Adaptive Duplexer Control for Wireless Transceivers

Raheleh Eslampanah

B. Eng., M.Sc.

College of Engineering and Science

Victoria University

Submitted in fulfillment of the requirements for the degree of

Doctor of Philosophy

March 2016

VICTORIA UNIVERSITY
MELBOURNE AUSTRALIA

To

my loving parents,

&

my dearest love, Moharram

Declaration

"I, Raheleh Eslampanah, declare that the PhD thesis entitled 'Pilot Subsystem for Adaptive Duplexerx' is no more than 100,000 words in length including quotes and exclusive of tables, figures, appendices, bibliography, references and footnotes. This thesis contains no material that has been submitted previously, in whole or in part, for the award of any other academic degree or diploma. Except where otherwise indicated, this thesis is my own work."

Raheleh Eslampanah

[redacted]

March 1, 2016

[redacted]

Abstract

The popularity of smart phones, cloud computing and the growing market for machine-to-machine communications is fuelling the growth of mobile broadband. More frequency bands are therefore being allocated to mobile services. Unfortunately, for historical reasons, these band allocations are fragmented, poorly harmonised between different countries and no longer large enough for the ever-increasing broadband data rates. Carrier aggregation, the operation on two or more bands at once, will be necessary. A future wireless terminal capable of global roaming would therefore have to handle 35 bands, and be capable of carrier aggregation. This is the research challenge.

Duplexing is a technique to simultaneously transmit and receive information from the same antenna. Frequency Division Duplex uses two frequency bands for uplink and downlink communications, enabling the transmitter and receiver to operate continuously. Duplexing filters are essential for isolating the receiver from the strong self transmitted signal. Unfortunately the filters are expensive, non tunable and not suitable for integration into silicon. Each new frequency band requires its own duplexing filter and a complex array of switches and filters is required for multi-band operation

This work presents an adaptive duplexer architecture compatible with silicon integration. It is tunable over a wide frequency range and capable of carrier aggregation. The structure uses a combination of a low isolation device with multiple analog cancelling loops controlled by a normalised least mean square algorithm to track changes in signal leakage. The control is a multi-input multi-output problem. However the use of the ”inverse plant control technique” enables orthogonalisation into multi single-input

single-output problems, simplifying the structure and reducing convergence time over previous search algorithms. Convergence takes 10ms from 'cold' and tracking has a latency of 1.5ms.

Low power pseudo noise pilot signals placed in the receive bands measure the cancelling error residues used as feedback to the controller. The pilot generator is integrated into silicon using the Peregrine UltraCMOS® GC process. The generator was programmable to four sequence lengths and 8 delays for correlation purposes. The maximum chip rate of 100Mchip/s was more than the 10Mchip/s required for an LTE channel.

Acknowledgements

I am greatly indebted to Prof. Mike Faulkner for his contributions to all aspects of my work on this thesis. This work could not be accomplished without his wisdom, endless support, valuable guidance, expertise, and patience. I deeply appreciate the time he has sacrificed for my thesis, especially the long overnight hours and weekends. Not forgetting his lovely wife, Naida for her unconditional love and constant support makes me feel I have another mom in Melbourne.

I would also like to thank Dr. Jean-Michel Redouté for all his valuable guidance, comments and feedback. His encouraging and positive personality inspired me during my candidature. I am also thankful to Dr. Shabbir Ahmed for his technical support and guidance during the system simulations.

I wish to thank the Peregrine Semiconductor Company, the silicon sponsor, for their generous offer to support this research by providing free manufacturing of the silicon designs. My special thanks to Mr. Dan Nobbe, vice-president of engineering at Peregrine, for all his assistant and technical support and the time he put in teleconferences.

I must express my gratitude to Ms. Elizabeth Smith in the Graduate Research Centre of Victoria University for her continuous support and friendly advice. Ms. Smith's support went beyond anything I would have expected. I would like to thank Prof. Chris Perera and Prof. Stephen Bigger for their kindness and understanding. I am also thankful to the security personnel at Victoria University for their kindness and concerns about my security during night works, which made me feel safe and secure to work alone in the lab. I acknowledge that this work was funded by the Victoria University Postgraduate Research Scholarship and also partially by Ericsson.

I am also grateful to the wonderful support given by my friends, old and new, and especially colleagues in CTME who have made my PhD journey more memorable and enjoyable. Special gratitude to Li Xiong and Sara Sangtarash who have helped me stay sane through these difficult years. I greatly value their friendship and their belief in me.

I owe a special thanks to my family, my mom and dad, my sister and my brothers for their love, encouragement, advice, and an infinite amount of support throughout everything.

Last but not least, my deepest gratitude goes to my love, Moharram who shared all the ups and downs of my research. Although we are a thousand miles apart, his endless love, continuous support and motivating words have given me the strength to move on till the end of this journey.

Contents

List of Figures

List of Tables

Abbreviations

1G	1th Generation
2G	2nd Generation
3G	3rd Generation
4G	4th Generation
5G	5th Generation
3GPP	3rd Generation Partnership Project
ACLR	Adjacent Channel Leakage Ratio
ACS	Adjacent Channel Selectivity
ADC	Analog-to-Digital Converter
AWGN	Additive White Gaussian Noise
BAW	Bulk Acoustic Wave
BB	Base Band
BP	Band Pass
BPF	Band Pass Filter
BW	Band Width
CA	Carrier Aggregation
CF	Cost Function
CR	Cognitive Radio
CDMA	Code Division Multiple Access
CMOS	Complementary Metal-Oxide-Semiconductor
DAC	Digital-to-Analog Converter
DFF	D Flip Flop
DFT	Discrete Fourier Transform
DSP	Digital Signal Processors

DSSS	Direct Sequence Spread Spectrum
DL	Down-link
DR	Dynamic Range
DSP	Digital Signal Processing
E-UTRAN	Evolved Universal Terrestrial Access Network
ETSI	Telecommunications Standards Institute
ESD	Electrostatic Discharge
EVM	Error Vector Magnitude
FDD	Frequency Division Duplex
GPA	Gain Phase Adjuster
GSG	Ground-Signal-Ground
GSM	Global System for Mobile communications
IC	Integrated Circuit
IDFT	Inverse Discrete Fourier Transform
IL	Insertion Loss
IM	Intermodulation
IMD	Intermodulation Distortion
ITU	International Telecommunication Union
LMS	Least Mean Square
LNA	Low Noise Amplifier
LFSR	Linear Feedback Shift Register
LO	Local Oscillator
LP	Low Pass
LTE	Long Term Evolution
LUT	Look-Up-Table
MIM	Metal-Insulator-Metal
MIMO	Multiple Input Multiple Output
MIMOC	Multiple Input Multiple Output Control
MLS	Maximum Length Sequence
NF	Noise Figure
NLMS	Normalised Least Mean Square

OOB	Out-Of-Band
OFDMA	Orthogonal Frequency Division Multiple Access
PA	Power Amplifier
PCB	Printed Circuit Board
PN	Pseudo Noise
PSD	Power Spectral Density
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase-Shift Keying
RF	Radio Frequency
RFFE	Radio Frequency Front End
SAW	Surface Acoustic Wave
SC-FDMA	Single Carrier Frequency Division Multiple Access
SDR	Software Defined Radio
SEM	Spectrum Emission Mask
SISO	Single Input Single Output
SISO	Single Input Single Output Control
SMD	Surface Mount Components
SNR	Signal-to-Noise Ratio
SOS	Silicon-on-Sapphire
SVD	Singular Value Decomposition
TDD	Time Division Duplexing
TD-SCDMA	Time Division Synchronous Code Division Multiple Access
TFBAR	Thin-Film Bulk Acoustic Wave
UL	Up-link
UMTS	Universal Mobile Telecommunication System
VCO	Voltage Controlled Oscillator

Chapter 1

Introduction

1.1 Background

The Radio Frequency (RF) spectrum is a limited resource and hence has to be utilized in an effective, efficient and economical manner. The growing demand for RF spectrum is caused by the increasing popularity of various mobile internet applications running on smart phones and tablet PCs. Long Term Evolution (LTE) or the Evolved Universal Terrestrial Access Network (E-UTRAN) was introduced in the 3rd Generation Partnership Project (3GPP) to offer higher data transmission rates, lower latencies and improved flexibility in time, frequency and bandwidth for wireless data communication. LTE uses high order modulation (up to 64QAM), large bandwidths (up to 20MHz), and spatial multiplexing for high throughput. A multi-carrier approach for multiple access improves radio spectral density (less wasted spectrum) by enabling efficient scheduling in both time and frequency domains. Hence, Orthogonal Frequency Division Multiple Access (OFDMA) is chosen for downlink transmissions and Single Carrier Frequency

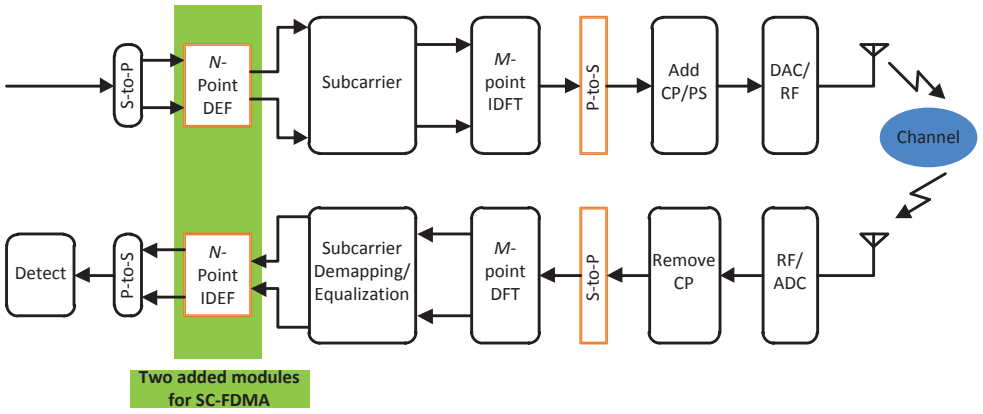


FIGURE 1.1: OFDMA and SC-FDMA. (cyclic prefix(CP), channel estimate pilots(PS))

Division Multiple Access (SC-FDMA) for uplink transmissions (shown in Fig. 1.1). The transmitter and receiver structures are similar with two extra blocks, (N-point DFT and IDFT) added for SC-FDMA. M-Total number of sub-carriers, set by the channel bandwidth (M=120 for Bch=20MHz). N-The number of sub-carriers in a resource block (N=12).

1.2 Concurrent Multichannel Operations

Carrier Aggregation (CA) technology, supported by the LTE-advanced standard enables transmission on multiple sub-bands to increase data rate and efficiency in spectrum usage [24]. Many channels each with bandwidths $\leqslant$20MHz are combined to give a maximum aggregated bandwidth of 100 MHz. Three schemes of carrier aggregation exist: intra-band contiguous, intra-band non-contiguous, as well as inter-band carrier aggregation are shown in Fig. 1.2. Intra-band contiguous combines adjacent channels while in non-contiguous allocation the channels are separated, but still in the same band. For inter-band non-contiguous operation the channels are in different operating frequency

bands [25]. Inter-band non-contiguous spectrum aggregation is required to accommodate todays fragmented band plans [26]. LTE transceivers also have to be able to accommodate these scenarios, and today's devices must be able to receive signals on two bands concurrently. Cognitive Radio (CR) is another system that could benefit from

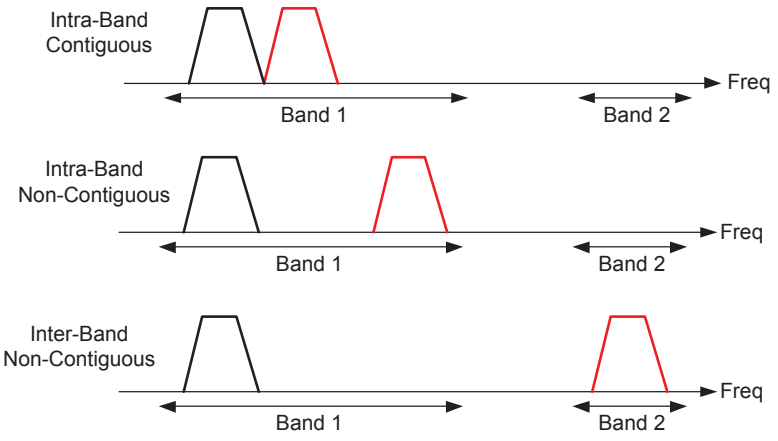


FIGURE 1.2: The specified carrier aggregation scenarios for LTE-Advanced.

the ability to receive on two channels at once. CR is an emerging technique to meet the shortage of frequency spectrum resources. A CR network facilitates the opportunistic usage of RF spectrum whenever or wherever it becomes available. It is able to adapt and optimise itself to the current radio environment by sensing which channels are un-occupied by primary (high priority) users. The original IEEE 802.22 CR protocol inserted blank periods in the transmission to allow CR users to sense for non activity on other channels. The loss in throughput due to the blank periods could be avoided if a CR could receive on two bands at once. one band could be used to receive the data (without blank periods) while the other band could be used for sensing purposes.

1.3 Software Defined Radio

Software Defined Radio (SDR) enables the implementation of CR by allowing the adjustment of radio parameters such as coding, modulation, carrier frequency and power. SDR provides an effective solution to the design of multi-mode, multi-band and multi-functional wireless communication transceivers. The SDR approach is to replace many of the traditional analogue hardware functions, such as channel selection, demodulation and synchronization, with software algorithms executed in high-speed Digital Signal Processors (DSPs). Configuring the radio to different frequency bands or air-interface standards necessitates the configurability provided by the software.

The remaining hardware functions in a SDR transceiver must be made generic (multi-band) and an important issue in SDR design is the replacement of fixed frequency components. These components cannot be tuned and only work for one band. Multiple components are therefore required for multiple band operation. Duplexing filters are the most notable of the fixed frequency components in today's wireless handsets. Their replacement is the topic of this thesis.

1.4 Duplexing

Duplexing is a technique to simultaneously transmit (Tx) and receive (Rx) information from the same antenna. Frequency Division Duplex (FDD) uses two frequency bands for uplink and downlink communication, enabling the transmitter and receiver to operate continuously. However, there is still a need to stop the very strong Tx signal

leaking into the receiver and corrupting its performance. This becomes more difficult to achieve with the increasing requirements on carrier aggregation and multi-band operation needed for modern wireless terminals.

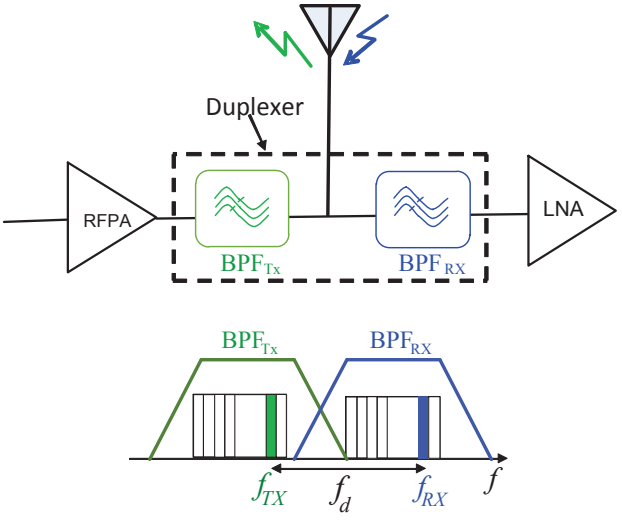


FIGURE 1.3: The conventional duplexer.

Isolating the receiver from the transmitter is traditionally achieved through the use of Passive Bandpass Filters (BPF) as shown in Fig. 1.3. The frequency difference between the paired transmit and receive channels is always constant over the band, and equal to the duplexing offset, f_d. The RFPA (radio frequency power amplifier) amplifies both the signal and the noise floor by a gain of G (Fig. 1.4). The noise PSD (power spectral density) of $K_B T F_t G$ is typically ~35dB above the receiver noise level. BPF_{Tx} stops the transmitter power amplifier's wideband noise from being injected into the antenna and receiver which would artificially increase the latter's noise figure in the receive band. On the other hand, BPF_{Rx} stops the high power transmit signal from driving the receiver circuits into intermodulation causing desensitisation and blocking of the received signal, as shown in Fig. 1.5.

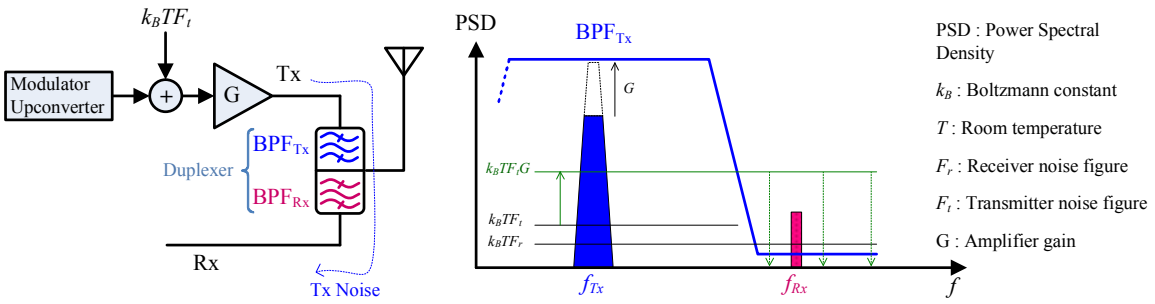


FIGURE 1.4: Transmitter noise in the Rx band is attenuated by BPF_{Tx}.

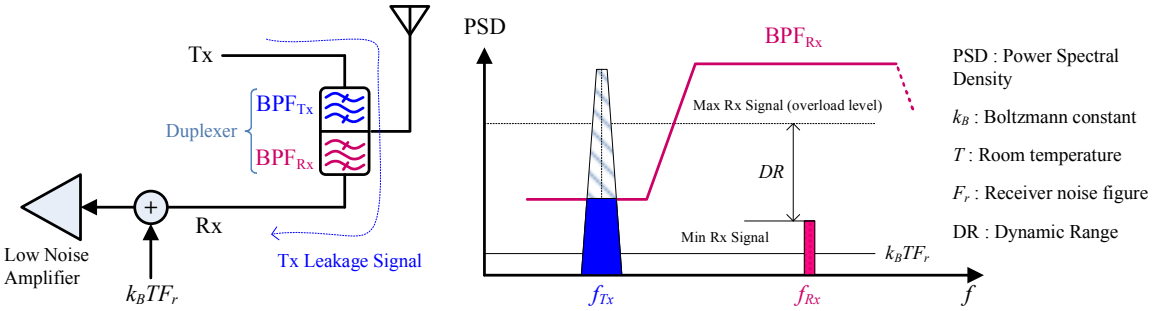


FIGURE 1.5: Receiver desensitisation due to overload by Tx signal unless attenuated by BPF_{Rx}.

Most of today's wireless terminals use Surface Acoustic Wave (SAW) or Thin Film Bulk Acoustic Resonator (TFBAR) devices to provide the required 45dB to 60dB of isolation between transmit and receive paths. The passive duplexer is one of the few remaining components that cannot currently be integrated into silicon, enabling a true single chip transceiver solution. Furthermore, these passive duplexers are non-tunable and so multiband FDD terminals require separate duplexing filters for each band and each aggregated band. A complex switching network is required to connect the appropriate duplexer to the antenna and the transceiver chip (Fig.1.7a). Moreover, SAW and FBAR duplexing filters are bulky devices which imply that putting an array of them into a cellular phone will result in increased size and cost of the handset [27].

1.5 LTE Operating Bands

An LTE world phone would be required to operate on some 31 duplexed bands as specified by the International Telecommunication Union (ITU). Fig. 1.6 shows the frequency allocations for 21 of these bands operating between 700MHz and 2.2GHz. Each band has an uplink and a downlink component separated by a frequency gap to allow for selectivity of the duplexing filters. It is not economic to make such a world phone; today's phones only operate on a subnet of these bands, depending on which geographic zone they are sold into. For example, the iphone 6 comes in four variants, one for Americas, one for Europe and two for Asia. Each variant will cover 16 of the 31 FDD bands.

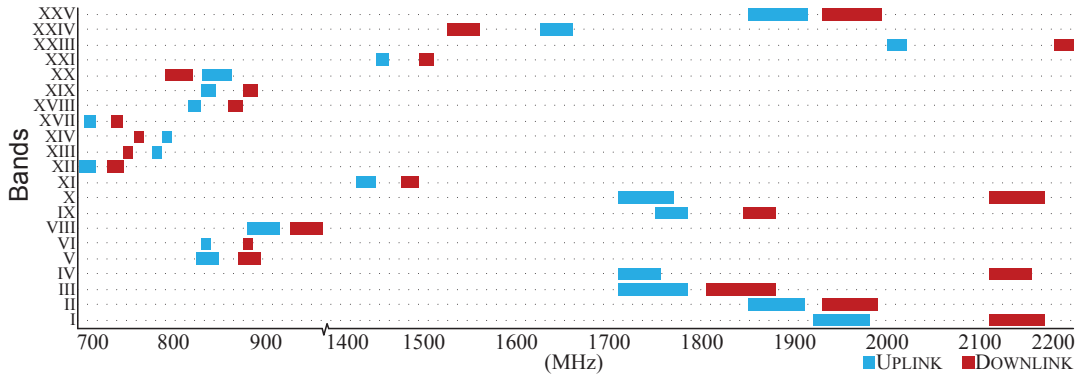


FIGURE 1.6: Band fragmentation; There are 31 FDD bands of which 21 are shown.

1.6 Adaptive Duplexers

To remove the size and flexibility restrictions imposed by passive duplexers on current RF transceivers, a method has been developed to replace the passive duplexer with an integrable adaptive duplexer [13]. An adaptive duplexer is a frequency flexible solution to the duplexing problem in the front end circuit of a low cost software defined

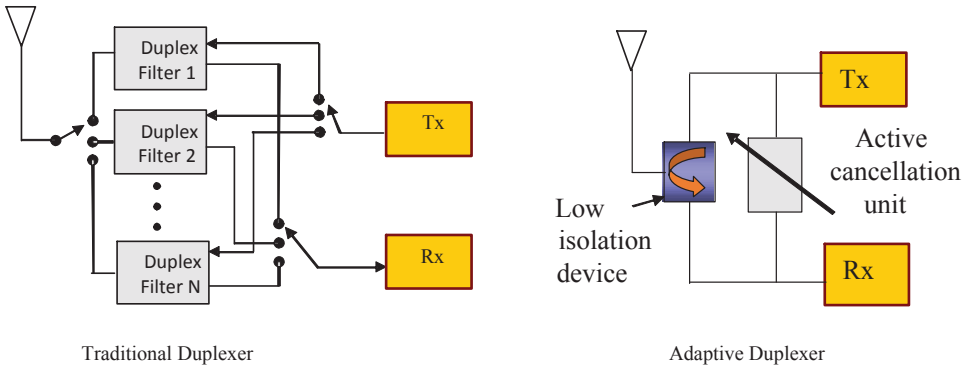


FIGURE 1.7: Multi-Band duplexer architectures. a) a traditional duplexer b) an adaptive duplexer.

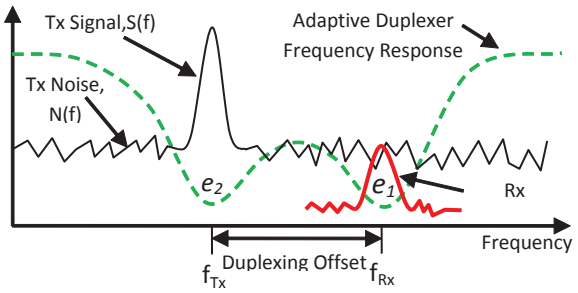


FIGURE 1.8: Adaptive duplexer cancels the TX signal and TX noise entering the receiver.

wireless terminal, and it is ideal for cognitive radio applications requiring simultaneous transmission and reception on unconstrained offset frequencies. Fixed frequency duplexing filters are replaced by adaptive cancelling loops. The adaptive wide-band architecture, shown in Fig. 1.7.b is based on a two stage process. In the first stage, a low isolation passive device is used to provide transmit-to-receive isolation of 20dB. This can be achieved using a wide-band circulator, directional couplers, or low-selectivity strip-line filters. This low isolation causes a significant transmit signal and noise in the receive signal path. In the second stage, an active double-loop cancelling technique is used to further increase the transmit to receive isolation by creating two nulls, one at the transmit frequency, f_{Tx} and the other at the receive frequency, f_{Rx}, shown in Fig. 1.8 to cancel the Tx signal leakage and noise [28].

Adjusting the cancelling unit to give two nulls requires an appropriate structure for the canceller and a means of measuring the residual Tx error signal, e_2, and a means of measuring the residual Tx noise error signal, e_1, in the Rx band for feedback control. The former (e_2) is easily obtained by using an auxiliary receiver tuned to f_{Tx}. The latter (e_1) is more difficult to obtain. A receiver at f_{Rx} already exists to receive the desired signal from the antenna. This receiver contains the Tx noise, to be minimised, masked by the Rx noise and the desired signal. A solution is to introduce a known pilot signal at the point in the system where the Tx noise is generated. Correlation can then be used to extract the pilot from the composite received signal, and give an estimate of e_1. Now with e_1 and e_2 known, an adaptive convergence scheme can be used to adjust the cancelling loops such that e_1 and e_2 are forced as close as possible to zero.

Currently, existing adaptation algorithms to null out the remnants of the Tx leakage and noise in two bands are too complex or not fast enough to track variations in the leakage path caused by antenna loading changes produced by skin contact or other disturbances. Further, existing adaptive duplexers do not generate multiple cancelling regions independently tunable over a wide frequency range. It should be noted that for 4G-LTE to meet it's specified 1Gbps downlink capacity, some 100MHz of spectrum is required. This can only be obtained by aggregating many channels (at least five for 20MHz channels).

1.7 Research Goals

- To extend the adaptive duplexing architecture to cater for multiple aggregated frequency bands by increasing the degrees of freedom of the cancelling process. Current adaptive duplexers only cater for one band.

- To develop a technique for measuring the performance of the adaptive duplexer in each aggregated receiver band.

- To develop a scalable control algorithm with low complexity and capable of tracking changes in leakage caused by variations in antenna environment. Current research is based on direct search adjustments which is slow when the number of adjustment parameters increase.

- To design, implement and test a integral pilot generator circuit for measuring noise leakage. The generator to provide both upconverted transmit signal and a reference signal for demodulation.

1.8 Research Contributions

The key contributions of the work are summarised:

✓ The use of multiple feed forward cancelling loops to extent the number of spectrum nulls.

✓ The additional nulls are used to increase the cancelling bandwidth. Carrier aggregation requires concurrent reception on multiple bands. Hence the canceller must

provide multiple nulls, which can be achieved by increasing the number of loops and therefore the degrees of freedom.

✓ The use of time shifted PN sequences to generate multiple orthogonal pilots without excessive complexity.

✓ PN sequences based on the well known m-sequence have an almost perfect autocorrelation function. This feature can be utilized to generate multiple (orthogonal) pilots from the one sequence. An additional feature of m-sequences is that when they are XORed with a delayed version of themselves another delayed version is generated. Multi-chip time shifted m-sequences can therefore be generated without additional shift registers, by using just a few extra gates.

✓ The development of a programmable pilot generator of four different lengths with RF and main and delayed outputs to enable measurement of amplitude and phase of the residual noise signal. A proof of concept demonstrator was designed in a Silicone on Sapphire (SOS) process.

✓ The use of ”inverse plant control” with integral feedback combined with NLMS tracking to enable low complexity convergence with latencies of ∼1.5ms for small adjustments. Convergence from ’cold’ is <10ms.

The research has led to the following publications:

- R. Eslampanah, S. Ahmed, M. Williamson, J.-M. Redouté, and M. Faulkner, ”LMS Inverse Control and Spread Spectrum Pilot Technique for Adaptive Duplexers”, Submitted to IEEE Transactions on Vehicular Technology on 11-02-2015.

- R. Eslampanah, M. Pourakbar, S. Ahmed, J.-M. Redouté, and M. Faulkner, "RF Modulator Design for a Low Level Pilot Sub-System", in Proc. of IEEE Australian Microwave Symposium (AMS), Melbourne, Australia, Jun. 2014,pp. 25–26.

- M. Pourakbar, R. Eslampanah, J.-M. Redouté, H. Sjöland, M. Törmänen, and M. Faulkner, "An SDR Duplex Filter in SOI Technology", in Proc. of IEEE Australian Microwave Symposium (AMS), Melbourne, Australia, Jun. 2014, pp. 26–27.

- R. Eslampanah, L. Linton, S. Ahmed, R. Amirkhanzadeh, M. Pourakbar, J.-M. Redouté, and M. Faulkner, "Active Duplexing for Software Defined Radio", in Proc. of IEEE International Symposium on Circuits and Systems (ISCAS), Melbourne, Australia, Jun. 2014, pp. 185–188 (*Selected among the 10 finalists for the student best paper award*).

- R. Eslampanah, M. Faulkner, "Pilot Signal for Adaptive Duplexers", in International Conference on Telecommunication and Technology and Applications., Proc. of CSIT., vol. 5, pp. 165-170, Jun 2011.

1.9 Thesis Outline

This thesis is organised in seven chapters. A brief description of each chapter is outlined below:

- Chapter 2 introduces the background of this research. The fundamental design requirements of a traditional wireless handset are presented. System design requirements and basic concepts are discussed.

- Chapter 3 reviews duplexer structures. Cancellation techniques and adaptation methods are discussed here.

- Chapter 4 is the main contribution and describes a solution to the convergence problem. The traditional two input two output control scheme is extended into a multi input multi output control (MIMOC) scheme for carrier aggregation. An adaptive duplexer with an inverse plant estimator based on the NLMS algorithm is demonstrated. A pilot subsystem is developed to measure noise leakage in the receive bands.

- Chapter 5 reviews the 'Silicon on Sapphire' process, and its advantages for system on chip integration.

- Chapter 6 describes the pilot subsystem circuit design. The pilot generator includes a variable length PN sequence associated with an adjustable delay line. A low output power RF modulator circuit is introduced.

- Finally, Chapter 7 summarizes the key outcomes and findings of the thesis, and gives direction for future work.

Chapter 2

Background Information

In this chapter, the fundamental design requirements of a traditional wireless handset are presented. The history of wireless systems from 1G to 4G and the required technologies for current wireless standards are explained in Section 2.1. Section 2.2 presents the direct conversion transceiver architecture. The concepts of these architectures such as noise, third order intermodulation, and the third intercept point are described. Section 2.3 and 2.4 provide system specifications to design an LTE transmitter and receiver respectively from which the duplexer isolation requirements are derived.

2.1 Wireless Systems

2.1.1 Wireless Systems History

The first generation (1G) of communication systems was analog and born in the 1980s and used 450MHz band. The systems had many problems such as capacity limitations,

low speech quality, and incompatibility with different country standards. Then second generation (2G), Global System for Mobile communications (GSM), appeared in the 1990s based on digital technology and it used 900MHz band, although later it started to operate in 1800MHz and 1900 MHz band. In between second and third, 2.5G with an unclear boundary with 2G came to provide more data capacity. Third generation (3G) provided a common platform for multiple wireless standards. These systems are mostly 200 times faster and support higher data rates than 2G systems, enabling services such as video, multimedia and internet. In 1999, the European Telecommunications Standards Institute (ETSI) employed 3G as the Universal Mobile Telecommunication Systems (UMTS) by choosing CDMA and W-CDMA standards. The 3rd Generation Partnership Project (3GPP) was established in 1999 to bring the telecommunications standards together. Fourth generation (4G) provides a variety of features such as different services, increased coverage and high speed data bandwidth. Recent trials have demonstrated data rates greater than 400 Mbps by aggregating up to 3 bands at once. Table. 2.1 shows a brief comparative study of the various generations of mobile equipment [22].

2.1.2 Wireless Today

3GPP introduces new features into the standard as ”Releases”. LTE, marketed as 4G LTE, is introduced in Release 8 and 9 to get higher data rates and lower latencies, to optimise packet switching, and to avoid unnecessary technology fragmentation between different jurisdictions [29]. Multicarrier and multiple antenna technologies are key fundamentals of LTE design.

TABLE 2.1: Mobile Generations [22].

Generation	Years	Mobile Technology /Switching Method	Data Bandwidth	Services Offered
1G	Early 1980s	Analog Cellular/ Circuit Switched	9.6kbps	Voice
2G	Early 1990s	Digital Cellular/ Circuit Switched	14.4kbps	Voice (main), SMS
2.5G	1996	Digital Cellular/ Circuit Switched Packet Enabled	144kbps	Voice and packet data introduced
3G	2000	Digital Cellular/ Circuit Switched Voice and Packet Switched Data	Up to 14Mbps	Packet data on high speed, voice, IMS enabled multimedia apps.
4G	2012-ongoing	Digital Cellular/ Circuit Switched, Packet Data Enable Packet based Voice	20 to 100 Mbps	Mobile broad-band,mobile TV,VOD, location based services, High speed data and security

2.1.2.1 Multicarrier Technology

In multicarrier technology, the downlink scheme is OFDMA, and the the uplink scheme is SC-FDMA (discussed in Section 1.1). Both provide multiple carrier flexibility, (Fig. 1.1). Flexibility enables wireless systems to have different bandwidths without any change in the fundamental system parameters. Further, it allows different users to have variable bandwidths, scheduled freely in the frequency domain. Moreover, fractional frequency operation is possible enabling interference coordination between cells [30].

2.1.2.2 Multiple Antenna Technology

Multiple Input Multiple Output (MIMO) systems are able to increase the system capacity without changing power or bandwidth. Traditionally, multiple antennas are used to enhance diversity to avoid fading [31]. Three main benefits of this technology are diversity gain, beamforming gain and spatial multiplexing (Fig. 2.1).

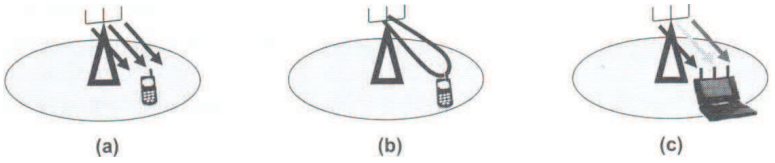


FIGURE 2.1: Main benefits of multiple antennas: a) diversity gain, b) array gain, c) spatial multiplexing [1].

2.2 Transceiver Architecture

A typical handset transceiver architecture is shown in Fig. 2.2. In the transmitter (Tx), a the input data is modulated in the DSP to form the Base Band (BB) signal which is converted to analog form and then upconverted to RF frequency. The Power Amplifier (PA) is used to amplify the signal to the desired level. In the receiver (Rx), the received signal is amplified and then downconverted to baseband. Then this signal passes through an Analog to a Digital Converter (ADC) and is demodulated in the DSP [32]. Any new adaptive duplexer (such as presented in this thesis) needs to be compatible with the Rx and Tx structures. The direct conversion transceiver architecture is commonly used for handheld transceivers. Other architectures are also possible. Sub-section 2.2.1 and 2.2.2 briefly explains the direct conversion transmitter and receiver respectively.

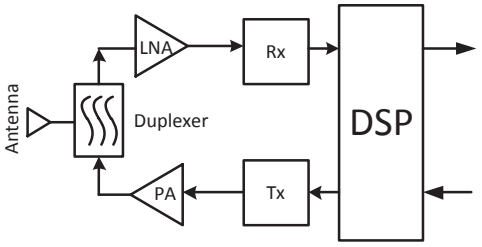


FIGURE 2.2: A typical transceiver architecture.

2.2.1 Direct Conversion transmitters

The transmitter's job is to deliver a BB signal to the antenna at a certain frequency, power and modulation in a specific band. In a direct conversion transmitter, the BB signal, normally centred to zero Hz, is converted to analog by two Digital-to-Analog Converters (DAC) that produce the in phase and quadrature components required for upconversion in a quadrature modulator. Then, the generated RF signal is amplified in the PA stage. A duplex filter providing band pass filtering is used to stop the transmitter noise from entering the receiver. A direct conversion transmitter architecture is shown in Fig. 2.3. Any deviation such as noise, mismatching, and distortion from the desired signal degrades the system performance. The main advantage of a direct conversion transmitter is its simple structure that makes it suitable for system on-chip [33].

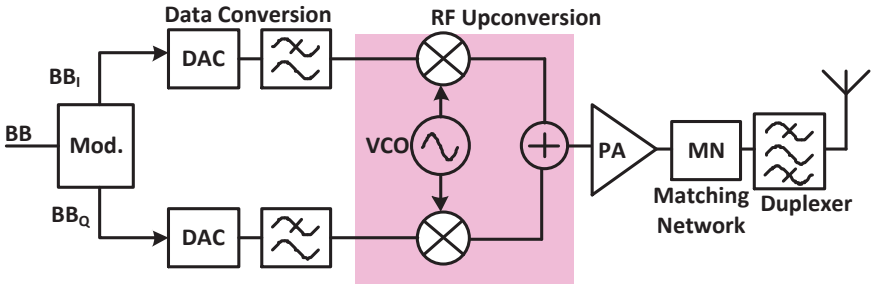


FIGURE 2.3: A direct conversion transmitter with quadrature modulation.

2.2.1.1 Transmitter Noise

Noise is a common problem in transmitter designs. The dominant noises are Local Oscillator (LO) phase noise and thermal noise (kTBxNF) generated by the electronics, where NF is the noise factor of the particular stage. These noises, along with the transmit signal, are boosted in the power amplifier.

- LO Phase Noise:

  Phase noise is caused by small frequency changes from the ideal oscillator set value. Phase noise varies with the amplitude of the carrier and affects the phase of a transmitted signal (Fig. 2.4). It is generally proportional to the oscillator frequency. If, for example, the carrier frequency is doubled the phase noise gets twice as big. Considering the constellation diagram of modulated signal, phase noise manifests itself as a circular error as shown in Fig. 2.5. Phase noise also produces adjacent channel interference in the transmitter. An out-of-band phase noise floor of -158 dBc/Hz often dominates the Tx output noise [34].

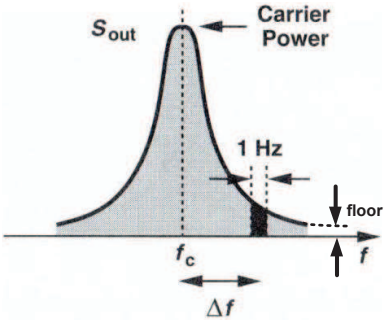


FIGURE 2.4: Phase noise diagram [2].

- Quantization Noise:

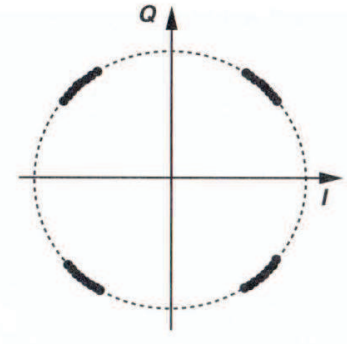


FIGURE 2.5: Constellation diagram with phase noise effect [2].

The mismatch between the the actual value and quantized DAC output value is called quantisation noise. The average noise power, assuming that the error is a white noise, is:

$$\sigma_q^2 = \int_{-q/2}^{+q/2} \frac{e^2}{q} de = \frac{q^2}{12} \tag{2.1}$$

Quantisation noise is not usually dominant in the output spectrum.

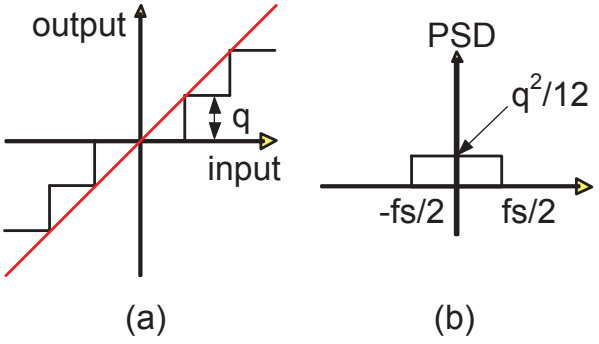


FIGURE 2.6: (a) Quantisation noise (b) The probability of error.

- RF Noise:

  RF noise comes from the RF components; the mixer and PA noise figures can be major contributors to the output spectral noise floor. The PA NF is usually in the 5dB to 10dB range [2].

2.2.1.2 Distortion

Non-linear distortion in the transmitter causes distortion in the transmitted constellation and unwanted transmissions in adjacent channels. The transmitter must meet a spectrum mask standard, and the requirements for LTE are described later in the chapter.

Distortion is caused by compression in the PA and by IQ (in-phase and quadrature) errors in the upconversion modulator. Techniques such as analog feedback or feed-forward linearisation or digital predistortion linearisation can correct such errors [35], [36].

2.2.2 Direct Conversion Receivers

A direct conversion receiver converts the received signal to the BB signal which can be filtered with a low pass filter. The architecture is depicted in Fig. 2.7. This simple architecture can be easily integrated. Further it has high selectivity with no image frequency problem, and utilises an LPF for channel selection. Unlike superheterodyne receivers, no RF image reject filter is needed. The performance of a RF receiver is limited by noise and nonlinearity effects such as third order intermodulation. These phenomenons will be discussed next [37].

2.2.2.1 Noise Figure

Noise Figure (NF) is a measure of degradation of the Signal to Noise ratio (SNR) in a system, and is the ratio between the noise output of the actual receiver and noise output of an ideal receiver with the same specs. It is defined as:

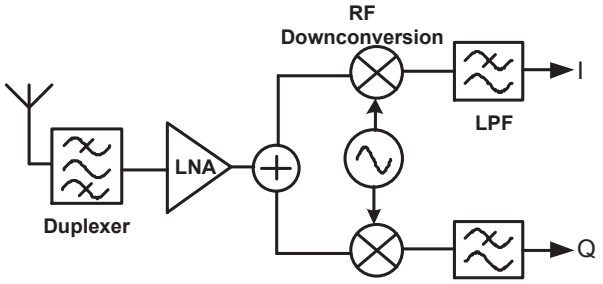


FIGURE 2.7: Direct conversion receiver.

$$NF = 10log\ \frac{SNR_{input}}{SNR_{output}} \tag{2.2}$$

The NF requirements for a handheld device in LTE is 9dB, and NF for the receiver circuits is 3-4dB [1].

2.2.2.2 Non-linearity Effect

Non-linearities in the active devices distort the amplified output signal. Desensitisation, blocking, and intermodulation which degrades the receiver sensitivity occur due to one or more strong input signals. In a FDD system signal leakage (through the duplexing device) can be one such strong signal. The output of a non-linear system is expressed as (2.3) which is a basic equation to analyse the non-linearity behaviour.

$$y(t) = a_1x(t) + a_2x^2(t) + \cdots + a_ix^i(t) \tag{2.3}$$

where x(t) is the input and $a_1, a_2, \cdots, a_i$ are coefficients defining the amplifier. In this sub-section the effects of non linearities in the presence of strong Tx leakages are

analysed.

- Receiver Desensitisation:

  A high Tx leakage reduces the gain of the desired signal and causes the signal-to-noise ratio at the detector input to drop. Assuming that $(A_{Rx}, \omega_{Rx}, \phi_{Rx})$ is a desired signal and $(A_{Tx}, \omega_{Tx}, \phi_{Tx})$ is the Tx leakage signal the input signal to the LNA becomes:

$$x(t) = A_{Rx}\cos(\omega_{Rx}t + \phi_{Rx}) + A_{Tx}\cos(\omega_{Tx}t + \phi_{Tx}) \tag{2.4}$$

  By substituting (2.4) in (2.3) and the fundamental component of the received signal is $(a_1 A_{Rx} + 3a_3 A_{Rx}^3/4 + 3a_3 A_{Rx} A_{Tx}^2/2)$. It can be easily observed that in the case of a strong Tx leakage signal and a weak received signal $(A_{Rx} \ll A_{Tx})$ the gain of the received signal reduces to $(a_1 A_{Rx} + 3a_3 A_{Rx} A_{Tx}^2/2)$ which is a function of the square of A_{Tx}. If the Tx signal is large enough the gain of the desired signal drops to zero and the signal is blocked. The difference (in dB) between the blocking signal level and the sensitivity level is called the receiver blocking ratio (Fig. 2.8). Note: a_3 in (2.3) is negative for amplifiers exhibiting compression at high signal levels.

- Intermodulation Distortion (IMD):

  Due to non-linearity behaviour of the LNA, a strong Tx leakage signal and any other interference signal (jammer) $(@f_j)$ in the Rx band creates spurious signals

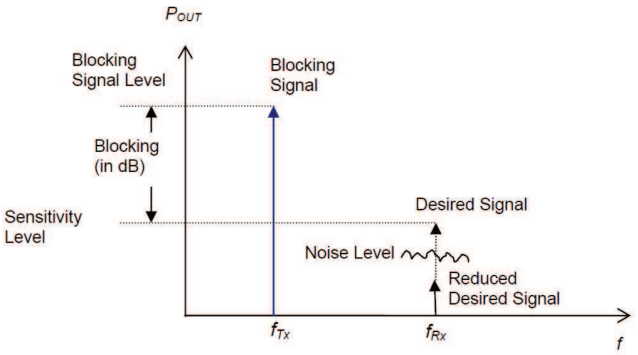


FIGURE 2.8: Receiver blocking.

at the output of the LNA known as intermodulation products [38]. The generated IMD frequencies are at: $\pm p\omega_{Tx} \mp q\omega_j$ where p and q are positive integers (Fig. 2.9).

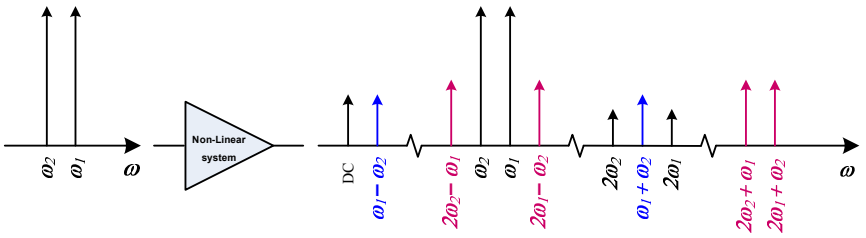


FIGURE 2.9: Intermodulation distortion.

- Third Order Intermodulation :

  Third order intermodulation is generated if a strong Tx leakage signal ($@f_{Tx}$) and a jamming signal ($@f_j$) corrupts the weak desired Rx signal. These products can fall in the Rx band and corrupt its performance, especially when the difference between f_{Tx} and f_j is small (Fig. 2.10). The figure shows an example of a narrowband jammer (eg a GSM signal) located approximately midway between a WCDMA transceiver's Tx and Rx bands.

  The Third order Intercept Point (IP3) is a specification used to compare the linearity of different circuits, and is the point at which the desired signal and the

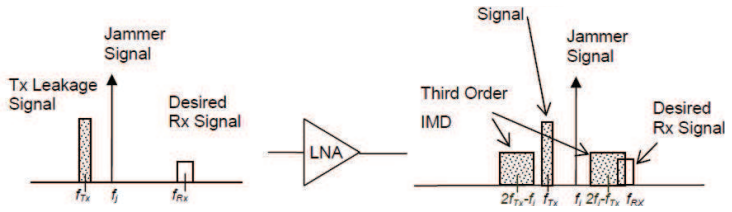


FIGURE 2.10: Distortion caused by Tx Leakage and a narrow band jammer

third order distortion products have the same amplitude. The combined Tx signal (A_{Tx}, ω_{Tx}) and jamming signal (A_j, ω_j) at the input is given by:

$$x(t) = A_{Tx} \cos \omega_{Tx} t + A_j \cos \omega_j t \tag{2.5}$$

IP3 is measured when $(A_{Tx} = A_j = A)$ and A is small so that higher non-linear orders are not taken into consideration. Then the third order intermodulation products are:

$$2\omega_{Tx} \pm \omega_j : \frac{3\alpha A^3}{4} \cos(2\omega_{Tx} + \omega_j) + \frac{3\alpha A^3}{4} \cos(2\omega_{Tx} - \omega_j) \tag{2.6}$$

$$2\omega_j \pm \omega_{Tx} : \frac{3\alpha A^3}{4} \cos(2\omega_j + \omega_{Tx}) + \frac{3\alpha A^3}{4} \cos(2\omega_j - \omega_{Tx}) \tag{2.7}$$

From (2.6) and (2.7), the third order intermodulation products are proportional to A^3 giving the distortion components a slope of 3 when plotted in dB. This concept is shown in Fig. 2.11 with the third order input intercept point (IIP3) and the third order output intercept point (OIP3). The plot is useful since it indicates how much back-off is required to give a predicted SDR (signal to distortion ratio). Typical receiver IIP3 values are between +0dBm and +10dBm [38].

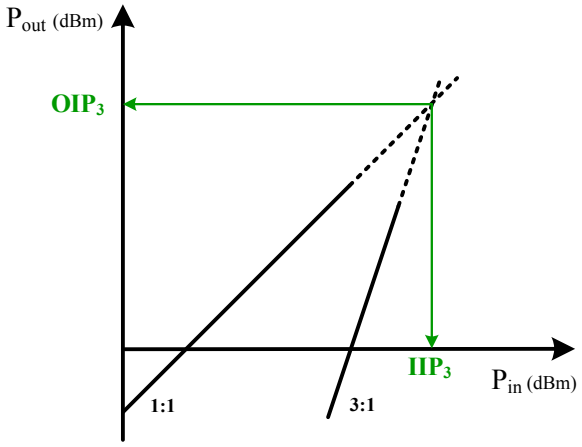


FIGURE 2.11: Definition of IP_3.

2.3 LTE Transmitter Requirements

Transmitter system specifications are related to the power level and the quality of the intended transmissions, and the acceptable level of unwanted emissions.

2.3.1 Requirements for the Intended Transmissions

- Signal Quality:

  Error Vector Magnitude (EVM): EVM is a measure of the modulation accuracy of the received signal and is caused by RF imperfections in the radio hardware. It is the magnitude difference between the actual received signal and the theoretical received signal. The maximum possible received SNR is determined by the EVM when the signal passes through an ideal noiseless channel. Shown in Fig. 2.12 is a one quadrant of a constellation diagram with the ideal signal and a signal with errors. The EVM requirement for Quadrature Phase Shift Keying (QPSK) is less than 17.5%, for 16 Quadrature Amplitude Modulation (16QAM) is less than 12.5%, and for 64QAM is less than 8% [1].

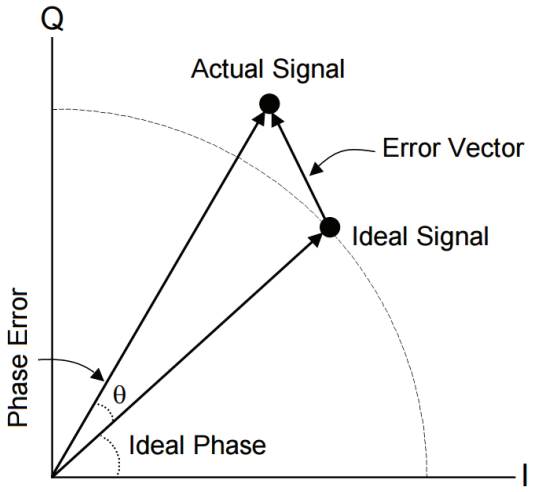


FIGURE 2.12: Error vector and Phase error.

- Transmit Output Power:

  The maximum transmit output power, measured at the antenna port, is 23dBm. The handset has to satisfy this requirement within a range of ±2dB [1].

2.3.2 Requirements for Unwanted Emissions

In reality, the transmitted signal emits outside of its transmission channel bandwidth. Out-of-band (OOB) emissions and spurious emissions are two categorises of undesired emissions shown in fig. 2.13. OOB emissions fall in a band close to the desired transmission, but spurious emissions can be at any other frequency, and the boundary between them depends on the details of the LTE specifications, meaning that it is not a fixed RF requirement.

- Out-of-Band Emissions:

  OOB emissions are mostly a product of the amplification and modulation processes and can normally be diminished by a reduction in transmitted power. In

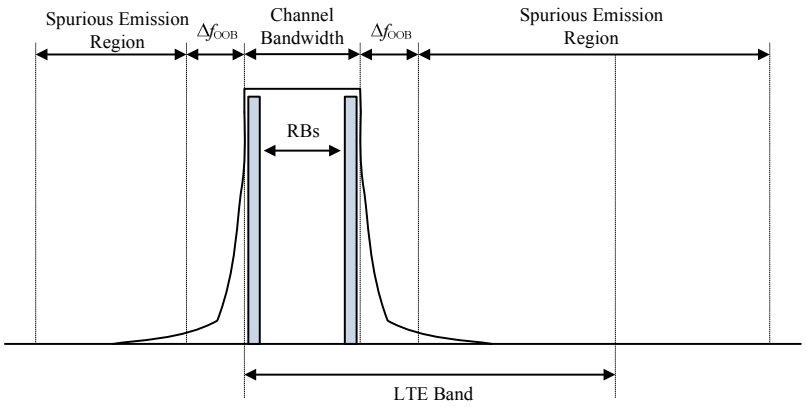


FIGURE 2.13: Transmitter spectrum in LTE [1].

LTE systems, OOB requirements are defined to the edge of the channel bandwidth by using the spectrum emission mask (SEM), and the adjacent channel leakage ratio (ACLR), a measure of unwanted power falling in nearby channels. The SEM's bandwidth is narrower than that of the ACLR and is related to the channel bandwidth of the transmitted signal as shown in Fig. 2.14.

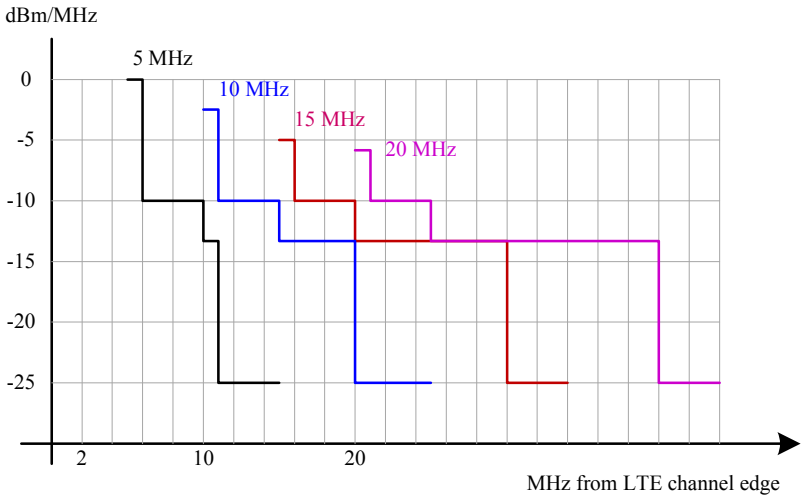


FIGURE 2.14: Spectrum Emission Mask for a user-end transmitter for different channel bandwidths [1].

ACLR estimates the effect of the OOB emissions on the neighbouring channels. The ACLR of a handheld is required to be >30dB for an adjacent LTE channel with 20MHz bandwidth, and >33dB for an adjacent UMTS channel in the closest

5MHz of the OOB spectrum, and >36dB for the adjacent UMTS channel in the next 5MHz [1].

- Spurious Emissions:

  Spurious emissions are caused by intermodulation distortion. These undesired emissions occur outside the channel. The basic specifications are given in Table. 2.2, but there are additional requirements in certain frequency bands.

TABLE 2.2: Spurious emissions requirements [1].

Frequency range	Maximum level (dBm)	Measurement bandwidth	PSD (dBm/Hz)
9KHz≤150KHz	-36	1KHz	-66
150KHz≤30MHz	-36	10KHz	-76
30MHz≤1000MHz	-36	100KHz	-86
1GHz≤12.75GHz	-30	1MHz	-90

Importantly, from a duplexing perspective, the spurious emissions from the transmitter of a handheld device that fall in its own receiver band are required to be ≤-47dBm/1MHz (-107 dBm/Hz) which is 21dB lower than specified in Table. 2.2. As the maximum transmit power of a handset is 23dBm, the spurious emissions needs to be -130dBc/Hz [1].

2.4 LTE Receiver Requirements

LTE receivers are similar to UMTS. The main difference is between the variable channel bandwidth and the new multiple access schemes.

2.4.1 Receiver General Requirements

The receive has to be able to meet several requirements. First, the antenna gain is assumed to have a gain of 0dBi. Then, the receiver has two antenna ports, and a four ports requirement maybe added in the future. Finally, for the purpose of specifications, the source of the signals are assumed to originate form an Additive White Gaussian Noise (AWGN) channel [39].

2.4.2 Reference Sensitivity

The reference sensitivity level is the minimum strength of Rx signal at the antenna port that can be successfully demodulated, and is defined as follows:

$$REFSENS = kTB + NF + SINR + IM - 3\ (dBm) \tag{2.8}$$

where kTB is thermal noise, NF is the total receiver noise figure, SINR is chosen based on the selected modulation and coding scheme, IM is the implementation margin, and the -3dB is for the diversity gain. Examples of the reference sensitivity are shown in Table. 2.3.

2.4.3 Transmit Noise leakage

In a FDD system, the receiver must operate in the presence of its own transmissions with only minimum performance loss. For both UMTS and LTE, when the transmitter

TABLE 2.3: REFSENS computation for Band 1 [1].

System	Modulation	Channel BW(MHz)	KTB (dBm)	NF (dB)	SINR (dB)	IM (dB)	REFSENS (dBm)
LTE UE	QPSK 1/3	5	-107.5	9	-1	2.5	-100
	QPSK 1/3	20	-101.4	9	-1	2.5	-94
	64QAM*3/4	5	-107.5	9	17.5	4	-80
	64QAM*3/4	20	-101.4	9	17.5	4	-74
LTE BS	QPSK 1/3	5	-107.5	5	1.5	2.5	-101.5
UMTS UE	QPSK 1/3	3.84	-108.2	9	1.2-21.1	2.5	-117

operates at full power the receiver sensitivity should be degraded by no more than 0.5dB [40]. The transmitter noise leaking into the receiver has to be 9dB less than the Noise Floor (-174dBm/Hz) measured at the antenna port which is -183dBm/Hz. If the spurious emissions are within the spectrum mask for the Rx band of -107dB/Hz (section 2.3.2), then the required duplex isolation is at least 76dB for LTE. In reality this is a high isolation requirement for transceivers in a confined space. Manufacturers, therefore have improved their spurious emissions by a further -26dB to -133dBm/Hz (Fig. 2.15 top) in which can the isolation requirements of the duplexing filters drop to 50dB(=183-133).

2.4.4 Transmit Signal Leakage

Isolation of the transmit signal itself is required to stop intermodulation in the low noise amplifier (LNA). The key specification is set by the intermodulation between the Tx leakage signal and a narrow band jamming signal located midway between the Tx and Rx bands at -44dBm. Such a jamming signal can come from a nearby GSM handset (see Fig. 2.10). A Tx leakage signal of below -25dBm is usually enough to meet this

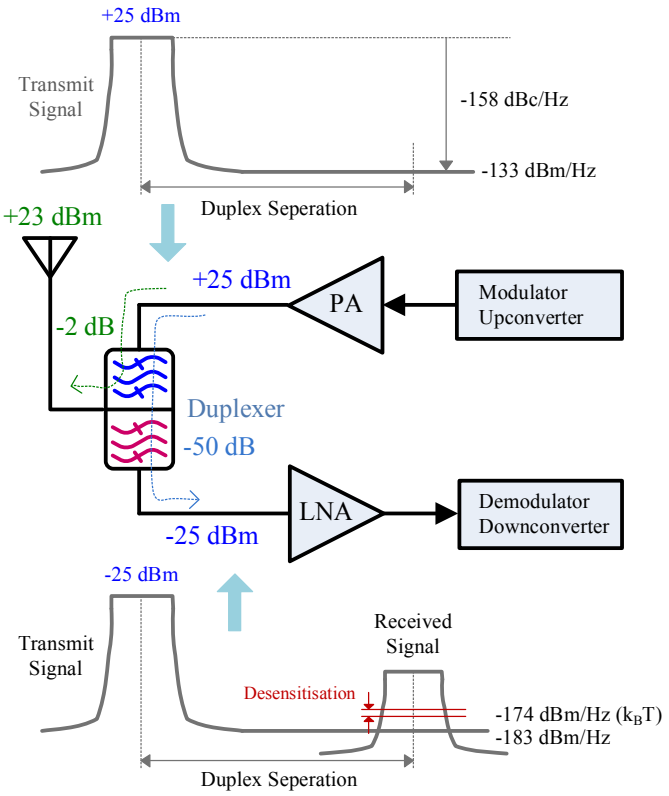


FIGURE 2.15: Tx leakage in the receiver in LTE [1].

requirement when the receiver IIP3 is ~+6dBm. Hence, if the transmit power delivered to the antenna port is 23dBm (25dBm at PA port -2dB transmission loss) and by assuming 50dB Tx-Rx isolation, the power at the LNA is -25dBm [1]. These scenarios are depicted in Fig. 2.15 [1].

2.5 Summary

After a brief overview of the history of cellular systems the chapter gives a basic overview of receiver and transmitter technologies, emphasising the direct conversion architecture. Distortion and noise issues are discussed with particular emphasis on the receiver since it's performance determines the isolation requirements of the duplexing filters. LTE transceiver requirements are then presented from which it is shown that the

duplexing filter must produce about 50dB of Tx to Rx isolation in both the Tx and Rx bands.

Chapter 3 reviews duplexers, adaptive duplexers, and adaptation schemes.

Chapter 3

A Review on Duplexer structures

The duplexer is a three-port device that isolates the receiver from the transmitter while permitting them to share a common antenna, and it is normally characterised by Tx and Rx insertion loss and isolation between Tx-Rx. As previously described, conventional duplexers consist of two bandpass filters and an impedance transforming circuit to allow both filters to connect to a common antenna port (Fig. 1.3). Duplexers are constructed in several ways which can lead to complex switching structures in multiband transceivers. Numerous researches have tried to overcome the complexity problem by proposing alternate structures which are either tunable or of a wideband nature. These are reviewed next. Section. 3.1 reviews duplexing filters and passive duplexers. Adaptive structures, recent published works, are discussed in Section 3.2. Existing adaptation methods are provided in Section 3.3, and a summary is given in Section 3.4.

3.1 Passive Duplexers

3.1.1 Duplexing Filters

Filters are the traditional method of providing Tx and Rx isolation in the FDD systems. They usually take the form of SAW, FBAR, ceramic and cavity structures. FDD systems use two singly terminated filters known as diplexers or duplexing filters. The two duplexing filters are connected at the terminating port (antenna) to form a three-terminal network. Handheld devices often use SAW filters which are based on the piezoelectric effect. A chip view is shown in Fig. 3.1 for an LTE band 1 frequency pair. The frequency response of the transmitter to antenna and antenna to receiver are depicted in Fig. 3.2, and a sample data sheet is briefed in Table. 3.1 [3]. From the response curves, insertion loss is less than 2dB and stop band isolation is > 40dB for the Tx noise at 2.14GHz, and > 50dB for the Tx signal at 1.95MHz. The goal of this research is to replace switched banks of these filters with an adaptive integrable system capable of similar specifications.

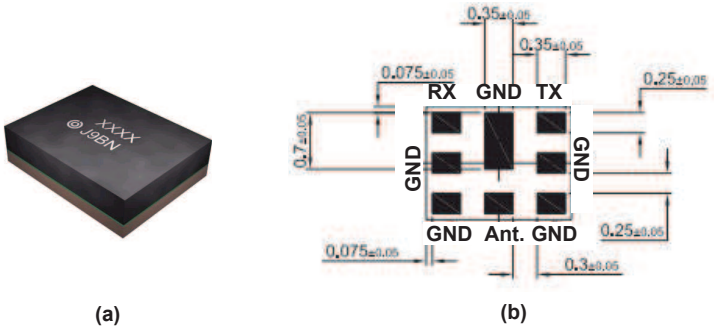


FIGURE 3.1: (a) SAW chip (b) a bottom view of the SAW layout [3].

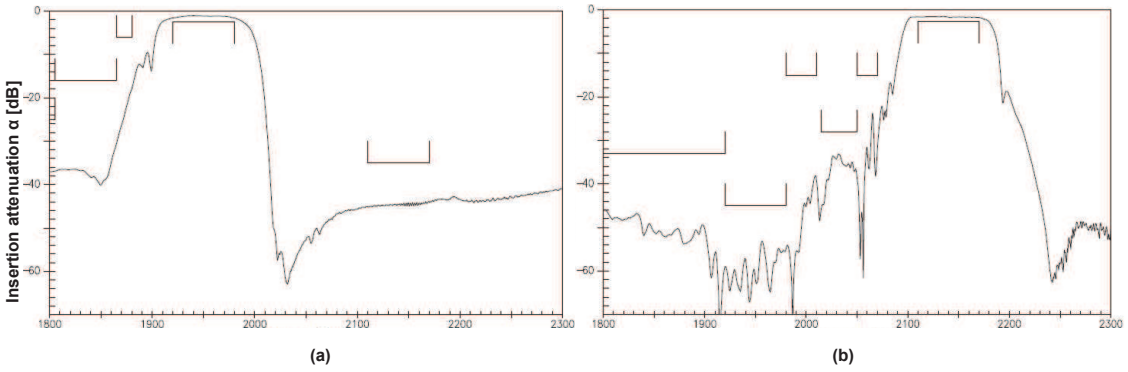


FIGURE 3.2: (a)Frequency response of Tx to antenna (b) Frequency response of antenna to Rx [3].

TABLE 3.1: SAW duplexer data sheet [3].

		B8653			
Characteristics TX-ANT		min.	typ. @ 25°C	max.	
Centre frequency	f_C	-	1950.0	-	MHz
Maximum insertion attenuation	α				
1920.48 ... 1979.52 MHz		-	1.7	2.4	dB
@ $f_{Carrier}$ 1922.40 ... 1977.60 MHz	α_{WCDMA}	-	1.6	2.4	dB
Amplitude ripple (p-p)	$\Delta\alpha$				
1920.48 ... 1979.52 MHz		-	0.4	0.8*	dB
Error Vector Magnitude					
@ $f_{Carrier}$ 1922.40 ... 1977.60 MHz	EVM	-	1.1	3**	%
Input VSWR (Tx port)					
1920.48 ... 1979.52 MHz		-	1.6	2	
Output VSWR (ANT port)					
1920.48 ... 1979.52 MHz		-	1.4	2	

* Over any 5 MHz in-band.

** EVM based on definition given in 3GPP TS 25.141.

3.1.2 Spatial Duplexing

Spatial duplexing with balanced Tx architectures has been proposed to provide the required Tx and Rx isolation. It needs one or multiple extra transmit antennas, which might not always be feasible in small form factor handsets. The architecture is shown in Fig. 3.3. By placing the Rx antenna midway between the two Tx antennas and inverting the feed to Tx1, the mutual couplings Tx1 to Rx and Tx2 to Rx cancel. A less

desirable side effect is the beam-forming inherent in such a structure. 55dB of isolation has been reported using this technique [4].

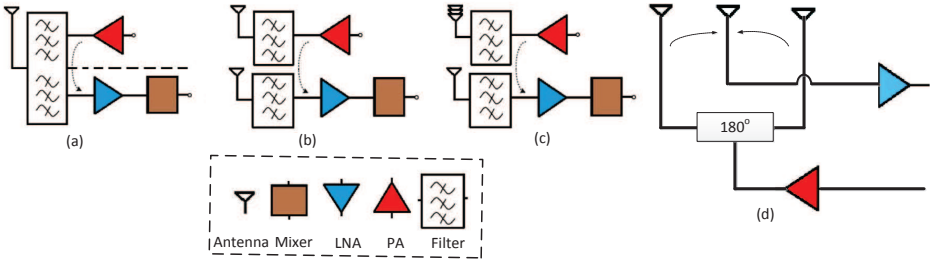


FIGURE 3.3: Evolution of the self-duplexing antenna. (a) Traditional duplexer. (b) Duall antenna. (c) Spatial duplexing filter (d) antenna placement [4].

3.1.3 Hybrid Transformer

The hybrid transformer can be used as a duplexer due to its simultaneously matched bi-conjugate ports that provide isolation between Rx and Tx, and split the incoming power between receiving and the balanced ports [5]. The autotransformer hybrid consists of a center taped coil with four ports labelled as Antenna, RX, TX, and Balance (Fig. 3.4(a),(b)). Hybrid transformers have the advantage of tunable wideband operation provided the balance network has the same impedance as the antenna at both Tx and Rx frequencies. Half the Tx power and half the Rx power ends up in this network. Apart from the 3dB insertion loss, there remains the problem of dynamically adjusting the balance network to track changes in antenna impedance with skin contact and changes in environment associated with a hand held device.

Related work [41] implemented a differential mode of the hybrid transformer to provide common mode Tx-Rx rejection at the receiver port as shown in Fig. 3.4.(c). A fully differential variation was also proposed by researchers from Qualcomm [42].

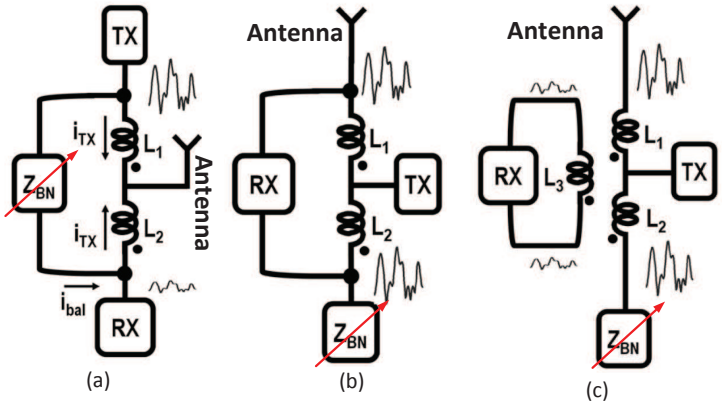


FIGURE 3.4: (a) Autotransformer. (b) Single ended RX. (c) Differential RX. [5].

3.1.4 Directional Coupler

Fig. 3.5 shows the directional coupler with matching network reported by [6]. The antenna port plays the input port role during receive, and the receiver port is the coupled port, where a percentage of the incoming power is delivered. An impedance matching network is placed at the coupled port to increase the isolation between transmitter and receiver during transmit. The scheme has similar losses and characteristics to the hybrid transformer of the previous section.

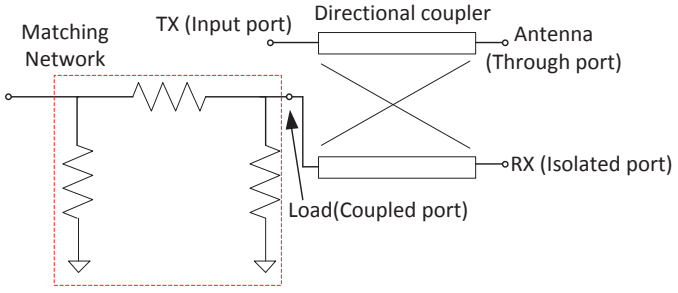


FIGURE 3.5: Directional Coupler Scheme [6].

3.1.5 Separated Antenna

A dual antenna scheme is proposed in [7] with analog canceller and digital adaptive filtering. In this example Tx and Rx share the same channel and must cancel all the co-channel interference from the Tx before the weak co-channel signal can be received. The use of separated antennas gives 20dB of leakage suppression. The cancelling loops provide the remaining isolation. Separated Tx and Rx antennas are a very efficient way of providing the initial isolation required to reduce signal levels to a power that can be handled by the receiver without generating distortion. The downside is the added real estate required for the second antenna. The scheme is shown in Fig. 3.6. The initial isolation of the separated antennas is boosted by cancelling loops in both the analog (RF) domain and the Digital Signal Processing (DSP) domain.

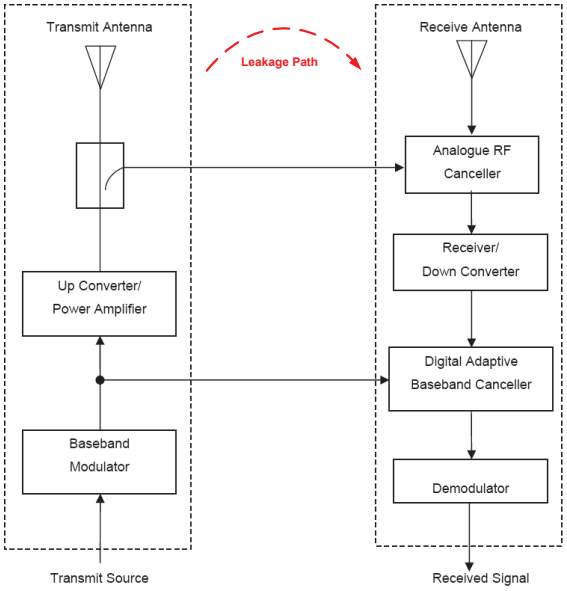


FIGURE 3.6: Block diagram of separated antenna RF communication system [7].

3.2 Adaptive Structures

An adaptive duplexer is a frequency flexible solution to the duplexing problem in the front end circuit of a low cost software defined wireless terminal. It consists of a low isolation device and an active cancellation unit to provide enough isolation between transmitter and receiver. Single-loop cancellation systems have been reported to improve the transmit band isolation of an existing duplexing filter, a circulator or a cross-polarised patch antenna [8]. Generally, single loop systems require unrealistic group delay matching with respect to the antenna path to achieve cancellation in both duplex bands. Therefore, it is necessary to employ a double-loop cancellation technique to achieve a double null with wideband cancellation.

3.2.1 Dual Antenna with Emulation Circuit

A dual antenna single loop system [8] used a delay line and antenna emulation circuit to increase the bandwidth of the cancelling loop (Fig. 3.7). An off-line adjustment procedure proposed the use of the receiver section to monitor (and minimise) a test signal (tone) transmitted at the required null frequency but no algorithm details or results were given [8].

3.2.2 A Feed-Forward Structure

A feed-forward cancelling scheme for the circulator has been proposed in [9] to enhance isolation. As is shown in Fig. 3.8, the input signal is sampled using a coupler, C1, and

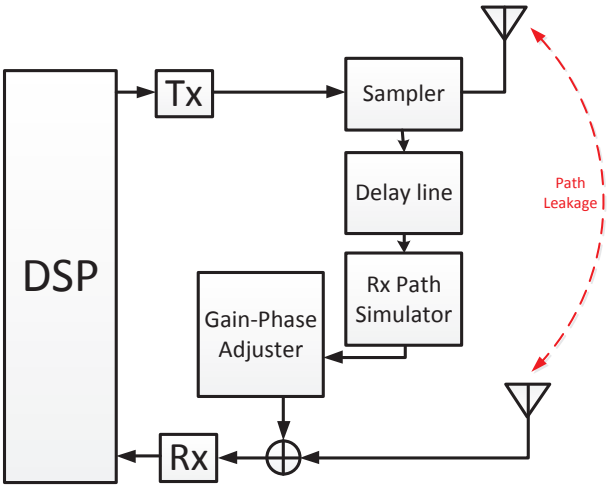


FIGURE 3.7: Block diagram of dual antenna with path emulation cancelling loop [8].

attenuated and phase shifted in a way that the amplitude of the sampled signal is the same as that of the leakage signal at the receiving port but 180° out of phase. Then the signals are combined in the receiver port. The scheme was improved in [43] to reach 60dB isolation over a small bandwidth by replacing the circulator with a dual polarised microstrip patch antenna. It should be noted that most feed-forward schemes suffer

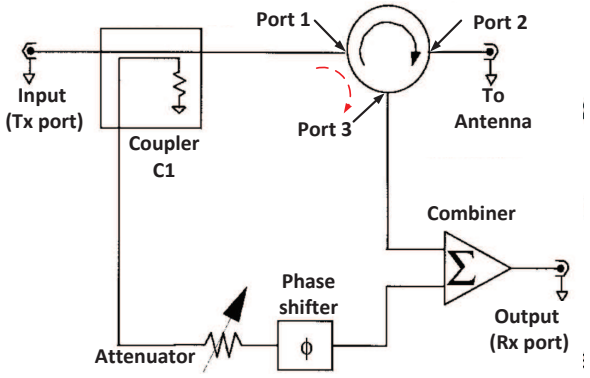


FIGURE 3.8: Feedforward scheme [9].

from the fact that the cancellation signal is a reproduction of the interference signal at one frequency only causing a narrow band null. Further, some of the feed forward

techniques increase the insertion loss from the transmitter to antenna because of the need for coupler C1.

3.2.3 Auxiliary Transmitter Cancelling Loop

Alternatively, [10] proposed to generate the Tx cancelling signal at baseband and up convert it with an auxiliary transmitter (Ax). The method has the potential of more accurate gain-phase matching over a wider bandwidth, but unfortunately exacerbates the transmitter noise falling in the receive band (Fig. 3.9). The noise from the "Ax" transmitter is added to the noise from the "Tx" transmitter. Filters at the Tx and Ax outputs must be added to solve the problem.

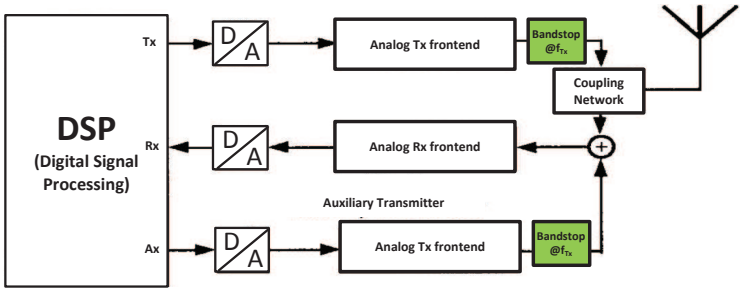


FIGURE 3.9: The TX/RX feedthrough applying an auxiliary transmitter [10].

3.2.4 Enhanced SAW Duplexing

Sullivan [11] used cancelling loops to improve the isolation of on existing SAW duplexing filters (Fig. 3.10). He found the cancelling null was very narrowband (2 MHz @20 dB) as a result of the large group delay mismatch between the leakage and feed-forward path. The acoustic nature of SAW filters impart considerable delay to the signal. Multiple cancelling loops were deployed to produce two or more nulls to widen the

cancellation bandwidth. This technique enhanced the noise rejection of the duplexer by 20dB over a 4.5MHz bandwidth, more than double the single loop bandwidth. The adjustment of the gain and phase coefficients in the cancelling loops was performed by look-up-table (LUT) following a pre optimisation calibration procedure. LUT's are only effective under stable conditions, a situation that does not exist in a handset environment due to continuous changes in antenna impedance.

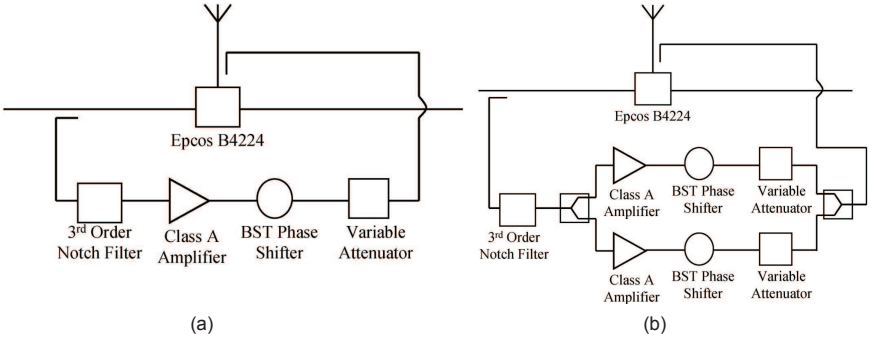


FIGURE 3.10: System topology for the adaptive duplexer. (a) single feedforward loop operation. (b)dual-path feedforward operation [11].

3.2.5 Analog Controlled Cancelling

An alternate adjustment technique [12] uses correlation to obtain the coefficient values for a single loop canceller. The cancellation process is applied after the LNA to reduce noise degradation due to noise generated by the cancelling loop. As shown in Fig. 3.11, the LMS control is fully implemented in analog. This structure is also for improving the duplexer performance and does not eliminate the need for duplexing filters. The main advantages of the work are the reduced area and eventual cost of the CDMA Rx, as well as low Rx SNR degradation. Furthermore, the circuit provides approximately 15dB additional cancellation in the case of a SAW duplexer. The bandwidth of 2 MHz is wide enough for a CDMA channel, but falls short of the 20MHz required for today's LTE

networks. The scheme correlates the output signal, $y(t)$, with the transmit signal, $r(t)$, to produce the cancelling signal, $z(t)$, such that the correlation $E(r(t), y^*(t))$ becomes zero.

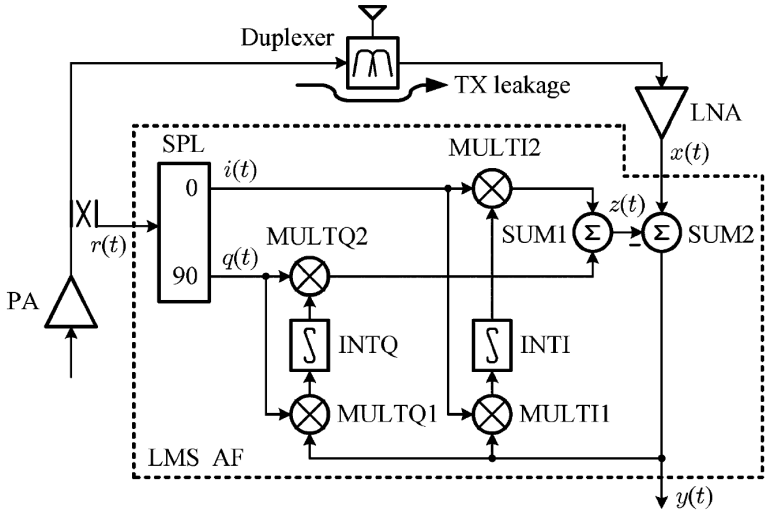


FIGURE 3.11: Analog LMS adaptive filter architecture used in [12].

All the above techniques focus on improving isolation in either transmit or receive bands separately, while relying on existing filters to provide the isolation for the other frequency bands.

3.2.6 Adaptive Duplexing

To remove the size and flexibility restrictions imposed by passive duplexers on current RF transceivers, a method has been developed to replace the passive duplexer with an integrable adaptive duplexer [13]. This solution combines a broadband low isolation circuit with a double loop cancellation technique that leads two nulls at the transmit frequency, f_{Tx}, and at the receive frequency, f_{Rx}. Cancellation of the transmit signal interference and the transmit noise in the receiver band is achieved simultaneously, without affecting the received signal. A simplified block diagram of the double loop cancellation process and the frequency domain representation of its desired response are shown

in Fig. 3.12, and Fig. 1.8, respectively. Each cancelling loop consists of a unique delay, τ_b, τ_c, and a vector modulator or gain-phase adjuster, h_1, h_2. The low isolation device in this case is circulator. These devices have relativity small group delay compared to SAW filters and so the null bandwidths are much wider, generally > 5MHz. Convergence was by way of a 1-dimensional search, will be discussed in section. 3.3.1. This structure forms the starting point for the research well work presented in this thesis.

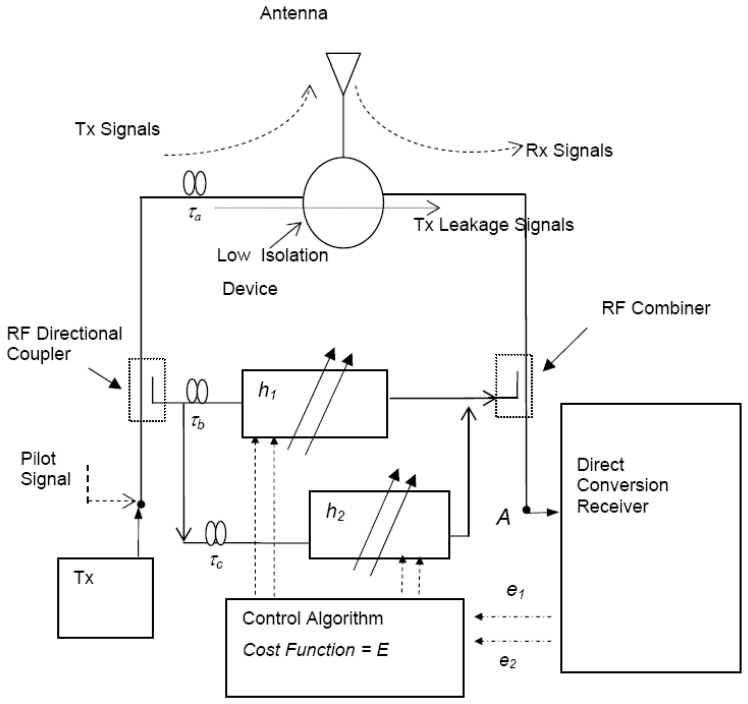


FIGURE 3.12: Double loop cancellation configuration presented by Kannangara [13].

3.2.7 Single Frequency Full Duplex

There has been recent interest in full-duplex communications operating on a common RF carrier. Since f_{Tx} and f_{Rx} are now the same frequency, the Tx signal must undergo a huge attenuation before reaching the receiver. Cancellation levels of over 100dB are required. This is a difficult task, however there is a benefit, and that is cancellation

is only required in one band, the Tx band. The work of [14] in single frequency full duplexing uses almost exactly the same architecture proposed in [13] to provide the initial cancellation, although only one frequency is involved, f_{Tx}. The analog cancellation circuit as shown in Fig. 3.13, employs fixed delays and tunable attenuators, so the cancelling signal is delayed by different fixed amounts and attenuated by different amounts. The delay lines are implemented with different lengths of wire causing a large circuit area. This method treats the circulator as a black box with known propagation delay and employs a large number of taps in the analog cancellation circuit. Digital cancellation loops provide a third stage of cancelling, and include non-linear as well as linear effects. Frequency division duplexers do not require this third stage of cancelling, since they have a more modest isolation target of ~50dB compared with 100dB for single frequency full duplex.

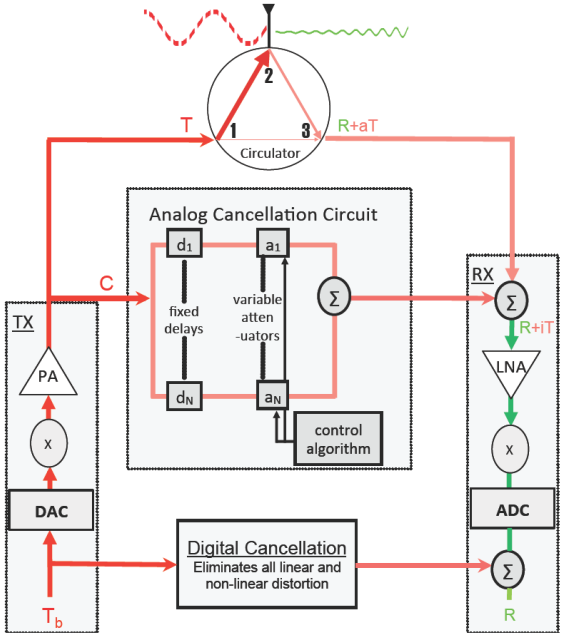


FIGURE 3.13: Full duplex configuration used in [14].

3.2.8 Full Duplex Using off-the-shelf Components

Duarte et al. in [15] implemented a single frequency full duplex prototype with a combination of antenna separation, analog and digital cancellation, and also by using the off the shelf MIMO radios. The RF model is shown in Fig. 3.14. The antenna separation distance between the Tx and Rx attenuates the signals and reduces self interference, h_{ab}. Their analog cancellation approach uses a separate transmit chain to create the cancellation signal, h_z. A number of factors limit the performance; including Tx noise, phase noise, and small channel variations.

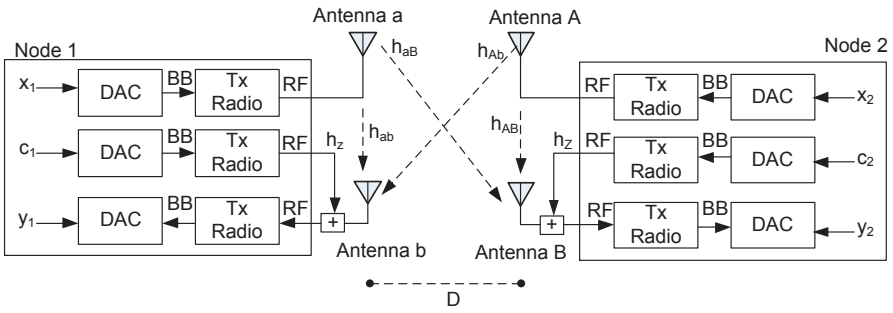


FIGURE 3.14: RF model of a full-duplex system [15].

3.2.9 Wiener Solution To Weight Adjustment

The work in [16], shown in Fig. 3.15, implemented the same RF scheme for the adaptive duplexer as [28] but used a Wiener filtering solution for weight adjustment in the passband. The scheme chooses the weights which minimise cancellation power. The suppression level of the cancellation is a function of the delay difference between taps and the number of them employed. A larger number of taps provides better suppression.

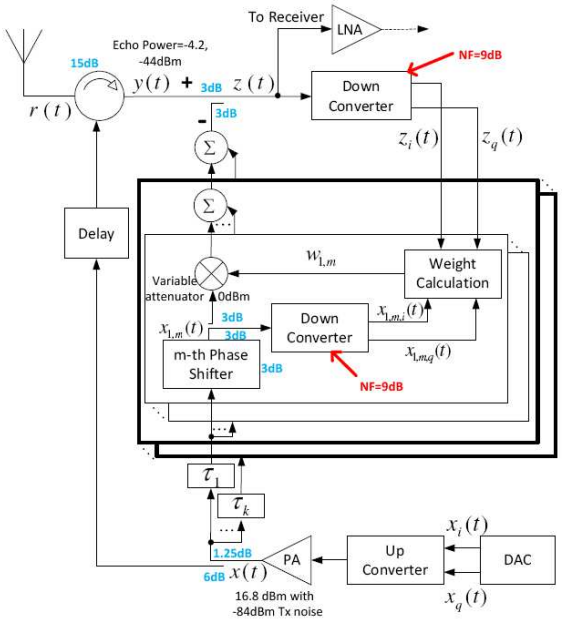


FIGURE 3.15: Full duplex configuration used in [16].

3.2.10 Adaptive Delay Filtering

Recent work [17] sought to reduce the number of taps by selecting the best delays. An adaptive delay filter normally used for baseband signals was customised for RF signals to reduce in-band and out of band Tx leakage. The proposed adaptation system (Fig. 3.16) estimates the delay and adaptation of each coefficient separately based on the error signals. On-line cancellation is by using the normalised least mean square algorithm. Tunable delay lines implement the delays which are calculated first, with the desired receive signal not present, (off-line). The coefficient weighting value, W_i is then estimated. The scheme was used to enhance an existing duplexing filter.

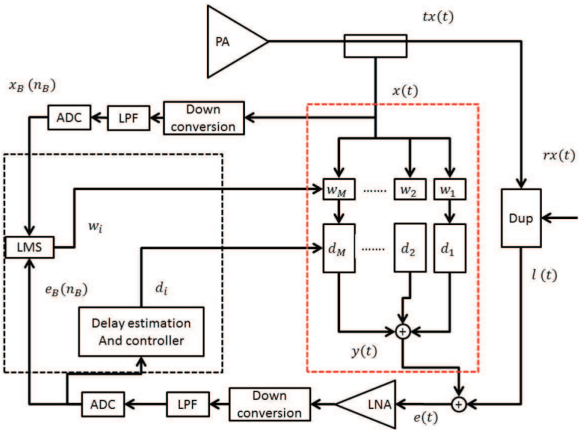


FIGURE 3.16: The structure of TX leakage canceller with adaptive delay filter proposed by [17].

3.3 Adaptation Algorithms

This section describes in more details the algorithms used to adjust the cancelling coefficients in an Adaptive Duplexer such as described in sub-section 3.2.6.

3.3.1 One-Dimensional Search

A one-dimensional (1-d) search is used to determine the minimum value of a function over a closed interval if the function has only one local minimum value. The algorithm involves a line search at each iteration. Assuming that $f(x^{(k)})$ is the function to be minimised, at the next iteration x^k is replaced with:

$$x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)} \tag{3.1}$$

where $x^{(0)}$ is a given initial point, $\alpha_k \geqslant 0$ is the step size, and $d^{(k)}$ the search direction. The vector, x^k, is incremented or decremented by a small step size, α_k, in the direction, d^k, to find the minimum value of the function (Fig. 3.17). The direction of search is then

changed and the process is repeated. It is normal to choose the direction, d^k, along one of the control axes.

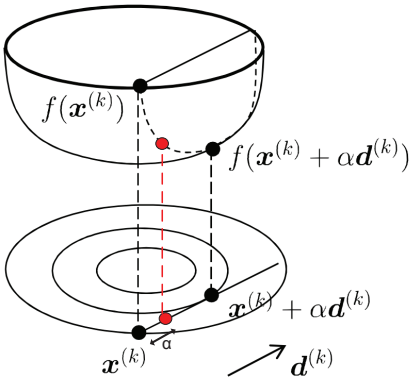


FIGURE 3.17: 1-d search; x^k is continuously incremented until the cost function is minimised in the direction of d^k. A new direction is then chosen [18].

Considering Shayma's work [28], explained in Section 3.2.6, the 1-d search was used to determine the minimum cost function to control the two error signals. It is based on searches along real (In-phase) and imaginary (Quadrature) parts of the control signals; h_1I, h_1Q, h_2I, h_2Q, such that an error cost function, E, is minimised.

$$E(h_1I, h_1Q, h_2I, h_2Q) = \beta|e_1|^2 + (1-\beta)|e_2|^2 \tag{3.2}$$

where β is a priority value, determining which band, $f_{Rx}(e_1)$ or $f_{Tx}(e_2)$, gets the most emphasis.

The cost function, E, is measured after each iteration and compared to the previous cost function to determine the direction of the cost function slope. The convergence time is reduced with larger step size, α^k, but this is at the expense of the final accuracy of the minimum point. It is normal to start the process with large α^k and finish with small

α^k (fine search). Even so convergence is generally slow compared to other techniques because the process of optimising each control signal must be repeated many times.

3.3.2 Singular Value Decomposition

Convergence time can be reduced if a multi-dimensional optimisation can be replaced by multi 1-d optimisations, all running in parallel. This can be done using a Singular Value Decomposition (SVD). The SVD of a matrix P, $P \in C^{m \times n}$, is its factorisation into three matrices; $P = U\Sigma V^H$. U and V are two unitary and orthogonal matrices, V^H is the conjugate transpose of V, and $VV^H = I$, the identical matrix. Σ is a real positive diagonal matrix containing the singular values. A general method [44] to compute SVD is factorization of the matrix A into Q and R and reduction of the upper triangular matrix R to a bidiagonal matrix B, and finally reduction of B to a diagonal matrix Σ by using an iterative method. This can be written as follows:

$$\begin{aligned} P &= QR \\ &= Q(U_1 B V_1) \\ &= Q(U_1(U_2 \Sigma V_2)V_1) \\ &= (QU_1U_2)\Sigma(V_2V_1) \\ &= U\Sigma V \end{aligned} \tag{3.3}$$

The SVD was used as a cross-talk cancellation method for the system in Fig. 3.12. A model of the feed forward cancelling loops and signal leakage path is shown in Fig. 3.18. The cancelling loops are modelled as a plant matrix, P, where the plant inputs are the

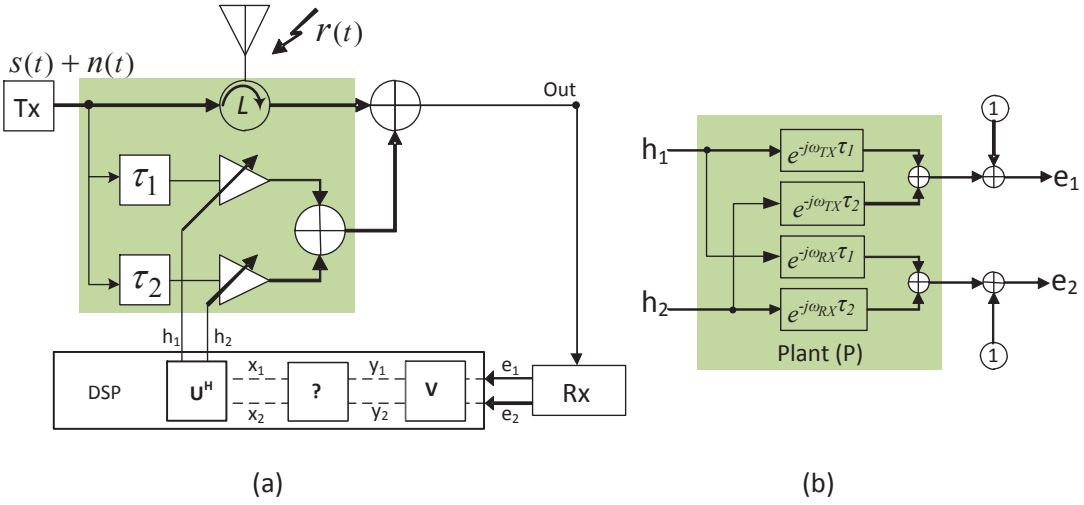


FIGURE 3.18: (a) adaptive duplexer with SVD (b) Model for input to output relationship [19].

control coefficients, h_1 and h_2, and the outputs are the residual error signals, e_1 and e_2, in the Rx band and Tx band respectively. The equations have been normalised such that $s(t)$ and $n(t)$ are equal to unity at f_{Tx} and f_{Rx} respectively. The leakage path through the circulator is $l_1@f_{Rx}$ and $l_2@f_{Tx}$. From the model these equations are derived:

$$E = L + PH \tag{3.4}$$

$$\begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} l_1 \\ l_2 \end{bmatrix} + \begin{bmatrix} e^{-j\omega_{Rx}\tau_1} & e^{-j\omega_{Rx}\tau_2} \\ e^{-j\omega_{Tx}\tau_1} & e^{-j\omega_{Tx}\tau_2} \end{bmatrix} \times \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \tag{3.5}$$

This equation is similar to that seen in a 2x2 Multiple Input Multiple Output (MIMO) wireless communication system [8]. One method of removing the effect of cross talk in a MIMO system is to use the SVD [19].

The SVD separates the matrix (P) into three matrices (U, Σ and V) as shown below:

$$P = U\Sigma V^H \tag{3.6}$$

$$\therefore U^H PV = \Sigma \tag{3.7}$$

Subsitiuting (3.6) into (3.4) and introducing new variables X and Y:

$$Y = U^H L + \Sigma X \tag{3.8}$$

where $H = VX$ and $Y = U^H E$.

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = U^H \begin{bmatrix} l_1 \\ l_2 \end{bmatrix} + \begin{bmatrix} \sigma_{11} & 0 \\ 0 & \sigma_{22} \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{3.9}$$

This linear equation now contains the complex controls x_1 and x_2, which, when multiplied with the transposed complex conjugate of matrix U^H generate h_1 and h_2 being sent to the gain and phase adjusters. As above, this combines with the plant to generate the error signals e_1 and e_2 which are then passed through matrix V to achieve the outputs y_1 and y_2. (3.9) shows the cross-coupling has been removed. The convergence simplifies as $y_1 \propto x_1$ and $y_2 \propto x_2$. x_1 can be adjusted to zero y_1 without having any effect on y_2. Convergence of y_1 and y_2 can now be performed independently using a single search method (1-d search for example) or directly by solving for x_1 and x_2 in (3.9) with $y_1 = y_2 = 0$. In matrix form : $X = \Sigma^{-1} U^H L$.

The manual calculation of the SVD for the plant in Fig. 3.18 outlined in (3.10)-(3.12):

$$\Sigma = \begin{bmatrix} \sqrt{2 + e^{\frac{1}{2}j\omega_D(\tau_1-\tau_2)} + e^{-\frac{1}{2}j\omega_D(\tau_1-\tau_2)}} & 0 \\ 0 & \sqrt{2 + e^{\frac{1}{2}j\omega_D(\tau_1-\tau_2)} + e^{-\frac{1}{2}j\omega_D(\tau_1-\tau_2)}} \end{bmatrix} \tag{3.10}$$

$$V = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{e^{\frac{1}{2}j\omega_D(\tau_1+\tau_2)}}{\sqrt{2}} & \frac{e^{\frac{1}{2}j\omega_D(\tau_1+\tau_2)}}{\sqrt{2}} \end{bmatrix} \tag{3.11}$$

$$U = \begin{bmatrix} \frac{e^{-j\omega_{TX}\tau_1-\frac{1}{4}j\omega_D(\tau_1-\tau_2)}}{\sqrt{2}} & -\frac{e^{-j\omega_{TX}\tau_1-\frac{1}{4}j\omega_D(\tau_1-\tau_2)}}{j\sqrt{2}} \\ \frac{e^{-j\omega_{TX}\tau_2+\frac{1}{4}j\omega_D(\tau_1-\tau_2)}}{\sqrt{2}} & \frac{e^{-j\omega_{TX}\tau_2+\frac{1}{4}j\omega_D(\tau_1-\tau_2)}}{j\sqrt{2}} \end{bmatrix} \tag{3.12}$$

It can be seen from the equations above for the SVD that each of the matrices has special conditions. Firstly the elements of matrix Σ are dependent on the duplexing frequency (ω_D), which is constant for a given FDD band, and the difference between delays τ_1 and τ_2. Secondly, for the matrix V the elements in the top row are independent and constant while the elements in the second row are dependent on ω_D and the sum of delays τ_1 and τ_2. The sum of the delays will not remain constant if the reference delay drifts due to antenna loading. Finally, in the U matrix all the elements are dependent on a single delay but they also depend on the Tx frequency as well as the duplexing frequency. The simulations of [19] were based on the theoretical values. Unfortunately, it is not possible to measure τ_1 and τ_2 due to their variation with antenna loading and circuit parameters. Any realistic solution must involve tracking the plant matrix, P, and

periodically performing an SVD to obtain the U, V, and Σ matrices. This adds to the complexity of the SVD system when operating under dynamic conditions.

3.3.3 LMS Adaptation Algorithm

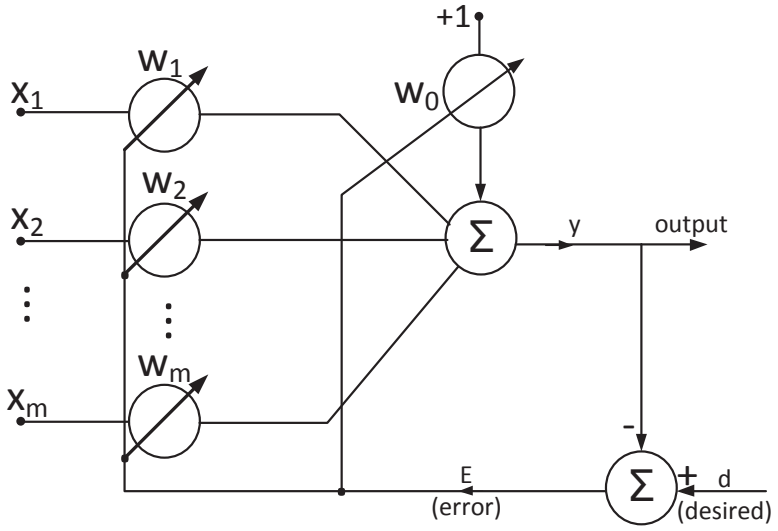


FIGURE 3.19: The adaptation algorithm structure [20].

The least mean suare (LMS) estimation algorithm was first proposed by Widrow and Hoff [20] in 1960 through their studies of pattern recognition, and is mostly used for adaptive filtering applications. The LMS structure is shown in Fig. 3.19 and its basic characterisation will be presented in this sub-section. The inputs are $x_1, x_2, \cdots, x_m$, and the W_is are weights. W_0 is the bias weight with a +1 constant input and is only necessary if the desired signal has a dc offset. The adaptation algorithm is similar to a neural network in terms of training. The desired response, d, is compared with the output response, y, and results in an error signal, e, which is used to adjust the weights, W to minimise the mean square of the error, $E(e^2)$. The maths that explain the algorithms is given here for completeness since the LMS algorithm plays an important part in the next chapter.

$$y = X^T W = W^T X \tag{3.13}$$

$$e = d - y = d - X^T W = d - W^T X \tag{3.14}$$

$$e^2 = d^2 - 2dX^T W + XX^T W^T W \tag{3.15}$$

$$MSE = \zeta = E[e^2] = E[d^2] - 2E[dX^T]W + W^T E[XX^T]W \tag{3.16}$$

$$\zeta = E[d^2] - 2R_{dX}^T W + W^T R_{XX} W \tag{3.17}$$

where $R_{dX} \triangleq E[dX]$ is the cross correlation between the desired signal and each of the input components and $R_{XX} \triangleq E[XX^T]$ is the autocorrelation matrix, and is a very useful indicator of the properties of input signal, X, and also the performance of the LMS algorithm.

So, the mean squared error, ζ, as a function of weights is quadratic and can be represented with a parabola as shown in Fig. 3.20. The weights for minimum mean square error occur at $W = W^*$. To get to this point, the bottom on the bowl shown in Fig. 3.20, the steepest descent technique can be used. The synapses the weights are updated as 3.18:

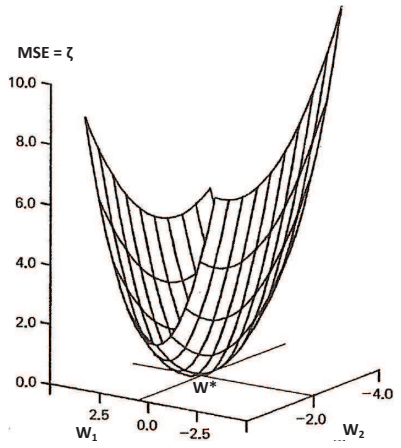


FIGURE 3.20: A two dimensional quadratic surface [20].

$$W_{j+1} = W_j + \mu(-\nabla) \tag{3.18}$$

where μ, a learning factor, controls the rate of descent down the steepest gradient, ∇. As in the real world we deal with the gradient estimate:

$$W_{j+1} = W_j + \mu(-\hat{\nabla}) \tag{3.19}$$

$$\nabla_j = \frac{\partial \zeta}{\partial W_j} = \begin{bmatrix} \frac{\partial \zeta}{\partial W_0} \\ \frac{\partial \zeta}{\partial W_1} \\ \vdots \\ \frac{\partial \zeta}{\partial W_m} \end{bmatrix} \tag{3.20}$$

$$\hat{\nabla}_j = \frac{\partial E(e^2)}{\partial W} = 2e\frac{\partial e}{\partial W} \tag{3.21}$$

and from 3.14:

$$\frac{\partial e}{\partial W} = -X \tag{3.22}$$

Finally the LMS algorithm by substitution of (3.22) and (3.20) into (3.19) results.

$$W_{j+1} = W_j + 2\mu e_j X_j \tag{3.23}$$

$$e_j = d_j - X_j^T W_j \tag{3.24}$$

and for stability: $\mu << 1/\lambda_{max}$ where λ_{max} is the greatest eigenvalue of R_{XX}. Considering (3.23), the implementation of the LMS algorithm is simple and efficient without a need for squaring, averaging, and differentiation.

Further, by substitution of 3.17 in 3.20:

$$\triangledown = 2R_{XX}W - 2R_{dX} \tag{3.25}$$

A primary concern with the LMS algorithm is its convergence to the optimum weight, where ζ is minimised. The minimum mean square error is the point where the gradient is zero and the weight vector is set to its optimal value, W^*:

$$\triangledown = 0 = 2R_{XX}W^* - 2R_{dX} \Rightarrow W^* = R_{XX}^{-1}R_{dX} \tag{3.26}$$

By substituting W^* for W in (3.17) the minimum mean square error is:

$$\begin{aligned}\zeta_{min} &= E[d^2] - 2R_{dX}^T W^* + {W^*}^T R_{XX} W^* \\ &= E[d^2] - 2R_{dX}^T [R_{XX}^{-1} R_{dX}] + [R_{XX}^{-1} R_{dX}]^T R_{XX} [R_{XX}^{-1} R_{dX}] \\ &= E[d^2] - R_{dX}^T R_{XX}^{-1} R_{dX} = E[d^2] - R_{dX}^T W^*\end{aligned} \tag{3.27}$$

Convergence speed of the LMS algorithm is another important parameter. This factor depends on the spread of eigenvalues of R_{XX}. The ratio, $\rho = \lambda_{max}/\lambda_{min}$ determines the speed the algorithm converges to a desired value. The speed goes up as ρ approaches to unity [45].

3.3.4 Normalised LMS Adaptation Algorithm

The normalised LMS algorithm has been developed as a solution for handling variable input signal powers which can lead to instability with large inputs or slow convergence for small inputs. It is hard to choose proper μ to guarantee a stable LMS with reasonable convergence time. This problem can be solved by normalising the LMS with the power of the input signals.

$$W_{j+1} = W_j + \frac{2\mu E_j X_j}{\|X_j\|^2} \tag{3.28}$$

Effectively μ is replaced by $\mu/\|X_j\|^2$, making the effective learning factor, μ_{eff}, reduce for large signals and increase for small signals. If $\|X_j\|^2$ is zero, then there is a

problem of instability and sensitivity to noise. A small constant added to the denominator: ($||X_j||^2 \rightarrow ||X_j||^2 + \varepsilon$) eliminates the problem.

3.4 Summary

This chapter reviews duplexing methods and includes a section on passive duplexers and adaptive structures. Passive duplexers (Section 3.1) included the filters, spatial duplexing, hybrid transformers, directional couplers, and separated antennas. Adaptive structures (Section 3.2) are based on various feed forward structures implemented in analog, DSP, and with separate auxiliary transmitters.

The adaptive duplexer described in section 3.2.6 is the basis of this work. It provides medium bandwidth cancellation for Tx interfering signals in the receiver and Tx noise signals in the receiver band and can cover most frequencies of interest.

The 1-d search, SVD, and LMS adaptation algorithms are discussed in Section 3.3. The speed of convergence using the 1-d search algorithm is limited by the bandwidth of the band pass filters. In the SVD work, ideal values were assumed, and therefore the system was limited to base station applications where conditions remain stable, and the system was not able to track the changes. The NLMS algorithm has the ability control its convergence rate based on the amplitude of the incoming signal. The control is a multi-input multi-output problem. The next chapter shows that modelling the inverse plant can orthogonalise the problem and dramatically improve convergence times with reduced complexity over previous search algorithms.

Chapter 4

System Level Design and Simulations

4.1 Introduction

The previous chapter overviewed various duplexer structures. In this chapter we generalise the two null adaptive duplexer structure of section 3.2.6 to a multiple null structure, where each transmission null can be independently controlled by a new convergence technique. The additional nulls can be used to increase the cancelling bandwidth for an existing transmission or increase the number of transmit and receive bands for concurrent multi-band operation as required by LTE-Advanced for carrier aggregation.

As mentioned, to remove the size and flexibility restrictions imposed by passive duplexers on current RF transceivers, a method has been developed to replace the passive duplexer with an integrable adaptive duplexer [13]. A broadband low isolation circuit combined with a two loop cancellation technique is used to produce nulls at the transmit frequency, f_{Tx}, and at the receive frequency, f_{Rx}. The structure is analogous to a finite

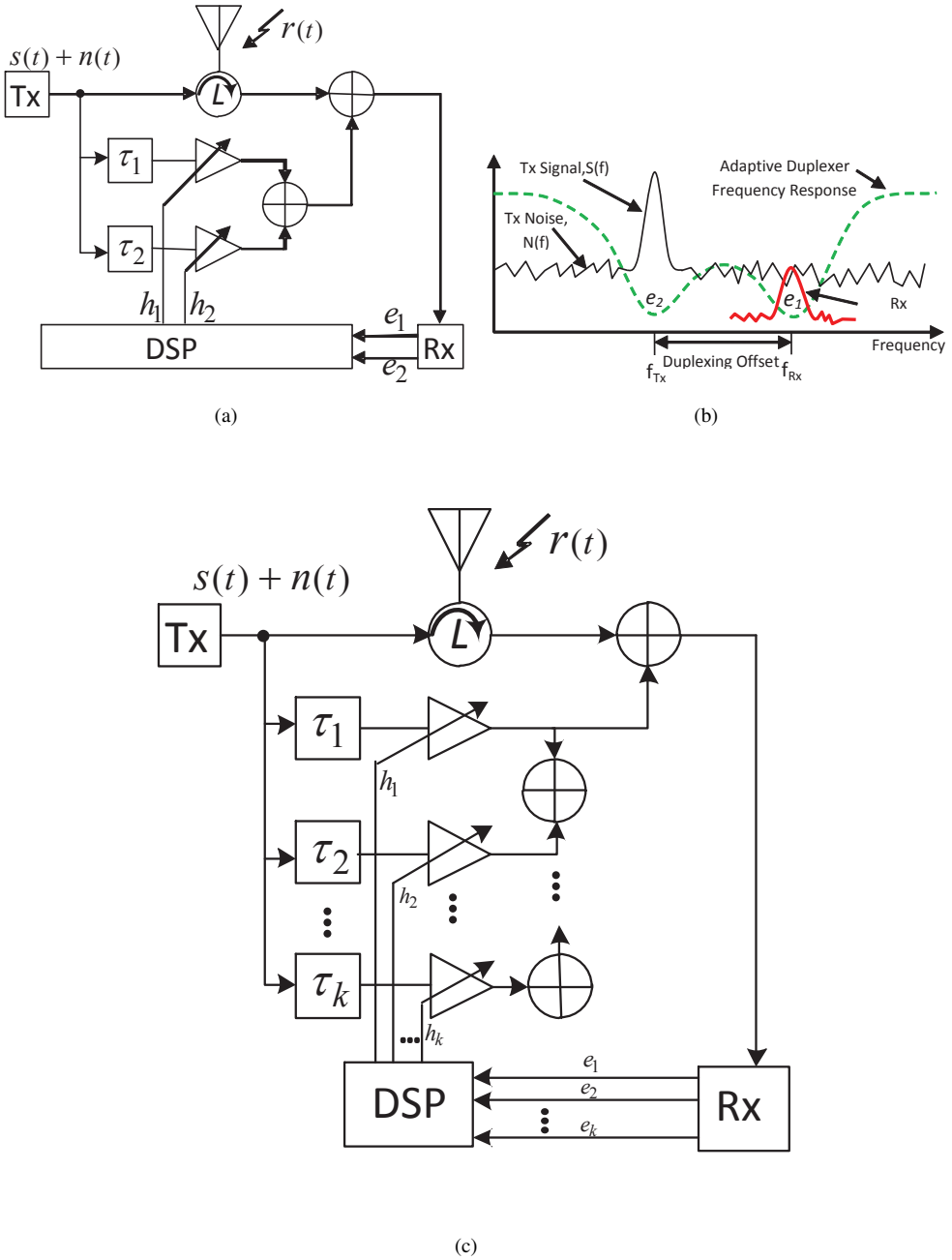


FIGURE 4.1: Adaptive duplexer architecture. (e_1 is measured in the transmit band and e_2 is measured in the receive band). (a) Dual loop canceller. (b) Frequency response (c) Proposed multiband [carrier aggregation].

impulse response (FIR) filter in DSP (digital signal processor). The two coefficients give two degrees of freedom that enable independent tuning of the two nulls. Cancellation of the transmit signal interference and the transmit noise in the receiver band is achieved simultaneously, without affecting the receive signal (Fig. 4.1).

This chapter is organised as follows: Section 4.2 gives an overview of the adaptive cancellation technique and extends it to carrier aggregated signals. Two potential convergence routines are also extended to handle the 'k' null requirement; including the SVD method and the 1-d iterative search method which is used as a reference for this work. A new inverse plant based solution is then presented in Section 4.3. It has some similarities with the SVD method, but avoids the need for a kxk matrix decomposition. Section 4.4 describes the pilot control subsystem for measuring isolation in the receive bands. Section V describes modifications to the LMS inverse control algorithm including the need for differential input signals. Simulation results are shown in Section 4.6 and Section 4.7 concludes the work.

4.2 Adaptive Cancellation with Carrier Aggregation

In the system diagram of Fig. 4.1(a) the band limited transmitted signal, $s(t)$, plus broadband transmitter noise, $n(t)$, is coupled via the low isolation device, L, to the antenna. Low isolation devices require about 20dB of isolation and can use various technologies for example; circulators, hybrids, tunable LC filters, separate Tx and Rx antennas etc [46]. Some of the signal leaks through the low isolation device into the receiver. Prior to the low isolation device a sample of the Tx signal and noise is passed through two cancellation paths made up of delays, τ_1 and τ_2, and gain-phase adjusters (GPA), h_1 and h_2. The single path canceller is represented by the special case $h_2 = 0$. The sum of these two paths produces a cancellation signal which is then combined with the reference path leakage signal. In the receiver the residual transmit signal and transmit noise leakages are measured e_1 ($@f_{Rx}$) and e_2 ($@f_{Tx}$) respectively) and passed

on to the DSP which adjusts the gain and phase of h_1 and h_2 until the cancellation signal is opposite to the leakage signal. The convergence goal is to reduce e_1 and e_2 to zero which produces the double null frequency response in Fig. 4.1(b). Carrier aggregation required additional receive band nulls; at least 1 null per receiver channel. Extra degrees of freedom are needed in the cancelling architecture, provided by additional loops each with a delay τ_k and a coefficient h_k; where k is the number of nulls. Following [13] the equation describing Fig. 4.1(c) is

$$E = L + PH \tag{4.1}$$

where the error signals, $E = [e_1, e_2, \cdots, e_k]^T$, comprise of leakages, $L = [l_1, l_2, \cdots, l_k]^T$, through the low isolation device and a cancelling signal controlled by the coefficients, $H = [h_1, h_2, \cdots, h_k]^T$. P is a kxk matrix describing the transfer function between H and E. An off-line adjustment procedure suggested by [8] was used to monitor and minimise a test signal transmitted at the required null frequency but no algorithm details or results were given. Two other methods will now be described.

4.2.1 Iterative Search

The k error signals $(e_1, e_2, \cdots, e_k)$ fed into the DSP (Fig. 4.1(c)) can be combined to create a single cost function that has a global minimum, $CF = \beta_1 {e_1}^2 + \beta_2 {e_2}^2 + \cdots + \beta_k {e_k}^2$, where β adjusts the emphasis between the nulls. The advantage of this cost function is that the error signal phase information is not required, which simplifies the error detection circuits to that of a simple rectifier. The complex coefficients, $h_1, h_2, \cdots, h_k$ can

be adjusted to minimise CF. The 1-d search technique adjusts h_{1I} by fixed increments to minimise CF, and then repeats the procedure for $h_{1Q}), h_{2I}, h_{2Q}, \cdots, h_{kI}, h_{kQ}$ (where h_I and h_Q are the real and imaginary components of h respectively). This process is further repeated a number of times before a global minimum is reached after which a smaller adjustment step size, α_k, is selected and the process repeated. From (4.1), it can be seen that a change in any one of either $h_{1I}, h_{1Q}, h_{2I}, h_{2Q}, \cdots, h_{kI}, h_{kQ}$ results in a change in all $e_1, e_2, \cdots, e_k$. This can result in a reduction in some of the error signals and an increase in the other error signals while still resulting in an improvement to the cost function [28].

4.2.2 Singular Value Decomposition (SVD)

While the 1-d search is stable and guaranteed to find the optimum solution the need for multiple rounds of improvement creates a long convergence time. The speed of convergence can be improved by creating controls that are orthogonal for the two error references to create two separate control loops. One technique for this would be to use the SVD which has been used for MIMO antenna systems.

As per Section 3.3.2 the SVD separates the plant matrix (P) into three matrices (U, Σ and V) each having special properties, however, the important property of the SVD is that the Σ matrix contains zeros in all elements except along the diagonals called the singular values. The components of the SVD are restated below for a system with k degrees of freedom [19].

$$P = U\Sigma V^H \tag{4.2}$$

$$\therefore U^H PV = \Sigma \tag{4.3}$$

Substituting (4.2) into (4.1) and then premultiplying both sides by U^H (where $(.)^H$ means the conjugate transpose.) gives:

$$Y = U^H L + \begin{bmatrix} \sigma_{11} & 0 & \cdots & 0 \\ 0 & \sigma_{22} & & \vdots \\ \vdots & & \cdots & 0 \\ 0 & \cdots & 0 & \sigma_{kk} \end{bmatrix} X \tag{4.4}$$

where $H = VX$ and $Y = U^H E$. The plant as seen by the new variables $X = [x_1, x_2, \cdots, x_k]^T$ and $Y = [y_1, y_2, \cdots, y_k]^T$ consists of k Single Input Single Output Control (SISOC) systems each with gains $\sigma_{11}, \cdots, \sigma_{kk}$ respectively. A simple integral controller on each SISOC channel will drive the error signals Y (and therefore E) to zero.

The effect of the SVD on the adaptive duplexer is to improve the convergence time but requires updating of the two unitary matrices as conditions such as delays and duplex frequency change. This was not considered in [19], since ideal U and V were assumed. In practice the plant must be continuously measured and decomposed using numerically difficult SVD operations. The decomposition complexity increases with $O(k^3)$ and becomes less attractive for systems with aggregated channels (large k).

4.3 Inverse Plant Modelling and Control

Our proposed method of removing the effect of cross talk in a Multiple Input Multiple Output (MIMOC) system is to use matrix inversion to find a new matrix (P^{-1}) that when pre multiplied with the plant results in the identity matrix. Inverting the P matrix is a slightly less numerically demanding operation than teh SVD, but still $O(k^3)$. We avoid the problem, however, by directly measuring the inverse plant, P^{-1} (rather than P) using the technique of "inverse plant modelling" [47], [48], [49]. The plant represents the cancelling loops and antenna port leakage (Fig. 4.2). A Model of the inverse plant, M, is estimated from the plant's inputs and outputs (H and E respectively) and then used as a pre-filter for the main (integral) control system.

The error signals will require both amplitude and phase estimates, which is difficult to obtain for the transmitter noise components falling in the receiver bands. Known pseudo noise pilot signals ($pn_1, pn_2, \cdots, pn_{k-1}$) are therefore added to the transmitter output at each receive frequency. Correlation (XC) is then used to identify the phase and amplitude of the residue pilot error signal ($e_1, e_2, \cdots, e_k$) in the output of each receive band. Suppression of the pilot error signals will also suppress the respective transmitter noise components ($n_1, \cdots, n_{k-1}$), since both pilot and noise travel the same paths. Fig. 4.3 shows the block diagram for a system with $(k-1)$ receive bands. The composite transmitted signal, $u(t)$, consists of the desired transmission, $s(t)$, noise, $n(t)$, and pilots:

$$u(t) = s(t)e^{j\omega_{Tx}t} + pn_1(t)e^{j\omega_{Rx1}t} + \cdots + pn_{k-1}(t)e^{j\omega_{Rx(k-1)}t} + n(t) \tag{4.5}$$

where $pn_1(t), pn_2(t), \cdots, pn_{k-1}(t)$ are pilot signals. The signal entering the receiver, $v(t)$,

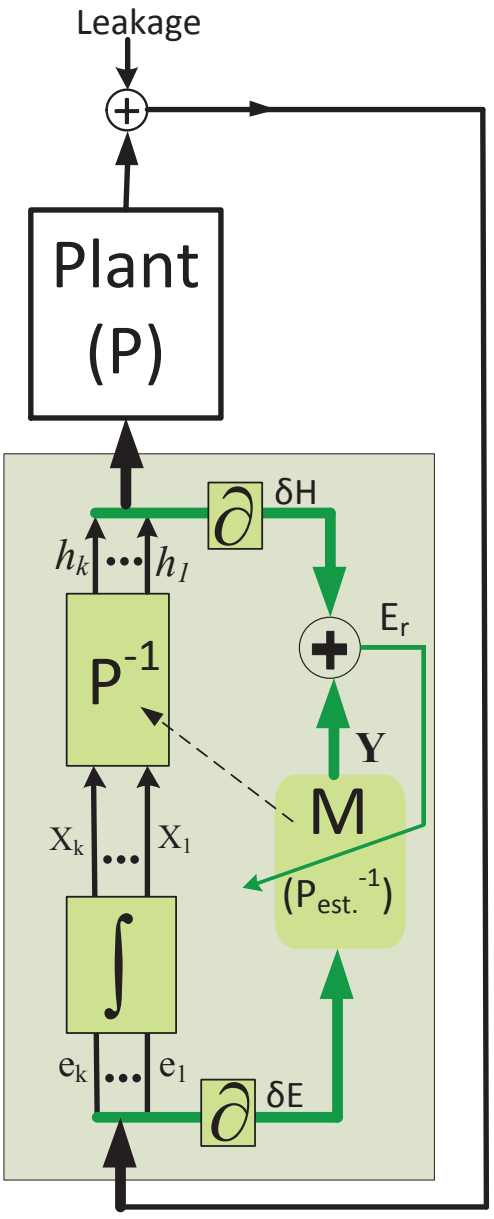


FIGURE 4.2: Modified model of inverse plant control. DSP sections are shaded.

consists of the leakage component and those passing through the delayed coefficients as well as the desired receive signal('s) from the antenna, $r(t)$.

$$v(t) = L(t)u(t-\tau_a) + h_1 u(t-\tau_b) + \cdots + h_k u(t-\tau_n) + r(t) \tag{4.6}$$

where $\tau_a, \tau_b, \cdots, \tau_n$ are the path delays. $L(t)$ is the frequency dependant leakage through the low isolation network.

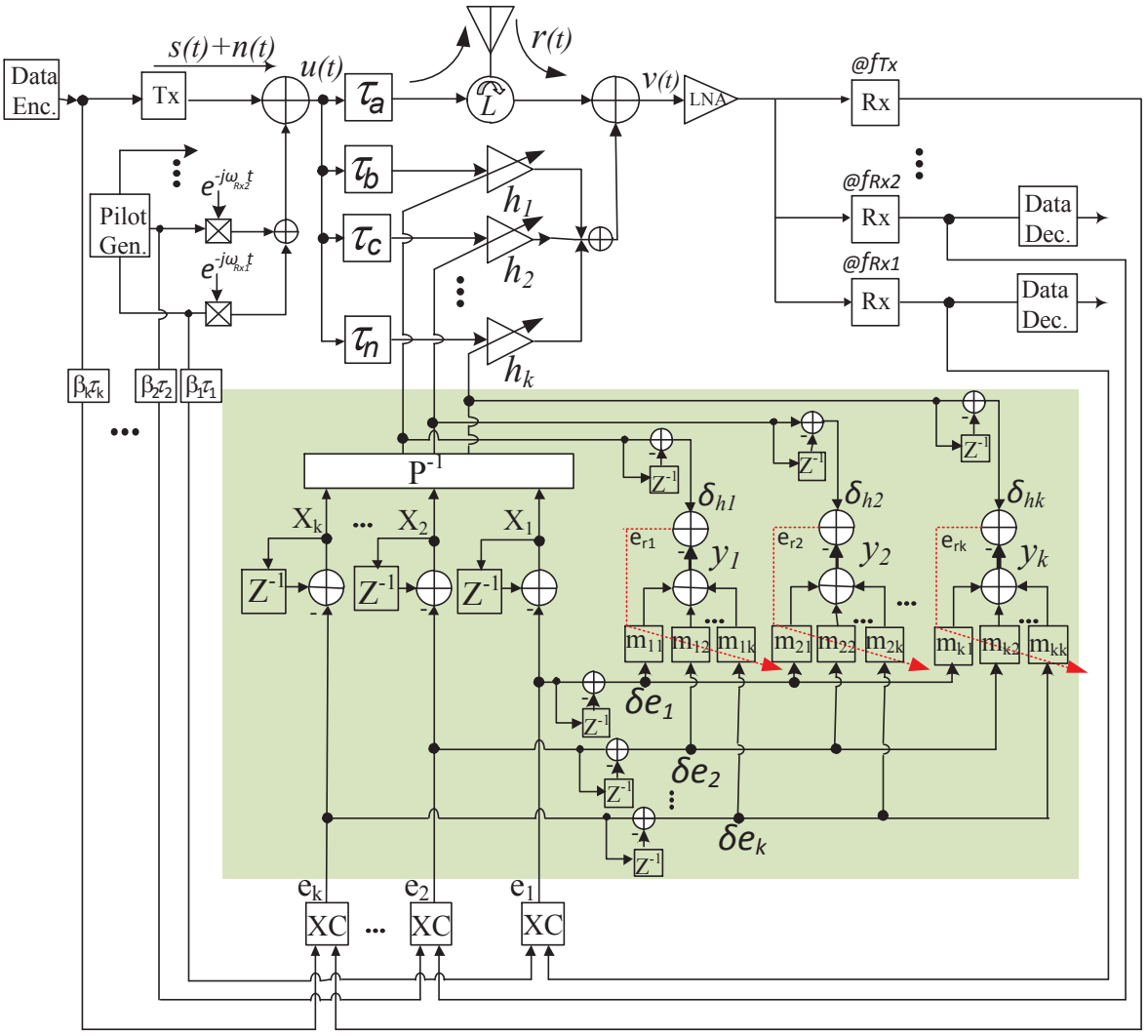


FIGURE 4.3: System block diagram detailed view. DSP control sections are shaded. The plant consists of all circuitry between the control coefficients, H, and the feedback error signals, E.

After the LNA (low noise amplifier) down-converters select and convert to baseband each of the receive bands as well as the transmit band. Cross-correlators isolate the pilot component from each composite receive signal to give the error signals required for the adaptation process.

$$e_1 = E[pn_1^H(t)v(t)] = l_1 + h_1 e^{j\omega_{Rx1}(\tau_a-\tau_b)} + h_2 e^{j\omega_{Rx1}(\tau_a-\tau_c)} + \cdots + h_k e^{j\omega_{Rx1}(\tau_a-\tau_n)} \tag{4.7}$$

$$e_2 = E[pn_2^H(t)v(t)] = l_2 + h_1 e^{j\omega_{Rx2}(\tau_a-\tau_b)} + h_2 e^{j\omega_{Rx2}(\tau_a-\tau_c)} + \cdots + h_k e^{j\omega_{Rx2}(\tau_a-\tau_n)} \tag{4.8}$$

$$\vdots$$

$$e_k = E[s^H(t)v(t)] = l_k + h_1 e^{j\omega_{Tx}(\tau_a-\tau_b)} + h_2 e^{j\omega_{Tx}(\tau_a-\tau_c)} + \cdots + h_k e^{j\omega_{Tx}(\tau_a-\tau_n)} \tag{4.9}$$

where the leakage terms $l_1, l_2, \cdots, l_k$ are assumed constant across a channel bandwidth, but are different between frequency bands. The delays have been normalised to that of the leakage path giving: $\tau_1(=\tau_a-\tau_b), \tau_2(=\tau_a-\tau_c), \cdots, \tau_k(=\tau_a-\tau_n)$, the Matrix form of the above equations is the same as (3.4) but with k degrees of freedom.

$$E = L + \begin{bmatrix} e^{j\omega_{Rx1}\tau_1} & e^{j\omega_{Rx1}\tau_2} & \cdots & e^{j\omega_{Rx1}\tau_k} \\ \vdots & \vdots & \vdots & \vdots \\ e^{j\omega_{Rx(k-1)}\tau_1} & e^{j\omega_{Rx(k-1)}\tau_2} & \cdots & e^{j\omega_{Rx(k-1)}\tau_{k-1}} \\ e^{j\omega_{Tx}\tau_1} & e^{j\omega_{Tx}\tau_2} & \cdots & e^{j\omega_{Tx}\tau_k} \end{bmatrix} H \tag{4.10}$$

where the terms in the bracket form the plant matrix, P. In the 'inverse plant' method instead of controlling the plant, P, by directly adjusting H, an indirect adjustment is used through a model, M, of the inverse plant. This decouples the paths.

$$H = MX \tag{4.11}$$

where $M = P_{est}^{-1}$, an estimate of P^{-1}.

Equation (4.1) is rewritten as:

$$\begin{aligned} E &= L + PMX \\ &= L + IX \end{aligned} \tag{4.12}$$

$$\begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_k \end{bmatrix} = \begin{bmatrix} l_1 \\ l_2 \\ \vdots \\ l_k \end{bmatrix} + \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} \tag{4.13}$$

The linear equation of (4.12) now contains the complex controls $x_1, x_2, \cdots, x_k$ which when multiplied with the matrix M generates $h_1, h_2, \cdots, h_k$ being sent to the GPAs. This combines with the plant to generate the error signals $e_1, e_2, \cdots, e_k$ with no cross-coupling from the new complex controls $x_1, x_2, \cdots, x_k$ described by equation (4.13). The inverse plant enables the two orthogonal outputs to form k SISO linear controllers. Given the correct inverse is applied then a change in one of the controls should cause an equal change in the corresponding output (i.e. by differentiating (4.13);

$$\begin{aligned} \triangle E &= I \triangle H \\ \delta e_k &= \delta x_k \end{aligned} \tag{4.14}$$

Using this property; an integral controller can be used to subtract the current error output e_k from the input x_k to cancel the leakage, l_k in the adaptive duplexer.

$$x_{k,n} = x_{k,(n-1)} - ce_{k,(n-1)} \tag{4.15}$$

where c is the update coefficient and subscript n is the iteration time index.

Assuming the plant remains constant, the inverse is perfect and there is no noise in the system; convergence can be obtained in one iteration by setting $c = 1$ in (4.15). In practise neither is true and so it is necessary to reduce the feedback error signals ($c < 1$) to improve robustness at the expense of convergence. A comparison of the integral controlled inverse algorithm with the 1-d iterative search technique [28] is shown in Fig. 4.4 in the presence of receiver noise.

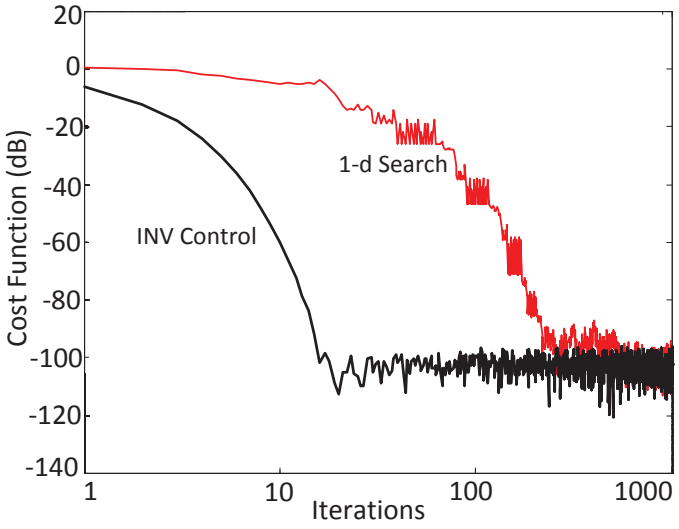


FIGURE 4.4: Inverse with integral control(INV) and one-dimensional search (1-d) algorithms. Conditions for this measurement are as follow: $k = 2, \tau_1 = -1ns, \tau_2 = 1ns, f_d = 45MHz, \beta = 0.5$, and cost function ($CF = \beta e_1^2 + (1-\beta)e_2^2$).

With the effect of cross talk removed, the number of iterations required for the system to converge to -60dB is 10 ($c = 0.5$), while the iterative search algorithm takes 150 iterations. We also note that the inverse algorithm is more stable in low signal to

noise situations. The iterative search algorithms can often jump from the converged solution due to the noise varying between samples, indicating that the cost function has reduced even though the two error components have in-fact increased. The effect is not as noticeable for the integral controlled inverse algorithm because the effective step size (correction) is not defined manually but by the size of the error signals.

The inverse plant for a double loop cancellation system can be manually calculated using elements from the plant divided by the plants determinant. The complexity quickly grows as the order of the plant increases. A kxk inverse requires $O(k^3)$ operations.

$$M = P^{-1} = \begin{bmatrix} \frac{e^{j\omega_{Tx}\tau_1}}{1-e^{-j\omega_D(\tau_1-\tau_2)}} & \frac{e^{j\omega_{Rx}\tau_1}}{1-e^{-j\omega_D(\tau_1-\tau_2)}} \\ \frac{e^{j\omega_{Tx}\tau_2}}{1-e^{-j\omega_D(\tau_1-\tau_2)}} & \frac{e^{j\omega_{Rx}\tau_2}}{1-e^{-j\omega_D(\tau_1-\tau_2)}} \end{bmatrix} \tag{4.16}$$

It can be seen from 4.16 that each element of the inverse plant is dependent on the delay of the antenna or reference path, that relatively defines τ_1 and τ_2, and the Tx and Rx frequency. The Tx and/or Rx frequencies will only change occasionally when, for example, the transceiver changes channels or shifts from one band of operation to another. These adjustments occur with the full knowledge of the transceiver allowing a Look-Up-Table (LUT) of applicable inverses to be stored and used when necessary.

The inverse plant look-up-table approach is only good when the antenna matching is stable such as for a base station. Unfortunately, of more importance, the reference path leakage will vary considerably in the mobile terminal because of the large variation in antenna loading. This will have a negative effect on the orthogonality of the inverse

system. As a result the inverse, $M = P_{est}^{-1}$, will need to be tracked to maintain the correct operation of the integral controller.

4.4 Pilot Control

The transmit signal itself ($s(t)$) is used to generate the error signal, e_k for the transmit band while separate direct sequence spread spectrum (DSSS) low level pilot signals ($pn_1, \cdots, pn_{k-1}$) are injected at the receive band frequencies to generate the remaining the error signals $e_1, \cdots, e_{k-1}$. The pilots have a low power spectral density that is at, or just below, the transmitter noise floor as labelled in Fig. 4.5.

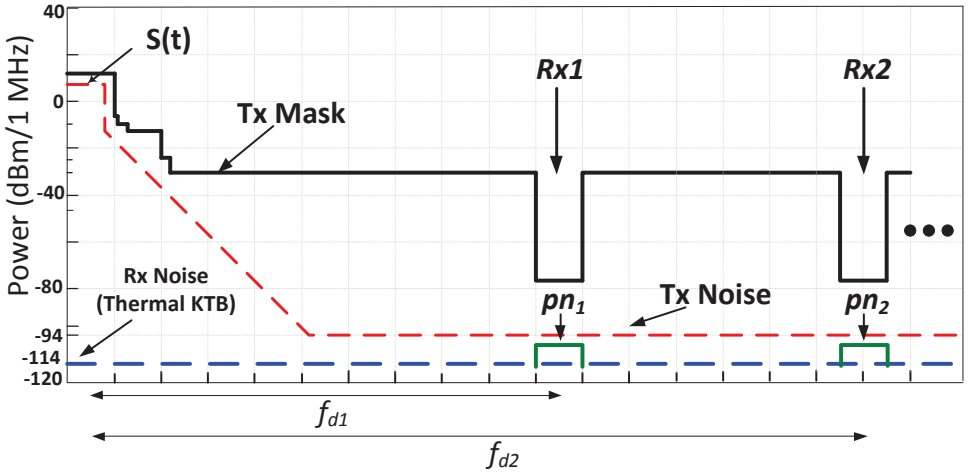


FIGURE 4.5: The transmitter output spectrum. The Tx mask is based on LTE specifications [1].

Both Tx noise and the pilots must be below the spectrum mask shown in black. The DSSS chip rate sets the cancellation bandwidth, which is normally made equal to the received channel bandwidth; for this work the chip rate is 10Mchips/s to set a

20MHz channel [50]. Multi-band reception requires each pilot to be orthogonal to minimise cross-contamination. Maximal length sequences (m-sequence) have good auto-correlation properties and so time shifted versions of the original sequence can be used for different bands. Although a shift of one chip is theoretically enough, it is preferable to use longer time shifts to avoid potential cross-talk from delayed replicas caused by filtering. Multi-chip time shifted m-sequences can be generated without additional shift registers, by using just a few extra XOR gates as illustrated in Fig. 4.6.

A simple algorithm to calculate the tap settings for the delayed sequence was developed by Kai Ping Yiu [51]. By considering $h(D)$ as the generator polynomial, γ as the required delay, n as the number of shift register bits, and $R_\gamma(D)$ as the remainder polynomial from the division of D^γ by $h(D)$ the procedure follows:

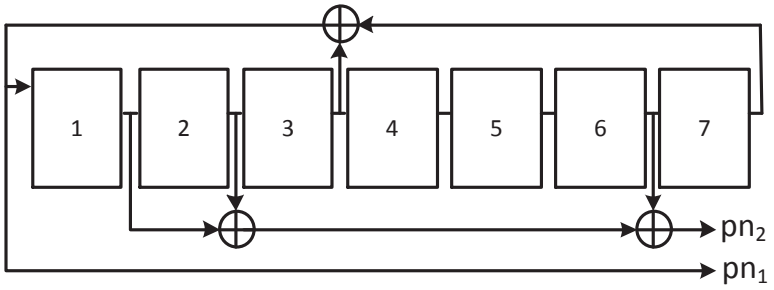


FIGURE 4.6: For this work PN length=127 and pn_3 is a delayed replica of pn_2 with $\beta = 61, \tau = 0.1\mu s$.

- Input: h(X); $D^\gamma \rightarrow R_\gamma(D)$.
- $s = [log_2 \frac{deg(R_\gamma(D))}{n}]$, $R_\gamma(D)\ mod\ h(D^{2s}) \rightarrow R_\gamma(D)$
- If deg $(R_\gamma(D)) \geqslant n$ then go the previous step, otherwise done!

where s is any positive integer, "*deg*" is the polynomial degree.

To get a delay of 61 ($\gamma = 61$), approximately half way along the 127 chip length sequence with generator polynomial $h(D) = D^7 + D^3 + 1$ the method is illustrated as follows:

- s=3, $R_\gamma(D) = D^{29} + D^5 = D^{61} \bmod h(D^6)$
- s=2, $R_\gamma(D) = D^{13} + D^5 + D = (D^{29} + D^5) \bmod h(D^4)$
- s=0, $R_\gamma(D) = D^6 + D^2 + D = (D^{13} + D^5 + D) \bmod h(D)$

which are the connections for pn_2 in Fig. 4.6. Therefore, the size, complexity and power consumption of the pilot sub-system is greatly reduced by avoiding the need for additional pilot generators or shift registers [52], [51].

The final element in the pilot sub-system is the cross-correlator. The cross-correlator, XC, separates e_{k-1} from the Tx noise leakage, n_{k-1}, the receiver noise and the received signal, r_{k-1}. The latter are potential interferers to the convergence process and inverse plant measurement, unless they are adequately suppressed. Greater suppression requires longer XC averaging times and therefore longer PN sequences. Implementing a variable length PN generator will be discussed later in the Thesis.

The goal of the convergence algorithm is to reduce e_k, such that no distortion is caused in the receiver circuits and to reduce $e_1, \cdots, e_{k-1}$ such that any remnant pilot signals (and hence transmitter noise leakage) are below the receiver noise floor. For this to happen the inverse plant must be estimated and tracked, which is discussed next.

4.5 LMS Inverse Control

From the previous sections the combination of a plant with its inverse greatly improves on the convergence time for the adaptive duplexer. However, the plant can vary due to changes in frequency, antenna loading, temperature, supply voltage and other environmental variations. A method is required to find the correct inverse for any given plant and to subsequently track the plant as it changes.

Inverse system identification is used to estimate the inverse plant model, M [20]. This can be done with the LMS adaptive algorithm to off-line calculate P^{-1} by using the plant output E as the *input* to the inverse model, M, and the plant input, H, as the desired reference *output*. The matrix M converges to inverse plant estimate, P_{est}^{-1}, and then the weights are copied across to P^{-1} for use in the integral control loop.

The block diagram of Fig. 4.3 shows the complete adaptive duplexer system with the adaptive inverse estimation and integral control shown shaded. When the inverse plant is correctly adjusted then the input to the GPAs (H) will reach a steady state as the error signals approach zero. As such, rather than the output following a given input, as is the case in many other inverse estimation problems, the adaptive duplexer requires the change in the output to follow the change in the input. A modification to the inverse plant modeller is therefore required. Differential signals now drive the LMS inverse plant modeller (Fig. 4.2). The input to the inverse model at the $n'th$ update is the change in error signal ΔE_n which produces an output estimate Y_n, that should be the same as the desired reference output ΔH_n, which is the change in GPA controls, $\delta h_1, \delta h_2, \cdots, \delta h_{k+1}$. The difference between the desired reference signal and the estimate output gives the

LMS error signal $E_R = [e_{r,1}, e_{r,2}, \cdots, e_{r,k+1}]^T$ for adjustment of the weights, M_n. The weights are then copied to the main control loop $P_n^{-1} = M_n$. This process is outlined in (4.17-4.21):

$$\Delta H_n = \begin{bmatrix} \delta h_1 \\ \delta h_2 \\ \vdots \\ \delta h_{k+1} \end{bmatrix} = \begin{bmatrix} h_{1,n} - h_{1,n-1} \\ h_{2,n} - h_{2,n-1} \\ \vdots \\ h_{k+1,n} - h_{k+1,n-1} \end{bmatrix} \tag{4.17}$$

$$\Delta E_n = \begin{bmatrix} \delta e_1 \\ \delta e_2 \\ \vdots \\ \delta e_{k+1} \end{bmatrix} = \begin{bmatrix} e_{1,n} - e_{1,n-1} \\ e_{2,n} - e_{2,n-1} \\ \vdots \\ e_{k+1,n} - e_{k+1,n-1} \end{bmatrix} \tag{4.18}$$

$$Y_n = M_n \Delta E_n \tag{4.19}$$

$$E_{Rn} = \Delta H_n - Y_n \tag{4.20}$$

$$M_{n+1} = M_n + 2\mu E_{Rn} (\Delta E_n)^H \tag{4.21}$$

The standard LMS algorithm (4.21) is used to adjust the weights of the inverse plant estimation, however, as the main loop converges the change in the GPA controls, ΔH_n,

and the change in the error outputs, ΔE_n, rapidly decrease, prematurely removing the activation signal needed to complete the convergence of the inverse plant, M. It should be noted that the integral controller will still converge even with appreciable cross-talk between the activation signals, $x_1, x_2, \cdots, x_k$ caused by a non-perfect inverse plant. The result is that the scaling on the LMS correction term in (4.21) becomes very small, even though the matrix M has not completely converged. This sensitivity to scaling can be overcome through the use of a NLMS algorithm, a generalisation is shown in [53], which normalises the input vector by its power, boosting the weight update value when the input signal is small:

$$M_{n+1} = M_n + 2\mu \frac{E_{Rn}(\Delta E_n)^H}{\Delta E_n^H \Delta E_n + \varepsilon} \tag{4.22}$$

In order to reduce susceptibility to noise when the input signal, ΔE_n, is very low, a dampening factor, ε, needs to be introduced to limit the normalisation of the input signal [54].

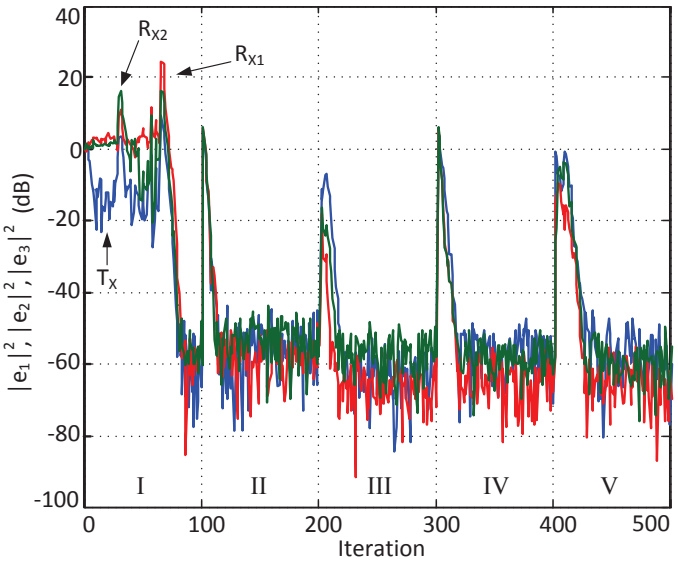
4.6 Simulation Results

The flexibility of this system as opposed to a passive duplexer is that it can operate at any number of frequencies and duplexing offsets with the only restriction coming from the capabilities of the hardware used for implementation. The plot in Fig. 4.7(a) shows how the adaptive duplexer with NLMS inverse control performs under changing conditions.

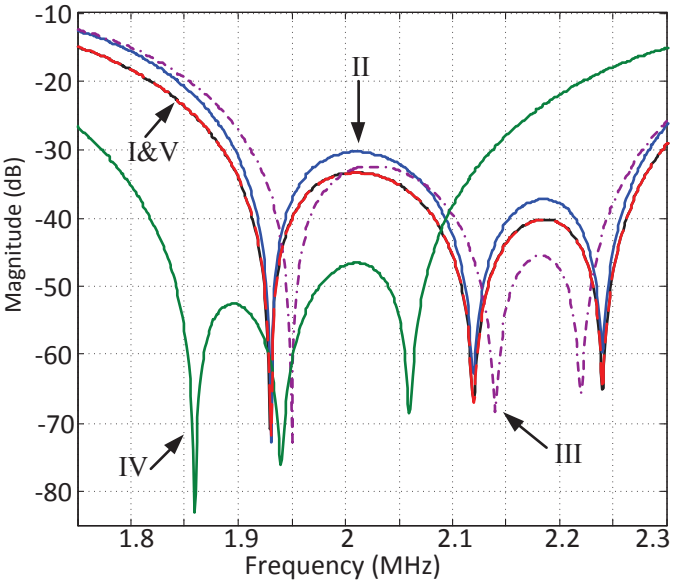
TABLE 4.1: Operating Conditions

	Iterations	Transmit frequency	Duplexing offset; f_{d1}	Duplexing offset; f_{d2}	Leakage Path Delay
I	100	1.93GHz	190MHz	310MHz	τ_a=1.56 ns
II	200	1.93GHz	190MHz	310MHz	τ_a=0.78 ns
III	300	1.95GHz	190MHz	270MHz	τ_a=0.78 ns
IV	400	1.86GHz	80MHz	200MHz	τ_a=0.78 ns
V	500	1.93GHz	190MHz	310MHz	τ_a=1.56 ns

The iteration time is dependant on the length of the PN sequence. For a PN sequence of length 127 chips running at 10Mchips/s, it takes 1.27ms for 100 iterations. Initially the transmit frequency, f_{Tx}, is 1.93GHz and the duplexing offsets, f_{d1} and f_{d2}, are 190MHz and 310MHz for the two Rx bands (f_{Rx1}=2.12GHz and f_{Rx2}=2.24GHz). After 100 iterations the delay in the reference path, τ_a, decreases to 0.78ns requiring a new inverse estimate. After 200 iterations a change in channels occur and after 300 iterations a change in the operating bands. Finally after 400 iterations changes in both band and τ_a occur, reverting the system to the original state (Table. 4.1). The plot in Fig. 4.7(a) shows how the cancelling loops of the adaptive duplexer perform from a convergence point of view while Fig. 4.7(b) displays the frequency response between the transmitter and receiver after each change of operating conditions. It takes about 80 iterations to reach -40dB from 'cold' ($H = [0, 0, 0]^T, P^{-1} = [0]_{k\times k}$) but only between 10 and 20 iterations to handle the changes thereafter. The long initial convergence is because the inverse plant, P^{-1}, must adapt from $[0]_{k\times k}$ and it takes some time to get a condition whereupon the integral control becomes stable. In terms of the spectrum, the nulls are in the correct position; curves I and V overlap each other as they should, and curve II shows a slight reduction in null bandwidth due to a delay change in the signal path.



(a) Dynamic Response. Power of (e_1, e_2 and e_3) vs iteration



(b) Tx to Rx frequency response after convergence

FIGURE 4.7: Adaptive duplexer performance (a) dynamic error vs iteration (b) frequency response after adaptation.

To further verify the operation, the system was simulated with a +20dBm OFDM signal having 20MHz bandwidth and a transmit noise floor of -80dBm/MHz. The low isolation device gives 20dB of initial isolation and receiver noise has been switched off so as not to mask the canceller performance. The resultant power spectrum into the receiver before convergence is given in Fig. 4.8 showing the transmitter signal and

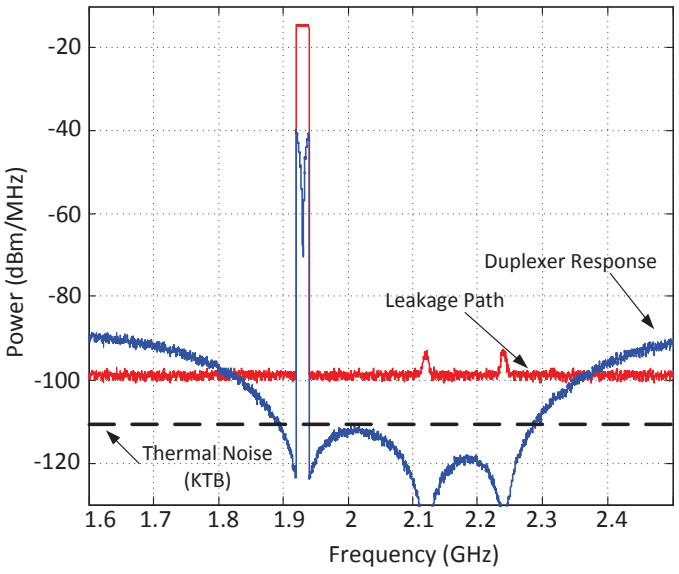


FIGURE 4.8: Rx input spectrum with Rx noise disabled. Before (red) and after (blue) convergence.

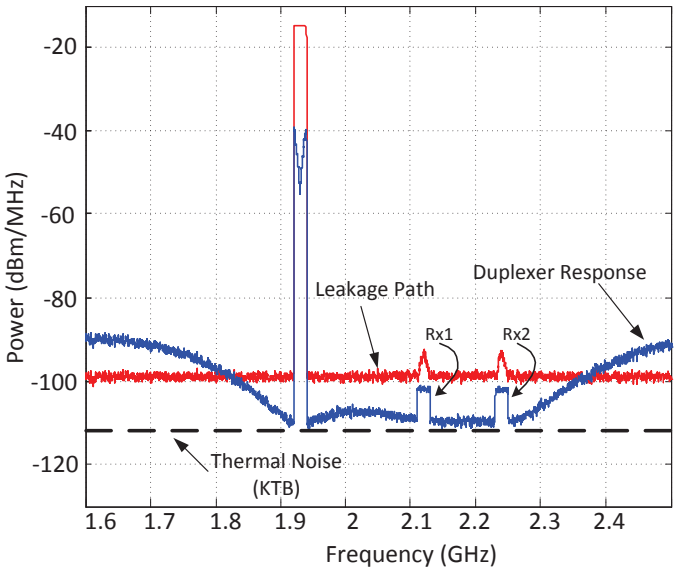


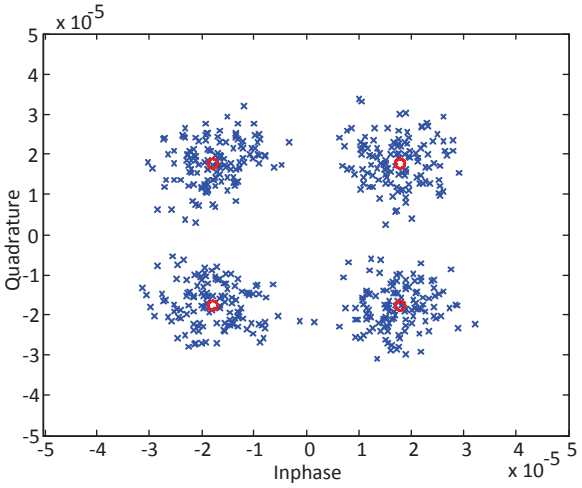
FIGURE 4.9: Rx input spectrum with two aggregated OFDM receive signals.

noise spectrum after the low isolation device. The two pilot signals at (2.12GHz and 2.24GHz) barely show above the Tx noise floor. After convergence the adaptive duplexer gives three nulls, as desired, at the Tx frequency and the two receive frequencies. The additional suppression from the cancelling loops is about 30dB. Note, both pilots and Tx noise have been suppressed below the receiver thermal KTB noise floor.

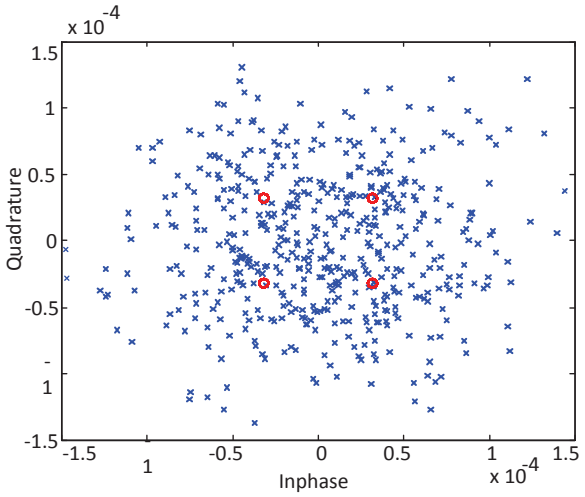
Any noise generated by the receiver that falls within the Rx signal band cannot be

removed, and this determines the practical cancellation limit. Additionally, the adaptive duplexer is required to operate in the presence of the desired receive signals. Fig. 4.9 shows the adaptive duplexer's behaviour in the presence of two 10dB signal to (receiver) noise ratio signals in the receiver bands. These signals are originally masked by the transmitter noise and pilot signals (red curve), but after convergence become completely visible almost to the receive noise limit (blue curve).

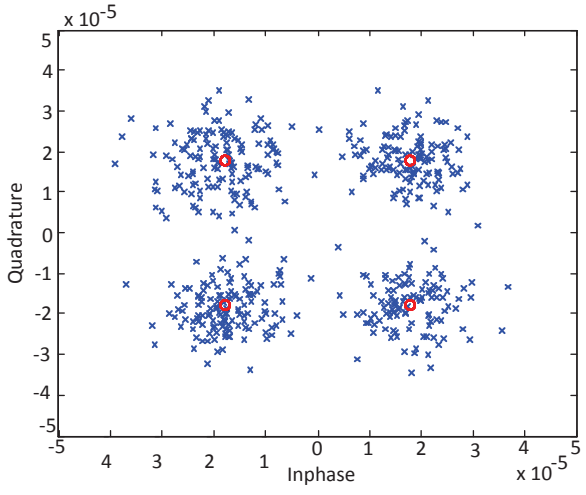
Fig. 4.10 shows the masking effect of the Tx noise and pilot on the desired Rx signals constellation. The ideal situation with both Tx and pilot signals disabled is shown in Fig. 4.10a; only Rx noise is present giving an SINR (=SNR) of 10dB. When the additional interference from the Tx noise and pilots is added, the QPSK signal is completely washed out (Fig. 4.10b) (note the change in scales). After convergence the Rx signal is again recoverable (Fig. 4.10c), but with SINR=7.5dB; a 2.5dB degradation due to non-perfect cancelling of the residual leakage. Increasing the length of the pilot sequences from 127 to 1023 will reduce this implementation loss to less than 1dB. If the receive signals are too large they act as jamming signals to the convergence process, by creating noise on the cross-correlator output error signals, E. The adaptive duplexer therefore has reduced cancellation capability and the received signal's error vector magnitude (EVM) deteriorates. The jammers can be further suppressed by increasing the length of the pilots' PN sequences and correlating over a longer period. Fig. 4.11 shows how the adaptive duplexer performs in the presence of a received signal with increasing signal to noise ratio for different PN sequence lengths. As expected, longer PN lengths produce lower EVMs. Note; the slope of the curves show EVM proportional to SNR at low Rx signal levels, indicating the dominance of receiver noise (Tx noise has been



(a) Rx noise only (SNR=10dB)



(b) Before convergence



(c) After convergence

FIGURE 4.10: QPSK constellation diagram for received OFDM signal, in the presence of Rx noise, Tx noise and pilot.

TABLE 4.2: Performance Comparison

Ref.	Loops	Channel Band-width	Bands	Leakage Suppres-sion	Convergence Time	Features
[16]	3	10MHz	1	60dB	0.3ms	Analog LMS-like
[14]	16	80MHz	1	45dB	1.0ms	LUT + Gradient Descent
This Work	3	20MHz (per band)	3	30dB	From Cold 1ms to 8ms Tracking 0.2 to 1.5ms	$\mathbf{P}^{-1}$ Modelling, NLMS, 2 Pilots

suppressed). For high values of Rx signal (SNR>25dB), the curves start to floor as the suppression of the Tx noise degrades due the jamming effect. Any increase in signal is now matched by an equal increase in Tx noise leaving EVM unchanged. The PN sequence length determines the EVM floor level. A length of 1023 chips gives an EVM floor of -25dB; enough to support 64QAM signals. The PN length is a trade-off between the pilot's ability to suppress interference from the error signals and the time it takes to converge. The 1023 length m-sequence would take 10ms to converge from cold and 1.5ms for small adjustments, still fast enough to track changes in antenna loading.

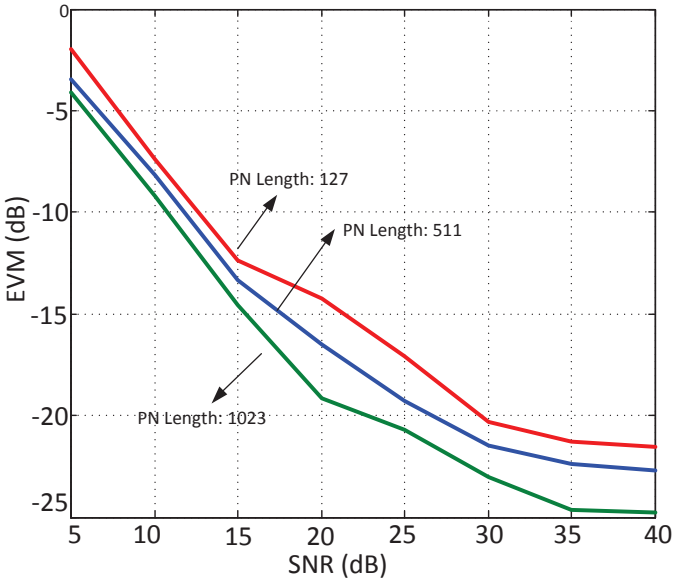


FIGURE 4.11: Receive channel EVM versus SNR for different PN length.

Table 4.2 compares two schemes with similar architectures to Fig. 4.1(a) but targeting the single frequency full duplex problem (RF component). These schemes concentrate all their degrees of freedom (nulls) into one band compared to three disjoint bands (for our FDD), which explains their improved cancelling performance. They also have strong signals on which to adapt, compared to pilots operating with negative SNRs, which mean they require less averaging. Despite these disadvantages our scheme obtains comparable convergence times particularly for small adjustments. The super fast convergence of [16] uses an all analog LMS-like solution which comes at the expense of extra hardware in the form of nine additional down-converter stages, while [14] uses a combination of look-up tables (LUT) and linear programming followed by a gradient search for the final adjustment. Our scheme is competitive, particularly for fine adjustments, but the need for long averaging of the pilots at high QAM levels makes the convergence from cold relatively slow. The use of a look-up table for $\mathbf{P}_{\mathbf{est}}^{-1}$ to start the process would make an improvement.

4.7 Summary

An adaptive duplexer for carrier aggregated systems has been described. It uses a combination of a broadband low isolation device and multiple cancelling loops to provide tunable nulls in the transmitter to receiver transfer function. The scheme provides up to 50dB of cancellation for a number of 20MHz bandwidth channels. The nulls are tuned by adjusting the loop coefficients using look-up tables or by using dynamic feedback control. An inverse plant model orthogonalises the control problem enabling each null to be independently controlled. Convergence from 'cold' takes 80 iterations using an

N-LMS algorithm, however this drops to about 15 iterations when tracking changes in one variable. Co-channel interference between pilot signals and the received signal limits the achievable receiver EVM. Improved EVMs require longer pilot sequences at the expense of slower convergence times. Simulations show that a 1023 chip pilot has an EVM limit of -25dB, enough to support 64QAM OFDM signals while needing 1.5ms to track small changes, such as variations in antenna loading.

As far as implementation is concerned the DSP adaptation (inverse plant model and integral control) can be implemented in a programmable device since the sample rate (f_{DSP}) is low compared to the channel band width (B_{ch}).

$$f_{s,DSP} = \frac{f_{chip}}{PNLength} = \frac{B_{ch}}{2 \times PNLength} \tag{4.23}$$

The pilot generator circuits operate at a much higher sample rate and have to be designed properly to get a low power sub-system. This implies an ASIC (Application Specific Integrated Circuit) solution. The pilot generator is implemented on Silicon on Sapphire (SOS) technology since this is the technology chosen for the RF gain-phase adjusters in the cancelling loops.

The next chapter discusses the development history of the SOS process and the features of the Peregrine company's UltraCMOS© process.

Chapter 5

Silicon-On-Sapphire Technology

5.1 Introduction

The previous chapter described the need for pilot signals to control the adaptation algorithm. These signals are fed directly into the RF signal lines and so should be co-located on the RF sub-system. Hence the choice of SOS process technology. In this chapter, Silicon on Insulator (SOI) and Silicon on Sapphire (SOS) processes are briefly described. They are a wafer substrate technology suitable for RF applications such as mobile phone base stations, smart phones, satellite communications, and mixed-mode IC designs. The technologies are replacing GaAs and the bulk CMOS processes in medium power RF circuits. SOS is part of the SOI family of CMOS processes, and provides all the benefits of SOI in addition to a few other advantages. The radiation hardened property of sapphire makes it appropriate for space and military applications and the transparency of the substrate makes it suitable for electro optical devices. Recent trends and growing demand for low cost RF circuits in portable devices has expanded the demand for

SOS solutions. SOS is now replacing expensive GaAs [55], [56] however the process is costly with respect to bulk CMOS.

Technology issues will be discussed in Section 5.2. Some essential Peregrine Ultra-CMOS process features are summarised in Section 5.3 and 5.4. Finally a summary is provided in Section 5.5.

5.2 Technology Issues

SOS devices have specific electronic applications because they can reduce noise and current leakage in MOSFET devices. SOS uses synthetic sapphire (Al_3O_2) as an insulating substrate [57]. Cross-sections of a CMOS inverter implemented in bulk CMOS, Peregrine SOS, and SiO_2 SOI processes are presented in Fig. 5.1. The SOS process, by taking advantage of its substrate, results in a simple structure and therefore there is a reduction in the number of fabricated masks compared with the bulk CMOS process. SOS MOSFETs are formed by growing a thin silicon epitaxial film on top of the sapphire substrate. Combining the well established CMOS technology with sapphire enables better RF performance from low cost manufacturing equipment.

The bottom cross section of Fig. 5.1 illustrates an inverter in SiO_2 SOI process. A thin film of SiO_2 separates the active circuits from a standard silicon substrate [21]. This improves isolation between devices compared to bulk CMOS, but not compared to SOS.

The main advantage of SOI is the insulation between the active and substrate area which increases the quality factor for passive devices and the performance of active

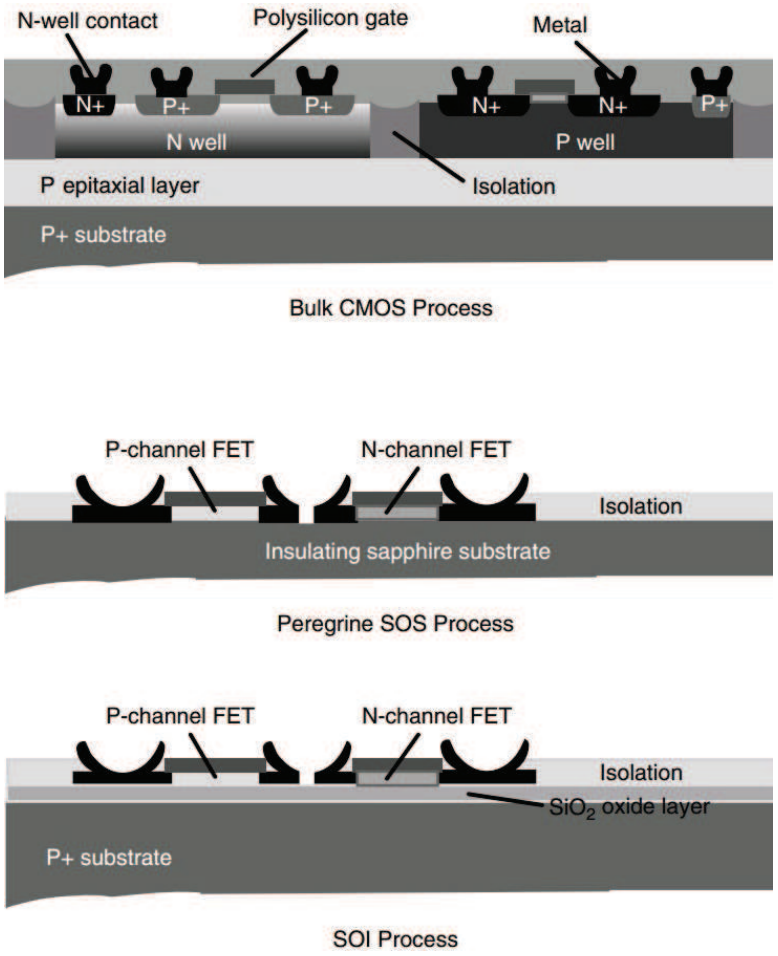


FIGURE 5.1: implementation of a CMOS inverter using bulk CMOS, SOS and SOI processes [21].

components by reducing coupling unto the conductive P^+ substrate. This advantage is enhanced in SOS devices because the substrate is completely replaced by a non-conductive substrate. SOS technology is not without its limitations. A downside of the SOS process is the minimum (now large) feature size of 250 nm which compares to < 15nm for today's CMOS. The achievable gain-bandwidth product, f_T, of SOS is therefore limited. In addition the "kink effect" phenomenon occurs because of the floating body. Impact ionization can occur at high drain voltages which then leads to an increase in the substrate potential which reduces the threshold voltage causing an increase in the Id current. A plot of a MOSFET I-V characteristic illustrating the kink

effect is shown in Fig. 5.2. Compared with SiO_2SOI, SOS devices are less prone to this effect leading to a reduction in the magnitude of the ”kink effect”, and an increase in the drain voltage that will trigger its onset [58]. The SOS process suffers from low channel mobility, and has poor thermal conductivity; almost one half that of silicon. Fabrication of sapphire wafers is more complex and 5 to 7 times more costly than the fabrication of silicon wafers [59].

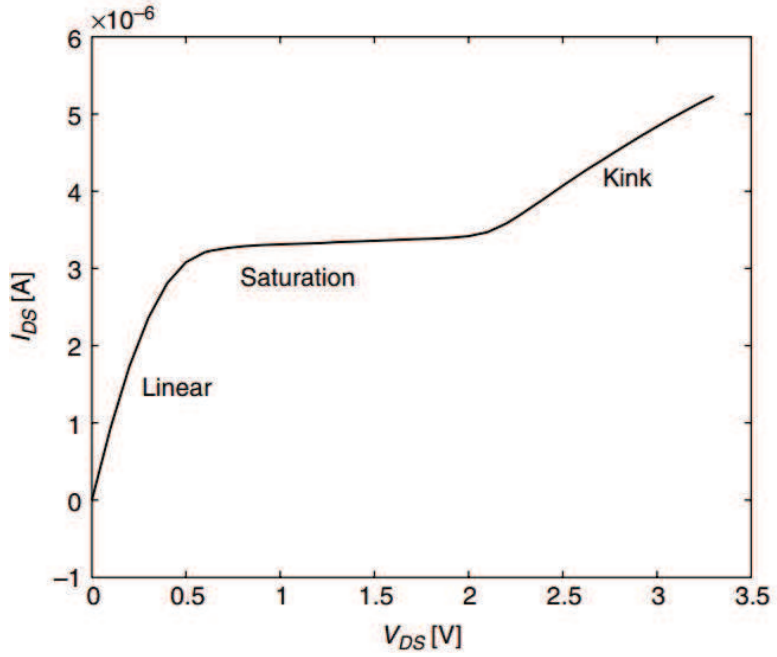


FIGURE 5.2: ”kink effect” from real data collected on a RN 1.2×5μm SOS device with VGS = 1V [21].

5.3 Peregrine UltraCMOS® Process

Peregrine semiconductor corporation markets its proprietary SOS process as UltraCMOS®, and its cross section is shown in Fig. 5.1. The process options are summarised in Table 5.1. The foundry process for this project is the GC 0.25μm SP3M (single poly triple metal). Metal-Insulator-Metal (MIM) capacitors are available in the GC process only;

a silicon dioxide dielectric layer is sandwiched between the top metal and the bottom plate. Further, the kink effect in the Peregrine process is strongest in the HN, and HP transistors, but is present at a much lesser degree in the RN, RN-RF, RP, IN, IN-RF and IP transistors. However the kink effect is not a problem for digital design where the increase in current makes the circuit faster [21]. The designed circuit in the next Chapter is predominantly digital.

TABLE 5.1: Technology Options of UltraCMOS® SOS Process [23]

Features Summary	Units	Process Variants					
		FA	FC	GA	GC	PA	PC
Generation		0.5 μm		0.25 μm / 0.35 μm		0.25 μm	
Release Status		Production		Production		Qualification	
Application		Com/Auto/Mil		Com/Auto/Mil		Com/Auto/Mil	
Supply Voltage	V	3.0/3.3		2.5/3.3		2.5/3.3	
Transistor Vts		3n/3p	3n/3p	3n/3p	3n/3p	3n/3p	3n/3p
Resistors		2n/1p/1polycide		2n		2n	
Interconnect Layers		3	3*	3	3*	3	3*
MIM Caps		No	Yes	No	Yes	No	Yes
Inductors		No	Yes	No	Yes	No	Yes
f_t (IN Device)	GHz	15		30		30	
f_{max}	GHz	45		90		90	

* Last layer interconnect is a thick metal layer for High Q Inductor construction

5.4 Digital Circuit Design in UltraCMOS®

The digital SPICE model parameter files include six device types: RN, RP, IN, IP, HN, HP, TIN, TRN, and TRP, and the specifications are shown in Table 5.2.

The simulation results of an inverter with NMOS device threshold (0.67V) for RN and PMOS devices threshold (0.7V) for RP are reported. Fig. 5.3.a shows the DC and AC simulation results of an inverter with both NMOS and PMOS devices with

TABLE 5.2: Transistors type descriptions for FC and GC Processes [23].

		Threshold Voltage (V)	Application
FC	RN	0.67	Digital and low leakage
	NL	0.13	High-performance digital and RF
	IN	-0.20	High-performance digital and RF
GC	RN	0.43	Digital and low leakage
	HN	0.70	Digital and low leakage
	IN	0.08	High-performance digital and RF

size [Width (W), Length (L)] of [5μm, 0.5μm]. The DC simulation shows a switching threshold of 1.6V. Fig. 5.3.b shows the maximum voltage gain is 3.1V/V at 1.6V DC input, and the bandwidth (f_T) at voltage gain of 1 is approximately 5 GHz. Fig. 5.3.c shows transient simulation results of the same inverter when the input frequency is at 1.6GHz.

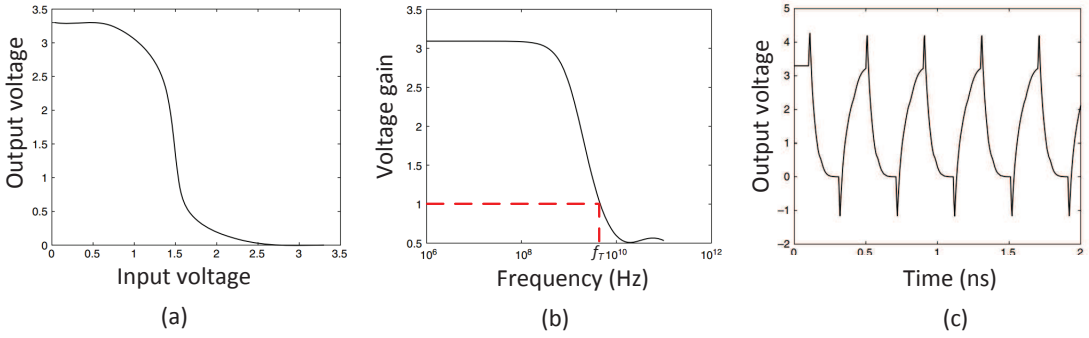


FIGURE 5.3: (a)DC simulation of SOS inverters., (b) AC simulation (c) Transient simulation [21].

5.5 Summary

The SOS process, with its insulating substrate, allows us to design isolated circuits in a single chip. Sapphire, as an insulating substrate, provides better isolation between circuit elements. The sapphire substrates have several distinct advantages such as high isolation between digital and analog RF circuits and low leakage current for transistors.

The substrate's high resistivity also decreases parallel substrate capacitance which has the effect of increasing the quality factor of integrated inductors. SOS fabrication is more complex and costly than silicon, but still much cheaper than GaAs. As such, its main application to-date has been the low cost replacement of GaAs devices in RF and microwave circuits.

As mentioned, in this work the GC process 0.25μm with RN, RP, HP and HN transistors for better RF switching performance are used. The pilot subsystem design using the above process is described in the next Chapter.

Chapter 6

Pilot Subsystem: System and Circuit Level Design

6.1 Introduction

The proposed pilot signal is generated at the receive frequency and is injected into the transmitter path at a level below the transmitter noise floor to measure the Tx noise falling in the Rx band. In this way interference with the incoming received signal is minimised. This chapter describes the transistor level design, post-layout simulation results, and the fabricated chip using the Peregrine SOS 0.25 μm GC process. A mathematical model of the pilot generator for a double loop cancellation is also reviewed.

Design overviews and the circuit block diagrams including mathematical model are presented in Section 6.2. Section 6.3 describes the main circuit implementation. An RF modulator circuit is discussed in Sub-section 6.3.1. Control blocks including switch

controls, Schmitt trigger as well as the ESD protection circuits, are discussed in Subsection 6.3.2. Section 6.4 presents the post-layout simulation results and finally the experimental work is presented in Section 6.5.

6.2 Pilot Generator Design overview

Fig. 6.1 shows a detailed block diagram of the signal and pilot processes needed to obtain the error signals e_1 and e_2 for a two loop (k=2) cancellation model. A direct sequence spread spectrum (DSSS) low level pilot, $p(t)$, is generated as a reference signal. It is up-converted to the receiver band frequency and added to the transmitter signal, $s(t)$, to give the composite output signal $u(t)$. Most of the signal, $u(t)$, is transmitted from the antenna but because of inevitable feed-through in the low isolation device,

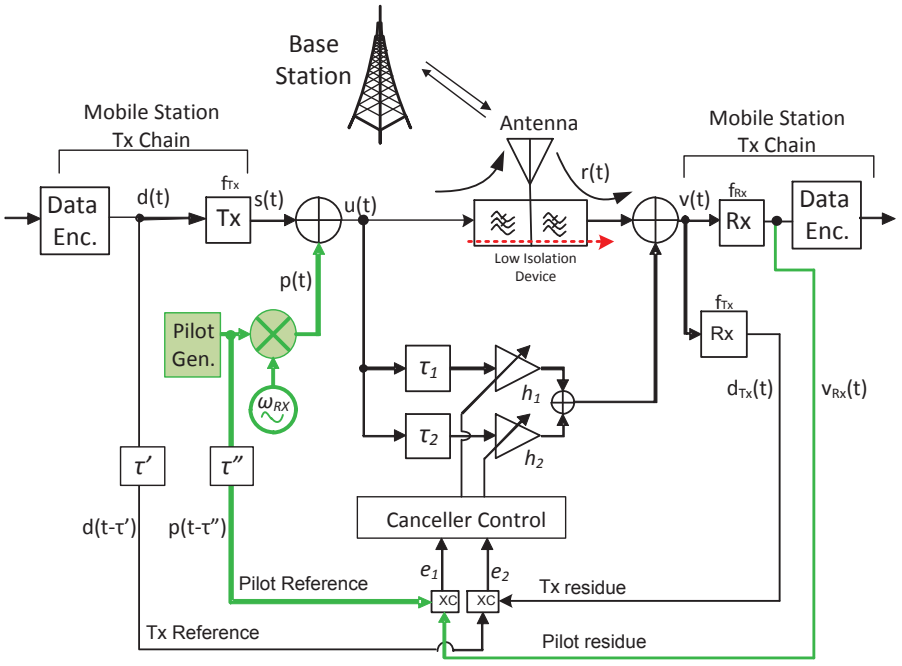


FIGURE 6.1: System Block Diagram. Shaded blocks are implemented in the Peregrine SOS process.

some of this signal leaks into the receiver circuits. Therefore, the output of the receiver, $v_{Rx}(t)$, consists of the wanted Rx signal, the Tx noise (falling in the Rx band), and the pilot signal. Cross correlating the reference pilot signal, $p(t - \tau'')$, with the Rx signal leads to error signal, e_1, which measures the amount of pilot residue in the Rx signal. The purpose is to use the pilot to adjust the feed forward coefficients to cancel both pilot and transmitter noise. The error signal e_2 represents the amount of Tx residue signal (in the Tx band) and is detected by a separate receiver and cross-correlation process.

Note, there is a need to delay the pilot and Tx reference signals by τ'' and τ' in order to match delays in the receivers caused by the low isolation device and receiver processing. Channel filtering is the dominant cause of these delays.

Cross correlation gives both amplitude and phase of the PN pilot residue and the transmit residue signals. The equations that explain the system processes in Fig. 6.1 are described in complex envelope notation. The analysis broadly follows Chapter. 4 equations (4.5–4.9). The transmitted signal, $s(t)$ is given by

$$s(t) = d(t)e^{-j\omega_d t} + n(t) \tag{6.1}$$

where $d(t)$ is the transmitted baseband data signal, ω_d is the duplex offset and $n(t)$ is the Tx noise. Note, the reference frequency for the complex envelop is set to the receive frequency, ω_{Rx} and so any transmit signal is rotating at the duplex offset frequency, f_d. After adding the pilot, $p(t)$, the Tx signal, $u(t)$, becomes:

$$u(t) = s(t) + p(t) = d(t)e^{-j\omega_d t} + n(t) + p(t) \tag{6.2}$$

The signal radiated from the antenna is $u(t)$. The composite signal, $v(t)$, at the receiver input becomes:

$$\begin{aligned} v(t) &= (l_1 + h_1 e^{j\omega_{Rx}\tau_1} + h_2 e^{j\omega_{Rx}\tau_2})(n(t) + p(t)) \\ &+ (l_2 + h_1 e^{j\omega_{Tx}\tau_1} + h_2 e^{j\omega_{Tx}\tau_2}) d(t) e^{-j\omega_d t} + r(t) \end{aligned} \tag{6.3}$$

where τ_1 and τ_2 are the delays of the cancellation paths with respect to the main path. A key assumption is that the delays τ_1 and τ_2 are small compared to 1/the signal bandwidth, in which case the dominant distortion is a fixed phase shift dependent on the signal frequency. Then $v(t)$ is down-converted to the baseband and filtered to the signal bandwidth prior to cross-correlation to give the error signals e_1 and e_2. The complex baseband output signals of the two down-conversion receivers become:

$$\begin{aligned} v_{Rx}(t) &= v(t) \\ &= (p(t - t^{''}) + n(t - t^{''}))(l_1 + h_1 e^{-j\omega_{Rx}\tau_1} + h_2 e^{-j\omega_{Rx}\tau_2}) + r(t) \end{aligned} \tag{6.4}$$

$$\begin{aligned} d_{Rx}(t) &= v(t) e^{j\omega_d t} \\ &= d(t - t^{'})(l_2 + h_1 e^{-j\omega_{Tx}\tau_1} + h_2 e^{-j\omega_{Tx}\tau_2}) e^{j\omega_d t} \end{aligned} \tag{6.5}$$

The low pass filtering in the two receivers delay the base-band waveforms, by $t^{'}$ for the Tx residue receiver and $t^{''}$ for the main mobile station receiver. The low level pilot signal, $p(t - t'')$, falls within the receive band and appears in the baseband of the receive

signal, $v_{Rx}(t)$. The cross-correlator output, e_1, measures the residual pilot buried in the received signal, $r(t)$ and any remaining noise. The goal is to reduce e_1, such that pilot and noise levels are below the received signal noise floor.

$$\begin{aligned} e_1 &= E\left\{v_{Rx}(t)p^*(t-\tau'')\right\} \\ e_1 &= E\left\{((p(t-t'')+n(t-t''))(l_1+h_1e^{-j\omega_{Rx}\tau_1}+h_2e^{-j\omega_{Rx}\tau_2})+r(t))p^*(t-\tau'')\right\} \end{aligned} \tag{6.6}$$

The noise, $n(t)$, and receive signal, $r(t)$ are uncorrelated with the pilot, $p(t)$ so the error signal becomes:

$$e_1 \approx (l_1+h_1e^{-j\omega_{Rx}\tau_1}+h_2e^{j\omega_{Rx}\tau_2})R_{pp}(t''-\tau'') \tag{6.7}$$

The pilot (auto) correlation, $R_{pp}(t''-\tau'')$, peaks and the detector is at its most sensitive when $\tau'' = t''$. The approximation is because the expectation is taken over a limited time period. Similarly for e_2 we have:

$$\begin{aligned} e_2 &= E\left\{d_{Rx}d^*(t-\tau')\right\} \\ e_2 &= E\left\{(d(t-t')(l_1+h_1e^{-j\omega_{Tx}\tau_1}+h_2e^{j\omega_{Tx}\tau_2})d^*(t-\tau')e^{j\omega_d t}\right\} \\ e_2 &\approx (l_2+h_1e^{-j\omega_{Tx}\tau_1}+h_2e^{j\omega_{Tx}\tau_2})R_{dd}(t'-\tau') \end{aligned} \tag{6.8}$$

which is at its most sensitive when $\tau' = t'$. It is now possible to adjust h_1 and h_2 to minimize both e_1 and e_2 concurrently using the algorithm of Chapter 4.

The delays t' and t'' are dominated by the receiver selectivity and therefore will vary with the channel bandwidth and the order of the receiver filters. For example, τ' and

τ'' will vary when the transceiver switches between 5MHz bandwidth WCDMA and 20MHz bandwidth LTE, but apart from that will remain relatively stable. Potential solutions for tracking τ' and τ'' exist and include LUT, delay locked loops and many others [60]. The best choice will depend on the stability and required accuracy with which τ'' should be matched to t''. As always, practical consideration such as circuit complexity and power consumption require the fidelity of τ'' as coarse as possible without overly compromising performance. Any loss in measurement sensitivity would have the effect of increasing the EVM floor by the same amount (Fig. 4.11). A loss of up to 1dB can probably be tolerated. The next sub-section considers this issue.

6.2.1 Accuracy of τ''

The pseudo noise signal is produced from an m-sequence generator based on a polynomial equation [60]. Its white noise-like spectrum and its good autocorrelation qualities are the most important characteristics for this project. For a pilot signal based on a maximal length sequence the auto-correlation is:

$$R_{pp}(\delta) = \frac{1}{NT_{chip}} \int_{t=0}^{(N-1)T_{chip}} p(t)p(t-\delta)dt \tag{6.9}$$

Where N is the sequence length. By mapping each binary zero to -1, the autocorrelation for the output waveform would be unity for $\delta = 0$ and $-1/N$, if δ is more than T_{chip} out of phase in either a positive or negative direction. The autocorrelation

for a 127 chip length m-sequence is shown in Fig. 6.2. The shape of the autocorrelation function between $-T_{chip}$ and $+T_{chip}$ is triangular, centred on $\delta = 0$. If the adjustment of δ in (6.9) is not correct, then there is a loss of measurement sensitivity given by $20log10((T_{chip} - \delta)/T_{chip})$. A 1dB loss in measurement sensitivity would require $\delta < 0.11T_{chip}$. By choosing $\delta = 1/16$ chip a maximum loss of 0.56dB would be required. To do this the PN sequence clock is oversampled by 8 to give delay resolution of $1/8T_{chip}$. The maximum delay error is therefore half of this. The total adjustment range is set by the minimum bandwidth of the received signal. The total range of adjustable delay varies between a maximum advance of $8T_{chip}$ to a maximum delay of $8T_{chip}$. In the simulation of Chapter 4, the range is enough to handle channel bandwidth down to 10MHz.

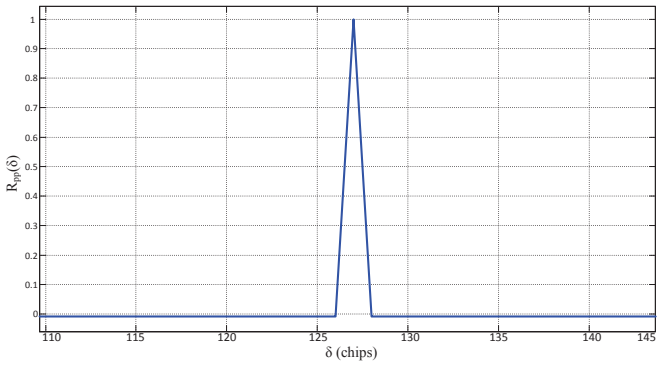


FIGURE 6.2: Autocorrelation of a n =7 register PN generator (127 chip).

6.3 Circuit Implementation

A full custom circuit which includes a variable length PN-sequence generator with reference and delayed outputs and accompanying RF up converter is fabricated using the Peregrine SOS 0.25 μm GC process. A Linear Feedback Shift Register (LFRS) with

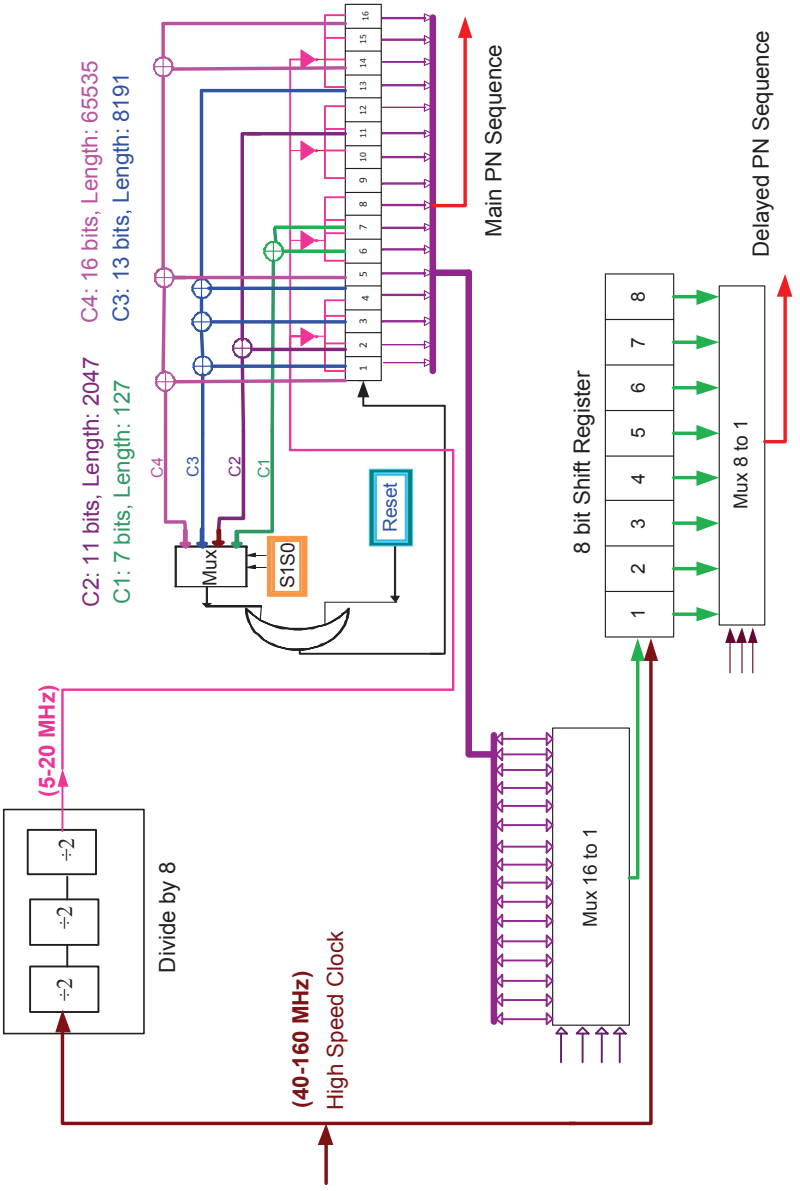


FIGURE 6.3: Block diagram of pilot generator logic.

specific feedback settings spanning n stages forms an economical, fast, and efficient method for pilot generation. The length of a PN sequence before repetition occurs depends on the feedback taps that are chosen and the initial state of the register. Table. 6.1 shows the polynomial equations used for this circuit. The all zero state must be avoided. By choosing the proper feedback taps a maximal length PN sequence, $N = 2^n - 1$ chips results. Prior works describe the tap settings of the LFSR for maximum length [61]

TABLE 6.1: Proposed programmable PN sequence equations.

Feedback Paths	Polynomial equations	Length
C1	$D^7 + D^6 + 1$	127
C2	$D^{11} + D^2 + 1$	2047
C3	$D^{13} + D^4 + D^3 + D^1 + 1$	8191
C4	$D^{16} + D^{14} + D^5 + D^1 + 1$	65535

The block diagram of the pilot circuit shown in Fig. 6.3 consists of: digital circuits; conventional falling edge D flip-flips (DFF); multiplexers; and logic gates. There are two outputs: the main which is upconverted to f_{Rx} in the modulator; and the delayed output, which can be adjusted for different advances or delays used in the cross-correlation. The proposed circuit is programmed for only four different lengths and up to ±8 delays to keep the circuit complexity low. Values for n are between 7 and 16 based on the selection of C1, C2, C3, or C4, feedback combinations given by Table. 6.1 and shown in Fig. 6.3. The shortest sequence is therefore 127 and the longest is 65,535. The length is a trade-off between the ability of the pilot to suppress interference from the receive signal in the control algorithm and the time it takes to converge.

PN sequence generators normally fall in the all zero state, and to prevent this, a narrow pulse is ORed with the feedback taps: C1, C2, C3 and C4. The divide-by-2 block is implemented using a conventional static ”D” Flip Flop (DFF) [2]. To provide bandwidths between 10-40MHz with a resolution of 1/8 chip, the operating input clock frequency of the DFFs has to be at least 40-160 MHz.

The layout of the pilot generator is shown in Fig. 6.4, and the post layout simulation results will be discussed in section 6.4.

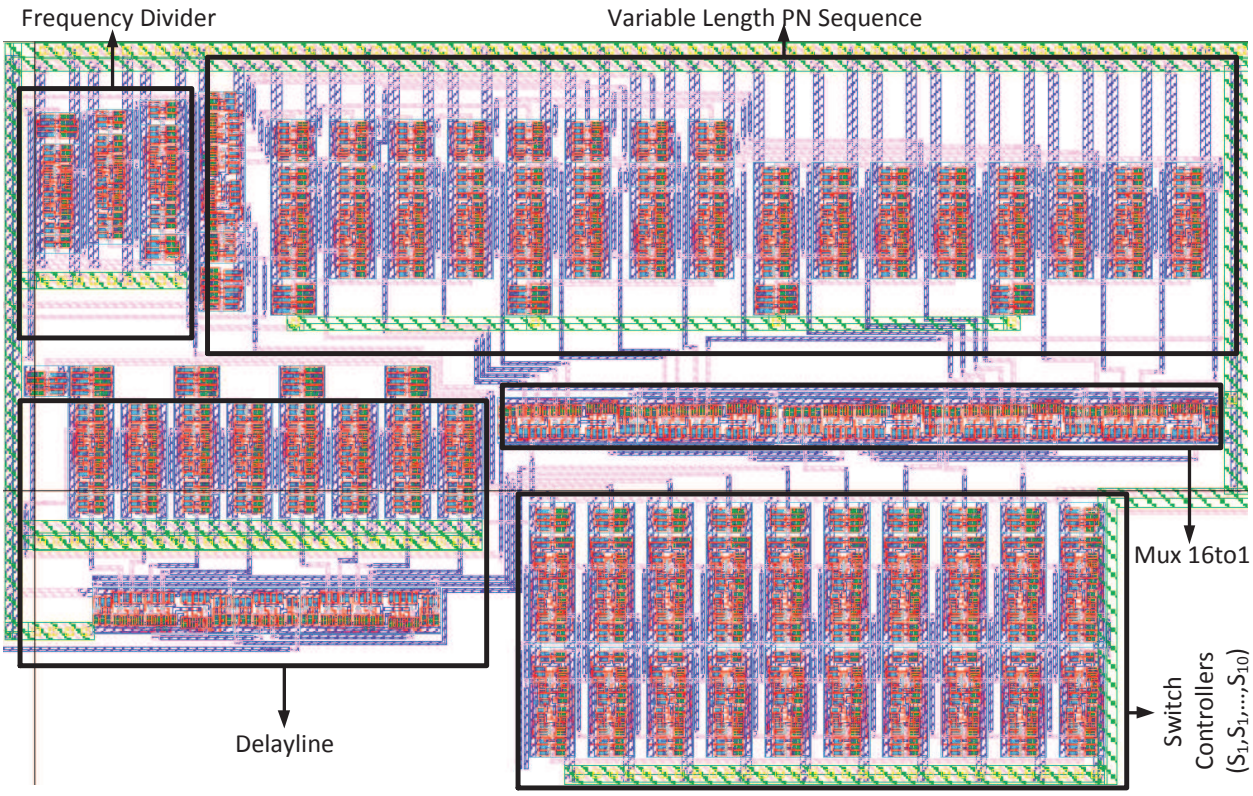


FIGURE 6.4: Layout of pilot generator circuit.

6.3.1 RF Modulator Circuit

The RF modulator circuit, used to up-convert the pilot signal to the receive frequency band, is shown in Fig. 6.5. The reference PN sequence signal, and its anti-phase signal are driving the gates of T_1 and T_2 hereby modulating the oscillator signal at f_c. The advantages of using a switched mixer for this application are its straightforward design, low power consumption and reduced occupied area compared with other mixer types.

The carrier signal is attenuated by Z1 and Z2 networks before it reaches the mixing stage to mitigate carrier feed through. After mixing, the signal is summed with the Tx output at a high impedance node (Zout) to attenuate the modulated signal even more and avoid loading the Tx signal.

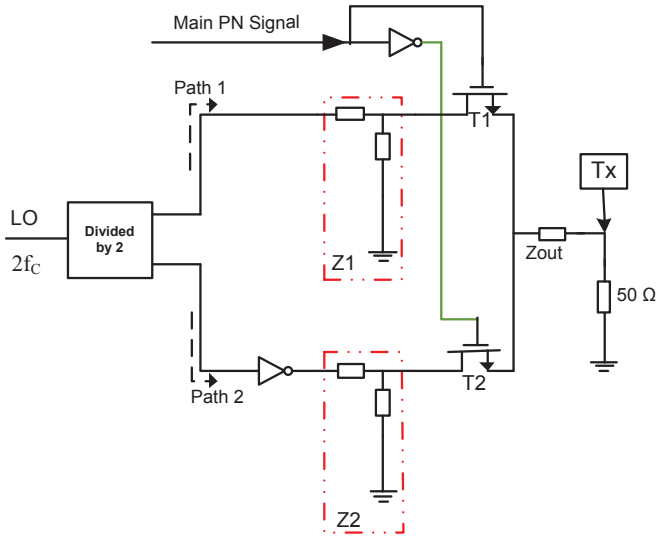


FIGURE 6.5: Proposed RF modulator.

The low output level means the pilot is particularly sensitive to carrier leakage generated by the up-conversion process. Local oscillator leakage components from direct conversion up-converters fall directly in the middle of the transmit channel. In some communication systems such as GSM, direct voltage controlled oscillator (VCO) modulation is used. The VCO generates a phase modulated RF signal without need for an up-conversion mixer [62] and so has no carrier leakage. In this work, advantage of the on-off nature of the PN pilot is taken and a direct conversion architecture is used in a non-conventional way. Instead of hard driving the oscillator port, the modulation port is hard driven. This reduces oscillator leakage at the expense of baseband PN leakage, which can be removed by filtering. In addition, the LO is provided at $2f_c$ and is divided down to f_c just prior to entering the modulator. Any unwanted coupling on the chip, IO wire-bonds or PCB will therefore be at $2f_c$ and also removable by filtering. Fig. 6.6 shows the circuit layout and the post layout simulation results are discussed in section 6.4.

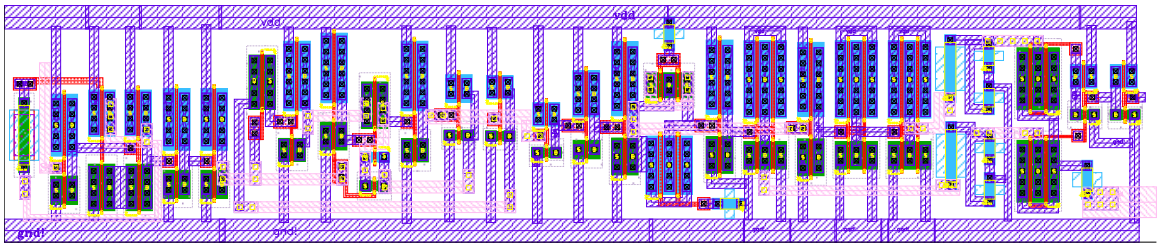


FIGURE 6.6: RF modulator layout.

Operating frequency and power consumption are key parameters for designing high speed digital circuits. To realise the divide-by-2 circuit a single edge triggered (SET) D-Flip Flop has been designed with a transmission gate which is depicted in Fig. 6.7.

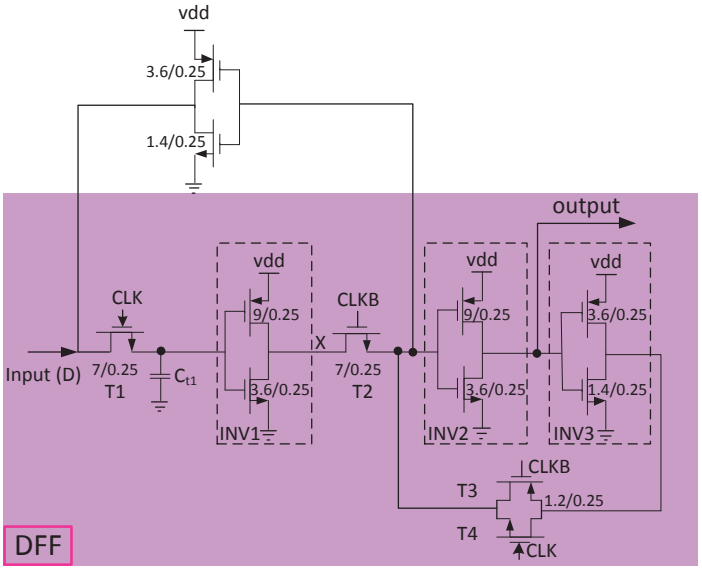
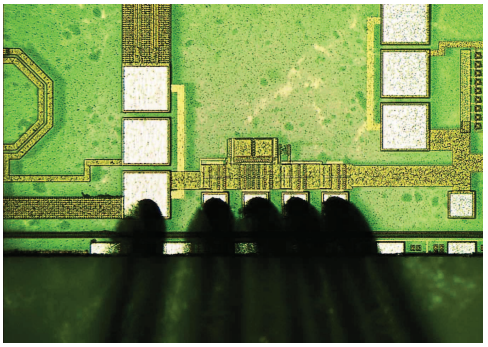


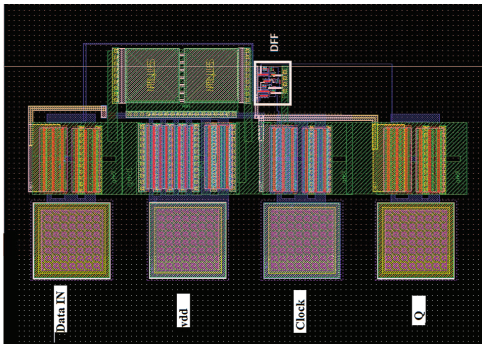
FIGURE 6.7: Frequency divider circuit level based on DFF.

The frequency divider uses a semi static master slave DFF [63]. In the state that the clock is high, the Master latch consisting of T1 and INVI is functional and the inverse of data is stored to an intermediate node X. In the state that the clock is low the slave latch consisting of T2 and a regenerative feedback becomes functional. The circuit is fast and suitable for RF applications (>2.5MHz) but cannot maintain the states at low

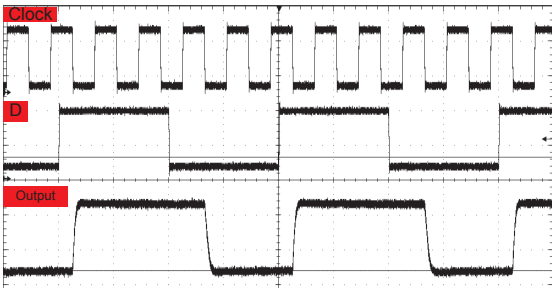
frequencies even though a weak regenerative feedback loop is used. Layout, photograph and measurement of the DFF are shown in Fig. 6.8.



(a) Chip micro photo.



(b) DFF layout.



(c) f_{Clock}=2.5MHz.

FIGURE 6.8: Chip die micrograph, layout, and measurement result of DFF.

6.3.2 Control Blocks

In addition to the primary circuit functions included on the prototype IC, several peripheral circuits are required to fully implement a testable chip. The circuits are: Serial to parallel IO, ESD protection and Schmitt trigger.

- Serial to Parallel IO:

  The circuit and layout shown in Fig. 6.9 is a 10 bit serial-in/ parallel-out shift register for reducing the number of IO pins. Each block includes two DFFs; ”DFF1” and ”DFF2”. The data transfers from ”Serial Input” to the DFF1’s by a

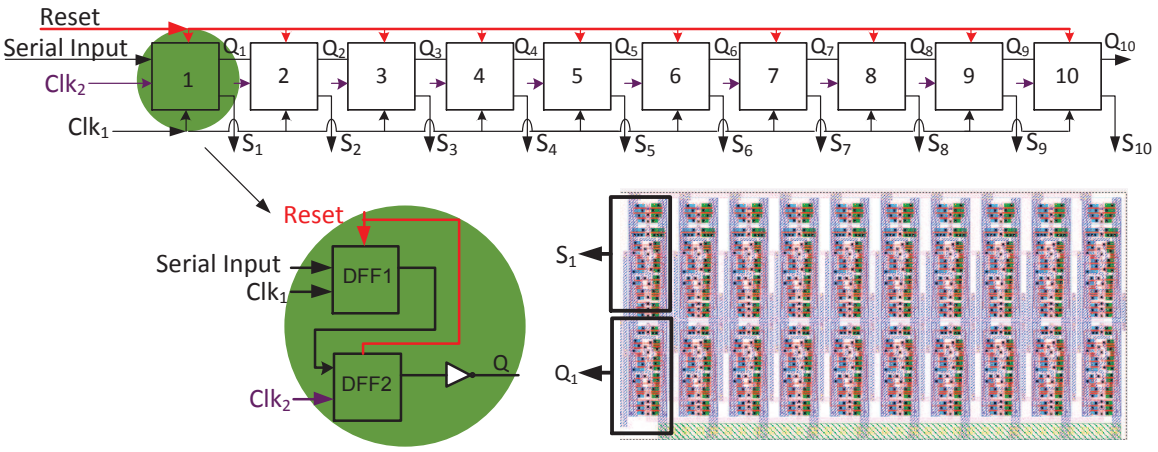


FIGURE 6.9: Clock distribution.

positive edge clock. The outputs of the DFF1s feed the DFF2s in the lower half in parallel. The lower register maintains the output while new data is being shifted into the upper register. The outputs drive the control lines selecting the PN length, and delay.

- Schmitt Trigger:

  The Schmitt circuit is a general inverter circuitry (double transistor inverter); to increase the hysteresis width by the additional two extra transistors [64]. Such circuits are used to reshape sloppy or distorted rectangular pulses and for noise rejection purposes in IC design. All clock inputs used a Schmitt trigger.

- ESD Circuit:

  A commonly observed phenomenon that causes IC failure is Electrostatic Discharge (ESD). It can happen during the manufacturing, shipping and testing. To avoid these failures, an on chip ESD protection has been implemented.

  The protection circuit chosen for this project are designed for use on digital lines with input voltages between 0 to 2V. Fig. 6.10.a illustrates a combination of a

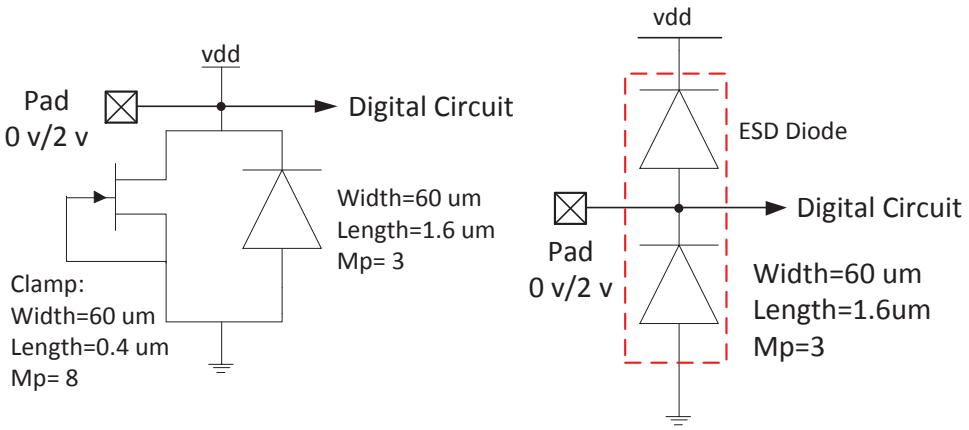


FIGURE 6.10: ESD circuit for control lines with (a) 2V supply and (b) Input/Output terminals

clamp and a diode to provide the ESD current paths for positive peaks (>2V) in the V_{DD} supply line and Fig. 6.10.b shows the ESD protection for the digital input and output pads and the RF pads. The layout has been designed with short current paths of very low ohmic resistance.

6.4 Post-Layout Simulation Results

The final layout of the complete pilot subsystem is illustrated in Fig. 6.11, where the circles are the solder pads.

The output of the serial to parallel controls the switches and enables us to select the desired length of the PN sequences as well as the required delay. The post layout simulation results are shown in Fig. 6.12–Fig. 6.14 to validate the prototype. Fig. 6.12 shows the ÷8 divider, that reduces the high speed clock to the chip rate. Fig. 6.13 shows the main and delayed outputs from the PN generator. Fig. 6.14 shows the operation of the up-convert modulator. The top trace is the LO input at $2f_c$, the middle trace shows

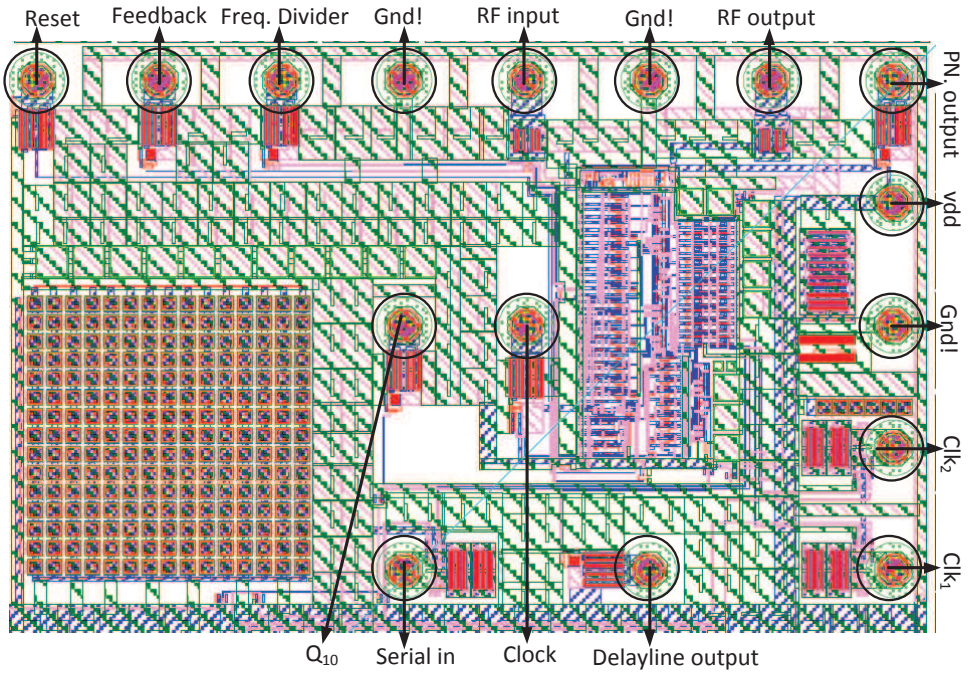


FIGURE 6.11: Final layout.

a change of state in the PN signal, the lower trace shows a 180^o phase reversal in the output modulated PN carrier at f_c.

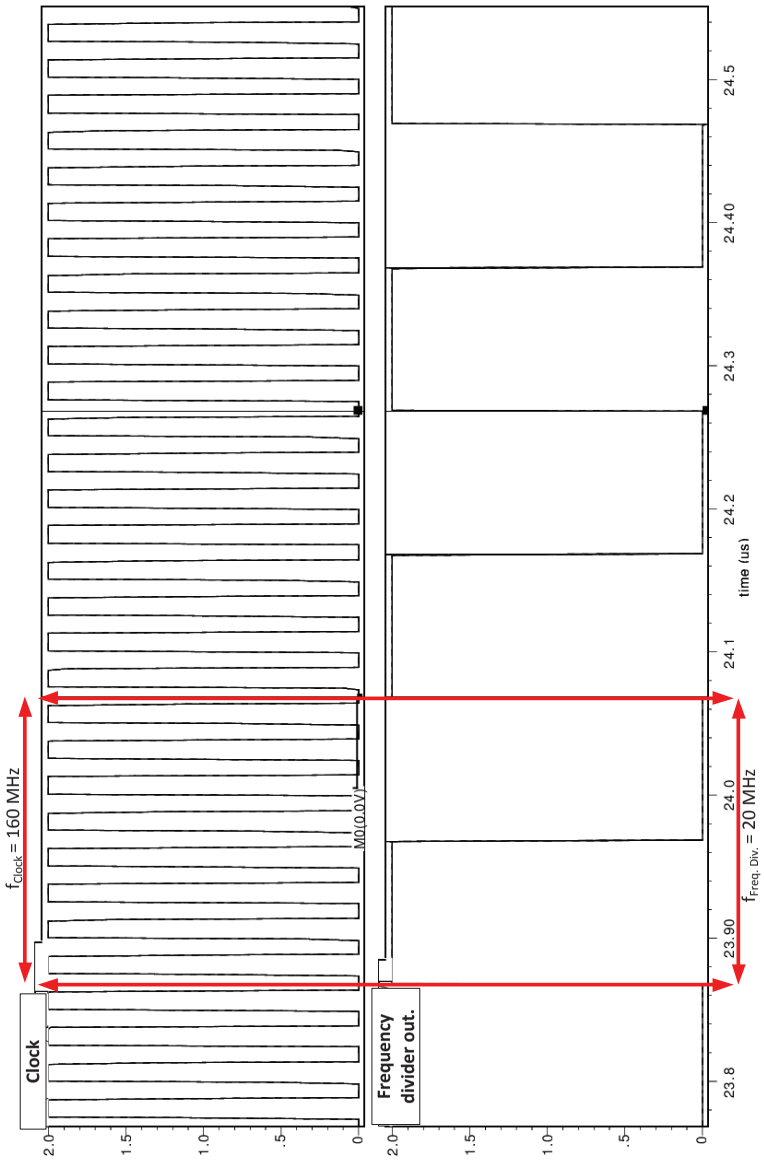


FIGURE 6.12: Output from divide-by-8 frequency divider f_{clock}=160MHz, $f_{divider}$=20MHz.

FIGURE 6.13: PN feedback, main output and delayed output. Delay=7 Chips

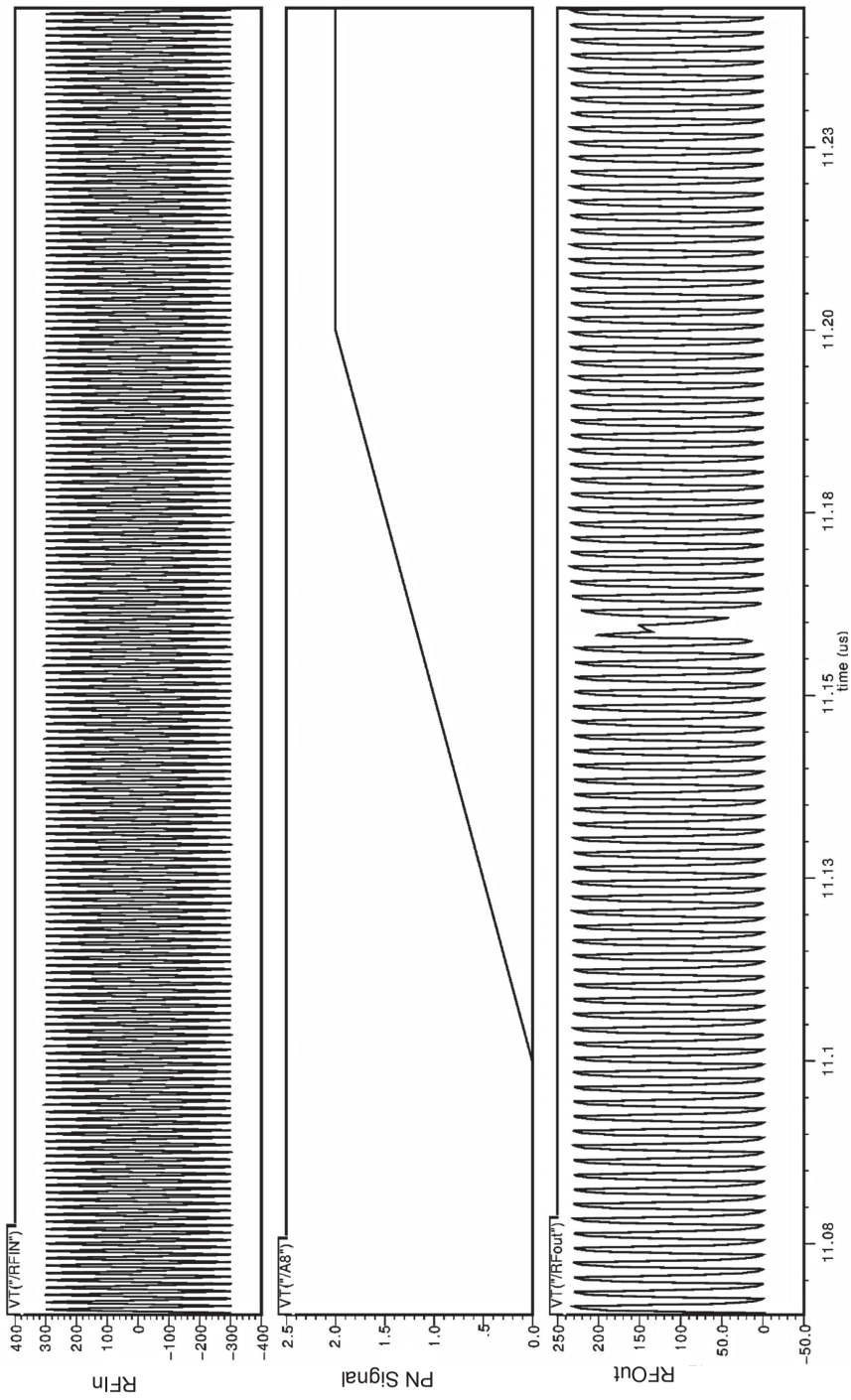
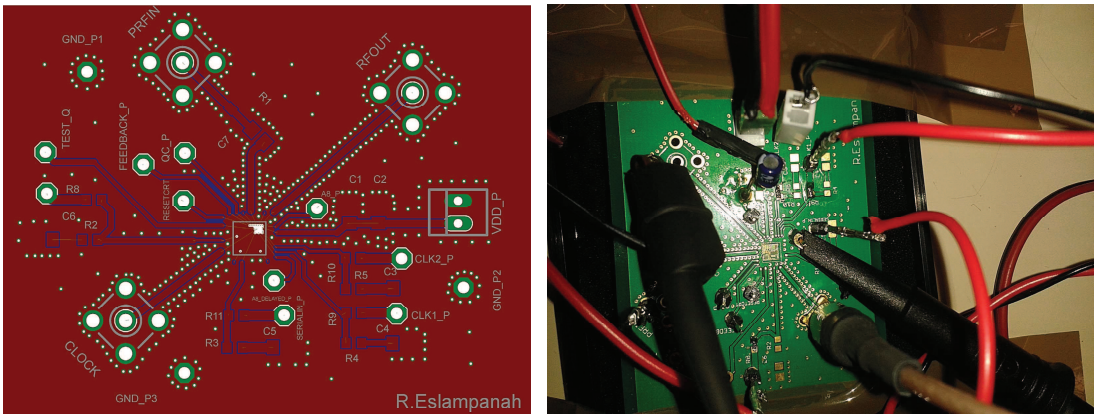


FIGURE 6.14: Simulation results of the RF modulator.

6.5 Testbed and Experimental Results

The pilot circuit was tested on a standard FR4 PCB with a 1.6 mm thickness and electroless nickel/immersion gold surfaces, where the wire bonds are to be attached (Fig. 6.15).



(a) (b) PCB layout.

FIGURE 6.15: PCB with surface mount components (SMD) and the chip. (a) naked board (b) Populated board under test

The Pilot subsystem prototype was fabricated in SOS 250-nm technology and characterized with a 2-V supply. Fig. 6.16 displays the die photograph of the pilot subsystem. The core area is $269\mu m \times 480\mu m$ and the device was originally designed for flip-chip mounting. This was found to be too costly and so the chip was wire bonded at the Melbourne Centre for Nano fabrication (MCN). Unfortunately the small active area for each pad ($35\mu m$ in diameter) caused most samples to fail. Ideally pad sizes should be greater than $60\mu m \times 60\mu m$ for wirebonding. Experimental work could not be done for the RF modulator, however the PN generator was successfully tested.

Fig. 6.17 shows the frequency divider measurement results. The plot shows the divide-by-8 frequency divider output with input clock rates of 20MHz and 450MHz respectively. The divider operates up to 950 MHz.

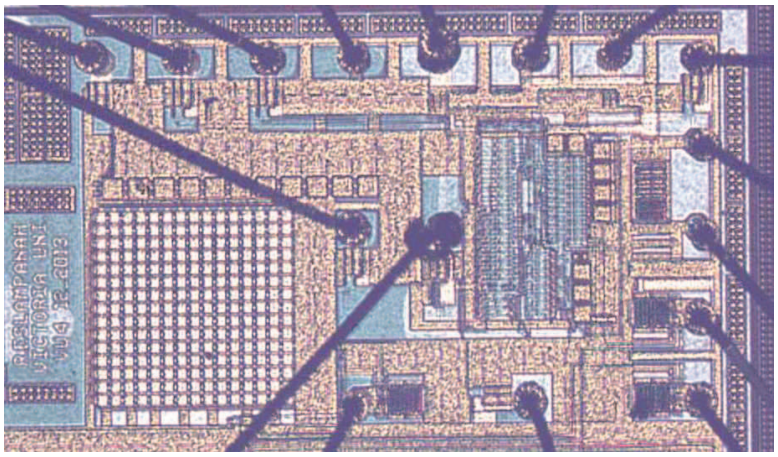
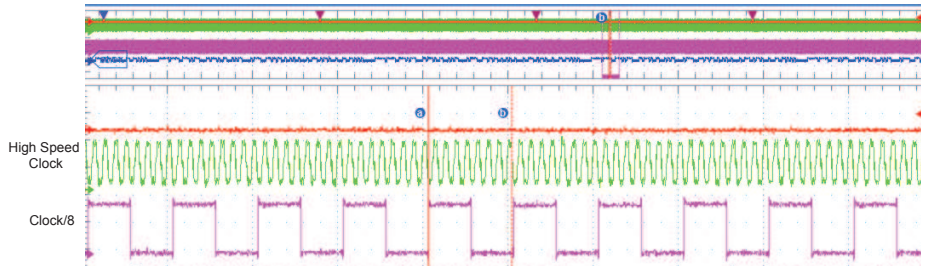
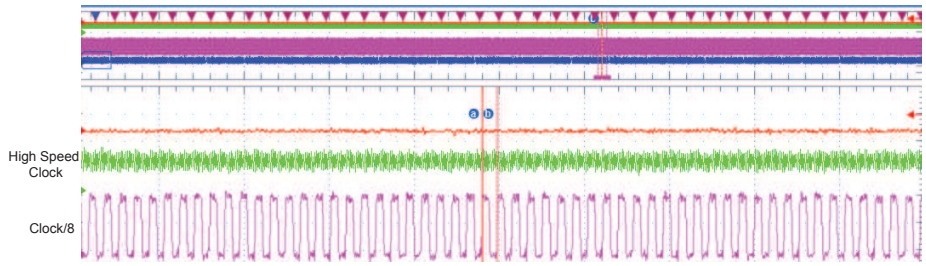


FIGURE 6.16: Die micrograph of the pilot subsystem.



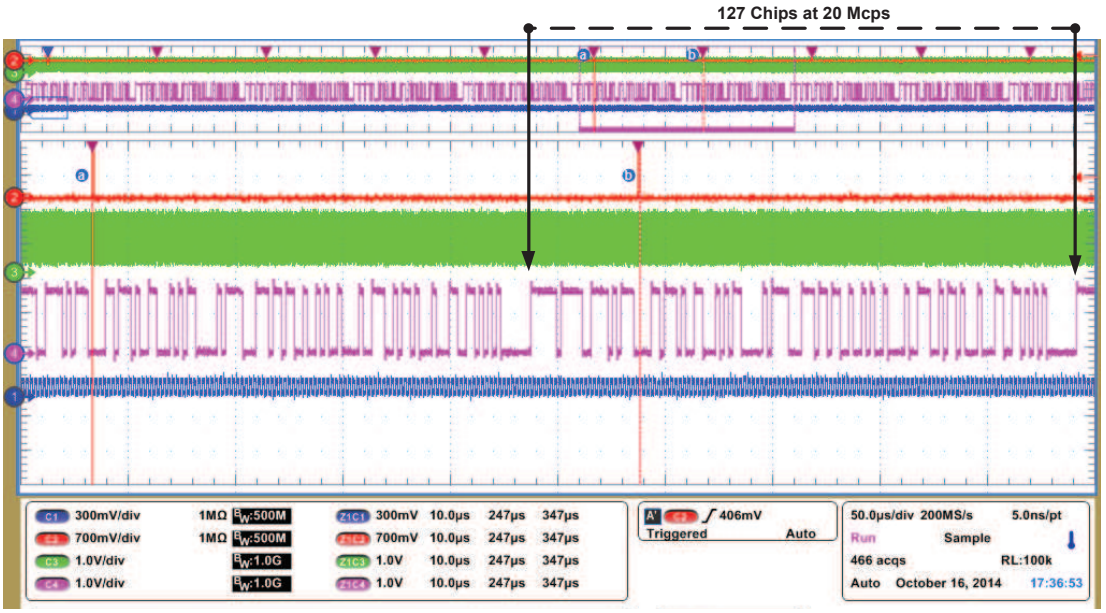
(a)



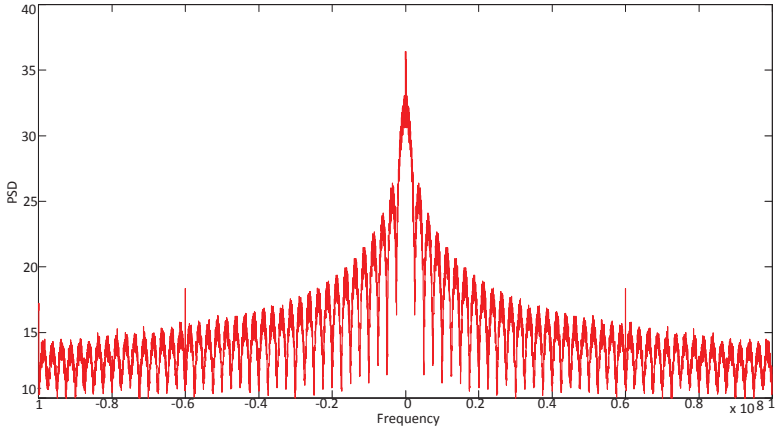
(b)

FIGURE 6.17: Divide by 8 with High Speed Clock at (a) 20 MHz (b) 450 MHz

Fig. 6.18 illustrates a sequence with length 127 and it's power spectral density (PSD). The PSD was calculated using an FFT routine of 100,000 samples downloaded from the digital oscilloscope. From the PSD, one can see that the main lobe is equal to twice the chip rate, 2.56 MHz with the nulls occurring at multiples of the chip rate. The DC component is due to the non-balanced logic levels of the output.



(a)



(b)

FIGURE 6.18: (a) Measurement results at 20 MHz Clock input. (b) The power spectrum.

6.6 Summary

The pilot signal is used to adjust feed forward coefficients to cancel the transmitter noise and leakage in an adaptive duplexer structure. In this chapter, the design, methodology, post layout simulation results and the measurement results of the pilot subsystem are explained. Although there was some problems connecting to the device, there is enough measurement evidence to show that the majority of functions in the device are working

as designed.

These results were published in [50], [65] and [66].

Chapter 7

Conclusion and Future Work

Todays wireless terminals are a mixture of software radio baseband circuits running on DSP cores and an analog circuit that does the up/down conversion to RF. The RF signal is switched through an appropriate off-chip duplexing network before reaching the antenna. Duplex filters, implemented in SAW or BAW technologies, separate out the Tx and Rx frequencies. They need to provide about 50dB isolation between the strong Tx signal and the sensitive receiver circuits. The filters are expensive external components in a radio communications solution and non tunable to support all bands. Further, multi-band flexibility is a need for future wireless handsets to have the capability of handling up to 35 bands. The adaptive duplexer is a potential solution to eliminate the need for multiple switched duplexers in multi-band systems. The scheme uses a low isolation circuit, such as a circulator, separated Tx and Rx antennas, tunable LC filters, or hybrid transformers etc., to give about 20dB isolation, and then uses two cancelling loops to obtain the remaining 30dB of isolation.

This Thesis extended the adaptive duplexing technique to cater for multiple aggregated frequency bands by increasing the number of cancelling loops. One loop is needed for each spectral null. The loops consist of a short delay and a gain-phase adjuster. A new algorithm was developed to control the gain-phase coefficients. The algorithm has k outputs controlling the gain-phase coefficients and k inputs that measure the residue leakage error in the k-1 Rx bands and the Tx band. The goal is to reduce the error signals to zero. The new algorithm uses an ”inverse plant control” strategy to orthogonalise the k-input k-output problem into k single-input single-output problems each controlled by an integral controller. The inverse plant estimation is performed by an NLMS algorithm using differential inputs and outputs.

The combination of the adaptive duplexer with inverse plant estimation and control is shown to reduce the transmit signal and transmitter noise leakage by up to 50 dB in both bands without the presence of noise and Rx signal in the receiver. When receiver noise is introduced (realistic case) the cancellation of the Tx noise leakage is reduced but is still well below the receiver noise floor. There is no effect on Tx signal leakage cancellation. Furthermore, the introduction of a (wanted) Rx signal jams the convergence process, further decreasing the systems ability to cancel the Tx noise leakage. The EVM reaches a floor as the wanted Rx signal increases. The floor can be reduced by increasing the length of the pilot PN sequence to improve the unwanted signal rejection of the cross-correlator. A pilot length of 1023 chips, will give an EVM floor of -25dB, suitable for 64QAM signals. Convergence takes 80 iterations (10ms) from cold, or 15 iterations (1.5ms) for changes in carrier frequency or antenna loading.

A variable length PN sequence and upconversion modulator was implemented in

silicon using an SOS process. The chip was wire bonded to a PCB and tested. Fig. 7.1 shows the test setup. Problems with the pad size caused many wirebonds to be shorted to ground and this inhibited a full test of the chip. The PN generator, and high speed clock divider were successfully measured, but the modulator failed.

7.1 Future Work

- Another adaptation scheme, such as recursive least squares (RLS) could reduce the number of iterations needed for convergence and tracking as it potentially offers faster convergence and smaller error with respect to the unknown system compared to LMS and NLMS [67] allowing the use of longer PN sequences in the cross-correlators. This could improve the EVM floor to a level suitable for 256QAM signals.

- A redesign of the pilot SOS chip with larger pads would improve the yield and enable the testing of more features.

- The cross correlators and additional down converter receivers are still needed for a complete pilot sub-system design. Implementation of the control algorithm in DSP or FPGA is also necessary.

- Other students have been involved in designing into silicon the Low Isolation Device and the gain-phase adjusters. Integration of these with the pilot sub-system would lead to a complete integrated adaptive duplexer. This is future work for someone else!

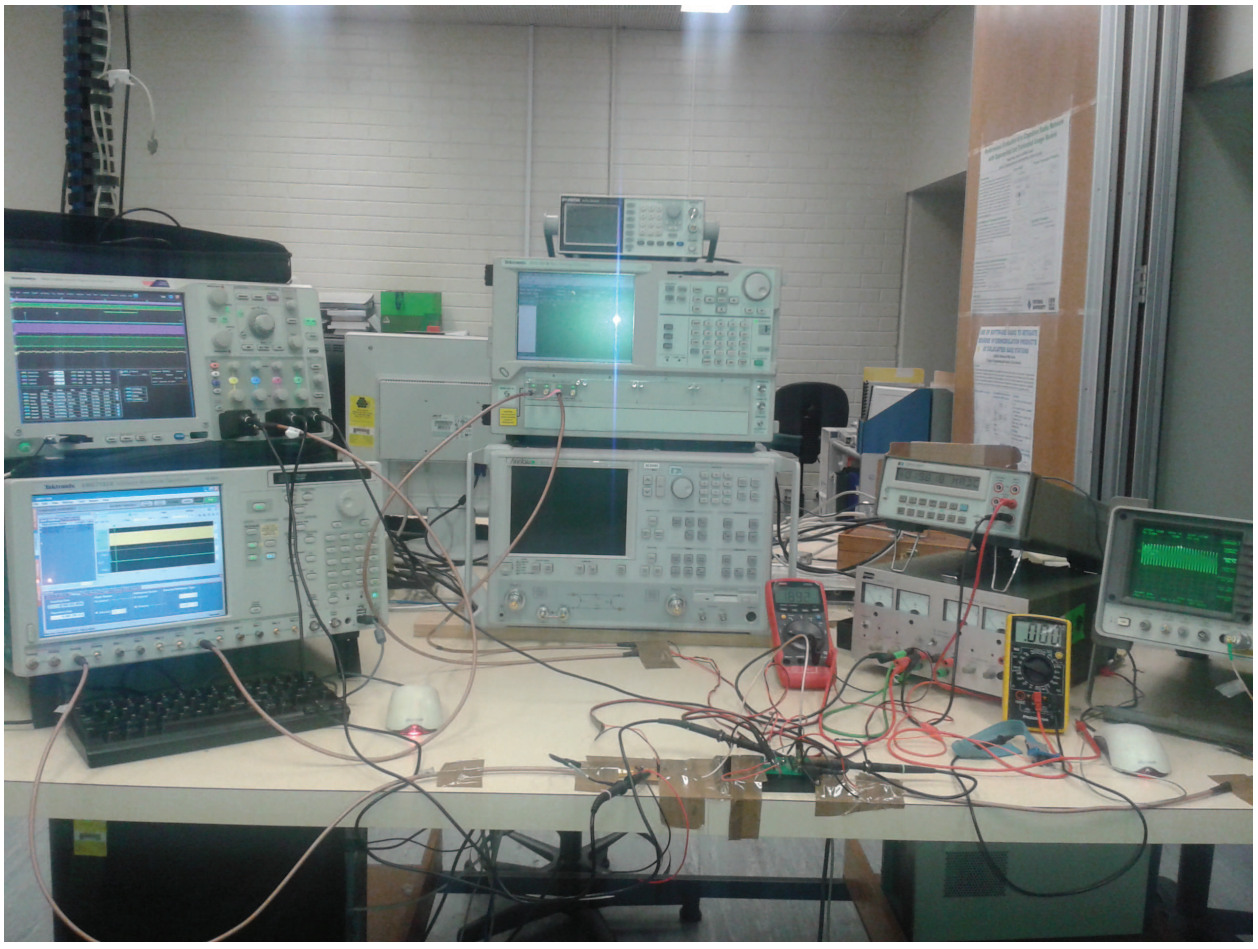


FIGURE 7.1: Photograph of the Experimental Setup.

References

[1] S. Stefania, T. Issam, and B. Matthew, "Lte, the umts long term evolution: from theory to practice," *A John Wiley and Sons, Ltd*, vol. 6, pp. 136–144, 2009.

[2] B. Razavi, *RF microelectronics*. Prentice Hall New Jersey, 1998, vol. 1.

[3] TDK Europe, "SAW Components, SAW Duplexer, LTE and WCDMA Band 1," Nov 2014. [Online]. Available: en.tdk.eu/inf/40/ds/mc/B8653.pdf

[4] O. Alrabadi, A. Tatomirescu, M. Knudsen, M. Pelosi, and G. Pedersen, "Breaking the transmitter receiver isolation barrier in mobile handsets with spatial duplexing," *Antennas and Propagation, IEEE Transactions on*, vol. 61, no. 4, pp. 2241–2251, April 2013.

[5] M. Mikhemar, H. Darabi, and A. Abidi, "A multiband rf antenna duplexer on cmos: Design and performance," *Solid-State Circuits, IEEE Journal of*, vol. 48, no. 9, pp. 2067–2077, Sept 2013.

[6] C. PRESTI, "Combined directional coupler and impedance matching circuit," Aug. 8 2013, wO Patent App. PCT/US2013/023,936. [Online]. Available: http://www.google.com/patents/WO2013116398A1?cl=en

[7] G. Kenworthy, "Self-cancelling full-duplex rf communication system," Nov. 25 1997, uS Patent 5,691,978. [Online]. Available: http://www.google.com/patents/US5691978

[8] S. Tiller, "Transmit signal cancellation in wireless receivers," Jun. 3 2004, uS Patent App. 10/655,748. [Online]. Available: http://www.google.com/patents/US20040106381

[9] S. Karode and V. Fusco, "Feedforward embedding circulator enhancement in transmit/receive applications," *Microwave and Guided Wave Letters, IEEE*, vol. 8, no. 1, pp. 33–34, Jan 1998.

[10] W. Schacherbauer, A. Springer, T. Ostertag, C. Ruppel, and R. Weigel, "A flexible multiband frontend for software radios using high if and active interference cancellation," in *Microwave Symposium Digest, 2001 IEEE MTT-S International*, vol. 2, May 2001, pp. 1085–1088 vol.2.

[11] T. O'Sullivan, R. York, B. Noren, and P. Asbeck, "Adaptive duplexer implemented using single-path and multipath feedforward techniques with bst phase shifters," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 53, no. 1, pp. 106–114, Jan 2005.

[12] V. Aparin, G. Ballantyne, C. Persico, and A. Cicalini, "An integrated lms adaptive filter of tx leakage for cdma receiver front ends," *Solid-State Circuits, IEEE Journal of*, vol. 41, no. 5, pp. 1171–1182, May 2006.

[13] S. Kannangara and M. Faulkner, "Adaptive duplexer for multiband transreceiver," in *Radio and Wireless Conference, 2003. RAWCON '03. Proceedings*, Aug 2003, pp. 381–384.

[14] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," *SIGCOMM Comput. Commun. Rev.*, vol. 43, no. 4, pp. 375–386, Aug. 2013. [Online]. Available: http://doi.acm.org/10.1145/2534169.2486033

[15] 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*, Nov 2010, pp. 1558–1562.

[16] Y.-S. Choi and H. Shirani-Mehr, "Simultaneous transmission and reception: Algorithm, design and system level performance," *Wireless Communications, IEEE Transactions on*, vol. 12, no. 12, pp. 5992–6010, December 2013.

[17] H. Su, G. Szczepkowski, and R. Farrell, "Wideband tx leakage cancellation using adaptive delay filter at rf frequencies," in *Irish Signals Systems Conference 2014 and 2014 China-Ireland International Conference on Information and Communications Technologies (ISSC 2014/CIICT 2014). 25th IET*, June 2014, pp. 396–401.

[18] E. K. Chong and S. H. Zak, *An introduction to optimization*. John Wiley & Sons, 2013, vol. 76.

[19] M. Williamson and M. Faulkner, "Effect of spread spectrum pilot on an svd based adaptive duplexer," in *Vehicular Technology Conference, 2006. VTC 2006-Spring. IEEE 63rd*, vol. 6, May 2006, pp. 2762–2766.

[20] S. Haykin and B. Widrow, *Least-mean-square adaptive filters*. John Wiley & Sons, 2003, vol. 31.

[21] E. Culurciello, *Silicon-on-sapphire circuits and systems*. McGraw-Hill, Inc., 2009.

[22] S. Dekleva, J. Shim, U. Varshney, and G. Knoerzer, "Evolution and emerging issues in mobile wireless networks," *Communications of the ACM*, vol. 50, no. 6, pp. 38–43, 2007.

[23] Peregrine Semiconductor Corp., "UltraCMOS® technology-the ultimate SOI," Apr 2014. [Online]. Available: http://www.psemi.compdf/sell_sheet-psg/DOC-35245-2web.pdf

[24] R. Ratasuk, D. Tolli, and A. Ghosh, "Carrier aggregation in lte-advanced," in *Vehicular Technology Conference (VTC 2010-Spring), 2010 IEEE 71st*, May 2010, pp. 1–5.

[25] J. Wannstrom, "Carrier aggregation explained," June 2013. [Online]. Available: http://www.3gpp.org/technologies/keywords-acronyms/101-carrier-aggregation-explained

[26] G. Ntouni, A.-A. Boulogeorgos, D. Karas, T. Tsiftsis, F. Foukalas, V. Kapinas, and G. Karagiannidis, "Inter-band carrier aggregation in heterogeneous networks: Design and assessment," in *Wireless Communications Systems (ISWCS), 2014 11th International Symposium on*, Aug 2014, pp. 842–847.

[27] M. Piette, "Ultra-wideband duplexers for low and high peak power applications," in *Ultra-Wideband Short-Pulse Electromagnetics 4, 1998*, 1998, pp. 187–193.

[28] S. Kannangara and M. Faulkner, "Analysis of an adaptive wideband duplexer with double-loop cancellation," *Vehicular Technology, IEEE Transactions on*, vol. 56, no. 4, pp. 1971–1982, July 2007.

[29] *3GPP SA TS 21.101, Technical Specifications and Technical Reports for a UTRAN-based 3GPP system*, `Version 10.x.y`. [Online]. Available: http://www.in2eps.com/3g21/tk-3gpp-21-101.html

[30] *3GPP TR 36.912, Feasibility study for Further Advancements for E-UTRA (LTE-Advanced)*. [Online]. Available: http://www.3gpp.org/dynareport/36912.htm

[31] G. Tsoulos, *MIMO System Technology for Wireless Communications*. CRC Press, 2006.

[32] P. Baudin, *Wireless Transceiver Architecture: Bridging RF and Digital Communications*. John Wiley and Sons, 2014.

[33] Q. Gu, *RF System Design of Transceivers for Wireless Communications*. Springer, 1936.

[34] M. Mikhemar, D. Murphy, A. Mirzaei, and H. Darabi, "A phase-noise and spur filtering technique using reciprocal-mixing cancellation," in *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2013 IEEE International*, Feb 2013, pp. 86–87.

[35] M. Faulkner and M. Johansson, "Adaptive linearization using predistortion-experimental results," *Vehicular Technology, IEEE Transactions on*, vol. 43, no. 2, pp. 323–332, May 1994.

[36] M. Faulkner, "Amplifier linearization using rf feedback and feedforward techniques," *Vehicular Technology, IEEE Transactions on*, vol. 47, no. 1, pp. 209–215, Feb 1998.

[37] V. Golikov, S. Hienonen, and P. Vainikainen, "Passive intermodulation distortion measurements in mobile communication antennas," in *Vehicular Technology Conference, 2001. VTC 2001 Fall. IEEE VTS 54th*, vol. 4, 2001, pp. 2623–2625 vol.4.

[38] F. Piazza, P. Orsatti, and Q. Huang, "A 0.25 mu;m cmos transceiver front-end for gsm," in *Custom Integrated Circuits Conference, 1998. Proceedings of the IEEE 1998*, May 1998, pp. 413–416.

[39] E. Atalla, A. Bellaouar, and P. Balsara, "Iip2 requirements in 4g lte handset receivers," in *Circuits and Systems (MWSCAS), 2013 IEEE 56th International Midwest Symposium on*, Aug 2013, pp. 1132–1135.

[40] A. Kiayani, L. Anttila, and M. Valkama, "Digital suppression of power amplifier spurious emissions at receiver band in fdd transceivers," *Signal Processing Letters, IEEE*, vol. 21, no. 1, pp. 69–73, Jan 2014.

[41] S. Abdelhalem, P. Gudem, and L. Larson, "Hybrid transformer-based tunable differential duplexer in a 90-nm cmos process," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 61, no. 3, pp. 1316–1326, March 2013.

[42] ——, "Tunable cmos integrated duplexer with antenna impedance tracking and high isolation in the transmit and receive bands," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 62, no. 9, pp. 2092–2104, Sept 2014.

[43] S. Karode and V. Fusco, "Dual polarised microstrip patch antenna using feed-forward isolation enhancement for simultaneous transmit/receive applications," in *Antennas and Propagation, 1999. IEE National Conference on.*, April 1999, pp. 49–52.

[44] G. Golub and W. Kahan, "Calculating the singular values and pseudo-inverse of a matrix," *Journal of the Society for Industrial and Applied Mathematics Series B Numerical Analysis*, vol. 2, no. 2, pp. 205–224, 1965. [Online]. Available: http://dx.doi.org/10.1137/0702016

[45] A. H. Sayed, *Fundamentals of adaptive filtering*. John Wiley & Sons, 2003.

[46] M. Pourakbar, R. Eslampanah, M. Faulkner, M. Tormanen, and H. Sjoland, "An sdr duplex filter in soi technology," in *Microwave Symposium (AMS), 2014 1st Australian*, June 2014, pp. 29–30.

[47] G. Plett, "Adaptive inverse control of linear and nonlinear systems using dynamic neural networks," *Neural Networks, IEEE Transactions on*, vol. 14, no. 2, pp. 360–376, Mar 2003.

[48] B. Widrow, "Adaptive inverse control," *IFAC Adaptive Systems in Control and Signal Processing*, 1986.

[49] B. Widrow and M. Bilello, "Adaptive inverse control," in *Intelligent Control, 1993., Proceedings of the 1993 IEEE International Symposium on*. IEEE, 1993, pp. 1–6.

[50] R. Eslampanah, L. Linton, S. Ahmed, R. Amirkhanzadeh, M. Pourakbar, J.-M. Redoute, and M. Faulkner, "Active duplexing for software defined radio," in *Circuits and Systems (ISCAS), 2014 IEEE International Symposium on*, June 2014, pp. 185–188.

[51] K. P. Yiu, "A simple method for the determination of feedback shift register connections for delayed maximal-length sequences," *Proceedings of the IEEE*, vol. 68, no. 4, pp. 537–538, April 1980.

[52] S. Tsao, "Generation of delayed replicas of maximal-length linear binary sequences," *Electrical Engineers, Proceedings of the Institution of*, vol. 111, no. 11, pp. 1803–1806, November 1964.

[53] S. Douglas, "A family of normalized lms algorithms," *Signal Processing Letters, IEEE*, vol. 1, no. 3, pp. 49–51, March 1994.

[54] R. Bitmead and B. Anderson, "Performance of adaptive estimation algorithms in dependent random environments," *Automatic Control, IEEE Transactions on*, vol. 25, no. 4, pp. 788–794, Aug 1980.

[55] G. Imthurn, "Silicon-on-sapphire, a replacement for gallium arsenide?" in *SOI Conference, 2009 IEEE International*, Oct 2009, pp. 1–3.

[56] M. Radmanesh, *Advanced RF & Microwave Circuit Design: The Ultimate Guide to Superior Design*. AuthorHouse, 2009.

[57] G. Imthurn, "The history of silicon-on-sapphire," *http://www.psemi.com/articles/History_SOS_73-0020-02.pdf*, Sep 2013.

[58] J. Roig, D. Flores, S. Hidalgo, J. Rebollo, and J. Millan, "Thin-film silicon-on-sapphire ldmos structures for rf power amplifier applications," *Microelectronics journal*, vol. 35, no. 3, pp. 291–297, 2004.

[59] J. Maurits, "Sos wafers ;some comparisons to silicon wafers," *Electron Devices, IEEE Transactions on*, vol. 25, no. 8, pp. 859–863, Aug 1978.

[60] R. C. Dixon, *Spread spectrum systems: with commercial applications*. John Wiley & Sons, Inc., 1994.

[61] R. Mutagi, "Pseudo noise sequences for engineers," *Electronics Communication Engineering Journal*, vol. 8, no. 2, pp. 79–87, Apr 1996.

[62] E. Hegazi and A. Abidi, "A 17-mw transmitter and frequency synthesizer for 900-mhz gsm fully integrated in 0.35- mu;m cmos," *Solid-State Circuits, IEEE Journal of*, vol. 38, no. 5, pp. 782–792, May 2003.

[63] M. Sharma, A. Noor, S. Tiwari, and K. Singh, "An area and power efficient design of single edge triggered d-flip flop," in *Advances in Recent Technologies in Communication and Computing, 2009. ARTCom '09. International Conference on*, Oct 2009, pp. 478–481.

[64] R. J. Baker, *CMOS: circuit design, layout, and simulation*. John Wiley & Sons, 2011, vol. 18.

[65] R. Eslampanah, M. Pourakbar, S. Ahmed, J.-M. Redoute, and M. Faulkner, "Rf modulator design for a low level pilot sub-system," in *Microwave Symposium (AMS), 2014 1st Australian*, June 2014, pp. 25–26.

[66] R. Eslampanah and M. Faulkner, "Pilot signal for adaptive duplexers," in *Proceedings of International Conference on Computer Communication and Management (ICCCM 2011)*, 2011.

[67] S. Ghauri and M. Sohail, "System identification using lms, nlms and rls," in *Research and Development (SCOReD), 2013 IEEE Student Conference on*, Dec 2013, pp. 65–69.