

# Optimal Signal Design for Multi-Parameter Estimation Problems

Hamza Soganci, *Member, IEEE*, Sinan Gezici, *Senior Member, IEEE*, and Orhan Arikan, *Member, IEEE*

**Abstract**—In this study, the optimal stochastic design of multiple parameters is investigated for an array of fixed estimators both in the absence and presence of an average power constraint. Two different performance criteria are considered: the total Bayes risk criterion and the maximum Bayes risk criterion. It is obtained that in the presence of  $K$  parameters and the average power constraint, the optimal stochastic parameter design results in randomization (time sharing) among at most two and  $(K + 1)$  different signals for the total Bayes risk and the maximum Bayes risk criteria, respectively. The average transmitted signal powers corresponding to the optimal parameter design approaches are specified, and the characterization of the optimal approaches is provided in various scenarios. In addition, sufficient conditions are derived to specify when the stochastic parameter design or the deterministic parameter design is optimal. Finally, numerical examples are presented to investigate the theoretical results, and to illustrate performance improvements achieved via the proposed approaches.

**Index Terms**—Parameter estimation, stochastic parameter design, Bayes risk, multi-parameter, minimax.

## I. INTRODUCTION

In many parametric estimation problems, the aim is to design the optimal estimator for an unknown parameter based on a given probability distribution of observations. The common estimators employed in such problems can be categorized into two groups based on the presence of prior information about the parameter to be estimated. If there exists prior information about the parameter, Bayesian estimators, such as the minimum mean-absolute error (MMAE) estimator and the minimum mean-squared error (MMSE) estimator, are commonly used [1]. On the other hand, when there is no prior information about the parameter, the minimum variance unbiased estimator (MVUE) or the maximum likelihood estimator (MLE) can be designed [2]. All these approaches involve the design of an optimal estimator under certain constraints. In a recent study, an alternative formulation is investigated by considering the stochastic design of a parameter when the estimator is fixed, where the aim is to improve the estimation performance by optimally designing the transmitted signal (which can be deterministic or stochastic) for each possible parameter value

[3]. It is shown that the performance of a given estimator can be enhanced by the optimal stochastic parameter design, which involves randomization (time sharing) between at most two different values for the signal transmitted for each parameter.

Randomization (time sharing) among different signal values has been utilized in various frameworks to improve performance of detection and estimation systems [4]-[17]. For example, performance of some detectors can be enhanced by the addition of a randomized noise component to the input (observation) without modifying the detector structure [4]-[10]. Such noise enhancement effects have been studied according to various criteria such as Neyman-Pearson (NP) [4], [5], Bayes [7], minimax [8], and restricted Bayes [9]. As another application of randomization, transmitting randomized signals for each information symbol can reduce the error probability of an average power constrained digital communication system in the presence of non-Gaussian noise [11], [12]. It is shown in [11] that the optimal strategy is to perform randomization (time sharing) among no more than three different transmitted signal values for each information symbol under second and fourth moment constraints. Randomization (time sharing) can be also utilized in jammer systems for improved jamming performance [18]-[20]. In [18], it is proved that a weak jammer employs on-off time sharing to maximize the average probability of error for a receiver operating in the presence of symmetric unimodal noise. On the other hand, for an average power constrained jammer that operates over an arbitrary additive noise channel, the detection probability of an instantaneously and fully adaptive receiver that employs the NP criterion is minimized via randomization between at most two different power levels [20]. In an estimation framework, benefits of randomization are observed in the context of noise enhanced estimation in [17], which proves that performance of some suboptimal estimators can be improved by adding randomized ‘noise’ to the observations before the estimation process.

In some estimation problems, the optimal estimator can be very complicated, and its implementation can be quite costly. In such scenarios, it can be reasonable to employ a suboptimal estimator with a low complexity, and try to employ alternative approaches for improving the performance of that suboptimal estimator. In [3], the optimal stochastic design of a single parameter is proposed in order to optimize the performance of a given (suboptimal) estimator. (Please see Section II of [3] for motivation and examples for stochastic parameter design.) It is shown that the Bayes risk of a given estimator can be reduced by performing randomization between at most two different signal levels for each parameter. In this manuscript, the aim is to propose a framework for the optimal stochastic design of *multiple* parameters. In this way, the approach in [3] for the single parameter case is extended to the multi-parameter

Copyright (c) 2015 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

H. Soganci is with TUBITAK-SAGE, Group of Electronic Systems and Flight Disciplines, 06261, Ankara, Turkey, Email: hso-ganci@ee.bilkent.edu.tr. S. Gezici and O. Arikan are with the Department of Electrical and Electronics Engineering, Bilkent University, Bilkent, Ankara, 06800, Turkey, Tel: +90-312-290-3139, Fax: +90-312-266-4192, Emails: {gezici,oarikan}@ee.bilkent.edu.tr.

This research was supported in part by the Distinguished Young Scientist Award of Turkish Academy of Sciences (TUBA-GEBIP 2013), and in part by the National Young Researchers Career Development Programme (project no. 110E245) of the Scientific and Technological Research Council of Turkey (TUBITAK). Part of this work was presented at IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), May 2014.

scenario in which there exist multiple parameters (each can be a scalar or a vector) and corresponding fixed estimators. That is, the optimal stochastic design of multiple parameters is performed in order to optimize the performance of an array of fixed estimators. It should be emphasized that the difference of the multi-parameter case investigated in this study from the single parameter case investigated in [3] is not only related to the number of parameters. The proposed multi-parameter formulation in this study also takes into account the possible interference among parameter related signals (cf. Fig. 1). In addition, two different performance criteria, the total Bayes risk and the maximum Bayes risk, are considered, and the probability distributions of the transmitted signals are specified for the optimal stochastic parameter design approaches. Furthermore, the average transmitted signal powers corresponding to the optimal parameter design approaches are determined, and the characterization of the optimal approaches is provided for various scenarios. Also, sufficient conditions are derived to specify when the stochastic parameter design or the deterministic parameter design (which involves no randomization) is optimal. Numerical examples are presented to investigate the theoretical results, and to illustrate performance improvements obtained via the proposed approaches.

The main contributions of this study can be summarized as follows:

- The optimal stochastic design of multiple parameters for improving the performance of a set of given estimators is studied for the first time.
- Both the total Bayes risk and the maximum Bayes risk criteria are considered, and it is shown that the optimal solution involves randomization between/among at most 2 and  $(K + 1)$  different signals for the total Bayes risk and the maximum Bayes risk criteria, respectively, under an average power constraint, where  $K$  is the number of parameters.
- For the total Bayes risk criterion, it is shown that when the optimal solution involves randomization between two different signals, the average transmitted power is always equal to the average power limit. Based on this result, a low complexity approach for obtaining the optimal solution is proposed.
- For the maximum Bayes risk criterion, a simple condition is derived in order to specify scenarios in which the optimal solution involves randomization between at most two different signals.
- Optimality conditions are derived to specify cases in which the stochastic parameter design or the deterministic parameter design is optimal.

#### A. Motivation

The main motivation behind the stochastic parameter design is to improve performance of a given (fixed) estimator at the receiver by performing optimal mapping (which can be stochastic in general) of parameter values at the transmitter [3]. This is especially useful when the optimal estimator is costly and a suboptimal estimator is employed at the receiver. In such cases, the stochastic parameter design provides a way of improving the accuracy of parameter estimation. In addition to the arguments provided in Section II of [3] for the

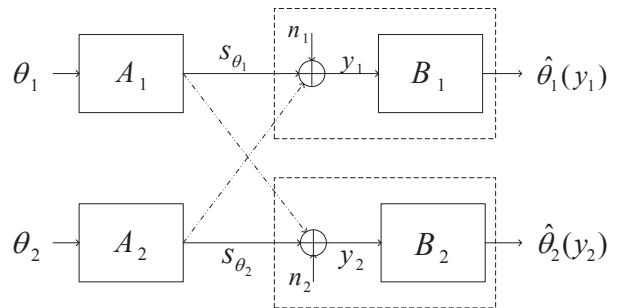


Fig. 1. System model for  $K = 2$ . Devices  $A_1$  and  $A_2$  transmit stochastic signals  $s_{\theta_1}$  and  $s_{\theta_2}$  for each value of parameters  $\theta_1$  and  $\theta_2$ , respectively. Devices  $B_1$  and  $B_2$  estimate  $\theta_1$  and  $\theta_2$  based on the noise and interference corrupted version of  $s_{\theta_1}$  and  $s_{\theta_2}$ , respectively.

stochastic design of a single parameter, additional motivations can also be provided for the multi-parameter case. Estimation of multiple parameters naturally arises in multiuser systems in which multiple devices send parameter related signals to multiple intended devices. For example, a wireless sensor network with multiple users, in which each user (Devices  $A_1$  and  $A_2$  in Fig. 1) aims to send a parameter value (such as temperature or pressure) to a corresponding device (Devices  $B_1$  and  $B_2$  in Fig. 1), can be considered. Since communications occur in the same environment, interference can also be observed at each receiving device, as shown in Fig. 1. In particular, when code division multiple access (CDMA) is employed, each user transmits its parameters via a waveform that depends on a specific spreading code for orthogonalization purposes.<sup>1</sup> However, in practical scenarios, waveforms of different users cannot be perfectly orthogonal (due to effects such as propagation delay) and some non-zero cross-correlations exist, which leads to multiuser interference [21]. Hence, the interference is determined by the cross-correlation properties of the employed spreading sequences in the system (cf. (1)). In addition, transmitters can obtain the knowledge of the probability distributions of the noise via feedback. Then, stochastic parameter design can be performed, and performance of the estimators at the receivers can be optimized.

#### B. Organization

The remainder of the manuscript is organized as follows: In Section II, the problem formulation is introduced and the optimal randomization strategies are obtained. In Section III, some properties of the optimal stochastic parameter design approaches are discussed. Sufficient conditions are derived in Section IV in order to specify when the stochastic parameter design or the deterministic parameter design is optimal. After the numerical examples in Section V, concluding remarks are made in Section VI.

<sup>1</sup>Note that the model in (1) provides an abstraction for all the operations in the system such as quantizer, encoder/decoder, modulator/demodulator, and additive noise channel [3].

## II. STOCHASTIC DESIGN FOR MULTI-PARAMETER ESTIMATION

In this section, we establish a framework for the stochastic design of multiple parameters for a given set of fixed estimators. Consider a parameter estimation scenario in which there exist  $K$  parameters denoted by  $\theta_1, \dots, \theta_K$ , where each parameter resides in  $\mathbb{R}^M$ . Information about parameter  $\theta_i$  is transmitted by device  $A_i$ , which can transmit any signal  $\mathbf{s}_{\theta_i} \in \mathbb{R}^M$  related to  $\theta_i$ , where  $i \in \{1, \dots, K\}$ . The transmitted signal  $\mathbf{s}_{\theta_i}$  is corrupted by both additive noise and the interference from other transmitted signals, and device  $B_i$  tries to estimate the unknown parameter  $\theta_i$  based on the noise and interference corrupted signal. An example system is depicted in Fig. 1 for  $K = 2$ . It should be emphasized that parameter  $\theta_i$  is not necessarily transmitted as it is; instead, device  $A_i$  can transmit any function of  $\theta_i$ , say  $\mathbf{s}_{\theta_i}$ . In addition, function  $\mathbf{s}_{\theta_i}$  can be of any type; it can be a deterministic function of  $\theta_i$ , or it can be a stochastic function. The aim of this study is to find the optimal  $\mathbf{s}_{\theta_i}$ , i.e., the optimal probability distribution of  $\mathbf{s}_{\theta_i}$ , for each  $\theta_i$ .

It is noted that the difference between the single parameter case studied in [3] and the multi-parameter case investigated in this manuscript is not only related to the number of parameters. The proposed multi-parameter formulation in this study also takes into account the possible interference among the parameter related signals, as shown by the dashed cross lines in Fig. 1. Considering  $K$  parameters, the received signal (observation) at device  $B_i$  can be expressed as

$$\mathbf{y}_i = \mathbf{s}_{\theta_i} + \sum_{\substack{j=1 \\ j \neq i}}^K \rho_{ij} \mathbf{s}_{\theta_j} + \mathbf{n}_i \quad (1)$$

for  $i \in \{1, \dots, K\}$ , where  $\rho_{ij}$  is the multiplier that is set according to the interference between the parameter related signals for the  $i^{\text{th}}$  and  $j^{\text{th}}$  parameters<sup>2</sup>, and  $\mathbf{n}_i$  represents the channel noise, which has a probability density function (PDF) denoted by  $p_{\mathbf{n}_i}(\cdot)$ . Each device  $B_i$  tries to estimate  $\theta_i$  based on the corresponding observation  $\mathbf{y}_i$  in (1). It is assumed that the devices employ fixed estimators specified by  $\hat{\theta}_i(\mathbf{y}_i)$  in order to estimate  $\theta_i$ . Let  $\boldsymbol{\theta}$  denote the overall parameter vector, which is defined as  $\boldsymbol{\theta} \triangleq [\theta_1^T \dots \theta_K^T]^T$ . The prior distribution of  $\boldsymbol{\theta}$  is represented by  $w(\boldsymbol{\theta})$ , and the parameter space in which  $\boldsymbol{\theta}$  resides is denoted by  $\Lambda$ . It should be emphasized that  $\mathbf{s}_{\theta_i}$  ( $\mathbf{s}_{\theta_j}$ ) in (1) can be any function of  $\theta_i$  ( $\theta_j$ ).

The aim is to obtain the optimal probability distributions of  $\mathbf{s}_{\theta}$  for each  $\boldsymbol{\theta} \in \Lambda$  in order to minimize a function of the Bayes risk for the given estimators, where  $\mathbf{s}_{\theta} \triangleq [\mathbf{s}_{\theta_1}^T \dots \mathbf{s}_{\theta_K}^T]^T$ . Since the parameters can interfere with each other, the optimization cannot be performed independently for each parameter in general; therefore, a joint optimization should be performed.

### A. Unconstrained Optimization

In this section, the optimal stochastic parameter design

<sup>2</sup>For the example of a CDMA system as in Section I-A, the  $\rho_{ij}$  terms in (1) can be determined by the cross-correlation properties of the employed spreading sequences in the system.

problem is formulated without any constraints as [3]

$$\{p_{\mathbf{s}_{\theta}}^{\text{opt}}, \boldsymbol{\theta} \in \Lambda\} = \arg \min_{\{p_{\mathbf{s}_{\theta}}, \boldsymbol{\theta} \in \Lambda\}} r(\hat{\boldsymbol{\theta}}) \quad (2)$$

where  $\{p_{\mathbf{s}_{\theta}}, \boldsymbol{\theta} \in \Lambda\}$  represents the set of PDFs for  $\mathbf{s}_{\theta}$  for all possible values of parameter  $\boldsymbol{\theta}$ , and  $r(\hat{\boldsymbol{\theta}})$  is the objective function for the overall system. For the single parameter case, the Bayes risk of the estimator was a natural choice for this objective function [3]. On the other hand, it is possible to consider various risk functions for the multi-parameter case. In this section, two different objective functions are considered. The first one is the sum of the Bayes risks of the  $K$  estimators in the system (called the *total Bayes risk*), and the second one is the maximum of the Bayes risks of the estimators (called the *maximum Bayes risk*). For both of these objective functions, the Bayes risk of each estimator should be calculated first. For the two parameter case, the Bayes risk of the first estimator is expressed as

$$r(\hat{\boldsymbol{\theta}}_1) = \int_{\Lambda_1} w(\boldsymbol{\theta}_1) \int p_{\mathbf{s}_{\theta}}(\mathbf{x}) \int C[\hat{\boldsymbol{\theta}}_1(\mathbf{y}_1), \boldsymbol{\theta}_1] \times p_{\mathbf{n}_1}(\mathbf{y}_1 - \mathbf{x}_1 - \rho_{12}\mathbf{x}_2) d\mathbf{y}_1 d\mathbf{x} d\boldsymbol{\theta}_1 \quad (3)$$

where  $C[\hat{\boldsymbol{\theta}}_1(\mathbf{y}_1), \boldsymbol{\theta}_1]$  denotes the cost of estimating  $\boldsymbol{\theta}_1$  as  $\hat{\boldsymbol{\theta}}_1(\mathbf{y}_1)$  [2], and  $\mathbf{x} = [\mathbf{x}_1^T \mathbf{x}_2^T]^T$ . (The Bayes risk of the second estimator can be expressed in a similar fashion.)

Defining an auxiliary function  $g_{\theta_1}(\mathbf{x})$  for the first estimator as

$$g_{\theta_1}(\mathbf{x}) \triangleq \int C[\hat{\boldsymbol{\theta}}_1(\mathbf{y}_1), \boldsymbol{\theta}_1] p_{\mathbf{n}_1}(\mathbf{y}_1 - \mathbf{x}_1 - \rho_{12}\mathbf{x}_2) d\mathbf{y}_1 \quad (4)$$

and a similar function for the second estimator, the total Bayes risk can be expressed as

$$r(\hat{\boldsymbol{\theta}}) = \int_{\Lambda} w(\boldsymbol{\theta}) \int p_{\mathbf{s}_{\theta}}(\mathbf{x}) (g_{\theta_1}(\mathbf{x}) + g_{\theta_2}(\mathbf{x})) d\mathbf{x} d\boldsymbol{\theta} = \int_{\Lambda} w(\boldsymbol{\theta}) E\{\tilde{g}_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\} d\boldsymbol{\theta} \quad (5)$$

with  $\hat{\boldsymbol{\theta}} = [\hat{\boldsymbol{\theta}}_1^T \hat{\boldsymbol{\theta}}_2^T]^T$ ,  $\boldsymbol{\theta} = [\boldsymbol{\theta}_1^T \boldsymbol{\theta}_2^T]^T$ ,  $\mathbf{s}_{\boldsymbol{\theta}} = [\mathbf{s}_{\boldsymbol{\theta}_1}^T \mathbf{s}_{\boldsymbol{\theta}_2}^T]^T$ , and

$$\tilde{g}_{\boldsymbol{\theta}}(\mathbf{x}) = g_{\theta_1}(\mathbf{x}) + g_{\theta_2}(\mathbf{x}). \quad (6)$$

For the  $K$  parameter case, similar expressions can be obtained by updating (3) and (4) in order to include the interference due to the other parameters as well. In that case, (5) still has the same form with the updated definition of  $\tilde{g}_{\boldsymbol{\theta}}$  which is given by  $\tilde{g}_{\boldsymbol{\theta}}(\mathbf{x}) = \sum_{i=1}^K g_{\theta_i}(\mathbf{x})$ .

Similarly to [3], it can be shown that the solution of the optimization problem in (2) for the total Bayes risk in (5) can be obtained as

$$p_{\mathbf{s}_{\boldsymbol{\theta}}}^{\text{opt}}(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{s}_{\boldsymbol{\theta}}^{\text{unc}}), \quad \mathbf{s}_{\boldsymbol{\theta}}^{\text{unc}} = \arg \min_{\mathbf{x}} \tilde{g}_{\boldsymbol{\theta}}(\mathbf{x}) \quad (7)$$

for all  $\boldsymbol{\theta} \in \Lambda$ , where  $\delta$  denotes the Dirac delta function. Hence, the deterministic parameter design is optimal and there is no need for stochastic modeling in this scenario. Also it can be observed from (7) that the solution is independent of the prior distribution  $w(\boldsymbol{\theta})$  as the optimal solution is obtained for each  $\boldsymbol{\theta}$  separately.

When the maximum Bayes risk criterion is considered, the objective function in (5) can be updated as

$$\begin{aligned} r(\hat{\theta}) &= \int_{\Lambda} w(\theta) \max_{i \in \{1, \dots, K\}} \left( \int p_{\mathbf{s}_\theta}(\mathbf{x}) g_{\theta_i}(\mathbf{x}) d\mathbf{x} \right) d\theta \\ &= \int_{\Lambda} w(\theta) \max_{i \in \{1, \dots, K\}} (E\{g_{\theta_i}(\mathbf{s}_\theta)\}) d\theta. \end{aligned} \quad (8)$$

Based on similar arguments to those employed above for the total Bayes risk criterion, it can be observed that the solution is independent of the prior distribution  $w(\theta)$  and the optimal solution can be obtained for each  $\theta$  separately. Hence, the optimization problem for the maximum Bayes risk criterion can be formulated as follows:

$$p_{\mathbf{s}_\theta}^{\text{opt}} = \arg \min_{p_{\mathbf{s}_\theta}} \max_{i \in \{1, \dots, K\}} E\{g_{\theta_i}(\mathbf{s}_\theta)\}. \quad (9)$$

The study in [8] considers an optimization problem that is in the same form as (9) (please see (13) in [8]). Hence, Proposition 1 in [8] also applies to the problem in (9), which implies that the optimal solution corresponds to a discrete random variable with at most  $K$  point masses for each  $\theta$  under some mild and practical conditions. Based on this result, the optimal stochastic parameter design problem for the maximum Bayes risk criterion can be expressed as

$$\begin{aligned} \min_{\{\lambda_{\theta,j}, \mathbf{s}_{\theta,j}\}_{j=1}^K} \max_{i \in \{1, \dots, K\}} \sum_{j=1}^K \lambda_{\theta,j} g_{\theta_i}(\mathbf{s}_{\theta,j}) \\ \text{subject to } \sum_{j=1}^K \lambda_{\theta,j} = 1 \\ \lambda_{\theta,j} \in [0, 1], \forall j \in \{1, \dots, K\} \end{aligned} \quad (10)$$

for  $\theta \in \Lambda$ , where  $\mathbf{s}_\theta$  takes the value of  $\mathbf{s}_{\theta,j}$  with probability  $\lambda_{\theta,j}$  for  $j = 1, \dots, K$ . Compared to (9), the formulation in (10) provides a significant reduction in computational complexity as it requires optimization over a finite number of variables instead of over all possible PDFs. Since generic cost functions and noise distributions are considered in the theoretical analysis, function  $g_{\theta_i}$  in (4) is generic as well; hence, the optimization problem in (15) can be nonconvex in general.

### B. Constrained Optimization

In this section, an average power constraint is considered [3]:

$$E\{\|\mathbf{s}_\theta\|^2\} \leq A_\theta \quad (11)$$

for  $\theta \in \Lambda$ , where  $\|\mathbf{s}_\theta\|$  is the Euclidean norm of vector  $\mathbf{s}_\theta$ , and  $A_\theta$  represents the average power limit for  $\theta$ . In general, constraint  $A_\theta$  can be a function of  $\theta$  as well. From (5) and (11), the optimal stochastic parameter design problem for the total Bayes risk criterion can be expressed as

$$\begin{aligned} \min_{\{p_{\mathbf{s}_\theta}, \theta \in \Lambda\}} \int_{\Lambda} w(\theta) E\{\tilde{g}_\theta(\mathbf{s}_\theta)\} d\theta \\ \text{subject to } E\{\|\mathbf{s}_\theta\|^2\} \leq A_\theta, \forall \theta \in \Lambda \end{aligned} \quad (12)$$

where  $\tilde{g}_\theta(\cdot)$  is as defined in (6). Due to the structure of the objective function and the constraint, the constrained

optimization problem in (12) can be solved individually for each  $\theta$  as

$$\min_{p_{\mathbf{s}_\theta}} E\{\tilde{g}_\theta(\mathbf{s}_\theta)\} \quad \text{subject to } E\{\|\mathbf{s}_\theta\|^2\} \leq A_\theta \quad (13)$$

for  $\theta \in \Lambda$ . Therefore, the solution does not depend on the prior distribution  $w(\theta)$ .

When the maximum Bayes risk criterion is considered, it can be obtained from (8) and (11) that the problem becomes

$$\min_{p_{\mathbf{s}_\theta}} \max_{i \in \{1, \dots, K\}} E\{g_{\theta_i}(\mathbf{s}_\theta)\} \quad \text{subject to } E\{\|\mathbf{s}_\theta\|^2\} \leq A_\theta \quad (14)$$

for  $\theta \in \Lambda$ . Similar optimization problems in the form of (13) and (14) have been investigated in the literature [3], [4], [5], [11]. The problem in (13) has the same form as the one considered in [3]. Therefore, the statistical behavior of the optimal solution is the same; that is, the optimal solution can be achieved by a randomization (time sharing) between at most two different values of  $\mathbf{s}_\theta$  for each  $\theta$ , as stated in Proposition 1 in [3]. Then, the optimal solution can be obtained based on a similar approach to that in [3]. Namely, the optimal stochastic parameter design problem for the total Bayes risk criterion can be expressed as

$$\begin{aligned} \min_{\{\lambda_{\theta,j}, \mathbf{s}_{\theta,j}\}_{j=1}^2} \sum_{j=1}^2 \lambda_{\theta,j} \tilde{g}_\theta(\mathbf{s}_{\theta,j}) \\ \text{subject to } \sum_{j=1}^2 \lambda_{\theta,j} \|\mathbf{s}_{\theta,j}\|^2 \leq A_\theta, \sum_{j=1}^2 \lambda_{\theta,j} = 1, \\ \lambda_{\theta,j} \in [0, 1], j \in \{1, 2\} \end{aligned} \quad (15)$$

for  $\theta \in \Lambda$ . That is, the optimal parameter design involves the use of at most two different signal values for each parameter according to the total Bayes risk criterion. On the other hand, the optimization problem in (14) has a different form than that in [3]. Based on arguments similar to those in [22], the following result can be obtained.

**Proposition 1:** *Suppose that functions  $g_{\theta_i}$  for  $i \in \{1, \dots, K\}$  are continuous, and each component of  $\mathbf{s}_\theta$  resides in a finite closed interval. Then, the optimal solution of (14) can be characterized by the following probability density:*

$$p_{\mathbf{s}_\theta}^{\text{opt}}(\mathbf{x}) = \sum_{j=1}^{K+1} \lambda_{\theta,j} \delta(\mathbf{x} - \mathbf{s}_{\theta,j}) \quad (16)$$

where  $\lambda_{\theta,j} \geq 0$  and  $\sum_{j=1}^{K+1} \lambda_{\theta,j} = 1$ .

Proposition 1 states that the optimal solution can be achieved by a randomization (time sharing) among at most  $K+1$  different values of  $\mathbf{s}_\theta$  for each  $\theta$ . Based on this result, the optimal stochastic parameter design problem for the maximum Bayes risk criterion can be expressed as

$$\begin{aligned} \min_{\{\lambda_{\theta,j}, \mathbf{s}_{\theta,j}\}_{j=1}^{K+1}} \max_{i \in \{1, \dots, K\}} \sum_{j=1}^{K+1} \lambda_{\theta,j} g_{\theta_i}(\mathbf{s}_{\theta,j}) \\ \text{subject to } \sum_{j=1}^{K+1} \lambda_{\theta,j} \|\mathbf{s}_{\theta,j}\|^2 \leq A_\theta, \sum_{j=1}^{K+1} \lambda_{\theta,j} = 1, \\ \lambda_{\theta,j} \in [0, 1], j \in \{1, \dots, K+1\} \end{aligned} \quad (17)$$

for  $\theta \in \Lambda$ .

From (15) and (17), it is concluded that randomization (time sharing) of transmitted signal values may offer improvements in the presence of an average power constraint for both the total Bayes risk and the maximum Bayes risk criteria. In addition, the optimization problems in (15) and (17) can be nonconvex in general since generic cost functions and noise distributions are considered in the theoretical analysis.

### III. CHARACTERIZATION OF OPTIMAL STOCHASTIC PARAMETER DESIGN IN THE PRESENCE OF AVERAGE POWER CONSTRAINT

In this section, some properties of the optimal stochastic parameter design approaches in the presence of average power constraints are discussed. Namely, the average transmitted signal powers corresponding to the optimal parameter design approaches are investigated, and the characterization of the optimal approaches is provided in various scenarios.

For the total Bayes risk criterion, the following two results are obtained when the *stochastic* parameter design is the solution of (15) (equivalently, (13)); that is, when the optimal solution involves randomization between two different signal values.<sup>3</sup>

**Lemma 1:** *Assume that the solution of (15) involves randomization between two different signals. Then, (i) one of the signals has a power below the average power limit, and the other signal has a power above the average power limit; (ii) the signal with the higher (lower) power has a lower (higher) risk than the other signal.*

**Proof:** Both results are proved via contradiction. For part (i), first assume that the powers of both signals are smaller than or equal to the average power limit. Then, the solution cannot be a randomization between these two signals since employing the signal with the lower risk (i.e., lower  $\tilde{g}_\theta$ ) exclusively achieves a lower total Bayes risk (see (15)) than performing randomization between these signals. Second, assume that either the powers of both signals are larger than the average power limit, or the power of one signal is equal to and that of the other is larger than the average power limit. In this scenario, the average power constraint in (15) is violated; hence, this cannot be a valid scenario. Therefore, it is concluded that if randomization between two different signals is the solution of (13), then one of the signals must have a power below the average power limit, and the other signal must have a power above the average power limit. For part (ii), if the signal with the lower power has a risk which is smaller than or equal to the risk of the other signal, then there is no need for randomization. In that case, employing this signal exclusively yields a lower risk; hence, randomization between these signals cannot be optimal. Therefore, if randomization between two signals is the solution of (13), then the signal with the higher (lower) power must have a lower (higher) risk than the other signal.  $\square$

Based on Lemma 1, the following result is obtained.

**Proposition 2:** *If the solution of (15) (equivalently, (13)) involves randomization between two different signals; that is,*

<sup>3</sup>In this study, the statement “the optimal solution involves randomization between two different signal values” is used to mean that there is no deterministic solution that achieves the same performance as the optimal stochastic solution.

*if stochastic parameter design is optimal, then the average signal power is equal to the average power limit; i.e., the solution operates at the average power limit.*

**Proof:** In order to prove the claim in the proposition, suppose that  $\{\lambda_{\theta,j}, \mathbf{s}_{\theta,j}\}_{j=1}^2$  is an optimal solution and utilizes a power strictly lower than the average power limit; i.e.,  $\lambda_{\theta,1}\|\mathbf{s}_{\theta,1}\|^2 + (1 - \lambda_{\theta,1})\|\mathbf{s}_{\theta,2}\|^2 < A_\theta$ . Without loss of generality, assume that  $\|\mathbf{s}_{\theta,1}\|^2 > A_\theta$  and  $\|\mathbf{s}_{\theta,2}\|^2 < A_\theta$  as a result of part (i) of Lemma 1. According to part (ii) of Lemma 1,  $\tilde{g}_\theta(\mathbf{s}_{\theta,1}) < \tilde{g}_\theta(\mathbf{s}_{\theta,2})$  is satisfied. Next, consider another solution  $\{\lambda'_{\theta,j}, \mathbf{s}_{\theta,j}\}_{j=1}^2$  with  $\lambda'_{\theta,1} = (A_\theta - \|\mathbf{s}_{\theta,2}\|^2)/(\|\mathbf{s}_{\theta,1}\|^2 - \|\mathbf{s}_{\theta,2}\|^2)$ . Note that the average power for this solution is equal to the average power limit; that is,  $\lambda'_{\theta,1}\|\mathbf{s}_{\theta,1}\|^2 + (1 - \lambda'_{\theta,1})\|\mathbf{s}_{\theta,2}\|^2 = A_\theta$ . In addition, it can be shown that  $\lambda'_{\theta,1} > \lambda_{\theta,1}$  as  $\|\mathbf{s}_{\theta,1}\|^2 > A_\theta$ ,  $\|\mathbf{s}_{\theta,2}\|^2 < A_\theta$ , and the average power of solution  $\{\lambda'_{\theta,j}, \mathbf{s}_{\theta,j}\}_{j=1}^2$  is larger than that of solution  $\{\lambda_{\theta,j}, \mathbf{s}_{\theta,j}\}_{j=1}^2$ . Since  $\tilde{g}_\theta(\mathbf{s}_{\theta,1}) < \tilde{g}_\theta(\mathbf{s}_{\theta,2})$  due to part (ii) of Lemma 1 and  $\lambda'_{\theta,1} > \lambda_{\theta,1}$ , it can be shown that solution  $\{\lambda'_{\theta,j}, \mathbf{s}_{\theta,j}\}_{j=1}^2$  achieves a lower total Bayes risk than solution  $\{\lambda_{\theta,j}, \mathbf{s}_{\theta,j}\}_{j=1}^2$ ; that is,

$$\lambda'_{\theta,1}\tilde{g}_\theta(\mathbf{s}_{\theta,1}) + (1 - \lambda'_{\theta,1})\tilde{g}_\theta(\mathbf{s}_{\theta,2}) < \lambda_{\theta,1}\tilde{g}_\theta(\mathbf{s}_{\theta,1}) + (1 - \lambda_{\theta,1})\tilde{g}_\theta(\mathbf{s}_{\theta,2}). \quad (18)$$

Based on (18), it is concluded that solution  $\{\lambda_{\theta,j}, \mathbf{s}_{\theta,j}\}_{j=1}^2$  cannot be optimal, which results in a contradiction. Hence, it is concluded that a solution with an average power lower than the average power limit cannot be optimal for the scenario in the proposition. That is, the solution of (15) operates at the average power limit when the stochastic parameter design is optimal.  $\square$

From Proposition 2, the solution of (15) can be obtained as stated in the following proposition.

**Proposition 3:** *The solution of (15) corresponds to either deterministic parameter design or stochastic parameter design, which can be obtained as follows:*

- **Deterministic Parameter Design:** *Transmit  $\mathbf{s}_\theta^{\text{det}}$  exclusively for  $\theta \in \Lambda$ , where*

$$\mathbf{s}_\theta^{\text{det}} = \arg \min_{\|\mathbf{s}_\theta\|^2 \leq A_\theta} \tilde{g}_\theta(\mathbf{s}_\theta). \quad (19)$$

- **Stochastic Parameter Design:** *Perform time sharing (randomization) between  $\mathbf{s}_{\theta,1}^{\text{opt}}$  and  $\mathbf{s}_{\theta,2}^{\text{opt}}$  with time sharing factors  $(A_\theta - \|\mathbf{s}_{\theta,2}^{\text{opt}}\|^2)/(\|\mathbf{s}_{\theta,1}^{\text{opt}}\|^2 - \|\mathbf{s}_{\theta,2}^{\text{opt}}\|^2)$  and  $(\|\mathbf{s}_{\theta,1}^{\text{opt}}\|^2 - A_\theta)/(\|\mathbf{s}_{\theta,1}^{\text{opt}}\|^2 - \|\mathbf{s}_{\theta,2}^{\text{opt}}\|^2)$ , respectively, where*

$$(\mathbf{s}_{\theta,1}^{\text{opt}}, \mathbf{s}_{\theta,2}^{\text{opt}}) = \arg \min_{\substack{\|\mathbf{s}_{\theta,1}\|^2 > A_\theta \\ \|\mathbf{s}_{\theta,2}\|^2 < A_\theta}} \frac{A_\theta - \|\mathbf{s}_{\theta,2}\|^2}{\|\mathbf{s}_{\theta,1}\|^2 - \|\mathbf{s}_{\theta,2}\|^2} \tilde{g}_\theta(\mathbf{s}_{\theta,1}) + \frac{\|\mathbf{s}_{\theta,1}\|^2 - A_\theta}{\|\mathbf{s}_{\theta,1}\|^2 - \|\mathbf{s}_{\theta,2}\|^2} \tilde{g}_\theta(\mathbf{s}_{\theta,2}) \quad (20)$$

for  $\theta \in \Lambda$ .

*The solution of (15) is the one ((19) or (20)) that results in the lower total Bayes risk.*

**Proof:** There exist two possible scenarios for the solution of (15). If no randomization is employed, the optimal solution can be obtained as in (19), which is called the deterministic parameter design. On the other hand, randomization between

two signals can be performed. As stated in Proposition 2, the average signal power must be equal to the average power limit in this scenario; that is,  $\lambda_{\theta,1}\|\mathbf{s}_{\theta,1}\|^2 + (1 - \lambda_{\theta,1})\|\mathbf{s}_{\theta,2}\|^2 = A_{\theta}$ . Therefore, the time sharing (randomization) factors can be calculated as  $\lambda_{\theta,1} = (A_{\theta} - \|\mathbf{s}_{\theta,2}\|^2)/(\|\mathbf{s}_{\theta,1}\|^2 - \|\mathbf{s}_{\theta,2}\|^2)$  and  $\lambda_{\theta,2} = 1 - \lambda_{\theta,1}$ . In addition, from part (i) of Lemma 1, one signal has a power higher than the average power limit and the other signal has a power lower than the average power limit. Hence, the optimization problem in (15) can be simplified as the one in (20). Finally, it is observed that the solution that achieves the lower risk in (19) and (20) becomes the solution of (15).  $\square$

Proposition 3 provides a simple approach for solving (15). Namely, the problems in (19) and (20) are solved, and the one that achieves the lower total Bayes risk becomes the solution of (15).

For the maximum Bayes risk criterion, the solution of (17) (equivalently, (14)) can be characterized as a special form under certain conditions. To that aim, the following lemma is presented first.

**Lemma 2:** Consider a set of functions,  $f_i$ , for  $i \in \{1, 2, \dots, K\}$ . If minimum value of a certain function, say  $f_m$ , is strictly higher than the values of the other functions at the same point, then this point is the solution of the minimax problem; that is,

$$\min_x \max_i f_i(x) = \min_x f_m(x). \quad (21)$$

**Proof:** Let  $x'$  denote the minimizer of  $f_m(x)$  and  $f_m(x') > f_i(x')$ ,  $\forall i \in \{1, 2, \dots, K\} \setminus \{m\}$ , as stated in the lemma. Suppose that  $x'$  is not the solution of the minimax problem, and consider another point  $x^*$  which yields a lower value for the minimax problem; that is,  $\max_i f_i(x^*) < f_m(x')$ . By definition,  $\max_i f_i(x^*) \geq f_m(x^*)$ . Combining the last two inequalities, it is obtained that  $f_m(x^*) < f_m(x')$ , which contradicts the fact that the value of  $f_m$  is strictly higher than the values of the other functions at  $x'$ . Hence, it is concluded that no other point,  $x^*$ , can yield a lower value for the minimax problem than  $x'$ .  $\square$

Based on Lemma 2, the following result is obtained about the solution of the optimal parameter design problem according to the maximum Bayes risk criterion.

**Proposition 4:** Consider the probability distribution of  $\mathbf{s}_{\theta}$  that minimizes the risk of the  $m^{\text{th}}$  estimator under the average power constraint, where  $m \in \{1, \dots, K\}$ . For that probability distribution, if the risk of the  $m^{\text{th}}$  estimator is strictly higher than the risks of the other estimators, then this distribution is the optimal solution of the minimax problem in (14) (equivalently, (17)) and it involves randomization between at most two signals.

**Proof:** Consider the minimax problem in (17). Let the minimum risk of the  $m^{\text{th}}$  estimator be strictly higher than the risks of the other estimators for the distribution of  $\mathbf{s}_{\theta}$  that minimizes the risk of the  $m^{\text{th}}$  estimator under the average power constraint; that is,

$$\begin{aligned} \min_{\{\lambda_{\theta,j}, \mathbf{s}_{\theta,j}\}} \sum_j \lambda_{\theta,j} g_{\theta_m}(\mathbf{s}_{\theta,j}) &\triangleq \sum_j \lambda'_{\theta,j} g_{\theta_m}(\mathbf{s}'_{\theta,j}) \\ &> \sum_j \lambda'_{\theta,j} g_{\theta_i}(\mathbf{s}'_{\theta,j}) \end{aligned} \quad (22)$$

for  $i \in \{1, \dots, K\} \setminus \{m\}$ , where  $\{\lambda'_{\theta,j}, \mathbf{s}'_{\theta,j}\}$  denotes the probability distribution of  $\mathbf{s}_{\theta}$  that minimizes the risk of the  $m^{\text{th}}$  estimator. In this scenario, Lemma 2 implies that the optimal solution for the  $m^{\text{th}}$  estimator is the solution of the minimax problem as well. Since the optimal solution for a single estimator corresponds to randomization between at most two signals (consider (14) and (16) as if  $K = 1$ ), the solution of the minimax problem in (17) is obtained via randomization between at most two signals under the conditions in the proposition.  $\square$

When the number of parameters is large, it can be difficult to solve the optimization problem in (17) since the dimension of the problem is high in that case. Proposition 4 offers a relatively simple test based on the solution of several low dimensional optimization problems before trying to solve this high dimensional optimization problem. If the conditions stated in the proposition are satisfied then there is no need for solving the high dimensional optimization problem.

#### IV. OPTIMALITY CONDITIONS

In this section, various conditions are derived in order to specify when the stochastic parameter design or the deterministic parameter design is optimal. In order to investigate such optimality conditions, the objective function to be considered should be identified first. In this study, two different objective functions, the total Bayes risk and the maximum Bayes risk, are considered, and the optimality conditions differ for these functions. For the total Bayes risk, the problem can be simplified to minimizing the expectation of a single function,  $\tilde{g}_{\theta}$ , as given in (13). As it was stated in Section II-B, this problem has the same form as the one studied in [3]. Therefore, the optimality conditions proposed in [3] are valid for the total Bayes risk criterion in this study as well. However, for the maximum Bayes risk criterion, the problem has a different form as given in (14); hence, the optimality conditions are different in this scenario. In this section, the optimality conditions are investigated for the maximum Bayes risk criterion.

The optimal parameter design problem presented in (14) does not necessarily yield a stochastic solution in all cases. In certain scenarios, the deterministic design is the optimal solution and in such cases the problem in (14) can be reformulated as

$$\min_{\mathbf{s}_{\theta}} \max_{i \in \{1, \dots, K\}} g_{\theta_i}(\mathbf{s}_{\theta}) \quad \text{subject to} \quad \|\mathbf{s}_{\theta}\|^2 \leq A_{\theta} \quad (23)$$

where  $\mathbf{s}_{\theta}$  is modeled as a deterministic quantity for each  $\theta$ . Let  $\mathbf{s}_{\theta}^{\text{det}}$  represent the minimizer of the optimization problem in (23). Then, the minimum Bayes risk achieved by the optimal deterministic parameter design is expressed as  $r_{\text{det}}(\theta) = \int_{\Lambda} w(\theta) \max_{i \in \{1, \dots, K\}} g_{\theta_i}(\mathbf{s}_{\theta}^{\text{det}}) d\theta$  (c.f. (8)). On the other hand, the minimum Bayes risk achieved by the optimal parameter design is denoted by  $r_{\text{sto}}(\hat{\theta}) = \int_{\Lambda} w(\theta) \int \max_{i \in \{1, \dots, K\}} g_{\theta_i}(\mathbf{x}) p_{\mathbf{s}_{\theta}}^{\text{opt}}(\mathbf{x}) d\mathbf{x} d\theta$ , where  $p_{\mathbf{s}_{\theta}}^{\text{opt}}$  is the optimal solution of (14) for a given  $\theta$ . If the stochastic parameter design is the optimal solution of the problem in (14), then  $r_{\text{sto}}(\hat{\theta})$  is strictly smaller than  $r_{\text{det}}(\hat{\theta})$ . Otherwise, it is concluded that the deterministic parameter design is the optimal solution and the stochastic design does not provide

any improvements; that is,  $r_{\text{sto}}(\hat{\theta}) = r_{\text{det}}(\hat{\theta})$ . In the following proposition, sufficient conditions presented for the second case.

**Proposition 5:** *For the maximum Bayes risk criterion, the stochastic parameter design cannot provide any improvements over the deterministic parameter design if at least one of the following conditions is satisfied for each  $\theta$ :*

- The solution of the unconstrained problem (see (9) or (10)) is deterministic (denoted by  $\mathbf{s}_{\theta}^{\text{unc}}$ ) and satisfies the power constraint; i.e.,  $\|\mathbf{s}_{\theta}^{\text{unc}}\|^2 \leq A_{\theta}$ .
- $g_{\theta_i}$  is a convex function for  $i \in \{1, \dots, K\}$ .

**Proof:** The first part of the proof can be obtained similarly to that of Proposition 2 in [3]. If the first condition in the proposition is satisfied, i.e., if the unconstrained problem has a deterministic solution and  $\|\mathbf{s}_{\theta}^{\text{unc}}\|^2 \leq A_{\theta}$ , then the solution of (23) is the same as that of the unconstrained problem in Section II-A; that is,  $\mathbf{s}_{\theta}^{\text{det}} = \mathbf{s}_{\theta}^{\text{unc}}$ . Therefore, the solution of the optimal stochastic parameter design problem in (14) is expressed as  $p_{\mathbf{s}_{\theta}}^{\text{opt}}(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{s}_{\theta}^{\text{unc}})$ . Hence, the deterministic parameter design is optimal in this case, and the stochastic parameter design cannot provide any improvements.

For the second condition in the proposition, it is note that, for any  $\mathbf{s}_{\theta}$ ,  $\mathbb{E}\{\|\mathbf{s}_{\theta}\|^2\} \geq \|\mathbb{E}\{\mathbf{s}_{\theta}\}\|^2$  holds due to Jensen's inequality as norm is a convex function. Therefore,  $\mathbb{E}\{\|\mathbf{s}_{\theta}\|^2\} \leq A_{\theta}$  in (14) implies that  $\|\mathbb{E}\{\mathbf{s}_{\theta}\}\|^2 \leq A_{\theta}$  must hold for any feasible PDF of  $\mathbf{s}_{\theta}$ . Let  $\mathbb{E}\{\mathbf{s}_{\theta}\}$  be denoted by  $\check{\mathbf{s}}_{\theta}$ ; that is,  $\check{\mathbf{s}}_{\theta} \triangleq \mathbb{E}\{\mathbf{s}_{\theta}\}$ . Since the minimizer of (23),  $\mathbf{s}_{\theta}^{\text{det}}$ , achieves the minimum value of  $\max_{i \in \{1, \dots, K\}} g_{\theta_i}(\mathbf{s}_{\theta})$  among all  $\mathbf{s}_{\theta}$  that satisfy  $\|\mathbf{s}_{\theta}\|^2 \leq A_{\theta}$ , the inequality  $\|\mathbb{E}\{\mathbf{s}_{\theta}\}\|^2 = \|\check{\mathbf{s}}_{\theta}\|^2 \leq A_{\theta}$  implies that

$$\begin{aligned} \max_{i \in \{1, \dots, K\}} g_{\theta_i}(\mathbb{E}\{\mathbf{s}_{\theta}\}) &= \max_{i \in \{1, \dots, K\}} g_{\theta_i}(\check{\mathbf{s}}_{\theta}) \\ &\geq \max_{i \in \{1, \dots, K\}} g_{\theta_i}(\mathbf{s}_{\theta}^{\text{det}}) \end{aligned} \quad (24)$$

holds. If  $g_{\theta_i}$ 's are convex functions, then

$$\begin{aligned} \max_{i \in \{1, \dots, K\}} \mathbb{E}\{g_{\theta_i}(\mathbf{s}_{\theta})\} &\geq \max_{i \in \{1, \dots, K\}} g_{\theta_i}(\mathbb{E}\{\mathbf{s}_{\theta}\}) \\ &\geq \max_{i \in \{1, \dots, K\}} g_{\theta_i}(\mathbf{s}_{\theta}^{\text{det}}) \end{aligned} \quad (25)$$

is obtained from Jensen's inequality and from (24). Therefore, when  $g_{\theta_i}$ 's are convex,  $\max_{i \in \{1, \dots, K\}} \mathbb{E}\{g_{\theta_i}(\mathbf{s}_{\theta})\}$  is never smaller than the minimum value of (23),  $\max_{i \in \{1, \dots, K\}} g_{\theta_i}(\mathbf{s}_{\theta}^{\text{det}})$ , for any PDF of  $\mathbf{s}_{\theta}$  that satisfies the average power constraint. For this reason, the minimum value of (14) cannot be smaller than  $\max_{i \in \{1, \dots, K\}} g_{\theta_i}(\mathbf{s}_{\theta}^{\text{opt}})$ , which means that it is always equal to  $\max_{i \in \{1, \dots, K\}} g_{\theta_i}(\mathbf{s}_{\theta}^{\text{opt}})$  as (14) covers (23) as a special case.

Overall, if at least one of the conditions in the proposition is satisfied for all  $\theta$ , the deterministic and stochastic parameter design approaches achieve the same minimum values for all parameters; that is,  $\max_{i \in \{1, \dots, K\}} g_{\theta_i}(\mathbf{s}_{\theta}^{\text{det}}) = \int \max_{i \in \{1, \dots, K\}} g_{\theta_i}(\mathbf{x}) p_{\mathbf{s}_{\theta}}^{\text{opt}}(\mathbf{x}) d\mathbf{x}$ ,  $\forall \theta$ .

Hence,  $r_{\text{det}}(\hat{\theta}) = \int_{\Lambda} w(\theta) \max_{i \in \{1, \dots, K\}} g_{\theta_i}(\mathbf{s}_{\theta}^{\text{det}}) d\theta$  and

$r_{\text{sto}}(\hat{\theta}) = \int_{\Lambda} w(\theta) \int \max_{i \in \{1, \dots, K\}} g_{\theta_i}(\mathbf{x}) p_{\mathbf{s}_{\theta}}^{\text{opt}}(\mathbf{x}) d\mathbf{x} d\theta$  are equal.  $\square$

For an example of Proposition 5, consider a scenario in which two scalar parameters  $\theta_1$  and  $\theta_2$  are to be estimated in the presence of zero-mean additive noise  $\mathbf{n}$ . The average power constraint is in the form of  $\mathbb{E}\{\|\mathbf{s}_{\theta}\|^2\} \leq A_{\theta}$  for all  $\theta$ , and the estimator is specified by  $\hat{\theta}(\mathbf{y}) = \mathbf{y}$ . Also, the cost function is modeled as  $C[\hat{\theta}(\mathbf{y}), \theta] = (\hat{\theta}_1(y_1) - \theta_1)^2 + (\hat{\theta}_2(y_2) - \theta_2)^2$ . In this case,  $g_{\theta_1}$  in (4) can be calculated as

$$\begin{aligned} g_{\theta_1}(x_1) &= \int_{-\infty}^{\infty} (y_1 - \theta_1)^2 p_{\mathbf{n}_1}(y_1 - x_1 - \rho_{12}x_2) dy_1 \\ &= \int_{-\infty}^{\infty} (y + x_1 + \rho_{12}x_2 - \theta_1)^2 p_{\mathbf{n}_1}(y) dy \\ &= (x_1 + \rho_{12}x_2 - \theta_1)^2 + \text{Var}\{\mathbf{n}_1\} \end{aligned} \quad (26)$$

where  $\text{Var}\{\mathbf{n}_1\}$  is the variance of the noise component for the first parameter. From (26), it is observed that  $g_{\theta_1}$  is a convex function for any value of  $\theta$ . Similarly it is possible to show that  $g_{\theta_2}$  is also a convex function for any  $\theta$ . Therefore, the second condition in Proposition 5 is satisfied for all  $\theta$ , which implies that the performance of the deterministic parameter design cannot be improved via the stochastic approach in this scenario.

In the following proposition, a modified version of Proposition 3 in [3] is obtained in order to present sufficient conditions that specify scenarios in which the stochastic parameter design provides improvements over the deterministic one.

**Proposition 6:** *For the maximum Bayes risk criterion, the stochastic parameter design achieves a lower Bayes risk than the deterministic parameter design if there exists  $\theta \in \Lambda$  for which all  $g_{\theta_i}(\mathbf{x})$ 's are second-order continuously differentiable around  $\mathbf{s}_{\theta}^{\text{det}}$ , and a real vector  $\mathbf{z}$  and a positive number  $k$  can be found such that*

$$\begin{aligned} (\mathbf{z}^T \mathbf{s}_{\theta}^{\text{det}}) (\mathbf{z}^T \nabla g_{\theta_i}(\mathbf{x})|_{\mathbf{x}=\mathbf{s}_{\theta}^{\text{det}}}) &< 0 \quad \text{and} \\ \|\mathbf{z}\|^2 &< (\mathbf{z}^T \mathbf{s}_{\theta}^{\text{det}}) (\mathbf{z}^T \mathbf{H}_{\theta_i} \mathbf{z} - \gamma_i/k) / (\mathbf{z}^T \nabla g_{\theta_i}(\mathbf{x})|_{\mathbf{x}=\mathbf{s}_{\theta}^{\text{det}}}) \end{aligned} \quad (27)$$

for  $i \in \{1, \dots, K\}$ , where  $\mathbf{s}_{\theta}^{\text{det}}$  is the solution of (23),  $\nabla g_{\theta_i}(\mathbf{x})|_{\mathbf{x}=\mathbf{s}_{\theta}^{\text{det}}}$  denotes the gradient of  $g_{\theta_i}(\mathbf{x})$  at  $\mathbf{x} = \mathbf{s}_{\theta}^{\text{det}}$ ,  $\mathbf{H}_{\theta_i}$  is the Hessian of  $g_{\theta_i}(\mathbf{x})$  at  $\mathbf{x} = \mathbf{s}_{\theta}^{\text{det}}$ , and  $\gamma_i \triangleq \max_{i} g_{\theta_i}(\mathbf{s}_{\theta}^{\text{det}}) - g_{\theta_i}(\mathbf{s}_{\theta}^{\text{det}})$ .

The conditions in Proposition 6 provide a relatively simple technique, which is based on the first and second order derivatives of  $g_{\theta_i}$ , for determining if the stochastic parameter design can provide improvements over the deterministic one. If the conditions are satisfied, the stochastic parameter design is guaranteed to outperform the deterministic parameter design, in which case the optimization problem in (15) can be solved to obtain the optimal solution. It should also be noted that there may exist scenarios in which the stochastic parameter design provides improvements over the deterministic one even though the the conditions in Proposition 6 are not satisfied, which is due to the fact that the conditions are sufficient but not necessary. In the next section, examples are presented for various scenarios.

## V. NUMERICAL RESULTS

In this section, numerical examples are presented in order to investigate the performance of the optimal parameter design

approach in various scenarios. Consider a wireless sensor network scenario in which two CDMA users aim to send information about two scalar parameters,  $\theta_1$  and  $\theta_2$ , to two intended devices as in Fig. 1. Then, parameter vector  $\boldsymbol{\theta} = [\theta_1 \ \theta_2]^T$  is to be estimated based on observation vector  $\mathbf{y} = [y_1 \ y_2]^T$ , which is modeled via (1) as

$$\mathbf{y} = \mathbf{s}_\theta + \rho(\mathbf{1} - \mathbf{I})\mathbf{s}_\theta + \mathbf{n} \quad (29)$$

where  $\mathbf{s}_\theta = [s_{\theta_1} \ s_{\theta_2}]^T$  consists of the transmitted (stochastic) signals for the two intended devices in the wireless sensor network for parameter values  $\theta_1$  and  $\theta_2$  ( $s_{\theta_1}$  and  $s_{\theta_2}$  can be any function of  $\theta_1$  and  $\theta_2$ , respectively),  $\mathbf{n} = [n_1 \ n_2]^T$  represents additive noise at the intended devices,  $\rho = \rho_{12} = \rho_{21}$  denotes the crosscorrelation parameter that is determined by the spreading sequences employed by the users (see Section I-A),  $\mathbf{I}$  is the identity matrix of size  $2 \times 2$ , and  $\mathbf{1}$  is the matrix of ones with the same size. The components  $n_1$  and  $n_2$  of the additive noise  $\mathbf{n}$  are independent and identically distributed Gaussian random variables, specified by PDFs  $p_{n_1}(\mathbf{n}) = p_{n_2}(\mathbf{n}) = \exp\{-(\mathbf{n} - \boldsymbol{\mu})^2 / (2\sigma^2)\} / (\sqrt{2\pi}\sigma)$ , which is a common model employed in wireless communication systems. The estimator is specified by  $\hat{\boldsymbol{\theta}}(\mathbf{y}) = \mathbf{y}$ , which estimates each parameter independently based on the corresponding observation. The cost function for each parameter is chosen as the uniform cost function, which is calculated as  $C[\hat{\theta}_i(y_i), \theta_i] = 1$  if  $|\hat{\theta}_i(y_i) - \theta_i| > \Delta$  and  $C[\hat{\theta}_i(y_i), \theta_i] = 0$  otherwise for  $i = 1, 2$ . Based on this model,  $g_{\theta_1}$  in (4) can be obtained as

$$g_{\theta_1}(\mathbf{x}) = Q\left(\frac{x_1 + \rho x_2 - \theta_1 + \mu + \Delta}{\sigma}\right) + Q\left(\frac{-x_1 - \rho x_2 + \theta_1 - \mu + \Delta}{\sigma}\right) \quad (30)$$

where  $Q(x) = (1/\sqrt{2\pi}) \int_x^\infty \exp\{-t^2/2\} dt$  denotes the  $Q$ -function. For each  $\boldsymbol{\theta}$ ,  $E\{\|\mathbf{s}_\theta\|^2\} \leq \|\boldsymbol{\theta}\|^2$  is employed as the constraint stated in (11). Similarly to (30),  $g_{\theta_2}$  for the second parameter can be obtained.

In the numerical examples, the parameter spaces for both parameters are specified as  $\Lambda_1 = \Lambda_2 = [-10, 10]$ . Also,  $s_{\theta_1}$  and  $s_{\theta_2}$  can take values in the interval  $[-10, 10]$  subject to the average power constraint,  $E\{\|\mathbf{s}_\theta\|^2\} \leq \|\boldsymbol{\theta}\|^2$ . Also, the Gaussian distribution of the noise is taken to be zero mean with  $\sigma = 0.5$  and  $\rho$  is chosen to be 0.25. Since the noise is a zero-mean random variable,  $\hat{\boldsymbol{\theta}}(\mathbf{y}) = \mathbf{y}$  can be considered as a practical estimator.<sup>4</sup> In addition,  $\Delta = 1$  is used for the uniform cost function described in the previous paragraph. To solve the optimization problems, the Multi-Start algorithm [23] is used in MATLAB, which employs a local solver from multiple start points to reach the global optimal solution of a non-convex problem.

In Fig. 2, the total Bayes risks for the stochastic parameter design, the unconstrained parameter design and the conventional parameter design (which transmits the parameters as they are; that is, employs  $s_{\theta_i} = \theta_i$ ) are illustrated. Also in Fig. 3, the total Bayes risks for the stochastic parameter design and the deterministic parameter design are compared. It is observed that the stochastic parameter design achieves

<sup>4</sup>Although this is not the optimal estimator, it can be used in practice due to its simplicity compared to the optimal estimator.

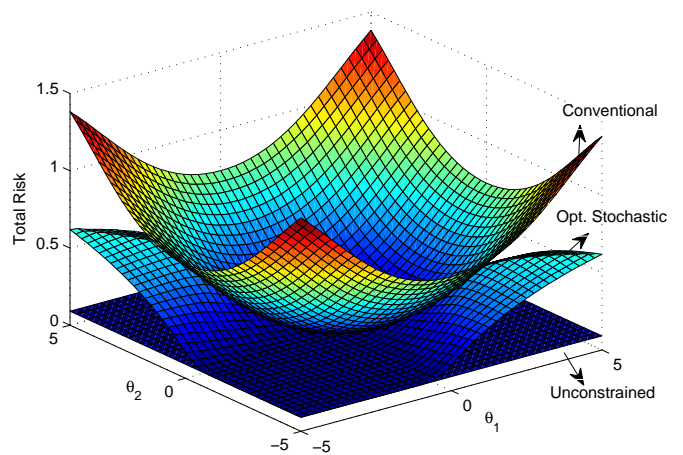


Fig. 2. Total Bayes risk versus  $\theta_1$  and  $\theta_2$ .

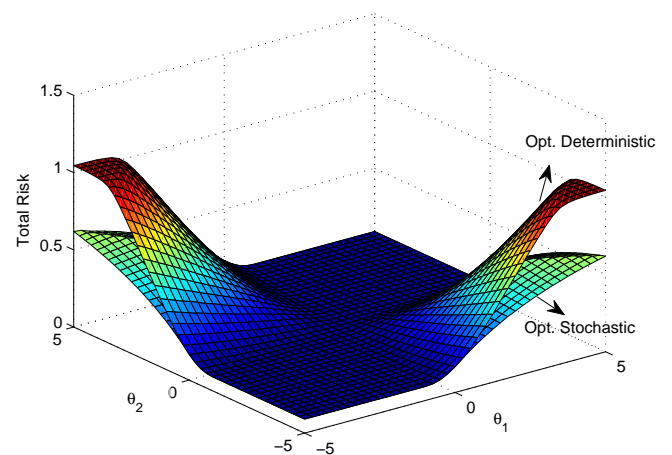


Fig. 3. Total Bayes risk versus  $\theta_1$  and  $\theta_2$ .

improvements over the deterministic and conventional designs. Also, for some values of  $\theta_1$  and  $\theta_2$ , the performance of the stochastic design is the same as the unconstrained one.

In Fig. 4, the maximum Bayes risks for the stochastic parameter design, the unconstrained parameter design and the conventional parameter design are plotted. Also, in Fig. 5, the maximum Bayes risks for stochastic parameter design and the deterministic parameter design are illustrated. Similar to the previous scenario, it is observed that the stochastic parameter design provides improvements over the conventional and deterministic parameter design approaches for certain range of parameter values.

In Table I, the total Bayes risk criterion is considered, and the optimal solutions for the stochastic, the deterministic and the unconstrained parameter design approaches are presented for various values of  $\boldsymbol{\theta}$ . It is observed from the table that the optimal stochastic parameter design can involve randomization between two different signals for certain values of  $\boldsymbol{\theta}$ , which corresponds to the cases in which the stochastic approach outperforms the deterministic parameter design, as can be verified from Fig. 3. Similarly, Table II presents the optimal solutions for the stochastic, the deterministic and the unconstrained parameter design approaches for the maximum



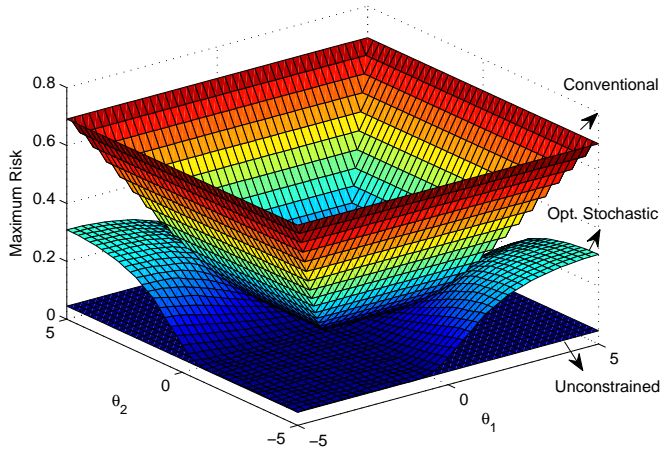


Fig. 4. Maximum Bayes risk versus  $\theta_1$  and  $\theta_2$ .

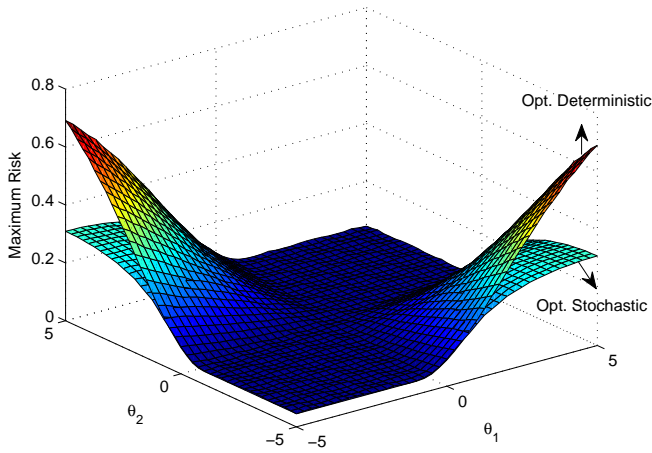


Fig. 5. Maximum Bayes risk versus  $\theta_1$  and  $\theta_2$ .

Bayes risk criterion. The main difference in this scenario is that randomization (time sharing) among up to three different signals can be performed for the optimal stochastic parameter design in accordance with Proposition 1.

Next, the maximum Bayes risk criterion is considered, and the conditions in Proposition 6 are studied. Namely, the existence of a real vector  $\mathbf{z}$  and a positive number  $k$  that satisfy the conditions in Proposition 6 for a certain value of  $\theta$  is investigated. Consider the parameter value  $\theta = [-5, 5]$ . If all the four conditions (two conditions for each estimator) are satisfied for this value of  $\theta$ , then it is guaranteed that the stochastic parameter design yields a lower maximum Bayes risk than the deterministic design. To test the first condition for each estimator, we need the value of  $\mathbf{s}_\theta^{\text{det}}$  and the gradients of  $g_{\theta_1}(\mathbf{x})$  and  $g_{\theta_2}(\mathbf{x})$  at  $\mathbf{x} = \mathbf{s}_\theta^{\text{det}}$ . As it can be observed from Table II,  $\mathbf{s}_\theta^{\text{det}} = [-5, 5]$  for  $\theta = [-5, 5]$ , and the gradients of  $g_{\theta_1}(\mathbf{x})$  and  $g_{\theta_2}(\mathbf{x})$  at  $\mathbf{x} = \mathbf{s}_\theta^{\text{det}}$  can be calculated based on the

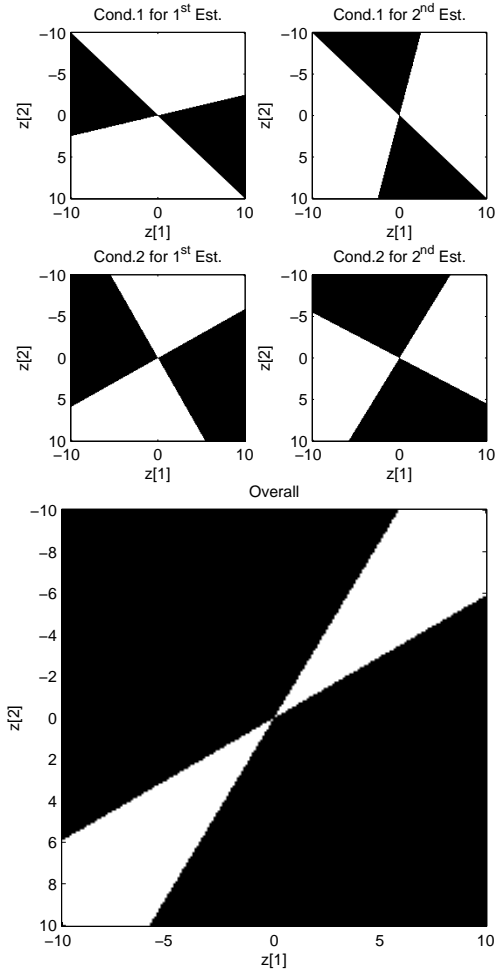


Fig. 6. Regions (white) in which the optimality conditions stated in Proposition 6 are satisfied for different values of  $\mathbf{z}$  for  $\theta = [-5, 5]$ .

following equations:

$$\begin{aligned} \nabla g_{\theta_1}(\mathbf{x}) &= \frac{-e^{-\frac{a_1^2}{2}} + e^{-\frac{a_2^2}{2}}}{\sqrt{2\pi}} [1, \rho]^T \\ \nabla g_{\theta_2}(\mathbf{x}) &= \frac{-e^{-\frac{b_1^2}{2}} + e^{-\frac{b_2^2}{2}}}{\sqrt{2\pi}} [1, \rho]^T \end{aligned} \quad (31)$$

where  $a_1 = (\mathbf{x}_1 - \rho\mathbf{x}_2 - \theta_1 + \mu + \Delta)/\sigma$ ,  $a_2 = (-\mathbf{x}_1 + \rho\mathbf{x}_2 + \theta_1 - \mu + \Delta)/\sigma$ ,  $b_1 = (\mathbf{x}_2 - \rho\mathbf{x}_1 - \theta_2 + \mu + \Delta)/\sigma$ , and  $b_2 = (-\mathbf{x}_2 + \rho\mathbf{x}_1 + \theta_2 + \mu + \Delta)/\sigma$ . Based on these equations, the first condition in Proposition 6 can be evaluated for each estimator. The first two plots in Fig. 6 illustrate the values of  $\mathbf{z}$  for which the first condition in Proposition 6 is satisfied for the first and the second estimator, respectively, for  $k = 1$  and  $\theta = [-5, 5]$ . Namely, in the white (black) regions, the conditions are satisfied (not satisfied). As observed from the figure, there are certain regions in which the first condition is satisfied for each estimator. Next, the second condition in Proposition 6 is tested. To that aim, the Hessians of  $g_{\theta_1}(\mathbf{x})$  and  $g_{\theta_2}(\mathbf{x})$  at  $\mathbf{x} = \mathbf{s}_\theta^{\text{det}}$  are calculated. The Hessians of these

TABLE I

UNCONSTRAINED SOLUTION  $p_{s_\theta}^{\text{opt}}(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{s}_\theta^{\text{unc}})$ , OPTIMAL DETERMINISTIC SOLUTION  $p_{s_\theta}^{\text{opt}}(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{s}_\theta^{\text{opt}})$ , AND OPTIMAL STOCHASTIC SOLUTION  $p_{s_\theta}^{\text{opt}}(\mathbf{x}) = \lambda_{\theta,1} \delta(\mathbf{x} - \mathbf{s}_{\theta,1}) + \lambda_{\theta,2} \delta(\mathbf{x} - \mathbf{s}_{\theta,2})$  FOR THE TOTAL BAYES RISK CRITERION.

$\theta_1$	$\theta_2$	$\mathbf{s}_\theta^{\text{unc}}$	$\mathbf{s}_\theta^{\text{opt}}$	$\lambda_{\theta,1}$	$\mathbf{s}_{\theta,1}$	$\lambda_{\theta,2}$	$\mathbf{s}_{\theta,2}$
-5	-5	(-4,-4)	(-4,-4)	1	(-4,-4)	-	-
-5	-2.5	(-4.667,-1.333)	(-4.667,-1.333)	1	(-4.667,-1.333)	-	-
-5	0	(-5.333,1.333)	(-4.889,1.049)	1	(-4.889,1.049)	-	-
-5	2.5	(-6,4)	(-4.694,3.035)	0.669	(-5.565,3.622)	0.331	(0.561,2.243)
-5	5	(-6.667,6.667)	(-5.950,3.820)	0.508	(1.138,4.553)	0.492	(-6.280,6.280)
-2.5	-5	(-1.333,-4.667)	(-1.333,-4.667)	1	(-1.333,-4.667)	-	-
-2.5	-2.5	(-2,-2)	(-2,-2)	1	(-2,-2)	-	-
-2.5	0	(-2.667,0.667)	(-2.444,0.527)	1	(-2.444,0.527)	-	-
-2.5	2.5	(-3.333,3.333)	(-2.5,2.5)	0.918	(-2.572,2.572)	0.082	(-2.023,-0.206)
-2.5	5	(-4,6)	(-3.035,4.694)	0.669	(-3.622,5.565)	0.331	(-2.243,-0.561)
0	-5	(1.333,-5.333)	(1.049,-4.889)	1	(1.049,-4.889)	-	-
0	-2.5	(0.667,-2.667)	(0.527,-2.444)	1	(0.527,-2.444)	-	-
0	2.5	(-0.667,2.667)	(-0.527,2.444)	1	(-0.527,2.444)	-	-
0	5	(-1.333,5.333)	(-1.049,4.889)	1	(-1.049,4.889)	-	-

TABLE II

UNCONSTRAINED SOLUTION  $p_{s_\theta}^{\text{opt}}(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{s}_\theta^{\text{unc}})$ , OPTIMAL DETERMINISTIC SOLUTION  $p_{s_\theta}^{\text{opt}}(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{s}_\theta^{\text{opt}})$ , AND OPTIMAL STOCHASTIC SOLUTION  $p_{s_\theta}^{\text{opt}}(\mathbf{x}) = \lambda_{\theta,1} \delta(\mathbf{x} - \mathbf{s}_{\theta,1}) + \lambda_{\theta,2} \delta(\mathbf{x} - \mathbf{s}_{\theta,2}) + \lambda_{\theta,3} \delta(\mathbf{x} - \mathbf{s}_{\theta,3})$  FOR THE MAXIMUM BAYES RISK CRITERION.

$\theta_1$	$\theta_2$	$\mathbf{s}_\theta^{\text{unc}}$	$\mathbf{s}_\theta^{\text{opt}}$	$\lambda_{\theta,1}$	$\mathbf{s}_{\theta,1}$	$\lambda_{\theta,2}$	$\mathbf{s}_{\theta,2}$	$\lambda_{\theta,3}$	$\mathbf{s}_{\theta,3}$
-5	-5	(-4,-4)	(-4,-4)	1	(-4,-4)	-	-	-	-
-5	-2.5	(-4.667,-1.333)	(-4.667,-1.333)	1	(-4.667,-1.333)	-	-	-	-
-5	0	(-5.333,1.333)	(-4.915,0.915)	1	(-4.915,0.915)	-	-	-	-
-5	2.5	(-6,4)	(-4.824,2.824)	0.637	(-5.498,3.416)	0.194	(0.503,2.167)	0.169	(-4.481,-1.094)
-5	5	(-6.667,6.667)	(-5,5)	0.487	(-6.264,6.313)	0.261	(-4.607,-0.995)	0.252	(0.794,4.682)
-2.5	-5	(-1.333,-4.667)	(-1.333,-4.667)	1	(-1.333,-4.667)	-	-	-	-
-2.5	-2.5	(-2,-2)	(-2,-2)	1	(-2,-2)	-	-	-	-
-2.5	0	(-2.667,0.667)	(-2.458,0.458)	0.998	(-2.458,0.458)	0.002	(-2.464,0.465)	-	-
-2.5	2.5	(-3.333,3.333)	(-2.5,2.5)	0.879	(-2.617,2.599)	0.066	(0.415,2.093)	0.055	(-2.011,-0.553)
-2.5	5	(-4,6)	(-2.824,4.824)	0.623	(-3.458,5.502)	0.194	(-2.211,-0.498)	0.183	(1.158,4.493)
0	-5	(1.333,-5.333)	(0.915,-4.915)	1	(0.915,-4.915)	-	-	-	-
0	-2.5	(0.667,-2.667)	(0.458,-2.458)	1	(0.458,-2.458)	-	-	-	-
0	2.5	(-0.667,2.667)	(-0.458,2.458)	1	(-0.458,2.458)	-	-	-	-
0	5	(-1.333,5.333)	(-0.915,4.915)	1	(-0.915,4.915)	-	-	-	-

functions can be found as follows:

$$\mathbf{H}_{\theta_1}(\mathbf{x}) = \frac{a_1 e^{-\frac{a_1^2}{2}} + a_2 e^{-\frac{a_2^2}{2}}}{\sqrt{2\pi}} \begin{bmatrix} 1 & \rho \\ \rho & \rho^2 \end{bmatrix}$$

$$\mathbf{H}_{\theta_2}(\mathbf{x}) = \frac{b_1 e^{-\frac{b_1^2}{2}} + b_2 e^{-\frac{b_2^2}{2}}}{\sqrt{2\pi}} \begin{bmatrix} \rho^2 & \rho \\ \rho & 1 \end{bmatrix} \quad (32)$$

where  $a_1$ ,  $a_2$ ,  $b_1$ , and  $b_2$  are as defined previously. Based on (31) and (32), the second condition in Proposition 6 can be evaluated for each estimator. The results of these evaluations are shown in the second and third plots in Fig. 6 for different values of  $\mathbf{z}$ . Similar to the first condition, the second condition is satisfied in certain range of  $\mathbf{z}$  values (white regions). The last plot in Fig. 6 shows the intersection of the regions in which the conditions are satisfied. As observed from the figure, the intersection is not an empty set, hence we can conclude that there exist a real vector  $\mathbf{z}$  and a positive number  $k$  for which all the conditions in Proposition 6 are satisfied. Therefore, in this scenario, it is guaranteed that the stochastic parameter design achieves a lower Bayes risk than the deterministic design for  $\theta = [-5, 5]$  as a result of Proposition 6. The applicability of Proposition 6 is also investigated for the whole parameter space,  $\theta \in [-10, 10] \times [-10, 10]$ , and the sets of parameter values for which the stochastic design provides (i) no improvements over the deterministic design, (ii) improve-

ments over the deterministic design but Proposition 6 does not apply, and (iii) improvements over the deterministic design and Proposition 6 applies are specified. The calculations show that, in the considered example, Proposition 6 provides sufficient conditions for improbability that are valid over a significant portion (about 58%) of the improbability region (in particular, for large values of the parameters in the improbability region) but the conditions are not necessary in general.

In order to gain intuition and further understanding, Fig. 7 and Fig. 8 visualize how the stochastic design approach provides performance improvements over the deterministic one for the total Bayes risk and the maximum Bayes risk criteria, respectively. In Fig. 7, the total Bayes risk  $\tilde{g}_\theta(\mathbf{x})$  is illustrated for  $\theta = [-5, 5]$ . The region inside the circle corresponds to the values of  $\mathbf{x}$  that satisfy the power constraint individually. Here it can be seen that the minimum value of  $\tilde{g}_\theta(\mathbf{x})$  is observed at a value of  $\mathbf{x}$  which does not satisfy the power constraint. In that case, the unconstrained solution simply picks that value of  $\mathbf{x}$  as  $\mathbf{s}_\theta^{\text{unc}}$ . On the other hand, the optimal deterministic solution, picks the value of  $\mathbf{x}$  residing on or inside the circle, which minimizes the value of  $\tilde{g}_\theta(\mathbf{x})$ , denoted as  $\mathbf{s}_\theta^{\text{det}}$ . Obviously, there is a performance gap between the unconstrained solution and the optimal deterministic solution, as can be observed from Fig. 2 and Fig. 3. The proposed stochastic design approach aims to achieve improvements over the deterministic solution.

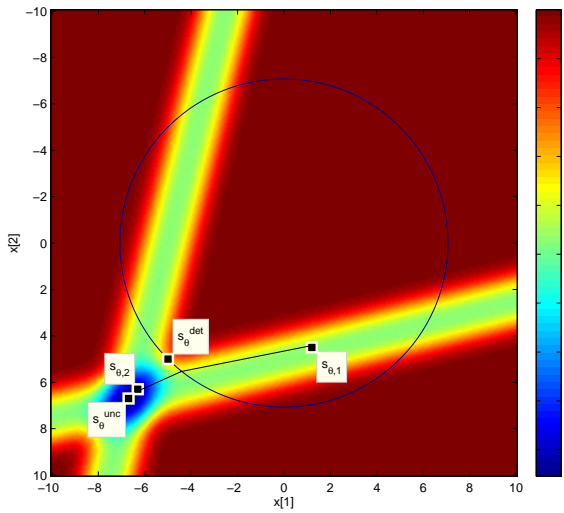


Fig. 7. The total Bayes risk  $\tilde{g}_\theta(\mathbf{x})$  for  $\theta = [-5, 5]$ .

To that aim, the stochastic design perform randomization (time sharing) between two signals, i.e., one that satisfies the power constraint with low performance, the other that does not satisfy the constraint but has high performance, as can be seen from Fig. 7. By randomizing between these two signals, it is possible to satisfy the power constraint on the average and to achieve a better performance than the optimal deterministic design. In Fig. 8, the Bayes risks for the estimators,  $g_{\theta_1}(\mathbf{x})$  and  $g_{\theta_2}(\mathbf{x})$ , are illustrated for  $\theta = [-5, 5]$ . As stated in (14), the stochastic design aims to minimize the maximum of the expectations of  $g_{\theta_1}(\mathbf{x})$  and  $g_{\theta_2}(\mathbf{x})$ . It is observed from Fig. 8 that the minimum values of  $g_{\theta_1}(\mathbf{x})$  and  $g_{\theta_2}(\mathbf{x})$  coincide outside the power constraint. Therefore, it is not possible to pick a single point which minimizes the maximum Bayes risk and satisfy the constraint at the same time. Similar to the total Bayes risk case, the stochastic design should perform randomization between some signals inside and outside the constraint to achieve improvements over the deterministic design. However, as seen in Fig. 8, the minimum values of  $g_{\theta_1}(\mathbf{x})$  and  $g_{\theta_2}(\mathbf{x})$  do not coincide inside the constraint and as a result any signal minimizing one of  $g_{\theta_1}(\mathbf{x})$  and  $g_{\theta_2}(\mathbf{x})$  maximizes the other one. Hence it is not possible to pick just one signal that satisfies the constraint and randomize it with another signal that does not satisfy the constraint. To overcome this problem, two signals satisfying the constraint should be chosen and these two signals should be randomized with a signal that does not satisfy the power constraint. As a result, the expectations of both  $g_{\theta_1}(\mathbf{x})$  and  $g_{\theta_2}(\mathbf{x})$  are minimized to a certain point, which makes it possible to minimize the maximum of these expectations.

Finally, in Fig. 9 and Fig. 10, the total Bayes risks (black) and the maximum Bayes risks (red) of the different approaches are plotted versus the standard deviation  $\sigma$  of the noise components and the cross-correlation parameter  $\rho$  in (29), respectively, for  $\theta = [-5, 5]$ , where  $\rho = 0.25$  in Fig. 9 and  $\sigma = 0.5$  in and Fig. 10. It is observed that the optimal stochastic design provides improvements over the optimal deterministic design, and the two algorithms have similar performance for small  $\rho$  and/or for large values of  $\sigma$  (i.e., in the noise-limited regime).

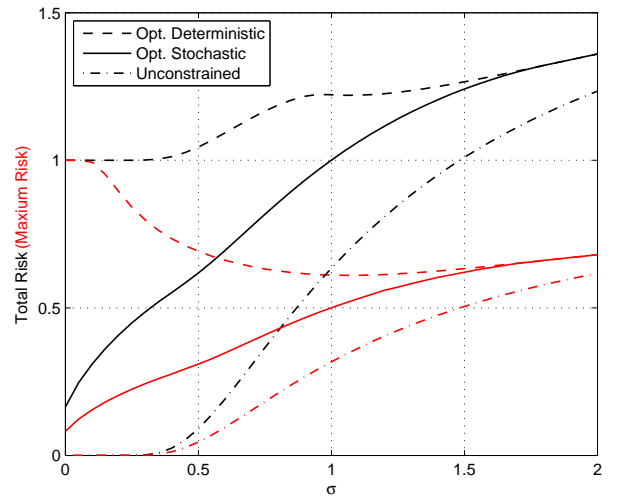


Fig. 9. The total Bayes risk (black) and the maximum Bayes risk (red) versus the noise standard deviation  $\sigma$  for  $\theta = [-5, 5]$  and  $\rho = 0.25$ .

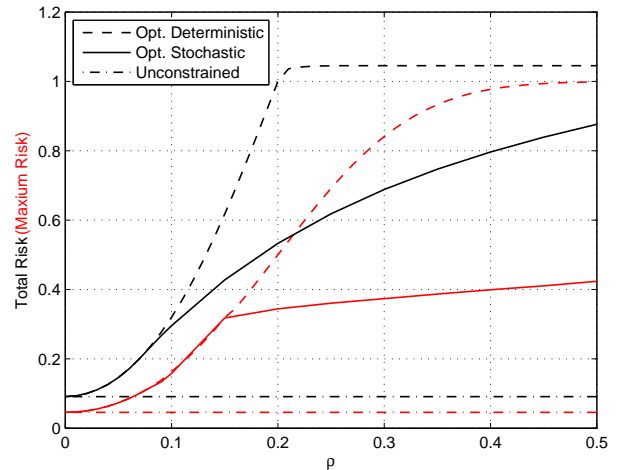


Fig. 10. The total Bayes risk (black) and the maximum Bayes risk (red) versus the cross-correlation parameter in (29) for  $\theta = [-5, 5]$  and  $\sigma = 0.5$ .

## VI. CONCLUSIONS

In this manuscript, the optimal stochastic design of multiple parameters has been studied for a given set of fixed estimators. Two different performance criteria have been considered; namely, the total Bayes risk criterion and the maximum Bayes risk criterion. It has been shown that, in the presence of  $K$  parameters, the optimal stochastic parameter design results in time sharing (randomization) among at most two and  $(K + 1)$  different signals values for the total Bayes risk and the maximum Bayes risk criteria, respectively. In addition, the average transmitted signal powers corresponding to the optimal parameter design approaches have been specified, and the characterization of the optimal approaches has been provided in various scenarios. Furthermore, various conditions have been derived in order to specify when the stochastic parameter design or the deterministic parameter design is optimal. Finally, the numerical examples have been presented to investigate the theoretical results, and to illustrate the amount of improvements achieved via the proposed approach.

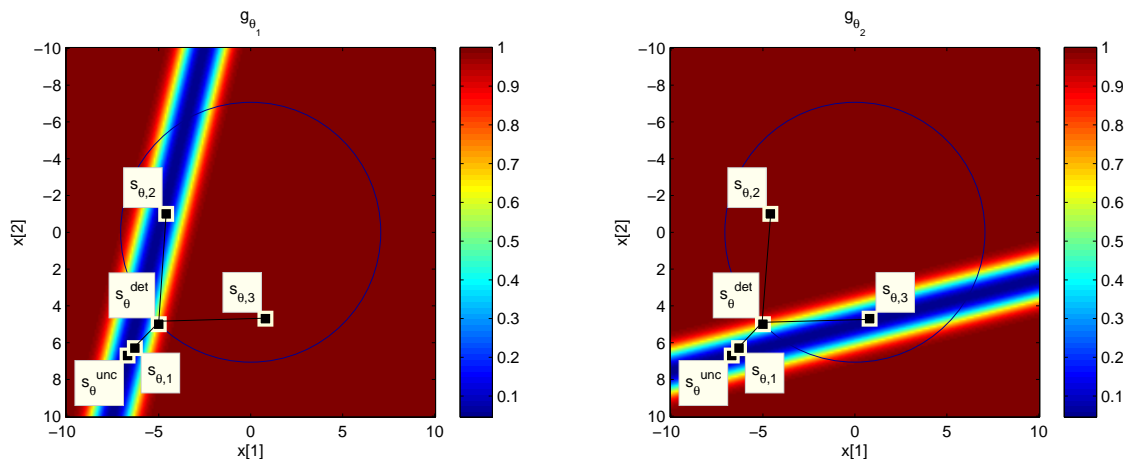


Fig. 8. The Bayes risk for the first estimator,  $g_{\theta_1}(\mathbf{x})$ , and the Bayes risk for the second estimator,  $g_{\theta_2}(\mathbf{x})$ , for  $\theta = [-5, 5]$ .

## REFERENCES

- [1] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice Hall, Inc., Upper Saddle River, NJ, 1993.
- [2] H. V. Poor, *An Introduction to Signal Detection and Estimation*, Springer-Verlag, New York, 1994.
- [3] H. Soganci, S. Gezici, and O. Arikan, "Optimal stochastic parameter design for estimation problems," *IEEE Trans. Sig. Processing*, vol. 60, no. 9, pp. 4950–4956, Sep. 2012.
- [4] H. Chen, P. K. Varshney, S. M. Kay, and J. H. Michels, "Theory of the stochastic resonance effect in signal detection: Part I—Fixed detectors," *IEEE Trans. Sig. Processing*, vol. 55, no. 7, pp. 3172–3184, July 2007.
- [5] A. Patel and B. Kosko, "Optimal noise benefits in Neyman-Pearson and inequality-constrained signal detection," *IEEE Trans. Sig. Processing*, vol. 57, no. 5, pp. 1655–1669, May 2009.
- [6] S. M. Kay, "Noise enhanced detection as a special case of randomization," *IEEE Sig. Processing Lett.*, vol. 15, pp. 709–712, 2008.
- [7] S. M. Kay, J. H. Michels, H. Chen, and P. K. Varshney, "Reducing probability of decision error using stochastic resonance," *IEEE Sig. Processing Lett.*, vol. 13, no. 11, pp. 695–698, Nov. 2006.
- [8] S. Bayram and S. Gezici, "Noise-enhanced  $M$ -ary hypothesis-testing in the minimax framework," in *Proc. International Conference on Signal Processing and Commun. Systems*, Omaha, Nebraska, Sep. 2009, pp. 31–36.
- [9] S. Bayram, S. Gezici, and H. V. Poor, "Noise enhanced hypothesis-testing in the restricted Bayesian framework," *IEEE Trans. Sig. Processing*, vol. 58, no. 8, pp. 3972–3989, Aug. 2010.
- [10] S. Bayram and S. Gezici, "Noise enhanced  $M$ -ary composite hypothesis-testing in the presence of partial prior information," *IEEE Trans. Sig. Processing*, vol. 59, no. 3, pp. 1292–1297, Mar. 2011.
- [11] C. Goken, S. Gezici, and O. Arikan, "Optimal stochastic signaling for power-constrained binary communications systems," *IEEE Trans. Wireless Commun.*, vol. 9, no. 12, pp. 3650–3661, Dec. 2010.
- [12] C. Goken, S. Gezici, and O. Arikan, "Optimal signaling and detector design for power-constrained binary communications systems over non-Gaussian channels," *IEEE Commun. Letters*, vol. 14, no. 2, pp. 100–102, Feb. 2010.
- [13] M.E. Tutay, S. Gezici, and O. Arikan, "Optimal randomization of signal constellations on the downlink of a multiuser DS-CDMA system," *IEEE Transactions on Wireless Communications*, vol. 12, no. 10, pp. 4878–4891, Oct. 2013.
- [14] B. Dulek and S. Gezici, "Detector randomization and stochastic signaling for minimum probability of error receivers," *IEEE Trans. Commun.*, vol. 60, no. 4, pp. 923–928, Apr. 2012.
- [15] M. E. Tutay, S. Gezici, and O. Arikan, "Optimal detector randomization for multiuser communications systems," *IEEE Trans. Commun.*, vol. 61, no. 7, pp. 2876–2889, July 2013.
- [16] B. Dulek, N. D. Vanli, S. Gezici, and P. K. Varshney, "Optimum power randomization for the minimization of outage probability," *IEEE Trans. Wireless Commun.*, vol. 12, no. 9, pp. 4627–4637, Sep. 2013.
- [17] H. Chen, P. K. Varshney, and J. H. Michels, "Noise enhanced parameter estimation," *IEEE Trans. Sig. Processing*, vol. 56, no. 10, pp. 5074–5081, Oct. 2008.
- [18] M. Azizoglu, "Convexity properties in binary detection problems," *IEEE Trans. Inform. Theory*, vol. 42, no. 4, pp. 1316–1321, July 1996.
- [19] B. Dulek, S. Gezici, and O. Arikan, "Convexity properties of detection probability under additive Gaussian noise: Optimal signaling and jamming strategies," *IEEE Trans. Signal Process.*, vol. 61, no. 13, pp. 3303–3310, July 2013.
- [20] S. Bayram, N. D. Vanli, B. Dulek, I. Sezer, and S. Gezici, "Optimum power allocation for average power constrained jammers in the presence of non-Gaussian noise," *IEEE Commun. Letters*, vol. 16, no. 8, pp. 1153–1156, Aug. 2012.
- [21] S. Verdú, *Multiuser Detection*, 1st ed. Cambridge, UK: Cambridge University Press, 1998.
- [22] S. Bayram and S. Gezici, "Stochastic resonance in binary composite hypothesis-testing problems in the Neyman-Pearson framework," *Digital Signal Processing*, vol. 22, no. 3, pp. 391–406, May 2012.
- [23] Z. Ugray, Leon Lasdon, John Plummer, Fred Glover, James Kelly, and Rafael Marti, "Scatter search and local NLP solvers: A multistart framework for global optimization," *INFORMS Journal on Computing*, vol. 19, no. 3, pp. 328–340, 2007.



**Hamza Soganci** received his B.Sc., M.Sc. and Ph.D. degrees in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 2007, 2009, and 2015, respectively. He is currently working at TÜBİTAK SAGE. His research interests are detection and estimation theory, wireless localization, and ultra-wideband systems.



**Sinan Gezici** (S'03, M'06, SM'11) received the B.S. degree from Bilkent University, Turkey in 2001, and the Ph.D. degree in Electrical Engineering from Princeton University in 2006. From 2006 to 2007, he worked at Mitsubishi Electric Research Laboratories, Cambridge, MA. Since 2007, he has been with the Department of Electrical and Electronics Engineering at Bilkent University, where he is currently an Associate Professor.

Dr. Gezici's research interests are in the areas of detection and estimation theory, wireless communications, and localization systems. Among his publications in these areas is the book *Ultra-wideband Positioning Systems: Theoretical Limits, Ranging Algorithms, and Protocols* (Cambridge University Press, 2008). Dr. Gezici is an associate editor for *IEEE Transactions on Communications*, *IEEE Wireless Communications Letters*, and *Journal of Communications and Networks*.



**Orhan Arikan** (M'91) was born in 1964 in Manisa, Turkey. In 1986, he received his B.Sc. degree in Electrical and Electronics Engineering from the Middle East Technical University, Ankara, Turkey. He received both his M.S. and Ph.D. degrees in Electrical and Computer Engineering from the University of Illinois, Urbana-Champaign, in 1988 and 1990, respectively. Following his graduate studies, he was employed as a Research Scientist at Schlumberger-Doll Research Center, Ridgefield, CT. In 1993

he joined the Electrical and Electronics Engineering Department of Bilkent University, Ankara, Turkey. Since 2011, he is serving as the department chairman. His current research interests include statistical signal processing, time-frequency analysis and remote sensing. Dr. Arikan has served as Chairman of IEEE Signal Processing Society Turkey Chapter and President of IEEE Turkey Section.