The Pennsylvania State University
The Graduate School
College of Engineering

JOINT SPECIFIC EMITTER IDENTIFICATION AND TRACKING USING DEVICE NONLINEARITY ESTIMATION

A Dissertation in
Electrical Engineering
by
Ming-Wei Liu

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2011

The dissertation of Ming-Wei Liu was reviewed and approved* by the following:

John F. Doherty
Professor of Electrical Engineering
Dissertation Advisor, Chair of Committee

W. Kenneth Jenkins
Head of the Department of Electrical Engineering
Professor of Electrical Engineering

John Metzner
Professor of Electrical Engineering

Allan Sonsteby
Professor of Information Sciences and Technology

*Signatures are on file in the Graduate School.

Abstract

In this thesis, we present radio frequency (RF) front-end nonlinearity estimators to perform joint specific emitter identification (SEI) and tracking. Our SEI systems discern radio emitters of interest through the estimation of transmitter nonlinearities caused by design and fabrication variations. These nonlinearity features provide unique signal signatures for each emitter, and we extract those characteristics through the estimation of transmitter nonlinearity coefficients. We first present a nonlinearity estimator which estimates the power series coefficients of nonlinear devices in the radio frequency (RF) front end by observing the spectral regrowth in additive white Gaussian noise (AWGN) channel. Then another robust algorithm is also provided by using alternative degrees of nonlinearities associated with symbol amplitudes for initial estimation, and then iteratively estimating the channel coefficients and distorted transmit symbols to overcome the inter-symbol interference (ISI) effect. The convergence and unbiasedness of the iterative estimator are demonstrated semi-analytically. Based on this analysis, we also trade error performance for complexity reduction using the regularity of the estimation process. The algorithm is applicable to a wide range of multi-amplitude modulation schemes, and we present an SEI system designed for an orthogonal frequency division multiplexing (OFDM) system over an empirical indoor channel model with associated numerical results. This technology is then adapted to provide location tracking in multipath environments, which locates the mobile stations (MS) based on the transmit power variation estimates. The method is simulated over a grid-based city map. In the last part of the thesis, complexity reduction methods are introduced to balance the convergence rate and identification performance.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would like to thank everyone who contributed to this thesis.

Foremost, I wish to thank my advisor, Professor John F. Doherty for his wise advice and support throughout my thesis work. His commitment to academic thoroughness and the sense of integrity continues to serve as a model upon which to fashion my own career. Thanks also go to the committee members for their feedback and guidance.

Many and heartfelt thanks go to all my friends, on both sides of the Pacific, for their constant encouragement, sense of humor, and willingness to listen.

Finally, my sincere thanks go to my wonderful parents and sister for the most unflinching love, the deepest attention, invaluable forbearance, and pure cares.

Chapter 1

Introduction

1.1 Motivation

The objective of this thesis is to research and develop algorithms capable of specific emitter identification (SEI) and specific emitter tracking (SET) based on external signal feature extraction which utilizes the information from transmitter nonlinearities estimates to improve situation awareness. Our algorithms can be applied to multiple communication systems of different modulation schemes with single or multiple antennas, and are robust to operate in multipath environments. SEI techniques are used to recognize individual electronic emitters through the precise measurement of selected signal external features, while SET provides geolocation tracking of the target. This capability leads to diverse applications such as cognitive radio [1], network intrusion detection, battlefield management, electronic support measurement system, and intelligence gathering [2–7].

In order to be identified by SEI technologies, a specific emitter must have one or more of these signal external features that are stable and unique. The external signal feature used for identification in this thesis is transmitter nonlinearities which satisfied our requirement in uniqueness and stability [4]. The main causes of this feature are different designs in power amplifier and fabrication randomness. The estimation algorithms are based on the measurement of power spectrum density and distorted signal constellation. To enhance the robustness of estimation, the algorithm iteratively estimates the channel coefficients and transmit symbols to suppress inter-symbol interference (ISI) effects which is new and can not be found in published research results. To complete the re-

connaissance ability, we utilize the relationship between the transmit power variation and the transmitter nonlinearities to perform SET. Moreover, several algorithms were developed to reduce the complexity for practical applications.

1.2 Contributions

A major issue for building SEI systems is to measure proper signal features that are consistent from one transmission to another for a given emitter but different from one emitter to another. The transmitter nonlinearities are introduced by alternative designs and fabrication randomness which provide unique characteristics of the transmitter [4], and this uniqueness is the base of the emitter identification. We first consider the transmit symbols as a sequence of random variables, and estimate the nonlinearity coefficients through the measurement of power spectrum density. This algorithm is presented in:
M.-W. Liu and J. F. Doherty, "Specific emitter identification using nonlinearity device es- timation," in *IEEE 2008 Sarnoff Symposium, Princeton*, NJ, 28-30 April 2008.

Another issue in SEI systems is that no research in the public arena is developed to suppress inter-symbol interference (ISI) effects which affect the performance deleteriously. In this dissertation, we develop iterative algorithms and utilize training sequences to suppress this effect to maintain robust estimation, and they are shown in:
M.-W. Liu and J. F. Doherty, "Nonlinearity estimation for specific emitter identification in multipath environment," in *IEEE 2009 Sarnoff Symposium*, March 2009.
M.-W. Liu and J. F. Doherty, "Wireless device identification in MIMO channels," in *IEEE Conference on Information Sciences and Systems 2009* , March 2009.
M.-W. Liu and J. F. Doherty, "Specific emitter identification in multipath channels," To appear in *IEEE Transactions on Information Forensics and Security*.

Moreover, the nonlinearity estimates also provide the information about the variation of transmit power. Since this information is location dependent, it is used to track the location of the transmitter in:
M.-W. Liu and J. F. Doherty. "Joint Specific Emitter Identification and Location Tracking for OFDM Systems." In *Symposium and Summer School on Wireless Comm.*, Virginia Tech., June 2010.
M.-W. Liu and J. F. Doherty. "On Joint Specific Emitter Identification and Tracking" Submitted to *IEEE Transactions on Information Forensics and Security*.

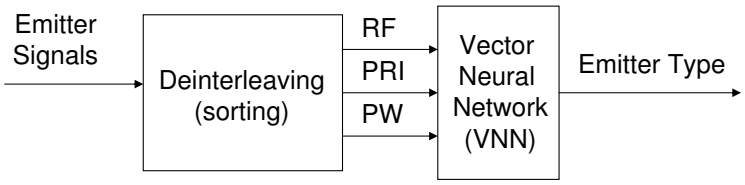


Figure 1.1. The flowchart of emitter signal classification

The estimation of transmitter nonlinearities can also be extended to channel estimation. In the majority of the channel estimation research, the transmitter nonlinearities are not considered. We develop a channel estimation algorithm for frequency-selective multiple-input multiple-output (MIMO) channel estimation including the transmitter nonlinearities, and the results show that the performance improves substantially. The results are published in:
M.-W. Liu and J. F. Doherty, "Frequency-selective multiple-input multiple-output channel estimation with transmitter non-linearities," *Signal Processing, IET*, vol. 3, no. 6, pp. 467 475, 2009.

1.3 Specific Emitter Identification Introduction

SEI systems recognize radio emitters of interest through only the measurement of external signal features. These features are selected based on the types of emitter sources, which can be loosely divided into two categories: radar and communications signals. In the scope of radar signals, there has been a great deal of research on performing SEI by aggregating decisions made by estimating characteristics of interpulse features such as pulse repetition interval (PRI) [8–10]. Another class of feature is intrapluse features including pulse width, pulse shape, intentional modulation on pulse (IMOP), and unintentional modulation on pulse (UMOP) [11].

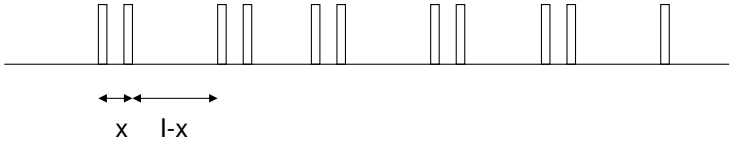
Fig. 1.1 shows a typical SEI block diagram, where the radar signals from multiple emitters are first deinterleaved; then the signal features are extracted and fed to the classifier in the next stage. Many of these deinterleaving and estimation technologies are based on the TOA difference histogram of the pulse train. If the potential PRI is identified, the algorithm looks for a group of pulses that form a periodic pulse train.

Further development to detect the fundamental PRIs using the cumulative difference

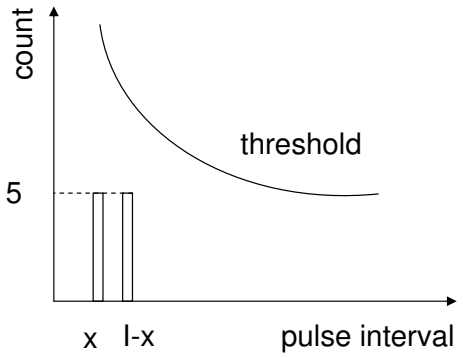
histogram (CDIF) is presented in [8] and the sequential difference histogram (SDIF) is proposed in [9]. Fig. 1.2(a) shows a simple example where two interleaved stable sequences with PRIs of I time units is presented or a two-level staggered signal with stagger frame rate equal to the PRI. CDIF algorithm first counts the histogram of the first difference of the signal as shown in 1.2(b). In this example, both stems are under the threshold, and the algorithm searches the second difference as shown in Fig. 1.2(c). Otherwise, if both stems are above the threshold, the algorithm determines the PRI and treats the second stem as the harmonics. In Fig. 1.2(c), since the third stem is above the threshold, it is considered as the PRI. By requiring the second harmonics, the algorithm keeps searching until the fourth difference as shown in Fig. 1.2(d) and 1.2(e). Once the PRI is determined, the algorithm removes the pulses of that PRI from the sample and the histogram resets. SDIF histogram is similar to CDIF but only the current differences exist. The purpose of this modification is to prevent the situation when many emitters are present, the first difference will produce a few values exceeding the threshold, and none of the corresponds to the correct PRI value.

These algorithms intend to avoid PRI harmonics by combining the histogram with sequence search techniques. Another identification method that exploits the properties of the matrix formed with the differences of TOA is presented in [10]. In addition, the algorithm in [12] also suppresses the PRI harmonics using PRI spectrum analysis. Other than signals with stable sequence intervals where a sharp narrow peak indicates a simple PRI, a staggered PRI gives many separated peaks in the histogram [13].

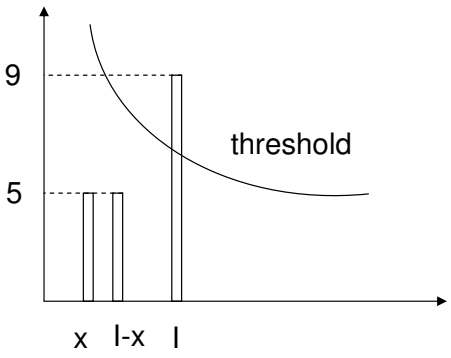
Another category of signal features uses the parameters of intentional pulse modulation (IMOP) found in pulse compression radars. Typical modulation on pulse includes frequency modulation on pulse (FMOP) and phase modulation on pulse (PMOP). In practice, linear frequency modulation (LFM) or chirp is used often, and the compressor linearly changes the carrier frequency with time. Moreover, phase modulation on pulse (PMOP) is also widely used. This technology broadens the bandwidth of the radar pulse by modulating the signal with code sequences. Electronic intercept systems can also distinguish the emitter of interest through the estimate of the frequency and the phase parameters. Other than IMOP, unintentional modulation on pulse (UMOP) signal features are widely used [3, 14]. UMOP introduced in the transmitted pulses is mainly due to the individual characteristics of the transmitter components. For instance, the envelope shape of a pulse depends on the interactions between the high voltage pulse



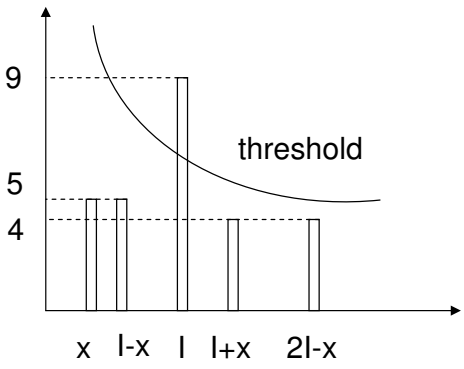
(a) Radar signal pulse train.



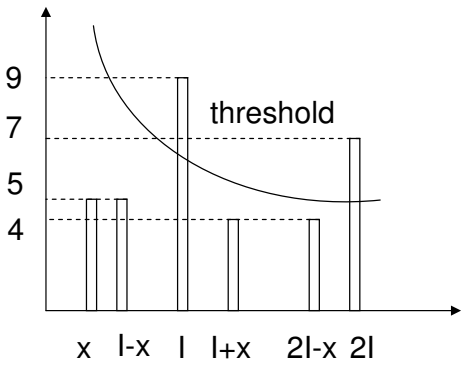
(b) First difference



(c) Second difference



(d) Third difference



(e) Fourth difference

Figure 1.2. A simple example of TOA histogram deinterleaving method.

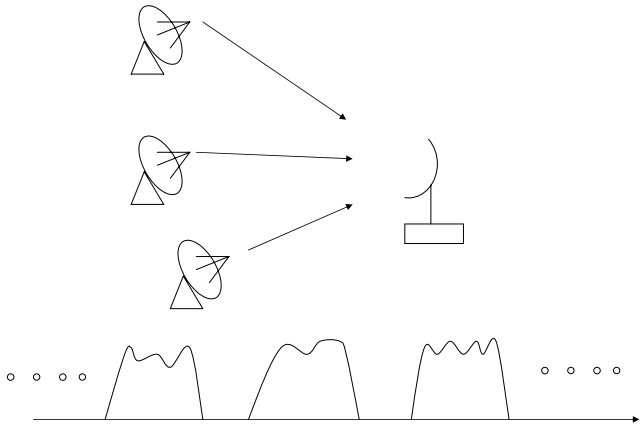


Figure 1.3. Emitter identifier receives signal from multiple radars.

generated by the individual high voltage power supply applied to the magnetron. Therefore, by analyzing the pulse shape generated by different transmitters as shown in Fig. 1.3, the SEI system can identify the transmitter of this pulse. Some common features of the amplitude waveform are rise time, fall time, rise angle, and fall angle [5]. The thermal condition of the amplifier also induces frequency shift, and thus the frequency deviation around the carrier also provides a characteristic feature [11].

In cases when communication systems are emitter sources, many signal characteristics have been investigated including channel-related and transmitter-specific features. The basic assumption of channel-related features is that the channel of each transmitter should represent unique characteristics of its location. In these location related features, [15] utilized the receive signal strength indication (RSSI) vector as the signal fingerprint to distinguish transmitters in a wireless local area network (WLAN) environment; [16] evaluated an approach that compares multipath coefficients of the channel; [17] proposed a physical layer authentication based on parameters from channel estimates. In the category of transmitter-specific features, transient amplitude and phase profiles of the radio frequency (RF) fingerprinting are the most researched type. In [6, 18], the authors sampled the instant waveforms captured from wireless devices, and the turn-on transients are determined when the receive signal power performs a sudden increase or is above a threshold. Then these samples are compared to the fingerprint

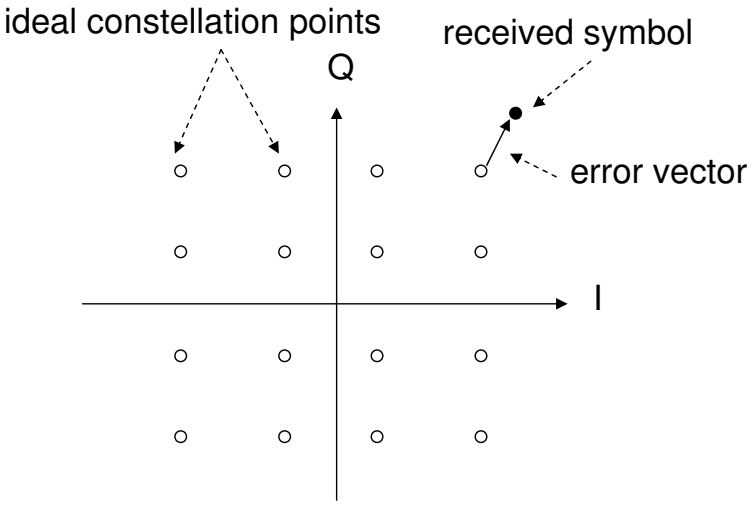


Figure 1.4. An example of error vector definition in 16-QAM constellation

features stored in the database for identification. Other features such as clock skew are studied in [19], while error vector magnitude (EVM) is used in [20, 21]. The definition of EVM is

$$EVM(dB) = 20log_{10}\left(\frac{P_{error}}{P_0}\right) \tag{1.1}$$

where P_{error} is the root mean square (RMS) power of the error vector and P_0 is average power of the ideal constellation as shown in Fig. 1.4. The error vector is produced by the imperfect design and fabrication which is considered unique to each transmitter. A typical allowed EVM in IEEE 802.11a specification is between -5 dB to -25 dB according to the data rate [22].

Another approach is to exploit frequency domain signal characteristics. For instance, the main lobe shape symmetry of power spectrum density (PSD) [4] as well as spectral correlation function [1] all provide device identification features.

1.3.1 Transmitter Nonlinearities Measurement

In this thesis, an SEI system is developed estimating the device nonlinearities of the transmitter. One major source of this unique feature is that each system design generates its own output response according to the input voltage, and another source comes from the fabrication randomness which varies in each manufacturing process [4].

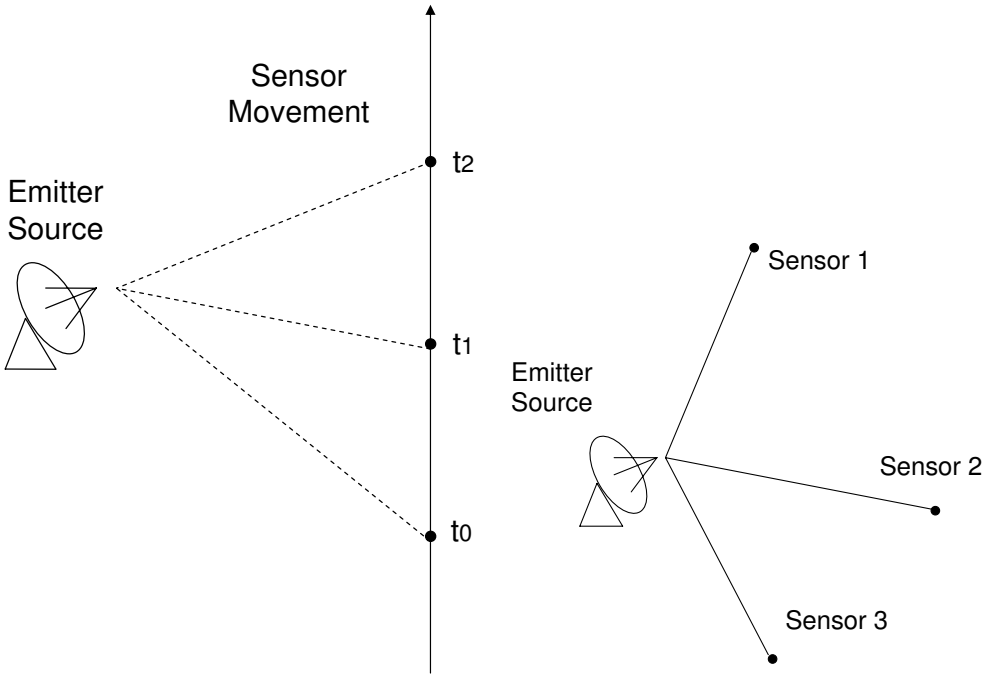
Our first measurement approach is to observe the power spectrum density (PSD) estimation [23]. It is known that transmitter nonlinearities change the shape of PSD [24,25], and the nonlinearity coefficients are extracted through the observation. Another nonlinearity estimation approach is to observe the distorted symbol constellation in multipath channels [26–28], and then the nonlinearities are extracted through the distortion results. The algorithm suppresses the ISI effects which affects the estimator and has not yet been researched in the open literature.

1.4 Specific Emitter Tracking Introduction

The capability of accurately performing emitter location has become a primary requirement for today's reconnaissance systems. In practice, one or more platforms can be used to perform geolocation using time and frequency information. These location systems can be classified into two broad categories: direction finding (DF) and range-based position location (PL) systems. DF systems estimate the position location of a emitter source by measuring the direction of arrival (DOA) or angle-of-arrival (AOA) of the emitter, and then overlay the bearings to form an interaction region as the potential emitter location. Range-based PL systems utilize the range between the emitter source and sensors, or range difference between the signal source and different sensors. The range difference measurement defines a hyperboloid of constant range, and the intersection of multiple range differences provides the location of the emitter. Moreover, with the help from terrain data, the receivers can estimate the range from the receive signal strength (RSS). Some indoor applications record the RSS of each sensor according to the transmitter location. Based on different terrain and channel condition, each emitter location brings a unique RSS vector which is then used as the location reference. Once the RSS is measured from each sensor, it is compared to the database to localize the emitter.

1.4.1 DF Measurements

The function of a DF receiver is to perform DOA measurement of the received signal which can be broadly classified into three categories: amplitude-comparison, phase interferometry, and subspace-based methods. In the case where only a single moving sensor is available as shown in Fig. 1.5(a), DF information is obtained through a se-



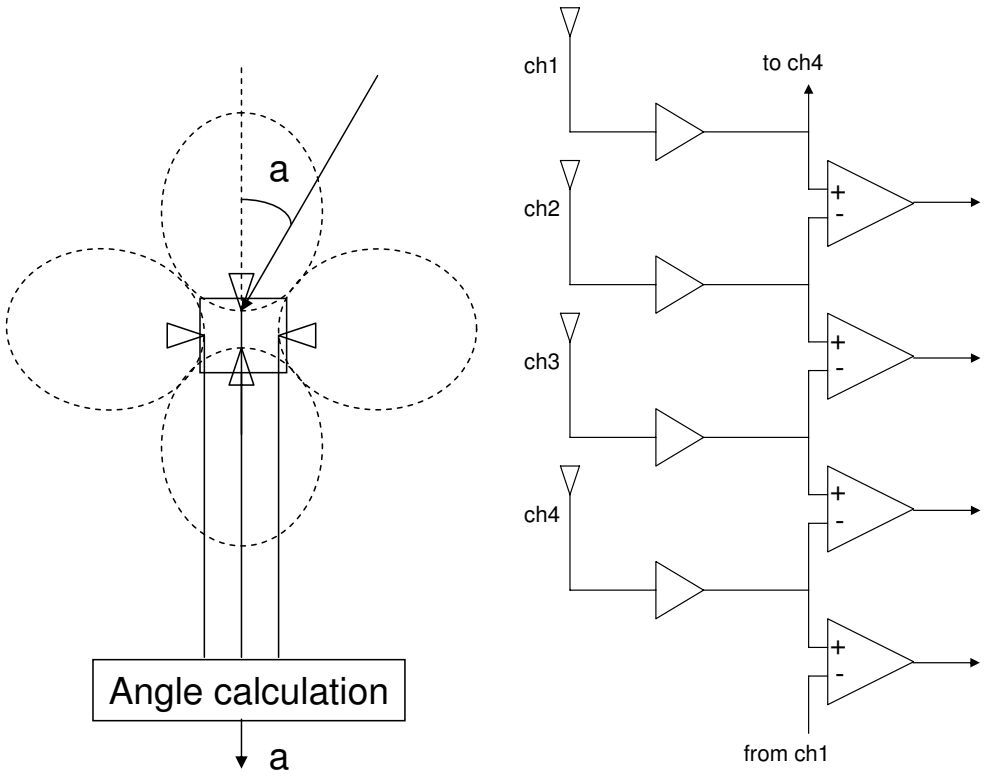
(a) Single-sensor emitter location finding using a moving sensor.

(b) Multiple-sensor emitter location finding using multiple stationary sensors.

Figure 1.5. Two emitter location finding schemes using AOA methods.

quence of observations which forms lines of bearings (LOBs) along which the emitter lies; then the position of the target can be estimated. Fig. 1.5(b) demonstrates emitter location finding with multiple cooperative platforms, where the LOBs from different platforms can be integrated to derive the emitter location.

One of the difficulties in locating communications emitters is their relatively lower frequencies which lead to longer wavelengths. For instance, very high frequency (VHF) ranges from 30M to 300 MHz, while ultra high frequency (UHF) ranges from 300 MHz to 3 GHz, and some wireless applications such as Bluetooth and wireless local area network (WLAN) in industrial, scientific and medical band use 900MHz, 2.5GHz, and 5.8GHz. The associated wavelength could be several centimeters and up to 10 meters, which increases the difficulty of design.



(a) Phase comparison direction finder.

(b) Angle calculator implemented by amplitude comparator.

Figure 1.6. Implementation of amplitude-comparison direction finder.

1.4.1.1 Amplitude-Comparison DF

One basic method in amplitude-comparison DF receiver is to use anisotropy in the reception pattern of an antenna. The beam of antenna is then rotated electronically or mechanically, and the direction corresponding to the maximum signal strength is taken as the direction of the transmitter. Another approach is to use the minimum of two stationary antennas with known antenna patterns as illustrated in Fig. 1.6(a). A simple way to implement the angle finder is to compare the signal strength received from each antenna yielding the transmitter direction as presented in Fig. 1.6(b). The system then knows the antennas with the strongest and next strongest receive signal strength. Direction finders with four antennas can provide a precision of 10 to 15 degree rms in 2-18 GHz band, while the precision improves to 4-6 degree rms with eight antennas [29].

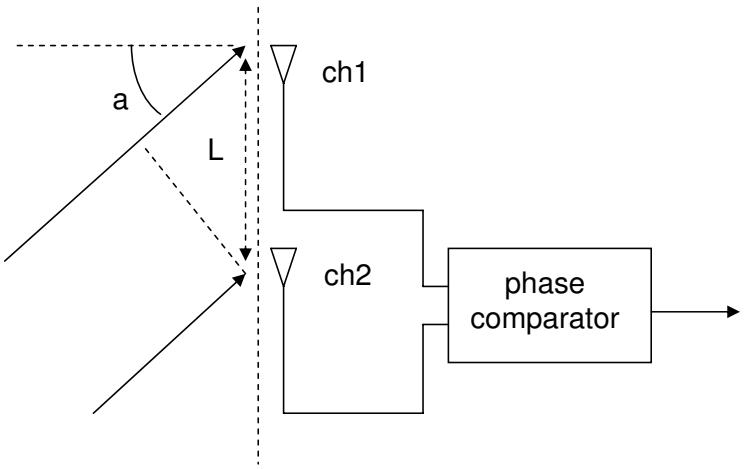


Figure 1.7. A phase-comparison DF implementation with two antennas.

1.4.1.2 Phase-Comparison DF

This technique derives the AOA measurements from the estimations of the phase differences in the arrival of a wave front. It typically requires an antenna array, in which the adjacent antenna elements are separated by a uniform distance. The principal of this method is shown in Fig. 1.7, and assuming the phase shift is less than 360 degree, a signal arriving from an off-axis direction a causes a phase shift

$$\varphi = \frac{L \sin a}{\lambda} 2\pi \tag{1.2}$$

where λ is the wave length. Phase interferometry AOA measurements rely on line-of-sight (LOS) path from the transmitter to the receiver. Multipath component may appear as a signal arriving from an entirely different direction and can lead to very different estimation results.

1.4.1.3 Subspace-Based Algorithms

The most well-known subspace-based algorithms are multiple signal classification (MUSIC) [30] and estimation of signal parameters by rotational invariance techniques (ESPRIT) [31, 32]. These eigenanalysis algorithms require multiple antennas to form a receive signal correlation matrix. After eigendecomposition of the correlation matrix, the vector space is divided into two orthogonal spaces, signal and noise space, which con-

tains the information of the direction. Then the algorithm then extracts the signal direction from these subspaces. These techniques can resolve signals with close frequencies; hence, are often referred to as super-resolution methods. Usually these methods fail in multipath environment. But using a spatially smoothed covariance matrix to replace the conventional covariance matrix, the modified algorithm can solve this problem [33].

1.4.2 Distance Related Measurements

Ranging-based systems estimate the emitter location based on the estimates of range related information between the transmitter and receivers. The range information comes from time, frequency estimates, and receive power. In this section, several methods are introduced such as time of arrival (TOA), time difference of arrival (TDOA), frequency difference of arrival (FDOA), and receive signal strength (RSS).

1.4.2.1 TOA

One approach in this category is the one-way propagation time measurement which estimates the difference between the sending at the transmitter side and receiving at the receiver side. It requires the synchronization of the local time at the receiver side and transmitter side. Another approach measures the difference between the time when a signal is sent to the receiver and returned to the transmitter. The advantage of this approach is that since the same clock is used, no synchronization is required. To illustrate the TOA concept, consider a 2-dimension localization system using three sensors is shown in Fig. 1.8. The time of arrival of a signal at each sensor is estimated by the relationship

$$R_i = cd_i = \sqrt{(x_i - X)^2 + (y_i - Y)^2},\ \forall i = 1,\ 2,\ 3 \tag{1.3}$$

where R_i is the distance between the emitter and sensor i, c is the signal propagation speed, d_i is the TOA estimate, (x_i, y_i) is the location coordinate of the sensor, and (X, Y) is the location of the radio emitter.

1.4.2.2 Distance Estimation via RSS

In this method, the location estimator combines the propagation model and RSS measurement to locate the emitter. A empirical propagation model of path loss and shadow-

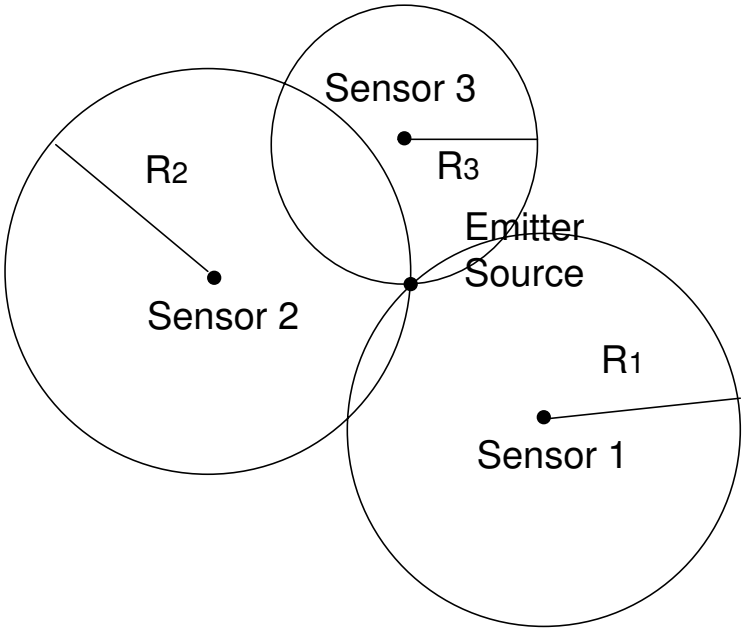


Figure 1.8. A ranging position location method example of one radio emitter and three sensors.

ing is

$$P_r(d) = P_0 - 10n_p log_{10}\frac{d}{d_0} + X_\sigma \tag{1.4}$$

where the received power is P_r, P_t is the reference power, n_p is the path loss exponent, X_σ is a zero mean Gaussian distributed random variable with standard deviation to account for shadowing [34]. RSS is measured by the receive signal power. This measurement is relatively inexpensive and simple to implement in hardware.

1.4.2.3 TDOA

TDOA measurements at multiple receivers are utilized with the knowledge of receiver location to estimate the location of the transmitter. Suppose that a signal propagates from a transmitter to receivers i and j where the local time is synchronized so that TDOA can be estimated. The range difference between the two paths is

$$\tau_{i,j} = (R_i - R_j)/c \tag{1.5}$$

where $\tau_{i,j}$ is the TDOA, R_i and R_j are the distances between the transmitter and sensor i and j, and c is the speed of propagation. For a given $\tau_{i,j}$ this equation becomes a

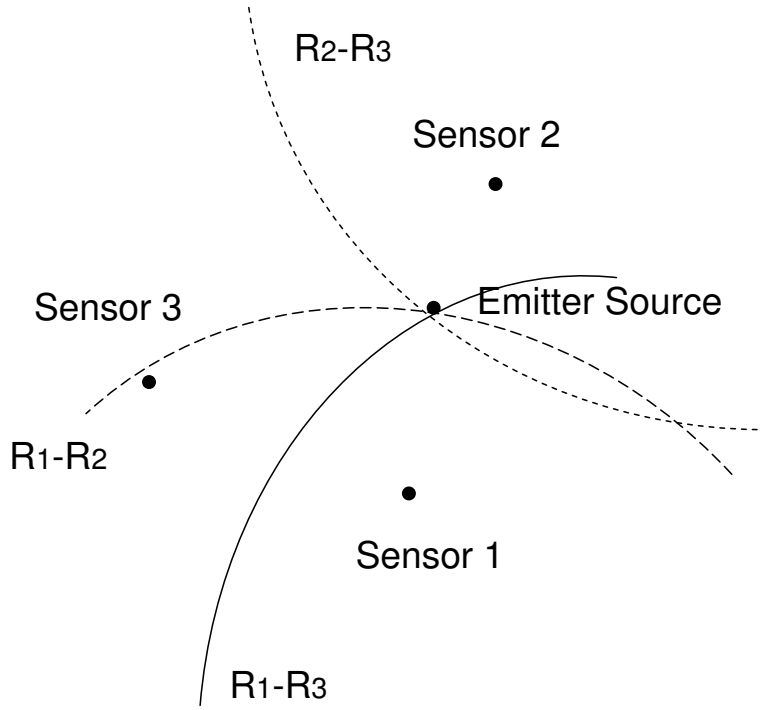


Figure 1.9. A TDOA example with three sensors and one emitter source.

hyperboloid on which the emitter resides. Therefore, if multiple estimates are given from more than three sensors, the intersection of the hyperboloid planes is the location of the emitter. Fig. 1.9 demonstrates an example of TDOA method with one emitter and three receivers. A major advantage of this TDOA method is that it does not require knowledge of the transmit time. However, clock synchronization is required for all sensors.

1.4.2.4 FDOA

FDOA methods involve the measurement of the difference between the received frequency at sensors from a single radio emitter. Since no movement leads to no Doppler shift, this method requires moving sensors or/and a moving transmitter. The difference between the frequency measurements at the sensors provides a FDOA measurement, and therefore the lack of carrier frequency information is not an issue. The difference between two receivers provides a surface where the emitter lies. TDOA and FDOA are sometimes combined to improve the accuracy since the estimates are uncorrelated [29].

1.4.3 RSS Profiling Measurements

The RSS profiling-based approach uses pre-measured map of the signal strength behavior in the coverage area. The map is obtained by simulating signal model or field measurement by a sniffing device at known location. This technique is attractive for wireless sensor networks and mostly used in WLANs. Once the sensor network is established, at each sample point, each sensor measures the RSS and the location estimator saves the data from all sensors as a vector. The collection of all these vectors provides a map of RSS profile. By referring to the RSS model, an emitter can be located using the RSS measurements.

1.4.4 Transmit Power Variation Measurements

To achieve joint specific emitter identification and tracking, we developed a transmit power variation estimator which derives the power variation through the observation of the nonlinearity estimates. In communication systems with a power control mechanism, the transmit power is location dependent. For instance, if the mobile station (MS) is closer to the base station (BS), the MS reduces its transmit power to reduce interference and save the power. On the other hand, once the MS is leaving the BS, it increases the transmit power to maintain constant receive power on the BS side to control the signal quality. To utilize this location dependent signature, our algorithm utilizes the transmitter nonlinearity estimation related to the transmit signal power variation. Considering the normalized transmitter nonlinearities, the amplitude of these features rises in proportional to the transmit power level. Therefore, by observing a sequence of this variation in time, the tracking system not only monitors the specific emitter but also obtains the movement of the MS. Moreover, the transmit variation profile also provides unique pattern of a route. By pre-measuring or software simulating the transmit power variation in each route, this information can be used as a reference to the location of the emitter.

1.5 Complexity Reduction

The major computation of the transmitter nonlinearity algorithms introduced is to solve a systems of linear equations. The complexity of the iterative algorithm grows with the size of systems of linear equations. Therefore, for complexity reduction, several

algorithms are introduced to increase the speed of convergence, reduce the hardware requirement, and at the same time maintain the noise floor. To increase the convergence rate, the algorithm segments the data matrix into smaller matrices and iterates among the smaller matrixes. However, our analysis shows that the greater the size of the matrix, the lower noise floor the algorithm achieves. Hence, we develop an algorithm that uses smaller matrix segments for initial estimation to increase the convergence rate and switch to larger matrix segments to reach lower noise floor results.

1.6 Organization

The rest of this thesis is organized as follows. In Chapter 2, we give a concise description of the SEI problem and estimate the transmitter nonlinearities through the observation of PSD. In Chapter 3, we consider the multipath channel scenario, and introduce an initial nonlinearity estimator and iterative algorithm to suppress the ISI effects. In Chapter 4, a MIMO channel estimator considering transmitter nonlinearities is presented. In Chapter 5, a location finding system is presented. In Chapter 6, several complexity reduction algorithms are provided to reduce the complexity of our iterative algorithms. Chapter 7 summarize the thesis.

Chapter 2

Nonlinearity Estimation using Bandwidth Regrowth

2.1 Introduction

In this chapter, a specific emitter identification method is proposed by estimating the nonlinearity of transmitter devices. Our algorithm can extract and analyze the nonlinearity coefficients based on the observation of spectral regrowth. We also developed an analytical model that could be used to study the detection method under AWGN channel environment. For different input power scenarios, we also obtained a model that can predict the trajectory of the moving features. Based on IEEE 802.11a/g specification, our simulation showed that by collecting data for a proper period of time, the proposed method can correctly identify detectable radio emitters in a typical SNR environment. Most of the work in this chapter is presented in [23].

2.2 On Baseband Representation of Bandpass Signal Nonlinearities

The response of memoryless nonlinearity can be modeled using the amplitude/amplitude (AM/AM) and amplitude/phase (AM/PM) characteristics, and this assumption also holds for wideband devices if the nonlinear response is not frequency dependent [35]. In this case, a complex power series expansion can be used to describe the AM/AM and

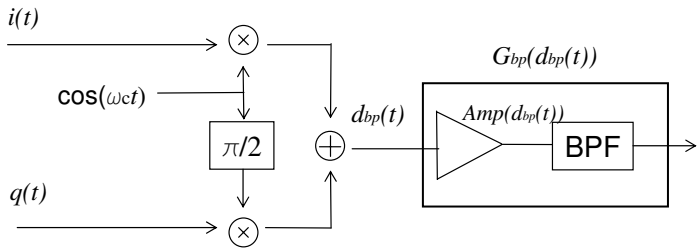


Figure 2.1. Block diagram of quadrature modulator and nonlinear devices.

AM/PM model [35]

$$Amp(d_{bp}(t)) = Re\left\{\theta_1 d_{bp}(t) + \theta_3 d_{bp}^3(t) + \ldots + \theta_M d_{bp}^M(t)\right\} \tag{2.1}$$

where $d_{bp}(t)$ is the bandpass input signal as shown in Fig. 2.1. For a bandpass signal, odd terms are the most important as they introduce intermodulation distortion within the usage band and produce spectral regrowth on the adjacent band, while even terms do not contribute to this effect.

Supposed the quadrature-modulated bandpass signal $d_{bp}(t)$ could be characterized by two independent zero-mean Gaussian processes as the following equation:

$$\begin{aligned} d_{bp}(t) &= i(t)\cos(2\pi f_c t) - q(t)\sin(2\pi f_c t) \\ &= \frac{1}{2}\left[d(t)e^{jw_c t} + d^*(t)e^{-jw_c t}\right], \end{aligned} \tag{2.2}$$

where $d(t) = i(t) + jq(t)$, $i(t)$ and $q(t)$ represent the in-phase and quadrature baseband signals as figure 2.1 and both signals are zero-mean Gaussian processes.

$$\begin{aligned} G_{dp}(d(t)) &= BPF(Amp(d_{bp}(t)) \\ &= Re\left[\sum_{k=0}^{\frac{M-1}{2}} \frac{\theta_{2m+1}}{2^{2m}} \binom{2m+1}{m+1} (d(t))^{m+1} (d^*(t))^m e^{j\omega_c t}\right] \end{aligned} \tag{2.3}$$

The baseband equivalent presentation of the nonlinear system is

$$\begin{aligned} G(d(t)) &= \sum_{m=0}^{\frac{M-1}{2}} \frac{\theta_{2m+1}}{2^{2m}} \binom{2m+1}{m+1} (d(t))^{m+1} (d^*(t))^m \\ &= a_1 \breve{d} + a_3 \breve{d} |\breve{d}|^2 + \ldots + a_M \breve{d} |\breve{d}|^{M-1} \\ &\equiv z(t) \end{aligned} \tag{2.4}$$

where

$$a_{2m+1} = \frac{\theta_{2m+1}}{2^{2m}} \binom{2m+1}{m+1} \sigma_d^{2m+1} \tag{2.5}$$

and $\breve{d}(t) = d(t)/\sigma_d$ is the power normalized version of input signal with σ_d^2 as the power of $d(t)$. Typical input signal powers of a WLAN device are 3 mW to 10 mW and the output powers are 25 mW to 35 mW, while the output signal power of a GSM device can be up to $2W$. Here we consider the actual input signal power as unknown; therefore we normalize it by its estimated power.

Since we only consider the spectrum around carrier frequency, it is sufficient to use $d(t)$ instead of $d_{bp}(t)$. Its AM/AM relationship could be explicitly shown as

$$|z(t)| = \left| \left(\sum_{k=0}^{\frac{M-1}{2}} a_{2m+1} d(t) \, |d(t)|^{2k} \right) \right|, \tag{2.6}$$

and the AM/PM could also be expressed as

$$\angle z(t) - \angle d(t) = \angle \sum_{k=0}^{\frac{M-1}{2}} a_{2m+1} \, |d(t)|^{2k} . \tag{2.7}$$

Since we are interested in the nonlinearities and the derivation of the linear gain is not in the scope of this research, we simply modified (2.4) as

$$z(t) = a_1 \left(\breve{d}(t) + b_3 \breve{d}(t) |\breve{d}(t)|^2 + \ldots + b_M \breve{d}(t) |\breve{d}(t)|^{M-1} \right). \tag{2.8}$$

where $b_{2m+1} = a_{2m+1}/a_1$ and $b_1 = 1$ by definition. In this dissertation we identify the emitter utilizing these nonlinearity parameters $\{b_3, b_5, \ldots, b_M\}$.

This assumption is valid in most OFDM system since the baseband signal is the combination of many frequencies and becomes a Gaussian process because of the central

limit theorem [36].

2.2.1 Bandpass Signal Power Spectrum Modeling

In this section, we utilize the power spectrum density (PSD) of the input signal of the transmitter to derive the device nonlinearities. The autocorrelation function of $z(t)$ and $z(t+\tau)$ is

$$R_{zz}(\tau) = \sum_{k=0}^{\frac{M-1}{2}} \sum_{m=0}^{\frac{M-1}{2}} a_{2k+1} a_{2m+1}^{*} E\left[\left(\check{d}_1\right)^{k+1} \left(\check{d}_1^{*}\right)^{k} \left(\check{d}_2^{*}\right)^{m+1} \left(\check{d}_2\right)^{m} \right] \tag{2.9}$$

where $R_{zz}(\tau) = E[z(t)z^{*}(t+\tau)]$, $\check{d}_1 = \check{d}(t)$ and $\check{d}_2 = \check{d}(t+\tau)$ and $E[\cdot]$ is the expected value operator.

The analysis of power spectrum could be derived by applying (2.8) to (2.4) and some extension work of [37] and [38]. As a result, a general expression of spectrum of the bandpass signal is the Fourier transform of (2.9) and is derived in the following equation:

$$S_{zz}(f) \;=\; \sum_{k=0}^{\frac{M-1}{2}} \sum_{m=0}^{\frac{M-1}{2}} a_{2k+1} a_{2m+1}^{*} FT\left\{ E\left[\check{d}_1 \check{d}_2^{*} \left(\check{d}_1 \check{d}_1^{*}\right)^{k} \left(\check{d}_2 \check{d}_2^{*}\right)^{m} \right] \right\}, \tag{2.10}$$

where $FT\{\cdot\}$ is the Fourier transform.

In case of complex Gaussian random process, [39] provides a result with different moments

$$E\left[(\check{d}_1 \check{d}_2^{*})^{k} \right] = k! R_{\check{d}\check{d}}^{k}(\tau) \quad k \in \text{odd integer}. \tag{2.11}$$

To associate the PSD with (2.10), we use the following equation

$$FT\left\{ R_{\check{d}\check{d}}^{2k+1}(\tau) \right\} = \underbrace{S_{\check{d}\check{d}}(f) * S_{\check{d}\check{d}}(f) * \ldots S_{\check{d}\check{d}}(f)}_{\text{2k+1 times}}, \tag{2.12}$$

where $*$ represents convolution [36]. Therefore the result from (2.10) can be extended to

$$S_{zz}(f) = \sum_{k=0}^{\frac{K-1}{2}} |a_{2k+1}|^2 (2k+1)! FT\left\{ R_{\check{d}\check{d}}^{2k+1}(\tau) \right\}. \tag{2.13}$$

Here we consider odd terms only in the derivation because of their contribution to the

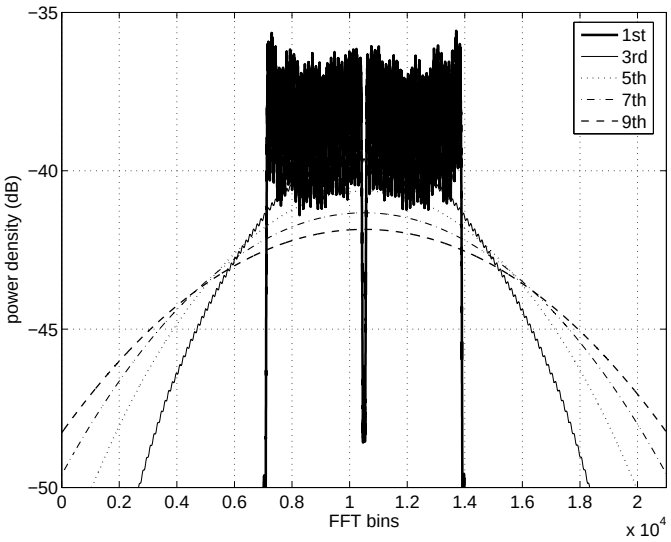


Figure 2.2. Estimated PSD of $\check{d}$ in different harmonic terms.

intermodulation nonlinear distortion.

This circuit-dependent distortion increases the bandwidth of the PSD, and this phenomenon is called spectral regrowth [38]. By extracting the features of the regrown spectrum, the degree of nonlinearity of each odd harmonic terms could be derived directly. The estimated PSD of each harmonic term of $\check{d}$ based on IEEE 802.11a/g specification is depicted in Figure 2.2.

2.3 SEI based on nonlinearity estimation

Suppose $(M-1)/2$ odd harmonic terms are sufficient enough to describe the nonlinearity model without significant loss of characters as in equation (2.4), we introduce $(M-1)/2$ features from the spectrum to estimate the coefficients. Instead of observing the amplitude of PSD at a particular frequency, we use total powers in different frequency bands and obtain a more reliable estimation. The estimated powers in the chosen frequency band can be noted as $P_m,\ m = 1\ldots(M-1)/2$. Assume additive white Gaussian noise (AWGN) channel environment with amplitude attenuation by factor of h, the receive signal is derived using (2.8) as

$$r(t) = hz(t) + n(t). \tag{2.14}$$

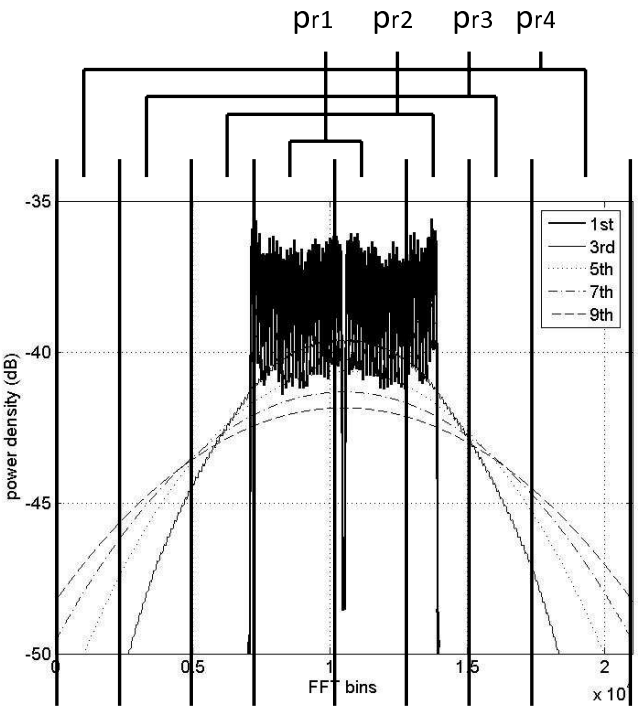


Figure 2.3. Receive signal PSD of an OFDM system is uniformly divided into four frequency bands. The received power of each band is shown as $P_{r1}, \ldots, P_{r4}$

.

The PSD of receive signal is therfore

$$S_{rr}(f) = |h|^2 S_{zz}(f) + \sigma_n^2 \tag{2.15}$$

where $S_{zz}(f)$ is defined in (2.10) and σ_n^2 is the additive noise power.

We uniformly divide the overall frequency band of overall bandwidth BW into $(M+1)/2$ bands, $\{FB_1,\ FB_2, \ldots,\ FB_{(M+1)/2}\}$, and the estimated power in these bands is derived by

$$\begin{aligned} p_{ri} &= \int\limits_{f \in FB_i} S_{rr}(f)df,\ i = 1,\ 2, \ldots,\ (M+1)/2 \\ &= |h|^2 \int\limits_{f \in FB_i} S_{zz}(f)df + \sigma_n^2 \frac{BW}{(M+1)/2}. \end{aligned} \tag{2.16}$$

An example of the definition in (2.16) is shown in Fig. 2.3.

Combining (2.15) and (2.16), the following relationship is derived

$$\begin{bmatrix} p_{r1} \\ p_{r2} \\ \vdots \\ p_{r\frac{M+1}{2}} \end{bmatrix} = \begin{bmatrix} C_{00} & C_{01} & \cdots & C_{0\frac{M-1}{2}} \\ C_{10} & C_{11} & \cdots & C_{0\frac{M-1}{2}} \\ \vdots & \vdots & \ddots & \vdots \\ C_{\frac{M-1}{2}0} & C_{\frac{M-1}{2}1} & \cdots & C_{\frac{M-1}{2}\frac{M-1}{2}} \end{bmatrix}_{\frac{M+1}{2}\times\frac{M+1}{2}} \begin{bmatrix} |h|^2 |a_1|^2 \\ |h|^2 |a_3|^2 \\ \vdots \\ |h|^2 |a_M|^2 \end{bmatrix} + \begin{bmatrix} \sigma_n^2 BW \frac{2}{M+1} \\ \sigma_n^2 BW \frac{2}{M+1} \\ \vdots \\ \sigma_n^2 BW \frac{2}{M+1} \end{bmatrix} \tag{2.17}$$

where C_{ij} is the power of i-th selected area contributed by j-th odd harmonic term, and that is

$$C_{ij} = (2j+1)! \int\limits_{f\in FB_i} \underbrace{S_{\check{d}\check{d}}(f) * S_{\check{d}\check{d}}(f) * \ldots S_{\check{d}\check{d}}(f)}_{\text{2j+1 times}} df \tag{2.18}$$

The above equation could be rewritten as

$$\mathbf{p}_r = \mathbf{C}\mathbf{a}_{sq,h} + \mathbf{n}. \tag{2.19}$$

Hence, the square of nonlinearity parameters $\mathbf{a}_{sq} = \left[|a_1|^2, |a_3|^2, \ldots, |a_M|^2\right]^T$ that characterizes the emitter transmitter can be extracted from (2.19) given $\mathbf{p}_r$, $\mathbf{C}$, $|h|^2$, and $\mathbf{n}$. There are many power spectral density estimators [40] which can be utilized as the power estimator $\hat{\mathbf{p}}_r$, $\mathbf{C}$ is derived using (2.18) with $S_{\check{d}\check{d}}(f)$. Moreover, we assume the noise power is known. Thus, a simple estimation $\hat{\mathbf{a}}_{sq,h}$ is defined as

$$\hat{\mathbf{a}}_{sq,h} = \mathbf{C}^{-1} \cdot (\hat{\mathbf{p}}_r - \mathbf{n}). \tag{2.20}$$

In case of knowledge about path loss $|h|$, the above equation provides $\hat{\mathbf{a}}_{sq}$ directly. However, if the this information is not available, the estimation scaled by $|h|^2$. Thus, we provide a pseudo-estimation for the square of nonlinearity parameters $\mathbf{b}_{sq} = [\ |b_3|^2, |b_5|^2, \ldots, |b_M|^2\ ]^T$ defined in (2.8). Since all of the elements in $\hat{\mathbf{a}}_{sq,h}$ are factored by $|h|^2$, the estimator is derived by dividing the second to last elements of $\hat{\mathbf{a}}_{sq,h}$ by its first

elements as

$$\hat{\mathbf{b}}_{sq} = \begin{bmatrix} \hat{a}_{sq,h,2} \\ \hat{a}_{sq,h,3} \\ \vdots \\ \hat{a}_{sq,h,((M+1)/2)} \end{bmatrix} / \hat{a}_{sq,h,1} \tag{2.21}$$

where $\hat{a}_{sq,h,i}$ is the ith element of $\hat{\mathbf{a}}_{sq,h}$.

2.3.1 Estimator Properties

Since the nonlinearity estimation utilizes PSD estimation, we introduce the performance of the PSD estimator first and use it to derive the nonlinearity estimator performance. In the following derivation, we assume periodogram approach is used for PSD estimator [40] as

$$\hat{S}(f) = \frac{1}{N} \left| \sum_{k=0}^{N-1} x[k] e^{-j2\pi fk} \right|^2 . \tag{2.22}$$

The expected value of the periodogram $\hat{S}(f)$ is equal to the true PSD as $N \to \infty$, but it is not a consistent estimator, meaning its variance does not approach zero with the increasing N. From [40], the expected value and covariance of periodogram are shown in Table 2.1 [40]. In Table 2.1, the expected value of periodogram result is the power

Table 2.1. Periodogram Properties

	Periodogram Properties
Expected Value	$E\left\{\hat{S}_{per}(f)\right\} = S(f) * W_B(f)$
Second-order moment	$E\left\{\hat{S}_{per}(f)\hat{S}_{per}(f)\right\} \approx P(f)P(f)\left\{1 + \left[\frac{\sin(N\pi(f_1-f_2))}{N\sin(\pi(f_1-f_2))}\right]^2\right\}$
Variance	$Var\left\{\hat{S}_{per}(f)\right\} \approx S^2(f)$

spectrum convolved with W_B, which is defined as the discrete time Fourier transform of a triangular window. Utilizing Table 2.1, the expected value of the estimated receive power in band FB_i is

$$E\left[p_{ri}\right] = E\left[\int_{f\in FB_i} \hat{S}_{per}(f)df\right] = \int_{f\in FB_i} E\left[\hat{S}_{per}(f)\right] df = \int_{f\in FB_i} S(f) * W_B(f)df \tag{2.23}$$

which is a biased estimator. Also, the variance is also expressed as

$$\begin{aligned} Var[p_{ri}] &= E\left[p_{ri}^2\right] - E\left[p_{ri}\right]^2 \\ &= E\left[\int_{f_1\in FB_i}\int_{f_2\in FB_i}\hat{S}(f_1)\hat{S}(f_2)df_1df_2\right] - \left[\int_{f\in FB_i}S(f)*W_B(f)df\right]^2 \\ &\approx \int_{f_1\in FB_i}\int_{f_2\in FB_i}S(f_1)S(f_2)\left\{1+\left[\frac{\sin\left(N\pi(f_1-f_2)\right)}{N\sin\left(\pi(f_1-f_2)\right)}\right]^2\right\}df_1df_2 \\ &\quad -\int_{f\in f_1}\int_{f\in f_2}S(f_1)*W_B(f_1)S(f_2)*W_B(f_2)df_1df_2 \end{aligned} \tag{2.24}$$

Considering the case when $N\to\infty$, W_B converges to an impulse; hence, periodogram approach is asymptotically unbiased meaning

$$\lim_{N\to\infty}E\left\{\hat{S}(f)\right\} = S(f). \tag{2.25}$$

From the derivation in 2.25, (2.23) becomes

$$\lim_{N\to\infty}E\left[P_i\right] = \int\limits_{f\in FB_i}E\left\{\hat{S}(f)\right\}df = \int\limits_{f\in FB_i}S(f)*W_B(f)df. \tag{2.26}$$

This shows that the estimated receive signal power is unbiased.

Note that for large N, in Table 2.1,

$$\left[\frac{\sin\left(N\pi(f_1-f_2)\right)}{N\sin\left(\pi(f_1-f_2)\right)}\right]^2 \to 0, \text{ when } |f_1-f_2| \gg 1/N, \tag{2.27}$$

which implies that there is little correlation between one frequency and another. At $f_1 = f_2$, we apply L'Hopital's rule to derive

$$\frac{\sin\left(N\pi(f_1-f_2)\right)}{N\sin\left(\pi(f_1-f_2)\right)} = \left.\frac{\frac{\partial}{\partial\triangle}\sin\left(N\pi\triangle\right)}{\frac{\partial}{\partial\triangle}N\sin\left(\pi\triangle\right)}\right|_{\triangle=f_1-f_2=0} = \left.\frac{(N\pi)\cos\left(N\pi\triangle\right)}{(N\pi)\cos\left(\pi\triangle\right)}\right|_{\triangle=0} = 1. \tag{2.28}$$

Substituting (2.26), (2.25), (2.27), and (2.28) into (2.24), the variance becomes

$$\begin{aligned}\lim_{N\to\infty} Var[P_i] &\approx \left[2\int_{f\in f_i} S^2(f)df + \int\int_{\substack{f_1,\, f_2\in f_i \\ \int f_1\neq f_2}} S(f_1)S(f_2)df_1df_2\right] \\ &- \int\int_{f_1,\, f_2\in f_i} S(f_1)S(f_2)df_1df_2 \\ &= \int_{f\in f_i} S^2(f)df. \end{aligned} \tag{2.29}$$

This result indicates that the variance of the estimate of received power does not converge to zero with the increasing size of observation window.

Fast Fourier transform (FFT) is a practical way to implement a periodogram. By using FFT, the estimation of the power within a frequency band becomes the summation of the PSD estimator bins with some normalized factor. This leads to the conclusion that $\hat{\mathbf{p}}_{ri}$ in (2.19) using FFT is a Gaussian random variable vector, and $\hat{\mathbf{a}}_{sq,h}$ is also a Gaussian vector.

The distribution of the nonlinearity estimation $\hat{\mathbf{b}}_{sq}$ is therefore given by the ratio of two Gaussian random variables. The precise PDF expression of the ratio of two Gaussian random variables is complicated, and we thus approximate it by the fact that $|a_1|^2$ is much greater than other nonlinearity parameters $|a_i|^2$, $i \neq 1$. This phenomenon is caused by the fact that most communication systems bound the degree of nonlinearities in a certain level. For instance, in IEEE 802.11a/g, the nonlinearity is less than 18 dB when 16 QAM modulation scheme is used. Assume $\hat{a}_{sq,h} \approx |a_1|^2$, which indicates the estimation noise is relatively low. Though the denominator term of $\hat{b}_{sq,2i+1} = \hat{a}_{sq,h,2i+1}/\hat{a}_{sq,h,1}$ is not a constant, it is large compared to the numerator and thus the ratio distribution is approximately Gaussian. As a result, when the estimation noise is low, $\hat{\mathbf{b}}_{sq}$ can be modeled as multivariate Gaussian with PDF of

$$p(\mathbf{b}_{sq}) \approx \frac{1}{(2\pi)^{d/2}\,|\Sigma|^{1/2}} exp\left[-\frac{1}{2}(\mathbf{b}_{sq}-\mu_{\mathbf{b}})^T\Sigma^{-1}(\mathbf{b}_{sq}-\mu_{\mathbf{b}})\right], \tag{2.30}$$

where the the element of ith row and jth column in Σ is

$$\sigma_{ij} = E\left[(b_{sq,i}-\mu_i)(b_{sq,j}-\mu_j)\right]$$

$$= E\left[(\mathbf{g}_i^T \cdot \mathbf{p} + \mathbf{g}_i^T \cdot \mathbf{n} - \mathbf{g}_i^T \cdot \mathbf{p})(\mathbf{g}_j^T \cdot \mathbf{p} + \mathbf{g}_j^T \cdot \mathbf{n} - \mathbf{g}_j^T \cdot \mathbf{p})\right] / \hat{a}_{1,sq,h}^2$$
$$= \mathbf{g}_i^T \mathbf{g}_j \cdot \sum_{k=1}^{\frac{M-1}{2}} \sigma_{nk}^2 / \hat{a}_{sq,h,1}^2 \tag{2.31}$$

where $b_{sq,i}$ is the i-th element of $\mathbf{b}_{sq}$, $\mathbf{g}_i^T$ is the i-th row of matrix $\mathbf{G}$, and σ_{nk}^2 is the noise variance of the k-th element in $\mathbf{n}$.

2.3.2 Effects of Input Power Variation

In the above discuss, the power of target transmitters is assumed to be fixed, and in this section the case about transmitters with variable powers would be analyzed. A common way to implement power control is to build an Automatic Gain Control (AGC) before the power amplifier, and change the input power from AGC. Since the signal suffers different degrees of nonlinearity when input power varies, the PSD would also accordingly. Figure 2.4 shows the power spectral density of the signal through the same nonlinear device with three different levels of input power. As can be seen, the spectral regrowth phenomenon is not the same from one to another. The signal with the highest output power has the highest spectral regrowth and suffers the highest degree of nonlinearity, while the signal with lowest output power has the lowest spectra regrowth. Figure 2.4 indicates that even for the same nonlinear device, its PSD will still change with the variation of the input power.

In the following discussion, we model the PSD as a function of the input power and show that the emitter is still identifiable when input power varies. From (2.4) and (2.8), the relationship between the nonlinearity parameters and signal power is shown as

$$b_{2m+1} \sim \sigma_d^{2m},\ m = 1,\ 2, \ldots,\ (M-1)/2 \tag{2.32}$$

where σ_d^2 is the power of output signal. As a result, if the transmit power increases, the values of nonlinearity parameters increases accordingly. Besides, (2.32) also indicates that nonlinearity parameters with higher degrees are more sensitive to the signal power. Since the nonlinearity parameter is a known function of input power, the variation is also predictable.

From the above derivation, the estimated values are functions of input signal power

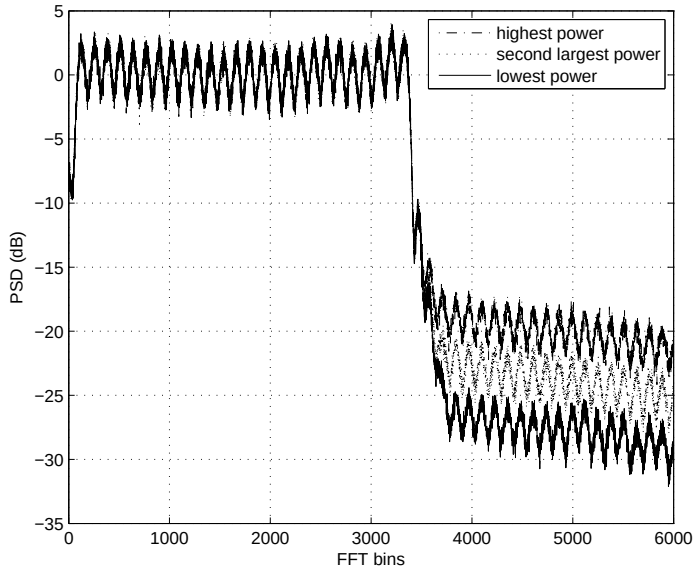


Figure 2.4. Estimated PSD of the one nonlinear device with different levels of input power. The maximum power of these three cases are normalized to the same level.

and we can obtain a general expression about the relationship between each feature as

$$\frac{|b_{2m+1}|}{|b_{2n+1}|} = c_{NL}\sigma_d^{2(m-n)} \tag{2.33}$$

where c_{NL} is a real value number describing the nonlinearities. Further expression of any two parameters is given by

$$|b_{2m+1}| = \frac{|\hat{b}_{2m+1}|}{|\hat{b}_{2k+1}|^{m/k}}|b_{2k+1}|^{m/k}. \tag{2.34}$$

The first term on the right-hand-side is any of the parameter pair of this device which is a constant that characterizes the nonlinear device. In short, the extracted features of a particular emitter would change with power variations but the trajectory of the change could be modeled in equation (2.33). In the following simulation, we will show the trajectory and how the estimated features move with the variation of power.

2.4 Simulation Results

In this section, simulation results are represented in two different cases. Firstly, the input power of each emitter is assumed to be fixed, and then the case of various input power levels will be shown. The simulation was based on WLAN IEEE 802.11a/g specification.

2.4.1 Fixed Power Case

The complex power series coefficients of emitter Ken1 is derived from a class C power amplifier in [35] p77, Ken2 from a class A power amplifier in [35] p76, and Gard1 from [38]. Figure 2.5 depicts two extracted features of the three emitters when input power of each device is fixed. As can be seen in the figure, three emitters can be distinguished in different feature clusters.

Each estimated sample was obtained by observing 315,000 OFDM symbols. In 802.11a case, the symbol period is 4μ seconds and therefore the optimal observation time is about 1.26 seconds for each sample. The subcarrier to noise power ratio is 10-dB, which is a typical required operation environment for BPSK mode, [41]. The probability density function has already been derived in (2.30).

2.4.2 Multiple Power Levels Case

Figure 2.6 shows the simulation results of three emitters, Ken1 and Ken2 as well as Gard1. The input power is stationary in each estimation but changes uniformly within some quantized steps for each estimation. The power is quantized by 1dB; therefore, the estimated samples even from the same emitter are clustered into groups along with a trajectory.

When the power is increased, the nonlinear device output signal suffers more nonlinearity, while it suffers less when the input power decreases. In consequence, the estimated sample might move to the more positive direction along with the line when the power increases and reversely when the power is reduced.

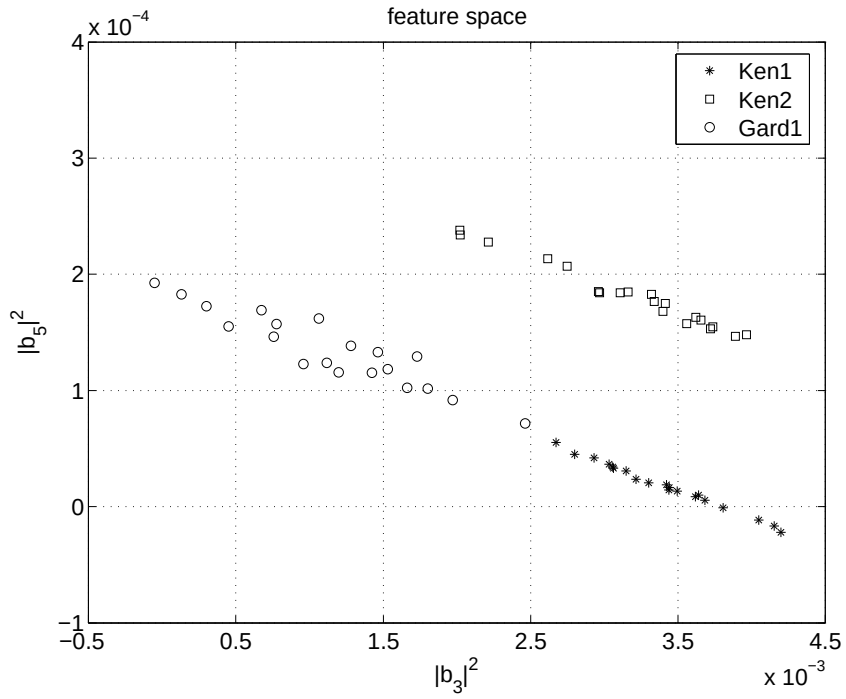


Figure 2.5. The 2-D graphic image of 3rd and 5th order harmonic estimation for three different emitters. $\left(|b_3|^2, |b_5|^2\right)$ of Ken1, Ken2, and Gard1 are supposed to be $\left(3.70 \times 10^{-3}, 6.59 \times 10^{-6}\right)$, $\left(2.61 \times 10^{-3}, 1.99 \times 10^{-4}\right)$, and $\left(6.96 \times 10^{-4}, 1.03 \times 10^{-4}\right)$

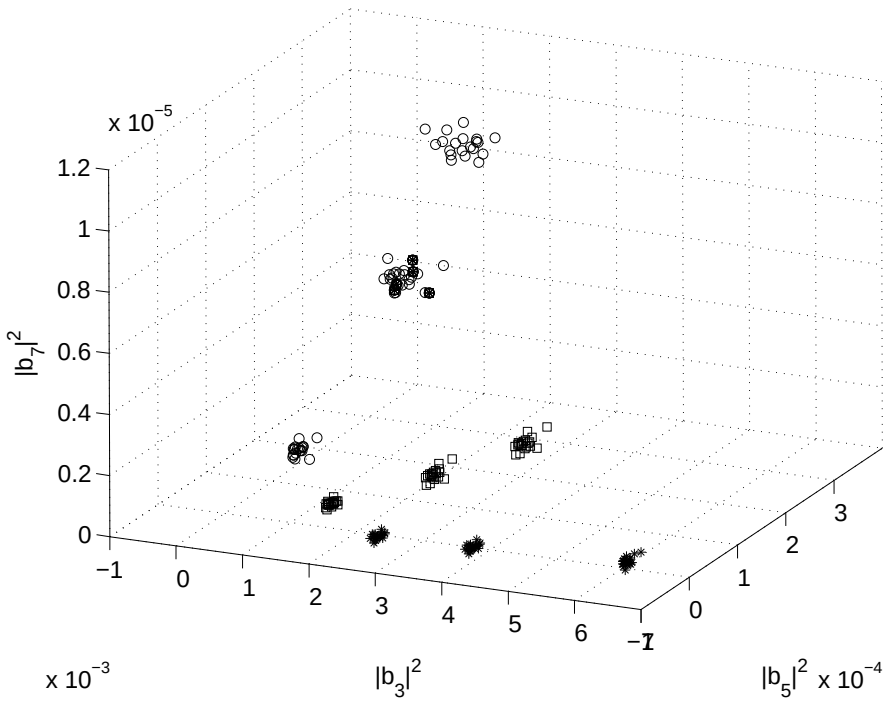


Figure 2.6. Three emitters are represented in 3-D feature space. Samples represented by $\star$ are from Ken1, $\square$ are from Ken2, and $\circ$ are from Gard1.

2.5 Conclusion

In this chapter, a specific emitter identification method is proposed by estimating the nonlinearity of transmitter devices. These features are unique from emitter to emitter due to alternative circuit designs and fabrication randomness. Our algorithm can extract and analyze the nonlinearity coefficients based on the observation of spectral regrowth. Later on, we developed an analytical model that could be used to study the sensitivity of the detection method under AWGN channel environment. For different input power scenarios, we also obtained a model that can predict the trajectory of the moving features. Based on IEEE 802.11a/g specification, our simulation showed that by collecting data for a proper period of time, the proposed method can correctly identify detectable radio emitters in a typical SNR environment.

Chapter 3

Specific Emitter Identification in Multipath Environment

3.1 Introduction

Specific emitter identification (SEI) systems discern wireless radio emitters of interest based only on the external signal feature measurements. This identification capability is used in applications such as battlefield management, electronic support measurement (ESM), intelligence gathering, cognitive radio, and network intrusion detection. This paper presents a SEI algorithm that uses the emitter nonlinearities as the identification features in multipath channels.

A major issue for building SEI systems is to measure proper signal features that are consistent from one transmission to another for a given emitter but different from one emitter to another. These features are selected based on the types of emitter sources, which can be loosely divided into two categories: radar and communications signals. Within the scope of radar signals, there has been a great deal of research on performing SEI by aggregating decisions made by estimating characteristics such as pulse width, the time of arrival (TOA), pulse shape, angle of arrival, carrier frequency, and pulse repetition interval (PRI) [2]. A comprehensive overview of these techniques is presented in [29]. In ESM applications, it is vital for the identification algorithm to estimate PRIs of an interleaved pulse train. Many of these deinterleaving and estimation technologies are based on the TOA difference histogram of the pulse train. For example, a sharp narrow peak indicates a simple PRI, while a staggered PRI gives many separated peaks

in the histogram [13]. Further development to detect the fundamental PRIs using cumulative difference histogram (CDIF) is presented in [8] and a sequential difference histogram (SDIF) approach is proposed in [9]. These algorithms intend to avoid PRI harmonics by combining histograms with sequence search techniques. Another identification method, which exploits the properties of the matrix formed with the differences of TOA, is presented in [10]. In addition, the algorithm in [12] also suppresses the PRI harmonics using PRI spectrum analysis. Other than PRI signal features, [3] and [14] demonstrate SEI technology by exploiting the unintentional modulation on pulse (UMOP) features. Also, the detailed structure in each pulse caused by both intentional and unintentional modulations can be used for SEI, as presented in [5].

In cases when communication systems are emitter sources, many signal characteristics have been investigated including transient amplitude and phase profiles of the radio frequency (RF) waveform [6,18], clock skew [19], I/Q offset, and error vector magnitude (EVM) [20,21]. Another approach is to exploit frequency domain signal characteristics. For instance, the main lobe shape symmetry of the power spectrum density (PSD) [4] as well as spectral correlation function [1] both provide useful identification features. Other than these signal features, [42] utilizes the correlation between the channel response and the emitter location to enhance the physical-layer authentication process in wireless networks to provide spoofing attack detection. A broad analytical framework for physical-layer authentication can also be found in [43]. This paper presents an extension of previous work, [23,26,27], in which the identification of radio emitters relies on the estimation of RF front-end nonlinearities, and the transmitter nonlinearities are unavoidable and unique due to alternative designs and fabrication variations [4].

In prior open SEI research, inter-symbol interference (ISI) has not been studied; however, to achieve reliable estimation, this channel effect has to be suppressed. Most prior results on equalization or channel estimation algorithms derive channel information by either sending training sequences [44] or using a known constellation structure, such as constant modulus algorithm (CMA), multimodulus algorithm (MMA), and reduced constellation algorithm (RCA) [45–48]. As a result, the referred knowledge usually comes from an ideal linear transmission system, and the desired nonlinear characteristics might be obscured when the received signal is forced to fit the referred linear structure. Since our identification depends on the nonlinearity features, applying these algorithms jeopardize the SEI process.

In this chapter, we propose a data-aided time domain algorithm to overcome the ISI problem for RF front-end nonlinearity estimation The estimator suppresses the ISI phenomenon and moreover preserves the nature of the signal features used for emitter identification. The algorithm first utilizes the different effect of the nonlinearity for each signal symbol, ignoring the nonlinearity effect for the symbols with lower amplitudes for the nonlinearity estimation based on symbols with higher amplitudes. After this initial estimation, the accuracy is further improved by an iterative algorithm, which iteratively estimates the transmitted symbol values and the channel coefficients to achieve asymptotically unbiased estimates. The main features of our approach can be summarized as follows.

1. The algorithm provides robust identification by suppressing the ISI effect. In our simulation, the algorithm discerned two transmitters over an empirical indoor model using one packet header from a commercial OFDM system standard.

2. The estimator is asymptotically unbiased and thus guarantees the reliability of identification under the assumption that the nonlinearity parameter set of an emitter is unique.

3. The algorithm is flexible to operate under various communication systems when multiple amplitude level schemes are used such as quadrature amplitude modulation (QAM), pulse amplitude modulation (PAM), and OFDM. Our simulation shows an example of an application on a widely used commercial OFDM system.

The rest of the chapter is organized as follows. In Section 3.2, we give a concise description of the SEI problem in the multipath channel setting. In Section 3.3, we develop our linear approximation algorithm for SEI. In Section 3.4, the iterative estimation algorithm is shown. In Section 3.5, a complexity reduction method is presented. In Section 3.6, numerical results using an OFDM system and a practical indoor channel model are presented. Section 3.7 summarizes the chapter.

3.2 Problem Statement

The objective of SEI is to identify transmitters based on the estimation of the nonlinear transmitter coefficients in the multipath environment. Throughout this paper, we as-

sume that the transmit sequence is known, the transmitter nonlinearity is static, and the channel is time invariant during the estimation.

Also, as a general convention, symbols for matrices are in uppercase bold letters, and vectors are in lowercase bold letters. The notations $(\cdot)^*$, $(\cdot)^{\mathrm{T}}$,$(\cdot)^{\dagger}$, $(\cdot)^{+}$, $E\{\cdot\}$, $||\cdot||$, $\mathbf{0}$, and $\mathbf{I}_q$ stand for conjugate, transpose, conjugate transpose, pseudo-inverse, expected value operator, 2-norm operator, zero vector/matrix, and $q \times q$ identity matrix respectively.

3.2.1 Nonlinearity Model

We first introduce the nonlinear transmitter model in a power-series expansion format, which provides enough degrees of freedom to describe the instantaneous amplitude/amplitude (AM/AM) and amplitude/phase (AM/PM) characteristics of a memoryless nonlinear system [23, 35, 38]. The nonlinear system is described by the input signal d, which is considered a random process in our case, and the nonlinearity coefficients $\{a_1, a_3, \ldots, a_M\}$, where M is an odd number and $(M+1)/2$ is the maximum number of coefficients. The baseband equivalent nonlinearity distortion model is express as $G(d) = \sum_{i=1}^{\frac{M+1}{2}} a_{2i-1} d^i (d^*)^{i-1}$ where both d and a_i are complex quantities, [35, 38]. It is sufficient to consider only the odd terms in the model because the transmitter bandpass filter removes the frequency components resulting from the even terms.

Since we are only interested in the nonlinearities, it is sufficient to use a normalized model

$$x = \sum_{i=1}^{\frac{M+1}{2}} b_{2i-1} \check{d}^i (\check{d}^*)^{i-1} \tag{3.1}$$

where x is the output of the nonlinear system, $b_1 = 1$ by definition, and $\check{d} = d/\sigma_d$ is the normalized version of d with σ_d as the standard deviation of d. Each emitter should contain a set of unique nonlinearity power series coefficients $\{b_3, b_5, \ldots, b_M\}$ due to variations in design, packaging, and fabrication. Our objective is to estimate the nonlinearity coefficients for emitter identification.

In cases where a known training sequence is used, the number of input symbols to the nonlinear system and the number of output symbols in the signal constellation are also determined. In other words, suppose the training sequence provides P different input symbols, $\{\check{d}_1, \check{d}_2, \ldots, \check{d}_P\}$, to the memoryless nonlinearity system, there will be a

maximum of P types of output symbols associated with each symbol as $\{x_1, x_2, \ldots, x_P\}$. For instance, in a 64-QAM modulation system, P can be a positive integer less than or equal to 64. The relationship of these two symbol sets associated with the nonlinearity coefficients can be derived from the nonlinearity model in (3.1):

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_P \end{bmatrix} = \begin{bmatrix} \check{d}_1 \\ \check{d}_2 \\ \vdots \\ \check{d}_P \end{bmatrix} + \begin{bmatrix} \check{d}_1^2\check{d}_1^* & \ldots & \check{d}_1^{\frac{M+1}{2}}(\check{d}_1^*)^{\frac{M-1}{2}} \\ \check{d}_2^2\check{d}_2^* & \ldots & \check{d}_2^{\frac{M+1}{2}}(\check{d}_2^*)^{\frac{M-1}{2}} \\ \vdots & \ddots & \vdots \\ \check{d}_P^2\check{d}_P^* & \ldots & \check{d}_P^{\frac{M+1}{2}}(\check{d}_P^*)^{\frac{M-1}{2}} \end{bmatrix} \begin{bmatrix} b_3 \\ b_5 \\ \vdots \\ b_M \end{bmatrix} \tag{3.2}$$

which has an equivalent compact matrix form

$$\mathbf{x} = \check{\mathbf{d}} + \check{\mathbf{D}}\mathbf{b}. \tag{3.3}$$

Therefore, once $\check{\mathbf{d}}$ is known and $\mathbf{x}$ is measured, the nonlinearity coefficients can be estimated from this relationship.

3.2.2 Nonlinearity Coefficients Derivation

The least-squares (LS) estimation of the nonlinearity coefficients in (3.3) is

$$\hat{\mathbf{b}} = \check{\mathbf{D}}^{+}\left(\hat{\mathbf{x}} - \check{\mathbf{d}}\right), \tag{3.4}$$

where $\hat{\mathbf{x}}$ is the vector of estimated transmit symbols and $\check{\mathbf{D}}^{+}$ is the pseudo-inverse of $\check{\mathbf{D}}$.

The criteria for a unique solution is that, $P \geq (M-1)/2$ and the matrix $\check{\mathbf{D}}$ is full column rank [49]. Taking the null space of $\check{\mathbf{D}}$ into consideration, (3.3) could be re-written $\mathbf{x} = \check{\mathbf{d}} + \check{\mathbf{D}}(\mathbf{b}+\mathbf{c})$ where $\mathbf{c}$ is in the null space of $\check{\mathbf{D}}$. If $\check{\mathbf{D}}$ is full rank, the null space is empty and $\check{\mathbf{D}}^{+}\check{\mathbf{D}} = \mathbf{I}_P$; hence a particular $\mathbf{x}$ maps to a unique $\mathbf{b}$.

3.2.3 Channel Model

In this chapter, we assume a coherent, synchronized symbol-spaced receiver front-end, such that the channel can be represented by an equivalent, discrete-time, baseband linear model [46]: $r[i] = \sum_{k=-\infty}^{\infty} x[k]h[i-k] + n[i]$ where $x[\cdot] \in \{x_1, x_2, \ldots, x_P\}$ is a sampled

transmit symbol, $h[\cdot]$ is the discrete-time channel impulse response, and and $n[\cdot]$ is zero mean Gaussian independent and identically distributed (i.i.d.) complex noise sample.

Assuming a time invariant channel of length L_h with channel coefficients $\mathbf{h} = [h_1, h_2, \ldots, h_{L_h}]^{\mathrm{T}}$, $r[i]$ is equivalent to

$$\begin{bmatrix} x[1] & 0 & \ldots & 0 \\ x[2] & x[1] & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ x[L_{ts}] & x[L_{ts}-1] & \ldots & x[L_{ts}-L_h+1] \\ 0 & x[L_{ts}] & \ldots & x[L_{ts}-L_h] \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & x[L_{ts}] \end{bmatrix} \mathbf{h} + \mathbf{n} = \begin{bmatrix} r[1] \\ \vdots \\ r[L_r] \end{bmatrix} \tag{3.5}$$

where $\mathbf{n} = [n[1], n[2], \ldots, n[L_r]]^{\mathrm{T}}$, $L_r = L_{ts} + L_h - 1$ is the number of received symbols, and L_{ts} is the length of the training sequence, which is typically greater than or equal to the size of the constellation, $P \leq L_{ts}$. Equation (3.5) could be rewritten in compact matrix notation

$$\mathbf{X}\mathbf{h} + \mathbf{n} = \mathbf{r}, \tag{3.6}$$

where $\mathbf{r} = [r[1], r[2], \ldots, r[L_r]]^{\mathrm{T}}$, and $\mathbf{X}$ is the convolution matrix. Straightforward rearrangement of the terms in (3.6) yields the alternate representation

$$\mathbf{H}\mathbf{x} + \mathbf{n} = \mathbf{r} \tag{3.7}$$

where $\mathbf{H}$ is the channel matrix populated by the channel coefficients corresponding to the transmit symbols, and $\mathbf{x}$ is defined in (3.3).

In principle, we estimate $\mathbf{b}$ based on the estimation of $\mathbf{x}$ given only the received signal $\mathbf{r}$, and training sequences, as well as channel length L_h.

3.3 Linear Approximation Approach

As indicated in Section 4.2, once the transmit symbol values $\mathbf{x}$ are provided, the nonlinearity coefficients can be obtained. In this section, we introduce a method that first derives the channel coefficients and then uses the result for transmit symbol estimation,

called the linear approximation (LA) approach. The method is based on the fact that, in multi-amplitude modulation schemes, symbols having lower amplitude levels usually suffer relatively lower level of distortion compared to those at higher levels. As long as the impact is small enough, these symbols with low-amplitude levels can be approximated by the corresponding part of the known training sequence. These approximated symbols are then utilized to estimate the channel coefficients, and the result is used to obtain the full transmit symbol values.

In practice the nonlinearity effects are limited in many communication systems, because of the constraints on power spectrum deviations and constellation errors. For example, one realization of the RF front end output constellation of the short training sequence symbols of IEEE 802.11a/g system [22] is shown in Fig. 3.1; the RF device model is extracted from [38] which satisfies the EVM requirement in [22]. As shown

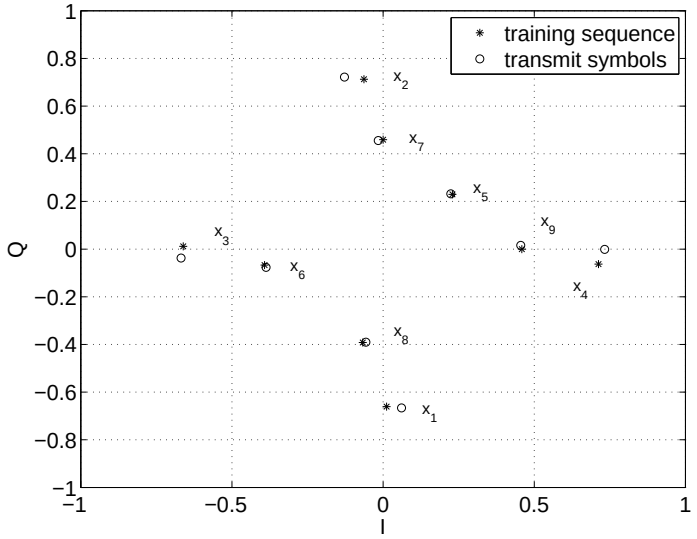


Figure 3.1. The 9-level quantization constellation of short training sequence

in Fig. 3.1, signals with higher amplitudes, such as $\{x_1, \ldots, x_4\}$, have higher nonlinear distortion, while symbols with low amplitudes have lower distortion and could be approximated as linear symbols, meaning $x_p \approx \check{d}_p,\ p = 5,\ \ldots, 9$ in this example. This phenomenon can also be described by the nonlinearity multipliers, $\alpha_p = x_p/\check{d}_p = 1 + b_3|\check{d}|_p^2 + b_5|\check{d}|_p^4 + \ldots + b_M|\check{d}|_p^{M-1}$. The multipliers deviate from 1 more if they represent higher nonlinearities, and the the multipliers of low-amplitude symbols can well be approximated by $\alpha_p = 1$.

Now consider that U of the P total possible symbols have higher nonlinear levels represented by $\mathbf{x}_U = [x_1,\, x_2, \ldots,\, x_U]^{\mathrm{T}}$, and the rest of the $P-U$ possible symbols are in lower nonlinear levels as $\mathbf{x}_L = [x_{U+1},\, x_{U+2}, \ldots,\, x_P]^{\mathrm{T}}$ which could be approximated as $\check{\mathbf{d}}_L = [\check{d}_{U+1},\, \check{d}_{U+2}, \ldots,\, \check{d}_P]^{\mathrm{T}}$. The above approximation also implies $\alpha_i \approx 1$, $i = U+1, \ldots,\, P$. For example, in Fig. 3.1, we can choose $U = 4$ and $P = 9$.

To derive the channel estimate, we can extend the channel coefficient vector to a new $(U+1)L_h \times 1$ vector: $\mathbf{h}_{LA} = \left[\mathbf{h}^{\mathrm{T}} |\, \alpha_1 \mathbf{h}^{\mathrm{T}} | \ldots |\, \alpha_U \mathbf{h}^{\mathrm{T}}\right]^{\mathrm{T}}$, where L_h is the channel impulse response length. By applying $\mathbf{h}_{LA}$ and $\alpha_i \approx 1$, $i = U+1, \ldots,\, P$, we can use (3.6) to approximate $\mathbf{Xh}$ by $\mathbf{D}_{LA}\mathbf{h}_{LA}$, where $\mathbf{D}_{LA}$ consists of elements in $\check{\mathbf{d}}$ only. Thus all the elements in $\mathbf{D}_{LA}$ are known and determined by the training sequence, and symbols are filled in according to their corresponding multipliers. In other words, $\check{d}_i$ should be arranged to a column associated with $\alpha_i \mathbf{h}$ when $i \in \{1,\, 2, \ldots,\, U\}$, and $\check{d}_j$ to $\mathbf{h}$ when $j \in \{U+1, \ldots,\, P\}$ because $\alpha_j \approx 1$. An example is provided in Appendix B.

The LS solution for $\mathbf{h}_{LA}$ is found using pseudoinversion as $\hat{\mathbf{h}}_{LA} = \mathbf{D}_{LA}^{+}\mathbf{r}$ [49]. The estimate of the channel coefficients $\hat{\mathbf{h}}^{(0)}$ is a $L_h \times 1$ vector, obtained by extracting the first L_h elements from $\hat{\mathbf{h}}_{LA}$, i.e., $\hat{\mathbf{h}}^{(0)} = \left[\begin{array}{c|c} \mathbf{I}_{L_h} & \mathbf{0} \end{array}\right] \hat{\mathbf{h}}_{LA}$, where $\mathbf{I}_{L_h}$ is a $L_h \times L_h$ identity matrix and $\mathbf{0}$ is a $L_h \times UL_h$ zero matrix. Once the channel coefficients are obtained, they are used to populate the channel matrix of (3.7). Firstly, we divide $\mathbf{Xh}$ into two terms, containing $\mathbf{x}_U$ and $\mathbf{x}_L$ separately, in which $\mathbf{x}_L$ can be approximated by $\check{\mathbf{d}}$. Now (3.7) can be rewritten as $\mathbf{H}_U\mathbf{x}_U + \mathbf{H}_L\mathbf{x}_L + \mathbf{n} = \mathbf{r}$, where $\mathbf{H}_U$ is the channel matrix associated with $\mathbf{x}_U$, and $\mathbf{H}_L$ is the channel matrix associated with $\mathbf{x}_L$. Here, both $\mathbf{H}_L$ and $\mathbf{H}_U$ are composed of the elements in $\hat{\mathbf{h}}^{(0)}$. Based on this result, the LS estimator of $\mathbf{x}_U$ is

$$\hat{\mathbf{x}}_U^{(0)} = \mathbf{H}_U^{(0)+}(\mathbf{r} - \mathbf{H}_L^{(0)}\check{\mathbf{d}}_L), \tag{3.8}$$

where the superscripts on $\mathbf{H}_U$ and $\mathbf{H}_L$ indicate that elements of the two matrices come from $\hat{\mathbf{h}}^{(0)}$. Also see Appendix B for an example.

The nonlinearity coefficients estimator is derived by modifying (3.4) and (3.8) as follows

$$\hat{\mathbf{b}}^{(0)} = \check{\mathbf{D}}_U^{+}(\hat{\mathbf{x}}_U^{(0)} - \check{\mathbf{d}}_U), \tag{3.9}$$

where $\check{\mathbf{D}}_U$ is the first U rows of matrix $\check{\mathbf{D}}$. The result from (3.9) is used as the initial estimate for the iterative estimation which will be discussed later.

We derive the performance of this approach as follows:

Theorem 1: In high signal-to-noise ratio (SNR) environment with $\alpha_{U+1} = \alpha_{U+2} = \ldots = \alpha_P = \alpha$, the LA estimate is

$$\hat{\mathbf{b}}^{(0)} \approx \mathbf{b} + \left(\frac{1}{\alpha} - 1\right)\left(\mathbf{b} + \check{\mathbf{D}}_U^{+}\check{\mathbf{d}}_U\right) + \frac{1}{\alpha}\check{\mathbf{D}}_U^{+}\mathbf{H}_U^{+}\mathbf{n}. \tag{3.10}$$

proof: See Appendix C.

The second term on the RHS of (3.10) is the estimation bias, which comes from the structure of the symbol constellation and the nonlinearity, and the third term shows the interference driven the additive noise. Furthermore, from the definition $\alpha_p = 1 + b_3|\check{d}|_p^2 + b_5|\check{d}|_p^4 + \ldots + b_M|\check{d}|_p^{M-1}$, we conclude that (3.10) not only provides the analysis of one particular case, it could also be utilized to approximate the performance of other situations, where low amplitude symbols have similar amplitudes.

3.4 Iterative Estimation Algorithm

In this section we will remove the bias in (3.10) using an iterative approach. The objective is to achieve an unbiased estimator, i.e., $E[\hat{\mathbf{b}}^{(k)}] = \mathbf{b}$, as $k \to \infty$, where k is the number of iterations. By removing the bias, the identification ambiguity is reduced; hence, the reliability of identification is enhanced.

Our previous approach provides an estimate of the nonlinearity coefficients, and this information is used to refine the transmit symbol estimation. This estimated symbol vector is treated as initial value to re-estimate the channel condition by substitution into (3.6), called the *h-step*. Then we use this new channel information to re-estimate the symbol coefficients, called the *x-step*, by substituting the information from the *h-step* into (3.7). This result provides new nonlinearity coefficient estimates by applying (3.4) if the iteration stops, or it is sent to the *h-step* for the next iteration. A more detailed description is discussed below.

3.4.1 Initial Value

Though the initial value of the nonlinearity outputs can be randomly assigned, we use the result from Section 3.3 by substituting (3.9) into (3.3) to derive $\hat{\mathbf{x}}^{(0)} = \check{\mathbf{d}} + \check{\mathbf{D}}\hat{\mathbf{b}}^{(0)}$. This estimate is then sent to the *h-step* as the first iteration.

3.4.2 h-step

For the kth iteration, the algorithm uses the symbol estimates in the $(k-1)$th iteration to re-estimate channel coefficients as:

$$\hat{\mathbf{h}}^{(k)} = \mathbf{X}^{(k-1)+}\mathbf{r}. \tag{3.11}$$

Here $\mathbf{r}$ is the received signal in (3.6), and $\mathbf{X}^{(k-1)+}$ is the pesudoinverse of $\mathbf{X}$ but composed of the estimated symbols from the $(k-1)$th iteration, $\hat{\mathbf{x}}^{(k-1)}$, instead of $\mathbf{x}$. This new estimate is then used in the next estimation of the transmit symbol values.

3.4.3 x-step

Analogous to the *h-step*, the estimated channel coefficients, $\hat{\mathbf{h}}^{(k)}$, are then used to populate matrix $\mathbf{H}$ of (3.7) to derive a new symbol estimate:

$$\hat{\mathbf{x}}^{(k)} = \mathbf{H}^{(k)+}\mathbf{r}. \tag{3.12}$$

At the kth iteration the nonlinearity coefficients may be computed from (3.4) and (3.12)

$$\hat{\mathbf{b}}^{(k)} = \check{\mathbf{D}}^{+}(\hat{\mathbf{x}}^{(k)} - \check{\mathbf{d}}). \tag{3.13}$$

Otherwise, the estimated symbols are sent to the *h-step* and the iterative process is repeated.

3.4.4 Estimation Bias and Error Convergence

There are two sources of error in the iterative estimator, one is the initial values error, and the other is driven by the additive noise which aggregates as the number of iterations increases. Since we are interested in an unbiased estimate, the additive noise is first assumed to have no effect on the expected value of (3.13), when analyzing the convergence criterion. This assumption is valid and proved in Appendix D.

For analysis, we set $\hat{\mathbf{x}}^{(k)} = \mathbf{x} + \mathbf{e}_x^{(k)}$, and $\hat{\mathbf{h}}^{(k)} = \mathbf{h} + \mathbf{e}_h^{(k)}$, where $\mathbf{e}_x^{(k)}$ is the error vector of the kth iteration in *x-step*, and $\mathbf{e}_h^{(k)}$ is that of *h-step*. At $k = 0$, the error vector $\mathbf{e}_x^{(0)}$ is inherited from the initial estimation, and our objective is to reduce $\mathbf{e}_x^{(0)}$ after a sufficient

number of iterations.

The evolution of the error vectors between two steps could be expressed as $\mathbf{e}_h^{(k)} = \mathbf{R}_x\mathbf{e}_x^{(k-1)}$, and $\mathbf{e}_x^{(k)} = \mathbf{R}_h\mathbf{e}_h^{(k)}$, where $\mathbf{R}_x$ and $\mathbf{R}_h$ are error transfer functions (see details in Appendix D). Based on this notation, the error vectors in kth iteration could be written as $\mathbf{e}_x^{(k)} = (\mathbf{R}_h\mathbf{R}_x)\mathbf{e}_x^{(k-1)} = (\mathbf{R}_h\mathbf{R}_x)^k\mathbf{e}_x^{(0)}$. Using eigenvalue decomposition [49], $(\mathbf{R}_h\mathbf{R}_x) = \mathbf{V}\sum\mathbf{V}^{-1}$, where matrix $\mathbf{V}$ contains eigenvectors and eigenvalue matrix $\sum = diag(\lambda_1,\ \lambda_2, \ldots,\ \lambda_r,\ 0, \ldots, 0)$. Furthermore, for the kth iteration, $(\mathbf{R}_h\mathbf{R}_x)^k = \mathbf{V}\sum^k\mathbf{V}^{-1}$. Thus we have $\lim_{k\to\infty}(\mathbf{R}_h\mathbf{R}_x)^k = \mathbf{0}$ if $\rho(\mathbf{R}_h\mathbf{R}_x) = max\{|\lambda_1|, \ldots, |\lambda_r|\} < 1$, [50]. Once the symbol estimator is unbiased, meaning $E[\hat{\mathbf{x}}^{(k)}] = \mathbf{x}, k \to \infty$, the nonlinearity estimation is also unbiased. This could be easily proved by simply substituting the estimated symbols into (3.13).

As a result, once the condition $\rho(\mathbf{R}_h\mathbf{R}_x) < 1$ is satisfied, the nonlinearity estimator is asymptotically unbiased.

3.4.5 i- step

Whenever $\rho(\mathbf{R}_h\mathbf{R}_x) = 1$, the error vector $\mathbf{e}_x^{(k)}$ will saturate at an equilibrium state which is related to the initial value. To solve this problem, we add an intermediate step after *h-* and *x-step* that reduces the maximum eigenvalue of the overall system and satisfies the convergence requirement.

A natural augmentation is to constrain the estimation result by utilizing the nonlinearity model in (3.3). Suppose we have more symbols, $\mathbf{x}$, than necessary to estimate the nonlinearity coefficients, then the estimated symbol values are projected to the solution plane of $\mathbf{b}$, removing the error perpendicular to the plane. Intuitively, if this projection provides a more accurate result, then the corresponding transmit symbol values provide a better channel estimation in the *h-step* and the procedure can be repeated. That is, after the *h-* and *x-step* are executed, $\hat{\mathbf{x}}^{(k)}$ is substituted into (3.4) and $\hat{\mathbf{b}}^{(k)}$ is calculated and substituted into (3.3), the projected result is $\check{\mathbf{x}}^{(k)} = \check{\mathbf{d}} + \mathbf{P}_{\check{D}}\left(\hat{\mathbf{x}}^{(k)} - \check{\mathbf{d}}\right)$ where $\mathbf{P}_{\check{D}} = \check{\mathbf{D}}\check{\mathbf{D}}^{+}$ is the projection matrix of $\check{\mathbf{D}}$ [49]. This estimate is sent to the *h-step* instead of the result from the *x-step* and the iteration is repeated.

After a sufficient number of iterations is taken, the estimator derives an estimate of distorted transmit symbol vector $\hat{\mathbf{x}}^{(k)}$ from (3.12), which is then used to obtain the transmitter nonlinearity coefficients using (3.13). If we replace $\hat{\mathbf{x}}^{(k)}$ by $\check{\mathbf{x}}^{(k)}$, the nonlinearity

estimates are identical (see Theorem D.1 in Appendix D). Hence, considering complexity, instead of $\check{\mathbf{x}}^{(k)}$, the estimator uses $\hat{\mathbf{x}}^{(k)}$ to derive the nonlinearity coefficients to save one $i-step$ operation.

To investigate the behavior of the nonlinearity estimator, the estimation error is given by (see Theorem D.2 in Appendix D):

$$\hat{\mathbf{b}}^{(k)} - \mathbf{b} = \mathbf{Q}_b^{(k)} \mathbf{e}_x^{(0)} + \mathbf{W}_b^{(k)} \mathbf{n}. \tag{3.14}$$

The first term on the RHS of (3.14) is contributed by the initial estimation error, which results in the estimation bias, and the second term on the RHS is caused by additive noise. Moreover, to achieve the unbiased estimator requirement $E[\hat{\mathbf{b}}^{(\infty)}] = \mathbf{b}$, the initial estimation error term in (3.14) must converge to a zero vector. As a result, the estimator is asymptotically unbiased if the spectral norm of $\mathbf{R}$ is less than 1, i.e., $\rho(\mathbf{R}) < 1$, where $\mathbf{R} = \mathbf{P}_{\check{D}} \mathbf{R}_h \mathbf{R}_x$. Once $\rho(\mathbf{R}) < 1$ is satisfied, the first term on the RHS of (3.14) converges to a zero vector, and the second term on the RHS of the same equation approaches a fixed vector (see Theorem D.3 in Appendix D). We note that from our experimental observations, the required number of symbols of different amplitudes usually has to be greater than M to satisfy the spectral norm criterion.

3.4.6 Stopping Criterion

In this section, we provide a stopping criterion to achieve an acceptable misclassification rate. Our stopping criterion is to stop the iteration when the difference of two estimation results is below a threshold. This threshold stops the iteration when the bias is small and the estimator nears the noise floor. It does not guarantee unbiased estimation, since it is an asymptotically unbiased estimator.

This criterion is based on the assumption that the initial estimation error is larger than the additive noise term at the beginning of the iteration. From (3.14) we know that the bias term is reduced with iteration number, while the additive noise term increases and then saturates (see Remarks of Theorem D.3 in Appendix D). Hence, once the difference is below a certain threshold, we expect the bias is sufficiently suppressed and the noise floor is also met.

We summarize the estimation procedure in Algorithm 1 , where the linear approximation approach is used as the initial estimation and the result is then utilized in the

iterative estimation.

Algorithm 1 Transmitter Nonlinearity Estimation

1: *Initialization:* Set the iteration number $k = 0$, the maximum number of iterations k_{max}, the stopping threshold $\varepsilon > 0$.
2: *Initial Estimation:* $\hat{\mathbf{x}}^{(0)} = \check{\mathbf{d}} + \check{\mathbf{D}}\hat{\mathbf{b}}^{(0)}$
3: **repeat**
4: $\quad k \leftarrow k+1$
5: $\quad$ *h-step:* $\hat{\mathbf{h}}^{(k)} = \mathbf{X}^{(k-1)+}\mathbf{r}$
6: $\quad$ *x-step:* $\hat{\mathbf{x}}^{(k)} = \mathbf{H}^{(k)+}\mathbf{r}$
7: $\quad$ *i-step:* $\check{\mathbf{x}}^{(k)} = \check{\mathbf{d}} + \mathbf{P}_{\check{D}}\left(\hat{\mathbf{x}}^{(k)} - \check{\mathbf{d}}\right)$
8: **until** $k = k_{max}$ or $\left\|\check{\mathbf{x}}^{(k)} - \check{\mathbf{x}}^{(k-1)}\right\| < \varepsilon$
9: *Output:* $\hat{\mathbf{b}}^{(k)} = \check{\mathbf{D}}^{+}(\hat{\mathbf{x}}^{(k)} - \check{\mathbf{d}})$

3.5 Complexity Reduction

We introduce a complexity reduction in the form of a gear-shifting method for high SNR conditions. The method first observes the trajectory of the previous estimated features, then predicts the next iteration result. After the prediction, the algorithm simply moves the current estimation to the prediction result. The reason that the feature trajectory is predictable is based on the following theorem.

Theorem: Assume the eigenvalues of $\mathbf{R}$ are $\{\lambda_1, \lambda_2, \ldots, \lambda_r\}$ with $\rho(\mathbf{R}) < 1$, and the kth power of the unique maximum eigenvalue is much greater than that of other eigenvalues, $\lambda_{max}^k \gg \lambda_i^k$ and $\lambda_{max} \neq \lambda_i$. Moreover, consider a high SNR environment with $\mathbf{Q}_b^{(k)}\mathbf{e}_x^{(0)} \gg \mathbf{W}_b^{(k)}\mathbf{n}$, then

$$\hat{\mathbf{b}}^{(k+2)} - \hat{\mathbf{b}}^{(k+1)} \approx \lambda_{max}(\hat{\mathbf{b}}^{(k+1)} - \hat{\mathbf{b}}^{(k)}). \tag{3.15}$$

Proof: Based on the same assumptions, we derive the kth power of the matrix $\mathbf{R}^k \approx \lambda_{max}^k \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$ where $\mathbf{V}$ is the eigenvector of matrix $\mathbf{R}$, and $\mathbf{\Lambda}$ is the diagonal eigenvalue matrix with a 1 in the location of the maximum eigenvalue of eigenvalue matrix $\mathbf{\Sigma}$ while the rest of the elements are zero. In the high SNR environment, the additive noise is weak, meaning $\mathbf{Q}_b^{(k)}\mathbf{e}_x^{(0)} \gg \mathbf{W}_b^{(k)}\mathbf{n}$; then, the difference of two estimated results is obtained by using (3.14): $\hat{\mathbf{b}}^{(k+2)} - \hat{\mathbf{b}}^{(k+1)} \approx \left(\lambda_{max}^{k+1} - \lambda_{max}^k\right)\left(\check{\mathbf{D}}^{+}\mathbf{R}_h\mathbf{R}_x\mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}\mathbf{e}_x^{(0)}\right)$.

Similar to the above equation, the difference between $(k+3)$th and $(k+2)$th iteration is $\hat{\mathbf{b}}^{(k+3)} - \hat{\mathbf{b}}^{(k+2)} \approx \lambda_{max}(\hat{\mathbf{b}}^{(k+2)} - \hat{\mathbf{b}}^{(k+1)})$. The ratio of these two differences is proportional to λ_{max}, and thus provides the predictability of the next estimation.

As the number of iterations increases, the $\mathbf{Q}_b^{(k)}\mathbf{e}_x^{(0)} \gg \mathbf{W}_b^{(k)}\mathbf{n}$ assumption might not hold, since $\mathbf{Q}_b^{(k)} \to \mathbf{0}$ when k is large.

To determine the starting iteration of prediction k_p that satisfies the above requirements, we observe the estimates trajectory. When (3.15) is satisfied, the ratio of the difference of two estimates is approximately λ_{max}; therefore, we can determine the starting iteration by checking if $\left\| \frac{\hat{\mathbf{b}}^{(k_p-1)} - \hat{\mathbf{b}}^{(k_p-2)}}{\hat{\mathbf{b}}^{(k_p-2)} - \hat{\mathbf{b}}^{(k_p-3)}} - \frac{\hat{\mathbf{b}}^{(k_p)} - \hat{\mathbf{b}}^{(k_p-1)}}{\hat{\mathbf{b}}^{(k_p-1)} - \hat{\mathbf{b}}^{(k_p-2)}} \right\| < \delta$, where δ is a positive small value.

Suppose the above requirement is satisfied at k_pth iteration, then the prediction of the features is

$$\hat{\mathbf{b}}^{(k)} := \hat{\mathbf{b}}^{(k)} + \mu\hat{\lambda}_{max}\left(\hat{\mathbf{b}}^{(k)} - \hat{\mathbf{b}}^{(k-1)}\right),\ k > k_p \tag{3.16}$$

where the positive real number $\mu < 1$ is the step size and $\hat{\lambda}_{max}$ is the estimated eigenvalue using the ratio of estimation differences described previously. This prediction is then used to derive the *i-step* result $\check{\mathbf{x}}^{(k)} = \check{\mathbf{d}} + \check{\mathbf{D}}\hat{\mathbf{b}}^{(k)}$. When $\mu = 0$, the algorithm is the same as our original iterative algorithm using only a single step size (SSS). When μ varies in different iterations this is referred to as multiple step size algorithm (MSS). In other words, the algorithm sets $\mu = 0$ before the feature trajectory is stable, and assigns another value to μ afterwards.

3.6 Numerical Results

In the simulation, a practical example using an OFDM system, IEEE 802.11a/g Specification [22], with the RF front-end device models from [38] as transmitter 1, and [35] as transmitter 2, which were used in our previous research [23] is shown. We apply an empirical indoor channel model, $h[n] = 0.9960\delta[n] + 0.0628\delta[n-1] + 0.0079\delta[n-2]$, extracted from an experimental result in [51]. The degree of nonlinearity is -16 dB EVM satisfying the specification [22] when operating in 16-QAM scheme. The output power is set to 25 dBm for each emitter, which is a typical average transmit power value in commercial market products, and the nonlinearity coefficients are shown in Table 3.1. Also, one metric we are interested in is the normalized mean squared error (NMSE) of

Table 3.1. Nonlinearity parameters of two transmitters

	Transmitter 1	Transmitter 2
b_3	$-0.0735-0.0114i$	$-0.0910+0.1580i$
b_5	$-0.0986+0.0590i$	$0.2503+0.0286i$
b_7	$-0.0547-0.0055i$	$0.0155+0.0025i$

the nonlinearity coefficients, which is defined as

$$\text{NMSE of } \mathbf{b} \equiv 10\log\left(E\left[\left\|\hat{\mathbf{b}}^{(k)}-\mathbf{b}\right\|^2/\|\mathbf{b}\|^2\right]\right)(\text{dB})\,. \tag{3.17}$$

The simulation shows the performance of NMSE and misclassification rate under different conditions such as SNR and number of iterations.

Our algorithm initially estimates transmit symbols using the LA approach based on the long training sequence [22] of length 160 samples containing 64 different transmit symbols where the nonlinearities of the 52 lowest amplitude symbols are considered insignificant and those symbols are approximated by linear symbols. After the initial estimation, the same training sequence is utilized for the iterative estimation. Based on this simulation environment, we assign parameters as $M=7$, $L_h=3$, $L_{ts}=160$, $P=64$, and $U=64-52=12$.

Fig. 3.2 shows the evolution of NMSE with number of iterations. The simulation result shows that the NMSE decreases for the first several iterations and then either slowly rises or saturates to a noise floor. The decrease of NMSE is mainly due to the reduction of initial estimation error; the aggregation of the additive noise increases the noise level and forms the noise floor as described in in Remarks of Theorem D.3 in Appendix D.

Also, as expected, the performance improves as the number of packet headers accumulates. From the simulation results, the minimum NMSE improves approximately 3 dB when the headers used are doubled, as shown in Fig. 3.3. Since the OFDM system operates using a 16QAM scheme, a general operating environment is 12 to 16 dB E_b/N_0 [41]. Fig. 3.3 depicts the minimum NMSE for different E_b/N_0 conditions, using an ensemble average of the 1000 realizations. To show the 3 dB improvement when the data used is doubled, we first use linear fit for 1 header experiment result and move the

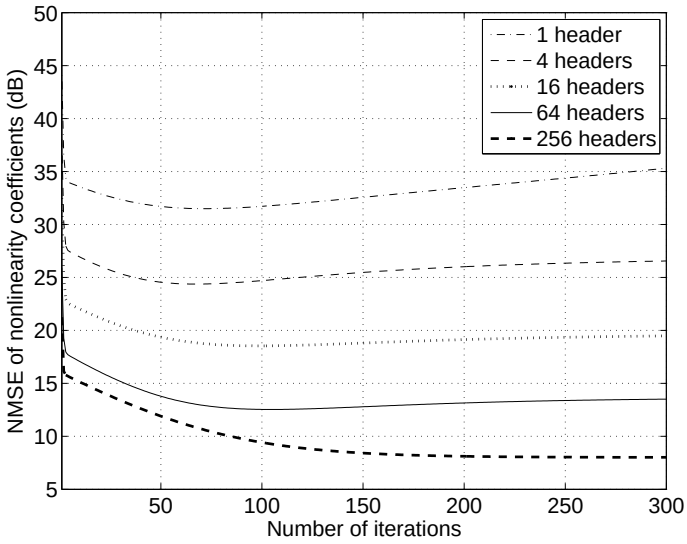


Figure 3.2. NMSE of the nonlinearity coefficients under E_b/N_0 of 10 dB environment with different number of packet headers used.

line 3 dB away to draw the linear fit for 2 headers case, then 3 dB lower for the 4 headers case, and then another 3 dB for the 8 and 16 headers case. Fig. 3.4 depicts the estimated

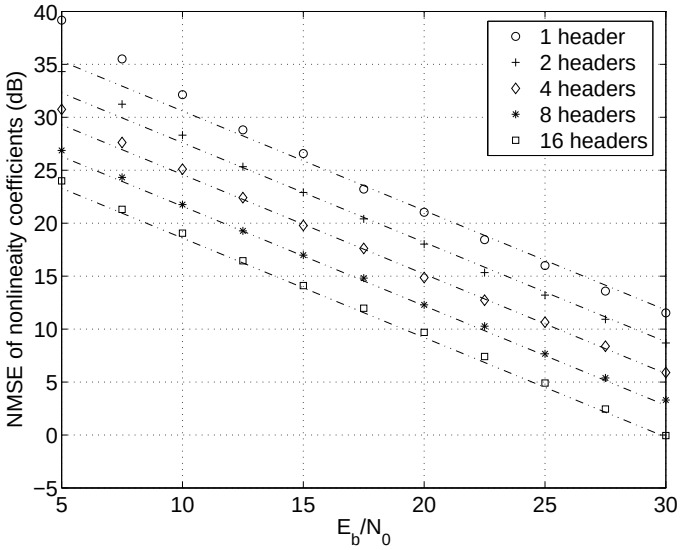


Figure 3.3. NMSE of the nonlinearity coefficients versus E_b/N_0 with different number of headers.

feature, b_3 from (3.1), for two different RF front-end models after 80 iterations, and the

results also show that the two emitters are identifiable.

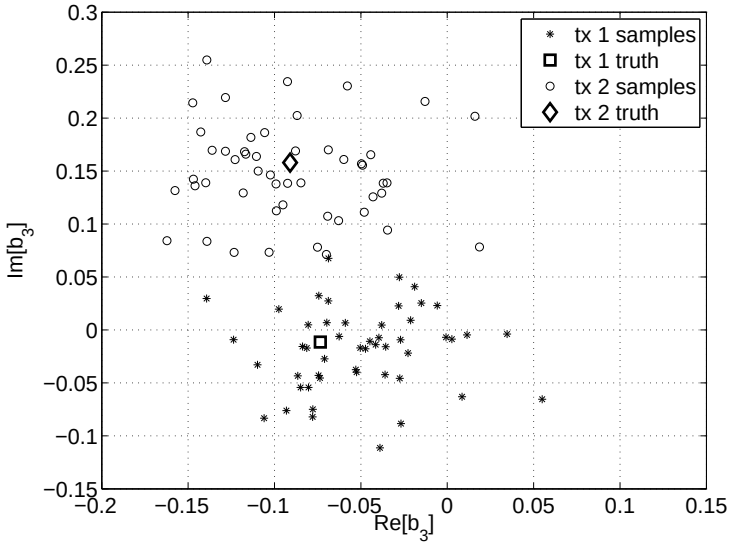


Figure 3.4. One of the estimated features observed from two transmitters in Table 3.1, with 30 dB E_b/N_0 and 16 packet headers for each sample.

Fig. 3.5 shows the misclassification rate versus number of iterations when $E_b/N_0 =$ 10 dB and 4 headers are used. Compared with Fig. 3.2, both misclassification rate and NMSE saturate at approximately the same number of iterations. In Fig. 3.5, we assumed the nonlinearity coefficients of both transmitters are known, and the NMSE at each iteration is given. Furthermore, the decision is determined simply by the minimum Euclidean distance between the estimated sample and nonlinearity coefficients. As a result, we stop the iterative process when NMSE meets the noise floor, and then classify the transmitter.

Fig. 3.6 presents the misclassification rates of a minimum error rate classifier [52]. Due to the correlation between estimate coefficients, a minimum distance classifier performs worse than a minimum error rate classifier. Thus we only show the numerical results of the minimum error rate classifier, which utilizes the covariance matrices and Bayes decision rule. The numerical results show that it requires 2 headers to achieve 30% misclassification rate when E_b/N_0 is 10 dB, while 16 headers are necessary for 10 percent misclassification rate. Also, to achieve 1% misclassification rate, each doubling of the headers provides 3 dB gain as shown in Fig. 3.6. For further evaluation, the Chernoff error Bound could be derived [52], and our simulation result satisfies this error

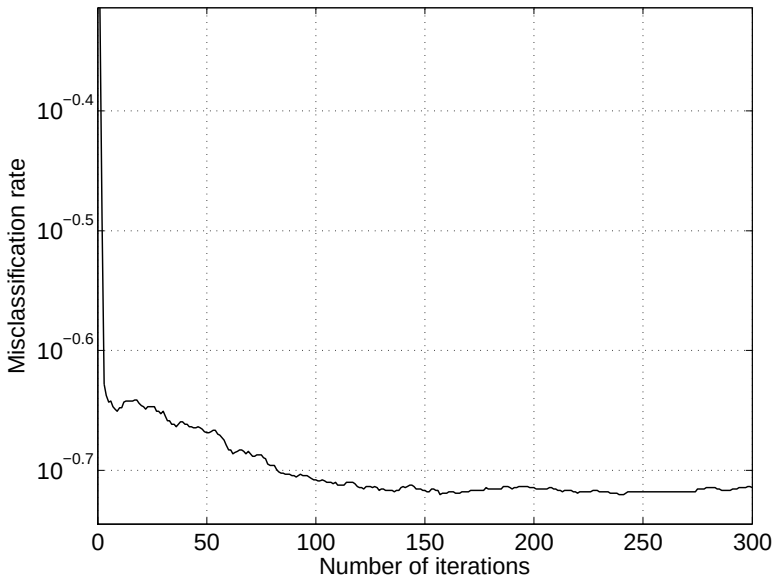


Figure 3.5. Misclassification rate of the two transmitters versus number of iterations.

bound.

As described in Section 6, both SSS and MSS scenarios are presented in the following simulation. Fig. 3.7 shows an example of using single and multiple step size results in E_b/N_0 of 10 dB environment with four headers data. SSS approach sets $\mu = 0$ from (3.16) throughout the estimation while the MSS approach uses $\mu = 0$ in the first 6 iterations and then switches to $\mu = 0.99$ afterwards. This example demonstrates the iteration result is predictable due to its regularity of the trajectory after the first several iterations. Also, the figure shows the reduction of complexity by depicting that MSS uses 12 iterations to achieve a similar result instead of 18 iterations. The comparison of the performance using alternative μ is shown in Fig. 3.8. After 300 iterations, the NMSE results of all three scenarios are within one dB difference, however the scenarios using MSS algorithm require less iterations to reach the lowest NMSE than using SSS.

3.7 Conclusion

In this chapter, we have presented an RF front-end nonlinearity estimator for SEI systems when multilevel modulation schemes are used. Our method overcomes the ISI effect and reduces the initial estimation error to arbitrarily small levels. The algorithm first utilizes the partitioning of symbol levels and assumes some of them are linear to

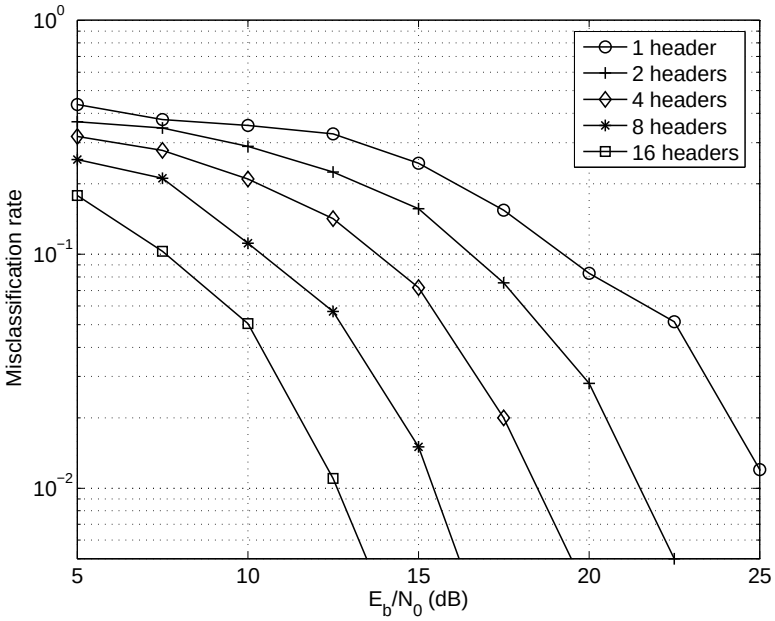


Figure 3.6. Misclassification rate in different E_b/N_0 situations. The classification is determined when the minimum NMSE iteration is achieved.

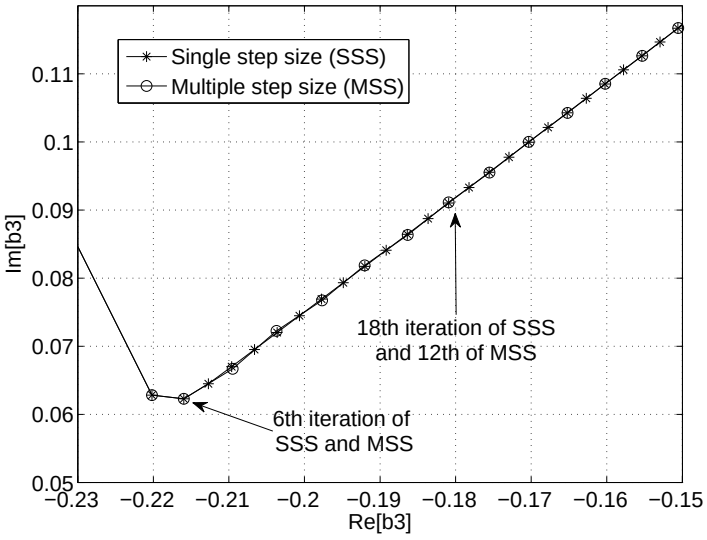


Figure 3.7. The estimated b_3 at each iteration applying SSS and MSS.

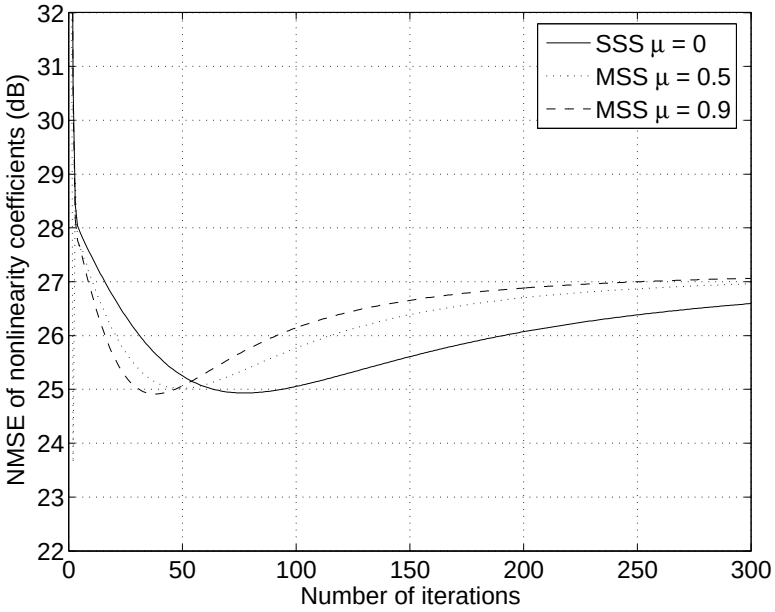


Figure 3.8. The estimated b_3 of each iteration applying SSS and MSS.

estimate the nonlinearity; further improvement is achieved by iterative estimation of the transmit symbol values and the channel coefficients, and the over-fitting of estimated nonlinearity coefficients to the model. We also provide a complexity reduction method which predicts the trajectory of signal features to reduce the number of iterations. Simulations have shown that the proposed method could utilize the time domain samples of an OFDM system as a multilevel modulation scheme, and two alternative emitters are identifiable with a 95% correct classification rate using 16 headers under an empirical indoor channel model.

Chapter 4

MIMO Channel Estimation with Transmitter Nonlinearities

4.1 Introduction

The demand for increasing capacity and quality of service in wireless communication has drawn extensive research on multiple-input multiple output (MIMO) systems that involves multiple antennas on the transmitter or/and receiver sides. Several related technologies such as space-time coding and beamforming are effective in improving the performance; however, the improvement depends highly on the precision of channel estimation available at both the transmitter and receiver side [24, 53–55]. While many channel estimators have been proposed, most analysis methods consider linear transmitter design under the assumption that the nonlinear properties are negligible. Though a highly linear Radio Frequency (RF) front-end is desirable, the design methods used in practice give rise to imperfections [35, 38, 56–58]. For example, a typical power amplifier affects both amplitude (AM) as AM/AM conversion and phase (PM) as AM/PM conversion in a wide range of operational environments [35]. As a result, various communication specifications mention upper limits of nonlinearity in terms of the Error-Vector-Measurement (EVM) and spectrum mask as the upper linearity limits [22, 59]. It is therefore relevant to study the nonlinear channel estimation caused by the RF front-end nonlinearity. Several techniques that combat the effects at the transmitter side have been proposed, and they consider either single input single output (SISO) systems, or flat fading MIMO systems [60–62].

In this chapter, we propose a data-aided asymptotically unbiased nonlinear channel estimator operating in frequency selective fading MIMO channels involving intersymbol interference (ISI) effects, and SISO channel is a special case for our algorithm. We assume that the channel is stationary during the estimation interval, and the transmitter nonlinearity is memoryless and static. Our method first exploits the differentiation of nonlinearity in each signal symbol, and ignores the nonlinearity of those symbols with lower amplitudes to derive our initial channel estimation. To accomplish asymptotically unbiased estimation, the accuracy is improved by an iterative estimation algorithm, which iteratively estimates the transmitted symbol values and the channel coefficients. The proposed algorithm, which overcomes the ISI phenomenon, is designed to operate under multiple amplitude level modulation schemes such as Quadrature Amplitude Modulation (QAM), Orthogonal Frequency Division Multiplexing (OFDM), and Pulse Amplitude Modulation (PAM).

The chapter is organized as follows. Section 4.2 gives the concise description of the problem including the nonlinearity and channel model. We then present the linear approximation algorithm in Section 4.3, and the iterative estimation algorithm in Section 4.4. In Section 4.5, numerical results using an MIMO OFDM system and a practical indoor MIMO channel model are shown. Section 4.6 gives the summary of the chapter.

4.2 Problem Statement

As indicated in the previous section, while the receiver is assumed linear, the description of the nonlinear channel in this chapter is composed of the linear channel model and the nonlinearity coefficients of the transmitter RF front-end. In this section, the memoryless nonlinear model used in the transmit side and linear MIMO channel model are introduced. Also, since, from the same transmitter set, the RF front-ends of all the antennas likely share the same design [63, 64], we assume the nonlinear properties are identical throughout this chapter. Further extensions using different nonlinear properties are possible.

4.2.1 Memoryless Nonlinear Model

The input signal, d, of the nonlinear system is assumed a random process, and then by applying the nonlinearity distortion analysis from [35], [38], [23], the memoryless nonlinear system could be described by its coefficients as:

$$G(d) = \sum_{i=1}^{\frac{M+1}{2}} a_{2i-1} d^i (d^*)^{i-1} \tag{4.1}$$

where both d and a_i are complex numbers with their normalized forms $\check{d}$ and $\check{a}_i$. The normalization is set to a reference structure with standard deviation of σ_d, and hence $\check{a}_i = a_i \cdot (\sigma_d)^i$ and $\check{d} = d/\sigma_d$. In this chapter M is assumed the maximum number of nonlinearity coefficients.

Our main concern in this chapter is the nonlinearities of the transmitters, and the linear gain is not in the scope of this research. Hence, the model of (4.1) is simplified by normalizing it as $x = \sum_{i=1}^{\frac{M+1}{2}} b_{2i-1} \check{d}^i (\check{d}^*)^{i-1}$ which is also shown in (3.1). where $b_{2j-1},\ j = 1,\ldots,(M+1)/2$ are the normalized nonlinearity coefficients in complex format, and b_1 is set to 1 by definition.

Since our proposed approaches are data-aided and the training sequence is known, the outcome of the nonlinear system is assumed finite. The input symbols belong to $\{\check{d}_1, \check{d}_2, \ldots, \check{d}_P\}$, and their one-to-one mapped outputs belong to $\{x_1, x_2, \ldots, x_P\}$, where P is dimensionality of the transmit symbols. For instance, $P \leq 64$ in a 64QAM scheme, and $P \leq 8$ in an 8PAM scheme. Our approach will estimate the nonlinear coefficients $\{b_3, b_5, \ldots, b_M\}$ which are identical for all the RF front-ends through the estimation of $\{x_1, x_2, \ldots, x_P\}$. The relationship between the input and output can further be expressed as (3.3) in matrix format $\mathbf{x} = \check{\mathbf{d}} + \check{\mathbf{D}}(\mathbf{b} + \mathbf{c})$ where $\mathbf{c}$ is in the null space of $\check{\mathbf{D}}$.

Assume there are more symbols than unknown nonlinearity coefficients, $P > \frac{M-1}{2}$, and the estimated symbols contain error as $\hat{\mathbf{x}} = \mathbf{x} + \mathbf{e}_x$, the least-squared estimation of nonlinearity coefficient vector $\hat{\mathbf{b}}$ is

$$\hat{\mathbf{b}} = \check{\mathbf{D}}^+ (\hat{\mathbf{x}} - \check{\mathbf{d}}) - \mathbf{c} = \mathbf{b} + \check{\mathbf{D}}^+ \mathbf{e}_x - \mathbf{c}. \tag{4.2}$$

As long as matrix $\check{\mathbf{D}}$ is full rank, the null space $\mathbf{c}$ is empty; hence a set of $\hat{\mathbf{x}}$ only maps to one unique $\hat{\mathbf{b}}$.

4.2.2 MIMO Channel Model

Consider wide-band linear MIMO channel model with n_T transmit antennas and n_R receive antennas. The signal at the receive antenna array are denoted $\mathbf{r}[\tau] = [r(1)[\tau], \ldots, r(n_R)[\tau]]^{\mathrm{T}}$, where $r(j)[\tau]$ is the signal at the j-th receive antenna and τ is the time index. Also, the signal from transmit array is denoted as $\mathbf{x}[\tau] = [x(1)[\tau], \ldots, x(n_T)[\tau]]^{\mathrm{T}}$, where $x(i)[\tau] \in \{x_1, x_2, \ldots, x_P\}$ is the signal at the i-th transmit antenna.

The linear dispersive channel can be represented by a baseband discrete-time equivalent model as [65–68]

$$\tilde{\mathbf{H}}[\tau] = \sum_{\tau_l=0}^{L_h-1} \bar{\mathbf{H}}[\tau_l]\delta[\tau - \tau_l] \tag{4.3}$$

where L_h is the channel length, and $\tilde{\mathbf{H}}[\tau] \in \mathbb{C}^{n_R \times n_T}$. Here $\bar{\mathbf{H}} \in \mathbb{C}^{n_R \times n_T}$ describes the linear transformation between the transmit antennas and receive antennas at delay τ_l, and is defined as

$$\bar{\mathbf{H}}[\tau_l] = \begin{bmatrix} h_{11}[\tau_l] & h_{12}[\tau_l] & \ldots & h_{1n_T}[\tau_l] \\ h_{21}[\tau_l] & h_{22}[\tau_l] & \ldots & h_{2n_T}[\tau_l] \\ \vdots & \vdots & \ddots & \vdots \\ h_{n_R1}[\tau_l] & h_{n_R2}[\tau_l] & \ldots & h_{n_Rn_T}[\tau_l] \end{bmatrix}, \tag{4.4}$$

where h_{ji} is the complex transmission coefficient from i-th transmit antenna to j-th receive antenna.

The relationship between the j-th receive element and the transmit symbols can be written as

$$r(j)[\tau] = \sum_{i=1}^{n_T} \sum_{\tau_l=0}^{L-1} h_{ji}[\tau_l]x(i)[\tau - \tau_l] + n(j)[\tau], \tag{4.5}$$

where $n(j)[\tau]$ is the additive zero mean white Gaussian noise. Furthermore, consider the input and output relationship of the MIMO system, (4.5) can be extended to

$$\mathbf{r}[\tau] = \tilde{\mathbf{H}}[\tau] * \mathbf{x}[\tau] + \mathbf{n}[\tau] = \sum_{\tau_l=0}^{L-1} \bar{\mathbf{H}}[\tau_l] \cdot \mathbf{x}[\tau - \tau_l] + \mathbf{n}[\tau] \tag{4.6}$$

where $*$ denotes convolution, and $\mathbf{n}[\tau] = [n(1)[\tau], \ldots, n(n_R)[\tau]]^{\mathrm{T}}$.

We set the overall channel coefficients as $\mathbf{h} = \left[\mathbf{h}_{11}^{\mathrm{T}} \ldots \mathbf{h}_{1n_T}^{\mathrm{T}} \middle| \mathbf{h}_{21}^{\mathrm{T}} \ldots \mathbf{h}_{2n_T}^{\mathrm{T}} \middle| \ldots \middle| \mathbf{h}_{n_R1}^{\mathrm{T}} \ \ldots \ \mathbf{h}_{n_Rn_T}^{\mathrm{T}}\right]^{\mathrm{T}}$ where $\mathbf{h}_{ji} = [\, h_{ji}[0], \ldots, h_{ji}[L-1] \,]^{\mathrm{T}}$. Moreover, the training sequence is as-

sumed given and of length L_{ts}, and it shall be greater or equal to the dimensionality of transmit symbol set, $L_{ts} \geq P$. As defined in (3.3), the training sequence symbol belongs to $\{x_1, x_2, \ldots, x_P\}$, and therefore we set $\mathbf{x} = [x_1, x_2, \ldots, x_P]^{\mathrm{T}}$ for convenience, where P is the dimensionality. From the above description, coefficient L_{ts} is greater or equal to P. The received symbol vector is constructed as $\mathbf{r} = [\mathbf{r}[1]^{\mathrm{T}}, \ldots, \mathbf{r}[L_r]^{\mathrm{T}}]^{\mathrm{T}}$ where $L_r = L_{ts} + L_h - 1$ and $\mathbf{r}[\tau]$ is defined in (4.6). Then (4.6) is unfolded to equation:

$$\mathbf{r} = \mathbf{H}\mathbf{x} + \mathbf{n} = \mathbf{X}\mathbf{h} + \mathbf{n} \tag{4.7}$$

where $\mathbf{H}$ is a $L_r n_R \times P$ matrix populated by channel coefficients corresponding to $\mathbf{x}$, $\mathbf{X}$ is a $L_r n_R \times L n_R n_T$ matrix populated by transmit symbols corresponding to $\mathbf{h}$, and $\mathbf{n}$ is the additive white Gaussian noise vector. Matrix $\mathbf{X}$ is obtained by combining the transmit symbols from all antennas, and the antennas do not necessarily transmit the same symbol at a given time index. Appendix E shows the details and examples of composing (4.7).

In principle, we seek to estimate the nonlinear symbols at the transmitter to improve the channel estimation given only the received signal, $\mathbf{r}$, and training sequences from each antenna.

4.2.3 Basic Least Squares (BLS) Channel Estimation

Here, a basic least squares channel estimator is introduced and will be used throughout this chapter. Given the training sequence, the BLS channel estimation of (4.7) can be derived by [68],

$$\hat{\mathbf{h}} = \arg\min_{\mathbf{h}} \|\mathbf{r} - \mathbf{X}\mathbf{h}\|^2 \tag{4.8}$$

where $\mathbf{X}$ is a function of known training sequence. Since Gaussian noise is our premise, the solution is simply $\hat{\mathbf{h}} = \mathbf{X}^{+}\mathbf{r}$, where superscript + denotes the pseudo-inverse operator. Unfortunately the training sequence suffers nonlinearity from the RF front-end, and thus the channel estimation result is degraded.

4.3 Linear Approximation Approach

The difficulty of parameter estimation here is that, all transmit symbols are affected by the nonlinear system and they are associated with each other due to the ISI effects, while

the nonlinearity coefficient estimation of the RF-front end requires the transmit symbol values $\mathbf{x}$. Since both $\mathbf{x}$ and $\mathbf{h}$ in (4.7) are unknown, it is not feasible to use the BLS approach, $\hat{\mathbf{h}} = \mathbf{X}^{+}\mathbf{r}$, for channel estimation. One the other hand, if we can provide sufficiently accurate approximation of some of the elements in $\mathbf{x}$, then this partial training sequence could be used to estimate the channel coefficients, $\mathbf{h}$.

Though vector $\mathbf{x}$ contains nonlinearity properties, the portion of the symbols containing less nonlinearities could be approximated to known symbols. In other words, in multi-amplitude modulation schemes, symbols having lower amplitude levels usually suffer relatively lower degree of nonlinearity compared to those at higher levels. As a result, those lower-amplitude symbols suffering small nonlinearity could be approximated as linear. This premise is based on the fact that the nonlinearity effects are constrained in many communication specifications, because of the power spectrum mask limitation and Relative Constellation Errors (RCE). For instance, Fig 3.1 shows one realization of a RF-front end output constellation of the short training sequence symbols of IEEE 802.11a/g system and some proposed 802.11n draft [22, 69]. In this example, signal $\{x_5, x_6, \ldots, x_9\}$ can be approximated to $\{\check{d}_5, \check{d}_6, \ldots, \check{d}_9\}$, where $\check{d}_i, i = 5 \sim 9$ is defined in (3.3). In this chapter, the nonlinearity model is extracted from [38] and coefficients are adjusted to meet the EVM or RCE requirement in [22]. Since the problem is generalized to Eq. (4.5) which is similar to Eq. (3.6) and (3.7), the linear approximation developed in Chapter 3 Section 3.3 applies.

4.4 Iterative Estimation Approach

The previous approach assumes the symbols in lower amplitudes as linear signals, and thus derives both the initial channel estimation and nonlinear transmit symbol estimation. In this section, we will further remove this approximation and improve the accuracy by an iterative algorithm. The initial nonlinear transmit symbol estimation derived from the linear approximation could be combined with (4.7) to estimate the linear channel coefficients, $\mathbf{h}$, called *h-step*. Then under the premise this newly estimated information is more accurate than the previous estimation, the new data can be utilized to re-estimate the transmit symbols by applying (4.7), called *x-step*. The result can be again sent to *h-step* and then *x-step* for a new iteration or stop to output the results. The iterative approach is based on the assumption that the new estimation supports more

accurate knowledge about the nonlinear channel than the previous one, and therefore improves the next estimation. When this premise holds, the accuracy will increase with the iterative procedure. The solution provided in Chapter 3 Section 3.3 is applied here.

4.5 Simulation Results

In this simulation, a practical example applying training sequences used by IEEE 802.11n standard proposals [69, 70] with RF front-end device models from [38], and a 2×2 indoor MIMO channel model extracted from [51,71,72] using correlation matrices will be shown. Here the indoor MIMO channel model utilizes the correlation matrices from Ped B model in [71] and the delay profile $\delta[\tau]+0.063\delta[\tau-1]+0.01\delta[\tau-2]$, obtained from [51]. The degree of nonlinearity is set by -17.5dB EVM satisfying the specification [22] when operating in 16-QAM scheme. Also, we use a common definition of SNR as the average power at the output of each receive antenna to the noise power ratio as in [55]. The SNR region is set between 15dB and 35dB which is a typical operation environment for 16-QAM scheme [41]. Furthermore, the Normalized Mean Squared Error (NMSE) of parameter vector $\mathbf{h}$ is defined by: NMSE of $\mathbf{h} \equiv 10\log E(|\hat{\mathbf{h}}(k)-\mathbf{h}|^2/|\mathbf{h}|^2)$(dB) .

Our algorithm initially uses the long training sequence of the one packet header from [22, 70] for linear approximation approach and then utilizes the same sequence for the iterative estimation. Here, channel estimation results using linear approximation (LA), LA and iterative estimation (LE), and Basic Least Squares (BLS) algorithms are compared. BLS in our example simply treats the transmitter as linear, and uses the LS algorithm as mentioned in Section 4.2.3. Figure 4.1 and 4.2 show the NMSE of linear channel coefficients, $\mathbf{h}$, and nonlinear symbols, $\mathbf{x}$. Figure 4.1 also indicates that the performance of a BLS estimator is limited by the nonlinearity effect even in high SNR situation, while our proposed method reduces the nonlinearity effect and increases the precision especially in high SNR environment. Moreover, our method not only estimates the wideband channel model, but achieves a precise estimation of transmit symbols as depicted in Fig 4.2.

To show the potential improvement on system performance, a simple equalizer that uses the estimated channel coefficients to reduce the ISI will be utilized. Three different scenarios are presented. The first situation is a linear transmitter, $\mathbf{b} = [1,0,0]^{\mathrm{T}}$, with BLS equalizer, and it provides the performance in absence of transmitter nonlinearities.

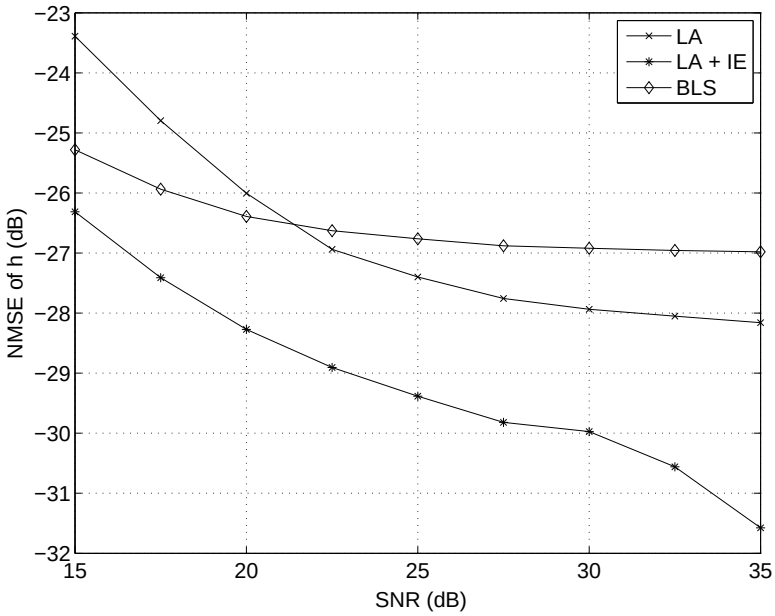


Figure 4.1. NMSE of the estimated channel coefficient $\hat{\mathbf{h}}$ versus SNR.

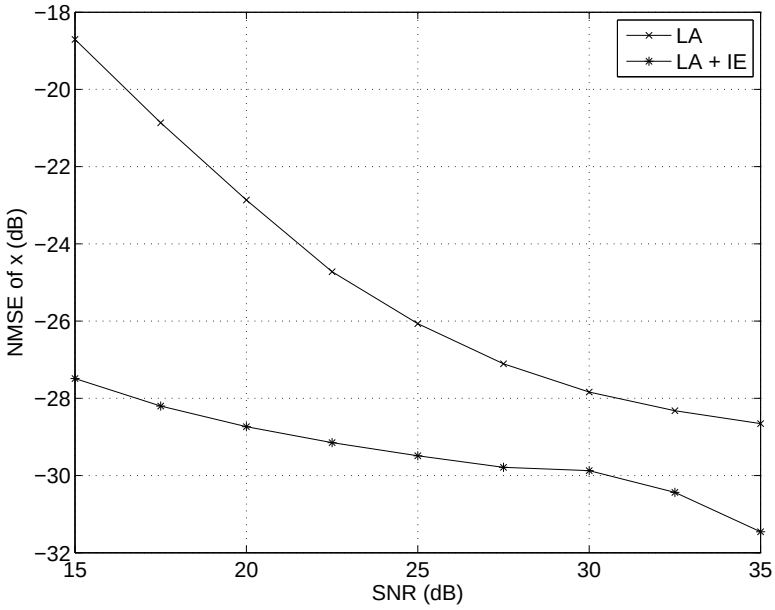


Figure 4.2. NMSE of the estimated nonlinear symbol vector $\check{\mathbf{x}}$ versus SNR.

The second situation is to use only BLS with nonlinear transmitters. The decreased performance, compared to the first situation, shows the effect of transmitter nonlinearities. The third situation is our proposed channel estimation algorithm with a simple nonlinearity compensator when nonlinear transmitter is used. Though the compensation of nonlinearity is not our focus in this chapter, a first order compensation method is applied for a basic comparison. Since the actual linear transmit symbols are unknown, the compensator merely utilizes the estimated transmit symbols of iteration k to construct $\check{\mathbf{D}}$ in (3.3) and estimate the linear symbols, $\hat{\mathbf{d}}$, by inserting the estimated $\hat{\mathbf{b}}$ from (3.3) as $\hat{\mathbf{d}} = \hat{\mathbf{x}}^{(k)} - \check{\mathbf{D}}(\hat{\mathbf{x}}^{(k)})\hat{\mathbf{b}}^{(k)}$, where the elements in $\check{\mathbf{D}}$ are substituted by $\hat{\mathbf{x}}^{(k)}$. In Fig 4.3, Bit Error Rate (BER) is derived for a 16-QAM scheme in a 64-bin FFT OFDM system. The simulation results show that in high SNR situation, using a simple compensation method can improve the performance by several dB. The noise floor in Fig 4.3 is caused by the nonlinearity which can be compared to the linear transmitter simulation result. Therefore, the difference between the noise floor and linear transmitter result provides an improvement margin for the nonlinearity compensator.

Figure 4.4 shows the improvement between BLS and the compensator. The difference between BLS and the compensator is provided, and also this difference is normalized to the BER of BLS to show the ratio of improvement. Here, EVM is introduced to show the degree of nonlinearity. For example, the nonlinearity is lower when EVM is lower. The BER improvement increases as the nonlinearity decreases, while the in the low nonlinearity region this improvement decreases due to the low BER rate in this region. The figure indicates that high nonlinearity degrades the compensator performance, and in region of -12.5 to -15dB the compensator removes the most error. As the figure illustrates, the normalized improvement increases as the nonlinearity decreases, meaning the compensator reduces the higher portion of the bit error in lower nonlinearity scenario.

4.6 Conclusion

In this chapter, we have proposed a channel estimator that estimates the linear MIMO channels and the nonlinearity caused by imperfect transmitter design at the same time. The algorithm first utilizes the differentiation of all symbol levels and assumes some of them are linear to estimate the nonlinearity; further improvement is achieved by itera-

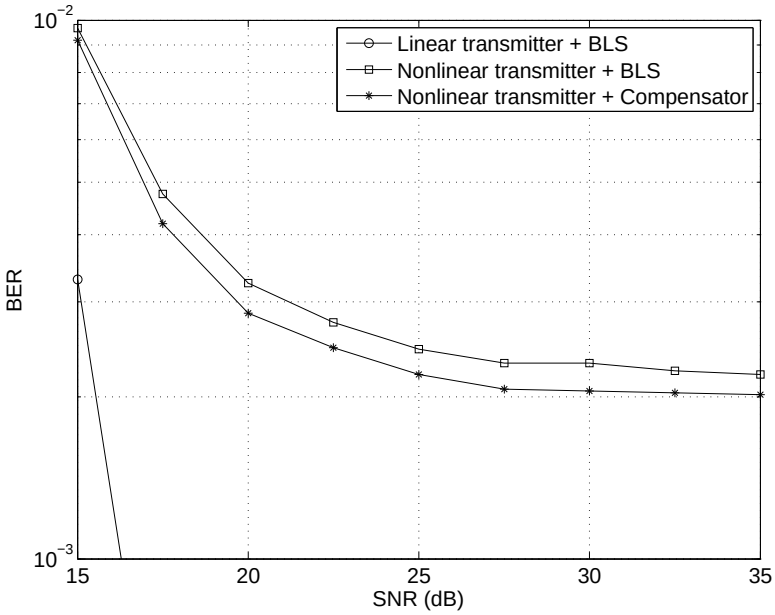


Figure 4.3. BER of a 2 $\times$ 2 MIMO OFDM system using 16-QAM modulation scheme.

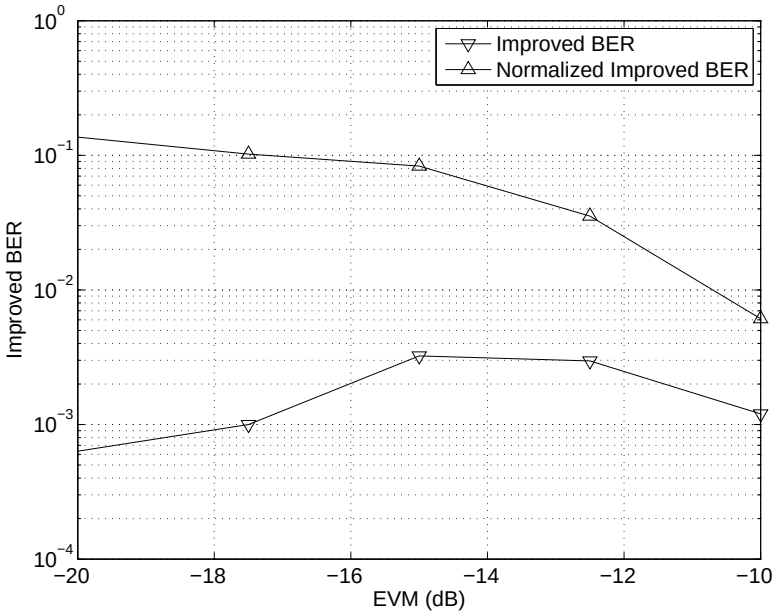


Figure 4.4. Improved BER over different degree of nonlinearity of a 2 $\times$ 2 MIMO OFDM system using 16-QAM modulation scheme. A typical EVM requirement for a 16-QAM OFDM system is about -16 to -19dB.

tive estimation of the transmit symbol values and the channel coefficients. The semi-analytical results of the performance in each iteration were presented. The numerical results show that the nonlinearity is a major limiting factor to channel estimation when using a basic LS estimator, while our approach can both overcome this limitation and improve the system performance.

Chapter 5

Emitter Tracking

5.1 Introduction

Transmitter nonlinearity estimates also provide the information about the change of transmit power, and we utilize this information to estimate the location of the transmitter. Since in most modern communication systems, power control is used to maintain the quality of the connection, it usually requires constant received power on the base station (BS) side. For instance, while the mobile station (MS) is moving away from BS, the MS might increase its transmit power to maintain constant received power on BS side. The transmit power is location dependent, and thus provides transmit power patterns for each possible MS route. Therefore, by estimating the nonlinearities, the location of the mobile station can be estimated. One of the advantages of this method is that the location tracking can be achieved using one single sensor.

5.2 Nonlinearity Model

Recall from (3.1) that the baseband equivalent nonlinearity distortion model is

$$\begin{aligned} z(d) &= a_1 d + a_3 d^2 d^* + \ldots a_M d^{\frac{M+1}{2}} (d^*)^{\frac{M-1}{2}} \\ &= (a_1 \sigma_d) \underbrace{\left[b_1 \check{d} + b_3 \check{d}^2 \check{d}^* + \ldots + b_M \check{d}^{\frac{M+1}{2}} (\check{d}^*)^{\frac{M-1}{2}} \right]}_{x} \end{aligned} \tag{5.1}$$

where $\breve{d} = d/\sigma_d$,

$$b_j = \frac{a_j}{a_1}\sigma_d^{j-1},\ j \in \{1,\ 3, \ldots,\ M\}\,. \tag{5.2}$$

Our goal is to use the nonlinearities for SEI and thus only the nonlinearity terms $b_j,\ j = 1,\ 3, \ldots,\ M$ are considered in the identification procedure and the linear gain a_1 is ignored. Moreover, (5.2) shows that the normalized nonlinearity coefficients are proportional to the input signal power. As a result, as the transmit power increases (decreases), the estimate values should increase (decreases) accordingly.

The objective of SEI is to identify transmitters based on the estimation of the nonlinear transmitter coefficients in the multipath environment. Throughout this chapter, we assume that the transmit sequence is known, the transmitter nonlinearity is static, and the channel is time invariant during the estimation.

5.2.1 Channel Model

In this chapter, we assume a coherent, symbol-spaced receiver front-end and precise knowledge of the symbol timing, such that the channel can be represented by an equivalent, discrete-time, baseband linear model [46]:

$$r[i] = \sum_{k=-\infty}^{\infty} z[k]h'[i-k] + n[i] \tag{5.3}$$

where $z[\cdot]$ is a sampled transmit symbol based on (5.2), $h'[\cdot]$ is the discrete-time channel impulse response, and and $n[\cdot]$ is zero mean Gaussian i.i.d. complex noise sample. Since we are interested in the nonlinearity of the system, the above equation is revised as

$$r[i] = \sum_{k=-\infty}^{\infty} x[k]h[i-k] + n[i] \tag{5.4}$$

where $x[\cdot] = z[\cdot]/(a_1\sigma_d) \in \{x_1,\ x_2, \ldots,\ x_P\}$, and $h[\cdot] = a_1\sigma_d h'[\cdot]$.

5.3 Location Tracking

Radio emitter identification and location tracking has various applications and extensive researches have been conducted in this area. In particular, specific emitter identification (SEI) systems discern wireless radio emitters of interest based only on the external signal feature measurement. This identification capability leads to diverse applications such as cognitive radio [1], network intrusion detection, battlefield management, electronic support measurement system, and intelligence gathering [2–6]. Our previous research uses the transmitter nonlinearity estimation for SEI [23,26–28] and in this chapter, we extend this approach for a mobile station location tracking method that enables joint SEI and location tracking. The major advantage of this algorithm is that the radio emitter identification and tracking are completed simultaneously on the same signal feature; therefore, the ambiguity of signals from different emitters are removed. Other advantages are that only one sensor is required and multipath mitigation capability.

There are various approaches to mobile station direction finding such as amplitude comparison, phase comparison, time-of-arrival, difference-time-of-arrival, and frequency-difference-of-arrival [73]. Locating the target mobile station usually requires another degree in spatial diversity which could be achieved by multiple observation sensors. In this chapter, we develop a single-sensor location finding technology utilizing the estimation of the transmitter nonlinearities cooperating with a pre-measured transmit power map.

The signal nonlinearities vary with changes of transmit power, and we connect these phenomena with location of the transmitter. Since in most modern wireless communication systems, power control is used to maintain the quality of the connection, resulting in constant receive power on base station (BS) side [34]. For instance, while the mobile station (MS) is moving away from the BS, the MS might increase its transmit power to maintain constant receive power on BS side. Moreover, the channel path-loss is also associated with distance, terrain features, and obstacles between the transmitter and the receiver. As a result, characteristics about the surrounding environment can be found in the propagation channel [74–81], and to compensate the channel, the power control also reflects this aspect [82]. As the mobile station moves along a particular path, its transmit power variation along the path contributes a location dependent pattern. To estimate the variation of transmit power, we utilize the transmitter nonlinearities estimation using

the fact that the increase or decay of normalized transmitter nonlinearities is associated with the transmit power. Therefore, by estimating the nonlinearities, the location of the mobile station can be estimated.

5.3.1 Estimator Properties

The performance of this iterative estimator is discussed in detail in [28], and we present some important conclusions in this section. At kth iteration, the nonlinearity estimation error is represented as (Appendix F.1)

$$\hat{\mathbf{b}}^{(k)} - \mathbf{b} = \mathbf{Q}_b^{(k)} \mathbf{e}_x^{(0)} + \mathbf{W}_b^{(k)} \mathbf{n}^{(0)} \tag{5.5}$$

where $\mathbf{Q}_b^{(k)}$ and $\mathbf{W}_b^{(k)}$ are the error transform matrix, $\mathbf{e}_x^{(0)}$ is the initial estimation error, and $\mathbf{n}^{(0)}$ is additive zero mean Gaussian noise. Two important remarks on (5.5) are (see details in Appendix D):

Remark 1 *If $\rho(\mathbf{R}) < 1$, then $\mathbf{Q}_b^{(k)} \mathbf{e}_x^{(0)}$ in (5.5) converges to a zero vector.*

Remark 2 *If $\rho(\mathbf{R}) < 1$, then $\mathbf{W}_b^{(k)} \mathbf{n}$ in (5.5) converges to a fixed vector.*

Here, $\mathbf{R} = \mathbf{P}_{\check{D}} \mathbf{R}_h \mathbf{R}_x$, $\rho(\mathbf{R}) = max\{|\lambda_1|, \ldots, |\lambda_r|\}$ is the maximum eigenvalues of $\mathbf{R}$. As a result, the first term on RHS of (5.5) associated the initial estimation error which asymptotically converges to zero, and the second term on the RHS contains Gaussian noise.

In the rest of the chapter, we consider a sufficient number of iterations such that $\mathbf{Q}_b^{(\infty)} \rightarrow \mathbf{0}$ and only the Gaussian noise term is left. $\mathbf{W}_b^{(\infty)}$ is associated with the noise floor, and it asymptotically converges to a fixed matrix.

5.3.2 Effects on Transmit Power Variation

In modern communication systems, power control is a commonly used technology; therefore of that it is crucial to study the effects of transmit power variation on the nonlinearity estimator. According to Remarks mentioned the nonlinearity estimates are zero mean Gaussian variables after sufficient iterations, and therefore we discuss the expected value and variance as following.

The expected value of the estimates also varies with the variation of transmit power based on (5.2) (see Appendix F.1). For instance, as the transmit power increases (decreases), the expected values of the transmitter nonlinearities also increase (decrease) accordingly. This result is concluded as

$$\frac{E[\hat{b}_j(i)]}{E[\hat{b}_j(i-1)]} = \frac{\sigma_d^{j-1}(i)}{\sigma_d^{j-1}(i-1)},\ j \in \{1,\ 3, \ldots,\ M\} \tag{5.6}$$

where i is the sample index. As a result, given a set of nonlinearity coefficients, the identifier is able to expect the possible estimates variation.

As far as the variance, using the definition in (5.4) the change of transmit power also affects the amplitude of channel coefficients $\mathbf{h} = \sigma_d \mathbf{h}'$ as in (5.4); hence the estimator variance is also proportional to the transmit power

$$\sigma_{\hat{\mathbf{b}}}^2 \sim \sigma_d^{-2} \tag{5.7}$$

where the definition of estimator variance is

$$\sigma_{\hat{\mathbf{b}}}^2 = E\left\{ \left\| \hat{\mathbf{b}} - \mathbf{b} \right\|^2 \right\} = \sum_{i=1}^{\frac{M-1}{2}} E[|\hat{b}_{2i+1} - b_{2i+1}|^2]. \tag{5.8}$$

This result indicates that when the transmit power increases, the accuracy of the estimation improves (see Appendix F.1).

Fig. 5.1 shows the simulation results of NMSE versus different transmit power, and as the transmit power increases, NMSE decreases as expected. In this simulation, eight training sequences from 32 packet headers of an OFDM system IEEE 802.11g [22] with 20 dB E_b/N_0 are used, the linear gain of the amplifier is set $a_1 = 1$, and the normalized mean squared error (MSE) is defined as

$$\text{NMSE of } \mathbf{b} \equiv 10\log\left(E\left[\left\| \hat{\mathbf{b}}^{(k)} - \mathbf{b} \right\|^2 \right] \right) (\text{dB})\ . \tag{5.9}$$

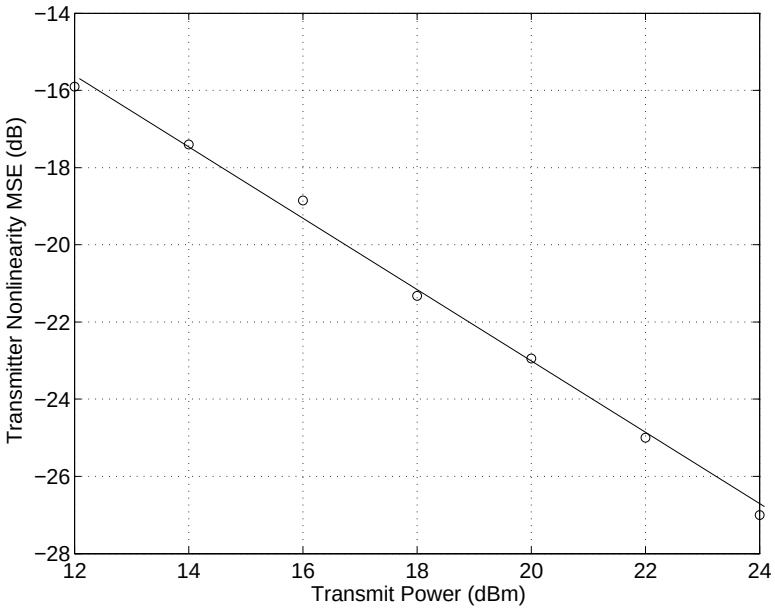


Figure 5.1. Normalized MSE of nonlinearity estimates as ∘ versus transmit power.

5.4 Location Tracking

The transmit power of MS is location dependent as described in Section 1, and each MS route contains a transmit power pattern. In general, MS adjusts its transmit power to maintain constant BS receive power to support consistent communication quality. If MS is moving away from BS, it raises its power to compensate for the increasing path loss, and similarly, while MS is approaching BS, its transmit power is decreased [82]. Then the MS locator compares the power variation estimate with candidate patterns of all routes, and find the most likely one. One advantage of this method is that the location tracking can be achieved using one single sensor. Assuming a moving MS with constant speed is observed, the variation of its transmit power estimates is defined as a series of transmit power ratio between two adjacent observations as

$$\triangle \mathbf{P} = [\triangle P(i), \triangle P(i+1), \ldots, \triangle P(i+N)]^{\mathrm{T}} \tag{5.10}$$

where $\triangle P(i) = \sigma_d^2(i)/\sigma_d^2(i-1)$ is the ratio of transmit power between ith and $(i-1)$th observation. For each route this vehicle takes, the observed transmit power ratio should be different from others. By comparing $\triangle \mathbf{P}$ with pre-measured map data, the location of the MS is then estimated.

5.4.1 Power Control

The simplest form of power control is to adjust the powers so that the received signal is at a constant level. A special case is when the power in the next power control interval can be increased or decreased by a fixed power control step size. For example, updated European telecommunication standard institute (ETSI) standard specifies 15 levels of power control for global system for mobile communication (GSM) 900 handset from 33 to 5 dBm with step size of 2 dB [82]. The emitter power can be as low as 5 dBm when it is close to the BS, while the maximum power is achieved when it is far away. Another example in WiMAX IEEE 802.16 standard is that the transmitter is required to provide a dynamic power control range of at least 50 dB in no less than -1 dB steps [83].

To obtain an estimate of $\triangle P(i)$, transmitter nonlinearity estimation is utilized, because the nonlinearities vary due to the change of transmit power. Based on (5.2), the expected value of the estimates are in proportion to the transmit signal variance σ_d^2 as (5.6). In another words, as the transmit power increases, the expected value of nonlinearity estimates also increase. Therefore, by recording the variation of the nonlinearity estimates, the changes of transmit power is also observed.

Therefore, the transmit power ratio of two adjacent observations is derived using the ratio of two estimated third-order nonlinearity estimates as

$$\triangle \hat{P}(i) = \frac{\hat{b}_3(i)}{\hat{b}_3(i-1)} \tag{5.11}$$

where $\hat{b}_3(i)$ is the ith estimate using the iterative nonlinearity estimator.

5.4.2 Location Tracking Estimator Properties

Based on (5.5), the nonlinearity estimates are Gaussian distributed, and therefore the transmit power variation in (5.11) is then the ratio of two Gaussian random variables. The ratio of two Gaussian random variables is a Cauchy-like distribution [84, 85]. In some cases, the probability density function (PDF) of this ratio is approximately a Gaussian distribution based on the coefficients of variation of the numerator and denominator [86]. This conclusion is modified to apply to the transmit power estimator in (5.11) as shown in Fig. 5.2(a) where the critical values of K.-S. Lilliefors Test [87] is presented. K.-S. Lilliefors Test is used to test the normality of data without specifying the expected

value and variance [87]. Fig. 5.2(a) shows the critical value of the test results using the ratio of two Gaussian distribution variables. The expected values of both the Gaussian variables are set to 0.04. Fig. 5.2(a), if the value is less than 0.886, then the data is accepted as normal distribution with 95% confidence interval. Fig. 5.2(a) suggests that when the NMSE of the estimation is less than $-14dB$, the estimates are normally distributed with expected value as [86]

$$E[\triangle\hat{P}] \approx \frac{\sigma_d^2(i)}{\sigma_d^2(i-1)} + Var[\hat{b_3}(i-1)]\frac{E[\hat{b}_3(i)]}{E[\hat{b}_3(i-1)]^3} \tag{5.12}$$

and variance as

$$Var[\triangle\hat{P}] \approx Var[\hat{b_3}(i-1)]\frac{E[\hat{b}_3(i)]^2}{E[\hat{b}_3(i-1)]^4} + \frac{Var[\hat{b_3}(i)]}{E[\hat{b_3}(i-1)]^2}. \tag{5.13}$$

Two histograms are also presented in Fig. 5.2(b) based on two scenarios in 5.2(a): acceptance and rejection region. Both numerical results are derived using Gaussian distributed variables with unit mean and ten thousand samples. The distribution derived from the acceptance region in Fig. 5.2(a) shows a Gaussian-like distribution while a Cauchy-like distribution is obtained in the rejection region.

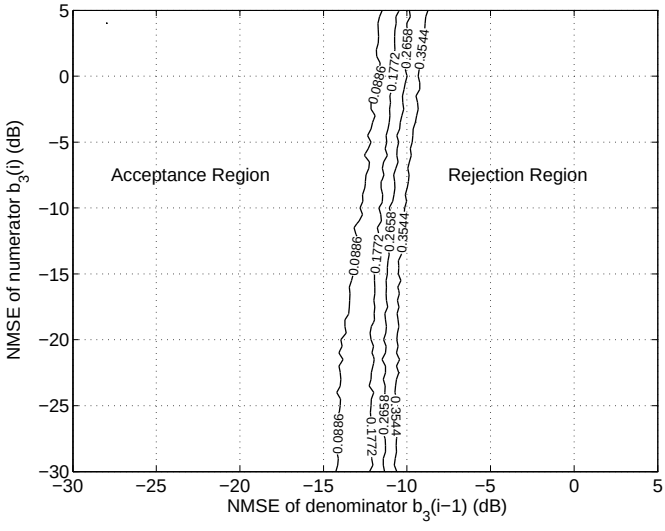
5.4.3 Effects on Channel Gain

To clarify the relationship between channel gain and the estimation result, a simple example using a 1-tap channel model is presented. In this example, it can be shown that the estimation variance defined in (5.8) is inversely proportional to the amplitude of channel as (see Appendix F.2)

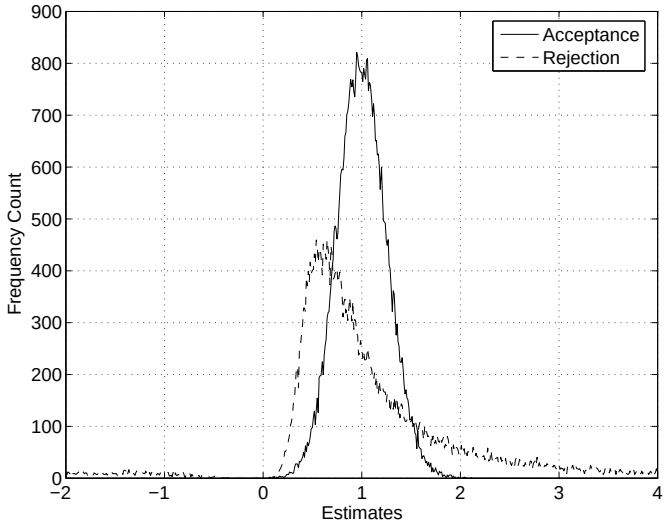
$$\sigma_{\hat{\mathbf{b}}}^2 \sim |h|^{-2}. \tag{5.14}$$

Therefore, as the channel gain increases, the estimation variance decreases. Considering the results from (5.7) and (5.14) as the estimation variance decreases, the bias term in the expected value in (5.12) shrinks, and similarly, the estimation variance in (5.13) also reduces.

Fig. 5.3 shows the simulation results of NMSE versus different channel gains as (5.14). In this simulation, eight packet headers are used and linear gain of the amplifier is set $a_1 = 1$.



(a) Region of acceptance of normality. The Monte Carlo simulation is conducted using 100 realizations with NMSE step size of 0.5 dB, and each realization includes 10000 data. The correlation coefficient of the numerator and denominator is zero.



(b) Histogram of 50000 estimates from two regions in Fig. 5.2(a). One setup is denominator -22.5 dB NMSE and numerator -12.5 dB, while the other setup is denominator NMSE -2.5 dB and numerator -12.5 dB.

Figure 5.2. Region of acceptance of normality and histogram examples on acceptance and rejection regions.

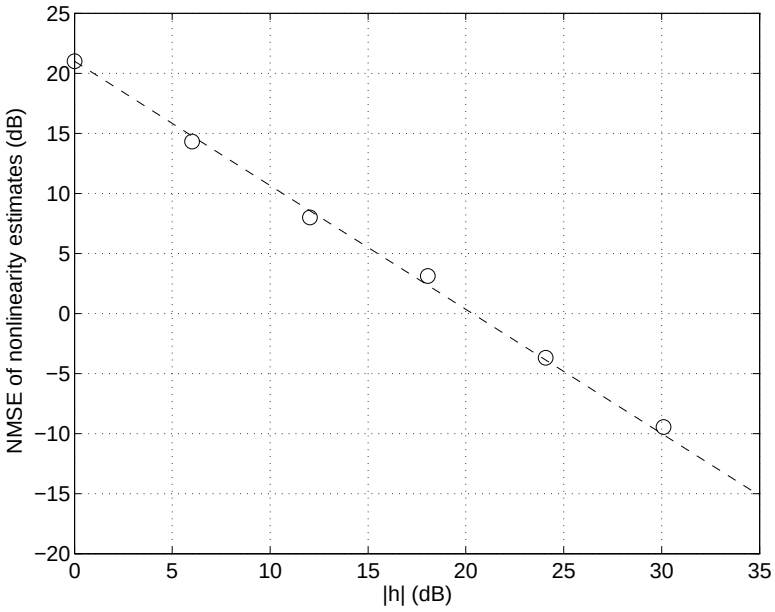


Figure 5.3. Normalized MSE of nonlinearity estimates as ∘ versus channel gain.

5.5 Numerical Results

In the simulations, an OFDM system, IEEE 802.11a/g Specification [22], using RF front-end device models from [38], which were used in our previous research [23, 26] is shown. We apply a one tap FIR channel model whose amplitude is a function of distance between MS and the receiver [75]. The maximum degree of nonlinearity is -16 dB EVM satisfying the specification [22] when operating in 16-QAM scheme. The maximum output power is set to 30 dBm for each emitter, which is a typical maximum transmit power value in commercial market products, and the nonlinearity coefficients are $b_3 = -0.0474 + 0.0823i$, $b_5 = 0.0678 + 0.0078i$, and $b_7 = 0.0022 + 0.0004i$. The simulation shows the misclassification rate improves as the number of headers used increases.

Our algorithm initially estimates transmit symbols using the LA approach [26] based on the long training sequence from [22] of length 160 samples containing 64 different transmit symbols, where the nonlinearities of the 52 lowest amplitude symbols are considered insignificant and those symbols are approximated by linear symbols in the LA method. After the initial estimation, the same training sequence is utilized for the iterative estimation. Based on this simulation environment, we can assign parameters as $M = 7$, $L_h = 3$, $L_{ts} = 160$, $P = 64$, and $H = 64 - 52 = 12$.

The contours in Fig. 5.4 show the power control contour in a finite grid-based city map, and the blocks in the figure indicate buildings. Two possible routes are considered in this simulation. The speed of MS is considered a known constant, and we show the transmit power of each route based on the power control in Fig. 5.5. As can be seen in Fig. 5.5 (a), the transmit power is reduced when MS is approaching BS, while in Fig. 5.5 (b) the zig-zag route shows a different pattern. The estimation error variances are shown in error bars in Fig. 5.5 (a) and (b). The variance at the beginning 10 meters of Route 1 is higher than that from 10 to 20 meters because the MS is further to the sensor at the beginning. This result is consistent to (5.14).

The accuracy at the beginning of Route 1 is higher than later on because the sensor is close to the MS for those measurements and therefore the channel gain is high. Thus based on (5.14), the estimation error variance is lower. Moreover, the transmit power is also higher at those measurements providing better signal strength as shown in (5.7).

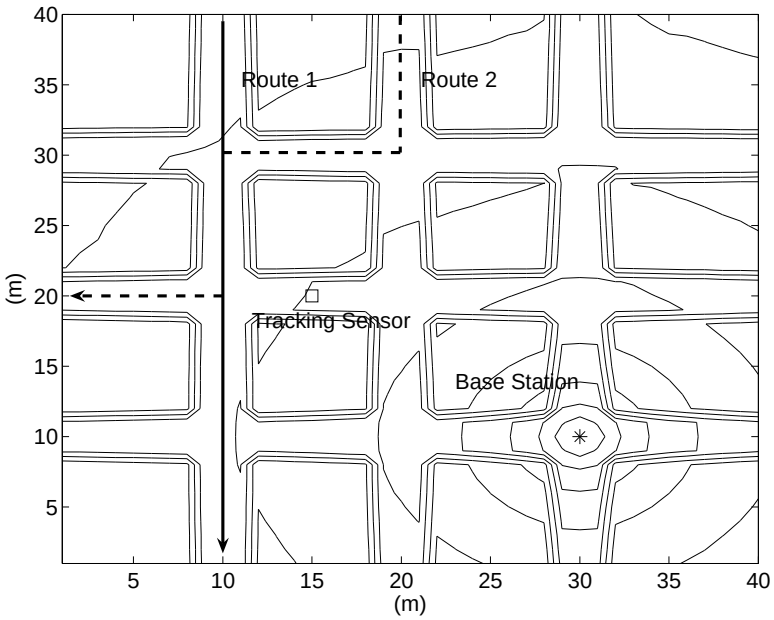


Figure 5.4. The $*$ in the figure shows the location of single sensor scenario.

Based on simulations shown above, the misclassification rate of two routes is calculated using different number of headers as shown in Fig. 5.6. In our simulation, the thermal noise variance of each sensor is assumed identical. MS is measured in eight of the spots uniformly distributed on both Route 1 and 2, where 16, 64, and 256 headers are used with minimum E_b/N_0 10 dB [22]. The sensors are placed in positions that

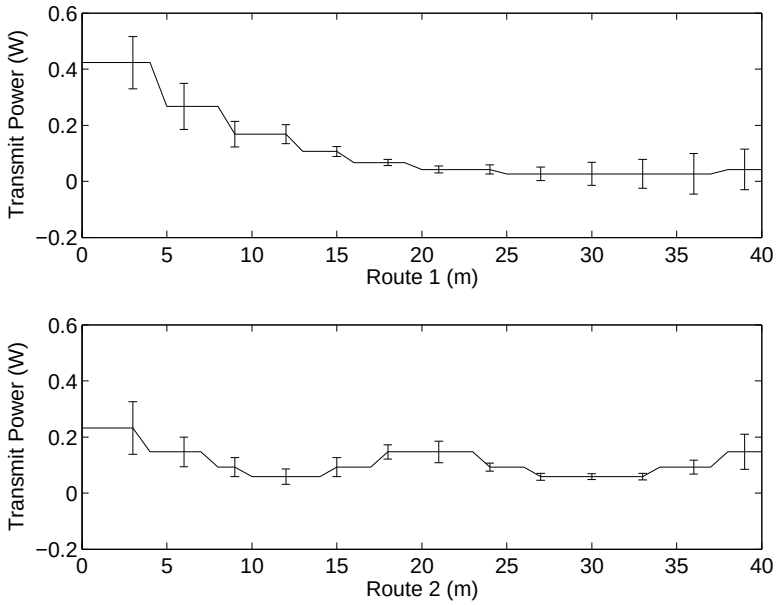


Figure 5.5. Transmit power on Route 1 and 2. The error bars show the variance of transmitter nonlinearity estimates.

minimize the expected value of nonlinearity estimation error variance. The simulation shows that to achieve 30% misclassification rate, headers with one sensor are required. To obtain 5% misclassification rate, more than 256 headers are required.

5.6 Conclusion

In this chapter, we have proposed a transmitter nonlinearity estimator for joint SEI and tracking systems. Also we have provided an unbiased linear minimum variance estimation fusion rule which outperforms selection combining estimation fusion. We discussed the impact of channel amplitude and showed that the estimation error variance is inversely proportional to the square of the channel amplitude. The effect of transmit power variation on nonlinearity estimation is also presented, and the result shows that the expected values of nonlinearity estimates are proportional to the transmit power. Since the transmit power is location dependent, its variation is compared to transmitter power profile of possible MS routes to identify the location of MS. Simulation results have shown that the proposed method provides less than 5% misclassification rate given two candidate routes when 32 headers in E_b/N_0 10 dB environment.

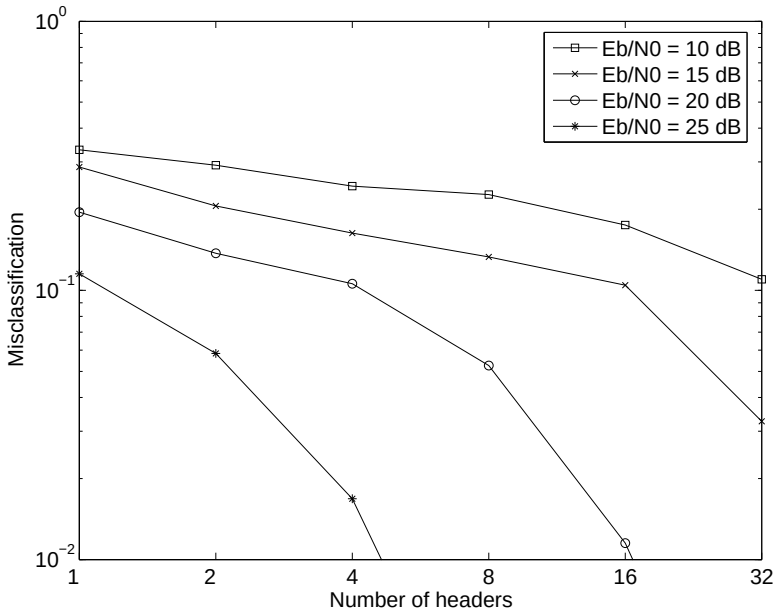


Figure 5.6. Misclassification rate between Route 1 and 2 in Fig. 5.4.

5.7 Centralized Fusion

In this section, multiple-sensor scenario is considered. The assumption is that all distributed sensors receive signals from one emitter and send the receive information back to a data fusion center (DFC). DFC receives data from its sensors, and after data rearrangement, it processes the SEI algorithm by iteratively estimating the channel coefficients and the transmit symbols. After sufficient number of iterations, the estimator utilizes the distorted transmit symbol constellation to derive the transmitter nonlinearities. In this section, we introduce the signal models which are later used for the iterative estimation.

We first derive the equation for channel coefficient estimation as follows. The receive data $\mathbf{r}_k$ could be rearranged as $\mathbf{r} = [\mathbf{r}_1^{\mathrm{T}}\ \mathbf{r}_2^{\mathrm{T}} \ldots \mathbf{r}_{N_s}^{\mathrm{T}}]^{\mathrm{T}}$, and $\mathbf{h} = \left[\mathbf{h}_1^{\mathrm{T}}, \mathbf{h}_2^{\mathrm{T}}, \ldots, \mathbf{h}_{N_s}^{\mathrm{T}}\right]^{\mathrm{T}}$ is the channel coefficient vector, $\mathbf{n}$ is zero mean additive white Gaussian noise. (3.6) can be modified to $\mathbf{X}\mathbf{h}_i + \mathbf{n}_i = \mathbf{r}_i$ to estimate $\mathbf{h}_i$, $\forall i$ individually, where $\mathbf{X}_i$ is composed by the estimated distorted transmit symbols from the ith sensor.

After the channel coefficients are derived from Eq. (3.6), this information is used to obtain the next estimation. Lets consider the multiple sensor case; Eq. (3.6) is further

rearranged as

$$\begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \\ \vdots \\ \mathbf{H}_{N_s} \end{bmatrix} \mathbf{x} + \begin{bmatrix} \mathbf{n}_1 \\ \mathbf{n}_2 \\ \vdots \\ \mathbf{n}_{N_s} \end{bmatrix} = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_{N_s} \end{bmatrix} \tag{5.15}$$

where $\mathbf{H}_k$ is the channel matrix, $\mathbf{x} = [x_1, \ldots, x_P]^{\mathrm{T}}$ is defined in Eq. (3.3), and $\mathbf{n}_k$ is the additive noise of sensor k. For instance, if $\mathbf{x}$ is a 4×1 vector, and the first few transmit symbols are $\{x[1], x[2], x[3], \ldots, x[L_t]\} = \{x_2, x_3, x_2, \ldots, x_1\}$. Then the channel matrix of sensor k is

$$\mathbf{H}_1 = \begin{bmatrix} 0 & h_{11} & 0 & 0 \\ 0 & h_{12} & h_{11} & 0 \\ 0 & h_{13}+h_{11} & h_{12} & 0 \\ \vdots & \vdots & \vdots & \vdots \\ h_{13} & 0 & 0 & 0 \end{bmatrix}. \tag{5.16}$$

A compact format for Eq. (5.15) is similar to (3.6) as $\mathbf{Hx} + \mathbf{n} = \mathbf{r}$.

5.7.1 Variable Step-Size LMS Algorithms

One major issue in this thesis is system identification, or solving a system of linear equations of the general form

$$\mathbf{Uw}^* + \mathbf{n} = \mathbf{d}. \tag{5.17}$$

A major concern is the complexity of solving Eq. (5.17), which depends on the size of the matrices. When the number of sensors increases, the size of matrices in Eq. (3.6) and (3.7) also rises. In this section, we introduce an adaptive filter using LMS algorithm to solve systems of linear equations for its low computational complexity.

Athough other approaches such as pseudo-inverse or total least square (TLS) [49,50] are also efficient, the computational complexity is prohibitive when the size of the matrices grows. The SEI algorithm has to calculate the inverse of huge matrices if pseudo-inverse is considered or the singular value decomposition (SVD) for TLS algorithm. Instead of computing the whole matrix, the family of LMS algorithms sequentially operates on one input vector at a time and requires less computational complexity and it is also capable to output intermediate results if necessary.

Table 5.1. LS Algorithm Complexity Comparison

LS Algorithm	Flop Count
Normal Equation	$mn^2 + n^3/3$
Householder Orthogonalization	$2mn^2 - 2n^3/3$
Modified Gram Schmidt	$2mn^2$
Given Orthogonalization	$3mn^2 - n^3$
Householder Bidiagonalization	$4mn^2 - 4n^3/2$
Golub-Reinsch SVD	$4mn^2 + 8n^3$
R-SVD	$2mn^2 + 11n^3$

The adaptive filtering is to try to adjust a set of filter weights so that the system output tracks a set of desired signals. Let the input vector to the system be denoted by $\mathbf{u}_k$, the filter weight vector is $\mathbf{w}(k)$ and the desired signal is a scalar d_k. The process is assumed to relate to

$$d_k = \mathbf{w}(k)^\dagger \mathbf{u}_k + e(k),\ \forall k \tag{5.18}$$

where $e(k)$ is the error between the adaptive filter output and the desired signal. The LMS algorithm sequentially updates the estimate $\mathbf{w}$ based on the estimation result of each input vector. The weight update recursion is given by

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu_k \mathbf{u}_k e(k)^*,\ \forall k \tag{5.19}$$

where $e(k) = d_j - \mathbf{w}(k)^\dagger \mathbf{u}_k$ is the disturbance mentioned above and $\mu(k)$ is the step size. The determination of step size is somewhat of an art, and a variety of methods have been developed to select the proper step sizes [44, 88–90].

To solve the system of equations in Algorithm 1 , the adaptive filter uses the row vector of $\mathbf{U}$ as input vector, and the desired signal is the corresponding element in $\mathbf{d}$. After the algorithm processes from the first row to the last row, it starts from the first row again until sufficient number of iterations are completed. This algorithm is summarized in Algorithm 2 where d_j is the jth element of $\mathbf{d}$ and $\mathbf{u}_j$ is the jth row of $\mathbf{U}$ in Eq. (5.17).

To further explain this process, we show a special case in Fig. 5.7 assuming a two-row matrix $\mathbf{U} = [\mathbf{u}_1\ \mathbf{u}_2]^{\mathrm{T}}$, a two-element vector $\mathbf{d} = [d_1\ d_2]^{\mathrm{T}}$, and $\mu(k) = 1,\ \forall k$. If prior knowledge about $\mathbf{w}$ is available, use that knowledge to select an appropriate initial value

Algorithm 2 Variable Step Size LMS Algorithm

1: *Initialization:* Set N as the number of rows in $\mathbf{U}$, $\mathbf{w}(0)$ and step size $\tilde{\mu}$
2: *Computation:*

$$\begin{aligned}
&\text{For } i = 0,\ 1, \ldots, \text{ compute} \\
&\text{For } j = 1,\ 2, \ldots,\ N \\
&e(iN+j) = d_j - \mathbf{w}^{\dagger}(iN+j-1)\mathbf{u}_j \\
&\mathbf{w}(iN+j) = \mathbf{w}(iN+j-1) + \mu(j)\mathbf{u}_j e^{*}(iN+j)
\end{aligned}$$

3: *Output:* The last estimation $\mathbf{w}$

for $\mathbf{w}(0)$. Otherwise, set $\mathbf{w}(0) = \mathbf{0}$. Then add a new vector to $\mathbf{w}(0)$ as an updated estimate $\mathbf{w}(1)$ that moves towards the estimation that best fits the first row of the observed data $\mathbf{u}_1$ and d_1.Then the algorithm exploits $\mathbf{u}_2$ and d_2 for $\mathbf{w}(2)$ update, and then the first row to update $\mathbf{w}(3)$. The algorithm sequentially applies the same principal to modify the estimation based on the following rows until the last row, and it repeats the same procedure from the first row using the estimate from the last row if required. Iteratively, this algorithm approaches the optimal estimation $\mathbf{w}_{opt}$.

The selection of step size is a state of art, and some discussion can be found in [44, 88–90]. In [88], the step size is updated by sign changes of successive samples of the gradient; [89] proposes the use of squared instantaneous errors to adapt the step-size of the algorithm to achieve faster tracking; [90] updates the filter through the squared autocorrelation of errors at adjacent time. Due to the fact that data might have very different power levels or noise levels from multiple sensors, variable step sizes can be utilized to fuse the heterogeneous data.

5.7.2 Normalized LMS Algorithm

Since the receive signal power of each sensor might vary substantially, the deviation of the filter update using LMS algorithm could be large. The classic normalized LMS (NLMS) algorithm is a modified LMS algorithm which minimizes the squared Euclidean norm of the filter parameter change:

$$\begin{aligned}
&\text{minimize} && \|\mathbf{w}(i+1) - \mathbf{w}(i)\|^2, \\
&\text{subject to} && d_j - \mathbf{w}^{H}(i+1)\mathbf{u}_j = 0
\end{aligned} \tag{5.20}$$

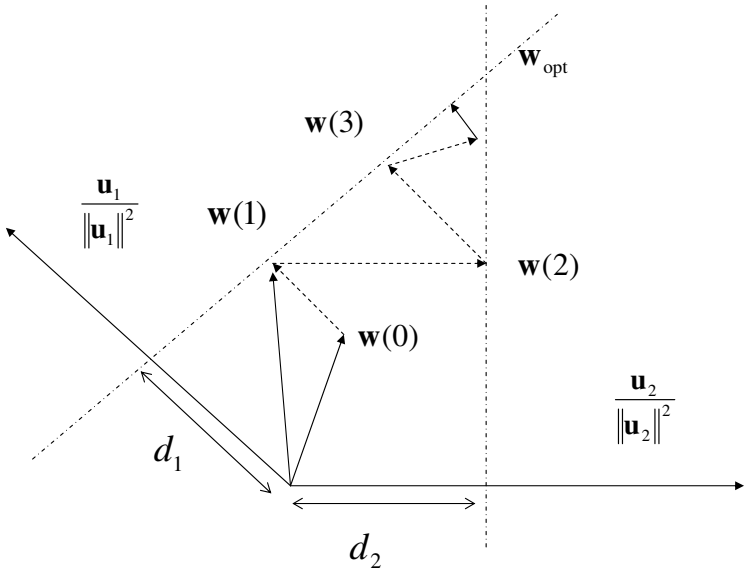


Figure 5.7. An example of data reusing algorithm.

where i is the current number of iteration and j is the current observation data used [44,91]. Besides, the NLMS algorithm also provides potentially-faster convergence rate for both correlated and whitened input data [91,92].

The time variant step size is then derived as

$$\mu(i) = \frac{\tilde{\mu}}{\|\mathbf{u}_i\|^2} \tag{5.21}$$

where $\tilde{\mu}$ is a user-specified convergence parameter and e_i is the estimation error [44].

5.7.3 Iterative Estimation

Based on Eq. (3.3), the transmitter nonlinearities are extracted from the distorted symbol constellation; hence we derive a robust iterative nonlinearity estimator which removes the ISI caused by the channel effects [26]. The Iterative Estimation (IE) algorithm first derives the channel coefficients as *h-step*, and then uses this result to obtain the transmit symbols under the constrain of the nonlinearity model as *x-step* and *i-step* step. This transmit symbol estimation is utilized to derive the transmit symbol matrix in *h-step* and begins the new iteration procedure. The iterative estimation is based on the assumption

that the new estimation provides a more accurate result, and it could furthermore improve the accuracy of the next estimation. Under this premise, the accuracy increases with the iterative procedure.

5.7.3.1 Initialization

In the initial step, set the iteration count $k = 0$. If prior knowledge about the transmit symbols is available, use this information to select initial values, otherwise set $\check{\mathbf{x}}^{(0)} = \check{\mathbf{d}}$ and $\hat{\mathbf{h}}^{(0)} = \mathbf{0}$. Since the values of transmit symbols are affected by the transmitter nonlinearities, the range of these values are specified by the specification such as Error Vector Measurement (EVM) in [22]. The requirement of the degree of nonlinearities is common in most of the commercial specifications in order to control the transmission quality. The initial value $\hat{\mathbf{h}}^{(0)} = \mathbf{0}$ is then used as the initial value of the filter coefficients.

5.7.3.2 h-step

In this kth iteration, the algorithm uses $\check{\mathbf{x}}^{(k-1)}$ to compose transmit symbol matrix $\mathbf{X}^{(k)} = diag(\tilde{\mathbf{X}}_1^{(k)}, \tilde{\mathbf{X}}_2^{(k)}, \ldots, \tilde{\mathbf{X}}_{N_s}^{(k)})$ in Eq. (3.6). In other words, the elements of $\check{\mathbf{x}}^{(k-1)}$ are used to compose $\tilde{\mathbf{X}}_i^{(k)}$, $i = 1, \ldots, N_s$. For the first iteration, the value of $\check{\mathbf{x}}^{(0)}$ comes from the initialization setup. The solution of this new system of equation is derived using the variable step size LMS filter mentioned in Algorithm 2. The new estimation is denoted as $\hat{\mathbf{h}}^{(k)}$ and we simplify this operation as

$$\hat{\mathbf{h}}^{(k)} = LMS\left(\mathbf{X}^{(k-1)}, \mathbf{r}, \hat{\mathbf{h}}^{(k-1)*}\right)^* \tag{5.22}$$

where $LMS(\cdot)$ denotes the LMS algorithm, $\mathbf{X}^{(k-1)}$ is the input matrix, $\mathbf{r}$ is the desired signal vector, and $\hat{\mathbf{h}}^{(k-1)*}$ is the initial value for the filter coefficients. We should notice that the conjugate operator $(\cdot)^*$ is involved in this equation due to the difference between the architecture of LMS algorithm and Eq. (3.6). The approach used to derive the step size is based on NLMS or WLMS mentioned above.

5.7.3.3 x-step

In this step, the transmit symbols are derived based on the knowledge of *h-step*. Similar to $h - step$, the algorithm utilizes $\hat{\mathbf{h}}^{(k)}$ to compose the channel coefficient matrix

$\mathbf{H}^{(k)}$ where the index k denotes the number of iteration. Therefore, the transmit symbol estimator could be written in a simple form as

$$\hat{\mathbf{x}}^{(k)} = LMS\left(\mathbf{H}^{(k)},\, \mathbf{r},\, \check{\mathbf{x}}^{(k-1)*}\right)^{*}. \tag{5.23}$$

This estimator is not constrained by the nonlinearity model in Eq. (3.3), and it is further refined by the next step.

5.7.3.4 i-step

In this step, the algorithm can either output the nonlinearity estimation result or refine the estimation for the next iteration. To derive the nonlinearity coefficients, the estimation result from $x-step$ is sent to Eq. (3.3) and could be written as

$$\hat{\mathbf{b}}^{(k)} = LMS\left(\mathbf{D},\, \hat{\mathbf{x}}^{(k)} - \check{\mathbf{d}},\, \hat{\mathbf{b}}^{(k-1)*}\right)^{*}. \tag{5.24}$$

If not outputing the nonlinearity estimation result, the algorithm can derive the distorted transmit symbols based on the new information as

$$\check{\mathbf{x}}^{(k)} = \check{\mathbf{d}} + \mathbf{D}\hat{\mathbf{b}}^{(k)}. \tag{5.25}$$

Then the index increases by one as $k \rightarrow k+1$ and $\check{\mathbf{x}}^{(k)}$ is sent to $h-step$ for the next iteration.

The above iterative algorithm is summarized in the following table.

5.7.4 Simulation Results

Simulation results of iterative nonlinearity estimator are presented using both NLMS and WLMS algorithms for a practical example. The device is assumed to be an OFDM system following IEEE 802.11a/g specification [22], with the RF front-end device models from [35], used in our previous research [23, 26]. The range of nonlinearities is -16 dB EVM satisfying the specification [22] when operating in 16-QAM scheme. The output power is 25 dBm for the emitter, which is a typical average transmit power value in commercial market products, and the nonlinearity coefficients are obtained as shown in Table 5.2 In our simulation, long training sequence given in [22] is utilized, and it

Algorithm 3 Iterative Nonlinearity Estimation

1: *Initialization:* Set $\check{\mathbf{x}}^{(0)}$ and $\hat{\mathbf{h}}^{(0)}$

2: *Computation:*

$$\begin{aligned}
&\text{For } k = 1,\ 2, \ldots \text{ compute} \\
&\hat{\mathbf{h}}^{(k)} = LMS\left(\mathbf{X}^{(k-1)},\ \mathbf{r},\ \hat{\mathbf{h}}^{(k-1)*}\right)^* \\
&\hat{\mathbf{x}}^{(k)} = LMS\left(\mathbf{H}^{(k)},\ \mathbf{r},\ \check{\mathbf{x}}^{(k-1)*}\right)^* \\
&\hat{\mathbf{b}}^{(k)} = LMS\left(\mathbf{D},\ \hat{\mathbf{x}}^{(k)} - \check{\mathbf{d}},\ \hat{\mathbf{b}}^{(k-1)*}\right)^* \\
&\check{\mathbf{x}}^{(k)} = \check{\mathbf{d}} + \mathbf{D}\hat{\mathbf{b}}^{(k)}
\end{aligned}$$

3: *Output:* $\hat{\mathbf{b}}^{(k)}$

Table 5.2. Nonlinearity coefficients

nonlinearity coefficients	*numberical values*
b_3	$-0.0910 + 0.1580i$
b_5	$0.2503 + 0.0286i$
b_7	$0.0155 + 0.0025i$

contains two OFDM symbols with total period of 8 μs. After matched filtering and sampling at the symbol rate, the training sequence provides 122 samples of 64 different symbols.

5.7.4.1 Comparison of NLMS and WLMS

In this simulation, we intend to show the performance comparison of two different variable step-size LMS algorithms. Eq. (3.7) is selected as the system of linear equations to identify. Assuming eight sensors are involved with their channel models $\mathbf{h}_i,\ i = 1, \ldots, 8$ shown in Table 5.3, and without loss of generality, channel lengths of each sensor are set to three taps.

The channel vector in Table 5.3 is distributed as

$$\mathbf{h}_i \sim N_{L_{hi},\,1}(\mathbf{0}_{L_{hi}\times 1}, \sigma_h \mathbf{I}_{L_{hi}}),\ \forall i. \tag{5.26}$$

In Eq. (3.7), the channel matrix $\mathbf{H}$ is constructed using the channel models in Table 5.3

Table 5.3. Channel coefficients of four different channels

	$\delta[n]$	$\delta[n-1]$	$\delta[n-2]$
$\mathbf{h}_1$	2.342	0.072	−1.231
$\mathbf{h}_2$	−0.601	0.038	0.014
$\mathbf{h}_3$	2.077	−0.673	0.001
$\mathbf{h}_4$	−0.018	0.503	−0.036
$\mathbf{h}_5$	−0.214	1.251	0.011
$\mathbf{h}_6$	0.008	−0.394	−1.358
$\mathbf{h}_7$	2.324	−0.429	1.110
$\mathbf{h}_8$	−0.906	0.716	0.050

with no error. However, the receive signal $\mathbf{r}$ is corrupted by additive white Gaussian noise with same variance to each sample as

$$\mathbf{n} \sim N_{L_r,\,1}(\mathbf{0}_{L_r\times 1}, \sigma_n \mathbf{I}_{L_r}), \tag{5.27}$$

where $L_r = 496$ and the standard deviation of noise σ_n is calculated by setting SNR of $\mathbf{r}$ 25 dB. The quantization noise is negligible in our simulation because the number of quantization levels is assumed high. Therefore, while additive white Gaussian noise (AWGN) is high, quantization noise could be ignored in the simulation. Here, SNR is defined as the ratio between the receive signal power and the noise power. In this scenario, the SNR in each sensor is different from that of another sensor.

Fig. 5.8 shows the convergence of the each algorithm based on Eq. (3.7) using eight sensors with channel vectors from Table 5.3 after fifty realizations. The result shows that WLMS algorithm outperforms NLMS algorithm in the final NMSE of $\mathbf{x}$ in this simulation. As can be seen, the WLMS algorithm actually converges to lower NMSE than the NLMS algorithm. Hence, the WLMS algorithm shows better result than NLMS.

5.7.5 Iterative Estimator

In this simulation, the performance of the iterative estimator is presented. The channel vector of sensor i is distributed as Eq. (5.26). In Fig. 5.9, eight sensors are involved in the simulation with SNR 20 dB, and the number of iterations in the figure indicates

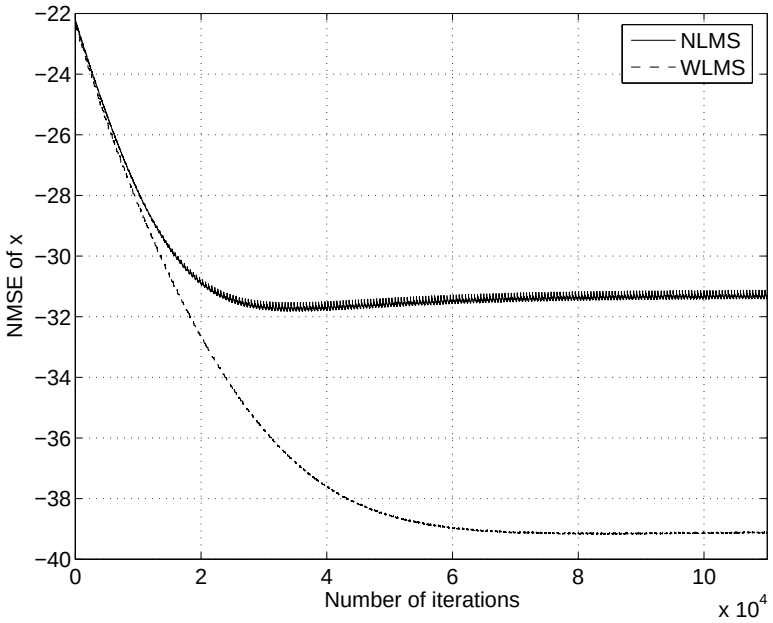


Figure 5.8. Comparison of NLMS and Weighted LMS algorithms with SNR 25 dB. The initial values are selected close to the true values.

the iteration of the IE algorithm. For each *h-* and *x-step*, the LMS algorithm executed 40000 iterations, and 640 iterations for *i-step*. As shown in the simulation, the NMSE of nonlinearity **b** decreases as the number of iterations increases, and then it meets the noise floor. Comparison of the NLMS and WLMS algorithm shows that 3 dB gain is derived using WLMS.

As the number of sensors increases, the performance is expected to improve. Figure 5.10 shows the simulation result when a different number of sensors are involved with SNR 10 dB and 20 dB cases. The simulation of SNR 10 dB shows that as the number of sensors increases, the nonlinearity estimation remains valid even with low SNR. Also, NMSE improves approximately about 3 dB if the number of sensors doubles.

Fig. 5.11 shows the result of iterative estimation using NLMS in different SNR of 50 realization. A linear fit to the data from the sixteen-sensor case is shown, and this line is shifted up 3 dB to fit the data from the eight-sensor case. This line is also shifted 4 dB and 8 dB to four-sensor and two-sensor cases. The results are also compared with the case when only data from the highest SNR sensor is utilized as Optimal SNR sensor (OptSr) algorithm. This algorithm performs better than two and four-sensor scenarios. In both Fig. 5.12 and 5.11, the performance of the two-sensor scenario in a low SNR

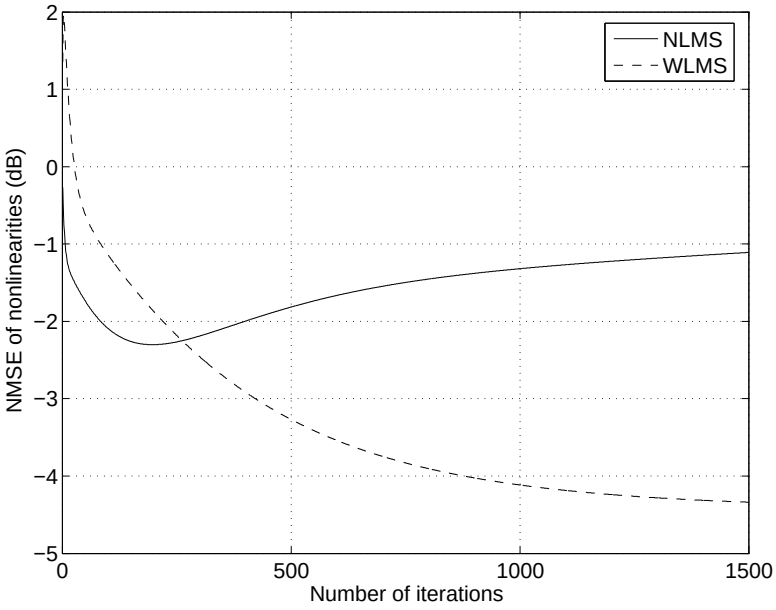


Figure 5.9. Performance of IE algorithm using NLMS and WLMS algorithm.

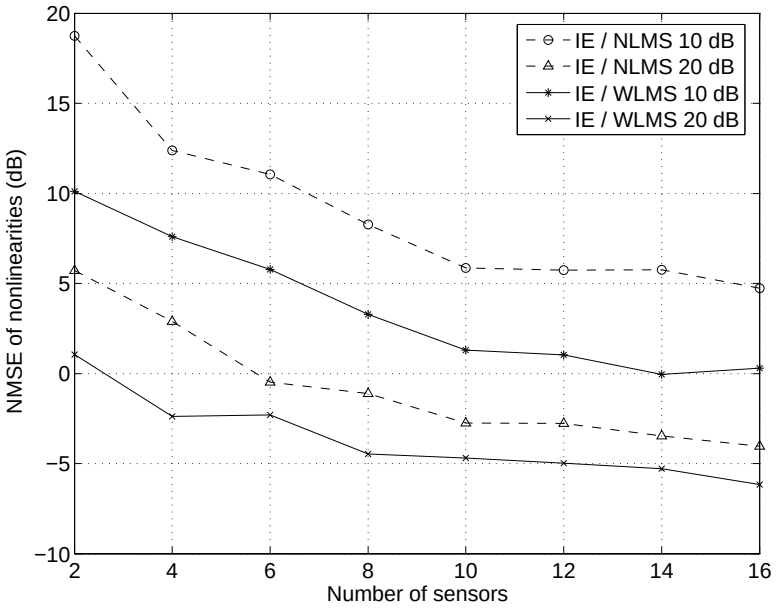


Figure 5.10. Performance with different number of sensors.

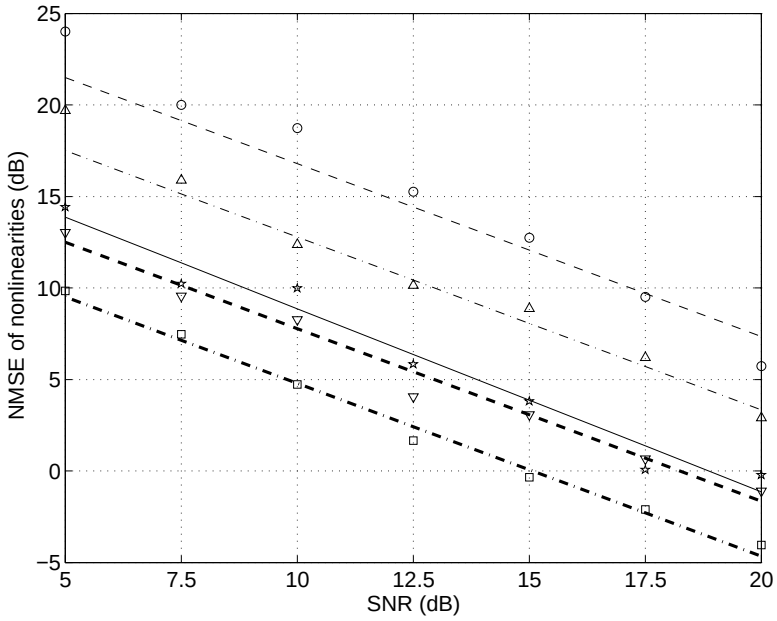


Figure 5.11. Performance of iterative estimation using NLMS with different number of sensors. In the figure, $\star$ denotes the data from OptSr algorithm, and the solid line is the linear fit of it. $\circ$ denotes 2-sensor case, $\triangle$ is 4-sensor case, $\bigtriangledown$ is 8-sensor case, and $\square$ is 16-sensor case.

environment deteriorates fast.

Fig. 5.12 shows the result of iterative estimation using WLMS for different SNR. In high SNR situation, the NMSE decreases approximately 3 dB as the number of sensors double. A linear fit for the four-sensor case is shown, and this line is shifted 3 dB higher and also 3 dB and 6 dB lower as the linear fit for two-, eight- and sixteen-sensor cases. The perofrmance of this algorithm is also shown to be better than OptSr algorithm.

5.8 Estimate-and-Forward Fusion

In this section, we develop a unbiased linear minimum variance estimation fusion rule to improve the accuracy of the estimates.

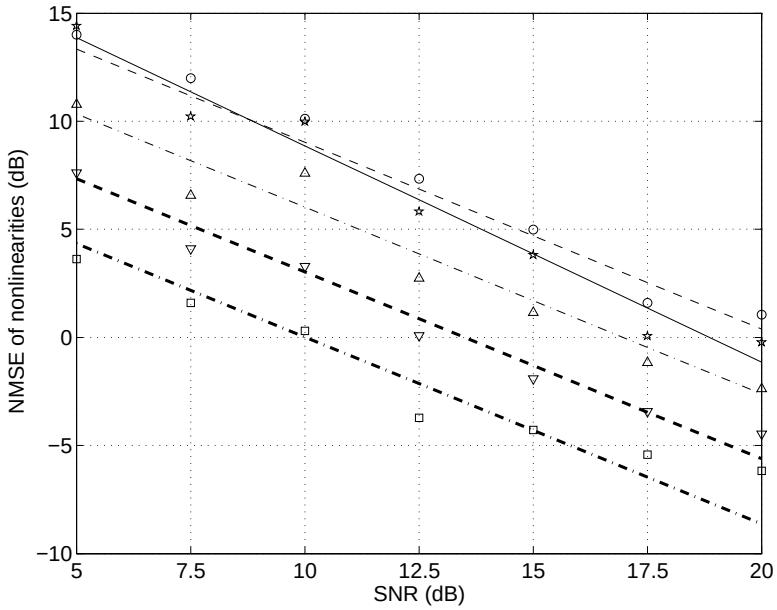


Figure 5.12. Performance of iterative estimation using WLMS with different number of sensors. $\star$ denotes the data from OptSr algorithm. In the figure, $\star$ denotes the data from OptSr algorithm, and the solid line is the linear fit of it. $\circ$ denotes 2-sensor case, $\triangle$ is 4-sensor case, $\bigtriangledown$ is 8-sensor case, and $\square$ is 16-sensor case.

5.8.1 Unbiased Linear Minimum Variance (ULMV) Estimation Fusion

To obtain linear unbiased estimation fusion rule, we consider the following estimation fusion

$$\hat{\mathbf{b}} = \sum_{q=1}^{N_s} \Theta(q)\hat{\mathbf{b}}(q) \tag{5.28}$$

where N_s is the number of sensors, $\Theta(q) = \text{diag}\,(\theta_3(q), \theta_5(q), \ldots, \theta_M(q))$ is the weighting matrix and $\hat{\mathbf{b}}(q)$ is the estimates of sensor q.

Our objective is to obtain the optimal weighting to minimize the estimation error variance while keeping the estimation results unbiased, and this objective is presented as

$$\min_{\{\Theta(1),\ldots,\Theta(N_s)\}} \sigma^2_{\hat{\mathbf{b}}}, \quad \text{s.t. } E\left\{\hat{\mathbf{b}}\right\} = \mathbf{b}, \tag{5.29}$$

where $\sigma^2_{\hat{\mathbf{b}}}$ is defined in (5.8). Since the expected values of nonlinearities are

$$E\left\{\hat{b}_{2i+1}\right\} = \sum_{q=1}^{N_s} \theta_{2i+1}(q)\hat{b}_{2i+1}(q),\ \forall i,\ q, \tag{5.30}$$

the summation of weighting coefficients is constrained as $\sum_{q=1}^{N_s} \theta_{2i+1}(q) = 1$. Using the method of Lagrangian multiplier, we thus derive the Lagrangian as

$$f = \left(\sum_{q=1}^{N_s}\sum_{i=1}^{\frac{M-1}{2}} \theta_{2i+1}(q)^2\sigma^2_{2i+1}(q)\right) + \sum_{i=1}^{\frac{M-1}{2}} \lambda_i \left(\sum_{q=1}^{N_s} \theta_{2i+1}(q) - 1\right).$$

After further derivation, the optimal weights are

$$\theta_{2i+1}(q) = \frac{\sigma^{-2}_{\hat{b}_{2i+1}}(q)}{\sum_{q=1}^{N_s} \sigma^{-2}_{\hat{b}_{2i+1}}(q)},\ \forall i,q. \tag{5.31}$$

Also, the estimation error variance after combining is

$$\begin{aligned}
\sigma^2_{\hat{\mathbf{b}}} &= \sum_{i=1}^{\frac{M-1}{2}}\sum_{q=1}^{N_s} \theta^2_{2i+1}(q)\sigma^2_{\hat{b}_{2i+1}}(q) \\
&= \sum_{i=1}^{\frac{M-1}{2}} \left[\sum_{q=1}^{N_s} \left(\frac{\sigma^{-2}_{\hat{b}_{2i+1}}(q)}{\sum_{m=1}^{N_s} \sigma^{-2}_{\hat{b}_{2i+1}}(m)}\right)^2 \sigma^2_{\hat{b}_{2i+1}}(q)\right] \\
&= \sum_{i=1}^{\frac{M-1}{2}} \left(\sum_{m=1}^{N_s} \sigma^{-2}_{\hat{b}_{2i+1}}(m)\right)^{-1}.
\end{aligned} \tag{5.32}$$

5.8.2 Comparison with Selection Combining

In this section, we prove that multi-sensor ULMV fusion rule performs better than using one sensor. Without loss of generality, assuming $0 < \sigma^2_{\hat{\mathbf{b}}}(1) \leq \ldots \leq \sigma^2_{\hat{\mathbf{b}}}(N_s)$, selection combining (SC) selects the branch that provides the maximum output signal power to error variance ratio [93] [94]. Hence the minimum error variance of the combiner output is $\sigma^2_{\hat{\mathbf{b}}}(1)$. Since the signal power of ULMV output and SC are the same, the performance depends on the error variance. Using (5.8) and (5.32), the performance of ULMV is

better than SC if

$$\frac{\sigma^2_{\hat{b}_{2i+1}}(1)}{\left(\sum_{m=1}^{N_s}\sigma^{-2}_{\hat{b}_{2i+1}}(m)\right)^{-1}} > 1, \forall i, \tag{5.33}$$

and this is true since the LHS equals $\left(1+\sigma^2_{\hat{b}_{2i+1}}(1)\left(\sum_{m=2}^{N_s}\sigma^{-2}_{\hat{b}_{2i+1}}(m)\right)\right)$ which is always greater than 1.

5.8.3 Diversity Gain using ULMV Estimation Fusion

In this section, we show the relationship between the diversity gain and the number of sensors. Channel coefficients can be deterministic or random variables, and we now consider the channel amplitude Rayleigh distributed. In case when the Gaussian noise for each sensor is the same, we assign the Gaussian error variance as σ_n^2, the pdf of channel power becomes a exponential distributed random variable. As a result, set random variable, of sensor q, $\alpha_1 = \sigma^{-2}_{\hat{b}_3}(q) = |h_i|^2/(c\sigma_n^2)$, its pdf becomes

$$f_{\alpha_1} = \frac{1}{2\sigma^2}e^{-\frac{\alpha_1}{2\sigma^2}} \tag{5.34}$$

where both c and σ are constant parameters.

From Eq. (5.32), the estimation error variance of ULMV estimation fusion is

$$\sigma^2_{\hat{b}_{2i+1}} = \left(\sum_{q=1}^{N_s}\sigma^{-2}_{\hat{b}_{2i+1}(q)}\right)^{-1}, \forall i. \tag{5.35}$$

When channel amplitudes are Rayleigh distributed random variables, from Eq.(5.34), $\sum_{q=1}^{N_s}\sigma^{-2}_{\hat{b}_3}(q)$ is a gamma distributed random variable $G(N_s, 2\sigma^2)$. Set $\alpha_2 = \sum_{q=1}^{N_s}\sigma^{-2}_{\hat{b}_{2i+1}}(q)$ then $\alpha_3 = 1/\alpha_2$ is inverse-gamma distribution

$$f_{\alpha_3}(\alpha_3,\, N_s,\, \beta) = \frac{\beta^{N_s}}{\Gamma(N_s)}\left(\frac{1}{\alpha_3}\right)^{N_s+1} e^{-\frac{\beta}{\alpha_3}} \tag{5.36}$$

where $\beta = 1/(2\sigma^2)$ is a channel model coefficient.

The expected value of inverse-gamma distribution is

$$E\left\{\sigma^2_{\hat{b}_{2i+1}}\right\} = \frac{\beta}{N_s - 1}, \text{ for } N_s > 1. \tag{5.37}$$

The diversity gain increases as the number of sensors increases. The above discussion shows that with the number of sensor increases, the performance also improves.

To place the sensors in proper positions, we consider sensor positions that minimize the expected value of estimation error variance as

$$\min_{\{v_1,\ldots,\, v_{N_s}\}} \sum_{m=1}^{N_{ms}} P_r(u_m)\sigma^2_{\hat{\mathbf{b}}}(\{v_1,\ldots,\, v_{Ns}\}, u_m) \tag{5.38}$$

where N_s is the number of sensors, $\{v_1,\ldots,v_{N_s}\}$ is the sensor positions, N_{ms} is the number of possible locations of the MS, u_m is the location of MS with probability $P_r(u_m)$, and $\sigma^2_{\hat{\mathbf{b}}}(\{v_1,\ldots,\, v_{Ns}\}, u_m)$ is the estimation error variance given sensor positions $\{v_1,\ldots,\, v_{N_s}\}$ and MS position u_m. Since the estimation error variance depends channel coefficients which can be derived through field measurements or simulations, the optimal sensor positions using (5.38) can be generating through numerical methods.

Based on simulation shown above, the misclassification rate of two routes is calculated using different number of sensors as shown in Fig. 5.13. In our simulation, the thermal noise variance of each sensor is assumed identical. MS is measured in eight of the spots uniformly distributed on both Route 1 and 2, where 16, 64, and 256 headers are used with minimum E_b/N_0 10 dB [22]. The simulation shows that to achieve 30% misclassification rate, 16 headers with two sensors or 64 headers with one sensor are required. To obtain 5% misclassification rate, 256 headers with two sensors are required.

5.9 Conclusion

We have proposed specific emitter location tracking algorithm in this chapter. This algorithm utilizes the variation of transmit power and track the emitter to the best candidate route. The effect of transmit power variation on nonlinearity estimation is also presented, and the result shows that the expected values of nonlinearity estimates are proportional to the transmit power. Since the transmit power is location dependent,

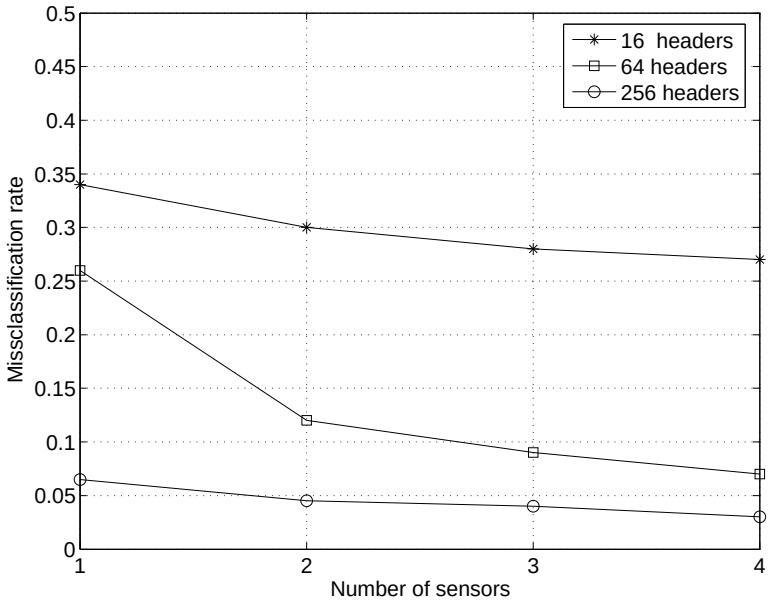


Figure 5.13. Misclassification rate between Route 1 and 2 in Fig. 5.4. ULMV is used as the estimation fusion rule.

its variation is compared to transmitter power profile of possible MS routes to identify the location of MS. Also we have provided a centralized data fusion algorithm and an estimate-and-forward algorithm. The fusion in the estimate-and-forward algorithm a uses unbiased linear minimum variance estimation fusion rule which outperforms selection combining estimation fusion. We discussed the impact of channel amplitude and showed that the estimation error variance is inversely proportional to the square of the channel amplitude. Simulation results have shown that the proposed method provides less than 5% misclassification rate given two candidate routes when there are 32 headers in E_b/N_0 10 dB environment in a single sensor scenario. While simulation results have shown that in the estimate-and-forward algorithm, the proposed method provides less than 5% misclassification rate given two candidate routes when 256 headers and two sensors are utilized.

Chapter 6

Complexity Reduction

6.1 Introduction

The major complexity of the iterative estimator is in proportional to the number of iterations required and the complexity of the inverse operation. In this chapter, both the number of iterations and complexity of the inverse operator are reduced. The complexity of the iterative algorithm is presented by the number of multipliers used. An example in Fig. 3.2 shows that under E_b/N_0 of 10 dB environment with sixty four transmit packet headers, approximate seventy five iterations are required to reach the noise floor. In the iterative algorithm, the major complexity is mainly in the pseudoinverse operations, and the matrices sizes in the pseudoinverse operations are $L_r \times P$, $L_r \times L_h$ and $P \times P$ where L_r is the length of receive symbols, L_h is the length of channel, and P is the constellation size. Therefore, based on Table 5.1 the total number of multipliers using Golub-Reinsch SVD is

$$N_{m,pinv} = N_i(4L_rP^2 + 8P^3 + 4L_rL_h^2 + 8L_h^3 + 12P^3) \tag{6.1}$$

where N_i is the number of iterations and in this case (6.2) is equal to 1.19×10^9. In Chapter 3 Section 6 a N_i reduction method is presented and the tradeoff between N_i and noise floor was discussed. In this chapter, other methods using different data sizes to process the iterative algorithm are shown. Also, LMS algorithms are also used in this chapter providing the complexity in order of

$$N_{m,LMS} = N_i(L_rP + L_rL_h + P^2) \tag{6.2}$$

which equal to 3.5×10^6 in this example.

6.2 Convergence Rate Improvement

(3.15) implies that the speed of the convergence is inversely proportional to the maximum eigenvalue of the error transform matrix. $\mathbf{Q}_b^{(k)} \approx \lambda_{max}^{k-1} \check{\mathbf{D}}^{+} \mathbf{R}_h \mathbf{R}_x \mathbf{V} \Lambda \mathbf{V}^{-1}$. Therefore, we utilize the structure of the receive data to decrease the λ_{max} coefficient.

Matrices $\mathbf{X}$ and $\mathbf{H}$ are separated by $\mathbf{X}^{\mathrm{T}} = [\mathbf{X}_1^{\mathrm{T}} | \ldots | \mathbf{X}_Z^{\mathrm{T}}]$, $\mathbf{H}^{\mathrm{T}} = [\mathbf{H}_1^{\mathrm{T}} | \ldots | \mathbf{H}_Z^{\mathrm{T}}]$, and vector $\mathbf{r}^{\mathrm{T}} = [\mathbf{r}_1^{\mathrm{T}} | \ldots | \mathbf{r}_Z^{\mathrm{T}}]$ according to the size of each segment. The iterative algorithm is then modified to estimate the first segment $\mathbf{X}_1$, $\mathbf{H}_1$ and $\mathbf{r}_1$ using the algorithm in Table 3 with only one iteration, and then this result is used as the initial value to process $\mathbf{X}_2$, $\mathbf{H}_2$ and $\mathbf{r}_2$. This process is continued until the last segment, and the result is used as the initial value for the first segment to repeat the same process again until the stopping decision is made. Semi-analytical results are shown in Appendix G.1 and the conclusion shows that the behavior of this new process in terms of NMSE can be presented as

$$\check{\mathbf{e}}_{x,Z}^{(k)} = (\mathbf{R}_{Z,1})^k \check{\mathbf{e}}_x^{(0)} + \sum_{j=0}^{k-1} \mathbf{R}_{Z,1}^j \left(\sum_{j=2}^{Z} \mathbf{R}_{Z,j} \mathbf{F}_{j-1}^{(k-i)} \mathbf{n}_{j-1} + \mathbf{F}_Z^{(k-i)} \mathbf{n}_Z \right) \tag{6.3}$$

where $\check{\mathbf{e}}_{x,Z}^{(k)}$ is the error of kth iterations from segment 1 to Z, $\mathbf{R}_{Z,1} = \mathbf{R}_Z \mathbf{R}_{Z-1} \ldots \mathbf{R}_1$, $\mathbf{R}_i$ is the error transfer matrix of segment i, and $\mathbf{F}_i^{(k)} = \mathbf{P}_D \left[\mathbf{R}_{h,i} \mathbf{X}_i^{(k-1)+} + \mathbf{H}_i^{(k)+} \right]$. The first term on the RHS of (6.3) converges while the second term increases as the iteration process. Comparing Eq. (6.3) and (3.14), the convergence behavior is similar, but the convergence rate of (3.14) is determined by $\rho(\mathbf{R})$ while that of (3.14) is determine by $\rho(\mathbf{R}_{Z,1})$. As a result, by decreasing $\rho(\mathbf{R}_{Z,1})$ the number of iterations might be reduced.

Since our algorithm collects data from multiple headers, matrices $\mathbf{H}$ and $\mathbf{X}$ have repeated structures. For instance, Appendix B is a one-header example, while $\mathbf{X}$ in a two-header example should be modified as

$$\mathbf{X} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & 0 & x_1 & x_2 & x_3 & x_4 & 0 \\ 0 & x_1 & x_2 & x_3 & x_4 & 0 & x_1 & x_2 & x_3 & x_4 \end{bmatrix}^{\mathrm{T}}. \tag{6.4}$$

In this example, the first half of the matrix, from the first to the fifth row, is identical to the latter half, from the sixth row to the last row, and similar situation can be seen

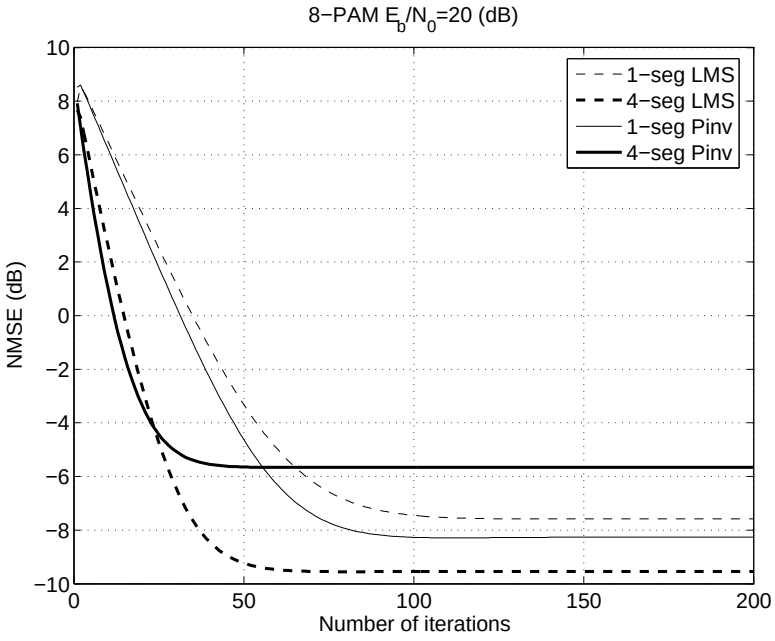


Figure 6.1. Comparison of different segment sizes using pseudoinverse and LMS algorithms.

in matrix $\mathbf{H}$. Based on this observation, if both matrices are separated by half, the the estimator processes the first half and the error transfer matrix is identical to that of processing the full block. As proved in Appendix G.2, $\mathbf{R}_{2,1} = \mathbf{R}_2\mathbf{R}_1$ where $\mathbf{R}_2 = \mathbf{R}_1$; hence $\mathbf{R}_1^{(k)} \approx \lambda_{max}^k \mathbf{V}\Lambda\mathbf{V}^{-1}$ and $\mathbf{R}_{2,1}^{(k)} \approx \lambda_{max}^{2k} \mathbf{V}\Lambda\mathbf{V}^{-1}$ if $\lambda_{max} \gg \lambda_j$, $\lambda_j \neq \lambda_{max}$. Fig. 6.1 shows the comparison of different segment sizes using both pseudoinverse and LMS algorithms. The result indicates that the algorithm using four smaller segments converges faster than that of one bigger segment using both pseudoinverse (Pinv) and LMS algorithm. The noise floor of four segments is higher than one segment in the pseudoinverse case but the LMS algorithm provides a lower noise floor. In Fig. 6.2 and 6.3, the number of iterations and multipliers required by using alternative segment sizes are shown. As a result, the more segments used, the less number of multipliers is required. Fig. 6.4 shows the NMSE performance using both pseudoinverse and LMS algorithms with different E_b/N_0 environments. As the number of segments increases, the NMSE performance worsens using pseudoinverse while the NMSE is better using the LMS algorithm.

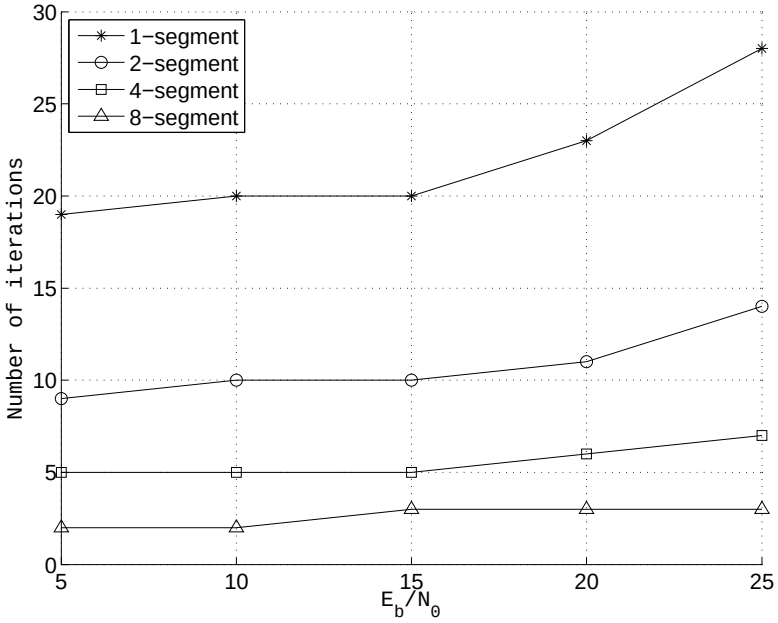


Figure 6.2. Number of iterations required using different segment sizes.

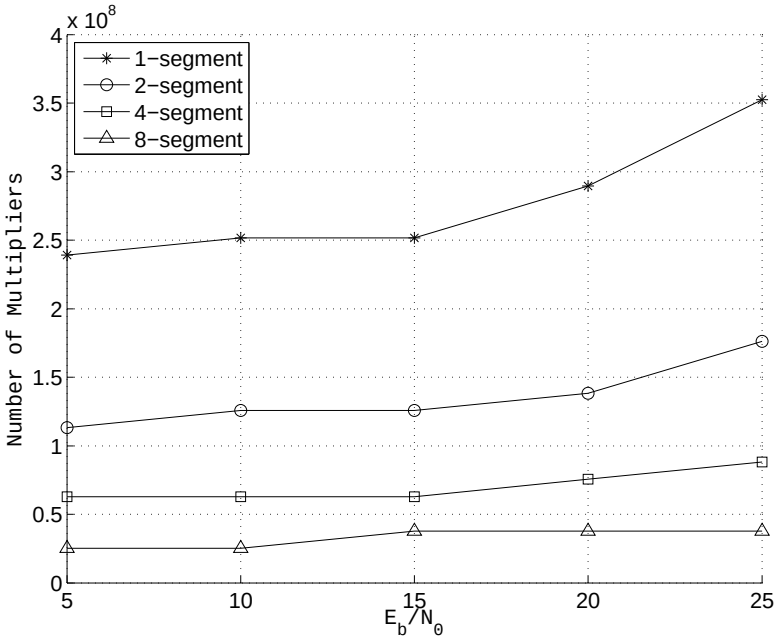


Figure 6.3. Number of multipliers required using different segment sizes and pseudoinverse.

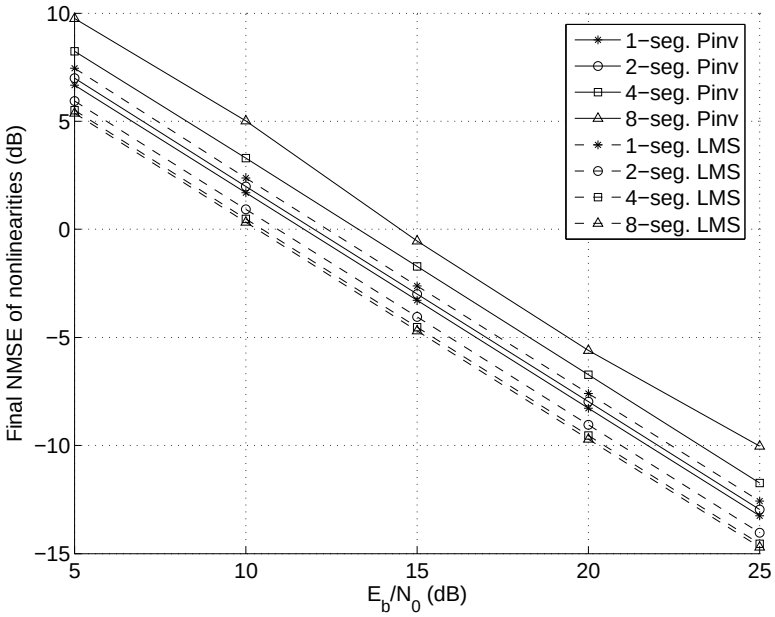


Figure 6.4. Final NMSE of nonlinearities with fixed-segment-size algorithms.

6.3 Variable Segment Size Algorithm

Once the segment size is reduced, the convergence rate improves but the noise floor also rises. Therefore, we utilize the advantage of using smaller segment sizes at the beginning of the algorithm and then use larger segment sizes afterwards.

Figure 6.5 shows the number of multipliers using full-segment size, one eighth-segment size, and variable-segment size. In the variable-segment size algorithm, the estimator starts as one eighth-segment size and switches to full-segment size after 10^8 multipliers. The number of iterations required is presented in Fig. 6.6. In this variable-segment-size example, the number of iterations used for each segment size is predetermined based on simulation results. Once the algorithm reaches the noise floor, it switches to another segment size.

However, in Fig. 6.7, LMS algorithm results show that simply using smaller segment sizes is sufficient to reach a better NMSE floor. Figure 6.8 also shows the number of multipliers required for the algorithm.

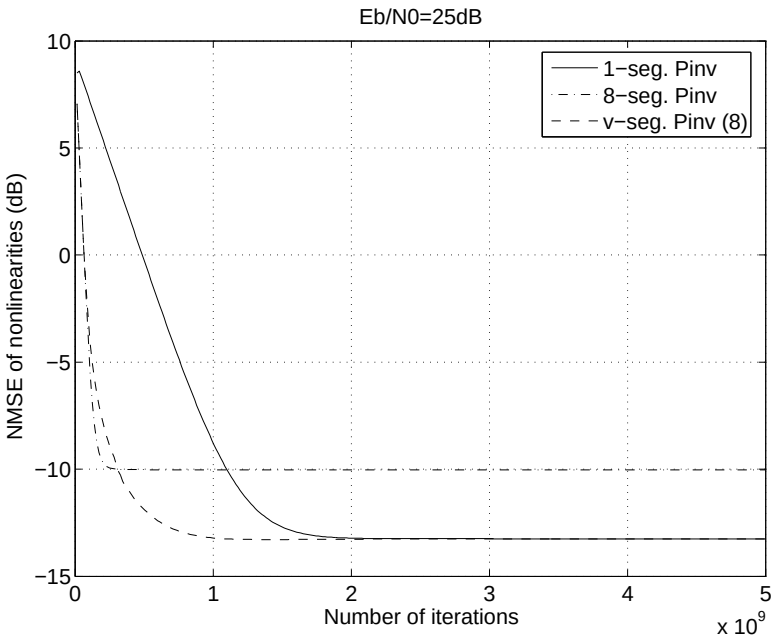


Figure 6.5. Number of multipliers required for fixed- and variable-segment-size algorithms using Golub-Reinsh LS algorithm.

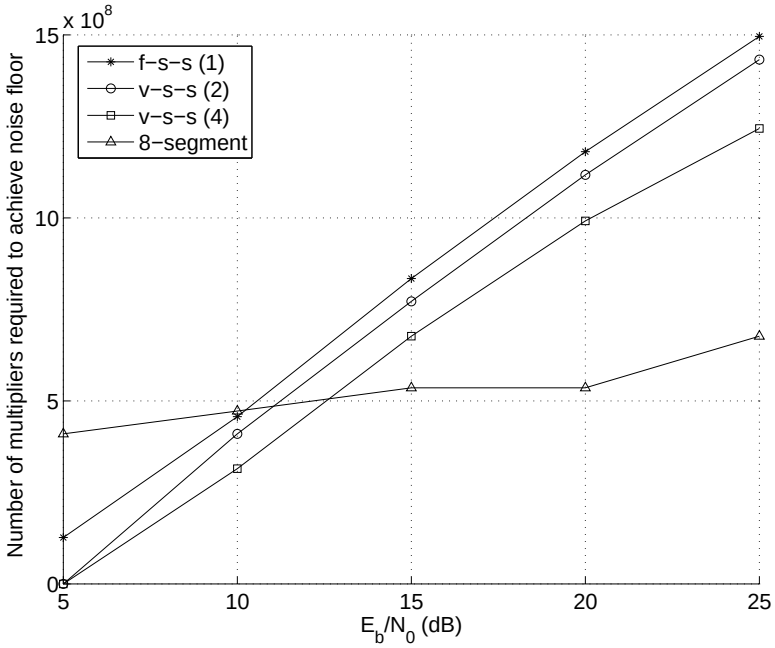


Figure 6.6. Number of multipliers required using pseudoinverse and variable segment-size algorithm. f-s-s (1) indicates fixed-segment-size of only one segment and v-s-s (i) is variable-segment-size starting from i segments to $i/2$ segments until one segment.

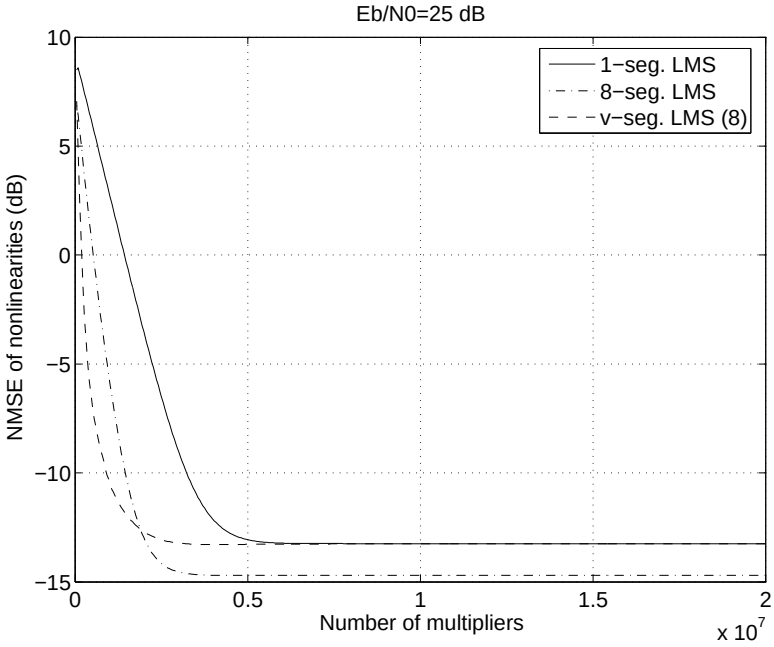


Figure 6.7. Number of multipliers required for fixed- and variable-segment-size algorithms using LMS algorithm.

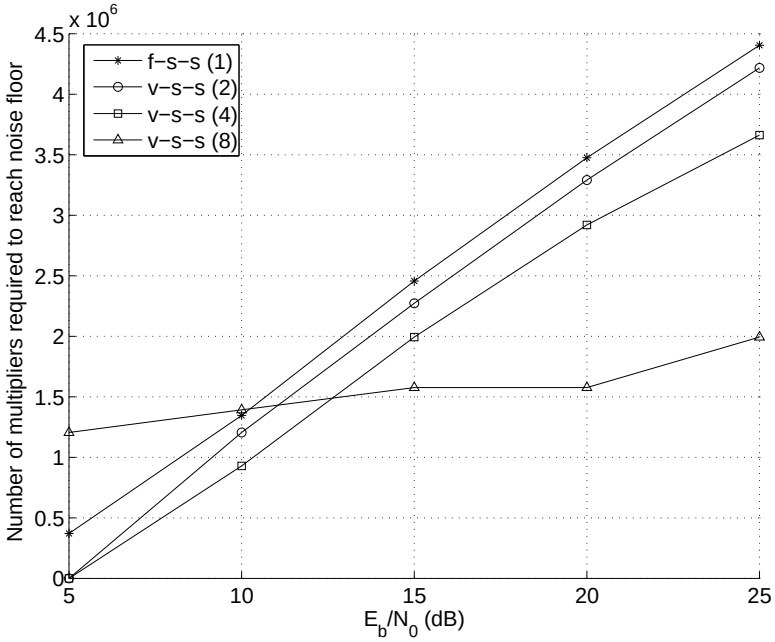


Figure 6.8. Number of multipliers required using LMS and variable segment-size algorithm.

6.4 Conclusion

In this section, we presented complexity reduction algorithms that segment the data into smaller sizes instead of using the whole data segment. The results show that using smaller segments increase the convergent rate. In terms of NMSE noise floor, smaller segment sizes though increases the rate but also rises the noise floor when pseudoinverse is used. While using classic LMS algorithms, the convergence rate is slower than that of pseudoinverse but the noise floor is lower in small-segment cases. In the last part of this chapter, variable-segment-size algorithm is also presented to utilize the fast convergent rate properties of small segment and the low noise floor of bigger segment.

Chapter 7

Summary

In this thesis, join SEI and SET algorithms were presented and simulations were conducted in practical scenarios. One major contribution in this thesis is to provide a robust transmitter nonlinearity estimator considering both stochastic signals or deterministic signals in multipath environments. Semi-analytical results of the algorithm performance were also provided for further investigation [1]. The nonlinearity estimator is also applied to channel estimation when transmitter nonlinearities are considered. The result shows BER saturates at a noise floor caused by nonlinear signal distortion and our channel estimator can provide more accurate estimates [2]. Another main contribution is the emitter tracking algorithms which jointly identify the emitter of interest and locate its current position. The algorithms utilize the power control profiles of the emitter and reduces inter-symbol interference effects. Several complexity reduction methods are provided in the last part to complete the thesis. The algorithms first reduce the number of iterations required in the iterative algorithm. Instead of processing the whole segment of data, smaller segment sizes are used to increase the convergent rate and therefore the number of iterations are reduced. Also, LMS algorithms are also explored, which change the complexity by one order compared to pseudo-inverse algorithms.

[1] **M.-W. Liu** and John F. Doherty, "Specific emitter identification in multipath channels," *IEEE Trans. Inf. Forensics Security*, Accepted for publication.

[2] **M.-W. Liu** and J. F. Doherty, "Frequency-selective multiple-input multiple-output channel estimation with transmitter non-linearities,"*Signal Processing, IET*, vol. 3, no. 6, pp. 467 475, 2009.

Appendix A

Periodogram Performance

The variance of the estimated spectra determines the performance and formulation of a classifier. The overall variance contains several sources such as the channel noise and the variance introduced in the spectra estimation procedure. The later factor depends on the methods of implementation. Spectrum estimation is a well developed area, some common used methods are periodogram, information theory approach, autocorrelation method etc.

$$E\left\{\hat{R}_{zz}(k)\right\} = \frac{1}{N}\sum_{n=0}^{N-1-k} E\left\{z(n+k)z^*(n)\right\} = \frac{1}{N}\sum_{n=0}^{N-1-k} r_{zz}(k) = \frac{N-k}{N} r_{zz}(k) \quad \text{(A.1)}$$

for $k \geq N$.

$$E\left\{\hat{R}_{zz}(k)\right\} = \begin{cases} \frac{N-|k|}{N} r_{zz}(k), & |k| \leq N \\ 0, & |k| > N \end{cases} \quad \text{(A.2)}$$

In this section we consider methods based on periodogram due to the computational efficiency of Fast Fourier transform (FFT). Supposed $r(t)$ is one realization of the stochastic process. Its periodogram can be easily computed using discrete Fourier Transform (DFT) or FFT as follows, [40]:

$$\hat{P}_{per}(l) = \frac{1}{N}\left|R_N(l)\right|^2 \quad \text{(A.3)}$$

where $R_N(l)$ is the DFT of of N-point data sequence $r_N(n)$ as defined in (2.14).

Appendix B

Example for Linear Approximation Approach

In this appendix, we introduce a simplified example of the LA method. Using the notations in Section 3.3, parameters are assigned as $U = 2$ and $P = 4$. Furthermore, the transmitter transmits 4 symbols from the alphabet $\{x_1,\, x_2,\, x_3,\, x_4\}$ through channel $\mathbf{h} = [h_1,\, h_2]^{\mathrm{T}}$. The number of received signal symbols is $L_r = L_{ts} + L_h - 1 = 4 + 2 - 1 = 5$. Therefore, the 5×2 convolution matrix $\mathbf{X}$ is composed as

$$\mathbf{X} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & 0 \\ 0 & x_1 & x_2 & x_3 & x_4 \end{bmatrix}^{\mathrm{T}}, \tag{B.1}$$

while the 5×4 channel matrix $\mathbf{H}$ and a 5×6 matrix $\mathbf{D}_{LA}$ are composed as

$$\mathbf{H} = \begin{bmatrix} h_1 & 0 & 0 & 0 \\ h_2 & h_1 & 0 & 0 \\ 0 & h_2 & h_1 & 0 \\ 0 & 0 & h_2 & h_1 \\ 0 & 0 & 0 & h_2 \end{bmatrix} \mathbf{D}_{LA} = \left[\begin{array}{cc|c|cc|cc} 0 & 0 & \check{d}_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \check{d}_1 & \check{d}_2 & 0 \\ \check{d}_3 & 0 & 0 & 0 & 0 & \check{d}_2 \\ \check{d}_4 & \check{d}_3 & 0 & 0 & 0 & 0 \\ 0 & \check{d}_4 & 0 & 0 & 0 & 0 \end{array}\right].$$

The estimation result is assumed to be $\hat{\mathbf{h}}^{(0)} = [\hat{h}^{(0)}[1],\, \hat{h}^{(0)}[2]]^{\mathrm{T}}$. Based on the same assumption, $\mathbf{x}_U = [x_1,\, x_2]^{\mathrm{T}}$ and $\mathbf{x}_L = [x_3,\, x_4]^{\mathrm{T}}$. Moreover, $\mathbf{H}_U^{(0)}$ is a 5×2 matrix and

$\mathbf{H}_L^{(0)}$ is also are 5×2 constructed as

$$\mathbf{H}_U^{(0)} = \begin{bmatrix} \hat{h}^{(0)}[1] & 0 \\ \hat{h}^{(0)}[2] & \hat{h}^{(0)}[1] \\ 0 & \hat{h}^{(0)}[2] \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \text{ and } \mathbf{H}_L^{(0)} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \hat{h}^{(0)}[1] & 0 \\ \hat{h}^{(0)}[2] & \hat{h}^{(0)}[1] \\ 0 & \hat{h}^{(0)}[2] \end{bmatrix}.$$

Appendix C

Proof for Theorem 1 in Chapter 3 Section 3.3

proof: From Section 3.3, the nonlinearity multiplier of the "nonlinear symbols" is defined as $\alpha_i = x_i/\check{d}_i$, $i = \{1,2,\ldots,P\}$. Hence, in the case when the amplitudes of low-amplitude symbols are identical; their nonlinearity multipliers are the same. Assuming $\alpha_i = \alpha$, $i = \{U+1, U+2, \ldots, P\}$, the following equation is derived using $\hat{\mathbf{h}}_{LA} = \mathbf{D}_{LA}^{+}\mathbf{r}$:

$$\mathbf{D}_{LA}\mathbf{h}_{LA} + \mathbf{D}_{LA}\begin{bmatrix} (\alpha-1)\mathbf{h} \\ \mathbf{0} \end{bmatrix} + \mathbf{n} = \mathbf{r}, \tag{C.1}$$

where $\mathbf{0}$ is a $L_h U \times 1$ zero vector, and therefore the estimated channel coefficient is derived as: $\hat{\mathbf{h}}^{(0)} = \begin{bmatrix} \mathbf{I}_{L_h} \,\big|\, \mathbf{0} \end{bmatrix} \hat{\mathbf{h}}_{LA} = \mathbf{h} + (\alpha-1)\mathbf{h} + \begin{bmatrix} \mathbf{I}_{L_h} \,\big|\, \mathbf{0} \end{bmatrix} \mathbf{D}_{LA}^{+}\mathbf{n}$. In the high SNR case, the additive noise is weak, which means $\hat{\mathbf{h}}^{(0)} \approx \alpha\mathbf{h}$. Substituting into (3.8),

$$\begin{aligned} \hat{\mathbf{x}}_U^{(0)} &\approx \left(\frac{1}{\alpha}\mathbf{H}_U^{+}\right)\left[\mathbf{r} - (\alpha\mathbf{H}_L)\check{\mathbf{d}}_L\right] \\ &\approx \frac{1}{\alpha}\mathbf{H}_U^{+}\left[\mathbf{H}_U\mathbf{x}_U + \mathbf{H}_L\mathbf{x}_L + \mathbf{n} - \mathbf{H}_L(\alpha\check{\mathbf{d}}_L)\right] \\ &\approx \frac{1}{\alpha}\left(\mathbf{x}_U + \mathbf{H}_U^{+}\mathbf{n}\right), \end{aligned} \tag{C.2}$$

where $\mathbf{x}_L = \alpha\check{\mathbf{d}}_L$. This result is substituted into (3.9),

$$\hat{\mathbf{b}}^{(0)} \approx \check{\mathbf{D}}_U^{+}\left[\frac{1}{\alpha}\left(\mathbf{x}_U + \mathbf{H}_U^{+}\mathbf{n}\right) - \check{\mathbf{d}}_U\right]. \tag{C.3}$$

After simple manipulation, (C.3) results in (3.10).

Appendix D

Proofs for Theorems in Chapter 3 Section 3.4

D.1

Theorem: The nonlinearity estimates before and after the *i-step* are identical.

Proof: Suppose the nonlinearity estimation after the *i-step* is $\check{\mathbf{b}}^{(k)}$ at kth iteration, then $\check{\mathbf{b}}^{(k)} = \check{\mathbf{D}}^{+}(\check{\mathbf{x}}^{(k)} - \check{\mathbf{d}})$. Then using Algorithm 3, we have $\check{\mathbf{x}}^{(k)} = \check{\mathbf{d}} + \mathbf{P}_{\check{D}}(\mathbf{x} + \mathbf{e}_x^{(k)} - \check{\mathbf{d}}) = \mathbf{x} + \mathbf{P}_{\check{D}}\mathbf{e}_x^{(k)}$. Substituting this equation into the previous equation, we obtain $\check{\mathbf{b}}^{(k)} = \check{\mathbf{D}}^{+}(\mathbf{x} + \mathbf{P}_{\check{D}}\mathbf{e}_x^{(k)} - \check{\mathbf{d}}) = \mathbf{b} + \check{\mathbf{D}}^{+}\mathbf{e}_x^{(k)}$. This is the same result when substituting $\hat{\mathbf{x}}^{(k)} = \mathbf{x} + \mathbf{e}_x^{(k)}$ into (3.13).

D.2

Theorem: The nonlinearity estimation error at the kth iteration is shown in (3.14) and

$$\mathbf{Q}_b^{(k)} = \check{\mathbf{D}}^{+}\mathbf{R}_h\mathbf{R}_x\mathbf{R}^{k-1}, \tag{D.1}$$

, where $\mathbf{e}_x^{(0)}$ is the initial estimation error, and

$$\mathbf{W}_b^{(k)} = \check{\mathbf{D}}^{+}\left\{\mathbf{R}_h\mathbf{R}_x\left[\sum_{i=0}^{k-1}\mathbf{R}^i\mathbf{P}_{\check{D}}\mathbf{R}_h\mathbf{X}^{(k-i-1)+} + \right.\right.$$

$$\sum_{i=0}^{k-1} \mathbf{R}^i \mathbf{P}_{\check{D}} \mathbf{H}^{(k-i)+} \Bigg] + \mathbf{R}_h \mathbf{X}^{(k-1)+} + \mathbf{H}^{(k)+} \Bigg\}. \tag{D.2}$$

proof: Assign $\hat{\mathbf{h}}^{(k)} = \mathbf{h} + \mathbf{e}_h^{(k)}$, $\hat{\mathbf{x}}^{(k)} = \mathbf{x} + \mathbf{e}_x^{(k)}$, and $\check{\mathbf{x}}^{(k)} = \mathbf{x} + \check{\mathbf{e}}_x^{(k)}$, where $\mathbf{e}_h^{(k)}$, $\mathbf{e}_x^{(k)}$, and $\check{\mathbf{e}}_x^{(k)}$ are the error vector containing both the estimator error and error driven by the additive noise in the kth step. Similar to the error vectors, we assign $\mathbf{X}^{(k)} = \mathbf{X} + \mathbf{E}_x^{(k)}$ and $\mathbf{H}^{(k)} = \mathbf{H} + \mathbf{E}_h^{(k)}$ in the kth step. In the following proof, L_r is the number of receive symbols, P is the length of vector $\mathbf{x}$, and L_h is the channel length as defined in Section 4.2.

For the kth *h-step*, the estimation is written in terms of error vector as

$$\begin{aligned}
\hat{\mathbf{h}}^{(k)} &= \mathbf{X}^{(k-1)+}\mathbf{r} = \mathbf{X}^{(k-1)+}(\mathbf{X}^{(k-1)}\mathbf{h} + \mathbf{E}_x^{(k-1)}\mathbf{h} + \mathbf{n}) \\
&= \mathbf{X}^{(k-1)+}\mathbf{X}^{(k-1)}\mathbf{h} + \mathbf{X}^{+}\mathbf{E}_x^{(k-1)}\mathbf{h} \\
&\quad + (\mathbf{X}^{(k-1)+} - \mathbf{X}^{+})\mathbf{E}_x^{(k-1)}\mathbf{h} + \mathbf{X}^{(k-1)+}\mathbf{n} \\
&\approx \mathbf{h} + (\mathbf{X}^{+}\mathbf{E}_x^{(k-1)}\mathbf{h} + \mathbf{X}^{(k-1)+}\mathbf{n}).
\end{aligned} \tag{D.3}$$

Here we assume both $\left\| \mathbf{X}^{(k-1)+} - \mathbf{X}^{+} \right\|$ and $\left\| \mathbf{E}_x^{(k-1)} \right\|$ are small; hence $(\mathbf{X}^{(k-1)+} - \mathbf{X}^{+})\mathbf{E}_x^{(k-1)} \rightarrow \mathbf{0}$. To further extract the error vector from (D.3), we expand the error matrix $\mathbf{E}_x^{(k-1)}$

$$\mathbf{E}_x^{(k-1)} = \sum_{i=1}^{P} (\check{e}_{x_i}^{(k-1)} \mathbf{I}_{L_r}) \mathbf{S}_{x,i} = \sum_{i=1}^{P} \left[\sum_{j=1}^{L_r} \boldsymbol{\psi}_{L_r,j} \check{\mathbf{e}}_x^{(k-1)\mathrm{T}} \boldsymbol{\Psi}_{x,i,j} \right] \mathbf{S}_{x,i}.$$

Here $\mathbf{S}_{x,i} \in \mathbb{R}^{L_r \times L_h}$ is a selection matrix for x_i, $i \in \{1, \ldots, P\}$, in which the elements of $\mathbf{S}_{x,i}$ are set to 1 only when x_i has contribution to $\mathbf{X}$ corresponding to those positions and the rest of the elements are set to zero. Also $\mathbf{I}_{L_r} \in \mathbb{R}^{L_r \times L_r}$ is an identity matrix, $\boldsymbol{\psi}_{L_r,j} \in \mathbb{R}^{L_r \times 1}$ sets only the jth element to 1 and the remaining elements zero, $\check{\mathbf{e}}_x^{(k-1)} = [\check{e}_{x_1}^{(k-1)}, \check{e}_{x_2}^{(k-1)}, \ldots, \check{e}_{x_P}^{(k-1)}]^{\mathrm{T}}$ is the error vector from *i-step* when $k > 1$ and $\check{\mathbf{e}}^{(0)} = \mathbf{e}^{(0)}$ from LA approach when $k = 1$, and $\boldsymbol{\Psi}_{x,i,j} \in \mathbb{R}^{P \times L_r}$ assigns 1 to the ith row and jth column while other elements of the matrix are set to zero. Then we can derive

$$\begin{aligned}
\mathbf{X}^{+}\mathbf{E}_x^{(k-1)}\mathbf{h} &= \mathbf{X}^{+} \left\{ \sum_{i=1}^{P} \left[\sum_{j=1}^{L_r} \boldsymbol{\psi}_{L_r,j} \check{\mathbf{e}}_x^{(k-1)\mathrm{T}} \boldsymbol{\Psi}_{x,i,j} \right] \mathbf{S}_{x,i} \right\} \mathbf{h} \\
&= \left\{ \sum_{i=1}^{P} \sum_{j=1}^{L_r} \left[\mathbf{X}^{+} \boldsymbol{\psi}_{L_r,j} \right] \left[\boldsymbol{\Psi}_{x,i,j} \mathbf{S}_{x,i} \mathbf{h} \right]^{\mathrm{T}} \right\} \check{\mathbf{e}}_x^{(k-1)}.
\end{aligned} \tag{D.4}$$

We further represent this equation by $\mathbf{X}^{+}\mathbf{E}_x^{(k-1)}\mathbf{h} \equiv \mathbf{R}_x\breve{\mathbf{e}}_x^{(k-1)}$, where $\mathbf{R}_x \in \mathbb{C}^{L_h \times P}$. Combining (D.3) and (F.2) , we have $\mathbf{e}_h^{(k)} = \mathbf{R}_x\breve{\mathbf{e}}_x^{(k-1)} + \mathbf{X}^{(k-1)+}\mathbf{n}$.

For the *x-step*, similar strategy is used to obtain $\hat{\mathbf{x}}^{(k)} - \mathbf{x} = \mathbf{e}_x^{(k)} \approx \mathbf{H}^{+}\mathbf{E}_h^{(k)}\mathbf{x} + \mathbf{H}^{(k)+}\mathbf{n}$. And the following result is also derived as:

$$\mathbf{H}^{+}\mathbf{E}_h^{(k)}\mathbf{x} = \left[\sum_{i=1}^{L_h}\sum_{j=1}^{L_r}\left(\mathbf{H}^{+}\boldsymbol{\psi}_{L_r,j}\right)\left(\Psi_{h,i,j}\mathbf{S}_{h,i}\mathbf{x}\right)^{\mathrm{T}}\right]\mathbf{e}_h^{(k)}, \tag{D.5}$$

where $\Psi_{h,i,j} \in \mathbb{R}^{L_h \times L_r}$ provides similar function assigning 1 to the ith row and jth column component while others set zero, and $\mathbf{S}_{h,i} \in \mathbb{R}^{L_r \times P}$ is the selection matrix for h_i, $i \in \{1, \ldots, L_h\}$. For convenience, we rewrite (F.3) as $\mathbf{H}^{+}\mathbf{E}_h^{(k)}\mathbf{x} \equiv \mathbf{R}_h \cdot \mathbf{e}_h^{(k)}$, where $\mathbf{R}_h \in \mathbb{C}^{P \times L_h}$.

In conclusion, the error vectors of *h-* and *x-step* are approximated as $\mathbf{e}_h^{(k)} \approx \mathbf{R}_x\mathbf{e}_x^{(k-1)} + \mathbf{X}^{(k-1)+}\mathbf{n}$, and $\mathbf{e}_x^{(k)} \approx \mathbf{R}_h\mathbf{e}_h^{(k)} + \mathbf{H}^{(k)+}\mathbf{n}$. Then using Theorem D.1, the $i-step$ error vector is $\breve{\mathbf{e}}_x^{(k)} = \mathbf{P}_{\breve{D}}\mathbf{e}_x^{(k)}$ and that of the nonlinearity estimate is $\mathbf{e}_b^{(k)} = \breve{\mathbf{D}}^{+}\mathbf{e}_x^{(k)}$. After inductive reasoning based on the above error vectors, this theorem is derived.

D.3

Theorem: The nonlinearity coefficients estimator is asymptotically unbiased if the spectral norm of $\mathbf{R}$ is less than 1, $\rho(\mathbf{R}) < 1$ where $\mathbf{R} = \mathbf{P}_{\breve{D}}\mathbf{R}_h\mathbf{R}_x$. By replacing $\mathbf{P}_{\breve{D}}$ by $\mathbf{I}_P$, this principal is also valid when the *i-step* is not used.

Proof: From (3.14) with $\rho(\mathbf{R}) < 1$, the bias term is $\mathbf{Q}_b^{(k)}\mathbf{e}_x^{(0)} = \mathbf{D}^{+}\mathbf{R}_h\mathbf{R}_x\mathbf{R}^{k-1}\mathbf{e}_x^{(0)} \rightarrow \mathbf{0}$ as $k \rightarrow \infty$. Once $\rho(\mathbf{R}) < 1$, matrix $\mathbf{R}$ will converge to a zero matrix asymptotically after sufficient number of iterations and thus the error caused by initial estimation will approach zero.

Remark: If $\rho(\mathbf{R}) < 1$, then $\mathbf{Q}_b^{(k)}\mathbf{e}_x^{(0)}$ in (3.14) converges to a zero vector.

Remark: If $\rho(\mathbf{R}) < 1$, then $\mathbf{W}_b^{(k)}\mathbf{n}$ in (3.14) converges to a fixed vector.

Proof: Assuming $\rho(\mathbf{R}) < 1$ and $\mathbf{R}^k \rightarrow \mathbf{0}\ \forall k > v$, the summations in (D.2) become fixed matrices when $k > v$. Therefore, as long as the spectral norm assumption holds, these terms will saturate after a sufficient number of iterations.

The MSE of the nonlinearity estimator also saturates under the same assumption. Assuming zero mean Gaussian noise and $\mathbf{e}_x^{(0)} = \boldsymbol{\varepsilon}_x^{(0)} + \mathbf{n}_x^{(0)}$ where $\boldsymbol{\varepsilon}_x^{(0)}$ is the bias term

and $\mathbf{n}_x^{(0)}$ is the additive noise term as described in Theorem 3.3. the following is obtained from (3.14):

$$E\left[\mathbf{e}_b^{(k)\dagger}\mathbf{e}_b^{(k)}\right] = \boldsymbol{\varepsilon}_x^{(0)\dagger}\mathbf{Q}_b^{(k)\dagger}\mathbf{Q}_b^{(k)}\boldsymbol{\varepsilon}_x^{(0)} + E\left[\mathbf{n}_x^{(0)\dagger}\mathbf{Q}_b^{(k)\dagger}\mathbf{Q}_b^{(k)}\mathbf{n}_x^{(0)}\right] + E\left[\mathbf{n}^{\dagger}\mathbf{W}_b^{(k)\dagger}\mathbf{W}_b^{(k)}\mathbf{n}\right] \tag{D.6}$$

where $\mathbf{e}_b^{(k)} = \hat{\mathbf{b}}^{(k)} - \mathbf{b}$. Furthermore, when $\rho(\mathbf{R}) < 1$, $\mathbf{Q}_b^{(\infty)} \to \mathbf{0}$ and $\mathbf{W}_b^{(\infty)}$ converges to a fixed matrix. As a result, as $k \to \infty$, the first and the second terms on the RHS of (D.6) asymptotically approach zero; the MSE of nonlinearity coefficient estimate is independent of the initial estimation value, and converges to a noise floor.

Appendix E

Generation of matrices X and H

To describe the details of generating matrices $\mathbf{X}$ and $\mathbf{H}$, several definitions are assigned for convenience. Define vector $\mathbf{x}(i)[\tau] = [x(i)[\tau],\ x(i)[\tau-2], \ldots,\ x(i)[\tau-L_h+1]]^{\mathrm{T}}$, where $x(i)[j] = 0$ when $j \leq 0$. Furthermore, a $\mathbb{C}^{n_R \times L_h n_T}$ matrix $\bar{\mathbf{X}}(i)[\tau]$ in (4.7) is defined as following

$$\bar{\mathbf{X}}(i)[\tau] = \begin{bmatrix} \mathbf{x}(i)[\tau] & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{x}(i)[\tau] & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \ldots & \mathbf{x}(i)[\tau] \end{bmatrix}, \tag{E.1}$$

where $\mathbf{0}$ is a $1 \times L_h$ zero vector. Therefore, we can define $\mathbf{X}$ as a $\mathbb{C}^{L_r n_R \times L_h n_R n_T}$ matrix:

$$\mathbf{X} = \begin{bmatrix} \bar{\mathbf{X}}(1)[1] & \bar{\mathbf{X}}(2)[1] & \ldots & \bar{\mathbf{X}}(n_T)[1] \\ \bar{\mathbf{X}}(1)[2] & \bar{\mathbf{X}}(2)[2] & \ldots & \bar{\mathbf{X}}(n_T)[2] \\ \vdots & \vdots & \ddots & \vdots \\ \bar{\mathbf{X}}(1)[L_r] & \bar{\mathbf{X}}(2)[L_r] & \ldots & \bar{\mathbf{X}}(n_T)[L_r] \end{bmatrix}. \tag{E.2}$$

where $Lr = L_{ts} + L_h - 1$ is number of the receive symbols, L_{ts} is the number of training sequence symbols, and L_h is the number of channel taps.

Here we provide an example to explain the composition of matrix $\mathbf{H} \in \mathbb{C}^{L_r n_R \times P}$ in (4.7). Given a 2×2 MIMO system, channel length $L_h = 3$, number of possible transmit symbols $P = 9$, and length of training sequence $L_{ts} = 160$. Hence, $\mathbf{x} = [x_1, \ldots, x_9]^{\mathrm{T}}$, and $\mathbf{H} \in \mathbb{C}^{324 \times 9}$. Suppose the first transmit element sends $x(1)[1],\ x(1)[2], \ldots,\ x(1)[160]$,

and assume they correspond to x_2, $x_2, \ldots,$ x_9. While the second transmit element sends $x(2)[1]$, $x(2)[2], \ldots,$ $x(2)[160]$ corresponding to x_1, $x_3, \ldots,$ x_3. Therefore, $\mathbf{H}$ in (4.7) becomes

$$\mathbf{H} = \begin{bmatrix} h_{12}[0] & h_{11}[0] & 0 & \ldots & 0 \\ h_{22}[0] & h_{21}[0] & 0 & \ldots & 0 \\ h_{12}[1] & h_{11}[0]+h_{11}[1] & h_{12}[0] & \ldots & 0 \\ h_{22}[1] & h_{21}[0]+h_{21}[1] & h_{22}[0] & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & h_{12}[2] & \ldots & h_{11}[2] \\ 0 & 0 & h_{22}[2] & \ldots & h_{21}[2] \end{bmatrix}. \tag{E.3}$$

From the above discussion, the population of channel coefficients in this matrix are related to the linear combination of channel coefficients and the transmit symbols.

Appendix F

Proofs for Chapter 5

F.1 Impact on the Transmit Power

This appendix shows the variation of the transmit power influences the estimation accuracy. The nonlinearity estimator error can be decomposed by the error caused by the initial estimation error and the additive white Gaussian noise. The result from [28] shows that the estimator error at kth iteration is

$$\hat{\mathbf{b}}^{(k)} - \mathbf{b} = \mathbf{Q}_b^{(k)} \mathbf{e}_x^{(0)} + \mathbf{W}_b^{(k)} \mathbf{n}, \tag{F.1}$$

where $\mathbf{Q}_b^{(k)}$ and $\mathbf{W}_b^{(k)}$ are defined in (D.1) and (D.2). Here, $\mathbf{R} = \mathbf{P}_{\check{\mathbf{D}}} \mathbf{R}_h \mathbf{R}_x$, $\mathbf{P}_{\check{\mathbf{D}}} = \check{\mathbf{D}} \check{\mathbf{D}}^{+}$,

$$\mathbf{R}_x = \sum_{i=1}^{P} \sum_{j=1}^{L_r} \left[\mathbf{X}^{+} \psi_{L_r,j} \right] \left[\Psi_{x,i,j} \mathbf{S}_{x,i} \mathbf{h} \right]^{\mathrm{T}}, \tag{F.2}$$

and

$$\mathbf{R}_h = \sum_{i=1}^{L_h} \sum_{j=1}^{L_r} \left[\mathbf{H}^{+} \psi_{L_h,j} \right] \left[\Psi_{h,i,j} \mathbf{S}_{h,i} \mathbf{x} \right]^{\mathrm{T}} \tag{F.3}$$

where $\mathbf{S}_{x,i} \in \mathbb{R}^{L_r \times L_h}$ is a selection matrix for x_i, $i \in \{1, \ldots, P\}$, in which the elements of $\mathbf{S}_{x,i}$ are set to 1 only when x_i has contribution to $\mathbf{X}$ corresponding to those positions and the rest of the elements are set to zero. $\mathbf{S}_{h,i} \in \mathbb{R}^{L_r \times P}$ is a selection matrix for $\mathbf{h}$, $\psi_{L_r,j} \in \mathbb{R}^{L_r \times 1}$ sets only the jth element to 1 and the remaining elements zero, and $\mathbf{e}_x^{(k-1)}$ is the error vector of $\mathbf{x}$ at $(k-1)$th iteration. $\Psi_{x,i,j} \in \mathbb{R}^{P \times L_r}$ assigns 1 to the ith row and jth column while other elements of the matrix are set to zero.

Therefore, if the maximum eigenvalue of $\mathbf{R}$ is less than 1, $\mathbf{R}^{(\infty)} \rightarrow \mathbf{0}$ and $\mathbf{Q}_b^{(\infty)} \rightarrow \mathbf{0}$. This indicates that the estimate is asymptotically unbiased. The estimator variance as $k \rightarrow \infty$ is

$$\sigma_{\hat{\mathbf{b}}^2} = diag(\mathbf{W}_b^{(\infty)\dagger}\mathbf{W}_b^{(\infty)})\sigma_n^2, \tag{F.4}$$

where $\sigma_{\hat{\mathbf{b}}^2}$ is defined in (5.8). Here we consider the scenario where the transmit power varies from σ_d^2 to $\gamma\sigma_d^2$ where γ is the ratio of the power variation and the channel $\mathbf{h}'$ in (5.3) is stationary in the observation time. Though the increase in transmit power effects the transmit symbol constellation, the variation in the nonlinearities is limited because most specifications of communication systems specify the upper limits of nonlinearity effects on transmit symbol constellation. For instance, the error vector magnitude requirement (EVM) in [22] is -16 dB and -21 dB in [83]. Compared to the limited nonlinearity variations, the degree of freedom in transmit power is higher. For example, the transmitter is required to provide dynamic power control range of at least 50 dB in no less than -1 dB steps in [83].

As a result, we assume that the transmit power constellation in (3.7) is approximately the same, namely $\mathbf{x}_\gamma \approx \mathbf{x}$, and we focus on the effects on the equivalent channel $\mathbf{h}_\gamma = \sqrt{\gamma}\mathbf{h}$ from (5.3). The performance of the nonlinearity estimation is derived in (F.4). The channel matrix with new transmit power is $\mathbf{H}_\gamma = \sqrt{\gamma}\mathbf{H}$, and the pseudoinverse becomes $\mathbf{H}_\gamma^+ = \mathbf{H}^+/\sqrt{\gamma}$. The relationship between the original and new error transform matrixes are also derived as $\mathbf{R}_{h,\gamma} = \mathbf{R}_h/\sqrt{\gamma}$, $\mathbf{R}_{x,\gamma} = \sqrt{\gamma}\mathbf{R}_x$, and $\mathbf{R}_\gamma = \mathbf{R}$. Based on the above derivation and (F.4), $\mathbf{Q}_b$ is identical even the transmit power changes which leads to $k \rightarrow \infty$ $\mathbf{Q}_b^{(\infty)} \rightarrow \mathbf{0}$, and therefore $\sigma_{b,\gamma}^{2(\infty)} = \sigma_b^{2(\infty)}/\gamma^2$.

F.2 Impact on Channel Gain

From (D.1), we know the estimation error variance is a function of $\mathbf{H}$, $\mathbf{R}_x$ $\mathbf{R}_h$ and $\mathbf{R}$ where $\mathbf{R}_x$ and $\mathbf{R}_h$ are the error transform matrices of *x*- and *h-step*, and $\mathbf{R} = \mathbf{P}_{\check{D}}\mathbf{R}_h\mathbf{R}_x$. Given $\mathbf{h} = h$, we can show that $\|\mathbf{H}\|_2 \sim |h|$ and $\|\mathbf{H}^+\|_2 \sim |h|^{-1}$. Using (F.2) and (F.3), we obtain that $\mathbf{R}_x \sim h$ and $\mathbf{R}_h \sim h^{-1}$. Furthermore, since $\mathbf{P}_{\check{D}}$ is fixed, then $\mathbf{R}$ is not a function of h. For the same transmitter, suppose as the number of iteration approaches infinity $\mathbf{x}^{(\infty)} \rightarrow \hat{\mathbf{x}}$ and $\hat{h}^{(\infty)} \rightarrow h$, then $\mathbf{W}_b^{(\infty)} \sim h^{-1}$. Therefore, the variance of nonlinearity estimation is in inversely proportional to the power of channel amplitude.

Appendix G

Proofs for Chapter 6

G.1 Proof for Error Accumulation

Assuming there are Z segments in a block, and $Z \geq 2$. In the iteration of the first segment generate the result as

$$\begin{aligned}
\check{\mathbf{e}}_{x,1}^{(1)} &= \mathbf{R}_1 \check{\mathbf{e}}_x^{(0)} + \mathbf{F}_1^{(1)} \mathbf{n}_1 \\
\check{\mathbf{e}}_{x,2}^{(1)} &= (\mathbf{R}_2 \mathbf{R}_1) \check{\mathbf{e}}_x^{(0)} + \mathbf{R}_2 \mathbf{F}_1^{(1)} \mathbf{n}_1 + \mathbf{F}_2^{(1)} \mathbf{n}_2 \\
&\vdots \\
\check{\mathbf{e}}_{x,Z}^{(1)} &= (\mathbf{R}_Z \dots \mathbf{R}_2 \mathbf{R}_1) \check{\mathbf{e}}_x^{(0)} + (\mathbf{R}_Z \dots \mathbf{R}_2) \mathbf{F}_1^{(1)} \mathbf{n}_1 + (\mathbf{R}_Z \dots \mathbf{R}_3) \mathbf{F}_2^{(1)} \mathbf{n}_2 + \dots \\
&\quad + \mathbf{R}_Z \mathbf{F}_{Z-1}^{(1)} \mathbf{n}_{Z-1} + \mathbf{F}_Z^{(1)} \mathbf{n}_Z
\end{aligned} \tag{G.1}$$

where $\mathbf{F}_i^{(k)} = \mathbf{P}_D \left[\mathbf{R}_{h,i} \mathbf{X}_i^{(k-1)+} + \mathbf{H}_i^{(k)+} \right]$, and $\mathbf{R}_i$ is the error transfer function of segment i. The summation of the above equation is

$$\check{\mathbf{e}}_{x,Z}^{(k)} = (\mathbf{R}_{1,Z})^k \check{\mathbf{e}}_x^{(0)} + \sum_{j=0}^{k-1} \mathbf{R}_{Z,1}^j \left(\sum_{j=2}^{Z} \mathbf{R}_{Z,j} \mathbf{F}_{j-1}^{(k-i)} \mathbf{n}_{j-1} + \mathbf{F}_Z^{(k-i)} \mathbf{n}_Z \right) \tag{G.2}$$

where $\mathbf{R}_{Z,1} = \mathbf{R}_Z \mathbf{R}_{Z-1} \dots \mathbf{R}_1$‘.

G.2 Proofs Error Transfer Matrices

The error transfer function of $\mathbf{X}_2^{\mathrm{T}} = [\mathbf{X}^{\mathrm{T}}|\mathbf{X}^{\mathrm{T}}]$ is

$$\begin{aligned}
\mathbf{R}_{x_2} &= \sum_{i=1}^{P}\sum_{j=1}^{2L_r}\left[[\frac{1}{2}\mathbf{X}^{+}|\frac{1}{2}\mathbf{X}^{+}]\psi_{2L_r,j}\right]\left[[\Psi_{x,i,j}|\Psi_{x,i,j}]\begin{bmatrix}\mathbf{S}_{x,i}\\ \overline{\mathbf{S}_{x,i}}\end{bmatrix}\mathbf{h}\right]^{\mathrm{T}} \\
&= \sum_{j=1}^{L_r}\sum_{i=1}^{P}\left[\frac{1}{2}\mathbf{X}^{+}\psi_{L_r,j}\right]\left[\Psi_{x,i,j}\mathbf{S}_{x,i}\mathbf{h}\right]^{\mathrm{T}} \\
&+ \sum_{j=L_r+1}^{2L_r}\sum_{i=1}^{P}\left[\frac{1}{2}\mathbf{X}^{+}\psi_{L_r,j}\right]\left[\Psi_{x,i,j}\mathbf{S}_{x,i}\mathbf{h}\right]^{\mathrm{T}} \\
&= \mathbf{R}_x. \qquad \text{(G.3)}
\end{aligned}$$

Similarly, the error transfer function of $\mathbf{H}_2^{\mathrm{T}} = [\mathbf{H}^{\mathrm{T}}|\mathbf{H}^{\mathrm{T}}]$ is also derived as $\mathbf{R}_{h_2} = \mathbf{R}_h$.

Bibliography

[1] K. Kim, C. M. Spooner, I. Akbar, and J. H. Reed, "Specific emitter identification for cognitive radio with application to IEEE 802.11," in *Proc. IEEE Global Telecommunications Conference*, New Orleans, LO, December 2008, pp. 1–5.

[2] C.-S. Shieh and C.-T. Lin, "A vector neural network for emitter identification," *IEEE Trans. Antennas Propag.*, vol. 50, no. 8, pp. 1120–1127, August 2002.

[3] L. E. Langley, "Specific emitter identification (SEI) and classical parameter fusion technology," in *Proc. IEEE Western Electronics Show and Conference*, San Francisco, CA, September 1993, pp. 377–381.

[4] K. A. Remley, C. A. Grosvenor, R. T. Johnk, D. R. Novotny, P. D. Hale, M. D. McKinley, A. Karygiannis, and E. Antonakakis, "Electromagnetic signatures of WLAN cards and network security," in *Proc. IEEE Signal Processing and Information Technology*, Athens, December 2005, pp. 484–488.

[5] A. Kawalec and R. Owczarek, "Specific emitter identification using intrapulse data," in *Proc. European Radar Conference*, Amsterdam, The Netherlands, October 2004, pp. 249–252.

[6] K. B. Rasmussen and S. Capkun, "Implications of radio fingerprinting on the security of sensor networks," in *Proc. Security and Privacy in Communications Networks and the Workshops*, Nice, France, September 2007, pp. 331–340.

[7] T.-W. Chen, W.-D. Jin, and J. Li, "Feature extraction using surrounding-line integral bispectrum for radar emitter signal," in *IEEE International Joint Conference on Neural Networks*, June 2008, pp. 294–298.

[8] H. K. Mardia, "New techniques for the deinterleaving of repetitive sequences," *IEE Proceedings F*, vol. 136, no. 4, pp. 149–154, August 1989.

[9] D. J. Milojevic and B. M. Popovic, "Improved algorithm for the deinterleaving of radar pulses," *IEE Proceedings F*, vol. 139, no. 1, pp. 98–104, February 1992.

[10] P. Ray, "A novel pulse TOA analysis technique for radar identification," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 34, no. 3, pp. 716–721, July 1998.

[11] F. Neri, *Introduction to Electronic Defense Systems*. Norwood, MA: Artech House, INC., 2001.

[12] K. Nishiguchi and M. Kobayashi, "Improved algorithm for estimating pulse repetition intervals," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, no. 2, pp. 407–421, April 2000.

[13] C. L. Davies and P. Hollands, "Automatic processing for esm," *IEE Proceedings F*, vol. 129, no. 3, pp. 164–171, June 1982.

[14] S. D'agostino, G. Foglia, and D. Pistoia, "Specific emitter identificaiton: Analysis on real radar signal data," in *Proceedings of the 6th European Radar Conference*, 2009, pp. 242–245.

[15] D. B. Faria and D. R. Cheriton, "Detecting identify-based attacks in wireless networks using signalprints," in *ACM WiSe*, 2006, pp. 43–52.

[16] N. Patwari and S. K. Kasera, "Robust location distinction using temporal link signatures," in *ACM Mobicom*, 2007, pp. 111–122.

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

[18] O. Ureten and N. Serinken, "Wireless security through RF fingerprinting," *The Canadian Journal of Electrical and Computer Engineering*, vol. 32, no. 1, pp. 27–33, Winter 2007.

[19] T. Kohno, A. Broido, and K. Claffy, "Remote physical device fingerprinting," *IEEE Trans. Dependable Secure Comput.*, vol. 2, no. 2, pp. 93–108, April-June 2005.

[20] V. Brik, S. Banerjee, M. Gruteser, and S. Oh, "Wireless device identification with radiometric signatures," in *ACM International Conference on Mobile Computing and Networking*, San Francisco, CA, 2008, pp. 116–127.

[21] A. Candore, O. Kocabas, and F. Koushanfar, "Robust stable radiometric fingerprinting for wireless devices," in *IEEE International Workshop on Hardware-Oriented Security and Trust*, 2009.

[22] *IEEE Std 802.11a-1999: Wireless LAN MAC and PHY Specification–High-Speed Physical Layer in the 5GHz Band*. IEEE, 2000.

[23] M.-W. Liu and J. F. Doherty, "Specific emitter identification using nonlinearity device estimation," in *Proc. IEEE 2008 Sarnoff Symposium*, Princeton, NJ, April 2008, pp. 1–5.

[24] C. van den Bos, M. H. L. Kouwenhoven, and W. A. Serdijn, "Effect of smooth nonlinear distortion on ofdm symbol error rate," *IEEE Trans. Commun.*, vol. 49, no. 9, pp. 1510–1514, September 2001.

[25] P. Banelli, "Theoretical analysis and performance of ofdm signals in nonlinear fading channels," *IEEE Trans. Wireless Commun.*, vol. 2, no. 2, pp. 284–293, March 2003.

[26] M.-W. Liu and J. F. Doherty, "Nonlinearity estimation for specific emitter identification in multipath environment," in *Proc. IEEE 2009 Sarnoff Symposium*, Princeton, NJ, March 2009, pp. 1–5.

[27] ——, "Wireless device identification in MIMO channels," in *Proc. IEEE Annual Conference on Information Sciences and Systems*, Baltimore, MD, March 2009, pp. 563–567.

[28] ——, "Specific emitter identification in multipath channels," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 1120–1127, September 2011.

[29] R. G. Wiley, *ELINT: The Interception and Analysis of Radar Signals*. Norwood, MA: Artech House, INC., 2006.

[30] R. Schmidt, "Multiple emitter location and signal parameter estimation," *Antennas and Propagation, IEEE Transactions on*, vol. 34, no. 3, pp. 276 – 280, March 1986.

[31] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 984–995, July 1989.

[32] A. Paulraj, R. Roy, and T. Kailath, "A subspace rotation approach to signal parameter estimation," *Proceedings of the IEEE*, vol. 74, no. 7, pp. 1044–1046, 1986.

[33] G. Xu and H. Liu, "An effective transmission beamforming scheme for frequency-division-duplex digital wireless communication systems," in *Acoustics, Speech, and Signal Processing, 1995. ICASSP-95., 1995 International Conference on*, vol. 3, May 1995, pp. 1729 –1732.

[34] A. Goldsmith, *wireless Communications*. New York: Cambridge, 2005.

[35] P. B. Kenington, *High-Linearity RF* Amplifier Design. Boston: Artech House, 2000.

[36] R. E. Ziemer and W. H. Tranter, *Principles of Communications*. New York: Wiley, 1995.

[37] M. C. Jeruchim, *Simulation of communication systems [electronic resource] : modeling, methodology, and techniques*. New York: Kluwer Academic/Plenum, 2000.

[38] K. G. Gard, L. E. Larson, and M. B. Steer, "The impact of RF front-end characteristics on the spectral regrowth of communications signals," *IEEE Trans. Microw. Theory Tech.*, vol. 53, no. 6, pp. 2179–2186, June 2005.

[39] I. S. Reed, "On a moment theorem for complex gaussian proccesses," *IEEE Trans. Inf. Theory*, vol. IT-8, pp. 194–195, April 1962.

[40] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*. New York: Wiley, 1996.

[41] J. Thomson, B. Baas, E. M. Cooper, J. M. Gilbert, G. Hsieh, P. Husted, A. Lokanathan, J. S. Kuskin, D. McCracken, B. McFarland, T. H. Meng, D. Nakahira, S. Ng, M. Rattehalli, J. L. Smith, R. Subramanian, Y.-H. W. L. Thon, R. Yu, and X. Zhang, "An integrated 802.11a baseband and MAC processor," in *IEEE International Solid-State Circuits Conference*, vol. 1, February 2002, p. 126.

[42] L. Xiao, L. Greenstein, N. Mandayam, and W. Trappe, "Channel-based detection of sybil attacks in wireless networks," *IEEE Trans. Inf. Forensics Security*, vol. 4, no. 3, pp. 492–503, September 2009.

[43] P. Yu, J. Baras, and B. Sadler, "Physical-layer authentication," *IEEE Trans. Inf. Forensics Security*, vol. 3, no. 1, pp. 38–51, March 2008.

[44] S. Haykin, *Adaptive Filter Theory*. New Jersey: Prentice Hall, 2002.

[45] D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication system," *IEEE Trans. Commun.*, vol. 28, no. 11, pp. 1867–1875, November 1980.

[46] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Trans. Inf. Theory*, vol. 40, no. 2, pp. 340–349, March 1994.

[47] J. Yang, J.-J. Werner, and G. A. Dumont, "The multimodulus blind equalization and its generalized algorithms," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 6, pp. 997–1015, June 2002.

[48] Y. Sato, "A method of self-recovering equalization for multilevel amplitude-modulation systems," *IEEE Trans. Commun.*, vol. 23, pp. 679–682, June 1975.

[49] G. Strang, *Linear Algebra and Its Applications*. California: Thomson Higher Education, 2006.

[50] G. H. Golub and C. F. V. Loan, *Matrix Computations*. Baltimore and London: John Hopkins, 1996.

[51] H. Hashemi, "The indoor radio propagation channel," in *Proceeding of the IEEE*, vol. 81, no. 7, July 1993, pp. 943–968.

[52] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. New York: Wiley-Interscience, 2000.

[53] P. Garg, K. R. Mallik, and H. M. Gupta, "Performance analysis of space-time coding with imperfect channel estimation," *IEEE Trans. Wireless Commun.*, vol. 4, no. 1, pp. 257–265, January 2005.

[54] V. Tarokh, A. Naguib, N. Seshadri, and A. R. Calderbank, "Space-time codies for high data rate wireless communication: Performance criteria in the presence of channel estimation errors, mobility, and multiple paths," *IEEE Trans. Commun.*, vol. 47, no. 2, pp. 199–207, February 1999.

[55] B. Vucetic and J. Yuan, *Space-Time Coding*. London: Wiley-Interscience, 2003.

[56] B. Razavi, *Design of analog* CMOS Integrated circuits. New York: McGraw Hill, 2003.

[57] T. C. W. Schenk, P. F. M. Smulders, and E. R. Fledderus, "Impact of nonlinearities in multiple-antenna OFDM transceivers," in *IEEE 2006 Symposium on Communications and Vehicular Technology*, Liege, November 2006, pp. 53–56.

[58] E. Costa, M. M. Midrio, and S. Pupolin, "Impact of amplifier nonlinearities on OFDM tranmission system performance," *IEEE Commun. Lett.*, vol. 3, pp. 37–39, February 1999.

[59] *IEEE Std 802.16e-2005: Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems*, 2005.

[60] M. Sanchez-Fernandez, M. de Prado-Cumplido, J. Arenas-Garcia, and F. Perez-Cruz, "SVM multiregression for nonlinear channel estimation in multiple-input multiple-output systems," *IEEE Trans. Signal Process.*, vol. 52, no. 8, pp. 2298–2307, August 2004.

[61] S. Chen, G. J. Gibson, C. F. N. Cowan, and P. M. Grant, "Adaptive equalization of finite nonlinear channels using multilayer perceptrons," *EURASIP Signal Process.*, vol. 10, pp. 107–119, 1990.

[62] F. Gregorio, S. Werner, T. I. Laakso, and J. Cousseau, "Receiver cancellation technique for nonlinear power amplifier distortion in SDMA-OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 56, no. 5, pp. 2499–2516, September 2007.

[63] Y. Palaska, A. Ravi, R. Pellerano, B. R. Carlton, M. A. Elmala, R. Bishop, G. Banerjee, R. B. Nicholls, S. K. Ling, N. Dinur, S. S. Taylor, and K. Soumyanath, "A 5-GHz 108-Mb/s 2 $\times$ 2 MIMO transceiver RFIC with fully integrated 20.5-dbm P_{1dB} power amplifiers in 90-nm cmos," *IEEE J. Solid-State Circuits*, vol. 41, no. 12, pp. 2746–2756, December 2006.

[64] D. G. Rahn, M. S. Cavin, F. F. Dai, N. H. W. Fong, R. Griffith, J. Macedo, A. D. Moore, J. W. M. Rogers, and M. Toner, "A fully integrated multiband MIMO WLAN transceiver RFIC," *IEEE J. Solid-State Circuits*, vol. 40, no. 8, pp. 1629–1641, August 2005.

[65] K. I. Pedersen, B. Andersen, J. P. Kermoal, and P. Mogensen, "A stocastic multiple-input-multiple-output radio channel model for evaluation of space-time coding algorithms," in *Vehicular Technology Conference*, vol. 2, Boston, September 2000, pp. 893–897.

[66] K. Yu, M. Bengtsson, B. Ottersten, D. McNamara, P. Karlsson, and M. Beach, "Modeling of Wide-Band MIMO radio channels based on NLoS indoor measurements," *IEEE Trans. Veh. Technol.*, vol. 53, no. 3, pp. 655–665, May 2004.

[67] D. Gesbert, M. Shafi, D.-S. Shiu, P. J. Smith, and A. Naguib, "From theory to practice: An overview of MIMO space-time coded wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 3, pp. 281–302, April 2003.

[68] S. Glisic, *Advanced Wireless Communications: 4G Technologies*. West Sussex: John wiley, 2004.

[69] E. Perahia, "IEEE 802.11n development: History, process, and technology," *IEEE Commun. Mag.*, vol. 46, pp. 48–55, July 2008.

[70] *Joint Proposal: High Throughput extension to the 802.11 Standard:* PHY. IEEE 802.11-05/1102r4, 2006.

[71] *WiMAX Forum Mobile Release 1.0 Channel Model*. WiMAX Forum, 2008.

[72] J. P. Kermoal, L. Schumacher, K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 6, pp. 1211–2002, August 2002.

[73] R. A. Poise, *Electronic Warfare Target Location Methods*. Norwood, MA: Artech House, INC., 2005.

[74] J. Andersen, T. Rappaport, and S. Yoshida, "Propagation measurements and models for wireless communications channels," *IEEE Commun. Mag.*, vol. 33, no. 1, pp. 42–49, January 1995.

[75] W. Lee and D. Lee, "Microcell prediction in dense urban area," *IEEE Trans. Veh. Technol.*, vol. 47, no. 1, pp. 246–253, February 1998.

[76] D. Har, H. Xia, and H. Bertoni, "Path-loss prediction model for microcells," *IEEE Trans. Veh. Technol.*, vol. 48, no. 5, pp. 1453–1462, September 1999.

[77] H. Bertoni, W. Honcharenko, L. Macel, and H. Xia, "UHF propagation prediction for wireless personal communications," *Proceedings of the IEEE*, vol. 82, no. 9, pp. 1333 –1359, September 1994.

[78] F. Ikegami, T. Takeuchi, and S. Yoshida, "Theoretical prediction of mean field strength for urban mobile radio," *IEEE Trans. Antennas Propag.*, vol. 39, no. 3, pp. 299 –302, March 1991.

[79] Y. Nagata, Y. Furuya, E. Moriyama, M. Mizuno, I. Kamiya, and S. Hattori, "Measurement and modeling of 2GHz-band out-of-sight radio propagation characteristics under microcellular environments," in *In Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications.*, September 1991, pp. 341–346.

[80] K. Rizk, J.-F. Wagen, and F. Gardiol, "Two-dimensional ray-tracing modeling for propagation prediction in microcellular environments," *IEEE Trans. Veh. Technol.*, vol. 46, no. 2, pp. 508 –518, May 1997.

[81] M. Hata, "Empirical formula for propagation loss in land mobile radio services," *IEEE Trans. Veh. Technol.*, vol. 29, no. 3, pp. 317–325, August 1980.

[82] J. Wiart, C. Dale, A. Bosisio, and A. Le Cornec, "Analysis of the influence of the power control and discontinuous transmission on RF exposure with GSM mobile phones," *IEEE Trans. Electromagn. Compat.*, vol. 42, no. 4, pp. 376–385, November 2000.

[83] *IEEE Standard for Local and Metropolitan Area Networks*. IEEE, 2005.

[84] D. V. Hinkley, "On the ratio of two correlated normal random variables," *Biometrika*, vol. 56, pp. 635–639, 1969.

[85] G. Marsaglia, "Ratios of normal variables and ratios of sums of uniform variables," *American Statistical Association Journal*, pp. 193–204, March 1965.

[86] J. Hayya, D. Armstrong, and N. Gressis, "A note on the ratio of two normally distributed variables," *Management Science*, vol. 21, no. 11, pp. 1338–1341, July 1975.

[87] H. W. Lilliefors, "On the Kolmogorov-Smirnov test for normality with mean and variance unknown," *Journal of the American Statistical Association*, no. 62, pp. 399–402, June 1967.

[88] R. Harris, D. Chabries, and F. Bishop, "A variable step (VS) adaptive filter algorithm," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 34, no. 2, pp. 309–316, April 1986.

[89] R. Kwong and E. Johnston, "A variable step size LMS algorithm," *Signal Processing, IEEE Transactions on*, vol. 40, no. 7, pp. 1633–1642, July 1992.

[90] T. Aboulnasr and K. Mayyas, "A robust variable step-size LMS-type algorithm: analysis and simulations," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 631–639, March 1997.

[91] S. Douglas, "A family of normalized LMS algorithms," *IEEE Signal Process. Lett.*, vol. 1, no. 3, pp. 49–51, March 1994.

[92] M. Tarrab and A. Feuer, "Convergence and performance analysis of the normalized LMS algorithm with uncorrelated gaussian data," *IEEE Trans. Inf. Theory*, vol. 34, no. 4, pp. 680–691, July 1988.

[93] M. Win and J. Winters, "Analysis of hybrid selection/maximal -ratio combining in rayleigh fading," vol. 1, 1999, pp. 6–10.

[94] A. Shah and A. Haimovich, "Performance analysis of maximal ratio combining and comparison with optimum combining for mobile radio communications with cochannel interference," *IEEE Trans. Veh. Technol.*, vol. 49, no. 4, pp. 1454–1463, July 2000.

Vita

Ming-Wei Liu

EDUCATION

Ph.D. in Electrical Engineering, Penn State University	2011
M.S. in Electronics Engineering, National Chiao Tung University, Taiwan	2004
B.E. in Electronics Engineering, National Chiao Tung University, Taiwan	2002

PROFESSIONAL WORK EXPERIENCE

Alcatel-Lucent Bell Labs	Summer 2007
EE Dept. at Penn State University	2006-2011
EE Dept. at National Chiao Tung University	2004-2005

SELECT PUBLICATIONS

1. **M.-W. Liu** and J. F. Doherty, "Specific emitter identification in multipath channels," IEEE Trans. Inf. Forensics Security, vol. 6, no. 3, September 2011.

2. **M.-W. Liu** and J. F. Doherty, "Frequency-selective multiple-input multiple-output channel estimation with transmitter non-linearities," *Signal Processing, IET*, vol. 3, no. 6, pp. 467-475, 2009.

3. **M.-W. Liu** and J. F. Doherty. "Joint specific emitter identification and location tracking for OFDM systems." In *Symposium and Summer School on Wireless Comm.*, Virginia Tech., June 2010.

4. **M.-W. Liu** and J. F. Doherty, "Nonlinearity estimation for specific emitter identification in multipath environment," in *IEEE 2009 Sarnoff Symposium*, March 2009.

5. **M.-W. Liu** and J. F. Doherty, "Wireless device identification in MIMO channels," in *IEEE Conference on Information Sciences and Systems 2009* , March 2009.

6. **M.-W. Liu** and J. F. Doherty, "Specific emitter identification using nonlinearity device estimation," in *IEEE 2008 Sarnoff Symposium, Princeton*, NJ, 28-30 April 2008.