CMOS Transceiver Front-Ends in Mobile Communication Handsets - Architectures and Building Blocks

Cijvat, Pieternella

2004

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

CMOS Transceiver Front-Ends in Mobile Communication Handsets

Architectures and Building Blocks

Ellie Cijvat

Department of Electroscience
Lund University
Sweden 2004

CMOS Transceiver Front-Ends in Mobile Communication Handsets

Architectures and Building Blocks

Ellie Cijvat

LUND
UNIVERSITY

Department of Electroscience

Lund, 2004

Department of Electroscience
Lund University
P.O. Box 118
221 00 Lund, Sweden

Printed in Sweden by *Tryckeriet i E-huset*, Lund
November 2004.

ISSN 1402-8662, No. 49.

ABSTRACT

For mobile communication systems in the low-GHz range, CMOS has increasingly become the technology of choice, and the level of integration in mobile handsets has risen. The use of off-chip components, which increases the handset assembly time and costs, is preferably avoided. However, integrating a complete transceiver on a single chip leads to disturbances between building blocks. This imposes new and more stringent requirements on building block and transceiver performance, as well as impacts the choice of transceiver architecture.

In the general introduction, an overview is given of front-end receiver and transmitter aspects as well as RF CMOS technology. This includes the impact of mobile communication system specifications on architectures and building blocks, transistor and monolithic inductor modeling, and disturbance issues. Special attention is given to power amplifiers, the most challenging building blocks in CMOS transceivers. Papers I, II and III address CMOS receiver front-end aspects and implementations, while in papers IV and V design and challenges of CMOS power amplifiers are described.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor Henrik Sjöland. Without him I would not have made it this far, and this thesis would not be in the shape it is now.

I thank my colleagues at the Dept. of Electroscience who made it a pleasant working environment. A special 'thank you' goes to the administative and technical staff at the department, for help with paper work, computers, programs and printed circuit boards. Anders Karlsson, Lars Olsson, Niklas Troedsson (Dept. of Electroscience) and Costantino Pala (Conexant) deserve credit for reading (part of) my thesis and giving useful feedback.

Furthermore I wish to thank some persons who have helped and supported me in the past: Dr. Mehran Mokhtari, now at Hughes Research Laboratory (HRL), for his supervision in the early days, and Prof. Asad A. Abidi, at the University of California, Los Angeles (UCLA), for giving me the opportunity to visit his research group. The list of people who helped and influenced me could be much larger, but I will stop here.

I would like to thank my friends for being there for me and showing me different perspectives on anything from mobile phones to world peace. And last but not least, I sincerely thank my parents and my sister for supporting me all these years.

Ellie Cijvat

Lund, December 2004

PREFACE

This thesis describes the work I have done in the field of RF CMOS design. The research was performed at several places: The Royal Institute of Technology, Stockholm, the University of California, Los Angeles and the Department of Electroscience at Lund University.

Both transmitter and receiver front-end architectures and building blocks are described. In part I, the general introduction, a framework is given of architectural considerations, mobile communication specifications and RF CMOS integrated circuit basics. These aspects were all used when performing the research that led to the papers presented in part II. Some topics are covered more extensively in the general introduction, such as the disturbance in a receiver caused by spurious signals, and power amplifiers.

Included Papers

I. E. Cijvat, "A 0.35 μm CMOS DCS Front-end with Fully Integrated VCO", in *Proceedings of the 8th IEEE International Conference on Electronics, Circuits and Systems (ICECS 2001)*, Malta, 2001, pp. 1595-1598.

II. S. Tadjpour, E. Cijvat, E. Hegazi and A. Abidi, "A 900 MHz Dual Conversion, Low-IF GSM Receiver in 0.35 μm CMOS", *IEEE Journal of Solid-State Circuits,* Vol. 36, No.12, pp. 1992 - 2002, Dec. 2001.

III. E. Cijvat, S. Tadjpour and A.A. Abidi, "Spurious mixing of off-channel signals in a wireless receiver and the choice of IF", *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, Vol. 49, No. 8, pp. 539 - 544, Aug. 2002.

IV. E. Cijvat and H. Sjöland, "A Fully Integrated 2.45 GHz 0.25 μm CMOS Power Amplifier", in *Proceedings of the 10th IEEE International Conference on Electronics, Circuits and Systems (ICECS 2003)*, United Arab Emirates, 2003, pp. 1094-1097.

V. E. Cijvat, N. Troedsson and H. Sjöland, "A Fully Integrated CMOS RF Power Amplifier with Internal Frequency Doubling", submitted to *Analog Integrated Circuits and Signal Processing.*

Related Publications

The following papers are not included but contain both overlapping and complementing material related to this work.

Licentiate of Engineering Thesis:

E. Cijvat, *A study of CMOS receiver architectures in Mobile Communication Handsets for GSM*. Tech. Lic. Thesis, Royal Institute of Technology, Dept. of Electronics, May 2000.

Conference Papers:

S. Tadjpour, E. Cijvat, E. Hegazi and A. Abidi, “A 900 MHz Dual Conversion, Low-IF GSM Receiver in 0.35 mm CMOS”, in *Technical Digest of the IEEE 2001 International Solid-State Circuits Conference (ISSCC)*, USA, 2001, pp. 292-293.

E. Cijvat, P. Eriksson, N. Tan and H. Tenhunen, “A 1.8 GHz Subsampling CMOS Downconversion Circuit for Integrated Radio Applications”, in *Proc. of the 5th IEEE International Conference on Electronics, Circuits and Systems (ICECS 1998)*, Portugal, 1998, pp. 149-152.

E. Cijvat, P. Eriksson, N. Tan and H. Tenhunen “A 1.8 GHz Subsampling CMOS Downconversion Circuit for Integrated Radio Circuits”, in *Proc. of the IEEE 1998 International Symposium on Circuits and Systems (ISCAS’98)*, USA, 1998, pp. II-65-68.

E. Cijvat, P. Eriksson, N. Tan and H. Tenhunen “A 1.8 GHz Subsampling CMOS Downconversion Circuit with Integrated A/D Converter”, in *Proc. of the 1997 15th Norchip Conference*, Estonia, 1997, pp. 352-359.

N. Tan, E. Cijvat, and H. Tenhunen, “Design and Implementation of High-Performance CMOS D/A Converter”, in *Proc. of the IEEE 1997 International Symposium on Circuits and Systems (ISCAS’97)*, Hong Kong, 1997, pp. 421-424.

CONTENTS

LIST OF ABBREVIATIONS

3G	third generation
ACLR	adjacent channel leakage power ratio
AD, A/D, ADC	analog-to-digital, analog-to-digital converter
AM-to-PM	amplitude modulation to phase modulation
BB	baseband
BER	bit error rate
BJT	bipolar junction transistor
BPF	band pass filter
BW	bandwidth
CALLUM	combined analog locked loop universal modulator
CDMA	code division multiple access
CMOS	complementary metal-oxide-semiconductor
C/(N+I)	carrier-to-noise and interference ratio
CS	common source
D/A, DAC	digital-to-analog, digital-to-analog converter
DC	direct current
DCS	digital cellular system
DDD	double doped drain
DIBL	drain induced barrier lowering
DSP	digital signal processing
EER	envelope elimination and restoration
F	noise factor
FDD	frequency division duplexing
FDMA	frequency division multiple access
FM	frequency modulation
FS	frequency synthesizer
FSK	frequency shift keying
GaAs	gallium arsenide
GMSK	Gaussian minimum shift keying
GSM	global system for mobile communications
HSDPA	high speed downlink packet access
I and Q	in-phase and quadrature phase
IC	integrated circuit
IF	intermediate frequency
IP_i	i-th order intercept point

IRF	image reject filter
LDD	lightly doped drain
LINC	linear amplification with non-linear components
LNA	low noise amplifier
LO	local oscillator
LPF	low pass filter
MN	matching network
MOSFET	metal-oxide-semiconductor field effect transistor
NF	noise figure
PA	power amplifier
PAE	power added efficiency
PGA	programmable gain amplifier
PWM	pulse width modulation
RF	radio frequency
RX	receiver
SH, S/H	sample and hold
Si	silicon
SNR	signal-to-noise ratio
SoC	system-on-chip
TDD	time division duplexing
TDMA	time division multiple access
TX	transmitter
UMTS	universal mobile telecommunication system
UWB	ultra wide band
VCO	voltage controlled oscillator
VLSI	very large scale integration
WiMax	worldwide interoperability for microwave access
η	efficiency

Part I

General Introduction

CHAPTER 1

INTRODUCTION

1.1. MOTIVATION

The maturing of communication systems such as GSM, Bluetooth and UMTS, operating in the low-GHz range, together with technological advancements for CMOS technologies, has resulted in increased research activities in so-called Radio Frequency CMOS circuits, i.e. analog circuits in a CMOS technology for frequencies up to several GHz. CMOS circuits for frequencies as high as 60 GHz have been reported. Most important for this development is the drive for highly-integrated, low-cost mobile handsets. If both the analog and digital part of a receiver/transmitter (transceiver) can be implemented in CMOS, a complete system may be implemented on a single chip. This System-on-Chip (SoC) development brings about issues of disturbance between the different parts or building blocks of the system, caused by coupling, either on-chip, through the substrate, or off-chip, through adjacent bondwires. These subjects will be briefly touched upon in this thesis.

One of the most challenging analog parts to implement in CMOS is the power amplifier (PA), for several reasons. Firstly, the transconductance to current ratio g_m/I of a CMOS device is generally lower than that of a bipolar or III-V device, implying that for the same gain a higher current is needed. Secondly, with decreasing device length (scaling) in current CMOS technologies the oxide thickness is decreasing as well, resulting in a lower breakdown voltage. On the other hand, the cut-off frequency f_T and the maximum frequency of oscillation f_{max} for CMOS devices are approaching the values for Si bipolar devices. In the general introduction of this thesis, special attention is given to PA properties such as linearity and efficiency, as well as to CMOS device modeling and scaling.

Another issue stemming from an increased level of integration and a drive for low-cost handsets is off-chip filters. These filters are costly, they take up considerable space and they complicate handset assembly. Therefore an important goal of integrated transceiver design is the reduction of the number of these filters. This

work addresses receiver and transmitter analog front-end architectures, and an analysis is presented of the effects of removing these filters from the transceiver.

1.2. MOBILE COMMUNICATION SYSTEMS

Before addressing transmitter and receiver properties in detail, some basic concepts in mobile communication systems will be described. A mobile communication system allows a user to communicate by means of a wireless link to a base station (see Fig. 1.1). Traditionally this was based on speech transmission, where relatively low data rates can be used while still maintaining a reasonable sound quality. Nowadays higher data rates are pursued, so that besides speech transmission high data rate transmission becomes a viable option as well. Third-generation (3G) wireless systems, capable of higher data rates, are now operated in many countries, and improved data rates of 3G (e.g., HSDPA [5]) are soon to be deployed. Other systems aiming at high data rates are for instance WiMax [3] and Ultra Wide Band (UWB) based systems [4].

As depicted in Fig. 1.1, a mobile communication system can consist of several cells with one or more base stations serving a number of users. The uplink is defined as the communication from the user to the base station, while the downlink entails the communication from the base station to the user. From a user's point of view the uplink is equivalent to transmission while the downlink is equivalent to reception [6].

For most modern communication systems different frequency bands are used for the uplink and downlink. This permits transmission and reception simultaneously, and is called *frequency division duplexing* (FDD) [6]. A disadvantage of FDD compared to *time division duplexing* (TDD), is that users cannot communicate with each other directly but have to communicate via the base station. A second disadvantage is that when both transmit and receive modes are on at the same time, the transmitted

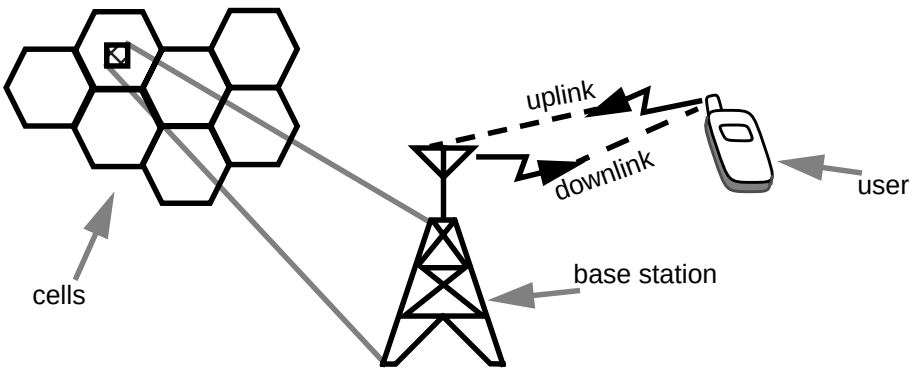


Fig. 1.1. A mobile communication system.

signal will leak to the receive path since the duplex filter located between the transceiver and the antenna has a finite transmitter-to-receiver isolation.

System Specifications

All mobile communication standards have more or less similar specifications, such as required *bit error rate* (BER), minimum detectable signal (sensitivity), blocking and interference performance, channel bandwidth, modulation scheme, output power range, frequency bands, et cetera. For each system the value of these requirements may differ. The maximum output power from the antenna for a mobile handset is for instance 30 dBm for GSM or DCS class 1 [7], 0 dBm for Bluetooth class 2 [8] and 24 dBm for UMTS (WCDMA) class 3 [9]. System level simulations must be performed to translate system specifications to building block specifications.

Multiple access methods

In order to accommodate several users simultaneously, different so-called multiple access schemes are in use. GSM and DCS use time division multiple access (TDMA), i.e. consecutive time slots are assigned to different users [6]. In this way eight users may use the same frequency band. Bluetooth uses a special kind of frequency division multiple access (FDMA), namely frequency hopping, i.e. each user is assigned a certain frequency band, but the assigned frequency band changes with the hopping rate. Finally, UMTS is a code division multiple access system (CDMA) where a pseudorandom code is used to spread the data over a wider frequency range before transmission. Many users may then send in the same frequency band simultaneously; at reception, each user's data can be extracted by using the code assigned to that user.

Transceiver architecture

In Fig. 1.2 a typical handset transmitter/receiver (or transceiver) architecture is shown. In the receive (RX) path the signal is generally amplified and downconverted from the *radio frequency* (RF) to the baseband frequency in one or several steps. The baseband signal is then digitized in the *analog to digital converter* (ADC) and demodulated in the *digital signal processing* (DSP) part. On a system level a decision must be taken regarding the amount of signal processing to be performed in the analog domain. A range of solutions is being or has been investigated, from directly sampling the RF signal with an AD converter, to full channel selection in the analog domain. This is a trade-off between complexity, performance and power consumption in the analog and digital part, respectively.

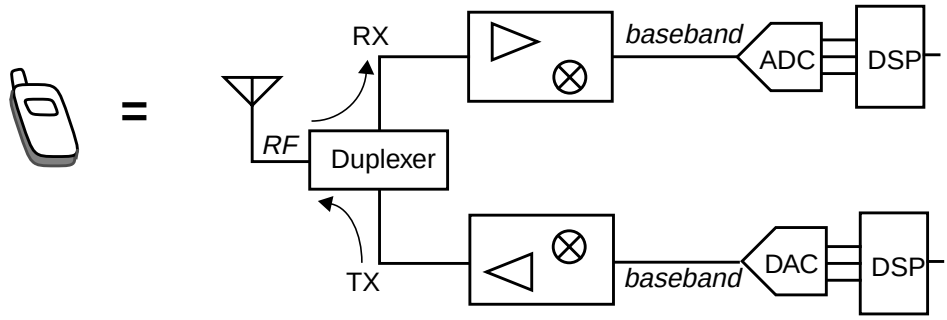


Fig. 1.2. A general transceiver architecture.

In the transmit (TX) part a baseband signal is modulated in the DSP, converted to an analog signal in the *digital to analog converter* (DAC) and then filtered and upconverted to RF in one or several steps. A power amplifier is used to give the RF signal the desired power. Similar trade-offs exist in the transmitter part as in the receiver part. In this thesis the focus will be on the front-end, which may roughly be defined as the part between the antenna and the baseband. Hereafter ’receiver front-end’ is meant when ’receiver’ is used, and similarly for ’transmitter’.

1.3. STRUCTURE OF THE THESIS

The thesis is divided in two parts: a general introduction in part I and a selection of relevant papers in part II.

Part I: General Introduction

In part I, basic aspects of communication systems are revised, divided in receivers, transmitters, and CMOS technology. In chapter 2 general receiver aspects are described, such as architectures, noise and linearity, system specifications and building block requirements, and frequency planning. Chapter 3 describes transmitters, including architectures, system specifications and building block requirements, frequency planning and power amplifiers including considerations on efficiency and linearity. Also, disturbance issues in highly integrated RF CMOS transceivers are addressed. In chapter 4 CMOS technology aspects are investigated, extended to RF operation, deep-submicron scaling issues, and noise. Inductors are also described, since they are key elements in several RF building blocks. Both MOSFET and inductor modeling are discussed briefly. Conclusions are presented in Chapter 5.

Much of the contents in chapter 2 and 4 has already been presented in [1]. However, the author has chosen to incorporate these parts since they complement the other chapters and cover many issues that come into play in integrated transceiver design.

Part II: Included Papers

The papers included in part II cover many of the aspects described in part I. Papers I, II and III focus on receiver front-ends. Paper I describes the design of a DCS receiver with a *low-noise amplifier* (LNA), a mixer and a *voltage controlled oscillator* (VCO) integrated on the same chip. Paper II describes the design, implementation and measurements of a low-IF GSM receiver front-end. In paper III an analysis is presented of spurious signals in a receiver and the effects on the choice of intermediate frequency (IF).

Papers IV and V both describe issues related to fully integrated CMOS transmitters. In paper IV the design of a class AB power amplifier with power control is presented, and in paper V the design and measurements of a class C power amplifier with internal frequency doubling and on-chip VCO is described.

CHAPTER 2

RECEIVERS FOR MOBILE HANDSETS

2.1. INTRODUCTION

In this chapter an overview is given of the most common receiver architectures, such as the direct-conversion receiver and the heterodyne receiver. Moreover, some less common receivers are described, such as image-reject receivers and subsampling receivers. Thereafter system specifications and parameters such as noise figure and intercept point are presented, as well as frequency planning in a receiver to avoid the impact of spurious signals.

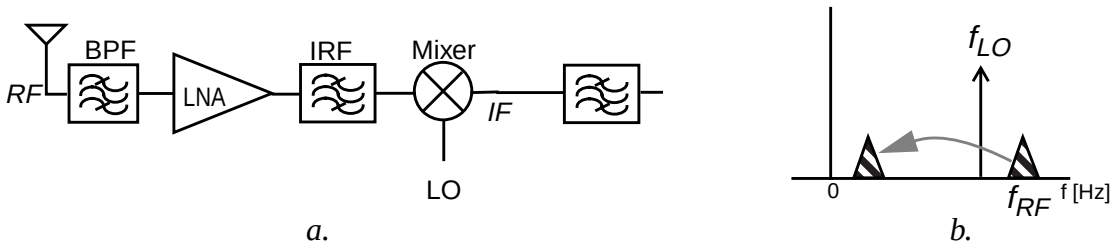
2.2. RECEIVER ARCHITECTURES

As shown in Fig. 1.2, the function of the receiver front-end is to convert the incoming RF signal to a lower frequency so that the AD converter (ADC) may digitize the signal. The front-end must ensure a sufficient Signal-to-Noise ratio (SNR). The receiver therefore typically consists of a Low-Noise Amplifier (LNA) to amplify the incoming signal without adding too much noise, one or more mixers to convert the RF signal to a lower frequency, and one or more filters to select the desired signal or frequency band and to block disturbing signals. Several methods of downconversion will be described here, such as direct conversion, heterodyne, and subsampling. Moreover, image reject receivers will be discussed briefly.

2.2.1. HETERODYNE RECEIVERS

One of the most common architectures is the heterodyne receiver [1] (see Fig. 2.1). The radio frequency (RF) signal passes through a *band pass filter* (BPF) which is

1. Historically the heterodyne receiver was a single downconversion receiver for which ω_{IF} was relatively low (in the audible frequency region). The superheterodyne receiver as patented by Armstrong had a higher $\omega_{IF,}$ but still used a single downconversion. Nowadays, generally no distinction is made between the ‘superheterodyne’ and ‘heterodyne’ receiver. This convention will be followed in this thesis as well.



a. *b.*

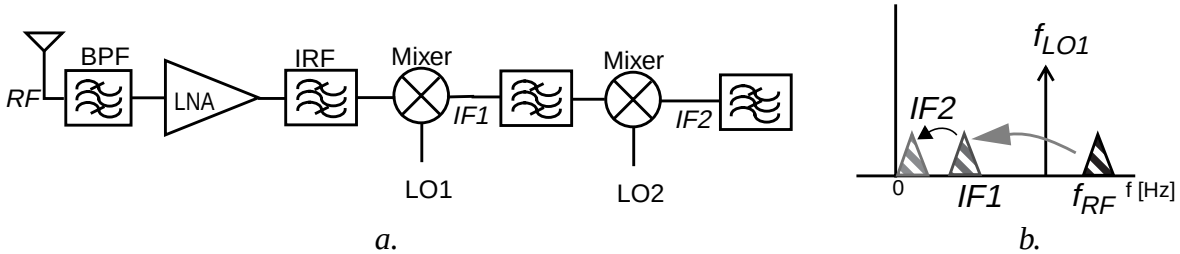
Fig. 2.1. A heterodyne receiver architecture, a). architecture, b). downconversion in the frequency domain.

used to suppress some of the undesired signals outside the receive band. Then the RF signal is amplified by a *low noise amplifier* (LNA) and sent through a second band pass filter, which is sometimes called *image reject filter* (IRF). The two filters together can provide an attenuation in the order of 100 dB for signals lying several hundreds of MHz away from the passband. In many - though not all - mobile communication receivers these two filters are used. Due to the high frequency and the steepness required, these filters typically are off-chip elements [10], [6], [11], [12], [13], [14].

The RF signal is downconverted in the mixer, which is also provided with a *local oscillator* (LO) signal. The *intermediate frequency,* or IF, is given by $\omega_{IF} = \omega_{RF} - \omega_{LO}$ [rad/s] where $\omega = 2\pi f$. For the case of $|\omega_{IF}| > 0$ - but ω_{IF} low enough to be considered a baseband signal - the term ‘low-IF downconversion’ may be used [15].

Often f_{LO} and f_{RF} are so far apart that the signal is not downconverted to baseband but to a relatively high IF. Therefore it is common to have a second downconversion stage, resulting in a double-conversion heterodyne architecture (see Fig. 2.2).

If $\omega_{RF} < \omega_{LO}$ then the IF signal will be at negative frequencies. If no special measures such as complex mixing are taken, the output signal will be indistinguishable from one having the same IF at positive frequencies.



a. *b.*

Fig. 2.2. A double-conversion heterodyne receiver architecture, a). architecture, b). downconversion in the frequency domain.

An advantage of double-conversion heterodyne receivers is that the total gain in the receive path can be distributed over several blocks, possibly accompanied by filters. The issue of gain distribution is related to receiver noise figure versus linearity, and will be covered more extensively in paragraph 2.3.1. Having the gain distributed over different frequencies also reduces the risk for self-oscillation.

The main disadvantage of the double-conversion heterodyne receiver is its complexity, including the fact that two LO signals are necessary, and the need for several filters which generally cannot be integrated on-chip.

Image Frequency

One of the main problems of the heterodyne receiver is the presence of an undesired signal at the so-called image frequency. This is explained first before methods to address this problem are discussed.

A downconversion block converts the desired signal at ω_{RF} to $\omega_{IF} = \omega_{RF} - \omega_{LO}$. For convenience it is assumed that $\omega_{RF} > \omega_{LO}$. Undesired signals that reach the downconverter input are downconverted as well. A signal situated at $\omega_{RF} - 2\omega_{IF}$ will be downconverted to $-\omega_{IF}$, as shown in Fig. 2.3, and can thus distort the desired signal at $+\omega_{IF}$ if positive and negative frequencies are not separated.

Two possible solutions for this problem exist: Firstly, the undesired signals can be prevented from reaching the downconverter input by means of filtering, and secondly, the positive and negative frequencies can be separated after downconversion by so-called image-reject mixing.

The former is highly dependent on filter characteristics, e.g. selectivity, in combination with the allowed blocking signal strengths in a communication system, while the latter is based on accurate complex mixing to separate the negative and positive frequency component. In paragraph 2.2.3 some image reject receiver architectures are described.

In order to get substantial image suppression from filtering, two conditions must be fulfilled (see Fig. 2.4):

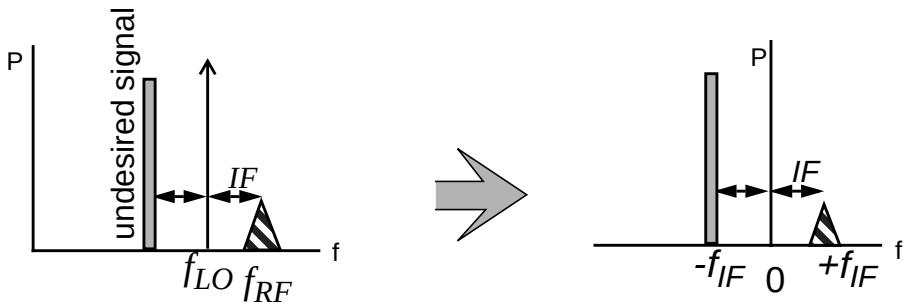


Fig. 2.3. The problem of the image frequency.

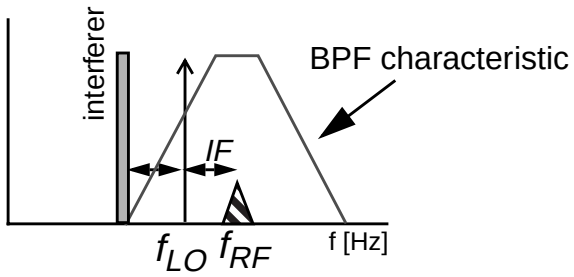


Fig. 2.4. A band pass filter utilized to provide image rejection.

- ω_{IF} must be relatively large, so that the distance between the desired signal and the image located at $\omega_{RF} - 2\omega_{IF}$ is substantial,
- the attenuation characteristic of the filter must be relatively sharp.

The filter requirements generally lead to the use of higher-order (> 4th) band pass filters. A discrete (off-chip), passive band pass filter (BPF) is commonly used [15]. A typical discrete filter characteristic (see [16]) gives a suppression of 20 dB at 30 MHz from the center frequency, increasing to maximum 60 dB at 200 MHz and more from the passband. So, if 110 dB of image rejection must be achieved, as in the GSM handset receiver, two discrete filters may be necessary, e.g. the BPF and IRF in Fig. 2.1.

Using a discrete filter between the LNA and the mixer has severe disadvantages. Both the output of the LNA and the input of the mixer have to be matched to the characteristic impedance of the filter, which may be 50Ω or 300Ω, in order for the filter to work properly. In the LNA generally an extra stage is added to achieve the LNA output matching, so that the power consumption will most likely increase. Moreover, because the filters generally are single-ended, either the building blocks need to be single-ended as well or differential-to-single-ended converters (baluns) must be used. Also, the filters suffer from losses in the passband, amounting to 2 - 3 dB for a typical BPF and 3 - 6 dB for a typical IRF.

2.2.2. DIRECT CONVERSION RECEIVERS

A direct conversion receiver has a single downconversion stage, as depicted in Fig. 2.5. In order to convert the desired channel to baseband, the LO frequency must be equal or close to the RF carrier frequency, the exact location of which is determined by the channel that is used. The receiver is called ‘homodyne’ if the input signal frequency ω_{RF} and LO frequency ω_{LO} are related such that $\omega_{RF} = \omega_{LO}$. In this case the IF is given by $\omega_{IF} = \omega_{RF} - \omega_{LO} = 0$.

The baseband signal can be filtered with a *low pass filter* (LPF). Depending on the amount of adjacent channel interference this LPF may need to have a sharp flank

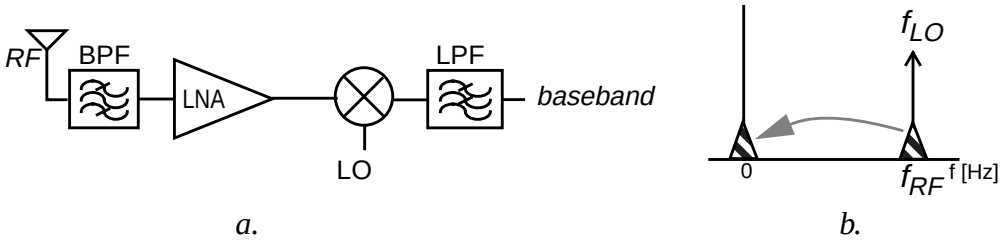


Fig. 2.5. A direct conversion receiver, a). architecture, b). direct downconversion in the frequency domain.

[6], corresponding to a high-order filter. However, due to the relatively low operating frequency this filter may be integrated.

The main advantage of direct conversion is the simple structure, since it has no image frequency problem, and the possibility of utilizing a low pass filter for channel selection. Thus, this architecture offers a high level of integration and a potentially low power consumption. However, complex mixing (I and Q, or *in-phase and quadrature,* having 90° LO phase difference [6]) is typically required, since in most communication systems information is also contained in the phase of the signal. Using a single mixer, this information is lost and cannot be recovered in the baseband.

Some serious drawbacks exist for this architecture [6], [17]. If $\omega_{IF} = 0$ the desired signal will be corrupted by any DC offset occurring in the circuitry. This offset can be minimized using careful design of the front-end, and can then to a large extent be cancelled in the baseband by further digital signal processing. The penalty is an increase in complexity. Another problem is the so-called *1/f* noise (see also paragraph 4.4), a low-frequency noise that can be very large at frequencies near DC, where the received signal will reside after downconversion. The *signal-to-noise ratio* (SNR) at baseband will be deteriorated if the designer does not take measures to reduce the effect of *1/f* noise. This problem is particularly prominent in CMOS transistors. Moreover, the LO signal may leak to either the antenna and thus become an in-band interferer for other users, or to the RF mixer input where this leaked signal is self-mixed with the LO signal, causing a large DC component at the mixer output.

2.2.3. IMAGE REJECTION RECEIVERS

As explained in section 2.2.1, the image constitutes a severe problem in heterodyne receivers. Besides filtering, image reject mixing may be used to reduce this problem. Several solutions exist:

- a Hartley or Weaver architecture
- the use of a polyphase filter
- keeping the I and Q paths separate throughout the front-end, performing the image rejection in the baseband.

In any image reject architecture, mismatch in the I and Q signal paths is a crucial issue. Generally the achievable image rejection using this strategy is in the range of 30 to 40 dB. If the I and Q paths are relatively long, reasonable image rejection will be difficult to achieve. Several highly integrated receivers achieving high image rejection have been reported [15], [18], [19].

Hartley and Weaver Image Reject Architectures

A Hartley downconversion architecture is depicted in Fig. 2.6. The RF input signal, after passing the BPF and LNA, is fed into two mixers driven by quadrature LO signals. The LO signal for each mixer generally is differential, implying that for a set of two quadrature mixers as in Fig. 2.6 four phases of the LO signal are needed.

After the RF signal has passed the mixer, the desired frequency component is selected by means of a low-pass filter (LPF). Thanks to the separation of the IF signal in two branches, an undesired component originating from the image frequency can be cancelled, so it is not seen at the output. This is done by means of a 90° phase shift network and an adder.

Assuming ideal components and an ideal phase shift network, the image is rejected completely. The 90° phase shift network can be split in two phase shift networks of +45° and -45°, one in each branch. If simple *RC* networks are used to implement this phase shift, the amplitude in the two paths is only equal for $\omega = 1/RC$. Therefore, this is a very narrowband solution sensitive to mismatches in *R* and *C*. Non-idealities typically limit the image rejection to 30 to 40 dB.

For a Weaver architecture, the phase shifting block in the Hartley architecture is replaced by a second set of mixers, with ω_{LO2} much lower than ω_{LO1} (see Fig. 2.7). This architecture suffers less from gain mismatch in the phase shifting blocks, but the secondary image frequency may cause problems. This image component is not suppressed when ω_{LO2} is small (see [6]). A Weaver architecture is used in [20].

A special case of image reject receivers, the ‘double-quadrature downconverter’, was presented in [15]. Here both the RF and LO signals are in quadrature. A higher

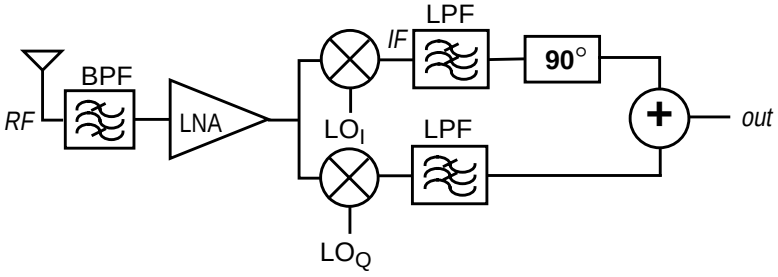


Fig. 2.6. The Hartley image reject receiver architecture.

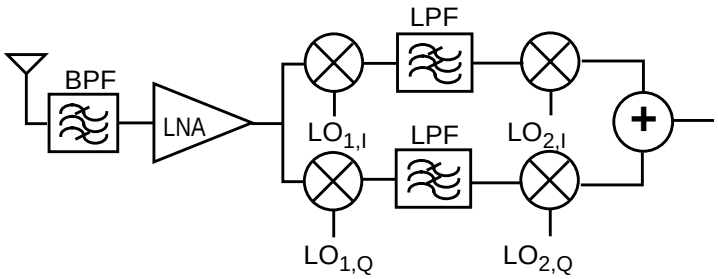


Fig. 2.7. The Weaver image reject receiver architecture.

image rejection is thus achieved at the expense of an increased complexity and higher power consumption.

Polyphase Filters

The low pass filters, phase shifters and adder of the Hartley architecture can be replaced by a single block, a polyphase filter, which may suppress the undesired component. The effective image rejection is dependent on the number of stages, and is related to the bandwidth of the polyphase network [21]. The most convenient way of making a polyphase network is to use 4 x 90° phase shifts, because they can readily be implemented as an *RC* combination. Therefore, each polyphase network consists of one or several stages, which in their turn consist of four *RC* combinations, as shown in Fig. 2.8. At a certain frequency and a certain phase difference between the four input ports, positive frequencies will be passed while negative frequencies will be suppressed, since they have a different phase relation. The negative frequency component suppression is at its maximum for $\omega_i = 1/R_iC_i$. By cascading several stages, each with a slightly different ω_i, a relatively broadband band stop filter for negative frequencies is created, with a bandwidth *BW* and a rejection *L*, as shown in

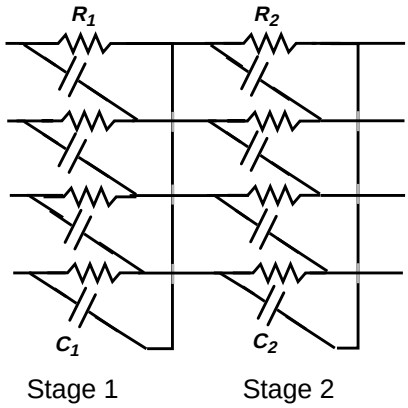


Fig. 2.8. An unloaded two-stage polyphase filter.

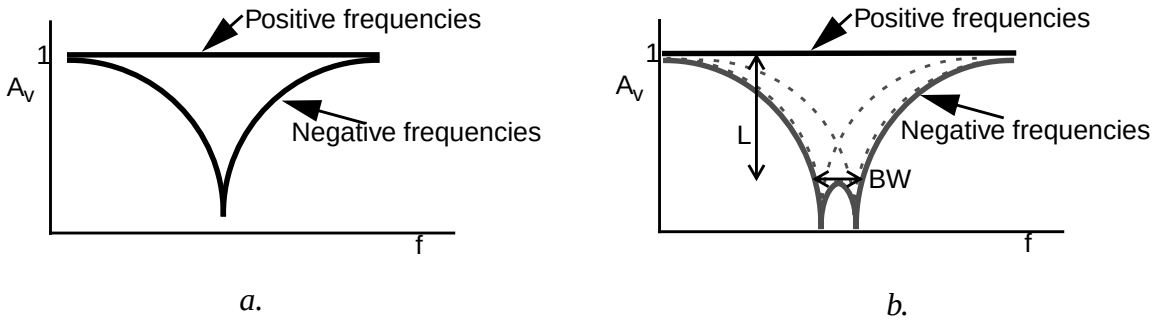


Fig. 2.9. Schematic polyphase characteristic, a). one stage, b). two stages.

Fig. 2.9b. One of the problems of polyphase filters is component accuracy related to processing spread; another problem is loading by subsequent building blocks [21].

2.2.4. SUBSAMPLING RECEIVER

Since most of the receive bands in communication systems are narrowband, the principle of subsampling can be used [22], [23], i.e., an RF signal, with a carrier frequency f_c, is sampled with a much lower frequency f_s. This is illustrated in Fig. 2.10.

In the case of ideal sampling, the bandlimited RF signal at the carrier frequency f_c is multiplied with an impulse train in the frequency domain. The result of ideal sampling is shown in Fig. 2.10c, where both positive and negative frequency components are shown on the same frequency axis. The term subsampling is used when the sampling frequency f_s is substantially lower than the highest frequency component in the sampled signal. If the sampling rate is higher than twice the *bandwidth*, *BW*, of the RF signal, the Nyquist criterion may still be met. More precise requirements are dependent on the bandwidth of the RF signal and its exact frequency location [24], [25].

After sampling the RF signal, filtering can be used to select the desired frequency component, which generally is the baseband component. In Fig. 2.11 a sub-

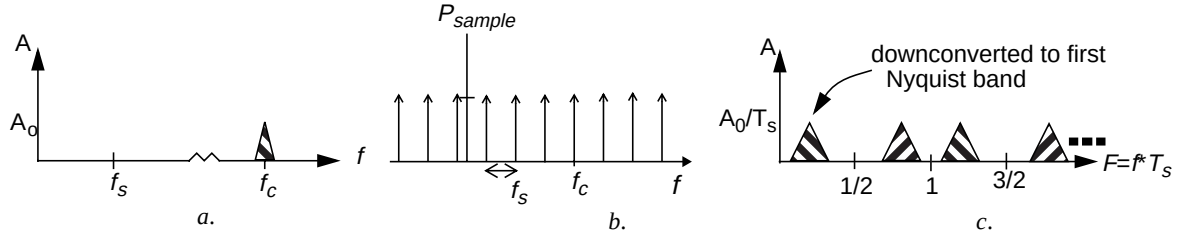


Fig. 2.10. Principle of subsampling: a). The RF input signal, where $f_c >> f_s$, b). Ideal sampling pulse train, c). Discrete-time signal after ideal sampling.

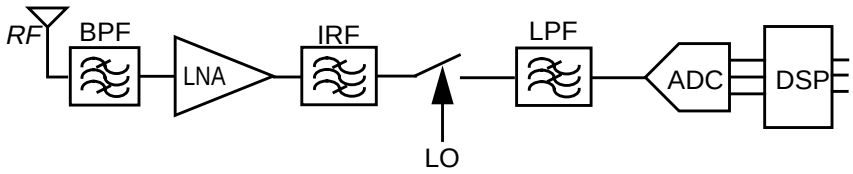


Fig. 2.11. A subsampling receiver architecture.

sampling receiver architecture is shown. A *sample-and-hold* (SH) downconversion circuit can perform both the sampling and selection of the baseband component [22], [23], which is otherwise done inside the A/D converter. The sampled signal is directly converted by the ADC. The receiver architecture is quite simple in comparison to for instance a heterodyne receiver.

An advantage of the subsampling downconverter, compared to for instance the heterodyne downconverter, is that the LO operates at a lower frequency, which with equal power consumption would allow for a better phase noise performance. However, it can be shown [26] that the single side-band noise power density to carrier ratio, $L(\Delta f)$, of the sampled signal and the LO (or clock) signal are related by:

$$L_{signal}(\Delta f) = L_{LO}(\Delta f) \times \left(\frac{f_c}{f_s}\right)^2 \qquad (2.1)$$

where Δf is the frequency offset from the RF carrier frequency f_c. It is then obvious that for large subsampling ratios f_c/f_s, the clock must have a very good phase noise performance. Since the thermal phase noise is dependent on the square of the oscillation frequency f_{LO} [27], a similar VCO power consumption will be needed for a similar overall phase noise performance. This advantage of the lower f_{LO} is thus nulled. Note that at larger frequency offsets the phase noise is determined by the relation of the oscillator output power and its noise floor and by the noise of any additional circuitry between the oscillator and the sampling switch, such as a clock driver [28]. It can therefore be expected that for applications which require good phase noise performance at larger frequency offsets, e.g receivers with stringent blocking requirements, the subsampling ratio f_c/f_s must be kept low.

The major disadvantage of the subsampling mixer is that noise and interfering signals that reach the mixer input, will be aliased to the baseband as well [24]. This results in a fundamentally higher noise figure for the subsampling mixer compared to for instance a Gilbert-cell based mixer, unless the sampling frequency is in the same order of magnitude as the designed analog bandwidth. To reduce the effect of the noise, a pre-select filter can be used [17].

2.3. RECEIVER REQUIREMENTS

The system parameters used in the various communication system specifications (e.g. [7], [8] and [9]) influence the requirements on the front-end building blocks. Before describing these relations the total receiver noise figure and linearity is linked to the noise figure and linearity of the individual blocks in the signal path.

2.3.1. TOTAL RECEIVER NOISE FIGURE AND IIP_3

The description of noise and linearity performance in terms of the total noise figure (*NF*) and *i*-th order input related intercept point, or IIP_i, of a receiver is general and widely used. Fig. 2.12 shows a system consisting of a chain of *m* building blocks, each with a certain available power gain A_{pi}, a voltage gain A_{vi}, a noise factor F_i and an input related 3rd order intercept point $IIP_{3,i}$.

According to Friis' equation [29] the total noise figure of this chain can be expressed as:

$$NF_{tot} = 10\log F_{tot}\ [dB] \qquad (2.2)$$

with the noise factor F_{tot} as

$$F_{tot} = 1 + (F_1 - 1) + \frac{F_2 - 1}{A_{p1}} + \ldots + \frac{F_m - 1}{A_{p1} \cdots A_{p(m-1)}} \qquad (2.3)$$

If the input- and output impedance of each block are matched, then A_v [dB] is equal to A_p [dB], taking $A_v[dB] = 20\log V_{out}/V_{in}$ and $A_p[dB] = 10\log P_{out}/P_{in}$. From the above equation it can be seen that when looking at the noise at the input, the noise of each building block is divided by the gain of preceding blocks. Therefore, it is crucial to have an amplifier contributing little noise, i.e. a low noise amplifier (LNA), as a first building block. Generally in CMOS integrated circuits the input and output impedances of building blocks are not matched, so that an error is made in the total noise figure calculation if Friis' equation is used. This error can be as large as several dB. Therefore, it is preferred to consider the integrated part of the

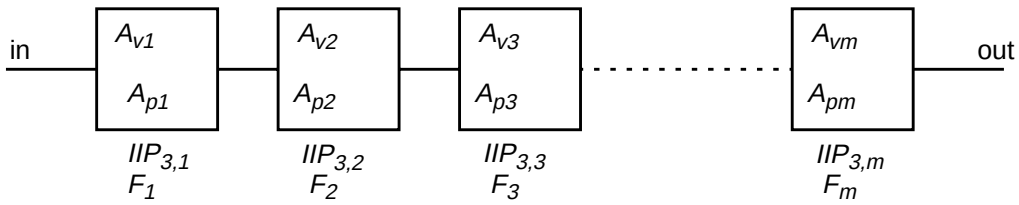


Fig. 2.12. Representation of a system as consisting of m *building blocks.*

receiver as a single unit, where the noise sources can be transferred to the input to calculate the equivalent voltage and current noise spectral density based on circuit analysis [29]. In chapter 4 the physical mechanisms of noise in a MOSFET will be described. For noise in receiver building blocks such as LNAs and mixers see [6], [30], [1].

The input-related 3rd order intercept point IIP_3 is defined as the input amplitude A_0 of two input tones $\Delta\omega$ apart, for which the output third order intermodulation product has the same magnitude as the fundamental output tones. Using simple algebra it can be shown that the third order intermodulation product increases with A_0^3, which in dB is shown in Fig. 2.13. Thus, the slope for the fundamental is equal to 1, while the slope for the 3rd order intermodulation product is equal to 3. However, only at low amplitudes the slopes of the fundamental and intermodulation product may be in accordance with the theory. In real circuits this approximation will not hold since the gain will be compressed and thus the slope of the fundamental will be less than 1 before the intercept point is reached. Moreover, the slope of the 3rd order intermodulation curve may deviate from 3 due to the influence of higher-order harmonics. For a more detailed analysis see [6].

From the plot one can find two ways to extract the IIP_3:

- Plotting the fundamental and the third-order intermodulation product at low amplitudes and extrapolate to the point where the two lines cross.
- A ‘short-cut’ method based on the slopes of 1 and 3: applying one two-tone input signal and using $IIP_3 = \frac{\Delta P}{2} + P_{in}$. Of course this input signal must be chosen with care.

If the IIP_3 of each block in the receiver chain is known, the total IIP_3 can be calculated as:

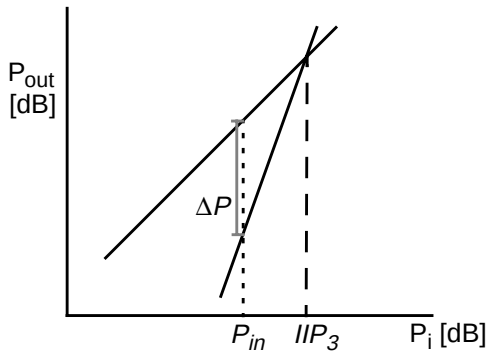


Fig. 2.13. The fundamental and 3rd order output signal for a non-linear block.

$$\frac{1}{(IIP_{3,\,tot})^2} = \frac{1}{(IIP_{3,\,1})^2} + \frac{A_{v1}{}^2}{(IIP_{3,\,2})^2} + \frac{A_{v1}{}^2 \cdot A_{v2}{}^2}{(IIP_{3,\,3})^2} + \ldots \qquad (2.4)$$

where A_{vi} is the voltage gain of block i (not in dB). For a receiver with a BPF, LNA, IRF, a mixer and some IF blocks (see Fig. 2.14), this leads to:

$$\frac{1}{(IIP_{3,\,tot})^2} = A_{BPF}{}^2 \cdot \left(\frac{1}{(IIP_{3,\,LNA})^2} + \frac{A_{vLNA}{}^2 \cdot A_{vIRF}{}^2}{(IIP_{3,\,mixer})^2} + \frac{A_{vLNA}{}^2 \cdot A_{vIRF}{}^2 \cdot A_{vmixer}{}^2}{(IIP_{3,\,IFblock})^2} + \ldots \right) \qquad (2.5)$$

where it is assumed that the filters are perfectly linear. Under the assumption that A_{vi} and A_{pi} behave similarly, it can be seen from Eq. 2.3 and Eq. 2.4 that in order to optimize both the noise performance and the linearity performance of a receiver, a trade-off must be sought for the distribution of the gain. In other words, if most of the gain is designed to be in the first few blocks of the chain, the noise performance is improved while the linearity performance lags. If most of the gain is put in the last few stages, the linearity performance is improved while the noise performance is degraded.

2.3.2. System Specifications vs. Receiver Requirements

In this section the GSM system is used as an example to illustrate the relation between system specifications and receiver requirements such as noise figure and IIP_i. Other standards such as Bluetooth [8] and UMTS [9], [31], [32] have similar specifications but different values.

A GSM mobile receiver must, under various specified test conditions such as the presence of blocking signals, maintain a maximum BER of 10^{-4} [7] [1]. In Fig. 2.15 and Fig. 2.16, the GSM blocking requirements are visualized. For the in-band blocking signals the offset relative to the carrier frequency f_c of the desired signal is given, while for the out-of-band blocking signals the absolute frequencies are used. The blocking requirements provide the worst-case test conditions under which the GSM

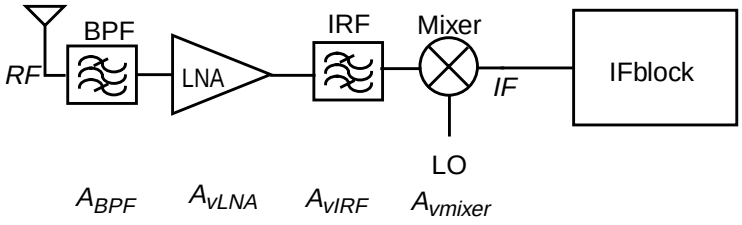


Fig. 2.14. A possible receiver front-end architecture.

1. For Bluetooth and UMTS this is 10^{-3} [8], [9].

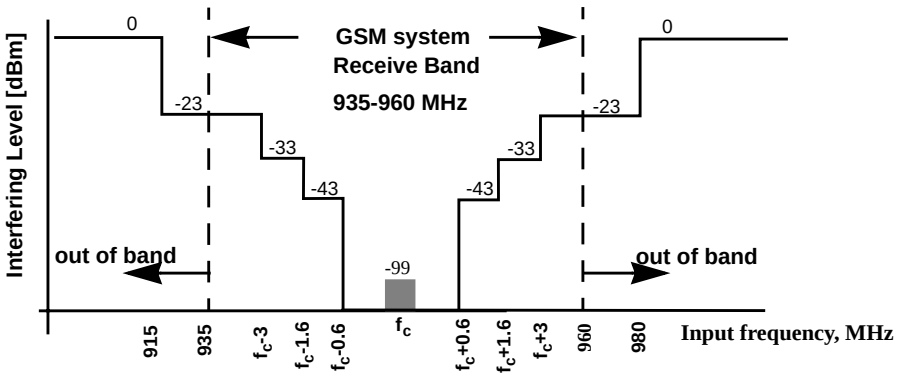


Fig. 2.15. In-band and out-of-band blocking levels for GSM.

receiver has to operate properly. The blocking signals are compared to a small desired signal of -99 dBm, while the co-channel and adjacent-channel interferers are compared to a desired signal of -82 dBm, according to the GSM test specifications [7].

The blocking signals can disturb the desired signal in a number of ways, such as through intermodulation, de-sensitization and reciprocal mixing. In-band blocking signals mainly determine the performance requirements of the frequency synthesizer and the channel-select filter, while out-of-band blocking levels mainly determine the *band pass filter* (BPF, or pre-select-filter) requirements at the input (see Fig. 2.1).

Desensitizing

One of the problems in existing circuits is gain compression, i.e. a reduction of the gain for large input signals. Due to non-linearities, higher order harmonic components exist which become more prominent for large signals (see Fig. 2.17).

For RF circuits a 1-dB compression point is usually characterized, i.e., the input power for which the gain is reduced by 1 dB. From algebraic analysis it follows that a fundamental relation exists between the 1-dB compression point and the IIP_3 (in dBm) [6]:

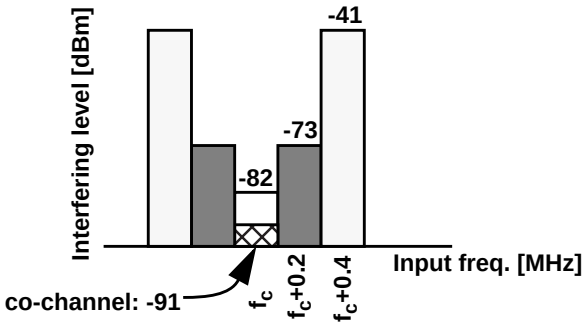


Fig. 2.16. Co-channel and adjacent channel blocking signals

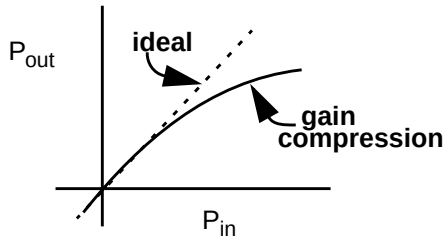


Fig. 2.17. Gain compression for large input signals.

$$IIP_3 \approx P_{-1dB} + 10dB \qquad (2.6)$$

so that if the compression point is known, the IIP_3 can be estimated. Deviation from this relation is due to the influence of higher-order odd harmonics.

If a small desired signal is accompanied by a large blocking signal, the blocking signal can cause gain compression for the desired signal, due to the non-linearity of a building block or element. This is called desensitization of the circuit. The noise figure will deteriorate and small signals can no longer be detected, and thus the sensitivity is reduced.

Harmonics and Intermodulation

Another effect of non-linearities in building blocks is the generation of spurious components. If a single input signal is fed into a non-linear system, higher order spurious signals (harmonics) will be seen at the output besides the fundamental component (see Fig. 2.20). The spurious signals can corrupt the desired signal at any stage in the receiver. This is highly dependent on receiver architecture and building block specifications, as will be shown in par. 2.4.

In Paper III an analysis is made of internal spurious signals, applied to a GSM receiver.

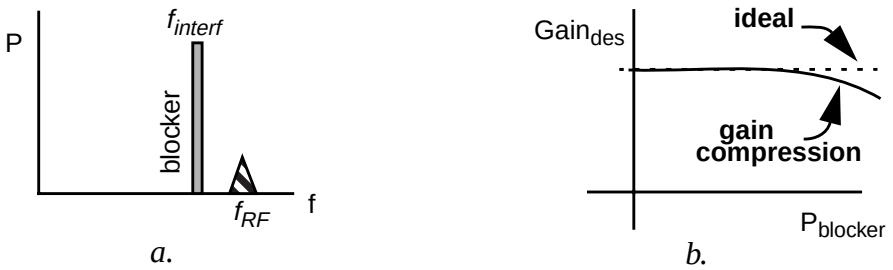


Fig. 2.18. Desensitizing due to a large blocking signal, a). input, b). gain compression for the desired signal due to the blocking signal.

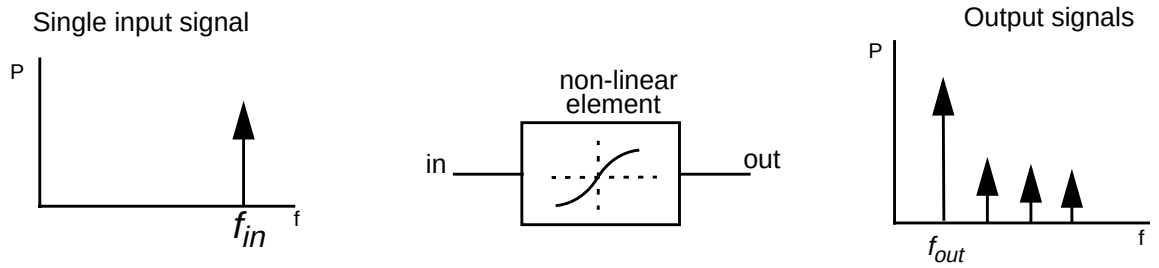


Fig. 2.20. Non-linear system with a single tone input signal.

Intermodulation is defined as the generation of spurious components when two or more signals are fed into a non-linear building block or element (see Fig. 2.19). The principle of intermodulation is valid both for a complete receiver and for a building block such as an LNA or a mixer.

The third-order intermodulation is generally considered to be the most harmful, because the resulting spurious components of two large blocking signals at f_1 and f_2 can fall directly on the desired signal, e.g. if $2f_1 - f_2 = f_c$. This situation occurs for all odd higher-order intermodulation, but the third-order non-linearity of an element is generally the most troublesome. Even-order non-linearity on the other hand results in intermodulation products at DC or low frequencies. These products may be troublesome in a low-IF or direct conversion receiver.

Only one intermodulation test is specified in the GSM specifications [7]: the intermodulation product of a sinusoidal signal of -49 dBm at ± 800 kHz and a modulated -49 dBm signal at ± 1600 kHz from the desired signal [1], is compared to a desired signal of -99 dBm. In order for the receiver to meet the BER requirement the intermodulation product must be at least 9 to 12 dB below the desired signal [2]. If for simplicity it is assumed that the two interfering signals are sinusoidal, the GSM intermodulation specification leads to a total receiver $IIP_{3,tot}$ of:

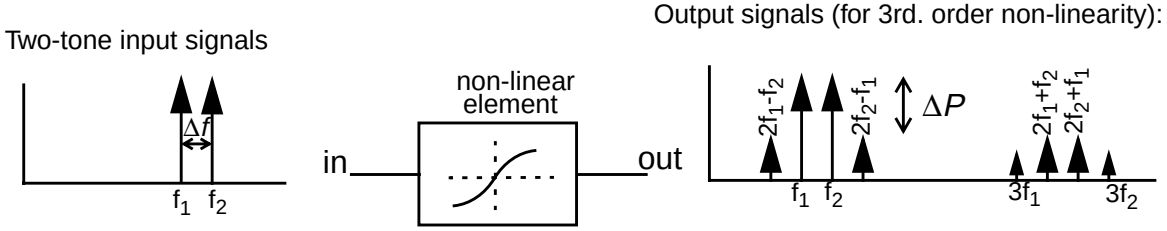


Fig. 2.19. Intermodulation in a non-linear element.

1. so that $|f_1-f_2| = 800$ kHz and $2f_1 - f_2 = f_c$
2. The minimum necessary carrier-to-noise and interference ratio $C/(N+I)_{min}$ is dependent on the type of GMSK detector that is used [33]. Therefore, various numbers can be found for $C/(N+I)_{min}$, ranging from 9 to 12 dB.

$$IIP_{3,\,tot} = \frac{\Delta P}{2} + P_{in} = \frac{(-49 + 99 + 12)}{2} - 49 = -18 \ \text{[dBm]} \qquad (2.7)$$

where ΔP is the difference in power [dBm] between the fundamental component and the intermodulation component at the output, as indicated in Fig. 2.19. In this case, P_{in} is the power of the interfering signals.

The single intermodulation test may appear to be somewhat superficial, since larger blocking signals can degrade the performance, while not mentioned in the GSM intermodulation test requirements. As an example, the GSM blocking spectrum could be applied to the intermodulation test for any two interfering signals satisfying $f_c = 2f_1 - f_2$. This would increase the $IIP_{3,tot}$ requirement dramatically. For instance, two interferers of -23 dBm - the largest in-band blocking signals - would give a minimum $IIP_{3,tot}$ of +21 dBm. If this situation occurs and the desired signal can not be detected, another channel must be chosen.

For third order intermodulation the difference ΔP between the fundamental component and the intermodulation product at the output can be calculated using the IIP_3, as was shown in Fig. 2.13:

$$\Delta P = 2(IIP_3 - P_{in}) = P_{fund,\,out} - P_{3,\,out} \qquad (2.8)$$

where P_{in} is the input power and $P_{fund,out}$ and $P_{3,out}$ are the fundamental and third order output component, respectively. All signal powers are in dBm. More generally, for *k*th order intermodulation the following relation is valid:

$$\Delta P = (k-1)(IIP_k - P_{in}) = P_{fund,\,out} - P_{k,\,out} \qquad (2.9)$$

For single-tone interferers a definition analog to IIP_k may be deduced, namely the input-related harmonic intercept point $P_{k,oi}$:

$$\Delta P_h = (k-1)(P_{k,\,oi} - P_{in}) = P_{fund,\,out} - P_{k,\,out}. \qquad (2.10)$$

The relationship between IIP_k and $P_{k,oi}$ is given by

$$P_{k,\,oi} = IIP_k + 10\log k. \qquad (2.11)$$

Since it is more customary to characterize a building block's linearity through IIP_k, this will be used hereafter, even when only single-tone interferers are considered.

Receiver Noise Figure

The noise figure of a receiver determines how much noise it contributes, and is specified at the input of the receiver. Depending on system specifications such as reference sensitivity and the minimum necessary carrier- to-noise and interference ratio, $C/(N+I)_{min}$, the maximum receiver noise figure may be expressed as

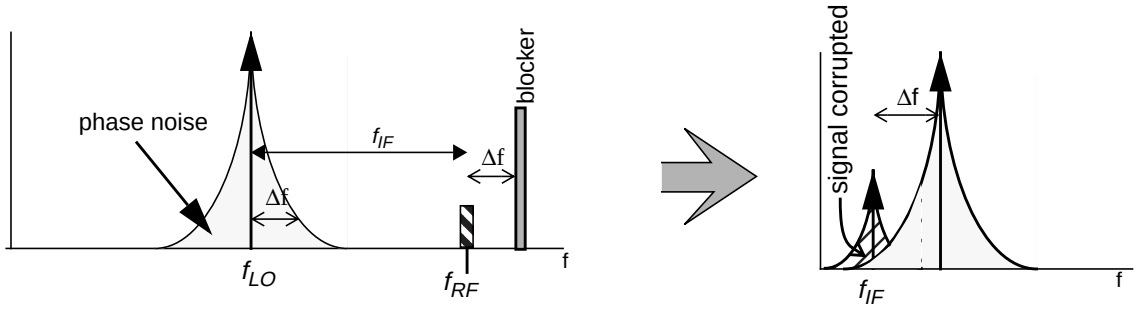


Fig. 2.21. Reciprocal mixing due to phase noise in the LO signal.

$$NF_{RX} < P_{des} - L_{bpf} - C/(N+I)_{min} + 174dBm - 10\log BW\,. \qquad (2.12)$$

Here P_{des} is the power of the desired signal in dBm, $C/(N+I)_{min}$ is the minimum necessary carrier-to-noise+interference ratio in dB, L_{bpf} is the insertion loss of the bandpass filter in dB and BW is the bandwidth of the signal in Hz. Note that this is the noise figure for the receiver excluding the BPF at the input. Using

$C/(N+I)_{min}$ = 9 - 12 dB

BW = 200 kHz

L_{bpf} = 2 dB

P_{des} = reference sensitivity (-102 dBm)

Eq. 2.12 gives

$$NF_{RX} < 5 \text{ [dB]}. \qquad (2.13)$$

Phase Noise Requirements

The mechanism that causes in-channel interference due to phase noise is called reciprocal mixing. The blocking signal is mixed with the LO signal including phase noise, and the tail of the resulting unwanted signal corrupts the desired signal at IF (see Fig. 2.21).

From the blocking and adjacent channel specifications the phase noise requirements for the LO signal of the receiver can be deduced as

$$L(\Delta f) = P_{des} - P_{int}(\Delta f) - C/(N+I)_{min} - 10\log(BW) \text{ [dBc/Hz]} \quad (2.14)$$

where $L(\Delta f)$ is given in [dBc/Hz] and is defined relative to the power of the carrier, i.e. the LO signal power. Furthermore, Δf is the offset frequency from the carrier and P_{int} is the power (in dBm) of the blocker which is located at Δf from the desired signal. In the derivation of the above equation it was assumed that the phase noise has a constant power spectral density over the bandwidth BW.

The blocking specifications as a function of the offset frequency as shown in Fig. 2.15, and the minimum desired signal P_{des} being -99 dBm, result in the following table of phase noise requirements.

TABLE 1. GSM RECEIVER PHASE NOISE REQUIREMENTS

Δf [MHz]	0.2	0.4	0.6 - 0.8	0.8 - 1.6	1.6 - 3	> 3	out-of-band
L [dBc/Hz]	-74	-106	-121	-121	-131	-141	-164 [a]

a. Assuming no filter is present at the receiver input

The requirements for $\Delta f > 3$ MHz and out-of-band are considered to be the most critical for a fully integrated CMOS VCO/frequency synthesizer [34].

Image Rejection

Image rejection can be seen as a special case of blocking signals, since the strength of the image signal is given by the blocking requirements. As was presented in paragraph 2.2.1, the image is located at $2\omega_{IF}$ from the desired signal at ω_{RF} (see Fig. 2.3).

From the GSM specifications it follows that an image rejection of 110 dB is necessary, assuming that the image frequency falls out of band, i.e. an IF of more than 10 MHz. Several measures can be taken to achieve this, besides using a highly selective band pass filter (BPF) preceding the LNA. Some alternatives, such as image reject mixing, were discussed in chapter 2. Another alternative strategy is to use the so-called ‘spurious response frequencies’ [7]. These are defined in the GSM specifications as exceptions for the blocking signals, with a non-predefined frequency, so that the blocking requirement decreases from 110 dB to approximately 80 dB. These exceptions can also be used as a solution for problems with the internal spurious response (see section 2.4). However, since the number of exceptions is limited, utilizing them for the aforementioned two problem areas may conflict.

Cross Modulation

If two signals are fed into a non-linear system, amplitude variations of one signal can appear in the other signal. The effect of this cross-modulation is similar to the effect of desensitization. The amplitude of the desired signal is modulated, possibly leading to this signal being blocked. Again, this is determined by the system non-linearity in combination with the blocking specifications.

The influence of system specifications on design parameters is summarized in table 2.

TABLE 2. SYSTEM SPECIFICATIONS VS. RECEIVER DESIGN ISSUES

	De-sensitizing	Image Rejection	Internal Spurious Response	Intercept Point	Cross Modulation	Noise Figure	Phase Noise
Blocking characteristics	X	X	X		X		X
Adjacent channel signals	X						
Intermodulation characteristics				X			
Reference sensitivity level	X	X	X	X	X	X	X
Bit Error Rate	X	X	X	X	X	X	X

2.4. FREQUENCY PLANNING

Receiver frequency planning related to intermodulation problems [20], [12], [13], [35], [36], [37] is analyzed in this section, taking into account the GSM in-band and out-of-band blocking specifications as shown in Fig. 2.15. The analysis is focused on the front-end of a heterodyne receiver, and a relatively high f_{IF} > 10 MHz is assumed. The analysis relates building block design parameters to both communication system specifications and predefined or predictable receiver parameters. The impact of the use of filters in a receiver and the consequences for the frequency planning of the receiver are clearly demonstrated.

2.4.1. A FRONT-END MODEL

A typical receiver front-end was depicted in Fig. 2.14. The IF block may consist of a programmable gain amplifier (PGA), an IF filter and a second downconversion mixer, followed by a channel select filter and an ADC, but is not specified further at this point. Three main harmonics problem areas can be defined in the front-end:

- Non-linearity in the LNA. This causes different frequency components to occur at the output of the LNA, at f_{in}, $2f_{in}$, $3f_{in}$ etc. where f_{in} can be an interfering signal or a desired signal.

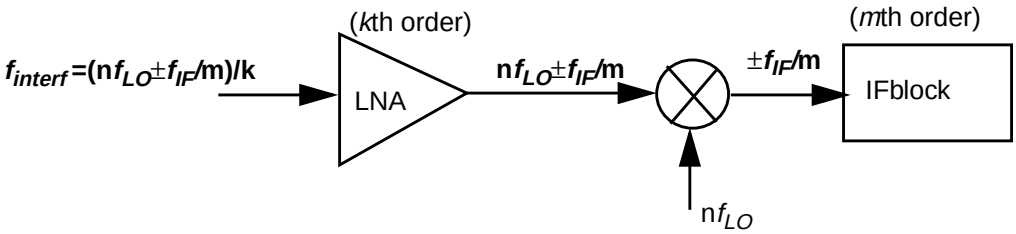


Fig. 2.22. Generation and mixing of harmonics in a receiver front-end.

- Non-idealities in the frequency synthesizer (such as harmonics of f_{LO} or of the reference frequency). Harmonics of f_{LO} appear at $2f_{LO}$, $3f_{LO}$, $4f_{LO}$ etc. The even-order LO harmonics are assumed to be 40 dB below the fundamental due to imperfections in the differential LO path. A square-wave differential LO has odd-order harmonic components with amplitudes of 1/3, 1/5, etc. (i.e. 10 dB, 14 dB lower harmonic output signal). The reference frequency from a crystal oscillator is ignored at this time.
- Non-linearity in the IF block. Mixer output signals at f_{mix} can generate harmonics at $2f_{mix}$, $3f_{mix}$, etc. The non-linearity of the mixer has been incorporated in the 'IF block', i.e., it is assumed that a signal is first downconverted after which non-linearity occurs. This is a simplification and an issue to be investigated further.

The above described sources of non-linearity are illustrated in Fig. 2.22. If the following equation is satisfied, the interferer falls on the desired signal [13]:

$$f_{interf} = (nf_{LO} \pm f_{IF}/m)/k \quad or \quad kf_{interf} = nf_{LO} \pm f_{IF}/m \qquad (2.15)$$

where f_{interf} is the frequency of an interfering signal, and k, m and n are integers representing the non-linearity of the LNA, the IF block and the LO signal, respectively.

In Fig. 2.23 a more extended model of the basic front-end is shown. The LNA input matching network, causing attenuation, has been incorporated in the term

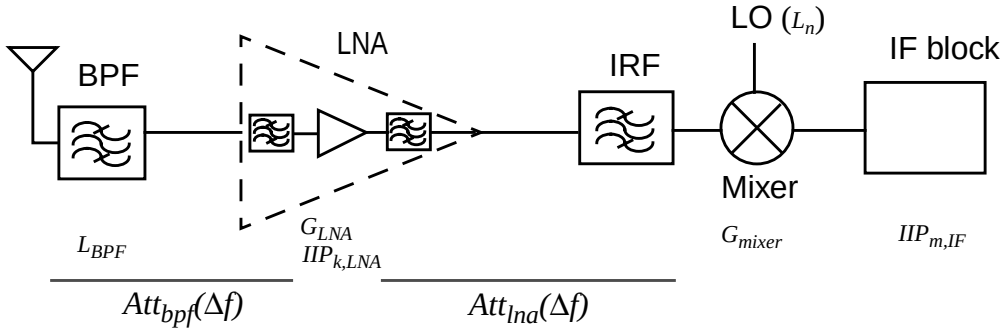


Fig. 2.23. Block schematic of a receiver front-end incorporating gain, attenuation and non-linearity of the building blocks.

$Att_{bpf}(\Delta f)$, while the attenuation caused by the output tuning network is included in the term $Att_{lna}(\Delta f)$. The amount of LNA input attenuation is dependent on the type of LNA; for a Common-Source amplifier the attenuation may be a few dB, while for a Common-Gate amplifier with its broadband input matching no significant attenuation will be seen. The loss of the IRF is incorporated in the mixer gain, $G_{mixer,dB}$. L_n represents the relative loss of the *n*th LO harmonic compared to the LO fundamental. For each combination of *n*, *k* and *m* the parameters indicated in this block schematic will be weighted differently, as will be shown in the next section.

2.4.2. Main Mechanisms for Disturbance from Spurious Signals

Different mechanisms can be distinguished for different combinations of *n*, *k* and *m* as defined in the previous section.

- Mechanism 1: $n>1$, $k>1$, $m=1$ or $kf_{interf} = nf_{LO} \pm f_{IF}$
  A harmonic of the interferer is downconverted with a harmonic of the LO signal. This mechanism depends on the linearity of the LNA, the suppression by various filters and the non-ideality of the frequency synthesizer.
- Mechanism 2: $n>1$, $k=1$, $m=1$ or $f_{interf} = nf_{LO} \pm f_{IF}$
  An interferer is downconverted with a harmonic of the LO signal. This mechanism is dependent on the frequency synthesizer non-ideality and the suppression by the filters before the mixer.
- Mechanism 3: $k=n=1$, $m>1$ or $f_{interf} = f_{LO} \pm f_{IF}/m$
  The non-linearity of the blocks after the mixer causes the downconverted interferer to fall on f_{IF}. This mechanism depends on the suppression by the filters and the linearity of the IFblock.
- Mechanism 4: $k>1$ (*n*, *m* arbitrary) or $kf_{interf} = f_{RF}$
  An interferer located at a subharmonic of f_{RF} produces a component at f_{RF} *before* mixing. This mechanism is only dependent on the non-linearity of the LNA and suppression by the filter before the LNA.

Various combinations of mechanisms can occur as well. However, the effect of these combinations will be negligible in practically every case, since with every mechanism the magnitude of the disturbing signal component is reduced. Therefore the analysis is limited to the aforementioned cases.

If $n=k=m=1$ in Eq. 2.15, it follows that $f_{interf} = f_{LO} \pm f_{IF}$ and the interferer is either a co-channel interferer or an image signal. These issues were discussed in the previous paragraph. Another special case is the so-called *half-IF problem* [6]. An interfering signal in between f_{RF} and f_{LO}, i.e. $f_{interf} = (f_{LO} + f_{RF})/2$ (see Fig. 2.24)

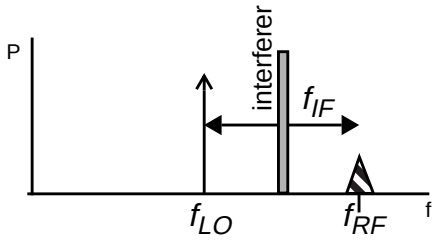


Fig. 2.24. Illustration of the 'half-IF' problem.

may fall on f_{IF}, by means of mechanism 1 (n=k=2, m=1) or mechanism 3 (n=k=1, m=2).

Mechanism 1: Mixing of LO Harmonics with Interferer Harmonics

For mechanism 1 the intermodulation products originating from the LNA non-linearity are multiplied with LO harmonics. The products of this multiplication can fall directly on f_{IF}. Eq. 2.15 becomes $kf_{interf} = nf_{LO} \pm f_{IF}$ (m=1). In order not to degrade the performance these terms must have less power than the desired signal by a margin of $C/(N+I)_{min}$, the carrier-to-noise and interference ratio:

$$P_{interf,IF} < P_{des,IF} - C/(N+I)_{min} \qquad (2.16)$$

The desired signal at the output of the mixer is given by:

$$P_{des,IF} = P_{des} - L_{BPF} + G_{LNA} + G_{mixer} \qquad (2.17)$$

where G_{LNA} and G_{mixer} are the power gain of LNA and mixer, respectively. The interfering signal, downconverted with mechanism 1 for k=3, is then given by:

$$P_{interf(IF)} = 3P_{interf} - 3L_{BPF} - 3Att_{bpf}(\Delta f) + G_{lna} - 2IIP_{3,LNA} - Att_{lna}(\Delta f_3) - L_n + G_{mixer} - 10dB \qquad (2.18)$$

where

P_{interf} = power of interfering signal

L_{bpf} = insertion loss of the band pass filter in the GSM receive band

$Att_{bpf}(\Delta f)$ = attenuation of the BPF (frequency dependent), including the attenuation of the LNA input matching network

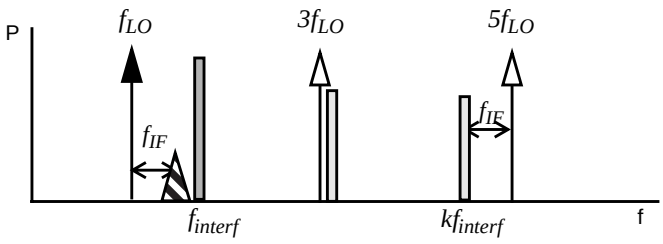


Fig. 2.25. Illustration of mechanism 1.

Δf = distance of interferer to the GSM receive band (935 - 960 MHz)

$Att_{lna}(\Delta f_3)$ = attenuation of the LNA output + image reject filter between LNA and mixer (frequency dependent)

Δf_3 = distance of 3rd order harmonic at LNA output to the GSM receive band

L_n = relative loss of the nth harmonic of the frequency synthesizer relative to the fundamental

P_{des} = smallest desired signal (-99 dBm)

More generally, for kth order intermodulation the interfering signal at the mixer output is

$$P_{interf,IF} = k(P_{interf} - L_{bpf} - Att_{bpf}(\Delta f)) + G_{\text{lna}} - (k-1)(IIP_k + 10\log k) - Att_{\text{lna}}(\Delta f_k) - L_n + G_{mixer} \quad (2.19)$$

An expression can now be derived for mechanism 1 giving a criterion for IIP_k of the LNA:

$$IIP_{k,LNA} > \frac{k}{k-1}P_{interf} - L_{bpf} - \frac{k}{k-1}Att_{bpf}(\Delta f) - \frac{1}{k-1}(Att_{\text{lna}}(\Delta f_k) + L_n) - 10\log(k) - \frac{1}{k-1}\left(P_{des} - \frac{C}{(N+I)_{min}}\right) \quad (2.20)$$

where $C/(N+I)_{min}$ is 9-12 dB for GSM. For a single-ended LNA, IIP_2 and IIP_3 are generally considered to be the most critical.

If the circuit parameters such as IIP_k are known, Eq. 2.20 can be used to determine the required BPF attenuation by rearranging it as:

$$Att_{bpf}(\Delta f) > P_{interf} - \frac{k-1}{k}L_{bpf} - \frac{1}{k}(Att_{\text{lna}}(\Delta f_k) + L_n) - \frac{k-1}{k}(IIP_k + 10\log k) - \frac{1}{k}\left(P_{des} - \frac{C}{(N+I)_{min}}\right) \quad (2.21)$$

If all circuit parameters including the BPF attenuation are known, the distance in frequency Δf can be considered to be the unknown variable. The desired signal and interferer are compared after downconversion, and both the attenuation of different building blocks and the interferer specifications are frequency dependent. Thus, it depends on the choice of IF whether the above requirements will be met. This may be seen from Fig. 2.25; if f_{IF} is smaller, $5f_{LO}$ will lie at a higher frequency, and moreover, the frequency of a possibly disturbing harmonic $3f_{interf}$ will lie closer to the new $5f_{LO}$ since f_{IF} is smaller, resulting in a higher frequency for $3f_{interf}$. Thus, f_{interf} will be higher as well, and this interfering signal will most likely have undergone a larger attenuation from the band pass filter.

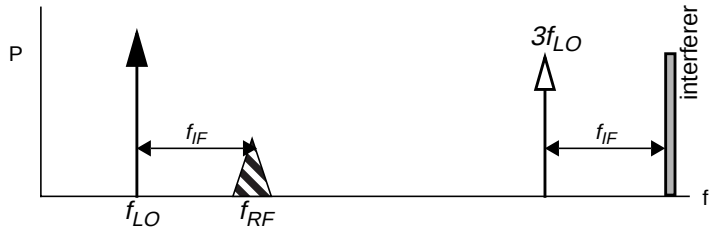


Fig. 2.26. Mechanism 2 illustrated.

This is elaborated further in Paper III.

Mechanism 2: Mixing of an Interferer with an LO Harmonic

An interferer, located at a relatively high frequency, may be downconverted with a harmonic of the LO, and appear at f_{IF}, according to $f_{interf} = nf_{LO} \pm f_{IF}$. This mechanism is independent of LNA or mixer non-linearity, because the fundamental of the interfering signal is involved. Therefore, a slightly different expression follows:

$$P_{interf} - Att_{bpf}(\Delta f) - Att_{\mathrm{lna}}(\Delta f) - L_n < P_{des} - \frac{C}{(N+I)_{min}} \qquad (2.22)$$

or

$$L_n > P_{interf} - Att_{bpf}(\Delta f) - Att_{\mathrm{lna}}(\Delta f) - P_{des} + \frac{C}{(N+I)_{min}}. \qquad (2.23)$$

The ability to deal with these interferers is dependent on the ideality of the frequency synthesizer and the attenuation of different parts (such as filters). With this mechanism Δf will be large, resulting in a large attenuation.

Mechanism 3: Harmonics of Downconverted Interferers

An interferer that is downconverted to f_{IF}/m, suffers from mth-order non-linearity *after* the mixer and produces a component at f_{IF}. Thus, k=n=1, m>1. The most critical interferer is situated in between f_{LO} and f_{RF}, closest to f_{LO}. An interferer located at the other side of f_{LO} could downconvert to the same $|f_{IF}|/m$, but this interferer

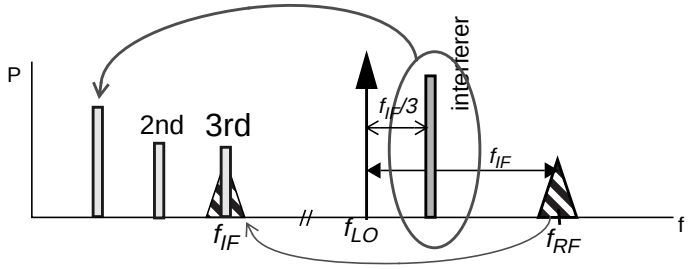


Fig. 2.27. Schematic representation of mechanism 3.

will be subject to more suppression from the filters since the distance in frequency to the RF passband is larger.

For this mechanism a similar expression as for mechanism 1 has been derived, expressing the required $IIP_{m,IF}$ of the IF block in other front-end parameters:

$$IIP_{m,\,IF} > \frac{m}{m-1}P_{interf} - L_{bpf} - \frac{m}{m-1}Att_{bpf}(\Delta f) - \frac{m}{m-1}Att_{\mathrm{lna}}(\Delta f) - 10\log(m) - \frac{1}{m-1}\left(P_{des} - \frac{C}{(N+I)_{min}}\right) + G_{mixer} + G_{LNA} \quad (2.24)$$

The IF blocks generally are fully differential (see for example [10]) so that even-order non-linearities are cancelled to the first order. Therefore, $IIP_{2,IF}$ will presumably be 20 dB higher than $IIP_{3,IF}$ and thus the case for m=2 will have considerably less impact.

Mechanism 4: An Interferer Located at a Subharmonic of the Desired Signal

This mechanism describes the case when the interferer is located at a subharmonic of the desired signal, and a harmonic appears at f_{RF} before the mixer (see Fig. 2.28). No downconversion is involved, and only the non-linearity of the LNA is of importance.

For this mechanism it has been derived that:

$$IIP_{k,\,LNA} > \frac{k}{k-1}P_{interf} - L_{bpf} - \frac{k}{k-1}Att_{bpf}(\Delta f) - 10\log(k) - \frac{1}{k-1}\left(P_{des} - \frac{C}{(N+I)_{min}}\right) \quad (2.25)$$

Δf will be larger for this mechanism than for mechanism 1 or 3, resulting in a larger attenuation. Since in most cases a BPF is used after the antenna, this mechanism has very limited impact.

Practical calculation examples including representative BPF and IRF and building block characteristics were presented in [1] and in Paper III.

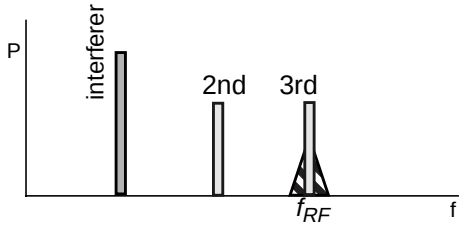


Fig. 2.28. Mechanism 4 depicted.

Summary

The mechanisms discussed above lead to distortion of the desired signal at different points in the receiver. In some cases, such as mechanism 1 and 3, this can be alleviated by the choice of intermediate frequency, IF. This determines the LO frequency as well as the operating frequency of the IF blocks. The filter attenuation characteristics have a large impact on the resulting IF frequency ranges, as have the building block non-linearities. Mechanism 2 may be the most critical one, since it is only dependent on filter attenuation and LO harmonics. The choice of a specific IF does not have any significant effect on this mechanism, although it is slightly better to choose $f_{LO} > f_{RF}$ than vice versa.

If an image-reject filter is used, then mechanism 1 and 2 will hardly play any role and mechanism 4 will not be that severe. However, due to other considerations, such as reduction of the number of external components, this is not the most desirable solution. For GSM, a possible solution may be to use the blocking exceptions (‘spurious response frequencies’ with an input level of -43 dBm [7]), where the exceptions can be redefined for each GSM channel.

In Paper III the above theory is described and applied to a GSM receiver, supported by measurement results.

CHAPTER 3

TRANSMITTERS FOR MOBILE HANDSETS

3.1. INTRODUCTION

In this chapter an overview is given of the most common transmitter architectures as well as some more unusual ones. It is followed by a brief discussion on the impact of communication system specifications on transmitter requirements, as well as transceiver disturbance issues and frequency planning. Special attention is given to power amplifiers, the most challenging building block of an integrated CMOS transmitter.

3.2. TRANSMITTER ARCHITECTURES

The main task of a transmitter is to submit a signal to the antenna at a certain frequency, with a certain power and a certain modulation. Moreover, the signal should not disturb communication in other bands or channels too much. These requirements are set in the communication system specifications. As in receiver design, a balance must be found between performance and issues such as complexity, number of external components and power consumption.

3.2.1. DIRECT CONVERSION TRANSMITTERS

A direct conversion transmitter has a single upconversion stage (see Fig. 3.1 and [6]). The baseband (BB) signal is first modulated, then upconverted to the desired RF frequency by means of a mixer stage driven by an LO signal. The RF signal is then amplified in the PA to get sufficient output power. A band-pass filter may be used between the mixer and the PA in order to prevent the LO harmonics from reaching the antenna. The matching network (MN) between the PA and antenna, which may be either integrated or off-chip, transforms the impedance of the antenna to an optimum load impedance for the PA. Also, a band-pass filter or duplex filter may be used between the matching network and the antenna. A duplex filter provides band pass

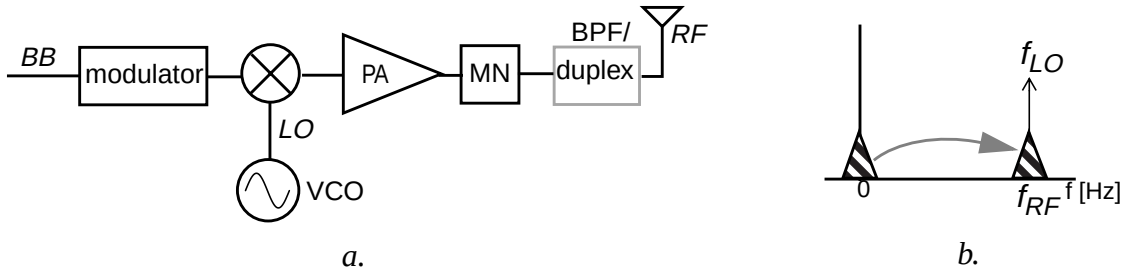


Fig. 3.1. A direct conversion transmitter, a). architecture, b). direct upconversion in the frequency domain.

filtering for both receiver and transmitter, and moreover, it provides isolation between the two parts.

In most communication systems the baseband signal is double side band and asymmetrical around 0 Hz, and is therefore modulated to consist of an *I*- and *Q*- part, so-called quadrature modulation. In that case the upconverter generally consists of two parallel mixers, driven by LO signals with 90° phase difference (quadrature mixers), and an adder, as depicted in Fig. 3.2.

The main advantage of a direct conversion transmitter is its simple structure, which makes it suitable for integration. A disadvantage of this architecture is that the PA output has the same frequency as the VCO (LO). Especially in cases when the two building blocks are parts of a single-chip transmitter and if the isolation between the VCO and PA is not perfect, the PA signal may disturb the VCO signal, a mechanism called injection pulling or injection locking (see par. 3.6, [6], [38] and [39]). This problem is more severe the closer the two signals are in frequency. Strategies to alleviate this problem are discussed in sections 3.2.2 and 3.6.

3.2.2. TWO-STEP TRANSMITTERS

In analogy to heterodyne receivers, also in transmitters an architecture with two mixer stages may be used (see Fig. 3.3), called a two-step transmitter [6]. The modulated baseband signal is first upconverted to a frequency ω_{IF}, where

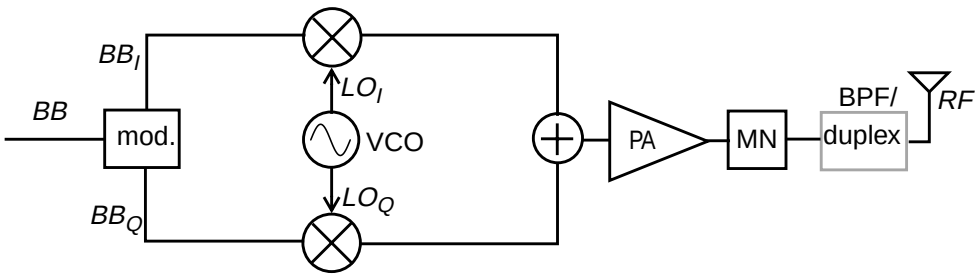


Fig. 3.2. A direct conversion transmitter with quadrature modulation.

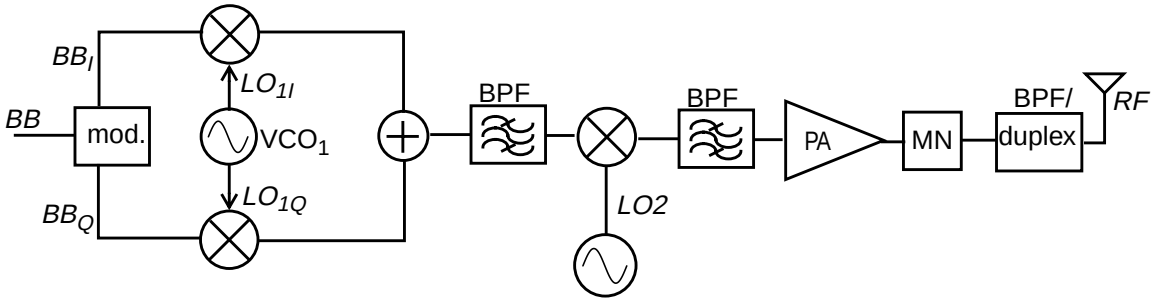


Fig. 3.3. A two-step transmitter architecture.

$\omega_{IF} = \omega_{BB} + \omega_{LO1}$. A band-pass filter is then typically used to suppress IF harmonics. The second mixer stage upconverts the IF signal to the RF frequency: $\omega_{RF} = \omega_{LO2} + \omega_{IF}$. However, a similar signal is created at the image frequency $\omega_{LO2} - \omega_{IF}$. This signal must be filtered out by the band-pass filter between the second mixer and the power amplifier. If ω_{LO1} is low, these two signals are close to each other in frequency, and thus a high-quality off-chip BPF may be necessary. Alternatively image-reject mixing may be used in analogy to image rejection in the receiver. Another problem is LO leakage; the LO signal must be suppressed by filtering or other means.

An advantage of this architecture is that the PA output is not operating at the same frequency as any of the VCOs. Moreover, quadrature upconversion and addition occur at a lower frequency (ω_{IF} instead of ω_{RF}), thus with possibly better I- and Q matching. Disadvantages are its complexity, the necessity for two LO signals and the large number of filters.

3.2.3. DIRECT-MODULATION TRANSMITTERS

In some communication systems such as GSM [40], [41] and Bluetooth [42], constant-envelope modulation schemes are used. This allows for so-called direct modulation, i.e., the baseband signal operates on the control voltage of the VCO. The VCO then generates a modulated RF signal directly, without any need for upconversion mixers. The number of filters may be reduced as well (see Fig. 3.4).

If a frequency synthesizer (FS) is used to stabilize the VCO output frequency, the direct modulation may be compromised by the feedback loop. Thus, in transmit

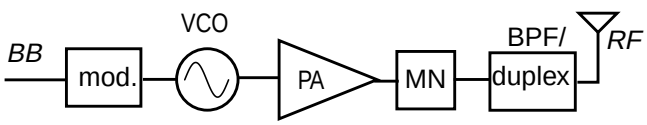


Fig. 3.4. A direct modulation transmitter architecture.

mode the FS loop must be open. Alternatively the baseband signal can be added either to the reference signal or in the feedback path, thus employing the feedback loop to modulate the output signal [40], [43]. In this way the wide-band noise at the reference input may be suppressed by the FS loop, reducing the need for a duplex filter so that a simple RF switch may be used instead [40].

3.2.4. HARMONIC TRANSMITTERS

Instead of using the fundamental LO signal and amplifying the resulting fundamental RF signal, higher-order harmonic signals may be used as the main LO signal. So-called harmonic mixing has been applied mainly in receivers, where a harmonic of the LO signal was used to drive the downconversion mixer. This harmonic was generated by a non-linear circuit, such as an anti-parallel diode pair [44], [45], [46], [47], or an amplifier [48]. In this architecture the VCO and PA will operate at a different frequency.

Another strategy is to use frequency multiplication in the PA. Since the multiplication occurs in the signal path, this may have a severe impact on the signal content. If frequency doubling is used in a narrowband FM system, i.e., $m\int x_{BB}(t)dt \ll 1$, the input signal is given by $x(t) = A\cos\omega_c t - Am\sin(\omega_c t)\int x_{BB}(t)dt$, where m is the modulation index, $x_{BB}(t)$ is the baseband signal and ω_c is the (angular) frequency of the carrier signal. The second order output signal may then be written as:

$$x^2(t) \approx A^2\left[DC + \frac{1}{2}\cos(2\omega_c t) - m\sin(2\omega_c t)\int x_{BB}(t)dt\right] \qquad (3.1)$$

Thus, for a narrow-band system the 2nd order harmonic carries the same signal information as the fundamental. Frequency doubling in the signal path is possible for FM-like modulation systems such as frequency shift keying (FSK) and Gaussian minimum shift keying (GMSK, used in GSM [7], DECT and Bluetooth [8]).

If an amplifier is used to generate the higher-order harmonic, then direct modulation may be used as shown in Fig. 3.5; also in this way the PA and VCO operate at different frequencies. One of the disadvantages is that a harmonic usually is smaller than the fundamental. A power amplifier with a strong 2nd order component was described in [49].

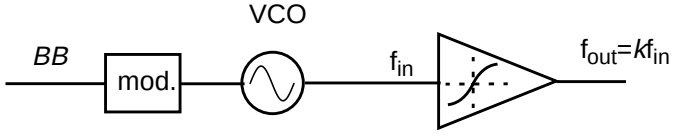


Fig. 3.5. A direct modulation transmitter with frequency multiplication.

In Paper V a CMOS power amplifier with internal frequency doubling is described.

3.3. TRANSMITTER REQUIREMENTS

Important communication system specifications for the transmitter, such as adjacent channel leakage power ratio (ACLR), spectral mask requirements, receive band leakage and spurious emissions, determine transmitter properties such as noise and linearity, and thus building block parameters such as VCO phase noise and PA noise and linearity.

3.3.1. TRANSMITTER NOISE AND NON-LINEARITY

Linearity Requirements

Power amplifier (PA) non-linearity largely determines the out-of-band emission of the transmitter, while non-linearity in other building blocks such as mixers may contribute as well. Like for the total receiver IIP_3 presented in par. 2.3.1, the same equation is valid for the transmitter chain:

$$\frac{1}{(IIP_{3,\,tot})^2} = \frac{1}{(IIP_{3,\,1})^2} + \frac{A_{v1}{}^2}{(IIP_{3,\,2})^2} + \frac{A_{v2}{}^2 \cdot A_{v3}{}^2}{(IIP_{3,\,3})^2} + \ldots \qquad (3.2)$$

In case of the transmitter the output related intercept point is of more interest:

$$\frac{1}{OIP_{3,\,tot}{}^2} = \frac{1}{\left(A_{v1}{}^2 \cdot A_{v2}{}^2 \cdot A_{v3}{}^2 \cdot \ldots\right)}\left(\frac{1}{(IIP_{3,\,1})^2} + \frac{A_{v1}{}^2}{(IIP_{3,\,2})^2} + \frac{A_{v1}{}^2 \cdot A_{v2}{}^2}{(IIP_{3,\,3})^2} + \ldots\right) \qquad (3.3)$$

where A_{vi} is the voltage gain and IIP_i the input-related 3rd order intercept point of block *i*. From this equation it can be seen that the non-linearity of the last block - the power amplifier - will dominate the transmitter non-linearity. It must be noted that for highly non-linear amplifiers there is no generally valid relationship between IIP_3 and adjacent channel leakage power ratio (ACLR) - the main UMTS transmitter specification related to linearity [9]. Generally, the PA is characterized by its ACLR, which should be a few dB below the total required ACLR in order to allow for some non-linearity in other building blocks.

Output Noise

Transmitter output noise can be described as a composition of thermal noise and phase noise, where thermal noise determines the noise floor and phase noise causes

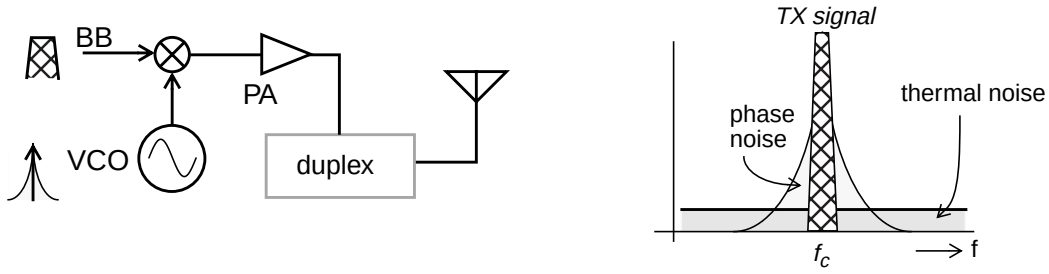


Fig. 3.6. Transmitter, a). common architecture, b). the output signal spectrum including noise.

additional noise around the output signal carrier frequency (see Fig. 3.6). Thermal noise is constant in the frequency domain (so-called 'white noise'), while phase noise originates from the VCO and forms 'skirts' around the carrier frequency having a certain frequency dependency [29], [27].

The transmitter output noise requirements are mainly bound by emission requirements, but also by the receiver noise floor, in case of a non-TDD system such as UMTS, where transmitting and receiving occur simultaneously. Taking UMTS as an example [9], assuming that no duplex filter is present and that no other noise or distortion is seen at the receiver input, the output noise power requirement for the transmitter becomes -167 dBm/Hz. With a duplex TX-to-RX isolation of 50 dB this will be -117 dBm/Hz. Using the out-of-band spurious emission specifications, the toughest requirement (from the frequency range 935-960 MHz) becomes -129 dBm/Hz without duplex filter. A duplex filter attenuation of 40 dB in this frequency band eases this requirement to -89 dBm/Hz. From the above figures it can be seen that a duplex filter is hard to avoid for UMTS. In paragraph 3.3.3 more calculations are presented.

3.3.2. System Specifications vs. Transmitter Requirements

Adjacent channel / spectral mask requirements

In constant-envelope systems such as GSM the transmitter power emission is generally limited by the spectral emission mask requirements (see for instance [7], [40]), while in variable-envelope systems such as UMTS the *adjacent channel leakage power ratio* (ACLR, see [9]) is given as well.

The ACLR is simply defined as

$$ACLR = \frac{P_{adj}}{P_{channel}} \qquad (3.4)$$

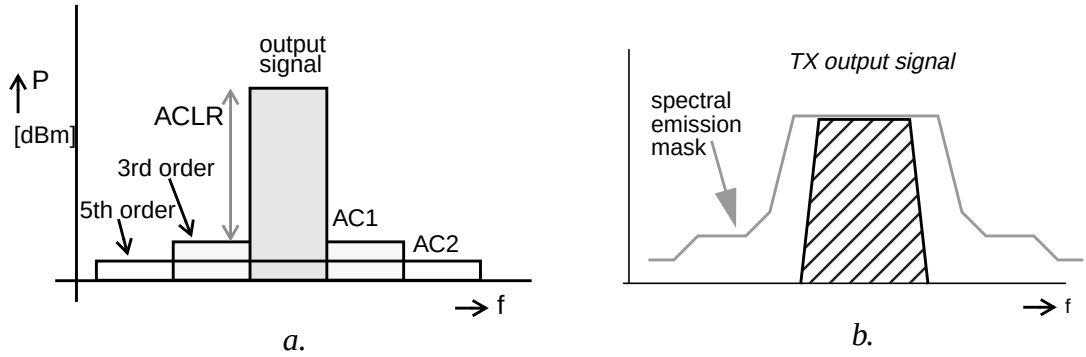


Fig. 3.7. a). Adjacent Channel Leakage Power Ratio illustration, b). spectral emission mask illustration.

where $P_{channel}$ is the power in a channel of a certain bandwidth according to the communication system specifications, and P_{adj} is the power in one of the adjacent channels, as indicated in Fig. 3.7a. It is a measure that comprehends all distortion appearing in the adjacent channel. Intermodulation contributes heavily to ACLR. Thus, when increasing the signal power, the power in the two adjacent channels (AC1 and AC2) will increase faster than that in the desired channel, in accordance with Fig. 2.13, and the ACLR decreases for increasing output power. This is also called 'spectral regrowth'. In system specifications, ACLR is therefore specified for a certain output power.

Spectral mask requirements limit the transmitter output signal spectrum close to the desired signal, as is illustrated in Fig. 3.7b. Both the ACLR and the spectral emission mask have an impact on the transmitter noise and linearity requirements. Calculation examples will be presented in section 3.3.3.

Out-of-band emissions

Besides limited disturbance to adjacent channels, the disturbance from the transmitter to other communication systems must be limited as well. This is ensured by means of the out-of-band emission requirements. For instance, UMTS has emission requirements for the GSM and DCS bands [9].

A mobile user may be disturbed by a neighboring handset, for instance when handset 1 is transmitting with high power, and handset 2 tries to receive a small signal at the same time. Since in some systems the two handsets may operate independently, the "send" time slot of handset 1 may coincide with the "receive" time slot of handset 2, even for a TDD system. This is illustrated in Fig. 3.8.

Because of the separation in frequency between the transmit and receive bands for GSM, this mechanism is determining the far-out phase noise requirement. In the Bluetooth system, communication does not pass through a base station. Instead each

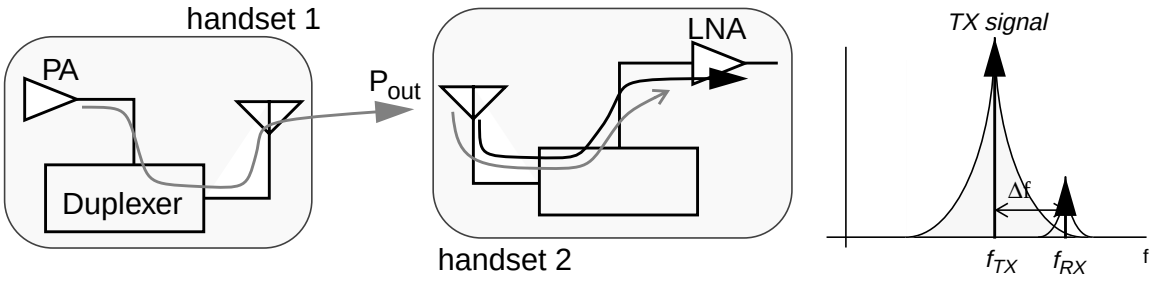


Fig. 3.8. a). The transmitted signal of handset 1 is disturbing the received signal of handset 2, b). Frequency domain illustration.

device can be a master or a slave in a point-to-point network. Transmitting and receiving occur in the same frequency band, and thus it is a non-FDD system.

In analogy to the receiver phase noise requirements, which are largely determined by blocking specifications, these out-of-band emission specifications play a large role for transmitter phase noise specifications.

Transmitter Feedthrough

Another problem of transceiver design is transmitter feedthrough: a high-power TX signal leaks through the duplexer and disturbs the signal to be received by the LNA (see Fig. 3.9). This phenomenon is only problematic for systems where the transmitter and receiver are operating simultaneously, i.e. non-TDD systems such as UMTS/FDD [9], CDMA-I and CDMA-II [50]. Communication systems such as GSM and Bluetooth use TDD and will thus not be affected by this problem. Since the feedthrough is disturbing the received signal, receiver specifications such as sensitivity level and maximum BER come into play.

The influence of system specifications on transmitter design parameters is summarized in Table 3. Moreover, system properties such as duplexing scheme (TDD/

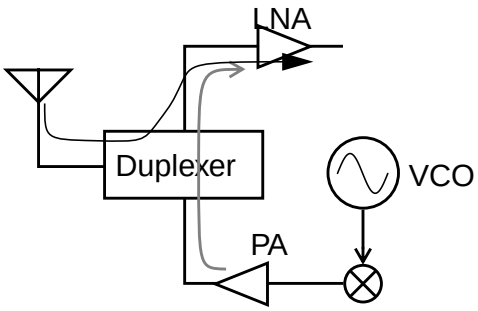


Fig. 3.9. Feedthrough of the transmitter signal into the receiver.

FDD) have an impact on the required transmitter phase noise and linearity, as will be shown in the examples below.

TABLE 3. SYSTEM SPECIFICATIONS VS. TRANSMITTER DESIGN ASPECTS

	TX phase noise	TX output noise floor	TX linearity
Adjacent channel leakage power ratio	X		X
Spectral emission mask	X		X
Out of band emission requirements	X	X	
RX reference sensitivity level	X	X	
Output power levels	X	X	X

3.3.3. TRANSCEIVER CALCULATION EXAMPLES

Noise and Distortion Related to Receiver Noise Floor

The effect of all noise and distortion related to the receiver noise floor will be shown by means of a UMTS example. The total noise and distortion in a 3.84 MHz channel, $P_{N,tot}$, at the receiver input must be lower than $P_{des} - E_b/N_t + PG - IL = -101dB$ where the duplexer insertion loss *IL* is estimated to be 2 dB, and the following parameters have been used [9]:

TABLE 4. KEY UMTS RECEIVER/TRANSMITTER SPECIFICATIONS

parameter	description	value
P_{des}	RX minimum desired signal [a]	-117 dBm
P_{out}	TX output power [b]	24 dBm
E_b/N_t	RX combined received energy per information bit to effective noise psd ratio necessary to meet BER specification	6 - 7 dB [51], [51]
PG	processing gain (chip rate/bit rate)	25 dB
BW	channel bandwidth	3.84 MHz

a. In the so-called Dedicated Physical Channel, DPCH.
b. The maximum transmitter Class 3 output power.

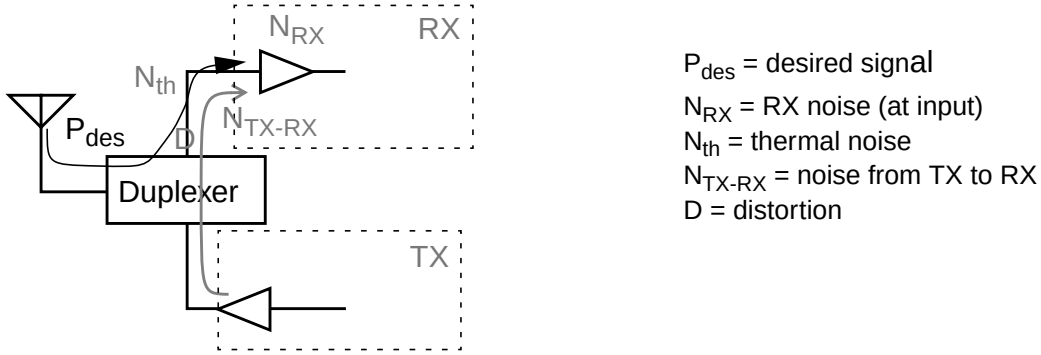


Fig. 3.10. The noise and distortion signals seen at the receiver input.

Taking into account receiver noise, leakage of transmitter output noise to the receiver input and distortion reaching the input, $P_{N,tot}$ can be expressed as (see Fig. 3.10):

$$P_{N,tot} = 10\log(1000kT) + 10\log(BW) + 10\log(F_{RX} + N_{TX-RX} + D) < -101dBm \qquad (3.5)$$

where F_{RX} is the noise factor of the receiver, $N_{TX\text{-}RX}$ is the noise leaked from the transmitter to the receiver input normalized to $4kTR_S$, and D is the normalized total distortion reaching the receiver output. k is Boltzmann's constant ($1.38\text{x}10^{-23}$ J/K) and T is the temperature of operation, assumed to be 300K. Thus, for the noise and distortion at the receiver input, i.e. $10\log(F_{RX} + N_{TX-RX} + D)$, 7 dB is left [51], [52].

$N_{TX\text{-}RX}$ is dependent on the isolation of the duplex filter, ISO_{dupl}, and may be expressed as

$$N_{TX-RX} = \frac{\overline{v_{tx}^2}}{ISO_{dupl} \cdot 4kTR_S} \qquad (3.6)$$

where $\overline{v_{tx}^2}$ (in V^2/Hz) is the noise at the transmitter output in the receiver frequency band. If it is assumed that the transmitter noise and distortion contribute equally and may each contribute 5% to the total noise at the receiver input, the transmitter thermal noise should be -180 dBm/Hz, or -130 dBm/Hz with a duplex filter isolation of 10^5 (50 dB). Moreover, the receiver noise figure should be less than 6.5 dB.

The transmitter to receiver leakage is problematic for systems where transmitting and receiving occur simultaneously, such as UMTS.

Transmitter Phase Noise Requirements to meet the Spectral Emission Mask

The spectral emission mask limits the emission close to the desired signal, and thus limits the allowed close-in phase noise. This is illustrated in Fig. 3.11. The phase noise requirements may be expressed as follows:

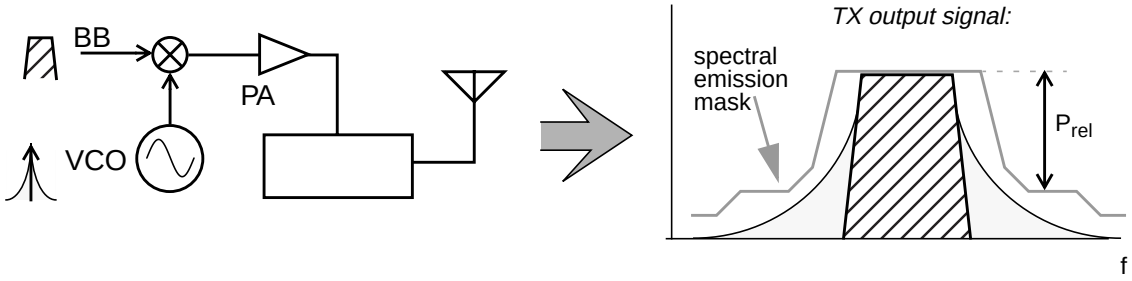


Fig. 3.11. Spectral emission specifications determining the phase noise requirements.

$$L(\Delta f) = P_{rel}(\Delta f) - 10\log(BW) \qquad (3.7)$$

where P_{rel} is the relative power to the desired signal power at a certain distance Δf from the carrier frequency, and BW is the measurement bandwidth as used in the specifications. For the GSM system, the phase noise requirements of the transmitter VCO are given in the table below.

TABLE 5. GSM TX PHASE NOISE REQUIREMENTS TO MEET THE SPECTRAL EMISSION MASK SPECIFICATIONS [7].

Δf [MHz]	0.2	0.4	1.8	3	6
L [dBc/Hz]	-75	-105	-113	-115	-121

The spectrum mask and spurious requirements for Bluetooth result in the following close-in phase noise requirements:

TABLE 6. BLUETOOTH TX CLOSE-IN PHASE NOISE REQUIREMENTS TO MEET THE SPECTRAL EMISSION MASK SPECIFICATIONS [8].

Δf [MHz]	0.5	1.5	≥ 2.5
L [dBc/Hz]	-80	-100	-120

However, it must be noted that spectral emissions are mainly due to non-linearities in the PA. The VCO phase noise should therefore be well below these limits.

Transmitter Phase Noise Requirements to meet the Out-of-Band Emission Limits

As explained before, the transmitter out-of-band emissions are limited especially in the receive band and in frequency bands where other systems operate. For GSM, the

most stringent requirements are in the GSM receive band. They result in certain far-out transmitter phase noise requirements, which are given in the table below:

TABLE 7. GSM TX PHASE NOISE REQUIREMENTS TO MEET OUT-OF-BAND EMISSION SPECIFICATIONS [7].

Δf [MHz]	10 [a]	20 [b]
L [dBc/Hz]	-122	-139

a. Emission specified for the frequency range 925-935 MHz.
b. Emission specified for the frequency range 935-960 MHz.

If a duplex filter is present, the VCO phase noise requirements are eased according to the attenuation from the duplex filter. For a typical GSM duplexer this would amount to more than 30 dB [16].

For UMTS the out-of-band emissions are even specified in the GSM and DCS bands. Both the resulting phase noise and output noise requirements are presented in the table below.[1]

TABLE 8. UMTS TX FAR-OUT PHASE NOISE AND OUTPUT NOISE REQUIREMENTS TO MEET OUT-OF-BAND SPURIOUS EMISSION SPECIFICATIONS [9].

f [MHz]	925-935	935-960	1805-1880 (Δf=40)
L [dBc/Hz]	-101	-113	-105
noise power P_N [dBm/Hz]	-77	-89	-81

Once again, other noise and distortion are not taken into account.

The out-of-band emissions are related to the problem of neighboring handsets. Taking GSM as an example, if one handset transmits the maximum allowable output power and a close-by handset wants to receive the smallest signal, the following equation may be used to determine the phase noise requirements:

$$L(\Delta f) = P_{des} - P_{out} + Att_{dupl}(\Delta f) - C/(N+I)_{min} - 10\log(BW). \quad (3.8)$$

Because of the separation in frequency between the transmit band and receive band for GSM and UMTS, this mechanism determines the far-out phase noise requirement. Relevant parameters are given in Table 9 [7].

1. A typical UMTS duplex filter is assumed, having an attenuation of 40 dB [16].

TABLE 9. GSM RX/TX SPECIFICATIONS

parameter	description	value
P_{des}	RX minimum desired signal	-102 dBm
P_{out}	TX maximum output power	33 dBm
$C/(N+I)_{min}$	RX required carrier to noise and interference ratio	9 - 12 dB
BW	Channel bandwidth	200 kHz
Δf	TX/RX band separation	20 MHz
$Att_{dupl}(\Delta f)$	Duplexer attenuation in RX frequency range	40 dB [16]

This results in a transmitter phase noise requirement of -160 dBc/Hz at 20 MHz. Without a duplex filter this decreases to -200 dBc/Hz.

Arguably the TX output power used in the equation above may be chosen as 0 dBm, since according to the GSM specifications the receiver has to function with a maximum blocking signal of 0 dBm. For that case the minimum desired signal that must be received increases to -99 dBm. Thus, the phase noise requirement without duplexer becomes -164 dBc/Hz at 20 MHz [40]. From these examples it can be seen that the spurious emission requirements do not include the worst-case situation described above.

Transmitter Phase Noise Requirements due to TX-RX feedthrough

Besides the output thermal noise, also the transmitter phase noise may disturb a small received signal, depending on the isolation from transmitter to receiver. For a CDMA system, the phase noise requirements may be expressed as:

$$L(\Delta f) = P_{des} - IL - P_{out} + Iso_{dupl}(\Delta f) - (E_b/N_t)_{min} + 10\log(PG) - 10\log(BW) \quad (3.9)$$

where Iso_{dupl} is the isolation from transmitter to receiver at receive band frequencies. For UMTS, the same parameters as in Table 4 were used. Additional parameters are:

TABLE 10. ADDITIONAL UMTS PARAMETERS

parameter	description	value
Δf	Minimum TX/RX band separation	135 MHz
IL	Insertion loss of duplexer to RX input	2 dB
$Iso_{dupl}(\Delta f)$	Duplexer isolation from TX to RX	50 dB [52]

These parameters result in a transmitter far-out phase noise requirement of -141 dBc/Hz at 135 MHz. Without a duplexer this would decrease to -191 dBc/Hz. This number is very unrealistic, both for a VCO and for a transmitter as a whole.

3.4. POWER AMPLIFIERS

The power amplifier is one of the most crucial building blocks in a transmitter. Its linearity determines spectral regrowth and spurious emissions while its efficiency determines much of the efficiency of the transmitter. For medium or high output powers it is a large contributor to the transceiver power consumption, and thus has a large impact on the battery lifetime.

In this section a brief overview will be given of basic PA parameters, starting with the PA classes, efficiency and linearity. For a more thorough description see [53], [54], [55], [56].

3.4.1. PA CLASSES

The number of classes for power amplifiers and their properties may seem to be exhaustive, with names such as A, B, C, AB, D, E, F, and S. However, these classes can be divided in three groups, starting with A, AB, B and C, then D, E and F, and treating class S individually.

Class A, AB, B and C

For the classes A through C the PA structure is the same (see Fig. 3.12), the only difference being the biasing point [55]. The DC current is supplied through the RF Choke inductor (RFC), while capacitor C_{DC} functions as a DC block, making sure no DC current reaches the output. The input voltage V_{in} will lie above the threshold voltage V_{th} for a limited time, depending on the biasing level at the gate of the MOSFET, as illustrated in Fig. 3.13. For class A, the input voltage exceeds V_{th} at all times, giving a conduction angle α of 2π. For class B, the conduction angle is exactly

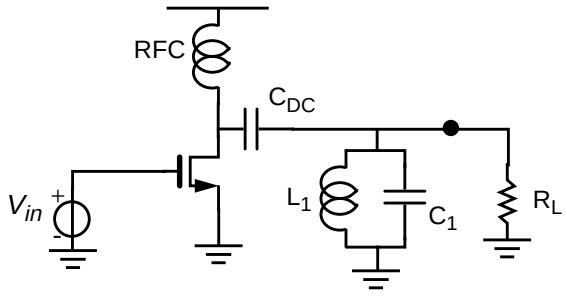


Fig. 3.12. A general class A through C power amplifier.

halved to π. A class AB amplifier has a conduction angle between these values, i.e. $2\pi < \alpha < \pi$, and a class C has a conduction angle less than π.

Equations for the DC current I_{DC}, the fundamental output current I_1 and second harmonic I_2 may be derived as a function of α, assuming the transistor can be modeled as an ideal transconductance [53]:

$$I_{DC} = \frac{I_{max}}{2\pi} \cdot \frac{2\sin\alpha/2 - \alpha\cos\alpha/2}{1 - \cos\alpha/2} \qquad (3.10)$$

$$I_1 = \frac{1}{\pi}\int_{-\alpha/2}^{\alpha/2} \frac{I_{max}}{1-\cos\alpha/2}[\cos\theta - \cos\alpha/2]\cos\theta d\theta \qquad (3.11)$$

$$= \frac{I_{max}}{2\pi} \cdot \frac{\alpha - \sin\alpha}{1 - \cos\alpha/2}$$

and

$$I_2 = \frac{1}{\pi}\int_{-\alpha/2}^{\alpha/2} \frac{I_{max}}{1-\cos\alpha/2}[\cos\theta - \cos\alpha/2]\cos 2\theta d\theta \qquad (3.12)$$

$$= \frac{I_{max}}{2\pi} \cdot \frac{1}{1-\cos\alpha/2}\left(-\frac{1}{3}\sin\frac{3\alpha}{2} + \sin\frac{\alpha}{2}\right)$$

Here I_{max} is the maximum current drawn by the MOSFET. The ideal DC, fundamental and second-order components, normalized to I_{max}, are plotted in Fig. 3.14. It can be seen that the fundamental output component is lower for class C ($\alpha < \pi$) than for class A ($\alpha = 2\pi$). The second-order component, however, is larger for class C than for class A. This is logical as the former is a non-linear amplifier while the latter ide-

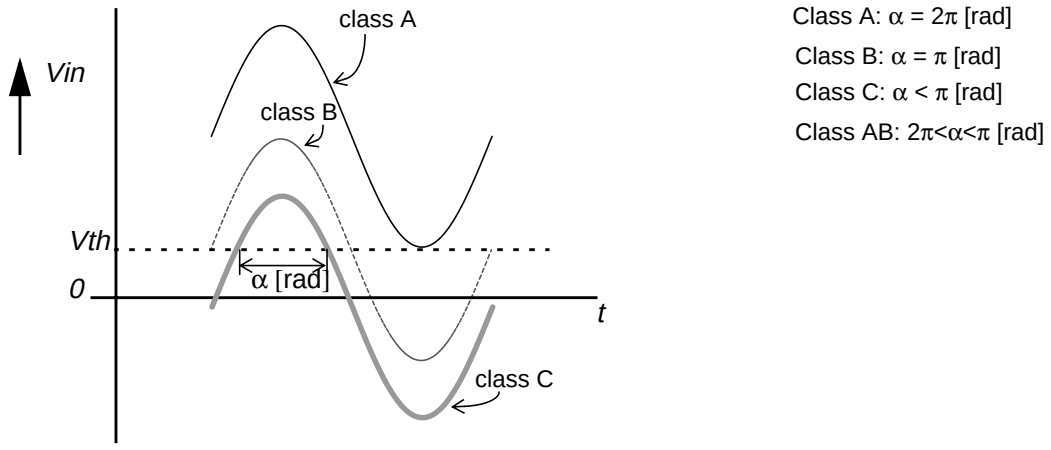


Fig. 3.13. PA input voltage waveforms indicating conduction angle α.

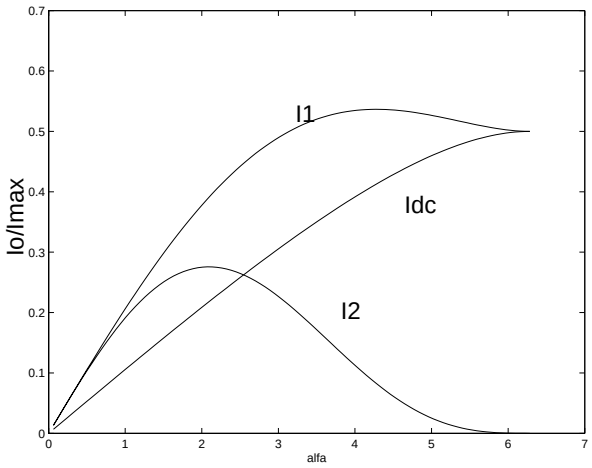


Fig. 3.14. Ideal normalized DC, fundamental and second order output current as a function of the conduction angle α.

ally is perfectly linear. The parallel *LC* circuit (L_1 and C_1 in Fig. 3.12) passes the desired output component while attenuating other tones. Ideally these tones do not reach the antenna.

Class D, E and F

A class D power amplifier (see Fig. 3.15) is based on the use of transistors as switches. The amplifier basically works as an inverter, with the PMOSFET supplying current for a low V_{in} and the NMOSFET supplying current for a high V_{in}. The signal at the drains of the two MOSFETs is a square wave. The series *LCR* network stops undesired harmonics from reaching the antenna. The theoretical maximum drain efficiency (see section 3.4.2) is 100%; non-zero ON-resistance of the switches and finite switching speed will, however, reduce the efficiency [53], [55]. Output power control or amplitude modulation is not possible without changing the supply voltage.

In a class E amplifier, a single MOSFET is used as a switch. The output network as shown in Fig. 3.16 gives the drain voltage a specific shape in the time domain, and the component values must be chosen with care [57]. The voltage over the MOSFET

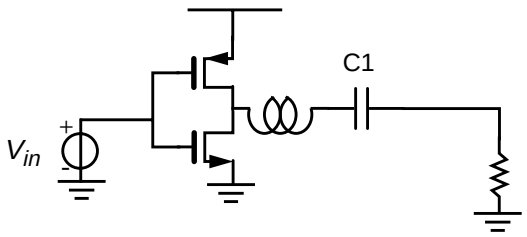


Fig. 3.15. A class D power amplifier.

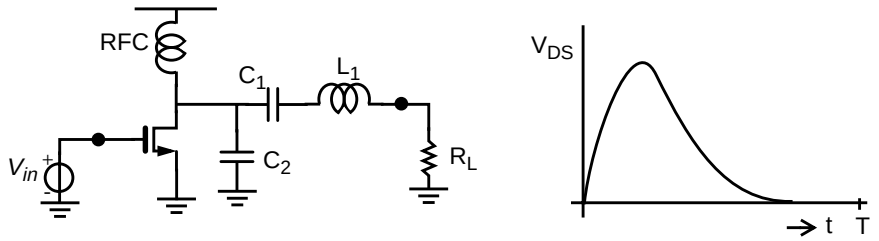


Fig. 3.16. A class E power amplifier and voltage waveform for one period T.

is shaped so that it has both a value of 0 and a slope of 0 when the switch turns on, thus reducing the power loss of the switch. The theoretical efficiency is 100%, depending on the switch ideality. The power handling capability is limited since the peak voltage is high, increasing the risk for breakdown in the MOSFET. Also, power control or amplitude modulation is in principle not possible at a constant supply voltage.

The principle behind a class F power amplifier is a boosting of the higher-order harmonics in order to decrease the peak voltage over the MOSFET when the current is maximum, thus decreasing the power consumption [58]. A simplified schematic and ideal waveforms are shown in Fig. 3.17.

Also for this PA the MOSFET operates as a switch. The theoretical maximum efficiency is 100%. Note that the only difference from a standard class A through C PA (see Fig. 3.12) is the section formed by L_2 and C_2, apart from the MOSFET behaving as a switch in this PA. In classical microwave class F amplifier design a transmission line is used instead of this section [30].

Class S

A class S PA basically is a class D PA with input *pulse width modulation* (PWM), where a sigma-delta modulator may be used for the modulation [59]. However, the

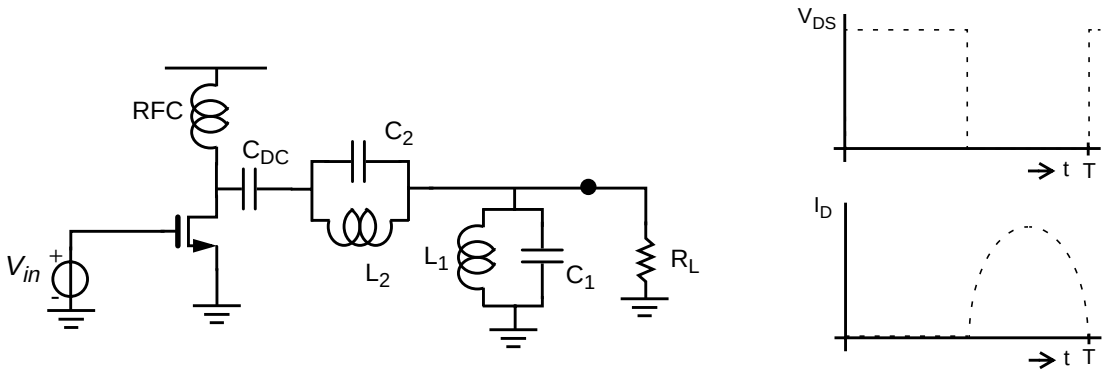


Fig. 3.17. A class F power amplifier with ideal voltage over and current through the MOSFET.

desired modulation puts higher speed requirements on the switches, which makes this class unsuitable for GHz range PAs. Technology advancements may allow for these switching speeds, but this will come with increased power consumption, decreasing the overall efficiency.

3.4.2. OTHER PA ISSUES

The knee voltage

The knee voltage, V_{knee}, is the drain-source voltage at which the MOSFET starts to operate as an amplifier. Assuming the PA consists of one single MOSFET (as in Fig. 3.12), V_{knee} is the same for the PA as for the MOSFET. If a cascode stage is used the PA knee voltage will be increased.

Traditionally the knee voltage is taken to be the voltage where the PA has reached 90% of its drain current in a typical I-V characteristic (see also chapter 4). However, if this is applied to a MOSFET, the knee voltage will be impairingly high. A typical MOSFET definition would be the drain-source voltage where the FET enters the saturation region.

One of the consequences of V_{knee} is a reduction of the maximum voltage swing at the drain, from $2V_{DD}$ to $2(V_{DD}-V_{knee})$. This will have a negative impact on the efficiency.

Efficiency

A key performance measure for power amplifiers is efficiency. Several definitions exist. First, the drain efficiency is defined as

$$\eta = \frac{P_{out}}{P_{DC}} \qquad (3.13)$$

i.e. the ratio of the output power to the DC power. Another definition is *power added efficiency* (PAE):

$$PAE = \frac{P_{out} - P_{in}}{P_{DC}} \qquad (3.14)$$

Here the input power is taken into account as well. For a PA with a high power gain, the drain efficiency and PAE will be almost equal.

For a class A amplifier the maximum fundamental output power $P_{out,max}$ is given by

$$P_{out,\,max} = \frac{\hat{i}_{max}\hat{v}_{max}}{2} = \frac{I_{max}V_{DD}}{4} \qquad (3.15)$$

where $\hat{i}_{max}$ and $\hat{v}_{max}$ are the maximum amplitude of the current and voltage swing at the drain. It is assumed that the PA has an optimum load impedance, that the knee voltage is 0 and that the MOSFET acts as a current source drawing a maximum current of I_{max}. The DC power is given by

$$P_{DC} = \frac{I_{max}V_{DD}}{2} \qquad (3.16)$$

so that the efficiency $\eta = 0.5 = 50\%$.

Similar derivations can be made for all classes. In the table below the theoretical maximum efficiency is given for each class [55].

TABLE 11. THEORETICAL MAXIMUM DRAIN EFFICIENCY FOR PA CLASSES

PA class	max. drain efficiency η [%]
A	50
AB	50-78.5
B	78.5
C	78.5-100
D	100
E	100
F	100
S	100

If the input power of the PA is lower than the input power associated with the maximum output power, maximum efficiency is not reached. Assuming an amplitude of A at the drain, which sees a load resistance R_{opt}, the output power of the PA is

$$P_{out}(A) = \frac{A^2}{2R_{opt}} \qquad (3.17)$$

Here the knee voltage is ignored, and thus $A \leq V_{DD}$ [1]. For class A amplifiers the DC power is given by Eq. 3.16, resulting in an amplitude-dependent efficiency of

$$\eta(A)_A = \frac{P_{out}(A)}{P_{DC,A}} = \frac{A^2}{R_{opt}I_{max}V_{DD}}. \qquad (3.18)$$

For class B, the DC power is

1. Assuming the output signal is voltage limited, not current limited.

$$P_{DC,B} = \frac{2}{\pi}\frac{A}{R_{opt}}V_{DD} \qquad (3.19)$$

resulting in an efficiency of

$$\eta(A)_B = \frac{P_{out}(A)}{P_{DC,B}} = \frac{\pi A}{4V_{DD}}. \qquad (3.20)$$

In Fig. 3.18 the amplitude-dependent efficiencies for class A and B are shown. Assuming the PA is backed off with 3 dB, the amplitude of the output voltage is reduced by a factor of $\sqrt{2}$ [53]. The efficiencies for class A and B are then:

$$\eta\left(\frac{V_{DD}}{\sqrt{2}}\right)_A = 25\,\% \text{ and } \eta\left(\frac{V_{DD}}{\sqrt{2}}\right)_B = 56\,\%. \qquad (3.21)$$

Linearity

As was described previously, PA non-linearity is an important issue in transmitter design, especially for systems with variable envelope modulation. The non-linearity causes various problems, such as gain compression, intermodulation distortion, and AM-to-PM (amplitude modulation to phase modulation) conversion. Both gain compression and intermodulation distortion were already discussed in chapter 2; they affect a power amplifier in a similar way, with some significant differences. The PA is operated with large signals, implying that it is very likely operating in the gain compression region. Thus, amplitude distortion will occur. Moreover, intermodulation behavior is quite complicated in this region and will most likely not follow a 3 dB slope for 3rd order intermodulation. Thus, the intercept point is of limited use for a PA.

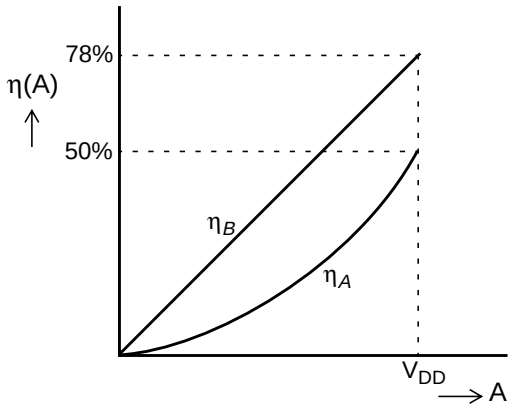


Fig. 3.18. The amplitude-dependent maximum efficiency for class A and B.

AM-to-PM conversion can be defined as an amplitude-dependent phase shift [6]. It is caused by signal-dependent device elements or non-linear resistances, such as the MOSFET input capacitance and resistances in the depletion layer and junctions (see also chapter 4). It is difficult to model these effects accurately, especially since the device is operated at large signals. AM-to-PM conversion results in distortion at the same frequencies as intermodulation distortion.

The ACLR for a transmitter was already discussed in section 3.3. Also for power amplifiers the ACLR is often characterized, since it is a measure of the PA linearity for which realistic signals are used.

3.5. TRANSMITTER LINEARIZATION

Depending on the communication system, the linearity of the transmitter may be a crucial issue. In a system such as UMTS, where a variable envelope modulation is used, a non-linear transmitter would cause a loss of signal information. In systems using a constant-envelope modulation (e.g. GSM, Bluetooth), however, no information is contained in the signal amplitude and linearity is not a critical issue [6], [53]. One way of achieving a better linearity is to back-off the power amplifier [53], i.e to reduce the input signal. This gives a severe reduction in efficiency, and thus other linearization techniques are preferred. In this paragraph a brief overview is given of such techniques.

Predistortion

If the non-linear characteristic of the PA is known in advance, a 'reverse distortion' may be added before the PA, as shown in Fig. 3.19. It must be noted that the non-linearity of a PA may be difficult to predict and is dependent on many factors such as process variations, temperature, impedances, etc. To have a fixed predistortion circuit may thus not yield sufficient improvement in linearity. However, in some applications even a few dB increase in linearity may be significant. More elaborate

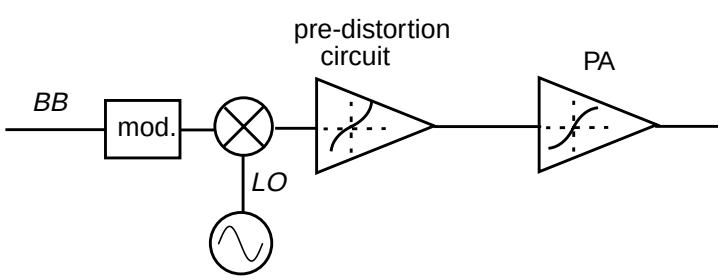


Fig. 3.19. Predistortion linearization technique.

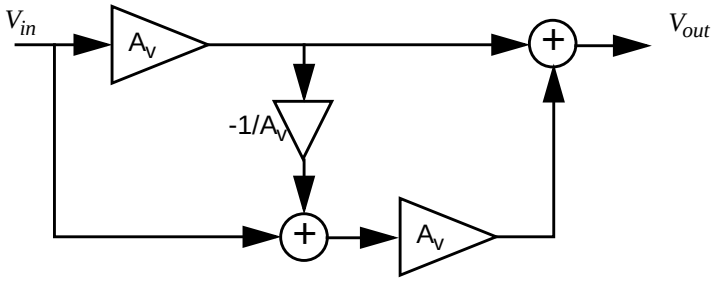


Fig. 3.20. Feedforward linearization technique.

predistortion techniques may be used, incorporating for instance feedback loops and look-up tables [53].

Feedforward

In a feedforward system, the difference between the undistorted input signal and a fraction of the output of the main PA is fed to an auxiliary PA, as shown in Fig. 3.20 [6]. This auxiliary amplifier thus only sees small input signals. Assuming equal gain in both amplifiers as well as ideal adders and connections, the output voltage $V_{out} = A_v V_{in}$. However, phase and gain mismatches in the amplifiers, interconnects and adders (or power combiners) limit the performance of the feedforward architecture. Also amplifier delay may cause problems [53].

Feedback

PA linearization feedback loops generally incorporate a downconversion mixer, so that the output signal can be compared to the baseband signal (see Fig. 3.21), and the loop bandwidth may be kept low. In this figure, θ is the excess phase used to ensure stability in the loop. It is dependent on PA output power, process variations and temperature, implying that it is undesirable to use a fixed value.

For quadrature modulated baseband signals the feedback loop must be expanded into I- and Q paths; the feedback is then called “Cartesian feedback” [6].

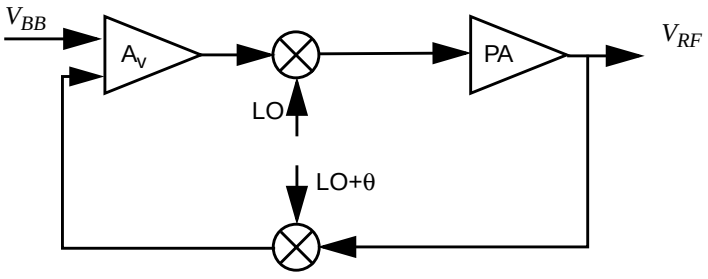


Fig. 3.21. Feedback linearization technique.

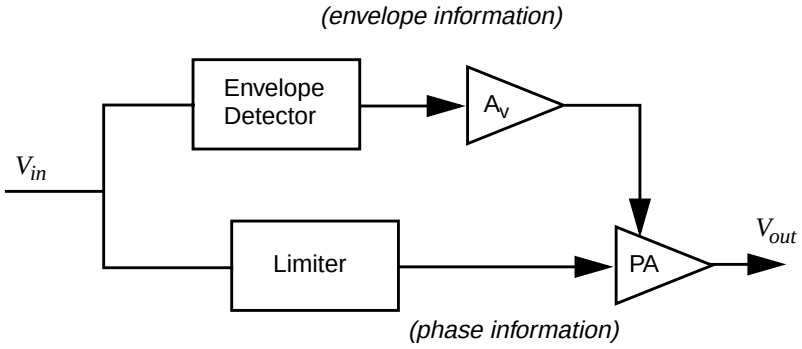


Fig. 3.22. Envelope Elimination and Restoration (EER) linearization technique.

Envelope Elimination and Restoration (EER)

With EER the RF signal is first split into a phase- and envelope signal. The phase signal drives a switching PA and the envelope signal controls the PA supply voltage. Thus, the two components are combined again in the output signal (see Fig. 3.22) [60], [61].

This technique does not require a linear PA. However, the performance is limited by non-ideality of the limiter, non-linearity in the envelope detector, unequal delay in the branches and supply-to-input isolation of the PA. Also, the PA may suffer from AM-to-PM and AM-to-AM conversion, which causes distortion. One of the challenges in this architecture is that the signals in the two branches are of a different nature; the envelope signal is low-frequency, while the phase signal is high-frequency.

Linear Amplification with Non-Linear Components (LINC)

LINC is based on the notion that an amplitude- and/or phase modulated bandpass signal may be decomposed in two constant-amplitude phase modulated signals (see

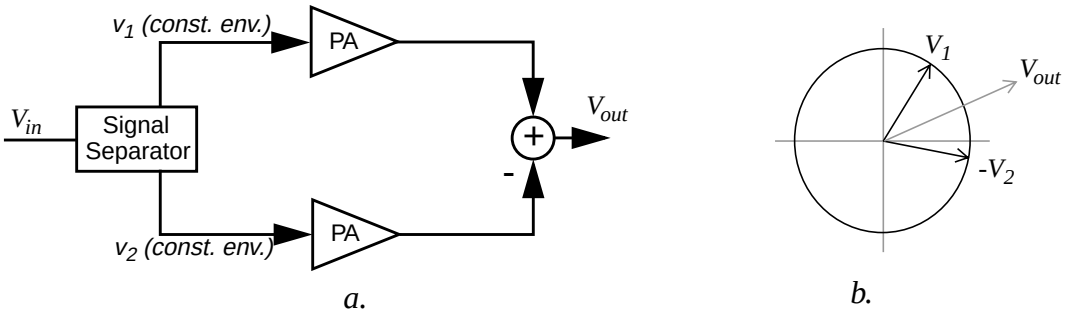


Fig. 3.23. LINC linearization technique, a). architecture, b). signal component vector diagram.

Fig. 3.23) [62], [63]. The most complex part is the signal separator. LINC performance is dependent on mismatch in the two signal paths as well as the isolation between the inputs of the adder. The latter is necessary due to the different phases of the signals in the two branches.

Combined Analog Locked Loop Universal Modulator (CALLUM)

A feedback variety of LINC is called CALLUM [64]. Like the Cartesian feedback architecture it incorporates downconversion. A signal separator is operating at baseband frequency, the two signals are upconverted, amplified, added and fed to the output. The output signal is then downconverted and fed back to the signal separator (see Fig. 3.24).

Advantages of CALLUM compared to LINC are that the signal separator is operating at a lower frequency. Moreover, negative feedback suppresses all distortion independent of origin, depending on the loop gain. Disadvantages are its increased complexity (e.g., several VCOs are necessary) and potential loop stability problems.

Cancellation

Besides transmitter linearization architectures, some work has been done on the building block level, i.e. amplifier linearization. One strategy is to have parallel devices with different bias and dimensions, so that intermodulation products may be nulled when summing the signals at the output [65]. Another strategy is to inject signals at the PA input that cancel the intermodulation products of the original signal at the output [66]. A disadvantage of such techniques is that nulling of intermodulation generally occurs only for a certain bias point or input level. Moreover, reliable simulation of these designs is highly dependent on the quality of the device model.

The PA described in Paper IV has a similar structure as the PA described in [65], but is more focused on efficiency improvement than linearity improvement.

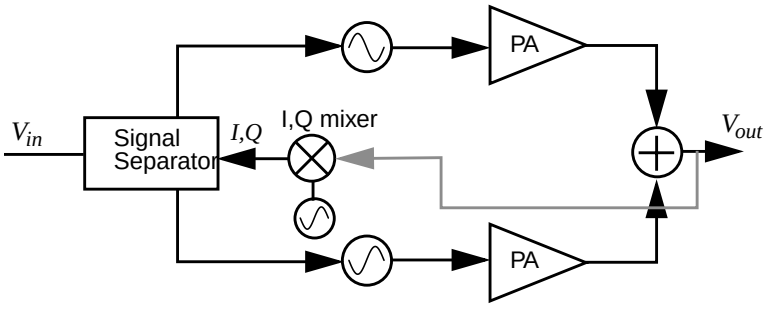


Fig. 3.24. CALLUM linearization technique.

3.6. DISTURBANCE ISSUES IN TRANSCEIVERS

One of the main issues in integrated transmitter or receiver design is the isolation between different signal nodes, since large signals are active on the same chip as sensitive desired signals. In a mixed-signal chip, large digital switching circuits may cause a large switching current, which creates transient voltages on the supply nodes (ground and V_{DD}). In the RF part, high-power PA or VCO signals may leak to other parts of the chip.

The isolation on a chip is limited since there are many paths for disturbances, for example through bondwire coupling or due to finite substrate resistivity. This section will give an introduction to some important disturbance mechanisms in RF circuits. Some consequences of this disturbance are described, as well as solutions, some of which are on the architectural level.

3.6.1. DISTURBANCE MECHANISMS

Substrate Coupling

An important source of disturbance in silicon technologies is coupling through the substrate [74]. Through capacitive coupling of the device to the substrate, disturbance is injected into the latter, which then may reach other circuit elements. Also disturbance on supply nodes may be transferred through the substrate. This is especially problematic for standard CMOS; if other technologies are used such as III-V or silicon-on-isolator, the substrate is highly resistive or isolating, and disturbance will not propagate as easily through the substrate.

Bondwire Coupling

Mutual coupling exists between bondwires, depending on the distance and angle between them, their length and the surrounding medium. If high-power and sensitive signals use adjacent bondwires, the mutual coupling can cause disturbance such as LO-to-RF feedthrough (see [6]) in case of an off-chip VCO, or power amplifier-to-LNA feedthrough.

3.6.2. EFFECTS OF DISTURBANCE

Influence on Devices

Firstly, MOSFETs may pick up disturbances directly from the substrate through the capacitances to substrate. A second mechanism is due to the body effect, described in paragraph 4.2.2, where the body (substrate) can be regarded as a second gate, a back-gate, of the transistor. This generates a drain current depending on the bulk

(body)-to-source voltage v_{BS}. If the source and bulk are connected on-chip, they will see the same disturbance, and v_{BS} will be 0.

Inductors and other passive devices are mainly exposed to disturbances through capacitive coupling to the substrate. While most passive structures will act as a medium to pass disturbance to other parts of the circuit, elements such as junction capacitors are voltage dependent (see paragraph 4.3.1) and can thus suffer from altered characteristics or component values due to the disturbance.

Influence on LNA and VCO

A transceiver as depicted in Fig. 3.25, with a power amplifier (PA), a duplexer for frequency domain duplexing (FDD), and an LNA at the receiver end, will suffer from feedthrough from the PA output to the LNA input. This is mainly due to the non-ideal filtering characteristic of the duplexer, which allows a small portion of the transmitted power through the duplexer to reach the LNA. This small portion may be large enough to desensitize the LNA, since the maximum output power of the PA is on the order of several tens of dBm [22]. Other leakage paths can be the package of the duplexer, bondwires or the printed circuit board on which the components are mounted [19].

Two sources can be pointed out for high power signals reaching the oscillator, namely a signal originating from the power amplifier, and a large interferer in the receive path, both indicated in Fig. 3.25. If the interfering signal frequency is relatively far away from the oscillation frequency f_{LO}, this frequency will not change, but the interfering signal will merely add to the oscillator phase noise.

If the interferer has a frequency close to f_{LO}, then f_{LO} will move towards the frequency of the interfering signal. This is called oscillator pulling and is described in [38]. The issue of oscillator pulling should be taken into account when considering the frequency planning of the transceiver.

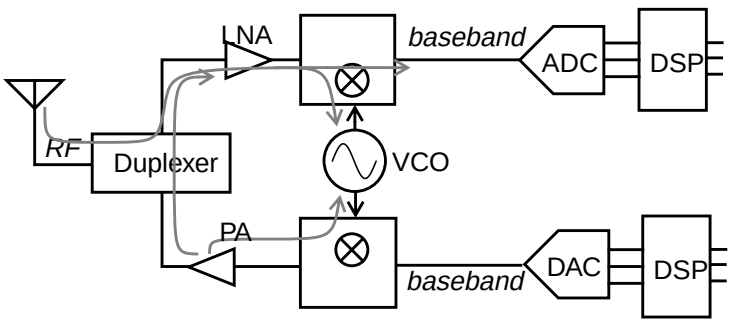


Fig. 3.25. A transceiver with disturbance paths indicated.

3.6.3. METHODS TO REDUCE THE EFFECT OF DISTURBANCE

Layout

On a physical layout level, the placement of sensitive parts is crucial, as is the use of guarding techniques. For a lightly doped silicon substrate (as is used for bipolar and BiCMOS technologies) so-called guard rings can largely reduce the disturbance, since they may act as a current sink for lateral currents. It is important to use a separate ground pin for the guard rings [74]. Physical separation of the circuits or building blocks plays a large role in reducing the effect of disturbances.

A heavily doped substrate behaves more as a single grounded node. The disturbance is much less dependent on physical distance, and eventually stays constant with increasing distance between two blocks [68].

Frequency Planning

The issue of frequency planning for a transmitter is quite different from receiver frequency planning, in that disturbing signals are not coming from the outside world but are generated on-chip or on-board. A critical issue is the coupling of the high-power PA signal to the VCO, especially if they operate at the same frequency. In order to prevent oscillator pulling in the transceiver, a frequency plan must be designed where both transmitter and receiver signals are taken into account. If the system is TDD the receiver and transmitter may share one VCO. However, for a system such as UMTS/FDD the transmitter and receiver may be operating simultaneously, with the possible result of having two VCO signals and one PA signal on at the same time.

Several strategies may be used to circumvent this problem, for instance using a two-step transmitter with a relatively high IF_1 and IF_2 (see Fig. 3.3). If a direct conversion transmitter architecture is used, the LO signal may be composed by mixing two VCO signals (so-called offset LO [6]), or by frequency multiplication [44]-[48]. Another alternative is to have the VCO operate at twice the desired LO frequency, and then use a divider to create the desired LO frequency. An additional advantage of this strategy is that quadrature signals may easily be generated [6].

In Paper V the issue of oscillator pulling is addressed by utilizing a PA with internal frequency doubling.

CHAPTER 4

RF CMOS TECHNOLOGY ASPECTS

4.1. INTRODUCTION

The utilization of CMOS VLSI technologies for RF applications is becoming more and more established. It is made possible mainly by the channel length decreasing to submicron sizes, providing an opportunity to increase the cut-off frequency f_T and maximum oscillation frequency f_{max}.

The advantages of scaling down the transistor dimensions are apparent in digital design, where a steady decrease is seen in the power-delay product [69], [71]. However, in analog design some disadvantages appear. One disadvantage is that the breakdown voltage decreases with reduced physical dimensions, so that lower supply voltages must be used and stacking of transistors is less efficient.

In this chapter CMOS technology and MOSFET operation are briefly discussed [69], [70], [71]. An overview will be given of the influence of scaling, and some problems related to deep-submicron technology are described. Finally inductors will be treated, since they are important for the performance of RF circuits such as LNAs, VCOs, and PAs.

4.2. BASIC MOSFET CHARACTERISTICS AND MODELS

4.2.1. BASIC OPERATION

For most modern standard CMOS technologies a p-type substrate is used. This is a heavily doped substrate (bulk), with a resistivity $\rho < 0.1\ \Omega$–cm, covered with a lightly doped epitaxial layer with a thickness of 5 to 15 μm. The heavily doped substrate is utilized to reduce the risk for so-called *latch-up*, the malfunctioning or destruction of a CMOS circuit through parasitic bipolar transistors forming a thyristor structure [69]. Other measures to prevent latch-up can be found in [69]. The epitaxial layer will be neglected further in this chapter.

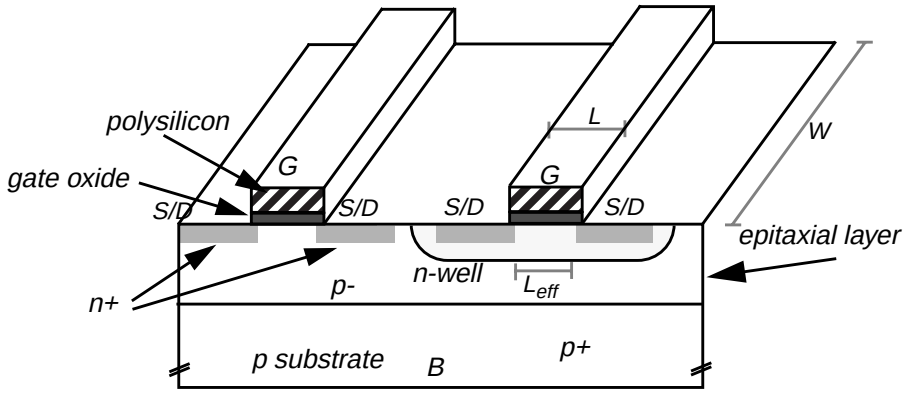


Fig. 4.1. An NMOS (left) and PMOS (right) transistor in a p-type substrate.

In Fig. 4.1 the two basic MOSFETs - n-channel and p-channel - are shown for a p-type substrate. The gate length L is the so-called drawn gate length, which is reduced during manufacturing to the effective gate length L_{eff} [70]. Due to the symmetry of the devices the source and drain are not defined until terminal voltages are applied.

Typically the p-type substrate is connected to the lowest supply voltage, thus keeping the NMOS source- and drain junction in reverse bias condition, while each n-well is generally connected to V_{DD}, but they could theoretically be biased at different voltages.

The n-channel MOSFET will be considered in more detail. The PMOSFET shows similar behaviour, but is not used as often in RF circuits. Depending on the gate-to-source voltage [1], v_{GS}, the MOSFET is said to operate in different 'inversion regions':

- $v_{GS} > V_H$: strong inversion region
- $V_M < v_{GS} < V_H$: moderate inversion
- $v_{GS} < V_M$: weak inversion

where $V_H - V_M \approx 0.5 - 0.6$ V, $V_M < V_{th} < V_H$, and V_{th} is the threshold voltage. This will not be explained further; for a more thorough description see [71]. Assuming $v_{DS} > 0$, the NMOSFET shows two basic regions of operation dependent on the drain-to-source voltage, v_{DS}, and v_{GS}:

- $v_{DS} < V_{eff}$: triode (or linear) region
- $v_{DS} > V_{eff}$: saturation region

1. Each terminal voltage v_{IJ} may consist of a constant bias voltage, V_{IJ}, and a time-varying signal voltage, v_{ij}: $v_{IJ} = V_{IJ} + v_{ij}$. For small signals the bias point and region of operation may be considered constant, but for large signals the bias point and region of operation may become signal-dependent.

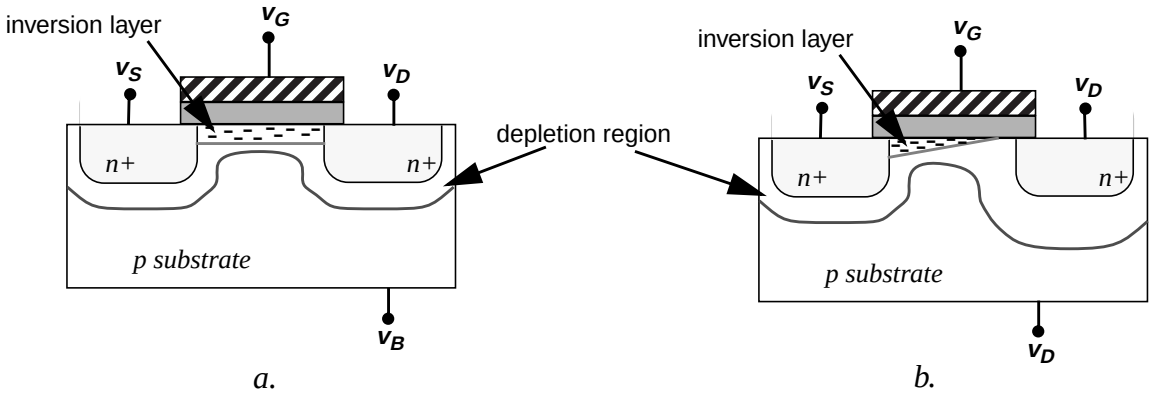


Fig. 4.2. Schematic representation of charges in the channel for a). triode, b). saturation region.

where V_{eff} is the overdrive voltage ($V_{eff} = v_{GS} - V_{th}$). For the most relevant regions of operation expressions can be derived for the drain current I_D and transconductance g_m. These expressions have been derived under certain ideality assumptions, such as constant carrier mobility in the channel, no reverse leakage current and a bulk-source voltage v_{BS}=0. Some of these second-order effects will be covered in section 4.2.2.

The Triode Region

Starting from $v_{DS} = 0$, if v_{GS} is large so that $V_{eff} > 0$ and $V_{eff} > v_{DS}$, the MOSFET is in triode region and an inversion layer is formed as shown in Fig. 4.2a. If a small v_{DS} is applied, drain current can flow:

$$I_D \approx \mu_n C_{ox} \frac{W}{L_{eff}} V_{eff} V_{DS} \qquad (4.1)$$

where

L_{eff} = the effective gate length (see Fig. 4.1)
μ_n = mobility for electrons
C_{ox} = oxide capacitance (see section 4.2.3)
W = transistor width (see Fig. 4.1).

The transconductance g_m, defined as $\partial I_D / \partial v_{GS}$, then becomes

$$g_m = \mu_n C_{ox} \frac{W}{L_{eff}} v_{DS} \qquad (4.2)$$

In this region of operation the MOSFET can be seen as a voltage controlled resistance, controlled by V_{eff}, where the resistance R_{ON} is given by [70]

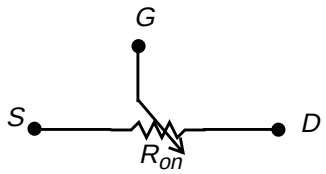


Fig. 4.3. A MOSFET in triode, modeled as a resistance R_{on}.

$$R_{ON} = \frac{1}{\mu_n C_{ox} \frac{W}{L_{eff}} V_{eff}} \qquad (4.3)$$

Note that only in deep triode region, i.e. $v_{DS} << V_{eff}$, the MOSFET can be expected to behave as a linear resistance. A simple triode region MOSFET model is shown in Fig. 4.3. For larger v_{DS}, but with the MOSFET still in triode region, the drain current is given by

$$I_D = \mu_n C_{ox} \frac{W}{L_{eff}} \left(V_{eff} v_{DS} - \frac{1}{2} v_{DS}^2 \right) \qquad (4.4)$$

The Saturation Region (strong inversion)

If the drain-source voltage v_{DS} is increased so that $v_{DS} > V_{eff}$, the MOSFET will enter the saturation region and the channel will be pinched off (see Fig. 4.2b). Raising V_{DS} further moves the pinch-off point slightly towards the source. In this region the drain current can be approximated as

$$I_D \approx \frac{1}{2} \mu_n C_{ox} \frac{W}{L_{eff}} V_{eff}^2 \qquad (4.5)$$

which appears to be independent of v_{DS}, assuming the effective channel length to be constant. In this region the transconductance g_m is given by

$$g_m = \frac{\partial I_D}{\partial v_{GS}} = \mu_n C_{ox} \frac{W}{L_{eff}} V_{eff} = \frac{2 I_D}{V_{eff}} = \sqrt{2 \mu_n C_{ox} \frac{W}{L_{eff}} I_D} \qquad (4.6)$$

In reality the channel length is not constant, but is dependent on v_{DS}. This so-called channel length modulation will be considered in the next section. For small signals and constant bias point, g_m is constant. The most basic small-signal model for the MOSFET in saturation is the hybrid-π model shown in Fig. 4.4.

The MOSFET can achieve higher gain in the saturation region than in the triode region. Therefore, in gain stages MOSFETs are usually biased in the saturation region.

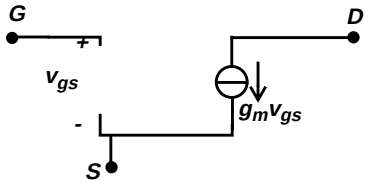


Fig. 4.4. Basic small-signal hybrid-π model of a MOSFET in saturation.

The Weak-Inversion Region

In the weak-inversion region, where V_{eff} is slightly below 0 V, the drain current shows exponential behavior and may be expressed as:

$$I_D \approx I_0 \exp\left(\frac{v_{GS}}{\zeta V_T}\right) \qquad (4.7)$$

where V_T is the thermal voltage defined as kT/q, which is about 26 mV at room temperature (T=300K). ζ is a factor larger than 1. The expression for the transconductance g_m becomes:

$$g_m = \frac{I_D}{\zeta V_T} \qquad (4.8)$$

and is similar to the expression for g_m in a BJT except for the factor ζ.

This region is not suitable for high-frequency operation, but is of use in some low-power/low-frequency applications such as hearing aids. The weak-inversion region points to a problem apparent in digital circuits, which is leakage current: the drain current is not cut off abrubtly for $v_{GS} < V_{th}$, but decreases exponentially. This is one of the main reasons for keeping the threshold voltage V_{th} at a relatively high level instead of scaling down with the technology (see paragraph 4.2.4).

4.2.2. SECOND-ORDER EFFECTS

Two important second-order effects are the *channel length modulation* and the *body effect*. As was already indicated in the previous paragraph, the effective channel length in the saturation region can not be assumed constant, but is dependent on v_{DS}. The drain current including channel length modulation may be given by [70]

$$I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L_{eff}}(V_{eff})^2(1 + \lambda v_{DS}) \qquad (4.9)$$

with $\lambda \propto 1/L$. A typical value for λ is 0.1. For short-channel MOSFETs this linear approximation is inaccurate [70].

In the previous paragraph it was assumed that the bulk - source voltage $v_{BS} = 0$. However, this approximation is not always valid. If the source voltage v_S is varied such that $v_{BS} < 0$ while all other voltage differences are kept constant- alternatively decreasing v_B while keeping all other voltages constant - it can be seen that the depletion region will become larger and thus the threshold voltage for creating an inversion layer will increase. This phenomenon is referred to as the *body effect* [70]. It is dependent on the substrate doping N_{sub} and the oxide capacitance C_{ox}, two important parameters in device scaling of modern technologies. An analysis based on charges can be found elsewhere [70], [71] and will not be given here.

Modeling of Second-Order Effects

In the small-signal model the body effect can be represented by an additional current source controlled by v_{bs}, parallel to the current source controlled by v_{gs} (see Fig. 4.5). The body small-signal transconductance g_{mb} is [70]

$$g_{mb} = g_m \cdot \frac{\gamma}{2\sqrt{2\phi_F + v_{SB}}} \qquad (4.10)$$

where ϕ_F is the Fermi potential and γ is the body effect coefficient. The channel length modulation can be modeled as a resistance r_{ds} at the output [70] with

$$r_{ds} \approx \sqrt{\frac{qN_{sub}}{2\varepsilon_{Si}}} \cdot L \cdot \frac{\sqrt{(v_{DS} - V_{eff})}}{I_D} \qquad (4.11)$$

Both the body effect and channel length modulation are taken into account in the small-signal model shown in Fig. 4.5.

A MOSFET quality parameter is $g_m r_{ds}$, which is the intrinsic gain of the transistor, or the maximum voltage gain that can be achieved in an amplifier using a single MOSFET [1]. This is given by

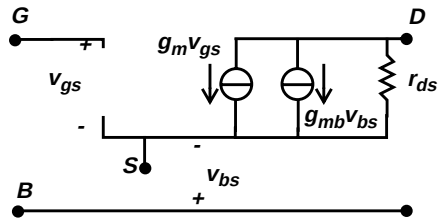


Fig. 4.5. A small-signal model where the body effect and channel length modulation are modeled with $g_{mb}v_{bs}$ and r_{ds}, respectively.

1. Ignoring the body effect; otherwise this is only valid for a CS configuration.

$$g_m r_{ds} = \sqrt{\frac{2qN_{sub}}{\varepsilon_{Si}}} \cdot L \cdot \frac{\sqrt{(v_{DS} - V_{eff})}}{V_{eff}} \qquad (4.12)$$

and is usually larger than 10 for modern technologies [70]. It is bias dependent, as can be seen from the above expression.

4.2.3. FREQUENCY DEPENDENT OPERATION

In this paragraph device capacitances and microwave parameters such as f_t and f_{max} will be considered. It must be stressed that short-channel effects are still not taken into account.

Device Capacitances

Based on a quasi-static analysis of the NMOSFET, the major capacitances in the device are shown in Fig. 4.6 (after [70]). C_{ox} is the capacitance between the channel and the polysilicon gate, given by

$$C_{ox} = WL_{eff}C_{ox}' \qquad (4.13)$$

with

$$C_{ox}' = \varepsilon_0\varepsilon_r / t_{ox} \qquad (4.14)$$

where ε_r is the relative permittivity of the silicon dioxide (about 3.9), ε_0 is the permittivity of free space, and t_{ox} is the oxide thickness. L_{eff} can be approximated as $L - 2L_D$, where L_D is the overlap of the gate on the source- and drain regions (see Fig. 4.6).

The overlap capacitances C_{ov} are due to the physical overlap between the source- and drain areas and the gate:

$$C_{ov} = WC_{ov}' \qquad (4.15)$$

where C_{ov}' is the overlap capacitance per unit width.

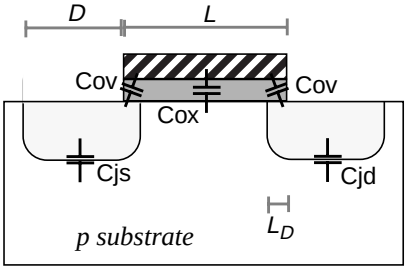


Fig. 4.6. An NMOSFET with parasitic capacitances.

C_{js} and C_{jd} are junction capacitances between the n-doped source and drain regions and the p-doped substrate. Both the bottom-plate and side-walls contribute to the capacitance:

$$C_{j,\,\text{side/bottom}} = \frac{C_{j0,\,\text{side/bottom}}}{(1 + V_{rev}/2\Phi_F)^{m(side,\,bottom)}} \qquad (4.16)$$

where V_{rev} is the reverse voltage across the junction. The value of m depends on the doping profiles, and is typically in the range 0.3-0.4 [70]. The bottom-plate capacitance $C_{j,bottom}$ has the unit of capacitance per unit area, and the total bottom-plate area is the transistor width multiplied by source- or drain- length D. The side-wall junction capacitance $C_{j,side}$ is given in capacitance per unit length, and the total side-wall length is equal to the transistor width W. This gives

$$C_{j(s,\,d)} = WDC_{j,\,bottom} + WC_{j,\,side} \qquad (4.17)$$

It is not straightforward to calculate the zero-bias capacitance C_{j0} for small dimensions. Thus, its value for both the side-wall and the bottom-plate junction capacitance is commonly determined by measurements. If the transistor is placed in an n-well (see Fig. 4.1) the junction capacitance of the well must also be taken into account.

If the device is on, the inversion layer acts as a shield between the gate and bulk [70]. Moreover, the capacitance between source and drain is very small and is usually neglected. The capacitances for the device in saturation can now be written as:

$$C_{GS} = \frac{2}{3}WL_{eff}C_{ox}' + WC_{ov}' \qquad (4.18)$$

$$C_{GD} = WC_{ov}'$$

$$C_{SB} \approx C_{DB} = WDC_{j,\,bottom} + WC_{j,\,side}$$

The basic model for a MOSFET in saturation may thus be extended to the model shown in Fig. 4.7.

In the triode region, the capacitances as shown in Fig. 4.8 are given by

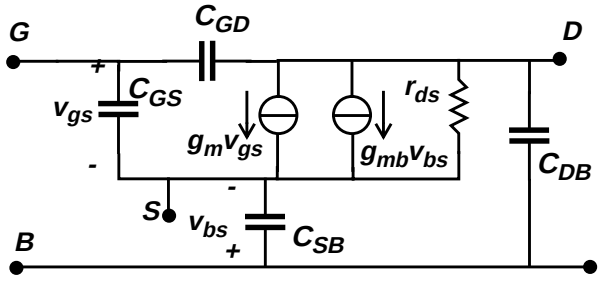


Fig. 4.7. Model of a MOSFET in saturation with device capacitances.

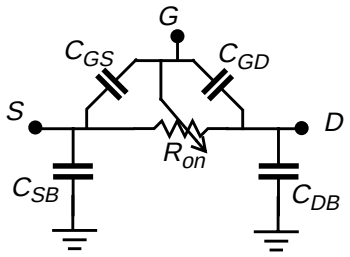


Fig. 4.8. A MOSFET in triode modeled as a resistance R_{on} with parasitic capacitances.

$$C_{GS} = C_{GD} = \frac{1}{2}C_{ox} + WC_{ov}' \qquad (4.19)$$

$$C_{DB} = C_{SB} \approx WDC_{j,\,bottom} + WC_{j,\,side}$$

Since the channel is not pinched off, it is assumed that the oxide capacitance (C_{ox} in Fig. 4.6) is equally divided between source and drain, so that $C_{GS} = C_{GD}$.

MOSFET Microwave Characteristics

A semiconductor technology can generally be characterized by parameters such as f_T, the cut-off frequency, and f_{max}, the maximum frequency of oscillation for a device [71], [72], [73]. The cut-off frequency f_T is defined as the frequency at which the small-signal current gain of the transistor with shorted output is equal to 1. Using Y or H parameters, this can be written as

$$\frac{|y_{21}|}{|y_{11}|} = 1 \quad \text{or} \quad |h_{21}| = 1 \qquad (4.20)$$

For CMOS, f_T can be calculated as [72]

$$f_T \approx \frac{g_m}{2\pi C_g} = \mu_{eff} \cdot V_{eff} \cdot \frac{1}{L^2} \qquad (4.21)$$

which is independent of the transistor width *W*. Here μ_{eff} is the effective mobility, influenced by velocity saturation (see paragraph 4.3).

The maximum oscillation frequency f_{max} is defined as the frequency at which the available power gain of the MOSFET is reduced to 1 [72], thus

$$\frac{|y_{21} - y_{12}|^2}{4(Re(y_{11}) \cdot Re(y_{22}) - Re(y_{12}) \cdot Re(y_{21}))} = 1 \qquad (4.22)$$

so that

$$f_{max} = \sqrt{\frac{f_T}{8\pi R_{g,i} C_{GD}}} \qquad (4.23)$$

where $R_{g,i}$ is the intrinsic gate resistance which can be expressed as

$$R_{g,i} = \frac{R_{poly,\square} W}{3L} \qquad (4.24)$$

The factor of 3 is due to the distributed nature of this resistance and $R_{poly,\square}$ is the sheet resistance of the polysilicon gate [72]. If the gate is contacted at both ends this factor will be 12. In many analog CMOS designs the gate resistance is neglected. However, it will not only affect f_{max} but will also add noise and will influence the input impedance of the device. It can be shown that f_{max} is inversely proportional to $W \cdot \sqrt{L}$ [72].

The model described above is useful for relatively low frequencies. For higher frequencies non-quasi-static effects must be considered. Firstly, in strong inversion the inversion layer will not be able to respond directly to fast changes in the terminal voltages. Thus, a change in the source or drain voltage will cause a delayed change in gate charge. Moreover, the drain current will not respond immediately to changes in the gate voltage. Similar notions are valid for the substrate (the back-gate). Secondly, distributed effects in the extrinsic part may be modeled by using several *RC* sections [71]. In the model shown in Fig. 4.9 both intrinsic and extrinsic effects are taken into account. The latter have the subscript *e*. τ represents a certain delay. For the model parameters see [71] and [72]. It must be noted that calculating the model parameters from the physical behaviour of the MOSFET is difficult. Typically these parameters are fit to measurement results, which reduces the general validity of the model.

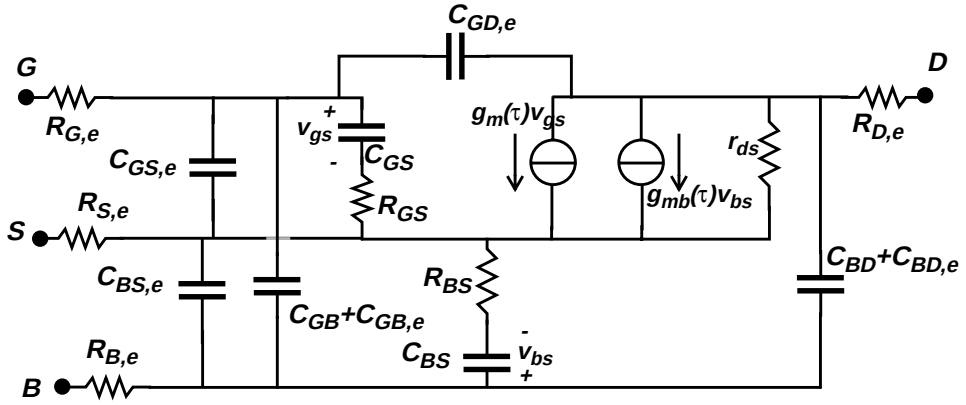


Fig. 4.9. A small-signal high-frequency model for a MOSFET in saturation.

4.2.4. SHORT-CHANNEL EFFECTS

There is an ongoing drive to shrink the channel length of MOSFETs [77]. Digital circuits benefit most, since total chip size shrinks for equal functionality and the power-delay product decreases. RF CMOS design benefits from this trend as well, due to improved high-frequency properties. Typical RF circuits that used to be designed in technologies such as silicon bipolar or gallium arsenide (GaAs) [74], [75] can now be implemented in CMOS [10], [15], [17], [20].

Scaling

In order to maintain similar performance or even increase it with decreased channel lengths, ideal MOSFET scaling (so-called constant field scaling) has been proposed [76], [77], [78] as presented in the table below, where K is a scaling factor larger than 1.

TABLE 12. IDEAL SCALING OF SHORT-CHANNEL DEVICES

parameter	scaling
L, gate length	$1/K$
W, gate width	$1/K$
t_{ox}, gate oxide thickness	$1/K$
x_j, junction depth	$1/K$
N_{sub}, substrate dopant concentration	K
V_{DD}, supply (drain) voltage	$1/K$
I_D, drain current	$1/K$
V_{th}, threshold voltage	$1/K$

The two parameters that deviate the most from ideal scaling are the supply voltage V_{DD} and the threshold voltage V_{th}. Scaling V_{th} increases the leakage current in digital circuits (see paragraph 4.2, weak-inversion region). Therefore neither V_{DD} can be scaled down as much as suggested in table 12.

The main problem with scaling down MOSFETs is the increased electric field in the devices [71]. Several problems arise. First of all, velocity saturation occurs for lower voltages. This can be modeled as a mobility reduction, where the effective mobility, μ_{eff}, depends on the field strength [71]:

$$\mu_{eff} = \frac{\mu_0}{1 + V_{eff} \cdot \frac{1}{L \cdot E_{crit}}} \qquad (4.25)$$

where E_{crit} is the critical field strength, above which the velocity has saturated. Typical values for E_{crit} are $8x10^3$ - $3x10^4$ V/cm for electrons and $2x10^4$ - 10^5 V/cm for holes [71]. As was seen in Eq. (4.21), f_T is related to μ_{eff}, so for decreasing channel lengths f_T will gradually lose its dependence on V_{eff} and will be inversely proportional to L rather than L^2. A second effect of the high field strengths is punchthrough, where, somewhat simplified, the source and drain depletion region are extended so that they touch each other [71]. This can be prevented by increased doping in the channel. A third problem is the hot-carrier effect. In the case of an NMOSFET, electrons in the channel gain enough energy to enter the gate oxide and thus charge it. This will eventually degrade the device, as V_{th} increases and g_m decreases. Moreover, some electrons may cross the gate oxide, resulting in a gate current. Also impact ionization may exist, resulting in more electron/hole pairs close to the drain (so-called 'weak avalanche'). The holes will be transported to the substrate, resulting in a drain-to-substrate current. Two possible remedies to prevent the hot carrier effect are a *lightly doped drain* area (LDD) and a *double doped drain* area (DDD) [69]. Fourthly, the depletion region is extended deeper into the substrate, and especially if $V_{DS} > 0$ the depletion region around the drain will extend towards the source depletion region (see Fig. 4.10). This causes barrier reduction, and is known as *drain-induced barrier lowering* (DIBL) [30], [71]. As a consequence the threshold voltage V_{th} is reduced and becomes dependent on v_{DS}. It affects the output impedance.

The gate oxide may break down if the electric field across it becomes too high. The reduction of oxide thickness has led to a lower breakdown voltage, which in its turn has forced a reduction of V_{DD}, although this reduction does not follow the scal-

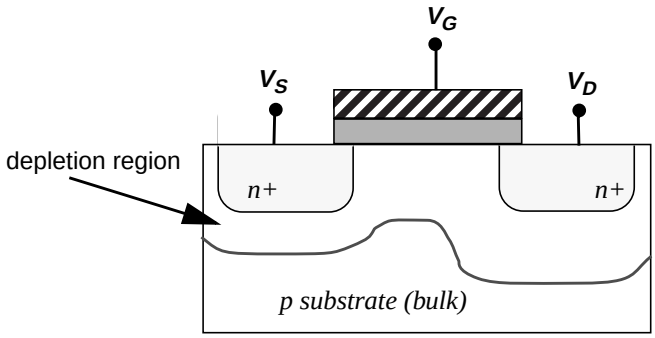


Fig. 4.10. Illustration of the depletion region in a short-channel MOSFET.

ing presented in table 12. For even thinner gate oxides direct tunneling may become a problem [77].

4.3. NOISE SOURCES IN A MOSFET

In this section noise sources in a MOSFET will be discussed as well as a MOSFET noise model.

4.3.1. MOSFET NOISE SOURCES

Since the channel resistance is in fact a physical resistance, a thermal noise current of

$$\overline{i_n^2} = 4kT\gamma g_{d0} \qquad (4.26)$$

can be associated to it [29], where g_{d0} is the channel conductance for V_{DS}=0, and $\overline{i_n^2}$ is the spectral noise power density in [A^2/Hz]. T is the absolute temperature and k is Boltzmann's constant. For a high-resistivity substrate one can assume g_{d0}=g_m, but for a low-resistivity substrate the ratio g_{d0}/g_m will become larger than 1 [29], implying an increased noise current for equal g_m.

In the triode region γ is equal to 1, so that with Eq. (4.26)

$$\overline{i_{n,\,triode}^2} = 4kTg_{d0} \qquad (4.27)$$

while in the saturation region γ can be as low as 2/3 for a long-channel device and as high as 4 for a short-channel device. In addition to channel length, this parameter is dependent on biasing conditions. Values of γ have been deduced from measurements [79].

The so-called *induced gate noise* also originates from the random movements of the charges in the channel, but is capacitively coupled from the channel to the gate [29], [71]. This noise can be expressed as [29]

$$\overline{i_{n,\,ig}^2} \approx 4kT\frac{\omega^2 C_{GS}^2}{3g_m} \qquad (4.28)$$

Because of the term ω^2 in the denominator this noise is a problem mainly for high frequencies, but may also be problematic in LNAs where high-Q inductive source degeneration is used [80].

The gate leakage current I_G tunneling through the reverse-biased source-substrate junction causes shot noise, which is best represented as a current source at the gate of value

$$\overline{i_{n,\,leak}^{\,2}} = 2qI_G \qquad (4.29)$$

Especially for long-channel devices this leakage current is usually quite small. Therefore this noise source is commonly neglected.

Flicker noise (or *1/f* noise) is believed to occur due to the Si-SiO_2 interface [70], [29]. Carriers can be trapped in empty bonds at the silicon surface, and released again randomly. Since the flicker noise power is dependent on the 'cleanness' of the interface, this noise can differ from wafer to wafer and from process to process. If there is no bias current, no *1/f* noise is seen either. The flicker noise can be modeled as a voltage source in series with the gate as

$$\overline{v_{n,\,1/f}^{\,2}} = \frac{K}{C_{ox}WL}\frac{1}{f} \qquad (4.30)$$

where K is a process dependent constant. This constant is larger for NMOSFETs than for PMOSFETs, and the latter thus have a better *1/f* noise performance. MOSFETs have more *1/f* noise than bipolar devices, and the noise is larger for modern CMOS technologies than for older ones. Even though this is a low-frequency phenomenon, it is important in RF circuits, because in non-linear circuits such as VCOs and mixers the noise is transformed to higher frequencies. Moreover, the 1/f noise may disturb the downconverted signal, as was pointed out in chapter 2. To circumvent the *1/f* noise problem passive CMOS mixers may be used.

Noise is also generated in the gate resistance, according to

$$\overline{v_{n,\,gate}^{\,2}} = 4kTR_g \qquad (4.31)$$

Note that this noise source is highly dependent on the layout: a multi-finger structure will reduce the gate resistance and thereby the noise voltage as well.

4.3.2. MOSFET Noise Model

In the MOSFET noise model the noise sources mentioned in paragraph 4.4.1 are taken into account: thermal noise due to the channel resistance, induced gate noise, 1/*f* noise (modeled as a current source), the gate leakage current noise and gate resistance noise (see Fig. 4.11). The thermal noise of the resistances r_s and r_d (see Fig. 4.9) are neglected here. The power spectral densities of the noise sources are repeated here:

$$\overline{i_{n,\,d}^{\,2}} = 4kT\gamma g_{d0} \qquad (4.32)$$

$$\overline{i_{n,\,1/f}^{\,2}} = \frac{K}{C_{ox}WL}\frac{1}{f}g_m^2 \qquad (4.33)$$

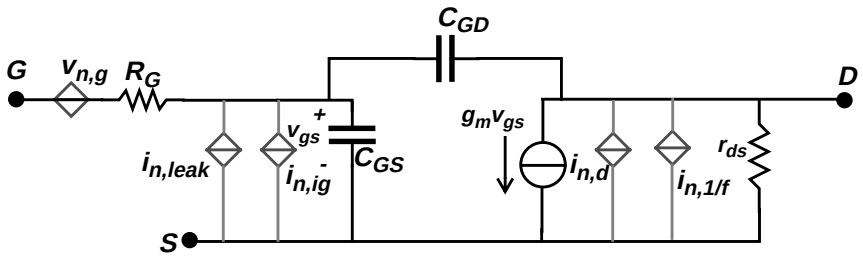


Fig. 4.11. A simplified MOSFET noise model.

$$\overline{v_{n,g}^{\,2}} = 4kTR_g \qquad (4.34)$$

$$\overline{i_{n,leak}^{\,2}} = 2qI_G \qquad (4.35)$$

$$\overline{i_{n,ig}^{\,2}} \approx 4kT\frac{\omega^2 C_{GS}^2}{3g_m} \qquad (4.36)$$

The flicker noise can be modeled as a voltage source in series with the gate (see previous section), but is more often modeled as a current noise source from drain to source. To indicate the power of the flicker noise relative to the thermal noise, a flicker noise corner frequency is defined as the crossing point of the spectral density curves, see Fig. 4.12. Using Eq. 4.32 and Eq. 4.33, this corner frequency f_c can be expressed as:

$$f_c = \frac{Kg_m^2}{4kT\gamma C_{ox}WLg_{d0}} \qquad (4.37)$$

From Eq. 4.32, Eq. 4.33 and Eq. 4.6 it can be seen that MOSFET design parameters *W* and *L* may be used to decrease the *1/f* noise while keeping the channel thermal noise constant, by increasing the product *WL* and keeping the ratio *W/L* constant. The reduced *1/f* noise will, however, come at the cost of increased parasitic capacitance (Eq. 4.18).

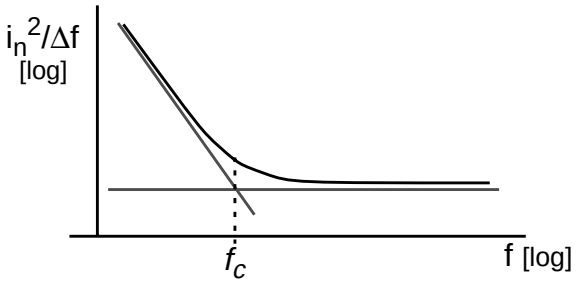


Fig. 4.12. The flicker noise corner frequency.

4.4. MONOLITHIC INDUCTORS IN CMOS TECHNOLOGY

Monolithic inductors are crucial elements in RF integrated circuits, where they typically are used in frequency selective circuits such as tuned LNAs, VCOs, PAs, and possibly mixers. High quality factors *(Q)* are hard to achieve in a standard CMOS technology. In this section the physical background will be discussed as well as basic inductor modeling.

4.4.1. MONOLITHIC INDUCTOR PHYSICS

Inductors may be realized in integrated circuits by a simple metal wire on a more or less insulating layer. The main problems of integrated inductors are the limited quality factor (*Q*) that may be achieved in CMOS technologies, limited inductance values, as well as the relatively large area they consume. Moreover, modeling of inductors is not straightforward, as will be shown in the next paragraph.

The inductance and *Q* depend on various factors such as frequency, geometry of the metal wire, properties of the metal and the insulating layer, and properties of the substrate. In a CMOS VLSI technology the Si substrate generally is heavily doped, i.e. lossy compared to a more lightly doped Si bipolar substrate [81]. This typically limits the *Q* to be less than 10, while discrete inductors can reach a *Q* of 100 and bondwires between 25 and 50.

The most common geometry for an integrated inductor is depicted in Fig. 4.13 [81], [30]. Between closely positioned segments with a current flow in the same direction, e.g. no. 1 and 5 in Fig. 4.13a, positive mutual inductance exists, which adds to the total inductance, while between segments with opposite current flow, e.g. no. 1 and 3 in Fig. 4.13a, negative mutual inductance exists which decreases the total inductance [82]. However, since segment 1 and 3 are quite far apart, the negative mutual inductance will be quite small. The inner turns of a planar inductor hardly

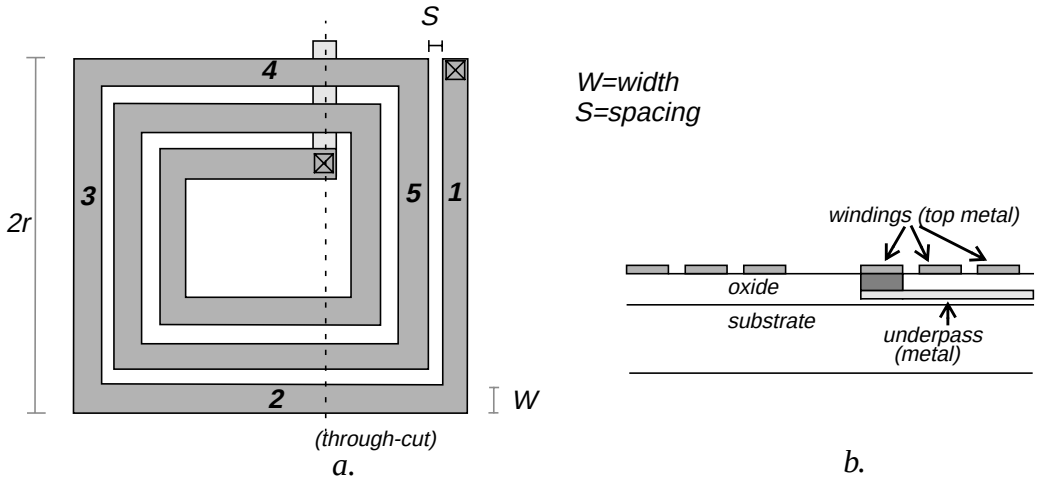


Fig. 4.13. A typical inductor layout, a). top view, b). through-cut view.

contribute to the total inductance, while they add to the parasitics. Therefore, so-called hollow spiral inductors are generally used [84].

The most important non-idealities are loss resistance in the metal wire, capacitance between the metal wire and the substrate, and finite substrate resistance which may cause ohmic losses in the substrate. Moreover, so-called eddy currents may counteract the magnetic field and cause more losses, thus reducing the inductance and the *Q*. The skin effect, i.e. concentration of current in the outer region of the conductor, causes an increasing resistance at high frequencies approximately proportional to $\sqrt{f}$.

Several techniques have been used to increase the *Q* of integrated inductors:

- Etching the substrate [83] - the capacitance and the substrate losses are reduced;
- Placing a patterned grounded shield underneath [85] - substrate losses and the image current are reduced;
- Using copper wires instead of aluminum [81] - the series resistance of the wire is reduced;
- Using shunted multilayer inductors [86] - the series resistance is reduced; However, the capacitances increase.
- Using stacked (series) multi-layer inductors [89], [87] - the mutual inductance is increased and the area for a certain inductance value is decreased; However, the capacitances increase.
- Using an extra thick top metal layer - the series resistance is reduced;
- Adding a BCB (Benzocyclobutene) layer [88] - substrate losses are reduced.

Some of these techniques require extra processing steps, which is undesirable if at all feasible.

In recent years, CMOS processes have developed in order to provide better RF performance, for instance by having an extra thick top metal layer and a more lightly doped substrate. Inductors with *Q* factors as high as 15 may then be designed.

4.4.2. MONOLITHIC INDUCTOR MODELING

For the circuit designer it is desirable to have a simple and accurate lumped-element inductor model, so that circuit simulations will not become too time-consuming. Moreover, a scalable model would give the designer flexibility in inductance value and other properties. However, as it is now, for each technology models are typically only available for a few inductor structures. These models are provided by the IC foundry, not chosen by the designer.

A suitable lumped-element inductor model is shown in Fig. 4.14. Several methods may be used to arrive at this model from the structure shown in Fig. 4.15. Firstly, electromagnetic field solvers may be used to simulate properties of wires or more complicated structures. Simulation time and memory requirements limit the accuracy of the simulation and the complexity of the structures that may be simulated. The parameters of a lumped-element model may then be fitted to the field solver result, or alternatively to measurement results. Secondly, the properties and dimensions of the metal wire, insulating layer and substrate may be used to estimate the model parameters. This will be discussed below.

For the planar rectangular inductor shown in Fig. 4.13, one way to estimate the total inductance L_S is to add the inductances of all segments

$$L_S = \sum_{i=1}^{Z} L_i + \sum M_+ - \sum M_- \qquad (4.38)$$

where L_i is the inductance of segment i, Z is the total number of segments, and $\sum M_+$ and $\sum M_-$ are the total positive and negative mutual inductance between all segments, respectively [82]. L_i (in nH) is given by

$$L_i = 0.2 \cdot l_i \left[\ln\left(\frac{2l_i}{W+d}\right) + 0.50049 + \frac{(W+d)}{3l_i} \right] \qquad (4.39)$$

where l_i is the length of segment i in mm. This is the self-inductance of a metal wire at low frequency, and a relative magnetic permeability μ_r of 1. For the mutual inductance equations can be found in [82]. Other equations to calculate the total inductance have been presented as well (see [30], [82]). The inductance value is dependent on the total wire length l, the spacing s between the windings, the width W of the wire, the thickness d of the metal layer, the number of turns N and the radius r of the inductor.

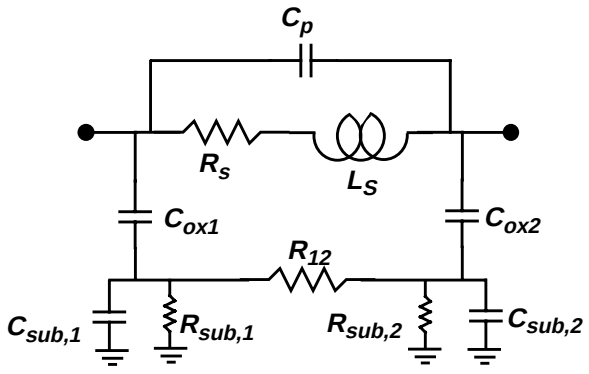


Fig. 4.14. A standard lumped-element inductor model.

Some simple equations will now be given to complete the model in Fig. 4.14. These equations do not take into account all aspects, but give an idea of how material properties and inductor dimensions influence the parasitic elements. For more elaborate inductor modeling see [90], [91], [92] and [93].

The series resistance R_S is the physical resistance of the wire, and is thus dependent on the width *W*, the length *l*, the metal thickness *d*, and the type of metal:

$$R_{S,0} = \frac{l}{\sigma W d} \qquad (4.40)$$

where σ is the conductivity of the metal. If the skin effect is included, this resistance can be expressed as

$$R_S = \frac{l}{2\sigma\delta(W+d)} \qquad (4.41)$$

where the (frequency dependent) skin depth δ is

$$\delta = \sqrt{\frac{1}{\pi f \mu \sigma}} \qquad (4.42)$$

and μ is the magnetic permeability. In Eq. 4.41 it is assumed that $\delta \ll d$ and $\delta \ll W$.

The capacitance C_P from port 1 to port 2 is related to the physical capacitance between the two ports, i.e. the underpass to take out the signal at the inner node (see Fig. 4.13):

$$C_p = (N-1)W^2 \cdot \frac{\varepsilon_{ox}}{t_{ox1}} \qquad (4.43)$$

where t_{ox1} is the oxide thickness between the two metal layers (see Fig. 4.15). The parasitic fringe capacitance [70], which often is larger than the capacitance of the underpass, is neglected here.

The two parasitic capacitances C_{ox1} and C_{ox2} are mainly due to the oxide capacitance between metal and substrate, which may be approximated by

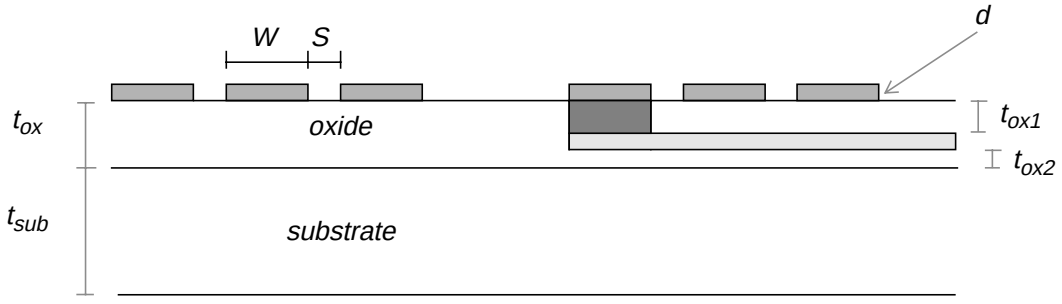


Fig. 4.15. Detailed through-cut view of a planar rectangular inductor.

$$C_{ox,\,tot} = \frac{\varepsilon_{ox}}{t_{ox}} \cdot A_{spiral} \qquad (4.44)$$

and

$$C_{underpass} = \frac{\varepsilon_{ox}}{t_{ox2}} \cdot A_{underpass} \qquad (4.45)$$

where A_{spiral} is the area of the top metal wire, t_{ox} is the oxide thickness between the top metal and the substrate, $A_{underpass}$ is the area of the underpass and t_{ox2} is the oxide thickness between the underpass and the substrate. Thus

$$C_{ox1} = \frac{1}{2}C_{ox,\,tot} + C_{underpass} \text{ and } C_{ox2} = \frac{1}{2}C_{ox,\,tot}. \qquad (4.46)$$

The substrate is modeled as:

$$C_{sub(1,\,2)} = \frac{\varepsilon_{sub}}{t_{sub}} \cdot \frac{1}{2} \cdot A_{spiral} \qquad (4.47)$$

$$R_{sub(1,\,2)} = \rho_{sub} \cdot \frac{2t_{sub}}{A_{spiral}} \qquad (4.48)$$

where ε_{sub} and t_{sub} are the permittivity and the thickness, respectively, of the substrate, and ρ_{sub} is the resistivity of the substrate. If an epitaxial layer is present the substrate parameters should be replaced by the parameters of the epitaxial layer.

The losses due to eddy currents are not part of the model. They can be included by adding a mutual inductance coupled to a resistor, with a value dependent on the inductor geometry and substrate conductivity.

Quality Factor and Self-Resonance

The inductor Q (see for example [85]) is determined by the metal losses and by the substrate losses. The plot of Q_{ind} versus frequency shows a peak at a frequency f_0, which can be tuned to the desired operating frequency, if the inductor is designed with care.

The self-resonance frequency f_{SR} is determined by the total capacitance at the inductor nodes and L_S. Often f_{SR} must be a factor 2 to 3 higher than the operating frequency of the circuit, in order to allow for other parasitic capacitances, or in case of a VCO, to allow for a certain tuning range. This requirement may contradict the desire to have maximum inductor Q at the operating frequency. For applications in the low-GHz range, such as GSM, Bluetooth, UMTS, et cetera, it is hard to get high Q values for inductors larger than 10 nH.

CHAPTER 5

CONCLUSIONS

As a result of the rapid development of mobile communication systems in the low GHz range and the high-frequency capabilities of modern CMOS technology, interest in CMOS transceiver front-ends is increasing, and building blocks that previously were made in other technologies are now more and more realized in CMOS. The level of integration is increasing, and the move towards true systems-on-chip (SoC) is clear. This brings about disturbance issues, as building blocks on the same chip may influence each other. Moreover, the use of high-quality external filters is avoided, which may lead to increased interference both from other transceiver building blocks and from external signals. In this work a brief overview was given of receiver and transmitter architectures, and the impact of system specifications and choice of architecture on building blocks was analyzed.

One of the most challenging building blocks in a CMOS transceiver is the power amplifier. A brief overview of PA issues such as efficiency and linearity was given. Transistor modeling is a critical aspect of RF CMOS design. The drive towards ever smaller devices complicates this, especially for analog, high-frequency and large-signal design. Also, in inductor modeling a more designer-friendly model is desirable.

The papers included in this work relate to many of the above aspects. Paper I, II and III are focused on receiver front-ends. Paper I describes the design of a DCS receiver with LNA, mixer and VCO integrated on the same chip. Paper II addresses the design, implementation and measurement of a low-IF GSM receiver front-end. In paper III an analysis is presented of spurious signals in a receiver and the effect on choice of intermediate frequency (IF).

Paper IV and V both deal with transmitter aspects. In paper IV the design of a class AB power amplifier with power control is presented, and in paper V the design

and measurement of a class C power amplifier with internal frequency doubling and on-chip VCO is described.

FUTURE WORK

The development of mobile communication systems is towards higher frequencies, as well as more broadband systems. On the other hand, old and new communication systems will coexist for quite some time. For CMOS transceiver design this implies a targeting of higher frequencies, facilitated by technology development, and implementation of flexible transceivers so that a user may have one device to communicate in various systems.

The decreasing supply voltage will demand new solutions for well-known problems. This process is, of course, on-going, and includes technology developments, new system specifications, and new design solutions for architectures and building blocks.

Transistor and inductor modeling remain challenging; good and reliable models are a necessity for the RF designer, even though architecture solutions such as feedback loops and control signals may alleviate some problems. However, this increase in complexity does not always come for free.

BIBLIOGRAPHY

[1]. E. Cijvat, *A Study of CMOS Receiver Architectures in Mobile Communication Handsets for GSM*. Tech. Lic. thesis, Royal Institute of Technology, Stockholm, Sweden, 2000.

[2]. J.R. Forest, "Communication Networks for the New Millennium", *IEEE Transactions on Microwave Theory and Techniques*, Vol. 47, No. 12, pp. 2195-2201, December 1999.

[3]. WiMax Forum (September 2004). *IEEE 802.16a Standard and WiMax - Igniting Broadband Wireless Access*. WiMax Forum. [Online] Available: http://www.wimaxforum.org/.

[4]. S. Roy, J.R. Foerster, V.S. Somayazulu and D.G. Leeper, "Ultrawideband Radio Design: The Promise of High-Speed, Short-Range Wireless Connectivity", *Proc. of the IEEE*, Vol. 92, No. 2, pp. 295-311, February 2004.

[5]. C. Nicol and M. Cooke, "Integrated Circuits for 3GPP Mobile Wireless Systems", in *Proc. of the IEEE 2002 Custom Integrated Circuits Conference (CICC)*, pp. 381-388, 2002.

[6]. B. Razavi, *RF Microelectronics*. Prentice Hall PTR, Upper Saddle River, NJ, USA, 1998.

[7]. *GSM Technical Specification 05.05*, European Telecommunications Standards Institute, Sophia Antipolis, France, 1992.

[8]. Bluetooth SIG (2001). *Specification of the Bluetooth System*, version 1.1, Bluetooth SIG, Inc. [Online]. Available: http://www.bluetooth.com.

[9]. 3GPP (1999). *UE Radio Transmission and Reception (FDD), 3G TS 25.101, v3.3.1*. 3rd Generation Partnership Project. [Online] Available: http://www.3gpp.org.

[10]. P. Orsatti, F. Piazza, and Q. Huang, "A 20-mA-Receive, 55-mA-Transmit, Single-Chip GSM Transceiver in 0.25-μm CMOS", *IEEE Journal of Solid State Circuits*, Vol. 34, No. 12, pp. 1869-1880, December 1999.

[11]. 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, December 1995.

[12]. S.M. Fitz, “Receiver Architectures for GSM Handsets”, in *Proceedings of the IEE Colloquium on Design of Digital Cellular Handsets*, UK, 1998, pp. 3/1-10.

[13]. T. Yamawaki, M. Kokubo, K. Irie, H. Matsui, K. Hori, T. Furuya, H. Shimizu, M. Katagishi, T. Endou, B. Henshaw and J.R. Hildersley, “A Dual-Band Transceiver for GSM and DCS1800 Applications”, in *Proceedings of the 1998 European Solid-State Circuits Conference* (ESSCIRC), The Netherlands, 1998, pp. 84-87.

[14]. J. Fenk, "Highly Integrated RF-ICs for GSM and DECT Systems - A Status Review", *IEEE Transactions on Microwave Theory and Techniques*, Vol. 45, No. 12, pp. 2531-2539, December 1997.

[15]. J. Crols and M. Steyaert, *CMOS Wireless Transceiver Design*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1997.

[16]. Murata dielectric filter specifications (2004). Murata Co. [Online]. Available: http://www.murata.com/.

[17]. A.A. Abidi, "Direct-Conversion Radio Transceivers for Digital Communication", *IEEE Journal of Solid State Circuits*, Vol. 30, No. 12, pp. 1399-1410, December 1995.

[18]. B. Razavi, "A 5.2-GHz CMOS Receiver with 62-dB Image Rejection", *IEEE Journal of Solid State Circuits*, Vol. 36, No. 5, pp. 810-815, May 2001.

[19]. S. Tadjpour, E. Cijvat, E. Hegazi and A. Abidi, “A 900 MHz Dual Conversion, Low-IF GSM Receiver in 0.35 um CMOS”, in *Technical Digest of the 2001 International Solid-State Circuits Conference*, USA, 2001, pp. 292-293.

[20]. S. Wu and B. Razavi, “A 900 MHz/1.8 GHz CMOS Receiver for Dual-Band Applications”, *IEEE Journal of Solid State Circuits*, Vol. 33, No. 12, pp. 2178-2185, December 1998.

[21]. 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, June 2001.

[22]. P.Y. Chan, A. Rofougaran, K.A. Ahmed and A.A. Abidi, "A Highly Linear 1-GHz CMOS Downconversion Mixer", in *Proceedings of the 1993 European Solid-State Circuits Conference* (ESSCIRC), Spain, 1993.

[23]. E. Cijvat, P. Eriksson, N. Tan and H. Tenhunen, "A 1.8 GHz Subsampling CMOS Downconversion Circuit for Integrated Radio Circuits", in *Proceedings of the 1998 IEEE International Symposium on Circuits and Systems* (ISCAS), USA, 1998, pp. II-65-68.

[24]. R.G. Vaughan, N.L. Scott and D.R. White, “The Theory of Bandpass Sampling”, *IEEE Transactions on Signal Processing*, Vol. 39 No. 9, pp. 1973-1984, September 1991.

[25]. N.D. Faulkner and E.V.I. Mestre, “Subharmonic Sampling for the Measurement of Short-term Stability of Microwave Oscillators”, *IEEE Transactions on Instrumentation and Measurement*, Vol. IM-32 No.1, pp. 208-213, March 1983.

[26]. D.H. Shen, C.-M. Hwang, B.B. Lusignan and B.A. Wooley, "A 900-MHz RF Front-End with Integrated Discrete-Time Filtering", *IEEE Journal of Solid State Circuits*, Vol. 31, No. 12, pp. 1945-1954, December 1996.

[27]. D.B. Leeson, "A Simple Model of Feedback Oscillator Noise Spectrum", *Proceedings of the IEEE*, Vol. 54, No. 2, pp. 329-330, February 1966.

[28]. P. Eriksson and H. Tenhunen, "Phase-Noise in Sampling and its Importance to Wideband Multicarrier Base Station Receivers" in *Proceedings of the 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing* (ICASSP), USA, 1999, pp. 2737-2740.

[29]. A. van der Ziel, *Noise in Solid State Devices and Circuits*. John Wiley & Sons, Inc., New York, NY, 1986.

[30]. T.H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*. Cambridge University Press, Cambridge, UK, 1998.

[31]. A. Springer, L. Maurer and R. Weigel, "RF System Concepts for Highly Integrated RFICs for W-CDMA Mobile Radio Terminals", *IEEE Transactions on Microwave Theory and Techniques*, Vol. 50, No. 1, pp. 254-267, January 2002.

[32]. O.K. Jensen, T.E. Kolding, C.R. Iversen, S. Laursen, R.V. Reynisson, J.H. Mikkelsen, E. Pedersen, M.B. Jenner and T. Larsen, "RF Receiver Requirements for 3G W-CDMA Mobile Equipment", *Microwave Journal*, Vol. 43, No. 2, February 2000.

[33]. K. Murota and K. Hirade, "GMSK Modulation for Mobile Radio Telephony", *IEEE Transactions on Communications*, Vol. com-29, No. 7, pp. 1044-1050, July 1981.

[34]. C. Samori, A.L. Lacaita, A. Zanchi and P.Vita, “Design Issues of LC Tuned Oscillators for Integrated Transceivers”, in *Proceedings of the 8th Great Lakes Symposium on VLSI*, USA, 1998, pp. 264-269.

[35]. R.C. Sagers, "Intercept Point and Undesired Responses", *IEEE Transactions on Vehicular Technology*, Vol. vt-32, No. 1, pp. 121-133, February 1983.

[36]. Z.-Q. Lang and S. Billings, "Evaluation of Output Frequency Responses of nonlinear Systems Under Multiple Inputs", *IEEE Transactions on Circuits and Systems-II: Analog and Digital Signal Processing*, Vol. 47, No. 1, pp. 28-38, January 2000.

[37]. J.C. Pedro and N. Borges de Carvalho, "On the Use of Multitone Techniques for Assessing RF Components’ Intermodulation Distortion", *IEEE Transactions on Microwave Theory and Techniques*, Vol. 47, No. 12, pp. 2393-2402, December 1999.

[38]. B. Razavi, "A Study of Injection Locking and Pulling in Oscillators", *IEEE Journal of Solid State Circuits*, Vol. 39, No. 9, pp. 1415-1424, September 2004.

[39]. B. Razavi, "RF Transmitter Architectures and Circuits", in *Proc. of IEEE 1999 Custom Integrated Circuits Conference (CICC 1999)*, USA, 1999, pp. 197-204.

[40]. E. Hegazi and A.A. Abidi, "A 17-mW Transmitter and Frequency Synthesizer for 900-MHz GSM Fully Integrated in 0.35-um CMOS", *IEEE Journal of Solid State Circuits*, Vol. 38, No. 5, pp. 782-792, May 2003.

[41]. W.T. Bax and M.A. Copeland, "A GMSK Modulator Using a DS Frequency Discriminator-Based Synthesizer", *IEEE Journal of Solid State Circuits*, Vol. 36, No. 8, pp. 1218-1227, August 2001.

[42]. H. Ishikuro, M. Hamada, K-I Agawa, S. Kousai, H. Kobayashi, D. Minh Nguyen, and F. Hatori, " A Single-Chip CMOS Bluetooth Transceiver with 1.5MHz IF and Direct Modulation", in *2003 IEEE International Solid-State Circuits Conference (ISSCC) Digest of Technical Papers*, USA, 2003, pp. 94-95.

[43]. G. Irvine, S. Herzinger, R. Schmidt, D. Kubetzko, and J. Fenk, "An Up-Conversion Loop Transmitter IC for Digital Mobile Telephones", in *1998 IEEE International Solid- State Circuits Conference (ISSCC) Digest of Technical Papers*, USA, 1998, pp. 364-365.

[44]. M. Cohn, J.E. Degenford and B.A. Newman, “Harmonic Mixing with an Antiparallel Diode Pair”, *IEEE Transactions on Microwave Theory and Techniques*, Vol. MTT-23, No. 8, pp. 667-673, August 1975.

[45]. T.F. McMaster, M.V. Schneider, and W.W. Snell, "Millimeter-Wave Receivers with Subharmonic Pump", *IEEE Transactions on Microwave Theory and Techniques*, Vol. MTT-24, No. 12, pp. 948-952, December 1976.

[46]. K. Itoh, T. Yamaguchi, T. Katsura, K. Sadahiro, T. Ikushima, R. Hayashi, F. Ishizu, E. Taniguchi, T. Nishino, M. Shimozawa, N. Suematsu, T. Tagaki and O. Ishida, “Integrated Even Harmonic Type Direct Conversion Receiver for W-CDMA Mobile Terminals”, *Proc. of 2002 IEEE Radio Frequency Integrated Circuits Symposium*, USA, 2002, pp. 263-266.

[47]. M.Cohn, R.G. Freitag, H. G. Henry, J.E. Degenford and D.A. Blackwell, "A 94 GHz MMIC Tripler using Anti-Parallel Diode Arrays for Idler Separation", in *1994 IEEE MTT-S Digest*, 1994, pp. 763-766.

[48]. R. Martin and F. Ali, “A Ku-Band Oscillator Subsystem Using a Broadband GaAs MMIC Push-Pull Amplifier/Doubler”, *IEEE Microwave and Guided Wave Letters*, Vol. 1 no. 11, pp. 348-350, November 1991.

[49]. H. Kondoh and A. Cognata, “A 20-50GHz MMIC Amplifier with 21dBm Output Power and its Application as a Frequency Doubler”, *Proc. of IEEE 1993 Microwave and Millimeter-Wave Monolithic Circuits Symposium*, 1993, pp. 35-38.

[50]. M. Zeng, A. Annamalai, and V.K. Bhargava, "Recent Advances in Cellular Wireless Communications", *IEEE Communications Magazine*, pp. 128-138, September 1999.

[51]. P. Madsen, O.K. Jensen, T. Amtoft, R.V. Reynisson, J.H. Mikkelsen, S. Laursen, T.E. Kolding, T. Larsen, and M.B. Jenner, "UTRA/FDD RF Transceiver Requirements", *Wireless Personal Communications*, Vol. 23 no. 1, pp. 55-66, October 2002.

[52]. G. Brenna, D. Tschopp, J. Rogin, I. Kouchev, and Q. Huang, "A 2-GHz Carrier Leakage Calibrated Direct-Conversion WCDMA Transmitter in 0.13-µm CMOS", *IEEE Journal of Solid State Circuits*, Vol. 39, No. 8, pp. 1253-1262, August 2004.

[53]. S.C. Cripps, *RF Power Amplifiers for Wireless Communication*. Artech House, Norwood, MA, USA, 1999.

[54]. R.S. Narayanaswami, *RF CMOS Class C Power Amplifiers for Wireless Communications*. PhD thesis, University of California, Berkeley, CA, USA, 2001.

[55]. C. Fallesen, *Design Techniques for Sub-Micron RF Power Amplifiers*. PhD thesis, Technical University of Denmark, Lyngby, Danmark, 2001.

[56]. M.M. Hella, *RF CMOS Power Amplifiers - Theory, Design and Implementation*. Kluwer Academic Publishers, Norwell, MA, USA, 2002.

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

[58]. F.H. Raab, "Class-F Power Amplifiers with Maximally Flat Waveforms", *IEEE Transactions on Microwave Theory and Techniques*, Vol. 45, No. 11, pp. 2007-2112, November 1997.

[59]. A. Jayaraman, P.F. Chen, G. Hanington, L. Larson and P. Asbeck, "Linear High-Efficiency Microwave Power Amplifiers Using Bandpass Delta-Sigma Modulators", *IEEE Microwave and Guided Wave Letters*, Vol. 8, No. 3, pp. 121-123, March 1998.

[60]. L.R. Kahn, "Single Sideband Transmission by Envelope Elimination and Restoration", *Proc. IRE*, Vol. 40, pp. 803-806, July 1952.

[61]. F.H. Raab and D.J. Rupp, "High-Efficiency Single-Sideband HF/VHF Transmitter Based upon Envelope Elimination and Restoration", in *Proc. of*

the 1994 IEE Conference on HF Radio Systems and Techniques, UK, 1994, pp. 21-25.

[62]. F.H. Raab, "Efficiency of Outphasing RF Power-Amplifier Systems", *IEEE Transactions on Communications*, Vol. com-33, No. 10, pp. 1094-1099, October 1985.

[63]. D.C. Cox, "Linear Amplification with Non-Linear Components", *IEEE Transactions on Communications*, Vol. 22, pp. 1942-1945, December 1974.

[64]. K.-Y. Chan and A. Bateman, "Linear Modulators Based on RF Synthesis: Realization and Analysis", *IEEE Transactions on Circuits and Systems - I: Fundamental Theory and Applications*, Vol. 42, No. 6, pp. 321-333, June 1995.

[65]. D. Webster, J. Scott and D. Haigh, "Control of Circuit Distortion by the Derivative Superposition Method", *IEEE Microwave and Guided Wave Letters*, Vol. 6 no. 3, pp. 123-125, March 1996.

[66]. C.S. Aitchison, M. Mbabele, M.R. Moazzam, D. Budimir, and F. Ali, "Improvement of Third-Order Intermodulation Product of RF and Microwave Amplifiers by Injection", *IEEE Transactions on Microwave Theory and Techniques*, Vol. 49 No. 6, pp. 1148-1154, June 2001.

[67]. R. Gharpurey, *Modeling and Analysis of Substrate Coupling in Integrated Circuits*. Ph.D. Thesis, University of California, Berkeley, CA, USA, 1995.

[68]. X. Aragonès, J.L.Gonzáles and A. Rubio, *Analysis and Solutions for Switching Noise Coupling in Mixed-Signal ICs*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1999.

[69]. S.M. Sze, *VLSI Technology*, 2nd ed. McGraw-Hill, Singapore, 1988.

[70]. B. Razavi, *Design of Analog CMOS Integrated Circuits*. McGraw-Hill, Singapore, 2000.

[71]. Y. Tsividis, *Operation and Modeling of the MOS Transistor*, 2nd ed. McGraw-Hill, Singapore, 1999.

[72]. T. Manku, "Microwave CMOS - Device Physics and Design", *IEEE Journal of Solid State Circuits*, Vol. 34, No. 3, pp. 277-285, March 1999.

[73]. C.C. Enz and Y. Cheng, "MOS Transistor Modeling for RF IC Design", *IEEE Journal of Solid State Circuits*, Vol. 35, No. 2, pp. 186-201, February 2000.

[74]. L.E. Larson, "Integrated Circuit Technology Options for RFIC's - Present Status and Future Directions", *IEEE Journal of Solid State Circuits*, Vol. 33, No. 3, pp. 387-399, March 1998.

[75]. M. Feng, S.-C. Shen, D.C. Caruth and J.-J. Huang, "Device Technologies for RF Front-End Circuits in Next-Generation Wireless Communications", *Proceedings of the IEEE*, Vol. 92, No. 2, pp. 354-374, February 2004.

[76]. B. Hoeneisen and C.A. Mead, "Fundamental Limitations in Microelectronics - I - MOS Technology", *Solid-State Electronics*, Vol. 15, pp. 819-829, 1972.

[77]. H. Iwai, "CMOS Technology - Year 2010 and Beyond", *IEEE Journal of Solid State Circuits*, Vol. 34, No. 3, pp. 357-366, March 1999.

[78]. R.H. Dennard, F.H. Gaensslen, H.-N. Yu, V.L. Rideout, E. Bassous and A.R. LeBlanc, "Design of Ion-Implanted MOSFET's with Very Small Physical Dimensions", *IEEE Journal of Solid State Circuits*, Vol. 9, pp. 256-268, 1974.

[79]. A.A. Abidi, "High-Frequency Noise Measurements on FET's with Small Dimensions", *IEEE Transactions on Electron Devices*, Vol. ed-33, No. 11, pp. 1801-1805, November 1986.

[80]. P. Andreani and H. Sjöland, "Noise Optimization of an Inductively Degenerated CMOS Low Noise Amplifier", *IEEE Transactions on Circuits and Systems - II: Analog and digital Signal Processing*, Vol. 48, No. 9, pp. 835 - 841, September 2001.

[81]. J.N. Burghartz, D.C. Edelstein, M. Suoyer, H.A. Ainspan and K.A. Jenkins, "RF Circuit Design Aspects of Spiral Inductors on Silicon", *IEEE Journal of Solid State Circuits*, Vol. 33, No. 12, pp. 2028-2034, December 1998.

[82]. H.M. Greenhouse, "Design of Planar Rectangular Microelectronic Inductors", *IEEE Transactions on Parts, Hybrids, and Packaging*, Vol. php-10, No. 2, pp. 101-109, June 1974.

[83]. 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, pp. 246-248, May 1993.

[84]. J. Craninckx and M.S.J. Steyaert, "A 1.8-GHz Low-Phase-Noise CMOS VCO Using Optimized Hollow Spiral Inductors", *IEEE Journal of Solid State Circuits*, Vol. 32, No. 5, pp. 736-744, May 1997.

[85]. C.P. Yue and S. Wong, "On-Chip Spiral Inductors with Patterned Ground Shields for Si-Based RF IC's", *IEEE Journal of Solid State Circuits*, Vol. 33, No. 5, pp. 743-752, May 1998.

[86]. M. Soyuer, J.N. Burghartz, K.A. Jenkins, S. Ponnapalli, J.F. Ewen and W.E. Pence, "Multilevel Monolithic Inductors in Silicon Technology", *Electronics Letters*, Vol. 31, No. 5, pp. 359-360, March 1995.

[87]. A. Zolfaghari, A. Chan and B. Razavi, "Stacked Inductors and Transformers in CMOS Technology", *IEEE Journal of Solid State Circuits*, Vol. 36, No. 4, pp. 620-628, April 2001.

[88]. K.J. Chen, X. Huo, L.L.W. Leung and P.C.H. Chan, "High-Performance Microwave Passive Components on Silicon Substrate", in *Proc. of the IEEE 2002 3rd International Conference on Microwave and Millimeter Wave Technology*, 2002, pp. 263-266.

[89]. R.B. Merrill, T.W. Lee, H. You, R. Rasmussen and L.A. Moberly, "Optimization of High *Q* Integrated Inductors for Multi-Level Metal CMOS", in *Technical Digest of the 1995 International Electron Devices Meeting* (IEDM), USA, 1995, pp. 983-986.

[90]. A.M. Niknejad and R.G. Meyer, "Analysis, Design and Optimization of Spiral Inductors and Transformers for Si RF IC's", *IEEE Journal of Solid State Circuits*, Vol. 33, No. 10, pp. 1470-1481, October 1998.

[91]. C.P. Yue, C. Ryu, J. Lau, T.H. Lee and S.S. Wong, "A Physical Model for Planar Spiral Inductors on Silicon", in *Technical Digest of the 1996 International Electron Devices Meeting* (IEDM), USA, 1996, pp. 155-158.

[92]. N. Troedsson and H. Sjöland, "A Distributed Capacitance Analysis of Co-Planar Inductors for a CMOS QVCO with Varactor Tuned Buffer Stage", accepted for publication in *Analog Integrated Circuits and Signal Processing*.

[93]. J.R. Long and M.A. Copeland, "The Modeling, Characterization, and Design of Monolithic Inductors for Silicon RF IC's", *IEEE Journal of Solid-State Circuits*, Vol. 32, No. 3, pp. 357-369, March 1997.

Part II

Included Papers

SUMMARY OF INCLUDED PAPERS

In this section a short summary will be given of each paper, and the author's contribution to each paper is listed.

I. In this paper the design of a CMOS receiver front-end is presented that meets the DCS requirements. It includes a low-noise amplifier, a mixer and a fully integrated voltage controlled oscillator.

   Contribution: Design, manuscript.

II. This paper addresses the design and measurement of a fully integrated GSM receiver front-end, which performs image rejection, channel selection, and resists GSM blocking signals. a controllable gain of more than 100 dB is achieved.

   Contribution: Design of some blocks (PGA, 2nd downconversion mixer), some analysis (interfering signals).

III. An analysis is made of the impact of spurious signals in a receiver. It leads to an optimal choice of intermediate frequency for a fully integrated receiver, and is supported by measurement results on a GSM receiver front-end.

   Contribution: Analysis, most of the manuscript.

IV. The design of a fully integrated CMOS power amplifier is described. The amplifier uses parallel output stages in order to optimize the efficiency for both high and low output power levels.

   Contribution: Design, manuscript.

V. In this paper the design and measurement results are presented of a fully integrated CMOS power amplifier, suitable for direct-conversion transmitters or low-IF upconversion. The amplifier uses internal frequency doubling.

   Contribution: Design, measurements, manuscript.

PAPER I

PAPER I

A 0.35 μm CMOS DCS Front-end with Fully Integrated VCO

Abstract

The design of a 0.35 μm CMOS front-end is presented which meets the DCS requirements. The front-end consists of a low-noise amplifier operating at 1.85 GHz, a mixer downconverting the RF signal to an IF of up to 250 MHz, and a fully integrated voltage controlled oscillator with a tuning range of 460 MHz. The total power consumption is 34 mW with a 2.5V supply, and the simulated maximum RF to IF gain is 26.6 dB.

Based on: E. Cijvat, "A 0.35 μm CMOS DCS Front-end with Fully Integrated VCO". © 2004 IEEE. Reprinted, with permission, from *Proceedings of the 8th IEEE International Conference on Electronics, Circuits and Systems (ICECS) 2001*, pp. 1595 - 1598. Malta, September 2001.

1. Introduction

Recently the market for wireless applications has shown tremendous growth, resulting in an increase in research and development of low-cost RF integrated circuits. It has been shown previously [1 - 4] that CMOS has the potential to be used for low-GHz cellular communication applications. The advantage of CMOS compared to other technologies, such as for instance silicon bipolar (Si BJT) or III - V based technologies, is that a higher level of integration can be achieved by having the radio frequency (RF), intermediate frequency (IF) and baseband part of a transceiver on a single chip.

An important aspect for a handset is battery lifetime. To extend this, overall power consumption must be minimized. Since external components weigh heavily on the costs of a handset, one of the goals of this design is to reduce the number of external components.

In this paper a CMOS downconversion front-end is described that meets the DCS-1800 specifications [5] for a handset receiver, which are to a large extent similar to the ones for GSM-900. Differences are for example the frequency band of operation (1805 to 1880 MHz for a DCS handset receiver) and details in the blocking specification.

The front-end comprises a low-noise amplifier (LNA), a downconversion mixer and a fully integrated CMOS voltage controlled oscillator (VCO). An 1850 MHz LNA with switchable gain is described, as well as a Gilbert-cell type mixer that downconverts the RF (radio frequency) signal to an IF (intermediate frequency) of up to 250 MHz. The third block described is a VCO with a tuning range of 460 MHz, which provides the local oscillator (LO) signal to the mixer. According to simulations this VCO meets the DCS phase noise requirements. A 0.35 μm BiCMOS technology, of which only CMOS transistors are used, is utilized for the design.

This paper is organized as follows: in Section 2 system aspects will be explained briefly, as well as their implications for the front-end subcircuits. In Section 3 the design of the LNA, mixer and VCO will be discussed. Simulation results will be presented in Section 4, and Section 5 contains the conclusions.

2. System Aspects

The channel bandwidth for the DCS system is 200 kHz. With direct downconversion, the *1/f* noise of the mixer would impair the signal-to-noise ratio (SNR). With dual downconversion, more gain can be implemented before downconversion to baseband, alleviating this problem. A second advantage of dual downconversion is

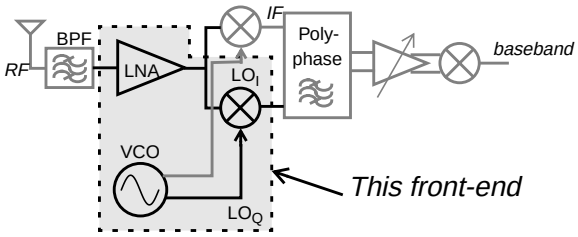


Fig. 1. Receiver architecture

that it facilitates gain distribution over the full receiver chain [6], so that noise and linearity can be better optimized and stability problems can be reduced.

To reduce the number of external components, no filter is used between the LNA and mixer. This implies that neither the output impedance of the LNA or the input impedance of the mixer needs to be matched to the impedance of a discrete filter. A discrete band pass filter (BPF) following the antenna is assumed to be present.

To achieve sufficient image rejection a quadrature mixer in combination with a polyphase filter can be used [1], as indicated in Fig. 1. The VCO quadrature signal can be generated by using a polyphase network [2, 3] as well.

Since the DCS noise requirements are stringent, high gain in the RF part of the receiver is important. This includes the mixer, which therefore should be active. However, a high gain in the RF part can compromise the total linearity of the receiver. System simulations have shown that a switchable LNA gain (5 or 20 dB gain) and a limited mixer gain (5 dB) are sufficient for the receiver to meet the DCS noise and linearity requirements, assuming reasonable noise and linearity performance of subsequent blocks.

High-side injection is chosen for the downconversion, so that the VCO must be tunable from roughly 1.95 to 2.15 GHz, leaving some flexibility to choose the IF.

The choice for high-side injection is based on two arguments. Firstly, if a GSM/DCS dual-mode receiver is designed, the VCO can be used for the downconversion of the GSM signal as well by increasing the tuning range beyond 2.15 GHz and using a divide-by-2 circuit. An advantage of this strategy is that the phase noise performance of the divided signal will improve with approximately 6 dB [6]. Secondly, since no filter is present between the LNA and mixer, interferers will not be suppressed except by the discrete BPF and the roll-off of the LNA. Thus the interferers may be downconverted with harmonics of the LO, possibly distorting the desired signal. High-side injection moves these LO harmonics to higher frequencies.

The front-end was modeled to analyse the effect of interferers and harmonic downconversion [7], taking into account the DCS interferer (blocker) spectrum, BPF

and LNA roll-off, building block linearity, LO harmonics etc. It was found that for an IF of 235 - 250 MHz, using high-side injection, the desired signal will not be distorted significantly, i.e. the required bit error rate (BER) for DCS will be met.

3. Circuit Design

3.1. The LNA

The LNA consists of a single-ended, cascoded, inductively degenerated common-source stage, with an off-chip input tuning network to achieve a 50 Ω input impedance [3, 4]. Parasitics at the input node (originating from bonding wire, pad, pin etc.) are included in this network. Assuming certain boundary conditions a trade-off exists between the noise figure and linearity of the LNA with respect to the quality factor Q of this network, since a high Q yields a better noise figure but decreases the IIP_3 assuming equal biasing conditions.

The LNA is loaded with an integrated inductor, tuning the maximum gain to 1.85 GHz. The parallel capacitor in Fig. 2 is formed by parasitic capacitances at the output node. A voltage gain of around 20 dB is pursued. A resistor, controlled by a switch, can be used to change the gain to a lower setting of about 5 dB (see Fig. 2).

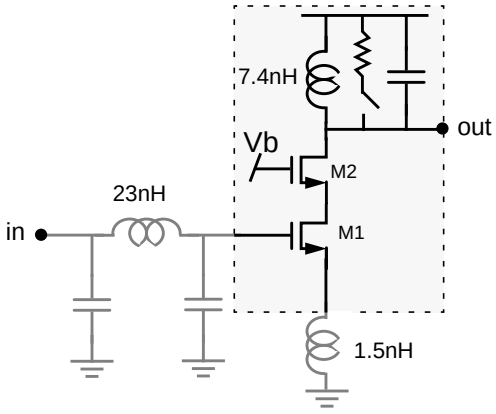


Fig. 2. Schematic of the 1.85 GHz LNA

Two parallel bondwires at the source of M1 provide inductive source degeneration.

3.2. The mixer

Two mixers have been compared: a single-balanced and a double-balanced Gilbert-cell type mixer. The main advantage of a double-balanced mixer is a first-order cancellation of LO feedthrough. Depending on the sensitivity of the IF building blocks for an undesired signal at f_{LO}, a double-balanced mixer - with one input AC grounded - may be preferable over a single-balanced mixer.

The main advantage of the single-balanced mixer is that a lower power consumption is needed to achieve similar gain, linearity and noise performance.

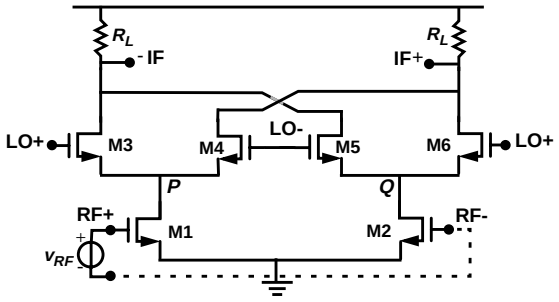


Fig. 3. Schematic of the double-balanced mixer

However, since no filter is present after the mixer, a double-balanced mixer is utilized in this front-end (see Fig. 3).

3.3. The VCO

The VCO consists of a differential switching NMOS pair with a tail current source (see Fig. 4). Issues of *1/f* noise vs. thermal noise, while considering the DCS phase noise requirements, have guided the choice for this architecture. The VCO is AC coupled to a buffer, which drives the LO ports of the mixer and provides some isolation between the VCO core and the mixer.

The inductive load is a differential octagonal inductor with a polysilicon grounded shield used to prevent the occurrence of image currents in the substrate [8]. With the available technology and careful layout a quality factor as high as 10 can be achieved.

Tuning is achieved partly by a varactor, partly by switch-controlled capacitors, the latter for discrete steps in carrier frequency. Thus the relative amount of varactor capacitance is kept small, so that the degradation in phase noise performance, due to degradation in quality factor of the total capacitance, will be small as well. The total tuning range is 460 MHz utilizing 4 discrete steps.

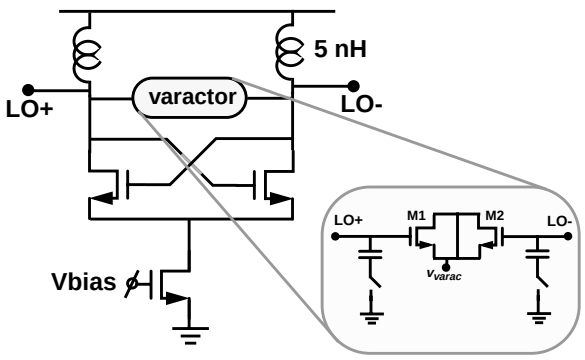


Fig. 4. Schematic of the VCO

4. Simulation Results

Both the complete front-end and the individual building blocks have been simulated using SpectreRF. With a mixer IIP_3 of 10 dBm (see Fig. 5), an LNA IIP_3 of 0 dBm and 21 dB of voltage gain preceding the mixer, the front-end IIP_3 is expected to be somewhat below -11 dBm. This was confirmed in front-end simulations.

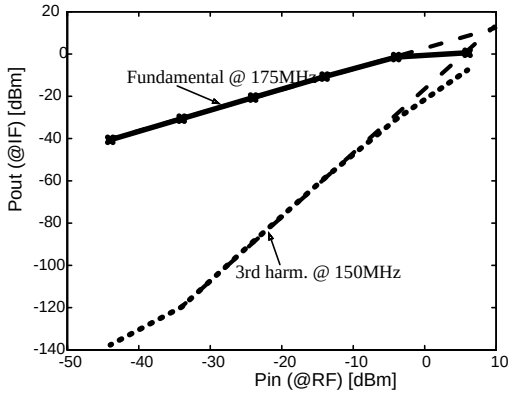


Fig. 5. Simulated IIP_3 of double-balanced mixer

Assuming the BPF has an insertion loss of 2.5 dB, this will result in a total front-end IIP_3 of -6.3 dBm. Using the low-gain LNA setting will result in a higher total IIP_3.

The VCO meets the phase noise requirements for the DCS system, as is illustrated in Fig. 6.

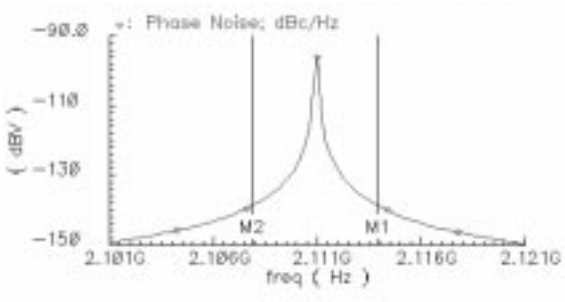


Fig. 6. VCO phase noise, with markers indicating the critical 3MHz frequency offset points

The maximum gain for the total front-end is 26.6 dB. Some loss occurs due to the capacitive coupling between the LNA and mixer. The total power consumption is 34 mW. The results for the front-end building blocks are summarized in table 1.

A layout of the front-end (see Fig. 7) will be sent to fabrication. For the purpose of measurements an open-drain buffer is added at the output of the mixer.

TABLE 1. FRONT-END SIMULATION RESULTS

LNA	gain (voltage) NF IIP_3 S_{11} power cons.	22/8 dB 1.8 dB (high gain) 0 dBm (high gain) < -12.7dB 8.8 mW
mixer	gain (voltage) NF (SSB) IIP_3 power cons.	5.6 dB 14 dB 9.9 dBm 10.5 mW
VCO	tuning range phase noise power cons.	460 MHz (1940-2400) -138.6 dBc/Hz @3MHz 14.8 mW (incl. buffer)

.

5. Conclusion

A 0.35 µm CMOS front-end has been described, meeting the requirements for a DCS receiver according to simulations. The front-end consists of a low-noise amplifier operating at 1.85 GHz, a mixer that downconverts the RF signal to an intermediate frequency of up to 250 MHz, and a voltage controlled oscillator with a tuning range of 460 MHz. Moreover, the VCO is fully integrated and meets the DCS phase noise requirements.

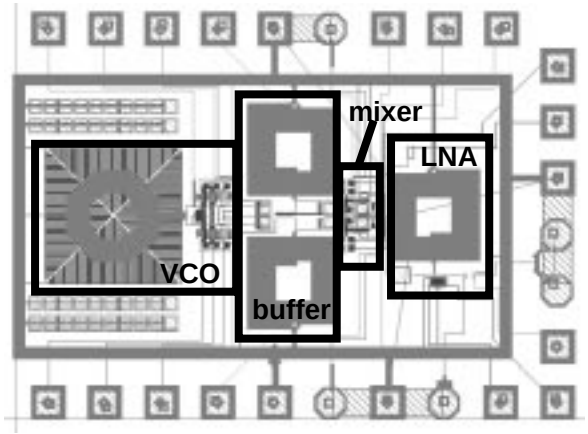


Fig. 7. Layout of the front-end

This front-end does not entail a - generally discrete - filter between LNA and mixer. Thus, the number of external components is reduced while dual downconversion is utilized.

ACKNOWLEDGEMENTS

The help of Shahrzad Tadjpour, Emad Hegazi and Jaesup Lee (University of California, Los Angeles, USA) and Henrik Sjöland (Lund University, Sweden) is greatly appreciated. The author is indebted to the Dept. of Microelectronics and Information Technology, Royal Institute of Technology (Stockholm, Sweden) for their financial support.

REFERENCES

[1]. F. Behbahani, Y. Kishigami, J. Leete and A.A. Abidi, “CMOS 10 MHz-IF downconverter with on-chip broadband circuit for large image-suppression”, in *Digest of Technical Papers of the 1999 Symposium on VLSI Circuits*, Kyoto, Japan, June 1999, pp. 83-6.

[2]. M. Steyaert, J. Janssen, B. De Muer, M. Borremans and N. Itoh, “A 2V CMOS Cellular Transceiver Front-End”, *IEEE Journal of Solid State Circuits*, Vol. 35, No. 12, December 2000, pp. 1895-1907.

[3]. S. Tadjpour, E. Cijvat, E. Hegazi and A.A Abidi, “A 900 MHz Dual Conversion, Low-IF GSM Receiver in 0.35-µm CMOS”, in *Technical Digest of the 2001 IEEE Solid State Circuits Conference (ISSCC)*, San Francisco, CA, February 2001, pp. 292-3.

[4]. P. Orsatti, F. Piazza, and Q. Huang, “A 20-mA-Receive, 55-mA-Transmit, Single-Chip GSM Transceiver in 0.25-µm CMOS”, *IEEE Journal of Solid State Circuits*, Vol. 34, No. 12, December 1999, pp. 1869-80.

[5]. *GSM Technical Specification 05.05*, European Telecommunications Standards Institute, Sophia Antipolis, France, 1992.

[6]. B. Razavi, *RF Microelectronics*, Prentice Hall PTR, Upper Saddle River, NJ, 1998.

[7]. R.C. Sagers, “Intercept Point and Undesired Responses”, *IEEE Transactions on Vehicular Technology*, Vol. VT-32, No. 1, February 1983, pp. 121-33.

[8]. C.P. Yue and S. Wong, “On-Chip Spiral Inductors with Patterned Ground Shields for Si-Based RF IC’s”, *IEEE Journal of Solid State Circuits*, Vol. 33, No. 5, May 1998, pp. 743-52.

PAPER II

PAPER II

A 900 MHz Dual Conversion, Low-IF GSM Receiver in 0.35 um CMOS

Abstract

A low-power fully integrated GSM receiver is developed in 0.35-um CMOS. This receiver uses dual conversion with a low IF of 140 kHz. This arrangement lessens the impact of the flicker noise. The first IF of 190 MHz best tolerates blocking signals. The receiver includes all of circuits for analog channel selection, image rejection, and more than 100-dB controllable gain. The receiver alone consumes 22 mA from a 2.5-V supply, to give a noise figure of 5 dB, and input IP3 of -16 dBm. A single frequency synthesizer generates both LO frequencies. The integrated VCO with on-chip resonator and buffers consume another 8 mA, and meet GSM phase-noise specifications.

Based on: S. Tadjpour, E. Cijvat, E. Hegazi and A. Abidi, "A 900 MHz Dual Conversion, Low-IF GSM Receiver in 0.35 μm CMOS". © 2004 IEEE. Reprinted, with permission, from *IEEE Journal of Solid-State Circuits,* Vol. 36, No.12, pp. 1992 - 2002, Dec. 2001.

1. Introduction

Mobile wireless transceivers face relentless pressure for low cost, low power, and small size. A highly integrated CMOS realization is therefore of interest. It is generally ac–cepted that RF-CMOS circuits are good for short-range communicators such as DECT or Bluetooth, but whether CMOS can ever deliver a fully integrated receiver with competitive performance for a challenging cellular standard such as GSM continues to be debated. Previous bipolar or CMOS GSM transceivers use off-chip components to reject the image and large blocking signals [1]-[3], or defer these tasks to the baseband sections [4].

This paper describes a fully integrated CMOS GSM receiver, which we believe is the more challenging half of the transceiver. The circuit includes on-chip image rejection and channel selection filter. It meets the requirements of GSM, yet consumes a power comparable to the lowest numbers reported in prior art. Section II describes the GSM system requirements and briefly discusses three different receiver architectures. Section III explains the proposed architecture and the rationale behind it. The circuit design is covered in detail in Section IV, and finally, the experimental results and the layout issues are presented in Section V.

2. System Description

The 900-MHz GSM standard [5] requires receive sensitivity of 102 dBm for a GMSK modulated signal that occupies a 200-kHz-wide bandwidth. To obtain a signal-to-noise ratio (SNR) of 9 dB at the demodulator at the target bit-error rate (BER), the noise figure (NF) at the antenna output must be at most 10 dB. Assuming 3-dB insertion loss, and therefore NF, for the RF bandpass filter prior to the receiver input, this yields an upper limit of 7 dB for NF of the receiver itself.

What distinguishes GSM from other well-known cellular standards is the required receiver immunity to very large in-band signals, called blockers. Fig. 1 shows the blocking profile for the GSM standard. The receiver must function at the target BER when the wanted channel is 3 dB above sensitivity level, and is accompanied by blockers as strong as 100 dB larger. Receiver LO phase noise, image rejection, and 1-dB compression point all may be deduced from the blocking profile. To pass the intermodulation test, a GSM signal 3 dB above the sensitivity level must be detected by the receiver in the simultaneous presence of two blockers: a static sine-wave, and a GMSK modulated signal at -49 dBm. With 9-dB SNR and a 3-dB margin, this means that receiver IIP3 must be at least -21 dBm. At the preselect filter input, this corresponds to an IIP3 of about -18 dBm.

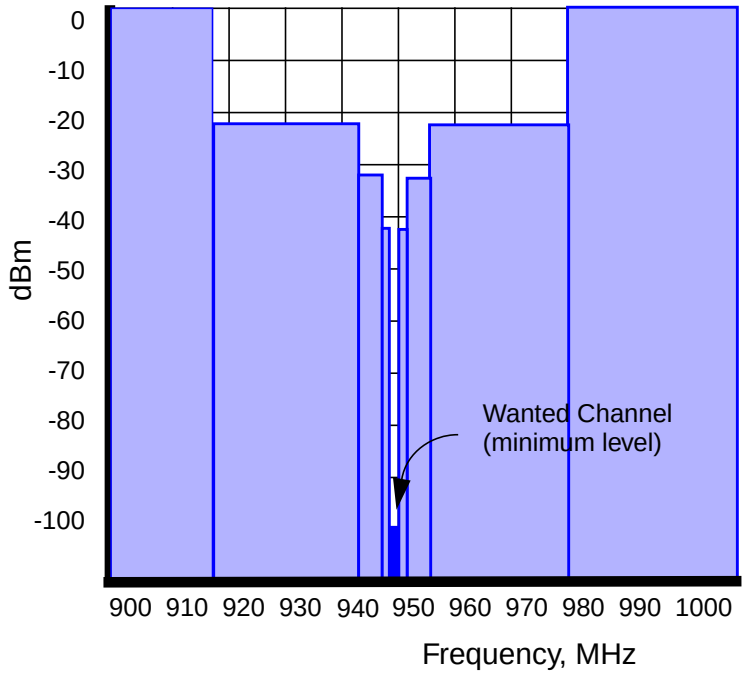


Fig. 1. GSM blocker template.

Any assessment of candidate architectures for integrated receivers should begin with direct conversion to zero IF. The direct conversion receiver eliminates image reject and off-chip IF filters. Instead, two integrated active low-pass filters in *I* and *Q* paths select the wanted channel. However, the narrow 200-kHz GMSK-modulated GSM channel at zero IF contains significant energy at very low frequency, and therefore cannot tolerate LO leakage, dc offset or second-order distortion products [6] without elaborate means to sense and remove these artifacts prior to demodulation.

Alternatively, downconversion to some low IF, such as 140 kHz, positions the signal above dc, obviating most of the aforementioned problems [4]. DC offset can be simply removed by a high-pass filter, and second-order distortion is not important since blockers only generate components near dc, which will fall below the downconverted signal spectrum. As in direct conversion, low-power active bandpass filters can still select the wanted channel. However, unlike direct conversion, the image of the downconverted adjacent or alternate channels can now dominate the channel of interest (Fig. 2). GSM limits the adjacent channel to be 9 dB larger than the desired signal, and the alternate channel to be 41 dB larger. The spectral tail of the GMSK modulated alternate channel dominates the image at this low IF, requiring an image rejection of 26 dB. Fortunately, this can be easily realized on-chip.

If low IF is attractive for a GSM receiver, what is the best frequency plan for downconversion? Single conversion from RF to low IF is certainly the simplest [4]. However, an important though not well-appreciated limitation of this plan arises

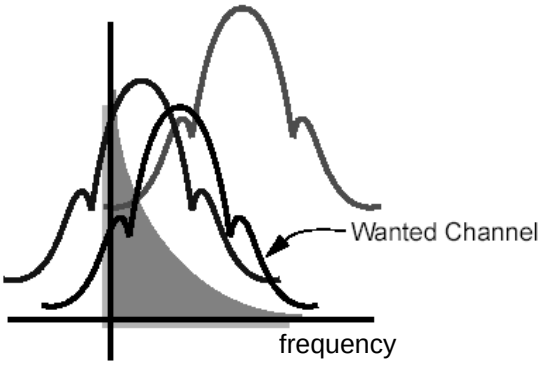


Fig. 2. Downconverted GSM channel, images of adjacent and alternate channels, and flicker noise.

from noise in the downconversion mixer. When translating a signal from a very high to a very low frequency, a current-commutating CMOS mixer always suffers from large flicker noise at its output [7]. This can submerge the weak received signal, which lies at low IF at the mixer output. Flicker noise at the output is inversely proportional to the product *S* x *T*, where *S* is the slope of the LO waveform at the instant of commutation, and *T* is the LO period. Therefore, to lower flicker noise, the LO period should be large and its waveform square-wave-like [7].

This favors a *dual-conversion* receiver. The entire GSM band is first downconverted to a first IF of, say, about 100 MHz. In a conventional receiver, an off-chip IF filter would partly suppress the in-band blocking signals to relax the required linearity in the following stages. On a single-chip receiver, this is not a viable option. Now strong blockers, only mildly suppressed by the antenna preselect filter, propagate through the receiver circuits until finally attenuated in a baseband channel-select filter. This poses some problems previously unknown to superheterodyne receivers, and ignored in previous work.

Consider, for example, dual conversion to zero IF. This suffers from second-order distortion products at dc due to the blockers. For the GSM receiver, where the in-band blocker 3 MHz away from the desired signal is 76 dB larger than the reference channel, this requires an IIP2 of at least 82 dBm in the baseband circuits [8].

Blockers may also translate to the IF through circuit nonlinearity and mixing with LO harmonics [9]. This is important in GSM, which allows large blockers. For example, a blocker lying at $(n\omega_{LO} + \omega_{IF})/k$, subject to *k*th-order nonlinearity in the front end mixes with the *n*th LO overtone and downconverts to IF. Front-end linearity, antenna filter selectivity, and the tuned load of the low-noise amplifier (LNA) all determine how important this effect will be (Fig. 3). Depending on *n* and *k* the choice of IF, the troublesome blockers might lie very close to the GSM band, in which case,

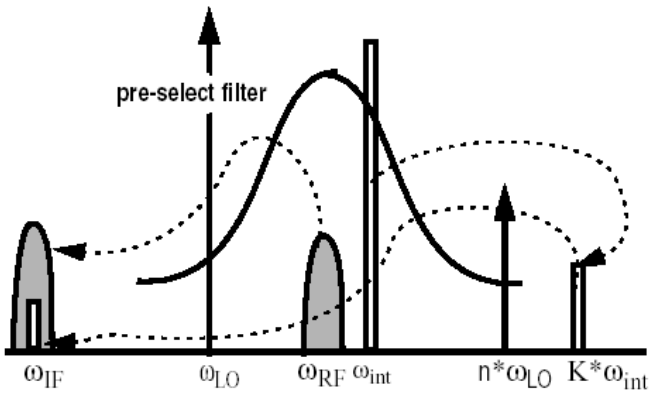


Fig. 3. Harmonic mixing and front-end nonlinearity translate the blocker (ω_{int}) to the same frequency as the desired signal.

preselect filtering before the receiver does not help. On the other hand, the correct IF guarantees that all troublesome blockers lie far away from the GSM band. The typical preselect filter and the dominant second- or third-order distortion in the front-end blocks imply that the combinations $k = 2$, $n = 3$ and $k = 3$, $n = 5$ are most problematic with low-side LO injection, and $k = 3$, $n = 2$ for high-side LO injection. It follows that with high-side LO injection, IF in the range 260-370 MHz is unsafe [10].

Incidentally, if the receiver front-end were perfectly linear, blockers at $n\omega_{LO} \mp \omega_{IF}$ could still mix with LO overtones and downconvert to IF. The case of $n = 1$ arises when the blocker lies at the image frequency. Nonlinearity in baseband circuit blocks prior to channel selection is also important. A blocker at $\omega_{LO} + \omega_{IF}/m$ downconverts to the IF subharmonic, and after suffering m^{th}-order circuit nonlinearity appears at IF. At low IF, this set of blockers may lie within the GSM band and pass through the antenna preselect filter unattenuated. This enforces a minimum linearity on the IF circuit blocks. For high-side LO injection, an IF higher than 170 MHz prevents this from occurring, and for low-side injection an IF higher than 260 MHz.

Based on these various considerations, with high-side LO injection the IF should lie between 170-260 MHz to best suppress spurious downconversion of blockers. There is no such safe region for low-side LO injection [10].

3. Receiver Architecture

Fig. 4 shows the proposed receiver, a dual conversion to low-IF architecture. The signal after low-noise amplification downconverts to a first IF (IF1). There, a programmable gain amplifier (PGA) boosts the wanted signal and close-in blockers, which downconvert to a second, low IF (IF2) for most of the amplification and filtering in the receiver.

The receiver is considerably simplified if the two LO frequencies driving mixers at the first and second IF are derived from a single frequency synthesizer. With IF2 chosen at 140 kHz, this consideration narrows down IF1 to 190 MHz if it is to lie in the safe region described above. Then, according to Fig. 1, the first image 380 MHz away from the desired signal must be attenuated by 110 dB. A typical antenna prefilter for GSM will suppress this image by about 55 dB in its stopband, which means that an IF1 polyphase filter and the tuned load of the LNA must suppress it by another 55 dB.

The gain of the single path (not quadrature) IF1 PGA is chosen so that blockers do not overload the subsequent circuits and desensitize the receiver. Next, *I* and *Q* mixers downconvert the GSM channel to IF2. Dual conversion improves receiver NF, first by providing larger gain before the high-noise baseband section, and second by driving the second mixers at a lower frequency, which reduces their output flicker noise. An active polyphase filter at the IF2 serves the dual purpose of selection of the desired channel, and rejection of the adjacent and image signals. This filter delivers the wanted channel alone at its output in differential quadrature format. Variable gain in the filter adjusts the signal level to lower dynamic range of the following simple A/D converter. Analog channel selection as described here contrasts with most other published GSM receiver ICs, which defer this task to a digital filter.

A high IF2 will lessen the impact of flicker noise in the baseband circuits, but the image signal will be very large. Also, the channel-select filter at high IF2 will consume more power. The second IF of 140 kHz is chosen as the best compromise between image rejection, flicker noise, and power dissipation.

The two LO frequencies are ganged. A divide-by-6 circuit generates the LO2 frequency from the RF VCO. As the receiver tunes across the GSM band, IF 1 varies

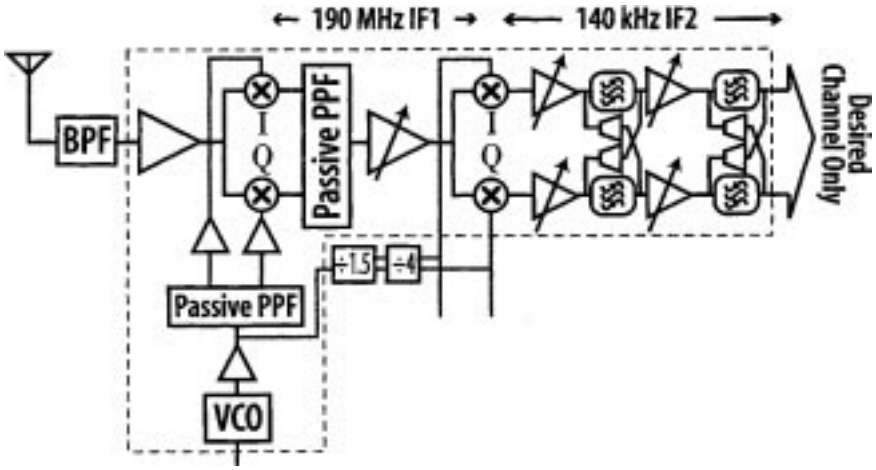


Fig. 4. Architecture of this GSM receiver prototype.

by ±2.5 MHz in order for the wanted channel at IF2 to fall into the fixed passband of the channel-select filter. This varying IF is only possible when the receiver does not depend on a fixed off-chip filter. To tune across the 25-MHz-wide GSM band, the first LO frequency varies by 30 MHz.

In this frequency plan, it is possible for the fifth overtone of LO2 to couple to the receiver input through wirebond inductors, the substrate, or otherwise poor reverse isolation. This overtone lies next to the desired signal, and comprises another strong blocker at the input. As long as it is not so large as to desensitize the receiver, this overtone downconverts to dc and along with other dc offset is rejected by the channel select filter.

Gain is distributed between the various blocks to maintain the cascade NF below 6 dB at minimum detectable signal, and the cascade input IP3 below -15 dBm. A large blocking signal at the input can potentially drive the receiver into compression, lowering average gain, and raising receiver NF. On detecting a blocker, the receiver gain is lowered to improve linearity. The circuits are designed so that cascade NF degrades by no more than 2.5 dB at low gain, which is better than the 3-dB degrada–tion allowed by the GSM standard.

4. CIRCUIT DESIGN

All circuits described here operate from a 2.5-V power supply.

A. Low-Noise Amplifier

The low noise amplifier is a single-ended inductively degenerated common-source stage (Fig. 5). A large voltage gain in the reactive input matching network overcomes FET noise, enabling low NF at low currents. The input impedance is sensitive to the package and parasitic capacitors to ground. Stray capacitance at the gate of the LNA transforms down the input resistance as follows:

$$Re[Z_{in}] \approx \frac{C_{gs}^2}{R^2 C_{gs}^2 C_p^2 \omega^2 + (C_{gs} + C_p)^2} \cdot R^2$$

where R is the real part of the LNA input impedance to the right of the parasitic, and C_{gs} is the gate-source capacitance. The larger the C_{gs}, the less the sensitivity to C_p. Therefore, the input FET W/L is chosen to be a large 350/0.35 um, whose C_{gs} is about 350 fF. With a wire-bond inductance of 2 nH, the real part of at the gate is 148 Ω. The estimated total parasitic capacitance to ground is about 1 pF, which lowers the input re–sistance to 17 Ω. An off-chip LC matching circuit serving as a narrow-band transformer restores this to 50 Ω. The large FET must be biased at low (V_{gs} -

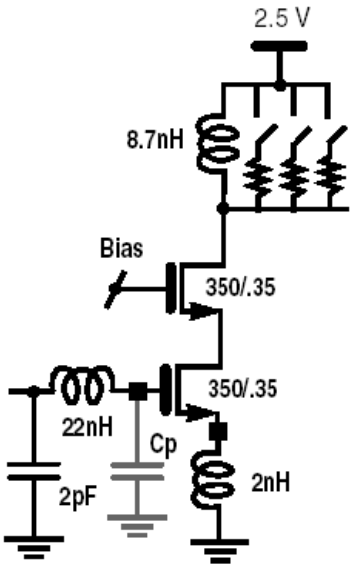


Fig. 5. Common-source LNA circuit.

V_t) to limit the LNA current, but too low a value will lead to small f_T and IP3. A bias current of 4.5 mA gives a satisfactory compromise between these quantities.

An on-chip single layer metal-4 spiral inductor tunes the FET drain load to 1 GHz. Inductor *Q* at this frequency is about 6. Three switch-selectable resistors placed across the inductor can vary the gain from 12 to 24 dB. The pMOS FET switches of 160/0.35 um give low ON resistance and also improve load linearity. This high-*Q* load inductor and the high-input matching network partly suppress the image 380 MHz away.

The cascode FET M2 lowers Miller multiplication of the input FETs C_{gd} and improves LNA reverse isolation. At high frequencies, junction capacitor C_{db} of M1 raises the high frequency output noise due to M2. This capacitance is lowered by merging the drain of the input device and the source of the cascode transistor, and deleting metal contacts which are not required [11]. To lower noise due to spreading resistance, the substrate is contacted close to the FET source junction.

The minimum LNA noise figure of 1.4 dB in the GSM band is obtained at a gain of 25 dB. It rises to 1.6 dB when gain is lowered to 20 dB. Measured LNA IIP3 is +2 dBm.

B. RF Mixers

The RF mixers downconvert the input signal to 190 MHz. It is most straightforward to drive a single-balanced mixer with the single-ended output of the LNA. A single-balanced mixer realized by a standard differential pair suffers from LO feedthrough

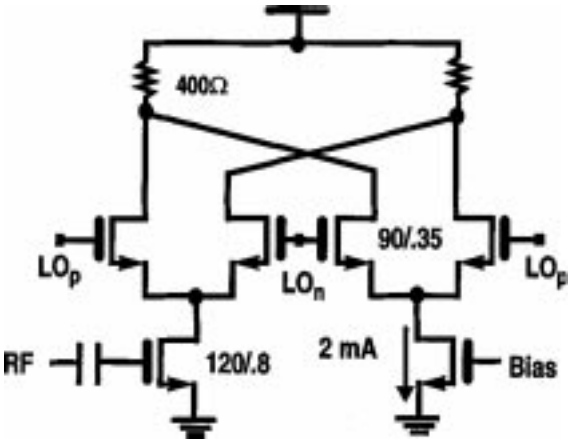


Fig. 6. Double-balanced mixer circuit.

at the output. The feedthrough tone is larger than the downcon–verted RF signal in the ratio $(V_{gs} - V_t)/2v_{RF}$, where v_{RF} is the RF input to the mixer (tens of microvolts) and V_{gs} - V_t is the gate overdrive on the mixer input FET (hundreds of millivolts). A low-order on-chip filter cannot adequately suppress large LO feedthrough at 1.14 GHz while passing the desired signal at 190 MHz. One possibility is to use current boosting [12], which lowers, although does not eliminate, LO feedthrough by steering bias current away from the switching differential pair; however, this is at the expense of linearity.

A double-balanced mixer does not suffer from LO feedthrough, but needs a differential input. A differential LNA was not considered because it requires a balun prior to the receiver, whose insertion loss directly adds to the RX noise figure. Alternatively, two RF phase splitting buffers can follow the LNA, but they consume power and degrade linearity. Instead, we chose to ac couple the single ended output of the LNA to one input each of two double-balanced mixers, whose other inputs are at ac ground (Fig. 6). Although each mixer's single-ended noise figure is 3 dB higher and the bias current is twice that of a single-ended mixer, this arrangement nulls LO feedthrough and additive noise on the LO signal lines. Long-channel grounded-source FETs at the mixer input improve linearity. Each mixer consumes 4 mA, and its double-side-band (DSB) noise figure is 10 dB with respect to the noise in 50 Ω. The subsequent polyphase filter rejects image noise in the mixer.

C. Polyphase Filter

Following the quadrature mixers, a passive RC polyphase filter passes the desired signal with an insertion loss of 9 dB, and rejects the image by more than 40 dB relative to the de–sired signal. Three stagger-tuned stages (Fig. 7) are required to guaran-

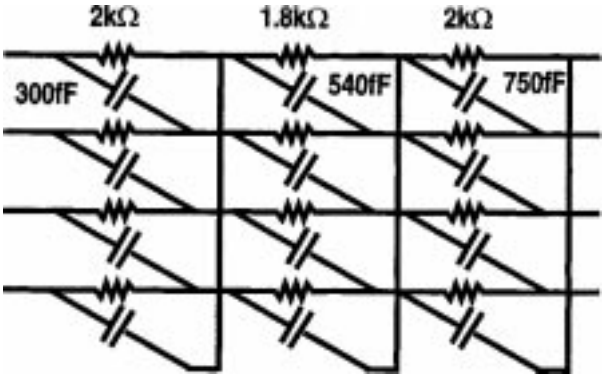


Fig. 7. Three-stage passive RC polyphase filter.

tee this image rejection across the GSM band, with a safety margin of ±30% to cover process spreads. The resistors in this filter are unsalicided polysilicon with the typical sheet resistance of 133 Ω/❑ and ±20% process variation, and capacitors are MIM with 30-nm-thick oxide, which gives a density of 1 fF/um. The capacitor value varies by about ±15% due to process spreads.

Passband loss in the polyphase filter is lowered by tapering up the resistance in the last stages [13]. Resistance of 2 kΩ in the last stage trades off noise and signal loss. The first stage uses a 2-kΩ resistor to prevent loading of the first mixer.

To reject the image by 40 dB, the components in the four branches of the filter must match to 1%. Capacitors of 200 fF or larger, and resistors occupying an area of at least 70 um^2 will match to this accuracy.

The LO commutates the RF signal current with a square wave. While the main harmonic of LOQ lags LOI in phase, the third harmonic of LOQ leads LOI. Certain out-of-band blockers will mix with LO harmonics and downconvert on top of the desired signal. The polyphase filter passes those mixer products possessing the same sequence [17] as the desired signal, but rejects mixer products with the opposite sequence, of which the image signal is the most obvious one. Specifically, with high-side LO injection, the polyphase filter rejects the image at $\omega_{LO} + \omega_{IF}$ and also blockers at $3\omega_{LO} - \omega_{IF}$, $5\omega_{LO} + \omega_{IF}\ldots$. The additional suppression of this subset of blockers is a welcome bonus.

D. Programmable Gain Amplifier

A programmable gain amplifier (PGA) at IF1 adjusts receiver gain to limit the signal level driving the second mixers. The PGA must be a wideband amplifier at the high IF of 190 MHz, which as a practical matter can provide a few coarse steps of gain. Fig. 8 shows the PGA circuit. The input FETs are sized as 120/0.35 um to lighten the capacitive loading on the polyphase filter. The PGA should be very linear at the

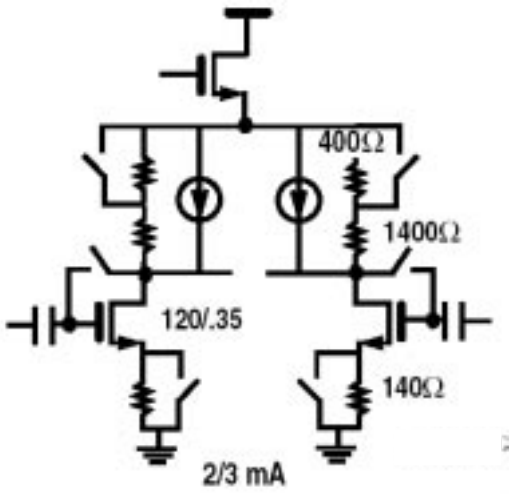


Fig. 8. Programmable gain amplifier circuit.

intermediate nominal gain of 10 dB. At this gain, the input differential pair is degenerated with linear resistors. Resistor loads further improve linearity and lower noise.

To step up gain to 20 dB, nMOS switches bypass the degeneration resistors. Bias current is steered away from the loads to improve headroom. A pMOS switch in parallel with the 400-Ω load resistor steps down gain to 10 dB. For 0-dB gain, switches short the differential input to the differential output, while another switch opens the connection to V_{DD}. The input switch FET size of 12/0.35 um limits capacitor loading on the polyphase filter, yet gives an acceptably low ON resistance of 180 Ω. The PGA NF is 6 dB at maximum gain of 20 dB. For 8-dB gain, the PGA NF is 9 dB and IIP3 is +19 dBm. The circuit drains 3 mA. The outputs of the PGA are ac coupled to the next stage by two picofarad capacitors.

E. Second Mixers

Quadrature double-balanced second mixers downconvert the wanted GSM channel to the IF2 of 140 kHz (Fig. 9). The 12-pF capacitor at each mixer's output removes high-frequency blockers as well as the signals upconverting to 380 MHz. The input FETs are biased at a gate overdrive of 450 mV to handle amplified blocking signals without gain compression. The outputs directly couple into the subsequent active filter.

A square-wave LO waveform switches the mixers to lower their output flicker noise. The mixer is laid out to minimize parasitic capacitance at the tail of each differential pair, which also helps to lower flicker noise [7]. The DSB noise figure is 11 dB relative to noise in a differential 100-Ω resistor, and flicker noise corner is below 40 kHz. The following active polyphase filter cancels image noise. Long-channel

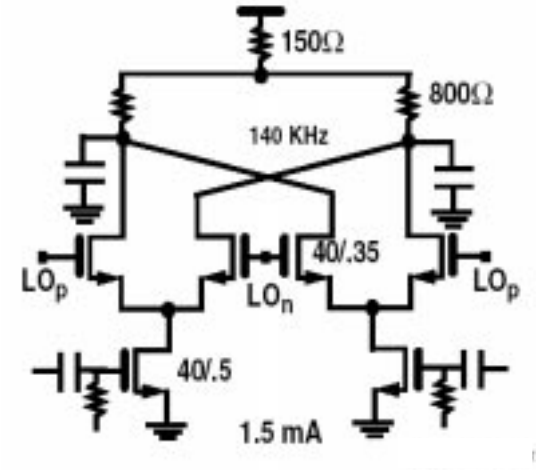


Fig. 9. Second mixer circuit.

grounded source input FETs improve input IIP3 of the mixer to +15 dBm. Each mixer drains 1.5 mA, and gives a conversion gain of 6 dB.

F. Active Polyphase Filter

At IF2, the image signal is only 280 kHz away from the wanted signal. After quadrature downconversion, both lie at the second IF of 140 kHz but in opposite sequences. The image consists of the lower adjacent and alternate channels, folded over and offset in frequency relative to the desired signal. Based on the modulated spectrum, the image must be rejected by 26 dB (Fig. 10).

A single *passive* polyphase filter stage can notch the image by 20 dB or more across a bandwidth as wide as 30% of the null frequency. Therefore, image rejection from 20 to 260 kHz which spans one GSM channel centered at 140 kHz and 30%

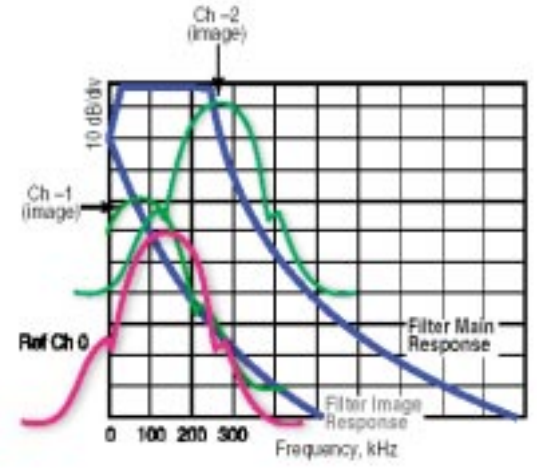


Fig. 10. GSM reference channel downconverted to IF of 140 kHz, and images of the left adjacent and left alternate channels that overlap it.

safety margin requires more than seven or eight stagger-tuned stages in cascade. This raises the passband insertion loss to more than 10 dB, which is unacceptable.

An alternative is the *active* polyphase filter [14]. This filter's frequency response is defined by a linear shift of a low-pass characteristic along the frequency axis, resulting in a bandpass filter with an arithmetically symmetric passband. A shift by ω_0 is obtained by transforming the Laplace variable s to $(s - j\omega_0)$. At the circuit level, this means that each capacitor in a low-pass active *RC* prototype is replaced by the composite structure shown in Fig. 11 to realize an admittance $(s - j\omega_0)C$. This requires quadrature signal paths.

In practice, component mismatch in the filter branches, or phase and amplitude mismatch at the input of the filter, will limit image rejection. Any signal at the input of the filter can be decomposed into two set of quadrature signals with opposite sequences. The filterpasses one sequence and rejects the opposite sequence. The frequency response to the wanted sequence is the shape of the filter passband on the positive frequency axis, while to the unwanted sequence, it is the transition band and stopband on the negative frequency axis. In case of quadrature errors in the input signals, the *image* signal can be decomposed into the unwanted sequence and a small wanted sequence. The frequency response to the image is now a linear superposition of the original filter response on the positive and negative frequency axes, weighted by the ratio of the two sequences.

Active *RC* and switched-capacitor filters are realized with high-gain opamps and linear resistors and capacitors, which makes them more linear than g_m-*C* filters. The active *RC* realization is chosen here to avoid an antialiasing filter. The sharper the filter transition band, the greater the image rejection 280 kHz away. Actual image rejection will also be limited by imbalance in the quadrature mixers, and component

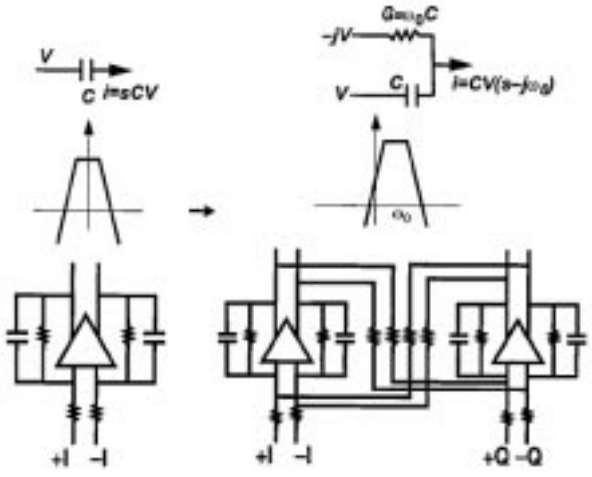


Fig. 11. Method to synthesize active bandpass polyphase filter from conventional low-pass prototype.

mismatch in the two branches. The fifth-order Chebychev type-I filter constellation gives 1-dB passband ripple, rejects the adjacent channel by 30 dB, and adequately suppresses image frequencies falling in its transition band. The filter is realized by cascading five stages of the type shown in Fig. 11.

Two-stage opamps drive the resistor loads (Fig. 12). Large FETs in the input stage lower flicker noise. A Miller capacitor and series resistor compensate the opamp to a unity-gain frequency of 33 MHz. The common-mode feedback (CMFB) cir–cuit is also shown in Fig. 12. The two nMOS transistors sense the output voltage, and the mirror forces equal current in the two pMOS transistors and drives the common-mode of the output nodes to the reference voltage. Long-channel pMOSFETs give high gain. Each branch drains a current of 15 uA. The 1.2-pF capacitors compensate the feedback loop consisting of the main opamp and CMFB amplifier.

The filter is tuned by a 5-bit array of switchable capacitors in parallel with a fixed capacitor, which gives ±50% tuning range with 3% accuracy. Interleaving gain and filtering in the cascade give the best results for the dynamic range. The adjustable gain before each stage can trade off filter noise and signal-induced overload. A switch-selectable array of resistors at the input of each stage scales filter gain. This prototype does not include an AGC loop, but makes available three pins for external gain control. The resistors are weighted to produce linear-in-dB gain. Total filter gain, determined by process-independent resistor ratios, varies by 60 dB. The filter has a minimum noise figure of 26 dB at maximum gain. At intermediate gain, its input IP3 to out-of-band tones which produce in-band intermodulation is +25 dBm.

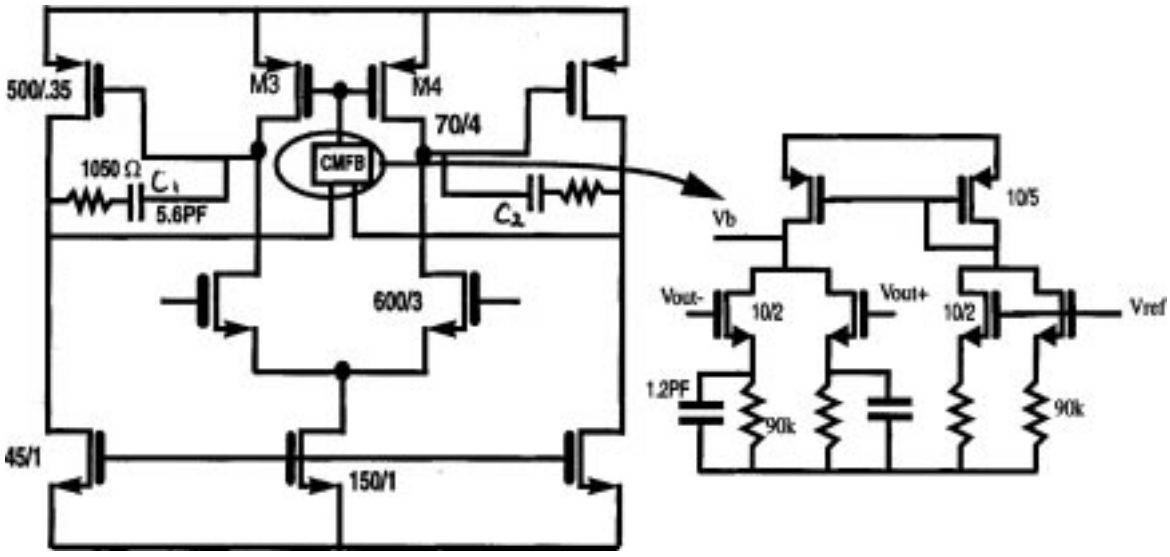


Fig. 12. Two-stage opamp used in filter, and accompanying common-mode feedback circuit.

G. Voltage Controlled Oscillator

A single voltage-controlled oscillator (VCO) generates both LO frequencies, and complies with GSM phase noise requirements. The VCO is a tail-biased differential *LC* oscillator (Fig. 13). MOS varactors tune the VCO frequency by 10%. The varactors are laid out as 160 fingers of 2/0.35-um nMOSFETs. The inductors are octagonal spirals with patterned polysilicon shields, giving a *Q* of 8 at 1140 MHz. The inductor is designed with a fast simulator that accurately captures all metal and substrate losses [15]. The oscillator circuit is optimized for best phase noise at least power consumption using an understanding of the physical mechanisms of phase noise [16]. To lower flicker noise, the tail current source *W* x *L* is scaled up while keeping *W/L* constant. FET V_{gs}-V_t is large for low transconductance, which lowers thermal noise in the drain current. The FET is laid out as an array of nine annular transistors in parallel to reduce the capacitance at the tail, which lowers flicker noise upconversion gain [15]. The circuit consumes 3.4 mA from 2.5 V, and its phase noise as simulated on Spectre-RF is -142 dBc/Hz at 3-MHz offset from 1.1 GHz.

Quadrature differential phases are needed to drive *I* and *Q* mixers. A quadrature VCO [19] automatically creates all these phases. For the same phase noise, though, the quadrature VCO must be biased at 12 mA. In this receiver, a passive polyphase filter follows the VCO to split the differential output into differ–ential quadrature phases. Unlike VCOs biased at high currents [4], here, the polyphase filter resistance significantly loads the oscillator, and its 1-pF input capacitance strongly pulls oscil–lation frequency. Lowering resonator inductance to restore the frequency raises power consumption.

Two grounded source FETs with inductive loads buffer the differential VCO output. A narrowband T-section capacitive transformer boosts the filter's input

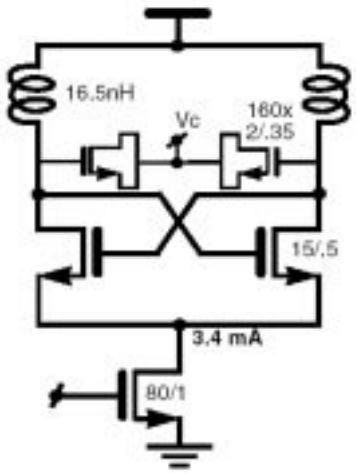


Fig. 13. 1.1-GHz VCO.

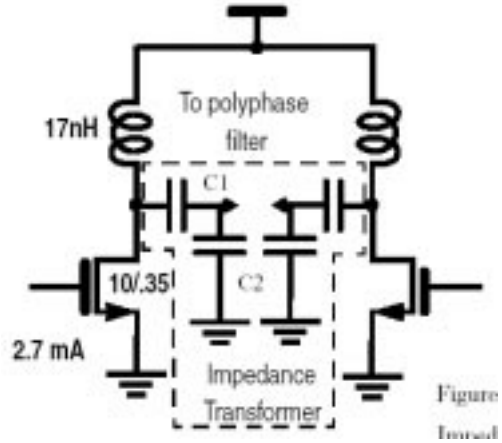


Fig. 14. VCO buffer with capacitive impedance transformer.

impedance (Fig. 14). The inductor load of the buffer resonates with the capacitive input of the passive polyphase filter (PPF) at 1.14 GHz to give a gain of 6 dB. The buffer drains 2.7 mA. The polyphase filter attenuates the LO amplitude by about 10 dB. Two sets of buffers, consisting of inductively loaded differential pairs with 46-nH series stacked inductors, boost the differential quadrature outputs of the polyphase filter to drive the mixers. The inductors yield 1 kΩ impedance at 1140 MHz to lower bias current to 0.9 mA. The entire LO generator consumes 7.9 mA.

5. Experimental Results

Fig. 15 shows SPECTRE-RF simulations of the entire receiver. The receiver gain from single-ended input to differential output is as high as 100 dB, and the cascade noise figure is below 5 dB. Flicker noise was simulated throughout the design evolution using the following expression for gate-referred noise voltage:

$$\overline{v^2}_{nf} = \frac{K}{f}\frac{1}{WLC_{ox}}$$

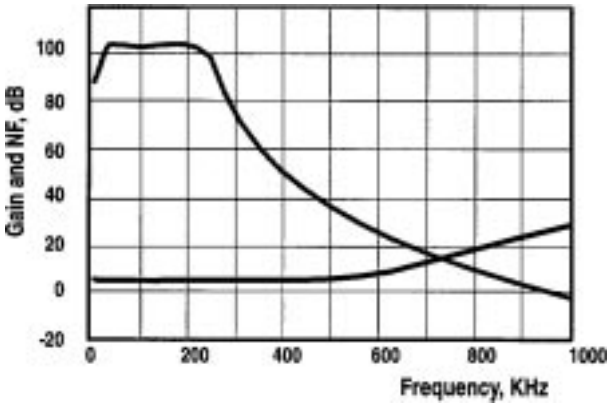


Fig. 15. Simulated frequency response and noise figure of receiver.

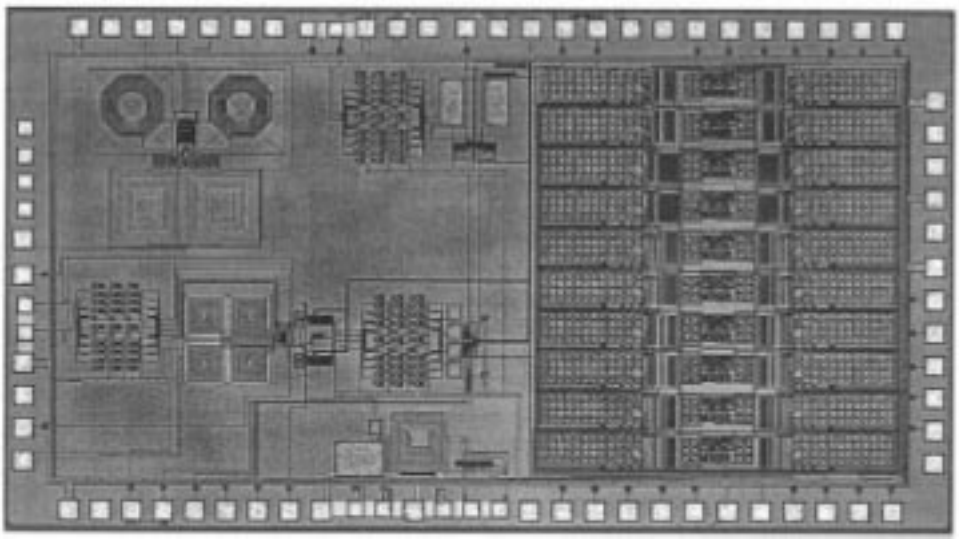


Fig. 16. Chip photograph.

In this technology, for nMOS is 8.2E -24 V^2F, and for pMOS is 1.9E-24 V^2F, and is assumed constant at all bias voltages [18].

A receiver prototype is fabricated in 0.35-um CMOS in the STMicroelectronics BiCMOS6M process. The chip (Fig. 16) consumes an area of 2.2 x 4 mm, the channel-select polyphase filter consuming almost half of this. Separate supply lines lower interaction between the LNA and the rest of the receiver. Several ground pads are down bonded to the package to lower stray inductance. Electrostatic discharge (ESD) protection comprises two reverse-biased diodes connected to each pin of the chip. The small diodes add little capacitance to the sensitive RF nodes, and are separated by ground and rings to reduce the possibility of latchup. Symmetric layout throughout the receiver guarantees required matching between and paths. Serpen–tine metal lines of variable pitch equalize interconnect capacitance between mixer

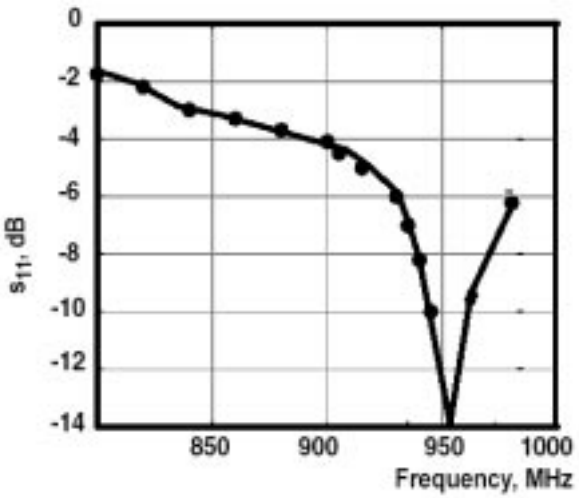


Fig. 17. Measured input reflection coefficient versus frequency.

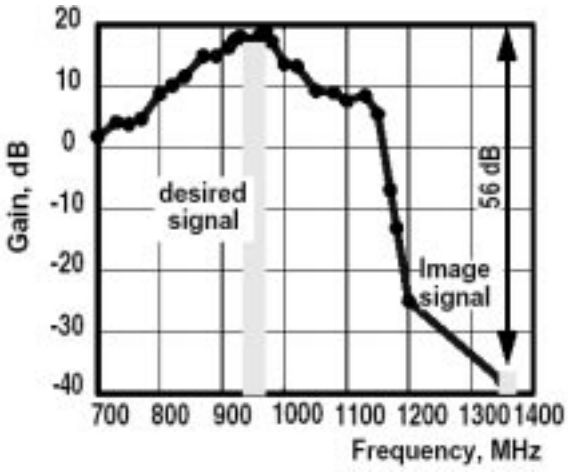


Fig. 18. Measured on-chip image rejection at first IF.

outputs and filter inputs, and inside the LO circuits. The back side of the chip is attached with conductive epoxy to a 40-pin surface-mount alumina/glass package.

The measured input return loss is 12 dB in the GSM band (Fig. 17). The front end comprising the LNA, first mixers, and the polyphase filter is first evaluated separately through test points. Front-end insertion gain is 20 dB in the GSM band, and it rejects the image on-chip by more than 55 dB (Fig. 18).

Next, receiver noise is measured up to the channel-select filter input using the HP8870B noise-figure meter. This instrument only operates at IF of 10 MHz and higher, but is the most accurate way to measure noise. It is used to calibrate a spectrum analyzer for noise measurement at IF of 140 kHz and below. In the white-noise region, measured receiver NF is 3.8 dB up to the second mixers. The channel-select filter degrades cascade NF by less than 0.5 dB. Due to onset of flicker noise, receiver NF at 240 kHz IF is 4 dB, rising to 5.7 dB at 40 kHz. The extrapolated 3-dB flicker corner frequency is below 20 kHz, where the spectral density of the wanted channel is down by more than 10 dB and the higher noise no longer matters (Fig. 19).

To test the immunity of the receiver to blockers, a single tone at each frequency listed in Table I is applied to the receiver input, and the output at IF2 is measured. The receiver gain to a -99 dBm tone in the wanted channel is also measured. The difference between the two is the blocking rejection. All major blockers, including the image at the first IF, are rejected by more than 110 dB. As explained earlier, the polyphase filter additionally attenuates a subset of blockers.

The channel-select filter gives the main amplification in the receiver. Its gain may be swept from about -2 dB to 58 dB, as shown in Fig. 20. The filter maintains its shape at all gain settings. The passband ripple is within 1 dB, and it rejects the adjacent channel 100 kHz away by more than 35 dB. The measured frequency response of the entire receiver at IF2, shown in Fig. 21, plots gain to the image at the second

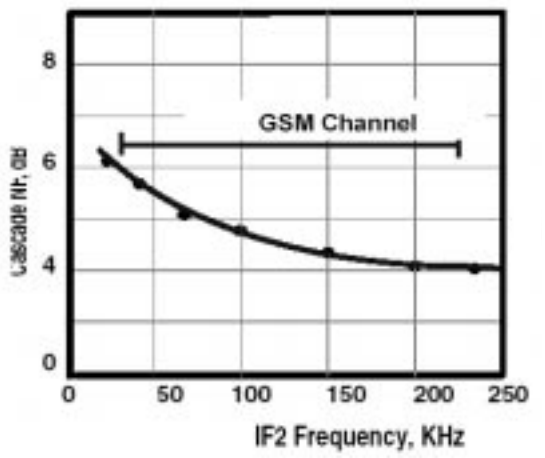


Fig. 19. Measured receiver cascode noise figure.

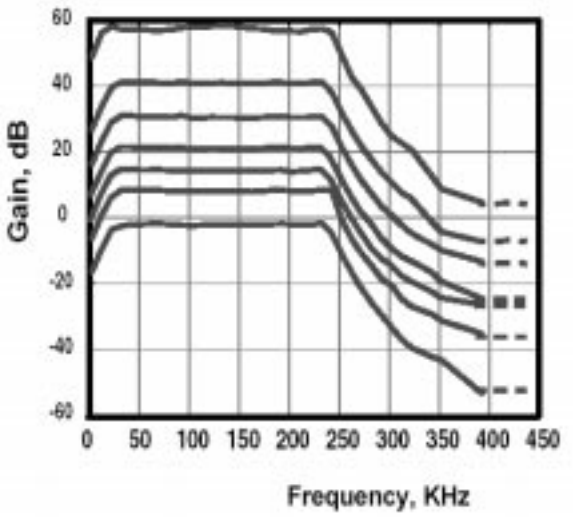


Fig. 20. Measured frequency response and gain range of channel select filter.

IF. At frequencies up to 70 kHz, the image response follows the folded-over filter transition band. At midchannel, quadrature inaccuracy in the second mixers limits

TABLE I. MEASURED REJECTION OF LIKELY BLOCKING SIGNALS

Physical Effect	**RF Input**	**Frequency (MHz)**	**On-chip Rejection (dB)**	**Rejection Incl. Prefilter**
1st Image	$(f_{LO}+f_{IF1})$	1328	56	111
Half IF	$(2f_{LO}+f_{IF1})/2$	1233	81.5	131
	$(2f_{LO}-f_{IF1})/2$	1043	72	114
	$(3f_{LO}+f_{IF1})/2$	1802	>91	141
	$(3f_{LO}-f_{IF1})/2$	1612	>109	159
Half IF	$(4f_{LO}+f_{IF1})/2$	2371	>93	148
	$(4f_{LO}-f_{IF1})/2$	2181	>93	148
HD3 in balanced IF1 circuits	$(2f_{LO}+f_{IF1})/3$	822	97	142
	$(2f_{LO}-f_{IF1})/3$	695	71	126
	$(3f_{LO}+f_{IF1})/3$	1201	76	126
	$(3f_{LO}-f_{IF1})/3$	1075	56	106
	$(5f_{LO}+f_{IF1})/3$	1960	>116	171
	$(5f_{LO}-f_{IF1})/3$	1833	>116	171
	$(3f_{LO}+f_{IF1})$	3604	106	161
	$(3f_{LO}-f_{IF1})$	3224	>103	158

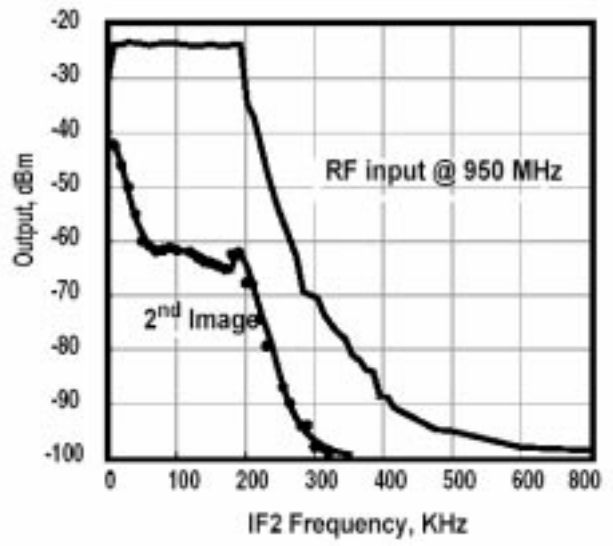


Fig. 21. Measured cascade frequency response of receiver to desired signal and second image.

image rejection; as explained earlier, the image response now resembles a scaled-down version of the passband. At 40 kHz, the image at IF2 is suppressed by 27 dB, and at 70 kHz by 37 dB. This is good enough.

The measured VCO phase noise at 1140 MHz is shown in Fig. 22. The VCO meets the GSM phase-noise requirements at all offset frequencies by some margin. At 3-MHz offset, the measured phase noise is -142 dBc/Hz and the flicker corner of the VCO is about 150 kHz. The measured results matches well with Spectre-RF simulations. The VCO tuning range is about 10%. The phase noise changes by less than 2 dB across the tuning range.

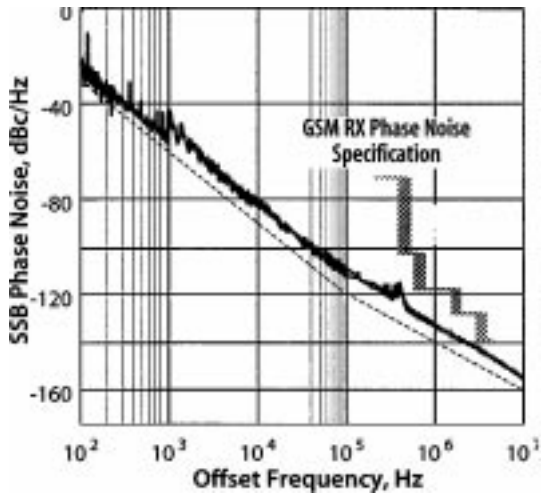


Fig. 22. Measured VCO phase noise.

TABLE II. RECEIVER PERFORMANCE SUMMARY

NF	5 dB
IIP3	- 16 dBm
On Chip Image Rejection	>55 dB @ 1st IF >35 dB @ 2nd IF
Input Return Loss	>12 dB (GSM band)
AGC	> 80 dB
Current Consumption @ 2.5 V	22mA (RX only) 8 mA (LO Generator
Area	2.2 x 4 mm

6. Conclusion

A fully integrated dual-conversion to low-IF RX is implemented in 0.35-um CMOS process. Dual conversion to low-IF lowers the impact of baseband noise. The right choice of IF for first downconversion substantially lowers the impact of the blockers. This IF also allows use of a single VCO to synthesize both LOs. A multifunction active polyphase filter selects the desired GSM channel and rejects the image signal by more than 27 dB across the filter passband. Table II summarizes receiver performance. The image signal at first IF is rejected by the LNA and three-stage polyphase filter by more than 40 dB. The minimum cascade noise figure is 5 dB and the cascade input IP3 is -16 dBm. The receiver consumes 22 mA from a 2.5-V power supply. The VCO and buffers drain another 8 mA.

This work illustrates the power of evolving a radio architecture with cognizance of the strengths and limitations of circuits. This design approach has enabled a highly integrated all-CMOS solution to the difficult GSM receiver problem with a power consumption among the lowest to date.

REFERENCES

[1]. T. Stetzler, I. Post, J. Havens, and M. Koyama, "A 2.7-4.5 V single chip GSM transceiver RF integrated circuit," *IEEE J. Solid-State Circuits*, vol. 30, pp. 1421-1429, Dec. 1995.

[2]. T. Yamawaki, M. Kokubo, K. Irie, H. Matsui, T. Endou, H. Hagisawa, T. Furuya, Y. Shimizu, M. Katagishi, and J. Hildersley, "A 2.7-V GSM RF transceiver IC," *IEEE J. Solid-State Circuits*, vol. 32, pp. 2089-2096, Dec. 1997.

[3]. P. Orsatti, F. Piazza, and Q. Huang, "A 20-mA-receive, 55-m A-transmit, single-

chip GSM transceiver in 0.25-um CMOS," *IEEE J. Solid-State Circuits*, vol. 34, pp. 232-233, Dec. 1999.

[4]. M. S. J. Steyaert, J. Janssens, B. de Muer, M. Borremans, and N. Itoh, "A 2-V CMOS cellular transceiver front-end," *IEEE J. Solid-State Circuits*, vol. 35, pp. 1895-1907, Dec. 2000.

[5]. "Digital cellular telecommunication system: Radio transmission and reception (GSM 5.05)," European Telecommunication Standards Inst. (ETSI), Sophia Antipolis, France.

[6]. A. A. Abidi, "Direct conversion radio transceivers for digital communications," *IEEE J. Solid-State Circuits*, vol. 30, pp. 1399-1410, Dec. 1995.

[7]. H. Darabi and A. Abidi, "Noise in CMOS mixers: A simple physical model," *IEEE J. Solid-State Circuits*, vol. 35, pp. 15-25, Jan. 2000.

[8]. J. Rudell, J. Weldon, J. Ou, L. Lin, and P. Gray, "An integrated GSM/DECT receiver: Design specification," UCB Electronics Laboratory Memo., Berkeley, CA, 1998.

[9]. S. Fitz, "Receiver architectures for GSM handsets," in *IEE Coll. Design of Digital Cellular Handsets*, London, U.K., 1998, pp. 3/1-3/10.

[10].E. Cijvat, "A study of CMOS receiver architectures in mobile commu–nication handsets for GSM," Ph.D. dissertation, KTH, Kista, Sweden, 2000.

[11].F. Stubbe, S. V. Kishore, C. Hull, and V. Della Torre, "A CMOS RF-receiver front-end for 1 GHz applications," in *Proc. Symp. VLSI Circuits*, Honolulu, HI, 1998, pp. 80-83.

[12].W. H. Sansen and R. G. Meyer, "Distortion in bipolar transistor variable gain amplifiers," *IEEE J. Solid-State Circuits*, vol. 8, pp. 275-282, Aug. 1973.

[13].F. Behbahani, Y. Kishigami, J. Leete, and A. A. Abidi, "CMOS 10 MHz-IF downconverter with on-chip broadband circuit for large image-suppression," in *Proc. Symp. VLSI Circuits*, Kyoto, Japan, 1999, pp. 83-86.

[14].J. Crols and M. Steyaert, *CMOS Wireless Transceiver Design*. Boston, MA: Kluwer, 1997.

[15].J. Lee, A. A. Abidi, and N. G. Alexopoulos, "Design ofspiral inductors on silicon substrates with a fast simulator," in *Proc. Eur. Solid-State Circuits Conf.*, The Hague, The Netherlands, 1998, pp. 328-331.

[16].J. J. Rael and A. A. Abidi, "Physical processes of phase noise in differential LC oscillators," in *Proc. IEEE Custom IC Conf.*, Orlando, FL, 2000, pp. 569-572.

[17].M. J. Gingell, "Single sideband modulation using sequence asymmetric polyphase networks," *Electric. Commun.*, vol. 48, no. 1-2, pp. 21-25, 1973.

[18].A. A. Abidi, C. R. Viswanathan, J. J.-M. Wu, and J. A. Wikstrom, "Flicker noise in CMOS: A unified model," in *Proc. Symp. VLSI Technology*, Karuizawa, Japan, 1987, pp. 85-86.

[19].A. Rofougaran, G. Chang, J. J. Rael, J. Y.-C. Chang, M. Rofougaran, P. J. Chang, M. Djafari, M. K. Ku, E. Roth, A. A. Abidi, and H. Samueli, "A single-chip 900-MHz spread-spectrum wireless transceiver in 1-um CMOS-Part I:

Architecture and transmitter design," *IEEE J. Solid-State Circuits*, vol. 33, pp. 515-534, Apr. 1998.

PAPER III

PAPER III

Spurious Mixing of Off-Channel Signals in a Wireless Receiver and the Choice of IF

Abstract

Circuit nonlinearity and LO harmonics can cause large interferers to translate on to the same intermediate frequency as the desired channel. The mechanisms responsible for spurious mixing, which are distinct from intermodulation distortion, are an–alyzed and catalogued. The analysis leads to an optimal choice of IF for a fully integrated 900 MHz GSM receiver that resists all blockers.

Based on: E. Cijvat, S. Tadjpour and A.A. Abidi, "Spurious mixing of off-channel signals in a wireless receiver and the choice of IF". © 2004 IEEE. Reprinted, with permission, from *IEEE Trans. on Circuits and Systems II: Analog and Digital Signal Processing*, Vol. 49, No. 8, pp. 539 - 544, August 2002.

1. Introduction

It is clear that today's wireless transceivers for mobile communication, especially those operating in the range of 1 to 5 GHz, increasingly must be more integrated.

A higher level of integration requires the circuit designer to eliminate off-chip filters: for instance, the passive image-reject filter after the low-noise amplifier (LNA), and the intermediate frequency (IF) filter used to suppress adjacent and out-of-band signals [1]-[3]. Now much stronger adjacent channels, variously called interferers or blockers, will traverse the receiver after only mild attenuation in the antenna preselect filter, eventually to be suppressed at IF or baseband. The blockers disturb the receiver in a number of well-known ways, through reciprocal mixing, gain desensitization and intermodulation distortion. These problems are lessened with low phase noise oscillators and more linear amplifiers, and are roughly independent of the numerical choice of IF. Unfortunately this is not all. Due to circuit nonlinearity and harmonic mixing, blockers may translate to exactly the same IF as the desired channel. This problem of spurious mixing is known only qualitatively to receiver experts [4]-[6]. The object of this paper is to clearly delineate all the important mechanisms of spurious mixing, and to relate them to circuit design and receiver architecture.

Investigating these mechanisms, we find that for a given profile of allowable blockers the problems of spurious mixing go away at certain intermediate frequencies. At these IFs it becomes possible to remove the customary off-chip filterbetween

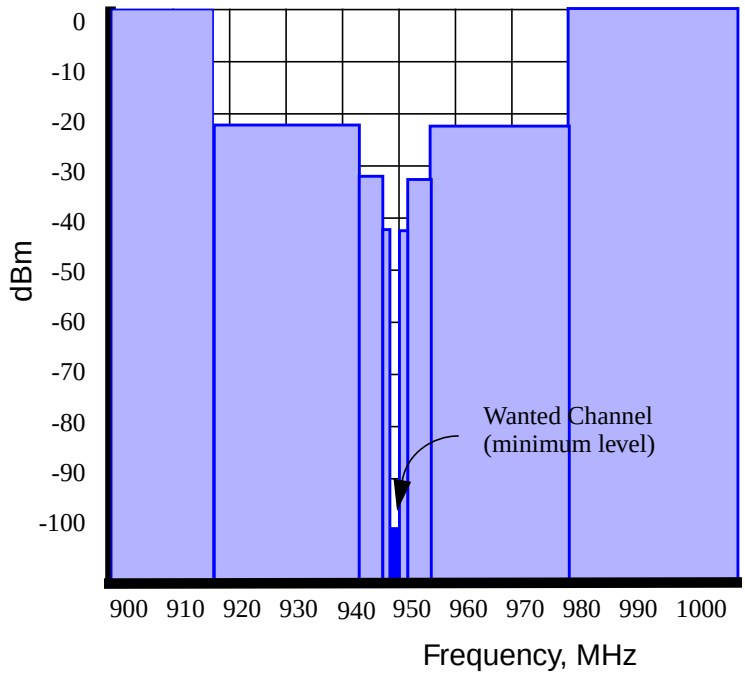


Fig. 1. GSM blocking signal template.

the LNA and mixer. This enables the complete integration of receivers for cellular systems such as GSM that allow very strong blockers. Fig. 1 shows the blocker template [7] for the 900-MHz GSM receive band. In a certified GSM receiver, the ratio of signal-to-noise and interference with the wanted channel 3 dB above the sensitivity level accompanied by a blocker up to 100 dB larger should be the same as with the wanted channel alone at the sensitivity level.

In Section II, we model the front-end and how it generates harmonics. Then in Section III different mechanisms of spurious mixing are described, and the results are applied to the GSM standard to find a good range for IF. Finally, Section IV presents the measured blocking tolerance of a fully integrated GSM receiver whose architecture evolved out of this analysis.

2. Modeling Receiver Front-End

Fig. 2 shows a generic front-end for a superheterodyne receiver. The RF preselect filter connected to the antenna passes the band of interest, but suppresses out-of-band blocking signals. Together with the LNAs tuned input matching network, it suppresses blocking signals by some amount $Att_{BPF}(\Delta f)$, where Δf is the frequency offset from the desired channel. Signals in the preselect filter's passband inevitably suffer a small insertion loss L_{BPF}. The term $Att_{LNA}(\Delta f)$ refers to how much the tuned load of the LNA and any image reject filter that follows attenuate blocking signals. L_n signifies how much lower the nth LO harmonic is when compared to the fundamental.

There are three important nonidealities in a receiver front–end.

1. LNA nonlinearity: This creates harmonics of an incoming RF signal and blockers, as well as intermodulation products.
2. Mixing with LO harmonics: For highest conversion gain, the mixer must completely commutate the RF signal with a square wave at the LO frequency. Therefore, the strong odd harmonics of a 50% duty cycle square-wave LO, given by

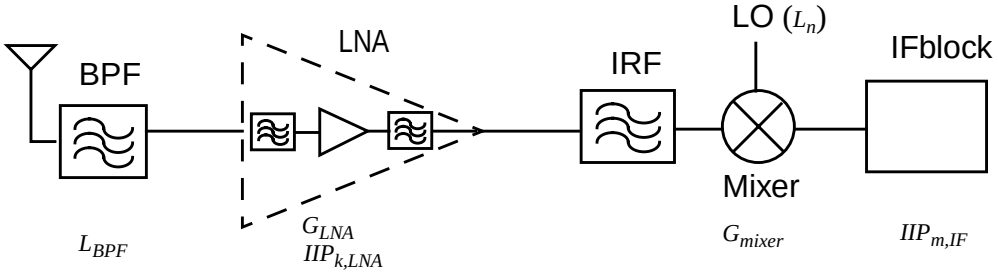


Fig. 2. Generic superheterodyne receiver front end.

$$sq(\omega_{LO}t) = \frac{4}{\pi}\left(\sin\omega_{LO}t + \frac{1}{3}\sin 3\omega_{LO}t + \ldots\right) \quad (1)$$

mix with the full span of the RF input spectrum, which in–cludes in addition to the channel of interest, various strong blocking signals.

3. Nonlinearity in IF circuits: This shifts interferers lying at subharmonics of IF onto the wanted channel.

Therefore, an input interferer at this frequency

$$f_{interf} = (nf_{LO} \pm f_{IF}/m)/k \quad (2)$$

where *k*, *m*, *n* are integers, after being subject to these three possible imperfections in the receiver, will coincide in frequency at IF with the desired signal. Here *k* refers to the *k*th-order nonlinearity of the LNA, *n* is the *n*th LO harmonic, and *m* represents the *m*th-order nonlinearity of the IF circuits after the mixer.

In any circuit with *k*th-order nonlinearity, the difference ΔP between the fundamental output signal and the kth-order intermodulated signal in decibels is [5]

$$\Delta P = (k-1)(IIP_k - P_{in}) = P_{fund,\,out} - P_{k,\,out} \quad (3)$$

where IIP_k is the *k*th-order input intercept point and P_{in} is the input power. For a wide-band receiver, the *k*th-order harmonic generated by a single-tone can be deduced using a similar definition

$$\Delta P_h = (k-1)(P_{k,\,oi} - P_{in}) = P_{fund,\,out} - P_{k,\,out} \quad (4)$$

where the fundamental is located at f_1 and the *k*th-order harmonic at kf_1. The relationship between IIP_k and $P_{k,oi}$ is

$$P_{k,\,oi} = IIP_k + 10\log k \quad (5)$$

Since it is customary to characterize nonlinearity in RF circuits in terms of intercept point IIP_k, from here on we use this to also represent distortion in single-tone blocking signals.

3. Mechanisms of Disturbance by Spurious Signals

Different mechanisms are distinguished by combinations of *n*, *k* and *m*. Three separate mechanisms may be identified as follows.

Mechanism 1: Harmonic Mixing, m = 1, n & k >1: An interferer lying at subject to kth-order nonlinearity in the RF frontend, mixes with the *n*th LO harmonic and downconverts to IF (Fig. 3). It overlays the desired signal, which has appeared at IF through the intended undistorted am–plification in the LNA, followed by downconversion by the fundamental frequency of the LO. For the receiver to function prop–

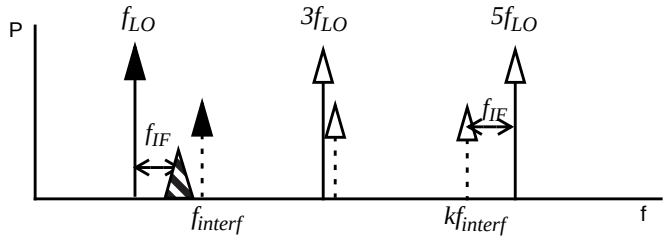


Fig. 3. Harmonic mixing with mechanism 1.

erly, an unintentionally downconverted interferer must lie below the desired signal by some specified carrier-to-interference ratio *C/I*

$$P_{interf,IF} < P_{des,IF} - \frac{C}{I} \tag{6}$$

The wanted channel is amplified in the LNA and mixer by power gains of G_{LNA} and G_{mxr}, respectively. At IF, the desired signal's power is then

$$P_{des,IF} = P_{des} - L_{bpf} + G_{LNA} + G_{mxr} \tag{7}$$

On the other hand, the interferer is subject to a different set of gains and losses. An interfering signal Δf away from the RF band is first suppressed in the RF preselect filter. Then its kth harmonic at kf_{interf}, calculated using (4), is additionally sup–pressed in the tuned load of the LNA. The interferer downconverts to IF by mixing with the *n*th LO harmonic. Finally at the mixer output, its strength in dB is

$$\begin{aligned} P_{interf,IF} = k(P_{interf} - L_{BPF} - Att_{BPF}(\Delta f)) + G_{LNA} - (k-1)(IIP_k + 10\log k) \\ - Att_{LNA}(\Delta f_k) - L_n + G_{mxr} \end{aligned} \tag{8}$$

where

P_{des} Minimum power of desired signal

P_{interf} power of interfering signal (from the antenna) according to the blocking profile

L_{BPF} insertion loss of the band pass filter in the RF band

$Att_{BPF}(\Delta f)$ attenuation of the antenna pre-select filter and the LNA input matching network at an offset frequency Δf

Δf distance of interferer to the RF receive band

$Att_{LNA}(\Delta f)$ attenuation by the tuned load of the LNA and/or the image reject filter at an offset frequency Δf

Δf_k distance of *k*th order harmonic of blocker at LNA output to the RF receive band

L_n loss of *n*th LO harmonic relative to the LO fundamental.

By inserting (7) and (8) into (6), we arrive at the minimum of the LNA that leads to the required carrier-to-interfer–ence ratio for a certain BER [7]

$$IIP_{k,LNA} > \frac{k}{k-1}P_{interf} - L_{BPF} - \frac{k}{k-1}Att_{BPF}(\Delta f) - \frac{1}{k-1}\cdot(Att_{LNA}(\Delta f_k) + L_n) - 10\log(k) - \frac{1}{k-1}\left(P_{des} - \frac{C}{I}\right) \quad (9)$$

If the circuit properties are known but unalterable, the equation may be rearranged to determine the required selectivity of the RF preselect filter

$$Att_{BPF}(\Delta f) > P_{interf} - \frac{k-1}{k}L_{BPF} - \frac{1}{k}(Att_{LNA}(\Delta f_k) + L_n) - \frac{k-1}{k}(IIP_k + 10\log k) - \frac{1}{k}\left(P_{des} - \frac{C}{I}\right) \quad (10)$$

Next, let us find the values of IF which satisfy this inequality for a GSM receiver. To start with, we must know the characteristics of the prefilter. Fig. 4 shows the frequency response of a typical commercially available 947-MHz SAW filter intended for the GSM receive band. Then, based on previous experience with RF CMOS circuits, we assume that in a well-designed LNA the IIP_3 is 0 dBm, and IIP_2 is +12 dBm. As defined by the relationship $kf_{interf} = nf_{LO} \pm f_{IF}$, the most troublesome interferers susceptible to third-order nonlinearity *(k = 3, n = 3)* lie in the frequency range 890 MHz < f_{interf} < 1025 MHz, and due to second-order nonlinearity *(k = 2, n = 2)* lie in the range 890 MHz < f_{interf} < 1020 MHz. After downconversion by the mechanisms described above, if the interferer is to *not* coincide at IF with the

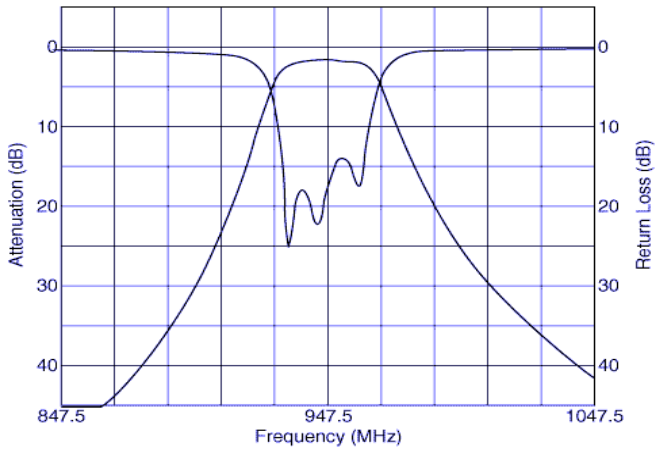


Fig. 4. Typical RF prefilter.

desired signal, the following must hold, respectively, for low-side LO injection where $f_{LO} = f_{RF} - f_{IF}$, and for high-side LO injection where $f_{LO} = f_{RF} + f_{IF}$:

$$f_{IF} \neq \frac{n f_{RF} - k f_{interf}}{n \pm 1} \quad or \quad f_{IF} \neq \frac{k f_{interf} - n f_{RF}}{n \mp 1} \qquad (11)$$

By stepping through all GSM channels in the range 935 MHz MHz, this leads to *undesired ranges* of f_{IF} shown in Table I for various *n* and *k*.

TABLE I. UNDESIRED F_{IF} REGIONS (MEGAHERTZ) DUE TO MECHANISM 1

	Low-Side Injection		**High-Side Injection**	
	k=2	*k*=3	*k*=2	*k*=3
n=2	13 - 47 40 - 140	-	13 - 57 40 - 170	265 - 292 340 - 377
n=3	171 - 285	15 - 53 30 - 105	-	15 - 68 30 - 135
n=4	340 - 376 382 - 412	148 - 180 199 - 225 247 - 300 332 - 375	-	-
n=5	-	269 - 310 322 - 355	-	-

The undesired responses invoked by the special case of $|n - k| = 1$ have been called "Able-Baker spurs" [6].

Mechanism 2: Mixing of Interferer With LO Harmonic, m = k = 1, n > 1: Even if the RF front-end is perfectly linear, LO harmonics in the mixer will downconvert interferers lying at $f_{interf} = n f_{LO} + f_{IF}$ to IF (Fig. 5). For the receiver to function normally in the presence of these interferers, the following inequality must hold:

$$P_{interf} - Att_{BPF}(\Delta f) - Att_{LNA}(\Delta f) - L_n < P_{des} - \frac{C}{I} \qquad (12)$$

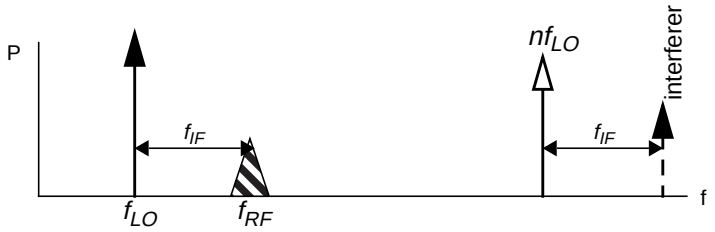


Fig. 5. Harmonic Mixing *m=k=1, n>1* (mechanism 2).

In this case, since the interferers of concern lie far away from the frequency band to which the receiver is tuned, they are subject to the full stopband loss of the RF prefilter and then further attenuation in the LNAs tuned load. GSM allows out-of band interferers to be as large as 0 dBm at the antenna; that is, 102 dB larger than the minimum detectable desired signal. Assuming that a practical preselect filter (Fig. 4) offers a stopband loss of 55 dB at best, the interferers must be suppressed roughly by another 55 dB on chip. The tuned input matching network and LNA load will each be partly responsible for this attenuation. For example, for $n = 3$ and an LO of 1 GHz with a reasonable IF of, say, 100 MHz, an LNA tuned by an on-chip inductor of moderate Q and impedance-matched at its input suppresses the interferer in question by more than 20 dB relative to the wanted signal. Then the interferer mixes with the third LO harmonic, which is one third the amplitude of the LO fundamental, and this translates the interferer to IF with one third (-10 dB) the conversion gain seen by the desired signal. Another 25 dB of attenuation is still needed. If this is not forthcoming, the interferer will jam the desired signal, and then can only be tolerated as one of the exceptions granted by GSM [7]. However, more than a few such interferers will soon exhaust the quota of exceptions.

With high-side LO injection, the interferers in question will lie yet further out in frequency. However, the filters described above will attenuate them by almost the same amount because, in practice, far away stopband loss tends to flattens out at some maximum value. In a superheterodyne receiver which uses off-chip filters, this problem is solved by placing an image reject filter, which is identical to the preselect filter, after the LNA. Now the cascaded stopband loss of the two filters pushes down all out-of-band interferers by more than 100 dB relative to pass-band signals.

It is clear that unless the IF is exceptionally high, mechanism 2 is largely independent of IF. Also that the classic image in wireless reception is a special case of this mechanism, when $n = m = k = 1$.

Mechanism 3: Subharmonics of IF, $k = n = 1$, $m > 1$: Interferers may be shifted onto the desired signal because of nonlinearity in the IF circuits. An interferer at $f_{LO} + f_{IF}/m$ mixes with the LO and downconverts to f_{IF}/m. Then, subject to IF circuit nonlinearity, its mth harmonic appears at f_{IF} (Fig. 6).

Depending on m, if the interferers in question initially lie in the RF filter's transition band they are only slightly attenuated before entering the receiver. For normal operation, the minimum linearity in the receiver's IF section, as characterized by intercept point, is

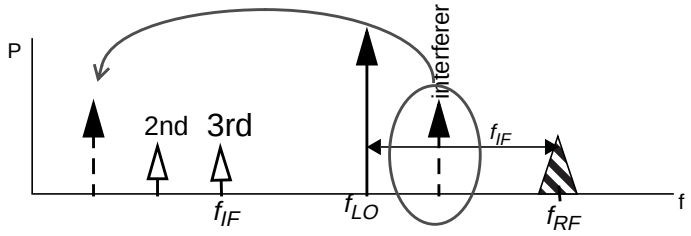


Fig. 6. Downconverted interferer appears at a subharmonic of IF (mechanism 3).

$$IIP_{m,\,IF} > \frac{m}{m-1}P_{interf} - L_{BPF} - \frac{m}{m-1}Att_{BPF}(\Delta f) - \frac{m}{m-1}\cdot Att_{LNA}(\Delta f) - 10\log(m) - \frac{1}{m-1}\left(P_{des} - \frac{C}{I}\right) + G_{mxr} + G_{LNA} \quad (13)$$

This expression imposes lower bounds on the IF. Once again, let us see how this mechanism works with GSM blocking signals. The IF circuits are usually fully differential; as a result they will largely suppress even-order harmonics. Based on circuit simulations, we assume an IIP_3 of +10 dBm at the IF input to the receiver, and supposing a 3% FET mismatch an IIP_2 of +40 dBm in the IF differential circuits. Then with 935 MHz < f_{RF} < 960 MHz, Table II presents the lower bounds on where the receiver is immune to this mechanism.

TABLE II. UNDESIRED f_{IF} REGIONS (MEGAHERTZ) DUE TO MECHANISM 3

	Low-Side Injection	**High-Side Injection**
m=2	0 - 150	0 - 160
m=3	0 - 113	0 - 105

Note that a blocker at $f_{LO} \pm f_{IF}/m$ can be subject to spurious mixing by the first mechanism with $n = k$, that is when m (mechanism 3) replaces k (mechanism 1). However, in the two cases the blocker appears at IF by entirely different processes. In the first mechanism, RF circuit nonlinearity shifts the interferer to a higher frequency, where it is attenuated by the tuned LNA load. In mechanism 3, however, the interferer is close to the RF band, and downconverts to a subharmonic of the IF.

The well-known half-IF interferer [3], $f_{LO} \pm f_{IF}/2$, is a special case of mechanisms 1 and 3.

With this background, let us study some previously published GSM receivers to understand their vulnerability to the spurious mixing of large interferers. It should be noted that the authors of these works make no mention of what steps they took, if any, to lessen their receivers' exposure to these vulnerabilities.

A. Example: GSM Receiver With IF of 71 MHz

This receiver [1] uses a passive RF filter before and after the LNA to cascade the stopband losses, and therefore strongly suppress out-of-band blockers from entering the mixer. Suppose that in the interest of higher integration, we remove the filter after the LNA. First consider mechanism 1. Assume the desired signal lies at 935 MHz, and that the LO is high-side injected at 1006 MHz. With $n = 3$, $k = 3$, $\mathrm{IIP}_{3,\mathrm{LNA}}$ = -7.5 dBm, L_3 = 10 dB, P_{des} = -99 dBm, P_{interf} = 0 dBm, L_{BPF} = 3 dB, C/I = 12 dB, and an approximate attenuation of 15 dB at $3f_{interf}$ due to the LNA load, it follows from (10) that the RF prefilter must attenuate an interferer at 983 MHz ($f_{LO} + f_{IF}/(3)$) by at least 28 dB. This is not possible with today's prefilters (Fig. 4).

Next consider mechanism 3 with $m = 3$. An interferer at 997 MHz suffers almost no attenuation through the LNA. Using (13) with Att_{BPF} = 26 dB, Att_{LNA} = 0 dB, G_{LNA} = 15.3 dB and G_{mxr} = 10 dB, the required $\mathrm{IIP}_{3,\mathrm{IF}}$ > +34 dBm. This is too high for a practical on-chip IF circuit, which means that the receiver is also vulnerable to blocking by this interferer.

These calculations show that with a 71 MHz IF a post-LNA filter is mandatory, and a GSM receiver using this IF cannot be fully integrated.

B. Example: GSM Receiver With Low IF of 100 kHz

Suppose the low-IF receiver for DCS-1800 [9] is re-tuned to 900 MHz GSM. The blocking templates for both standards are very similar. With the desired signal at 950 MHz, and assuming high-side injection f_{LO} = 950.100 MHz, mechanism 3 exposes the receiver to an interferer at $f_{LO} + f_{IF}/m$, for example, with $m = 2$, at 950.150 MHz or 950.050 MHz. However GSM limits the strength of interferers within 600 kHz of the desired signal to less than -43 dBm. The resulting downconverted blocker on to the desired channel is inconsequential.

If mechanism 1 is at work, and assuming $n = k$, the candidate interferer lies at $f_{LO} \pm f_{IF}/k$. This amounts to merely replacing m above by k, and again the pertinent interferers pose no threat. For example, assuming comparable $\mathrm{IIP_k}$ of the LNA as the one used in the receiver described in [8], the ranges of undesired IF are similar to the ones in Table I, and therefore the low IF is safe.

However, for mechanism 2 and with $n = 3$, an interferer 0 dBm strong at roughly 2850 MHz + 100 kHz ($f_{interf} = nf_{LO} + f_{IF}$) might overwhelm the desired signal. As it propagates through the receiver, this interferer is attenuated in the preselect filter by, say, Att_{BPF} = 60 dB. It is further attenuated in the tuned load of the LNA. If this load is a parallel LCR circuit with quality factor Q and assuming that $f_{IF} << f_{LO}$,

it follows that its gain to an interferer close to nf_{LO} compared to its gain to the wanted signal close to f_{LO} is in the ratio *1 : nQ*. With a *Q* of, say, five this amounts to Att_{LNA} = 23 dB. In mixing with the third LO harmonic whose L_3 = 10 dB, the left-hand side of (12) amounts to -93 dBm while the right-hand side may add up to -111 dBm. This violates the inequality in (12), exposing a vulnerability.

4. Experimental Validation on a Receiver with Optimal IF

The results of the analysis above were used to develop the frequency plan of a fully integrated 900-MHz CMOS GSM receiver [8]. The prototype receiver was fabricated in 0.35-um CMOS. Considerations of flicker noise at the output of the first mixer lead to the choice of *dual conversion to low-IF*. Given the GSM template of allowed channel strengths, it is relatively easy to choose the second IF close to 100 kHz because now the image need only be rejected by about 20 dB. However, what of the first IF? This is where the analysis above gives guidance.

According to the analysis, with high-side LO injection a first IF in the range of 170-265 MHz is immune to interference caused by mechanisms 1 and 3 (Tables I and II). It turns out that for lowside LO injection there is no "good" IF. Interest–ingly, Rohde and Bucher in their textbook [4] have come to the same conclusion. On page 112, they note that generally "the difference mixer with high-side oscillator provides fewer spurious responses."

Mechanism 2 remains a problem. The responsible interferers cannot be adequately rejected by the prefilter alone, and must be additionally suppressed by the frequency-dependent gain in the LNA. As was explained earlier in the example of the low IF receiver, the antenna prefilter and the LCR tuned load of the LNA do not adequately suppress the interferers that GSM allows at 3 x the frequency of the receive band (3420 x 190 MHz in this case). Therefore, a notch in the LNA frequency response must be realized at these frequencies. The most convenient way to do so is by resonating a capacitor in parallel with the LNA FETs source degeneration inductor at the desired notch frequency (Fig. 7). This inductor is realized with a bondwire [8], which resonates at around 3 GHz with the capacitance of the bondpad and PN junction present at the source terminal. Simulations suggest that with the high *Q* of the bondwire inductance, the notch may be as much as 100 dB below the gain peak.

The receiver uses a high-side LO and a first IF of about 190 MHz [8]. The second IF is selected to be 120 kHz. This choice of IFs gives a simple integer ratio of 6 x between the required LO frequencies at the two stages of downconversion, and allows both to be synthesized with a single VCO and a simple divider.

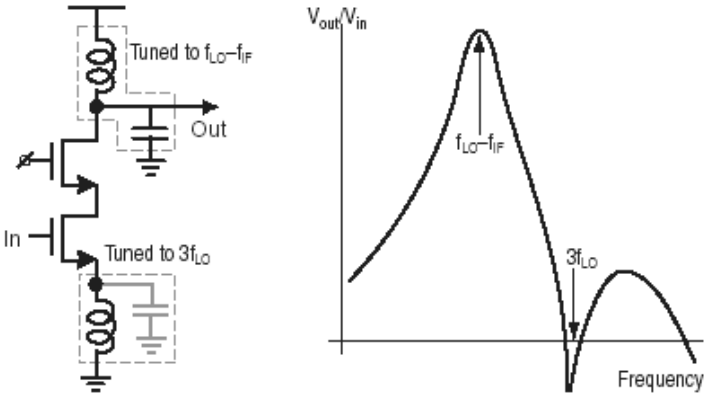


Fig. 7. LNA circuit with notch filter to suppress third harmonic, and associated frequency response.

The prototype receiver circuit was subjected to a comprehensive range of GSM blockers. The test consists of applying a tone at each of the frequencies listed in Table III, accompanied by a reference signal 3 dB above the sensitivity level in the middle of the GSM receive band. The amplitude of the interferer tone is raised until

TABLE III. MEASURED BLOCKING SIGNAL REJECTION FOR A GSM RECEIVER [8]

Physical Effect	**RF Input**	**Frequency (MHz)**	**On-chip Rejection (dB)**	**Rejection Incl. Prefilter**
1st Image	$(f_{LO}+f_{IF1})$	1328	56	111
Half IF	$(2f_{LO}+f_{IF1})/2$	1233	81.5	131
	$(2f_{LO}-f_{IF1})/2$	1043	72	114
	$(3f_{LO}+f_{IF1})/2$	1802	>91	141
	$(3f_{LO}-f_{IF1})/2$	1612	>109	159
Half IF	$(4f_{LO}+f_{IF1})/2$	2371	>93	148
	$(4f_{LO}-f_{IF1})/2$	2181	>93	148
HD3 in balanced IF1 circuits	$(2f_{LO}+f_{IF1})/3$	822	97	142
	$(2f_{LO}-f_{IF1})/3$	695	71	126
	$(3f_{LO}+f_{IF1})/3$	1201	76	126
	$(3f_{LO}-f_{IF1})/3$	1075	56	106
	$(5f_{LO}+f_{IF1})/3$	1960	>116	171
	$(5f_{LO}-f_{IF1})/3$	1833	>116	171
	$(3f_{LO}+f_{IF1})$	3604	106	161
	$(3f_{LO}-f_{IF1})$	3224	>103	158

its in-band spurious response overtakes the reference signal. Measurements show that the receiver rejects all potential interferers, including the image signal, by more than 105 dB. An on-chip circuit in this receiver attenuates the image channel, $f_{LO} + f_{IF}$, at the first IF. It also suppresses image channels downconverted by the LO harmonics, of which the principal one is $3f_{LO} - f_{IF}$ - the sign reversal is due to the fact that the quadrature sequence is reversed for the third harmonic of the LO.

These results show the practical outcome of the analysis in the design of an integrated receiver robust to interferers.

5. Conclusion

This paper presents a systematic and comprehensive analysis of how circuit imperfections cause adjacent and out-of-band signals incident on a wireless receiver to translate to the same frequency as the desired signal. The comprehensive analysis reveals mechanisms in wireless receivers other than the well–understood intermodulation distortion, whereby adjacent and out-of-band signals become "interferers." Some of these signals, although initially distant from the wanted channel but nu–merically related in frequency, mix via spurious mechanisms on to it at IF and interfere with its detection.

Knowing the profile of adjacent channel strengths, characteristics of the antenna filter, and the typical nonlinearity in the RF circuits, we can predict the strength of spurious mixing products. In some cases, such as mechanism 1 and 3, the only way to counter the resulting problems is by proper choice of intermediate frequency, IF. Mechanism 2 is often most important, since here, blocker suppression depends exclusively on filter attenuation and LO harmonics.

In a fully integrated receiver which dispenses with fixed external filters to save power and physical volume, the problems of spurious mixing are particularly pressing. Based on this analysis, a fully integrated GSM receiver with a unique frequency plan was developed. Measurements on the prototype, which uses only one off-chip filter for RF preselection, verify that it suppresses all blockers by more than 105 dB.

We expect that an awareness of spurious mixing will add yet another criterion that guides the choice of receiver architectures in the future.

REFERENCES

[1]. P. Orsatti, F. Piazza, and Q. Huang, "A 20-mA-receive, 55-mA-transmit, single-chip GSM transceiver in 0.25-um CMOS," *IEEE J. Solid State Circuits*, vol. 34, pp. 1869-1880, Dec. 1999.

[2]. T. Yamawaki, M. Kokubo, K. Irie, H. Matsui, K. Hori, T. Furuya, H. Shimizu, M. Katagishi, T. Endou, B. Henshaw, and J. R. Hildersley, "A dual-band trans-

ceiver for GSM and DCS1800 applications," *Proc. Eur. Solid-State Circuits Conf. (ESSCIRC)*, pp. 84-87, 1998.

[3]. 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 J. Solid State Circuits*, vol. 30, pp. 1421-1429, Dec. 1995.

[4]. U. Rohde and T. Bucher, *Communications Receivers: Principles and Design*. New York: McGraw-Hill, 1988.

[5]. R. C. Sagers, "Intercept Point and Undesired Responses," *IEEE Trans. Veh. Technol*., vol. VT-32, pp. 121-133, Feb. 1983.

[6]. K. Hansen and A. Nougeras, "Receiver RF design considerations for wireless communications systems," *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, pp. 93-96, 1996.

[7]. *GSM Technical Specification 05.05*, European Telecommunications Standards Institute, 1992.

[8]. S. Tadjpour, E. Cijvat, E. Hegazi, and A. Abidi, "A 900 MHz dual con–version, low-IF GSM receiver in 0.35-um CMOS," *IEEE J. Solid State Circuits*, vol. 36, pp. 1992-2002, Dec. 2001.

[9]. M. S. J. Steyaert, J. Janssens, B. de Muer, M. Borremans, and N. Itoh, "A 2-V CMOS cellular transceiver front-end," *IEEE J. Solid State Circuits*, vol. 35, pp. 1895-1907, Dec. 2000.

PAPER IV

PAPER IV

A Fully Integrated 2.45 GHz 0.25 um CMOS Power Amplifier

Abstract

A fully integrated differential class-AB power amplifier has been designed in a 0.25um CMOS technology. It is intended for medium output power ranges such as Bluetooth class I, and has an operating frequency of 2.45GHz. By using two parallel output stages that can be switched on or off, a high efficiency can be achieved for both high and low output power levels. The simulated maximum output power is 22.7 dBm, while the maximum power-added efficiency is 22%.

Based on: E. Cijvat and H. Sjöland, "A Fully Integrated 2.45 GHz 0.25 μm CMOS Power Amplifier". © 2004 IEEE. Reprinted, with permission, from *Proceedings of the 10th IEEE International Conference on Electronics, Circuits and Systems (ICECS) 2003*, pp. 1094-1097. Sharjah, United Arab Emirates, December 2003.

1. Introduction

With the recent emergence of short-range communication standards such as Bluetooth, the research interest for highly integrated power amplifiers (PAs) has increased [1-13]. For frequencies up to several GHz and low to medium output power, CMOS may be an alternative to stand-alone power amplifiers, offering a higher level of integration in a relatively cheap technology, in exchange for less efficiency and a lower maximum output power.

In most communication systems transmitter output power control is required. In order to increase the battery lifetime, it is important to have a relatively high efficiency over the whole PA output power range, i.e. for both lower and higher output power, since the PA is more likely to operate at lower than higher output power.

For the Bluetooth standard the highest output power is 20 dBm (class I, [14]) which is feasible for CMOS implementation (see [1-13]). Moreover, a constant envelope modulation scheme is used, implying that linearity of the PA is not a critical issue for this standard.

In this work a class-AB power amplifier is described that consists of two stages, with the output stage comprising two parallel blocks that may be switched *on* or *off* (see fig. 2). In this way the efficiency may be optimized for different output power settings. The output impedance transformation network is fully integrated.

The paper is structured as follows: First some PA theory is described, then the design itself is presented. Simulation results are shown in Section 4, and finally conclusions are presented.

2. Analysis

In fig. 1 a current source with impedance transformation network *T* is shown. This serves as a model for an ideal output stage, where the transistor operates as a controlled current source driving R_{opt}, the transformed load impedance R_L. The maximum signal amplitude is V_{DD}, and the ideal maximum output power is given by

$$P_{out,\,max,\,ideal} = \frac{V_{DD}^2}{2R_{opt}} \qquad (1)$$

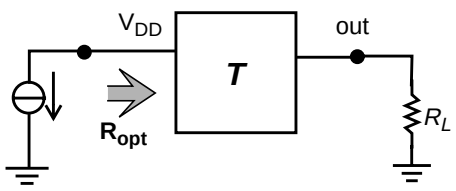


Figure 1. *Principle of impedance transformation.*

Thus, for a given V_{DD}, R_{opt} determines $P_{out,max}$ assuming a maximum voltage swing. For $P_{out,max}$ = 22 dBm and V_{DD} = 3.3 V, the optimum load resistance R_{opt} is equal to 34 Ω.

The power added efficiency (PAE) is defined as

$$PAE = \frac{P_{out} - P_{in}}{P_{DC}} \qquad (2)$$

where P_{DC} is the power supplied by the battery, which is signal-dependent for most types of PAs. The PAE typically is maximum for an output power close to $P_{out,max}$ and decreasing fast for lower output powers [5]. Therefore, for a high average efficiency, the PA may be designed to have different $P_{out,max}$ by changing R_{opt} or V_{DD}. In this work the former strategy is used.

Non-idealities

For high voltage swings the transistor will enter the linear region, and no longer behave as an ideal current source. This is commonly modeled with the knee voltage V_{knee} [15]. The output voltage swing is reduced to V_{DD} - V_{knee}, and the maximum output power may be written as

$$P_{out,\,max} = \frac{(V_{DD} - V_{knee})^2}{2R_{opt}} \qquad (3)$$

Assuming an output stage with an integrated current-supplying inductor and a switch at V_{DD}, allowing the whole stage to be turned *on* or *off*, other non-idealities may be identified such as the series resistance of the inductor and the on-resistance of the supply switch, both reducing the bias voltage at the drain. Moreover, the finite output impedance of the transistor and the finite quality factor *Q* of the passives in the matching network will cause power loss. To compensate for these losses, a PA is generally designed for a higher $P_{out,max}$ by reducing R_{opt} in Eq. 3.

3. Design

A fully integrated 2.45 GHz PA was implemented in a 0.25um CMOS technology. In order to implement different output power settings and increase the average efficiency, two parallel output blocks were used (see fig. 2). Due to die size considerations the number of parallel blocks was limited to two.

The matching network was chosen so that a maximum output power of 22 dBm was achieved with both blocks *on*. When one stage is *off*, the matching network provides a higher R_{opt} and thus a lower $P_{out,max}$. In this way a relatively high average efficiency over the total PA output power range can be achieved.

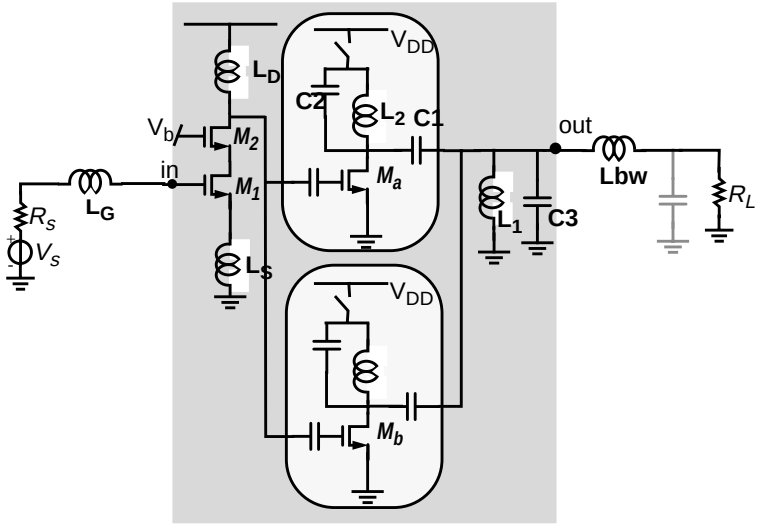


Figure 2. *Simplified single-ended schematic of the PA with two parallel output stages.*

For the output stage no cascode transistors are used, since this would increase the knee voltage and thus decrease $P_{out,max}$. A switch in the signal path at the output would also have a negative impact on the output power and thus on the efficiency, and was therefore avoided as well.

The input capacitance of the output stage forms a large part of the total capacitance at the drain of M_2 which is parallel to L_D, see fig. 2. When switching *off* one stage (by changing the gate bias to V_{DD} and opening the V_{DD} switch), the input capacitance will change significantly. In order to decrease the impact on input stage tuning, a rather small AC coupling capacitance is used between the input- and output stage.

By connecting the two output nodes as shown in fig. 2, the two parallel stages share L_1, C_3 and L_{bw}. Only C_1, L_2 and C_2 are available to separately design R_{opt} for each stage. Moreover, due to numerous parasitic capacitances and size restrictions of integrated passives, the impedance transformation ratio cannot be varied over a wide range.

Generally when parallel output stages are used, power combining is implemented either through a transformer [3] or transmission lines [5]. In this design, however, the two stages are not isolated, meaning that the network of one stage has an impact on the impedance transformation of the other.

The matching network for one output stage may be drawn as shown in fig. 3. The two output stages are connected at point *P.*

When for example the lower stage in fig. 2 is switched *off,* FET M_b is brought into the linear region, thus providing a low impedance. Assuming this to be a short, one can see from fig. 2 that the equivalent capacitance parallel to C_3 will be larger.

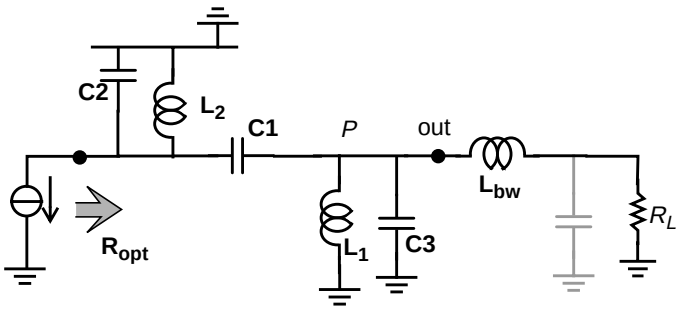


Figure 3. *The matching network for one stage, with the FET represented as an ideal current source.*

This decreases the transformation ratio for the upper stage, thus increasing R_{opt}, which is desirable when only one stage is *on*.

The two stages are unequal, having different FET widths and different capacitance and inductance values, and thus different transformation ratios and gains.

Comparing fig. 2 with fig. 3 one can see that the matching network includes parasitic capacitance at the drain of the FET (incorporated in C_2), as well as the on-chip output node (in C_3).

The 5M1P 0.25um CMOS technology offers thick-metal inductors with quality factors ranging from approximately 5 to 15. For L_2 and L_S inductors provided by the manufacturer were used. L_D and L_1 were designed using Fast Henry [16] and ASITIC [17]. For the integrated matching network, MOM (metal-oxide-metal) capacitors with highest quality (Q) factor available in this technology were used. The FETs M_a and M_b in fig. 2 do not have minimum gate length, but 0.32μm, and have a higher breakdown voltage.

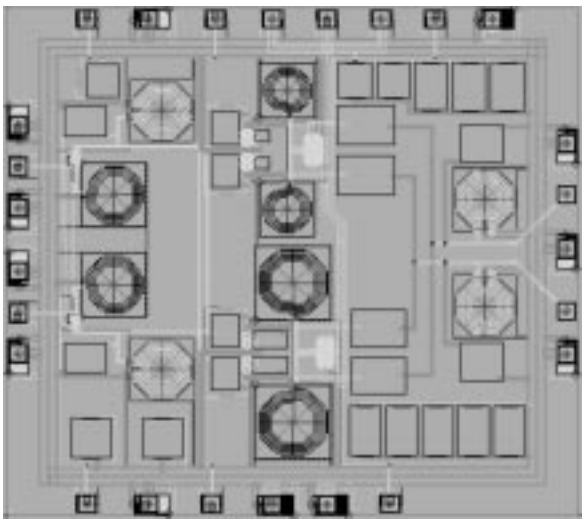


Figure 4. *Plot of the PA layout.*

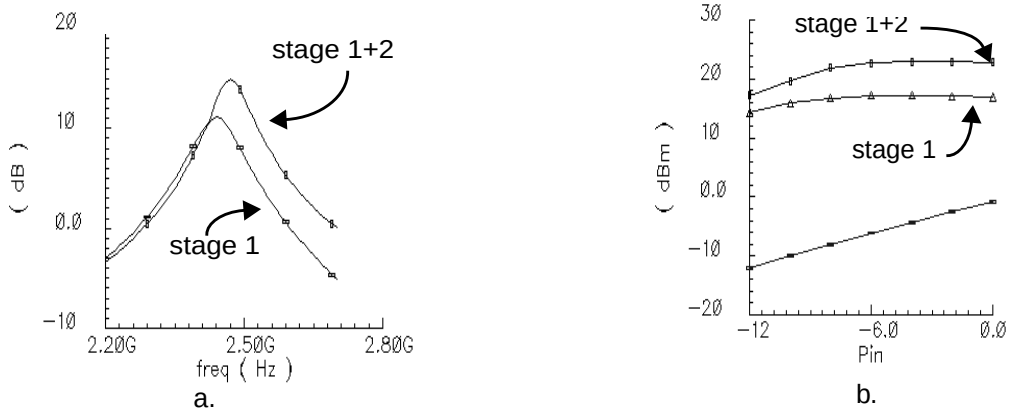


Figure 5. *Simulation results, a). Frequency response, b).* P_{out} *as a function of* P_{in}.

4. Simulation Results

The above described design was simulated using SpectreRF, with BSIM3v3 models. Post-layout parasitics were taken into account. The layout is shown in fig. 4. A large area is occupied by the integrated passives, and a substantial area saving may be achieved by using differential inductors [18].

In fig. 5 some simulation results are shown. The maximum PAE (22 %) is achieved for P_{out} slightly below $P_{out,max}$ (which is about 23 dBm, see fig. 5.b). The center frequency for both cases (stage 1+2 and stage 1 only) is approximately 2.45GHz. The simulation results are summarized in table 1.

TABLE I. SIMULATION RESULTS, SUMMARY

PAE_{max}	22%	
P_{out} (differential)	22.7 dBm	@2.45GHz @PAE_{max}
power gain	28 dB	@PAE_{max}
total die area	6.25 mm^2	including pads
P_{out} (stage1 only)	17.2 dBm	PAE=15.8%
P_{out} (stage 2 only)	12.4 dBm	PAE=6.1%

From table 2 a comparison can be made between this PA and previously published work. It can be seen that the PA presented in this work performs quite well, given the limitations of an on-chip matching network and class-AB. Moreover, the PA includes measures to improve the average efficiency.

TABLE II. COMPARISON OF MEDIUM POWER PAS

	Pout (dBm)	frequency (GHz)	efficiency (max)	technology	output matching	power control	class
[1]	15 (@P_{-1dB})	0.9	<30% (η)	1um CMOS	off-chip	steps, 5dB	C
[2]	23.5 (max)	1.9	35% (PAE)	0.35 um (Bi)CMOS	off-chip	-	AB
[3]	33.4 (@max PAE)	2.4	31% (PAE)	0.35um (Bi)CMOS	on-chip	-	E/F3
[4]	28.2	1.9	30% (PAE)	30GHz BiCMOS	off-chip	-	AB
[5]	24.8 (max)	1.4	49% (PAE)	0.25um CMOS	off-chip (transm. lines)	3 parallel stages	F
[6]	31.8 (max)	0.9	43% (PAE)	0.2um CMOS	off-chip	-	F
[7]	30 (@max. PAE)	1.8	45% (PAE)	0.35um CMOS	off-chip	-	AB
[8]	20 (max)	1.9	16% (η)	0.8um CMOS	on-chip	-	F?
[8]	22 (max)	2.4	44% (η)	0.25um CMOS	off-chip	-	F?
[9]	17.5 (max)	2.4	16.4% (PAE)	0.35um CMOS	partly on-chip	-	A
[10]	23 (max)	2.4	42% (PAE)	0.18um CMOS	off-chip	-	AB
[11]	30 (max)	0.7	62% (PAE max)	0.35um CMOS	partly on-chip	-	E
[12]	9 (@P_{-5dB})	2.4	16% (P_{-5dB})	0.18um CMOS	partly on-chip	-	AB?
[13]	18.6 (@max PAE)	0.9	30% (PAE max)	0.6um CMOS	on-chip	-	C
this work	22.7 (@max PAE)	2.45	22% (PAE max)	0.25um CMOS	on-chip	2 settings	AB

5. Conclusions

A 2.45 GHz power amplifier has been designed in a 0.25um CMOS technology. The PA is fully integrated, including output matching network. Simulations show that an output power of 22.7 dBm may be achieved with a maximum PAE of 22%. The average efficiency has been improved by using two parallel output stages.

ACKNOWLEDGEMENTS

The authors are grateful to Niklas Troedsson, MSc (Dept. of Electroscience, Lund University) for help with integrated inductor design and modeling.

REFERENCES

1. M. Rofougaran, A. Rofougaran, C. Ølgaard, and A.A. Abidi, “A 900 MHz CMOS RF Power Amplifier with Programmable Output”, *1994 Symposium on VLSI Circuits Digest*, pp. 133-134, June 1994.
2. A. Giry, J-M. Fournier, and M. Pons, “A 1.9GHz Low Voltage CMOS Power Amplifier for Medium Power RF Applications”, *Proceedings of 2000 IEEE Radio frequency Integrated Circuits (RFIC) Symposium*, pp. 121-124, May 2000.
3. I. Aoki, S.D. Kee, D.B. Rutledge, and A. Hajimiri, “Fully Integrated CMOS Power Amplifier Design using the Distributed Active-Transformer Architecture”, *IEEE Journal of Solid-State Circuits*, Vol. 37, no.3, March 2002.
4. S. Luo and T. Sowlati, “A monolithic Si PCS-CDMA Power Amplifier with 30% PAE at 1.9GHz using a Novel Biasing Scheme”, *IEEE Transactions on Microwave Theory and Techniques*, Vol. 49, no. 9, September 2001.
5. A. Shirvani, D.K. Su, and B.A. Wooley, “A CMOS RF Power Amplifier with Parallel Amplification for Efficient Power Control”, *IEEE Journal of Solid-State Circuits*, Vol. 37, no. 6, June 2002.
6. T.C. Kuo and B.B. Lusignan, “A 1.5W Class-F RF Power Amplifier in 0.2um CMOS Technology”, *2001 International Solid-State Circuits Conference (ISSCC) Digest*, pp. 154-155, February 2001.
7. C. Fallesen and P. Asbeck, “A 1W 0.35um CMOS Power Amplifier for GSM-1800with 45% PAE”, *2001 International Solid-State Circuits Conference (ISSCC) Digest*, pp. 158-159, February 2001.
8. Y.J.E. Chen, M. Hamai, D. Heo, A. Sutono, S. Yoo, and J. Laskar, “RF Power Amplifier Integration in CMOS Technology”, *2000 IEEE Microwave Symposium (MTT-S) Digest*, pp. 545-548, May 2001.
9. C.C. Hsiao, C-W. Kuo, and Y-J. Chan, “Integrated CMOS Power Amplifier and Down-Converter for 2.4GHz Bluetooth Applications”, *Proceedings of the 2001 IEEE Radio and Wireless Conference (RAWCON)*, pp. 29-32, August 2001.
10. T. Sowlati and D. Leenaerts, “A 2.4GHz 0.18um CMOS Self-Biased Cascode Power Amplifier with 23 dBm Output Power”, *2002 International Solid-State Circuits Conference (ISSCC) Digest*, pp. 294-295, February 2002.

11. K. Mertens and M.S.J. Steyaert, “A 700MHz 1-W Fully Differential CMOS Class-E Power Amplifier”, *IEEE Journal of Solid-State Circuits*, Vol. 37, no. 2, February 2002.
12. K. Yamamoto, T. Heima, A. Furukawa, M. Ono, Y. Hashizume, H. Komurasaki, S. Maeda, H. Sato, and N. Kato, “A 2.4GHz-Band 1.8-V Operation Single-Chip Si-CMOS T/R-MMIC Front-End with a Low Insertion Loss Switch”, *IEEE Journal of Solid-State Circuits*, Vol. 36, no. 8, august 2001.
13. R. Gupta, B.M. Ballweber, and D.J. Allstot, “Design and Optimization of CMOS RF Power Amplifiers”, *IEEE Journal of Solid-State Circuits*, Vol. 36 no. 2, February 2001.
14. The Bluetooth Special Interest Group, *“Specification of the Bluetooth System”*, Vol. 1, version 1.1. www.bluetooth.com, 2001.
15. S.C. Cripps, “*RF Power Amplifiers for Wireless Communication*”. Artech House, Norwood, MA, USA, 1999.
16. M. Kamon, M.J. Tsuk, and J. White, “FastHenry: A Multipole-Accelerated 3-D Inductance Extraction Program”, *Proc. of the 30th ASM/IEEE Design Automation Conference (DAC)*, pp. 678-683, 1993.
17. A.M. Niknejad, “Modeling of Passive Elements with ASITIC”, *Proc. of the 2002 IEEE Radio Frequency Integrated Circuits (RFIC) Symposium*, pp. 303-306, May 2002.
18. M. Danesh and J.R. Long, “Differentially Driven Symmetric Microstrip Inductors”, *IEEE Trans. on Microwave Theory and Techniques*, Vol. 50, no. 1, January 2002.

PAPER V

PAPER V

A Fully Integrated CMOS RF Power Amplifier with Internal Frequency Doubling

Abstract

The design and measurement results of a fully integrated 0.18μm 1P6M CMOS power amplifier with internal frequency doubling are presented. Two designs were measured, one stand-alone PA and one PA with a VCO on the same chip. Since the PA and VCO operate at different frequencies, this configuration is suitable for direct-conversion transmitters or low-IF upconversion as oscillator pulling will be reduced. The maximum output power is 15 dBm, and the maximum drain efficiency is 10.7% at an output operating frequency of 2.4 GHz. An analysis of the efficiency is given.

Based on: E. Cijvat, N. Troedsson and H. Sjöland, "A Fully Integrated CMOS RF Power Amplifier with Internal Frequency Doubling". Submitted to *Analog Integrated Circuits and Signal Processing.*

1. Introduction

In recent years, the wireless communications market has grown significantly. Systems such as GSM have matured, and new systems such as UMTS, Bluetooth and IEEE802.11 are now commercially available. In handsets the level of integration has increased, in order to reduce the number of chips per handset as well as the number of external components. In many cases CMOS is the technology of choice for all parts except the power amplifier.

However, for frequencies up to several GHz and low to medium output power, the research interest for highly integrated CMOS power amplifiers (PAs) has increased [1-7]. They may be an alternative to stand-alone power amplifiers, offering a higher level of integration in a relatively cheap technology, generally in exchange for less efficiency and a lower maximum output power.

When having the PA on the same chip as a VCO (voltage controlled oscillator), one of the problems occurring is oscillator pulling, i. e. the VCO oscillation frequency changes due to the PA output signal. This is a problem if the PA and VCO are operating at nearly equal frequencies. for instance in a direct-conversion or low-IF transmitter. Several strategies can be used to alleviate this problem, such as composing the LO signal from two VCOs in a direct conversion transmitter or by upconversion in several steps [8]. The approach used in this work is to have a built-in upconverter in the power amplifier by cancelling the fundamental signal and utilizing the second harmonic, thus making sure that the VCO and PA are operating at different frequencies.

In mixers, so-called even-harmonic conversion has been widely used [9-10], especially for millimeter-wave or receiver applications. With regard to power amplifiers the technique has been described and used previously [11-13], but has to the authors' knowledge not been used in CMOS integrated PAs.

For fully integrated power amplifiers, which include the output impedance transformation network, the limited quality factor of the integrated passives diminishes the efficiency. An efficiency analysis of the PA with transformation network as presented in this work will be given.

Since many wireless communication systems operate with an output power in the 10 - 20 dBm range for the handset, this PA was designed for an output power in that range, aiming at a high efficiency. Moreover, linearity was not considered to be a critical issue since many wireless systems use a constant envelope modulation scheme. Therefore the PA was biased in class C.

2. Power Amplifier Analysis

One of the frequency components automatically generated in a PA biased in non-class A is the second-order harmonic. Thus, internal frequency doubling may be achieved if this harmonic is large enough compared to other output harmonics. When the drains of two stages are tied together to the same load, as shown in Fig. 1a, the odd harmonics are cancelled if the inputs are driven with a differential signal. This may be referred to as a push-push configuration [11].

An important issue when using frequency doubling in the PA is signal contents. A narrowband FM input signal can be written as $x(t) = A\cos\omega_c t - Am\sin(\omega_c t)\int x_{BB}(t)dt$, where $m\int x_{BB}(t)dt \ll 1$. The second order output signal is given by:

$$x^2(t) \approx A^2\left[DC + \frac{1}{2}\cos(2\omega_c t) - m\sin(2\omega_c t)\int x_{BB}(t)dt\right] \qquad (1)$$

As can be seen, the 2nd order harmonic carries the same signal information as the fundamental. This approximation is valid for most current commercial mobile communication systems such as for instance GSM, DECT and Bluetooth. From the above equation it can also be seen that a phase shift in the input signal will result in a double phase shift in the output signal. Thus, if a quasi-differential amplifier is used in order to get a differential output signal, the second stage must be driven with a 90° phase shift relative to the first stage, as shown in Fig. 1b. The PA is thus driven with quadrature signals.

Ideal PA response

Assuming that the output stage of the PA can be modeled as an ideal current source with a maximum voltage swing of $2V_{DD}$ at the drain node, the DC, fundamental and

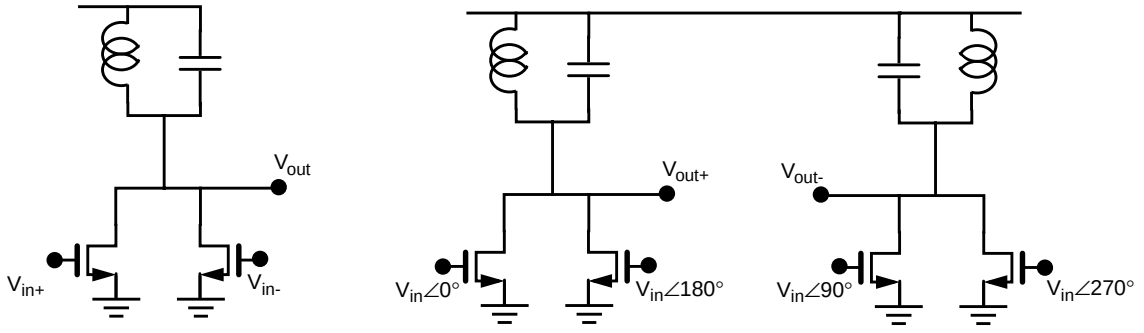


Fig. 1. a). A single-ended PA stage with cancelled odd harmonics, b). differential version.

2nd order response of the PA depend on the conduction angle or class in which the PA is biased [14]. The DC current is given by

$$I_{DC} = \frac{I_{max}}{2\pi} \cdot \frac{2\sin\alpha/2 - \alpha\cos\alpha/2}{1-\cos\alpha/2} \qquad (2)$$

where I_{max} is the maximum output current and α is the conduction angle (in radians).

The magnitude of the second order harmonic may be written as

$$I_2 = \frac{1}{\pi}\int_{-\alpha/2}^{\alpha/2} \frac{I_{max}}{1-\cos\alpha/2}[\cos\theta - \cos\alpha/2]\cos 2\theta d\theta \qquad (3)$$

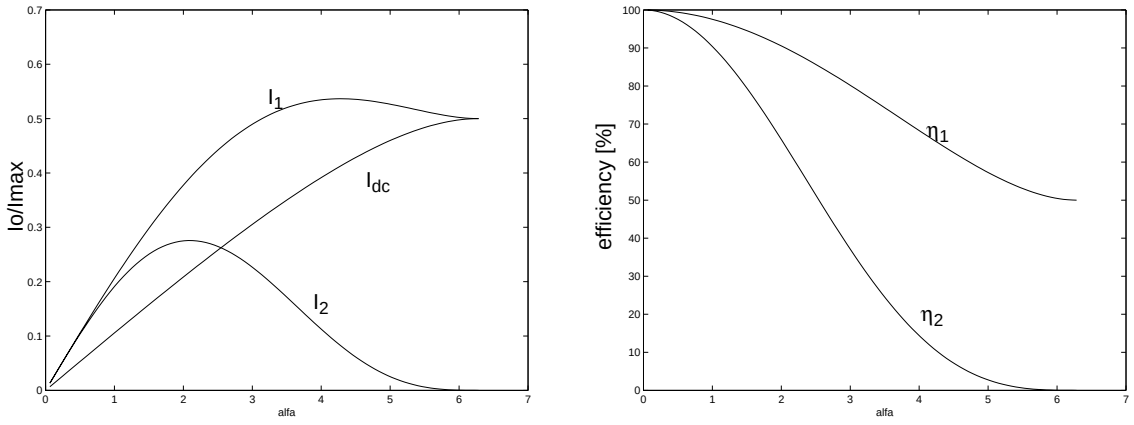
$$= \frac{I_{max}}{2\pi} \cdot \frac{1}{1-\cos\alpha/2}\left(-\frac{1}{3}\sin\frac{3\alpha}{2} + \sin\frac{\alpha}{2}\right)$$

This results in a maximum efficiency for the second order harmonic of

$$\eta_2 = \frac{P_2}{P_{DC}} = \frac{V_{DD}I_2}{2V_{DD}I_{DC}} = \frac{1}{2}\frac{\left(\sin\frac{\alpha}{2} - \frac{1}{3}\sin\frac{3\alpha}{2}\right)}{(2\sin\alpha/2 - \alpha\cos\alpha/2)} \qquad (4)$$

For this ideal PA both the fundamental and second order responses (output currents normalized to the maximum output current I_{max}), as well as the efficiencies η_1 [14] and η_2 are depicted in Fig. 2, as a function of the conduction angle α.

From Fig. 2 it can be seen that the 2nd order response is approaching the fundamental for small conduction angles, i.e. for a PA biased deep in class C. Moreover, the efficiency is high in that region, also for the 2nd order output signal. A disadvantage of deep class C is, however, that the maximum output power is relatively low.



a. b.

Fig. 2. a). Ideal normalized DC, fundamental and second order output current as function of the conduction angle, b). drain efficiency for the fundamental and second-order output.

In this deduction it is assumed that the PA does not have a so-called knee voltage V_{knee} [14], that only the relevant harmonic is seen at the output, and that the load resistance R_{opt} is always optimized for maximum voltage swing ($2V_{DD}$) with maximum current swing I_{max}.

Maximum output power

The maximum linear output power for an ideal PA, modeled as a current source with a certain drain bias voltage V_{DD}, is given by

$$P_{out,max} = \frac{V_{DD}^2}{2R_{opt}} \qquad (5)$$

where R_{opt}, the impedance seen at the drain, is the load resistance transformed by a transformation network (see Fig. 3) designed to achieve maximum current and voltage swing simultaneously.

The MOSFET DC current is supplied through L_1. The section formed by L_{bw} (the output bondwire inductance) and C_3 can be seen as a low-pass up-transformation L-section, and the section C_1 and C_2 as a down-transformation stage. In theory this provides us with enough degrees of freedom to ensure sufficient bandwidth while being able to freely choose the transformation ratio. However, both parasitics and size limitations for integrated passive components limit the impedance transformation ratio. Parasitic capacitances from the pad may be included in C_3, while parasitic capacitances at the drain are included in C_1. For the PA with internal frequency doubling the network must be tuned so that the desired impedance transformation is achieved at twice the PA input frequency.

Assuming all passive components are ideal, the output power will be the same as the power at the drain. For a linear output power of around 15 dBm (32 mW) and a supply voltage of 3.3V, R_{opt} becomes 172 Ω, implying an upward impedance trans-

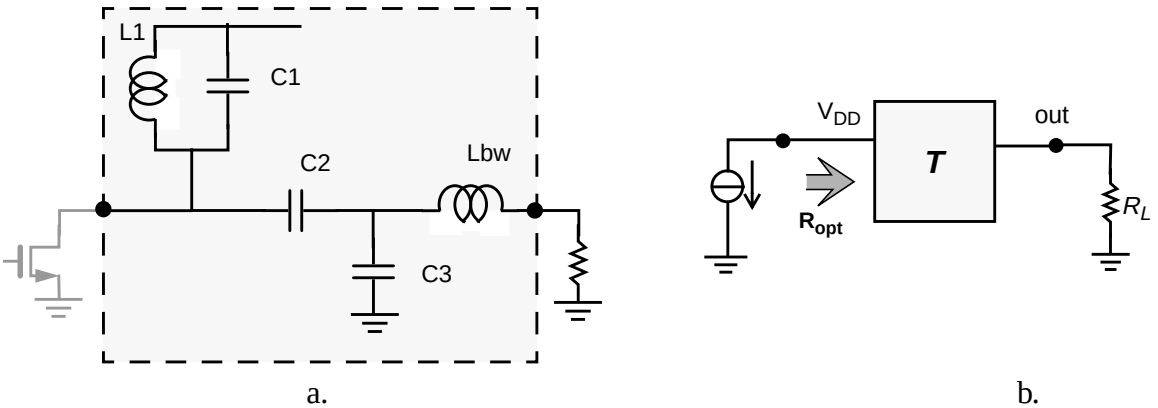


Fig. 3. a). A PA with transformation network (shaded area), b). an ideal model.

formation with respect to the standard 50 Ω load. However, this is highly dependent on both the desired output power and the supply voltage. Due to non-idealities in the PA and transformation network, R_{opt} must be chosen with some margin for power loss, i.e. a lower R_{opt} is chosen. In the next session we will look more into non-idealities of the transformation network.

3. Implementation

The quasi-differential PA as described in the previous section was implemented in a 0.18μm 1P6M CMOS technology with the option of 3.3V supply. Since we aimed at output powers of 15 - 20 dBm, this supply voltage was chosen, and thus non-minimum length MOSFETs were used in the final stage. One test chip included a passive polyphase filter, a pre-amplifier and the PA, as shown in Fig. 4, while a second test chip included both a quadrature VCO and the PA (consisting only of a final stage), as shown in Fig. 5.

The motivation for the circuit in Fig. 4 was to test the concept of internal frequency doubling, as well as the output power and efficiency of the final PA stage. The polyphase filter provides the driving stage with I- and Q- signals, but has a significant signal loss. The driving stage then amplifies the signal again.

For the VCO core in the circuit of Fig. 5 PMOSFETs were used, so that the VCO output voltage swings around ground, the desired bias voltage of the PA input. The two blocks can then be connected without buffers or coupling capacitors; a disadvantage is that this fixes the PA input bias voltage, while an advantage is that no

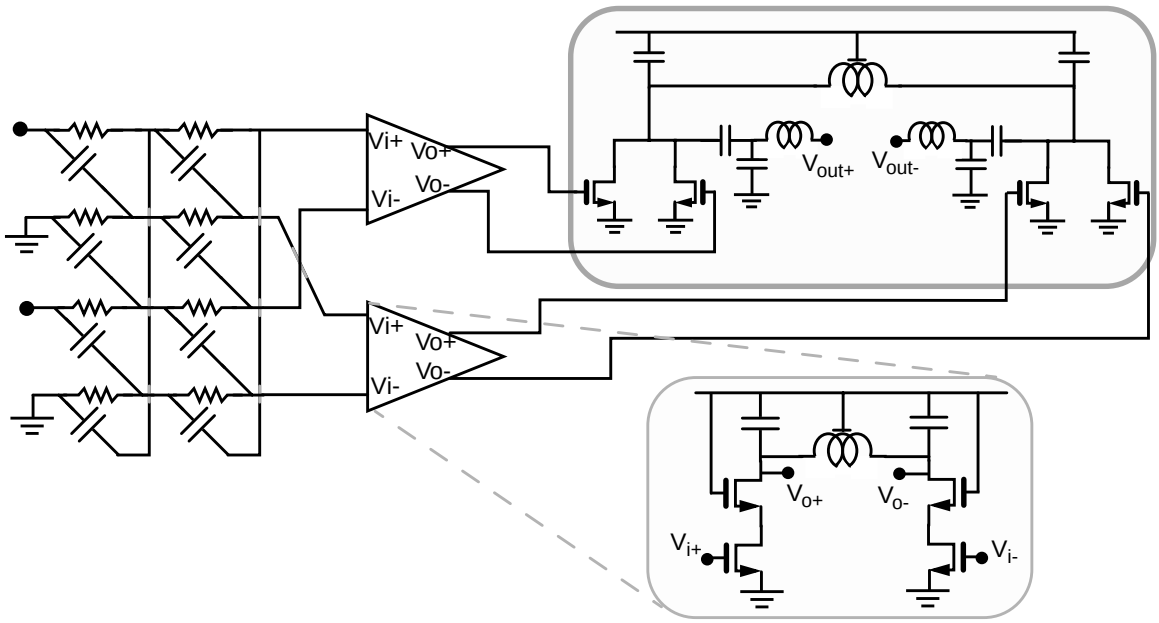


Fig. 4. Simplified schematic of the stand-alone PA including polyphase filter and driving stages.

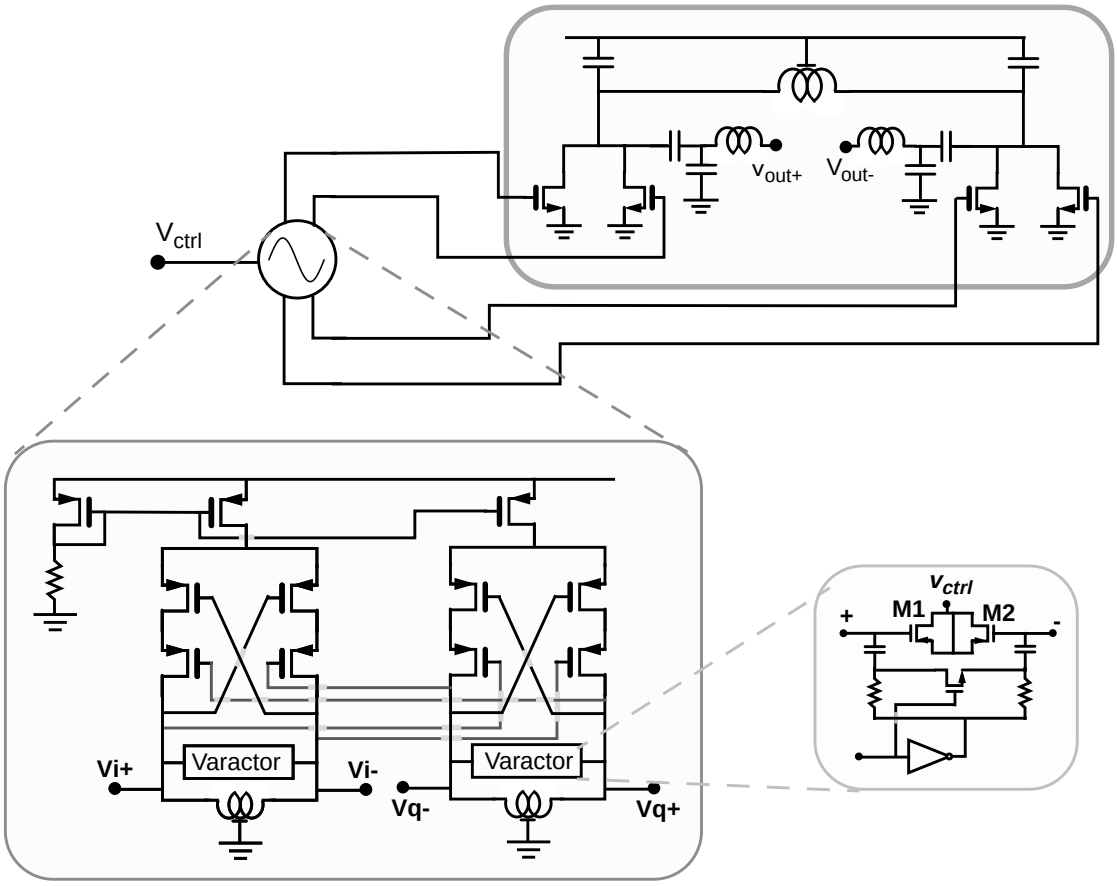


Fig. 5. Schematic of the VCO and PA, including the varactor with continuous and discrete tuning.

signal degradation will take place between the VCO and PA. The switched tuning of the varactor as shown in Fig. 5 has been described earlier [15].

Non-idealities of the PA implementation

The equation for the maximum output power, Eq. 5, is valid under the assumption that the transformation network is ideal (i.e. lossless), and that the PA transistor does not have a knee voltage V_{knee}. First we will look into the non-idealities of the final stage of the PA.

One of the most crucial non-idealities in any PA implementation is the knee voltage, i.e. the minimum drain voltage necessary to have the PA operating as an amplifier. V_{knee} will reduce the maximum voltage swing and thus the maximum output power, according to

$$P_{out,\,max} = \frac{(V_{DD} - V_{knee})^2}{2R_{opt}} \qquad (6)$$

For deep submicron CMOS technologies V_{knee} may be up to 50% of V_{DD}, using the common definition for the knee voltage [14]. This may be considered too high; also in the triode region the final stage may still operate as an amplifier. The knee voltage reduces the efficiency with a factor of

$$\eta_{knee} = \frac{V_{DD} - V_{knee}}{V_{DD}} \qquad (7)$$

Non-idealities of the transformation network

The integrated passives that make up the impedance transformation network, including the "RF choke" inductor L_1, all have a limited quality factor. In this design both poly-insulator-silicon and metal-insulator-metal capacitors are used. The inductors are differential, saving area and yielding a higher Q factor. One of the inductors is a so-called stacked inductor with 3 metal layers, while for the other one only the top layer is used. The inductor design and modeling was done in-house [16]. The transformation network was designed to transform the 50 Ω antenna impedance to approximately 90 Ω at the frequency of operation.

For the transformation network of Fig. 3 with suitable capacitor and inductor values, including a Q factor of 12 for the inductor L_1, MATLAB simulations show an efficiency of the transformation network of 75%, i.e. the power loss from drain to output will be about 1.2 dB. Moreover, in a real impedance transformation network also capacitor non-idealities and interconnect parasitics will cause power losses.

Combined efficiency example

With a conduction angle α=2.1 (giving the maximum 2nd order output current, see Fig. 2), V_{knee}/V_{DD}=0.1 and a Q factor of 12 for inductor L_1, the combined efficiency following from Eqs. 4 and 7 becomes:

$$\eta_{tot} = \eta_2 \eta_{knee} \eta_Q = 0.64 \times 0.9 \times 0.75 = 0.43 = 43\% \qquad (8)$$

In the above analysis the impact of FET non-linearity (apart from the knee voltage) and efficiency was not taken into account.

4. Results

In Fig. 6 the chip micrograph of the PA with VCO is shown. For the measurements both the stand-alone PA and the PA with VCO were packaged in an LCC package and attached to a PCB.

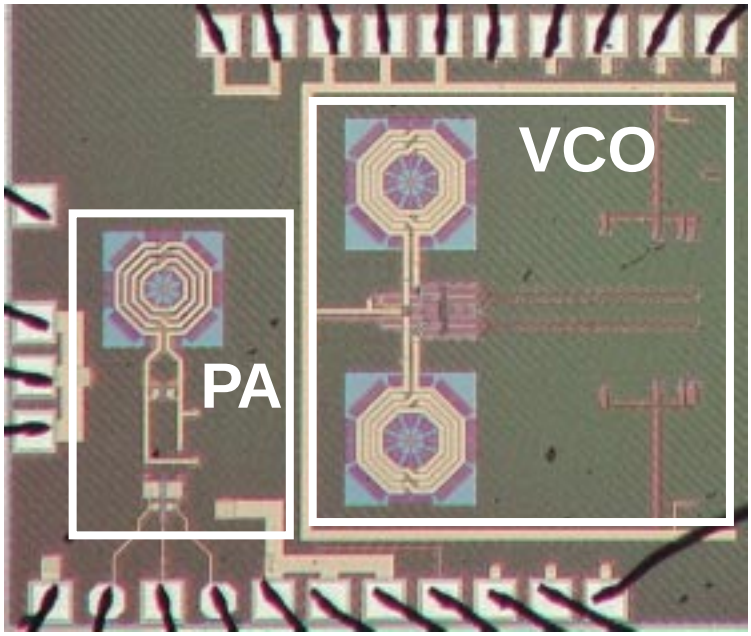
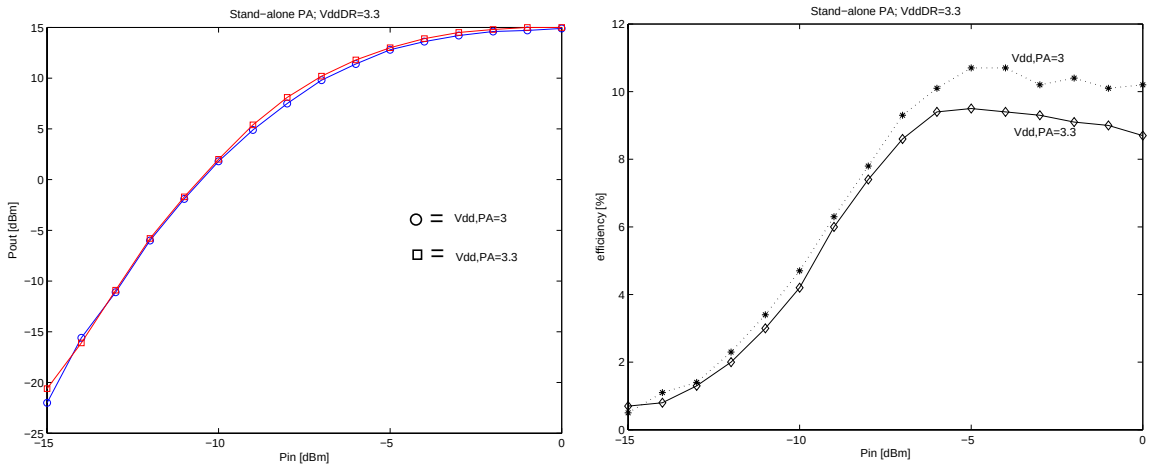


Fig. 6. A chip micrograph of the PA with VCO.

The stand-alone PA

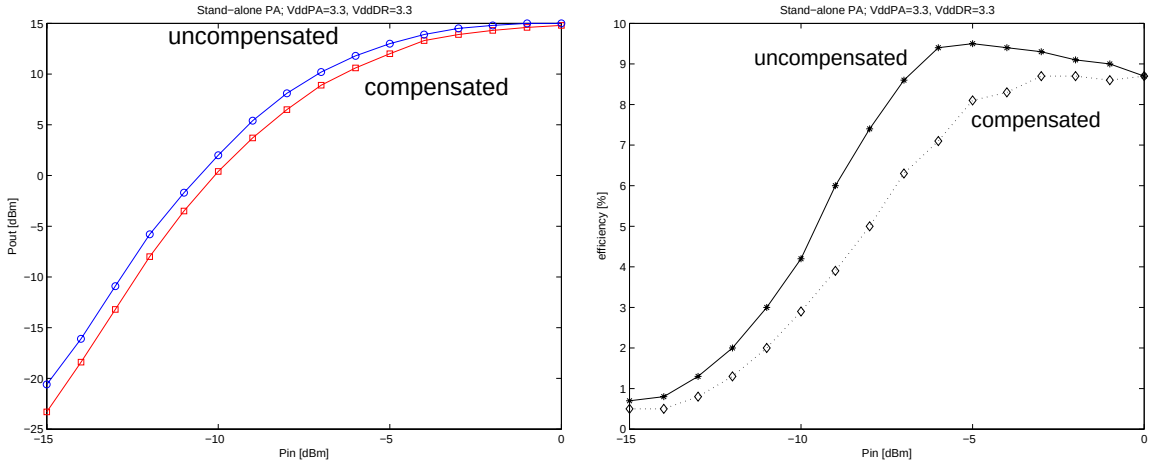
The purpose of this circuit was to verify that the internal frequency doubling worked, and to characterize the final stage. Due to the on-chip polyphase network used to generate the quadrature phases the gain is low. In the measurements the circuit was therefore preceded by a Mini-Circuits amplifier.

In Fig. 7 measurement results for the stand-alone PA are shown,. compensated for the power loss in output cables which added up to approximately 1.8 dB at 2.3 GHz.



a. b.

Fig. 7. Measured results for the stand-alone PA for Vddpa=3.3 and 3 V, a). output power, b). drain efficiency.



a. b.

Fig. 8. Results for uncompensated and compensated outputs, a). output power, b). drain efficiency.

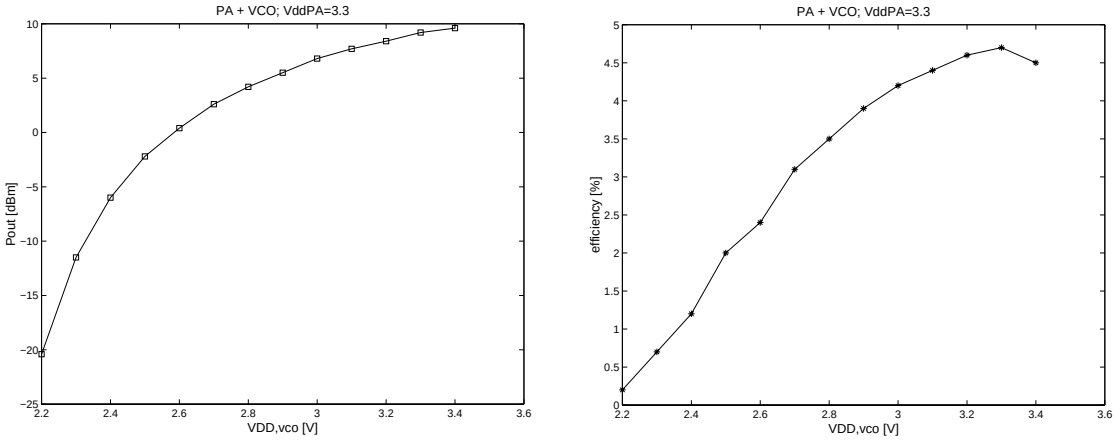
Parasitics related to the bondwires, package and board were estimated and taken into account in the design. S_{22} measurements showed a phase difference between the signals at V_{out+} and V_{out-}, caused by inequality in bondwire parasitics for the symmetrical differential output. This problem was more prominent when the PA was operating with large signals. The mismatch in S_{22} was largely corrected by adding discrete components at the outputs V_{out+} and V_{out-}. The measurement results are shown in Fig. 8.

The correction resulted in a slight decrease in maximum output power and drain efficiency.

The PA with VCO

The differential outputs of this circuit were corrected in order to have approximately equal S_{22} and suitable R_{opt} at the frequency of operation. The matching was worse than for the stand-alone PA, mainly due to the discrete values of the external components.

In Fig. 9 the output power and efficiency as function of $V_{dd,vco}$ is shown. Increasing $V_{dd,vco}$ is equivalent to increasing the input power of the PA stage. The maximum output frequency was 3.27 GHz. The VCO used in this circuit was not optimized for the purpose. Therefore, it could not drive the PA sufficiently which had a negative impact on the efficiency and output power. Moreover, the oscillation frequency was not in agreement with the optimum PA operating frequency. Therefore, passive off-chip components were used to tune the PA output to the desired frequency.



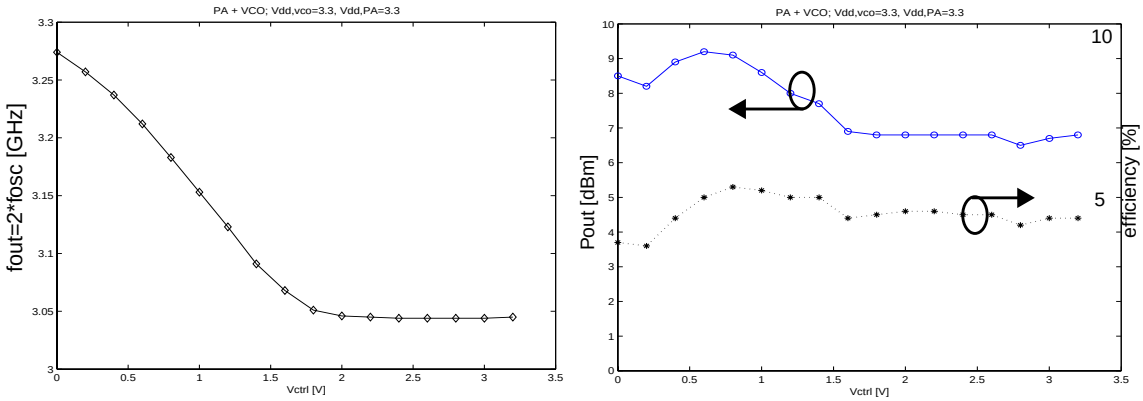
a. b.

Fig. 9. Measured results for the PA with VCO as a function of $V_{dd,vco}$, a). output power, b). drain efficiency.

In Fig. 10 the output frequency vs. control voltage is shown, as well as the output power and drain efficiency.

The output frequency was measured for a VCO control word of (1111). Maximum VCO oscillation frequency was 1.637 GHz, the minimum oscillation frequency 1.522 GHz, resulting in a VCO tuning range of 115 MHz or 7.3% for one control word setting.

Fig. 10b shows that the output power and thus drain efficiency is varying with V_{ctrl}. This is partly due to the PA frequency response, as shown in Fig. 11, and partly to the dependence of the VCO tank quality factor on V_{ctrl}. The output power is almost constant in the frequency range 3.05-3.23 GHz with a variation of 0.7 dB, while decaying for higher frequencies. For this measurement the VCO drive was not



a. b.

Fig. 10. Measurements for varying V_{ctrl}, a). output frequency, b). output power and drain efficiency.

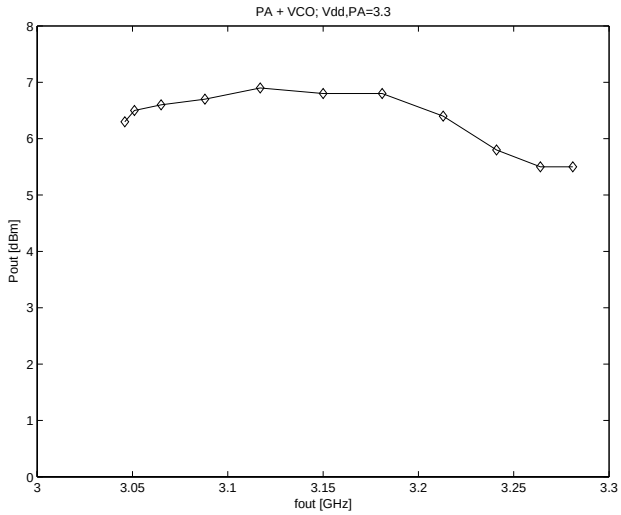
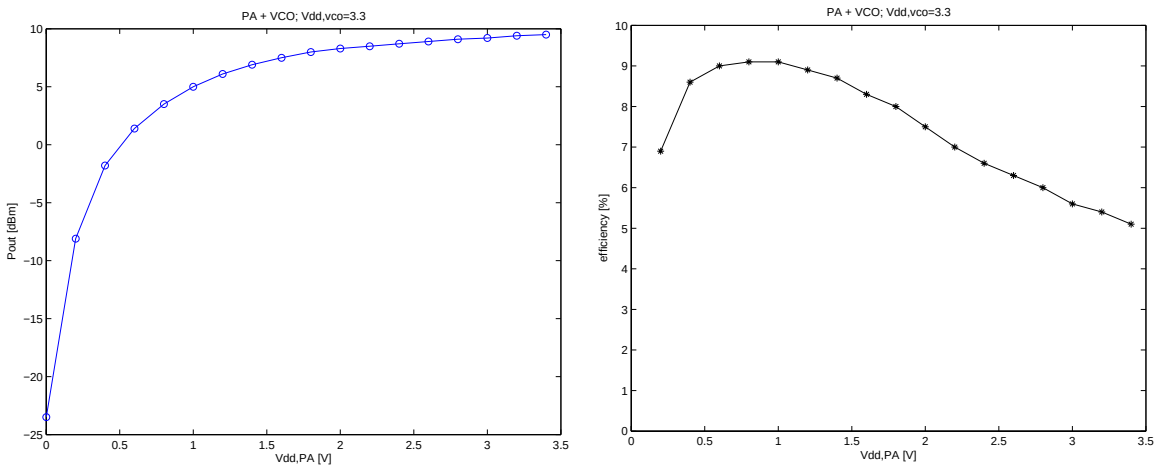


Fig. 11. Frequency response of the PA output stage.

maximum, thus the output power is lower than for the measurements shown in Fig. 9.

Fig. 12 shows the output power as function of $V_{dd,PA}$, illustrating the knee effect, and the drain efficiency. The maximum efficiency occurs for relatively low supply voltages; this illustrates the fact that for equal driving power, the PA with low supply voltage has a clipped output voltage waveform, which may be beneficial for the efficiency [14]. For even lower supply voltages ($V_{dd,PA}$ < 0.6 V, see Fig. 12b) the MOSFET will be in triode and thus have a lower g_m, resulting in a decreasing efficiency.

Measurements of both circuits showed that the gain in each block is less than expected from simulations. This may be due to a lower transistor g_m or lower inductor Q value. This has a large impact on the total gain, but also on the efficiency of the



a. b.

Fig. 12. Results for varying $V_{dd,PA}$, a). output power, b). drain efficiency.

PA and power consumption of the VCO. In Table 1 the measurement results are summarized.

TABLE I. SUMMARY OF MEASUREMENT RESULTS

maximum PA output power	15 dBm
maximum PA drain efficiency	10.7%
VCO frequency range	1.522-1.637GHz
active area: stand-alone PA VCO + PA	1.44 mm^2 2.25 mm^2

In table 2 this work is compared to other fully integrated CMOS power amplifiers. It must be noted that this is the only PA with internal frequency doubling. As explained in Section II, the efficiency and the output power are degraded since the 2nd harmonic is taken instead of the fundamental.

TABLE II. COMPARISON OF FULLY INTEGRATED CMOS PAS

	P_{out} (dBm)	freq. (GHz)	η (max)	technology	output matching	other	class
[2]	33.4 (@PAE_{max})	2.4	31% (PAE)	0.35um (Bi)CMOS	on-chip	-	E/F3
[3]	20 (max)	1.9	16% (η)	0.8um CMOS	on-chip	-	F?
[4]	17.5 (max)	2.4	16.4% (PAE)	0.35um CMOS	partly on-chip	-	A
[5]	9 (@P_{-5dB})	2.4	16% (P_{-5dB})	0.18um CMOS	partly on-chip	-	AB?
[6]	18.6 (@PAE_{max})	0.9	30% (PAE_{max})	0.6um CMOS	on-chip	-	C
this work	15 (max)	2.4	10.7% (η)	0.18um CMOS	on-chip	freq. doubling	C

5. CONCLUSIONS

A fully integrated class-C power amplifier with internal frequency doubler has been designed and measured. A 0.18μm 1P6M CMOS technology was used. The circuit is suitable for a direct-conversion or low-IF transmitter, since the PA and VCO do not operate at the same frequency. Thus, VCO pulling will be reduced. The maximum output power is 15 dBm, with a maximum drain efficiency of 10.7%. The concept of internal frequency doubling was tested both in a stand-alone PA and a PA with VCO on the same chip.

REFERENCES

[1]. M. Rofougaran, A. Rofougaran, C. Ølgaard, and A.A. Abidi, “A 900 MHz CMOS RF Power Amplifier with Programmable Output”, *1994 Symposium on VLSI Circuits Digest*, pp. 133-134, June 1994.

2. I. Aoki, S.D. Kee, D.B. Rutledge, and A. Hajimiri, “Fully Integrated CMOS Power Amplifier Design using the Distributed Active-Transformer Architecture”, *IEEE Journal of Solid-State Circuits*, Vol. 37, no.3, March 2002.
3. Y.J.E. Chen, M. Hamai, D. Heo, A. Sutono, S. Yoo, and J. Laskar, “RF Power Amplifier Integration in CMOS Technology”, *2000 IEEE Microwave Symposium (MTT-S) Digest*, pp. 545-548, May 2001.
4. C.C. Hsiao, C-W. Kuo, and Y-J. Chan, “Integrated CMOS Power Amplifier and Down-Converter for 2.4GHz Bluetooth Applications”, in *Proceedings of the 2001 IEEE Radio and Wireless Conference (RAWCON)*, pp. 29-32, August 2001.
5. K. Yamamoto, T. Heima, A. Furukawa, M. Ono, Y. Hashizume, H. Komurasaki, S. Maeda, H. Sato, and N. Kato, “A 2.4GHz-Band 1.8-V Operation Single-Chip Si-CMOS T/R-MMIC Front-End with a Low Insertion Loss Switch”, *IEEE Journal of Solid-State Circuits*, Vol. 36, no. 8, august 2001.
6. R. Gupta, B.M. Ballweber, and D.J. Allstot, “Design and Optimization of CMOS RF Power Amplifiers”, *IEEE Journal of Solid-State Circuits*, Vol. 36 no. 2, February 2001.
7. M.M. Hella, *RF CMOS Power Amplifiers - Theory, Design and Implementation*. Kluwer Academic Publishers, Norwell, MA, 2002.
8. B. Razavi, *RF Microelectronics*. Prentice Hall PTR, Upper Saddle River, NJ, USA,1998.
9. K. Itoh, T. Yamaguchi, T. Katsura, K. Sadahiro, T. Ikushima, R. Hayashi, F. Ishizu, E. taniguchi, T. Nishino, M. Shimozawa, N. Suematsu, T. Tagaki and O. Ishida, “Integrated Even Harmonic Type Direct Conversion Receiver for W-CDMA Mobile Terminals”, *Proc. of 2002 IEEE Radio Frequency Integrated Circuits Symposium*, pp. 263-266, June 2002.
10. M. Cohn, J.E. Degenford and B.A. Newman, “Harmonic Mixing with an Antiparallel Diode Pair”, *IEEE Trans. on Microwave Theory and Techniques*, Vol. MTT-23 no. 8, August 1975.
11. R. Martin and F. Ali, “A Ku-Band Oscillator Subsystem Using a Broadband GaAs MMIC Push-Pull Amplifier/Doubler”, *IEEE Microwave and Guided Wave Letters*, Vol. 1 no. 11, November 1991.
12. H. Kondoh and A. Cognata, “A 20-50GHz MMIC Amplifier with 21dBm Output Power and its Application as a Frequency Doubler”, *Proc. of IEEE 1993 Microwave and Millimeter-Wave Monolithic Circuits Symposium*, pp. 35-38, June 1993.
13. J. Smith, *Modern Communication Circuits*. McGraw-Hill, Singapore, 1986.
14. S.C. Cripps, “*RF Power Amplifiers for Wireless Communication*”. Artech House, Norwood, MA, USA, 1999.
15. H. Sjöland, “Improved Switched Tuning of Differential CMOS VCOs”, *IEEE Trans. on Circuits and Systems II: Analog and Digital Signal Processing*, Vol. 49 no. 5, May 2002.
16. N. Troedsson and H. Sjöland, “A Distributed Capacitance Analysis of Co-Planar Inductors for a CMOS QVCO with Varactor Tuned Buffer Stage”, accepted for publication in *Analog Integrated Circuits and Signal Processing*.