Perceiving Accurate CSI Phases with Commodity WiFi Devices

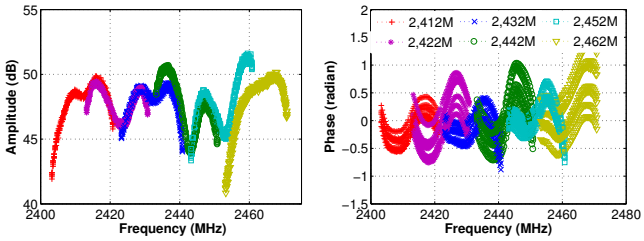
Yiwei Zhuo[1], Hongzi Zhu[1], Hua Xue[1], Shan Chang[2]
[1]Shanghai Jiao Tong University, China
[2]Donghua University, China
{zyw081285, hongzi, howardsid}@sjtu.edu.cn, changshan@dhu.edu.cn

***Abstract*—WiFi technology has gained a wide prevalence for not only wireless communication but also pervasive sensing. A wide variety of emerging applications leverage accurate measurements of the Channel State Information (CSI) information obtained from commodity WiFi devices. Due to hardware imperfection of commodity WiFi devices, the frequency response of internal signal processing circuit is mixed with the real channel frequency response in passband, which makes deriving accurate channel frequency response from CSI measurements a challenging task. In this paper, we identify non-negligible non-linear CSI phase errors and report that IQ imbalance is the root source of non-linear CSI phase errors. We conduct intensive analysis on the characteristics of such non-linear errors and find that such errors are prevalent among various WiFi devices. Furthermore, they are rather stable along time and the received signal strength indication (RSSI) but sensitive to frequency bands used between a transmission pair. Based on these key observations, we propose new calibration methods to compensate both non-linear and linear CSI phase errors. We demonstrate the efficacy of the proposed methods by applying them in CSI splicing. Results of extensive real-world experiments indicate that accurate CSI phase measurements can significantly improve the performance of splicing and the stability of the derived power delay profiles (PDPs).**

Index Terms*—*Channel State Information (CSI); non-linear phase errors; rotation phase error; empirical study; CSI splicing

I. Introduction

Ubiquitous WiFi technology has fostered a broad range of applications beyond a vehicle for communication. In recent years, fast conceptualization and continuous revolution of myriad emerging applications, e.g., seeing through-walls [1], gesture recognition [2, 17, 22], line-of-sight (LOS) identification [4, 18], indoor localization [6–10], detecting movements of an object [5, 11, 12], secure communication [9, 10], continuously revolutionize the horizon [14]. Such applications rely heavily on accurate measurements of the Channel State Information (CSI), which refers to the channel properties such as channel frequency responses of a communication link in a special frequency band. This information describes how a signal propagates from the transmitter to the receiver and represents the combined effect of, for example, scattering, fading, and power decay with distance. Theoretically, the frequency domain responses can also be transformed lossless to the time domain Power Delay Profile (PDP) through IFFT (Inverse Fast Fourier Transform). A PDP fully characterizes a multipath channel, and has been recently used for various motion- or location-based applications. As a result, accurate CSI measurements are of great significance to tremendous applications.



(a) (b)

Fig. 1: (a) raw CSI amplitudes obtained from six 20MHz 802.11n bands in a typical indoor environment; (b) the corresponding raw CSI phases.

To obtain a CSI, commodity WiFi network interface cards (NICs) such as Intel 5300 and Atheors AR9380 can be easily used. Deriving accurate CSIs directly from such NIC readings, however, is challenging as the obtained CSI measurements describe not only channel properties in passband but also the signal processing circuit properties in baseband. For example, Figure 1 illustrates both amplitude and phase errors in raw CSIs measured in six 20MHz 802.11n bands in a typical indoor environment using Atheors AR9380 NICs. Previous studies [6, 8, 9, 15, 16] have pointed out the following sources of CSI measurement errors due to hardware imperfection in the wireless signal processing, including power control uncertainty, packet detection delay (PDD), sampling frequency offset (SFO), carrier frequency offset (CFO), random initial phase offset, and phase ambiguity. The impacts of above error sources to CSI measurements are three-fold: 1) power control uncertainty causes a CSI amplitude offset; 2) packet detection delay and SFO, essentially equivalent to a time delay, cause CSI phase rotation errors; 3) the rest would respectively cause an identical CSI phase offset error on each measured sub-carriers. Consequently, these sources can only introduce linear phase errors expressed as a rotation error proportional to the sub-carrier index plus an offset in the measured CSI phases.

Authorized licensed use limited to: NANJING UNIVERSITY OF SCIENCE AND TECHNOLOGY. Downloaded on May 07,2025 at 08:05:16 UTC from IEEE Xplore. Restrictions apply.
978-1-5090-5336-0/17/$31.00 ©2017 IEEE

According to previous work [12], the CSI amplitude offsets in individual bands can be easily removed by averaging the sufficient number of CSI measurements obtained within the channel coherence time. As for CSI phase linear errors, several state-of-the-art strategies have been proposed. For example, a linear transform on the raw CSI phase can be conducted [5, 18], in the way that the mean of phases on all sub-carriers is forced to zero, and the phase slope between the first sub-carrier and last sub-carrier is forced to zero too. Another example is to search a linear fitting [6, 10] and subtract the fitted linear function from the raw CSI phase. Recent work [16] obtains CSIs from different frequency bands, averages raw CSI phase measures from the same individual frequency band to mitigate the rotation error due to PDD, and search an identical rotation among individual frequency bands to compensate the rotation error due to SFO. All strategies above are based on an assumption that all the notable CSI phase errors except measurement noise are linear. In contrast, it is obvious to see from Figure 1(b) that phases measured on sub-carriers especially for those at both ends of a band are severely distorted in a non-linear way, which suggests there exists an unknown source of non-linear CSI phase errors with commodity WiFi devices.

In this paper, we focus on achieving accurate CSI phase measurements and conduct extensive empirical study using commodity WiFi NICs. In addition to verifying those linear-error sources mentioned above, we find non-linear CSI phase errors across all sub-carriers in all WiFi bands are prevalent in commodity WiFi devices and verify that the root source of such non-linear errors stem from the IQ imbalance issue of direct down conversion receivers. We analyze the characteristics of non-linear CSI phase errors, and have the following two key observations: 1) non-linear CSI phase errors are rather stable over time and different received signal strength indication (RSSI) conditions; 2) such errors are sensitive to different frequency bands used between a transmission pair. Based on these observations, we propose a novel scheme to estimate parameters of our non-linear phase error mode and compensate non-linear CSI phase errors in multipath environments. Moreover, leveraging the insight that, when the channel is stable, the channel phase response for one specific frequency in passband should be the same, we propose to use the method of ordinary least squares on overlapping bands to further remove residual linear phase errors in each band. To verify the efficacy of the proposed calibration schemes for both non-linear and linear CSI phase errors, we conduct a case study and apply our proposed schemes in CSI splicing. We conduct extensive real-world experiments in three different indoor environments with light-of-sight (LOS) and non-light-of-sight (NLOS) conditions. Results demonstrate that accurate CSI phase measurements can be achieved, which significantly improves the performance of CSI splicing and the stability of the derived power delay profiles (PDPs).

In the remainder of this paper, we first introduce some preliminary knowledge about the channel frequency response, the current signal processing design used in commodity WiFi devices and the reported CSI measurement error sources in Section II. Section III elaborates our empirical studies on CSI measurements, where non-linear CSI phase errors are identified and analyzed. We then propose schemes to eliminate both non-linear and linear CSI phase errors in Section IV and evaluate the performance of our scheme in Section V. Section VI presents related work and we conclude in Section VII.

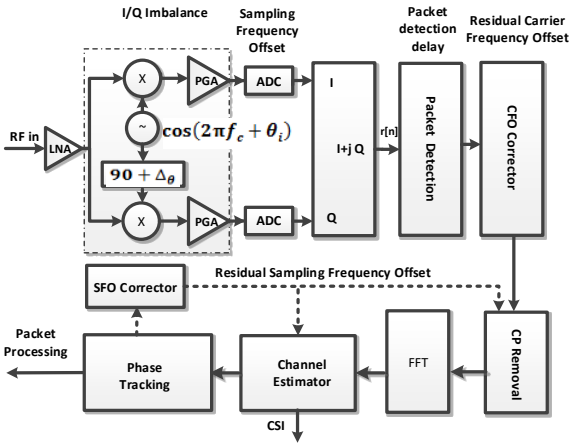


Fig. 2: Illustration of signal processing in 802.11n.

II. Preliminaries

A. Theoretical Foundation

According to [19, 20], the channel frequency response $h(f)$ for multipath scenario can be expressed as:

$$h(f) = \sum_{l=1}^{N} \alpha_l \cdot e^{-j \cdot 2\pi \cdot f \cdot \tau_l} \quad (1)$$

where N is the total number of multipaths, α_l and τ_l represent the attenuation and the propagation delay of the signal through path l, respectively. For each CSI entry, the channel frequency responses for all sub-carriers and all transmission pairs are organized as one CSI matrix. Each frequency response is complex, so it can be expressed with amplitude and phase.

For single direct path scenario, since different sub-carriers in the same frequency band undergo the same time-of-flight, the phase difference between sub-carriers m and n can be expressed as:

$$\Delta_{m,n} = -2\pi \cdot (f_m - f_n) \cdot \tau_1 mod 2\pi \quad (2)$$

where f_m and f_n are the frequency of sub-carriers m and n in passband.

B. Signal Processing at an 802.11 Receiver

A typical WiFi 2.4GHz receiver with direct down conversion architecture is shown in Figure 2. An incoming radio fre-

Authorized licensed use limited to: NANJING UNIVERSITY OF SCIENCE AND TECHNOLOGY. Downloaded on May 07,2025 at 08:05:16 UTC from IEEE Xplore. Restrictions apply.

quency (RF) signal is first amplified by a low noise amplifier (LNA), then mixed with a pair of quadrature sinusoidal signals to perform the so-called quadrature down conversion in order to get the in-phase (I) and the quadrature (Q) baseband signals. After that, a programmable gain filter/amplifiers (PGA) and an Analog-to-Digital convertor (ADC) are applied to the parallel I and Q branches. After sampling, the discrete time domain signal $r[n]$ is passed through the packet detector, which performs energy detection or correlation between $r[n]$ and a pre-defined 802.11 preamble pattern to confirm an incoming packet. Because the existence of CFO will seriously degrade the performance of OFDM, once the packet is detected, the CFO is estimated and corrected to minimize the effects of ICI in the later stages. The channel estimator estimates the instantaneous CSI and the subsequent equalization module (not shown) acts as channel corrector to compensate attenuation and phase errors prior to the packet decoding. Note that, the extracted CSI characterizes not only the frequency response of the external wireless channel in passband, but also the frequency response of the inner circuit mainly in baseband.

C. Reported CSI Measurement Error Sources

Since we aim to sense the external environment with CSIs extracted from commodity WiFi NICs, in this paper, all frequency responses of the inner signal processing circuit are regarded as errors. Besides measurement noise, previous studies [6, 8, 9, 15, 16] have reported the sources of CSI measurement errors as follows.

Power amplifier uncertainty (PAU). Due to the resolution limitation of hardware, for example, 0.5dB for Atheros 9380, the total gain achieved from LNA and PGA cannot perfectly compensate the signal amplitude attenuation to the transmitted power level. The measured CSI amplitude equals to the compensated power level, mixed with a power amplifier uncertainty error, which causes a CSI amplitude offset.

Carrier Frequency Offset (CFO). The central frequencies of a transmission pair cannot be perfectly synchronized. The carrier frequency offset is compensated by the CFO corrector of the receiver, but due to the hardware imperfection, the compensation is incomplete. Signal still carries residual CFO, which leads to a time-varying CSI phase offset across sub-carriers.

Sampling frequency offset (SFO). The sampling frequencies of the transmitter and the receiver exhibit an offset due to non-synchronized clocks, which can cause the received signal after ADC a time shift with respect to the transmitted signal. After the SFO corrector, residual SFO leads to a rotation error. Because clock offsets are relatively stable within a short time (e.g., in the order of minutes [10]), such phase rotation errors are nearly constant.

Packet detection delay (PDD). Packet detection delay stems from energy detection or correlation detection which occurs in digital processing after down conversion and ADC sampling. Packet detection introduces another time shift with respect to the transmitted signal [13, ref21], which leads to packet-varying phase rotation error.

PLL Phase Offset (PPO). The phase-locked loop (PLL) is responsible for generating the center frequency for the transmitter and the receiver, starting at random initial phase [8]. As a result, the CSI phase measurement at the receiver is corrupted by an additional phase offset.

Phase ambiguity (PA). When examining the phase difference between two receiving antennas, recent work [9] validates a so called four-way phase ambiguity existence in Intel 5300 when working on 2.4GHz. Generally speaking, if the phase difference between the first receiving antenna and the second antenna should be $\theta \in (0, \pi/2)$, the four-way phase ambiguity can lead the phase difference to be θ, $\theta+\pi/2$, $\theta-\pi/2$ or $\theta-\pi$. As for Atheros 9380, we similarly discover a two-way phase ambiguity. As a result, phase ambiguity will lead to another phase offset.

From the above known error sources, the measured CSI phases are mainly distorted with various phase ration errors and/or phase offset errors. For a transmission pair, the phase measurement $\phi_(i,k)$ for sub-carrier k in band i can be expressed as

$$\phi_{i,k} = \theta_{i,k} - 2\pi \cdot k \cdot f_s \cdot \delta_i + \beta_i + Z \tag{3}$$

where k ranges from -28 to 28 (index 0 is reserved for carrier frequency) in IEEE 802.11n for 20MHz band width, $\theta_(i,k)$ denotes the true phase, δ_i is the timing offset at the receiver, including time shift due to PDD and SFO, f_s is the sub-carrier spacing between two adjacent sub-carriers (i.e. 312.5KHz), β_i is the total phase offset, and Z is the additive white Gauss measurement noise. Note that, except for Z, other reported phase errors are linear with sub-carrier indexes.

III. Identifying Non-linear CSI Phase Error and Its Root Source

In this section, we conduct empirical study on CSI measurements and describe the non-linear errors and their characteristics with respect to both amplitude and phase.

A. Observing Non-linear CSI Phase Errors

In 802.11n, a channel sounding mechanism is defined, with which a transmitter can trigger CSI estimation at a receiver by setting an appropriate flag in the transmitted packet [23, 24]. We adopt Atheros AR9380 and Intel 5300 NICs, which support 802.11n with 20MHz/40MHz bands at the 2.4GHz/5GHz frequency bands and have three antennas on each NIC. In specific, we setup two pairs of HP desktops running Linux OS with one pair installed with Atheros AR9380 NICs and the other installed with Intel 5300 NICs. With the help of the open source software *hostapd*, we configure one desktop in each pair to acts as AP to transmit packets and the other one as the receiver to extract CSI measurements. We also modify the drivers of both NIC drivers so that receivers can report an estimated CSI to the user space once a packet is received.

Authorized licensed use limited to: NANJING UNIVERSITY OF SCIENCE AND TECHNOLOGY. Downloaded on May 07,2025 at 08:05:16 UTC from IEEE Xplore. Restrictions apply.

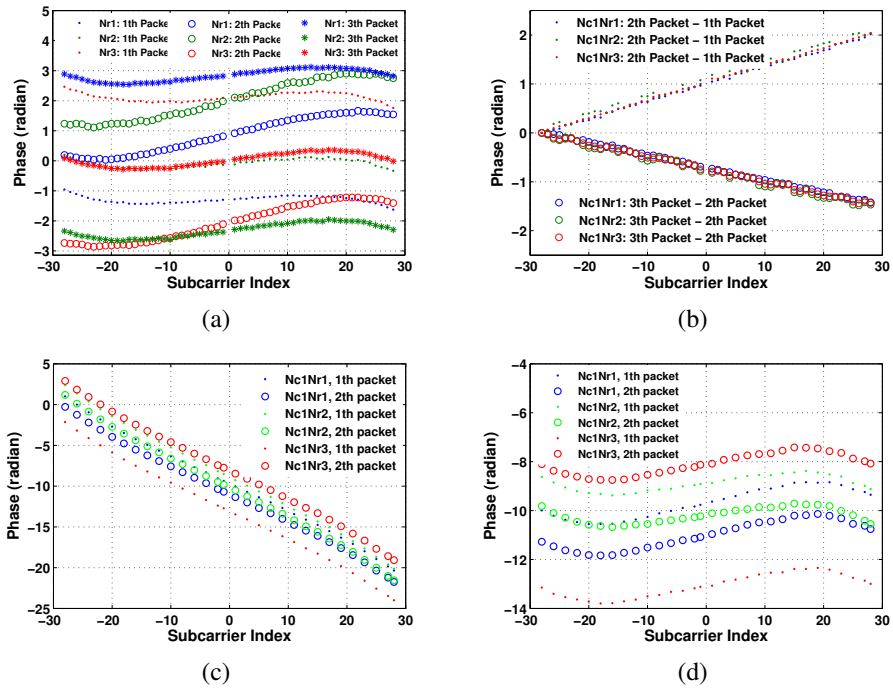


Fig. 3: (a) three groups of unwrapped CSI phases measures from strong LOS scenario with Atheros AR9380; (b) the CSI phase differences of each transmission pair between two consecutive packets, with Nc1Nr1, Nc1Nr2, and Nc1Nr3 denoting transmitting pairs between the first antenna of the transmitter and the first, the second and the third antenna of the receiver, respectively; (c) two unwrapped CSI phase measures between one transmission pair with Intel 5300 NICs; (d) the phase measures after compensating another phase rotation corresponding to a time shift of 200ns.

Packets in all experiments have the minimum payload (to ensure a short transmission delay, i.e., about 0.2ms in our experiment). When working in a 20MHz band with Atheros (Intel) NICs, there are 56 (30) complex numbers in one CSI measurement for each transmission pair.

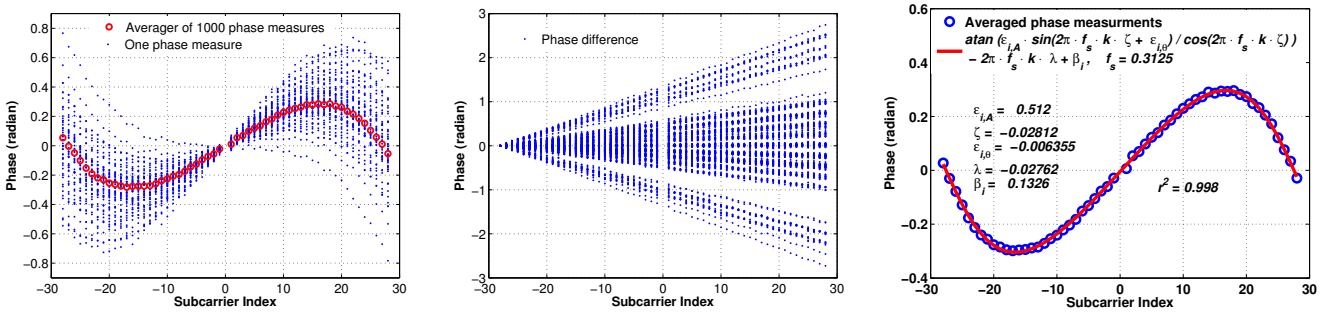
We conduct an experiment in a typical indoor environment with the length and width of the room being 12 meters and 10 meters, respectively. We arrange the transmitter and the receiver in strong line-of-sight (LOS) condition with distance of 0.5 meter, and make the transmitter to transmit with its first antenna, denoted as *Nc1*, with a fixed transmitting power of 5dBm and the receiver to receive with all of its three antennas denoted as *Nr1*, *Nr2* and *Nr3* respectively. We collect CSIs when the environment is stable.

Figure 3(a) illustrates three groups of unwrapped CSI phase measurements for three consecutive packets, with each group containing CSI phases for 1 by 3 transmission pairs. Intuitively, in such strong LOS scenarios, the direct path component dominates all multipath components in the total power of the received signal. According to (2), the ideal phases on different sub-carriers should be almost linear with the sub-carrier indexes. We observe, however, obvious non-linear distortions in all unwrapped phase measurements. We repeat such experiment in an indoor gymnasium with length of 50 meter and width of 30 meter, and get similar results. According to previous work [4, 16], if the wireless channel is stable, the unwrapped phase differences of two consecutive packets for the same transmission pair are almost linear. After removing the phase offset at sub-carrier #-28 from each CSI phase measurement, we calculate the CSI phase differences of each transmission pair between any two consecutive packets using the same CSIs in Figure 3(a) and plot the results in Figure 3(b). It can be clearly seen that the unwrapped phase differences of two consecutive packets for the same transmission pair are almost linear with the sub-carrier index, indicating that the environment is quite stable. In addition, it also suggests that the non-linear CSI phase errors seem to be constant between different measurements.

We repeat the experiment except that we change to use Intel 5300 NICs and draw the unwrapped CSI phase measures of two packets in Figure 3(c). At the first glance, it seems that the CSI phases are pretty linear with sub-carrier indexes. According to previous work [25], the packet detection delay can span hundreds of nanoseconds for Intel 5300. After compensating 4 sampling periods, i.e., 200 ns, we plot the corrected CSI phases in Figure 3(d). It can be seen that the envelopes of phase measures are similar to Figure 3(a).

To further confirm the existence of non-linear CSI errors, we conduct more intensive measurements. In specific, we use a RF cable of 30cm and an attenuator of 50dB to connect the

Authorized licensed use limited to: NANJING UNIVERSITY OF SCIENCE AND TECHNOLOGY. Downloaded on May 07,2025 at 08:05:16 UTC from IEEE Xplore. Restrictions apply.



(a) (b) (c)

Fig. 4: (a) 100 CSI phase measurements in a 20MHz WiFi band at the 2.4GHz frequency band between a transmission pair obtained in a stable and approximate single direct path, with the mean of each measurement removed to zero; (b) the phase differences of 100 phase measures, after removing a particular phase offset respectively; (c) illustration of the least-square regression on the phases of an averaged CSI example

first radio chains of both the transmitter and the receiver. The transmitter sends 1,000 packets within three seconds each time with a fixed transmission power of 15dBm in a 20MHz band with a central frequency of 2,412MHz. We random select 100 CSI measurements, remove the mean from each CSI phase measurement, and plot the unwrapped CSI phases and the phase differences for any two consecutive phase measures in Figure 4(a) and (b), respectively. We have three observations as follows: 1) the envelopes of unwrapped phases are not linear but symmetrical and analogous to some form of trigonometric function; 2) the phase differences of consecutive packets are linear with sub-carrier index, which makes one envelope easy to rotate to another. The default assumption that only notable linear phase error exists cannot hold and an unrevealed non-linear phase error exists, which cannot be mitigated through existing methods. To make matter worse, obviously this non-linear error is orders-of-magnitude higher than the ground truth [1] phase and thus non-negligible. We augment the CSI phase error model as

$$\phi_{i,k} = \theta_{i,k} + \varphi_{i,k} - 2\pi \cdot k \cdot f_s \cdot \delta_i + \beta_i + Z \qquad (4)$$

where $\varphi_{i,k}$ denotes the non-linear error as a function of the sub-carrier index k in band i, with other parameters the same as in (3).

B. Root Source of Non-Linear CSI Errors

Commodity WiFi 2.4GHz receivers normally adopt the direct down conversion architecture as shown in Figure 2. According to previous work [26, 27], there is a universal performance issue, named *IQ imbalance*, in the design of direct down conversion receivers. A direct conversion receiver uses two quadrature sinusoidal signals to perform the quadrature down conversion. This process requires shifting the local oscillator (LO) signal by 90 degrees to produce a quadrature sinusoidal component. When mismatches exist between the gain and phase of the two sinusoidal signals and/or along the two branches of down-conversion mixers, amplifiers, and low-pass filters, the quadrature baseband signals will be corrupted. Once I/Q imbalance exists, after sampling and FFT, the NIC would estimate and report an anamorphic CSI.

When there is only one path between a transmission pair, we assume the averaged phase measurement $\phi_{i,k}$ of subcarrier k in band i as:

$$\phi_{i,k} = atan\left(\epsilon_{i,A} \cdot \frac{sin(2\pi \cdot f_s \cdot k \cdot \zeta + \epsilon_{i,\theta})}{cos(2\pi \cdot f_s \cdot k \cdot \zeta)}\right) - 2\pi \cdot f_s \cdot k \cdot \lambda + \beta_i \quad (5)$$

where $\epsilon_{i,A}$ and $\epsilon_{i,\theta}$ denote the gain mismatch and the phase mismatch for band i respectively due to the IQ imbalance problem, ζ is an unknown timing offset, λ is the equivalent timing delay caused by time-of-flight, PDD and SFO, and β_i is a phase offset error.

To verify the validity of (5), we then apply the least-square regression analysis to the average of the 1,000 CSIs measured via a short RF cable as described in above subsection. The significance of the regression is measured by the coefficient of determination r^2, defined as $r^2 \equiv 1 - \frac{\sum_i (y_i - \bar{y})^2}{\sum_i (y_i - f_i)^2}$, where y_i is the averaged CSI phase with mean $\bar{y}$ and f_i is the modeled/fitted value.

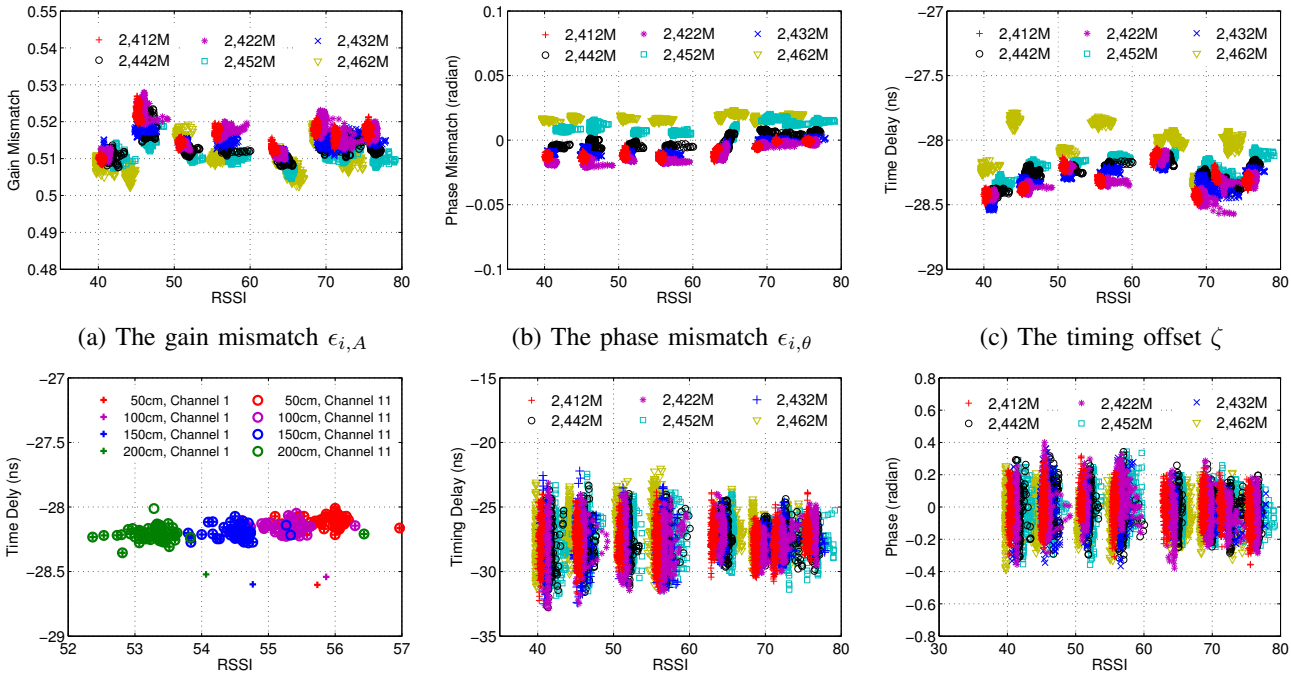
As shown in Figure 4(c), the averaged CSI phase measurements are very well approximated ($r^2 > 0.998$) by the model in (5). We repeat this exercise in all bands and with all NICs and obtain similar results. As a result, we claim that the IQ imbalance problem is the root source of non-linear CSI phase errors.

C. Characteristics of Non-Linear and Linear Phase Errors

We study the characteristics of non-linear CSI phase errors and conduct more intensive CSI measurements. In specific, we use combinations of different attenuators of 30/40/50/60 dB

[1] With a 30cm RF cable, the ground truth of CSI phases is a line with the slope being about 0.002 rad/sub-carrier index

Authorized licensed use limited to: NANJING UNIVERSITY OF SCIENCE AND TECHNOLOGY. Downloaded on May 07,2025 at 08:05:16 UTC from IEEE Xplore. Restrictions apply.



(a) The gain mismatch $\epsilon_{i,A}$ (b) The phase mismatch $\epsilon_{i,\theta}$ (c) The timing offset ζ

(d) The timing offset ζ (e) The equivalent timing delay λ (f) The phase offset error β_i

Fig. 5: Estimates of parameters related to non-linear CSI phase errors, obtained in different RSSI conditions and 6 bands. In addition, the timing offset ζ is further studied with different lengths of RF cables (i.e., time-of-flight) and the result is presented in (d).

and transmitting powers of 15/10/5 dBm to achieve various signal strength. In addition, the transmitter and receiver hop synchronously among six different bands once 1,000 CSIs are collected and averaged on one band. For each configuration, we repeat the data collection for 200 times in a duration of two weeks and for each time we conduct the least-square regression analysis to the averaged CSI to derive all parameters in (5).

Figure 5 (a) to (d) plot the derived parameters related to non-linear CSI phase errors. We have the following four main observations: 1) on a particular band and in a relatively stable environment, the gain mismatch $\epsilon_{i,A}$, the phase mismatch $\epsilon_{i,\theta}$ and the unknown time delay ζ are rather stable along time; 2) on a particular band but in different RSSI conditions, the gain mismatch $\epsilon_{i,A}$, the phase mismatch $\epsilon_{i,\theta}$ and the unknown time delay ζ slightly vary but are still stable as each parameter tends to fluctuate around a horizontal line as RSSI changes; 3) the phase mismatch $\epsilon_{i,\theta}$ is sensitive to the frequency bands as they diverge clearly when measured on different bands but in relatively stable RSSI conditions; 4) from Figure 5(d), it can be seen that the unknown time delay ζ is stable when changing bands and RF cables of different length, which indicates that ζ is independent of frequency bands and the time-of-flight of signal.

Figure 5(e) and (f) plot the derived parameters related to linear CSI phase errors. We have the following four main observations: 1) the timing delay δ_i introduced by PDD and SFO in (4) [2] is independent of frequency bands and RSSI conditions; 2) on a particular band and in a relatively stable RSSI environment, the timing delay δ_i follows a nonzero-mean Gaussian distribution; 3) the variance of the Gaussian distribution of δ_i is large; 4) the phase offset error β_i is analogous to the timing delay δ_i except the mean of its Gaussian distribution is zero.

IV. Perceiving Accurate CSI Phase Measurements

A. Removing Non-linear Phase Errors

From the above study, we have one key observation that non-linear CSI phase errors caused by IQ imbalance are relatively stable over time and various RSSI conditions but sensitive to frequency bands. If the parameters of $\epsilon_{i,A}$, $\epsilon_{i,\theta}$ and ζ are known, non-linear phase errors can be removed. On one hand, if there is only one dominant path between a transmission pair, least-square regression analysis as described in Subsection III-B can be conducted but in real world multipath is inevitable. One straightforward method is to connect the transmission pair via an RF cable but it is infeasible in most cases. On the other hand, if a measured CSI phases can perfectly fit the model in (5), it means that either only one

[2] δ_i can be derived by subtracting the known time-of-flight from the equivalent timing delay λ, when only the direct path exists.

Authorized licensed use limited to: NANJING UNIVERSITY OF SCIENCE AND TECHNOLOGY. Downloaded on May 07,2025 at 08:05:16 UTC from IEEE Xplore. Restrictions apply.

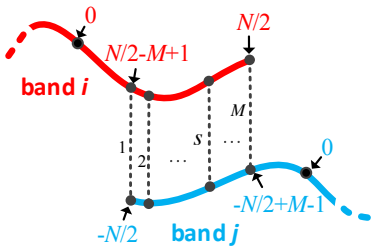


Fig. 6: Illustration of overlapping frequencies in two bands.

dominant path exists (e.g., in a strong LOS and weak multipath environment) or multipath is counteracted.

With this inference, we propose to conduct a *utility test* on raw CSIs and a CSI is said to be positive if it passes the test. In the test, we apply the least-square regression analysis to the phases of a CSI and the condition for this CSI to pass the test is that the significance of the regression measured by the coefficient of determination r^2 is larger than a threshold. In practice, positive CSIs can always be obtained when putting the transmission pair in a strong LOS and weak multipath environment. With sufficient positive CSIs, parameters of $\epsilon_{i,A}$, $\epsilon_{i,\theta}$ and ζ associated with specific bands can be accurately estimated and used to remove future non-linear CSI phase errors.

B. Removing Linear Phase Errors

From previous analysis in Subsection III-C, though the linear (or rotation) CSI phase errors introduced by PDD and SFO are Gaussian distributions, due to large variance, it is infeasible to get the mean by averaging a small number of CSIs measured within the channel coherence time. As a result, each averaged CSI still has its own residual phase rotation error.

In order to eliminate phase rotation errors, we leverage the key insight that, given that the wireless channel is stable, the channel phase response for one specific frequency in passband should be the same even when it is measured from different bands. As illustrated in Figure 6, suppose there are M overlapping subcarriers between band i and band j both of which contain N non-zero indexed subcarriers exposed in the CSI measurements. The $\theta_{i,\frac{N}{2}-M+s}$ and $\theta_{j,-\frac{N}{2}+s-1}$, for $s \in [1, M]$, should be identical. According to (4), the measurement noise Z can be ignored for averaged CSIs and we have

$$\phi_{i,\frac{N}{2}-M+s} - \varphi_{i,\frac{N}{2}-M+s} + 2\pi \cdot (\frac{N}{2} - M + s) \cdot f_s \cdot \delta_i - \beta_i = \phi_{j,-\frac{N}{2}+s-1} - \varphi_{i,-\frac{N}{2}+s-1} + 2\pi \cdot (-\frac{N}{2} + s - 1) \cdot f_s \cdot \delta_j - \beta_j. \quad (6)$$

Given that $\phi_{i,\frac{N}{2}-M+s}$ and $\phi_{j,-\frac{N}{2}+s-1}$ are averaged CSI phases and the non-linear phase errors $\varphi_{i,\frac{N}{2}-M+s}$ and $\varphi_{i,-\frac{N}{2}+s-1}$ can be estimated, there are only 4 unknown parameters, i.e., δ_i, β_i, δ_j, and β_j for M equations. For overdetermined equations (i.e., M is larger than 4 for commodity WiFi devices), we adopt the method of ordinary least squares (OLS) to find an approximate solution. After compensating with both non-linear and linear phase errors, a good estimation of the channel phase response in a band i, i.e., $\theta_{i,k}$, can be obtained by calculating the $\phi_{i,k} - \varphi_{i,k} + 2\pi \cdot k \cdot f_s \cdot \delta_i - \beta_i$ for not only those overlapping subcarriers but also non-overlapping subcarriers.

V. Performance Evaluation

A. Methodology

We use Atheors AR9380 NICs and collect CSIs on six 20MHz bands (i.e., band 1, 3, 5, 7, 9 and 11) at the 2.4GHz frequency bands. We fix the transmitter and change the position of the receiver randomly selected from 10 LOS locations and 10 NLOS locations, respectively, in three different indoor environments, i.e., a 12m×10m laboratory room, a 50m×3m corridor and a gym. At each location, after collecting a batch of 20 CSIs within 4ms on one band, the transmission pair both switch to another band within 5ms. It takes about 50ms [3] to iterate all considered bands and a group of CSIs over all bands can be obtained. We repeat the iteration for 60 times (about 3s) at one location and then move the receiver to the next location. In order to verify the stability of PDPs derived from CSIs, we keep the environment as static as possible. The batches of 20 CSIs collected on each individual bands in each group are first averaged.

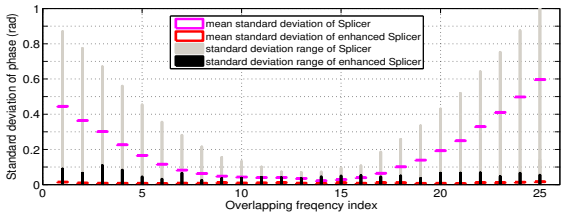
We then conduct the utility tests for all averaged CSIs collected in LOS conditions with the threshold of the coefficient of determination r^2 set to 0.995. For each band, we randomly select 50 positive averaged CSIs to learn the empirical values of $\epsilon_{i,A}$, $\epsilon_{i,\theta}$ and ζ and take the average for each parameter. After that ,the learned $\epsilon_{i,A}$, $\epsilon_{i,\theta}$ and ζ are used to remove non-linear phase errors of averaged CSIs measured on the corresponding band i. Finally, the OLS method is adopted to find the optimal solution of all δ_i and β_i and compensate phase rotation errors for each group of averaged CSIs of six bands. In addition, we take the same procedure to splice CSI amplitudes as introduced in Splicer [16] and derive high-resolution power delay profiles with spliced CSI amplitudes and phases.

We evaluate and compare the performance of Splicer and our sheme using the following two metrics:

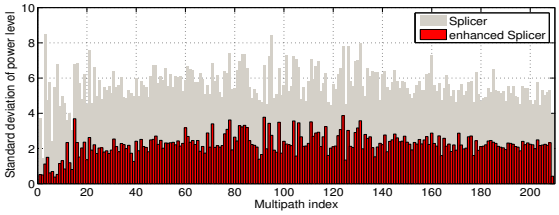
- *Phase differences at overlapping frequencies.* It is known that the phase responses should be identical for subcarriers of the same frequency in two bands. For a group of six corrected CSIs, there are five overlapping frequency bands with each having 25 overlapping subcarriers, i.e., $M = 25$ as in Figure 6. For each $s \in [1, 25]$, we calculate the difference of two corrected phases in each overlapping band of a group before splicing and calculate the standard deviation.
- *Stability of power delay profiles (PDPs) obtained in static environments.* A set of PDPs are accurate and obtained

[3]The channel coherence time when human mobility exists is around 50ms [16].

Authorized licensed use limited to: NANJING UNIVERSITY OF SCIENCE AND TECHNOLOGY. Downloaded on May 07,2025 at 08:05:16 UTC from IEEE Xplore. Restrictions apply.



(a) Statistics of standard deviations of phase difference on overlapping frequencies



(b) Standard deviations of power level in an example location

Fig. 7: Performance of the Splicer and the enhanced Splicer.

in a static environment, they should be very similar if not identical. For each location in our experiment, we calculate the standard deviation of the power levels for each multipath over all 60 PDPs derived from corresponding groups of CSIs.

B. Performance comparison

Figure 7(a) depicts the mean and the range of standard deviations calculated over all groups of CSIs and all locations. Its clear that the mean, the minimal and the maximal deviations of the enhanced Splicer are all much smaller than these of the original Splicer. In addition, it is interesting to see that the original Splicer prefers to align phases at subcarriers in the middle of overlapping bands, leaving subcarriers at both ends badly aligned. In contrast, the enhanced Splicer can align all overlapping frequencies perfectly.

Figure 7(b) depicts the standard deviation of power levels over each multipath in a NLOS condition. It can be seen that the PDPs derived by the enhanced Splicer are more stable than those derived by the original Splicer. Moreover, we also notice that the standard deviations of the first 12 paths is just about 1dB for the enhanced Splicer. On one hand, the small deviation indicates the environment is static. On the other hand, we explain that this 1dB deviation is because of the small amplitude differences (around 1dB) among spliced CSIs.

VI. Related Work

A. Reported CSI phase errors

Besides measurement noise, prior studies also notice that the CSIs reported by WiFi NICs contain phase errors introduced by hardware. Previous studies [10, 15] explicitly point out SFO can cause a phase rotation error, and other studies [5–8, 16, 18, 27] concern this rotation phase error too. Another phase rotation error can be caused by PDD [13, 21, 24], and studies [5, 7, 8, 16, 18] pay much attention to its existence. The authors of work [15] give a good description of the phase offset error caused by CFO, and points out the residual CFO is small after CFO corrector clearly. Recent work [8] observes and tries to mitigate a phase offset error caused by PPO. Another phase offset error due to PA is firstly validated by most recent work [9] in the using of Intel 5300, we observe this error exists similarly in Atheros 9380. However, all above phase errors are linear with sub-carrier indexes.

B. CSI phase calibration

As for CSI phase linear error, to the state of art, there are following strategies: Previous studies including [7, 18] recommend to perform a linear transform on the raw CSI phase. After transforming, both the mean of one phase measure and the phase slope between the first sub-carrier and last sub-carrier are forced to zero. After the transformation, the CSI phase measure can be used as fingerprint for some applications. However, such a brute transform just adds or subtracts another linear error. Studies [5, 6] search a linear fitting and subtract the fitting linear from the raw CSI phase. However, its common to over subtraction. MegaMIMO aims to explicitly correct linear phase errors [27]. However, it requires both nanosecond-level synchronization and the access to the raw signal at PHY layer, which are not available on commodity NICs. Splicer [16] obtains CSIs from different frequency bands, averages raw CSI phase measures for the same individual frequency band expecting to mitigate the rotation error due to the PDD to same level, and cluster an identical rotation to compensate all phase measures from different bands. However, its almost impossible to collect sufficient CSIs within the restriction of strict coherence time to guarantee the residua rotation error to be the same level. In order to remove random initial phase offset, the authors in work [8] propose to collect and process CSIs both from the transmitter and receiver for the same instant. However, even if CSIs can be collected at the instant, there is no guarantee for total phase offset error to be the same. All strategies above are designed for kinds of linear error, and none of them can eliminate rotation phase error well. Meanwhile, they are all based on an assumption that all the notable phase errors except measurement noise are linear with subcarrier indexes.

VII. Conclusion

In this paper, we focus on obtaining accurate CSI phase measurements with commodity WiFi devices. Non-linear phase errors caused by the IQ imbalance issue are identified. In addition, such errors are independent of time and channel conditions but sensitive to frequency bands. We propose two novel schemes to remove non-linear CSI phase errors and residual phase rotation errors in indoor multipath environment. Results of extensive real-world experiments in various indoor environments demonstrate that accurate CSI phase measurements can be achieved, which significantly improves

Authorized licensed use limited to: NANJING UNIVERSITY OF SCIENCE AND TECHNOLOGY. Downloaded on May 07,2025 at 08:05:16 UTC from IEEE Xplore. Restrictions apply.

the performance of CSI splicing and the stability of the derived power delay profiles.

ACKNOWLEDGEMENTS

This research was supported in part by National Natural Science Foundation of China (Grants No. 61472255, 61420106010, 61300199, 61672151).

REFERENCES

[1] F. Adib and D. Katabi, *See Through Walls with Wi-Fi!*, 2013.

[2] L. Sun, S. Sen, D. Koutsonikolas, and K.-H. Kim, "WiDraw: Enabling Hands-free Drawing in the Air on Commodity WiFi Devices," in *Proceedings of ACM MobiCom*, 2015.

[3] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.

[4] Z. Zhou, Z. Yang, C. Wu, W. Sun, and Y. Liu, "LiFi: Line-Of-Sight Identification with WiFi," in *Proceedings of IEEE INFOCOM*, 2014.

[5] Y. Wang, J. Liu, Y. Chen, M. Gruteser, J. Yang, and H. Liu, "E-eyes: Device-free location-oriented activity identification using fine-grained WiFi signatures," in *Proceedings of ACM MobiCom*, 2014.

[6] M. Kotaru, K. Joshi, D. Bharadia, and S. Katti, "SpotFi: Decimeter Level Localization Using WiFi," in *Proceedings of ACM SIGCOMM*, 2015.

[7] S. Sen, B. Radunovic, R. R. Choudhury, and T. Minka, "You are Facing the Mona Lisa: Spot Localization using PHY Layer Information," in *Proceedings of ACM SIGCOMM*, 2012.

[8] D. Vasisht, S. Kumar, and D. Katabi, "Decimeter-Level Localization with a Single WiFi Access Point," in *Proceedings of USENIX NSDI*, 2016.

[9] A. Tzur, O. Amrani, and A. Wool, "Direction Finding of rogue Wi-Fi access points using an off-the-shelf MIMO-COFDM receiver," *Physical Communication*, vol. 17, pp. 149–164, 2015.

[10] S. Jana and S. K. Kasera, "On fast and accurate detection of unauthorized wireless access points using clock skews," *IEEE Transactions on Mobile Computing*, vol. 9, no. 3, pp. 449–462, 2010.

[11] C. Han, K. Wu, Y. Wang, and L. M. Ni, "WiFall: Device-free fall detection by wireless networks," in *Proceedings of IEEE INFOCOM*, 2014.

[12] V. P. G. Jimenez, M.-G. Garcia, F. G. Serrano, and A. G. Armada, "Design and implementation of synchronization and AGC for OFDM-based WLAN receivers," *IEEE Transactions on Consumer Electronics*, vol. 50, no. 4, pp. 1016–1025, 2004.

[13] L. He, L. Fu, L. Zheng, Y. Gu, P. Cheng, J. Chen, and J. Pan, "Esync: An energy synchronized charging protocol for rechargeable wireless sensor networks," in *Proceedings of ACM MobiHoc*, 2014.

[14] Z. Zhou, Z. Yang, C. Wu, L. Shangguan, H. Cai, Y. Liu, and L. M. Ni, "WiFi-Based Indoor Line-of-Sight Identification," *IEEE Transactions on Wireless Communications*, vol. 14, no. 11, pp. 6125–6136, 2015.

[15] J. K. Tan, "An Adaptive Orthogonal Frequency Division Multiplexing Baseband Modem for Wideband Wireless," Ph.D. dissertation, Citeseer, 2006.

[16] Y. Xie, Z. Li, and M. Li, "Precise Power Delay Profiling with Commodity WiFi," in *Proceedings of ACM MobiCom*, 2015.

[17] Q. Pu, S. Gupta, S. Gollakota, and S. Patel, "Whole-Home Gesture Recognition Using Wireless Signals," in *Proceedings of ACM MobiCom*, 2013.

[18] C. Wu, Z. Yang, Z. Zhou, K. Qian, Y. Liu, and M. Liu, "PhaseU: Real-time LOS Identification with WiFi," in *Proceedings of IEEE INFOCOM*, 2015.

[19] T. S. Rappaport *et al.*, *Wireless Communications: principles and practice*. Prentice Hall PTR New Jersey, 1996, vol. 2.

[20] Y. Xie, "Atheros CSI Tool," Website, 2016, http://pdcc.ntu.edu.sg/wands/Atheros/.

[21] J. Gjengset, G. McPhillips, and K. Jamieson, "Arrayphaser:Enabling signal processing on WiFi access points," *RN*, vol. 14, no. 04, p. 04, 2014.

[22] G. Wang, Y. Zou, Z. Zhou, K. Wu, and L. M. Ni, "We can hear you with Wi-Fi!" in *Proceedings of ACM MobiCom*, 2014.

[23] "Ieee 802.11n-2012 standard. 2012," Website, 2016, http://standards.ieee.org/findstds/standard/802.11n-2012.html.

[24] H. Rahul, H. Hassanieh, and D. Katabi, "SourceSync: A Distributed Wireless Architecture for Exploiting Sender Diversity," in *Proceedings of ACM SIGCOMM*, 2010.

[25] K.-Y. Sung and C.-c. Chao, "Estimation and compensation of I/Q imbalance in OFDM direct-conversion receivers," *IEEE Journal in Signal Processing*, vol. 3, no. 3, pp. 438–453, 2009.

[26] M. Petit and A. Springer, "Analysis of a Properness-Based Blind Adaptive IQ. IEEE Wireless Communications," *IEEE Transactions on Wireless Communications*, vol. 15, no. 1, pp. 781–793, 2016.

[27] H. Rahul, S. S. Kumar, and D. Katabi, "MegaMIMO:Scaling Wireless Capacity with User Demands," in *Proceedings of ACM SIGCOMM*, 2012.

Authorized licensed use limited to: NANJING UNIVERSITY OF SCIENCE AND TECHNOLOGY. Downloaded on May 07,2025 at 08:05:16 UTC from IEEE Xplore. Restrictions apply.