Optimal Parameter Design for Estimation Theoretic Secure Broadcast

Cagri Goken, *Student Member, IEEE*, and Sinan Gezici, *Senior Member, IEEE*

***Abstract*—In this letter, estimation theoretic secure broadcast of a random parameter is investigated. In the considered setting, each receiver device employs a fixed estimator and carries a certain security risk such that its decision can be available to a malicious third party with a certain probability. The encoder at the transmitter is allowed to use a random mapping to minimize the weighted sum of the conditional Bayes risks of the estimators under secrecy and average power constraints. After formulating the optimal parameter design problem, it is shown that the optimization problem can be solved individually for each parameter value and the optimal mapping at the transmitter involves a randomization among at most three different signal levels. Sufficient conditions for improvability and non-improvability of the deterministic design via stochastic encoding are obtained. Numerical examples are provided to corroborate the theoretical results.**

***Index Terms*—Parameter estimation, broadcast channel, secrecy, optimization.**

I. Introduction

As an alternative to traditional information theoretic secrecy, estimation theoretic secrecy is investigated in a wide variety of settings to design low-complexity, practical, and secure systems, where the main goal is to securely transmit data to an intended receiver in the presence of a malicious third party such as an eavesdropper or a hijacker [1]–[8]. To achieve this goal, encoding the message/parameter at the transmitter can be an effective strategy. In [1], binary stochastic encoding, i.e., bit flipping, is applied on the quantized version of a noisy measurement related to a deterministic parameter to achieve secure communication. In [2] and [3], the optimal deterministic encoding of a random scalar parameter is investigated to minimize the expectation of conditional Cramér-Rao bound (ECRB) and the worst-case Fisher information of the parameter at the intended receiver, respectively, while ensuring a certain estimation error at the eavesdropper. In [4], secure transmission of a vector parameter is investigated and practical encoding strategies are introduced. In [5], estimation theoretic security is investigated when the encoder at the transmitter is allowed to use a randomized mapping with two functions and the eavesdropper is fully aware of the encoding strategy.

Secure broadcast of data to multiple users is a critical issue in the secrecy literature [8]–[11]. In [8], beamforming schemes are developed to ensure that legitimate users meet individual estimation error targets whereas the eavesdropper is deliberately jammed by an artificial noise component. In [11], security via regularized channel inversion precoding is investigated in a broadcast channel with confidential messages, where the transmitter broadcasts data to multiple users including potentially malicious ones and external eavesdroppers.

In certain scenarios, malicious third parties can directly hijack the devices in the system or can access decoded/estimated data. In this letter, we consider the broadcast of a parameter to a number of low-complexity receivers with fixed estimators, where each receiver carries a certain risk of being compromised. Our goal is to obtain an optimal parameter encoding strategy to minimize the average estimation performance at the receivers under secrecy and power constraints. To this end, each parameter is mapped using a stochastic function. In the literature, stochastic encoding of random parameters is studied for estimation problems [12], [13]; however, secrecy constraints are not considered, which become highly critical in modern systems. We show that an optimal signal design involves randomization among at most three different signal levels for each parameter value. We also provide sufficient conditions to specify when randomization can or cannot improve the optimal deterministic signaling approach. Numerical examples illustrate the benefits of randomization and the impact of the constraints to the overall estimation performance.

II. Optimal Parameter Design

Consider a system in which parameter $\boldsymbol{\theta} \in \Lambda$ is broadcasted to K different devices, where the channel for each device is modeled as an additive noise channel. The transmitter can send a random function of the parameter, that is, $\mathbf{s}_{\boldsymbol{\theta}}$, for each value of $\boldsymbol{\theta}$. Then, the received signal at the kth device can be written as

$$\mathbf{y}_k = \mathbf{s}_{\boldsymbol{\theta}} + \mathbf{n}_k\,, \quad k \in \{1 \dots K\}\,, \tag{1}$$

where $\mathbf{n}_k$ denotes the channel noise, which has a generic probability density function (PDF) represented by $p_{\mathbf{n}_k}(\cdot)$. Also, the prior distribution of the parameter is denoted by $w(\boldsymbol{\theta})$, and $\mathbf{s}_{\boldsymbol{\theta}}$ and $\mathbf{n}_k$ are independent for all $\boldsymbol{\theta}$. It is assumed that each receiving device employs a fixed estimator $\hat{\boldsymbol{\theta}}_k(\mathbf{y}_k)$ based on their observation $\mathbf{y}_k$. (Note that the estimators of the devices can be different.) Also, each device in the system has a certain assessed security risk probability γ_k to be compromised such that the estimate of the parameter at device k becomes available to a malicious third party with probability γ_k. It is important to emphasize that in the secrecy literature, the common assumption is that eavesdroppers employ optimal estimators/decoders to obtain the secret message since such an assumption (and knowledge of the encoding strategy at the eavesdropper in some

Manuscript received October 28, 2019; revised December 23, 2019; accepted January 16, 2020. Date of publication January 23, 2020; date of current version February 13, 2020. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Yunlong Cai. *(Corresponding author: Cagri Goken.)*

The authors are with the Dept. of Electrical and Electronics Engineering, Bilkent University, Bilkent, Ankara 06800, Turkey (e-mail: cgoken@ee.bilkent.edu.tr; gezici@ee.bilkent.edu.tr).

Digital Object Identifier 10.1109/LSP.2020.2969019

scenarios) is required to obtain fundamental limits of secure communications. In our setting, it is assumed that the malicious parties may hijack the estimators/devices instead of designing their own, and it is assumed that the receivers in the systems are simple, low-complexity devices employing potentially suboptimal estimators. It is important to note that these two assumptions are independent. It means that the devices are not vulnerable due to their simplicity but due to possible proximity to adversarial attacks and security measures against them. Such scenarios can be encountered in wireless sensor networks, public safety, and tactical communications scenarios in practice.

The main goal at the transmitter is to find the optimal probability distribution of $\mathbf{s}_{\boldsymbol{\theta}}$, that is, $p_{\mathbf{s}_{\boldsymbol{\theta}}}$, for each $\boldsymbol{\theta} \in \Lambda$ in order to minimize the weighted sum of Bayes risks of the estimators in the system under a security constraint on each value of $\boldsymbol{\theta}$. For a given value of $\boldsymbol{\theta}$, the conditional Bayes risk of the estimator at the kth device, $R_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}}_k)$, is given by

$$R_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}}_k) = \int C[\hat{\boldsymbol{\theta}}_k(\mathbf{y}_k), \boldsymbol{\theta}] p_{\boldsymbol{\theta}}(\mathbf{y}_k) d\mathbf{y}_k \,, \tag{2}$$

where $C[\hat{\boldsymbol{\theta}}_k(\mathbf{y}_k), \boldsymbol{\theta}] \geq 0$ represents a cost function [14], and $p_{\boldsymbol{\theta}}(\mathbf{y}_k)$ denotes the conditional PDF of $\mathbf{y}_k$ for a given value of parameter $\boldsymbol{\theta}$. Note that $p_{\boldsymbol{\theta}}(\mathbf{y}_k)$ can be expressed in terms of the PDF of $\mathbf{n}_k$ and the probability distribution of $\mathbf{s}_{\boldsymbol{\theta}}$ as $p_{\boldsymbol{\theta}}(\mathbf{y}_k) = \int p_{\mathbf{s}_{\boldsymbol{\theta}}}(\mathbf{x}) p_{\mathbf{n}_k}(\mathbf{y}_k - \mathbf{x}) d\mathbf{x}$ since $\mathbf{s}_{\boldsymbol{\theta}}$ and $\mathbf{n}_k$ are independent. Then, (2) becomes

$$R_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}}_k) = \int C[\hat{\boldsymbol{\theta}}_k(\mathbf{y}_k), \boldsymbol{\theta}] p_{\boldsymbol{\theta}}(\mathbf{y}_k) d\mathbf{y}_k = \int p_{\mathbf{s}_{\boldsymbol{\theta}}}(\mathbf{x}) \times \int C[\hat{\boldsymbol{\theta}}_k(\mathbf{y}_k), \boldsymbol{\theta}] p_{\mathbf{n}_k}(\mathbf{y}_k - \mathbf{x}) d\mathbf{y}_k d\mathbf{x} = E\{f_{\boldsymbol{\theta}}^{(k)}(\mathbf{s}_{\boldsymbol{\theta}})\} \tag{3}$$

where $f_{\boldsymbol{\theta}}^{(k)}(\mathbf{x}) \triangleq \int C[\hat{\boldsymbol{\theta}}_k(\mathbf{y}_k), \boldsymbol{\theta}] p_{\mathbf{n}_k}(\mathbf{y}_k - \mathbf{x}) d\mathbf{y}_k$ and the expectation operator in (3) is over the PDF of $\mathbf{s}_{\boldsymbol{\theta}}$ for a given value of $\boldsymbol{\theta}$. Also, the Bayes risk of the estimator at the kth device is given by

$$r(\hat{\boldsymbol{\theta}}_k) = \int_{\Lambda} w(\boldsymbol{\theta}) R_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}}_k) d\boldsymbol{\theta} \,. \tag{4}$$

In order to measure the estimation performance of the whole system, we consider the weighted sum of Bayes risks of the estimators at the devices, where each Bayes risk is weighted by $c_k(\gamma_k)$, which is a non-negative scalar function of γ_k. Then, the objective function becomes $\sum_{k=1}^{K} c_k(\gamma_k) r(\hat{\boldsymbol{\theta}}_k)$, which is expressed, via (3) and (4), as

$$\begin{aligned} \sum_{k=1}^{K} c_k(\gamma_k) r(\hat{\boldsymbol{\theta}}_k) &= \int_{\Lambda} w(\boldsymbol{\theta}) \sum_{k=1}^{K} c_k(\gamma_k) E\{f_{\boldsymbol{\theta}}^{(k)}(\mathbf{s}_{\boldsymbol{\theta}})\} d\boldsymbol{\theta} \\ &= \int_{\Lambda} w(\boldsymbol{\theta}) E\left\{ \sum_{k=1}^{K} c_k(\gamma_k) f_{\boldsymbol{\theta}}^{(k)}(\mathbf{s}_{\boldsymbol{\theta}}) \right\} d\boldsymbol{\theta} \\ &= \int_{\Lambda} w(\boldsymbol{\theta}) E\{F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\} d\boldsymbol{\theta} \end{aligned} \tag{5}$$

where $F_{\boldsymbol{\theta}}(\mathbf{x}) \triangleq \sum_{k=1}^{K} c_k(\gamma_k) f_{\boldsymbol{\theta}}^{(k)}(\mathbf{x})$.

Furthermore, the security constraint on each value of $\boldsymbol{\theta}$ is modeled as the weighted sum of the conditional Bayes risks of the estimators in the system, where each (non-negative) weight is denoted by $b_k(\gamma_k)$,[1] that is, $\sum_{k=1}^{K} b_k(\gamma_k) R_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}}_k) = \sum_{k=1}^{K} b_k(\gamma_k) E\{f_{\boldsymbol{\theta}}^{(k)}(\mathbf{s}_{\boldsymbol{\theta}})\}$. Then, the security constraint is in the form of

$$E\{G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\} \geq \eta_{\boldsymbol{\theta}} \,, \quad \forall \boldsymbol{\theta} \in \Lambda \tag{6}$$

where $G_{\boldsymbol{\theta}}(\mathbf{x}) \triangleq \sum_{k=1}^{K} b_k(\gamma_k) f_{\boldsymbol{\theta}}^{(k)}(\mathbf{x})$, and $\eta_{\boldsymbol{\theta}}$ is the secrecy limit for each value of $\boldsymbol{\theta}$. In $G_{\boldsymbol{\theta}}(\mathbf{x})$, the estimation performance of the devices that are more likely to be compromised is prioritized as compared to that of the safer devices via proper weighting. The physical meaning behind the constraint in (6) is that the total estimation accuracy of the vulnerable, high-risk devices is limited by a security target. In practical systems, there is also an average power constraint on the encoded version of the parameter in the form of $E\{||\mathbf{s}_{\boldsymbol{\theta}}||^2\} \leq A_{\boldsymbol{\theta}}$, where $||\mathbf{s}_{\boldsymbol{\theta}}||$ is the Euclidean norm of vector $\mathbf{s}_{\boldsymbol{\theta}}$ and $A_{\boldsymbol{\theta}}$ is the average power limit for $\boldsymbol{\theta}$. Therefore, based on (5) and (6), the optimal parameter design problem can be proposed as

$$\begin{aligned} &\min_{p_{\mathbf{s}_{\boldsymbol{\theta}}}, \boldsymbol{\theta} \in \Lambda} \int_{\Lambda} w(\boldsymbol{\theta}) E\{F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\} d\boldsymbol{\theta} \\ &\text{s.t.} \quad E\{G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\} \geq \eta_{\boldsymbol{\theta}} \,, \; E\{||\mathbf{s}_{\boldsymbol{\theta}}||^2\} \leq A_{\boldsymbol{\theta}} \,, \quad \forall \boldsymbol{\theta} \in \Lambda \end{aligned} \tag{7}$$

where $F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$ and $G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$ are as defined before. Note that as the constraints in (7) are defined for each value of $\boldsymbol{\theta}$, the optimization problem can be solved individually for each $\boldsymbol{\theta}$; hence, the solution does not depend on the prior distribution $w(\boldsymbol{\theta})$. In particular, (7) becomes

$$\min_{p_{\mathbf{s}_{\boldsymbol{\theta}}}} E\{F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\} \text{ s.t. } E\{G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\} \geq \eta_{\boldsymbol{\theta}}, \; E\{||\mathbf{s}_{\boldsymbol{\theta}}||^2\} \leq A_{\boldsymbol{\theta}} \tag{8}$$

for $\boldsymbol{\theta} \in \Lambda$. The optimization problems in the form of (8) have extensively been studied in the literature [12], [15], [16]. It can be shown that if $F_{\boldsymbol{\theta}}(\mathbf{x})$ and $G_{\boldsymbol{\theta}}(\mathbf{x})$ are continuous and each component of $\mathbf{x}$ belongs to a finite closed interval, an optimal solution of (8) involves randomization among at most 3 different values of $\mathbf{s}_{\boldsymbol{\theta}}$ due to Carathéodory's theorem [17].[2] Hence, the optimal parameter design problem in (8) can be solved via the following problem:

$$\begin{aligned} &\min_{\{\lambda_{\boldsymbol{\theta},j}, \mathbf{s}_{\boldsymbol{\theta},j},\}_{j=1}^{3}} \sum_{j=1}^{3} \lambda_{\boldsymbol{\theta},j} F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta},j}) \\ &\text{s.t.} \quad \sum_{j=1}^{3} \lambda_{\boldsymbol{\theta},j} G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta},j}) \geq \eta_{\boldsymbol{\theta}}, \quad \sum_{j=1}^{3} \lambda_{\boldsymbol{\theta},j} ||\mathbf{s}_{\boldsymbol{\theta},j}||^2 \leq A_{\boldsymbol{\theta}}, \\ &\sum_{j=1}^{3} \lambda_{\boldsymbol{\theta},j} = 1, \; \lambda_{\boldsymbol{\theta},j} \; \in [0,1], j = 1,2,3. \end{aligned} \tag{9}$$

It is noted that the optimization problem in (9) is much simpler to solve compared to (8) as it involves optimization over 6 variables instead of PDFs. In some cases, the optimal solution

[1] It is reasonable to select $c_k(\gamma_k)$ to be a decreasing function of γ_k and $b_k(\gamma_k)$ to be increasing with γ_k. Two example selections for $c_k(\gamma_k)$ and $b_k(\gamma_k)$ are $c_k(\gamma_k) = 1 - \gamma_k$ and $b_k(\gamma_k) = \gamma_k$ or $c_k(\gamma_k) = 1\{\gamma_k < \tau\}$ and $b_k(\gamma_k) = 1\{\gamma_k \geq \tau\}$, where τ is the risk threshold and $1\{\cdot\}$ is the indicator function.

[2] In general, if there were N_c constraints in (8) involving $E\{\tilde{H}_{\boldsymbol{\theta}}^{(i)}(\mathbf{s}_{\boldsymbol{\theta}})\}$ for $i = 1, \ldots, N_c$ with continuous functions $\tilde{H}_{\boldsymbol{\theta}}^{(i)}$, then the solution of the optimization problem would involve randomization among at most $N_c + 1$ points. A proof of the statement for $N_c = 1$ and how Carathéodory's theorem is utilized is available in [12].

may not involve randomization and a deterministic solution can be sufficient to obtain the optimal solution. However, if the deterministic solution is improvable, this result implies that it is sufficient to randomize the signal by using at most 3 different levels.

The deterministic solution corresponds to the solution of the following problem:

$$\min_{\mathbf{s}_{\boldsymbol{\theta}}} \quad F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}}) \quad \text{s.t.} \quad G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}}) \geq \eta_{\boldsymbol{\theta}}, \quad ||\mathbf{s}_{\boldsymbol{\theta}}||^2 \leq A_{\boldsymbol{\theta}}. \tag{10}$$

The deterministic solution is improvable by the stochastic solution if there exists $p_{\mathbf{s}_{\boldsymbol{\theta}}}$ such that $E\{F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\} < F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}}^{det})$ with $E\{G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\} \geq \eta_{\boldsymbol{\theta}}$ and $E\{||\mathbf{s}_{\boldsymbol{\theta}}||^2\} \leq A_{\boldsymbol{\theta}}$.

Remark 1: The optimization problem in (9) turns out to be non-convex in most cases, and it is required to utilize global optimization techniques such as particle swarm optimization (PSO) or approximation techniques such as convex relaxation [12]. In this work, we utilize the Global Optimization Toolbox of MATLAB to obtain the solution of the optimization problems. For some specific cost functions and noise PDFs, the optimal solution can also be obtained directly. As an example in the scalar case, when the cost function is $C[\hat{\theta}_k(\mathbf{y}_k), \theta] = (\hat{\theta}_k(\mathbf{y}_k) - \theta)^2$ and the noise component is zero-mean, i.e., $E\{n_k\} = 0$ for all $k = 1, \ldots, K$, the problem simplifies and the solution can be obtained without global optimization techniques. The problems in (8)–(10) are not necessarily feasible in all cases. One generic sufficient condition for the existence of the solutions is $\eta_{\boldsymbol{\theta}} \leq \max_{||\mathbf{s}_{\boldsymbol{\theta}}||^2 \leq A_{\boldsymbol{\theta}}} G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$.

The following proposition provides a sufficient condition for the nonimprovability of the deterministic solution.

Proposition 1: If $F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$ is a convex and $G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$ is a concave function of $\mathbf{s}_{\boldsymbol{\theta}}$ for each $\boldsymbol{\theta}$, then the deterministic solution cannot be improved via the stochastic solution.

Proof: Due to Jensen's inequality, for any $\mathbf{s}_{\boldsymbol{\theta}}$, $||E\{\mathbf{s}_{\boldsymbol{\theta}}\}||^2 \leq E\{||\mathbf{s}_{\boldsymbol{\theta}}||^2\} \leq A_{\boldsymbol{\theta}}$, where the second inequality is due to the average power constraint. Therefore, for any feasible PDF of $\mathbf{s}_{\boldsymbol{\theta}}$ ($p_{\mathbf{s}_{\boldsymbol{\theta}}}$) for the problem in (8), $||E\{\mathbf{s}_{\boldsymbol{\theta}}\}||^2 \leq A_{\boldsymbol{\theta}}$. Similarly, $\eta_{\boldsymbol{\theta}} \leq E\{G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\} \leq G_{\boldsymbol{\theta}}(E\{\mathbf{s}_{\boldsymbol{\theta}}\})$ due to the concavity of $G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$. Let $\mathbf{s}_{\boldsymbol{\theta}}^{\dagger} = E\{\mathbf{s}_{\boldsymbol{\theta}}\}$; therefore, for any feasible $p_{\mathbf{s}_{\boldsymbol{\theta}}}$, $||\mathbf{s}_{\boldsymbol{\theta}}^{\dagger}||^2 \leq A_{\boldsymbol{\theta}}$ and $G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}}^{\dagger}) \geq \eta_{\boldsymbol{\theta}}$. As $\mathbf{s}_{\boldsymbol{\theta}}^{\dagger}$ is a feasible deterministic point, and $F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$ is convex, $F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}}^{det}) \leq F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}}^{\dagger}) \leq E\{F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\}$, where $\mathbf{s}_{\boldsymbol{\theta}}^{det}$ denotes the optimal deterministic solution. So, when $F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$ is convex and $G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$ is concave, $E\{F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\}$ in (8) cannot be lower than the optimal value of (10) for any feasible PDF of $\mathbf{s}_{\boldsymbol{\theta}}$. ■

The main idea behind Proposition 1 is as follows: Under the conditions in the proposition, for any candidate stochastic solution of (8), we can obtain the deterministic solution $\mathbf{s}_{\boldsymbol{\theta}}^{\dagger} = E\{\mathbf{s}_{\boldsymbol{\theta}}\}$, which outperforms the stochastic solution and satisfies the constraints in (8). Next, a sufficient condition for the improvability of the deterministic solution is provided.

Proposition 2: The deterministic solution can be improved via the stochastic solution for a given $\boldsymbol{\theta} \in \Lambda$, if there exists real vectors $\mathbf{x}$ and $\mathbf{z}$ such that $F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$ and $G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$ are second-order partial differentiable around $\mathbf{s}_{\boldsymbol{\theta}} = \mathbf{x}$, $||\mathbf{x}||^2 \leq A_{\boldsymbol{\theta}}$, and the following inequality is satisfied:

$$F_{\boldsymbol{\theta}}(\mathbf{x}) + (\eta_{\boldsymbol{\theta}} - G_{\boldsymbol{\theta}}(\mathbf{x})) \frac{\mathbf{z}^T\mathbf{H}_f\mathbf{z} - \frac{\mathbf{z}^T\tilde{\mathbf{f}}}{\mathbf{z}^T\mathbf{x}}||\mathbf{z}||^2}{\mathbf{z}^T\mathbf{H}_g\mathbf{z} - \frac{\mathbf{z}^T\tilde{\mathbf{g}}}{\mathbf{z}^T\mathbf{x}}||\mathbf{z}||^2} < F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}}^{det}) \tag{11}$$

where $\mathbf{s}_{\boldsymbol{\theta}}^{det}$ is the solution of (10), $\tilde{\mathbf{f}}$ and $\tilde{\mathbf{g}}$ denote the gradients of $F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$ and $G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$ at $\mathbf{s}_{\boldsymbol{\theta}} = \mathbf{x}$, respectively, and $\mathbf{H}_f$ and $\mathbf{H}_g$ are the Hessian matrices of $F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$ and $G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$ at $\mathbf{s}_{\boldsymbol{\theta}} = \mathbf{x}$, respectively.

Proof: Consider a value of $\boldsymbol{\theta}$ for which the conditions in the proposition are satisfied. The main goal is to show that randomization around $\mathbf{x}$ can achieve a strictly lower objective value than that of the deterministic solution while satisfying the constraints. Suppose that stochastic signaling involves randomization between two values, that is, $\mathbf{x} + \boldsymbol{\epsilon}_1$ and $\mathbf{x} + \boldsymbol{\epsilon}_2$. For sufficiently small $\boldsymbol{\epsilon}_1$ and $\boldsymbol{\epsilon}_2$, the following expressions can be written by using Taylor's series expansion around $\mathbf{s}_{\boldsymbol{\theta}} = x$: $||\mathbf{x} + \boldsymbol{\epsilon}_i||^2 \approx ||\mathbf{x}||^2 + 2\boldsymbol{\epsilon}_i^T\mathbf{x} + ||\boldsymbol{\epsilon}_i||^2$, $F_{\boldsymbol{\theta}}(\mathbf{x} + \boldsymbol{\epsilon}_i) \approx F_{\boldsymbol{\theta}}(\mathbf{x}) + 2\boldsymbol{\epsilon}_i^T\tilde{\mathbf{f}} + \boldsymbol{\epsilon}_i^T\mathbf{H}_f\boldsymbol{\epsilon}_i$, and $G_{\boldsymbol{\theta}}(\mathbf{x} + \boldsymbol{\epsilon}_i) \approx G_{\boldsymbol{\theta}}(\mathbf{x}) + 2\boldsymbol{\epsilon}_i^T\tilde{\mathbf{g}} + \boldsymbol{\epsilon}_i^T\mathbf{H}_g\boldsymbol{\epsilon}_i$ for $i = 1, 2$.

In order to show that the stochastic solution with PDF $p_{\mathbf{s}_{\boldsymbol{\theta}}}(\mathbf{s}_{\boldsymbol{\theta}}) = \lambda\delta(\mathbf{s}_{\boldsymbol{\theta}} - (\mathbf{x} + \boldsymbol{\epsilon}_1)) + (1 - \lambda)\delta(\mathbf{s}_{\boldsymbol{\theta}} - (\mathbf{x} + \boldsymbol{\epsilon}_2))$ improves the deterministic solution, it is sufficient to satisfy the following conditions: $\lambda||\mathbf{x} + \boldsymbol{\epsilon}_1||^2 + (1 - \lambda)||\mathbf{x} + \boldsymbol{\epsilon}_2||^2 = ||\mathbf{x}||^2 \leq A_{\boldsymbol{\theta}}$, $\lambda F_{\boldsymbol{\theta}}(\mathbf{x} + \boldsymbol{\epsilon}_1) + (1 - \lambda)F_{\boldsymbol{\theta}}(\mathbf{x} + \boldsymbol{\epsilon}_2) < F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}}^{det})$, and $\lambda G_{\boldsymbol{\theta}}(\mathbf{x} + \boldsymbol{\epsilon}_1) + (1 - \lambda)G_{\boldsymbol{\theta}}(\mathbf{x} + \boldsymbol{\epsilon}_2) = \eta_{\boldsymbol{\theta}}$.

If we insert the relations in the first paragraph of the proof into those in the second paragraph in order, and let $\boldsymbol{\epsilon}_1 = \alpha\mathbf{z}$ and $\boldsymbol{\epsilon}_2 = \beta\mathbf{z}$, then the relations become

$$\begin{aligned} &k\mathbf{z}^T\mathbf{x} = -||\mathbf{z}||^2, \\ &k\mathbf{z}^T\tilde{\mathbf{f}} + \mathbf{z}^T\mathbf{H}_f\mathbf{z} < (F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}}^{det}) - F_{\boldsymbol{\theta}}(\mathbf{x}))/k_2, \\ &k\mathbf{z}^T\tilde{\mathbf{g}} + \mathbf{z}^T\mathbf{H}_g\mathbf{z} = (\eta_{\boldsymbol{\theta}} - G_{\boldsymbol{\theta}}(\mathbf{x}))/k_2, \end{aligned} \tag{12}$$

where $k = k_1/k_2$ with $k_1 = 2(\lambda\alpha + (1 - \lambda)\beta)$ and $k_2 = \lambda\alpha^2 + (1 - \lambda)\beta^2$. Also note that $k = -||\mathbf{z}||^2/\mathbf{z}^T\mathbf{x}$ and $k_2 = (\eta_{\boldsymbol{\theta}} - G_{\boldsymbol{\theta}}(\mathbf{x}))/(k\mathbf{z}^T\tilde{\mathbf{g}} + \mathbf{z}^T\mathbf{H}_g\mathbf{z})$ due to the first and third equalities in (12). If they are inserted in the second relation in (12), the sufficient condition corresponds to that given in (11). ■

The idea in Proposition 2 is to provide conditions under which randomization around a real vector leads to an improvement over the optimal deterministic solution. To illustrate this idea, consider a scenario in which $F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$ and $G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$ are both convex functions. In this scenario, randomization increases the value of the objective and secrecy functions due to Jensen's inequality. Therefore, suppose that there exists a point $\mathbf{x}$ satisfying the average power constraint with $F_{\boldsymbol{\theta}}(\mathbf{x}) < F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}}^{det})$ and $G_{\boldsymbol{\theta}}(\mathbf{x}) < \eta_{\boldsymbol{\theta}}$. The condition in Proposition 2 implies that randomization around $\mathbf{x}$ ensures that the security constraint is satisfied, i.e., $E\{G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\} = \eta_{\boldsymbol{\theta}}$, while the increase in the objective value due to randomization is still sufficiently small to improve the deterministic solution, i.e., $E\{F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\} < F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}}^{det})$. Also, even though the derivation of Proposition 2 is based on the similar idea and techniques presented in [12], we manage to reduce the number of equations in the sufficient condition compared to [12], by allowing the randomization to be around any feasible point $\mathbf{x}$ with $||\mathbf{x}||^2 \leq A_{\boldsymbol{\theta}}$ and letting the candidate stochastic solution satisfy the secrecy constraint with equality.Note that Propositions 1 and 2 are valid regardless of the uniqueness of the solutions to the problems in (8)–(10).

Special Case With No Average Power Constraint: As a special case, we consider the problem in (8) with only the secrecy constraint. In this case, the optimization problem can be expressed as $\min_{p_{\mathbf{s}_{\boldsymbol{\theta}}}} E\{F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\}$ s.t. $E\{G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\} \geq$

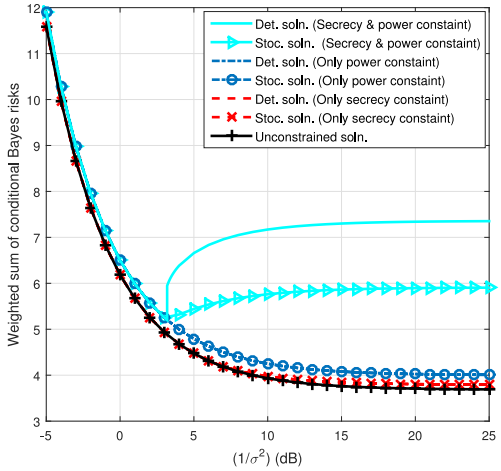

Fig. 1. Weighted sum of conditional Bayes risks versus $1/\sigma^2$.

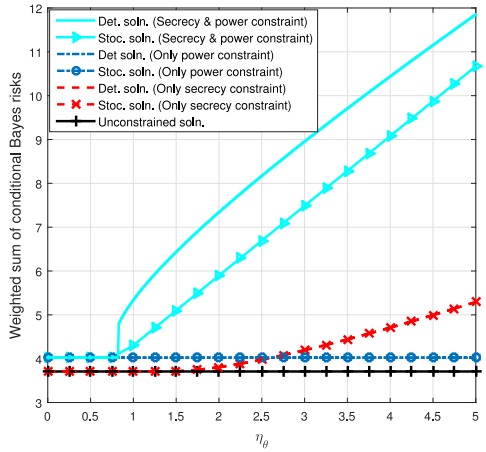

Fig. 2. Weighted sum of conditional Bayes risks versus η_θ.

$\eta_{\boldsymbol{\theta}}$. The solution of this problem involves randomization between at most 2 different values; hence, the simplified problem becomes $\min_{\{\lambda,\mathbf{s}_{\boldsymbol{\theta},1},\mathbf{s}_{\boldsymbol{\theta},2}\}} \lambda F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta},1}) + (1-\lambda)F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta},2})$ s.t. $\lambda G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta},1}) + (1-\lambda)G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta},2}) \geq \eta_{\boldsymbol{\theta}}$, $\lambda \in [0,1]$. Note that for given $(\mathbf{s}_{\boldsymbol{\theta},1}, \mathbf{s}_{\boldsymbol{\theta},2})$, finding the optimal λ based on the simplified problem is straightforward as all the functions can be calculated directly and have scalar real values.

III. Numerical Results and Conclusions

In the numerical examples, we consider the transmission of a scalar parameter θ to $K = 5$ devices, all of which employ the fixed estimator given by $\hat{\theta}_k(y_k) = y_k$ and the cost function is selected as the squared error function, i.e., $C[\hat{\theta}_k(\mathbf{y}_k), \theta] = (\hat{\theta}_k(\mathbf{y}_k) - \theta)^2$ for $k = 1, \ldots, K$. The security risk probabilities, γ_k's, of the devices are assessed as $(1, 0.75, 0.5, 0.25, 0)$ and $c_k(\gamma_k) = 1 - \gamma_k$ and $b_k(\gamma_k) = \gamma_k$ are used. The noise of the kth device (user) is modeled by Gaussian mixture noise with two mass points such that its PDF is given by $p_{n_k}(x) = \nu_k e^{-\frac{(x-\mu_k)^2}{\sigma^2}}/\sqrt{2\pi\sigma^2} + (1-\nu_k)e^{-\frac{(x+\mu_k)^2}{\sigma^2}}/\sqrt{2\pi\sigma^2}$, and the noise parameters are taken as $\boldsymbol{\nu} = [0.2, 0.25, 0.3, 0.4, 0.9]$ and $\boldsymbol{\mu} = [0, 0.4, 0.8, 1.2, 1.6]$, where $\boldsymbol{\nu} = [\nu_1, \nu_2, \nu_3, \nu_4, \nu_5]$ and $\boldsymbol{\mu} = [\mu_1, \mu_2, \mu_3, \mu_4, \mu_5]$. In the examples, the stochastic and deterministic solutions are considered for the original problem with the average power and secrecy constraints in (9) and (10), respectively. Also, the performance results are presented when there exists only the average power constraint, only the secrecy constraint and no constraints for comparison purposes as they yield various lower bounds for the original problem.

In Fig. 1, the weighted sum of the conditional Bayes risks (i.e., $E\{F_\theta(s_\theta)\}$) is plotted versus $1/\sigma^2$ for $\theta = 1$, $\eta_\theta = 2$, and $A_\theta = 1$. It is observed that the stochastic solution improves the deterministic solution especially for lower values of σ^2. It is also noted that the stochastic and deterministic parameter designs have the same performance when one of the constraints is removed, and their performance is close to the unconstrained solution (which is the solution of (8) in the absence of the constraints) in this particular scenario. However, when both of the constraints are imposed, the performance of the deterministic design starts to deteriorate severely in the low σ^2 region due to the interference components present in the Gaussian mixture noise. The randomization via stochastic signaling alleviates the effects of the interference resulting in an improved solution compared to the deterministic one. Although the deterministic solution is an attractive alternative due to its simplicity and achieves the optimal solution when the noise variance is large, there is no guarantee that the deterministic solution is a good approximation to the optimal stochastic solution in general.

In Fig. 2, the weighted sum of the conditional Bayes risks is plotted versus η_θ for $\theta = 1$, $1/\sigma^2 = 20$ dB, and $A_\theta = 1$. When both of the constraints are considered, it is observed that the Bayes risks start to rise as the security demand increases, especially for $\eta_\theta \geq 0.75$, and the stochastic solution improves the deterministic solution similarly to Fig. 1. Note that the unconstrained solution and the solution of the problem with only the average power constraint is constant as they do not consider the secrecy constraint. The solution of the unconstrained problem (s_θ^{unc}) satisfies the secrecy constraint until a certain point ($\eta_\theta \approx 1.75$); however, for larger η_θ, the secrecy constraint becomes effective leading to a slight increase in the Bayes risks. When one or both of the constraints are removed, the deterministic and stochastic designs have the same performance similarly to Fig. 1.

The improvement via stochastic signaling can theoretically be justified based on Proposition 2 when both constraints are considered. For example, when $\eta_\theta = 2$ and $1/\sigma^2 = 20$ dB, the optimal deterministic solution, s_θ^{det}, is 0.151 yielding 7.340 as the weighted sum of Bayes risks. The inequality condition given in (11) is satisfied when $z = 1$ and for any $x \in [0.151, 1]$. Since this is a sufficient condition for improvability, it is known that the deterministic solution can be improved via the stochastic solution. In fact, the stochastic solution represented by $p_{s_\theta}(s_\theta) = \lambda_\theta\, \delta(s_\theta - s_{\theta,1}) + (1-\lambda_\theta)\, \delta(s_\theta - s_{\theta,2})$ with $\lambda_\theta = 0.693$, $s_{\theta,1} = 1.196$, and $s_{\theta,2} = -0.168$ yields 5.762 as the weighted sum of Bayes risks. Note that for this solution $E\{s_\theta\} = 0.777$; hence, randomization around this point improves the deterministic solution as predicted by Proposition 2. In general, one possible way to check the sufficient condition in Proposition 2 is to fix $\mathbf{z}$, and then perform the search over $\mathbf{x}$ in the closed ball $||\mathbf{x}||^2 \leq A_{\boldsymbol{\theta}}$ while checking the inequality in (11). If there is no $\mathbf{x}$ satisfying the condition for a given $\mathbf{z}$, then another $\mathbf{z}$ can be selected until a preset number of maximum trials is reached.

REFERENCES

[1] T. C. Aysal and K. E. Barner, "Sensor data cryptography in wireless sensor networks," *IEEE Trans. Inf. Forensics Secur.*, vol. 3, no. 2, pp. 273–289, Jun. 2008.

[2] C. Goken and S. Gezici, "ECRB-based optimal parameter encoding under secrecy constraints," *IEEE Trans. Signal Process.*, vol. 66, no. 13, pp. 3556–3570, Jul. 2018.

[3] C. Goken and S. Gezici, "Optimal parameter encoding based on worst case Fisher information under a secrecy constraint," *IEEE Signal Process. Lett.*, vol. 24, no. 11, pp. 1611–1615, Nov. 2017.

[4] C. Goken, S. Gezici, and O. Arikan, "Estimation theoretic optimal encoding design for secure transmission of multiple parameters," *IEEE Trans. Signal Process.*, vol. 67, no. 16, pp. 4302–4316, Aug. 2019.

[5] C. Goken and S. Gezici, "Estimation theoretic secure communication via encoder randomization," *IEEE Trans. Signal Process.*, vol. 67, no. 23, pp. 6105–6120, Dec. 2019.

[6] A. Ozcelikkale and T. M. Duman, "Cooperative precoding and artificial noise design for security over interference channels," *IEEE Signal Process. Lett.*, vol. 22, no. 12, pp. 2234–2238, Dec. 2015.

[7] J. Zhang, R. S. Blum, and H. V. Poor, "Approaches to secure inference in the Internet of Things: Performance bounds, algorithms, and effective attacks on IoT sensor networks," *IEEE Signal Proc. Mag.*, vol. 35, no. 5, pp. 50–63, Sep. 2018.

[8] M. Pei, J. Wei, K. Wong, and X. Wang, "Masked beamforming for multiuser MIMO wiretap channels with imperfect CSI," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 544–549, Feb. 2012.

[9] R. Liu, T. Liu, H. V. Poor, and S. Shamai, "Multiple-input multiple-output Gaussian broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4215–4227, Sep. 2010.

[10] S. A. A. Fakoorian and A. L. Swindlehurst, "On the optimality of linear precoding for secrecy in the MIMO broadcast channel," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1701–1713, Sep. 2013.

[11] G. Geraci, S. Singh, J. G. Andrews, J. Yuan, and I. B. Collings, "Secrecy rates in broadcast channels with confidential messages and external eavesdroppers," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2931–2943, May 2014.

[12] H. Soganci, S. Gezici, and O. Arikan, "Optimal stochastic parameter design for estimation problems," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4950–4956, Sep. 2012.

[13] H. Soganci, S. Gezici, and O. Arikan, "Optimal signal design for multiparameter estimation problems," *IEEE Trans. Signal Process.*, vol. 63, no. 22, pp. 6074–6085, Nov. 2015.

[14] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York, Berlin, Germany: Springer, 1994.

[15] A. Patel and B. Kosko, "Optimal noise benefits in Neyman–Pearson and inequality-constrained signal detection," *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1655–1669, May 2009.

[16] S. Bayram, S. Gezici, and H. V. Poor, "Noise enhanced hypothesis-testing in the restricted Bayesian framework," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 3972–3989, Aug. 2010.

[17] R. T. Rockafellar, *Convex Analysis*. Princeton, NJ, USA: Princeton Univ. Press, 1968.