Robust Secret Key Extraction from Channel Secondary Random Process

Robust Secret Key Extraction from Channel Secondary Random Process

Ahmed Badawy*†, Tamer Khattab†, Tarek ElFouly‡, Carla-Fabiana Chiasserini *, Amr Mohamed‡ and Daniele Trinchero*

*Politecnico di Torino, DET. (ahmed.badawy, daniele.trinchero, chiasserini@polito.it)

†Qatar University, Electrical Engineering Dept. (tkhattab@qu.edu.qa)

‡Qatar University, Computer Engineering Dept. (tarekfouly,amrm@qu.edu.qa)

***Abstract*—The vast majority of existing secret key generation protocols exploit the inherent randomness of the wireless channel as a common source of randomness. However, independent noise added at the receivers of the legitimate nodes affect the reciprocity of the channel. In this paper, we propose a new simple technique to generate the secret key that mitigates the effect of noise. Specifically, we exploit the estimated channel to generate a secondary random process (SRP) that is common between the two legitimate nodes. We compare the estimated channel gain and phase to a preset threshold. The moving differences between the locations at which the estimated channel gain and phase exceed the threshold are the realization of our SRP. We study the properties of our generated SRP and derive a closed form expression for the probability mass function of the realizations of our SRP. We simulate an orthogonal frequency division multiplexing (OFDM) system and show that our proposed technique provides a drastic improvement in the key bit mismatch rate (BMR) between the legitimate nodes when compared to the techniques that exploit the estimated channel gain or phase directly. In addition to that, the secret key generated through our technique is longer than that generated by conventional techniques. Moreover, we compute the conditional probabilities used to estimate the secret key capacity.**

***Index Terms*—hysical layer security; Secret key generation; Bit mismatch rate; Channel estimation; OFDM systems.hysical layer security; Secret key generation; Bit mismatch rate; Channel estimation; OFDM systems.P**

I. INTRODUCTION

Orthogonal frequency division multiplexing (OFDM) is a multi-carrier modulation scheme that has been widely adopted in many wireless communication systems such as Long Term Evolution (LTE) systems [1]. It provides many advantages over the single-carrier modulation schemes, including: high data rate, immunity to selective fading, resilience to inter-symbol interference and higher spectrum efficiency [2].

As in any wireless communication system, security of OFDM wireless system is a critical issue. Currently, security relies on cryptographic techniques and protocols that lie at the upper layers of the wireless network. One main drawback of these solutions is the necessity of a complex key management scheme in the case of symmetric ciphers and high computational complexity in the case of asymmetric ciphers. On the other hand, physical layer security relies on the randomness of the communication channel and has a much lower computational complexity.

Unlike conventional cryptographic techniques, physical layer security relies on a source of randomness that is common between the legitimate communicating nodes and not shared with malicious nodes. This common source of randomness is typically a physical layer specific characteristic such as channel estimates, which is the most commonly exploited characteristic for secret key generation (SKG). The secret key is then used to encrypt and decrypt the exchanged data. Channel based SKG techniques mainly rely on channel reciprocity assumption. An identical signal that is exchanged between two antennas across a linear and isotropic channel, will be the same at the two receiving sides of the nodes. This is because of the reciprocity of the radiating and receiving antenna pattern [3], [4].

In [5]–[8], channel measurements were exploited to solve the problem of SKG. In [5] the authors observed that the maximum size of the generated secret key mainly depends on the mutual information between the channel estimates at the two legitimate nodes. They also derived an expression for the mutual information for a general multipath channel. The most popular feature of the fading channel characteristics, which is used extensively in the literature, is channel gain, mainly because of its ease of implementation [7], [9]. Others exploit the channel phase to generate the secret key as in [10], [11]. Unlike channel gain, channel phase is uniformly distributed in narrowband fading channels. The authors in [10] were able to generate a *long* key as compared to the conventional cryptographic techniques from the estimated channel phase, while in [11], they extend their system to the use of relay nodes. Exploiting the channel estimates to generate a secret key has also been investigated under multiple antenna scenarios [12] and relaying scenarios [13]. Other physical layer characteristics used to generate the secret key include distance between the two legitimate users as in [14], [15] and angle of arrival as in [16].

In [9], [17], the authors presented a popular technique to extract a secret key that is based on level crossing of the estimated channel gain. The main advantage of their level crossing technique is that it achieves a low bit mismatch rate (BMR) between the key generated at the legitimate nodes. The authors studied the channel probing rate effect on the secret key rate for different Doppler shifts. They found that secret key rate increases as the probing rate increases and saturates at 20 KHz probing rate for the worst case Doppler shift they assumed. The smaller the Doppler shift the smaller the probing rate required to saturate the secret key rate. In [7], the authors observed that as the carrier frequency increases, the probing rate should increase to achieve a suitable key rate. This is mainly because the channel temporal variation increases at higher carrier frequencies.

One main advantage of exploiting channel estimates to generate the secret key is its high key generation rate. However, a major downside of using the channel reciprocity for SKG is that the additive white Gaussian noise (AWGN) at both receivers affects the reciprocity of the channel measurements [18]. This drawback causes the BMR between the legitimate nodes to rise, which affects the operation of SKG based on channel estimates, particularly, at low and medium signal to noise ratio (SNR) scenarios. This issue was stated as one of the challenges of physical layer security in [19].

To address the latter drawback of physical layer security techniques, we propose a robust SKG technique to mitigate the effect of AWGN. We propose a SKG technique, which we apply on the estimated channel gain only, channel phase only and combined gain and phase, which enhances the performance of the SKG system at low and medium SNR levels. In our technique, the estimated channel is considered our primary random process, from which we derive a secondary random process (SRP) that is then used to generate the secret key. The primary random process, which is either the estimated channel gain or phase, is compared to a preset threshold. The locations of the realizations at which the primary random process exceeds the threshold are stored. The moving differences, which are the differences between each two adjacent locations, are the realizations of our SRP. Those realizations are then used to generate the secret key. We derive a closed form expression for the probability mass function of those realizations. Our proposed technique improves the BMR drastically and achieves a longer key length than the conventional techniques. The entropy rate achieved through our technique is comparable to that achieved by conventional techniques. In addition, we numerically compute the conditional probabilities used in secret key capacity estimation.

The rest of this paper is organized as follows: In Section II the system model is presented. Related existing techniques are addressed in Section III. Our proposed channel SRP for SKG technique is presented in Section IV. The properties of our generated SRP are discussed in Section V. The capacity of our SRP secret key is presented in Section VI. We evaluate the performance of our solution in Section VII. The paper is then concluded in Section VIII.

II. System Model

We assume that there exist two legitimate nodes, named Alice and Bob, trying to secure a communicating link, and that each of them uses OFDM for transmission/reception. In particular, consider an OFDM system where each OFDM symbol consists of N orthogonal subcarriers. After modulating the input serial data streams, a serial to parallel converter converts serial data symbols to parallel streams. N_t pilots, denoted by x_t, are then inserted for the measurement of channel conditions. This results in a vector $X[k]$ for $k = 0, 1, ..., N-1$. $X[k]$ is then used as input to an N-point Inverse Fast Fourier Transform (IFFT). The time domain signal is now:

$$x[n] = \text{IFFT}\left\{X[k]\right\} \qquad\qquad n = 0, 1, 2, \cdots, N-1. \quad (1)$$

A guard interval of length N_d, also known as cyclic prefix, is appended according to:

$$x_f[n] = \begin{cases} x[n+N], & n = -N_d, -N_d+1, \cdots, -1, \\ x[n], & n = 0, 1, \cdots, N-1. \end{cases} \quad (2)$$

$x_f[n]$ is then passed through a parallel to serial converter and digital to analog converter, and it is then transmitted to the other node. The received signal at Alice and Bob is given by:

$$y_f^A[n] = x_f^B[n] \otimes h[n] + w_A[n], \quad (3)$$
$$y_f^B[n] = x_f^A[n] \otimes h[n] + w_B[n], \quad (4)$$

where x_f^B is the transmitted signal from Bob to Alice, x_f^A is the transmitted signal from Alice to Bob, h is a random process that describes the wireless channel between Alice and Bob and w_A and w_B are the additive white Gaussian noise (AWGN) at Alice and Bob's receivers, respectively. Note that the pilots, also known as training signals or reference signal, within x_f^A and x_f^B are identical. The guard interval is then removed from the received signal yielding $y[n] = y_f[n]$ for $n = 0, 1, \cdots, N-1$. $y[n]$ is then passed through an N-point FFT yielding the frequency domain signal $Y[k] = \text{FFT}\{y[n]\}$ $k = 0, 1, ..., N-1$. The pilots, whose locations are already known, are then extracted from $Y[k]$ yielding Y_t, where $t = 1, \cdots, N_t$. Note that the signal exchange between Alice and Bob is performed during the coherence time of the channel.

For simplicity, we estimate the channel through the least squares (LS) estimator in the frequency domain. The LS estimator minimizes the squared error as [20]:

$$\hat{H} = \arg\min ||Y_t - X_t H||. \quad (5)$$

The estimated channel at both Alice and Bob can be given by:

$$\hat{H}_{LS}^A = (X_t)^{-1} Y_t^A, \quad (6)$$
$$\hat{H}_{LS}^B = (X_t)^{-1} Y_t^B, \quad (7)$$

where X_t is the diagonal matrix defined as $X_t = \text{diag}(x_1, \cdots, x_{N_t})$ and Y_t has a dimension of $N_t \times 1$. Since the entries $(x_1, \cdots, x_{N_t})$ are non-zero, the matrix X_t is invertible. The estimated channel at the pilot locations are then interpolated to estimate the channel across the entire OFDM symbol. The estimated channel gains at Alice and Bob $|\hat{H}_{LS}^A|$ and $|\hat{H}_{LS}^B|$ as well as the phases, which are the angles of $\hat{H}_{LS}^A$ and $\hat{H}_{LS}^B$, are the common sources of randomness which are typically used to generate the secret key and from which we will derive our SRP.

In our adversary model, we assume that an eavesdropper (Eve) can listen to all the exchanged signals between the two legitimate communicating nodes (Alice) and (Bob). Moreover, Eve can estimate the channel between itself and both Alice and Bob. However, Eve can not be within a few wavelengths of either of the two communicating nodes, Alice and Bob, which ensures that her estimated channel between either of them is independent of that between Alice and Bob. In addition, we assume that Eve is a passive adversary, that is not interested in active attacks.

III. Review of Existing Techniques

The most typical steps employed in SKG techniques are presented in Figure 1. In the first step, Alice and Bob exchange beacon signals, from which each estimates the physical layer characteristics that are used as common sources of randomness. In our case, they estimate the channel gain and phase. The channel measurements are then quantized and converted into a stream of bits. This is followed by an information reconciliation as well as a privacy amplification step to be applied on the two streams of bits.

It is well known that the major advantage of uniform quantization is its ease of implementation. However, increasing the number of quantization bits dramatically degrades the performance of the SKG technique. This is due to the quantization error that increases as more quantization levels are added. This leads to a higher BMR between Alice and Bob. In [8], an encoding algorithm is proposed to address this issue where each uniformly quantized value is encoded with multiple values. It is worth noting that a lower BMR after the quantization step leads to a longer key, which increases the SKG technique's efficiency.

Another popular technique to address the BMR is presented in [9], [17]. Their solution is based on level crossing of the estimated channel gain. They first use the statistics of the estimated channel gain to compute two thresholds (C_+ and C_-). Alice determines the locations of her estimated channel gain, which is stored in a vector G_A, that are above C_+ or below C_- for a duration of m successive estimates. Alice then sends those locations to Bob. Bob then compares his estimated channel gain at the locations in G_A to determine G_B at which the estimated channel gains are higher than C_+ or below C_- for a duration of $m-1$ successive estimates. Bob's estimated locations G_B, which is a subset of G_A are sent back to Alice. The channel estimates at the locations G_B at both Alice and Bob are then quantized and converted into bitstreams. The main difference between the level crossing technique and the traditional techniques is that the information reconciliation step is performed before the quantization and the bitstream generation. This leads to a much better BMR but at the cost of much shorter key length. To address this drawback, the authors of [9], [17] have proposed to increase the propping rate of the channel.

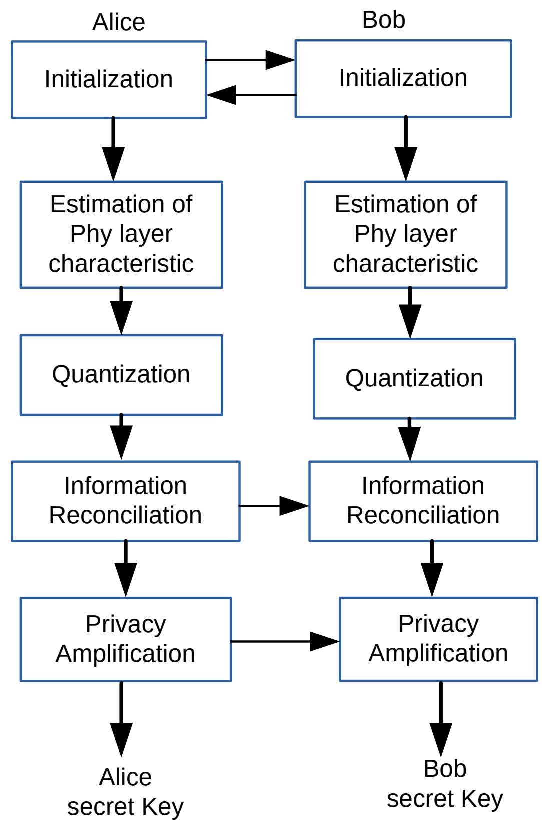

Fig. 1: Typical steps for SKG.

IV. Proposed SRP Technique

We propose a simple SKG technique exploiting, *indirectly*, the estimated channel. Our technique can be applied on the channel gain only, phase only or a combination of the channel gain and phase as we will show later. It is assumed that Alice and Bob have exchanged signals within the coherence time of the channel. They then have estimated the channel using (6) and (7). They applied an interpolation technique on their channel estimates at the pilot locations to estimate the channel across the entire OFDM symbol. It is worth noting that our technique is not exclusive to OFDM systems, rather it can be applied on the estimated channel in presence of any other system.

A. *Creating a secondary random process*

Due to the reciprocity of the channel, the channel estimates at Alice and Bob, $\hat{H}^A_{LS}$ and $\hat{H}^B_{LS}$, are supposed to be identical. However, because of the AWGN added at the two receivers, $\hat{H}^A_{LS}$ and $\hat{H}^B_{LS}$ are not identical. To address the BMR issue explained earlier, we generate a *secondary* random process from the channel estimates. This SRP is then used as common source of randomness to generate the secret key. The steps, which can be applied on the estimated channel gain or phase, are reported below. The steps are reported for the channel gain and apply similarly to the phase. For simplicity, we limit the description below to the case in which they are applied to the estimated channel gain. The steps to generate our SRP are:

1) Both Alice and Bob use their estimated channel gain to estimate a threshold (γ_g) as:

$$\gamma_g^A = \mathbb{E}[|\hat{H}_{LS}^A|] + \alpha\,\mathrm{std}(|\hat{H}_{LS}^A|) \quad (8)$$
$$\gamma_g^B = \mathbb{E}[|\hat{H}_{LS}^B|] + \alpha\,\mathrm{std}(|\hat{H}_{LS}^B|), \quad (9)$$

   where $\mathbb{E}[.]$ is the mean operation, $\mathrm{std}(.)$ is the standard deviation operation and α is a design parameter $\in [-1 : 1]$.
2) Both Alice and Bob compare their channel gain, recursively to the preset thresholds γ_g^A and γ_g^B, respectively.
3) If the channel estimate is higher than the preset threshold, the location, i.e, the index (x-axis) is stored in a vector S initialized to all zeros. Alice and Bob estimate their vectors as S_g^A and S_g^B.
4) Alice and Bob then estimate the moving difference of their estimated locations J_g^A and J_g^B for channel gain, which are computed as:

$$J_g^A[i] = S_g^A[i+1] - S_g^A[i], \qquad i = 1, ..., N-1, \quad (10)$$
$$J_g^B[i] = S_g^B[i+1] - S_g^B[i], \qquad i = 1, ..., N-1. \quad (11)$$

A flow chart of the SRP of the channel gain is presented in Figure 2 for Alice. The realizations in the vectors J_g^A and J_g^B constitute the entries of our *secondary* random process. In other words, we have created two SRPs, one for the channel gain and another for the channel phase. These SRPs are considered our new common sources of randomness which are then used by Alice and Bob to generate the secret key. In Section VII, we provide an example of our SRP. Alice and Bob can use SRP extracted from channel gain only, channel phase only or a combination of the two for the SKG.

B. Quantization

Now that we have our *secondary* common source of randomness estimated at both Alice and Bob, the following step is to quantize it into a bit stream suitable for SKG. We use, as stated earlier, the most popular technique for quantization, which is the uniform quantization. In uniform quantization, the spaces along the x-axis are equal. Similarly, the spaces along the y-axis, which represents the estimated *secondary* common source of randomness, are uniformly distributed. When using n_q bits as the number of quantization bits, there will exist 2^{n_q} levels to quantize the *secondary* common sources of randomness. The quantized decimal valued are then converted into bits.

C. Information Reconciliation and Privacy Amplification

The produced streams of bits at Alice and Bob will have some discrepancy, particularly at low SNR levels. This is due to several causes that include interference, AWGN, hardware limitations and quantization error. An information reconciliation technique such as the one presented in [21] will be used to reduce the discrepancy. Both Alice and Bob first permute their bit streams in the same manner. Then they divide the permuted bit stream into small blocks. Alice then sends permutations and parities of each block to Bob. Bob compares the received parity information with the ones he already processed. In case of a parity mismatch, Bob changes his bits in this block to match the received ones.

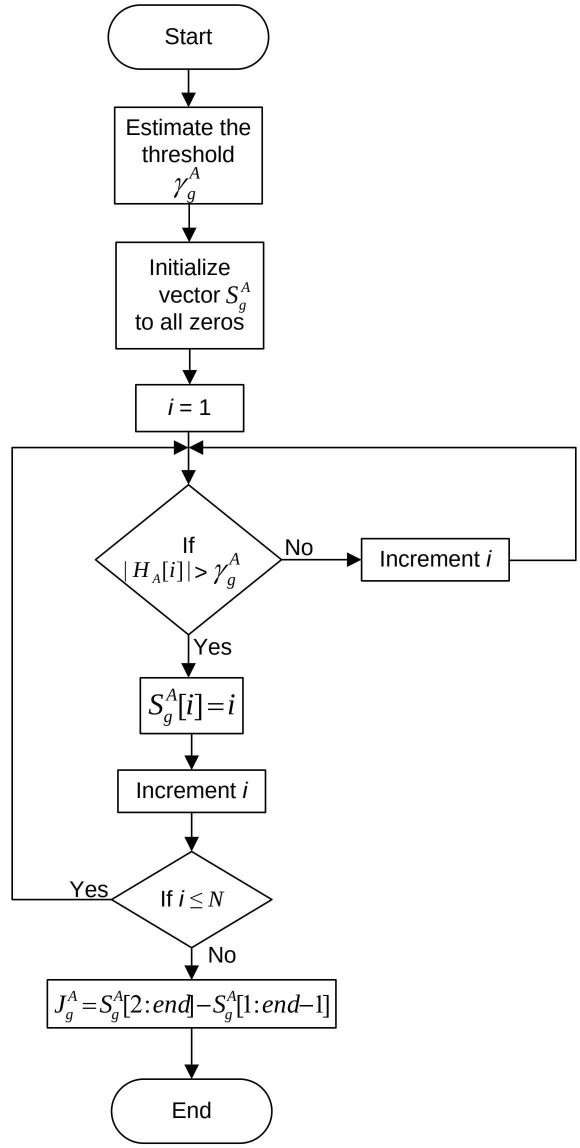

Fig. 2: Flow chart of SRP creation for channel gain at Alice.

The information reconciliation step is designed in such a way that minimizes the information leaked to the adversary. However, some information about the secret key might be leaked during the communication between Alice and Bob at this stage. The eavesdropper can still use this leaked information to guess the rest of the secret key. Privacy amplification addresses this issue by reducing the length of the output bit stream. The generated bit stream is shorter in length but higher in entropy. To do so, both Alice and Bob apply a universal hash function selected randomly from a set of hash functions known by both Alice and Bob. Alice sends the number of the selected hash function to Bob so that Bob can use the same hash function. Further details on exploiting universal hash function for privacy amplifcation is presented in [22].

Our SKG technique is summarized in Algorithm 1 for the channel gain. It is assumed that Alice and Bob have already estimated the channel. Same steps can be applied to the channel phase.

Algorithm 1 SRP SKG Technique for Channel Gain

Step 1: Creating secondary random process
Alice and Bob estimate their thresholds using (8) and (9), respectively.
Both Alice and Bob apply the following steps on $|\hat{H}_{LS}^{A}|$ and $|\hat{H}_{LS}^{B}|$.
for *i = 1: length*($|\hat{H}_{LS}^{A}|$) **do**
  if $|\hat{H}_{LS}| > \gamma_g$ **then**
    $S[i] = i$
  else
    $S[i] = 0$
  end if
end for
Both Alice and Bob estimate $J_g^A = S_g^A[i+1] - S_g^A[i]$ and $J_g^B = S_g^B[i+1] - S_g^B[i]$.
Step 2: Uniform Quantization
Alice and Bob use n_q bits to quantize J_g^A and J_g^B.
Alice and Bob convert their quantized values into bit-streams.
Step 3: Information Reconciliation
Alice and Bob permute the bit streams and divide them into small blocks.
Alice sends the permutation and parities to Bob.
Bob compares the received parity information with his own.
In case of mismatch, Bob corrects his bits accordingly.
Step 4: Privacy Amplification
Alice sends the number of the hash function to Bob.
Alice and Bob apply the hash function to the bit stream.

V. Properties of SRP

In this section, we study the characteristics of our generated SRP. The first step in our SRP creation is to compare the estimated channel gain or phase to a preset threshold. This process can be considered as independent and identically distributed Bernoulli trials. For the channel gain, the success is defined as $|\hat{H}_{LS}[i]| > \gamma_g$ and the failure defined as $|\hat{H}_{LS}[i]| \leq \gamma_g$. The probability of success for the channel gain, p_g, is given by

$$\begin{aligned} p_g &= Pr(|\hat{H}_{LS}[i]| > \gamma_g) \\ &= 1 - q_g \\ &= 1 - Pr(|\hat{H}_{LS}[i]| \leq \gamma_g), \end{aligned} \quad (12)$$

where q_g is the probability of failure. The channel gain follows a Rayleigh distribution with probability density function defined as:

$$f(r) = \frac{r}{\Omega^2} \exp\left(-\frac{r^2}{2\Omega^2}\right), \quad \text{for} \quad r \geq 0 \quad (13)$$

where r is the envelope amplitude of the received signal and $2\Omega^2$ is the average power of multipath signal prior to envelope detection. Hence,

$$p_g = \exp\left(-\frac{\gamma_g^2}{2\Omega^2}\right). \quad (14)$$

Similarly, for channel phase, the success is defined as $\angle\hat{H}_{LS}[i] > \gamma_{ph}$ and the failure defined as $\angle\hat{H}_{LS}[i] \leq \gamma_{ph}$, where γ_{ph} is the threshold for the channel phase. The probability of success for the channel phase is

$$\begin{aligned} p_{ph} &= Pr(\angle\hat{H}_{LS}[i] > \gamma_{ph}) \\ &= 1 - q_{ph} \\ &= 1 - Pr(\angle\hat{H}_{LS}[i] \leq \gamma_{ph}), \end{aligned} \quad (15)$$

where q_{ph} is the probability of failure. The channel phase, θ, follows a uniform distribution with probability density function defined as:

$$f(\theta) = \frac{1}{2\pi}, \quad \text{for} \quad 0 \leq \theta \leq 2\pi \quad (16)$$

Hence,

$$p_{ph} = 1 - \frac{\gamma_{ph}}{2\pi}. \quad (17)$$

Remember that the vectors S_g and S_{ph} are initialized to all zeros. We search for the locations at which the estimated channel gain or phase exceeds the threshold. These locations are the nonzero entries in S_g and S_{ph}. They are estimated as the number of trials, v, needed to achieve u successes. Therefore, these locations, V_g, follow a negative binomial (NB) distribution according to $V_g \sim \mathcal{NB}(u_g, p_g)$ for the channel gain and $V_{ph} \sim \mathcal{NB}(u_{ph}, p_{ph})$ for the channel phase. The probability mass function of V_g is given by:

$$\begin{aligned} l_g(v_g, u_g) &= Pr(V_g = v_g) \\ &= \begin{pmatrix} v_g - 1 \\ u_g - 1 \end{pmatrix} (1-p_g)^{v_g - u_g} p_g^{u_g}. \end{aligned} \quad (18)$$

$l_{ph}(v_{ph}, u_{ph})$ is defined similarly for the channel phase. Thus, the probability of overwriting the initial zero in S_g is given by (18) and the probability that it remains zero is $l'_g(v_g, u_g) = 1 - l_g(v_g, u_g)$. Also $l'_{ph}(v_{ph}, u_{ph})$ is described in the same manner. The entries in the vectors J_g and J_{ph} are the moving differences between each two consecutive entries in S_g and S_{ph}, respectively. Hence, each entry in J_g and J_{ph} has four possibilities as follows. We present the cases for the channel gain only. The four cases for the channel phase are similar with the probabilities assigned to the channel phase vector entries.

- Case 1: the two consecutive entries in S_g are zeros. Consequently, the entry in J_g is zero with probability $l'_g(v_g, u_g)\ l'_g(v_g+1, u_g)$.
- Case 2: the two consecutive entries in S_g are the values of the NB random variables (v_g and $v_g + 1$). Consequently, the entry in J_g is 1 with probability $l_g(v_g, u_g)\ l_g(v_g + 1, u_g + 1)$.
- Case 3: the first (out of the two producing J_g entry) entry is zero and the second is a value of the NB random

variable. Consequently, the entry in J_g is the same value of the NB random variable (v_g) with probability $l'_g(v_g,u_g)\ l_g(v_g+1,u_g+1)$.

- Case 4: the first entry is a value of the NB random variable and the second is zero. Consequently, the entry in J_g is the negative of the value of the NB random variable ($-v_g$) with probability $l_g(v_g,u_g)\ l'_g(v_g+1,u_g)$.

To find a closed form expression for the probability mass function of each entry in J_g, which we denote by $P(J_g[i]=j_g)$, we use the Lagrange interpolating polynomial formula [23]. Lagrange interpolating polynomial method finds the polynomial of degree $\leq n_{lg}-1$ which passes through n_{lg} points $((x_{lg_1},y_{lg_1}),(x_{lg_2},y_{lg_2}),\cdots,(x_{lg_{n_{lg}}},y_{lg_{n_{lg}}}))$. It is defined as

$$D(x_{lg})=\sum_{i_{lg}=1}^{n_{lg}} T_{lg}(x_{lg}), \tag{19}$$

with

$$T_{lg}(x_{lg})=y_{lg_{i_{lg}}}\prod_{\substack{k_{lg}=1\\k_{lg}\neq i_{lg}}}^{n_{lg}}\frac{x_{lg}-x_{lg_{k_{lg}}}}{x_{lg_{i_{lg}}}-x_{lg_{k_{lg}}}}. \tag{20}$$

Using the four cases explained above, the probability mass function of each entry in J_g for $j_g\in\{-v_g,0,1,v_g\}$ can be given by

$$\begin{aligned}
&P(J_g[i]=j_g)\\
&=\frac{l_g(v_g,u_g)l_g(v_g+1,u_g+1)j_g(v_g+j_g)(v_g-j_g)}{(v_g-1)(v_g+1)}\\
&-\frac{l'_g(v_g,u_g)l'_g(v_g+1,u_g)(j_g-1)(v_g+j_g)(v_g-j_g)}{v_g^2}\\
&+\frac{l_g(v_g,u_g)l'_g(v_g+1,u_g)j_g(v_g-j_g)(j_g-1)}{2v_g^2(v_g+1)}\\
&+\frac{l'_g(v_g,u_g)l_g(v_g+1,u_g+1)j_g(v_g+j_g)\ (j_g-1)}{2v_g^2(v_g-1)}.
\end{aligned} \tag{21}$$

The probability mass function of each entry in J_g is zero otherwise. The mean, $\mathbb{E}[J_g[i]]$, is then:

$$\begin{aligned}
\mathbb{E}\left[J_g[i]\right]&=\sum_{j_g} j_g\ P(J_g[i]=j_g)\\
&=l_g(v_g,u_g)\ l_g(v_g+1,u_g+1)\\
&+v_g\ l'_g(v_g,u_g)\ l_g(v_g+1,u_g+1)\\
&-v_g\ l_g(v_g,u_g)\ l'_g(v_g+1,u_g),
\end{aligned} \tag{22}$$

and

$$\begin{aligned}
\mathbb{E}\left[J_g^2[i]\right]&=\sum_{j_g} j_g^2\ P(J_g[i]=j_g)\\
&=l_g(v_g,u_g)\ l_g(v_g+1,u_g+1)\\
&+v_g^2\ l'_g(v_g,u_g)\ l_g(v_g+1,u_g+1)\\
&+v_g^2\ l_g(v_g,u_g)\ l'_g(v_g+1,u_g).
\end{aligned} \tag{23}$$

Hence, the variance of $J_g[i]$ can be given by:

$$\begin{aligned}
\mathrm{var}\left[J_g[i]\right]&=\mathbb{E}\left[J_g^2[i]\right]-\left[\mathbb{E}\left[J_g[i]\right]\right]^2\\
&=l_g(v_g,u_g)\ l_g(v_g+1,u_g+1)\\
&+v_g^2\ l'_g(v_g,u_g)\ l_g(v_g+1,u_g+1)\\
&+v_g^2\ l_g(v_g,u_g)\ l'_g(v_g+1,u_g)\\
&-\bigg(l_g(v_g,u_g)\ l_g(v_g+1,u_g+1)\\
&+v_g\ l'_g(v_g,u_g)\ l_g(v_g+1,u_g+1)\\
&-v_g\ l_g(v_g,u_g)\ l'_g(v_g+1,u_g)\bigg)^2.
\end{aligned} \tag{24}$$

The probability mass function for the channel phase, $P(J_{ph}[i]=j_{ph})$ is defined similarly.

VI. Secret Key Capacity

Since the entries in our generated SRPs are independent and identically distributed (i.i.d.), our secret key rate after the information reconciliation and privacy amplification exhibits the same generic results presented in [24]. The upper and lower bounds for the channel gain SRP are given by [24]:

$$\begin{aligned}
R_g^U(J_g^A[i];J_g^B[i]||J_g^E[i])\leq\min\Big[&I(J_g^A[i];J_g^B[i]),\\
&I(J_g^A[i];J_g^B[i]|J_g^E[i])\Big],
\end{aligned} \tag{25}$$

$$\begin{aligned}
R_g^L(J_g^A[i];J_g^B[i]||J_g^E[i])\geq\max\Big[&I(J_g^B[i];J_g^A[i])-\\
&I(J_g^E[i];J_g^A[i]),I(J_g^A[i];J_g^B[i])\\
&-I(J_g^E[i];J_g^B[i])\Big],
\end{aligned} \tag{26}$$

where $I(J_g^A[i];J_g^B[i])$ is the mutual information between $J_g^A[i]$ and $J_g^B[i]$ and $I(J_g^A[i];J_g^B[i]|J_g^E[i])$ is the mutual information between $J_g^A[i]$ and $J_g^B[i]$ given $J_g^E[i]$ for the eavesdropper, Eve. The supremum of the secret key rate is considered the secret key capacity C_g:

$$\begin{aligned}
C_g&=\max_{P(J_g^A[i])} I(J_g^A[i];J_g^B[i]||J_g^E[i])\\
&\leq\min\Big[\max_{P(J_g^A[i])} I(J_g^A[i];J_g^B[i]),\\
&\max_{P(J_g^A[i])} I(J_g^A[i];J_g^B[i]|J_g^E[i])\Big].
\end{aligned} \tag{27}$$

However, in the definitions above, it was assumed that Eve has access to the primary random process, i.e., channel estimates. In order for Eve to collect correlated channel measurements, she has to be within a half wavelength apart from either Alice or Bob. In other words, Eve has to place herself within a close proximity (typically a few centimeters) of either of them to obtain useful channel estimates, which is very unlikely to occur. Therefore, as in [25], we disregard the feasibility of eavesdropping. Consequently, the secret key capacity for the channel gain SRP can be given by

$$C_g = \lim_{N\to\infty} \frac{1}{N} I\left(J_g^A[i]; J_g^B[i]\right). \tag{28}$$

The mutual information is defined as

$$I\left(J_g^A[i]; J_g^B[i]\right) = \sum_{j_g^A\in[-v_g,0,1,v_g]} \sum_{j_g^B\in[-v_g,0,1,v_g]} \Bigg[P\bigg(J_g^A[i]=j_g^A, J_g^B[i]=j_g^B\bigg) \log\left(\frac{P(J_g^A[i]=j_g^A, J_g^B[i]=j_g^B)}{P(J_g^A[i]=j_g^A)P(J_g^B[i]=j_g^B)}\right)\Bigg], \tag{29}$$

where $P\left(J_g^A[i]=j_g^A, J_g^B[i]=j_g^B\right)$ is the joint probability mass function of $J_g^A[i]$ and $J_g^B[i]$, while $P(J_g^A[i]=j_g^A)$ and $P(J_g^B[i]=j_g^B)$ are the probability mass functions of $J_g^A[i]$ and $J_g^B[i]$, respectively, which are defined by (21). $P\left(J_g^A[i]=j_g^A, J_g^B[i]=j_g^B\right)$ can be given by

$$P\left(J_g^A[i]=j_g^A, J_g^B[i]=j_g^B\right) = P\left(J_g^A[i]=j_g^A | J_g^B[i]=j_g^B\right) P(J_g^B[i]=j_g^B). \tag{30}$$

Since the two vectors $J_g^A[i]$ and $J_g^B[i]$ are highly correlated, the probability that the entry at J_g^B is identical to the entry at J_g^A is high. We denote this probability by p_g^o. It is defined as $p_g^o = P\left(J_g^A[i]=j_g^B | J_g^B[i]=j_g^B\right)$[1]. The probability that an error occurred, i.e., the entry at J_g^B is different from the entry J_g^A is defined as $p_g^e = P\left(J_g^A[i]\neq j_g^B | J_g^B[i]=j_g^B\right)$. The error can happen in two cases. The first case occurs if either one of the entries in S_g^A, which are used to generate the entry J_g^A, is different from its counterpart in S_g^B. We denote this probability by p_g^{e1}. The second case occurs if the two entries in S_g^A are different from their counterparts in S_g^B. We denote this probability by p_g^{e2}. The relation between the three probabilities follow $p_g^o > p_g^{e1} > p_g^{e2}$ at medium and high SNR levels. Based on these probabilities, we define $P\left(J_g^A[i]=j_g^A | J_g^B[i]=j_g^B\right)$ for all possible values of j_g^A and j_g^B in Table I. Similarly, the secret key capacity for the channel phase, C_{ph}, is defined in the same manner with the probabilities p_{ph}^o, p_{ph}^{e1} and p_{ph}^{e2}. We compute the values of both channel gain and phase probabilities in Section VII.

[1]Even if the two entries of S_g^A and S_g^B were different and resulted in $J_g^A[i]=j_g^B|J_g^B[i]=j_g^B$, we still consider that as a success since j_g^B is the value that will be used to generate the secret key and it should be equal at both Alice and Bob. However, we would like to state that having the two entries in S_g^A and S_g^B different and resulting in a success shall constitute a very small percentage of p_g^o because the two vectors S_g^A and S_g^B are highly correlated.

TABLE I: $P\left(J_g^A[i]=j_g^A | J_g^B[i]=j_g^B\right)$

j_g^A \ j_g^B	$-v_g$	0	1	v_g
$-v_g$	p_g^o	p_g^{e1}	p_g^{e1}	p_g^{e2}
0	p_g^{e1}	p_g^o	p_g^{e2}	p_g^{e1}
1	p_g^{e1}	p_g^{e2}	p_g^o	p_g^{e1}
v_g	p_g^{e2}	p_g^{e1}	p_g^{e1}	p_g^o

VII. Performance Evaluation

To evaluate the performance of our technique, we simulate an entire OFDM system and estimate the channel using the LS estimator. Table II summarizes our simulation parameters for the subsequent figures. We simulate the conventional channel gain and phase techniques, level crossing technique, and proposed SRP technique for channel gain only and for channel phase only. Then we obtain the combined SRP by concatenating bitstreams from channel gain and phase SRPs. Our combined vectors are given by

$$J_c^A = [J_g^A[1], J_p^A[1], J_g^A[2], J_p^A[2], \cdots, J_g^A[N], J_p^A[N]], \tag{31}$$

$$J_c^B = [J_g^B[1], J_p^B[1], J_g^B[2], J_p^B[2], \cdots, J_g^B[N], J_p^B[N]]. \tag{32}$$

We first present an example of our generated SRP. To show the effect of our proposed SRP technique on the BMR, we simulate all techniques up to the quantization and bitstream generation step. For a fair comparison, the level crossing technique is simulated without the information reconciliation step. In other words, channel estimates at the locations G_A and G_B are quantized and converted into bitstreams. We plot the BMR for all techniques. We then compute the secret key capacity probabilities for both channel gain and phase SRPs. Afterwards, we estimate the entropy rate of the generated key for our techniques versus existing techniques. The secret key length is then presented.

TABLE II: Simulation parameters

Parameter	Value
No. of subcarriers	1024
No. of FFT point	1024
Subcarrier spacing	15 KHz
Number of pilots	16.7%=171
Cyclic prefix length	25%=256
Modulation scheme	QPSK
Channel type	Rayleigh
Doppler shift	100 Hz
Chan. Estimation	LS
Interpolation type	Linear
α	-0.2
m for Level crossing	4
n_q	8 bits
Number of iterations	10000

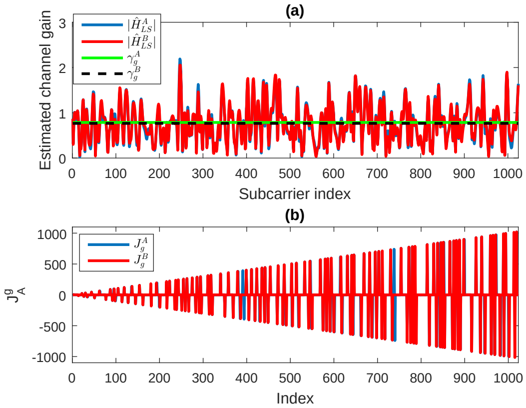

Fig. 3: (a) Estimated channel gain at Alice and Bob with γ_g^A and γ_g^B and (b) our estimated J_A and J_B.

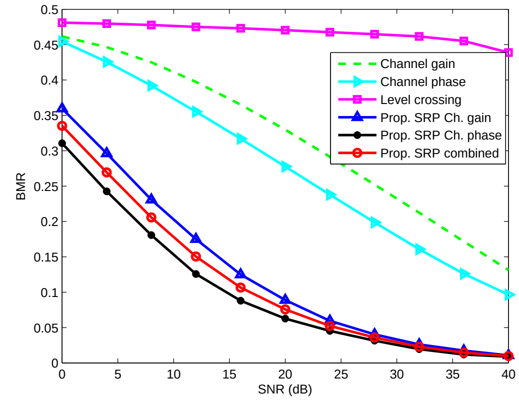

Fig. 4: BMR as a function of SNR for our scheme vs. existing techniques.

A. SRP

In Figure 3-(a), we plot the estimated channel gain at both Alice and Bob, for SNR = 20 dB and the thresholds estimated from (8) and (9). We then follow the steps in Section IV-A to estimate J_g^A and J_g^B and plot them in Figure 3-(b). The estimated channel gain at Alice and Bob is almost identical with some discrepancy in the value of the gain (y-axis) due to the effect of the AWGN. Note that SNR = 20 dB can be considered a moderately high SNR level. The effect of AWGN at lower SNR levels is more severe. On the other hand, since our SRP depends on the locations (x-axis), the effect of AWGN on our channel gain SRP is tolerable. The same conclusion is drawn for the channel phase SRP.

B. BMR

We plot the BMR between the secret keys generated at Alice and Bob for all the techniques in Figure 4. Our proposed SRP techniques drastically improve the BMR achieving a BMR that is ranging from 10-15% at low and high SNR levels to 25% at medium SNR levels less than that of the conventional channel gain and phase. In addition to that, our proposed SRP is achieving a BMR that is ranging from 12% at low SNR levels to 40% at medium and high SNR levels less than that of the level crossing technique. It is worth noting that on average the worst BMR achieved is 0.5 which is equivalent to random guessing. The level crossing technique is performing the worst; achieving the highest BMR, which indicates that the strength of the level crossing algorithm comes from the information reconciliation step. The combined SRP technique achieves a BMR that is average between the SRP channel gain and phase. Also, as expected, as the SNR increases, the BMR for all techniques improves.

C. Probabilities for secret key capacity

We compute the probabilities, p_g^o, p_g^{e1} and p_g^{e2} numerically in Figure 5 for the channel gain SRP and p_{ph}^o, p_{ph}^{e1} and p_{ph}^{e2} in Figure 6 for the channel phase SRP for SNR ranging from 0 to 40 dB. As expected, since $J_g^A[i]$ and $J_g^B[i]$ are highly correlated, p_g^o is much higher than p_g^{e1} and p_g^{e2}, particularly at medium and high SNR levels. As SNR increases, p_g^o increases, while p_g^{e1} and p_g^{e2} decrease. In addition, $p_g^{e1} > p_g^{e2}$ at medium and high SNR levels since it is more likely for one entry in S_g to change rather than the two entries. The same result is obtained for the channel phase. Note that $p_g^o + 2\ p_g^{e1} + p_g^{e2} = 1$. In addition $p_{ph}^o > p_g^o$ at low SNR levels, which suggests exploiting the channel phase SRP over channel gain SRP at low SNR levels should be preferred.

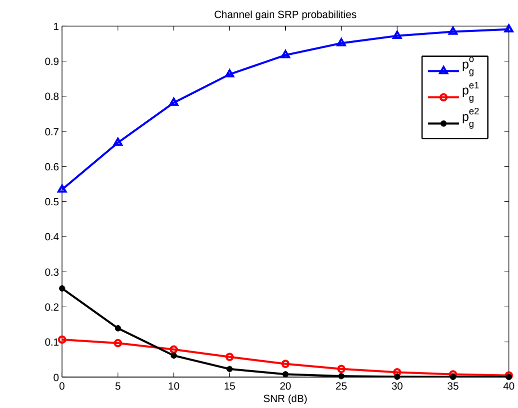

Fig. 5: Probabilities for channel gain SRP.

D. Entropy

Entropy is a measure of the level of randomness of the generated key. For example, for our SRP channel gain, the entropy of a secret key generated from Alice's estimated channel gain is defined as $\mathcal{H}(J_g^A[i]) = \log\left(1/P(J_g^A[i])\right)$. The average entropy is then $\mathbb{E}[\mathcal{H}(J_g^A)]$. As expected from Figure 3-(b), the average entropy of our SRP secret key will be less than that of the channel gain. We plot the achieved entropy rate of all techniques in Figure 7. Our entropy rate for the channel gain is consistent with the results obtained in [26]. Our SRP channel gain and phase exhibit less entropy than all other techniques. To address this drawback, we proposed the combined

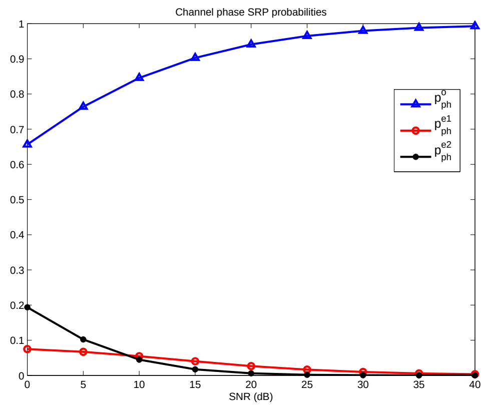

Fig. 6: Probabilities for channel phase SRP.

channel gain and phase SRP algorithm, which improved the entropy rate of the generated secret key. We sacrifice a bit of entropy (15%) to greatly improve the BMR. Also, it is worth nothing that the combined SRP technique does not increase the complexity of the system since both channel gain and phase can be calculated from the channel estimates. In addition to that, it only requires a simple concatenation operation.

The reduction in entropy resulting from our method which is associated with significant reduction in BMR has the advantage that less exchange of messages is needed in the subsequent phases of information reconciliation and privacy amplifications. Knowing that more exchange of messages for information reconciliation results in more side information available to Eve, which in turn will mean less entropy of the final key after privacy amplification [27], we can argue in a qualitative manner that we achieve a performance very close to classical key extraction methods in terms of final key entropy. However, in this work we are not addressing the subsequent phases mentioned above and we stop at showing that BMR is reduced.

E. Key Length

Figure 8 shows the simulated key length of all techniques normalized to the length of the secret key generated through the conventional channel gain technique. Our proposed SRP channel gain and phase is achieving approximately the same key length as of that of the channel gain and phase techniques, while SRP combined is achieving twice that length. On the contrary, the level crossing technique is performing the worst achieving a normalized key length of 30%. This implies that for the level crossing rate technique to achieve a reasonable key length, the frequency of channel propping should increase which decreases the throughput of the system.

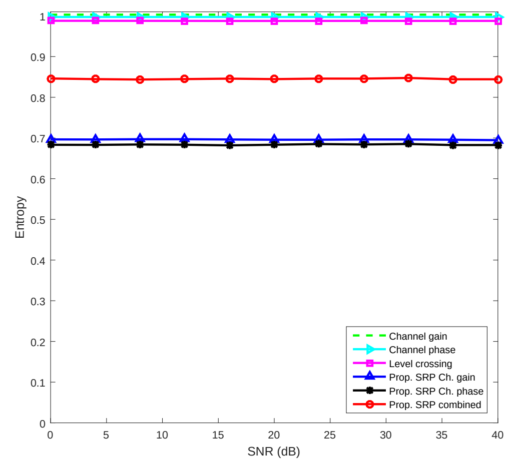

Fig. 7: Entropy as a function of SNR for our scheme vs. existing techniques.

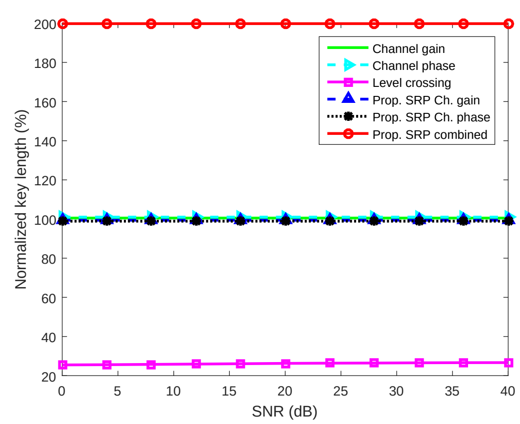

Fig. 8: Normalized key length as a function of SNR for our scheme vs. existing techniques.

VIII. Conclusion

We proposed a simple yet robust technique to extract a secret key from a secondary random process that is derived from the channel estimates. Our SRP technique can be applied on the channel gain only, channel phase only as well as a combination of the two. We derived a closed form expression for the probability mass function of an entry of the SRP vector and simulated our technique using a complete OFDM system. Compared to existing techniques, our SRP solution provided a drastic improvement in the BMR, and achieved comparable entropy and a much longer key length in the case of the combined SRPs. We computed the conditional probabilities used to estimate the secret key capacity for both the channel gain and phase SRP. In addition, our SRP solution is easy to implement and does not increase the complexity of the system.

Acknowledgment

This research was made possible by NPRP 5-559-2-227 grant from the Qatar National Research Fund (a member of The Qatar Foundation). The statements made herein are solely the responsibility of the authors.

References

[1] 3GPP, "The 3rd generation partnership project, url = http://www.3gpp.org."

[2] T. Hwang, C. Yang, G. Wu, S. Li, and G. Li, "Ofdm and its wireless applications: A survey," *Vehicular Technology, IEEE Transactions on*, vol. 58, no. 4, pp. 1673–1694, May 2009.

[3] C. A. Balanis, *Antenna theory: analysis and design*. John Wiley & Sons, 2012.

[4] G. Smith, "A direct derivation of a single-antenna reciprocity relation for the time domain," *Antennas and Propagation, IEEE Transactions on*, vol. 52, no. 6, pp. 1568–1577, 2004.

[5] R. Wilson, D. Tse, and R. Scholtz, "Channel identification: Secret sharing using reciprocity in ultrawideband channels," *Information Forensics and Security, IEEE Transactions on*, vol. 2, no. 3, pp. 364–375, 2007.

[6] B. Azimi-Sadjadi, A. Kiayias, A. Mercado, and B. Yener, "Robust key generation from signal envelopes in wireless networks," in *Proceedings of the 14th ACM Conference on Computer and Communications Security*, ser. CCS '07, 2007, pp. 401–410.

[7] Z. Li, W. Xu, R. Miller, and W. Trappe, "Securing wireless systems via lower layer enforcements," in *Proceedings of the 5th ACM Workshop on Wireless Security*, ser. WiSe '06, 2006, pp. 33–42.

[8] J. Zhang, S. Kasera, and N. Patwari, "Mobility assisted secret key generation using wireless link signatures," in *INFOCOM, 2010 Proceedings IEEE*, 2010, pp. 1–5.

[9] S. Mathur, W. Trappe, N. Mandayam, C. Ye, and A. Reznik, "Radio-telepathy: Extracting a secret key from an unauthenticated wireless channel," in *Proceedings of the 14th ACM International Conference on Mobile Computing and Networking*, ser. MobiCom '08, 2008, pp. 128–139.

[10] A. A. Hassan, W. E. Stark, J. E. Hershey, and S. Chennakeshu, "Cryptographic key agreement for mobile radio," *Digital Signal Processing*, vol. 6, no. 4, pp. 207 – 212, 1996. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1051200496900238

[11] Q. Wang, K. Xu, and K. Ren, "Cooperative secret key generation from phase estimation in narrowband fading channels," *Selected Areas in Communications, IEEE Journal on*, vol. 30, no. 9, pp. 1666–1674, October 2012.

[12] J. Wallace and R. Sharma, "Automatic secret keys from reciprocal mimo wireless channels: Measurement and analysis," *Information Forensics and Security, IEEE Transactions on*, vol. 5, no. 3, pp. 381–392, Sept 2010.

[13] H. Zhou, L. Huie, and L. Lai, "Secret key generation in the two-way relay channel with active attackers," *Information Forensics and Security, IEEE Transactions on*, vol. 9, no. 3, pp. 476–488, March 2014.

[14] O. Gungor, F. Chen, and C. Koksal, "Secret key generation via localization and mobility," *Vehicular Technology, IEEE Transactions on*, vol. 64, no. 6, pp. 2214–2230, June 2015.

[15] A. Badawy, T. Khattab, T. ElFouly, A. Mohamed, and D. Trinchero, "Secret key generation based on channel and distance measurements," in *Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), 2014 6th International Congress on*, Oct 2014, pp. 136–142.

[16] D. T. T. E. Ahmed Badawy, Tamer Khattab and A. Mohamed, "A simple aoa estimation scheme," *CoRR*, vol. abs/1409.5744, 2014. [Online]. Available: http://arxiv.org/abs/1409.5744

[17] C. Ye, S. Mathur, A. Reznik, Y. Shah, W. Trappe, and N. B. Mandayam, "Information-theoretically secret key generation for fading wireless channels," *Information Forensics and Security, IEEE Transactions on*, vol. 5, no. 2, pp. 240–254, June 2010.

[18] N. Patwari, J. Croft, S. Jana, and S. Kasera, "High-rate uncorrelated bit extraction for shared secret key generation from channel measurements," *Mobile Computing, IEEE Transactions on*, vol. 9, no. 1, pp. 17–30, 2010.

[19] K. Zeng, "Physical layer key generation in wireless networks: challenges and opportunities," *Communications Magazine, IEEE*, vol. 53, no. 6, pp. 33–39, June 2015.

[20] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993.

[21] G. Brassard and L. Salvail, "Secret-key reconciliation by public discussion." Springer-Verlag, 1994, pp. 410–423.

[22] C. H. Bennett, G. Brassard, C. Crepeau, and U. M. Maurer, "Generalized privacy amplification," *IEEE Transactions on Information Theory*, vol. 41, no. 6, pp. 1915–1923, Nov 1995.

[23] E. Sli and D. F. Mayers, *An Introduction to Numerical Analysis*. Cambridge University Press, 2003, cambridge Books Online. [Online]. Available: http://dx.doi.org/10.1017/CBO9780511801181

[24] U. Maurer, "Secret key agreement by public discussion from common information," *Information Theory, IEEE Transactions on*, vol. 39, no. 3, pp. 733–742, 1993.

[25] Y. Liu, S. Draper, and A. Sayeed, "Exploiting channel diversity in secret key generation from multipath fading randomness," *Information Forensics and Security, IEEE Transactions on*, vol. 7, no. 5, pp. 1484–1497, Oct 2012.

[26] J. Erin, D. Croft, E. D. Croft, S. K. Kasera, R. rong Chen, C. Furse, and J. Regehr, "Shared secret key establishment using wireless channel measurements," 2011.

[27] C. Cachin and U. M. Maurer, "Linking information reconciliation and privacy amplification," *Journal of Cryptology*, vol. 10, no. 2, pp. 97–110. [Online]. Available: http://dx.doi.org/10.1007/s001459900023