Estimation Theoretic Secure Communication via Encoder Randomization

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

***Abstract*—Estimation theoretic secure transmission of a scalar random parameter is investigated in the presence of an eavesdropper. The aim is to minimize the estimation error at the receiver under a secrecy constraint at the eavesdropper; or, alternatively, to maximize the estimation error at the eavesdropper for a given estimation accuracy limit at the receiver. In the considered setting, the encoder at the transmitter is allowed to use a randomized mapping between two one-to-one and continuous functions and the eavesdropper is fully aware of the encoding strategy at the transmitter. For small numbers of observations, both the eavesdropper and the receiver are modeled to employ linear minimum mean-squared error (LMMSE) estimators, and for large numbers of observations, the expectation of the conditional Cramér-Rao bound (ECRB) metric is employed for both the receiver and the eavesdropper. Optimization problems are formulated and various theoretical results are provided in order to obtain the optimal solutions and to analyze the effects of encoder randomization. In addition, numerical examples are presented to corroborate the theoretical results. It is observed that stochastic encoding can bring significant performance gains for estimation theoretic secrecy problems.**

***Index Terms*—Estimation, secrecy, Gaussian wiretap channel, optimization, Internet of Things (IoT).**

I. Introduction and Motivation

A. Literature Review

In a secure communication system, the main goal is to secretly transmit data to an intended receiver in the presence of a malicious third party such as an eavesdropper. As the age of Internet of Things (IoT), smart homes and cities, self-driving cars, and wireless sensor networks with a vast number of nodes has already arrived, it is necessary to find ways to ensure secure communication of data in such systems. Massive deployments of sensors, the nature of wireless links across a network, and the sensitivity of data collected by sensors present serious security challenges. Traditionally, key-based cryptographic approaches have been employed in many applications for secure communication [1], [2]. However, the management of key generation and distribution can be very challenging in heterogenous and dynamic networks with vast numbers of connections [3], [4]. Furthermore, as many nodes in sensor networks are low-cost with limited battery power and bandwidth and have strict latency requirements, it might not be suitable to consider cryptographic solutions as the only layer of security in such systems [5].

Based on these motivations, there has been a renewed interest in physical layer secrecy to develop alternative or complementary layers of security technologies. Physical layer secrecy is based on the idea of exploiting the randomness in wireless channel conditions to ensure secure communication [6]. In this regard, information theoretic metrics and tools, such as capacity, have been employed in a multitude of studies for various channel models such as fading channels [7], [8], Gaussian wiretap, broadcast and interference channels [9]–[12]. In the literature, alternative metrics and frameworks have also been utilized to quantify secrecy levels. For example, in [13] and [14], secure communication problem is investigated based on the signal-to-noise ratio (SNR) metric in the quality-of-service (QoS) framework. In [15], the secrecy constrained distributed detection problem is studied under Bayesian and Neyman-Pearson frameworks. Alternatively, secrecy levels can be measured via estimation theoretic tools and metrics, such as Fisher information and mean-squared error (MSE), where the aim is the design of low-complexity, practical, and secure systems [16]–[29].

Estimation theoretic secrecy has been studied in a wide variety of settings. In [16], the secret communication problem is considered for Gaussian interference channels with vector parameters in the presence of eavesdroppers. The problem is formulated to minimize the total minimum mean-squared error (MMSE) at the intended receivers while keeping the MMSE at the eavesdroppers above a certain threshold, where joint artificial noise and linear precoding schemes are used to satisfy the secrecy requirements. In [17], privacy of households using smart meters is considered in the presence of adversary parties who estimate energy consumption based on data gathered in smart meters. The Fisher information is employed as a metric of privacy for both scalar and multivariable parameter cases, and the optimal policies for the utilization of batteries are derived to minimize the Fisher information to achieve privacy. Both [18] and [19] investigate secrecy in a distributed inference framework, where the information coming to a fusion center from various sensor nodes can also be observed by eavesdroppers. In [18], the estimation problem of a single point Gaussian source in the presence of an eavesdropper is analyzed for the cases of multiple transmit sensors with a single antenna and a single sensor with multiple transmit antennas. Optimal transmit power allocation policies are derived to minimize the average MSE for the parameter of

Manuscript received May 23, 2019; revised September 19, 2019; accepted October 20, 2019. Date of publication November 4, 2019; date of current version November 25, 2019. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Stefano Tomasin. *(Corresponding author: Cagri Goken.)*

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

Digital Object Identifier 10.1109/TSP.2019.2951231

interest while guaranteeing a target MSE at the eavesdropper. In [19], the asymptotic secrecy and estimation problem is studied when the sensor measurements are quantized and the channel between sensors and receivers are assumed to be binary symmetric channels. The sensor quantization thresholds are designed to ensure perfect secrecy when the number of sensors is very large. In [20], the secure inference problem is investigated for deterministic parameters in IoT systems under spoofing and man-in-the-middle-attack (MIMA). For MIMAs, necessary and sufficient conditions are derived to decide when the attacked data can or cannot improve the estimation performance in terms of the Cramér-Rao bound. For spoofing attacks, effective attack strategies are described with a guaranteed performance in terms of Cramér-Rao bound (CRB) degradation and it is shown that quantization imposes a limit on the robustness of the system against such attacks.

Stochastic encryption has been used as a defense mechanism against eavesdropper attacks in the estimation theoretic security framework [21]–[24]. In [22], stochastic encryption is performed based on the 1-bit quantized version of a noisy sensor measurement of a deterministic parameter to achieve secret communication, where both symmetric and asymmetric bit flipping strategies are considered under the assumptions that the intended receiver is aware of the flipping probabilities and the eavesdropper is unaware of the encryption. It is shown that it is possible to create biased estimation and large errors at the eavesdropper via this simple scheme. In [23], the binary stochastic encryption (BSE) approach proposed in [22] is extended to non-binary stochastic encryption (NBSE) to facilitate vector parameter estimation. In [24], secrecy provided by stochastic encryption is studied under the assumptions that the eavesdropper is aware of the particular technique, e.g., BSE, NBSE, employed in the transmitter, uses an unbiased estimator, and does not know the encryption key and quantizer regions. It is shown that such a scheme is secure in the domain of unbiased estimators.

While the aforementioned studies focus on the stochastic encryption of a quantized measurement of a deterministic parameter, [25] and [26] focus on the secrecy problem for a random parameter in the Bayesian estimation setting. In [25], the optimal deterministic encoding of a scalar random parameter is investigated based on the minimization of expectation of the conditional Cramér-Rao bound (ECRB) in order to guarantee a certain level of estimation accuracy at the intended receiver while keeping the estimation error at the eavesdropper above a certain level. In [26], a robust parameter encoding approach is developed and the optimization is based on the worst-case CRB of the parameter in order to guarantee a certain level of estimation accuracy at the intended receiver. The results in [25] are extended to vector parameter estimation scenarios in [27]. The common assumption in [25]–[27] is that the encoding function is not available to the eavesdropper; hence, it acts like a secret key similarly to the assumption of flipping probabilities not being available to the eavesdropper in [22] and [24]. On the other hand, for determining fundamental security limits of many systems (such as those investigated in the classical information theoretical framework), it is a common practice to assume that the eavesdropper has the full knowledge of the encoding strategy at the transmitter. For example, in a Gaussian wiretap channel, the positive secrecy capacity is possible even though the eavesdropper knows the encoding scheme [12]. In particular, data is kept private as a result of the condition that the noise present in eavesdropper's received signal is stronger than the noise at the intended receiver. In that setting, the key ingredient is to apply stochastic encoding at the transmitter to achieve a positive rate with no data leakage to the eavesdropper. The encoder is used to confuse the eavesdropper with the cost of a reduced communication rate. Inspired from this classical setting, in this manuscript, estimation theoretic secure transmission of a scalar random parameter is investigated in a Gaussian wiretap channel under the Bayesian framework, which has not been investigated in the literature. As the encoding strategy is available to the eavesdropper, the encoder randomization is allowed to increase ambiguity to possibly enhance security. The work in this manuscript is distinguished from [25]–[27] as it assumes that the mapping strategy is available to both the eavesdropper and the receiver (i.e., not secret), allows stochastic encoding in the transmitter, considers multiple observations rather than a single one, and employs different performance metrics leading to a distinct optimization problem. It is also different from those studies (such as [22], [23]) that allow stochastic encryption as it considers direct encoding of a random parameter rather than a measured deterministic one.

B. Contributions

In this manuscript, estimation theoretic secure transmission of a scalar random parameter is investigated in the presence of an eavesdropper in a Gaussian wiretap channel. The aim is to achieve accurate estimation of the parameter at the intended receiver while keeping the estimation error at the eavesdropper above a certain level; or, alternatively, to ensure that the estimation error at the eavesdropper is as large as possible while satisfying an estimation accuracy constraint at the intended receiver. To enhance security, stochastic encoding is employed at the transmitter, and the encoder is modeled to perform randomization between two one-to-one, continuous encoding functions, which should be designed. It is assumed that the mapping at the encoder is fully available to the eavesdropper and the receiver. For small numbers of channel observations, both the eavesdropper and the receiver are modeled to employ linear MMSE (LMMSE) estimators, and for large numbers of observations, the ECRB metric is employed both in the receiver and the eavesdropper [30]. This is because of the fact that even though the optimal estimator in terms of the MSE metric is the MMSE estimator, the calculations for its MSE have high computational complexity and do not yield closed-form expressions in general. LMMSE and ECRB tightly approximate the optimal metric for small and large numbers of observations (e.g., see Figs. 2–4), respectively, in our setting, and they facilitate theoretical analyses with intuitive explanations based on closed-form expressions. Therefore, based on these metrics, the optimization problems are formulated to perform optimal encoding for small and large numbers of observations separately. Both generic and affine functions are considered in the proposed encoding scheme, and a number of theoretical

results on the solutions of the problems are provided. Finally, numerical examples are presented to illustrate the theoretical results for both small and large numbers of observations. The main contributions and novelty in this manuscript can be summarized as follows:

- The problem of parameter encoding via encoder randomization is analyzed to ensure estimation theoretic secure communication under the assumption that the encoding scheme is available to the eavesdropper.
- For small numbers of observations, a closed form expression for the MSE of the LMMSE estimator is derived for both the receiver and the eavesdropper for the considered transmission and encoding scheme. The optimization problems to minimize the MSE at the intended receiver for a given secrecy target at the eavesdropper and to maximize the MSE at the eavesdropper for a given estimation accuracy limit at the receiver are formulated. The relationship between the solutions of those problems is characterized. An optimal solution of the optimization problems is obtained theoretically when the channel of the eavesdropper is noisier than the channel of the intended receiver. It is also shown that a simple deterministic affine function can attain the optimal value. For the case of affine functions, the monotonicity behavior of the MSE is obtained with respect to the randomization probability when the encoding functions are fixed.
- For large numbers of observations, the optimization problems to minimize the ECRB at the intended receiver for a given secrecy target at the eavesdropper and to maximize the ECRB at the eavesdropper for a given estimation accuracy limit at the receiver are formulated. The optimizations problems are theoretically solved when only deterministic encoding is considered. It is also shown that under symmetric mapping, the ECRB is maximized when the randomization probability is $1/2$. Also, the monotonicity behavior of the ECRB is obtained with respect to the randomization probability when the encoding functions are fixed for this case, as well.

II. System Setup

Consider the transmission of a scalar parameter $\theta \in \Lambda$ to an intended receiver in the presence of an eavesdropper who wants to estimate parameter θ. Both the intended receiver and the eavesdropper obtain n-dimensional observations over their respective additive noise channels. The aim is to achieve accurate estimation of the parameter at the intended receiver while keeping the estimation error at the eavesdropper above a certain level; or, alternatively, to ensure that the estimation error at the eavesdropper is as large as possible while satisfying an estimation constraint at the intended receiver. To that aim, the parameter is encoded by an encoding function $f : \Lambda \to \Gamma$. Let $f(\theta)$ denote the encoded version of the parameter. Hence, the ith observation at the intended receiver can be written as

$$Y_i = f(\theta) + V_i, \quad i = 1, 2, \ldots, n. \tag{1}$$

where the noise V_i is modeled as a zero-mean Gaussian random variable with variance σ_V^2, and V_i and θ are assumed to be independent [12]. On the other hand, the ith observation at the eavesdropper is

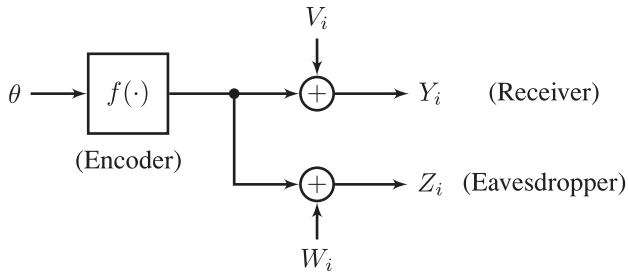

Fig. 1. System model for the parameter encoding problem.

$$Z_i = f(\theta) + W_i, \quad i = 1, 2, \ldots, n. \tag{2}$$

where W_i is zero-mean Gaussian noise with variance σ_W^2, which is independent of θ for $i = 1, 2, \ldots, n$. Also, the prior information on parameter θ is represented by a probability density function (PDF) denoted by $p_\theta(\theta)$ for $\theta \in \Lambda$. The signal model in (1) and (2) can also be employed for flat-fading channels assuming perfect channel estimation and appropriate equalization [31]. The intended receiver aims to estimate parameter θ based on observations $\mathbf{Y} \triangleq [Y_1, Y_2, \ldots, Y_n]^T$ whereas the eavesdropper uses observations $\mathbf{Z} \triangleq [Z_1, Z_2, \ldots, Z_n]^T$ for estimating θ. The system model is illustrated in Fig. 1.

The considered system model is also known as the Gaussian wiretap channel [9], [12], and has been studied extensively via information theoretical tools, as mentioned in Section I. In that framework, it is assumed that the eavesdropper knows the codewords (mapping) in the encoder and has unlimited resources/time for computation. Therefore, the encoder applies a stochastic mapping from messages to codewords to ensure that the message can be kept unknown to the eavesdropper by exploiting the degradedness of eavesdropper's channel while still being able to transmit the message to the intended receiver at a certain rate.[1] Motivated from such a setting, the following assumptions are made for the rest of this study:

- The encoding function at the transmitter is fully available to the eavesdropper and the receiver. Therefore, it is possible that both the eavesdropper and the receiver can utilize optimal estimators according to a certain metric.
- To enhance security, stochastic encoding is employed and the encoder is modeled to perform the following mapping:

$$f(\theta) = \begin{cases} f_1(\theta), & \text{with probability } \gamma \\ f_2(\theta), & \text{with probability } 1-\gamma \end{cases} \tag{3}$$

  where $f_k(\theta) : \Lambda \to \Gamma$ is a continuous and one-to-one function for $k = 1, 2$ and $\gamma \in [0, 1]$. [2]
- Each observation is corrupted by independent and identically distributed noise components. Therefore, based on this

[1] Unlike the classical Gaussian wiretap channel [9], [12], we consider a scenario in which the channel of the eavesdropper is not necessarily worse than that of the intended receiver.

[2] The stochastic encoder in (3) both facilitates practical implementations and allows for theoretical investigations. Note that it can also be represented as $f(\theta) = f_{2-X}(\theta)$, where X is a Bernoulli random variable with parameter γ and X is statistically independent of all other variables.

and the previous assumption, the conditional PDF of the n observations at the receiver given θ, denoted by $p(\mathbf{y}|\theta)$, can be expressed as

$$p(\mathbf{y}|\theta) = \prod_{i=1}^{n} p(y_i|\theta) \tag{4}$$

where $\mathbf{y} \triangleq [y_1, y_2, \ldots, y_n]^T$, $p(y_i|\theta) = \gamma\, p_V(y_i - f_1(\theta)) + (1-\gamma)\, p_V(y_i - f_2(\theta))$ and $p_V(x) = \frac{1}{\sqrt{2\pi}\sigma_V} \exp\{-\frac{x^2}{2\sigma_V^2}\}$. Similarly, the conditional PDF of the n observations at the eavesdropper given θ, $p(\mathbf{z}|\theta)$, can be stated as

$$p(\mathbf{z}|\theta) = \prod_{i=1}^{n} p(z_i|\theta) \tag{5}$$

where $\mathbf{z} \triangleq [z_1, z_2, \ldots, z_n]^T$, $p(z_i|\theta) = \gamma\, p_W(z_i - f_1(\theta)) + (1-\gamma)\, p_W(z_i - f_2(\theta))$ and $p_W(x) = \frac{1}{\sqrt{2\pi}\sigma_W} \exp\{-\frac{x^2}{2\sigma_W^2}\}$.

In this setting, the encoder should be designed in such a way that the estimation errors at the eavesdropper or, alternatively, at the intended receiver satisfy the constraints. It is noted that the secrecy capacity in information theory is an asymptotic metric and assumes that $n \to \infty$. In practice, it is also important to investigate how much secrecy can be achieved in the finite regime with a small number of observations. For example, [32] provides new achievability results and converse bounds for the maximal secret communication rate of wiretap channels for a given finite blocklength n. Similarly, we focus on the optimal encoding design in the non-asymptotic region for both small and large numbers of observations in this work.

It is known that the optimal estimator for Bayesian parameter estimation in terms of the MSE metric is the MMSE estimator. However, in most scenarios, the MSE of the optimal MMSE estimator does not have a closed form expression. Therefore, even though the encoding operation can be performed with such an approach by using numerical methods, it does not allow theoretical investigations for achieving intuitive understanding of the parameter encoding problem. It is known that for a large number of observations, the MSE of the MMSE estimator converges to the ECRB [30], and for a small number of observations, the MSE of the LMMSE estimator is a close approximation to the optimal MMSE (see Figs. 2–4 for an illustration). (Note that the LMMSE estimator would actually be the optimal MMSE estimator if the parameter of interest and the observations were jointly Gaussian random variables.) Therefore, instead of the optimal MMSE, the ECRB and the LMMSE estimator will be considered in the rest of the manuscript.

Remark 1: The main reason for employing the MSE metric in both the receiver and the eavesdropper is that we focus on a parameter estimation problem in the Bayesian setting in the presence of an eavesdropper and the MSE metric is widely used in practice with or without secrecy concerns in such problems. For example, estimation theoretic secrecy based on the MSE metric has been considered in various channel scenarios such as Gaussian interference channel [16], multiuser MIMO broadcast channel [28], sensor network systems with eavesdroppers [18] and MIMO Gaussian wiretap channel [29]. In addition to parameter estimation problems, the MSE metric is also utilized to design practical and implementable methods to degrade performance of eavesdroppers for enhancing security as an additional layer.

III. Small Number of Observations

In this section, it is assumed that a small number of observations are available to the intended receiver and the eavesdropper to estimate θ. As motivated in the previous section, both the eavesdropper and the intended receiver are modeled to employ LMMSE estimators for a given number of observations n.

A. Generic Encoding Functions

First, generic encoding functions are considered at the transmitter. To that end, as motivated in [25], the parameter space and the intrinsic constraints on the functions $f_1(\theta)$ and $f_2(\theta)$ are specified as follows:

- $\theta \in \Lambda = [a, b]$.
- $f_k(\theta) \in [a, b]$ for $k = 1, 2$.
- $f_1(\theta)$ and $f_2(\theta)$ are continuous and one-to-one functions.

The LMMSE estimator at the intended receiver can explicitly be written for given observations $\mathbf{y}$ as

$$\hat{\theta}_r = E(\theta) + \mathbf{\Sigma}_{\theta,\mathbf{Y}} \mathbf{\Sigma}_{\mathbf{Y}}^{-1} (\mathbf{y} - E(\mathbf{Y})), \tag{6}$$

and the corresponding MSE can be obtained as

$$\epsilon_r = MSE = Var(\theta) - \mathbf{\Sigma}_{\theta,\mathbf{Y}} \mathbf{\Sigma}_{\mathbf{Y}}^{-1} \mathbf{\Sigma}_{\theta,\mathbf{Y}}^{T}. \tag{7}$$

where $\mathbf{\Sigma}_{\theta,\mathbf{Y}} = [Cov(\theta, Y_1), Cov(\theta, Y_2) \ldots Cov(\theta, Y_n)]$ and $\mathbf{\Sigma}_{\mathbf{Y}} = E((\mathbf{Y} - E(\mathbf{Y}))(\mathbf{Y} - E(\mathbf{Y}))^T)$. Similarly, the MSE of the LMMSE estimator at the eavesdropper, ϵ_e, can be obtained for given observations $\mathbf{z}$ by using $\mathbf{Z}$ instead of $\mathbf{Y}$ in (7). Based on these MSE expressions, the optimization problems can be proposed as follows:

$$\min_{\gamma, f_1(\theta), f_2(\theta)} \epsilon_r \quad \text{s.t.} \quad \epsilon_e \geq \alpha_1 \tag{8}$$

and

$$\max_{\gamma, f_1(\theta), f_2(\theta)} \epsilon_e \quad \text{s.t.} \quad \epsilon_r \leq \alpha_2 \tag{9}$$

where α_1 and α_2 denote, respectively, the secrecy target for the first problem and the estimation accuracy (error) limit at the intended receiver for the second problem. The following proposition provides a closed form expression for the MSE of the LMMSE estimator at the intended receiver.

Proposition 1: The MSE (ϵ_r) of the LMMSE estimator at the intended receiver for the encoding model specified in (3) with given $f_1(\theta)$, $f_2(\theta)$ and γ is

$$\epsilon_r = Var(\theta) - \frac{n\,(\gamma\, c_1 + (1-\gamma)\, c_2)^2}{(n-1)x + \tau - nt} \tag{10}$$

where

$$\begin{aligned} x &\triangleq \gamma^2\, r_1 + (1-\gamma)^2 r_2 + 2\,\gamma\,(1-\gamma) E(f_1(\theta)\, f_2(\theta)) \\ \tau &\triangleq \gamma\, r_1 + (1-\gamma)\, r_2 + \sigma_V^2 \\ t &\triangleq (\gamma\, m_1 + (1-\gamma)\, m_2)^2 \end{aligned} \tag{11}$$

with $m_i = E(f_i(\theta))$, $r_i = E(f_i(\theta)^2)$ and $c_i = Cov(f_i(\theta), \theta)$ for $i = 1, 2$.

Proof: Note that $\mathbf{\Sigma_Y} = E(\mathbf{Y}\,\mathbf{Y}^T) - E(\mathbf{Y})E(\mathbf{Y})^T$. Also, $E(Y_k|\theta) = \gamma\, f_1(\theta) + (1-\gamma)\, f_2(\theta)$. Then, $E(Y_k) = E(E(Y_k|\theta)) = \gamma\, m_1 + (1-\gamma)\, m_2$ for $k = 1, 2, \ldots, n$. Therefore, $E(\mathbf{Y}) = (\gamma\, m_1 + (1-\gamma)\, m_2)\mathbf{1}$, where $\mathbf{1}$ denotes the $n \times 1$ column vector of ones. Thus, $E(\mathbf{Y})E(\mathbf{Y})^T = (\gamma\, m_1 + (1-\gamma)\, m_2)^2 \mathbf{1}\mathbf{1}^T = t\mathbf{1}\mathbf{1}^T$.

In addition, $E(Y_k^2|\theta) = \gamma\,(f_1(\theta)^2 + \sigma_V^2) + (1-\gamma)\,(f_2(\theta)^2 + \sigma_V^2)$; hence, $E(Y_k^2) = \gamma\, r_1 + (1-\gamma)\, r_2 + \sigma_V^2 = \tau$ for $k = 1, 2, \ldots, n$. Similarly, $E(Y_j Y_k|\theta) = E(Y_j|\theta)E(Y_k|\theta) = (\gamma\, f_1(\theta) + (1-\gamma)\, f_2(\theta))^2$. Then, $E(Y_j Y_k) = \gamma^2\, r_1 + (1-\gamma)^2\, r_2 + 2\,\gamma\,(1-\gamma)\, E(f_1(\theta)\, f_2(\theta)) = x$ for $j, k = 1, 2, \ldots, n$ and $j \neq k$. Overall, the value of the diagonal elements of $\mathbf{\Sigma_Y}$ is $\tau - t$ and the rest of the elements are $x - t$.

Furthermore, $\mathbf{\Sigma}_{\theta,\mathbf{Y}} = Cov(\theta, Y_k)\mathbf{1}^T$ and $Cov(\theta, Y_k) = E(\theta Y_k) - E(\theta)E(Y_k)$. Note that $E(\theta\, Y_k) = E(E(\theta Y_k|\theta)) = E(\theta\, E(Y_k|\theta)) = \gamma\, E(\theta f_1(\theta)) + (1-\gamma)E(\theta f_2(\theta))$. Then, $Cov(\theta, Y_k) = \gamma\,(E(\theta f_1(\theta)) - E(\theta)E(f_1(\theta))) + (1-\gamma)(E(\theta f_2(\theta)) - E(\theta)E(f_2(\theta))) = \gamma\, c_1 + (1-\gamma)c_2$. Therefore, the MSE becomes $Var(\theta) - \mathbf{\Sigma}_{\theta,\mathbf{Y}}\mathbf{\Sigma}_{\mathbf{Y}}^{-1}\mathbf{\Sigma}_{\theta,\mathbf{Y}}^T = Var(\theta) - (\gamma\, c_1 + (1-\gamma)\, c_2)^2 \mathbf{1}^T\mathbf{\Sigma}_{\mathbf{Y}}^{-1}\mathbf{1}$. Note that the sum of the elements in each row of $\mathbf{\Sigma_Y}$ is the same; therefore, $\mathbf{\Sigma_Y}\mathbf{1} = \lambda\mathbf{1}$, where $\lambda = (n-1)x + \tau - nt$. As λ is an eigenvalue of $\mathbf{\Sigma_Y}$ with a corresponding eigenvector $\mathbf{1}$, $\mathbf{\Sigma}_{\mathbf{Y}}^{-1}\mathbf{1} = (1/\lambda)\mathbf{1}$ holds. Then, $\mathbf{1}^T\mathbf{\Sigma}_{\mathbf{Y}}^{-1}\mathbf{1} = (1/\lambda)\,\mathbf{1}^T\mathbf{1} = n/\lambda$. Hence, the MSE becomes $Var(\theta) - (\gamma\, c_1 + (1-\gamma)\, c_2)^2 n/\lambda$, and inserting the value of $\lambda = (n-1)x + \tau - nt$ concludes the proof. ■

Proposition 1 provides a tool to calculate the MSE for any given prior information $p_\theta(\theta)$, encoding scheme ($f_1(\theta)$, $f_2(\theta)$, γ) and number of observations n. Note that Proposition 1 can similarly be derived for the eavesdropper by using σ_W^2 instead of σ_V^2 whenever necessary. It can be observed that the MSE in (10) increases when the noise variance increases; therefore, $\epsilon_r < \epsilon_e$ when $\sigma_V^2 < \sigma_W^2$.

It is noted that the optimization problems in (8) and (9) are related such that the expressions for ϵ_r and ϵ_e differ only in the noise variance terms. Therefore, it is possible to find a relationship between the solutions of (8) and (9), as stated in the following proposition.

Proposition 2: Suppose that $\mathcal{S} = \{(\gamma^*, f_1^*, f_2^*)\}$ is the set of optimal solutions to (8). Let the optimal value of (8) be denoted as ϵ_r^*. If α_2 is set as $\alpha_2 = \epsilon_r^*$ in (9), then the optimal solutions of (9) satisfy the constraint in (9) with equality, and $\epsilon_e^\dagger = \max_{(\gamma, f_1, f_2)\in\mathcal{S}} \epsilon_e$, where $\epsilon_e^\dagger$ is the optimal value of (9). Similarly, let $\bar{\mathcal{S}} = \{(\gamma^\dagger, f_1^\dagger, f_2^\dagger)\}$ denote the set of optimal solutions to (9). If $\alpha_1 = \epsilon_e^\dagger$ in (8), then the optimal solutions to (8) satisfy the constraint in (8) with equality, and $\epsilon_r^* = \min_{(\gamma, f_1, f_2)\in\bar{\mathcal{S}}} \epsilon_r$.

Proof: We provide a proof only for the first statement as the second one can be shown in a similar fashion. Let the MSEs of the intended receiver and the eavesdropper be denoted, respectively, as $\epsilon_r = T(\gamma, f_1, f_2, \sigma_V^2)$ and $\epsilon_e = T(\gamma, f_1, f_2, \sigma_W^2)$ for given γ, f_1, and f_2. Suppose that $(\gamma^\dagger, f_1^\dagger, f_2^\dagger)$ is an optimal solution to (9) with $T(\gamma^\dagger, f_1^\dagger, f_2^\dagger, \sigma_V^2) < \alpha_2 = \epsilon_r^*$. Then, $(\gamma^\dagger, f_1^\dagger, f_2^\dagger)$ cannot be in the feasible set of (8) as $\alpha_2 = \min \epsilon_r$ for $\epsilon_e \geq \alpha_1$ in (8), implying that $T(\gamma^\dagger, f_1^\dagger, f_2^\dagger, \sigma_W^2) < \alpha_1$. Note that any $(\gamma^*, f_1^*, f_2^*) \in \mathcal{S}$ satisfies $T(\gamma^*, f_1^*, f_2^*, \sigma_W^2) \geq \alpha_1 > T(\gamma^\dagger, f_1^\dagger, f_2^\dagger, \sigma_W^2)$, which shows that $(\gamma^\dagger, f_1^\dagger, f_2^\dagger)$ cannot be an optimal solution to (9). Therefore, the optimal solution to (9) should satisfy $T(\gamma^\dagger, f_1^\dagger, f_2^\dagger, \sigma_V^2) = \alpha_2 = T(\gamma^*, f_1^*, f_2^*, \sigma_V^2) = \epsilon_r^*$, and it needs to be in $\mathcal{S}$. Hence, the sufficient space to search for the optimal solution of (9) reduces to $\mathcal{S}$, and $\epsilon_e^\dagger = \max_{(\gamma, f_1, f_2)\in\mathcal{S}} \epsilon_e$. ■

The following corollaries immediately follow from Proposition 2.

Corollary 1: If (γ^*, f_1^*, f_2^*) is a unique solution to (8) with the optimal value ϵ_r^*, then it is also a unique solution to (9) for $\alpha_2 = \epsilon_r^*$.

Corollary 2: If all the optimal solutions to (8) satisfy the constraint in (8) with equality, then the optimal value of (9), $\epsilon_e^\dagger$, is equal to α_1 for $\alpha_2 = \epsilon_r^*$.

Corollary 3: If $(\gamma^\dagger, f_1^\dagger, f_2^\dagger)$ is a unique solution to (9) with the optimal value $\epsilon_e^\dagger$, then it is also a unique solution to (8) for $\alpha_1 = \epsilon_e^\dagger$.

Corollary 4: If all the optimal solutions to (9) satisfy the constraint in (9) with equality, then the optimal value of (8), ϵ_r^*, is equal to α_2 for $\alpha_1 = \epsilon_e^\dagger$.

As the optimization problems in (8) and (9) require a search over functions, characterizing the set of optimal solutions in every case may not be possible. However, Proposition 1 provides the required expressions to evaluate the objective and constraint functions for given σ_W^2 and σ_V^2. Based on those expressions, the following proposition provides a closed form expression for an optimal solution to (8) and (9) when the channel of eavesdropper is noisier than that of the intended receiver; that is, $\sigma_W^2 > \sigma_V^2$.

Proposition 3: If $\sigma_W^2 > \sigma_V^2$, an optimal solution to (8) is a deterministic affine function, denoted by $f^*(\theta) = k_1^*\theta + k_2^*$, where

$$k_1^* = \pm\sqrt{\frac{\sigma_V^2}{n}\left(\frac{1}{\alpha_1} - \frac{1}{Var(\theta)}\right)} \tag{12}$$

and k_2^* can be anything as long as $f^*(\theta) \in [a, b]$. Then, the optimal value of (8) is

$$\epsilon_r^* = \frac{\sigma_V^2\, Var(\theta)\, \alpha_1}{\sigma_W^2(Var(\theta) - \alpha_1) + \sigma_V^2\alpha_1}\,. \tag{13}$$

Similarly, an optimal solution to (9) is a deterministic affine function, $f^\dagger(\theta) = k_1^\dagger\theta + k_2^\dagger$, where

$$k_1^\dagger = \pm\sqrt{\frac{\sigma_W^2}{n}\left(\frac{1}{\alpha_2} - \frac{1}{Var(\theta)}\right)} \tag{14}$$

and $k_2^\dagger$ can be anything as long as $f^\dagger(\theta) \in [a, b]$. Then, the optimal value of (9) is

$$\epsilon_e^\dagger = \frac{\sigma_W^2\, Var(\theta)\, \alpha_2}{\sigma_V^2(Var(\theta) - \alpha_2) + \sigma_W^2\alpha_2}\,. \tag{15}$$

Proof: First, we focus on the optimization problem in (9). The denominator of the second term in (10) can be rewritten as $n(x-t) + \tau - x$, where $x - t = Var(\gamma f_1(\theta) + (1-\gamma)f_2(\theta))$ and $\tau - x = \gamma\,(1-\gamma)E(|f_1(\theta) - f_2(\theta)|^2) +$

σ_V^2. Also, the numerator of the second term in (10) can be expressed as $n\,Cov(\gamma f_1(\theta)+(1-\gamma)f_2(\theta),\theta)^2$. Therefore, ϵ_e and ϵ_r become

$$\epsilon_e = Var(\theta) - \frac{n\,Cov(\tilde{f},\theta)^2}{n\,Var(\tilde{f})+\gamma\,(1-\gamma)E\left(|f_1(\theta)-f_2(\theta)|^2\right)+\sigma_W^2}$$

$$\epsilon_r = Var(\theta) - \frac{n\,Cov(\tilde{f},\theta)^2}{n\,Var(\tilde{f})+\gamma\,(1-\gamma)E\left(|f_1(\theta)-f_2(\theta)|^2\right)+\sigma_V^2}$$

respectively, where $\tilde{f} \triangleq \gamma f_1(\theta)+(1-\gamma)f_2(\theta)$. It is noted that unless we have the trivial case of $\tilde{f}=0$, the following equation holds:

$$\frac{\epsilon_r - V}{\epsilon_e - V} = \frac{\Delta+\sigma_W^2}{\Delta+\sigma_V^2}$$

where $V = Var(\theta)$ and $\Delta \triangleq n\,Var(\tilde{f})+\gamma\,(1-\gamma)E(|f_1(\theta)-f_2(\theta)|^2)$. Then, for all feasible $\gamma, f_1(\theta), f_2(\theta)$,

$$\epsilon_e = V-(V-\epsilon_r)\frac{\Delta+\sigma_V^2}{\Delta+\sigma_W^2} \le V-(V-\alpha_2)\frac{\Delta+\sigma_V^2}{\Delta+\sigma_W^2} \le V-(V-\alpha_2)\frac{\Delta^*+\sigma_V^2}{\Delta^*+\sigma_W^2} \tag{16}$$

where $\Delta^* = \min_{\gamma,f_1,f_2}\Delta$ s.t., $\epsilon_r \le \alpha_2$. Note that the first inequality in (16) is due to the fact that $\epsilon_r \le \alpha_2$ in the feasible region, and the second inequality is due to the fact that $(\Delta+\sigma_V^2)/(\Delta+\sigma_W^2)$ is an increasing function of Δ as $\sigma_W^2 > \sigma_V^2$ with $\Delta \ge 0$. As (16) provides a global upper bound for ϵ_e, if there exists a feasible (γ, f_1, f_2) such that ϵ_e attains the global bound, then it is concluded that ϵ_e is maximized with it. A sufficient condition for the existence of such a case is that the solution of $\min_{\gamma,f_1,f_2,\epsilon_r\le\alpha_2}\Delta$ satisfies the constraint with equality, i.e., $\epsilon_r = \alpha_2$. Therefore, we aim to obtain the solution of the following problem:

$$\min_{\gamma,f_1(\theta),f_2(\theta)} nVar(\tilde{f})+\gamma(1-\gamma)E\left(|f_1(\theta)-f_2(\theta)|^2\right) \text{ s.t. } \frac{n\,Cov(\tilde{f},\theta)^2}{n\,Var(\tilde{f})+\gamma\,(1-\gamma)E\left(|f_1(\theta)-f_2(\theta)|^2\right)+\sigma_V^2} \ge V-\alpha_2 \tag{17}$$

Note that for any possible $\tilde{f}$, which is obtained using a feasible (γ, f_1, f_2), there are infinitely many alternative ways of constructing it with other feasible (γ, f_1, f_2)'s. Among all constructions, choosing $\tilde{f} = f_1 = f_2$ yields a smaller objective value and a larger value for the left side of the constraint in (17), implying that it is the optimal selection. Therefore, the problem reduces to

$$\min_{\tilde{f}}\ Var(\tilde{f}) \text{ s.t. } V - \frac{n\,Cov(\tilde{f},\theta)^2}{n\,Var(\tilde{f})+\sigma_V^2} \le \alpha_2 \tag{18}$$

The constraint in (18) can be expressed as

$$\frac{n\left(Var(\theta)Var(\tilde{f})-Cov(\tilde{f},\theta)^2\right)+\sigma_V^2 Var(\theta)}{n\,Var(\tilde{f})+\sigma_V^2} \le \alpha_2$$

Note that $Var(\theta)Var(\tilde{f})-Cov(\tilde{f},\theta)^2 \ge 0$ for any $\tilde{f}$ due to Cauchy-Schwarz inequality. Therefore, $Var(\tilde{f}) \ge \sigma_V^2(Var(\theta)-\alpha_2)/(n\alpha_2)$ for any $\tilde{f}$. This global lower bound can be achieved via $\tilde{f}(\theta) = k_1^\dagger\theta + k_2^\dagger$ with $k_1^\dagger$ being given by (14) and $k_2^\dagger$ being selected as any value to guarantee $\tilde{f}(\theta) \in [a,b]$. It is noted that when (9) is a feasible problem, $|k_1^\dagger| \le 1$. For such an encoding, $\Delta^* = \sigma_V^2(Var(\theta)-\alpha_2)/\alpha_2$ and $\epsilon_r = \alpha_2$, i.e., the constraint is satisfied with equality in (17). Therefore, an optimal solution of (17), which is a deterministic affine function, is also an optimal solution of (9), which yields the optimal value of $\epsilon_e^\dagger = \frac{\sigma_W^2\,Var(\theta)\,\alpha_2}{\sigma_V^2(Var(\theta)-\alpha_2)+\sigma_W^2\alpha_2}$.

Based on the preceding discussion and Corollary 4, it can be argued that an optimal solution to (8) is a deterministic affine function when $\sigma_W^2 > \sigma_V^2$. First, notice that any optimal solution to (9) should satisfy the constraint with equality, i.e., $\epsilon_r = \alpha_2$. This is due to the fact for any other solution which does not satisfy the constraint with equality, the inequality in (16) would strictly be implying a gap between ϵ_e and the global bound, and it is already shown that this bound can actually be achieved. Therefore, the result of Corollary 4 can be applied to connect the solutions of (8) and (9) and to imply that the deterministic affine functions solve (8) as well under the conditions of Proposition 3. Via Corollary 4 and (15), the expression in (13) can be obtained after a rearrangement. ■

There are some interesting observations regarding the result in Proposition 3. First, randomization between two functions does not bring any benefits over deterministic encoding when the intended receiver has already a less noisy channel than the eavesdropper, and the encoding function can be selected as a simple affine function. Second, for a given α_1 (or, α_2) value, ϵ_r^* (and $\epsilon_e^\dagger$) does not depend on n; however, the slope of the deterministic affine optimal function decays with $1/\sqrt{n}$. This means that the transmit power per channel use should be decreased as n increases such that the total transmitted signal power to send θ with n channel uses stays constant. Also, the constant term in the deterministic affine optimal function does not have any effects; hence, it can be chosen freely as long as the function remains in the feasible set.

Even though Proposition 3 provides a closed-form expression for an optimal solution when $\sigma_W^2 > \sigma_V^2$, it does not bring any conclusions into the case of $\sigma_W^2 < \sigma_V^2$. In order to obtain the solutions of the optimization problems in (8) and (9) in this case, the solution methods provided in [25] can be adopted, and ϵ_e and ϵ_r can directly be calculated using (10). In this study, the piecewise linear approximation method described in [25] is utilized to obtain the optimal solutions when $\sigma_W^2 < \sigma_V^2$. In particular, for $f_i(\theta)$, the increment in the kth interval in $[a,b]$ is defined as $\Delta x_k^{(i)} \triangleq f_i(a+k\Delta\theta) - f_i(a+(k-1)\Delta\theta)$ for $k=1,\dots M$, and the optimization is performed over $2M+1$ variables, that is, $[\Delta x_1^{(1)}, \Delta x_2^{(1)}, \dots, \Delta x_M^{(1)}, \Delta x_1^{(2)}, \Delta x_2^{(2)}, \dots, \Delta x_M^{(2)}, \gamma]$, by using the Global Optimization Toolbox of MATLAB. In the

numerical examples, M is taken to be 25, which seems to provide a good trade-off between accuracy and complexity.

Next, we investigate a special case in which the encoding function is restricted to be affine.

B. Affine Encoding Functions

In this section, it is assumed that encoding is performed via affine encoding functions such that $f_1(\theta) = k_1\theta + k_2$ and $f_2(\theta) = s_1\theta + s_2$.[3] For this case, the MSE of the intended receiver (and the eavesdropper by using σ_W^2) can be expressed in terms of k_1, k_2, s_1 and s_2 as a corollary to Proposition 1.

Corollary 5: The MSE (ϵ_r) of the LMMSE estimator at the intended receiver for the encoding model specified in (3) when $f_1(\theta) = k_1\theta + k_2$ and $f_2(\theta) = s_1\theta + s_2$ is

$$\epsilon_r = Var(\theta)\frac{\gamma\,(1-\gamma)\,\kappa + \sigma_V^2}{n\,Var(\theta)(\gamma\,k_1 + (1-\gamma)s_1)^2 + \gamma\,(1-\gamma)\,\kappa + \sigma_V^2} \tag{19}$$

where

$$\kappa \triangleq E\left(((k_1 - s_1)\theta + (k_2 - s_2))^2\right). \tag{20}$$

Proof: For the given f_1 and f_2, c_1 and c_2 defined in Proposition 1 become $k_1\,Var(\theta)$ and $s_1\,Var(\theta)$, respectively. Hence, the numerator of the second term in (10) becomes $n(\gamma\,k_1 + (1-\gamma)\,s_1)^2\,Var(\theta)^2$. Also, the denominator of (10) can be rewritten as $n(x-t) + \tau - x$, where x, τ and t are as defined in (11). Note that $(x-t) = \gamma^2 k_1^2 Var(\theta) + (1-\gamma)^2 s_1^2 Var(\theta) + 2\gamma(1-\gamma)k_1 s_1 Var(\theta) = (\gamma k_1 + (1-\gamma)s_1)^2\,Var(\theta)$, and $\tau - x = \gamma\,(1-\gamma)\kappa + \sigma_V^2$, where κ is as defined in (20). After arranging the terms, the final expression in (19) is obtained. ■

When the encoding functions are restricted to affine functions, the optimization problems in (8) and (9) involve a search over only 5 variables instead of functions. Let $\mathbf{x_a} \triangleq [\gamma, k_1, k_2, s_1, s_2]$ and $T_a(\mathbf{x_a}, \sigma_V^2) \triangleq \epsilon_r$, where ϵ_r is as defined in (19). Then, the optimization problems can be written as

$$\min_{\mathbf{x_a}} T_a(\mathbf{x_a}, \sigma_V^2) \quad \text{s.t.} \quad T_a(\mathbf{x_a}, \sigma_W^2) \geq \alpha_1 \tag{21}$$

$$\max_{\mathbf{x_a}} T_a(\mathbf{x_a}, \sigma_W^2) \quad \text{s.t.} \quad T_a(\mathbf{x_a}, \sigma_V^2) \leq \alpha_2 \tag{22}$$

where $T_a(\mathbf{x_a}, \sigma_W^2) \triangleq \epsilon_e$. It is noted that the optimization problems in (21) and (22) are much easier to solve than those in the case of encoding with generic functions.

Finally, as the closed form expression for the MSE with affine encoding can be calculated based on given encoding coefficients, it is also possible to investigate its behavior as γ changes. Namely, the aim is to provide regions of $\gamma \in [0,1]$ in which the MSE increases or decreases with respect to γ. Such a characterization is helpful for both theoretical analysis and gaining intuition on the benefits of randomization. In addition, it facilitates the specification of the exact optimal solution of γ for the given encoding functions, i.e., k_1, k_2, s_1, s_2, and secrecy target. The following proposition characterizes the behavior of the MSE with respect to γ, where γ is taken as a real number (the case of $\gamma \in [0,1]$ immediately follows as a corollary).

Proposition 4: Define $\nu(\gamma) \triangleq \nu_2\gamma^2 + \nu_1\gamma + \nu_0$ with

$$\begin{aligned}\nu_2 &\triangleq -\kappa(k_1^2 - s_1^2)\\ \nu_1 &\triangleq -2\,\kappa\,s_1^2 - 2\sigma_V^2(k_1 - s_1)^2\\ \nu_0 &\triangleq \kappa s_1^2 - 2\sigma_V^2(k_1 - s_1)s_1\end{aligned} \tag{23}$$

where κ is as defined in (20). Then,

- if $\nu_2 = 0$ and $\nu_1 > 0$, then ϵ_r is an increasing (a decreasing) function of γ for $\gamma > -\nu_0/\nu_1$ ($\gamma < -\nu_0/\nu_1$);
- if $\nu_2 = 0$ and $\nu_1 < 0$, then ϵ_r is a decreasing (an increasing) function of γ for $\gamma > -\nu_0/\nu_1$ ($\gamma < -\nu_0/\nu_1$);
- if $\nu_2 > 0$, then ϵ_r is a decreasing function of γ when γ is in between the roots of $v(\gamma) = 0$, which are $\frac{\kappa s_1 - 2\sigma_V^2(k_1 - s_1)}{\kappa(k_1 + s_1)}$ and $\frac{-s_1}{k_1 - s_1}$, and an increasing function elsewhere;
- if $\nu_2 < 0$, then ϵ_r is an increasing function of γ when γ is in between the roots of $v(\gamma) = 0$, and a decreasing function elsewhere;
- if $\nu_1 = \nu_2 = 0$, then ϵ_r is constant with respect to γ.

Proof: From (19), the MSE can be expressed as $\epsilon_r = Var(\theta)h(\gamma)/(\xi g(\gamma)^2 + h(\gamma))$, where $h(\gamma) = \gamma(1-\gamma)\kappa + \sigma_V^2$, $g(\gamma) = (k_1 - s_1)\gamma + s_1$, and $\xi = nVar(\theta) > 0$. Consider the derivative of the MSE with respect to γ, i.e., $d\epsilon_r/d\gamma$. As the denominator of $d\epsilon_r/d\gamma$ is always positive, it is enough to characterize the sign of its numerator with respect to γ. Let $\hat{v}(\gamma)$ denote the numerator of $d\epsilon_r/d\gamma$.[4] Then,

$$\begin{aligned}\hat{\nu}(\gamma) &= h'(\gamma)\left(\xi g(\gamma)^2 + h(\gamma)\right) - h(\gamma)\left(2\xi g(\gamma)g'(\gamma) + h'(\gamma)\right)\\ &= \xi g(\gamma)\left(h'(\gamma)g(\gamma) - 2h(\gamma)g'(\gamma)\right) \triangleq \xi\, v(\gamma)\end{aligned} \tag{24}$$

where $h'(\gamma) = (1-2\gamma)\kappa$ and $g'(\gamma) = k_1 - s_1$. After inserting these into (24), $\nu(\gamma)$ becomes

$$\begin{aligned}\nu(\gamma) &= ((k_1 - s_1)\gamma + s_1)\\ &\quad \times \left(-\kappa(k_1 + s_1)\gamma + \kappa s_1 - 2\sigma_V^2(k_1 - s_1)\right)\\ &= \nu_2\gamma^2 + \nu_1\gamma + \nu_0\end{aligned} \tag{25}$$

where ν_2, ν_1, and ν_0 are as given in (23). As the roots of $v(\gamma)$ are $\frac{\kappa s_1 - 2\sigma_V^2(k_1 - s_1)}{\kappa(k_1 + s_1)}$ and $\frac{-s_1}{k_1 - s_1}$, the conclusions in the proposition can be obtained by applying the sign test to $v(\gamma)$. ■

The result in Proposition 4 can be used to find the optimal γ directly when k_1, k_2, s_1 and s_2 are fixed. For example, consider a scenario with a single observation ($n = 1$), $\sigma_V = 0.01$, $\sigma_W = 0.5$, and a secrecy target of $\alpha_1 = 0.08$. If $f_1(\theta) = \theta$ and $f_2(\theta) = 1 - \theta$, where θ is uniformly distributed in [0,1], then $\nu_2 = 0$ and $\nu_1 < 0$ with $-\nu_0/\nu_1 = 1/2$ for both ϵ_r and ϵ_e. Therefore, when $\gamma > 1/2$, the MSE is a decreasing function of γ and when $\gamma < 1/2$ it is an increasing function of γ according to Proposition 4. Due to the symmetry in this specific problem, it is possible to restrict γ to $\gamma \in [0, 1/2]$. Therefore, when γ increases, the MSEs (both ϵ_r and ϵ_e) increase monotonically until $\gamma = 1/2$, as well. As the goal is to minimize ϵ_r, it is

[3] k_1 and k_2 should be such that $k_1\theta + k_2 \in [a,b]$ for all $\theta \in [a,b]$. Similarly, $s_1\theta + s_2$ needs to be in $[a,b]$ for all $\theta \in [a,b]$. Note that this requires $|k_1| \leq 1$ and $|s_1| \leq 1$.

[4] The $Var(\theta)$ term is omitted in the expression as it is always positive.

obvious that γ should be increased until it yields $\epsilon_e = \alpha_1 = 0.08$ but no more. Finally, $\gamma = 0.3$ can be obtained as the optimal probability, and the corresponding MSE at the intended receiver becomes $\epsilon_r = 0.07$.

IV. Large Number of Observations

In this section, it is assumed that a large number of observations are available to the intended receiver and the eavesdropper to estimate θ.[5] As motivated in Section II, the ECRB metric is employed for both the intended receiver and the eavesdropper in this scenario. The constraints on the parameter space and the encoding functions are the same as in the previous section.

The ECRB is defined as the expectation of the conditional CRB with respect to the unknown parameter [30], which is expressed as

$$E_\theta((I^{(n)}(\theta))^{-1}) = \int_a^b p_\theta(\theta)\frac{1}{I^{(n)}(\theta)}\,d\theta \triangleq ECRB \tag{26}$$

where $p_\theta(\theta)$ is the prior PDF of θ, $I^{(n)}(\theta)^{-1}$ corresponds to the conditional CRB for estimating θ and $I^{(n)}(\theta)$ denotes the Fisher information based on n observations. Therefore, for the intended receiver, $I_r^{(n)}(\theta)$ can be expressed as

$$I_r^{(n)}(\theta) = \int \left(\frac{\partial \log p(\mathbf{y}|\theta)}{\partial\theta}\right)^2 p(\mathbf{y}|\theta)\,d\mathbf{y} \tag{27}$$

with $p(\mathbf{y}|\theta)$ representing the conditional PDF of the n observations for a given value of θ [33]. Also, due to (4), $I_r^{(n)}(\theta) = nI_r(\theta)$, where $I_r(\theta)$ is the Fisher information based on $p(y|\theta) = \gamma\, p_V(y - f_1(\theta)) + (1-\gamma)\, p_V(y - f_2(\theta))$. Therefore,

$$I_r(\theta) = \int_{-\infty}^{\infty} \frac{u(\theta)^2}{p(y|\theta)}\,dy \tag{28}$$

where

$$\begin{aligned} u(\theta) = {} & \gamma\frac{1}{\sqrt{2\pi}\sigma_V} e^{\frac{-(y-f_1(\theta))^2}{2\sigma_V^2}} \frac{(y-f_1(\theta))}{\sigma_V^2} f_1'(\theta) \\ & + (1-\gamma)\frac{1}{\sqrt{2\pi}\sigma_V} e^{\frac{-(y-f_2(\theta))^2}{2\sigma_V^2}} \frac{(y-f_2(\theta))}{\sigma^2} f_2'(\theta) \end{aligned} \tag{29}$$

and

$$p(y|\theta) = \frac{\gamma}{\sqrt{2\pi}\sigma_V} e^{\frac{-(y-f_1(\theta))^2}{2\sigma_V^2}} + \frac{1-\gamma}{\sqrt{2\pi}\sigma_V} e^{\frac{-(y-f_2(\theta))^2}{2\sigma_V^2}} \tag{30}$$

In addition, when (28) is employed in (26), the ECRB at the intended receiver, E_r, is obtained as

$$E_r = \frac{1}{n}\int_a^b p_\theta(\theta)\frac{1}{I_r(\theta)}\,d\theta\,. \tag{31}$$

Similarly, the ECRB at the eavesdropper can be obtained by defining Fisher information $I_e(\theta)$ based on $p(z|\theta) = \gamma\, p_W(z - f_1(\theta)) + (1-\gamma)\, p_W(z - f_2(\theta))$, which can be calculated as in (28)–(30). Then, the ECRB at the eavesdropper, E_e, is

$$E_e = \frac{1}{n}\int_a^b p_\theta(\theta)\frac{1}{I_e(\theta)}\,d\theta\,. \tag{32}$$

Therefore, similarly to (8) and (9), the optimization problems can be proposed as follows:

$$\min_{\gamma, f_1(\theta), f_2(\theta)} E_r \quad \text{s.t.} \quad E_e \geq \eta_1 \tag{33}$$

$$\max_{\gamma, f_1(\theta), f_2(\theta)} E_e \quad \text{s.t.} \quad E_r \leq \eta_2 \tag{34}$$

where η_1 and η_2 denote the secrecy target for the first problem and the estimation accuracy limit at the intended receiver for the second problem. Even though the simplification to (28) may not be possible for the generic case, calculating the ECRB is still easier and more practical for a large number of observations than calculating the MSEs of estimators such as the MAP or MMSE estimators.

Remark 2: Similarly to the results in Proposition 2 and Corollary 1–4, the exact relationship between the solutions of (33) and (34) can be obtained based on a similar approach, which is not repeated here for brevity.

It is noted that if the encoding function is deterministic, then simplification is possible for both E_r and E_e. The following proposition provides the solutions to the optimization problems in (33) and (34) in the absence of randomization.

Proposition 5: Suppose that a deterministic encoding function $f(\theta)$ is employed at the transmitter. For a given feasible secrecy target η_1, the optimal value of the optimization problem in (33) is $\eta_1\,\sigma_V^2/\sigma_W^2$. Furthermore, any $f(\theta)$ with $(\sigma_W^2/n)\int_a^b p_\theta(\theta)/f'(\theta)^2 d\theta = \eta_1$ is an optimal deterministic encoding function for (33). Similarly, for a given estimation accuracy limit η_2, the optimal value of the optimization problem in (34) is $\eta_2\,\sigma_W^2/\sigma_V^2$. Furthermore, any $f(\theta)$ with $(\sigma_V^2/n)\int_a^b p_\theta(\theta)/f'(\theta)^2 d\theta = \eta_2$ is an optimal deterministic encoding function for (34).

Proof: When a deterministic encoding function $f(\theta)$ is employed at the transmitter, $I_r(\theta)$ in (28) simplifies to $I_r(\theta) = f'(\theta)^2/\sigma_V^2$ [25]. Similarly, $I_e(\theta) = f'(\theta)^2/\sigma_W^2$. Then, the optimization problem in (33) becomes

$$\begin{aligned} \min_{f(\theta)} \quad & \frac{\sigma_V^2}{n}\int_a^b p_\theta(\theta)\frac{1}{f'(\theta)^2}\,d\theta \\ \text{s.t.} \quad & \frac{\sigma_W^2}{n}\int_a^b p_\theta(\theta)\frac{1}{f'(\theta)^2}\,d\theta \geq \eta_1\,. \end{aligned} \tag{35}$$

As the integral term is identical in both the objective and the constraint functions, the argument in Proposition 5 follows by choosing an encoding function that satisfies the constraint with equality. The result for (34) can be justified similarly. ■

Proposition 5 shows that if there is no randomization in the encoding function, then the ratio of E_r/E_e depends only on the noise variances in the channels of the eavesdropper and the intended receiver. Therefore, any deterministic encoding function can be used at the transmitter as long as it satisfies the constraints. Also, it is noted that the only difference between using a generic deterministic encoding function and an affine

[5] It should be emphasized that the ECRB approaches the MSE of the MMSE estimator in the asymptotic region, which refers to either a large number of observations or high SNR/SINR scenarios [30]. When stochastic encoding is employed, there exists a certain interference term in the received signal limiting the effective SINR. Therefore, the ECRB metric is not reliable for a small number of observations even for a small noise variance.

deterministic encoding function is that the former may support a larger set of feasible η_1 (or, η_2) values.

Finally, it is possible to obtain some theoretical and intuitive results for the generic stochastic encoding scheme in (3) by using the convexity of the Fisher information with respect to the conditional distribution [34]. Specifically, let the Fisher information based on $p_1(y|\theta)$ and $p_2(y|\theta)$ be denoted by $I_1(\theta)$ and $I_2(\theta)$, respectively. If $p_3(y|\theta) = \gamma p_1(y|\theta) + (1-\gamma)p_2(y|\theta)$, then the Fisher information $I_3(\theta)$ based on $p_3(y|\theta)$ satisfies $I_3(\theta) < \gamma I_1(\theta) + (1-\gamma)I_2(\theta)$ given that $\gamma \in (0,1)$ and $p_1(y|\theta) \neq p_2(y|\theta)$. This implies that $I_3(\theta)$ is also a convex function of γ for any given $\theta \in [a,b]$, and it always remains below the linear line connecting $I_1(\theta)$ and $I_2(\theta)$.

This convexity property is helpful for providing a few intuitive and analytical results. For example, a lower bound for the ECRB can be obtained when $f_1(\theta)$ and $f_2(\theta)$ correspond to affine encoding. To that end, consider the affine encoding scheme described in Section III-B. Then, $I_1(\theta) = k_1^2/\sigma^2$ and $I_2(\theta) = k_2^2/\sigma^2$. Then, $I_3(\theta) < (\gamma k_1^2 + (1-\gamma)\, k_2^2)/\sigma^2 \ \ \forall \theta \in [a,b]$. Therefore, for the ECRB of the intended receiver, it is obtained that $E_r > \frac{\sigma_V^2}{n(\gamma k_1^2 + (1-\gamma)\, k_2^2)}$ and for the ECRB of the eavesdropper, it is obtained that $E_e > \frac{\sigma_W^2}{n(\gamma k_1^2 + (1-\gamma) k_2^2)}$. The following proposition provides a result for symmetric encoding:

Proposition 6: Consider the symmetric mapping with $f_1(\theta) = g(\theta)$ and $f_2(\theta) = g_0 - g(\theta)$ such that $g(\theta) \in [a,b]$ and $g_0 - g(\theta) \in [a,b]$ for all $\theta \in [a,b]$. Then, the ECRB is maximized at $\gamma = 1/2$.

Proof: Let $\gamma = \gamma_0 \in [0,1]$. For the given model, $I(\theta) = g'(\theta)^2 \int_{-\infty}^{\infty} \hat{u}(\theta)^2/p(y|\theta)\, dy$, where

$$\begin{aligned}\hat{u}(\theta) &= \gamma_0 \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(y-g(\theta))^2}{2\sigma^2}} \frac{(y-g(\theta))}{\sigma^2} \\ &\quad - (1-\gamma_0)\frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(y+g(\theta)-g_0)^2}{2\sigma^2}} \frac{(y+g(\theta)-g_0)}{\sigma^2} \\ &\triangleq m(y,\theta,\gamma_0) \end{aligned} \tag{36}$$

and

$$\begin{aligned} p(y|\theta) &= \gamma_0 \frac{1}{\sqrt{2\pi}\sigma_V} e^{\frac{-(y-g(\theta))^2}{2\sigma_V^2}} \\ &\quad + (1-\gamma_0)\frac{1}{\sqrt{2\pi}\sigma_V} e^{\frac{-(y+g(\theta)-g_0)^2}{2\sigma_V^2}} \triangleq d(y,\theta,\gamma_0). \end{aligned} \tag{37}$$

If the change of variables with $g_0 - y = \hat{y}$ is applied in the integration for $I(\theta)$, it is obtained that $I(\theta) = g'(\theta)^2 \int_{-\infty}^{\infty} \frac{m(\hat{y},\theta,1-\gamma_0)^2}{d(\hat{y},\theta,1-\gamma_0)}\, d\hat{y}$. Therefore, $I(\theta)$ attains the same value for $\gamma = \gamma_0$ and $\gamma = 1-\gamma_0$; hence, it is a symmetric function of γ around $\gamma = 1/2$ for any $\theta \in [a,b]$. Due to this fact and the convexity of $I(\theta)$ with respect to γ, its minimum occurs at $\gamma = 1/2$ for all $\theta \in [a,b]$, implying that the ECRB is maximized at $\gamma = 1/2$. ■

Finally, the behavior of the ECRB with respect to γ can be investigated for the general encoding scheme in (3) based on the convexity property, as stated in the following proposition. (Similar results can also be derived for $I_e(\theta)$.)

Proposition 7: Let $\frac{dI_r(\theta)}{d\gamma}|_{\gamma=0^+} \triangleq d_0$ and $\frac{dI_r(\theta)}{d\gamma}|_{\gamma=1^-} \triangleq d_1$. Then,

- if $d_1 < 0$ for all $\theta \in [a,b]$, $I_r(\theta)$ is monotone decreasing with γ, implying that the ECRB is monotone increasing with $\gamma \in (0,1)$;
- if $d_0 > 0$ for all $\theta \in [a,b]$, $I_r(\theta)$ is monotone increasing with γ, implying that the ECRB is monotone decreasing with $\gamma \in (0,1)$;
- if $d_0 < 0$ and $d_1 > 0$ for a given $\theta \in [a,b]$, $I_r(\theta)$ has a minimum $\gamma^* \in (0,1)$. Furthermore, if γ^* minimizes $I_r(\theta)$ for all $\theta \in [a,b]$, then E_r is maximized at $\gamma = \gamma^*$

Proof: Due to the strict convexity of $I_r(\theta)$ with respect to γ, $\frac{d^2 I_r(\theta)}{d\gamma^2} > 0$ holds for $\gamma \in (0,1)$. If $d_1 < 0$ for all $\theta \in [a,b]$, then $\frac{dI_r(\theta)}{d\gamma} < 0$ for all $\gamma \in (0,1)$ as the value of the derivative only increases as γ increases. Hence, $I_r(\theta)$ is a monotone decreasing function of γ for all $\theta \in [a,b]$, which implies that E_r is monotone increasing. Similarly, if $d_0 > 0$ for all $\theta \in [a,b]$, $\frac{dI_r(\theta)}{d\gamma} > 0$ for all $\gamma \in (0,1)$; hence, $I_r(\theta)$ is a monotone increasing function of γ for all $\theta \in [a,b]$, which implies that E_r is monotone decreasing. Finally, if $d_0 < 0$ and $d_1 > 0$, then via a similar argument, there exists a $\gamma = \gamma^*$ such that $\frac{dI_r(\theta)}{d\gamma}|_{\gamma=\gamma^*} = 0$, and it is the minimum for $I_r(\theta)$, and the rest of the arguments in the proposition follow from (31). ■

The following point should be noted related to γ^* in Proposition 7. Even though there may not exist such a γ^* which is the minimum for all $\theta \in [a,b]$ in general, E_r can still have a maximizer in $\gamma \in (0,1)$. Hence, it is only a sufficient condition, and the symmetric mapping given in Proposition 6 is an example in which this condition is satisfied.

Remark 3: The monotonicity results are important to gain intuition about the benefits of randomization and provide a practical tool and guide to obtain the optimal value of γ for given functions $f_1(\theta)$ and $f_2(\theta)$. For example, if the designer fixes the encoding functions to decrease system complexity, then the problem reduces to finding the optimal γ to satisfy the secrecy targets. (In some other scenarios, it may help reduce the search space.) However, in order to obtain the solutions of the optimization problems in (33) and (34) in general, similarly to the previous section, the piecewise linear approximation method described in [25] can be utilized, and E_e and E_r are calculated based on (26)–(32).

Remark 4: Even though the ECRB metric is also utilized in [25], the current problem setup is significantly different as it considers encoder randomization, multiple observations ($n > 1$), and the availability of encoding information at the eavesdropper. ECRB is only an optimization metric for the performance of the estimator at the receiver in [25], i.e., optimizing it *implies* improved overall performance. However, in this study, ECRB is used only when n is sufficiently large; hence, it is rather directly a tight approximation of the optimal MSE value in the asymptotic region. Also, in [25], different metrics are utilized in the receiver (ECRB) and the eavesdropper (MSE of LMMSE estimator) whereas in this section, ECRB is utilized both in the intended receiver and the eavesdropper. Due to these reasons, most of the theoretical discussions in [25] cannot be applied to the current study.

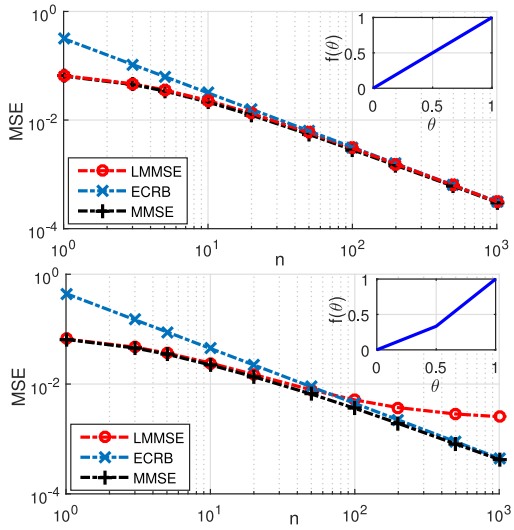

Fig. 2. ECRB, LMMSE and MMSE versus n for two simple encoding scenarios.

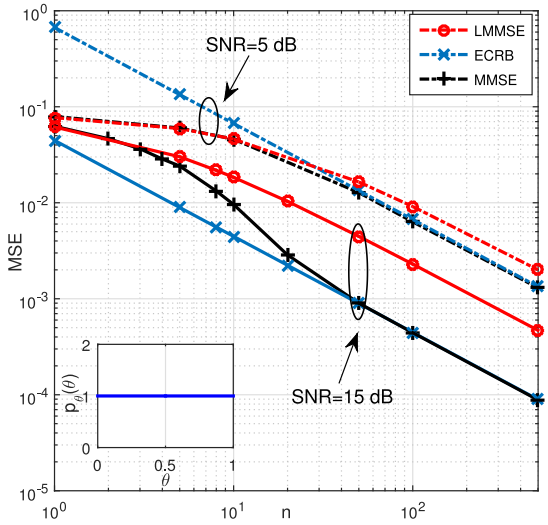

Fig. 3. ECRB, LMMSE and MMSE versus n, where θ has uniform distribution in [0,1].

V. Numerical Results

In this section, numerical examples are provided to investigate the theoretical results and the solution of the optimization problems proposed in Sections III and IV.

A. Justification for LMMSE Estimator and ECRB Metric

In this section, we provide numerical examples to illustrate the motivation behind using different approaches for the cases of small and large numbers of observations. In all examples, the corresponding ECRB and the MSEs for the MMSE and LMMSE estimators are plotted versus the number of observations n. The SNR is defined as $10\log_{10}(1/\sigma^2)$, where σ^2 is the variance of the zero-mean Gaussian noise. In the first example, we consider a simple scenario in which the parameter is not encoded, i.e., $f(\theta)=\theta$. In the second example, the parameter is encoded by a simple piecewise linear deterministic encoding function such that $f(\theta)=2\theta/3$ for $\theta\in[0,0.5]$ and $f(\theta)=(4\theta-1)/3$ for $\theta\in[0.5,1]$. In both examples, it is assumed that θ has uniform distribution in $\theta\in[0,1]$ and the SNR is set to 5 dB. The results are shown in Fig. 2 (top and bottom figures), and the corresponding encoding functions are provided in the upper right corner of each figure. It is observed that the MSEs of the LMMSE and MMSE estimators are close to each other when n is small whereas the ECRB converges to the MSE of the MMSE estimator for large n values in both figures. In the absence of encoding, the MSE performance of the MMSE and LMMSE estimators is almost the same for large numbers of observations, as well. However, the performance of the LMMSE estimator deviates from that of the MMSE estimator and the ECRB for large numbers of observations in the second example (with nonlinear encoding function), which motivates the use of ECRB in this regime in the general case. It is also noted that the ECRB is not a lower bound, and it rather identifies the optimal estimator behavior in asymptotic scenarios.

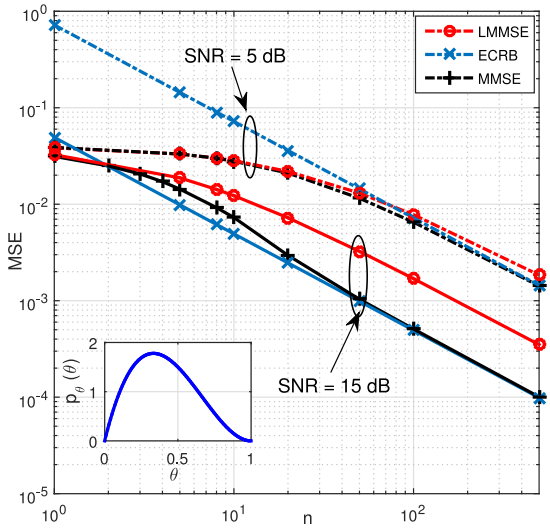

Fig. 4. ECRB, LMMSE and MMSE versus n, where θ has beta distribution with parameters (2,3) in [0,1].

Next, we provide two numerical examples in Figs. 3 and 4 under stochastic encoding as modeled in (3). In both of the examples, it is assumed that $\gamma=0.8$, $f_1(\theta)=\theta$, and $f_2(\theta)=1-\theta$ and $\theta\in[0,1]$. Also, θ has uniform distribution in Fig. 3, and beta distribution with parameters (2,3), i.e., $p_\theta(\theta)=12\,\theta\,(1-\theta)^2$, in Fig. 4. It is observed that for both SNR values in the figures, the MSE of the LMMSE estimator and the ECRB are close to the MSE of the MMSE estimator when n is small and large, respectively.[6] Another important observation is that as the noise variance decreases, the ECRB also reduces rapidly. For small

[6]At high SNRs, the MSE of the MMSE estimator may be in between the ECRB and the MSE of the LMMSE estimator for medium values of n; hence, a more conservative approach can be taken and the ECRB can be used for the eavesdropper and the LMMSE metric can be used for the intended receiver in such a case.

values of n, the ECRB cannot capture the interference effect on the error due to the randomization employed in the encoder, and it can yield optimistic values for the MSE, which motivates the use of the LMMSE estimator in such scenarios. On the other hand, there is a performance gap between the LMMSE and MMSE estimators for large values of n. This is due to the fact that practical estimators start correctly deciding which mode of encoding (f_1 or f_2) is employed with larger observations. However, the LMMSE is unable to achieve such a decision, motivating the use of the ECRB in such scenarios as it is very tight in that region. Therefore, the LMMSE estimator and the ECRB can be utilized for small and large numbers of observations, respectively, at both the receiver and the eavesdropper.

Note that the MMSE solutions in these examples are obtained based on the following approach: For a given θ, n–dimensional realizations $\mathbf{y}$ are obtained empirically at each run of Monte-Carlo simulations, and the conditional MSE is obtained. Then, the MMSE estimator $\hat{\theta}(\mathbf{y}) = E(\theta|\mathbf{Y} = \mathbf{y})$ is analytically calculated for a given $\mathbf{y}$ at each run. Finally, the MSE is obtained by taking the expectation of the conditional MSE over $p_\theta(\theta)$ analytically. The total number of Monte-Carlo runs is set to 10^5.

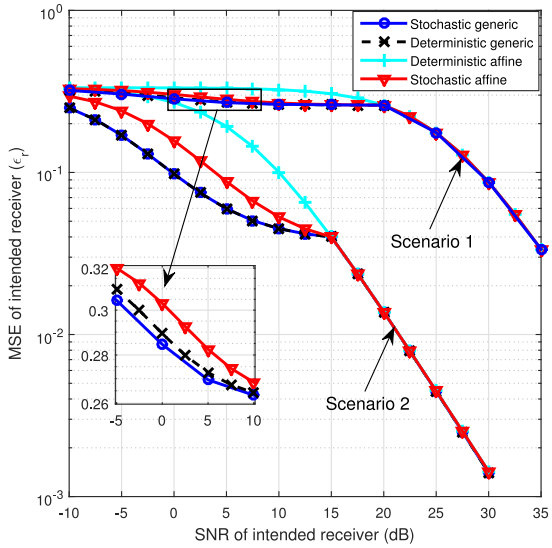

Fig. 5. MSE of intended receiver (ϵ_r) versus SNR of intended receiver for two different scenarios.

B. Small Number of Observations

In this section, numerical results are provided for the case of small number of observations. In all of the examples in this section, it is assumed that the number of observations is 5, i.e., $n = 5$, and θ is uniformly distributed in [0,2]. The SNRs of the intended receiver and the eavesdropper are defined as $10\log_{10}(1/\sigma_V^2)$ and $10\log_{10}(1/\sigma_W^2)$, where σ_V^2 and σ_W^2 are the variances of the zero-mean Gaussian noise at each observation of the intended receiver and the eavesdropper, respectively. The following strategies are evaluated in the examples:

Stochastic generic: This strategy corresponds to the solution of (8) (and alternatively (9)), which provides optimal generic encoding functions $f_1(\theta)$ and $f_2(\theta)$, and the probability γ.

Stochastic affine: This strategy corresponds to the solution of (21) (and alternatively (22)), which provides the optimal affine encoding functions $f_1(\theta) = k_1\theta + k_2$ and $f_2(\theta) = s_1\theta + s_2$, and the probability γ.

Deterministic generic: This strategy corresponds to the solution of (8) (and alternatively (9)) when a deterministic generic encoding function $f(\theta)$ is employed at the transmitter.

Deterministic affine: This strategy corresponds to the solution of (21) (and alternatively (22)) when a deterministic encoding function $f(\theta) = k_1\theta + k_2$ is employed at the transmitter.

First, we consider the minimization of the MSE at the intended receiver for a given secrecy level at the eavesdropper, i.e., the optimization problems in (8) and (21).

In the first example, two different scenarios are considered, and the MSE of the intended receiver is plotted versus the SNR of the intended receiver. In Scenario 1, the SNR of the eavesdropper is 20 dB, and the secrecy target $\alpha_1 = 0.26$ and in Scenario 2, the SNR of the eavesdropper is 15 dB, and the secrecy target $\alpha_1 = 0.04$. In Fig. 5, it is observed that when the SNR of the intended receiver is higher than the SNR of the eavesdropper, all strategies yield the same performance in both scenarios. This result is actually proved formally in Proposition 3, and the optimal value for the MSE of the intended receiver can be achieved by using a simple deterministic affine function. For example, when the SNR of the intended receiver is 30 dB, $f(\theta) = 0.013\,\theta$ is an optimal encoder for Scenario 1, yielding $\epsilon_r^* = 0.0872$, and $f(\theta) = 0.0663\,\theta$ is an optimal encoder for Scenario 2, yielding $\epsilon_r^* = 0.0014$ according to (12) and (13). It is also observed in Fig. 5 that when the SNR of the intended receiver is lower than that of the eavesdropper, there is a performance gap between different strategies. In that region, the deterministic affine functions perform worse than the other strategies, and applying randomization to affine functions brings significant performance gains. Also, the generic functions yield lower MSE values than affine functions. In Scenario 1, stochastic generic functions bring a small performance gain over deterministic generic functions. However, stochastic and deterministic generic functions yield the same performance in Scenario 2, implying that randomization is not necessary if a generic function is employed in that scenario. Also, the MSE of the intended receiver is equal to α_1 for all strategies when the SNRs of the intended receiver and the eavesdropper are the same.

In Fig. 6, the MSE of the intended receiver is plotted versus the secrecy target at the eavesdropper when the SNRs of the eavesdropper and the intended receiver are 15 and 5 dB, respectively. Obviously, as the secrecy target becomes larger, the MSE of the intended receiver increases, as well. When the secrecy target is very small (≈ 0) or very ambitious ($\approx Var(\theta)$), all the strategies have similar performance. For medium values of α_1, it is observed that the deterministic affine function strategy performs significantly worse than the other strategies. However, the stochastic affine strategy has significantly closer performance to that of generic functions. When α_1 is less than 0.24, randomization does not bring any improvements over the deterministic generic strategy. However, as α_1 gets larger (that

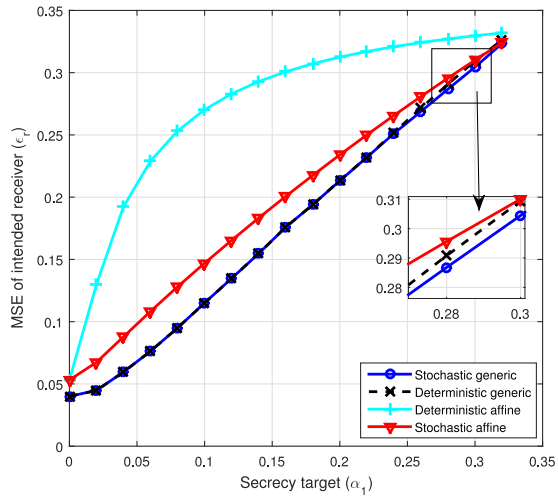

Fig. 6. MSE of intended receiver (ϵ_r) versus secrecy target (α_1) when SNRs of eavesdropper and intended receiver are 15 and 5 dB, respectively.

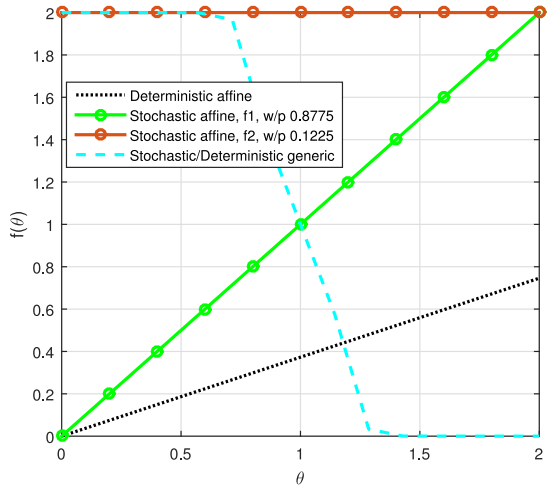

Fig. 8. Optimal encoding functions for different strategies when SNRs of eavesdropper and intended receiver are 15 and 5 dB, respectively, and secrecy target α_1 is 0.04.

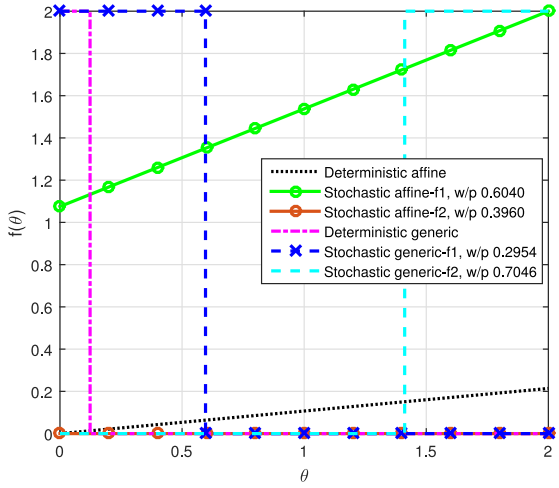

Fig. 7. Optimal encoding functions for different strategies when SNRs of eavesdropper and intended receiver are 10 and 0 dB, respectively, and secrecy target α_1 is 0.28.

is, a relatively large MSE is required at the eavesdropper), stochastic generic functions have slightly better performance than deterministic ones. This implies that it is not possible to claim that deterministic generic functions are an optimal class of functions in all settings even though their performance is not far from that of stochastic generic functions.

In Fig. 7, the optimal encoding functions for different strategies are plotted when the SNRs of the eavesdropper and the intended receiver are 10 and 0 dB, respectively, and the secrecy target α_1 is 0.28. Some important observations can be made from the figure related to the optimal functions. First, it is noticed that the deterministic affine function maps $\theta \in [0, 2]$ to a smaller interval ($[0, 0.213]$) to solve the optimization problem and has a low degrees of freedom in the mapping operation. On the other hand, the stochastic affine strategy sends an affine function $f_1(\theta) = 0.4625\,\theta + 1.075$ with probability 0.604 and nothing (i.e., $f_2(\theta) \approx 0$) with probability 0.396. Furthermore, the characteristics of the generic functions are quite different from those of the affine functions. The optimal deterministic generic function is $f(\theta) \approx 2$ if $\theta < 0.1232$, and $f(\theta) \approx 0$ otherwise.[7] This implies that the optimal deterministic function actually converges to a non-uniform quantizer such that θ values are mapped to 0 and 2. Furthermore, the stochastic generic function strategy randomizes between two *quantizer-like* generic functions to outperform the optimal deterministic encoding function strategy. The intuition behind such a scheme is that a quantizer-like encoder already assigns ≈ 2 and ≈ 0 for a set of θ values and provides one layer of ambiguity. Then, randomization over these two quantizer-like functions provides an extra layer of ambiguity about the parameter to achieve required secrecy targets for the eavesdropper.

In Fig. 8, the optimal encoding functions for different strategies are plotted when the SNRs of the eavesdropper and the intended receiver are 15 and 5 dB, respectively, and the secrecy target α_1 is 0.04. In this case, the secrecy constraint is not as ambitious as the previous one. Similarly to the previous case, the deterministic affine function maps $\theta \in [0, 2]$ to a smaller interval ($[0, 0.746]$). The stochastic affine approach sends the original value of the parameter with probability 0.8775 but it maps θ to ≈ 2 with probability 0.1225. According to Fig. 5, the deterministic and stochastic affine approaches yield the MSE values of 0.1923 and 0.088, respectively, illustrating the benefits of randomization. In addition, the optimal deterministic generic function (and also the optimal stochastic generic function) has different characteristics than the one in Fig. 7. In particular, it has three different regions; namely, $f(\theta) \approx 2$ for $\theta < 0.57$, $f(\theta) \approx 0$ for $\theta > 1.43$, and $f(\theta)$ decreases monotonically for

[7]Note that the encoding functions are required to be one-to-one functions in this study; therefore, even though they are not allowed to stay constant over an interval, it is easy to make sure that they are arbitrarily close to being constant and still do not violate the one-to-one assumption. Also note that if $f(\theta)$ is an optimal deterministic solution, then $\bar{f}(\theta) = \bar{f}_0 \pm f(\theta)$ is also an optimal solution as long as $\bar{f}(\theta) \in [a, b]$, where $\bar{f}_0$ is a constant.

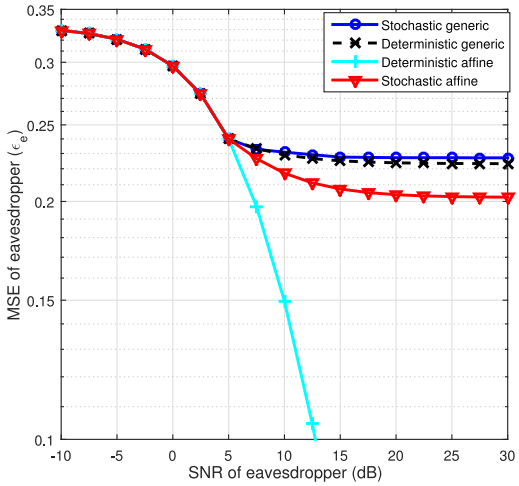

Fig. 9. MSE of eavesdropper (ϵ_e) versus SNR of eavesdropper when SNR of intended receiver is 5 dB, and estimation accuracy limit α_2 is 0.24.

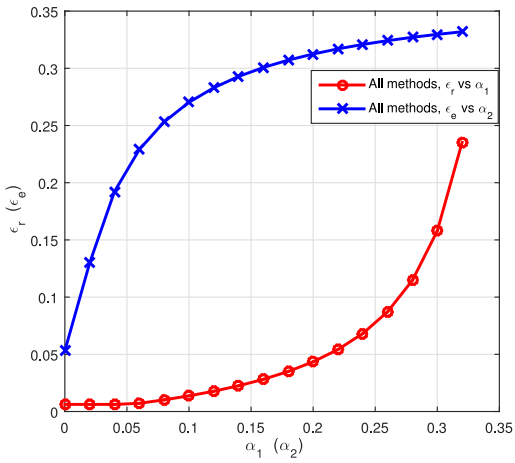

Fig. 10. ϵ_r versus α_1 and ϵ_e versus α_2 when SNRs of eavesdropper and intended receiver are 5 and 15 dB, respectively.

$0.57 \leq \theta \leq 1.43$, yielding an MSE value 0.0597. This implies that when the secrecy target is not very high, the deterministic generic encoding function does not actually behave like a non-uniform quantizer.

Proposition 4 can be utilized to derive the probability values for the stochastic affine strategy theoretically for given affine functions f_1 and f_2. For example, if the parameters of Fig. 8 are used in Proposition 4, it is obtained that $\nu_2 < 0$ for the MSEs of both the eavesdropper and the intended receiver, and according to the root test given in the proposition, the MSE decreases as γ increases when $\gamma \in [0, 1]$. For $\gamma = 0$, ϵ_e is found as $1/3 > \alpha_1 = 0.04$; hence, γ has to be increased until $\epsilon_e = \alpha_1 = 0.04$ to minimize ϵ_r. After some algebra, γ can be obtained as 0.8775.

We also provide an example for the problem of maximizing the MSE at the eavesdropper for a given estimation accuracy limit at the intended receiver (i.e., the optimization problems in (9) and (22)). In Fig 9, the MSE of the eavesdropper is plotted versus the SNR of the eavesdropper when the SNR of the intended receiver is 5 dB and the estimation accuracy limit α_2 is 0.24. It is observed that when the SNR of the eavesdropper is lower than the SNR of the intended receiver, all the solutions have the same performance; that is, using an optimal deterministic affine function is sufficient as claimed in Proposition 3. However, when the SNR of the eavesdropper increases, the MSE of the eavesdropper keeps decreasing for the deterministic affine strategy. Performing randomization over affine functions stops such a decline in the MSE and creates an MSE floor at the eavesdropper. Using generic functions yields even a higher MSE floor, where the stochastic approach performs slightly better than the deterministic one.

Finally, in Fig. 10, the MSE of the intended receiver (ϵ_r) is plotted versus the secrecy target α_1, and the MSE of the eavesdropper (ϵ_e) is plotted versus the estimation accuracy limit α_2 when the SNRs of the eavesdropper and the intended receiver are 5 and 15 dB, respectively. In this scenario, it is already established that all the methods have the same performance. Note that ϵ_r can be kept at relatively low levels for $\alpha_1 < 0.2$; then, it increases rapidly as the secrecy demand becomes more ambitious. Also, ϵ_e increases at a high rate when α_2 is lower than 0.15, but further relaxing the estimation accuracy limit at the intended receiver does not bring significant benefits in terms of the MSE level at the eavesdropper.

It is noted that Proposition 2 and Corollary 1-4 establish the direct relationship between the optimization problems in (8) and (9). Also, based on Proposition 3, it has already been established that the conditions of Corollary 2 and 4 are satisfied when the SNR of the intended receiver is higher than the SNR of the eavesdropper; hence, their results can be applied. This can also be verified in Fig. 10. For example, given a secrecy level of $\alpha_1 = 0.2$, the minimum MSE value at the intended receiver is obtained as $\epsilon_r = 0.043$ after solving (8). Furthermore, for a given estimation accuracy limit of $\alpha_2 = 0.043$, the maximum MSE value at the eavesdropper becomes $\epsilon_e = 0.2$ after solving (9). A similar relationship is also observed when the SNR of the intended receiver is lower than the SNR of the eavesdropper according to Figs. 6 and 9.

C. Large Number of Observations

In this section, the numerical examples are provided for a large number of observations. In all the examples in this section, it is assumed that the number of observations is 1000, i.e., $n = 1000$. Similarly to the previous section, it is assumed that θ is uniformly distributed in [0,2] and the SNRs are defined in the same way. Also, the stochastic generic, stochastic affine and deterministic function strategies are evaluated in a similar fashion. The stochastic generic strategy corresponds to the solution of (33) and alternatively (34). The stochastic affine strategy also solves (33) or (34) with the additional assumption that the encoding functions are affine; that is, $f_1(\theta) = k_1\theta + k_2$ and $f_2(\theta) = s_1\theta + s_2$. Based on Proposition 5, there will be no deterministic affine and deterministic generic strategies separately in this section, and the solution of the deterministic strategy is directly evaluated via Proposition 5.

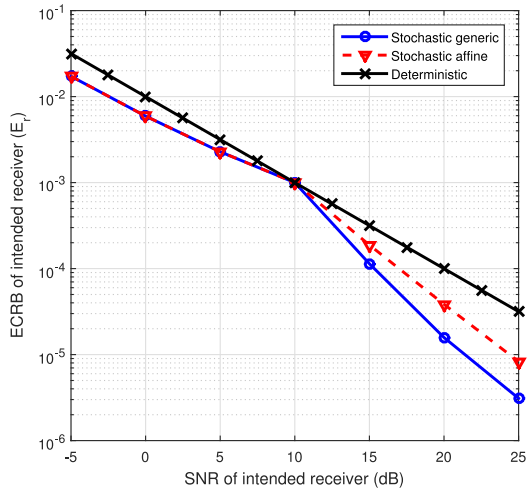

Fig. 11. ECRB of intended receiver (E_r) versus SNR of intended receiver when SNR of eavesdropper is 10 dB, and target secrecy level η_1 is 0.001.

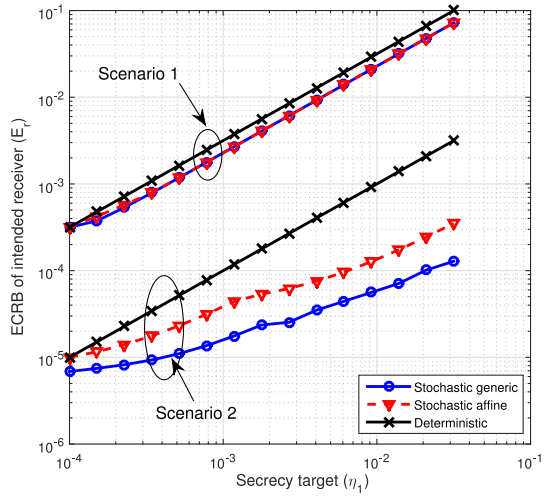

Fig. 13. ECRB of intended receiver (E_r) versus secrecy target (η_1) for two different scenarios.

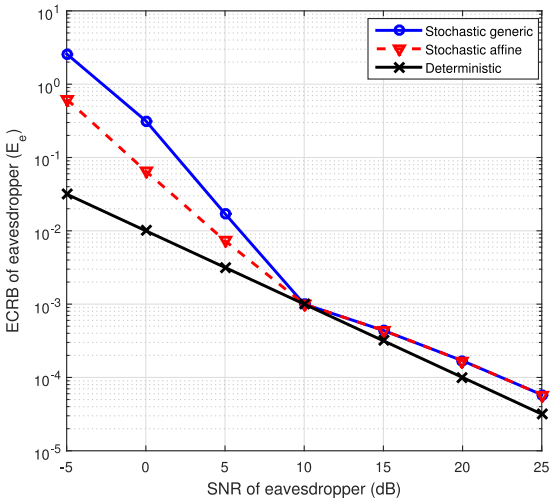

Fig. 12. ECRB of eavesdropper (E_e) versus SNR of eavesdropper when SNR of intended receiver is 10 dB, and estimation accuracy limit η_2 is 0.001.

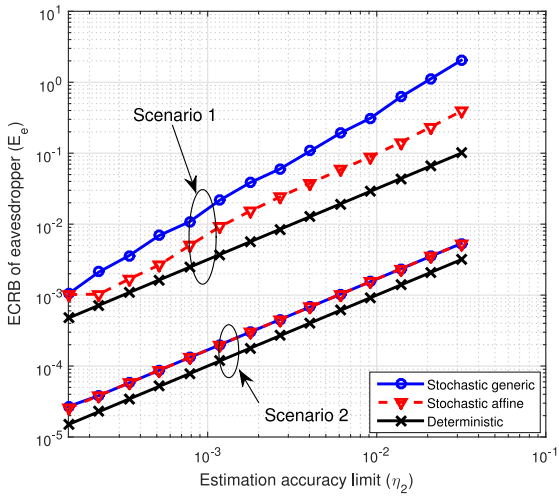

Fig. 14. ECRB of eavesdropper (E_e) versus estimation accuracy limit (η_2) for two different scenarios.

In this part, we consider the minimization (maximization) of the ECRB at the intended receiver (eavesdropper) for a given secrecy level (estimation accuracy limit) at the eavesdropper (intended receiver) in Figs. 11 and 13 (Figs. 12 and 14). First, the ECRB of the intended receiver (eavesdropper) is plotted versus the SNR of the intended receiver (eavesdropper) when the SNR of the eavesdropper (intended receiver) is 10 dB, and the secrecy target $\eta_1 = 0.001$ (and the estimation accuracy limit $\eta_2 = 0.001$). In Fig. 11 (Fig. 12), it is observed that the deterministic functions yield the worst performance and randomization is beneficial at all SNR values of the intended receiver (eavesdropper) for a large number of observations, which was not the case for a small number of observations. Note that the stochastic generic and affine functions have the same performance when the SNR of the intended receiver is lower than that of the eavesdropper. However, the stochastic generic functions outperform the stochastic affine functions when the SNR of intended receiver is higher than the SNR of the eavesdropper. Note that the ECRB versus SNR curve for the deterministic functions is a linear line as explained in Proposition 5. Also, the ECRB of the intended receiver (eavesdropper) is equal to η_1 (η_2) for all the strategies when the SNRs of the intended receiver and the eavesdropper are same.

Next, in Fig. 13 (Fig. 14), the ECRB of the intended receiver (eavesdropper) is plotted versus the secrecy target (estimation accuracy limit) for two different scenarios. In both scenarios, the SNR of the eavesdropper (receiver) is 10 dB and the SNR of the intended receiver (eavesdropper) is 5 and 20 dB in the first and second scenarios, respectively. In the first (second) scenario in Fig. 13 (Fig. 14), the performances of the stochastic strategies are almost the same and they are better than the deterministic solution. Furthermore, in the second (first) scenario in Fig. 13 (Fig. 14), the stochastic generic solution has better performance

than the stochastic affine solution; hence it has the overall best performance. In that case, it is interesting to note that as η_1 (η_2) increases, the performance gap between the stochastic solutions and the simple deterministic solution increases, as well. This shows that randomization can bring significant performance improvements over the deterministic solution in the case of a large number of observations.

Finally, Proposition 6 and 7 can be utilized in the numerical examples to further analyze the results. For example, in Fig. 11, when the SNR of the eavesdropper is 10 dB, and the SNR of the intended receiver is 15 dB, with $\eta_1 = 0.001$, the solution of the optimal stochastic affine encoding strategy is found as $f_1(\theta) = 0.4824\theta + 1.0352$, $f_2(\theta) = 0.9648 - 0.4824\theta$, and $\gamma = 0.5$. Note that according to Proposition 6, this is a symmetrical mapping; therefore, the ECRB of the eavesdropper is maximized at $\gamma = 0.5$. Also, as this encoding function satisfies the secrecy constraint with equality, Proposition 6 implies that other γ values would be infeasible for this particular f_1 and f_2. Also, again in Fig. 11, when the SNR of the intended receiver is 5 dB, the solution of the optimal stochastic affine encoding strategy is found as $f_1(\theta) = 0.4274\,\theta + 0.2597$, $f_2(\theta) = 0.4274\,\theta + 0.8989$, and $\gamma = 0.5$. In order to employ Proposition 7, it can be shown that $d_0 < 0$ and $d_1 > 0$ for all $\theta \in [0, 2]$. Actually, $I(\theta)$ is constant for a given γ for this given f_1 and f_2, and $\gamma = 0.5$ minimizes $I(\theta)$ (as it is constant, basically for all $\theta \in [a, b]$). Therefore, the ECRB of the eavesdropper is maximized at $\gamma = 0.5$.

It is important to mention that the closed-form expressions (e.g., Proposition 1, Corollary 5, and eqns. (26)–(30)) obtained in the theoretical parts (Sections III and IV) are used to calculate the LMMSE and ECRB values in the numerical examples. The performance of the theoretically optimal solutions (e.g., Proposition 3 and 5) is compared with the simulations for verification and the same performance results are obtained. However, the curves are not duplicated in the figures for brevity/clarity of presentation.

D. Computational Complexity

The dimension of the search space and the number of multiplications required to calculate the constraint and objective functions are both important factors about the complexity of the proposed methods. In the case of the stochastic generic function approach, the optimization is performed over $2M + 1$ variables, where M is the number of piecewise regions. For the deterministic generic solutions, the optimization is over M variables. The affine solutions require optimization over five and two variables for the stochastic and deterministic cases, respectively. When Proposition 3 and 5 are utilized, no search is required and the solutions can be obtained directly. Also, the intuition provided by Proposition 4 can reduce the search space to four variables for the stochastic affine solutions in the small number of observations case. For large numbers of observations, the search space for the stochastic generic solution can be reduced to $2M$ based on Proposition 6 when the conditions of the proposition hold.

For small numbers of observations, we use the expressions in Proposition 1 to calculate the MSE. In the calculations, the most costly terms are the expectation terms such as $E(f_1(\theta)\,\theta)$ and $E(f_2(\theta)\,\theta)$. To calculate these terms, which include one-dimensional integrals, one of the possible ways is to employ Riemann sums, each of which includes S terms for a given step size. Then, when the stochastic and deterministic generic functions are used, calculating the objective function requires $O(14S)$ and $O(5S)$ multiplications, respectively. For the affine solutions, we do not have any of these terms, which implies a complexity of $O(1)$. As the only difference between the objective and constraint is the noise variance term, the complexity does not double for calculating both functions. It is important to note that the computational complexity does not depend on n.

For large numbers of observations, the overall expression requires double integration and complexity of $O(14S_1S_2)$, where the Riemann sums have S_1 and S_2 terms. Even though the ECRB calculation is more complex than calculating the MSE of the LMMSE estimator, it also does not depend on n. Note that the optimal MMSE expression would require $n + 1$ integrals instead of two; hence, it is possible to tightly approximate the optimal MSE performance by using the ECRB with a much lower complexity. Finally, when the conditions of Corollaries 1–4 are satisfied, it is possible to connect the optimization problems in (8) and (9) (or, (33) and (34)) so that it is sufficient to solve one of the problems to obtain the solutions of both.

VI. Concluding Remarks

Estimation theoretic secure transmission of a random scalar parameter has been investigated in a Gaussian wiretap channel model, and various constrained optimization problems have been proposed in terms of estimation accuracy performance of the intended receiver and the eavesdropper. The results have shown that for small numbers of observations, when the SNR of the intended receiver is higher than that of the eavesdropper, the deterministic affine solution forms a class of optimal functions, which verifies the theoretical results. When the SNR of the intended receiver is lower than that of the eavesdropper, stochastic generic functions have the best performance in general; however, depending on the target secrecy/accuracy value, deterministic generic functions can provide an optimal solution, as well. Stochastic affine functions can provide significant performance gains over deterministic affine functions, and they can be an attractive alternative solution to generic functions. For large numbers of observations, deterministic generic/affine functions have worse performance than stochastic solutions at all SNRs and in all the considered scenarios. Therefore, stochastic encoding is also attractive in this region of operation. Similarly to the previous case, stochastic generic functions have the best performance in general; however, stochastic affine functions can also provide an optimal solution in certain scenarios. Intuitively, the main factor that determines whether the stochastic methods bring performance gains or not is the quality and quantity of the measurements available to the eavesdropper given the secrecy target. If the eavesdropper has a large number of observations or a small number of observations with a better SNR than the intended receiver, then it is encoder's task to make estimation more challenging for the eavesdropper; hence, stochastic encoding provides performance gains especially in such scenarios. As

a relevant future work, it would be interesting to investigate the MSE-based and information theoretically optimal solutions in a common and fair framework to provide theoretical comparisons and connections.

REFERENCES

[1] W. Diffie and M. Hellman, "New directions in cryptography," *IEEE Trans. Inf. Theory*, vol. IT-22, no. 6, pp. 644–654, Nov. 1976.

[2] R. L. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," *Commun. ACM*, vol. 21, no. 2, pp. 120–126, Feb. 1978.

[3] J. Yang, L. Wang, G. Geraci, M. Elkashlan, J. Yuan, and M. Renzo, "Safeguarding 5G wireless communication networks using physical layer security," *IEEE Commun. Mag.*, vol. 53, no. 4, pp. 20–27, Apr. 2015.

[4] A. Mukherjee, "Physical-layer security in the Internet of Things: Sensing and communication confidentiality under resource constraints," *Proc. IEEE*, vol. 103, no. 10, pp. 1747–1761, Oct. 2015.

[5] Y. Wang, G. Attebury, and B. Ramamurthy, "A survey of security issues in wireless sensor networks," *IEEE Commun. Surveys Tut.*, vol. 8, no. 2, pp. 2–23, Apr.–Jun. 2006.

[6] A. Yener and S. Ulukus, "Wireless physical-layer security: Lessons learned from information theory," *Proc. IEEE*, vol. 103, no. 10, pp. 1814–1825, Oct. 2015.

[7] Y. Liang, H. V. Poor, and S. Shamai, "Secure communication over fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2470–2492, Jun. 2008.

[8] P. K. Gopala, L. Lai, and H. E. Gamal, "On the secrecy capacity of fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4687–4698, Oct. 2008.

[9] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 3, pp. 339–348, May 1978.

[10] H. Weingarten, Y. Steinberg, and S. S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.

[11] Y. Liang, A. Somekh-Baruch, H. V. Poor, S. Shamai, and S. Verdu, "Capacity of cognitive interference channels with and without secrecy," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 604–619, Feb. 2009.

[12] S. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian wire-tap channel," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 4, pp. 451–456, Jul. 1978.

[13] H. Shen, W. Xu, and C. Zhao, "QoS constrained optimization for multi-antenna AF relaying with multiple eavesdroppers," *IEEE Signal Process. Lett.*, vol. 22, no. 12, pp. 2224–2228, Dec. 2015.

[14] W. Liao, T. Chang, W. Ma, and C. Chi, "QoS-based transmit beamforming in the presence of eavesdroppers: An optimized artificial-noise-aided approach," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1202–1216, Mar. 2011.

[15] J. Guo, U. Rogers, X. Li, and H. Chen, "Secrecy constrained distributed detection in sensor networks," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 4, no. 2, pp. 378–391, Jun. 2018.

[16] 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.

[17] F. Farokhi and H. Sandberg, "Fisher information as a measure of privacy: Preserving privacy of households with smart meters using batteries," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 4726–4734, Sep. 2018.

[18] X. Guo, A. S. Leong, and S. Dey, "Estimation in wireless sensor networks with security constraints," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 2, pp. 544–561, Apr. 2017.

[19] J. Guo, H. Chen, and U. Rogers, "Asymptotic perfect secrecy in distributed estimation for large sensor networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Mar. 2017.

[20] 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.

[21] B. Kailkhura, V. S. S. Nadendla, and P. K. Varshney, "Distributed inference in the presence of eavesdroppers: A survey," *IEEE Commun. Mag.*, vol. 53, no. 6, pp. 40–46, Jun. 2015.

[22] 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.

[23] A. N. Samudrala and R. S. Blum, "Asymptotic analysis of a new low complexity encryption approach for Internet of Things, smart cities and smart grid," in *Proc. IEEE Int. Conf. Smart Grid Smart Cities*, Jul. 2017, pp. 200–204.

[24] A. N. Samudrala, "On the estimation and secrecy capabilities of stochastic encryption for parameter estimation in IoT," in *Proc. 52nd Annu. Conf. Inf. Sci. Syst.*, Mar. 2018, pp. 1–6.

[25] 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.

[26] 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.

[27] 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.

[28] 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.

[29] H. Reboredo, J. Xavier, and M. R. D. Rodrigues, "Filter design with secrecy constraints: The MIMO Gaussian wiretap channel," *IEEE Trans. Signal Process.*, vol. 61, no. 15, pp. 3799–3814, Aug. 2013.

[30] H. L. V. Trees and K. L. B . Eds., *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*. Hoboken, NJ, USA: Wiley, 2007.

[31] A. J. Goldsmith, *Wireless Communications*. New York, NY, USA: Cambridge Univ. Press, 2005.

[32] W. Yang, R. F. Schaefer, and H. V. Poor, "Wiretap channels: Nonasymptotic fundamental limits," *IEEE Trans. Inf. Theory*, vol. 65, no. 7, pp. 4069–4093, Jul. 2019.

[33] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York, NY, USA: Springer-Verlag, 1994.

[34] M. Cohen, "The Fisher information and convexity (corresp.)," *IEEE Trans. Inf. Theory*, vol. IT-14, no. 4, pp. 591–592, Jul. 1968.

Cagri Goken (S'10) received the B.S. and M.S. degrees in electrical engineering from Bilkent University, Ankara, Turkey, and the M.A. degree in electrical engineering from Princeton University, Princeton, NJ, USA, in 2009, 2011, and 2014, respectively. He is currently working toward the Ph.D. degree with the Bilkent University. Since 2016, he has been with the Aselsan Inc. Ankara, Turkey, where he is currently a Senior Design Engineer. His research interests include detection and estimation theory, wireless communications, and physical layer secrecy.

Sinan Gezici (S'03–M'06–SM'11) received the B.S. degree from Bilkent University, Ankara, Turkey, in 2001, and the Ph.D. degree in electrical engineering from Princeton UniversityPrinceton, NJ, USA, in 2006. From 2006 to 2007, he was with the Mitsubishi Electric Research Laboratories, Cambridge, MA, USA. Since 2007, he has been with the Department of Electrical and Electronics Engineering, Bilkent University, where he is currently a Professor. His research interests are in the areas of detection and estimation theory, wireless communications, and localization systems. Among his publications in these areas is the book *Ultra-wideband Positioning Systems: Theoretical Limits, Ranging Algorithms, and Protocols* (Cambridge University Press, 2008). He was an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE WIRELESS COMMUNICATIONS LETTERS, and *Journal of Communications and Networks*.