ECRB based Optimal Parameter Encoding under Secrecy Constraints

Çağrı Göken, *Student Member, IEEE,* and Sinan Gezici, *Senior Member, IEEE*

***Abstract*—In this paper, optimal deterministic encoding of a scalar parameter is investigated in the presence of an eavesdropper. The aim is to minimize the expectation of the conditional Cramér-Rao bound (ECRB) at the intended receiver while keeping the mean-squared error (MSE) at the eavesdropper above a certain threshold. First, the optimal encoding function is derived in the absence of secrecy constraints for any given prior distribution on the parameter. Next, the optimization problem is formulated under a secrecy constraint and various solution approaches are proposed. Also, theoretical results on the form of the optimal encoding function are provided under the assumption that the eavesdropper employs a linear minimum mean-squared error (MMSE) estimator. Numerical examples are presented to illustrate the theoretical results and to investigate the performance of the proposed solution approaches.**

***Index Terms*—Parameter estimation, Cramér-Rao bound (CRB), secrecy, optimization.**

I. Introduction and Motivation

Security has been a crucial issue for communications. In a secure communication system, the aim is to secretly transmit secret data to an intended receiver in the presence of an eavesdropper. Cryptographic protocols based on secret keys have extensively been employed to prevent any third parties from extracting secret data [1], [2]. In [3], Shannon proved that the cryptographic approach known as one-time-pad can achieve the perfect secrecy; that is, the original message and the cypher text become independent, if the number of different keys is at least as high as the number of messages. On the other hand, physical layer secrecy relies on the characteristics of the wireless channel and tries to ensure secret communications by exploiting varying channel conditions. In [4], Wyner proved that when the channel between the transmitter and the eavesdropper is a degraded version of the channel between the transmitter and the intended receiver, then reliable communication can be achieved without information leakage to the eavesdropper. One common approach to measure the amount of achieved secrecy is to use information theoretical metrics and tools, such as mutual information, and to examine the highest rates at which the transmitter can encode a message while maintaining a certain equivocation level at the eavesdropper. Following Wyner's work, a multitude of studies have been performed based on this approach for various channel and transmission scenarios [5]–[12]. In the literature, there also exist quality-of-service (QoS) frameworks based on the signal-to-noise ratio (SNR), which is used as a metric for physical layer security [13], [14]. For example, in [14], a cooperative jamming scenario is considered for multiple-input multiple-output (MIMO) broadcast channels with multiple receivers and eavesdroppers, and the optimal friendly jammer strategy is designed to keep signal-to-interference-plus-noise ratio (SINR) at the eavesdroppers below a certain threshold to ensure secrecy.

As a common alternative approach, secrecy levels can be quantified based on estimation theoretic metrics. In this case, the aim is to optimize the estimation accuracy performance of the estimator at the intended receiver, while keeping the minimum mean-squared error (MMSE) at the eavesdropper above a certain target. This setting has been employed in a wide variety of problems [15]–[23]. In [15], the output Y of a channel for a given input X is encoded by a random mapping $P_{Z|Y}$ in order to ensure that the MMSE for estimating Y based on Z is minimized while the MMSE for estimating X based on Z is above $(1-\epsilon)Var(X)$ for a given $\epsilon \geq 0$, where $Var(X)$ denotes the variance of X. In [16], the secret communication problem is considered for Gaussian interference channels in the presence of eavesdroppers. The problem is formulated to minimize the total 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.

Another application area of the estimation theoretic secrecy is distributed inference networks, where the information coming to a fusion center (FC) from various sensor nodes can also be observed by eavesdroppers. The secrecy for distributed detection and estimation can be ensured via various techniques such as design of sensor quantizers and decision rules, stochastic encoding, artificial noise to confuse eavesdroppers, and MIMO beamforming [17]. In [18], the estimation problem of a single point Gaussian source in the presence of an eavesdropper is investigated 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 mean-squared error (MSE) for the parameter of interest while guaranteeing a target MSE at the eavesdropper. Furthermore, in [19], the secrecy problem in a distributed inference framework is investigated in terms of distortion (and secrecy) outage, which is the probability that the MMSE at the FC (eavesdropper) is above (below) certain distortion levels. The optimal transmit power allocation policies are derived to minimize the distortion outage at the FC under an average transmit power and a secrecy outage constraint at the eavesdropper. In [20], stochastic encryption is performed based on the 1-bit quantized version of a noisy sensor measurement to achieve secret communication, where both symmetric and asymmetric bit flipping strategies are

The authors are with the Dept. of Electrical and Electronics Engineering, Bilkent University, Bilkent, Ankara 06800, Turkey, Tel: +90-312-290-3139, Fax: +90-312-266-4192, E-mails: {cgoken,gezici}@ee.bilkent.edu.tr.

considered under the assumptions that the transmitter is aware of the flipping probabilities and the eavesdropper is unaware of the encryption. The effects of the flipping probabilities on the Cramér-Rao bound (CRB) and the maximum likelihood (ML) estimator at the fusion center, and on the bias and the MSE at the eavesdropper are investigated [20]. In [21], 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 house utilizes the batteries to mask the real energy consumption. The Fisher information is employed as a metric of privacy and the optimal policies for the utilization of batteries are derived to minimize the Fisher information to achieve privacy.

For estimation theoretic approaches, the Cramér-Rao bounds provide useful theoretical limits for assessing performance of estimators. It is known that when the parameter to be estimated is non-random, the conditional CRB states that, under some regularity conditions, the MSE of any unbiased estimator is bounded by the inverse of the Fisher information for each given value of the parameter [24]. On the other hand, if the parameter to be estimated is random with a known prior distribution, then the extended versions of the CRB, such as the Bayesian Cramér-Rao bound (BCRB) and the expectation of the conditional Cramér-Rao bound (ECRB), can be employed [25]. Even though the BCRB effectively takes the prior information into account and can provide a useful lower bound for the maximum a-posterior probability (MAP) estimator in the low signal-to-noise ratio (SNR) regime, it does not exist for some prior distributions due to the violation of an assumption in its derivation. For example, the BCRB does not exist when the parameter has a uniform prior distribution over a closed set [25]–[27]. More importantly, when the conditional CRB is a function of the unknown parameter, which is commonly the case, the BCRB does not present a tight bound in the high SNR regime.[1] Therefore, for the parameter encoding problem in this paper, the use of the BCRB as the objective function may be misleading and can result in trivial bounds in some cases. For these reasons, the ECRB is employed in this study, which has widely been utilized in a variety of applications in the literature; e.g., [29]–[31], [42]. The ECRB is known to provide a tight limit for the MAP estimator asymptotically, and converges to the Ziv-Zakai bound (ZZB) in the high SNR regime [25]. Therefore, the optimization of parameter encoding according to the ECRB metric leads to close-to-optimal performance for practical MAP estimators in the high SNR regime. Although the ZZB can provide a tight limit for all SNRs, it has high computational complexity compared to the ECRB [25], [28] and does not allow theoretical investigations for achieving an intuitive understanding of the parameter encoding problem.

In this paper, we consider the transmission of a scalar parameter to an intended receiver in the presence of an eavesdropper. In order to ensure secret communications, we utilize an encoding function (continuous and one-to-one) applied on the original parameter. The aim is to minimize the ECRB at the intended receiver while ensuring a certain MSE target at the eavesdropper. It is assumed that the eavesdropper uses a linear MMSE estimator without being aware of the encoding. An optimization problem is formulated to obtain the optimal encoding function for given target MSE levels. At the first step, the secrecy requirements are omitted and the optimization problem is solved under no constraints. In that case, a closed-form analytical solution is provided for the optimal encoding function for any given prior distribution. Next, the MSE constraint for the eavesdropper is included and various solution approaches, such as polynomial approximation, piecewise linear approximation, and linear encoding are proposed. Also, theoretical results are derived related to the structure of the optimal encoding function under some assumptions. Then, numerical results are provided for both uniform and nonuniform prior distributions. The main contributions in this paper can be summarized as follows:

- The problem of optimal parameter encoding is proposed by considering an ECRB metric at the intended receiver and an MSE target level at the eavesdropper.
- Considering a generic prior distribution, a closed-form expression is derived for the optimal encoding function under no secrecy constraints.
- A closed form expression for $E(|\hat{\beta}(Z)-\theta|^2)$ is provided when the eavesdropper employs the linear MMSE estimator without being aware of the encoding, where $\hat{\beta}(Z)$ is the estimator of the eavesdropper and θ is the true value of the parameter. It is shown that the corresponding ECRB and MSE value do not change if the domain of the function is shifted. It is also proved that if the prior distribution is symmetric on the domain, the search for optimal encoding functions can be limited to decreasing functions. In addition, a closed-form expression is derived for the supremum of $E(|\hat{\beta}(Z)-\theta|^2)$ over all feasible encoding functions when the prior distribution is uniform.
- Three solution approaches are proposed to find the optimal encoding function. The polynomial and piecewise linear approximations are used to calculate the optimal encoding functions numerically, and linear functions are employed to develop a suboptimal encoding scheme. It is shown that the optimal linear encoding function can be obtained simply by finding the roots of a polynomial equation. In addition, solutions are provided based on power functions in the numerical examples.
- Via numerical examples, the optimal ECRB values and encoding functions are obtained based on the proposed approaches for the case of a varying target MSE level when eavesdropper's channel quality is fixed, and for the case of a varying eavesdropper's channel quality when the target MSE level is fixed.

The rest of the paper is organized as follows: The optimal parameter encoding problem is formulated in Section II. Optimal encoding functions with and without secrecy constraints are investigated in Section III. The solution approaches for the optimal encoding problem are proposed in Section IV. The numerical results are presented in Section V, and the concluding remarks are given in Section VI.

[1]This is also a problem for the weighted Cramér-Rao bound (WCRB), which is a generalized version of the BCRB using a weighting function, and can be employed for the cases in which the BCRB does not exist [25], [27].

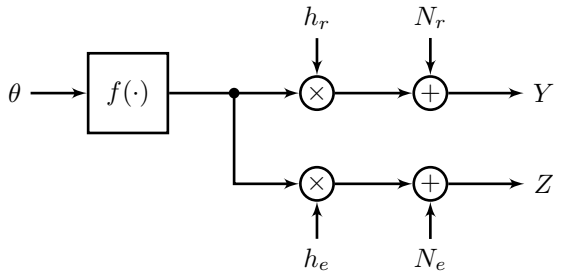

Fig. 1: System model for the parameter encoding problem.

II. Problem Formulation

Consider the transmission of a scalar parameter $\theta \in \Lambda$ to an intended receiver over a noisy and fading channel, where the noise is denoted by N_r and the instantaneous fading coefficient of the channel is denoted by the constant h_r. It is also assumed that there exists an eavesdropper trying to estimate parameter θ. 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. To that aim, the parameter is encoded by a continuous, real valued, and one-to-one function $f : \Lambda \to \Gamma$. Hence, the received signal at the intended receiver can be written as

$$Y = h_r f(\theta) + N_r \tag{1}$$

where N_r is modeled as a zero-mean Gaussian random variable with variance σ_r^2, and N_r and θ are assumed to be independent. On the other hand, the eavesdropper observes

$$Z = h_e f(\theta) + N_e \tag{2}$$

where h_e is the fading coefficient for the eavesdropper, and N_e is zero-mean Gaussian noise with variance σ_e^2, which is independent of θ and N_r. Also, the prior information on parameter θ is represented by a probability density function (PDF) denoted by $w(\theta)$ for $\theta \in \Lambda$. The intended receiver tries to estimate parameter θ based on observation Y whereas the eavesdropper uses observation Z for estimating θ. The system model is illustrated in Fig. 1. It is assumed that the channels are slowly fading; that is, the channel coefficients are constant during the transmission of the parameter.[2]

The following assumptions are made about the eavesdropper's strategy:

- f acts like a secret key between the transmitter and the intended receiver and is not known by the eavesdropper. Hence, the estimator at the eavesdropper actually tries to estimate $f(\theta) \triangleq \beta$ without the knowledge of f based on observation $Z = h_e f(\theta) + N_e$.
- The eavesdropper observes a scaled and noise corrupted version of $f(\theta)$ (not θ) and it can only obtain prior information related to $f(\theta)$ (e.g., based on previous observations). It is assumed that the eavesdropper knows only the mean and the variance of $f(\theta)$, which are quite easy to obtain compared to the PDF of $f(\theta)$.
- Based on the previous assumption, the eavesdropper employs the linear MMSE estimator, which requires the prior knowledge of the mean and variance of $f(\theta)$ due to the independence of θ and N_e (see (24) and (25)).

According to this strategy, the MSE at the eavesdropper can be written as $E(|\hat{\beta}(Z) - \theta|^2)$, where $\hat{\beta}(Z)$ is the estimator of the eavesdropper and θ is the true value of the parameter.

For quantifying the estimation accuracy at the intended receiver, the ECRB will be used in this study, as motivated in Section I. The ECRB is defined as the expectation of the conditional CRB with respect to the unknown parameter [25], which is expressed as

$$E_\theta\big(I(\theta)^{-1}\big) = \int_\Lambda w(\theta)\frac{1}{I(\theta)}d\theta = ECRB \tag{3}$$

where $w(\theta)$ is the prior PDF of θ, $I(\theta)^{-1}$ corresponds to the conditional CRB for estimating θ,[3] and $I(\theta)$ denotes the Fisher information, i.e.,

$$I(\theta) = \int \left(\frac{\partial \log p_{Y|\theta}(y)}{\partial \theta}\right)^2 p_{Y|\theta}(y)dy \tag{4}$$

with $p_{Y|\theta}(y)$ representing the conditional PDF of Y for a given value of θ [24].

The aim is to minimize the ECRB at the intended receiver over the encoding function $f(\cdot)$. However, the estimation performance at the eavesdropper, which tries to estimate the parameter by using its observation Z, should also be considered. Therefore, the aim becomes the minimization of the ECRB for θ at the intended receiver while keeping the estimation error at the eavesdropper above a certain limit. Therefore, when deciding on the encoding scheme by using a one-to-one and continuous function in the presence of an eavesdropper, the average error at the eavesdropper should be considered, as well. Hence, the overall optimization problem is proposed as follows:

$$f_{opt} = \arg\min_f \int_\Lambda w(\theta)\frac{1}{I(\theta)}d\theta \;\; s.t. \;\; E\left(\left|\hat{\beta}(Z) - \theta\right|^2\right) \geq \alpha \tag{5}$$

where α is the MSE target at the eavesdropper and the expectation is over the joint distribution of θ and Z. In addition, the parameter space and the intrinsic constraints on the encoding function f are specified as follows:

- $\theta \in \Lambda = [a, b]$.
- $f(\theta) \in [a, b]$.
- f is a continuous and one-to-one function.

Namely, it is assumed that the parameter space is a closed set in $\mathbb{R}$ and the encoder function is an endofunction; that is, the domain and the codomain of the encoder function are the same. This is due to the practical concern that the transmitter should use the same hardware structure in the presence and absence of encoding. Furthermore, the endofunction assumption implies the peak power constraint on the encoder and it guarantees that the identity mapping $f(\theta) = \theta$ (i.e., no encoding) is a legal encoding function. It also preserves the maximum

[2]Considering a block fading scenario in which the channel coefficients are constant for a block of transmissions [6], [32]–[34], the parameter encoding function should be designed for each block.

[3]The conditional CRB presents a lower limit on the MSE of any unbiased estimator of θ based on Y for every $\theta \in \Lambda$.

range of the parameter, $b-a$. Note that it is actually possible to impose different constraints (e.g., average power constraint, boundedness) or assumptions (e.g., stochastic encoding) on the encoding function depending on the design choice and application.

The use of the ECRB as the performance metric for the design of optimal encoding functions can be justified as follows: (i) For sufficiently high SNRs, the MSE of the MAP estimator converges to the ECRB [25]. (For low SNRs, the MAP estimator depends mainly on the prior information; hence, parameter encoding becomes ineffective.) (ii) Unlike the MSE metric, the ECRB metric does not depend on a specific estimator structure. (iii) The use of the ECRB facilitates theoretical investigations for achieving intuitive understanding of the parameter encoding problem.

III. Optimal Encoding Function

In this section, the optimization problem in (5) is investigated in detail. To that aim, the MSE of the eavesdropper in the constraint of (5) is analyzed first.

$$E\left(\left|\hat{\beta}(Z)-\theta\right|^2\right) = E\left(\left|\hat{\beta}(Z)-f(\theta)+f(\theta)-\theta\right|^2\right) \tag{6}$$

$$= E\left(\left|\hat{\beta}(Z)-f(\theta)\right|^2\right) + E\left(|f(\theta)-\theta|^2\right) + 2E\left(\left(\hat{\beta}(Z)-f(\theta)\right)\left(f(\theta)-\theta\right)\right). \tag{7}$$

It is noted from (7) that the MSE of the eavesdropper is determined by both the estimation error for estimating $f(\theta)$ (that is, $\hat{\beta}(Z)-f(\theta)$) and the distortion due to the encoding function (that is, $f(\theta)-\theta$). The last term in (7) can be written as

$$E\left(\left(\hat{\beta}(Z)-f(\theta)\right)\left(f(\theta)-\theta\right)\right)$$

$$= E_\theta E_{Z|\theta}\left(\left(\hat{\beta}(Z)-f(\theta)\right)\left(f(\theta)-\theta\right)|\,\theta\right) \tag{8}$$

$$= E_\theta\left(\left(f(\theta)-\theta\right)E_{Z|\theta}\left(\hat{\beta}(Z)-f(\theta)\right)\right) \tag{9}$$

where E_θ denotes the expectation with respect to θ and $E_{Z|\theta}$ represents the conditional expectation with respect to Z given θ. As a special case, if the estimator of the eavesdropper, $\hat{\beta}(Z)$, satisfies $E_{Z|\theta}\left(\hat{\beta}(Z)-f(\theta)\right)=0$, $\forall\theta$, then the term in (9) becomes zero. This condition actually corresponds to the definition of an unbiased estimator for estimating $f(\theta)$ based on Z; i.e., $E_{Z|\theta}\left(\hat{\beta}(Z)\right)=f(\theta)$, $\forall\theta$. In other words, when the estimator of the eavesdropper is unbiased, its MSE in (6) simply becomes the sum of the MSE for estimating $f(\theta)$ (the first term in (7)) and the mean-squared distortion to θ due to the encoding function f (the second term in (7)).

The observations in the previous paragraph lead to an intuitive explanation of the proposed problem formulation. For example, suppose that the transmitter is to send parameter θ which is either 0 or 1 with equal probabilities, where $h_e=h_r=\sigma_e^2=\sigma_r^2=1$. In addition, the estimator at the eavesdropper is given by

$$\hat{\beta}(Z)=\begin{cases}1, & \text{if } Z\geq 0.5\\ 0, & \text{otherwise}\end{cases}. \tag{10}$$

If the transmitter sends the parameter without any encoding; that is, if $f(\theta)=\theta$, then the MSE of the estimator at the eavesdropper can be calculated from (7) and (10) as $Q(0.5)=0.309$ (the second and the third terms in (7) are zero), where $Q(x)=(1/\sqrt{2\pi})\int_x^\infty e^{-u^2/2}du$ represents the Q-function. On the other hand, if the transmitter employs an encoding function specified by $f(\theta)=1-\theta$, then the MSE at the eavesdropper becomes $1-Q(0.5)=0.691$ (the first term in (7) is the same as in the previous case, but the second term is 1 and the third term is $-2\,Q(0.5)$). Hence, the eavesdropper has a higher MSE as a result of secret encoding, which is not known by the eavesdropper (i.e., the eavesdropper thinks that the transmitted value is the original parameter θ). The encoding function is known by the intended receiver, which can use this information to design its estimator accordingly. However, for a generic encoding function, there can occur a penalty at the intended receiver in terms of the estimation performance. Hence, in the design of the encoding function, the trade-off between the MSE at the eavesdropper and the estimation accuracy at the intended receiver should be considered.

To specify the Fisher information in (5), the conditional PDF of Y given θ is expressed from (1) as

$$p_{Y|\theta}(y)=\frac{1}{\sqrt{2\pi\sigma_r^2}}\,e^{-\frac{(y-h_r f(\theta))^2}{2\sigma_r^2}}. \tag{11}$$

Then, the Fisher information for parameter θ can be calculated via (4) and (11) as follows:

$$I(\theta)=\frac{h_r^2\,f'(\theta)^2}{\sigma_r^2} \tag{12}$$

where $f'(\theta)$ denotes the derivative of $f(\theta)$.

Based on (7) and (12), the optimization problem in (5) can be analyzed. However, before tackling the problem in (5), the unconstrained version of it is investigated in the next section to provide initial theoretical steps towards the analysis of the generic case.

A. Optimization without Secrecy Constraints

Consider the optimization problem in (5) without the secrecy constraint; that is, by omitting the presence of the eavesdropper. Then, the optimization problem is formulated as

$$f_{opt}=\arg\min_f \int_a^b w(\theta)\frac{1}{I(\theta)}d\theta \tag{13}$$

where $\Lambda=[a,b]$ is employed as specified in Section II. Based on (12), the problem in (13) can be rewritten, by removing the constant terms, as

$$f_{opt}=\arg\min_f \int_a^b w(\theta)\frac{1}{f'(\theta)^2}d\theta\,. \tag{14}$$

The solutions of (14) are specified by the following proposition.

Proposition 1: *The optimal encoding functions in the absence of an eavesdropper are given by*

$$f(\theta)=a+\int_a^\theta g(\theta)d\theta \quad \textit{and} \quad f(\theta)=b-\int_a^\theta g(\theta)d\theta \tag{15}$$

where

$$g(\theta) \triangleq \frac{(b-a)w(\theta)^{1/3}}{\int_a^b w(\theta)^{1/3}d\theta}\,. \tag{16}$$

Proof: Since f is one-to-one and continuous, consider a monotonically increasing (decreasing) function with $f'(\theta) \geq 0$ ($f'(\theta) \leq 0$), $\forall \theta \in [a,b]$.[4] Also, due to the facts that $f(\theta)$ is monotone and $f(\theta) \in [a,b]$, the following relation can be obtained: $\int_a^b \frac{df}{d\theta}d\theta = f(b)-f(a) \leq b-a$ ($f(b)-f(a) \geq a-b$). Then, defining $g(\theta) \triangleq f'(\theta)$ ($g(\theta) \triangleq -f'(\theta)$), the problem in (14) becomes

$$\min_{g} \int_a^b w(\theta)\frac{1}{g(\theta)^2}\,d\theta \tag{17}$$

$$\text{s.t.} \int_a^b g(\theta)d\theta \leq b-a \tag{18}$$

$$g(\theta) \geq 0,\ \forall \theta \in [a,b] \tag{19}$$

Note that for all $\theta \in [a,b]$, increasing the value of $g(\theta)$ does not increase the value of the objective function; hence, the constraint in (18) is satisfied with equality. Now, in order to solve the optimization problem in (17)–(19), the calculus of variations is employed, and the problem is expressed in the form of

$$\min_{g\geq 0} \left\langle w, \frac{1}{g^2} \right\rangle \quad \text{s.t.} \quad \langle g, 1\rangle = b-a\,. \tag{20}$$

Then, the Lagrangian is obtained as

$$L(g,\epsilon,t,\lambda) = \left\langle w+\epsilon t, \frac{1}{(g+\epsilon t)^2} \right\rangle + \lambda\langle g+\epsilon t, 1\rangle \tag{21}$$

where ϵ, t, and λ represent the perturbation, the test function and the Lagrange multiplier, respectively. The optimal solution must satisfy $\frac{\partial L}{\partial \epsilon}\big|_{\epsilon=0} = 0\ \forall t$ [35], [36]. Hence, the following optimality condition is obtained:

$$\left\langle w, \frac{-2t}{(g+\epsilon t)^3} \right\rangle + \lambda\langle t, 1\rangle\Big|_{\epsilon=0} = 0 \tag{22}$$

which leads to $\langle t, \lambda + \frac{-2w}{g^3}\rangle = 0$. In order for this to hold for all t, $g = kw^{1/3}$ must be satisfied for some constant $k \geq 0$. From the equality constraint, the constant can be calculated as $k = (b-a)/\int_a^b w(\theta)^{1/3}d\theta$. Note that this $g(\theta)$ is valid, as θ takes values in $[a,b]$; hence, $w(\theta)$ is not 0 over a closed interval in $[a,b]$. Since $g(\theta) = f'(\theta)$ and $g(\theta) = -f'(\theta)$ for the monotone increasing and the monotone decreasing scenarios, respectively, the solutions can be obtained as in (15) and (16). ■

Proposition 1 states that either of the two functions given in (15) is an optimal solution for the minimization problem in (14). As a corollary to Proposition 1, if the prior distribution of the parameter is uniform over $[a,b]$, the optimal encoding functions can be found via (15) and (16) as $f(\theta) = \theta$ and $f(\theta) = a+b-\theta$. In other words, for the uniform prior, parameter encoding is not needed for reducing the ECRB at the intended receiver.

[4]Note that $f'(\theta)$ can be zero at certain points; however, it is not 0 for a closed interval in $[a,b]$ due to the one-to-one property.

B. Optimization with Secrecy Constraints

In this part, the optimization problem in (5) is considered without omitting the secrecy constraint, where the parameter space is specified by $\Lambda = [a,b]$ as before. Although the linear MMSE estimator is assumed to be employed at the eavesdropper in this study (see Section II), a corollary to Proposition 1 is presented first for the case in which the eavesdropper employs the MMSE estimator, defined as $\hat{\beta}(z) = E(\beta|Z=z)$ with $\beta = f(\theta)$.

Corollary 1: *Suppose that the eavesdropper employs the MMSE estimator for a given encoding function $f(\theta)$. Denote the corresponding MSE at the eavesdropper as $R(f^+)$ when the encoding function is $f(\theta) = a + \int_a^\theta g(\theta)d\theta \triangleq f^+$, and as $R(f^-)$ when the encoding function is $f(\theta) = b - \int_a^\theta g(\theta)d\theta \triangleq f^-$, where $g(\theta)$ is as defined in Proposition 1. Then, the following statements hold:*

a) *If the target MSE of the eavesdropper, α in* (5)*, satisfies $\alpha \leq \min\{R(f^+), R(f^-)\}$, then both f^+ and f^- are optimal encoding functions.*

b) *If $\min\{R(f^+), R(f^-)\} \leq \alpha \leq \max\{R(f^+), R(f^-)\}$, then the optimal encoding function is f^+ if $R(f^+) > R(f^-)$ and it is f^- otherwise.*

Proof: Proposition 1 implies that if f^+ or f^- is admissible by the constraint, it becomes the minimizer of the objective function. When the eavesdropper employs the MMSE estimator, $\hat{\beta}(z) = E(\beta|Z=z)$, the MSE at the eavesdropper can be calculated from (7) for a given encoding function. For the special cases of encoding functions f^+ and f^-, the corresponding MSE values are denoted by $R(f^+)$ and $R(f^-)$, respectively. If α is less than both of $R(f^+)$ and $R(f^-)$, then f^+ and f^- do not violate the constraints and solve (5). If α is less than only one of $R(f^+)$ or $R(f^-)$, then still one of f^+ and f^- is admissible; hence, the optimal encoding function. ■

It is noted that when $\alpha \geq \max\{R(f^+), R(f^-)\}$, the shortcut provided in Corollary 1 cannot be used, and it is required to design another encoding function to satisfy the secrecy constraint.

Remark 1: The statement in Corollary 1 in fact holds for any estimator at the eavesdropper since the proof is not specific to the MMSE estimator. In other words, as long as any of the encoding functions in Proposition 1 results in an MSE at the eavesdropper that is higher than the target MSE α, that encoding function is also optimal for the problem in (5). Since the MMSE estimator achieves the minimum MSE among all estimators, it is concluded that if one of the encoding functions in Proposition 1 is optimal when the eavesdropper employs the MMSE estimator, then that encoding function is in fact optimal for any other estimator at the eavesdropper.

Even though the MMSE estimator is the optimal estimator according to the MSE metric, for implementing the MMSE estimator, the eavesdropper must know the prior PDF of $f(\theta)$, which can be difficult to obtain (learn). In this study, it is assumed that the eavesdropper has the knowledge of the mean and variance of $f(\theta)$. Therefore, the eavesdropper is assumed to employ the linear MMSE estimator to estimate $\beta = f(\theta)$ based on Z, as noted in Section II. It is known that the linear

MMSE estimator is the optimal linear estimator according to the MSE metric [37]. Furthermore, it would actually be the optimal MMSE estimator to estimate β based on Z, $E(\beta|Z = z)$, if β and Z were jointly Gaussian random variables [24]. For the system model in this study, the MMSE estimator and the linear MMSE estimator will have similar performance at low SNRs if the prior is uniformly distributed.

When the linear MMSE estimator is employed at the eavesdropper, $\hat{\beta}(z)$ can be expressed as

$$\hat{\beta}(z) = k_0 + k_1 z \tag{23}$$

where k_0 and k_1 are chosen to minimize $E\left(\left|\hat{\beta}(Z) - \beta\right|^2\right) = E\left(\left|k_0 + k_1 Z - \beta\right|^2\right)$ as the eavesdropper does not know the encoding. The resulting coefficients for the eavesdropper's estimator are given as (see Appendix A for the derivation)

$$k_1 = \frac{h_e Var(\beta)}{h_e^2 Var(\beta) + \sigma_e^2} \tag{24}$$

$$k_0 = (1 - k_1 h_e) E(\beta). \tag{25}$$

Then, the resulting MSE between the estimate of the eavesdropper and the true value of parameter θ can be derived from (23)–(25) and (7) as (see Appendix B for the derivation)

$$E\left(\left|\hat{\beta}(Z) - \theta\right|^2\right) = \frac{h^2 V(V - 2C)}{h^2 V + 1} + (E(\beta) - E(\theta))^2 + Var(\theta) \tag{26}$$

where $\beta = f(\theta)$, $V = Var(\beta)$, $C = Cov(\beta, \theta)$, and $h = h_e/\sigma_e$.

It is observed that the MSE value at the eavesdropper corresponding to the linear MMSE estimator depends on both the encoding function and the channel quality h at the eavesdropper. It is noted that for a given encoding function with $V - 2C > 0$, the first term in (26) is positive, and the MSE at the eavesdropper becomes an increasing function of h^2. This means that as the channel quality for the eavesdropper improves, the resulting MSE at the eavesdropper increases in that scenario. This seemingly counterintuitive result is simply due to the fact that the estimator of the eavesdropper is based on the noisy observation of the distorted version of the original parameter. Hence, one can transmit the inflicted distortion more efficiently to the eavesdropper under good channel conditions leading to a higher MSE. If the eavesdropper knew the prior distribution of the original parameter and realized that the transmitter sends the encoded version, it would simply stop using the observation and set $\hat{\beta}(Z) = E(\theta)$, resulting in an MSE of $Var(\theta)$, which is lower than the value in (26) for the case of $V - 2C > 0$. However, the eavesdropper does not have that knowledge and the channel observation is the only information it can use to estimate the parameter, which is utilized by the transmitter.

Remark 2: In the considered setting, the eavesdropper employs the linear MMSE estimator and the transmitter is aware of this situation. Then, to obtain the optimal encoding function based on (5), (12), and (26), the transmitter should have the knowledge of the prior PDF of the parameter and the channel quality parameter h_e^2/σ_e^2 for the eavesdropper. In practice, it can be challenging for the transmitter to have an accurate knowledge of the channel quality for the eavesdropper. In such cases, a conservative approach can be taken by either increasing the MSE target α in (5) or considering the worst-case (minimum) value of the MSE at the eavesdropper according to the uncertainty in the channel quality parameter.

The following proposition presents a shift invariance property for the considered problem.

Proposition 2: *Suppose that the unknown parameter θ resides in $[a, b]$ with a prior distribution specified by $w(\theta)$, and the encoding function $f(\theta) : [a, b] \to [a, b]$ results in a certain ECRB at the intended receiver and a certain MSE at the eavesdropper, which employs the linear MMSE estimator. If the parameter θ were defined in $[0, b-a]$ with the prior distribution $\hat{w}(\theta) = w(\theta + a)$, then the use of the encoding function $\hat{f}(\theta) : [0, b-a] \to [0, b-a]$ such that $\hat{f}(\theta) = f(\theta+a)-a$ would result in the same MSE at the eavesdropper and the same ECRB at the intended receiver as in the original scenario.*

Proof: The ECRB in the original scenario can be expressed from (12) and (13) as

$$\frac{\sigma_r^2}{h_r^2} \int_a^b w(\theta) \frac{1}{f'(\theta)^2} d\theta \tag{27}$$

which is equivalent to

$$\frac{\sigma_r^2}{h_r^2} \int_0^{b-a} w(\theta + a) \frac{1}{((f(\theta + a) - a)')^2} d\theta \tag{28}$$

since $(f(\theta + a) - a)' = f'(\theta + a)$. As the expression in (28) corresponds to the ECRB in the second scenario, the equivalence of the ECRBs is established. To prove that the MSE at the eavesdropper does not change, it is noted that the parameter defined in $[0, b - a]$ with the prior distribution $\hat{w}(\theta) = w(\theta+a)$ corresponds to shifting the original parameter as $\theta - a$. Also, let $\bar{\beta}$ and β denote the random variables for the encoded versions of the shifted and original parameters via encoding functions $\bar{f}(\theta)$ and $f(\theta)$, respectively. Then, $\bar{\beta} = \beta - a$ holds. Furthermore, it is noted that shifting the specified random variables (θ and $\beta = f(\theta)$) just changes their means by the amount of the shift without modifying the second order statistics V and C. Hence, (26) reveals that the MSE at the eavesdropper stays the same as in the original scenario after the shift operations. ■

Based on Proposition 2, the estimation of a parameter in $\theta \in [0, b - a]$ can be considered without loss of generality for the case of the linear MMSE estimator at the eavesdropper (see Proposition 4).

The next proposition states that when the prior PDF of $\theta \in [a, b]$ is symmetric around $(a + b)/2$, parameter encoding via a strictly decreasing function is more desirable than that via a strictly increasing one.

Proposition 3: *Suppose that the eavesdropper employs the linear MMSE estimator and $w(\theta)$ is symmetric around $(a+b)/2$. Then, for any given continuous and strictly increasing encoding function, there exists a corresponding continuous and strictly decreasing encoding function that yields the same ECRB at the intended receiver with a higher MSE at the eavesdropper.*

Proof: Consider two encoding functions $f(\theta)$ and $s(\theta) = f(a+b-\theta)$, where $\theta \in [a,b]$ and $f(\theta)$ is a continuous and monotonically increasing function. Since $w(\theta) = w(a+b-\theta)$ due to the symmetry assumption and $s'(\theta) = -f'(a+b-\theta)$ by definition, both encoding functions result in the same ECRB, which can be proved via (14) as follows:

$$\int_a^b w(\theta)\frac{1}{s'(\theta)^2}d\theta = \int_a^b w(a+b-\theta)\frac{1}{f'(a+b-\theta)^2}d\theta = \int_a^b w(\theta)\frac{1}{f'(\theta)^2}d\theta \quad (29)$$

where the final expression is obtained via a change of variables. To compare the MSEs corresponding to the two encoding functions, define $\beta_f \triangleq f(\theta)$ and $\beta_s \triangleq s(\theta)$, and let $p_{\beta_f}(x)$ and $p_{\beta_s}(x)$ represent the PDFs of β_f and β_s, respectively. Then, it is noted that $p_{\beta_f}(x) = p_{\beta_s}(x)$ for $x \in [a,b]$ since $w(\theta) = w(a+b-\theta)$ due to symmetry. Hence, both β_f and β_s have the same expectation and the variance. For the covariance, $Cov(\beta,\theta) = E(\beta\theta) - E(\beta)E(\theta)$, the following expression can be obtained:

$$E(\beta_f\theta) - E(\beta_s\theta) = \int_a^b w(\theta)f(\theta)\theta d\theta - \int_a^b w(\theta)f(a+b-\theta)\theta d\theta \quad (30)$$

$$= \int_a^b w(\theta)f(\theta)(2\theta - a - b)d\theta. \quad (31)$$

where (30) follows from the definitions of β_f and β_s, and (31) is obtained from the symmetry of $w(\theta)$. Since $f(\frac{a+b}{2} - x) < f(\frac{a+b}{2} + x)$ for $x \in (0, \frac{a+b}{2}]$, $E(\beta_f\theta) - E(\beta_s\theta) > 0$. Then, $Cov(\beta_f,\theta) > Cov(\beta_s,\theta)$ and $E(|\hat{\beta}_f - \theta|^2) < E(|\hat{\beta}_s - \theta|^2)$ according to (26). Therefore, it is always possible to achieve a higher MSE by employing $s(\theta)$ instead of $f(\theta)$ while keeping the ECRB the same. ■

Proposition 3 implies that it is sufficient to search for the optimal encoding function among strictly decreasing functions if the prior distribution of the parameter satisfies the symmetry condition (e.g., the uniform distribution). This is based on the idea that for any given increasing encoding function that solves (5), there exists a legitimate decreasing function obtained by a simple transformation, which yields the same optimal ECRB value with an increased MSE at the eavesdropper. Hence, from a practical point of view, the search space for the optimal encoding function can be confined to strictly decreasing functions under the conditions in the proposition.

1) Special Case: Uniform Prior Distribution: For the special case of a uniform prior distribution, the following result characterizes the optimal encoding function when the eavesdropper employs the linear MMSE estimator.

Corollary 2: *Suppose that the parameter has uniform prior distribution over $[a,b]$ and the eavesdropper employs the linear MMSE estimator. Then, if the target MSE α satisfies $\alpha \leq \frac{V_u}{h^2V_u+1}$, then $f(\theta) = \theta$ is an optimal encoding function, where $V_u \triangleq (b-a)^2/12$. On the other hand, if $\alpha \leq \frac{4h^2V_u^2+V_u}{h^2V_u+1} + (a+b-2E(\theta))^2$, then $f(\theta) = a+b-\theta$ is an optimal encoding function.*

Proof: From the expressions in Proposition 1, it can be shown, for the uniform prior distribution, that either of $f(\theta) = \theta$ or $f(\theta) = a+b-\theta$ is an optimal encoding function in the absence of the constraint (i.e., in the absence of the eavesdropper). When the eavesdropper employs the linear MMSE estimator, the use of $f(\theta) = \theta$ leads to an MSE of $\frac{V_u}{h^2V_u+1}$ and the use of $f(\theta) = a+b-\theta$ results in an MSE of $\frac{4h^2V_u^2+V_u}{h^2V_u+1} + (a+b-2E(\theta))^2$, which can be derived based on (26). Then, based on similar arguments to those in Corollary 1, it is deduced that if the MSE corresponding to $f(\theta) = \theta$ is larger than or equal to α, $f(\theta) = \theta$ is an optimal encoding function. Similarly, if the MSE for $f(\theta) = a+b-\theta$ is larger than or equal to α, $f(\theta) = a+b-\theta$ is an optimal encoding function. ■

The following proposition provides an upper bound on the MSE at the eavesdropper, which employs the linear MMSE estimator, when the parameter has uniform prior distribution.

Proposition 4: *If the eavesdropper employs the linear MMSE estimator and θ has uniform distribution over $[0,\gamma]$, then*

$$\sup_f E\left(\left|\hat{\beta}(Z) - \theta\right|^2\right) = \begin{cases} \frac{\gamma^2}{3}, & \gamma \leq \frac{2}{h} \\ \frac{h^2\gamma^4}{2h^2\gamma^2+8} + \frac{\gamma^2}{12}, & \gamma > \frac{2}{h} \end{cases} \quad (32)$$

where $f(\theta) : [0,\gamma] \to [0,\gamma]$ is a continuous and one-to-one function.

Proof: For an encoding function $f(\theta) = \beta$, let $V = Var(\beta)$, $C = Cov(\beta,\theta)$, and $\mu = E(\beta)$. It can be shown that for a random variable defined on the bounded interval of $[0,\gamma]$, the following relations hold for $\mu \in [0,\gamma]$:

$$0 \leq V \leq \mu(\gamma - \mu) \leq \frac{\gamma^2}{4}\cdot \quad (33)$$

In addition, C can be expressed as

$$C = \frac{1}{\gamma}\int_0^\gamma f(\theta)\left(\theta - \frac{\gamma}{2}\right)d\theta.$$

For a given continuous endofunction on $[0,\gamma]$, it can be shown that C is in $(-\gamma^2/8, \gamma^2/8)$. Also, from (26), the MSE at the eavesdropper can be stated as

$$E\left(\left|\hat{\beta}(Z) - \theta\right|^2\right) = \frac{h^2V(V-2C)}{h^2V+1} + \left(\mu - \frac{\gamma}{2}\right)^2 + \frac{\gamma^2}{12} \leq \frac{h^2V(V-2C)}{h^2V+1} - V + \frac{\gamma^2}{3} = \frac{-V(1+2h^2C)}{h^2V+1} + \frac{\gamma^2}{3} \quad (34)$$

where the inequality holds for any continuous encoding function defined on $[0,\gamma]$. Therefore, the upper bound on $E(|\hat{\beta}(Z) - \theta|^2)$, specified in (34), holds for all possible encoding functions. Next, the maximum of this generic upper bound is to be found over the PDF of β, denoted by p_β, where $\beta = f(\theta)$. It is observed that if $(1+2h^2C) > 0$ for any given p_β, the first term in (34) is nonpositive as $V \geq 0$ and $h > 0$; hence, the maximum of the upper bound is $\gamma^2/3$. When $(1+2h^2C) \leq 0$, then the first term in (34) is maximized by increasing V and decreasing C at the same time. Therefore, if $V = \gamma^2/4$ and $C = -\gamma^2/8$, the maximum of the upper bound

is achieved. Thus, for $\gamma \leq 2/h$, $E(|\hat{\beta}(Z)-\theta|^2) \leq \gamma^2/3$ holds, and for $\gamma > 2/h$,

$$E(|\hat{\beta}(Z)-\theta|^2) \leq \frac{h^2\gamma^4}{2h^2\gamma^2+8} + \frac{\gamma^2}{12} \tag{35}$$

is obtained. Furthermore, in the first case (i.e., $\gamma \leq 2/h$), $\tilde{f}(\theta) = 0$ (or, $\tilde{f}(\theta) = \gamma$) for $\theta \in [0, \gamma]$ attains the upper bound on the MSE, that is, $\gamma^2/3$. In the second case (i.e., $\gamma > 2/h$), it is possible to achieve the upper bound in (35) by using $\tilde{f}(\theta)$ defined as

$$\tilde{f}(\theta) = \begin{cases} \gamma, & 0 \leq \theta \leq \gamma/2 \\ 0, & \gamma/2 < \theta \leq \gamma \end{cases}. \tag{36}$$

Even though the maximum values for the upper bounds are obtained and it is argued that they are exactly attained by using $\tilde{f}(\theta)$, it should be noted that $\tilde{f}(\theta)$'s are not in the feasible function set as they do not satisfy the one-to-one and continuity properties. However, it is possible to approach arbitrarily close to $\tilde{f}(\theta)$ while staying in the feasible function set (e.g., take $\delta > 0$, set $f(\theta) = \gamma - \delta\theta$, and let $\delta \to 0$ for the first case). Furthermore, the objective is continuous functional acting on the encoding function. Hence, the upper bound values for the MSE cannot exactly be achieved; however, one can get arbitrarily close to them by employing one-to-one and continuous functions, which yield them as the supremum values for the MSE at the eavesdropper, resulting in the expression in (32). ■

In addition to providing a closed form upper bound on the distortion at the eavesdropper, Proposition 4 plays another important role by helping us gain practical intuition about the behavior of the optimal encoding function (*variance minimizing mode* or *variance maximizing mode*). As argued in Proposition 3, if there are two alternative encoding functions with the same ECRB, then it is better to choose the one which yields the higher MSE. Note that given an encoding function $f(\theta)$, one can shuffle the increments of the given function and end up with an alternative encoding function with the same ECRB. The alternative encoding function will possibly have different (μ, V, C). Therefore, it is important to understand how the MSE behaves as μ, V, and C change. Note that in (32), the supremum changes depending on the value of the channel quality of the eavesdropper h for a given γ ($h \leq 2/\gamma$ or $h \geq 2/\gamma$). Let us investigate those two cases:

- If the channel quality h is small enough, one can let $\beta \to 0$ (or γ) to maximize its MSE. This strategy is equivalent to minimizing the variance and letting $E(\beta) \to 0$ (or γ).
- If the channel quality h is large enough, then one can increase the variance V and decrease C at the same time to maximize its MSE (effectively maximize $E(|\beta-\theta|^2)$).

This discussion becomes clearer at the extreme cases of the value of h^2. For example, suppose that h^2 is very small. Then, (26) reveals that $E(|\hat{\beta}(Z)-\theta|^2) \approx (E(\beta)-E(\theta))^2 + Var(\theta)$; hence, it is possible to generate a larger MSE by making $E(\beta)$ as close to the boundaries 0 and γ as possible. This behavior can be regarded as the *variance minimizing mode*. If h^2 is very large, then $E(|\hat{\beta}(Z) - \theta|^2) \approx (V - 2C) + (E(\beta) - E(\theta))^2 + Var(\theta) = E((\beta-\theta)^2) + E(\theta)^2$; hence, it is possible to generate a higher MSE by maximizing $E((\beta-\theta)^2)$. This behavior can be regarded as the *variance maximizing mode*. Of course, if the resulting MSE is higher than the target MSE α for a given h, then one can use linear encoding $f(\theta) = \gamma - \theta$ for minimizing the ECRB.

It is important to note that Proposition 4 does not have any constraints on the ECRB. The original problem tries to minimize the ECRB for a target MSE. Among two candidates with the same ECRB, the one yielding the larger MSE at the eavesdropper is preferred in the search of the optimal encoder. The feasible set for (μ, V, C) is specific to that ECRB value. For example, one might not be able to let $\mu \to 0$ anymore. However, one can generate a larger MSE by making μ very close to its limit in the feasible set for sufficiently small h values (e.g., by using a decreasing concave function) or by making $E(|\beta-\theta|^2)$ as large as possible for sufficiently high h values (e.g., by using a decreasing concave function for $\beta < \gamma/2$ and a decreasing convex function for $\beta > \gamma/2$). Hence, the optimal encoding function will be in either of the modes (variance minimizing or maximizing) described above.

Remark 3: The optimal value of the optimization problem in (5) can be named as $G(\alpha)$ since the optimal ECRB value depends on the target MSE α. That is,

$$G(\alpha) \triangleq \frac{\sigma_r^2}{h_r^2} \int_a^b w(\theta) \frac{1}{f'_{opt}(\theta)^2} d\theta \tag{37}$$

where $f_{opt}(\theta)$ is a solution to (5). Note that the optimal value of the ECRB in the case of optimization without secrecy constraints can be denoted by $G(0)$. Then, $G(\alpha)$ has the following properties:

- $G(\alpha)$ is constant between $0 \leq \alpha \leq \alpha_{th}$ with $\alpha_{th} = \max\{R(f^+), R(f^-)\}$, where $R(f^+)$ and $R(f^-)$ can similarly be defined as in Corollary 1 except that the linear MMSE is employed at the eavesdropper.
- $G(\alpha)$ is a non-decreasing function between $\alpha_{th} \leq \alpha < \alpha_{max}$, where $\alpha_{max} = \sup_f E\left(\left|\hat{\beta}(Z) - \theta\right|^2\right)$.
- $G(\alpha) \to \infty$ as $\alpha \to \alpha_{max}$.

The second property follows from the following argument: Let S_{α_1} and S_{α_2} be the feasible sets for α_1 and α_2, respectively. If $\alpha_1 \geq \alpha_2$, then $S_{\alpha_1} \subseteq S_{\alpha_2}$; hence, $G(\alpha_1) \leq G(\alpha_2)$. Note that a closed form expression for α_{max} is provided in Proposition 4 for the special case of uniform prior distribution.

IV. Solution Approaches

In general, the optimal parameter encoding problem formulated by (5), (12), and (26) is a difficult optimization problem as it requires a search over functions. Although the theoretical results in the previous section can lead to closed-form solutions or reductions in the search space in certain scenarios, it may still be necessary to solve the problem directly in some cases. Therefore, various solution approaches are developed in this section for obtaining suboptimal solutions of (5). In the proposed approaches, it is assumed that the encoding function f is picked among a family of functions characterized by a certain number of parameters. Then, the optimization problem becomes easier to solve as it involves minimization over a few variables (instead of functions), which

also leads to some analytical solutions, as discussed below. However, the obtained encoding function will be suboptimal in general since the actual solution of (5) may not be a function from the assumed family of functions.

A. Linear Encoding Functions

One suboptimal encoding scheme is to employ a linear encoding function to minimize the ECRB at the intended receiver while satisfying the MSE constraint at the eavesdropper. To obtain analytical results for generic prior PDFs, the eavesdropper is modeled to employ the linear MMSE estimator as before, and the encoding function is assumed to be a decreasing linear function. However, the analysis can also be performed easily for increasing linear functions in a similar fashion, which yields similar analytical results to those in Proposition 5 and afterwards. (In practice, it is advised to solve the encoding problem restricted to decreasing linear functions and to increasing linear functions separately, and select the one with the lower objective value. However, when the prior PDF of the parameter, $w(\theta)$, is symmetric around $(a+b)/2$, where $\theta \in [a,b]$, it is sufficient to consider decreasing functions only, as shown in Proposition 3.)

For the considered model, the linear encoding function can be expressed as

$$f(\theta) = c_0 + m(b-\theta) \tag{38}$$

where $m \in (0,1]$, $c_0 \geq a$, and $c_0 + m(b-a) \leq b$. In other words, for a fixed m, c_0 can be any real number in $[a, b-m(b-a)]$. In addition, the random variable $\beta = f(\theta)$ has the following PDF: $p_\beta(x) = \frac{1}{m} w\big(\frac{c_0+mb-x}{m}\big)$ for $x \in [c_0, c_0+m(b-a)]$. For example, if $w(\theta)$ is the uniform PDF over $[a,b]$, then β will have uniform distribution over $[c_0, c_0+m(b-a)]$; hence, its amplitude is $\frac{1}{m(b-a)}$ inside that interval and 0 elsewhere. Also, the value of c_0 does not change this amplitude but only causes a shift in the domain of β. First, the following proposition is presented about c_0 for any given input distribution $w(\theta)$.

Proposition 5: *When the eavesdropper employs the linear MMSE estimator, the MSE at the eavesdropper for the linear encoding function $f(\theta) = c_0 + m(b-\theta)$ is a convex function of c_0 for a fixed $m > 0$. Hence, the MSE is maximized either at $c_0 = a$ (if $E(\theta) > (a+b)/2$) or at $c_0 = b - m(b-a)$ (if $E(\theta) < (a+b)/2$).*

Proof: The variance and the mean of $\beta = f(\theta)$ can be calculated as $Var(\beta) = m^2 Var(\theta)$ and $E(\beta) = c_0 + mb - mE(\theta)$. Also, the covariance of β and θ can be obtained as $Cov(\beta,\theta) = -mVar(\theta)$. In (26), only the second term depends on c_0. In addition, $(E(\beta)-E(\theta))^2 = (c_0-(E(\theta)(1+m)-mb))^2$ is a convex function of c_0 for a fixed m, and it is equal to $(a-E(\theta)-m(E(\theta)-b))^2$ at $c_0 = a$ and $(b-E(\theta)-m(E(\theta)-a))^2$ at $c_0 = b-m(b-a)$. Hence, for a given $m \in (0,1)$, the MSE is maximized either at $c_0 = a$ if $E(\theta) > (a+b)/2$ or at $c_0 = b-m(b-a)$ if $E(\theta) < (a+b)/2$. (If $m = 1$ or $E(\theta) = (a+b)/2$, it has the same value at both of the boundaries, hence, there exist two maximizers in that case.) ■

Proposition 5 leads to the closed-form solution for the optimal linear encoding function as follows: Since the ECRB expression depends only on the derivative of the encoding function (see (5) and (12)), it is proportional to $1/m^2$ for the linear encoding function in (38); hence, it does not depend on c_0. Therefore, c_0 can be chosen to maximize the MSE at the eavesdropper based on Proposition 5, which implies that c_0 is equal to either a or $b-m(b-a)$ (which corresponds to either $f(b) = a$ or $f(a) = b$). Based on these observations, it is sufficient to perform a search only over parameter m in order to determine the optimal linear encoding function. Suppose that $E(\theta) > (a+b)/2$ and model the linear encoding function as $f(\theta) = a + m(b-\theta)$ (see Proposition 5). (The case of $E(\theta) < (a+b)/2$ and $f(\theta) = b + m(a-\theta)$ can be treated similarly.) Then, the optimization problem specified by (5) and (12) can be rewritten to find the optimal m as follows:

$$m_{opt} = \arg\min_m \frac{1}{m^2} \text{ s.t. } E\Big(\big|\hat{\beta}(Z)-\theta\big|^2\Big) \geq \alpha,\ 0 < m \leq 1 \tag{39}$$

where $E(|\hat{\beta}(Z)-\theta|^2) = \frac{h^2m^2V(m^2V+2mV)}{h^2m^2V+1} + V + (a-E(\theta)-m(E(\theta)-b))^2$ with $V = Var(\theta)$ due to (26). Obviously, the optimal m is the largest m that satisfies the constraints. After some algebra, the first constraint can be expressed as

$$\begin{aligned} &\big(tV^2+\kappa_1^2 tV\big)m^4 + \big(2tV^2+2tV\kappa_1\kappa_2\big)m^3 \\ &+ \big(tV^2+(\kappa_2^2-\alpha)tV+\kappa_1^2\big)m^2 + (2\kappa_1\kappa_2)m \\ &+ \big(\kappa_2^2+V-\alpha\big) \geq 0 \end{aligned} \tag{40}$$

where $t \triangleq h^2$, $\kappa_1 \triangleq b-E(\theta)$, and $\kappa_2 \triangleq a-E(\theta)$. Hence, the optimal m is the largest m in $(0,1]$ satisfying (40). This optimal value can be obtained algebraically by finding the roots of the fourth degree polynomial in (40). For example, when $h = 1$, $a = 0$, $b = 1$, $w(\theta)$ is uniform, and $\alpha = 0.15$, (40) becomes $m^4 - m^3 + 9.55m^2 - 18m + 6.6 \geq 0$. This polynomial has roots at 1.3001, 0.4915, and $-0.3958 \pm 3.1895i$, implying that the constraint holds when $m \geq 1.3001$ or $m \leq 0.4915$; thus, the optimal m is given by $m = 0.4915$. Overall, it is concluded that considering an encoding function among the family of linear functions, the optimal solution can be obtained by finding the roots of a polynomial equation without performing any functional optimization.

Remark 4: One alternative approach could be to consider an encoding function in the form of $f(\theta) = a + p\big(\frac{b-\theta}{b-a}\big)^q$, where the function is parameterized by p and q. Hence, instead of trying to optimize over functions, one can try to use this family of *power functions*, and perform optimization over $p \in (0, b-a]$ and $q \in (0, 3/2)$. Even though this will lead to a suboptimal encoding function, it is still easier to perform optimization via a 2-dimensional search than optimizing over functions as in (5). On the other hand, this approach will have higher computational complexity than the one that employs (38).

B. Polynomial Approximation

The second approach for obtaining a suboptimal solution of (5) is to use a polynomial approximation method. Approximating a function via polynomials is a well-known numerical analysis method [39]–[41]. To apply this method to the parameter encoding problem, it is assumed that the

encoding function is in the form of a polynomial. In fact, any continuous real-valued function defined on $[a,b]$ can be uniformly approximated by polynomials in that interval [38]. That is, for a given continuous and bounded function $f(x)$ and $\epsilon > 0$, there exists a polynomial $P(x)$ on $[a,b]$ such that $\sup_x |f(x) - P(x)| < \epsilon$. Motivated by this fact, the encoding function is expressed by Kth degree polynomials, i.e., $P(x) = \sum_{n=0}^{K} c_n x^n$, and the aim becomes the calculation of the optimal coefficients c_n for $n = 0, 1, \ldots, K$. Hence, by using $f(\theta) = \sum_{n=0}^{K} c_n \theta^n$, the optimization problem specified by (5) and (12) can be rewritten to find the optimal coefficients as follows:

$$\begin{aligned} \mathbf{c}^{\text{opt}} = \underset{c_0,c_1,\ldots,c_K}{\arg\min} & \int_{\Lambda} w(\theta) \left(\sum_{n=0}^{K} n c_n \theta^{n-1} \right)^{-2} d\theta \\ \text{s.t.} \quad & E\left(\left| \hat{\beta}(Z) - \theta \right|^2 \right) \geq \alpha \end{aligned} \tag{41}$$

After finding the optimal coefficients, the encoding function can be written as $f_{\text{opt}}(\theta) = \sum_{n=0}^{K} c_n^{\text{opt}} \theta^n$, where c_n^{opt} represents the nth element of $\mathbf{c}^{\text{opt}}$. Note that the resulting encoding function should also satisfy the implicit conditions, that is, $f(\theta) \in [a,b]$ and the monotonicity.

C. Piecewise Linear Approximation:

Finally, a third approach is proposed, which is based on the idea that any continuous bounded function can be uniformly approximated by piecewise linear functions. Therefore, the parameter space $[a,b]$ is partitioned into M intervals and the optimal increment (or, decrement) is found in each interval, which results in an approximation of the encoding function f via a piecewise linear function. In particular, the increments/decrements are defined as $\Delta x_k = f(a + k\Delta\theta) - f(a + (k-1)\Delta\theta)$, and the optimization is performed over M variables, $\Delta x_1, \Delta x_2, \ldots, \Delta x_M$. As M increases, more accurate approximation is achieved; however, the computational complexity of solving the optimization problem increases, as well. Note that, for $M = 1$, this approach reduces to the linear encoding function case in Section IV-A. The optimization problem specified by (5) and (12) can be stated to find the optimal increments as follows:

$$\begin{aligned} \mathbf{\Delta x}_{\text{opt}} = \underset{\Delta x_1, \Delta x_2, \ldots, \Delta x_M}{\arg\min} & \sum_{k=1}^{M} \frac{1}{\Delta x_k^2} \int_{a+(k-1)\Delta\theta}^{a+k\Delta\theta} w(\theta) d\theta \\ \text{s.t.} \quad & E\left(\left| \hat{\beta}(Z) - \theta \right|^2 \right) \geq \alpha \end{aligned} \tag{42}$$

Similar to the previous case, the resulting encoding function should also satisfy the implicit conditions, that is, $f(\theta) \in [a,b]$ and the monotonicity. For example, if a decreasing encoding function is used, then all the elements in $\mathbf{\Delta x}_{\text{opt}}$ should be negative. In order to solve the problems given in (41) and (42), we have used the Global Optimization Toolbox of MATLAB. As the initial point, the linear solution, which is calculated analytically, can be used. It is noted that the objective function given in (14) is a convex operation on f; however, the feasible set does not need to be convex. This discussion holds for both of the problems in (41) and (42).

Remark 5: Most of the theoretical results in this paper can be extended, under certain conditions, to scenarios in which the eavesdropper employs an arbitrary affine estimator, $\hat{\beta}(z) = R_0 + R_1 z$, instead of the linear MMSE estimator. In this case, after some manipulation, the MSE of the eavesdropper can be obtained for given R_1 and R_0 as

$$\begin{aligned} E\left(\left| \hat{\beta}(Z) - \theta \right|^2 \right) = & R_1^2 (h_e^2 V + \sigma_e^2) - 2R_1 h_e C + Var(\theta) \\ & + (R_1 h_e E(\beta) - E(\theta) + R_0)^2 \end{aligned} \tag{43}$$

where $V = Var(\beta)$ and $C = Cov(\beta, \theta)$. Then, the results can be extended as follows:

- Proposition 2 does not hold for general R_1 and R_0. However, for the special case of $R_1 = 1/h_e$, it holds for any R_0, and for $R_0 = E(\beta)(1 - R_1 h_e)$, it holds for any R_1. It is noted that the second case implies that $E(\hat{\beta}(z)) = E(\beta)$.
- Proposition 3 holds if $R_1 h_e > 0$. If $R_1 h_e < 0$, then the reverse of the argument holds; that is, for a given strictly decreasing function, one can find a simple transformation such that the resulting encoding function has a lower MSE. Corollary 2 can also be generalized in a similar fashion.
- Proposition 4 is particular to the assumption of the linear MMSE estimator; hence, it cannot be generalized directly for arbitrary R_1 and R_0. However, an upper limit can be found as follows by considering R_1 and R_0 as given constants:

$$\begin{aligned} \sup_f E\left(\left| \hat{\beta}(Z) - \theta \right|^2 \right) = \sup_f \Big(& E\left(\left| R_1 h_e \beta - \theta \right|^2 \right) \\ & + 2R_1 R_0 h_e E(\beta) + g(R_0, R_1) \Big) \end{aligned}$$

  where $g(R_0, R_1) = R_1^2 - 2R_0 E(\theta)$. Next, let $R_1 h_e = k$ and $k > 0$. Then, for a fixed $E(\beta) = \alpha$ with $\alpha \in [0, \gamma]$, $E\left(\left| k\beta - \theta \right|^2 \right)$ is maximized if $\beta = \gamma$ for $\theta < \alpha$ and 0 otherwise. Then, the analysis can be completed by finding the optimal α.
- Finally, if $R_1 h_e > 0$, Proposition 5 can also be generalized. Namely, the MSE is a convex function of c_0 for a fixed $m > 0$ and is maximized either at $c_0 = a$ or $c_0 = b - m(b-a)$.

V. Numerical Results

In this section, numerical examples are provided to investigate the theoretical results in Section III and to compare the proposed approaches in Section IV. Throughout the simulations, h_r and σ_r^2 are set as $h_r = \sigma_r^2 = 1$.

First, we consider a scenario in which the channel parameters for the receiver and the eavesdropper are fixed, and investigate the relation between the ECRB and the secrecy limit α by using different encoding strategies. It is assumed that the parameter θ has uniform distribution over $[0,1]$ and $h = h_e/\sigma_e = 1$. Also, the eavesdropper employs the linear MMSE estimator for the encoded parameter $\beta = f(\theta)$. The theoretical results derived in Section III can be applied for this example. In particular, based on Proposition 1, it is known that if there is no secrecy constraint, either $f(\theta) = \theta$

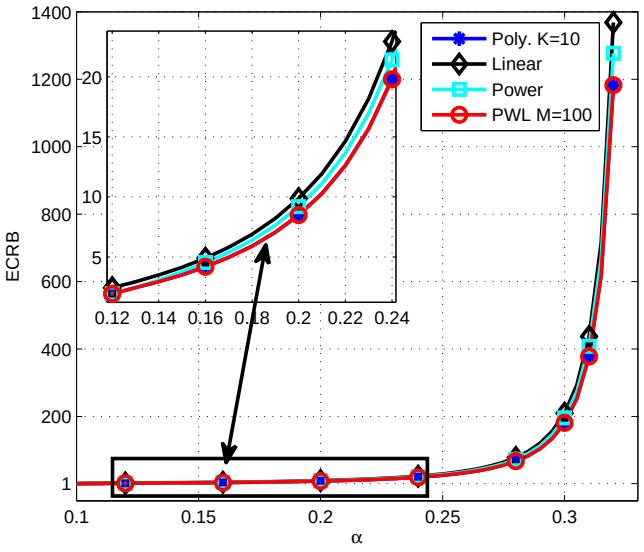

Fig. 2: ECRB versus α for various solution approaches, where $h = 1$ and $0.1 \leq \alpha \leq 0.32$.

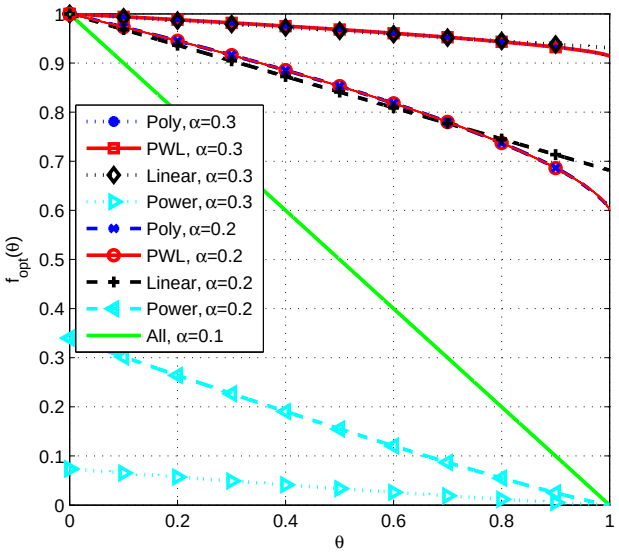

Fig. 3: $f_{opt}(\theta)$ versus θ for various solution approaches, where $\alpha = 0.1$, 0.2, and 0.3.

or $f(\theta) = 1 - \theta$ is an optimal encoding function. Also, Proposition 3 states that the optimal encoding function can be searched among monotonically decreasing functions as the uniform distribution satisfies the symmetry condition. In addition, Corollary 2 reveals that if $\alpha \leq 4/39 = 0.1026$, then $f(\theta) = 1 - \theta$ is the optimal encoding function since such a secrecy level can be guaranteed by using $f(\theta) = 1 - \theta$. Furthermore, Proposition 4 claims that it is not possible to achieve a secrecy limit α higher than $1/3$ as $\gamma = 1 < 2/h_e = 2$ in this scenario.

For obtaining the encoding function based on the proposed approaches in Section IV, the linear and power encoding functions, and the polynomial and piecewise linear (PWL) approximations are considered. For the linear encoding, $f(\theta) = 1 - m\theta$ is used due to Proposition 5. Then, (39) provides a simple tool for the solution. For the power encoding function, $f(\theta) = p(1-\theta)^q$ is employed, and the optimal p and q values are found for a given target α value (see Remark 4). For the polynomial approximation (with a degree of $K = 10$) and the piecewise linear approximation (with $M = 100$ intervals), the formulations in (41) and (42) are utilized, respectively. In Fig. 2, the relation between the target level α and the optimal ECRB value can be observed. When $\alpha = 0.10$, it is noted that the optimal ECRB is 1, which can be achieved with $f(\theta) = 1-\theta$. As α increases, the optimal ECRB increases exponentially. For example, when $\alpha = 0.25$, the optimal ECRB is found to be 25.06 and it becomes 1182.3 when $\alpha = 0.32$ for the piecewise linear approximation. Hence, the ECRB goes to infinity as α goes to the theoretical bound of $1/3$, as expected.[5] In Fig. 3, the encoding functions corresponding to the proposed solution approaches are presented for various values of α. It is observed that the polynomial approximation and the piecewise linear approximation yield almost the same function, which can also be deduced from the performance graph in Fig. 2. It is also seen that when $\alpha = 0.1$, all the methods lead to $f(\theta) = 1 - \theta$. When $\alpha = 0.2$, the difference between the solutions of the linear encoding and the approximation methods becomes noticeable. Since the approximation methods can use higher degrees of freedom than the linear encoding, they can achieve lower ECRBs. However, the linear encoding provides a simple solution for this scenario. For example, when $\alpha = 0.2$, the optimal linear encoding function can be obtained by finding the largest $m \in (0, 1]$ that satisfies $m^4 - m^3 + 9.4m - 18m + 4.8 \geq 0$, yielding $m = 0.3184$ due to (40); hence, $f(\theta) = 1 - 0.3184\,\theta$. It is also observed that the performance of the optimal power encoding approach in terms of the ECRB and the computational complexity is in between those of the optimal linear encoding and the other two approaches.

Next, the effects of the channel quality h of the eavesdropper on the optimal ECRB and encoding function are investigated for a given value of α. For this purpose, $\alpha = 0.15$ is used and the ECRB performance is evaluated versus $h = h_e/\sigma_e$ in Fig. 4. As discussed before, as h increases, the distortion due to encoding is transmitted to the eavesdropper more effectively and the intended MSE can be generated with a lower ECRB. Some interesting observations can be made in Fig. 4. First, three different regions are noted for the ECRB. In the first region, the ECRB slowly decreases as h increases for all the solution approaches. In the second region, for the power and the approximation approaches, the ECRB decreases more rapidly and finally when h is above some threshold value, $f(\theta) = 1 - \theta$ becomes sufficient to generate the MSE value of $\alpha = 0.15$ at the eavesdropper. Actually, this threshold can be calculated analytically based on Corollary 2. For the parameters in the considered scenario, $V_u = 1/12$, $E(\theta) = 1/2$, and $\alpha = 0.15$; hence, $h_{th} = \sqrt{48/11} = 2.09$. It is observed that the performance

[5] In this example, the optimal ECRB value should not directly be taken as equal to the MSE at the estimator of the intended receiver since h_r/σ_r is not sufficiently high. Here, the ECRB is merely used as an objective function to represent generic estimation accuracy.

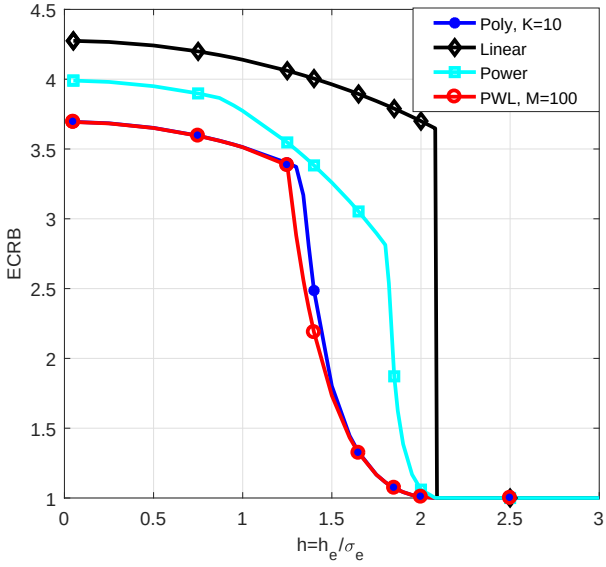

Fig. 4: ECRB versus h for various solution approaches when $\alpha = 0.15$ with uniform prior distribution.

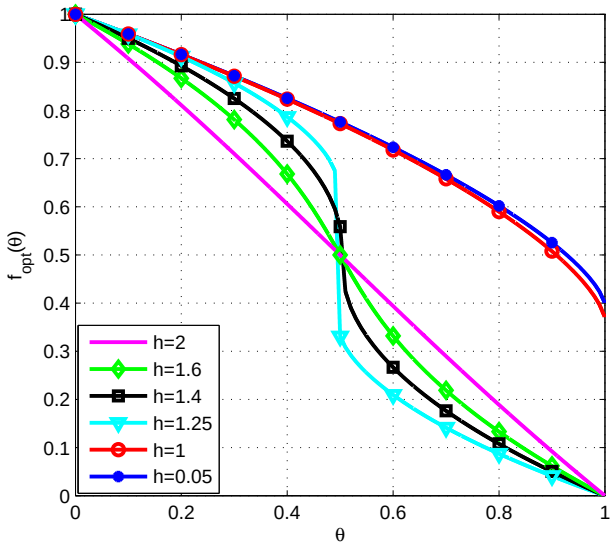

Fig. 5: $f_{opt}(\theta)$ versus θ for the piecewise linear approximation when $\alpha = 0.15$ with uniform prior distribution.

of the polynomial approximation is very similar to that of the piecewise linear approximation; however, in the second region, it is slightly worse than that of the piecewise linear approximation. The optimal encoding function corresponding to the piecewise linear approximation approach is presented in Fig. 5, which reveals that the encoding function changes characteristics as h increases. This also explains why the polynomial approximation is slightly worse than the piecewise linear approximation for medium values of h. Namely, for the polynomial approximation, it is harder to correctly implement the sudden decrease around $\theta = 0.5$ while it has sufficient degrees of freedom to produce an encoding function required for smaller h values as it can also be observed in Fig. 3. It is also noted that the encoding function is still continuous; that is, it has a finite but large derivative around $\theta = 0.5$. In addition, it is seen that when $h = 2$, the encoding function is almost linear.

Next, a scenario with a nonuniform prior distribution is considered, and the prior PDF of parameter θ is modeled as $w(\theta) = 2\theta$ for $\theta \in [0, 1]$. Similar to the uniform distribution case, the characteristics of the optimal encoding function are investigated for the fixed α and fixed h cases. First, it is assumed that $h = 1$ and the optimal encoding function is presented for various α values in Fig. 6 by using the piecewise linear approximation approach. The theoretical optimal solution $f(\theta) = 1 - \theta^{4/3}$ for the no constraint case is also shown in the figure, which is calculated based on Proposition 1 for the given prior distribution. It is observed that when the target level is small; i.e., $\alpha = 0.1$, the optimal encoding function calculated via the piecewise linear approximation is exactly the same as the theoretical solution. As α increases, in order to satisfy the target secrecy level, the optimal encoding function maps θ to lower values. It is noted that higher target α levels are achievable for this prior distribution as compared to the uniform distribution when $h = 1$. In particular, the secrecy limit is $1/2$ instead of $1/3$ in this example. Then, α is fixed as $\alpha = 0.34$, and the ECRB performance is investigated with respect to h in Fig. 7. It is noted that the performance trends of the different solution approaches are similar to those in the uniform case presented in Fig. 4; however, unlike the uniform distribution case, a sharp decrease to the minimum ECRB does not exist in this scenario (see Fig. 7). This is mainly due to the fact the optimal functions for the various function families yielding that minimum ECRB in the absence of an eavesdropper actually could not satisfy the secrecy requirement even if h gets large. For example, if the linear encoding with $m = 1$ is used, it can be shown that as $h \to \infty$, the resulting MSE is $1/3$, and if the theoretical solution for the no constraint case (that is, $f(\theta) = 1 - \theta^{4/3}$) is used, the resulting MSE is 0.318 as $h \to \infty$. It is known that the linear encoding with $m = 1$ would yield an ECRB value of 1 and $f(\theta) = 1 - \theta^{4/3}$ would yield an ECRB value of $27/32 = 0.844$. However, unlike the previous example, since the target α value is too high to achieve with those encoding functions, these minimum ECRB values cannot be attained even if h gets arbitrarily large; hence, a slow decay with a floor is observed in the ECRB instead of a sudden decrease. The ECRB floor values are found to be 1.5625, 1.0482, and 0.8835 for the linear encoding, the power encoding, and the piecewise linear approximation, respectively. Also, it is noted that the performance differences between the different solution approaches are small when h is low, which become more significant for medium values of h. Finally, the optimal solutions via the piecewise linear approximation are provided for various h values in Fig. 8. It is noted that the characteristics of the optimal encoding function are different for small, medium, and large values of h. One interesting observation is that for medium values of h, it is seen that the sudden decrease in the optimal encoding function does not necessarily happen at 0.5 unlike the uniform prior distribution case.

Finally, we provide the simulation times for obtaining the solutions of the various methods and the resulting ECRB

Solution Method	ECRB	Time (ms.)	Solution Method	ECRB	Time (ms.)
Linear Encoding	4.1395	0.35	Power Encoding	3.7730	35
Poly. App. ($K=2$)	3.6170	33	PWL App ($M=5$)	3.5634	159
Poly. App. ($K=4$)	3.5263	142	PWL App ($M=10$)	3.5289	302
Poly. App. ($K=6$)	3.5163	763	PWL App ($M=25$)	3.5159	750
Poly. App. ($K=8$)	3.5139	4680	PWL App ($M=50$)	3.5134	1483
Poly. App. ($K=10$)	3.5135	5540	PWL App ($M=100$)	3.5125	6220
Poly. App. ($K=14$)	3.5129	18102	PWL App ($M=200$)	3.5123	23687

TABLE I: ECRB values and simulation times for various approaches, where $\alpha = 0.15$.

.

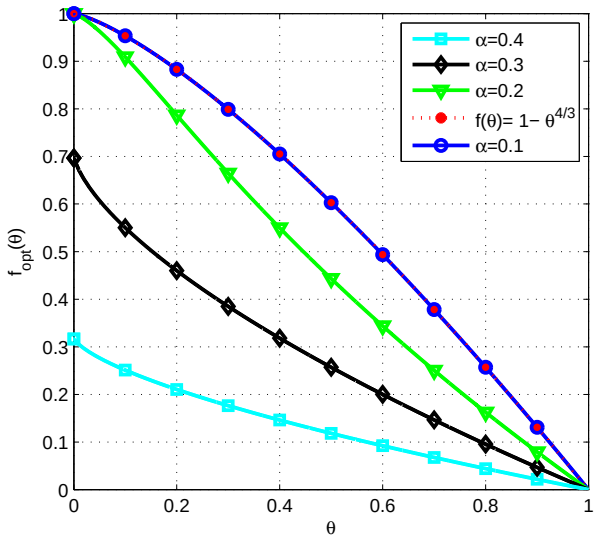

Fig. 6: $f_{opt}(\theta)$ versus θ for piecewise linear approximation ($M = 100$), where $\alpha = 0.1$, 0.2, 0.3, and 0.4. $f(\theta) = 1-\theta^{4/3}$ is the optimal function under no secrecy constraints according to Proposition 1.

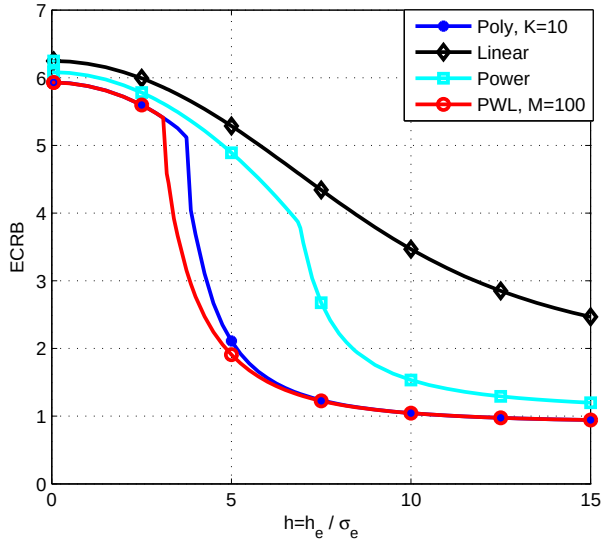

Fig. 7: ECRB versus h for various solution approaches when $\alpha = 0.34$ for $w(\theta) = 2\theta$ for $\theta \in [0, 1]$.

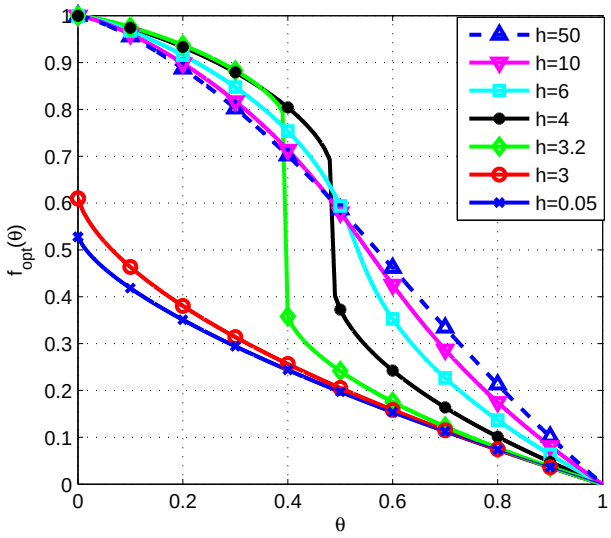

Fig. 8: $f_{opt}(\theta)$ versus θ for piecewise linear approximation when $\alpha = 0.34$ with $w(\theta) = 2\theta$ for $\theta \in [0, 1]$.

values in Table I for the scenario considered in Fig. 2 with $\alpha = 0.15$.[6] We observe that the linear and power encoding approaches have shorter solution times while they provide suboptimal solutions. For the polynomial and piecewise linear approximations, as K and M increase, the simulation times increase and lower ECRB values can be obtained. However, it is observed that after a certain value, the improvement in the ECRB is not significant. Therefore, it makes sense to choose the values of these parameters considering the solution times as well. In this study, we have used $K = 10$ and $M = 100$.

VI. Conclusion and Future Work

The optimal parameter encoding problem has been studied in the presence of an eavesdropper, where the aim is to minimize the ECRB at the intended receiver under the constraint of a target MSE value at the eavesdropper. A closed-form expression has been derived for the optimal encoding function when there is no secrecy constraint. When a certain secrecy level is to be guaranteed at the eavesdropper, first a sufficient condition has been provided for the case in which the optimal encoding function under no secrecy constraints is still optimal. Next, a closed-form expression for the MSE of the eavesdropper has been derived under the assumption

[6]The simulations are performed with Intel Core i5-4590 CPU 3.30 GHz processor and Matlab R2017B.

that the eavesdropper employs the linear MMSE estimation. Based on this result, the shift invariance property has been shown for generic prior PDFs, and it has been proved that it is sufficient to restrict the search to decreasing encoding functions if the prior distribution of the parameter has a certain symmetry property. In addition, an upper limit has been obtained for the MSE of the eavesdropper for the uniform prior distribution. This result implies that the optimal encoding function either maximizes or minimizes the variance of the encoded parameter depending on the channel quality parameter and the length of the range interval for the encoding function. In order to calculate the optimal encoding function numerically, various solution approaches have been considered; namely, linear encoding, polynomial approximation, and piecewise linear approximation. It has also been shown that the optimal solution for the linear encoding function can be obtained algebraically. Numerical results have considered both uniform and nonuniform parameter distributions, and provided the optimal solutions based on the proposed techniques. The future work is to investigate the extension of the analytical results to the cases in which the eavesdropper employs the MMSE estimator. Another interesting extension would be to formulate the problem in a game theoretic framework, where the eavesdropper has some partial information about transmitter's strategy and the transmitter considers this possibility in the design of the encoding function.

APPENDIX

A. Derivation of (24) and (25)

The linear MMSE estimator $\hat{\beta}(Z)$ to estimate β based on Z can be expressed as [37]

$$\hat{\beta}(Z) = E(\beta) + \frac{Cov(\beta, Z)}{Var(Z)}(Z - E(Z)) \tag{44}$$

From $Z = h_e\beta + N_e$, the following relations are obtained:

$$\begin{aligned}\hat{\beta}(Z) &= E(\beta) + \frac{Cov(\beta, h_e\beta + N_e)}{Var(h_e\beta + N_e)}(Z - h_eE(\beta)) \qquad (45)\\ &= E(\beta) + \frac{h_eVar(\beta)}{h_e^2Var(\beta) + \sigma_e^2}(Z - h_eE(\beta))\\ &= \frac{h_eVar(\beta)}{h_e^2Var(\beta) + \sigma_e^2}Z + \left(1 - h_e\frac{h_eVar(\beta)}{h_e^2Var(\beta) + \sigma_e^2}\right)E(\beta)\end{aligned}$$

where the second inequality is due to the independence of β and N_e.

B. Derivation of (26)

The eavesdropper is modeled to employ the linear MMSE estimator specified by $\hat{\beta}(z) = k_0 + k_1 z$, where k_1 and k_0 are given by (24) and (25), respectively. Defining $\beta = f(\theta)$, $V = Var(\beta)$, $C = Cov(\beta, \theta)$, and $h = h_e/\sigma_e$, the MSE at the eavesdropper can be written as

$$\begin{aligned}&E\left(\left|\hat{\beta}(Z) - \theta\right|^2\right) = E\left((k_1Z + k_0 - \theta)^2\right) \qquad (46)\\ &= E\left(k_1^2Z^2 + 2k_1k_0Z + k_0^2 + \theta^2 - 2(k_1Z + k_0)\theta\right) \qquad (47)\\ &= k_1^2E\left(h_e^2\beta^2 + 2h_e\beta N + N^2\right) + 2k_1k_0h_eE(\beta)\\ &+ k_0^2 + E(\theta^2) - 2k_1h_eE(\theta\beta) - 2k_0E(\theta) \qquad (48)\end{aligned}$$

where (48) follows from that facts that $Z = h_e\beta + N$ and $E(Z) = h_eE(\beta)$. In addition, it is known that $E(N^2) = \sigma_e^2$, and θ and N are independent random variables with $E(N) = 0$; hence, $E(\beta N) = 0$. Then, the expression in (48) is further processed as follows:

$$\begin{aligned}&E\left(\left|\hat{\beta}(Z) - \theta\right|^2\right)\\ &= k_1^2h_e^2E(\beta^2) + k_1^2\sigma_e^2 + 2k_1k_0h_eE(\beta)\\ &+ k_0^2 + E(\theta^2) - 2k_1h_eE(\theta\beta) - 2k_0E(\theta) \qquad (49)\\ &= k_1^2h_e^2E(\beta^2) + k_1^2\sigma_e^2 + 2k_1(1 - k_1h_e)h_eE(\beta)^2\\ &+ E(\beta)^2(1 + k_1^2h_e^2 - 2k_1h) + E(\theta^2)\\ &- 2k_1h_eE(\theta\beta) - 2(1 - k_1h_e)E(\beta)E(\theta) \qquad (50)\\ &= k_1^2h_e^2(E(\beta^2) - E(\beta)^2) + k_1^2\sigma_e^2 + E(\theta^2)\\ &+ E(\beta)^2 - 2k_1h_e(E(\beta\theta) - E(\beta)E(\theta))\\ &- 2E(\beta)E(\theta) = k_1^2h_e^2V + k_1^2\sigma_e^2 - 2k_1h_eC\\ &+ E(\theta^2) - E(\theta)^2 + E(\theta)^2 + E(\beta)^2 - 2E(\beta)E(\theta) \qquad (51)\\ &= k_1^2(h_e^2V + \sigma_e^2) - 2k_1h_eC + Var(\theta) + (E(\beta) - E(\theta))^2 \qquad (52)\\ &= \frac{h_e^2V^2 - 2h_e^2VC}{h_e^2V + \sigma_e^2} + Var(\theta) + (E(\beta) - E(\theta))^2 \qquad (53)\\ &= \frac{(h_e/\sigma_e)^2V(V - 2C)}{(h_e/\sigma_e)^2V + 1} + Var(\theta) + (E(\beta) - E(\theta))^2 \qquad (54)\\ &= \frac{h^2V(V - 2C)}{h^2V + 1} + Var(\theta) + (E(\beta) - E(\theta))^2 \qquad (55)\end{aligned}$$

where (49) follows directly from (48), (50) is obtained by inserting (25) into (49), (51) follows by rearranging the terms and adding and subtracting $E(\theta)^2$ in (50), (53) is obtained by inserting (24) into (52), and finally (55) is due to the use of $h = h_e/\sigma_e$ in (54). ■

REFERENCES

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

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

[3] C. E. Shannon, "Communication theory of secrecy systems," *Bell System Technical Journal*,, vol. 28, no. 4, pp. 656–715, 1949.

[4] A. D. Wyner, "The wire-tap channel," *Bell System Technical Journal*, vol. 54, no. 8, pp.: 1355-1387, 1975.

[5] Y. Liang, H. V. Poor and S. Shamai, "Secure communication over fading channels," *IEEE Trans. on Inform. Theory*, vol. 54, no. 6, pp. 2470-2492, June 2008.

[6] P. K. Gopala, L. Lai and H. El Gamal, "On the secrecy capacity of fading channels," *EEE Trans. on Inform. Theory*, vol. 54, no. 10, pp. 4687-4698, Oct. 2008.

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

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

[9] O. O. Koyluoglu, H. El Gamal, L. Lai, and H. V. Poor, "Interference alignment for secrecy," *IEEE Trans. on Inform. Theory*, vol. 57, pp. 3323-3332, June 2011.

[10] J. Zhu, J. Mo, and M. Tao, "Cooperative secret communication with artificial noise in symmetric interference channel," *IEEE Commun. Letters*, vol. 14, pp. 885-887, Oct. 2010.

[11] R. Bassily et. al., "Cooperative security at the physical layer: A summary of recent advances," *IEEE Signal Proc. Mag.*, vol. 30, no. 5, pp. 16-28, Sep. 2013.

[12] Y. W. P. Hong, P. C. Lan and C. C. J. Kuo, "Enhancing physical-layer secrecy in multiantenna wireless systems: An overview of signal processing approaches," *IEEE Signal Proc. Mag.*, vol. 30, no. 5, pp. 29-40, Sep. 2013.

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

[14] J. Yang, I. M. Kim and D. I. Kim, "Optimal cooperative jamming for multiuser broadcast channel with multiple Eavesdroppers," *IEEE Trans. Wireless Comm.*, vol. 12, no. 6, pp. 2840-2852, June 2013.

[15] S. Asoodeh, F. Alajaji and T. Linder, "Privacy-aware MMSE estimation," *IEEE International Symposium on Information Theory (ISIT)*, Barcelona, July 2016, pp. 1989-1993.

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

[17] B. Kailkhura, V. S. Siddhardh Nadendla and P. K. Varshney, "Distributed inference in the presence of eavesdroppers: a survey," *IEEE Communications Magazine*, vol. 53, no. 6, pp. 40-46, June 2015.

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

[19] X. Guo, A. S. Leong and S. Dey, "Distortion outage minimization in distributed estimation with estimation secrecy outage constraints," *IEEE Trans. on Signal and Inform. Proccessing on Networks*, vol. 3, no. 1, pp. 12-28, March 2017.

[20] T.C. Aysal and K. E. Barner, "Sensor data cryptography in wireless sensor networks," *IEEE Trans. Info. Forensics and Security*, vol. 3, no. 2, pp. 273-289, June 2008.

[21] F. Farokhi and H. Sandberg, "Fisher information as a measure of privacy: Preserving privacy of households with smart meters using batteries," *IEEE Trans. on Smart Grid*, vol. PP, no. 99, pp. 1-1. doi: 10.1109/TSG.2017.2667702.

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

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

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

[25] H. L. Van Trees and K. L. Bell (editors), *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking.* New Jersey: Wiley-Interscience, 2007.

[26] R. D. Gill and B. Y. Levit, "Application of the Van Trees inequality: A Bayesian CramérRao bound," *Bernoulli*, no. 1, pp. 5979, 1995.

[27] E. Gonendik and S. Gezici, "Fundamental limits on RSS based range estimation in visible light positioning systems," *IEEE Communications Letters*, vol. 19, no. 12, pp. 2138-2141, Dec. 2015.

[28] M.F. Keskin, E. Gonendik and S. Gezici, "Improved lower bounds for ranging in synchronous visible light positioning systems," *Journal of Lightwave Technology*, vol. 34, no. 23, pp. 5496–5504, Dec. 2016.

[29] A. A. Nasir et al., "Optimal training sequences for joint timing synchronization and channel estimation distributed communication networks," *IEEE Trans. on Communications*, vol. 61, no. 7, pp. 3002–3015, July 2013.

[30] A. Shahmansoori, R. Montalban, and G. Seco-Granados, "Effect of channel variability on pilot design for joint communications and positioning in OFDM," *11th International Symposium on Wireless Communications Systems (ISWCS)*, pp. 292–296, Barcelona, 2014.

[31] H. Gazzah and J. P. Delmas, "Direction finding antenna arrays for the randomly located source," *IEEE Trans. on Signal Proc.*, vol. 60, no. 11, pp. 6063-6068, Nov. 2012.

[32] A. Khisti, G. Wornell, A. Wiesel and Y. Eldar, "On the Gaussian MIMO Wiretap Channel," *2007 IEEE International Symposium on Information Theory*, Nice, 2007, pp. 2471-2475.

[33] A. Guillen i Fabregas and G. Caire, "Coded modulation in the block-fading channel: coding theorems and code construction," *IEEE Trans. on Inform. Theory*, vol. 52, no. 1, pp. 91-114, Jan. 2006.

[34] R. Knopp and P. A. Humblet, "On coding for block fading channels," *IEEE Trans. on Inform. Theory*, vol. 46, no. 1, pp. 189-205, Jan 2000.

[35] I. M. Gelfand and S. V. Fomin, *Calculus of Variations*, Prentice-Hall, 1963.

[36] David G. Luenberger, *Optimization by Vector Space Methods*, John Wiley & Sons, Inc., 1997

[37] E. W. Karmen and J. K. Su, *Introduction to Optimal Estimation*, London, U.K.: Springer-Verlag, 1999.

[38] H. Jeffreys and B. S. Jeffreys, "Weierstrass's Theorem on approximation by polynomials," *Methods of Mathematical Physics*, pp. 446-448, Cambridge University Press, 3rd ed., 1988.

[39] K. Atkinson, *An Introduction to Numerical Analysis*, John Wiley, 2nd edition, 1989.

[40] T. J. Rivlin, *An Introduction to the Approximation of Functions*, Dover, New York, 1981.

[41] G. G. Lorentz, *Approximation of Functions*, Chelsea, New York, 1986.

[42] R. Montalban, J. A. López-Salcedo, G. Seco-Granados and A. L. Swindlehurst, "Power allocation method based on the channel statistics for combined positioning and communications OFDM systems," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Vancouver, BC, 2013, pp. 4384–4388.

Cagri Goken received the B.S. and M.S degrees from Bilkent University, Ankara, Turkey and the M.A. degree from Princeton University, NJ, USA, all in electrical engineering, in 2009, 2011 and 2014 respectively. He is currently working towards his Ph.D. degree at Bilkent University. He has also been with Aselsan Inc. since 2016, 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, Turkey in 2001, and the Ph.D. degree in Electrical Engineering from Princeton University in 2006. From 2006 to 2007, he worked at Mitsubishi Electric Research Laboratories, Cambridge, MA. Since 2007, he has been with the Department of Electrical and Electronics Engineering at Bilkent University, where he is currently a Professor. Dr. Gezici's research interests are in the areas of detection and estimation theory, wireless communications, and localization systems. Among his publications in these areas is the book Ultra-wideband Positioning Systems: Theoretical Limits, Ranging Algorithms, and Protocols (Cambridge University Press, 2008). Dr. Gezici was an associate editor for IEEE Transactions on Communications, IEEE Wireless Communications Letters, and Journal of Communications and Networks.