at the receiver under various uncertainty models. Compared
with [26], our improvements include: i) we consider colored noise instead of white noise; ii) we jointly optimize both the transmitter (waveforms) and the receiver (estimator) instead of limiting only to the transmitter side; and iii) we account for the uncertainty for both target and noise PSDs instead of assuming perfectly known noise PSD. Results show that the MI- and NMSE-based robust designs are built on an identical least favorable set (LFS), which differs from the LFS of the MMSE-based designs.

The organization of this paper is as follows. The system model is given in Section II. The three design criteria (MI, MMSE and NMSE) and their corresponding optimum waveform designs in the presence of colored noise are introduced in Section III. The joint estimator and power loading robust designs are discussed in Section IV for all three criteria. Numerical results are given in Section V, followed by the concluding remarks in Section VI.

Notation: We will use boldface lowercase $a$ and uppercase $A$ letters for vectors and matrices, respectively. $E\{\cdot\}$ will be used for expectation, $\text{tr}\{\cdot\}$ for trace of matrix $\mathbf{A}$, and $|\mathbf{A}|$ for determinant of $\mathbf{A}$. $\otimes$ refers to Kronecker product, superscript $\{\cdot\}^H$ for matrix Hermitian and $\{\cdot\}^T$ for matrix transposition. $\mathbf{I}_N$ denotes the $N \times N$ identity matrix, and $|a|^\dagger \triangleq \max\{0, a\}$.

II. SYSTEM MODEL

Despite the many similarities between MIMO communications and MIMO radar, they have some fundamental differences. Take the transmitted signal optimization as an example, where we consider a simple setup of $M = 2$ transmit antennas and $N = 2$ receive antennas.

In communications, the objective is to optimize the transmitted signals for better estimation of themselves. Assume that the duration of the transmitted signals is $L$, length of the channel delay is $K$, and the white noise $\xi$ is zero-mean Gaussian distributed. The system representation can be expressed as:

$$
\begin{bmatrix}
\mathbf{r}_1 \\
\mathbf{r}_2
\end{bmatrix} = \mathbf{H} \begin{bmatrix}
\mathbf{x}_1 \\
\mathbf{x}_2
\end{bmatrix} + \xi = \begin{bmatrix}
\mathbf{H}_{1,1} & \mathbf{H}_{1,2} \\
\mathbf{H}_{2,1} & \mathbf{H}_{2,2}
\end{bmatrix} \begin{bmatrix}
\mathbf{x}_1 \\
\mathbf{x}_2
\end{bmatrix} + \xi ,
$$

(1)

where $\mathbf{x}_i$ is the $L \times 1$ transmitted signal vector emitted from the $i$th transmitter, $\mathbf{r}_j$ is the $L \times 1$ received signal vector at the $j$th receiver, and the $L \times L$ Toeplitz matrix $\mathbf{H}_{j,i}$ represents the channel response from the $i$th transmitter to the $j$th receiver. Denote the covariance matrices as $\mathbf{\Sigma}_x$ for the transmitted signals, and $\mathbf{\Sigma}_\xi = \sigma_\xi^2 \mathbf{I}$ for the white noise. Then the mutual information (MI) between the transmitted and received signals is:

$$
\text{MI} = \log |\sigma_\xi^{-2} \mathbf{\Sigma}_x \mathbf{H}^H \mathbf{H} + \mathbf{I}_{2L}| ,
$$

(2)

and the resultant MMSE after the MMSE estimator is

$$
\text{MMSE} = \text{tr} \left\{ (\sigma_\xi^{-2} \mathbf{H}^H \mathbf{H} + \mathbf{\Sigma}_x)^{-1} \right\} .
$$

(3)

In communications, the classical optimum transmitted signal design maximizing the MI or minimizing the MMSE is achieved when $\left( \sigma_\xi^{-2} \mathbf{\Sigma}_x \mathbf{H}^H \mathbf{H} + \mathbf{I}_{2L} \right)$ or $\left( \sigma_\xi^{-2} \mathbf{H}^H \mathbf{H} + \mathbf{\Sigma}_x^{-1} \right)$ is a diagonal matrix (see e.g., [18]--[21]). The $2L$ eigenvalues of $\mathbf{\Sigma}_x$ are allocated according to the strength of the corresponding channel eigenvalues of $\mathbf{H}^H \mathbf{H}$.

On the other hand, for a $2 \times 2$ radar system with widely separated transmit and receive array elements, the total $M \times N = 4$ viewing aspects of the target can be acquired, and the signal model is given by [25]:

$$
\begin{bmatrix}
\mathbf{r}_1 \\
\mathbf{r}_2
\end{bmatrix} = \begin{bmatrix}
\mathbf{X} & \mathbf{0}_{L \times 2K} \\
\mathbf{0}_{L \times 2K} & \mathbf{X}
\end{bmatrix} \begin{bmatrix}
\mathbf{g}_1 \\
\mathbf{g}_2
\end{bmatrix} + \xi ,
$$

(4)

where $\mathbf{X} = [\mathbf{X}_1 \mathbf{X}_2]$ and $\mathbf{X}_i$ is the $L \times K$ Toeplitz signal matrix emitted from the $i$th transmit antenna, $\mathbf{r}_j$ is the $L \times 1$ signal vector at the $j$th receive antenna, and $\mathbf{g}_j = [\mathbf{g}_{j,1}^T \mathbf{g}_{j,2}^T]^T$ with $\mathbf{g}_{j,i}$ being the $K \times 1$ target viewing aspect from the $i$th transmit antenna to the $j$th receive antenna. Accordingly, the MI between the target response and the received signal is:

$$
\text{MI} = \log |\sigma_\xi^{-2} \mathbf{\Sigma}_g (\mathbf{I}_2 \otimes \mathbf{X}^H \mathbf{X}) + \mathbf{I}_{4K}| ,
$$

(5)

where $\mathbf{\Sigma}_g$ is the covariance matrix of the target response, and the resulting MMSE after the MMSE estimator is:

$$
\text{MMSE} = \text{tr} \left\{ (\sigma_\xi^{-2} \mathbf{I}_2 \otimes \mathbf{X}^H \mathbf{X} + \mathbf{\Sigma}_g)^{-1} \right\} .
$$

(6)

In order to maximize the MI or minimize the MMSE, the optimum strategy again requires $\left( \sigma_\xi^{-2} \mathbf{\Sigma}_g (\mathbf{I}_2 \otimes \mathbf{X}^H \mathbf{X}) + \mathbf{I}_{4K} \right)$ or $\left( \sigma_\xi^{-2} (\mathbf{I}_2 \otimes \mathbf{X}^H \mathbf{X}) + \mathbf{\Sigma}_g^{-1} \right)$ to be an optimum diagonal matrix, as detailed in [25]. However, unlike the communications case, the special repeated structure of $\mathbf{I}_2 \otimes \mathbf{X}^H \mathbf{X}$ makes such a condition impossible. In other words, one cannot design the 2 transmitted signals such that the estimation of 4 viewing aspects (namely $g_{11}$, $g_{12}$, $g_{21}$ and $g_{22}$) are optimized simultaneously -- there is simply a lack of sufficient degrees of freedom.

This simple comparison suggests that the transmitted signal optimization problem is ill-formation for the widely separated MIMO radar setup in [25]. To avoid this problem, we consider a mixed MIMO structure (see Fig. 1), which is equipped with a widely separated $M$-element transmit array and a closely spaced $N$-element receive array. For an extended target of interest, the $M$ transmitted waveforms can impinge distinct scatterers from different angles. On the other hand, for each of the $M$ reflected signals, the receiver acquires $N$ coherent returns, the only difference among which is a phase shift. One can then combine them coherently to obtain a processing gain of $N$. As a result, the target response is captured by the mixed MIMO setup from totally $M$ viewing aspects. Having $M$ transmitted signals to design, this mixed MIMO setup provides sufficient degrees of freedom for the signal design optimizing the estimation of all $M$ viewing aspects.

Bearing the goal of comparing the MI- and MSE-based radar waveform designs, we will borrow the “mode” space signal model from [25]. Though originally developed for the widely separated MIMO radar, it can be readily modified for our mixed MIMO setup. The coherent combining of the arrival signals at the closely spaced receive array is rather straightforward. Hence we set $N = 1$ here without loss of generality. As a result, the received waveform can be expressed as

$$
\mathbf{r}_1 = \mathbf{X} \mathbf{g} + \xi ,
$$

(7)
where $X = [X_1, \cdots, X_M]$, $g = [g_1^T, \cdots, g_M^T]^T$, $r_1$ is the \(L \times 1\) signal vector at the receive antenna, and \(\xi\) the \(L \times 1\) zero-mean Gaussian noise vector. To facilitate the target response estimation, it is required that \(L\) cannot be less than \(MK\). We set \(L = MK\) in our paper.

While the target response is assumed to be Gaussian distributed with full rank covariance matrix \(\Sigma_g\) [25], the zero-mean non-Gaussian noise has covariance matrix \(\Sigma_\xi\). Through eigenvalue decomposition, \(\Sigma_g\) and \(\Sigma_\xi\) can be diagonalized as

\[
\Sigma_g = U_g \Lambda_g U_g^H, \quad \Sigma_\xi = U_\xi \Lambda_\xi U_\xi^H,
\]

where the entries of the diagonal matrices \(\Lambda_g\) and \(\Lambda_\xi\) can be regarded as their corresponding PSD samples [25], respectively. In [25], the maximum MI or minimum MMSE is achieved when 

\[
X = \Psi D^2 U_g^H, \quad \Psi \text{ being an } L \times MK \text{ matrix with orthonormal columns and } D \text{ a diagonal matrix}
\]

Having \(L = MK\), we can generalize this result to the colored noise case by constraining the arbitrary matrix \(\Psi\) as \(U_\xi\), namely

\[
X = U_\xi D^2 U_g^H,
\]

which gives rise to the “mode” space system representation:

\[
y = D^2 h + \eta,
\]

where \(y \triangleq U_\xi^H r_1\) is defined as the \(MK \times 1\) observed signal in the mode space, \(h \triangleq U_g^H g\) is the \(MK \times 1\) “mode” vector capturing the response of the extended target, and \(\eta \triangleq U_\xi^H \xi\) the \(MK \times 1\) Gaussian noise vector in the mode space. Clearly, the covariance matrices for \(h\) and \(\eta\) are \(\Lambda = \Lambda_\xi\) and \(\Sigma_\eta = \Lambda_\xi\) diagonal matrices, respectively. \(D = \text{diag}\{d_1, \cdots, d_{MK}\}\) is the power allocation matrix with transmit power \(d_i\) allocated to the corresponding mode space waveform, subject to the total transmit power constraint \(\sum_{i=1}^{MK} d_i = P_0\).

With this representation, the waveform design problem simplifies to a power allocation problem, where the total power will be optimally assigned to \(MK\) orthogonal waveforms in the mode space. It is worth noting that, though we adopt the mode space representation in [25] for comparison with the results therein, this model is actually very general. Our model and the results hereafter can be readily generalized to cover “MIMO” radar systems resulted not only from multiple spatial viewing aspects, but also by alternative means such as frequency agility. Additionally, various receiver beamforming techniques (conventional [14], minimum variance distortionless response (MVDR) [6], and other adaptive algorithms [23]) can also be readily accommodated.

### III. Optimum Power Allocation in Colored Noise

The optimum power allocation problem has been studied in [25], under the white Gaussian noise assumption for which \(\Sigma_\eta = \sigma_\eta^2 I_{MK}\). In a sensing scenario, however, unwanted interferences including jammers and antenna effects are often inevitable, suggesting the necessity of incorporating more general colored noise. Throughout our analysis, we will consider colored Gaussian noise with zero-mean and covariance matrix \(\Sigma_\eta = \text{diag}\{\sigma_{\eta_1}^2, \cdots, \sigma_{\eta_{MK}}^2\}\), where the diagonal elements can be different.

In this section, we assume that both the target and noise PSDs are known exactly at both the transmitter and the receiver. This assumption will be relaxed in the next section, where uncertainty of such knowledge will be taken into account.

#### A. MI and MMSE Criteria

The MI between the observed signal \(y\) and the target mode response \(h\) given power allocation matrix \(D\) is:

\[
I(y; h|D) = \log |\Delta D \Sigma_\eta^{-1} I_{MK}| = \sum_{i=1}^{MK} \log (\sigma_{\eta_i}^{-2} \lambda_i d_i + 1) \quad (9)
\]

The logarithm is base-2 unless otherwise indicated, and the unit for MI is bit. Note that, instead of the white noise with a flat PSD in [25], we consider non-flat colored noise here.

**Proposition 1 (MI-based Optimum Power Allocation):** The optimum power allocation maximizing the MI in the presence of colored Gaussian noise has the following water-filling form:

\[
d_i = \left[\gamma_{MI} - \frac{\sigma_{\eta_i}^2}{\lambda_i}\right]^+, \quad \text{for } i = 1, \cdots, MK, \quad (10)
\]

where \(\gamma_{MI}\) is a constant satisfying the total power constraint \(\sum_{i=1}^{MK} d_i = P_0\).

**Proof:** The problem of optimum power allocation for maximizing the MI in (9) subject to the total power constraint can be formulated as:

\[
\max_{\{d_i\}} \sum_{i=1}^{MK} \log (\sigma_{\eta_i}^{-2} \lambda_i d_i + 1), \quad \text{subject to } \sum_{i=1}^{MK} d_i = P_0, \text{ and } d_i \geq 0, \text{ for } i = 1, \cdots, MK. \quad (11)
\]

This constrained optimization problem can be solved by the Lagrange multiplier method. Specifically, we first construct the objective function:

\[
J_{MI} = \sum_{i=1}^{MK} \log (\sigma_{\eta_i}^{-2} \lambda_i d_i + 1) + \gamma'_{MI} \left(\sum_{i=1}^{MK} d_i - P_0\right), \quad (12)
\]

and then differentiate it with respect to \(d_i\) and set it to zero. As a result, we get:

\[
d_i = \left[\gamma_{MI} - \frac{\sigma_{\eta_i}^2}{\lambda_i}\right]^+, \quad \text{for } i = 1, \cdots, MK, \quad (13)
\]

where \(\gamma_{MI} = -1/\gamma'_{MI}\) can be determined from the total power constraint.

---

*Fig. 1. Illustration of the mixed MIMO structure.*
For the MSE-based designs, one needs to first specify the MMSE estimator, denoted by $\Phi$, as follows:

$$\hat{h} = \Phi y = \left(D\Sigma^{-1}_\eta + \Lambda^{-1}\right)^{-1}D^\top\Sigma^{-1}_\eta y.$$  \hfill (14)

Accordingly, the MMSE is given by

$$\text{MMSE} = \text{tr} \left\{ \mathbb{E} \left\{ (h - \hat{h}) (h - \hat{h})^\top \right\} \right\} = \text{tr} \left\{ \left(\Sigma^{-1}_\eta D + \Lambda^{-1}\right)^{-1} \right\} = \sum_{i=1}^{MK} \frac{\lambda_i}{\sigma_i^2 + d_i} + 1.$$  \hfill (15)

**Proposition 2 (MMSE-based Optimum Power Allocation):**
The optimum power allocation minimizing the MMSE in the presence of colored Gaussian noise has the following form:

$$d_i = \lambda_i \left[ \frac{\sigma_i^2}{\lambda_i} - \frac{\sigma_i^2}{\lambda_i} \right], \text{ for } i = 1, \ldots, MK,$$  \hfill (16)

where $\gamma_{\text{MMSE}}$ is a constant ensuring the total power constraint.

**Proof:** The problem of optimum power allocation for minimizing the MMSE in (15) subject to the total power constraint can be formulated as:

$$\min_{\{d_i\}} \sum_{i=1}^{MK} \frac{\lambda_i}{\sigma_i^2 + d_i} + 1,$$  \hfill (17)

subject to $\sum_{i=1}^{MK} d_i = P_0$, and $d_i \geq 0$, for $i = 1, \ldots, MK$.

Using the Lagrange multiplier method, by differentiating the following cost function with respect to $d_i$:

$$J_{\text{MMSE}} = \sum_{i=1}^{MK} \frac{\lambda_i}{\sigma_i^2 + d_i} + \gamma_{\text{MMSE}} \left( \sum_{i=1}^{MK} d_i - P_0 \right),$$  \hfill (18)

and then setting it to zero, we get:

$$d_i = \left[ \gamma_{\text{MMSE}} \sqrt{\frac{\sigma_i^2}{\lambda_i} - \frac{\sigma_i^2}{\lambda_i}} \right]^+, \text{ for } i = 1, \ldots, MK,$$  \hfill (19)

where $\gamma_{\text{MMSE}} = \sqrt{1/\gamma_{\text{MMSE}}}$ is a constant ensuring the total power constraint.

From Propositions 1 and 2, we see that in the presence of colored Gaussian noise, the MMSE criterion does not lead to the water-filling solution as in [25]. In other words, the MI and MMSE criteria now are not equivalent as observed in [25], when the additive noise is colored.

**B. NMSE Criterion**

While the MMSE design minimizes the sum of the target mode estimation errors, there is no guarantee on the MSES of individual modes. Additionally, it is possible for the weakest modes to be discarded due to the total power constraint. In a radar problem, however, these weak modes may assume significant information useful in describing the target [2]. A natural amendment to this problem is to introduce the normalized MSE (NMSE) criterion, which is a common exercise in various estimation problems (see e.g., [11], [17], [24]).

Specifically, normalizing the individual MSES with respect to their average strength, we obtain the following expression:

$$\text{NMSE} \triangleq \text{tr} \left\{ \mathbb{E} \left\{ \Lambda^{-1/2} \left(h - \hat{h}\right) \left(h - \hat{h}\right)^\top \Lambda^{-1/2} \right\} \right\} = \text{tr} \left\{ \left(\Sigma^{-1}_\eta D + \Lambda^{-1}\right)^{-1} \right\} = \sum_{i=1}^{MK} \frac{1}{\sigma_i^2 + d_i} + 1.$$  \hfill (20)

**Proposition 3 (NMSE-based Optimum Power Allocation):**
The optimum power allocation minimizing the NMSE in the presence of colored Gaussian noise has the following form:

$$d_i = \left[ \gamma_{\text{NMSE}} \sqrt{\frac{\sigma_i^2}{\lambda_i} - \frac{\sigma_i^2}{\lambda_i}} \right]^+, \text{ for } i = 1, \ldots, MK,$$  \hfill (21)

where $\gamma_{\text{NMSE}}$ is a constant ensuring the total power constraint.

**Proof:** The problem of optimum power allocation based on the NMSE criterion under a total power constraint can be formulated as:

$$\min_{\{d_i\}} \sum_{i=1}^{MK} \frac{1}{\sigma_i^2 + d_i} + 1,$$  \hfill (22)

subject to $\sum_{i=1}^{MK} d_i = P_0$, and $d_i \geq 0$, for $i = 1, \ldots, MK$.

As in proofs of Propositions 1 and 2, we construct the cost function:

$$J_{\text{NMSE}} = \sum_{i=1}^{MK} \frac{1}{\sigma_i^2 + d_i} + \gamma_{\text{NMSE}} \left( \sum_{i=1}^{MK} d_i - P_0 \right),$$  \hfill (23)

and differentiate it to obtain:

$$d_i = \left[ \gamma_{\text{NMSE}} \sqrt{\frac{\sigma_i^2}{\lambda_i} - \frac{\sigma_i^2}{\lambda_i}} \right]^+, \forall i,$$  \hfill (24)

where $\gamma_{\text{NMSE}} = \sqrt{\frac{1}{\gamma_{\text{NMSE}}}}$ can be determined from the total power constraint.

Results summarized in Propositions 1–3 show that the optimum power allocation obtained via maximizing the MI differs from those by minimizing the MMSE or NMSE, when colored noise is taken into account. These results are based on the exact knowledge of both the target and noise PSDs. In practice, however, this knowledge can only be acquired on the exact knowledge of both the target and noise PSDs.

IV. JOINT ESTIMATOR AND POWER ALLOCATION ROBUST DESIGN

When the PSDs of the target and the colored noise are not precisely available, some robust approaches need to be utilized to design the probing waveforms at the transmitter, as well as the estimator at the receiver. [26] adopted the minimax approach to address the bounded target PSD uncertainty, while assuming that the white noise PSD is perfectly known. Here, we not only consider a colored noise PSD which includes
unwanted interferences such as jamming signals often encountered by a sensing system, but also allow for various uncertainty levels in the noise PSD. Another significant difference from [26] is: [26] focuses only on the robust design for the transmitted waveforms while assuming that the exact target PSD is available at the receiver; whereas we jointly design the transmitted waveforms and the MMSE estimator via a robust procedure.

A. Robust Minimax Design Criteria

In our minimax problem formulation, the estimator and the power allocation are jointly designed to: i) provide the optimum performance for the least favorable set (LFS) of the target and noise PSDs; and ii) provide equivalent or better performance for all other possible sets within the uncertainty regions. Mathematically, we jointly design the MMSE estimator matrix $\Phi$ and the power allocation matrix $D$ such that [12], [13]:

$$
\max_D \left\{ \inf_{\Lambda, \Sigma} \text{MI}(\Lambda, \Sigma; D) \right\}, \quad \text{MI-based},
$$

$$
\min_{\Phi, D} \left\{ \sup_{\Lambda, \Sigma} \text{MMSE}(\Lambda, \Sigma; \Phi, D) \right\}, \quad \text{MMSE-based}, \quad (25)
$$

$$
\min_{\Phi, D} \left\{ \sup_{\Lambda, \Sigma} \text{NMSE}(\Lambda, \Sigma; \Phi, D) \right\}, \quad \text{NMSE-based}.
$$

To solve the above problems, we will look for a saddle point for each criterion by finding the LFS $\left\{ \Lambda^R = \text{diag}\{\lambda_1^R, \ldots, \lambda^R_{M_K}\}, \Sigma^R = \text{diag}\{\sigma^R_{\eta_1}, \ldots, \sigma^R_{\eta_{M_K}}\} \right\}$, the robust estimator $\Phi^R$ and the robust power allocation matrix $D^R = \text{diag}\{d_1^R, \ldots, d^R_{M_K}\}$ satisfying:

$$
\min_{\Lambda, \Sigma} \text{MI}(\Lambda, \Sigma; D^R) = \text{MI}(\Lambda^R, \Sigma^R; D^R)
= \max_D \text{MI}(\Lambda^R, \Sigma^R; D),
$$

$$
\max_{\Lambda, \Sigma} \text{MMSE}(\Lambda, \Sigma; \Phi^R, D^R) = \text{MMSE}(\Lambda^R, \Sigma^R; \Phi^R, D^R)
= \min_{\Phi, D} \text{MMSE}(\Lambda^R, \Sigma^R; \Phi, D),
$$

$$
\max_{\Lambda, \Sigma} \text{NMSE}(\Lambda, \Sigma; \Phi^R, D^R) = \text{NMSE}(\Lambda^R, \Sigma^R; \Phi^R, D^R)
= \min_{\Phi, D} \text{NMSE}(\Lambda^R, \Sigma^R; \Phi, D).
$$

Note that once the LFSs $\Lambda^R$ and $\Sigma^R$ are determined, the second equalities in the above saddle point condition can be easily achieved through the optimum design that has been introduced in Section III. So the minimax robust procedure will focus on the first equalities only, resulting in an equivalent saddle point condition:

$$
\text{MI}(\Lambda, \Sigma; D^R) - \text{MI}(\Lambda^R, \Sigma^R; D^R) \geq 0
$$

$$
\text{MMSE}(\Lambda, \Sigma; \Phi^R, D^R) - \text{MMSE}(\Lambda^R, \Sigma^R; \Phi^R, D^R) \leq 0
$$

$$
\text{NMSE}(\Lambda, \Sigma; \Phi^R, D^R) - \text{NMSE}(\Lambda^R, \Sigma^R; \Phi^R, D^R) \leq 0
$$

Next, the explicit expressions of the differences on the left hand sides of (26) will be derived for the MI, MMSE and NMSE criteria to facilitate the identification of their respective LFSs.

**MI-based:**

As defined in (9), the MI formula has nothing to do with the receiver design. To calculate the MI difference in (26), one can substitute with (9) as follows:

$$
\text{MI}(\Lambda, \Sigma; D^R) - \text{MI}(\Lambda^R, \Sigma^R; D^R)
= \sum_{i=1}^{M_K} \log \left( \frac{\sigma^R_{\eta_i} \lambda_i d_i^R + 1}{\sigma^R_{\eta_i} \lambda_i d_i^R + 1} \right)
= \sum_{i=1}^{M_K} \log \left( \frac{\sigma^R_{\eta_i} \lambda_i d_i^R + 1}{\sigma^R_{\eta_i} \lambda_i d_i^R + 1} \right),
$$

where the robust power allocation $D^R$ is simply the optimum one for the LFS [c.f. (10)]:

$$
d_i^R = \lambda_i^R \left( d_i^R - \frac{\sigma^R_{\eta_i} \lambda_i^R}{\sigma^R_{\eta_i} \lambda_i} \right), \quad \forall i.
$$

In Section IV-B, we will find the LFS for MI using (27) and (28).

**MMSE-based:**

Calculation of the MMSE difference is more complicated than the MI case, because one now needs to jointly consider the power allocation $D^R$ and the estimator $\Phi^R$.

As mentioned before, the robust power allocation matrix $D^R$ is optimum for the LFS; that is [c.f. (16)]:

$$
d_i^R = \lambda_i^R \left( d_i^R - \frac{\sigma^R_{\eta_i} \lambda_i^R}{\sigma^R_{\eta_i} \lambda_i} \right), \quad \forall i.
$$

Likewise, the robust estimator $\Phi^R$ is also optimum (MSE-minimizing) for the LFS. From (14), we have:

$$
\Phi^R = \left( D^R \left( \Sigma^R_{\eta} \right)^{-1} + \left( \Lambda^R \right)^{-1} \right)^{-1} \left( D^R \right)^{\frac{1}{2}} \left( \Sigma^R_{\eta} \right)^{-1},
$$

which results in the MMSE for the LFS [c.f. (15)]:

$$
\text{MMSE}(\Lambda^R, \Sigma^R; \Phi^R, D^R) = \sum_{i=1}^{M_K} \frac{\lambda_i^R}{\sigma^R_{\eta_i} \lambda_i d_i^R + 1} + \left( \frac{\sigma^R_{\eta_i} \lambda_i^R}{\sigma^R_{\eta_i} \lambda_i} \right)^2 d_i^R
$$

where the last equality seems redundant but will make our later computations easier.

In the lack of exact knowledge of the true target and noise PSDs, one should always use the robust power allocation design $D^R$ at the transmitter, and the robust estimator design $\Phi^R$ at the receiver. Accordingly, for target response $h$ with arbitrary PSD $\Lambda$ and noise $\eta$ with $\Sigma_{\eta}$, we have:

$$
\text{MMSE}(\Lambda, \Sigma_{\eta}; \Phi^R, D^R)
= \text{tr} \left\{ \text{E} \left\{ (h - \hat{h}) (h - \hat{h})^H \right\} \right\}
= \text{tr} \left\{ \text{E} \left\{ (h - \Phi^R (D^R h + \eta)) (h - \Phi^R (D^R h + \eta))^H \right\} \right\}
$$

$$
= \sum_{i=1}^{M_K} \frac{\lambda_i}{\left( \sigma^R_{\eta_i} \lambda_i d_i^R + 1 \right)^2} + \frac{\left( \sigma^R_{\eta_i} \lambda_i^R d_i^R + 1 \right)^2 \sigma^R_{\eta_i}}{\left( \sigma^R_{\eta_i} \lambda_i^R d_i^R + 1 \right)^2 \sigma^R_{\eta_i}}.
$$
Subtracting (31) from (32), we obtain the MMSE difference in (26) as:

\[
\text{MMSE}(\Lambda, \Sigma_{\eta}; \Phi^R, D^R) - \text{MMSE}(\Lambda^R, \Sigma_{\eta}; \Phi^R, D^R) = \sum_{i=1}^{MK} \left( \frac{\lambda_i - \lambda_i^R}{\sigma_{\eta_i}^2 R \lambda_i^R d_i^R + 1} + \frac{\sigma_{\eta_i}^2 R \lambda_i^R d_i^R + 1}{(\sigma_{\eta_i}^2 R \lambda_i^R d_i^R + 1)^2} \right) \right) .
\]

**NMSE-based:**

Using the robust MMSE estimator \( \hat{\Phi}^R \), we obtain the NMSE for the LFS as [c.f. (20)]:

\[
\text{NMSE}(\Lambda, \Sigma_{\eta}; \hat{\Phi}^R, D^R) = \text{tr} \left\{ \mathbb{E} \left\{ (\hat{\Phi}^R)^{-1} (h - \hat{h}) (h - \hat{h})^\ast (\hat{\Phi}^R)^{-1} \right\} \right\}
\]

\[
= \sum_{i=1}^{MK} \frac{1}{\sigma_{\eta_i}^2 R \lambda_i^R d_i^R + 1} + \frac{\sigma_{\eta_i}^2 R \lambda_i^R d_i^R + 1}{(\sigma_{\eta_i}^2 R \lambda_i^R d_i^R + 1)^2} \lambda_i^R \right) ,
\]

where \( d_i^R \)'s come from the optimum NMSE-based power allocation matrix \( D^R \) for the LFS [c.f. (21)]:

\[
d_i^R = \left[ \gamma_{\text{NMSE}}^R \left( \frac{\sigma_{\eta_i}^2 R}{\lambda_i^R} - \frac{\sigma_{\eta_i}^2 R}{\lambda_i^R} \right) \right] , \quad \forall i .
\]

For target response and noise with arbitrary PSDs, we have:

\[
\text{NMSE}(\Lambda, \Sigma_{\eta}; \Phi^R, D^R) = \text{tr} \left\{ \mathbb{E} \left\{ (\Phi^R)^{-1} (h - \hat{h}) (h - \hat{h})^\ast (\Phi^R)^{-1} \right\} \right\}
\]

\[
= \sum_{i=1}^{MK} \left( \frac{1}{\sigma_{\eta_i}^2 R \lambda_i^R d_i^R + 1} + \frac{(\sigma_{\eta_i}^2 R \lambda_i^R)^2 d_i^R}{(\sigma_{\eta_i}^2 R \lambda_i^R d_i^R + 1)^2} \lambda_i^R \right) . \]

As a result, the NMSE difference is given by:

\[
\text{NMSE}(\Lambda, \Sigma_{\eta}; \Phi^R, D^R) - \text{NMSE}(\Lambda^R, \Sigma_{\eta}; \Phi^R, D^R) = \sum_{i=1}^{MK} \left( \frac{(\sigma_{\eta_i}^2 R \lambda_i^R)^2 d_i^R}{(\sigma_{\eta_i}^2 R \lambda_i^R d_i^R + 1)^2} \lambda_i^R \right) . \]

Recall that our robust minimax designs based on MI, MMSE and NMSE criteria are described by the three inequalities in (26), respectively. Now with the specific expressions available in (27), (33) and (37), we will next find the LFSs \((\Lambda^R, \Sigma_{\eta}^R)\) satisfying these inequalities.

**B. Joint Robust Designs**

We have seen in the preceding subsection that the robust transmitter and receiver designs \( D^R \) and \( \Phi^R \) will be uniquely specified once the LFS is determined. The LFS is not only determined by the inequalities in (26), but also heavily dependent on the uncertainty model. In this subsection, we will consider separately two models allowing for uncertainty in both target and noise PSDs.

**Uncertainty Model I:**

In [26], a banded uncertainty model for the target PSD is considered: the exact target PSD is unknown, but lies within a band whose upper and lower bounds are known. Here, we adopt this model for both the target and noise; that is,

\[
\lambda_i^L \leq \lambda_i \leq \lambda_i^U , \quad \sigma_{\eta_i}^{2L} \leq \sigma_{\eta_i}^2 \leq \sigma_{\eta_i}^{2U} , \quad \forall i . \]

**Proposition 4 (LFS for Uncertainty Model I):** When the uncertain target and noise PSDs fall within the banded regions with known upper and lower limits and with no other constraint, the LFS for the joint robust designs consists of:

- \( \{ \lambda_i^L \}_{i=1}^{MK} , \{ \sigma_{\eta_i}^{2L} \}_{i=1}^{MK} \) for MI criterion;
- \( \{ \lambda_i^L \}_{i=1}^{MK} , \{ \sigma_{\eta_i}^{2U} \}_{i=1}^{MK} \) for MMSE criterion; and
- \( \{ \lambda_i^L \}_{i=1}^{MK} , \{ \sigma_{\eta_i}^{2U} \}_{i=1}^{MK} \) for NMSE criterion.

**Proof:** See Appendix A.

We notice that the LFS for MI- and NMSE-based designs is identical, but differs from that of the MMSE-based design.

**Uncertainty Model II:**

In this model, we assume that the target PSD is known for simplicity and incorporate an average power ratio constraint on the noise PSD uncertainty. That is,

\[
\sigma_{\eta_i}^{2L} \leq \sigma_{\eta_i}^2 \leq \sigma_{\eta_i}^{2U} , \quad \forall i , \quad \text{and} \quad \frac{1}{MK} \sum_i \sigma_{\eta_i}^2 = \rho . \]

**Proposition 5 (LFS for Uncertainty Model II):** When the uncertain noise PSD falls within the banded region with known upper and lower limits under the average power ratio constraint, the LFS is given as follows:

\[
\sigma_{\eta_i}^{2R} = \begin{cases} 
\sigma_{\eta_i}^{2L} , & \text{if } \sigma_{\eta_i}^{2L} > k_n \lambda_i \\
\sigma_{\eta_i}^2 , & \text{if } \sigma_{\eta_i}^{2L} < k_n \lambda_i \\
\frac{k_m \lambda_i}{\lambda_i} , & \text{otherwise}
\end{cases}
\]

for MI and NMSE criteria, and

\[
\sigma_{\eta_i}^{2R} = \begin{cases} 
\sigma_{\eta_i}^{2L} , & \text{if } \sigma_{\eta_i}^{2L} > k_n \lambda_i^2 \\
\sigma_{\eta_i}^{2U} , & \text{if } \sigma_{\eta_i}^{2L} < k_n \lambda_i^2 \\
k_m \lambda_i^2 , & \text{otherwise}
\end{cases}
\]

for MMSE criterion, where \( k_n \) and \( k_m \) (typically \( k_n \neq k_m \)) are constants ensuring the average power ratio constraint and the target PSD is assumed to be known.

**Proof:** See Appendix B.

Proposition 5 is illustrated in Fig. 2. Under the average power ratio constraint, the MI and NMSE criteria again give rise to the same LFS. In fact, the LFS given by (40) is one where \( \{ \sigma_{\eta_i}^2 / \lambda_i \} \) is made as flat as possible. This result is
very intuitive since the worst interference PSD is the one that perfectly matches the target PSD. On the other hand, MMSE-based design suggests a different LFS where the noise PSD is maximally matched to the target PSD square. This result is consistent with the NMSE-based one, considering that the latter is obtained by normalizing the MSE with respect to the target PSD.

V. NUMERICAL RESULTS

In this section, we provide simulation results to verify our analytical conclusions and to provide further comparisons among the MI, MMSE and NMSE criteria.

A. Optimum Power Allocation in Colored Noise

In this simulation, we consider an extended target described by five modes. Fig. 3 gives the target and noise PSDs \( \{\lambda_i\} \) and \( \{\sigma^2_i\} \), for \( i = 1, \cdots, 5 \). The total power constraint is \( P_0 = 10 \) dB. Fig. 4 shows the optimum power allocation schemes for all three criteria in the presence of colored noise. Recall that under the assumption of white Gaussian noise, the MMSE and MI criteria lead to the same water-filling power allocation [25]. In colored noise, however, we can see from the figure that while the MI-based solution remains water-filling, the MMSE-based is not. In addition, they are both different from the NMSE-based solution. In terms of the optimum power allocation under the colored noise, no connection between the MSEs and the MI is observed.

The performance curves are plotted in Figs. 5, 6 and 7 for the three merits (MI, MMSE, and NMSE) based on the three criteria. Evidently, in terms of MI performance, the MI-based design is optimum, while the MMSE- and NMSE-based designs both exhibit performance losses. Similar observations
can be made for the MMSE and NMSE performance. These results agree well with our analysis that all the criteria are different in the presence of colored noise.

**B. Joint Estimator and Power Allocation Robust Design**

Since the LFS for the robust design with uncertainty model I is straightforward, here we will only verify the case with uncertainty model II. The target PSD is assumed to be available as shown in Fig. 3(a). Uncertainty in the colored noise normalized by target PSD \( \{\lambda_i\} \) is modeled as in Fig. 8(a), which has upper bound \( \sigma_{\eta_i}^2/\lambda_i \) and lower bound \( \sigma_{\eta_i}^{2L}/\lambda_i \) subject to the average power ratio constraint \( \rho = 1 \). The same noise uncertainty normalized by target PSD square \( \{\lambda_i^2\} \), is given in Fig. 8(b). A set of arbitrary nominal values is randomly chosen to make performance comparisons. The MI and NMSE criteria share the identical normalized noise LFS, which is supposed to be as flat as possible. In Fig. 8(a), when the constant \( k_n = 1 \), the straight line lies between the upper and lower bounds. According to the MI and NMSE criteria, the normalized noise LFS which satisfies the average power ratio constraint should be \( \sigma_{\eta_i}^{2R}/\lambda_i = 1, \forall i \). For MMSE criterion, on the other hand, the straight line \( k_m = 5\rho \sum \lambda_i = 1.0101 \) is lying between its corresponding upper and lower bounds in Fig. 8(b). Therefore the noise LFS for the MMSE criterion is \( \sigma_{\eta_i}^{2R}/\lambda_i^2 = 1.0101, \forall i \).

We plot the MMSE, NMSE and MI curves in Figs. 9, 10 and 11, respectively, to show how the robust design optimizes the worst case scenario. There are four curves for each case: A) nominal PSD with nominal design, which is the best achievable performance if there is no uncertainty; B) LFS PSD with nominal design, which is the case when one assumes the nominal PSD in design but encounters the LFS scenario; C) LFS PSD with robust design, which indicates the worst-case
performance for the robust design; and D) nominal PSD with robust design, which corresponds to the actual performance when the nominal PSD is applied but the robust design is used. For MMSE and NMSE criteria, large gaps between B (LFS PSD with nominal design) curves and C (LFS PSD with robust design) curves illustrate the significant improvement provided by the robust design for the worst case. This “best worst case” performance provides a performance lower bound. That is, for any PSD within the uncertainty region, the performance cannot be worse than this lower bound. This is evidenced by the D (nominal PSD with robust design) curves, which are bounded by C (LFS PSD with robust design) curves. For the MI-based design, we see that the performance is mainly determined by the actual PSD (nominal or LFS) but hardly affected by the power allocation design (nominal or robust), especially at high SNR. To find an explanation for this, let us go back to the power allocation scheme \( d_i = \left[ \gamma_{mi} - \frac{\sigma_{\eta i}^2}{\lambda_i} \right] + \) [c.f. (10)]. Subject to the average power ratio constraint \( \rho \), the water-filling level \( \gamma_{mi} = \frac{P_0}{MK\lambda_i} \) is the same for any possible PSD. Especially when \( P_0 \) is large, \( d_i \) is dominated by \( \gamma_{mi} \) and remains pretty much the same for both nominal and robust designs. That is the reason why the MI performance comes mainly from the different PSDs, but not from the different designs. We can remark that the robust design does not help much in this case. Despite the small quantitative improvement, Fig. 11 confirms that the robust design does improve the LFS performance as well as push up the MI lower bound.

VI. CONCLUSIONS

In this paper, we studied the optimum waveform design problem for target parameter estimation. Different from existing works, we considered a mixed MIMO radar setup for which the waveform optimization problem is meaningful, took into account the colored noise, incorporated the NMSE as a design criterion in addition to the MI and MMSE, and derived joint robust designs for both the transmitter (waveforms) and the receiver (estimator) under various uncertainty models. The analytical and numerical results suggest that: i) the equivalence between the MI and MMSE criteria does not hold when the noise is colored; and ii) compared to MMSE criterion, the NMSE-based criterion (due to the known target PSD assumption):  

\[
\text{NMSE}(\Lambda, \Sigma_{\eta i}, \Phi_R, D^R) - \text{NMSE}(\Lambda, \Sigma_{\eta i}^R, \Phi_R, D^R)
\]

Next, we will evaluate (43) in three different cases.

Case 1: \( d^R_i > 0, \forall i \). This means that we can remove the “+” sign in (43) and simplify it to:

\[
\text{NMSE}(\Lambda, \Sigma_{\eta i}, \Phi_R, D^R) - \text{NMSE}(\Lambda, \Sigma_{\eta i}^R, \Phi_R, D^R)
\]

\[
= \sum_{i=1}^{MK} \frac{\gamma_{mi}^2 \sigma_{\eta i}^2}{\lambda_i} \left[ \frac{\sigma_{\eta i}^2}{\lambda_i} - \frac{\sigma_{\eta i}^{2R}}{\lambda_i} \right]^2 + \left( \frac{\sigma_{\eta i}^2}{\lambda_i} - \frac{\sigma_{\eta i}^{2R}}{\lambda_i} \right) \right]^2 \right] + \left( \frac{\sigma_{\eta i}^2}{\lambda_i} - \frac{\sigma_{\eta i}^{2R}}{\lambda_i} \right) \right)^2 \right] \right)
\]

Define \( \Delta = \frac{\gamma_{mi}^2 \sigma_{\eta i}^2}{\lambda_i} \). As observed in Fig. 2(a), for all \( i \in \theta_0 \), \( \frac{\sigma_{\eta i}^2}{\lambda_i} = k_n \), and thus \( \Delta = \frac{\gamma_{mi}^2 \sigma_{\eta i}^2}{\lambda_i} \) remains a constant. Hence the summation in (44) over set \( \theta_0 \) is:

\[
\sum_{i=\theta_0} \frac{\gamma_{mi}^2 \sigma_{\eta i}^2}{\lambda_i} \left( \frac{\sigma_{\eta i}^2}{\lambda_i} - \frac{\sigma_{\eta i}^{2R}}{\lambda_i} \right) = c \sum_{i=\theta_0} \left( \frac{\sigma_{\eta i}^2}{\lambda_i} - \frac{\sigma_{\eta i}^{2R}}{\lambda_i} \right) \right].
\]

For all \( i \in \theta_+, \frac{\sigma_{\eta i}^2}{\lambda_i} - \frac{\sigma_{\eta i}^{2R}}{\lambda_i} \geq 0, \frac{\sigma_{\eta i}^{2R}}{\lambda_i} > k_n \), and thus \( \Delta < c \). So the summation in (44) over set \( \theta_+ \) satisfies

\[
\sum_{i=\theta_+} \frac{\gamma_{mi}^2 \sigma_{\eta i}^2}{\lambda_i} \left( \frac{\sigma_{\eta i}^2}{\lambda_i} - \frac{\sigma_{\eta i}^{2R}}{\lambda_i} \right) \leq c \sum_{i=\theta_+} \left( \frac{\sigma_{\eta i}^2}{\lambda_i} - \frac{\sigma_{\eta i}^{2R}}{\lambda_i} \right) \right].\]
For all \(i \in \theta_-, \frac{\sigma_n^2}{\lambda_i} - \frac{\sigma_T^2}{\lambda_i} \leq 0, \frac{\sigma_T^2}{\lambda_i} < k_n, \) and hence \(\Delta > c.\) We can obtain the summation over set \(\theta_-\) as:

\[
\sum_{i \in \theta_-} \frac{\gamma_{\text{MSE}}^R}{\sigma_n^2} - 1 \left( \frac{\sigma_n^2}{\lambda_i} - \frac{\sigma_T^2}{\lambda_i} \right) \leq c \sum_{i \in \theta_-} \frac{\sigma_n^2}{\lambda_i} \left( \frac{\sigma_T^2}{\lambda_i} \right).
\]  

(47)

Adding up (45), (46) and (47) will give us the right hand side of (44) as the following:

\[
\sum_{i=1}^{MK} \frac{\gamma_{\text{MSE}}^R}{\sigma_n^2} - 1 \left( \frac{\sigma_n^2}{\lambda_i} - \frac{\sigma_T^2}{\lambda_i} \right) \leq c \sum_{i=1}^{MK} \frac{\sigma_n^2}{\lambda_i} \left( \frac{\sigma_T^2}{\lambda_i} \right) = 0,
\]

where the last equality holds due to the average power ratio constraint. In other words, (44) becomes:

\[
\text{NMSE}(\Lambda, \Sigma_\eta, \Phi^R, D^R) - \text{NMSE}(\Lambda, \Sigma_\eta, \Phi^R, D^R) \leq 0. \quad (48)
\]

Case 2: Some \(d_i^R = 0\) on part of the \(\theta_+\) set where the weakest modes reside. This happens when the weakest modes are discarded due to the total power constraint. Proof for this case is the same as Case 1 for \(i \in \theta_0\) and \(i \in \theta_-\). The only difference occurs for those \(i \in \theta_+\), where (46) become

\[
\sum_{i \in \{\theta_+, \text{and } d_i^R > 0\}} \frac{\gamma_{\text{MSE}}^R}{\sigma_n^2} - 1 \left( \frac{\sigma_n^2}{\lambda_i} - \frac{\sigma_T^2}{\lambda_i} \right) \leq c \sum_{i \in \theta_+} \frac{\sigma_n^2}{\lambda_i} \left( \frac{\sigma_T^2}{\lambda_i} \right). \quad (49)
\]

This case can be easily extended to \(d_i^R = 0\) on the whole \(\theta_+\) set and part of the \(\theta_0\) subset.

Case 3: The total power is very limited so that \(d_i^R = 0\), for all \(i \in \theta_+ \cup \theta_0^c\) and for some \(i \in \theta_0^c\) or even \(i \in \theta_-\). In this case, (43) becomes:

\[
\sum_{i \in \{\theta_0^c \cup \theta_-\} \cup \theta_0^c \cup \theta_+} \frac{\gamma_{\text{MSE}}^R}{\sigma_n^2} - 1 \left( \frac{\sigma_n^2}{\lambda_i} - \frac{\sigma_T^2}{\lambda_i} \right) \leq 0. \quad (50)
\]

Summarizing all three cases, we have verified that the LFS for NMSE criterion in (40) satisfies the inequality condition in (26).

**MI-based:** Substitute (28) into (27), let \(\lambda_T^R = \lambda_i\) due to the known target PSD assumption, and exchange the order of the difference. Then (27) becomes:

\[
\text{MI}(\Lambda, \Sigma_\eta; D^R) - \text{MI}(\Lambda, \Sigma_\eta; D^R)
\]

\[
= \frac{1}{\ln 2} \sum_{i=1}^{MK} \frac{\lambda_i}{\sigma_n^2} \log d_i^R + \frac{\sigma_n^2}{\lambda_i} \log \frac{\sigma_n^2}{\lambda_i} + \frac{\sigma_T^2}{\lambda_i} \log \frac{\sigma_T^2}{\lambda_i} + 1
\]

\[
= \frac{1}{\ln 2} \sum_{i=1}^{MK} \left( \frac{\sigma_n^2}{\lambda_i} \log \frac{\sigma_n^2}{\lambda_i} + \frac{\sigma_T^2}{\lambda_i} \log \frac{\sigma_T^2}{\lambda_i} + 1 \right)
\]

\[
= \frac{1}{\ln 2} \sum_{i=1}^{MK} \left( \frac{\sigma_n^2}{\lambda_i} - \frac{\sigma_T^2}{\lambda_i} \right) \left( \frac{\sigma_T^2}{\lambda_i} - 1 \right)^+.
\]  

(51)

Similar to the NMSE-based proof, on the set partition \(\theta_0, \theta_+\) and \(\theta_-\), we need to evaluate (51) for LFS in three different cases.

**Case 1:** \(d_i^R > 0, \forall i\). This means that we can remove the “+” sign in (51):

\[
\text{MI}(\Lambda, \Sigma_\eta; D^R) - \text{MI}(\Lambda, \Sigma_\eta; D^R)
\]

\[
= \frac{1}{\ln 2} \sum_{i=1}^{MK} \frac{\sigma_n^2}{\lambda_i} \left( \frac{\sigma_T^2}{\lambda_i} - 1 \right) \left( \frac{\gamma_{\text{MSE}}^R}{\sigma_n^2} - 1 \right).
\]  

(52)

Define \(\Delta = \frac{\gamma_{\text{MSE}}^R}{\sigma_n^2} - 1\). As observed in Fig. 2(a), for all \(i \in \theta_0, \frac{\sigma_T^2}{\lambda_i} = k_n\), and thus \(\Delta = \frac{\gamma_{\text{MSE}}^R}{\sigma_n^2} - 1\) is a constant. Hence the summation in (52) over set \(\theta_0\) is

\[
\frac{1}{\ln 2} \sum_{i \in \theta_0} \frac{\sigma_n^2}{\lambda_i} \left( \frac{\gamma_{\text{MSE}}^R}{\sigma_n^2} - 1 \right). \quad (53)
\]

For all \(i \in \theta_+, \frac{\sigma_n^2}{\lambda_i} - \frac{\sigma_T^2}{\lambda_i} \leq 0, \frac{\sigma_n^2}{\lambda_i} > k_n, \) and thus \(\Delta < c.\) So the summation in (44) over set \(\theta_+\) satisfies

\[
\frac{1}{\ln 2} \sum_{i \in \theta_+} \frac{\sigma_n^2}{\lambda_i} \left( \frac{\gamma_{\text{MSE}}^R}{\sigma_n^2} - 1 \right).
\]  

(54)

Adding up (53), (54) and (55) will give us the right hand side
of (52) as the following:

\[
\frac{1}{\ln 2} \sum_{i=1}^{MK} \left( \frac{\sigma_{n_i}}{\lambda_i} - \frac{\sigma_{n_i}R}{\lambda_i} \right) \left( \frac{\gamma R}{\mu_i} \lambda_i - 1 \right) \leq \frac{c}{\ln 2} \sum_{i=1}^{MK} \left( \frac{\sigma_{n_i}^2}{\lambda_i} - \frac{\sigma_{n_i}R}{\lambda_i} \right) = 0,
\]

where the last equality holds due to the average power ratio constraint. In other words, (52) becomes:

\[
\text{MI}(\Lambda, \Sigma_n; D^R) - \text{MI}(\Lambda, \Sigma_n; D^R) \leq 0. \tag{56}
\]

Due to the space limit, here we only give the simple but tedious proof for Case 1 to illustrate the similarity and difference from the NMSE-based proof. Case 2 and Case 3 can be readily derived by analogy, which all result in the inequality condition \(\text{MI}(\Lambda, \Sigma_n; D^R) - \text{MI}(\Lambda, \Sigma_n; D^R) \leq 0\).

**MMSE-based:** A new partition should be defined for the mode index \(i\) based on the MMSE criterion, as marked in Fig. 2(b):

\[
\vartheta_0 \triangleq \left\{ i \in \{1, \ldots, MK\} : \frac{\sigma_{n_i}^2}{\lambda_i} \leq k_m \leq \frac{\rho_{n_i}^2}{\lambda_i} \right\}
\]

\[
\vartheta_+ \triangleq \left\{ i \in \{1, \ldots, MK\} : \frac{\sigma_{n_i}^2}{\lambda_i} > k_m \right\}
\]

\[
\vartheta_- \triangleq \left\{ i \in \{1, \ldots, MK\} : \frac{\sigma_{n_i}^2}{\lambda_i} < k_m \right\}.
\]

(57)

Similar to the NMSE- and MI-based cases, this new partition has the following property: \(\forall i \in \vartheta_+, \sigma_{n_i}R = \sigma_{n_i}^2\), and therefore \(\frac{\sigma_{n_i}^2}{\lambda_i} \geq \sigma_{n_i}^2; \forall i \in \vartheta_-\), \(\sigma_{n_i}R = \sigma_{n_i}^2\), and therefore \(\frac{\sigma_{n_i}^2}{\lambda_i} \leq \sigma_{n_i}^2\); and the set \(\vartheta_0\) can be further divided into two subsets \(\vartheta_0^+ = \left\{ i \in \vartheta_0 : \frac{\sigma_{n_i}^2}{\lambda_i} \geq \sigma_{n_i}^2 \right\}\) and \(\vartheta_0^- = \left\{ i \in \vartheta_0 : \frac{\sigma_{n_i}^2}{\lambda_i} < \sigma_{n_i}^2 \right\}\).

Substituting (29) into (33) and letting \(\lambda_i^R = \lambda_i\), we will obtain:

\[
\text{MMSE}(\Lambda, \Sigma_n; \Phi^R, D^R) - \text{MMSE}(\Lambda, \Sigma_n; \Phi^R, D^R) = \sum_{i=1}^{MK} \left( \frac{\sigma_{n_i}^2}{\lambda_i} \right)^2 \frac{R}{2} \text{MMSE} \left( \frac{\sigma_{n_i}^2}{\lambda_i} - \frac{\sigma_{n_i}^2R}{\lambda_i} \right).
\]

(58)

The procedure is similar to the one used for the NMSE proof and will be omitted here.

**REFERENCES**


Wenshu Zhang received the B.Sc. and M.Sc. degrees from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 2005 and 2007, respectively, both in Electrical Engineering. She is currently pursuing the Ph.D degree with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, Florida. Her research interests are in the areas of wireless communications and related fields.

Liuqing Yang (S’02-M’04-SM’06) received her M.Sc. and Ph.D. degrees in Electrical and Computer Engineering from the University of Minnesota, Minneapolis, in 2002 and 2004, respectively. Since August 2004, she has been with the Department of Electrical and Computer Engineering at the University of Florida, Gainesville, where she is now an Associate Professor. Her current research interests include signal processing, communications theory and networking.

Dr. Yang received the Best Dissertation Award in the Physical Sciences & Engineering from the University of Minnesota in 2005, the Office of Naval Research Young Investigator Award in 2007, and the National Science Foundation Faculty Early Career Development Award in 2009. She is an Associate Editor of the IEEE Transactions on Communications, IEEE Transactions on Wireless Communications and the IEEE Transactions on Intelligent Transportation Systems. She is also a member of the Editorial Board of PHYCOM: Physical Communications. Dr. Yang is presently the vice chair of the IEEE Section of Gainesville, FL, and a co-chair of IEEE Intelligent Transportation Systems Society technical committee on Mobile Communication Networks.