ULTRA-WIDEBAND NONLINEAR ECHO-CANCELLATION

A Thesis Submitted to the
College of Graduate and Postdoctoral Studies
in Partial Fulfillment of the Requirements
for the Degree of Master of Science
in the Department of Electrical and Computer Engineering
University of Saskatchewan
Saskatoon

by

Lance Pitka

Permission to Use

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department of Electrical and Computer Engineering or the Dean of the College of Graduate and Postdoctoral Studies at the University of Saskatchewan. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Disclaimer

Reference in this thesis/dissertation to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement, recommendation, or favoring by the University of Saskatchewan. The views and opinions of the author expressed herein do not state or reflect those of the University of Saskatchewan, and shall not be used for advertising or product endorsement purposes.

Request for permission to copy or to make any other use of material in this thesis in whole or in part should be addressed to:

Head of the Department of
Electrical and Computer Engineering
University of Saskatchewan
3B78 Engineering Building
57 Campus Drive
Saskatoon, Saskatchewan S7N 5A9
Canada

OR

Dean
College of Graduate and Postdoctoral Studies
University of Saskatchewan
116 Thorvaldson Building
110 Science Place
Saskatoon, Saskatchewan S7N 5C9
Canada

Acknowledgments

I would like to thank my co-supervisors Dr. Eric Salt and Dr. Brian Berscheid for their help and guidance throughout my research and the writing process. Specifically, Dr. Salt spent a great deal of time preparing short courses to help me with my project, and Dr. Berscheid assisted in finding solutions to the most challenging problems I faced. This thesis would not have been possible without these two professors.

I would also like to thank my parents for their love and support throughout my university career. I wouldn't be the person I am today without them. My success throughout university is largely accredited to their encouragement.

Finally, I would like to thank my girlfriend, Kelly, for always being there for me, for reminding me to enjoy life outside of school, and for supporting all of my endeavors. Completing this master's degree would have been much more difficult without her.

Abstract

Hybrid fiber coaxial (HFC) networks are used around the world to distribute cable television and broadband internet services to customers. These networks are governed by the Data-Over-Cable Service Interface Specification (DOCSIS) family of standards, with the most recent version at the time of this writing being DOCSIS 3.1. A frequency division duplex (FDD) spectrum is used in DOCSIS 3.1, where the upstream and downstream signals are separated in frequency to eliminate interference. A possible method to increase signal bandwidths is to use a full-duplex (FDX) spectrum, in which the US and DS signals use the same frequencies at the same time.

A main challenge faced when implementing FDX in a DOCSIS node is eliminating the interference in the received US signal caused by the transmitted DS signal. One possible method for eliminating the interference is utilizing an echo-canceling algorithm, which predicts the self-interference (SI) based on the known DS signal and cancels it from the received US signal.

Although echo-cancellation algorithms exist for fundamentally similar applications, the DOCSIS FDX case is more complicated for two main reasons. First, the DOCSIS node uses a nonlinear power amplifier to amplify the DS signal. Second, the DS signal is an ultra-wideband signal spanning a frequency range of up to 1.2 GHz. Most of the amplifier modeling techniques discussed in the literature were designed for narrowband wireless signals and will have limited performance when used with ultra-wideband signals.

This thesis develops an algorithm to characterize the power amplifier and to predict the harmonics it generates for a given DS signal. These predicted harmonics can be used to cancel the SI signal in a full duplex DOCSIS system. The algorithm, which is referred to as the ultra-wideband memory polynomial (UWB-MP) model, is based on the well-known memory polynomial model with adaptations which allow the model to predict harmonics for ultra-wideband signals.

Since a direct implementation of the UWB-MP model in an FPGA would result in very high resource usage, system architecture recommendations are provided. Our proposed im-

plementation of the model compensates for harmonics up to and including the 3^{rd} order, which has a power spectrum extending above 3600 MHz. Using the techniques discussed in this thesis, it is shown that a sampling rate of 4 GHz allows for cancellation of the SI signal while providing a reasonable balance between performance and resource usage.

Matlab simulations of a DOCSIS node with various parameters and PA simulation models were conducted. The simulations showed that over 75 dB of cancellation of the SI signal is possible in an idealized hardware setup. It is also demonstrated that AWGN injected into the received signal does not reduce the ability of the model to estimate the PA harmonics, although the noise itself cannot be canceled. Further simulations showed that the UWB-MP model could cancel harmonics whose power is much higher than that specified in DOCSIS. Although the UWB-MP model was designed with memory polynomial type PAs in mind, simulation results show that significant cancellation is possible with PAs that are represented by Wiener models as well. Based on the simulation results, we recommend using a filter of length 20 coefficients for each harmonic in the UWB-MP model, and 60 iterations with 500 samples for estimating the coefficients with the least squares method.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Ω	Frequency (radians/second)
ω	Frequency (radians/sample)
ω_c	Carrier frequency (radians/sample)
Ω_s	Sampling rate ($2\pi\times$samples/second)
ω_{DS}	Downstream carrier frequency (rad/sample)
ω_{US}	Upstream carrier frequency (rad/sample)
F	Frequency (cycles/second)
f	Frequency (cycles/sample)
f_c	Carrier frequency (cycles/sample)
F_s	Sampling rate (samples/second)
ADC	Analog-to-Digital Converter
AM	Amplitude Modulation
AM-AM	Amplitude Modulation - Amplitude Modulation
AM-PM	Amplitude Modulation - Phase Modulation
ASIC	Application-Specific Integrated Circuit
BPF	Band Pass Filter
CDMA	Code-Division Multiple Access
CM	Cable Modem
CMTS	Cable Modem Termination System
CS	Compressed Sensing
CT	Continuous-Time
DAC	Digital-to-Analog Converter

dB	Decibel
dBm	Decibel-milliwatts
DOCSIS	Data Over Cable Service Interface Specification
DPD	Digital Predistortion
DS	Downstream
DSL	Digital Subscriber Line
DSP	Digital Signal Processing
DT	Discrete-Time
DTFT	Discrete-Time Fourier Transform
EC	Echo Cancellation
EMNE	Effectively Memoryless Nonlinear Element
FDD	Frequency-Division Duplex
FDX	Full-Duplex
FFT	Fast Fourier Transform
FPGA	Field-Programmable Gate Array
FTTH	Fiber-To-The-Home
Gbps	Gigabits per second
GHz	Gigahertz
HFC	Hybrid Fiber-Cable
HPF	High Pass Filter
Hz	Hertz
IBO	Input Back-Off
IDTFT	Inverse Discrete-Time Fourier Transform
IFFT	Inverse Fast Fourier Transform
IP	Input Power

ISP	Internet Service Provider
LNA	Low Noise Amplifier
LPF	Low Pass Filter
LS	Least Squares
LTI	Linear Time-Invariant
Mbps	Megabits per second
MHz	Megahertz
MNE	Memoryless Nonlinear Element
NEWM	Nonlinear Element With Memory
NMSE	Normalized Mean Square Error
OFDM	Orthogonal Frequency-Division Multiplexing
OFDMA	Orthogonal Frequency-Division Multiple Access
OP	Output Power
PA	Power Amplifier
PAPR	Peak-to-Average Power Ratio
PHY	Physical Layer
PM	Phase Modulation
PSD	Power Spectral Density
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
RF	Radio Frequency
Rx	Receive
SI	Self-Interference
SNR	Signal-to-Noise Ratio
SOI	Signal Of Interest

SSPA	Solid State Power Amplifier
Tx	Transmit
US	Upstream

1. Introduction

1.1 Cable Television History

Cable television, which was once referred to as Common Antenna Television (CATV), started as a one-way television network from a single high gain receiving antenna over a coaxial cable network to several televisions [1]. Over the decades as the demand for higher data rates and the need for two-way communication grew, so did the technology supporting the transmission and reception. The network evolved to support two-way communication and data communication.

The cable industry changed with the advancement of the internet. Initially, telephone companies offered internet service with digital subscriber line (DSL) technology, but as the cable industry transitioned from analog to digital and implemented hybrid-fiber coaxial (HFC) networks, it was able to compete with telephone companies as an internet service provider (ISP) [1].

At the time, no industry-wide standards were governing the equipment used to provide cable and internet services. The lack of standards resulted in the service providers being forced to purchase proprietary equipment at whatever cost the equipment provider demanded. Furthermore, the lack of standards meant that equipment provided by competing hardware vendors was frequently incompatible, causing service providers to be "locked in" to a particular vendor. This lead to the services providers collaborating to form CableLabs [1], a not-for-profit organization which creates open standards for cable and internet equipment to ensure compatibility between different equipment manufacturers. One essential standard developed by CableLabs is the Data-Over-Cable Service Interface Specification, referred to as DOCSIS.

1.2 History of DOCSIS

DOCSIS 1.0 was released by CableLabs in 1997. The physical network described by DOCSIS consists of a cable modem termination system (CMTS) connected to multiple cable modems (CMs) via an HFC network as shown in Figure 1.1. The CMTS in the service provider's headend connects to many optical nodes via fiber optic cables. The optical node receives the optical signal from the fiber and modulates it as a voltage signal on the copper coaxial line connecting to multiple taps. The taps send a portion of the signal power on the cable to the CMs in subscribers' homes.

A DOCSIS network enables bidirectional communication. Downstream (DS) transmission refers to transmission from the CMTS to the CMs, and upstream (US) transmission refers to transmission from the CMs to the CMTS. Over the years CableLabs has updated the standard to improve both US and DS data rates. However, the physical network has remained relatively unchanged. Transmission rates between the node and the CM's via coaxial cable is the bottleneck in the system since fiber allows for higher transmission rates, so updates to the DOCSIS standard typically involve increasing the transmission rate between the node and CMs. DOCSIS updates have included versions 1.1, 2.0, 3.0, and 3.1, which was released in 2013 and is the most recent version.

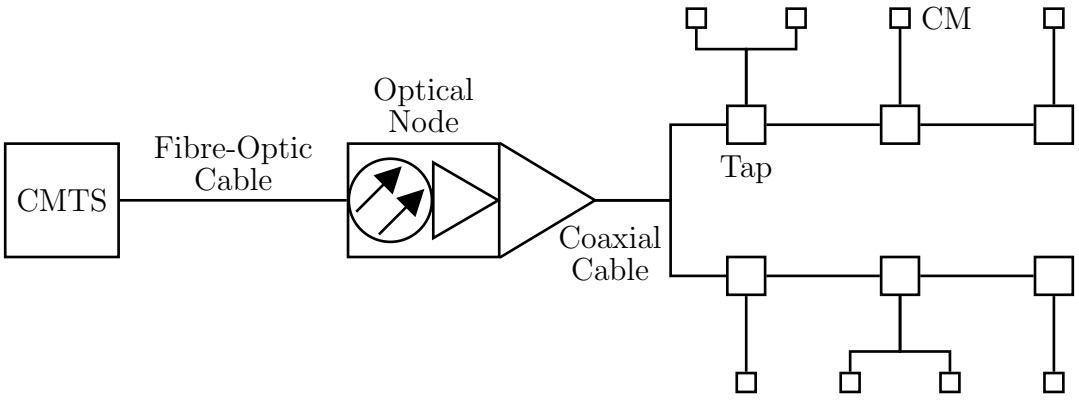


Figure 1.1: Hybrid fiber-cable network architecture

It should be noted that the update from version 1.0 [2] to 1.1 [3] was focused on quality of service (QoS), and these two versions are often collectively referred to as DOCSIS 1.x. The

main changes from DOCSIS 1.x to 2.0 [4] included enhanced upstream data rates and the addition of synchronous code-division multiple access (S-CDMA) in the upstream direction. The upgrade from DOCSIS 2.0 to 3.0 [5] significantly increased the data rates in both the upstream and downstream directions. The increases were mainly due to significant improvements to the modulation techniques used to transmit data over the coaxial network.

1.3 DOCSIS 3.1

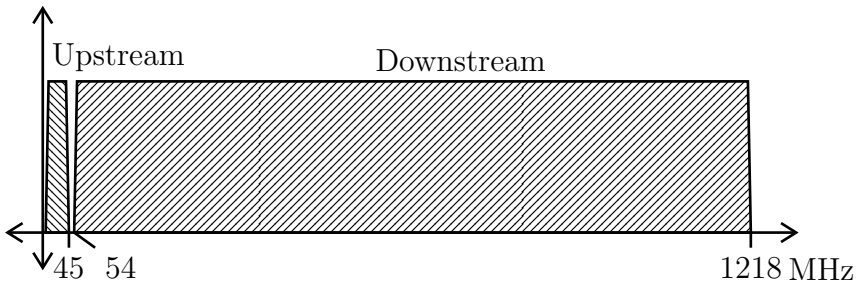
DOCSIS 3.1 [6] includes significant improvements over DOCSIS 3.0. Most notably, the maximum downstream capacity increased from 1.2 Gbps to 10 Gbps, and the maximum upstream capacity increased from 100 Mbps to 1 Gbps. The frequency ranges allocated to transmission in the US and DS directions in the coaxial cable network have increased considerably as well. The US signals can occupy frequencies between 5 MHz to 204 MHz, and the DS signals can occupy frequencies in the range of 54 MHz to 1218 MHz (and up to 1800 MHz, but this is not commonly used).

Since the US and DS signals coexist on the same coaxial cable, there is potential for them to interfere with each other if the system is not carefully designed. DOCSIS avoids this problem by using a scheme called frequency-division duplexing (FDD), in which the US and DS signals are separated into non-overlapping frequency bands. DOCSIS 3.1 provides three different options for this partitioning to accommodate a variety of network architectures. A summary of the different frequency band options allowed in DOCSIS 3.1 is given in Table 1.1 and illustrated in Figure 1.2.

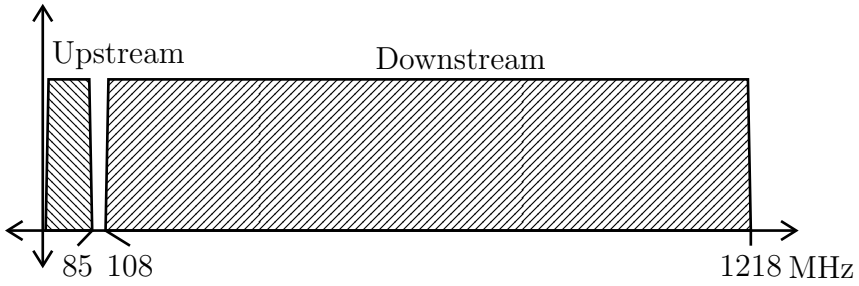
Option #	Upstream Frequency Band (MHz)	Downstream Frequency Band (MHz)
1	5-45	54-1218
2	5-85	108-1218
3	5-204	258-1218

Table 1.1: DOCSIS 3.1 frequency band options

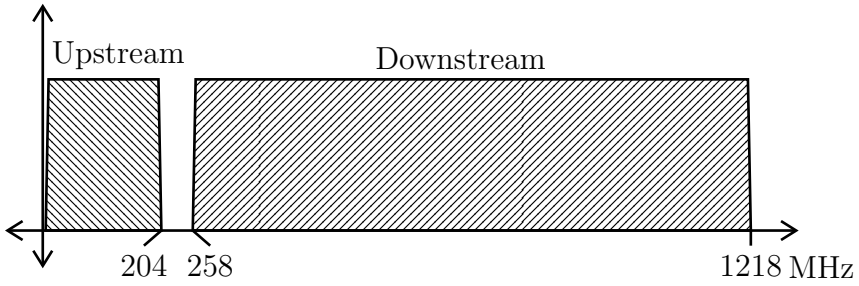
The increase in data rates in DOCSIS 3.1 is made possible by implementing orthogonal



(a) FDD spectrum option 1



(b) FDD spectrum option 2



(c) FDD spectrum option 3

Figure 1.2: DOCSIS 3.1 frequency division duplex spectrum options

frequency-division multiplexing (OFDM) modulation in the downstream, and orthogonal frequency-division multiple access (OFDMA) in the upstream. OFDM is a digital multi-carrier modulation scheme that transmits over a large number of closely spaced narrow orthogonal subcarriers instead of transmitting at a high rate over one wider subcarrier.

An OFDM signal is created by encoding the data to be transmitted (Tx) data into frequency components of a signal, then converting the signal to its time domain equivalent by

using an inverse fast Fourier transform (IFFT). The time domain signal is then transmitted via coax to a receiver, in which it is converted back to frequency domain with a fast Fourier transform (FFT). The original data is then recovered from the FFT output. An IFFT is defined by

$$x_n = \sum_{k=0}^{N-1} X_k e^{\jmath 2\pi kn/N} \tag{1.1}$$

where X_k is a complex frequency domain signal, x_n is a complex time domain signal, and N is the IFFT order, or the number of subcarriers defined. Similarly, an FFT is defined by

$$X_k = \sum_{n=0}^{N-1} x_n e^{-\jmath 2\pi kn/N}. \tag{1.2}$$

OFDMA is similar to OFDM except it assigns subcarriers to multiple devices, allowing for simultaneous data transmission from multiple devices over the same medium. OFDM on the other hand only allows for one device to be transmitting at a time over the entire used frequency spectrum.

One side effect of using multiple narrow subcarriers in OFDM and OFDMA signals is that the output of the N-point IFFT is the sum of N independent random variables which leads to the time domain signal having an amplitude distribution which is approximately Gaussian. This causes the OFDM and OFDMA signals to have a high peak-to-average power ratio (PAPR), primarily if a large number of subcarriers are used. In a DOCSIS 3.1 system, either 4096 or 8192 subcarriers are used in an OFDM block, potentially resulting in extremely high signal peaks if most or all of the subcarriers align in time. Fortunately, the probability of such events is extremely low.

A simplified diagram of the optical node in a DOCSIS 3.1 network is shown in Figure 1.3. The physical layer (PHY) device is what communicates with the CMTS in the head-end. The downstream PHY converts downstream digital data from the CMTS and converts it to an analog RF signal to be transmitted over a coaxial cable and sent to cable modems [7]. In Figure 1.3 the digital-to-analog converter (DAC) and the DS PHY are shown separately, but they may be contained in the same integrated circuit. Before the RF signal can be sent to the coaxial cable (coax), it first has to be adjusted to compensate for the attenuation of

the coax network. Long coax cables attenuate high frequencies more than low frequencies. To compensate for this attenuation, a device is used to add a tilt in the frequency domain to the transmitted signal with an opposite slope compared to the attenuation in the cable. CableLabs standards call for a 21 dB tilt from 108 MHz to 1218 MHz. After the tilt compensator, the RF signal is passed through a power amplifier to increase the power to a high enough level such that the signal can be received by the CMs located farthest away.

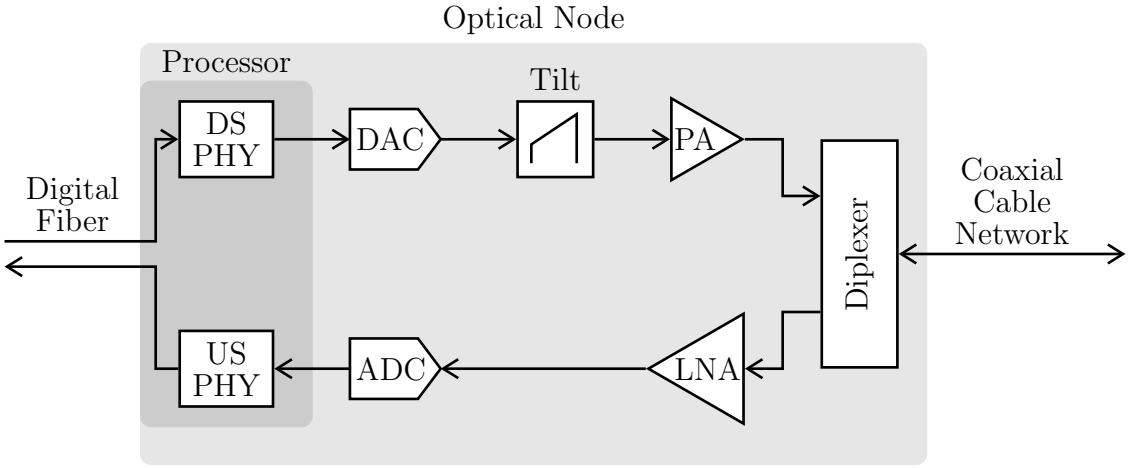


Figure 1.3: DOCSIS 3.1 optical node to coax interface

In both the optical node and the cable modems, a device called a diplexer is used to separate the upstream and downstream signals that jointly occupy the coaxial cable. The diplexer acts as a high pass filter (HPF) for the downstream data, and serves as a low pass filter (LPF) for the upstream data, as illustrated in Figure 1.4 [6]. The combination of the high pass and low pass filters prevents any part of the downstream signal from interfering with the upstream signal. The power loss between the input and output ports of the diplexer, referred to as the insertion loss, is usually around 0.5 dB [8].

The upstream data follows an opposite path in the node compared to the downstream data, as shown in Figure 1.3. The US data is sent from the diplexer through a low noise amplifier (LNA) to boost the received signal before it is sampled by the analog-to-digital converter (ADC). The US PHY demodulates the received signal and sends the data to the CMTS via a digital fiber optic line.

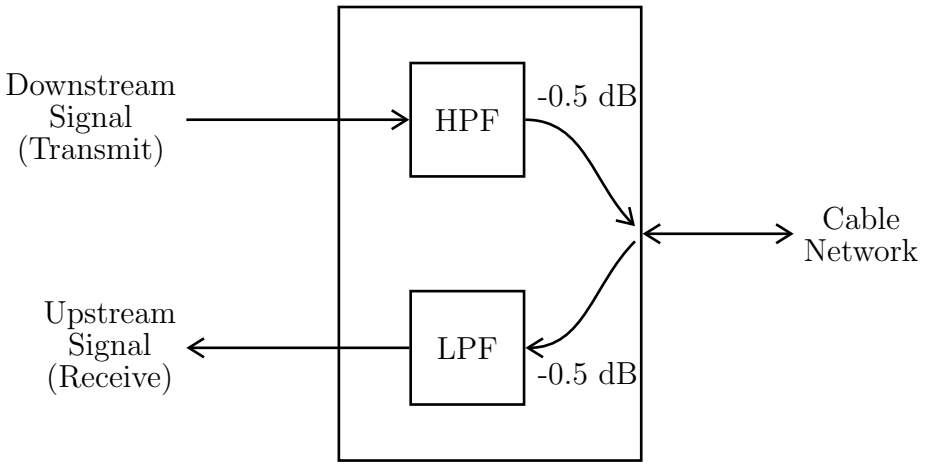


Figure 1.4: Diplexer

1.4 DOCSIS 3.1 Full-Duplex

Cable and internet subscribers continue to demand higher data-rates, especially in the upstream direction as services such as video conferencing and gaming become more popular. Cable companies need to increase the throughput of their networks to remain competitive with fiber-to-the-home (FTTH) networks, which run a fiber line all the way to the customer's location and eliminate the coax bottleneck. One method for increasing the upstream capacity is to simultaneously transmit upstream and downstream on the same frequency band at the same time, which is known as full-duplex (FDX). Full-duplex communication has the potential to vastly increase the upstream data rates compared to DOCSIS 3.1 by allowing more of the spectrum to be used by the upstream signal without significantly reducing the achievable DS data rates. FDX is advantageous over FTTH because it can use the HFC network that is already in service without requiring costly fiber installations.

In October 2017, CableLabs released the first version of the DOCSIS 3.1 Full-Duplex standard. The standard dictates that for FDX operation, the downstream signal will occupy frequencies from 108 MHz to 1218 MHz and the upstream signal will occupy 5 MHz to 684 MHz. This implies that there is a region in the spectrum (between 108 MH and 684 MHz) which carries both US and DS data simultaneously, as illustrated in 1.5.

The main change in hardware components in implementing FDX compared to FDD is

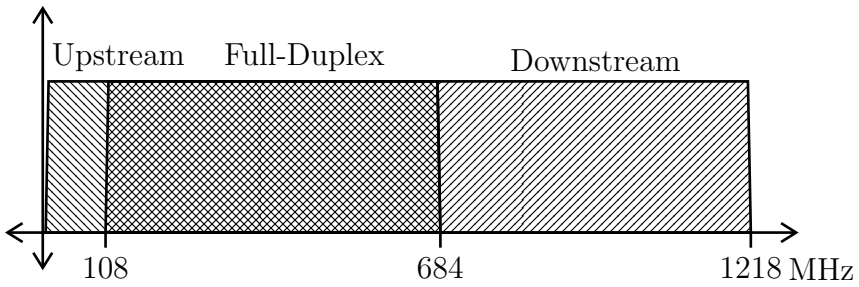


Figure 1.5: Full-duplex spectrum

the diplexer can no longer be used since the upstream and downstream signals will occupy the same frequencies. Instead, a 2-wire to 4-wire directional coupler will have to be used to couple the upstream and downstream signals in the node to the coaxial cable network. The 2-wire side of the coupler is the coaxial cable, and the 4-wire side consists of the upstream and downstream signals in the node. A diagram of a coupler is shown in Figure 1.6.

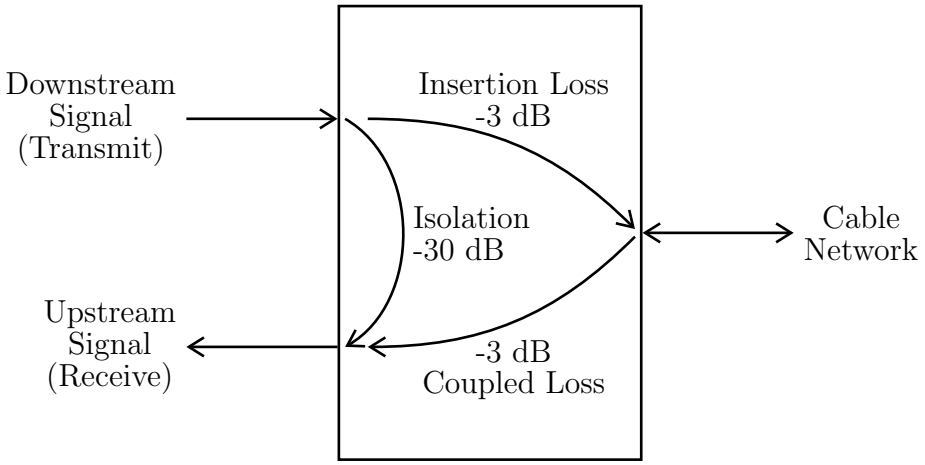


Figure 1.6: Directional coupler

A directional coupler is a passive RF device used to couple a specific portion of the power in one transmission line out through another connection or port. In our particular case, the coupler is used to couple the upstream signal and the downstream signal to the cable network. The power propagating down the coaxial cable network toward the node enters the 2-wire side of the coupler. The coupler splits the power and delivers portions to each of

the two ports on the 4-wire side of the coupler[1]. The power delivered to the transmit port is consumed by the output resistance and therefore wasted. This power loss causes about a 3 dB attenuation between the coupled and received port, referred to as the coupled loss. The power entering the coupler from the transmit port on the 4-wire side is split with half that power being delivered to the cable and the other half delivered to an internally terminated port (not shown in the diagram). This causes about a 3 dB loss from the transmit to coupled port, which is referred to as insertion loss. In theory, the two ports on the 4-wire side are completely isolated, but in practice, there is a leakage path that couples the transmit port to the receive port. In practice, the leakage is about 30 dB below the transmit level [9].

The other change in hardware involves inserting a LPF after the LNA and before the analog-to-digital converter (ADC). This low pass filter is necessary because the coupler does not act as a LPF to the US data as the diplexer did. The LPF removes unwanted high-frequency components that would otherwise alias when sampled by the ADC. A diagram of the FDX node is shown in Figure 1.7.

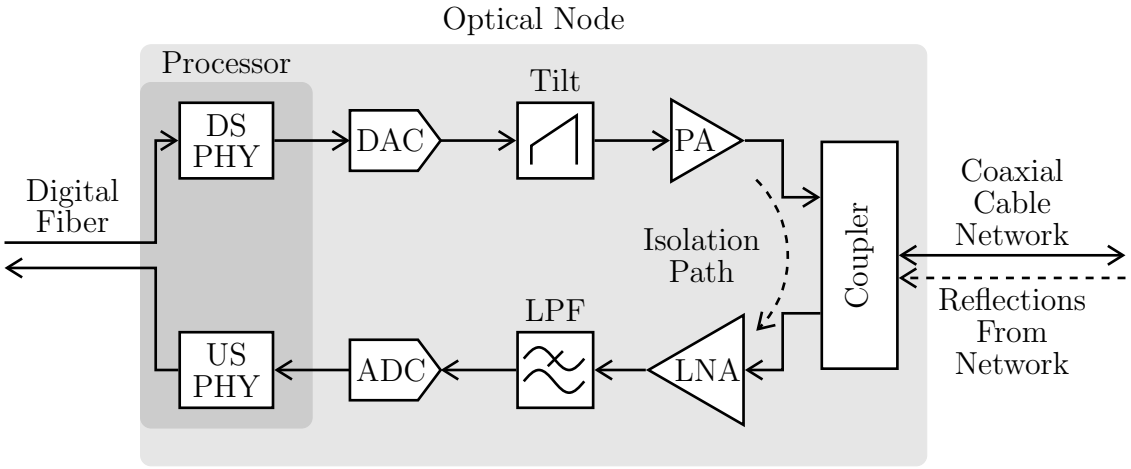


Figure 1.7: DOCSIS 3.1 full-duplex optical node to coax interface

Using a directional coupler instead of a diplexer adds two main challenges in implementing full-duplex. The first challenge is dealing with the self-interference (SI) created when the DS

[1]Couplers with various "coupling factors" are available, allowing the portions of power going to the US and DS ports to be controlled. In the following discussion, a 3 dB coupler is assumed, meaning the power is split equally between the two ports

signal leaks into the US path. The self-interference comes from both the isolation path in the coupler and from the DS signal being reflected due to imperfections in the cable network and returning through the coupler into the US path, as shown in Figure 1.7.

Since the DS and US data occupy the same frequencies in the FDX band, the self-interference overlaps with the upstream signal and cannot be directly filtered out. An alternative way to deal with this self-interference is digital echo cancellation (EC). Digital echo cancellation involves digitally modeling the path the known DS signal takes from the transmitter to the receiver and using the model to construct a replica of the SI signal seen in the receiver. If done correctly, the modeled signal and the received SI signal should be nearly identical, and the SI can be canceled by the model signal leaving only the US signal of interest for the receiver.

The other primary challenge imposed by the directional coupler is the increase in insertion loss compared to the diplexer. The diplexer has 0.5 dB insertion loss and the coupler has about 3 dB insertion loss, meaning the downstream signal must be transmitted with 2.5 dB more power to compensate. To achieve this, the power amplifier (PA) in the node must output 2.5 dB more power. However, driving the PA at a higher power has trade-offs, as discussed in the following section.

1.5 Basic Amplifier Theory

The purpose of a power amplifier is to increase the power of an input signal. In the context of a DOCSIS network, this allows the signal to travel a larger distance to reach a CM while maintaining an adequate signal level. Ideally, the power amplifier would be a linear device in which the output is equal to the input multiplied by a constant, i.e., $v_o(t) = Gv_i(t)$. However, in practical amplifiers, the output becomes nonlinear as the input signal level increases. The output is then not only linearly dependent on the input but also on higher orders of the input, referred to as harmonics. For example, for input $v_i(t)$, the output can be represented with the Taylor Series expansion [10]

$$v_o(t) = a_0 + a_1 v_i(t) + a_2 v_i^2(t) + a_3 v_i^3(t) + \dots \tag{1.3}$$

where a_n is the coefficient for the n^{th} order harmonic. The constant term, with coefficient a_0, leads to rectification, converting the AC input signal to DC. The linear term, a_1, is the linear gain of the nonlinear device, and therefore $a_1 > 1$ for an amplifier. The higher order coefficients $(a_2, a_3, ...)$ represent the dominance of the higher order output voltages created by the PA, which will be referred to as harmonics.

The function defining $v_o(t)$ differs from device to device based on the dominance of particular terms in the Taylor expansion. Practical nonlinear devices usually have a series expansion containing many non-zero terms. Some of the critical impacts of the power amplifier nonlinear behaviour will be discussed next.

1.5.1 Gain Compression

First consider the case where a single frequency sinusoid is applied to the input of an amplifier:

$$v_i = V_0 cos(\omega_0 t) \tag{1.4}$$

Equation (1.3) gives the output voltage as

$$\begin{aligned} v_o &= a_0 + a_1 V_0 cos\omega_0 t + a_2 V_0^2 cos^2\omega_0 t + a_3 V_0^3 cos^3\omega_0 t + \dots \\ &= \left(a_0 + \frac{1}{2}a_2 V_0^2\right) + \left(a_1 V_0 + \frac{3}{4}a_3 V_0^3\right) cos\omega_0 t + \frac{1}{2}a_2 V_0^2 cos2\omega_0 t \\ &+ \frac{1}{4}a_3 V_0^3 cos3\omega_0 t + \dots \end{aligned} \tag{1.5}$$

The resulting voltage gain of the signal component at frequency ω_0 is:

$$G_v = \frac{v_o^{\omega_0}}{v_i^{\omega_0}} = \frac{a_1 V_0 + \frac{3}{4}a_3 V_0^3}{V_0} = a_1 + \frac{3}{4}a_3 V_0^2 \tag{1.6}$$

where only terms up to the third order have been considered.

The result of Equation (1.6) contains the coefficient of the linear term, a_1, as expected, but also an additional term proportional to the square of the input amplitude. In most power amplifiers, a_3 will have the opposite sign of a_1 and thus the output of the amplifier will be reduced from the expected linear gain for large values of input voltage V_0. This effect is called *gain compression* or *saturation* [10].

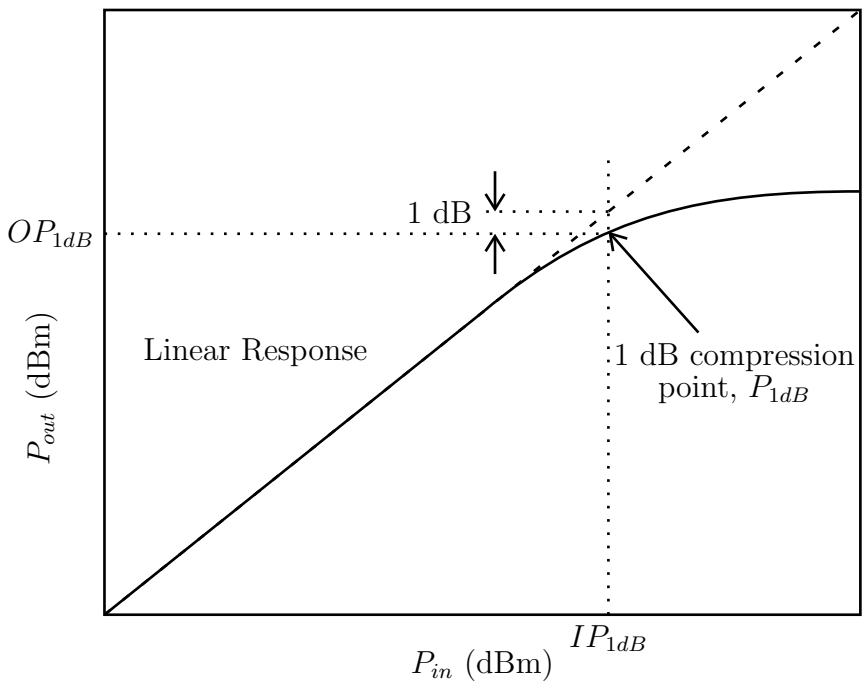


Figure 1.8: Power amplifier 1dB compression point

A typical plot of the output power versus input power for an amplifier is shown in Figure 1.8. An ideal amplifier would have a linear slope of a_1 from equation (1.3). However practical amplifiers saturate as illustrated in Figure 1.8. The linear operating range of a device is quantified by defining the *1 dB compression point* as the power level where the power has decreased by 1 dB from the ideal linear characteristic [10]. This power level is denoted by either the input power (IP_{1dB}) or output power (OP_{1dB}).

The nonlinear effects causing gain compression also leads to amplitude modulation (AM-AM) and phase modulation (AM-PM). AM-AM is the amplitude conversion as a function of the input amplitude to the amplifier, and AM-PM is phase conversion as a function of the input amplitude [10]. In other words, if the amplitude at the input of an amplifier is A, then the amplitude at the output of the amplifier is $G(A)$, where $G(\cdot)$ is the AM-AM conversion of the amplifier. This conversion is not linear as is shown in Figure 1.8. Likewise, the phase response of the amplifier is $F(A)$ where $F(\cdot)$ is the AM-PM conversion and is a function of the input amplitude. $F(\cdot)$ is not necessarily a linear function of A either.

1.5.2 Harmonic and Intermodulation Distortion

A nonlinear device such as a power amplifier with a single input frequency ω_0 will generate an output with harmonics of the input frequency at $n\omega_0$, for $n = 0, 1, 2, \ldots$. However, the situation is different when the input signal consists of two closely spaced frequencies. Note that an OFDM signal has many closely spaced frequencies.

Consider a two-tone input voltage, consisting of two closely spaced frequencies ω_1 and ω_2:

$$v_i = V_0(cos\omega_1 t + cos\omega_2 t) \tag{1.7}$$

From equation (1.3) the output is

$$\begin{aligned}
v_o &= a_0 + a_1 V_0(cos\omega_1 t + cos\omega_2 t) + a_2 V_0^2(cos\omega_1 t + cos\omega_2 t)^2 \\
&\quad + a_3 V_0^3(cos\omega_1 t + cos\omega_2 t)^3 + \ldots \\
&= a_0 + a_1 V_0 cos\omega_1 t + a_1 V_0 cos\omega_2 t + \frac{1}{2}a_2 V_0^2(1 + cos2\omega_1 t) + \frac{1}{2}a_2 V_0^2(1 + cos2\omega_2 t) \\
&\quad + a_2 V_0^2 cos((\omega_1 - \omega_2)t) + a_2 V_0^2 cos((\omega_1 + \omega_2)t) \\
&\quad + a_3 V_0^3\left(\frac{3}{4}cos\omega_1 t + \frac{1}{4}cos3\omega_1 t\right) + a_3 V_0^3\left(\frac{3}{4}cos\omega_2 t + \frac{1}{4}cos3\omega_2 t\right) \\
&\quad + a_3 V_0^3\left[\frac{3}{2}cos\omega_2 t + \frac{3}{4}cos((2\omega_1 - \omega_2)t) + \frac{3}{4}cos((2\omega_1 + \omega_2)t)\right] \\
&\quad + a_3 V_0^3\left[\frac{3}{2}cos\omega_1 t + \frac{3}{4}cos((2\omega_2 - \omega_1)t) + \frac{3}{4}cos((2\omega_2 + \omega_1)t)\right] + \ldots
\end{aligned} \tag{1.8}$$

where standard trigonometric identities have been used to expand the initial expression. It is seen that the output spectrum consists of harmonics at frequencies

$$m\omega_1 + n\omega_2 \tag{1.9}$$

with $m, n = 0, \pm 1, \pm 2, \ldots$. The combinations of these frequencies are called *intermodulation products* and the order of a given product is defined as $|m| + |n|$ [10]. The spectrum of two-tone intermodulation of the second and third order is illustrated in Figure 1.9. Each frequency component in the figure is labeled with its associated harmonic order.

Further expanding Equation (1.8) reveals that odd order products lie in odd multiples of the center frequency ($\omega_c = \frac{\omega_1 + \omega_2}{2}$), and even order products lie in even multiples of the

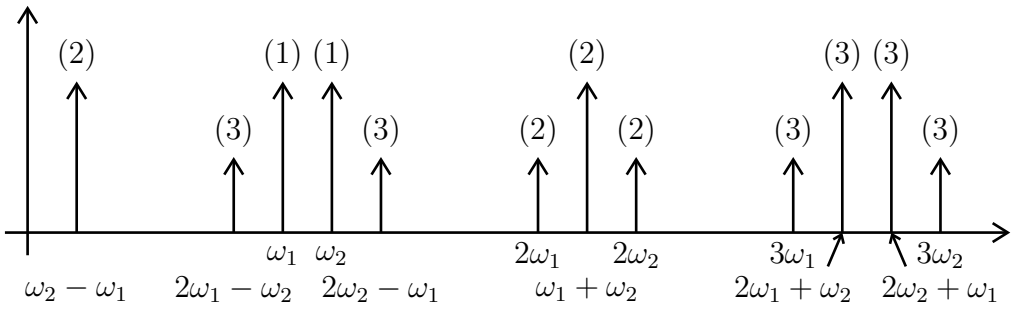


Figure 1.9: Intermodulation spectrum of second and third order two-tone signals

center frequency. For example, a 5^{th} order harmonic has frequency components at ω_c, $3\omega_c$, and $5\omega_c$, and a $4th$ order harmonic has frequency components at $0\omega_c$, $2\omega_c$, and $4\omega_c$. Further, the frequency components belonging to a particular harmonic order are spread across a range of frequencies proportional to the harmonic order. For example, the 2^{nd} order harmonics are twice as wide in frequency as the original signal, and the 3^{rd} order harmonics are three times as wide as the original signal. These frequency translation and spreading effects are illustrated in Figure 1.10[2] for harmonics up to the 5th order.

1.5.3 Memory Effects

Power amplifiers typically exhibit memory effects, meaning the output is not only dependent on the current input but also on past inputs. Regarding memory, nonlinear devices can be classified into three categories: memoryless nonlinear elements (MNE's), effectively memoryless nonlinear elements (EMNE's), and nonlinear elements with memory (NEWM's) [12].

MNE's respond to the input signals without delay. They can be characterized by:

$$v(t) = f[Acos(\omega t + \phi(t))] \tag{1.10}$$

where $Acos(\omega t + \phi(t))$ is the sinusoidal input voltage, $v(t)$ is the output voltage, and $f[\cdot]$ is the function defining the power amplifier. $v(t)$ can be put into the alternate form,

$$v(t) = G(A)cos(\omega t + \phi(t)) \tag{1.11}$$

[2]Figure derived from [11]

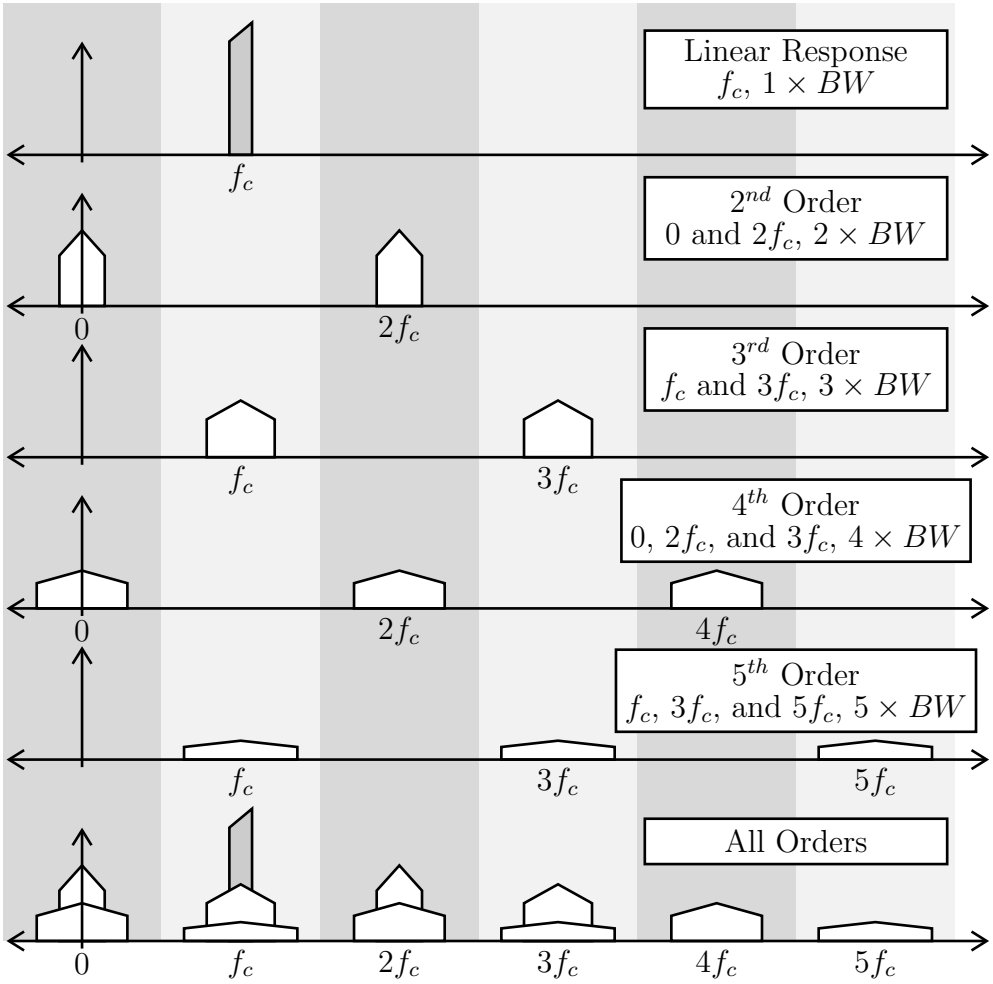


Figure 1.10: Individual harmonics up to the 5^{th} order

where $G(A)$ is the AM/AM transfer characteristic of the MNE [12]. In both Equations (1.10) and (1.11), the PA output $v(t)$ will have frequency components at integer multiples of the input frequency w due to the nonlinear nature of PA's as discussed in the previous section. These additional frequency components are considered by $f[\cdot]$ and $G(A)$ respectively. MNE's however do not exist in practical devices because they would have to respond instantaneously and experience no phase shift in the output.

EMNE's on the other hand experience both AM-AM and AM-PM transfer characteristics. Thus, $v(t)$ in Equation (1.11) would take the form

$$v(t) = G(A)cos(\omega t + \phi(t) + F(A)) \tag{1.12}$$

where $F(A)$ is the AM-PM transfer characteristic which represents an envelope-dependent phase shift in the output signal $v(t)$ [12].

A nonlinear element with a bandwidth much wider than the signal passing through it can be treated as an EMNE. However, nonlinear elements with bandwidths comparable to that of the signal such as the solid state power amplifiers (SSPA's) used in DOCSIS must be treated as nonlinear elements with memory [12]. The effect of the memory on the power amplifier output is to make the AM-AM and AM-PM conversions become frequency dependent. In this case, Equation (1.12) becomes

$$v(t) = G(A, \omega)cos(\omega t + \phi(t) + F(A, \omega)) \tag{1.13}$$

where $G(A, \omega)$ and $F(A, \omega)$ are the frequency dependent AM-AM and AM-PM conversions respectively. The frequency dependent effect on the power amplifier output results in a non-flat gain across the passband frequencies and a nonlinear phase response.

1.6 Power Amplifier Operation

The harmonics in the output of the PA act as noise in the signal. Noise which falls outside of the US frequency band, or out-of-band (OOB) noise can be filtered out and is therefore not a concern but the in-band noise is an issue. One way this issue can be resolved is by operating the amplifier strictly within the linear region defined as up to the 1 dB compression point. However, this method results in extremely low power efficiencies in the PA [10]. If the PA is to operate entirely in its linear region, then taking into account the very high peak to average power ratio of the OFDM signals (typically 15 dB) the amplifiers would need to operate on average 15 dB below the start of compression. This mode of operation is referred to as input back-off (IBO) and it ensures that no significant signal compression occurs even at the peaks of the signal.

Some noise, however, is acceptable. The DOCSIS standard requires the noise level to be about 40 dB below the signal level at the PA output to enable high data rates. Therefore the input only has to be backed off to the point where the noise reaches this level. The power amplifier operating point is illustrated in Figure 1.11.

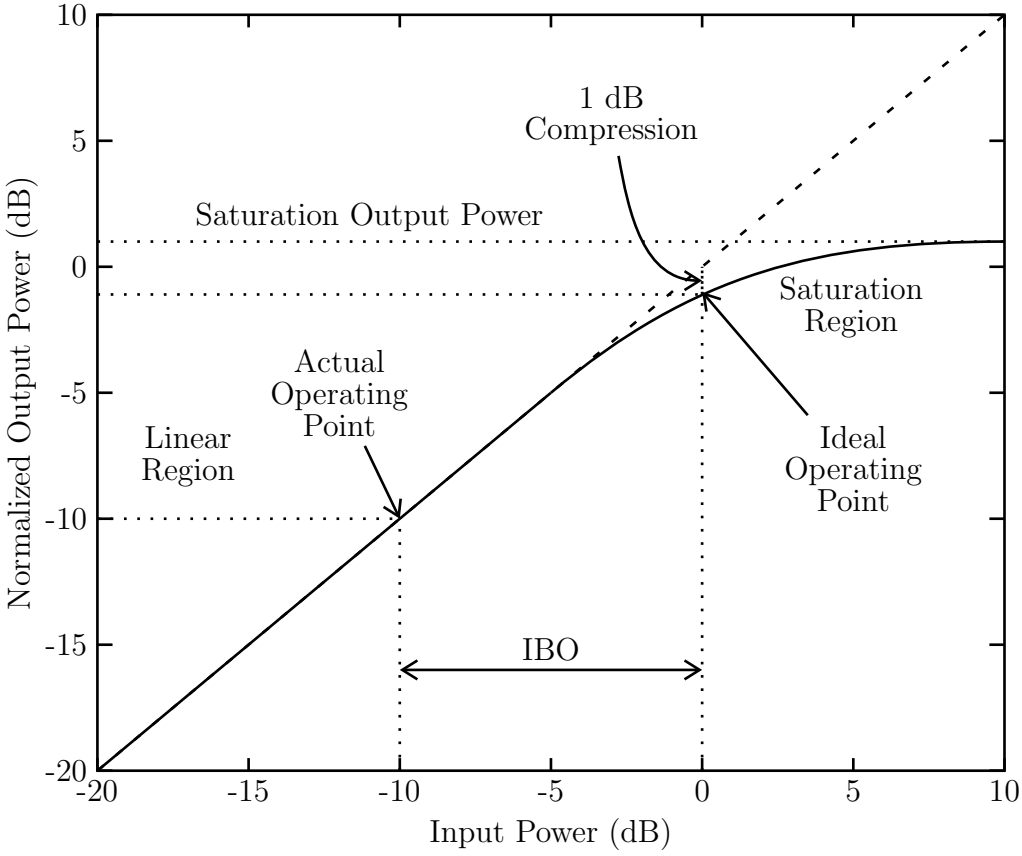


Figure 1.11: Non-linear power amplifier distortion

1.7 Problem Definition

In the full-duplex node, the downstream signal from the output of the power amplifier leaks into the upstream signal path through both the coupler and reflections in the cable network. This self-interference acts like noise in the upstream signal because the US and DS signals occupy the same frequencies in FDX, and there is no longer a diplexer to completely isolate the signals. Without cancellation, this SI noise makes it impossible to demodulate high data rate signals, which are highly susceptible to noise. To eliminate this self-interference, both the signal at the output of the power amplifier and the characteristics of the channel that this signal takes to get to the receiver must be very accurately known.

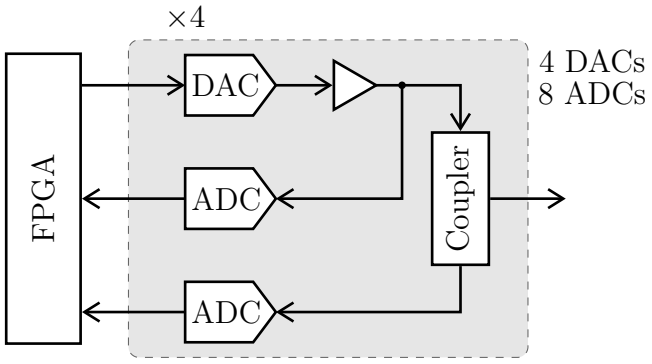
The purpose of this thesis is to model the power amplifier to predict its output for an ultra-wideband input signal such as that present in a DOCSIS node. Such a model is necessary

to facilitate self-interference cancellation, thereby enabling FDX DOCSIS operation. To predict the PA output, a nonlinear model must be used since the PA is a nonlinear device. Estimating the SI channel is out of the scope of this thesis.

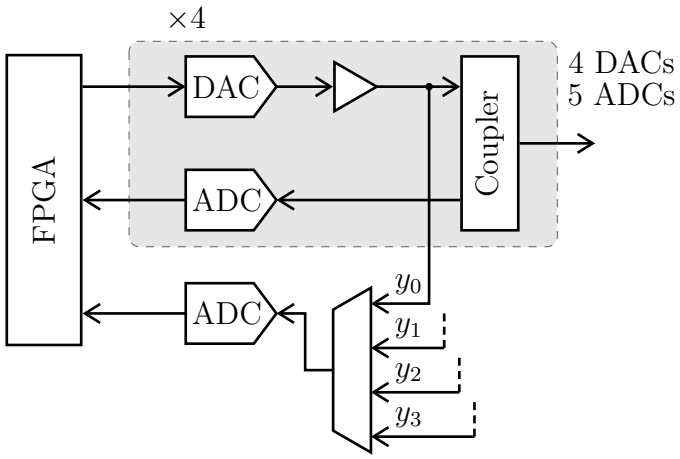
Perhaps an obvious and less complicated alternative EC structure compared to modeling the PA output is to directly measure the PA output with an ADC and subtract this measured signal from the received US signal to cancel the self-interference. Doing so would eliminate the need for a nonlinear model. Although such a method would work, the trade-off is one extra ADC per amplifier is required. A typical DOCSIS node has four amplifiers, meaning four additional ADC's along with the ADC's for each upstream signal would be needed. Including additional ADC's into the node increases its cost and power consumption, both of which could potentially make the node undesirable to customers. There is potential to reduce cost and power consumption by instead modeling the SI signal to perform EC. These two potential node designs are shown in Figure 1.12.

The level of accuracy needed in predicting the PA output stems from the SNR required in the received upstream signal in the node. To reliably receive the high data rate signals used in DOCSIS 3.1, an SNR of at least 40 dB is required [13]. This level of SNR requires the self-interference signal from the coupler to be attenuated by 60 dB. The reasoning for this is illustrated in Figure 1.13 and discussed in the following paragraphs.

From the DOCSIS 3.1 Full-Duplex standard, the downstream signal transmitted from 108 to 1218 MHz has a linear tilt in the frequency domain of 21 dB to compensate for the frequency response of the coaxial cable. At 108 MHz the DS signal will have a PSD of 37 dBmV/6MHz, and at 1218 MHz it will have a PSD of 58 dBmV/6MHz. The coupler provides about 30 dB isolation between the US and DS signal, leaving 7 dBmV/6MHz PSD at 108 MHz and 28 dBmV/6MHz at 1218 MHz at the US receiver input. This signal is shown as 'PA leakage' in Figure 1.13. The PA generates nonlinearities at a power level 40 dB below the DS output signal, shown as 'PA noise' in Figure 1.13. This noise signal also experiences leakage through the coupler, and arrives at the US receiver input with a PSD of -33 dBmV/6MHz at 108 MHz and -12 dBmV/6MHz at 1218 MHz, shown as 'PA noise leakage' in Figure 1.13.



(a) Noise sampling



(b) Noise prediction

Figure 1.12: Noise cancellation options

In order to achieve a 40 dB SNR at the US receiver, the 'PA leakage' and 'PA noise leakage' signals must be canceled at least to the level indicated as 'Receiver noise floor for 40 dB SNR' in Figure 1.13. For an expected US input level of 3 dBmV/6MHz, this floor would be at -37 dBmV/6MHz. However, we will specify the required noise floor to be -40 dBmV/6MHz to leave 3dB of headroom for additional sources of noise.

Note that we only need to cancel the portion of the signal which falls below the upper edge of the US signal, as the remainder can be filtered away. The upper edge of the US signal is at 684 MHz, which corresponds to a maximum SI PSD of about 18 dBmV/6MHz. For this self-interference signal at the upper edge to be attenuated down to the -40 dBmv/6MHz

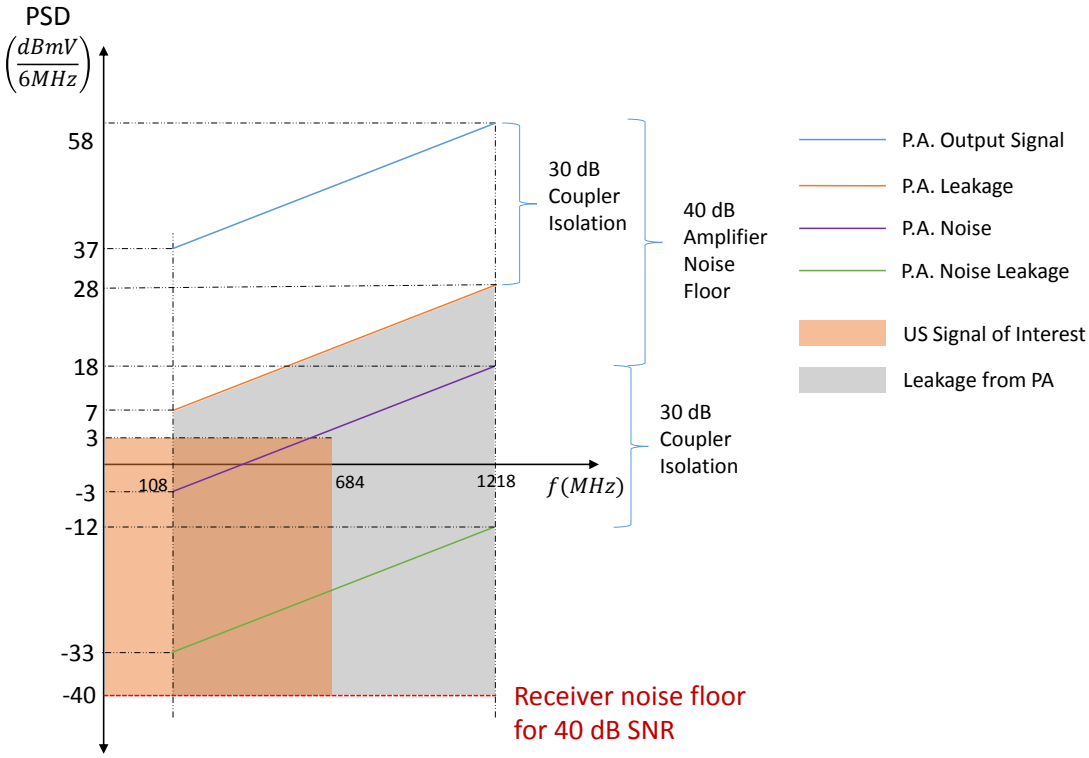


Figure 1.13: Full-duplex signal power levels

noise floor in the receiver, it will have to be canceled by approximately 60 dB.

1.8 Thesis Outline

This thesis is organized as follows. Chapter 2 will discuss previous work in amplifier modeling. Various amplifier models will be considered and the most promising models will be looked at in greater detail. The limiting factors in these models will be discussed which will demonstrate the need for a new model for the case of an ultra-wideband signal.

Chapter 3 discusses the theory of ultra-wideband power amplifier harmonic distortion. A new model will be developed for the ultra-wideband case and the challenges of implementing the model will be addressed. Chapter 4 will explain how to implement the proposed model from Chapter 3 and will investigate practical limitations. Methods for reducing the sampling rate and hardware resource requirements will also be discussed. Chapter 5 will show simulation results and evaluate the performance of the newly developed model. Finally, Chapter

6 will conclude the thesis by summarizing the main developments.

2. Previous Work in Amplifier Modeling

Power amplifier modeling has been widely researched for the use of digital predistortion (DPD) and echo cancellation (EC). Research into DPD techniques accelerated because conventional power amplifiers in most radio-frequency (RF) applications must be backed off considerably from their peak operating power level to control the in and out of band harmonic noise caused by the nonlinear effects. This large back-off typically means that expensive amplifiers with very high maximum power ratings must be used to transmit adequate levels of output power. A more desirable alternative is to utilize digital predistortion ahead of the amplifier to compensate for the nonlinear effects, therefore allowing it to operate closer to its maximum power output while maintaining lower power in the harmonics.

The purpose of echo cancellation is to remove self-interference from the received signal in a full-duplex transceiver. Echo cancellation is functionally similar to DPD in that it attempts to model the output of the power amplifier to cancel the unwanted signal in the receiver. A common application of EC is speakerphone applications where you can talk and listen at the same time on the same frequency (the frequency of your voice), without feedback occurring. More recently, EC has become increasingly popular in research towards FDX in cellular communication, such as in [14–18] to name a few recent publications.

To explain the operation of digital predistortion and echo cancellation, some background on complex signals and transmission/reception in a transceiver must first be reviewed.

2.1 Complex Signals

Complex signals can occupy positive and negative frequencies independently, but real signals cannot independently occupy positive and negative frequencies. Instead, the distri-

bution of power across the frequency spectrum for real signals is symmetric about 0 Hz. The reasoning for this fact can be shown using a Fourier transform, which converts a time domain signal to a frequency domain signal. First consider a real signal $\cos(2\pi Ft)$ where F is the frequency in Hz. From the well-known Euler formula,

$$\cos(2\pi Ft) = \frac{1}{2}e^{\jmath 2\pi Ft} + \frac{1}{2}e^{-\jmath 2\pi Ft} \tag{2.1}$$

the Fourier transform of $\cos(2\pi Ft)$ can be calculated,

$$\mathcal{F}\left\{\cos(2\pi Ft)\right\} = \frac{1}{2}\left[\delta(f-F) + \delta(f+F)\right] \tag{2.2}$$

where $\delta(\cdot)$ is the Dirac delta function [19]. This result means that $\cos(2\pi Ft)$ has frequency components at F and $-F$, both with amplitudes of $1/2$. Next, consider the complex signal $\cos(2\pi Ft) + \jmath\sin(2\pi Ft)$. From the Euler formula,

$$\cos(2\pi Ft) + \jmath\sin(2\pi Ft) = e^{\jmath 2\pi Ft} \tag{2.3}$$

the Fourier transform can then be calculated,

$$\mathcal{F}\left\{\cos(2\pi Ft) + \jmath\sin(2\pi Ft)\right\} = \delta(f-F) \tag{2.4}$$

which has an amplitude of 1 at only F [19]. This example shows how a complex signal can define a positive frequency independently from the negative frequency, as opposed to a real signal, which has symmetric amplitudes about 0 Hz.

2.2 Complex Signal Transmission

In most modern systems that communicate over an analog medium such as cable and cellular, the signals are processed digitally at baseband (centered around 0 Hz) as complex numbers and are transmitted/received in the analog domain as purely real signals at passband (centered around a carrier frequency). The complex baseband transmit (Tx) signal ($\widetilde{x}(n)$) is up-converted to passband through multiplication with a complex sinusoid at the carrier frequency, resulting in a real complex passband signal. Since only real signals can be transmitted over a physical transmission medium, the imaginary portion of the complex passband signal is discarded, yielding a real passband signal. The functional diagram of the

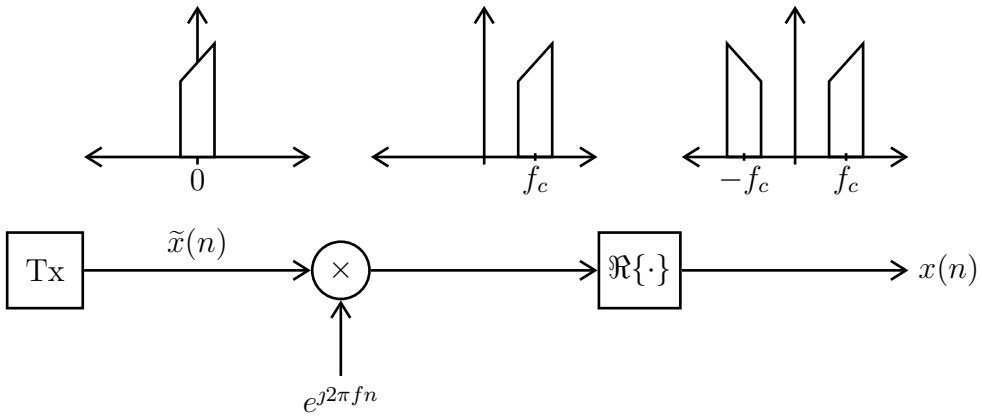


Figure 2.1: Up-conversion using complex multiplication

up-conversion is shown in Figure 2.1. As shown in the diagram, discarding the imaginary portion of the signal creates a frequency spectrum that is symmetric about 0 Hz.

A complex digital baseband signal $\widetilde{x}(n)$ is up-converted to the real passband signal $x(n)$ by

$$x(n) = \Re\left\{\widetilde{x}(n)e^{\jmath 2\pi fn}\right\} \tag{2.5}$$

where f is the passband (carrier) frequency in cycles/sample, $f = F/F_s$ where F is the passband frequency in Hz and F_s is the sampling rate in samples/second, n is the sample number (unit samples), and $\Re\{\cdot\}$ is an operator which returns the real portion of its argument. Equation (2.5) can be rewritten as

$$\begin{aligned} x(n) &= \Re\left\{(\widetilde{x}_I(n) + \jmath\widetilde{x}_Q(n))(\cos(2\pi fn) + \jmath\sin(2\pi fn))\right\} \\ &= \Re\{\widetilde{x}_I(n)\cos(2\pi fn) - \widetilde{x}_Q(n)\sin(2\pi fn) \\ &+ \jmath(\widetilde{x}_I(n)\sin(2\pi fn) + \widetilde{x}_Q(n)\cos(2\pi fn))\} \\ &= \widetilde{x}_I(n)\cos(2\pi fn) - \widetilde{x}_Q(n)\sin(2\pi fn) \end{aligned} \tag{2.6}$$

where $\widetilde{x}_I(n)$ is the inphase (real) part of the complex baseband signal and $\widetilde{x}_Q(n)$ is the quadrature (imaginary) part of the complex baseband signal. The result of Equation (2.6) is commonly used in hardware because taking only the real result of the complex multiplication requires only two multiplies, whereas the full result of the complex multiplication requires four multiplies.

The hardware equivalent of Equation (2.6) is illustrated in Figure 2.2. In lower frequency applications, this up-conversion can happen digitally if the sampling rate exceeds the Nyquist rate, which is twice the highest frequency present in the signal after the up-conversion. The sinusoids used to generate the carrier frequencies are digitally generated by a numerically controlled oscillator (NCO) which stores the sinusoids in random access memory (RAM).

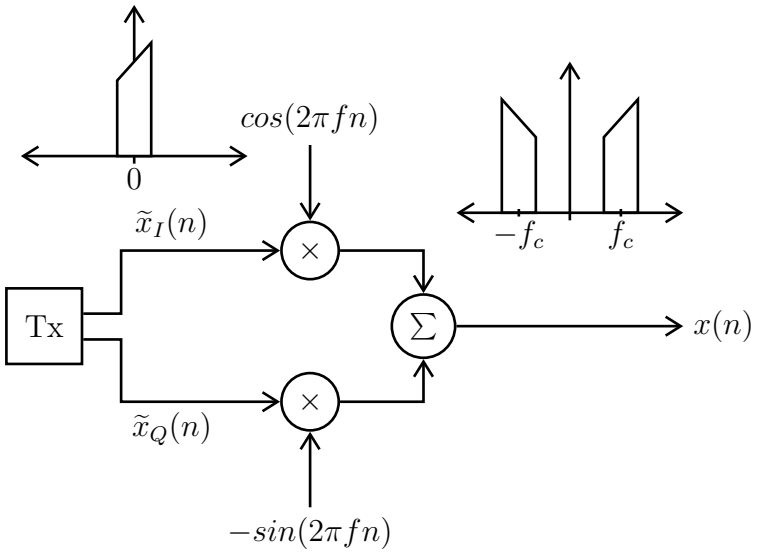


Figure 2.2: Up-conversion using quadrature multiplication

In high-frequency applications where the digital sampling rate cannot meet the Nyquist rate for the passband signal, the complex baseband signals are first converted to analog then multiplied by the output of a voltage controlled oscillator (VCO) and added together to be up-converted to real passband. In the analog case, the n and f in Equation (2.6) would be replaced by t and F respectively, where t is time in seconds and F is the carrier frequency in Hz.

In the receiver, the signal is converted back down to complex baseband in a very similar fashion. The received real passband signal $y(n)$ has the form

$$y(n) = A\widetilde{x}_I(n)\cos(2\pi fn + \theta) - A\widetilde{x}_Q(n)\sin(2\pi fn + \theta) \tag{2.7}$$

where A is an arbitrary amplitude and θ is an arbitrary phase shift dependent on the length and properties of the transmission medium. $y(n)$ is then converted to complex baseband $\widetilde{y}(n)$

through multiplication with a complex exponential with a frequency equal to the negative of the carrier frequency, resulting in

$$\begin{aligned}\widetilde{y}(n) &= (A\widetilde{x}_I(n)\cos(2\pi fn+\theta) - A\widetilde{x}_Q(n)\sin(2\pi fn+\theta))e^{-\jmath 2\pi fn} \\ &= A\widetilde{x}_I(n)\cos(2\pi fn+\theta)\cos(2\pi fn) \\ &- A\widetilde{X}_Q(n)\sin(2\pi fn+\theta)\cos(2\pi fn) \\ &+ \jmath\left[-A\widetilde{x}_I(n)\cos(2\pi fn+\theta)\sin(2\pi fn)\right. \\ &+ \left. A\widetilde{x}_Q(n)\sin(2\pi fn+\theta)\sin(2\pi fn)\right]. \end{aligned} \tag{2.8}$$

Using the following trigonometric identities,

$$\cos(x)\cos(y) = \frac{1}{2}(\cos(x-y)+\cos(x+y)) \tag{2.9}$$

$$\cos(x)\sin(y) = \frac{1}{2}(\sin(x+y)-\sin(x-y)) \tag{2.10}$$

$$\sin(x)\sin(y) = \frac{1}{2}(\cos(x-y)-\cos(x+y)) \tag{2.11}$$

Equation (2.8) becomes

$$\begin{aligned}\widetilde{y}(n) &= A\widetilde{x}_I(n)\frac{1}{2}(\cos(\theta)+\cos(4\pi fn+\theta)) \\ &- A\widetilde{x}_Q(n)\frac{1}{2}(-\sin(\theta)+\sin(4\pi fn+\theta)) \\ &+ \jmath\left[-A\widetilde{x}_I(n)\frac{1}{2}(-\sin(\theta)+\sin(4\pi fn+\theta))\right. \\ &\left. +A\widetilde{x}_Q(n)\frac{1}{2}(\cos(\theta)-\cos(4\pi fn+\theta))\right] \end{aligned} \tag{2.12}$$

The double frequency components at $4\pi f$ in Equation (2.12) can be easily removed using a LPF. In addition, the receiver must have circuitry to correct for the arbitrary phase offset θ, making it equal to zero. Assuming $\theta = 0$ and the double frequency components are filtered away, $\widetilde{y}(n)$ becomes

$$\widetilde{y}(n) = \frac{A}{2}\widetilde{x}_I(n) + \frac{A}{2}\jmath\widetilde{x}_Q(n) \tag{2.13}$$

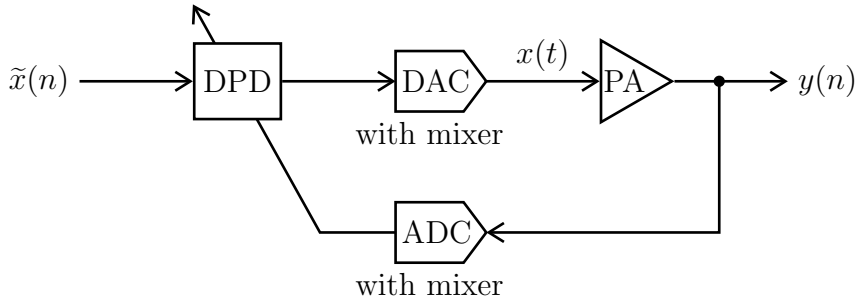
which is equal to the original baseband Tx signal with a scaled magnitude.

2.3 Digital Predistortion

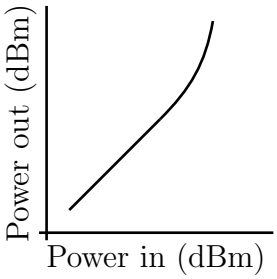
The idea behind a digital predistorter is quite straightforward. The DPD circuit attempts to model the inverse of the power amplifier's nonlinear response, such that the cascade of the

DPD circuit and the power amplifier will result in a linear gain. The operation of a DPD-amplifier cascaded system is illustrated in Figure 2.3. Figure 2.3a contains a block diagram indicating where the DPD fits into the feedback system. Note that many modern DACs and ADCs contain circuitry to perform the up-conversion and down-conversion discussed in section 2.2. We assume that the data converters being used include this functionality.

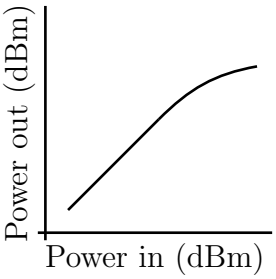
The DPD operates on the baseband signal $\widetilde{x}(n)$ to create the inverse PA response for the passband signal $x(t)$ as shown in Figure 2.3b. Figure 2.3c shows the AM-AM curve of the power amplifier, and finally Figure 2.3d shows the AM-AM curve of the cascade of the DPD and PA. As this cascaded result becomes linear, the power in the harmonic noise will decrease.



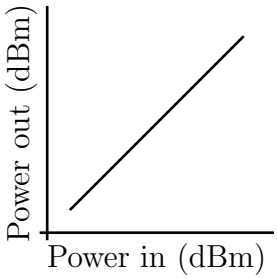
(a) DPD block diagram



(b) DPD AM-AM



(c) PA AM-AM



(d) DPD-PA AM-AM

Figure 2.3: Digital predistortion operation

Usually, a DPD device will model the PA using the baseband transmitted and received signals. An ADC can be used to tap into the output of the power amplifier to sample and down-convert the signal to complex baseband for use by the DPD algorithms. Using the

baseband equivalents of the input and output to the power amplifier, the DPD will construct a nonlinear model of the PA, then calculate the inverse of the model and apply the inverse characteristics to the Tx signal. A parameter estimation technique such as least squares is used to estimate the parameters of the nonlinear model [20–22].

Applying DPD at complex baseband as opposed to real passband has many benefits. The main reason the modeling and predistortion are done at baseband is sampling rate. In high-frequency applications such as 4G (cellular), the carrier frequency can be up to nearly 6 GHz [23] which would require a sampling rate of at least 12 GHz to capture all of the signal without aliasing, which is beyond the capabilities of current technology. Also, when a baseband signal is raised to a higher order, the resulting signal contains only the in-band harmonics. In contrast, a passband signal which is raised to a higher order will generate both in-band and out-of-band harmonics. The distinction between in-band and out-of-band harmonics will be discussed in more detail later in this chapter. Since out-of-band harmonics are often filtered out and not reflected in the final signal, a baseband representation may be preferable. The downside to processing at complex baseband is that one complex multiplication at passband requires four multipliers whereas one real multiplication requires only one multiplier.

2.4 Echo Cancellation

Echo cancellation, like DPD, involves modeling the power amplifier. However, it does not compute the inverse of the PA model. Instead, it directly calculates an estimate of the output of the PA and uses this estimate to cancel the signals reflected back to the receiver (echos). Like a DPD, EC usually takes place at complex baseband because of sampling rates and the ability to calculate only in-band harmonics.

An echo canceler uses the complex baseband equivalents of the input and output of the PA to construct a nonlinear model. Using these signals along with a parameter estimation technique such as least squares, the PA can be modeled, and its output can be predicted. An ADC can be used to tap into the output of the PA to sample it and down-convert it to baseband for use during the model parameter estimation process.

For EC to be successful, along with the output of the PA, the path that the self-interference signal takes from the PA to the receiver must be known. This path is referred to as the SI channel. The SI channel can be represented with a linear model since the channel typically has a linear response. The number of samples needed to model the linear SI channel will depend on how long it takes for the last echo to return to the receiver, as well as the sampling rate. For example, if the sampling rate is 100 MHz and if the last echo returns in 3 μs, the SI channel model must include at least $(100 \times 10^6 \text{ samples/sec}) \times (3 \times 10^{-6} \text{ sec}) = 300$ samples.

A block diagram of an echo canceler is shown in Figure 2.4. In the figure, $\widetilde{x}(n)$ is the transmitted (Tx) baseband signal, $\widetilde{y}(n)$ is the received (Rx) baseband signal, and $\widetilde{d}(n)$ is the self-interference signal down-converted to baseband. The SI, as mentioned previously, includes the leakage of the PA output through the coupler as well as echos returning from the channel. As you can see in Figure 2.4, one ADC is used to receive the signal from the channel (transmitted from the cable modems in the case of DOCSIS), and another ADC is used to tap into the output of the power amplifier. The digital signals at the output of both ADCs feedback into the echo canceling circuit to update the power amplifier and channel model. Once the echo canceler has converged, it will calculate an estimate of the SI signal $(\widehat{\widetilde{d}}\,(n))$ which would ideally be equal to the actual SI signal $(\widetilde{d}(n))$. The estimated SI signal is then used to cancel the actual SI signal, ideally leaving only the desired received signal $\widetilde{y}(n)$ to be demodulated.

2.5 Power Amplifier Models

Previous work in modeling amplifiers falls into two main categories: frequency independent and frequency dependent models. Frequency independent models assume the gain of the amplifier is constant across the entire passband and there are no memory effects. Frequency-dependent models, on the other hand, compensate for non-flat passband gains as well as memory effects. It was noted in Chapter 1 that power amplifiers with wideband signal inputs often experience frequency dependent AM-AM and AM-PM effects, implying their models must include memory samples. Frequency independent models will not be discussed

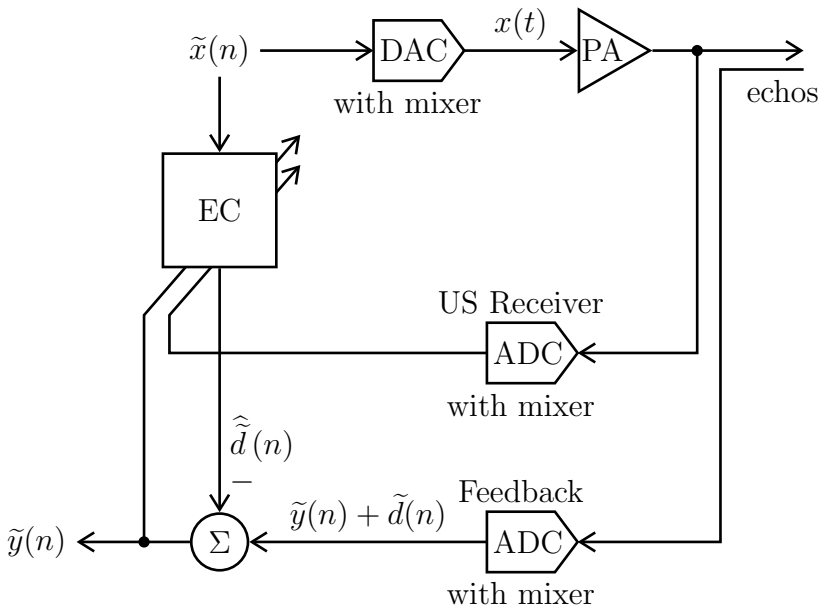


Figure 2.4: Block diagram of an echo canceler

here since they do not apply to the wide-band signals seen in DOCSIS.

Frequency-dependent models are widely used for power amplifier modeling, usually at complex baseband for the use of pre-distortion and echo cancellation. For notational purposes, a signal name with a tilde, i.e., $\widetilde{x}$, will represent a complex baseband signal and a signal name with no tilde, i.e., x, will represent a real passband signal. The most prominent models from the literature are discussed in the following sections.

2.5.1 Volterra Series

The Volterra series is a nonlinear behavioral model developed by Vito Volterra in the late 1800's. The Volterra series is similar to the Taylor series except it also includes memory effects. The Taylor series can be used to approximate the output of a nonlinear system if the output depends only on the input at that given time, whereas the Volterra series output depends on the current input and all past inputs.

The Volterra series is defined by Equation (2.14). For an input $x(n)$ the output $y(n)$ is

given by [24]:

$$\begin{aligned} y(n) = h_0 & \\ & + \sum_{m_1=0}^{\infty} h_1(m_1)x(n-m_1) \\ & + \sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty} h_2(m_1,m_2)x(n-m_1)x(n-m_2) \\ & + \sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty}\cdots\sum_{m_p=0}^{\infty} h_p(m_1,m_2,...,m_p)x(n-m_1)x(n-m_2)\cdots x(n-m_p) \\ & + \cdots \end{aligned} \tag{2.14}$$

where $h_p(m_1, m_2, ..., m_p)$ is known as the p-th order Volterra kernel. The zeroth-order kernel h_0 is a constant, the first-order kernel h_1 is a linear filter, and the rest of the kernels are higher order convolutions. Without any loss of generality, one can assume that the Volterra kernels are symmetric about the diagonal. For example, $h_2(a, b) = h_2(b, a)$ since $x(n-a)x(n-b) = x(n-b)x(n-a)$.

The Volterra series expansion can be written more compactly by defining the pth-order Volterra operator $h_p(x(n))$ as

$$h_p(x(n)) = \sum_{m_1=0}^{M-1}\sum_{m_2=0}^{M-1}\cdots\sum_{m_p=0}^{M-1} h_p(m_1,m_2,...,m_p)x(n-m_1)x(n-m_2)\ldots x(n-m_p). \tag{2.15}$$

Substituting Equation (2.15) into Equation (2.14) yields

$$y(n) = h_0 + \sum_{p=1}^{\infty} h_p(x(n)). \tag{2.16}$$

The structure for the Volterra series is illustrated in Figure 2.5.

Since the infinite series given in Equation (2.14) cannot be practically implemented, one

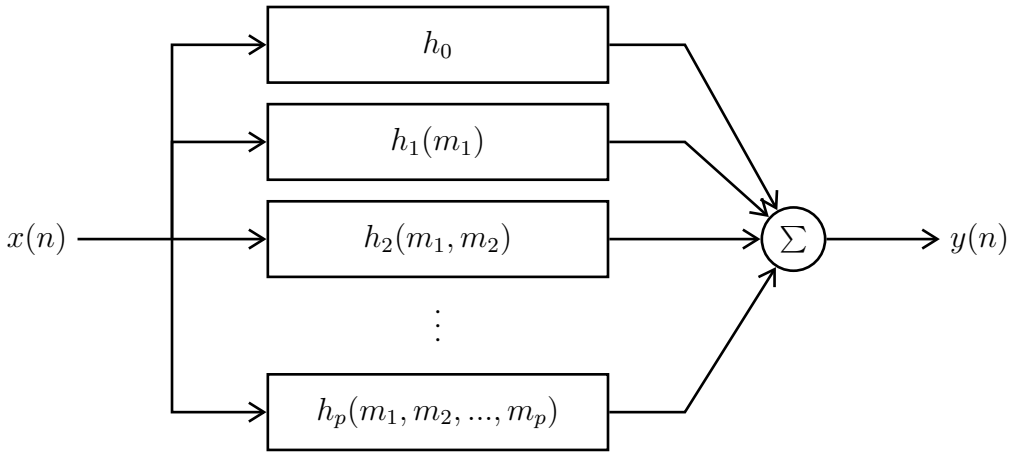


Figure 2.5: Block diagram of Volterra series

must work with the truncated form [24]

$$\begin{aligned} y(n) &= h_0 + \sum_{p=1}^{P} h_p(x(n)) \\ &= h_0 \\ &+ \sum_{m_1=0}^{M-1} h_1(m_1)x(n-m_1) \\ &+ \sum_{m_1=0}^{M-1}\sum_{m_2=0}^{M-1} h_2(m_1,m_2)x(n-m_1)x(n-m_2) \\ &+ \sum_{m_1=0}^{M-1}\sum_{m_2=0}^{M-1}\cdots\sum_{m_p=0}^{M-1} h_p(m_1,m_2,...,m_p)x(n-m_1)x(n-m_2)\ldots x(n-m_p) \end{aligned} \tag{2.17}$$

where P is the highest order of non-linearity of the Volterra series expansion, M is referred to as its length, and $M-1$ as its memory. Figure 2.6 illustrates the computation of a Volterra series with $M=2$ and $P=2$. The Volterra series has a computational complexity of $O(M^P)$, i.e., the number of coefficients in this polynomial expansion is proportional to M^P. This means that implementing this system will result in very large filters for even moderate values of M and P. Consequently, various alterations to the Volterra series have been developed to make implementation more practical.

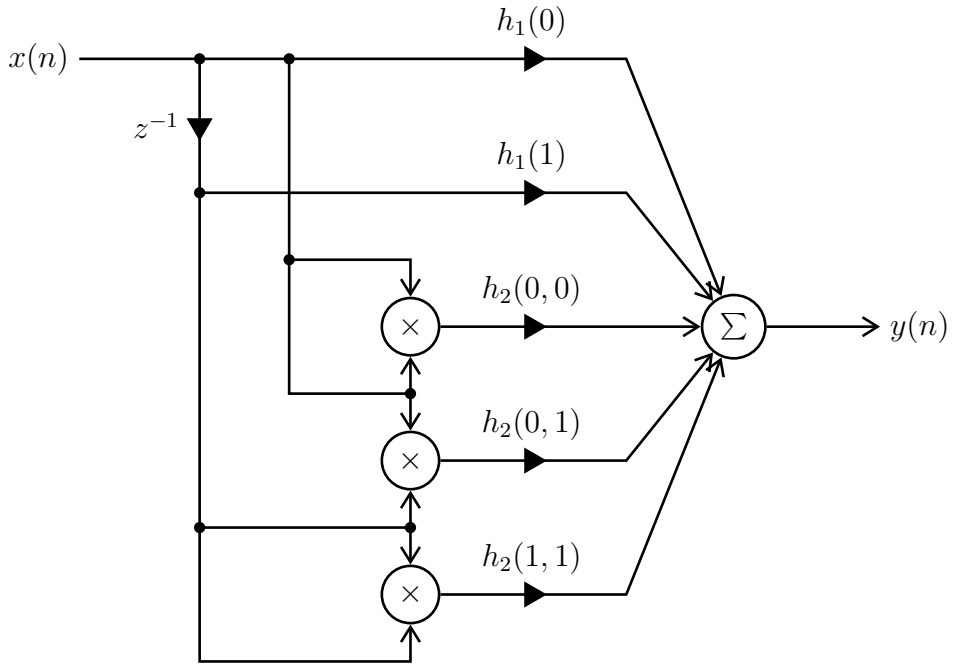


Figure 2.6: Example 2nd-order Volterra series with length 2

2.5.2 Wiener Model

A special case of the truncated Volterra series is the Wiener model. The Wiener model consists of a linear filter (h) followed by a memoryless nonlinearity [24] and is equivalent to taking

$$h_k(m_1, m_2, \cdots, m_k) = a_k h(m_1)h(m_2)\cdots h(m_k), k = 1, 2, ..., K \tag{2.18}$$

in (2.17), where a_k are the complex polynomial coefficients of the nonlinearity. Thus, the Wiener model can be written as [21]

$$y_{\mathrm{W}}(n) = \sum_{k=1}^{K} a_k \left[\sum_{m=0}^{M-1} h(m)x(n-m) \right]^k . \tag{2.19}$$

where M is the memory length and K is the highest order of nonlinearity.

Equation (2.19) results in the input $x(n)$ first being convolved with the linear filter h, followed by a nonlinear calculation where the output of the convolution is raised to the power of k and multiplied by the corresponding coefficient a_k. The initial convolution accounts for memory effects in a nonlinear system. The block diagram of this operation is shown in Figure

2.7.

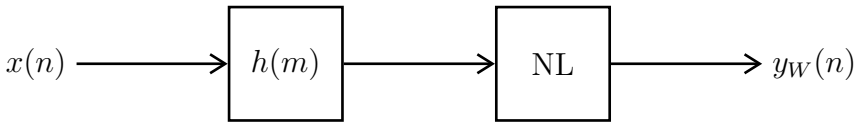


Figure 2.7: Wiener diagram

The Wiener model has been studied for traveling-wave tube (TWT) amplifiers for satellite communication in [25] and [26] and for general nonlinear system identification in [27] and [28]. The Wiener approach is a simple model, but its effectiveness in modeling most PA's is limited [21], and the output in (2.19) depends nonlinearly on the coefficients $h(m)$, making the system identification more difficult and costly.

2.5.3 Hammerstein Model

Another simple nonlinear model with memory is the Hammerstein model. The Hammerstein model is formed by a nonlinearity followed by a linear filter (h), which is the same structure as the Wiener model in reverse order. Thus, for the Hammerstein model we have [21]

$$y_{\mathrm{H}}(n) = \sum_{m=0}^{M-1} h(m) \sum_{k=1}^{K} a_k x^k(n-m). \tag{2.20}$$

This model is also straightforward in that the result is formulated by convolving the output of a memoryless nonlinear system with a linear filter. The model has been studied for general nonlinear system identification in [28–31], as well as for modeling power amplifiers, most notably in [32–34]. The block diagram of the Hammerstein model is shown in Figure 2.8. This structure is more desirable compared to the Wiener model because it is linear in the parameters $h(m)$ and a_k, which makes the parameters easier to estimate.

It should be noted that Wiener and Hammerstein models can cascade into the Wiener-Hammerstein model, which consists of a linear filter followed by a memoryless nonlinearity, followed by another linear filter [21]. Although the resulting filter is more general than either one of the Wiener or Hammerstein models, it is still nonlinearly dependent on the first linear

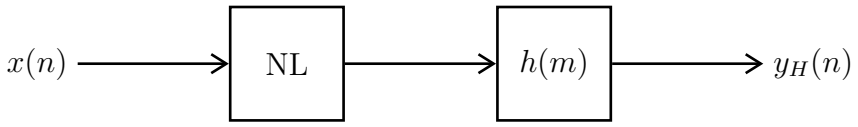


Figure 2.8: Hammerstein diagram

filter in the cascade and hence has only been studied as a general analysis tool [21].

2.5.4 Memory Polynomial Model

The memory polynomial model is a widely researched and commonly used method for modeling power amplifiers for both DPD and EC. It is a modification of the Volterra series that significantly reduces the complexity yet has superior performance compared to the other models discussed. The memory polynomial model is defined by assuming all Volterra kernels to be zero except for the diagonal [22], resulting in the equation

$$y_{MP}(n) = \sum_{k=1}^{K} \sum_{m=0}^{M-1} a_{km} x^k(n-m). \tag{2.21}$$

This structure involves K linear convolutions, each with memory depth M, where the input to each convolution is the current sample $x(n)$ raised to the power of k. This model is similar to the Taylor series of Equation (1.3), with the extension of memory samples for each order k. A diagram of the memory polynomial structure is given in Figure 2.9. Notice that the MPM model can be thought of as a generalized Hammerstein model. If the coefficients for each order of nonlinearity are the same, it is equivalent to a Hammerstein model.

2.6 Baseband Models

The models discussed in Section 2.5 accept passband input signals. However, since DPD and EC are typically performed at baseband, the models can be modified for narrowband applications at baseband. The updated models can be derived by first defining $x(n)$ in terms of its complex baseband equivalent. Let $\widetilde{x}(n) = \widetilde{x}_I(n) + \jmath\widetilde{x}_Q(n)$, and let the passband carrier

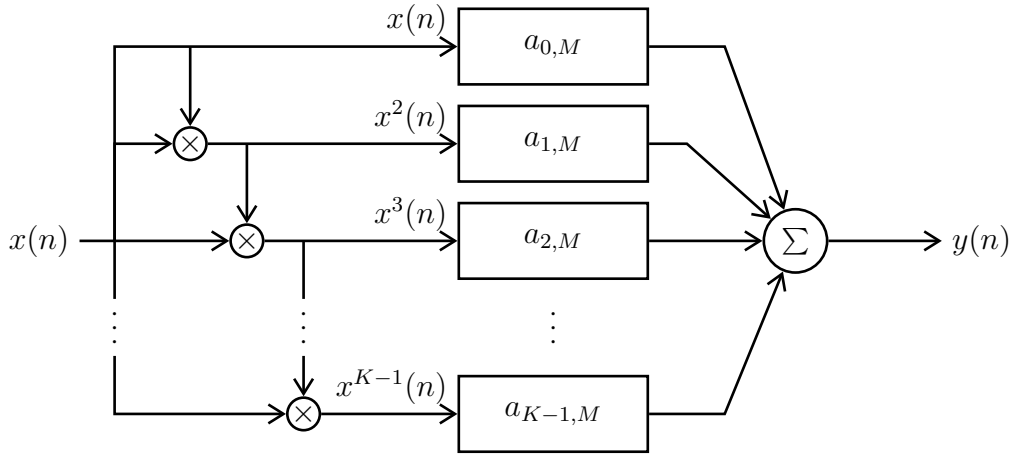


Figure 2.9: Memory polynomial structure

frequency be ω_c.

$$\begin{aligned}
x(n) =& \Re\left\{\widetilde{x}(n)e^{\jmath\omega_c n}\right\} \\
=& \Re\left\{\widetilde{x}(n)(\cos\omega_c n + \jmath\sin\omega_c n)\right\} \\
=& \Re\left\{(\widetilde{x}_I(n) + \jmath\widetilde{x}_Q(n))(\cos\omega_c n + \jmath\sin\omega_c n)\right\} \\
=& \widetilde{x}_I(n)\cos\omega_c n - \widetilde{x}_Q(n)\sin\omega_c n \\
=& \widetilde{x}_I(n)\left(\frac{1}{2}e^{\jmath\omega_c n} + \frac{1}{2}e^{-\jmath\omega_c n}\right) - \widetilde{x}_Q(n)\left(\frac{1}{2\jmath}e^{\jmath\omega_c n} - \frac{1}{2\jmath}e^{-\jmath\omega_c n}\right) \\
=& \frac{1}{2}\left(\widetilde{x}_I(n)\left(e^{\jmath\omega_c n} + e^{-\jmath\omega_c n}\right) + \jmath\widetilde{x}_Q(n)\left(e^{\jmath\omega_c n} - e^{-\jmath\omega_c n}\right)\right) \\
=& \frac{1}{2}\left((\widetilde{x}_I(n) + \jmath\widetilde{x}_Q(n))e^{\jmath\omega_c n} + (\widetilde{x}_I(n) - \jmath\widetilde{x}_Q(n))e^{-\jmath\omega_c n}\right) \\
=& \frac{1}{2}\left(\widetilde{x}(n)e^{\jmath\omega_c n} + \widetilde{x}^*(n)e^{-\jmath\omega_c n}\right)
\end{aligned} \tag{2.22}$$

Where $\widetilde{x}^*(n)$ is the complex conjugate of $\widetilde{x}(n)$.

Using the result from (2.22) we can square $x(n)$ to find the second order passband harmonics in terms of the baseband signal,

$$x^2(n) = \frac{1}{4}\left((\widetilde{x}(n))^2 e^{j2\omega_c n} + 2|\widetilde{x}(n)|^2 + (\widetilde{x}^*(n))^2 e^{-j2\omega_c n}\right) \tag{2.23}$$

which has harmonics at 0 Hz (DC) and at $2\omega_c$, and no harmonics in-band at ω_c. Upon

down-conversion to baseband, the harmonics will have the form

$$x^2(n)e^{-\jmath\omega_c n} = \frac{1}{4}\left((\widetilde{x}(n))^2 e^{j\omega_c n} + 2|\widetilde{x}(n)|^2 e^{-j\omega_c n} + (\widetilde{x}^*(n))^2 e^{-3j\omega_c n}\right) \tag{2.24}$$

which has no frequency components at DC and hence will have no harmonic noise in the signal, assuming a narrowband $x(n)$. The baseband memory-polynomial model therefore generally ignores second order harmonics.

Similarly, we can calculate the third order harmonics,

$$\begin{aligned} x^3(n) =& \frac{1}{8}\left(2|\widetilde{x}(n)|^2\widetilde{x}(n)e^{j\omega_c n} + (\widetilde{x}(n)e^{j\omega_c n})^3 + |\widetilde{x}(n)|^2\widetilde{x}^*(n)e^{-j\omega_c n}\right. \\ &\left. + 2|\widetilde{x}(n)|^2\widetilde{x}^*(n)e^{-j\omega_c n} + |\widetilde{x}(n)|^2\widetilde{x}(n)e^{j\omega_c n} + (\widetilde{x}^*(n)e^{-j\omega_c n})^3\right) \\ =& \frac{1}{8}\left(3|\widetilde{x}(n)|^2\widetilde{x}(n)e^{j\omega_c n} + (\widetilde{x}(n))^3 e^{j3\omega_c n} + 3|\widetilde{x}(n)|^2\widetilde{x}^*(n)e^{-j\omega_c n} + (\widetilde{x}^*(n))^3 e^{-j3\omega_c n}\right). \end{aligned} \tag{2.25}$$

which has harmonic components at both positive and negative ω_c and $3\omega_c$. The third order harmonic down-converted to baseband will then be

$$\begin{aligned} x^3(n)e^{-\jmath\omega_c n} =& \frac{1}{8}\left(3|\widetilde{x}(n)|^2\widetilde{x}(n) + (\widetilde{x}(n))^3 e^{j2\omega_c n}\right. \\ &\left. + 3|\widetilde{x}(n)|^2\widetilde{x}^*(n)e^{-2j\omega_c n} + (\widetilde{x}^*(n))^3 e^{-j4\omega_c n}\right). \end{aligned} \tag{2.26}$$

Since a harmonic does land at the carrier frequency, after down conversion to baseband in (2.26) the term $|\widetilde{x}(n)|^2\widetilde{x}(n)$ will remain as baseband noise and therefore must be calculated by the baseband polynomial model. Following the same pattern for fourth, fifth, and higher order harmonics it can be seen that only odd order harmonics fall in-band, and when down-converted to baseband these harmonics will take the form

$$x^k(n) \rightarrow \widetilde{x}(n)|\widetilde{x}(n)|^{k-1}, \qquad k \text{ odd} \tag{2.27}$$

for odd values of k. We can use the results from Equation (2.27) to adapt the passband nonlinear models to be used at baseband. The Wiener model is excluded from the baseband analysis because of the nonlinear dependence on the $h(m)$ coefficients and is therefore not as simple to adapt to baseband.

2.6.1 Baseband Hammerstein Model

The Hammerstein model given in Equation (2.20) can be adapted to be used for narrowband applications at baseband by substituting in Equation (2.27) and only calculating odd order harmonics. The resulting equation is

$$\widetilde{y}_{\mathrm{H}}(n) = \sum_{m=0}^{M-1} h(m) \sum_{\substack{k=0 \\ k \text{ even}}}^{K-1} a_k \widetilde{x}(n)|\widetilde{x}(n)|^k, \tag{2.28}$$

where $k = 0$ corresponds to the 1^{st} order harmonic, $k = 2$ corresponds to the 3^{rd} order harmonic, and so on.

2.6.2 Baseband Memory Polynomial Model

The Memory polynomial model in Equation (2.21) can be used for narrowband signals at baseband by substituting in Equation (2.27). The resulting equation is

$$\widetilde{y}_{MP}(n) = \sum_{\substack{k=0 \\ k \text{ even}}}^{K-1} \sum_{m=0}^{M-1} a_{km} \widetilde{x}(n-m)|\widetilde{x}(n-m)|^k. \tag{2.29}$$

Although only even values of k (hence odd powers) are commonly used, some researchers have reported mild performance improvements when using a more general model including odd values of k (even powers of $\widetilde{x}$) [35]. Also, it should be noted that it is not necessary to have the same memory depth for all powers of k. A diagram of the baseband memory polynomial model is shown in Figure 2.10.

The baseband memory polynomial model has been researched for digital predistortion, most notably in [20, 21, 36] and echo cancellation in [35]. It has proven to provide a good balance of complexity and performance.

2.7 General Memory Polynomial Model

An even higher performance model derived from the baseband memory polynomial model is the general memory polynomial model, developed in [21]. This model uses kernels in the Volterra series offset from the diagonal in both the positive and negative direction, in addition

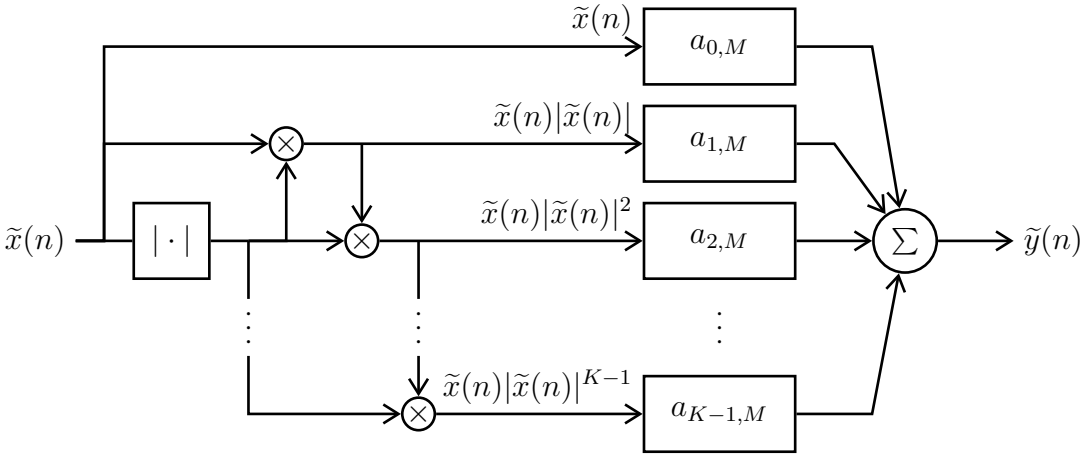


Figure 2.10: Baseband memory polynomial diagram including all values of k

to the diagonal kernels. The result therefore includes the MP model as well as products of the current sample $\widetilde{x}(n)$ with delayed and advanced samples. The resulting equation is

$$\begin{aligned}\widetilde{y}_{\text{GMP}}(n) = &\sum_{k=0}^{K_a-1}\sum_{l=0}^{L_a-1} a_{kl}\widetilde{x}(n-l)|\widetilde{x}(n-l)|^k \\ &+ \sum_{k=0}^{K_b-1}\sum_{l=0}^{L_b-1}\sum_{m=1}^{M_b} b_{klm}\widetilde{x}(n-l)|\widetilde{x}(n-l-m)|^k \\ &+ \sum_{k=0}^{K_c-1}\sum_{l=0}^{L_c-1}\sum_{m=1}^{M_c} c_{klm}\widetilde{x}(n-l)|\widetilde{x}(n-l+m)|^k \end{aligned} \tag{2.30}$$

where K_a, K_b, and K_c are the highest orders of nonlinearities for the aligned envelope (memory polynomial), lagging envelope, and leading envelope respectively. L_a, L_b, and L_c are the memory depths for the aligned, lagging, and leading signals respectively and finally M_b and M_c are the number of cross terms to include in the lagging and leading signals respectively.

Here, K_aL_a is the number of coefficients for the aligned signal and envelope (memory polynomial); $K_bL_bM_b$ is the number of coefficients for signal and lagging envelope; and $K_cL_cM_c$ is the number of coefficients for signal and leading envelope. This cross-term model, like the memory polynomial model, is also linearly dependent on the coefficients and can therefore be estimated using a linear method such as least-squares.

In many cases, not all of the coefficients are required in (2.30). For instance, odd-order nonlinearities usually dominate so one may want to consider only even multiples of k. Also, it is not necessary to have the same memory depth for all powers of k. A diagram of the general memory polynomial is given in Figure 2.11, where z^{-1} is a delay, z^1 is an advance, and MP is the memory polynomial model. In a causal system, future samples cannot be taken. Instead, $\widetilde{x}(n)$ can first be delayed by M_c samples to ensure the most advanced value is not a future sample in time.

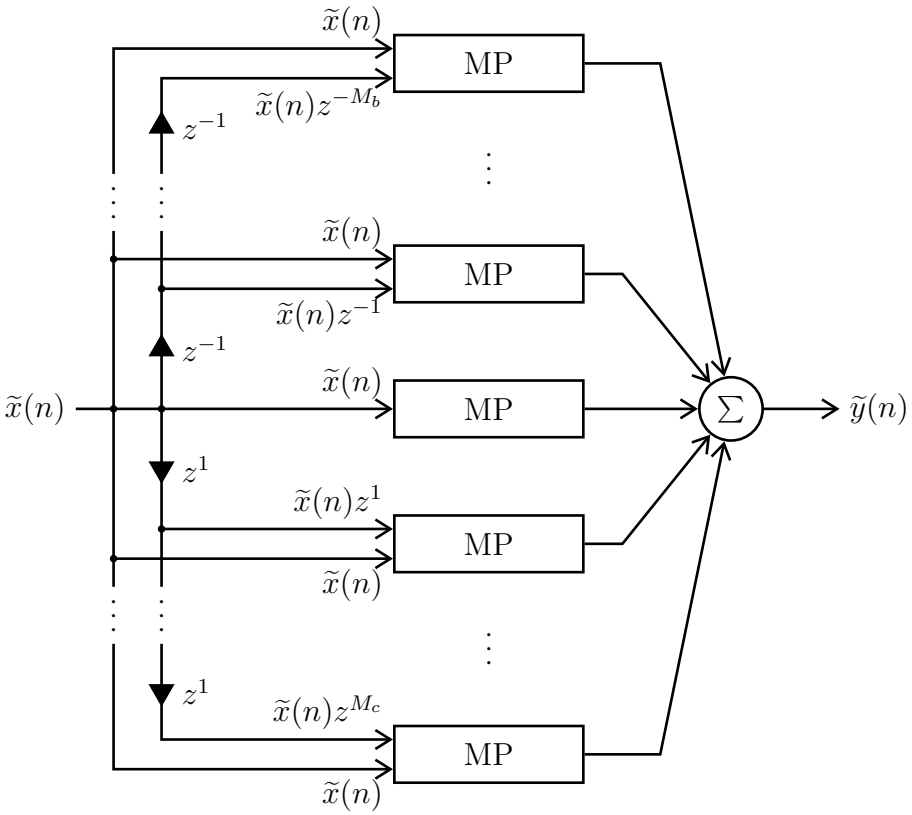


Figure 2.11: Generalized memory polynomial diagram

The general memory polynomial model scales the number of coefficients compared to the memory polynomial model by the total number of cross terms used. Even for a small number of cross terms, the complexity will significantly increase. However, there are applications where performance can be more important than complexity to an extent. The GMP model was first used in [21] as a predistorter for a wireless CDMA signal, and in [37] the GMP

model was found to have the highest accuracy in modeling PA's when complexity is not a concern.

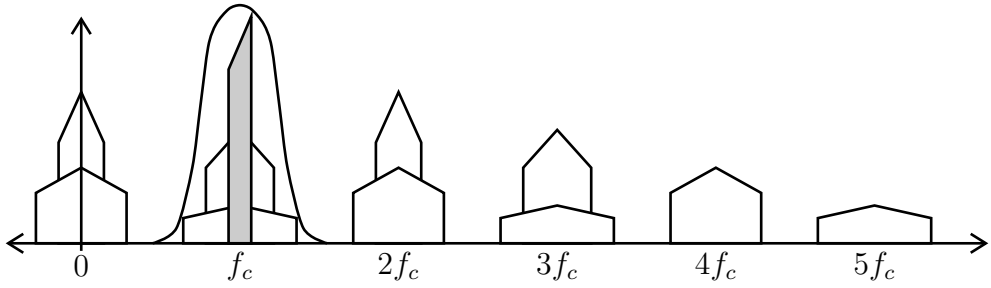
2.8 Limitations of Models

There are implementation issues with the passband models and conceptual problems with the baseband models. First, consider a 250 MHz passband signal which is a low enough frequency that it could be processed digitally at passband. If, for example, the 5th order products had to be calculated at passband, the highest frequency harmonic would extend past 1.25 GHz, requiring a sampling rate of at least 2.5 GHz. Such high sampling rates become very expensive regarding cost, power consumption, and resource usage, if even possible to implement.

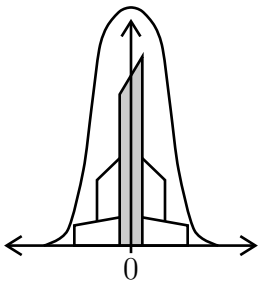
The baseband models on the other hand only work well if the signal bandwidth is narrow enough such that the out-of-band harmonics can be filtered away before down-converting to baseband, such as in Figure 2.12[1]. Figure 2.12a shows how a bandpass filter can isolate the 3^{rd} and 5^{th} order harmonics, so that after a down-conversion to baseband as in 2.12b, a nonlinear model such as the memory polynomial model can accurately calculate the in-band harmonics. This is the method used in many modern wireless systems [20–22, 35, 38].

However, for a wideband signal such as the example illustrated in Figure 2.13, the harmonics with center frequencies at multiples of the carrier frequency overlap with the main signal and will therefore interfere with the main signal when filtered and converted to baseband. These out-of-band harmonics cannot be calculated with the current baseband models and thus cannot be accounted for in a DPD or EC circuit. As discussed in Chapter 3, this is the scenario faced when performing DPD or EC on the wideband output signal from a DOCSIS 3.1 node. Consequently, modifications to the traditional amplifier models are necessary for this application.

[1]Figure derived from [11]



(a) Out-of-band harmonics filtered away



(b) Resulting baseband signal after down conversion.

Figure 2.12: Baseband representation of narrow-band harmonics

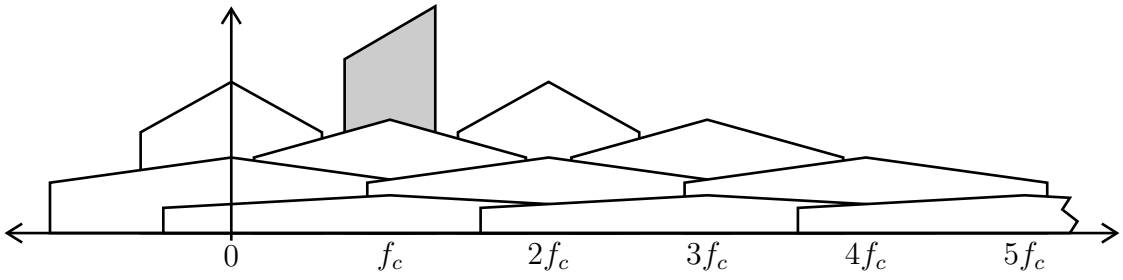


Figure 2.13: Ultra-wideband intermodulation distortion

3. Ultra-Wideband Power Amplifier Modeling

3.1 Introduction

The nonlinear models discussed in Chapter 2 cannot compensate for the in-band distortions from harmonics of ultra-wideband signals, and therefore these models must be adapted. This chapter will develop a new model which is suitable for DOCSIS 3.1 signals. The memory polynomial model is used as the starting point for the development of this new model since it has relatively good performance and reasonable resource usage. Before discussing the new model, sampling theory will be reviewed as it pertains to the challenges in developing this model. Next, the theory of ultra-wideband harmonics and their effects on the baseband processing will be discussed. Then the difficulties with processing UWB signals at baseband will be investigated, and finally, the proposed solution will be developed.

3.2 Digitally Processing an Analog Signal

Information is conveyed over coaxial cable networks through analog voltage signals. However, these signals are typically processed digitally in the transmitter and receiver, requiring conversion between the analog and digital domains. This section discusses analog-to-digital conversion, some essential digital signal processing (DSP) operations, and digital-to-analog conversion.

3.2.1 Analog-to-Digital Conversion

Analog (continuous-time, CT) signals are sampled and stored as discrete-time (DT) digital signals to be digitally processed. If the sampling period is T_s seconds/sample, then the sequence of digital samples $x(n)$ is obtained from the continuous-time signal $x_c(t)$ according

to the relationship

$$x(n) = x_c(nT_s), \qquad -\infty < n < \infty \tag{3.1}$$

where n is an integer representing the sample number. The sampling rate, F_s, is the reciprocal of the sampling period, i.e. $F_s = 1/T_s$ samples/second. The sampling frequency can also be expressed as $\Omega_s = 2\pi F_s$ when we want to use the unit radians/second. Note that DT signals will have the unit samples as its index, such as $x(n \text{ samples})$, and CT signals will have the unit time as its index, such as $x(t \text{ seconds})$, however the unit is not usually written.

It is convenient to represent the sampling process mathematically in two stages [19]. The first stage is an impulse train modulator, and the second is a conversion of the impulse train to a sequence. The periodic impulse train is

$$s(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT_s), \tag{3.2}$$

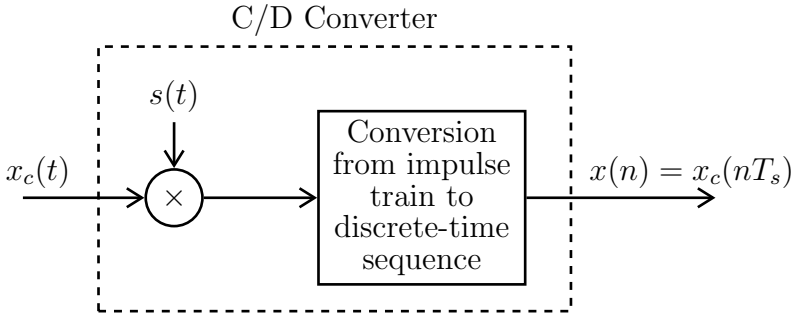
where $\delta(t)$ is the Dirac delta function. The product of $s(t)$ and the CT input $x_c(t)$ is

$$\begin{aligned} x_s(t) =& x_c(t)s(t) \\ =& x_c(t) \sum_{n=-\infty}^{\infty} \delta(t - nT_s) = \sum_{n=-\infty}^{\infty} x_c(t)\delta(t - nT_s), \end{aligned} \tag{3.3}$$

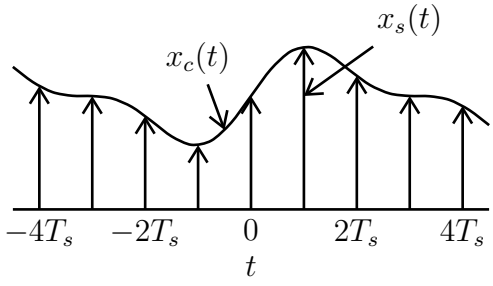
where the area (size) of $x_s(t)$ at time $t = nT_s$ is equal to the value of the CT signal $x_c(t)$ at that time. Since $\delta(t)$ is zero for all t except $t = 0$, $x_s(t)$ can be expressed as

$$x_s(t) = \sum_{n=-\infty}^{\infty} x_c(nT_s)\delta(t - nT_s). \tag{3.4}$$

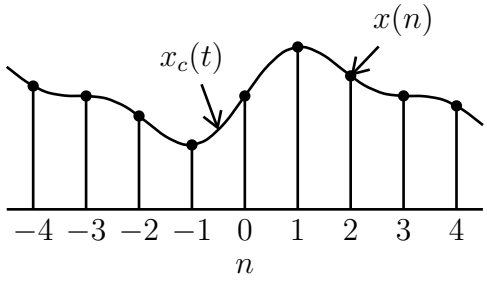
Figure 3.1 shows a CT signal $x_c(t)$ being sampled with an impulse train then converted to the DT signal $x(n)$. The essential difference between $x_s(t)$ and $x(n)$ is that $x_s(t)$ is a continuous time signal that is zero except at integer multiples of T_s where the area of the impulse is equal to $x_c(t)$ at integer multiples of T_s. $x(n)$ on the other hand is a digital signal whose value at sample n is equal to the value of the CT signal at time nT_s, and is indexed on the integer variable n, which, in effect, introduces a time normalization. This means that the sequence of numbers $x(n)$ contains no information about the sampling period. Note that the value of $x(n)$ is undefined for non-integer values of n.



(a) C/D converter



(b) Sampled signal



(c) Discrete-time signal

Figure 3.1: Continuous-to-discrete conversion model

It is essential to note that Figure 3.1 is strictly a mathematical representation of the sampling process and does not reflect how an actual analog-to-digital converter operates. The model is useful however to derive some important insights about sampled signals.

3.2.2 Frequency Domain Representation of a Sampled Signal

The frequency domain representation of the discrete signal $x(n)$ is important for understanding sampling rate requirements, up/down sampling, and digital-to-analog conversion. To derive the frequency domain relation between the input and output of an ideal analog-to-digital converter, consider the Fourier transform of $x_s(t)$. From (3.3), $x_s(t)$ is the product of $x_c(t)$ and $s(t)$, and therefore the Fourier transform of $x_s(t)$ is the convolution of the Fourier transforms of $x_c(t)$ and $s(t)$ scaled by $\frac{1}{2\pi}$. The Fourier transform of the periodic impulse

train $s(t)$ is

$$S(\jmath\Omega) = \frac{2\pi}{T_s} \sum_{k=-\infty}^{\infty} \delta(\Omega - k\Omega_s), \tag{3.5}$$

where $\Omega_s = 2\pi/T_s$ is the sampling frequency in radians/second. Since

$$X_s(\jmath\Omega) = \frac{1}{2\pi} X_c(\jmath\Omega) * S(\jmath\Omega), \tag{3.6}$$

where $*$ denotes convolution. It follows that [19]

$$X_s(\jmath\Omega) = \frac{1}{T_s} \sum_{k=-\infty}^{\infty} X_c(\jmath(\Omega - k\Omega_s)). \tag{3.7}$$

Equation (3.7) states that the Fourier transform of the sampled signal ($X_s(\jmath\Omega)$) consists of periodically repeated copies of $X_c(\jmath\Omega)$. These copies are shifted by integer multiples of the sampling frequency to produce the Fourier transform of $x_s(t)$. Figure 3.2 illustrates the sampling process.

An effect referred to as aliasing occurs when the sampling rate is less than double the highest frequency signal being sampled ($\Omega_s < 2\Omega_N$). Aliasing will corrupt the sampled signal, which leads to the Nyquist sampling theorem which states that the sampling frequency must be at least double that of the highest frequency signal to fully capture the signal [19]. Figure 3.2d illustrates an example of aliasing.

Conversion from the CT sampled signal $x_s(t)$ to the DT signal $x(n)$ involves frequency normalization. If the sampling rate is F_s in samples/second and the frequency of the CT signal being sampled is F_o cycles/second (Hz), then the sampled signal $x(n)$ has a discrete time frequency of $f_o = F_o/F_s$ cycles/sample. Alternatively, if the CT signal is expressed in unit radians/second instead of cycles/second, the CT frequency would be $\Omega_o = 2\pi F_o$ radians/second, and the discrete time frequency would be $\omega_o = \Omega_o/F_s$ radians/sample. Conversions between the units radians and cycles maybe be performed by utilizing the factor 1 cycle $= 2\pi$ radians. Hence $\omega_o = 2\pi f_o$.

The Fourier transform of the DT signal, $X(e^{\jmath\omega})$ is shown in Figure 3.3. When processing a DT signal, we usually only look at the frequencies $-\pi < \omega_o < \pi$ since the spectrum is periodic with period 2π, which means that $-\pi$ to π encompasses all of the unique frequency content of the signal.

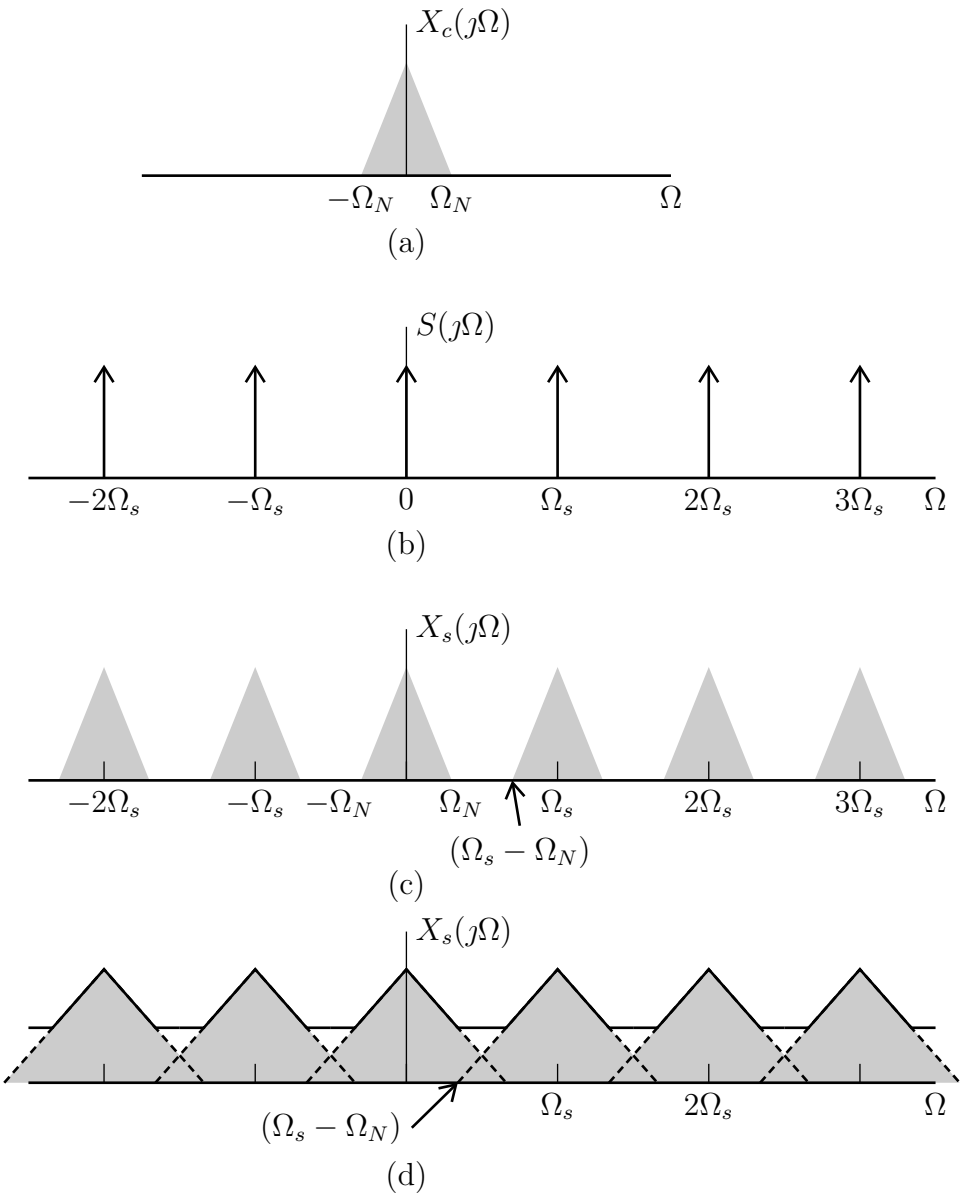
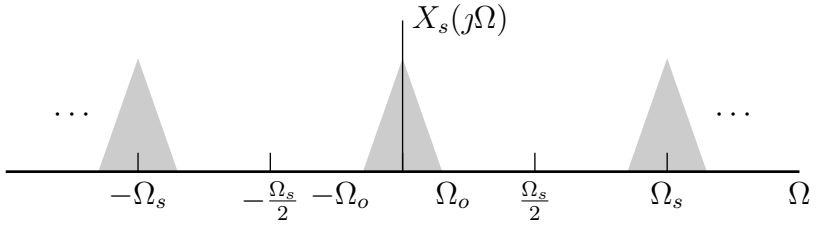


Figure 3.2: Frequency domain representation of sampling in the time domain. (a) Spectrum of the original signal. (b) Fourier transform of the sampling function. (c) Fourier transform of the sampled signal with $\Omega_s > 2\Omega_N$. (d) Fourier transform of the sampled signal with aliasing ($\Omega_s < 2\Omega_N$)

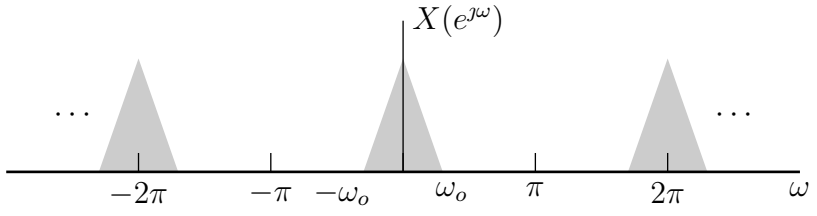
3.2.3 Up-Sampling

Up-sampling involves increasing the sampling rate of a DT signal. Consider the originally sampled signal

$$x(n) = x_c(nT_s). \tag{3.8}$$



(a) Sampled-signal spectrum



(b) Discrete-time signal spectrum

Figure 3.3: Continuous-time and discrete-time signal spectrums

The up-sampled signal would then be

$$x_L(n) = x(n/L) = x_c(nT_s/L), \qquad n = 0, \pm L, \pm 2L, \ldots. \tag{3.9}$$

and $x_L(n) = 0$ for all other values of n. $x_L(n)$ can be alternatively represented by sampling the original discrete signal $x(n)$ with an impulse train

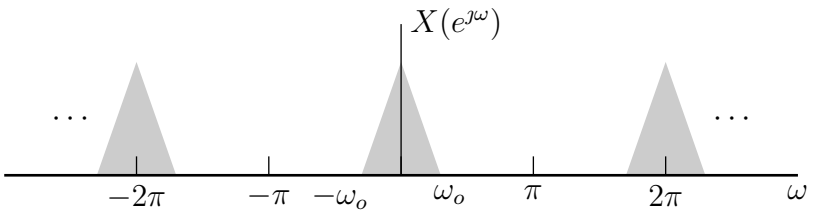
$$x_L(n) = \sum_{k=-\infty}^{\infty} x(k)\delta(n - kL). \tag{3.10}$$

which has the effect of inserting $L - 1$ zeros in-between samples of the original signal. The Fourier transform of $x_L(n)$ can then be described as [19]

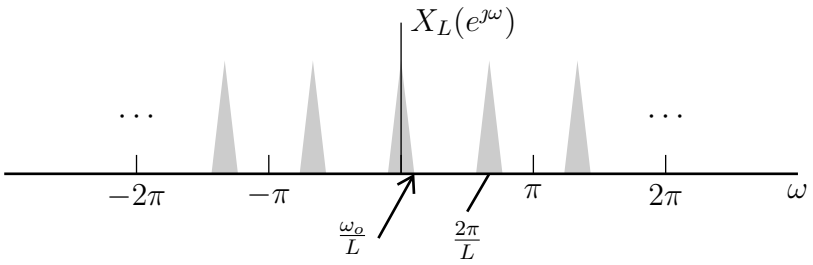
$$\begin{aligned} X_L(e^{\jmath\omega}) &= \sum_{n=-\infty}^{\infty} \left(\sum_{k=-\infty}^{\infty} x(k)\delta(n - kL) \right) e^{-\jmath\omega n} \\ &= \sum_{k=-\infty}^{\infty} x(k)e^{-\jmath\omega Lk} \\ &= X(e^{\jmath\omega L}) \end{aligned} \tag{3.11}$$

Thus, the Fourier transform of the output of an up-sampler is a frequency-scaled version of the Fourier transform of the original signal, i.e. ω is replaced by ωL.

The Fourier transforms of $X(e^{j\omega})$ and $X_L(e^{j\omega})$ are plotted in Figure 3.4. You can see that the effect of up-sampling compresses the signal in the frequency domain, and brings in copies (images), of the signal into the $\omega = -\pi$ to π frequency range, as shown in Figure 3.4b. These images must be filtered away to reconstruct the signal properly.



(a) Original Signal



(b) Signal up-sampled by $L = 3$

Figure 3.4: Frequency-domain representation of a signal up-sampled by a factor of $L = 3$

3.2.4 Down-Sampling

Down-sampling a DT signal $x(n)$ by an integer factor M involves keeping every M^{th} sample from the signal and discarding the rest, and reducing the sampling rate F_s by the same factor. The down-sampled signal $x_M(n)$ can be expressed as

$$x_M(n) = x(nM) = x_c(nMT_s). \tag{3.12}$$

First, recall that the DTFT of $x(n) = x_c(nT_s)$ is

$$X(e^{j\omega}) = \frac{1}{T_s} \sum_{k=-\infty}^{\infty} X_c \left[\jmath \left(\frac{\omega}{T_s} - \frac{2\pi k}{T_s} \right) \right]. \tag{3.13}$$

Similarly, the DTFT of $x_M(n) = x(nM) = x_c(nMT_s)$ is

$$X_M(e^{\jmath\omega}) = \frac{1}{MT_s}\sum_{r=-\infty}^{\infty} X_c\left[\jmath\left(\frac{\omega}{MT_s} - \frac{2\pi r}{MT_s}\right)\right]. \tag{3.14}$$

To see the relationship between between Equations (3.13) and (3.14), note that the summation index in (3.14) can be expressed as

$$r = i + kM \tag{3.15}$$

where k and i are integers such that $-\infty < k < \infty$ and $0 \leq i \leq M-1$. Equation (3.14) can now be expressed as

$$X_M(e^{\jmath\omega}) = \frac{1}{M}\sum_{i=0}^{M-1}\left\{\frac{1}{T_s}\sum_{k=-\infty}^{\infty} X_c\left[\jmath\left(\frac{\omega}{MT_s} - \frac{2\pi k}{T_s} - \frac{2\pi i}{MT_s}\right)\right]\right\}. \tag{3.16}$$

The terms inside the curly brackets in (3.16) is recognized from (3.13) as

$$X(e^{\jmath(\omega-2\pi i)/M}) = \frac{1}{T_s}\sum_{k=-\infty}^{\infty} X_c\left[\jmath\left(\frac{\omega - 2\pi i}{MT_s} - \frac{2\pi k}{T_s}\right)\right]. \tag{3.17}$$

Thus, we can express Equation (3.17) as [19]

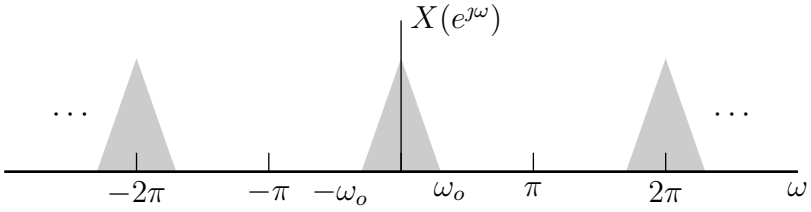
$$X_M(e^{\jmath\omega}) = \frac{1}{M}\sum_{i=0}^{M-1} X(e^{\jmath(\omega/M - 2\pi i/M)}) \tag{3.18}$$

Equation (3.18) tells us that the spectrum of the down-sampled signal ($X_M(e^{\jmath\omega})$) is equal to a scaled version of the spectrum of the original signal ($X(e^{\jmath\omega})$) with the spectral shape widened by a factor of M, and with images at integer multiples of 2π as is the case for all DT signals.

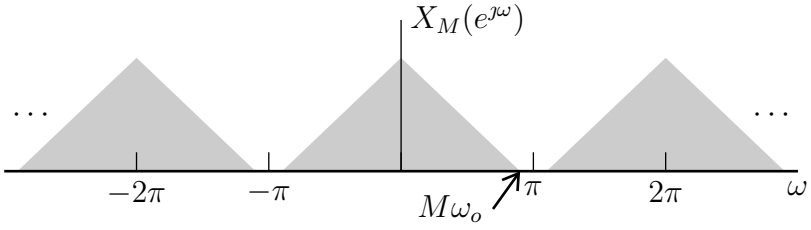
The Fourier transform of a DT signal before and after down-sampling is shown in Figure 3.5. It is clear that down-sampling expands the range of occupied frequencies by a factor of M. It is also important to note that the signal being down-sampled must be band-limited such that $\omega_o < \pi/M$ otherwise aliasing will occur, as shown in Figure 3.5c.

3.2.5 Discrete-to-Continuous Conversion

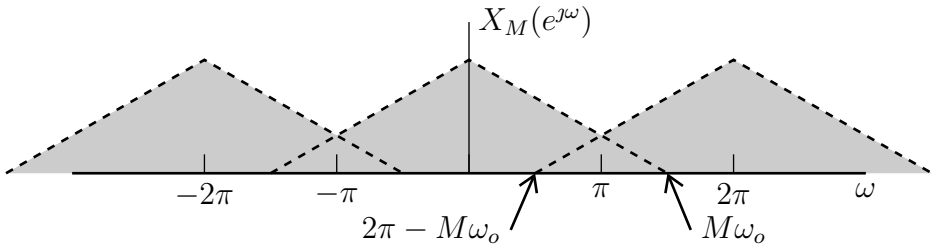
According to the Nyquist sampling theorem, a continuous-time signal can be entirely reconstructed from its samples if the samples were taken at a rate greater than twice its



(a) Original Signal



(b) Signal down-sampled by $M = 3$



(c) Aliasing resulting from down-sampling when $\omega_o > \pi/M$

Figure 3.5: Frequency spectrum of a signal down-sampled by a factor of $M = 3$

frequency [19]. The mathematical model of discrete-to-continuous conversion first involves creating the sampled signal $x_s(t)$ discussed in the previous section, followed by a CT low pass filter to remove the high-frequency signal copies.

The sampled signal $x_s(t)$ can be formulated by multiplying the sequence $x(n)$ by an impulse train,

$$y_s(t) = \sum_{n=-\infty}^{\infty} x(n)\delta(t - nT_s), \tag{3.19}$$

where the n^{th} sample is associated with the impulse at $t = nT_s$, and T_s is the sampling period associated with $x(n)$. This impulse train is the input to an ideal low pass continuous-time

reconstruction filter with frequency response $H_r(\jmath\Omega)$ and impulse response $h_r(t)$. The output of the filter will be the reconstructed CT signal

$$y_r(t) = \sum_{n=-\infty}^{\infty} x(n)h_r(t - nT_s) \tag{3.20}$$

which in theory will exactly equal the originally sampled signal $x_c(t)$. The D-to-C process is depicted in Figure 3.6.

3.3 Ultra-Wideband Harmonic Theory

Calculating the harmonics in an ultra-wideband system is similar to the narrow-band calculations in Chapter 2, however the larger bandwidths present new challenges. We want to model the harmonics seen at baseband created from an ultra-wide passband signal to perform baseband echo-cancellation.

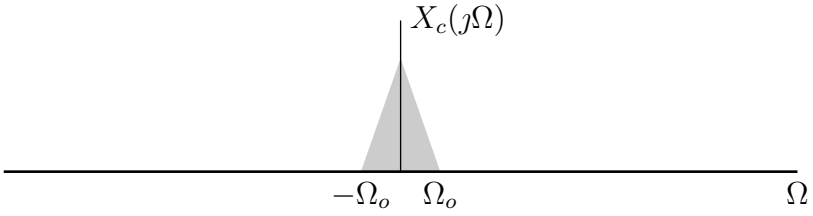
An ultra-wide signal is defined as a signal whose absolute bandwidth is large relative to its center frequency. The metric used to make this classification is the fractional bandwidth, defined by [39]

$$\text{fractional bandwidth } = \frac{f_n - f_1}{f_c}, \tag{3.21}$$

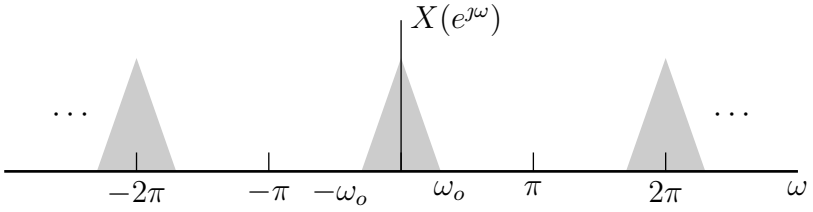
where f_n, f_1, and f_c is the highest, lowest, and center frequency of the passband signal. When the fractional bandwidth is above 0.2 [40][1], the application is deemed ultra-wideband. For the DOCSIS 3.1 FDX application, recall from Figure 1.5 that the downstream signal spans from 108 MHz to 1218 MHz and has a center frequency of 663 MHz. Thus, the fractional bandwidth of this signal is $(1218 - 108)/663 = 1.7$ and the DOCSIS 3.1 DS signal is therefore considered an ultra-wideband signal. In contrast, typical wireless applications (the focus of most existing PA modeling literature) are very narrowband, with fractional bandwidths mostly in the range 0.01 to 0.04 [23].

For narrowband applications, only harmonics with center frequencies at f_c must be modeled as was shown in Figure 2.12 in Chapter 2 since the harmonics with center frequencies other than f_c do not interfere with the signal of interest. However, with ultra-wideband

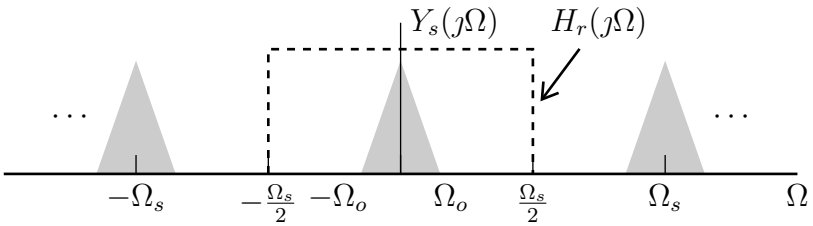
[1]The definition of an UWB signal is vague and depends on the application. See the Appendix for further details.



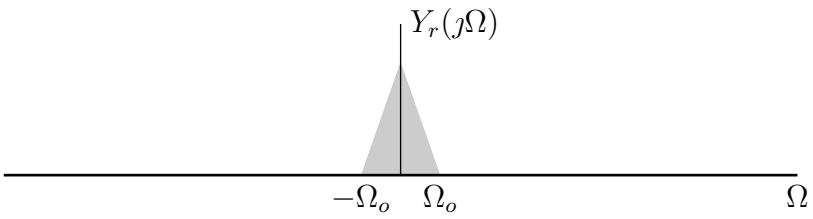
(a) Original Signal



(b) Discrete time signal



(c) New CT signal with an Ideal LPF



(d) Reconstructed CT signal

Figure 3.6: Reconstructing a CT signal from a DT signal

signals like the one shown in Figure 2.13, harmonics with center frequencies other than f_c, such as DC, $2f_c$, and $3f_c$, will interfere with the SOI and therefore must also be modeled.

With ultra-wideband signals, the harmonics can grow so wide in frequency that the

harmonics centered at positive frequencies can spread past 0 Hz and into the negative frequencies. Since real signals are symmetric about 0 Hz, this situation results in the spectrum "folding" back into the positive frequencies, as illustrated in Figure 3.7. This "folded" back signal can also interfere with the SOI, however by computing the harmonic correctly it can still be accounted for.

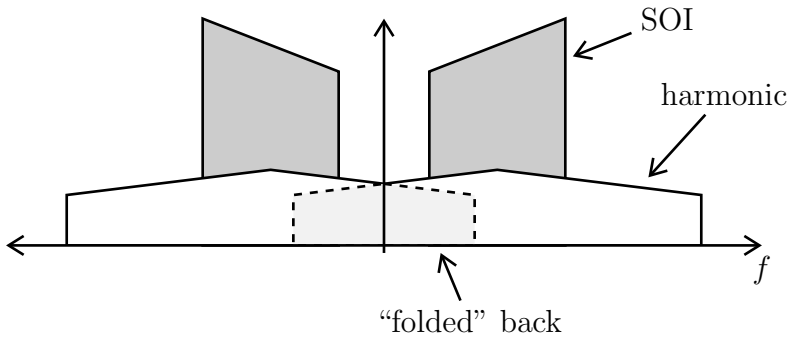


Figure 3.7: Example of a harmonic "folding" back into the signal of interest

An obvious starting point for computing the OOB harmonics might be to raise the passband signal to various powers to calculate all of the harmonics. However, this would unnecessarily calculate the high-frequency harmonics that do not interfere with the signal of interest, and it would require a higher sampling rate to handle the high-frequency signals. A more desirable approach is to calculate only the harmonics that interfere with the SOI. The SOI in the DOCSIS FDX case is the upstream signal. Since the US signal frequency spreads from 5 MHz to 684 MHz, only the harmonics that overlap with these frequencies need to be calculated.

For echo-cancellation to execute at baseband, the harmonics should be calculated from the baseband transmitted signal. Recall from Equation (2.22) that the real passband Tx signal is given from the complex baseband Tx signal by

$$x(n) = \Re\left\{\widetilde{x}(n)e^{\jmath\omega_c n}\right\} = \frac{1}{2}\left(\widetilde{x}(n)e^{\jmath\omega_c n} + \widetilde{x}^*(n)e^{-\jmath\omega_c n}\right), \tag{3.22}$$

where $\omega_c = 2\pi f_c$, and f_c is the carrier frequency. Now we can square $x(n)$ to find the second order harmonics,

$$x^2(n) = \frac{1}{4}\left((\widetilde{x}(n)e^{j\omega_c n})^2 + 2|\widetilde{x}(n)|^2 + (\widetilde{x}^*(n)e^{-j\omega_c n})^2\right). \tag{3.23}$$

Using Equation (3.22) and the property of complex conjugates that $(\widetilde{x}^*)^2 = (\widetilde{x}^2)^*$, we can write $x^2(n)$ as

$$x^2(n) = \frac{1}{2}\Re\left\{|\widetilde{x}(n)|^2 + (\widetilde{x}(n)e^{j\omega_c n})^2\right\}. \tag{3.24}$$

We can now calculate the third order harmonics:

$$\begin{aligned} x^3(n) =& x^2(n)x(n) \\ =& \frac{1}{8}\left((\widetilde{x}(n)e^{\jmath\omega_c n})^2 + 2|\widetilde{x}(n)|^2 + (\widetilde{x}^*(n)e^{-\jmath\omega_c n})^2\right)(\widetilde{x}(n)e^{\jmath\omega_c n} + \widetilde{x}^*(n)e^{\jmath\omega_c n}) \\ =& \frac{1}{8}\left((\widetilde{x}(n)e^{\jmath\omega_c n})^3 + 3|\widetilde{x}(n)|^2\widetilde{x}(n)e^{\jmath\omega_c n} + 3|\widetilde{x}(n)|^2\widetilde{x}^*(n)e^{-\jmath\omega_c n} + (\widetilde{x}^*(n)e^{-\jmath\omega_c n})^3\right) \\ =& \frac{1}{4}\Re\left\{(\widetilde{x}(n)e^{\jmath\omega_c n})^3 + 3|\widetilde{x}(n)|^2\widetilde{x}(n)e^{\jmath\omega_c n}\right\}. \end{aligned} \tag{3.25}$$

By following the same process used to generate Equations (3.24) and (3.25), the K^{th} order passband harmonics are generalized by Equation (3.26). Note that the scaling constants corresponding to the individual harmonic terms have been omitted from (3.26) for brevity. The reason this omission is justified is that the PA modeling process applies unknown coefficients to each harmonic component in order to model device specific variations, so the specific constants in (3.26) are of little consequence.

$$x^K(n) = \begin{cases} \Re\left\{|\widetilde{x}(n)|^K + |\widetilde{x}(n)|^{K-2}(\widetilde{x}(n)e^{j\omega_c n})^2 \right. \\ \qquad \left. +|\widetilde{x}(n)|^{K-4}(\widetilde{x}(n)e^{j\omega_c n})^4 + \cdots + (\widetilde{x}(n)e^{j\omega_c n})^K\right\}, & K \text{ even} \\ \Re\left\{|\widetilde{x}(n)|^{K-1}(\widetilde{x}(n)e^{j\omega_c n}) + |\widetilde{x}(n)|^{K-3}(\widetilde{x}(n)e^{j\omega_c n})^3 \right. \\ \qquad \left. +|\widetilde{x}(n)|^{K-5}(\widetilde{x}(n)e^{j\omega_c n})^5 + \cdots + (\widetilde{x}(n)e^{j\omega_c n})^K\right\}, & K \text{ odd} \end{cases} \tag{3.26}$$

For $K \geq 1$. Equation (3.26) can be compressed using summation notation:

$$x^K(n) = \begin{cases} \sum\limits_{\substack{r=0 \\ r \text{ even}}}^{K} |\widetilde{x}(n)|^{K-r}\Re\left\{(\widetilde{x}(n)e^{j\omega_c n})^r\right\}, & K \text{ even} \\ \sum\limits_{\substack{r=1 \\ r \text{ odd}}}^{K} |\widetilde{x}(n)|^{K-r}\Re\left\{(\widetilde{x}(n)e^{j\omega_c n})^r\right\}, & K \text{ odd} \end{cases}, \quad K \geq 1 \tag{3.27}$$

Equation (3.27) can be used to calculate the individual passband harmonics using the baseband Tx signal. Figure 3.8 shows the position of each harmonic in a DOCSIS FDX

application with the bandwidths drawn to scale in frequency. Each harmonic is labeled with its equation in terms of the complex baseband Tx signal given in (3.27), with the $\Re\{\cdot\}$ omitted. Only the DC and positive frequencies are shown in the figure, but recall that each signal will also have a negative frequency component because the passband signal is purely real.

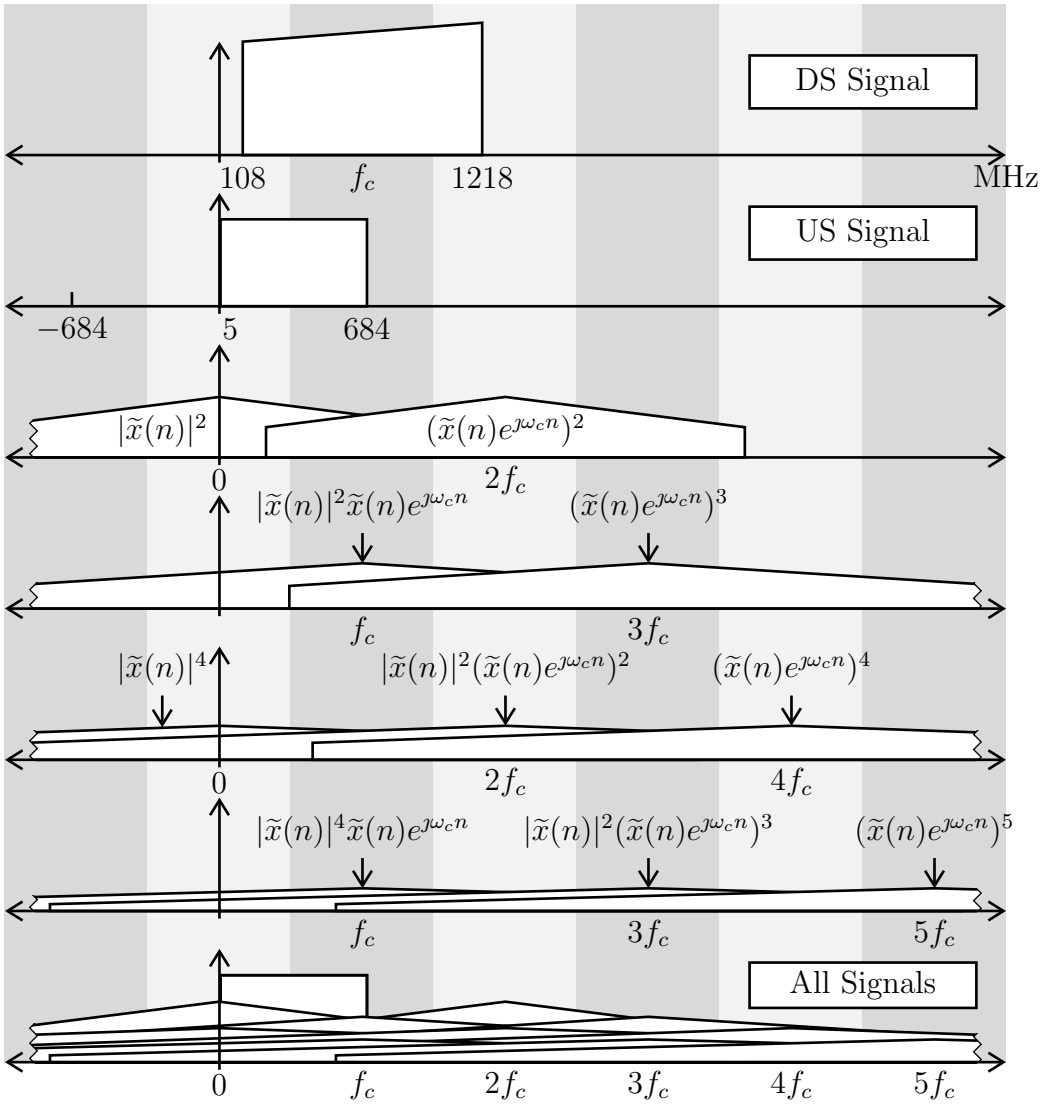


Figure 3.8: Harmonics in a DOCSIS FDX application (bandwidths drawn to scale)

For DOCSIS FDX echo-cancellation, we are concerned with the harmonics that interfere

with the upstream frequency band from 5 MHz to 684 MHz. Figure 3.8[2] provides a good starting point for identifying which individual harmonics will need to be calculated. However, the power in each harmonic differs from one device to the next, and will also depend on the operating point (output power, temperature, etc.) of the power amplifier. A method for determining which harmonics need to be canceled could be canceling one at a time and observing the reduction in error power in the received signal. A significant reduction in error power implies the harmonic in question has substantial power in the US signal frequency band.

Upon calculating the passband harmonics, the harmonics will have to be down-converted to baseband as would the received US signal. Down-conversion is executed by multiplying each harmonic by $e^{-\jmath\omega_c n}$ as described in Chapter 2. After down-conversion, all of the baseband signals should be in the correct locations for echo-cancellation. Unfortunately, practical hardware has limitations on how high of a sampling rate it can handle. The impact of these limitations on the ultra-wideband PA modeling process is discussed in the following sections.

3.4 Sampling Rate Limitations

Practical devices such as application-specific integrated circuits (ASICs) and field-programmable gate arrays (FPGAs) have physical limitations that cap their maximum sampling rate. For example, a top of the line FPGA such as the Intel Arria 10 has a maximum clock rate of 800 MHz [41][3]. Thus, only sampling rates of up to 800 MHz can directly be used, corresponding to signals with bandwidths of up to 400 MHz (according to the Nyquist theorem). However, operations can happen in parallel in an FPGA, thereby providing an effectively higher sampling rate. For example, the sampling rate could be increased from 800 MHz to 3200 MHz by executing four parallel operations at each clock pulse. The trade-off in doing so is at least four times the hardware resources will be used inside the FPGA, and the implementation complexity will significantly increase.

[2]Figure derived from [11]

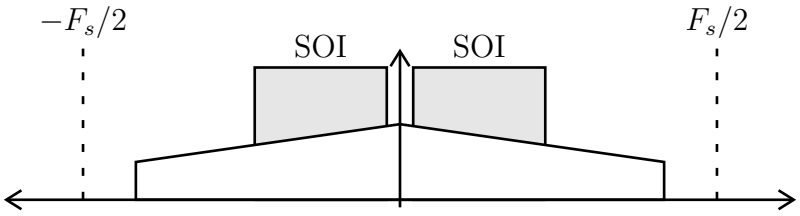
[3]In practice, FPGA designs typically operate will below this maximum due to difficulties in meeting register setup and hold time requirements when the maximum clock rate listed in the datasheet is approached.

Although a harmonic from a power amplifier might extend to a very high frequency, the entirety of the harmonic does not have to be known to perform echo-cancellation. Only the portion of the harmonic interfering with the US signal must be known since that is the part we need to cancel, and the remaining portion of the harmonic can be filtered away. From Equation (3.27), we see that each individual harmonic can be calculated at baseband prior to up-conversion by first evaluating all terms with $\widetilde{x}(n)$, then multiplying by $(e^{\jmath\omega_c n})^k$ to shift the harmonic to its proper center frequency.

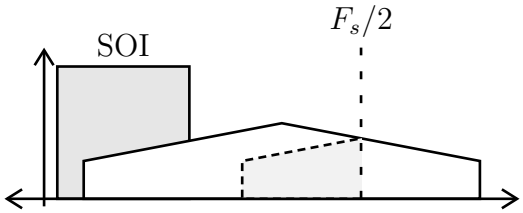
Recall that if the frequency of a sampled signal extends past the Nyquist rate, or $F_s/2$, the signal will alias. The aliased signal is mirrored about $F_s/2$, and the portion of the signal that is aliased is corrupted. Fortunately, we are not concerned about a corrupted signal outside of the US signal frequency band because it will eventually be filtered away. Therefore, when the harmonic at baseband is up-converted to passband, we can allow it to alias as long as the bandwidth of the signal beyond $F_s/2$ is less than the frequency spread between $F_s/2$ and the upper edge of the US signal (684 MHz). In this case, the portion of the harmonic overlapping with the US signal will remain correct.

Two different situations need to be considered. The first situation arises when the baseband version of harmonic fits within the Nyquist rate like that in Figure 3.9a. If the up-converted harmonic aliases such that the aliased part of the signal does not interfere with the SOI such as the example in Figure 3.9b, then the in-band harmonic will be correct.

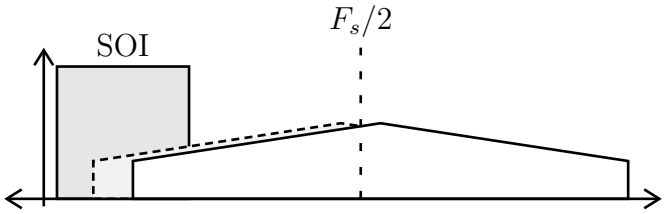
If, on the other hand, the harmonic is aliased such as that in Figure 3.9c where aliasing occurs within the SOI, the calculated harmonic will not be correct. In this case, there are two options for solving the problem. The first is to increase the sampling rate such that the in-band signal no longer aliases on top of the SOI. The second option is to apply a complex LPF to the signal at baseband to band-limit it as illustrated in Figure 3.10a. The impulse response of the filter must be complex so that it can attenuate positive frequencies that would alias into the SOI while allowing the negative frequencies that overlap with the SOI at passband to pass. A filter with a real-valued impulse response would attenuate both positive and negative frequencies. The result of up-converting the filtered harmonic to passband is illustrated in Figure 3.10b.



(a) Baseband harmonic is not aliased



(b) Up-converted harmonic *does not* alias into the SOI

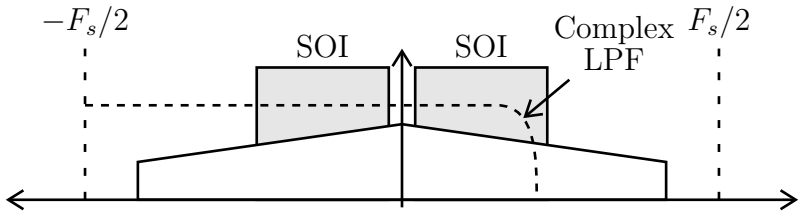


(c) Up-converted harmonic *does* alias into the SOI

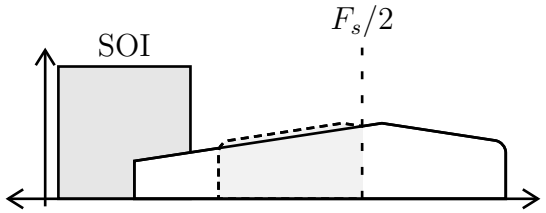
Figure 3.9: Examples of aliased harmonics (gray dashed signals are aliases)

Another situation to consider occurs when the baseband harmonic does not fit within the Nyquist rate as shown in Figure 3.11. Upon up-converting the harmonic, if the aliased portions of the signal do not interfere with the SOI (both positive and negative frequencies), then the aliasing is not a concern as it will eventually be filtered away in the receiver. However, if the aliased portions do interfere with the SOI, then it is necessary to increase the sampling rate.

Finally, with our understanding of how to calculate ultra-wideband harmonics and the associated sampling rate limitations, we can devise an ultra-wideband echo-cancellation model for the DOCSIS FDX application.



(a) Complex LPF applied to baseband harmonic prior to up-conversion



(b) Up-converted filtered harmonic no longer aliases into the SOI

Figure 3.10: Examples of aliased harmonics (gray dashed signals are aliases)

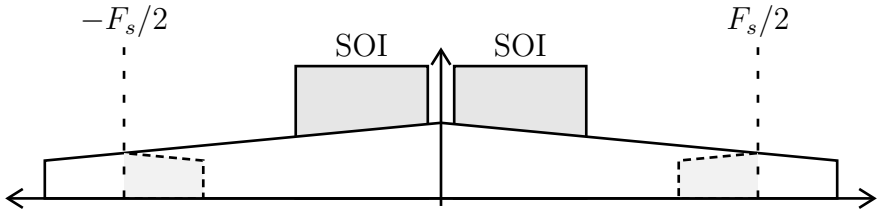


Figure 3.11: Baseband harmonic aliases

3.5 Ultra-Wideband Harmonic Model

Given the baseband transmitted signal $\widetilde{x}(n)$, we want to approximate the baseband equivalent of the output of the power amplifier $\widetilde{y}(n)$. First, we must decide on a nonlinear model to adapt into an ultra-wideband model. The performance and resource usage of various models have been compared in many papers including [37, 42, 43]. In general, polynomial-based models exhibit superior performance with reasonable complexity. The researchers in [37, 42] compare the MP and GMP models directly and find that the GMP has slightly better performance for the same number of coefficients, however, the GMP model has a more complex structure and is, therefore, more challenging to implement. In this thesis,

we will adapt the memory polynomial model to calculate individual out-of-band harmonics due to its adequate performance and simple structure, which simplifies the derivation of the ultra-wideband model.

To begin the formulation of the UWB-MP model, recall from Eq. (3.27) that the K^{th} order passband harmonics calculated from the baseband signal $\widetilde{x}(n)$ are given by

$$x^K(n) = \begin{cases} \sum\limits_{\substack{r=0 \\ r \text{ even}}}^{K} |\widetilde{x}(n)|^{K-r} \Re\left\{(\widetilde{x}(n)e^{j\omega_c n})^r\right\}, & K \text{ even} \\ \sum\limits_{\substack{r=1 \\ r \text{ odd}}}^{K} |\widetilde{x}(n)|^{K-r} \Re\left\{(\widetilde{x}(n)e^{j\omega_c n})^r\right\}, & K \text{ odd} \end{cases}, \quad K \geq 1.$$

This equation can be compressed by allowing r to be an element of the set $\{R_K\}$, where the set includes the values of r for the given harmonic order K. In general, $R_K = \{0, 2, 4, ..., K\}$ for even K, and $R_K = \{1, 3, 5, ..., K\}$ for odd K. However, R_K can be selective, meaning not all values of r need to be included in the model, depending on the application. For example, consider the 5^{th} order harmonics. Perhaps only the the harmonics at f_c and $3f_c$ are to be included and the harmonic at $5f_c$ is to not be included. In this case, $R_5 = \{1, 3\}$. Re-writing Eq. (3.27) with this new notation results in

$$x^K(n) = \sum_{r \in R_K} |\widetilde{x}(n)|^{K-r} \Re\left\{(\widetilde{x}(n)e^{\jmath\omega_c n})^r\right\}, \qquad K \geq 1. \tag{3.28}$$

Now let us form a summation of passband harmonics for orders up to and including K. We do so by replacing K with k in Equation (3.28) and summing k from 1 to K. The resulting Equation is

$$\sum_{k=1}^{K} x^k(n) = \sum_{k=1}^{K} \sum_{r \in R_k} |\widetilde{x}(n)|^{k-r} \Re\left\{(\widetilde{x}(n)e^{\jmath\omega_c n})^r\right\}, \qquad K \geq 1. \tag{3.29}$$

Finally, by including memory samples as well as a set of unknown coefficients for each harmonic term, we arrive at the ultra-wideband memory polynomial model for *passband* harmonics:

$$y_{\text{UWB-MP}}(n) = \sum_{k=1}^{K} \sum_{r \in R_k} \sum_{m=0}^{M-1} b_{krm} |\widetilde{x}(n-m)|^{k-r} \Re\left\{\left(\widetilde{x}(n-m)e^{\jmath\omega_c(n-m)}\right)^r\right\} \tag{3.30}$$

where,

- K is the highest order of nonlinearity,
- R_k is the set of individual passband harmonics to be included for the order k, $R_k = \{0, 2, 4, ..., k\}$, k even,
  $R_k = \{1, 3, 5, ..., k\}$, k odd,
  R_k is selective, i.e. not all values have to be used,
- M is the number of samples (memory) to include,
- b_{krm} is the real coefficient corresponding to the r^{th} harmonic of order k with a delay of m samples.

To complete the baseband power amplifier model, the passband harmonics in (3.30) must be down-converted to baseband and filtered with a LPF. The down-conversion involves multiplication by $e^{-\jmath\omega_c n}$ which produces the desired baseband signal along with a double frequency component. Note that again we will drop the coefficient b_{krm} since it is for the passband harmonics, and instead we will introduce a coefficient for the baseband harmonics when completing the baseband model. The down-converted unfiltered signal is given by

$$e^{-\jmath\omega_c n} \sum_{k=1}^{K} \sum_{r\in R_k} \sum_{m=0}^{M-1} |\widetilde{x}(n-m)|^{k-r} \Re\left\{ \left(\widetilde{x}(n-m)e^{\jmath\omega_c(n-m)}\right)^r \right\}. \tag{3.31}$$

The $e^{-\jmath\omega_c n}$ term can be moved to inside the summation. After doing so, (3.31) is re-written as:

$$\sum_{k=1}^{K} \sum_{r\in R_k} \sum_{m=0}^{M-1} e^{-\jmath\omega_c n} |\widetilde{x}(n-m)|^{k-r} \Re\left\{ \left(\widetilde{x}(n-m)e^{\jmath\omega_c(n-m)}\right)^r \right\}. \tag{3.32}$$

Finally, the down-converted signal in (3.32) is filtered to remove the double-frequency component with the following convolution, where $h(l)$ are the coefficients for a low pass filter and L is the length of the filter:

$$\sum_{l=0}^{L-1} h(l) \sum_{k=1}^{K} \sum_{r\in R_k} \sum_{m=0}^{M-1} e^{-\jmath\omega_c n} |\widetilde{x}(n-m-l)|^{k-r} \Re\left\{ \left(\widetilde{x}(n-m-l)e^{\jmath\omega_c(n-m-l)}\right)^r \right\}. \tag{3.33}$$

As was done with the $e^{-\jmath\omega_c n}$ term, we will move the $\sum_{l=0}^{L-1} h(l)$ term to be inside the other summations. The resulting equation is:

$$\sum_{k=1}^{K} \sum_{r\in R_k} \sum_{m=0}^{M-1} \sum_{l=0}^{L-1} h(l) e^{-\jmath\omega_c n} |\widetilde{x}(n-m-l)|^{k-r} \Re\left\{ \left(\widetilde{x}(n-m-l)e^{\jmath\omega_c(n-m-l)}\right)^r \right\}. \tag{3.34}$$

Equation (3.34) can be simplified by substituting

$$\psi(n,k,r,m) = \sum_{l=0}^{L-1} h(l) e^{-\jmath\omega_c n} |\widetilde{x}(n-m-l)|^{k-r} \Re\left\{ \left(\widetilde{x}(n-m-l) e^{\jmath\omega_c (n-m-l)} \right)^r \right\}. \quad (3.35)$$

Finally, the ultra-wideband polynomial model for *baseband* signals is given in Equation (3.36), where the coefficients a_{krm} have been re-introduced for the baseband model.

$$\widetilde{y}_{\text{UWB-MP}}(n) = \sum_{k=1}^{K} \sum_{r \in R_k} \sum_{m=0}^{M-1} a_{krm} \psi(n,k,r,m) \quad (3.36)$$

The benefit of moving the down-conversion multiplication and LPF inside the summations is it allows us to down-sample the signals before convolving them with the a_{krm} coefficients. Down-sampling enables this convolution to happen at a lower clock rate, therefore using fewer resources.

A block diagram showing the sequence of operations in generating $\widetilde{y}_{\text{UWB-MP}}(n)$ from $\widetilde{x}(n)$ is given in Figure 3.12.

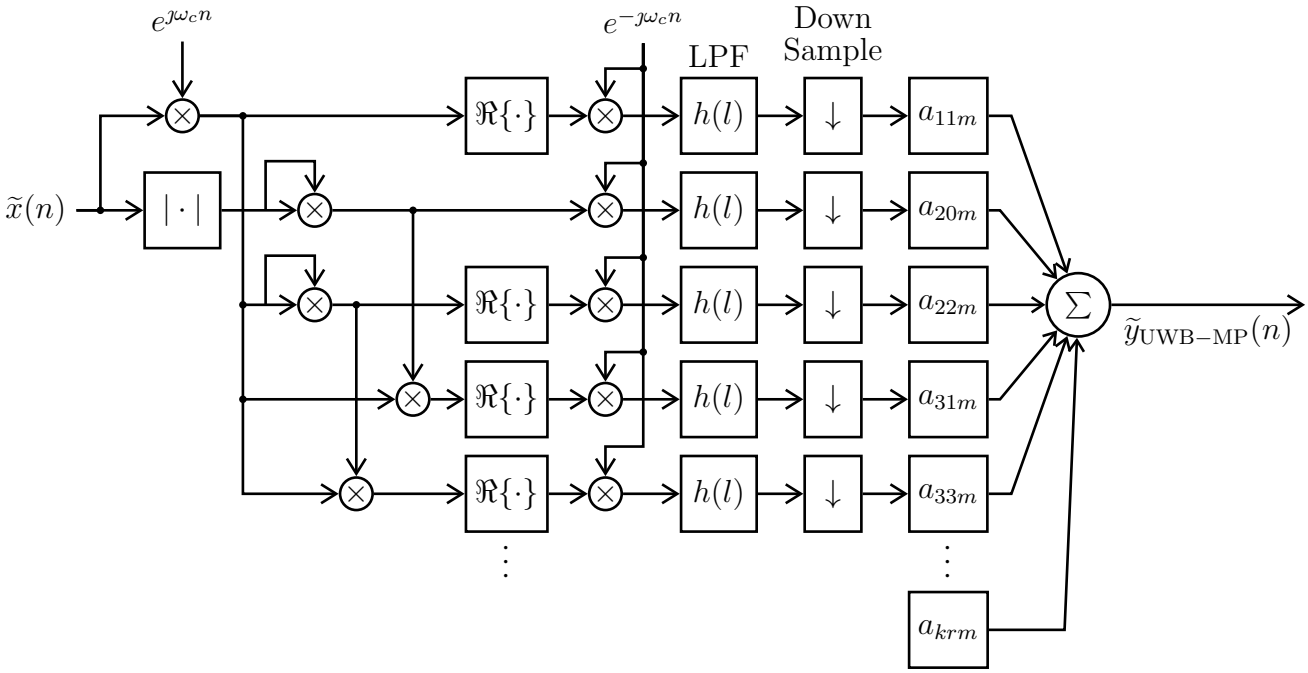


Figure 3.12: Ultra-wideband memory polynomial block diagram

Equations (3.36) and (3.35) provide a full mathematical representation of how the received baseband signal from the output of a power amplifier can be computed given the

baseband input to the amplifier. Although it might look daunting, it arises from the cascade of a few simple operations. To summarize, the $|\widetilde{x}(n-m)|^{k-r}\Re\left\{\left(\widetilde{x}(n-m)e^{\jmath\omega_c(n-m)}\right)^r\right\}$ term comes from the passband UWB-MP model from Equation (3.30), the $e^{-\jmath\omega_c n}$ term down-converts the harmonics to baseband, and the $h(l)$ term and substitution of n with $n-l$ results from the low pass filtering operation after the down-conversion.

The next step in predicting the output of an ultra-wideband PA is to use a parameter estimation technique to solve for the unknown a_{krm} coefficients in Equation (3.36).

3.6 Parameter Estimation

A parameter estimation technique must be used to estimate the coefficients a_{krm} for the ultra-wideband memory polynomial model in Equation (3.36). Since the output $\widehat{\widetilde{y}}_{UWB-MP}(n)$ is linearly dependent on a_{krm}, a linear parameter estimation technique can be used. In literature, the most common technique used for estimating the parameters for a memory polynomial model is least squares (LS) [21, 22, 35, 44], so the LS method will also be used here. First, we will go through the development of the LS algorithm, then apply it to the MP model, and finally extend it to the UWB-MP model.

3.6.1 Least Squares

The basic idea of the method of least squares is to fit a curve to a set of measured data in a manner which minimizes the sum of squares of the difference between the measured and calculated outputs [45].

Consider a physical phenomenon that is characterized by two sets of variables, $\{y(n)\}$ and $\{x(n)\}$, where $y(n)$ at sample n is a function of the subset of variables $x(n), x(n-1), ..., x(n-M+1)$ applied as inputs. If this relationship is linear, we may express $y(n)$ as [45]

$$y(n) = \sum_{i=0}^{M-1} a_i x(n-i) + e(n) \tag{3.37}$$

where a_i are unknown parameters of the model, and $e(n)$ represents the measurement error. In the absence of evidence to the contrary, it is customary to assume that the measurement

error is a white noise process with zero mean.

The problem we want to solve is to estimate the unknown parameters a_i given the two observable sets of data $\{x(n)\}$ and $\{y(n)\}$, $n = 0, 1, ..., N-1$. In this thesis, we will denote an estimated value with an overhead hat $(\widehat{\cdot})$, so the estimates of the unknown parameters are denoted $\widehat{a}_i$. Once the $\widehat{a}_i$ parameters have been estimated, they can be used to generate estimates of the output data $\widehat{y}(n)$ based on the known input samples $x(n)$ as follows:

$$\widehat{y}(n) = \sum_{i=0}^{M-1} \widehat{a}_i x(n-i). \tag{3.38}$$

The error in the estimated output samples is then calculated by:

$$e(n) = y(n) - \widehat{y}(n). \tag{3.39}$$

In the method of least squares, the estimated coefficients $\widehat{a}_i$ are chosen to minimize the sum of squared errors where M is the memory depth (number of coefficients) and N is the number of equations to solve (number of observations of $x(n)$ and $y(n)$). The LS algorithm generally operates on an over-determined system, meaning there are more equations than there are variables ($N > M$).

From [45], the LS algorithm using the inputs and outputs of a system can be expressed as follows. Equation (3.38) can be put into the matrix form

$$\mathbf{y} = \mathbf{X}\mathbf{A} \tag{3.40}$$

where the desired response over some measurement interval of length N can be written in matrix form as an $(N \times 1)$ column matrix,

$$\mathbf{y} = [y(0), y(1), ..., y(N-1)]^T, \tag{3.41}$$

the coefficient matrix as an $(M \times 1)$ column matrix,

$$\mathbf{A} = [a_0, a_1, ..., a_{M-1}]^T, \tag{3.42}$$

and the input written as an $(N \times M)$ rectangular matrix,

$$\mathbf{X} = \begin{bmatrix} x(0) & x(1) & \dots & x(M-1) \\ x(1) & x(2) & \dots & x(M) \\ \vdots & \vdots & \ddots & \vdots \\ x(N-1) & x(N) & \dots & x(N+M-2) \end{bmatrix}. \tag{3.43}$$

The coefficients can then be estimated as [45]

$$\widehat{\mathbf{A}} = (\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H\mathbf{y} \tag{3.44}$$

where $(\cdot)^H$ denotes a Hermitian (complex conjugate transpose) matrix transform. In a dynamic system such as a power amplifier, the true coefficients can change over time. The estimated coefficients in this case can be updated using a gradient-based approach [45]

$$\widehat{\mathbf{A}}^{(m)} = \widehat{\mathbf{A}}^{(m-1)} + \mu(\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H\mathbf{e} \tag{3.45}$$

where μ is the step size, $\mathbf{e} = \mathbf{y} - \mathbf{X}\widehat{\mathbf{A}}^{(m-1)}$ is the $(N \times 1)$ error matrix between the desired signal $\mathbf{y}$ and the LS estimation $\mathbf{X}\widehat{\mathbf{A}}^{(m-1)}$ from the previous iteration, and $\widehat{\mathbf{A}}^{(0)}$ is the initial estimation from (3.44). For a small step size μ, the algorithm converges more slowly but will result in a more accurate final estimate, whereas a larger μ will converge more quickly but will have more error in the converged solution. It is usually a good idea to begin with a larger μ then reduce μ as the estimation converges.

3.6.2 Least Squares for Nonlinear Systems

The least squares algorithm can be used to solve nonlinear systems as long as the system is *linearly* dependent on the coefficients. As we saw in Chapter 2, the output of the memory polynomial model is nonlinearly dependent on the input, and *linearly* dependent on the coefficients, so the LS algorithm is suitable to calculate the coefficients.

Here, the LS algorithm will first be used to estimate the coefficients for the memory polynomial model, then further adapted for the ultra-wideband memory polynomial model.

Recall from (2.29) the equation of the memory polynomial model with all values of k

included,

$$\widetilde{y}(n) = \sum_{k=0}^{K-1} \sum_{m=0}^{M-1} a_{km} \widetilde{x}(n-m) |\widetilde{x}(n-m)|^k. \tag{3.46}$$

The MP model can be put into matrix form [35]:

$$\mathbf{y} = \mathbf{X}\mathbf{A} \tag{3.47}$$

where $\mathbf{y}$ is the $(N \times 1)$ matrix,

$$\mathbf{y} = [\widetilde{y}(0), \widetilde{y}(1), ..., \widetilde{y}(N-1)]^T, \tag{3.48}$$

$\mathbf{A}$ is the $(KM \times 1)$ coefficient matrix,

$$\mathbf{A} = [a_{00}, ..., a_{km}, ..., a_{K-1,M-1}]^T, \tag{3.49}$$

and $\mathbf{X}$ is the $(N \times KM)$ rectangular input matrix,

$$\mathbf{X} = [\Psi_0, \Psi_1, ..., \Psi_{N-1}]^T, \tag{3.50}$$

where Ψ_n is the $(KM \times 1)$ matrix,

$$\Psi_n = [\psi(n,0,0), ..., \psi(n,k,m), ..., \psi(n,K-1,M-1)]^T, \tag{3.51}$$

and

$$\psi(n,k,m) = \widetilde{x}(n-m)|\widetilde{x}(n-m)|^k. \tag{3.52}$$

Thus, for an over-determined system ($N > M$), the coefficient matrix is estimated by the LS equations given in (3.44) and (3.45).

Now we can further extend the LS algorithm to calculate the coefficients for the ultra-wideband memory polynomial model. In order to do so, we need to define a new constant S which equals the total number of values in all $\{R_k\}$ sets. For example, if $K = 2$ and $R_1 = \{1\}, R_2 = \{0, 2\}$, then $S = 3$. This can be represented by the following equation,

$$S = \sum_{k=1}^{K} \overline{\overline{R}}_k, \tag{3.53}$$

where $\overline{\overline{R}}_k$ is the cardinal number for R_k, which is equal to the number of elements in the set.

Now we put the UWB-MP model into matrix form to solve for the coefficients using the LS method. From Equation (3.36), $\widetilde{y}_{\mathrm{UWB-MP}}(n)$ can be put into the matrix form,

$$\mathbf{y} = \mathbf{X}\mathbf{A} \tag{3.54}$$

where $\mathbf{y}$ is the $(N \times 1)$ matrix,

$$\mathbf{y} = [\widetilde{y}(0), \widetilde{y}(1), ..., \widetilde{y}(N-1)]^T, \tag{3.55}$$

$\mathbf{A}$ is the $(SM \times 1)$ coefficient matrix,

$$\mathbf{A} = [a_{110}, a_{200}, a_{220}, a_{310}, ..., a_{krm}, ..., a_{K,R_K,M-1}]^T, \tag{3.56}$$

and $\mathbf{X}$ is the $(N \times SM)$ rectangular input matrix,

$$\mathbf{X} = [\Psi_0, \Psi_1, ..., \Psi_{N-1}]^T, \tag{3.57}$$

where Ψ_n is the $(1 \times SM)$ matrix,

$$\Psi_n = [\psi(n,1,1,0), \psi(n,2,0,0), \psi(n,2,2,0), \psi(n,3,1,0), ..., \psi(n,k,r,m), ..., \psi(n,K,R_K,M-1)], \tag{3.58}$$

and,

$$\psi(n,k,r,m) = \sum_{l=0}^{L-1} h(l) e^{-\jmath\omega_c n} |\widetilde{x}(n-m-l)|^{k-r} \Re\left\{ \left(\widetilde{x}(n-m-l) e^{\jmath\omega_c (n-m-l)}\right)^r \right\}. \tag{3.59}$$

Finally, with the above equations the coefficient matrix is estimated by the LS equations given in (3.44) and (3.45).

4. Implementing the Ultra-Wideband Model

This chapter covers the requirements for implementing the ultra-wideband memory polynomial model. Considering the sampling rate limitations discussed in Chapter 3, we will investigate techniques for modeling harmonics with the lowest possible sampling rate for the specific case of DOCSIS FDX.

4.1 Calculating UWB Harmonics

As discussed in Chapter 3, maximum device clock rates limit the maximum frequency of signals that can be represented straightforwardly. Although parallel operations can be used to achieve an effective sampling rate that is a multiple of the clock rate, doing so uses proportionally more hardware and is, therefore, more expensive. The purpose of this section is to find the lowest possible sampling rates required for accurately modeling each harmonic present at the power amplifier output, thus providing recommendations for the sampling rates and implementation structures to be used when constructing the DOCSIS 3.1 UWB-MP model in hardware.

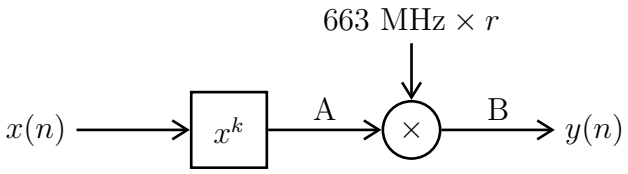
First, we must decide which harmonics are to be calculated and canceled with echo-cancellation. Referring to Figure 3.8, all harmonics up to and including the 5^{th} order interfere with the signal of interest from 5 MHz to 684 MHz. Therefore we *might* have to compensate for each of these harmonics. The power in each harmonic will vary from device to device and based on operating point, so the only way to know for sure which harmonics must be compensated for is testing the EC algorithm with an actual PA and observing how much additional cancellation is obtained when additional harmonics are included.

A good starting point is to begin by modeling the harmonics requiring the lowest sampling

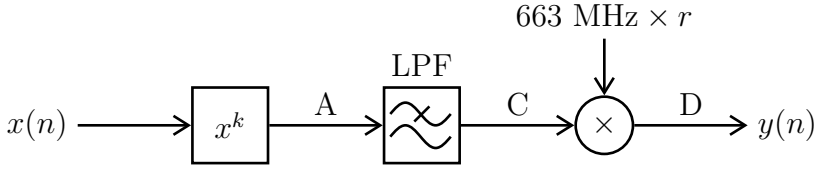
rates, and including additional harmonics as necessary until satisfactory echo-cancellation is achieved. We will start with the 1^{st}, 2^{nd}, and the 3^{rd} order harmonics, as these are traditionally the most significant harmonics in CATV amplifiers [46]. The process for extending past the 3^{rd} order will be very similar, but with higher sampling rate requirements.

To simplify the hardware implementation of the UWB-MP model, it is desirable to use a single sampling rate for all harmonics. Therefore, the strategy pursued in this section is to first determine the required sampling rate for each harmonic, then select the common sampling rate as the largest of the individual rates.

Each harmonic must be analyzed individually to find its minimum sampling requirements. Where advantageous, sampling rates will be reduced by applying the technique discussed in Section 3.4 which involves the use of complex filters. Therefore, there are two methods of calculating the harmonics at baseband: one which includes a complex filter and one which does not. Diagrams of the two options are given in Figure 4.1. The following sections analyze each harmonic using both options. We first consider the general case of a harmonic of arbitrary order and center frequency in Section 4.1.1. Then, Sections 4.1.2 and 4.1.3 apply the analysis to the specific cases of 2^{nd} and 3^{rd} order harmonics, respectively.



(a) Option 1 - no baseband filter



(b) Option 2 - with baseband filter

Figure 4.1: Two options for computing harmonics at baseband

In Figure 4.1, the x^k block represents calculating the harmonics at baseband. Recall the block diagram of the UWB-MP model in Figure 3.12. In this diagram, the harmonics were calculated at passband directly. To calculate the harmonics at baseband with optional filters, then up-convert the harmonics to their proper center frequencies, the diagram of the UWB-MP model would take the form presented in Figure 4.2.

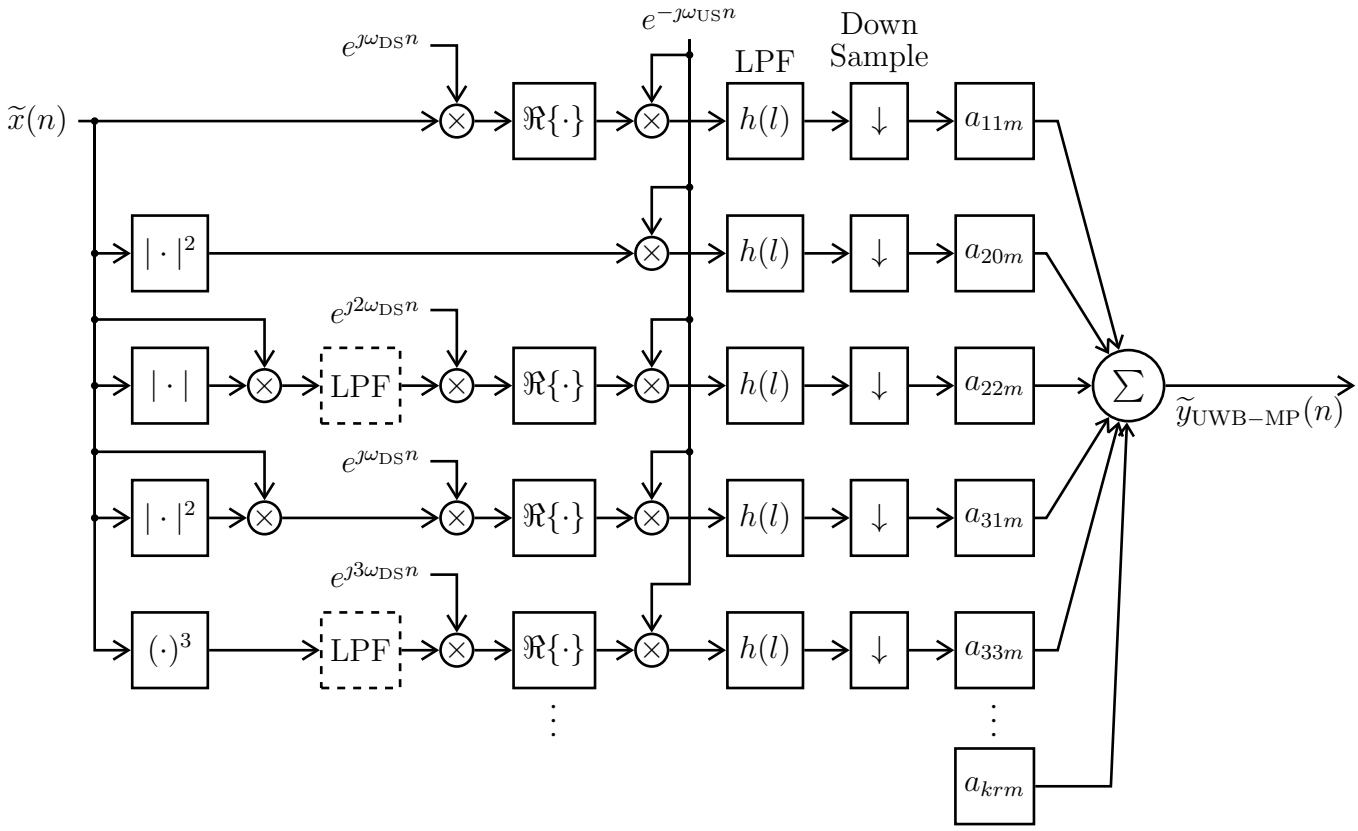


Figure 4.2: UWB-MP diagram with harmonics calculated at baseband and optional filters

4.1.1 Modeling Arbitrary Harmonics

Let each harmonic be denoted by H_{kr}, where k is the harmonic order and r is the multiple of the carrier frequency at which the harmonic is centered. For example, H_{20} is the 2^{nd} harmonic at DC, and H_{33} is the 3^{rd} order harmonic at $3f_c$. Recall that the downstream signal is centered at $f_c = (1218 + 108)/2 = 663$ MHz, and has a bandwidth of $BW = 1218 - 108 = 1110$ MHz. The center frequencies and bandwidths of each harmonic

up to the 5^{th} order are summarized by Table 4.1, where all values are in units of MHz.

Harmonic	BW	Center	Lower Edge	Upper Edge
H_{20}	2220	0	-1110	1110
H_{22}	2220	1326	216	2436
H_{31}	3330	663	-1002	2328
H_{33}	3330	1989	324	3654
H_{40}	4440	0	-2220	2220
H_{42}	4440	1326	-894	3546
H_{44}	4440	2652	432	4872
H_{51}	5550	663	-2112	3438
H_{53}	5550	1989	-786	4764
H_{55}	5550	3315	540	6090

Table 4.1: Harmonic spectrum centers and spans (all frequencies in MHz)

Before calculating the passband harmonics in Table 4.1, they can first be calculated at baseband in our model. The method for determining whether or not a baseband complex filter is beneficial (option 2) is described here. The specific cases of calculating the 2^{nd} and 3^{rd} order harmonics are discussed as examples in Sections 4.1.2 and 4.1.3.

Recall the spectrum of the downstream signal $x(n)$ at baseband spans from -555 to 555 MHz, and the US spectrum at baseband (which is the signal of interest - SOI) spans from -339.5 to 339.5 MHz. At passband, the US signal spans the frequency range from 5 to 684 MHz. However, since the passband US signal is real, it also occupies the negative frequencies from -684 to -5 MHz. Correctly modeling the effect of the passband harmonics created by the PA requires an accurate representation of the entire US regions at passband. Thus, there can be no aliasing into the region spanning -684 to 684 MHz.

Points A, B, C, and D in the following discussion will be referring to Figure 4.1. At point A, the spectrum spans from $-k \times 555$ to $k \times 555$ MHz, and at point B the spectrum spans from $-k \times 555 + 663 \times r$ to $k \times 555 + 663 \times r$ MHz.

At point B, aliasing into the US region will occur if

$$684 < F_s - k \times 555 - r \times 663, \tag{4.1}$$

where all values are in units of MHz. Therefore, for no aliasing to occur in the US region at point B,

$$F_s > 684 + k \times 555 + r \times 663. \tag{4.2}$$

For no aliasing of the DS signal on top of itself at point A,

$$F_s > 2k \times 555. \tag{4.3}$$

The spectrum at point C depends on the LPF. The filter should be designed to allow the portion of the signal that will overlap with the SOI to pass and remove as much of the rest as possible. The portion of the signal that will overlap with the SOI is the part that will be in the -684 to 684 MHz region after up-conversion. Let us define the baseband equivalent of this region as the set of frequencies Q:

$$Q = \{\, F \mid -684 - r \times 663 < F < 684 - r \times 663 \,\} \tag{4.4}$$

We will also define the set of frequencies S that is occupied by the baseband harmonic at point A:

$$S = \{\, F \mid -k \times 555 < F < k \times 555 \,\} \tag{4.5}$$

The output of the baseband filter should be the portion of the signal which falls within Q. Therefore the baseband filter should have a passband which is the intersection of the sets Q and S, i.e. $Q \bigcap S$. The intersection is not empty (sets overlap) if

$$\max(-k \times 555, -684 - r \times 663) < \min(k \times 555, 684 - r \times 663). \tag{4.6}$$

If the above inequality is true, then the intersection of the sets spans from

$$\max(-k \times 555, -684 - r \times 663) \text{ to } \min(k \times 555, 684 - r \times 663). \tag{4.7}$$

Equations (4.6) and (4.7) tell us two important details. First, (4.6) tells us if the H_{kr} harmonic interferes with the signal of interest. If so, it will contribute to the SI corrupting

our US signal. Second, (4.7) gives the frequency range of the baseband harmonic that will interfere with the SOI at passband. This region will be the passband of the filter if Option 2 from Figure 4.1 is used to calculate the harmonic.

Designing the complex LPF for option two involves a trade-off between transition bandwidth and filter length. Bellanger's rule for digital filter design [47] indicates that the transition width of the filter is inversely proportional to the length of the filter, and therefore the cost of the filter. In our specific case, using a narrow transition band allows for a lower sampling rate, but at a higher cost. Whereas a wide transition width requires a higher sampling rate but allows a lower filter cost. In the end, a common sampling rate will be chosen for calculating all harmonics, so it is reasonable to use the widest allowable transition bandwidth for each filter, given the sampling rate. The next sections apply the logic from this section in determining the minimum sampling rate required for calculating the 2^{nd} and 3^{rd} harmonics.

4.1.2 Calculating 2nd Order Harmonics

The spectrum of the 2^{nd} order harmonic at DC, H_{20}, is illustrated in Figure 4.3, with the US signal band also labeled.

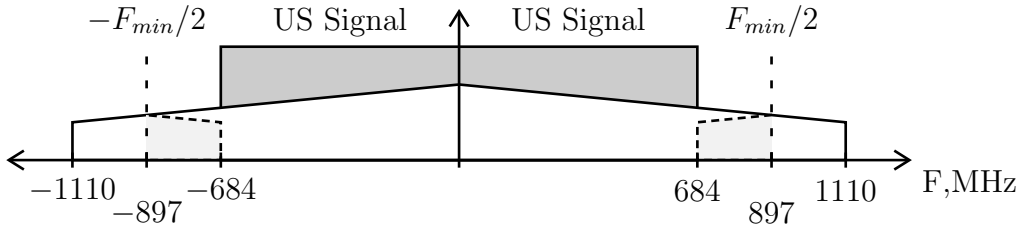


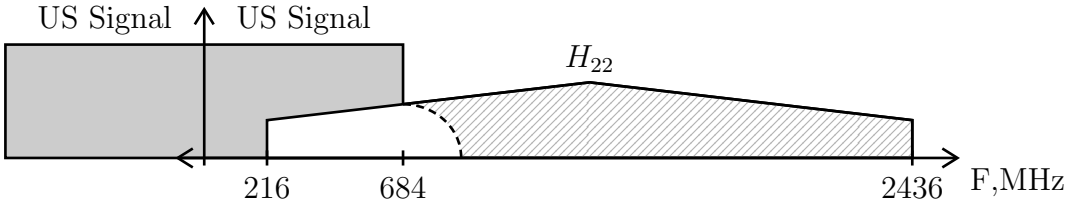
Figure 4.3: H_{20} harmonic with minimum sampling rate

Since this is a baseband signal it is not up-converted and therefore filtering will not help reduce the sampling rate. From Equation (4.2), the minimum sampling rate for calculating

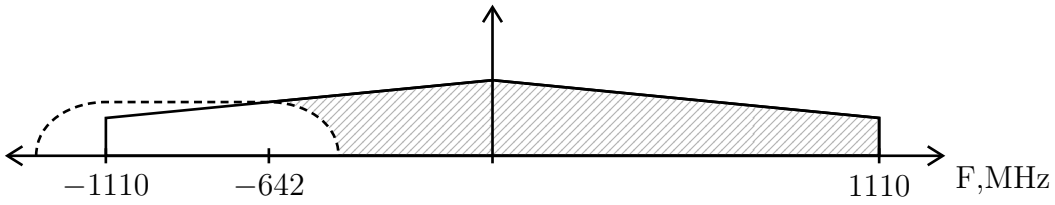
H_{20} is

$$F_{min} > 684 + 2 \times 555 + 0 \times 663$$
$$F_{min} > 1794 \tag{4.8}$$

The second order harmonic at $2f_c$, H_{22}, is first calculated at baseband then up-converted to passband. The passband and baseband spectra of H_{22} are given in Figure 4.4. Figures 4.4a and 4.4b illustrate the harmonic both with and without the use of a baseband filter. The hatched out area is the portion of the signal the filter would remove, and the dotted line represents the filter's frequency response.



(a) H_{22} harmonic at passband



(b) H_{22} at baseband prior to up-conversion

Figure 4.4: H_{22} at passband and baseband with a complex LPF

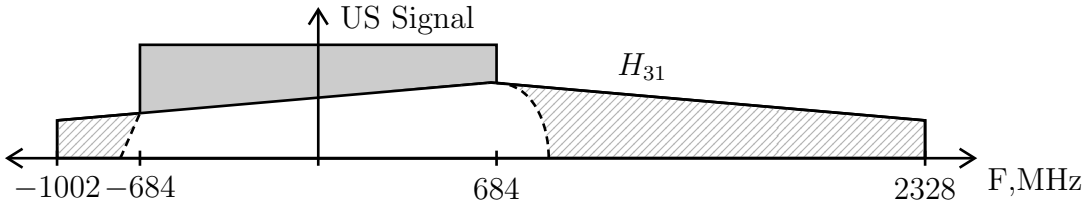
From Figure 4.4a, it is seen that only the portion of the passband signal spectrum from 216-684 MHz interferes with the US signal. Therefore, only this portion of the spectrum needs to be captured by the model. The remainder can be filtered away at baseband to reduce the sampling rate requirements. If filtered at baseband, the minimum sampling rate required to represent H_{22} at passband is $2 \times 684 = 1368$ MHz plus the baseband filter transition width.

H_{22} at baseband is shown in Figure 4.4b. The portion of this spectrum interfering with the US signal at passband is the most negative part of the spectrum, from -642 to -1110 MHz. Thus, a complex LPF with a cutoff frequency as low as -642 MHz can be used to pre-filter the signal. However, the minimum sampling rate required to calculate this baseband signal is $2 \times 1110 = 2220$ MHz since no part of the signal can alias before the filter without corrupting the signal of interest. This means that the minimum sampling rate for the calculation of H_{22} is 2220 MHz if a filter is used.

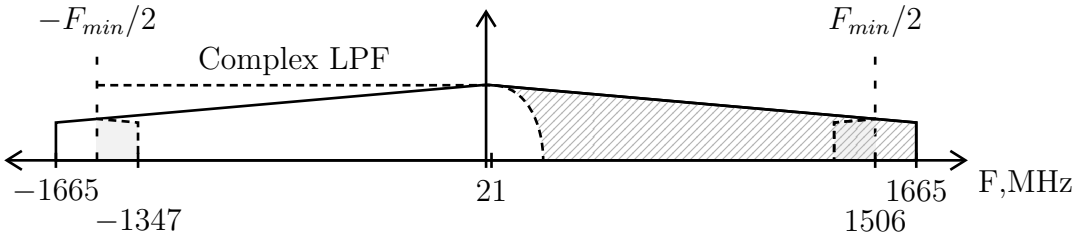
If H_{22} is not pre-filtered, then from (4.2) the lowest sampling rate at passband would be $F_{min} > 2436 + 684 = 3120$ MHz. Therefore, pre-filtering the harmonic can save up to 900 MHz in sampling rate requirements, at the expense of implementing a complex filter.

4.1.3 Calculating 3rd Order Harmonics

The spectrum of the 3^{rd} order harmonic centered at f_c, H_{31}, is given in Figure 4.5a. From Figure 4.5a, frequencies that will be above the US signal band at passband can be



(a) H_{31} harmonic at passband



(b) H_{31} at baseband prior to up-conversion with a complex filter

Figure 4.5: H_{31} at passband and baseband with a complex LPF

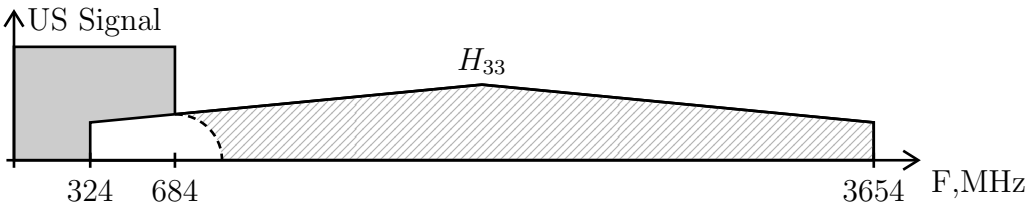
filtered away at baseband to reduce the sampling rate requirement. Also, frequencies that

will fall below the negative US signal band at passband may be allowed to alias up to the point where the aliased signal reaches -684 MHz. If H_{31} is filtered at baseband, then the minimum sampling rate at passband is $2 \times 684 = 1368$ MHz plus the filter transition width.

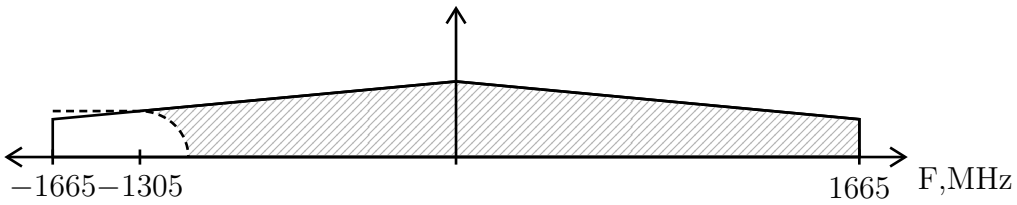
The baseband complex filter and aliasing is illustrated in Figure 4.5b. The -1347 to 11 MHz band at baseband corresponds to the US signal band at passband. The minimum sampling rate at baseband required to calculate the signal of interest from -1347 to 11 MHz is the rate where the alias interferes up to -1347 MHz. This rate is calculated by $F_{min} > 1665 + 1347 = 3012$ MHz.

If H_{31} is not pre-filtered, then from (4.2) the minimum sampling rate would be $F_{min} > 684 + 3 \times 555 + 1 \times 663 = 3012$ MHz. In this case, pre-filtering does not reduce sampling rate requirements.

H_{33} is calculated in a similar fashion, with the signal's spectrum given in Figure 4.6. From Figure 4.6a, the passband spectrum from 324 to 684 MHz must be calculated as it



(a) H_{33} harmonic pre-filtered



(b) H_{33} prior to up-conversion with a complex filter

Figure 4.6: H_{33} at passband and baseband with a complex LPF

interferes with the US signal, and the remaining portion can be filtered away at complex baseband. Thus, the minimum sampling rate required to represent H_{33} at passband with

pre-filtering at baseband is $2 \times 684 = 1368$ MHz plus the filter roll-off.

H_{33} at baseband is given in Figure 4.6b. As in the case of H_{22}, the most negative part of the spectrum from -1305 to -1665 MHz at baseband interferes with the US signal at passband. Thus, a complex LPF with a cutoff frequency as low as -1305 MHz can be used to pre-filter the signal. The minimum sampling rate required to calculate this baseband signal before the filter is $2 \times 1665 = 3330$ MHz since we cannot allow aliasing to occur before the filter.

If H_{33} is not pre-filtered, then from (4.2) the lowest sampling rate at passband would be $F_{min} > 684 + 3 \times 555 + 3 \times 663 = 4338$ MHz. Thus, pre-filtering the harmonic can save 1008 MHz in sampling rate requirements.

Calculating higher order harmonics follows the same process as the calculations for the 2^{nd} and 3^{rd} order harmonics.

4.1.4 Summary of Calculations

A summary of the required sampling rates for calculating the harmonics discussed in Section 4.1 is presented here.

Table 4.2 gives the minimum sampling rate in MHz required for calculating each harmonic, both with and without a baseband filter. These sampling rates do not account for the filter transition width, so the sampling rate used will have to be adequately higher than the listed rate to allow for a reasonable filter transition band.

Harmonic	F_{min} No Filter	F_{min} With Filter
H_{20}	1794	1794
H_{22}	3120	2220
H_{31}	3012	3012
H_{33}	4338	3330

Table 4.2: Summary of minimum sampling rates for calculating harmonics

Upon examining the minimum required sampling rates, it is reasonable to conclude that we should calculate H_{20}, H_{22}, and H_{31} without a filter since H_{33} requires the highest sampling rate. H_{33}, however, should be calculated using a baseband filter since this signal is the limiting factor for the system.

If we implement a complex baseband filter when calculating H_{33}, the filter transition width will equal the increase in sampling rate frequency required. For example, if the filter's transition width is 200 MHz, then the minimum sampling rate would be $3330 + 200 = 3530$ MHz. A mathematical optimization problem could be set up here to minimize hardware cost based on sampling rate and filter length, but that is left as future work.

4.2 Down-Converting Harmonics

The passband harmonics calculated in the previous section model the self-interference signal leaking into the upstream signal path. These modeled signals must follow the same operations done on the actual received signal as they are down-converted to baseband. Following the same operations as the actual received signal, the modeled baseband signals should match the received baseband signals, allowing echo-cancellation to take place.

The actual received signal goes through an analog low pass filter before the ADC to band-limit the signal to prevent aliasing. In the DOCSIS FDX case, the analog LPF will likely have a passband corner at 684 MHz such that the upstream signal is preserved while eliminating out-of-band power that could otherwise saturate the ADC. This spectrum with the LPF is illustrated in Figure 4.7.

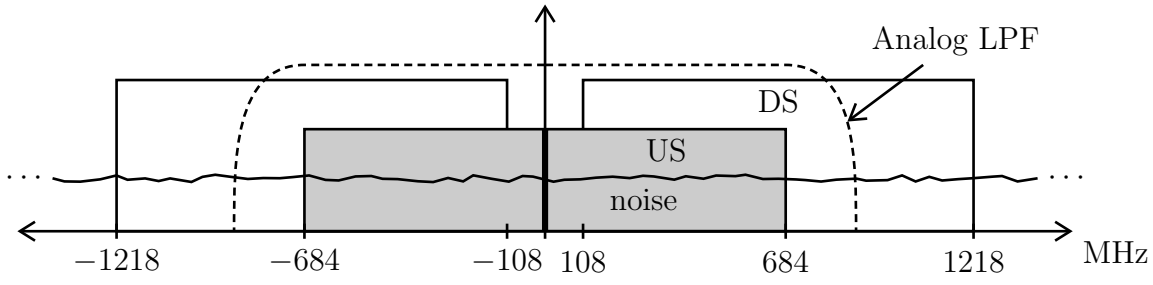


Figure 4.7: Received signal with analog LPF

After the filter, the signal is sampled by the ADC and down-converted to baseband. Following the down-conversion, the double-frequency component is filtered away as discussed in Section 2.2. The complex baseband result of the down-conversion and filtering is illustrated in Figure 4.8. After the down-conversion and filtering, the signal is down-sampled before demodulation.

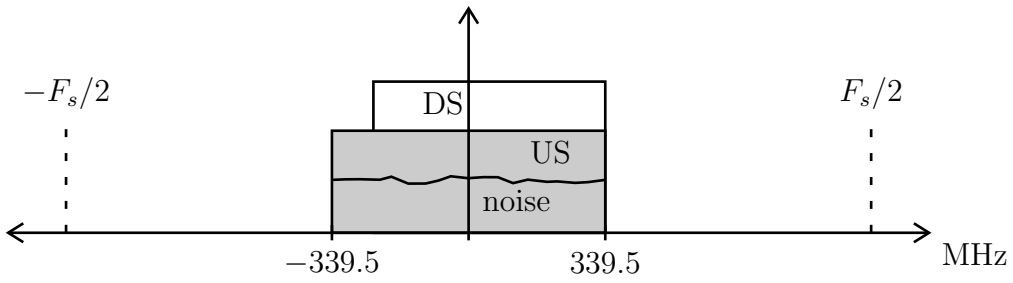


Figure 4.8: Complex down-converted received signal

Note that the analog LPF does not necessarily have to be modeled for our echo-cancellation system. Consider a case where we do not apply a filter in our model corresponding to the analog LPF seen in the actual system. The system would proceed to down-convert the output of the PA, including the high-frequency harmonics. After the down-conversion, the signal would be digitally filtered. This digital filter would remove the out-of-band harmonics that would have been removed by the analog LPF seen in the actual system.

Therefore we do not need to model the analog filter. The trade-off here is that when the signal is down-converted, the power in the high-frequency harmonics aliases and adds power to the spectrum that is being filtered by the baseband digital LPF after the down-conversion. There would be more remaining OOB power at the input of this filter if it was not removed by an equivalent analog LPF. Therefore the digital LPF must have a better out-of-band attenuation compared to the case where we model the analog filter.

Finally, after the modeled harmonics are represented at baseband, they can be down-sampled and used in conjunction with the least squares algorithm to estimate the UWB-MP model coefficients.

4.3 Echo-Cancellation

Using the baseband down-sampled modeled harmonics developed in Section 4.2, along with the least squares algorithm for computing coefficients as discussed in Section 3.6, echo-cancellation can take place. Assuming that the amplifier characteristics do not significantly change over time and temperature, the LS calculations may be performed offline. Optionally, the computational workload of the FPGA can be further reduced be offloading these calculations to a processor. In such an implementation, the processor calculates the coefficients for the UWB-PM model and sends them back into the FPGA to be applied to the modeled baseband signals. The result is subtracted from the received signal, hence performing echo-cancellation.

The coefficients are repeatedly updated until the error power is below a threshold, at which point the coefficients can stop updating until the PA response changes substantially. Such a change would be detected via a significant increase in the error power. An illustration of the completed echo-cancellation operation is provided in Figure 4.9.

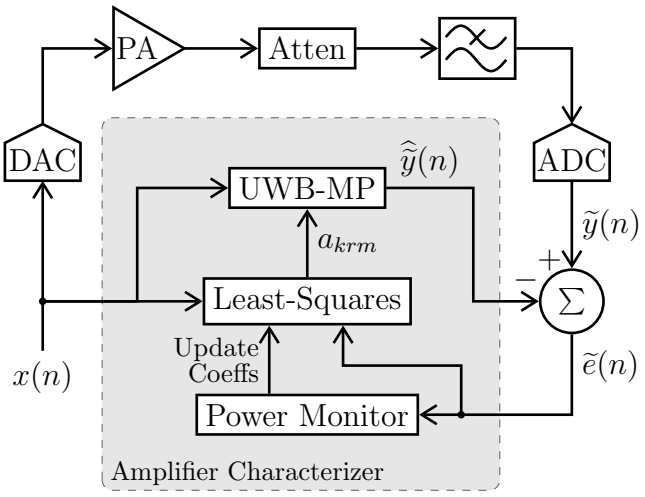


Figure 4.9: Echo cancellation block diagram

Of course, the system in Figure 4.9 can only be used to characterize the amplifier. To perform full echo-cancellation, the FDX echo channel response must first be estimated. The estimated PA output from the UWB-MP model can then be convolved with the (linear)

FDX channel estimate to produce an estimate of the self-interference seen at the US receiver. Finally, this SI estimate is subtracted from the signal at the US receiver. In the ideal case, where all of the estimates are perfect, only the US signals transmitted from cable modems would remain after the cancellation. Implementing the FDX channel estimation is out of the scope of this thesis, however, Figure 4.10 provides a recommended structure derived from the narrowband EC structure presented in [35].

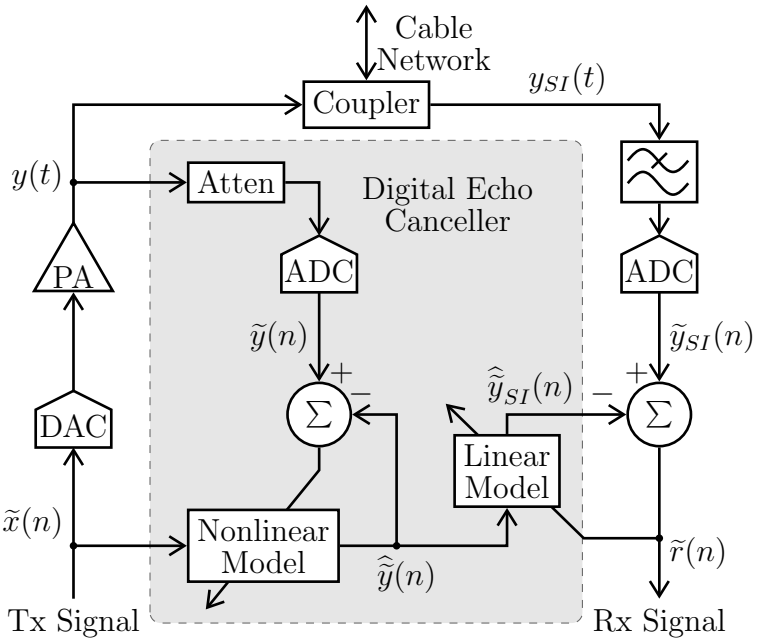


Figure 4.10: Full echo-cancellation system diagram

The echo-canceler given in Figure 4.10 requires only a single feedback ADC for EC purposes, regardless of the number of amplifiers in the node because as soon as the nonlinear model converges to sufficiently characterize the amplifier, that model can stop updating, and the ADC can be used to characterize another PA. When all PAs have been characterized, the ADC can be shutdown until an amplifier response changes.

5. Results

This chapter tests the performance of the ultra-wideband memory polynomial (UWB-MP) echo-cancellation model. The chapter is organized as follows. First, the simulation procedure is described, and a diagram of the simulation model is given. Next, the performance metrics used to compare results are developed. Lastly, simulations are demonstrated, and the results are discussed.

Since no other publicly available[1] echo-cancellation algorithms exist for our particular problem, we cannot compare our newly developed algorithm against others. Instead, we will compare the performance of the newly developed UWB-MP model against a linear model. Doing so will justify the cost and complexity of the UWB-MP model by showing the performance difference compared to using only a linear model, which is unable to predict or represent the harmonics generated by the PA.

Further simulations will compare the impact on cancellation performance of changing the number of coefficients used in the UWB-MP model, the number of samples used with the LS algorithm, and the number of iterations of the LS algorithm. We will also investigate the change in performance based on the power amplifier simulation model used.

The simulations in this chapter serve multiple purposes. First, the simulations verify the mathematical representation of the UWB-MP derived in Chapter 3. Additionally, we will justify the increase in cost and complexity associated with canceling harmonics with the UWB-MP model opposed to using a simple model only able to cancel the linear signal.

[1]While a number of companies in the cable/DOCSIS industry have developed or are working to develop ultra-wideband models for this application, these models are proprietary. To the best of our knowledge, no suitable models have been disclosed in the open literature.

Lastly, the results of the simulations tell us if the amount of cancellation required to support DOCSIS FDX is possible. The simulations are not intended to find an optimal hardware implementation of the algorithm. Chapter 6 draws conclusions from the simulations and lays out potential future work on the topic of DOCSIS full-duplex echo-cancellation.

5.1 Simulation Procedure

This section describes the simulations of the ultra-wideband power amplifier characterization model in Matlab. The downstream signal is simulated with a band-limited AWGN signal. Such signals have been shown to be reasonable approximations of OFDM signals [48]. The bandwidth of the signal will closely represent the frequencies used in DOCSIS FDX. The reason for using a band-limited AWGN signal instead of an actual OFDM signal conforming to DOCSIS 3.1 is for simplicity. An actual DOCSIS signal consists of many channels and different OFDM blocks [6]. Simulating such a signal would make the simulations less configurable, slower to run, and add unnecessary complexity.

The PA models chosen take the same structures as those found in the literature. However, most PA models used in research only account for in-band harmonics for narrow-band signals, and the coefficients are often given for a baseband equivalent power amplifier. In our simulations, we use a passband PA model with an UWB signal. Therefore, the coefficients of the PA models in literature are not applicable for our simulations. Instead, we will use PA model architectures found in literature, but with our coefficients at passband.

The coefficients are chosen such that the power in the harmonics corresponds to the power we expect to see in the output of an actual DOCSIS PA. From the DOCSIS specification, the in-band distortion of the DS signal in the spectrum below 684 MHz must have power at least 39 dB down from the carrier power, i.e., 39 dBc [7]. Referencing the datasheet of the Qorvo RFPD3580 power amplifier, which is a DOCSIS compliant amplifier, we see that the worst case MER for the input power discussed in Chapter 1 is about 40 dB [49]. Thus, in our simulations, we adjust the coefficients such that the total power in the harmonic distortion is about 40 dBc.

Further, the Qorvo datasheet indicates that the power in the second and third order distortions is roughly equal. Therefore in our PA models, we will make the total power in each of the second and third order harmonics the same.

5.1.1 Simulation Model

Figure 5.1 provides a diagram of the Matlab simulation used in the following section.

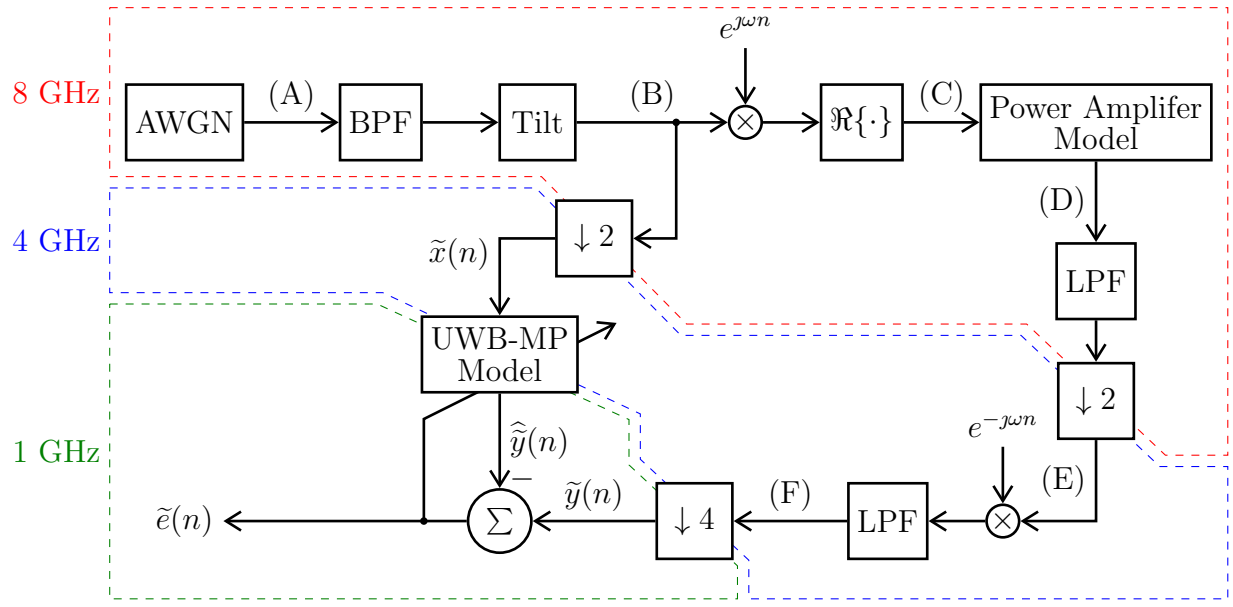


Figure 5.1: Matlab simulation model

The simulation of the downstream signal operates as follows. The DS signal is generated at complex baseband with AWGN generators creating the real and imaginary signals (point A).

This Tx baseband signal is filtered to band-limit it from -555 to 555 MHz. This band-limited signal then has a 21 dB tilt applied to it (point B) using an FIR filter to correspond to the tilt applied to the DOCSIS DS signal, as discussed in Chapter 1. Although in an actual FDX system the filter which produces the tilt will have to be adapted to since it is applied in the analog domain, we assume here that the coefficients used to create the tilt are already known. This assumption is realistic because, in a real system, the response of the

tilt filter could be measured during manufacturing. Therefore, estimating these coefficients is out of the scope of this project.

The output of the tilt filter is then up-converted to passband, where the signal's frequency spectrum spans from 108 to 1218 MHz (point C). This represents a transmitted signal in the DOCSIS FDX system. The signal is put through a power amplifier model, which includes memory and nonlinear terms (point D). This operation happens at a high enough sampling rate to capture all of the high-frequency harmonics without aliasing to accurately represent what the signal would experience in the analog domain.

The PA output is passed through a LPF which represents the analog antialiasing filter the signal would see before the ADC in an actual US demodulator. The output of the filter is down-sampled to the system sampling rate (point E), down-converted to complex baseband, then filtered to remove the double frequency component (point F). Finally, the signal is down-sampled again to produce $\widetilde{y}(n)$ which can be used with the UWB-MP model and the LS algorithm to characterize the simulated power amplifier model. The signal $\widetilde{y}(n)$ represents the self-interference present in the FDX node.

In parallel, the simulation applies the UWB-MP model to estimate the self-interference. Harmonics up to and including the 3^{rd} order are calculated in the simulations as discussed in Chapter 4. The process for including higher order harmonics would be the same, but a higher sampling rate would be required. The input to the UWB-MP model is $\widetilde{x}(n)$ from Figure 5.1, which is the down-sampled output of the tilt filter. A down-sampler is used because the data is generated at a higher rate for the power amplifier model. However, the digital baseband system runs at a lower rate to reduce the hardware requirements. This is consistent with the expected implementation in an actual system.

The output of the UWB-MP model $\widehat{\widetilde{y}}(n)$ is the approximation of the self-interference signal $\widetilde{y}(n)$, and is subtracted from $\widetilde{y}(n)$ to produce the error signal $\widetilde{e}(n)$. The coefficients used in the UWB-MP model to calculate $\widehat{\widetilde{y}}(n)$ are computed using the LS algorithm, as described in Section 3.6

The sampling rate used to calculate the harmonics in Matlab was chosen to be 4 GHz.

This rate is used for a number of reasons. First, referring back to Table 4.2, we found that the minimum sampling rate required to calculate the 1^{st}, 2^{nd}, and 3^{rd} order harmonics was dictated by the calculation for the 3^{rd} order harmonic at $3f_c$, referred to as H_{33}. To first calculate this harmonic at baseband, the minimum required sampling rate was 3330 MHz. To calculate H_{33} at passband without first filtering it at baseband, the minimum required sampling rate was 4338 MHz.

However, if H_{33} is filtered at baseband before up-converting, the minimum required sampling rate was reduced to 3330 MHz. Thus, for the simulations, we choose to use a 4 GHz sampling rate and apply a complex filter to H_{33} at baseband before up-conversion. Therefore we can calculate all of the 1^{st}, 2^{nd}, and 3^{rd} order harmonics without the aliases corrupting the signal of interest. A slightly lower sampling rate could be used at the expense of using a larger filter on H_{33}. However, this trade-off will not be analyzed here.

As previously mentioned, the input signal is processed at a higher sampling rate so the PA can be accurately modeled. This rate was chosen to be 8 GHz for two reasons. First, this rate allows the PA output to be modeled without aliasing. Recall from Chapter 4 that the highest frequency harmonic H_{33} harmonic extends to 3654 MHz. Therefore, the chosen sampling rate of 8 GHz can capture this signal entirely without aliasing. Second, a rate of 8 GHz allows the signal to be matched to the system sampling rate of 4 GHz through a simple down-sample by 2 operation.

After the harmonics are down-converted to baseband, the sampling rate is reduced by a factor of 4, which results in a sampling rate of 1000 MHz. This sampling rate is theoretically capable of handling signals with frequencies as high as 500 MHz, which makes it sufficient for calculating the baseband US signal approximation which has frequencies as high as 339.5 MHz.

Although there are other combinations of sampling rate and down-sampling factors that would work, we found the values chosen to be realistic regarding the hardware implementation of the model. Furthermore, they provide a reasonable trade-off between performance and resource usage.

5.2 Performance Metrics

To quantify the performance of an echo-cancellation algorithm, the normalized mean square error (NMSE) is commonly used, and is defined as [37]:

$$\text{NMSE} = \frac{\sum_n |\widetilde{y}(n) - \widehat{\widetilde{y}}(n)|^2}{\sum_n |\widetilde{y}(n)|^2}. \tag{5.1}$$

where $\widetilde{y}(n)$ and $\widehat{\widetilde{y}}(n)$ are the measured and approximated SI signals respectively from Figure 5.1.

The NMSE is useful for quantifying the cancellation over the entire frequency band from $-\pi$ to π rad/sample. However, for FDX EC, we are more concerned with the cancellation in the upstream frequency band (from -339.5 to 339.5 MHz after down-conversion), as any power remaining out of band can be filtered away. The signal-to-error power in the US signal band can be determined by taking the ratio of the signal power in the US signal band with the error power in the same band.

The signal power in the US band is found by calculating the DFT of the measured signal, and summing together the powers in the frequency components of the US band. The total power of a signal $x(n)$ over a range of discrete frequencies is calculated by

$$\sum_{k=k_{\text{start}}}^{k_{\text{end}}} \left| \sum_{n=0}^{N-1} x(n) e^{-\jmath \frac{2\pi k}{N} n} \right|^2, \tag{5.2}$$

where k is an integer between 0 and $N-1$, and N is the total number of samples in the signal. The indices k_{start} and k_{end} are chosen to define the range of frequencies over which the power is to be summed. This calculation is accurate as long as the frequency step size $(2\pi/N)$ is sufficiently small, i.e. N is sufficiently large. Finally, the total cancellation of the self-interference signal will be given in log scale, calculated by

$$\text{Total in-band cancellation} = 10\log_{10}\left(\frac{\text{total in-band signal power}}{\text{total in-band error power}}\right). \tag{5.3}$$

5.3 Simulations

The following parameters will be held constant through all of the simulations. The DS signal will occupy the spectrum from 108 to 1218 MHz, and the US signal will occupy 5 to 684

MHz. Echo-cancellation will take place at complex baseband. The results of each simulation show the received self-interference signal, the resulting error signal after the cancellation, and the total in-band cancellation.

In all simulations, nonlinear terms up to and including the 3^{rd} order are included in both the PA simulation model and the UWB-MP cancellation model. The first simulation changes the number of memory terms used in the UWB-MP model, the number of data points included in the LS algorithm, and the number of LS iterations to compare the impact of these parameters upon the cancellation performance. The remainder of the simulations uses a constant set of parameters to compare the performance of the UWB-MP model against different power amplifier models. The least-squares step factor is set to $\mu = 0.1$. Increasing μ typically leads to a faster convergence (if the coefficients don't diverge), but will have more error when they are converged. Our tests indicate that $\mu = 0.1$ is a good compromise between these conflicting goals. Ideally μ would start larger then decrease as the coefficients converge. However, for simplicity we keep μ fixed.

5.3.1 Simulation 1 - Memory Polynomial PA Model

First, we will simulate the power amplifier using a passband memory polynomial model derived from [20]. In [20], the model was used at baseband and included only odd-order in-band harmonics. Instead, we will implement the model at passband with both 2^{nd} and 3^{rd} order harmonics, as would be experienced with an actual DOCSIS power amplifier.

The model takes the form:

$$y(n) = \sum_{k=1}^{K} \sum_{q=0}^{Q-1} c_{kq} z(n-q)^{k-1}. \tag{5.4}$$

The coefficients are chosen such that the PA has a linear gain of 1 ($c_{10} = 1$), followed by random memory terms. The memory terms for the 2^{nd} and 3^{rd} order harmonic are also randomly chosen. The criteria for selecting the noise terms is such that the total power in the 2^{nd} and 3^{rd} order harmonics is about 40 dB down from the total power in the DS signal. This difference in power is observed by plotting the power spectrums of the linear signal and harmonics. The total number of memory terms for each harmonic is chosen to be

$Q = 10$. In Matlab, the memory terms are chosen by $0.01 \times \text{randn}(1, 10)$, where randn(1,10) is 10 samples chosen independently from a Gaussian distribution and 0.01 is the variance. Note that this is a pessimistic model, as the memory terms are likely to be correlated in a practical system [50]. The resulting c_{kq} for this simulation is:

$$\begin{array}{lll}
c_{10} = 1 & c_{20} = -0.0156 & c_{30} = -0.0125 \\
c_{11} = 0.0064 & c_{21} = 0.0167 & c_{31} = -0.0054 \\
c_{12} = -0.0021 & c_{22} = 0.0047 & c_{32} = -0.0071 \\
c_{13} = 0.0013 & c_{23} = 0.0097 & c_{33} = 0.0105 \\
c_{14} = 0.0048 & c_{24} = -0.0068 & c_{34} = 0.0018 \\
c_{15} = 0.0069 & c_{25} = -0.0062 & c_{35} = 0.0182 \\
c_{16} = -0.0131 & c_{26} = 0.0256 & c_{36} = -0.0039 \\
c_{17} = -0.0101 & c_{27} = -0.0056 & c_{37} = 0.0012 \\
c_{18} = -0.0128 & c_{28} = 0.0029 & c_{38} = -0.0133 \\
c_{19} = -0.0006 & c_{29} = -0.0129 & c_{39} = -0.0019
\end{array} \tag{5.5}$$

The passband PA output is shown in red in Figure 5.2, with the noise from the 2^{nd} and $3rd$ harmonics plotted in blue.

Cancellation will be performed with the UWB-MP model from Equation (3.36) with memory depths of M = 10, 20, and 30. The number of LS data points used will be N = 250 and 500, and the number of iterations used will be 20, 40, and 60.

Figure 5.3 plots an example of the spectrums of the SI signal $\widetilde{y}(n)$, remaining error signal $\widetilde{e}(n)$ after the cancellation, and the extracted harmonic noise. When generating the plot, the parameters used were M=20, LS data points = 500, and number of iterations = 60. The total in-band cancellation in this example is 74.9 dB. Table 5.1 provides a summary of the total in-band cancellation for each combination of the given parameters.

Simulation 1 results can be compared against the amount of cancellation achieved using only a linear model to justify the added complexity involved with using the UWB-MP model to predict the SI signal, as opposed to a simple linear model. By setting $K = 1$ in Equation

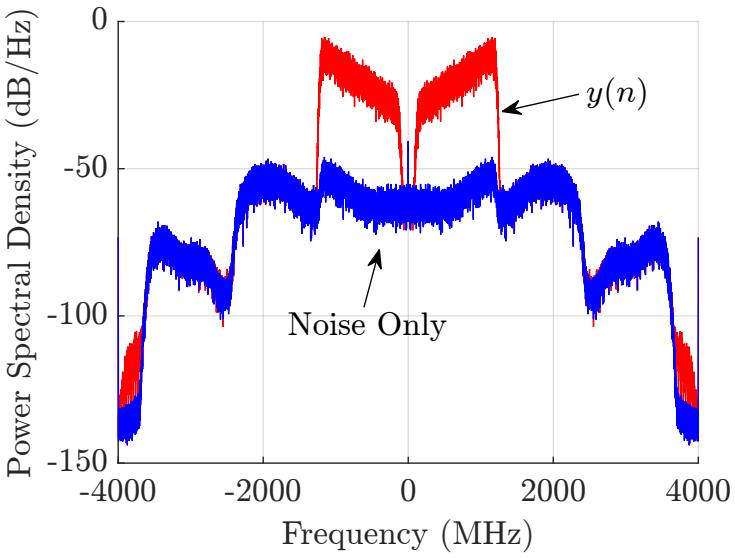


Figure 5.2: Passband PA output with noise

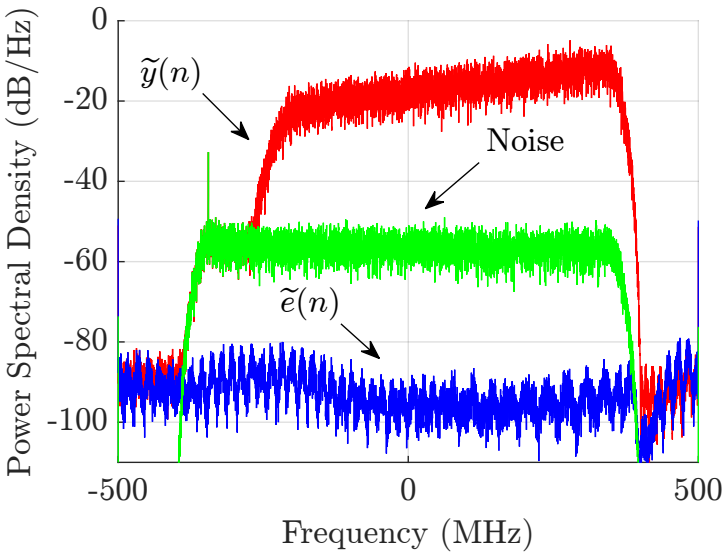


Figure 5.3: Simulation 1 results with M=20, N=500, iterations=60

(3.36), the UWB-MP model is reduced to a simple linear model. The resulting cancellation from this simplified model is 40.1 dB, with the resulting spectrum shown in Figure 5.4. Obviously, the performance of the linear model is limited by the total power in the harmonic distortion, as the linear model is unable to cancel the harmonic distortion.

To determine the limitations of the UWB-MP model, we repeated simulation 1 with

		Iterations		
M	N	20	40	60
10	250	51.1	68.7	71.7
10	500	59.7	71.0	71.7
20	250	48.6	66.3	74.5
20	500	58.1	72.8	74.9
30	250	42.8	60.9	73.5
30	500	56.2	71.9	74.9

Table 5.1: Simulation 1 total in-band cancellation (dB)

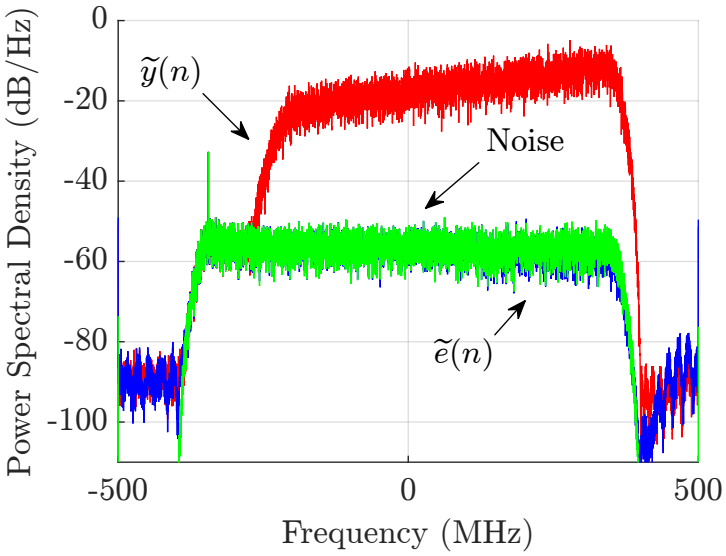


Figure 5.4: Simulation 1 results with a linear model

various numbers of coefficients in the MPM for the power amplifier. These simulations showed the cancellation performance begins degrading around 130 coefficients for each order in the PA model. Given that all known PA models require significantly fewer than 130 coefficients per path, this result considerably increases our confidence in the applicability of the UWB-MP model to practical amplifiers.

5.3.2 Simulation 2 - Memory Polynomial PA Model With AWGN

Simulation 2 uses the same power amplifier model as simulation 1, but it adds zero-mean white Gaussian noise to $\widetilde{y}(n)$ in Figure 5.1. The UWB-MP cancellation model parameters for this simulation and the remainder of the simulations will be set to M=20, N=500, and the number of iterations = 60. The power of the AWGN was set to be 60 dB less than the power in the received signal $\widetilde{y}(n)$.

Figure 5.5 gives the results of the simulation. Even with the noise floor raised due to the Gaussian noise, the algorithm still cancels the signal down to the noise floor. This indicates that the model was able to accurately estimate and model the harmonic distortions, even in the presence of AWGN. The total in-band cancellation in this simulation is 59.8 dB.

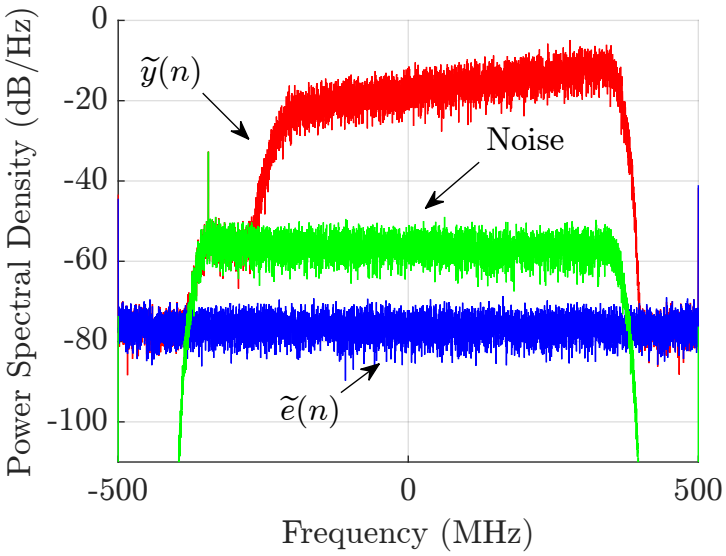


Figure 5.5: Simulation 2 results

5.3.3 Simulation 3 - Arbitrary Memory Polynomial PA Model

Simulations 1 and 2 assumed the MPM followed that of a typical PA, with the gain of the first linear coefficient (c_{10}) being large, followed by random small coefficients for the memory and harmonic terms. The simulation in this section will use the same parameters as Simulation 1, but with completely arbitrary coefficients for the PA model, meaning all of

them (including c_{10}) will be generated randomly. This means that the signal to distortion ratio at the PA output will likely be much worse than the 40 dB chosen in Simulation 1. The test will demonstrate the ability of the UWB-MP model to cancel the output of an arbitrary ultra-wideband nonlinear model, opposed to a PA-specific model. The simulation results are plotted in Figure 5.6, and the total in-band cancellation achieved is 76.3 dB. For comparison, the linear only model achieved a cancellation of only 9.3 dB. This proves that the UWB-MP model can provide significant cancellation even when the power in the harmonic terms is much larger than dictated by DOCSIS 3.1. Note that the amount of cancellation using the linear model will vary greatly from run to run, depending on the magnitude of the linear coefficients. However, the cancellation of the UWB-MP model has been observed to be consistently in the 75-77 dB range (essentially the noise floor), regardless of the PA coefficients used.

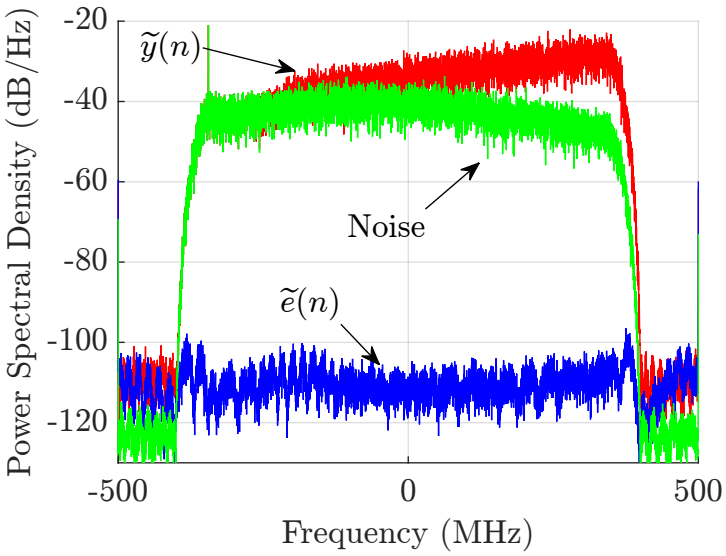


Figure 5.6: Simulation 3 results

5.3.4 Simulation 4 - Wiener-Hammerstein PA Model

Simulation 4 uses a Wiener-Hammerstein power amplifier model derived from the model presented in [20], which consists of a linear time-invariant (LTI) system, followed by a memoryless nonlinearity, followed by another LTI system. The W-H model block diagram is given

in Figure 5.7.

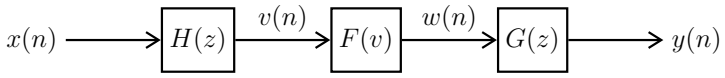


Figure 5.7: W-H block diagram

The LTI blocks in the W-H model from Figure 5.7 denoted $H(z)$ and $G(z)$, were extracted from an actual Class AB PA in [20], and are given to be

$$H(z) = \frac{1 + 0.5z^{-2}}{1 - 0.2z^{-1}}, \qquad G(z) = \frac{1 - 0.1z^{-2}}{1 - 0.4z^{-1}}. \tag{5.6}$$

The model in [20] assumed a high frequency low bandwidth signal, so the memoryless nonlinearity in their W-H model only included odd order terms for a baseband model. For our case, we will include nonlinearities of order 2 and 3 at passband. The memoryless nonlinear portion of the W-H model in Figure 5.7 is

$$w(n) = \sum_{k=1}^{K} b_k v(n)^k, \tag{5.7}$$

where $v(n)$ and $w(n)$ are respectively the input and output of the memoryless nonlinear block. The coefficients chosen based on our desired total harmonic power of 40 dB are

$$\begin{aligned} b_1 &= 1 \\ b_2 &= 0.1 \\ b_3 &= -0.1 \end{aligned} \tag{5.8}$$

The baseband Rx self-interference signal and the remaining error signal after cancellation are given in Figure 5.8. The total in-band cancellation for this simulation is 59.4 dB. In comparison, a linear model achieves 44.4 dB of cancellation.

5.4 Discussion

Through the Matlab echo-cancellation simulations, we verified the mathematical UWB-MP model by observing significant cancellation of the UWB nonlinear signals. In general,

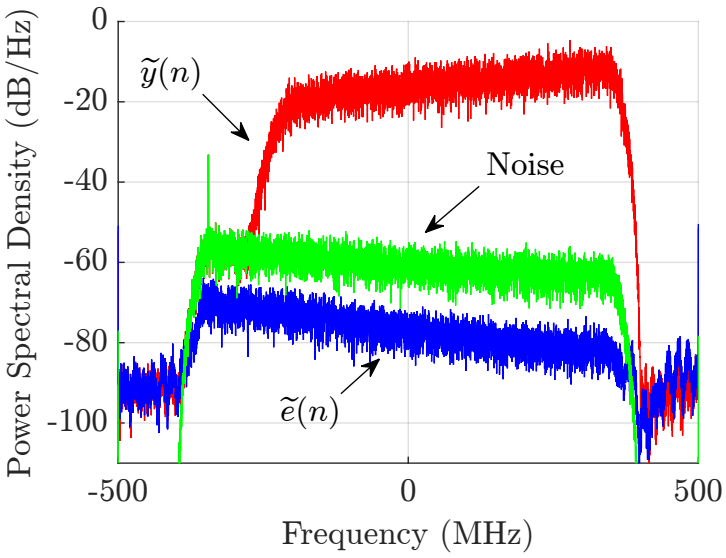


Figure 5.8: Simulation 4 results

the UWB-MP model is capable of canceling a significant amount of the self-interference signal.

The performance of the UWB-MP model is dependent on the PA model. The UWB-MP model performs exceptionally well with cancellation down to the noise floor when the PA model involves a nonlinearity followed by memory terms. The UWB-MP model does not perform quite as well if the PA model requires memory terms before the nonlinearity. The behavior is expected since the UWB-MP model extends from the MPM which assumes a nonlinear term followed by memory terms. Recall from Chapter 2 that Hammerstein models also follow this pattern. However, Wiener models assume memory terms before nonlinear terms. It is reasonable to assume that the closer the characteristics of a PA match that of an MPM or Hammerstein model, the better the cancellation will be when the UWB-MP model is used. Our simulation results confirm this expectation.

In Simulation 1, we found that the simulated UWB-MP model achieves excellent cancellation performance when the memory length of the PA model is less than 130 samples for each nonlinearity. Since the UWB-MP model used in the simulations used 20 memory samples at 1 GHz, and the PA was simulated at 8 GHz, the theoretical maximum number of memory samples that could be accounted for is $20 \times 8 = 160$. The reason that performance

began degrading before 160 samples is likely because the UWB-MP used an approximation of the analog LPF instead of using a LS algorithm or similar method to adapt the coefficients more accurately to the filter.

This observation highlights the necessity to perform the cancellation at the lowest sampling rate possible. Naturally, running at a lower sampling rate uses fewer resources on an FPGA for a given UWB-MP length. An additional benefit is that running the UWB-MP model at a lower speed would require fewer coefficients to predict the SI signal from a PA of fixed memory length. With fewer coefficients, the model is both cheaper to implement in the cancellation stage and faster to converge during the adaptation stage.

Simulation 2 demonstrates that injecting AWGN does not hinder the ability of the UWB-MP and least-squares adaptation to characterize the power amplifier. As seen in the simulation results, the EC system can characterize and cancel the self-interference signal down to the level of the AWGN, which would be the noise floor in most applications.

The results from the simulation indicate that sufficient cancellation for FDX communication may be feasible, since the cancellation obtained in simulation 1 was nearly 75 dB, whereas the necessary cancellation for FDX as stated in Section 1.7, is 60 dB. However, the performance of the UWB-MP model for modeling an actual DOCSIS power amplifier can only be determined by testing with the actual amplifier. Furthermore, a practical FDX node will have many additional sources of noise and hardware imperfections that could limit the cancellation performance. For example, the ADCs and DACs in the system add quantization noise. Only when testing with a full FDX node is conducted will it be known if the UWB-MP model is capable of providing sufficient cancellation to allow high data-rate full-duplex communication.

6. Conclusions

6.1 Summary

Demand for higher data rates from cable providers requires more bandwidth through HFC networks. The DOCSIS 3.1 standard has been updated by CableLabs to support full-duplex operation which increases the upstream and downstream signal bandwidths by simultaneously sharing part of the spectrum.

Implementing FDX in a DOCSIS node involves significant challenges that must be addressed. A primary challenge includes eliminating the self-interference from the downstream signal that leaks into the upstream signal through the coupler in the node. A significant source of the SI signal is the power amplifier, which tends to product nonlinear distortions when the output power level is high. Many echo-cancellation algorithms exist for canceling SI signals in other FDX systems. However, the case of full duplex DOCSIS is particularly challenging because the bandwidth of the downstream signal is extremely wide (over 1 GHz).

Although algorithms have been developed to characterize power amplifiers to predict nonlinear SI signals, these algorithms are typically for the case of high frequency, narrowband signals in which only odd-order harmonics centered at the carrier frequency interfere with the signal of interest. In the DOCSIS FDX case, the downstream signal is ultra-wideband, which results in many orders and center frequencies of harmonics interfering with the signal of interest.

This thesis provided an echo-cancellation algorithm for DOCSIS FDX communication involving ultra-wideband signals. Further, suggestions were given to reduce the implementation cost of this algorithm on an FPGA in a DOCSIS node. Performance analysis of the

new algorithm was done using Matlab simulations.

6.2 Contributions

The primary contribution of this work is the development of the ultra-wideband memory polynomial model to be used for echo-cancellation in a DOCSIS FDX node. The UWB-MP model is an extension of the well-known memory polynomial model and is unique from other models in the literature in that it characterizes harmonic distortions from power amplifiers with ultra-wideband signals. The model accepts a baseband copy of the outgoing downstream signal and computes estimates of the harmonic components that will be generated by the power amplifier. The model contains a set of coefficients which characterize the amplifier response. We have shown how a least-squares algorithm can be used to find a suitable set of coefficients for representing a particular amplifier.

Further, implementation guidelines for reducing hardware costs and complexity were given. These guidelines are an essential bridge between the simulation model and the design of a hardware implementation of the model. The trade-off between hardware implementation complexity and model performance was investigated through simulations, and recommendations for the size and structure of the model were produced. Analysis showed that these guidelines could significantly reduce the hardware implementation costs of the model.

6.3 Results and Conclusions

The UWB-MP echo-cancellation model was tested by simulating an FDX node in Matlab and observing the amount of cancellation of the self-interference signal. Since the problem of ultra-wideband EC has not been thoroughly studied in the literature, there are no benchmark algorithms to compare our performance or resource usage with. Instead, the UWB-MP model was compared against a linear model which can only compensate for 1^{st} order harmonics. Doing so justifies the cost and complexity of using the UWB-MP model by illustrating the performance difference compared to the linear model, which would be the default implementation for a cable node.

The simulation of the FDX node involved various system settings and different power amplifier simulation models. We found that with memory polynomial type amplifier models, the in-band cancellation reached 75 dB (approximately down to the noise floor). With Wiener type amplifier models, the in-band cancellation achieved about 59 dB. Injecting an AWGN signal into the system did not affect the performance of the UWB-MP cancellation algorithm. However, the cancellation was limited to the new noise floor created by the AWGN signal.

The amount of cancellation achieved by the UWB-MP model is not limited by the power in the harmonic distortions. The simulations demonstrated that the algorithm could sufficiently cancel harmonics even at much higher levels than the -40 dBc expected in a DOCSIS system. Further, the performance did not begin to degrade until the number of memory samples in the PA model exceeded 130 samples. This length is much longer than one would expect for an actual PA, which means our algorithm should be able to handle the memory effects associated with most power amplifiers.

From the results of the simulations, we conclude that sufficient cancellation of the SI signal for FDX communication may be possible. In addition, we can also conclude that the cancellation can take place at baseband, which significantly reduces resource requirements when implementing the model. The maximum cancellation achieved in the simulations was 75 dB, whereas the necessary cancellation for FDX to run at the desired data rate is around 60 dB. However, the cancellation performance in an actual node with a real power amplifier can only be determined by hardware testing. Furthermore, a practical FDX node will have additional noise sources that could limit cancellation performance, such as quantization noise in the ADC and DAC. Only after testing with an actual FDX node takes place will we know if sufficient cancellation using the UWB-MP model is achievable to allow high data-rate full-duplex communication.

6.4 Future Work

The most relevant future work involves hardware testing with an actual DOCSIS power amplifier and node. The amount of cancellation possible with the UWB-MP model should

be measured to see if the proposed scheme is capable of supporting full duplex communication. Note that there are additional sources of noise present in hardware that was not taken into account in the simulations such as quantization noise and phase noise. Next, a channel estimation algorithm could be used in conjunction with the UWB-MP model to fully characterize the self-interference from reflections in the channel along with the interference from leakage within the node. An entire full duplex system could then be thoroughly tested.

Further research involves performance and resource usage. After testing with an actual PA, if the amount of cancellation is not sufficient, a more robust power amplifier model could be adapted to ultra-wideband signals. If this turns out to be the case, the general memory polynomial model is an alternative worth considering, although it would entail an increase in implementation cost.

We strongly recommend investigating the use of a compressed sensing (CS) based technique for reducing the hardware costs of the echo-canceling algorithm. CS algorithms attempt to determine which coefficients in a model (in our case, the UWB-MP model and channel model) are necessary and which are not based on measurements in the actual system. Doing so creates a sparse representation of the system which may use significantly fewer coefficients than the full system while experiencing only a minor performance degradation.

Appendix

Defining an Ultra-Wideband Signal

The characterization of a signal as ultra-wideband is typically governed by its fractional bandwidth. Recall from Section 3.3, the fractional bandwidth of a signal is defined as

$$\text{fractional bandwidth } = \frac{f_n - f_1}{f_c}.$$

The Federal Communications Commission (FCC) defines an UWB signal as having a fractional bandwidth ≥ 0.2, or an absolute bandwidth ≥ 500 MHz, regardless of the fractional bandwidth [40]. Our DOCSIS 3.1 DS signal has a fractional bandwidth of 1.67 and an absolute bandwidth of 1110 MHz, placing it solidly in the category of ultra-wideband signals by either definition.

The purpose of defining a fractional bandwidth in this thesis is to distinguish between the behaviour of signals when applied to a nonlinear device. For ultra-wideband signals, the out-of-band harmonics will interfere with the signal of interest, which is not the case for narrowband signals. In other words, harmonics with center frequencies at $2f_c$, $3f_c$, etc., will have a wide enough bandwidth to interfere with the original signal centered at f_c. In this case, narrowband DPD and EC algorithms will not be able to attenuate the harmonics interfering with the SOI. One could argue that very high order of harmonics at multiples of the carrier frequency would eventually interfere with the SOI for even narrowband signals, however the power in such harmonics is typically insignificant. For ultra-wideband signals, even the low order harmonics tend to interfere with the SOI.

References

[1] B. Fortosky, "Docsis 3.1 cable modem and upstream channel simulation in matlab," Master's thesis, University of Saskatchewan, 2016.

[2] Cable Television Laboratories, *Data-Over-Cable Service Interface Specifications DOCSIS 1.0*, radio frequency interface specification ed., November 2001.

[3] Cable Television Laboratories, *Data-Over-Cable Service Interface Specifications DOCSIS 1.1*, radio frequency interface specification ed., September 2005.

[4] Cable Television Laboratories, *Data-Over-Cable Service Interface Specifications DOCSIS 2.0*, radio frequency interface specification ed., April 2009.

[5] Cable Television Laboratories, *Data-Over-Cable Service Interface Specifications DOCSIS 3.0*, physical layer specification ed., December 2017.

[6] Cable Television Laboratories, *Data-Over-Cable Service Interface Specifications DOCSIS 3.1*, physical layer specification ed., May 2018.

[7] Cable Television Laboratories, *Data-Over-Cable Service Interface Specifications DOCSIS 3.1*, remote phy specification ed., May 2018.

[8] MACOM, *MAFL-008098-CD0550 Datasheet*, catv diplex filter: 5-42/54-1100 mhz ed. Rev. V6.

[9] MACOM, *MACP-010250-C808A0 Datasheet*, docsis 3.1 coupler 5-1200mhz 8db smt ed., Jan 2018.

[10] D. M. Pozar, *Microwave Engineering 3e.* 2005.

[11] P. Pratt and F. Kearney, "Ultrawideband digital predistortion (dpd): The rewards (power and performance) and challenges of implementation in cable distribution systems," July 2017. VOL 51.

[12] X. T. Vuong and A. F. Guibord, "Modelling of nonlinear elements exhibiting frequency-dependent am/am and am/pm transfer characteristics," *Canadian Electrical Engineering Journal*, vol. 9, pp. 112–116, July 1984.

[13] M. Huang, "4096-ofdm implementation on the hfc plant with fiber deep and distributed access architecture."

[14] F. Asharif, S. Tamaki, M. R. Alsharif, and H. G. Ryu, "Application of full-duplex wireless communication system on echo cancellation," in *Wireless Communications and Networking Conference (WCNC), 2013 IEEE*, pp. 3626–3631, IEEE, 2013.

[15] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1637–1652, 2014.

[16] M. Duarte and A. Sabharwal, "Full-duplex wireless communications using off-the-shelf radios: Feasibility and first results," in *Signals, Systems and Computers (ASILOMAR), 2010 Conference Record of the Forty Fourth Asilomar Conference on*, pp. 1558–1562, IEEE, 2010.

[17] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Transactions on Wireless Communications*, vol. 11, no. 12, pp. 4296–4307, 2012.

[18] Z. Zhan and G. Villemaud, "Combination of digital self-interference cancellation and aarfsic for full-duplex ofdm wireless," in *Communications in China (ICCC), 2014 IEEE/CIC International Conference on*, pp. 593–597, IEEE, 2014.

[19] V. O. Alan, W. S. Ronald, and R. John, "Discrete-time signal processing," *New Jersey, Printice Hall Inc*, 1989.

[20] L. Ding, G. T. Zhou, D. R. Morgan, Z. Ma, J. S. Kenney, J. Kim, and C. R. Giardina, "A robust digital baseband predistorter constructed using memory polynomials," *IEEE Transactions on communications*, vol. 52, no. 1, pp. 159–165, 2004.

[21] D. R. Morgan, Z. Ma, J. Kim, M. G. Zierdt, and J. Pastalan, "A generalized memory polynomial model for digital predistortion of rf power amplifiers," *IEEE Transactions on signal processing*, vol. 54, no. 10, pp. 3852–3860, 2006.

[22] 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, 2001.

[23] European Telecommunications Standards Institute, 650, Route des Lucioles Sophia Antipolis 06560 Valbonne FRANCE, *LTE Technical Specification*, 14.3.0 ed., April 2017.

[24] V. Mathews, "Adaptive polynomial filters," *IEEE Signal Processing Magazine*, vol. 8, pp. 10–26, July 1991.

[25] 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, 1999.

[26] H. W. Kang, Y. S. Cho, and D. H. Youn, "Adaptive precompensation of wiener systems," *IEEE Transactions on Signal Processing*, vol. 46, no. 10, pp. 2825–2829, 1998.

[27] P. Celka, N. J. Bershad, and J.-M. Vesin, "Stochastic gradient identification of polynomial wiener systems: Analysis and application," *IEEE transactions on signal processing*, vol. 49, no. 2, pp. 301–313, 2001.

[28] A. E. Nordsjo and L.-H. Zetterberg, "Identification of certain time-varying nonlinear wiener and hammerstein systems," *IEEE transactions on signal processing*, vol. 49, no. 3, pp. 577–592, 2001.

[29] K. Narendra and P. Gallman, "An iterative method for the identification of nonlinear systems using a hammerstein model," *IEEE Transactions on Automatic control*, vol. 11, no. 3, pp. 546–550, 1966.

[30] F. Chang and R. Luus, "A noniterative method for identification using hammerstein model," *IEEE Transactions on Automatic Control*, vol. 16, no. 5, pp. 464–468, 1971.

[31] E.-W. Bai and D. Li, "Convergence of the iterative hammerstein system identification algorithm," *IEEE Transactions on automatic control*, vol. 49, no. 11, pp. 1929–1940, 2004.

[32] P. Gilabert, G. Montoro, and E. Bertran, "On the wiener and hammerstein models for power amplifier predistortion," in *Microwave Conference Proceedings, 2005. APMC 2005. Asia-Pacific Conference Proceedings*, vol. 2, pp. 4–pp, IEEE, 2005.

[33] M. Isaksson, D. Wisell, and D. Ronnow, "A comparative analysis of behavioral models for rf power amplifiers," *IEEE transactions on microwave theory and techniques*, vol. 54, no. 1, pp. 348–359, 2006.

[34] T. Liu, S. Boumaiza, and F. M. Ghannouchi, "Augmented hammerstein predistorter for linearization of broad-band wireless transmitters," *IEEE Transactions on Microwave Theory and Techniques*, vol. 54, no. 4, pp. 1340–1349, 2006.

[35] Y. Liu, X. Quan, W. Pan, and Y. Tang, "Digitally assisted analog interference cancellation for in-band full-duplex radios," *IEEE Communications Letters*, 2017.

[36] O. Hammi, F. M. Ghannouchi, and B. Vassilakis, "A compact envelope-memory polynomial for rf transmitters modeling with application to baseband and rf-digital predistortion," *IEEE Microwave and Wireless Components Letters*, vol. 18, no. 5, pp. 359–361, 2008.

[37] A. S. Tehrani, H. Cao, S. Afsardoost, T. Eriksson, M. Isaksson, and C. Fager, "A comparative analysis of the complexity/accuracy tradeoff in power amplifier behavioral models," *IEEE Transactions on Microwave Theory and Techniques*, vol. 58, no. 6, pp. 1510–1520, 2010.

[38] S. Benedetto and E. Biglieri, "Nonlinear equalization of digital satellite channels," *IEEE Journal on Selected Areas in Communications*, vol. 1, no. 1, pp. 57–62, 1983.

[39] M. Welborn and J. McCorkle, "The importance of fractional bandwidth in ultra-wideband pulse design," in *Communications, 2002. ICC 2002. IEEE International Conference on*, vol. 2, pp. 753–757, IEEE, 2002.

[40] G. Breed, "A summary of fcc rules for ultra wideband communications," in *High Frequency Electronics*, pp. 42–44, Federal Communications Commission (FCC), Summit Technical Media, 2005.

[41] Intel, *Intel Arria 10 Device Datasheet*, June 2017.

[42] T. Gotthans, G. Baudoin, and A. Mbaye, "Comparison of modeling techniques for power amplifiers," in *Radioelektronika (RADIOELEKTRONIKA), 2013 23rd International Conference*, pp. 232–235, IEEE, 2013.

[43] A. Melendez-Cano, S. Juárez-Cázares, J. Galaviz-Aguilar, J. Cárdenas-Valdez, M. Garcia-Ortega, A. Calvillo-Téllez, P. Roblin, and J. Núñez-Pérez, "Behavioral modeling for power amplifiers comparing mpm, wiener and hammerstein with fpga-based implementation," in *Mechatronics, Electronics and Automotive Engineering (ICMEAE), 2016 International Conference on*, pp. 119–124, IEEE, 2016.

[44] F. M. Ghannouchi and O. Hammi, "Behavioral modeling and predistortion," *IEEE Microwave Magazine*, vol. 10, no. 7, pp. 52–64, 2009.

[45] S. Haykin, *Adaptive Filter Theory*. Prentice Hall, second ed., 1991.

[46] D. Large and J. Farmer, *Broadband cable access networks: the HFC plant*. Morgan Kaufmann, 2008.

[47] M. Bellanger, *Digital processing of signals*. Wiley New York, 1989.

[48] T. Araujo and R. Dinis, "On the accuracy of the gaussian approximation for the evaluation of nonlinear effects in ofdm signals," *IEEE Transactions on Communications*, vol. 60, no. 2, pp. 346–351, 2012.

[49] Qorvo, *RFPD3580 Datasheet*, 2016.

[50] A. S. Tehrani, T. Eriksson, and C. Fager, "Modeling of long term memory effects in rf power amplifiers with dynamic parameters," in *Microwave Symposium Digest (MTT), 2012 IEEE MTT-S International*, pp. 1–3, IEEE, 2012.