UNIVERSITY OF CALIFORNIA
RIVERSIDE

Self-Interference Cancellation in Full-Duplex Radio

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Electrical Engineering

by

Yifan Li

March 2017

Dissertation Committee:

Dr. Yingbo Hua, Chairperson
Dr. Ilya Dumer
Dr. Gang Chen

The Dissertation of Yifan Li is approved:

Committee Chairperson

University of California, Riverside

Acknowledgments

I would like to thank my advisor Dr. Yingbo Hua for his guidance during my Ph.D. study. He is so knowledgeable and kind that whenever I have any concern or question in my research I can always seek for help from him.

I would like to thank my committee members: Dr.Ilya Dumer and Dr.Gang Chen. Without their help, I would not succeed.

I would like to thank Dr. Ertem Tuncel and Dr. Anastasios Mourikis for their assistance during my Ph.D. study

I would like to thank Yiming Ma, Qian Gao, Lei Chen, Armen Gholian and Ali Cirik for their kind encouragement and help as my seniors.

I would like to thank Qiping Zhu, Habib Gharavi, Chaitanya Mauskar and Reza Sohrabi for their company and encourage as my labmates.

I would like to thank Pavle Kirilov for his assistance with lab equipments.

To my parents for all the support.

ABSTRACT OF THE DISSERTATION

Self-Interference Cancellation in Full-Duplex Radio

by

Yifan Li

Doctor of Philosophy, Graduate Program in Electrical Engineering
University of California, Riverside, March 2017
Dr. Yingbo Hua, Chairperson

With increasing demand of mobile wireless communications, the available radio spectrum becomes ever more crowded and hence efficient spectral usage becomes ever more crucial. The full-duplex technology that allows simultaneous reception and transmission at the same frequency can double the spectral efficiency and thus has attracted much attention in the past a few years. To realize a full-duplex radio, the main technical challenge is the self-interference problem. This dissertation focuses on the principles behind several self-interference cancellation (SIC) approaches. Radio frequency (RF) impairments such as in-phase and quadrature (I/Q) imbalances and phase noises in practical RF mixers and demixers are considered in designing and analyzing SIC algorithms. In particular, this dissertation investigates deeper into SIC at the RF frontend by further developing a quadratic-model based method and an affine-model based method. Adaptive algorithms are developed and shown to yield substantial improvement over prior algorithms.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Most of the communication applications used in the market now are half-duplex which means the information transmission works in either different frequency bands (FD), using different time slots (TD) or other division such as Code division multiple access (CDMA) used in 3G communication. However, with more and more mobil devices used, the spectrum becomes very crowed and it is very important to find a more efficient way to achieve the wireless communication.

One answer to this problem is full-duplex (FD) which has attracted lots of attention in the past ten years. Full-duplex communication means transmit and receive the signal in the same time and using the same frequency. If it is realized, the spectrum efficiency of the data transmission can be doubled theoretically. However, implementing the full-duplex faces a big challenge: self-interference. Since the full-duplex radio uses the same time slot and spectrum to transmit and receive data, along with the desired signal, the receive node will also receive an interference which comes from its own transmit node. The real scattered

environment will cause multiple reflection channels. Each channel will introduce a random delay and attenuation to the transmitted signal. Combine together at the receiver and form the self-interference.

Due to the fact that the signal attenuation increases as a longer transmission distance it experiences, the self interference is much stronger than the desired signal. Considering a Wifi system, the typical noise floor is around -90 dBm, a received desired signal with -70 dBm power level is usually enough to achieve successful transmission with 20 dB signal to noise ratio (SNR). However, for a small cell system, the transmission power can be up to 20 dbm. Assume the antenna isolation is 30 dB, without implementing self-interference cancellation the received self-interference can be -10 dBm which is much stronger than the desired signal. The desired signal is totally buried in the self-interference and the system can not decode it. Therefore, to achieve full-duplex communication, self-interference cancellation is needed to reduce the self-interference to the noise floor level first.

In the recent ten years, researchers have proposed multiple self interference cancellation techniques in the literatures, which can be divided into two categories: passive cancellation and active cancellation.

Among the passive self-interference cancellation literatures, the passive cancellation is achieved by antenna placement in [1–4]. It is proposed in [1] that two transmit antennas are placed with distance d and $d + \frac{\lambda}{2}$ away from the receive antenna where λ is the wavelength, to create a cancellation signal with same attenuation but opposite phase. In [2], the author placed 4 antennas as an equilateral triangle, where 3 transmit antennas are

placed on the vertices and the receive antenna is place on the centroid. [3] proposed a two level antenna cancellation method. In [4], three different antenna placement configurations are tested on the actual devices to compare the practical isolation performance.

Apart from the antenna placement, [5–7] discussed how the beamforming weights can be utilized to selective cancel the self-interference at the receiver side. In [8], dual polarized antennas are applied to reduce the self-interference. [9, 10] used the directional antennas to exploit the directional diversity and reduce the self-interference level. [11] proposed a balanced RF front-end circuit based antenna cancellation scheme and [12] proposed a Multi-Reconfigurable Antenna cancellation scheme. [13] is based on antenna subset selection and null-space projection to achieve the spatial suppression.

The largest amount of passive self-interference cancellation reported in the previous literatures is 70 dB. However, it is only valid for some particular scenario and assumption. Also the passive self-interference cancellation is usually accompany with a trade off using more antennas in the system. So, passive cancellation has its own limitation and it is hard to reduce the self-interference to noise floor by only using the passive cancellation method. Therefore, active self-interference cancellation is needed and researchers have come up with lots of schemes of it.

Based on the resource signal type and the location of the cancellation happens, common active cancellation methods can be divided into digital cancellation, analog cancellation and hybrid cancellation which is shown in Fig. 1.1.Digital cancellation starts from the baseband signal, using digital signal processing to generate a copy of the self interference and cancel after the analog-to-digital converter (ADC) at the receiver side. The diagram is

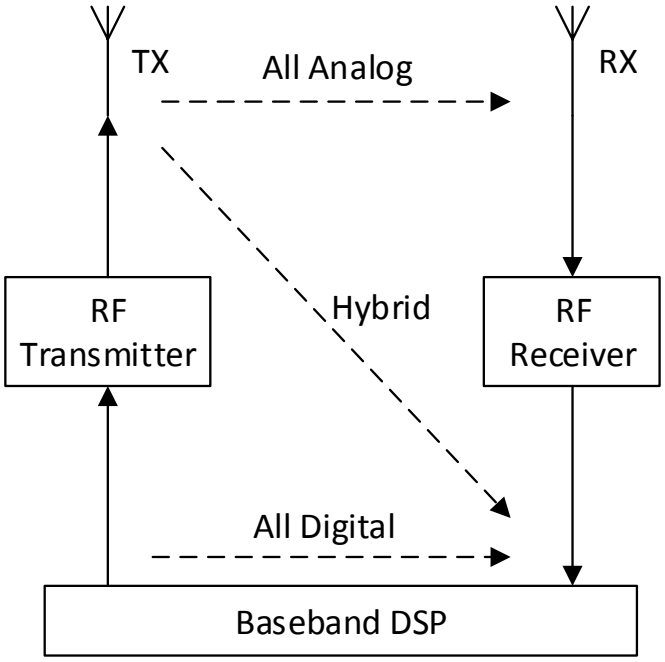


Figure 1.1: Common active self-interference cancellation methods

show in Fig. 1.2. The main idea of the digital cancellation is using the knowledge of the input baseband signal and the received interference to estimate the interference channel. Then based on the estimated channel generate the cancellation sequence.

Multiple ways to estimate the interference channel are proposed. Among which, the most popular one is the least square (LS) estimation [14–18]. In [14], OFDM signal is considered and the least square estimation is operated in the frequency domain. In [15,16], a widely linear model is proposed to handle the impairments introduce by the I/Q imbalance. The authors state their model can recover the linear relationship. [17, 18] discussed both linear and nonlinear signal processing. The nonlinear part in the received self-interference is modeled as higher order terms. Correspondingly, the higher order terms of the baseband input are used in the channel estimation.

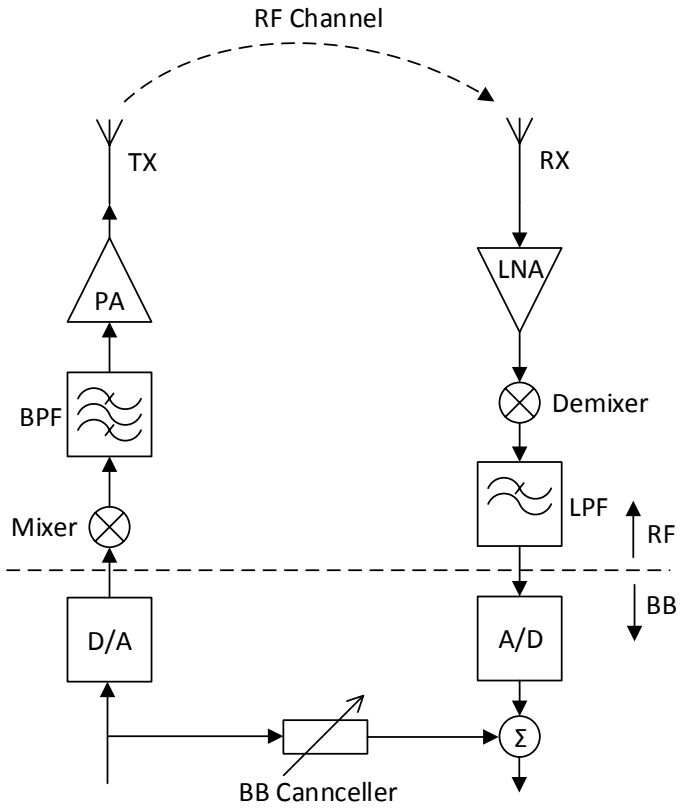


Figure 1.2: Diagram of digital self-interference cancellation

Besides the LS estimation, many other estimation algorithms are also proposed to estimate the interference channel. In [19], sparse signal recovery algorithm is used to do the channel estimation. In [20, 21], the least mean square (LMS) algorithm is applied. Compared to the LS algorithm, the LMS algorithm has a lower computational complexity. However, the upper bound of the LMS estimation is limited by the LS estimation. A recursive least square (RLS) based channel estimation algorithm is introduced in [22] for the passband signal. While in the latest literature, combined with the analog cancellation, the author used the maximum-likelihood (ML) channel estimator in [23] and subspace-based algorithm in [24] to do the channel estimation and handle the remained residual.

The advantage of digital self-interference cancellation method is that it is easy to implement and does not require additional hardware. However, due to the large power of the self interference, the ADC in the receive chain faces the saturation problem and can not provide a proper quantization. Another problem is, the digital SIC also suffers from the nonlinearity in the transmit chain such as I/Q imbalance, phase noise, PA nonlinearity and etc. So as it is shown in [23, 24], the digital cancellation is usually used accompany with analog self-interference cancellation. The analog cancellation cancels the self-interference in the RF domain, reduce the power of the received interference to avoid the ADC saturation, then use the digital cancellation to handle the residual after the analog cancellation and further reduce it to the noise floor.

Similar to the digital cancellation, the hybrid cancellation shown in Fig. 1.3 cancels the self-interference after the ADC at the receiver side. The difference is that the hybrid cancellation uses an auxiliary receive chain to get the digital copy of the transmitted signal and uses the estimated signal as the inputs in the estimation algorithm. Since the linear relationship of the system is modeled from the estimation of the transmitted RF signal to the baseband self-interference. By implementing the hybrid cancellation, the nonlinear components in the transmit chain can be totally lumped into the transmitted RF signal and therefore has no impact on the linear relationship in system model.

In the hybrid cancellation literatures, most of the authors used the LS algorithm to estimate the interference channel [25–29]. In [25–27], the authors considered a full-duplex system that both self-interference (near-end) signal and the desired (far-end) signal are existing, the channel response of the near-end and far-end are estimated jointly by using

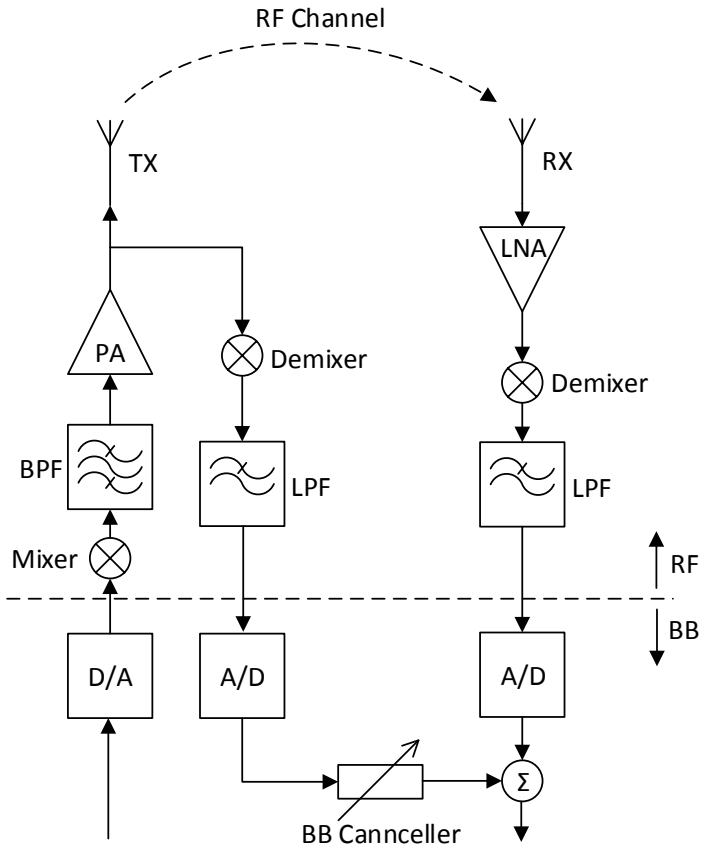


Figure 1.3: Diagram of hybrid self-interference cancellation

the knowledge of transmitted near-end and far-end signals through the LS estimator. Once the channel response of the self-interference channel is estimated, a cancellation channel can be implemented using the FIR filter to reduce the power of the self-interference signal which is similar as the idea used in the digital cancellation. [28] only considered the self-interference and analyzed the system model in the frequency domain. [29] extended the previous work to the MIMO full-duplex system. In [30], a prototype using the hybrid cancellation scheme is presented and the cancellation performance is evaluated.

Since both digital cancellation and hybrid cancellation work after the ADC in the receiver chain, the ADC saturation problem is a critical challenge due the large power of

the self-interference signal. To solve that, the analog cancellation method was proposed and caused great attention. In the analog cancellation, the self interference is cancelled before the ADC in the analog domain at the receiver end. In the early stage, researchers considered to cancel the self-interference using an auxiliary transmit chain [31–34]. With the same baseband transmitted symbols before being converted by the digital to analog converter (DAC) at the transmitter side, the auxiliary transmit chain generates a cancellation signal which has the opposite phase of the self-interference signal and cancels it in the RF domain before the demodulation. However, using an auxiliary transmit chain require additional hardware cost. The extra antenna in the scheme is only used for self-interference cancellation rather transmitting data, the design itself is not efficient in wireless communication perspective. Therefore, in the later of the study, the researchers focused more on the all analog self-interference cancellation design.

All analog cancellation uses a RF canceller to generate the cancellation signal as it is shown in Fig. 1.4. A coupler or splitter is used to get a copy of the RF signal. The RF signal then passes the RF canceller which is formed with delay lines, attenuators and phase shifter. The purpose of this design is using the RF canceller to mimic the practical interference channel. Fig. 1.5 shows three architectures of the RF cancellers that has been proposed. [35, 36] analysed the performance of the RF canceller shown in 1.5(a). In [35], the author assume the frequency response of the interference channel is known and used the proposed structure to match the interference channel with variable attenuations and variable phase shifting. The Gaussian error model is used in [36] to represent the estimation error of the phase and attenuation. Numerical evaluation is performed to show the effect of the

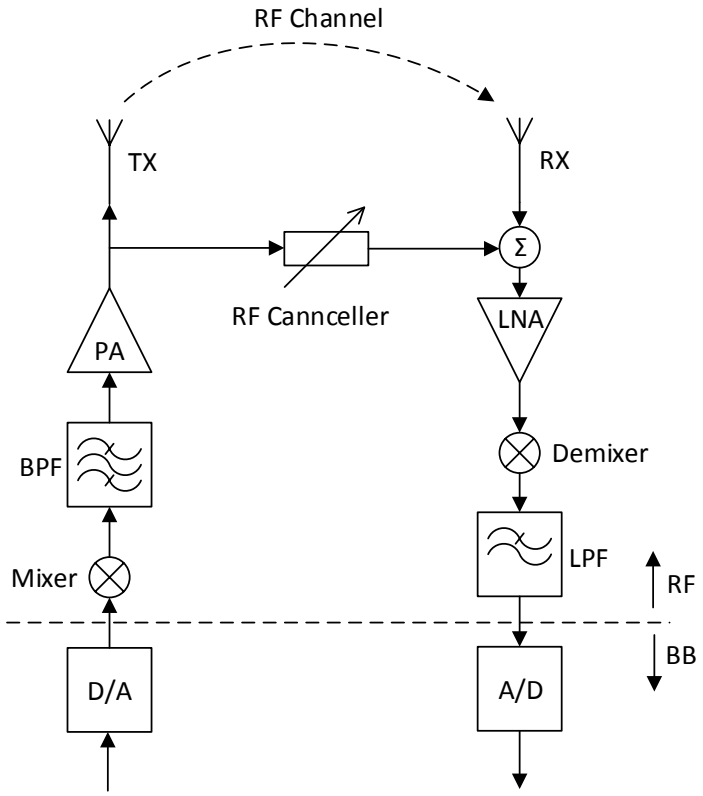


Figure 1.4: Diagram of all analog self-interference cancellation

phase and attenuation estimation errors on the cancellation performance. The variable phase shifters and the attenuators provide more freedom for the tuning algorithm to match the self-interference channel.

However, the practical RF devices will introduce non ideal components to the system. Therefore, the author simplified the structure to 1.5(b) in [14, 37–40]. In the new design of the RF canceller, the variable phase shifters are removed and it only consists fixed delay and variable attenuators. [14, 37] only proposed to use passive attenuations and delays to setup the cancellation channel and no detailed structure is provided. Also the presented seif-interference cancellation degrades fast as the bandwidth increase. So it is

only valided for narrowband signals. In [38, 39], the authors provided a detailed design of the RF canceller using fixed delays and variable attenuators. An analog cancellation board is presented in the paper and the author used it to do the experimental test for the cancellation. The author stated the test results proves their design work for the wideband signal and can make the full duplex feasible and practical. However, it is not showed in the paper how the tuning algorithm is designed and how to find the optimal setting of the RF canceller. The authors extended their work to full-duplex MIMO radio in [40]. A cascaded cancellation design is proposed to reduce the number of taps (paths of delay and attenuators) used to achieve MIMO full-duplex. Same as the previous two papers, again there is no detail introduction of the tuning algorithm used showing how the RF canceller is tuned in the paper.

In [41], the authors named the structure in Fig. 1.5(b) as the uniform structure and demonstrated the shortage of the uniform structure is that it is sensitive to the carrier frequency. To solve this problem, a clustered architecture is proposed and the structure of the clustered canceller is shown in Fig. 1.5(c). In this design, 4 attenuators are connected with cascaded 90 degree phase shifters and form a clustered tap. Each tap is connected via a fixed delay line. In the paper, a detailed comparison of the uniform structure and the clustered structure was performed and the results showed the performance of the clustered structure is as good or better than the uniform structure. Also, the clustered is compatible to different carrier frequencies therefore showing an advantage over the uniform structure.

Using the clustered structure, detailed tuning algorithms are proposed in [42, 43]. In [42], a quadratic model based tuning algorithm is presented. The algorithm used the

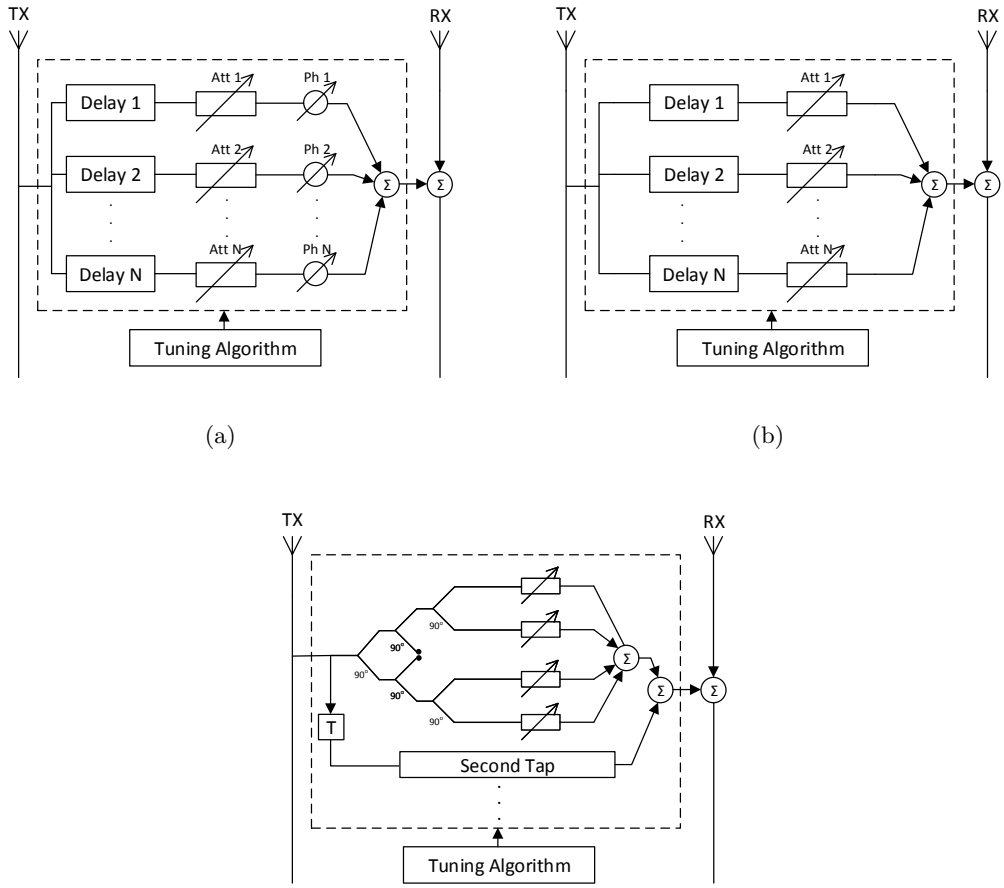


Figure 1.5: Architectures of the RF cancellers.(a), RF canceller consisting of N paths with fix time delay, variable attenuators and variable phase shifters. (b), RF canceller consisting of N paths with fix time delay and variable attenuators. (c), RF canceller consisting of taps, each tap is formed with 90 degree phase shifter and 4 attenuators, a fixed time delay T is used to connect different taps.

power of the residual signal to find the optimal setting of the RF canceller. First, pre-designed training matrix is used to collect the output residual powers with different settings of the RF canceller. Based on the collected information, the coefficients of the quadratic model can be estimated and the optimal setting of the RF canceller can be found easily then. In the paper, a real valued model is also proposed to handle the I/Q imbalance problem. An affine model based blind tuning algorithm is proposed in [43]. Compared to

the quadratic model, the affine model uses the whole receive residual sequence to estimate the coefficients. A PCB board is designed and presented in the paper which implementing the clustered taps. Also, in the analysing, the practical attenuator is considered. Different from the ideal ones, the practical attenuator will introduce phase shifting caused by the attenuation. The phase shift increases as the attenuation decreases in dB and will destroy the linear relationship in the system. To solve this, the paper used the LS algorithm to find the optimal solution considering the equivalent complex attenuation of one clustered tap. Find the optimal setting of the practical attenuator for each tap through brute force by searching for the combination that matches the complex attenuation best.

Meanwhile others considered to least mean square algorithm to tune the RF canceller. [44] demonstrate the feasibility of the LMS tuning algorithm. The author analysed the system using the uniform tap idea. The algorithm find the optimal weights of the RF canceller through iterations and the information of the RF input of each tap is request in the updating algorithm. To collect these information, large number of down-converting chains are need which becomes the main challenge of this design. In [45, 46], the authors presented the RF canceller demonstration boards to show the validation of their design. These demonstrate boards are further integrated with digital self-interference cancellation and the overall self-interference cancellation results for the whole system is shown in [47, 48]. [49] proposed an adaptive least mean square algorithm, however ideal integrators have to be implemented with active circuit. In [50], the authors criticized the previous LMS algorithm for the above two shortages and proposed a passive resistor-capacitor (RC) circuits based design to reduce the hard ware complexity. The impacts of the practical attenuators on

analog cancellation is investigated in [51]. [52] proposed an analog baseband approach and the performance of channel matching is presented. Table 1.1 summarizes the cancellation scheme used in each paper.

Category	Reference
Passive Cancellation	[1], [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13] [19], [35]
Digital Cancellation	[14], [15], [16], [17], [18], [20], [21], [22], [23], [24], [38]
Hybrid Cancellation	[25], [26], [27], [28], [29], [30]
Analog Cancellation	[14], [23], [31], [32], [33], [34], [35], [36], [37], [38], [39] [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50] [51]

Table 1.1: Cancellation Scheme used in the reference paper

To the perspective of the optimization algorithm used in the papers, it can be mainly separated into least square (LS) algorithm, least mean square (LMS) algorithm and recursive least square (RLS) algorithm. LS estimation is an off-line batch algorithm and uses N known equations to estimate M unknowns and the necessary condition is $N \geq M$. The optimal solution minimizes the sum of the squares of the errors in every single equation. Therefore, in the self-interference cancellation, a batch version of the input signal sequence is needed to do the estimation. LS square estimation is robust to instantaneous data and has a high efficiency in the estimation. However, the LS algorithm is not an adaptive algorithm so that it is not robust to the time variant component in the system. Also, among the three algorithms, LS has the highest computational complexity ($O(N^3)$). Both of the LMS algorithm and the RLS algorithm are adaptive algorithm. Contrast to the LS algorithm that do instantaneous estimation, the adaptive algorithms require time and iterations to get converged. The LMS algorithm is popular for its low computational complexity ($O(N)$).

However, it is well known that the convergence speed of LMS algorithm is low. Especially in the full duplex radio, implementing LMS algorithm requires additional down-converting circuits and integrating circuits. The RLS algorithm has a very fast convergence speed and it is robust to the time variant components in the full-duplex system. However, the trade off is that its computational complexity ($O(N^2)$) is higher than the LMS algorithm. Table 1.2 summarizes the optimization algorithm applied in the reference paper.

Category	Reference
Least Square	[14], [15], [16], [17], [18], [25], [26], [27], [28], [29], [31]
Least Mean Square	[20], [21], [44], [45], [46], [47], [48], [49], [50]
Recursive Least Square	[22]

Table 1.2: Optimization algorithm used in the reference paper

In many of the work such as [14, 18, 19, 21, 32, 33, 35–38], ideal components are assumed and the authors proved the validation of their theory by mathematical analysing. However, in a practical full-duplex system, non-ideal components will introduce nonlinearity to the system which will harm the channel estimation and self-interference cancellation since most of the models used in the self-cancellation literatures are linear.

Generally, these non-ideal components includes the phase noise, I/Q imbalance, power amplifier (PA) nonlinearity, ADC resolution and etc. Among them, the I/Q imbalance and the phase noise are proved to be the main challenges for the self-interference cancellation. In [15, 16, 18, 24, 25, 28, 31, 42, 43], the effect of the I/Q imbalance is considered. Widely linear model and real valued model are proposed in [16] and [42] separately to handle the I/Q imbalance problem in the full-duplex system. It has been proved that these two model are equivalent to each other mathematically. Both of them can solve the

nonlinear problem caused by the I/Q imbalance. The advantage of the real valued model is that it has a lower computational complexity in the channel estimation.

In other works like [53–59] addressed the impact of phase noise in different perspective and regarded it as the bottle neck of the self-interference cancellation. Phase noise estimation and compensation algorithms are proposed in [60–62] which are analyzed in both time and frequency domain. During evaluating the cancellation performance of the self-interference cancellation theories, many authors also take the phase noise into account and provided more practical simulation results [23, 27, 28, 34, 44].

It is worth noting that in the previous, the I/Q imbalance and phase noise are analyzed separately. Therefore, the estimation and compensation algorithm are designed with considering only one of them as well. However, since it has been proved that both of these two non-ideal component will affect the self-interference cancellation performance heavily, it is important to analyze the situation both of them are existed and re-evaluate the proposed compensation algorithms.

The effects of other non-ideal component are also considered. In [18, 63], the PA induced nonlinearity is discussed and it is suppressed by both linear and nonlinear signal processing. Also, recall that both hybrid and analog cancellation can avoid the impact of the PA nonlinearity inherently. ADC resolution and quantization limit is evaluated in [64, 65]. [66] analyzed the impact of thermal noise on a full-duplex system. However, these non-ideal component are not as critical as the phase noise and I/Q imbalance so that can be neglected. Table 1.3 shows the nonlinear component considered in the reference paper.

Category	Reference
Ideal	[14], [19], [21], [18], [32], [33], [35], [36], [37], [38]
I/Q Imbalance	[15], [16], [18], [24], [25], [28], [31], [42], [43]
Phase Noise	[53], [54], [55], [56], [57], [58], [59], [60], [61], [62], [23] [27], [28], [34], [44]
Others	[18], [63], [64], [65], [66]

Table 1.3: Nonlinear component considered in the reference paper

We have proposed several novel architectures and algorithms for analog cancellation. A blind tuning algorithm is proposed in [42] where clustered taps are used in the cancellation channel and quadric model is used to model the relationship of input/output of the system. The quantization error induced by the RF step attenuators is considered in [43]. In this paper, an affine model is used to model the system. Blind tuning and brute force searching are applied together to find optimal values for the step attenuators. In my upcoming paper, an optimal training matrix selection algorithm is developed for the quadratic model. The importance of considering I/Q imbalance and phase noise simultaneously in full-duplex system is demonstrated. Also a adaptive blind tuning algorithm which is robust to both I/Q imbalance and phase noise is proposed in that paper. The following chapters will show my contributions in details and the rest of this thesis is organized as follow:

In Chapter 2, RF impairments which induce nonlinearity to the full-duplex radio system will be introduced, the impact of the phase noise and I/Q imbalance is emphasized and analyzed in detail. The real valued representation is proposed to solve the I/Q imbalance nonlinearity. It is compared with the widely linear representation. The results show its advantage on computational complexity.

In Chapter 3, a review of the structure and model of the digital self-interference cancellation and hybrid self-interference cancellation are presented. The least square al-

gorithm is used to estimate the impulse response of the interference channel. With the estimated channel response, the cancellation signal is generated using the FIR filter.

In Chapter 4, a quadratic model based analog self-interference cancellation is presented. A novel blind tuning algorithm is proposed which has two procedures, training and optimization. In the training period, a heuristic training matrix is designed in the early stage and then the optimal training matrix selection is developed to minimize the estimation mean square error with least square (LS) estimation and linear minimum mean square error (LMMSE) estimation.

Chapter 5 will show affine model based analog self-interference cancellation. Practical RF attenuators are considered in the cancellation channel. The impact of quantization error induced by the step attenuators is demonstrated through the simulation. A brute force searching is used to find optimal setting values in dB for the non-ideal RF attenuators. Next, a recursive least square based adaptive blind tuning algorithm is presented to solve the nonlinearity problem from I/Q imbalance and phase noise. Computer simulations are performed to prove the validation of the proposed theories. The results show the proposed adaptive algorithm is robust to both I/Q imbalance and phase noise which outperforms the existing method.

Finally, the conclusion is given in Chapter 6 at the end of this thesis and potential future work is mentioned.

Chapter 2

RF Impairments in Full-duplex Radio System

In the practical full-diplex radio system, signals are distorted by RF impairments. In most of the self-interference cancellation algorithms, the system is modeled by the linear mathematical model. Then, the linear part of the system can be handled easily however the nonlinear part will become a big challenge and affect the cancellation performance. In this chapter, PA nonlinearity, I/Q imbalance and phase noise will be introduced and the effects of them will be discussed.

2.1 PA Nonlinearity

In the transmit chain, the power amplifier (PA) drives the antenna to send out RF signals. It converts a low power RF signal to a higher power signal. Ideally, the output and the input signal of the power amplifier hold the linear relationship. However, the practical

PAs present impairments and are generally modeled by AM/AM and AM/PM. For example, assume the input of the PA is $x(t) = a(t)e^{j\phi(t)}$, the output signal can be express as

$$y_{PA}(t) = AM(a(t))e^{j(\phi(t)+PM(a(t)))} \tag{2.1}$$

$AM(a(t))$ and $PM(a(t))$ describe the output amplitude and phase of the PA as a function of the input signal amplitude separately. Equation (2.1) shows that the $AM(a(t))$ causes nonlinear amplitude distortion and $PM(a(t))$ leads to nonlinear phase shift. In full-duplex radio analyzing, a widely used PA nonlinearity model is the Rapp Model. The Rapp model only deals with the AM/AM component and models $PM(a(t)) = 0$. The $AM(a(t))$ is expressed as

$$AM(a(t)) = \frac{G|x(t)|}{[1 + (\frac{G|x(t)|}{V_{sat}})^{2p}]^{1/2p}} \tag{2.2}$$

where it is recommended to choose $p = 3$ as the smoothness factor, G denotes the small signal gain and V_{sat} is the saturation level.

There are many ways to deal with the PA nonlinearity. One is proposed in [18] by using linear and nonlinear estimation. Another is using the hybrid self-interference cancellation and analog self-interference cancellation. In these two methods, the cancellation starts from the output of the Tx antenna. The distortion from the power amplifier will merge into the transmit signal which will not affect the accuracy of the future optimization algorithm. This becomes one of the advantages of the hybrid cancellation and analog cancellation which will be further discussed later.

2.2 I/Q Imbalance and Phase Noise

For RF system with known carrier frequency, the signals can be equivalently represented in the baseband. Fig.2.1 shows an illustration of the self-interference in the dull-duplex radio using one antenna through the circulator. Compared to the structures shown in Fig.1.2-1.4, the implementation of the circulator can reduce the number of the antenna used in the full-duplex radio since it uses the same antenna to transmit and receive signal. In this section, we present these representations in the context of RF mixer and RF demixer considering the effects of both I/Q imbalance and phase noise.

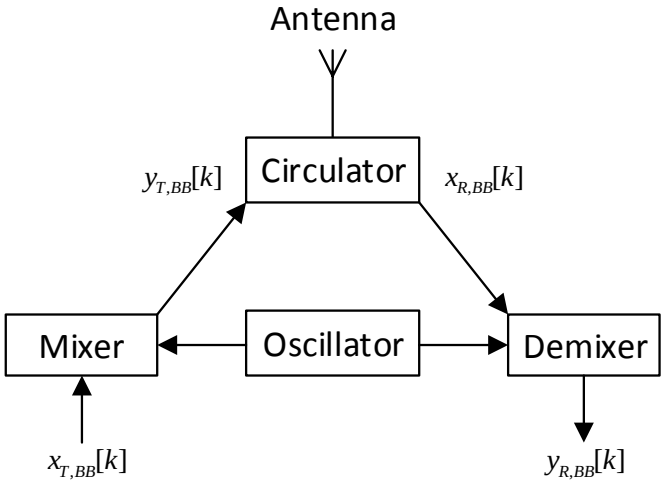


Figure 2.1: An illustration of self-interference in full-duplex radio

2.2.1 I/Q Imbalance and Phase Noise Caused by RF Mixer

At the transmitter side, the transmit chain uses a RF mixer (also called RF modulator) to convert a baseband signal into RF signal as it shows in Fig.2.2. The baseband input signal of the RF mixer has two real-valued components and can be represented by a

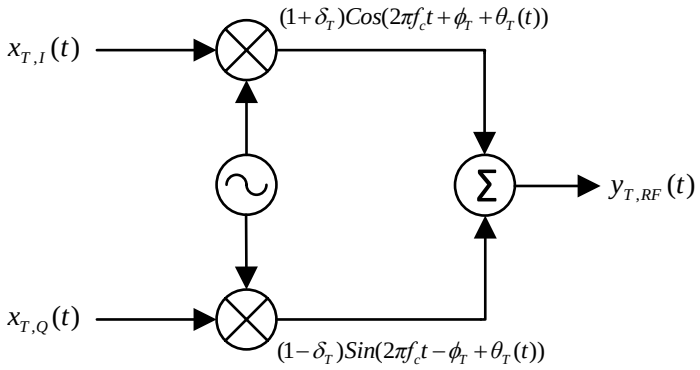


Figure 2.2: The RF mixer I/Q imbalance and phase noise signal model

complex signal: $x_{T,BB}(t) = x_{T,I}(t) + jx_{T,Q}(t)$ where $x_{T,I}(t)$ and $x_{T,Q}(t)$ are the real-valued I and Q components. For a practical RF system, the imperfect RF mixer will introduce I/Q imbalance and phase noise. Then, the RF signal produced by the RF mixer can be expressed as

$$\begin{aligned} y_{T,RF}(t) =&(1+\delta_T)x_{T,I}(t)cos(2\pi f_c t + \phi_T + \theta_T(t)) \\ &- (1-\delta_T)x_{T,Q}(t)sin(2\pi f_c t - \phi_T + \theta_T(t)) \end{aligned} \tag{2.3}$$

where f_c is the carrier frequency, δ_T and ϕ_T are the amplitude and phase components of the I/Q imbalance, $\theta_T(t)$ is the phase noise. The baseband equivalent form of $y_{T,RF}(t)$ is defined as $y_{T,BB}(t)$ that satisfies

$$y_{T,RF}(t) = Re\{y_{T,BB}(t)e^{j2\pi f_c t}\} \tag{2.4}$$

which follows that

$$y_{T,BB}(t) =(1+\delta_T)x_{T,I}(t)e^{j\phi_T + j\theta_T(t)} + j(1-\delta_T)x_{T,Q}(t)e^{-j\phi_T + j\theta_T(t)} \tag{2.5}$$

which is complex-valued. We see if the I/Q imbalance is zero, then $y_{T,BB}(t) = [x_{T,I}(t) + jx_{T,Q}(t)]e^{j\theta_T(t)} = x_{T,BB}(t)e^{j\theta_T(t)}$ which shows a linear relationship between $y_{T,BB}(t)$ and $x_{T,BB}(t)$.

However, if the I/Q imbalance is not zero, the linear dependency between $y_{T,BB}(t)$ and $x_{T,BB}(t)$ no longer holds. However, we can write (2.5) as

$$y_{T,BB}(t) = [a_T x_{T,BB}(t) + b_T x^*_{T,BB}(t)]e^{j\theta_T(t)} \tag{2.6}$$

where $a_T = 0.5[(1+\delta_T)e^{j\phi_T}+(1-\delta_T)e^{-j\phi_T}] = \cos\phi_T + j\delta_T \sin\phi_T$ and $b_T = 0.5[(1+\delta_T)e^{j\phi_T} - (1-\delta_T)e^{-j\phi_T}] = \delta_T \cos\phi_T + j\sin\phi_T$. Both a_T and b_T are complex. The expression of (2.6) shows a linear dependency of $y_{T,BB}(t)$ on $x_{T,BB}(t)$ and $x^*_{T,BB}(t)$ which is proposed in [16] and they call it widely linear model.

Another method proposed in [42] that can re-generate the linear relationship between the $y_{T,BB}(t)$ and $x_{T,BB}(t)$ is the real value model. Apply the real value model and denote $\mathbf{y}_{T,BB}(t) = [y_{T,I}(t), y_{T,Q}(t)]^T$, $\mathbf{x}_{T,BB}(t) = [x_{T,I}(t), x_{T,Q}(t)]^T$. We can rewrite (2.5) as

$$\mathbf{y}_{T,BB}(t) = \mathbf{F}_T(t)\mathbf{C}_T\mathbf{x}_{T,BB}(t) \tag{2.7}$$

where

$$\mathbf{F}_T(t) = \begin{bmatrix} \cos(\theta_T(t)) & -\sin(\theta_T(t)) \\ \sin(\theta_T(t)) & \cos(\theta_T(t)) \end{bmatrix}$$

and

$$C_T = \begin{bmatrix} (1+\delta_T)\cos(\phi_T) & (1-\delta_T)\sin(\phi_T) \\ (1+\delta_T)\sin(\phi_T) & (1-\delta_T)\cos(\phi_T) \end{bmatrix}$$

Note that a_T and b_T can be easily expressed as a linear combination of the entries in C_T and versa vise. Therefore, it is revealed that mathematically the real valued model and widely linear model are equivalent. However, the real value model shows a computational advantage comparing with the widely linear model which is shown as follow.

Assume the phase noise is zero, using N consecutive samples of the received self-interference, in complex model we can express it as

$$\mathbf{y}_C = \mathbf{X}_C \mathbf{h}_C + \mathbf{w}_C \tag{2.8}$$

where $\mathbf{y}_C$ and $\mathbf{w}_C$ represent the self-interference, receive noise separately and both are $N \times 1$ complex valued vectors, complex matrix $\mathbf{X}_C \in C^{N \times L}$ is a $(N \times L)$-block Toeplitz matrix of the transmit signal, and $\mathbf{h}_C$ represents the overall interference channel response and is a $L \times 1$ complex valued vector.

Rewrite (2.8) into widely linear model, we have

$$\begin{aligned} \mathbf{y}_{WL} =& \mathbf{X}_C \mathbf{h}_1 + \mathbf{X}_C^* \mathbf{h}_2 + \mathbf{w}_{WL} \\ =& \mathbf{X}_{WL} \mathbf{h}_{WL} + \mathbf{w}_{WL} \end{aligned} \tag{2.9}$$

where $\mathbf{X}_{WL} = (\mathbf{X}_C, \mathbf{X}_C^*)$ is a $N \times 2L$ complex matrix, $\mathbf{h}_{WL} = [\mathbf{h}_1^T, \mathbf{h}_2^T]^T$ is a $2L \times 1$ complex valued vector. A common and robust method used in the channel estimation in full-duplex is the least square method. So the estimation of $\mathbf{h}_{WL}$ is

$$\hat{\mathbf{h}}_{WL} = (\mathbf{X}_{WL}^H \mathbf{X}_{WL})^{-1} \mathbf{X}_{WL}^H \mathbf{y}_{WL} \tag{2.10}$$

Rewrite (2.8) into real value model, we have

$$\mathbf{y}_R = \mathbf{X}_R \mathbf{h}_R + \mathbf{w}_R \tag{2.11}$$

where $\mathbf{y}_R$ and $\mathbf{w}_R$ are $2N \times 1$ real vectors, $\mathbf{X}_R$ is a $2N \times 2L$ real matrix and the relationship between $\mathbf{X}_R$ and $\mathbf{X}_C$ is

$$\mathbf{X}_R = \begin{bmatrix} \mathbf{X}_{R,11} & \cdots & \mathbf{X}_{R,1N} \\ \cdots & \cdots & \cdots \\ \mathbf{X}_{R,L1} & \cdots & \mathbf{X}_{R,LN} \end{bmatrix} \tag{2.12}$$

each $\mathbf{X}_{R,ij}$ is a 2×2 matrix composed by $x_{c,ij}$ which is the ith row jth column element in $\mathbf{X}_C$ where

$$\mathbf{X}_{R,ij} = \begin{bmatrix} Re(x_{c,ij}) & -Im(x_{c,ij}) \\ Im(x_{c,ij}) & Re(x_{c,ij}) \end{bmatrix} \tag{2.13}$$

$\mathbf{h}_R$ is a $2L \times 1$ real vector. The channel estimation of $\mathbf{h}_R$ is

$$\hat{\mathbf{h}}_R = (\mathbf{X}_R^T \mathbf{X}_R)^{-1} \mathbf{X}_R^T \mathbf{y}_R \tag{2.14}$$

Compare(2.14) to (2.10), when the dimension of the matrix $\mathbf{X}_C$ is large, the computational complexity in the channel estimation mainly comes from the matrix inverse computation ($O(N^3)$). Notice that $(\mathbf{X}_{WL}^H \mathbf{X}_{WL})$ is a $2L \times 2L$ complex valued matrix, $(\mathbf{X}_R^H \mathbf{X}_R)$ is a $2L \times 2L$ real valued matrix. Although they have the same dimension, the complexity of doing matrix inverse for a complex matrix is much higher than that for the pure real matrix.

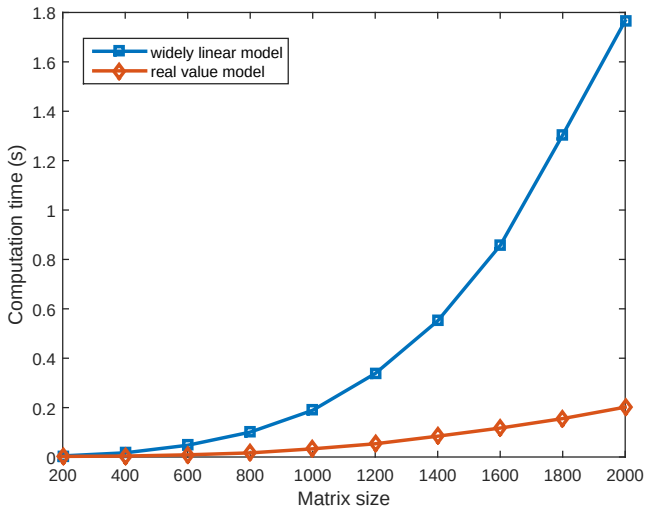


Figure 2.3: Matrix inverse computation time for real model and widely linear model

We simulated the computation speed for the matrix inverse to compare the computational complexity of the widely linear model and real valued model. By generating a random $N \times L$ complex matrix $\mathbf{X}_C$, we can get $\mathbf{X}_{WL}$ and $\mathbf{X}_R$ using (2.9), (2.12) and (2.13). We pick $L = 100, 200, \cdots, 1000$ and counted the corresponding $2L \times 2L$ real and complex matrices inverse computation time. From Fig.2.3, we can see that the computation time keeps at a relative low level for the real valued model. However, the time consuming for the complex matrix inverse increases significantly as the matrix dimension becomes large. With the same dimension of the matrices, the computation speed of the complex valued matrix is much slower than the real valued matrix. This difference becomes obvious especially when the matrix size is large. Therefore, the real value model has a computational advantage of the widely linear model proposed in [16].

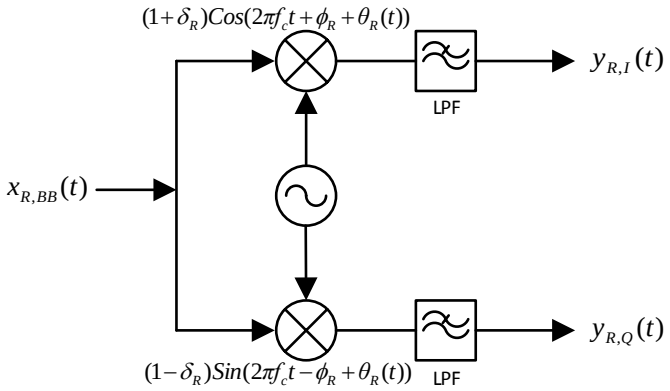


Figure 2.4: The RF demixer I/Q imbalance and phase noise signal model

2.2.2 I/Q Imbalance and Phase Noise Caused by RF Demixer

At the receiver side. the the receive chain uses a RF demixer (also called RF demodulator) to convert the RF signal to a baseband signal as shown in Fig.2.4. Let $x_{R,RF}(t)$ represent the RF input signal of the demixer and $y_{R,BB}(t)$ be the complex baseband output signal generated by the demixer. We can also write $x_{R,RF}(t) = Re\{x_{R,BB}(t)e^{j2\pi f_c t}\}$ with $x_{R,BB}(t) = x_{R,I}(t) + jx_{R,Q}(t)$ being the baseband equivalent form of $x_{R,RF}(t)$. at the demixer, the I channel output signal is achieved by demodulating the input signal by $(1+\delta_R)cos(2\pi f_c t+\phi_R+\theta_R(t))$ and the Q channel output signal is achieved by demodulating the input signal by $(1-\delta_R)sin(2\pi f_c t-\phi_R+\theta_R(t))$. Here, δ_R and ϕ_R are the amplitude and phase components of the I/Q imbalance in the demixer and $\theta_R(t)$ is the phase noise of the demixer. Then it follows that

$$\begin{aligned} y_{R,BB}(t) =& y_{R,I}(t) + jy_{R,Q}(t) \\ =& (1+\delta_R)Re\{x_{R,BB}(t)e^{-j\phi_R - j\theta_R(t)}\} \\ &+ j(1-\delta_R)Im\{x_{R,BB}(t)e^{j\phi_R - j\theta_R(t)}\} \end{aligned} \tag{2.15}$$

For a widely linear model, from (2.15) we have

$$y_{R,BB}(t) = a_R e^{-j\theta_R(t)} x_{R,BB}(t) + b_R e^{j\theta_R(t)} x^*_{R,BB}(t) \tag{2.16}$$

where $a_R = 0.5[(1+\delta_R)e^{-j\phi_R} + (1-\delta_R)e^{j\phi_R}]$ and $b_R = 0.5[(1+\delta_R)e^{j\phi_R} - (1-\delta_R)e^{-j\phi_R}]$. Here, $y_{R,BB}(t)$ is also a linear function of $x_{R,BB}(t)$ and $x^*_{R,BB}(t)$ with complex coefficients.

For the real valued representation, denote $\mathbf{y}_{R,BB}(t) = [y_{R,I}(t), y_{R,Q}(t)]^T$ and $\mathbf{x}_{R,BB}(t) = [x_{R,I}(t), x_{R,Q}(t)]^T$. We can rewrite a linear relationship between $\mathbf{x}_{R,BB}(t)$ and $\mathbf{y}_{R,BB}(t)$ from (2.15) as

$$\mathbf{y}_{R,BB}(t) = \mathbf{C}_R \mathbf{F}_R(t) \mathbf{x}_{R,BB}(t) \tag{2.17}$$

where

$$\mathbf{F}_R(t) = \begin{bmatrix} \cos(\theta_R(t)) & \sin(\theta_R(t)) \\ -\sin(\theta_R(t)) & \cos(\theta_R(t)) \end{bmatrix}$$

and

$$C_R = \begin{bmatrix} (1+\delta_R)\cos(\phi_R) & (1+\delta_R)\sin(\phi_R) \\ (1-\delta_R)\sin(\phi_R) & (1-\delta_R)\cos(\phi_R) \end{bmatrix}$$

It worths noticing the different between (2.7) and (2.17). Also the matrices $\mathbf{C}_T$, $\mathbf{C}_R$, $\mathbf{F}_T(t)$ and $\mathbf{F}_R(t)$ have slice difference due the practical structure and function of the RF mixer and demixer.

2.3 Self-Interference Channel with RF Impairments

Now consider the self-interference channel between the RF frontend from the transmitter to the receiver. Recall previous section, $y_{T,BB}(t)$ denotes the baseband equivalent form of the output of the RF mixer which is also the input of the self-interference channel. $x_{R,BB}(t)$ is the baseband equivalent form of the input of the RF demixer which is also the output of the self-interference channel. Given a bandwidth of interest W and a sampling rate meeting the Nyquist condition (i.e., $T_s \leq \frac{1}{W}$), the discrete time form of $y_{T,BB}(t)$ and $x_{R,BB}(t)$ are denoted by $y_{T,BB}[n]$ and $x_{R,BB}[n]$ separately. Then, apply the real model we have

$$\mathbf{x}_{R,BB}[n] = \sum_{l=0}^{L-1} \mathbf{H}_i[l]\mathbf{y}_{T,BB}[n-l] \tag{2.18}$$

where

$$\mathbf{H}_i[l] = \begin{bmatrix} Re\{h_i[l]\} & -Im\{h_i[l]\} \\ Im\{h_i[l]\} & Re\{h_i[l]\} \end{bmatrix} \tag{2.19}$$

and $h_i[l], l = 0, 1, \cdots, L-1$ is the impulse response of the self-interference channel. Note that the interference channel is typically linear and hence the exact linearity holds between $\mathbf{y}_{T,BB}[n]$ and $\mathbf{x}_{R,BB}[n]$.

Combining the transmitter and receiver RF chain, in another word, combining (2.7), (2.17) and (2.18), we can write the overall baseband input-output relationship of the self-interference channel as follow

$$\mathbf{y}_{R,BB}[n] = \sum_{l=0}^{L-1} \mathbf{C}_R\mathbf{F}_R[l]\mathbf{H}_i[l]\mathbf{F}_T[n-l]\mathbf{C}_T\mathbf{x}_{T,BB}[n-l] \tag{2.20}$$

In (2.20), we can see the sequence $\mathbf{y}_{R,BB}[n]$ is still a linear function of the sequence $\mathbf{x}_{T,BB}[n]$. However, this function is time-varying due the phase noise. Also, the practical structure of the RF mixer and demixer make the effect of I/Q imbalance and phase noise has a cascaded expression. In previous literatures, when only considering the I/Q imbalance, the real model and widely linear model are proposed to solve it. While it can be seen in (2.20) the existence of time variant phase noise will destroy the performance. In the other case, when only considering the phase noise, the property $\mathbf{F}^T[n]\mathbf{F}[n] = \mathbf{I}$ can be utilized for phase noise estimation and compensation whose detail procedure can be found in Apendix A. While the cascaded structure will obstruct it in practical. Therefore, the model and analysis under the assumption of considering only one factor is not precise. So in the rest of this paper, we will based on the model shown in (2.20). To further simplify the problem, we use the second order approximation to deal with it. When the phase noise is small, using the second order approximation for the phase noise, (2.20) can be approximated as

$$\mathbf{y}_{R,BB}[n] = \sum_{l=0}^{L-1} \mathbf{H}_{int}[l]\mathbf{x}_{T,BB}[n-l] + \mathbf{w}_{R,BB}[n] \tag{2.21}$$

where $\mathbf{H}_{int}[l] = \mathbf{C}_R\mathbf{H}_i[l]\mathbf{C}_T$ for $l = 0, 1, \cdots, L-1$ is the overall self-interference channel response and $\mathbf{w}_{R,BB}[k]$ is small perturbation due to the phase noise. More specifically, $\mathbf{w}_{R,BB}[n] = \sum_{l=0}^{L-1} \mathbf{C}_R\partial\mathbf{F}_R[n]\mathbf{H}_i[l]\mathbf{C}_T\mathbf{x}_{T,BB}[n-l] + \sum_{l=0}^{L-1} \mathbf{C}_R\mathbf{H}_i[l]\partial\mathbf{F}_T[n-l]\mathbf{C}_T\mathbf{x}_{T,BB}[n-l]$ where

$$\partial\mathbf{F}_R[n] = \begin{bmatrix} 0 & \theta_R[n] \\ -\theta_R[n] & 0 \end{bmatrix}$$

$$\partial \mathbf{F}_T[n] = \begin{bmatrix} 0 & -\theta_T[n] \\ \theta_T[n] & 0 \end{bmatrix}$$

After the approximation, we can see the effect of phase noise can be modeled as small additive perturbation which has zero mean and linear relationship holds between the input and output.

Chapter 3

Interference Channel Estimation Based Self-Interference Cancellation

In this section, digital self-interference cancellation and hybrid self-interference cancellation technologies will be reviewed. In both of these two methods, after the analog interference signal is converted digital signal by the ADC at the receiver end, the self-interference cancellation is implemented and works in the digital domain. By estimating the channel impulse response of the self-interference channel, the finite impulse response (FIR) filter can be applied to generate the cancellation signal. Then the interference channel estimation becomes the key point and the general estimation algorithm is shown as below.

3.1 Digital Self-interference Cancellation

The diagram of the digital self-interference cancellation is shown in Fig.1.2. In [16], the authors proposed a digital self-interference cancellation algorithm. In this section, we will review this method based on the new approaching.

Recall (2.21), to estimate the channel coefficients, let us rewrite it as

$$\mathbf{y}_{R,BB}[n] = \sum_{l=0}^{L-1} \mathbf{X}_{T,BB}[n-l]\mathbf{h}_{int}[l] + \mathbf{w}_{R,BB}[n] \tag{3.1}$$

where

$$\mathbf{X}_{T,BB}[n-l] = \begin{bmatrix} x_{T,I}[n-l] & -x_{T,Q}[n-l] \\ x_{T,Q}[n-l] & x_{T,I}[n-l] \end{bmatrix} \tag{3.2}$$

and $\mathbf{h}_{int}[l] = [Re\{h_{int}[l]\}, Im\{h_{int}[l]\}]^T$. Using N consecutive samples to form the vector $\mathbf{y}_{real} \in R^{2N\times 1}$, we then have

$$\mathbf{y}_{real} = \mathbf{X}_{real}\mathbf{h}_{real} + \mathbf{w}_{real} \tag{3.3}$$

where $\mathbf{X}_{real} \in R^{2N\times 2L}$ is a $(2N \times 2L)$- block Toeplitz matrix defined as

$$\mathbf{X}_{real} = \begin{bmatrix} \mathbf{X}_{T,BB}[0] & \mathbf{X}_{T,BB}[-1] & \cdots & \mathbf{X}_{T,BB}[-L+1] \\ \mathbf{X}_{T,BB}[1] & \mathbf{X}_{T,BB}[0] & \cdots & \mathbf{X}_{T,BB}[-L+2] \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{X}_{T,BB}[N-1] & \mathbf{X}_{T,BB}[N-2] & \cdots & \mathbf{X}_{T,BB}[N-L] \end{bmatrix} \tag{3.4}$$

and $\mathbf{h}_{real} = [\mathbf{h}_{int}[0]^T, \mathbf{h}_{int}[1]^T, \cdots, \mathbf{h}_{int}[L-1]^T]^T \in R^{2L\times 1}$. The least square solution of $\mathbf{h}_{real}$ is given by

$$\hat{\mathbf{h}}_{real} = (\mathbf{X}_{real}^T \mathbf{X}_{real})^{-1} \mathbf{X}_{real}^T \mathbf{y}_{real} \tag{3.5}$$

Note once the self-interference channel is estimated, the cancellation signal can be generated through digital signal processing. We can apply a Lth order FIR filter and the generated cancellation signal can be expressed as

$$\mathbf{y}_{can} = \mathbf{X}_{real} \hat{\mathbf{h}}_{real} \tag{3.6}$$

In (3.5), it shows that to estimate the interference channel, it requires the knowledge of $\mathbf{X}_{real}$ and $\mathbf{y}_{real}$. $\mathbf{y}_{real}$ is the baseband output of the system which can be measured easily. If the baseband input signal $\mathbf{X}_{real}$ is directly used as the source signal, the whole cancellation occurs in the digital domain and it belongs to the digital self-interference cancellation. However, the performance of this method is heavily limited by the PA nonlinearity. Also the additive perturbation introduced by the impact of phase noises will cause large estimation error which will consequently further limit the cancellation performance. Nonlinear estimation methods are used to solve this problem. As another solution, the hybrid self-interference cancellation is proposed.

3.2 Hybrid Self-interference Cancellation

Hybrid self-interference cancellation uses an auxiliary receive chain to collect the baseband equivalent $\hat{x}_{T,BB}[n]$ of the output of the transmit node as it is shown in Fig.1.3.

The linear relationship of the system is assumed from the estimated signal $\hat{x}_{T,BB}[n]$ to the output interference $\mathbf{y}_{real}$. Then, replace $x_{T,BB}[n]$ with $\hat{x}_{T,BB}[n]$ in (3.3), we have the new interference channel model

$$\mathbf{y}_{real} = \hat{\mathbf{X}}_{real}\mathbf{h}_{real} + \mathbf{w}_{real} \tag{3.7}$$

The impacts of the hybrid cancellation method are twofold. On one hand, it introduces another receive chain which increases complexity of the system and the cost of hardware. On the other hand, since the cancellation channel operates from the output of the transmit chain to the output of the ADC in the receive chain, it can avoid the PA nonlinearity in the transmit chain. Also, the auxiliary receive chain will not introduce extra nonlinearity so the linear relationship still holds in (3.7). If we further assume $\hat{x}_{T,BB}$ and $y_{R,BB}$ share the same phase noise, this method can reduce the level of the perturbation caused by the phase noise. The least square solution of $\mathbf{h}_{real}$ is given by

$$\hat{\mathbf{h}}_{real,h} = (\hat{\mathbf{X}}_{real}^T\hat{\mathbf{X}}_{real})^{-1}\hat{\mathbf{X}}_{real}^T\mathbf{y}_{real} \tag{3.8}$$

and the cancellation channel can be expressed as

$$\mathbf{y}_{can,h} = \hat{\mathbf{X}}_{real}\hat{\mathbf{h}}_{real,h} \tag{3.9}$$

In this section, we reviewed the digital self-interference cancellation and hybrid self-interference cancellation methods. Compared with the digital cancellation, hybrid cancellation can avoid the PA nonlinearity in the transmit chain. However, since both the cancellations of these two methods happen after the ADC, they will suffer from the ADC saturation if the input

self-interference signal is large. In the following sections, we will mainly focus on the analog self-interference cancellation method which the cancellation happens before the ADC therefore, reduce the input signal power and avoid the risk of saturation.

Chapter 4

Optimal Training Matrix Selection for Quadratic Model Based Analog Self-Interference Cancellation

In the following the two chapters, the quadratic model and affine model used in the analog self-interference cancellation will be discussed. Analog cancellation uses the RF canceller to generate the cancellation signal. The RF canceller is controlled by the tunable attenuator, so the key problem is to find the optimal setting of the attenuators that can minimize the residual after the cancellation.

4.1 Quadratic Model Based Analog Self-Interference Cancellation

4.1.1 System Model

Consider a system configuration for quadratic model based analog self-interference cancellation shown in Fig.4.1. In the figure, H_2 represents the interference channel, G represents the RF canceller formed with clustered taps shown in (1.5(c)), H_1 and H_3 represent the transmit and receive chain respectively. $x[n]$ and $y[n]$ are baseband digital waveforms of the source signal and observation separately. $w[n]$ and $v[n]$ denote the noises from the transmit chain and receive chain.

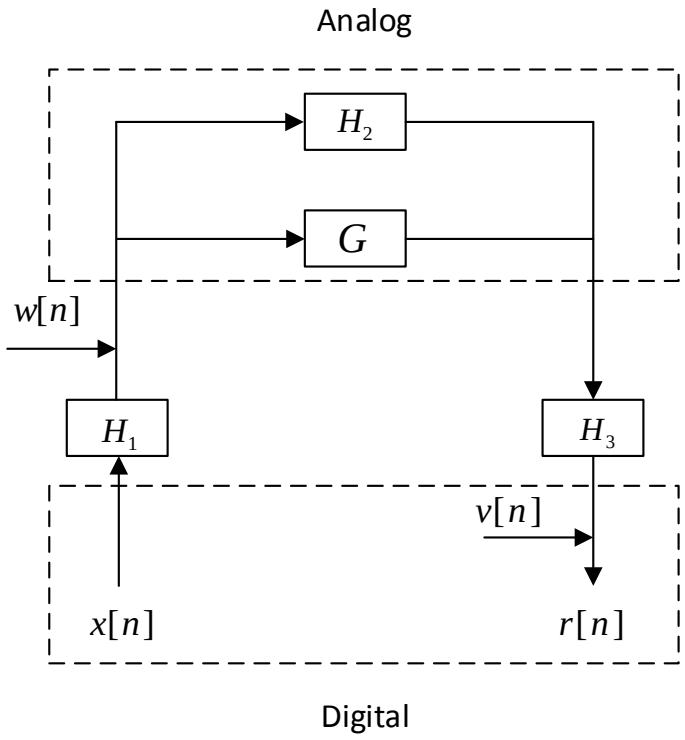


Figure 4.1: System configuration of quadratic model based analog self-interference cancellation

Using the real value model, the relationship between the source signal and the observation vectors can be expressed as

$$\mathbf{r} = \mathbf{y}_{int} + \mathbf{y}_{can} + \mathbf{v} \tag{4.1}$$

with

$$\mathbf{y}_{int} = \mathbf{T}_{H_3}\mathbf{T}_{H_2}\mathbf{w} + \mathbf{T}_{H_3}\mathbf{T}_{H_2}\mathbf{T}_{H_1}\mathbf{x} \tag{4.2}$$

and

$$\mathbf{y}_{can} = \mathbf{T}_{H_3}\mathbf{T}_{\mathbf{g}}\mathbf{w} + \mathbf{T}_{H_3}\mathbf{T}_{\mathbf{g}}\mathbf{T}_{H_1}\mathbf{x} \tag{4.3}$$

where $\mathbf{T}_{H_i}$, $i = 1, 2, 3$ are the linear operators corresponding to the channels. $\mathbf{g}$ is the vector of the controllable attenuator settings. $\mathbf{T}_{\mathbf{g}}$ is the linear operators representing the cancellation channel which is a linear function of $\mathbf{g}$.

Furthermore, we can write $\mathbf{T}_{\mathbf{g}}\mathbf{w} = \mathbf{T}_{\mathbf{w}}\mathbf{g}$ where $\mathbf{T}_{\mathbf{w}}$ is a linear function of $\mathbf{w}$. Applying the same idea, we can write $\mathbf{T}_{\mathbf{g}}\mathbf{T}_{H_1}\mathbf{x} = \mathbf{T}_{H_{1,\mathbf{x}}}\mathbf{g}$. Note $\mathbf{T}_{H_{1,\mathbf{x}}}$ is governed by both the H_1 channel and $\mathbf{x}$, which is not necessarily a linear function of $\mathbf{x}$. Then (4.3) becomes

$$\begin{aligned} \mathbf{y}_{can} =& \mathbf{T}_{H_3}\mathbf{T}_{\mathbf{w}}\mathbf{g} + \mathbf{T}_{H_3}\mathbf{T}_{H_{1,\mathbf{x}}}\mathbf{g} \\ =& \mathbf{T}\mathbf{g} \end{aligned} \tag{4.4}$$

For fixed sequence $\mathbf{x}$, $\mathbf{r}$ is random since there are noise in transmit chain and receive chain. To reduce the effect of the noise, we need collect multiple observations with the same $\mathbf{x}$ and calculate the average power of the observation. Then what we get is

$$
\begin{aligned}
e =& E\{\|\mathbf{r}\|^2\} \\
=& E\{\|\mathbf{Tg} + \mathbf{y}_{int} + \mathbf{v}\|^2\} \\
=& E\{\mathbf{g}^T\mathbf{T}^T\mathbf{Tg} + 2(\mathbf{y}_{int} + \mathbf{v})^T\mathbf{Tg} + \mathbf{y}_{int}^T\mathbf{y}_{int} + 2\mathbf{y}_{int}^T\mathbf{v} + \mathbf{v}^T\mathbf{v}\}
\end{aligned} \tag{4.5}
$$

where E denotes the expectation and assume $\mathbf{v}$ has a zero mean, it can be simplified as

$$
\begin{aligned}
e =& \mathbf{g}^T\{\mathbf{T}^T\mathbf{T}\}\mathbf{g} + 2\{\mathbf{y}_{int}^T\mathbf{T}\}\mathbf{g} + E\{\mathbf{y}_{int}^T\mathbf{y}_{int}\} + E\{\mathbf{v}^T\mathbf{v}\} \\
=& \mathbf{g}^T\mathbf{Ag} + \mathbf{b}^T\mathbf{g} + c + w \\
=& \mathbf{g}_{big}^T\mathbf{p} + w
\end{aligned} \tag{4.6}
$$

where e denotes the power of the residual interference, $\mathbf{g} = [g_1, g_2, \cdots, g_{N_g}]^T$ is the controllable attenuation values, $\mathbf{A}$, $\mathbf{b}$ and c are unknowns and need to be estimated, w represents the perturbation from the noise. $\mathbf{g}_{big}^T = [(\mathbf{g}^T \otimes \mathbf{g}^T)\mathbf{S}^T, \mathbf{g}^T, 1]$ and $\mathbf{p}^T = [vec(\mathbf{A})^T\mathbf{S}^T\mathbf{D}, \mathbf{b}^T, c]$. Here, $\otimes$ denotes the Kronecker product. $\mathbf{S}$ represents a selection matrix that $\mathbf{S}vec(\mathbf{A})$ selects the lower triangular elements of $\mathbf{A}$, $\mathbf{D}$ is a diagonal matrix that doubles the selected off-diagonal elements of $\mathbf{A}$. The purpose of including $\mathbf{S}$ and $\mathbf{D}$ is to reduce the dimension of $\mathbf{g}_{big}$ and $\mathbf{p}$, therefore, reduce the computation complexity in the optimization.

For example, if there is only one clustered tap used in the RF canceller, the dimension of $\mathbf{g}$ is 4×1, then

$$\mathbf{S} = \begin{bmatrix} \mathbf{J}_{4,4} & & & \\ & \mathbf{J}_{4,3} & & \\ & & \mathbf{J}_{4,2} & \\ & & & \mathbf{J}_{4,1} \end{bmatrix} \tag{4.7}$$

where $\mathbf{J}_{N,M}$ with $M \leq N$ is the last M rows of the $N \times M$ identity matrix and

$$\mathbf{D} = diag[1\ 2\ 2\ 2 \mid 1\ 2\ 2 \mid 1\ 2 \mid 1] \tag{4.8}$$

From (4.6), we see there is a quadratic relationship between the $\mathbf{g}$ and residual power e and we call it quadratic system model. Based on the quadratic model, the algorithm to find optimal $\mathbf{g}$ that minimize the residual power e is as follow. First, we use some training vectors of $\mathbf{g}$ to estimate the $\mathbf{A}$, $\mathbf{b}$ and c. Second, we use the estimated $\mathbf{A}$, $\mathbf{b}$ and c to find the optimal $\mathbf{g}$ that minimizes the output e.

To estimate the $\mathbf{A}$, $\mathbf{b}$ and c, we can estimate $\mathbf{p}$ instead. A set of training vectors $\mathbf{g}$ denoted as $\mathbf{g}_1, \cdots, \mathbf{g}_N$ are needed in the learning period. Corresponding to each $\mathbf{g}_i$, we have the observation of the output $\mathbf{y}_i$ and its power e_i, $i = 1, \cdots, N$. Define $\mathbf{e} = [e_1, e_2, \cdots, e_N]^T$, we have

$$\mathbf{e} = \mathbf{G}\mathbf{p} + \mathbf{w} \tag{4.9}$$

where

$$\mathbf{G} = \begin{bmatrix} \mathbf{g}_{big,1}^T \\ \vdots \\ \mathbf{g}_{big,N}^T \end{bmatrix} \tag{4.10}$$

Assume there are N_g attenuators in the system, the dimension of the $\mathbf{g}$ is $N_g \times 1$ and $\mathbf{p}$ is $N_p \times 1$ where $N_p = \frac{N_g(N_g+1)}{2} + N_g + 1$. To have a unique solution $\mathbf{p}$ from (4.9), N is chosen to be $N = N_g(N_g+1)/2 + N_g + 1$ to make sure $\mathbf{G}$ has full column rank. In [42], one heuristic choice of $\mathbf{G}$ is constructed by the following N training vectors of $\mathbf{g}$:

- $\mathbf{g}_1 = \mathbf{0}$.
- For $i = 2, 3, \cdots, N_g + 1$, $\mathbf{g}_i = \alpha \mathbf{e}_{N_g, i-1}$ where $\mathbf{e}_{N_g,i}$ represents the ith column of the $N_g \times N_g$ identity matrix.
- For $i = N_g + 2, N_g + 3, \cdots, 2N_g + 1$, $\mathbf{g}_i = \beta \mathbf{e}_{N_g, i-1}$ with $\beta \neq \alpha$.
- For $i = 2N_g + 2, 2N_g + 3, \cdots, N$, $\mathbf{g}_i = \alpha \mathbf{e}_{N_g,k} + \alpha \mathbf{e}_{N_g,l}$ where $1 \leq k < l \leq N_g$.

Although, it is proved this choice is sparse and well conditioned, it is still not the optimal selection. In the next part, we will discuss the way to figure out the optimal $\mathbf{G}$ choice which has the minimum estimation error of $\mathbf{p}$.

Once we collect all the observations of residual power $\mathbf{e}$, we can estimate $\mathbf{p}$ by $\hat{\mathbf{p}} = \mathbf{G}^{+}\mathbf{e}$. With the estimation of $\mathbf{p}$, we can extract $\mathbf{A}$, $\mathbf{b}$ and c from the $\hat{\mathbf{p}}$. The optimization problem becomes

$$\min_{\mathbf{g}} \{\mathbf{g}^T \mathbf{A} \mathbf{g} + \mathbf{b}^T \mathbf{g} + c + w\} \tag{4.11}$$

and the solution is

$$\mathbf{g}_{opt} = -\frac{1}{2} \mathbf{A}^{+} \mathbf{b} \tag{4.12}$$

4.1.2 Simulation for quadratic model self-interference cancellation

We have simulated the performance of the quadratic model based analog self-interference cancellation shown in Section 4.1.1. In the simulation, we model the relationship between the input and output of the interference channel H_2 as

$$\tilde{y}_{int}(t) = \sum_{i=0}^{I-1} a_i \tilde{x}_{int}(t - \tau_i) \tag{4.13}$$

where $\tilde{x}_{int}(t)$ is the RF input signal and $\tilde{y}_{int}$ is the RF output signal. Considering the real value model, the baseband equivalent of (4.13) is expressed as

$$\mathbf{y}_{int}(t) = \sum_{i=0}^{I-1} a_i \begin{bmatrix} \cos(2\pi f_c \tau_i) & \sin(2\pi f_c \tau_i) \\ -\sin(2\pi f_c \tau_i) & \cos(2\pi f_c \tau_i) \end{bmatrix} \mathbf{x}_{int}(t - \tau_i) \tag{4.14}$$

The attenuation a_i of the multipaths have the form

$$a_i = \frac{\epsilon \alpha_i}{(d + c\tau_i)^2} \tag{4.15}$$

where $a_0 = \epsilon/d^2$,$c = 3 \times 10^8$ m/s, $d = 0.3$ m, $0 \leq \tau_i \leq 10$ ns (random), $0 \leq \alpha_i \leq 1$ (random) and $\epsilon = 8 \times 10^{-4}$.

For the cancellation channel, we use the clustered taps shown in Fig.1.5(c). The relationship of the RF input/output of the G channel is

$$\tilde{y}_{can}(t) = \sum_{n=0}^{N_T-1} \sum_{l=0}^{3} g_{n,l} \tilde{x}_{can}(t - nT - l\eta) \tag{4.16}$$

where $\tilde{x}_{can}(t)$ is the RF input of the G channel and $\tilde{y}_{can}(t)$ is the RF output of the G channel. Then the baseband equivalent is

$$\mathbf{y}_{can}(t) = \sum_{n=0}^{N_T-1} \sum_{l=0}^{3} g_{n,l} \mathbf{G}_{n,l} \mathbf{x}_{can}(t - nT - l\eta) \tag{4.17}$$

where

$$\mathbf{G}_{n,l} = \begin{bmatrix} \cos(2\pi f_c(nT + l\eta)) & \sin(2\pi f_c(nT + l\eta)) \\ -\sin(2\pi f_c(nT + l\eta)) & \cos(2\pi f_c(nT + l\eta)) \end{bmatrix} \tag{4.18}$$

We choose the center frequency as $f_c = 2.5$ GHz, the delay between two clustered taps as $T = \frac{1}{20W}$, the bandwidth of interest as $W = 40$ MHz and $\eta = \frac{1}{4f_c}$

As the inputs of the interference channel and the cancellation channel share the same RF input, it means $\tilde{x}_{int}(t) = \tilde{x}_{can}(t)$. Also the input to H_3 channel is the combination of the outputs of the interference channel and the cancellation channel $\tilde{y}_{int}(t) + \tilde{y}_{can}(t)$. In the blind tuning algorithm, we do not need any direct access to these waveforms.

To simulate the analog channel, we apply a higher sampling rate with the sampling interval $T_L = T_s/L$ with a large L to approximate (4.13) and (4.16) by discrete operations. Let $T_s = 1/W$ be the sampling rate of the digital parts. Then, the H_1 channel is modeled by a discrete-time channel $H_{1,D}$ with rate $1/T_s$, the H_3 channel is modeled by a discrete-time channel $H_{3,D}$ with rate $1/T_s$. The RF interference channel and cancelation channel are modeled by discrete channels with rate $1/T_L$. Where L is chosen to be $L = 500$.

$H_{1,D}$ is modeled by an FIR lowpass filter with a double-sided bandwidth W. Considering the transmit I/Q imbalance the input/ouput relationship of $H_{1,D}$ is

$$\mathbf{y}_1[n] = \sum_{l=-M_h}^{M_h} \mathbf{H}_{1,D}[l] \mathbf{x}_1[n - l] \tag{4.19}$$

with

$$\mathbf{H}_{1,D}[n] = \begin{bmatrix} h_{T,r}[n] & -h_{T,i}[n] \\ h_{T,i}[n] & h_{T,r}[n] \end{bmatrix} \begin{bmatrix} (1+\delta_T)\cos(\phi_T) & (1-\delta_T)\sin(\phi_T) \\ (1+\delta_T)\sin(\phi_T) & (1-\delta_T)\cos(\phi_T) \end{bmatrix} \tag{4.20}$$

where $h_{T,r}[n] = h_{T,i}[n] = w_h[n]c_h(nT_s)$, $c_h(t) = sinc(Wt) = sinc(\pi Wt)/\pi Wt$ and $w_h[n] = 0.54 + 0.46\cos(2\pi n/2M_h + 1)$. We choose $M_h = 20$.

Similarly, $H_{3,D}$ is also modeled as an FIR lowpass filter and its input/output relationship is

$$\mathbf{y}_3[n] = \sum_{l=-M_h}^{M_h} \mathbf{H}_{3,D}[l]\mathbf{x}_3[n-l] \tag{4.21}$$

with

$$\mathbf{H}_{3,D}[n] = \begin{bmatrix} (1+\delta_R)\cos(\phi_R) & (1+\delta_R)\sin(\phi_R) \\ (1-\delta_R)\sin(\phi_R) & (1-\delta_R)\cos(\phi_R) \end{bmatrix} \begin{bmatrix} h_{R,r}[n] & -h_{R,i}[n] \\ h_{R,i}[n] & h_{R,r}[n] \end{bmatrix} \tag{4.22}$$

where $h_{R,r}[n] = h_{R,i}[n] = w_h[n]c_h(nT_s)$. The amplitude imbalance and phase imbalance if the I/Q imbalance are assumed to follow uniform distribution. The range is set to be $\mathcal{U}(-0.05, 0.05)$.

During the training period, the interference channel is assumed to be fixed. The input is chosen to be $x[n] = \delta[n]$. For one training vector $\mathbf{g}_i$, we simulate N_r realizations. Each realization, we apply a random noise $w(mT_L)$ and $v[n]$ and collect the corresponding output residual $r[n]$. Then we compute

$$e_i = \frac{1}{N_r} \sum_{k=1}^{N_r} \sum_{n} \|r^{(k,i)}[n]\|^2 \tag{4.23}$$

where $r^{(k,i)}[n]$ is the kth realization of $r[n]$ corresponding to $\mathbf{g} = \mathbf{g}_i$. With the collected e_i, we can determine $\mathbf{A}$, $\mathbf{b}$ and c.

With the estimated $\mathbf{A}$, $\mathbf{b}$ and c, we compute optimal $\mathbf{g}$ using (4.12). With the optimal solution $\mathbf{g}_{opt}$, we simulate a new realization with another random $w(mT_L)$ and $v[n]$ and collect the output residual $r[n]$. Then the power (in dB) difference of residual before cancellation and after cancellation is the self-interference cancellation performance of the RF canceller. Obviously, the cancellation performance is limited by the level of the noise in the transmit chain and transmission SNR is denoted as SNR_T.

In Fig.4.2, we simulate cancellation performance subject to a large transmission SNR. In this scenario, it illustrates the effect of the number of clustered tap N. It shows with more clustered taps used in the RF canceller, the cancellation performance becomes better. $N > 3$ will not further improve the performance.

In Fig.4.3, we simulate cancellation performance subject to a typical transmission SNR. In this case, we want to demonstrate the impact of number of realizations N_r. When $N_r = 1$, the existence of transmission noise totally destroys the system and the self-interference cancellation is almost not working. When we increase the training realizations, the performance improves. When $N_r = 1000$, the average amount of self-interference cancellation is about 40 dB with two clustered taps used in the RF canceller.

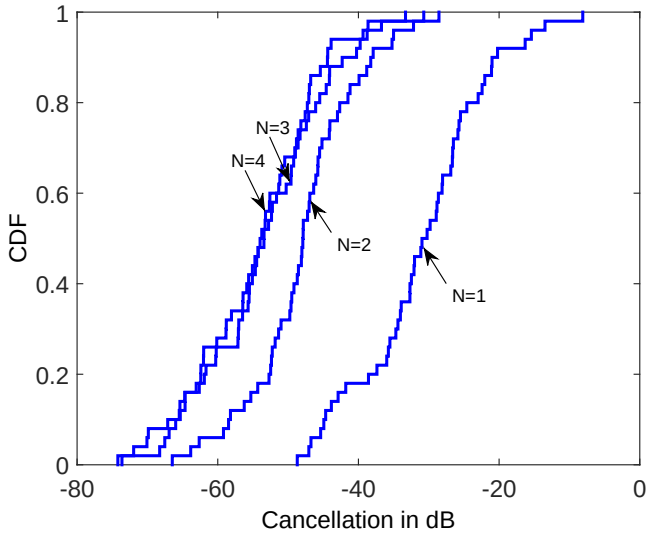


Figure 4.2: CDF of self-interference cancellation performance under $SNR_T = 100$ dB

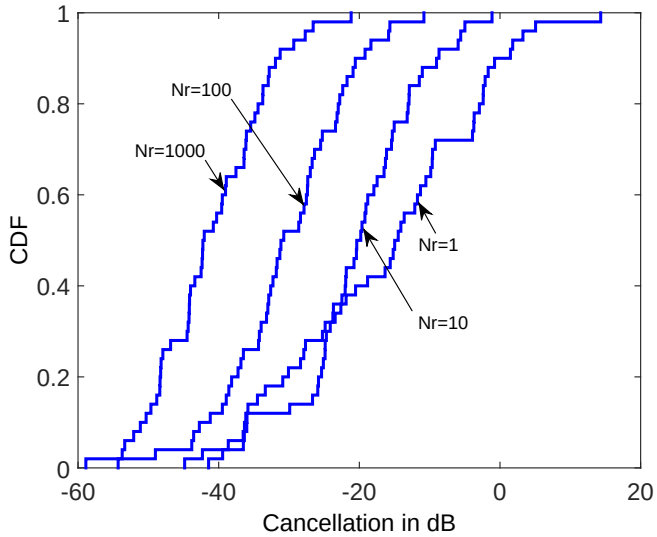


Figure 4.3: CDF of self-interference cancellation performance under $SNR_T = 30$ dB with different number of realization N_r used for training

4.2 Optimal Training Matrix Selection for Channel Estimation

As it is mentioned in the previous section, to find the optimal $\mathbf{g}$, we need do the channel estimation $\mathbf{p}$ using the training matrix $\mathbf{G}$ first. To analyze the relationship between the estimation error $\Delta\hat{\mathbf{p}}$ and $\mathbf{G}$, least square (LS) estimator and linear minimum mean square error (LMMSE) estimator can be applied.

4.2.1 LS Estimation

Using LS method, the estimation of $\mathbf{p}$ can be expressed as

$$\hat{\mathbf{p}}_{LS} = \mathbf{G}^{+}\mathbf{e} \tag{4.24}$$

Substitute (4.9) into (4.24), the estimation error becomes

$$\begin{aligned} \Delta\hat{\mathbf{p}}_{LS} &= \mathbf{G}^{+}(\mathbf{G}\mathbf{p} + \mathbf{w}) - \mathbf{p} \\ &= \mathbf{G}^{+}\mathbf{w} \end{aligned} \tag{4.25}$$

Define the cost function as

$$\begin{aligned} J_{LS} &= E[\|\Delta\hat{\mathbf{p}}_{LS}\|^2] \\ &= Tr[\mathbf{G}^{+}\mathbf{w}\mathbf{w}^{H}\mathbf{G}^{+H}] \\ &= \sigma_w^2 \cdot Tr[(\mathbf{G}^{H}\mathbf{G})^{-1}] \end{aligned} \tag{4.26}$$

where σ_w^2 represents the variance of the noise. The target is to find the optimal training matrix $\mathbf{G}$ to minimize the estimation error. Then, the optimization problem becomes

$$\min_{\mathbf{G}} \sigma_w^2 \cdot Tr[(\mathbf{G}^H \mathbf{G})^{-1}] \tag{4.27}$$

A general idea to solve (4.27) is the gradient decent method. However, since training matrix $\mathbf{G}$ has its own structure shown in (4.10), the optimal solution from gradient decent may not be able to guarantee every row of $\mathbf{G}$ has the structure $\mathbf{g}_{big,i}^T = [(\mathbf{g}_i^T \otimes \mathbf{g}_i^T)\mathbf{S}^T, \mathbf{g}_i^T, 1]$. To solve this, we rewrite (4.27) to an equivalent optimization problem

$$\min_{\mathbf{g}_s} \sigma_w^2 \cdot Tr[(\mathbf{G}^H \mathbf{G})^{-1}] \tag{4.28}$$

where

$$\mathbf{g}_s = \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \\ \vdots \\ \mathbf{g}_N \end{bmatrix} = \begin{bmatrix} g_{1,1} \\ \vdots \\ g_{1,N_g} \\ g_{2,1} \\ \vdots \\ g_{2,N_g} \\ \vdots \\ g_{N,1} \\ \vdots \\ g_{N,N_g} \end{bmatrix} \tag{4.29}$$

Using the gradient decent, we have

$$\mathbf{g}_s^{(k+1)} = \mathbf{g}_s^{(k)} - \mu_1 \frac{\partial J_{LS}^{(k)}}{\partial \mathbf{g}_s^{(k)}} \tag{4.30}$$

μ_1 represents the step size. For every iteration, after calculate the updated $\mathbf{g}_s^{(k+1)}$, the corresponding training vectors $\mathbf{g}_1^{(k+1)}, \mathbf{g}_2^{(k+1)}, \cdots, \mathbf{g}_N^{(k+1)}$ can be extracted. Then the new $\mathbf{g}_{big,i}^{(k+1)}$ and $\mathbf{G}^{(k+1)}$ can be constructed using the equation (4.10). Finally, the new cost function $J_{LS}^{(k+1)}$ can be calculated and used for the next iteration through the equation (4.26).

Following the previous step, the specific structure of $\mathbf{G}$ is promised and the optimal $\mathbf{g}_s$ and corresponding $\mathbf{G}$ could be found. In the equation (4.30), $\frac{\partial J_{LS}^{(k)}}{\partial \mathbf{g}_s^{(k)}}$ is calculated as

$$\begin{aligned}
\partial J_{LS} &= \sigma_w^2 \cdot Tr[\partial((\mathbf{G}^H\mathbf{G})^{-1})] \\
&= \sigma_w^2 \cdot Tr[-(\mathbf{G}^H\mathbf{G})^{-1}\partial(\mathbf{G}^H\mathbf{G})(\mathbf{G}^H\mathbf{G})^{-1}] \\
&= -\sigma_w^2 \cdot Tr[(\mathbf{G}^H\mathbf{G})^{-1}((\partial\mathbf{G}^H)\mathbf{G} + \mathbf{G}^H(\partial\mathbf{G}))(\mathbf{G}^H\mathbf{G})^{-1}] \\
&= -2\sigma_w^2 \cdot Tr[(\mathbf{G}^H\mathbf{G})^{-1}(\partial\mathbf{G}^H)\mathbf{G}(\mathbf{G}^H\mathbf{G})^{-1}]
\end{aligned} \tag{4.31}$$

and according to (4.29)

$$\frac{\partial J_{LS}}{\partial \mathbf{g}_s} = \begin{bmatrix} \frac{\partial J_{LS}}{\partial g_{1,1}} \\ \frac{\partial J_{LS}}{\partial g_{1,2}} \\ \vdots \\ \frac{\partial J_{LS}}{\partial g_{N,N_g}} \end{bmatrix} \tag{4.32}$$

To get the result of (4.32), we can calculate every term one by one using $\frac{\partial J_{LS}}{\partial g_{i,j}} = -2\sigma_w^2 \cdot Tr[(\mathbf{G}^H\mathbf{G})^{-1}\frac{\partial \mathbf{G}^H}{\partial g_{i,j}}\mathbf{G}(\mathbf{G}^H\mathbf{G})^{-1}]$ and $\frac{\partial \mathbf{G}}{\partial g_{i,j}} = \partial\mathbf{G}|_{\partial g_{i,j}=1, other \partial g_{k,l}=0}$ where

$$\partial\mathbf{G} = \begin{bmatrix} (\partial\mathbf{g}_1^T)\otimes\mathbf{g}_1^T\mathbf{S}^T + \mathbf{g}_1^T\otimes(\partial\mathbf{g}_1^T)\mathbf{S}^T, & \partial\mathbf{g}_1^T, & 0 \\ (\partial\mathbf{g}_2^T)\otimes\mathbf{g}_2^T\mathbf{S}^T + \mathbf{g}_2^T\otimes(\partial\mathbf{g}_2^T)\mathbf{S}^T, & \partial\mathbf{g}_2^T, & 0 \\ \vdots & \vdots & \vdots \\ (\partial\mathbf{g}_N^T)\otimes\mathbf{g}_N^T\mathbf{S}^T + \mathbf{g}_N^T\otimes(\partial\mathbf{g}_N^T)\mathbf{S}^T, & \partial\mathbf{g}_N^T, & 0 \end{bmatrix} \tag{4.33}$$

and $\partial\mathbf{g}_i^T = [\partial g_{i,1}, \partial g_{i,2}, \cdots, \partial g_{i,N_g}]$.

With the steps (4.30)-(4.33), we can continuously do the iteration in (4.30) until it converges. Then, extract $\mathbf{g}_1, \cdots, \mathbf{g}_N$ from $\mathbf{g}_s$ and we can get the optimal training matrix $\mathbf{G}$ using (4.10).

4.2.2 LMMSE Estimation

Using the LMMSE method, if $\mathbf{p}$ has a mean of $\bar{\mathbf{p}}$ and covariance matrix of $\mathbf{R_{pp}}$, the estimation of $\mathbf{p}$ can be expressed as

$$\hat{\mathbf{p}}_{LMMSE} = \bar{\mathbf{p}} + \mathbf{R_{pe}}\mathbf{R_{ee}}^{-1}(\mathbf{e} - \mathbf{G}\bar{\mathbf{p}}) \tag{4.34}$$

where $\mathbf{R_{pe}}$ and $\mathbf{R_{ee}}$ represent the covariance matrix correspondingly. Define $\Delta\mathbf{p} = \mathbf{p} - \bar{\mathbf{p}}$ and $\Delta\mathbf{e} = \mathbf{e} - \bar{\mathbf{e}} = \mathbf{e} - \mathbf{G}\bar{\mathbf{p}}$, the estimation error is

$$\begin{aligned}\Delta\hat{\mathbf{p}}_{LMMSE} &= \bar{\mathbf{p}} + \mathbf{R_{pe}}\mathbf{R_{ee}^{-1}}(\mathbf{e} - \mathbf{G}\bar{\mathbf{p}}) - \mathbf{p} \\ &= \mathbf{R_{pe}}\mathbf{R_{ee}^{-1}}\Delta\mathbf{e} - \Delta\mathbf{p}\end{aligned} \tag{4.35}$$

Since it is easy to show $\mathbf{R_{ee}} = \mathbf{R_{\Delta e\Delta e}}$ and $\mathbf{R_{pe}} = \mathbf{R_{\Delta p\Delta e}}$, as in [67], the cost function can be written as

$$\begin{aligned}J_{LMMSE} =& E[\|\Delta\hat{\mathbf{p}}_{LMMSE}\|^2] \\ =& Tr[\mathbf{R_{\Delta p\Delta p}} + \mathbf{R_{pe}}\mathbf{R_{ee}^{-1}}\mathbf{R_{\Delta e\Delta e}}\mathbf{R_{ee}^{-1}}\mathbf{R_{ep}} \\ & - \mathbf{R_{pe}}\mathbf{R_{ee}^{-1}}\mathbf{R_{\Delta e\Delta p}} - \mathbf{R_{\Delta p\Delta e}}\mathbf{R_{ee}^{-1}}\mathbf{R_{ep}}] \\ =& Tr[\mathbf{R_{pp}} - \mathbf{R_{pe}}\mathbf{R_{ee}^{-1}}\mathbf{R_{ep}}] \\ =& Tr[(\mathbf{R_{pp}^{-1}} + \frac{\mathbf{G}^H\mathbf{G}}{\sigma_w^2})^{-1}]\end{aligned} \tag{4.36}$$

similar to (4.27)-(4.28), considering the structure constrain, we solve the following optimization problem

$$\min_{\mathbf{g}_s} Tr[(\mathbf{R_{pp}^{-1}} + \frac{\mathbf{G}^H\mathbf{G}}{\sigma_w^2})^{-1}] \tag{4.37}$$

the solution can be found following the same idea shown in the equation (4.30)-(4.33)

$$\mathbf{g}_s^{(k+1)} = \mathbf{g}_s^{(k)} - \mu_2\frac{\partial J_{LMMSE}^{(k)}}{\partial \mathbf{g}_s^{(k)}} \tag{4.38}$$

where

$$\begin{aligned}\partial J_{LMMSE} &= Tr[\partial((\mathbf{R_{pp}^{-1}} + \frac{\mathbf{G}^H\mathbf{G}}{\sigma_w^2})^{-1})] \\ &= Tr[-(\mathbf{R_{pp}^{-1}} + \frac{\mathbf{G}^H\mathbf{G}}{\sigma_w^2})^{-1}\partial(\mathbf{R_{pp}^{-1}} + \frac{\mathbf{G}^H\mathbf{G}}{\sigma_w^2})(\mathbf{R_{pp}^{-1}} + \frac{\mathbf{G}^H\mathbf{G}}{\sigma_w^2})^{-1}] \\ &= -\frac{2}{\sigma_w^2}Tr[(\mathbf{R_{pp}^{-1}} + \frac{\mathbf{G}^H\mathbf{G}}{\sigma_w^2})^{-1}(\partial\mathbf{G}^H)\mathbf{G}(\mathbf{R_{pp}^{-1}} + \frac{\mathbf{G}^H\mathbf{G}}{\sigma_w^2})^{-1}]\end{aligned} \tag{4.39}$$

4.2.3 Estimation Error versus. Cancellation Residual

Once we know the knowledge of $\mathbf{p}$, we know $\mathbf{A}, \mathbf{b}, c$. Assume the real true estimation is $\mathbf{p}_T, \mathbf{A}_T, \mathbf{b}_T, c_T$. Based on the equation (4.12), the optimal $\mathbf{g}_{opt,T}$ that minimizes the residual power e can be calculated as

$$\mathbf{g}_{opt,T} = -\frac{1}{2}\mathbf{A}_T^{-1}\mathbf{b}_T \tag{4.40}$$

and the corresponding minimal residual power is

$$e_{opt,T} = -\frac{1}{4}\mathbf{b}_T^T\mathbf{A}_T^{-1}\mathbf{b}_T + c_T + w \tag{4.41}$$

However, since we have noise in the system, we can never get a perfect estimation of $\mathbf{p}$. For the estimated $\hat{\mathbf{A}}, \hat{\mathbf{b}}, \hat{c}$, we have

$$\begin{aligned}
\mathbf{g}_{opt} &= -\frac{1}{2}\hat{\mathbf{A}}^{-1}\hat{\mathbf{b}} \\
&= -\frac{1}{2}(\mathbf{A}_T + \Delta\hat{\mathbf{A}})^{-1}(\mathbf{b}_T + \Delta\hat{\mathbf{b}}) \\
&\cong -\frac{1}{2}(\mathbf{A}_T^{-1} - \mathbf{A}_T^{-1}\Delta\hat{\mathbf{A}}\mathbf{A}_T^{-1})(\mathbf{b}_T + \Delta\hat{\mathbf{b}}) \\
&= -\frac{1}{2}\mathbf{A}_T^{-1}\mathbf{b}_T - \frac{1}{2}\mathbf{A}_T^{-1}\Delta\hat{\mathbf{b}} + \frac{1}{2}\mathbf{A}_T^{-1}\Delta\hat{\mathbf{A}}\mathbf{A}_T^{-1}\mathbf{b}_T + \frac{1}{2}\mathbf{A}_T^{-1}\Delta\hat{\mathbf{A}}\mathbf{A}_T^{-1}\Delta\hat{\mathbf{b}}
\end{aligned} \tag{4.42}$$

where $\Delta\hat{\mathbf{A}}$ and $\Delta\hat{\mathbf{b}}$ represent the estimation error. Denote $\mathbf{x} = \mathbf{A}_T^{-1}\mathbf{b}_T$, $\mathbf{y} = \mathbf{A}_T^{-1}\Delta\hat{\mathbf{b}} - \mathbf{A}_T^{-1}\Delta\hat{\mathbf{A}}\mathbf{A}_T^{-1}\mathbf{b}_T$ and $\mathbf{z} = \mathbf{A}_T^{-1}\Delta\hat{\mathbf{A}}\mathbf{A}_T^{-1}\Delta\hat{\mathbf{b}}$. Then $\mathbf{g}_{opt} = -\frac{1}{2}(\mathbf{x} + \mathbf{y} - \mathbf{z})$. Apply it to the equation (4.6), the residual under the optimal training sequence is

$$
\begin{aligned}
e =& \mathbf{g}_{opt}^T \mathbf{A}_T \mathbf{g}_{opt} + \mathbf{b}_T^T \mathbf{g}_{opt} + c_T + w \\
=& \frac{1}{4}(\mathbf{x}+\mathbf{y}-\mathbf{z})^T \mathbf{A}_T (\mathbf{x}+\mathbf{y}-\mathbf{z}) - \frac{1}{2}\mathbf{b}_T^T(\mathbf{x}+\mathbf{y}-\mathbf{z}) + c_T + w
\end{aligned} \tag{4.43}
$$

since $\mathbf{x}^T\mathbf{A}_T = \mathbf{b}_T^T\mathbf{A}_T^{-1}\mathbf{A}_T = \mathbf{b}_T^T$, the first order perturbation term

$$
\frac{1}{2}\mathbf{x}^T\mathbf{A}_T\mathbf{y} - \frac{1}{2}\mathbf{b}_T^T\mathbf{y} = 0 \tag{4.44}
$$

so, we consider the second order perturbation and e becomes

$$
\begin{aligned}
e =& \frac{1}{4}\mathbf{x}^T\mathbf{A}_T\mathbf{x} - \frac{1}{2}\mathbf{b}_T^T\mathbf{x} + c_T + w + \frac{1}{4}(-2\mathbf{x}^T\mathbf{A}_T\mathbf{z} + \mathbf{y}^T\mathbf{A}_T\mathbf{y}) + \frac{1}{2}\mathbf{b}_T^T\mathbf{z} \\
=& e_{opt,T} + \frac{1}{4}(-2\mathbf{x}^T\mathbf{A}_T\mathbf{z} + \mathbf{y}^T\mathbf{A}_T\mathbf{y}) + \frac{1}{2}\mathbf{b}_T^T\mathbf{z} \\
=& e_{opt,T} + \frac{1}{4}\mathbf{y}^T\mathbf{A}_T\mathbf{y} \\
=& e_{opt,T} + \frac{1}{4}(\mathbf{A}_T^{-1}\Delta\hat{\mathbf{b}} - \mathbf{A}_T^{-1}\Delta\hat{\mathbf{A}}\mathbf{A}_T^{-1}\mathbf{b}_T)^T\mathbf{A}_T(\mathbf{A}_T^{-1}\Delta\hat{\mathbf{b}} - \mathbf{A}_T^{-1}\Delta\hat{\mathbf{A}}\mathbf{A}_T^{-1}\mathbf{b}_T) \\
=& e_{opt,T} + \frac{1}{4}(\hat{\Delta}\mathbf{b}^T\mathbf{A}_T^{-1}\Delta\hat{\mathbf{b}} - 2\Delta\hat{\mathbf{b}}^T\mathbf{A}_T^{-1}\Delta\hat{\mathbf{A}}\mathbf{A}_T^{-1}\mathbf{b}_T + \mathbf{b}_T^T\mathbf{A}_T^{-1}\Delta\hat{\mathbf{A}}\mathbf{A}_T^{-1}\Delta\hat{\mathbf{A}}\mathbf{A}_T^{-1}\mathbf{b}_T)
\end{aligned} \tag{4.45}
$$

then denote the residual error Δe which is caused by the estimation error as

$$
\Delta e = \frac{1}{4}(\Delta\hat{\mathbf{b}}^T\mathbf{A}_T^{-1}\Delta\hat{\mathbf{b}} - 2\Delta\hat{\mathbf{b}}^T\mathbf{A}_T^{-1}\Delta\hat{\mathbf{A}}\mathbf{A}_T^{-1}\mathbf{b}_T + \mathbf{b}_T^T\mathbf{A}_T^{-1}\Delta\hat{\mathbf{A}}\mathbf{A}_T^{-1}\Delta\hat{\mathbf{A}}\mathbf{A}_T^{-1}\mathbf{b}_T) \tag{4.46}
$$

(4.46) shows the relationship between the estimation error $\Delta\hat{\mathbf{A}}$ and $\Delta\hat{\mathbf{b}}$ and the perturbation of the residual power Δe. Considering the statistic performance, its expectation is

$$\begin{aligned}\mathbf{E}[\Delta e] =& \frac{1}{4}(\sum_{g=1}^{N_g}\sum_{f=1}^{N_g}(\mathbf{A}_T^{-1})_{gf}\mathbf{E}[\hat{\Delta}b_g\hat{\Delta}b_f] \\ &- 2\sum_{k=1}^{N_g}\sum_{j=1}^{N_g}\sum_{i=1}^{N_g}\sum_{h=1}^{N_g}(\mathbf{A}_T^{-1})_{kj}(\mathbf{A}_T^{-1})_{ih}b_h\mathbf{E}[\hat{\Delta}b_k\Delta\hat{\mathbf{A}}_{ji}] \\ &+ \sum_{r=1}^{N_g}\cdots\sum_{m=1}^{N_g}\sum_{l=1}^{N_g}b_r(\mathbf{A}_T^{-1})_{rq}(\mathbf{A}_T^{-1})_{pn}(\mathbf{A}_T^{-1})_{ml}b_l\mathbf{E}[\Delta\hat{\mathbf{A}}_{qp}\Delta\hat{\mathbf{A}}_{nm}])\end{aligned} \tag{4.47}$$

where $\mathbf{E}[\Delta b_g \Delta b_f]$, $\mathbf{E}[\Delta b_k \Delta \hat{\mathbf{A}}_{ji}]$ and $\mathbf{E}[\Delta \hat{\mathbf{A}}_{qp} \Delta \hat{\mathbf{A}}_{nm}]$ can be found in the covariance matrix of $\Delta\hat{\mathbf{p}}$. For LS case, it is calculated as

$$\mathbf{E}[\Delta\hat{\mathbf{p}}\Delta\hat{\mathbf{p}}^T]_{LS} = \sigma_w^2(\mathbf{G}^H\mathbf{G})^{-1} \tag{4.48}$$

For LMMSE case, it is expressed as

$$\mathbf{E}[\Delta\hat{\mathbf{p}}\Delta\hat{\mathbf{p}}^T]_{LMMSE} = (\mathbf{R}_{\mathbf{pp}}^{-1} + \frac{\mathbf{G}^H\mathbf{G}}{\sigma_w^2})^{-1} \tag{4.49}$$

4.2.4 Quantization Error

Since the real RF Digital Step Attenuators (DSA) are used in the system, quantization would occur in the both learning and optimization periods.

In the learning period, due to the dynamic range of the attenuator, although the initial and optimal choices of $\mathbf{g}_s$ and $\mathbf{G}$ are already known, we can not access to that setting in most cases. Therefore, we can only use the quantized $\mathbf{g}_{s,q}$ to get $\mathbf{G}_q$ as our training matrix. Rather than the quantization error, this is the hardware limitation. Due to the

limitation of the hardware, we can only use a suboptimal choice $\mathbf{G}_q$ which has a larger estimation MSE to do the training instead of using the optimal $\mathbf{G}$.

In the optimization period, once we find the optimal $\mathbf{g}$, we need to quantize it, then we have

$$\mathbf{g}_{opt,q} = \mathbf{g}_{opt} + \Delta\mathbf{g}_q \tag{4.50}$$

$\Delta\mathbf{g}_q$ is the quantization error and $\mathbf{g}_{opt,q}$ is the final optimal setting for the cancellation channel. Therefore, in the equation (4.46), we need to replace $\mathbf{g}_{opt}$ by $\mathbf{g}_{opt,q}$, the residual power e becomes

$$\begin{aligned} e_q =& \mathbf{g}_{opt,q}^T \mathbf{A}_T \mathbf{g}_{opt,q} + \mathbf{b}_T^T \mathbf{g}_{opt,q} + c_T + w \\ =& \frac{1}{4}(\mathbf{x}+\mathbf{y}-\mathbf{z}+2\Delta\mathbf{g}_q)^T \mathbf{A}_T (\mathbf{x}+\mathbf{y}-\mathbf{z}+2\Delta\mathbf{g}_q) \\ & - \frac{1}{2}\mathbf{b}_T^T(\mathbf{x}+\mathbf{y}-\mathbf{z}+2\Delta\mathbf{g}_q) + c_T + w \end{aligned} \tag{4.51}$$

similar to (4.44), the first order perturbation of $\Delta\mathbf{g}_q$

$$\mathbf{x}^T \mathbf{A}_T \Delta\mathbf{g}_q - \mathbf{b}_T^T \Delta\mathbf{g}_q = 0 \tag{4.52}$$

consider the second order perturbation, the equation (4.45) becomes

$$e_q = e_{opt,T} + \Delta e + \Delta\mathbf{g}_q^T \mathbf{A}_T \Delta\mathbf{g}_q \tag{4.53}$$

In (4.53), the residual power e_q can separate to three parts. $e_{opt,T}$ represents the idea minimal residual power mentioned in (4.41), Δe is the estimation error based residual error

in (4.46). Denote $\Delta e_q = \Delta \mathbf{g}_q^T \mathbf{A}_T \Delta \mathbf{g}_q$ as the residual error caused by the quantization error. The expectation of Δe_q is

$$\mathbf{E}[\Delta e_q] = \sum_{j=1}^{Ng} \sum_{i=1}^{Ng} \mathbf{A}_{T,ji} \mathbf{E}[\Delta g_{q,j} \Delta g_{q,i}] \tag{4.54}$$

Since the quantization error of $\mathbf{g}_{opt}$ in dB follows the uniform distribution, $\Delta g_{q,dB,i} \sim U(-\frac{s}{2}, \frac{s}{2})$ for $i = 1, 2, \cdots, N_g$ and s represents the step size the of the DSA.

$$\partial g_{q,dB,i} = \frac{10}{\ln 10} \cdot \frac{1}{g_{opt,q,i}} \partial g_{q,i} = h(\partial g_{q,i}) \tag{4.55}$$

Since we know the PDF of $\Delta g_{q,dB,i}$ is $f_{dB}(\Delta g_{q,dB,i}) = \frac{1}{\frac{s}{2} - (-\frac{s}{2})} = \frac{1}{s}$, the PDF of $\Delta g_{q,i}$ can be expressed as

$$\begin{aligned} f(\Delta g_{q,i}) &= f_{dB}(h(\Delta g_{q,i})) \cdot \frac{\partial h(\Delta g_{q,i})}{\partial (\Delta g_{q,i})} \\ &= \frac{1}{s} \cdot \frac{10}{\ln 10} \cdot \frac{1}{g_{opt,q,i}} \end{aligned} \tag{4.56}$$

Denote $D = \frac{s \cdot g_{opt,q,i} \ln 10}{10}$, in the equation (4.54), when $j = i$

$$\begin{aligned} \mathbf{E}[\Delta g_{q,j} \Delta g_{q,i}] =& \mathbf{E}[\Delta g_{q,i}^2] \\ =& \int_{-\frac{D}{2}}^{\frac{D}{2}} [f(\Delta g_{q,i}) \cdot \Delta g_{q,i}^2] d\Delta g_{q,i} \\ =& \frac{2}{3} (\frac{D}{2})^3 \frac{10}{s g_{opt,q,i} \ln 10} \\ =& \frac{D^2}{12} \end{aligned} \tag{4.57}$$

when$j \neq i$, $\Delta g_{q,j}$ and $\Delta g_{q,i}$ are independent

$$\mathbf{E}[\Delta g_{q,j} \Delta g_{q,i}] = (\mathbf{E}[\Delta g_{q,i}])^2 = 0 \tag{4.58}$$

4.2.5 Simulation for optimal training matrix selection

In this section, simulation results are presented to compared the mean square error and the residual power error of the initial training matrix and the optimal training matrix. The initial training matrix $\mathbf{G}$ is constructed by the N training vectors of $\mathbf{g}$ shown in Section 4.1.1 as the heuristic choice with $\alpha = 1$ and $\beta = 0.5$. We set the noise variance $\sigma_w^2 = 1$, the covariance matrix $\mathbf{R_{pp}}$ is modeled as $\mathbf{R_{pp}} = \sigma^2\mathbf{I}$ and σ^2 is chosen to be $\sigma^2 = 50$. Using (4.30) and (4.38), after the quantization of the converged $\mathbf{g}_s$, we can find the optimal training matrix $\mathbf{G}$. Then we compared the performance of the initial and optimal choice as follow.

Fig.4.4 shows the MSE J_{LS} and J_{LMMSE} of the LS and LMMSE estimator with the initial and optimal training sequence versus the number of the attenuators N_g where $N_g = 1, 2, \cdots, 10$. From the figure, we see that the MSE increase with the increase of N_g since the size of the training matrix and the complexity of the system is becoming larger for both cases. As expected, compared with the initial choice, the optimal one can reduce the MSE obviously and the difference becomes with the number of attenuators increasing.

Fig.4.5 shows the residual power error of the LS estimator. We choose $N_g = 12$. The x label represents 100 different realizations of the random multi-paths interference channel. For every realization, we compared the expectation of the residual power errors

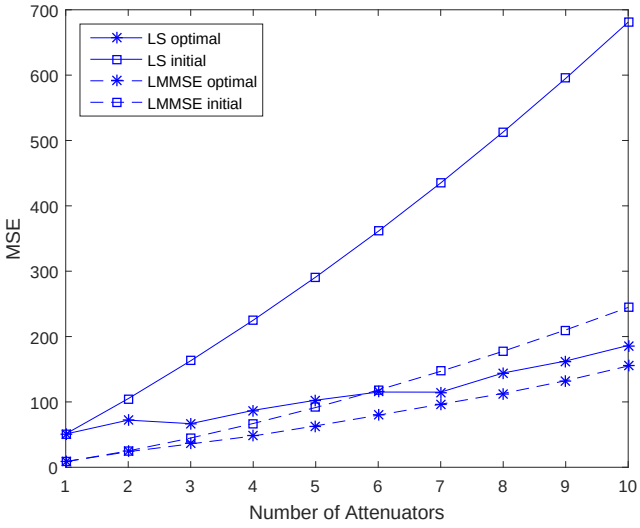


Figure 4.4: Channel estimation MSE of LS and LMMSE estimators

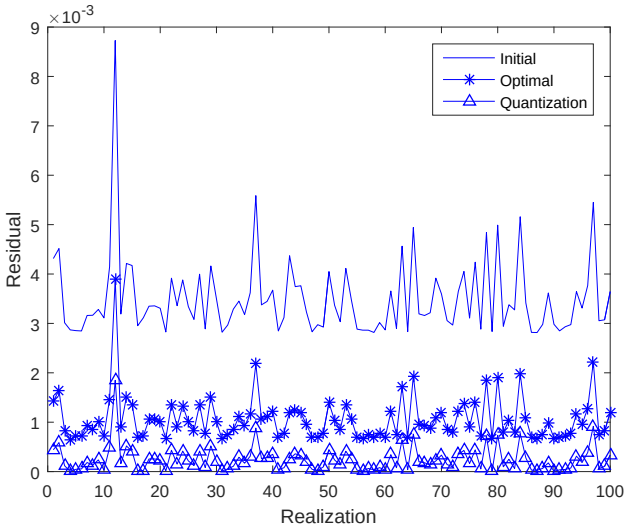


Figure 4.5: Residual power error of LS estimator

Δe for both initial and optimal training matrix. Also we simulated the expectation of quantization error based residual error $\Delta \mathbf{g}_q$ for the optimal choice. If the quantization error dominates the residual error, the optimal channel matrix selection is useless because even we have the estimation error based residual error component, it is negligible compared to the quantization error part. In the figure, we can see the the quantization error based residual error is much smaller than the residual error part. We know the training matrix selection is meaningful. Also, compared to the initial choice, with the optimal training matrix, the estimation error based residual error level reduces significantly for every realization. This shows the necessity of optimal training matrix selection.

Chapter 5

Adaptive Blind Tuning for Affine Model Based Analog Self-Interference Cancellation

5.1 Architecture of All-Analog Cancellation Channel

The architecture of the all-analog cancellation channel is crucial for the full-duplex radio which can determine the self-interference cancellation performance. A well designed architecture should have a high capacity to match the interference channel.

The impulse response of the RF interference channel can be written as

$$h_{int,RF}(t) = \sum_{i=0}^{I-1} a_i \delta(t - \tau_i) \tag{5.1}$$

where I is the number of the multipaths, a_i the propagating attenuation of the ith path and τ_i is the delay of the ith path. denote f_c as the carrier frequency and W as the bandwidth, the baseband equivalent of (5.1) is

$$h_{int}(t) = \sum_{i=0}^{I-1} a_i e^{-j2\pi f_c \tau_i} sinc(W(t-\tau_i)) \tag{5.2}$$

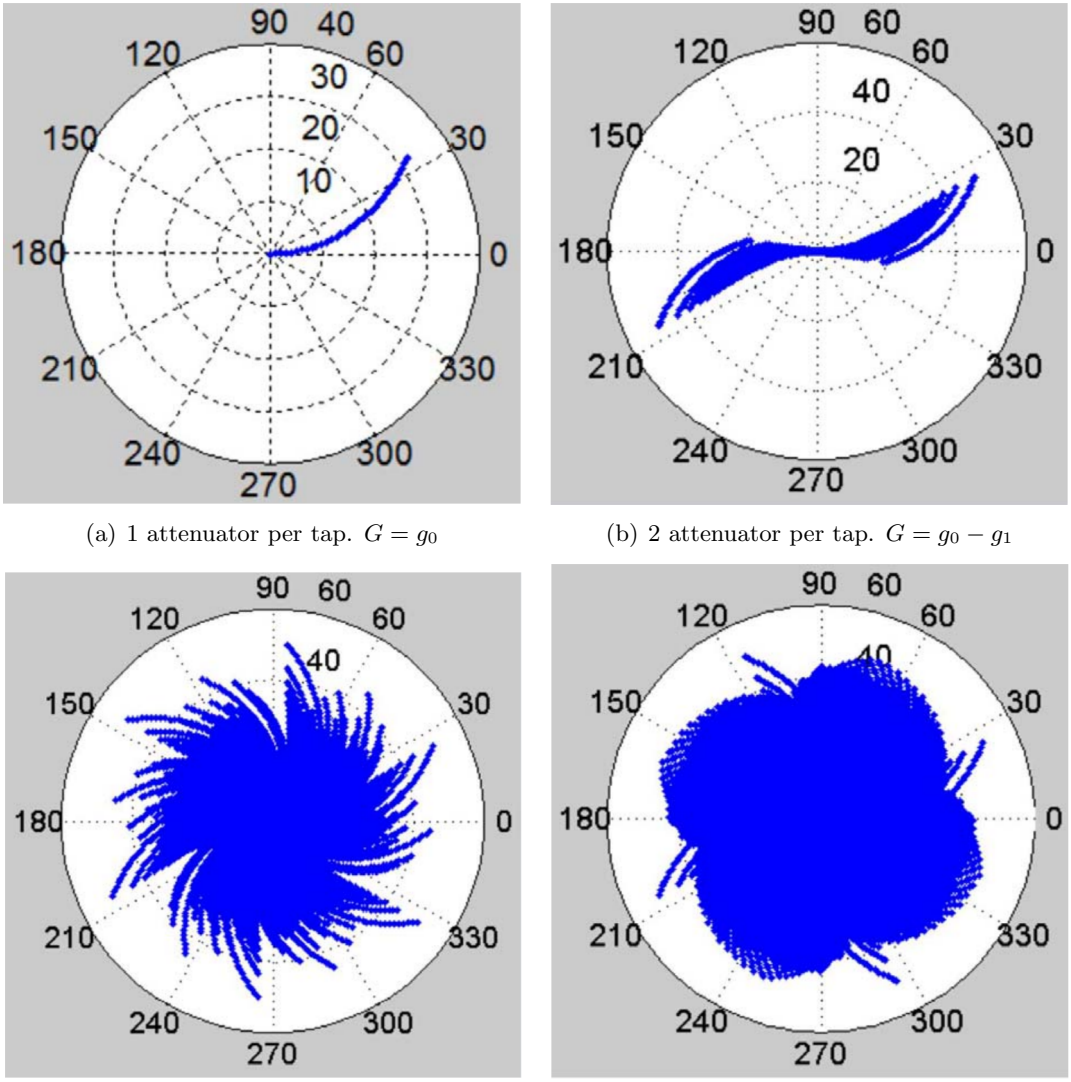
An architecture of the RF canceller that can match the above interference channel properly is shown in Fig.1.5(c). In each branch, same number of 90-degree phase shifters are used to balance the insertion losses and phase shift of the four paths in each clustered tap. The impulse response of this RF canceller is

$$h_{can,RF}(t) = \sum_{l=0}^{N_T-1} (P_0^3 g_{l,0} + P_0^2 P_{\pi/2} g_{l,1} + P_0 P_{\pi/2}^2 g_{l,2} + P_{\pi/2}^3 g_{l,3}) \delta(t-lT) \tag{5.3}$$

where P_0 is the transfer function of the 0-degree output of the 90-degree phase shifter, $P_{\pi/2}$ is the transfer function of the 90-degree output of the 90-degree phase shifter. $g_{l,k}$ is the attenuation of the kth attenuator in the lth clustered tap. Subject to the same carrier frequency and bandwidth of interest, the baseband equivalent of (5.3) is

$$h_{can}(t) = \sum_{l=0}^{N_T-1} P_0^3 G_l e^{-j2\pi f_c lT} sinc(W(t-lT)) \tag{5.4}$$

where $G_l = g_{l,0} - j g_{l,1} - g_{l,2} + j g_{l,3}$.



(a) 1 attenuator per tap. $G = g_0$

(b) 2 attenuator per tap. $G = g_0 - g_1$

(c) 3 attenuator per tap. $G = g_0 + e^{-j2\pi/3}g_1 + e^{j2\pi/3}g_2$

(d) 4 attenuator per tap. $G = g_0 - jg_1 - g_2 + jg_3$

Figure 5.1: Impedance coverage of a single clustered tap with different number of attenuators. Each attenuator has the attenuation range from 0dB to 32dB with 1dB step size. Each attenuator also has an attenuation-dependent phase

To demonstrate that the design of clustered tap is a good choice, the impedance of coverage of the tap is shown in 5.1 for different number of attenuators evenly distributed in phase per tap. It is clear that 5.1(d) covers most of the area and therefore it is the most desirable choice.

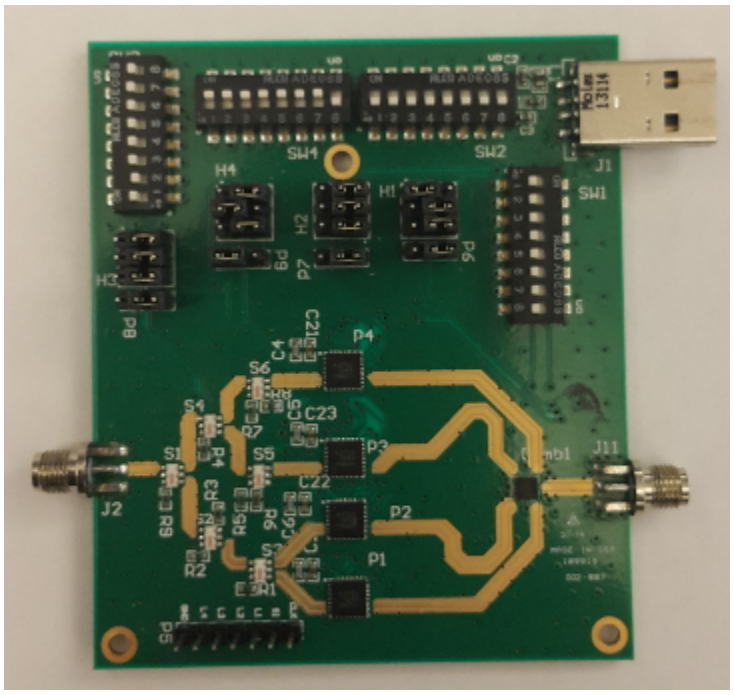


Figure 5.2: A prototype of clustered tap using four step-attenuators (PE43703), six (could be five) 90-degree splitters (QCN27) and power combiner (WP4U1+)

Fig.5.2 shows a custom fabricated single clustered tap PCB board which is based on the structure in Fig.1.5(c).

5.2 Blind Tuning Algorithm

In the previous chapter, the power of the residual is regarded as the observation and the quadratic model is derived. In this chapter, we assume the observation of the interference is the entire waveforms after the cancellation and we have

$$\begin{aligned} \mathbf{r} &= \mathbf{y} + \mathbf{M}\mathbf{g} + \mathbf{w} \\ &= \mathbf{X}\mathbf{v} + \mathbf{w} \end{aligned} \tag{5.5}$$

where $\mathbf{r} = [Re\{r(1)\}, Im\{r(1)\}, \cdots, Re\{r(K)\}, Im\{r(K)\}]^T$ is the observed waveforms, the training vector is $\mathbf{g} = [\mathbf{g}_0^T, \cdots, \mathbf{g}_{N_T-1}^T]^T$ with $\mathbf{g}_l = [Re(g_{l,0}), Im(g_{l,0}), \cdots, Re(g_{l,3}), Im(g_{l,3})]^T$, $\mathbf{w}$ is the receiver noise. $\mathbf{X} = [\mathbf{y}, \mathbf{M}] \in R^{2K \times (8N_T+1)}$ and $\mathbf{v} = [1, \mathbf{g}]^T \in R^{(8N_T+1)\times 1}$. Both $\mathbf{M}$ and $\mathbf{y}$ are treated unknown.

Assume a sequence of N_g training vectors of $\mathbf{g}$, For the ith training vector $\mathbf{g}_i$, the corresponding measurement is subject to the Tx random noise. Therefore, similar to the idea applied in the quadratic model, we measure multiple realization with the same training vector and calculate the averaged observation. Then, with either large time of averaging, we can regard $\mathbf{X}$ as time-invariant.

Let $\mathbf{R} = [\mathbf{r}_1, \cdots, \mathbf{r}_{N_g}]$ and $\mathbf{G} = [\mathbf{v}_1, \cdots, \mathbf{v}_{N_g}]$, we have

$$\mathbf{R} = \mathbf{X}\mathbf{G} \tag{5.6}$$

If $\mathbf{G}$ has a full row rank, the solution of (5.6) is

$$\begin{aligned} \mathbf{X} &= \mathbf{R}\mathbf{G}^{+} \\ &= \mathbf{R}\mathbf{G}^T(\mathbf{G}\mathbf{G}^T)^{-1} \end{aligned} \tag{5.7}$$

From the estimated $\mathbf{X}$, we can obtain the estimated $\mathbf{M}$ and $\mathbf{y}$. Then the optimal $\mathbf{g}$ is

$$\begin{aligned} \mathbf{g} &= -\mathbf{M}^{+}\mathbf{y} \\ &= -(\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T\mathbf{y} \end{aligned} \tag{5.8}$$

Considering the practical RF step attenuator, i.e., PE43703, the practical gains of the devices have both the attenuation and an attenuation-dependent phase. Assume there

is only one clustered tap used in the cancellation channel, an example of the training matrix $\mathbf{G}$ is

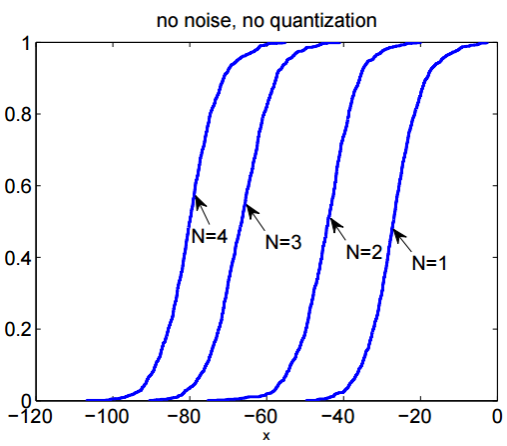
$$\mathbf{G} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0.15 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.05 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0.15 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.05 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0.15 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.05 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0.15 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.05 \end{bmatrix} \tag{5.9}$$

where each entry represents a real or imaginary component of a practical complex gain except of the first row. Value 1 corresponds to the 0 dB attenuation, value 0 corresponds to the maximum attenuation (32 dB in simulation) and value $0.15 + j0.05$ corresponds to the practical gain of 16 dB attenuation with attenuation-dependent phase. It can be seen the presented $\mathbf{G}$ is full rank and sparse, therefore it is a good choice for training.

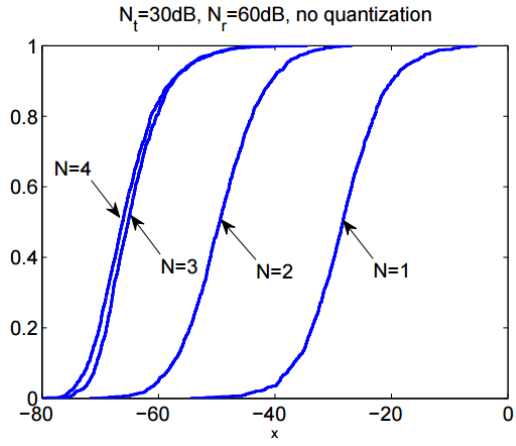
However, the optimal solution from (5.8) is not following the phase constrain in most cases so we can not directly implement the values to the system. One solution for this problem is using brute force search to find the solution for the following equation

$$g_{l,0} - jg_{l,1} - g_{l,2} + jg_{l,3} = \hat{g}_{l,0} - j\hat{g}_{l,1} - \hat{g}_{l,2} + j\hat{g}_{l,3} \tag{5.10}$$

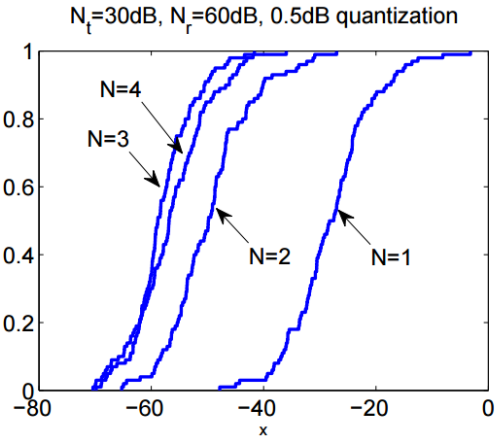
where left hand side term are from the optimal solution from (5.8), right hand side terms are the practical complex gains of the step attenuators following the step size and phase constrains.



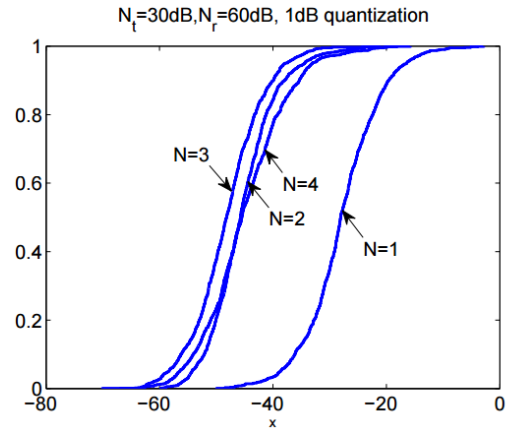
(a) $SNR_T = \infty$, $SNR_R = \infty$, step size=0 dB



(b) $SNR_T = 30$ dB, $SNR_R = 60$ dB, step size=0 dB



(c) $SNR_T = 30$ dB, $SNR_R = 60$ dB, step size=0.5 dB



(d) $SNR_T = 30$ dB, $SNR_R = 60$ dB, step size=1 dB

Figure 5.3: CDF of normalized power in dB of residual interference after self-interference cancellation

Fig.5.3 shows the CDF of the cancellation performance in dB for different scenarios. Fig.5.3(a) is the ideal case and there is no noise in the system and no quantization errors. Fig.5.3(b) shows the effect of the noise. With noise from the Tx/Rx chains, the cancellation

is reduced. The more clustered taps used in the cancellation channel, the more sensitive to the noise it is. Fig.5.3(c) and 5.3(d) show the impact of the step attenuators. When using the step attenuator, the optimal values require quantization and quantization errors are induced. We see the step size has a major impact and limits the cancellation performance.

5.3 Receiver Phase Noise Estimator

Considering the receiver phase noise, we can rewrite (5.5) as

$$\mathbf{r}_k = \mathbf{F}_k \mathbf{X} \mathbf{v}_k + \mathbf{w}_k \tag{5.11}$$

where $\mathbf{F}_k = diag[\mathbf{F}_{k,1}, \mathbf{F}_{k,2}, \cdots, \mathbf{F}_{k,N_e}]$ and $\mathbf{F}_{k,n} = \begin{bmatrix} \cos\theta_{k,n} & \sin\theta_{k,n} \\ -\sin\theta_{k,n} & \cos\theta_{k,n} \end{bmatrix}$. $\theta_{k,n}$ is the random phase noise of the nth samples, kth realization. Note that property $\mathbf{F}_k^{-1} = \mathbf{F}_k^T$ holds for all $\theta_{k,n}$.

Our objective is that given $\mathbf{r}_k$, $k = 1, 2, \cdots, K$, obtain an estimate of $\mathbf{v}_{opt}$ where

$$\begin{aligned} \mathbf{v}_{opt} &= \min_{\mathbf{v}, \|\mathbf{v}\|=1} \|\mathbf{X}\mathbf{v}\|^2 \\ &= \min_{\mathbf{v}, \|\mathbf{v}\|=1} \mathbf{v}^T \mathbf{X}^T \mathbf{X} \mathbf{v} \end{aligned} \tag{5.12}$$

Assuming that $\mathbf{w}_k \forall k$ is white Gaussian, the maximum likelihood estimation (MLE) of $\mathbf{X}$ and other parameters is given by the following:

$$\begin{aligned}\arg \min_{\mathbf{X},\theta_{k,n}\forall k\forall n} \sum_k \|\mathbf{r}_k - \mathbf{F}_k\mathbf{X}\mathbf{v}_k\|^2 \\ = \arg \min_{\mathbf{X},\theta_{k,n}\forall k\forall n} \sum_k \|\mathbf{F}_k^T\mathbf{r}_k - \mathbf{X}\mathbf{v}_k\|^2\end{aligned} \tag{5.13}$$

Let $\mathbf{R}_{new} = [\mathbf{F}_1^T\mathbf{r}_1, \cdots, \mathbf{F}_K^T\mathbf{r}_K]$, the MLE of $\mathbf{X}$ is

$$\begin{aligned}\hat{\mathbf{X}} =& \mathbf{R}\mathbf{V}^+ \\ =& \mathbf{R}\mathbf{V}^T(\mathbf{V}\mathbf{V}^T)^{-1}\end{aligned} \tag{5.14}$$

Notice that $\mathbf{F}_k$ is another unknown matrix, to find the estimation of $\mathbf{X}$, we need to estimate $\mathbf{F}_k$ first. The MLE of the angles is given by

$$\min_{\theta_{k,n}\forall k\forall n} \sum_k \|\mathbf{F}_k^T\mathbf{r}_k - \mathbf{E}\mathbf{V}^+\mathbf{v}_k\|^2 \tag{5.15}$$

Note that $\mathbf{V}^+\mathbf{v}_k$ is the kth column of the (symmetric) projection matrix $\mathbf{U} \doteq \mathbf{V}^+\mathbf{V}$ to be denoted by $\mathbf{u}_k$. Now let $\mathbf{u}_k = [u_{k,1}, \cdots, u_{k,1}]^T \in R^{K\times 1}$. The above problem becomes

$$\min_{\theta_{k,n}\forall k\forall n} \sum_k \|\mathbf{F}_k^T\mathbf{r}_k - \sum_{j=1}^{K} \mathbf{F}_j^T\mathbf{r}_j u_{k,j}\|^2 \tag{5.16}$$

or equivalently, for each n,

$$\min_{\theta_{k,n}\forall k} \sum_k \|\mathbf{F}_{k,n}^T\mathbf{r}_{k,n} - \sum_{j=1}^{K} \mathbf{F}_{j,n}^T\mathbf{r}_{j,n} u_{k,j}\|^2 \tag{5.17}$$

where $\mathbf{r}_k = \left[\mathbf{r}_{k,1}^T, \cdots, \mathbf{r}_{k,N}^T\right]^T$ and $\mathbf{r}_{k,n} \in R^{2\times 1}$. Furthermore, we can write the above problem as follows:

$$\min_{\theta_{k,n}\forall k} \sum_{k=1}^{K} \|e^{-j\theta_{k,n}} c_{k,n} - \sum_{l=1}^{K} u_{k,l} e^{-j\theta_{l,n}} c_{l,n}\|^2 \tag{5.18}$$

where $c_{k,n}$ is the complex scalar constructed from $\mathbf{r}_{k,n}$.

Define $\mathbf{c}_n = [c_{1,n}, \cdots, c_{K,n}]^T$ and $\mathbf{\Lambda}_n = diag[e^{-j\theta_{1,n}}, \cdots, e^{-j\theta_{K,n}}]$, we have

$$\min_{\theta_{k,n}\forall k} \|\mathbf{\Lambda}_n \mathbf{c}_n - \mathbf{U}\mathbf{\Lambda}_n \mathbf{c}_n\|^2 \tag{5.19}$$

If $\mathbf{U} = \mathbf{I}$, the above cost function is invariant to all angles. In general, the above cost is invariant to at least one of $\theta_{1,n}, \cdots, \theta_{K,n}$ for each n. So, without loss of generality, we can set $\theta_{1,n} = 0 \forall n$.

To simplify the notations, we will consider the following problem:

$$\min_{\theta_k \forall k} \|\mathbf{\Lambda}\mathbf{c} - \mathbf{U}\mathbf{\Lambda}\mathbf{c}\|^2 \tag{5.20}$$

where $\mathbf{c} \in C^{K\times 1}$ and $\mathbf{U} \in R^{K\times K}$ are given, and $\mathbf{\Lambda} = diag[1, e^{-j\theta_2}, \cdots, e^{-j\theta_K}]$ depends on the unknown angles. (We know that both $\mathbf{c}$ and $\mathbf{\Lambda}$ are dependent on n while $\mathbf{U}$ is not.) A Newton's iterative algorithm to search for the optimal angles that minimize the cost $J = \|\mathbf{\Lambda}\mathbf{c} - \mathbf{U}\mathbf{\Lambda}\mathbf{c}\|^2$ is derived next.

Let $\boldsymbol{\theta} = [0, \theta_2, \cdots, \theta_K]^T$ and $\mathbf{\Lambda} = \mathbf{\Lambda}(\boldsymbol{\theta})$. Then, the first order Taylor series expansion of $\mathbf{\Lambda}(\boldsymbol{\theta})$ around $\boldsymbol{\theta} = \boldsymbol{\theta}_i$ is

$$\mathbf{\Lambda}(\boldsymbol{\theta}) \approx \mathbf{\Lambda}_i - j\mathbf{\Lambda}_i \mathbf{D}_{\Delta\boldsymbol{\theta}_i} \tag{5.21}$$

where $\mathbf{\Lambda}_i = \mathbf{\Lambda}(\boldsymbol{\theta}_i)$, $\Delta\boldsymbol{\theta}_i = \boldsymbol{\theta} - \boldsymbol{\theta}_i$, and $\mathbf{D_a} = diag(\mathbf{a})$ with the first diagonal element set to zero. It follows that

$$\begin{aligned} J \approx & \|\mathbf{\Lambda}_i\mathbf{c} - j\mathbf{\Lambda}_i\mathbf{D}_{\Delta\boldsymbol{\theta}_i}\mathbf{c} - \mathbf{U}\mathbf{\Lambda}_i\mathbf{c} + j\mathbf{U}\mathbf{\Lambda}_i\mathbf{D}_{\Delta\boldsymbol{\theta}_i}\mathbf{c}\|^2 \\ = & \|(\mathbf{I}-\mathbf{U})\mathbf{\Lambda}_i\mathbf{c} - j(\mathbf{I}-\mathbf{U})\mathbf{\Lambda}_i\mathbf{D_c}\Delta\boldsymbol{\theta}_i\|^2 \end{aligned} \tag{5.22}$$

Here, we have used $\mathbf{D}_{\Delta\boldsymbol{\theta}_i}\mathbf{c} = \mathbf{D_c}\Delta\boldsymbol{\theta}_i$ where the first diagonal elements of $\mathbf{D}_{\Delta\boldsymbol{\theta}_i}$ and $\mathbf{D_c}$ are zero.

Note that the (minimum norm) solution to $\min_{\mathbf{r}} \|\mathbf{a} - \mathbf{B}\mathbf{r}\|^2$ where $\mathbf{r}$ is real but $\mathbf{a}$ and $\mathbf{B}$ are complex is given by $\mathbf{r} = Re^{+}\{\mathbf{B}^H\mathbf{B}\}Re\{\mathbf{B}^H\mathbf{a}\}$. Also note that $(\mathbf{AB})^+ = \mathbf{B}^+\mathbf{A}^+$, and $\mathbf{P}^+ = \mathbf{P}$ if $\mathbf{P}$ is a projection matrix.

Then, it follows that the optimal (minimum norm) solution of $\Delta\boldsymbol{\theta}_i$ is given by

$$\begin{aligned} \Delta\boldsymbol{\theta}_i = & Re^{+}\{\mathbf{D_c}^H\mathbf{\Lambda}_i^H(\mathbf{I}-\mathbf{U})\mathbf{\Lambda}_i\mathbf{D_c}\}Re\{-j\mathbf{D_c}^H\mathbf{\Lambda}_i^H(\mathbf{I}-\mathbf{U})\mathbf{\Lambda}_i\mathbf{c}\} \\ = & Re^{+}\{\mathbf{D_c}^H\mathbf{\Lambda}_i^H(\mathbf{I}-\mathbf{U})\mathbf{\Lambda}_i\mathbf{D_c}\}Im\{\mathbf{D_c}^H\mathbf{\Lambda}_i^H(\mathbf{I}-\mathbf{U})\mathbf{\Lambda}_i\mathbf{c}\} \end{aligned} \tag{5.23}$$

The Newton's iteration is

$$\theta_{i+1} = \theta_i + \Delta\theta_i \tag{5.24}$$

The above steps show the method to estimate the angle for each phase. A necessary condition for the angles to be obtained exactly by MLE is $K \geq 2M - 1$.

Fig.5.4 proves the validation of the proposed phase estimator. From the figure, we can see the phase noise estimator provides around 15dB improvement of the self-interference cancellation. Noth that the simulation is done under the assumption that other RF impairments are also existing which will affect the estimation result. If we only consider the

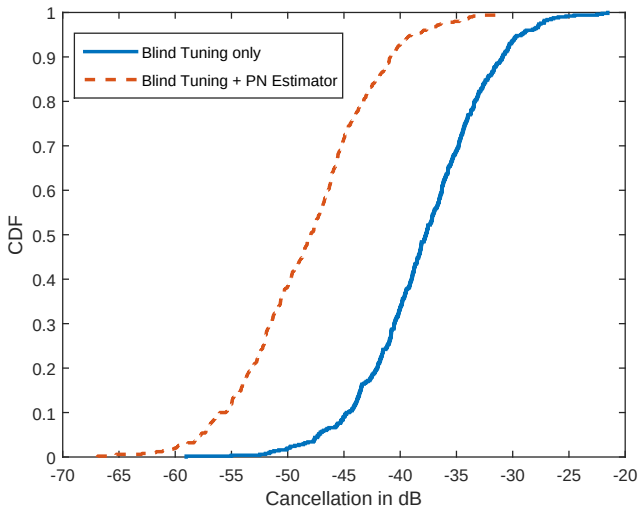


Figure 5.4: Improvement of the amount of self-interference cancellation in dB using the phase noise estimator for affine model blind tuning algorithm.

receiver phase noise and no other impairment, the proposed estimator can estimate the phase noise correctly and compensate all the phase noise according to our simulation.

5.4 Adaptive Blind Tuning Algorithm

In this section, without loss of the generality, the analog cancellation channel is assumed to be the structure shown in Fig.5.5. In the RF canceller, two attenuators and one 90 degree phase shifter forms a tap. N_T represents the number of taps used in the cancellation channel and there is a fixed delay T between tap and tap. So totally, the number of attenuator is $N_g = 2 * N_T$.

Since the canceller works in the RF domain, the exact input RF signal of the cancellation channel is unknown. So we propose to use the blind tuning to find the optimal setting of the attenuator based on an affine system model. By tuning the known input of

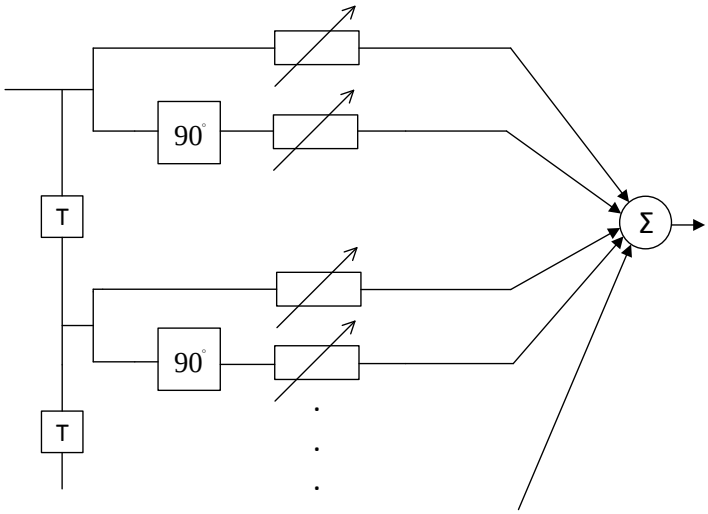


Figure 5.5: Structure of RF Canceller

the attenuation $\mathbf{g} = [g_{1,2}, g_{1,2}, \cdots, g_{N_T,1}, g_{N_T,2}]^T \in R^{N_g \times 1}$, we can have the corresponding output residual $\mathbf{r} = [\mathbf{r}[1]^T, \mathbf{r}[2]^T, \cdots, \mathbf{r}[N_r]^T]^T \in R^{2N_r \times 1}$ where $\mathbf{r}[n] = [r_I[n], r_Q[n]]^T$ for $n = 1, 2, \cdots, N_r$ and it can be written as.

$$\mathbf{r} = \mathbf{y}_{int} + \mathbf{y}_{can} + \mathbf{w}_{thermal} \tag{5.25}$$

where $\mathbf{y}_{int} = [\mathbf{y}_{R,BB}[1]^T, \mathbf{y}_{R,BB}[2]^T, \cdots, \mathbf{y}_{R,BB}[N_r]^T]^T \in R^{2N_r \times 1}$ is the unknown output of the self-interference channel, $\mathbf{y}_{can} = [\mathbf{y}_{can,BB}[1]^T, \mathbf{y}_{can,BB}[2]^T, \cdots, \mathbf{y}_{can,BB}[N_r]^T]^T$ is the output of the cancellation channel, $\mathbf{w}$ represents the thermal noise. For the RF canceller, a linear relationship holds between the output and input. So considering the phase noise and I/Q imbalance, similar to (2.21), we have

$$\mathbf{y}_{can} = \mathbf{T}_g \mathbf{x}_{T,BB} + \mathbf{w}_{can} \tag{5.26}$$

where $\mathbf{T}_g$ represents the overall cancellation channel response which is a linear function of $\mathbf{g}$, $\mathbf{x}_{T,BB}$ is the input signal which we do not need to know what exact it is, $\mathbf{w}_{can}$ is again the perturbation introduced by the phase noise. Furthermore, we can rewrite $\mathbf{T}_g\mathbf{x}_{T,BB} = \mathbf{M}\mathbf{g}$ where linear operator $\mathbf{M} \in R^{2N_r \times N_g}$ is governed by the RF canceller structure as well as the source signal $\mathbf{x}_{T,BB}$. In the training period, we assume the source signal $\mathbf{x}_{T,BB}$ is fixed. For $k = 1, 2, \cdots, K$ different settings of attenuation $\mathbf{g}$, we can rewrite (5.25) as

$$\begin{aligned} \mathbf{r}_k =& \mathbf{y}_{int,k} + \mathbf{M}\mathbf{g}_k + \mathbf{w}_{can,k} + \mathbf{w}_{thermal,k} \\ =& \mathbf{X}\mathbf{v}_k + \mathbf{w}_k \end{aligned} \tag{5.27}$$

where

$$\mathbf{X} = [\mathbf{y}_{int}, \mathbf{M}] \tag{5.28}$$

and

$$\mathbf{v}_k = \begin{bmatrix} 1 \\ \\ \mathbf{g}_k \end{bmatrix} \tag{5.29}$$

$\mathbf{w}_k = \mathbf{w}_{int,k} + \mathbf{w}_{can,k} + \mathbf{w}_{thermal,k} \in R^{2N_r \times 1}$ is unknown zero mean noise which has three components, perturbation from interference channel due to phase noise, perturbation from cancellation channel due to phase noise and thermal noise. Then, let $\mathbf{R} = [\mathbf{r}_1, \cdots, \mathbf{r}_K]$, $\mathbf{V} = [\mathbf{v}_1, \cdots, \mathbf{v}_K]$ and $\mathbf{W} = [\mathbf{w}_1, \cdots, \mathbf{w}_K]$. We have

$$\mathbf{R} = \mathbf{X}\mathbf{V} + \mathbf{W} \tag{5.30}$$

If $\mathbf{V}$ has a full row rank which means $K \geq N_g + 1$, the LS estimation for (5.30) is

$$\hat{\mathbf{X}} = \mathbf{R}\mathbf{V}^{+} = \mathbf{R}\mathbf{V}^{T}(\mathbf{V}\mathbf{V}^{T})^{-1} \tag{5.31}$$

where + represents the pseudo inverse of a matrix. For the choice of V, assume $N_g = 4$, an example is

$$\mathbf{V} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{5.32}$$

for the setting of the attenuator, 1 represents the 0dB attenuation, 0 represents the maximum attenuation. From $\hat{\mathbf{X}}$, we can extract $\hat{\mathbf{y}}_{int}$ and $\hat{\mathbf{M}}$. Then, the optimal $\mathbf{g}_{LS}$ for this model is obtained by

$$\mathbf{g}_{LS} = -\hat{\mathbf{M}}^{+}\hat{\mathbf{y}}_{int} = -(\hat{\mathbf{M}}^{T}\hat{\mathbf{M}})^{-1}\hat{\mathbf{M}}^{T}\hat{\mathbf{y}}_{int} \tag{5.33}$$

We call the above algorithm blind tuning (LS estimator). Notice that, in (5.27) $\mathbf{y}_{int,k}$ is subject to the random phase noise. In the LS estimation, we regard it as time invariant and lump the perturbations into the noise $\mathbf{w}_k$. However, due to the impact of the phase noise, this noise can be relative large and make the estimation result not promising. Therefore, we propose a recursive least square (RLS) [68] based adaptive blind tuning algorithm.

In the adaptive algorithm, we regard (5.27) as the system model and detailed structure is as follow:

1. First, we use $K = N_g + 1$ different setting of $\mathbf{g}$ to train the system (an example is given in (5.32)) and get corresponding residual observations $\mathbf{r}$.

2. With the knowledge of $\mathbf{g}$ and $\mathbf{r}$ from the training, we can find the estimation of the coefficient matrix $\hat{\mathbf{X}}$ through (5.29)-(5.31). We regards it as the initial point of the iteration and denote $\mathbf{X}^{(0)} = \hat{\mathbf{X}}$.

3. Once we know $\mathbf{X}^{(0)}$, we regard it fixed and estimate $\mathbf{g}$. We can extract $\mathbf{y}_{int}^{(0)}$ and $\mathbf{M}^{(0)}$ using (5.28) and calculate $\mathbf{g}^{(1)} = -\mathbf{M}^{(0)+}\mathbf{y}_{int}^{(0)}$ following (5.33).

4. Next, we regard $\mathbf{g}^{(1)}$ fixed and apply the setting of $\mathbf{g}^{(1)}$ we can have a new observation of the residual $\mathbf{r}^{(1)}$.

5. With the knowledge of $\mathbf{X}^{(0)}$, $\mathbf{g}^{(1)}$ and $\mathbf{r}^{(1)}$, we update $\mathbf{X}^{(0)}$ to $\mathbf{X}^{(1)}$ row by row. Denote $\mathbf{X}^{(0)} = [\mathbf{x}_1^{(0)^T}; \mathbf{x}_2^{(0)^T}; \cdots; \mathbf{x}_{2N_r}^{(0)^T}]$. For each row, we have $r_i = \mathbf{x}_i{}^T\mathbf{v}$. Regard $\mathbf{x}_i{}^T$ as the weight of a system, $\mathbf{v}$ as the input and r_i as the output residual, we can apply RLS algorithm to update $\mathbf{x}_i{}^T$ with the new input and observation. After updated all the $\mathbf{x}_i{}^T$ for $i = 1, \cdots 2Nr$, we can get the updated matrix $\mathbf{X}^{(1)}$.

6. After we get the updated $\mathbf{X}^{(1)}$, we go back to the step 3 and repeat the step 3 to step 5 to update $\mathbf{X}$ until it is converged (in other word, $\mathbf{g}$ is converged).

7. With the converged $\mathbf{g}$, we find the optimal setting of the attenuator which provides the best self-interference cancellation performance and minimize the output residual $\mathbf{r}$.

The detailed procedure presented by the mathematical expression is shown below as Algorithm 1.

Algorithm 1 RLS-based Algorithm

Input:

The parameters $\hat{\mathbf{X}}^{(0)}$; $\mathbf{v}^{(1)}$; $\mathbf{r}^{(1)}$; forget factor λ;
Accuracy threshold ε; Maximum iteration N.

$\Phi^{(0)^{-1}} = \delta^{-1}\mathbf{I}$,
$n = 1$,
while $(\Delta > \varepsilon \;\&\&\; n \leq N)$ **do**
  $\mathbf{k}^{(n)} = \frac{\Phi^{(n-1)^{-1}}\mathbf{v}^{(n)}}{\lambda + \mathbf{v}^{(n)^T}\Phi^{(n-1)^{-1}}\mathbf{v}^{(n)}}$,
  for $i = 1, 2, \cdots, 2N_r$
  $\xi_i^{(n)} = r_i^{(n)} - \mathbf{x}_i^{(n-1)^T}\mathbf{v}^{(n)}$,
  $\mathbf{x}_i^{(n)} = \mathbf{x}_i^{(n-1)} + \xi_i^{(n)}\mathbf{k}^{(n)}$,
  end for
  $\hat{\mathbf{X}}^{(n)} = [\mathbf{x}_1^{(n)^T}; \mathbf{x}_2^{(n)^T}; \cdots; \mathbf{x}_{2N_r}^{(n)^T}]$,
  $\Phi^{(n)^{-1}} = \lambda^{-1}\Phi^{(n-1)^{-1}} - \lambda^{-1}\mathbf{k}^{(n)}\mathbf{v}^{(n)^T}\Phi^{(n-1)^{-1}}$,
  extract $\hat{\mathbf{y}}_{int}^{(n)}$ and $\hat{\mathbf{M}}^{(n)}$ from $\hat{\mathbf{X}}^{(n)}$,
  $\mathbf{g}^{(n+1)} = -\hat{\mathbf{M}}^{(n)+}\hat{\mathbf{y}}_{int}^{(n)}$,
  $\mathbf{v}^{(n+1)} = [1, \mathbf{g}^{(n+1)^T}]^T$
  $\Delta = \frac{\|\mathbf{v}^{(n+1)} - \mathbf{v}^{(n)}\|}{\|\mathbf{v}^{(n)}\|}$,
  observe $\mathbf{r}^{(n+1)}$ based on $\mathbf{g}^{(n+1)}$,
  $n = n + 1$.
end while
return $\mathbf{g}_{RLS} = \mathbf{g}^{(n)}$

5.5 Simulation

In this section, simulation results are shown to compare the performance of the self-interference cancellation. In most of the precious work, the authors use the baseband discrete time model to model and simulate the full-duplex system. For example, they model the interference channel as

$$y_i[n] = x[n] \otimes h_i[l] \tag{5.34}$$

where $x[n]$ is the samples of the input signal, $h[l]$ is the baseband impulse response of the interference channel, $\otimes$ denotes the convolution operation. The sampling rate T_s of $x[n]$ and $h[l]$ satisfies the Nyquist sampling theorem. Assume the interested signal has a double side bandwidth W, then $T_s \geq \frac{1}{W}$.

However, the practical RF interference channel is continues in time and can be express as

$$\tilde{y}_i(t) = \sum_{i=0}^{I-1} a_i \tilde{x}(t - \tau_i) \tag{5.35}$$

where $\tilde{x}(t)$ is the RF input of the interference channel, $\tilde{y}_i(t)$ is the Rf output of interference channel. Using the real valued system model, the baseband equivalent of (5.35) is

$$\mathbf{y}_i(t) = \sum_{i=0}^{I-1} a_i \begin{bmatrix} \cos(2\pi f_c \tau_i) & \sin(2\pi f_c \tau_i) \\ -\sin(2\pi f_c \tau_i) & \cos(2\pi f_c \tau_i) \end{bmatrix} \mathbf{x}(t - \tau_i) \tag{5.36}$$

where $I = 100$ is number of the multipaths, a_i is the real valued attenuation factor for each path and is modeled as $a_0 = \frac{\epsilon}{d^2}$ and for $i = 1, \cdots, I-1$

$$a_i = \frac{\epsilon \alpha_i}{(d + c\tau_i)^2} \tag{5.37}$$

where $c = 3 \times 10^8$ m/s is the speed of light; $d = 0.3$ m represents the distance between the transmitter and receive antennas; $0 \leq \tau_i \leq \tau_{max}$ (uniform random). When $\tau_{max} = 10$ ns it represents the maximum delay of the interference channel is 3 m and $\tau_{max} = 100$ ns corresponding to 30 m; $0 \leq \alpha_i \leq 1$ (uniform random) and $\epsilon = 0.01$.

For the all analog cancellation channel, we use the structure shown in Fig.5.5. The relationship between the RF input and output can be expressed as

$$\tilde{y}_c(t) = \sum_{n=0}^{N_T-1} \sum_{l=0}^{1} g_{n,l} \tilde{x}(t - nT - l\eta) \tag{5.38}$$

where $\tilde{y}_c(t)$ is the RF output of the cancellation channel and N_T is the number of the taps used in the cancellation channel. Similarly to (5.36), the real form baseband equivalent of (5.38) is expressed as

$$\mathbf{y}_c(t) = \sum_{n=0}^{N_T-1} \sum_{l=0}^{1} g_{n,l} \mathbf{T}_{n,l} \mathbf{x}(t - nT - l\eta) \tag{5.39}$$

where

$$\mathbf{T}_{n,l} = \begin{bmatrix} \cos(2\pi f_c(nT + l\eta)) & \sin(2\pi f_c(nT + l\eta)) \\ -\sin(2\pi f_c(nT + l\eta)) & \cos(2\pi f_c(nT + l\eta)) \end{bmatrix} \tag{5.40}$$

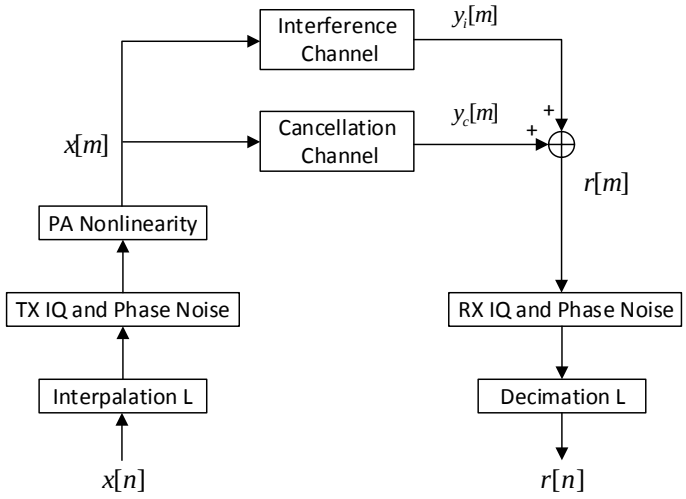


Figure 5.6: Virtual simulation system structure diagram for analog cancellation

In the simulation, we choose the carrier frequency $f_c = 2.4$ GHz; the bandwidth of interest $W = 40$ Mhz; the fixed delay between two taps T is chosen to be a small faction of $\frac{1}{W}$ which varies in the simulation depends on the selection of τ_{max} and the equivalent delay of the 90 degree phase shifter $\eta = \frac{1}{4f_c}$. The baseband sampling rate is choose as $f_s = \frac{1}{T_s} = 2W$ and we use index n to represent the discrete baseband signal as $x[n] = x(nT_s)$.

Since in the analog channel, the random delay τ_i is continues in time which is uniform distributed in the time interval $(0 - \tau_{max})$, to make the simulation more accurate, we use a higher sampling rate $f_L = \frac{1}{T_L} = \frac{L}{T_s}$, $L = 500$ to mimic the continuous analog cancellation and interference channel. For the higher sampling rate, we use index m to represent the samples with $x[m] = x(mT_L)$. Fig.5.6 shows the diagram of the simulation for analog cancellation. Once the baseband signal $x[n]$ is generated, we apply a interpolator with factor of L and add all the impairments in the transmit chain to get the signal $x[m]$ with higher sampling rate. Then apply $x[m]$ to (5.36) and (5.39) by replacing $\mathbf{x}(t - \tau_i)$ to $\mathbf{x}[m - m_i]$ where $m_i = round(\frac{\tau_i}{T_L})$ and replacing $\mathbf{x}(t - nT - l\eta)$ to $\mathbf{x}[m - round(\frac{nT + l\eta}{T_L})]$ to get the received interference $\mathbf{y}_i[m]$ and cancellation signal $\mathbf{y}_c[m]$. Combining interference and cancellation signals, we can get the residual $r[m]$ with higher sample frequency. Next, passed by a decimator with factor L, we finally get the residual signal $r[n]$ after the cancellation.

For the hybrid cancellation, the virtual diagram of the simulation procedure is shown in Fig.5.7. We apply the same structure to simulation the transmit chain and get $x[m]$. First we assume there is no cancellation channel, same as the analog cancellation, we can get the received interference $\mathbf{y}_i[m]$ using (5.36). Then we simulated the auxiliary receive chain to get the estimation of the baseband signal $\hat{x}[n]$ therefore we can get $\hat{\mathbf{X}}$ and use

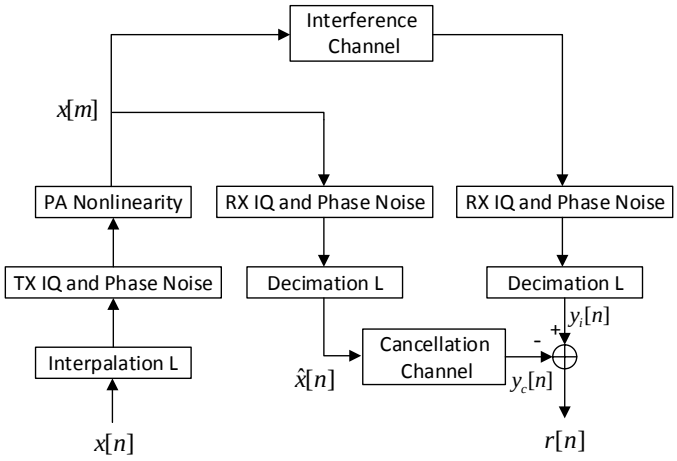


Figure 5.7: Virtual simulation system structure diagram for hybrid cancellation

(3.8) to estimate impulse response of the interference channel using finite impulse response filter. In the simulation, we generated one random receive phase noise and introduced it to both the receiver chain and auxiliary to get $\mathbf{y}_i[m]$ and $\hat{x}[n]$. Note the order of the impulse response of the interference channel can be calculate as $L = \frac{\tau_{max}}{T_s}$. In the simulation for hybrid cancellation, the maximum delay of the interference channel τ_{max} is set to be 100 ns, which means the equivalent

In the following simulations, the I/Q imbalance is set to be uniform distribution with $\mathcal{U}(-0.05, 0.05)$ and the phase noise to modeled as a Gaussian distribution with $\mathcal{N}(0, 0.05)$ if it is considered. We followed (2.7) and (2.17) to introduce the I/Q imbalance and phase noise in the transmitter and receiver chain. In each simulation, we have simulated 1000 random realizations (fixed input signal with random I/Q imbalance and/or random phase noise) to the statistic of the cancellation performance.

First, we want to demonstrate the impact of the phase noise. In Fig.5.8, we simulated a system considering I/Q imbalance only and considering both I/Q imbalance

and phase noise. We tested the performance of the analog blind tuning algorithm for these two scenarios. It can be seen that only considering the I/Qimbalnce the LS estimator can provide a good cancellation performance. However, the cancellation performance will reduce heavily if we also include the phase noise in the full-duplex system. This shows the regular real model and widely linear based LS estimator can not handel the effect of the phase noise properly.

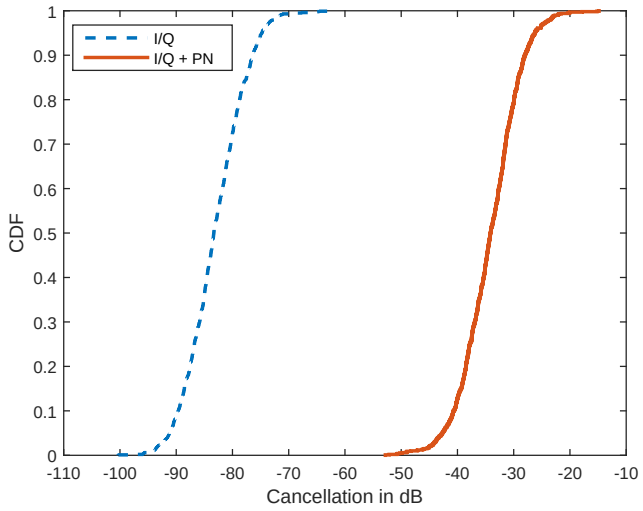


Figure 5.8: CDF of the amount of self-interference cancellation in dB using the blind tuning algorithm.

Next, we demonstrate the impact of the I/Q imbalance. In Fig.5.9, we only considered the phase noise and ignored the I/Q imbalance first. It can be seen the phase noise estimator shown in the Appendix A can improve the cancellation performance by 15dB since it can only compensate the phase noise at the receiver end and the phase noise at the transmitter side will limit the performance. Furthermore, once the I/Q imbalance is also considered, the performance of the phase noise estimator will become even worse due to

the cascaded mathematical relation of the impact of the phase noise and I/Q imbalance. Combined with the previous simulation result, We know it is necessary and important to consider the I/Q imbalance and phase noise together when analyzing the full-duplex system.

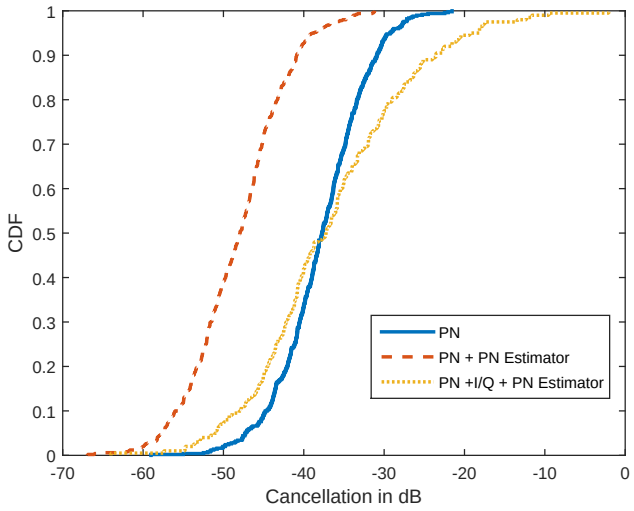


Figure 5.9: CDF of the amount of self-interference cancellation in dB using the phase noise estimator.

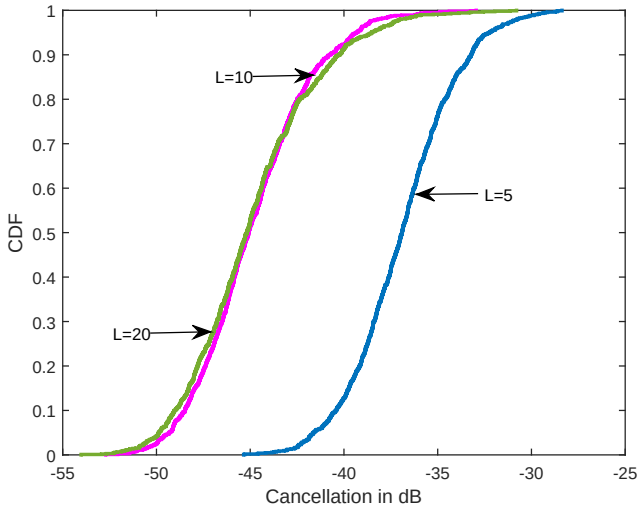


Figure 5.10: CDF of the amount of self-interference cancellation in dB using the hybrid cancellation method.

In Fig.5.10, we simulated the statistic cancellation performance for the hybrid cancellation method with different Lth order FIR filter. In the simulation, we considered both phase noise and I/Q imbalance and picked the choices of $L = 5, 10, 20$. From the figure, we can see that with the increasing of the order L of the FIR filter, the baseband canceler provides a better cancellation performance. When choose $L = 10$, the hybrid cancellation method can reduce the self-interference by an average of 45dB. Also, beyond $L = 10$, the performance will not improve obviously when further increasing the number of L which is determined by the maximum delay of the interference channel. Since $L = 10$ already reaches the maximum order of the equivalent baseband impulse response of the interference channel,there is no need to choice a larger L with will introduce redundancy to the system.

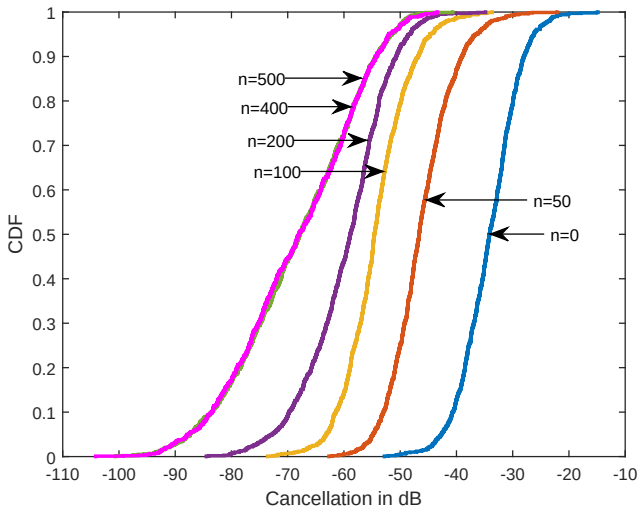


Figure 5.11: CDF of the amount of self-interference cancellation in dB. Both I/Q imbalance and phase noise are considered. $\tau_{max} = 10$ ns; $T = \frac{1}{10W}$.

Finally, we want to show the feasibility of our proposed analog adaptive blind tuning algorithm on overcoming the impact of both I/Q imbalance and phase noise. In the

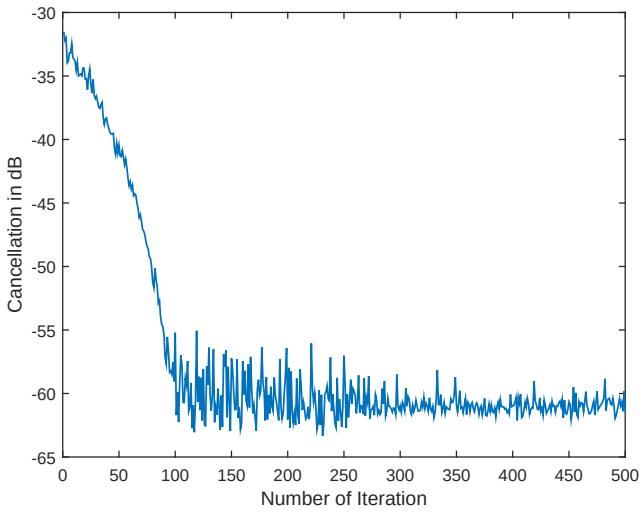


Figure 5.12: An example of the trend of convergence for the adaptive blind tuning algorithm. $\tau_{max} = 10$ ns; $T = \frac{1}{10W}$.

simulation, we considered both I/Q imbalance and phase noise in our system and the level of which are the same as what is used in the hybrid cancellation simulation. We consider two scenario of the maximum delay in the interference channel τ_{max}. In Fig.5.11-5.12, we choose $\tau_{max} = 10$ ns, the fix delay T between the taps in the RF canceller is set to be $T = \frac{1}{10W}$ and in the RF canceller we use $N = 4$ taps. Fig.5.11 shows the statistic results of the cancellation performance with 1000 random realizations after n times iterations. From the figure we can see, when $n = 0$ which means the initial case that achieved by the blind tuning method, the average cancellation is around 35dB since it is limited by the phase noise. After implementing the adaptive algorithm, with the increasing of the number of the iteration n, the cancellation performance becomes better and better. Notice that the curves representing 400 iterations and 500 iterations almost overlap with each other. This shows that the cancellation performance becomes statistic stable after 400 times iterations and provides an average 65dB self-interference cancellation which is still 45dB for the worst

case. By comparing the curves of $n = 0$ and $n = 400$, we see the proposed adaptive blind tuning algorithm improves the cancellation performance by around 35dB. In Fig.5.12, we pick an example from the 1000 realizations to demonstrate the trend of the cancellation performance changing with the n increases. It shows that the proposed adaptive algorithm converges fast and it is robust to the nonlinear components after the convergence since in every iteration, we added random phase noise and thermal noise to the system in our simulation.

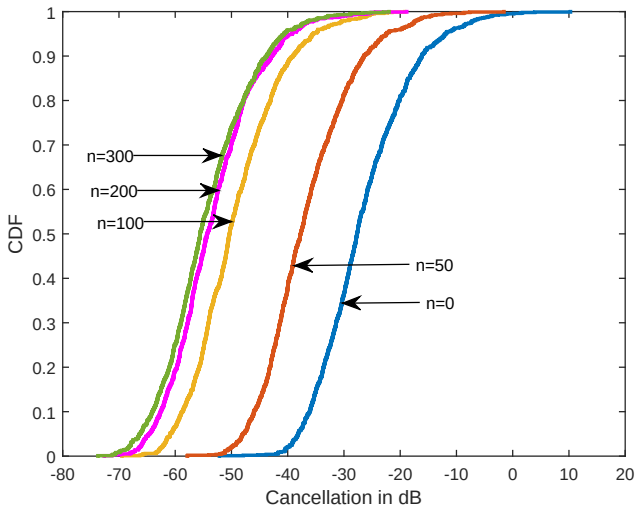


Figure 5.13: CDF of the amount of self-interference cancellation in dB. Both I/Q imbalance and phase noise are considered. $\tau_{max} = 100$ ns; $T = \frac{1}{5W}$.

In Fig.5.13-5.14, we choose $\tau_{max} = 100$ ns, $T = \frac{1}{5W}$ and $N = 4$. Under this assumption, we simulated a exactly same system as the hybrid cancellation and we want to compare the performance of these two self-interference cancellation methods. From the figure we can see that when the maximum delay of the interference channel is larger, we can still achieve a 55dB self-interference cancellation with very few number of taps. Also, after

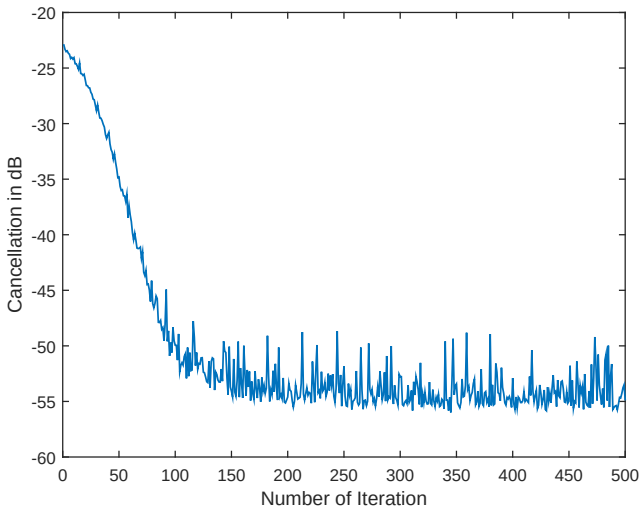


Figure 5.14: An example of the trend of convergence for the adaptive blind tuning algorithm. $\tau_{max} = 100$ ns; $T = \frac{1}{5W}$.

200 iterations, the performance does not have much further improvement which means the converging speed becomes faster.

Compared to the hybrid cancellation method in Fig.5.10 which the average cancellation is around 45dB, the adaptive blind tuning method can provide a better cancellation performance under the same assumption. Also, recall that the adaptive blind tuning is an analog cancellation method which works before the ADC. It can effectively reduce the risk of the ADC saturation problem and has congenital advantage compare to the hybrid cancellation. Therefore, our proposed adaptive blind tuning algorithm is a promising self-interference cancellation method which can handle the I/Q imbalance and phase noise simultaneously.

Chapter 6

Conclusions

In this dissertation, a full picture of the theories proposed for the self-interference cancellation in the full-duplex radio is presented. An insight of the practical RF impairments and their impacts are given. Especially, we reviewed the impact of the I/Q imbalance and phase noise for the full-duplex system. We demonstrated the importance of considering both I/Q imbalance and phase noise in modeling the full-duplex system.

Real valued model is presented to solve the nonlinearity problem induced by the I/Q imbalance. We compared the proposed model with the widely linear model which is also proposed for the same purpose. We have demonstrated both of these two methods are valid and mathematically they are equivalent. However, the real valued model has the advantage on the computational complexity and therefore is a better choice.

For the self-interference cancellation structure, this dissertation mainly focused on the analog self-interference cancellation method. Two system models: quadratic model and affine model are proposed to model the relationship between the output and input

of the full-duplex system. Blind tuning algorithm which includes training and optimizing procedures is used to find the optimal setting of the attenuators that minimized the residual interference after the self-interference cancellation.

To reduce the estimation error in the training period, an optimal training training matrix selection algorithm is propose. It is shown that with the optimal training matrix, the estimation error is reduced. Corresponding, the optimal solution found in the optimizing period is more accurate and therefore the self-interference cancellation performance achieved is better.

Furthermore, practical RF step attenuators are considered to be used in the cancellation channel. Due to the step size constrain (quantization constrain), the self-interference cancellation is limited. We ran multiple simulations to demonstrate the impacts of the quantization constrain with different step sizes. A brute force searching method is also presented to solve the attenuation-dependent phase problem induced by the RF step attenuators.

Last, we provided a RLS based adaptive blind tuning algorithm which is robust to both the I/Q imbalance and phase noise. Simulation results show the proposed algorithm can give an average of 65dB self-interference cancellation in the analog domain with both I/Q imbalance and phase noise considered in the system. Therefore, the proposed adaptive blind tuning algorithm can obviously reduce the self-interference before it is converted to the digital signal and solve the ADC saturation problem.

Bibliography

[1] Jung Il Choi, Mayank Jain, Kannan Srinivasan, Phil Levis, and Sachin Katti. Achieving single channel, full duplex wireless communication. In *Proceedings of the Sixteenth Annual International Conference on Mobile Computing and Networking*, MobiCom '10, pages 1–12, New York, NY, USA, 2010. ACM.

[2] Bo Chen, Vivek Yenamandra, and Kannan Srinivasan. Flexradio: Fully flexible radios and networks. In *Proceedings of the 12th USENIX Conference on Networked Systems Design and Implementation*, NSDI'15, pages 205–218, Berkeley, CA, USA, 2015. USENIX Association.

[3] Mohammad A. Khojastepour, Karthik Sundaresan, Sampath Rangarajan, Xinyu Zhang, and Sanaz Barghi. The case for antenna cancellation for scalable full-duplex wireless communications. In *Proceedings of the 10th ACM Workshop on Hot Topics in Networks*, HotNets-X, pages 17:1–17:6, New York, NY, USA, 2011. ACM.

[4] Achaleshwar Sahai, Gaurav Patel, and Ashutosh Sabharwal. Pushing the limits of full-duplex: Design and real-time implementation. *CoRR*, abs/1107.0607, 2011.

[5] E. Tsakalaki, E. Foroozanfard, E. de Carvalho, and G. F. Pedersen. A 2-order mimo full-duplex antenna system. In *The 8th European Conference on Antennas and Propagation (EuCAP 2014)*, pages 2546–2550, April 2014.

[6] E. Foroozanfard, O. Franek, A. Tatomirescu, E. Tsakalaki, E. D. Carvalho, and G. F. Pedersen. Full-duplex mimo system based on antenna cancellation technique. *Electronics Letters*, 50(16):1116–1117, July 2014.

[7] Y. Hua, P. Liang, Y. Ma, A. C. Cirik, and Q. Gao. A method for broadband full-duplex mimo radio. *IEEE Signal Processing Letters*, 19(12):793–796, Dec 2012.

[8] B. Debaillie, D. J. van den Broek, C. Lavn, B. van Liempd, E. A. M. Klumperink, C. Palacios, J. Craninckx, B. Nauta, and A. Prssinen. Analog/rf solutions enabling compact full-duplex radios. *IEEE Journal on Selected Areas in Communications*, 32(9):1662–1673, Sept 2014.

[9] E. Everett, M. Duarte, C. Dick, and A. Sabharwal. Empowering full-duplex wireless communication by exploiting directional diversity. In *2011 Conference Record of the*

Forty Fifth Asilomar Conference on Signals, Systems and Computers (ASILOMAR), pages 2002–2006, Nov 2011.

[10] E. Everett, A. Sahai, and A. Sabharwal. Passive self-interference suppression for full-duplex infrastructure nodes. *IEEE Transactions on Wireless Communications*, 13(2):680–694, February 2014.

[11] Aimin Tang and Xudong Wang. Balanced rf-circuit based self-interference cancellation for full duplex communications. *Ad Hoc Netw.*, 24(PA):214–227, January 2015.

[12] E. Ahmed, A. M. Eltawil, Z. Li, and B. A. Cetiner. Full-duplex systems using multireconfigurable antennas. *IEEE Transactions on Wireless Communications*, 14(11):5971–5983, Nov 2015.

[13] T. Riihonen, S. Werner, and R. Wichman. Mitigation of loopback self-interference in full-duplex mimo relays. *IEEE Transactions on Signal Processing*, 59(12):5983–5993, Dec 2011.

[14] Mayank Jain, Jung Il Choi, Taemin Kim, Dinesh Bharadia, Siddharth Seth, Kannan Srinivasan, Philip Levis, Sachin Katti, and Prasun Sinha. Practical, real-time, full duplex wireless. In *Proceedings of the 17th Annual International Conference on Mobile Computing and Networking*, MobiCom '11, pages 301–312, New York, NY, USA, 2011. ACM.

[15] D. Korpi, L. Anttila, and M. Valkama. Feasibility of in-band full-duplex radio transceivers with imperfect rf components: Analysis and enhanced cancellation algorithms. In *2014 9th International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*, pages 532–538, June 2014.

[16] D. Korpi, L. Anttila, V. Syrjl, and M. Valkama. Widely linear digital self-interference cancellation in direct-conversion full-duplex transceiver. *IEEE Journal on Selected Areas in Communications*, 32(9):1674–1687, Sept 2014.

[17] D. Korpi, S. Venkatasubramanian, T. Riihonen, L. Anttila, S. Otewa, C. Icheln, K. Haneda, S. Tretyakov, M. Valkama, and R. Wichman. Advanced self-interference cancellation and multiantenna techniques for full-duplex radios. In *2013 Asilomar Conference on Signals, Systems and Computers*, pages 3–8, Nov 2013.

[18] L. Anttila, D. Korpi, E. Antonio-Rodrguez, R. Wichman, and M. Valkama. Modeling and efficient cancellation of nonlinear self-interference in mimo full-duplex transceivers. In *2014 IEEE Globecom Workshops (GC Wkshps)*, pages 777–783, Dec 2014.

[19] M. A. Khojastepour and S. Rangarajan. Wideband digital cancellation for full-duplex communications. In *2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, pages 1300–1304, Nov 2012.

[20] N. Li, W. Zhu, and H. Han. Digital interference cancellation in single channel, full duplex wireless communication. In *2012 8th International Conference on Wireless Communications, Networking and Mobile Computing*, pages 1–4, Sept 2012.

[21] J. Kim, K. Shamaileh, S. Adusumilli, and V. Rao. Digital interference cancellation for multimedia transmission in full duplex communication link. In *2013 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, pages 1–5, June 2013.

[22] J. R. Krier and I. F. Akyildiz. Active self-interference cancellation of passband signals using gradient descent. In *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pages 1212–1216, Sept 2013.

[23] A. Masmoudi and T. Le-Ngoc. A maximum-likelihood channel estimator for self-interference cancelation in full-duplex systems. *IEEE Transactions on Vehicular Technology*, 65(7):5122–5132, July 2016.

[24] A. Masmoudi and T. Le-Ngoc. Channel estimation and self-interference cancelation in full-duplex communication systems. *IEEE Transactions on Vehicular Technology*, 66(1):321–334, Jan 2017.

[25] S. Li and R. D. Murch. Full-duplex wireless communication using transmitter output based echo cancellation. In *2011 IEEE Global Telecommunications Conference - GLOBECOM 2011*, pages 1–5, Dec 2011.

[26] Micael Bernhardt, Fernando H Gregorio, and Juan E Cousseau. A robust wireless ofdm echo cancellation system. *XV Reunin de Trabajo en Procesamiento de la Informacin y Control, RPIC*, 2013.

[27] S. Li and R. D. Murch. An investigation into baseband techniques for single-channel full-duplex wireless communication systems. *IEEE Transactions on Wireless Communications*, 13(9):4794–4806, Sept 2014.

[28] E. Ahmed and A. M. Eltawil. All-digital self-interference cancellation technique for full-duplex systems. *IEEE Transactions on Wireless Communications*, 14(7):3519–3532, July 2015.

[29] D. Korpi, L. Anttila, and M. Valkama. Reference receiver based digital self-interference cancellation in mimo full-duplex transceivers. In *2014 IEEE Globecom Workshops (GC Wkshps)*, pages 1001–1007, Dec 2014.

[30] S. Enserink, M. P. Fitz, K. Goverdhanam, Changyi Gu, T. R. Halford, I. Hossain, G. Karawasy, and O. Y. Takeshita. Joint analog and digital interference cancellation. In *2014 IEEE MTT-S International Microwave Symposium (IMS2014)*, pages 1–3, June 2014.

[31] Y. Hua, Y. Ma, P. Liang, and A. Cirik. Breaking the barrier of transmission noise in full-duplex radio. In *MILCOM 2013 - 2013 IEEE Military Communications Conference*, pages 1558–1563, Nov 2013.

[32] Z. Zhan, G. Villemaud, and J. M. Gorce. Design and evaluation of a wideband full-duplex ofdm system based on aasic. In *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pages 68–72, Sept 2013.

[33] J. H. Lee. Self-interference cancelation using phase rotation in full-duplex wireless. *IEEE Transactions on Vehicular Technology*, 62(9):4421–4429, Nov 2013.

[34] A. Balatsoukas-Stimming, P. Belanovic, K. Alexandris, and A. Burg. On self-interference suppression methods for low-complexity full-duplex mimo. In *2013 Asilomar Conference on Signals, Systems and Computers*, pages 992–997, Nov 2013.

[35] J. G. McMichael and K. E. Kolodziej. Optimal tuning of analog self-interference cancellers for full-duplex wireless communication. In *2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 246–251, Oct 2012.

[36] Z. He, S. Shao, Y. Shen, C. Qing, and Y. Tang. Performance analysis of rf self-interference cancellation in full-duplex wireless communications. *IEEE Wireless Communications Letters*, 3(4):405–408, Aug 2014.

[37] Steven S. Hong, Jeffrey Mehlman, and Sachin Katti. Picasso: Flexible rf and spectrum slicing. In *Proceedings of the ACM SIGCOMM 2012 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication*, SIGCOMM '12, pages 37–48, New York, NY, USA, 2012. ACM.

[38] Dinesh Bharadia, Emily McMilin, and Sachin Katti. Full duplex radios. *SIGCOMM Comput. Commun. Rev.*, 43(4):375–386, August 2013.

[39] Dinesh Bharadia, Kiran Joshi, and Sachin Katti. Robust full duplex radio link. *SIGCOMM Comput. Commun. Rev.*, 44(4):147–148, August 2014.

[40] Dinesh Bharadia and Sachin Katti. Full duplex mimo radios. In *11th USENIX Symposium on Networked Systems Design and Implementation (NSDI 14)*, pages 359–372, Seattle, WA, 2014. USENIX Association.

[41] A. Gholian, Y. Ma, and Y. Hua. A numerical investigation of all-analog radio self-interference cancellation. In *2014 IEEE 15th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 459–463, June 2014.

[42] Yingbo Hua, Yiming Ma, Armen Gholian, Yifan Li, Ali Cagatay Cirik, and Ping Liang. Radio self-interference cancellation by transmit beamforming, all-analog cancellation and blind digital tuning. *Signal Process.*, 108(C):322–340, March 2015.

[43] Y. Hua, Y. Li, C. Mauskar, and Q. Zhu. Blind digital tuning for interference cancellation in full-duplex radio. In *2014 48th Asilomar Conference on Signals, Systems and Computers*, pages 1691–1695, Nov 2014.

[44] Y. S. Choi and H. Shirani-Mehr. Simultaneous transmission and reception: Algorithm, design and system level performance. *IEEE Transactions on Wireless Communications*, 12(12):5992–6010, December 2013.

[45] T. Huusari, Y. S. Choi, P. Liikkanen, D. Korpi, S. Talwar, and M. Valkama. Wideband self-adaptive rf cancellation circuit for full-duplex radio: Operating principle and measurements. In *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*, pages 1–7, May 2015.

[46] J. Tamminen, M. Turunen, D. Korpi, T. Huusari, Y. S. Choi, S. Talwar, and M. Valkama. Digitally-controlled rf self-interference canceller for full-duplex radios. In *2016 24th European Signal Processing Conference (EUSIPCO)*, pages 783–787, Aug 2016.

[47] D. Korpi, Y. S. Choi, T. Huusari, L. Anttila, S. Talwar, and M. Valkama. Adaptive nonlinear digital self-interference cancellation for mobile inband full-duplex radio: Algorithms and rf measurements. In *2015 IEEE Global Communications Conference (GLOBECOM)*, pages 1–7, Dec 2015.

[48] D. Korpi, J. Tamminen, M. Turunen, T. Huusari, Y. S. Choi, L. Anttila, S. Talwar, and M. Valkama. Full-duplex mobile device: pushing the limits. *IEEE Communications Magazine*, 54(9):80–87, September 2016.

[49] J. Wang, H. Zhao, and Y. Tang. A rf adaptive least mean square algorithm for self-interference cancellation in co-frequency co-time full duplex systems. In *2014 IEEE International Conference on Communications (ICC)*, pages 5622–5627, June 2014.

[50] X. Huang and Y. J. Guo. Radio frequency self-interference cancellation with analog least mean-square loop. *IEEE Transactions on Microwave Theory and Techniques*, PP(99):1–15, 2017.

[51] H. Zhao, J. Liu, J. Wang, W. Guo, and Y. Tang. Impacts of attenuators on analog self-interference cancellation in full duplex radios. In *2016 IEEE Globecom Workshops (GC Wkshps)*, pages 1–5, Dec 2016.

[52] B. Kaufman, J. Lilleberg, and B. Aazhang. An analog baseband approach for designing full-duplex radios. In *2013 Asilomar Conference on Signals, Systems and Computers*, pages 987–991, Nov 2013.

[53] T. Riihonen, P. Mathecken, and R. Wichman. Effect of oscillator phase noise and processing delay in full-duplex ofdm repeaters. In *2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, pages 1947–1951, Nov 2012.

[54] A. Sahai, G. Patel, C. Dick, and A. Sabharwal. On the impact of phase noise on active cancelation in wireless full-duplex. *IEEE Transactions on Vehicular Technology*, 62(9):4494–4510, Nov 2013.

[55] E. Ahmed, A. M. Eltawil, and A. Sabharwal. Self-interference cancellation with phase noise induced ici suppression for full-duplex systems. In *2013 IEEE Global Communications Conference (GLOBECOM)*, pages 3384–3388, Dec 2013.

[56] S. Shao, X. Quan, Y. Shen, and Y. Tang. Effect of phase noise on digital self-interference cancellation in wireless full duplex. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2759–2763, May 2014.

[57] V. Syrjala, M. Valkama, L. Anttila, T. Riihonen, and D. Korpi. Analysis of oscillator phase-noise effects on self-interference cancellation in full-duplex ofdm radio transceivers. *IEEE Transactions on Wireless Communications*, 13(6):2977–2990, June 2014.

[58] V. Syrjl, K. Yamamoto, and M. Valkama. Analysis and design specifications for full-duplex radio transceivers under rf oscillator phase noise with arbitrary spectral shape. *IEEE Transactions on Vehicular Technology*, 65(8):6782–6788, Aug 2016.

[59] X. Quan, Y. Liu, S. Shao, C. Huang, and Y. Tang. Impacts of phase noise on digital self-interference cancellation in full-duplex communications. *IEEE Transactions on Signal Processing*, 65(7):1881–1893, April 2017.

[60] E. Ahmed, A. M. Eltawil, and A. Sabharwal. Self-interference cancellation with nonlinear distortion suppression for full-duplex systems. In *2013 Asilomar Conference on Signals, Systems and Computers*, pages 1199–1203, Nov 2013.

[61] V. Syrjaelae and K. Yamamoto. Self-interference cancellation in full-duplex radio transceivers with oscillator phase noise. In *European Wireless 2014; 20th European Wireless Conference*, pages 1–6, May 2014.

[62] E. Ahmed and A. M. Eltawil. On phase noise suppression in full-duplex systems. *IEEE Transactions on Wireless Communications*, 14(3):1237–1251, March 2015.

[63] L. Anttila, D. Korpi, V. Syrjl, and M. Valkama. Cancellation of power amplifier induced nonlinear self-interference in full-duplex transceivers. In *2013 Asilomar Conference on Signals, Systems and Computers*, pages 1193–1198, Nov 2013.

[64] T. Riihonen and R. Wichman. Analog and digital self-interference cancellation in full-duplex mimo-ofdm transceivers with limited resolution in a/d conversion. In *2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, pages 45–49, Nov 2012.

[65] D. Korpi, T. Riihonen, V. Syrjl, L. Anttila, M. Valkama, and R. Wichman. Full-duplex transceiver system calculations: Analysis of adc and linearity challenges. *IEEE Transactions on Wireless Communications*, 13(7):3821–3836, July 2014.

[66] Z. Zhan, G. Villemaud, and J. M. Gorce. Analysis and reduction of the impact of thermal noise on the full-duplex ofdm radio. In *2014 IEEE Radio and Wireless Symposium (RWS)*, pages 220–222, Jan 2014.

[67] M. Biguesh and A. B. Gershman. Training-based mimo channel estimation: a study of estimator tradeoffs and optimal training signals. *IEEE Transactions on Signal Processing*, 54(3):884–893, March 2006.

[68] Simon Haykin. *Adaptive Filter Theory (3rd Ed.)*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1996.