Energy-Efficient Distributed Estimation by Utilizing a Nonlinear Amplifier

by,

Robert Santucci

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved November 2013 by the
Graduate Supervisory Committee:

Andreas Spanias, Chair
Cihan Tepedelenlioğlu
Bertan Bakkaloglu
Kostas Tsakalis

ARIZONA STATE UNIVERITY

December 2013

ABSTRACT

Distributed estimation uses many inexpensive sensors to compose an accurate estimate of a given parameter. It is frequently implemented using wireless sensor networks. There have been several studies on optimizing power allocation in wireless sensor networks used for distributed estimation, the vast majority of which assume linear radio-frequency amplifiers. Linear amplifiers are inherently inefficient, so in this dissertation nonlinear amplifiers are examined to gain efficiency while operating distributed sensor networks.

This research presents a method to boost efficiency by operating the amplifiers in the nonlinear region of operation. Operating amplifiers nonlinearly presents new challenges. First, nonlinear amplifier characteristics change across manufacturing process variation, temperature, operating voltage, and aging. Secondly, the equations conventionally used for estimators and performance expectations in linear amplify-and-forward systems fail. To compensate for the first challenge, predistortion is utilized not to linearize amplifiers but rather to force them to fit a common nonlinear limiting amplifier model close to the inherent amplifier performance. This minimizes the power impact and the training requirements for predistortion. Second, new estimators are required that account for transmitter nonlinearity. This research derives analytically and confirms via simulation new estimators and performance expectation equations for use in nonlinear distributed estimation.

An additional complication when operating nonlinear amplifiers in a wireless environment is the influence of varied and potentially unknown channel gains. The impact of these varied gains and both measurement and channel noise sources on estimation performance are analyzed in this paper. Techniques for minimizing the estimate variance are developed. It is shown that optimizing transmitter power allocation to minimize estimate variance for the most-compressed parameter measurement is equivalent to the problem for linear sensors. Finally, a method for operating distributed estimation in a multipath environment is presented that is capable of developing robust estimates for a wide range of Rician *K*-factors.

This dissertation demonstrates that implementing distributed estimation using nonlinear sensors can boost system efficiency and is compatible with existing techniques from the literature for boosting efficiency at the system level via sensor power allocation. Nonlinear transmitters work best when channel gains are known and channel noise and receiver noise levels are low.

To my friends and family,
who have understandingly made sacrifices
and pushed me on,
enabling me to complete my degree.

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the work, encouragement, and sacrifices of many others. First off, I would like to thank my advisor Dr. Andreas Spanias and post-doctoral research associate Dr. Mahesh Banavar. Not only has their keen insight pointed me in the right direction when I have become stuck while solving specific problems in the course of completing this dissertation, but they have gone above and beyond their job requirements to make themselves available off campus, on weekends, and during evenings around my work schedule. They have helped me navigate through the process of completing a dissertation. I would also like to thank my committee, including Dr. Cihan Tepedelenlioğlu, Dr. Bertan Bakkaloglu, and Dr. Kostas Tsakalis for their advice on and pointers. I also have to thank my undergraduate professors at Embry-Riddle including Dr. Matt Jaffe for the fundamentals that have enabled me to succeed in graduate school.

While my graduate school professors have been extremely helpful, thanks are also due to my managers from Intel who have both paid my tuition from their budgets and allowed me to modify my work schedule to attend graduate classes during the daytime. I especially have to thank Bart McDaniel and Dr. Waleed Khalil who supported me when I began this journey. They have also provided me the opportunities to apply what I have learned at Arizona State University to industry projects. I also would like to thank Terry and Bill Wilhoit and the ARCS Foundation whose scholarship helped me out when I was a full-time student.

While at Arizona State University, I have to thank the great community of students and professors with whom I have gotten to interact. I have taken numerous classes from both Dr. Tolga Duman and Dr. Antonia Papandreou. I have had many great discussions over the last few years with my now graduated classmates, Dr. Karthikeyan Ramamurthy and Dr. Jayaraman Thiagarajan, and fellow students Deepta Rajan and Girish Kalyanasundaram. I have to thank Dayna Peterson for helping me practice my defense presentation and providing helpful input.

A healthy rivalry with my sister Debbie to see who can finish their degree first has helped make sure I did finish, so thank you to her for the rivalry. Finally, I'd like to thank my parents Bill and Nancy. My parents and grandparents made sure my education started well before kindergarten with reading and math exercises when I was little. They provided needed encouragement when I was frustrated over trying to complete school work while missing out on other activities. Without their support and understanding, this would have been an even tougher effort.

TABLE OF CONTENTS

Page

Page

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

Figure Page

1 INTRODUCTION

Distributed estimation is the production of an estimate of one or more physical quantities based on the measurements of a set of sensors. These sensors are individually low-precision devices, allowing them to be used in any given scenario. By aggregating the output of many sensors, it is possible to produce a detailed estimate of the quantities being measured over some region.

An area of active research is the implementation of distributed estimation systems using wireless sensor networks [1]. These sensors are often used to perform surveillance on locations that are physically difficult to access. For these scenarios, the wireless sensors should be inexpensive enough to be considered disposable. Additionally, the wireless sensors should be self-contained with their own power source, typically a battery. For battery-powered estimation systems, overall utility is determined by the system's ability to produce quality estimates and the battery lifetime over which quality estimates can be attained. Extending the battery life over which accurate estimates can be attained requires increasing the energy-efficiency of both the sensors and the algorithms used to produce distributed estimates.

In other applications, wireless sensors may be required by the application but the endurance is required to exceed the lifetime of any available batteries. Passive wireless sensors can be built that obtain their energy from a directional high-frequency field transmitted by a master, similar to operation of radio-frequency identification (RFID) tags [2]. Examples of systems with such requirements would be wireless sensors used to

monitor stress across a wind-turbine blade. The amount of power that can be coupled into the "passive" sensor is small, and thus both the sensor and the sensing algorithm must be energy-efficient to avoid long charging times and large charge-storage devices.

There are a variety of ways to transmit data from each sensor to the fusion center. Each individual sensor can transmit its data to the fusion center using an orthogonal signaling scheme, which makes each sensor's measurement independently discernible, but also requires a total system bandwidth that grows with the total number of sensors. This increasing bandwidth requirement limits the attainable sensor sampling rate. This may be acceptable if the number of sensors is small, but may not be applicable when the number of sensors is large. Several papers have proposed deactivating sensors [3] or truncating digitally transmitted values [4,5,6] to reduce the amount of transmitted data and thus boost energy-efficiency of the distributed estimation system.

A different technique is to use an estimation algorithm which does not need to attain the values of individual sensor measurements orthogonally, but instead operates on an aggregate received value. One estimation algorithm that can operate using a shared bandwidth is Amplify-and-forward (AF) [1,7,8]. Optimization has been performed to minimize the sum of transmitter powers subject to achieving some steady-state error power or to minimize the error power subject to some sum of transmitter powers [9].

An issue with the power optimizations proposed in the reference literature is that it has been assumed that the transmitter is linear with constant gain and that power consumption is proportional to the transmitted power. These assumptions do not hold

with conventional amplifiers, which are only linear when operating in the inefficient class-A region of operation and consume a large amount of battery power to maintain required biasing [10]. This dissertation investigates this tradeoff and develops a modified amplify-and-forward algorithm aimed at extending sensor battery life while minimizing impact to error power.

Another assumption commonly made with AF systems is that the transmitting sensors are phase-aligned. This dissertation also investigates that assumption and proposes an alternate system for implementing AF distributed estimation in a dynamic scattering environment.

1.1 Problem Statement

An assumption made in AF systems is that the transmitter is either inherently linear or has been linearized by a predistorter. Though linearity is desirable for accurate estimation by keeping the equations of [7] and [8] tractable, it significantly reduces the power efficiency of the transmitting sensors. Many papers have been published on maximizing the energy-efficiency of wireless sensor networks at the system-level by optimally allocating transmitted output power among the sensors [1,6,9,11] or by phase-modulation [12,13], but none of these examines the battery power required to generate the output. In this research, an algorithm is developed that uses the nonlinear region of amplifier operation in combination with a modified amplify-and-forward estimation to operate with higher power efficiency from the sensor's power source.

The key component in a transmitter that affects its efficiency is the power amplifier (PA) [14]. PA design involves a tradeoff between linearity and efficiency. When an amplifier is operating linearly as typically assumed by AF algorithms, the amplifier's efficiency is low. Efficiency improves as input signal power increases, but at the cost of introducing gain-compression, as shown in Figure 1-1. One common metric for amplifier efficiency is called power-added efficiency (PAE). PAE is defined as the ratio between signal power out and the sum of signal power and supply power into the amplifier:

$$\text{PAE} = \frac{P_{signalOut}}{P_{supplyIn}+P_{signalIn}} \tag{1-1}$$

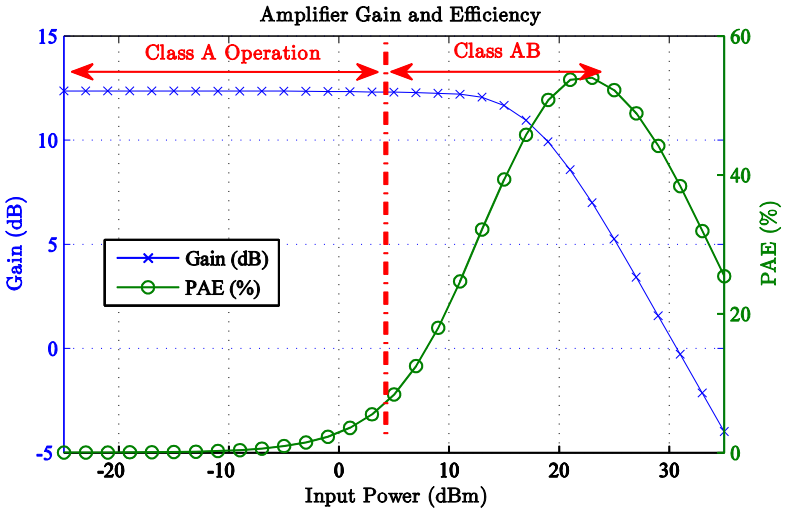


Figure 1-1: Class AB Amplifier. PAE peaks when gain compression is present, but the amplifier is not yet saturated.

When gain compression is uncorrected, the equations for linear AF from [7] and [8] fail. To add further implementation complexity, the exact gain-compression behavior can differ between sensors based on manufacturing process, supply voltage, and

temperature (PVT) variations. The gain-compression problem cannot be solved by assuming all sensors are identically nonlinear.

Linear AF operation in the gain compressed region can be accomplished by linearizing each amplifier individually. This linearization is done by adding a gain-expansive predistorter [15] preceding the amplifier in each transmitter. It is possible to operate linearly a few decibels into the class AB region of operation by using linearization as shown in Figure 1-2. However, the amplifier system still requires a linear dynamic range equal to the dynamic range of the noisy sensed measurements. The highly efficient operating region will only be utilized at the largest measurement values. Most measurements will still occur in the inefficient class A region.

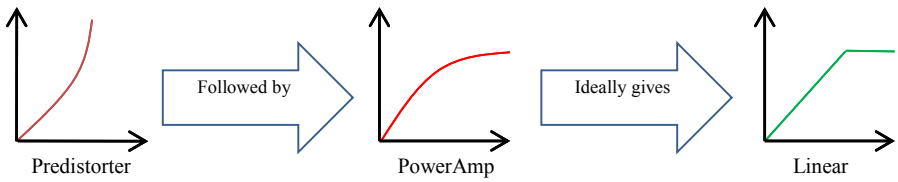


Figure 1-2: Predistortion (Linearization) System Gain.

The objective of this research is to develop methodologies to extend the battery life of sensors in AF distributed estimation systems by enabling the use of high-efficiency nonlinear amplifiers. For the proposed methodologies, performance expectations are developed and related to the conventional linear AF distributed estimation system. The performance of proposed distributed estimation systems is also investigated in the presence of unequal gain channels. In summary, the problems investigated by this dissertation are:

- Amplify-and-forward distributed estimation over a coherent multiple-access channel is a common technique used to estimate a measured parameter using a large number of inexpensive sensors.
- A lot of research has been published on predicting estimate variance attainable using amplify-and-forward distributed estimation systems and on optimal transmitted power allocation. However, the vast majority of this research assumes linear transmit amplifiers.
- Linear transmit amplifiers are less efficient than their nonlinear counterparts.
- Practical implementations of distributed estimation over coherent multiple-access networks suffer from the need to have accurate phase estimation.

1.2 Contributions

This research proposes a new approach for boosting efficiency of AF systems where receiver and channel noise are low relative to received power. Predistortion is utilized not to linearize an amplifier, but rather to force all sensor amplifiers to fit a common limiting-amplifier model. Modified estimators are designed that exploit the more efficient, nonlinear, class AB region of amplifier operation for a wider range of sensor measurements than the linear AF system allows. The performance of the proposed technique is analyzed in terms of asymptotic variance and power added efficiency.

Analytical backing of the proposed algorithms is produced using a hyperbolic tangent limiting amplifier model and uniform sensing noise. Modeling and simulation is

used to validate the algorithms with simulated data from a real amplifier for both perfect and more practical methods of digital predistortion. Significant improvements in power-added efficiency are demonstrated at the cost of small amounts of additional estimation variance. The modified estimator is proven to be effective in the presence of limited additive receiver noise, and sensitivities not present in the linear AF scheme are discussed. Imperfections in realizing predistortion are shown to have similar performance degradation. The performance of the proposed estimator is investigated for uniform and non-uniform additive white Gaussian noise channels. For operation in dynamic scattering environments, a new approach utilizing orthogonal-frequency division multiple access is been proposed.

Numeric simulations have been conducted to verify the performance of the proposed distributed estimation system utilizing the less tractable Cann limiting amplifier model [16], and using Gaussian and Cauchy sensing noise sources. Consistent estimators are demonstrated for Cauchy sensing noise, which are not possible using the linear AF system. Additionally, numeric solutions have been utilized to solve analytically posed optimization problems for sensor powers in a wireless network with widely varied channel gains.

As a part of the implementation of the proposed distributed estimation system for scattering environments, this work proposes two developments of digital predistortion that improve performance under specific conditions. The first development is an improvement to help reduce the convergence time of gain-based look-up-table

incorporating memory-effect compensation predistortion. This improvement has been done by including a binned block least-mean-squares (LMS) gradient-descent algorithm to accelerate convergence by using more samples when there is long transmitter-receiver loop latency in the modem. The second improvement is done to reduce the number of required weights to be trained in neural network predistortion. This is done when the power amplifier can be nearly modeled by a memory polynomial, neglecting some memory effects. The memory polynomial terms are all input to the neural network. The neural network is then already nearly fit to the amplifier at the start of training, thus leaving the existing neurons free to be trained for matching more complex interactions not already incorporated by the memory polynomial.

In summary, this dissertation provides the following contributions to the literature:

- Proposes a new method for performing distributed estimation over a coherent multiple access channel using nonlinear amplifiers. Nonlinear amplifiers are capable of provided better power efficiency than linear amplifiers.
- Provides Analytical derivations for both the estimator and the estimate variance attainable using nonlinear distributed estimation system in both identical and non-identical gain additive white Gaussian noise channels, for both known and unknown channel gains.

- Provides a technique for optimizing sensor power allocation for nonlinear distributed estimation systems. Simulated data demonstrates that this in most scenarios optimization yields one of two two-complexity sensor power allocation strategies.
- Proposes a technique for using orthogonal-frequency division multiple-access to allow real-time changeover between coherent multiple-access channels and orthogonal distributed estimation.
- Proposes predistortion techniques for use in OFDMA systems to improve performance with nonlinear amplifiers.

1.3 Dissertation Organization

In the second chapter of this dissertation, a literature review of both amplify-and-forward distributed estimation and of predistortion is presented. Refinements developed for digital predistortion are presented in chapter three. In chapter four, an energy-efficient form of distributed estimation is presented in the case of either uniform gain or equalized additive white Gaussian noise channels. In the fifth chapter, the proposed technique is extended to non-uniform gain channels, including the proposal of an orthogonal frequency division multiple-access technique for use in dynamic scattering environment. Finally, in the last chapter conclusions and opportunities for future work are discussed.

2 LITERATURE REVIEW

2.1 Literature Review of Distributed Estimation

In distributed estimation, noisy measurements from multiple sensors are processed in order to estimate the value of a set of parameters. The measurements of these sensors are then jointly processed in some manner to form an estimate of the single quantity they all measure. Each sensor must be self-contained, with its own power source and transmitter. For the cases examined in this research, data is transmitted from the sensor to a fusion center which processes the measurements of each sensor into a single estimate. Other techniques are available for doing estimation using ad-hoc wireless networks [17]. The sensors themselves are small and have generally low data processing capability. Current trade publications are discussing proposed systems for implementing wireless sensor networks using sensors that do not even utilize a proper battery, deriving all required power from a radio transmitter in a manner similar to an radio-frequency identification (RFID) tag [2]. A typical generic system for doing distributed estimation using a fusion center is shown in Figure 2-1.

One of the key decisions is how to transmit data to the fusion center. Data transmission can be scheduled so that each sensor transmits its signal orthogonally from the other sensors or simultaneously over a multiple-access channel [5]. Data can be transmitted from the sensors either digitally [6] or in an analog format [1]. The scheduling and format by which sensor measurements are transmitted to the fusion center have a significant effect on the implementation required to boost the energy-efficiency of

the overall system. When designing a distributed estimation system one of two approaches is generally taken, one can design either to minimize the estimate variance for a given power constraint or one can design to minimize the power consumed in meeting a variance constraint. Power constraints can be either a system-level total power constraint (TPC) or a per-sensor individual power constraint (IPC) [18].

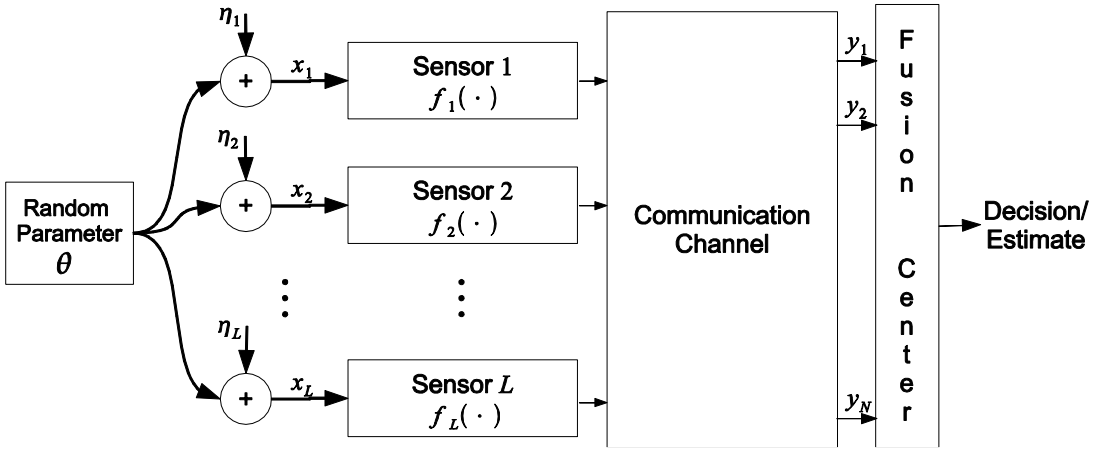


Figure 2-1: Generic distributed estimation system using a fusion center, modified from [8]

This literature review of distributed estimation is divided into three parts. In section 2.1.1, basic amplify-and-forward concepts are reviewed. In section 2.1.2, algorithms for energy-efficient estimation using orthogonal channels are reviewed. In section 2.1.3, algorithms for energy-efficient estimation when using coherent multiple-access channels (MAC) are reviewed.

2.1.1 Amplify-and-Forward Distributed Estimation

Consider a set of L sensors, each with its own power supply transmitting over a Gaussian multiple-access channel [19] to a fusion center as shown in Figure 2-2:

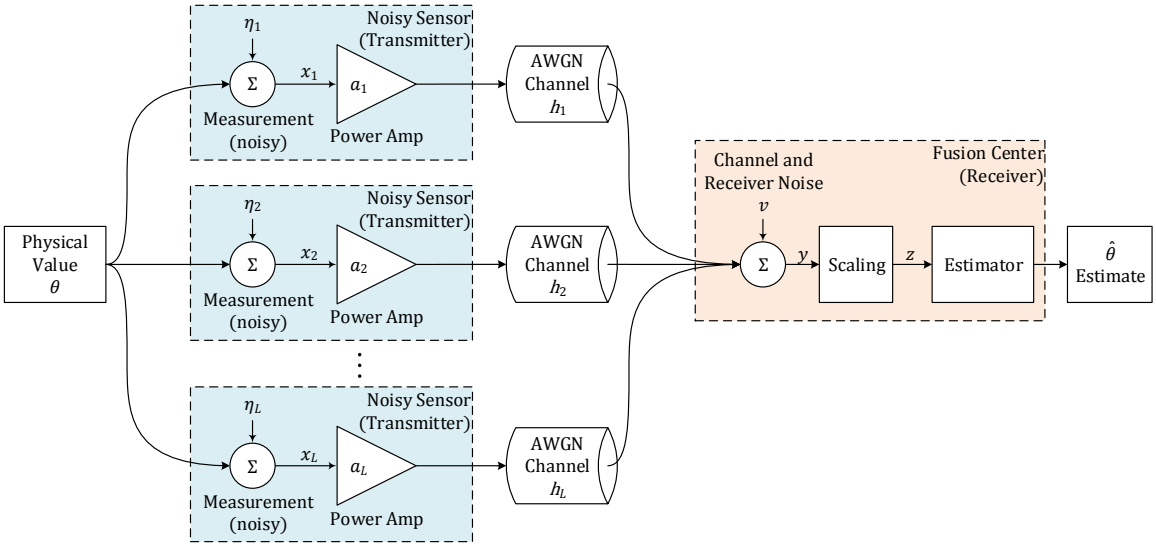


Figure 2-2: Distributed Estimation System Topology.

For simplicity, the channels between the sensors and the fusion center are assumed to be additive white Gaussian noise (AWGN) channels. Each of the sensors is measuring $\theta \in \mathbb{R}^+$, in some additive noise η_l. The measurement, x_l, including observed noise at each sensor is:

$$x_l = \theta + \eta_l. \tag{2-1}$$

For the reference techniques presented in [7] and [8], the value transmitted by the sensor l is a perfectly scaled version of the noisy sensor measurement $a_l\, x_l$, where $a_l \in \mathbb{R}^+$ is a hardware determined constant determining the largest gain factor of the input signal range that can be applied while keeping the amplifier linear. Initially, it is assumed that all sensors have the same gain $a_l = a$, and all channels have the same gain $h_l = 1$. The received signal, y, is a sum of the transmitted signals from each sensor, subject to some receiver noise, v:

$$y = \sum_{l=1}^{L} ax_l + v. \tag{2-2}$$

At the receiver, this is scaled by the number of sensors to provide:

$$z = \frac{1}{L} y = \frac{1}{L} \sum_{l=1}^{L} ax_l + \frac{v}{L} \, . \tag{2-3}$$

Evaluating the limit of z as $L \to \infty$ gives the asymptotic approximation, $\zeta(\theta)$, of the value received z as the number of sensors becomes large:

$$\zeta(\theta) = \lim_{L \to \infty} z = \lim_{L \to \infty} \left(\frac{1}{L} \sum_{l=1}^{L} ax_l \right). \tag{2-4}$$

Note that equations (2-2) through (2-4) add the baseband equivalent signals and the resulting output signal produces a coherent sum of the baseband signals. This statement makes the assumption that the transmitters have all been phase synchronized including accounting for their channel state information (CSI). Making this assumption implies that other than distributed estimation signals, there is a channel estimation system running prior to and during the measurement estimation process and for that estimation to work that the channels and signals are slowly varying [9]. Examples of synchronization systems are shown in [20] and [21]. It is demonstrated in [20] and [22] that even with some phase error in the synchronization, reasonable estimates can be developed.

In (2-4), the receiver noise v is ignored due to the assumption that as the number of sensors transmitting, L, becomes large, the received power from the sensors becomes much greater than the fusion center's receiver noise floor. By applying the central limit

theorem, it can be seen that $\zeta(\theta)$ in (2-4) approximates a Gaussian random variable with mean $\bar{\zeta}(\theta)$ and variance $\Sigma(\theta)$, when $\bar{\zeta}(\theta)$ and $\Sigma(\theta)$ exist [23].

The asymptotic mean of received value $\zeta(\theta)$ as $L \to \infty$ is subject to variation induced by the random sensor measurement noise. It is useful to develop an estimate of the expected value received given sensor noise distributions $p(\eta_1)p(\eta_2)\cdots p(\eta_l) = p(\boldsymbol{\eta})$:

$$\bar{\zeta}(\theta) = E[\zeta(\theta)] = a\int_{\boldsymbol{\eta}} p(\boldsymbol{\eta})\left(\theta + \frac{1}{L}\sum_{l=1}^{L}\eta_l\right)d\boldsymbol{\eta}. \tag{2-5}$$

The sum term from (2-5) can be simplified using the central limit theorem [24], provided each sensor noise has similar distribution and is of zero mean and defined variance, σ_η^2, to obtain $\bar{\zeta}(\theta)$, the asymptotic mean of the received value for a given actual parameter θ:

$$\bar{\zeta}(\theta) = E[\zeta(\theta)] = a\theta. \tag{2-6}$$

Although the estimator with uniform sensor noise, perfect CSI, and linear amplifiers can be found using a conventional means, this research takes the more general approach to developing the estimator for smallest variance for a received signal in Gaussian noise [23]. The estimator with smallest variance is found by solving the minimization problem in (2-7) utilizing the actual observed received value, z, assuming small receiver noise, $L \to \infty$, and the expected asymptotic mean and variance of received signal z, given by $\bar{\zeta}(\theta)$ and $\Sigma(\theta)$ respectively [23,25]:

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}}\left(\left(z - \bar{\zeta}(\theta)\right)\Sigma^{-1}(\theta)\left(z - \bar{\zeta}(\theta)\right)\right). \tag{2-7}$$

The variance of the received value, $\Sigma(\theta)$, asymptotically converges to:

$$\Sigma(\theta) = a^2\sigma_\eta^2. \tag{2-8}$$

Using (2-8), the $\Sigma(\theta)$ can be evaluated for either the Gaussian or uniform distributed sensing noise by substituting the value of variance, σ_η^2, appropriate for that distribution. This yields the values:

$$\begin{aligned} &\text{If } \eta_l \sim \mathcal{N}(0,\sigma^2) \rightarrow \Sigma(\theta) = a^2\sigma_\eta^2 \\ &\text{If } \eta_l \sim U(-b,b) \rightarrow \Sigma(\theta) = a^2b^2/3. \end{aligned} \tag{2-9}$$

Solving the equation (2-7) utilizing either the Gaussian or uniform distribution yields an identical result:

$$\hat{\theta} = \frac{z}{a}\ . \tag{2-10}$$

From [23,25], the asymptotic variance of the optimal estimate (2-10) can be calculated as:

$$\text{AsVar}\left(\hat{\theta}\right) = \frac{\Sigma(\theta)}{\left[\frac{\partial}{\partial\theta}\bar{\zeta}(\theta)\right]^2}\ , \tag{2-11}$$

resulting in:

$$\begin{aligned} &\text{If } \eta_l \sim \mathcal{N}(0,\sigma^2) \rightarrow \text{AsVar}\left(\hat{\theta}\right) = \sigma^2 \\ &\text{If } \eta_l \sim U(-b,b) \rightarrow \text{AsVar}\left(\hat{\theta}\right) = b^2/3. \end{aligned} \tag{2-12}$$

From (2-12) it can be seen that in the absence of receiver noise and channel noise, if the sensor measurement noise is distributed in such a manner that the central limit theorem can be applied, the estimate has asymptotically the same variance as the quantity being measured.

Using Cauchy distributed sensor measurement noise yields an undefined mean and variance. Consequently the reference amplify-and-forward schemes do not work for sensors with Cauchy distributed noise. This is because the conventional amplify-and-forward technique implicitly computes the sample mean of the sensor observations via (2-3). For heavy-tailed distributions such as the Cauchy distribution, the sample mean does not produce a meaningful value.

In chapter 4, an algorithm is proposed that provides accurate estimates in the presence of nonlinear amplifiers. It also provides consistent estimates with Cauchy distributed sensing noise, while linear amplifiers do not.

2.1.2 Energy-Efficient Estimation: Orthogonal MAC Sensor Transmissions

Orthogonal transmission of sensor data to the fusion center places full burden of estimation on the fusion center. One of benefits of implementing orthogonal sensor transmissions is that it allows reusing existing wireless digital communications techniques such as coding to minimize the rate of bit errors occurring between the sensor transmitters and the fusion center. Thus, issues that must be solved for analog orthogonal or multiple-access channel implementations such as channel-state equalization and phase synchronization can be largely mitigated. Common orthogonal transmission schemes for distributed estimation include time-division multiple access (TDMA) [6], code-division multiple-access (CDMA), or frequency-division multiple-access (FDMA) [9].

In many papers including [5] and [11] orthogonal digital transmission is used and the total power transmitted by the sensors is minimized subject to an estimate variance constraint. The individual sensor powers are often controlled by adjusting the number of bits to which each sensor reading is quantized. To determine power, it is assumed each sensor transmits its reading as a single *M*-level quadrature amplitude modulation (QAM) constellation point. The number of bits of to which the l^{th} sensor is quantized is Q_l. Each sensor measures a value ranging from $-W$ to $+W$. The average power, P_l, to transmit a measurement of Q_l bits from the l^{th} sensor is [11]:

$$P_l = \frac{c_l}{h_l} B_s \ln\left(\frac{2}{p_b^l}\right)(2^{Q_l} - 1), \tag{2-13}$$

where h_l is the path or channel gain from the l^{th} sensor, p_b^l is the bit error probability at the l^{th} sensor, B_s is the sampling rate, and c_l is a constant determined in part by the receiver noise figure and the single-sided power spectral density of channel noise. The point of (2-13) is that transmitting the same information in a channel with high channel loss, also known as low channel gain, takes much more power than transmitting the same information in a low loss channel. The quantized sensors inject quantization noise into the estimate beyond the sensor measurement noise already present. The quantization noise power, δ_l^2, can be characterized by [26]:

$$\delta_l^2 = \frac{W^2}{(2^{Q_l} - 1)^2} \ . \tag{2-14}$$

From the quantization noise power and the sensor measurement noise power itself, and with a bounded number of bit errors, estimate variance $\mathrm{Var}(\hat{\theta})$ can be upper-bounded by [27]:

$$\mathrm{Var}(\hat{\theta}) \leq (1+p_0)^2 \left(\sum_{l=1}^{L} \frac{1}{\sigma_{\eta_l}^2 + \delta_l^2} \right), \tag{2-15}$$

where p_0 is a constant dependent upon bit error rate and sensor noise. Assuming all channels have similar noise characteristics c_l and sampling bandwidth B_s, a total sensor transmitted power minimization problem is setup in [27] to attain estimate variance of less than a desired target Υ using digital quantization levels of Q_l bits for the l^{th} sensor:

$$\min_{Q_l \in \mathbb{Z}} \sum_{l=1}^{L} \frac{(2^{Q_l}-1)^2}{h_l^2},$$

Subject to:

$$\mathrm{Var}(\hat{\theta}) = (1+p_0)^2 \left(\sum_{l=1}^{L} \frac{1}{\sigma_l^2 + \delta_l^2} \right) \leq \Upsilon \tag{2-16}$$

$$\delta_l^2 = \frac{W^2}{(2^{Q_l}-1)^2} \ , \ \ l = 1, \dots, L$$

$$Q_l \geq 0, \quad l = 1, \dots, L.$$

Solving this minimization yields optimal quantization levels of:

$$M_l^{opt} = \begin{cases} 0 & h_l \tau_0 \leq 1 \\ \log_2 \left(1 + \dfrac{W}{\sigma_{\eta_l}} \sqrt{\tau_0 h_l - 1} \right) & h_l \tau_0 > 1 \end{cases}, \tag{2-17}$$

where the threshold τ_0 is a function of the desired estimate variance, sensor noise variance, and the channel gains for activated sensors. Equation (2-17) indicates that

higher levels of quantization, which corresponds to higher transmitter power, is provided to sensors that have better channels with higher channel gain, h_l. Additionally it indicates that there is some channel gain $h_l < 1/\tau_0$ below which it does not make sense to transmit any power in order to achieve the desired distortion target. .

In [6], the opposite problem is solved. In this setup of orthogonal digital sensors, the estimate variance is minimized subject to constraints limiting both the total power transmitted by all sensors and on the power transmitted by any individual sensors. For analog sensors, similar minimizations can be performed indicating that transmit power is allocated to each sensor based on the combined quality of the sensor measurement noise and channel loss. Sensors with the lowest measurement noise and the lowest channel loss transmit the measurements with the most power [11]. A limiting aspect on performance in orthogonal schemes is that because the sensor measurements are processed individually and that each measurement introduces some independent sensing noise; consequently when transmit power is finite, even with $K \to \infty$ channels, estimation variance does not approach zero [1].

2.1.3 Energy-Efficient Estimation: Coherent MAC Sensor Transmissions

Amplify-and-forward is an example of coherent MAC sensor transmission. In this system, all sensor measurements are transmitted simultaneously and received by the fusion center in a manner where individual sensor transmissions cannot be isolated. The operation of an amplify-and-forward link in with perfect and identical channel state

information (CSI) is discussed in section 2.1.1. Optimal power allocation is discussed in [1] as a minimization of estimate variance, $\mathrm{Var}(\hat{\theta})$, subject to a constraint on the total power transmitted by all sensors. In [1], the measured parameter θ is unknown, distributed according to $\theta \sim \mathcal{N}(0, \sigma_\theta^2)$ and observed in Gaussian noise $\eta_l \sim \mathcal{N}(0,1)$ to produce noisy sensor measurement x_l. The transmitter amplifies the noisy sensor measurement by a factor a to produce output y. For the l^{th} sensor, output is found to be:

$$y_l = a_l(\sigma_{\theta_l}\theta + \eta_l). \tag{2-18}$$

The power transmitted by the sensor is:

$$P_l = a_l^2(\sigma_{\theta_l}^2 + 1). \tag{2-19}$$

The estimate variance, $\mathrm{Var}(\hat{\theta})$, is found in [1] according to the formula:

$$\frac{1}{\mathrm{Var}(\hat{\theta})} = 1 + \left(1 + \sum_{l=1}^{L} h_l\, a_l^2\right)^{-1} \left(\sum_{l=1}^{L} h_l^2 a_l^2 \sigma_{\theta_l}^2\right)^2, \tag{2-20}$$

where h_l is the channel gain for the l^{th} sensor and L is the number of sensors. The optimization of estimate variance given a total transmitted power constraint of P is done according to:

$$\begin{aligned} &\max_{a_l : 1 \le l \le L} 1 + \left(1 + \sum_{l=1}^{L} h_l\, a_l^2\right)^{-1} \left(\sum_{l=1}^{L} h_l^2 a_l^2 \sigma_{\theta_l}^2\right)^2 \\ &\text{subject to } \sum_{l=1}^{L} (\sigma_{\theta_l}^2 + 1) a_l^2 \le P. \end{aligned} \tag{2-21}$$

In minimizing distortion using coherent estimation, unlike for orthogonal transmissions, no sensors are turned off except for the trivial case of zero channel gain. Sensor is allocated optimally to each power based on the formula [1]:

$$P_l^{opt} = c_l P, \;\; l = 1, \ldots, L$$

Where:

$$c_l = c\frac{h_l^2\sigma_{\theta_l}^2\left(\sigma_{\theta_l}^2+1\right)}{\left(\sigma_{\theta_l}^2+1+h_l^2P\right)^2}, \qquad c = \left(\sum_{l=1}^{L}\frac{h_l^2\sigma_{\theta_l}^2\left(\sigma_{\theta_l}^2+1\right)}{\left(\sigma_{\theta_l}^2+1+h_l^2P\right)^2}\right)^{-1}. \tag{2-22}$$

The key asymptotic finding here is that unlike orthogonal sensor transmissions, estimate variance does approach zero as the number of sensors approaches infinity, even with a finite amount of total sensor power. Even with infinite power, if the number of channels is finite the estimator will still have non-zero variance [1].

2.2 Literature Review of Power Amps and Predistortion

2.2.1 Power Amplifier Background

Mobile communication systems are subject to two severe physical limitations: power and bandwidth. First transmitters are not typically connected to the power grid; therefore power consumption is critical to battery life and device usability. The second limitation is that bandwidth available over the air is both limited and expensive; therefore both a signal's amplitude and phase must be modulated to achieve high data rate within a given bandwidth. Power amplifiers are a requirement for any wireless communications system. There are many different power amplifier topologies [10]. Most of these topologies address different levels of the tradeoff between amplifier linearity and

amplifier efficiency. In a wireless communications system, the power amplifier is typically the single largest energy consumer. Consequently, its efficiency has a tremendous effect on the overall power efficiency of the communications system.

When the communication method being utilized involves changing, or modulating, the amplitude of the signal, it is important the amplifier have a constant gain over the range of expected signal amplitudes to avoid distorting the signal. The region of constant gain can be called the linear region of the amplifier. If the gain is not constant across the region of operation, distortion can result as shown in Figure 2-3. It can be seen that for a quadrature amplitude modulation (QAM) scheme, amplifier non-linearity reduces the distance between constellation points and thus degrades demodulation performance in the presence of noise.

Power amplifier efficiency is determined largely by the amplifier class. An amplifier's class is determined by the fraction of each sine wave spent conducting power. Typically, a full sinusoid cycle is represented by 2π. An amplifier that's conducting all the time would have a conduction angle of 2π, called class A. Class A amplifiers have theoretical PAE of up to 50% percent at maximum unclipped output power [10], and get much more inefficient when operating at lower power levels that may be required based on the input dynamic range. An amplifier that conducts during either the positive or negative half cycle would have a conduction angle of π, known as class B. If the amplifier cuts off some of its negative peaks but conducts during the entire positive peak,

than it would have a conduction angle between π and 2π (class AB) as shown in Figure 2-4.

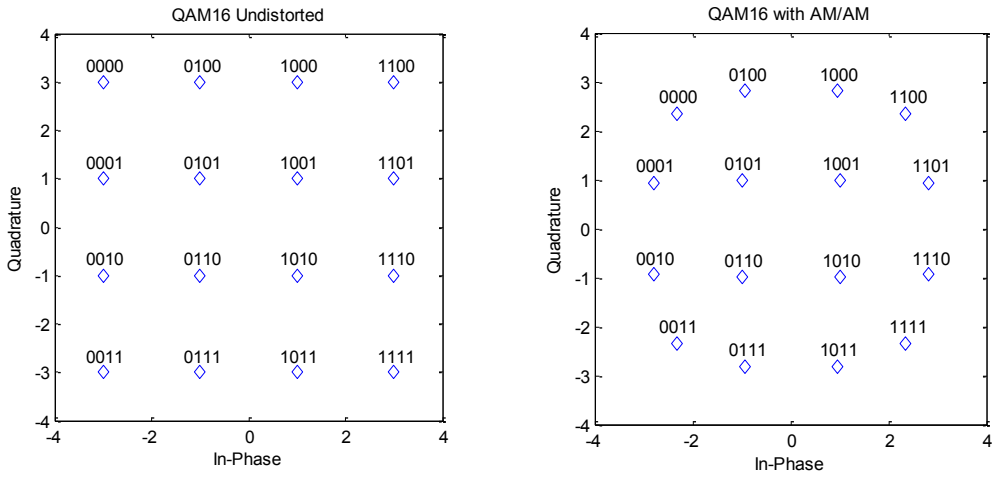


Figure 2-3: QAM16 (a) Ideal and (b) if amplifier gain is reduced at large amplitude.

When the conduction angle is 2π, the amplifier is conducting all the time. Amplifiers of this characteristic tend to have linear behavior and low distortion. This is because they do not clip the input waveform. However, a bias current is running all the time and is significant compared to the output current. Reducing the conduction angle decreases gain, but it also significantly reduces bias current and thus boosts efficiency [10].

When the conduction angle is reduced, the output waveform is a clipped sinusoid. For most communications systems, the carrier frequency is much faster than the modulation frequency. Consequently, the waveform is generally highly correlated cycle-to-cycle and can be considered nearly periodic over a period much shorter than the symbol time. This local (near) periodicity means it can be (nearly) represented by

Fourier series, or a sum of complex exponentials with frequencies at integer multiples of the fundamental frequency. This is the Fourier Series representation of a function $f(t)$, where T is the period of the fundamental, t, is time, c_n is the Fourier Series Coefficient of the n^{th} harmonic:

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{j2\pi t/T}. \tag{2-23}$$

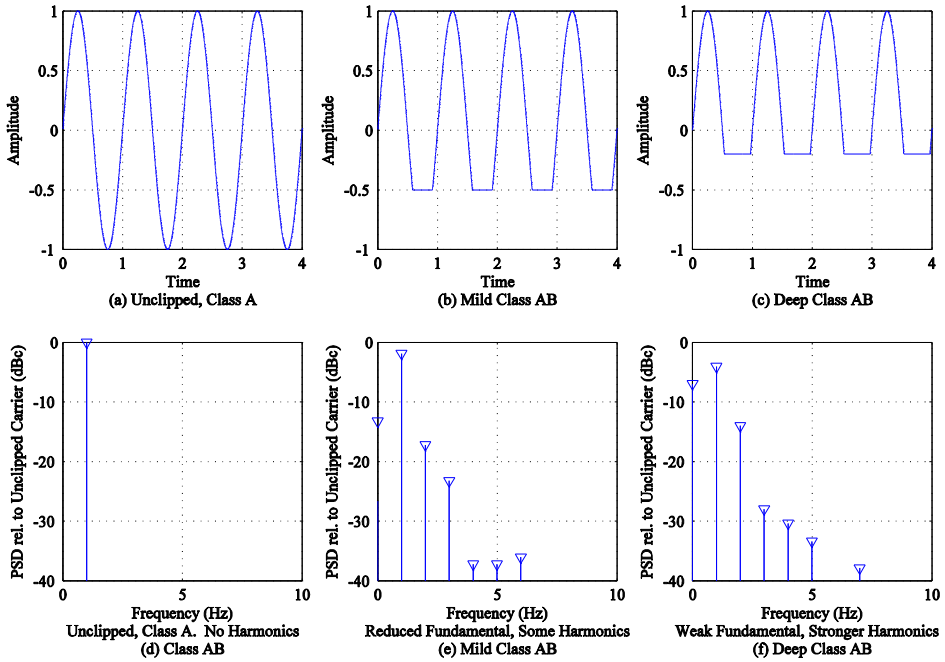


Figure 2-4: Various amounts of clipping and the resulting harmonic spectrums.

Figure 2-4 illustrates the effect of various amounts of amplifier clipping on a sinusoidal input. These 4 cycles of input sinusoid can be considered as the near stationary output of the amplifier within a given modulation symbol. It can be seen that as amplifier clipping is increased, the fundamental gain is reduced and more energy is

shifted to the harmonics. The shifting of energy from the fundamental to the harmonics under clipping mandates the use of high-order low-insertion loss filters at the output of a transmitter. Conventional filtering networks do a good job of removing harmonics of the carrier frequency.

A more difficult problem is introduced by the change in amplifier gain with clipping. For an amplifier with a fixed clipping level, the percentage of waveform cycle clipped varies with the input signal amplitude [10]. When a signal is amplitude modulated, its input level and consequently the gain of its primary harmonic changes. This results in a slowly time varying change in the amplifier gain. The change in gain is illustrated by Figure 2-5. Here, the input signal level causes the amplifier to effectively change conduction angle.

This time-varying amplifier gain multiplying the modulated input signal can be viewed in the frequency domain as the convolution of the time-varying amplifier gain spectrum and the original modulated signal spectrum. Thus, driving a class AB amplifier with an amplitude modulated signal into clipping (compression) causes spectral spreading on the order of the modulation frequency. This spectral spreading cannot be alleviated with simple low-pass filters as can harmonics of the carrier frequency. Spectral spreading can cause a transmitter to fail requirements, called masks, specifying how much power can be leaked into adjacent frequencies. Because the gain changes are introduced by changes in the input signal amplitude, modulation systems where the input has large swings in amplitude are more susceptible to distortion due to amplifier non-

linearity. One common modulation system in use today which has peak-to-average power variations (PAR) of around 10dB is orthogonal frequency-division multiplexing (OFDM). This system is used in IEEE standards 802.11g wireless LAN, 802.16e WiMAX, and Long-Term Evolution (LTE).

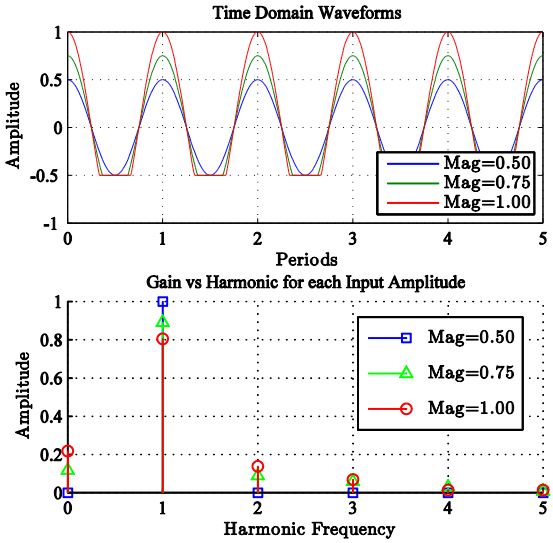


Figure 2-5: Change in Gain versus Input Magnitude for Class AB Amplifier.

When a multi-tone signal is fed into the amplifier, mixing products are produced as the harmonics are multiplied. Two sets of these mixing products produce problems that are difficult to filter away. First, mixing products that produce tones around the input signal frequency are problematic. These mixing products produce spectral spreading. Secondly, mixing products are produced near DC. The resulting low-frequency envelope can be responsible for memory effects, the dependency of amplifier outputs on previous inputs [28]. Slowly-varying low-frequency signals can be at high amplitudes and low-amplitudes for longer durations than the amplifier package thermal

time-constant, which can change the temperature, and thus the electrical characteristics, of the amplifier. Additionally, as signal bandwidth increases, the analog filters used to maintain a constant DC bias on the amplifier must be more broadband, requiring better performance that can be difficult to achieve at several MHz. The resulting bias movement causes a dependency on previous symbol points. A common symptom of memory effects is an asymmetric frequency spectrum [28].

Two of the metrics more commonly used metrics used to measure transmitter performance that directly indicate the linearity of the amplifier system are error-vector magnitude (EVM) and adjacent channel power ratio (ACPR). For OFDM data, the performance metrics EVM and ACPR are defined over the *N* subcarriers in the frequency domain:

$$EVM_{dB} = 10\log_{10}\left(\frac{\|\mathbf{V_{RX}} - \mathbf{V_{DETECTED}}\|^2}{\|\mathbf{V_{DETECTED}}\|^2}\right), \tag{2-24}$$

$$ACPR_{dB} = 10\log_{10}\left(\frac{\sum_{k=N+1}^{2N}\|\mathbf{V_{RX}}(k)\|^2}{\sum_{l=1}^{N}\|\mathbf{V_{RX}}(l)\|^2}\right). \tag{2-25}$$

Here, the $\mathbf{V_{RX}}$ is the column vector of all subcarrier's after the Fast-Fourier Transform (FFT) of the received packet data is taken. $\mathbf{V_{DETECTED}}$ indicates the subcarrier values within the intended channel, after they have been mapped to the nearest constellation point.

One common metric for amplifier efficiency is called power-added efficiency (PAE). PAE is defined as the ratio between signal power out and the sum of signal power and supply power into the amplifier:

$$PAE = \frac{P_{signalOut}}{P_{supplyIn} + P_{signalIn}} \ . \tag{2-26}$$

A table showing both gain and power-added efficiency for an actual class AB amplifier simulation is shown in Figure 1-1.

2.2.2 Power Amplifier Modeling

Many papers describe methods for developing analytical models for power amplifiers, as summarized in [29]. Most of these models fit the relationship of the baseband equivalent signal into and out-of the amplifier. A common family of models used for representing amplifiers used to transmit data signals are composed of full or reduced versions of the Volterra series [30,31]. Another common family of amplifier models used to represent amplifiers transmitting data is based on nonlinear neural networks [32]. These models often include memory-effects, which become very significant as modulated data signal increases to megahertz bandwidths. There are also amplifier models which place more emphasis on modeling amplifier compression behavior [33]. These limiting amplifier models are more useful in distributed estimation applications.

2.2.2.1 Volterra Series-Based Power Amplifier Models

A Volterra series model can be used to represent the non-linear transmitting behavior of power amplifiers operating with memory effects. The Volterra series can be viewed as a weighted combination of the products of various lag terms of an amplifier. A

discrete Volterra series with a memory of M taps and non-linearity up to order-K non-linearity of up to can be represented as [34]:

$$v_{out}(n) = \sum_{k=0}^{K} \left(\sum_{m_1=0}^{M-1} \sum_{m_2=0}^{M-1} \dots \sum_{m_k=0}^{M-1} a_k(m_1, m_2, \dots, m_k) v_{in}(n-m_1) v_{in}(n-m_2) \dots v(n-m_k) \right) \quad (2\text{-}27)$$

The problem with the Volterra series is that the number of complex weights, a, that must be estimated becomes impractically large, taking $\sum_{k=0}^{K} M^k$ separate weights to train. This makes it difficult to estimate weights in a manner that is computationally efficient and numerically robust. Aside from being expensive to compute a valid set of weights, once a set of weights is computed, a different set of $v_{in}(n)$ used to validate the robustness of the model may not yield a good fit.

Techniques have been developed to reduce the number of terms to be trained in the Volterra series. Some special cases of the Volterra series include parallel summations of cascaded memoryless nonlinearities and linear time-invariant (LTI) filters, with the cascaded elements known as either Wiener and/or Hammerstein models [35,36,37] as shown in Table 2-1. Where the linear filter has finite impulse response, the relation to the Volterra series is straightforward. When the linear filter has infinite impulse response, the relation is more difficult as $m_K = \infty$.

Table 2-1: Cascaded Elements of Wiener and Hammerstein Models

Reduced Volterra Model	Stage 1	Stage 2	Stage 3
Wiener	LTI	Memoryless Nonlinearity	Not used
Hammerstein	Not used	Memoryless Nonlinearity	LTI
Wiener-Hammerstein	LTI	Memoryless Nonlinearity	LTI

A power series is another model commonly fit to the amplifier's memoryless non-linear transfer function using least-squares techniques [37,38]. This power series models the changes invoked on the baseband equivalent signal. The fitting yields amplifier output as a power series of the baseband input signal $v_{in}(n)$ with weighting parameters a_k where K is the maximum order of nonlinearity:

$$v_{out}(n) = \sum_{k=1}^{K} a_k v_{in}(n)|v_{in}(n)|^{k-1}\,. \tag{2-28}$$

To incorporate memory effects into the power series, the weighted power series of the current power amplifier input with the weighted power series of previous amplifier inputs. In this power-series resembling model, called the memory-polynomial model [38,39], products between different memory lags are assumed to be 0, yielding for an amplifier with maximum memory M,

$$v_{out}(n) = \sum_{m=0}^{M} \sum_{k=1}^{K} a_{k,m} v_{in}(n-m)|v_{in}(n-m)|^{k-1}\,, \tag{2-29}$$

where $a_{k,m}$ is the complex weight of the m^{th} lag at the k^{th} order nonlinearity.

Finally, the memory polynomial is generalized in [35] to include cross-terms between a region of previous amplifier inputs and complex envelopes from M_b nearby terms. This model is known as the generalized memory polynomial, and can be represented as:

$$v_{out}(n) = \sum_{m=0}^{M} \sum_{k=1}^{K} \sum_{i=-M_b}^{M_b} a_{k,m,i} v_{in}(n-m) |v_{in}(n-m+i)|^{k-1}, \qquad (2\text{-}30)$$

The choice between amplifier models is made based on the accuracy and precision requirements of the system being designed around the amplifier. While more generalized models yield more accurate performance predictions, they require significantly more computation to generate. Among the most accurate and most expensive ways to model amplifier performance when such a model is required, is to use a periodic-steady-state circuit simulator to predict the baseband-equivalent amplifier output for a given amplifier input signal.

2.2.2.2 Memoryless Limiting Amplifier Models

Models are also available which pay more attention to the amplifier's compression characteristics than its memory effects [33]. These models are of greater interest when working with distributed estimation applications where data is usually transmitted more slowly, making memory effects less significant. The Cann model [16] has a parameter, $s_C \in \mathbb{R}^+$, to control sharpness of the gain curve "knee" where the amplifier enters the class AB region, and a parameter, $c \in \mathbb{R}^+$, indicating the amplifier's maximum, or limiting, value:

$$v_{out}(n) = \frac{c}{\sqrt[s_C]{1 + \left(c / v_{in}(n)\right)^{s_C}}}. \qquad (2\text{-}31)$$

A more analytically convenient and common model for soft saturation is the hyperbolic tangent function [40], where $c \in \mathbb{R}^+$ is again the maximum output amplitude.

$$v_{out}(n) = c \tanh\big(v_{in}(n)\big). \tag{2-32}$$

The hyperbolic tangent function model is convenient because its derivatives are defined, continuous functions and its inverse function can be determined analytically. These properties are utilized when analyzing distributed estimation through an amplifier in soft compression, as done in chapter 4.

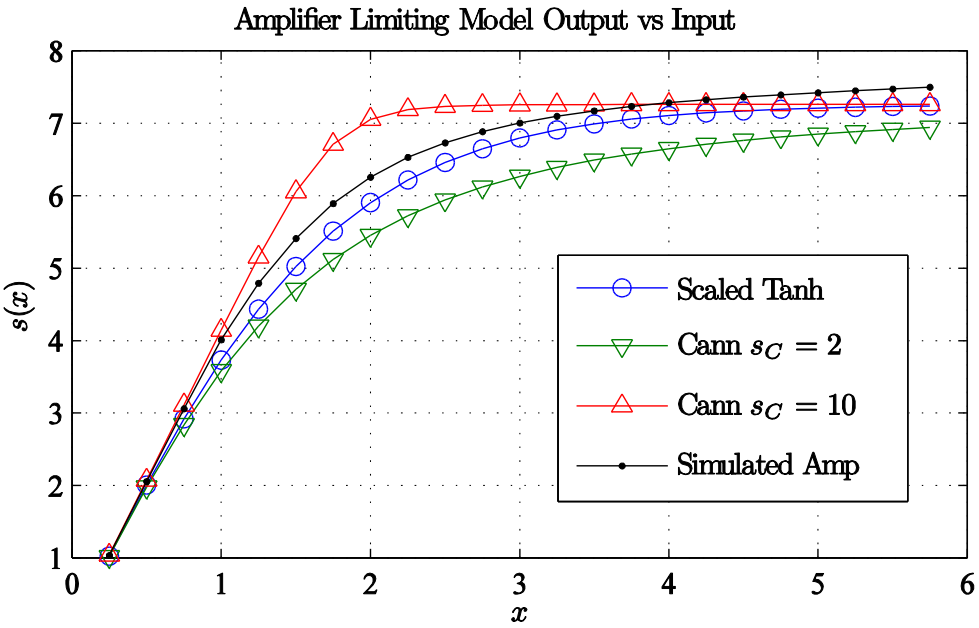


Figure 2-6: Limiting Amplifier Models

A comparison between the Cann model utilizing various of knee sharpness, s_C, the hyperbolic tangent model, and a simulated class AB amplifier circuit is shown in Figure 2-6. It can be seen from the figure that hyperbolic tangent model behavior can be close to that of a simulated amplifier, despite the relative lack of parameters available for

tuning its behavior. Furthermore, where tuning the sharpness of the transition to more closely match amplifier compression is beneficial, than the Cann model can be utilized. It can also be observed that as $s_C \to \infty$, the Cann model approximates the behavior of the ideal perfectly linearized power amplifier.

2.2.3 Digital Predistortion Overview

Predistortion can be implemented either via analog circuitry or digital circuitry. Digital predistortion is more prevalent due to its easier control and adaptation. A typical digital predistortion system is shown in Figure 2-7. In digital predistortion, the actual transmitter output is monitored by a co-located receiver. By comparing received values to the desired to be transmitted, an adaptation algorithm can be used to modify the predistortion function such that actual outputs match the values intended to be transmitted accounting for the process, voltage, and temperature (PVT) variation. Two of the methods used for digital predistortion include gain-based look-up-table (LUT) [15] and neural network techniques [41].

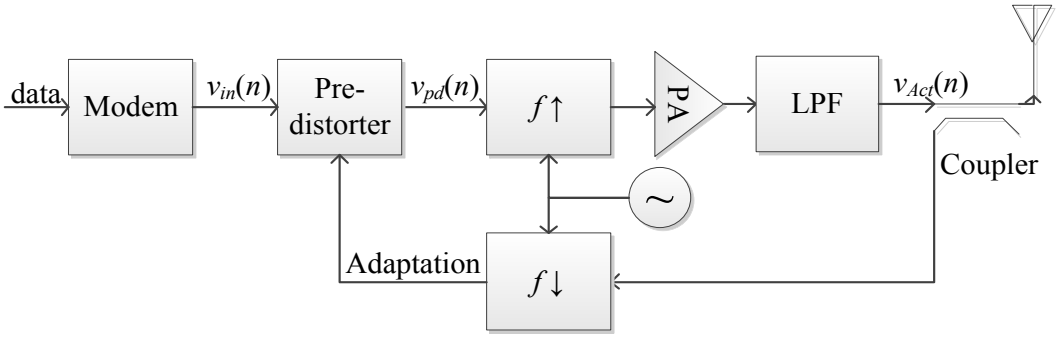


Figure 2-7: Typical predistortion system, modified from [15].

2.2.3.1 Gain Based Look-up-Table Predistortion for Memoryless PA

The gain-based predistorter [15], shown in Figure 2-8, is a common predistorter that compensates for variable amplifier gain at the fundamental signal frequency. It dramatically extended the capabilities, decreased the learning time, and reduced the memory requirements from the original mapping predistorter developed in [42]. The gain-based predistorter topology is based on two assumptions. First, it assumes compression occurs primarily as a result of input power. Second, it assumes adjacent input power levels have highly similar power amplifier complex gains. Consequently, sensor outputs are assigned look-up-table (LUT) bins based on their input power level. The exact method used for allocation of these bins is discussed in [43].

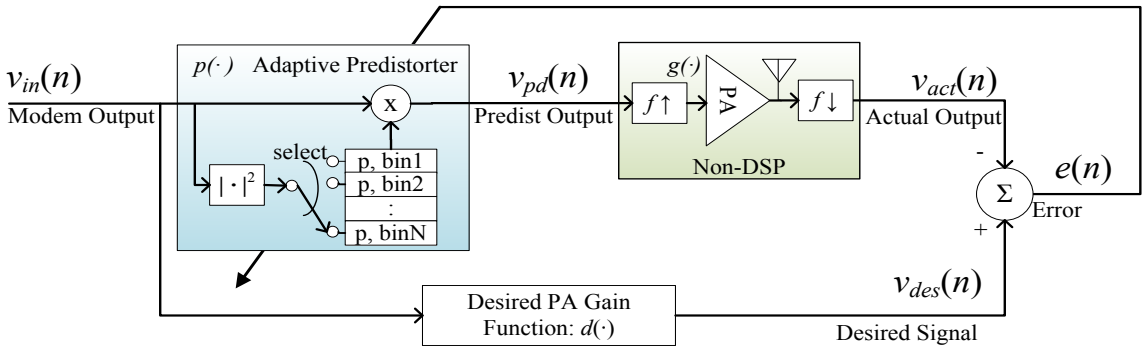


Figure 2-8: Gain-based predistortion.

Each bin has a correction factor which scales the sensor output, with the intent of bringing the combined gain of the predistorter and the nonlinear amplifier to the total desired gain. The correction factors are learned via the least-mean-squares (LMS) gradient descent algorithm [44]. It is assumed that increasing the correction factor

magnitude increases the output magnitude of the amplifier, so that the amplifier's output is monotonically increasing with output power.

For a given bin, the error, *e*, can be found by finding the difference between the actual signal transmitted, $v_{act}(n)$, and the desired signal, $v_{des}(n)$ according to (2-33). Actual amplifier gain is $g(\cdot)$, predistorter gain is $p(\cdot)$, and the desired gain resulting from the soft-saturation function is $d(\cdot)$.

$$e(n) = v_{des}(n) - v_{act}(n) = d\big(v_{in}(n)\big) - g\big(p(n)v_{in}(n)\big). \tag{2-33}$$

The error, $e(n)$, the current predistortion correction factor for input points within the corresponding bin, $p_{bin}(n)$, and the input value $v_{in}(n)$, can then be used to compute an updated correction factor $p_{bin}(n+1)$ according to the LMS algorithm as shown:

$$p_{bin}(n+1) = p_{bin}(n) + 2\mu e(n)v_{in}^{*}(n) \tag{2-34}$$

Many different customizations have been made to the LUT technique to aid with modeling accuracy given specific nonlinear characteristics, memory effects, or to improve learning time in the presence of significant delay between when the modem sends the transmit signal and has the time-synchronized signal from the receiver available for adaptation [43].

2.2.3.2 Gain Based Look-up-Table Predistortion for PA with Memory Effects

The gain-based LUT has been extended to compensate for power amplifier with memory effects. In one extension, proposed by [45], a second-dimension is added to the gain-based LUT. The second dimension is indexed by a metric of previous symbol

powers. Three different symbol power metrics are utilized in [45]. In the first metric, the ratio of current symbol averaging the symbol power over the M previous symbols is used. The second metric shown took a weighted average of the symbol powers, where each symbol was weighted by an exponential decay, D^{-m}, where m is the amount of delay from the current sample and $0 < D < 1$, to provide more weighting to the most recent terms. The third metric used is a special case of the first where M=1 delay term, resulting in the ratio of current symbol power to previous symbol power. As a variation of the approach utilized in [45], previous symbol amplitudes instead of powers are used similarly in [46]. A convenient side effect of using amplitudes instead of powers is a reduction in required computations in the digital signal processor.

In another extension of gain-based LUT predistortion developed in [39], it is assumed the PA with memory effects is modeled as a power series function with finite non-linear memory, M. The LMS algorithm is still split into bins based on the current input symbol power. These bins now consist of finite impulse response (FIR) filters, not simple gain factors. The assumptions about transmitter gain being locally similar and monotonic from memoryless gain-based predistortion are still required. Additionally the weighted higher-order terms of the memory lag are assumed to remain small compared to the current sample, otherwise they will dominate the response and a linear filter may not provide good performance. The update equation now becomes a vector equation,

$$\mathbf{b}(n+1) = \mathbf{b}(n) + 2\mu e(n)\mathbf{v}_{\mathbf{in}}^{*}(n), \tag{2-35}$$

where $\mathbf{b}(n)$ is the column vector of FIR coefficients for a given bin at time n, and $\mathbf{v_{in}}(n)$ is the column vector of current and previous input points, current input being at time n:

$$\mathbf{b}(n) = \begin{bmatrix} b_0(n) \\ b_1(n) \\ \vdots \\ b_M(n) \end{bmatrix}, \qquad \mathbf{v_{in}}(n) = \begin{bmatrix} v_{in}(n) \\ v_{in}(n-1) \\ \vdots \\ v_{in}(n-M) \end{bmatrix}. \tag{2-36}$$

Similar with the gain-based technique, the selected FIR filter coefficients within a selected bin, $\mathbf{b}(n)$, are used to calculate the predistorter output, $v_{pd}(n)$, for the next sample within a given LUT bin after the update operation is completed. The FIR equation is:

$$v_{pd}(n) = \mathbf{v_{in}^T}(n)\mathbf{b}(n). \tag{2-37}$$

The major weakness of the LMS gain-based predistorters is also that a significant assumption is made about the structure of the linearizer required and thus the behavior of the power amplifier; specifically, the linearizer can be represented accurately by a linear FIR filter within each given bin. The non-linear aspect of behavior comes from the selection of different linear filters based on the current input symbol amplitude.

2.2.3.3 Neural Network Based Predistortion

Another approach not utilizing LUTs that can be taken to implanting digital predistortion is to either learn a function by which the power amplifier can be modeled by a known model form and then invert it as is done in [47], or by learning the inverse function of the power amplifier itself [37]. These approaches use conventional power amplifier models such as Weiner-Hammerstein or memory polynomial models;

consequently, like the LMS based approaches make significant assumptions about the behavior of the power amplifier. A neural-network based method of identifying the inverse function allows estimation of the power amplifier function with less restrictive assumptions about its form. This may be able to provide a better estimate of the power amplifier function and its inverse.

First, a network is developed that estimates the inverse function of the power amplifier with the desired gain removed. Next, a copy of this network is placed immediately after the modem as the predistorter. This copied network then warps, or predistorts, the value to be transmitted such that after the amplifier distortion the transmitter output is as originally intended by the modem

The simplified topology used in the neural network algorithms is the indirect learning architecture described in [37] shown in Figure 2-9. Note that the desired gain of the PA is removed when identifying g^{-1}, the inverse function of PA gain.

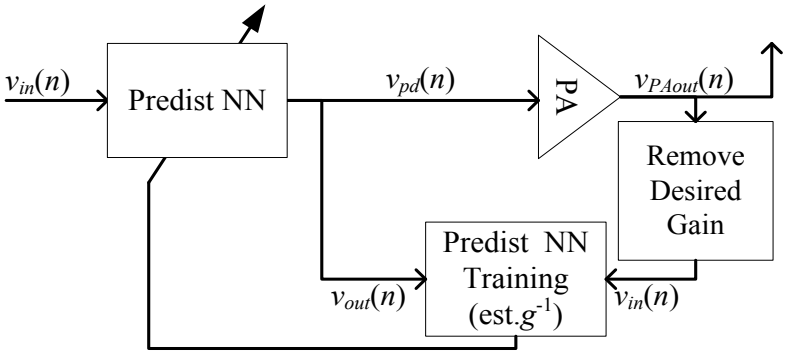


Figure 2-9: Neural Net Predistortion by Indirect Learning

When identifying the baseband-equivalent function of a power amplifier, all terms involved are complex valued. But, many of the existing tools and code for neural network training and evaluation assumes real numbers [48]. By separating out the baseband equivalent inputs and outputs to the in-phase and quadrature (I and Q) components while leaving them connected internally, it is implied in [48] that the modified neural network can remain fully real but yet accommodate for operations such as phase rotation. Specifically, in [48] a neural network estimates the forward function of a power amplifier with memory effects. The inputs to the network are the in-phase and quadrature components of the current and (M-1) previous input symbol coordinates, and the outputs are the current in-phase and quadrature output coordinates.

In [32,41,49], neural networks using similar topologies are used to identify the inverse function of a power amplifier excluding the nominal desired gain of the amplifier. The real-valued neural networks all consist of a series of interconnected neuron layers. The neuron layers each take weighted inputs from the previous stage, sum them with a weighted bias, and pass the sum through an activation function to produce an output. In both [32,41], the activation functions used are linear for the first layer, tan-sigmoid for the second layer, and linear for the third layer. The weights are initialized with random values. The neural network is trained using back-propagation to model the inverse power amplifier function. The current and lag values of baseband I and Q components measured at the transmit antenna are the inputs to the network and the baseband modem feed is the expected output.

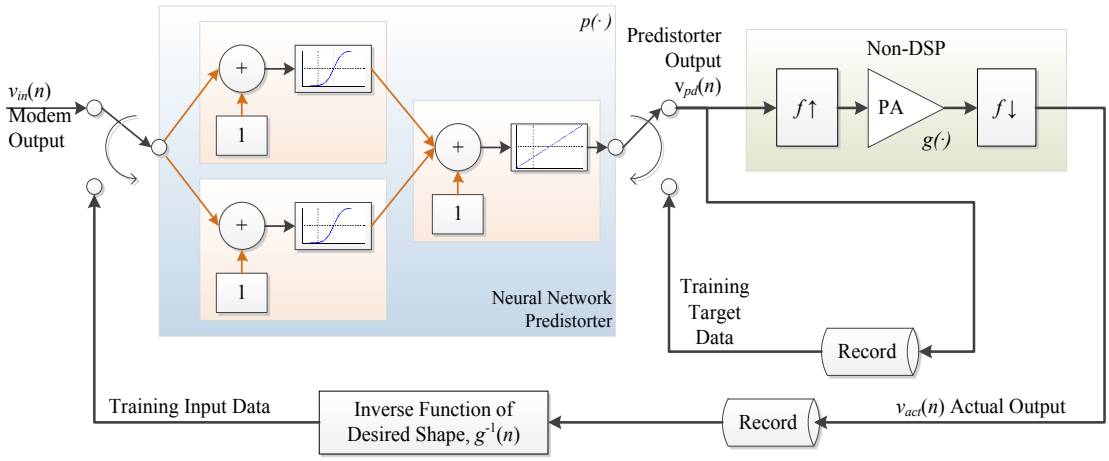


Figure 2-10: Neural-Network Based Predistortion System.

From Figure 2-10, it can be determined that if the predistorter has function $p(n)$ and the combination of up-converter and power amplifier have function, $g(n)$, then the following relationships hold:

$$v_{act}(n) = g\left(v_{pd}(n)\right) = g\left(p\left(v_{in}(n)\right)\right). \tag{2-38}$$

The following substitution can be made to make $v_{act}(n)$ a predetermined function of input voltage:

$$v_{act}(n) = d\left(v_{in}(n)\right). \tag{2-39}$$

Solving (2-38) for $v_{in}(n)$ and substituting back into (2-39) yields:

$$v_{act}(n) = g\left(p\left(d^{-1}(v_{act}(n))\right)\right). \tag{2-40}$$

Applying the inverse of function g to each side of (2-40) yields:

$$g^{-1}\left(v_{act}(n)\right) = p\left(d^{-1}(v_{act}(n)\right). \tag{2-41}$$

The function $d^{-1}(\cdot)$ can be determined provided the amplifier system desired function $d(\cdot)$ is invertible. Furthermore, $g^{-1}\left(v_{act}(n)\right)$ is the input to the amplifier,

$v_{pd}(n)$, that produces output $v_{act}(n)$. The predistortion function $p(\cdot)$ is the function that takes as input $d^{-1}\left(v_{act}(n)\right)$ and produces output $v_{pd}(n)$. When measurements of power amplifier output are available, this function can be fit to a neural network. Such a network is numerically fit using a Levenberg-Marquadt neural network training algorithm [50] on a neural network consisting of typically two layers. The first layer uses non-linear hyperbolic tangent functions capable of modeling non-linear combinations and the second layer is a linear node as a combiner.

The neural network based predistortion techniques are computationally more intensive then the LMS based approaches. The linearization computation is done between transmitted packets. The memory requirements can be large due to the number of weights involved in the network. For the neural network, the number of weights in the network, *w*, can be computed by:

$$w = \sum_{m=0}^{M} c_m(c_{m-1} + 1)\,, \tag{2-42}$$

where *M* is the number layers in the neural network, c_m is the number of neurons in the m^{th} layer of the network, and c_0 are the number of inputs to the networks first layer.

2.2.3.4 Considerations for Implementing Practical Digital Predistortion

When implementing predistortion systems, some practical concerns must be met. First, to cancel the spectral broadening introduced by nonlinear amplification, the predistortion system itself is equally broadband in most predistortion systems. This

requires oversampling the transmitted data by twice the order of the highest nonlinear term [51]. This oversampling increases power requirements in the baseband analog converters and the digital signal processor of a conventional in-phase/quadrature modulated system. Additionally, feeding data back into the DSP for performing predistortion requires an additional receive chain to be implemented in the predistorter. This shift of complexity from the RF components to the baseband and digital domain has two benefits. First, digital circuits can be inexpensively manufactured to operate deterministically across process corners, while the performance of analog devices varies. Second, the power consumption and size of digital integrated circuits improves much faster than analog devices, a realized trend predicted by in 1965 as Moore's Law [52]. Despite the trend towards increase digital performance at lower power levels, the additional power penalty imposed by the increased DSP and baseband requirements must be considered when choosing if to implement predistortion.

Care must be taken in designing the quadrature modulator and demodulator to minimize errors. Errors introduced in these devices in include both gain and phase imbalance between the in-phase and quadrature components, and DC offset. Methods for quantifying the performance impact due these imperfections are discussed in [53]. Techniques for correcting these imperfections are discussed in [54]. Further design tradeoffs for gain-based LUT and neural network based predistortion are discussed in [55,56,57]

3 PREDISTORTION

In this chapter, two different modifications to existing predistortion techniques are proposed. The first technique, outlined in section 3.1 is a block-LMS based technique intended to demonstrate a reduction in convergence time required to train the gain-based LUT's FIR filters [58]. The second technique, outlined in section 3.2, is designed to reduce the number of weights that must be trained for neural network based predistortion [59]. The proposed technique to reduce neural network weight count is implemented by starting the neural network with inputs approximating a memory-polynomial based predistortion.

3.1 Block Adaptive Gain-based LUT Predistortion

3.1.1 Algorithm Description

In an attempt to reduce measurement noise and to decorrelate the input data, a modified block LMS (BLMS) algorithm was developed. The difference from conventional BLMS algorithms [60] is due to the power binned nature of the gain-based predistorter. As with standard BLMS approaches, this BLMS computes a gradient based on averaging *N* individual measurements of error as shown:

$$\mathbf{b}(n+1) = \mathbf{b}(n) + 2\frac{\mu}{N}\left[\mathbf{v}_{\mathbf{in}}^{*}(n_1) \quad \mathbf{v}_{\mathbf{in}}^{*}(n_2) \quad \dots \quad \mathbf{v}_{\mathbf{in}}^{*}(n_N)\right]\begin{bmatrix} e(n_1) \\ e(n_2) \\ \vdots \\ e(n_N) \end{bmatrix} \tag{3-1}$$

In the standard BLMS algorithm, the measurements are taken for subsequent points where $n_2 = n_1 + 1$, implying a vertical shift of one element between subsequent

columns. Due to the binning behavior of the gain-based LUT, this relationship does not hold for measurements in this BLMS implementation. As shown in Figure 3-1 subsequent measurements may be assigned to different bins based on their input power magnitude.

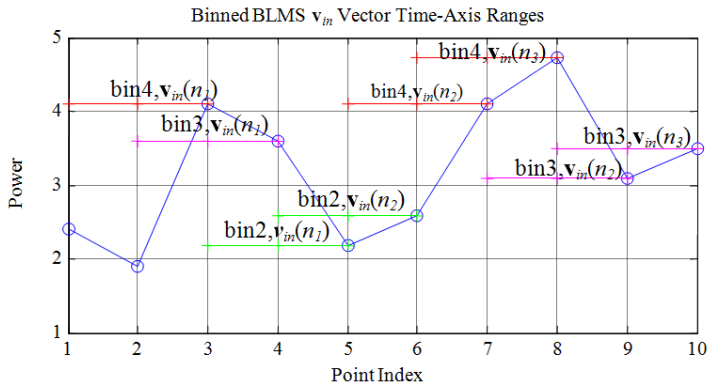


Figure 3-1: Block LMS Gain Based Predistorter Binning

Consequently, the indexes within the BLMS matrix with $N = 3$ measurements for bin 4 with an FIR memory $M = 2$ could be:

$$\mathbf{b}_{bin4}(9) = \mathbf{b}_{bin4}(8) + 2\frac{\mu}{N}\begin{bmatrix} v_{in}^*(3) & v_{in}^*(7) & v_{in}^*(8) \\ v_{in}^*(2) & v_{in}^*(6) & v_{in}^*(7) \\ v_{in}^*(1) & v_{in}^*(5) & v_{in}^*(6) \end{bmatrix}\begin{bmatrix} e(3) \\ e(7) \\ e(8) \end{bmatrix} \tag{3-2}$$

The BLMS algorithm is evaluated to determine if it can converge in fewer points for a transceiver with significant latency from predistorter output to output data measurement. In such a system, all data samples between the time an adjustment is made in the predistorter and the reception of the first data point using that adjustment are of limited use. The cost of making an adjustment is that all measurements taken for points within that specific LUT bin during the latency duration are discarded. Thus it is critical that when an adjustment is made, it is made with upmost accuracy. Implementation of

this modified BLMS requires maintenance of a matrix for each LUT bin along with a trigger condition to trigger evaluation when the matrix is full and ready for gradient calculation.

3.1.2 Simulation Environment

To simulate the proposed digital predistortion algorithm, a MATLAB simulation environment was written. The simulation consists of a signal generator, a parameterized digital predistorter, a PA model, and a signal analyzer.

3.1.2.1 OFDM Signal Generator and Signal Analyzer

The signal used to test the algorithm is an OFDM signal with 2048 subcarriers. Each subcarrier is BPSK modulated with amplitude of 1 and a random data bit. This signal is oversampled by a factor of 4. There is a cyclic prefix of 80 points at the 4x oversampled rate. The points used for the cyclic prefix are the last 80 points of the transmitted data burst. Additionally, every fourth subcarrier is used as a pilot tone for tracking channel gain variation by the receiver.

The OFDM receiver starts by locating the received signal time-offset that maximizes the autocorrelation of points separated by the word length in time-domain data points. This trick is exploitable because the cyclic prefix is the same as the transmitted data word itself. Because the amplifier model being used is constant, a filter tracking movement of the signal time-offset is not required.

The receiver trims off all data, including the cyclic prefix, not part of the modulated data burst. It takes the FFT of the modulated data word to recover the individual data subcarriers. The pilot tones are used to reverse gain variation across the channel bandwidth, with simple linear interpolation used to correct subcarriers between each pilot tone. After the correction, a measurement of EVM and ACPR is taken. Finally, the received data is compared to the transmitted data to check for any bit errors within the burst.

3.1.2.2 PA Model

The power amplifier model presented by [37] is used in this simulation. The power amplifier model incorporates memory effects. It is implemented in the memory polynomial format indicated by (2-29). The maximum valid magnitude for the inputs of this polynomial while attaining compressive characteristics is 0.75. Above this, output was hard-clipped to. The nominal gain is 1. The coefficients used were:

Table 3-1: Power amplifier memory-polynomial model coefficients

$a_{1,0} = 1.0513 + j\,0.0904$	$a_{5,1} = -0.2451 - j\,0.3735$
$a_{3,0} = -0.0542 - j\,0.2900$	$a_{1,2} = 0.0289 - j\,0.0054$
$a_{5,0} = -0.9657 - j\,0.7028$	$a_{3,2} = -0.0621 - j\,0.0932$
$a_{1,1} = -0.0680 - j\,0.0023$	$a_{5,1} = 0.1229 + j\,0.1508$
$a_{3,1} = 0.2234 + j\,0.2317$	all other terms 0

3.1.3 Block LMS Predistortion Performance

The block LMS predistorter was simulated with average transmit power backed off from peak amplifier by values from 10dB to 18dB, stepping by 2dB. For each power

level, 20 packets of OFDM data were simulated. To quantify the performance differences caused by the predistorter, ACPR was measured for each packet for each of 4 different predistortion configurations. The first simulation run was with predistortion disabled, shown in Figure 3-2(a). In this configuration, large values of ACPR indicate the amplifier is experiencing significant spectral leakage due to non-linearity. It also indicates the leakage is as expected, stronger for lower power back-off values such as 10dB. The second configuration was a predistorter with a block size of 1, corresponding to a conventional non-block LMS and shown in Figure 3-2(b). For 10dB and 12dB power back-off, a significant reduction in ACPR is observed. Larger back-offs indicate an increase in ACPR, indicating a probably need to normalize against μ. The modified block LMS algorithm can be seen in Figure 3-2(c) and (d) with block sizes of 4 and 16 respectively. An increase in convergence speed and decrease in sensitivity at larger power back-off values is observed.

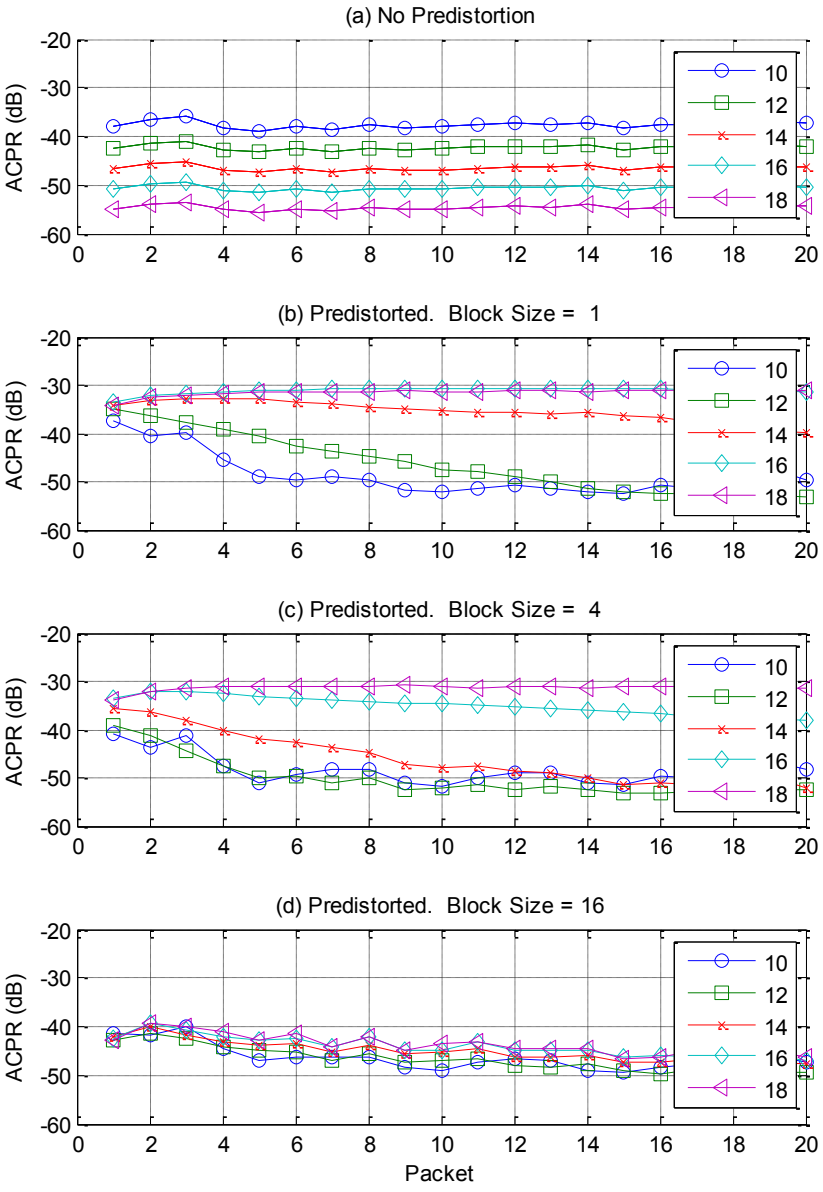


Figure 3-2: ACPR at different power back-off values versus sequential packet with (a) no predistortion, (b)conventional predistortion, and block LMS predistortion with block sizes of (c) 4 and (d) 16.

3.2 Neural Networks w/ Higher-Order Terms as Inputs

3.2.1 Proposed Change

A major issue with the neural network predistortion is the amount of computation required to train the network. This research attempts to improve performance by employing a hybrid between the neural network topology, which is ignorant about the function being modeled, and the LMS which assumes a specific equation form for the predistortion function. This decision was made with the objective of reducing the number of nodes in the neural network and thus lighten the computational and memory requirements. To implement such a topology, it was assumed that the primary function involved in a power amplifier is roughly a memory polynomial. A memory polynomial is a simplification of the Volterra series where the higher-order products between different lag terms are all insignificant [37]. Additionally, it was assumed second order terms were also insignificant. Third power terms of the in-phase and quadrature input series, $v_{in_I}(n)$ and $v_{in_Q}(n)$, were assumed significant and thus $\left|v_{in_I}(n)\right|^2 v_{in_I}(n)$ and $\left|v_{in_Q}(n)\right|^2 v_{in_Q}(n)$ input magnitudes were added to the list of network inputs. Where other terms or cross-terms are required, it is desired to let the neural network learn them.

3.2.2 Simulation Results

To validate the proposed neural network predistortion, two different neural network datasets were simulated. In the first data set, a fairly large neural network configured as specified in Table 3-2 topology (a) was utilized. Utilizing this predistortion with only first-ordered lag terms yielded an improvement in ACPR levels relative to the case of no predistortion across a large range of power back-off values, as shown in Figure 3-3 for topology (a). This confirms the methods described in [32] and [41] are capable of linearizing amplifiers. By including third-order terms to the neural network, and thus mimicking the structure of a memory polynomial, performance was increased as indicated by a further reduction in ACPR for the 3rd order predistortion of topology (a) in Figure 3-3. However this further reduction of ACPR was attained at the cost of introducing new weights to the neural network, boosting the number of weights from an already large 707 to an even larger 857 weights.

To validate the worthiness of the proposed approach, the linearization performance would have to improve while simultaneously reducing the number of weights in the neural network more than what is possible for the first-order terms only network. This is demonstrated using the neural network topology specified in Table 3-2 simulation set (b). Using this set, it can be seen that predistortion using the reference technique required 229 weights, but failed to provide a large boost in linearization at largest generally usable back-off of around 10dB. However, by including the third-

ordered terms, topology (b) was able to outperform the reference technique from topology (a) utilizing only 299 weights instead of 707 weights.

Table 3-2: Neural Network Simulation Set Topologies

Network Topology	(a)	(b)
Layer 1 Neurons	15	7
Layer 1 Function	Tan-Sigmoid	Tan-Sigmoid
Layer 2 Neurons	30	15
Layer 2 Function	Tan-Sigmoid	Tan-Sigmoid
Layer 3 Neurons	2	2
Layer 3 Function	Linear	Linear
Total Weights, First Order Terms Only	707	229
Total Weights, First and Third Order Terms	857	299

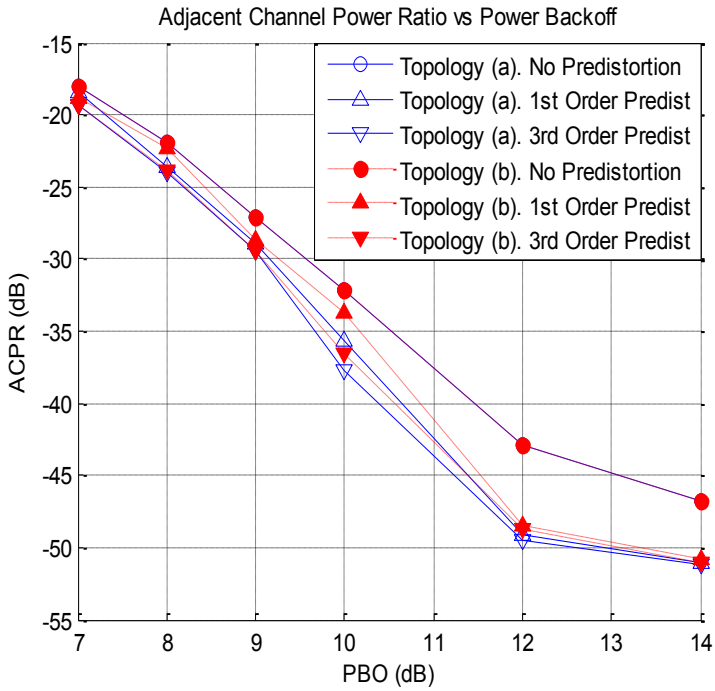


Figure 3-3: Improved Performance with Lower Weight Count

4 ESTIMATION IN SOFT COMPRESSION, AWGN

Traditional amplify-and-forward (AF) distributed estimation techniques [11] assume the entire dynamic range of noisy sensor measurements is within the transmitter's linear region. Although this assumption may be permissible when measuring a parameter of small dynamic-range and transmitting via a linearized class AB amplifier, this assumption limits the efficiency of higher dynamic range systems. The limitation is imposed because assuming the transmitter is linear over the full measured parameter dynamic range implies that much of the time the amplifier will be operating well backed-off of its maximum amplitude and thus in the inefficient class A region of operation. In this chapter, an alternate approach is proposed to allow utilizing the amplifier more efficiently in the gain-compressed operating region.

In the proposed approach, rather than trying to linearize away the mathematically inconvenient gain compression, it is utilized to attain higher sensor energy efficiency. Specifically, the soft-limiting behavior of amplifiers is utilized to compress the dynamic range of the sensor measurements, and thus largely avoid using amplifiers in the inefficient deep class-A region of operation. Implementation of the proposed system does impose new design challenges. First, because amplifiers are being used in a non-linear manner, spectral widening can be a problem if the sensors are measuring signals of arbitrarily high bandwidth. For simplicity, it is assumed that the sensors are measuring slowly-varying physical quantities such as temperature or moisture for environmental monitoring thus largely avoiding mixing products. Second, introducing nonlinearities to

the amplify-and-forward process makes deriving estimators much more complex. Finally, a mechanism needs to be developed to ensure all sensors have a similarly profiled soft-limiting behavior.

Implementing and analyzing a distributed estimation system utilizing the soft-limiting behavior of amplifiers requires three things: a model for the soft-limiting behavior of amplifiers, a method to ensure all amplifiers in the sensor network adhere to the behavior specified in the model across PVT variation, and an algorithm for computing and quantifying the validity of estimates based on the value received at the fusion center from the sensors. The proposed system shows increases in sensor power-added-efficiency are attainable in exchange for a predictable and tolerable penalty in estimate variance. The next two subsections describe the first two requirements: limiting amplifier models and a method to make sure amplifiers match the models.

4.1 Limiting Amplifier Model Selection

When implementing predistortion, the first decision that must be made is to decide what function it is desired to have the amplifiers model. The conventional choice when implementing predistortion is to linearize the amplifier, but that need not be the only option. To utilize the gain compressed region of operation, limiting amplifier models are utilized as the target function to which the transmitter is matched via predistortion. When selecting a limiting amplifier model, the selected model must be chosen such that the sensor network produces consistent estimates.

Using techniques for distributed estimation with bounded transmissions outlined in [25], it is asserted that if a non-linear transmitter shaping function satisfies certain criteria, a consistent estimate of θ can be attained. The required criteria for this non-linear function from [25] are in place to ensure that if the non-linearity, $s(x)$, is one-to-one and monotonic, then the expected value at the receiver, $\bar{\zeta}(\theta)$, is also one-to-one and monotonic. The required criteria are:

- $s(x)$ is differentiable everywhere in the domain of x $(\forall x)$;
- $s(x)$ is bounded by a finite constant, c, such that $|s(x)| \leq c < \infty, \forall x$;
- $\lim_{|x|\to\infty} |s(x)| \to c$;
- $s(x)$ is monotonic in x.

For the sake of completeness, a proof of the required criteria from [25] is summarized. Lemma 1 of the proof demonstrates that mean of all transmissions received at the estimator in the fusion center can be found to be a function, $\bar{\zeta}(\theta)$, independent of the individual sensing noise distributions, η_l, distributed according to probability density function $p(\eta_l)$ when the number of sensors, $L \to \infty$, as shown:

$$\lim_{L\to\infty} \frac{1}{L} \sum_{l=1}^{L} s(\theta + \eta_l) = \zeta(\theta). \tag{4-1}$$

To prove lemma 1, it is demonstrated in [61] that the mean, $\zeta(\theta)$, of a set of random variables with finite variances approaches its expectation, $\bar{\zeta}(\theta)$, almost surely as the number of random variables, $L \to \infty$ where:

$$\bar{\zeta}(\theta) = E[\zeta(\theta)] = \lim_{L\to\infty} \frac{1}{L} \sum_{l=1}^{L} \int_{-\infty}^{\infty} s(\theta + \eta_l)\, p(\eta_l) d\eta_l \,, \tag{4-2}$$

This assumption can be made with regard to random variables $S_l = s(\theta + \eta_l)$ whenever the variance of $\mathrm{Var}[s(\theta + \eta_l)] \leq c^2$, which holds when the function is bounded by $|s(x)| \leq c, \ \forall x$. This bounding is consequently listed in the $s(x)$ requirement. This bounding holds even distributions for which a mean does not exist, such as the Cauchy distribution.

Once a function, $\bar{\zeta}(\theta)$, can indicate the value expected at the fusion center's estimator, z, for a given θ, the inverse function can be used to compute the an estimate, $\hat{\theta}$ according to:

$$\hat{\theta} = \bar{\zeta}^{-1}(\theta). \tag{4-3}$$

The inverse function is most useful if there is a one-to-one relationship between parameter, θ, and expected value received at the fusion center $\bar{\zeta}(\theta)$, and thus between the observed value received at the fusion center, z, and parameter estimate $\hat{\theta}$. A one-to-one relationship will hold in $z = \bar{\zeta}(\theta)$, if each component of the sum in equation (4-2) is itself differentiable and monotonic. For each component $\varsigma_l(\theta) = E[s(\theta + \eta_l)]$ to be monotonic, it is shown in [25] that using the assumptions that $s(x)$ is differentiable everywhere and bounded by $\pm c$ in the domain of x yields the requirement that $s(x)$ should itself be monotonic. Usually monotonically increasing functions are used by convention as shown in (4-4):

$$\frac{\partial \varsigma_l(\theta)}{\partial \theta} = \int_{-\infty}^{\infty} \frac{\partial s(\theta + \eta_l)}{\partial \theta} p(\eta_l)\, d\eta_l > 0, \tag{4-4}$$

although implementation using monotonically decreasing functions is also possible.

Both of the limiting amplifier models presented in section 2.2.2.2, meet the criteria required for consistent distributed estimation from [25]. The limiting amplifier models are parameterized to allow fitting them to a performance consistently achievable across all sensors. For the Cann model from (2-31), the new parameters added is a measurement scaling factor applied to the noisy sensor measurement, k. The existing parameters for maximum compressed output, c, and the sharpness factor, s_C, yield a transmitter function, $s(x)$ where $x > 0$:

$$s(x) = \frac{c}{\left[1 + \left(\frac{c}{x/k}\right)^{s_C}\right]^{\frac{1}{s_C}}}. \tag{4-5}$$

While the Cann model does have the benefit of a sharpness parameter, s_C, that allows tuning to match the model abruptness to the compression characteristics of more amplifiers, it has a drawback that it is not very analytically convenient for conducting mathematical analysis of the estimation algorithm performance, though such analysis has been completed and is discussed in section 4.4. The function is difficult to invert, making exact determination of the estimator difficult. For most analysis, the scaled hyperbolic tangent function was used as it is a more common and tractable limiting amplifier model.

The scaled hyperbolic tangent function from (2-32) is an analytically tractable model that both satisfies the mathematical requirements for a consistent estimator to exist

and can accurately approximate the response of a transmit amplifier, as shown in Figure 2-6. The scaled hyperbolic tangent function is mathematically convenient in that it is both invertible and has continuous derivatives. A new parameter has been added to scale the noisy sensor measurement x by $1/k \in \mathbb{R}^+$ and give the amplifier has peak output amplitude $c = \alpha k \in \mathbb{R}^+$ with amplifier gain-scaling constant $\alpha \in \mathbb{R}^+$ to give:

$$s(x) = \alpha k \tanh\left(\frac{x}{k}\right) \tag{4-6}$$

Determination of the proper values of α, k, and c is discussed in section 4.2.

If the channel gains between each sensor and the fusion center are unequal, this must be compensated for by channel estimation. It is assumed in this chapter that the transmitter's power control dynamic range system can bias the amplifier such that peak amplifier output as seen at the receiver is similar for all amplifiers, and thus amplifier efficiency is similar across all amplifiers for similar sensor measurements. Also, it is assumed that the predistorter can shape the amplifier response such that all amplifiers appear to be transmitting the same function through the same channel loss when observed from the fusion center. For the remainder of this chapter unless otherwise noted, the assumption is that channel estimation is working and each transmitter appears to be operating through an identical AWGN channel with a gain of 1.

The implementation of the proposed predistortion is similar to the reference techniques reviewed in section 2.2.3 except for one main difference: the target value to be produced by the combined predistorter and amplifier system is not be linear, but rather is

the limiting amplifier model behavior. On the implementation side, for the gain-based LUT predistortion technique, a linear LMS convergence rate was observed to have greater stability at a cost of some additional convergence time compared to the secant-based algorithm used in [15]. For the neural network predistorter, the desired transmitter output function, $d(\cdot)$, was set equal to the limiting amplifier model $s(\cdot)$. The training was done in accordance with Figure 2-10. For the hyperbolic tangent limiting amplifier model, the predistortion function $p(\cdot)$ was learned utilizing training inputs $(k\tanh^{-1}(v_{act}(n)/(\alpha k)))$ and training outputs $v_{pd}(n)$ according to (4-7):

$$v_{pd}(n) = p\left(k\tanh^{-1}\left(\frac{v_{act}(n)}{\alpha k}\right)\right). \tag{4-7}$$

4.2 Estimator & Performance Analysis

In the following subsection, the estimator for transmitters operating in soft-compression is derived for distributed estimation systems. The techniques for distributed estimation using bounded transmissions outlined in [25] are the starting point for developing the proposed distributed estimation system. First the assumptions made for linear AF systems are reviewed to see if they can be maintained for nonlinear AF systems. Then, estimators and performance estimates for general non-linear AF systems are derived [62,63]. Finally, specific equations are derived using measurements with a uniform sensing noise transmitted through a limiting amplifier fit to a scaled hyperbolic tangent function as described in sections 2.2.2.2 and 4.1.

When quantifying the estimate variance for linear AF systems, an assumption made was that the number of sensors $L \to \infty$. This assumption allows analyzing the signal at the receiver as a Gaussian distribution. In practical systems, there will be a finite, but large, number of sensors, $L_{\mathcal{N}}$, for which the distribution is effectively Gaussian. In a linear AF system with $L_{\mathcal{N}}$ sensors and uniform-gain AWGN channels, the receiver divides the received value by the number of sensors, $L_{\mathcal{N}}$, and the linear gain, a, to attain the estimate. Because a is generally a large constant, the receiver noise ν is generally insignificant compared to the sensing noise contribution to estimate variance and thus ignored.

When nonlinear transmitters are used, the sensor gain is no longer fixed at a generally large constant a. Instead, the gain changes as a function of the sensor measurement. For large sensor measurements operating in the highly-compressed region of amplifier operation, effective sensor gain could be very small. Thus, a small amount of noise coupled at the receiver input could result in a large estimation error. Therefore, the receiver noise is no longer ignored for a finite number of sensors, $L_{\mathcal{N}}$. The distributed estimation equations from section 2.1.1 are reanalyzed now assuming the receiver noise is Gaussian with mean μ_ν and variance σ_ν^2, such that $\nu \sim \mathcal{N}(\mu_\nu, \sigma_\nu^2)$.

The value received at the fusion center is the scaled sum of the transmissions due noisy sensor measurements, x_l, using limiting amplifier model $s(\cdot)$ plus receiver noise ν:

$$\zeta(\theta) = \frac{\nu}{L_{\mathcal{N}}} + \frac{1}{L_{\mathcal{N}}} \sum_{l=1}^{L_{\mathcal{N}}} s(x_l) \,. \tag{4-8}$$

The expected value at the fusion center, $\bar{\zeta}(\theta)$, can be found given the distribution of sensing noise at each sensor, assuming each sensor has independent identically-distributed sensing noise distributed according to $p(\eta)$, and the mean receiver noise is μ_ν:

$$\bar{\zeta}(\theta) = E[\zeta(\theta)] = \int_{\eta} p(\eta) s(\theta + \eta) d\eta + \frac{\mu_\nu}{L_{\mathcal{N}}} \tag{4-9}$$

The choice of $L_{\mathcal{N}}$ is such that the summation in (4-8) converges according to the law of large numbers, but the channel noise does not go to 0. The variance of the expected received value can be calculated as [23]:

$$\Sigma(\theta) = \int_{\eta} p(\eta) s^2(\theta + \eta) d\eta - \bar{\zeta}^2(\theta) + \frac{\sigma_\nu^2}{L_{\mathcal{N}}} \,. \tag{4-10}$$

Using [23], the asymptotic variance of the optimal estimate can be calculated to be:

$$\mathrm{AsVar}\left(\hat{\theta}\right) = \frac{\Sigma(\theta)}{\left[\frac{\partial}{\partial\theta}\bar{\zeta}(\theta)\right]^2} \quad . \tag{4-11}$$

Intuitively in (4-11), when the amplifier is highly compressed, the change in value expected at the receiver due to changes in θ becomes small. When the derivative term is small for large θ and the variance due to receiver noise is constant, the estimation variance for large θ will become very large.

As an example case, a scaled hyperbolic tangent limiting amplifier model as shown in (4-6), a uniform distribution of sensing noise bounded by $\pm b$, $\eta \sim U(-b, b)$,

and zero-mean receiver noise, $\mu_v = 0$ is solved. Under these assumptions, the value observed at the receiver is:

$$\begin{aligned}\bar{\zeta}(\theta) &= \int_{-b}^{b} \frac{1}{2b}\left[\alpha k \tanh\left(\frac{\theta+\eta}{k}\right)\right] d\eta \\ &= \frac{k^2\alpha}{2b} \ln\left(\cosh\left(\frac{b+\theta}{k}\right)\operatorname{sech}\left(\frac{b-\theta}{k}\right)\right)\end{aligned} \tag{4-12}$$

The estimate $\hat{\theta}$ can be found by replacing the expected receiver value $\bar{\zeta}(\theta)$ with the actual received value z and replacing actual θ with the estimate $\hat{\theta}$, then solving for $\hat{\theta} \in \mathbb{R}^+$. These steps find the inverse function, $\bar{\zeta}^{-1}(z)$, of the expected value, $\bar{\zeta}(\theta)$. For the hyperbolic tangent limiting amplifier model, defining constant $m = \exp(2bz/(\alpha k^2))$, and for $z \in [0, \alpha k]$:

$$\hat{\theta} = k \cosh^{-1}\left(\frac{\sinh\left(\frac{b}{k}\right)(1+m)}{\sqrt{2m\cosh\left(\frac{2b}{k}\right) - 1 - m^2}}\right). \tag{4-13}$$

It should be noted here that the estimate in (4-13) is the same solution to the optimization problem in (2-7), and the condition $z \in [0, \alpha k]$, is satisfied almost surely for a sufficiently large number of sensors $L_{\mathcal{N}}$ and small consequently $\sigma_v^2/L_{\mathcal{N}}$.

The variance of the received value at the fusion center's estimator, $\Sigma(\theta)$, can be computed for the uniformly distributed sensing noise as:

$$\begin{aligned}\Sigma(\theta) &= \int_{-b}^{b} \frac{1}{2b}\left[\alpha k \tanh\left(\frac{\theta+\eta}{k}\right)\right]^2 d\eta - \bar{\zeta}^2(\theta) + \frac{\sigma_v^2}{L_{\mathcal{N}}} \\ &= \frac{(\alpha k)^2}{2b}\left(2b - k\tanh\left(\tfrac{b-\theta}{k}\right) - k\tanh\left(\tfrac{b+\theta}{k}\right)\right) - \bar{\zeta}^2(\theta) + \frac{\sigma_v^2}{L_{\mathcal{N}}}.\end{aligned} \tag{4-14}$$

The derivative of $\bar{\zeta}(\theta)$ can be found using (4-12) to be:

$$\frac{\partial\bar{\zeta}(\theta)}{\partial\theta} = \frac{\alpha k}{2b}\left[\tanh\left(\frac{b+\theta}{k}\right) + \tanh\left(\frac{b-\theta}{k}\right)\right]. \tag{4-15}$$

Substituting (4-14) and (4-15) into (4-11) yields the asymptotic variance of the estimate calculated using a scaled hyperbolic tangent function at specific values of θ measured in the presence of uniformly distributed zero-mean sensing noise:

$$\mathrm{AsVar}\left(\hat{\theta}\right) = \frac{2b\left[2b - k\tanh\left(\frac{b-\theta}{k}\right) - k\tanh\left(\frac{b+\theta}{k}\right)\right] - k^2\left[\ln\left(\cosh\left(\frac{b+\theta}{k}\right)\mathrm{sech}\left(\frac{b-\theta}{k}\right)\right)\right]^2 + \frac{4b^2}{(\alpha k)^2}\frac{\sigma_v^2}{L_{\mathcal{N}}}}{\left[\tanh\left(\frac{b+\theta}{k}\right) + \tanh\left(\frac{b-\theta}{k}\right)\right]^2} \tag{4-16}$$

The scaled hyperbolic tangent function becomes equivalent to the linear function when $\alpha = a$ and $k \to \infty$. Evaluating from (4-16), $\lim_{k\to\infty} \mathrm{AsVar}\left(\hat{\theta}\right) = b^2/3$, yielding as expected the same estimate variance result as attained by using the perfectly linear model.

A design tradeoff exists when determining the proper value of k for sensor measurements. Smaller k allows the sensor to operate at higher output powers and thus attain higher efficiency, but at the cost of increased estimation variance. As demonstrated previously large k selection approximates linear amplification, and thus ultimately yield estimation accuracy equivalent to a linear amplifier, but the highly efficient compressed region of operation will not be utilized. As k is decreased from ∞, at first the measurement noise will dominate estimate variance. Eventually, when k

reaches some threshold, the receiver noise, ν, will dominate and the asymptotic variance increases dramatically. The best point of operation from an efficiency standpoint is the smallest k for which estimate variance is determined by the sensing noise. The k value at which this transition occurs will depend on the specific sensing noise distribution. For a given amplifier type, the maximum amplifier output, c, is typically fixed. Thus, the α can be determined given k via $\alpha = c/k$.

For Gaussian and Cauchy sensing noise sources, equation (4-11) holds but the integrals cannot be easily evaluated in a closed form. Simulations showing the performance attainable with Gaussian and Cauchy sensor noise distributions are shown in section 4.3.4, using numerical integration to determine asymptotic variance of the estimate for those sensing noise distributions.

4.3 Verification and Results

To verify the proposed estimation algorithm, simulations were performed in MATLAB. Initial simulations to validate the analytical solutions (4-13) to (4-16) were performed with uniform noise passed through perfectly predistorted amplifiers. This "perfect" predistortion was attained by using the targeted scaled hyperbolic tangent function itself as the amplifier model.

To attain a realistic profile of PAE versus amplifier output power, an example class-AB amplifier from the PA design guide of Agilent Advanced Design System (ADS) was simulated using the ADS harmonic balance simulator [64]. PAE was calculated by

taking the ratio of output RF fundamental-frequency power to the sum of input RF and input DC power as input RF power was swept. PAE was assembled into a list indexed by the output power at the fundamental frequency. In the MATLAB estimation simulations, PAE is determined by performing a table look-up based on the fundamental-frequency output power. PAE and fundamental-frequency gain for this amplifier simulation are shown in Figure 1-1.

4.3.1 Selection of Hyperbolic Tangent Scaling Factors

To use the scaled hyperbolic tangent model to do nonlinear estimation, the proper values to use for α and k must be determined. First, the limiting value of the amplifier $c = \alpha k$ can be selected to be just below the amplifiers maximum output amplitude. In this case, based on the actual amplifier simulated in ADS, the value of $c = \alpha k = 7.26$.

In simulation, θ was chosen to range from $^{6}/_{32}$ to 6, with sensing noise, η, chosen to be uniform bounded by $\pm b$, with $b = {}^{3}/_{32}$. These selections were made to roughly mimic 5-bits of accuracy in a converter. It is demonstrated via simulation in subsection 4.3.4 that this performance is equivalent to analog sensor over the same range with Gaussian noise variance $b^2/3$. The receiver noise was Gaussian with variance such that $\sigma_v^2/L_{\mathcal{N}} = 10^{-7}$. The sensor measurements were allowed to be equally likely across the anticipated range to determine expected PAE.

To find the best value for measurement scaling factor k, simulations were run across the full range of expected θ with many realizations of sensing noise. It is known

that for large k the sensor approximates linear operation, and the estimate asymptotic variance should approach $b^2/3$. This was confirmed by simulation for large k, as shown in Figure 4-1, where at $k = 40$, $\mathrm{AsVar}(\hat{\theta}) = b^2/3$. The rapid rise in estimate asymptotic variance at low k given non-zero σ_v^2 is predicted by (4-16), which is experimentally verified in section 4.3.2. In Figure 4-1 under the simulated conditions, the lowest value of k that maintains worst case $\mathrm{AsVar}(\hat{\theta}) \approx b^2/3$ over the entire expected range of θ is $k = 1.75$, thus setting the nominal k value used throughout this chapter unless otherwise noted. Using $k = 1.75$ with $c = 7.26$, requires that $\alpha = 4.15$ because $c = \alpha k$.

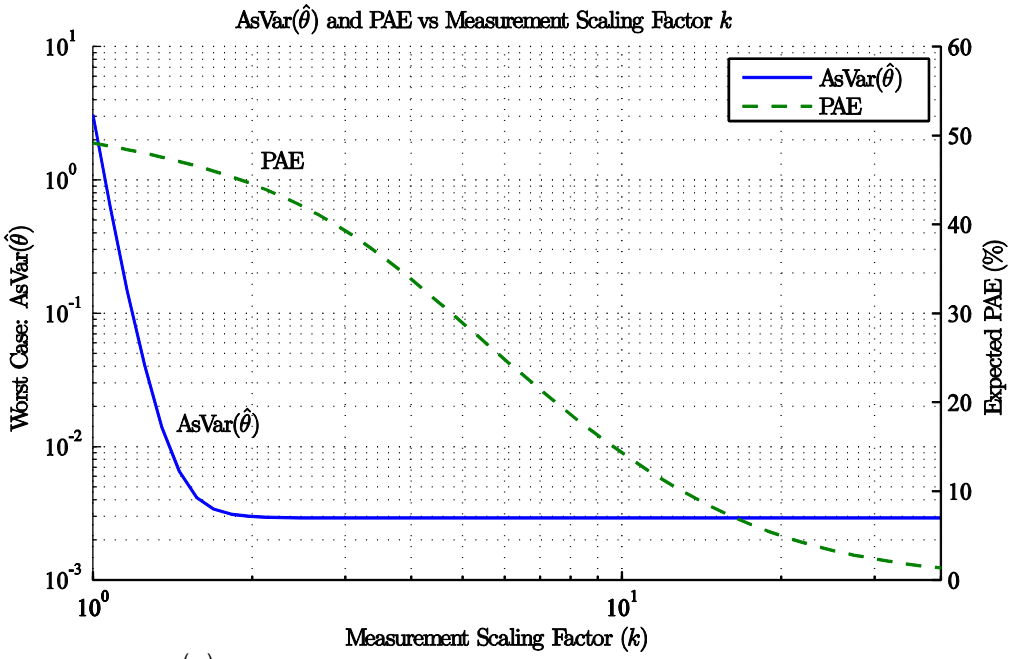


Figure 4-1: $\mathrm{AsVar}(\hat{\theta})$ and PAE versus scaling k. Optimal PAE is attained at smallest k where estimate variance is sufficiently small.

4.3.2 Experimental Verification of Analytical Results

Simulations with both $L = 10$ and $L = 1000$ sensors were used to verify the expected value received at the fusion center, $\bar{\zeta}(\theta)$, and asymptotic variance of the value received at the fusion center, $\Sigma(\theta)$, predicted by equations (4-12) and (4-14). The sensor measurement noise, η_l, used for each sensor was uniformly distributed from $-b$ to b with $b = 0.25$ in this simulation. The resulting simulations are shown in Figure 4-2. It is seen that the experimentally generated values of $\bar{\zeta}(\theta)$ match the analytically determined values. Similarly, the experimentally determined value $L\,\mathrm{Var}(z)$ matches the analytically determined asymptotic variance.

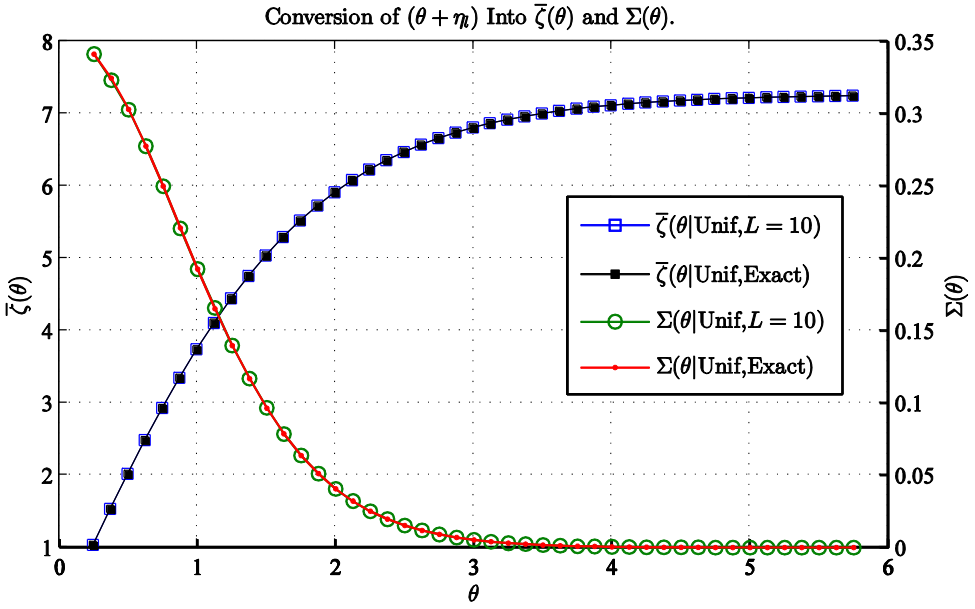


Figure 4-2: $\bar{\zeta}(\theta)$ and $\Sigma(\theta)$ with Uniform Sensing Noise

Simulations were performed using perfect predistortion to experimentally verify the analytic formulae (4-13) to (4-16). Unless otherwise noted, the simulations used α, k,

b, and $\sigma_\nu^2/L_{\mathcal{N}}$ values as specified in section 4.3.1. The values attained from simulation using $L_{\mathcal{N}} = 10$ sensors proved identical to the analytic formulae, as can be seen from Figure 4-3 and Figure 4-4. In these figures, markers indicate the values attained from simulation with $L_{\mathcal{N}} = 10$ sensors, and the solid lines indicate values attained from simulation. These plots are used to demonstrate the tradeoffs between k, $\sigma_\nu^2/L_{\mathcal{N}}$, and achievable estimate asymptotic variance $\mathrm{AsVar}(\hat{\theta})$.

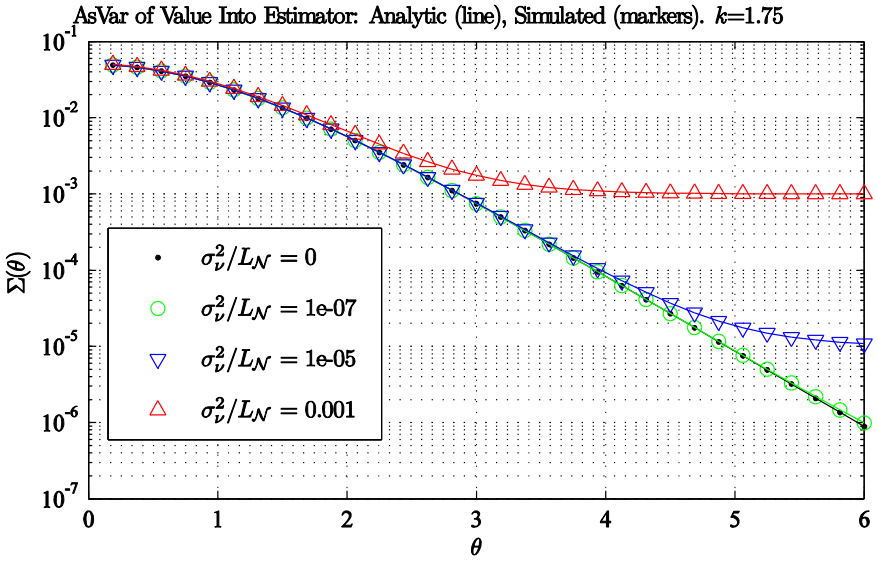


Figure 4-3: Asymptotic Variance of Value Received for Estimator, $\Sigma(\theta)$. Note that receiver noise establishes a noise floor not present when only using sensing noise.

Some key observations can be made from these plots. First, from (4-14) it can be seen that receiver noise introduces an additive term, which is not a function of θ in $\Sigma(\theta)$, that establishes a floor on the asymptotic variance $\Sigma(\theta)$ of the quantity $\zeta(\theta)$ provided to the estimator. This noise floor is demonstrated for several different receiver noise strengths in Figure 4-3, and serves to limit the maximum θ for which accurate estimates

may be attained. A plot of estimate asymptotic variance, $\mathrm{AsVar}(\hat{\theta})$, for the same conditions is shown in Figure 4-4. It can be seen that once the receiver noise floor determines $\Sigma(\theta)$, accurate estimates are no longer attainable. These plots show both results obtained from the analytic formulas derived in section 4.2 as solid lines, and results obtained via simulation as markers.

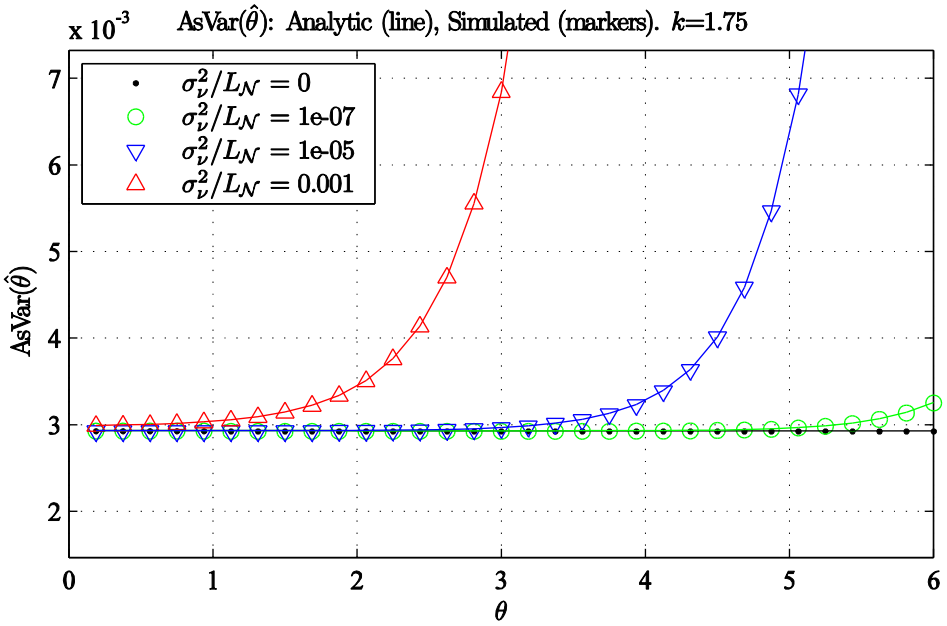


Figure 4-4: Asymptotic Variance of the Estimate. The flat receiver noise causes trivial changes in $\mathrm{AsVar}(\hat{\theta})$ until some threshold value of θ above which large increases in $\mathrm{AsVar}(\hat{\theta})$ occur.

Aside from reducing $\sigma_\nu^2/L_\mathcal{N}$, estimate asymptotic variance, $\mathrm{AsVar}(\hat{\theta})$, can be reduced by increasing the measurement scaling factor, k, as shown in Figure 4-5. This has the effect of moving amplifier operation back closer to the linear region. From a performance standpoint, it is equivalent to scaling down the values of θ, but this improvement in $\mathrm{AsVar}(\hat{\theta})$ comes at a cost of decreased PAE as shown previously in

section 4.3.1. Here again the values attained from the analytic equation (4-16) match values attained from simulation.

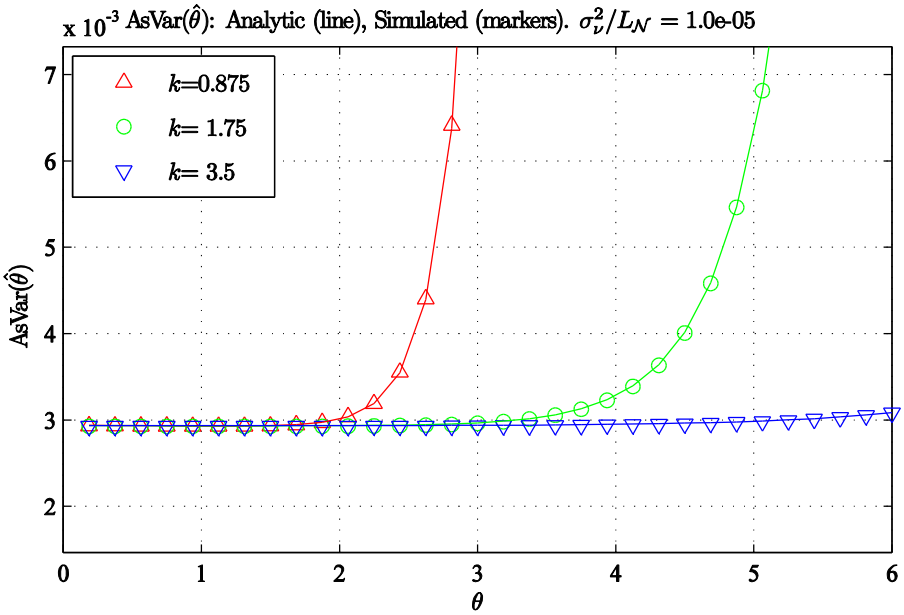


Figure 4-5: AsVar$\left(\hat{\theta}\right)$ Improves at a fixed receiver noise level by increasing the scaling factor, k, by which measurements are divided.

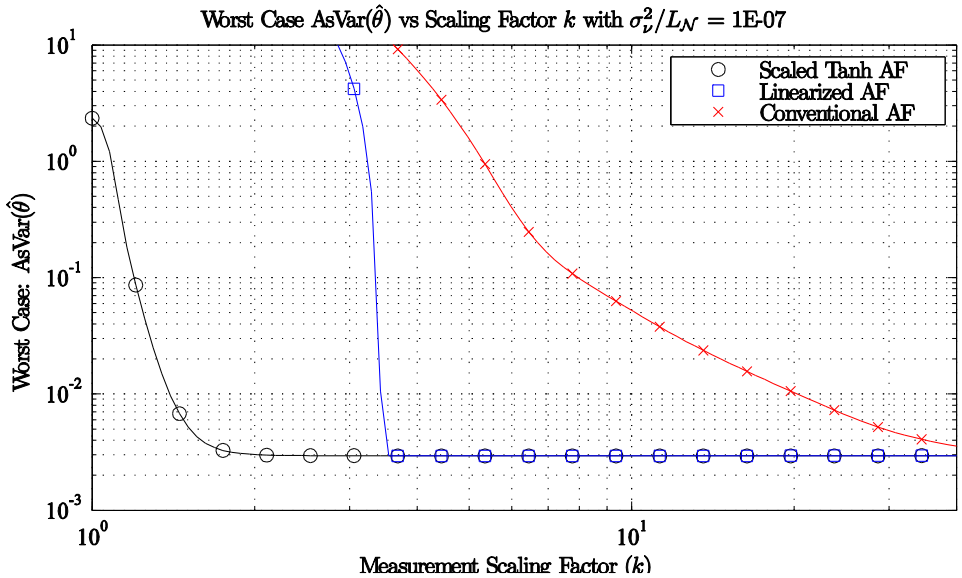
4.3.3 Efficiency Performance Vs. Conventional Techniques

The major benefit of the proposed technique is the ability to obtain accurate estimates while pushing further into the non-linear region of the amplifier. To demonstrate this, simulations were run to determine the asymptotic variance of $\hat{\theta}$ for three different configurations of AF distributed sensing networks. The first configuration simulates the reference technique: conventional AF over an uncompensated amplifier. The estimator assumes linear transmitters. In the second technique, the amplifier has been linearized to make it perform perfectly linearly up to its maximum output power. The estimator still assumes linear transmitters. The third technique simulated is the

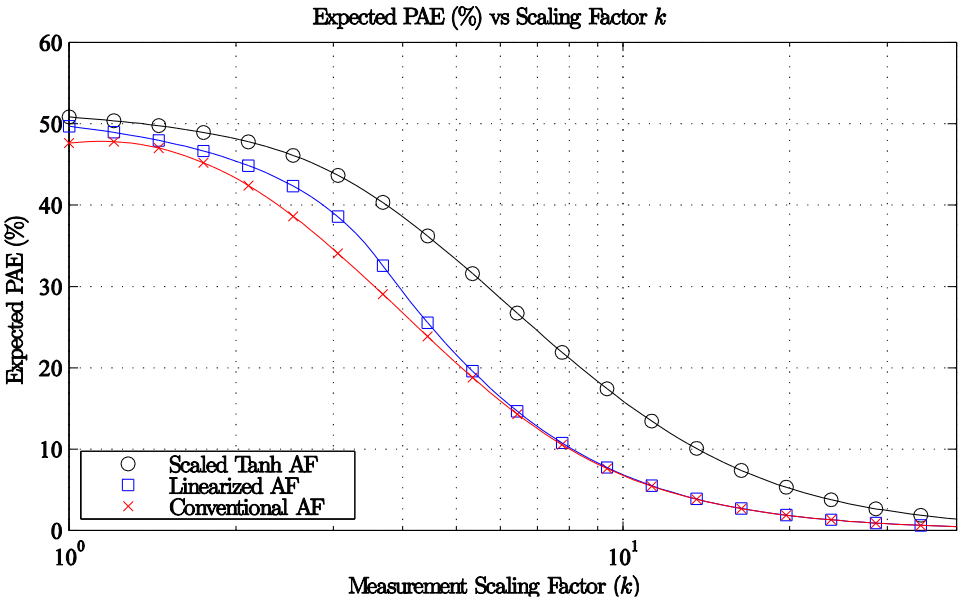
proposed technique, with the amplifier predistorted to fit a scaled hyperbolic tangent and estimation done according to (4-13).

Each of the three techniques was simulated using the same receiver noise characteristics to indicate the effect of receiver noise sensitivity which is a potential vulnerability of the proposed approach relative to the two reference techniques. Additionally, the measurement scaling factor k was swept to make sure the comparison included the best scaling factor possible for each of the three techniques to ensure a fair comparison. The performance at each scaling factor was quantified via two metrics. The first metric was the worst-case asymptotic variance of $\hat{\theta}$ across the range of θ. This value should be kept as low as possible. The second metric was the power-added efficiency across θ, obtained by the ratio of all power transmitted at the fundamental frequency to all power consumed across θ for each scaling factor. PAE should be as high as possible.

It can be seen in Figure 4-6 that the proposed technique is able to provide $\mathrm{AsVar}(\hat{\theta})$ limited by sensing measurement noise while achieving a PAE around 48% with the simulated class AB amplifier when receiver noise $\sigma_v^2/L_{\mathcal{N}} = 10^{-7}$. Achieving estimate variance with a perfectly linearized amplifier, only 33% PAE was achievable. Without linearization, using the amplifier in deep class A by using large k results in PAE around 2%.



(a) Worst Case Asymptotic Variance vs. Measurement Scaling k



(b) PAE vs. Measurement Scaling k

Figure 4-6: Performance across k and Estimation Algorithm. Best performance is for k with AsVar$\left(\hat{\theta}\right)$ low while PAE is high.

In cases where the sensor network can tolerate $\mathrm{AsVar}(\hat{\theta})$ greater than the $b^2/3$ limit imposed by the sensing noise, performance of the proposed technique degrades more slowly with decreases in k than does the linearized AF technique. The performance degradation is shown by the slope of the worst-case $\mathrm{AsVar}(\hat{\theta})$ shown in Figure 4-6(a), with steeper slopes indicating more sudden performance degradation. The sudden decrease in performance for linearized amplifiers is because once linearized amplifiers reach their maximum output amplitude, any additional information about how far beyond maximum power would have been transmitted is fully lost. In the proposed algorithm using the limiting amplifier models, the performance gracefully degrades as $\partial\bar{\zeta}(\theta)/\partial\theta$ decreases.

Conventional linear AF systems have the slowest performance degradation with decrease in k, but power-added efficiency is unacceptably low. This is because once any amplifier reaches compression, the assumptions made by the estimator start to fail. When the amplifier first encounters compression, the impact on performance is light as the amount of compression is small. As compression increases, the performance falls further away from the desired value. Compression can only be avoided by increasing the amplifier scaling factor k, which decreases the power the amplifier is transmitting and in turn decreases PAE.

When receiver noise $\sigma_v^2/L_{\mathcal{N}}$ is increased, the margin between linearized AF and the proposed technique decreases, but PAE remains a few percent higher. The performance of the two techniques becomes similar at approximately $\sigma_v^2/L_{\mathcal{N}} = 10^{-3}$.

4.3.4 Performance for Gaussian and Cauchy Sensor Noise

To verify the performance of the estimation technique for other sensor noise distributions, simulations were performed with both Gaussian and Cauchy distributed sensing noise sources. The parameters $\bar{\zeta}(\theta)$ and $\Sigma(\theta)$ determine the asymptotic variance of the best attainable estimator using (4-11). These two values are plotted and used to attain $\mathrm{AsVar}(\hat{\theta})$.

For Cauchy distributed noise sources in conventional AF systems, it was determined in section 2.1.1 that a consistent estimator does not exist. The expected value of data at the receiver with the proposed scheme is shown in Figure 4-7. Using the proposed limiting amplifier models with Cauchy sensing noise, $\bar{\zeta}(\theta)$ was consistent and well-defined. This is an advantage over the reference techniques.

From Figure 4-7, it can be seen that with comparable sensor noise variances, changes in the value expected at the receiver, $\bar{\zeta}(\theta)$, versus distribution are small. Thus, rough estimates for θ can be attained using the analytical estimator (4-13) from uniformly distributed sensor noise with comparable variance.

From Figure 4-8, it can be seen that with $L = 10$ sensors, the uniformly distributed and Gaussian distributed sensing noises of the same sensing noise variance

have almost identical asymptotic variance at the receiver, $\Sigma(\theta)$, as alluded to in section 4.2. Thus, similar estimator performance is to be expected, and this is confirmed in Figure 4-9. For the Cauchy distribution, $\Sigma(\theta)$ is much higher. Consequently, the performance of its estimator is worse and is only valid for small θ. This restriction requiring low θ can be bypassed by using a higher measurement scaling value k, but at a power-added efficiency cost.

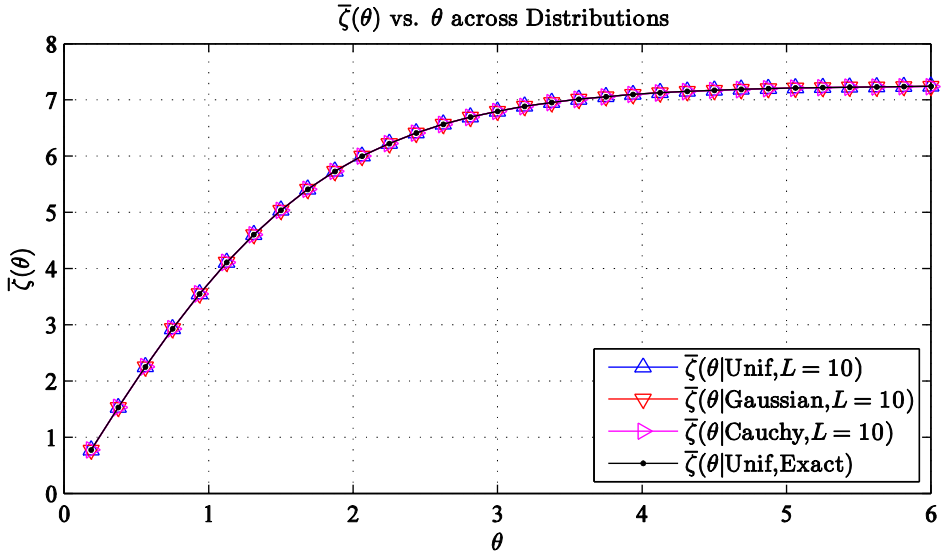


Figure 4-7: $\bar{\zeta}(\theta)$ for different sensor noise distributions. Sensor noise distribution does not induce large change in $\bar{\zeta}(\theta)$

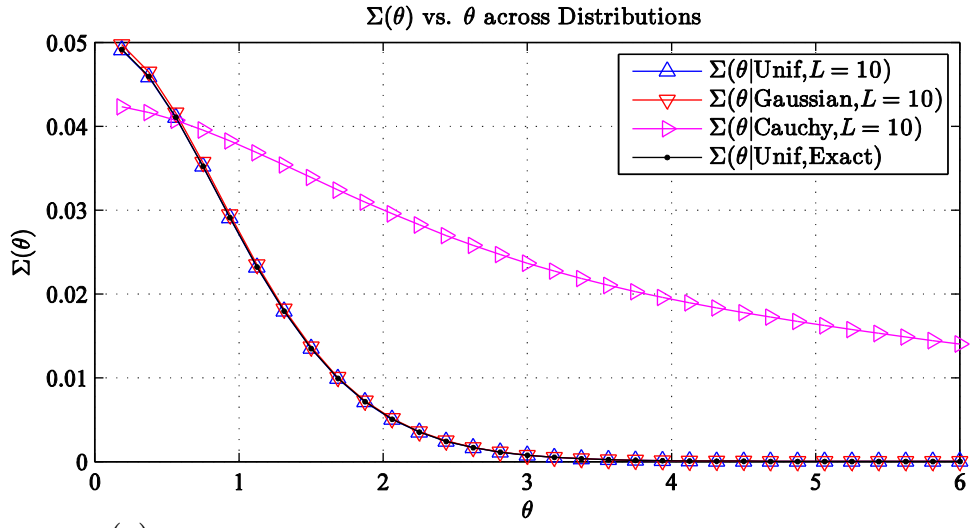


Figure 4-8: $\Sigma(\theta)$ for different sensing noise distributions. Gaussian and Uniform sensing noises of similar variance have similar $\Sigma(\theta)$.

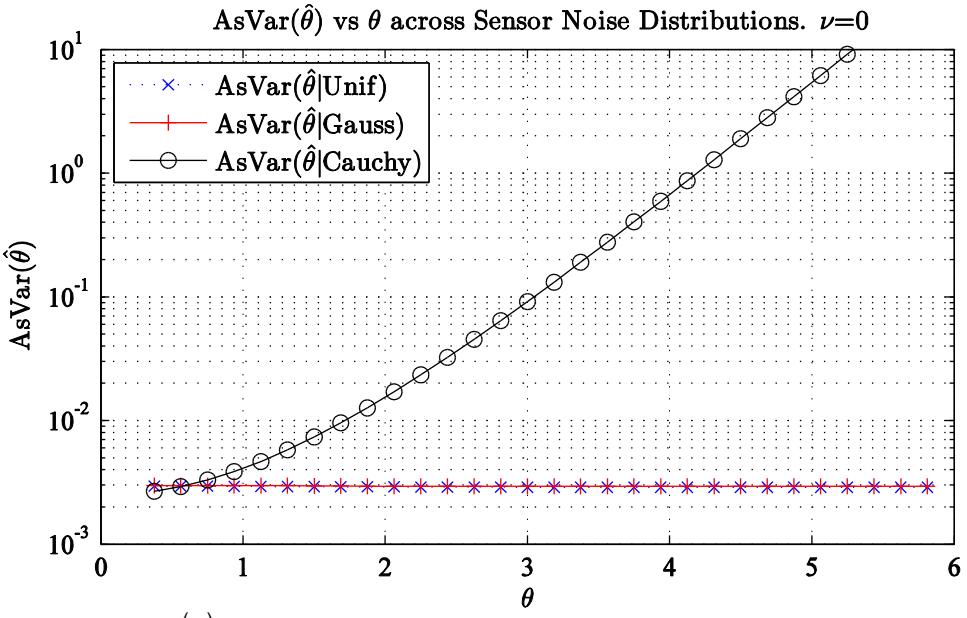


Figure 4-9: AsVar$\left(\hat{\theta}\right)$ vs θ for different sensor noise distributions. Cauchy estimates can be seen to be less accurate at larger θ than Gaussian or uniform noise distributions.

4.3.5 Predistortion Application & Efficiency Gains

In this section it is verified that it is possible to fit amplifier behavior to a scaled hyperbolic tangent function using predistortion, and thus extend the range of sensor values that can be successfully estimated well into the amplifier's more power efficient non-linear class AB region of operation. To achieve this, the previously discussed ADS amplifier model was utilized.

Three different predistorter configurations were simulated. The first two configurations used the gain-based look-up-table technique, the first with 32 bins and the second with 128 bins. The third configuration used a neural network where the first layer consisted of 3 hyperbolic tangent neurons and the second layer was a single linear neuron. The performance of each predistorter implementation is shown in Figure 4-10. The predistorter imperfections manifest themselves similarly to modest receiver noise.

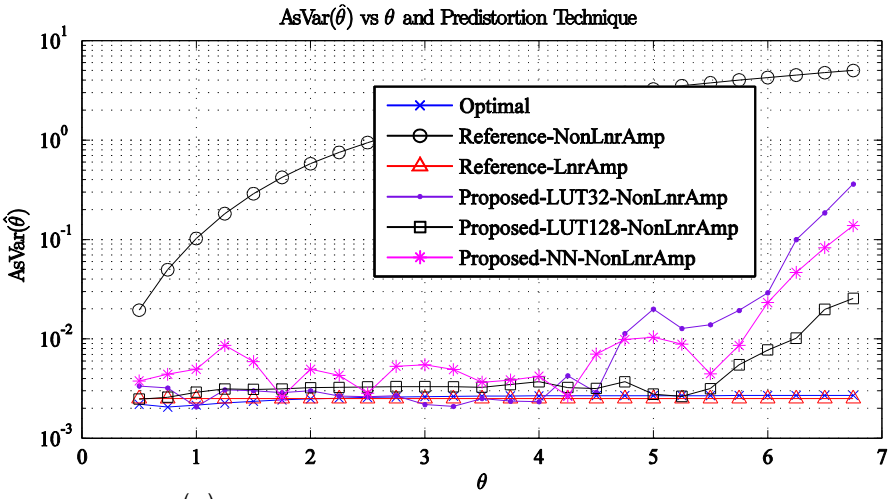


Figure 4-10: AsVar$\left(\hat{\theta}\right)$ vs. θ. Good estimates are still possible even using realizable predistortion techniques.

4.4 Estimation Using a Cann Limiting Amplifier Model

As previously discussed in sections 4.1 and 4.2, the Cann model is another limiting amplifier that meets the requirements for use in amplify-and-forward distributed estimation outlined in section 4.1. This model has three parameters. Two of them are similar to the scaled hyperbolic tangent function, the limiting value $c = \alpha k$ and uncompressed gain $1/k$ also known as measurement scaling factor k. The third parameter, s_C, is specific to the Cann model [16,65] and controls the sharpness of the transition from the linear region to the limited region. The actual amplifier modeling function is:

$$s(x) = \frac{c}{\left[1 + \left(\frac{c}{x/k}\right)^{s_C}\right]^{\frac{1}{s_C}}} = \frac{x/k}{\left[1 + \left(\frac{x/k}{c}\right)^{s_C}\right]^{\frac{1}{s_C}}} \quad where\ x > 0 \tag{4-17}$$

Similarly, to the scaled hyperbolic tangent function approaches, the expected value received at the fusion center using a Cann limiting amplifier model and sensing noise uniformly distributed from –*b* to *b* can be computed according to (4-9) as [66]:

$$\begin{aligned} &\bar{\zeta}(\theta) \\ &= \frac{1}{4bk}(\theta + b)^2\ {}_2F_1\left(\frac{1}{s_C}, \frac{2}{s_C}, \frac{2 + s_C}{s_C}; -\left[\frac{1}{ck}(\theta + b)\right]^{s_C}\right) \\ &- \frac{1}{4bk}(\theta - b)^2\ {}_2F_1\left(\frac{1}{s_C}, \frac{2}{s_C}, \frac{2 + s_C}{s_C}; -\left[\frac{1}{ck}(\theta - b)\right]^{s_C}\right), \end{aligned} \tag{4-18}$$

where ${}_2F_1(t, u, v; w)$ is the ordinary hypergeometric function as defined as [67]:

$$ {}_2F_1(t,u,v;w) = \frac{\Gamma(v)}{\Gamma(t)\Gamma(u)} \sum_{n=0}^{\infty} \frac{\Gamma(t+n)\Gamma(u+n)}{\Gamma(v+n)} \frac{w^n}{n!}. \tag{4-19} $$

The variance expected about the value received at the fusion center subject to the same uniform sensing noise can be computed according to (4-10) as:

$$ \begin{aligned} &\Sigma(\theta) \\ &= \frac{1}{6bk^2}(\theta+b)^3 \, {}_2F_1\left(\frac{2}{s_C}, \frac{3}{s_C}, \frac{3+s_C}{s_C}; -\left[\frac{1}{ck}(\theta+b)\right]^{s_C}\right) \\ &- \frac{1}{6bk^2}(\theta-b)^3 \, {}_2F_1\left(\frac{2}{s_C}, \frac{3}{s_C}, \frac{3+s_C}{s_C}; -\left[\frac{1}{ck}(\theta-b)\right]^{s_C}\right) - [\bar{\zeta}(\theta)]^2 \\ &+ \frac{\sigma_v^2}{N}. \end{aligned} \tag{4-20} $$

Finally, the derivative of the expected value received at the fusion center is required to compute the asymptotic variance of the estimator utilizing (4-11). The derivative is:

$$ \frac{\partial \bar{\zeta}(\theta)}{\partial \theta} = \frac{(\theta+b)}{2bk\left(1+\left[\frac{1}{ck}(\theta+b)\right]^{s_C}\right)^{\frac{1}{s_C}}} - \frac{(\theta-b)}{2bk\left(1+\left[\frac{1}{ck}(\theta-b)\right]^{s_C}\right)^{\frac{1}{s_C}}} \tag{4-21} $$

Unlike the scaled hyperbolic tangent model, the expected received value function was not straightforward to invert, and thus a table look-up or other function approximation is to be used for determining the estimate $\hat{\theta}$ from the received value *z*.

If the estimator is simulated without receiver noise, with perfect predistortion, and noiseless unity-gain channels, all of the techniques can produce good estimates across the full-range of θ until both $\partial\bar{\zeta}(\theta)/\partial\theta$ and $\Sigma(\theta)$ become so nearly zero that the asymptotic variance calculation becomes numerically unreliable. In the example shown in Figure

4-11, $b = 0.25$, meaning the variance of what is being measured itself is $\sigma_\eta^2 = b^2/3 = 0.0208$.

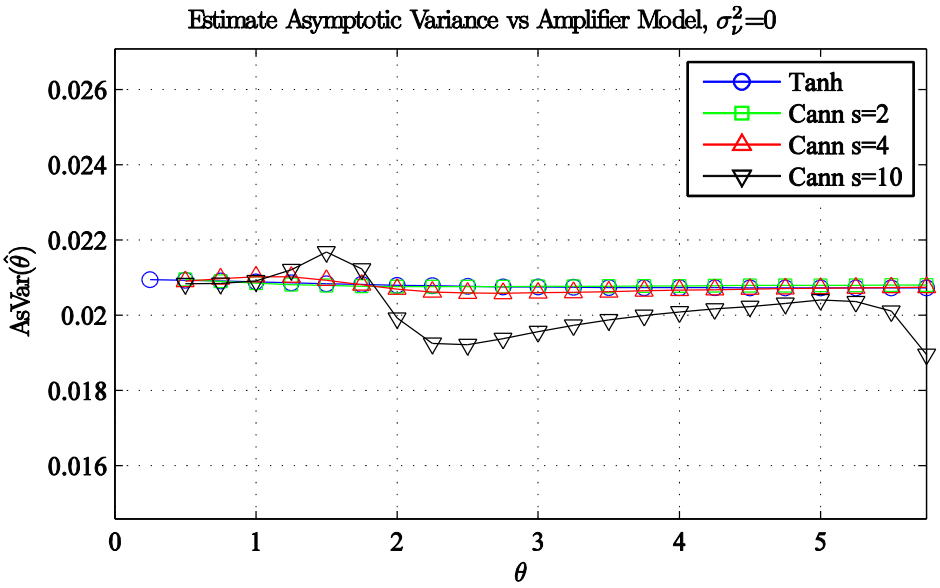


Figure 4-11: Estimate Asymptotic Variance when $\sigma_v^2 = 0$.

Assuming appropriate choices are made for measurement scaling k and amplifier limiting value c, then the asymptotic variance of the estimator is determined largely by the receiver noise. The best limiting amplifier model function to select is the sharpest one that still enables identification. An example plot showing how receiver noise impairs estimation performance is shown in Figure 4-12.

According to Figure 4-12, it can be seen that subject to receiver noise the Cann model with sharpness parameter s_C=10, the signal is too high up into the compressed region for the estimator to distinguish values in the presence of receiver noise. Accordingly, increasing the measurement scaling k may enable operation with low

estimate asymptotic variance over the full-range of anticipated sensor measurements θ. To verify this, simulations were run to determine PAE and asymptotic variance for many different $1/k$ values across the entire anticipated range of θ. The best model type and gain combination is the one with an acceptable worst case of estimate asymptotic variance while maintaining the highest PAE. For this specific model, the best would be the Cann model with a low sharpness parameter of s_C=2, as shown in Figure 4-13

To verify the analytical formulations made in (4-18) through (4-21), the analytically determined values from the equations were compared to experimentally determined values. It can be seen that the values align nicely for the Cann model across a large range of sharpness parameters.

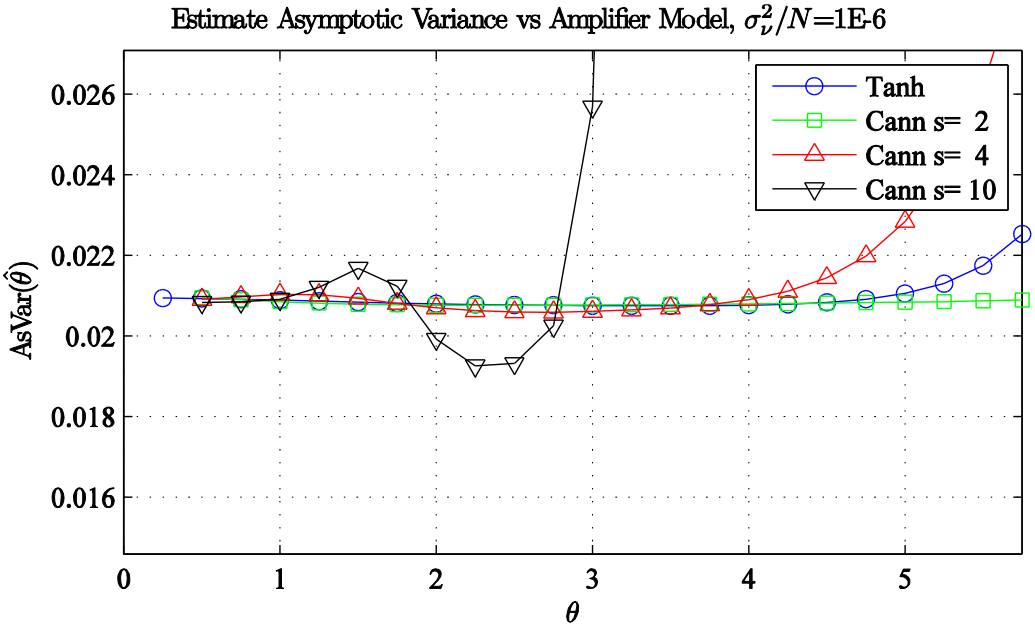


Figure 4-12: Estimate Asymptotic Variance when $\sigma_\upsilon^2 / L \quad = 10^{-6}$

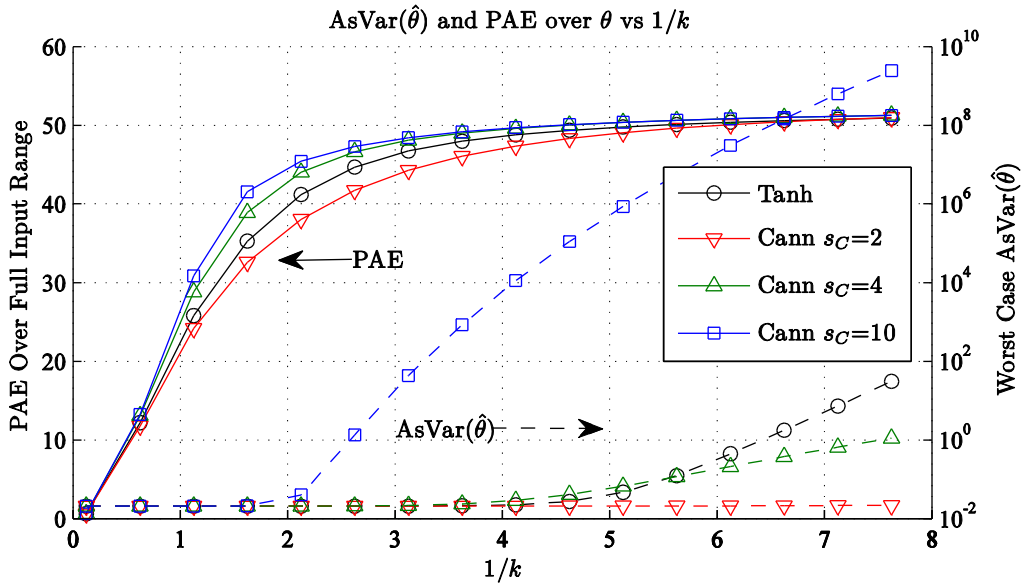


Figure 4-13: Estimate Asymptotic Variance and PAE when $\sigma_v^2 / L \quad = 10^{-6}$

5 ESTIMATION WITH NON-UNIFORM CHANNELS

Distributed estimation uses multiple inexpensive sensors to estimate a single quantity. In previous literature, many methods have been developed for transmitting measured data back to a fusion center [1,5,6,17]. For each of the methods, much analysis has been done to determine how to allocate power amongst the sensors for a variety of channel conditions [18]. One common algorithm used in the literature is amplify-and-forward (AF) distributed estimation, where sensors transmit a value based on a measured noisy parameter. When these amplifiers transmit simultaneously over a shared channel, that method is called AF over a coherent multiple access channel (MAC) [7,8,19]. This technique is summarized in section 2.1 and illustrated in Figure 2-2. For most analyzed AF systems, linear transmit amplifiers have been assumed. While linear amplifiers make analytical optimization of transmitted power tractable, they have poor power-added efficiency (PAE). In chapter 4, a technique utilizing efficient nonlinear amplifiers in an AF system with coherent MAC was demonstrated when the channels had been equalized to appear as equivalent-gain AWGN channels.

In this chapter, estimators and performance expectations are derived for utilizing nonlinear transmitters in an AF system operating over AWGN channels that have not been equalized to have constant gain. This approach merges the transmitter power allocation techniques from existing literature [1,5,6,11] with the techniques for utilizing efficient nonlinear amplifiers introduced in chapter 4. Real world systems utilizing AWGN channels of varying gains may exist where the fusion center has line-of-sight

paths to each sensor being the dominant path of signal propagation. These environments are assumed to have low multipath levels. In the first two subsections, estimator performance for nonlinear AF transmissions over non-equal AWGN channels is analyzed. In subsection 5.1, estimator performance is derived assuming only gain statistics, and full phase-information is available at the transmitter for the corresponding channel. In subsection 5.2, performance and power allocation is analyzed when full channel gain statistics are available.

In the last scenario, it is assumed that sensors are operated in a dynamic multipath-fading channel environment. In this environment, the channel characteristics must be derived as the data is being transmitted. This may be the case of transmitters operating in a busy urban environment, where assumptions about low multipath and slowly varying path gains may not hold. In section 5.3, a technique utilizing OFDM with multiple access (OFDMA) to simultaneously implement both coherent multiple-access and orthogonal-signaling AF distributed estimation is introduced. It is possible in this technique to implement amplify-and-forward distributed estimation and determine, in real time, if better performance can be expected from coherent multiple access or orthogonal signaling techniques. Additionally, the pilot tones used in this OFDMA system may render it possible to localize estimates in AF, utilizing wireless sensor network localization techniques [68]. Depending on the spatial resolution required, it may even be possible to dynamically allocate sensors to separate common-subcarrier pools once sensor location has been established.

5.1 Phase-Only Channel State Information

5.1.1 Network Topology, Assumptions, and Analysis

A typical distributed estimation topology for utilizing AF distributed estimation over a coherent MAC is shown in Figure 5-1. In this scenario, it is assumed that only channel gain statistics are available and not individual path gains. It is also assumed that full state information is available at each sensor for achieving phase alignment at the fusion center. Each transmitter amplifies a noisy sensor measurement, x_l, through a common limiting amplifier model meeting requirements described in 4.2, $s(x_l)$, and output power capability determining gain, g_l.

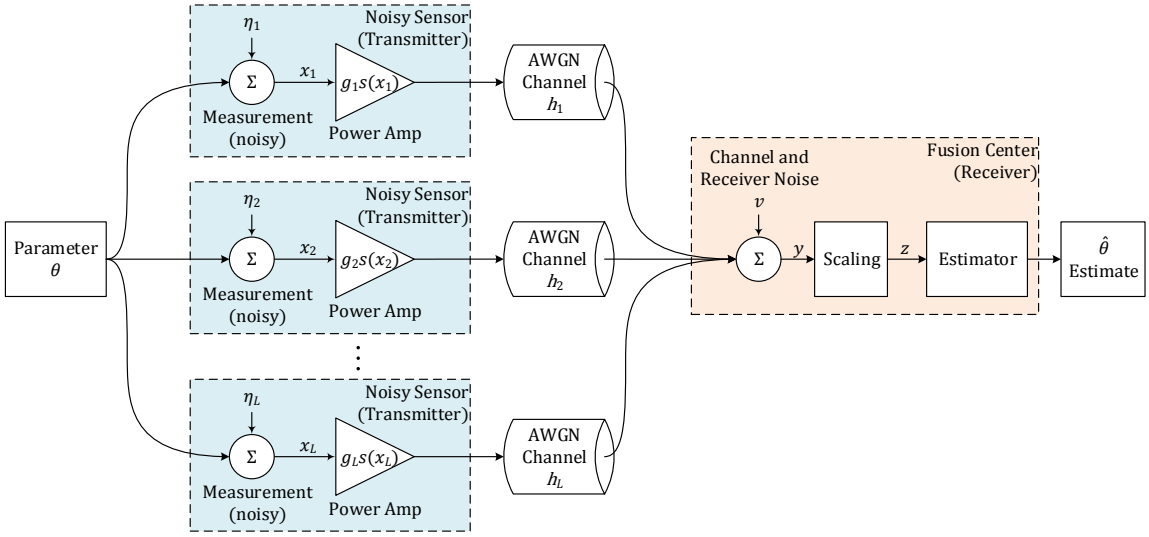


Figure 5-1 :Distributed estimation topology with random channels and sensor gains

For this scenario, the value received at the fusion center is:

$$y(\theta) = \sum_{l=1}^{L} h_l g_l s(\theta + \eta_l) + \nu \quad , \tag{5-1}$$

where g_l and h_l are assumed not to be independent across the same transmitter, but are independent across different transmitters. Each sensing measurement noise, η_l, is iid and assumed to be independent from channel noise, ν. The expected value received at the fusion center is:

$$E[y(\theta)] = \sum_{l=1}^{L} E[h_l g_l] E[s(\theta)] + \mu_\nu \ , \tag{5-2}$$

where μ_ν is the mean of the channel noise, usually zero. Assuming all channels have iid total-gain distributions $h_l g_l \sim hg$, and the channel and receiver noise is zero mean, $\mu_\nu = 0$:

$$E[y(\theta)] = L\, E[hg] E[s(\theta)]. \tag{5-3}$$

The derivative of $E[y(\theta)]$ with respect to θ is:

$$\frac{\partial E[y(\theta)]}{\partial \theta} = LE[hg] E\left[\frac{\partial s(\theta)}{\partial \theta}\right]. \tag{5-4}$$

The variance of the expected received value at the fusion center is:

$$\mathrm{Var}[y(\theta)] = \sigma_\nu^2 + L(E[h^2 g^2] E[s^2(\theta)] - E^2[hg] E^2[s(\theta)]) \, . \tag{5-5}$$

Consequently, the variance of the parameter estimate expected at the fusion center is according to [25] :

$$\mathrm{Var}\left(\hat{\theta}\right) = \frac{\mathrm{Var}[y(\theta)]}{\left[\frac{\partial}{\partial\theta}E[y(\theta)]\right]^2}\ . \tag{5-6}$$

To have a fair metric for comparing estimator performance independent of the number of sensors, asymptotic expectation and variance values were developing using $z = y/L$. Using this substitution and methods similar to those used in section 4.2, it can be found that:

$$\bar{\zeta}(\theta) = E[z] = E[hg]E[s(\theta)], \tag{5-7}$$

$$\Sigma(\theta) = L\,\mathrm{Var}[z] = E[h^2g^2]E[s^2(\theta)] - E^2[hg]E^2[s(\theta)] + \frac{\sigma_v^2}{L}. \tag{5-8}$$

By swapping the order of the expectation and derivative, which are taken with respect to different variables, it can be found that:

$$\frac{\partial\bar{\zeta}(\theta)}{\partial\theta} = E[hg]E\left[\frac{\partial s(\theta)}{\partial\theta}\right]. \tag{5-9}$$

By substitution of (5-8) and (5-9) into (4-11), a formula can be determined for asymptotic variance of the estimate.

$$\mathrm{AsVar}\left(\hat{\theta}\right) = \frac{\Sigma(\theta)}{\left[\frac{\partial\bar{\zeta}(\theta)}{\partial\theta}\right]^2} \tag{5-10}$$

$$\mathrm{AsVar}\left(\hat{\theta}\right) = \frac{E[h^2g^2]E[s^2(\theta)] - E^2[hg]E^2[s(\theta)] + \frac{\sigma_v^2}{L}}{E^2[hg]E^2\left[\frac{\partial s(\theta)}{\partial\theta}\right]} \tag{5-11}$$

5.1.2 Analytic Solution & Verification for Hyperbolic Tangent Amplifier

For verifying the formula specified above, simulations were run. For initial simulations used to validate the accuracy of the predicted $\mathrm{AsVar}(\hat{\theta})$ formulation, the channel and the sensor gains were both assumed to be uniformly distributed from 0 to 1 to mimic channels that may have obscured paths. A more realistic model may be to use a Rician distribution of gains with a random *K*-factor. The hyperbolic tangent model was used as the limiting amplifier model, but unlike in section 4.2 the maximum nominal amplitude of the limiting amplifier model, c, was left at 1. This simplification was done only because in this simulation, no attempt was made to match a specific circuit simulated amplifiers to develop estimates of power-added efficiency. The limiting amplifier nominal output was scaled by a gain factor, g_l, controllable in the amplifier biasing circuitry.

The steps outlined in section 5.1.1 were used to determine an estimator and its performance for the hyperbolic tangent limiting amplifier model. The amplifier model differs from the formulation in (4-6) in that the limit setting term, $c = \alpha k$, has been incorporated into the sensor gain value g_l leaving only the hyperbolic tangent of the scaled measurement as the amplifier model function $s(x)$ as shown in (5-12):

$$s(x) = \tanh(x/k). \tag{5-12}$$

Substituting $s(x)$ into (5-7) yields an expectation of the value at the fusion center's estimator, $\bar{\zeta}(\theta)$, as shown in (5-13):

$$\bar{\zeta}(\theta) = E[hg]\frac{k}{2b}\log\left(\cosh\left(\frac{b+\theta}{k}\right)\operatorname{sech}\left(\frac{b-\theta}{k}\right)\right). \quad (5\text{-}13)$$

To develop the estimator, the inverse equation to (5-13) is found in a process similar to that in section 4.2. The estimate $\hat{\theta}$ can be found by replacing the expected receiver value $\bar{\zeta}(\theta)$ with the actual received value z and replacing the actual θ with the estimate $\hat{\theta}$, then solving for $\hat{\theta} \in \mathbb{R}^+$. For the hyperbolic tangent limiting amplifier model, defining constant $m = \exp\left(\frac{2bz}{kE[hg]}\right)$, and for $z \in [0, E[hg]]$:

$$\hat{\theta} = k\cosh^{-1}\left(\frac{\sinh\left(\frac{b}{k}\right)(1+m)}{\sqrt{2m\cosh\left(\frac{2b}{k}\right) - 1 - m^2}}\right). \quad (5\text{-}14)$$

To analyze the performance of the estimator, the variance of the value received at the fusion center, $\mathrm{AsVar}(\hat{\theta})$, is determined using (5-10). The variance divided by the number of sensors, in this example also known as the asymptotic variance, can be determined according to:

$$\Sigma(\theta) = E[h^2g^2]\frac{2b - k\tanh\left(\frac{b-\theta}{k}\right) - k\tanh\left(\frac{b+\theta}{k}\right)}{2b} - \bar{\zeta}^2(\theta) + \frac{\sigma_v^2}{L}. \quad (5\text{-}15)$$

The derivative of the value expected is the other component required to determine estimator performance.

$$\frac{\partial\bar{\zeta}(\theta)}{\partial\theta} = E[hg]\frac{\tanh\left(\frac{b-\theta}{k}\right) + \tanh\left(\frac{b+\theta}{k}\right)}{2b} \quad (5\text{-}16)$$

To validate that values calculated in (5-13) through (5-16) are correct, simulations were run in MATLAB to find experimentally determined values for $\Sigma(\theta)$ and $\mathrm{AsVar}(\hat{\theta})$

for random channels of known statistics. For these simulations both the sensor gains and the channel gains were randomized to verify the robustness of the equations. Both sensor gains and channel gains were chosen to be uniformly distributed between 0 and 1, thus $g \sim U[0,1]$ and $h \sim U[0,1]$. These values were compared to those attained from simulations with identical constant channel gains, with $g = 0.5$ and $h = 0.5$. For both simulations, channel noise was set to $\nu = 0$ as the influence of channel noise is the same between the two topologies. As shown in Figure 5-2, theoretical results match the experimental results for both the identical constant channels case and the channel-gain statistics cases.

It can be seen from Figure 5-2, that the mean and derivative against θ of the value received at the fusion center's estimator is identical between the two cases. The key observation from Figure 5-2 is therefore that variance observed at the estimator is much higher when only channel-gain statistics are known than it is when the channels are all identical and known. Within the only channel-gain statistics cases, it can also be seen that variance increases with increasing θ, whereas for identical channels variance decreases with increasing θ. This increasing received value variance is especially problematic as the derivative of the received value is also decreasing, causing estimate variance to expand as specified by (5-10). This rapidly increasing estimate variance with θ is confirmed via the simulated and theoretical results shown in Figure 5-3.

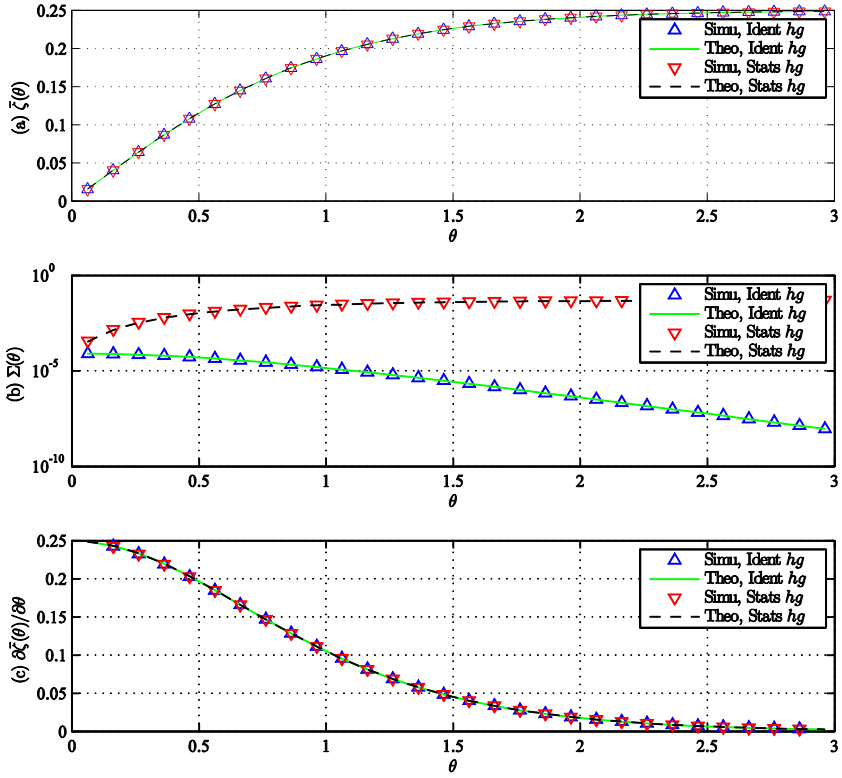


Figure 5-2: Verification of phase-only CSI gain-statistic equations, and comparison of AF systems with gain-statistics available versus systems with identical channels.

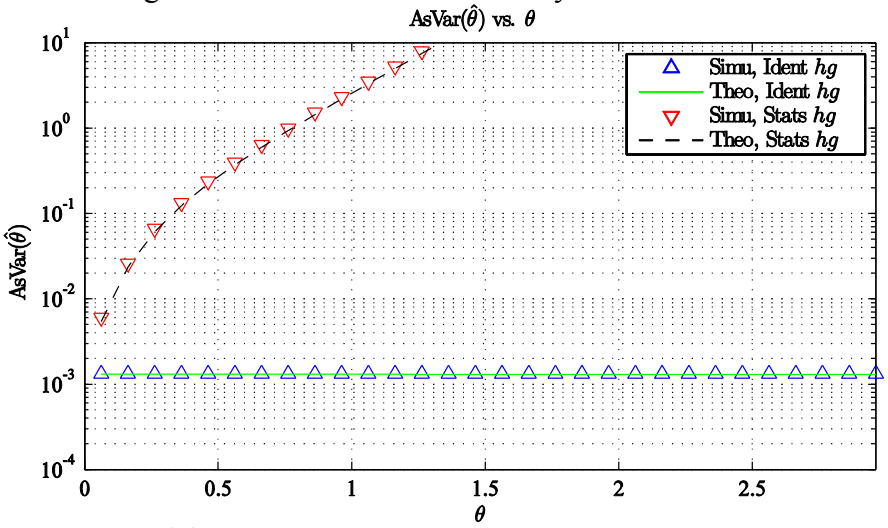


Figure 5-3: AsVar$\left(\hat{\theta}\right)$ for Identical Constant and Statistical Gain Channels versus θ.

5.1.3 Issues with Nonlinear AF Using Phase-Only CSI

Figure 5-4 shows histograms of the actual values received, z, for a finite number of sensors, L. For these simulations, the hyperbolic tangent limiting amplifier model has been used with $k = 1$, $\theta = 1.25$, $b = 3/32$, and $\nu = 0$. The channel gains, h, were chosen to be uniformly distributed between 0 and 1, $h_l \sim U[0,1]$, and the sensor gains were set to 1, $g_l = 1$, thus $E[hg] = 0.5$. The primary problem in estimating z with only channel gain-statistics is that variation in z becomes much greater than for identical constant channels as the transmitter becomes more nonlinear. This is due to the exaggerated effect on received value of channel gain changes when the amplifier is operating in the compressed region. This point is illustrated in by histograms values received by estimator for both (a) identical constant gain channels and for (b) the random phase-only CSI channels are shown in Figure 5-4. In these two plots, the variance for constant identical channels at $L = 20$ sensors is still lower than the variance of the gain-statistic channels at $L = 2000$ sensors. This corresponds to a 100x increase in sensor power without improvement in estimate variance compared to the known statistics case.

A second problem illustrated by Figure 5-4 is that it is likely that even with a practically large L, such as 20, the value presented to the estimator $z \notin [0, E[hg]]$. When $z \notin [0, E[hg]]$, the estimator in (5-14) provides a complex-valued estimate and thus fails. Therefore, it is important to make sure the number of channels is large enough to make the probability $\Pr[z > E[hg]] \to 0$ when implementing a nonlinear distributed

estimation system, as also demonstrated in [69]. This requirement is easily violated when using nonlinear distributed estimation with phase-only channel-state information or when channel noise is high

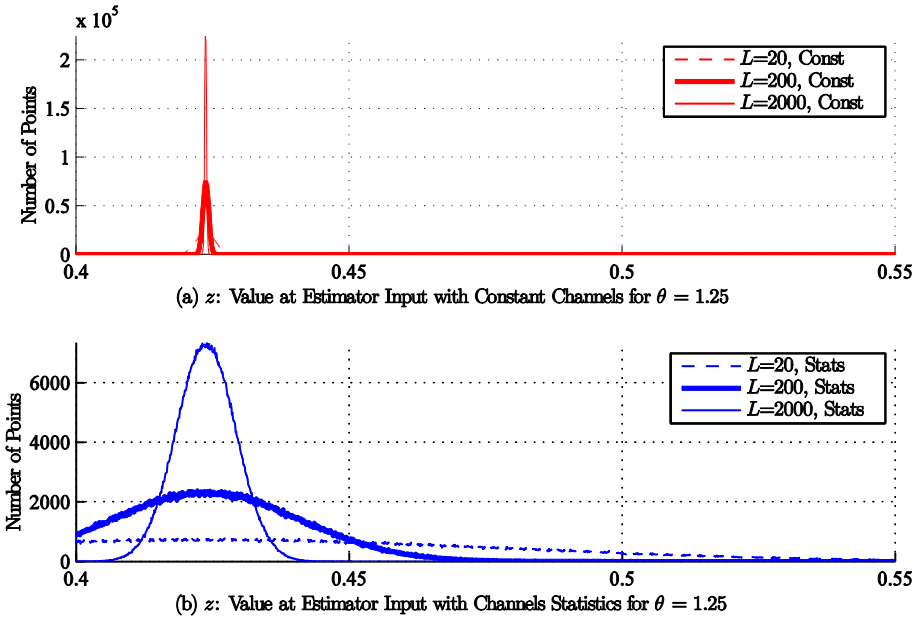


Figure 5-4: Histogram of values received at fusion center's estimator input.

The increase in number of sensors that is required to attain similar variance at the fusion center's estimator, $\mathrm{Var}(z)$, when the channel gains are only statistically known is shown in Figure 5-5. In this plot, the limiting amplifier model used for transmitting data is the hyperbolic tangent function. The channels in this plot have uniformly random gains between 0 and 1. At $\theta = 0.25$, the hyperbolic tangent function is nearly linear. When transmitters are operating nearly linear, it can be seen that similar performance can be attained to constant channels when the number of sensors is increased by a factor of 10; however, when the transmitters are operating nonlinearly, it can be seen that a large

increase in the number of sensors is required. In the case of $\theta = 1.00$, it can be seen that 333 times more sensors are required to achieve the same variance. Because the power consumed in transmitters scales with the number of sensors, this system of nonlinear transmitters operating through statistically-known phase-synchronized channels cannot be regarded as very efficient.

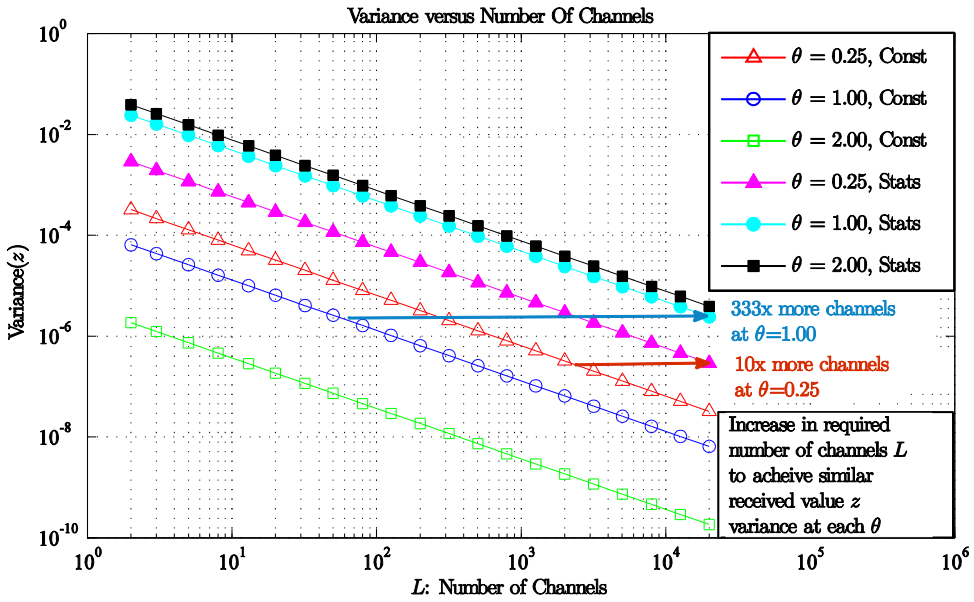


Figure 5-5: Variance vs. number of channels for hyperbolic-tangent shaped transmitters operating through constant channels and random statistical channels.

5.2 Full Channel-State Information

In this section, the performance of nonlinear distributed estimation is investigated when the channel state information is fully known at the transmitters. First, analytical expressions are derived for the asymptotic variance of the estimate when the channel state information, h_l, is known for all channels. Because h_l is known, it is also assumed

that g_l is also known, and will be chosen strategically during system implementation to achieve a desired performance in terms of estimate variance.

5.2.1 Derivation of Estimate Variance

Because the h_l and g_l are now known values, the analysis of section 5.1 must be reconsidered. Starting from the same system model in (5-1), the expectation can be recalculated assuming non-random sensor and channel gain values:

$$E[y(\theta)] = E[s(\theta)] \sum_{l=1}^{L} h_l g_l + \mu_v \,. \tag{5-17}$$

A slightly different definition of z is utilized here than in previous sections:

$$z = \frac{y}{\sum_{l=1}^{N} h_l g_l} \,. \tag{5-18}$$

With this definition of z, the expected value when channel noise, v, has zero mean is:

$$\bar{\zeta}(\theta) = E[z(\theta)] = E[s(\theta)]. \tag{5-19}$$

The estimate can therefore be computed using the inverse of the expected value of the non-linear amplifier model, *i.e.*,

$$\hat{\theta} = \bar{\zeta}^{-1}(z). \tag{5-20}$$

The variance of the received value, z, with preconditioning on known h_l and g_l can be found to be:

$$\mathrm{Var}(z) = \frac{\sum_{l=1}^{L} h_l^2 g_l^2}{(\sum_{l=1}^{L} h_l g_l)^2} (E[s^2(\theta)] - E^2[s(\theta)]) + \frac{\sigma_v^2}{(\sum_{l=1}^{L} h_l g_l)^2} \,. \tag{5-21}$$

Using (5-6), the estimator variance can be found to be:

$$\mathrm{Var}\left(\hat{\theta}\right) = \frac{(E[s^2(\theta)] - E^2[s(\theta)])(\sum_{l=1}^{L} h_l^2 g_l^2) + \sigma_v^2}{E^2[\partial s(\theta)/\partial\theta\,](\sum_{l=1}^{L} h_l g_l)^2}. \tag{5-22}$$

For the specific case of a hyperbolic tangent limiting amplifier model, as shown in (5-12), with sensing noise distributed uniformly and iid from $-b$ to b, $\eta_l \sim U[-b, b]$, and fully known channel and sensor gain information $h_l g_l$, the definitions of $E[s^2(\theta)]$, $E[s(\theta)]$, and $\mathrm{E}[\partial s(\theta)/\partial\theta]$ representing the random components of sensing can be found to be:

$$\bar{\zeta}(\theta) = E[s(\theta)] = \frac{k}{2b}\log\left(\cosh\left(\frac{b+\theta}{k}\right)\mathrm{sech}\left(\frac{b-\theta}{k}\right)\right), \tag{5-23}$$

$$E[s^2(\theta)] = 1 - \frac{k}{2b}\tanh\left(\frac{b-\theta}{k}\right) - \frac{k}{2b}\tanh\left(\frac{b+\theta}{k}\right), \tag{5-24}$$

$$\frac{\partial\bar{\zeta}(\theta)}{\partial\theta} = E\left[\frac{\partial s(\theta)}{\partial\theta}\right] = \frac{1}{2b}\tanh\left(\frac{b-\theta}{k}\right) + \frac{1}{2b}\tanh\left(\frac{b+\theta}{k}\right) \tag{5-25}$$

The estimator can be found very similarly to (4-13) and (5-14) by solving for the inverse of the expectation function. Where $m = \exp(2bz/k)$ and $m \in [0,1]$, then:

$$\hat{\theta} = k\cosh^{-1}\left(\frac{\sinh\left(\frac{b}{k}\right)(1+m)}{\sqrt{2m\cosh\left(\frac{2b}{k}\right) - 1 - m^2}}\right). \tag{5-26}$$

To validate the variance of the received value and estimate, simulations were performed and the results were compared to the analytically-determined theoretical values as shown in Figure 5-7. To perform the side-by-side comparison more clearly, the measurements of the full exact channel state information case was scaled by $E[hg] = 0.25$ for the

expected values received and by the variance $E^2[hg] = 0.25^2$. This was done as the uniform channels technique did not incorporate the channel gains into the measurement scaling.

Figure 5-7 shows that the performance in the presence of full CSI, the received value variance is much closer to the performance of uniform identical channels than are the phase-only CSI simulations from Figure 5-2. In agreement with the decrease in received value variance, the variance of the estimate is also much closer to the performance of the constant identical channels than the phase-only CSI. The estimate variance can be seen in Figure 5-6.

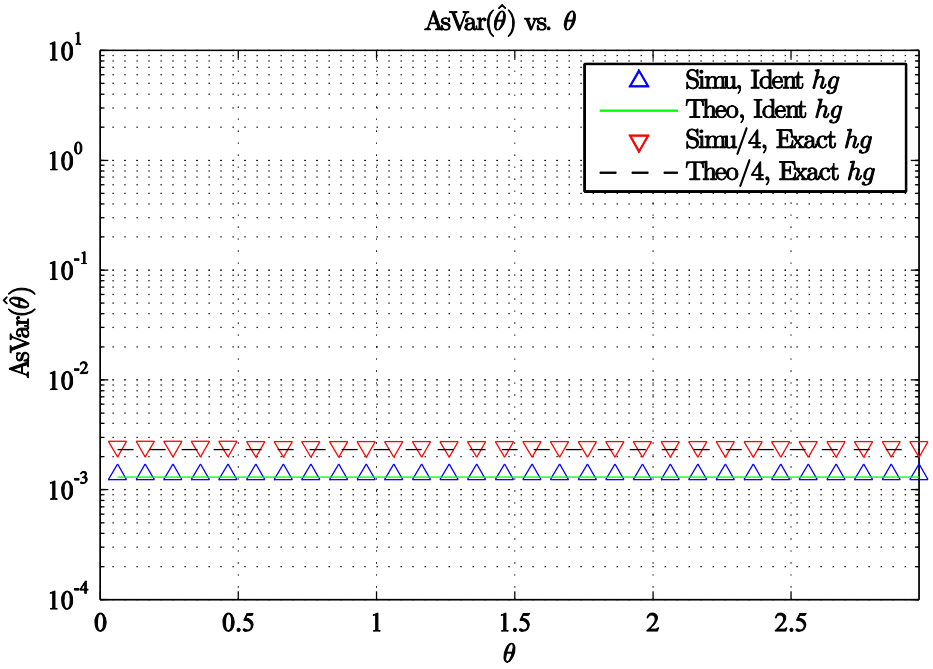


Figure 5-6: AsVar$\left(\hat{\theta}\right)$ for Identical Constant and Full CSI Channels versus θ.

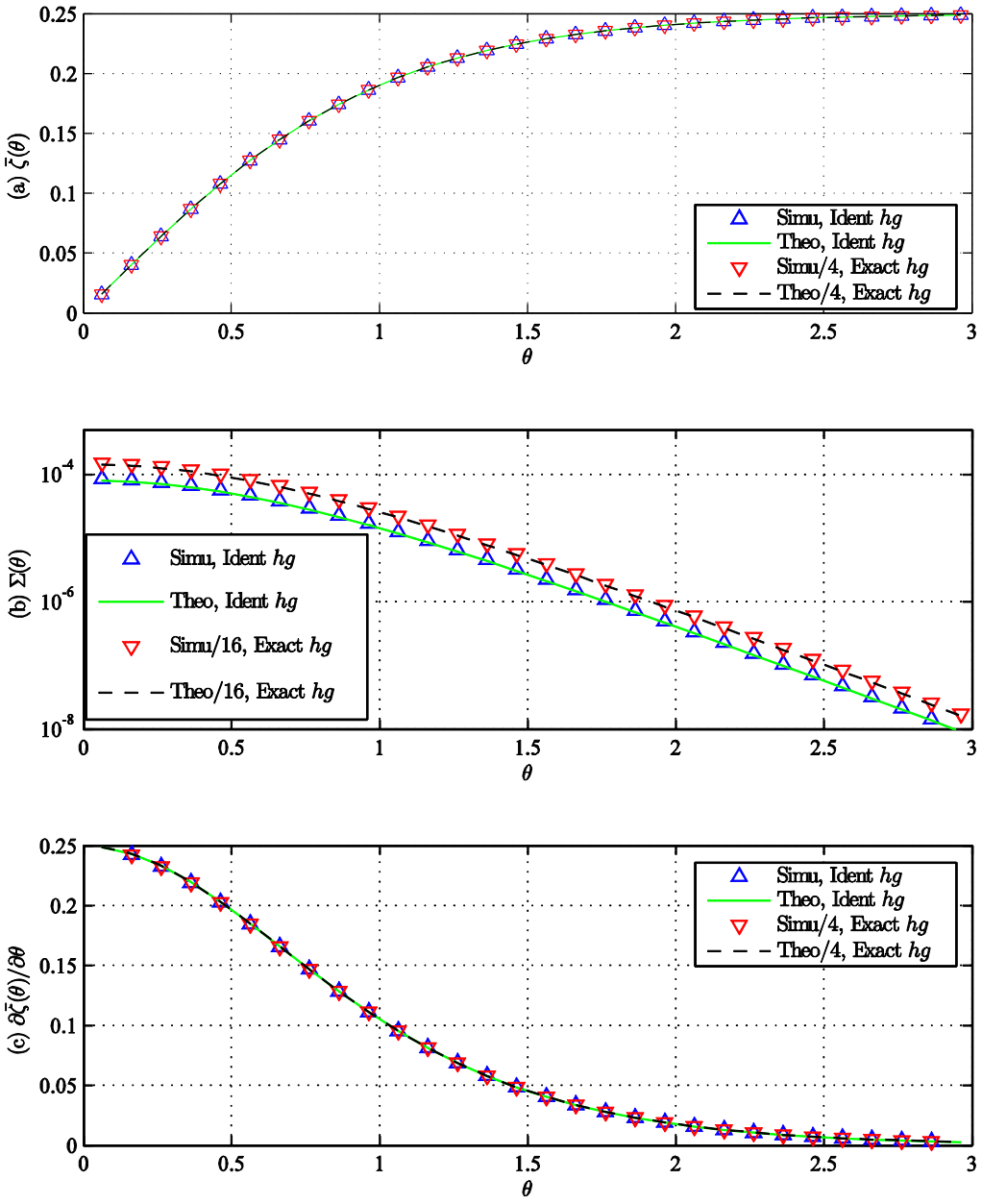


Figure 5-7: Verification of full CSI gain-statistic equations, and comparison of AF systems with full CSI available versus systems with identical channels.

5.2.2 Optimal Power Allocation Strategies

With an analytically verified solution for the estimate variance for AF distributed estimation with full channel-state information, a minimization problem can be formed to intelligently allocate power between the individual transmitting sensors in order to achieve the best performance while maintaining a specific performance level. The problem can be posed as an optimization problem of the sensor gains. Given an expected probability distribution function of the underlying physical values to be measured $p(\theta)$, a function $P_{amp}(g_l, x_l)$ which is the power consumed by a sensor measuring quantity $x_l = \theta + \eta_l$ with sensing transmitter gain of g_l, and a total consumed power budget of P_{budget} the best average estimate variance possible is:

$$\begin{aligned} &\underset{\{g_1,\ldots,g_L\}}{\text{minimize}} \int_\theta \frac{(E[s^2(\theta)] - E^2[s(\theta)])(\sum_{l=1}^{L} h_l^2 g_l^2) + \sigma_v^2}{E^2\left[\frac{\partial s(\theta)}{\partial \theta}\right](\sum_{l=1}^{L} h_l g_l)^2} p(\theta) d\theta \\ &\text{subject to: } \int_\theta \left(\sum_{l=1}^{L} \int_{\eta_l} P_{amp}(g_l, \theta + \eta_l) d\eta_l \right) p(\theta) d\theta = P_{budget}. \end{aligned} \tag{5-27}$$

This optimization is not easily evaluated using convex optimization techniques, so as part of making the optimization easier, a very similar dual problem was posed similar to the techniques in [8]. The dual problem minimizes the power consumed subject to a specified average variance level, Υ_t:

$$\underset{\{g_1,\ldots,g_L\}}{\text{minimize}} \int_{\theta} \left(\sum_{l=1}^{L} \int_{\eta_l} P_{amp}(g_l, \theta + \eta_l) d\eta_l \right) p_\theta(\theta) d\theta$$
$$\text{subject to: } \int_{\theta} \frac{(E[s^2(\theta)] - E^2[s(\theta)])(\sum_{l=1}^{L} h_l^2 g_l^2) + \sigma_v^2}{E^2\left[\frac{\partial s(\theta)}{\partial \theta}\right] (\sum_{l=1}^{L} h_l g_l)^2} p_\theta(\theta) d\theta = \Upsilon_t. \tag{5-28}$$

This optimization problem is itself still computationally intensive to solve using convex optimization techniques, so the optimization problem was again reposed with a slightly different target. In this revision of the dual problem, it is decided to optimize against the worst case variance produced by the desired estimator. It is observed consistently throughout chapters 4 and 5 that the worst case estimator performance, both with and without considering channel noise, ν, occurs when the limiting amplifier model is most compressed. This compression occurs at the largest value of θ. Given all noise measurements $x_l = \theta + \eta_l$ are relatively similar, it is assumed that the amplifier's efficiency will be similar for all sensors, and thus, the function $P_{amp}(g_l, x_l)$ can be reduced to a constant factor, ϕ. Similarly, if only the worst case of θ is considered, the probability density function, $p(\theta) = \delta(\theta_{wc})$, where θ_{wc} is the maximum value in the expected range of θ to be measured, and Υ_{wc} is the worst case estimate variance. This reduced problem can be considered as:

$$\underset{\{g_1,\ldots,g_L\}}{\text{minimize}} \sum_{l=1}^{L} \phi g_l^2$$
$$\text{s.t.: } \frac{(E[s^2(\theta_{wc})] - E^2[s(\theta_{wc})])(\sum_{l=1}^{L} h_l^2 g_l^2) + \sigma_v^2}{E^2\left[\frac{\partial s(\theta_{wc})}{\partial \theta}\right](\sum_{l=1}^{L} h_l g_l)^2} = \Upsilon_{wc} \tag{5-29}$$

By introducing a slack variable, ρ, and ignoring the constant factor ϕ which has no bearing on the minimization, the problem can be rearranged to a similar form as is solved in [8]. In this form the minimization does not directly find the total power consumed, but rather a scaled version of it due to the removal of the $P_{amp}(g_l, x_l)$.

$$\underset{\{g_1,\ldots,g_L,\rho\}}{\text{minimize}} \sum_{l=1}^{L} g_l^2$$
$$\text{subject to: } (E[s^2(\theta_{wc})] - E^2[s(\theta_{wc})])\left(\sum_{l=1}^{L} h_l^2 g_l^2\right) + \sigma_v^2 = \rho^2 \Upsilon_{wc}\,. \tag{5-30}$$
$$E\left[\frac{\partial s(\theta_{wc})}{\partial \theta}\right]\left(\sum_{l=1}^{L} h_l g_l\right) = \rho$$

The minimization in (5-30) has been solved in [8]. A few specific cases are investigated to demonstrate performance and sensor gain allocation in specific scenarios. In the first scenario, $L = 12$ sensors were used with $h_1 = h_2 = \cdots = h_9 = 5.7735 \times 10^{-4}$, and $h_{10} = h_{11} = h_{12} = 0.57735$, making $\sum_{l=1}^{12} h_l^2 = 1$. In this scenario, the split between how power was allocated between strong and weak channels was investigated. Solutions to the Karush-Kuhn-Tucker (KKT) conditions [70] were computed using Mathematica. The sensing noise, b, has been chosen to be uniformly distributed such

that it is bounded by $\pm\ ^1/_{32}$. The channel and receiver noise, v, has been chosen to have Gaussian distribution with zero-mean and variance $10^{-7}/L$.

We have compared two extreme cases with the solution of the optimization problem. In one extreme case, a *water-filling* like solution is utilized where power is allocated to the sensors in proportion to their channel strengths. A power allocation allocating power to the strongest channels is best at overcoming limitations when channel and receiver noise is the leading cause of received value variance. However, when sensor measurement noise is a limiting factor, the proportional allocation of power is poor at averaging multiple noisy sensor measurements to achieve consensus. Depending on the noise sources, one of these two low-complexity cases is very close to the performance of the optimized solution with much lower computation requirements. If the crossover point can be found, it may be possible to get near optimal performance from the two low-complexity approaches.

For the given distribution of channels gains, estimate variance with a uniform power allocation strategy should be similar to a case with 3 uniform, identical gain channels since 9 of the channels are very weak. Using that with negligible channel noise, $L_e\ \mathrm{Var}(\hat{\theta}) = b^2/3$, where L_e is the effective number of channels, the expected variance for the estimate $\mathrm{Var}(\hat{\theta}) = 1.08 \times 10^{-4}$ should be the best performance possible. In cases where the tolerable gain variance is much greater than $\mathrm{Var}(\hat{\theta}) = b^2/(3L_e)$, the total transmitter power can be decreased to where channel noise is the primary

contributor to measurement variance. This is confirmed by the simulations as regardless of the sum of transmitter powers this is the best performance possible, as confirmed in Figure 5-8.

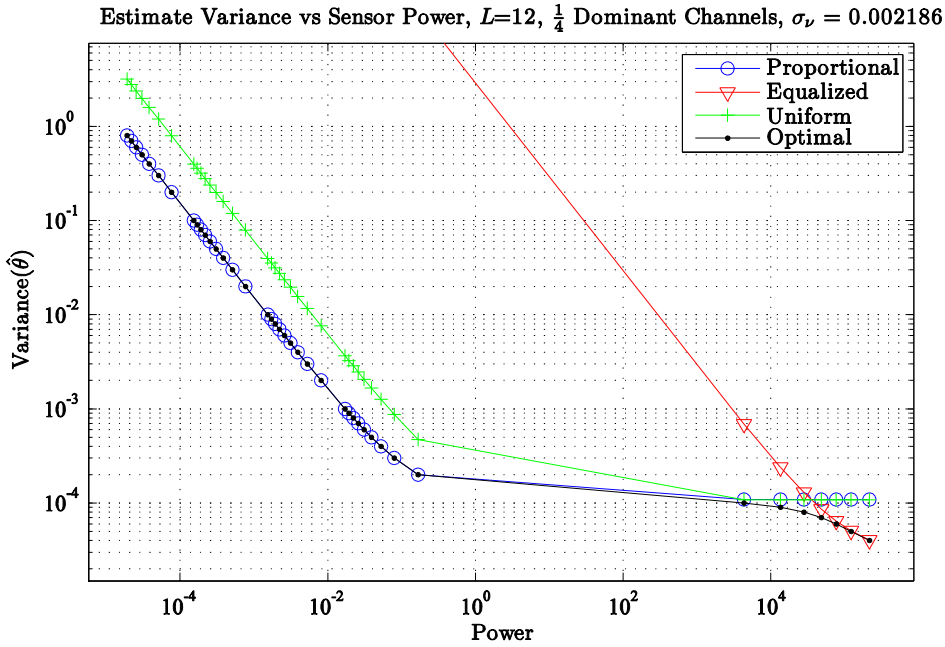


Figure 5-8: Estimate variance vs. scaled total sensor power allocation for various allocations of sensor powers with faded channels.

The allocation of power to strongest channels works well until such a point that the sensing measurement bound for using the strongest channels has been attained. To perform better than that bound requires moving towards a situation where all channels arrive at the receiver's FC with similar strength, and thus increase the effective number of channels, L_e, by taking the expectation over more instances of the sensing noise. In order to accomplish this in the presence of weak channels requires placing a considerably higher gain on weaker channels to make its arrival strength similar to those of the strong channels. In this opposite extreme case of power allocation, the sensor gains are set

inversely proportional to the channel gains, the objective being to make $h_l g_l$ equal across all sensors making $L_e = L$. This case achieves the best performance in the event that sensor measurement noise solely limits estimator performance rather than channel and receiver noise. It can be seen from Figure 5-8 that when weak channels exist, the equalizing technique uses a lot of power and has poor performance when channel and receiver noise limits estimator performance. However by averaging all measurements the equalized technique can achieve better estimation accuracy than can the proportional power allocation technique.

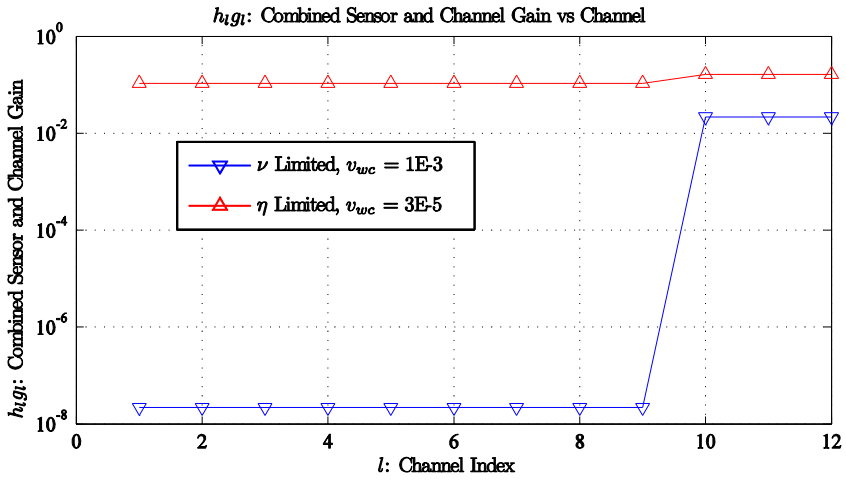


Figure 5-9: Combined gains for sensing-noise limited and channel-noise limited scenarios.

It can be seen from Figure 5-8 that the optimal solution for worst case variance tracks the proportional allocation of sensor gains where the power budget is low and thus performance is limited by the fixed channel and receiver noise, ν. It can also be seen that where power budget is high and the sensor measurement noise limits estimator performance, channel equalization as a power allocation strategy yields the best

performance. This can be confirmed when the combined channel gains are plotted for two different power budgets, as shown in Figure 5-9.

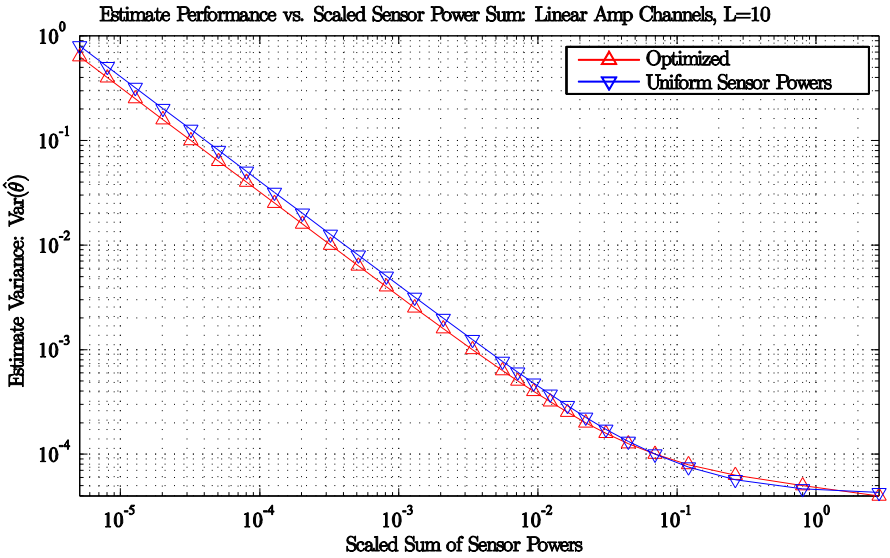


Figure 5-10: Estimate variance vs. scaled total sensor power allocation for uniformly distributed and optimized distribution of sensor powers with progressively increasing channel gains.

The return that can be expected by optimizing the sensor gain powers depends on the number of channels that are deeply faded. In a second simulated scenario, $L = 10$ channels were used with channel gains $h_l = l/L$ with $l = 1, \ldots, 10$. The measurement noise is uniformly distributed $b \sim U[-1/32, 1/32]$ and channel noise is normally distributed with variance $10^{-7}/L$, Simulations results are shown in Figure 5-10. Even though channel gains are unequal, the reduction in power consumed by the sensors in while maintaining the same estimate variance is typically 20% for this distribution of

channel gains. Whether or not it makes sense to optimize the allocation of sensor powers in the presence of full CSI will depend primarily on the expected distribution of channel gains, especially how many channels are known to be deeply faded and thus a wastful place to allocate sensor power.

5.3 OFDMA with a Continuous AM Carrier

5.3.1 Proposed Technique

A preliminary investigation has been conducted into the use of orthogonal frequency-division multiple access (OFDMA) with an amplify-and-forward distributed estimation scheme. This estimation system was designed not for power efficiency, but rather to operate in a heavily scattered environment where synchronizing the transmitters and developing channel statistics has proven difficult to achieve. The key difference from previous discussions is that the channels between the transmitter and the carrier contain multiple-paths with time lags, amplitudes, and phase rotations, denoted in Figure 5-11 by the fading channel. Orthogonal-frequency division multiplexing has been proven to be an effective modulation scheme in the presence of multipath channels, and thus has been utilized in these experiments.

In the proposed system, there are two types of data-transporting subcarriers per each sensor: a common subcarrier and an independent subcarrier. Each subcarrier is modulated with a "continuous-AM" signal, which is an amplitude-modulated signal with amplitude equal to a linearly-scaled copy of the noisy sensor measurement, x_l. In

practice, the resolution of the "continuous-AM" signals would be dependent upon the resolution of each sensor's output digital-to-analog converter. Measurement quantization noise can be modeled by incorporating it into the sensor measurement noise distribution. Each data-transporting subcarrier is enclosed by pilot subcarriers of both lower and higher frequencies for channel estimation. The channel estimation is done according to comb-type pilot arrangement in [71]. Because the OFDMA system is heavily utilized commercially, its selection for this distributed estimation system provides opportunities for system implementation utilizing inexpensive, commodity hardware.

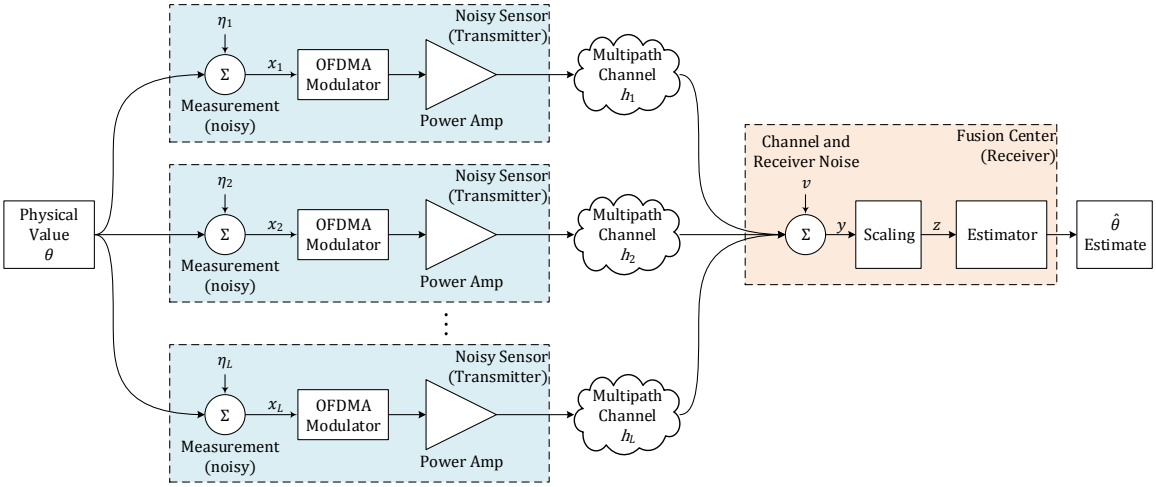


Figure 5-11: Distributed Estimation over Fading Channels.

The common subcarrier and its surrounding tones is a shared-access channel, allowing estimator asymptotic variance performance less than the variance of individual measurements provided a strong enough signal is received. The design intent of the common subcarrier is that with each sensor transmitting similar data, excepting measurement noise, on the same subcarrier, that each sensor's transmission looks like

multipath data received from a single equivalent transmitter. The independent subcarriers from each sensor allow both forming an estimate by consensus from the individual measurements, though estimate asymptotic variance is limited to the variance of individual sensors. The consensus estimate for orthogonal signals has been formed by taking the weighted average of the orthogonal sensor measurements where the weighting has been done based on the channel power estimates such that:

$$\hat{\theta} = \frac{\sum_{l=1}^{L} \hat{\theta}_l |H_{est}(l)|^2}{\sum_{l=1}^{L} |H_{est}(l)|^2}. \tag{5-31}$$

The effect of randomness in the channel gains due to multipath is determined by examining the inverse fast-Fourier transform (IFFT) used for demodulating the received signal. According to this formula, the received signal is $X(k)$ is:

$$X(k) = \frac{1}{N} \sum_{n=0}^{N-1} h(n)x(n)e^{-j2\pi kn/N}\ , \tag{5-32}$$

where $x(n)$ is randomly distributed due to the noisy sensor measurements and $h(n)$ is randomly distributed due to the channel noise. By the central limit theorem [24], for channel noise distributions with defined mean and variance, the noise in $X(k)$ is a sum of complex random numbers of random phase and is thus can thus be approximated by a complex Gaussian distribution.

A known performance limiter for OFDMA systems is frequency mismatch between individual nodes in the network. Where frequency mismatch or excessive phase noise occurs, the subcarriers are no longer completely orthogonal and introduce inter-

carrier interference (ICI) [72,73]. A formula by given for SNR degradation, D_{SNR} for due to ICI in terms of the original signal to noise ratio, E_s/N_0, the number of subcarriers N, the modulated bandwidth W, and the frequency mismatch ΔF [73]:

$$D_{SNR}(\text{dB}) \approx \frac{10}{3\ln 10}\left(\pi \frac{N\,\Delta F}{W}\right)^2 \left(\frac{E_s}{N_0}\right) \tag{5-33}$$

In the event of time-synchronization errors, the effects of this mismatch are exaggerated farther away from the carrier frequency [74], thus for this system the common subcarrier and its pilots were transmitted on subcarriers indexed between 1 and 3. Six subcarriers were also reserved for peak-power reduction by method of tone reservation at the subcarriers farthest away from the carrier for similar reason. Tone reservation is outlined in [75] and [76].

The use of independent pilot tones from each sensor opens additional opportunities including the use of multiple fusion centers to provide triangulation data for determining the location and the exact measurement from each sensor [68]. Opportunities for future optimization to help obtain estimation efficiency typically lost when using an OFDM transmission system include both PAPR reduction systems and advanced amplifier predistortion systems.

5.3.2 Simulation Results

Simulations were performed to analyze the performance of the system. In this system, each transmitted signal passed through several independent paths to arrive at the

fusion center. The amplitude of each path were independently distributed. The lag of each path were also randomly distributed.

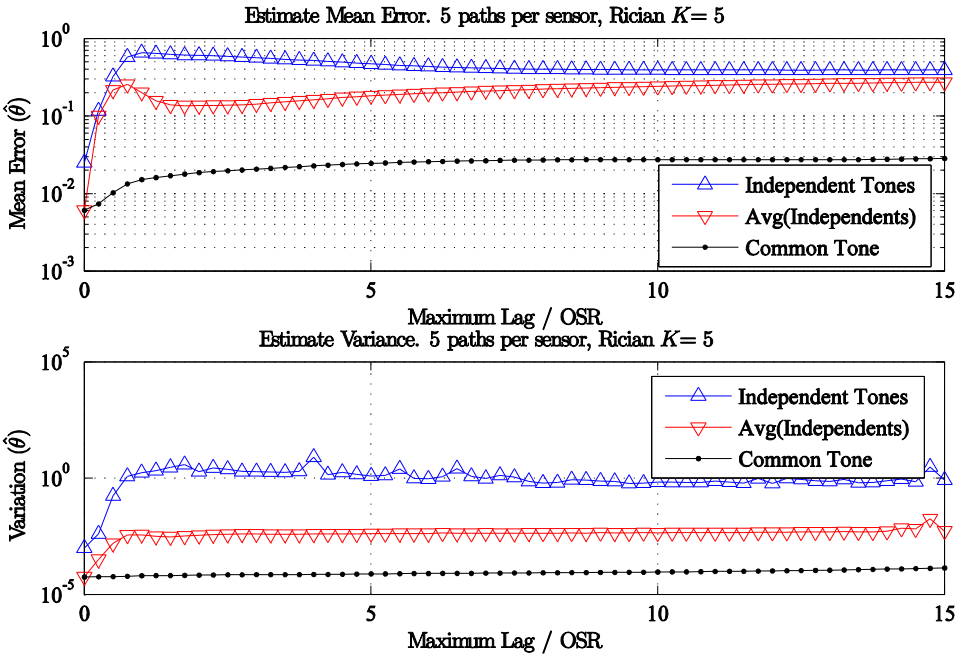


Figure 5-12: OFDMA channel estimation performance vs estimation method for various multipath channel lags, Rician K=5 channel gains.

For the first set of simulations, the maximum time difference in OFDMA multipath delay times were investigated. The simulation utilized 18 sensors. Each sensor had 5 paths to the fusion center. For the initial simulations, the individual path gains were all set to have a Rician distribution with K = 5. The maximum path delay time, t_{max} was swept between 0 to 15 sampling rate symbol times or 60 oversampled symbols. For each path, the delay for sensor l path p was distributed according to a uniform distribution with $t_{l,p} \sim U[0, t_{max}]$. The cyclic prefix length was 10 symbols prior to oversampling, 40 after oversampling. It can be seen from Figure 5-12 that anticipated

performance is better when the variation in the multipath delays were minimized. More importantly, when performance was being estimated in the presence of moderate delays on the order of 5 symbols prior to oversampling, the multiple access subcarrier yielded better performance with smaller mean error and variance than could be attained by taking the weighted mean of orthogonal measurements.

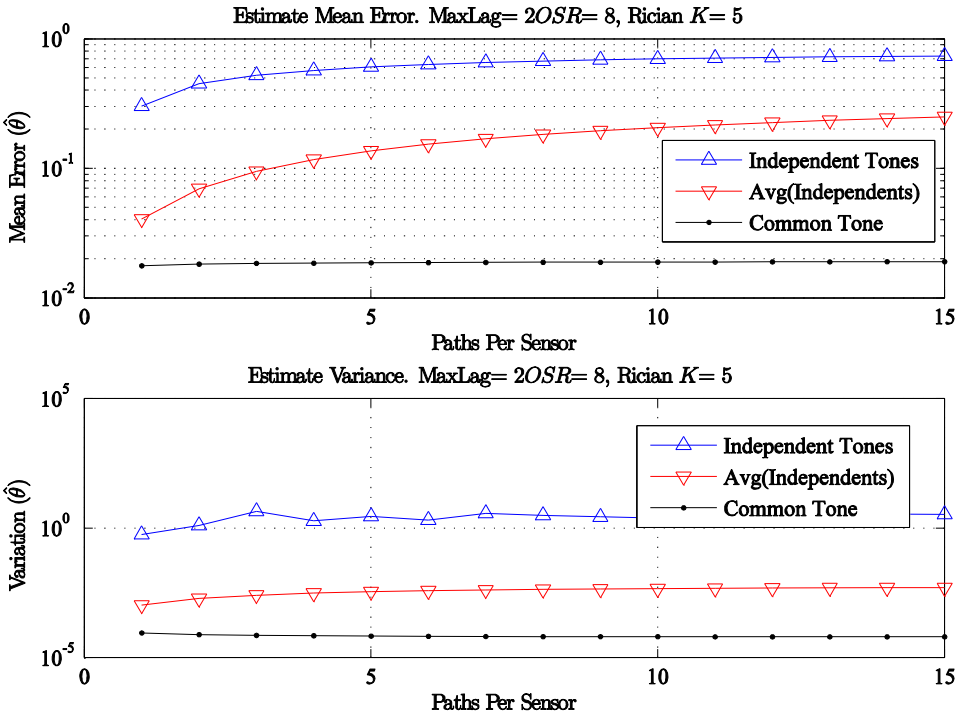


Figure 5-13: OFDMA channel estimation performance vs estimation method for various path counts, Rician K=5 channel gains.

The second simulation performed was to estimate the performance of the sensor network when more multipath components present in signal. In this simulation, the individual path gains were all set to have a Rician distribution with K = 5. The maximum path delay time was 2 symbols, or 8 oversampled symbols. In this simulation, the

number of paths was swept from 1 to 15. It can be seen from Figure 5-13 via this technique that the performance using the multiple-access channel exceeds that attainable from the weighted mean of orthogonal sensor estimates. While the variance is again slightly lower for the weighted average, the mean error was higher and thus the coherent MAC signaling was better when there was a significant line-of-sight component such as when multipath channel gains are distributed with a Rician distribution with K=5.

The third simulation compared the performance of the proposed techniques in the presence or absence of a strong line-of-sight channel. This was accomplished by sweeping the Rician K-factor from 0.01, corresponding to a near Rayleigh distribution, up to 100, corresponding to a near line-of-sight channel gain distribution. This was accomplished by setting the gains of each individual paths to have a complex Gaussian distribution where P_o is the power per sensor path, and Rician K-factor:

$$h_{l,p} \sim \mathcal{CN}\left(\sqrt{\frac{KP_o}{(1+K)}}, \frac{P_o}{2(1+K)}\right) \tag{5-34}$$

The simulation results, shown in Figure 5-14, demonstrate that when the Rician K-factor is small, that better performance is attainable via the weighted mean of the independent subcarrier results. However, when a strong line-of-sight component exists, the performance attained is better using the shared multiple-access channel with a single subcarrier.

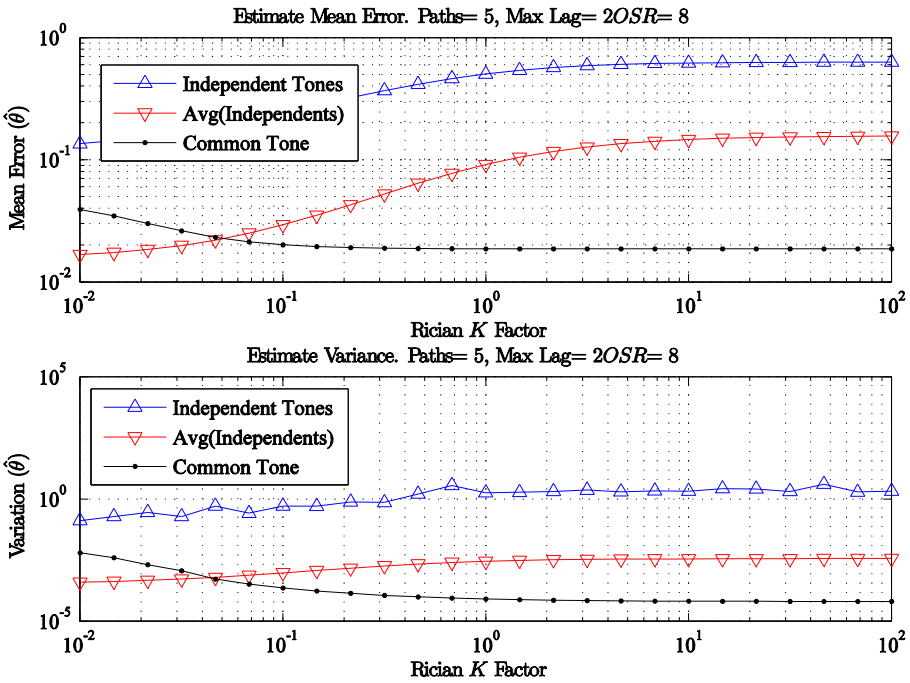


Figure 5-14: OFDMA Channel Estimation Performance vs Estimation Method for Various Path Gain Distributions.

Some odd behavior is observed in the current simulation environment with regard to what subcarrier selected as the common subcarrier. This could indicate weakness in the timing synchronization. This symptoms of this problem include a ringing response to different maximum path lag times. The frequency of the ringing is roughly proportional to the subcarrier index of the common subcarrier. These effects are shown in Figure 5-15.

Opportunities for future work exist in expanding the capabilities of this OFDMA technique. One opportunity is developing techniques for estimating the K-factor to determine when to use the weighted mean of independent subcarriers and when to use the common multiple-access subcarrier for composing the estimate. Existing techniques for

estimating Rician K-factors have been presented in [77]. Another opportunity is to localize the sensors using the pilot tones and techniques similar to those in [68]. Once localized, the sensors could be ganged into several different groups sharing a common subcarrier where such allocations would help either reduce estimate variance or bandwidth requirements.

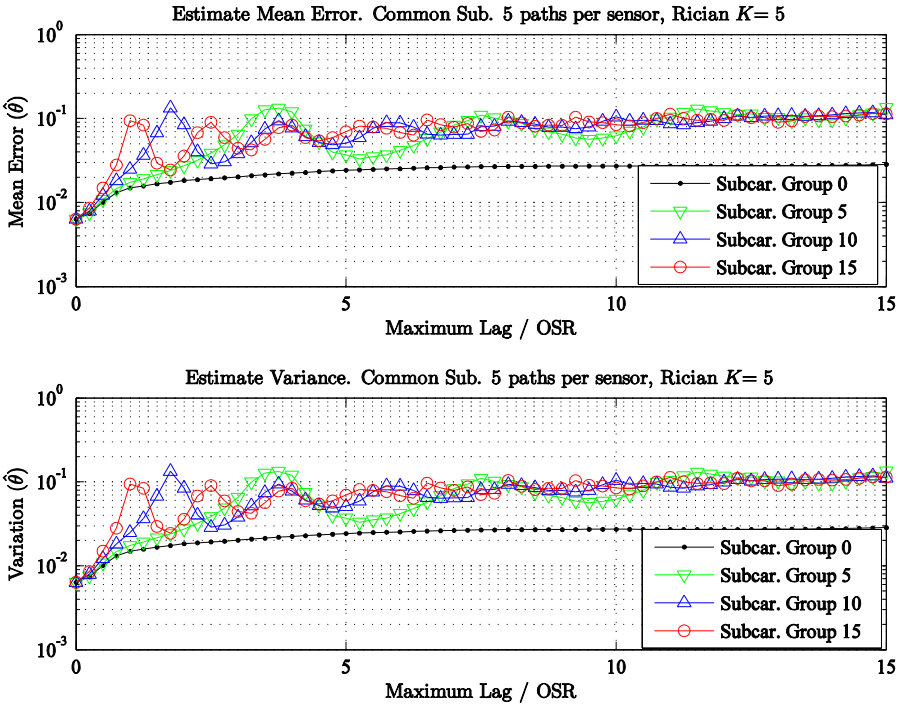


Figure 5-15: Difference in estimation accuracy against maximum lag by subcarrier.

5.4 Summary

In this chapter, it has been demonstrated that it is possible to perform distributed estimation over a coherent multiple access channel when full channel state information is known. With only channel-phase state information, the use of nonlinear transmitters

drives up the number of required sensors required to attain accurate estimates to unrealistically high levels. It has been demonstrated that when full channel state information is available, good performance is attainable with nonlinear transmitters. It was found via simulation that in most cases where allowable estimate variance is greater than the measurement sensing noise and thus limited by the channel and receiver noise, it is best to allocate power to the strongest channels. Where the estimate variance desired is less than that attainable via the independent combination of the strongest sensors, better performance is possible by equalizing the channels to average out more sensor measurements, but this gain comes at a substantial cost in power.

A new method has been presented for implementation of distributed estimation possible for implementation using OFDMA hardware. The estimation technique opens opportunities for sensor localization. Once localized, opportunities exist for splitting the sensors up into separate groups of subcarriers to perform coherent multiple-access distributed estimation of groups of sensors measuring quantities that may differ between the two separate locations.

6 CONCLUSIONS

6.1 Summary

This research has proposed a new algorithm for amplify-and-forward distributed estimation utilizing non-linear amplifiers that extends the battery life of sensors by boosting transmitter efficiency [63]. Analytical backing has been provided for the proposed algorithm over equalized channels, known channels of unequal gain, and for channels with known phase information and gain statistics. Modeling and simulation was used to validate the algorithms using simulated data from a real amplifier for both perfect and more practical methods of digital predistortion. The proposed approach demonstrated significant improvements over linear amplify-and-forward distributed estimation systems in overall efficiency for situations where channel gains were known and levels of channel and receiver noise were low. For sensing scenarios where channel gains were fully known, optimizing to minimize the estimate variance of a sensor measurement, θ, for a given transmitted power budget at the most compressed region of amplifier operation was shown to be an equivalent problem to the case of linear amplify-and-forward operation. Thus, the proposed technique to boost amplifier efficiency is compatible with existing systems in the literature to minimize transmitted signal power.

While the proposed approach is better when full channel-state information is available, its performance suffers relative to linear amplify-and-forward when only aggregate channel gain statistics are available. In these scenarios, an individual sensor

with a low-probability abnormally large channel gain can overwhelm the results of many individual sensors with nominal channel gain measuring a value near their compressed maximum. Simulations have demonstrated the increase in power required maintain the estimation accuracy in the case of channel gain-statistics only increases as the sensors operate more deeply into compression.

Transmitting power amplifiers do not have consistent gain and output power capability across process, voltage, temperature, or manufacturing variation. As such, a method is required with the proposed approach to force amplifiers to have a similar characteristic. In this approach, rather than attempting to linearize the amplifier, predistortion was utilized to match the amplifier to a similarly shaped non-linear limiting amplifier model. The compression present in the non-linear amplifier model reduces the dynamic range of the incoming signal and allows maintaining the amplifier in more efficient class AB operation more often than linearized amplifiers. Selection of a proper amplifier model is required for the proposed estimation technique to produce consistent estimates. The amplifier model should provide a gradual compression based on the sensing measurement, channel, and receiver noise distributions present in the network. Ideally, the model should be close to the inherent behavior of the amplifier to minimize the effort and power required to implement the predistortion system.

The techniques proposed for implementing a distributed estimation system over coherent multiple access channels can work well when a strong line-of-sight path is present. But there may be scenarios where the validity of such an assumption may be

questionable. Preliminary work has been done for a proposed distributed estimation technique to be used channel state information is unknown and a multipath environment is present. The proposed approach uses the orthogonal frequency-division multiple access scheme to transmit both orthogonal and coherent measurements, with included pilot tones to allow for channel estimation simultaneously with the reception of incoming sensor measurements. For near line-of-sight environments, the shared subcarrier coherent multiple-access approach yields the best performance. When the channel gain is Rayleigh, the independent measurements yield best performance. Thus, it is possible to make a decision in the receiver about the approach that yields the best performance if the channel environment Rician K-factor is known.

When the OFDMA approach is used, it will likely be required to utilize a high-quality predistortion system due to the high peak-to-average ratios inherent to orthogonal-frequency division technique, even though the number of carriers allocated per sensor is relatively low. In this research, we have proposed techniques for improving the performance and computational complexity of digital predistortion when memory-effects are present relative to existing techniques from the literature [59].

A block LMS (BLMS) technique for calculating predistortion coefficients in a transceiver with long loop-latency was presented [58,59]. This technique allows for faster convergence by allowing more points to be incorporated into each update of the LMS gradient calculation, thus discarding a lower fraction of measurements due to an update already being in progress. Initial simulations show good promise in terms of both

ACPR and EVM reductions. Further areas for exploration are made possible by using the block structure. These areas include optimizing μ by using non-constant values and re-use of certain data measurements. Implementations using weighted frequency-domain BLMS algorithms are an interesting area to pursue to allow explicitly biasing performance to reduce either ACPR or EVM. These approaches are complicated due to the non-standard sampling of the time-domain BLMS input matrix.

A modified neural network based predistorter was presented [59] that enabled reduction of the number of neurons, and thus weights, used in the predistorter neural network by incorporating as inputs the terms of a memory polynomial that is a rough match for the required predistortion function.

6.2 Future Work

The proposed OFDMA based distributed estimation opens numerous opportunities for future research projects to be completed within the group. OFDMA is a modulation standard used in many of the fourth-generation cellular communication standards. Thus it may be possible to implement the proposed system utilizing commercial off-the-shelf hardware at commodity design prices, and utilize it in real-world trials. Second, because pilot tones are used for the channel estimation, it may be possible to localize individual sensors. In doing so, the sensors could perhaps be allocated dynamically to different pools to allow measuring different systems. Opportunities also exist for research into peak-to-average power ratio reduction via tone reservation.

Room for future innovation exists in tuning both the high and low amplitude sides of the amplifier region being utilized. Right now the modified amplify-and-forward technique compresses the dynamic range of the transmitted signal at the high-power ranges. An intriguing problem to try and develop a further modified amplify-and-forward scheme that introduces both an offset, along with measurement and sensor gain scaling coefficients. In this way, the amplifier could be designed to allow increased estimate variance by both compressing the signal dynamic range and eliminating to as much as the variance requirement allows operation within the linear region of operation. Other opportunities could include scheduling different sensor gains during different times to extend the battery life of the whole system.

REFERENCES

[1] J. J. Xiao, S. Cui, Z. Q. Luo, and A. J. Goldsmith, "Linear Coherent Decentralized Estimation," *IEEE Transactions on Signal Processing*, vol. 56, no. 2, pp. 757-770, Feb 2008.

[2] A. Abedi, "Wireless Sensors Without Batteries," *High Frequency Electronics*, vol. 11, no. 11, pp. 22-26, November 2012.

[3] C. Rago, P. Willett, and Y. Bar-Shalom, "Censoring sensors: a low-communication-rate scheme for distributed detection," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 32, no. 2, pp. 554-568, April 1996.

[4] A. Ribiero and G. B. Giannakis, "Bandwidth-constrained distributed estimation for wireless sensor Networks-part I: Gaussian case," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 1131-1143, March 2006.

[5] J. J, Xiao, A. Ribeiro, Z. Q. Luo, and G. B. Giannakis, "Distributed compression-estimation using wireless sensor networks," *IEEE Signal Processing Magazine*, vol. 23, no. 4, pp. 27-41, July 2006.

[6] V. Aravinthan, S. Jayaweera, and K. Al Tarazi, "Distributed Estimation in a Power Constrained Sensor Netowrk," in *IEEE Vehiular Technology Conference*, vol. 3, Melbourne, 2006, pp. 1048-1052.

[7] M Gastpar and M Vetterli, "Source-channel communication in sensor networks," in *Proc.of 2nd International Workshop of Info Processing Sensor Networks*, Palo Alto, 2003, pp. 162-177.

[8] M. K. Banavar, C. Tependelenlioglu, and A. Spanias, "Estimation Over Fading Channels with Limited Feedback Using Distributed Sensing," *IEEE Transactions on Signal Processing*, vol. 58, no. 1, pp. 415-425, 2010.

[9] A. Leong, S. Dey, and J. Evans, "Asymptotics and power allocation for state estimation over fading channels," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 1, pp. 611-633, January 2011.

[10] S. Cripps, *RF Power Amplifiers for Wireless Communications*. Norwood, MA: Artech House, 1999.

[11] S. Cui, J. J. Xiao, A. J. Goldsmith, Z. Q. Luo, and H.V. Poor, "Estimation Diversity and Energy Efficiency in Distributed Sensing," *IEEE Transactions on Signal Processing*, vol. 55, no. 9, pp. 4683-4695, September 2007.

[12] C. Tepedelenlioglu, M. Banavar, and A. Spanias, "On the Asymptotic Efficiency of Distributed Estimation Systems With Constant Modulus Signals Over Multiple-Access Channels," *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 7125-7130, October 2011.

[13] C. Tepedelenlioglu and A. Narasimhamhurthy, "Universal Distributed Estimation Over Multiple Access Channels With Constant Modulus Signaling ," *IEEE Transactions on Signal Processing*, vol. 58, no. 9, pp. 4783-4794, September 2010.

[14] R. A. Brockbank and C. A. A. Wass, "Non-Linear Distortion in Transmission Systems," *Journal of the Institution of Electrical Engineers - Part III: Radio and Communication Engineering*, vol. 92, no. 17, pp. 45-56, 1945.

[15] J. K. Cavers, "Amplifier linearization using a digital predistorter with fast adaptation and low memory requirements," *IEEE Transactions on Vehicular Technology*, vol. 39, no. 4, pp. 374-385, November 1990.

[16] A. J. Cann, "Nonlinearity Model with Variable Knee Sharpness," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 16, no. 6, pp. 874-877, 1980.

[17] Z. Q. Luo, "An isotropic universal decentralized estimation scheme for a bandwidth constrained ad hoc sensor network," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 4, pp. 735-744, April 2005.

[18] Feng Li, Jamie Evans, and Subhrakanti Dey, "Design of distributed detection schemes for multiple access channels," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 2, pp. 1552-1569, April 2012.

[19] T. M. Cover and J. A. Thomas, *Elements of Information Theory*.: John Wiley and Sons, 1991.

[20] R. Mudumbai, G. Barriac, and U. Madhow, "On the feasibility of distributed beamforming in wireless networks," *IEEE Transactions on Wireless Communications*, vol. 6, no. 5, pp. 1754-1763, May 2007.

[21] J. A. Bucklew and W. A. Sethares, "Convergence of a class of decentralized beamforming algorithms," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2280-2288, June 2008.

[22] W. Li and H. Dai, "Distrubed detection in wireless sensor networks using a multiple access channel," *IEEE Transactions on Signal Processing*, vol. 55, no. 3, pp. 822-833, March 2007.

[23] B. Porat, *Digital Processing of Random Signals: Theory and Methods*. Englewood Cliffs, NJ: Prentice-Hall, 1994.

[24] A. Papoulis and S. U. Pillai, *Probability, Random Variable, and Stochastic Processes*. New York: McGraw-Hill, 2002.

[25] S. Dasarathan and C. Tepedelenlioglu, "Distributed Estimation and Detection with Bounded Transmissions over Gaussian Multiple Access Channels", 2013, Available from Arxiv.org.

[26] J. J. Xiao, S. Cui, Z. Q. Luo, and A. J. Goldsmith, "Joint estimation in sensor networks under energy constraints," in *IEEE First Conference on Sensor and Ad Hoc Communications and Networks*, Santa Clara, 2004, pp. 264-271.

[27] J. J. Xiao, S. Cui, Z. Q. Luo, and A. J. Goldsmith, "Power scheduling of universal decentralized estimation in sensor networks," *IEEE Transactions on Signal Processing*, vol. 54, no. 2, pp. 413-422, Feb 2006.

[28] J. H. K. Vuolevi, T. Rahkonen, and J. P. A. Manninen, "Measurement Techniques for characterizing memory effects," *IEEE Transactions on Microwave Theory and Techniques*, vol. 49, no. 8, pp. 1383-1389, Aug 2001.

[29] J. C. Pedro and S. A. Maas, "A comparative overview of microwave and wireless power-amplifier behavioral modeling approaches," *IEEE Transactions on MIcrowave Theory and Techniques*, vol. 53, no. 4, pp. 1150-1163, 2005.

[30] A. Zhu, J. C. Pedro, and T. J. Brazil, "Dynamic Deviation Reduction-Based Volterra Behavrioal Modeling of RF Power Amplifiers," *IEEE Transactiosn on Microwave Theory and Techniques*, vol. 54, no. 12, pp. 4323-4332, December 2006.

[31] A. Zhu and T. Brazil, "Behavioral Modeling of RF Power Amplifiers Based on Pruned Volterra Series," *IEEE Microwave and Wireless Component Letters*, vol. 14, no. 12, pp. 563-565, December 2004.

[32] F. Mkadem, M. Ayed, S. Boumaiza, J. Wood, and P. Aaen, "Behavioral Modeling and digital predistortion of power amplifiers with memory using two hidden layers artificial neural networks," in *Proceedings of International Microwave Symposium*, Anaheim, 2010, pp. 656-659.

[33] K. Gard, L. Larson, and M. Steer, "The Impact of RF Front-End Characteristics on the Spectral Regrowth of Communications Signals," *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, no. 6, pp. 2179-2186, 2005.

[34] I. W. Sandberg, "Uniform Approximation with Doubly-Finite Volterra Series," in *IEEE International Sympoisum on Circuits and Systems*, Singapore, June 1991, pp. 754-757.

[35] D. R. Morgan, Z. Ma, J. Kim, M. G. Zierdt, and J. Pastalan, "A Generalized Memory Polynomial for Digital Predistortion of RF Power Amplifiers," *IEEE Transactions on Signal Processing*, vol. 54, no. 10, pp. 3852-3860, October 2006.

[36] C. J. Clark, G. Chrisikos, M. S. Muha, A. A. Moulthrop, and C. P. Silva, "Time-Domain Envelope Measurement Technique with Application to Wideband Power Amplifier Modeling," *IEEE Transactions on Microwave Theory and Techniques*, vol. 46, no. 12, pp. 2531-2540, December 1998.

[37] L. Ding et al., "Memory Polynomial Predistorter Based on the Indirect Learning Architecture," in *Global Telecomunnications Conferece*, Taipei, 2002, pp. 967-971.

[38] J. Kim and K. Konstantinou, "Digital Predistortion of wideband signals based on power amplifier model with memory," *Electronics Letters*, vol. 37, no. 23, pp. 1417-1418, November 2001.

[39] P. Jardin and G. Baudoin, "Filter Lookup Table Method for Power Amplifier Linearization," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 3, pp. 1076-1087, 2007.

[40] J. H. Van Vleck and D. Middleton, "The spectrum of clipped noise," *Proceedings of the IEEE*, vol. 54, no. 1, pp. 2-19, 1966.

[41] M. Rawat, K. Rawat, and F. Ghannouchi, "Adaptive digital predistortion of wireless power amplifiers/transmitters using dynamic real-valued focused time-delay line neural networks," *IEEE Transactions on Microwave Theory and Techniques*, vol. 58, no. 1, pp. 95-104, 2010.

[42] Y. Nagata, "Linear Amplification Technique for Digital Mobile Communications," in *39th IEEE Vehicular Technology Conference*, San Francisco., 1989, pp. 159-164.

[43] K. J. Muhonen, M. Kavehrad, and R. Krishnamoorthy, "Look-up table techniques for adaptive digital predistortion: a development and comparison," *IEEE Transactions on Vehicular Technology*, vol. 49, no. 5, pp. 1995-2002, 2000.

[44] B. Widrow, J. McCool, and M. Ball, "The Complex LMS Algorithm," *Proceedings of the IEEE*, vol. 63, no. 4, pp. 719-720, April 1975.

[45] H. Zhi-yong, G. Jian-hua, G. Shu-jian, and W. Gang, "An improved look-up table predistortion technique for HPA with memory effects in OFDM systems," *IEEE Transactions on Broadcasting*, vol. 52, no. 1, pp. 87-91, March 2006.

[46] B. Ai, Z. X. Yang, C. Y. Pan, S. G. Tang, and T. T. Zhang, "Analysis on LUT Based Predistortion Method for HPA with Memory," *IEEE Transactions on Broadcasting*, vol. 53, no. 1, pp. 127-131, March 2007.

[47] H. W. Kang, Y. S. Cho, and D. H. Youn, "On compensating nonlinear distortions of an OFDM system using an efficient adaptive predistorter," *IEEE Transactions on Communications*, vol. 47, no. 4, pp. 522-526, April 1999.

[48] T. Liu, S. Boumaiza, and F. Ghannouchi, "Dynamic behavioral modeling of 3G power amplifiers usign real-valued time-delay neural networks," *IEEE Transactions on Microwave Theory and Techniques*, vol. 3, no. 52, pp. 1025-1033, March 2004.

[49] S Boumaiza and F. Mkadem, "Wideband RF power amplifier predistortion using real-valued time-delay neural networks," in *Proceedings of the European Microwave Conference*, Rome, Italy, 2009, pp. 1449-1452.

[50] M. T. Hagan and M. B. Menhaj, "Training feedforward networks with the Marquardt algorithm," *IEEE Transactions on Neural Networks*, vol. 5, no. 6, pp. 989-993, 1994.

[51] H. Koeppl and P. Singerl, "An Efficient Scheme for Nonlinear Modeling and Predistortion in Mixed Signal Systems," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 53, no. 12, pp. 1368-1372, December 2006.

[52] G. E. Moore, "Cramming More Components onto Integrated Circuits," *Electronics*, pp. 114-117, April 1965.

[53] J. K. Cavers, "The Effect of Quadrature Modulator and Demodulator Errors on Adaptive Digital Predistiortion for Amplifier Linearization," *IEEE Transactions on Vehicular Technology*, vol. 46, no. 2, pp. 456-466, May 1997.

[54] J. K. Cavers, "New Methods for Adaption of Quadrature Modulators and Demodulations in Amplifier Linearization Circuits," *IEEE Transactions on Vehicular Technology*, vol. 46, no. 3, pp. 707-716, August 1997.

[55] R. Santucci and A. Spanias. (2012, November) Java-DSP for RF power amplifiers. [Online]. http://jdsp.engineering.asu.edu/RFPAweb/

[56] R. Santucci, A. Spanias, T. Gupta, and M. Shah, "Advanced functions of Java-DSP for use in electrical and computer engineering courses," in *American Society for Engineering Education*, Louisville., 2010.

[57] R. W. Santucci and A. S. Spanias, "Use of Java-DSP to Demonstrate Power Amplifier Linearization Techniques," in *American Society for Engineering Education*, Vancouver, 2011.

[58] R. Santucci and A. Spanias, "A block adaptive predistortion algorithm for transceivers with long transmit-receive latency ," in *2010 4th International*

Symposium on Communications, Control and Signal Processing (ISCCSP), Limassol, 2011.

[59] R. Santucci and A. Spanias, "Block Adaptive and Neural Network based Digital Predistortion and Power Amplifier Performance," in *Proc. of IASTED Signal Processing, Pattern Recognition, and Applications*, Innsbruck, 2011.

[60] G. Clark, S. Mitra, and S. Parker, "Block-implementation of adaptive digital filters," *Proceedings of the IEEE*, vol. 63, no. 4, pp. 719-720, April 1975.

[61] W. Feller, "The Strong Law of Large Numbers," in *An Introduction to Probability Theory and its Applications*. New York, USA: Wiley, 1968, ch. 10.7, pp. 243-245.

[62] R. Santucci, M. K. Banavar, C. Tepedelenlioğlu, and A. Spanias, "Estimation for Amplify-and-Forward Transmissions with Nonlinear Amplifiers," in *Proceedings of IET 2013 Constantinides International Workshop on Signal Processing*, London, 2013.

[63] R. Santucci, M. K. Banavar, C. Tepedelenlioglu, and A. Spanias, "Energy-efficient distributed estimation by utilizing a nonlinear amplifier," *IEEE Transactions on Circuits and Systems I: Regular Papers*, p. Accepted for Publication, January 2014.

[64] Agilent. (2012, November) Agilent Amplifier DesignGuide. [Online]. http://edocs.soco.agilent.com/display/ads2009/Amplifier+DesignGuide

[65] A. J. Cann, "Improved Nonlinearity Model with Variable Knee Sharpness," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 4, pp. 3937-3646, October 2012.

[66] R. Santucci, M. K. Banavar, A. Spanias, and C. Tepedelenlioglu, "Design of limiting amplifier models for nonlinear amplify-and-forward distributed estimation," in *2013 18th International Conference on Digital Signal Processing*, Thira, 2013, pp. 1-6.

[67] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*. Washington DC, United States: United States Government Printing Office, 1972.

[68] X. Zhang et al., "Performance for Comparison of Localization Tehcniques for Sequential WSN Discovery," in *Proceeding of Sensor Signal Processing for Defence*, London, UK, 2012, pp. 1-5.

[69] M. K. Banavar, C. Tepedelenlioglu, and A. Spanias, "Distributed SNR Estimation with Power Constrained Signaling over Gaussian Multiple-Access Channels," *IEEE Transactions on Signal Processing*, vol. 60, no. 6, pp. 3289-2394, June 2012.

[70] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, MA: Cambridge University Press, 2004.

[71] S. Coleri, M. Ergen, A. Puri, and A. Bahai, "Channel Estimation Techniques Based on Pilot Arrangement in OFDM Systems," *IEEE Transactions on Broadcasting*, vol. 48, no. 3, pp. 223-229, September 2002.

[72] A. Narasimhamurthy, M. K. Banavar, and C. Tepelenlioglu, *OFDM Systems for Wireless Communications*. United States of America: Morgan & Claypool, 2010.

[73] L. Ye and L. J. Cimini, "Bounds on the Interference of OFDM in Time-Varying Impairments," *IEEE Transactions on Communications*, vol. 49, no. 3, pp. 401-404, March 2001.

[74] O. Edfors, M. Sandell, J. J. van de Beek, D. Landstrom, and F. Sjoberg, "An Introduction to Orthogonal Frequency Division Multiplexing," Lulea University of Technology, Lulea, Sweden, 1996.

[75] S. H. Han and J. H. Lee, "An Overview of Peak-to-Average Power Ratio Reduction Techniques for Multicarrier Transmission," *IEEE Wireless Communications*, vol. 12, no. 2, pp. 56-65, April 2005.

[76] B. S. Krongold and D. L. Jones, "An Active-Set Approach for OFDM PAR Reduction via Tone Reservation," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 495-509, February 2004.

[77] C. Tepedelenlioglu, A. Abdi, and G. B. Giannakis, "The Rician K Factor: Estimation and Performance Analysis," *IEEE Transactions on Wireless Communications*, vol. 2, no. 4, pp. 799-810, July 2003.