Error Probability of FDD-based Secret Key Generation using Multiple Linear Regression

1st Ehsan Olyaei Torshizi
School of Computer Science and Engineering
Constructor University Bremen
Bremen, Germany
eolyaei@constructor.university

2nd Werner Henkel, *senior member, IEEE*
School of Computer Science and Engineering
Constructor University Bremen
Bremen, Germany
werner.henkel@ieee.org

***Abstract*—In frequency division duplexing (FDD) systems, the uplink and downlink transmit information in different frequency bands, so it is difficult to use channel reciprocity to generate secret keys. Since the reciprocity holds for the same frequency ranges, if we consider the FDD bands very close to each other, we can anticipate continuity between uplink and downlink bands which can guarantee the required reciprocity. In this paper, we use phase differences between neighboring antennas in an antenna array to construct reciprocal channel features. Moreover, we increase the reciprocity by applying polynomial curve-fitting on the measurements so that two users can generate highly similar keys in FDD systems. Exploiting an effective pre-processing procedure, our proposed scheme achieves competitive performance in terms of efficiency and key disagreement rate (KDR). In addition, we present a detailed statistical analysis to determine the error probability for generated keys from both sides. Numerical simulation results are presented to verify the feasibility and effectiveness of the proposed scheme.**

***Index Terms*—Secret key generation, physical layer security, FDD, error probability, multiple linear regression.**

I. Introduction

Supporting high transmission rates for the extensive range of wireless devices can facilitate the fast spreading of the Internet-of-Things (IoT) for a wide collection of smart industries. The intrinsically shared nature of wireless transmissions and large-scale IoT environments with potentially untrusted nodes give rise to a large number of security threats and vulnerabilities [1]. Wireless physical layer security, as the first line of defense against eavesdropping, aims to keep the information transmitted between legitimate partners safe from adversarial eavesdropping and intervention. Secret key generation (SKG) is a promising candidate which relies on the reciprocity, spatial decorrelation, and temporal variation of the wireless channel to generate the symmetric shared key between two communication partners.

In FDD systems, unlike time-division duplexing (TDD) systems, the uplink and downlink transmit over different carrier frequencies and experience different fading. Accordingly, most of the mutually attainable channel parameters which can be used in TDD systems, may not be the same in FDD systems between the uplink and downlink [2]. It is thus challenging to find reciprocal channel parameters in FDD systems. Since FDD is dominant in existing cellular communications, such as narrowband IoT and Long Term Evolution (LTE), key generation for such FDD-based systems is demandable.

Several key generation methods are developed for FDD systems which are based on using the angle and delay of path [3], channel covariance matrix [4], loop-back mechanisms [5] and [6], and received signal strength indicator (RSSI) [7]. They come with some limitations such as security problems, too large complexity, and high overhead. In this paper, in line with our previous works [8] and [9], we directly employ the phase differences between neighboring antennas derived from scattering parameters S12 and S21 bidirectional measurements between a circular antenna array and a single dipole counterpart. Since we consider FDD bands very close to each other and the reciprocity holds for the same frequency range, the existing continuity between frequency bands can guarantee the required reciprocity. The proposed idea does not need large commputational cost or any additional reverse channel training. In addition, we extract an error probability relation for the presented SKG model by providing a detailed statistical analysis and employing a multiple linear regression model.

The remainder of the paper is organized as follows. In Section II, we describe the system setup, our testbed, and all components of each measurement set. The proposed SKG scheme is illustrated in Section III. In sections IV and V, error probability relation and estimation of parameters using multiple linear regression are analyzed, respectively. Conclusions are drawn in Section VI.

II. System Model

We consider the basic key generation model in which Alice and Bob are two legitimate counterparts intending to transmit data securely over a wireless channel in the presence of an adversary Eve acting as a passive attacker trying to eavesdrop confidential information exchanged between them. In our setup, Alice is realized as a circular antenna array with 40 antenna positions. Moreover, we consider single dipoles for Bob and Eve. Consequently, after each round of measurements, we have 40 phase differences between consecutive neighboring antennas and we consider each of them as a measurement set that can generate a secret key between Alice and Bob. We

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – HE 3654/27-1.

use a standard vector network analyzer to measure scattering parameters S_{12} and S_{21} in a remotely controlled fashion. We considered many scenarios in wireless indoor environments including office, home, basement, corridor, etc. In order to construct the channel profiles, we measure S_{12} and S_{21} in two closely neighboring $\Delta f = 5$ MHz frequency bands on both sides of a central (carrier) frequency of $f_0 = 2.19$ GHz such that the uplink and downlink frequency width, satisfies that $\Delta f \ll f_0$. Each FDD band separately consists of 801 frequency samples and the collection of these samples from both sides constitutes a measurement set. After completing all the required steps for key generation, which we will explain in the next section, the rightmost point of S_{12} (frequency sample 801) and the leftmost point of S_{21} (frequency sample 802) can be considered as two Gaussian random variables that based on which quantization interval they are in can provide the keys from Alice and Bob's side, respectively.

III. Proposed SKG for FDD systems

In this section, we elaborate on the proposed SKG scheme which employs the phase differences between neighboring antennas derived from scattering matrix parameters S_{12} and S_{21} bidirectional measurements for FDD systems. The proposed scheme in its most complete form consists of five phases: determine phase differences between neighboring antennas from measured scattering parameters S_{12} and S_{21}, pre-process the measurements, quantization, key reconciliation, and finally privacy amplification. In the end, identical secret keys shall be generated.

Our comprehensive investigations on the resulting measurement sets showed that the presence of 2π jumps on the original phase differences led to generating dissimilar keys and consequently increase the key disagreement rate (KDR). Employing unwrapping, the most common solution to prevent them, not only could not provide us with clean data in the correct interval between $-\pi$ to π but also caused a dramatic increase in the variance of the data which lead to not satisfying our threshold value for variance and reducing the efficiency. Hence, we employ a two-stage pre-processing step which includes a jump-removal technique that is followed by an outlier-correction stage [8] to provide clean data with minimum variance in comparison with both original and unwrapped versions. Afterwards, we curve-fit the resulting phase differences to improve channel reciprocity. Employing second-order polynomial curve fitting can approximate the frequency behavior of the phase difference between the transmission characteristics of the neighboring antennas well [8]. Ideally, due to the reciprocity, we expect a direct continuity between the S_{12} and S_{21} phase difference spectral segments $\big(S_{12}(801) \approx S_{21}(802)\big)$. An example of the original phase difference measurements for S_{12} and S_{21} measured from a circular array along with the corresponding unwrapped and pre-processed versions are illustrated in Fig. 1. In order to generate primary secret keys, the midpoint phase difference estimate should be quantized. We employ a linear quantization scheme that divides the complete 2π phase range into 2^M

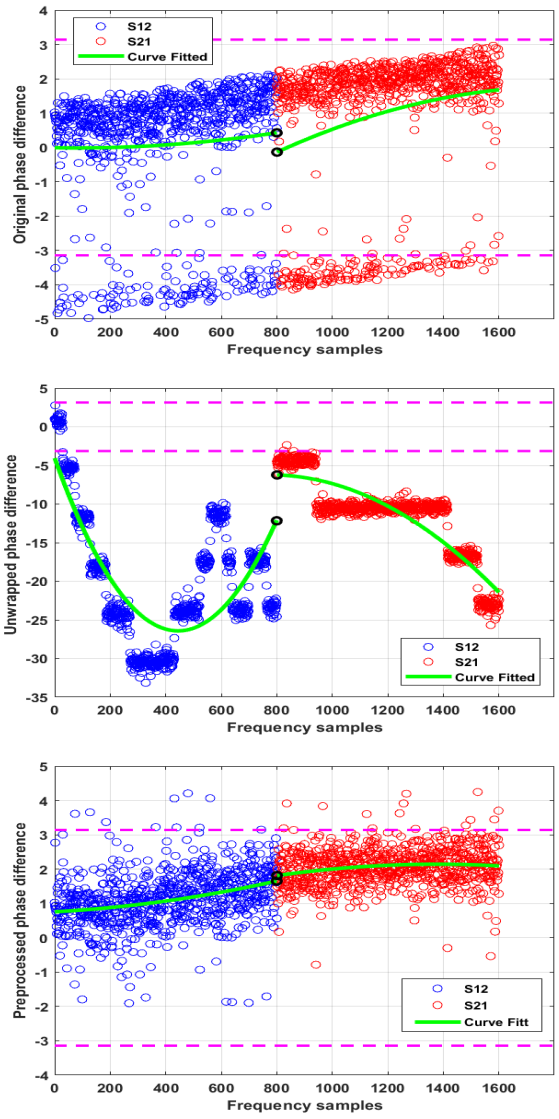

Figure 1. (Phase difference measurements for S12 and S21) Top: Original, Middle: Unwrapped, Low: Pre-processed

equal quantization intervals. Applying a Gray coding scheme leads to allocating an M-bit binary codeword to each quantization level which we interpret as primary keys between Alice and Bob. The Gray coding ensures a single-bit change when crossing quantization boundaries. Moreover, to reduce the key disagreement rate, the primary keys of Alice and Bob should be reconciled. We force the quantized measurements from one side to be at the midpoint of the resulting quantization intervals and consider this as the first stage of key reconciliation. This amount of shift is publicly communicated, such that especially the legitimate counterpart can likewise adjust the quantization grid. Further key reconciliation steps like Slepian-Wolf coding-based approaches and possibly following privacy amplification can be applied to avoid leakage to an eavesdropper. Figure 2 shows the allocated Gray code after quantization and reconciliation.

In order to further clarify the importance of data pre-

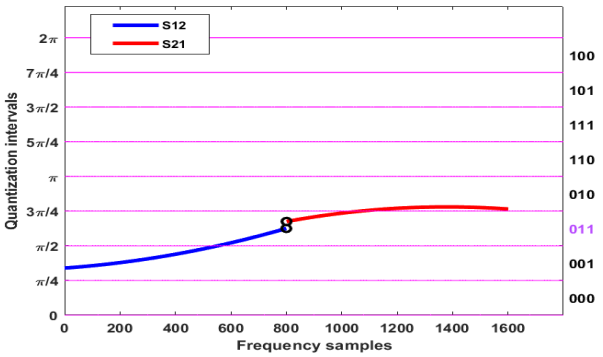

Figure 2. linearly quantized phase difference measurements

processing, we consider the phase differences resulting from a complete measuring round including 40 sets. Unusable measurements we conclude from a very noisy phase. In our processing, we rejected measurements when the variance of the phase relative to the fitted polynomial exceeded "1" radian. Figure 3 provides a variance comparison possibility for original, unwrapped, and pre-processed versions of the measurements. As we can clearly see, employing unwrapping decreases the variance in sets $\{2, 3, 28, 29, 30, 33, 37, 38\}$ for S_{12} and sets $\{24, 25\}$ for S_{21}, but it leads to an increase in the variance of sets $\{10, 11\}$ and $\{34, 35\}$ for S_{12} and S_{21}, respectively, and consequently, the elimination of them due to not satisfying the variance threshold. Moreover, it dramatically increases the variance of the sets $\{4, 7, 8, 9, 15, 16\}$ for S_{12} and $\{10, 11, 15, 16, 28\}$ for S_{21} which were clean data originally and this causes efficiency decreasing conspicuously. The pre-processing can correct the variance of all measurement sets and provide a very good efficiency. To demonstrate the performance of our reconciliation method, the resulting keys for both, Alice and Bob, for all 40 previously considered measurement sets are presented in Fig. 4. Presented results in figs. 1 and 2 are related to the thirty-fifth set where one can clearly recognize from Fig. 3 that unwrapping was not able to correct, but as Fig. 4 shows, the pre-processed version indeed generates the same key "011" from both sides.

IV. Error Probability Derivation

Consider $X \sim N(\mu_x, \sigma_x^2)$ and $Y \sim N(\mu_y, \sigma_y^2)$ as two Gaussian random variables which represent the rightmost point of the interpolated S_{12} and the leftmost point of the interpolated S_{21}, respectively. We determine the probability that X and Y end up in different quantization intervals which means an error or key disagreement. Let N be the number of quantization levels. Then the width of each quantization interval would be $\Delta = \frac{2\pi}{N}$ considering linear quantization. The probability that a Gaussian RV X falls in between two consecutive thresholds at $(m-1)\frac{2\pi}{N}$ and $m\frac{2\pi}{N}$, where m is an integer between $1 - \frac{N}{2}$ to $\frac{N}{2}$, can be computed as

$$P\Big((m-1)\frac{2\pi}{N} \le X \le m\frac{2\pi}{N}\Big) = \frac{1}{2}\Big[\mathrm{erf}\Big(\frac{2\pi m - N\mu_x}{\sqrt{2}N\sigma_x}\Big) - \mathrm{erf}\Big(\frac{2\pi(m-1) - N\mu_x}{\sqrt{2}N\sigma_x}\Big)\Big] . \quad (1)$$

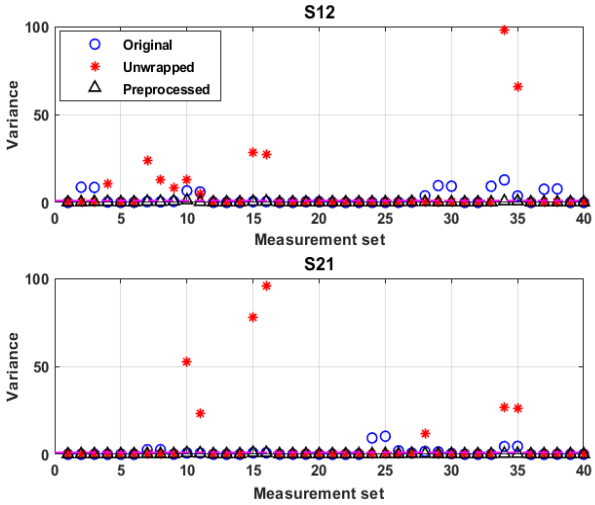

Figure 3. A variance comparison between all kinds of measurements

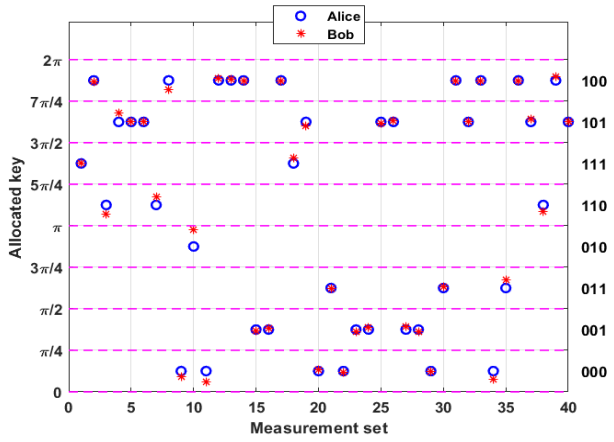

Figure 4. The resulting keys from Alice and Bob for all measurement sets of a complete measuring round

Depending on which quantization interval the midpoint is at, we have two M-tuple vectors from left and right for S_{12} and S_{21}, respectively. Ending up in two different quantization intervals usually just means a single-bit error between the two resulted keys from left and right. Consequently, the error probability would be equal to the probability of the situation in which two RVs X and Y have different Gray codes, i.e.,

$$P_{error} = P(X_{GC} \neq Y_{GC}) . \quad (2)$$

and the corresponding BER would be P_{error}/M.

In our case, considering 8 quantization intervals and 3 Bit Gray coding, we have

$$P_{error} = P(X_{GC} = 000) \cdot P(error|X_{GC} = 000) + \cdots + P(X_{GC} = 100) \cdot P(error|X_{GC} = 100) . \quad (3)$$

Since the conditional probability of error given a specific Gray sequence for X is the same as the probability that the allocated

Gray code to Y is not equal to that specific sequence for X, we can express the error probability equation as

$$P_{error} = P(X_{GC} = 000) \cdot \big(1 - P(Y_{GC} = 000)\big) + \cdots + P(X_{GC} = 100) \cdot \big(1 - P(Y_{GC} = 100)\big) \ . \quad (4)$$

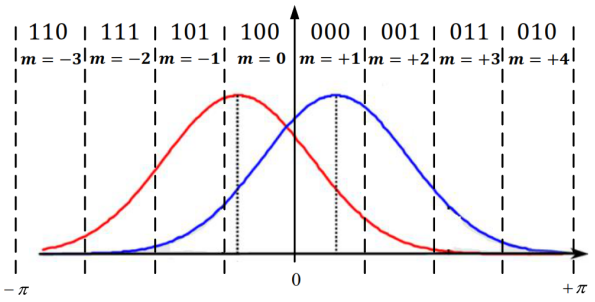

Figure 5. A representation of two Gaussian RVs X and Y with corresponding quantization thresholds and Gray codes

According to the Gray codes assigned to each quantization interval in Fig. 5, the error probability relation can be written as

$$P_{error} = \sum_{m=1-\frac{N}{2}}^{\frac{N}{2}} \Big[P\Big((m-1)\frac{2\pi}{N} \le X \le m\frac{2\pi}{N}\Big)\Big] \cdot \Big[1 - P\Big((m-1)\frac{2\pi}{N} \le Y \le m\frac{2\pi}{N}\Big)\Big] \ . \quad (5)$$

By inserting Eq. (1) into Eq. (5), the final error probability relation can be expressed as

$$P_{error} = \sum_{m=1-\frac{N}{2}}^{\frac{N}{2}} \left[\frac{1}{2}\Big[\mathrm{erf}\Big(\frac{2\pi m - N\mu_x}{\sqrt{2}N\sigma_x}\Big) - \mathrm{erf}\Big(\frac{2\pi(m-1) - N\mu_x}{\sqrt{2}N\sigma_x}\Big)\Big]\right] \cdot \left[1 - \frac{1}{2}\Big[\mathrm{erf}\Big(\frac{2\pi m - N\mu_y}{\sqrt{2}N\sigma_y}\Big) - \mathrm{erf}\Big(\frac{2\pi(m-1) - N\mu_y}{\sqrt{2}N\sigma_y}\Big)\Big]\right] . \quad (6)$$

V. Multiple Linear Regression

In this section, we determine the mean and variance of both aforementioned RVs X and Y to use in the extracted error probability.

A. Data Model

We can consider the multiple linear regression model in the most general case for our data as follows:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \ , \quad (7)$$

where $\mathbf{y}$ and $\mathbf{X}$ are a $n \times 1$ vector of n observations of the study variable and a $n \times (k+1)$ matrix of n observations on each of the $k+1$ explanatory variables, which is often referred to as the design matrix, respectively. Moreover, $\boldsymbol{\beta}$ is a $(k+1) \times 1$ vector including fixed but unknown model parameters representing regression coefficients and a $n \times 1$ vector of $\boldsymbol{\epsilon}$ is related to random error components which can be assumed $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I}_n)$. Moreover, $\mathbf{X}$ is assumed as a non-stochastic matrix such that $\mathrm{rank}(\mathbf{X}) = k$. Spelling out the components of Eq. (7), this reads

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 & \cdots & x_1^k \\ 1 & x_2 & \cdots & x_2^k \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^k \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \ . \quad (8)$$

B. Estimation of parameters

The general procedure for the estimation the regression coefficient vector for $k = 2$ results from minimization of a metric M

$$\sum_{i=1}^{n} \mathrm{M}(\varepsilon_i) = \sum_{i=1}^{n} \mathrm{M}(y_i - \beta_0 - x_i\beta_1 - x_i^2\beta_2) \ . \quad (9)$$

Choosing $\mathrm{M}(x) = x^2$ for the above metric leads to the ordinary least-squares method. Let us consider $\mathcal{B}$ as the $(k+1)-$dimensional real Euclidean space consisting of the set of all possible vectors of $\boldsymbol{\beta}$. The objective is to find a $(k+1)-$tuple vector $\hat{\boldsymbol{\beta}} = (\hat{\beta}_0, \hat{\beta}_1, ..., \hat{\beta}_k)$ from $\mathcal{B}$ that minimizes the sum of squared deviations of $\boldsymbol{\epsilon}^{\mathrm{T}}$ for given $\mathbf{y}$ and $\mathbf{X}$ as

$$\mathrm{S}(\mathbf{b}) = \sum_{i=1}^{n} \varepsilon_i^2 = \mathbf{y}^{\mathrm{T}}\mathbf{y} + \mathbf{b}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{X}\mathbf{b} - 2\mathbf{b}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{y}. \quad (10)$$

To find the desired vector, we should have $\frac{\partial\ \mathrm{S}(\mathbf{b})}{\partial \mathbf{b}} = 2\mathbf{X}^{\mathrm{T}}\mathbf{X}\mathbf{b} - 2\mathbf{X}^{\mathrm{T}}\mathbf{y} = 0$ which implies that $\mathbf{X}^{\mathrm{T}}\mathbf{X}\mathbf{b} = \mathbf{X}^{\mathrm{T}}\mathbf{y}$. If $\mathbf{X}$ is full rank, we have $\mathrm{rank}(\mathbf{X}) = k+1$, then $\mathbf{X}^{\mathrm{T}}\mathbf{X}$ is positive definite and consequently, the unique solution of (9) is

$$\hat{\boldsymbol{\beta}} = (\hat{\beta}_0, \hat{\beta}_1, ..., \hat{\beta}_k) = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y} = \mathbf{b} \ . \quad (11)$$

since $\partial^2\ \mathrm{S}(\mathbf{b})/\partial\mathbf{b}^2 = 2\mathbf{X}^{\mathrm{T}}\mathbf{X}$, at least, is non-negative definite, the aforementioned obtained $\hat{\boldsymbol{\beta}}$ minimizes $\mathrm{S}(\mathbf{b})$. In case $\mathbf{X}$ is not full rank, the solution of Eq. (9) is as follows:

$$\mathbf{b} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{\mathrm{g}}\ \mathbf{X}^{\mathrm{T}}\mathbf{y} + [\ \mathbf{I} - (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{\mathrm{g}}\ \mathbf{X}^{\mathrm{T}}\mathbf{X}\ ]\ \mathbf{w} \quad (12)$$

where $(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{\mathrm{g}}$ represents the generalized inverse of $\mathbf{X}^{\mathrm{T}}\mathbf{X}$ and $\mathbf{w}$ can be considered as an arbitrary vector. If we consider $\mathbf{b}$ as the estimate of $\boldsymbol{\beta}$, then clearly the fitted values are

$$\hat{\mathbf{y}} = \mathbf{X}\mathbf{b} \ , \quad (13)$$

and in the case of $\mathbf{b} = \hat{\boldsymbol{\beta}}$, for the fitted values, we have

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y} \ . \quad (14)$$

by defining $\mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}$ as a matrix $\mathbf{H}$, we obtain

$$\hat{\mathbf{y}} = \mathbf{H}\mathbf{y} \ . \quad (15)$$

The $\mathbf{H}$ matrix maps the vector of observed values (dependent variable values) to the vector of fitted values, and its diagonal elements are defined as the leverages, which describe the influence each response value has on the fitted value for that same observation. This matrix is symmetric and idempotent and we have

$$\mathrm{trace}(\mathbf{H}) = \mathrm{trace}(\mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}) = \mathrm{trace}(\mathbf{X}^{\mathrm{T}}\mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}) = \mathrm{trace}(\mathbf{I}_{k+1}) = k+1 \ . \quad (16)$$

Moreover, we can define the residuals as difference between the observed and fitted values of the study variable as

$$\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{H}\mathbf{y} = (\mathbf{I} - \mathbf{H})\mathbf{y} = \bar{\mathbf{H}}\mathbf{y} \; . \tag{17}$$

The matrix $\bar{\mathbf{H}}$ is symmetric and idempotent and we have

$$\mathrm{trace}(\bar{\mathbf{H}}) = \mathrm{trace}(\mathbf{I}_n) - \mathrm{trace}(\mathbf{H}) = n - (k+1) \; . \tag{18}$$

Theorem 1: If $\mathbf{X}$ is full rank, then $\mathrm{E}(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta}$ and $Cov(\hat{\boldsymbol{\beta}}) = \sigma^2(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}$.

proof.

$$\begin{aligned} \mathrm{E}(\hat{\boldsymbol{\beta}}) &= \mathrm{E}[(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y}] = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}E[\mathbf{y}] \\ &= (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{X}\boldsymbol{\beta} = \boldsymbol{\beta} \; . \end{aligned} \tag{19}$$

and

$$\begin{aligned} Cov(\hat{\boldsymbol{\beta}}) &= Cov[(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y}] \\ &= (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}Cov[\mathbf{y}]\big((\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\big)^{\mathrm{T}} \\ &= (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}Cov[\mathbf{y}]\mathbf{X}\big((\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\big)^{\mathrm{T}} \\ &= (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}Cov[\mathbf{y}]\mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1} \\ &= \sigma^2(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1} \\ &= \sigma^2(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}. \end{aligned} \tag{20}$$

Theorem 2: If $\mathbf{A}$ is an $n \times n$ matrix of constants and $\mathbf{y}$ is an n-dimensional random vector such that $\mathrm{E}(\mathbf{y}) = \boldsymbol{\mu}$ and $Cov(\mathbf{y}) = \boldsymbol{\Sigma}$, then $E(\mathbf{y}^{\mathrm{T}}\mathbf{A}\mathbf{y}) = \mathrm{trace}(\mathbf{A}\boldsymbol{\Sigma}) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{A}\boldsymbol{\mu}$.

proof. See [10]

Theorem 3: If X is full rank, then $\mathrm{E}\{\mathrm{S}(\hat{\boldsymbol{\beta}})\} = \sigma^2\big(n-(k+1)\big)$.

proof. Since

$$\begin{aligned} \mathrm{S}(\hat{\boldsymbol{\beta}}) &= (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})^{\mathrm{T}}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \\ &= \mathbf{y}^{\mathrm{T}}\mathbf{y} - 2\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{y} + \hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{X}\hat{\boldsymbol{\beta}} \\ &= \mathbf{y}^{\mathrm{T}}\mathbf{y} - \hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{y} = \mathbf{y}^{\mathrm{T}}\mathbf{y} - \mathbf{y}^{\mathrm{T}}\mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y} \\ &= \mathbf{y}^{\mathrm{T}}\mathbf{y} - \mathbf{y}^{\mathrm{T}}\mathbf{H}\mathbf{y} = \mathbf{y}^{\mathrm{T}}\bar{\mathbf{H}}\mathbf{y}. \end{aligned} \tag{21}$$

Using Theorem 2 implies that

$$\begin{aligned} \mathrm{E}[\mathrm{S}(\hat{\boldsymbol{\beta}}))] &= \mathrm{trace}\big((\mathbf{I}_n - \mathbf{H})(\sigma^2\mathbf{I})\big) + (\mathbf{X}\boldsymbol{\beta})^{\mathrm{T}}(\mathbf{I} - \mathbf{H})(\mathbf{X}\boldsymbol{\beta}) \\ &= \sigma^2\mathrm{trace}(\mathbf{I}_n - \mathbf{H}) + \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}(\mathbf{I} - \mathbf{H})(\mathbf{X}\boldsymbol{\beta}) \\ &= \sigma^2\mathrm{trace}(\mathbf{I}_n - \mathbf{H}) + \boldsymbol{\beta}^{\mathrm{T}}\big(\mathbf{X}^{\mathrm{T}}\mathbf{X} - \mathbf{X}^{\mathrm{T}}\mathbf{X}\big)\boldsymbol{\beta} \\ &= \sigma^2\mathrm{trace}(\mathbf{I}_n - \mathbf{H}) \; . \end{aligned} \tag{22}$$

using Eq. (18), $\mathrm{E}\{\mathrm{S}(\hat{\boldsymbol{\beta}}))\} = \sigma^2\big(n - (k+1)\big)$.

Finally, σ^2 can be estimated as sum of squares of the residuals i.e.

$$\hat{\sigma}^2 = \frac{1}{n - (k+1)} \sum_{i=1}^{n} (y_i - \hat{y_i})^2 \; . \tag{23}$$

In our case, employing second order polynomial curve fitting we have $y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2$ for $i = 1, 2, ..., n$. Employing Theorem 1, linear regression coefficients would be normal as $\hat{\boldsymbol{\beta}} \sim N(\boldsymbol{\beta}, \sigma^2(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1})$. Hence one can determine the required mean and variance of the resulting Gaussian distribution at the merging point for random variable X as $X \sim N(\beta_0 + \beta_1 x + \beta_2 x^2, \sigma^2_{\beta_0} + x^2\sigma^2_{\beta_1} + x^4\sigma^2_{\beta_2})$ and the procedure for random variable Y is same in which $\sigma^2_{\beta_0}$, $\sigma^2_{\beta_1}$, and $\sigma^2_{\beta_2}$ are the first, second, and the third entries on the main diagonal of covariance matrix of $\hat{\boldsymbol{\beta}}$, respectively.

From Fig. 4 we can clearly recognize that the key distribution over all quantization intervals is not uniform. Using a non-linear quantization scheme [9], this challenge could be addressed easily. As an alternative solution, one can apply arithmetic coding after linear quantization. Mean and variance of the aforementioned measurement set, which is shown earlier in Figs. 1 and 2, are $\mu_x = 1.6551$, and $\sigma_x^2 = 0.5847$ for the rightmost point of S_{12} and $\mu_y = 1.8126$, and $\sigma_y^2 = 0.2939$ for the leftmost point of S_{21}. Using Eq. (6) the corresponding error probability for this set is determined as $P_{error} = 8.79\ E-2$. The average error probability for all 40 measurement sets of Fig. 4, is determined as less than $5\ \%$.

VI. Conclusions

This paper has provided a secret key generation scheme for FDD systems that can simultaneously reach low complexity and good performance metrics. Using neighboring frequency bands very close to each other can guarantee the required reciprocity for such systems. It was shown that employing robust pre-processing steps on the measurements led to getting better results in terms of efficiency and KDR. We provided a detailed statistical analysis for the random variables at the merging point of two FDD bands using a multiple linear regression model. Moreover, we extracted an error probability relation for the generated keys from both sides.

References

[1] Z. Ji, Y. Zhang, Z. He, P. L. Yeoh, B. Li, H. Yin, Y. Li, and B. Vucetic, "Wireless secret key generation for distributed antenna systems: A joint space-time-frequency perspective," *IEEE Internet of Things Journal*, vol. 9, no. 1, pp. 633–647, 2021.

[2] X. Zhang, G. Li, J. Zhang, A. Hu, Z. Hou, and B. Xiao, "Deep-learning-based physical-layer secret key generation for FDD systems," *IEEE Internet of Things Journal*, vol. 9, no. 8, pp. 6081–6094, 2021.

[3] W. Wang, H. Jiang, X. Xia, P. Mu, and Q. Yin, "A wireless secret key generation method based on chinese remainder theorem in FDD systems," *Science China Information Sciences*, vol. 55, no. 7, pp. 1605–1616, 2012.

[4] B. Liu, A. Hu, and G. Li, "Secret key generation scheme based on the channel covariance matrix eigenvalues in FDD systems," *IEEE Communications Letters*, vol. 23, no. 9, pp. 1493–1496, 2019.

[5] A. M. Allam, "Channel-based secret key establishment for FDD wireless communication systems," *Commun. Appl. Electron*, vol. 7, no. 9, pp. 27–31, 2017.

[6] D. Qin and Z. Ding, "Exploiting multi-antenna non-reciprocal channels for shared secret key generation," *IEEE Transactions on information forensics and security*, vol. 11, no. 12, pp. 2693–2705, 2016.

[7] T. Aono, K. Higuchi, T. Ohira, B. Komiyama, and H. Sasaoka, "Wireless secret key generation exploiting reactance-domain scalar response of multipath fading channels," *IEEE Transactions on Antennas and Propagation*, vol. 53, no. 11, pp. 3776–3784, 2005.

[8] E. O. Torshizi, U. Uprety, and W. Henkel, "Highly efficient FDD secret key generation using esprit and jump removal on phase differences," in *2022 IEEE Conference on Communications and Network Security (CNS)*, pp. 1–6, IEEE, 2022.

[9] E. O. Torshizi and W. Henkel, "Reciprocity and secret key generation for FDD systems using non-linear quantization," in *2022 IEEE Globecom Workshops (GC Wkshps)*, pp. 927–932, IEEE, 2022.

[10] A. C. Rencher and G. B. Schaalje, *Linear models in statistics*. John Wiley & Sons, 2008.