Integrated Transmitters for Cellular User Equipment–Wideband CMOS Power Amplifiers and Antenna Impedance Tuners

Lindstrand, Jonas

2019

Document Version:
Publisher's PDF, also known as Version of record

Integrated Transmitters for Cellular User Equipment–Wideband CMOS Power Amplifiers and Antenna Impedance Tuners

JONAS LINDSTRAND
DEPARTMENT OF ELECTRICAL AND INFORMATION TECHNOLOGY
FACULTY OF ENGINEERING | LTH | LUND UNIVERSITY

Integrated Transmitters for Cellular User Equipment–Wideband CMOS Power Amplifiers and Antenna Impedance Tuners

Jonas Lindstrand

LUND
UNIVERSITY

DOCTORAL DISSERTATION

by due permission of the Faculty of Engineering, Department of Electrical and Information Technology, Lund University, Sweden.

To be defended on Friday, November 8, 2019 at 9.15 in lecture hall E:1406, Department of Electrical and Information Technology, Ole Römers väg 3, 223 63 Lund Sweden

Faculty opponent

Professor Timo Rahkonen, University of Oulu, Oulu, Finland

Organization	Document name
LUND UNIVERSITY Department of Electrical and Information Technology P.O. Box 118 221 00 Lund Sweden	DOCTORAL DISSERTATION
	Date of issue: October 2019
Author: Jonas Lindstrand	Sponsoring organization: VINNOVA

Title and subtitle: Integrated Transmitters for Cellular User Equipment–Wideband CMOS Power Amplifiers and Antenna Impedance Tuners

Abstract:

The digital cellular systems era started about thirty years ago with the release of the first digital cellphones. These first digital cellphones were very different from today's slim and esthetic cellular pocket computers. They were not mass-produced in million units a day, and they were designed for radio performance rather than appearance. Today, all components are integrated inside the mobile phone to enable a product for the masses and not only the lucky few. For the radio performance this makes a large difference, especially the cellphones interaction with the user, which has a tendency to load the integrated antennas. This loading of the antennas means that the electronics inside the cellphone works sub-optimally, and a decrease in radio performance is inevitable, resulting in increased power consumption and reduced data rates. This problem can, however, be reduced by a concept called adaptive antenna impedance matching. This compensates for antenna loading effects, so that the electronics inside the cellphone can still operate with a close to optimum impedance, although the antenna impedance is changed due to user interaction. For adaptive impedance matching, the key component is the so called impedance tuner, which is studied, designed, and evaluated in this thesis. The requirements on this impedance tuner are very high, with low insertion loss, in-band distortion, out-of-band distortion, high tunability, and good power handling. The cost should also be as low as possible, which means that it should be implemented in a CMOS based technology suitable for mass-production. In this thesis, an impedance tuner is therefore designed and implemented in a CMOS-SOI technology. It has been verified to fulfill the requirements for use in a modern cellphone, with all measurements of key merits indicating high performance. Finally, it is worth to mention that this impedance tuner has also been used in a different project, where adaptive impedance matching was used in MIMO channel measurements with real cellphone users, but that project is outside the scope of this thesis.

The range of frequencies used for cellular communication has increased over the years, and today a large part of the so called sub-6 GHz frequency range is used. Most of the wireless services we have today use this decade wide frequency range (~600-6000 MHz), and although it is a wide frequency range, the spectrum is congested with a high density of communication. The circuits used to communicate in the sub-6 GHz bands must therefore have high RF-performance, and they should also be low cost since a large number of circuits is used to cover the complete frequency range. Difficulties reducing the cost per frequency band has drastically increased the cost of today's cellphones. This thesis therefore proposes an alternative concept for the power amplifier, the key component in the transmitter of the mobile phone, with the goal to reduce the cost of and the number of power amplifier circuits required to cover the complete sub-6 GHz range. The core of the concept was first designed and verified by measurements, an injection-locked power amplifier with supply modulation and dynamic transistor bias, resulting in high efficiency and bandwidth. To further reduce the cost of the cellphone more of the transmitter functionality, i.e. the frequency up-conversion, was added to the power amplifier circuit, which also improved the overall transmitter performance. Furthermore, a new version of polar modulation is proposed, to reduce the baseband signal bandwidth expansion, which polar modulation is notorious for. The reduction in bandwidth expansion decreases the overall power consumption of the transmitter, since the baseband circuits can then have lower bandwidth and clock-frequency. To further reduce the number of power amplifier circuits needed to cover the entire sub-6 GHz range, the bandwidth of the circuit was improved using a new higher order output matching network, together with a dual output power amplifier, resulting in a circuit that can operate with high performance over the complete sub-6 GHz frequency range. The proposed solutions in this thesis can reduce the number of ICs in cellular devices, which benefits not only the production cost, but also has positive effects on the environment.

Key words: CMOS, LTE, Impedance tuner, Injection lock, Power amplifier, Efficiency, Polar bandwidth expansion, and Multiband

Classification system and/or index terms (if any)

Supplementary bibliographical information	Language: English
ISSN and key title: Series of licentiate and doctoral dissertations 1654-790X; No. 127	*ISBN (digital)*: 978-91-7895-307-3 *ISBN (printed)*: 978-91-7895-306-6
Recipient's notes	Number of pages 138 / Price
	Security classification

Integrated Transmitters for Cellular User Equipment–Wideband CMOS Power Amplifiers and Antenna Impedance Tuners

Doctoral Thesis

Jonas Lindstrand

LUND
UNIVERSITY

Department of Electrical and Information Technology
Faculty of Engineering, Lund University
Lund, Sweden 2019

Department of Electrical and Information Technology
Lund University
P.O Box 118
ISBN 978-91-7895-306-6
ISSN 1654-790X; No. 127

Printed in Sweden by *Tryckeriet i E-huset,* Lund University
Lund 2019

…To the People that have always believed in me and my abilities…

Abstract

The digital cellular systems era started about thirty years ago with the release of the first digital cellphones. These first digital cellphones were very different from today's slim and esthetic cellular pocket computers. They were not mass-produced in million units a day, and they were designed for radio performance rather than appearance. Today, all components are integrated inside the mobile phone to enable a product for the masses and not only the lucky few. For the radio performance this makes a large difference, especially the cellphones interaction with the user, which has a tendency to load the integrated antennas. This loading of the antennas means that the electronics inside the cellphone works sub-optimally, and a decrease in radio performance is inevitable, resulting in increased power consumption and reduced data rates. This problem can, however, be reduced by a concept called adaptive antenna impedance matching. This compensates for antenna loading effects, so that the electronics inside the cellphone can still operate with a close to optimum impedance, although the antenna impedance is changed due to user interaction. For adaptive impedance matching, the key component is the so called impedance tuner, which is studied, designed, and evaluated in this thesis. The requirements on this impedance tuner are very high, with low insertion loss, in-band distortion, out-of-band distortion, high tunability, and good power handling. The cost should also be as low as possible, which means that it should be implemented in a CMOS based technology suitable for mass-production. In this thesis, an impedance tuner is therefore designed and implemented in a CMOS-SOI technology. It has been verified to fulfill the requirements for use in a modern cellphone, with all measurements of key merits indicating high performance. Finally, it is worth to mention that this impedance tuner has also been used in a different project, where adaptive impedance matching was used in MIMO channel measurements with real cellphone users, but that project is outside the scope of this thesis.

The range of frequencies used for cellular communication has increased over the years, and today a large part of the so called sub-6 GHz frequency range is used. Most of the wireless services we have today use this decade wide frequency range (~600-6000 MHz), and although it is a wide frequency range, the spectrum is congested with a high density of communication. The circuits used to communicate in the sub-6 GHz bands must therefore have high RF-performance, and they should also be low cost since a large number of circuits is used to cover the complete frequency range. Difficulties reducing the cost per frequency band has drastically increased the cost of today's cellphones. This thesis therefore proposes an alternative concept for the power amplifier, the key component in the transmitter of the mobile phone, with the goal to reduce the cost of and the number of power amplifier circuits required to cover the complete sub-6 GHz range. The core of the concept was first designed and verified by measurements, an injection-locked power

amplifier with supply modulation and dynamic transistor bias, resulting in high efficiency and bandwidth. To further reduce the cost of the cellphone more of the transmitter functionality, i.e. the frequency up-conversion, was added to the power amplifier circuit, which also improved the overall transmitter performance. Furthermore, a new version of polar modulation is proposed, to reduce the baseband signal bandwidth expansion, which polar modulation is notorious for. The reduction in bandwidth expansion decreases the overall power consumption of the transmitter, since the baseband circuits can then have lower bandwidth and clock-frequency. To further reduce the number of power amplifier circuits needed to cover the entire sub-6 GHz range, the bandwidth of the circuit was improved using a new higher order output matching network, together with a dual output power amplifier, resulting in a circuit that can operate with high performance over the complete sub-6 GHz frequency range. The proposed solutions in this thesis can reduce the number of ICs in cellular devices, which benefits not only the production cost, but also has positive effects on the environment.

Populärvetenskaplig sammanfattning

Under de tre senaste årtiondena har mobilindustrin utvecklats och utvecklat trådlösa system som blivit nödvändiga i vår vardag. Man kan säga att denna resa började med 2G, vilket faktiskt är den andra generationen för den mobila kommunikationen, men den första generationen som baserades på signalbehandling i den digitala domänen. När resan sedan fortsatte har digitalisering av mobila kommunikationssystem varit en stark trend, vilket innebär att det mesta av funktionaliteten nu är digital, och denna trend inom mobilindustrin är en del av trenden i dagens samhälle med genomgående digitalisering. Men själva de trådlösa signalerna är fortfarande analoga radiovågor som kommuniceras över etern med en fysisk antenn som bygger på samma principer som en vanlig antenn till radion eller TVn. Inuti mobiltelefonen finns dessa antenner kvar, men de syns inte från utsidan utan är inbyggda i telefonen, vilket har strömlinjeformat det estetiska utseendet. Denna integrering av antenner i dagens tunna mobiltelefoner har gett en ny utmaning, och det är hur antennerna fungerar vid beröring från användarens händer och huvud. Detta gör att elektroniken inte arbetar som det är tänkt, och telefonen kan då dra mer ström så att den måste laddas oftare. Men det finns en metod som kallas adaptiv matchning som kan minimera detta problem. Adaptiv matchning innebär att antennen är oförändrad ur elektronikens synvinkel, det vill säga för den elektronik som finns inuti mobiltelefonen. Ett huvudspår i denna avhandling har varit att undersöka så kallade impedanstuners, som är den viktigaste elektronikkomponenten för adaptiv matchning, och slutresultaten visar på en signifikant vinst i minskad effektförbrukning eller ökad datahastighet. De viktigaste egenskaperna för en impedanstuner, förutom att kompensera för antennens variationer, är att den ska kosta så lite pengar så möjligt vilket betyder att den måste byggas med halvledare i CMOS (som är billiga vid massproduktion), och att elektroniken i impedanstunern endast minimalt får förändra de analoga radiovågorna i form av dämpning och förvrängning av signalen. Denna avhandling uppvisar en impedanstuner i CMOS som uppfyller dessa krav.

Frekvensspektrumet för mobil kommunikation utökas ständigt, men delas i dagsläget mellan 2G, 3G, 4G, och snart kommer också 5G, och det kan därmed sägas att spektrumet är väldigt fullt. Detta betyder att kraven på elektroniken i en mobiltelefon är mycket höga, och att det dessutom behövs mer elektronik för att täcka många frekvensområden. Som en direkt konsekvens har totalkostnaden per mobiltelefon ökat. För att försöka minska kostnaden har därför en alternativ lösning för den dyraste komponenten föreslagits i denna avhandling. Komponenten är effektförstärkaren, som dessutom är den viktigaste byggstenen i en mobil sändare, och en av de största energiförbrukarna i en mobiltelefon och vars prestanda därför är centrala för att öka tiden mellan laddningarna. Effektförstärkaren måste med

precision förstärka radiosignalerna innan de omvandlas till radiovågor i antennen, vilket måste ske med så lite strömförbrukning som möjligt, och komponenten i sig måste också ha en så låg kostnad som möjligt. Den alternativa lösning för effektförstärkare som tagits fram i avhandlingen har därför utvecklats för låg strömförbrukning samt för att uppfylla precisionskraven för mobil kommunikation, och CMOS har använts för att den potentiella kostnaden ska vara minimal vid massproduktion. För att vidareutveckla konceptet har ett elektroniksystem för sändare konstruerats som innehåller mer funktionalitet än endast den hos effektförstärkaren. Detta system ger generellt lägre strömförbrukning och leder också till att färre komponenter behövs vilket minskar produktionskostnaderna. Till systemet har också en ny signalbehandlingsalgoritm tagit fram vars huvudsyfte är att ytterligare sänka strömförbrukningen. Ett annat sätt att minska kostnaden är att minimera antalet dyra komponenter i mobiltelefonen genom att minska antalet effektförstärkare. Detta har också gjorts i denna avhandling genom att konstruera och utvärdera en effektförstärkare som har likvärdig prestanda i alla frekvensband där det finns mobil kommunikation i dagsläget. Som slutsats kan det nämnas att de kretslösningar som tagits fram har en hög komplexitet men också hög potential ur en kostnadssynvinkel och ur ett miljöperspektiv, för när antalet elektronik-komponenter minimeras per konsumentprodukt kommer kostnaden och miljö-påverkan från själva produktionen att minska.

Table of Contents

Preface

My journey into the world of electrical engineering started already in my boy room growing up, and it involved music, electric guitars, and electric guitar gear. I did not play the instruments myself, the fascination for me was all the equipment including amplifiers, speakers, and effects based on distortion and delays. My keen interest in these things pointed me in the direction of Analog electronics and later in my undergraduate studies towards Radio electronics in CMOS technology. The grand idea of me getting a good education was always spoken of by my maternal grandmother and she always encouraged me to go down this path, which I apparently did. Her belief started a summer evening over thirty years ago, among some stunned family friends and a grand design built from LEGO.

As this journey ends and a new one begins, there is some time for reflection. This journey in science and engineering resulted in a handful of publications included as the main body of this thesis (Papers [**I-V]**), and also in additional publications where I have contributed to the research of others (Papers [VI-IX]). Furthermore, this project was funded by the Vinnova Industrial Excellence Center – System Design on Silicon (SoS) and the Excellence Center at Linköping – Lund in Information Technology (ELLIIT), which have a close relation to the Swedish industry. Through this relation two patent applications were filed, and as of today both applications are granted patents [i, ii]. I would like to end with a citation, well-known to me even long before I started this journey of circuits and systems:

"*Success is a journey, not a destination. The doing is often more important than the outcome.*"

-Arthur Ashe, American Athlete

Included publications

I. J. Lindstrand, I. Vasilev, and H. Sjöland, "A low band cellular terminal antenna impedance tuner in 130nm CMOS-SOI technology," *in Proc. European Solid-State Circuit Conference*, Venice, Italy, Sept. 2014, pp. 459-462.

II. J. Lindstrand, C. Bryant, M. Törmänen, and H. Sjöland, "A 1.6-2.6GHz 29dBm injection-locked power amplifier with 64% peak PAE in 65nm CMOS," *in Proc. European Solid-State Circuit Conference*, Helsinki, Finland, Sept. 2011, pp. 299–302.

III. J. Lindstrand, M. Törmänen, and H. Sjöland, "An injection-locked power up-converter in 65-nm CMOS for cellular applications," *in IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 3, pp. 1065–1077, March 2019.

IV. J. Lindstrand, M. Törmänen, and H. Sjöland, "Origin attraction – a technique to reduce signal bandwidth in polar transmitters," Submitted to *IEEE Transactions on Circuits and Systems I: Regular Papers*.

V. J. Lindstrand, M. Törmänen, and H. Sjöland, "A Decade Frequency Range CMOS Power Amplifier for Sub-6 GHz Cellular Terminals," Accepted for Publication in *IEEE Microwave and Wireless Components Letters.*

Related publications

VI. I. Vasilev, J. Lindstrand, V. Plicanic, H. Sjöland and B. K. Lau, "Experimental Investigation of Adaptive Impedance Matching for a MIMO Terminal With CMOS-SOI Tuners," *in IEEE Transactions on Microwave Theory and Techniques*, vol. 64, no. 5, pp. 1622-1633, May 2016.

VII. C. Bryant, J. Lindstrand, H. Sjöland and M. Törmänen, "A 70 and 210 GHz LO generator in 65nm CMOS," *in Proc. 2012 IEEE International Symposium on Radio-Frequency Integration Technology*, Singapore, 2012, pp. 195-197.

VIII. M. Törmänen, J. Lindstrand and H. Sjöland, "A 13dBm 60GHz-band injection locked PA with 36% PAE in 65nm CMOS," *in Proc. Asia-Pacific Microwave Conference 2011*, Melbourne, Australia , 2011, pp. 1-4.

IX. M. Törmänen, A. Axholt, J. Lindstrand and H. Sjöland, "A 2GHz Tx LO generation circuit with active PPF and 3/2 divider in 65nm CMOS," *in Proc. 2011 International Symposium on Integrated Circuits*, Singapore, 2011, pp. 208-211.

Approved Patents

i. J. Lindstrand, C. Bryant, and H. Sjöland, "High efficiency power amplifier," Okt. 2013. US Patent 8,554,162

ii. C. Bryant, J. Lindstrand, and H. Sjöland, "Transceiver front-end," Sept. 2017. US Patent 9,774,365

Acknowledgements

First of all I would like to extend my gratitude to my supervisor, Professor Henrik Sjöland. He always appears to know something about pretty much everything, a natural in technical solutions, and he has real prowess in technical and academic writing. It should also be mentioned that Professor Sjöland and his connections provided me a very rare and game changing master thesis opportunity and he also gave me the chance to pursue a PhD degree, and I am always grateful to him for these long term investments in me. I would also like to thank my assisting supervisor, Associate Professor Markus Törmänen. For teaching me some very important skills in the beginning of my PhD-studies, also for being a good friend, for all your support, and for always believing in me. My sincerest gratitude goes to the head of the department, Professor Daniel Sjöberg, without your guidance and support this thesis would never have been written. To the people I have cooperated with in the adaptive matching project, the pleasure was all mine and thank you for everything, Professor Buon Kiong (Vincent) Lau and my PhD student friend Doctor Ivaylo Vasilev. I would also like to acknowledge my two favorite educators, Associate Professor Pietro Andreani and Lecturer Göran Jönsson, especially for passing on your invaluable knowledge and for being a big inspiration. To my past and present fellow PhD students and group members, some often enclosed in your own world, myself often included, and to all of you that I interacted with over the years, you have my sincerest appreciation and some of you are really good and genuine people. I can add that Earth is a big place, but somehow the world is yet so small and hopefully I will interact with some of you in the future. My grateful thanks goes to Elisabeth Nordström, Andreas Johansson, Göran Jönsson, Erik Jonsson, Stefan Molund, and Sirvan Abdollah Poor of the technical and administrative staff for all the help over the years. I would also like to thank the retired members of the technical and administrative staff, Pia Bruhn and Lars Hedenstjerna, whom always have a special place in my heart.

I send all my love to my parents, Berit and Jan, for always being there, for all your support, and you have really been good and fair parents. Your Swedish way rarely makes you speak your mind, but you always show your appreciation with real actions and real commitments, and these are the things that really matters at the end of the day. I should also add without being humble, that I inherited my mother's determination and patience, my father's creativity and innovative ways to problem solving, and without these characteristics it would have been impossible for me to reach this level of education within the field of engineering. To my brother, Philip, I always try to send as much love as possible, he is the biggest catalyst of me starting this journey. It is his love and passion for music and electric guitars that sparked my love and passion for electric devices and their details.

I would also like to thank the American side of my family and although you live on the other side of the planet, you are always here with me in spirit. To Eileen, David, and Paul, your door is always open and your homes have been a refuge for me throughout the years, a place away from PhD student life and its implications. My sincerest appreciation to Inger Elman, a very close family friend and a big inspiration, energetic supporter, and role model.

Last but not least to my love, Elise, you have always rocked my world and when I am with you I always feel like a winner. You are my biggest inspiration to say the least, your open-mindedness, your intelligence, your efficiency, and your ability to learn and solve the impossible. Throughout the years, your love and support has made completion of this journey possible, and even through all these highs and lows. Somehow, as the turmoil of daily life has pushed us very close to the physical and mental breaking point at times, but our relationship and love for each other has always helped us prevail and only to make our love and relationship even stronger.

Jonas Lindstrand

Lund, 2019-09-19

List of Abbreviations

2G	2nd Generation cellular systems
2-Way	Two-Way combiner
3G	3rd Generation cellular systems
3GPP	3G Partnership Project
4G	4th Generation cellular systems
ACLR	Adjacent Channel Leakage Ratio
ADC	Analog to Digital Converter
AM-AM	Amplitude-Modulation to Amplitude-Modulation distortion
AM-PM	Amplitude-Modulation to Phase-Modulation distortion
Aux	Auxiliary
BALUN	BALanced to UNbalanced
BAW	Bulk Acoustic Wave
BB	Base-Band
BER	Bit Error Rate
BOM	Bill Of Material
BS	Base-Station
BW	BandWidth
BWE	BandWidth Expansion
CM	Common Mode
CMOS	Complementary Metal Oxide Semiconductor
CMRR	Common Mode Rejection Ratio
CP	Compression Point
C-Sens	Current-Sensor/Sensing
DAC	Digital to Analog Converter
DCO	Digitally Controlled Oscillator
DET	envelope DETector
DL	Down-Link

DPD	Digital Pre-Distortion
DUT	Device Under Test
EER	Envelope Elimination and Restoration
EM	Electro-Magnetic
ET	Envelope Tracking
E-UTRA	Evolved UMTS Terrestrial Radio Access
EV	Error Vector
EVM	Error Vector Magnitude
F	noise Factor
FDD	Frequency Division Duplex
FFP	Feed-Forward Path
G_P	Power Gain
GaAs	Gallium Arsenide
GMSK	Gaussian Minimum Shift Keying
GSM	Global System for Mobile communication
Gen	signal Generator
H-Band	High-Band
HD	Harmonic Distortion
H-EER	Hybrid Envelope Elimination and Restoration
I_{BB}	In-phase Base-Band
IC	Integrated Circuit
ICP	Input referred Compression Point
ICP_1	1 dB Input referred Compression Point
IIP_x	Input referred Intercept Point of the x^{th} order
IL	Insertion-Loss
ILPA	Injection-Locked PA
IM_2	2^{nd}-order Inter-Modulation product
IM_3	3^{rd}-order Inter-Modulation product
IMD	Inter-Modulation Distortion

IMD_2	2^{nd}-order Inter-Modulation Distortion
IMD_3	3^{rd}-order Inter-Modulation Distortion
IMD_5	5^{th}-order Inter-Modulation Distortion
IP	Intercept Point
IP_3	3^{rd}-order Intercept Point
ISI	Inter Symbol Interference
ISO	ISOlation
L-Band	Low-Band
LIM	LIMiter
LINC	LInear amplification using Nonlinear Components
L-ISO	Low-ISOlation
LNA	Low Noise Amplifier
LO	Local Oscillator
LPA	Linear class Power Amplifier
LTE	Long Term Evolution
LUT	Look-Up Table
MEMS	Micro-Electro-Mechanical Systems
MIMO	Multiple Input Multiple Output
ML	Matching-Loss
MMPA	Mixed-Mode/class Power Amplifier
MN	Matching Network
MODEM	MOdulator-DEModulator
MOM	Metal-Oxide-Metal
NF	Noise Figure
OCP	Output referred Compression Point
OFDM	Orthogonal Frequency-Division Multiplexing
OIP_x	Output referred Intercept Point of the x^{th} order
PA	Power Amplifier
PAE	Power Added Efficiency

PAPR	Peak-to-Average Power Ratio
PC	Personal Computer
PCB	Printed Circuit Board
PLL	Phased Locked Loop
PPA	Pre-Power Amplifier
PSK	Phase Shift Keying
PWM	Pulse Width Modulation
Q	Quality factor
Q_{BB}	Quadrature-phase Base-Band
QAM	Quadrature Amplitude Modulation
QVCO	Quadrature Voltage Controlled Oscillator
R_X	Receiver
RF	Radio Frequency
RL	Resistive-Loss
RRC	Root Raised Cosine
SAW	Surface Acoustic Wave
SC-FDMA	Single-Carrier Frequency-Division Multiple Access
SCS	Signal Component Separator
SiGe	Silicon Germanium
SM	Supply Modulator
SMPA	Switched-Mode Power Amplifier
SNR	Signal-to-Noise Ratio
SOC	System On Chip
SOI	Silicon On Isolator
SOS	Silicon On Sapphire
SRF	Self-Resonance Frequency
STM	ST-Microelectronics
T	Transformer
T_X	Transmitter

T_X-BB	Transmitter Base-Band
TDD	Time Division Duplex
T/R	Transmit/Receive
T_X-Synth	Transmitter Synthesizer
UE	User Equipment
UL	Up-Link
UMTS	Universal Mobile Telecommunications System
UTRA	UMTS Terrestrial Radio Access
VCO	Voltage Controlled Oscillator
VNA	Vector Network Analyzer
VSWR	Voltage Standing Wave Ratio
W-CDMA	Wideband Code Division Multiple Access
Wi-Fi	Wireless Fidelity
WLAN	Wireless Local Area Network
XMD	CROSS-Modulation Distortion
ZVS	Zero-Voltage Switching

List of Symbols

Γ	Reflection coefficient
η_{PA}	Power amplifier drain efficiency
η_{SM}	Supply modulator drain efficiency
η_{tot}	Total drain efficiency
$\phi(t)$	Time-variant phase signal
ω	Angular frequency
$A(t)$	Time-variant amplitude signal
$A_{Duplex}^{f=Rx}$	Duplexer attenuation in the receive band
C_{ds}	Drain-source capacitance
C_{gd}	Gate-drain capacitance
C_{gs}	Gate-source capacitance
$Env(t)$	Time-variant envelope signal
$EV(I,Q)$	Error Vector in the Cartesian domain
f_{max}	Maximum oscillation frequency
f_T	Transit frequency
F_{Tot}	Total noise factor
g_d	Small-signal drain conductance
G_m	Large-signal transconductance
g_m	Small-signal transconductance
P_{DC}	DC power consumption
$\overline{P_{EV}}$	Average power of the Error Vector
$P_{Noise-Tx}^{f=Rx}$	Transmitter output noise in the receive band
$\overline{P_{Ref}}$	Average power of the Reference Vector
$P_{Tx}^{f=Tx}$	Transmitter output power

r_{ds}	Small-signal drain-source resistance
R_L	Load resistance
r_n	Small-signal negative resistance
R_{off}	Off-resistance
R_{on}	Triode on-resistance
t_s	Symbol time

Part I

General Introduction

CHAPTER 1

1 Introduction

During the last three decades digital wireless communication has evolved rapidly, with a user maximum data rate going from 270 kilobits per second (kbps) in the 2nd Generation cellular systems (2G) or Global System for Mobile communication (GSM) [1, 2] to today's approximately 100-400 megabits per second (Mbps) in the 4th Generation cellular systems (4G) or Long Term Evolution (LTE) [3]. The data rates have thus increased about 1000 times. The main driving force in this dramatic development has been Moore's Law [4], i.e. the exponential down-scaling of features-sizes in the most significant semiconductor technology for electronic consumer products, which is the Complementary Metal Oxide Semiconductor (CMOS) technology. This exponential down-scaling in features-sizes means that the digital functionality per unit area is doubled every two years, and circuit speed is also increased, best observed as the increase in Personal Computer (PC) processor clock frequency during the 1990's [5]. In the beginning of digital wireless communication, the CMOS-scaling only effected the digital blocks since all the analog blocks were implemented in specialized technologies like as GaAs[6] and SiGe[7], which had superior analog performance during those days, and still have better performance in some key aspects like gain, noise, linearity and voltage breakdown. However, during the 3rd Generation (3G) or Universal Mobile Telecommunications System (UMTS)/Wideband Code Division Multiple Access (W-CDMA) era in the late 1990's, due to scaling of CMOS it was now possible to use the technology to also implement the analog- and Radio Frequency (RF)-blocks, and the main reason was that the CMOS processes now had enough speed (f_t and f_{max}). By implementing all building blocks in CMOS, complete System On Chip (SOC) solutions were possible and the main feature was that digital, mixed-signal, analog, and RF circuits all could be implemented on the same die. This opened-up the door for wireless products to the main advantage of CMOS technology, the ability to mass-produce chips and also to reduce the number of chips in each User Equipment (UE)/Cellphone. The first fully integrated RF SOC in CMOS was developed by Ericsson in the late 1990's, it is known in the world as Bluetooth [8], and as of today Bluetooth is used in the widest range and largest quantities of consumer products ranging from videogames to bed lamps. This is only possible by

implementing the chip in a CMOS process, where the high level of integration and large volume of fabrication makes each sold unit cheap. If other more exotic technologies were to be used the cost would be orders of magnitude higher for the required Integrated Circuit (IC) production. Finally, the main misconception of CMOS though, that it only cost "nickels" and "dimes", is not true. Only with mass-production can each chip unit cost "nickels" and "dimes", and then only if millions and millions of chips are sold.

As the development of analog/RF-CMOS circuits continued, the main approach to achieve sufficient performance was circuit design innovation exploiting opportunities and avoiding weaknesses of the technology, and it is still this way today. This is all in the nature of the technology and the market, CMOS is cheap to mass-produce, and yes, speed has increased with Moore's Law, but gain, noise, linearity and voltage breakdown are all getting worse with each CMOS scaling step. The designer has to work with what is available, in Receivers (R_X) the Low Noise Amplifier (LNA) was a big issue in CMOS, but to achieve sufficient gain and low noise, a noise canceling approach was first disclosed in [9]. Further, the designer can use the main advantage of CMOS down-scaling, which is the high performance of transistors operated as switches. High performance, low noise and high linearity, mixers for frequency conversion in receivers and Transmitters (T_X) can then be realized. The mixers can be of different types, active [10] or passive [11], current mode [12] or voltage mode [13], and can also include harmonic rejection [14] to counteract the harmonics of the switching waveform. Switches can also be used to create programmable and reconfigurable circuits, switching in and out components or entire building blocks. For instance a Digitally Controlled Oscillator (DCO) [15] can be created by switching in and out capacitors of the resonator according to a digital control number, altering the oscillation frequency. For transmitters Power Amplifiers (PA) with high efficiency can be realized using transistors operating as switches, called Switched-Mode Power Amplifiers (SMPA) [16]. Still, it is very hard to design a power amplifier in CMOS technology. In scaled down technology, where transistors have high speed and excellent switch performance, the voltage breakdown is very low, making it hard to reach high output power. An often used circuit solution is to use stacked transistors, in a so-called stacked-cascode configuration [17]. Two or more transistors stages are then stacked on top of each other, so that the maximum supply voltage of the PA can be the sum of the maximum supply voltage of all stacked stages, i.e. the PA supply voltage is divided among several transistors. Then a higher supply voltage compatible with higher output power can be used, even though the individual transistors cannot handle such voltage levels. By using a stacked-cascode configuration, the performance of each individual transistor is not improved, but the result yields an increased performance for each cascode element added.

From a system point of view, one of the biggest breakthroughs has been the elimination of the Surface Acoustic Wave (SAW) and/or Bulk Acoustic Wave (BAW) filter in both R_X and T_X. This elimination became possible by the use of circuit understanding and design, the solution was a passive mixer [11], which has a feature called frequency-translation [18]. In the frequency-translation phenomenon, if an impedance is connected to the mixer low frequency terminal, at the RF mixer terminal the same impedance will appear, but up-converted to RF frequency. The frequency difference, i.e. how much the impedance is translated in frequency, is set by the Local Oscillator (LO) frequency. If the impedance has a low-pass characteristic, like a parallel RC circuit, and if the LO frequency is much higher than the BandWidth (BW) of the low frequency filter, the result is a very narrow RF bandpass filter, which would not be possible to realize on-chip without this technique. Further, by using this solution, tunable selectivity can also be introduced and by changing LO frequency the passband frequency of the filter can be changed as well.

These are examples of only a few circuit solutions/concepts developed over the two and a half decades when CMOS been a serious contender for RFIC implementation.

In the meantime, the wireless system technology has evolved as well, at first GSM only used Gaussian Minimum Shift Keying (GMSK), a type of frequency modulation, but as throughput requirements increased higher order modulation schemes like 8 Phase Shift Keying (PSK), 16 Quadrature Amplitude Modulation (QAM), 64QAM, 256QAM, were required to increase spectral efficiency, i.e. data rate for a certain bandwidth. To further increase the data-rate the bandwidth has also increased, and to handle the frequency variations of wireless channels Orthogonal Frequency-Division Multiplexing (OFDM) was introduced [19]. Effectively the data is then carried by many closely spaced narrow-banded individually modulated so-called subcarriers. The higher order modulation schemes and OFDM increased spectral efficiency and data rates [20], but the requirements on the hardware in the Modulator-DEModulator (MODEM) also became more stringent. Further, to meet the ever increasing requirements for higher data rate, a multiple antenna technique called Multiple Input Multiple Output (MIMO) was introduced [21]. From basic MIMO theory, the theoretical throughput increases linearly with the number of antenna elements, but this is only true if the antenna signals are uncorrelated [22]. Circuit hardware, on the other hand, must be added to each individual antenna element, increasing cost, power consumption, and complexity. This is a big problem for the handheld UE, powered by batteries. Finally, the main issue with a handheld UE utilizing MIMO is the human touch and the fact that people touch the UE when they use it. To clarify, today's wireless UE has most of its surface area covered by antennas, due to MIMO and the wide range of carrier-frequencies/bands used. If the user touches the UE, it will load the different antennas, and this results in an increase in correlation between the different antennas and an increase in RFIC power consumption, due to the mismatched antenna impedance [23].

In Fig. 1.1 the evolution of the cellular wireless spectrum over the past three decades is illustrated, 30 years ago GSM mainly used four bands (850 MHz, 900 MHz, 1800 MHz, and 1900 MHz), and in rare cases also 400 MHz and 450 MHz [24], and today with LTE sixty-six (66) bands are used ranging from 450 MHz to 5850 MHz [25]. Further, the LTE standard also has to co-exist with 2G, 3G, and Wireless Local Area Network (WLAN)/Wireless Fidelity (Wi-Fi) which means that the utilized spectrum (450 MHz to 5850 MHz) is at max capacity today. Looking at the cost of cellular devices, the Bill Of Material (BOM) of the MODEM has increased, mainly due to the amount of hardware/components required to cover all the utilized frequency bands, also supporting MIMO. The second problem, since the frequency spectrum is full all unwanted spectral emissions, both in-band and out-of-band, must be kept at a minimum. The third major issue, all the hardware has to be connected on a Printed Circuit Board (PCB) to each antenna, with high quality supply and ground connections, such complex PCB having substantial cost. Finally, all these complex devices have to co-exist in a crowded frequency spectrum, where the margin of error can be considered as small as the feature-size of the individual devises on the CMOS die.

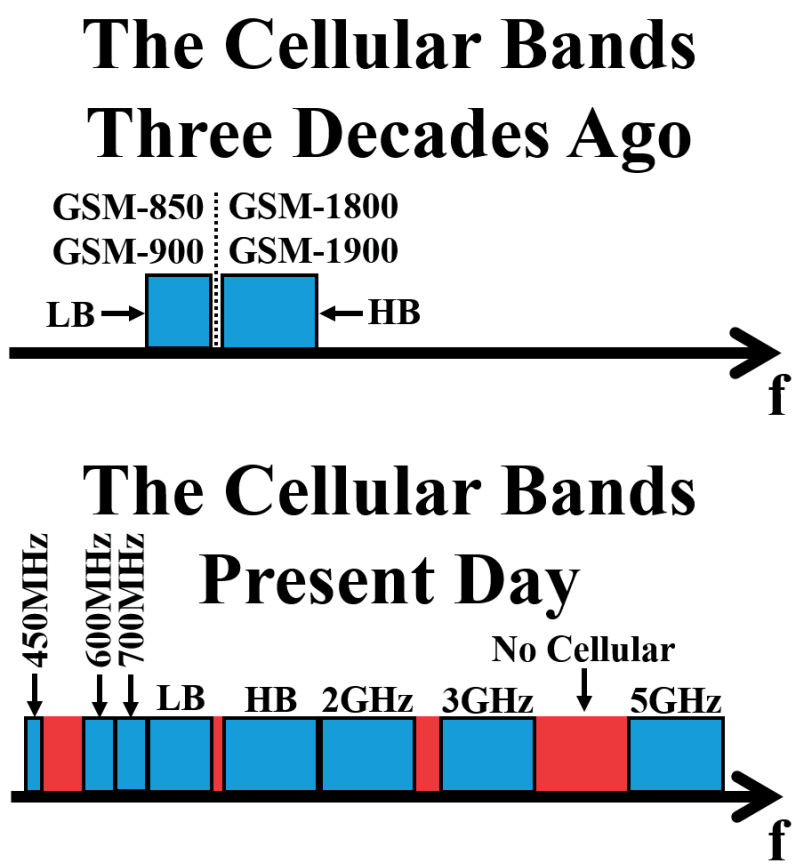

Fig. 1.1. The spectral-evolution of wireless cellular communication.

CHAPTER 2

2 Cellular UE Transcievers

The core of every MODEM is the transceiver, which consist of one or many R_X and T_X pairs. In today's transceivers all frequency band are covered by one IC, which can be realized by using separate R_X and T_X chains for each band [26], or by using a wideband R_X and T_X solution [27]. Further, the major difference between today's transceiver and the transceivers of the past is the frequency conversion, and in present day it is dominated by homodyne/direct conversion architectures [28], whereas in the past heterodyne/IF architectures [29] dominated the transceiver implementation. The main differences are illustrated in Fig 2.1, the homodyne transceiver operates with an LO frequency equal to the RF frequency whereas the heterodyne structure has an LO frequency lower or higher than the RF frequency.

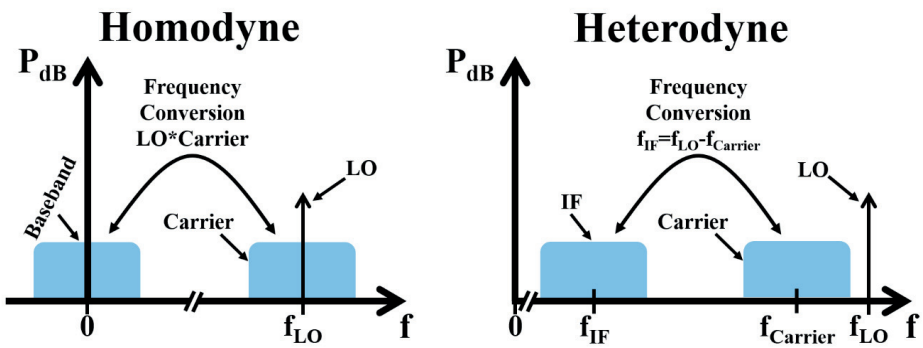

Fig. 2.1. Comparison of homodyne and heterodyne frequency conversion.

The main benefit of the homodyne is that it requires the lowest amounts of off-chip filtering, whereas the heterodyne has a natural isolation between RF and LO (not operating at the same frequency). The main drawback of the homodyne is that it has issues with CROSS-Modulation Distortion (XMD) due to interaction of LO signal leakage and signal-blockers [30] and it is also very susceptible to oscillator pulling [31], whereas the heterodyne requires higher amounts of sharp filters and also more frequency conversion steps. Overall, the homodyne transceiver is more straight-forward and it also requires the least amount of off-chip hardware.

2.1 Multi-band Transceivers

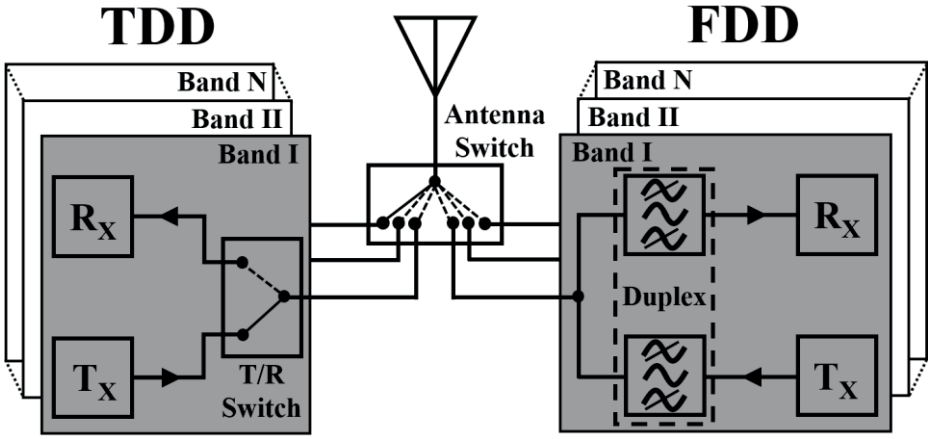

Fig. 2.2. The TDD and FDD multi-band cellular transceiver systems.

Fig. 2.2 shows a simplified block diagram of a multi-band cellular transceiver, the R_X (Down-Link [DL]) and T_X (Up-Link [UL]) can operate at different frequency bands according to both Time Division Duplex (TDD) (for GSM/LTE) and Frequency Division Duplex (FDD) (for W-CDMA/LTE) schemes. In TDD the R_X and T_X do not operate simultaneously, and a Transmit/Receive (T/R) switch can thus be used to connect them alternatingly to the antenna. In FDD, on the other hand, the T_X and R_X operate simultaneously, but at different frequencies. In order to isolate the signals from the T_X to the antenna and from the antenna to R_X, so that strong T_X signals do not saturate the R_X, an off-chip duplex filter (duplexer) is generally used in FDD. Such off-chip (SAW or BAW) duplexers are fixed to a specific frequency band. Therefore, switches are used to change operating bands and each duplexer provides the required isolation for each band. The duplexer has three system functions, to isolate the sensitive R_X from the strong T_X-carrier, to filter-off the noise in the R_X-band generated by the T_X, and to filter-off noise and spurious content to and from the antenna. At the R_X input the high powered (+27dBm [0.5W], Power Class 2 in LTE) T_X-carrier signal must be attenuated to not compress the R_X, the noise from the T_X is caused mainly by the phase-noise in the local oscillator signal (LO_{TX}) and thermal noise in the T_X mixers and some of it ends up in the R_X-band, and if not suppressed this noise will reduce the ability to receive weak signals, i.e. it will desensitize the R_X.

In Fig. 2.3 the frequency characteristic of a SAW duplexer is shown, ISOlation (ISO) of both T_X-carrier and T_X-noise, and Insertion-Loss (IL) of both the R_X and T_X path to the antenna are shown. As can be seen, the T_X-carrier isolation (in the T_X-band) is high (about 50 dB) and the T_X-noise in the R_X-band is attenuated by about 45 dB, so if the MODEM did not include a duplexer, the transceiver system requirements would be impossible to meet with a reasonable power consumption. Active duplexers and balancing duplexers for implementation on-chip have been sought to replace off-chip SAW/BAW devices, and examples are shown in [33] and [34-37], respectively.

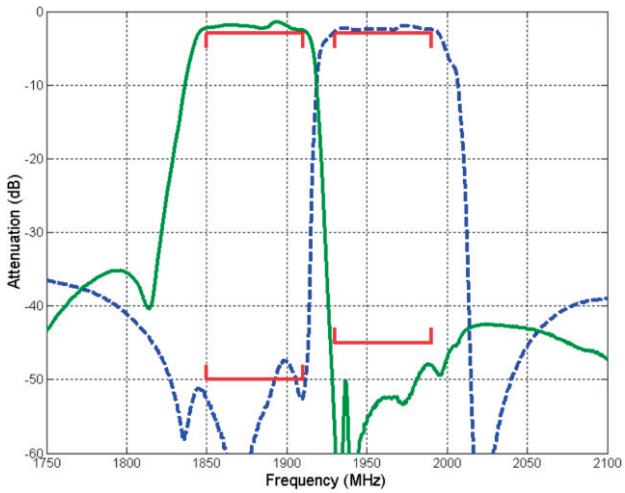

Fig. 2.3. Frequency-characteristics of a SAW-duplexer [32].

2.1.1 Circuit's for the Duplex Function

The block diagram of an active duplexer is shown in Fig. 2.4. The core of the active duplexer is the Low- ISOlation (L-ISO) device, which connects the antenna to the R_X input and T_X output. For the implementation of the L-ISO device, low Quality (Q)-factor tunable resonators are used, which are easier to implement and also provides a wider frequency coverage. The L-ISO device supplies enough isolation for the R_X input not to significantly load the T_X output, the main additional T_X load being due to the two Feed-Forward Paths (FFP) [FFP_1 and FFP_2].

The function of FFP_1 and FFP_2 is to null-out the transfer from the T_X to the R_X port at both the R_X and T_X frequencies, which is done by adding a phase shifted (anti-phase) and properly attenuated version of the T_X signal to the R_X port. The attenuation and phase of the FFPs should be adjusted so that the T_X signal fed forward will cancel the T_X signal leaking through the L-ISO at the center frequencies of the T_X and R_X signals. The individual FFP consist of a tunable Phase-Shifter (φ) and a tunable Attenuator (A). Control range, insertion loss and linearity are the key metrics for these building blocks. The most important metric, however, is the power

breakdown of the passive (inductors [L] and capacitors [C]) and active devices, and it is therefore common that circuits close to the antenna are implemented in either Silicon On Sapphire (SOS) or Silicon On Isolator (SOI) processes. The substrate can then be high-resistive, about 1000 Ωcm, compared with a bulk RF CMOS process of 10 Ωcm at most. This enables good passive components with high Q-values and Self Resonance Frequencies (SRF), but also the option of isolating the individual transistors from each other, which is very beneficial when designing stacked transistors in switches and power amplifiers (more stacked devices can be used than in bulk CMOS).

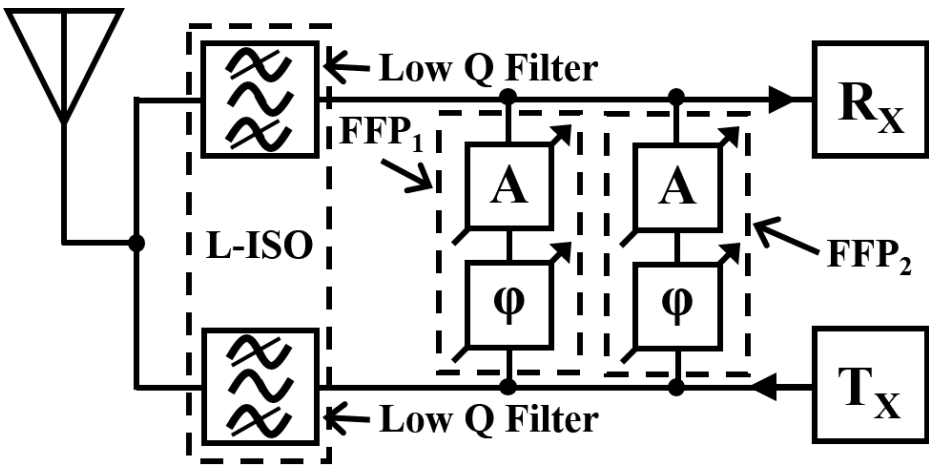

Fig. 2.4. The Block diagram of an active duplexer system.

In Fig. 2.5 several versions of balancing duplexers are shown. Fig. 2.5a shows a balancing duplexer, which is very similar from a system point of view to the active duplexer in Fig. 2.4. The balancing duplexer consists of a Transformer (T) connected to the antenna, R_X input, and T_X output, which is fairly close to the L-ISO device, and the resistor R and the capacitor C is equivalent to a single FFP used in the active duplexer. The differences, on the other hand, is that the R and C in the balancing duplexer also helps with the impedance matching of the antenna, and the R and C path also has much larger current amplitudes, causing losses contributing to increased IL.

The balancing duplexer in Fig. 2.5b uses a single-ended "Dummy-Load". The main concept is to cancel the T_X-carrier and the T_X-noise at the R_X input by the means of balancing. If the dummy load, represented by R_{ant} and C_{ant} in the figure, is equal to the antenna impedance, then equal T_X currents (in both magnitude and phase) will flow from the PA connected to the center tap to each of the antenna and dummy load terminals. Since the two currents will then flow in opposite directions through equal transformer primary windings, no signal will be induced in the transformer

secondary, where the R_X is connected. With a perfectly balanced structure, i.e. symmetric transformer, and perfectly matched dummy load to the antenna, the T_X to R_X isolation becomes infinite. It should be noted, however, that due to capacitive coupling between the transformer primary and secondary side, some T_X signal will appear in Common Mode (CM) at the LNA input, requiring a certain CM Rejection Ratio (CMRR) of the LNA, but the required CM voltage/power range will be more important at full T_X power. Further, the PA feeds equal current to antenna and dummy load, so fundamentally half the power goes to the antenna and the other half goes to the dummy load. This feature results in a loss of 3dB, in addition to losses of the transformer. Furthermore, the noise of the dummy load will equal that of the antenna at the R_X input, so the NF of the R_X is also at least 3dB. Finally, another bottleneck is associated with the tunable dummy load, whose tunable impedance has to be equal to the antenna impedance at all times over the frequency ranges of interest, or the balancing technique will fail. So with today's antennas and together with user interactions, the tunable dummy load must cover most of the possible antenna impedances and the tunable circuit elements in the load must also have high Q-values, good tunability, and very high voltage and current breakdown levels. The normal time-frame for user interaction is around 1 ms [23], so the tunable load does not have to be varied at a fast pace. To reduce the impedance tuning range of the dummy load an adaptive impedance tuner can be placed between the duplexer and antenna, but at the expense of increased IL.

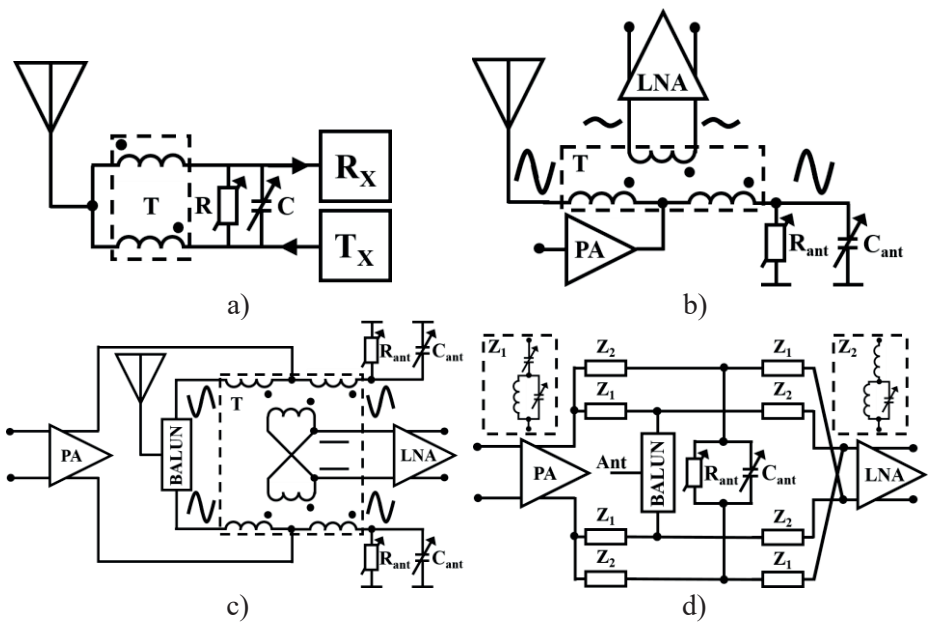

Fig. 2.5. Balancing duplexers. (a) Narrow-band version [34].
(b) Wideband duplexer [35]. (c) Full balanced wideband duplexer [36].
(d) Hybrid-filter version with low loss [37].

In Fig 2.5c a fully balanced balancing duplexer is shown. The main difference from the balancing duplexer circuit in Fig. 2.5b is that the whole structure is differential, and thereby the common-mode voltage of the R_X input due to capacitive coupling in the transformer is cancelled. The power handling capabilities are thereby greatly improved. However, all other bottlenecks associated with the structure in Fig. 2.5b still exist, and finally an extra BALanced to UNbalanced (BALUN) component is required between the duplexer and antenna.

The differential balancing duplexer in Fig. 2.5d includes filters, which eliminates most of the power dissipation in the dummy load (R_{ant}, C_{ant}). This structure uses two types of resonators/filters, Z_1 and Z_2, the first resonator/filter Z_1 being tuned to a low impedance at the T_X-carrier frequency and a high impedance at the R_X, whereas the resonator/filter Z_2 is tuned to a high impedance at the T_X-carrier frequency and a low at the R_X. The PA is connected differentially to the antenna port (BALUN) through Z_1 and to the dummy load through Z_2. This means that T_X-carrier at the antenna is filtered/suppressed at the R_X-band frequency, and the T_X signal across R_{ant} and C_{ant} is filtered/suppressed at the T_X-carrier frequency. A lower T_X signal amplitude is then present at the dummy load, which reduces the power dissipated. The LNA is connected differentially to the dummy load through Z_1 and to the antenna through Z_2, which means that the LNA is isolated from R_{ant} and C_{ant} at the R_X-carrier frequency, and that there is a low loss path from the antenna to the input of the LNA, which means that the LNA can be designed with conventional input impedance. To further improve the isolation between the PA and LNA, the balancing concept is also used in this solution. The strong T_X-carrier with noise in the R_X-band is canceled at the LNA input if the dummy load impedance is equal to the antenna impedance, since the cross-coupled and non-cross-coupled signal paths from PA to LNA then become equal.

2.2 The Homodyne Receiver

2.2.1 Basic Noise Properties

The overall Noise Figure (*NF*) is considered the key figure of merit of the R_X, since for a certain signal strength, in absence of strong interference, this will set the Signal-to-Noise Ratio (SNR) level in the detector and thus the Bit Error Rate (BER) of the received data. A good starting point for understanding the noise figure (*NF*)/noise Factor (*F*) in an R_X, is provided by Friis' formula (2.1).

$$F_{Tot} = F_1 + \frac{F_2-1}{G_1} + \frac{F_3-1}{G_1 G_2} + \cdots + \frac{F_n-1}{G_1 G_2 \cdots G_{n-1}} \approx F_1 + \frac{F_{rest}-1}{G_1} \qquad (2.1)$$

The equation describes the total noise factor of a cascade of circuit blocks, where F_1, G_1 are the noise factor and gain of the first block, F_2, G_2 of the second, etc. Let us now consider the R_X as a cascade of such blocks. If a circuit with high gain (G_1) is placed in the beginning of the R_X-chain, the equation tells us that the gain of this first block (G_1) will suppress the noise factor (F) of the circuit blocks further down the R_X-chain. The total noise factor will then be dominated by that of the first stage (F_1), so by having the first stage gain high and noise factor low, the entire receiver will have high noise-performance. However, there is always another side to the story, and to place a high gain circuit in the beginning of the R_X-chain might cause linearity problems, due to large signals further down the R_X-chain. The noise figure (NF) is closely related to the noise factor (F), it is merely a conversion to dB units (2.2).

$$NF(dB) = 10 \cdot log_{10}(F) \tag{2.2}$$

2.2.2 Basic Linearity Properties

The other important performance of the R_X scenario is linearity. Lack of linearity, i.e. non-linearity, gives rise to distortion. Depending on the signals involved and type of non-linearity we can distinguish different types of distortion. One type of such distortion is called signal compression. In Fig. 2.6 signal compression and the measurement of Compression Point (CP) is illustrated. The CP is defined as the Input referred CP (ICP) or Output referred CP (OCP), these CP are also defined by the amount of signal compression in dB by the index, and as an example: a 1dB Input referred CP is referred to as ICP_1. The meaning of ICP_1 is that this input power (P_{In}, Fig. 2.6) due to compression reduces the Power Gain (G_P) by 1 dB compared to the small-signal value.

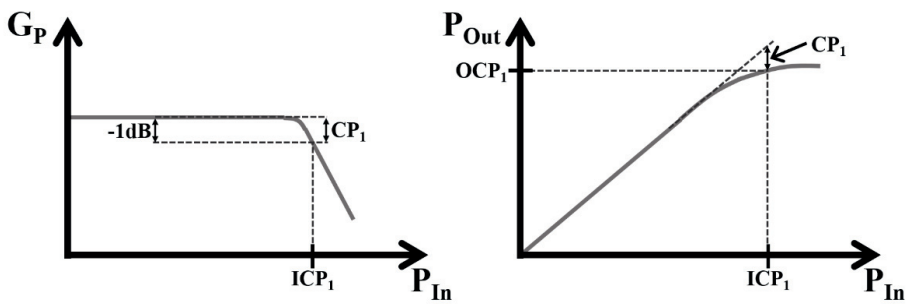

Fig. 2.6. Illustration defining 1dB compression point.

The decrease in G_P will consequently decrease the output power of the fundamental-signal at the circuit output. This is due to a non-linear transfer characteristic, where the gain/slope is lower at higher input signal levels. Fitting a polynomial to such a non-linearity gives odd-order coefficients with a negative sign. In general, however, there are also non-zero even-order coefficients. The polynomial transfer will depend on signal level and the coefficients give rise to Harmonic Distortion (HD). This distortion occur at new frequencies, at multiples of the frequencies of the signal components of the input signal.

The HD is typically not a major issue for the modern R_X, because the distortion occurs at frequencies outside the band of interest, but there may be cases where interferers occur at say a third of the received frequency so that third-order harmonics may disturb reception. A worse problem, however, is typically the so-called Inter-Modulation Distortion (IMD), which occurs when two or more signals with different frequencies enter a non-linearity, and the signals then mix with each other to generate distortion at new signal frequencies. The main difference between HD and IMD is that the new IMD tones are often/generally close to the original input signals, making them hard to suppress with filters.

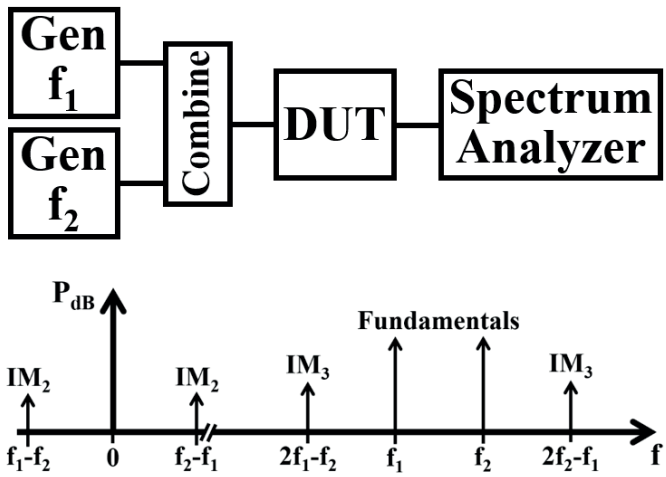

Fig. 2.7. The test bench and visual concept of a two-tone-test.

In order to benchmark the IMD performance, a two-tone-test can be performed either by measurement or by simulation, as shown in Fig. 2.7. The two tones have the same power/amplitude, but different frequencies f_1 and f_2. IMD tones at different frequencies will then be generated due to the non-linearity of the Device Under Test (DUT). The most investigated IMD tones are the third-order ones at $2f_2$-f_1 and $2f_1$-f_2, and the second-order one at f_2-f_1. During the measurement/simulation the power of the two tones are increased in logarithmic/dB steps.

In this way the non-linear behavior of the DUT can be quantified and illustrated in a graphical plot as shown in Fig. 2.8. It is important that the diagram has a dB scale on both the horizontal and vertical axis. At low signal levels, well below compression, the slope of the fundamentals (f_1 and f_2) should then be close to one, the slope of the IM_2 close to two, and the slope of the IM_3 close to three, which can be seen in Fig. 2.8. The curves from the measurement/simulation, i.e. the fundamental, the IM_2, and the IM_3, are all linearly extrapolated from the curve-portions where the slope is constant, and this extrapolation is also shown in Fig 2.8. Finally, the intersection of the extrapolated lines are found, the Intercept Point (IP) between the fundamental and the IM_2 is called the second-order IP (IP_2), and the IP between the fundamental and the IM_3 is called the third-order IP (IP_3). The IP can also be Input- and Output- referred, named the IIP_x and OIP_x, respectively, where x is the IMD order. By using these linearity benchmarks, system design and specifications can be well defined and connected to circuit design.

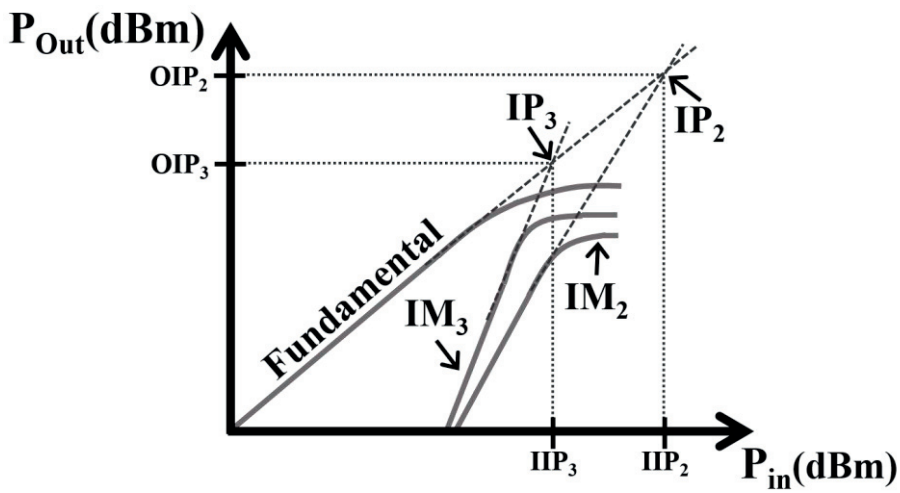

Fig. 2.8. Intercept point extraction from a two-tone measurement.

2.2.3 Cross-Modulation Distortion in the Homodyne R_X

The cellular frequency spectrum of today is a cocktail of wireless standards, with both narrow and wideband operation. For these standards to all co-exist in the same spectrum, the mobile standard protocols and the frequency bands are managed/controlled by the 3G Partnership Project (3GPP) community [38]. Achieving co-existence still becomes increasingly difficult due to the ever more crowded spectrum. This makes the R_X design a real challenge, especially the circuits close to the antenna, which operate with the high frequency RF-carrier. These building blocks, the LNA and the down-conversion mixer, are also known as the R_X front-end. Fig. 2.9 illustrates the R_X front-end in a scenario with a strong blocker

signal and LO-leakage in co-existence with the received RF-carrier at the R_X input. The linearity of the LNA is of foremost importance. If its compression point is not high enough, the strong blocker will cause gain compression of the LNA. If the blocker then has amplitude modulation, it will modulate the gain of the LNA. The LO-leakage at the output of the LNA will then be amplitude modulated. This mechanism is called cross modulation (XMD). Since in a homodyne R_X the LO-leakage is located at the center of the channel to receive, if the LO-leakage is not low enough and the LNA compression point high enough, the XMD can result in a total corruption of the received signal. If severe, the information in the RF-carrier might be undetectable (BER=50%), but a smaller blocker will cause less BER degradation.

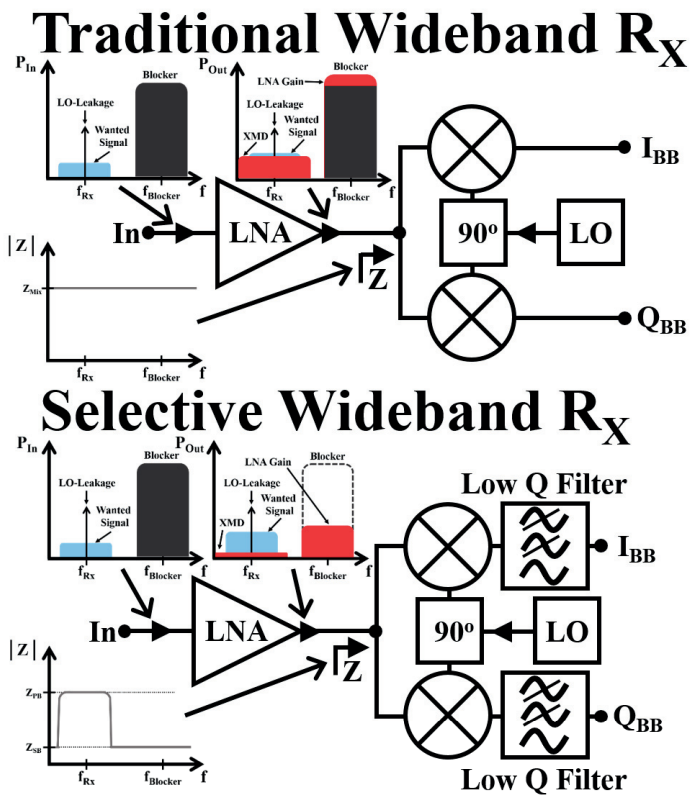

Fig. 2.9. Receiver cross-modulation scenario, due to LO-leakage and interferer.

When using a wideband R_X-chain, the LNA should have a good wideband NF, G_P, ICP_1, IIP_3, and reverse-isolation. This is required since the received carrier is very small compared to the blocker and LO signal. However, to aid the LNAs linearity, down-conversion can be performed with a current mode passive mixer [12] or a voltage mode passive mixer with frequency-translation [18]. The mixer will then

have a low input impedance at the blocker-frequency, loading the LNA output reducing its blocker voltage level, reducing the risk for gain compression. The example in Fig. 2.9 uses a voltage mode passive mixer with frequency-translation. This will increase the ICP_1, IIP_3, and also the filtering of frequencies where the wanted R_X signal is not present. Finally, by improving the selectivity, G_P of the LNA can be increased, improving the in-band noise-performance. However, the noise-performance in presence of blockers is also dependent on R_X-LO phase-noise.

2.2.4 Reciprocal Mixing and Mixer Issues in the Homodyne R_X

When a strong blocker is present at the RF-input of the down-conversion mixer, it causes the phase-noise of the LO_{Rx} at the blocker-frequency to be down-converted to DC. This phenomenon is called reciprocal mixing and is illustrated for a homodyne R_X in Fig. 2.10. It occurs when in the mixer the phase noise of the LO_{Rx} is multiplied with the strong blocker signal. The phase-noise within the frequency range $f_{Blocker}$-f_{BB} to $f_{Blocker}$+f_{BB} will after this reciprocal mixing fall on top of the received channel and degrade the noise-performance of the R_X. Although just illustrated for homodynes in Fig. 2.10, the problem of reciprocal mixing is present in all R_X architectures, i.e. also in heterodynes.

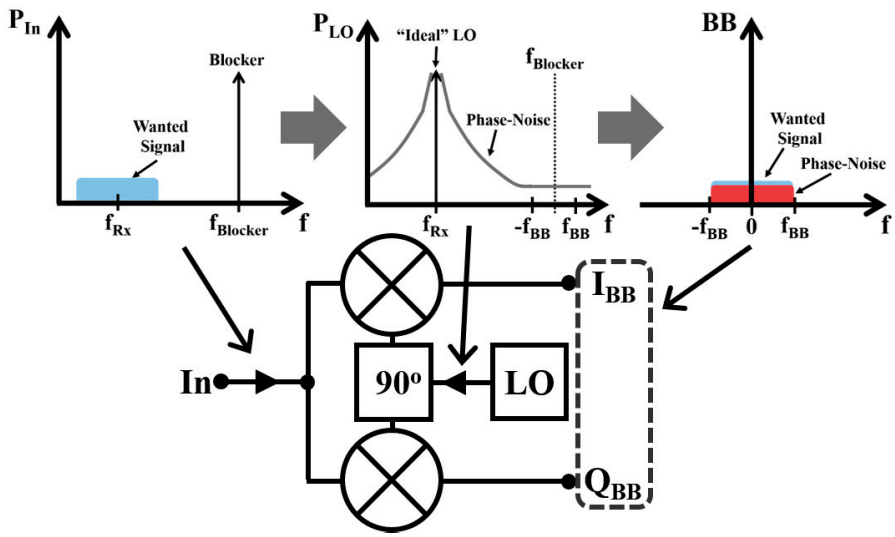

Fig. 2.10. Receiver reciprocal mixing caused by LO phase-noise and a large blocker.

LO-leakage and second order non-linearity (IIP_2) are of particular importance in down-conversion mixers for homodyne receivers. For this reason differential mixer implementations are preferred. The LO-leakage and IIP_2 are then caused by circuit

mismatch, since a perfectly symmetric mixer driven by perfectly differential signals will not have any LO-leakage or even-order non-linearity. Some degree of mismatch is, however, inevitable. The resulting LO-leakage at the mixer input will then propagate through the LNA and end up at the LNA input. In a homodyne receiver the LO-leakage will be in the passband of the antenna filter, and will thus reach the antenna with low attenuation. High reverse isolation in the LNA is thus important. The LO-leakage at the antenna must be below spectral emission limits, and additional limits on the leakage level are often put by XMD and DC-offset. Finally, the IIP_2 of the mixer is of highest concern when the transceiver operates in FDD. The self-interference from the T_X is then the main issue. The duplexer in Fig. 2.3 is used to attenuate the T_X-carrier at the R_X input, but for economic reasons and to achieve small footprint the isolation of the duplexer is limited. The strongest interferer is then typically the own T_X signal, which together with blockers must pass through the LNA and down-conversion mixer without generating significant distortion in the channel of interest. Due to mismatch, the differential mixer will produce unwanted IM_2 signals, which may end up within the Base-Band (BB) channel of the signal to receive, preventing or degrading the reception. For a single strong modulated interferer, like the own T_X signal, a second-order non-linearity will produce distortion components with twice the original interferer bandwidth, centered at DC, like the homodyne baseband signal. The mixer IM_2 requirements in the homodyne R_X are thus stringent.

2.3 The Homodyne T_X

The traditional linear T_X-chain is shown in Fig. 2.11, consisting of the digital T_X-BB, Digital to Analog Converters (DACs), two mixers also known as a quadrature modulator, a Pre-PA (PPA), and a PA. In the digital T_X-BB, BB-signals with proper pulse-shapes and filtering are generated, the digital in-phase (I_{BB}) and quadrature-phase (Q_{BB}) BB signals. The digital signals are converted to analog form in the DACs, including reconstruction filtering to remove high frequency aliases. The analog BB-signals are then frequency up-converted to RF in the quadrature modulator, by the use of the T_X-LO signal (LO_{TX}), which is generated by the T_X

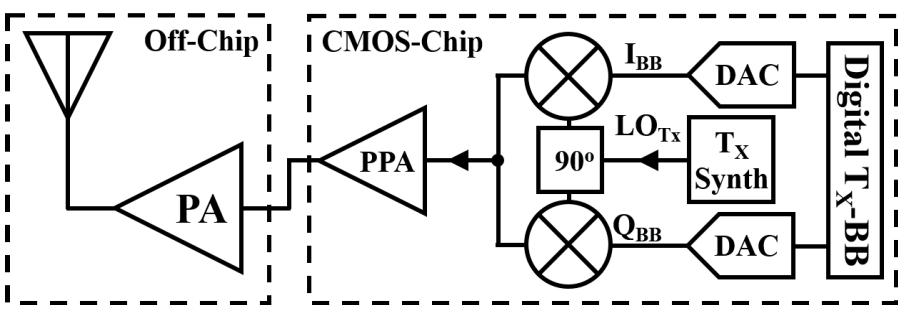

Fig. 2.11. Linear cellular homodyne transmitter.

Synthesizer (T_X-Synth) in Fig. 2.11. Further, the generated modulated RF-carrier is amplified by the PPA, which often has a variable/programmable gain to adjust the transmit power level. The PA is typically off-chip, with associated parasitic impedances of the PA/PPA interface. The PA then requires sufficient input matching, and the PPA sufficient drive capability, but by implementing the PA in a semiconductor technology optimized for that purpose, higher performance can be reached than for an on-chip CMOS PA.

In an FDD-system, the T_X transmits a signal at the same time as the R_X receives, but at different frequencies (T_X at UL and R_X at DL frequency). By operating the T_X and R_X simultaneously, the requirements on both signal-chains become more stringent, foremost in terms of T_X out-of-band noise and R_X linearity The main sources of the T_X out-of-band noise is the T_X frequency synthesizer (T_X-Synth) and the frequency up-conversion mixers. The output signal (LO_{Tx}) of the T_X-Synth has phase-noise (thermal-noise converted into time jitter inside the oscillator in T_X-Synth), which the mixers add to the signal in the frequency up-conversion and they also add their own thermal noise. The remainder of the T_X-chain amplifies the up-converted noise and adds its own noise as well, but the noise contribution of the PPA and PA is typically not dominant.

2.3.1 Self-Interference from the T_X in FDD

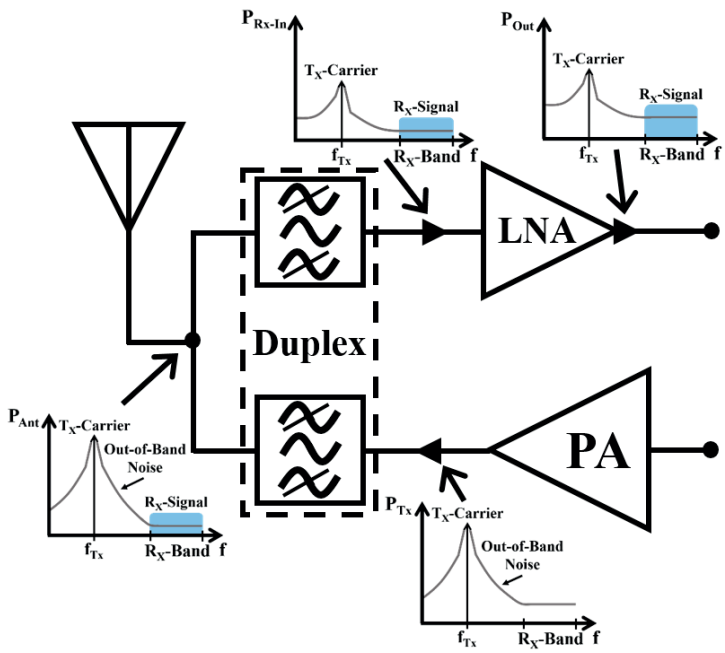

Fig. 2.12. Effects of the transmitter noise in the R_X-band.

In Fig. 2.12, T_X out-of-band noise in the R_X-band and the effects on the R_X front-end are illustrated. As can be seen there is T_X out-of-band noise at the frequency of the wanted R_X-carrier, so leakage of this noise through the duplexer can increase the noise-floor at the R_X input and degrade the R_X SNR. To not de-sensitize the R_X-chain, the T_X-noise issues must then be addressed at the source of the problem, which is the circuits that generate the modulated T_X-carrier. The T_X output noise requirement is comparably straight-forward to determine, the noise-floor in the R_X is set by the thermal noise level and the noise figure (NF_{R_X}) [39], the duplexer R_X-band suppression is 45 dB at minimum, and the T_X out-of-band output noise requirement can then be determined as (2.3). Note the degradation of the R_X noise figure (ΔNF_{R_X}) in (2.4), caused by the noise ratio (NR_{R_X}), a ratio between the noise-floor from the T_X at the LNA input ($P_{Noise-Tx-Rx}^{f=Rx}$) and the noise-floor without the T_X at the LNA input ($P_{Noise-LNA}^{f=Rx}(dBm)$, [-174 (dBm/Hz) + NF_{R_X}]). By calculating a degradation of the noise figure of approx. 1dB with (2.4), resulting in a finit noise-floor caused by the T_X, which must be approx. 6dB below the noise-floor of the R_X.

$$P_{Noise-Tx}^{f=Rx}\left(\frac{dBc}{Hz}\right) = P_{Noise-LNA}^{f=Rx}(dBm) + NR_{R_X-1dB}(dB) + A_{Duplex}^{f=Rx} - P_{Tx}^{f=Tx} = \\ = -174 + 3 + (-6) + A_{Duplex}^{f=Rx} - P_{Tx}^{f=Tx} = -177 + A_{Duplex}^{f=Rx} - P_{Tx}^{f=Tx} \quad (2.3)$$

$$\Delta NF_{R_X}(dB) = 10 \cdot log_{10}(1 + NR_{R_X}) = 10 \cdot log_{10}\left(1 + \frac{P_{Noise-Tx-Rx}^{f=Rx}}{P_{Noise-LNA}^{f=Rx}}\right) \quad (2.4)$$

A more complex situation is self-interference together with an external blocker, as shown in Fig. 2.13. The worst case situation is when the UE is located far-away from the Base-Station (BS), and the received signal is small and close to the thermal-noise-floor. Far from the BS the UE will also generate a strong/high-power T_X-carrier. The T_X-carrier is suppressed by about 50 dB in the duplexer before it reaches the R_X-input. The blocker, however, is close in frequency to the received signal and is therefore not attenuated significantly before reaching the R_X-input. The blocker is also of relatively strong/high-power, with respect to the small received signal, and the blocker and T_X-carrier will cross-modulate to form XMD in the LNA. Further, all the signals are now present at the R_X-mixer input, i.e. the wanted received signal, the T_X-carrier, the blocker, and the XMD. Now as the mixer down-converts the weak wanted signal, it also generates IMD of the second-order (IMD_2), caused mainly by the T_X-carrier, which is stronger than the blocker. This scenario sets the linearity requirements of circuit blocks in the R_X front-end, and that is the IIP_3 of the LNA and the IIP_2 of the down-conversion mixer. Finally, to ease the effects of the self-interference, cancelation paths like the FFP of the active duplexer (Fig. 2.4) can be used, as shown in [40]. This is possible because the T_X-carrier is generated locally in the UE and it is hence available for cancellation.

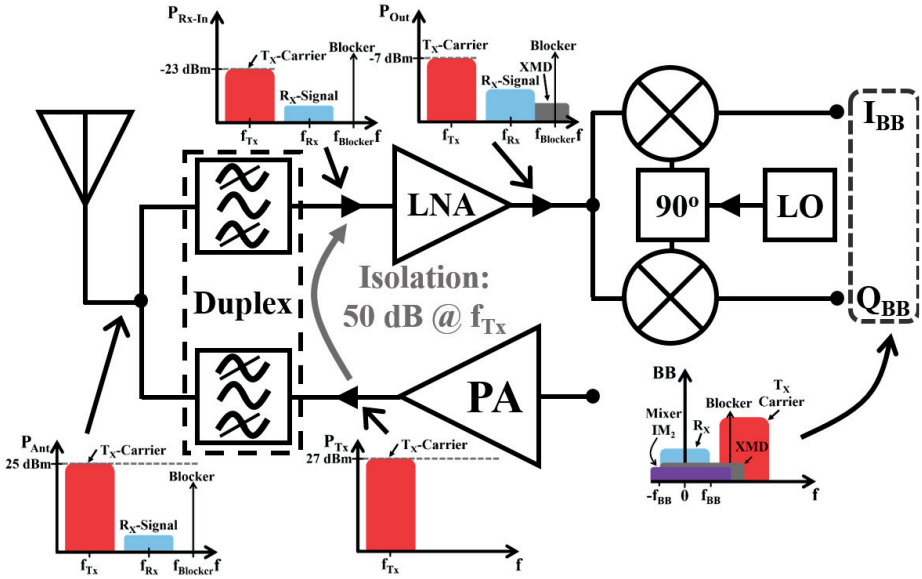

Fig. 2.13. Transceiver scenario with self-interference and strong blocker.

2.3.2 T_X Linearity and Noise

When discussing T_X linearity, the distortion can be divided into three main categories, distortion within the carrier bandwidth, spectral-regrowth around the carrier, and out-of-band distortion. The distortion within the carrier corrupts the transmitted information and hence increases the BER at the receiver. The spectral-regrowth around the carrier is caused by IMD and will disturb nearby channels within the same frequency-band. The out-of-band distortion is mainly due to HD generated within the T_X-chain, but also spurious tones from the LO-generation, like the reference or fractional spurs from the Phased Locked Loop (PLL) or leakage in frequency dividers/multipliers. This out-of-band distortion is often less of a concern, since it is less expensive to filter-off from a BOM and RF-performance perspective.

The main measure for in-band carrier distortion is the Error Vector Magnitude (EVM), illustrated in Fig. 2.14 and clarified mathematically in (2.5)-(2.7). In Fig. 2.14, the transmitted signal is represented by the vector named "Real" and that vector is compared to the reference/ideal vector named "Ref". The result of the comparison is an Error Vector (EV) named "EV". The magnitude of the EV is normalized with the magnitude of the "Ref"-vector, resulting in the *EVM* over total symbol time (T_S), and this commonly displayed in either percent (2.8) or in decibel (2.9). Finally, a more in-depth representation of the in-band carrier distortion is the Amplitude-Modulation to Amplitude-Modulation (AM-AM) and Amplitude-Modulation to Phase-Modulation (AM-PM) distortion, see example in Fig. 2.15.

$$EV(I,Q) = (I_{Real} - I_{Ref}, Q_{Real} - Q_{Ref}) \tag{2.5}$$

$$\overline{P_{EV}} = \frac{1}{T_S}\sum_{t_S=1}^{T_S} \frac{|EV(I,Q)|_{t_S}^2}{2R_L} \tag{2.6}$$

$$\overline{P_{Ref}} = \frac{1}{T_S}\sum_{t_S=1}^{T_S} \frac{|Ref(I,Q)|_{t_S}^2}{2R_L} \tag{2.7}$$

$$EVM(\%) = \sqrt{\frac{\overline{P_{EV}}}{\overline{P_{Ref}}}} \cdot 100 \tag{2.8}$$

$$EVM(dB) = 10 \cdot log_{10}\left(\frac{\overline{P_{EV}}}{\overline{P_{Ref}}}\right) \tag{2.9}$$

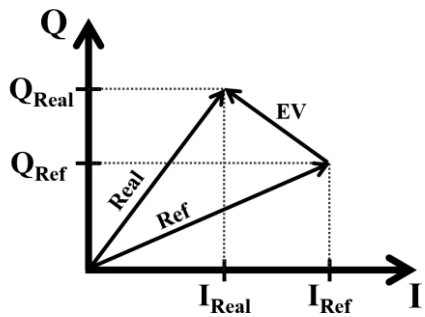

Fig. 2.14. Graphical representation of the error-vector.

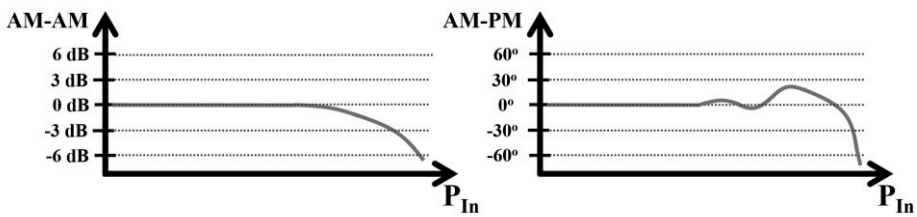

Fig. 2.15. Example of AM-AM and AM-PM charteristic of a power amplifier.

This example shows how a linear amplifier operates with constant gain at the lower P_{In} levels, but as input power is increased, the amplifier starts to compress and the gain decreases, which can be observed as the negative trend in the AM-AM characteristic. Simultaneously, the AM-PM curve has a constant phase at low input power levels, but as the power levels increase the AM-PM starts to shift due to signal amplitude dependent parasitic capacitances inside the active circuit elements of the amplifier.

The spectral-regrowth around the carrier is due to odd-order non-linearity, typically dominated by IMD of third-order (IMD_3) and fifth-order (IMD_5). To quantify the effect of spectral-regrowth in the radio environment, the Adjacent Channel Leakage Ratio (ACLR) can be measured, with the basic example illustrated in Fig. 2.16. For the older radio standards such as 2G and 3G, the ACLR-measurement was straightforward, and the spectral power of each adjacent channel (high-side and low-side) were then measured and compared to the spectral power of the main carrier.

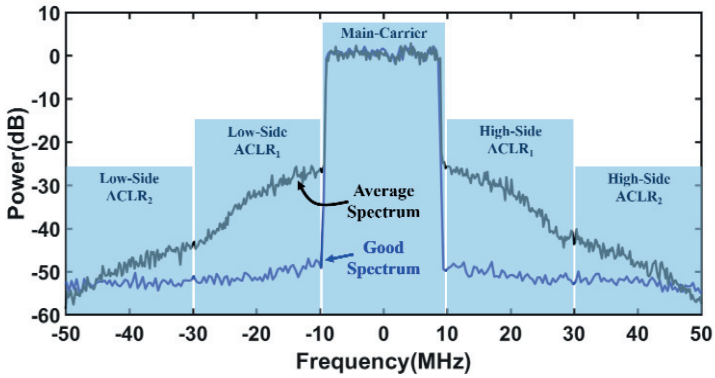

Fig. 2.16. Example of traditional ACLR measurement.

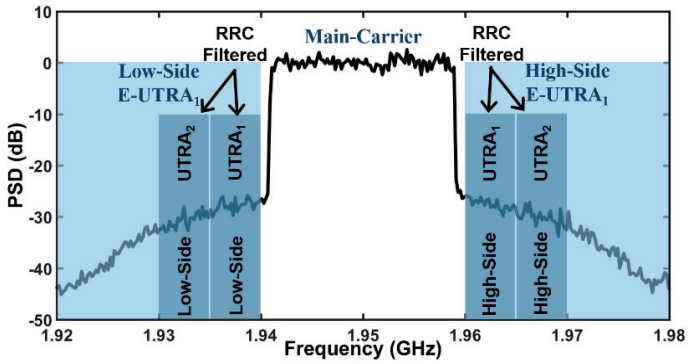

Fig. 2.17. The output spectral measurement of an LTE transmitter in band I.

For LTE (4G) there are two categories of ACLR to measure, UMTS Terrestrial Radio Access (UTRA) and Evolved UTRA (E-UTRA) [41]. This is to ensure that LTE can coexist with both W-CDMA (UTRA) and with itself (E-UTRA). The most common ACLR measurement for LTE UE is the E-UTRA measurement, which compares the average power in the transmitted carrier with the average power of the adjacent-channels. This provides a measure of how much the transmitter affects the adjacent LTE channels, when received in the BS. On the other hand, the measured ACLR for coexistence with W-CDMA systems is a bit more complex and requires a Root Raised Cosine (RRC)-filter, which is used to limit Inter Symbol Interference (ISI) in W-CDMA. The RRC-filter is applied to the adjacent channels where ACLR is measured, and where W-CDMA communication is assumed to take place. The average filtered signal power is then compared with the average power of the LTE carrier. The measurement of E-UTRA and UTRA on an LTE Single-Carrier Frequency-Division Multiple Access (SC-FDMA) 20MHz bandwidth T_X output signal is shown in Fig. 2.17.

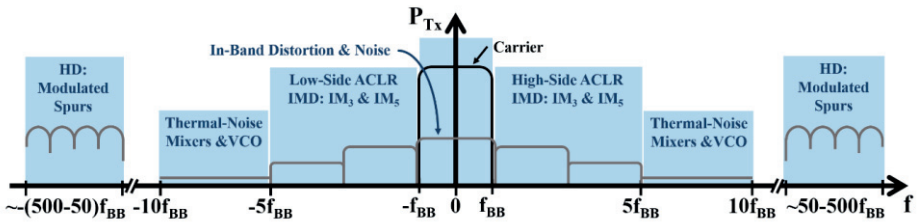

Fig. 2.18. Frequency-offset regions for transmitter non-idealities.

The out-of-band spectral distortion is commonly of less concern, if it is mainly HD. The thermal noise, on the other hand, is of more concern when the T_X is used in FDD. The maximum allowed power of any spurious tone is about -60dBc (-29dBm) at the antenna and the noise requirement in the R_X-band is about -160dBc/Hz, both at peak T_X-power. In Fig. 2.18, the frequency-offset dependence on different non-idealities is shown. Within the carrier, noise and in-band distortion dominates, affecting the final BER in the receiver. Just outside the carrier, the IMD dominates, but further out the thermal noise from the LO-synthesizer and mixers dominates. The out-of-band spurs, caused by HD is the furthest away from the carrier.

CHAPTER 3

3 User-Interaction and Adaptive Matching in UE

In the modern cellphone, using multiple frequency bands and MIMO, antennas are everywhere, and it is thus impossible to avoid that some antennas become affected by the user grip. Especially the sub-1 GHz antenna/antennas, which are integrated as a part of physical cellphone structure, are susceptible to user-interaction, which changes not only the impedance [23], but also the radiation pattern [42]. For the impedance variation, the antenna can be designed to be more resilient to loading-effects caused by its surroundings. To further mitigate the antenna impedance variation, a tunable impedance matching network can be used, i.e. an impedance tuner. To provide good matching with minimum effects on the overall RF-system performance, the impedance tuner must, in addition to sufficient impedance coverage/tuning range, also have low loss and high linearity. When it comes to the radiation pattern, it can be changed by using either beamforming [43] or by switching/changing antennas based on the user-interaction. The beamforming approach is more suitable for mm-waves than sub-1 GHz, since it requires an array of antennas and each antenna has to be approximately a quarter wave-length ($\lambda/4$) and the antenna spacing is also wave-length dependent. At 1 GHz the wavelength is 30 cm and an antenna array for beamforming would thus not fit inside a cellphone.

3.1 The Effect of User-Interaction on Antenna Impedance

The impedance (Z) (3.1) has a real part, the resistance (R), and an imaginary part, the reactance (jX). To display the impedance (Z), the Smith-Chart [44] is the norm in RF-design. It displays impedance normalized to Z_O, which is often equal to 50 Ω for single-ended and 100 Ω for differential impedances (50 Ω per side). The Smith-Chart is shown in Fig. 3.1a, with the real 50 Ω impedance located in the center, corresponding to a normalized impedance equal to 1. When the Smith-Chart is used in measurements/simulations of impedance matching, the reflection (3.2) of the radio-wave that encounters the impedance is the key parameter of interest. The reflection is generally quantified by the magnitude in dB of S_{11} (3.3) [44] and this represents the part of the incident radio-wave power (P_{in}) that is reflected back (P_r), see Fig. 3.1b. The amount of reflection can be seen in the Smith-Chart as the distance from the center, where the middle point is perfectly matched in a 50Ω

system, with zero reflection, whereas impedance points at the very edge of the Smith-Chart correspond to total reflection, i.e. all incident signal power is reflected.

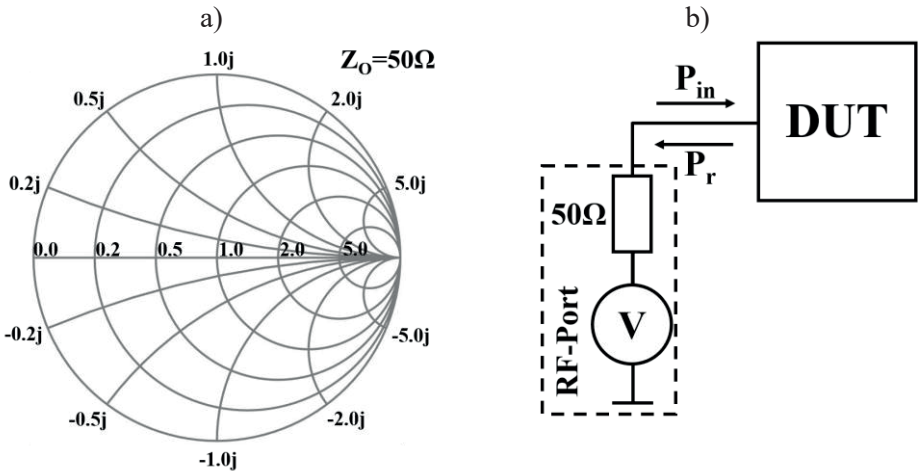

Fig. 3.1. The RF-impedance. (a) The Smith-Chart.
(b) Verification of input-matching (S_{11}).

$$Z = R + jX \tag{3.1}$$

$$\Gamma = \frac{V_r}{V_{in}} = -\frac{I_r}{I_{in}} = \frac{Z - Z_O}{Z + Z_O} = \frac{Z_{DUT} - Z_O}{Z_{DUT} + Z_O} \tag{3.2}$$

$$S_{11} = 10 \cdot log_{10}\left(\left|\frac{1}{\Gamma^2}\right|\right) = -20 \cdot log_{10}(|\Gamma|) = -10 \cdot log_{10}\left(\frac{P_r}{P_{in}}\right) \tag{3.3}$$

To clarify, when a radio-wave is propagating through a medium, the impedance at both ends has to be conjugate matched or reflection will occur. When there is a reflection a standing wave will occur, which is exactly the same phenomenon as the standing acoustic waves well-known from basic physics, but this will only happen when the electrical length of the mediums/components is at least one tenth of the radio-wave's wave-length. For severe mismatch at the DUT input, the DUT cannot effectively use the power from the incoming radio-wave, because there will not be sufficient power transferred to the DUT input. Finally, a different merit to measure the amount of wave-reflection, which is commonly used in antenna design, is the Voltage Standing Wave Ratio (VSWR). The *VSWR* (3.4) is defined as the ratio of the standing wave voltages, i.e. the maximum voltage amplitude (3.5) and the minimum voltage amplitude (3.6) of the standing waves, all equations from [44]. The equivalent S_{11} and *VSWR* values in a Smith-Chart is shown in Fig. 3.2.

$$VSWR = \frac{|V_{Max}|}{|V_{min}|} = \frac{1+|\Gamma|}{1-|\Gamma|} = \frac{1+\sqrt{\frac{P_r}{P_{in}}}}{1-\sqrt{\frac{P_r}{P_{in}}}} \tag{3.4}$$

$$|V_{Max}| = V_{in} + V_r = (1 + |\Gamma|) \cdot V_{in} = \left(1 + \sqrt{\frac{P_r}{P_{in}}}\right) \cdot V_{in} \tag{3.5}$$

$$|V_{min}| = V_{in} - V_r = (1 - |\Gamma|) \cdot V_{in} = \left(1 - \sqrt{\frac{P_r}{P_{in}}}\right) \cdot V_{in} \tag{3.6}$$

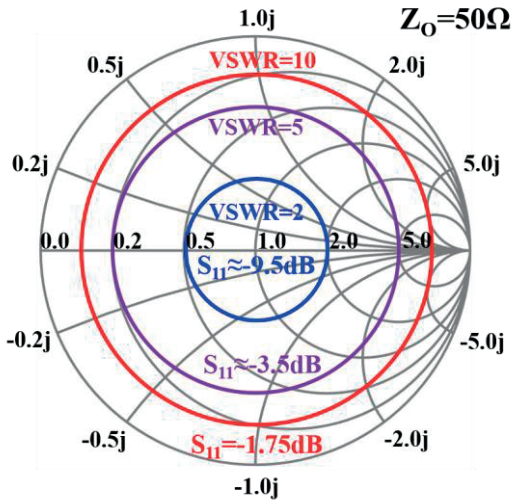

Fig. 3.2. The VSWR-circles with S_{11} equivalent values in the Smith-Chart.

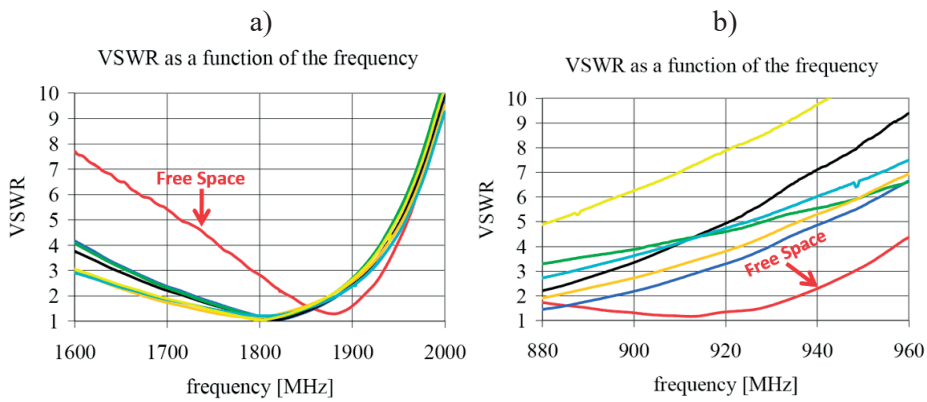

Fig. 3.3. Antennas with and without user.
(a) High-band antenna. (b) Low-band antenna.

For a UE antenna, achieving a favorable "free-space"/unloaded antenna impedance is of less concern, since anything like the "free-space" scenario is very unlikely to occur in practical user cases. As the user interacts with the UE, loading the antennas with their hands and head, the loading will affect antenna behavior. In Fig. 3.3 are two examples from [45], the first showing the loading behavior of a High-Band (H-Band) antenna (Fig. 3.3a), and the second is an antenna for the Low-Band (L-Band) (in Fig. 3.3b). The antenna in the first example is affected less by the user since it has less surface area and it is less accessible for the user. In the second example, the "free-space"/unloaded antenna has a better than average impedance match, but the impedance is more affected by the user, since the antenna has a larger surface area.

3.2 Adaptive Impedance Tuners

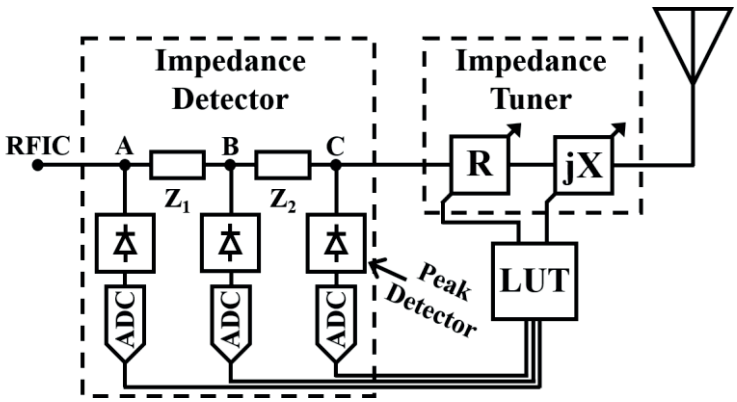

Fig. 3.4. The adaptive antenna impedance tuner system.

The antenna impedance varies due to changes/movements in the close antenna surroundings. These variations are not fast, it takes parts of seconds, not milliseconds, for the user to change grip or position. Therefore, the circuitry controlling the impedance tuner to mitigate the variable antenna impedance does not have to be fast. In Fig. 3.4 the adaptive impedance tuner system is shown, the core of the system is an impedance tuner and an impedance detector, and the impedance tuner is controlled based on measurements from the impedance detector. The impedance tuner is controlled by a digital word, where the number of bits set the maximum impedance resolution, and the impedance coverage is set by the tunability of each tuning-element and the number and/or type of elements in the impedance tuner. The impedance detector can either be a directional coupler (big and bulky), or as shown in Fig. 3.4, a lumped detector consisting of two known impedances elements (Z_1 and Z_2) and three peak detectors [45]. The voltage signal levels in nodes A, B, and C are detected by peak detectors connected to each node, whose outputs are then converted into digital data by Analog to Digital Converters

(ADCs). The control of the impedance tuner can then be performed by a Look-Up Table (LUT).

3.3 Impedance Tuners in CMOS-SOI

In the choice of technology for the impedance tuner in [**I**], the main requirements are good power-handling capability and sufficient linearity, as well as high isolation and low loss, for both passive and active/switch-components. Low cost was also a key parameter, putting the focus on silicon technologies. The main options investigated were Micro-Electro-Mechanical Systems (MEMS), standard CMOS, and CMOS-SOS or CMOS-SOI. MEMS was considered in the beginning of the project, but reliability was of great concern and advances in the other technologies were rapid, largely caused by circuit design innovations. For the standard CMOS technology, if compared with CMOS-SOS or CMOS-SOI, it did not provide as high performance, which could also be concluded by looking at the groups' previous work [23] and [46]. CMOS-SOS and CMOS-SOI were then the only options left, both technologies yielding high performance and both being mature enough for mass-production. The processes were used in antenna-switches, CMOS-SOS in [47] and CMOS-SOI in [48], which is as close to an impedance tuner as it gets in terms of circuit requirements. Through our cooperation with Dr. Andreia Cathelin at ST-Microelectronics (STM), Grenoble, France, we have access to CMOS-SOI technology and fabrication, and for this project we decided to use an STM 130nm CMOS-SOI process. The passive components in the process have low losses, and capacitor Q-values ranging in several hundreds and inductor Q-values as high as 40 can be achieved (from Electro-Magnetic [EM]-simulations). However, to save chip area and to keep the signal losses at an absolute minimum, off-chip inductors were used, which "consume no chip area" and have Q-values in the range of 100. The bulk/body of each transistor can be individually isolated, making transistor-stacking reliable and effective. The physical structure of a standard CMOS and CMOS-SOI device is shown in Fig. 3.5.

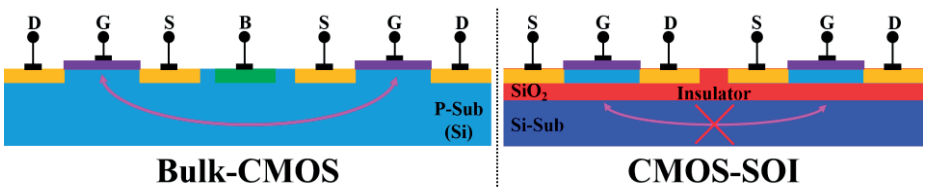

Fig. 3.5. The difference in physical structure between Bulk-CMOS and CMOS-SOI.

3.3.1 Circuit Design

The first step was to select the architecture of the impedance tuner, and a double-π network [23, 46] was chosen as shown in Fig. 3.6. To note, all the variable capacitors are shunted to signal ground, which means that at least one terminal has low signal voltage. However, "floating" variable series capacitors are also possible to implement in CMOS-SOI, since the body/bulk of the devices can be isolated from signal ground.

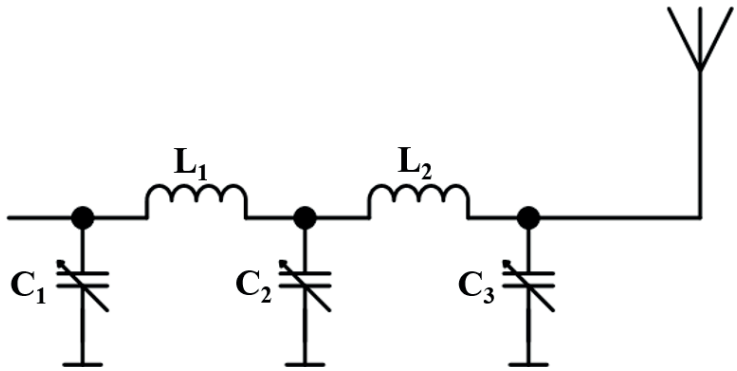

Fig. 26. The double-π impedance tuner topology.

In the next step, the voltage/power-handling and linearity properties decided the number of stacked-devices/transistors in each unit-cell in the variable capacitor. It was based on the peak RF input power, but also considering the internal peak Q-values of the impedance tuner, which is more important when the antenna impedance is up to one order of magnitude higher than the nominal 50 Ω. The voltage stress of each individual thick-oxide transistor was conservatively limited to a drain-source voltage (V_{DS}) of 2.5 V, resulting in eight stacked thick-oxide transistors per unit-cell. Further, the voltage/power handling of the capacitors was also considered. The integrated capacitors voltage handling is dependent on the physical spacing between plates, and the chosen capacitor topology was a fingered Metal-Oxide-Metal (MOM) structure, custom designed with a distance between the plates to handle the peak voltage of the circuit. All the metal-layers in the CMOS-SOI process were used, which is possible due to the high resistive substrate making the substrate losses small/non-existent. The required distance between the plates (fingers) was calculated based on a silicon-oxide maximum electric field strength of 100 V/μm [49]. However, the final breakdown voltage was higher than required, due to the minimum spacing of the top metal-layer, which was used in order to maximize the Q-value of the capacitor.

The variable/programmable capacitors could then be designed, and the schematic of a unit-cell is shown in Fig. 3.7a. The unit-cell consists of an eight-stack of thick-oxide transistors (M_1-M_8), each transistor dimensioned with the same channel-Width (W) and channel-Length (L), to make V_{DS} the same for all transistors. To ensure the same V_{DS} of all transistors, all the gate-source-voltages (V_{GS}) must also be equal, and also, all gate-drain-voltages (V_{GD}) must be equal. To get V_{GS} and V_{GD} of all transistors equal, the gate-terminals must have high-impedance (in relation to the rest of the elements in the unit-cell) to signal ground at RF. This is achieved by making the gate-resistor (R_G) large, which also improves the Q-value in the OFF-state. To ensure an even voltage distribution between each transistor terminal, for maximum voltage handling capability, the sizes of C_{gs} and C_{gd} must be equal in the OFF-state, and by using transistors with an odd number of fingers both the drain and source junctions will have the same gate-overlap and feature-size, resulting in a good match between C_{gs} and C_{gd}. To note, in contrast, in most analog-IC design an even-number of fingers is beneficial from a parasitic drain-capacitance point of view, since all drain areas can then be shared between two gates [50]. In this specific case, however, the symmetry is of foremost importance, which results in an equal voltage-division between gate-source and gate-drain capacitance for each transistor. The simplified schematic in the OFF-state is shown in Fig. 3.7b, consisting of parasitic capacitances and gate resistors.

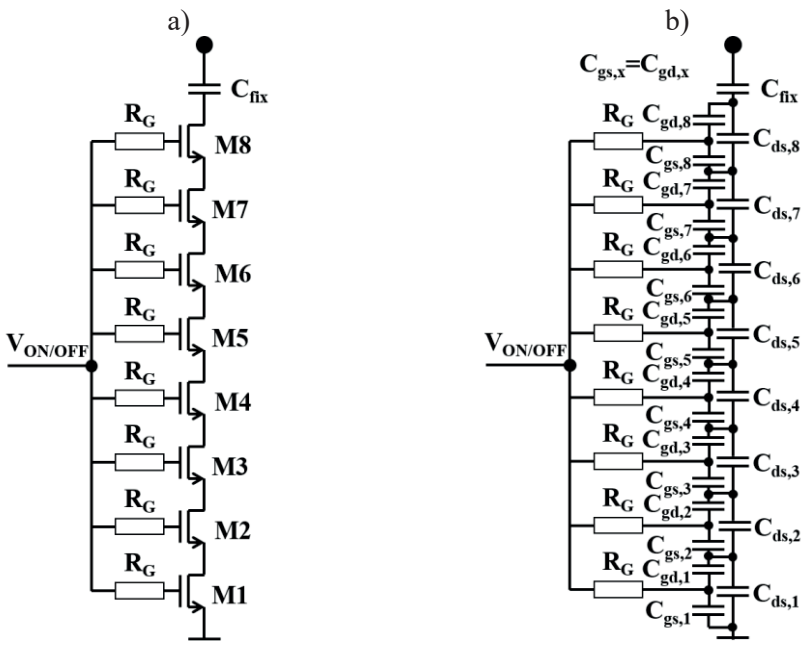

Fig. 3.7. The switched-capacitor unit-cells.
(a) Circuit schematic. (b) Circuit schematic with parasitic capacitance.

When sizing the fixed-capacitor (C_{fix}, Fig. 3.7) and transistors in the unit-cell (M_1-M_8, Fig. 3.7), the following metrics are considered, the maximum to minimum capacitance-ratio (C_{on}/C_{off}), the ON-state Q-value (Q_{on}), the OFF-state Q-value (Q_{off}), and the required impedance coverage area in the Smith-Chart. The C_{on}/C_{off} ratio is set by the ratio of C_{fix} (C_{on}, ON-State) and C_{fix} in series with the parasitic capacitance of the transistors in the eight-stack (C_{off}, OFF-state), which can be seen as a series connection of capacitors (3.7). Simulation results are shown in Fig. 3.8, and as can be seen the C_{on}/C_{off} ratio decreases by increasing the transistors width (W), as that increases the parasitic capacitances. Further, the value of Q_{on} is set by the reactance of C_{fix} and the total on-resistance of the transistors M_1-M_8 ($R_{on,tot}$), eq. (3.8), with simulation results in Fig. 3.9.

$$\begin{cases} C_{\text{off}} = \left(\frac{1}{C_{fix}} + \frac{1}{C_{Par,1}} + \frac{1}{C_{Par,2}} + \cdots + \frac{1}{C_{Par,8}}\right)^{-1} \\ C_{Par,x} = \left(\frac{1}{C_{gs,x}} + \frac{1}{C_{gd,x}}\right)^{-1} + C_{ds,x} \end{cases} \tag{3.7}$$

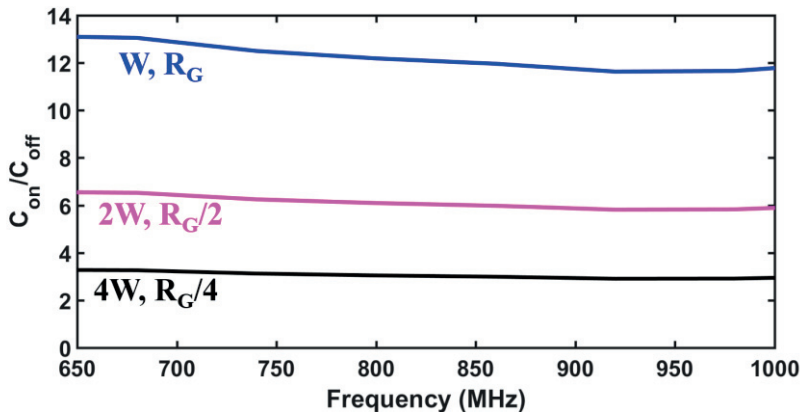

Fig. 3.8. Simulation of the C_{on}/C_{off} capacitance-ratio for different transistor width.

$$Q_{on} = \frac{1}{2\pi f_{RF} \cdot C_{fix} \cdot R_{on,tot}} \tag{3.8}$$

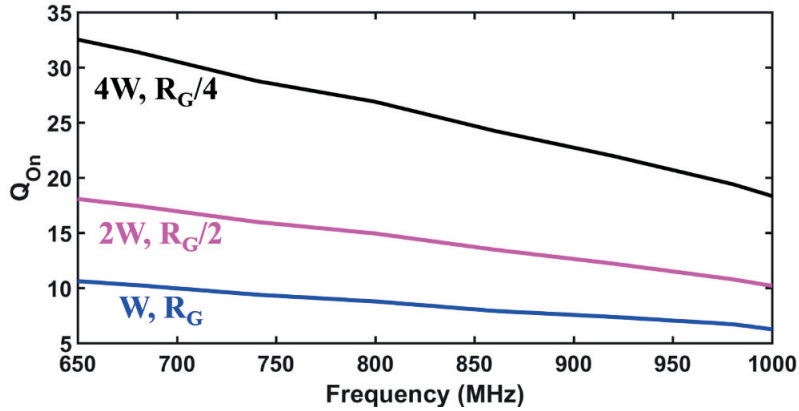

Fig. 3.9. Simulation of the Q-value in the ON-state
with variable transistor width.

The value of Q_{off} depends on the size of R_G and the size of the total parasitic capacitance (C_{off}) of all the unit-cell transistors. By assuming a large C_{on}/C_{off} ratio, all the signal voltage is located across the N transistors in the OFF-state. Further, if R_G is large the signal voltage over the transistors is divided evenly across each device and the voltage swing at each gate terminal is n/N, where n is the index number of the transistor (see Fig. 3.7). The average power dissipated in all the R_G resistors ($P_{R_G,tot}$) can then be approximated by (3.9). The average reactive power in C_{off} ($P_{C_{\text{off}}}$) is calculated with (3.10), and the Q_{off} value then becomes (3.11). Finally, the circuit simulation results of Q_{off} for different R_G values is shown in Fig. 3.10.

$$P_{R_G,tot}(N) \approx \frac{N}{3} \cdot \frac{V_{rms}^2}{R_G} \tag{3.9}$$

$$P_{C_{\text{off}}} = 2\pi f_{RF} C_{\text{off}} V_{rms}^2 \tag{3.10}$$

$$Q_{Off} = \frac{P_{C_{\text{off}}}}{P_{R_G,tot}} \approx \frac{2\pi f_{RF} C_{\text{off}} V_{rms}^2}{\frac{N}{3} \cdot \frac{V_{rms}^2}{R_G}} = \frac{6\pi f_{RF} C_{\text{off}} R_G}{N} \tag{3.11}$$

Finally, the impedance coverage area is set by the C_{off} to C_{on} range together with the inductances (L_1 and L_2). The main question is if the impedance tuner is to be designed as an off-the-shelf generic component or as a customized impedance tuner for a particular antenna with a well-known behavior. The customized version will generally have better Q-values, i.e. lower losses, since the coverage area and shape can be optimized to what is needed. The off-the-shelf version has to have a larger C_{on}/C_{off} ratio, so the overall Q-value will be lower or it may require a more expensive semiconductor technology.

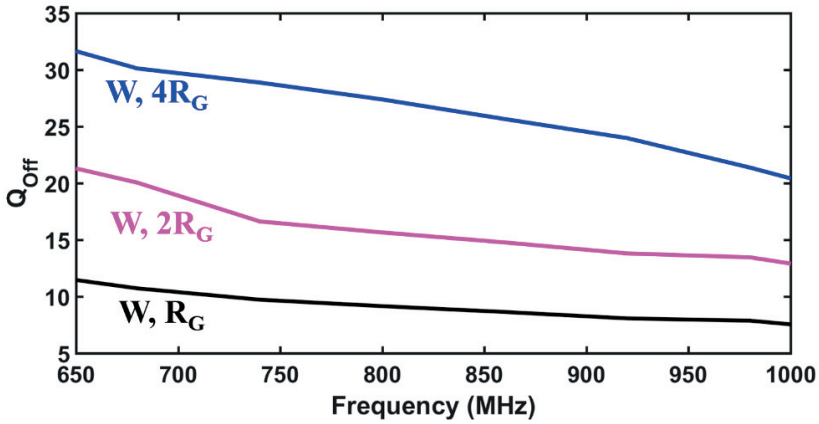

Fig. 3.10. Simulation of the Q-value in the OFF-state vs. gate resistor value.

3.3.1 Circuit Verification by Measurement

To measure and verify an impedance tuner, can be easy if all merits are measured in the nominal 50-Ω impedance state, or very complex if measured over all the impedance states, i.e. not only at the nominal 50-Ω. The main source of complexity is that neither the input/antenna nor the output/RFIC port are impedance matched to the measurement equipment in the non-50Ω states, and the internal losses in the impedance tuner give a non-uniform matching of the input and output port. This non-uniform matching means, if one port is matched, the other port can have an impedance error in relation to the internal losses of the impedance tuner. In contrast, this effect does not exist in lossless matching networks, in which, if one port is matched, the other per default is simultaneously matched.

The first measurement is the most straight-forward, to measure the impedance coverage of the impedance-tuner/DUT, and this is done with a Vector Network Analyzer (VNA), while all impedance states of the DUT are swept. This measurement can be very time-consuming, however, if a large number of impedance states are used, and an automated scripted measurement is therefore required. The measured losses consist of both the IL and the losses associated with the impedance mismatch at the ports, and this is not the right way to evaluate the internal losses in the DUT, which is also dependent on the different impedance states.

To measure and quantify the losses, two additional impedance tuners (see Fig. 3.11) are required, and they should have much lower losses and higher linearity than the DUT. The losses are divided into two metrics, the Matching Losses (ML), caused by the matching discrepancy and the Resistive Losses (RL). To measure and categorize the losses of the DUT in each specific digital state, the antenna port has

to be conjugate matched, and this is done by connecting and tuning one of the additional impedance tuners. The losses are now measured with the forward path (S_{21}) of the VNA as shown in Fig. 3.12a, the measured value includes the ML, RL, but also the RL from the added impedance tuner and the wanted value is only the ML and RL of the DUT. To eliminate the RL of the added impedance tuner, a term called de-embedding is applied, which is to find/measure the losses from components that are not the actual DUT and then deduct them from the measurement result. For the additional impedance tuner, a measurement of the total reflection was used to determine its losses. The measured value after de-embedding then only includes ML and RL, and the ML in this case is the losses due to the impedance mismatch of the RFIC port.

Fig. 3.11. The two additional impedance tuners, which are tuned manually.

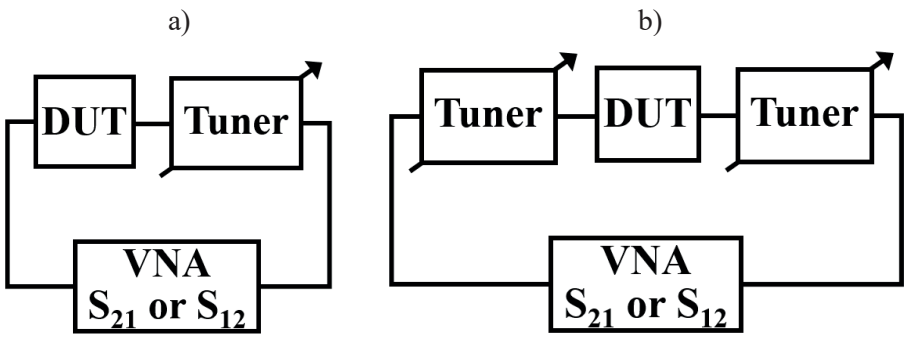

Fig. 3.12. Measurement setups. (a) To verify the matching-loss.
(b) To verify the insertion-loss.

The next measurement step is then to remove the ML component, so the only loss measured is the RL of the DUT, and this is done by matching the RFIC port with the second of the additional impedance tuners, as shown in Fig. 3.12b. Further, now the measured value includes the RL component and the RL from the newly added tuner. However, by changing the source impedance of the RFIC port, the impedance changes at the antenna port. Therefore, both of the tuners have to be tuned simultaneously to ensure that both ports are conjugate matched. The de-embedding of the additional tuners is then performed by measuring their total reflection. The main benefit of this method is that accurate measurements of other merits, like linearity, can now be performed at different impedance states of the DUT.

Fig. 3.13. The base-station amplifier with high linearity and high output power.

For the linearity measurements of an impedance tuner, mainly the IMD, a very linear and consequently high-power amplifier is required, and Fig. 3.13 shows the PA/amplifier used. To be exact, the amplifier is a PA of a BS and can deliver a very linear RF signal at very high power levels. This amplifier was borrowed from Aleksandar Mitrovic at MaxLab, Lund University, Lund, Sweden, and without it the linearity measurements presented below would not have been possible. To measure IMD with a two-tone-test, two RF-signals are required, and optimally the two signals are generated and amplified in two separate signal-chains, and combined in a combiner connected to the DUT. In this case, however, only one PA/amplifier could be obtained and for obvious reasons. Therefore, the two tones had to be combined prior to the PA/amplifier, and this can only be done if the OIP_3 of the PA/amplifier is at least one order of magnitude larger than the IIP_3 of the DUT, which it was in this case.

The IP3 measurement was performed for different impedance states of the tuner, for which both input and output port had to be matched by the additional impedance tuners, and the measurement setup is shown in Fig. 3.14. The losses of the added impedance tuners were measured, and also needed to determine the input- and output power of the two tones, which is key when measuring the distortion performance. Finally, the actual two-tone-test was performed in the traditional way, that is by increasing the power of the two tones in incremental steps. Note that the power levels used are relatively high and external attenuators (A_1 and A_2) are thus required to protect the spectrum analyzer, and these are also included in Fig. 3.14.

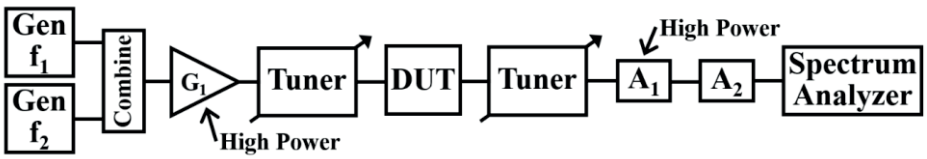

Fig. 3.14. The measurement setup for a high-power two-tone-test in a non-50Ω impedance case.

The measurement of harmonic distortion, at first glance seems straight-forward, but with the required input-power levels, the linearity requirements of the equipment generating the RF-signal is tested to the max. An additional 6th-order Butterworth-filter had to be used to suppress harmonic distortion from the amplifier and signal generator. This Butterworth-filter also has its own losses, so the output of the amplifier had to be increased, and the measurement setup is shown in Fig. 3.15. The measurement was performed by increasing the DUT input power until the maximum allowed spurious-tone/HD was obtained. The value of the input power was then retrieved, which is the highest allowed input power meeting the spectrum emission requirements.

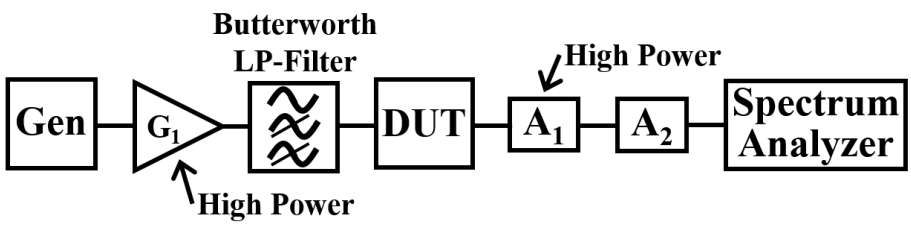

Fig. 3.15. The measurement setup to verify the harmonic distortion in an impedance tuner.

3.4 User-Interactions in a MIMO Environment

In the modern cellphone, antennas at the sub-1 GHz bands are integrated as a part of physical structure of the phone, and when MIMO is then considered, antenna-correlation comes to mind. As the user interacts with the MIMO antenna system, the non-correlation/isolation of the individual antennas can be disturbed, and the theory states that each antenna/channel-path has to be independent from any of the other antennas/channel-paths [22]. If this is not the case, the maximum gains of using MIMO cannot be achieved, and this is the case when the antenna signals become correlated. To evaluate a possibility to mitigate this effect by using adaptive impedance matching together with MIMO, several experiments were performed in [VI], which also provide a practical context for the impedance tuners in this work.

3.4.1 The Experimental MIMO UE

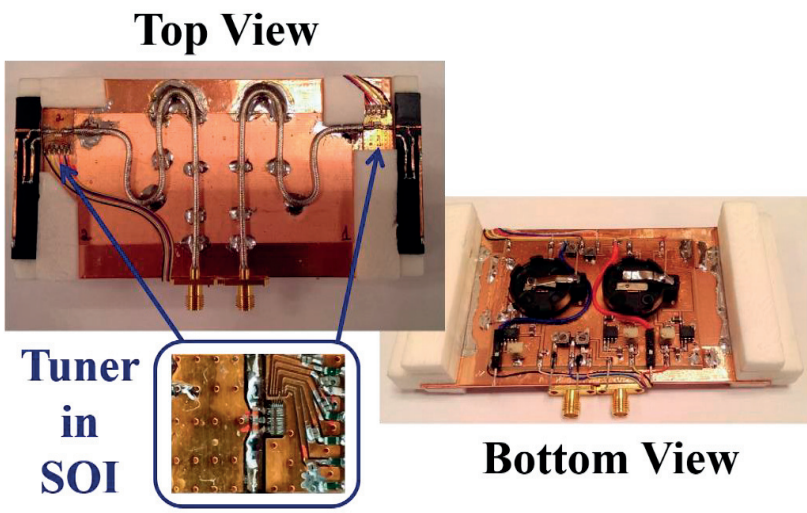

Fig. 3.16. The experimental UE with impedance tuners, used in MIMO channel measurements.

The main body of the experimental MIMO UE consists of two patch-antennas from [52]. Two impedance tuners in CMOS-SOI are added to the structure to reduce the impedance mismatch of the antennas when a human interacts with the experimental UE. This means that the impedance tuner, which consists of a tuner-chip and passive components on a PCB is mounted/assembled on to the structure of the patch-antenna. Further, all non-RF signals fed by cable, and these low-frequency signals of the impedance tuners were both decoupled by capacitors and high-impedance isolated with RF-Chokes, which are large inductors that have a high impedance at RF. This must be done, because the RF-signals are present and radiated through the

largest surface-area of the structure. The common ground plane was not used in the traditional way, since it is a major part of the L-Band antenna and the RF signals are present at the surface of the ground plane. In order to power the circuits, an additional battery driven power supply was also added to the UE. Shown in Fig. 3.16 is the top of the structure with two CMOS-SOI impedance tuners added, and the bottom of the structure with the battery supply.

3.4.2 The MIMO-Channel Measurement

To measure the behavior of the experimental MIMO UE, a channel-measurement campaign was performed including real human users. A 4-port VNA was used, with two ports connected to the handheld UE structure and the other two connected to two antennas mounted on a stand, which represented the BS. Mainly, a two-handed grip was used and the performance was measured with and without impedance matching, and the two measurements were compared. The basic principle of the measurement is shown in Fig. 3.17, and the results can be found in [VI]. To conclude, the MIMO performance was significantly improved, by up to 42%, equivalent to 3.5 dB, by the use of impedance matching with this impedance tuner [VI].

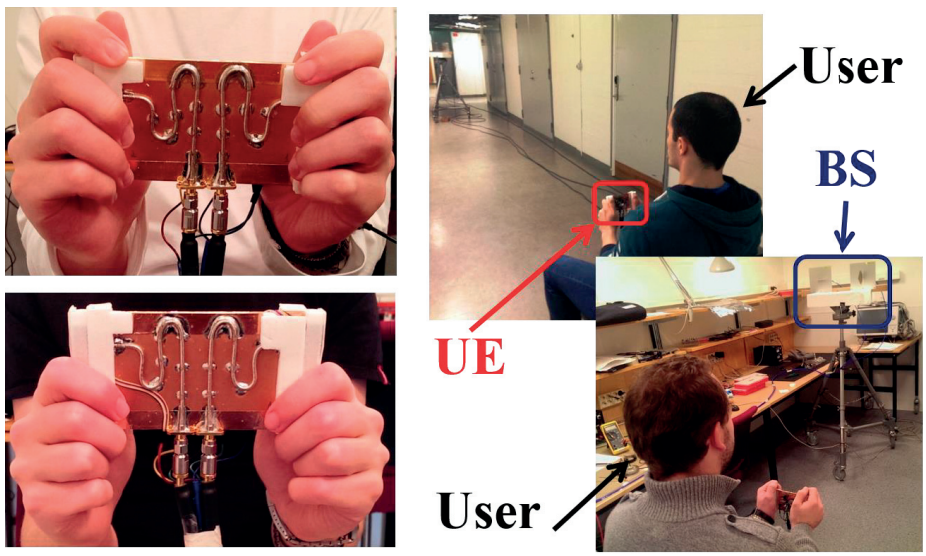

Fig. 3.17. MIMO channel measurements with different users and grips.

CHAPTER 4

4 Efficiency Enhancement Techniques for UE Transmitters

The efficiency of a battery powered UE T_X is of utmost importance, since having to charge the battery very frequently is annoying and out of convenience. To save power in the transmitter, segmentation of the individual circuit blocks has been performed and is the standard today. The PPA, the up-converting mixers, the LO-drivers and the analog-BB are then all divided into unit-cells that can be switched ON or OFF depending on output power level. The amount of T_X output power that is needed depends mainly on the distance between the UE and the BS. This segmentation can also be performed in the RFPA, but is it enough? A cellular RFPA is a major power consumer requiring in the range of 1-10W at peak power. To save power most cellular RFPAs are biased with a current that depends on output power level, and also the supply voltage is reduced to save power when the UE is close to the BS. By combining this with segmentation, even more power can be saved at low output power levels. However, this is no longer sufficient, since higher order modulation is used in modern cellular standards. Such modulation schemes require high transmitter linearity, reducing the achievable efficiency. A key obstacle to achieving high efficiency is the non-constant signal envelope. A measure of the level of envelope variation is the Peak-to-Average Power Ratio (*PAPR*) (4.1), which is the ratio between the peak output power (P_{Peak}) and the average modulated output power (P_{Avg}). The power amplifier must be biased so that it can amplify the signal peaks with sufficient accuracy, while it on average delivers much less output power. If for example PAPR is 10dB, the average power is just 10% of the peak power. The bias of the PA must then be set so that the signal compression is acceptable at peak power, which causes significant power dissipation, while the amplifier on average only produces 10% of the peak power. It is thus difficult to achieve high efficiency when PAPR is high. A high PAPR implies that the RFPA/T_X rarely operates at peak output power, where the efficiency is highest.

$$PAPR = 10 \cdot log_{10}\left(\frac{P_{Peak}}{P_{Avg}}\right) \tag{4.1}$$

This chapter provides an overview of known techniques to decrease power consumption in transmitters and power amplifiers operating with high PAPR

signals. Note that there is still room for improved techniques, a solution where output power and consumed power scale linearly is still missing.

4.1 Introduction to Power Amplifier Classes

This is to provide a short introduction to power amplifier classes, since more extensive treatments have been presented numerous times already, for instance in [53]. A general PA circuit is shown in Fig. 4.1, consisting of a transistor (NMOS-device [M_1]), a current-source (RF-Choke [L]), and a Matching Network (MN) containing the load [R_L]. The transistor operation is either in the current-saturation region (as a transconductance) or in the triode region (as a switch/resistor). When the transistor is operated in the current-saturation region the amplifier is considered linear, and when the transistor is operated in the triode region, the amplifier is in a switched/non-linear-mode. These definitions are based on the amplifiers ability to respond to input signal amplitude changes, and also in so-called linear operation substantial distortion can be produced.

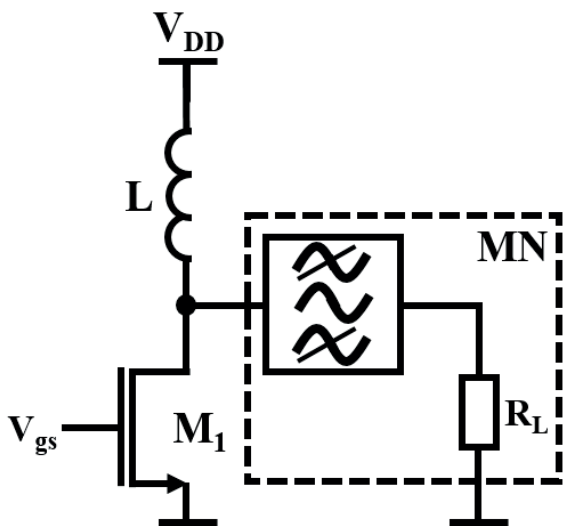

Fig. 4.1. General CMOS radio frequency power amplifier.

Among the linear classes of amplification, the most linear is the class A. In class A the transistor is biased so that the drain-source current never turns off, even at the negative peaks of the signal. This means that the gate-voltage exceeds the threshold voltage at all times. This is the most common way to bias transistors in amplifiers for lower power levels, since the efficiency is then less important, and the distortion will be low. In a PA, however, the always on transistor current may result in too low efficiency. To improve the efficiency, the class B amplifier is biased just at the point

where the transistor conducts (V_{GS}=V_{th}). The transistor then conducts for half the signal period instead of the full period, resulting in considerable efficiency improvements. There will, however, be cross-over distortion when the transistor crosses between conduction and non-conduction. There will also be a gain drop at low amplitudes, as the effective transconductance of the transistor is then lower. Class AB attempts to combine the advantages of both class A and B, by using a low bias current. The efficiency is then better than for class A, as the bias current is lower, and the gain variation at low amplitudes is reduced compared to class B, as the presence of a bias current sets the transconductance for small signals.

Higher efficiency still can be achieved in class C, where the transistor is biased to only conduct when a large input signal is present and for less than half the signal period. Since the transistor operates in current-saturation this is categorized as a linear class of operation, but class C is rather non-linear, not responding to low input signals. An important feature is that class C operation can generate a gain expansion when the input signal power is high. This is the opposite behavior to class A, B, and AB, which all compress at high input power.

For the switched classes, the transistor operates as the name of the class indicates, as a switch, which can be obtained by compressing/saturating the transistor with a large input signal. The bias condition is not that important in the switched-mode classes, but it controls the transition point, i.e. at which input signal amplitude the transistor changes state from ON to OFF or from OFF to ON. The class D amplifier can implemented with a single transistor (Fig. 4.1). It should also be noted that a complementary/inverter stage [54] can be used in the class D amplifier, and that the MN should then have the same properties as in the single-transistor case. When the transistor operates as a switch, a square wave voltage is produced at the drain terminal. To support that, the MN must supply a high-impedance for the odd harmonics at the drain of the transistor. The MN must also block these harmonics from reaching the load. The square wave at the drain terminal enables high efficiency, since it allows the transistor to have a low drain-source voltage when on, resulting in low power dissipation. The shaping of the drain voltage through the drain impedance is what sets the different switched classes apart.

The class E amplifier was first shown in [55]. This class takes the transistor drain voltage shaping to another level, and a concept called Zero-Voltage-Switching (ZVS) is used. The ZVS means that at the point in time where the transistor turns ON, the voltage across the transistor is zero (V_{ds}=0 V), and also the time derivative is zero ($V'_{ds}(t)$=0 V/s) [56].

$$P_{dyn} = C_{Par} f V_{ds}^2 / 2 \tag{4.2}$$

This type of pulse-shaping does not only minimize the voltage-current overlap, it also minimizes the dynamic losses (4.2) [57], since no electric charges are present at the drain of the transistor during state transition from OFF to ON. The major issue with the class E amplifier, however, is high voltage stress on the transistor due to large peak drain voltages, which can be as high as a theoretical $3.56V_{DD}$ [56], when perfect ZVS is obtained.

Last but not least of the switched-mode amplifiers, is the class F (voltage mode in [57] and current mode in [58]), which uses more aggressive/sharp shaping of the impedance at the drain of the transistor than the class D. The class F uses separate resonators to tune the fundamental and the odd harmonic frequencies individually, and to achieve low loss the used passive components require high Q-values.

As a final remark, the switched-mode power amplifiers can still be used in their natural mode in some rare cases, but today when modulation with non-constant envelope and a wide power control range is the norm, their dynamic-range is just too low. This can be solved by designing an RFPA, which has the ability to operate as a linear class amplifier at low output power levels, and as a switch class amplifier at high output powers. The concept is called mixed-class operation, or as in [59], a class AB/D amplifier, and the main idea is to use the amplifier in linear class when maximum Power Added Efficiency (*PAE*) (4.3) is not a must, and in switched class when it is. The switched-mode of operation is obtained by compressing a linear class amplifier, and the class of operation in the switched-mode is set by the MN.

$$PAE = \frac{P_{out} - P_{in}}{P_{DC}} \tag{4.3}$$

4.2 Efficiency Enhancement Circuits and Techniques

This chapter introduces different efficiency enhancement techniques for transmitters operating with non-constant envelope signals, starting with efficiency enhancement techniques that are more linear by nature, followed by more non-linear schemes, and also some key circuits and techniques that are needed to support the efficiency enhanced transmitters.

4.2.1 The Doherty Amplifier

The Doherty amplifier from 1936 [60], named after its inventor William H. Doherty, is a well-used and celebrated PA solution [61], which is still used in wireless communication today [62]. This amplifier system, shown in Fig. 4.2, combines a "Main" linear class (A/B/AB) amplifier with an "Auxiliary" (Aux) amplifier operating in class C. The Main amplifies smaller signals, and when it starts to enter compression, the Aux starts to provide output signal (gain expansion, class C) to the

system to maintain output power and linearity. The system operates with a 90° transmission line at the output, and a 90° phase shift at the input, both indicated as 90° blocks in Fig. 4.2. Acting as an impedance inverter, the quarter-wave transmission line at the output lowers the output impedance of the Main when the Aux is conducting, counteracting compression of the Main. The input phase-shifter is often implemented as a quarter-wave-hybrid [44], and the output as a quarter-wave transmission line (λ/4) [44] or a transmission line emulated by lumped components [44]. This limits the operation to a single frequency band, but with a higher order network implemented with transmission lines more frequency bands can be covered [63].

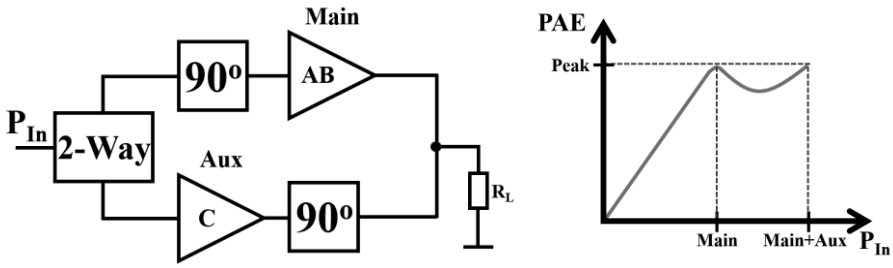

Fig. 4.2. The classic Doherty RF power amplifier and its PAE characteristic.

Further, the phase-shifters and transmission lines have a physical size that depends on the operation frequency, and for cellular frequencies the phase-shifters become too large for efficient implementation on chip in a modern CMOS process. However, fully integrated Doherty amplifiers have been shown in CMOS [64] and these solutions use transformer structures to implement the Doherty functionality (Main load impedance modulation by Aux signal), but the losses of the transformers are problematic. Finally, the characteristic shape of the efficiency (PAE) curve is also shown in Fig. 4.2. The extra peak in the efficiency curve is the reason to use a Doherty amplifier with high PAPR signals. This peak can be designed to coincide with the average output power of the modulated signal, greatly improving the overall efficiency of the system.

4.2.2 The Envelope Tracking System

The Envelope Tracking (ET) system [65-67] is shown in Fig. 4.3. In an ET system, a Linear class PA (LPA) is used, and its supply voltage (V_{DD}) is generated by a Supply Modulator (SM), which is a fast DC-DC converter. The main idea of ET is that since the efficiency of the LPA is highest close to the compression point, if we vary its V_{DD} with the signal envelope so that it operates closer to the compression point for different signal levels, the efficiency will be increased. At low signal levels we then use a low supply voltage, reducing the compression point and the power dissipation, whereas for a higher signal level we use a higher supply voltage yielding

a higher compression point. The idea can also be seen as for each P_{Out} using the V_{DD} curve that yields the highest PAE in Fig. 4.3.

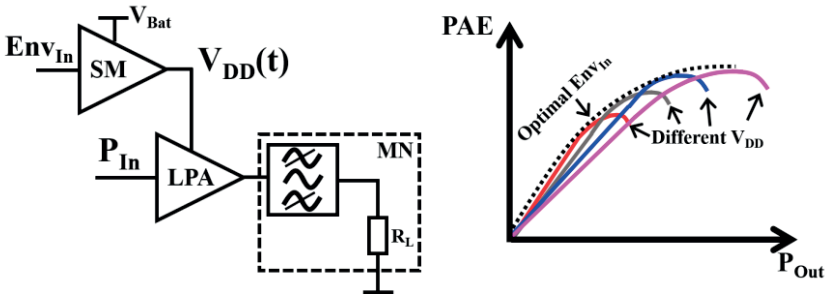

Fig. 4.3. The envelope tracking for efficient RF power amplification.

To characterize the amplifier for ET, first a constant value of V_{DD} is applied, and the compression point (1dB or other) is then found for this specific V_{DD}-value. This is then repeated for a series of V_{DD} values, and the compression point values are stored. Inverting this series of measurements gives a function for controlling the V_{DD}, which maintains a low gain variation for all powers levels. At high back-off levels the linear amplifier operates with a low constant supply voltage without any compression. ET reduces the power dissipation by maintaining a constant and minimal headroom over the larger amplified RF signal, compared to a varying headroom in a linear amplifier supplied with a constant high V_{DD}. This comparison is illustrated in Fig. 4.4. However, the ET system also requires an SM, which has its own power dissipation affecting the overall system power consumption and efficiency (4.4). All the supply-power fed to the linear amplifier has to be delivered by the SM, connected to the battery (V_{Bat}), and in all practical scenarios it is the consumption from the battery that counts.

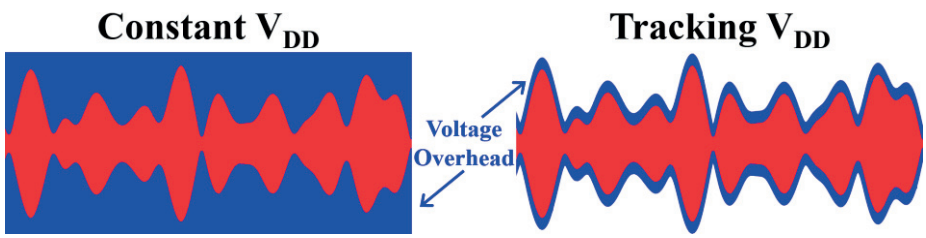

Fig. 4.4. The theoretical power dissipation of constant and dynamic supplies.

$$\eta_{tot} = \frac{P_{out}}{P_{Bat}} = \eta_{PA} \cdot \eta_{SM} \tag{4.4}$$

Finally, the supply modulation also has some negative effects on the linearity of the transmitter. These are mostly caused by the voltage dependent parasitic capacitances of the transistor. By changing the supply voltage at the drain of the transistor, a change in parasitic capacitance occurs, and also in the power dissipation of the transistor. The capacitance variation results in AM-PM non-linearity, and the variable power dissipation causes a varying device temperature. As it takes time for the transistor to heat up and cool down there is a memory-effect associated with the temperature. These effects have been studied, analyzed and quantified in [68]. To note, the memory-effects due to varying temperature causes one of the most difficult distortions to mitigate in the ET system.

4.2.3 The Envelope Elimination and Restoration System

The Envelope Elimination and Restoration (EER) system [69], invented by Kahn in 1952 [70], is often also called the Kahn technique or the pure polar system. The system is shown in Fig. 4.5, and as can be seen it uses an SM like ET. It also uses a switched-mode PA (SMPA), and an optional envelope DETector (DET) and a LIMiter (LIM), which are used only if an RF input signal is used with the system. Otherwise, polar signals can be generated from the Cartesian baseband signals (I and Q) inside the RFIC, the amplitude/envelope-signal to be used by the SM is then readily generated as the length of the IQ vector (4.5). The input signal to the SMPA is an RF signal with the phase according to the IQ vector (4.6), and it can be generated by IQ up-conversion in a quadrature modulator (mixer), or by a phase locked loop. With an RF input signal, however, the system separates the envelope

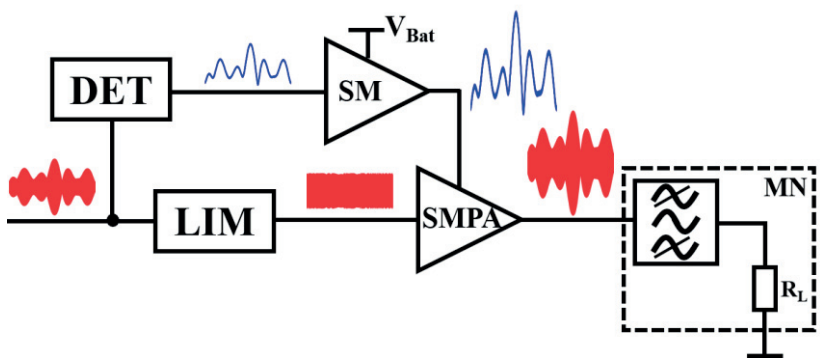

Fig. 4.5. The envelope elimination and restoration system.

$$A(t) = Env(t) = \sqrt{I(t)^2 + Q(t)^2} \tag{4.5}$$

$$\phi(t) = \begin{cases} arctan\left(\frac{Q(t)}{I(t)}\right), if\ I > 0 \\ arctan\left(\frac{Q(t)}{I(t)}\right) + \pi, if\ I < 0 \end{cases} \tag{4.6}$$

signal from the input signal with the DET. The envelope is then removed/eliminated from the input signal in the LIM, and left is the RF-signal without any envelope information, which is the RF-carrier with phase information. After this separation of amplitude and phase modulation, the amplitude/envelope-signal is amplified by the SM and fed to the SMPA supply, and the RF-carrier with phase is fed to the input of the SMPA. Inside the SMPA, the supply voltage is multiplied by the RF-carrier with phase, and at the output of the SMPA we obtain an amplified and restored version of the input signal.

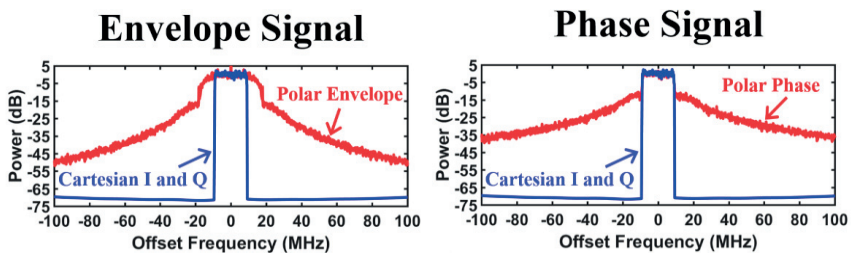

Fig. 4.6. Spectral comparison of polar and Cartesian coordinates.

EER can be considered the "Holy-Grail" of PA systems, since the efficiency can be very high and the power consumption scales very well the RF output power, but the system did not really come of age until the late 1990s when [71] was published. However, the EER system has bottlenecks that are fundamentally hard to solve when the modulation bandwidth of the signal is wide. The polar baseband signals (amplitude and phase) suffer from BandWidth Expansion (BWE) due to the non-linearity of the coordinate transformation (4.5-4.6) [72], and the BW of the two paths (amplitude and phase) must thus be wide to avoid distortion, and the synchronization/time-alignment of the two paths must also be very precise [72]. The transistor in the SMPA also has a very non-linear behavior, in terms of output amplitude and phase versus supply voltage, due to the parasitic gate-drain capacitor (C_{gd}) inside the transistor [73], and it also suffers from memory-effects due to temperature variations. As the main problem with EER is the BWE, a spectral comparison between polar and Cartesian baseband coordinates, with a 20MHz UL LTE signal, is shown in Fig. 4.6. The most severe bandwidth expansion occurs when the Cartesian baseband signals crosses or passes close to the origin (0, 0) in the IQ-plane. These effects can be minimized by avoiding the origin, which is called vector-hole punch/modulation [74]. However, if the avoidance of the origin is too aggressive, the errors introduced can cause large amounts of distortion (spectral-regrowth) and the ACLR requirements might be in jeopardy. The requirements on the synchronization/time-alignment and the BW of the two paths are mostly set by the BWE, and consequently, a reduction in BWE eases these requirements, improving the entire system.

A good approach to improve the EER system is to reduce the non-linear effects of the SMPA by using a cascode amplifier to isolate the input from the output [50], which was shown in [75] with high performance. Finally, some of the problems of EER are also present in the ET system, but since the input signal of the SMPA (EER) has no envelope/amplitude-information, and the SM thus directly sets the output amplitude, all errors of the amplitude path are directly fed to the output and there is no mitigation of distortion due to imperfect combination of amplitude and phase signals. At low signal levels the SMPA input signal is also much larger than the output, making it sensitive to the feed-forward through C_{gd}. By keeping the envelope/amplitude-information in the PA input signal in ET, most BWE is avoided, and some degree of freedom still exists in supply voltage overhead, which can be used to tone down or mitigate the non-linear effects inherent in supply-modulated systems.

4.2.4 The Hybrid and Mixed-Mode Systems with Supply Modulation

The Hybrid-EER (H-EER) system [76], is a hybrid of ET and EER, and it can be regarded as an EER system without the limiter (LIM). In this system the switched-mode PA input signal thus has a time varying envelope. This eases the requirements on the BW and synchronization/time-alignment of the signal paths, the high efficiency of EER is maintained and the system is almost as linear as the ET system. The dynamic range is still inferior to ET, however, since the distortion in the transistor at low output powers is severe, caused by the low supply voltage pushing the transistor into the non-linear triode region.

Another system, called Mixed-Mode H-EER [**II**], is a hybrid between ET and H-EER. It is shown in Fig. 4.7, and the key building blocks are a Mixed-Mode/class PA (MMPA) and an SM. The input signal to the MMPA is the original RF-carrier with all modulation information, as in ET and H-EER, and the amplitude/envelope-signal is similar to the one used in ET. However, in ET the V_{DD}-function is selected to only give a modest amount of compression (CP≈1-3dB), but in Mixed-Mode H-EER the amount of compression is larger and it increases with input power, according to trade-offs between gain, linearity, and system efficiency. By increasing the amount of compression, the gain is decreased and linearity becomes worse, but the system efficiency is improved. To further improve the system performance, the MMPA should also be gate-biased dynamically [**II, V**], which increases the control of the current-flow through the transistor/transistors and adds more degrees of freedom to the system.

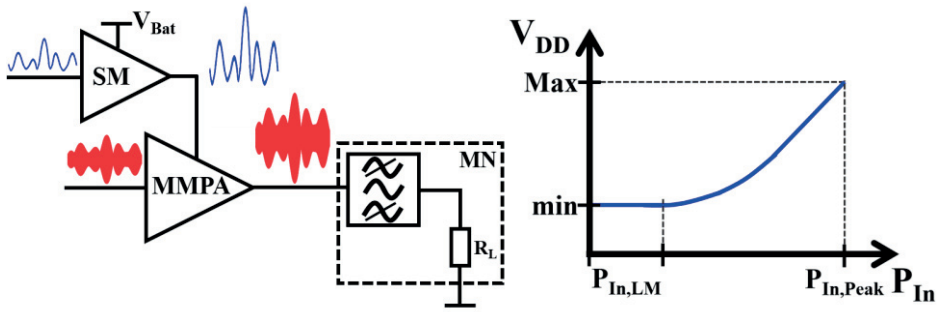

Fig. 4.7. Mixed-mode hybrid envelope elimination and restoration system.

The MMPA operates in linear class with a constant low V_{DD} when the output power is low, and as the power increases it starts to compress, and to add expansion/headroom the V_{DD} is then increased. But as the output power further increases, the V_{DD} signal is increased less than in a corresponding ET system, pushing the MMPA deeper into high-compression/switched-mode operation. When the MMPA is in full saturation, it operates as a purely switched class amplifier, and now the V_{DD}-signal fully controls the amplitude/envelope, like in EER, due to the large compression in the MMPA. The PA operates in linear class or in switched class depending on signal level, hence the name mixed-mode power amplifier (MMPA).

4.2.5 The Supply Modulator

In all the supply modulated systems, a high performance SM is an absolute necessity, and the key features are high efficiency, sufficient output power (both voltage and current), low amounts of spurious-tones/distortion/switching-noise, and sufficient speed (bandwidth and slew-rate). The mainly used concept is the hybrid-switching or master-slave modulator [77], a DC-DC-converter that combines a linear class AB amplifier with a switching class D amplifier, and the system is shown in Fig. 4.8. By combining a linear class AB voltage-mode amplifier with the class D stage, its low output impedance suppresses the switching-noise/distortion from the class D stage at the output of the SM. This not only eases the off-chip filter requirements (L in Fig. 4.8) of the class D amplifier, it also lowers the required switching-frequency/BW of the class D stage. The higher frequency signals are then delivered mainly by the class AB stage, which is possible since the higher frequency signals have smaller signal amplitudes. Different methods have been shown for controlling the class D amplifier, the most common being the self-oscillating system [77], but also clock driven approaches exist [78]. The self-oscillating system uses only a current-sensor (C-Sens) at the output of the class AB stage. When the class AB stage cannot deliver all the needed current to the load (R_L), the class D amplifier is activated to aid delivering current, and as this occurs the current from the class

AB can be reduced, this will be sensed by C-Sens and the class D stage will then be stopped from supplying current. Therefore, the current from the class D stage will turn on and off in a self-oscillation, if hysteresis is present in the current-sensor (C-Sens) or in the Control.

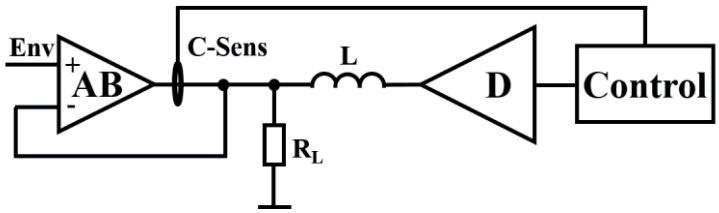

Fig. 4.8. A hybrid switching supply modulator.

For the clock driven system, the current-sensor is still used, but without any hysteresis, and in this case for forming a stable control loop. A Pulse Width Modulation (PWM) carrier is supplied by a separate clock circuit involving a triangle wave generator and a comparator. The switching noise is still suppressed by the class AB amplifier, the power of the switching frequency fundamental and its harmonics are more defined and less spread in frequency, but hard to suppress by filtering. In other more traditional clock driven DC-DC converters [79], a triangle wave generator with randomized switching frequency has been shown in [80] with very good results. This randomized clock generator spreads the power of the switching noise over a wider spectrum (thermal noise like), decreasing the amount of power at each instance of frequency, lowering the suppression requirements. This makes the hybrid switching SM with a randomized triangle wave generator attractive in the important trade-off between switching noise and power efficiency.

4.2.6 Linear Amplification using Nonlinear Components

The LInear amplification using Nonlinear Components (LINC) [81], shown in Fig. 4.9, is a system that can use switched-mode power amplifiers to amplify a signal with varying envelope. The input signal is first divided into two constant envelope signals using the Signal Component Separator (SCS). The amplitude information is within the phase difference between the two signals, and the constant envelope signals are then amplified individually by SMPAs. The output signals from the two SMPAs are then combined (2-Way) at the output of the system. The combination results in an output signal containing the amplitude information. The combination can be seen as an addition of two vectors of equal length, where the resulting amplitude depends on the phase difference (see Fig. 4.9). The LINC system suffers from similar issues as the EER system, i.e. the individual signals have large amounts of BWE due to non-linear coordinate transformations.

The BW requirement of each individual signal path is high, and the output combiner must have low amplitude and phase imbalance, low loss, and preferably also high isolation. There is, however, a trade-off between combiner loss and isolation, and by sacrificing isolation less loss can be obtained. However, with low isolation the SMPAs will experience an amplitude dependent load impedance, and unless the SMPAs have a very low output impedance this will cause non-linear distortion. If the SMPAs are implemented as a complementary/inverter based class D amplifier as was done in [54], they can have a low output impedance, so that a low isolation combiner can be used with limited effect on the overall system performance.

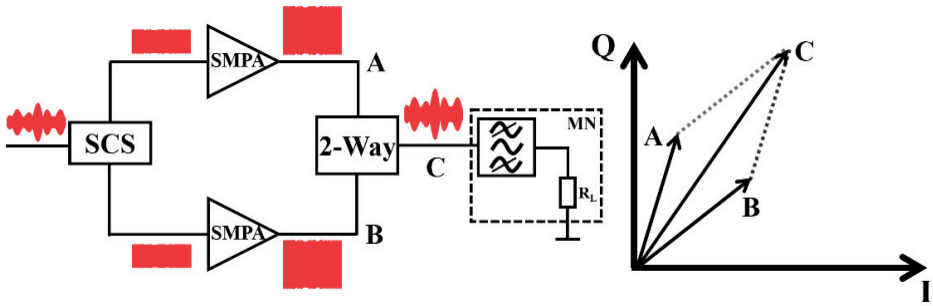

Fig. 4.9. The concept of linear amplification using nonlinear components (LINC).

4.2.7 Digital Pre-Distortion

In most RF transmitters designed for high efficiency, memory-effects and/or IMD are a problem. A common technique to mitigate that today is to use memory or memory-less pre-distortion at baseband in the digital domain, which is called Digital Pre-Distortion (DPD) [82] or baseband pre-distortion. Fig. 4.10 shows the basic principles, and as can be seen the system can be either open-loop or closed-loop. The main idea is that a signal producing an output that is the opposite of the expected error of the DUT/PA is added to the input signal. The result will then be a close to error-free output signal. An amplifier with negative feedback has a similar mechanism, the non-idealities detected by the feedback at the output are added to the input signal with opposite phase, and the non-idealities cancel inside the amplifier, which gives an output signal that is close to a perfectly amplified replica of the input signal (if the loop-gain is high). However, to use negative feedback in RFPA systems is hard, especially if off-chip matching is used. The main concern is loop stability, and sufficient gain in the loop is possible only if the loop is very close/local to the RFPA [83]. The main advantages of DPD is that stability issues are non-existent, even in a closed-loop setup all the processing can be performed at baseband, in the digital domain, and can be as simple or complex as necessary depending on the required level of distortion and the properties of the RFPA system.

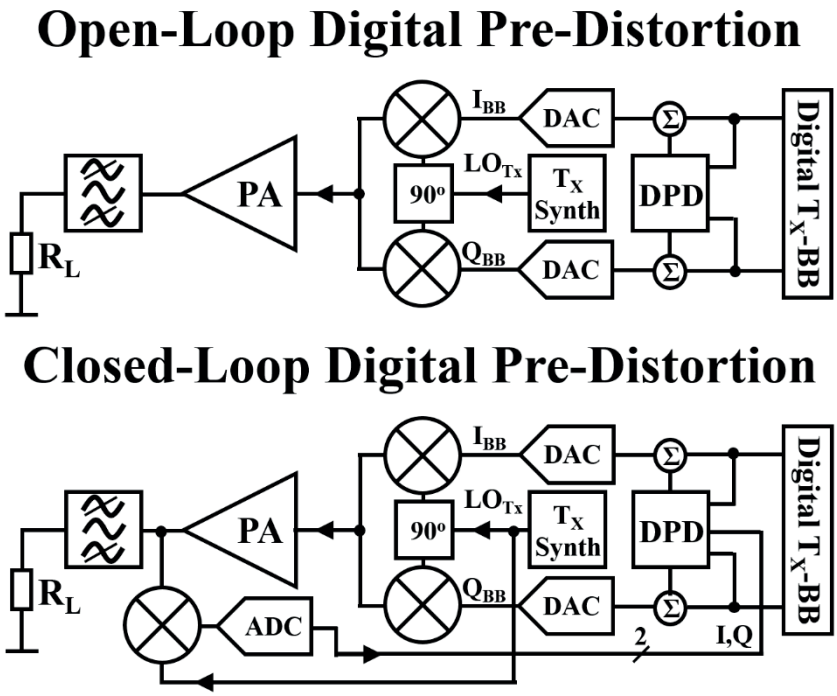

Fig. 4.10. Digital pre-distortion, open-loop and closed-loop architectures.

For systems that utilize switched-mode RFPAs, the input to output transfer of the amplitude information can be very low, and amplitude error corrections applied at the input then has very low or no effect. If the system is supply modulated, the amplitude errors can instead be corrected by applying the amplitude error corrections (e_{LF}+e_{HF}) to the envelope signal. Errors far out from the carrier frequency (e_{HF}) can then be corrected if the SM has sufficient BW. The DPD system used in this thesis work ([**III**] and [**V**]) employs this technique, with the system block diagram shown in Fig. 4.11. The system is used with a mixed-mode H-EER transmitter (MMPA), so at low output powers the PA is linear, but at high output power levels it is in deep compression. To generate the DPD signals, the output signal is measured and the transmitter distortion is extracted in time-domain using MatLab over several modulation symbols. The same modulated signal can then be applied to the circuit again, this time also applying the extracted DPD signals. By using this simplified method, the complexity is low and memory-effects are naturally included. This method has been sufficient for circuit verification in the research lab, but it is not practical in its current form. In a practical system the DPD needs to create a model of the transmitter, so it can generate appropriate correction signals for all input signals. The Matlab algorithm for generating the DPD signals for circuit verification first extracts corrections for the errors close to and within the

main carrier. The corrections are applied to both the SMPA input signal and to the envelope signal (e_{LF}, in Fig. 4.11). A target is to not introduce to much new AM-AM and AM-PM distortion, due to bias changes of the SMPA. The process is then iterated until sufficient EVM and spectral purity is obtained. The far-out spectral distortion is then mitigated using additional corrections to the envelope signal, e_{HF} in Fig. 4.11. This cancels the error at the output by adding amplitude modulation [84], but it is only completely effective if the output spectrum is symmetric, i.e. both sidebands have approximately the same level.

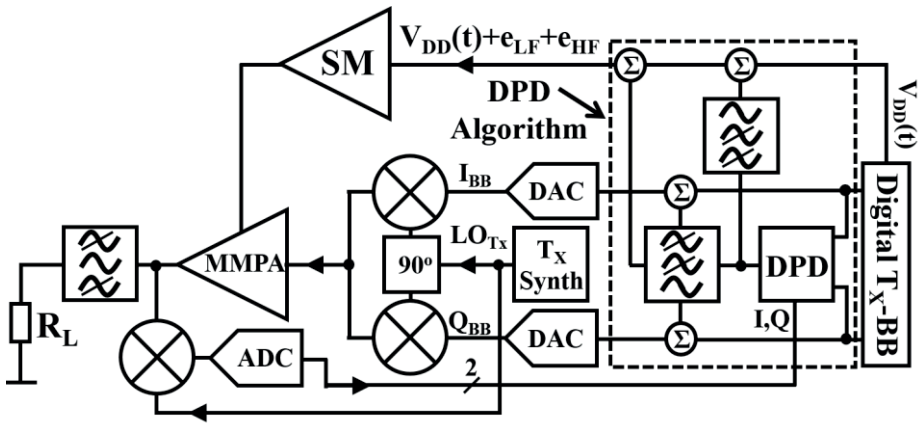

Fig. 4.11. Digital pre-distortion for mixed-mode hybrid-EER transmitters.

CHAPTER 5

5 Injection-Lock Theory and Techniques

5.1 Introduction

The phenomenon of injection locking has been observed by most people, but few know it. It can be seen when visiting a clock store, and watching pendulum-clocks that share a wall. As can be seen, all the pendulums are then synchronized in frequency, and in phase, so that they all reach the endpoints at the same time. This is what injection-lock is, a small part of the pendulum resonant movement of each clock will couple to the wall and travel through the wall to the other clocks, coupling to their pendulums. Since the pendulums are resonant at almost the same frequency, not much coupling is needed to make all pendulums lock together, to a frequency mainly set by the biggest pendulum (having the most energy). If all pendulums are started in an unsynchronized way, in phase and frequency, over time they will still lock together, to the same phase and frequency. This is called injection locking, or entrained as B. van der Pol called it [85]. If the coupling is not strong enough to obtain phase and frequency lock, but it still affects the behavior, it is called injection pulling. Injection pulling is typically undesired. The injection pulling makes homodyne radio transceivers hard to design, as the strong disturbance from the PA pulls the Voltage Controlled Oscillators (VCOs) in the frequency synthesizers. This results in a degraded spectral purity of the LO, which impairs T_X spectral emissions and R_X selectivity.

In electronics injection pulling and locking have been studied since the 1920's. The complexity of the phenomenon has led to theoretical studies on fairly simple circuits, and more complex circuits have only been investigated by simulations and/or measurements. The first major theoretical breakthrough was in 1946 when R. Adler derived the famous equations in [86]. However, nearly 60 years later, in 2003/2004, Prof. Razavi derived more accurate equations for large-signal operation in [87], which is typically the operation mode of injection-locked circuits. Finally, a personal favorite is the paper written by K. Kurokawa in 1973 [88], and this paper deals with a broader range of injection-lock phenomena and circuits, ranging from the classic solid-state oscillators to negative-resistance amplifiers. There are many others that have contributed to the theoretical development, but the aforementioned publications are considered the most important.

Injection-lock can be used in different oscillator based circuits, for example in Quadrature VCOs (QVCO) [89, 90], relaxation oscillators [91], frequency-dividers [92], and with oscillator arrays for beamforming [93]. Injection-lock can also be used for power amplification, and a high power gain and power added efficiency (PAE) can then be obtained, since a high gain PA can be made from relatively large transistors, which in a regular PA would cause issues with stability and drive requirements. The Injection-Locked PA (ILPA) is intentionally unstable, and by using self-oscillating amplification stages the power gain can be increased by approximately 10dB per stage compared to a stable PA. Due to the higher gain, the ILPA has less drive signal requirements, and the input signal drives an injection transistor which is often 2 to 5 times smaller than the actual power transistor. The gate capacitance is thus 2 to 5 times smaller, reducing one of the major bottlenecks in multistage stable PAs, the interstage matching between the final amplification stages, where gate impedance can be very low. The CMOS-ILPA has been used with GSM signals in [94], with supply modulation in [**II**], [**III**] and [**V**], and for mm-waves in [VIII].

The major constraint in injection-locked circuits is the complexity, since they involve several non-linear components, and if a proper lock is not achieved spurious distortion will be severe. For larger circuits the analysis must be limited due to complexity. On the other hand, there are large benefits in reduced power consumption and input drive requirements of the injection-locked circuits. The pre-scaler in mm-wave PLL-systems is thus often an injection-locked divider, because the input of the pre-scaler operates at the highest frequency of the system (i.e. the frequency of the oscillator). Frequency multipliers at mm-wave are also typically injection-locked. The injection locking saves power and may also enable operation at higher frequencies than otherwise reachable.

5.2 Dynamic Biasing and Supply Modulation of the ILPA

The bias voltages and currents are of fundamental importance to both linearity and power dissipation of the PA. By using a dynamic bias-scheme, where all internal circuit currents and voltages are controlled as a function of the signal power, operation can be optimized for all power levels [95]. For the ILPA [**II**, **III**] and [**V**], the dynamic biasing controls the output power, and also the operating mode to either linear or injection-locked mode. The basic idea of this PA concept is to use linear operation with constant low supply and bias at low output power levels, and switched-mode injection-locked operation with supply modulation and dynamic bias at higher output power levels, where supply and bias voltages/currents are all functions of the output power.

In Fig. 5.1 the ILPA with dynamic biasing for the different operation modes is shown. Starting with the high power mode, the whole circuit is used during this operation, and the supply voltage is set as a function of the signal power level, and so is the gate-bias voltage $V_{G,O}$ of the cascode transistor (M_O). The cascode gate-voltage sets what amount of current can flow to the load through the channel (drain to source) of M_O, through M_S and I_{inj}. This also sets the local loop-gain and negative output conductance of M_S. The dynamic bias voltages ($V_{G,O}$ and V_{DD}) are both functions of the targeted output voltage level. By tailoring these functions the output signal level will be close to the targeted, and a high linearity can be obtained.

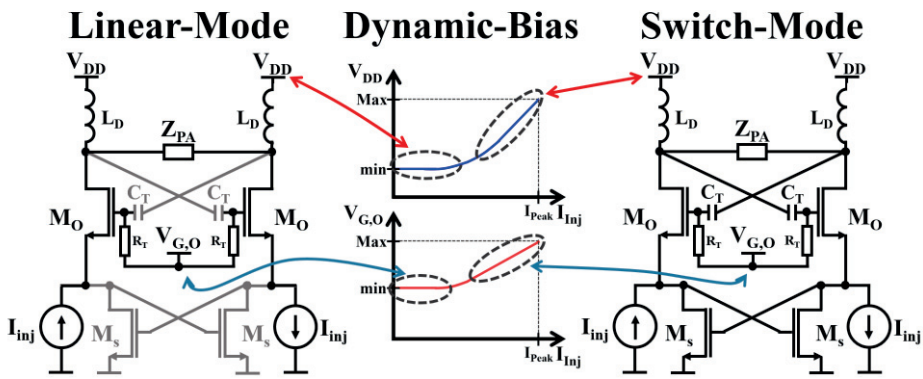

Fig. 5.1. The dynamic biasing of the ILPA in different modes of operation.

The linear-mode of operation is also shown in Fig. 5.1. Then the gate-bias ($V_{G,O}$) of the cascode is set to a constant voltage, and that voltage ensures that M_O operates in current-saturation and that the positive loop-gain of the ILPA is low (less than one), i.e. the circuit is stable. The voltage is low enough to ensure that M_S is in the OFF-state, i.e. the gate-source voltage of M_S is below the threshold-voltage (V_{th}) for the whole RF-signal. When M_S is off the circuit is stable, and the cross-coupling of M_O becomes ineffective, but the parasitic capacitances of M_O are still there, which ensures correct output matching. Finally, in the linear-mode, the transistors are very large in relation to the output power and the parasitic elements are also consequently large, but the low voltage swings in each node minimize the losses due to parasitic capacitance.

5.3 *Theory for the Injection-Locked PA with Supply Modulation*

An analysis of the circuits in [**II**, **III**] and [**V**] will now be performed based on the equations used in [87], for a more detailed explanation of the fundamental theory see Appendix A. The analyzed circuit schematic is shown in Fig. 5.2. The output matching networks are replaced by a general PA load impedance $Z_{PA}(\omega)$, which loads the drain of the output transistor M_O. To model the differential cross-coupling with capacitive voltage-division of M_O, a gain of $-\alpha$ is used in the single-ended equivalent circuit in Fig. 5.2, where α is the voltage-division ratio between the gate capacitance and the added capacitance (C_T in Fig. 5.1). The source of M_O is connected to the drain of the cross-coupled transistor M_S (with gain of -1 to model the differential circuit cross-coupling to its gate), and to the injection current-source with the injection frequency ω_{inj}. Finally, the total parasitic capacitance in the injection node is modelled as C_{Par}, since this node does not contain a physical resonator and dynamic switching-losses will then be present.

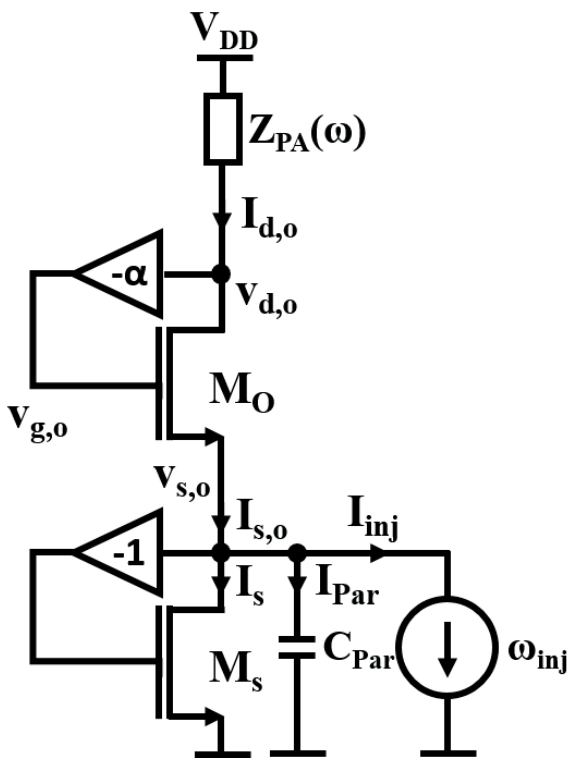

Fig. 5.2. Injection-locked power amplifier, analyzed schematic.

a) b)

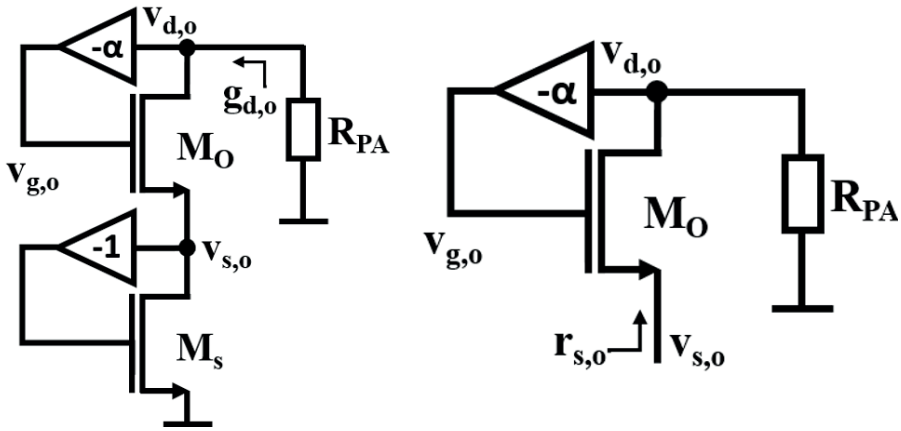

Fig. 5.3. Circuits for calculation of: (a) Loop-gain.
(b) Cascode source resistance.

The basic behavior of the circuit in Fig. 5.2 was analyzed in [**III**], without including effects of frequency dependence, assuming $Z_{PA}(\omega)$ to be in resonance and presenting a resistive load R_{PA}. To derive the startup loop-gain (5.1) the small-signal model was then used on Fig. 5.3a. The input resistance r_{so} of the output/cascode transistor when in active region (5.2) and in triode (5.3) [ON-State] and (5.4) [OFF-State] was derived from Fig. 5.3b, for details see [**III**].

$$A_V = g_{d,o}R_{PA} = \frac{\left(\alpha g_{m,o} - \frac{1}{r_{ds,o}}\right)R_{PA}}{1-\left(g_{m,o} - \frac{1}{r_{ds,o}}\right)r_{n,s}} \approx \frac{\alpha g_{m,o}R_{PA}}{1-g_{m,o}r_{n,s}} \tag{5.1}$$

$$r_{s,o} = \frac{\left(1-g_{m,o}\alpha R_{PA}+\frac{R_{PA}}{r_{ds,o}}\right)}{\left(g_{m,o}+\frac{1}{r_{ds,o}}\right)} \approx \frac{1}{g_{m,o}} - \alpha R_{PA} \tag{5.2}$$

$$R_{s,o} = R_{on,o} + R_{PA}\,, \qquad \left(V_{GS,O} - V_{th}\right) \geq V_{DS,O}, V_{GS,O} > V_{th} \tag{5.3}$$

$$R_{s,o} = R_{off,o} = \infty\,, \qquad V_{GS,O} < V_{th} \tag{5.4}$$

$$Z_{s,o}(\omega) \approx \left(\frac{1}{g_{m,o}} - \alpha Z_{PA}(\omega)\right) // \frac{1}{j\omega C_{Par}} \tag{5.5}$$

For the source resistance (5.2) of M_O, when frequency dependencies are considered, all parasitic capacitances of the devices can either be included in the output impedance $Z_{PA}(\omega)$ or in the cascode source node capacitance, as C_{Par} in Fig. 5.2. In order to derive the source impedance from (5.2), R_{PA} is replaced by the impedance $Z_{PA}(\omega)$, and the cascode source node capacitance (C_{Par}) is also included. The positive feedback ($-\alpha$) connected to the gate of M_O generates a voltage at the gate ($V_{g,o}$, in Fig. 5.2) with an opposite phase to the output voltage. The frequency dependent cascode source impedance is shown in (5.5), the term caused by the feedback includes the frequency dependent load impedance ($Z_{PA}(\omega)$), and an additional contribution from the impedance of the source node capacitance (C_{Par}) is also added.

From Fig. 5.2 important relations between circuit currents and voltages can readily be derived (5.6)-(5.10).

$$V_{d,o} = Z_{PA}(\omega) \cdot I_{d,o} \tag{5.6}$$

$$V_{s,o} = Z_{s,o}(\omega) \cdot I_{s,o} \tag{5.7}$$

$$V_{g,o} = -\alpha \cdot V_{d,o} = -\alpha \cdot Z_{PA}(\omega) \cdot I_{d,o} \tag{5.8}$$

$$I_{d,o} = I_{s,o} \tag{5.9}$$

$$I_{s,o} = I_s + I_{inj} \tag{5.10}$$

To analyze the functionality of the injection node, a deeper look at the theory in Appendix A is needed. The tank current I_T in Appendix A is determined by the current to voltage transfer (impedance) of the LC-tank, and I_T is a linear function of the LC-tank voltage. Now by observing the ILPA circuit in Fig. 5.2, the tank current ($I_{d,o}$) and voltage ($v_{d,o}$) have the same transfer as the basic case in [87], but the ILPA tank current ($I_{d,o}= I_{s,o}$) goes from source to drain of the cascode transistor M_O. Further, the drain-source current through M_O is controlled by the gate-source voltage $V_{gs,o}$ and the gate-source voltage of M_O must thus be in phase with the current ($I_{d,o}= I_{s,o}$) through $Z_{PA}(\omega)$. The gate-voltage ($V_{g,o}$) is the output resonator voltage ($V_{d,o}$) multiplied by the positive feedback factor α (5.8), so the phase shift of Z_{PA} (ϕ_0) is present at the gate terminal. Therefore, the source-voltage ($V_{s,o}$) of M_O must operate with a phase shift that keeps V_{gs} in-phase with the steady-state drain-source current ($I_{ds,o}= I_{d,o}= I_{s,o}$) and this is achieved by the positive feedback which sets the output to a steady-state oscillation as long as the loop-gain (5.1) and the local cascode loop-gain are high enough (local loop-gain of M_O: $A_{V,O}=\alpha G_{m,o} Z_{PA}(\omega)$). The source current to voltage transfer (impedance), includes the cascode source impedance (5.5) ($Z_{s,o}(\omega)$) and the drain-resistance of M_S ($-1/g_{m,s}$) (parasitic capacitance of M_S is absorbed in C_{Par}). In the steady-state, the phase shift is only dependent on $Z_{s,o}(\omega)$ (5.5), but it should be redefined for large-signal and in polar

form as (5.11) (effects of C_{Par} assumed small). The current-injection source (I_{inj}) is replaced by a transconductance with an input voltage V_{inj} and a voltage vector representation is shown in Fig. 5.4.

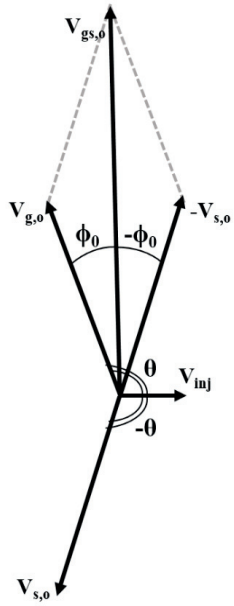

Fig. 5.4. Vector representation of the ILPA node voltages.

$$Z_{s,o}(\omega) \approx \frac{1}{G_{m,o}}\left(1 - \alpha Z_{PA}(\omega) G_{m,o}\right) = \frac{1}{G_{m,o}}\left(1 + \alpha |Z_{PA}(\omega)| G_{m,o} e^{j[\pi - \phi_0]}\right) \quad (5.11)$$

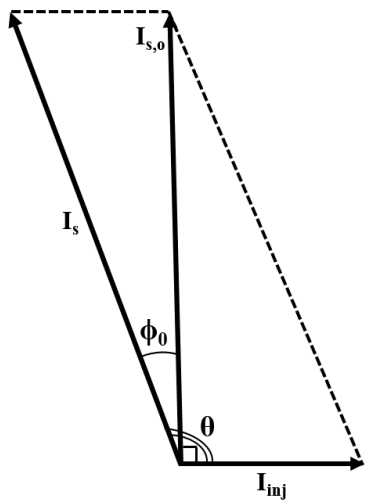

Fig. 5.5. Vector representation of the currents in the injection node.

The vector voltage representation in Fig. 5.4, shows the voltage vector representation for all node voltages of the circuit in Fig. 5.2. By assuming $\alpha|Z_{PA}(\omega)|G_{m,o} \gg 1$ in (5.11), the phase shift of the vector representing $-V_{s,o}$ becomes $-\phi_0$, due to the phas shift in $Z_{s,o}$ caused by the cross-coupling of M_O. This means, if the output load $Z_{PA}(\omega)$ is capacitive (with negative phase shift), then the cascode source impedance $Z_{s,o}(\omega)$ will be inductive (the same phase shift but with a positive sign), and vice versa. Similar cross-coupling techniques are often found in differential amplifier designs. For this ILPA, it sets the gate-source voltage to drain-source current relation. To illustrate the phase relation not only between the voltages, the vector representation of also the currents in the injection node is shown in Fig. 5.5.

The current representations in Fig. 5.5 indicate that the locking-range will be in the same range as the large-signal locking-range in [87], but this requires a few assumptions to be valid. The local loop-gain of M_O, $\alpha G_{m,o} Z_{PA}(\omega)$, must be much larger than unity, $G_{m,o}$ must be large enough to dominate over parasitic capacitances in the source node, and $V_{gs,o}$ must be in-phase with $I_{ds,o}$. The magnitude of the output load impedance reduces with distance from center frequency, and at the -3dB bandwidth the magnitude has dropped by a factor $\sqrt{2}$, which will directly lower both the loop-gain of the complete system and the local loop-gain of M_O. It can be observed that if the output load impedance has a high bandwidth, then a circuit with very high and wideband performance can be obtained, and the limiting factor then becomes the injection node capacitance C_{Par}. A wideband load typically also has low effective reactance, and hence low phase shift, and the phases of all current and voltages will then be well aligned for maximum power transfer and efficiency.

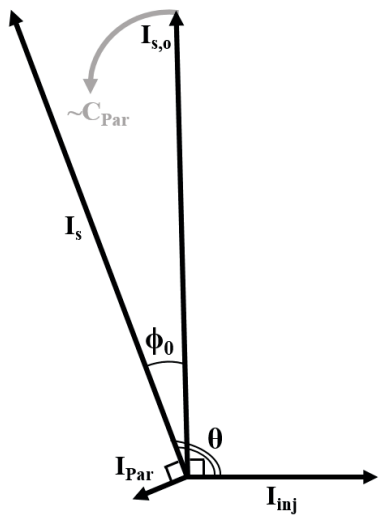

Fig. 5.6. Vector representation of the injection node currents with parasitic capacitance.

The capacitance C_{Par} of the injection node will cause dynamic power losses ($P_{dyn} = C_{Par} f V_{s,o}^2/2$)), but these will be limited since the voltage swing in the node is relatively low. The effects of C_{Par} on the injection-lock mechanism are on the other hand harder to predict. To better understand these a current vector representation at the edge of the locking-range, with the current through C_{Par} included, is shown in Fig. 5.6. As can be seen the current due to C_{Par} is almost in opposite phase to I_{inj}, which effectively reduces the locking-range. Since the current is proportional to the voltage swing of the injection node, maintaining a low cascode source impedance will minimize this effect.

5.4 Conclusion

The main conclusion is that RF PA circuits using injection-lock can have higher PAE and gain than stable solutions. The higher gain comes from the fact that the injection-locked circuits are to a large extent self-driven. By not having to design for stability and high gain, there are more degrees of freedom to design the circuitry for high efficiency. Injection-lock is particularly attractive in wide-band designs and in designs operating at frequencies approaching the limits of the semiconductor technology, making high gain difficult to achieve in a stable amplifier.

An ILPA with dynamic bias can achieve very low AM-AM and AM-PM distortion. This is based on the assumptions by Razavi in [87], that when the injected signal is within the locking-range the phase purity of the injected signal sets that of the complete system output. In [**III**], this was used to great advantage for both phase-linearity and phase-noise. The theory of the ILPA with dynamic biasing is based on that work.

To conclude the theory of the ILPA with dynamic biasing, through positive feedback the capacitive cross-coupled cascode devices operate with a gate-source voltage that is in phase with the LC-tank current. This also means that under normal operation, the large-signal locking-range is equivalent to the one derived in the past by [87], while injecting the current in the source node of the cascode. The source node injection technique was first used in frequency dividers, but a main difference is that the ILPA injects the current in differential-nodes and not a common-node (tail), and that the injection is at the fundamental frequency and not at a harmonic.

Finally, the main disadvantage of injection locking is not the technique itself, but rather the complexity of the mechanism. By also including dynamics into the circuits, complexity is increased to a large extent and can in some cases even explode. This is the main bottleneck of injection-locked dynamic PAs, and may be the main reason why most researchers just build classic linear or switched-mode circuits when dynamics is introduced due to signals with amplitude modulation.

CHAPTER 6

6 Summary with Conclusions and Future Work

To summarize, an impedance tuner [**I**], a novel dynamically biased ILPA [**II**], a novel ILPU [**III**] with a new approach to polar modulation [**IV**], and a new low cost wideband ILPA [**V**] were analyzed, designed, implemented, and verified as a part of this PhD-project. All circuits have been implemented in CMOS technology, the impedance tuner in CMOS-SOI and all the RF power electronics in standard bulk CMOS. The main advantages of the included circuits are their high efficiency and low cost in mass production, important in reducing BOM of the UE and to increase the time between battery charging.

For the impedance tuner, the measured performance met the requirements for use in a mobile phone, with low loss and low distortion, both in-band (intermodulation) and out of band (harmonic distortion). The circuit was verified by measurement at both the nominal 50Ω antenna impedance and non-50Ω impedances. Measurements such as IL, ML, RL, and OIP_3 were published in [**I**], and to the authors knowledge such measurements were never presented before in a publication. To demonstrate the capabilities in a real user environment, MIMO channel measurements were also performed and published in [VI], showing a significant improvement of MIMO capacity using the tuner from [**I**]. The main drawback, from one point of view, is that the impedance coverage was optimized for the specific antenna used in [VI], but this is natural for a custom made solution.

To further improve the impedance tuner, a newer technology node could be used, but the breakdown voltage of the transistor used was appropriate for this context. However, the impedance detector could be implemented as well, and the digital LUT could also be implemented on the same die, but implementation of diodes for the detector and digital-logic for the LUT are less advantageous in a 130nm CMOS-SOI process. A re-design of the impedance tuner could also be performed, were an off-the-shelf component would be the main objective, or a test/verification of the existing tuner with more general/broader set of antennas could be interesting and would have a lower barrier to entry. This could verify what IL, ML, RL, and OIP_3 could be achieved when the impedance tuner is not optimized for the specific antenna, and what system performance gain could then be obtained, although it might be lower than for a co-optimized antenna and impedance tuner. Finally, a

fully integrated solution could be possible, by using the same CMOS-SOI technology, but the all the required (large) inductors would occupy a large chip area, and the losses would be higher compared to using off-chip inductors.

The power amplifier circuits in this work [**II**, **III**, **V**] are based on the injection locking technique, together with control by the use of dynamic transistor biasing. By combining these techniques, power amplification with excellent PAE was obtained, and although the output power was lower than in some state-of-the-art CMOS PAs, this originated from the comparably low supply voltage used. The linearity was sufficient both for a W-CDMA UL signal and for a SC-FDMA LTE UL signal, when DPD was used, the requirement of DPD being the main bottleneck of this circuit approach. With all the different flavors of this concept, either in-band linearity in [**III**] or BW in [**V**] could be improved compared to the first version of the circuit in [**II**]. Finally, a new algorithm to reduce the BW expansion in polar modulation was proposed in [**IV**] and verified by utilizing it while measuring the circuit in [**III**]. The main difficulty when measuring with the 20MHz UL LTE signal, is the time-alignment between the polar signal paths, which requires a time resolution of about 1ns ($\sim 1/100 f_{BB}$).

The injection-lock based circuits were also analyzed theoretically and the most important results are shown in this thesis. The analysis shows that the functionality of this specific circuit is equivalent to the traditional IL-oscillators, while also including a tapped-cascode transistor, enabling higher supply voltage without significantly increasing the overall PA power dissipation.

In future work, this PA concept should be integrated with an SM on the same die. This would enable a study of the possibility for co-existence, with regards to the switching-noise from the SM and the different on-chip propagation paths. The elimination of requirement of DPD from digital baseband should be the main goal, or at least to minimize the requirement by the use of a simple analog path, either an amplitude feedforward or an amplitude feedback-loop. This would ease the BW requirement of the baseband DACs, if it would be sufficient to apply just in-band DPD instead of the wideband DPD (e_{HF}) applied to the baseband signal $V_{DD}(t)$ in this work. For the ILPA/ILPU, the lack of output power could be mitigated by using more stacked transistors for the cascode so that the supply voltage could be increased. This would increase the output power, but the final PAE would be lower, especially at higher back-off levels. Finally, there is one more way to improve the PAE of the ILPA, by further reducing the dynamic power dissipation of the injection node by using an inductor to put it in resonance. However, this improvement would limit the maximum achievable BW and also increase the required chip-area. The main upside, on the other hand, is an RFPA that can be designed/optimized for higher back-off levels, which would improve both linearity and efficiency.

Appendix A

Appendix A: Basic Injection-Lock Theory

The injection-lock mechanism is very non-linear, as can be seen in the example in Fig. A.1. In this example there is an oscillator with an oscillation frequency of f_o, and a signal with the frequency $f_o+\Delta f_o$ is injected into the oscillator. First the injected signal power is low, and then it is increased until the oscillator becomes locked to the injected signal. When the injected signal is weak, the output signal spectrum has two sidebands, at $f_o-\Delta f_o$ and $f_o+\Delta f_o$, and these represent an amplitude modulation component. In the time-domain it is a sinusoidal amplitude modulation with the modulation frequency Δf_o. As the injected signal power increases, the signal starts to pull the frequency of the oscillator towards $f_o+\Delta f_o$, and eventually when the signal is strong enough the oscillator will lock to the frequency of the injected signal ($f_o+\Delta f_o$).

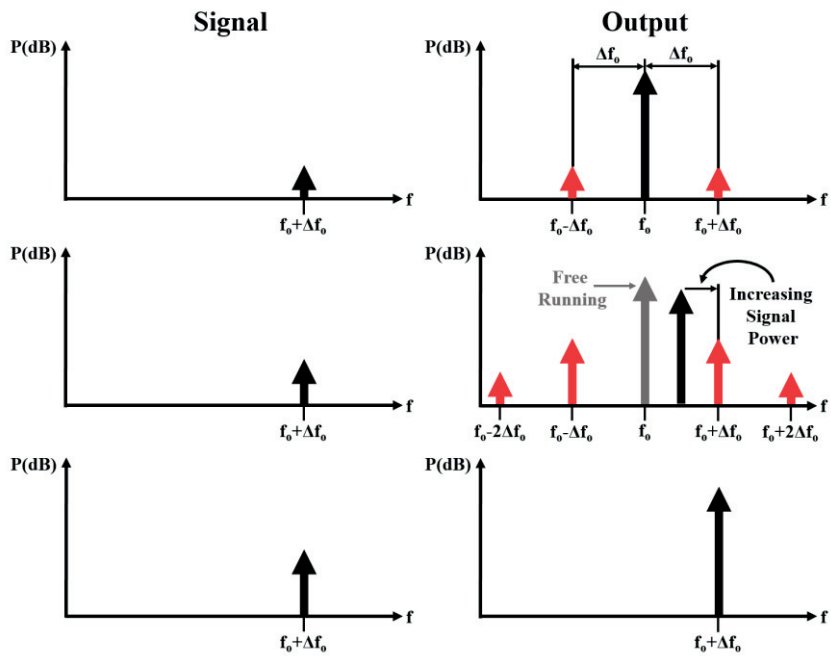

Fig. A.1. Spectrum of oscillator injected with different signal strengths.

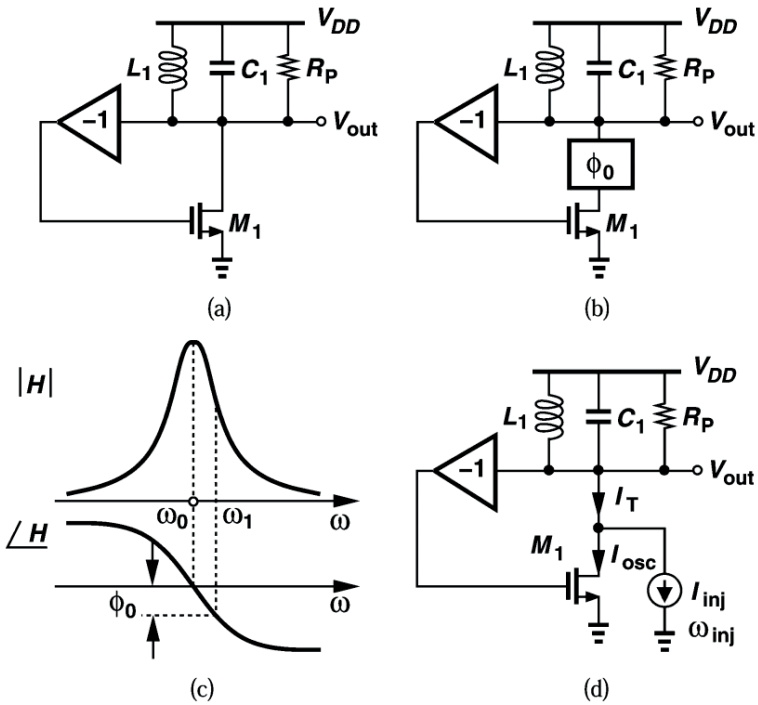

Fig. A.2 [87]. (a) Conceptual oscillator. (b) Frequency shift due to additional phase shift. (c) Open-loop characteristics. (d) Frequency shift by injection.

This section is based on Prof. Razavi's analysis in [87] and starts from Fig. A.2, also from [87]. Fig. A.2a illustrates an ideal oscillator with an LC-tank (L_1, C_1 and R_P) and a positive feedback (-1 and M_1). By investigating the startup loop-gain (A_v) (A.1), it can be seen that it is the product between g_m and the LC-tank impedance at the oscillation frequency (R_p). For large signals the loop-gain can be defined as the product of the large-signal transconductance (G_m) and the LC-tank impedance at the oscillation frequency ($Z_T(\omega_o)$). Further, to model the signal injection in the oscillator a phase shift of ϕ_0 is added to the feedback current, as can be seen in Fig. A.2b. When this phase shift is added to the oscillator feedback loop, for oscillation to still occur the opposite phase shift must be present in the other parts of the loop, so that the complete feedback loop has no phase shift. The phase shift of the LC-tank is frequency dependent, and the oscillator loop will thus compensate for the added phase shift by changing the output frequency to ω_1, which is best shown in Fig. A.2c. The loop-gain can then be derived as (A.2), where the phase shift has been added and the LC-tank impedance is at ω_1, as $Z_T(\omega_1)$. As explained, this change in frequency causes the phase shift in the LC-tank impedance, and the oscillator does not operate at the LC-tank resonance frequency when the phase shift is present. The magnitude of the LC-tank impedance is then lower, and the voltage amplitude

is reduced. It is therefore appropriate to convert the tank impedance into polar coordinates and (A.2) then becomes (A.3). Finally, the way the phase shift occurs in an injection-locked oscillator is shown in Fig. A.2d, and also shown are the currents in the injection node, I_T, I_{osc}, and I_{inj}. The phase shift of the current I_T with respect to I_{osc} can then be analyzed using the vector representation in Fig. A.3.

$$A_v = g_m R_P > 1 \tag{A.1}$$

$$A_v = G_m e^{j\phi_0} Z_T(\omega_1) = 1 \tag{A.2}$$

$$A_v = G_m e^{j\phi_0} |Z_T(\omega_1)| e^{-j\phi_0} = 1 \tag{A.3}$$

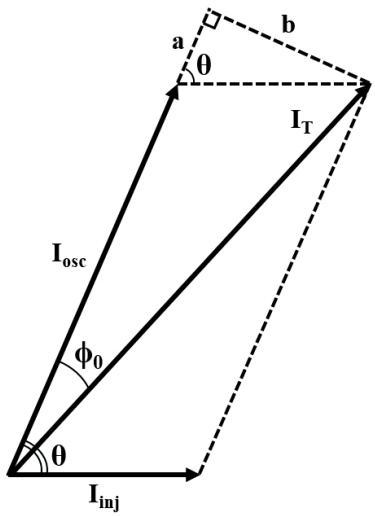

Fig. A.3. Injection-lock, a vector representation.

In order to derive the geometry of the vectors in Fig. A3, as was done in Adler's, Razavi's, and Kurokawa's papers [86-88], an angle (θ) between I_{osc} and I_{inj} is assumed. Furthermore, to solve the geometry an extra tringle can be added (a, b, and I_{inj}), which is used to derive (A.4) and (A.5), while (A.6) comes from substituting (A.5) into (A.4). Then by applying KCL ($I_T = \left| I_{osc} e^{j\theta} + I_{inj} \right|$) to (A.6), (A.7) is obtained.

$$sin(\phi_0) = \frac{b}{I_T} \tag{A.4}$$

$$sin(\theta) = \frac{b}{I_{inj}} \tag{A.5}$$

$$sin(\phi_0) = \frac{I_{inj}}{I_T} sin(\theta) \tag{A.6}$$

$$sin(\phi_0) = \frac{I_{inj}}{|I_{osc}e^{j\theta} + I_{inj}|} sin(\theta) \tag{A.7}$$

The extreme values of angles (ϕ_0 and θ) can be derived from Fig. A.4, illustrating the situation at the edge of the locking-range. (A.8) is derived readily from the figure.

$$sin(\phi_{0,max}) = \frac{I_{inj}}{I_{osc}} \tag{A.8}$$

At the edge of the locking-range the angle between I_{osc} and I_{inj} is equal to $\pi/2 + \phi_{o,max}$, from [87], and this can also be seen in Fig. A.4.

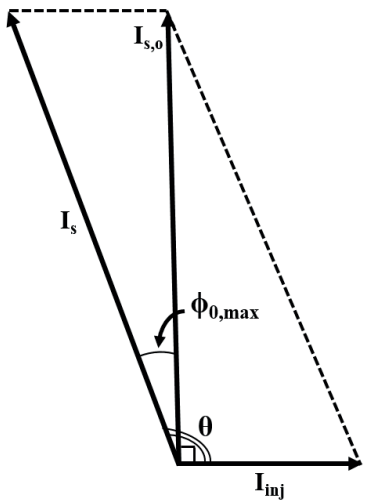

Fig. A.4. Vector representation at the edge of the large-signal locking-range.

The LC-tank phase shift is (A.9), from Adler [86] and Razavi [87]. From Fig. A.4 it is found that $tan(\phi_o) = I_{inj}/I_T$ and the (inverse) Pythagorean theorem gives $I_T = \sqrt{I_{osc}^2 - I_{inj}^2}$. Applying this to (A.9) we get (A.10), and finally the maximum locking-range is obtained in (A.11).

$$tan(\phi_0) \approx \frac{2Q}{\omega_0}(\omega_0 - \omega) \tag{A.9}$$

$$tan(\phi_0) \approx \frac{2Q}{\omega_0}(\omega_0 - \omega_{inj}) = \frac{I_{inj}}{I_T} = \frac{I_{inj}}{\sqrt{I_{osc}^2 - I_{inj}^2}} \tag{A.10}$$

$$\omega_{L,max} = \left(\omega_0 - \omega_{inj}\right) = \frac{\omega_0}{2Q} \cdot \frac{I_{inj}}{I_{osc}} \cdot \frac{1}{\sqrt{1-\frac{I_{inj}^2}{I_{osc}^2}}} \tag{A.11}$$

For weak signal injection locking, the mechanism was theoretically described by Adler in [86], which was the first significant theoretical break-through in the analysis of injection locking. For weak injection-lock it is assumed that $I_{inj}<<I_{osc}$ and (A.6) is then assumed approximately equal to (A.9), which gives (A.12) and (A.13). While the injection-lock is weak, the phase shift between I_{osc} and I_{inj} is assumed to be $\pi/2$ *(sin(θ)=1)* at the edge of the locking-range, and therefore, the weak injection-lock range becomes (A.14).

$$sin(\phi_0) = \frac{I_{inj}}{I_T} sin(\theta) \approx tan(\phi_0) \approx \frac{2Q}{\omega_0}\left(\omega_0 - \omega_{inj}\right) \tag{A.12}$$

$$sin(\theta) = \frac{2Q}{\omega_0}\frac{I_T}{I_{inj}}\left(\omega_0 - \omega_{inj}\right) \tag{A.13}$$

$$\omega_L = \left(\omega_0 - \omega_{inj}\right) = \frac{\omega_0}{2Q} \cdot \frac{I_{inj}}{I_T} \tag{A.14}$$

References

[1] J. Eberspächer, H.-J. Vögel, C. Bettstetter, and C. Hartmann, "GSM - Architecture, Protocols and Services," *3rd edition, New York, NY, USA: Wiley,* 2009.

[2] https://www.ericsson.com/en/about-us/history/changing-the-world/world-leadership/gsm-and-ericsson

[3] E. Dahlman, S. Parkvall, and J. Sköld, "4G, LTE-Advanced Pro and the Road to 5G," *1st edition, Cambridge, MA, USA: Academic Press,* 2016.

[4] G. E. Moore, "Cramming more components onto integrated circuits," *Electronics 38*, pp.114–117, 1965.

[5] S. H. Fuller and L. I. Millett, "Computing Performance: Game Over or Next Level?," in *Computer*, vol. 44, no. 1, pp. 31-38, Jan. 2011.

[6] H. Fukui, "Determination of the basic device parameters of a GaAs MESFET," in *The Bell System Technical Journal*, vol. 58, no. 3, pp. 771-797, March 1979.

[7] H. -. Chen, K. -. Chen, G. -. Huang, and C. -. Chang, "Small-Signal Modeling of SiGe HBTs Using Direct Parameter-Extraction Method," in *IEEE Transactions on Electron Devices*, vol. 53, no. 9, pp. 2287-2295, Sept. 2006.

[8] F. O. Eynde *et al*., "A fully-integrated single-chip SOC for Bluetooth," *2001 IEEE International Solid-State Circuits Conference. Digest of Technical Papers. ISSCC (Cat. No.01CH37177)*, San Francisco, CA, USA, 2001, pp. 196-197.

[9] F. Bruccoleri, E. A. M. Klumperink, and B. Nauta, "Wide-band CMOS low-noise amplifier exploiting thermal noise canceling," in *IEEE Journal of Solid-State Circuits*, vol. 39, no. 2, pp. 275-282, Feb. 2004.

[10] P. J. Sulivan, B. A. Xavier, and W. H. Ku, "Low voltage performance of a microwave CMOS Gilbert cell mixer," in *IEEE Journal of Solid-State Circuits*, vol. 32, no. 7, pp. 1151-1155, July 1997.

[11] E. S. Atalla, F. Zhang, P. T. Balsara, A. Bellaouar, S. Ba, and K. Kiasaleh, "Time-Domain Analysis of Passive Mixer Impedance: A Switched-Capacitor Approach," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 64, no. 2, pp. 347-359, Feb. 2017.

[12] A. Mirzaei and H. Darabi, "Analysis of Imperfections on Performance of 4-Phase Passive-Mixer-Based High-Q Bandpass Filters in SAW-Less Receivers," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 58, no. 5, pp. 879-892, May 2011.

[13] A. Mirzaei, H. Darabi, J. C. Leete, and Y. Chang, "Analysis and Optimization of Direct-Conversion Receivers With 25% Duty-Cycle Current-Driven Passive Mixers," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 57, no. 9, pp. 2353-2366, Sept. 2010.

[14] J. A. Weldon *et al*., "A 1.75-GHz highly integrated narrow-band CMOS transmitter with harmonic-rejection mixers," in *IEEE Journal of Solid-State Circuits*, vol. 36, no. 12, pp. 2003-2015, Dec. 2001.

[15] R. B. Staszewski, Chih-Ming Hung, N. Barton, Meng-Chang Lee, and D. Leipold, "A digitally controlled oscillator in a 90 nm digital CMOS process for mobile phones," in *IEEE Journal of Solid-State Circuits*, vol. 40, no. 11, pp. 2203-2211, Nov. 2005.

[16] S. Cuk and R. W. Erickson, "A conceptually new high-frequency switched-mode amplifier technique eliminates current ripple", *Proc. Fifth National Solid-State Power Conversion Conf. (Powercon5)*, May. 1978, pp. G3.l-G3.22.

[17] S. Leuschner, J. Mueller, and H. Klar, "A 1.8GHz wide-band stacked-cascode CMOS power amplifier for WCDMA applications in 65nm standard CMOS," *2011 IEEE Radio Frequency Integrated Circuits Symposium*, Baltimore, MD, 2011, pp. 1-4.

[18] C. Andrews and A. C. Molnar, "A Passive Mixer-First Receiver With Digitally Controlled and Widely Tunable RF Interface," in *IEEE Journal of Solid-State Circuits*, vol. 45, no. 12, pp. 2696-2708, Dec. 2010.

[19] R. W. Chang, "Synthesis of band-limited orthogonal signals for multichannel data transmission," in *The Bell System Technical Journal*, vol. 45, no. 10, pp. 1775-1796, Dec. 1966.

[20] A. F. Molisch, "Wireless Communications," *2nd edition, New York, NY, USA: Wiley,* 2010.

[21] A. Paulraj and T. Kailath, "Increasing capacity in wireless broadcast systems using distributed transmission/directional reception (DTDR)," Sept. 1994. US Patent 5,345,599

[22] Da-Shan Shiu, G. J. Foschini, M. J. Gans and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," in *IEEE Transactions on Communications*, vol. 48, no. 3, pp. 502-513, March 2000.

[23] P. Sjöblom and H. Sjöland, "An adaptive impedance tuning CMOS circuit for ISM 2.4-GHz band," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 52, no. 6, pp. 1115-1124, June 2005.

[24] http://niviuk.free.fr/gsm_band.php

[25] http://niviuk.free.fr/lte_band.php

[26] M. Nilsson *et al*., "A 9-band WCDMA/EDGE transceiver supporting HSPA evolution," *2011 IEEE International Solid-State Circuits Conference*, San Francisco, CA, 2011, pp. 366-368.

[27] M. Ingels *et al*., "A 5 mm^{2}40 nm LP CMOS Transceiver for a Software-Defined Radio Platform," in *IEEE Journal of Solid-State Circuits*, vol. 45, no. 12, pp. 2794-2806, Dec. 2010.

[28] A. A. Abidi, "Direct-conversion radio transceivers for digital communications," in *IEEE Journal of Solid-State Circuits*, vol. 30, no. 12, pp. 1399-1410, Dec. 1995.

[29] J. C. Rudell *et al*., "A 1.9-GHz wide-band IF double conversion CMOS receiver for cordless telephone applications," in *IEEE Journal of Solid-State Circuits*, vol. 32, no. 12, pp. 2071-2088, Dec. 1997.

[30] V. Aparin and L. E. Larson, "Analysis and reduction of cross-modulation distortion in CDMA receivers," in *IEEE Transactions on Microwave Theory and Techniques*, vol. 51, no. 5, pp. 1591-1602, May 2003.

[31] M. E. Heidari and A. A. Abidi, "Behavioral models of frequency pulling in oscillators," *2007 IEEE International Behavioral Modeling and Simulation Workshop*, San Jose, CA, 2007, pp. 100-104.

[32] S. Marksteiner *et al*., "A miniature BAW duplexer using flip-chip on LTCC," *IEEE Symposium on Ultrasonics, 2003*, Honolulu, HI, USA, 2003, pp. 1794-1797 Vol. 2.

[33] S. Kannangara and M. Faulkner, "Analysis of an Adaptive Wideband Duplexer With Double-Loop Cancellation," in *IEEE Transactions on Vehicular Technology*, vol. 56, no. 4, pp. 1971-1982, July 2007.

[34] M. Mikhemar, H. Darabi, and A. Abidi, "A tunable integrated duplexer with 50dB isolation in 40nm CMOS," *2009 IEEE International Solid-State Circuits Conference - Digest of Technical Papers*, San Francisco, CA, 2009, pp. 386-387,387a.

[35] M. Mikhemar, H. Darabi, and A. A. Abidi, "A Multiband RF Antenna Duplexer on CMOS: Design and Performance," in *IEEE Journal of Solid-State Circuits*, vol. 48, no. 9, pp. 2067-2077, Sept. 2013.

[36] S. H. Abdelhalem, P. S. Gudem, and L. E. Larson, "Hybrid Transformer-Based Tunable Differential Duplexer in a 90-nm CMOS Process," in *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 3, pp. 1316-1326, March 2013.

[37] I. u. Din, J. Wernehag, S. Andersson, and H. Sjöland, "Two Tunable Frequency Duplexer Architectures for Cellular Transceivers," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 64, no. 11, pp. 2988-2999, Nov. 2017.

[38] https://www.3gpp.org/

[39] D. Montanari *et al*., "An FDD Wireless Diversity Receiver With Transmitter Leakage Cancellation in Transmit and Receive Bands," in *IEEE Journal of Solid-State Circuits*, vol. 53, no. 7, pp. 1945-1959, July 2018.

[40] T. Zhang, A. R. Suvarna, V. Bhagavatula, and J. C. Rudell, "An Integrated CMOS Passive Self-Interference Mitigation Technique for FDD Radios," in *IEEE Journal of Solid-State Circuits*, vol. 50, no. 5, pp. 1176-1188, May 2015.

[41] "3GPP TS 36.104 V11.3.1", *3rd Generation Partnership Project; LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Base Station (BS) Radio Transmission and Reception (Release 11)*, Feb. 2013.

[42] M. Pelosi, O. Franek, M. B. Knudsen, M. Christensen, and G. F. Pedersen, "A Grip Study for Talk and Data Modes in Mobile Phones," in *IEEE Transactions on Antennas and Propagation*, vol. 57, no. 4, pp. 856-865, April 2009.

[43] B. D. Van Veen and K. M. Buckley, "Beamforming: a versatile approach to spatial filtering," in *IEEE ASSP Magazine*, vol. 5, no. 2, pp. 4-24, April 1988.

[44] L. Sundström, G. Jönsson, and H. Börjeson "Radio Electronics, " *2015 edition, Lund University, Lund, Sweden,* 2015.

[45] E. L. Firrao, A. Annema and B. Nauta, "An Automatic Antenna Tuning System Using Only RF Signal Amplitudes," in *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 55, no. 9, pp. 833-837, Sept. 2008.

[46] P. Sjöblom and H. Sjöland, "Measured CMOS Switched High-Quality Capacitors in a Reconfigurable Matching Network," in *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 54, no. 10, pp. 858-862, Oct. 2007.

[47] R. A. Johnson *et al*., "Silicon-on-sapphire MOSFET transmit/receive switch for L and S band transceiver applications," in *Electronics Letters*, vol. 33, no. 15, pp. 1324-1326, 17 July 1997.

[48] C. Tinella, J. M. Fournier, D. Belot, and V. Knopik, "A high-performance CMOS-SOI antenna switch for the 2.5-5-GHz band," in *IEEE Journal of Solid-State Circuits*, vol. 38, no. 7, pp. 1279-1283, July 2003.

[49] A. A. Tamez, J. A. Fredenburg and M. P. Flynn, "An integrated 120 volt AC mains voltage interface in standard 130 nm CMOS," *2010 Proceedings of ESSCIRC*, Seville, 2010, pp. 238-241.

[50] P. R. Gray, R. G. Meyer, "Analysis and Design of Analog Integrated Circuits," 3rd *edition*, *New York, NY, USA: Wiley*, 1993

[51] T. H. Lee, "The Design of CMOS Radio-Frequency Integrated Circuits," *1st edition, Cambridge, England, United Kingdom: Cambridge University Press*, 1998

[52] I. Vasilev, V. Plicanic, and B. K. Lau, "Impact of Antenna Design on MIMO Performance for Compact Terminals With Adaptive Impedance Matching," in *IEEE Transactions on Antennas and Propagation*, vol. 64, no. 4, pp. 1454-1465, April 2016.

[53] S. C. Cripps, "Advanced Techniques in RF Power Amplifier Design," *1st edition, Norwood, MA: Artech House,* 2002.

[54] J. Fritzin, C. Svensson, and A. Alvandpour, "Analysis of a 5.5-V Class-D Stage Used in $+$ 30-dBm Outphasing RF PAs in 130- and 65-nm CMOS," in *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 59, no. 11, pp. 726-730, Nov. 2012.

[55] N. O. Sokal and A. D. Sokal, "Class E-A new class of high-efficiency tuned single-ended switching power amplifiers," in *IEEE Journal of Solid-State Circuits*, vol. 10, no. 3, pp. 168-176, June 1975.

[56] F. Raab, "Idealized operation of the class E tuned power amplifier," in *IEEE Transactions on Circuits and Systems*, vol. 24, no. 12, pp. 725-735, Dec. 1977.

[57] J. M. Rabaey, A. Chandrakasan, B. Nikolic, "Digital Integrated Circuits: A Design Perspective," *2nd edition, London, United Kingdom: Pearson*, 2003

[58] P. Wright, A. Sheikh, C. Roff, P. J. Tasker, and J. Benedikt, "Highly efficient operation modes in GaN power transistors delivering upwards of 81% efficiency and 12W output power," *2008 IEEE MTT-S International Microwave Symposium Digest*, Atlanta, GA, USA, 2008, pp. 1147-1150.

[59] F. H. Raab, "The class BD high-efficiency RF power amplifier," in *IEEE Journal of Solid-State Circuits*, vol. 12, no. 3, pp. 291-298, June 1977.

[60] W. H. Doherty, "A New High Efficiency Power Amplifier for Modulated Waves," in *Proceedings of the Institute of Radio Engineers*, vol. 24, no. 9, pp. 1163-1182, Sept. 1936.

[61] R. Pengelly, C. Fager, and M. Özen, "Doherty's Legacy: A History of the Doherty Power Amplifier from 1936 to the Present Day," in *IEEE Microwave Magazine*, vol. 17, no. 2, pp. 41-58, Feb. 2016.

[62] P. Saad, R. Hou, R. Hellberg, and B. Berglund, "Ultra-Wideband Doherty-Like Power Amplifier," *2018 IEEE/MTT-S International Microwave Symposium - IMS*, Philadelphia, PA, 2018, pp. 1215-1218.

[63] P. Saad, P. Colantonio, L. Piazzon, F. Giannini, K. Andersson, and C. Fager, "Design of a Concurrent Dual-Band 1.8–2.4-GHz GaN-HEMT Doherty Power Amplifier," in *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 6, pp. 1840-1849, June 2012.

[64] E. Kaymaksut and P. Reynaert, "Dual-Mode CMOS Doherty LTE Power Amplifier With Symmetric Hybrid Transformer," in *IEEE Journal of Solid-State Circuits*, vol. 50, no. 9, pp. 1974-1987, Sept. 2015.

[65] F. Wang, D. F. Kimball, D. Y. Lie, P. M. Asbeck, and L. E. Larson, "A Monolithic High-Efficiency 2.4-GHz 20-dBm SiGe BiCMOS Envelope-Tracking OFDM Power Amplifier," in *IEEE Journal of Solid-State Circuits*, vol. 42, no. 6, pp. 1271-1281, June 2007.

[66] H. Harju, T. Rautio, S. Hietakangas, and T. Rahkonen, "Envelope tracking power amplifier with static predistortion linearization," *2007 18th European Conference on Circuit Theory and Design*, Seville, 2007, pp. 388-391.

[67] B. Park *et al*., "High-Performance CMOS Power Amplifier With Improved Envelope Tracking Supply Modulator," in *IEEE Transactions on Microwave Theory and Techniques*, vol. 64, no. 3, pp. 798-809, March 2016.

[68] J. Vuolevi, T. Rahkonen, "Distortion in RF Power Amplifiers," *1st edition, Norwood, MA: Artech House,* 2003.

[69] L. R. Kahn, "Comparison of Linear Single-Sideband Transmitters with Envelope Elimination and Restoration Single-Sideband Transmitters," in *Proceedings of the IRE*, vol. 44, no. 12, pp. 1706-1712, Dec. 1956.

[70] L. R Kahn, "Single sideband transmission envelope elimination and restoration," in *Proceedings of the IRE,* vol 40, no. 7, pp. 803-806, July 1952.

[71] D. K. Su and W. J. McFarland, "An IC for linearizing RF power amplifiers using envelope elimination and restoration," in *IEEE Journal of Solid-State Circuits*, vol. 33, no. 12, pp. 2252-2258, Dec. 1998.

[72] F. H. Raab, "Intermodulation distortion in Kahn-technique transmitters," in *IEEE Transactions on Microwave Theory and Techniques*, vol. 44, no. 12, pp. 2273-2278, Dec. 1996.

[73] P. Reynaert and M. S. J. Steyaert, "A 1.75-GHz polar modulated CMOS RF power amplifier for GSM-EDGE," in *IEEE Journal of Solid-State Circuits*, vol. 40, no. 12, pp. 2598-2608, Dec. 2005.

[74] J. Wang, A. Zhu, X. Zhu, and T. J. Brazil, "Vector Hole Punching Technique for OFDM Signals Using Circle-Tangent Shift and Unused Tones," in *IEEE Transactions on Microwave Theory and Techniques*, vol. 57, no. 11, pp. 2682-2691, Nov. 2009.

[75] M. Apostolidou, M. P. van der Heijden, D. M. W. Leenaerts, J. Sonsky, A. Heringa, and I. Volokhine, "A 65 nm CMOS 30 dBm Class-E RF Power Amplifier With 60% PAE and 40% PAE at 16 dB Back-Off," in *IEEE Journal of Solid-State Circuits*, vol. 44, no. 5, pp. 1372-1379, May 2009.

[76] F. Wang *et al*., "An Improved Power-Added Efficiency 19-dBm Hybrid Envelope Elimination and Restoration Power Amplifier for 802.11g WLAN Applications," in *IEEE Transactions on Microwave Theory and Techniques*, vol. 54, no. 12, pp. 4086-4099, Dec. 2006.

[77] R. Shrestha, R. van der Zee, A. de Graauw, and B. Nauta, "A Wideband Supply Modulator for 20 MHz RF Bandwidth Polar PAs in 65 nm CMOS," in *IEEE Journal of Solid-State Circuits*, vol. 44, no. 4, pp. 1272-1280, April 2009.

[78] T. Kwak, M. Lee, and G. Cho, "A 2 W CMOS Hybrid Switching Amplitude Modulator for EDGE Polar Transmitters," in *IEEE Journal of Solid-State Circuits*, vol. 42, no. 12, pp. 2666-2676, Dec. 2007.

[79] M. C. W. Hoyerby and M. A. E. Andersen, "Envelope tracking power supply with fully controlled 4th order output filter," *Twenty-First Annual IEEE Applied Power Electronics Conference and Exposition, 2006. APEC '06*, Dallas, TX, 2006, pp. 8 pp.-.

[80] C. Tao and A. A. Fayed, "A Low-Noise PFM-Controlled Buck Converter for Low-Power Applications," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 59, no. 12, pp. 3071-3080, Dec. 2012.

[81] D. Cox, "Linear Amplification with Nonlinear Components," in *IEEE Transactions on Communications*, vol. 22, no. 12, pp. 1942-1945, December 1974.

[82] Y. Liu, W. Pan, S. Shao, and Y. Tang, "A General Digital Predistortion Architecture Using Constrained Feedback Bandwidth for Wideband Power Amplifiers," in *IEEE Transactions on Microwave Theory and Techniques*, vol. 63, no. 5, pp. 1544-1555, May 2015.

[83] S. Kousai, K. Onizuka, T. Yamaguchi, Y. Kuriyama, and M. Nagaoka, "A 28.3mW PA-closed loop for linearity and efficiency improvement integrated in a +27.1dBm WCDMA CMOS power amplifier," *2012 IEEE International Solid-State Circuits Conference*, San Francisco, CA, 2012, pp. 84-86.

[84] W. F. Egan, "Frequency Synthesis by Phase Lock," *2nd edition, New York, NY, USA: Wiley,* 1999.

[85] B. van der Pol, "The Nonlinear Theory of Electric Oscillations," in *Proceedings of the Institute of Radio Engineers*, vol. 22, no. 9, pp. 1051-1086, Sept. 1934.

[86] R. Adler, "A Study of Locking Phenomena in Oscillators," in *Proceedings of the IRE*, vol. 34, no. 6, pp. 351-357, June 1946.

[87] B. Razavi, "A study of injection locking and pulling in oscillators," in *IEEE Journal of Solid-State Circuits*, vol. 39, no. 9, pp. 1415-1424, Sept. 2004.

[88] K. Kurokawa, "Injection locking of microwave solid-state oscillators," in *Proceedings of the IEEE*, vol. 61, no. 10, pp. 1386-1410, Oct. 1973.

[89] A. Rofougaran, J. Rael, M. Rofougaran, and A. Abidi, "A 900 MHz CMOS LC-oscillator with quadrature outputs," *1996 IEEE International Solid-State Circuits Conference. Digest of Technical Papers, ISSCC*, San Francisco, CA, USA, 1996, pp. 392-393.

[90] P. Andreani, A. Bonfanti, L. Romano, and C. Samori, "Analysis and design of a 1.8-GHz CMOS LC quadrature VCO," in *IEEE Journal of Solid-State Circuits*, vol. 37, no. 12, pp. 1737-1747, Dec. 2002.

[91] A. Kitipongwatana, P. Koseeyaporn, J. Koseeyaporn, and P. Wardkein, "Enhanced mathematical model for studying fundamental behaviors of injection locked relaxation oscillator," *Proceedings of the International MultiConference of Engineers and Computer Scientists II, IMECS 2015*, Hong Kong, March 2015.

[92] W. Chen *et al*., "A 53.6 GHz direct injection-locked frequency divider with a 72% locking range in 65 nm CMOS technology," *2013 IEEE MTT-S International Microwave Symposium Digest (MTT)*, Seattle, WA, 2013, pp. 1-3.

[93] R. A. York and T. Itoh, "Injection- and phase-locking techniques for beam control [antenna arrays]," in *IEEE Transactions on Microwave Theory and Techniques*, vol. 46, no. 11, pp. 1920-1929, Nov. 1998.

[94] K.-C. Tsai and P. R. Gray, "A 1.9-GHz, 1-W CMOS class-E power amplifier for wireless communications," in *IEEE Journal of Solid-State Circuits*, vol. 34, no. 7, pp. 962-970, July 1999.

[95] S. Jin *et al*., "A Highly Efficient CMOS Envelope Tracking Power Amplifier Using All Bias Node Controls," in *IEEE Microwave and Wireless Components Letters*, vol. 25, no. 8, pp. 517-519, Aug. 2015.

Part II

Included papers

Summary of included papers

Paper I

J. Lindstrand, I. Vasilev, and H. Sjöland, "A low band cellular terminal antenna impedance tuner in 130nm CMOS-SOI technology," *in Proc. European Solid-State Circuit Conference*, Venice, Italy, Sept. 2014, pp. 459-462.

Contribution:
This paper presents an impedance tuner in CMOS-SOI technology. The circuit was verified as state-of-the-art with high power handling (2W/+33dBm), high linearity (IP_3:56-60dBm), and low insertion loss (<1dB). The internal loss mechanisms were also characterized over the different impedance states, which to the knowledge of the authors was the first time such measurements were presented.

I contributed with the design of the impedance tuner and PCB, assembly of the CMOS-SOI tuner and PCB, the measurement methods in a non-50Ω environment, performed the measurements, and wrote the published paper.

Paper II

J. Lindstrand, C. Bryant, M. Törmänen, and H. Sjöland, "A 1.6-2.6GHz 29dBm injection-locked power amplifier with 64% peak PAE in 65nm CMOS," *in Proc. European Solid-State Circuit Conference*, Helsinki, Finland, Sept. 2011, pp. 299–302.

Contribution:
The paper presents the first ILPA, to the authors' knowledge, that uses supply-modulation and dynamic biasing. The circuit was verified with an output power of +29dBm (high for 65nm CMOS) and a peak PAE of 64%. This work showed that ILPAs can provide high performance also when amplifying modulated signals with a non-constant envelope.

The first author contributed with the design of the CMOS-IC for the ILPA, the idea of the concept, designed the off-chip output matching network, verified the concept by measurement and wrote the published paper.

Paper III

J. Lindstrand, M. Törmänen, and H. Sjöland, "An injection-locked power up-converter in 65-nm CMOS for cellular applications," *in IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 3, pp. 1065–1077, March 2019.

Contribution:
Presented in this paper is a circuit that inputs a BB signal and LO, and outputs a high power modulated RF-carrier, which can be used to drive the antenna of a cellphone. The journal article includes IC design, design of an off-chip matching network for the flip-chip assembled CMOS chip, verification with a world record in PAE (for a CMOS PA at 2GHz), and also a theoretical analysis of the key merits of the circuit.

My contributions were the design of the CMOS-IC, matching network, measurements (also with LTE modulation and DPD), paper writing, and also all the theoretical calculations, simulations, and evaluations.

Paper IV

J. Lindstrand, M. Törmänen, and H. Sjöland, "Origin attraction – a technique to reduce signal bandwidth in polar transmitters," Submitted to *IEEE Transactions on Circuits and Systems I: Regular Papers.*

Contribution:
The paper presents a novel algorithm, in different versions, to reduce the bandwidth expansion of polar modulated signals. The main differences between versions are related to the complexity/amount of hardware required. The algorithm is compared with the traditional BW reduction technique called “Vector-hole-modulation” and there is a significant improvement in bandwidth, especially in the phase-signal, which is the major bottleneck in polar modulated systems. The algorithm was also used in the verification of the circuit in Paper III.

I designed the novel algorithm, performed all the required simulations in MatLab, implemented the algorithm into the measurement hardware, and wrote the paper.

Paper V

J. Lindstrand, M. Törmänen, and H. Sjöland, "A Decade Frequency Range CMOS Power Amplifier for Sub-6 GHz Cellular Terminals," Accepted for Publication in *IEEE Microwave and Wireless Components Letters.*

Contribution:
Demonstrated in this paper is a CMOS PA with a frequency range of one decade, with high verified RF performance over the entire frequency range. To achieve this, a novel MN is introduced along with a dual-output structure. The measured PAE is close to 60% and the output power is +28dBm, all from a single wideband CMOS circuit with two outputs. The circuit was verified with a 20MHz 16QAM SC-FDMA modulated signal in four different bands within the sub-6 GHz frequency range.

My contributions were the design of the CMOS-IC, inventing and designing the new matching network, verifying the matching network with special PCB probes, verifying the complete circuit solution, theoretical calculations and analysis of the new matching network, and writing the paper.