entropy

MDPI

Article

Optimal Detection under the Restricted Bayesian Criterion

Shujun Liu [1], Ting Yang [1] and Hongqing Liu [2,*]

[1] College of Communication Engineering, Chongqing University, Chongqing 400044, China; liusj@cqu.edu.cn (S.L.); yandtt2008@126.com (T.Y.)
[2] Chongqing Key Lab of Mobile Communications Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
* Correspondence: hongqingliu@outlook.com; Tel.: +86-023-6346-0781

Received: 8 May 2017; Accepted: 18 July 2017; Published: 19 July 2017

Abstract: This paper aims to find a suitable decision rule for a binary composite hypothesis-testing problem with a partial or coarse prior distribution. To alleviate the negative impact of the information uncertainty, a constraint is considered that the maximum conditional risk cannot be greater than a predefined value. Therefore, the objective of this paper becomes to find the optimal decision rule to minimize the Bayes risk under the constraint. By applying the Lagrange duality, the constrained optimization problem is transformed to an unconstrained optimization problem. In doing so, the restricted Bayesian decision rule is obtained as a classical Bayesian decision rule corresponding to a modified prior distribution. Based on this transformation, the optimal restricted Bayesian decision rule is analyzed and the corresponding algorithm is developed. Furthermore, the relation between the Bayes risk and the predefined value of the constraint is also discussed. The Bayes risk obtained via the restricted Bayesian decision rule is a strictly decreasing and convex function of the constraint on the maximum conditional risk. Finally, the numerical results including a detection example are presented and agree with the theoretical results.

Keywords: restricted Bayesian; hypothesis-testing; Bayes risk

1. Introduction

Bayesian, Minimax, and Neyman-Pearson (NP) decisions are three common approaches in the applications of signal detection and processing [1–8]. For instance, a Bayesian approach is proposed in [2] for the signal detection in compressed sensing (CS). In [4], a Minimax framework is introduced for multiclass classification, which can be applied to general data including imagery and other types of high-dimensional data. In order to detect ultra-wideband signals in the presence of dense multipath channels and ambient noise, the NP theorem is used to derive the ultra-wideband (UWB) signal detector [5]. The NP and Bayesian framework are also utilized to access the performance of channel-aware binary-decision fusion over a shared Rayleigh flat-fading channel with multiple antennas at the Decision Fusion Center (DFC) [8], and the DFC is widely used in the spectrum sensing for cognitive radio scenarios [9].

In the classical hypothesis testing problem, a relevant part of uncertainty is usually represented by the prior probability distributions [10]. The aim of the Bayesian criterion is to minimize the Bayes risk [11] and the Bayesian decision rule is determined based on posterior probabilities, where the prior information is assumed to be completely known. On the other hand, no prior information is considered for the Minimax decision rule [12], which minimizes the maximum of the conditional risks given over the parameter space. In addition, the Neyman-Pearson decision rule can be also conceived in the presence of prior distributions [13–15]. Therefore, the Bayesian, Neyman-Pearson, and Minimax approaches can be viewed as three different ways of exploiting prior information. In

the former two cases, prior information is considered to be completely available, whereas no prior information is required in the latter. In practice, however, the extreme cases rarely happen, and the partial prior information or the prior information with uncertainty is often available, such as the adaptive vector subspace signal detection in partially homogeneous Gaussian disturbance and structured (unknown) deterministic interference within the framework of invariance theory [16].

In most practical cases, the prior distributions are available with certain degree of uncertainty because they are usually obtained based on the previous decisions and experiences [8]. As a result, the Bayesian and Neyman-Pearson approaches are ineffective due to the absence of complete prior information, and the Minimax approach achieves a poor performance since the available partial prior information is ignored. In order to utilize the partial information and to achieve a better performance, several studies have been conducted [8,17–25], for example, maximum entropy (ME), Γ-minimax, restricted Bayes, and restricted Neyman-Pearson approaches, to name but a few. For instance, the ME method is utilized in [23] to translate the information contained in the known form of the likelihood into a prior distribution for Bayesian inference. The restricted NP approach is applicable for a binary composite hypothesis-testing problem in the presence of prior information uncertainty [10]. In [25], the group minimax is obtained through the emergent theory of quantizing with Bregman divergences and a closed form Stolarsky mean expression is obtained by optimizing the minimax Bayes risk error divergence for the binary simple hypothesis testing with a partially known prior probability.

To the best of our knowledge, no previous work has focused on the optimal decision rule under the restricted Bayesian criterion for a binary composite hypothesis-testing problem in the presence of prior distribution uncertainty, where the uncertainties may exist not only in the prior probability of the null hypothesis, but also in the distribution probability of each parameter value under the null and the alternative hypotheses. In this paper, we utilize the constraints that the conditional risks should be less than a predefined value to reduce the negative effects of the uncertain prior information, thereby the focus of this paper is to find the optimal decision rule to minimize the Bayes risk on the basis of the constraints and then to explore the relations between the Bayes risk and the predefined value.

In order to solve the optimization problem, the Lagrange duality is applied to convert the minimization of the Bayes risk under the constraints on conditional risks to an unconstrained one. In doing so, the minimization of the Bayes risk under the constraints on conditional risks is equivalent to the minimization of the Bayes risk with a modified prior distribution based on this conversion. Finally, the corresponding theorems and algorithms are developed to search the restricted Bayes decision rule. If there is no uncertainty in the prior distribution, the constraint on the conditional risks is not necessary. In such a case, the classical Bayesian decision rule is applicable, which minimizes the Bayes risk without any constraints. If the value of the constraint is larger than the maximum conditional risk obtained in the classical Bayesian approach, the constraint on the conditional risks is ineffective and the restricted Bayesian decision rule is identical to the classical Bayesian decision rule. On the other hand, if the prior information is full of uncertainty, the Minimax decision rule can be utilized to minimize the worst-case (maximum) conditional Bayes risk. It should be noted that the lowest constraint achieved in the restricted Bayesian approach is equal to the Bayes risk obtained via the Minimax decision rule. Therefore, the classical Bayesian and Minimax approaches are two extreme cases of the restricted Bayesian approaches. In addition, the Bayes risk is a strictly decreasing and convex function of the constraint on conditional risks in the restricted Bayesian approach. The main contributions of this paper are summarized as follows:

- Formulation of the restricted Bayesian framework, which aims to minimize the Bayes risk under the constraint on the conditional risks.
- Derivations of the restricted Bayes decision rule.
- Algorithm for searching the restricted Bayes decision rule.
- Characteristics of the Bayes risk versus the constraint on the conditional risks.

The remainder of this paper is organized as follows: in Section 2, a restricted Bayesian framework is formulated for a binary composite hypothesis-testing problem, which aims to

minimize the Bayes risk under the constraint on the conditional risks. The optimal restricted Bayesian decision rule is explored in Section 3 and the corresponding algorithm is provided. Furthermore, the relation between the Bayes risk and the constraint is discussed in Section 4. Finally, numerical examples are presented in Section 5 to illustrate the theoretical results and conclusions are made in Section 6.

2. Problem Formulation

In the theory of signal detection, the detection problems such as radar and communication signal detection are usually formulated as a hypothesis testing problem, and the corresponding framework is developed to provide a theoretical and analytical basis for the detection of useful signal. In this paper, we consider a binary composite hypothesis-testing problem with partially known prior distribution, given by:

$$H_i : X \mid \theta \sim P_\theta, \ \ \theta \in \Lambda_i, \ \ i = 0,1, \tag{1}$$

where H_0 and H_1 are the null and the alternative hypotheses, respectively, X is a random variable with the sample space Γ and a K-dimensional observation vector $\mathbf{x} \in \mathbb{R}^K$, $p_\theta(\mathbf{x})$ denotes the pdf of $\mathbf{x}$ for a given parameter value $\Theta = \theta$, Λ_0 and Λ_1 are the respective sets of all possible parameter values of Θ under H_0 and H_1. Intuitively, the union of Λ_0 and Λ_1 forms the parameter set Λ, i.e., $\Lambda = \Lambda_0 \bigcup \Lambda_1$, and $\Lambda_0 \bigcap \Lambda_1 = \varnothing$. In addition, the prior distribution of Θ is denoted by $\omega(\theta)$, which is usually estimated based on previous observations and known up to a given degree of uncertainty due to the estimation errors.

If the Bayes risk is calculated based on the estimated prior distribution and the minimization of the Bayes risk is performed under the classical Bayesian criterion, then it means that the estimation errors are ignored directly. In doing so, a poor performance is obtained if the estimated distribution differs significantly from the correct one. On the other hand, if the Minimax criterion is utilized and the maximum conditional risk is minimized, then it fails to take the advantage of the useful prior information contained in the estimated prior distribution. In order to utilize the estimated prior distribution and alleviate the negative impact caused by the mismatch between the estimated prior distribution and the correct one, the restricted Bayesian criterion is applied in this paper. More specifically, this paper aims to minimize the Bayes risk, calculated based on the estimated prior distribution, under the constraint that the maximum conditional Bayes risk stays below a significance level that can be adjusted based on the degree of uncertainty in the estimated prior distribution.

Accordingly, the restricted Bayes optimization problem can be formulated by:

$$\min_{\phi} \int_{\Lambda} R_\theta(\phi)\omega(\theta)d\theta \tag{2}$$

subject to:

$$\max_{\theta \in \Lambda} R_\theta(\phi) \le \alpha, \tag{3}$$

where $r(\phi) = \int_{\Lambda} R_\theta(\phi)\omega(\theta)d\theta = E\{R_\theta(\phi)\}$ denotes the Bayes risk, ϕ represents the decision rule which maps the observation vector to 0 or 1, α is the upper limit on the maximum conditional risk, and $R_\theta(\phi)$ represents the conditional risk of $\phi(\cdot)$ for $\Theta = \theta$ and $\theta \in \Lambda$. Prior to calculating the conditional risk $R_\theta(\phi)$, a cost function $C[i,\theta]$ is used to assign costs to the decision results, where $C[i,\theta]$ denotes the cost of choosing H_i when $\Theta = \theta$ and $\theta \in \Lambda$. The $R_\theta(\phi)$ can be calculated as the average cost of decision rule $\phi(\mathbf{x})$ for $\Theta = \theta$, given by:

$$R_\theta(\phi) = E_\theta\{\mathrm{C}[\phi(\mathbf{x}),\theta]\} = \int_{\Gamma} \mathrm{C}[\phi(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}. \tag{4}$$

In order to solve the constrained optimization problem in (2) and (3), we introduce a regularization parameter λ to construct an unconstrained optimization problem as below:

$$\min_{\phi}\left\{\lambda\int_{\Lambda}\omega(\theta)\int_{\Gamma}\mathrm{C}[\phi(\mathbf{x}),\theta]p_{\theta}(\mathbf{x})d\mathbf{x}d\theta+(1-\lambda)\max_{\theta\in\Lambda}\int_{\Gamma}\mathrm{C}[\phi(\mathbf{x}),\theta]p_{\theta}(\mathbf{x})d\mathbf{x}\right\}, \tag{5}$$

which is also a transformation of the Lagrangian of the inequality-constrained optimization problem in (2) and (3), where λ is designed according to α and $0 \le \lambda \le 1$. In particular, λ decreases as α decreases, and this fact can be utilized to adjust the value of λ. Accordingly, the unconstrained optimization problem in (5) with a suitable λ is an alternative representation of the constrained optimization problem in (2) and (3) according to the Lagrange duality, the equivalence of them will be proved in the next section.

3. Restricted Bayes Decision Rule

In this section, based on the problem formulation in Section 2, the characteristics of the optimal decision rule under the restricted Bayes criterion are investigated first, and then an algorithm for finding the restricted Bayes decision rule is developed.

3.1. Characteristics of the Restricted Bayes Decision Rule

According to the formulation in (5), the following theorems are developed to characterize the restricted Bayes decision rule under certain conditions.

Theorem 1. *Let $g(\theta)=\lambda\omega(\theta)+(1-\lambda)v(\theta)$, where $v(\theta)$ is any valid pdf. If ϕ^* is the classical Bayes decision rule for the modified prior density $g(\theta)$ and satisfies following equality:*

$$\int_{\Lambda}v(\theta)\int_{\Gamma}\mathrm{C}[\phi^*(\mathbf{x}),\theta]p_{\theta}(\mathbf{x})d\mathbf{x}d\theta=\max_{\theta\in\Lambda}\int_{\Gamma}\mathrm{C}[\phi^*(\mathbf{x}),\theta]p_{\theta}(\mathbf{x})d\mathbf{x}, \tag{6}$$

then ϕ^* is a solution of the optimization problem in Equation (5).

The proof is presented in the Appendix. Theorem 1 indicates that the solution of the unconstrained optimization problem in (5) is calculated as a classical Bayes decision rule for a modified prior density under certain conditions. In other words, the optimization problem in (5) can be equivalent to a classical Bayesian optimization problem for the modified prior distribution $g(\theta)=\lambda\omega(\theta)+(1-\lambda)v(\theta)$ if $g(\theta)$ satisfies the equality in (8).

Next, in order to illustrate the equivalence between the optimization problem in (5) and that in (2) and (3), a proposition is developed as follows.

Proposition 1. *Under the conditions in Theorem 1, ϕ^* is also the solution of the optimization problem in Equations (2) and (3) if $\max_{\theta\in\Lambda}\int_{\Gamma}\mathrm{C}[\phi^*(\mathbf{x}),\theta]p_{\theta}(\mathbf{x})d\mathbf{x}=\alpha$.*

The proof is given in the Appendix. Proposition 1 implies that if the decision rule ϕ^* defined in Theorem 1 meets the constraint in (3) with equality, then it also provides a solution for the restricted Bayes optimization problem in (2) and (3). In other words, the minimum Bayes risk is achieved when the maximum conditional risk for $\theta \in \Lambda$ is equal to the upper limit α. Theorem 1 and Proposition 1 also build a relationship between λ and α. For any λ, the achievable upper limit α can be calculated by the equality in Proposition 1.

In addition, the modified prior distribution $g(\theta)$ specified in Theorem 1 is the least favorable distribution among a family of pdfs, which have the same form with $g(\theta)$. Explicitly, the achievable minimum Bayes risk corresponding to $g(\theta)$ specified in Theorem 1 is greater than or equal to that corresponding to any other distribution with the pdf form of $\hat{g}(\theta)=\hat{\lambda}\omega(\theta)+(1-\hat{\lambda})\hat{v}(\theta)$, where $\hat{\lambda}\ge\lambda$ and $\hat{v}(\theta)$ is any valid pdf. Theorem 2 further states this case.

Theorem 2. *Under the conditions in Theorem 1, the* $g(\theta)$ *specified in Theorem 1 maximizes the Bayes risk among all probability distributions in the form of* $\hat{g}(\theta)=\hat{\lambda}\omega(\theta)+(1-\hat{\lambda})\hat{v}(\theta)$*, where* $\hat{\lambda}\geq\lambda$ *and* $\hat{v}(\theta)$ *is any valid pdf. Equivalently:*

$$\int_{\Lambda} g(\theta)\int_{\Gamma} \mathrm{C}[\phi^*(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}d\theta \geq \int_{\Lambda} \hat{g}(\theta)\int_{\Gamma} \mathrm{C}[\hat{\phi}^*(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}d\theta\,, \tag{7}$$

where ϕ^* *and* $\hat{\phi}^*$ *are the classical Bayes decision rules corresponding to* $g(\theta)$ *and* $\hat{g}(\theta)$*, respectively.*

The proof is presented in the Appendix. As pointed out in Theorem 2, since $g(\theta)$ is the least favorable distribution, this property of $g(\theta)$ can be utilized to search its explicit expression. From the definitions of $g(\theta)$ and $\hat{g}(\theta)$, λ is a special case of $\hat{\lambda}$ and only the case of $\hat{\lambda}=\lambda$ is concerned in practical applications.

For the rest of this section, we first present the solution to the classical Bayes decision rule, and then develop the algorithm for finding the $g(\theta)$ specified in Theorem 1 and the optimal decision rule for the restricted Bayesian optimization problem.

3.2. Classical Bayes Decision Rule

As discussed above, the classical Bayes decision rule ϕ^* corresponding to $g(\theta)$ can be expressed by:

$$\phi^* = \min_{\phi}\int_{\Lambda} g(\theta)\int_{\Gamma} \mathrm{C}[\phi(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}d\theta = \min_{\phi}\int_{\Lambda} g(\theta)R_\theta(\phi)d\theta\,, \tag{8}$$

where $g(\theta)=\lambda\omega(\theta)+(1-\lambda)v(\theta)$. Since $R_\theta(\phi)=E_\theta\{\mathrm{C}[\phi(\mathbf{x}),\theta]\}=E_\theta\{\mathrm{C}[\phi(\mathbf{x}),\hat{\Theta}]\,|\,\hat{\Theta}=\theta\}$, where $\hat{\Theta}\sim g(\theta)$, the modified Bayes risk $\hat{r}(\phi)$ is calculated by:

$$\hat{r}(\phi)=\int_{\Lambda} g(\theta)R_\theta(\phi)d\theta = E\{E_\theta\{\mathrm{C}[\phi(\mathbf{x}),\hat{\Theta}]\,|\,\hat{\Theta}=\theta\}\} = E\{\mathrm{C}[\phi(\mathbf{x}),\hat{\Theta}]\}\,, \tag{9}$$

where the last equality holds due to the use of the iterated expectation of $E\{B\}=E\{E\{B\,|\,A\}\}$. Therefore, $\hat{r}(\phi)$ can be simply viewed as the cost of using ϕ averaged over $\hat{\Theta}$ and $\mathbf{x}$. By using the iterated expectation again, $\hat{r}(\phi)$ is rewritten as:

$$\hat{r}(\phi)=E\{E\{\mathrm{C}[\phi(\mathbf{x}),\hat{\Theta}]\,|\,\mathbf{x}\}\}\,. \tag{10}$$

According to the Bayes lemma [11], the optimal decision rule to minimize $\hat{r}(\phi)$ is expressed by:

$$\phi^*(\mathbf{x})=\begin{cases} 1 & \text{if } E\{\mathrm{C}[1,\hat{\Theta}]\,|\,\mathbf{x}=x\}<E\{\mathrm{C}[0,\hat{\Theta}]\,|\,\mathbf{x}=x\} \\ 0 \text{ or } 1 & \text{if } E\{\mathrm{C}[1,\hat{\Theta}]\,|\,\mathbf{x}=x\}=E\{\mathrm{C}[0,\hat{\Theta}]\,|\,\mathbf{x}=x\} \\ 0 & \text{if } E\{\mathrm{C}[1,\hat{\Theta}]\,|\,\mathbf{x}=x\}>E\{\mathrm{C}[0,\hat{\Theta}]\,|\,\mathbf{x}=x\} \end{cases}. \tag{11}$$

In many cases, the costs over sets Λ_0 and Λ_1 are uniform, i.e.:

$$\mathrm{C}[i,\hat{\Theta}]=\mathrm{C}_{ij},\hat{\Theta}\in\Lambda_j\,. \tag{12}$$

Usually, the cost of a right decision is less than that of a wrong one, i.e., $\mathrm{C}_{11}<\mathrm{C}_{01}$ and $\mathrm{C}_{00}<\mathrm{C}_{10}$. Then (11) is simplified as:

$$\phi^*(\mathbf{x})=\begin{cases} 1 & \text{if } L(\mathbf{x}) > l_0 \\ 0 \text{ or } 1 & \text{if } L(\mathbf{x}) = l_0 \\ 0 & \text{if } L(\mathbf{x}) < l_0 \end{cases}, \tag{13}$$

where $L(\mathbf{x}) = P(\hat{\Theta} \in \Lambda_1 \mid \mathbf{x} = x) / P(\hat{\Theta} \in \Lambda_0 \mid \mathbf{x} = x)$, $l_0 = (\mathrm{C}_{10} - \mathrm{C}_{00}) / (\mathrm{C}_{01} - \mathrm{C}_{11})$, and $P(\hat{\Theta} \in \Lambda_j \mid \mathbf{x} = x)$ represents the conditional probability for $\hat{\Theta} \in \Lambda_j$ given that $\mathbf{x} = x$. Based on the total probability formula, $P(\hat{\Theta} \in \Lambda_j \mid \mathbf{x} = x)$ is calculated by:

$$P(\hat{\Theta} \in \Lambda_j \mid \mathbf{x} = x) = \frac{p(x \mid \hat{\Theta} \in \Lambda_j) P(\hat{\Theta} \in \Lambda_j)}{p(x)}, \tag{14}$$

where $p(x) = \sum_{j=0}^{1} p(x \mid \hat{\Theta} \in \Lambda_j) P(\hat{\Theta} \in \Lambda_j)$, and $P(\hat{\Theta} \in \Lambda_j) = \int_{\Lambda_j} g(\theta) d\theta = p(H_j)$ denotes the probability of $\hat{\Theta} \in \Lambda_j$, $j = 0,1$. Based on the notation introduced in [11], $p(x \mid \hat{\Theta} \in \Lambda_j)$ is given as:

$$p(x \mid \hat{\Theta} \in \Lambda_j) = \int_{\Lambda} p_\theta(x) g_j(\theta) d\theta, \tag{15}$$

where $g_j(\theta)$ denotes the pdf of $\hat{\Theta}$ given that $\hat{\Theta} \in \Lambda_j$, and it is:

$$g_j(\theta) = \begin{cases} 0 & \text{if } \theta \notin \Lambda_j \\ g(\theta) / p(\hat{\Theta} \in \Lambda_j) & \text{if } \theta \in \Lambda_j \end{cases}. \tag{16}$$

Accordingly, Equation (13) is rewritten as:

$$\phi^*(\mathbf{x}) = \begin{cases} 1 & \text{if } \int_{\Lambda} p_\theta(\mathbf{x}) g_1(\theta) d\theta > \gamma \int_{\Lambda} p_\theta(\mathbf{x}) g_0(\theta) d\theta \\ 0 \text{ or } 1 & \text{if } \int_{\Lambda} p_\theta(\mathbf{x}) g_1(\theta) d\theta = \gamma \int_{\Lambda} p_\theta(\mathbf{x}) g_0(\theta) d\theta \\ 0 & \text{if } \int_{\Lambda} p_\theta(\mathbf{x}) g_1(\theta) d\theta < \gamma \int_{\Lambda} p_\theta(\mathbf{x}) g_0(\theta) d\theta \end{cases}, \tag{17}$$

where $\gamma = (p(H_0)(\mathrm{C}_{10} - \mathrm{C}_{00})) / (p(H_1)(\mathrm{C}_{01} - \mathrm{C}_{11}))$. As a result, the classical Bayes decision rule is specified by the observation densities of (15), which depends on the probability density of the observation $\mathbf{x}$ and the conditional densities $g_j(\theta)$, and the decision threshold is determined by the prior probabilities $p(H_j)$ and the costs over sets Λ_0 and Λ_1.

3.3. Algorithm for Finding the Restricted Bayes Decision Rule

Based on the analysis in Section 3.1, the function of $g(\theta)$ and the corresponding classical Bayes decision rule ϕ^* specified in Theorem 1 are required in order to solve the constrained optimization problem in (2) and (3). It also implies that the explicit expression of $v(\theta)$ and the value of λ should be determined in advance. To achieve this, the condition in Theorem 1 can be expressed as:

$$\int_{\Lambda} R_\theta(\phi^*) v(\theta) d\theta = \max_{\theta \in \Lambda} R_\theta(\phi^*), \tag{18}$$

with $R_\theta(\phi^*) = \int_{\Gamma} \mathrm{C}[\phi^*(\mathbf{x}), \theta] p_\theta(\mathbf{x}) d\mathbf{x}$. It is obvious that $v(\theta)$ assigns non-zero probabilities only for the values of θ that correspond to the global maximum of $R_\theta(\phi^*)$ from Equation (18). In fact, there could be unique, multiple or infinite number of maximizers that achieve global maximum. Therefore, in this work, the solution of the optimization problem in Equations (2) and (3) is discussed under three cases according to the number of the maximizers in the following subsection.

First, assume that $R_\theta(\phi^*)$ has a unique maximizer that achieves the global maximum, then $v(\theta)$ can be expressed by:

$$v(\theta) = \delta(\theta - \theta_1), \tag{19}$$

where θ_1 denotes the unique value of θ corresponding to the maximum $R_\theta(\phi^*)$ and δ denotes the Dirac delta function, i.e., $\delta(x) = \begin{cases} \infty, & x = 0 \\ 0, & x \neq 0 \end{cases}$ and $\int \delta(x) dx = 1$. If ϕ^* satisfies the conditions in

Theorem 1 and Proposition 1, ϕ^* is the solution of the restricted Bayesian optimization problem in (2) and (3). Based on this assumption, Algorithm 1 is developed to find the restricted Bayes decision rule ϕ^* and θ_1.

Algorithm 1. Restricted Bayes decision rule.

1: Initialize $\lambda_{\inf} = 0$, $\lambda_{\sup} = 1$ and $\lambda = (\lambda_{\inf} + \lambda_{\sup})/2$.

2: Obtain $R_\theta(\phi^*_{\theta_1}) = \int_\Gamma \mathrm{C}[\phi^*_{\theta_1}(\mathbf{x}), \theta] p_\theta(\mathbf{x}) d\mathbf{x}$ for all $\theta_1 \in \Lambda$, where $\phi^*_{\theta_1}$ denotes the classical Bayes decision rule corresponding to $g(\theta) = \lambda\omega(\theta) + (1-\lambda)\delta(\theta - \theta_1)$ as determined in Section 3.2.

3: Calculate:

$$\theta_1^* = \arg\max_{\theta_1 \in \Lambda} \hbar(\theta_1), \tag{20}$$

where $\hbar(\theta_1) = \lambda \int_\Lambda R_\theta(\phi^*_{\theta_1})\omega(\theta) d\theta + (1-\lambda) R_{\theta_1}(\phi^*_{\theta_1})$ and $R_{\theta_1}(\phi^*_{\theta_1}) = \int_\Gamma \mathrm{C}[\phi^*_{\theta_1}(\mathbf{x}), \theta_1] p_{\theta_1}(\mathbf{x}) d\mathbf{x}$.

4: If $R_{\theta_1^*}(\phi^*_{\theta_1^*}) \neq \max_{\theta \in \Lambda} R_\theta(\phi^*_{\theta_1^*})$, terminate this algorithm, reset $v(\theta)$ and restart the algorithm. When $R_{\theta_1^*}(\phi^*_{\theta_1^*}) = \max_{\theta \in \Lambda} R_\theta(\phi^*_{\theta_1^*})$, if the difference between $\max_{\theta \in \Lambda} R_\theta(\phi^*_{\theta_1^*})$ and α is less than the predefined precision value, $\phi^*_{\theta_1^*}$ is a solution of the restricted Bayes optimization problem in (2) and (3), and terminate the algorithm. Otherwise, continue to the next step.

5: If $\max_{\theta \in \Lambda} R_\theta(\phi^*_{\theta_1^*}) > \alpha$, λ and $\lambda_{\sup}$ are replaced with $(\lambda + \lambda_{\inf})/2$ and λ, respectively. If $\max_{\theta \in \Lambda} R_\theta(\phi^*_{\theta_1^*}) < \alpha$, λ and $\lambda_{\inf}$ are replaced with $(\lambda + \lambda_{\sup})/2$ and λ, respectively.

6: If the difference between $\lambda_{\sup}$ and $\lambda_{\inf}$ is less than the predefined precision value, terminate the algorithm and reset $v(\theta)$ to restart the algorithm; else go to Step 2.

In Algorithm 1, Step 3 follows from Theorem 2 and Step 4 is developed according to Theorem 1 and Proposition 1. In Step 3, $\hbar(\theta_1)$ can be treated as the minimum Bayes risk without any constraints on conditional risks corresponding to $g(\theta) = \lambda\omega(\theta) + (1-\lambda)\delta(\theta - \theta_1)$. It should be noted that the value of λ is related to α through Proposition 1, and α increases with the increase of λ. Hence, the constraint α can be achieved by adjusting the value of λ. For any θ_1^* that satisfies (20), if the corresponding classical Bayes decision rule $\phi^*_{\theta_1^*}$ satisfies $R_{\theta_1^*}(\phi^*_{\theta_1^*}) = \max_{\theta \in \Lambda} R_\theta(\phi^*_{\theta_1^*})$ and the difference between $R_{\theta_1^*}(\phi^*_{\theta_1^*})$ and α is less than the predefined precision value, the θ_1^* is the maximizer that achieves the global maximum of $R_\theta(\phi^*)$ and $\phi^*_{\theta_1^*}$ can be viewed as the solution of the restricted Bayes optimization problem in (2) and (3). In addition, since the classical Bayes decision rule can be obtained based on Section 3.2 for any valid $g(\theta)$, the solution of the restricted optimization problem in (2) and (3) exists as long as α is properly defined. The effective interval of α will be discussed in Section 4.

Second, suppose that there are multiple values of θ corresponding to the global maximum of $R_\theta(\phi^*)$, $v(\theta)$ should be expressed as:

$$v(\theta) = \sum_{l=1}^{L} \eta_l \delta(\theta - \theta_l), \tag{21}$$

where $\eta_l \geq 1$, $\sum_{l=1}^{L} \eta_l = 1$, L and θ_l denote the number and the value of θ that achieve the maximum $R_\theta(\phi^*)$, respectively. For the convenience of analysis, let ℓ denote the vector of all unknown components of $v(\theta)$, i.e.:

$$\ell = [(\eta_1, \theta_1), \ldots, (\eta_L, \theta_L)]. \tag{22}$$

In order to solve the restricted Bayes optimization problem in (2) and (3), the Step 3 should be replaced by:

$$\ell^* = \arg\max_{\ell} \hbar(\ell) = \arg\max_{\ell} \left\{ \lambda \int_{\Lambda} R_\theta(\phi_\ell^*)\omega(\theta)d\theta + (1-\lambda)\sum_{l=1}^{L} \eta_l R_{\theta_l}(\phi_\ell^*) \right\}, \tag{23}$$

where ϕ_ℓ^* represents the classical Bayes decision rule for the modified prior distribution $g(\theta) = \lambda\omega(\theta) + (1-\lambda)\sum_{l=1}^{L}\eta_l\delta(\theta-\theta_l)$. Correspondingly, the condition in Step (4) should be updated by replacing $R_{\theta_1^*}(\phi_{\theta_1^*}^*)$ and $\max_{\theta\in\Lambda} R_\theta(\phi_{\theta_1^*}^*)$ with $R_{\ell^*}(\phi_{\ell^*}^*)$ and $\max_{\ell} R_\ell(\phi_{\ell^*}^*)$, respectively. Obviously, compared with the case where $R_\theta(\phi^*)$ has only one maximizer, the computational complexity of multiple maximizers increases significantly. In such case, the global optimization algorithms such as the particle swarm optimization algorithm (PSO) and the ant colony algorithm (ACO) can be utilized to find ℓ^* [26–28].

Finally, if there are infinitely many values of θ that correspond to the global maximum of $R_\theta(\phi^*)$, it is difficult to obtain an accurate solution of the optimization problem in (2) and (3). In this case, an approximate solution can be obtained by employing the Parzen window density estimation to approximate the form of $v(\theta)$, where $v(\theta)$ is expressed approximately by a convex combination of many window functions, given by:

$$v(\theta) = \sum_{l=1}^{W} \tau_l \varphi(\theta-\theta_l), \tag{24}$$

where $\tau_l \geq 1$, $\sum_{l=1}^{W}\tau_l = 1$, and φ denotes a window function which could be a Rectangular, Gauss, or Cosine window function. In order to solve the optimization problem in (2) and (3), Algorithm 1 should be updated by replacing ℓ and ϕ_ℓ^* with $\ell = [(\tau_1,\theta_1),\ldots,(\tau_W,\theta_W)]$ and the classical Bayes decision rule corresponding to $g(\theta) = \lambda\omega(\theta) + (1-\lambda)\sum_{l=1}^{W}\tau_l\varphi(\theta-\theta_l)$, respectively.

In real applications, it may be difficult to know all the exact values of θ that achieve the global maximum of $R_\theta(\phi^*)$ in advance. The practical way is that we start with the assumption that there is a unique θ that achieves the global maximum. Based on the assumption and Algorithm 1, θ_1^* and the corresponding $\phi_{\theta_1^*}^*$ is obtained. If θ_1^* and $\phi_{\theta_1^*}^*$ satisfy the condition in Step 3, the restricted Bayes optimization problem in (2) and (3) is solved. Otherwise, the number of θ which are assumed to achieve the global maximum of $R_\theta(\phi^*)$ is increased incrementally until the ϕ_ℓ^* calculated through the updated Algorithm 1 is verified as the solution of the restricted Bayes optimization problem.

4. Relationship between Bayes Risk and the Constraint

In this section, the relationship between Bayes risk and the constraint on the conditional risks is explored. As the analysis in introduction, the constraint on the conditional risks is introduced from the uncertainty of $\omega(\theta)$. Generally speaking, the classical Bayesian criterion is considered to minimize the average Bayes risk when $\omega(\theta)$ is completely known, and the Minimax decision rule can be utilized to minimize the worst-case (maximum) conditional Bayes risk if the prior information is full of uncertainty. Therefore, the classical Bayesian and Minimax approaches can be viewed as two extreme cases of the restricted Bayesian approaches.

For the notational simplicity, here we respectively denote the classical Bayes and the Minimax decision rules by ϕ_b^o and ϕ_m^o, i.e.:

$$\phi_b^o = \arg\min_{\phi} \int_{\Lambda} R_\theta(\phi)\omega(\theta)d\theta, \tag{25}$$

$$\phi_m^o = \arg\min_{\phi} \max_{\theta\in\Lambda} R_\theta(\phi) . \tag{26}$$

Then define $B = \max_{\theta\in\Lambda} R_\theta(\phi_b^o)$ and $M = \max_{\theta\in\Lambda} R_\theta(\phi_m^o)$ as the maximum conditional risks achieved by the classical Bayesian and Minimax decision rules, respectively. In addition, the restricted Bayes decision rule where $\max_{\theta\in\Lambda} R_\theta(\phi) \le \alpha$ is denoted by ϕ_r^α. From the definitions of ϕ_b^o, ϕ_m^o and ϕ_r^α, it should be noted that $r(\phi_b^o) \le r(\phi_r^\alpha) \le r(\phi_m^o)$ and $M \le \max_{\theta\in\Lambda} R_\theta(\phi_r^\alpha) \le B$.

According to the definition of the Minimax criterion, M is the achievable minimum of the maximum conditional risk, which means that there is no solution for the optimization problem in (2) and (3) if the value of α is less than M. On the other hand, it is obvious that the constraint on the conditional risks enforced by α is ineffective if α is greater than B. In other words, the restricted Bayes optimization problem in (2) and (3) is equivalent to the Bayes optimization problem in (25) for any $\alpha \ge B$. Namely $\phi_r^\alpha = \phi_b^o$ and $r(\phi_b^o) = r(\phi_r^\alpha)$ if $\alpha \ge B$. As a result, α should be defined in the interval of $[M, B]$; otherwise the constraint is ineffective or meaningless.

For the restricted Bayes optimization problem formulated in (2) and (3), the maximum conditional risk $\max_{\theta\in\Lambda} R_\theta(\phi_r^\alpha)$ and the Bayes risk $r(\phi_r^\alpha)$ are closely related to the value of the constraint α. When $\alpha \in [M, B]$, the following conclusion holds that:

$$\max_{\theta\in\Lambda} R_\theta(\phi_r^\alpha) = \alpha . \tag{27}$$

A simple contradiction method can be used here to show this conclusion. Suppose that $\max_{\theta\in\Lambda} R_\theta(\phi_r^\alpha) < \alpha$, and let $\phi = \xi\phi_r^\alpha + (1-\xi)\phi_b^o$, where $0 \le \xi \le 1$ and ϕ_b^o denotes a classical Bayes decision rule obtained in (25), $\max_{\theta\in\Lambda} R_\theta(\phi_b^o) = B > \alpha$ and $r(\phi_b^o) < r(\phi_r^\alpha)$. There must exist a ξ such that $\max_{\theta\in\Lambda} R_\theta(\phi) = \alpha$ and $r(\phi) = \xi r(\phi_r^\alpha) + (1-\xi) r(\phi_b^o) < r(\phi_r^\alpha)$. Obviously, it contradicts with the definition of ϕ_r^α. Therefore, the conclusion in (27) is true.

Based on the discussions above, the relationship between the restricted Bayesian decision rule $r(\phi_r^\alpha)$ and the value of the constraint on the conditional risks is presented by the Theorem 3.

Theorem 3. *The Bayes risk $r(\phi_r^\alpha)$ obtained by the restricted Bayes decision rule is a strictly decreasing and convex function of α for $\alpha \in [M, B]$.*

The proof is provided in the Appendix. Theorem 3 implies the relationship between Bayes risk and the constraint on conditional risks. Moreover, the value of α is closely related to the uncertainty of $\omega(\theta)$. In general, a smaller value of α should be specified for a greater degree of the uncertainty. Therefore, based on these characteristics, one can predefine a suitable value of α to obtain the expected Bayes risk in practice. In addition, the restricted optimization problem in (2) and (3) is analyzed and solved in Section 3 from a special perspective that its solution is identical to a classical Bayes decision rule with a modified prior distribution as presented in Theorem 1 and Proposition 1. For any fixed prior distribution, the corresponding classical Bayes decision rule can be obtained from Section 3.2, thereby the solution of the restricted optimization problem in (2) and (3) exists as long as α is properly defined in $[M, B]$ and can be obtained further by the algorithm in Section 3.3.

5. Numerical Results

In this section, a binary hypothesis-testing problem is studied to illustrate theoretical results. The hypotheses are defined as:

$$\begin{cases} H_0 : x = \Theta + n, \Theta \in \Lambda_0 \\ H_1 : x = \Theta + n, \Theta \in \Lambda_1 \end{cases}, \tag{28}$$

where $x \in \mathbb{R}$, Λ_0 and Λ_1 denote the sets of all possible values of parameter Θ under H_0 and H_1, respectively, which are specified as $\Lambda_0 = \{0\}$ and $\Lambda_1 = \{A, -A\}$, where A is a known constant. In addition, n denotes a zero mean noise that is a mixture of Rayleigh distributed components; that is, $p_n(n) = \sum_{i=1}^{N} m_i \varphi_i(n - \mu_i)$, where $m_i \geq 0$ for $i = 1, \ldots, N$, $\sum_{i=1}^{N} m_i = 1$, and:

$$\varphi(n) = \begin{cases} \frac{n}{\sigma_i^2} \exp(-\frac{n^2}{2\sigma_i^2}), n \geq 0 \\ 0, \qquad\qquad n < 0 \end{cases}. \tag{29}$$

In the numerical results, the same variance is assumed, i.e., $\sigma_i = \sigma$ for $i = 1, \ldots, N$. Furthermore, the parameters are specified as $N = 4$, $\mu_1 = 0.2$, $\mu_2 = 1$, $\mu_3 = -2\sigma\sqrt{\frac{\pi}{2}} - 0.2$, $\mu_4 = -2\sigma\sqrt{\frac{\pi}{2}} - 1$, and $m_i = 0.25$ for $i = 1, \ldots, 4$. Then the conditional pdf of x for a given value of $\Theta = \theta$ can be expressed by:

$$p_\theta(x) = \sum_{i=1}^{N} m_i \varphi_i(x - \mu_i - \theta). \tag{30}$$

With the assumption, the detection problem in (28) is equivalent to the detection of a signal that employs binary phase shift keying (BPSK). Accordingly, the probability distribution of Θ under H_0 is directly expressed as $\omega_0(\theta) = \delta(\theta)$. On the other hand, the probability distribution of Θ under H_1 can be modeled as:

$$\omega_1(\theta) = \rho\delta(\theta - A) + (1 - \rho)\delta(\theta + A), \tag{31}$$

where ρ is estimated based on the previous experiences, and it is known with some uncertainty due to the presence of estimation errors. As a result, the probability distribution of Θ can be determined by:

$$\omega(\theta) = \sum_{i=0}^{1} p(H_i)\omega_i(\theta), \tag{32}$$

where $p(H_i) \geq 0$, $\sum_{i=0}^{1} p(H_i) = 1$ and $p(H_i)$ denotes the prior probability of H_i for $i = 0, 1$.

In order to obtain the optimal restricted Bayesian decision rule, the prior distribution $\omega(\theta)$ should be modified via a pdf $v(\theta)$ according to Theorem 1. Since $\Lambda = \Lambda_0 \cup \Lambda_1 = \{0, A, -A\}$, the complete form of $v(\theta)$ is expressed as:

$$v(\theta) = \sum_{i=1}^{3} t_i \delta(\theta - \theta_i), \tag{33}$$

where t_1, t_2 and t_3 are the respective weights assigned for $\theta_1 = 0$, $\theta_2 = A$ and $\theta_2 = A$, $t_i \geq 0$ for $i = 1, \ldots, 3$ and $\sum_{i=1}^{3} t_i = 1$.

First, it is assumed that any two of the three assigned weights are zero, i.e., $v(\theta) = \delta(\theta - \theta_i)$. By employing the algorithm based on (20), the optimal decision rule is obtained if the condition in the third step is satisfied. Otherwise, we assume that only one of the three assigned weights is zero, i.e., $v(\theta) = t\delta(\theta - \theta_i) + (1 - t)\delta(\theta - \theta_j)$, where $t = t_i$, $\theta_i, \theta_j \in \Lambda$, and $\theta_i \neq \theta_j$. In such a case, as t is the only one unknown parameter in $v(\theta)$, the algorithm based on (21) and (22) is used to find the value of t and the corresponding $g(\theta) = \lambda\omega(\theta) + (1 - \lambda)v(\theta)$ that maximizes the Bayes risk. If the corresponding decision rule satisfies the condition in (23), it is the optimal restricted Bayesian rule that we seek. Finally, if the condition is still not met in the second case, the $v(\theta)$ should be determined as $v(\theta) = t_1\delta(\theta) + t_2\delta(\theta - A) + t_3\delta(\theta + A)$ and $t_i \neq 0$ for $i = 1, \ldots, 3$. If we find the values of t_1 and t_2 that maximize the Bayes risk corresponding to $g(\theta)$, the optimal restricted Bayesian rule is also obtained.

In Figure 1, the achievable minimum Bayes risks obtained by the restricted Bayesian decision rule for different values of ρ are plotted against α, where $p(H_0)=0.45$, $A=2$ and $\sigma=0.5$. As analyzed in Section 4, when α is equal to the Bayes risk of the Minimax decision rule, the restricted Bayesian decision rule is identical to the Minimax decision rule. When α is smaller than the Bayes risk of the Minimax decision rule, there is no restricted Bayesian decision rule that satisfies the limit established by α. On the other hand, the detection performance of the restricted Bayesian decision rule is the same with that of the classical Bayesian decision rule when α is greater than and equal to the maximum conditional risk of the classical Bayesian decision rule. As expected, the lowest Bayes risks are obtained by the classical Bayesian decision rule, but it also leads to the highest maximum conditional risk that are 0.3456, 0.3657, 0.3816, 0.4131 and 0.4325 for $\rho=0.6$, 0.65, 0.686, 0.75 and 0.8, respectively. On the contrary, the Minimax decision rule achieves the lowest maximum conditional risk, but it produces the worst Bayes risk. It should be noted that the conditional risks of the Minimax decision rule are equal to each other, which are 0.2381.

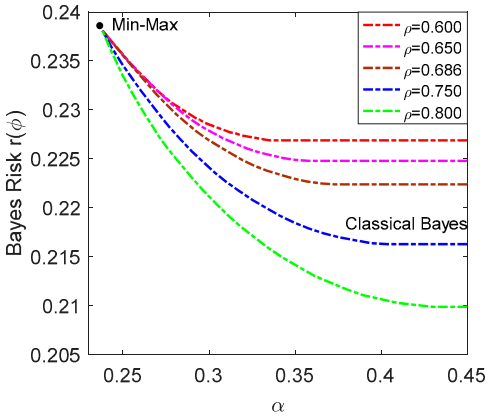

Figure 1. Bayes risk versus α of the restricted Bayesian decision rule for $\rho=0.6$, 0.65, 0.686, 0.75 and 0.8, where $p(H_0)=0.45$, $A=2$ and $\sigma=0.5$.

For the restricted Bayesian decision rule, the maximum conditional risk is equal to α except for the linear part that corresponds to the classical Bayesian decision rule. In fact, the restricted Bayes decision rule makes a tradeoff between the maximum conditional risk and Bayes risk, and generalizes the Minimax and the classical Bayesian decision rules. It is also shown from the figure that with the increase of ρ, the performance difference of the classical Bayesian and the Minimax decision rules increases for $\rho \geq 0.5518$. Specifically, the maximum conditional risk of the classical Bayesian decision rule increases and the corresponding Bayes risk decreases as ρ increases. Furthermore, when α is greater than the maximum conditional risk of the Minimax decision rule and smaller than that of the classical Bayesian decision rule, the Bayes risk is a strictly decreasing and convex function of α. This agrees with the conclusion made in Theorem 3. In addition, Figure 1 acts a guideline for the design of α in practice by observing the corresponding Bayes risk for each α. Hence, instead of assigning a value for α arbitrarily, Figure 1 can be utilized to choose a more appropriate α in practical problems.

Figure 2 compares the achievable lowest maximum conditional risk obtained by the restricted Bayes decision rule and the corresponding Bayes risk versus λ in (5) for $\rho=0.6$, 0.686 and 0.8, where $p(H_0)=0.45$, $A=2$ and $\sigma=0.5$. The maximum conditional risk is equal to the Bayes risk for $0 \leq \lambda \leq 0.7036$ when $\rho=0.8$, for $0 \leq \lambda \leq 0.8207$ when $\rho=0.6$ and $\rho=0.686$. This shows that the restricted Bayesian decision rules becomes identical to the Minimax decision rule when λ reduces to less than or equal to a certain value. For convenience, this value is denoted by λ_m, and it is seen that $\lambda_m=0.7036$, 0.8207 and 0.8207 for $\rho=0.6$, 0.686 and 0.8, respectively. In fact, the value of λ_m increases from 0.7035 to 0.8207 when ρ decreases from 0.8 to 0.686 and maintains at 0.8207 when $\rho \in [0.6, 0.686]$. In addition, when $\lambda \geq \lambda_m$, the achievable lowest maximum conditional risk

increases and the corresponding Bayes risk decreases with the increase of λ. In order to further illustrate the results in Figure 2, Tables 1–3 are provided to show the parameters of t_1, t_2 and t_3 in $v(\theta)$, the maximum conditional and Bayes risks of the restricted Bayesian decision rule for different values of λ for $\rho = 0.6$, 0.686 and 0.8.

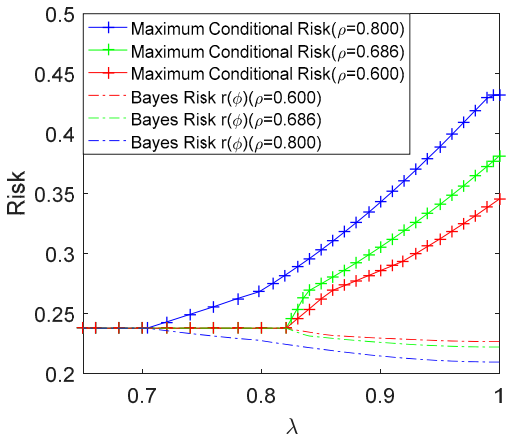

Figure 2. The maximum conditional and Bayes risks of the restricted Bayesian decision rule versus λ for $\rho = 0.6$, 0.686 and 0.8, where $p(H_0) = 0.45$, $A = 2$ and $\sigma = 0.5$.

Table 1. The parameters of $v(\theta)$ and different risks for various λ when $\rho = 0.6$.

λ	$t_1 / t_2 / t_3$	$R_0 / R_A / R_{-A}$	$\max_{\theta\in\Lambda} R_\theta$	$r(\phi^*)$
1.0000	--/--/--	0.1188/0.2951/0.3456	0.3456	0.2269
0.9250	0/0/1	0.1474/0.2938/0.2974	0.2974	0.2287
0.8207	0/0.2162/0.7838	0.2381/0.2381/0.2381	0.2381	0.2381
0.7036	0.1777/0.2612/0.5611	0.2381/0.2381/0.2381	0.2381	0.2381
0.6500	0.2194/0.2717/0.5089	0.2381/0.2381/0.2381	0.2381	0.2381

Table 2. The parameters of $v(\theta)$ and different risks for various λ when $\rho = 0.686$.

λ	$t_1 / t_2 / t_3$	$R_0 / R_A / R_{-A}$	$\max_{\theta\in\Lambda} R_\theta$	$r(\phi^*)$
1.0000	--/--/--	0.1340/0.2632/0.3657	0.3456	0.2224
0.9250	0/0/1	0.1739/0.2402/0.3227	0.2974	0.2246
0.8207	0/0/1	0.2381/0.2381/0.2381	0.2381	0.2381
0.7036	0.1777/0.1489/0.6734	0.2381/0.2381/0.2381	0.2381	0.2381
0.6500	0.2194/0.1839/0.5967	0.2381/0.2381/0.2381	0.2381	0.2381

Table 3. The parameters of $v(\theta)$ and different risks for various λ when $\rho = 0.8$.

λ	$t_1 / t_2 / t_3$	$R_0 / R_A / R_{-A}$	$\max_{\theta\in\Lambda} R_\theta$	$r(\phi^*)$
1.0000	--/--/--	0.1839/0.1808/0.4325	0.4325	0.2099
0.9250	0/0/1	0.1964/0.1763/0.4049	0.4049	0.2105
0.8207	0/0/1	0.2212/0.1692/0.3583	0.3583	0.2134
0.7036	0.1777/0/0.8223	0.2381/0.2381/0.2381	0.2381	0.2381
0.6500	0.2194/0.0674/0.7131	0.2381/0.2381/0.2381	0.2381	0.2381

From the Tables 1–3, it is observed that the maximum conditional risk is always equal to the conditional risk for $\Theta = -A$. Actually, the maximum conditional risk can also be viewed as the achievable minimum α. Since $\Theta = 0$ is independent on ρ, the weight t_1 of $\Theta = 0$ in $v(\theta)$ is same for each λ when $\rho = 0.6$, 0.686 and 0.8. Conversely, parameters of t_2 and t_3 change with ρ. In the simulation results, the distribution of Θ based on the Minimax criterion can be calculated as $\omega_m(\theta) = 0.3693\delta(\theta) + 0.3096\delta(\theta - A) + 0.3211\delta(\theta + A)$. In order to achieve the same

performance with the Minimax decision rule, the $\omega(\theta)$ in (31) should be modified by making $g(\theta)$ equal to $\omega_m(\theta)$, i.e., $g(\theta)=\lambda\omega(\theta)+(1-\lambda)v(\theta)=\omega_m(\theta)$. For instance, as listed in Tables 1–3, $v(\theta)=0.2162\delta(\theta-A)+0.7838\delta(\theta+A)$ and $\lambda=0.8207$ for $\rho=0.6$, $v(\theta)=\delta(\theta+A)$ and $\lambda=0.8207$ for $\rho=0.686$, $v(\theta)=0.1777\delta(\theta)+0.8223\delta(\theta+A)$ and $\lambda=0.7036$ for $\rho=0.8$. Therefore, Figure 2 and Tables 1–3 also provide a guideline to select a suitable value of λ.

Figure 3a plots the achievable minimum α obtained via the restricted Bayesian decision rule versus σ for $\lambda=1$, 0.8 and 0.6 when $p(H_0)=0.45$ and $A=2$, and the corresponding minimum Bayes risks are presented in Figure 3b. In general, with the increase of σ, the achievable minimum α and the Bayes risk increase first, and then decrease to troughs, and increase gradually again, all regardless of the values of λ. In fact, when $\lambda=1$, the restricted Bayesian decision rule is identical to the classical Bayesian decision rule. In addition, the maximum conditional risk is equal to α for any λ, which agrees with the conclusion in (27). As expected, the classical Bayesian decision rule achieves the highest maximum conditional risk and the minimum Bayes risk. It is also seen that the maximum conditional risk decreases and the Bayes risk increases with the decrease of λ. In this case, when $\lambda=0.6$, it is the same with that of the Minimax decision rule when $\sigma\in(0.26,0.61)$, which implies $\lambda_m\geq 0.6$ for $\sigma\in(0.26,0.61)$. Similarly, the cases for $\lambda\in(0.6,\lambda_m)$ are the same to that when $\sigma\in(0.26,0.61)$. From Figure 3a,b, compared with the performance of the classical Bayesian decision rules, the maximum conditional risks obtained via the restricted Bayesian decision rules decrease significantly with much lower increase of the corresponding Bayes risks, especially for $\sigma>0.92$ and $\lambda=0.8$. To clearly investigate the results in Figure 3, the restricted Bayes decision rules for different cases are presented. It should be noted beforehand that the form of the optimal decision rule can be determined as follows:

$$\phi^*=\begin{cases}1, x\in\Gamma_1\\ 0, otherwise\end{cases}, \tag{34}$$

where Γ_1 represents a part of the sample space. The restricted Bayes decision rules, the constraint α and the Bayes risks for different values of σ are provided in Tables 4–6 when $\lambda=1$, 0.8 and 0.6.

Table 4. The restricted Bayes decision rules, the achievable minimum α and the corresponding Bayes risk for different σ when λ = 1.

σ	Γ_1	α	$r(\phi)$
0.10	(−∞,−1.2508) ∪ (−0.9744,−0.4508) ∪ (0.7492,1.0086) ∪ (1.5492, ∞)	0.0087	0.0076
0.32	(−∞,−1.8020) ∪ (0.1986,0.2724) ∪ (0.9990,1.0486) ∪ (2.2000, ∞)	0.5001	0.2749
0.90	(−∞,−3.1154) ∪ (−0.9326,0.9838) ∪ (2.6262, ∞)	0.2147	0.1700
2.00	(−∞,−5.4212) ∪ (−2.5198,1.1780) ∪ (3.6814, ∞)	0.4523	0.2864

Table 5. The restricted Bayes decision rules, the achievable minimum α and the corresponding Bayes risk for different σ when λ = 0.8.

σ	Γ_1	α	$r(\phi)$
0.10	(−∞,−1.2508) ∪ (−0.9776,−0.4508) ∪ (0.7492,1.0122) ∪ (1.5492, ∞)	0.0079	0.0077
0.32	(−∞,−1.8014) ∪ (0.1980,0.2832) ∪ (0.9990,1.0486) ∪ (2.2000, ∞)	0.5001	0.2750
0.90	(−∞,−3.0248) ∪ (−0.9674,0.9198) ∪ (2.6970, ∞)	0.1975	0.1715
2.00	(−∞,−4.9886) ∪ (−2.4148,1.4090) ∪ (3.7878, ∞)	0.3661	0.2916

Table 6. The restricted Bayes decision rules, the achievable minimum α and the corresponding Bayes risk for different σ when λ = 0.6.

σ	Γ_1	α	$r(\phi)$
0.10	(−∞,−1.2508) ∪ (−0.9776,−0.4508) ∪ (0.7492,1.0122) ∪ (1.5492, ∞)	0.0079	0.0077
0.32	(−∞,−1.8012) ∪ (−1.3404,−1.0020) ∪ (−0.5420,0.4810) ∪ (0.9984,1.2714) ∪ (2.2000, ∞)	0.3323	0.3323
0.90	(−∞,−2.9516) ∪ (−0.9956,0.8524) ∪ (2.8002, ∞)	0.1867	0.1762
2.00	(−∞,−4.7802) ∪ (−2.1130,1.5476) ∪ (4.0936, ∞)	0.3337	0.3059

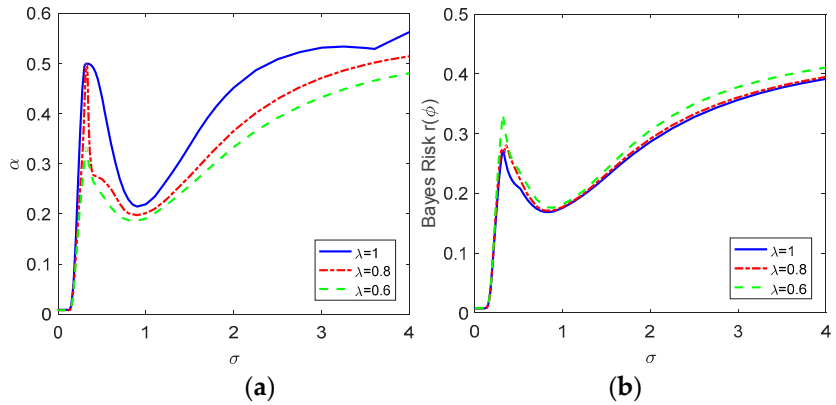

Figure 3. The achievable minimum α and the corresponding minimum Bayes risk as function of σ for $\lambda = 1$, 0.8 and 0.6 in (**a**,**b**), respectively, when $p(H_0) = 0.45$ and $A = 2$.

In Figure 4a, the achievable minimum values of α are plotted versus A for $\lambda = 1$, 0.8 and 0.6 when $p(H_0) = 0.45$ and $\sigma = 0.5$, where the corresponding minimum Bayes risks are illustrated in Figure 4b. Similarly, the decision rule for $\lambda = 1$ is identical to the classical Bayesian decision rule, and the maximum conditional risk is equal to α. When A is small such as $A \in (0, 0.2)$, the maximum conditional risk decreases significantly with a small increase of the Bayes risk via the restricted decision rule for both $\lambda = 0.8$ and 0.6. In general, a lower maximum conditional risk and a higher Bayes risk can be obtained for a smaller value of λ, but it is not true if $\lambda \in (\lambda, \lambda_m)$ according to the results shown in Figure 2. Especially, the Bayes risk obtained via the restricted Bayes decision rule when $\lambda = 0.8$ is almost equal to that obtained via the classical Bayes decision rule for some values of A, while it decreases significantly when $\lambda = 0.6$. The difference between $\lambda = 0.8$ and 0.6 is especially noticeable when $A = 2.45$. In addition, when $A > 2.45$, α and the Bayes risk decrease gradually to zero as A increases.

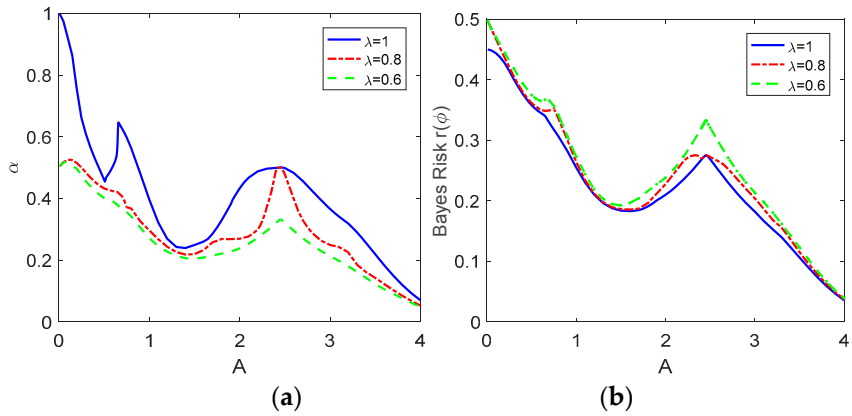

Figure 4. The achievable minimum α and the corresponding minimum Bayes risk versus A for $\lambda = 1$, 0.8 and 0.6 in (**a**,**b**), respectively when $p(H_0) = 0.45$ and $\sigma = 0.5$.

6. Conclusions

In this paper, the restricted Bayes decision rule is developed for a binary composite hypothesis-testing problem with a partially known prior distribution. Generally, the prior distribution is estimated based on previous information and is not completely accurate due to the estimate errors. In order to utilize the useful prior information, the Bayes risk is calculated based on the estimated prior distribution. On the other hand, a constraint on the maximum conditional risk is applied to alleviate the negative impact caused by the mismatch between the estimated prior distribution and the correct one. By applying the Lagrange duality, the restricted Bayesian optimization problem is transformed to an unconstrained optimization problem. Based on the

transformation, Theorems are derived to provide theoretical supports for the algorithms to find the optimal restricted Bayes decision rule. Specially, an additional conclusion is made that the Bayes risk obtained via the restricted Bayes decision rule is a strictly decreasing and convex function of the constraint on the maximum conditional risk. In addition, the constraint should be defined in an appropriate interval, otherwise the constraint is not effective. Finally, the numerical examples are provided to demonstrate the performance of the proposed decision rule, which are consistent with theoretical results.

The classical Bayes and Minimax decision rules are usually used in two different scenarios from the restricted Bayesian optimization problem, but they can be realized as two special cases of the restricted Bayes decision rule by defining two proper values of the constraint. In fact, the classical Bayesian decision rule achieves the lowest Bayes risk but with the highest maximum conditional risk, whereas the Maximum decision rule obtains the lowest maximum conditional risk but with a worst Bayes risk. By adjusting the constraint, the restricted Bayesian decision rule provides the ability to keep the balance between the maximum conditional risk and the Bayes risk.

Acknowledgments: This research is partly supported by the Basic and Advanced Research Project in Chongqing (Grant No. cstc2016jcyjA0134, No. cstc2016jcyjA0043), graduate scientific research and innovation foundation of Chongqing, China (Grant No. CYB17041), the National Natural Science Foundation of China (Grant No. 61501072, No. 41404027, N0. 61571069, No. 61675036, No. 61471073) and the Fundamental Research Funds for the Central Universities (Grant No. 106112017CDJQJ168817, No. 106112017CDJQJ168819) supported by the Fundamental Research Funds for the Central Universities.

Author Contributions: Shujun Liu raised the idea of the framework to find a suitable decision rule for a binary composite hypothesis-testing problem. Ting Yang and Shujun Liu contributed to the drafting of the manuscript, interpretation of the results, some experimental design and checked the manuscript. Hongqing Liu contributed to analyze the optimal restricted Bayesian decision rule and develop the corresponding algorithm. Ting Yang contributed to the proofs of the theories developed in this paper. All authors have read and approved the final manuscript.

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

Appendix A

Proof of Theorem 1. It is obvious that:

$$\begin{aligned}
&\lambda\int_{\Lambda}\omega(\theta)\int_{\Gamma}\mathrm{C}[\phi(\mathbf{x}),\theta]p_{\theta}(\mathbf{x})d\mathbf{x}d\theta+(1-\lambda)\max_{\theta\in\Lambda}\int_{\Gamma}\mathrm{C}[\phi(\mathbf{x}),\theta]p_{\theta}(\mathbf{x})d\mathbf{x}\\
&\geq\lambda\int_{\Lambda}\omega(\theta)\int_{\Gamma}\mathrm{C}[\phi(\mathbf{x}),\theta]p_{\theta}(\mathbf{x})d\mathbf{x}d\theta+(1-\lambda)\int_{\Lambda}v(\theta)\int_{\Gamma}\mathrm{C}[\phi(\mathbf{x}),\theta]p_{\theta}(\mathbf{x})d\mathbf{x}d\theta\\
&=\int_{\Lambda}(\lambda\omega(\theta)+(1-\lambda)v(\theta))\int_{\Gamma}\mathrm{C}[\phi(\mathbf{x}),\theta]p_{\theta}(\mathbf{x})d\mathbf{x}d\theta\\
&=\int_{\Lambda}g(\theta)\int_{\Gamma}\mathrm{C}[\phi(\mathbf{x}),\theta]p_{\theta}(\mathbf{x})d\mathbf{x}d\theta
\end{aligned}, \tag{A1}$$

where the first inequality holds as $\max_{\theta\in\Lambda}\int_{\Gamma}\mathrm{C}[\phi(\mathbf{x}),\theta]p_{\theta}(\mathbf{x})d\mathbf{x}\geq\int_{\Lambda}v(\theta)\int_{\Gamma}\mathrm{C}[\phi(\mathbf{x}),\theta]p_{\theta}(\mathbf{x})d\mathbf{x}d\theta$. Since ϕ^{*} is the classical Bayes decision rule, which minimizes the Bayes risk without any constraints for the modified prior distribution of θ denoted by $g(\theta)$, one obtains:

$$\int_{\Lambda}g(\theta)\int_{\Gamma}\mathrm{C}[\phi(\mathbf{x}),\theta]p_{\theta}(\mathbf{x})d\mathbf{x}d\theta\geq\int_{\Lambda}g(\theta)\int_{\Gamma}\mathrm{C}[\phi^{*}(\mathbf{x}),\theta]p_{\theta}(\mathbf{x})d\mathbf{x}d\theta\ . \tag{A2}$$

If $\int_{\Lambda}v(\theta)\int_{\Gamma}\mathrm{C}[\phi^{*}(\mathbf{x}),\theta]p_{\theta}(\mathbf{x})d\mathbf{x}d\theta=\max_{\theta\in\Lambda}\int_{\Gamma}\mathrm{C}[\phi^{*}(\mathbf{x}),\theta]p_{\theta}(\mathbf{x})d\mathbf{x}$, the following equality holds

$$\begin{aligned}
&\lambda\int_{\Lambda}\omega(\theta)\int_{\Gamma}\mathrm{C}[\phi^{*}(\mathbf{x}),\theta]p_{\theta}(\mathbf{x})d\mathbf{x}d\theta+(1-\lambda)\int_{\Lambda}v(\theta)\int_{\Gamma}\mathrm{C}[\phi^{*}(\mathbf{x}),\theta]p_{\theta}(\mathbf{x})d\mathbf{x}d\theta\\
&=\lambda\int_{\Lambda}\omega(\theta)\int_{\Gamma}\mathrm{C}[\phi^{*}(\mathbf{x}),\theta]p_{\theta}(\mathbf{x})d\mathbf{x}d\theta+(1-\lambda)\max_{\theta\in\Lambda}\int_{\Gamma}\mathrm{C}[\phi*(\mathbf{x}),\theta]p_{\theta}(\mathbf{x})d\mathbf{x}
\end{aligned}. \tag{A3}$$

Therefore, ϕ^{*} is a solution of the optimization problem in (5). Especially, the value of the expression in (5) is always greater than or equal to (A3). □

Proof of Proposition 1. Based on the assumption in Theorem 1, ϕ^* is the optimal decision rule to minimize the value of the objective function in (5), which means:

$$\begin{aligned}&\lambda\int_\Lambda \omega(\theta)\int_\Gamma \mathrm{C}[\phi^*(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}d\theta+(1-\lambda)\max_{\theta\in\Lambda}\int_\Gamma \mathrm{C}[\phi^*(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}\\ &\le \lambda\int_\Lambda \omega(\theta)\int_\Gamma \mathrm{C}[\phi(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}d\theta+(1-\lambda)\max_{\theta\in\Lambda}\int_\Gamma \mathrm{C}[\phi(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}\end{aligned}. \tag{A4}$$

Since $\max_{\theta\in\Lambda}\int_\Gamma \mathrm{C}[\phi(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}\le\alpha$ holds due to the constraint on the conditional risks in (6) and $\max_{\theta\in\Lambda}\int_\Gamma \mathrm{C}[\phi^*(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}=\alpha$ holds based on the assumption in the corollary, $\int_\Lambda \omega(\theta)\int_\Gamma \mathrm{C}[\phi(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}d\theta$ should be greater than or equal to $\int_\Lambda \omega(\theta)\int_\Gamma \mathrm{C}[\phi^*(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}d\theta$ in order to satisfy the inequality in (A4). Therefore, ϕ^* is the optimal decision rule that minimizes the Bayes risk under the constraint on the conditional risks if $\max_{\theta\in\Lambda}\int_\Gamma \mathrm{C}[\phi^*(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}=\alpha$. In other words, ϕ^* is a solution of the optimization problem in (2) and (3). □

Proof of Theorem 2. Under the conditions in Theorem 1, the minimum Bayes risk corresponding to $g(\theta)$ is expressed as:

$$\begin{aligned}&\int_\Lambda g(\theta)\int_\Gamma \mathrm{C}[\phi^*(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}d\theta\\ &=\lambda\int_\Lambda \omega(\theta)\int_\Gamma \mathrm{C}[\phi^*(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}d\theta+(1-\lambda)\int_\Lambda v(\theta)\int_\Gamma \mathrm{C}[\phi^*(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}d\theta,\\ &=\lambda\int_\Lambda \omega(\theta)\int_\Gamma \mathrm{C}[\phi^*(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}d\theta+(1-\lambda)\max_{\theta\in\Lambda}\int_\Gamma \mathrm{C}[\phi^*(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}\end{aligned} \tag{A5}$$

where the second equality holds according to the condition (6) in Theorem 1. Since $\int_\Lambda \omega(\theta)\int_\Gamma \mathrm{C}[\phi^*(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}d\theta\le\max_{\theta\in\Lambda}\int_\Gamma \mathrm{C}[\phi^*(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}$, one finds:

$$\begin{aligned}&\int_\Lambda g(\theta)\int_\Gamma \mathrm{C}[\phi^*(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}d\theta\\ &\ge\hat{\lambda}\int_\Lambda \omega(\theta)\int_\Gamma \mathrm{C}[\phi^*(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}d\theta+(1-\hat{\lambda})\max_{\theta\in\Lambda}\int_\Gamma \mathrm{C}[\phi^*(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}\\ &\ge\hat{\lambda}\int_\Lambda \omega(\theta)\int_\Gamma \mathrm{C}[\phi^*(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}d\theta+(1-\hat{\lambda})\int_\Lambda \hat{v}(\theta)\int_\Gamma \mathrm{C}[\phi^*(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}d\theta\\ &=\int_\Lambda(\hat{\lambda}\omega(\theta)+(1-\hat{\lambda})\hat{v}(\theta))\int_\Gamma \mathrm{C}[\phi^*(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}d\theta\\ &=\int_\Lambda \hat{g}(\theta)\int_\Gamma \mathrm{C}[\phi^*(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}d\theta\\ &\ge\int_\Lambda \hat{g}(\theta)\int_\Gamma \mathrm{C}[\hat{\phi}^*(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}d\theta\end{aligned} \tag{A6}$$

for any $\hat{\lambda}\ge\lambda$, where the second inequality follows from $\max_{\theta\in\Lambda}\int_\Gamma \mathrm{C}[\phi^*(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}\ge\int_\Lambda \hat{v}(\theta)\int_\Gamma \mathrm{C}[\phi^*(\mathbf{x}),\theta]p_\theta(\mathbf{x})d\mathbf{x}d\theta$, and the last inequality holds based on the definition of $\hat{\phi}^*$. Therefore, $g(\theta)$ specified in Theorem 1 is the least favorable distribution that maximizes the Bayes risk among all the probability distributions in the form of $\hat{g}(\theta)=\hat{\lambda}\omega(\theta)+(1-\hat{\lambda})\hat{v}(\theta)$ for $\hat{\lambda}\ge\lambda$. □

Proof of Theorem 3. According to the definition of the restricted Bayes optimization problem formulated in (A), $r(\phi_r^\alpha)$ is a non-increasing function in terms of α. First, we shall use the contradiction method to prove the strictly decreasing property of $r(\phi_r^\alpha)$. Suppose that $M\le\alpha_1\le\alpha_2\le B$ and $r(\phi_r^{\alpha_1})=r(\phi_r^{\alpha_2})$. Actually, $\phi_r^{\alpha_2}$ is a restricted Bayes rule corresponding to α_1, which means $\max_{\theta\in\Lambda}R_\theta(\phi_r^{\alpha_2})=\alpha_1<\alpha_2$. Obviously, it contradicts with the conclusion in (27). Hence, $r(\phi_r^{\alpha_1})\ge r(\phi_r^{\alpha_2})$ must be satisfied.

Next, we shall prove the convexity of $r(\phi_r^\alpha)$ over $\alpha\in[M,B]$ by defining a decision rule ϕ as follows:

$$\phi = \xi\phi_r^{\alpha_1} + (1-\xi)\phi_r^{\alpha_2}, \tag{A7}$$

where $0 \le \xi \le 1$, $M \le \alpha_1 \le \alpha_2 \le B$, $\phi_r^{\alpha_1}$ and $\phi_r^{\alpha_2}$ denote the restricted Bayes decision rules corresponding to the constraint $\alpha = \alpha_1$ and $\alpha = \alpha_1$, respectively. The conditional risks for $\theta \in \Lambda$ and the Bayes risk corresponding to ϕ are respectively calculated by:

$$R_\theta(\phi) = \xi R_\theta(\phi_r^{\alpha_1}) + (1-\xi) R_\theta(\phi_r^{\alpha_2}), \tag{A8}$$

$$r(\phi) = \xi r(\phi_r^{\alpha_1}) + (1-\xi) r(\phi_r^{\alpha_2}), \tag{A9}$$

From (A8), the maximum conditional risk is upper bounded by:

$$\begin{aligned} \max_{\theta\in\Lambda} R_\theta(\phi) &= \max_{\theta\in\Lambda}\{\xi R_\theta(\phi_r^{\alpha_1}) + (1-\xi) R_\theta(\phi_r^{\alpha_2})\} \\ &\le \xi \max_{\theta\in\Lambda} R_\theta(\phi_r^{\alpha_1}) + (1-\xi)\max_{\theta\in\Lambda} R_\theta(\phi_r^{\alpha_2}), \\ &= \xi\alpha_1 + (1-\xi)\alpha_2 \end{aligned} \tag{A10}$$

where the last equality holds according to (27). Let $\alpha_o = \max_{\theta\in\Lambda} R_\theta(\phi)$ and $\hat{\alpha} = \xi\alpha_1 + (1-\xi)\alpha_2$, one obtains:

$$r(\phi) \ge r(\phi_r^{\alpha_o}) \ge r(\phi_r^{\hat{\alpha}}), \tag{A11}$$

where the first inequality is satisfied since $\phi_r^{\alpha_o}$ is the optimal decision rule that minimizes the Bayes risk under the constraint $\alpha = \alpha_o$ on the conditional risks, and the last inequality holds since $r(\phi_r^{\alpha})$ is non-increasing with respect to (w.r.t.) α. Therefore, $r(\phi_r^{\alpha})$ is a strictly decreasing and convex function of α for $\alpha \in [M, B]$. □

References

1. Madadi, Z.; Anand, G.V.; Premkumar, A.B. Signal detection in generalized gaussian noise by nonlinear wavelet denoising. *IEEE Trans. Circuits Syst. I* **2013**, *60*, 2973–2986.
2. Cao, J.; Lin, Z. Bayesian signal detection with compressed measurements. *Inf. Sci.* **2014**, *289*, 241–253.
3. Higger, M.; Akcakaya, M. Nezamfar, H.; LaMountain, G.; Orhan, U.; Erdogmus, D. A Bayesian framework for intent detection and stimulation selection in SSVEP BCIs. *IEEE Signal Process. Lett.* **2015**, *22*, 743–747.
4. Cheng, Q.; Zhou, H.; Cheng, J.; Li, H. A Minimax framework for classification with applications to images and high dimensional data. *IEEE Trans. Pattern Anal. Mach. Intell.* **2014**, *36*, 2117–2130.
5. Alhakim, R.; Raoof, K.; Simeu, E. Detection of UWB signal using dirty template approach. *Signal Image Video Process.* **2014**, *8*, 549–563.
6. Shevlyakov, G.; Shin, V.; Lee, S.; Kim, K. Asymptotically stable detection of a weak signal. *Int. J. Adapt. Control Signal Process.* **2014**, *28*, 848–858.
7. Morey, R.D.; Wagenmakers, E.J. Simple relation between Bayesian order-restricted and point-null hypothesis tests. *Stat. Probab. Lett.* **2014**, *92*, 121–124.
8. Ciuonzo, D.; Romano, G.; Rossi, P.S. Channel-aware decision fusion in distributed MIMO wireless sensor networks: Decode-and-fuse vs. decode-then-fuse. *IEEE Trans. Wirel. Commun.* **2012**, *11*, 2976–2985.
9. Rossi, P.S.; Ciuonzo, D.; Romano, G. Orthogonality and Cooperation in Collaborative Spectrum Sensing through MIMO Decision Fusion. *IEEE Trans. Wirel. Commun.* **2013**, *12*, 5826–5836.
10. Bayram, S.; Gezici, S. On the restricted Neyman–Pearson approach for composite hypothesis-testing in presence of prior distribution uncertainty. *IEEE Trans. Signal Process.* **2011**, *59*, 5056–5065.
11. Poor, H.V. *An Introduction to Signal Detection and Estimation*; Springer: New York, NY, USA, 1994.
12. Strawderman, W.E. Minimaxity. *J. Am. Stat. Assoc.* **2000**, *95*, 1364–1368.
13. Lehmann, E.L. Some history of optimality. *Lect. Notes Mono-Graph Ser.* **2009**, *57*, 11–17.
14. Begum, N.; King, M.L. A new class of test for testing a composite null against a composite alternative hypothesis. In Proceedings of the Australasian Meeting of the Econometric Society, Brisbane, Australia, 3–6 July 2007.

15. Van Trees, H.L. *Detection, Estimation, and Modulation Theory: Part I*, 2nd ed.; Wiley: New York, NY, USA, 2001.
16. Ciuonzo, D.; Maio, A.D.; Orlando, D. On the Statistical Invariance for Adaptive Radar Detection in Partially Homogeneous Disturbance Plus Structured Interference. *IEEE Trans. Signal Process.* **2016**, *65*, 1222–1234.
17. Blum, J.R.; Rosenblatt, J. On partial a priori information in statistical inference. *Ann. Math. Stat.* **1967**, *38*, 1671–1678.
18. Hodges, J.L., Jr.; Lehmann, E.L. The use of previous experience in reaching statistical decisions. *Ann. Math. Stat*. **1952**, *23*, 396–407.
19. Robbins, H. An empirical Bayes approach to statistics. In Proceedings of the 3rd Berkeley Symposium on Mathematical Statistics and Probability 1, Berkeley, CA, USA, December 1954–July & August 1955; pp. 157–164.
20. Robbins, H. The empirical Bayes approach to statistical decision problems. *Ann. Math. Stat.* **1964**, *35*, 1–20.
21. Savage, L.J. *The Foundations of Statistics*, 2nd ed.; Dover: New York, NY, USA, 1972.
22. Watson, S.R. On Bayesian inference with incompletely specified prior distributions. *Biometrika* **1974**, *61*, 193–196.
23. Caticha, A.; Preuss, R. Maximum entropy and Bayesian data analysis: Entropic prior distributions. *Phys. Rev. E* **2014**, *70*, doi:10.1103/PhysRevE.70.046127.
24. Palmieri, F.A.N.; Ciuonzo, D. Objective priors from maximum entropy in data classification. *Inf. Fusion* **2013**, *14*, 186–198.
25. Varshney, K.R.; Varshney, L.R. Optimal Grouping for Group Minimax Hypothesis Testing. *IEEE Trans. Inf. Theory* **2014**, *60*, 6511–6521.
26. Parsopoulos, K.E.; Vrahatis, M.N. *Particle Swarm Optimization Method for Constrained Optimization Problems*; IOS Press: Amsterdam, The Netherlands, 2002; pp. 214–220.
27. Hu, X.; Eberhart, R. Solving constrained nonlinear optimization problems with particle swarm optimization. In Proceedings of the World Multiconference on Systemics, Cybernetics and Informatics, Orlando, FL, USA, 14–18 July 2002.
28. Price, K.V.; Storn, R.M.; Lampinen, J.A. *Differential Evolution: A Practical Approach to Global Optimization*; Springer: New York, NY, USA, 2005.

© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).