Microwave CMOS Beamforming Transmitters

Johan Wernehag

Lund 2008

LUND INSTITUTE OF TECHNOLOGY
Lund University

Department of Electrical and Information Technology
Lund University
P.O. Box 118
SE-221 00 LUND
SWEDEN

No. 12
ISSN 1654-790X
Series of licentiate and doctoral theses

Produced using LaTeX Documentation System.
Printed in Sweden by *Tryckeriet i E-huset*, Lund.
November 2008.

Abstract

The increase of the consumer electronics market the last couple of decades has been one of the main drivers of IC process technology development. The majority of the ICs are used in digital applications, and for these CMOS is the choice of technology. The urge to squeeze more transistors on to a given area has led to shrinking feature sizes. It has resulted in higher transition frequencies and reduced supply voltage. During the last decade the increasing transition frequency has enabled CMOS to be used in RF applications, as well. Unfortunately, the decreasing supply voltage that, until recently, has accompanied the reduced feature sizes makes it more difficult to build power amplifiers that can deliver the amount of power needed to transmit the radio signal over the desired distance. In the receiver, the reduced supply voltage has resulted in reduced signal swing, which compromises linearity and dynamic range.

In this thesis new topologies for the power amplifier is investigated, and the approach to combine the power from multiple power amplifiers is taken. In this way, despite the low supply voltage, the transmitted power by the IC can still be high. The increased transition frequency of CMOS technology can be used to increase the operating frequency to tens of GHz. The possibility for small sized phased antenna arrays then reveals, giving high directivity of the antenna and the potential for electrical beam steering. This both reduces interference to nearby receivers through spatial selectivity, and increases the equivalent isotropic radiated power. Power amplifiers with digital 360° phase control and antenna arrays have been investigated.

In recent years applications at high operating frequencies have attained much focus from both academia and industry, such as automotive radar at 77 GHz and WLAN at 60 GHz. Even though the shrinking feature sizes of CMOS transistors have resulted in transit frequencies above 150 GHz, the high frequency required by many applications is still a great challenge for the CMOS designer. Therefore, in Paper IV and Paper VI different approaches to keep the on chip frequency lower than the RF carrier frequency as long as possible have been taken. In Paper IV two different frequency doubling 60 GHz power amplifier topologies are presented, and in Paper VI a subharmonic mixer with 30 GHz radio frequency and 15 GHz differential local oscillator is presented.

Many transceiver architectures rely on quadrature signals driving the down- or upconversion mixers. The power amplifiers in Paper I and II need quadrature signals to implement the digital phase control. Therefore, in Paper V a three-stage active polyphase filter with quadrature output signals, high operation frequency, and wide bandwidth is analyzed. Analytical equations for both voltage gain and phase transfer function of a loaded stage are derived. The filter shows robustness against process parameter spread and achieves high quadrature signal quality from 6 GHz to 14 GHz.

Contents

Preface

This thesis summarizes my academic work in the Analog/RF Group at the Department of Electroscience and later the Department of Electrical and Information Technology, Faculty of Engineering, Lund University, for the Ph. D. degree in Circuit Design.

The content of the thesis is divided into two parts, where the first part is an introduction to the research area. General receiver and transmitter architectures are discussed, followed by different power amplifier topologies. CMOS transistor behavior and modeling is described in Chapter 3. In Chapter 4 link budget calculations are performed for two examples and in Chapter 5 the concept of phased array antennas for beam steering applications is illustrated. In Chapter 6 a few transmitter architectures suitable for integrated beamforming transmitters are discussed. The second part contains the included research papers.

Part I: Introduction

Chapter 1. Starts with a short motivation to the research area and then lists the main research contributions of the thesis.

Chapter 2. Begins with a short description of what a radio transceiver is. The most common transceiver architectures are then described. The heterodyne, the direct conversion, and the image rejection receivers are described as well as the direct conversion, the two-step, and the direct-modulation transmitters. The last section is dedicated to the power amplifier, which is, perhaps, the most critical and challenging block in the transmitter chain.

Chapter 3. The operation of the n-channel MOS transistor is described. A low and medium frequency small signal model is described in strong inversion. Some small dimension effects are treated, e. g. channel length modulation, drain induced barrier lowering, and velocity saturation. The main noise sources in the MOS transistor are also discussed.

Chapter 4. A short introduction to link budget calculations is presented and illustrated by two examples. The first example is a UMTS up-link budget calculation and the second example is a 60 GHz WPAN link budget calculation. In the second example, the use of an antenna array illustrates how to achieve high output power (EIRP) also in nanometer CMOS technologies.

Chapter 5. In this chapter the fundamentals of electrically scanned antennas are discussed. Topics such as directivity, array factor, and grating lobes are treated for phased array antennas. This chapter introduces the theory of how a beamforming transmitter should behave.

Chapter 6. A few transmitter architectures suitable for integrated beamforming transmitters are discussed in this chapter. The following architectures are touched upon: baseband phase shifting, local oscillator phase shifting,

offset local oscillator phase shifting, variable ring oscillator phase shifting, and radio frequency phase shifting.

Chapter 7. A short summary of the included research papers.

Chapter 8. Different aspects of the research results of the thesis are discussed and future work based on these results are elaborated on.

Part II: Included Research Papers
The main contribution to the thesis is derived from the following publications.

[1] Johan Wernehag, Henrik Sjöland, "An 8-GHz Beamforming Transmitter IC in 130-nm CMOS", *IEEE Radio Frequency Integrated Circuits Symposium*, pp. 577–580, June 3–5, 2007, Honolulu, HI, USA.

[2] Johan Wernehag, Henrik Sjöland, "A 24-GHz Automotive Radar Transmitter with Digital Beam Steering in 130-nm CMOS", *PhD Research in Microelectronics and Electronics*, pp. 481–484, June 11–16, 2006, Lecce Italy.

[3] Johan Wernehag, Henrik Sjöland, "Antenna Array for a 24-GHz Automotive Radar with Dipole Antenna Element Patches", *International Symposium on Antennas and Propagation*, November 1–4, 2006, Singapore Singapore.

[4] Johan Wernehag, Henrik Sjöland, "60 GHz 130-nm CMOS Second Harmonic Power Amplifiers", *IEEE Asia Pacific Conference on Circuits and Systems*, November 30–December 3, 2008, Macao China.

[5] Johan Wernehag, Henrik Sjöland, "Analysis of a High Frequency and Wide Bandwidth Active Polyphase Filter Based on CMOS Inverters", *Analog Integrated Circuits and Signal Processing*, Springer Science, 2008, submitted.

[6] Johan Wernehag, Henrik Sjöland, "A 30 GHz 90-nm CMOS Passive Subharmonic Mixer with 15 GHz Differential LO", *IEEE Proceedings of the 26th Norchip Conference*, November 17–18, 2008, Tallinn Estonia.

Related Publications: The following publications are not included but contain overlapping, complementing and some additional material:

Johan Wernehag, Henrik Sjöland, "60 GHz Second Harmonic Power Amplifiers in 130-nm CMOS", *Proceedings of the 8th Swedish System-on-Chip Conference*, May 5–6, 2008, Södertuna slott Sweden.

Vara Prasad Goluguri, Johan Wernehag, Henrik Sjöland and Niklas Troedsson, "A Quad-Core 130-nm CMOS 57-64 GHz VCO", *Proceedings from the 9th GigaHertz Symposium*, March 5–6, 2008, Göteborg Sweden.

Andreas Axholt, Filip Oredsson, Tony Petersson, Johan Wernehag and Henrik Sjöland, "A Low-Power Class-D Audio Amplifier in 0.35μm CMOS", *Proceedings of the 25th Norchip Conference*, pp. 1–4, November 19–20, 2007, Aalborg Denmark.

Johan Wernehag, Henrik Sjöland, "Measurement Considerations for a Phased Array Beamforming Transmitter", *RF Measurement Technology Conference*, September 11–12, 2007, Gävle Sweden.

Johan Wernehag, Henrik Sjöland, "Second Harmonic 60-GHz Power Amplifiers in 130-nm CMOS", *PhD Research in Microelectronics and Electronics*, pp. 149–152, July 2–5, 2007, Bordeaux France.

Johan Wernehag, Henrik Sjöland, "Beamforming Transmitter IC at 8-GHz in 130-nm CMOS", *Proceedings of the 7th Swedish System-on-Chip Conference*, May 14–15, 2007, Fiskebäckskil Sweden.

Johan Wernehag, Henrik Sjöland, "Automotive Radar Transmitter at 24-GHz with Digital Beam Steering in 130-nm CMOS", *Proceedings of the 6th Swedish System-on-Chip Conference*, May 4–5, 2006, Kolmården Sweden.

Niklas Troedsson, Johan Wernehag, Henrik Sjöland, "Differential Measurement and Parameter Extraction of Symmetrical Inductors", *Proceedings of the 23th Norchip Conference*, pp. 289–292, November 21–22, 2005, Oulu Finland.

Johan Wernehag, Henrik Sjöland, "A 90 nm CMOS 10 GHz Beam Forming Transmitter", *International Symposium on Signals, Circuits and System*, Volume I, pp. 375–378, July 14–15, 2005, Iasi Romania.

Johan Wernehag, Henrik Sjöland, "A Comparison of Two 10 GHz Beam Forming Transmitters, in 90 nm and 130 nm CMOS", *Proceedings of the 5th Swedish System-on-Chip Conference*, April 18–19, 2005, Tammsvik Sweden.

Johan Wernehag, Henrik Sjöland, "Low Power Radio Transmitter", *Proceedings from the Seventh GigaHertz Symposium*, November 4–5, 2003, Linköping Sweden.

Published Pedagogical Papers and Teaching Material: The author's interest in undergraduate teaching at the department has resulted in a new exercise material in the course *Analog Electronics*, mandatory for the junior Electrical Engineering classes and a couple of pedagogical publications presented at conferences.

Martin Anderson, Johan Wernehag, Henrik Sjöland, "Teaching Top Down Design of Analog/Mixed Signal ICs Through Design Projects", *Proceedings 37th Annual Conference Frontiers in Education*, pp. T1C-1–T1C-4, October 10–13, 2007, Milwaukee, WI, USA.

Martin Anderson, Johan Wernehag and Henrik Sjöland, “IC-Project and Verification Course: Teaching Top Down Analog/Mixed Signal Design”, *Proceedings of the 7th Swedish System-on-Chip Conference*, May 14–15, 2007, Fiskebäckskil Sweden.

Andreas Axholt, Tony Pettersson, Filip Oredsson, Johan Wernehag, and Henrik Sjöland, “IC-Project and Verification Course: An Integrated Class-D Audio Amplifier”, *Proceedings of the 7th Swedish System-on-Chip Conference*, May 14–15, 2007, Fiskebäckskil Sweden.

Martin Anderson, Johan Wernehag, Roland Strandberg, *Analog Electronics, Exercises*, Department of Electroscience, Lund University, P.O. Box 118, SE-221 00 Lund, Sweden 2003.

Acknowledgment

During my time as a postgraduate student I have met many people that have contributed in one way or another to this thesis, but to mention them all here is impractical and I restrict myself to those that have had substantial impact on this work.

I have been fortunate to have a very competent and enthusiastic supervisor, Professor Henrik Sjöland, who has inspired my work from master thesis to this doctoral thesis. His experience and knowledge in analog- and RF design cannot be overestimated.

I would like to thank all the colleagues at the department for a nice and relaxed atmosphere and some fruitful discussions in various fields, in particular the colleagues in the analog/RF group whom it have been nice to 'waist' time with, without mentioning names I want you all to feel included. The cooperation within the group has also, in a few cases, resulted in co-authored papers: Niklas Troedsson, Martin Anderson, and Andreas Axholt, thanks for bearing with me in the pursuit for perfection. A special thanks to Fredrik Tillman, for being there to discuss everything from process aspects to architectural considerations in the beginning of my PhD studies. I would also like to thank all the different people that I have been traveling with to different conferences, their companionship have made the trips more enjoyable: Peter Sjöblom, Henrik Sjöland, Markus Törmänen, and Andreas Axholt. I would also like to express my gratitude to all the anonymous conference organizers for choosing such nice locations.

I am very thankful for the help and support from the technical staff and administration at the department over the years. A special thanks to Erik Jonsson for quick computer support '24–7', Stefan Molund for Cadence support and making sure that the simulation tools were working, Lars Hedenstjerna for all kinds of mechanical assistance, such as bonding the chips, and last Pia Bruhn who has helped me with all different kinds of administration tasks and travel arrangements. A special thanks goes to Lars Olsson for proofreading this thesis.

Last but not least I would like to thank my fiancée, Christina Larsson, for her love and support, and excepting my sometimes distracted mind over the last years. Now I am back! To my parents, Rolf and Marianne Wernehag, thanks for believing in me throughout the years, always supporting me in the pursuit for higher education. My sisters, Camilla Wernehag, it was an inspiration and challenge trying to follow your results in school. In later days always good to talk to about everything, thanks for being there and listening.

I am grateful for the financial support from the Swedish Agency for Innovation Systems (Vinnova) and Competence Center for Circuit Design (CCCD).

Johan Wernehag

...I knew you had to be tough or die...
"A Boy Named Sue" — Johnny Cash

List of Acronyms

AC	Alternating Current
ACC	Adaptive Cruise Control
AICC	Adaptive Intelligent Cruise Control
AF	Array Factor
ADC	Analog-to-Digital Converter
ADS	Advanced Design System, a simulation software from Agilent
BB	Baseband
BPF	Band Pass Filter
BSF	Band Select Filter
BSIM	Berkeley Short-channel IGFET Model
CG	Conversion Gain
CLM	Channel Length Modulation
CMOS	Complementary Metal Oxide Semiconductor
CP	Compression Point
CSF	Channel Select Filter
DC	Direct Current
DI	Differential Input
DIBL	Drain Induced Barrier Lowering
DUT	Device Under Test
EBT	Ericsson Technology Licensing
EIRP	Equivalent Isotropic Radiated Power
EM	Electro Magnetic
ETSI	European Telecommunication Standards Insitute
F	Noise Factor
FCC	Federal Communication Commission
FDD	Frequency Division Duplexing
FET	Field Effect Transistor
FoM	Figure of Merit

FoV	Field of View
FS	Frequency Synthesizer
Gbps	Giga bit per second
GSG	Ground Signal Ground
GSM	Global System for Mobile Communications
HPBW	Half Power Beam Width
IBM	International Business Machines
IC	Integrated Circuit
IEEE	Institute of Electrical and Electronics Engineers
IF	Intermediate Frequency
IGFET	Insulated Gate Field Effect Transistor
IIP_3	3rd order Input referred Intercept Point
IP	Intercept Point
IRF	Image Reject Filter
IRR	Image Rejection Ratio
ISS	Impedance Standard Substrate
ITRS	International Technology Roadmap for Semiconductors
LNA	Low Noise Amplifier
LO	Local Oscillator
LPF	Low Pass Filter
MIM	Metal-Insulator-Metal
MIMO	Multiple Input Multiple Output
MN	Matching Network
MOS	Metal Oxide Semiconductor
NF	Noise Figure
nMOS	N-channel MOS transistor
OFDM	Orthogonal Frequency-Division Multiplexing
OIP_3	3rd order Output referred Intercept Point
PA	Power Amplifier
PAE	Power-Added Efficiency

PCB	Printed Circuit Board
PLL	Phase Locked Loop
pMOS	P-channel MOS transistor
QVCO	Quadrature Voltage Controlled Oscillator
RF	Radio Frequency
RFC	Radio Frequency Choke
RFIC	Radio Frequency Integrated Circuit
rms	Root Mean Square
SGS	Signal Ground Signal
SHM	Subharmonic Mixer
SI	Single-ended Input
SiGe	Silicon-Germanium
SiO_2	Silicon dioxide
SNR	Signal-to-Noise Ratio
SoC	System-on-Chip
SpectreRF	Spectre Radio Frequency, a simulation software from Cadence
SRR	Short Range Radar
SSB	Single Sideband
SWR	Standing Wave Ratio
TDD	Time Division Duplexing
TL	Transmission Line
Tx	Transmitter
UMC	United Microelectronics Corporation
UMTS	Universal Mobile Telecommunication System
US	United States
UWB	Ultra WideBand
VCO	Voltage Controlled Oscillator
VNA	Vector Network Analyzer
WLAN	Wireless Local Area Network
WPAN	Wireless Personal Area Network

List of Symbols

α	(rad) Conduction angle of an amplifier *or* charge distribution profile
A_{em}	Maximum effective area of an antenna
C_{DC}	(F) DC blocking capacitance
C'_{ov}	(F/m) Capacitance per unit length from gate to source/drain
C'_{ox}	(F/m^2) Capacitance per unit area
C_{tg}	(F) Capacitance from trace-to-ground of an inductor
C_{tt}	(F) Capacitance from trace-to-trace of an inductor
$\mathcal{E}_{sat}$	(V/m) Critical field strength for velocity saturation
D	Duty cycle
D	Directivity of an antenna
ϵ	Amplitude error in an image reject architecture
ϵ_x	(Vs/Am) Permittivity of material x
η	(i) The ratio of C_{max} to C_{min} in a varactor *or* (ii) Degree of non-saturation
η_D	Drain efficiency
f_c	(Hz) Carrier frequency of a radio signal *or* (Hz) Corner frequency between white and 1/f-noise
f_{sr}	(Hz) Self-resonance frequency
f_T	(Hz) Transition frequency of a transistor
f_x	(Hz) Frequency of the signal called x
γ	($\sqrt{\text{V}}$) The body effect coefficient
I+	In phase positive signal
I−	In phase negative signal
$\Im m$	Imaginary part
k	(VC/K) Boltzmann's constant
λ	(m) Wavelength of a electromagnetic signal
l_p	(m) Shortening of the channel length due to high V_{DS}
L_D	(m) The source and drain under-diffusion in a MOS transistor

L_s	(H) Series inductance
μ	(i) (m^2/Vs) Mobility of electrons *or* (ii) Mean value of a data set
n	Number of turns of an inductor
N_A	(m^{-3}) Acceptor dopant concentration
N_D	(m^{-3}) Donor dopant concentration
ϕ_0	(V) Surface potential of two-terminal MOS structure in strong inversion
ϕ_F	(V) Fermi potential
ϕ_t	(V) Thermal voltage, equal to 26 mV in room temperature
ψ_s	(V) Surface potential in the MOS channel
$\Re e$	Real part
R_s	(Ω) Series resistance
R_{sh}	(Ω) Effective resistance of an inductor shield
$R_\square$	(Ω) Resistance per square of a material
ω_0	(rad/s) Angular resonance frequency
ω_x	(rad/s) Angular frequency of the signal called x
σ	Standard deviation of a data set
θ	($^\circ$) Phase error in an image reject architecture
θ_b	(V^{-1}) Mobility degradation parameter due to vertical field in a MOSFET
P_x	(W) Power of signal x
Q	Quality factor
Q+	Quadrature phase positive signal
Q−	Quadrature phase negative signal
v_d	(m/s) Average velocity of charge carriers in the channel
$V_{DS}^{'}$	(V) The value of V_{DS} at onset of saturation
V_{FB}	(V) The flat band potential
V_{knee}	(V) Voltage at the output of an amplifier in saturation
V_T	(V) Threshold voltage of a transistor
q	(C) The negative charge of a single electron

Introduction

Introduction

Chapter 1

Motivation and Research Contributions

1.1 Motivation

There has always been a need for people to communicate and in many independent civilizations ways to communicate over long distances have evolved. The native Americans as well as Australian aboriginals communicated by smoke puffs. The aboriginals also communicated with different sounds, while people in woodland areas often used drums to communicate.

In China 500 BC a large canal construction started and along the canals watchtowers were built [7]. In these towers light sources was placed and by exposing and de-exposing the light in the direction of the next watchtower they were able to send messages down the canal in a 'Morse-like' fashion. This helped the emperor to communicate and control his large empire. This technique to communicate was still in use during the second world war on vessels at sea.

In the beginning of the 20th century wireless radio communication began with Guglielmo Marconi successful transmission of radio signals across the Atlantic Ocean [8, 9]. Since then radio communication development has been intense, and still is. The challenges today differ from the ones encountered by the pioneers, like Marconi and Popov, but still challenges remain and research in this area is therefore highly motivated. The existing public wireless communication standards like GSM, UMTS, Bluetooth etc. put high demands on the technology, especially at the time when they were introduced. The coming standards will also do this, as they too are trying to get the highest possible performance out of the system. But even though the performance requirements are tough, low cost is the main driving force.

Most of the building blocks in consumer electronics today work in the digital domain, and when treating digital signals the dominating integrated circuit (IC) technology is complementary metal oxide semiconductor (CMOS). A major part of the cost of consumer electronics is associated with assembling the different circuits. By implementing the analog and radio frequency parts in CMOS together with the digital circuits, the number of chips can be reduced. The ultimate reduction leads to a sin-

gle chip solution, usually called a system-on-chip (SoC). Another advantage of using the dominant technology is that the price per chip area is cheap in mass production leading to a more cost-efficient solution.

The possibility to integrate analog and digital circuits motivates a continued research of CMOS radio frequency circuits. Already today CMOS is successfully used in many systems, but demands on higher data rates will force some future systems to be more wideband. These larger bandwidths are available at high frequencies, e. g. at 60 GHz. The question is if CMOS will then still be a good choice of technology?

The decreasing supply voltage in sub-micrometer CMOS technology nodes makes the implementation of power amplifiers particularly difficult, therefore the focus in this thesis has been on power amplifiers. At high frequencies, a beamforming transmitter architecture is possible since the phased array antennas decrease in size. The demand on output power from the power amplifier decreases in a beamforming transmitter, thanks to increased antenna directivity and the possibility to use multiple power amplifiers in parallel.

1.2 Research Contribution

The main contributions are summarized in this section.

- ♢ A power amplifier with a full 360° range phase control has been designed and verified by measurements. The output phase control is achieved by summation of weighted quadrature signals. The weighting is performed by means of binary weighted transistor banks. An array of such power amplifiers can be put on a single chip, were the output phase of each of the power amplifiers is steerable individually. The different phases of the power amplifiers can be set by a digital control word, clocked in by a shift register.

- ♢ An antenna array of dipole antenna element patches has been simulated to assess the suitability of such an antenna array in an integrated application, using power amplifiers with differential outputs.

- ♢ Two different frequency doubling power amplifier architectures at 60 GHz have been implemented and measured, showing that the concept of frequency conversion at the antenna interface is possible. The benefits are larger tuning range for the on-chip voltage controlled oscillator, since it can operate at half the output frequency, and less losses thanks to smaller parasitics.

- ♢ A passive subharmonic mixer with differential local oscillator at 15 GHz and radio frequency at 30 GHz has been investigated and taped-out. The new topology uses a filter between the two mixer cores. As a result mixing is accomplish with a differential local oscillator instead of a quadrature one.

  In a beamforming transmitter with multiple mixers it is beneficial to have a differential local oscillator, both regarding area and crosstalk. A combination of a

frequency doubling power amplifier and a subharmonic mixer further increases the tuning range of the local oscillator and reduces the losses and sensitivity to parasitics.

◇ An active polyphase filter was analyzed and analytic equations for the voltage gain and phase transfer function of a loaded stage have been derived. The quadrature phase error over wide bandwidth and for large input differential phase error has been examined, showing good performance. A thorough evaluation of robustness against process parameter spread, voltage variation, and temperature has been conducted.

Chapter 2

Radio Transceiver Fundamentals

2.1 What is a Radio Transceiver?

A transceiver consists of two parts; a transmitter and a receiver. The word transceiver is a fusion of these two words, *trans*mitter and re*ceiver*. In the transmitter, the information to be transmitted is converted to a modulated radio frequency (RF) signal, which is amplified and fed to an antenna. The receiver does the opposite, it receives the RF signal through an antenna, amplifies, downconverts and demodulates it, and if everything is successful the original information is then retrieved. In Fig. 1 this is illustrated in a simple block diagram. The figure is meant as an illustration, and there are other ways this can be implemented, for instance the modulation and demodulation can also take place in the baseband (BB) regime.

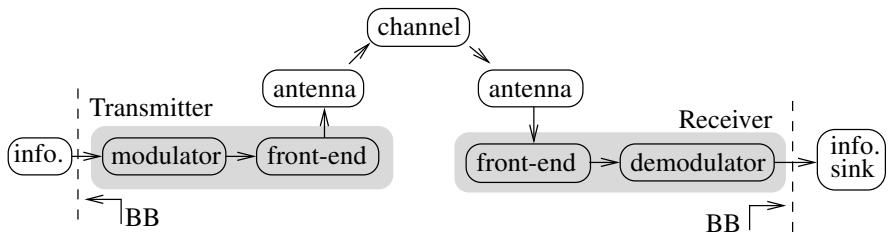


Figure 1: A simple block diagram of a radio transmission

2.2 Receiver

This section contains a brief discussion of receiver architectures. A more detailed description can be found in e. g. [10, 11].

The main objective of the receiver is to filter out and amplify the channel containing the information, and also to downconvert it to lower frequencies. This is a challenging task, since the radio frequency spectrum is full of interfering signals, some of which are close in frequency (in-band) and others more distant (out-of-band). The

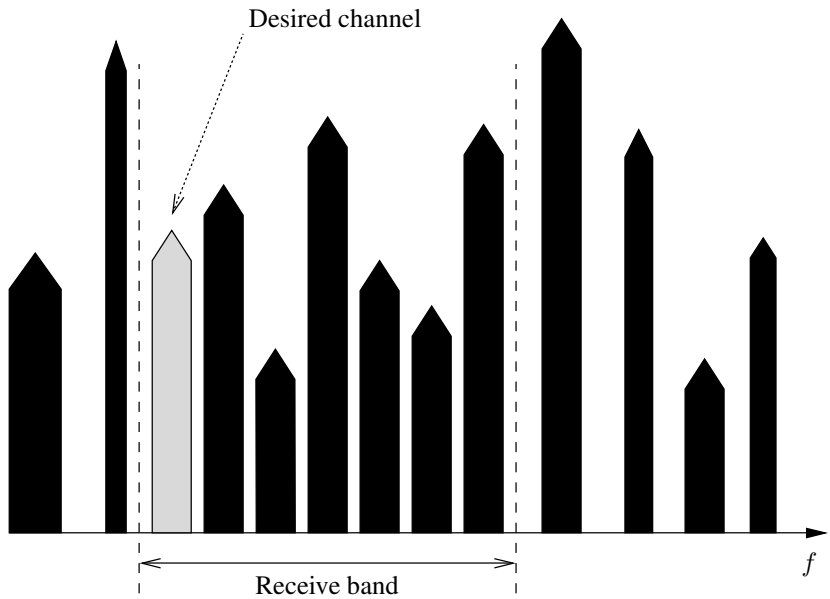


Figure 2: A typical radio spectrum that a receiver must be able to cope with

signal level of the different interferers can be much higher than that of the desired signal, see Fig. 2.

The most common architectures like the heterodyne receiver[1], the image reject receiver, and the direct conversion receiver will be described in the following sections. Also a short section on how to calculate the overall receiver linearity and sensitivity is included.

2.2.1 Receiver Architectures

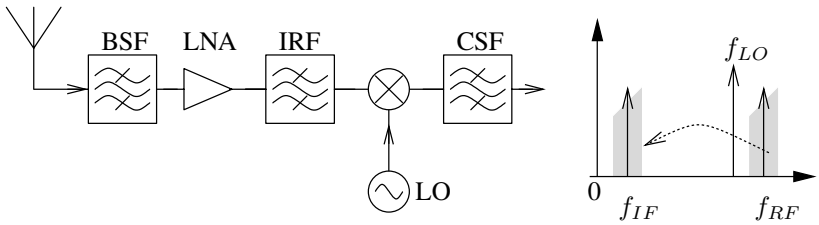
Heterodyne Receiver

The heterodyne receiver is perhaps the most common architecture[2] [15, 16].

The heterodyne receiver architecture (Fig. 3) has a band select filter (BSF) close

[1]The word heterodyne is derived from the Greek roots hetero-, 'different', and -dyne, 'power' [12].

[2]Their are two types of of heterodyne receivers, firstly the heterodyne which historically had one down-conversion to a rather low intermediate frequency (IF), in the audible frequency region. The heterodyne principle was invented by Reginald Aubrey Fessenden on Chrismas Eve 1906. He transmitted the first audio radio broadcast in history from Brant Rock, Massachusetts. The second type is the supersonic heterodyne receiver, or as it is more commonly called, the superheterodyne receiver. It also has a single down-conversion, but to a higher IF. It was patented by Edwin Howard Armstrong in 1918. Nowadays generally no distinction is made between the heterodyne and the superheterodyne, and that practice will be followed in this thesis as well [13, 14].



(a) Architecture of the heterodyne receiver

(b) Down-conversion in the frequency domain

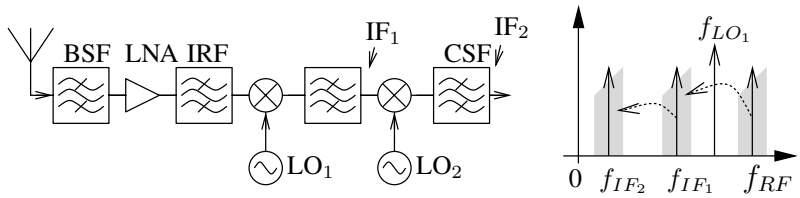
Figure 3: (a) The architecture of a heterodyne receiver. (b) Down-conversion in the frequency domain

to the antenna, which provides attenuation of out-of-band interferers. The radio frequency (RF) signal is then amplified by a low noise amplifier (LNA) followed by a second band pass filter (BPF). This filter is sometimes called image reject filter (IRF). These two filters can together provide an attenuation of 100 dB at frequencies far from the channel. The requirements on these filter are often tough and therefore they typically have to be implemented off-chip, which makes them bulky and expensive [9–11, 15, 17]. The second filter precedes the mixer, where the RF signal is downconverted to an intermediate frequency (IF) by mixing with a local oscillator (LO) signal. After the mixer, a channel select filter (CSF) filters out the channel to be demodulated. The selected frequency is at $f_{IF} = |f_{RF} - f_{LO}|$ in Fig. 3(b).

Sometimes the frequency down-conversion is performed in two steps i. e. with two separate mixer stages, resulting in the double-conversion heterodyne architecture, see Fig. 4. There are a few advantages with a double conversion, the gain and filtering can be distributed over more frequencies, which is beneficial when it comes to linearity, stability, and noise trade-offs, see Section 2.2.2. The gain being distributed over several frequencies reduces the risk for the receiver to self oscillate. The drawback is that the architecture is more complicated, e. g. it requires two different LOs.

The main problem of the heterodyne architecture is that the mixer is unable to separate the wanted channel from the image channel, as depicted in Fig. 5. The image is located on the opposite side of the f_{LO} compared to f_{RF}, and the distance between RF signal and image is $2 \cdot f_{IF}$. If the image is not rejected/suppressed before the mixer, it can not be distinguished from the wanted signal, since it will be downconverted to the same IF. The IRF is therefore placed before the mixer to take care of this issue.

The LNA has to be matched to the impedance of the off-chip filters, often 50 Ω, and the same goes for the input of the downconverter. The problem with the image leads to a trade-off in choosing f_{IF}. A high f_{IF} results in large spacing between the RF and the image, simplifying the image reject filter, but a high f_{IF} results in more difficulties with the CSF, and vice versa [9, 10, 17]. The double-conversion heterodyne receiver has an advantage here. The f_{IF_1} can be high leading to relaxed demands on



(a) Architecture of the double-conversion heterodyne receiver

(b) Down-conversion in the frequency domain

Figure 4: (a) The architecture of the double-conversion heterodyne receiver. (b) Down-conversion in the frequency domain

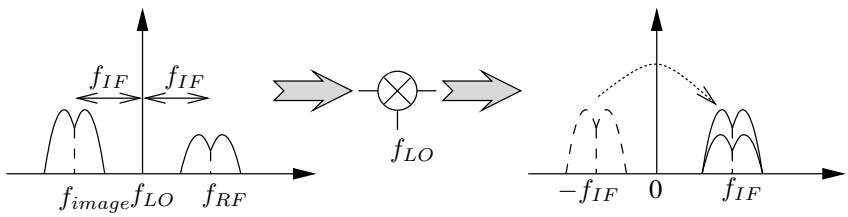


Figure 5: The image frequency problem

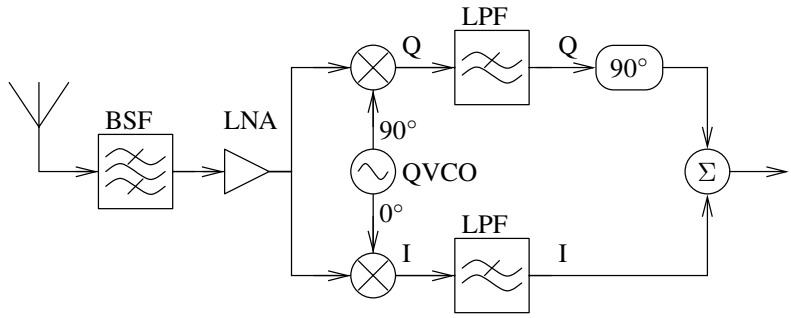
the IRF. At the same time the second IF can be low, and also the demands on CSF can be relaxed. It should also be noted, however, that one additional filter is needed in the double-conversion receiver.

Image Rejection Receiver

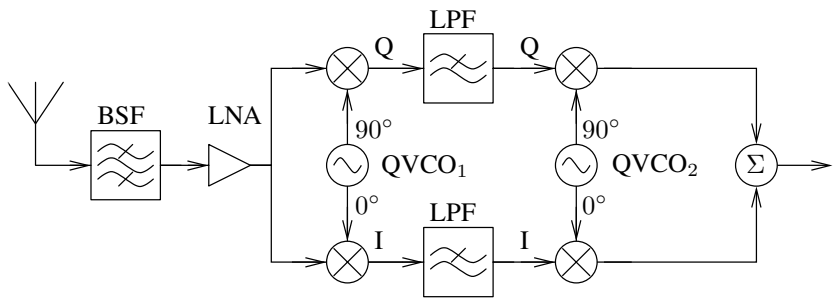
As described in the previous section, the image frequency is a problem in the heterodyne architectures. To deal with this issue, a frequency down-conversion mixer structure which suppresses the image signal can be used. There are two main image reject topologies, the Weaver[3] and the Hartley[4], see Fig 6. Another possibility is to keep the I and Q paths separate in the analog part and perform the image rejection in the digital baseband, but this will not be further treated in this thesis. In the image rejection architectures it is very important that the phase and amplitude errors are minimized, since errors result in a reduced image rejection ratio (IRR) [20]. An equation (1) for IRR is derived in [10]:

[3]The Weaver topology was first published in 1956 [18], the title of the paper were "A Third Method of Generation and Detection of Single-sideband Signals".

[4]Ralph Hartley filed a patent for his idea in April 1928 [19], he then called it a "Single-sideband Modulator".



(a) The Hartley image reject architecture



(b) The Weaver image reject architecture

Figure 6: The two most common image reject receivers, (a) Hartley and (b) Weaver

$$IRR = \frac{P_{RF}}{P_{image}} = \frac{1 + 2(1+\epsilon)\cos\theta + (1+\epsilon)^2}{1 - 2(1+\epsilon)\cos\theta + (1+\epsilon)^2} \approx \left.\frac{4}{\epsilon^2 + \theta^2}\right|_{\epsilon \text{ and } \theta \text{ small}} \quad (1)$$

where ϵ is the relative amplitude error and θ is the phase error (in radians).

To get a 26 dB IRR, the structures require a maximum of 5% amplitude mismatch between the two branches and 5° phase error, and to get a more desirable IRR of, say 60 dB, 0.1% and 0.1° would be needed, requiring calibration circuitry. In the literature [21, 22] two different CMOS circuits operating at 2.4 GHz are reported with an IRR of 45 dB and 30 dB, respectively. In [23, 24] calibration circuits are used in a Weaver architecture, and IRR of 57 and 59 dB are reported at 2 GHz and 1.8 GHz, respectively.

In the image reject architectures, the RF signal first passes a BSF and an LNA, as in a heterodyne architecture. The signal is then downconverted in two mixers with a quadrature LO signal, followed by a low pass filters (LPF). In the Hartley architecture a 90 degree phase shift is introduced in one of the signal paths, usually implemented

as a $+45$ degrees in one path and a -45 degrees in the other. The two paths are then summed together and the downconverted image signals in the two branches ideally cancel each other, while the desired signals add in phase. In the Weaver architecture, the 90° phase shift is performed using a second set of mixers also fed with quadrature LO signals.

To examine how the image cancellation occurs, a brief walk-through in the Hartley case will follow. For a more detailed description see [10]. Assume that the RF and image signals are sinusoids and that $f_{RF} > f_{LO}$. In the upper path in Fig. 6(a), the RF and image signal are multiplied with another sinusoid. After the LPF, there are two cosines left with frequency $f_{RF} - f_{LO} = f_{IF}^{RF}$ and $f_{image} - f_{LO} = -f_{IF,}^{image}$ since cosine is an even function, negative and positive arguments give equal results (2a). These cosine signals are phase shifted 90 degrees, which means that the cosine is replaced by a sine (2b),

$$\cos\left(f_{IF}^{RF}\right) + \cos\left(-f_{IF}^{image}\right) = \cos\left(f_{IF}^{RF}\right) + \cos\left(f_{IF}^{image}\right) \tag{2a}$$

$$\text{(2a)} \; + \; 90^\circ \,\text{phase shift} \; \longrightarrow \sin\left(f_{IF}^{RF}\right) + \sin\left(f_{IF}^{image}\right). \tag{2b}$$

In the lower path, the input RF and image signals are multiplied by a cosine and after the LPF two sinusoids are left (3),

$$\sin\left(f_{IF}^{RF}\right) + \sin\left(-f_{IF}^{image}\right) = \sin\left(f_{IF}^{RF}\right) - \sin\left(f_{IF}^{image}\right). \tag{3}$$

After the summation of the right hand side of (2b) and (3) it is now clear that only the RF signal contributes to the output.

Direct Conversion Receiver

The direct conversion receiver is also called the homodyne[5] receiver or zero-IF receiver. The idea of the homodyne receiver is to downconvert the RF signal directly to baseband. This is accomplished by mixing the RF signal with an LO signal of the same frequency as the RF carrier, f_{RF}. The center of the band is then translated to a zero IF (4), see Fig. 7. The mixer is followed by a LPF that performs the channel select filtering.

$$\omega_{IF} = \omega_{LO} - \omega_{RF} = 0 \tag{4}$$

The requirements on the adjacent channel interference sets the minimum sharpness and thus the order of the low pass filter [10]. With the rather low corner frequency of the filter, it can typically be integrated as an active filter [27]. The main advantage of the direct conversion receiver is its simplicity, see Fig. 8, giving the architecture good ability for integration and a potentially low power consumption.

[5]The homodyne receiver dates back to 1924 when it first was published by Colebrook [25, 26].

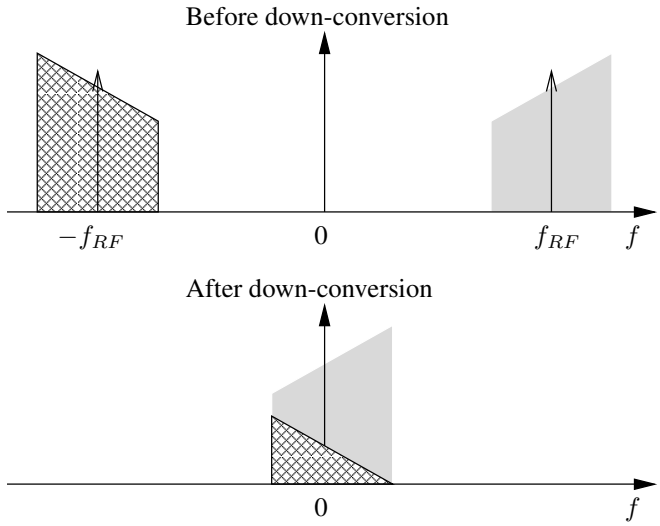


Figure 7: Spectrum before and after down-conversion with a direct conversion receiver

To be able to recover the phase information of the signal, complex mixing has to be applied, which makes it similar to the image reject architectures see Fig. 6 and 8. A difficulty with the heterodyne receiver and other architectures with a non zero-IF is the image frequency. With a zero-IF architecture the signal is its own image, thus avoiding the problem of image signals being stronger than the signal to receive, see Fig. 7.

For the complex mixing, a quadrature LO is needed, which can be generated in a number of different ways. A quadrature voltage controlled oscillator (QVCO) can be used. This oscillator has two resonance tanks, which means that it requires two inductors and hence occupies a large die area [28]. Another approach is to use a VCO followed by a passive or active polyphase filter, which converts the differential signal to a quadrature one [5,29–31]. The polyphase filter can be made small, but the passive polyphase filter attenuates the signal and may therefore require amplifying stages. The active polyphase filter can on the other hand be power hungry. A differential VCO signal can be frequency divided, obtaining a quadrature signal at half the VCO frequency. This approach is common in the low GHz range (GSM, EDGE, etc.). The frequency dividers are small but can nevertheless consume large current. A higher frequency makes the VCO smaller and the separation in VCO and power amplifier frequency makes it more robust against pulling.

The drawback of the direct conversion receiver is that the architecture is sensitive to DC offsets, flicker noise ($1/f$-noise), and even-order non-linearities, which all cause disturbances at, or close to, DC, the center frequency of the downconverted signal. DC

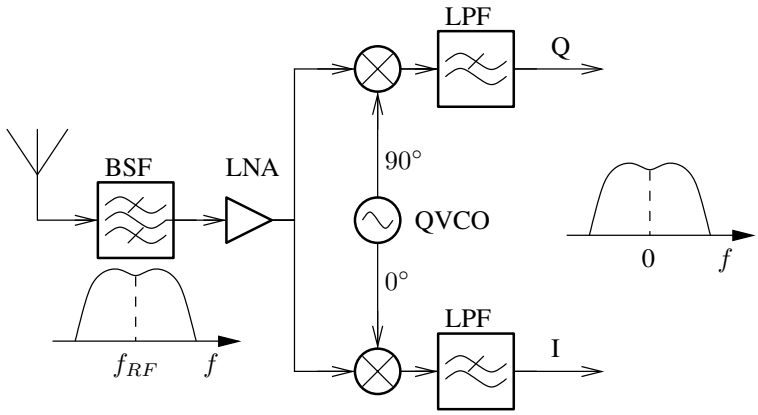


Figure 8: A direct conversion receiver with complex mixing

offsets can be caused by self-mixing, i. e. the LO can leak to the LNA input or the RF mixer input, thus being mixed with itself resulting in a DC component. The leaking can be through substrate coupling, capacitive coupling, or if the LO is off-chip bond wire coupling. Strong interferers at the LNA input can also leak to the LO input of the mixer and be mixed by themselves, which is illustrated in Fig. 9(b). The gain after the mixer is typically about 70 dB, [9], so even a rather small DC offset can easily saturate the following amplifiers and analog-to-digital converter (ADC). The flicker noise increases with decreasing frequency and can therefore corrupt the information in a direct conversion receiver. The flicker noise is also inversely proportional to the area of the device, thus there is a trade-off between bandwidth and $1/f$-noise. The I/Q mismatch is not as big of a problem in a direct conversion receiver, since the image is never stronger than the signal to receive.

As mentioned even-order distortion can also be a problem. If two strong interferers, f_1 and f_2, are located close together in the receive band, they can be mixed by an even-order non-linearity in the mixer and end up close to DC, $f_2 - f_1 \approx 0$. This signal can then mask the wanted signal, see Fig. 9(c). This can also be caused by a single modulated interferer, containing signal energy at f_1 and f_2 [27].

2.2.2 Receiver Linearity and Noise

Two important performance measures of a receiver are the linearity and noise figure. Noise limits the sensitivity of the receiver, while linearity limits how well the receiver can cope with interferers. The requirements on these parameters can often be calculated from the system specifications. As seen in the previous sections the receiver is composed of several building blocks. It would, therefore, be valuable to know the influence of different blocks on the performance of the complete receiver. This can be

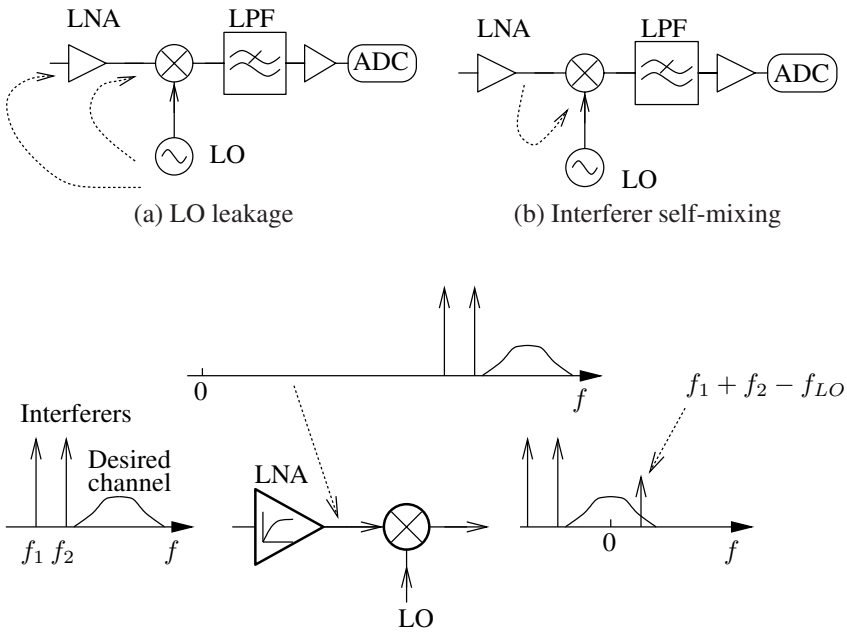


Figure 9: (a) DC offset caused by LO leakage. (b) DC offset by interferer. (c) Even-order distortion in mixer

found in various books, e. g. [10, 11, 17]. With this knowledge, the specification of the different blocks can be decided.

There are many other requirements on a receiver than low noise and high linearity. For instance, it must have sufficient selectivity. Since this thesis is mainly concerned with transmitters, however, the reader is directed to e. g. [10, 32] for a more complete treatment of receivers.

Noise

Assume that the receiver consists of m different blocks, see Fig. 10, where F_x is the noise factor of block x counted from the antenna, $A_{p,x}$ is the available power gain of block x, and $A_{v,x}$ is the voltage gain.

The noise factor of the chain is calculated according to (5) [10, 11, 17, 33], which

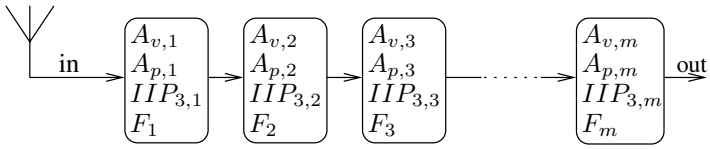


Figure 10: m different blocks building up the receiver

is also known as Friis' formula[6].

$$F_{tot} = 1 + (F_1 - 1) + \frac{F_2 - 1}{A_{p,1}} + \frac{F_3 - 1}{A_{p,1}A_{p,2}} + \cdots + \frac{F_m - 1}{A_{p,1}A_{p,2}...A_{p,m-1}} \tag{5}$$

With adjacent blocks matched to the same impedance, $A_{v,x} = A_{p,x}$.

In an integrated CMOS receiver, however, the blocks are typically mismatched on-chip, and $A_{p,x}$ and $A_{v,x}$ are usually not the same. The equivalent noise of the entire front-end is therefore a more common way to characterize the performance in integrated solutions, but still Friis' formula gives the designer a tool to optimize the front-end for noise performance. High gain and low noise in the first stage is the most important insight. This is seen in all architectures as they all begins with an LNA.

Linearity

The linearity measure most often used is the intercept points (IPs). Those are defined as the points where the extrapolated intermodulation product has the same power as the extrapolated desired signal. This is illustrated in Fig. 11 for the third order intercept point. The magnitude of the third order intermodulation grows with the cube of the input magnitude, seen in Fig. 11 as the slope is 3 for the third order intermodulation due to the log scale and 1 for the desired signal until the receiver goes into compression. If the power level is referred to the input it is called input IP (IIP), correspondingly when it is referred to the output is is called output IP (OIP). For large input signals the receiver saturates, it enters compression. Since this happens before the IP is reached, the IP is defined as an extrapolation from small input signals. The odd-order intermodulation products may be in the wanted frequency band, e. g. $2f_1 - f_2$ for the third order. Also higher order intermodulation exist, but the third is typically the most troublesome. The even-order intermodulation are close to DC or at $2f_{RF}$ etc. The low frequency intermodulation could cause problems in direct conversion receivers, which therefore typically have stringent requirements on IP_2.

As the input signal gets larger the output of the amplifier can not follow the input signal linearly, the amplifier saturates. A measure of the deviation from linear operation is the 1-dB compression point, which is the point where the linear extrapolation

[6]Harald T. Friis was born in Denmark 1893 and moved to the US in 1919 [34]. He developed the famous formula in 1942 working at Bell Labs [35].

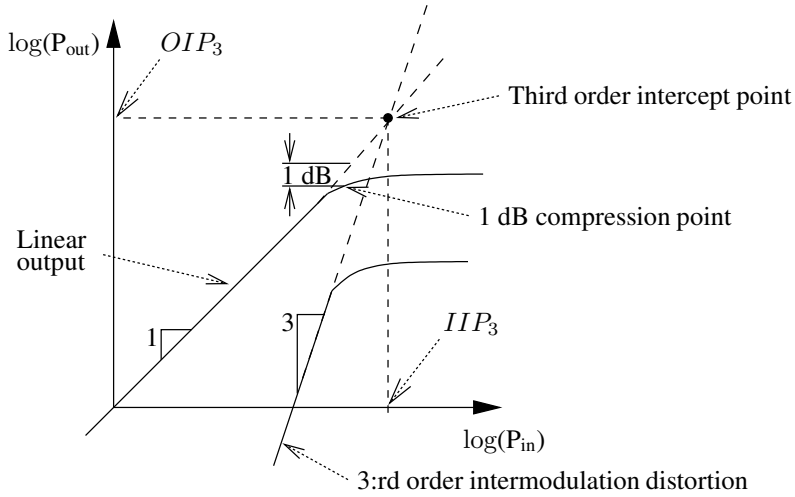


Figure 11: A diagram of the third order intercept point

of, and the actual output differ by 1 dB. The compression point (CP) indicates an upper limit to how large an input signal can be for linear reception. The dynamic range of the receiver is often defined as the distance from the noise floor up to the CP. In amplifiers with weak non-linearity, where the 3rd order non-linearity dominates up to the compression point, there is a 10 dB difference between the 1-dB CP and the IP_3.

When the IP_3 and the gain of the different building blocks are known the total receiver IP_3 can be calculated according to (6) [10, 11, 17].

$$\frac{1}{IIP_{3,tot}^2} = \frac{1}{IIP_{3,1}^2} + \frac{A_{v,1}^2}{IIP_{3,2}^2} + \cdots + \frac{(A_{v,1}A_{v,2}\cdots A_{v,m-1})^2}{IIP_{3,m}^2} \tag{6}$$

As can be seen, the gain of the first stage should be low to get a high IIP_3 , i. e. just the opposite to the noise optimization. Thus it is not possible to simultaneously optimize them both.

2.3 Transmitter

In this section an overview of the most common transmitter architectures is given. The direct conversion, the two-step, and the direct-modulation transmitter are treated. Then the perhaps most challenging block in an integrated CMOS transmitter, the power amplifier (PA), is dealt with in more detail.

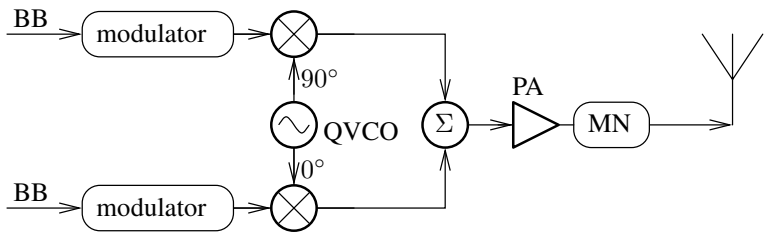


Figure 12: The quadrature version of the direct conversion transmitter

2.3.1 Transmitter Architectures

Direct Conversion Transmitter

In the direct conversion transmitter, the transmitted carrier and the LO are at the same frequency. The frequency up-conversion takes place in a single conversion stage, see Fig. 12, where the quadrature version of the direct conversion transmitter is shown [10, 11, 17, 36]. The simplicity of the transmitter makes it an attractive choice for integration.

The BB signal is modulated and then upconverted to the RF carrier by a mixer driven by the quadrature LO signals. The RF signal is then amplified in the PA to obtain the desired output power. The antenna impedance is transformed by a matching network (MN) so that the impedance seen by the PA is close to its optimum load impedance. The MN can be implemented both on- and off-chip. To avoid harmonics and wideband noise reaching the PA, a BPF is often inserted.

Since the VCO and the PA operate at the same frequency, there is a risk that the strong output signal of the PA can disturb the VCO, this phenomenon is known as injection pulling or injection locking [36, 37]. The coupling from the PA output to the VCO can take place through different paths; the substrate, bond wires if the VCO is off-chip, capacitive coupling, etc. Injection pulling is when the modulated PA output signal pull the VCO frequency back and fourth. Locking is when the disturbance is so strong that the VCO locks to the same frequency as the PA output, see Fig. 13. Disturbance levels of -40 dBc compared to the VCO can create big problems in a radio tranceiver [36].

The problem can be reduced by avoiding operating the VCO and the PA at the same frequency. A different flavor of the direct conversion transmitter has therefore been invented, the offset-LO direct conversion transmitter [15, 36]. In this topology the QVCO signal in Fig. 12 is generated by two LOs with frequencies, f_1 and f_2. These are mixed and filtered by a BPF, where the sum of the frequencies is passed through. The frequency $f_1 + f_2$ is then converted to a quadrature signal and fed to the mixers in Fig. 12. The drawback is that two different LOs are required.

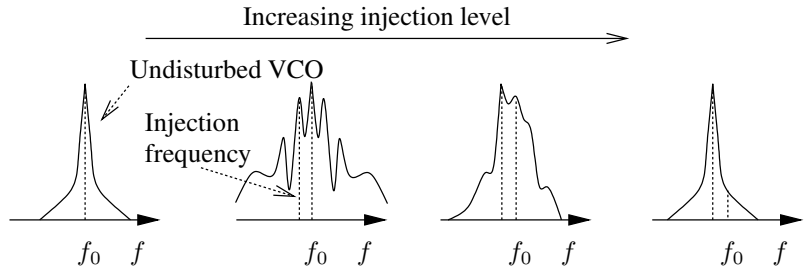


Figure 13: VCO injection pulling and locking

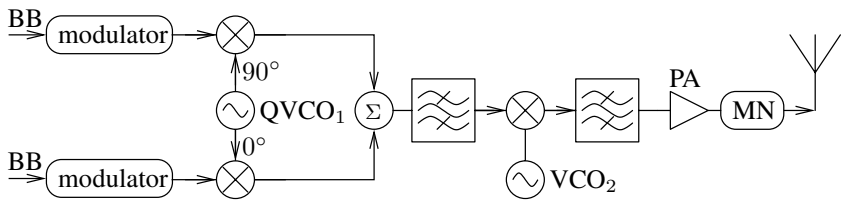


Figure 14: The two-step transmitter

Two-Step Transmitter

The up-conversion can also take place in two steps, like in the heterodyne receiver architecture, see Fig. 14. This is called the two-step transmitter [10]. The BB signal is then first upconverted to an intermediate frequency by quadrature mixers. The I- and Q-paths are then added together and bandpass filtered by an IF filter. The IF signal is then upconverted to RF by a second mixer, resulting in frequencies $f_{RF} = f_{LO,2}+f_{IF}$ and $f_{image} = f_{LO,2} - f_{IF}$. The image frequency must be filtered out before reaching the PA. That is done by a second BPF. If f_{IF} is low the distance in frequency between f_{RF} and f_{image} is small and the requirements on the BPF become tough. With a high IF on the other hand, the quadrature matching of the first QVCO and the IF filter sharpness become more troublesome to implement. This corresponds to the case of a heterodyne receiver, where there is a trade-off between requirements on the IF and the RF filters.

A key advantages of this architecture is that the I and Q mixers operate at low frequencies, compared to the direct conversion transmitters, which makes matching easier. Also the fact that the radio frequency is separated from both the LO frequencies makes injection pulling or locking less severe. The drawback is that it requires two different LOs, just like the offset-LO direct conversion architecture.

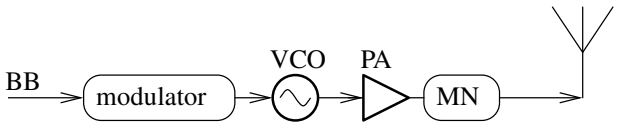


Figure 15: The direct modulation transmitter

Direct-Modulation Transmitter

In a direct-modulation transmitter, the low frequency BB signal is applied to the control voltage of the VCO, thereby directly modulating the frequency of the transmitters output, see Fig. 15. In this way no mixers are needed. This type of transmitter can be used in constant-envelope modulation schemes, such as GSM [38] and Bluetooth [39]. Sometimes also AM modulation is required. This can then be added to the signal after the VCO by e. g. modulating the supply voltage to the power amplifier, or having a variable gain amplifier before it.

The architecture is well suited for integration. Since the VCO produces a signal with low noise no filter is needed before the PA.

The instability of the free running VCO output frequency is a problem. The VCO is therefore often controlled by a phase locked loop (PLL) frequency synthesizer (FS). The FS controls the output frequency and as a result it can be necessary to disable it during transmission so that output frequency follows the modulation instead. Another approach is to compensate for the transfer functions of the different blocks in the FS, as in e. g. [40]. In [40] it is shown that the transfer to the output is low-pass from one injection point and high-pass from another, thus by injecting the control signals at both points an all-pass filter response is obtained and the FS can remain in operation during transmission as well.

2.3.2 Transmitter Requirements

The frequency spectrum is a very valuable asset. Different systems operate in different frequency bands, and one system must not disturb another by emissions in its band. This is accomplished by transmitters with high linearity and low noise. The linearity and noise requirements can, therefore, be derived from system specifications on allowed out-of-band emission level, but also from the in-band requirements. To obtain the specifications of different blocks it is necessary to understand how the different blocks in the transmitter affect the linearity and noise of the whole transmitter. This is described in a number of textbooks, e. g. [10, 11, 17].

Linearity

The spectrum leakage is largely dependent on the non-linearity of the last amplifying stage, i. e. the PA. The other non-linear blocks contribute as well to the total

transmitter non-linearity, but in a well designed transmitter the PA should dominate. The equation for the receiver non-linearity (6) is valid also for transmitters (7), where block m in this case is closest to the antenna.

$$\frac{1}{IIP_{3,tot}^2} = \frac{1}{IIP_{3,1}^2} + \frac{A_{v,1}^2}{IIP_{3,2}^2} + \cdots + \frac{(A_{v,1}A_{v,2}\cdots A_{v,m-1})^2}{IIP_{3,m}^2} \tag{7}$$

In transmitters it is typically more convenient to have the IP_3 output instead of input referred, since the output power is of interest. This is given in (8) [10, 11, 17], where $A_{v,x}$ is the voltage gain of stage x and $OIP_{3,x}$ is the output referred third order intercept point of stage x.

$$\begin{aligned}\frac{1}{OIP_{3,tot}^2} &= \frac{1}{(A_{v,1}\cdots A_{v,m})^2}\left[\frac{1}{IIP_{3,1}^2} + \frac{A_{v,1}^2}{IIP_{3,2}^2} + \cdots + \frac{(A_{v,1}A_{v,2}\cdots A_{v,m-1})^2}{IIP_{3,m}^2}\right] \\ &= \frac{1}{OIP_{3,m}^2} + \frac{1}{A_{v,m}OIP_{3,m-1}^2} + \cdots + \frac{1}{(A_{v,1}A_{v,2}\cdots A_{v,m-1})^2 OIP_{3,1}^2}\end{aligned} \tag{8}$$

As can be seen, to obtain a high transmitter linearity, the last stage (stage m) should have high linearity and gain.

Noise

The noise of the transmitter can degrade the sensitivity of the receiver and it is therefore an important factor in transmitter design. Systems with frequency division duplexing (FDD) are specially vulnerable to transmitter noise since both transmitter and receiver in the same device are operating at the same time and the isolation between them is finite. In time division duplexing (TDD) systems the transmit and receive path are not active at the same time, then the isolation between the blocks and the transmitter noise performance is less critical. Emission regulations then sets requirements on how large noise emissions are allowed in different frequency bands.

The noise at the output of the transmitter consists of a frequency independent noise, called white noise, and a frequency dependent phase noise, see Fig. 16. The phase noise around the RF carrier comes from the VCO (frequency synthesizer) and is transferred to the signal in the frequency up-conversion [33, 41].

2.4 Power Amplifiers

The power amplifier is the most power consuming building block in an RF transceiver, and thus it is a crucial block to optimize, especially for battery powered devices. As shown in Section 2.3.2, the linearity of the transmitter is determined to a large extent by the PA. If the PA is not linear enough, a linearization technique must be employed.

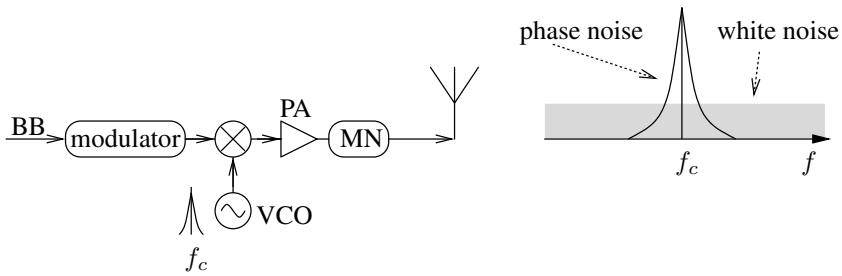


Figure 16: Noise at the output of the transmitter

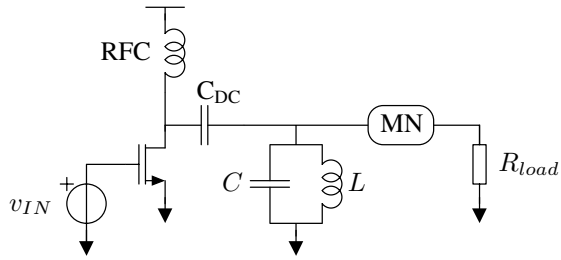


Figure 17: The architecture of class A–C power amplifiers

2.4.1 Power Amplifier Classes

PAs are divided into different classes depending on how they operate. In this section short descriptions of the classes A to F are presented. For further treatment there are many books on power amplifier design, e. g. [11, 42].

Classes A–C

Classes A–C constitute a family of amplifiers using the same circuit structure, see Fig. 17. They differ only in the choice of bias point of the amplifying transistor. The classes are usually called A, AB, B, and C [11, 42, 43].

The DC current is supplied through the RF choke (RFC), whose purpose is to block RF current and let DC trough. The DC blocking capacitance (C_{DC}) has the opposite purpose, it blocks DC current so that no DC power is lost in the MN and load. The resonance tank, which is comprised of L and C, is tuned to the operating frequency to suppress harmonics.

The bias level of the input voltage, v_{IN}, determines in which class the amplifier operates, see Fig. 18. In class A the input voltage is always larger than the threshold voltage (V_T). This means that current will always flow in the transistor. It will never be shut off. In class B the input is above V_T for half of the period, so current will flow

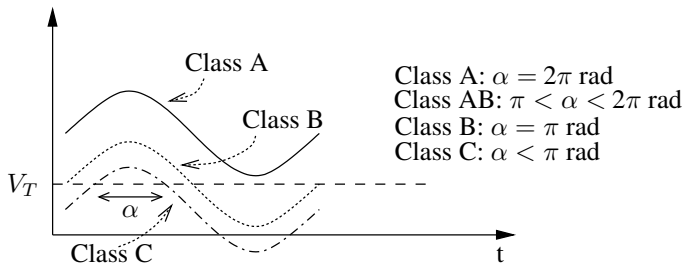


Figure 18: The bias levels for class A–C power amplifiers

in the transistor for half the period. Less current is therefore consumed compared to class A. Class AB is in between class A and B, which means that the transistor is on for more than half but less than the whole period. Finally, in class C the conduction time is less than half the period. It is the conduction time that defines to which class the PA belongs. The conduction time per period is usually called duty cycle (D), defined in (9a). The duty cycle can also be expressed as an angle, called conduction angle, α, defined in (9b).

$$D \equiv \frac{T_{on}}{T_{period}} \tag{9a}$$

$$\alpha \equiv 2\pi \cdot D \tag{9b}$$

The linearity improves with increased bias level. Unfortunately, so does also the power consumption, and the drain efficiency decreases. To illustrate this, a simulation has been performed. A simple class A–C amplifier has been simulated in a 130 nm CMOS process with a supply voltage of 1.2 V. The operating frequency was 2 GHz, which is far below f_T of the transistor. A 50 Ω load was used, all the passive components were ideal, and a resonant tank at the drain of the transistor was tuned to 2 GHz. The threshold voltage of the transistor is 334 mV. The bias voltage at the transistor gate is in class A equal to 767 mV, in class B 334 mV, and in class C 84 mV. The transistor sizes are chosen to achieve maximum efficiency at the 1-dB compression point. In Fig. 19, the amplifier's third order intercept diagram is plotted (left y-axis) together with the drain efficiency (right y-axis). The intermodulation distortion was simulated with a two-tone test, the two tones being placed at 1.9 GHz and 2.1 GHz. The OIP_3 degrades as the transistor gate bias level is decreased, achieving 22 dBm for class A and 5 dBm for class B. The class C amplifier enters class C operation at $\mathrm{P_{in}} = -2$ dBm $\left(\hat{V}_{in} > V_T\right)$. The amplifier is highly non-linear after this point and no extrapolation point for OIP_3 is visible. The drain efficiency on the other hand goes in the opposite direction, class C reaches up to 80% as the output power saturates at

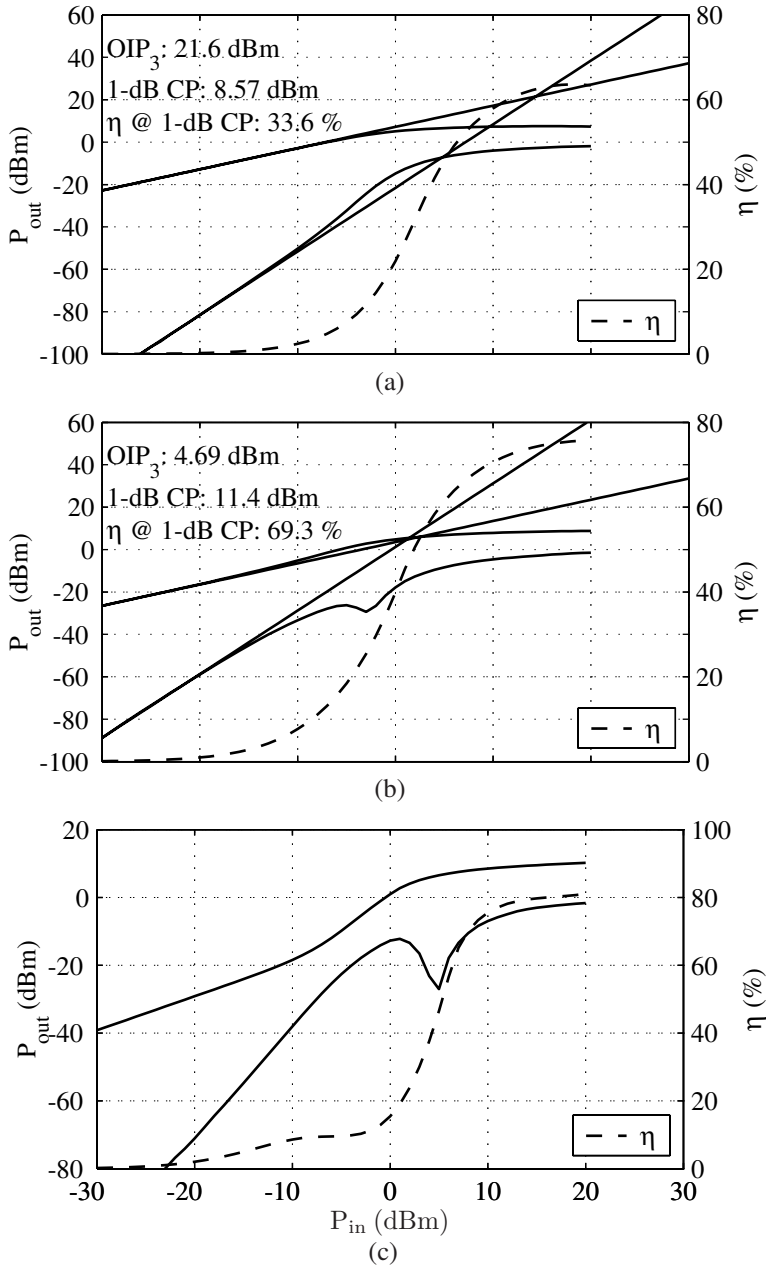


Figure 19: Simulation of class A–C amplifier, drain efficiency and intermodulation distortion. (a) Class A, (b) Class B, and (c) Class C

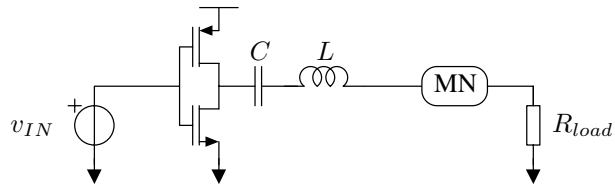


Figure 20: The architecture of the class D amplifier

10 dBm, while class A and B achieves 34% and 69% at their 1-dB compression points, respectively.

Classes D–F

The class D to F amplifiers use the transistors as switches [44]. An ideal switch either has zero voltage across source and drain (closed) or zero current (open), resulting in zero power dissipation. The theoretical maximum efficiency therefore becomes 100%. The class D amplifier basically functions as an inverter, see Fig. 20. When v_{IN} is high the n-channel MOS (nMOS) transistor is on, connecting the output to ground potential, and when the input is low the p-channel MOS (pMOS) is on, connecting the output to the supply potential. A square wave is thereby created at the drain of the devices. The LC series network, tuned to the operating frequency, passes the fundamental current and prevents DC and harmonics of the square wave from reaching the load. The theoretical maximum efficiency of a class D amplifier is 100%, but losses in the switches and finite switching speed will reduce the efficiency that can be obtained in practice [11]. Due to the finite switching speed the class D amplifier should not be used at high frequencies approaching the speed capabilities of the transistor used[7].

The class E PA is shown in Fig. 21. The higher order reactive network at the output of the transistor makes it possible to force the drain voltage to zero with a slope of zero when the transistor turns on. The reactive component values must be chosen with care, see [11,45] for more details and design equations. The power handling capabilities of the class E amplifier is not as high as that of the class D due to larger peak voltage. This increases the risk of breakdown in the transistor if the supply voltage is not reduced, resulting in less output power.

Also, the class F amplifier uses reactive components to shape the waveforms at the transistor drain, see Fig. 22, [11,46,47]. The L_1C_1 resonator is tuned to the fundamental. The parallel resonator comprised of L and C is tuned to the third harmonic,

[7] The switching losses, due to that the transistor conducts current and simultaneously has a non-zero voltage drop between source and drain, will be a larger and larger percentage of the period when approaching f_T and thus decreasing the efficiency.

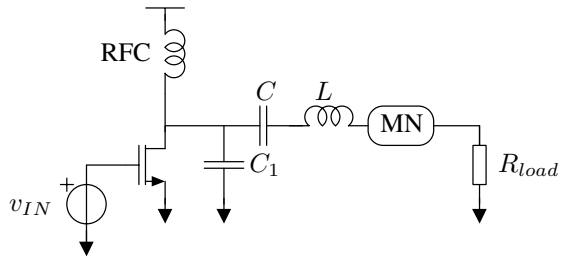


Figure 21: The architecture of the class E amplifier

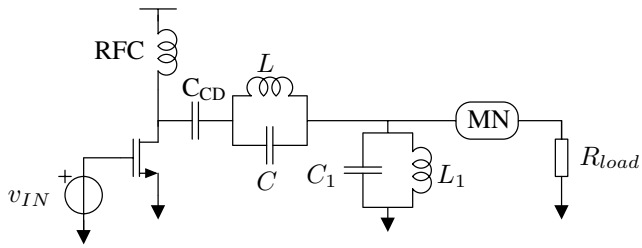


Figure 22: The architecture of the class F amplifier

preventing it from reaching the antenna. There can also be higher order odd harmonic filters attached thereafter if necessary. The odd harmonics then see a high impedance at the drain of the nMOS transistor and the waveform becomes more and more like a square wave, since the even-order harmonics are grounded through the resonance circuits.

In microwave design, the parallel resonance circuit(s) connected in series with the signal in Fig. 22 is usually replaced by a $\lambda/4$ transmission line. The transmission line acts as a high impedance for all odd-order harmonics and shorts the even-order ones.

2.4.2 PA Efficiency

The PA handles the signal with the largest power in the transceiver and its efficiency is therefore very important to the overall power consumption. The efficiency is typically highest near peak output power, but not only the efficiency at maximum power is of importance. Also the efficiency at lower power levels is important, since that is what is used most of the time. The modulation may also have a built-in power variation, the crest factor[8] or peak-to-average ratio. The average output power is always lower than the maximum. The drain efficiency is defined as (10) and will simply be referred to as

[8]The crest factor is the peak signal value divided by the rms value.

efficiency below.

$$\eta_D \equiv \frac{P_{RF}}{P_{DC}} \tag{10}$$

For the classes A–C, the maximum efficiency with a sinusoidal signal is given in (11), which is derived in [11], α is the conduction angle (9b).

$$\eta_{D,\,max} = \frac{\alpha - \sin(\alpha)}{4\left[\sin(\frac{\alpha}{2}) - \frac{\alpha}{2}\cos(\frac{\alpha}{2})\right]} \quad \text{Classes A–C} \tag{11}$$

The efficiency as a function of output voltage is also derived in [11] for class A (12) and class B (13).

$$\eta_D(v_{out}) = \frac{v_{out}^2/(2R_{load})}{I_{DC}V_{DD}} = \frac{v_{out}^2}{2R_{load}I_{DC}V_{DD}} \quad \text{Class A} \tag{12}$$

$$\eta_D(v_{out}) = \frac{v_{out}^2/(2R_{load})}{2V_{DD}v_{out}/(\pi R_{load})} = \frac{\pi}{4}\frac{v_{out}}{V_{DD}} \quad \text{Class B} \tag{13}$$

As can be seen, the class B amplifier has an efficiency that depends linearly on v_{out} and thus faster achieves a high efficiency compared to class A, which instead has a square dependency. The maximum efficiency of the class A is 1/2, whereas it is π/4 for the class B amplifier. The class C amplifier efficiency approaches 100% as $\alpha \to 0$, unfortunatly the output power then approaches zero.

The efficiency of the class D amplifier has been derived in [48] (14):

$$\eta_D(v_{out}) = \frac{R_{load}}{R_{load} + r_{ds,on} + R_L + \frac{2I_r^2(r_{ds,on}+R_L)R_{load}^2}{v_{out}}} \quad \text{Class D} \tag{14}$$

Here $I_r = \frac{V_{DD}}{4\sqrt{3}Lf_{IN}}$ is the current ripple through the inductor, $r_{ds,on}$ is the on resistance of the transistor, and R_L is the series resistance of the inductor L. As can be seen in (14), the efficiency has a rather weak dependency on v_{out}, which gives the amplifier good efficiency even at low output power levels. If switching ripple is neglected, $I_r = 0$, the efficiency is reduced to the well known resistive division between the load and the losses in the inductor and the switch.

The class E efficiency has been derived in [49]. The rms currents used to calculate the component losses are derived with lossless components, and assumed constant with losses, which is a good approximation if the losses are low. The efficiency is given by (15):

$$\eta_D(D) = \left\{1 + \frac{R_C + R_L + R_{C_1} + (r_{ds,on} + R_{C_1})\left[D + \frac{1}{4\pi}\sin(2\phi) - \sin(4\pi D + 2\phi)\right]}{R_{load}} + \frac{\left[R_{L_{RFC}} + r_{ds,on}(2-D) + R_{C_1}\right]\left[\cos(2\pi D + \phi) - \cos(\phi)\right]^2}{2\pi^2(1-D)^2 R_{load}}\right\}^{-1}, \text{Class E} \tag{15}$$

Table 1: Summary of the efficiency for the different PA classes

	η_D	$\eta_{D,\,max}$
Class A	(12)	50%
Class B	(13)	$\pi/4 \approx 78.5\%$
Class C	—	100%
Class A–C	—	$\frac{\alpha-\sin(\alpha)}{4\left[\sin(\frac{\alpha}{2})-\frac{\alpha}{2}\cos(\frac{\alpha}{2})\right]}$
Class D	(14)	100%
Class E	(15)	100%
Class F	(17)	100%

where $R_{L_{RFC}}$ is the series resistance of the L_{RFC} inductor, D is the duty cycle, R_{C_x} is the series resistance of capacitor C_x, and ϕ is the phase angle by which the output current lags the input current (16), as derived in [50]:

$$\phi = \arctan\left[\frac{1-\cos(2\pi D)}{2\pi(1-D)+\sin(2\pi D)}\right] = \arctan\left[\frac{1-\cos(\alpha)}{2\pi-\alpha+\sin(\alpha)}\right]. \tag{16}$$

The efficiency of the class F amplifier depends on how many odd harmonics are present at the drain, see [47] for a detailed description. The waveform represents a square wave better if many odd harmonics are present, and thus a higher efficiency can be achieved. The efficiency of the class F amplifier is expressed in (17):

$$\eta_D(\gamma_V,\gamma_I) = \frac{\gamma_V\gamma_I}{2}, \qquad \text{Class F} \tag{17}$$

where γ_V is a 'quality factor' of the fundamental voltage, $\gamma_V \equiv V_{fund.}/V_{DD}$. For an infinite number of odd harmonics γ_V approaches $4/\pi$. How well the current waveform represents a half sinusoid is represented by γ_I and approaches $\pi/2$ for an infinite number of even harmonics in the drain-source current. Of course component losses will reduce the efficiency to a value less than (17).

A summary of the efficiencies is presented in Table 1.

2.4.3 Output Power Issues in CMOS RF PAs

The knee voltage, V_{knee}, of the CMOS class A–C amplifier is defined as the voltage at the output where the amplifier starts to act as a voltage-controlled current source, as illustrated in Fig. 23. This is the smallest voltage that has to be present at the output for the amplifier to work properly. If a cascode is inserted, that extra stacked transistor also needs some voltage, thereby increasing V_{knee}. This diminishes the maximum peak-to-peak voltage from $2V_{DD}$ to $2(V_{DD} - V_{knee})$, so the full supply voltage can not be utilized. This degrades both maximum output power and efficiency .

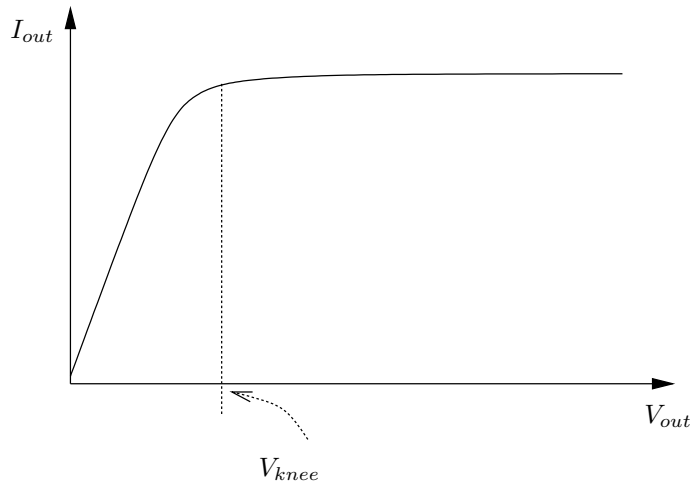


Figure 23: How the knee voltage of a CMOS amplifier is defined

Furthermore in today's downscaling of the CMOS feature sizes, also the supply voltage must be downscaled to prevent the electric fields in the device from increasing too much. This makes the design of high power CMOS amplifiers difficult, and obtaining the power level of about 1 W required by for instance cell phones, is extremely challenging using standard transistors in a scaled CMOS process.

Chapter 3

Some Aspects on CMOS Technology in High Frequency Design

Before going into how the metal oxide semiconductor (MOS) transistor works, let us take a look back on how it all began with the birth of the first transistor and later the CMOS technology.

3.1 The Birth of the Transistor and CMOS Technology

As commonly known the transistor was invented in December 1947 at Bell Labs by John Bardeen and Walter H. Brattain. Not as commonly known is that in Europe in June 1948 two researchers, Herbert Mataré and Heinrich Welker[9] [51, 52], invented the transistron, a device very much like the transistor invented only six months earlier. Both devices made amplification possible and were composed of a germanium sliver with three contacts placed close together. One could say that the transistor was not invented once but twice.

In 1962 at Fairchild Semiconductor Research and Development a young Ph. D., Frank M. Wanlass, was working on the development of a complementary field effect transistor. In June 18, 1963, he filed a patent application [53, 54] of a CMOS inverter, a NOR gate, and a NAND gate. The inverter is depicted in Fig. 24. The large benefit of the CMOS technology was that the static power consumption was reduced by six orders of magnitude compared to bipolar and p-channel MOS (pMOS) logic. In one step the power consumption was reduced from milliwatts to nanowatts. This made the technology ideal for battery powered products. Neither Wanlass nor Fairchild Semiconductor became rich from his patent, however, because in those days the companies traded their patent portfolios.

In the mid-1970s analog CMOS circuits were developed. The most well-known circuit is the switched-capacitor circuit techniques [55], which paved the road for mixed-signal design in CMOS [56, 57]. It was not until the mid-1990s, however, that the first radio frequency (RF) circuits were reported [55]. The first published RF CMOS amplifier appeared in 1993, see Fig. 25. The circuit was a balanced tuned amplifier operating at 770 MHz with 14 dB gain. In the 2 μm digital process a noise figure of 6 dB was achieved, with a power consumption of 7 mW from a 3 V supply [58].

[9] Mataré and Welker were to researchers from Germany who worked in France after World War II. Mataré worked in the German radar program and Welker was a solid-state physicist.

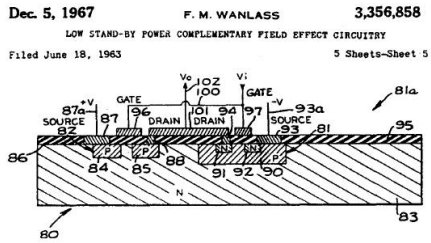


Figure 24: Wanlass's patent portray an integrated CMOS inverter

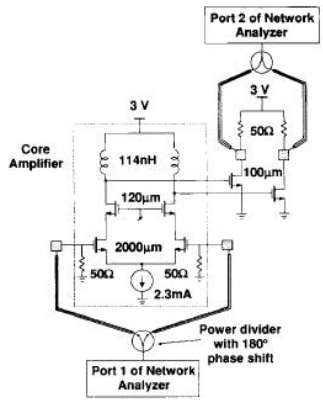


Figure 25: The very first published RF CMOS circuit

3.2 The MOS Transistor

This chapter contains a description of the n-channel MOS (nMOS) transistor, but the principle of operation is the same for the p-channel MOS (pMOS) device. The interested reader is referred to [59] for more detailed information on the pMOS transistor. All equations derived or cited in this chapter are for the nMOS transistor.

In the nMOS transistor the electrons transport the current, but in the pMOS it is instead the 'holes' that perform the transport. In an nMOS transistor the drain and source areas are doped with a group V-material (donors), while in pMOS they are instead doped with a group III-material (acceptors). A group V-material has five valence electrons thus making a tight bond to the silicon lattice by four of them, leaving one electron loosely bound, hence the name donors. In the same manner the group III-material lacks one electron to form a tight bound and this creates what is called a hole in the lattice, which can travel when a voltage is applied just like an electron. That is why atoms from group III is called acceptors, since they accept an extra electron.

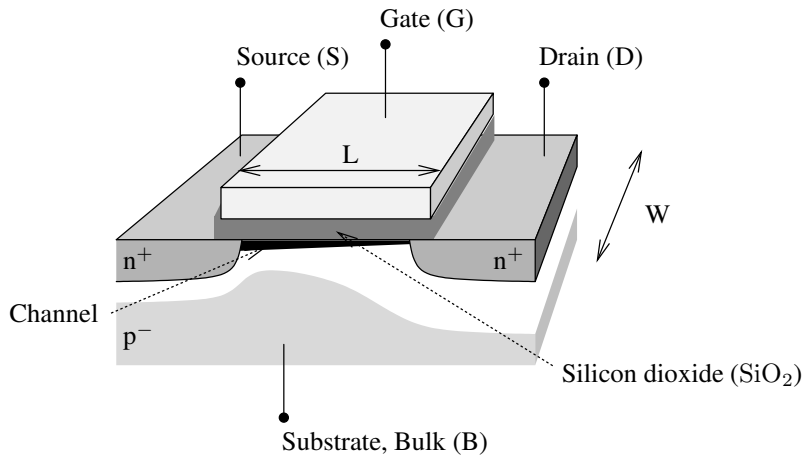


Figure 26: A cross section view of an nMOS transistor

Despite these differences, the operation of the nMOS and pMOS is very similar, and from this point on, only nMOS will therefore be described.

A cross section of an nMOS transistor is shown in Fig. 26. The transistor is usually fabricated on a highly doped p-substrate resulting in a low-ohmic bulk (B), which makes latch-up less of a problem. Latch-up can be caused if an n^+p^--junction in a transistor becomes forward biased (see Fig. 26). If this happens the whole circuit can be destroyed, and to release the latch-up the supply has to be disconnected [60,61]. To avoid forward biasing the n^+p^--junctions the bulk is connected to the lowest potential in the circuit.

The length of a transistor channel is denoted L, and the width W. The length is measured across the poly-silicon gate (Fig. 26). Under the gate (G) is an oxide, in a silicon transistor technology it is usually silicon dioxide (SiO_2)[10]. The oxide isolates the gate from the channel, which is between source (S) and drain (D). The thin oxide layer also isolates the gate from the three other terminals meaning that no significant DC current can flow into the gate terminal. The MOS transistor is symmetrical[11] and there is no difference between the source and the drain terminal. The terminal with the lowest potential is usually called source.

Before going into the operation of the transistor, let's define some nomenclature of potentials and currents. Capital letter of a quantity and its subscript means that it is

[10]Other gate oxide materials have been introduced recently in the academia and industry, called high-k oxides. With these high-k materials the physical gate oxide can be made thicker, still having control of the charges in the channel, thus reducing the tunneling current and increasing the breakdown voltage of the nanometer thin gate oxide.

[11]To reduce drain induced barrier lowering, the drain side of the channel can have a different doping level, more about that in Section 3.4.

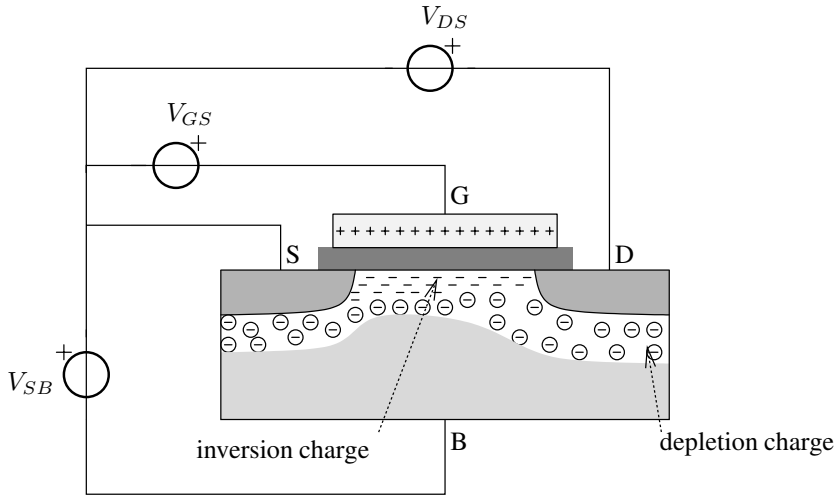


Figure 27: A static model of the nMOS transistor in the strong inversion region

a DC quantity, small letter of the quantity and its subscript means small signal, and a small letter in combination with a capital subscript indicates that it is the total value. Primed quantities are fixed to a value, constants.

3.2.1 Basic Transistor Operation

To get a better understanding of the operation of the transistor, a short description of the nMOS static behavior is presented. The starting point is the static model of the transistor shown in Fig. 27.

Let us assume that the bulk and source are connected to the lowest potential in the circuit, i. e. $V_{SB} = 0$, and that the gate and drain potential are higher than the source, $V_{GS} > 0$ and $V_{DS} > 0$. With sufficiently high gate potentials electrons will be attracted to the surface just under the gate, forming an excess of electrons in the channel called an inversion layer. It is called inversion because the p-substrate has an excess of holes, but the channel now instead has an excess of electrons, the inverse of holes. The electrons in the channel are supplied from the highly n-doped source and drain areas. The n-regions form n^+p^--junctions with the bulk, creating depletion regions under the n-regions. The depletion regions almost entirely stretch in to the substrate, since this is not as highly doped as the n-regions [60].

The number of free electrons in the channel is mainly determined by the potential

Table 2: Boundaries of the different inversion regions

	Weak inversion	Moderate inversion	Strong inversion	
			linear	saturation
V_{GS}	$V_{LB} \leq V_{GS} < V_{MB}$	$V_{MB} \leq V_{GS} < V_{HB}$	$V_{HB} \leq V_{GS}$	
V_{DS}	—	—	$0 < V_{DS} \leq V_{DS}^{'}$	$V_{DS}^{'} < V_{DS}$

$$V_{SB} = 0 \tag{18a}$$

$$V_{CB}(x) = V_{SB} + \frac{V_{GS} - V_T}{\alpha}\left[1 - \sqrt{1 - \frac{x}{L}(1-\eta^2)}\right] \tag{18b}$$

$$V_{CB}(0) = V_{SB} \tag{18c}$$

$$V_{LB} = V_L + V_{CB} = V_{FB} + \phi_F + \gamma\sqrt{\phi_F + V_{CB}} + V_{CB} \tag{18d}$$

$$V_{MB} = V_M + V_{CB} = V_{FB} + 2\phi_F + \gamma\sqrt{2\phi_F + V_{CB}} + V_{CB} \tag{18e}$$

$$V_{HB} = V_H + V_{CB} = V_M + V_Z + V_{CB} \tag{18f}$$

$$V_{DS}^{'} = \frac{V_{GS} - V_T}{\alpha} \tag{18g}$$

difference between gate and source. The transistor is said to operate in different inversion regions, weak, moderate, and strong inversion, depending on the number of free charge carriers in the channel. The boundary between the regions are summarized in Table 2 [59], where V_{FB} is the flat band voltage[12], ϕ_F the Fermi potential, and V_{CB} the voltage difference between channel and bulk. V_Z is typically about 0.5 V–0.6 V at room temperature with a small dependency of V_{CB}. The exact value depends on doping levels and oxide thickness. The parameter γ is defined in (21), V_T in (20a), and α is the charge distribution profile of the channel. V_L, V_M, and V_H are the voltages between gate and channel at the onset of weak, moderate, and strong inversion, respectively. V_{LB}, V_{MB}, and V_{HB} are the voltages when the bulk is at different potential than the channel.

Strong inversion has been, and still is, the most common region for analog design, but with the downscaling of the CMOS technology and the accompanying reduction of supply voltage, the strong inversion region becomes smaller and smaller. In weak inversion the power consumption is the lowest among the three regions, which has made that region of operation attractive for ultra low power applications, e. g. hearing aid products, pacemakers, and built-in defibrillators.

In weak inversion the current is mainly transported by diffusion, and in strong inversion mainly by drift [62]. However, in moderate inversion both drift and diffusion contribute roughly the same amount to the current, which makes it more difficult to

[12]The flat band voltage is the voltage that has to be applied between gate and bulk for the energy levels to align, typically -1 V for an nMOS transistor [59].

find simple expressions for the current in that region. Weak and moderate inversion will not be treated in the following sections. For more details about operation regions see e.g. [59], from here on only strong inversion will be considered.

3.2.2 DC Transistor Operation

If a voltage is applied between source and drain, the electrons will be attracted to the drain area, and a current will flow in the opposite direction. Also in this section we assume that $V_{SB} = 0$. When a voltage is applied between source and drain, the inversion layer will be different at the source and the drain side of the channel. The region of inversion for the transistor refers to the source side of the channel, which is most inverted. For a small V_{DS} the current is proportional to voltage applied (19b) [59].

$$I_{DS} = \frac{W}{L_{eff}}\mu C_{ox}^{'}\left[\left(V_{GS} - V_T\right)V_{DS} - \frac{\alpha}{2}V_{DS}^2\right] \qquad 0 < V_{DS} \leq V_{DS}^{'} \quad (19a)$$

$$\approx \frac{W}{L_{eff}}\mu C_{ox}^{'}\left[\left(V_{GS} - V_T\right)V_{DS}\right], \qquad V_{DS} \text{ small} \quad (19b)$$

where L_{eff} is the effective length of the channel, which differs from the drawn gate length due to under-diffusion, and at large V_{DS} also due to channel length modulation. To the first approximation α is equal to one, and (19a) then reduces to a more common expression for the drain current [43, 63]. V_T is the threshold voltage (20a), μ the mobility of the charge carriers in the channel, $V_{DS}^{'}$ the upper limit of the linear region (24), and $C_{ox}^{'}$ the gate oxide capacitance per unit area (22), where ϵ_r is the permittivity of the oxide and t_{ox} the thickness of the same oxide.

$$V_T = V_{T0} + \gamma\sqrt{\phi_0 + V_{SB}} - \sqrt{\phi_0} \quad (20a)$$

$$V_{T0} = V_{FB} + \phi_0 + \gamma\sqrt{\phi_0} \quad (20b)$$

where ϕ_0 is the surface potential[13] and γ is the body effect coefficient, defined as (21).

$$\gamma \equiv \frac{\sqrt{2q\epsilon_s N_A}}{C_{ox}^{'}} \quad (21)$$

$$C_{ox}^{'} = \epsilon_0\epsilon_r/t_{ox}, \quad (22)$$

As can be seen in (19b), for small V_{DS} the transistor acts as a resistor with a resistance controllable by V_{GS}. Up to moderate values of V_{DS} equation (19a) is valid, this is called the linear region. For an illustration see Fig. 28, where the current is plotted according to (19a) and (23).

[13] ϕ_0 is often taken as $2\phi_F$ but that is a rather large under-estimation as has been shown in [59] and should be more like $\phi_F + 6\phi_t$ (ϕ_t being the thermal voltage, 25 mV at room temperature) for a uniform substrate. These are only approximations and ϕ_0 is a function of V_{GB}.

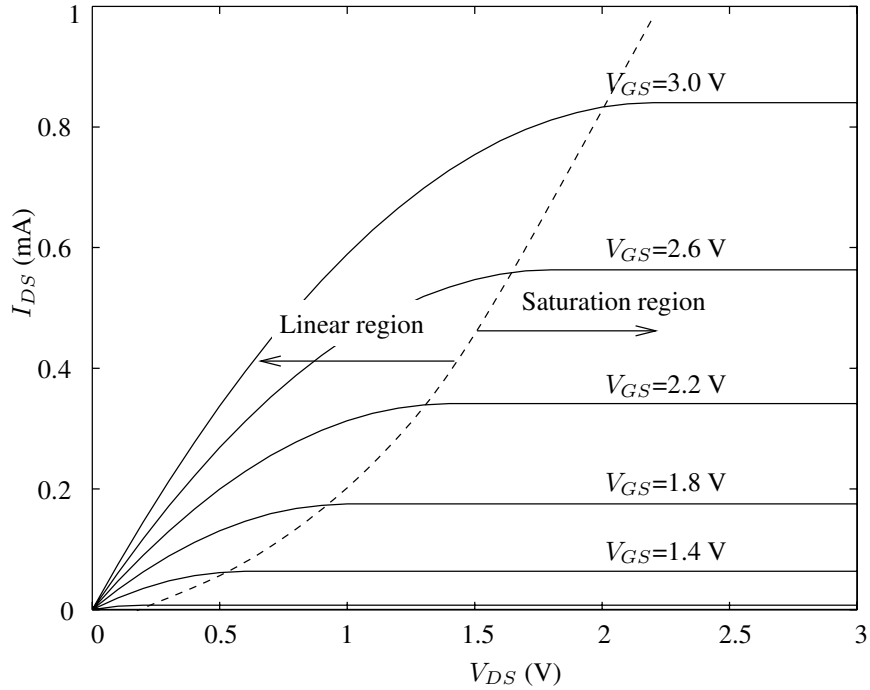


Figure 28: Output characteristics of an nMOS transistor in strong inversion, with $\alpha = 1$

Increasing V_{DS} further will empty the channel from free charges at the drain side. Since there are no free charges at the drain, V_{DS} will no longer influence the current in the channel. The transistor is said to be in saturation, and a further increase of V_{DS} will only result in a voltage drop in the so called pinched off region of the channel, see Fig. 29. In saturation the transistor behaves as a voltage-controlled current source (23).

$$I_{DS} = \frac{W}{L_{eff}} \mu C'_{ox} \frac{(V_{GS} - V_T)^2}{2\alpha} \qquad V'_{DS} < V_{DS} \quad (23)$$

V'_{DS} is defined as the voltage at maximum I_{DS} of (19a), which results in (24).

$$V'_{DS} \equiv \frac{V_{GS} - V_T}{\alpha} \quad (24)$$

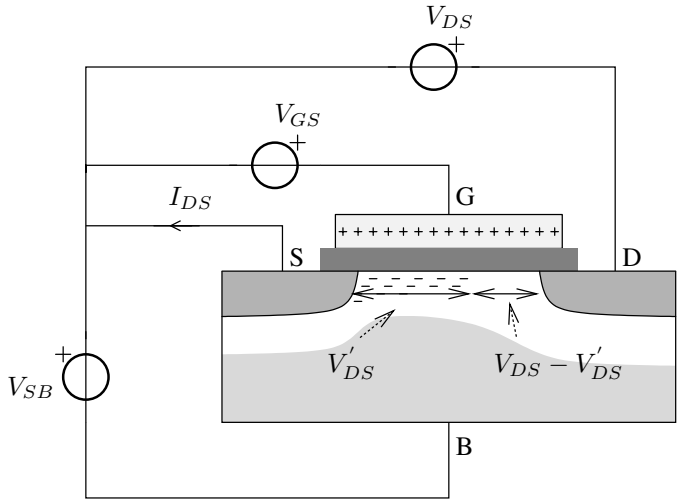


Figure 29: A cross section view of charges in the channel in saturation region

3.3 Small Signal Modeling

In the previous section the large signal DC characteristics were given. When it comes to the design of analog circuits, however, small signal models are needed. The small signal model is a linearisation of the large signal model around a quiescent point. To achieve this the large signal equation is differentiated with respect to the different voltages, resulting in (25), which depends linearly on terminal voltages.

$$i_{ds} \approx \left(\frac{\partial I_{DS}}{\partial V_{GS}}\bigg|_{V_{SB},\, V_{DS}} \right) \cdot v_{gs} + \left(\frac{\partial I_{DS}}{\partial V_{SB}}\bigg|_{V_{GS},\, V_{DS}} \right) \cdot v_{sb} + \left(\frac{\partial I_{DS}}{\partial V_{DS}}\bigg|_{V_{SB},\, V_{GS}} \right) \cdot v_{ds} \tag{25}$$

The voltages and direction of the currents are defined in Fig. 29. The transistor model considered in this section is the long-channel approximation with infinitesimal source and drain areas. Deviations from this approximation will be treated in Section 3.4. The section covers low and medium frequency models of the transistor. For high frequency models, sometimes called non-quasi-static model, refer to [59, 64] for details.

3.3.1 Transconductances and Output Conductance

As was just described, the large signal drain current equation should be differentiated with respect to different voltages to achieve the small signal model (25). The result is the gate transconductance, g_m, the bulk transconductance, g_{mb}, and the output conductance, g_{sd} (26). Equation (25) can now be expressed in a more compact way (27).

$$g_m \equiv \left.\frac{\partial I_{DS}}{\partial V_{GS}}\right|_{V_{SB},\,V_{DS}},\ g_{mb} \equiv \left.\frac{\partial I_{DS}}{\partial V_{SB}}\right|_{V_{GS},\,V_{DS}},\ g_{sd} \equiv \left.\frac{\partial I_{DS}}{\partial V_{DS}}\right|_{V_{SB},\,V_{GS}} \tag{26}$$

$$i_{ds} \approx g_m \cdot v_{gs} + g_{mb} \cdot v_{sb} + g_{sd} \cdot v_{ds} \tag{27}$$

Taking the derivative of the drain current with respect to DC voltages in strong inversion results in two sets of equations depending on in which region the transistor is working. For g_m this results in (28a) and (28b), where $V_{DS}^{'}$ is the same as in (24).

$$\begin{aligned} g_m &= \frac{W}{L_{eff}}\mu C_{ox}^{'} V_{DS}^{'} = \frac{2I_{DS}}{V_{GS}-V_T} = \sqrt{\frac{W}{L_{eff}}\frac{\mu C_{ox}^{'}}{\alpha} I_{DS}} \\ &= \frac{W}{L_{eff}}\frac{\mu C_{ox}^{'}}{\alpha}\left(V_{GS}-V_T\right) \qquad V_{DS}^{'} < V_{DS} \quad (28a) \\ g_m &= \frac{W}{L_{eff}}\mu C_{ox}^{'} V_{DS} \qquad 0 < V_{DS} \le V_{DS}^{'} \quad (28b) \end{aligned}$$

The gate transconductance, g_m, is the most important parameter, as it represents the gain of the device. The bulk transconductance, g_{mb}, only has an effect when there is a signal voltage between the source and substrate. This is the case in some commonly used structures, such as the common gate amplifier stage. Then also g_{mb} has to be incorporated in parallel with g_m in the small signal model, see Fig 30. Taking the derivative of the current results in (29a) and (29b).

$$\begin{aligned} g_{mb} &= \left(\frac{\gamma}{\sqrt{V_{DS}^{'}+V_{SB}+\phi_0}+\sqrt{V_{SB}+\phi_0}}\right) g_m \qquad V_{DS}^{'} < V_{DS} \quad (29a) \\ g_{mb} &= \left(\frac{\gamma}{\sqrt{V_{DS}+V_{SB}+\phi_0}+\sqrt{V_{SB}+\phi_0}}\right) g_m \qquad 0 < V_{DS} \le V_{DS}^{'} \quad (29b) \end{aligned}$$

For low V_{GS} and V_{DS} a simplification of the above equations can be made (30), which gives an approximation of the impact of g_{mb}

$$\frac{g_{mb}}{g_m} \approx \frac{\epsilon_s}{\epsilon_{ox}}\frac{t_{ox}}{d_{Bm}} \tag{30}$$

where d_{Bm} is the depth of the depletion region. Typical values of this ratio are in the range 0.1 to 0.3 [43].

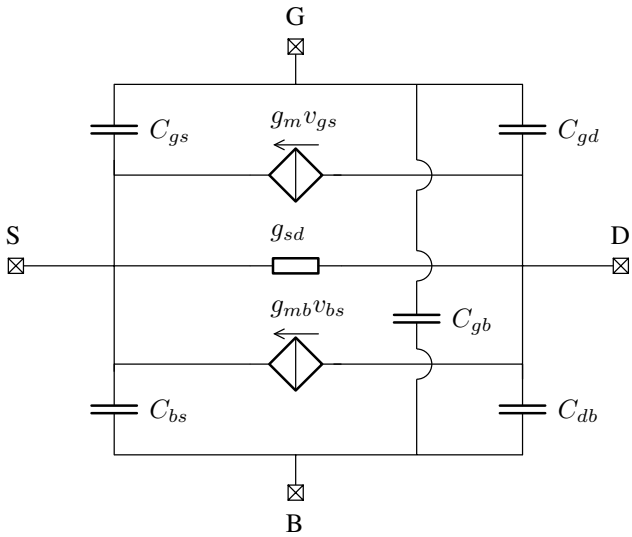


Figure 30: Small signal model of the MOS transistor at medium frequencies

The source-drain output conductance in the saturation region is zero with the long-channel approximation; deviations will be discussed in Section 3.4. In the linear region the conductance is:

$$g_{sd} = \frac{W}{L_{eff}} \mu C_{ox}^{'} \left[V_{GS} - V_T - \alpha V_{DS} \right] \qquad 0 < V_{DS} \leq V_{DS}^{'} \quad (31)$$

The output conductance is important, since it limits the intrinsic voltage gain of a transistor.

3.3.2 Capacitive Effects

To incorporate charge storage effects in the model, five capacitances can be added between the four terminals. There is a capacitance between all terminal pairs except source and drain as illustrated in Fig. 30. The capacitances are defined according to (32).

$$C_{gs} \equiv -\left.\frac{\partial Q_G}{\partial V_S}\right|_{V_G,\, V_D,\, V_B} \quad C_{bs} \equiv -\left.\frac{\partial Q_B}{\partial V_S}\right|_{V_G,\, V_D,\, V_B} \quad C_{gd} \equiv -\left.\frac{\partial Q_G}{\partial V_D}\right|_{V_G,\, V_S,\, V_B}$$
$$C_{bd} \equiv -\left.\frac{\partial Q_B}{\partial V_D}\right|_{V_G,\, V_S,\, V_B} \quad C_{gb} \equiv -\left.\frac{\partial Q_G}{\partial V_B}\right|_{V_G,\, V_D,\, V_S} \quad (32)$$

Expressions for the capacitances can be found in [59], repeated here for conve-

nience.

$$C_{gs} = C_{ox}\frac{2(1+2\eta)}{3(1+\eta)^2} \tag{33a}$$

$$C_{bs} = (\alpha-1)C_{ox}\frac{2(1+2\eta)}{3(1+\eta)^2} = (\alpha-1)C_{gs} \tag{33b}$$

$$C_{gd} = C_{ox}\frac{2(\eta^2+2\eta)}{3(1+\eta)^2} \tag{33c}$$

$$C_{bd} = (\alpha-1)C_{ox}\frac{2(\eta^2+2\eta)}{3(1+\eta)^2} = (\alpha-1)C_{gd} \tag{33d}$$

$$C_{gb} = \frac{\alpha-1}{3\alpha}C_{ox}\left(\frac{1-\eta}{1+\eta}\right)^2 \tag{33e}$$

Where C_{ox} is the total oxide capacitance, $C_{ox} = C_{ox}^{'}WL$, and η is defined as:

$$\eta \equiv \begin{cases} 1-\frac{V_{DS}}{V_{DS}^{'}}, & 0 < V_{DS} \leq V_{DS}^{'} \\ 0, & V_{DS} > V_{DS}^{'} \end{cases} \tag{34}$$

In the saturation region, where $\eta = 0$, (33a)–(33e) this is reduced to (35a)–(35e).

$$C_{gs} = \frac{2}{3}C_{ox} \tag{35a}$$

$$C_{bs} = \frac{2}{3}(\alpha-1)C_{gs} \tag{35b}$$

$$C_{gd} = 0 \tag{35c}$$

$$C_{bd} = 0 \tag{35d}$$

$$C_{gb} = \frac{\alpha-1}{3\alpha}C_{ox} \tag{35e}$$

The dominant capacitance is clearly the gate-source capacitance, C_{gs}, and the capacitances associated with the drain are both equal to zero. Note that this is for the intrinsic transistor. The situation will be changed somewhat by the extrinsic effects described in the next section.

3.3.3 Some Extrinsic Effects

Gate Resistance

The gate is made of silicided poly-silicon in most CMOS processes and has a resistance of about 5-10 Ω per square [63, 65, 66]. In recent 45 nm CMOS processes, however, metal gates with much lower resistance are used [67, 68]. In the commonly used BSIM3 transistor model gate resistance is not included. This resistance influences the noise performance of the transistor. The noise generated by the resistance is

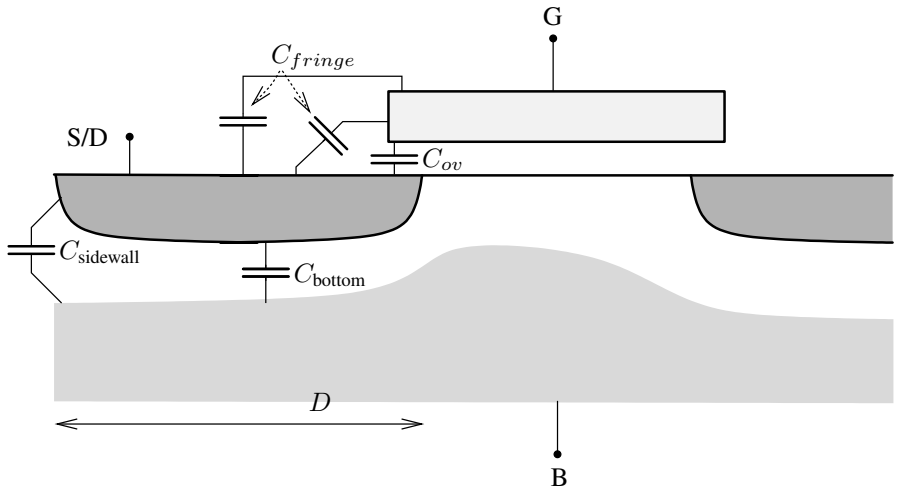


Figure 31: Capacitances arising from finite drain and source area

amplified by the transistor and added to the output current noise. At some bias conditions and transistor geometries, it can be dominant and should therefore always be included when simulating an LNA. Contacting the gate on one side the equivalent gate resistance, R_{gate}, can be calculated as (36), where $R_{\square}$ is the resistance per square of the gate. If the gate instead is contacted on both sides the resistance is reduced by a factor of four [69, 70].

$$R_{gate} = \frac{1}{3}\frac{W}{L}R_{\square} \tag{36}$$

The gate resistance is distributed, forming small RC-links with the channel capacitance. If the transistor is wide, also a phase shift can thus be introduced along the gate. This phase shift can affect the stability of feedback circuits at high frequencies. The low pass characteristic will also reduce the gain of the device at high frequencies [70].

Junction and Overlap Capacitances

In a real transistor the drain and source regions have finite area, which was not modeled in Section 3.3.2. Fig. 31 shows a cross section of an nMOS transistor with junction and extrinsic gate capacitances. The overlap capacitance, C_{ov}, is due to under-diffusion when depositing the n-doped source and drain regions, using the gate as a mask (self-alignment).

The overlap and fringe capacitances can be calculated from basic electrostatics, if the geometries are known. The under-diffusion length, L_D, is not easy to determine,

so the capacitance from gate to source/drain in Fig. 31 is usually characterized by measurements. The total capacitance is proportional to the transistor width, and can be calculated as (37), where $C_{ov}^{'}$ is the overlap capacitance per unit length.

$$C_{ov} = C_{ov}^{'} W \tag{37}$$

As mentioned in Section 3.2.1 there is a depletion region under the source and drain. The depth of the region depends on the reverse voltage across the junction, V_{rev}, and the capacitance is an ordinary junction capacitance. To model the capacitance accurately, a model with a bottom-plate capacitance and a side-wall capacitance is typically used [11, 43, 63]. The total capacitance can then be expressed as in (38), where D is the length of the drain and source diffusion regions, and the parameter $m_{j,x}$ depends on the doping profile and is in the range of 0.3–0.4 [10].

$$\begin{aligned} C_{\text{junc.}} &= C_{\text{sidewall}} + C_{\text{bottom}} \\ &= \frac{C_{j0,\text{sidewall}}}{\left(1 + V_{rev}/2\phi_F\right)^{m_{j,\text{sidewall}}}} \cdot (2D + W) + \frac{C_{j0,\text{bottom}}}{\left(1 + V_{rev}/2\phi_F\right)^{m_{j,\text{bottom}}}} \cdot DW \end{aligned} \tag{38}$$

The zero-bias capacitances per unit length or unit area, $C_{j0,\text{sidewall/bottom}}$, are not trivial to calculate and are therefore also usually found by measurements.

3.4 Small Dimension Effects

3.4.1 Channel Length Modulation

As described in Section 3.2.2 an increase of V_{DS} beyond $V_{DS}^{'}$ results in a region near the drain where the channel is not inverted, see Fig. 29. This causes a drain voltage dependent effective channel length reduction, resulting in a finite output conductance of the device [43]. This effect thus makes the drain current dependent on the output voltage, even in the saturation region, and the output characteristics of Fig. 28 are not valid any more. The output characteristics are instead as in Fig. 32. The behavior is often modeled with a factor in the current equation that is linearly dependent on V_{DS} (39), as is typically done in bipolar processes [43].

$$I_{DS} = \frac{W}{L_{eff}} \mu C_{ox}^{'} \frac{\left(V_{GS} - V_T\right)^2}{2\alpha} \left(1 + \lambda V_{DS}\right), \qquad V_{DS}^{'} < V_{DS} \tag{39}$$

where λ sets the slope of the output current in the saturation region.

The channel length modulation (CLM) is present in all MOS technologies, and it is listed in this section since it is a small dimension effect. The reduction of the channel length is independent of total gate length, and thus the relative increase in current becomes more and more significant the shorter the gate is.

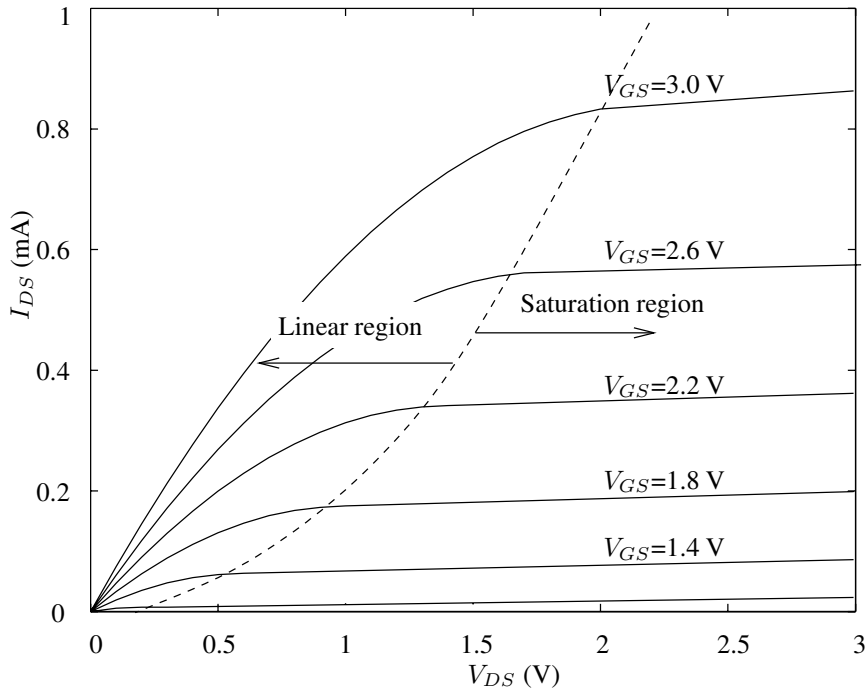


Figure 32: Output characteristics of a nMOS transistor, with $\alpha = 1$ and channel length modulation

3.4.2 Threshold Voltage Lowering

In a short device the drain and source areas are closely spaced. As illustrated in Fig. 33 the depletion regions from drain and source can then extend under a large part of the channel. The extended depletion region under the channel increases the surface potential, ψ_s, in the channel making it easier for electrons to enter[14]. The transistor then needs a smaller V_{GS} to start conducting. The effect can be modeled as a lowering of the threshold voltage of the device [71, 72], see Fig. 34.

As discussed in Section 3.3.3, the depth of the depletion region changes with the reverse voltage applied across the junction. An increased V_{DB} or V_{SB} thus leads to an increased depletion region. The increased depletion region results in an increased surface potential in the channel, which can be significant in a short device. Thus the effective threshold voltage of the device decreases and the drain current becomes

[14]For a fixed positive V_{GB} a larger (negative) voltage occurs over the depletion region. This causes the potential difference between gate and channel to increase, resulting in a higher surface potential. This effect is sometimes called barrier lowering since it lowers the potential barrier for the electrons to enter the channel.

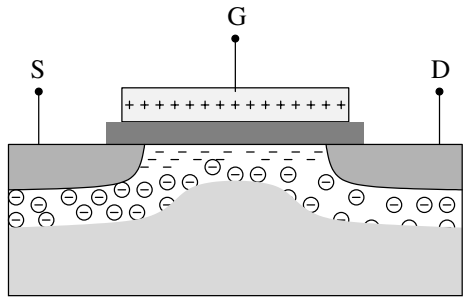


Figure 33: A short channel MOS transistor

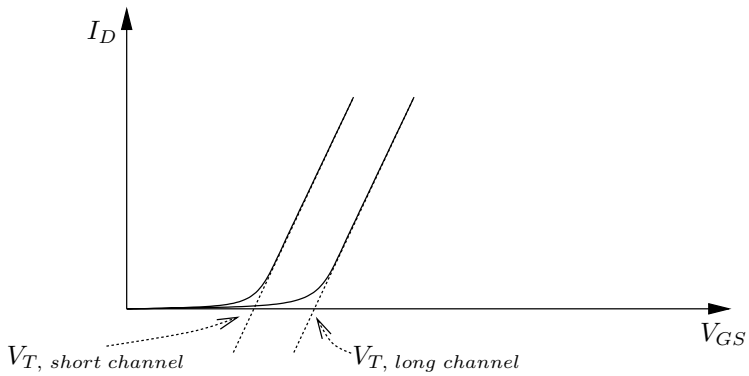


Figure 34: Threshold lowering effect in a short channel MOS transistor

dependent on $V_{SB/DB}$. This effect is usually called drain-induced barrier lowering (DIBL), an analytical expression is derived in [72]. The drain-bulk dependency of the threshold voltage may cause the transistor to start conducting even though V_{GS} is well below V_{T0}. This is particularly harmful in digital circuits, where it results in increased leakage currents.

3.4.3 Horizontal Field Effects

Velocity Saturation

When the equations for the current in Section 3.2.2 were derived, the horizontal field in the channel was assumed to be small, giving a constant mobility with a charge carrier velocity, v_d, proportional to the electric field. As the horizontal field increases, this is no longer true, and the velocity of the charge carriers begins to saturate as illustrated

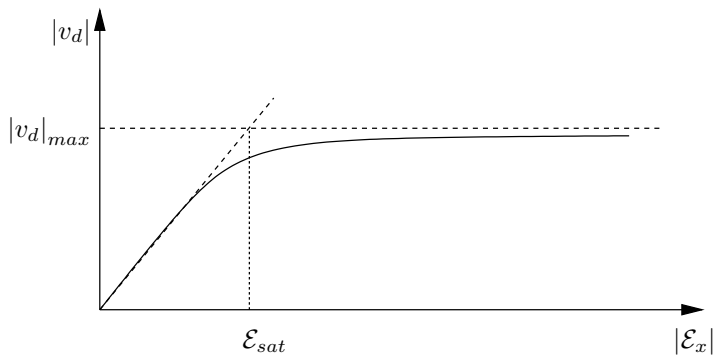


Figure 35: Velocity saturation

in Fig. 35 [73]. From this curve a parameter called the critical field strength, $\mathcal{E}_{sat}$, is extracted. This is the extrapolated intersection between the portion of the curve where the mobility is constant and the part where the velocity has saturated (40). In silicon the critical field strength is at about 1 V/μm and the corresponding velocity is 0.1 Mm/s [11].

$$\mathcal{E}_{sat} = \frac{|v_d|_{max}}{\mu} \tag{40}$$

A variety of fitting functions exists to approximate the curve [59]. A commonly used function is (41) [74], where $\mathcal{E}_x$ is the horizontal field. This fitting function is popular since it results in simple transistor equations. The channel current derived without the velocity saturation effect then only has to be modified by a multiplicative factor, $1/\left(1 + V_{DS}/L\mathcal{E}_{sat}\right)$, where V_{DS} should be replaced by $V_{DS}^{'}$ in the saturation region.

$$|v_d| = |v_d|_{max} \frac{|\mathcal{E}_x| \,/\, |\mathcal{E}_{sat}|}{1 + |\mathcal{E}_x| \,/\, |\mathcal{E}_{sat}|} \tag{41}$$

Hot Carriers

Even though the mean carrier velocity saturates at $|v_d|_{max}$ there are some carriers moving at greater speed, called hot carriers. They can have much higher kinetic energy. These hot carriers can cause impact ionization when colliding with the silicon lattice atoms, knocking more tightly bound electrons free, thereby creating electron hole pairs. These create a drain to bulk current. The hot electrons may also damage the gate oxide if they collide with it. They can also be trapped inside the oxide, resulting in a lowered threshold voltage. More information about hot carriers can be found in [75, 76].

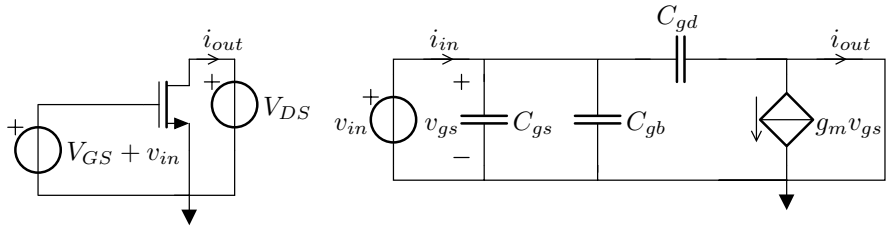


Figure 36: Test bench and small signal model for calculating f_T

3.4.4 Vertical Field

The vertical field effect is not a short channel effect, since it is present also in long devices, but it is a high field effect.

In the MOS structure of Fig. 29 the strong gate to channel field attracts the charge carriers towards the oxide interface. The roughness of the surface causes the charge carriers to scatter and lose velocity and thus reduces the mobility [43, 59]. The effect of the mobility degradation is usually modeled as an effective mobility reduction (42), where μ_0 is the mobility without vertical field, and θ_b is inversely proportional to the gate oxide thickness and is found by measurements.

$$\mu_{eff} = \frac{\mu_0}{1 + \theta_b \left(V_{GS} - V_T\right)} \tag{42}$$

3.4.5 Transition Frequency

The transition frequency, f_T, is the most commonly used figure of merit to measure the speed of a transistor. The transition frequency is defined as the frequency at which the small signal current gain of the transistor equals unity, when its output is small signal shorted (43) as depicted in Fig. 36.

$$f_T \equiv \frac{1}{2\pi} \frac{i_{out}}{i_{in}}\bigg|_{V_{ds}=0} = \frac{1}{2\pi} \frac{g_m}{C_{gs} + C_{gb} + C_{gd}} \tag{43}$$

Using the long channel equations results in (44).

$$\begin{aligned} f_T^{long} &= \frac{1}{2\pi} \frac{g_m}{C_{gs} + C_{gb} + C_{gd}} \approx \frac{1}{2\pi} \frac{g_m}{C_{gs}} = \frac{1}{2\pi} \frac{\mu C'_{ox} \frac{W}{L_{eff}} \left(V_{GS} - V_T\right)}{\frac{2}{3} W L C'_{ox}} \\ &\approx \frac{1}{4} \frac{\mu \left(V_{GS} - V_T\right)}{L^2} \end{aligned} \tag{44}$$

Taking into account small dimensional effects such as velocity saturation the transconductance has to be rederived. The drain to source current in the strong inversion saturation region can be expressed according to (45a), where $V_{DS,sat}$ is as in

(45b) [11].

$$I_{DS} = \frac{\mu C_{ox}^{'}}{2}\frac{W}{L_{eff}}\left(V_{GS}-V_T\right)^2 = \frac{\mu C_{ox}^{'}}{2}\frac{W}{L_{eff}}\left(V_{GS}-V_T\right)V_{DS,sat} \tag{45a}$$

$$V_{DS,sat} \approx \left(V_{GS}-V_T\right) \,||\, \left(L_{eff}\mathcal{E}_{sat}\right) = \frac{V_{GS}-V_T}{1+\frac{V_{GS}-V_T}{L_{eff}\mathcal{E}_{sat}}} \tag{45b}$$

The ratio of $V_{GS}-V_T$ to $L\mathcal{E}_{sat}$ determines how large the velocity saturation effect of the transistor is. If $V_{GS}-V_T$ is much larger than $L\mathcal{E}_{sat}$ it behaves like a short channel device, and (45a) and (45b) can be approximated as (46a) and (46b).

$$I_{DS} = \frac{\mu C_{ox}^{'}}{2}W\left(V_{GS}-V_T\right)\mathcal{E}_{sat} \tag{46a}$$

$$V_{DS,sat} = L_{eff}\mathcal{E}_{sat} \tag{46b}$$

The drain to source current now no longer depends on the device length. This can be seen as the electrons reaching $|v_d|_{max}$ as soon as they enter the channel. The charge in the channel and the transit time are then both proportional to L, and thus the current is independent of the length. Furthermore, the current only depends linearly on the overdrive voltage, $V_{GS}-V_T$. Under these circumstances the transconductance, g_m, is independent of both L (47) and the overdrive voltage.

$$g_m \equiv \left.\frac{\partial I_{DS}}{\partial V_{GS}}\right|_{V_{SB},\,V_{DS}} = \frac{\mu C_{ox}^{'}}{2}W\mathcal{E}_{sat} \tag{47}$$

The transition frequency can be expressed as (48). The transition frequency is then proportional to $1/L$, and not $1/L^2$ as in the long channel approximation (44).

$$f_T^{short} = \frac{1}{2\pi}\frac{g_m}{C_{gs}+C_{gb}+C_{gd}} \approx \frac{g_m}{C_{gs}} = \frac{1}{2\pi}\frac{\frac{\mu C_{ox}^{'}}{2}W\mathcal{E}_{sat}}{\frac{2}{3}WLC_{ox}^{'}} = \frac{1}{8}\frac{\mu\mathcal{E}_{sat}}{L} \tag{48}$$

3.4.6 Maximum Frequency of Oscillation

The maximum frequency of oscillation, f_{max}, is another common figure of merit for high frequency performance of transistors. It is the highest frequency at which an oscillator can be designed using the transistor. It is defined as the frequency where the unilateral power gain is unity. Unilateral power gain is the power gain when both input and output are conjugate-matched to the source and the load impedances. The input power is dependent on the gate resistance, which can be reduced by proper layout, see Section 3.3.3. The power delivered to the load is also dependent on careful layout. Minimizing the resistance in both drain and source is therefore important. Of course the intrinsic output conductance should be as large as possible. An approximate

Table 3: Reported f_T and f_{max} in CMOS process nodes

f_T (GHz)	39	70	115	155	170	150	180	360	280
f_{max} (GHz)	45	150	80	250	240	200	255	420	350
Process node (nm)	250	180	130	90	90	90	65	65	45
L_{eff} (nm)	—	—	—	63	65	70	55	29	36
Reference	[77]	[78]	[79]	[80]	[81]	[82]	[80]	[83]	[80]

expression for f_{max} is given in (49a) [11]. Taking also the output conductance into account the equation in (49b) results [59]. Adding the source and drain resistance as well, f_{max} can be written as (49c), where R_i is the resistance associated with the finite quality factor of the gate-to-source capacitance [84, 85].

$$f_{max} \approx \frac{1}{4\pi}\sqrt{\frac{\omega_T}{R_{gate}C_{gd}}} \tag{49a}$$

$$f_{max} \approx \frac{1}{4\pi}\frac{\omega_T}{\sqrt{R_{gate}\left(g_{sd}+\omega_T C_{gd}\right)}} \tag{49b}$$

$$f_{max} \approx \frac{g_m}{2\pi C_{gs}}\frac{1}{2\sqrt{\left(R_{gate}+R_s+R_i\right)\left(g_{sd}+g_m\frac{C_{gd}}{C_{gs}}\right)}} \tag{49c}$$

In Table 3 a summary of reported f_T and f_{max} for different CMOS process nodes and effective gate lengths is given.

3.4.7 Intrinsic Voltage Gain

The intrinsic voltage gain, a_v, is the maximum voltage gain of a single transistor, defined in (50), and it is an important parameter in amplifier design.

$$a_v \equiv \frac{g_m}{g_{sd}} \tag{50}$$

With the long channel approximation in Section 3.3.1, the intrinsic voltage gain in strong inversion saturation region is infinite, since $g_{sd} = 0$.

With small dimension effects, the output conductance is no longer zero. The output current, when taking CLM into account, is given by (39) and the output conductance derived from that equation is (51).

$$g_{sd} = \frac{W}{L_{eff}}\mu C_{ox}^{'}\frac{\left(V_{GS}-V_T\right)^2}{2\alpha}\lambda = \lambda I_{DS} \qquad \text{CLM, } V_{DS}^{'} < V_{DS} \tag{51}$$

Resulting in an intrinsic gain (52), where g_m is given in (28a). The intrinsic gain thus reduced with increased overdrive voltage.

$$a_v = \frac{\frac{2I_{DS}}{V_{GS-V_T}}}{\lambda I_{DS}} = \frac{2}{\lambda\left(V_{GS}-V_T\right)} \qquad \text{CLM, } V_{DS}^{'} < V_{DS} \quad (52)$$

As mentioned earlier the threshold voltage is lowered if the depletion regions from drain and source extend under the channel. Applying a large V_{DB} will cause a change in the depletion region depth, from $V_{DB} = 0$, and thus modulate the threshold voltage and drain current, also causing the output conductance to increase. The intrinsic gain under DIBL is derived in [59] and cited below.

$$\begin{aligned} a_v &= \frac{\frac{W}{L_{eff}}\frac{\mu C_{ox}^{'}}{\alpha}\left[V_{GS}-V_T\left(V_{DS}\right)\right]}{\frac{W}{L_{eff}}\frac{\mu C_{ox}^{'}}{\alpha}\left[V_{GS}-V_T\left(V_{DS}\right)\right]\left(-\frac{\partial V_T}{\partial V_{DS}}\right)} \\ &= -\frac{1}{\frac{\partial V_T}{\partial V_{DS}}} \approx 2\frac{\epsilon_{ox}}{\epsilon_s}\frac{L}{t_{ox}} \qquad \text{DIBL, } V_{DS}^{'} < V_{DS} \quad (53) \end{aligned}$$

Depending on the modeling of $V_T\left(V_{DS}\right)$ somewhat different expressions for the intrinsic gain can be derived, for more details see [59]. From (53) it can be concluded that the intrinsic gain improves with increasing length of the transistor, i. e. for high gain make the transistor as long as possible. This also reduces the channel length modulation.

The above mentioned effects are both present at the same time. If the transistor is biased at a small overdrive voltage, $V_{GS} - V_T$ small, the DIBL effect dominates. For small overdrive voltage the output current changes rapidly with a change in threshold voltage. For larger overdrive voltage the CLM effect is instead dominant.

To get an estimation on the intrinsic gain[15] of a transistor, four different CMOS processes, 0.35 μm, 130 nm, 90 nm, and 65 nm are simulated. The simulation is conducted with a drain-source voltage that is half the recommended maximum supply voltage of the process, for three transistor lengths, L_{min}, $2L_{min}$, and $5L_{min}$, see Table 4. The width to length ratio has been kept constant at 10. For the older process, 0.35 μm, the intrinsic gain is around hundred already at L_{min}, ranging up to several hundred at $5L_{min}$, while in the newer processes the intrinsic gain starts at 10 and ranges up to 40 at $5L_{min}$.

3.5 Noise

Low noise is very important in a receiver since it determines the level of the smallest signal detectable. This, of course, puts stringent requirements on the building blocks of the receiver, but also the transmitter must be 'quiet' and not leak too much noise to the receiver or antenna.

[15]The intrinsic gain is calculated from the DC operating point in Cadence simulations.

Table 4: Simulated intrinsic gain for four different CMOS processes

node	65 nm	90 nm	130 nm	0.35 μm
$V_{GS} - V_T$ (mV)	200	200	200	200
W/L (μm/μm)	10	10	10	10
a_v @ L_{min} (V/V)	5.9	12.7	11.8	71.6
a_v @ $2L_{min}$ (V/V)	18.4	38.3	29.7	276.6
a_v @ $5L_{min}$ (V/V)	38.2	43.2	46.4	561.0

A typical noise power spectrum, S_i, for the drain current of a MOS transistor is depicted in Fig. 37. Two different frequency regions can be identified, one to the left of the corner frequency, f_c, and one to the right. The noise spectrum is frequency dependent to the left of f_c and has a slope of approximately minus one in the log-log plot. This is called $1/f$-noise, or flicker noise. On the right hand side the noise is frequency independent up to very high frequencies (THz) [59]. This is called white noise.

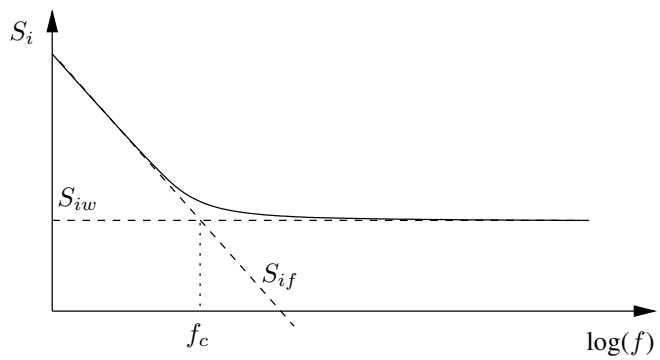


Figure 37: Typical drain current noise spectrum plot of the MOS transistor

3.5.1 White Noise

White noise is due to random thermal movements of electrons and holes [86]. Since the origin of white noise is thermal movement it exists in all devices. In transistors the white noise has different levels in different inversion regions, strong inversion is treated briefly below. For a thorough investigation see [33, 59]. The white noise is

Table 5: Flicker noise trends for the two theories, ΔN and $\Delta\mu$

	ΔN	$\Delta\mu$
Strong inversion, Linear Region	$\propto \frac{I_D^2}{WLV_{GS}^2}$	$\propto \frac{I_D^2}{WLV_{GS}}$
Strong inversion, Saturation Region	$\propto \frac{I_D}{L^2}$	$\propto \frac{I_D^{3/2}}{WL^3}$

often model as a noise-current source between source and drain.

In strong inversion the white noise can be expressed as (54), which has been derived in [59].

$$S_{iw} = 4kT\left[g_{sd0}\frac{2}{3}\frac{1+\eta+\eta^2}{1+\eta}\right] \tag{54}$$

where η is defined in (34), k is Boltzmann's constant, and T is the absolute temperature. The factor 2/3 in (54) comes from long channel approximation and can due to hot electrons be 3–5 times larger in short channel devices [11, 87, 88].

3.5.2 Flicker Noise

Flicker noise causes an increase of the noise at low frequencies. The location of the corner frequency, see in Fig. 37, varies with technology. In MOSFETs it may vary from tens of kilohertz to several megahertz [11]. This noise source is very hazardous in zero-IF receivers and in oscillators, where the low frequency noise is transformed to the carrier frequency by non-linearities in the oscillator.

The origin of the noise is still not completely understood. Two main theories exist today [89, 90],

1. (ΔN) The noise comes from traps in the lattice which captures charge carriers and then randomly releases them again [91]

2. $(\Delta\mu)$ The mobility of the charge carriers is randomized due imperfections (scattering) in the lattice [92]

The dependency of the flicker noise on current and transistor size varies between the two models and with inversion region [89]. The dependency, from [89], on I_D and V_{GS} has been summarized in Table 5.

From Table 5 it is safe to say that the noise increases with the drain current and with decreasing length of the transistor. If the circuit is sensitive to flicker noise, the current should be kept low and the transistors long. The problem is that this also results in reduced speed.

3.5.3 Gate Induced Noise

The random fluctuation of electrons in the channel, which causes drain current noise, also couple capacitively to the gate of the transistor. This results in induced gate noise. An equation for the gate induced noise was derived in [33] and is cited below

$$\overline{i_{ng}^2} = 4kT\delta g_g \Delta f \tag{55}$$

$$g_g = \frac{\omega^2 C_{gs}^2}{5 g_{d0}} \tag{56}$$

The gate induced noise can be modeled as a current source injecting the noise current across C_{gs}. The constant δ has been derived for the long channel transistor, and is 4/3 (2γ) [33].

The noise behavior in long channel operation is fairly well understood but the behavior of δ in short channel regime is unknown at present. The origin of γ and δ are both random movement of electrons in the channel and they are thus correlated. As an approximation the same relationship ($\delta = 2\gamma$) is commonly assumed also in short channel operation [11, 59].

Despite the quite complex nature of the gate induced noise the resulting equation is almost the same as the noise of a conductance, g_g, the factor δ is the only difference.

The gate induced noise increases with the square of frequency and is thus important to model in circuits working at high frequencies. It can also be harmful in circuits with high Q at its input, e. g. in source degenerated LNAs [93].

Chapter 4

Link Budget

Link budget calculations are performed for all radio systems to estimate the required transmit power and receiver sensitivity to achieve the desired coverage area. The link budget contains estimates of losses in the transmitter and receiver, the antennas, and the path loss. The principle is illustrated in Fig. 38, where $\mathrm{P_{TX}}$ is the transmitter output power, $\mathrm{G_{TX}}$ the gain of the transmit antenna, $\mathrm{L_{FS}}$ free space loss, and $\mathrm{L_M}$ loss margin for fading. $\mathrm{G_{RX}}$ is the gain of the receiving antenna, and $\mathrm{P_{RX}}$ the received power.

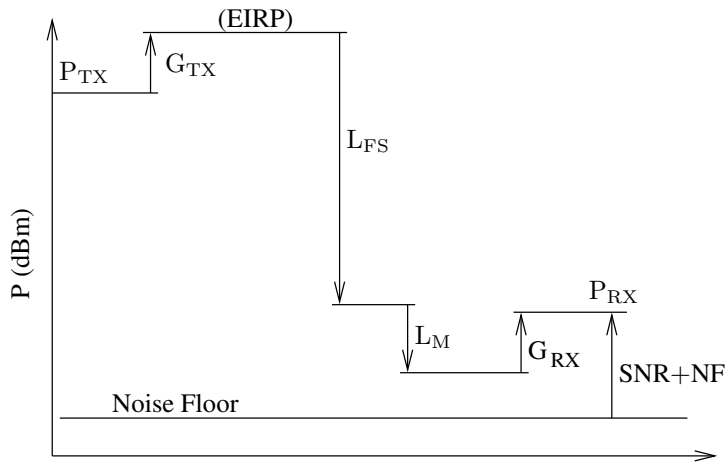


Figure 38: A principle link budget diagram

The transmitter and receiver losses are feeding losses to the antennas, due to filters, interconnect, and connectors. They have been left out in Fig. 38 for clarity. The minimum signal-to-noise ratio (SNR) required by the receiver determines how much stronger a received signal must be than the noise floor to be correctly detected. The SNR required is dependent on the modulation and coding used to transmit the data. The antenna can have either gain or loss depending on the design and orientation. A highly directional transmit antenna pointing towards the receiver can have a large

gain, when compared to an omnidirectional antenna. The path loss is the loss between the transmit and receive antennas. The path loss consists of different parts, where free space loss (57) [94] is one, absorption in molecules, and diffraction are others. In equation (57a) d is distance between transmitter and receiver. Line of sight is assumed, and the received power decreases as the square of the distance. For long distances (57a) gives a too optimistic estimation and thus has to be modified to also incorporate the ground reflection (57b). Long distance is in this case $\approx 4h_{TX}h_{RX}/\lambda$, where $h_{TX/RX}$ is the height above ground of the transmit and receive antennas, and λ is the wavelength. These are two simple equations to estimate the free space loss. To get higher accuracy more complicated or empirical models can be used [94].

$$P_{RX}(d) = P_{TX}G_{TX}G_{RX}\left(\frac{\lambda}{4\pi d}\right)^2 \tag{57a}$$

$$P_{RX}(d) \approx P_{TX}G_{TX}G_{RX}\left(\frac{h_{TX}h_{RX}}{d^2}\right)^2 \tag{57b}$$

The gain of an antenna depends on its directivity, D, which can be written as (58), where A is the aperture size of the antenna. The gain of an antenna equals the directivity if the antenna is ideal, i. e. loss-less and perfectly matched. The directivity is a measure of how much more effectively the antenna radiates in a certain direction compared to an isotropic antenna. When antennas are designed for a higher frequency, the aperture size and the wavelength will both decrease, keeping the directivity roughly constant. The received power, however, will decrease by the square of frequency (57a) dramatically shortening the distance at which communication can be upheld.

$$D = \frac{4\pi A}{\lambda^2} \tag{58}$$

If instead the antenna occupies the same area even though the frequency is increased, i. e. constant A, then the directivity instead will increase. To make this work the beam of the transmit antenna must be directed towards the transceiver and vice versa. This can be accomplished using beamforming techniques with antenna arrays, more on that in Section 4.1.2.

Fading originates from multi path propagation, the signal reaches the receiver from more than one direction. The received signals can then be delayed or reflected differently, resulting in constructive or destructive interference. In worst case a total cancellation can occur. How deep fading the system can manage is called the fading margin, $\mathrm{L_M}$, and is taken into account in the link budget.

4.1 Link Budget Calculation

In this section two simple link budget calculations illustrate the procedure. The first example is the WCDMA system for mobile phones, see Table 6. The second is

Table 6: Link budget calculation for the WCDMA up-link (∼2 GHz)

Receiver		
Noise floor ($10\log(kTB)$)	-108.2 dBm	$B = 3.84$ MHz
SNR	-20 dB	FDD mode [96]
NF	4 dB	Base station
Antenna gain	17 dBi	
Cable and Connector	3 dB	
Receiver sensitivity	-138.2 dBm	
Transmitter		
Output power	21 dBm	Maximum hand-held
Antenna gain	2 dBi	
EIRP	23 dBm	
Distance		
Loss margin	15 dB	Fading etc.
Free space	146.2 dB	
Distance	1.8 km	Okamura-Hata model

an envisioned system for short range wireless communication in the unlicensed frequency band at 60 GHz, where the benefits of multiple power amplifiers combined with phased array antennas in high frequency systems will be shown, see Table 7.

4.1.1 WCDMA Example

The following example is based on the WCDMA standard, numbers are taken from [95]. The calculation is for the up-link, see Table 6.

The receiver noise floor is at $kTB = -108.2$ dBm, where B is the channel bandwidth of 3.84 MHz. The receiver SNR can be low, roughly -20 dB in FDD mode [96], and the noise figure can be made low since it is the base station receiver, 4 dB [97, 98]. The negative SNR can be used thanks to the de-spreading gain in WCDMA. Feeding losses (3 dB) and antenna gain (17 dBi for the receiver) are also in the budget. The maximum output power from the power amplifier in the hand-held unit is 21 dBm, antenna gain and feeding losses can be about 2 dBi, giving an EIRP of 23 dBm. The difference between EIRP and receiver sensitivity, 161.2 dB, is distributed to path loss and loss margin. A loss margin of 15 dB, including fading and interference, is reasonable. Using the Okamura-Hata model to calculate the path loss, $\mathrm{L_{FS}} = 137.4 + 35.2\log(\,\mathrm{R(km)}\,)$, results in a cell radius of 1.8 km.

4.1.2 60 GHz WPAN Example

In the US a 7 GHz wide license free band is located around 60 GHz. A bandwidth of 7 GHz opens up for unprecedented data rates for wireless transmission, over 3 Gbps, even when using simple modulation schemes. There is a 5 GHz world wide overlap, making this band commercially attractive. Applications where high data rates are needed are for example file transfer, kiosk movie download, and wireless HDTV [99, 100].

One of the reasons this frequency band is license free is that the oxygen molecule has a resonance at 60 GHz contributing to a high absorption in air. The high absorption makes long distance communication unfeasible. In the same time it also makes wireless personal area network more robust, since interference from neighboring cells is attenuated. The large attenuation of signals penetrating walls of buildings also contributes to this.

To succeed in the consumer electronics market, the products must be cheap. That is one of the reasons why the work in [1, 101] is focused on CMOS implementations. As the feature sizes of the CMOS transistor decrease, the transit frequency and maximum frequency of oscillation increase, and thus higher frequencies of operation can be used, see Table 3. Unfortunately the supply voltage must be reduced, which limits the amount of output power the transistor can deliver to a load. One way to increase the output power is to use multiple power amplifiers. The power from the different power amplifiers can then either be combined in a power combiner on the chip or PCB, or by using multiple antennas, combined in the air. The drawback with a power combiner is that it is inherently lossy. At 60 GHz the antennas as well as the distance between them in an array can be made quite small; $\lambda/2$ equals 2.5 mm. Furthermore, using an antenna array the transmitted power can be directed towards the receiver, further increasing the equivalent isotropic radiated power (EIRP). Table 7 is an example of EIRP and size of an antenna array at 60 GHz[16]. The output power of each PA is assumed to be 10 dBm, and the array consists of n dipole antenna elements, with an antenna element spacing of $\lambda/2$. As can be seen the size of the array is implementable in most consumer electronic devices.

A link budget estimation is performed in the same manner as in the previous section. The noise floor is higher $kTB = -84$ dBm, since the bandwidth is larger, 1 GHz. The SNR is taken to be 10 dB[17] [84] and the noise figure of the receiver 10 dB [84, 103, 104]. The antenna gain and thereby also the number of transmit antennas needed, depends on the number of parallel receiver paths. Lets assume 4 parallel paths, i. e. antenna gain 8.2 dBi, with variable phase delays so that the receiving beam is steerable and the signals add coherently. The effective receiver noise figure is also reduced by the antenna gain due to that noise in the different receivers is uncorrelated but the signals are correlated, which gives 6 dB less system noise figure [84, 105, 106].

[16] The calculation of the antenna gain is a bit conservative since coupling between antenna elements are not taken into account. This could roughly give 1 dB extra gain [102].

[17] An SNR of 10 dB corresponds to a bit error rate of 0.001 with frequency shift keying coding.

Table 7: Antenna array approximations at 60 GHz

n	1	2	3	4	5	6	7
P_{out} (dBm)	10.0	13.0	14.8	16.0	17.0	17.8	18.5
Antenna gain (dBi)	2.1	5.2	6.9	8.2	9.1	9.9	10.6
EIRP (dBm)	12.1	18.2	21.7	22.4	26.1	27.7	29.0
Size (mm)	—	2.5	5.0	7.5	10.0	12.5	15.0

The antenna gain is calculated as: $G_A = D_{dipole} + 10\log(\mathrm{n})$, where $D_{dipole} = 2.14$ dBi.

To have a WPAN that can cover a room or an office area distances of up to 10 m between receiver and transmitter should be allowed. Indoor propagation measurements at 60 GHz show that the path loss can be up to 90 dB [107–109] for a distance of 10 m. The loss margin is again set to 15 dB, since the placement of transmitter and receiver has a large influence. The measurements in [107] show dips of 30 dB depending on reflection material and height difference between receiver and transmitter, in good agreement with the two ray pass model (59) [94, 110]:

$$\frac{P_R}{P_T} = \left(\frac{\lambda}{4\pi d}\right)^2 \left|1 + \Gamma \exp\left(j\frac{2\pi}{\lambda}\frac{2h^2}{d}\right)\right|^2 , \tag{59}$$

where d is the distance between RX and TX, h is the height difference, and Γ is the reflection coefficient of the reflected beam.

This results in an EIRP of 26.8 dBm, which is below the maximum allowed limits by FCC [111, 112]. Looking in Table 7, a 6 element antenna array results in 27.7 dBm EIRP, giving one extra dB loss margin, see summary in Table 8. The antenna beam of the transmitter has to be steerable to point at the receiver, an electrical steerable transmitter therefore needs individual output phases to the antenna elements. In Chapter 6 different architectures to implement individually controllable phase shifts are discussed, and in [1, 2] the author's approach is presented.

Table 8: Link budget calculation for 60 GHz WPAN

Receiver		
Noise floor (10log(kTB))	−84.0 dBm	B = 1 GHz
SNR	10 dB	
NF	10 dB	CMOS [84]
NF gain	−6[1] dB	Array of 4 elements
Antenna gain	8.2 dBi	Array of 4 elements
Receiver sensitivity	−78.2 dBm	
Distance		
Loss margin	15 dB	Fading etc.
Free space	90 dB	Indoor [107–109]
Distance	10 m	Office area
Transmitter		
EIRP	27.7 dBm	
Output power	17.8 dBm	6 PAs at 10 dBm
Antenna gain	9.9 dBi	Array of 6 elements

[1]The noise at the receiving antennas is uncorrelated but the signals are correlated, if line of sight reception.

Chapter 5

Phased Array Antenna

This chapter contains a brief overview of phased array antennas used in electrically steerable beamforming systems, their benefits, problems, and limitations. The alternative to electrical steering is a mechanical scanning antenna, but this is not included in this thesis. The theory of how to feed signals to a phased array antenna in order to direct the antenna lobe in a certain direction, and what the resulting lobes will look like are described. The feeding of an antenna array puts requirements on the beamforming transmitter circuitry, presented in Chapter 6. For further information about phased array antennas see [102, 113–116].

5.1 Advantages of Phased Array Antennas

Phased array antennas are used in a large number of applications, such as radar, satellite communications, and base stations. Fig. 39 shows an example of a large phased array antenna, which is part of the ballistic missile early warning system and is placed in Alaska to monitor the northern sky. Evidently, a phased array antenna can be both large and complicated, as well as small and complicated.

With phased arrays the antenna directivity, D, can be made larger by increasing the number of elements of the antenna (60) [113],

$$D = \frac{4\pi}{\lambda^2} A_{em}, \tag{60}$$

where A_{em} is the maximum effective area of the antenna, which of course increases with number of elements in the array.

One benefit of antenna arrays is that the directivity can be made high by increasing the number of antenna elements. With the phased array comes another benefit, the direction of the antenna lobe can be steered by changing the individual phases of signals to/from the antenna elements. The beam can then scan through a certain part of a sphere, depending on the geometry of the array, without changing its physical position. This is called electrical scanning, more on that in Section 5.2 and in the literature [114, 115].

Figure 39: Figure of a phased array. This large, 27 m in diameter, phased array is placed in Alaska to monitor the northern sky as a part of the ballistic missile early warning system. Courtesy of Alaska District

5.2 Linear Phased Array Antenna

A linear phased array consists of a line of unit elements, with a centre to centre distance between neighboring elements of Δx. The number of unit elements and their separation sets the physical size of the array, and the larger the size the higher the directivity (60). A linear phase array consisting of 18 antenna elements was simulated in Paper III. Fig. 40 shows a principle sketch of a phased array antenna; the distances and angles defined in the figure will be used in equations throughout this section. The maximum scanning angle of the main beam is called θ_s.

The electric far field pattern of a dipole element is given by (61) [102]. The unit element can be any type of antenna, but in this case the dipole was chosen. The reason for using dipole is to have a concrete example, and the ability to use differential signals, which is often used in integrated circuits. For other element patterns see [102, 113, 114]. The type of unit element limits the maximum scan angle of the array, since the scanning angle can not be made wider than the main lobe of the unit element.

$$E_{element}\left(\theta\right) = \frac{\cos\left[\pi/2 \cdot \sin\left(\theta\right)\right]}{\cos\left(\theta\right)} \tag{61}$$

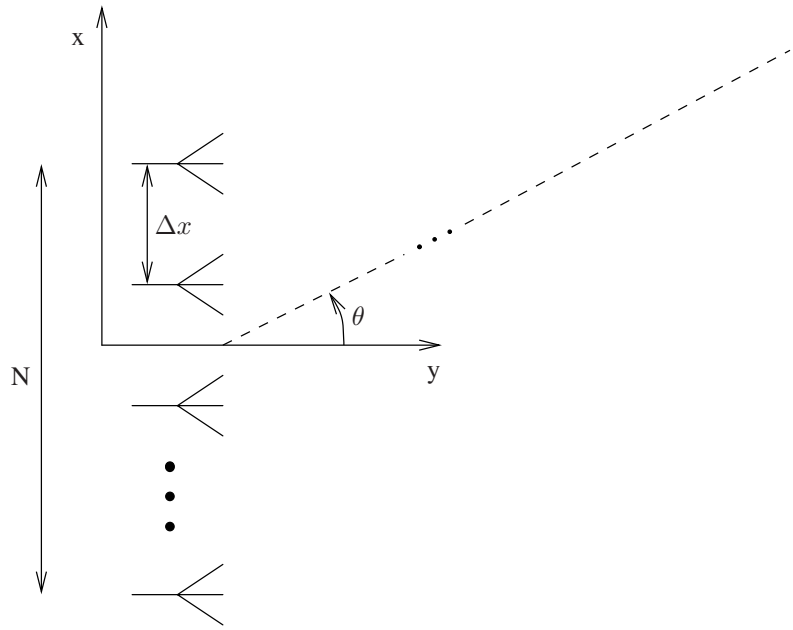


Figure 40: A principle sketch of a phased array antenna

To calculate the electric far field of the array, the field contributed by each of the elements has to be summed. This can be done using the principle of pattern multiplication [102]:

> *The field pattern of an array of nonisotropic but similar point sources is the product of the pattern of the individual source and the pattern of an array of isotropic sources having the same locations, relative amplitudes, and phase as the nonisotropic point sources.*

This means that in the far field, where the antenna elements behave as point sources, the assumption can be used if the sources are similar. There are situations where this principle of pattern multiplication can not be used. For instance, if the array is receiving a very short pulse (e. g. an UWB-pulse) at a certain angle, then the elements are not hit simultaneously, and thus the sources are not similar. There are other cases as well where the sources are not similar and the individual responses have to be summed to get the array response. From now on, however, the sources will be assumed similar if not stated otherwise.

A general expression for the far field pattern of an array of isotropic point sources, called array factor, AF, can then be used to calculate the total array pattern. For an N-element linear phased array with uniform spacing and equal amplitudes, the

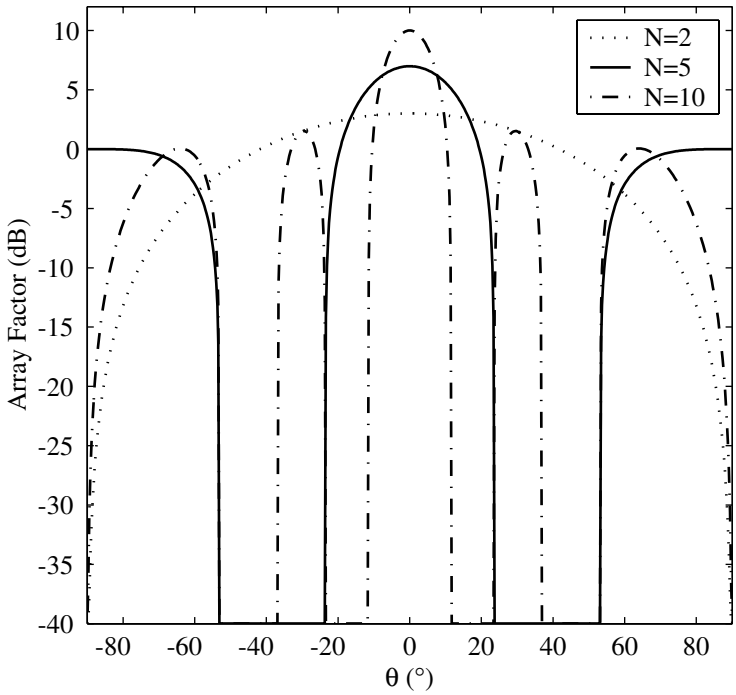


Figure 41: Array factor for equal spacing (λ/2) and amplitude with 2, 5, and 10 elements. The elements are in-phase

array factor is given in (62a), where δ is the relative phase shift between the antenna elements. The array factor for other stimuli, i. e. non-uniform amplitude, spacing, and phase are given in the literature [102, 113, 117].

$$AF = \frac{\sin\left(\frac{N}{2}\psi\right)}{\sin\left(\frac{1}{2}\psi\right)} \approx \frac{\sin\left(\frac{N}{2}\psi\right)}{\frac{1}{2}\psi} \tag{62a}$$

$$\psi = \beta\Delta x \cdot \sin(\theta) + \delta \tag{62b}$$

$$\beta = \frac{2\pi}{\lambda} \tag{62c}$$

In Fig. 41 the array factor is plotted for three different numbers of antenna elements. In Fig. 42 the array factor is plotted for three different phase offsets, δ, between the elements. The lobe is pointing at 0, -6.3, and -16.2 degrees for offset phases of 0, 20, and 50 degrees, respectively. The side lobes in Fig. 42 are 8.5 dB lower than the main lobe. With non-uniform amplitude distribution the side lobe suppression can be made larger at the expense of reduced gain. In applications where suppression of side

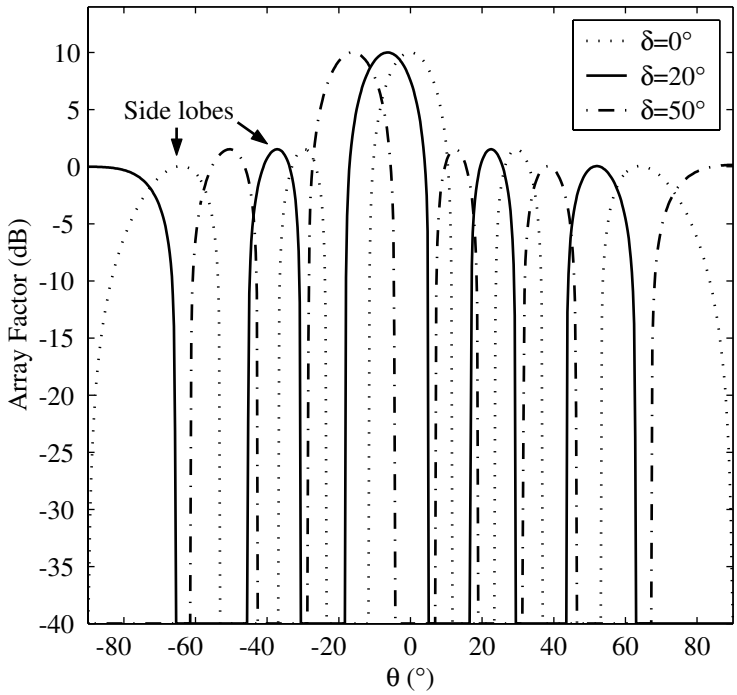


Figure 42: Array factor for equal spacing ($\lambda/2$) and amplitude with three different offset phases. The number of antenna elements is 10

lobes are of utmost importance a binomial distribution can be applied, as this theoretically eliminates the side lobes altogether. The antenna gain is reduced with binomial amplitude distribution, compared to uniform distribution. With an array consisting of eight isotropic elements with $\lambda/2$ spacing, the gain is reduced by 2.2 dB [102].

The far field pattern of the array now becomes:

$$E_{tot} = AF \cdot E_{element} \tag{63}$$

This means that if the array factor is known, the antenna array response only depends on the element pattern. Different types of antenna elements are discussed in [114,115].

5.2.1 Grating Lobes

An inherent problem with linear phased arrays is that they have grating lobes, that is the pattern has more than one main lobe. This springs from that the antenna element signals can add coherently in more than one direction. A definition of grating lobes is given in [113] and restated below.

A grating lobe is defined as a lobe, other than the main lobe, produced by an array antenna when the inter element spacing is sufficiently large to permit the in-phase addition of radiated fields in more than one direction.

To put it in equations, if the main beam is pointing at a direction, θ_0, then a grating lobe occurs at (64b):

$$\beta \Delta x \left(\sin\left(\theta\right) - \sin\left(\theta_0\right)\right) = 2\pi p \tag{64a}$$

$$\frac{\Delta x}{\lambda} \left(\sin\left(\theta\right) - \sin\left(\theta_0\right)\right) = p \tag{64b}$$

for all integer p. Solving (64b) for all angles θ_p results in (65b):

$$\sin\left(\theta_p\right) = \sin\left(\theta_0\right) + \frac{\lambda p}{\Delta x} \tag{65a}$$

$$\theta_p = \arcsin\left(\frac{\lambda p}{\Delta x}\right) + \theta_0 \tag{65b}$$

If the maximum scanning angle is known (65b) can be used to calculate the maximum element distance to eliminate the risk of grating lobes (66):

$$\Delta x \leq \frac{\lambda}{1 + \sin\left(\theta_s\right)} \tag{66}$$

where the first grating lobe occurs at $-90°$, giving the 1 in the denominator. In (66) p has been set to -1, since that gives the closest set of grating lobes. In some equations describing grating lobes, p is neglected altogether, since usually the interest is only when the first grating lobe enters the scanning area.

The array factor was plotted to validate equation (66) as well, see Fig. 43, with $\Delta x = 2.2\lambda$. It can be seen that the first grating lobe is at 66°, and the main lobe at 0°. From expression (66) the maximum grating lobe free scanning angle should be 33°.

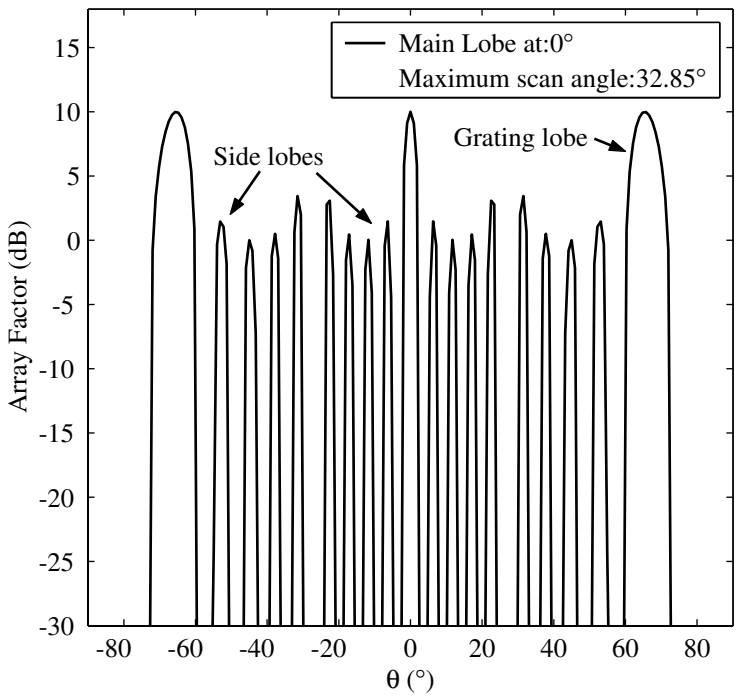


Figure 43: Array factor with equal spacing (2.2λ), phase, and amplitude. The grating lobe free area is 66°, which gives a maximum grating lobe free scanning of 33°. The number of antenna elements is 10

Chapter 6

Beamforming Transmitter Architectures

This section describes some beamforming transmitter architectures, appropriate for integrated circuit implementation. This excludes some well established topologies for discrete implementations of phased array transmitters, e. g. space fed and multi module arrays [118–120]. The aim here is to discuss topologies suitable for implementation in consumer products, such as automotive radar (24 GHz and 77 GHz) and 60 GHz WLAN.

To achieve electrical beamforming, the phase of the signal to each antenna element in the array has to be set individually [114–116]. If also the power to each antenna element can be set individually, a larger set of patterns can be produced and the sidelobe level can be reduced compared to uniform power distribution.

6.1 Baseband Phase Shifting

In the baseband phase shifting architecture the phases and amplitudes of the signals are created in the digital baseband. The phase control can be very precise, but the architecture requires a complete signal path from baseband to the antenna for each element (Fig. 44). The architecture is also called digital array [121], since the beamforming takes place in the digital domain. This architecture results in a large hardware cost and power consumption due to the multiple signal paths, but also in great flexibility. As a consequence this architecture is probably too complex for automotive radar and 60 GHz WLAN. The flexibility of the architecture with parallel paths can, however, also be used to transmit separate information in different directions in multiple input multiple output (MIMO) systems [122, 123].

6.2 Local Oscillator Phase Shifting

Phase shifting can also take place in the LO path [105, 116, 124–127], see Fig. 45. It is also possible to use phase shifters in the signal path, either at IF or RF. Performing the phase shift at LO or RF results in the same amount of hardware, as the phase shifters are only placed at different positions. Placing them in the LO path makes amplitude variation between different phase settings less important if the mixers are driven hard. Then amplitude variation in the LO path will not have much effect on the signal path. To implement the phase shift in the LO path is thus easier. Different approaches can be taken, one of which is exemplified below.

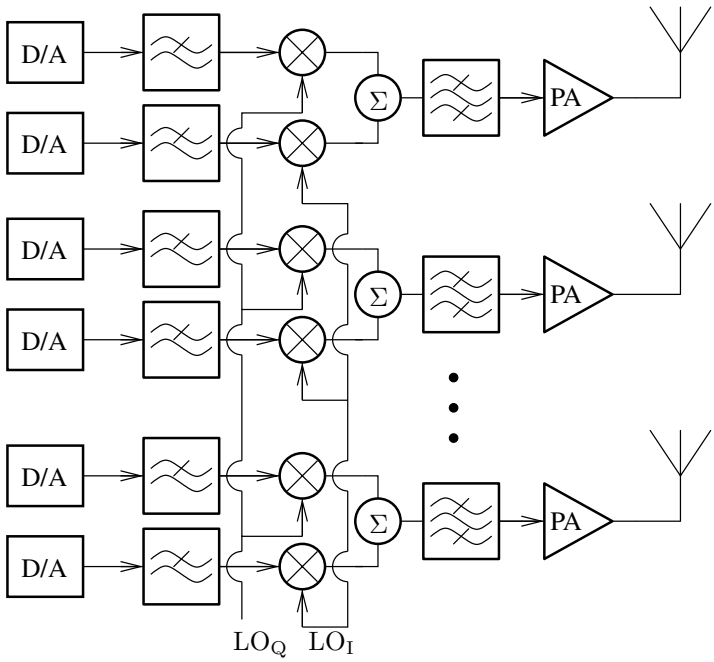


Figure 44: Transmitter architecture for baseband phase shifting

♢ A vector modulator can be used to achieve a variable phase shift in the LO path. The modulator can be implemented with a quadrature LO where each of the quadrature phases are weighted and combined to a resulting vector. This means that if the I+ and Q+ are weighted equally while I− and Q− are discarded the resulting vector will point in 45 degrees. The phase of the vector can be set arbitrarily with a constant amplitude, depending on the implementation of the weighting function. The same quadrature LO can be used to drive all the vector modulators in the beamforming transmitter. The circuit presented in Paper I is one such example [1]. In Paper V, an active polyphase filter is presented which can be used to produce the quadrature phases [5]. An alternative would be to implement a quadrature VCO on chip.

6.3 Offset Local Oscillator Phase Shifting

Operating the power amplifier and local oscillator at the same frequency can result in injection pulling, see Section 2.3.1. It might be difficult to assure adequate isolation in order to avoid that the PA corrupts the oscillator signal. To mitigate this on an architectural level, offset LO phase shifting can be used [128,129] as shown in Fig. 46.

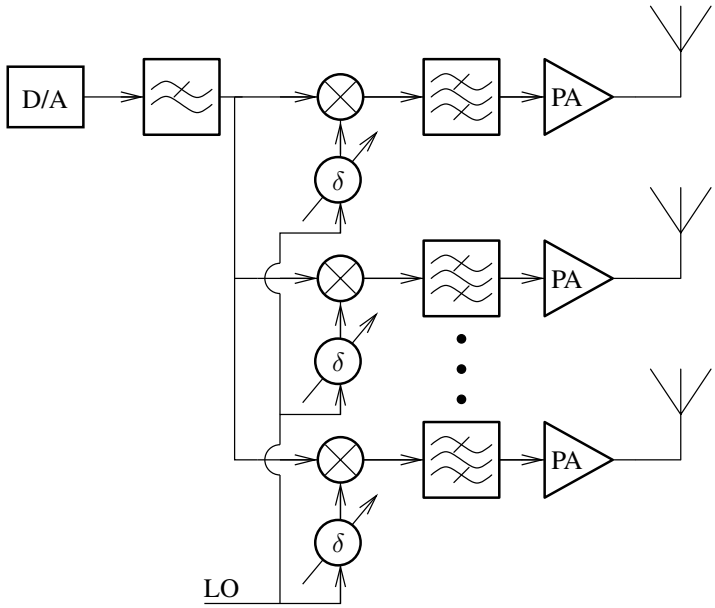


Figure 45: Transmitter architecture for phase shifting in the local oscillator path, polar modulation

Applications for beamforming transmitters like automotive radar (24 GHz and 77 GHz) and WLAN (60 GHz) are located at high frequencies. It is beneficial to use lower frequencies as much as possible on the chip, and multiply the frequency close to the PA. A reduced VCO frequency makes it possible to achieve a wider tuning range, and the MOS varactor quality factor is increased [84] ($\sim$1–2 @ 60 GHz versus $\sim$3–6 @ 20 GHz). Furthermore, the capacitively coupled losses to the bulk decreases linearly with frequency. A couple of architectures are described below.

♢ The offset LO approach in [128, 129] (Fig. 46) is based on an LO signal that is frequency divided by 2 and later mixed with the original LO signal, thus achieving a 1.5 times higher LO frequency. The phase shifting can preferably take place after the divide-by-2 circuit. The frequency is lowest there, one third of the carrier frequency, which makes the implementation more robust and less sensitive to parasitics. In the frequency division quadrature phase can be extracted from a differential LO, which can be used to implement the phase shift in a vector modulator circuit.

♢ In publications from Professor Hajimiri's group at Caltech [105, 124–126] another approach is used. A multi-phase LO has been demonstrated in silicon-germanium (SiGe) and CMOS at 24 GHz, both for receiver and transmitter.

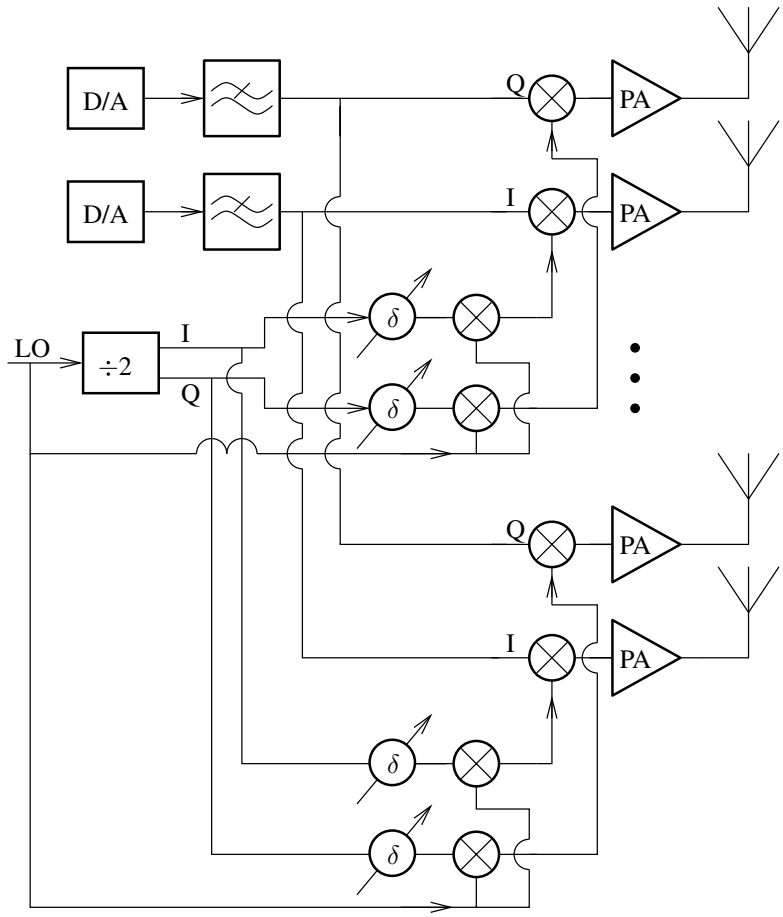


Figure 46: Offset local oscillator phase shifting for beamforming transmitter

A multi-phase LO at 19.2 GHz is produced by a tuned ring oscillator with 16 phases. The baseband signal is converted to a 24 GHz RF in two steps, the LO frequency is divided by 4 converting the baseband signal to 4.8 GHz IF in the first mixer. In the second up-converting stage, a phase selector decides which of the 16 LO phases to be used and thus in which direction the transmitted beam is directed.

The concept has been shown to perform well when it is integrated in silicon. One limiting factor of the design is the discrete number of phases. The phase step between two adjacent transmit paths sets the resolution in the beam direction, with 16 phase 22.5 degrees apart the resolution is 7 degrees[18]. To implement more phases, however, makes the selection network, and of course the oscillator, more complicated.

♢ Another implementation of the LO path is achieved by combining the circuits of Paper I and Paper IV [1,4]. A quadrature LO at half the carrier frequency is then used. The vector modulator in Paper I implements the phase shifting, and the the frequency doubler in Paper IV doubles the frequency before it is fed to the mixer at the PA input. An alternative is to perform the frequency multiplication after the mixer prior to the PA, or to replace the PA altogether. This would allow the mixer to run at half the carrier frequency as well. Using the latter approach, however, the amplitude in the signal path is non-linear [6].

6.4 Ring Oscillator Based Phase Shifting

This architecture uses a ring oscillator with a tunable phase shift between the oscillating elements [130–132], see Fig. 47. Each of the tuned oscillators in the ring is detuned from its center frequency. The LC-loads can sustain up to $\pm 90^\circ$ phase shift [130, 133]. The phase shift around the ring must always be equal to 360 degrees, or a multiple thereof. If each oscillating element is non-inverting, and no excess phase shift is introduced in the loop, the phase shift between consecutive elements is zero degrees. Introducing an excess phase shift of Kδ degrees will result in a phase shift of δ degrees in each of the K identical oscillators in the loop. This is the principle of this architecture, for further reading see [130, 131, 133].

In the literature several beamforming transmitters based on this architecture are reported, three of which are briefly presented below.

♢ The Hossein Hashemi Group at University of Southern California has suggested a variable phase ring oscillator system such as the one depicted in Fig. 47 [131]. The phase delay is achieved as explained above. In the transmit mode, the power amplifiers are connected directly to the tuned oscillators in the loop. If the floorplanning of the ring oscillator is optimized, the distribution network of the high

[18]The antenna spacing also effects the beam direction (62a). The 7 degrees resolution is for half a wavelength element spacing, which is commonly used since the grating lobes are then diminished.

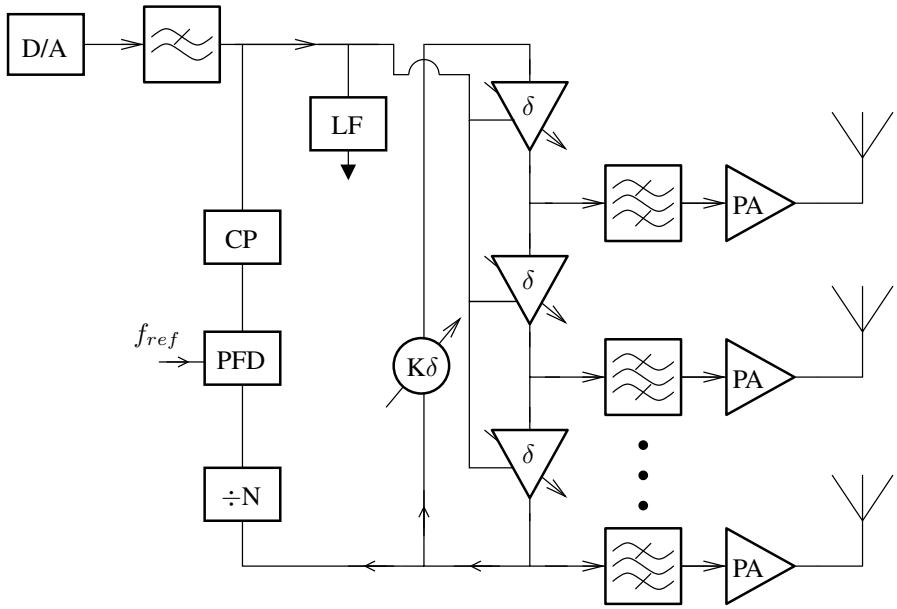


Figure 47: Transmitter architecture for variable phase ring oscillator in a phase locked loop

frequency oscillator signals can be minimized, which reduces complexity and losses. By placing the ring oscillator in a phase locked loop (PLL) the center frequency can be locked to a stable reference and the information signal can be added to the control voltages of the tuned elements (loop filter), to modulate the output frequency directly. The PLL response has to be slower than the information to be transmitted, otherwise the loop will cancel the information. Worth noticing in Fig. 47 is the absence of mixers, reducing the complexity and saving current.

♢ From the same group at University of Southern California that presented the variable phase ring oscillator, a slightly different approach is taken in [130], where two PLLs are used to transmit a UWB signal. The first PLL is locked to a crystal oscillator and thus produces a stable reference at the ring oscillator frequency. The UWB baseband signal is upconverted using the crystal locked PLL. The second PLL is designed as in Fig. 47, without frequency division. The second PLL has a wider bandwidth than the UWB signal. Thus the ring oscillator follows the UWB signal. The ring oscillator tuning signal can also be used to tune the following blocks in the RF path, allowing narrow band circuits

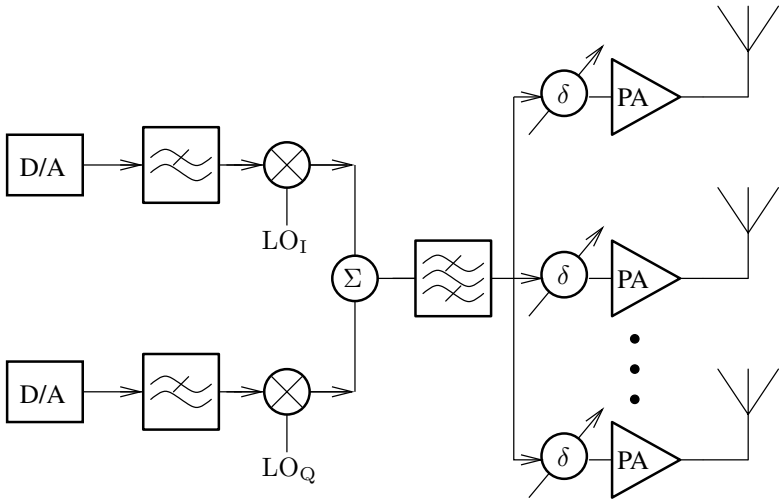


Figure 48: Transmitter architecture for phase shifting in the radio frequency path

to transmit a wideband signal. The architecture works well for FM modulated signals.

♢ In [134] an alternative implementation was presented. Here the phase shift is generated by a coupled array of oscillators, where one master oscillator couples to the others in the array. The phase shift is introduced by changing the control voltage of the slave oscillators.

Placing the master oscillator in a PLL the oscillation frequency will be stable, and the baseband phase information can be injected in the PLL just as in [130, 131].

6.5 Radio Frequency Phase Shifting

The most hardware efficient phase shifting, counting number of building blocks, is to perform it just before the power amplifier. The only circuit component that needs to be duplicated is the power amplifiers, see Fig. 48. The drawback is that the phase shifting takes place at the highest frequency and signal level in the system [135, 136]. The linearity of the phase shifters might be an issue if an envelope modulation scheme is used, but noise is of less importance since the power level is high.

Implementing the phase shifters at the highest frequency can actually be beneficial. If transmission lines are used as discrete phase shifters, they become shorter with frequency. This is a common architecture in radar systems (III-V semiconductor implementation). A set of fixed phase shifts is then implemented and switches controlled

by a selection logic choses the phase shift. Of course the transmission lines are linear so these phase shifters can very well be used in envelope modulated systems. Another advantage is that the delay is constant over a wide bandwidth.

Other RF phase shift architectures are described below, suitable for integration in CMOS.

- ♢ A beamforming transmitter for WLAN IEEE 802.11a/b/g is presented in [135]. In the RF phase shifters, the input signal is connected to two branches. In the upper branch a series inductance is placed before a source degenerated amplifying stage, and in the lower branch a series capacitance is used before another source degenerated stage. The series inductance and capacitance produce a 90 degree phase shift between the signals at the gates of the two amplifiers. The output of the two amplifiers are connected together in a common load. The difference in gain of the two amplifiers can be used to control the phase shift over a 90° range. Four of these vector modulators are placed in series to achieve a full 360° range. A look-up table is used to set the bias currents in the two branches, which determine the phase shift. A driver and power amplifier follows each phase shifter.

  Another common way to implement the phase shifters is to feed two variable gain amplifiers with a quadrature signal directly. With variable gain amplifier capable of both positive and negative gain only one vector modulator stage is needed to achieve a full 360° range [137, 138]. To generate the quadrature phases e. g. a polyphase filter can be used, see Paper V.

- ♢ In Paper I and Paper II an RF phase shifting topology is presented [1, 2], which uses vector modulators, in which quadrature phases are binary weighted and combined. Depending on the number of bits in the weighting function different phase resolution and amplitude accuracy can be achieved. The output phase is steerable over the full 360° range having a differential or single-ended output. Using these phase shifter at the highest frequency may introduce a too large a load capacitance, depending on the carrier frequency. Then the modulator could be at half the carrier frequency, reducing the capacitive load to half, and be followed by the frequency doubling amplifier described in Paper IV. No additional mixer or LO signal is introduced by this approach. The frequency doubling can be performed in the driver of the power amplifier.

Chapter 7

Summary of Included Papers

In this section a summary of the papers included in the next part is given. If not stated otherwise at the beginning of the included paper, the changes made compared to the original manuscript are restricted to syntax changes and redrawing of some figures.

- ♢ **Paper I: An 8-GHz Beamforming Transmitter IC in 130-nm CMOS [1]**
  An 8 GHz beamforming transmitter has been implemented in a 130 nm CMOS process and verified by measurements. It consists of a quadrature voltage controlled oscillator (QVCO) and two power amplifiers with 360° phase control range. The QVCO is an LC oscillator with cross coupled differential pairs biased by a tail current source. The high frequency current source noise is filtered out by a shunt capacitance, and the sources of the differential pairs are tuned to twice the resonance frequency by an inductor. The oscillator is isolated from the PAs by a buffer. The buffer and QVCO are tuned to the same frequency, which maximizes the voltage swing for a given bias current.

  To perform beamforming two power amplifiers with 360 degrees steerable output phase follows the buffer. The variable output phase is achieved by using binary weighted transistor banks on each of the quadrature phases. The weights are decided by a digital control word, and the outputs of the banks are summed. The PAs are intended to be connected to separate antennas. The power combination is thereby performed in the air. By combining the signals in the air an on-chip power combiner can be omitted, resulting in increased efficiency.

- ♢ **Paper II: A 24-GHz Automotive Radar Transmitter with Digital Beam Steering in 130-nm CMOS [2]**
  The paper introduces a transmitter architecture suitable for CMOS implementation of an automotive radar. To be able to deliver the power needed in a radar application using CMOS technology, multiple transmitting power amplifiers are necessary. In this paper it is shown that an array of 18 antenna elements is sufficient to achieve a 5 degree beam width, which is required by the application. The directivity is 12 dBi. The individual power amplifiers then need to transmit 0 dBm each to the antenna element.

  The topology features a QVCO and an active distribution network to the power amplifiers. The active distribution tree-network consists of transmission lines arranged as I+, I−, ground, Q+, Q− with buffers where the signal splits. The

buffers are implemented as self biased inverters. The power amplifier topology is the same as in Paper I. The power amplifiers occupy three sides of the chip to which the branches of tree-network reach. The power amplifiers have differential output signals.

♢ **Paper III: Antenna Array for a 24-GHz Automotive Radar with Dipole Antenna Element Patches [3]**
The paper presents 3D electromagnetic simulations of an antenna array. The antenna array consists of 24 dipole antenna element patches resonant at 24 GHz with $\lambda/2$ spacing, giving the array a physical size of 150 mm. A finite ground plane of 200×100 mm^2 has been assumed in simulations. The dipole patches are fed in the center by a differential signal. The simulations are performed assuming a standard off the shelf substrate, Arlon AD 1000. The directivity is 9 dBi with a standing wave ratio of less than 3 from 22.5 GHz to 24.5 GHz, when matched to 60 Ω. This impedance was chosen to fit the PAs of Paper II. The antenna element input impedance and standing wave ratio were optimized for a single element prior to the array simulations. Beam steering can be accomplished by changing the phase to the individual antenna elements, e. g. with a chip as the one described in Paper II.

♢ **Paper IV: 60 GHz 130-nm CMOS Second Harmonic Power Amplifiers [4]**
Two different frequency doubling power amplifiers have been measured, one with differential and one with single-ended input, but both with single-ended output tuned to 60 GHz. The amplifiers have been implemented in a 1p8M 130-nm CMOS process. The resonant nodes are tuned to 30 GHz or 60 GHz using on-chip transmission lines, which have been simulated in ADS and Momentum. The frequency doubling takes place in the first stage, using its non-linearity. The differential input topology uses the fact that the second harmonics of the differential signals are in phase. The drains of the first stage are therefore connected together, which efficiently performs the multiplication and at the same time cancels the fundamental and odd harmonics. The second stage is a 60 GHz common source amplifier.

The measured input impedance of the single-ended PA is high (250 Ω), and the layout of the differential input is similar, making the PA a suitable load for an oscillator in a fully integrated transmitter. The measured saturated output power delivered into 50 Ω for the single-ended and the differential input PA is 1 dBm and 3 dBm, respectively, both with a drain efficiency of 8%.

♢ **Paper V: Analysis of a High Frequency and Wide Bandwidth Active Polyphase Filter Based on CMOS Inverters [5]**
An active polyphase filter capable of high frequency quadrature generation has been investigated. The active polyphase filter is implemented with transconductors instead of resistors as in a passive polyphase filter implementation.

Analytical expressions for the voltage gain and the phase transfer function of a loaded stage have been derived. These expressions have been verified through transistor level simulations. By simplifying the model, assuming small signal and low frequency, the expressions have also been verified by comparison with results from an earlier publication [31]. A non-linear Verilog-A model has also been implemented to evaluate the sensitivity to transconductor phase delay on the performance of the active polyphase filter.

The investigation shows that high frequency quadrature generation is possible with high bandwidth, $\pm 40\%$ with less than 1° phase error. The robustness against process parameter spread was estimated by corner simulations. The quadrature phase error was highest, 4°, when both the n- and the pMOS transistors were slow. Statistical simulations over 500 runs yield a mean phase error of 0.2° with an associated standard deviation of 1.9°. The far out phase noise was less than that of the eqvivalent passive polyphase filter for offset frequencies above 10 MHz.

◇ **Paper VI: A 30 GHz 90-nm CMOS Passive Subharmonic Mixer with 15 GHz Differential LO [6]**

A new passive subharmonic mixer topology has been investigate and fabricated in a 90 nm CMOS technology. The mixer uses a differential local oscillator to perform the mixing function. This is accomplished by low pass filtering the inter-stage signal by a second order LC-filter, which suppresses frequencies above the LO (15 GHz). The new topology is compared to an established quadrature LO passive subharmonic mixer topology. The new topology suffers a few dBs in noise figure but has somewhat higher conversion gain. The higher conversion gain comes from the voltage gain in the interstage filter.

The new topology seems to be beneficial in architectures with multiple transmit/receive paths, where quadrature LO feeds can be both area and power consuming. Thus, beamforming transmitters would be an attractive application for this mixer topology.

Chapter 8

Discussion and Future Work

The results of the research in this thesis cover building blocks and architectures for CMOS beamforming transmitters at high frequencies. A power amplifier with 360° output phase at 8 GHz has been manufactured and measured. The measurements show that the circuit topology works satisfactory. The output phase setting is controllable digitally. The phase steering technique, used in the power amplifiers, is also suitable to use in a vector modulator to achieve arbitrary phases in the LO path of the transmitter. Two different frequency doubling power amplifier topologies at 60 GHz have been manufactured and measured. One topology with single-ended input, and the other with differential input, both with single-ended output. The frequency doubling is performed using the non-linearities of the transistors in the input common source stage. The differential input topology has an inherent fundamental and odd-order harmonic suppression, reducing the leakage from the local oscillator to the antenna.

When it comes to power amplifier design an interesting path to pursue would be the full integration of a multiple phase controlled power amplifier at 30–40 GHz. The amplifiers should be measured both electrically and with an antenna array in an anechoic chamber. In this thesis circuits with two power amplifiers (8 GHz) have been measured, and a circuit with 18 amplifiers has been manufactured (24 GHz). The results these show are promising, with an output phase that can be controlled arbitrarily to enable beam steering. A natural continuation would be to add the frequency doubling power amplifier, with more stages at 77 (60) GHz compared to [4], trying to fulfill the requirements of a 77 GHz automotive radar. Also the approach with ring oscillator phase shifting seems promising, since both transmitter and receiver can then use the same hardware, if time division duplex is used in the communication.

In Paper VI a passive subharmonic mixer topology is presented, requiring only a differential local oscillator. Once the mixer is verified it would be interesting to build a transmitter with the architecture outlined in the same paper. Both automotive radar and 60 GHz WLAN applications would be good targets. The attractiveness is an LO at a quarter of the carrier frequency and that only differential phase shifted signals need to be distributed across the chip to the frequency doubling power amplifiers.

The high frequency active polyphase filter could be further investigated by manufacturing prototypes for different frequency bands. The equations could then be compared to measured results as well as to the simulated [5]. An interesting challenge would be to design a multiband filter, which could be tuned to operate in different frequency bands, while maintaining good reliability and small quadrature mismatch.

Appendix A

Other Fabricated Chips

Most of the fabricated chips designed during the PhD studies are published and described in the paper section of this thesis or cited in the preface. Some chips, however, have not been presented before, and are therefore described in this appendix.

A.1 10-GHz 130-nm CMOS Dual Power Amplifier

A circuit consisting of a passive polyphase filter and two power amplifiers with beamforming capability was fabricated, see the chip microphotograph in Fig. 49. The power amplifiers are identical to those in Paper I. The passive polyphase filter consists of two stages to achieve a rather wide bandwidth, converting a differential input to a quadrature output, see Fig. 50. The first stage was tuned to a frequency slightly lower than 10 GHz, the second stage slightly higher. In front of the polyphase filter an L-match network was implemented to get a 50 Ω interface to the off-chip measurement equipment.

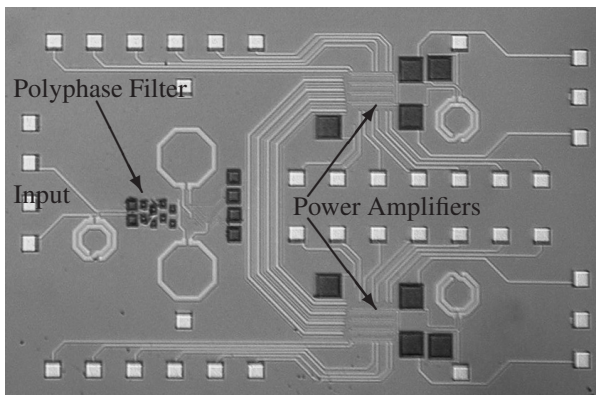


Figure 49: Microphotograph of the 10-GHz dual power amplifier

During the measurements it was observed that the chip did not transmit any signal from the input. The attenuation from input to output was in the order of 50–70 dB,

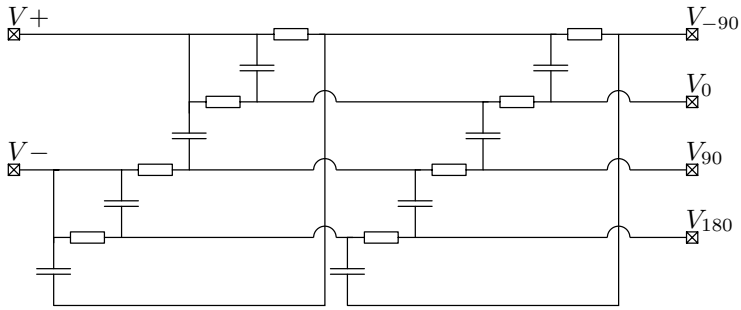


Figure 50: Schematic of the passive polyphase filter

independent of the power amplifier bias. The same behavior was observed over a frequency range from 5 GHz to 20 GHz. It is probably leakage, through the bulk and other paths, that is measured at the output.

The DC current through the power amplifiers change as expected with gate bias level, so the power amplifiers seem to work. Probably there is a short to ground somewhere between the input and the end of the polyphase filter. The input reflection coefficient, $S11$, was measured with a network analyzer and did not show good matching anywhere in the frequency range from 10 MHz to 20 GHz.

The layout has been thoroughly examined without finding any source for this behavior.

A.2 24-GHz 130-nm CMOS Power Amplifier

This chip was built using the same blocks as in Section A.1, but tuned to a higher input frequency. The output phase of the power amplifier is changed with a digital control word, which is switched in by a shift register in this design. The chip microphotograph can be seen in Fig. 51.

Also in this case no signal reached the output of the transmitter. The DC current from the supply changed with gate bias of the PA, and also the shift register function is verified through DC current measurements. A unit step in the current is observed as one more cascode transistor in the power amplifier is turned on.

This chip was intended as a test chip for a transmitter with 18 power amplifiers, see Fig. 52. These 18 power amplifier transmitters were intended as a prototype for an automotive radar transmitter [2, 3].

A new chip with an on-chip 24 GHz QVCO has been designed and fabricated, trying to circumvent the problem in the previous design. Measurements are ongoing.

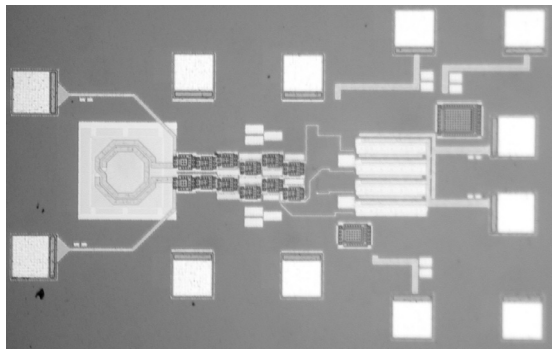


Figure 51: Microphotograph of the 24-GHz power amplifier

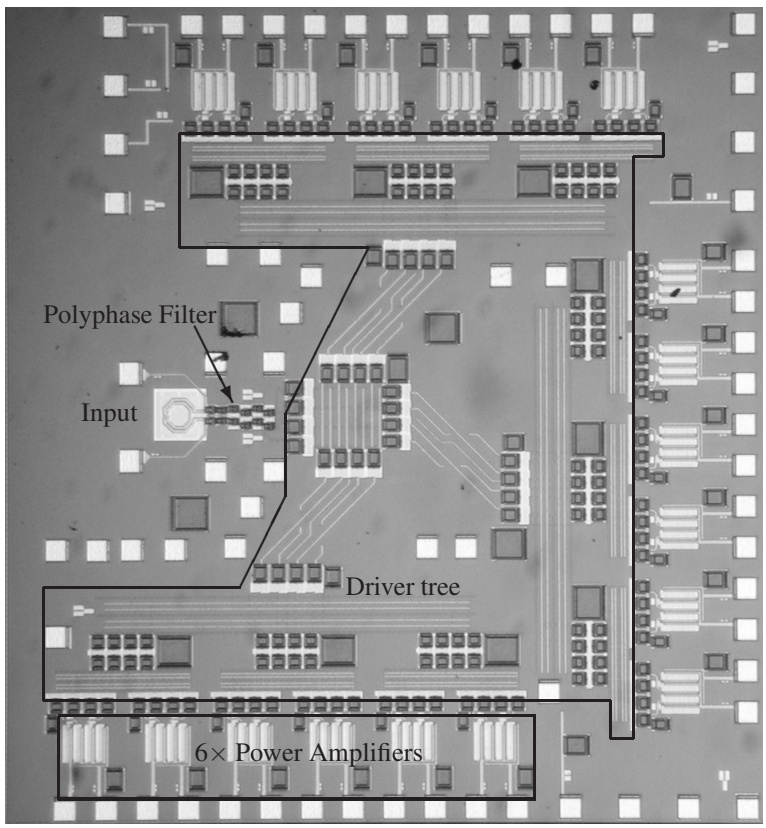


Figure 52: Microphotograph of the chip with 18 power amplifiers at 24-GHz

References

[1] J. Wernehag and H. Sjöland, "An 8-GHz Beamforming Transmitter IC in 130-nm CMOS," in *IEEE Radio Frequency Integrated Circuits Symposium*, July 2007, pp. 577–580.

[2] ——, "A 24-GHz Automotive Radar Transmitter with Digital Beam Steering in 130-nm CMOS," in *PhD Research in Microelectronics and Electronics*, June 2006, pp. 481–484.

[3] ——, "Antenna Array for a 24-GHz Automotive Radar with Dipole Antenna Element Patches," in *International Symposium on Antennas and Propagation*, Nov. 2006.

[4] ——, "60 GHz 130-nm CMOS Second Harmonic Power Amplifiers," in *IEEE Asia Pacific Conference on Circuits and Systems (APCCAS)*, Dec. 2008.

[5] ——, "Analysis of a High Frequency and Wide Bandwidth Active Polyphase Filter Based on CMOS Inverters," *Analog Integrated Circuits and Signal Processing*, Submitted.

[6] ——, "Differential Measurement and Parameter Extraction of Symmetrical Inductors," in *Proceedings of the Norchip Conference*, Nov. 2008.

[7] A. Henriksson and H. Tsu-Yü, *Kinesisk Historia*, 1st ed. Bonnier, 1978.

[8] G. Marconi, "Syntonic Wireless Telegraphy," *The Electrician*, no. 47, pp. 172–175 and 211–213, May 1901.

[9] B. Razavi, "Challenges in Portable RF Transceiver Design," *IEEE Circuits and Devices Magazine*, vol. 12, pp. 12–25, Sep. 1996.

[10] ——, *RF Microelectronics*. Upper Saddle River, New Jersey, USA: Prentice Hall, 1998.

[11] T. H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*, 1st ed. The Edinburgh Building, Cambridge CB2 2Ru, UK: Cambridge University Press, Inc, 1998.

[12] NationMaster, http://www.nationmaster.com/encyclopedia/Heterodyne, Oct. 2008.

[13] IEEE, "An Unsung hero: Reginald f."

[14] ——, "IEEE Communications Society Edwin Howard Armstrong Achievement Award," http://www.comsoc.org/ awards/armstrong.html, Oct. 2008.

[15] T. D. Stetzler, I. G. Post, J. H. Havens, and M. Koyama, "A 2.7-4.5 V Single Chip GSM Transceiver RF Integrated Circuit," *IEEE Journal of Solid-State Circuits*, vol. 30, no. 12, pp. 1421–1429, Dec. 1995.

[16] F. Arams, B. Peyton, F. Pace, R. Lange, and A. DiNardo, "Packaged 10.6-Micron Heterodyne 1-GHz Bandwidth Receiver," in *IEEE International Electron Devices Meeting, Technical Digest*, vol. 14, 1968, pp. 60–62.

[17] P. H. Young, *Electronic Communication Techniques*, 4th ed. Upper Saddle River, New Jersey, USA: Prentice Hall, 1999.

[18] D. K. Weaver, "A Third Method of Generation and Detection of Single-Sideband Signals," *Proceedings of the IRE*, vol. 44, no. 12, pp. 1703–1705, 1956.

[19] R. Hartley, "Single-sideband Modulator," *U.S. Patent*, no. 1,666,206, Apr. 1928.

[20] A. A. Abidi, "Direct-Conversion Radio Transceiver for Digital Communication," *IEEE Journal of Solid-State Circuits*, vol. 30, pp. 1399–1410, Dec. 1995.

[21] P. B. Khannur and K. S. Ling, "A 2.45GHz Fully-Differential CMOS Image-Reject Mixer for Bluetooth Application," *IEEE MTT-S International Microwave Symposium Digest*, vol. 1, pp. 549–552, June 2002.

[22] C.-C. Tang, C.-H. Wu, K.-H. Li, T.-C. Lee, and S.-I. Liu, "CMOS 2.4 GHz Receiver Front End with Area-Efficient Inductors and Digitally Calibrated 90° Delay Network," *IEE Proceedings Circuits, Devices and Systems*, vol. 150, pp. 439–444, Oct. 2003.

[23] L. Der and B. Razavi, "A 2-GHz CMOS Image-Reject Receiver With LMS Calibration," *IEEE Journal of Solid-State Circuits*, vol. 38, pp. 167–175, Feb. 2003.

[24] M. A. I. Elmala and S. H. K. Embabi, "Calibration of Phase and Gain Mismatches in Weaver Image-Reject Receiver," *IEEE Journal of Solid-State Circuits*, vol. 39, pp. 283–289, Feb. 2004.

[25] D. G. Tucker, "The History of the Homodyne and Synchrodyne," *The Journal of the British Institution of Radio Engineers*, vol. 14, pp. 143–154, Apr. 1954.

[26] F. M. Colebrook, "Homodyne," *Wireless World and Radio Review*, vol. 13, pp. 645–648, Feb. 1924.

[27] B. Razavi, "Design Considerations for Direct-Conversion Receivers," *IEEE Transactions on Circuits and Systems—Part II: Express Briefs*, vol. 44, pp. 428–435, June 1997.

[28] P. Andreani, A. Bonfanti, L. Romanò, and C. Samori, "Analysis and Design of a 1.8-GHz CMOS LC Quadrature VCO," *IEEE Journal of Solid-State Circuits*, vol. 37, pp. 1737–1747, Dec. 2002.

[29] J. D. K. West, "The application of the asymmetric polyphase filter in an SSB transceiver," in *Proceedings of the South African Communications and Signal Processing*, Aug. 1991, pp. 85–92.

[30] J. Crols and M. Steyaert, "An Analog Integrated Polyphase Filter for a High Performance Low-IF Receiver," in *IEEE Symposium on VLSI Circuits*, June 1995, pp. 87–88.

[31] F. Tillman and H. Sjöland, "A Polyphase Filter based on CMOS Inverters," in *Proceedings of the Norchip Conference*, Nov. 2005, pp. 12–15.

[32] J. J. Carr, *The Technician's Radio Receiver Handbook: Wireless and Telecommunication Technology*, 1st ed. Linacre House, Jordan Hill, Oxford OX2 8DP: Newnes, 2000.

[33] A. van der Ziel, *Noise in Solid State Devices and Circuits*. New York, USA: John Wiley and Sons, Inc., 1986.

[34] IEEE History Center, http://www.ieee.org/web/aboutus/history_center/biography/friis.html, Oct. 2008.

[35] Microwaves101, http://www.microwaves101.com/encyclopedia/hallofame2.cfm, Oct. 2008.

[36] B. Razavi, "RF Transmitter Architecture and Circuits," in *Proceedings of the IEEE Custom Integrated Circuits Conference, CICC*, May 1999, pp. 197–204.

[37] R. Kurorawa, "Injection Locking of Microwave Solid-State Oscillators," *Proceedings of the IEEE*, vol. 61, no. 10, pp. 1386–1410, Oct. 1973.

[38] E. Hegazi and A. A. Abidi, "A 17-mW Transmitter and Frequency Synthesizer for 900-MHz GSM Fully Integrated in 0.35-μm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 38, pp. 782–792, May 2003.

[39] H. Ishikuro, M. Hamada, K. ichi Agawa, S. Kousai, H. Kobayashi, D. M. Nguyen, and F. Hatori, "A Single-Chip CMOS Bluetooth Transceiver with 1.5MHz IF and Direct Modulation Transmitter," in *IEEE International Solid-State Circuits Conference, Digest of Technical Papers*, vol. 1, 2003.

[40] E. Löfgren and A. Mannesson, "Behaviour Models of a Bluetooth Transmitter A Top-Down Approach," Master's thesis, Lund University, Faculty of Engineering, Nov. 2005.

[41] D. Leeson, "A Simple Model of Feedback Oscillator Noise Spectrum," *Proceedings of the IEEE*, vol. 54, pp. 329–330, Feb. 1966.

[42] S. C. Cripps, *Rf Power Amplifiers For Wireless Communications*. Norwood MA, USA: Artech House Books, 1999.

[43] P. R. Gray, P. J. Hurst, S. H. Lewis, and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 4th ed. New York, USA: John Wiley and Sons, Inc., 2001.

[44] W. J. Chudobiak and D. F. Page, "Frequency and Power Limitations of Class-D Transistor Amplifiers," *IEEE Journal of Solid-State Circuits*, vol. 4, no. 1, pp. 25–37, Feb. 1969.

[45] N. O. Sokal and A. D. Sokal, "Class E-A New Single-Ended Class of High-Efficiency Tuned Switching Power Amplifiers," *IEEE Journal of Solid-State Circuits*, vol. 10, no. 3, pp. 168–176, June 1975.

[46] N. Aoki, Y. Asano, M. Minowa, and T. Takano, "Improved Output Spectrum of Linearized Class-F Power Amplifier for Digital Cellular Mobile Communications," in *IEEE International Conference on Universal Personal Communications*, Sep. 1992, pp. 195–199.

[47] F. H. Raab, "Class-F Power Amplifiers with Maximally Flat Waveforms," *IEEE Transactions on Microwave Theory and Techniques*, vol. 45, no. 11, pp. 2007–2012, Nov. 1997.

[48] S. Burrow and D. Grant, "Efficiency of Low Power Audio Amplifiers and Loudspeakers," *IEEE Transactions on Consumer Electronics*, vol. 47, no. 3, pp. 622–630, Aug. 2001.

[49] D. J. Kessler and M. K. Kazimierczukt, "Power Losses and Efficiency of Class-E Power Amplifier at Any Duty Ratio," *IEEE Transactions on Circuits and Systems—Part I: Regular Papers*, vol. 51, no. 9, pp. 1675–1689, Sep. 2004.

[50] M. K. Kazimierczukt, "Analysis of Class E Zero-Voltage-Switching Rectifier," *IEEE Transactions on Circuits and Systems*, vol. 37, no. 6, pp. 747–755, June 1990.

[51] M. Riordan, "How Europe Missed The Transistor," *IEEE Spectrum*, no. 11, pp. 46–51, Nov. 2005.

[52] A. V. Dormael, "The 'French' Transistor," *IEEE History Center*.

[53] M. J. Riezenman, "Wanlass's CMOS Circuit," *IEEE Spectrum*, no. 5, p. 44, May 1991.

[54] F. M. Wanlass, "Low Stand-By Power Complementary Field Effect Circuitry," *U.S. Patent*, no. 3,356,858, Dec. 1967.

[55] A. A. Abidi, "RF CMOS Comes Of Age," *IEEE Journal of Solid-State Circuits*, vol. 39, no. 4, Apr. 2004.

[56] F. H. Musa and R. C. Huntington, "A CMOS 3-1/2-Digit A/D Converter," in *IEEE International Solid-State Circuits Conference, Digest of Technical Papers*, vol. 19, Feb. 1976, pp. 144–145.

[57] W. Black, J. R. H. McCharles, and D. A. Hodges, "CMOS Process for High-Performance Analog LSI," in *IEEE International Electron Devices Meeting, Technical Digest*, vol. 22, 1976, pp. 331–334.

[58] J. Y.-C. Chang, A. A. Abidi, and M. Gaitan, "Large Suspended Inductors on Silicon and Their Use in a 2-μm CMOS RF Amplifier," *IEEE Electron Device Letters*, vol. 14, no. 5, May 1993.

[59] Y. P. Tsividis, *Operation and Modeling of the MOS Transistor*, 2nd ed. 198 Madison Avenue, New York: Oxford University Press, Inc, 1999.

[60] B. G. Streetman, *Solid State Electronic Devices*, 4th ed. Prentice-Hall Inc, 1995.

[61] J. M.Rabaey, *Digital Integrated Circuits, A Design Perspective*, 1st ed. Upper Saddle River, New Jersey 07458: Prentice-Hall Inc, 1996.

[62] C. Turchetti, "Relationships for the drift and the diffusion components of the drain current in a MOS transistor," *Electronics Letters*, vol. 19, pp. 960–962, Nov. 1983.

[63] K. Martin and D. A. Johns, *Analog Integrated Circuit Design*, 1st ed. John Wiley and Sons, Inc., 1997.

[64] M. Bagheri and Y. Tsividis, "A Small Signal dc-to-High-Frequency Nonquasistatic Model for the Four-Terminal MOSFET Valid in All Regions of Operation," *IEEE Transactions on Electron Devices*, vol. 32, no. 11, pp. 2383–2391, Nov. 1985.

[65] C. H. Doan, S. Emami, A. M. Niknejad, and R. W. Brodersen, "Millimeter-wave CMOS device modeling and simulation," in *Proceedings of the IEEE International Symposium on Circuits and Systems, ISCAS*, vol. 5, May 2004, pp. V–524–V–527.

[66] J. A. Kittl, "Salicides and alternative technologies for future ICs: Part 1," *Journal Solid State Technology*, pp. 81–92, June 1999.

[67] IBM, "IBM Advancement to Spawn New Generation of Chips," http://www-03.ibm.com/press/us/en/pressrelease/20980.wss, Jan. 2007.

[68] Intel, "Intel's Transistor Technology Breakthrough Represents Biggest Change to Computer Chips In 40 Years," http://www.intel.com/press-room/archive/releases/20070128comp.htm, Jan. 2007.

[69] B. Razavi, R.-H. Yan, and K. F. Lee, "Impact of Distributed Gate Resistance on the Performance of MOS Devices," *IEEE Transactions on Circuits and Systems—Part I: Regular Papers*, vol. 41, pp. 750–754, Nov. 1994.

[70] R. R. J. Vanoppen, J. A. M. Geelen, and D. B. M. Klaassen, "The High-Frequency Analogue Performance of MOSFETs," in *IEEE International Electron Devices Meeting, Technical Digest*, Dec. 1994, pp. 173–176.

[71] R. R. Troutman and A. G. Fortino, "Simple Model for Threshold Voltage in Short-Channel IGFET," *IEEE Transactions on Electron Devices*, vol. ED-24, pp. 1266–1268, Oct. 1977.

[72] R. R. Troutman, "VLSI Limitations from Drain-Induced Barrier Lowering," *IEEE Journal of Solid-State Circuits*, vol. SC-14, pp. 383–391, Apr. 1979.

[73] F. N. Trofimenkoff, "Field-Dependent Mobility Analysis of the Field-Effect Transistor," *Proceedings of the IEEE*, vol. 53, pp. 1765–1766, Jan. 1965.

[74] B. Hoeneisen and C. A. Mead, "Current-Voltage Characteristics of Small Size MOS Transistor," *IEEE Transactions on Electron Devices*, vol. ED-19, pp. 382–383, 1972.

[75] G. Groeseneken, "Hot Carrier Degradation and ESD in Submicron CMOS Technologies: How do they Interact?" in *IEEE Proceedings of the Electrical Overstress/Electrostatic Discharge Symposium*, Sep. 2000, pp. 276–286.

[76] Y. Leblebici and S.-M. Kang, "Modeling of nMOS Transistors for Simulation of Hot-Carrier-Induced Device and Circuit Degradation," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 11, no. 2, Feb. 1992.

[77] Chartered, "Chartered's RF CMOS Manufacturing Process Achieves First-Pass Success for AMIC Technology's 2.4 GHz RF Transceiver," http://ir.charteredsemi.com/phoenix.zhtml?c=93866&p=irol-newsArticle&t=Regular&id=457943&, Sep. 2001.

[78] J. C. Guo, C. H. Huang, K. T. Chan, W. Y. Lien, C. Wu, and Y. Sun, "0.13μm Low Voltage Logic Based RF CMOS Technology with 115GHz f_t and 80GHz f_{max}," in *European Microwave Conference*, Oct. 2003, pp. 683–686.

[79] L. Tiemeijer, H. Boots, R. Havens, A. Scholten, P. de Vreede, P. Woerlee, A. Heringa, and D. Klaassen, "A record high 150 GHz f_{max} realized at 0.18 μm gate length in an industrial RF-CMOS technology," in *IEEE International Electron Devices Meeting, Technical Digest*, Dec. 2001, pp. 10.4.1–10.4.4.

[80] H. Li, B. Jagannathan, J. Wang, T.-C. Su, S. Sweeney, J. J. Pekarik, Y. Shi, D. Greenberg, Z. Jin, R. Groves, L. Wagner, and S. Csutak, "Technology Scaling and Device Design for 350 GHz RF Performance in a 45nm Bulk CMOS Process," in *Symposium on VLSI Technology, Digest of Technical Papers*, June 2007, pp. 56–57.

[81] W. Jeamsaksiri, D. Linten, S. Thijs, G. Carchon, J. Ramos, A. Mercha, X. Sun, P. Soussan, M. Dehan, T. Chiarella, R. Venegas, V. Subramanian, A. Scholten, P. Wambacq, R. Velghe, G. Mannaert, N. Heylen, R. Verbeeck, W. Boullart, I. Heyvaert, M. I. Natarajan, G. Groeseneken, I. Debusschere, S. Biesemans, and S. Decoutere, "A low-cost 90nm RF-CMOS platform for record RF circuit performance," in *Symposium on VLSI Technology, Digest of Technical Papers*, June 2005, pp. 60–61.

[82] W. Jeamsaksiri, A. Mercha, J. Ramos, D. Linten, S. Thijs, S. Jenei, C. Detcheverry, P. Wamhacq, R. Velghe, and S. Decoutere, "Integration of a 90nm RF CMOS technology (200GHz f_{max} – 150GHz f_t NMOS) demonstrated on a 5GHz LNA," in *Symposium on VLSI Technology, Digest of Technical Papers*, June 2004, pp. 100–101.

[83] I. Post, M. Akbar, G. Curello, S. Gannavaram, W. Hafez, U. Jalan, K. Komeyii, J. Lin, N. Lindert, J. Park, J. Rizk, G. Sacks, C. Tsai, D. Yeh, P. Bai, and C.-H. Jan, "A 65nm CMOS SOC Technology Featuring Strained Silicon Transistors for RF Applications," in *IEEE International Electron Devices Meeting, Technical Digest*, Dec. 2006, pp. 1–3.

[84] A. M. Niknejad and H. Hashemi, Eds., *mm-Wave Silicon Technology, 60GHz and Beyond*, 1st ed. Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA: Springer, Feb. 2008.

[85] G. Dambrine, C. Raynaud, D. Lederer, M. Dehan, O. Rozeaux, M. Vanmackelberg, F. Danneville, S. Lepilliet, and J.-P. Raskin, "What are the Limiting Parameters of Deep-Submicron MOSFETs for High Frequency Applications?" *IEEE Electron Device Letters*, vol. 24, no. 3, pp. 189–191, Mar. 2003.

[86] S. Franco, *Design with Operational Amplifiers and Analog Integrated Circuits*, 3rd ed. The McGraw-Hill Companies, Schaum's, 11 West 19th Street, New York: McGraw-Hill, 2002.

[87] A. S. Roy and C. C. Enz, "Compact Modeling of Thermal Noise in the MOS Transistor," *IEEE Transactions on Electron Devices*, vol. 52, no. 4, pp. 611–614, Apr. 2005.

[88] A. A. Abidi, "High-Frequency Noise Measurements on FETŠs with Small Dimension," *IEEE Transactions on Electron Devices*, vol. 33, no. 11, pp. 1801–1805, Nov. 1986.

[89] M. Valenza, A. Hoffmann, D. Sodini, A. Laigle, F. Martinez, and D. Rigaud, "Overview of the impact of downscaling technology on $1/f$ noise in p-MOSFETs to 90nm," *IEE Proceedings Circuits, Devices and Systems*, vol. 151, no. 2, pp. 102–110, Apr. 2004.

[90] L. K. J. Vandamme, X. Li, and D. Rigaud, "l/f Noise in MOS Devices, Mobility or Number Fluctuations?" *IEEE Transactions on Electron Devices*, vol. 41, no. 11, pp. 1936–1945, 1994.

[91] H. E. Maes, S. H. Usamani, and G. Groeseneken, "Correlation between $1/f$ noise and interface state density at the Fermi level in field effect transistors," *Journal of Applied Physics*, vol. 57, no. 10, pp. 4811–4813, 1985.

[92] L. Vandamme, "Model for $1/f$ Noise in MOS Transistors Biased in Linear Region," *IEEE Solid-State Electronics*, vol. 23, pp. 317–323, 1980.

[93] P. Andreani and H. Sjöland, "Noise Optimization of an Inductively Degenerated CMOS Low Noise Amplifier," *IEEE Transactions on Circuits and Systems—Part II: Express Briefs*, vol. 48, no. 9, pp. 835–841, Sep. 2001.

[94] A. F. Molisch, *Wireless Communications*, 1st ed. West Sussex, PO19 8SQ, England: John Wiley and Sons, Inc., 2006.

[95] UMTS World, http://www.umtsworld.com/technology/linkbudget.htm, Apr. 2008.

[96] H. Fischer, F. Henkel, M. Engels, and P. Waldow, "UMTS/GSM Multi Mode Receiver Design," in *Proceedings of the IEEE International Symposium on Circuits and Systems, ISCAS*, vol. 4, no. 8, May 2004, pp. 261–264.

[97] Jussila, J. Ryynänen, K. Kivekäs, L. Sumanen, A. Pärssinen, and K. A. I. Halonen, "A 22-mA 3.0-dB NF Direct Conversion Receiver for 3G WCDMA," *IEEE Journal of Solid-State Circuits*, vol. 36, no. 12, pp. 2025–2029, Dec. 2001.

[98] J. Ryynanen, K. Kivekas, J. Jussila, A. Parssinen, and K. Halonen, "A Dual-Band RF Front-End for WCDMA and GSM Applications," *IEEE Journal of Solid-State Circuits*, vol. 36, no. 8, pp. 1198–1204, Aug. 2001.

[99] B. Razavi, "Gadgets Gab at 60 GHz," *IEEE Spectrum*, vol. 45, no. 2, pp. 40–45, Feb. 2008.

[100] S. K. Moore, "Cheap Chips for Next Wireless Frontier," *IEEE Spectrum*, vol. 43, no. 6, pp. 8–9, June 2006.

[101] B. Razavi, "A Millimeter-Wave CMOS Heterodyne Receiver With On-Chip LO and Divider," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 2, pp. 477–485, Feb. 2008.

[102] J. D. Kraus and R. J. Marhefka, *Antennas: For All Applications*, 3rd ed. New York: McGraw-Hill, 2002.

[103] B. Razavi, "A 60-GHz CMOS Receiver Front-End," *IEEE Journal of Solid-State Circuits*, vol. 41, pp. 17–22, Jan. 2006.

[104] P. Smulders, "Exploiting the 60 GHz Band for Local Wireless Multimedia Access: Prospects and Future Directions," *IEEE Communications Magazine*, vol. 40, pp. 140–147, Jan. 2002.

[105] H. Hashemi, X. Guan, and A. Hajimiri, "A Fully Integrated 24GHz 8-Path Phased-Array Receiver in Silicon," in *IEEE International Solid-State Circuits Conference, Digest of Technical Papers*, Feb. 2004, pp. 390–391.

[106] H. Krishnaswamy and H. Hashemi, "Effect of Process Mismatches on Integrated CMOS Phased Arrays Based on Multiphase Tuned Ring Oscillators," *IEEE Transactions on Microwave Theory and Techniques*, vol. 56, no. 6, pp. 1305–1315, June 2008.

[107] T. Hirose, "Study of mm wave propagation modeling to realize WPANs," *IEEE802.15-03/0365*, Sep. 2003.

[108] H. Yang, M. H. A. J. Herben, and P. F. M. Smulders, "Impact of Antenna Pattern and Reflective Environment on 60 GHz Indoor Radio Channel Characteristics," *IEEE Antennas and Wireless Propagation Letters*, vol. 4, pp. 300–303, 2005.

[109] P. Smulders and L. Correia, "Characterisation of propagation in 60 GHz radio channels," *Electronics & Communication Engineering Journal*, vol. 9, pp. 73–80, Apr. 2007.

[110] M. Klepal, "Antennas and Propagation for Wireless Communication," http://www.aws.cit.ie/mklepal/AntennasAndPropagation/two_ray_model.htm.

[111] C. H. Doan, S. Emami, D. A. Sobel, A. M. Niknejad, and R. W. Brodersen, "Design considerations for 60 GHz CMOS radios," *IEEE Communications Magazine*, vol. 42, pp. 132–140, Dec. 2004.

[112] Federal Communications Commission, *The FCC rules and regulations are codified in Title 47 of the Code of Federal Regulations*, 47CFR15.255, Oct. 2001.

[113] C. A. Balanis, *Antenna Theory: Analysis and Design*, 2nd ed. New York: John Wiley and Sons, Inc., 1996.

[114] R. J. Mailloux, *Phased Array Antenna Handbook*, 2nd ed. Norwood, MA: Artech House Books, 2005.

[115] ——, "Phased Array Theory and Technology," *Proceedings of the IEEE*, vol. 70, no. 3, pp. 246–291, Mar. 1982.

[116] A. Hajimiri, H. Hashemi, A. Natarajan, X. Guan, and A. Komijani, "Integrated Phased Array Systems in Silicon," *Proceedings of the IEEE*, vol. 93, no. 9, pp. 1637–1655, Sep. 2005.

[117] J. E. Rhodes, *Antenna Handbook*. Quantico: Marine Corps Doctrinal Publications Status, 1999.

[118] D.-C. Chang, C.-I. Hung, and W.-K. Huang, "Development of X-band Space Feed Phased Array Antenna System," in *International Symposium Antennas and Propagation Society, 'Merging Technologies for the 90's'. Digest.*, vol. 2, May 1990, pp. 706–709.

[119] R. J. Mailloux, "Antenna Arrav Architecture," *Proceedings of the IEEE*, vol. 80, no. 1, pp. 163–172, Jan. 1992.

[120] D. Busuioc, M. Shahabadi, A. Borji, and S. Safavi-Naeini, "G/T of Multi-Segment Active Arrays-Dependence on Array Configuration," in *International Symposium Antennas and Propagation Society*, vol. 1, June 2004, pp. 133–136.

[121] D. J. Rabideau and P. Parker, "Ubiquitous MIMO Multifunction Digital Array Radar," in *Conference on Signals, Systems and Computers, Conference Record of the Thirty-Seventh Asilomar*, vol. 1, Nov. 2003, pp. 1057–1064.

[122] D. G. Rahn, M. S. Cavin, F. F. Dai, N. H. W. Fong, R. Griffith, J. Macedo, A. D. Moore, J. W. M. Rogers, and M. Toner, "A Fully Integrated Multiband MIMO WLAN Transceiver RFIC," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 8, pp. 1629–1641, Aug. 2005.

[123] Y. Palaskas, A. Ravi, S. Pellerano, B. R. Carlton, M. A. Elmala, R. Bishop, G. Banerjee, R. B. Nicholls, S. K. Ling, N. Dinur, S. S. Taylor, and K. Soumyanath, "A 5-GHz 108-Mb/s 2 2 MIMO Transceiver RFIC With Fully Integrated 20.5-dBm P1dB Power Amplifiers in 90-nm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 12, pp. 2746–2756, Dec. 2006.

[124] A. Natarajan, A. Komijani, and A. Hajimiri, "A Fully Integrated 24-GHz Phased-Array Transmitter in CMOS," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 12, pp. 2502–2514, Dec. 2005.

[125] H. Hashemi, X. Guan, A. Komijani, and A. Hajimiri, "A 24-GHz SiGe Phased-Array Receiver-LO Phase-Shifting Approach," *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, no. 2, pp. 614–626, Feb. 2005.

[126] A. Natarajan, A. Komijani, and A. Hajimiri, "A Fully Integrated 24-GHz Eight-Element Phased-Array Receiver in Silicon," *IEEE Journal of Solid-State Circuits*, vol. 39, no. 12, pp. 2311–2320, Dec. 2004.

[127] T. Yamaji, D. Kurose, O. Watanabe, S. Obayashi, and T. Itakura, "A Four-Input Beam-Forming Downconverter for Adaptive Antennas," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 10, pp. 1619–1625, Oct. 2003.

[128] S. Gueorguiev, S. Lindfors, and T. Larsen, "A 5.2 GHz CMOS I/Q Modulator With Integrated Phase Shifter for Beamforming," *IEEE Journal of Solid-State Circuits*, vol. 42, no. 9, pp. 1953–1962, Sep. 2007.

[129] ——, "A CMOS transmitter for 802.11a WLAN with beam forming capability," in *Midwest Symposium on Circuits and Systems*, vol. 2, Aug. 2005, pp. 1605–1608.

[130] H. Krishnaswamy and H. Hashemi, "A 4-Channel 24-27 GHz UWB Phased Array Transmitter in 0.13μm CMOS for Vehicular Radar," in *Proceedings of the IEEE Custom Integrated Circuits Conference, CICC*, June 2007, pp. 753–756.

[131] ——, "A Fully Integrated 24GHz 2-Channel Phased-Array Ttranceiver in 0.13μm CMOS Based on a Variable-Phase Ring Oscillator and PLL Architecture," in *IEEE International Solid-State Circuits Conference, Digest of Technical Papers*, Feb. 2007, pp. 753–756.

[132] ——, "Variable-Phase Ring-Oscillator Arrays, Architectures, and Related," *USPatent*, no. US2007/02909 A1, Dec. 2007.

[133] K. D. Stephan, "Inter-Injection-Locked Oscillators for Power Combining and Phased Arrays," *IEEE Transactions on Microwave Theory and Techniques*, vol. 34, no. 10, pp. 1017–1025, Oct. 1986.

[134] R. A. York and T. Itoh, "Injection- and Phase-Locking Techniques for Beam Control," *IEEE Transactions on Microwave Theory and Techniques*, vol. 46, no. 11, pp. 1920–1929, Nov. 1998.

[135] D. R. Banbury, N. Fayyaz, S. Safavi-Naeini, and S. Nikneshan, "A CMOS 5.5/2.4 GHz dual-band smart-antenna transceiver with a novel RF dual-band phase shifter for WLAN 802.11a/b/g," in *IEEE Radio Frequency Integrated Circuits Symposium*, June 2004, pp. 157–160.

[136] S. Hamedi-Hagh and C. A. T. Salama, "CMOS Wireless Phase-Shifted Transmitter," *IEEE Journal of Solid-State Circuits*, vol. 39, no. 8, pp. 1241–1252, Aug. 2004.

[137] S. Gueorguiev, "Integrated Circuit Techniques and Architectures for Beamforming Radio Transmitters," Ph.D. dissertation, Aalborg University, Aalborg University, Fredrik Bajers Vej 5, P.O. Box 159, DK–9100 Aalborg, Denmark, Aug. 2006.

[138] C. Cao, Y. Ding, X. Yang, J.-J. Lin, H.-T. Wu, A. K. Verma, J. Lin, F. Martin, and K. K. O, "A 24-GHz Transmitter With On-Chip Dipole Antenna in 0.13-μm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 6, pp. 1394–1402, June 2008.

Paper I

Paper I

An 8-GHz Beamforming Transmitter IC in 130-nm CMOS

Abstract

An 8-GHz beamforming transmitter IC has been designed in a 130-nm CMOS process. Two power amplifiers with independently controllable phase enable the beamforming. The phases are digitally controllable over the full 360° range, which is accomplished by binary weighting of quadrature phase signals in the power amplifiers. The quadrature phase signals are generated by a quadrature voltage controlled oscillator followed by a buffer, which serves as an isolation between the power amplifiers and the oscillator. The chip contains seven on-chip differential inductors, and consumes a total of 47 mA from a 1.0 V supply. The measured output power is −3 dBm for each power amplifier.

Based on: Johan Wernehag and Henrik Sjöland, "An 8-GHz Beamforming Transmitter IC in 130-nm CMOS", *IEEE Radio Frequency Integrated Circuits Symposium*, pp. 577–580, June 3–5, 2007, Honolulu, HI, USA.

1 Introduction

This work is part of a project directed to find transceiver architectures and circuit topologies that can cope with the trend of scaling CMOS technology, which leads to shorter gate lengths, higher transit frequencies, lowered supply voltages, decreased dynamic range for receivers, and less output power for transmitters. These changes put new demands designing analog circuits, but also enables the use of CMOS at higher frequencies. New applications for RF CMOS could be automotive radar at 24 and 77 GHz [1–3] and indoor WLAN at 60 GHz [4–6]. This project is aimed at 60 GHz CMOS WLAN.

A 130-nm CMOS 8-GHz beamforming transmitter IC is presented, consisting of a quadrature voltage controlled oscillator (QVCO), a tuned buffer, and two power amplifiers (PAs) (Fig. 1). The output phase of each PA is controllable through 360°

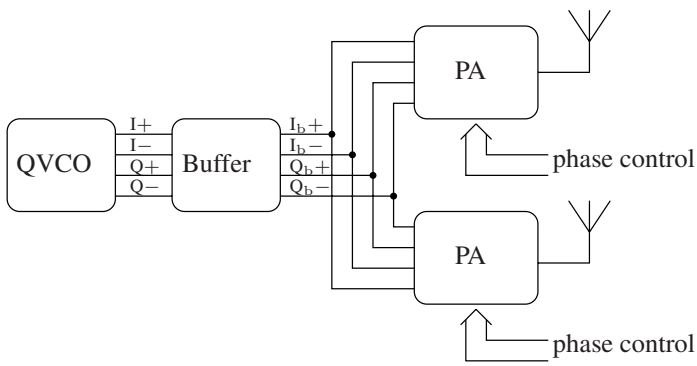


Figure 1: Block diagram of the beamforming transmitter

to enable beamforming. Beamforming and electronic scanning are suitable for radar system and can also be used in WLAN systems. In a WLAN system it is beneficial to steer the antenna lobe of the transmitter towards the intended receiver. Less power then needs to be transmitted, saving battery energy and relaxing the requirements on the PAs. Less interference to other communication devices is also generated, increasing the total system capacity.

2 Circuit Topology

The PAs (Fig. 2) are designed to have high gain and isolation without having too large input capacitance, since this capacitance has to be driven by the buffer (included in its resonance tank).

The buffer (Fig. 3(b)) is needed to ensure that the QVCO (Fig. 3(a)) has a fixed load impedance. Without it the QVCO would be affected by the output signal of

the PAs, because the large signal swings of the PAs drive the cascode transistors into triode region for parts of the signal cycle, diminishing the isolation of the cascodes. A changing antenna impedance caused by a changing environment would thus effect the QVCO. Furthermore, the input impedance of a PA changes as its phase is changed by a new phase control word. All this would degrade the accuracy of the quadrature signal generated by the QVCO, and thereby also the accuracy of the beamforming. An even worse problem is that a shifting load would shift the frequency of the QVCO, and of the entire transmitter. This effect could, however, be significantly suppressed by locking the frequency of the QVCO to a crystal reference using a PLL frequency synthesizer.

The QVCO operates at a rather high amplitude. This brings down the phase noise and ensures a reliable start-up. In between each building block there are capacitive taps, which decreases the loading of the previous stage and increases the isolation. The tap consists of the input capacitance of the following stage and an additional series capacitance, which also performs an AC-coupling.

2.1 Power Amplifier

The PA consists of binary weighted transistors for each of the four quadrature phases. They can be switched on and off by cascode transistors. The cascode transistors can also be used to select the direction of the signal current, towards the positive or negative differential output (Fig. 2). By switching on and off and directing the current of the transistors in different combinations, the phase of the output is controllable through 360°, and thus the antenna lobe is steerable in any direction. When the current of transistor V_{x_y} is directed towards the positive output branch, the current of transistor V_{x+180_y} should be directed towards the negative output branch, see Fig. 2. This switching scheme ensures a differential output of the PA.

This circuit has a control word length of three (transistors weighted by 1, 2 and 4) leading to a rather large phase discretization. This makes it manageable to sweep the phase directly without using a shift register or digital encoding of control words. There is no restriction to increasing the resolution by adding smaller transistors. The output power will vary somewhat for different phase settings, and an increased phase resolution also has the benefit of reducing this power variation.

2.2 Quadrature Voltage Controlled Oscillator and Buffer

The topology of the oscillator and buffer is almost the same as in [7]. This topology (Fig. 3) is chosen because it is robust and has a high quadrature accuracy. The phase noise performance is also good thanks to the series connected transistors [8], and the inductor tuning the source node of the VCO to twice the frequency of oscillation [9].

A difficulty when designing an oscillator at this high frequency is that the reactance of the resonance tank is quite low. It is further reduced by the capacitive load of the buffer. The inductance of the resonance tank must therefore be very small. The

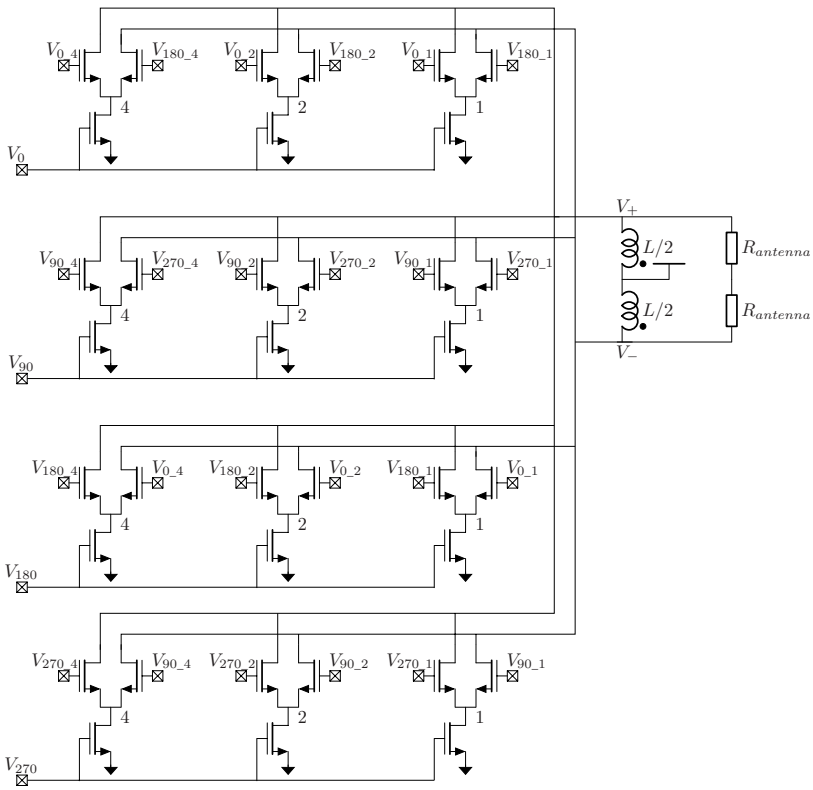


Figure 2: Schematic of the power amplifier with three binary weighted transistors per bank

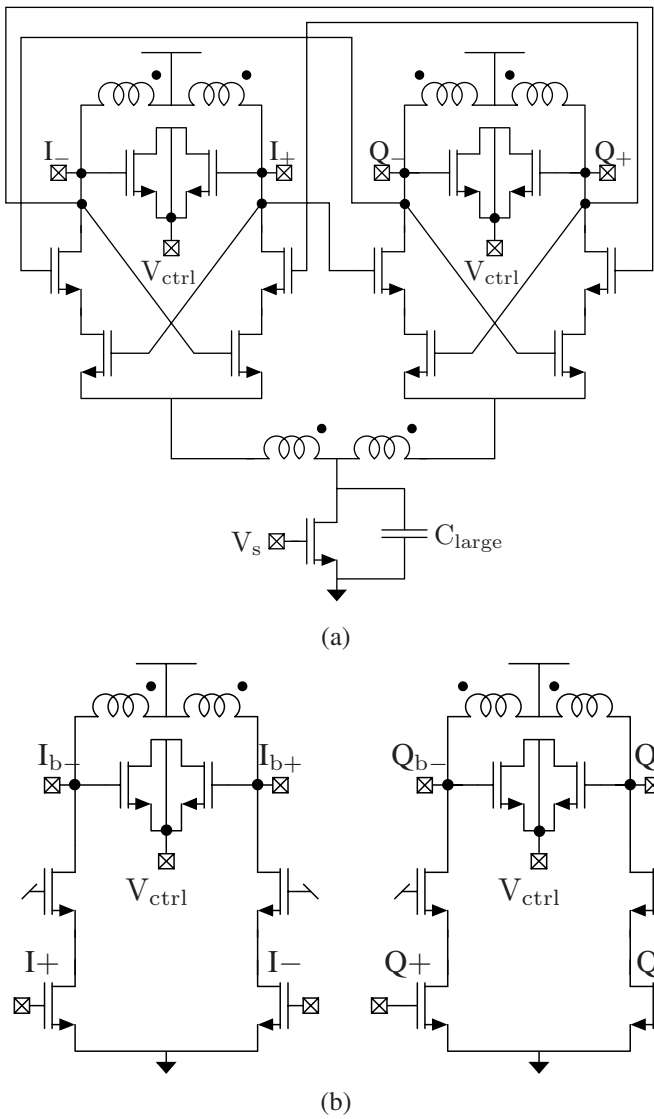


Figure 3: (a) Topology of the QVCO (b) Topology of the tuned buffer

limit to how small it can be made is set by the parasitic inductance. Since the tank is fully integrated using short on-chip wires, a 630 pH differential inductor could be used.

The buffer is separately tuned (Fig. 3(b)) with the PAs as load. The inductance is slightly smaller due to the heavy loading of the gate capacitance of the PAs. A varactor is used also in the buffer to maximize its output voltage swing over the operating frequency range, and to reduce its current consumption, [7]. The same control voltage, V_{ctrl}, is used in both the QVCO and the buffer.

2.3 On-Chip Inductors

The process has eight metal layers in copper, above a lightly doped substrate. Only differential inductors are used in this fully differential circuit. Two inductors at the output of the PAs form resonance tanks with the parasitic drain-bulk capacitance of the cascode transistors. The drain current to the PAs is also fed through these inductors. Two inductors each are used in the QVCO and buffer resonance tanks. Finally one inductor is used at the source nodes of the switching pairs in the QVCO.

The quality factor (Q) of the differential inductors is quite high thanks to the excellent process, resulting in a low power consumption of the QVCO and buffer. The in-house inductor optimization software, Indentro [10], was used to find the geometry and metal layers to use. It is a very quick tool where one can sweep different parameters and put restrictions on others, e. g. $f_{srf} \geq$20 GHz. After finding a promising geometry an .ind-file can be exported to FastHenry [11] and simulated there as well. The agreement is quite good, but FastHenry typically gives a somewhat lower Q and L. Since a high resistivity substrate is used the Eddy current in the substrate can be neglected. The capacitively coupled substrate losses are blocked by a patterned ground shield. The substrate therefore does not need to be included in the simulations, and FastHenry gives accurate results. A .cif-file can also be exported from Indentro to the Cadence layout environment. The .cif-file contains all metal traces and vias for interconnect. The key for all simulations to be accurate is a correct technology file, where all thicknesses and electrical properties of the different layers of the semiconductor process are specified.

3 Results and Measurement

The circuit, Fig. 4, was flip chipped onto a gold plated printed circuit board (PCB). This minimizes the parasitics going on an off chip. Three different chips were mounted on PCBs and tested.

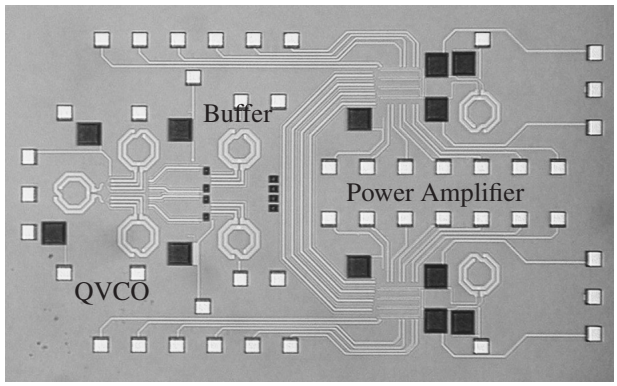


Figure 4: Chip microphotograph

3.1 Quadrature Voltage Controlled Oscillator

The phase noise of the QVCO was measured[1], see Fig. 5, and compared with the simulated phase noise. As one can see the phase noise agrees well between simulation and measurement. The figure of merit (FoM) was calculated according to (1) and found to be, in average over the three chips, 181 dB @ 3 MHz offset frequency, which is the same as in the simulation. The measured current consumption from a 1 V supply was 15.5 mA compared to 17.5 mA simulated.

$$\mathrm{FoM} = \left(\frac{f_0}{f_m}\right)^2 \frac{1}{\mathcal{L}(f_m) V_{DD} I_{VCO}(mA)} \tag{1}$$

A measurement of the quadrature error of the QVCO is not possible, since the QVCO output is not directly accessible. Instead the quadrature error has been measured after the PAs, that is the whole transmitter quadrature error has been measured. The error was 3.2° on average.

3.2 Power Amplifier

To measure the power amplifier, its differential outputs were connected to a balun[2], which in turn was connected to a spectrum analyzer. The spectrum analyzer was calibrated with a power meter. The PA not connected to the spectrum analyzer was terminated in 50 Ω after the balun. The measured power was compensated for the total

[1]The phase noise was measured with a PN9000 measurement system by Aeroflex Europtest.

[2]Balun: Krytar 6–26.5GHz, spectrum analyzer: Rhode & Schwartz FSU 20Hz–50GHz, power meter: Hewlett Packard 437B, mixer: Mini-Circuits ZMX-10GHz, oscilloscope: Tektronix 4GHz, 40GS/s, network analyzer: Agilent Technologies E8361A 10MHz–67GHz.

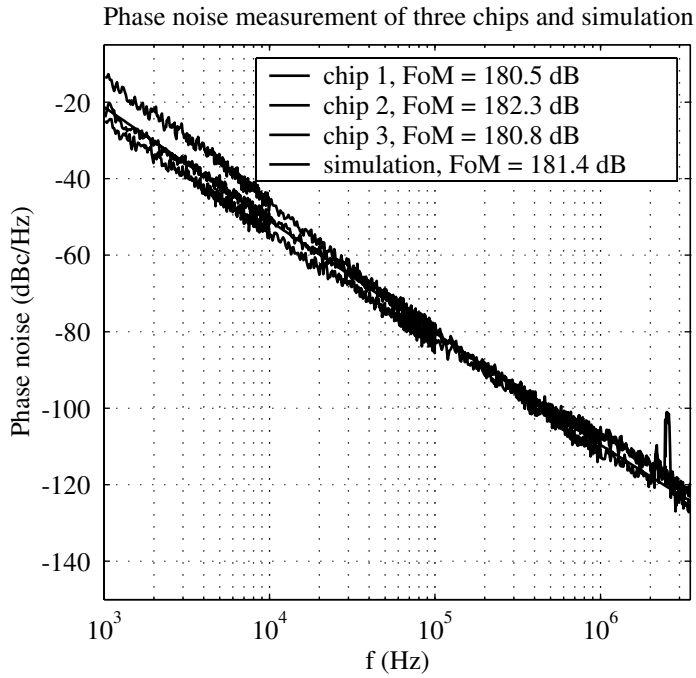


Figure 5: Measured and simulated phase noise of the QVCO

losses in connectors, cables, balun, and PCB, which were measured to 5 dB[3] using a Network Analyzer.

The maximum output power was found by adjusting the bias conditions for the QVCO, buffer, and PA. It was measured over the full tuning range for the three chips, see Fig. 6. As can be seen chip 1 and 3 have approximately the same output power and frequency, difference less than a dB over the full range, and they have quite constant output power over the full tuning range.

To be able to measure the phase at the output of the PAs the frequency was downconverted using two mixers. The intermediate frequency (IF) was chosen to 50–100 MHz. The IF signals were measured with a digital oscilloscope and from the measurement data the phase difference was calculated. The phase difference between different phase control words was measured as the phase change between the two PA outputs. It is necessary to measure the phase of one PA relative to that of the other, since the QVCO is free-running and not locked in a phase locked loop. The waveforms at the oscilloscope have noise and harmonics, thus a Fourier expansion was performed, and the phase of the fundamental tone was used.

[3]It is worth mentioning that there are many different connectors needed going from the spectrum analyzer to an SMA connector of the right gender.

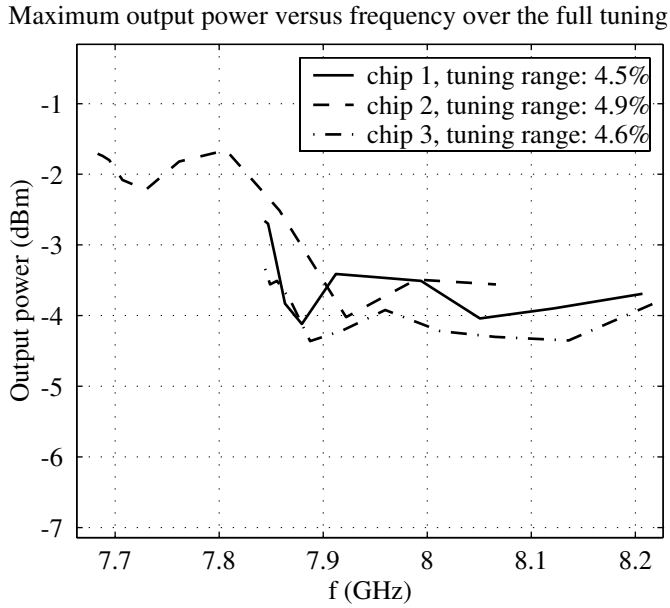


Figure 6: Maximum output power versus frequency over the full tuning range for the three chips

The phase has been measured over one quadrant for all three chips (Fig. 7). The measured and theoretical phase correspond well, with an average phase discrepancy of 3.2°. A slight curvature can be observed, however, which is due to an amplitude mismatch between I and Q. Performing a full 360° sweep at reduced PA bias (class C) increases the effect. The mismatch is therefore believed to be located prior to the PAs in the signal chain.

3.3 Summary

Table 1 is a summary of the measurement results in comparison to the simulated results. The current consumption is lower for all building blocks in the measurement. The oscillator FoM is the same in both simulations and measurements, but the tuning range differs quite a bit, which could spring from inaccurate modeling of the nMOS varactor. The output power is slightly lower than simulated.

4 Conclusion

An 8-GHz CMOS beamforming transmitter circuit in 130-nm CMOS is presented, featuring a quadrature oscillator with low phase noise (-120 dBc/Hz @ 3 MHz offset)

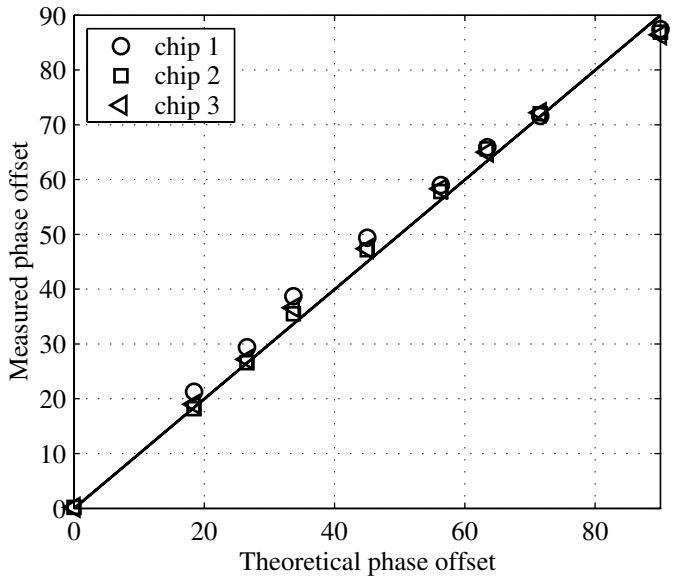


Figure 7: Output phase of the three PAs over one quadrant

and two phase controlled power amplifiers, each with an output power of -3 dBm. To be able to deliver the amount of power needed to transmit high date rates at high frequency (60 GHz), with CMOS technology, multiple power amplifiers and beamforming will be needed. Also at lower frequencies (IEEE 802.11) beamforming can be valuable to reduce interference by mainly transmitting in the wanted direction, thus allowing more units to operate in the same cell with good SNR. This work shows a robust PA design with 360° output phase control range suitable for beamforming applications.

Acknowledgment

The authors would like to thank Niklas Troedsson for using his inductor optimization software, Indentro. We would like to thank united microelectronics corporation (UMC) for giving us the opportunity to work with a state of the art 130 nm CMOS process, and the Knut and Alice Wallenberg foundation for the 60-GHz measurement equipment. Last but not least thanks to the Swedish Agency for Innovation Systems (Vinnova) for funding this project, 'Techniques for Low Cost 60 GHz WLAN'.

Table 1: Comparison between Measurements and Simulations for the Transmitter

	Simulated	Measured
V_{DD} (V)	1.0	1.0
I_{DC} (QVCO) (mA)	17.5	15.5
I_{DC} (Buffer) (mA)	19.0	17.3
I_{DC} (PA) (mA)	18.3	13.8
Phase Noise @ 3 MHz (dBc/Hz)	-123	-122
FoM @ 3 MHz (dB)	181	181
Quadrature phase error ($^\circ$)	—	3.2
Tuning Range (%)	2.2	4.7
Tuning Range (MHz)	190	370
Output Power (dBm)	-0.42	-3

References

[1] ETSI EN 302 288-1, *Electromagnetic compatibility and Radio spectrum Matters (ERM); Short Range Devices; Road Transport and Traffic Telematics (RTTT); Short range radar equipment operating in the 24 GHz range; Part 1: Technical requirements and methods of measurement*, European Telecommunications Standards Institute, Dec. 2005.

[2] ETSI TR 101 983 V1.1.1, *Electromagnetic compatibility and Radio spectrum Matters (ERM); Radio equipment to be used in the 76 GHz to 77 GHz band; System Reference Document for Short-Range Radar to be fitted on road infrastructure*, European Telecommunications Standards Institute, July 2001.

[3] 47cfr15.253, *The FCC rules and regulations are codified in Title 47 of the Code of Federal Regulations*, Code of Federal Regulations, Oct. 2005.

[4] P. Smulders, "Exploiting the 60 GHz Band for Local Wireless Multimedia Access: Prospects and Future Directions," *IEEE Communications Magazine*, vol. 40, pp. 140–147, Jan. 2002.

[5] E. R. C. (ERC), "The European Table of Frequency Allocations and Utilisations Covering the Frequency Range 9 kHz to 275 GHz," ERC REPORT 25, http://www.ero.dk/documentation/docs/doc98/official/pdf/ ERCREP025.PDF, Copenhagen: 2004.

[6] F. C. C. (FCC), "Amendment of Parts 2, 15 and 97 of the Commission's Rules to Permit Use of Radio Frequencies Above 40 GHz for New Radio Applications," ftp://ftp.fcc.gov./pub/Bureaus/Engineering_Technology/Orders/1995/ fcc95499.txt, 1995.

[7] N. Troedsson and H. Sjöland, "A Distributed Capacitance Analysis of Co-Planar Inductors for a CMOS QVCO with Varactor Tuned Buffer Stage," *Analog Integrated Circuits and Signal Processing*, vol. 42, pp. 7–19, 2005.

[8] P. Andreani, "A Low-Phase-Noise Low-Phase-Error 1.8GHz Quadrature CMOS VCO," in *IEEE International Solid-State Circuits Conference, Digest of Technical Papers*, vol. 2, Dec. 2002, pp. 228–229.

[9] E. Hegazi, H. Sjöland, and A. A. Abidi, "Filtering Technique to Lower LC-oscillator Phase Noise," *IEEE Journal of Solid-State Circuits*, vol. 36, pp. 1921–1930, Dec. 2001.

[10] N. Troedsson, http://www.indentro.com.

[11] Fast Field Solvers, http://www.fastfieldsolvers.com.

Paper II

Paper II

A 24-GHz Automotive Radar Transmitter with Digital Beam Steering in 130-nm CMOS

Abstract

In this paper simulations of a 130-nm CMOS 24-GHz automotive radar transmitter with digital beam steering is presented. The beam steering is performed by multiple PAs connected to separate antenna elements. The output phases of the PAs are individually controllable through 360° by binary weighting of quadrature phases. The circuit contains 18 PAs, each delivering 0 dBm to the antenna, resulting in a combined output power of 13 dBm. The 18 element antenna array will at 24 GHz be 11 cm, and have a directivity of 12 dBi and a half power beam width of 5 degrees.

Based on: Johan Wernehag and Henrik Sjöland, "A 24-GHz Automotive Radar Transmitter with Digital Beam Steering in 130-nm CMOS", *IEEE PhD Research in Microelectronics and Electronics*, pp 481–484, June 11–16, 2006, Lecce Italy.

1 Introduction

The car industry and the legislators are very interested in an automotive radar system. The injuries from car collisions cost the society a lot both in medical bills and in human tragedies. In the United States (US) alone motor vehicle crashes accounted for 42,000 deaths, more than 5.3 million injuries, and over $231 billion in economic losses in 2000 [1]. Already today there is a system called adaptive cruise control (ACC), or sometimes AICC, where I stands for Intelligent. The ACC system measures the distance to the closest vehicle ahead, either by radar or laser [2–4]. If the distance becomes too short the car closes its throttle, or if necessary the system will apply the brakes. If the vehicle ahead changes lane or accelerates, the ACC system will open the throttle and accelerate up to the speed preset by the cruise control. This system adds about $1,500 – $3,000 to the cost of the car [2]. The price of radar solutions has to go down if they are going to be available in all cars. This is the main reason for choosing standard CMOS technology. The cost of the system can then also be further reduced by integrating much of the digital functionality on the same chip.

The European telecommunications standards institute (ETSI) has a temporary standard [5] for short range radar (SRR) operating in the frequency band from 24.05 GHz to 24.25 GHz. The maximum effective isotropic radiated power (EIRP) is 20 dBm, limiting the maximum allowed output power. Furthermore, both ETSI and the federal communications commission (FCC) have a license free UWB frequency band in this range[1], which also can be used for automotive radar applications. A permanent frequency band in Europe and US is opened at 76–77 GHz with a maximum EIRP of 40 dBm and 48 dBm respectively [6, 7]. In Europe there is also a permanent location for anti-collision SRR at 79 GHz with an EIRP of 50 dBm [8]. With today's CMOS technologies 24 GHz operating frequency is possible, enabling low-cost implementation of 24 GHz SRR systems. In the coming years, CMOS technology feature sizes are predicted to decrease making also low-cost implementations of 77 GHz systems possible. According to the international technology roadmap for semiconductors (ITRS) [9] the 65 nm node and beyond will have transition frequencies higher than 170 GHz making them possible for 77 GHz implementation.

A typical specification for a 77 GHz radar front end can be seen in Table 1, [3, 10–12]. Since our aim is to find a transmitter architecture that can be migrated to 77 GHz when sufficiently fast CMOS technology becomes available, we use this specification for our design, although the EIRP is a bit high.

2 Air Interface

The antenna has to have a high directivity and a small half power beam width (HPBW) to be able to fulfill the specification. In addition to that it should be steerable and mechanically stable. A linear array needs to be at least $10\lambda_0$ to achieve the required

[1] In the US the band is 22 GHz–29 GHz and in Europe 22.65 GHz–25.65 GHz.

Table 1: Specification for a 77 GHz front end

Transmitter	
Frequency	76–77 GHz
Modulation	FM-CW
Tx Phase Noise	< -80 dBc/Hz @ 100 kHz offset
Sweep Width	300 MHz
Range	$2 \rightarrow 150$ m
Linearity	$<0.5\%$
Transmit Power	10–15 dBm
Reciever	
Beam Width	3°
Field Of View	$8^\circ - 20^\circ$
Beam Overlap	0.5°
Relative Speed	$-200 \rightarrow +100$ km/h
Calculation Frequency	10 Hz
Range (for 1 m^2 target)	1–100 m

HPBW. This corresponds to 4 cm at 77 GHz and 13 cm at 24 GHz. A linear array of $\lambda/2$ dipole patches has been assumed, fed in the center by the differential signal from the PAs, see Fig. 1. The supply voltage to the PAs is inserted at the signal ground at end of the antenna patches. The physical size of the dipole antenna is $\lambda_{sub}/2$, where λ_{sub} is the wavelength in the patches on the substrate.

In Fig. 2 the radiation pattern is plotted when the beam is swept over field of view (FoV). To achieve a FoV of $\pm 13^\circ$ up to $\pm 40^\circ$ phase difference of adjacent antenna elements is required. The PAs are not restricted to this FoV, however, since they can deliver any phase difference between the antenna elements.

The HPBW and directivity have been simulated, see Fig. 3. The phase between different elements was swept from 40° to -40° giving $\pm 13^\circ$ beam steering. As can be seen, the directivity is larger than 12 dBi over the complete range and the HPBW is below 4.3 degrees. To achieve a HPBW of 3 degree or better over the full range the size of the antenna array would have to be increased. An increased antenna element spacing of $0.65\lambda_0$ would give 3 degree HPBW with 18 antenna elements.

3 Design and Simulation

With today's CMOS processes, such as the 130 nm used in this paper, 77 GHz is a very high frequency. The aim is therefore set on the 24 GHz frequency band for car

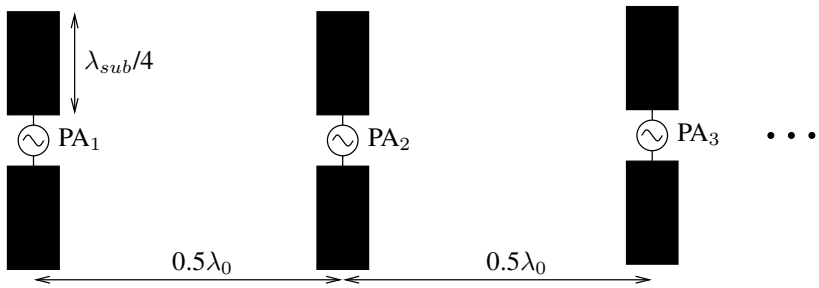


Figure 1: The dipole patch antenna array with the PA feeding indicated

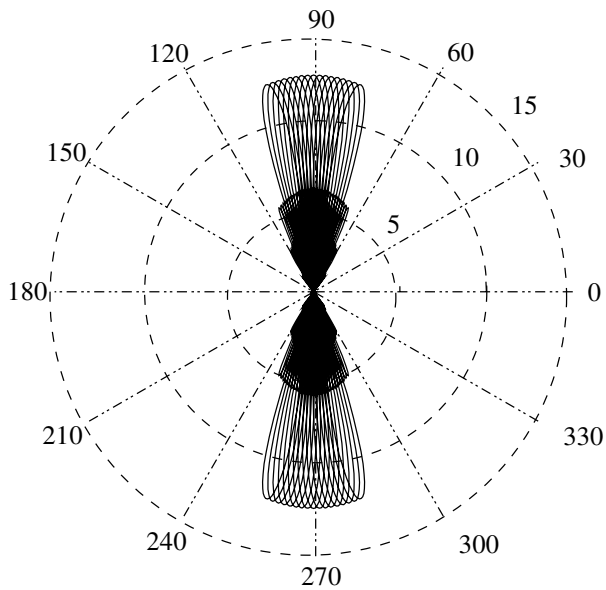


Figure 2: Radiation patterns for $\lambda/2$ spacings of the antenna elements

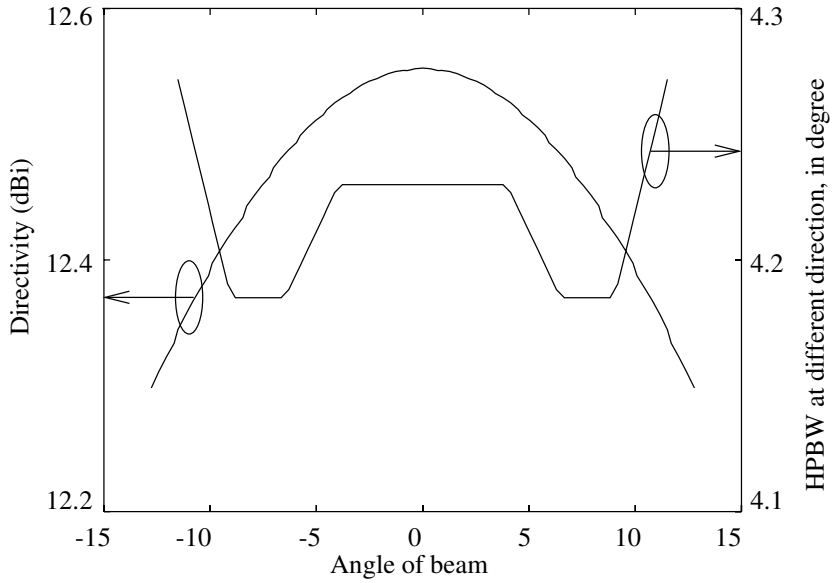


Figure 3: HPBW and directivity with 18 antennas in a linear patch array and $0.5\lambda_0$ spacing

radar applications. The specifications in Table 1 are used for this band as well.

A block diagram of the beam steering multiple PA circuit can be seen in Fig. 4. Distributing the quadrature signal to the 18 PAs therefore requires long interconnect lines which has to be considered during simulation.

In the following sections the different parts of the circuit are described from left to right in Fig. 4.

3.1 Polyphase Filter

To convert the externally applied differential input signal to a quadrature signal, a passive polyphase filter [13] is used, see Fig 5. The filter has two links, tuned to a lower and a higher frequency. The filter is thereby rather broadband. For a $\pm 2^\circ$ quadrature phase error the band is from 18 GHz to 29 GHz, and the voltage loss of the filter within that band is less than 11 dB, when loaded by the inverter tree.

To reach the high operating frequencies, the resistances of the filter must be low, 45 Ω and 60 Ω in the two stages. This results in an input impedance of about $10-j14\,\Omega$. To get a 50 Ω input impedance of the circuit, better suited for the measurement equipment, an on-chip L-match network is used in-front of the polyphase filter.

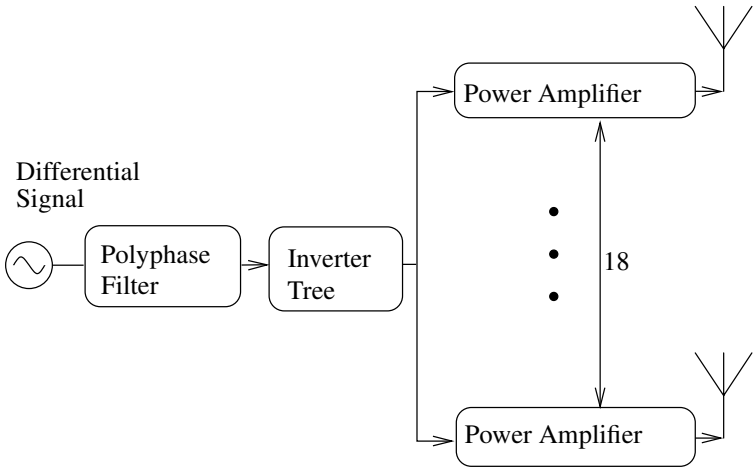


Figure 4: Block diagram the car radar transmitter

3.2 Inverter Tree Driver

Since the loading of 18 PAs is high (6.4 pF) a tree structure has been chosen for the driver circuit. The driver consists of five stages of inverters, see Fig. 6. Each inverter symbol in Fig. 6 has four inverters, one for each phase of the quadrature signal. To stabilize the inverter output DC voltage at $V_{DD}/2$, a high ohmic resistive feedback is applied locally, and a DC-block capacitor is connected in series with the input to isolate the stages.

The purpose of the first stage in the inverter tree is to reduce the loading of the polyphase filter and to drive the three following inverters. The routing from the first to the third stage is long, thus an second stage is inserted to drive the interconnect wire and the inverters of the third stage. The capacitance of one of the interconnect wires is roughly 18 fF, and the series resistance in the same wire is below 1 Ω. From the third to the fifth stage the routing is local and thus the tree starts to grow, ending in 18 output branches.

The ratio of the output current and input current of the inverter tree, when placed in the transmitter circuit, is 3.3 times, achieving an amplitude of 112 mV at the PA input.

3.3 Power Amplifier

A 360° phase steering PA fed by a quadrature signal has been designed, see Fig. 7. The binary weighted transistor banks gives the possibility to weight the phases differently by changing the control voltages, V_{x_y}. Then the transmitted signal from each PA can

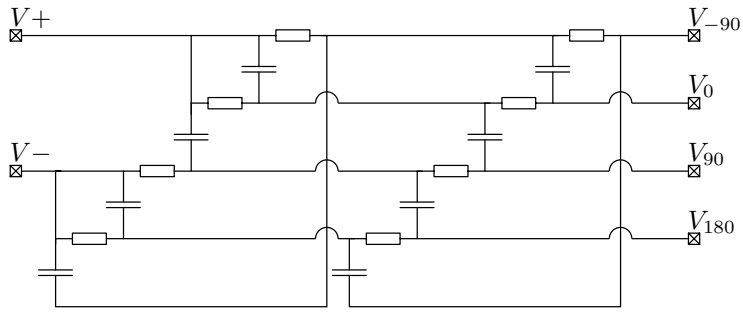


Figure 5: The passive two stage polyphase filter

have an arbitrary phase[2].

The power delivered to the antenna array should be 10 dBm, distributed among the 18 PAs. Each PA should then deliver at least -3 dBm, and is therefore designed to output 0 dBm into a 60 Ω load, giving a 3 dB margin for losses in antenna feeding networks. The 60 Ω patch resistance is obtained by tuning the feeding point, x, of the patch antenna according to (1a). $R_{in}(x = 0)$ is made high by making the patch width, W, small (1b) [14].

$$R_{in}(x) = R_{in}(x = 0)\cos^2\left(\frac{\pi}{L}x\right) \tag{1a}$$

$$R_{in}(x = 0) = \frac{1}{2G} = 45\left(\frac{\lambda_0}{W}\right)^2 \qquad W \ll \lambda_0 \tag{1b}$$

Thus the patch can be made small and matched to the PA at the same time.

It should be stressed that the inductors in Fig. 7 are on the antenna substrate and not on-chip, they are to be realized as stubs.

The phasor of one of the PAs has been swept over a quadrant and the result is plotted in Fig. 8 together with the ideal points. The other PA, which loads the same inverter one stage back, is set to 45° during the sweep.

3.4 Shift Register

To change the phase setting of the 18 PAs a shift register has been designed to serially clock in the digital phase control word that gives the new position of the beam. The total time for each beam position is about 10 ms, see Table 1[3]. In this time slot a new

[2]This is true if the discretization depth is large ($:\infty$). In this case there are 40 different output phases with an output power within $\pm 11\%$.

[3]To sweep the FoV approximately 10 phase settings is required. This should be performed 10 times per second, thus giving 10 ms per phase setting.

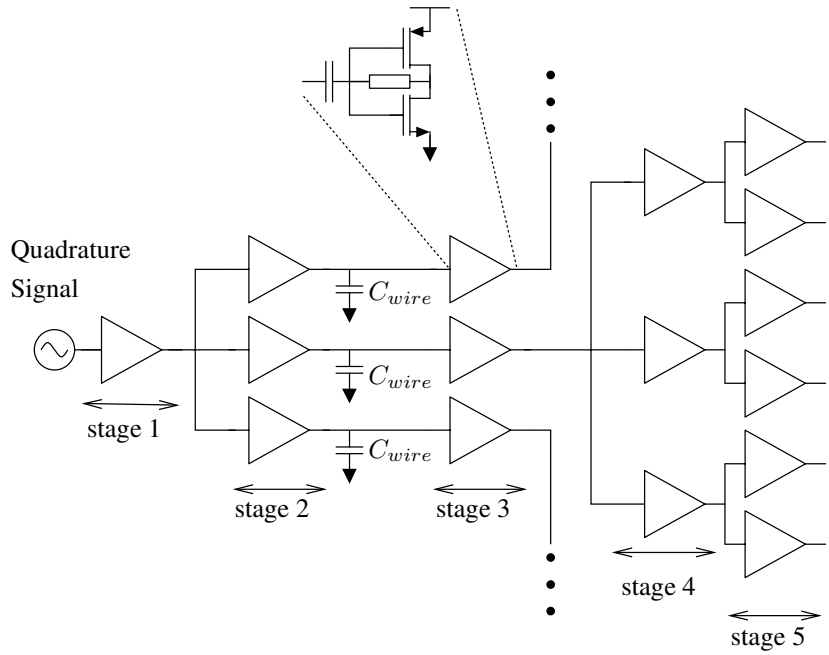


Figure 6: Structure of the driver circuit for the PAs

phase setting must be clocked in, the circuit has to stabilize, and a new measurement be performed. To clock in the 216 (12×18) control signals does not take more than a couple of microseconds with a 100 MHz clock, which is not considered high in modern digital CMOS processes. The shift register is built with the cell shown in Fig. 9.

The last inverter in the stage is larger than the preceding one. It is necessary to have a large driving capability of the last inverter to force a change in the next stage. In this case the last inverters are made 4/3 times larger.

4 Conclusion

The results of the simulations support the idea that a low cost automotive radar system in CMOS can be built, enabling increased safety on the roads.

It has been shown that the transmit power needed to get sufficient distance coverage is possible to deliver with 130 nm CMOS technology at 24 GHz, by use of multiple PAs. The multiple PAs deliver 13 dBm in total. They also support digital beam steering with good accuracy. Electrical beam steering makes the radar design more robust and less expensive than mechanical solutions.

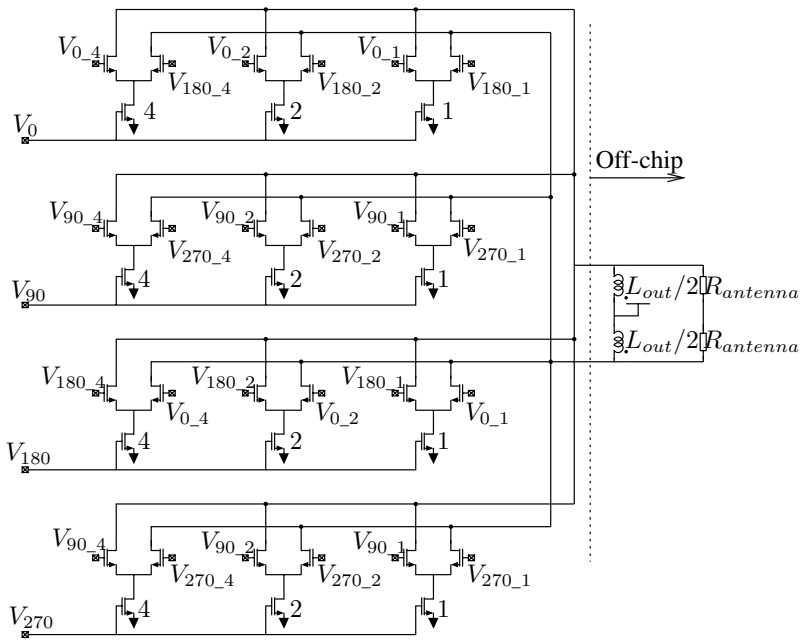


Figure 7: Schematic of the power amplifier with three binary weighted transistors per bank

Acknowledgment

The authors would like to thank united microelectronics corporation (UMC) for giving us the opportunity to work with a state of the art 130 nm CMOS process. They would also like to thank the Swedish Agency for Innovation Systems (Vinnova) for funding this project, which is a part of 'Techniques for Low Cost 60 GHz WLAN'.

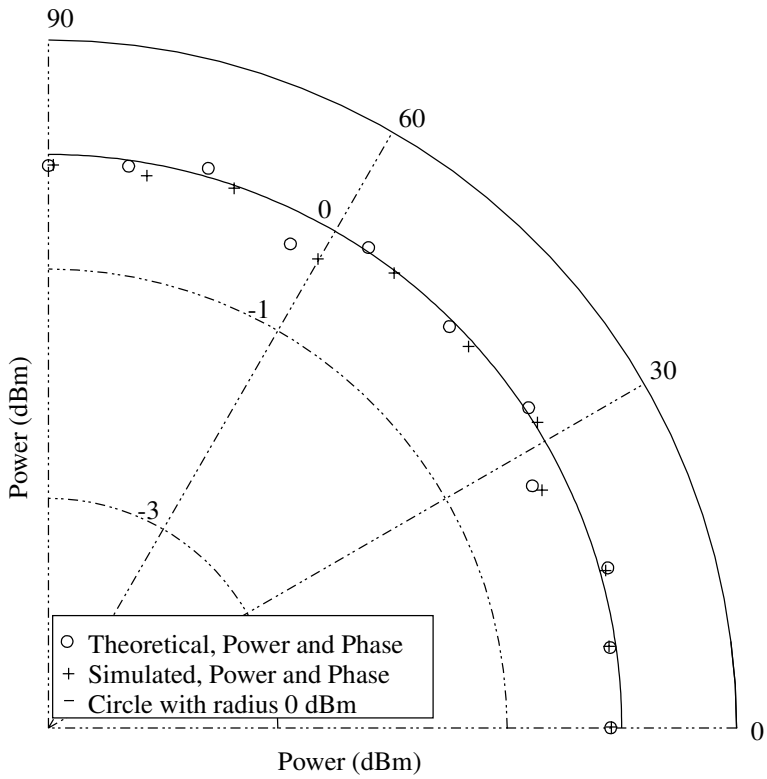


Figure 8: The output phasors of a PA in one quadrant

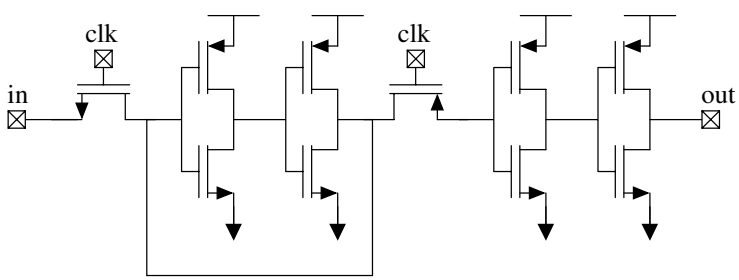


Figure 9: Schematic of the shift register cell

References

[1] L. Blincoe, A. Seay, E. Zaloshnja, T. Miller, E. Romano, S. Luchter, and R. Spicer, "The Economic Impact of Motor Vehicle Crashes, 2000," *National Highway Traffic Safety Administration*, May 2002, report NO. DOT HS 809 446.

[2] W. D. Jones, "Keeping Cars from Crashing," *IEEE Spectrum*, Sep. 2001.

[3] J. Robinson, D. K. Paul, J. Bird, D. Dawson, T. Brown, D. Spencer, and B. Prime, "A Millimetric Car Radar Front End for Automotive Cruise Control," in *IEE Colloquium on Automotive Radar and Navigation Techniques*, no. 1, 1998, pp. 1–8.

[4] G. R. Widmann, M. K. Daniels, L. Hamilton, L. Humm, B. Riley, J. K. Schiffmann, D. E. Schnelker, and W. H. Wishon, "Comparison of Lidar-Based and Radar-Based Adaptive Cruise Control Systems," *SAE Technical Paper Series*, 2000-01-0345.

[5] ETSI EN 302 288-1, *Electromagnetic compatibility and Radio spectrum Matters (ERM); Short Range Devices; Road Transport and Traffic Telematics (RTTT); Short range radar equipment operating in the 24 GHz range; Part 1: Technical requirements and methods of measurement*, European Telecommunications Standards Institute, Dec. 2005.

[6] ETSI TR 101 983 V1.1.1, *Electromagnetic compatibility and Radio spectrum Matters (ERM); Radio equipment to be used in the 76 GHz to 77 GHz band; System Reference Document for Short-Range Radar to be fitted on road infrastructure*, European Telecommunications Standards Institute, July 2001.

[7] Federal Communication Commission, *The FCC rules and regulations are codified in Title 47 of the Code of Federal Regulations, 47cfr15.253*, Code of Federal Regulations, Oct. 2005.

[8] ETSI TR 102 263 V1.1.2, *Electromagnetic compatibility and Radio spectrum Matters (ERM); Road Transport and Traffic Telematics (RTTT); Radio equipment to be used in the 77 GHz to 81 GHz band; System Reference Document for automotive collision warning Short Range Radar*, European Telecommunications Standards Institute, Feb. 2004.

[9] International Technology Roadmap for Semiconductors (ITRS), http://public.itrs.net/.

[10] I. Gresham, N. Jain, T. Budka, A. Alexanian, N. Kinayman, B. Ziegner, S. Brown, and P. Steacker, "A Compact Manufacturable 76-77 GHz Radar Module for ACC Applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 49, no. 1, pp. 44–58, Jan. 2001.

[11] A. Kawakubo, S. Tokoro, Y. Yammada, K. Kuroda, and T. Kawasaki, "Electronically-Scanning Millimeter-Wave RADAR for Forward Object Detection," *SAE Technical Paper Series*, 2004-01-1122.

[12] Prismark Discovery Series, "Automotive radar, cruising at 77 while stopping crashes at 24," *Prismark Partners LLC*, Dec. 2004.

[13] F. Behbahani, Y. Kishigami, J. Leete, and A. A. Abidi, "CMOS Mixers and Polyphase Filters for Large Image Rejection," *IEEE Journal of Solid-State Circuits*, vol. 36, no. 6, pp. 873–887.

[14] C. A. Balanis, *Antenna Theory: Analysis and Design*, 2nd ed. New York: John Wiley and Sons, Inc., 1996.

Paper III

Paper III

Antenna Array for a 24-GHz Automotive Radar with Dipole Antenna Element Patches

Abstract

In this paper 3D electromagnetic simulations of an antenna array have been performed. The array is intended for automotive radar applications at 24 GHz. It is constructed from dipole antenna element patches, which are fed in the center by a differential signal.

The dipole antenna element patches are simulated assuming a standard off the shelf substrate. They have a standing wave ratio less than 3 from 22.5 GHz to 24.5 GHz, when matched to 60 Ω. The directivity of the dipole patch is 9 dBi.

The array consists of 24 elements giving it a physical size of 150 mm, with a groundplane of 200×100 mm^2. An antenna of that size is easy to integrate in a car. Beam steering can be accomplished by changing the phases of the signals to the different elements [1], thus making the arrangement mechanically robust since the antenna does not have to move. The directivity for the antenna array is larger than 9.4 dBi for a steering angle of $\pm 7^\circ$ and the half power beam width is smaller than 6° over the same steering angle.

Based on: Johan Wernehag and Henrik Sjöland, "Antenna Array for a 24-GHz Automotive Radar with Dipole Antenna Element Patches", *International Symposium on Antennas and Propagation*, November 1–4, 2006, Singapore Singapore.

1 Introduction

The car industry and the legislators are very interested in an automotive radar system. The injuries from car collisions cost the society a lot both in medical bills and in human tragedies. In the United States (US) alone motor vehicle crashes accounted for 42,000 deaths, more than 5.3 million injuries, and over $231 billion in economic losses in 2000 [2].

Today there are already automotive radars available [3–5]. These system are either based on microwave signals or laser. Today's systems add about $1,500–$3,000 to the cost of a car [3]. This is too much if the radar is going to be an every car commodity and thus decrease the injuries, deaths, and the cost of car crashes. The price tag is why the focus in [1] was directed to digital CMOS processes, which are predicted by international technology roadmap for semiconductors (ITRS) to have transition frequencies (f_T) and maximum oscillation frequencies (f_{max}) in excess of 200 GHz and 310 GHz, respectively, in the coming 5–10 years [6]. This will enable implementation of automotive radar systems at 77 GHz in CMOS. At 77 GHz there is a frequency band allocated both in Europe [7], Japan [8], and the US [9]. The European telecommunications standards institute (ETSI) also has a temporary standard [10] for short range radar (SRR) operating in the frequency band from 24.05 GHz to 24.25 GHz. Furthermore, both ETSI and the federal communications commission (FCC) have a license free ultra wideband (UWB) frequency band in this range[1], which can also be used for automotive radar applications. In [1] a 24 GHz automotive radar transmitter circuit topology was presented. The antenna array of this paper is designed to fit that circuit topology.

The chip area in CMOS is relatively cheap compared to III–V devices and the ability to integrate digital signal processing on the same chip as the transmitter further reduces the overall system cost. A typical specification for a 77 GHz radar front end can be seen in Table 1, [4,8,11–13]. This specification is used for the design at 24 GHz as well since the main antenna[2] requirements are the same.

2 Radar Transmitter Circuit

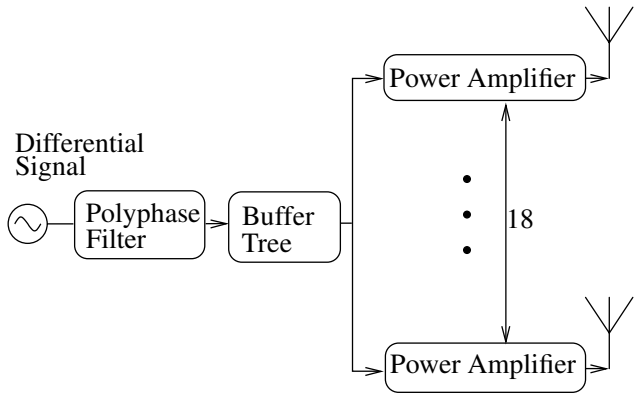
With today's CMOS processes, such as the 130 nm used in [1], 77 GHz is a very high frequency. The aim is therefore set on the 24 GHz frequency band for automotive radar applications.

A block diagram of the beam steering multiple power amplifier (PA) circuit can be seen in Fig. 1(a).

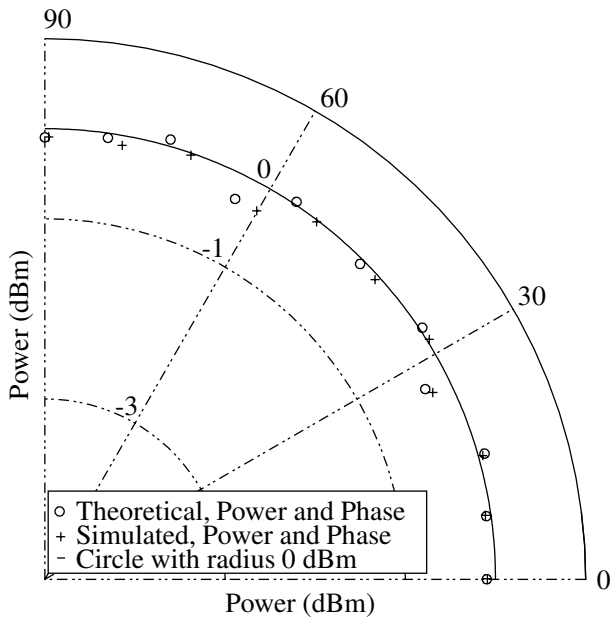
The PAs in Fig. 2 are fed by quadrature signals. A digital control word binary weights the four phases to the output and thus the output phase is controllable through

[1] In the US the ultra wideband is between 22 GHz–29 GHz and in Europe 22.65 GHz–25.65 GHz.

[2] That is the requirement on the resolution is the same and thus the antenna design parameters are the same, such as beam width, field of view, etc.



(a)



(b)

Figure 1: (a) Block diagram the automotive radar transmitter, (b) The output phasor of a PA swept over one quadrant

Table 1: Specification for a 77 GHz front end

Frequency	76–77 GHz
Modulation	FM-CW
Tx Phase Noise	< -80 dBc/Hz @ 100 kHz offset
Bandwidth	300 MHz
Transmit Power	10–15 dBm
Beam Width	4°
Field Of View	8°–20°
Beam Overlap	0.5°
Relative Speed	$-200 \rightarrow +100$ km/h
Update Frequency	10 Hz
Range (for 1 m^2 target)	1–100 m

360°. The simulated phasor tip from one PA is plotted in Fig. 1(b). For a more detailed description of the transmitter see [1].

3 Antenna Design

The antenna array must have a high directivity and a small half power beam width (HPBW) to be able to fulfill a specification like the one in Table 1. In addition, the lobe should be steerable and the antenna mechanically robust. A linear array needs to be at least $10\lambda_0$ to achieve the required HPBW. This has been deduced from ideal radiation expressions [14], hereafter referred to as ideal. That size corresponds to 4 cm at 77 GHz and 13 cm at 24 GHz. A linear array of $\lambda/2$ dipole patches has been assumed, fed in the center by the differential signal from the PAs, see Fig. 3.

The supply voltage to the PAs is inserted at the signal ground at the end of the antenna patches, eliminating the need of separate RF-chokes for feeding the DC-current to the PAs. The physical size of the dipole antenna is roughly $\lambda_{sub}/2$, where λ_{sub} is the wavelength in the substrate. The antenna substrate has a relative dielectric constant of 10.2[3].

To investigate the antenna system further, not assuming ideal expressions, a 3D electro magnetic (EM) solver, SEMCAD [15], was used.

First one antenna element was simulated and optimized and later the full array. Broadband simulations were used for the antenna element to investigate its input impedance resonance, field patterns, and surface current distributions. Harmonic simulations of the full array were then performed at 24 GHz. To perform broadband

[3]This corresponds to substrates available off the shelf, e. g. Arlon AD 1000. The dimensions used in the following section are within the range for this substrate.

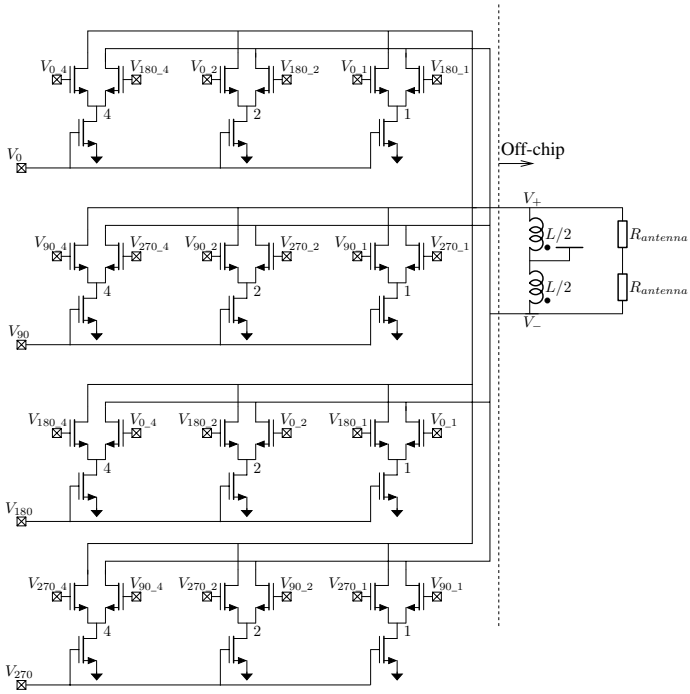


Figure 2: Schematic of the power amplifier with three binary weighted transistors per bank

simulations of the full array would not be practical, due to excessive computer memory requirements. An advantage of harmonic simulations is that the grid can be made tighter.

3.1 Antenna Element

The geometry of one patch in a dipole element is defined in Fig. 4(a). The distance between the two patches of the differential antenna element is 100 μm. The length, L, of the patch was tuned to give a first resonance at 24 GHz, which occurred for a length equal to 1.3 mm. From this the effective relative dielectric constant at 24 GHz was calculated to be equal to 5.8. The width was chosen to make the input impedance as close as possible to 60 Ω over the frequency band. A wide patch results in a narrow band resonance with a high corresponding impedance. For a narrow patch with a width that is a fraction of the length, the impedance at resonance decreases and multiple resonances occur, thus a more broadband response is attained. A width of 0.6 mm

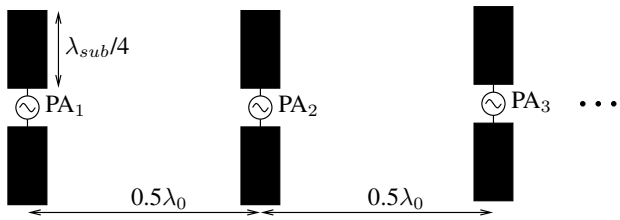


Figure 3: The dipole patch antenna array with the PA feeding indicated

gave the best result, Fig. 5(a–b). At last the thickness of the substrate was investigated to see how thin the substrate could be made. When the thickness was less than 2.8 mm surface waves started to propagate. A substrate thickness of 3.2 mm was therefore chosen to create some margin.

The radiation pattern and surface current distribution for an antenna element with these dimensions are plotted in Fig. 4(b). As seen in Fig. 5(a–b) the input resistance is close to 60 Ω from 22.5 GHz to 24.5 GHz, resulting in a standing wave ration (SWR) less than 3 over that frequency range. The electric field in two vertical cuts for 5 different frequencies is plotted in Fig. 5(c–d). The two cuts are the xz-plane and the yz-plane, according to the coordinate system in Fig. 4(b). As can be seen the directivity is about 9 dBi over the entire frequency range, with a maximum of 10.0 dBi and a minimum of 7.0 dBi, which is in alignment with [16–18]. In Fig. 5(d) there is horizontal emission at one frequency, 24.5 GHz. The total simulated ($\eta_{\mathrm{mismatch}} \times \eta_{\mathrm{radiation}}$) antenna element efficiency is 95% at 24 GHz.

3.2 Antenna Array

The array is made of twentyfour elements with the dimensions found in Section 3.1. A larger substrate of the array than 206×103 mm could not be simulated due to memory limitations (4 GB) of the computer in combination with high requirements on grid resolution.

The three dimensional far-field pattern and surface current distribution is plotted in Fig. 6(a). The antenna efficiency at 2° steering angle is 75%. To validate that the simulation grid is tight enough it has been further tightened without any significant changes in the obtained simulation results. This indicates that the grid is tight enough to predict the far-field behavior correctly. The electric field pattern in the xz- and yz-direction[4] is plotted in Fig. 6(c–d). The field patterns in Fig. 6(c–d) are from 2° and

[4]We use the term direction instead of plane for the array, since when steering out the lobe from zero degrees also the yz-direction of interest will steer out with that angle, theta≠0°. The peak directivity in the xz-direction is not at zero degree either, see the dips in the yz-direction (Fig. 6(d)), so the xz-direction is tilted to the first peak in the yz-direction.

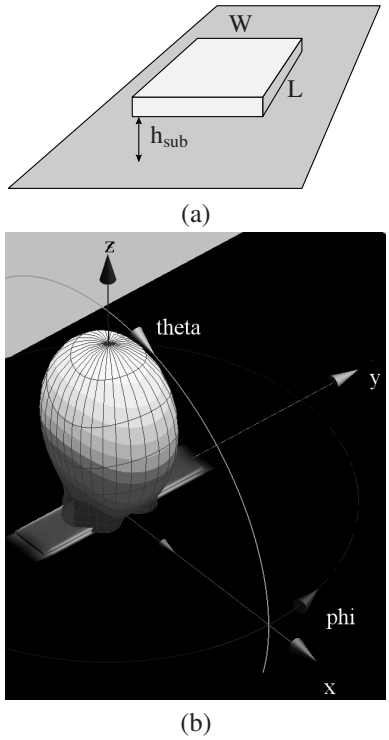


Figure 4: (a) Geometry of one patch in a dipole element, (b) Radiation pattern and surface current distribution of the element

$-6°$ steering angles. The sidelobe supression is larger than 7 dB over the full steering range.

The directivity ranges from 12.3 dBi at 2° to 9.4 dBi at 4° steering angle. In the yz-direction there are dips of about 10 dB originating from the one dimensional structure of the array. The xz-direction is tilted 13° to align the cut with the first peak in the yz-plane (Fig. 6(a)) for 2° steering angle.

The directivity and HPBW were simulated as a function of steering angle and are plotted in Fig. 6(b). The directivity is larger than 9 dBi and the HPBW is below 6° over the $\pm 7°$ steering range. A comparison is also made between SEMCAD, Array Pattern Visualizer [19], and the ideal expressions.

4 Conclusion

The antenna simulations performed in this paper support the idea that an automotive radar antenna array at 24 GHz can be constructed from dipole patches. The simulated directivity ($\approx$10 dBi) and HPBW (6°) are satisfactory. The physical size of the array is practically implementable in a car (20×10 cm).

Together with the 24 GHz CMOS automotive radar transmitter circuit topology presented by the authors in [1], a low-cost and mechanically robust solution can be accomplished.

Acknowledgment

The authors would also like to thank the Swedish Agency for Innovation Systems (Vinnova) for funding this research, which is a part of the project 'Techniques for Low Cost 60 GHz WLAN'. For SEMCAD introduction and help the authors would like to acknowledge Dr. Anders Johansson.

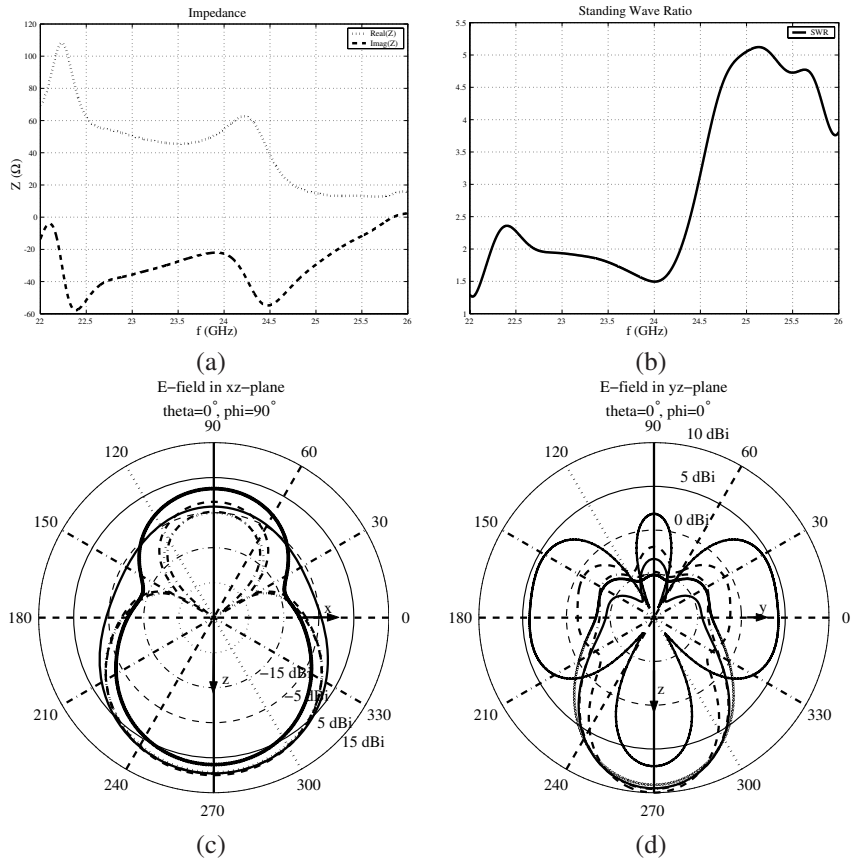


Figure 5: (a) Input impedance with length 1.3 mm and width 0.6 mm, (b) SWR with length 1.3 mm and width 0.6 mm, (c) and (d) Polar plots of the E-field in vertical cuts at frequencies: 22.5 GHz, 23.0 GHz, 23.5 GHz, 24.0 GHz, and 24.5 GHz

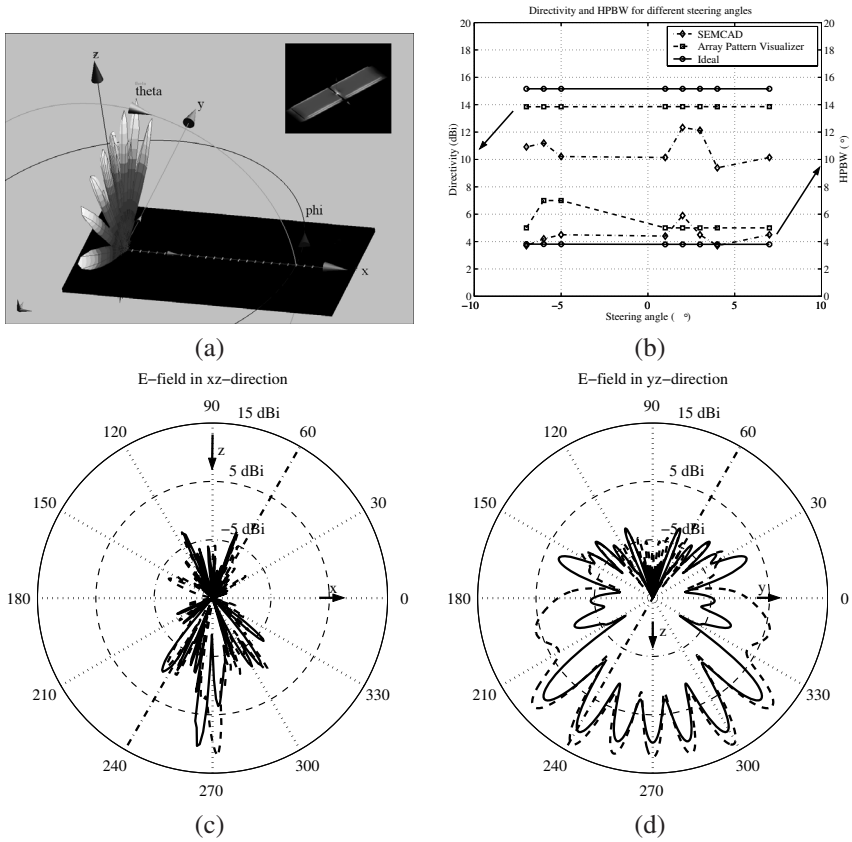


Figure 6: (a) Radiation pattern and surface current distribution of the antenna array consisting of 24 dipole antenna patches, (b) Directivity and HPBW as a function of steering angle, (c–d) Polar plots of the E-field in vertical cuts at 24 GHz at steering angle 2 and -6 degrees

References

[1] J. Wernehag and H. Sjöland, "A 24-GHz Automotive Radar Transmitter with Digital Beam Steering in 130-nm CMOS," in *PhD Research in Microelectronics and Electronics*, June 2006, pp. 481–484.

[2] L. Blincoe, A. Seay, E. Zaloshnja, T. Miller, E. Romano, S. Luchter, and R. Spicer, "The Economic Impact of Motor Vehicle Crashes, 2000," *National Highway Traffic Safety Administration*, May 2002, report NO. DOT HS 809 446.

[3] W. D. Jones, "Keeping Cars from Crashing," *IEEE Spectrum*, Sep. 2001.

[4] J. Robinson, D. K. Paul, J. Bird, D. Dawson, T. Brown, D. Spencer, and B. Prime, "A Millimetric Car Radar Front End for Automotive Cruise Control," in *IEE Colloquium on Automotive Radar and Navigation Techniques*, no. 1, 1998, pp. 1–8.

[5] G. R. Widmann, M. K. Daniels, L. Hamilton, L. Humm, B. Riley, J. K. Schiffmann, D. E. Schnelker, and W. H. Wishon, "Comparison of Lidar-Based and Radar-Based Adaptive Cruise Control Systems," *SAE Technical Paper Series*, 2000-01-0345.

[6] International Technology Roadmap for Semiconductors (ITRS), http://public.itrs.net/, 2005.

[7] ETSI TR 102 263 V1.1.2, *Electromagnetic compatibility and Radio spectrum Matters (ERM); Road Transport and Traffic Telematics (RTTT); Radio equipment to be used in the 77 GHz to 81 GHz band; System Reference Document for automotive collision warning Short Range Radar*, European Telecommunications Standards Institute, Feb. 2004.

[8] Australian Communications Authority, *A review of automotive radar systems - devices and regulatory frameworks*.

[9] 47cfr15.253, *The FCC rules and regulations are codified in Title 47 of the Code of Federal Regulations*, Code of Federal Regulations, Oct. 2005.

[10] ETSI EN 302 288-1, *Electromagnetic compatibility and Radio spectrum Matters (ERM); Short Range Devices; Road Transport and Traffic Telematics (RTTT); Short range radar equipment operating in the 24 GHz range; Part 1: Technical requirements and methods of measurement*, European Telecommunications Standards Institute, Dec. 2005.

[11] I. Gresham, N. Jain, T. Budka, A. Alexanian, N. Kinayman, B. Ziegner, S. Brown, and P. Steacker, "A Compact Manufacturable 76-77 GHz Radar Module for ACC Applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 49, no. 1, pp. 44–58, Jan. 2001.

[12] A. Kawakubo, S. Tokoro, Y. Yammada, K. Kuroda, and T. Kawasaki, "Electronically-Scanning Millimeter-Wave RADAR for Forward Object Detection," *SAE Technical Paper Series*, 2004-01-1122.

[13] Prismark Discovery Series, "Automotive radar, cruising at 77 while stopping crashes at 24," *Prismark Partners LLC*, Dec. 2004.

[14] C. A. Balanis, *Antenna Theory: Analysis and Design*, 2nd ed. New York: John Wiley and Sons, Inc., 1996.

[15] SEMCAD, http://www.semcad.com.

[16] J. Granholm and K. Woelders, "Dual Polarization Stacked Microstrip Patch Antenna Array With Very Low Cross-Polarization," *IEEE Transactions on Antennas and Propagation*, no. 10, pp. 1393–1402.

[17] N. Chamberlain, L. Amaro, E. Oakes, R. Hodges, S. Spitz, and P. A. Rosen, "Microstrip Patch Antenna Panel for Large Aperture L-band Phased Array," in *IEEE Aerospace Conference, Digest*, Mar. 2005, pp. 1185–1192.

[18] P. Salonen and H. Hurme, "A Novel Fabric WLAN Antenna For Wearable Applications," in *IEEE Antennas and Propagation Society International Symposium*, June 2003, pp. 700–703.

[19] J. Kraus and R. Marhefka, "Array Pattern Visualizer," http://esl.eng.ohio-state.edu/ rjm/antennas/standalone.htm.

Paper IV

Paper IV

60 GHz 130-nm CMOS Second Harmonic Power Amplifiers

Abstract

Two different frequency doubling power amplifiers have been measured, one with differential and one with single-ended input, both with single-ended output at 60 GHz. The amplifiers have been implemented in a 1p8M 130-nm CMOS process. The resonant nodes are tuned to 30 GHz or 60 GHz using on-chip transmission lines, which have been simulated in ADS and Momentum.

The measured input impedance of the single-ended PA is high at 250 Ω, and the differential input is similar, making the PA a suitable load for an oscillator in a fully integrated transmitter. The single-ended and differential input PA delivers 1 dBm and 3 dBm, respectively, of measured saturated output power to 50 Ω, both with a drain efficiency of 8%.

Based on: Johan Wernehag and Henrik Sjöland, "60 GHz 130-nm CMOS Second Harmonic Power Amplifiers", *IEEE Asia Pacific Conference on Circuits and Systems (APCCAS)*, November 30 – December 3, 2008, Macao China.

1 Introduction

The drive for gigabit per second transmission rates in coming WLAN and WPAN applications forces the design of new systems with wider bandwidth. The federal communications commission (FCC) has opened a 7 GHz wide spectrum from 3–10 GHz for UWB communications, and a 7 GHz wide unlicensed band at 57–64 GHz [1] is also available. Both these spectrum allocations will permit communications at several gigabits per second. Also in Japan and Europe [2] frequency bands at 60 GHz are opened for unlicensed WLAN communications, providing a 5 GHz world wide overlap.

To meet the stringent cost requirements of the consumer electronics market, the 60 GHz transceivers must be realized in low cost CMOS technology [3, 4]. To use CMOS at this high frequency is extremely challenging, both in receiver and transmitter. In this work we have therefore realized a frequency doubling power amplifier (PA). This will reduce the difficulties in designing the rest of the transmitter, which can then operate at 30 GHz instead of 60 GHz. The power consumption and the losses are then reduced. A harmonic amplifier also makes the frequency planning easier, since the voltage controlled oscillator (VCO) and PA output signal are separated in frequency, reducing the risk of oscillator pulling. Furthermore, at 30 GHz center frequency the achievable tuning range of the VCO increases. The tuning range is limited at high frequencies by the performance of the varactors, and The quality factor of the varactor is approximately doubled at 30 GHz compared to 60 GHz [5].

2 Circuit Topologies

To show the concept of 60 GHz CMOS frequency doubling power amplifiers, two different frequency doubling PA topologies have been investigated, Fig. 1. In cases where higher output power and more gain is needed, more stages can be used [6, 7]. Both amplifiers convert a 30 GHz input signal to a 60 GHz output. The frequency doubling takes place in the first stage, and the drain node of M1 (and M3) is tuned to 60 GHz, passing the second harmonic and suppressing the 30 GHz fundamental. The second stage is a 60 GHz common source amplifier. At the output of the second stage an L-match is used to transform the 50 Ω load to a higher impedance seen by the PA. The L-match consists of the drain-bulk capacitance of M2 and the transmission line from M2 to the output pad. The input of the amplifiers are high-ohmic, making the PAs suitable to proceed the oscillator in a fully integrated transmitter.

The first stage of the differential input amplifier [8] (Fig. 1(b)) conducts at both input polarities, contributing to a larger effective g_m of that stage compared to its single-ended counterpart. The drain-bulk capacitance, however, is correspondingly larger, making the first stage gain approximately equal for the two topologies. The fundamental tone at the drains of M1 and M3 is 180° out of phase, but the second harmonic is in phase. Combining the drains then results in a signal with a suppressed

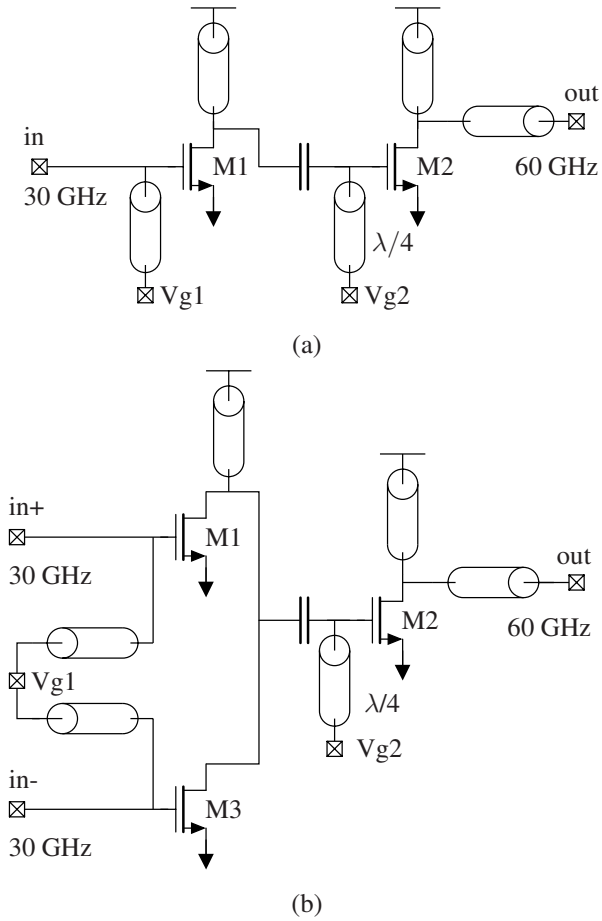


Figure 1: The frequency doubling power amplifier topologies investigated. (a) Single-ended input (b) Differential input

fundamental and a strong tone at twice the frequency of the input. The suppression of the fundamental makes LO leakage through the power amplifier less of a problem, see Table 1. Furthermore, not just the fundamental, but all odd harmonics are rejected [8]. The differential input amplifier is also suitable for on-chip implementation, where differential signal schemes dominate thanks to their high common mode noise/interference suppression. In addition to that, the power amplifier produces a single-ended output, eliminating the need of an output balun.

2.1 Transmission Lines

All the resonant nodes are tuned by transmission lines (TLs) terminated by short circuits. A model of the silicon back end of line (metal stack) has been used to simulate the TLs in advanced design system (ADS) [9]. The 2.0 μm thick top metal layer was used for the TLs, and metal 1 for ground plane. The width of the TLs was chosen to give a 50 Ω characteristic impedance, which also resulted in a high Q-value, see [10] and [6] for a more thorough investigation of TLs in silicon processes.

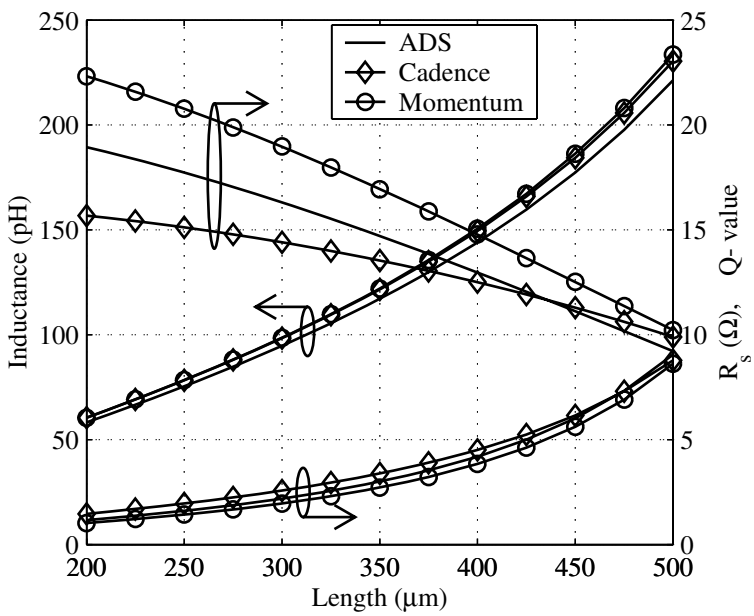


Figure 2: Comparison of transmission line simulation in ADS, Momentum, and Cadence at 60 GHz, width 10.75 μm

The ADS and Momentum simulation results have been compared to the component *tline3* from the library *rfExamples* in Cadence (Fig. 2). The inductance, series resistance, and Q-value have been plotted as a function of TL length at 60 GHz. The

Cadence component agrees well with ADS and Momentum for the inductance, but overestimates the losses by 30% for short TLs. Despite this discrepancy *tline3* has been used to simulate the amplifiers, leading to a somewhat pessimistic gain, but still a correct resonance frequency. The Q-values and inductances are in the same range as in [6, 10].

2.2 Layout Considerations

Metal-insulator-metal (MIM) capacitors were used to provide a good AC-ground at the end of each TL. An array of small, high Q-value, capacitors with a total capacitance of 1 pF was placed closest to the TL, and further out larger capacitors were placed. In total more than 11 pF was used for each AC-ground, see the chip microphotographs, Fig. 3(a) and (b). Under each TL a 20 μm wide metal 1 return path was placed, minimizing current loops in the ground plane. The ground plane was also made in metal 1, and put across the entire chip to provide a low-impedance ground.

To verify that the decoupling was sufficient, simulations including parasitics and bond wires were performed. A π-model consisting of 3 nH series inductance with $Q = 10$ and 50 fF shunting capacitances was used for the bond wires. The circuits showed no stability problems, indicating sufficient decoupling.

3 Measurement Results

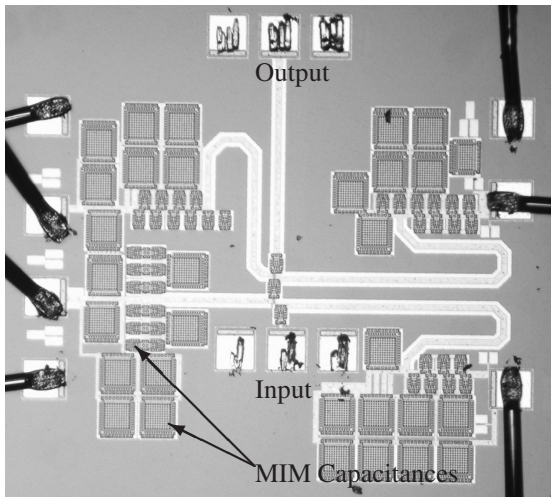
The two amplifiers were measured in a probe station using ground-signal-ground (GSG) and signal-ground-signal (SGS) RF probes[1]. The DC pads were bonded to a PCB providing supply and bias voltages.

Using a network analyzer, S11 of the single-ended input (SI) PA was measured. It was then converted to Z11, input impedance, shown in Fig. 4 together with simulated impedance for both single-ended and differential input (DI). As can be seen the measured input resonates at 32 GHz, with a real valued input impedance of 210 Ω to 250 Ω at resonance. The measured real valued input impedance is 80% of the simulated one. The differential input impedance is therefore estimated to be 275 Ω, also that 80% of the simulated value[2].

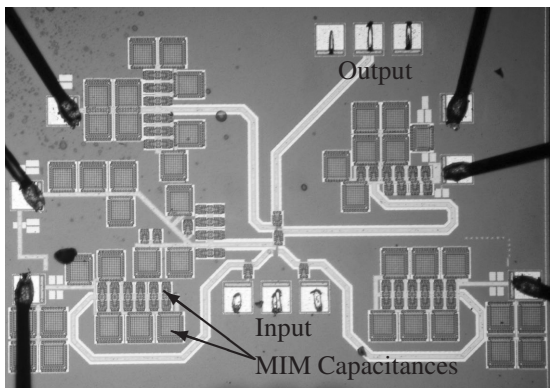
The output power versus input voltage was also measured, Fig. 5, achieving 1.2 dBm saturated output power with a drain efficiency of 7.8% for the SI PA. Deriving the input voltage from the above assumption about the input impedance, the output power versus input voltage for the DI PA was also plotted in Fig. 5. It achieves 3.1 dBm saturated output power with a drain efficiency of 8.2%. As can be seen in

[1] Spectrum analyzer: Rohde&Schwartz FSU 20Hz-50GHz, Harmonic Mixer: Rohde&Schwartz FS-Z75 50-75 GHz, Signal generator: Agilent Technologies E8257D 250kHz-40GHz, Probes: Cascade Infinity and Picoprobe Dual Microwave Probe, VNA: Agilent Technologies E8361A 10MHz-67GHz, Calibration: Cascade Microtech Impedance Standard Substrate and Wincal

[2] The differential input PA could not be measured with the VNA due to calibration difficulties.



(a)



(b)

Figure 3: Chip microphotographs. (a) Single-ended input (b) Differential input

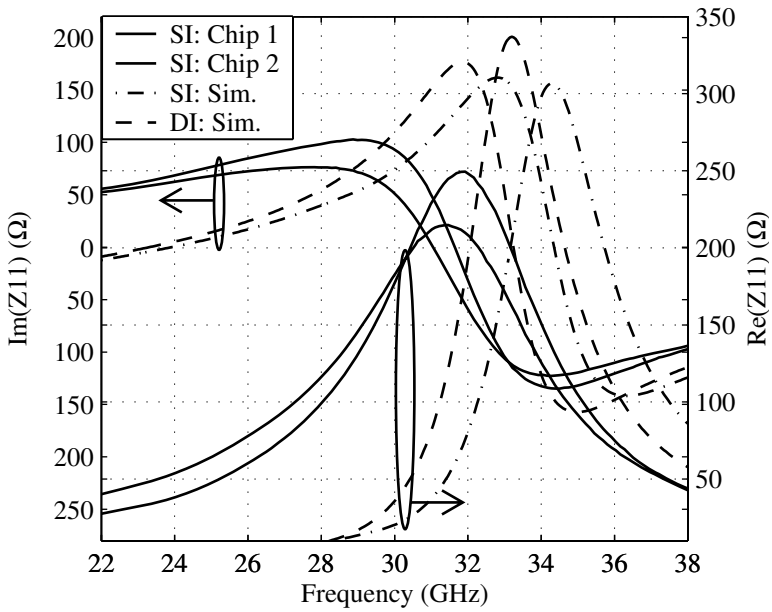


Figure 4: Input impedance, Z11. SI PA: measured and simulated, DI PA: simulated

the figure, the DI PA reaches its saturated output power for lower input amplitudes, making it easier to drive for an oscillator.

The fundamental frequency suppression versus input signal phase difference was measured for the DI PA, using two frequency locked signal generators, showing a good agreement with simulated results, Fig. 6. The measured maximum suppression is 37 dB and it is achieved with differential input. The measured fundamental frequency suppression for the SI PA is larger than 22 dB over the full frequency range.

3.1 Summary

Table 1 is a summary of the simulated and measured results compared to two references [6, 7], both at 60 GHz and in CMOS, but not frequency converting. The authors would like to make a comment about the power gain. In simulations the first stage has a conversion loss of about 4 dB, which should be added to the overall gain when comparing to non frequency converting PAs. Other published 60 GHz power amplifiers in CMOS show about 2–4 dB gain per stage [7, 11, 12], which our amplifiers also have in the last stage. For higher gain one or two more stages at 30 GHz or 60 GHz can be added to the PA.

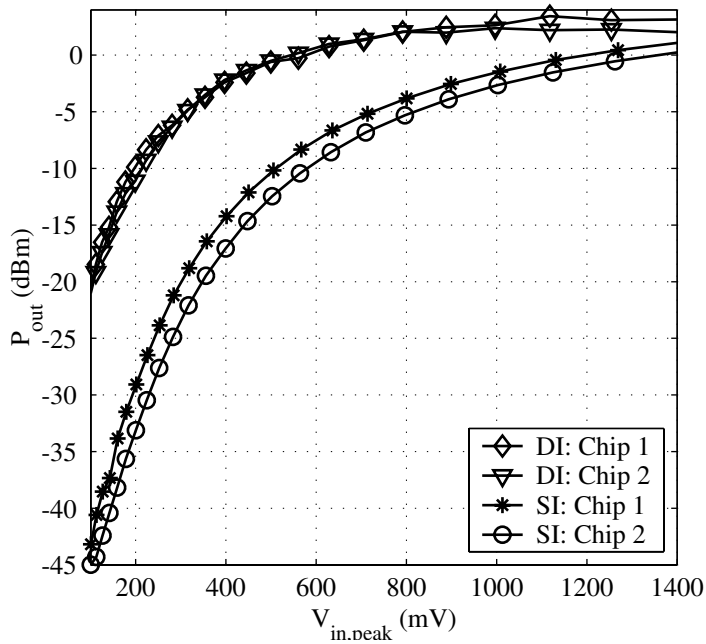


Figure 5: Output power versus input voltage

4 Conclusion

Two 60 GHz frequency doubling power amplifiers in 130-nm CMOS have been presented, one with single-ended and one with differential input. The frequency doubling allows the rest of the transmitter to operate at 30 GHz instead of 60 GHz, which eases the design and lowers its power consumption.

The amplifiers use transistor non-linearity to perform the frequency doubling from 30 GHz to 60 GHz. The differential input amplifier makes use of the fact that the second harmonics at the transistor drains are in phase, thus connecting the drains together (Fig. 1(b)) effectively produces a frequency doubling. A cancellation of the fundamental frequency and the odd harmonics is also achieved by this connection, reducing the feedthrough.

The SI PA shows a 3-dB bandwidth of 12 GHz, well covering the unlicensed frequency band at 57–64 GHz. The SI PA and DI PA reaches 1.2 dBm and 3.1 dBm of saturated output power, respectively, both with a drain efficiency of 8%.

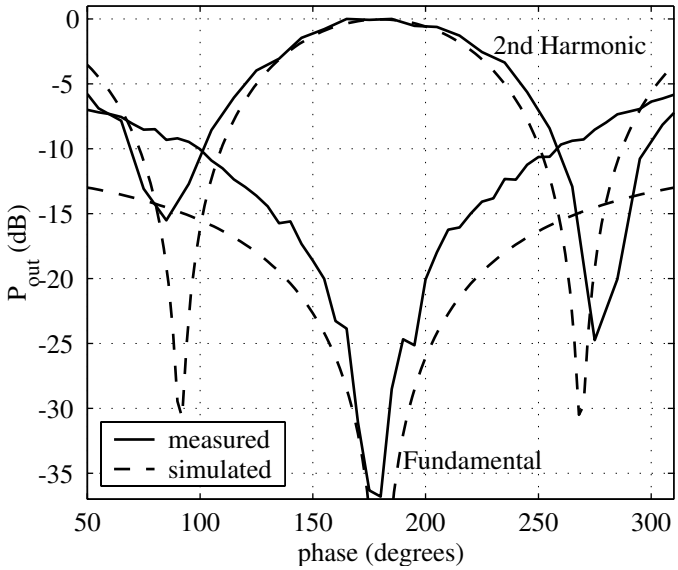


Figure 6: Fundamental tone suppression for the DI PA

5 Acknowledgment

The authors would like to thank united microelectronics corporation (UMC) for giving us the opportunity to work with a state-of-the-art 130-nm CMOS process, and the Knut and Alice Wallenberg foundation for the 60-GHz measurement equipment. Last but not least thanks to the competence center for circuit design (CCCD) and the Swedish Agency for Innovation Systems (Vinnova) for funding this project, 'Techniques for Low Cost 60 GHz WLAN'.

Table 1: Comparison between this Work and Two References in CMOS at 60 GHz

	SI sim.	SI meas.	DI sim.	DI meas.	[6]	[7]
V_{DD} (V)	1.3	1.3	1.2	1.2	1.5	1.5
I_{DC} (mA)	20.5	13.1	27.7	15.7	36	26.5
Max. P_{out} (dBm)	4.0	1.2	5.0	3.1	—	9.3
Drain efficiency (%)	8.6	7.8	9.5	8.3	3[3]	21.4
Power Gain (dB)	−1.1	−3.9	0.4	−0.5	12	5.2
3-dB Bandwidth (GHz)	6.0	12.4	6.8	—	14	13
Fundamental Tone Suppression (dB)	13.8	22.5	—	37	—	—
Number of stages @ 60 GHz	1	1	1	1	3	3
Frequency doubling	yes	yes	yes	yes	no	no
Total Area (mm^2)	—	0.52	—	1.26	1.30	0.15
Process node (nm)	130	130	130	130	130	90

[3] Authors calculation from P_{DC} and OCP1dB.

References

[1] Federal Communications Commission, *Amendment of Parts 2, 15 and 97 of the Commission's Rules to Permit Use of Radio Frequencies Above 40 GHz for New Radio Applications*, FCC 95-499, ET Docket No. 94-124, RM-8308, Dec. 1995.

[2] European Radiocommunications Committee (ERC), *The European Table of Frequency Allocations and Utilisations Covering the Frequency Range 9 kHz TO 275 GHz*, ERC REPORT 25, http://www.ero.dk/documentation/docs/doc98/official/pdf/ ERCREP025.PDF, Copenhagen: 2004.

[3] B. Razavi, "Gadgets Gab at 60 GHz," *IEEE Spectrum*, vol. 45, no. 2, pp. 40–45, Feb. 2008.

[4] S. K. Moore, "Cheap Chips for Next Wireless Frontier," *IEEE Spectrum*, vol. 43, no. 6, pp. 8–9, June 2006.

[5] A. M. Niknejad and H. Hashemi, Eds., *mm-Wave Silicon Technology, 60GHz and Beyond*, 1st ed. Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA: Springer, Feb. 2008.

[6] C. H. Doan, S. Emami, A. M. Niknejad, and R. W. Brodersen, "Millimeter-Wave CMOS Design," *IEEE Journal of Solid-State Circuits*, vol. 40, pp. 144–155, Jan. 2005.

[7] T. Yao, M. Gordon, K. Yau, M. Yang, and S. P. Voinigescu, "60-GHz PA and LNA in 90-nm RF-CMOS," in *IEEE Radio Frequency Integrated Circuits Symposium*, June 2006, pp. 147–150.

[8] E. Cijvat, N. Troedsson, and H. Sjöland, "A 2.4GHz CMOS Power Amplifier Using Internal Frequency Doubling," in *Proceedings of the IEEE International Symposium on Circuits and Systems, ISCAS*, vol. 3, May 2005, pp. 2683–2686.

[9] Agilent, Advanced Design System (ADS), *http://eesof.tm.agilent.com/products/ads_main.html*.

[10] T. Yao, M. Gordon, K. Yau, M. Yang, and S. P. Voinigescu, "A 60-GHz CMOS Receiver Front-End," *IEEE Journal of Solid-State Circuits*, vol. 41, pp. 17–22, Jan. 2006.

[11] C. H. Doan, S. Emami, D. A. Sobel, A. M. Niknejad, and R. W. Brodersen, "Design considerations for 60 GHz CMOS radios," *IEEE Communications Magazine*, vol. 42, pp. 132–140, Dec. 2004.

[12] C. H. Doan, S. Emami, A. M. Niknejad, and R. W. Brodersen, "Design of CMOS for 60GHz Applications," in *IEEE International Solid-State Circuits Conference, Digest of Technical Papers*, vol. 1, Feb. 2004, pp. 440–538.

Paper V

Paper V

Analysis of a High Frequency and Wide Bandwidth Active Polyphase Filter Based on CMOS Inverters

Abstract

An active polyphase filter capable of high frequency quadrature signal generation has been analyzed. The resistors of the classical passive polyphase filter have been replaced by transconductors, CMOS inverters [1, 2].

A three-stage 0.13 μm CMOS active polyphase filter has been designed. Simulations with a differential input signal show a quadrature error less than 1° for the full stable input voltage range for frequencies from 6 GHz to 14 GHz. Phase errors in the differential input signal are suppressed at least 3 times at the output. Corner simulations at 10 GHz show a maximum phase error of 3° with both n- and pMOS slow, in all other cases the error is less than 0.75°. The three-stage filter consumes 34 mA from a 1.2 V supply.

To investigate the robustness of the filter to changes in inverter delay, an inverter model was implemented in Verilog-A. Linear c_{in} and g_{in} were used, whereas g_m, c_{out}, and g_{out} were non-linear. It was found that the filter could tolerate substantial delays. Up to 40° phase shift resulted in less than 1.5° quadrature phase error at the output.

Based on: Johan Wernehag and Henrik Sjöland, "Analysis of a High Frequency and Wide Bandwidth Active Polyphase Filter Based on CMOS Inverters", *Analog Integrated Circuits and Signal Processing*, Springer Science, 2008, Submitted.

1 Introduction

Popular architectures for integrated radio transceivers, like direct conversion and low-IF, require quadrature local oscillator signals. As the carrier frequency increases in the quest for more bandwidth, generating accurate quadrature signals becomes a major challenge. There are a number of ways to generate high frequency quadrature signals, some are: (i) Quadrature oscillator, (ii) Frequency division of a differential oscillator running at twice the frequency. After the frequency division quadratures phases can be obtained, (iii) Differential oscillator followed by a passive polyphase filter, and (iv) Differential local oscillator + active polyphase filter [1, 2].

To build quadrature LC oscillators at high frequencies is possible but more difficult than at lower frequencies, since the coupling of the oscillators adds capacitance to the oscillating nodes and thus reduces the tuning range. Despite this some high frequency LC quadrature voltage controlled oscillators can be found in [3–6]. In ring oscillators the quadrature phases can be obtained more or less for free, but the phase noise is much higher than in the CMOS LC counterparts [7–9]. Using other processes like SiGe the obtainable phase noise will be of the same level as for CMOS LC oscillators [10]. To divide the frequency of an oscillator running at twice the frequency is also possible, but might be difficult with low phase noise at high frequencies. Using a passive polyphase filter and a differential local oscillator is a well proven way of generating quadrature signals. At high frequencies, however, losses become problematic. At frequencies of tens of GHz the component values must be made small, resulting in poor matching and low input impedance. An active polyphase filter has the same demands on low resistance values. These can, however, be implemented by increasing the transconductance, g_m, of the active part in the filter, see Fig. 17. The active elements eliminate the problem of signal attenuation, and also increase the input impedance.

Also after downconversion, at intermediate frequency (IF), quadrature signals are often used, and active or passive polyphase filters can be used to suppress signals at the image frequency [1, 2, 11–14].

In this paper the capability of using an active polyphase filter based on CMOS inverters for high frequency quadrature generation is investigated. Compared to a passive polyphase filter an active one is more complicated, as it contains non-linear devices in the signal path, and although the architecture is known [1] more investigations in this area are needed.

The rest of this paper is organized as follows: in Section 2 the passive polyphase filter is shortly recaptured. In Section 3 the transconductor is examined and a transconductor model is derived and verified against transistor level simulations. Parameters of the model are stretched and the effects on the active polyphase filter are observed. Section 4 shows simulated results of the active polyphase filter in a 130-nm 1p8M CMOS process. Section 5 investigates the stability of the active polyphase filter through PVT simulations.

2 Passive Polyphase Filter

We start with a short summarized of the passive polyphase filter, Fig. 1. It is well known [1, 15, 16] that the voltage gain of this filter is $\sqrt{2}$ for an unloaded stage and $1/\sqrt{2}$ for a stage in a cascaded chain of identical stages. Assuming all R and C to be perfectly matched, a differential input signal at V_I^+ and V_I^- will be converted to a perfect quadrature output signal, if the frequency is exactly equal to $1/2\pi RC$. If quadrature signals must be generated over a wide frequency range, or if process variations are significant, several stages can be cascaded, tuned to different frequencies. Good matching of the components in each stage is essential, as it limits the achievable quadrature accuracy.

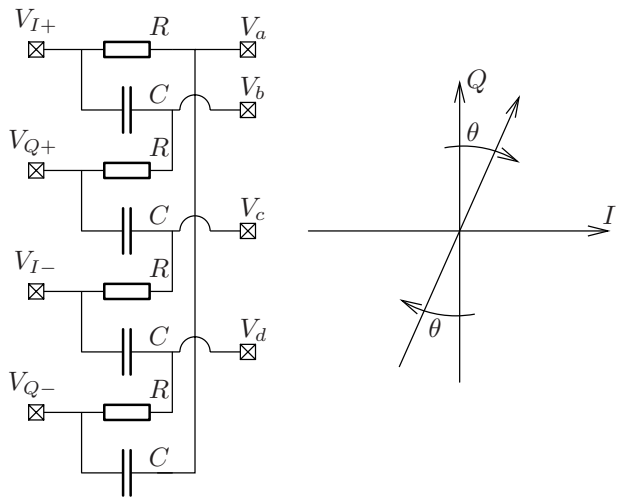


Figure 1: A passive RC polyphase filter

The image rejection ratio (IRR) of a quadrature receiver depends on the quadrature accuracy, and can be written as:

$$IRR \approx \left.\frac{4}{\epsilon^2 + \theta^2}\right|_{\text{Small } \epsilon \text{ and } \theta}, \tag{1}$$

where ϵ denotes the relative amplitude error and θ the quadrature error in radians between the I and Q phases.

In a receiver architecture relying on quadrature downconversion, the achievable image rejection ratio is typically in the range of 30–40 dB [13, 17–19], without calibration schemes. In order to increase the image rejection, very accurate quadrature signals, both phase and amplitude, must be generated. As can be seen in [12] a phase

error of $\pm 5^\circ$ leads to a maximum image rejection of 30 dB, whereas an amplitude error of $\pm 12\%$ is alloved for the same rejection ratio. In practice typically switching mixers are used, and the amplitude error is then less important.

3 Active Polyphase Filter Analysis

In this paper the resistors of the passive polyphase filter are replaced by transconductors. With ideal transconductors g_m should be equal to $1/R$ not to change the frequency of operation. The gain of the active filter can be larger than unity and the load on the preceding stage reduced thanks to the higher input impedance of the transconductor compared to the resistor. Since the active polyphase filter is a feedback structure, with many different feedback paths, the stability is examined in Section 3.1.

The transconductor investigated in this paper is designed using a 1p8M 130-nm CMOS process and an extraction frequency of 10 GHz, but the procedure and derived equations are not dependent on the process, only some numerical values calculated.

3.1 Stability of the Active Polyphase Filter

A three-stage active polyphase filter was designed and simulated. It was found that the filter becomes unstable at small input signal amplitudes. The input amplitude should be larger than 200 mV to stop the filter from self oscillating and instead lock to the input signal. At 280 mV input amplitude the full saturated output amplitude of 600 mV was reached.

When instability occurs the filter oscillates in a differential mode. The inverters in the second stage then forms a 4-stage ring oscillator. The same happens in the third stage. Due to the internal structure of the transconductor, see Fig. 2(a), also a ring with an even number of stages can self-oscillate. When the input signal is too small to force the inverters to switch, the positive feedback of the ring will make them saturate at high or low. The internal feedback through R_f then slowly changes the gate potentials, through the RC time-constant between the feedback resistor and the capacitance at the gate, until the inverters switch. This causes the whole ring to change state. The periodicity is set by the feedback resistance and the capacitance at the gate. The loop gain at the self oscillation frequency decreases with increasing input amplitude and the filter eventually stops oscillating and instead locks to the input signal.

Not having the internal feedback and DC-block, DC-offsets would be amplified and the filter could saturate at high or low potential. Thus there is a trade-off between voltage gain, feedback resistance, and minimum input amplitude. However, this high frequency polyphase filter is intended to transform a differential oscillator signal into a quadrature one, and generally the amplitude from an oscillator is quite large. Thus instability should not be a problem in this application.

3.2 Transconductor

The transconductor is to be used as a component in the high frequency polyphase filter, therefore the fewer internal nodes the better. The transconductor used is this analysis a CMOS inverter with input DC-block and bias voltage feedback (Fig. 2). The purpose of the DC-block and bias feedback is to set the DC potential at the outputs to $V_{DD}/2$, so that any DC offset that might occur does not get amplified in the filter chain.

The analysis is not restricted to this implementation of the transconductor, however, since an extracted model is used in all the analysis and such a model can be extracted for any implementation of the transconductor. However, when numerical results are given these are, of course, for this transconductor.

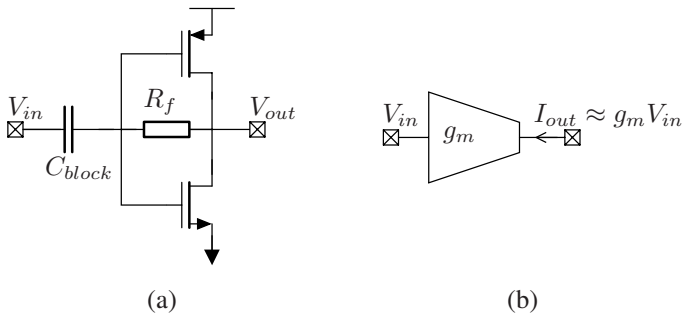
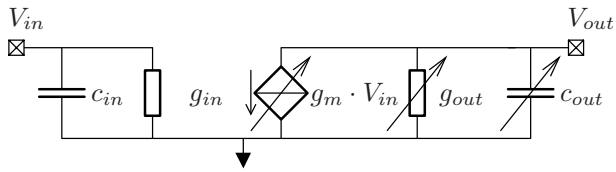


Figure 2: (a) Schematic of the transconductor with internal feedback and DC-block capacitance (b) Symbol of the transconductor

The objective is to extract a model which captures the non-linearities of the transconductor when both input and output amplitude are large, and yet is not more complex than necessary. The approach we took was to start from the transistor behavior. From that it was assumed that the most non-linear elements are the transconductance, g_m, and the output admittance, y_{out}, which is modeled as g_{out} in parallel with c_{out}. The other elements in the model are considered linear in the analysis. The transconductor model is shown in Fig. 3 together with the values extracted at 10 GHz (the extraction frequency $f_0 = 10$ GHz). This model was implemented in a Verilog-A transconductor used to verify the validity of the extraction. The extraction was performed in three steps using transient simulations:

1. The output admittance, amplitude and phase, was simulated versus output voltage amplitude while the input was at its DC bias level. The output conductance and capacitance were fitted to third order polynomials, see Fig. 4.

2. g_m was extracted by stimulating the transconductor with a sinusoidal input voltage and loading the output with a small signal short circuit. The transient output



(a)

$c_{in} = 106.12$ fF
$g_{in} = 0.67$ mS
$|g_m| = 30.51{+}1.27{\cdot}\hat{V}_{in}{-}61.42{\cdot}\hat{V}_{in}^2{+}68.22{\cdot}\hat{V}_{in}^3$ mS
$\angle g_m = 173.71{-}5.31{\cdot}\hat{V}_{in}{+}4.45{\cdot}\hat{V}_{in}^2{-}2.05{\cdot}\hat{V}_{in}^3$ degree
$g_{out} = 6.68{+}4.27{\cdot}\hat{V}_{out}{-}25.62{\cdot}\hat{V}_{out}^2{+}40.87{\cdot}\hat{V}_{out}^3$ mS
$c_{out} = 101.36{+}49.98{\cdot}\hat{V}_{out}{-}324.76{\cdot}\hat{V}_{out}^2{+}581.48{\cdot}\hat{V}_{out}^3$ fF

(b)

Figure 3: Model extracted for the transconductor at $f_0 = 10$ GHz. (a) A lumped model (b) The corresponding parameters

current and input voltage were then used to calculate the g_m polynomial, see Fig. 5.

3. The small signal c_{in} and g_{in} were extracted from a transient simulation. By observing the magnitude and phase relation of the input voltage and current the complex admittance was calculated and g_{in} and c_{in} extracted.

The voltage gain of the transconductor was simulated (Fig. 6). There is a good agreement between the extracted model and transistor up to about 50% of full swing. The error at large amplitudes is due to the extraction procedure, where large signals were not simultaneously applied to input and output. To capture this a more complicated model would be necessary, but for our purposes the level of accuracy achieved is sufficient.

To give an indication of the level of non-linearity, the four non-linear parameters, normalized with their small signal values, are plotted in Fig. 7. The magnitude of g_m is within 20% of its small signal value over the amplitude range, g_{out} and c_{out} are within 20% of their small signal values, up to an amplitude of 510 mV, which corresponds to 85% of the full voltage swing. The phase of g_m only changes by a few degrees over the full range (Fig. 5).

3.3 Voltage Gain

To calculate the voltage gain of a stage in the active polyphase filter Fig. 8 was used. By applying Kirchoff's current law in node a, by inspection one gets (2a). Inserting the expressions for the currents (2b)–(2e) and solving (2a), results in the voltage

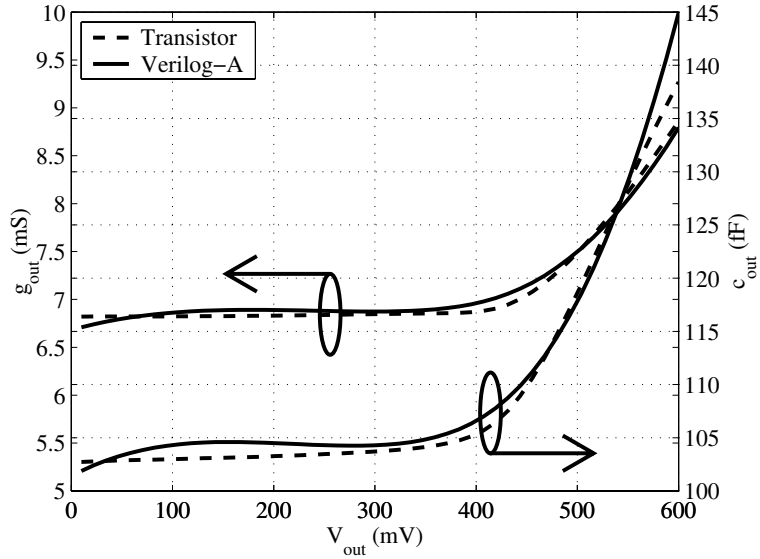


Figure 4: Output admittance extracted as g_{out} and c_{out} compared to the transistor model

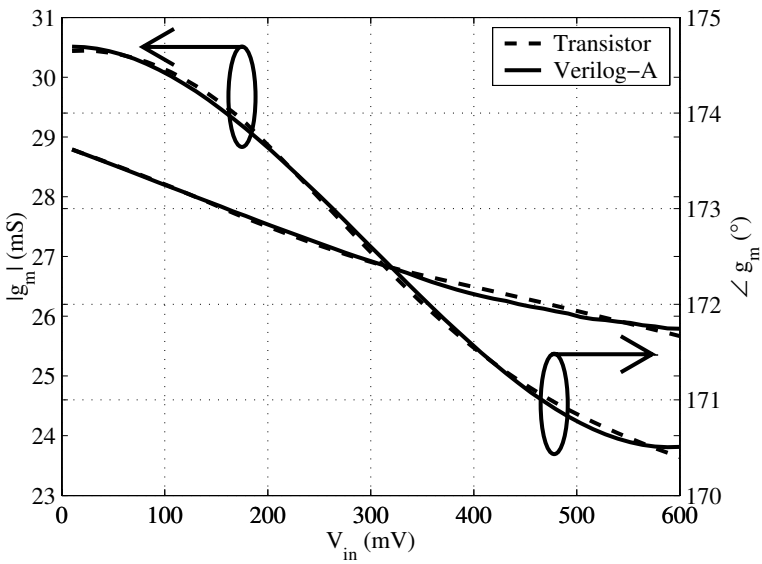


Figure 5: The transconductance fitted to the Verilog-A model and compared to the transistor model

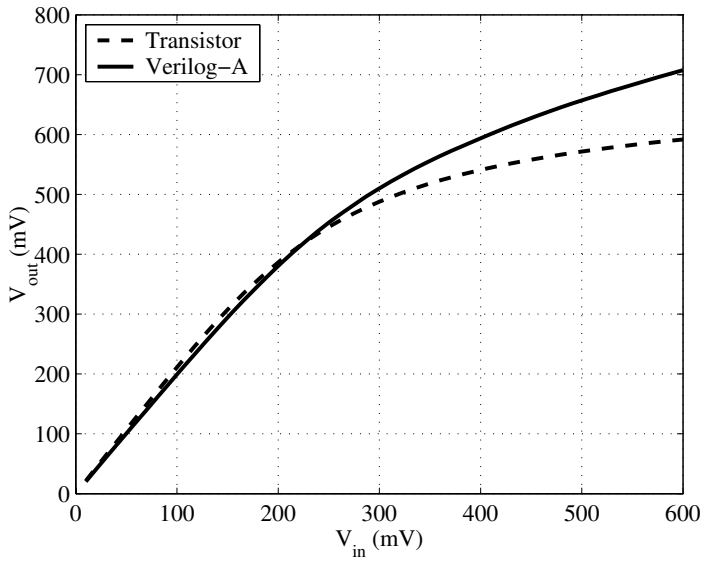


Figure 6: Output voltage magnitude versus input magnitude for both transistor and Verilog-A implementation. Full voltage swing is 600 mV

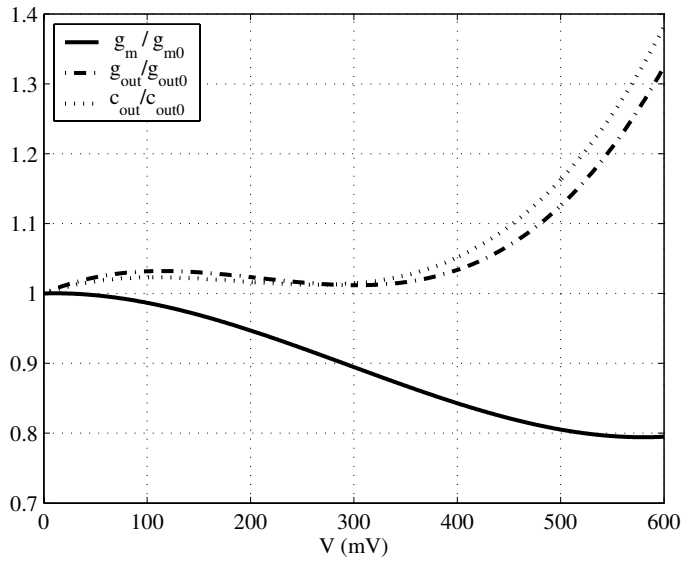


Figure 7: Plot of three of the non-linear transconductor parameters as they change with the voltage at input or output of the transconductor

gain (3). At the designed frequency, where $g_m = \omega_0 C$, (3) can be simplified to (4). The unloaded voltage gain with ideal transconductor ($Z_{out} = \infty$) can then be found to be equal to 2, as in [1]

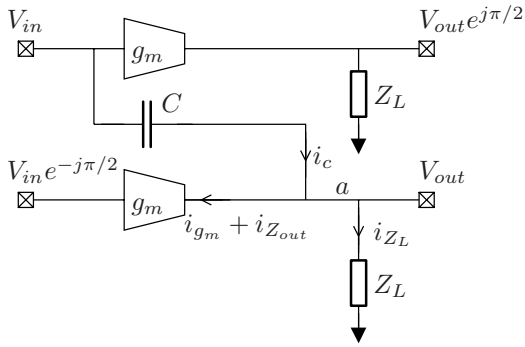


Figure 8: Schematic of a part of the polyphase filter suitable to calculate the voltage gain of a stage

$$0 = i_c - i_{g_m} - i_{Z_{out}} - i_{Z_L} \tag{2a}$$
$$i_c = (V_{in} - V_{out})\, j\omega C \tag{2b}$$
$$i_{g_m} = g_m V_{in} e^{-j\frac{\pi}{2}} \tag{2c}$$
$$i_{Z_{out}} = V_{out}/Z_{out} \tag{2d}$$
$$i_{Z_L} = V_{out}/Z_L \tag{2e}$$

$$A_V = \frac{V_{out}}{V_{in}} = j\frac{\omega C + g_m}{j\omega C + 1/Z_{out} + 1/Z_L} \tag{3}$$

$$A_V|_{\omega C = g_m} = j\frac{2g_m}{jg_m + 1/Z_{out} + 1/Z_L} \tag{4}$$

Equation (2e) through (4) contain the load impedance, Z_L. In a chain of cascaded stages, the load impedance can be calculated as Z_{in} of the transconductor in parallel with the impedance looking into the filter capacitance, C. Making the observation that the voltage is in phase[1] at both sides of the capacitance leads to an equivalent

[1]The input of the second transconductor in Fig. 17 is 270° out of phase with respect to the one above, the transconductor inverts the signal to the output changing the phase to 90°. The output current is integrated reducing the phase by 90°. Thus the signals are in phase.

impedance of the filter capacitance:

$$Z_C = \frac{1}{j\omega C\left(A_V - 1\right)} \tag{5}$$

The equivalent impedance of the filter capacitance is determined by the voltage gain of the stage. When the voltage gain is high the capacitance is increased. When the filter is saturated ($A_V = 1$) the capacitance become a virtual open circuit. The effect can be expressed as charge efficiency,

$$\frac{C_{in}}{C_{eqv.}} = \frac{C_{in}}{\left(A_V - 1\right)C}, \tag{6}$$

which is a measure of the energy it takes to make a transition for the transconductor compared to charging the filter capacitances. The charge efficiency is simulated for a three-stage active polyphase filter in Section 3.5. When the polyphase filter is saturated the charge efficiency becomes infinite.

With (5) the load impedance, or rather admittance, can be expressed:

$$Y_L = Y_{in} + Y_C = Y_{in} + \left(A_V - 1\right) j\omega C \tag{7}$$

Assuming the stage is loaded by an identical one, the voltage gain (3) can be solved by inserting (7) and solving for A_V:

$$A_V = -\frac{Y_{out} + Y_{in}}{2j\omega C} \pm \sqrt{\left(\frac{Y_{out} + Y_{in}}{2j\omega C}\right)^2 + 1 + \frac{g_m}{\omega C}} \tag{8a}$$

$$= -\frac{g_{out} + g_{in} + j\omega\left(c_{out} + c_{in}\right)}{2j\omega C} \pm$$

$$\sqrt{\left(\frac{g_{out} + g_{in} + j\omega\left(c_{out} + c_{in}\right)}{2j\omega C}\right)^2 + 1 + \frac{g_m}{\omega C}} \tag{8b}$$

The negative root in (8) is discarded since it yields a negative voltage gain. Imposing small signal and low frequency ($SSLF$) conditions reduces (8b) to:

$$A_V|_{SSLF} = -\frac{g_{out} + g_{in}}{2j\omega C} + \sqrt{\left(\frac{g_{out} + g_{in}}{2j\omega C}\right)^2 + 1 + \frac{g_m}{\omega C}} \tag{9}$$

From (9) the $SSLF$ voltage gain at the designed centre frequency ($g_m = \omega C$) and with ideal transconductor, that is $g_{in} = g_{out} = 0$, equals $\sqrt{2}$. From (4) the unloaded voltage gain, when Z_L is infinite, is equal to 2, also that with ideal transconductor. This agrees with the findings in [1].

In Fig. 9(a) the voltage gain (8), is plotted together with the $SSLF$ voltage gain (9), and in Fig. 9(b) the simulated voltage gain is compared to (8) for an input amplitude of

440 mV[2]. In Fig. 9(a) the achievable voltage gain is limited by the $SSLF$ curve and can approach that for an ideal transconductor. For the implemented transconductor the small signal gain is just above 1.2 at the centre frequency. The correspondence between the simulated voltage gain and (8) is excellent from 8 to beyond 14 GHz, see Fig. 9(b).

3.4 Phase Analysis

The output phases from a stage in the filter are analyzed using the schematic in Fig. 7. It is assumed that Z_L is equal to Z_{in}, which is valid inside a saturated filter. The output voltages, V_A and V_B, can then easily be expressed as (10a)–(10b),

$$\begin{aligned} V_A =& \frac{1}{Z+Z_C}\left(XZ_Ce^{-j\theta} + Ze^{-j\theta} - ZZ_Cg_me^{-j\pi/2}\right)\cdot \\ & V_{in}e^{-j\pi/2} = \frac{1}{Z+Z_C}Q_AV_{in}e^{-j\pi/2} \qquad (10a) \\ V_B =& \frac{1}{Z+Z_C}\left(XZ_C - ZZ_Cg_me^{-j(\pi/2+\theta)} + Z\right)V_{in} \\ =& \frac{1}{Z+Z_C}Q_BV_{in} \qquad (10b) \\ Q_A =& XZ_Ce^{-j\theta} + Ze^{-j\theta} - ZZ_Cg_me^{-j\pi/2} \qquad (10c) \\ Q_B =& XZ_C + Z - ZZ_Cg_me^{-j(\pi/2+\theta)}, \qquad (10d) \end{aligned}$$

where X and Z are given by (11a)–(11b) and θ is defined in Fig. 1 as the input quadrature phase error. For quadrature output signals the phase of Q_A and Q_B should be equal, which happens only if $\theta = 0$.

$$X = \frac{Z_{out} \parallel Z_{in}}{Z_{out} \parallel Z_{in} + Z_C} \qquad (11a)$$

$$Z = Z_{out} \parallel Z_{in} \qquad (11b)$$

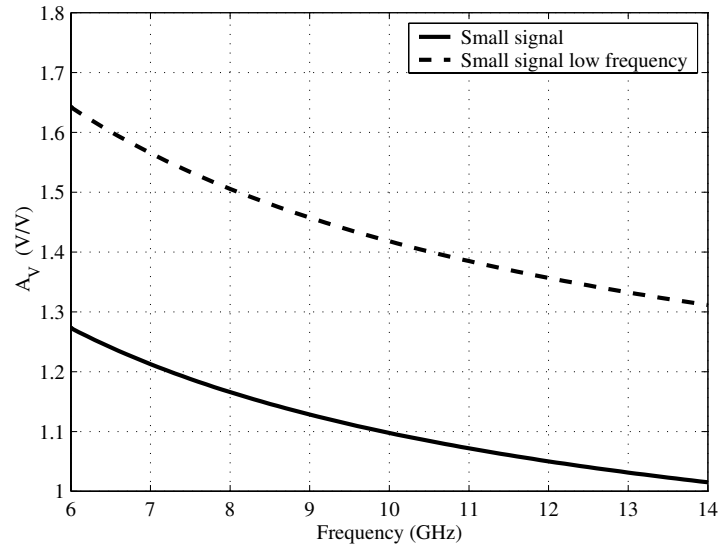
The phase difference, $\angle Q_A - \angle Q_B$, as a function of input quadrature phase error is plotted in Fig. 11(a).

To verify the analytic expressions a comparison to simulated phase error was performed for two different input amplitudes, Fig. 11(b). The agreement is very good, justifying the assumption of Z_L being equal to Z_{in} made in the derivation.

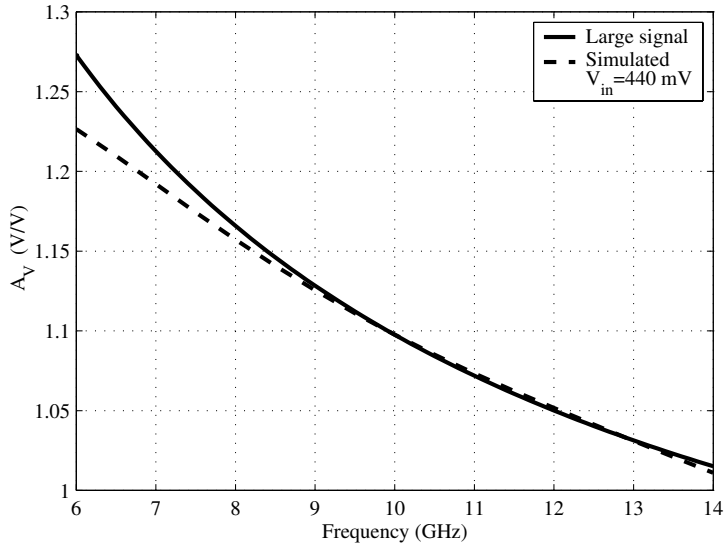
3.5 Model Simulations of the Polyphase Filter

In this section the three-stage active polyphase filter (Fig. 17) is simulated using the Verilog-A transconductor model from Section 3.2. Five aspects are investigated, the

[2]An iteration has been performed to accurately calculate the large signal g_{out} and c_{out}.



(a)



(b)

Figure 9: (a) Small signal voltage gain of the active polyphase filter (8) and small signal low frequency approximation (9) (b) Analytic large signal voltage gain (8) (solid), and simulated voltage gain (dashed), $V_{in} = 440$ mV

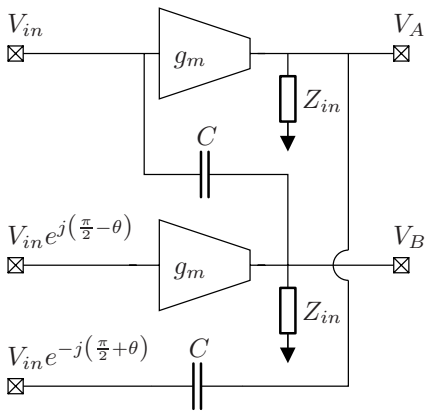


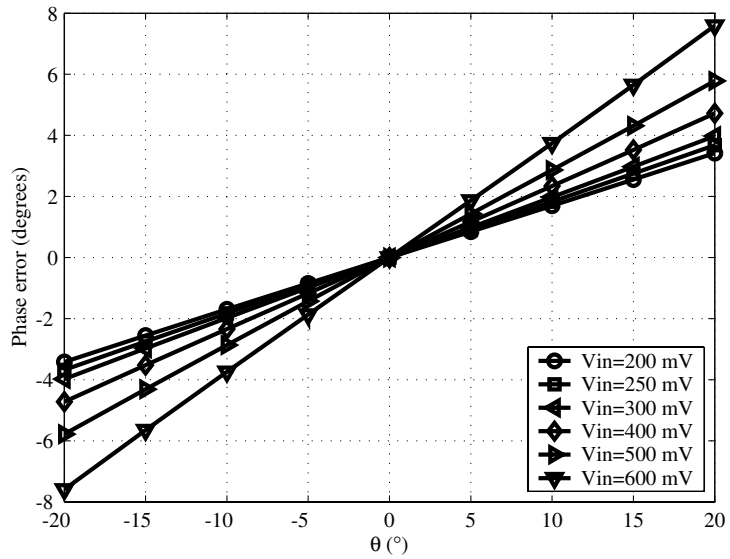
Figure 10: Schematic used to calculate the output phase after a filter stage

first three with quadrature input signal: (i) voltage gain versus transconductor phase delay, (ii) effective capacitance in the filter versus phase delay, (iii) charge efficiency versus frequency with scaled filter capacitance, (iv) quadrature phase error after three stages versus phase delay with differential input signal, and (v) voltage gain after the three stage filter versus phase delay with differential feed.

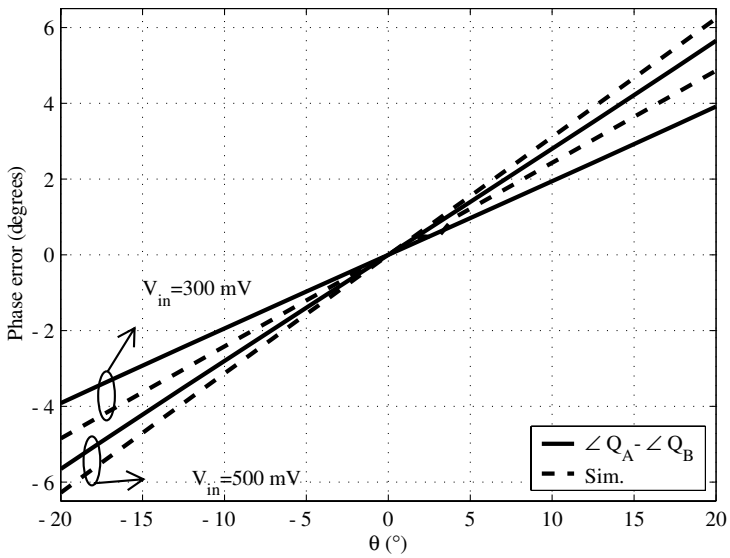
The voltage gain versus phase delay of the three-stage active polyphase filter is plotted in Fig. 12. At low input amplitudes the filter is unstable for large phase delays. The curves are therefore cut. The extra phase delay is added to the extracted phase of g_m in the previous section. The voltage gain reaches a maximum at about 30° phase delay for small input signals, and up to 70° for large signals.

The filter capacitances were set to $C = g_{m0}/\omega_0 = 486$ fF. The effective capacitance was calculated from the reactive part of the impedance looking into a filter capacitance. The phase delay was swept from 0 to 90 degrees for different input amplitudes. For phase delays above 50 degrees the effective capacitance becomes negative for low input amplitudes and the curves are cut, see Fig. 13. The curves in Fig. 12 and Fig. 13 show an interesting resemblance.

The effective capacitance increases with phase delay until it reaches a peak, then decreases and becomes negative. The explanation is that with no additional phase delay the current injected by g_m to the output load is almost 90 degree out of phase with respect to the voltage at the other side of the filter capacitance, see Fig. 17. When the current is fed into the mainly capacitive impedance at the output, the output voltage will be almost in phase with the voltage at the other side of the filter capacitance. If the gain is close to unity the signal voltage over the capacitance is then small, and so is the effective capacitance. A low effective capacitance will result in a reduced power



(a)



(b)

Figure 11: (a) Sensitivity of the phases of Q_A and Q_B with respect of the quadrature offset, θ (b) Comparison of analytic expression (solid) and simulation (dashed)

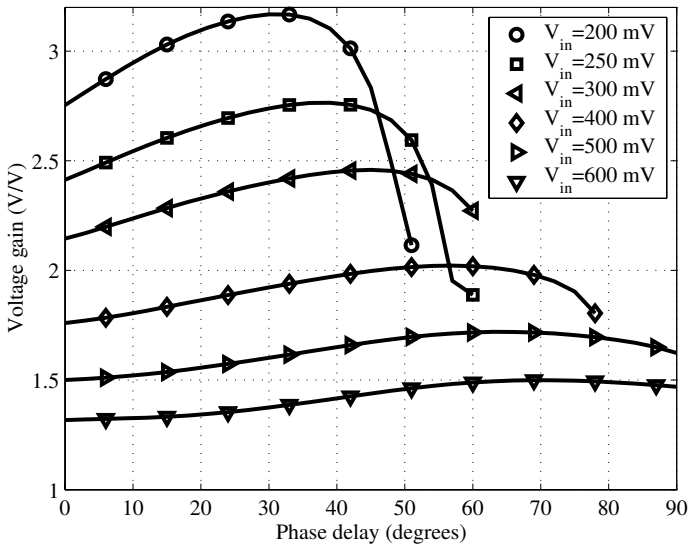


Figure 12: Voltage gain of the three-stage active polyphase filter versus phase delay of the transconductor for six different input amplitudes

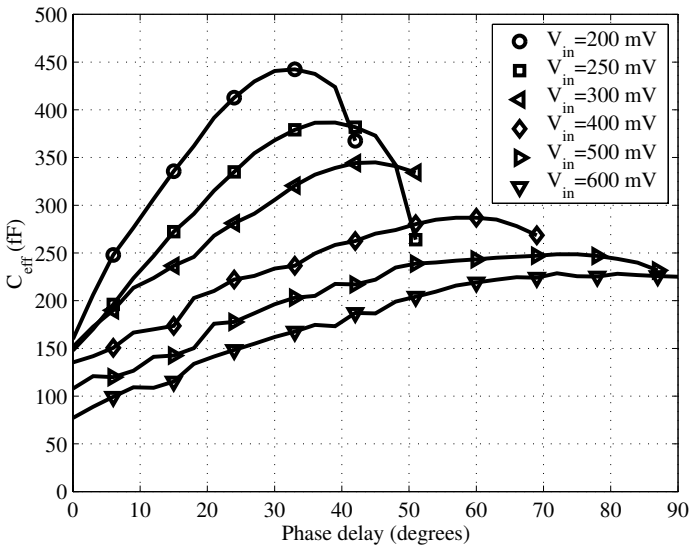


Figure 13: Effective filter capacitance of the three-stage active polyphase filter versus phase delay of the transconductor for six different input amplitudes

consumption and an increased capability of high frequency operation. The effective capacitance is not zero at zero additional delay, however, since there is a phase delay in g_m at 10 GHz, and that even at small amplitudes the effect of g_{out} and g_{in} is not negligible, preventing the nodes from being purely capacitive. The effect of g_{out} increases with amplitude, which is the reason why the effective capacitance does not drop until a larger phase delay is applied.

The charge efficiency is defined as the ratio between the input capacitance of the transconductor, c_{in}, and the effective capacitance in the filter, $C_{eff} = C\,(A_V - 1)$. The charge efficiency was simulated versus frequency, where the filter capacitance was scaled with frequency, $C = g_{m0}/\omega$, and the effective capacitance was extracted. Since the frequency was changed also the phase delay had to be changed accordingly. The delay in the transistor can be expressed as, [20]:

$$y_m = \frac{g_m}{1 + j\omega\tau_1 + (j\omega\tau_2)^2 + \cdots}, \tag{12}$$

where all time-constants, τ, are $\propto 1/f_T$. Staying well below f_T, (12) can be reduced to (13):

$$y_m \approx \frac{g_m}{1 + j\omega\tau_1}, \tag{13}$$

where $\tau_1 = \frac{4}{15 \cdot \omega_t}$. The transit frequencies were simulated for the transistors constituting the transconductor. The f_T used in (13) is the 43 GHz f_T of the pMOS. Equation (13) then gives a delay of 3.55° compared to the modeled 6.29°. In the simulation $\angle g_{m0}$, where g_{m0} is the small signal component of g_m, were given as $\omega \cdot \tau_1$ instead of the extracted value and the simulations were run up to 19 GHz. At this frequency τ_1/τ_2 is 13, still making the approximation (13) valid.

The simulation shows that the charge efficiency varies from 0.8 to infinity with frequency and input amplitude, see Fig. 14. The charge efficiency is smaller (1.0–2.0) at midrange frequencies and larger at higher and lower frequencies. As expected from the effective capacitance plot the charge efficiency is larger for large input amplitudes. The high charge efficiency at high frequencies indicates that the active polyphase technique can be used at frequencies close to the maximum of a chain of cascaded inverters. The addition of the cross-coupling capacitors will not slow the inverters down significantly. This also means that the additional power consumption due to the capacitors will be small.

The influence of phase delay in the transconductor on the quadrature accuracy was investigated with the three-stage active polyphase filter fed by a differential input signal (Fig. 15). The extra phase delay was, as before, added to the delay of g_m. For input amplitudes larger than 300 mV (V_{DD}/4) the quadrature error first increases to a maximum at 15 to 35 degrees delay and then decreases. The quadrature error for smaller signals starts close to zero and then increases with the phase delay.

A sweep of the delay over a full 360° revolution was then performed. Also the voltage gain was plotted this time (Fig. 16). The gain is still high with differential feed,

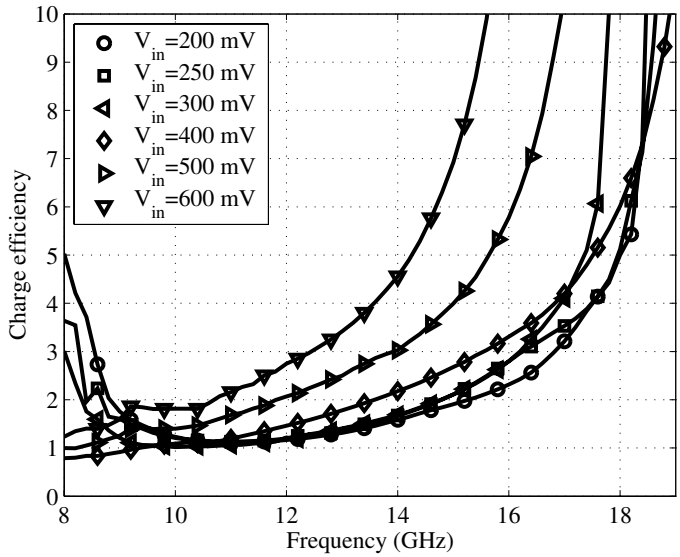


Figure 14: Charge efficiency of the three-stage active polyphase filter for a frequency range up to approximately $f_T/2$

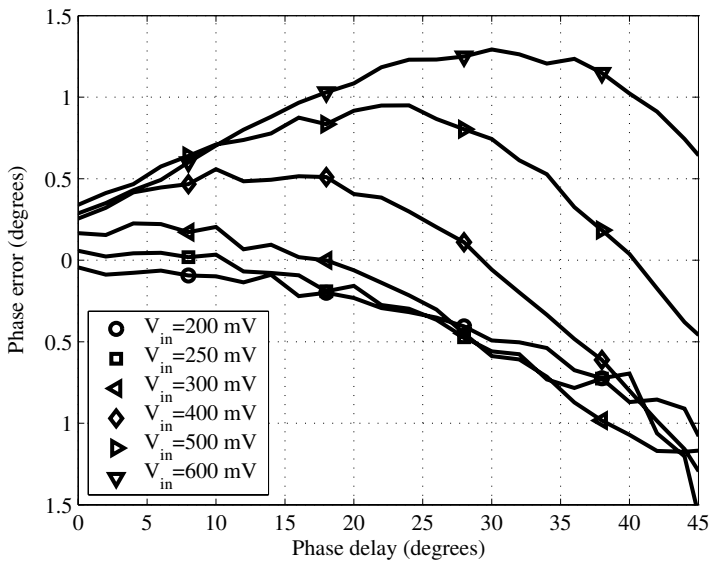


Figure 15: Quadrature phase error vs. phase delay

it is reduced by less than 15% compared to quadrature feed (Fig. 15). Only the stable parts where the filter is locked to the input signal is shown. For phase delays in ranges above 0 and 180 degrees the polyphase filter is locked and achieves a quadrature error of less than 4°. When the filter is locked at 180° phase delay the polarities of the outputs are reversed. Surprisingly, the filter is unstable for phase delays just below 360 and 180 degrees.

4 Active Polyphase Filter Simulation

The three-stage active polyphase filter (Fig. 17) was simulated on transistor level to evaluate the performance and verify some of the findings of the previous section. The capacitances in the stages are decreasing in size, tuning the three stages to 8, 10, and 12 GHz by weighting the capacitances as {1.2 1.0 0.8} with respect to the nominal value found at the extraction frequency of 10 GHz. The power consumption of the three-stage filter was 40.6 mW from a 1.2 V supply.

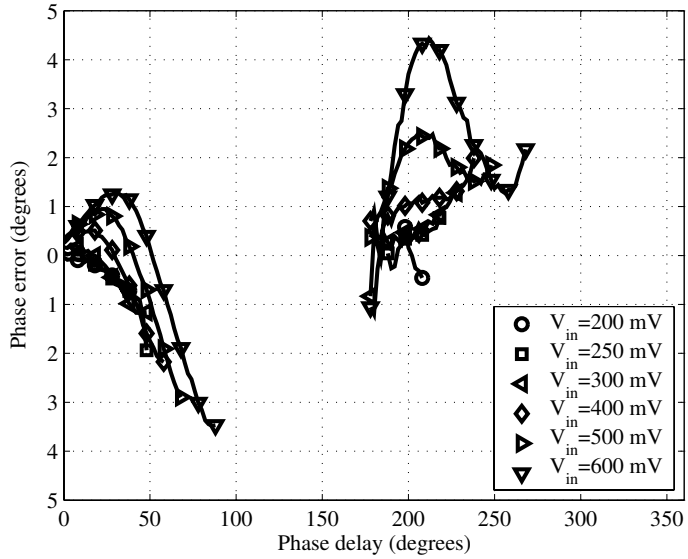
Transient simulations in Spectre were used to evaluate the filter. The phase relations were calculated from the time of the zero crossings of the voltage waveforms, that is their crossings with the bias level, after that the initial transients had settled for 100 periods. All phases are relative to the phase of V_A in Fig. 17, called I+ hereafter.

The following simulations were performed:

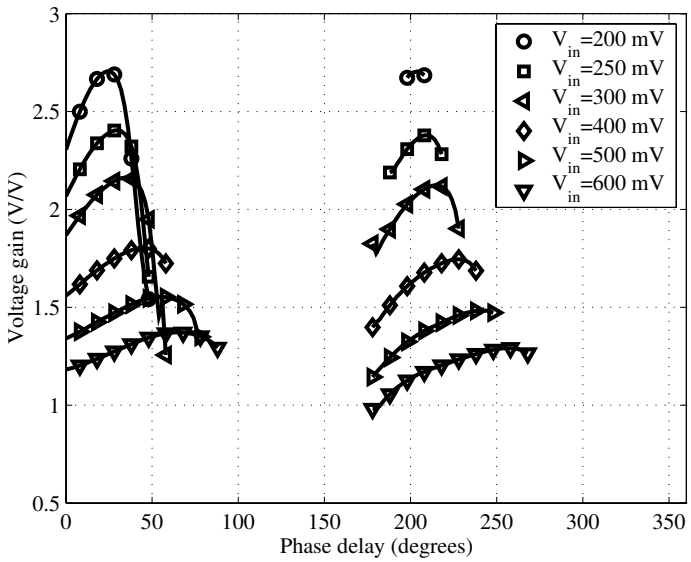
- Quadrature phase error versus input voltage and frequency with differential feed of the filter, Fig. 18.
- The voltage gain with differential feed for the same frequency and amplitude range as above, Fig. 19.
- Quadrature phase error versus input voltage and input differential phase error, δ, with differential feed of the filter (Fig. 20(a)).
- To verify the robustness of the active filter, corner simulations were performed for five process corners, Fig. 22. The five corners are: tt – typical n- and pMOS, ss – slow n- and pMOS, ff – fast n- and pMOS, sf – slow nMOS and fast pMOS, and fs – fast nMOS and slow pMOS. For each corner the input amplitude was swept at 10 GHz.
- The phase noise was simulated using the SpectreRF pnoise analysis and compared to an equivalent passive polyphase filter at 10 GHz, Fig. 21. Equivalent means that the passive filter resistances were set to $R = 1/g_{m0}$.

The differential feed was realized by connecting the positive phase of the input signal to the 0° and 90° inputs in Fig. 17 and the negative to the two other inputs.

As can be seen in the plot in Fig. 18(a) the quadrature phase error is less than 1° from 6 GHz to 14 GHz over the stable input voltage range. The non-stable range has



(a)



(b)

Figure 16: Simulation of (a) Quadrature phase error (b) Voltage gain versus phase delay

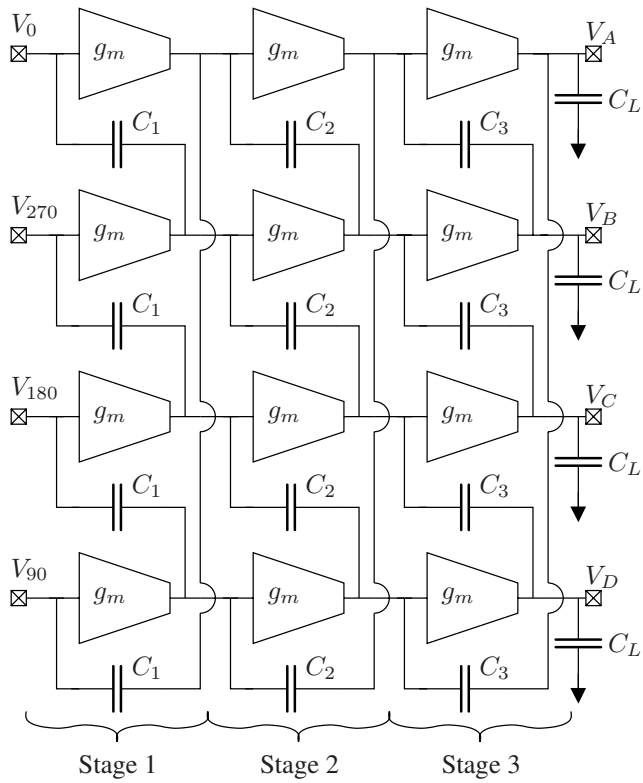
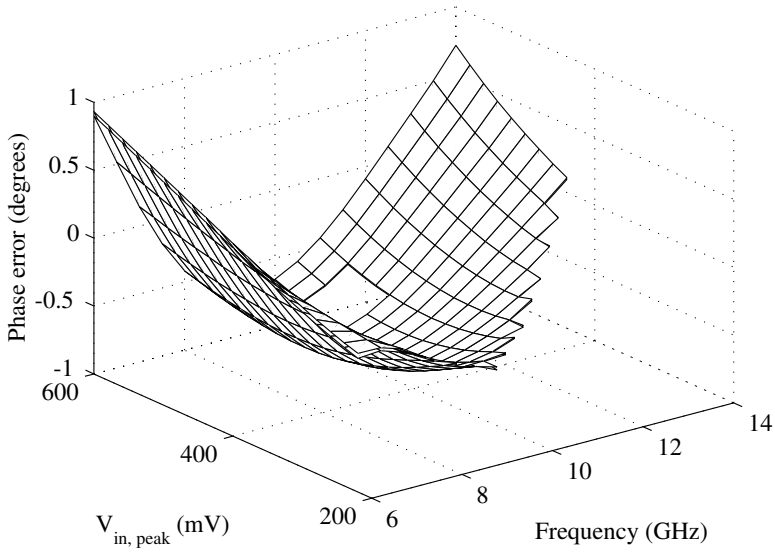


Figure 17: An active three-stage polyphase filter, where the capacitances $C_{1,2,3}$ are weighted as $\{1.2\ 1.0\ 0.8\}\times\ g_{m0}/\omega_0$

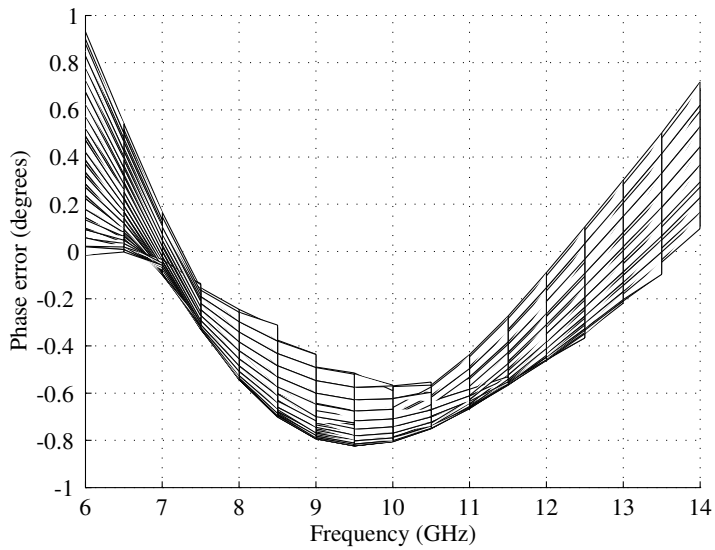
been left out in Fig. 18 to Fig. 20, hence the absence of data for some parts of the plots. There are also some missing data points due to simulation convergence issues. The phase error is frequency dependent, which makes it interesting to view Fig. 18(a) in the phase error-frequency-plane, Fig. 18(b). One can clearly see that the quadrature error is below 1.0° for all stable amplitudes and frequencies.

As can be seen in Fig.19, the voltage gain at 10 GHz ranges from 5.2 to 0.2 dB when the filter is stable, ($V_{in} = 280$–600 mV). The filter is stable for $V_{in} = 200$ mV up to 6.5 GHz with a gain of more than 9 dB. At higher frequencies, larger amplitudes are required to lock the filter to the input signal. At 14 GHz 440 mV is needed, and the gain is then also lower, see Fig. 20(b).

In Fig. 20(a) the output quadrature error is plotted versus input differential errors and input amplitude. Over the full stable input voltage range the output phase error is below $\pm 6.5°$ for differential errors δ up to $\pm 20°$, that is the attenuation of errors



(a)



(b)

Figure 18: Transistor simulation results of the active polyphase filter. (a) Quadrature error vs. V_{in} and frequency (b) Quadrature error vs. V_{in} for the full frequency range. Error of Q$-$ and Q$+$ phases shown with respect to I$+$

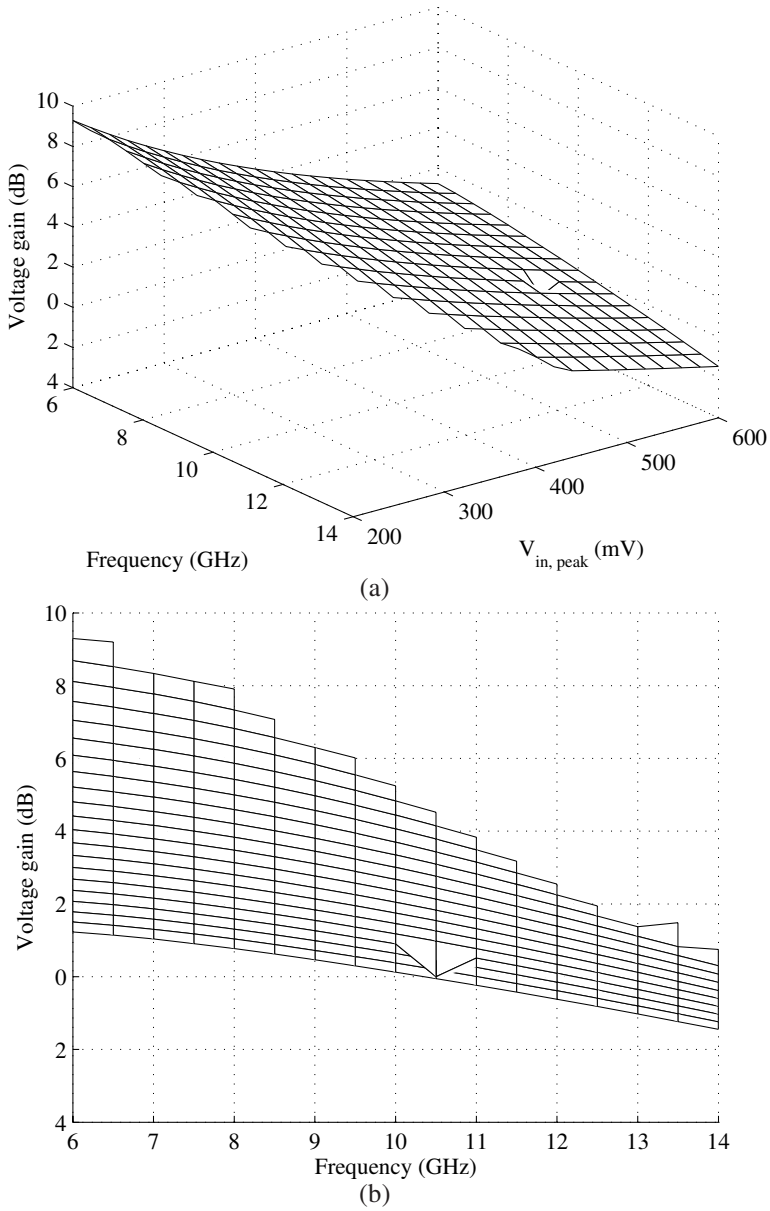


Figure 19: Transistor simulation results of the active polyphase filter. (a) Voltage gain vs. frequency and input V_{in} (b) Voltage gain orthogonal to the frequency axis

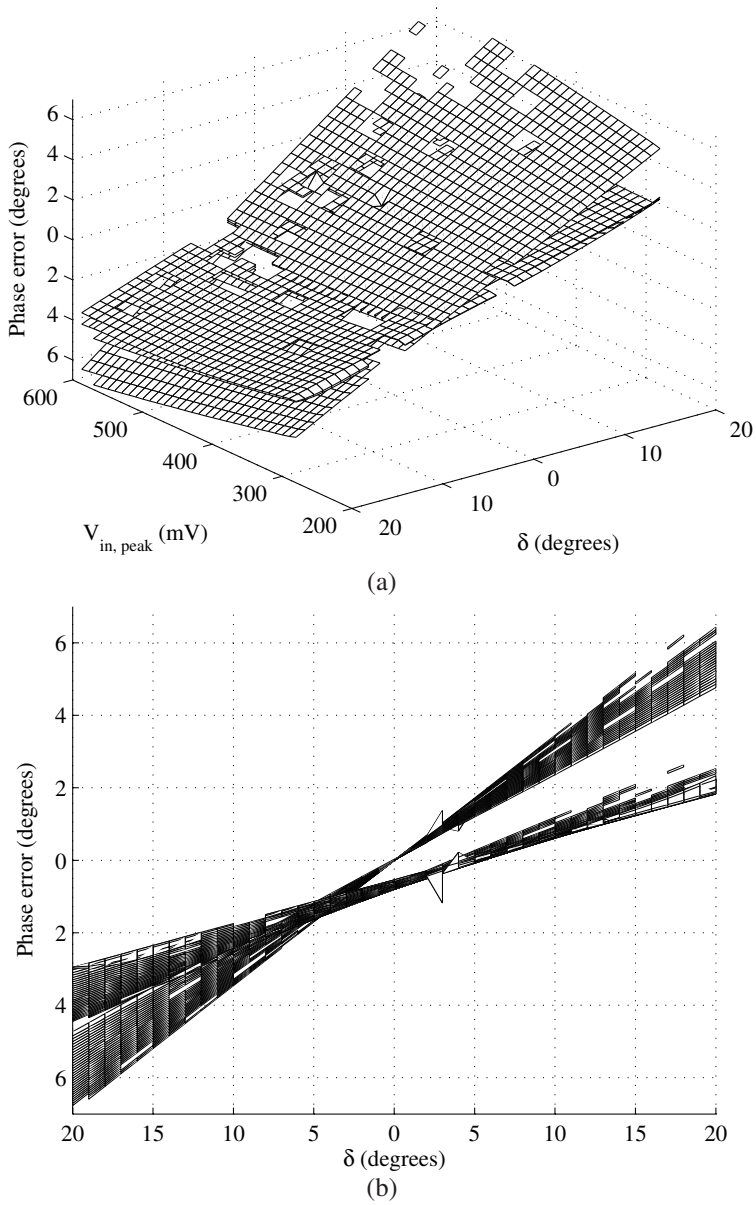


Figure 20: (a) Quadrature phase error vs. V_{in} and δ at 10 GHz (b) The quadrature error for all three phases vs. δ for full range input amplitudes

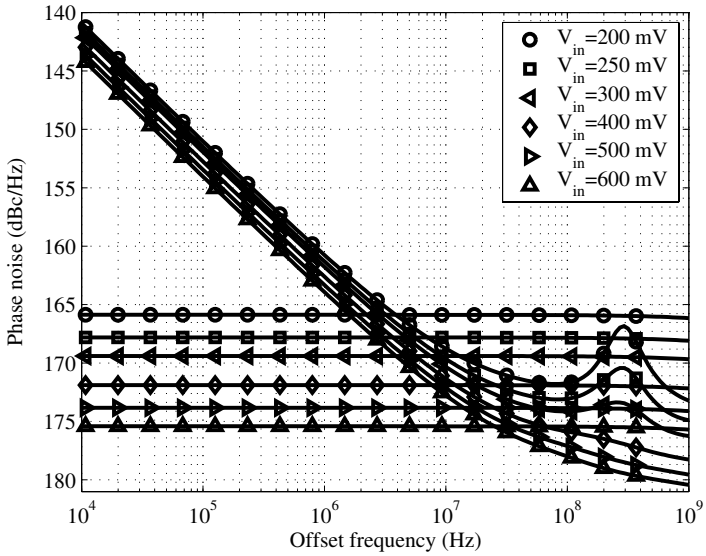


Figure 21: Noise simulation of the active three-stage polyphase filter and its passive counterpart at 10 GHz

is at least 3 times. The error increases somewhat with the input amplitude, approximately by 2° for large mismatches. This is illustrated by the thickness of the curves in Fig. 20(b), which is Fig. 20(a) viewed in the phase error-δ-plane. There one can also clearly see that the quadrature phase error is within $\pm 6.5°$ for $\delta = \pm 20°$ and full input amplitude range.

The phase noise of the passive and active polyphase filters are plotted in Fig. 21. The noise of the passive filter is constant and depends solely on the resistances in the filter. Thus the phase noise level is constant versus offset frequency and decreases as the input amplitude increases, and as expected there is a 6 dB difference between $V_{in} = 300$ mV and $V_{in} = 600$ mV. For the active polyphase filter the phase noise has a $1/f$ behavior below 10 MHz offset.

The far out phase noise of the active polyphase filter is about 5 dB better than the passive implementation, flattening out at -173 dBc/Hz ($V_{in} = 200$ mV) to -180 dBc/Hz ($V_{in} = 600$ mV).

5 Active Polyphase Filter Robustness Simulations

In this section the robustness of the active polyphase filter is examined through different simulations, usually called process, voltage, and temperature (PVT) simulations.

Two different process simulations are performed (i) corner simulation and (ii) statistical simulations. Since there are no Monte Carlo data available to us, the physical dimensions of the transistors have been varied in a random fashion in the statistical simulations.

5.1 Process Variations

The effect of process variations on the phase error is investigated by corner and statistical simulations of the active three-stage polyphase filter at its centre frequency.

The variations of V_T, $I_{d_{sat}}$, and f_T in the different process corners are given in Table 1. The quadrature error of the Q−-phase is plotted in Fig. 22 at 10 GHz, for the five corners including the typical in the process. The Q−-phase was chosen since it shows the largest quadrature error in Fig. 18(b). The simulations show that the largest quadrature phase error is 4°, when both n- and pMOS are slow. For all other simulated corners the error is below 1.5°. It is worth noticing that the quadrature errors for the corners are in the same order as the speed(↑), current(↑), and threshold(↓) for the corners, that is high speed results in low quadrature error and vice versa.

Table 1: Corner parameters, variation relative to tt

		pMOS	nMOS			pMOS	nMOS
ss	V_T	22%	18%	sf	V_T	−11%	9%
	$I_{d_{sat}}$	−24%	−15%		$I_{d_{sat}}$	8%	−7%
	f_T	−22%	−19%		f_T	10%	−11%
fs	V_T	11%	−10%	ff	V_T	−22%	−19%
	$I_{d_{sat}}$	−14%	7%		$I_{d_{sat}}$	18%	16%
	f_T	−14%	8%		f_T	19%	15%

The statistical simulations were performed by changing the length and width of each transistor and each capacitance of the filter randomly. The physical dimensions of the transistor was changed by ±10% of its minimum feature size, e. g.

$$L = L_{nom} + Rand \cdot L_{min}, \tag{14}$$

where $Rand$ is a rectangular distributed from −0.1 to 0.1 and L_{nom} is the desired value. The filter capacitances were altered ±10% in the same manner. This can be considered rather conservative, as with good layout optimized for matching the standard deviation should not exceed a few %. A run of 500 simulations was performed at 10 GHz with input amplitude of {400, 500, 600} mV. The mean phase error, μ, was $\{-0.1, 0.1, 0.5\}$ degrees with an associated standard deviation, σ, of {1.91, 1.99, 1.90} degrees. In Fig. 23 the phase error distribution is plotted.

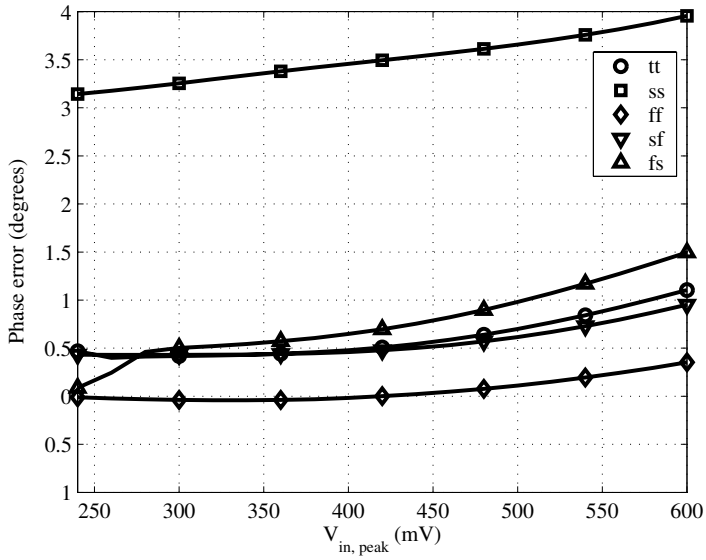


Figure 22: Corner simulation of the active polyphase filter at 10 GHz

A comparison to the passive polyphase filter counterpart was made by varying the capacitance and resistance by $\pm 10\%$. The mean phase error was 0 degrees with an associated standard deviation of 3.8 degrees.

5.2 Supply Voltage Variation

The supply voltage was changed from 0.8 V to 1.8 V for four different input amplitudes at 10 GHz (Fig. 24) to investigate the quadrature sensitivity. For small input amplitudes the filter needs almost 1.2 V supply[3] to achieve quadrature output. As the input amplitude increases also the supply variation can increase without penalty in quadrature error. With 400 mV input amplitude the supply can be down to 1 V.

5.3 Temperature Variation

A temperature sweep from -40 to 120 degrees Celsius was performed with four different input amplitudes at 10 GHz and 1.2 V supply, see Fig. 25. The temperature variation does not significantly affect the quadrature phase error. It changes by no more than 0.7° over the full temperature range.

[3]The datapoints where the filter was not producing quadrature output have been cut out of the figure.

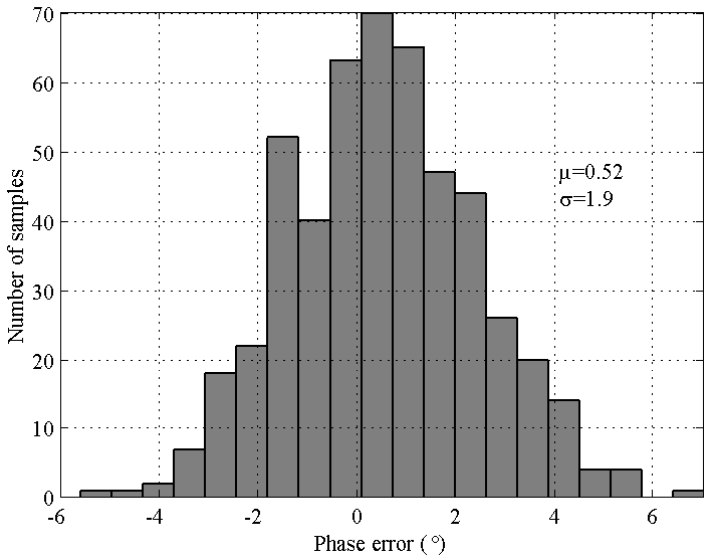


Figure 23: Statistical simulation of the active polyphase filter at 10 GHz. Input amplitude 600 mV and 500 runs. Uniform distribution used

6 Conclusion

A high frequency active polyphase filter has been investigated through model and transistor simulations. The filter uses transconductors instead of resistors as in a passive polyphase filter. When a transconductor is used the input impedance is larger and the filter can amplify the signal. In the simulations the transconductor was a CMOS inverter with a local bias feedback.

Simulations show a robust filter over wide frequency range (6–14 GHz) with an output quadrature error less than 1° for stable input amplitudes. The filter also shows robustness against input differential phase error and inverter delay. The potential of the technique has been demonstrated using the concept of effective capacitance. This indicates that high frequency filters can be realized with low power consumption. Corner simulations have been performed at 10 GHz, and for both n- and pMOS slow the phase error is below 4°, but for all other corners the error is less than 1.5°. The simulated far out phase noise is below that of the corresponding passive filter.

Acknowledgment

The authors would like to thank United Microelectronics Corporation (UMC) for giving us the opportunity to work with a state-of-the-art 130-nm CMOS process.

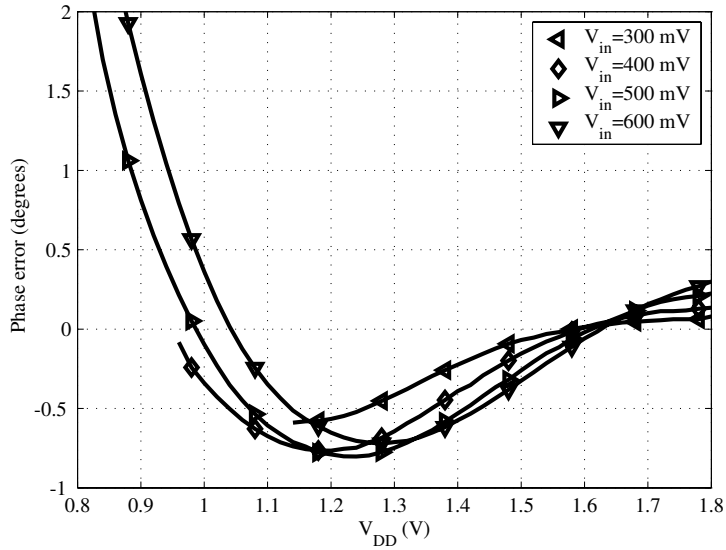


Figure 24: Supply voltage variation of the active polyphase filter at 10 GHz

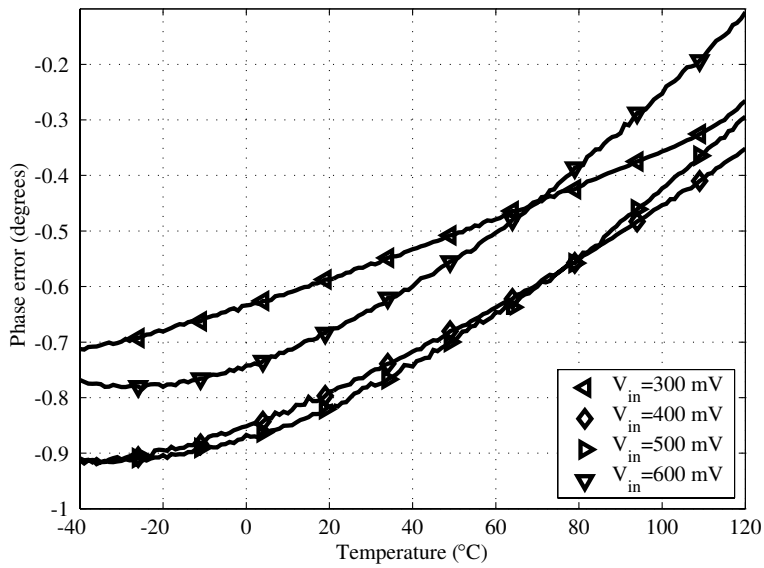
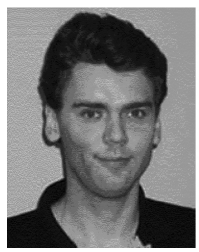


Figure 25: Temperature dependence of the phase error at 10 GHz

Biography

Johan Wernehag received the M.Sc. degree in Electrical Engineering from Lund University, Lund, Sweden in December 2002. He is currently a Ph.D. student at the Analogue/RF Group at Lund University, working towards the Ph.D. degree. His research interests include design of beam steering transmitters for WLAN/WPAN networks aiming at the frequency band round 60 GHz (IEEE P802.15), automotive radar, and quadrature generation. During the years 1998–2003 he was a part-time teaching assistant at the Dept. of Physics, Lund University and since 2003 at the Dept. of Electrical and Information Technology (former Electroscience), Lund University.



Henrik Sjöland received the M.S. degree in electrical engineering and the Ph.D. in applied electronics from Lund University, Lund, Sweden, in 1994 and 1997, respectively. He is a Professor in Analog Circuit Design at Lund University. His research interests include the design and analysis of analog integrated circuits, feedback amplifiers, and RF CMOS. In 1999, he spent one year as a Fulbright Postdoctoral with the Abidi group at UCLA. He is the author of Highly Linear Integrated Wideband Amplifiers (Boston, MA: Kluwer, 1999).

References

[1] F. Tillman and H. Sjöland, "A polyphase filter based on CMOS inverters," *Analog Integrated Circuits and Signal Processing*, vol. 50, no. 1, pp. 7–12, Jan. 2007.

[2] ——, "A Polyphase Filter based on CMOS Inverters," in *Proceedings of the Norchip Conference*, Nov. 2005, pp. 12–15.

[3] O. Casha, I. Grech, and J. Micallef, "Comparative study of gigahertz CMOS LC quadrature voltage-controlled oscillators with relevance to phase noise," *Analog Integrated Circuits and Signal Processing*, vol. 52, no. 1–2, pp. 1–14, Aug. 2007.

[4] A. W. L. Ng and H. C. Luong, "A 1-V 17-GHz 5-mW CMOS Quadrature VCO Based on Transformer Coupling," *IEEE Journal of Solid-State Circuits*, vol. 42, no. 9, pp. 1933–1941, Sep. 2007.

[5] W.-Z. Chen, C.-L. Kuo, and C.-C. Liu, "10 GHz Quadrature-Phase Voltage Controlled Oscillator and Prescaler," in *Proceedings of the European Solid-State Circuits Conference*, Sep. 2003, pp. 361–364.

[6] M. Törmänen and H. Sjöland, "A 26-GHz LC-QVCO in 0.13-um CMOS," in *Proceedings, Asian-Pacific Microwave Conference*, Dec. 2007, pp. 1769–1772.

[7] H. Q. Liu, W. L. Goh, and L. Siek, "A 0.18-μm 10-GHz CMOS Ring Oscillator for Optical Transceivers," in *Proceedings of the IEEE International Symposium on Circuits and Systems, ISCAS*, vol. 2, May 2005, pp. 1525–1528.

[8] Y. Eken and J. Uyemura, "The design of a 14 GHz I/Q ring oscillator in 0.18 μm CMOS," in *Proceedings of the IEEE International Symposium on Circuits and Systems, ISCAS*, May 2004, pp. 133–136.

[9] A. A. Abidi, "CMOS Mixers and Polyphase Filters for Large Image Rejection," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 8, pp. 1803–1816, Aug. 2006.

[10] R. Kodkani and L. Larson, "A 25 GHz quadrature voltage controlled ring oscillator in 0.12μm SiGe HBT," in *Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems*, Jan. 2006, pp. 383–386.

[11] P. Andreani and S. Mattisson, "On the Use of NautaŠs Transconductor in Low-Frequency CMOS g_m-C Bandpass Filters," *IEEE Journal of Solid-State Circuits*, vol. 37, no. 2, pp. 114–124, Feb. 2002.

[12] C.-Y. Chou and C.-Y. Wu, "The Design of Wideband and Low-Power CMOS Active Polyphase Filter and its Application in RF Double-Quadrature Receivers," *IEEE Transactions on Circuits and Systems—Part I: Regular Papers*, vol. 52, no. 5, pp. 825–833, May 2005.

[13] M. A. I. Elmala and S. H. K. Embabi, "Calibration of Phase and Gain Mismatches in Weaver Image-Reject Receiver," *IEEE Journal of Solid-State Circuits*, vol. 39, no. 2, pp. 283–289, Feb. 2004.

[14] E. Stikvoort, "Polyphase Filter Section With OPAMPs," *IEEE Transactions on Circuits and Systems—Part II: Express Briefs*, vol. 50, no. 6, pp. 376–378, June 2003.

[15] F. Behbahani, Y. Kishigami, J. Leete, and A. A. Abidi, "Phase Noise and Jitter in CMOS Ring Oscillators," *IEEE Journal of Solid-State Circuits*, vol. 36, no. 6, pp. 873–887, June 2001.

[16] H. Song, "A General Method to VLSI Polyphase Filter Analysis and Design for Integrated RF Applications," in *IEEE International SOC Conference*, Sep. 2006, pp. 31–34.

[17] B. Razavi, "Design Considerations for Direct-Conversion Receivers," *IEEE Transactions on Circuits and Systems—Part I: Regular Papers*, vol. 44, pp. 428–435, June 1997.

[18] ——, *RF Microelectronics*. Upper Saddle River, New Jersey, USA: Prentice Hall, 1998.

[19] X. Guan and A. Hajimiri, "A 24-GHz CMOS Front-End," *IEEE Journal of Solid-State Circuits*, vol. 39, no. 2, pp. 368–373, Feb. 2004.

[20] Y. P. Tsividis, *Operation and Modeling of the MOS Transistor*, 2nd ed. 198 Madison Avenue, New York: Oxford University Press, Inc, 1999.

Paper VI

Paper VI

A 30 GHz 90-nm CMOS Passive Subharmonic Mixer with 15 GHz Differential LO

Abstract

A new passive subharmonic mixer topology is presented and compared to a previously published passive topology. The comparison is conducted using simulations at 30 GHz with a 90-nm CMOS design kit.

The advantage of the new passive subharmonic mixer is that it only requires a differential local oscillator (LO) signal, compared to the previously published mixer that requires a quadrature LO signal. The mixer consists of two cascaded passive mixers with an interstage second order filter suppressing harmonics while providing some 10 dB of voltage gain at the LO frequency.

The noise performance of the differential mixer is slightly worse than for the quadrature one, with a simulated down-conversion SSB NF of 10 dB compared to 7 dB. The voltage conversion gain is -1 dB for both mixers, all with a 1 V LO amplitude.

Based on: Johan Wernehag and Henrik Sjöland, "A 30 GHz 90-nm CMOS Passive Subharmonic Mixer with 15 GHz Differential LO", *IEEE Proceedings of the 26th Norchip Conference*, November 17–18, 2008, Tallinn Estonia.

1 Introduction

A wide range of applications use frequency bands located at several tens of gigahertz, e. g. automotive radar (24 GHz and 77 GHz) and WLAN/WPAN (60 GHz). To make these applications penetrate the mass market the cost of the chip sets must be reduced, which is a drive and motivation to use CMOS technology also when implementing the analog and RF parts [1, 2].

The motivation to integrate automotive radar, also in the low cost segment cars, is high. The injuries from car collisions cost the society a lot both in medical bills and in human tragedies. Just in the United States (US) alone motor vehicle accidents accounted for 42,000 deaths, more than 5.3 million injuries, and over $231 billion in economic losses in year 2000 [3]. The consumer electronics, WLAN/WPAN applications, is extremely cost sensitive, and to succeed a low cost solution is a must.

A drawback when it comes to high speed CMOS processes is the low supply voltage. It reduces the available dynamic range of receivers and the achievable output power of the transmitters. To overcome these disadvantages a beamforming transceiver can be used [4, 5]. Combining a beamforming transceiver and phased array antenna the system will have an increased antenna directivity ($\propto$ the number of antennas) compared to a single antenna element [6]. The increased antenna directivity increases the strength of the signal to receive and at the same time reduces the level of interferers from other directions.

In Fig. 1 the proposed beamforming transmitter architecture is presented. The phase shifting to control the direction of the beam is performed in the local oscillator (LO) path, by means of a quadrature to differential vector modulator (VM) [7, 8], one modulator for each transmitter path. A subharmonic mixer, presented in this paper, upconverts the phase shifted signal to twice the LO frequency. A power amplifier (PA) driver and a frequency doubling PA then converts the signal to four times the LO frequency and transforms the differential to a single-ended signal [9]. The LO thus runs at a quarter of the carrier frequency, which increases the tuning range of the LO and the robustness to parasitics. Furthermore, the vector modulator could be simplified, as 360° at the the output corresponds to 90° at the LO, and thus it is sufficient to be able to steer the phase across one quadrant. The frequency (phase) modulation is inserted on the control voltage of the QVCO. The IF port of the subharmonic mixer is used to control the output power, and also to reduce the level of the side lobes.

Two of the three main building blocks have been presented earlier [7–9]. In this paper the third block, the subharmonic mixer (SHM), is addressed. A subharmonic passive mixer driven with quadrature LO was presented in [10]. It shows good performance but it is not suitable for a beamforming transmitter due to the quadrature LO needed. In a beamforming transmitter there are many transmit paths, and distributing the quadrature (phase shifted) LO across the chip to the mixers with sufficient signal quality is non-trivial and requires significant chip area and power consumption [11]. Distributing a differential signal is much easier. The demands on the vector modulators are also relaxed, since implementing vector modulators with differential output is

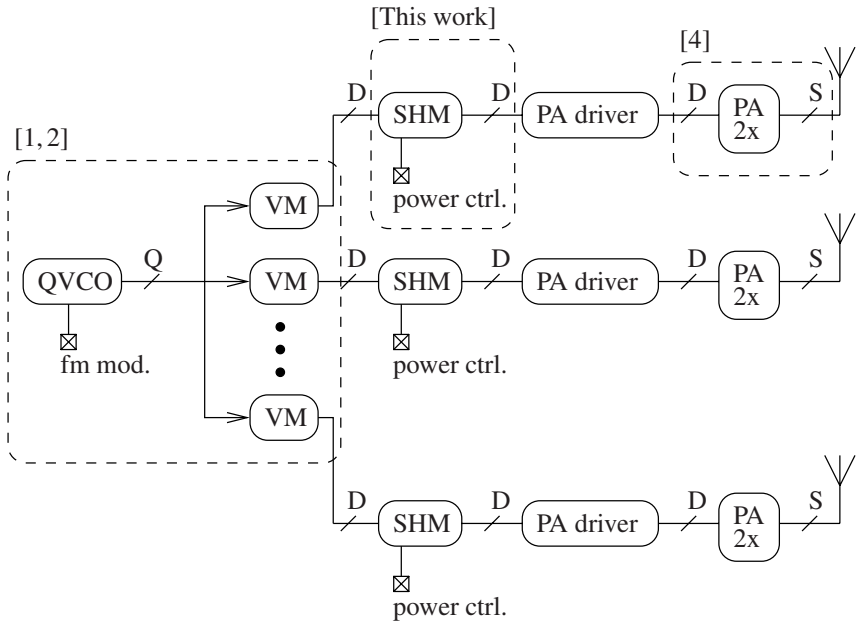


Figure 1: Proposed beamforming architecture with LO phase shifting. Q means quadrature signals, D differential, and S single-ended

less difficult than with quadrature output. An active SHM with differential LO was presented in [12]. The design uses two inductors, as does the proposed topology in Fig. 2(b). The hardware cost for the two mixers are about the same but the passive has less power consumption and is less complex. When used in a direct conversion receiver the passive mixer also has a $1/f$ noise advantage.

Therefore a comparison between the topology in [10], Fig. 2(a), and a novel topology presented in Fig. 2(b) is conducted.

2 Subharmonic Mixer Topologies

The comparison is performed using a 90-nm CMOS design kit with BSIM4.3 transistor models [13]. The non-quasi static model of the transistors is used, Table 1 shows the BSIM4 user switch settings. All simulations were performed with the Cadence SpectreRF simulator.

To determine the dimensions of the transistors and the LO bias voltage, parametric sweeps were performed and the voltage conversion gain (CG) and the single sideband noise figure (SSB NF) were plotted. The finger width was fixed to 2 μm and the number of fingers was changed to change the width of the transistor. The transistor

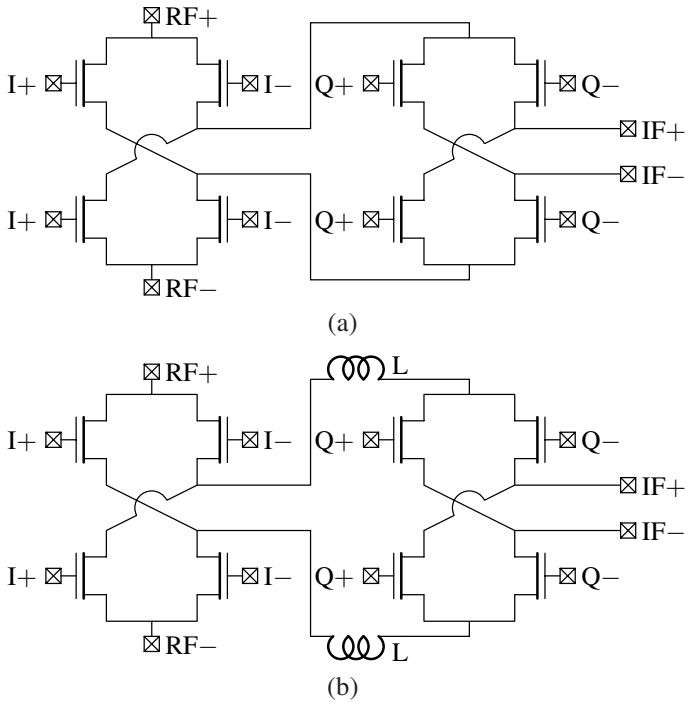


Figure 2: The two passive subharmonic mixer topologies. (a) Quadrature LO. (b) Differential LO

length was the minimum, 90 nm. To fit into the beamforming architecture presented in the introduction the LO frequency was 15 GHz, RF was 30 GHz, and the IF was zero. The LO was applied through a 5 pF DC-block capacitance while the bias was fed through a 10 kΩ resistor with an effective resistance of 3.8 kΩ at 15 GHz. The LO amplitude was held constant at 1 V peak, which can be achieved in a 1.2 V process. The higher the LO amplitude the better performance of the mixer. When simulating down-conversion a differential capacitive load of 500 fF was applied at the IF side, and a resistive load of 300 Ω was applied at the RF side in up conversion simulations.

2.1 Quadrature LO SHM

The dimensions of the transistors and the LO bias voltage were determined through a two-dimensional parametric sweep. In Fig. 3 the SSB NF and CG are plotted. The width was swept from 10 μm to 100 μm in steps of 10 μm. The NF decreases with increasing width, while the CG just drops slightly. Only the CG at 10 μm and 100 μm are plotted. At 100 μm the decrease in NF has flattened out, and the width of the

Table 1: BSIM4.3 Transistor Model Switches

acnqsmod = 1	trnqsmod = 1	rgatemod = 0	mobmod = 2
rdsmod = 1	igcmod = 1	igbmod = 1	capmod = 2
rbodymod = 1	diomod = 2	pemod = 1	geomod = 3
rgeomod = 1			

transistors was thus chosen to 100 μm. The CG is largest at zero LO bias while the NF reaches its minimum at 420 mV, thus a trade-off has to be made. Making the degradation from their optimum equal, a bias level of 210 mV was chosen, with a corresponding degradation of 0.6 dB. It can also be seen that the LO bias voltage and

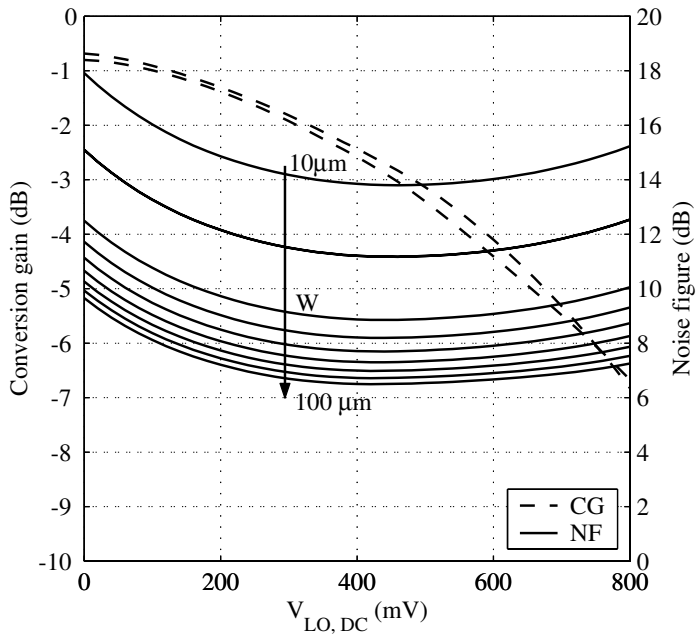


Figure 3: Simulation of down-conversion voltage gain and SSB noise figure of the quadrature LO SHM, at an IF of 10 MHz. The width is swept from 10 μm to 100 μm in steps of 10 μm

transistor size are orthogonal with respect to CG and NF.

With the transistor sizes and LO bias voltage now set, CG and NF versus frequency, and linearity were investigated. The 1-dB compression point referred to the RF side is 0.8 dBm, while CG and SSB NF remain constant at -1.4 dB and 7.0 dB,

respectively, up to 1 GHz IF frequency the degradation is less than 1 dB. The up-conversion comparison between the two topologies with respect to noise, gain, and compression point is presented in Section 2.3.

An LO signal with perfect quadrature is difficult to achieve and route across the chip. The sensitivity against quadrature phase error, θ, was therefore simulated, see Fig. 4. The quadrature phase error stretches from 0 to 360 degrees and one can see that the order of Q+ and Q− is arbitrary (works well at 180° error), which can be valuable if the locking order of the QVCO is difficult to assure. It is also clear that the CG deteriorates as θ approaches 90 and 270 degrees, thus a modification of the mixer is required to work with differential LO signals.

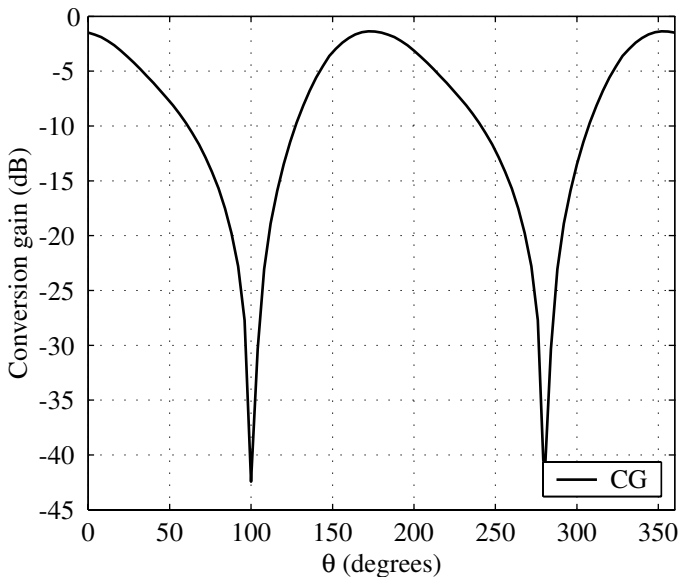


Figure 4: Conversion gain vs. quadrature phase error

2.2 Differential LO SHM

In the previous section it was observed that the effects of transistor width and LO bias voltage on CG and NF were orthogonal. Thus the transistor width and inductance are chosen first through a two dimensional parametric simulation. The LO bias voltage is then chosen through a separate simulation. The inductance and transistor width are connected together in the interstage filter function and thus have to be changed simultaneously. The inductors are spiral inductors supplied by the foundry, swept from 200 pH to 1 nH. Their Q-value at 15 GHz is above 18 for all sizes. Two-dimensional

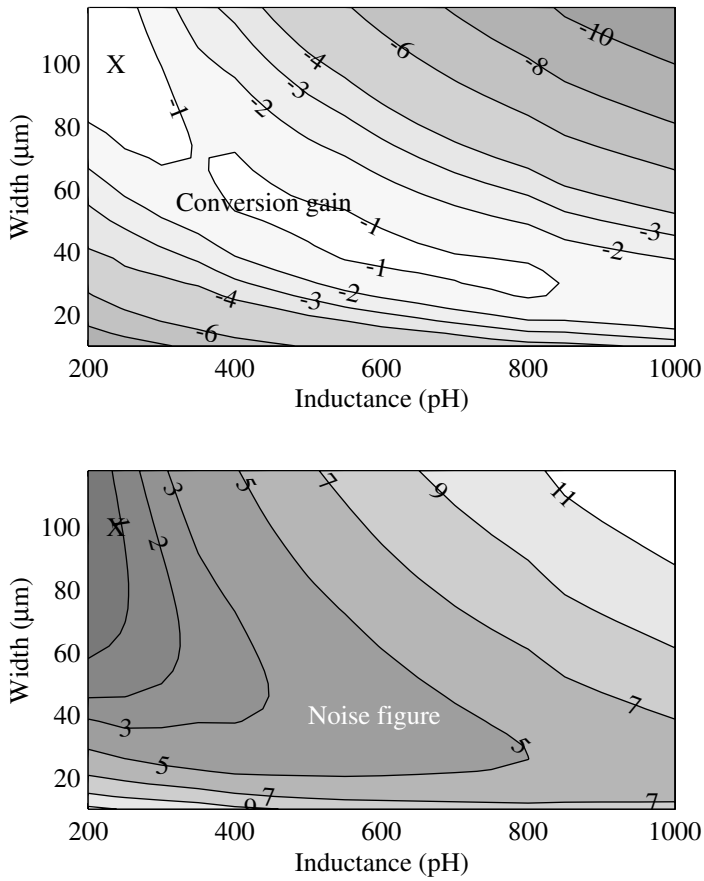


Figure 5: Simulation of the differentially driven SHM as the transistor width and the inductance are swept, at 10 MHz IF. Top: Relative voltage conversion gain. Bottom: Relative single sideband NF

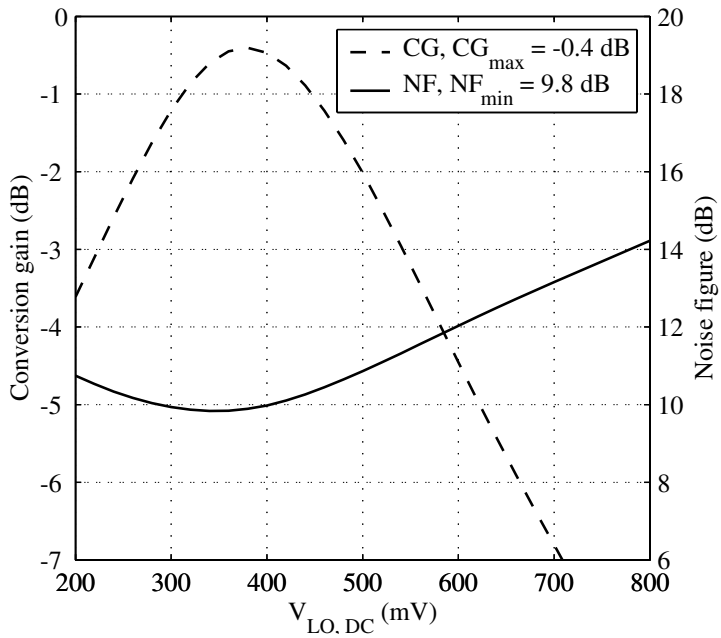


Figure 6: Down conversion CG and SSB NF for the differentially driven SHM versus LO bias voltage, IF is 10 MHz

contour plots of the SSB NF and CG are presented in Fig. 5. The plots are normalized relative to the optimum value since the LO bias was not yet optimized. The conversion gain is best for a seemingly constant LC-product, corresponding to a peak in the filter function at the LO frequency, see Fig. 7. The SSB NF is also in this case the lowest for large transistors. A transistor width of 100 μm and an inductance of 225 pH were chosen, indicated by the cross (X) in Fig. 5.

The LO bias level was swept and then set to 370 mV, which gives a CG and SSB NF of -0.43 dB and 9.9 dB, respectively, see Fig. 6. For the bias point selected the CG and NF deviates less than 0.5 dB from their optimum values. The noise figure of this mixer is 2.9 dB worse than the quadrature LO one.

The compression point at the RF side is -6.2 dBm. The CG and SSB NF are rather constant up to 1 GHz IF frequency, deviates less than 1 dB.

2.3 Up-Conversion Comparison

The up-conversion performance for both topologies is evaluated with respect of noise, gain, and 1-dB compression point (Fig. 8). The load at the RF side is 300 Ω, emulating the load of a tuned 30 GHz PA driver following the SHM. The SSB noise figure

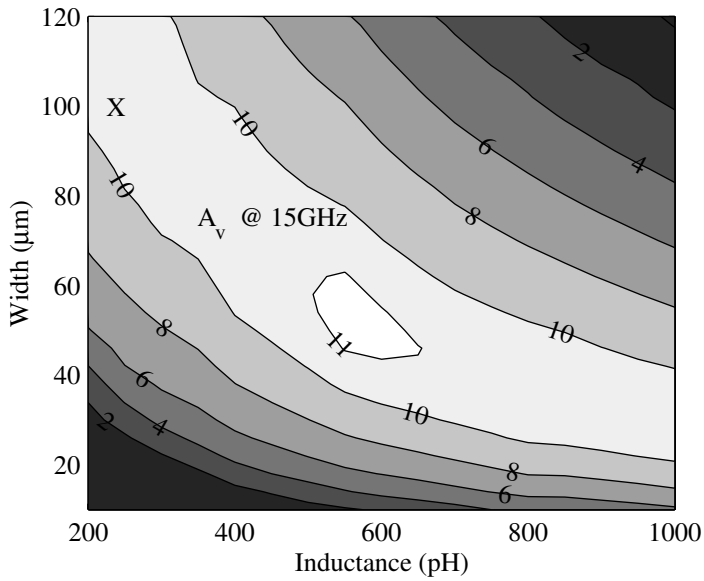


Figure 7: Voltage gain for the 15 GHz tone of the interstage second order filter

is 4 dB higher for up-conversion than down-conversion, 11 dB and 14 dB for the quadrature and differential LO SHM respectively. The 3-dB difference between the two topologies remain.

The voltage conversion gain, simulated with one RF tone and a DC IF input voltage, is 0.14 dB and -1.59 dB for the quadrature and differential LO SHM respectively. In Fig. 8 output power vs. input DC voltage is plotted and the 1-dB compression points are extracted, -2.7 dBm and 0.17 dBm for the quadrature and differential LO corresponding to an output voltage of 220 mV and 320 mV respectively over a 300 Ω load.

2.4 Summary

The performance and design parameters of the two SHM are summarized in Table 2.

3 Conclusion

A comparison between two passive SHM has been performed, one with quadrature LO and one with differential LO. The new mixer topology with differential LO uses a second order interstage filter to suppress high order harmonics and provide some 10 dB of voltage gain for the desired signal at the LO frequency. In a beamforming

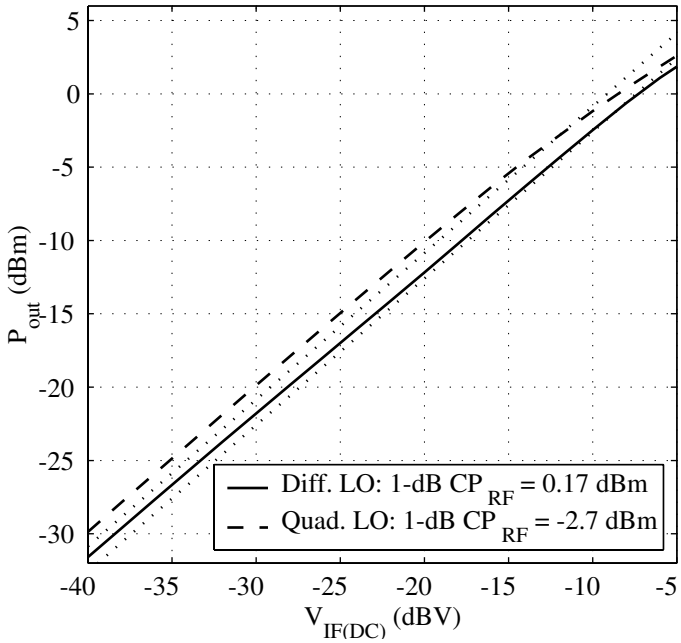


Figure 8: Up-conversion output power vs. DC IF input voltage

transceiver, with multiple receive and transmit paths, the LO generation (with or without phase shift) is performed at one place on the chip. The LO therefore needs to be distributed to all the mixers in the transceiver (Fig. 1), which is much easier with a differential than with a quadrature LO signal.

The differential subharmonic mixer, however, has some penalties in noise figure and chip area, but in a beamforming application the advantage of having of having differential LO signals is so large that these penalties in most cases can be accepted.

4 Acknowledgment

The authors would like to thank United Microelectronics Corporation (UMC) for giving us the opportunity to work with a state-of-the-art 90-nm CMOS process and the VINNOVA industrial excellence center System Design on Silicon for funding the research.

Table 2: Performance and Design Parameter Summary

	Down-Conversion		Up-Conversion	
	Quad. LO	Diff. LO	Quad. LO	Diff. LO
SSB NF (dB)	7	10	11	14
CG (dB)	-1.4	-0.4	0.14	-1.6
RFCP_1 (dBm)	0.8	-6.2	-2.7	0.2
$\mathrm{V}_{\mathrm{LO,\,DC}}$ (mV)	210	370	150	270
$\mathrm{V}_{\mathrm{LO,\,amp}}$ (V)	1	1	1	1
Width (μm)	100	100	100	100
Inductance (pH)	—	225	—	225

References

[1] B. Razavi, "Gadgets Gab at 60 GHz," *IEEE Spectrum*, vol. 45, no. 2, pp. 40–45, Feb. 2008.

[2] S. K. Moore, "Cheap Chips for Next Wireless Frontier," *IEEE Spectrum*, vol. 43, no. 6, pp. 8–9, June 2006.

[3] L. Blincoe, A. Seay, E. Zaloshnja, T. Miller, E. Romano, S. Luchter, and R. Spicer, "The Economic Impact of Motor Vehicle Crashes, 2000," *National Highway Traffic safety Administration*, May 2002, report NO. DOT HS 809 446.

[4] A. Natarajan, A. Komijani, and A. Hajimiri, "A Fully Integrated 24-GHz Phased-Array Transmitter in CMOS," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 12, pp. 2502–2514, Dec. 2005.

[5] A. Hajimiri, H. Hashemi, A. Natarajan, X. Guan, and A. Komijani, "Integrated Phased Array Systems in Silicon," *Proceedings of the IEEE*, vol. 93, no. 9, pp. 1637–1655, Sep. 2005.

[6] R. J. Mailloux, *Phased Array Antenna Handbook*, 2nd ed. Norwood, MA: Artech House Books, 2005.

[7] J. Wernehag and H. Sjöland, "An 8-GHz Beamforming Transmitter IC in 130-nm CMOS," in *IEEE Radio Frequency Integrated Circuits Symposium*, July 2007, pp. 577–580.

[8] ——, "A 24-GHz Automotive Radar Transmitter with Digital Beam Steering in 130-nm CMOS," in *Ph. D. Research in Microelectronics and Electronics*, June 2006, pp. 481–484.

[9] ——, "60 GHz 130-nm CMOS Second Harmonic Power Amplifiers," in *IEEE Asia Pacific Conference on Circuits and Systems (APCCAS)*, Dec. 2008.

[10] R. M. Kodkani and L. E. Larson, "A 24-GHz CMOS Passive Subharmonic Mixer/Downconverter for Zero-IF Applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 56, no. 5, pp. 1247–1256, May 2008.

[11] E. Mensink, D. Schinkel, E. A. M. Klumperink, E. van Tuijl, and B. Nauta, "Optimal Positions of Twists in Global On-Chip Differential Interconnects," *IEEE Transactions on very Large Scale Integration (VLSI) Systems*, vol. 15, no. 4, pp. 438–446, Apr. 2007.

[12] A. Parsa and B. Razavi, "A 60GHz CMOS Receiver Using a 30GHz LO," in *IEEE International Solid-State Circuits Conference, Digest of Technical Papers*, Feb. 2008, pp. 190–606.

[13] BSIM Research Group, “BSIM4.3.0 MOSFET Model,” http://www-device.eecs.berkeley.edu/∼bsim3/bsim4.html.