Estimation Theoretic Optimal Encoding Design for Secure Transmission of Multiple Parameters

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

***Abstract*—In this paper, optimal deterministic encoding of a vector parameter is investigated in the presence of an eavesdropper. The objective is to minimize the expectation of the conditional Cramér–Rao bound at the intended receiver, while satisfying an individual secrecy constraint on the mean-squared error of estimating each parameter at the eavesdropper. The eavesdropper is modeled to employ the linear minimum mean-squared error estimator based on the noisy observation of the encoded parameter without being aware of encoding. First, the problem is formulated as a constrained optimization problem in the space of vector-valued functions. Then, two practical solution strategies are developed based on nonlinear individual encoding and affine joint encoding of parameters. Theoretical results on the solutions of the proposed strategies are provided for various scenarios on channel conditions and parameter distributions. Finally, numerical examples are presented to illustrate the performance of the proposed solution approaches.**

***Index Terms*—Fisher information matrix (FIM), parameter estimation, Cramér-Rao bound (CRB), secrecy, optimization.**

I. Introduction

SECURE transmission of data to an intended receiver in the presence of an eavesdropper has been a crucial problem for communications. Physical layer secrecy is based on the idea of exploiting the randomness in wireless channels to ensure secure communication. In recent years, there has been a renewed interest in the physical layer secrecy with the advances in wireless communication systems. As the age of Internet of Things (IoT), smart homes and cities, and wireless sensor networks with vast amount of nodes has already arrived, ensuring the security of data in such networks appears to be a challenging task. Key-based cryptographic approaches such as [1] and [2] have been employed in many applications to ensure confidential communication, and they may still be a valuable option and even necessary for certain applications such as military communications. However, as the management of key generation and distribution can be very challenging in heterogeneous and dynamic networks with a vast number of device connections, cryptographic approaches may no longer be the most suitable solution [3], [4].

Manuscript received October 20, 2018; revised May 6, 2019 and June 30, 2019; accepted July 7, 2019. Date of publication July 22, 2019; date of current version July 31, 2019. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Laura Cottatellucci. *(Corresponding author: Cagri Goken.)*

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

Digital Object Identifier 10.1109/TSP.2019.2929921

Traditionally, information theoretical metrics such as mutual information have been employed to quantify the secrecy levels in physical layer security over wireless networks [5]–[10]. In particular, 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 [5]. Alternatively, estimation theoretic tools such as mean-squared error (MSE) and Fisher information have recently been used to measure security performance of communication systems to design low-complexity, practical and secure systems [11]–[22].

Estimation theoretic security has found applications in a wide variety of problems. For example, such tools can be employed in distributed inference networks, where the information coming to a fusion center from various sensor nodes can also be observed by eavesdroppers [12]–[14]. In [11], the secret communication problem is investigated for Gaussian interference channels in the presence of eavesdroppers for vector parameters. 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 level, where joint artificial noise and linear precoding schemes are used to satisfy the secrecy constraint. In [15], privacy of households using smart meters is considered in the presence of adversary parties. The Fisher information is employed as a metric of privacy for both scalar and multivariable case and the optimal policies for the utilization of batteries are derived to achieve privacy. In [16], a decentralized estimation problem is considered in an insecure sensor network environment, where each sensor network performs stochastic encryption based on the 1-bit quantized version of a noisy sensor measurement to achieve secret communication. In [17], the optimal deterministic encoding of scalar parameters is investigated based on the minimization of the expectation of 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 [18], a robust parameter encoding approach is developed and the optimization is based on the worst-case CRB (equivalently, the worst-case Fisher information) of the parameter in order to guarantee a certain level of estimation accuracy at the intended receiver.

In the estimation theoretic secrecy framework, Fisher information and Cramér-Rao bounds provide crucial metrics to evaluate performance of estimators and have been employed in various security problems [15]–[18]. Even though the CRB

and the Fisher information for a given value of a parameter of interest have very clear interpretations as a measure of estimation efficiency, they are not directly applicable in the Bayesian framework. In such a case, the expectation of the conditional Cramér-Rao bound (ECRB), can be utilized as a metric of estimation accuracy, when the prior information about transmitted parameters is available [23]. The ECRB has been employed in various different contexts in the literature [24], [25], and utilized as a metric to quantify estimation accuracy in security/privacy problems [15], [17]. In particular, the ECRB facilitates theoretical investigations for achieving intuitive understanding of the parameter encoding problem and it does not assume any fixed estimator structure in order to be calculated [17]. Also, the MSE of the MAP estimator converges to the ECRB in the high SNR region [23]; hence, ECRB-based optimization also guarantees optimizing the performance of certain practical estimators. Based on all these reasons, the ECRB is employed in this study, as well.

Even though the optimal parameter encoding problem has been investigated for scalar parameters in [17] and [18] from a CRB-based optimization perspective, it is possible that the channel input can contain multiple parameters in many practical scenarios such as [11], [15], [19]–[22]. Estimation of multiple parameters is required in many applications such as in localization [26] and joint frequency and phase estimation [23]. Secure transmission of multiple parameters has also been investigated in the literature for different applications and scenarios. In [19], the filter design with secrecy constraints is studied for a multiple-input multiple-output (MIMO) Gaussian wiretap channel, where the parameter of interest is a vector, each component of which is zero mean with a unit variance and is independent of others. In [20], a beamforming scheme is proposed for a downlink multiuser MIMO system for secure communication, where the vector parameter carries the unit-energy data symbols of each user. In [21], the binary stochastic encryption introduced in [16] is extended to the vector parameter estimation case. Another important use-case for the secure multiple parameter estimation problem occurs in smart grids/homes and internet of things (IoT) systems [22]. For example, the vector parameter carries the state of the grid, i.e., the voltage angles and magnitudes at each of the buses, in the scenario of state estimation problem in a smart-grid system. In another example, the parameter is the state of the position and velocity of an autonomous vehicle. In a further example, the parameter represents the pollutant concentration over an entire city in an air monitoring system in a smart city, where each individual component of the vector can represent the pollutant concentration in a certain neighborhood [22].

Based on the preceding motivations, we focus on a secure multi-parameter transmission scenario in this study. Similarly to [17] and [18], the parameter is encoded using an encoding function prior to transmission. It is important to emphasize that the difference of the multiparameter scenario investigated in this manuscript from the single parameter case studied in [17] is not only based on the number of parameters. In the encoding of a scalar parameter, a single scalar valued function is utilized as an encoder. In this manuscript, as the parameter of interest is a random vector, the encoding function becomes a vector valued function, which generates different opportunities compared to the scalar case during the encoding operation such as joint encoding of parameters using a nonlinear function. As a simple example, consider a scenario in which the parameter involves the coordinates of the location of a target. Then, before sending the true coordinate, a simple shuffle of the coordinates can create a considerable amount of localization error at the eavesdropper as the eavesdropper is not aware that such a secret-key is employed. This means that the problem of optimal encoding of multiple parameters requires new analyses and theoretical investigations as the theoretical analysis and tools employed in [17] are not able to cover it directly in general. When the encoding function is assumed to be an affine function as a special case, it corresponds to employing a linear precoding matrix strategy, which has been employed in various studies to ensure security [11], [12].

In this work, the objective of encoding design is to minimize the ECRB, which is defined as the average of the trace of the inverse Fisher Information Matrix (FIM). The eavesdropper is modeled to employ the linear MMSE (LMMSE) estimator based on the noisy observation of the encoded parameter without being aware of encoding. Compared to other studies in the estimation theoretic security literature, the proposed formulation is a novel approach for problems involving multiple parameters. Also, the possible correlations among the parameters and the correlations in the noise components of intended receiver/eavesdropper are taken into account, which is not applicable in [17]. First, the optimization problem is formulated to obtain the optimal encoding function for a given target MSE level based on the assumption that the joint encoding approach is applied via a nonlinear encoding function. Based on this formulation, two special cases of the generic form of the encoding function is studied to develop practical encoders. In the first approach, each element of the vector parameter is encoded individually by a nonlinear scalar function. For this strategy, it is shown that when the transmitted parameters are independent and the channel noise for the eavesdropper is white, the optimization problem decouples into individual scalar problems, which are investigated in [17]. Then, the case for colored Gaussian noise for the eavesdropper is investigated, where the optimization problem cannot be decoupled. For the two-parameters case, fundamental insights are provided about the optimal solution of the multiple parameter case by considering the correlation in the noise components, which cannot be obtained by studying the single parameter case. In the second approach, the encoding function is assumed to be an affine function. This method allows for joint encoding, or simple shuffle and scale of the parameters, which cannot be utilized in the single parameter case. Therefore, all the theoretical analyses related to this approach are new contributions. For this strategy, first the secrecy requirements are omitted, and an optimal solution is derived theoretically when the channel noise for the intended receiver is white. Next, the MSE constraint for the eavesdropper is considered and several theoretical results are provided regarding the form of the optimal affine joint encoder. Finally, numerical examples are provided to investigate various scenarios for both nonlinear individual encoding and affine joint encoding strategies. The

main contributions in this manuscript can be summarized as follows:

- The optimal encoding of multiple parameters is proposed by utilizing the ECRB metric at the intended receiver and a MSE target at the eavesdropper. Two practical encoding strategies, nonlinear individual encoding and affine joint encoding, are introduced as possible encoding solutions.
- For nonlinear individual encoding, it is shown that the optimization problem can be decoupled into independent problems if the channel noise for the eavesdropper is white and parameters are independent. It is also proved that if the prior distribution of a given parameter is symmetric on the domain, then the corresponding encoding function can be limited to decreasing functions.
- For affine joint encoding, the optimal encoding function is provided when there is no secrecy constraints and the channel noise for intended receiver is white.
- It is shown that the search for the optimal affine encoding strategy can be converted to a precoding matrix search; that is, the constant term can be eliminated from the optimization problem.

The rest of the manuscript is organized as follows: The optimal encoding problem for multiple parameters is formulated in Section II. The nonlinear individual encoding strategy and affine joint encoding strategies are studied in Sections III and IV, respectively. Numerical results are presented in Section V and concluding remarks are given in Section VI.

II. Problem Formulation

Consider a scenario in which N-dimensional random vector parameter $\boldsymbol{\theta} = [\theta_1\, \theta_2 \cdots \theta_N]^T \in \Lambda$ is to be transmitted to an intended receiver over N channels, and $w(\boldsymbol{\theta})$ denotes the joint probability density function (PDF) of $\boldsymbol{\theta}$. A block fading channel model is assumed such that the instantaneous fading coefficient at each channel is independent and denoted by constant $h_{r,i}$ for $i = 1, 2, \ldots, N$. As this model considers a slowly fading channel, it is assumed that the channel coefficients are constant during the transmission of the parameters. In addition to the transmitter and the intended receiver, there exists an eavesdropper that tries to estimate the parameter $\boldsymbol{\theta}$. The objective is to perform accurate estimation of the parameter at the intended receiver while keeping the estimation error at the eavesdropper above a certain level [17]. Therefore, vector parameter $\boldsymbol{\theta}$ is encoded by using a vector-valued encoding function $\boldsymbol{f} : \Lambda \to \Gamma$ before the transmission of the parameter.[1] Let $\boldsymbol{\beta} \in \Gamma$ be the encoded version of the parameter, which is defined as

$$\boldsymbol{\beta} \triangleq \boldsymbol{f}(\boldsymbol{\theta}) = \begin{bmatrix} f_1(\theta_1, \theta_2, \ldots, \theta_N) \\ f_2(\theta_1, \theta_2, \ldots, \theta_N) \\ \vdots \\ f_N(\theta_1, \theta_2, \ldots, \theta_N) \end{bmatrix}. \tag{1}$$

[1]The encoder is designed for each transmission block and should be updated when the channel realization changes.

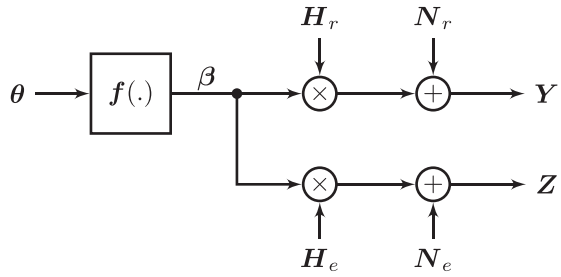

Fig. 1. System model.

Then, the received signal at the intended receiver is expressed as

$$\boldsymbol{Y} = \boldsymbol{H}_r \boldsymbol{\beta} + \boldsymbol{N}_r \tag{2}$$

where $\boldsymbol{H}_r = \mathrm{diag}\{h_{r,1}, h_{r,2}, \ldots, h_{r,N}\}$ is an $N \times N$ diagonal matrix of channel coefficients and $\boldsymbol{N}_r$ is the N-dimensional channel noise which is modeled as a zero-mean Gaussian random vector with covariance matrix $\boldsymbol{\Sigma}_r$ and is independent of $\boldsymbol{\theta}$. On the other hand, the eavesdropper observes

$$\boldsymbol{Z} = \boldsymbol{H}_e \boldsymbol{\beta} + \boldsymbol{N}_e \tag{3}$$

where $\boldsymbol{N}_e$ is zero-mean Gaussian noise with covariance matrix $\boldsymbol{\Sigma}_e$, which is also independent of $\boldsymbol{\theta}$, and $\boldsymbol{H}_e = \mathrm{diag}\{h_{e,1}, h_{e,2}, \ldots, h_{e,N}\}$ is an $N \times N$ diagonal matrix representing the channel between the transmitter and the eavesdropper under a block fading channel model. The intended receiver tries to estimate parameter $\boldsymbol{\theta}$ based on observation $\boldsymbol{Y}$ whereas the eavesdropper employs observation $\boldsymbol{Z}$ for estimating $\boldsymbol{\theta}$, as illustrated in Fig. 1. Note that the eavesdropper is not aware of encoding; hence, it effectively tries to estimate $\boldsymbol{\beta}$.

In order to measure estimation accuracy at the intended receiver, the expectation of Cramer-Rao bound (ECRB) is employed similarly to [17]. It is also assumed that the eavesdropper employs the LMMSE estimator $\hat{\boldsymbol{\beta}}(\boldsymbol{Z})$ whose coefficients are selected to estimate $\boldsymbol{\beta} = \boldsymbol{f}(\boldsymbol{\theta})$ based on $\boldsymbol{Z}$. The secrecy goal is achieved when the MSE at the eavesdropper for each θ_i is above a certain threshold. The ECRB for vector parameters can be expressed as [23]

$$E_{\boldsymbol{\theta}}\big(\boldsymbol{I}(\boldsymbol{\theta})^{-1}\big) = \int_{\Lambda} w(\boldsymbol{\theta})\; \boldsymbol{I}(\boldsymbol{\theta})^{-1} d\boldsymbol{\theta} = ECRB \tag{4}$$

where $\boldsymbol{I}(\boldsymbol{\theta})$ represents the Fisher information matrix (FIM), which is given by

$$I(\boldsymbol{\theta}) = E\left(\left(\frac{\partial p_{\boldsymbol{Y}|\boldsymbol{\theta}}(\mathbf{y}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right)\left(\frac{\partial p_{\boldsymbol{Y}|\boldsymbol{\theta}}(\mathbf{y}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right)^T\right) \tag{5}$$

with $p_{\boldsymbol{Y}|\boldsymbol{\theta}}(\mathbf{y}|\boldsymbol{\theta})$ representing the conditional PDF of $\boldsymbol{Y}$ for a given value of $\boldsymbol{\theta}$ [26]. Also, the error covariance matrix at the eavesdropper, who is unaware of the encoding, based on the estimate of the eavesdropper $\hat{\boldsymbol{\beta}}(\boldsymbol{Z})$ and the true value of the parameter $\boldsymbol{\theta}$ is defined as

$$\boldsymbol{\Sigma}_{err} = E\Big(\big(\hat{\boldsymbol{\beta}}(\boldsymbol{Z}) - \boldsymbol{\theta}\big)\big(\hat{\boldsymbol{\beta}}(\boldsymbol{Z}) - \boldsymbol{\theta}\big)^T\Big). \tag{6}$$

The expression in (4) is a matrix with each diagonal element representing the estimation accuracy limit for an individual parameter. Therefore, to determine the optimal encoding function for the overall vector parameter, the cost function is based on the sum of the diagonal elements of the inverse FIM, and the optimal parameter encoding problem is proposed as follows:

$$\begin{aligned}\boldsymbol{f}_{opt} = \arg\min_{\boldsymbol{f}} \int_{\Lambda} w(\boldsymbol{\theta})\ tr\{\boldsymbol{I}(\boldsymbol{\theta})^{-1}\} d\boldsymbol{\theta} \\ \text{s.t. } \boldsymbol{\Sigma}_{err}(i) \geq \eta_i\,, \quad i = 1, 2, \ldots, N. \end{aligned} \tag{7}$$

where $tr\{\cdot\}$ denotes the trace operator, $\boldsymbol{\Sigma}_{err}(i)$ is the ith diagonal element of $\boldsymbol{\Sigma}_{err}$, and η_i is the MSE target for θ_i at the eavesdropper.

It is important to emphasize that (7) involves optimization in the space of vector-valued functions with multiple inputs, hence it is difficult to solve in general. In the following sections, two special cases of the generic form of the encoding function given in (1) are considered as practical solution approaches.

Remark 1: Note that the closed-form expression for $\boldsymbol{\Sigma}_{err}$ can be derived in the following way (similarly to the derivation for the scalar case in [17]). The LMMSE estimator $\hat{\boldsymbol{\beta}}(\boldsymbol{Z})$ is expressed as $\hat{\boldsymbol{\beta}}(\boldsymbol{Z}) = \mathbf{A}\boldsymbol{Z} + \boldsymbol{b}$, where $\mathbf{A}$ and $\boldsymbol{b}$ are chosen to minimize $E(||\hat{\boldsymbol{\beta}}(\boldsymbol{Z}) - \boldsymbol{\beta}||^2)$, as the eavesdropper is unaware of the encoding, and are given by

$$\mathbf{A} = \boldsymbol{\Sigma}_{\boldsymbol{\beta},\boldsymbol{Z}} \left(\boldsymbol{H}_e \boldsymbol{\Sigma}_{\boldsymbol{\beta}} \boldsymbol{H}_e^T + \boldsymbol{\Sigma}_e\right)^{-1}, \tag{8}$$

and

$$\boldsymbol{b} = (\mathbf{I} - \mathbf{A}\boldsymbol{H}_{\mathbf{e}})\, E(\boldsymbol{\beta}), \tag{9}$$

with

$$\boldsymbol{\Sigma}_{\boldsymbol{\beta},\boldsymbol{Z}} = E\left((\boldsymbol{\beta} - E(\boldsymbol{\beta}))\left(\boldsymbol{Z} - E(\boldsymbol{Z})^T\right)\right). \tag{10}$$

Based on (8)–(10), $\boldsymbol{\Sigma}_{err}$ can be obtained as

$$\begin{aligned}\boldsymbol{\Sigma}_{err} = \boldsymbol{\Sigma}_{\boldsymbol{\beta}} \mathbf{R} \boldsymbol{\Sigma}_{\boldsymbol{\beta}}^T - \boldsymbol{\Sigma}_{\boldsymbol{\beta}} \mathbf{R} \boldsymbol{\Sigma}_{\boldsymbol{\beta},\boldsymbol{\theta}} - \boldsymbol{\Sigma}_{\boldsymbol{\beta},\boldsymbol{\theta}}^T \mathbf{R} \boldsymbol{\Sigma}_{\boldsymbol{\beta}}^T + \boldsymbol{\Sigma}_{\boldsymbol{\theta}} \\ + \left((E(\boldsymbol{\beta}) - E(\boldsymbol{\theta}))\,(E(\boldsymbol{\beta}) - E(\boldsymbol{\theta}))^T\right), \end{aligned} \tag{11}$$

where

$$\begin{aligned}\boldsymbol{\Sigma}_{\boldsymbol{\beta}} &= E\big(\boldsymbol{\beta}\boldsymbol{\beta}^T\big) - E(\boldsymbol{\beta})E(\boldsymbol{\beta})^T, \\ \boldsymbol{\Sigma}_{\boldsymbol{\beta},\boldsymbol{\theta}} &= E\big(\boldsymbol{\beta}\boldsymbol{\theta}^T\big) - E(\boldsymbol{\beta})E(\boldsymbol{\theta})^T, \\ \boldsymbol{\Sigma}_{\boldsymbol{\theta}} &= E\big(\boldsymbol{\theta}\boldsymbol{\theta}^T\big) - E(\boldsymbol{\theta})E(\boldsymbol{\theta})^T, \\ \mathbf{R} &= \boldsymbol{H}_e^T \left(\boldsymbol{H}_e \boldsymbol{\Sigma}_{\boldsymbol{\beta}} \boldsymbol{H}_e^T + \boldsymbol{\Sigma}_e\right)^{-1} \boldsymbol{H}_e. \end{aligned} \tag{12}$$

III. Nonlinear Individual Encoding

In this section, the proposed problem in Section II is investigated for an encoding approach such that each parameter θ_i is encoded *individually* by a nonlinear scalar function such that

$$\boldsymbol{\beta} \triangleq \boldsymbol{f}(\boldsymbol{\theta}) = \begin{bmatrix} f_1(\theta_1) \\ f_2(\theta_2) \\ \vdots \\ f_N(\theta_N) \end{bmatrix}. \tag{13}$$

Furthermore, as motivated in [17], the parameter space and the intrinsic constraints on each encoding function $f_i(\theta_i)$ are specified as follows:

- $\theta_i \in [a_i, b_i]$ for $i = 1, 2, \ldots, N$.
- $\beta_i = f_i(\theta_i) \in [a_i, b_i]$ for $i = 1, 2, \ldots, N$.
- f_i is a continuous and one-to-one function.

Under these assumptions, the optimal encoding problem in (7) can be written as

$$\begin{aligned}\boldsymbol{f}_{opt} = \arg \min_{f_1(\theta_1),\ldots,f_N(\theta_N)} \int_{\Lambda} w(\boldsymbol{\theta})\ tr\{\boldsymbol{I}(\boldsymbol{\theta})^{-1}\} d\boldsymbol{\theta} \\ \text{s.t. } \boldsymbol{\Sigma}_{err}(i) \geq \eta_i\,, \quad i = 1, 2, \ldots, N. \end{aligned} \tag{14}$$

In the remainder of this section, the solution of the problem in (14) is investigated. To that end, $tr\{\boldsymbol{I}(\boldsymbol{\theta})^{-1}\}$ for parameter $\boldsymbol{\theta}$ is derived for the system model specified by (2) and the error covariance matrix in (11) is employed. Note that for a fixed $\boldsymbol{f}$ and channel matrix $\boldsymbol{H}_r$, $\boldsymbol{Y}$ is a Gaussian random vector with mean $\mu(\boldsymbol{\theta})$ expressed as

$$\mu(\boldsymbol{\theta}) = \boldsymbol{H}_r \boldsymbol{\beta} = \begin{bmatrix} h_{r,1}\, f_1(\theta_1) \\ h_{r,2}\, f_2(\theta_2) \\ \vdots \\ h_{r,N}\, f_N(\theta_N) \end{bmatrix} \tag{15}$$

and covariance matrix $\boldsymbol{\Sigma}_r$. Accordingly, each element of $\boldsymbol{I}(\boldsymbol{\theta})$ can explicitly be written as [27]

$$[\boldsymbol{I}(\boldsymbol{\theta})]_{i,j} = [\boldsymbol{\Sigma}_r^{-1}]_{i,j} \left(h_{r,i} \frac{df_i(\theta_i)}{d\theta_i}\right) \left(h_{r,j} \frac{df_j(\theta_j)}{d\theta_j}\right), \tag{16}$$

where $[\boldsymbol{\Sigma}_r^{-1}]_{i,j}$ denotes the (i,j)th element of $\boldsymbol{\Sigma}_r^{-1}$. Note that if $\alpha_i \triangleq h_{r,i} \frac{df_i(\theta_i)}{d\theta_i}$, then $[\boldsymbol{I}(\boldsymbol{\theta})]_{i,j} = \alpha_i \alpha_j [\boldsymbol{\Sigma}_r^{-1}]_{i,j}$; thus, the FIM can simply be expressed as $\boldsymbol{I}(\boldsymbol{\theta}) = \mathrm{diag}\{\alpha_1, \alpha_2, \ldots, \alpha_N\} \boldsymbol{\Sigma}_r^{-1} \mathrm{diag}\{\alpha_1, \alpha_2, \ldots, \alpha_N\}$. Therefore, the following expression is obtained:

$$tr\{\boldsymbol{I}(\boldsymbol{\theta})^{-1}\} = \sum_{i=1}^{N} \frac{\sigma_{r,i}^2}{\alpha_i^2} = \sum_{i=1}^{N} \frac{\sigma_{r,i}^2}{{h_{r,i}}^2 f_i'(\theta_i)^2} \tag{17}$$

where $f_i'(\theta_i)$ denotes the derivative of $f_i(\theta_i)$. Note that (17) implies that even though the effective noise is not necessarily white, $tr\{\boldsymbol{I}(\boldsymbol{\theta})^{-1}\}$ can still be written as the sum of individual scalar inverse Fisher information corresponding to different parameters. Then, the cost function in (14) becomes

$$\begin{aligned}&\int_{a_1}^{b_1} \int_{a_2}^{b_2} \cdots \int_{a_N}^{b_N} w(\boldsymbol{\theta}) \sum_{i=1}^{N} \frac{\sigma_{r,i}^2}{{h_{r,i}}^2 f_i'(\theta_i)^2} d\theta_1 d\theta_2 \ldots d\theta_N \\ &= \sum_{i=1}^{N} \int_{a_1}^{b_1} \int_{a_2}^{b_2} \cdots \int_{a_N}^{b_N} w(\boldsymbol{\theta}) \frac{\sigma_{r,i}^2}{{h_{r,i}}^2 f_i'(\theta_i)^2} d\theta_1 d\theta_2 \ldots d\theta_N \\ &= \sum_{i=1}^{N} \frac{\sigma_{r,i}^2}{{h_{r,i}}^2} \int_{a_i}^{b_i} w_i(\theta_i) \frac{1}{f_i'(\theta_i)^2} d\theta_i. \end{aligned} \tag{18}$$

It is observed that the overall cost function is actually the sum of individual ECRB values for any generic $w(\boldsymbol{\theta})$. Based on (11) and (18), one can calculate the cost function and the constraints

in (14) for any given $w(\boldsymbol{\theta})$, $\boldsymbol{\beta}$ and channel statistics. In the following, two specific scenarios are investigated in more detail.

A. Independent Parameters & White Gaussian Noise for Eavesdropper

We first consider the scenario in which the channel noise is zero-mean white Gaussian for the eavesdropper,[2] that is, $\boldsymbol{\Sigma}_e = \text{diag}\{\sigma_{e,1}^2, \sigma_{e,2}^2, \ldots, \sigma_{e,N}^2\}$ and the parameters, θ_i's, are independent of each other with marginal distributions denoted by $w_i(\theta_i)$ for $i = 1, 2, \ldots, N$. (Note that $w(\boldsymbol{\theta}) = \prod_{i=1}^{N} w_i(\theta_i)$ in this scenario.) Under this setting, the following proposition reveals that the optimization problem be decoupled into independent scalar problems.

Proposition 1: If the parameters are independent and the channel noise for the eavesdropper is white Gaussian, the optimization problem in (14) can be decoupled into independent problems as follows:

$$f_{i,opt} = \arg\min_{f_i} \int_{a_i}^{b_i} w_i(\theta_i) \frac{1}{f_i'(\theta_i)^2} d\theta_i \quad \text{s.t. } \boldsymbol{\Sigma}_{err}(i) \geq \eta_i\,, \quad i = 1, 2, \ldots, N. \tag{19}$$

where

$$\boldsymbol{\Sigma}_{err}(i) = \frac{h_i^2 V_i (V_i - 2C_i)}{h_i^2 V_i + 1} + Var(\theta_i) + \big(E(f_i(\theta_i)) - E(\theta_i)\big)^2 \tag{20}$$

$V_i = Var(f_i(\theta_i))$, $C_i = Cov(f_i(\theta_i), \theta_i)$ and $h_i = h_{e,i}/\sigma_{e,i}$.

Proof: First, we focus on the error covariance matrix $\boldsymbol{\Sigma}_{err}$. Note that $\boldsymbol{\Sigma}_{\boldsymbol{\beta}} = \text{diag}\{V_1, V_2, \ldots, V_N\}$ with $V_i = Var(f_i(\theta_i))$, $\boldsymbol{\Sigma}_{\boldsymbol{\beta},\boldsymbol{\theta}} = \text{diag}\{C_1, C_2, \ldots, C_N\}$ with $C_i = Cov(f_i(\theta_i), \theta_i)$ and $\boldsymbol{\Sigma}_{\boldsymbol{\theta}} = \text{diag}\{Var(\theta_1), Var(\theta_2), \ldots, Var(\theta_N)\}$ due to the independence of θ_i's. Also,

$$\mathbf{R} = \text{diag}\left\{ \frac{{h_{e,1}}^2}{{h_{e,1}}^2 V_1 + \sigma_{e,1}^2}, \frac{{h_{e,2}}^2}{{h_{e,2}}^2 V_2 + \sigma_{e,2}^2}, \ldots, \frac{{h_{e,N}}^2}{{h_{e,N}}^2 V_N + \sigma_{e,N}^2} \right\}$$

due to the independence of θ_i's and the white Gaussian noise assumption for the eavesdropper. Therefore $\boldsymbol{\Sigma}_{err} = \text{diag}\{\boldsymbol{\Sigma}_{err}(1), \boldsymbol{\Sigma}_{err}(2), \ldots, \boldsymbol{\Sigma}_{err}(N)\}$, where

$$\boldsymbol{\Sigma}_{err}(i) = \frac{h_i^2 V_i (V_i - 2C_i)}{h_i^2 V_i + 1} + Var(\theta_i) + \big(E(f_i(\theta_i)) - E(\theta_i)\big)^2 \tag{21}$$

and $h_i = h_{e,i}/\sigma_{e,i}$. Based on (18) and (21), the generic optimization problem in (14) reduces to

$$\boldsymbol{f}_{opt} = \arg \min_{f_1, f_2, \ldots, f_N} \sum_{i=1}^{N} \frac{\sigma_{r,i}^2}{{h_{r,i}}^2} \int_{a_i}^{b_i} w_i(\theta_i) \frac{1}{f_i'(\theta_i)^2} d\theta_i \quad \text{s.t. } \boldsymbol{\Sigma}_{err}(i) \geq \eta_i\,, \quad i = 1, 2, \ldots, N. \tag{22}$$

[2]Note that there is no further assumption on the noise statistics for the intended receiver, as it does not effect the constraint and the cost function according to (11) and (17).

Note that the constraints are independent of each other and each element of the sum in the objective function has no effect on the others. Therefore, the optimization problem can be decoupled and each θ_i can be optimized individually, where the decoupled problems can be expressed as in (19). ■

Remark 2: The optimization problem in (19) has been investigated in [17] in detail and the results and the solution methods proposed in that study can directly be applied to the vector parameter problem, when the channel noise for the eavesdropper is white Gaussian and the parameters are independent of each other. Also, when the parameters are not independent, the constraints given in (14) include cross terms even if the eavesdropper has white Gaussian noise; therefore, the optimization problem needs to be solved based on (14) for correlated parameters.

B. Independent Parameters & Colored Gaussian Noise Vectors

In this part, we again assume that the parameters are independent of each other, i.e., $w(\boldsymbol{\theta}) = \prod_{i=1}^{N} w_i(\theta_i)$; however, we suppose that $\boldsymbol{\Sigma}_e$ is a symmetric, positive definite matrix which is not necessarily diagonal. Due to the independence of parameters, $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}$, $\boldsymbol{\Sigma}_{\boldsymbol{\beta},\boldsymbol{\theta}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ take diagonal forms as in Section III-A. Then, the ith diagonal element $\boldsymbol{\Sigma}_{err}(i)$ of $\boldsymbol{\Sigma}_{err}$ can be written as

$$\boldsymbol{\Sigma}_{err}(i) = h_{e,i}^2 V_i (V_i - 2C_i)\gamma_i + Var(\theta_i) + \big(E(f_i(\theta_i)) - E(\theta_i)\big)^2 \tag{23}$$

where V_i and C_i are as defined previously. Also, γ_i is the ith diagonal element of matrix $(\tilde{\mathbf{D}} + \boldsymbol{\Sigma}_e)^{-1}$, where $\tilde{\mathbf{D}} = \text{diag}\{h_{e,1}^2 V_1, h_{e,2}^2 V_2, \ldots, h_{e,N}^2 V_N\}$. Note that γ_i depends on $\boldsymbol{H}_e$ and the encoding function $\boldsymbol{f}$. Due to the cross terms in the constraints, the optimization problem cannot be decoupled anymore, hence it should be solved using (14) based on (17) and (23). However, it is possible to derive some theoretical results about the form of the solution in the considered scenario. Lemma 1 generalizes Proposition 3 in [17] for the multivariable case.

Lemma 1: Suppose that the eavesdropper employs the linear MMSE estimator and $w_i(\theta_i)$ is symmetric around $(a_i + b_i)/2$. Then, for any given encoding function $\boldsymbol{f}(\boldsymbol{\theta})$ which consists of continuous and strictly increasing encoding functions $f_i(\theta_i)$, there exists a corresponding encoding function $\boldsymbol{s}(\boldsymbol{\theta})$ consisting of continuous and strictly decreasing encoding functions $s_i(\theta_i)$ that yields the same ECRB at the intended receiver with a higher MSE for the individual parameters at the eavesdropper.

Proof: By using the arguments in [17], we consider two encoding functions $f_i(\theta_i)$ and $s_i(\theta_i) = f_i(a_i + b_i - \theta_i)$, where $\theta_i \in [a_i, b_i]$ and $f_i(\theta_i)$ is a continuous and monotonically increasing function. Since $s_i'(\theta_i) = -f_i'(a_i + b_i - \theta_i)$ by definition and due to the symmetry in $w_i(\theta_i)$, both encoding functions result in the same $tr\{\boldsymbol{I}(\boldsymbol{\theta})^{-1}\}$, which is given in (17). Furthermore, as shown in [17], $Cov(f_i(\theta_i), \theta_i) > Cov(s_i(\theta_i), \theta_i)$ and two encoders yield the same variance and expectation for the encoded version of the parameter. Also, $\boldsymbol{\Sigma}_e$ is a positive definite matrix and $\tilde{\mathbf{D}}$ has positive entries. Therefore, $(\tilde{\mathbf{D}} + \boldsymbol{\Sigma}_e)^{-1}$ is also a positive definite matrix[3] and $\gamma_i > 0$ always holds. Combining

[3]Since $\boldsymbol{\Sigma}_e$ is a positive definite symmetric matrix, it can be expressed as $\boldsymbol{\Sigma}_e = \sum_{k=1}^{N} \lambda_k \mathbf{v}_k \mathbf{v}_k^T$ and since $\tilde{\mathbf{D}}$ is diagonal, $(\tilde{\mathbf{D}} + \boldsymbol{\Sigma}_e)^{-1} = \sum_{k=1}^{N} \frac{1}{\lambda_k + h_{e,k}^2 V_k} \mathbf{v}_k \mathbf{v}_k^T$ can be obtained.

these results and via (23), it is obtained that a larger MSE for parameter θ_i, i.e., $\mathbf{\Sigma}_{err}(i)$, can be achieved by employing $s_i(\theta_i)$ instead of $f_i(\theta_i)$ while keeping the ECRB the same. ■

Lemma 1 has an important practical implication that the search space for the optimal encoding function for the ith parameter can be restricted to strictly decreasing functions when the sufficient condition given in the lemma is satisfied. Note that Lemma 1 can be applied if θ_i has a symmetric distribution on its domain. Some examples of continuous symmetric distributions on a bounded interval satisfying the condition include uniform distribution, beta distribution with both parameters of $1/2$, and raised cosine distribution.

1) Two-Parameter Case ($N = 2$): In this part, we investigate the case of $N = 2$; that is, $\boldsymbol{\theta} = [\theta_1, \theta_2]^T$. Therefore, the channel noise $\boldsymbol{N}_e$ for the eavesdropper can be modeled as zero-mean Gaussian with covariance matrix $\mathbf{\Sigma}_e = \left[\begin{smallmatrix} \sigma_{e,1}^2 & \rho \\ \rho & \sigma_{e,2}^2 \end{smallmatrix}\right]$. For this particular case, γ_i in (23) can explicitly be written as

$$\gamma_1 = \frac{h_{e,2}^2 V_2 + \sigma_{e,2}^2}{(h_{e,1}^2 V_1 + \sigma_{e,1}^2)(h_{e,2}^2 V_2 + \sigma_{e,2}^2) - \rho^2} \tag{24}$$

and γ_2 can be obtained by replacing the numerator in (24) with $h_{e,1}^2 V_1 + \sigma_{e,1}^2$. After some manipulation, $\mathbf{\Sigma}_{err}(1)$ can be derived as

$$\begin{aligned}\mathbf{\Sigma}_{err}(1) = \lambda\ & E(|\beta_1 - \theta_1|^2) \\ & + (1-\lambda)\left((E(\beta_1) - E(\theta_1))^2 + Var(\theta_1)\right)\end{aligned} \tag{25}$$

where

$$\lambda = \frac{h_1^2 V_1}{h_1^2 V_1 + 1 - r_2(\rho)}$$

with

$$r_2(\rho) = \frac{\rho^2/\sigma_{e,1}^2}{h_{e,2}^2 V_2 + \sigma_{e,2}^2}$$

and $h_1 = h_{e,1}/\sigma_{e,1}$.

It is possible to gain practical intuition about the behavior of the optimal encoding function as a closed-form expression for $\mathbf{\Sigma}_{err}(1)$ (and $\mathbf{\Sigma}_{err}(2)$) is available. There are several important observations related to (25).

- For a fixed $r_2(\rho)$, if we let $h_1^2 \to \infty$, then $\mathbf{\Sigma}_{err}(1) \approx E(|\beta_1 - \theta_1|^2)$; hence, it is maximized when $E(|\beta_1 - \theta_1|^2)$ is maximized. This mode can be called as the *variance maximizing mode* as in [17]. If we let $h_1^2 \to 0$, then $\mathbf{\Sigma}_{err}(1) \approx (E(\beta_1) - E(\theta_1))^2 + Var(\theta_1)$; therefore, it is maximized if $\beta_1 \to a_1$ or $\beta_1 \to b_1$. This mode can be called as the *variance minimizing mode* [17].
- For a fixed h_1 (and relevant parameters for θ_2), as ρ^2 increases, $r_2(\rho)$ and λ also increase. According to (25), if λ is small enough, the encoder is in the variance minimizing mode; however, as λ increases and becomes large enough, maximizing $E(|\beta_1 - \theta_1|^2)$ becomes the priority. As ρ increases, after a certain threshold, which can be denoted as ρ_0, the mode of operation can change and the encoder can get into the variance maximizing mode when $\rho > \rho_0$.

Note that in the analysis above h_1^2 can be viewed as the signal-to-noise ratio (SNR) for the channel of θ_1 to the eavesdropper. As the SNR of this channel increases, the distortion due to encoding is transmitted to the eavesdropper more effectively and the main factor to create a large MSE at the eavesdropper is the distortion to the parameter via encoding in the variance maximizing mode. Also, when $h_1 \to 0$, this means that the channel is very noisy; hence, the only information available to the eavesdropper through its observation is the mean of the encoded version of the parameter. Therefore, the encoder tries to ensure that the mean of the encoded version is away from the true mean. Note that in practice, even if the SNR values are not necessarily in absolute limits, we can still observe the aforementioned behavior in the encoding functions (see Figs. 3 and 5). Hence, it can be concluded that the form of encoding function depends on the parameters of the channel and the correlation between eavesdropper's noise components. Finally, we note that a similar derivation and analysis can be performed for $\mathbf{\Sigma}_{err}(2)$ based on γ_2 and (23).

IV. Affine Joint Encoding Strategy

In this section, the encoding operation is assumed to be an affine function. Namely, the vector parameter $\boldsymbol{\theta}$ is encoded by using an $N \times N$ precoding matrix $\boldsymbol{P}$ and an N-dimensional constant vector $\boldsymbol{r}$ prior to transmission such that $\boldsymbol{\beta} = \boldsymbol{P}\boldsymbol{\theta} + \boldsymbol{r}$. Under this assumption, the optimal parameter encoding problem can be expressed as follows:

$$\begin{aligned}[\boldsymbol{P}_{opt}, \boldsymbol{r}_{opt}] = \arg\min_{\boldsymbol{P},\boldsymbol{r}} & \int_{\Lambda} w(\boldsymbol{\theta})\ tr\{\boldsymbol{I}(\boldsymbol{\theta})^{-1}\} d\boldsymbol{\theta} \\ \text{s.t.} \quad & \mathbf{\Sigma}_{err}(i) \geq \eta_i\,, \quad i = 1, 2, \ldots, N.\end{aligned} \tag{26}$$

As in the previous section, the parameter space is specified as $\theta_i \in [a_i, b_i]$, for $i = 1, 2, \ldots, N$ for this strategy. If we define $a \triangleq \min\{a_1, a_2, \ldots a_N\}$ and $b \triangleq \max\{b_1, b_2, \ldots b_N\}$, then $\theta_i \in [a, b]$, for $i = 1, 2, \ldots, N$. In this section, it is assumed that the generalized domain of the parameters, i.e., $[a, b]$, needs to be preserved after the encoding operation; hence, it is assumed that $\beta_i \in [a, b]$, for $i = 1, 2, \ldots, N$. This condition can be guaranteed if the sum of the absolute values of the elements in each row of $\boldsymbol{P}$ is less than or equal to 1. This can formally be expressed as $\|\boldsymbol{P}^T \boldsymbol{e}_j\|_1 \leq 1$ for $j = 1, 2, \ldots, N$, where $\boldsymbol{e}_j$'s are standard basis vectors.[4] Finally, the precoding matrix $\boldsymbol{P}$ is taken to be full rank (invertible).

In the remainder of this section, the solution of the problem in (26) is investigated. First, $tr\{\boldsymbol{I}(\boldsymbol{\theta})^{-1}\}$ for parameter $\boldsymbol{\theta}$ is derived for the given system model and encoding strategy. Note that $\boldsymbol{Y}$ is a Gaussian random vector with mean $\mu(\boldsymbol{\theta}) = \boldsymbol{H}_r\boldsymbol{\beta} = \boldsymbol{H}_r\boldsymbol{P}\boldsymbol{\theta} + \boldsymbol{H}_r\boldsymbol{r}$ and covariance matrix $\mathbf{\Sigma}_r$ for fixed $\boldsymbol{P}$, $\boldsymbol{r}$ and channel matrix $\boldsymbol{H}_r$. Therefore, each element of $\boldsymbol{I}(\boldsymbol{\theta})$ can explicitly be written as

$$\begin{aligned}[\boldsymbol{I}(\boldsymbol{\theta})]_{i,j} &= \left(\frac{d\boldsymbol{\mu}(\boldsymbol{\theta})}{d\theta_i}\right)^T \mathbf{\Sigma}_r^{-1} \left(\frac{d\boldsymbol{\mu}(\boldsymbol{\theta})}{d\theta_i}\right) \\ &= \boldsymbol{p}_i^T \boldsymbol{H}_r \mathbf{\Sigma}_r^{-1} \boldsymbol{H}_r \boldsymbol{p}_j\end{aligned} \tag{27}$$

[4] $\|\mathbf{x}\|_1 \triangleq \sum_{i=1}^{N} |x_i|$ is called the l_1 norm of vector $\mathbf{x}$.

where $\boldsymbol{p}_i$ denotes the ith column of precoding matrix $\boldsymbol{P}$. Accordingly, the FIM can be expressed as

$$\begin{aligned}\boldsymbol{I}(\boldsymbol{\theta}) &= \boldsymbol{P}^T\boldsymbol{H}_r\boldsymbol{\Sigma}_r^{-1}\boldsymbol{H}_r\boldsymbol{P} \\ &= \boldsymbol{P}^T\boldsymbol{D}\boldsymbol{P} \end{aligned} \tag{28}$$

where $\boldsymbol{D} \triangleq \boldsymbol{H}_r\boldsymbol{\Sigma}_r^{-1}\boldsymbol{H}_r$. Note that $\boldsymbol{D}$ and $I(\boldsymbol{\theta})$ are positive definite, invertible and symmetric matrices. Also, $\boldsymbol{I}(\boldsymbol{\theta})$ is not a function of $\boldsymbol{\theta}$. Therefore, the objective function in (26) simplifies to

$$\int_{\Lambda} w(\boldsymbol{\theta})\ tr\{\boldsymbol{I}(\boldsymbol{\theta})^{-1}\}d\boldsymbol{\theta} = tr\left\{\left(\boldsymbol{P}^T\boldsymbol{D}\boldsymbol{P}\right)^{-1}\right\}. \tag{29}$$

Note that the objective function depends only on $\boldsymbol{P}$ and the constant factor $\boldsymbol{r}$ in the encoding operation does not effect its value. Furthermore, if the zero-mean Gaussian random noise $\boldsymbol{N}_r$ in the received signal has independent components, then $\boldsymbol{D}$ becomes a diagonal matrix with its ith diagonal element being given by $h_{r,i}^2/\sigma_{r,i}^2$, where $\sigma_{r,i}^2$ is the variance of the ith noise component in $\boldsymbol{N}_r$.

The following proposition provides an optimal solution to the affine joint encoding problem without any secrecy constraints for a diagonal $\boldsymbol{D}$.

Proposition 2: Assume $\boldsymbol{D}$ is a diagonal matrix. In the absence of secrecy constraints on the eavesdropper, any signed permutation matrix[5] is an optimal solution. Furthermore, any other precoding matrix with a different form is not optimal.

Proof: In the absence of secrecy constraints, the optimization problem can be formulated as

$$\begin{aligned}\boldsymbol{P}_{opt} = \ &\arg\min_{\boldsymbol{P}} tr\left\{\left(\boldsymbol{P}^T\boldsymbol{D}\boldsymbol{P}\right)^{-1}\right\} \\ &\text{s.t. } \|\boldsymbol{P}^T\boldsymbol{e}_j\|_1 \le 1\,, \quad j = 1, 2, \ldots, N. \end{aligned} \tag{30}$$

Then, a lower bound for any given feasible $\boldsymbol{P}$ can be obtained as follows:

$$\begin{aligned} tr\left\{\left(\boldsymbol{P}^T\boldsymbol{D}\boldsymbol{P}\right)^{-1}\right\} &= tr\left\{\left(\boldsymbol{P}^{-1}\boldsymbol{D}^{-1}\boldsymbol{P}^{-T}\right)\right\} \\ &= \left\|\boldsymbol{P}^{-1}\boldsymbol{D}^{-1/2}\right\|_F^2 \\ &= \sum_{j=1}^{N}\frac{1}{\lambda_j}\|\boldsymbol{m}_j\|_2^2 \end{aligned} \tag{31}$$

where $\boldsymbol{M} \triangleq \boldsymbol{P}^{-1}$, $\boldsymbol{m}_j$ is the jth column of $\boldsymbol{M}$ and $\boldsymbol{D} = diag\{\lambda_1, \lambda_2, \ldots, \lambda_N\}$. Note that $\boldsymbol{P}\boldsymbol{M} = \mathbf{I}$, thus $\boldsymbol{p}_j^{(r)}\boldsymbol{m}_j = 1$ for $j = 1, 2, \ldots N$, and $\boldsymbol{p}_j^{(r)} = \boldsymbol{e}_j^T\boldsymbol{P}$ is the jth row of $\boldsymbol{P}$. As the sum of the absolute values of the elements in each row cannot be greater than 1, $\|\boldsymbol{p}_j^{(r)}\|_2 \le \|\boldsymbol{p}_j^{(r)}\|_1 \le 1$. Also, via Cauchy-Schwarz inequality, it can be obtained that $1 = |\boldsymbol{p}_j^{(r)}\boldsymbol{m}_j|^2 \le \|\boldsymbol{p}_j^{(r)}\|_2^2\|\boldsymbol{m}_j\|_2^2$; hence, as $\|\boldsymbol{p}_j^{(r)}\|_2 \le 1$, $\|\boldsymbol{m}_j\|_2 \ge 1$ for $j = 1, 2, \ldots, N$. Therefore,

$$tr\left\{\left(\boldsymbol{P}^T\boldsymbol{D}\boldsymbol{P}\right)^{-1}\right\} = \sum_{j=1}^{N}\frac{1}{\lambda_j}\|\boldsymbol{m}_j\|_2^2 \ge \sum_{j=1}^{N}\frac{1}{\lambda_j} \tag{32}$$

[5]A signed permutation matrix is defined as a matrix whose every row and column has exactly one non-zero entry, which can be either 1 or -1.

for any given feasible $\boldsymbol{P}$. Note that this lower bound can exactly be attained when $\|\boldsymbol{m}_j\|_2 = 1$, which implies $\|\boldsymbol{p}_j^{(r)}\|_2 = 1$ for an optimal solution. Also, due to the relation $1 = \|\boldsymbol{p}_j^{(r)}\|_2 \le \|\boldsymbol{p}_j^{(r)}\|_1 \le 1$ for $j = 1, 2, \ldots, N$, $\|\boldsymbol{p}_j^{(r)}\|_2 = \|\boldsymbol{p}_j^{(r)}\|_1 = 1$. This is satisfied if and only if $\boldsymbol{p}_j^{(r)}$ contains an element with a value of $+1$ or -1 and the rest of its elements are zero. Due to the rank constraint, each $\boldsymbol{p}_j^{(r)}$ should have the non-zero element at a different location and this is satisfied if and only if the precoding matrix is a signed permutation matrix. ■

Proposition 2 reveals that if there is no secrecy constraint for a given diagonal $\boldsymbol{D}$, then a signed permutation matrix can be used as the optimal precoding matrix.

Next, the optimal affine joint encoding problem is considered in the presence of secrecy constraints. The error covariance matrix $\boldsymbol{\Sigma}_{err}$ in the constraint of (26) can be calculated based on the procedure in Remark 1. Specifically, it can be obtained by using the equations given in (11) and (12) and inserting $\boldsymbol{\Sigma}_{\boldsymbol{\beta}} = \boldsymbol{P}\boldsymbol{\Sigma}_{\boldsymbol{\theta}}\boldsymbol{P}^T$ and $\boldsymbol{\Sigma}_{\boldsymbol{\beta},\boldsymbol{\theta}} = \boldsymbol{P}\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$. Note that only the last term in (11) depends on $\boldsymbol{r}$. As only the diagonal terms are taken into consideration for the secrecy targets, they can explicitly be calculated. The following lemma is provided regarding the relationship between $\boldsymbol{\Sigma}_{err}$ and $\boldsymbol{r}$ for any given $\boldsymbol{P}$ and $w(\boldsymbol{\theta})$.

Lemma 2: When the eavesdropper employs the linear MMSE estimator, then $\boldsymbol{\Sigma}_{err}(i)$, (i.e., the ith diagonal element of $\boldsymbol{\Sigma}_{err}$) for the encoding operation $\boldsymbol{\beta} = \boldsymbol{P}\boldsymbol{\theta} + \boldsymbol{r}$ is a convex function of r_i, i.e., the ith element of $\boldsymbol{r}$ for a fixed $\boldsymbol{P}$.

Proof: Consider the expression for $\boldsymbol{\Sigma}_{err}$ in Remark 1 (see (11) and (12)). It is noted that only the last term in (11) depends on $\boldsymbol{r}$, which can be written as

$$\begin{aligned}&\left(\left(E(\boldsymbol{\beta}) - E(\boldsymbol{\theta})\right)\left(E(\boldsymbol{\beta}) - E(\boldsymbol{\theta})\right)^T\right) \\ &= (\boldsymbol{P} - \mathbf{I})\,E(\boldsymbol{\theta})E(\boldsymbol{\theta})^T\,(\boldsymbol{P} - \mathbf{I})^T + \boldsymbol{r}E(\boldsymbol{\theta})^T\,(\boldsymbol{P} - \mathbf{I})^T \\ &\quad + (\boldsymbol{P} - \mathbf{I})\,E(\boldsymbol{\theta})\boldsymbol{r}^T + \boldsymbol{r}\boldsymbol{r}^T\,. \end{aligned} \tag{33}$$

For a given $\boldsymbol{P}$, the contribution of (33) (i.e., the last term of $\boldsymbol{\Sigma}_{err}$) to $\boldsymbol{\Sigma}_{err}(i)$, denoted as $g(i)$, can be calculated as

$$g(i) = \left(r_i + \boldsymbol{p}_i^{(r)}E(\boldsymbol{\theta}) - E(\theta_i)\right)^2, \tag{34}$$

where $\boldsymbol{p}_i^{(r)}$ is the ith row of $\boldsymbol{P}$. As the other terms of $\boldsymbol{\Sigma}_{err}$ does not depend on $\boldsymbol{r}$ (see (11)) and $\frac{d^2g(i)}{dr_i^2} = 2 > 0$, the convexity claim in the lemma holds. ■

As a result of Lemma 2, $\boldsymbol{\Sigma}_{err}(i)$ is maximized either at $r_i^{\min}$ or $r_i^{\max}$, where $r_i^{\min}$ and $r_i^{\max}$ are, respectively, the lowest and highest possible values of r_i for a given $\boldsymbol{P}$, while ensuring that the ith element of $\boldsymbol{P}\boldsymbol{\theta} + \boldsymbol{r}$, i.e., β_i, is in $[a, b]$. For example, if $\theta_1, \theta_2 \in [0, 1]$ and $\boldsymbol{P} = \left[\begin{smallmatrix} 0.1 & 0.5 \\ 0 & -0.8 \end{smallmatrix}\right]$, then $0 \le r_1 \le 0.4$ and $0.8 \le r_2 \le 1$ to ensure $\beta_1, \beta_2 \in [0, 1]$. Therefore, $r_1^{\min} = 0, r_1^{\max} = 1$, $r_2^{\min} = 0.8$ and $r_2^{\max} = 1$ for this particular example. Among $r_i^{\min}$ or $r_i^{\max}$, the one that yields a higher $\boldsymbol{\Sigma}_{err}(i)$ can be selected. As the objective function in (26) does not depend on $\boldsymbol{r}$, it can freely be selected to maximize $\boldsymbol{\Sigma}_{err}(i)$ for a given $\boldsymbol{P}$; therefore, it is sufficient to search over precoding matrices for the optimal strategy.

Corollary 1: Suppose that eavesdropper's noise has independent components, and $\beta_i = w_i\theta_j + r_i$ for some $i \neq j$. If either of $E(\theta_i)$ or $E(\theta_j)$ is equal to $\frac{a+b}{2}$, then, the sign of w_i does not effect $\mathbf{\Sigma}_{err}(i)$.

Proof: We prove the statement for the case of $E(\theta_i) = (a+b)/2$, as it can be shown for $E(\theta_j) = (a+b)/2$ in a similar fashion. First, we note that $\mathbf{\Sigma}_{err} = \mathbf{\Sigma}^{(1)}_{err} + \mathbf{\Sigma}^{(2)}_{err}$ such that $\mathbf{\Sigma}^{(1)}_{err}$ represents the first four terms of the sum in (11) and $\mathbf{\Sigma}^{(2)}_{err}$ denotes the last term. Under the condition in the corollary, w_i's appear in the form of w_i^2's in the diagonals of $\mathbf{\Sigma}^{(1)}_{err}$. Therefore, the sign of w_i does not have any effect on $\mathbf{\Sigma}^{(1)}_{err}$. For $\mathbf{\Sigma}^{(2)}_{err}$, if $\beta_i = w_i\theta_j + r_i$, then we know that $\mathbf{\Sigma}^{(2)}_{err}(i) = (r_i + w_i E(\theta_j) - E(\theta_i))^2$. As $\mathbf{\Sigma}^{(2)}_{err}(i)$ is maximized either at $r_i^{\min}$ or $r_i^{\max}$ due to Lemma 2, we have

$$\mathbf{\Sigma}^{(2)}_{err}(i) = \max\left\{\left(\frac{b-a}{2} + \alpha(E(\theta_j) - b)\right)^2, \left(\frac{a-b}{2} + \alpha(E(\theta_j) - a)\right)^2\right\}$$

for $w_i = \alpha > 0$ and

$$\mathbf{\Sigma}^{(2)}_{err}(i) = \max\left\{\left(\frac{b-a}{2} - \alpha(E(\theta_j) - a)\right)^2, \left(\frac{a-b}{2} - \alpha(E(\theta_j) - b)\right)^2\right\}$$

for $w_i = -\alpha < 0$. Note that the $\mathbf{\Sigma}^{(2)}_{err}(i)$ expressions are exactly the same for both sign options for w_i as long as $|w_i|$ does not change. Therefore, $\mathbf{\Sigma}_{err}(i)$ does not depend on the sign of w_i. ■

Lemma 3: Suppose the encoding matrix $\boldsymbol{P}$ has the form of $\boldsymbol{P} = \boldsymbol{W}_1\boldsymbol{W}_2$, where $\boldsymbol{W}_1 = diag\{w_1, w_2, \ldots w_N\}$ is a diagonal matrix and $\boldsymbol{W}_2$ is a permutation matrix. Then, $tr\{(\boldsymbol{P}^T\boldsymbol{D}\boldsymbol{P})^{-1}\}$ does not depend on the signs of the elements in $\boldsymbol{P}$.

Proof: Note that if $\boldsymbol{P} = \boldsymbol{W}_1\boldsymbol{W}_2$, then

$$\begin{aligned} tr\left\{(\boldsymbol{P}^T\boldsymbol{D}\boldsymbol{P})^{-1}\right\} &= tr\left\{(\boldsymbol{W}_2^T\boldsymbol{W}_1\boldsymbol{D}\boldsymbol{W}_1\boldsymbol{W}_2)^{-1}\right\} \\ &= tr\{\boldsymbol{W}_2^T\hat{\boldsymbol{W}}_1\boldsymbol{D}^{-1}\hat{\boldsymbol{W}}_1\boldsymbol{W}_2\} \\ &= tr\{\boldsymbol{W}_2\boldsymbol{W}_2^T\hat{\boldsymbol{W}}_1\boldsymbol{D}^{-1}\hat{\boldsymbol{W}}_1\} \\ &= tr\{\hat{\boldsymbol{W}}_1\boldsymbol{D}^{-1}\hat{\boldsymbol{W}}_1\} \\ &= \sum_{j=1}^{N}\frac{\hat{d}_j}{w_j^2} \end{aligned} \tag{35}$$

where $\hat{\boldsymbol{W}}_1 = \boldsymbol{W}_1^{-1} = diag\{1/w_1, 1/w_2, \ldots, 1/w_N\}$ and $\hat{d}_j$ is the jth diagonal element of $\boldsymbol{D}^{-1}$. As $tr\{(\boldsymbol{P}^T\boldsymbol{D}\boldsymbol{P})^{-1}\}$ is the sum of squares, the signs of w_i's do not effect its value. ■

Corollary 1 and Lemma 3 imply that if the encoder applies the method of simple shuffle and scale, then the sign of the scaling factor does not matter in terms of the cost and objective of the optimization. Therefore, optimal scaling factors can be assumed to be positive without loss of generality, which reduces the search space.

Remark 3: By Proposition 2, we know that when $\boldsymbol{D}$ is a diagonal matrix, permutation matrices (with $+1$ or -1 as nonzero elements) are optimal precoding matrices. Also, the optimal precoder belongs to this family of matrices up to a certain secrecy target level $\eta^\dagger$ for each parameter. In other words, if the secrecy target for a given parameter is larger than $\eta^\dagger$, then the objective will be larger and the optimal precoder will not be a permutation matrix anymore. The exact value of $\eta^\dagger$ can be found by solving the following optimization problem:

$$\eta^\dagger = \max_{\boldsymbol{P}\in\mathcal{P}} \min_i \mathbf{\Sigma}_{err}(i) \tag{36}$$

where $\mathcal{P}$ denotes the set of permutation matrices with $+1$ or -1 as non-zero elements and $\mathbf{\Sigma}_{err}$ is as given in (11). Note that there are $2^N N!$ elements in $\mathcal{P}$; therefore, as N gets larger, it gets challenging to solve the optimization problem in (36). However, for small N's, it can be solved and provides a practical limit for the secrecy level that can be satisfied without increasing the ECRB values of the case without any secrecy concerns.

V. Numerical Results

In this section, numerical results are provided for both strategies proposed in Section III and Section IV.

A. Nonlinear Individual Encoding

In all the numerical examples for the individual encoding strategy, $\boldsymbol{\theta}$ is modeled as $\boldsymbol{\theta} = [\theta_1\ \theta_2]^T$, where both θ_1 and θ_2 are uniformly distributed in [0, 1] and are independent of each other. The channel parameters for the intended receiver are taken to be $h_{r,1} = h_{r,2} = 2$ and $\sigma^2_{r,1} = \sigma^2_{r,2} = 1$. As the conditions in Lemma 1 are satisfied, the optimal encoding functions are searched among decreasing functions. For the first example, the eavesdropper fading coefficients are taken as $h_{e,2} = 1.5$ and $h_{e,1} \in \{1, 1.2\}$. The channel noise for the eavesdropper is modeled as zero-mean multivariate Gaussian random variable with the covariance matrix $\mathbf{\Sigma}_e = \left[\begin{smallmatrix}\sigma^2_{e,1} & \rho \\ \rho & \sigma^2_{e,2}\end{smallmatrix}\right]$, where $\sigma^2_{e,1} = \sigma^2_{e,2} = 1$. The target secrecy levels are $\eta_1 = \eta_2 = 0.15$. In order to solve the optimization problem in (14), the approximation methods described in [17] can be used. In this study, the piecewise linear approximation method is employed. Namely, for each $f_i(\theta_i)$, $\Delta x_k^{(i)} \triangleq f_i(a_i + k\Delta\theta_i) - f_i(a_i + (k-1)\Delta\theta_i)$ is defined, and the optimization is performed over MN variables; that is, the increments/decrements for each parameter ($\Delta x^{(i)} = [\Delta x_1^{(i)}, \Delta x_2^{(i)}, \ldots, \Delta x_M^{(i)}]$ for $i = 1, 2, \ldots, N$) are obtained. For the numerical results, M is taken to be 50 and Global Optimization Toolbox of MATLAB is used.

In Fig. 2, the total and individual ECRB values for θ_1 and θ_2 are plotted for various ρ values. It is observed that as ρ increases, the total and individual ECRB values decrease, which implies that the correlation between the noise components of the eavesdropper for each parameter is useful for our design purposes. Also, the ECRB for θ_1 decreases very slightly until a certain value of ρ_0 (i.e., $\rho_0 \approx 0.2$ and 0.6 for $h_{e,1} = 1.2$ and 1, respectively), and then a sharper decrease in the ECRB is

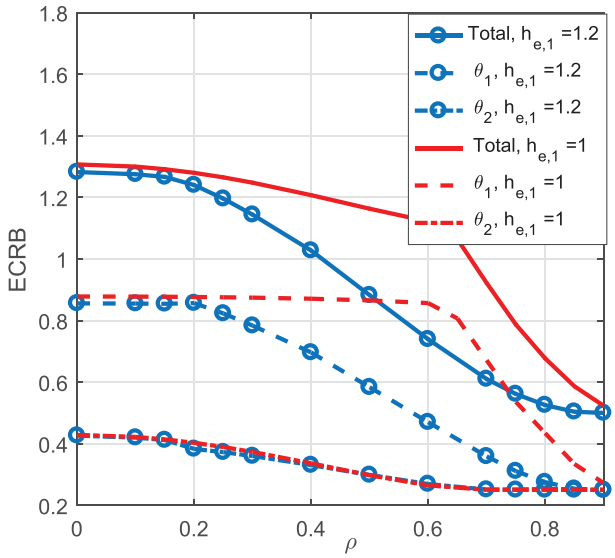

Fig. 2. Total and individual ECRB values versus ρ for $h_{e,1} = 1$ and $h_{e,1} = 1.2$.

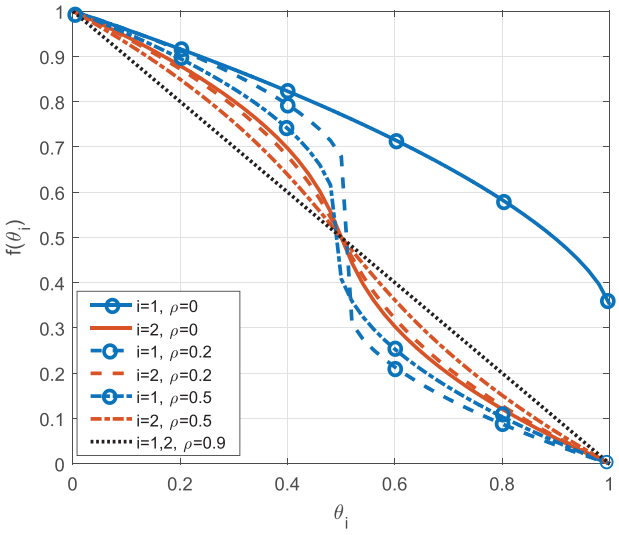

Fig. 3. The optimal encoding functions for θ_1 and θ_2 for $\rho = \{0, 0.2, 0.5, 0.9\}$ when $h_{e,1} = 1.2$.

observed. This is due to the fact that the encoding mode for θ_1 changes as explained in Section III-B1. Another interesting observation is that for $h_{e,1} = 1.2$, the total and individual ECRB for θ_1 is lower than that in the case of $h_{e,1} = 1$ and the ECRB for θ_2 stays almost the same. The reason for having a lower total ECRB for a larger $h_{e,1}$ is the fact that the eavesdropper is unaware of encoding; hence, the distortion due to the encoding function is transmitted more effectively to the eavesdropper. Also, for larger values of ρ, the ECRB values for both parameters converge to each other.

In Fig. 3, the optimal encoding functions for θ_1 and θ_2 are presented for $\rho \in \{0, 0.2, 0.5, 0.9\}$ when $h_{e,1} = 1.2$. This figure explains some of the behaviors observed in Fig. 2. For example, when $\rho = 0$, $f_1(\theta_1)$ is in the variance minimizing mode and

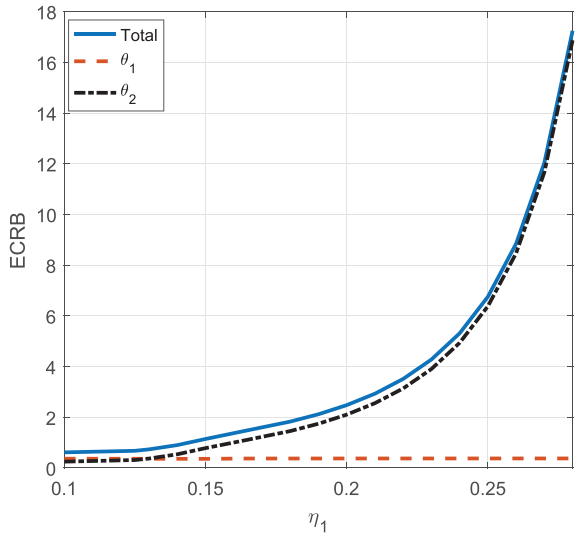

Fig. 4. Total and individual ECRB values versus η_1.

$f_2(\theta_2)$ is in the variance maximizing mode[6]. As ρ increases, the changes in $f_2(\theta_2)$ are not significant and there is no mode change. On the other hand, the characteristics of $f_1(\theta_1)$ change when ρ increases, and it gets into the variance maximizing mode for $\rho \in \{0.2, 0.5, 0.9\}$. Also, both encoding functions are linear, $f_i(\theta_i) = 1 - \theta_i$, for $\rho = 0.9$, yielding the same ECRB.

For the second example, $h_{e,1} = 1.2$, $h_{e,2} = 1.5$, $\sigma^2_{e,1} = \sigma^2_{e,2} = 1$ and $\rho = 0.3$. The target secrecy level for θ_2 is fixed to be $\eta_2 = 0.15$, and the target secrecy level for θ_1 is increased starting from 0.1. In Fig. 4, the total and individual ECRB values for θ_1 and θ_2 are plotted for various η_1 values. Note that the change in the secrecy target for θ_1 does not have any significant effect on the ECRB performance of θ_2. However, the ECRB for θ_1 and the total ECRB increase exponentially as η_1 increases. The reason of this can be deduced from Fig. 5. In Fig. 5, the optimal encoding functions for θ_1 and θ_2 are given for $\eta_1 \in \{0.1, 0.15, 0.2, 0.25\}$. It is observed that when $\eta_1 = 0.1$, $f_1(\theta_1) = 1 - \theta_1$. When $\eta_1 = 0.15$, $f_1(\theta_1)$ operates in the variance maximizing mode, and for $\eta_1 = 0.2$ and 0.25, it is in the variance minimizing mode. Note that as η_1 increases, $f_1(\theta_1)$ approaches to 1. (Note that as $f_1(\theta_1) \to 1$, the ECRB goes to ∞). Also, note that the encoding function for θ_2 is insensitive to changes in η_1; that is, $f_2(\theta_2)$ does not change even though η_1 increases, and it is the same for all values of η_1 in this example.

In order to demonstrate the advantages of the proposed encoding scheme, the solution based on [17] is selected as a benchmark scheme, and a direct performance comparison between the optimal solution based on NIE and the solution based on [17] is provided in Fig. 6. Note that the individual encoding functions are obtained independently for each element of the vector parameter in the benchmark scheme as [17] provides a solution method for scalar problems. In this scenario, the ECRB is plotted versus η_2 for the solution based on [17] and NIE when

[6]Practically, in the variance minimizing mode, the encoder effectively decreases the transmitted signal power to hide the parameter; and in the variance maximizing mode, it has a two-level *quantizer-like* behavior to ensure secrecy.

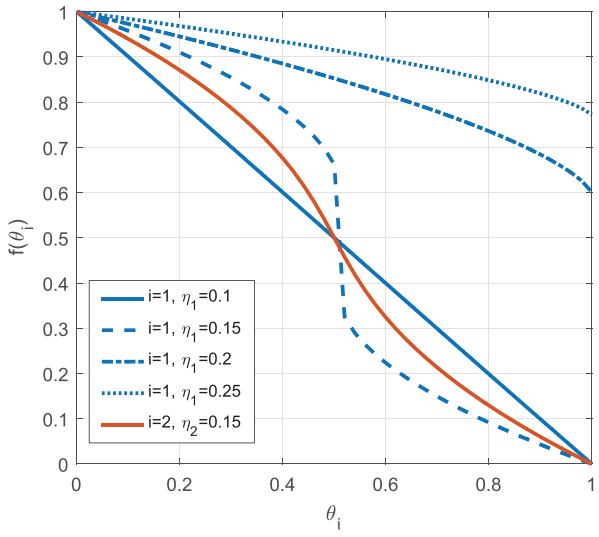

Fig. 5. The optimal encoding functions for θ_1 and θ_2 for $\eta_1 \in \{0.1, 0.15, 0.2, 0.25\}$ and $\eta_2 = 0.15$.

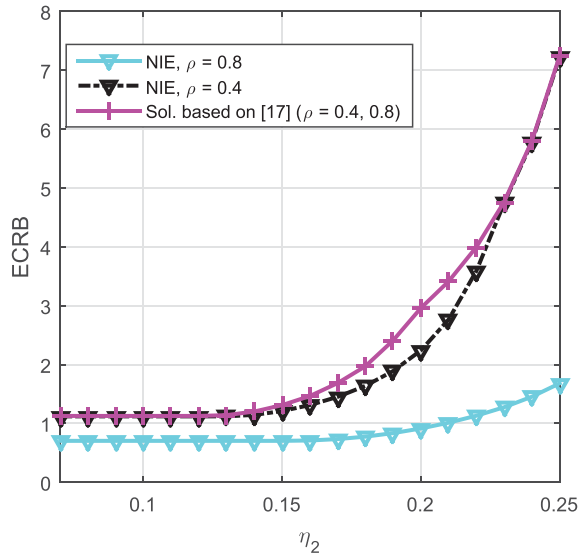

Fig. 6. Total ECRB values versus η_2 for different approaches.

$\rho = 0.4$ and $\rho = 0.8$ and the parameters are set to $h_{e,1} = 1$, $h_{e,2} = 1.5$, and $\eta_1 = 0.15$. Note that the solution based on [17] is the same for both ρ values, as it does not take ρ into account. It is observed that NIE has better performance than the solution based on [17], and the performance gap dramatically increases when the noise components have high correlation in this scenario. This is intuitive as optimizing the encoders in a joint manner makes sense in a correlated environment. However, if the correlation is decreased, the performance of NIE will converge to that of the solution based on [17] as proven in Proposition 1. Note that this can be observed in Fig. 2 as well. The performance of NIE and the solution based on [17] would be same for $\rho = 0$, and as ρ increases, ECRB of NIE starts to decrease in Fig. 2, however the solution based on [17] would stay constant, yielding a non-negligible performance difference especially in scenarios with medium and high correlation in the noise components.

Finally, the maximum estimation error values at the the eavesdropper are given in Table I when the parameters are directly sent to the channel without any encoding, i.e., $f_i(\theta_i) = \theta_i$ for $i = 1, 2$, to further emphasize the importance of the encoding operation. If there exists no eavesdroppers, not applying any encoding is a logical option, as the encoding operation can cause a loss in receiver's estimation accuracy. However, under secrecy constraints, lack of encoding can compromise the security, and a limited error can be caused at the eavesdropper. It is observed from Table I that the achievable target error levels are around 0.07 or lower for the simulation parameters considered in this study; however, larger error values are possible if NIE is applied as illustrated in the examples.

B. Affine Joint Encoding

In this part, we investigate the affine joint encoding strategy and obtain the optimal precoding matrix $\boldsymbol{P}$ to satisfy certain secrecy constraints. In all the numerical examples, $\boldsymbol{\theta}$ is modeled as $\boldsymbol{\theta} = [\theta_1\ \theta_2]^T$, and θ_1 and θ_2 are assumed to be independent of each other with $\theta_1, \theta_2 \in [0, 1]$. Also, the channel parameters for the intended receiver are taken to be $h_{r,1} = h_{r,2} = 2$. The precoding matrix is expressed as $\boldsymbol{P} = \left[\begin{smallmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{smallmatrix}\right]$. Note that $|p_{11}| + |p_{12}| \leq 1$ and $|p_{21}| + |p_{22}| \leq 1$ should be satisfied to ensure $\beta_1, \beta_2 \in [0, 1]$. The strategies considered in the numerical results are given as follows:

- *Affine Joint Encoding (AJE):* This approach refers to the solution of the optimization problem in (26).
- *Nonlinear Individual Encoding (NIE):* This approach refers to the solution of the optimization problem in (14).
- *Affine Individual Encoding (AIE):* This is a simplified version of the AJE approach. In particular, precoding matrix $\boldsymbol{P}$ has the form of $\boldsymbol{P} = \boldsymbol{W}_1\boldsymbol{W}_2$, where $\boldsymbol{W}_1 = \text{diag}\{w_1, w_2, \ldots, w_N\}$ is a diagonal matrix and $\boldsymbol{W}_2$ is a permutation matrix. The AIE approach can further be grouped as follows:
  1) *AIE without permutation:* This refers to special case with $\boldsymbol{W}_2 = \mathbf{I}$. For $N = 2$, we assume $p_{12} = p_{21} = 0$.
  2) *AIE with permutation:* This refers to the scenario with $\boldsymbol{W}_2 \neq \mathbf{I}$. For $N = 2$, we assume $p_{11} = p_{22} = 0$.

We provide five different examples to investigate the affine joint encoding strategy numerically. In the examples, different values for eavesdropper's fading coefficients and prior distributions for θ_1 and θ_2 are used in order to show the advantages and disadvantages of certain encoding strategies over each other in terms of their performance and to corroborate the theoretical results provided in the manuscript. For the first four examples, the channel noise for the eavesdropper and the intended receiver is taken to be zero-mean Gaussian random variables with independent components of unit variance, i.e., $\boldsymbol{\Sigma}_e = \boldsymbol{\Sigma}_r = \mathbf{I}$. In the first example, θ_1 and θ_2 are assumed to be uniformly distributed and the secrecy target for the second parameter, η_2, is set to be 0.15. Also, the eavesdropper fading coefficients are taken as $h_{e,1} = 1.2$ and $h_{e,2} = 1.5$. In Fig. 7, the total optimal ECRB values for θ_1 and θ_2 are plotted for various η_1

TABLE I
MAXIMUM SECRECY TARGET LEVEL VALUES FOR θ_1 AND θ_2, WHEN $f_i(\theta_i) = \theta_i$ FOR $i = 1, 2$

$h_{e,1} = 1, h_{e,2} = 1.5$	η_1	η_2	$h_{e,1} = 1.2, h_{e,2} = 1.5$	η_1	η_2
$\rho = 0$	0.0769	0.0702	$\rho = 0$	0.0744	0.0702
$\rho = 0.3$	0.0764	0.0692	$\rho = 0.3$	0.0738	0.0692
$\rho = 0.5$	0.0754	0.0670	$\rho = 0.5$	0.0723	0.0671
$\rho = 0.7$	0.0730	0.0621	$\rho = 0.7$	0.0692	0.0625
$\rho = 0.9$	0.0660	0.0478	$\rho = 0.9$	0.0605	0.0497

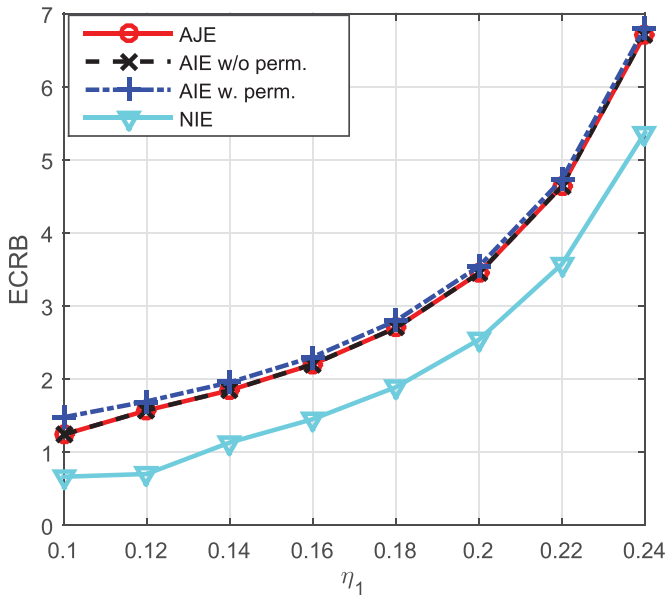

Fig. 7. Total ECRB versus η_1 for different approaches.

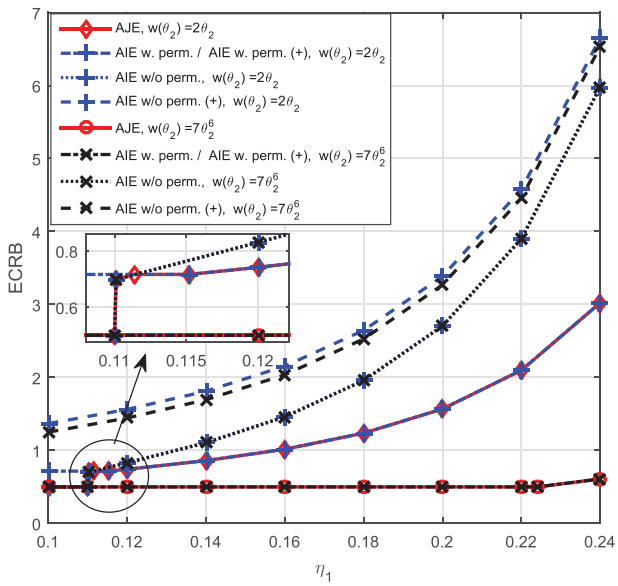

Fig. 8. Total ECRB versus η_1 for different approaches.

values. It is observed that NIE provides improved performance compared to the affine encoding options for this scenario. Also, the optimal AJE solution is the same as the optimal AIE without permutations and they perform slightly better than AIE with permutations.

For the second example, we investigate affine encoding strategies in more detail. The simulation parameters are the same as the first example except that the distribution of θ_2 is taken to be $w(\theta_2) = 2\theta_2$ and $w(\theta_2) = 7\theta_2^6$. The secrecy target for the second parameter, η_2, is set to 0.15. In Fig. 8, the total optimal ECRB values for θ_1 and θ_2 versus η_1 are plotted for various affine encoding strategies. For AIE with and without encoding strategies, we also study the case in which the coefficients of the matrix are restricted to be positive and this is illustrated in the legend of Fig. 8 with (+) next to the name of the corresponding strategy, e.g., AIE w/o perm. (+). When $w(\theta_2) = 7\theta_2^6$, the solutions for the optimal AJE, AIE with permutation and AIE with permutation with positive coefficients are the same and yield the best performance, whereas AIE without permutation with positive coefficients gives the worst performance. AIE without permutation provides a moderate performance except for $\eta_1 < 0.11$, where it also provides the optimal performance. When $w(\theta_2) = 2\theta_2$, AIE with permutation and AIE with permutation with positive coefficients have the same performance, and they perform better than AIE without permutation when $\eta_1 > 0.111$; however, AIE without permutation is better when $\eta_1 < 0.111$. The optimal AJE solution achieves the minimum of these three strategies at all η_1 values. AIE without permutation with positive coefficients yields the worst performance in this case, as well. Note that Corollary 1 and Lemma 3 can be applied in this example for AIE with permutation strategy. As $E(\theta_1) = 1/2$, and eavesdropper's noise is white, Corollary 1 and Lemma 3 imply together that for the AIE with permutation strategy, the matrix elements can be restricted to be positive without loss of generality. Therefore, it is not a coincidence that AIE with permutation and AIE with permutation with positive coefficients yield the same performance in this example.

For the third example, θ_1 is assumed to be uniformly distributed and the distribution of θ_2 is taken to be $w(\theta_2) = 4\theta_2^3$. The secrecy targets for both parameters are set to 0.15. In Fig. 9, the total optimal ECRB values for θ_1 and θ_2 are plotted for various $h_{e,1}$ values when $h_{e,2} = 1.5$. It is observed that the performance of AIE with permutation and AIE with permutation with positive coefficients are the same as $E(\theta_1) = 1/2$ for this example, as well. Their performance stays constant as $h_{e,1}$ increases. The performance of AIE without permutation is initially worse than that of AIE with permutation; however, it improves as $h_{e,1}$ increases and performs better when $h_{e,1} > 2.57$. AIE without permutation with positive coefficients yields the worst

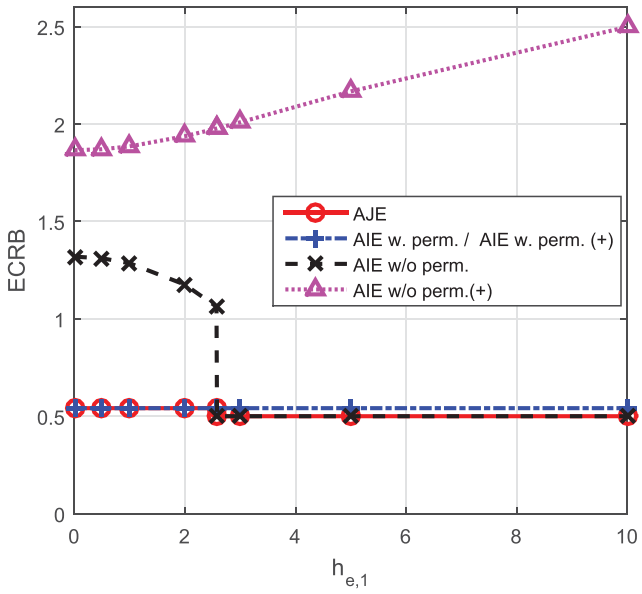

Fig. 9. Total ECRB versus $h_{e,1}$ for different approaches.

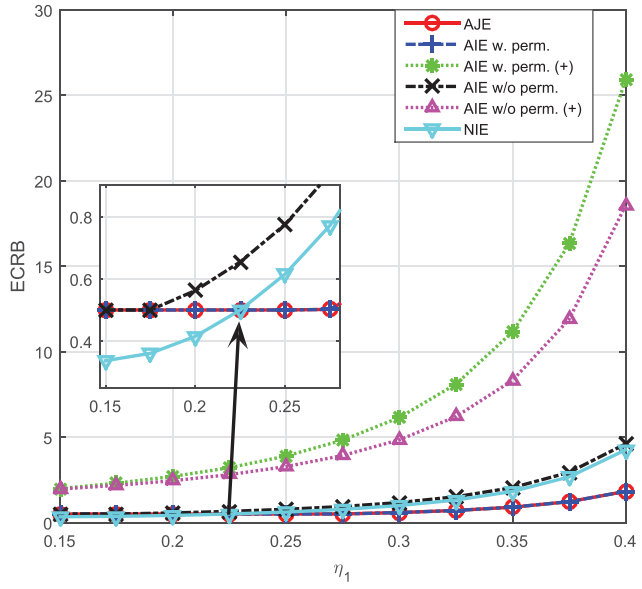

Fig. 10. Total ECRB versus η_1 for different approaches.

performance, and its performance gets even worse as $h_{e,1}$ increases. The different responses of the strategies to the increase of $h_{e,1}$ are due to the fact that the structure of $\mathbf{\Sigma}_{err}$ varies as the encoding strategy changes. The optimal AJE solution is the same as AIE with permutation when $h_{e,1} < 2.57$ and it is same as AIE without permutation when $h_{e,1} \geq 2.57$.

For the fourth example, the distribution of θ_1 is taken to be $w(\theta_1) = 2\theta_1$ and the distribution of θ_2 is given by $w(\theta_2) = 4\theta_2^3$. The secrecy target for the second parameter, η_2, is set to 0.2. In Fig. 10, the total optimal ECRB values for θ_1 and θ_2 are plotted for various η_1 values. It is observed that when $\eta_1 < 0.225$, the best performance is obtained by employing NIE; however, after $\eta_1 > 0.225$, the optimal AJE solution, which has the same performance as AIE with permutation, starts to yield the best

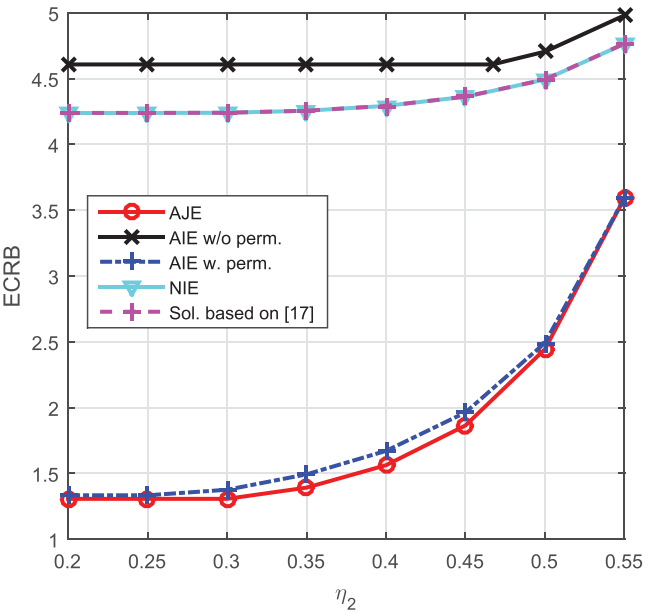

Fig. 11. Total ECRB versus η_2 for different approaches.

performance. This shows that the simple flip and scale approach may be better than the individual nonlinear encoding function strategy in certain scenarios. AIE without permutation performs slightly worse than NIE. AIE with/without permutation with positive coefficients do not achieve a good performance in this scenario. As the conditions given in Corollary 1 are no longer satisfied, there is a significant performance gap between the optimal AIE solutions and the AIE solutions which are restricted to positive coefficients.

In all the four examples, we have observed that the optimal AJE solution has the form of one of the AIE solutions. However, this does not have to be the case in all scenarios and the fifth example provides such an example. In this example, eavesdropper's fading coefficients are taken as $h_{e,1} = 0.8$ and $h_{e,2} = 1.25$. The channel noise for the eavesdropper is modeled as zero-mean multivariate Gaussian random variable with the covariance matrix $\mathbf{\Sigma}_e = \left[\begin{smallmatrix} \sigma_{e,1}^2 & \rho_e \\ \rho_e & \sigma_{e,2}^2 \end{smallmatrix}\right]$, where $\sigma_{e,1}^2 = \sigma_{e,2}^2 = 1$ and $\rho_e = -0.5$ and the channel noise for the eavesdropper is also modeled as zero-mean multivariate Gaussian random variable with the covariance matrix $\mathbf{\Sigma}_r = \left[\begin{smallmatrix} \sigma_{r,1}^2 & \rho_r \\ \rho_r & \sigma_{r,2}^2 \end{smallmatrix}\right]$, where $\sigma_{r,1}^2 = \sigma_{r,2}^2 = 1$ and $\rho_r = 0.7$. The distribution of θ_1 is taken to be $w(\theta_1) = 2\theta_1$ and the distribution of θ_2 is given by $w(\theta_2) = 5\theta_2^4$. The secrecy target for the first parameter, η_1, is set to be 0.4, and the total optimal ECRB values for θ_1 and θ_2 are plotted for various η_2 values. In Fig. 11, it is observed that the optimal AJE solution is better than both the optimal AIE with and without permutation solutions. For example, when $\eta_2 = 0.35$, the optimal precoding matrix for the AIE with permutation solution is $\left[\begin{smallmatrix} 0 & -0.4787 \\ -0.7807 & 0 \end{smallmatrix}\right]$, yielding an objective value of 1.5012. On the other hand, the optimal precoding matrix for the AJE strategy is $\left[\begin{smallmatrix} 0 & -0.48 \\ -0.6578 & -0.2439 \end{smallmatrix}\right]$, yielding an objective value of 1.4008.[7] Therefore, it is possible

[7]The corresponding optimal $\boldsymbol{r}$ values for the AIE with permutation solution and the AJE solution can be found as $\boldsymbol{r}^T = [0.4787\ 0.7807]$ and $\boldsymbol{r}^T = [0.48\ 0.9017]$ respectively.

TABLE II
MAXIMUM SECRECY TARGET LEVEL VALUES FOR θ_1 AND θ_2 WHEN $\boldsymbol{P} = \mathbf{I}$ AND $\boldsymbol{r} = \mathbf{0}$

Parameters	η_1	η_2
$h_{e,1} = 1.2, h_{e,2} = 1.5, w(\theta_1) = 1, w(\theta_2) = 1$	0.0744	0.0702
$h_{e,1} = 1.2, h_{e,2} = 1.5, w(\theta_1) = 1, w(\theta_2) = 2\theta_2$	0.0744	0.0494
$h_{e,1} = 1.2, h_{e,2} = 1.5, w(\theta_1) = 1, w(\theta_2) = 7\theta_2^6$	0.0744	0.0118
$h_{e,1} = 1, h_{e,2} = 1.5, w(\theta_1) = 1, w(\theta_2) = 4\theta_2^3$	0.0769	0.0252
$h_{e,1} = 3, h_{e,2} = 1.5, w(\theta_1) = 1, w(\theta_2) = 4\theta_2^3$	0.0476	0.0252
$h_{e,1} = 5, h_{e,2} = 1.5, w(\theta_1) = 1, w(\theta_2) = 4\theta_2^3$	0.0270	0.0252
$h_{e,1} = 10, h_{e,2} = 1.5, w(\theta_1) = 1, w(\theta_2) = 4\theta_2^3$	0.0089	0.0252
$h_{e,1} = 1.2, h_{e,2} = 1.5, w(\theta_1) = 2\theta_1, w(\theta_2) = 4\theta_2^3$	0.0514	0.0252
The parameters of Fig. 11	0.0531	0.0191

that joint encoding of parameters can outperform individual encoding depending on the channel and parameter statistics. It is also observed that NIE and the solution based on [17] have almost the same performance, and even though they are better than AIE without permutation, they perform worse than AIE with permutation and AJE. This implies that, in this particular scenario, the main source of performance improvement is to exploit the fact that there are multiple elements in the vector by shuffling the order of the elements or even jointly encoding them rather than individual encoding via a nonlinear function. Therefore, there might be cases in which it is not very critical to take the correlation in noise components into account in NIE, as the performance improvement can be negligible.

The maximum secrecy target levels with no encoding are provided for this encoding scheme as well; that is, $\boldsymbol{P} = \mathbf{I}$ and $\boldsymbol{r} = \mathbf{0}$ in Table II for all the considered scenarios. It is observed that the achievable secrecy levels are much lower than those of the AJE scheme. It is also interesting to note that as $h_{e,1}$ increases, the secrecy levels decrease in Table II. This is because of the fact that the channel of the eavesdropper gets better and the error performance improves when the original parameter is transmitted. Such an issue does not occur if the optimal AJE is applied, and this can even be turned into an advantage according to Fig. 9 due to the secret encoder. Also, for the parameters of Fig. 11, the maximum error levels for θ_1 and θ_2 are 0.0531 and 0.0191, respectively; however, the optimal AJE can reach $\eta_1 = 0.4$ and $\eta_2 = 0.55$ (and possibly more) according to the fifth example. This shows the clear advantage of the proposed schemes as compared to not utilizing any encoder in the presence of an eavesdropper.

C. Computational Complexity

One of the main factors determining the computational complexity of the proposed algorithms is the dimension of the space in which the search is performed. When we use the piecewise linear approximation (PWL) method to obtain the optimal solution for nonlinear individual encoding, the search is performed over MN variables as described in Section V-A. As M increases, lower ECRB values can be obtained. However, it increases the search dimension and the complexity. For affine joint encoding, the original optimization problem in (26) requires a search over $\boldsymbol{P}$ and $\boldsymbol{r}$, yielding a search space over $N^2 + N$ variables. However, it is shown in Lemma 2 that it is enough to calculate $\boldsymbol{P}$ for optimal encoding reducing the space to N^2 variables.

Another important factor related to the computational complexity of encoder optimization is the number of multiplications at the calculation of the cost and objective functions for a given candidate encoder. For NIE, both the objective and cost functions require a calculation of an N dimensional integral. Let X denote the terms in the Riemann sum for a given step size. Then, the objective function requires $\mathcal{O}(NX)$ multiplications. To calculate $\boldsymbol{\Sigma}_{err}$, each of $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\beta},\boldsymbol{\theta}}$ needs $\mathcal{O}(N^2X)$ and $E(\boldsymbol{\beta})$ needs $\mathcal{O}(NX)$ calculations. Then, the overall complexity to calculate (11) becomes $\mathcal{O}(N^2X) + \mathcal{O}(N^3)$. For AJE, the complexity of calculating the cost function and the objective function are both $\mathcal{O}(N^3)$. Therefore, AJE has lower computational complexity especially if N is not very large. However, if N is large, then the optimal matrix calculation can become more costly than the NIE algorithm. Note that AJE is a type of a precoding based encoding strategy; hence it has a comparable complexity to the beamforming strategies in the literature, which are employed in different problems.

As a special case of AJE, AIE is also considered in the numerical examples. If AIE without permutation is employed, the search space reduces to N from N^2, and the complexity of the cost and objective function calculations also decreases relatively. For AIE with permutation, the search space is $N + 1$, where the extra variable indicates the permutation order. Note that when N increases, the possible values for the permutation order increases very quickly. However, it is always possible to prune the size of this set to a practical maximum size, and to choose the permutation order from it.

D. General Observations

We have investigated the optimal encoding of multiple parameters for secure communication for the two proposed practical encoding approaches. For the NIE scheme, it is observed that as the correlation between eavesdropper's noise components increases, the total ECRB cost decreases for a given target secrecy level implying that such a correlation is useful for the parameter encoding task. It is also observed that the encoding function is in either the variance minimizing or maximizing mode depending on the channel quality and the correlation values of the parameter. In the second part, the affine joint and individual encoding schemes are compared with each other for various parameter distributions. It is observed that in many scenarios, the solution of the AJE scheme is in the form of the AIE solution, which can be with or without permutations. This implies that individually encoding each parameter can be good enough to solve the optimization problem in most cases. However, it is important to emphasize that this is not a theory as it is possible to find counter examples. Also, when AJE and NIE are compared to each other, it is observed that one can have better performance than the other depending on the scenario. This means that in certain scenarios, simple permutation or/and scaling of the parameters can be the effective security solution

and in some cases, using a nonlinear function without utilizing any permutation brings more benefits.

We note that the main goal behind the encoding operation is to achieve the desired secrecy levels by also providing a certain estimation quality at the receiver. It is observed that both of the proposed approaches have the ability to achieve large estimation errors at the eavesdropper, which could not have been possible if there were no encoding utilized. Another important contribution of this study is that we provide useful theoretical simplifications (and sufficient conditions to apply them) to obtain optimal encoders in practice and they have been used in all the examples. Proposition 1 is utilized when $\rho = 0$ and decoupled optimization problems are solved. Similarly, Lemma 1 is also applied and even though the encoding functions are obtained jointly for $\rho > 0$, they are searched over decreasing functions as it can be observed in Figs. 3 and 5. For AJE, Lemma 2 is utilized to solve the optimization problem and the search is restricted to the optimal precoding matrix as the constant term can be found theoretically for a given encoder. It is observed that even though signed permutation matrices are optimal when there is no secrecy constraint according to Proposition 2, they are not optimal under secrecy requirements in general and either joint encoding and/or scaling of the parameters is required. Finally, Corollary 1 and Lemma 3 are utilized to show that in certain scenarios, matrix coefficients can be restricted to be positive as can be observed in Figs. 8 and 9.

VI. Conclusions

In this paper, optimal encoding of multiple parameters has been investigated in the presence of an eavesdropper. An optimization problem has been proposed with an objective to minimize the ECRB at the intended receiver while satisfying the MSE targets at the eavesdropper. Two practical encoding schemes, i.e., NIE and AJE, have been proposed. It has been observed that both schemes are able to create large estimation errors at the eavesdropper, which is not possible when no encoding is applied, and they can be employed as a security measure. The performance of both schemes has also been compared to each other and it has been observed that one can have better performance than the other depending on the scenario. Another observation is that the optimal encoding function for NIE is in either the variance minimizing or maximizing mode, and in many scenarios individually encoding the parameters with an affine function (with or without permutation of the parameters) has as good performance as that of AJE. Also, theoretical results derived in the manuscript prove useful in simplifying the optimal encoding problem.

Finally, we note that it is entirely possible that two strategies proposed in this paper can be combined to further optimize the encoding function. For example, the first encoding block can perform nonlinear individual encoding and the second encoding block can perform affine joint encoding to the output of the first block. In that case, both the nonlinear individual encoding functions and the precoding matrix should be optimized jointly and despite the increase in computational complexity, the performance can further be improved. As future work, we aim to investigate scenarios in which the eavesdropper has full or partial knowledge of the encoder and the transmitter employs a stochastic encoder to possibly enhance security.

References

[1] W. Diffie and M. Hellman, "New directions in cryptography," *IEEE Trans. Inf. 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," *Mag. Commun. ACM*, vol. 21, no. 2, pp. 120–126, Feb. 1978.

[3] N. Yang, L. Wang, G. Geraci, M. Elkashlan, J. Yuan, and M. D. 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] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, 1975.

[6] 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.

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

[8] A. Khisti, A. Tchamkerten, and G. W. Wornell, "Secure broadcasting over fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2453–2469, Jun. 2008.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] X. Guo, A. S. Leong, and S. Dey, "Distortion outage minimization in distributed estimation with estimation secrecy outage constraints," *IEEE Trans. Signal Inf. Proccess. Netw.*, vol. 3, no. 1, pp. 12–24, Mar. 2017.

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] M. Pei, J. Wei, K.-K. Wong, and X. Wang, "Masked beamforming for multiuser MIMO wiretap channels with imperfect CSI," *IEEE Trans. Wireless Commun.*, vol. 11, pp. 544–549, Feb. 2012.

[21] 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*, 2017, pp. 200–204.

[22] Y. Chen, S. Kar, and J. M. F. Moura, "The Internet of Things: Secure distributed inference," *IEEE Signal Process. Mag.*, vol. 35, no. 5, pp. 64–75, Sep. 2018.

[23] H. L. Van Trees and K. L. Bell, Eds., *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*. Hoboken, NJ, USA: Wiley, 2007.
[24] M. F. Keskin, E. Gonendik and S. Gezici, "Improved lower bounds for ranging in synchronous visible light positioning systems," *J. Lightw. Technol.*, vol. 34, no. 23, pp. 5496–5504, Dec. 2016.
[25] A. A. Nasir *et al.*, "Optimal training sequences for joint timing synchronization and channel estimation distributed communication networks," *IEEE Trans. Commun.*, vol. 61, no. 7, pp. 3002–3015, Jul. 2013.
[26] H.V. Poor, *An Introduction to Signal Detection and Estimation*. New York, NY, USA: Springer, 1994.
[27] Steven M. Kay, *Fundamentals of Statistical Signal Processing—Estimation Theory*, vol. 1. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.

Cagri Goken (S'10) received the B.S. and M.S. degrees in electrical engineering from Bilkent University, Ankara, Turkey, in 2009 and 2011, respectively, and the M.A. degree in electrical engineering from Princeton University, NJ, USA, in 2014. He is currently working toward the Ph.D. degree at Bilkent University. Since 2016, he has been with 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 University, Princeton, NJ, USA, in 2006. From 2006 to 2007, he was with 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*.

Orhan Arikan (M'91) was born in 1964, in Manisa, Turkey. He received the B.Sc. degree in electrical and electronics engineering from the Middle East Technical University, Ankara, Turkey, in 1986, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Illinois Urbana-Champaign, Champaign, IL, USA, in 1988 and 1990, respectively. Following his graduate studies, he was a Research Scientist for three years with Schlumberger-Doll Research Center, Ridgefield, CT, USA. During this time, he was involved in the inverse problems and fusion of multiple modality measurements. In 1993, he joined the Department of Electrical and Electronics Engineering, Bilkent University, Ankara, Turkey, where he is currently a Professor and the Chairman. His research interests are in the areas of statistical signal processing and remote sensing. He was the recipient of the Distinguished Teaching Award of Bilkent University in 1998, and the Young Investigator Award in Engineering from Turkish Scientific and Technical Research Foundation in 2002. He was the Chairman of IEEE Signal Processing Society, Turkey Section in 1995–1996 and was the President of IEEE Turkey Section in 2000–2001.