# Fast and Practical Secret Key Extraction by Exploiting Channel Response

Hongbo Liu*, Yang Wang*, Jie Yang†, Yingying Chen*
*Dept. of ECE, Stevens Institute of Technology
Castle Point on Hudson, Hoboken, NJ 07030
*{hliu3, ywang48, yingying.chen}@stevens.edu*
†Dept. of CSE, Oakland University
Rochester, Michigan 48309
*yang@oakland.edu*

***Abstract*—Securing wireless communication remains challenging in dynamic mobile environments due to the shared nature of wireless medium and lacking of fixed key management infrastructures. Generating secret keys using physical layer information thus has drawn much attention to complement traditional cryptographic-based methods. Although recent work has demonstrated that Received Signal Strength (RSS) based secret key extraction is practical, existing RSS-based key generation techniques are largely limited in the rate they generate secret bits and are mainly applicable to mobile wireless networks. In this paper, we show that exploiting the channel response from multiple Orthogonal Frequency-Division Multiplexing (OFDM) subcarriers can provide fine-grained channel information and achieve higher bit generation rate for both static and mobile cases in real-world scenarios. We further develop a Channel Gain Complement (CGC) assisted secret key extraction scheme to cope with channel non-reciprocity encountered in practice. Our extensive experiments using WiFi networks in both indoor as well as outdoor environments demonstrate that our approach can achieve significantly faster secret bit generation rate at $60 \sim 90$bit/packet, and is resilient to malicious attacks identified to be harmful to RSS-based techniques including predictable channel attack and stalking attack.**

## I. Introduction

Exploiting physical layer information for secret key establishment between wireless devices has attracted much attention recently. The basic idea of physical layer based secret key extraction is that, through public information exchange, a pair of wireless devices (e.g., Alice and Bob) can obtain reciprocal observations on the temporal and spatial randomness of the wireless channel between them, which are served as the basis for secret key generation. Different from traditional cryptographic-based methods relying on computational hardness, the essential security of the secret key generated from physical layer information of a radio channel is guaranteed by the fact that the wireless channel between two devices is uncorrelated from other channels [1]–[4]. It thus appears promising that physical layer based secret key generation methods can be deployed as alternatives of existing encryption methods for mobile wireless devices with limited resources or without key management infrastructures (e.g., peer-to-peer association, neighborhood devices changing frequently).

Table I summarizes the practical secret key extraction methods. Existing implementations mainly use Received Signal Strength (RSS) and channel response extracted from a single frequency to perform key extraction. And most recent studies focus on improving the secret bit generation rate by exploiting temporal and spatial variations of radio channel [5], multiple antenna diversity [6], and multiple frequencies [7]). However, since both RSS and channel response extracted from a single frequency can only provide coarse-grained information of the radio channel (e.g., each wireless packet can only provide a single RSS value), the current implementations are largely limited in their real-world deployment even with the assistance of multi-bit quantization.

In this work, we take a different view point by exploring fine-grained physical layer information made available from Orthogonal Frequency-Division Multiplexing (OFDM). The channel response from multiple subcarriers of OFDM provides detailed Channel State Information (CSI), which can be utilized to achieve higher secret bit generation rate and make the secret key extraction approaches (based on physical-layer characteristics) more practical. In particular, we show that by using the Intel 5300 WiFi card [10], multiple subcarriers information can be extracted from a single 802.11 wireless packet to provide the diversity of physical layer channel information, indicating the feasibility of using OFDM-enabled CSI to provide a fast and practical way for secret key generation.

The detailed channel information obtained from multiple subcarriers in OFDM is useful, however, utilizing this extracted information to generate secret keys exhibits unique challenges. For example, the CSI obtained by a pair of wireless devices within the coherence time of the channel may hardly be reciprocal due to different electrical characteristics of wireless devices, especially for antenna gain and RF front attenuation. This non-reciprocity component embedded in the CSI measurements prevents us from extracting secret bits with low bit mismatch rate. To address this issue, we propose a novel channel gain complement (CGC) algorithm that can mitigate the CSI disparity between a pair of wireless devices by removing the non-reciprocity component learned from a small number of probe packets. Extensive experimental results in both indoor and outdoor environments have confirmed the effectiveness and efficiency of the proposed CGC assisted secret key extraction algorithm by leveraging the detailed channel information. Specifically, we make the following contributions:

- We investigate the practical application of utilizing CSI to perform secret key extraction by exploiting OFDM subcarriers, which could provide fine-grained channel response information to facilitate significantly higher bit generation rate at $60 \sim 90$bit/packet as compared with many exiting studies such as the popular RSS based

978-1-4673-5946-7/13/$31.00 ©2013 IEEE

| Existing Work | Device | Physical Modality | Technique | BMR | BGR |
|---|---|---|---|---|---|
| Mathur: Mobicom 08 [4] | commercial 802.11a/b/g modem IP | RSS<br>CIR | Level-crossing | $10^{-7}$<br>$10^{-7}$ | < 1 bit/pkt<br>< 1 bit/pkt |
| Jana: Mobicom 09 [8] | Intel 3945ABG 802.11g WiFi card | RSS | Adaptive Secret Bit Generation | $\sim 3\% - 6\%$ | 2-3 bit/pkt |
| Zeng: INFOCOM 10 [6] | Dell e5400 laptops | RSS | Multi-antenna | $0 - 12\%$ | < 1 bit/pkt |
| Patwari: TMC 10 [5] | Crossbow TelosB wireless sensors | RSS | Multi-bit Adaptive Quantization | $0.04\% - 2.2\%$ | 3 bit/pkt |
| Liu: INFOCOM 11 [9] | MICAz sensor motes | RSS | Group Key Extraction | $\sim 3\%$ | 2-4 bit/pkt |

TABLE I
SUMMARY OF EXISTING PRACTICAL SECRET KEY EXTRACTION SYSTEMS (BMR: BIT MISMATCH RATE; BGR: BIT GENERATION RATE).

methods.

- Our experiments in WiFi networks show the feasibility and effectiveness of using OFDM subcarriers for fast key generation in both indoor and outdoor environments.
- To mitigate the non-reciprocity of CSI caused by the disparity of electrical characteristics between different wireless devices in practice, we develop a Channel Gain Complement assisted secret key extraction scheme, which is highly effective in both static and mobile environments.
- By leveraging the detailed channel response information, our proposed approach is resilient to attack scenarios which have been identified harmful to secret key extraction when using RSS, including predictable channel attack and stalking attack.

The rest of the paper is organized as follows. We place our work in the context of related research in Section II. We provide a feasibility study of using fine-grained channel-response information for secret key extraction and present the attack model in Section III. We then present the proposed channel gain complement assisted key extraction scheme and the corresponding analysis in Section IV. In Section V, we describe the experimental methodology and evaluation metrics. We next evaluate the performance of our channel response based secret key extraction approach in Section VI. We discuss the resilience of our approach to two types of attacks in Section VII. Finally, we conclude our work in Section VIII.

## II. RELATED WORK

The randomness of radio channel's physical layer characteristics have been theoretically explored for secret key generation. A number of studies [1], [11], [12] have proposed to use the phase change of received signals to generate secret keys. Sayeed *et. al.* [1] and Wilson *et. al.* [12] exploit the randomness of phase for secret key extraction in OFDM and UWB systems respectively, whereas Wang *et. al.* [11] propose a phase-based scalable and efficient secret key generation scheme. Tope *et. al.* [2] utilize the randomness of received signal's envelope to share the secrecy between two parties. Similarly, secret bits have been extracted from the deep fades of channel gain caused by multipath [3]. All of these investigations are based on theoretical analysis and only provide simulation results.

RSS has been widely used in the work proposing practical secret key generation methods, because it is readily available in existing wireless infrastructures. Previous studies on RSS based methods mainly focus on exploiting temporal and spatial variations of radio channel [4], [5], [8], and multiple antenna diversity [6] for secret bit extraction. Since RSS can only provide coarse-grained channel information, RSS based methods suffer from low secret bit generation rate. Channel response has also been exploited to generate secret keys. For example, Mathur *et. al.* [4] utilize the channel impulse response (CIR) extracted from a single frequency to generate at most one secret bit per second. Recent work using frequency selectivity of channel fading shows the feasibility of generating secret keys in static wireless sensor networks [7], however, they only evaluate the entropy of the secret keys and do no provide the secret bit generation rate in practical environments.

We note that there are discussions questioning the efficiency and security of the secret key generation based on physical layer features of a radio channel [13]–[15]. Although we agree that the wireless nature may reduce the security level of the secret keys generated from radio channel measurements, we argue that such degradation in security is tolerable as long as the key generation method implemented in practice can generate a sufficiently long key in an efficient way.

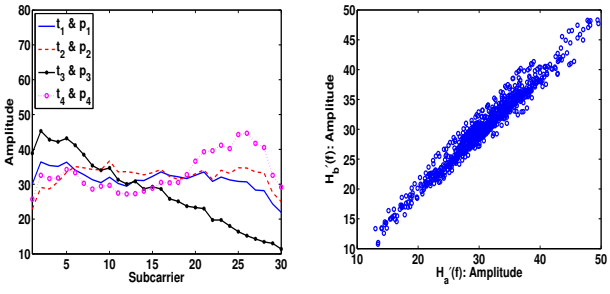
Different from the aforementioned work, in this paper, we seek to exploit the fine-grained channel response information provided by OFDM to improve the practical usage of secret key extraction based on physical layer features. While the feasibility of using channel state information in OFDM system to generate secret keys has been explored [9], [16], these studies have set the theoretic basis and do not provide any practical solution. Our approach investigates the non-reciprocities and unknown fading statistics encountered in real environments. Additionally, our method is resilient to the malicious attacks which are harmful to RSS based secret key generation.

## III. FEASIBILITY STUDY AND ATTACK MODEL

In this section, we first introduce the background of OFDM. We then discuss the feasibility of extracting secret keys by using the CSI measured from OFDM subcarriers. We next present two types of attacks considered in our work.

### A. Preliminaries

OFDM is being commonly used in wireless communication systems, such as IEEE 802.11 a/g/n, WiMAX and 3G LTE, to improve the communication performance by exploiting both space and frequency diversity. In OFDM, a single stream of data is split into multiple parallel streams, each of which is coded and modulated on to a subcarrier. The frequency on each sub-carrier is chosen such that the subcarriers are orthogonal to each other resulting in minimal interference during transmission. For example, the 802.11 a/g/n physical layer is based on OFDM, in which the relatively wideband




(a) (b)

Fig. 1. (a) Example for channel state information of OFDM; (b) Illustration of channel reciprocity.

$20MHz$ 802.11 channel (or carrier) is partitioned into multiple subcarriers, such that each subcarrier can be thought of as a narrowband channel. This inspires us to exploit the channel state information extracted from OFDM subcarriers for secret key generation, which may harvest more randomness and consequently achieve higher secret bit generation rate in practice. Figure 1 (a) depicts the amplitude of channel response across 30 subcarriers at four different time points and positions extracted from an Intel WiFi 5300 card in a laptop [10].

### B. Feasibility Study

A pair of wireless devices, Alice and Bob, are going to establish shared secret key via the reciprocal CSI. Assuming the wireless channel response recorded at Alice and Bob are $H_a(f)$ and $H_b(f)$ respectively, where $f$ represents one particular subcarrier of OFDM. The signals successively received by Alice and Bob can be expressed as:

$$\begin{aligned} \hat{R}_a(f) &= S(f)H_a(f) + Z_a(f) \\ \hat{R}_b(f) &= S(f)H_b(f) + Z_b(f) \end{aligned} \tag{1}$$

where $S(f)$ is the known probe signal at frequency $f$, $Z_a$ and $Z_b$ are independent noise at Alice and Bob respectively. Based on the received signal, Alice and Bob compute (noisy) estimates of $H_a(f)$ and $H_b(f)$:

$$\begin{aligned} H'_a(f) &= H_a(f) + N_a(f) \\ H'_b(f) &= H_b(f) + N_b(f) \end{aligned} \tag{2}$$

where $N_a(f)$ and $N_b(f)$ represent the noise term (due to $Z_a$ and $Z_b$) after processing the function that estimates $H_a(f)$ and $H_b(f)$.

To obtain effective secret keys, the measured channel response information at Alice and Bob should satisfy two requirements: *randomness* and *reciprocity*. Figure 1 (a) shows an example of the measured amplitude of channel response across 30 subcarriers at four different time points and locations, $[t_i, p_i], i = 1, 2, 3$ and 4. We observe that the channel response at different subcarriers is different due to frequency diversity. Further, the channel response of a specific subcarrier is different when measuring at different location and time due to space and time diversity. Thus, the randomness presented in the measured CSI ensures an adversary cannot easilty predict the channel state information between Alice and Bob by merely eavesdropping the wireless communication.

Furthermore, in typical multipath environments, the wireless channel between Alice and Bob produces a time-varying, stochastic mapping between the transmitted and received signals. This mapping is identical on both directions of the wireless link theoretically. If Alice and Bob send probe packets to each other within or similar to the channel's coherence time, the estimation of CSI: $H'_a(f)$ and $H'_b(f)$ showed in equation (2) should be highly correlated in practice. Figure 1 (b) demonstrates a linear relationship between the measured amplitude of $H'_a(f)$ and $H'_b(f)$. We observe that the channel response from Alice and Bob are highly correlated. An increasing value for $H'_a(f)$ results in increasing value for $H'_b(f)$, and vice versa. The measured CSI at Alice and Bob thus provides reciprocal information for secret key generation.

### C. Attack Model

We further consider two types of attacks that have been identified harmful to RSS-based secret key extraction methods in real environments.

**Predictable Channel Attack** [8]**:** When both Alice and Bob are stationary, the wireless channel between them is relatively stable. An adversary Eve, however, can use planned movements to cause desired and predictable changes in the channel measurements between Alice and Bob, referred as predictable channel attack. For example, it can be easily inferred that when the Line-of-Sight (LOS) between Alice and Bob is blocked (e.g., Eve is intentionally crossing the wireless links between Alice and Bob), the transmitted signal may suffer sharp attenuation. We assume that Eve does not possess the prior knowledge of CSI between two arbitrary positions that Alice and Bob reside, since such information requires big efforts to obtain and is environmental sensitive.

**Stalking Attack** [17]**:** In this attack, a passive adversary, called Stalker, follows the trajectory of either Alice or Bob during the secret key establishment and eavesdrops all the legitimate communication between them. The Stalker is able to measure the wireless channels between itself to Alice or Bob when Alice and Bob are exchanging probe packets. Moreover, Stalker also has the knowledge of the secret key extraction algorithm and corresponding parameters for secret key generation. We assume Stalker cannot be too close to either Alice or Bob (i.e., at least half of wavelength away, which is approximately $6cm$ at $2.4GHz$), otherwise it increases the chances to expose himself to be detected.

## IV. Channel Gain Complement Assisted Secret Key Extraction

### A. Motivation

Although theoretically the channel response should be identical on both directions of the wireless link, a practical secret key generation scheme must consider the non-reciprocity presented in the measured CSI due to additive noise, half-duplex nature of the wireless channel, and hardware differences, especially the difference in the antenna gain and RF front attenuation at wireless devices. This is because the non-reciprocity presented in the measurements directly affect the

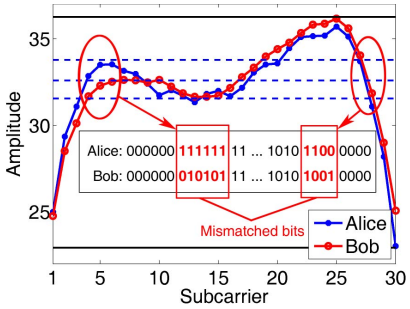



Fig. 2. Channel Response of Alice and Bob before channel gain complement (The correlation between the two CSI measurements is 87.8%).

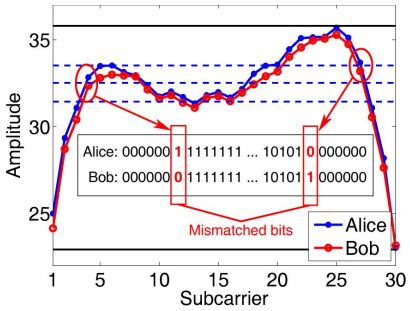



Fig. 3. Channel Response of Alice and Bob after channel gain complement (The correlation between the two CSI measurements is 96.9%).

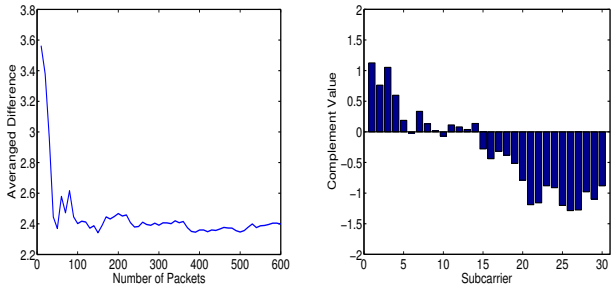
bit mismatch rate (i.e., bits that do not match between two generated keys at Alice and Bob), which is critical to the secret key agreement as a high bit mismatch rate leads to increased number of probe packets exchanging between Alice and Bob or even a failure to establish secret keys.

To investigate the non-reciprocity component in the measured CSI, we expand the equation (2) as:

$$\begin{aligned} H'_a(f) &= H_a(f) + N_a(f) = \widetilde{H}(f) + \rho_a(f) + N_a(f) \\ H'_b(f) &= H_b(f) + N_b(f) = \widetilde{H}(f) + \rho_b(f) + N_b(f) \end{aligned} \tag{3}$$

where $\widetilde{H}(f)$ is the reciprocal component of channel response, and $\rho_a(f)$ and $\rho_b(f)$ are the non-reciprocal components measured at Alice and Bob respectively. We show in the following that while this non-reciprocal component has significant impact on the measured CSI, it is statistically stable for each subcarrier, which is different from the additive noise on the wireless channel.

Figure 2 presents the amplitude of channel response across 30 subcarriers measured at two wireless devices (Alice and Bob) when the measurements are taken within the coherence time of the wireless channel. We observe that although these two shapes of the amplitude across 30 subcarriers measured at Alice and Bob are similar, there are differences existing at each subcarrier. While the similar shape of CSIs provides the basis for exploiting the reciprocal measurements for secret key generation (with 87.8% correlation), the disparity of the channel gain at each subcarrier leads to a certain degree of bit disagreement. This motivates us to develop a scheme to complement the non-reciprocity so that to reduce the bit mismatch rate while maintaining a high speed of secret bit generation rate in practical environments.




(a) (b)

Fig. 4. (a) Averaged differences of one particular subcarrier along the time; (b) complemented value across different subcarriers at one particular time.

To explore how to complement the non-reciprocity, we study the statistical characteristics of the CSI difference at each subcarrier between two wireless devices. The mean and variance of $m$ samples of CSI difference are defined as:

$$\begin{aligned} u_f &= \frac{1}{m}\sum_{i=1}^{m}(H_a^{t_i}(f) - H_b^{t_i}(f)) \\ v_f &= \frac{1}{m-1}\sum_{i=1}^{m}(H_a^{t_i}(f) - H_b^{t_i}(f) - u_f)^2 \end{aligned} \tag{4}$$

where $H_a^{t_i}(f)$ and $H_b^{t_i}(f)$ are the channel state information measured at Alice and Bob at time instant $t_i$, respectively. Figure 4(a) shows the mean value of CSI difference for 600 samples. We find the mean value of the first 100 samples change sharply due to the randomness of channel noise. However, since the additive noise at both Alice and Bob are independently and identically distributed, the impact of additive noise on CSI difference becomes less and less and only the component caused by non-reciprocity remains as more samples are collected. This provides an important insight: the mean values of CSI difference becoming stable after a short time period (e.g., 100 samples) indicating that we can learn the disparity of channel response caused by non-reciprocity (e.g., the difference in the antenna gain and RF front attenuation) between wireless devices. After obtaining the statistical information of CSI difference, we can mitigate the disparity of channel response by complementing the channel gain at each subcarrier to produce a lower bit mismatch rate.

### B. Scheme Overview

In order to reduce the bit mismatch rate while achieving a higher bit generation rate in practice, we propose a channel gain complement (CGC) scheme to learn the non-reciprocity to reduce the disparity of channel response measured at Alice and Bob. The basic idea is to mitigate the non-reciprocity component by learning the channel response from a small number of probe packets. To implement the channel gain complement, we first collect a small number of channel responses from probe packets to learn the non-reciprocity component $\mu_f$ of each subcarrier. We then use $\mu_f$ to mitigate the impact of non-reciprocity component to achieve a low bit mismatch rate while maintaining high bit generation rate when using multi-level quantization method.

### C. Secret Key Extraction Flow

The steps that incorporate CGC method in secrete key extraction mainly include: non-reciprocity learning, channel gain complement, filtering, and quantization. Each step is presented in details in the following.

**Non-reciprocity Learning:** To learn the non-reciprocity, we need to measure the CSI on both directions of the wireless links between Alice and Bob. Since the channel responses measured at Alice and Bob satisfy the principle of reciprocity only if they are measured within the coherence time. To achieve this point, Alice and Bob exchange probe packets, and extract the CSI (i.e., $H'_a(f)$ and $H'_b(f)$) at each OFDM subcarrier from each probe packet. Once both Alice and Bob have collected a certain number of samples of CSI, say $m'$, Bob sends the extracted CSI samples together with the corresponding time stamps to Alice. Alice then compares her own time stamps $t_i^a$ of measured channel response with $t_i^b$ for $i = 1, \cdots, N^a$ where $N^a$ is the total number of CSI samples collected by Alice during this short time period to make sure the measured CSI at Alice and Bob are within coherence time. Only the samples with time stamps on both sides satisfying the following requirement are utilized for non-reciprocity learning:

$$\|t_i^a - t_i^b\| < \delta \tag{5}$$

where $\delta$ is the threshold of the coherence time.

Based on equation (3), the difference of non-reciprocal component between Alice and Bob can be obtained as:

$$\begin{aligned} \rho_{a,b}(f) &= H'_a(f) - H'_b(f) \\ &= \widetilde{H}(f) + \rho_a(f) + N_a(f) - (\widetilde{H}(f) + \rho_b(f) + N_b(f)) \\ &= \rho_a(f) - \rho_b(f) + (N_a(f) - N_b(f)) \end{aligned} \tag{6}$$

We assume both $N_a(f)$ and $N_b(f)$ follow Gaussian distribution $N(u, \sigma^2)$, we can then get $\rho_{a,b}(f) \sim N(\rho_a(f) - \rho_b(f), 2\sigma^2)$. Therefore, by averaging a number of the difference of channel state information between $H'_a(f)$ and $H'_b(f)$ over time, we can obtain $\mu_f$ based on equation (4). Thus, the obtained $\mu_f$ is the expected value of $\rho_{a,b}(f)$. Similarly, the variance of the non-reciprocity component $v_f$ can be obtained as well based on equation (4).

**Channel Gain Complement:** After non-reciprocity learning, we can mitigate the non-reciprocity of channel response by deducting $\rho_{a,b}(f)$ from $H'_a(f)$. The channel state information $H'_a(f)$ at Alice is updated as:

$$\begin{aligned} H'_a(f) &= H'_a(f) - \rho_{a,b}(f) = \widetilde{H}(f) + \rho_a(f) - \rho_{a,b}(f) + N_a(f) \\ &= \widetilde{H}(f) + \rho_b(f) + N_a(f) \end{aligned} \tag{7}$$

The estimation of $\rho_{a,b}(f)$ (i.e., $\mu_f$) is called complement value at the subcarrier $f$. Figure 4(b) shows an example on the complemented values at each subcarrier.

After the probe packet exchanging for non-reciprocity learning, Alice and Bob start to exchange probe packets for key extraction (i.e., the CSI extracted from these packets will be used to quantize for secret bit generation). Then, Alice and Bob also exchange the time stamp of probe packets they received. Similarly, only the CSI measurements with time stamps satisfying the requirement in equation (5) are used for secret key extraction between Alice and Bob. The channel state information $H_a^{t'_i}(f_j)$ measured from these probe packets at Alice at time instant $t'_i$ will be complemented by using the learned non-reciprocity component $u_f$ (i.e., complement value):

$$H_a^{t'_i}(f) = H_a^{t'_i}(f) - u_f \tag{8}$$

After channel complement, the quantization method is applied to the updated channel response $H_a^{t'_i}(f)$. Figure 3 shows the amplitude of channel response measured at Alice and Bob after channel gain complement. Comparing Figure 3 to Figure 2, we observe the correlation of the measurements between Alice and Bob improved from $87.8\%$ to $96.9\%$ after channel gain complement and the mismatched bits are largely reduced. This indicates that the proposed channel gain complement approach can effectively reduce the bit mismatch rate.

**Filtering:** To further reduce the impact of noise, the moving average based filtering technique is adopted. For example, for a particular subcarrier $f$, given a frequency-window size $w$ (e.g., $w = 3$), the channel response $H_a^{t'_i}(f)$ is obtained as following:

$$H_a^{t'_i}(f) = \sum_{k=f-\lfloor w/2 \rfloor}^{f+\lfloor w/2 \rfloor} H_a^{t'_i}(k) \tag{9}$$

**Quantization:** Since most existing studies only use single dimensional physical layer information (e.g., RSS or phase), the quantization can only be applied in time domain (e.g., RSS sequence or phase sequence). Extracting secret bits based on OFDM subcarriers provides information in frequency domain (i.e., frequency diversity), which enables a new dimension of information to be utilized. We thus explore the quantization level in frequency domain, i.e., quantizing the amplitude of CSI across different subcarriers, to boost the secret bit generation rate.[1]

After channel gain complement, the quantization mismatch between Alice and Bob is resulted from the remaining channel noise. Since the noise at Alice and Bob has been assumed to follow Gaussian distribution, the difference of channel response between Alice and Bob, $\rho_H(f)$, can be obtained as:

$$\rho_H(f) = H_a^{t'_i}(f) - H_b^{t'_i}(f) \sim N(0, 2\sigma^2) \tag{10}$$

The variance $\sigma^2$ can be estimated from $v_f$ from the non-reciprocal learning process. And $v_f$ helps to determine the number of quantization levels. If $v_f$ is large, fewer quantization levels should be chosen so that to reduce the quantization mismatch; otherwise, more quantization levels can be used to improve secret bit generation rate. We provide the quantization mismatch analysis in the next subsection for determining the number of quantization levels based on $v_f$, the estimated value of the variance of non-reciprocal component difference.

Once the number of quantization levels $n$ is determined based on $v_f$, the amplitude of $H_a^{t'_i}(f)$ is then quantized into

[1]We also tried to quantize the amplitute of CSI in time domain. However, it results in much higher bit mismatch rate compared with quantizing in frequency domain.

$n$ quantization levels according to the distributions of $H_a^{t_i'}(f)$ across all subcarriers in a single packet. Taking Alice as an example, the value of each quantization level is determined based on the Cumulative Distribution Function (CDF) of $H_a^{t_i'}(f)$, $F(q_k) = P[H_a^{t_i'}(f) < q_k]$. The $k$th quantization level $q_k$ is calculated as:

$$q_k = F^{-1}(\frac{k}{n}), k = 1, \cdots, n-1 \tag{11}$$

and $q_0 = min(H_a^{t_i'}(f))$ and $q_n = max(H_a^{t_i'}(f))$. The $k$th quantization bin is then defined as the interval $[q_{k-1}, q_k]$ for $k = 1, \cdots, n$. $H_a^{t_i'}(f)$ is equally distributed in each quantization bin.

After quantization, Alice and Bob can establish secret keys by following the traditional steps:

- **Encoding:** After quantization, if $H_a^{t_i'}(f)$ falls into the quantization bin $[q_{k-1}, q_k]$, gray coding techniques [18] are employed for extracting $log_2 n$ bits from each subcarrier of $H_a'(f)$.
- **Information Reconciliation:** After the secret bit extraction from channel response, Alice and Bob end up with two bit sequences, $K_a$ and $K_b$, respectively. To reconcile the bit discrepancies resulted from noise, interference, etc., existing information reconciliation techniques, such as error correction code, BCH code [11] and low-density parity-check (LDPC) codes [6], is employed.
- **Privacy Amplification:** Since the information during the reconciliation stage can also be heard by Eve in the public channel, partial information about the secret key between Alice and Bob may be exposed to Eve. To ensure the shared secret key completely unknown to Eve, the technique of privacy amplification [11] can be used to solve this problem.

### *D. Analysis on Quantization Mismatch*

In this subsection, we provide the analysis on the probability of quantization mismatch, which is used to help determining the number of quantization levels based on the estimation of variance $v_f$.

The quantization mismatch between Alice and Bob is resulted from the channel noise as shown in equation (13). Given the quantization levels $[q_0, q_1, q_2, \cdots, q_n]$, where $q_0 = min(H_a^{t_i'}(f))$ and $q_n = max(H_a^{t_i'}(f))$, we assume the channel response measured at Alice, $H_a^{t_i'}(f)$, is located in the $k$th quantization bin $[q_{k-1}, q_k]$. The probability of quantization mismatch can be expressed as:

$$\begin{aligned} P_e^k(H_a^{t_i'}(f)) &= P(H_b^{t_i'}(f) > q_k | H_a^{t_i'}(f) \in [q_{k-1}, q_k]) \\ &\quad + P(H_b^{t_i'}(f) \le q_{k-1} | H_a^{t_i'}(f) \in [q_{k-1}, q_k]) \\ &= 1 - \int_{H_a^{t_i'}(f) - q_{k-1}}^{q_k - H_a^{t_i'}(f)} \frac{1}{2\sqrt{\pi}\sigma} e^{-\frac{x^2}{4\sigma^2}} dx \end{aligned} \tag{12}$$

Across $L$ subcarriers, the averaged probability of quantization mismatch on CSI between Alice and Bob using $n$ quantization levels is:

$$\begin{aligned} P_e &= \frac{1}{L} \sum_{f=1}^{L} \sum_{k=1}^{n} P_e^k(H_a^{t_i'}(f)) P_k(H_a^{t_i'}(f)) \\ &= 1 - \frac{1}{L} \sum_{f=1}^{L} \sum_{k=1}^{n} \int_{H_a^{t_i'}(f) - q_{k-1}}^{q_k - H_a^{t_i'}(f)} \frac{1}{2\sqrt{\pi}\sigma} e^{-\frac{x^2}{4\sigma^2}} dx P_k(H_a^{t_i'}(f)) \end{aligned} \tag{13}$$

where

$$P_k(H_a^{t_i'}(f)) = \begin{cases} 1, & \text{if } H_a^{t_i'}(f) \in [q_{k-1}, q_k] \\ 0, & \text{if } H_a^{t_i'}(f) \notin [q_{k-1}, q_k] \end{cases}$$

The variance $\sigma^2$ is estimated from $v_f$ during the non-reciprocity learning process. The analysis of quantization mismatch probability provides the useful information to choose appropriate quantization levels to satisfy the error correction tolerance during information reconciliation for key extraction.

## V. Experimental Methodology

**Experiment Setup:** We conduct experiments using WiFi network in both typical indoor multi-path and outdoor environments. Two Lenovo laptops, T500 and T61, both equipped with Intel WiFi Link 5300 wireless card are acting as Alice and Bob, respectively. These two laptops run Ubuntu 10.04 LTS with the 2.6.36 kernel and exchange probe packets at the rate of 5 pkt/sec for secret key extraction, which ensures that the sampling interval is larger than the coherence time of the channels at the frequency spectrum of 802.11n. For each packet, we extract CSI for 30 subcarrier groups, which are evenly distributed in the 56 subcarriers of a 20MHz channel [10]. Concurrently, we record the RSS value from each packet.

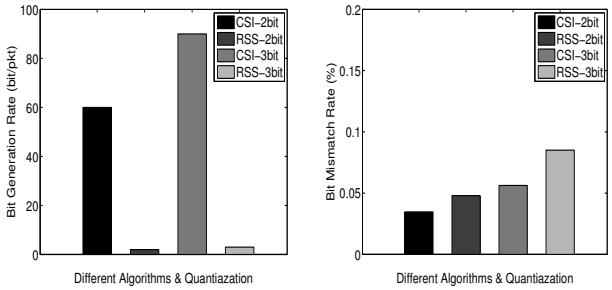
**Testing Scenarios:** We experiment with two different scenarios, i.e., mobile and static, in both indoor and outdoor environments. For mobile scenario, two laptops are moving together at a normal walking speed, which is around 1-2 meters per second with the distance of around 2 meters between them. Whereas in static scenario, both laptops are stationary while people are moving around. The distance between two laptops is about 3 meters. We conduct experiments in two places for outdoor environments: the *tennis court* and the *Babbio square* at Stevens Institute of Technology. Whereas the indoor experiments are conducted in the *student lab* of Burchard building at Stevens Institute of Technology, equipped with furniture and cubicle dividers.

**Metrics:** To evaluate the performance of secret key extraction using channel response, we use the following metrics.

*Bit Generation Rate (BGR):* The bit generation rate is defined as the number of secret bits extracted from each packet, whose measured CSI is used for secret bit quantization.

*Bit Mismatch Rate (BMR):* The bit mismatch rate is defined as the number of bits that do not match between two devices divided by the total number of bits extracted before information reconciliation and privacy amplification.

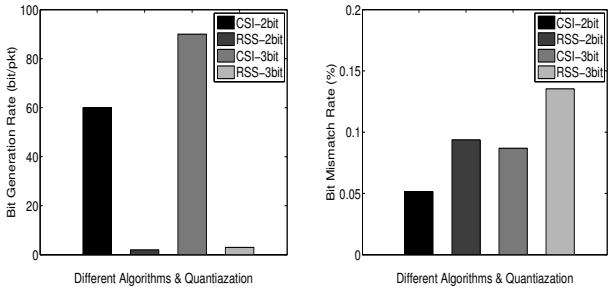
*Randomness:* The standard NIST test suite is employed to measure the randomness of the generated secret bit string.




(a) BGR, indoor




(b) BMR, indoor




(c) BGR: outdoor




(d) BMR: outdoor

Fig. 5. Comparison of bit generation rate and bit mismatch rate between CSI and RSS based methods under different quantization levels and environments.

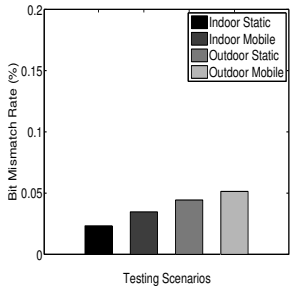
## VI. Performance Evaluation

In this section, we first compare the performance of CSI based key extraction to that of existing studies using RSS [8]. We next study the effectiveness and scalability of the channel gain complement algorithm. Finally, we present the randomness test results on the generated secret keys.

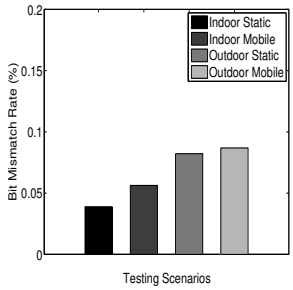
### A. Performance of CSI Based Secret Key Generation

We first compare the performance of the proposed CGC assisted secret key generation using CSI extracted from OFDM subcarriers to that of existing secret key generation method using RSS [8]. Figure 5 shows the experimental results under different quantization levels for mobile scenario in both *student lab* (indoor) and *Babbio square* (outdoor) environments. The number of quantization level used are 4 and 8, which means the amplitude of channel response at each subcarrier will be quantized to 2-bits and 3-bits respectively. From Figure 5(a) and (c), we observe that CSI based method achieves a very high bit generation rate at around 60 bit/packet and 90 bit/packet for 2-bits and 3-bits quantization respectively, whereas the RSS based method only obtains the bit generation rate as low as 2 bit/packet and 3 bit/packet. Figure 5(b) and (d) show that incorporating CGC method for secret key generation achieves much lower bit mismatch rate than that of exsiting method using RSS in both indoor and outdoor environments under different quantization levels. Specifically, our proposed method achieves significant improvements from 27% to 47% in bit mismatch rate. These results show that the proposed CGC assisted secret key generation using CSI is more efficient and reliable than that of RSS based method.

Comparing the results in indoors to those from outdoors, we observe the bit mismatch rate in indoors is lower. In particular, for the 2-bit quantization case shown in Figure 5(b) and (d),




(a) 2-bits quantization




(b) 3-bits quantization

Fig. 6. Bit mismatch rate in mobile and static testing scenarios for both indoors and outdoors.

the bit mismatch rate of the CSI based method is 3.5% in indoors, whereas it is about 5% (i.e., 32% worse) in outdoor environments. Similarly, the bit mismatch rate of RSS based method increases from 4.8% in indoors to 9.4% in outdoors. This is mainly because the multi-path effect of an indoor wireless channel is heavier than that of an outdoor wireless channel. And the heavier multi-path effect leads to lower bit mismatch rate.

### B. Evaluation of Channel Gain Complement

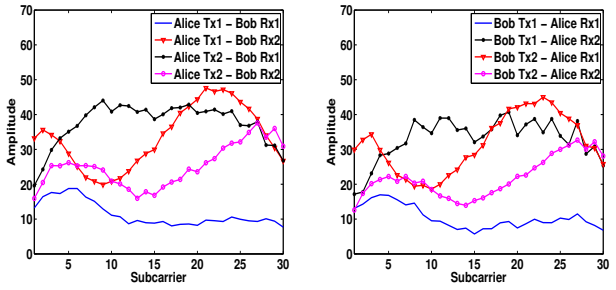
We next compare the performance of CSI based secret key generation method with and without applying channel gain complement. The comparison of bit mismatch rate under different environments is provided in Table II. Overall, the bit mismatch rate has significant improvement after using CGC method. Particularly, in *tennis court*, the bit mismatch rate has 48% improvement; for *Babbio square*, the improvement is about 40%. Whereas in *student lab*, the bit mismatch rate after complement outperforms that before complement over 33%. These results show that the channel gain complement can effectively mitigate non-reciprocity of wireless channel. Thus, it can significantly reduce the bit mismatch rate, which increases the potential for practical deployment.

### C. Scalability Study

Figure 6 compares the bit mismatch rate of our CSI based secret key generation method under static and mobile scenarios in both indoor and outdoor environments. We observe that the bit mismatch rate in mobile scenarios is higher than that in static scenarios. Particularly, in indoors, the bit mismatch rate are 2.3% and 3.5% for static and mobile case respectively, whereas they are 4.4% and 5.1% in outdoor environments. This is because the Doppler effect caused by fast movements in mobile scenarios results in shorter coherence time than in static scenarios [19], which may result in higher bit mismatch rate in mobile scenarios than that in static ones.

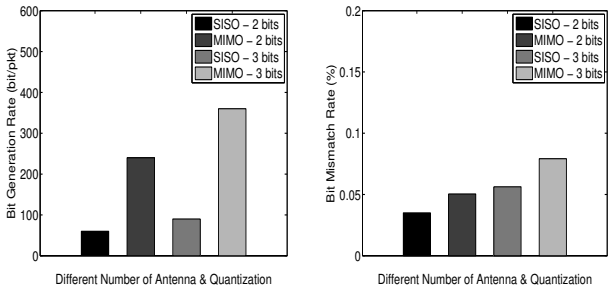
| Bit Mismatch Rate | w/o CGC | CGC | Improvment |
|---|---|---|---|
| Tennis Court (outdoor) | 11.52% | 5.92% | 48.61% |
| Babbio Square (outdoor) | 8.51% | 5.1% | 40.42% |
| Student Lab (indoor) | 5.25% | 3.5% | 33.33% |

TABLE II
Comparison of BMR with and without CGC in different testing scenarios with 2-bit quantization.




(a) CSI measured by Alice (b) CSI measured by Bob

Fig. 7. Channel state information of MIMO from both Alice and Bob.




(a) Bit generation rate (b) Bit mismatch rate

Fig. 8. Performance of CSI based method when exploiting MIMO system.

Furthermore, we study the performance of the CSI based method with Multiple-Input and Multiple-Output (MIMO) system in mobile scenario in indoors. In this experiment, 2 testing laptops with 2 antennas each form a $2 \times 2$ MIMO system. Therefore, we can observe four CSIs from different transmit-receiver antenna pairs between Alice and Bob as illustrated in Figure 7. By leveraging the antenna diversity from a $2 \times 2$ MIMO system, the bit generation rate can achieve four times over that in a Single-Input and Single-Output (SISO) system. Particularly, we can achieve the secret bit generation rates of 240 bit/pkt and 360 bit/pkt for 2-bits and 3-bits quantization respectively as presented in Figure 8. In the meanwhile, the MIMO system has slightly higher bit mismatch rate than SISO system due to the radio interference between integrated antennas of laptops. We note that even with the increased interference in MIMO system, with CGC algorithm, the bit mismatch rate is still comparable to that of using RSS.

### D. Randomness

To ensure that the secret key generated is substantially random, the standard randomness test suite from NIST [20] is employed to verify the effectiveness of the secret bits extracted after secret key reconciliation and privacy amplification [11]. Since the bit length generated from our experiments should meet the recommended size of the NIST tests, we run 8 NIST tests and calculate their p-values. The results of these tests for 4 different experimental scenarios: A) indoor static, B) indoor mobile, C) outdoor static, and D) outdoor mobile, are listed in Table III. All the cases pass the test, and have the p-value much larger than 0.01, which is the threshold to pass the test.

| Test | A | B | C | D |
|---|---|---|---|---|
| Freq. | 0.396 | 0.887 | 0.777 | 0.571 |
| Block Freq. | 0.868 | 0.787 | 0.765 | 0.858 |
| Cum. sums (Fwd) | 0.405 | 0.834 | 0.973 | 0.704 |
| Cum. sums (Rev) | 0.574 | 0.939 | 0.892 | 0.892 |
| Runs | 0.846 | 0.555 | 0.319 | 0.508 |
| Longest run of 1s | 0.572 | 0.701 | 0.906 | 0.8 |
| FFT | 0.516 | 0.516 | 0.516 | 0.516 |
| Approx. Entropy | 0.766 | 0.898 | 0.825 | 0.882 |
| Serial | 0.498<br>0.817 | 0.498<br>0.183 | 0.498<br>0.498 | 0.9<br>0.965 |

TABLE III
NIST STATISTICAL TEST SUITE RESULTS IN DIFFERENT SCENARIOS

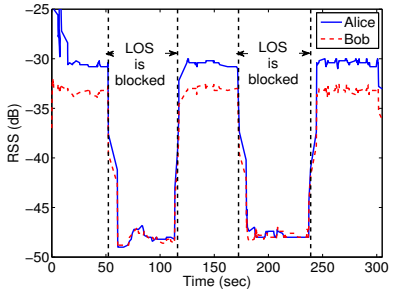



Fig. 9. RSS measurements when an intermediate object moving between Alice and Bob.

## VII. Resilience to Attacks on Secret Key Extraction

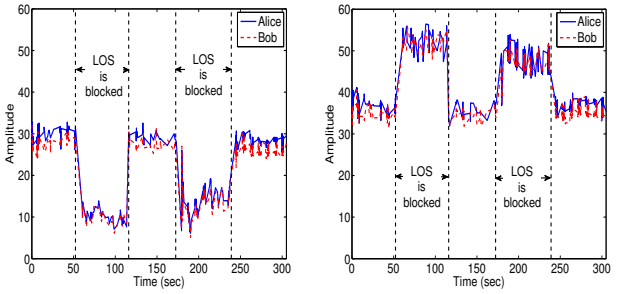
In this section, we show that the proposed secret key generation using CSI measured from OFDM subcarriers is resilient to attack scenarios which have been identified harmful for secret key extraction when using RSS, including predictable channel attack and stalking attack.

### A. Coping with Predictable Channel Attack

Received Signal Strength is usually dominated by the Line-of-Sight (LOS) signal. The attacker Eve can deploy planned movements to block the LOS between Alice and Bob such that the secret key extracted from the RSS measurements with desired changes becomes predictable when both Alice and Bob are stationary. In fact, the attacker can even predict the RSS changes by just observing arbitrary objects blocking the LOS between Alice and Bob.

We show this kind of attack by experimenting with two stationary laptops acting as Alice and Bob. The separation between Alice and Bob is about 3 meters, and the intermediate object periodically blocks the LOS of Alice and Bob for 60 seconds. The collected RSS readings at Alice and Bob is shown in Figure 9. It is obvious that the attacker can predict the changes of RSS measurements of Alice and Bob by observing intermediate objects blocking their LOS, that is, when the LOS between Alice and Bob is blocked, the RSS value drops significantly; when LOS is clear, the RSS value rises.

The advantage of using CSI based key extraction is that the channel response of different subcarriers within CSI does not follow the same trend as RSS does when the LOS is blocked. We show the channel response of two subcarriers, $f = 5$ and 25, for illustration when conducting the same experiment as using RSS. As observed in Figure 10, different subcarriers respond differently when the LOS is blocked, i.e., the channel response of subcarrier $f = 5$ drops significantly




(a) Subcarrier $f$=5 (b) Subcarrier $f$=25

Fig. 10. CSI measurements of subcarriers $f$ = 5 and 25 when an intermediate object moving between Alice and Bob.

similar to that of RSS readings, however, the channel response of subcarriers $f = 25$ increases when the LOS is blocked. It is thus much more difficult for the attacker to predict the fluctuation of CSI at each subcarrier. Consequently, the attacker cannot extract identical secret bits as Alice and Bob by performing predictable channel attack.

### B. Coping with Stalking Attack

RSS is a measurement of the power presented in a received radio signal, and it hides fine-grained information about the wireless signal received at the receiver. On the other hand, CSI includes channel response across multiple OFDM subcarriers, and provides rich information on the wireless channel. Therefore, it is much harder for an adversary to observe the same readings when using CSI than that of using RSS under the stalking attack, in which the attacker follows the trajectory of either Alice or Bob and eavesdrops all the communication.

We experiment with one laptop playing as the stalker following the trajectory of Alice and measuring the CSI of the wireless channel between itself and Bob. Based on the CSI of the wireless channel between the stalker and Bob, the stalker generates secret keys using the same key extraction algorithm and parameters. As shown in figure 11, the bit mismatch rate between the secret keys generated by Alice and the stalker is around 50%, which shows that the stalker's generated key is roughly a random guess.

## VIII. Conclusion

In this paper, we show that it is practical to exploit channel state information (CSI) measured from OFDM subcarriers for secret key extraction. The CSI of wireless channel provides fine-grained channel response information and facilitates high bit generation rate as compared with exiting studies, for example, the popular Received Signal Strength (RSS) based methods. To reduce the bit mismatch rate, we propose the Channel Gain Complement (CGC) assisted secret key extraction scheme to mitigate the non-reciprocity of CSI caused by the disparity of electrical characteristics between different wireless devices in practice. To evaluate the proposed approach, we conduct extensive experiments using WiFi devices under static and mobile scenarios in both indoor and outdoor environments. Results from real implementation show that our approach has significantly faster secret key generation rate, and the CGC method improves the bit mismatch rate significantly.

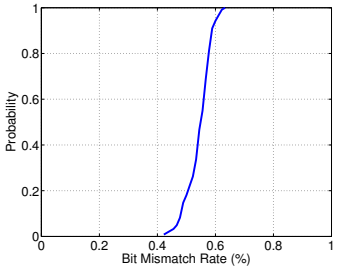



Fig. 11. Bit mismatch rate between Stalker and Alice: the median bit mismatch rate is around 50% with range from around 42% to 62%.

Specifically, while past work generates up to 4bit/packet, our method achieves a secret bit generate rate of 60-90bit/packet. In addition, we show that our proposed method is resilient to the attacks that are harmful to existing secret key extraction methods based on RSS, including predictable channel attack and stalking attack.

## References

[1] A. Sayeed and A. Perrig, "Secure wireless communications: Secret keys through multipath," in *IEEE ICASSP*, 2008.
[2] M. Tope and J. McEachen, "Unconditionally secure communications over fading channels," in *IEEE MILCOM*, 2001.
[3] B. Azimi-Sadjadi and et. al., "Robust key generation from signal envelopes in wireless networks," in *ACM CCS*, 2007.
[4] S. Mathur, W. Trappe, N. Mandayam, C. Ye, and A. Reznik, "Radio-telepathy: extracting a secret key from an unauthenticated wireless channel," in *ACM MobiCom*, 2008.
[5] N. Patwari, J. Croft, S. Jana, and S. Kasera, "High-rate uncorrelated bit extraction for shared secret key generation from channel measurements," *IEEE Transactions on Mobile Computing*, 2009.
[6] K. Zeng and et. al., "Exploiting multiple-antenna diversity for shared secret key generation in wireless networks," in *IEEE INFOCOM*, 2010.
[7] M. Wilhelm, I. Martinovic, and J. Schmitt, "Secret keys from entangled sensor motes: implementation and analysis," in *ACM Wisec*, 2010.
[8] S. Jana and et. al., "On the effectiveness of secret key extraction from wireless signal strength in real environments," in *ACM MobiCom*, 2009.
[9] Y. Liu, S. Draper, and A. Sayeed, "Exploiting channel diversity in secret key generation from multipath fading randomness," *IEEE Transactions on Information Forensics and Security*, 2012.
[10] D. Halperin and et. al., "Tool release: Gathering 802.11n traces with channel state information," *ACM SIGCOMM CCR*, 2011.
[11] Q. Wang, H. Su, K. Ren, and K. Kim, "Fast and scalable secret key generation exploiting channel phase randomness in wireless networks," in *IEEE INFOCOM*, 2011.
[12] R. Wilson, D. Tse, and R. Scholtz, "Channel identification: Secret sharing using reciprocity in ultrawideband channels," *IEEE Transactions on Information Forensics and Security*, 2007.
[13] N. Döttling and et. al., "Vulnerabilities of wireless key exchange based on channel reciprocity," *Information Security Applications*, 2011.
[14] M. Zafer, D. Agrawal, and M. Srivatsa, "Limitations of generating a secret key using wireless fading under active adversary," *IEEE/ACM Transactions on Networking*, 2012.
[15] M. Edman, A. Kiayias, and B. Yener, "On passive inference attacks against physical-layer key extraction?" in *EUROSEC*, 2011.
[16] T.-H. Chou, S. Draper, and A. Sayeed, "Impact of channel sparsity and correlated eavesdropping on secret key generation from multipath channel randomness," in *IEEE ISIT*, 2010.
[17] H. Liu, J. Yang, Y. Wang, and Y. Chen, "Collaborative secret key extraction leveraging received signal strength in mobile wireless networks," in *IEEE INFOCOM*, 2012.
[18] C. Ye, A. Reznik, and Y. Shah, "Extracting secrecy from jointly gaussian random variables," in *IEEE ISIT*, 2006.
[19] Joint Technical Committee (JTC) on Wireless Access, "Final Report on RF Channel Characterization," *JTC(AIR)/93.09.23-238R2*, 1993.
[20] A. Rukhin and et. al., "A statistical test suite for random and pseudorandom number generators for cryptographic applications," *National Institute of Standards and Technology*, 2001.