Design of Radio Frequency Power Amplifiers for Cellular Phones and Base Stations in Modern Mobile Communication Systems

**Von der Fakultät Informatik, Elektrotechnik und Informationstechnik
der Universität Stuttgart
zur Erlangung der Würde eines Doktor-Ingenieurs
(Dr.-Ing.) genehmigte Abhandlung**

Vorgelegt von

Lei Wu

aus der Volksrepublik China · Shandong

Hauptberichter: Prof. Dr.-Ing. Manfred Berroth
Mitberichter: Prof. Dr.-Ing. Joachim Burghartz
Tag der mündlichen Prüfung: 29.06.2009

Institut für Elektrische und Optische Nachrichtentechnik
2009

Acknowledgements

My foremost appreciation goes to Professor Berroth who has given me the chance to finish my PhD at the Institute of Electrical and Optical Communications Engineering, University of Stuttgart. Thanks for his numerous and professional instructions in the last years. I also would like to thank my colleagues and my students for their great technical and personal help through the years.

Thanks also go to my parents who have supported me for so long time. They are not with me in Germany; however, they have given me very strong mental supports for many many years. I also would like to thank my dear wife who is just alone in China and waiting for me patiently. Her existence and her love let me always know that I must work hard to win a wonderful future. Finally, I wish to acknowledge family Mueller for the great private help in the last ten years.

Symbols

Symbol	Description	Unit
a_0, $a_1 \ldots a_n$	Taylor coefficients	
C	capacitance	F
C_b	bypass capacitor	F
C_c	coupling capacitors	F
C_s	series capacitor	F
C_{ds}	drain-source-capacitance	F
C_{gd}	gate-drain capacitance	F
C_{gs}	gate-source capacitance	F
C_{sb}	source-bulk capacitance	F
C_p	parallel capacitance	F
C_{sh}	shunt capacitance	F
C_{ox}	gate oxide capacitance per unit area	F/m^2
C_{th}	thermal capacitance	F
C_i	center of the input stability circle	
C_o	center of the output stability circle	
c	speed of light	m/s
D	distance of the metal traces	m
D_{ds}	drain-source diode	
E_c	energy loss per cycle	J
E_C	conduction band	eV
E_V	valence band	eV
E_F	fermi-level	eV
f	operating frequency	Hz
f_{res}	resonant frequency	Hz
f_0	center frequency	Hz
Δf	band width	Hz
G	gain	dB
G_T	transducer power gain	dB
G_P	operating power gain	dB
G_A	available power gain	dB
g_m	transconductance	S
g_{mb}	transconductance caused by bulk potential	S
$g_{m,d}$	the transconductance of the differential pair	S
g_{ds}	drain-source admittance of the transistor	S
h	substrate thickness	m
I_q	quiescent current	A
I_{dc}	DC component of the current	A
I_n	*n*th harmonic of the current	A
I_1	fundamental component of the current	A
I_d	drain current	A
I_D	current distribution	$mA/\mu m^2$
I_{max}	the maximum value of the current	A
I_R	current on the shunt resistor	A
I_S	current source	A
K	parameter of the Rollet's condition	
L	inductance	H
L_s	series inductor	H

L_w	wiring-inductor	H
L_n	channel length of the n-channel transistor	m
l	length of the microstrip line	m
N	number of sections of quarter-wave transmission lines	
P_L	power delivered to the load	W
P_{1dB}	1 dB compression point	dBm
P_{AVS}	power available from the source	W
P_{IN}	power input to the network	W
P_{in}	RF input power; RF drive power	dBm
P_{out}	RF output power	dBm
Q	quality factor	
R_{ds}	drain-source resistance	Ω
$R_{C_{gd1}}$	series resistance facing C_{gd1}	Ω
R_{opt}	optimum load	Ω
R_{th}	thermal resistance	K/W
R_g	source resistance	Ω
R_L	load resistance to be matched	Ω
R_l	load resistance	Ω
$R_{load.opt}$	optimum load in the class E amplifier	Ω
R_i	radius of the input stability circle	
R_o	radius of the output stability circle	
R_p	shunt resistance	Ω
r_{ds-p}	drain-source resistance of PMOS transistors	Ω
T_n	Chebyshev polynomials	
t	conductor thickness	m
t_{ox}	oxide thickness	m
U_{dd}	supply voltage	V
U_{ds}	drain-source voltage	V
U_{d6}	drain voltage of the top device	V
U_i	input voltage	V
U_o	output voltage	V
U_{gg}	biasing voltage	V
U_{gs}	gate-source voltage	V
U_{ds-ON}	“on” drain-source voltage	V
U_K	knee voltage	V
U_{max}	maximum voltage	V
U_{th}	threshold voltage	V
U_B	bulk voltage	V
U_{br}	breakdown voltage	V
v_p	phase velocity	m/s
v_i	input voltage	V
v_o	output voltage	V
W	conductor width	m
W_n	channel width of the n-channel transistor	m
$W_{D,cell}$	width of the metal lines connected at the drain	m
$W_{S,cell}$	width of the metal lines connected at the source	m
X_L, X_C	reactance of the reactive elements	Ω
Z_0	characteristic impedance	Ω
Z_L	load impedance to be matched	Ω
Z_{opt}	optimum impedance	Ω
Z_{in_opt}	input optimum impedance	Ω

Z_{out_opt}	output optimum impedance	Ω
$Z_1 \ldots Z_n$	characteristic impedance of transmission lines	Ω
α	conduction angle	degree
β	imaginary part of the propagation constant	1/m
Δ	auxiliary parameter of the Rollet's condition	
ΔG	gain offset	dB
ε_0	vacuum dielectric constant	F/m
ε_{eff}	effective relative dielectric constant	
η	output efficiency	%
Γ	voltage reflection coefficient	
Γ_{in}	input reflection coefficient	
Γ_{out}	output reflection coefficient	
Γ_S	source reflection coefficient	
Γ_L	load reflection coefficient	
Γ_m	maximum voltage reflection coefficient	
Γ_n	partial reflection coefficients	
λ	channel-length modulation coefficient	1/V
μ	unconditional stability factor	
μ_n	mobility of the electrons	$m^2/V \cdot s$
ω	angular frequency	Hz
ω_0	center frequency	Hz
π-θ_m	upper edges of the passband	
τ	period of the sinusoidal signal	
τ_i	time constants	s
θ	an arbitrary angle between $-\alpha/2$ and $\alpha/2$	
θ_m	lower edges of the passband	

Glossary

2DEG	two-dimensional electron gas
3GPP	3rd generation partnership project
π/4-DQPSK	π/4 differential quadrature phase shift keying
ACLR	adjacent channel leakage ratio
ACPR	adjacent channel power ratio
AGC	automatic gain control
AlGaAs	aluminium-gallium-arsenide
CDMA	code division multiple access
CM	common-mode
CMOS	complementary metal-oxide-semiconductor
CS	common-source
DC	directional current
DCW	digital control words
ETSI	european telecommunications standardisation institute
FET	field-effect transistor
FDD	frequency division duplex
FDMA	frequency division multiple access
GaAs	gallium-arsenide
GMSK	Gaussian minimum shift keying
GPRS	general packet radio service
GSM	global system for mobile communication

HBT	heterojunction bipolar transistor
HEMT	high electron mobility transistor
HiVP	high voltage/high power
IF	intermediate frequency
IMD	intermodulation distortion
IMD3	third order intermodulation distortion
IMD5	fifth order intermodulation distortion
InP	indium phosphide
IP3	third order intercept point
IIP3	input IP3
LDMOS	lateral double diffused metal-oxide-semiconductor
MIM	metal-insulator-metal
M_I	input matching network
M_O	output matching network
MODEM	modulator/demodulator
MOS	metal-oxide-semiconductor
MOSFET	metal-oxide-semiconductor field-effect transistor
M_1; M_2	metal traces
NADC	north American digital cellular
NMOS	n-type metal-oxide-semiconductor
OCTC	open-circuit time constants
OIP3	output third order intercept point
PA	power amplifier
PAE	power added efficiency
PAR	peak to average ratio
PCB	printed circuit board
PCS	personal communication system
PGA	programmable gain amplifier
PMOS	p-type metal-oxide-semiconductor
RF	radio frequency
SCTC	short-circuit time constants
SiGe	silicon germanium
SMA	straight medium adaptors
SMD	surface mount devices
T	transistor
TEM	transverse electromagnetic mode
TD-CDMA	time division code division multiple access
TDMA	time division multiple access
UMTS	universal mobile telecommunication system
VGA	variable gain amplifier
VSWR	voltage standing wave radio
WCDMA	wideband code division multiple access
ZCS	zero current switching
ZVS	zero voltage switching

Zusammenfassung

Die Bereitstellung von Informationen und Dienstleistungen zu jeder Zeit an jedem Ort wird durch eine rasante Weiterentwicklung der Mobilkommunikation unterstützt. Wo immer drahtlose Kommunikation existiert, gibt es Funksender und viele Funksender benötigen Leistungsverstärker. Deshalb wurden Hochfrequenz (HF)-Leistungsverstärker in den letzten Jahren weltweit intensiv untersucht. In dieser Arbeit werden Leistungsverstärker and ihre Vorverstärker nicht nur für die Basisstationen, sondern auch für die Mobiltelefone entworfen und gemessen, die in den Mobilkommunikationssystemen verwendet werden. Sowohl die hybriden als auch die integrierten Verstärker werden mit MOS-Transistoren entworfen, da sie wegen billiger Materialien und relativ einfacher Herstellungsprozesse als kostengünstigste Technologie angesehen werden. Dank der Entwicklung der modernen Halbleitertechnologie sind die MOS-Transistoren nun genügend schnell, um in HF-Schaltungen eingesetzt zu werden, die bei einigen Gigahertz arbeiten. Aber die MOS-Transistoren zeigen auch Nachteile bei manchen HF-Schaltungen. Zum Beispiel ist es schwierig, ein logarithmisches Verhalten in einer CMOS-Technologie zu realisieren. Außerdem nimmt die Durchbruchspannung der CMOS-Transistoren mit der verkleinerten Dimension stark ab. Die Fähigkeit, bei hohen Spannungen betrieben zu werden, sinkt und die Schwierigkeit, einen Leistungsverstärker in der CMOS Technologie zu entwerfen, steigt. Um die Durchbruchspannung der MOS Transistoren zu erhören, wird ein zusätzliches niedrigdotiertes n-Diffusionsgebiet bei dem Drainanschluss in einem LDMOS-Transistor verwendet. Die Durchbruchspannung dieser Transistoren wird sogar bis auf 70 V erhöht, so dass die Leistungsverstärker, die in den Basisstationen verwendet werden, auch in einer MOS-Technologie realisiert werden können. Aber auf der anderen Seite ist die optimale Ausgangsimpedanz von den LDMOS-Transistoren meistens sehr niedrig. Sie sind deshalb nicht geeignet für breitbandige Leistungsverstärker. Um die oben genannten Probleme zu lösen, werden einige neu entwickelte Schaltungskonzepte in dieser Arbeit präsentiert. Deren Funktionsprinzipien werden bei den Schaltungsentwürfen ausführlich diskutiert. Die Simulationsergebnisse und die Messergebnisse der entworfenen Schaltungen werden verglichen.

Heute wird die komplementäre MOS-Technologie (CMOS) weitgehend für integrierte Schaltungen verwendet. Sie ist bereits, bedingt durch extrem hohe Integrationsdichte und die Möglichkeit einer kostengünstigen Fertigung, in digitalen Schaltungen dominant. In dieser Arbeit wird nun versucht, die in den Mobiltelefonen verwendete Leistungsverstärker und ihre Vorverstärker auch in den Standard-CMOS-Technologien zu entwerfen.

Die Verstärkungsvariation ist eine unverzichtbare Funktion für viele Kommunikationssysteme. Diese Funktion kann durch einen einstellbareren Verstärker (PGA) realisiert werden. Es gibt in der Praxis PGAs, die entweder bei einer Zwischenfrequenz (üblicherweise bei 70 MHz) oder direkt bei der Operationsfrequenz des Systems funktionieren. Man nennt sie IF-PGA bzw. RF-PGA. Der Letztere kann auch als Vorverstärker des Leistungsverstärkers verwendet werden, wenn er genügend hohe Ausgangsleistung liefert. In dieser Arbeit wird ein logarithmischer CMOS-RF-PGA präsentiert. Ein MOS-Transistor hat eine exponentielle Funktion nur in dem Bereich der schwachen Inversion. Da der Strom in diesem Bereich jedoch sehr klein ist, ist es schwierig, einen logarithmischen PGA mit hoher Ausgangsleistung und großem Variationsbereich der Verstärkung in einer CMOS-Technologie zu realisieren. In dieser Arbeit wird ein neues Schaltungskonzept entworfen, bei welchem mehrere Verstärkerzellen parallel verschaltet sind. Der Schaltplan ist in Abb. 3.8 dargestellt. Die parallelen Verstärkerzellen werden durch einen Demultiplexer kontrolliert. Es wird jeweils nur eine Zelle eingeschaltet. Die unterschiedliche Verstärkung und der adaptive Leistungsverbrauch werden erzeugt, indem die Verstärkerzellen verschiedene

Transistorgrößen besitzen. Um eine hohe Immunität gegen das Rauschen von der Umgebung zu erlangen, wird die differentielle Operation in den Verstärkerzellen eingesetzt. Die Kaskode-Konfiguration wird auch in der Schaltung verwendet, um den Miller-Effekt zu unterdrücken. Statt der aktiven Last wird ein LC-Schwingkreis als die gemeinsame Last für alle Verstärkerzellen eingesetzt. Die Simulation zeigt, dass der LC-Schwingkreis einen viel kleineren Spannungsabfall als die aktive Last benötigt, so dass eine höhere Ausgangsleistung erzielt wird. Der Testchip wurde mit einer 0,12-µm CMOS-Technologie hergestellt. Die Messergebnisse zeigen, dass die Verstärkung des PGAs zwischen - 43 dB und 8 dB mit Kontrollschritten von 3 dB geändert werden kann. Es wird somit ein Variationsbereich von 51 dB erreicht. Das ist der größte HF-logarithmische Variationsbereich für die Verstärkung, von dem bisher berichtet wird. Außerdem wird auch ein adaptiver Leistungsverbrauch erhalten. Die maximale Ausgangsleistung erreicht 9 dBm, während der 1-dB Kompressionspunkt bei 8 dBm liegt. Der oIP3 Punkt liegt bei 22 dBm.

Des Weiteren wird in dieser Arbeit ein HF-CMOS-Klasse-A-Leistungsverstärker demonstriert. Die größte Schwierigkeit beim Entwurf eines Leistungsverstärkers in einer modernern CMOS-Technologie ist, dass die Durchbruchspannung ständig mit der verkleinerten Transistordimension sinkt, aber auf der anderen Seite die Ausgangsspannung genügend groß bleiben muss. Deshalb gibt es bisher weltweit ganz wenige Versuche, Leistungsverstärker in einer Technologie moderner als 0,18-µm CMOS zu realisieren. Um dieses Problem zu lösen, wird die High Voltage/High Power (HiVP) Technik (s. Abb. 2.19) in dieser Arbeit eingesetzt, wo mehrere identische Transistoren DC und HF übereinander verbunden sind. Der größte Vorteil der HiVP Struktur ist, dass eine hohe Ausgangsspannung auf mehrere Transistoren verteilt werden kann, damit die Spannung zwischen zwei beliebigen Anschlüssen jedes Transistors nicht höher als die Durchbruchspannung wird. Jedoch auch die Beschränkungen der HiVP Struktur werden in dieser Arbeit ausführlich diskutiert. Erstens kann die Anzahl der Transistoren, die in der HiVP Struktur verwendet werden, nicht unbegrenzt erhöht werden. Die DC-Simulation zeigt, dass die HiVP Struktur ein ähnliches Verhalten wie ein einzelner Transistor hat. Aber der maximale Schwingungsbereich für die HF-Spannung wird verkleinert, wenn die Anzahl der Transistoren sich erhöht (s. Abb. 3.21). Ferner erhöht sich die Schwierigkeit mit der ansteigenden Transistoranzahl, die Versorgungsspannung gleichmäßig auf alle Transistoren zu verteilen. Zweitens wird die HF-Eigenschaft des Leistungsverstärkers dramatisch verschlechtert, wenn ein Serienwiderstand zum Source-Anschluss der HiVP Struktur vorkommt. Da der Strom, der durch die Transistoren fließt, sehr groß ist (mehr als 1 A), kann ein kleiner Serienwiderstand einen starken Abfall der Gate-Source-Spannung aller Transistoren verursachen. Die Transistoren arbeiten dann im Abschnürbereich. Die Ausgangsleistung wird dementsprechend stark reduziert. Schließlich wird eine weitere 0,13-µm CMOS-Technologie für den geplanten Leistungsverstärker ausgewählt, bei welcher drei Wannen verfügbar sind. Die NMOS-Transistoren können durch P-Wanne von dem Substrat isoliert werden und deshalb wird der Leistungsverlust durch das Substrat verkleinert. Außerdem kann der Bulk-Anschluss eines Transistors mit dem Source verbunden werden, so dass die Bulk-Spannung mit der Source-Spannung schwingen kann. Ein anderes wichtiges Thema in dieser Arbeit ist, ein kompaktes Layout für die großen Transistoren zu zeichnen. Gleichzeitig muss der große Strom, der durch die HiVP Struktur fließt, berücksichtigt werden. Alle sechs Metallschichten, die in der CMOS-Technologie zur Verfügung stehen, werden deshalb in dem Layout verwendet. In Abbildung 3.24 wird ein neues gestaffeltes Transistorlayout vorgestellt, das die gesamte Chipfläche effektiv verkleinert. Dieses neue Layoutkonzept hat noch weitere Vorteile, wie z.B. keinerlei Überlappung zwischen dem Drainanschluss und dem Sourceanschluss, gleichmäßige Verteilung des aktiven Bereichs, etc. Eine so genannte H-Struktur wird in diesem Layout für die Gateleitungen eingesetzt. Der Phasenunterschied zwischen den Eingängen der Transistorzellen wird dadurch minimiert. Bei

dem erstellten Testchip wird eine maximale Ausgangsleistung von 29,5 dBm (0,9 W) bei einer Frequenz von 900 MHz gemessen, während die Klein-Signal-Verstärkung ca. 11 dB beträgt. Die maximale Effizienz erreicht 34,5 %. Ferner wird in dieser Arbeit die Methode vorgestellt, mit der die Effizienz des Leistungsverstärkers bei niedrigen Leistungen erhöht wird. Ein Transistor mit großer Gateweite kann in mehrere Zweige aufgeteilt werden (s. Abb. 3.37). Jeder Zweig kann auf digitale Weise ein- oder ausgeschaltet werden. Im Fall, dass die Ausgangsleistung klein ist, werden die meisten Zweige deaktiviert. Der gesamte DC-Strom des Verstärkers wird verkleinert und die Effizienz wird dann erhöht.

In dieser Arbeit werden Leistungsverstärker auch für die Basisstationen in den Mobilkommunikationssystemen entworfen. Heute werden die LDMOS-Transistoren weitgehend in den Basisstationen verwendet wegen der niedrigen Kosten und der besserer Linearität verglichen mit den anderen konkurrierenden Technologien. Der LDMOS-Transistor MRF21030SR3 wird in dieser Arbeit genutzt. Die geplanten Leistungsverstärker sollen breitbandig sein, damit sie sowohl im UMTS-System als auch im GSM1800-System in Europa eingesetzt werden können.

Zuerst wird ein einphasiger Klasse-AB-Leistungsverstärker entworfen. Die Stabilität der Schaltung gehört zu den wichtigsten Themen beim Entwurf eines breitbandigen Leistungsverstärkers, da die meisten Transistoren in einem großen Frequenzband potentiell instabil sind. In dieser Arbeit wird eine neue Methode für die Stabilitätsverbesserung vorgestellt, bei der ein serieller Kondensator mit einem parallelen Widerstand direkt am Drain des LDMOS-Transistors verwendet wird (s. Abb. 4.5). Dieser Kondensator schützt den Widerstand gegen einen großen DC-Strom, damit wird die Stabilitätsverbesserung gewährleistet, die durch den parallelen Widerstand erzielt wird. Außerdem wird der gesamte Leistungsverbrauch der Verstärkerschaltung durch den Kondensator verringert. Die Leistungseffizienz wird erhöht. Die Durchbruchspannung eines LDMOS-Transistors wird wegen des zusätzlichen nieder-dotierten Bereiches beim Drainanschluss dramatisch erhöht. Aber auf der anderen Seite wird eine große Ausgangskapazität durch diesen Dotierungs-bereich erzeugt. Die optimale Ausgangsimpedanz des Transistors wird durch diese Ausgangskapazität auf einen sehr niedrigen Wert transformiert. Eine breitbandige Anpassung ist aus diesem Grund für einen LDMOS-Transistor schwer zu erreichen. Um dieses Problem zu beheben, werden verschiedene Anpassungsnetzwerke in dieser Arbeit entworfen und jeweils in die Verstärkerschaltung eingesetzt. Anhand der S-Parameter-Simulation wird schließlich festgelegt, dass nur die Ein- und Ausgangsanpassungsnetzwerke, die aus Multi-Sektions-Übertragungsleitungen bestehen, eine genügende Bandbreite bieten können. Die Struktur und die Entwurfsparameter der optimalen Ein- und Ausgangsanpassungsnetzwerken werden in Abb. 4.7 und in der Tabelle 4.2 präsentiert. Eine λ/4-Übertragungsleitung wird in dem Versorgungsnetzwerk verwendet. Die Länge der Leitung wird bei einer Frequenz von 2 GHz festgelegt. Die zweite Harmonische des Ausgangssignals, die die größte Signal-verzerrung verursacht, wird durch diese λ/4-Leitung effektiv unterdrückt. Der entworfene Leistungsverstärker wird auf dem Substrat RO4003 aufgebaut und gemessen. Mit einer Versorgungsspannung von 26 V und einer Gate-Source-Spannung von 3,8 V wird eine 3-dB Bandbreite von etwa 1 GHz bei der Messung erzielt. Die Mittenfrequenz liegt bei etwa 2,1 GHz. Das ist die größte Bandbreite, die bisher von einem LDMOS-Leistungsverstärker berichtet wird. Die Breitbandigkeit wird auch bei der Großsignal-Messung nachgewiesen. Eine maximale Ausgangsleistung von 43,5 dBm wird sowohl in dem UMTS-Band als auch in dem GSM1800-Band erzielt. Eine Leistungseffizienz höher als 30 % bzw. 37 % wird bei 1,85 GHz und bei 2,14 GHz erzielt. Hohe Linearität, die mit einem Nachbarkanal-Leistungsverhältnis (ACPR) niedriger als - 40 dB gekennzeichnet ist, wird in den gesamten Leistungsebenen erreicht.

Ein wesentlicher Nachteil von einem einphasigen Leistungsverstärker ist, dass die Reflektionskoeffizienten sowohl am Eingang als auch am Ausgang schlecht sind. Grund dafür ist, dass die Ein- und Ausgangsanpassungsnetzwerke für die Leistungsanpassung und nicht für die konjugiert komplexe Anpassung entworfen werden. Um bessere Reflektionskoeffizienten zu erreichen und deshalb eine hohe Kaskade-Fähigkeit von der Schaltung zu gewinnen, wird ein balancierter Leistungsverstärker in dieser Arbeit entwickelt. Die maximal erreichbare Ausgangsleistung soll weiterhin erhöht werden. Zwei identische Transistoren sind mit zwei hybriden Kopplern parallel verbunden (s. Abb. 4.20). Die zwei von den beiden einzelnen Transistoren reflektierten Signale haben am Eingang des gesamten Leistungsverstärkers die gleiche Amplitude und 180° Phasenunterschied. Sie kompensieren sich gegeneinander, deshalb soll das gesamte reflektierte Signal am Eingang des balancierten Leistungsverstärkers einen Betrag von null haben. Genau das Gleiche gilt auch für die reflektierten Signale am Ausgang. Die gleichen Transistoren, die gleichen Arbeitspunkte und die gleichen Anpassungsnetzwerke, die beim Design des einphasigen Leistungsverstärkers benutzt wurden, werden auch hier in den beiden individuellen Leistungsverstärken verwendet. Der balancierte Leistungsverstärker wird auch auf RO4003 aufgebaut. Bei der Messung wird eine Bandbreite von etwa 830 MHz mit der Mittenfrequenz von 2 GHz erzielt. Die maximale Ausgangsleistung erreicht 50 W (47 dBm) sowohl im UMTS-Band als auch in dem GSM1800-Band. Eine Leistungseffizienz höher als 35 % bzw. 42 % wird jeweils bei 1,85 GHz und bei 2,14 GHz erreicht. Hohe Linearität, die mit einem Nachbarkanal-Leistungsverhältnis (ACPR) niedriger als - 40 dB gekennzeichnet ist, wird in den gesamten Leistungsebenen erzielt. Der in dieser Arbeit erreichte balancierter Leistungsverstärker ist ein aussichtsreicher Kandidat für Multistandard-Multiband-Mobilkommunikationssysteme.

1. Introduction

The mobile radio communication has begun with Guglielmo Marconi's and Alexander Popov's experiments with ship-to-shore communication in the 1890's. Land mobile radio telephone systems have been used since the Detroit City Police Department installed the first wireless communication system in 1921 [1]. Since that time, radio systems have become more and more important for both voice and data communication.

The modern mobile communication systems are mainly designed in high frequency ranges due to the larger available bandwidth at these frequencies. Today, the mostly used mobile communication systems in the United States are cellular telephone systems operating at 800 – 900 MHz and personal communication systems (PCS) at 1800 – 2000 MHz. In Europe, these include the Global System for Mobile Communication (GSM) and Universal Mobile Telecommunications System (UMTS). China now has GSM/GPRS and Code Division Multiple Access (CDMA) networks. For the third generation services, China has been planning a 3G standard called Time Division Synchronous CDMA (TD-SCDMA) since 1999, which is planned to operate at 2010 MHz - 2025 MHz.

In this work, attentions are paid on the uplink and downlink applications in the GSM and the UMTS systems adopted in Europe. The frequency bands of these mobile communication systems are depicted in Fig. 1.1 and listed in Table 1.1.

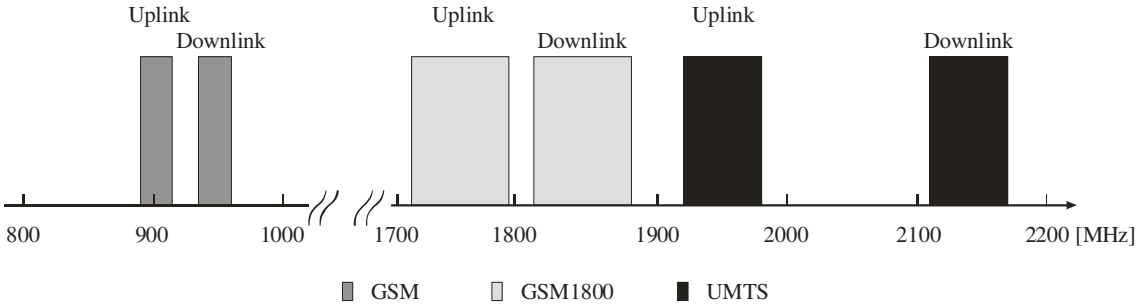


Fig. 1.1. Frequency bands of the mobile communication systems in Europe.

Table 1.1
Frequency bands of the mobile communication systems in Europe

	GSM	**GSM1800**	**UMTS**
Uplink [MHz]	890 - 915	1710 - 1785	1920 - 1980
Downlink [MHz]	935 - 960	1805 - 1880	2110 - 2170

The GSM system is commonly referred to as the second generation mobile phone standard. In the GSM system, time division multiple access (TDMA) is applied. It allows several users to share the same frequency by dividing it into different time slots. The synchronisation of the mobile phones is achieved by sending timing offset commands from the base station which instructs the mobile phones to transmit earlier or later. In radio systems, TDMA is almost always used alongside frequency division multiple access (FDMA) and frequency division duplex (FDD); the combination is referred to as FDMA/TDMA/FDD. FDMA is used in radio systems to share the radio spectrum. This means that users are allocated with different carrier frequencies of the radio spectrum.

Using the code division multiple access (CDMA) modulation process, the UMTS system is aimed as the third generation of mobile communication system, whose standards have first been developed by European Telecommunications Standardisation Institute (ETSI). CDMA is

a form of multiplexing and a method of multiple accesses that does not divide up the channel by time, or frequency, but by encoding the data with a special code associated with each channel and uses the constructive interference properties of the special codes to perform the multiplexing. It usually also refers to digital cellular telephony systems. Compared to the GSM system, the UMTS system has more advantages like longer talk time due to lower power consumption, better quality, larger capacity and cheaper equipment etc. Therefore, the UMTS system is capable of offering new and much more services including multimedia and the access to internet.

No matter which system is discussed, a wireless communication link usually includes a transmitter, a receiver, and a channel as shown in Fig. 1.2 [2]. The functions of the quantization, of the coding and of the decoding are only performed in digital systems. Most links are fully duplex and include a transmitter and a receiver or a transceiver at each end of the link. Obviously, to send or receive large enough signals, power amplifiers and their driving amplifiers are necessary on both sides of the link.

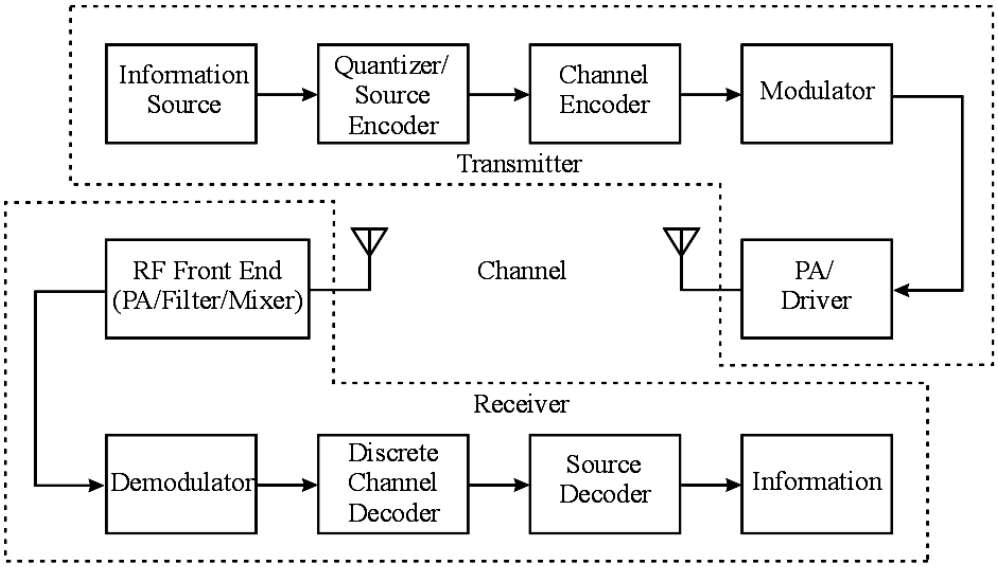


Fig. 1.2. The block diagram of a wireless communication link.

A radio frequency (RF) power amplifier (PA) is a circuit for converting directional current (DC) input power into a significant amount of RF output power. One of the principal differences between a small-signal amplifier design and a power amplifier design is that the main purpose of the latter is the maximum output power, not the maximum gain. However, a power amplifier cannot simply be regarded as a small-signal amplifier driven into the saturation. There is a great variety of different power amplifiers, while most of them employ techniques beyond simple linear amplification. In other words, RF power can be generated by a wide variety of techniques using a wide variety of devices.

In this work, the fundamental theories used for the design of RF power amplifiers are systematically introduced. Using these theories, power amplifier circuits are designed both for the base stations and for the cellular phones adopted in the modern mobile communication systems in Europe. Not only current source mode but also switching mode power amplifiers are considered and attempted.

The basic techniques for RF power amplification are classes A, B, C, D, E, and F modes. In sections 2.1 and 2.2, the functional principles of these different power amplifier modes are first introduced. The most important component of an RF power amplifier circuit is the transistor used as the power device. Today, LDMOS, GaAs, HBT transistors are widely

adopted as commercial power devices for the base station applications, which offer an output power level up to several hundred Watt. In cellular phones, GaAs technology is widely used due to its low loss substrate and high breakdown voltage. On the other hand, standard digital technology sees a breakthrough in both performance and size in case that deep-submicron CMOS devices are used. For certain highly-integrated, low-power, wireless transceivers operating at several GHz, CMOS is now very attractive. However, the most significant drawback of the deep-submicron CMOS transistors, namely the low breakdown voltage, is a limitation for the power amplifier design. To solve this problem, the high voltage/high power (HiVP) structure is adopted, which can also be regarded as a single active device. The features of these semiconductor technologies and constructions are introduced in section 2.3. In this work, MOSFETs are applied for the developments of both the hybrid and the integrated power amplifier circuits. Several conventional design concepts for MOSFET circuits, for instance, the common-source stage, the differential amplifier and the cascode stage, are introduced in section 2.4. In order to obtain the maximum output power, the reference impedance (usually 50 Ohm) must be transformed to the optimum input and output impedance of the selected transistor. Matching networks are therefore necessary at the input and at the output of a power amplifier circuit. The design methods for different matching concepts, such as discrete matching networks or distributed matching networks are provided in section 2.5. Another important issue for a power amplifier design is to select the proper quiescent point and holding the quiescent point constant over variation in transistor parameters and temperature. This can be realized by selecting a suitable biasing network. Moreover, the other topics must also be taken into account in the designing of the biasing network, for example, the output power, amount of distortion, the efficiency, the voltage headroom, the gain of the stage, the noise of the stage, and the class of operation, etc. Additionally, a biasing network is also helpful for suppressing the higher harmonics and controlling the DC power consumption in dependence on the power level. All these considerations are introduced in section 2.6. Finally, in section 2.7, the design parameters in terms of developing the microwave power amplifiers are described. They are the power gain, the stability, the 3-dB bandwidth, the 1-dB compression point, the power added efficiency (PAE), the intermodulation distortion (IMD) and the adjacent channel power ratio (ACPR).

Today, the design and analysis of integrated RF and microwave circuits is receiving a considerable interest by the research community due to the continuous growth in the wireless telecommunication market. In particular, many ongoing efforts are focused on the integration of RF circuits in standard CMOS technologies. This is necessary in order to allow the implementation of RF front-ends with digital signal processors and enable low-cost single-chip fully integrated solutions. The evolution of CMOS technologies and the high level of integration they offer have made the CMOS circuits attractive candidates for RF and microwave applications.

In section 3.1, the design of an RF CMOS logarithmic programmable gain amplifier (PGA) is first demonstrated which is realized in the Infineon 0.12-µm CMOS technology. CMOS devices perform an exponential or logarithmic function only in the subthreshold region. Since the current in this region is very small, it is generally difficult to realize a logarithmic PGA with high output power in CMOS technology. This work presents a novel circuit concept employing parallel amplifier cells. In this manner, the logarithmic gain variation and adaptive power consumption can easily be obtained. Moreover, a large gain control range of 51 dB, high output power of 9 dBm as well as high linearity are also achieved at the radio frequencies.

Section 3.2 presents the design of an RF class A CMOS power amplifier. To overcome the problem of low breakdown voltage, the HiVP structure is employed, in which several transistor devices are connected DC and RF in series. Therefore, the large output voltage can be divided by all the cascaded devices. In this section, the design approaches to obtain an

equal voltage division and to draw a compact layout for the large transistors used in the HiVP structure are described in detail; the simulation and experimental results of the proposed power amplifier are presented. A small signal gain of 11 dB is obtained in the measurement, whereas the maximum output power reaches 29.5 dBm. The maximal power added efficiency reaches 34.5 %. They are comparable with the other works reported over the world in recent years, or even better. On the other hand, the limitation of using the HiVP structure is also described, while lots of methods to improve the HiVP-performance are provided in this chapter, not only for the circuit design but also for more adopted transistor layout. Finally, the design concept for transistors with adjustable gate width is introduced in this work, so that the power added efficiency of the HiVP power amplifier can significantly be increased in the low power levels.

In chapter 4, the design of broadband power amplifiers for base station applications is demonstrated. These amplifiers can simultaneously be used in the GSM1800 and in the UMTS systems in Europe. Motorola LDMOS transistors are employed for these works.

Section 4.1 presents the design of a single-ended class AB power amplifier. Firstly, a novel modification for the stability improvement is introduced, with which the conditional stability of the proposed power amplifier circuit is obtained in the whole frequency band. Simulations also show that this modification can greatly reduce the total DC power consumption; hence the power added efficiency is enhanced. Secondly, different impedance matching networks described in chapter 2 are attempted in order to fulfil the specification of large bandwidth. The S-parameters of power amplifiers with these different matching networks are simulated. Comparing the simulation results, the impedance matching networks consisting of multi-section transmission lines are determined to be the best candidate to realize a broadband power amplifier. Finally, the proposed power amplifier circuit is fabricated in the laboratory and measured. Good agreements between the simulations and the measurements are obtained. The measurement results show that this power amplifier has a 3-dB bandwidth of 1 GHz at the center frequency of 2.1 GHz. To the best of the author's knowledge, this is the largest bandwidth reported so far at these frequencies in an LDMOS technology. The maximum output power of this circuit reaches 43.5 dBm. High efficiency and high linearity are also achieved over large frequency range.

Using the same LDMOS transistors and the same matching networks obtained in the design of the single-ended one, a broadband LDMOS balanced class AB power amplifier is presented in section 4.2. Two power devices are connected in parallel, hence an even higher output power up to 50 W (47 dBm) is achieved. Simultaneously, the balanced structure also provides perfect voltage standing wave ratio (VSWR) performance. Measurement results show that S_{11} and S_{22} smaller than -10 dB are obtained in large frequency range.

Finally, the most important design processes, technical innovations and simulation as well as experimental results for developments of RF power amplifiers and their driving stages obtained in this work are summarized and presented in chapter 5.

2. Fundamentals to the Power Amplifier Design

The general design concept of a power amplifier is given in Fig. 2.1. Ordinarily, it consists of the transistor T, the input and output matching networks M_I and M_O, the biasing networks and the alternating current (AC) couplings C_g and C_d. U_i and U_o in Fig. 2.1 indicate the input and the output voltages, while U_{dd} and U_{gg} are the supply voltage and the biasing voltage, respectively.

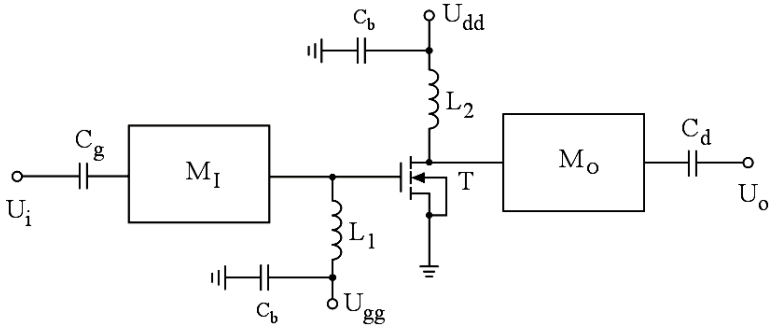


Fig. 2.1. General design concept of a RF power amplifier.

The main purpose of a power amplifier design is to provide sufficiently high output power, while another very important aim is to achieve high efficiency. Therefore, there are generally two types of power amplifiers: the current source mode power amplifiers and the switching mode power amplifiers. Different kinds of each mode of power amplifiers and their functional principles are introduced in detail in sections 2.1 and 2.2. The active devices are the basic components in a power amplifier circuit. Different technologies, e.g. CMOS, LDMOS, etc are used in practice. They are concisely described in section 2.3. Moreover, the HiVP structure, which can also be regarded as a single active device is demonstrated here. In this work, MOS transistors are used as the active devices. Several typical circuit concepts using MOS devices are introduced in section 2.4. In section 2.5, different matching networks used at the input and the output of power amplifier circuits are described. For a current source mode power amplifier, a so-called load-line matching [3] at the output is often used, which is a compromise to extract the maximum power from the RF transistors and at the same time to keep the RF voltage swing within the available DC supply. On the other hand, a complex conjugate matching is usually used at the input. For a switching mode power amplifier, fixed structures of the output matching network have been developed, in order to obtain determined waveforms of the output voltage and the output current. Theoretically, there is no overlap between the output voltage and output current in the time domain, so that a high efficiency of the power amplifier could be ensured. In Fig. 2.1, the inductors L_1 and L_2 serve as RF choke which block the RF signals and simultaneously feed DC power to the device. This is the simplest configuration of biasing networks. Various structures of biasing networks are provided in section 2.6, which have different functions, as well as different benefits and drawbacks. The AC coupling C_g and C_d pass the RF signals and cut off the power supply. In section 2.7, the most important design parameters for a power amplifier circuit are presented, which are frequently used to describe the specifications of the power amplifiers.

2.1 Current Source Mode Power Amplifiers

Obviously, in a current source mode power amplifier, the power device is regarded as a current source, which is controlled by the input signal. The most important current source

mode power amplifiers are class A, class B, class AB and class C power amplifiers. They differ from each other in the operating points. Figure 2.2 illustrates the different classes of current source mode power amplifiers in the transfer characteristic of a field-effect transistor (FET) device. The drain current I_d exhibits pinch-off, when the channel is completely closed by the gate-source voltage U_{gs} and reaches the saturation, in which further increase of gate-source voltage results in no further increase in drain current.

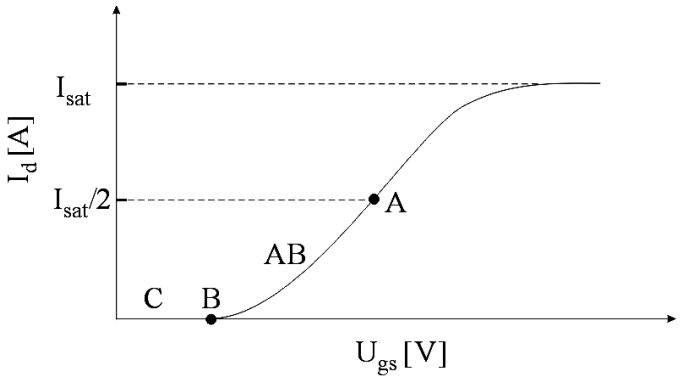


Fig. 2.2. Operating points of the different classes of current mode power amplifiers [3].

The other very important concept to define the different classes of current source mode power amplifier is the conduction angle α. As illustrated in Fig. 2.3, the conduction angle depicts the proportion of the RF cycle for which conduction occurs, where ω is the angular frequency and ωt is the radian. I_q in Fig. 2.3 denotes the quiescent current, while I_{max} denotes the maximum value of the current. The conduction angles of different classes are summarized in Table 2.1.

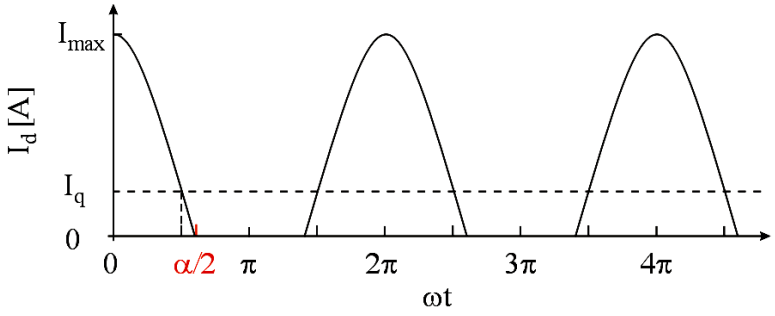


Fig. 2.3. The conduction angle of the current source mode power amplifiers [3].

Table 2.1
Conduction angle of the different classes of current mode power amplifiers [3]

Class	Conduction angle
A	2π
AB	π - 2 π
B	π
C	0 - π

The general schematic of a current source mode power amplifier is already depicted in Fig. 2.1. Usually, matching networks are used to realise a conjugate complex matching at the input

and a load-line matching at the output. The load-line matching at the output takes both the maximum current I_{max} that the device can supply and the maximum voltage that the device can sustain across its terminals into consideration. It has the task of transforming the load impedance to the optimum output resistance R_{opt}, which can be calculated using the load-line method as shown below.

Fig. 2.4 shows the output characteristic of a FET, where U_{br} and U_K denote the breakdown voltage and the knee-voltage of the transistor, respectively.

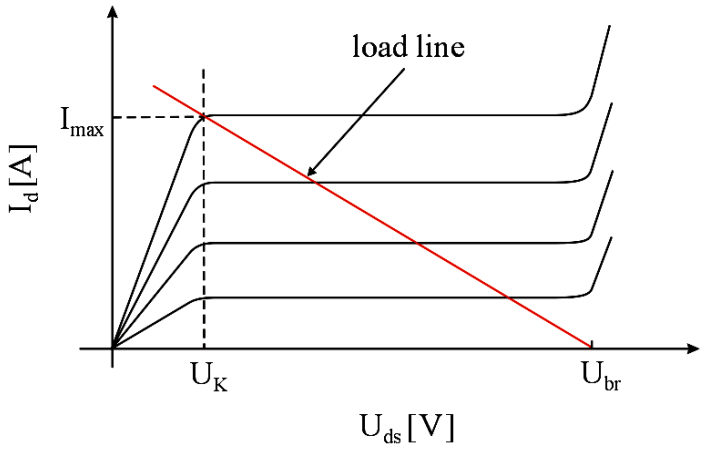


Fig. 2.4. The load-line method.

The optimum output resistance R_{opt} can be obtained using the following equation.

$$R_{opt} = \frac{(U_{br} - U_K)}{I_{max}} \tag{2.1}$$

However, the calculated R_{opt} can normally not directly be used in practice, since it will be transformed by the output capacitance of the power devices (e.g. the drain-source-capacitance C_{ds} of a MOS transistor) and the output wiring-inductance L_w to the practical optimum impedance Z_{opt}, which has unfortunately a much smaller magnitude than R_{opt}. Z_{opt} for different frequencies of the commercial power devices are usually provided by the manufacturers, which are extended to the frequency-dependent optimum input impedance Z_{in_opt} and optimum output impedance Z_{out_opt} determined through the load-pull measurements [3]. The load-pull data has been the mainstay of RF and microwave power amplifier design for many years. It gives the power amplifier designer a simple target area in the Smith chart on which to base the strategy for suitable matching network design. Usually, output matching networks are designed as low pass filters; hence the harmonics generated by the transconductive nonlinearities will be greatly attenuated.

2.1.1 The Class A Power Amplifier

As shown in Fig. 2.2, the operating point of a class A power amplifier is just located at the midpoint of the region between the cutoff and the saturation. The conduction angle of a class A power amplifier is 2π. Therefore, the output waveforms of a class A power amplifier using FET devices can be shown in Fig. 2.5. The drain current having an amplitude of I_{max} /2 sweeps the entire range from zero to the maximum current I_{max}. The drain-source voltage can swing over its maximum range of zero to $2U_{dd}$ [3], where U_{dd} is the supply voltage of the

power amplifier circuit as shown in Fig. 2.1. The drain voltage drops while the drain current increases, and vice versa.

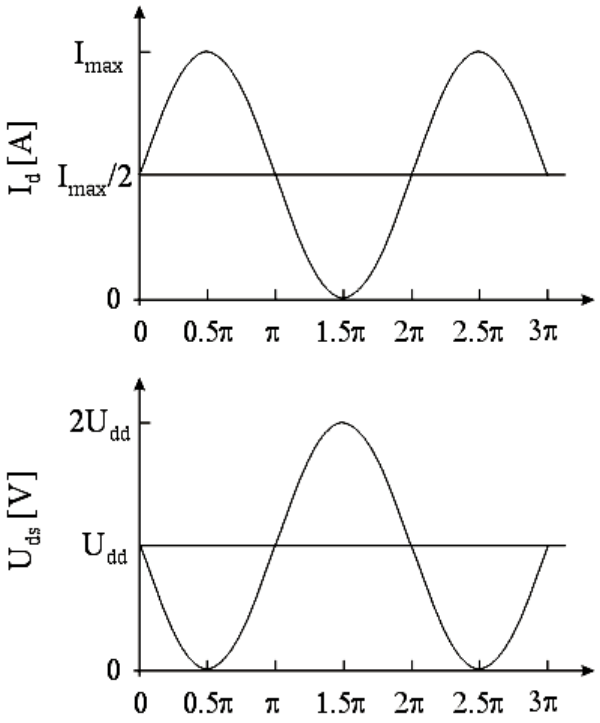


Fig. 2.5. Waveforms of a class A power amplifier.

Ideally, if a sinusoidal signal is fed at the input, the waveform of the output current remains also sinusoidal for a limited range and no harmonic occurs. Therefore, a class A power amplifier is theoretically the most linear power amplifier. However, in practice, the linear region contains weak nonlinearities. On the other side, it is not negligible that a class A power amplifier consumes the most DC power resulting in a lower efficiency compared to the other classes of power amplifiers. The maximum drain efficiency of a class A power amplifier is theoretically 50 % [3]. The heat dissipation is therefore an essential problem which must be considered.

2.1.2 The Class B Power Amplifier

As shown in Fig. 2.2, the operating point of a class B power amplifier is just located at the pinch off point in the transfer characteristic of a FET device. The conduction angle of a class B power amplifier is π, just half of that of a class A amplifier. The output waveforms of a class B power amplifier using FET devices are shown in Fig. 2.6. Clearly, the negative part of its drain current is cut off, while the drain-source voltage can also swing over its maximum range of zero to $2U_{dd}$ [3].

Due to the much lower operating point, the DC power consumption of a class B power amplifier can significantly be reduced. Much higher efficiency can therefore be expected. The maximum drain efficiency of a class B power amplifier is theoretically 78.5 % [3]. However, due to the loss of the negative part of its drain current, harmonics of the output current occur, which should be diminished by suitable output matching network or biasing network as shown later.

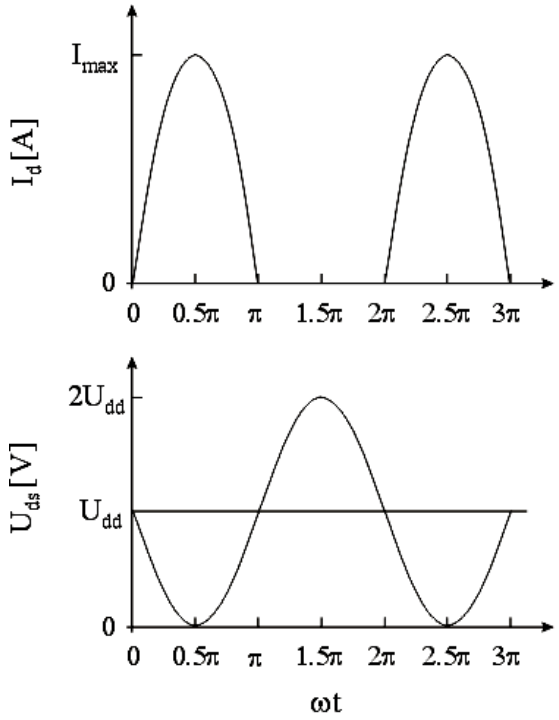


Fig. 2.6. Waveforms of a class B power amplifier.

In practice, class AB power amplifier is widely used. It has the operating point between those of the class A and class B operation modes; therefore, a class AB power amplifier has moderate waveforms and features. Part of the negative drain current is cut off. The drain-source voltage swings from zero to the voltage level even higher than $2U_{dd}$ [3], depending on the resonance circuit used at the output. Theoretically, the linearity of a class AB power amplifier is worse than that of a class A power amplifier. However, since the section of the transfer characteristic of the power devices, which locates between the cutoff and the saturation points, is usually not linear in practice, simulations and also the experiments often show that the linearity of a class AB power amplifier can even be better than that of a class A power amplifier. On the other hand, due to the relatively low operating point, class AB power amplifiers have higher efficiency than class A power amplifiers. The maximum drain efficiency of a class AB power amplifier can theoretically also reach 78.5 % [3]. Therefore, as a good compromise of linearity and efficiency, class AB power amplifiers are very popular and widely adopted in the practical applications.

2.1.3 The Class C Power Amplifier

As shown in Fig. 2.2, the operating point of a class C power amplifier is located between zero and the pinch-off point in the transfer characteristic of an enhancement FET device. The conduction angle of a class C power amplifier is between 0 and π. The output waveform of a class C power amplifier using FET devices is shown in Fig. 2.7. Clearly, the drain-source voltage can also swing over its maximum range of zero to $2U_{dd}$ [3]. On the other hand, the entire negative part and a fraction of the positive part of the drain current are cut off; the current waveform is reduced to a train of short pulses, which have lower DC component compared to the other classes of power amplifiers mentioned above, but also a lower fundamental RF component. Consequently, very high efficiencies can be obtained, but at the expense of lower RF output power and heavy input drive requirements. The maximum drain

efficiency of a class C power amplifier can even reach 100 % [3], if the operating points close to the zero point are selected.

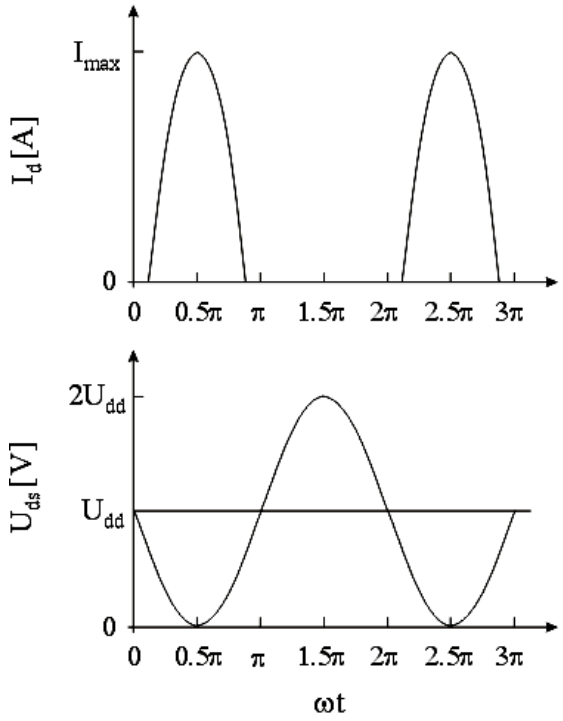


Fig. 2.7. Waveforms of a class C power amplifier.

One of the major problems with utilizing class C modes in solid state applications is the large negative swing of the input voltage, which coincides with the output voltage peaks. This is precisely the worst condition for reverse breakdown in any kind of transistor, and even small amounts of leakage current flowing at this point of the cycle can degrade the efficiency. For this reason, class C power amplifiers are not often used in solid state amplification at higher RF and microwave frequencies.

2.1.4 Comparison of the Different Classes of Power Amplifiers

As mentioned above, different classes of the current source mode power amplifier have different operating points. Observing Fig. 2.3, it is intuitive that the mean component will decrease as the conduction angle is reduced. Additionally, harmonics will occur as well. Using Fourier analysis, the mean current, i.e. the DC component I_{dc} can be given as follows [3]

$$I_{dc} = \frac{I_{\max}}{2\pi} \cdot \frac{2\sin(\alpha/2) - \alpha \cdot \cos(\alpha/2)}{1 - \cos(\alpha/2)}. \tag{2.2}$$

The *n*th harmonic I_n can be described by

$$I_n = \frac{1}{\pi} \int_{-\alpha/2}^{\alpha/2} \frac{I_{\max}}{1 - \cos(\alpha/2)} \left[\cos\theta - \cos(\alpha/2)\right] \cdot \cos n\theta \, d\theta, \tag{2.3}$$

where θ is an arbitrary angle between –α/2 and α/2. The fundamental component I_1 is therefore

$$I_1 = \frac{I_{max}}{2\pi} \cdot \frac{\alpha - \sin\alpha}{1 - \cos(\alpha/2)}. \tag{2.4}$$

The amplitudes of the DC component and of the different harmonics |A| in dependence on the conduction angle is shown in Fig. 2.8. Obviously, the DC component decreases monotonically as the conduction angle is reduced. According to equation (2.2), the DC components I_{dc} of the class A and the class B power amplifiers can be calculated, resulting in the results of I_{dc}(class A)= $I_{max}/2$ and I_{dc}(class B)= I_{max}/π, respectively. Therefore, it is clear that the class B power amplifier has the same fundamental component as the class A power amplifier, while its DC supply power is reduced by a factor of π/2, hence the efficiency increases from 1/2 in the class A mode to π/4 (about 78.5%) in class B. Compared to the class A operation, the class AB operation shows benefits of even higher fundamental component and lower DC power consumption.

As mentioned above, the linearity of the class B operation is degraded, compared to the class A power amplifier. Note in Fig. 2.8 that throughout the class AB range and up to the midway class B condition the only significant harmonic, other than the fundamental signal, is the second harmonic. This however, can be shorted by using suitable low pass or band pass matching networks at the output of the power amplifier circuit. Therefore, the class AB operation sometimes could have even better linearity than the class A operation, depending on the selected operating point and the actual shape of the transfer characteristic.

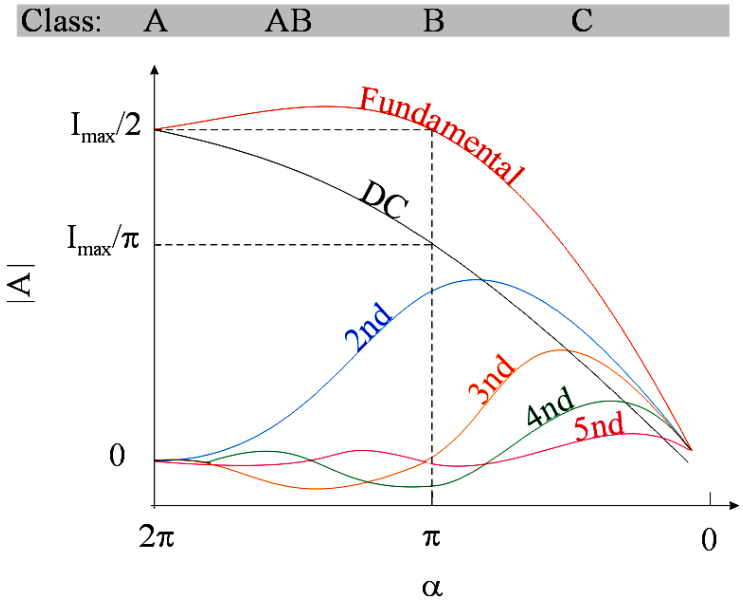


Fig. 2.8. The Fourier analysis of the DC component and the harmonics.

For conduction angles lower than π, corresponding to class C operation, the DC component continues to drop, but the fundamental component of the current also starts to drop below the level of class A and class B operations. However, it can mathematically be proven that the efficiency of the class C operation continues to increase.

2.2 Switching Mode Power Amplifiers

The switching mode power amplifiers are widely used in different frequency ranges and output power levels beginning from several kilowatts at low frequencies and up to one watt at microwaves. In these power amplifiers, the transistors operate as an on-to-off switch and the waveforms of the output current and the output voltage fulfill the condition that they do not overlap at any given time. Therefore, the power dissipation is minimized and the efficiency is maximized. Such an operation mode can be realized by an appropriate choice of the values for the reactive elements in the output load network.

The first possibility of increasing the efficiency of the single-ended power amplifier by modifying of the output matching circuit was experimentally described by Lohrmann in 1966 [4]. Three years later Artym [5] and Gruzdev [6] provided the theoretical analysis of the operating conditions of the single-ended switching mode power amplifiers with the calculation of their circuit parameters. Analysis is carried out for the operation conditions of switching mode power amplifiers not only with the shunt capacitance but also with the resonant circuit tuned on the fundamental to provide the sinusoidal signal flowing into the load. Later, the generalized analysis of the electrical performance and circuit parameters of the single-ended switching mode power amplifiers with shunt capacitance C and series inductance L, as well as with a parallel LC circuit was presented by Popov [7] and Kozyrev [8]. The schematics of them are separately shown in Fig. 2.9 (a) and Fig. 2.9 (b). These two general design models have been used even up to now.

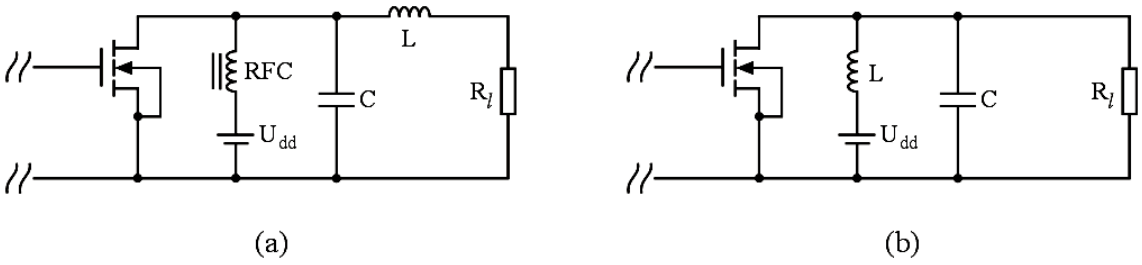


Fig. 2.9. Switching-mode power amplifiers with
(a) shunt capacitance and series inductance and (b) parallel LC circuit.

The most significant benefit of switching mode power amplifier is the high efficiency of up to 100 %. But since the transistors are used as switches, the largest disadvantage of a switching mode power amplifier is the nonlinearity. Therefore the signal spectrum can be degraded and severe adjacent channel interference will occur.

The typical switching mode power amplifiers are class D and class E power amplifiers. The class F power amplifiers are often used as the driving stages of the class E power amplifiers.

2.2.1 The Class D Power Amplifier

The original class D power amplifier is the voltage mode class D power amplifier. Using a MOSFET device, the simplified circuit of this power amplifier and its waveforms of the drain voltage and the drain current are shown in Fig. 2.10. Two transistors are connected in parallel and a series filter comprising the inductor L and the capacitor C is employed. This filter is tuned at the fundamental frequency. The two transistors T1 and T2 are driven 180° out-of-phase. That means, the input connection guarantees that only one transistor is driven on at a given time, with one transistor handling the positive half-cycles and the other one the negative half-cycles. It works just like a push-pull class B power amplifier. The difference here is that

the transistors are driven hard enough and hence act like switches. The voltage across the transistors is a square wave and the transistor current becomes a half-wave rectified sine wave. A Zero-Current-Switching (ZCS) is realized, which means that the drain current is zero, if a transistor is turned on. Ideally, there is no overlap of the drain voltage and the drain current. Therefore, the DC power consumption should theoretically be zero.

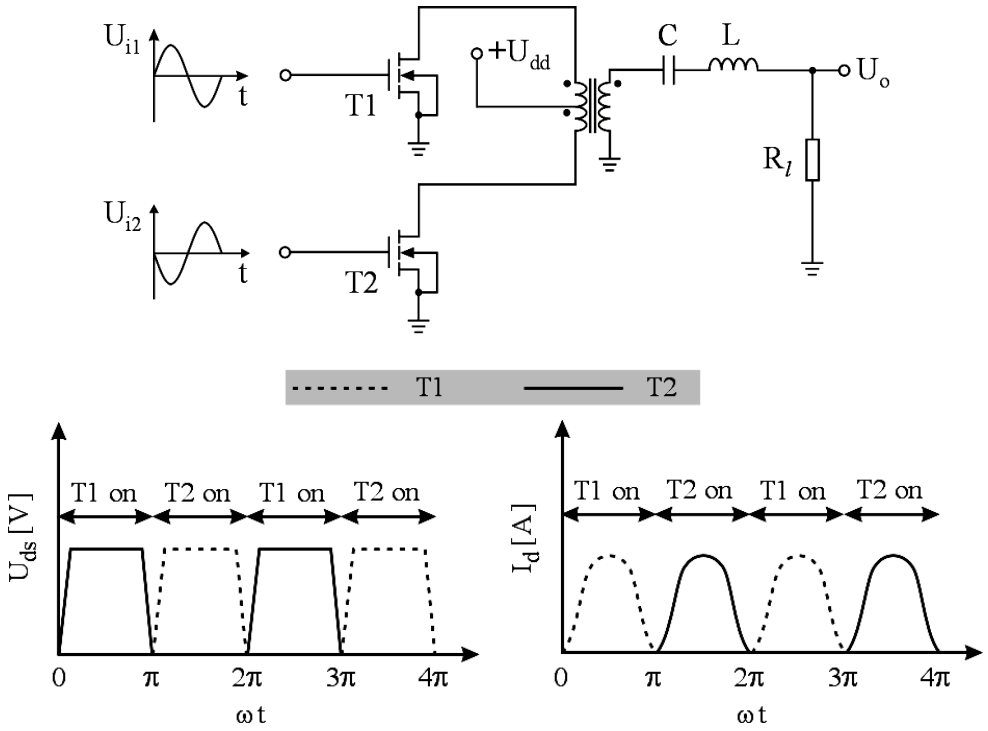


Fig. 2.10. The voltage mode class D power amplifier and its waveforms.

Unfortunately, a real transistor has always parasitic capacitors, e.g. the drain-source capacitor in a field-effect transistor. The transistor must be charged or discharged to the supply voltage U_{dd} or ground through the transistor. That means, the voltage waveform can not have a perfect square wave shape and some transient current spikes occur when the transistor turns on. The overlap of voltage and current can not be avoided. The energy loss E_c per cycle can be calculated with the following equation [9]

$$E_c = \frac{1}{2} C_{ds} \cdot U_{ds-ON}^2, \tag{2.5}$$

where C_{ds} is the drain-source capacitance and U_{ds-ON} is the drain-source voltage when the transistor is turned on. This energy loss on the output capacitance is independent on the channel resistance of the transistor and becomes the dominant loss mechanism at high frequencies.

To overcome the problem of energy loss, a new design concept of class D power amplifier, namely current-mode class D power amplifier is given, which is shown in Fig. 2.11. Two transistors are also connected in parallel, but instead of a voltage source we use a current source in this design. Instead of a series filter a shunt filter is employed, which is also tuned at the fundamental frequency. The voltage across the transistors is a half-wave rectified sine wave and the transistor current becomes a square wave. Due to the filter resonance, there is no voltage across the transistors when a transistor is turned on; therefore, a so-called Zero-

Voltage-Switching (ZVS) is realized. Even if the transistors have some output capacitance, they can be regarded as a part of the parallel filter, so that an overlap of the drain voltage and the drain current can theoretically be avoided.

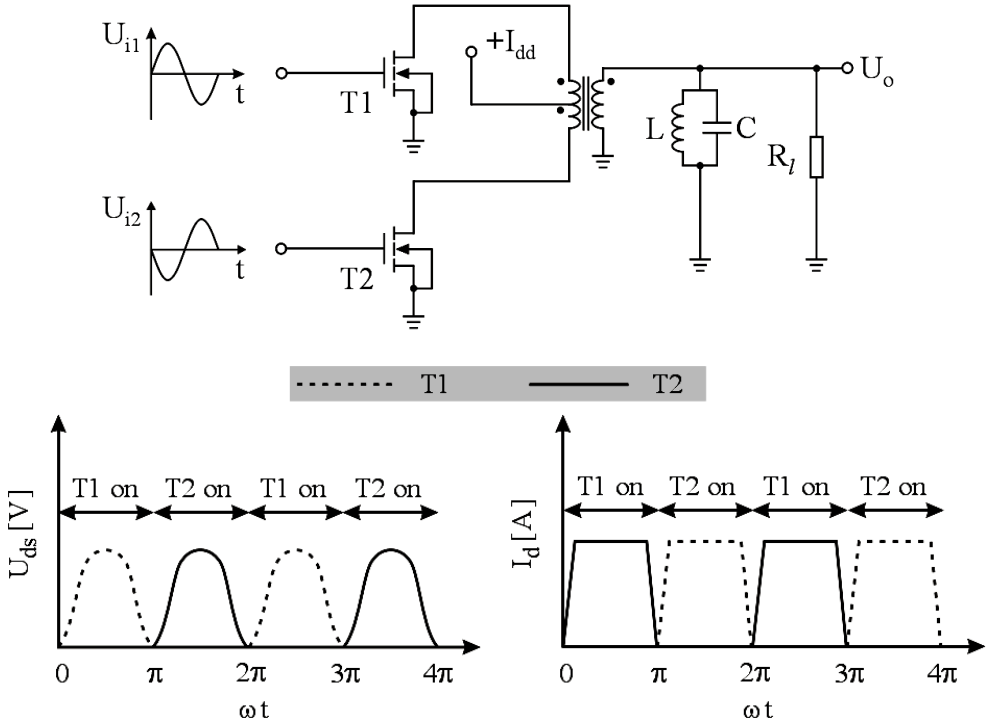


Fig. 2.11. The current mode class D power amplifier and its waveforms.

2.2.2 The Class E Power Amplifier

Another important kind of switching-mode power amplifier is the class E power amplifier, which is introduced by Socal in 1975 [10]. The schematic of a class E power amplifier is shown in Fig. 2.12. A shunt capacitor C_p is used to provide the charge-discharge function of the switching-mode operation. It ensures that the voltage across the switch still remains relatively low, when the switch is turned off, until the drain current reaches zero.

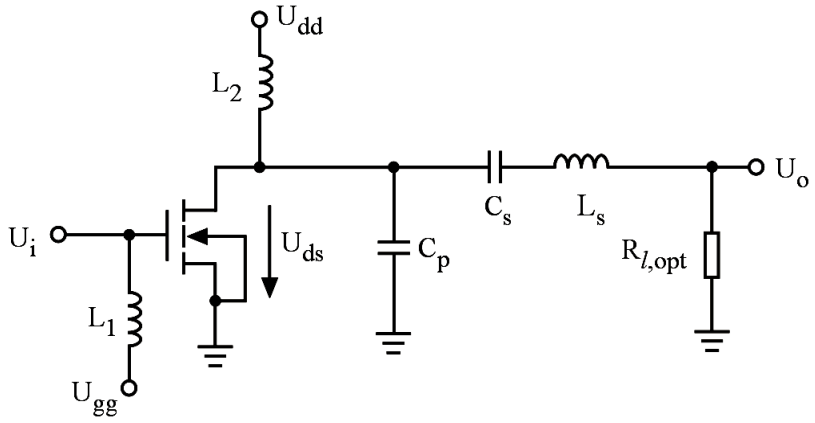


Fig. 2.12. The schematic of a class E power amplifier.

A series filter is adopted which consists of a series inductor L_s and a series capacitor C_s. This filter is tuned at the fundamental frequency to ensure a sine waveform on the load. $R_{l,opt}$ is the optimum load of the class E power amplifier, to which the 50 Ohm reference load should be transformed.

The waveforms of the voltage across the transistor and the current flowing through it are shown in Fig. 2.13. Clearly, the amplitude of the drain-source voltage can even reach $4.5U_{dd}$ [3]; therefore, the supply voltage of a class E power amplifier is usually lower than those used in the other classes.

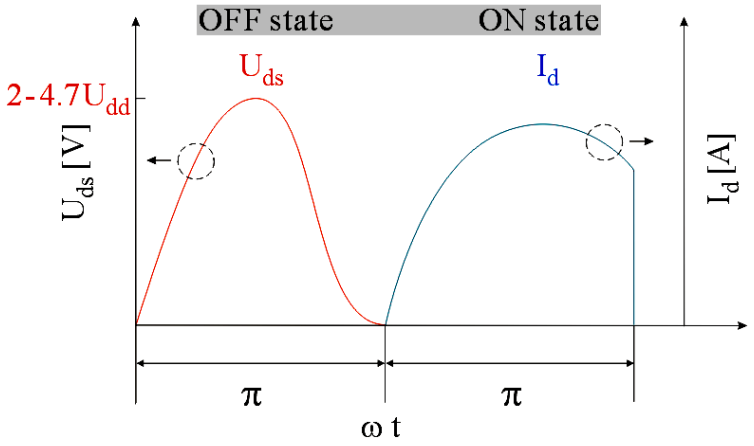


Fig. 2.13. Voltage and current waveforms for an optimum class E power amplifier.

The switch alternately opens and closes at the operating frequency, and the common choice of the duty cycle is 50 %. During the ON state, the transistor is overdriven and should provide a resistance as low as possible, while sustaining the current running through it. The voltage across the switch is zero and the current flows totally through the switch. Conversely, during the OFF state the transistor is in cut-off region and should provide very high impedance. It should also be able to sustain the voltage rise across its terminals. The entire current is flowing through the shunt capacitor C_p, charging and discharging it, and the switch voltage has a characteristic asymmetric waveform. The voltage and the current waveforms never simultaneously have non-zero values, so that no power is dissipated in the switch and the efficiency of this operation can theoretically be set to 100 %. However, on the other hand, this operation is clearly an extremely non-linear regime.

To realize the so-called soft switching, which means that the shunt capacitor will not be discharged through the switch resulting power loss, two conditions should be fulfilled [11], [12]. The first one is that the voltage returns to zero at the switch turn-on. The second one is zero voltage slope at the switch turn-on. These two conditions can mathematically be described by the following two equations:

$$U_{ds}(t = \tau/2) = 0\,, \tag{2.6}$$

$$\left.\frac{dU_{ds}}{dt}\right|_{t=\tau/2} = 0\,, \tag{2.7}$$

where U_{ds} is the drain-source voltage of the transistor as shown in Fig. 2.12 and t denotes the time; the time constant τ is equal to $2\pi/\omega$. These two conditions can be fulfilled by an appropriate choice of the reactive elements applied in the output network of the proposed amplifier circuit. These reactive elements can be described by the following equations [13]:

$$R_{l,opt} \approx 0.577 \frac{U_{dd}^2}{P_{out}}, \tag{2.8}$$

$$C_p = \frac{0.1837}{\omega \cdot R_{l,opt}}, \tag{2.9}$$

$$L_s = \frac{Q \cdot R_{l,opt}}{\omega}, \tag{2.10}$$

and

$$C_s = \frac{1}{\omega^2 \cdot L_s}. \tag{2.11}$$

Therefore, the mostly used reference impedance, namely 50 Ohm should first be transformed to $R_{l,opt}$ given in equation (2.8), according to the required output power. The other components can then be calculated in turn. The values of L_s and C_s are not uniquely determined; they depend on the chosen quality factor Q. The shunt capacitance C_p in equation (2.9) includes the output capacitance of the active device.

Unfortunately, the transistors used as switches always have a finite on-resistance, and the transition times from the off-state to the on-state and vice-versa are not negligible. In order to avoid power loss from the finite on-resistance of the switch, the FET is designed with a large gate width. On the other hand, the transition time of the input signal is increased in this way due to the enhanced parasitic capacitance. To obtain the shortest transition time from one state to the other, it is necessary to use a driving stage in the power amplifier, which feeds a squarewave-like signal to the input of the class E power amplifier. Class F power amplifiers are good candidates to generate such squarewave-like periodic signal.

In order to yield good performance in a class E power amplifier, suitable technology should be selected. Due to low substrate power loss and low parasitic capacitance and resistance, GaAs FET is frequently used. On the other hand, standard digital technology has seen a breakthrough in both performance and size through the use of device scaling to deep-submicron CMOS. For highly integrated, low-power, wireless transceivers operating near 1 - 2 GHz, CMOS may also prove to be a useful process to realize switching-mode power amplifiers. A CMOS class E power amplifier is proposed in this work.

2.2.3 The Class F Power Amplifier

The class F power amplifier was introduced at the end of the 1950s [14]. The basic idea of this operation mode is loading the active device output with appropriate terminations at fundamental and harmonic frequencies to improve the efficiency. Basically, open- and short-circuit terminations at odd and even harmonics are presented at the output of the transistors to shape the drain waveforms. This can be done with an output network including a load and harmonic resonators. Ideally, infinite number of resonators should be used, which is however not feasible in practice. Fig. 2.14 shows the schematic of a class F power amplifier using a field-effect transistor as the active device. This circuit has a band rejection filter tuned to the second harmonic in the output matching network, which consists of the inductors L_1, L_2, L_3 and the capacitors C_1, C_2 and C_3. Such a configuration will give a close approximation to a

square wave for the output voltage [15], hence can be used as the driving stage for a class E power amplifier.

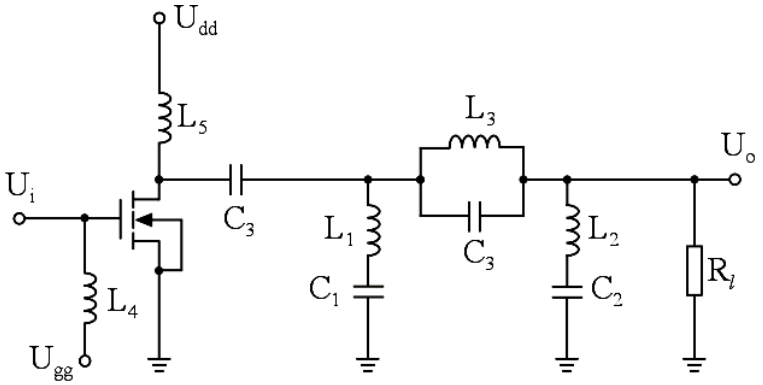


Fig. 2.14. The schematic of a class F power amplifier.

Another possibility to shorten several determined harmonics is to use the corresponding resonators in the biasing networks as shown later.

The waveforms of the drain voltage and the drain current are illustrated in Fig. 2.15 for the ideal case [16], i.e. the output matching network includes infinite number of resonators. Obviously, the drain-source voltage has a square waveform, while the drain current regularly presents a half sine waveform. In Fig. 2.15, the peak value of the drain-source current I_m is equal to πIq, where Iq is the quiescent current of the amplifier circuit. The quiescent current is defined using the following equation:

$$A_S = A_q\,, \tag{2.12}$$

where A_S is the area covered by the half sine waveform and A_q is the area covered by the quadrate shown in Fig. 2.16. The area A_S can be calculated as following

$$\begin{aligned} A_S &= \int_0^{\pi} I_m \sin x dx \\ &= I_m \cdot -\cos x\Big|_0^{\pi} \\ &= I_m(-\cos(\pi)+\cos(0)) \\ &= 2I_m\,. \end{aligned} \tag{2.13}$$

On the other side, the quadrate area A_q can be calculated as following:

$$A_q = I_q \cdot 2\pi\,. \tag{2.14}$$

Using (2.13) and (2.14) in (2.12), we obtain

$$I_m = I_q \cdot \pi\,. \tag{2.15}$$

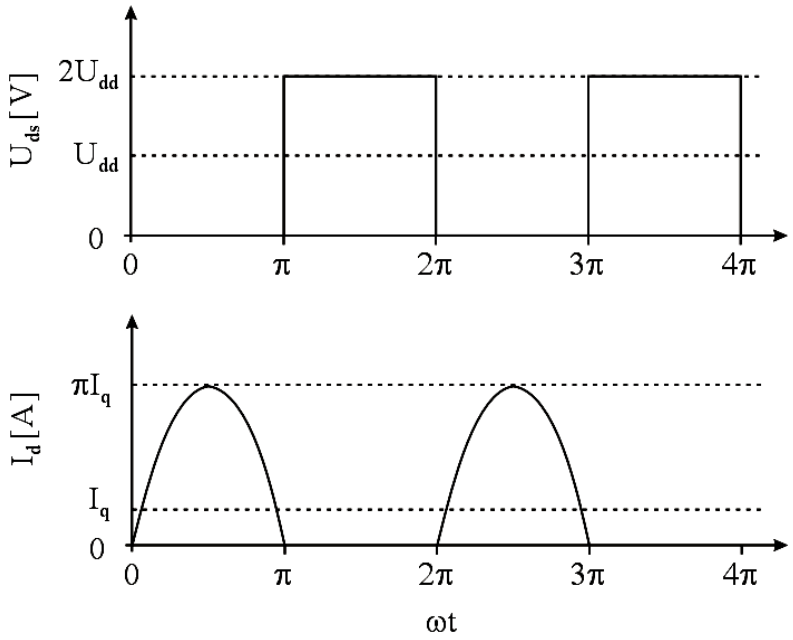


Fig. 2.15. Ideal waveforms of a class F power amplifier.

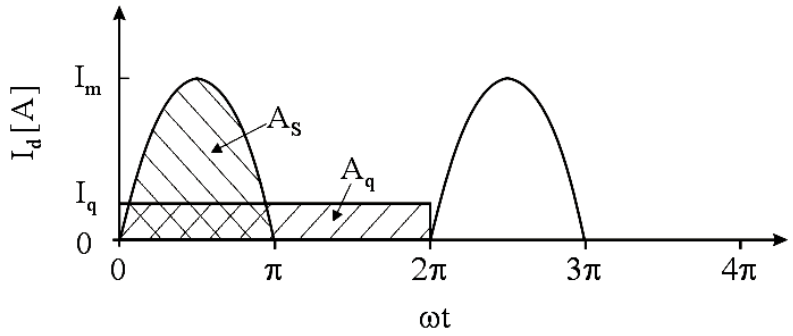


Fig. 2.16. Calculation of the quiescent current.

A class F power amplifier has in practice only limited number of resonators; hence the drain-source voltage waveform includes one or more odd harmonics and approximates a squarewave-like waveform. The current waveform includes even harmonics and approximates a half-wave rectified sine wave. Theoretically, the efficiency of a class F power amplifier increases as more harmonics are tuned. It approaches the ideal value of 100% in case that all the harmonics are rejected.

2.3 Active Devices for the Power Amplifier Design

A wide range of semiconductor technologies are being applied for RF power transistor applications between 500 MHz - 2.5 GHz. In practice, CMOS, LDMOS, HBT, and GaAs-HEMT are the mostly used technologies for the microwave power amplifier development. In this section, concise introductions of some of these technologies are provided. The novel HiVP configuration is also applied in practice for the power amplifier design; especially in case that a high supply voltage is used. The HiVP configuration can be considered as a single transistor as shown later, hence an analytical overview of it is given in this section.

2.3.1 CMOS Devices

The concept of metal-oxide-silicon field-effect transistors (MOSFETs) was generated in the early 1930s [17] [18]. In the mid-1960s, the complementary MOS (CMOS) devices (i.e. both n-type and p-type transistors) were introduced [19] [20], initiating a revolution in the semiconductor industry. Today, CMOS technology dominates the digital market, due to its low fabrication cost and low power consumption which happens only during the switching. The dimensions of MOS devices can be scaled down; hence their speed can be increased. CMOS technology is attractive for modern communication also due to the possibility of the monolithic-integration of the digital and analog circuits, which improve the overall performance and reduce the cost of packaging.

Fig. 2.17 illustrates a simplified structure of an n-type MOS transistor. Fabricated on a p-type substrate (also called the "bulk" or the "body"), the device consists of two heavily-doped n-regions forming the source (S) and drain (D) terminals. Note that the structure is symmetric with respect to S and D. A heavily-doped piece of polysilicon operates as the gate (G), which is insulated from the substrate by a layer of silicon dioxide (SiO_2) with a thickness of t_{ox}. L_n and W_n are the effective gate length and gate width, respectively. In reality, the substrate potential must generally be considered, which also influences the device characteristics. The connection to the substrate is usually realized through an additional p^+ region indicated as the bulk (B) in Fig. 2.17. Therefore, a MOSFET is normally regarded as a four-terminal device. In CMOS technologies, both NMOS and PMOS transistors are available. The PMOS devices are obtained by negating all of the doping types. They are usually placed in a "local substrate" called n-well, since they are fabricated on the same wafer as the NMOS devices [21].

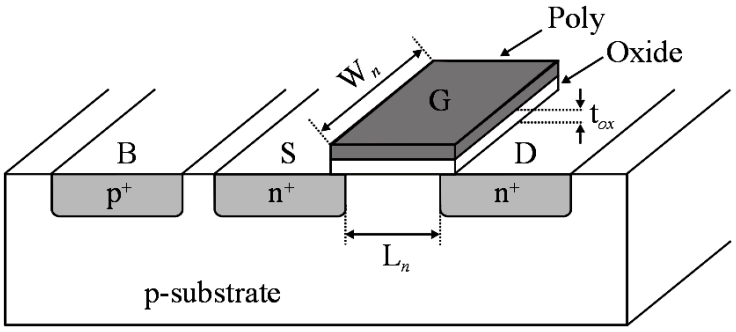


Fig. 2.17. The structure of an n-channel MOS device.

The output characteristic of an enhancement NMOS transistor [22], which indicates the relationship between the drain current I_d and the voltages can be described in Fig. 2.18. Obviously, there are two different regions, namely the triode region and the saturation region, before the device enters the breakdown region. They can be distinguished by $U_{ds} < U_{gs} - U_{th}$

and $U_{ds} > U_{gs} - U_{th}$, where U_{gs} and U_{ds} denote the voltage difference between the gate and source terminals, as well as between the drain and source terminals, respectively. U_{th} is the threshold voltage of the MOSFET. The threshold voltage of a MOSFET is usually defined as the gate voltage where an inversion layer forms at the interface between the insulating layer (oxide) and the substrate of the transistor. In case of $U_{gs} = 0\ V$, there is no conductive connection between the drain and the source. With a positive gate voltage $U_{gs} > U_{th}$ an n-conductive channel is generated, hence the current flows between the drain and the source, which rises with an increased U_{ds}. However, this channel is "pinched off" [22], if U_{ds} is larger than U_{gs}-U_{th}. That means the drain current remains constant as U_{ds} further increased.

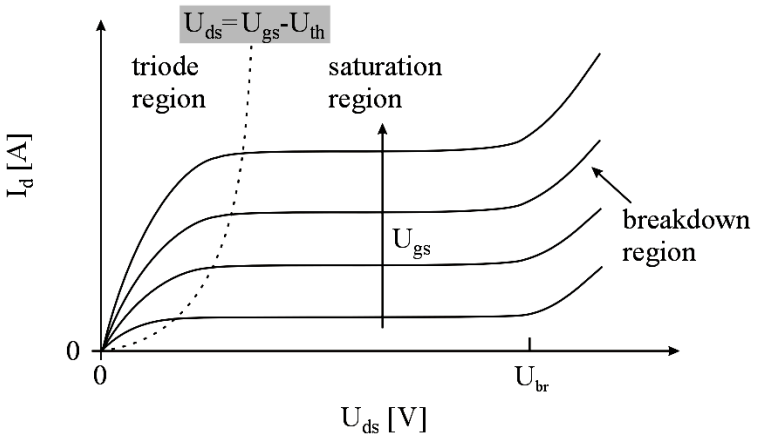


Fig. 2.18. The output characteristic of an enhancement MOS device.

In the triode region, the drain current I_d is given by [22]

$$I_d = \mu_n C_{ox} \frac{W_n}{L_n}\left[\left(U_{gs} - U_{th}\right)\cdot U_{ds} - \frac{1}{2}U_{ds}^2\right], \tag{2.16}$$

where μ_n is the mobility of the electrons and C_{ox} is the gate oxide capacitance per unit area. $U_{gs}-U_{th}$ is usually called as the "overdrive voltage", while W_n/L_n denotes the "aspect ratio". Obviously, in the triode region, I_d increases with the enhanced U_{gs} and U_{ds}, if U_{gs} excesses the threshold voltage.

In the saturation region, I_d can be expressed by

$$I_d = \frac{1}{2}\mu_n C_{ox} \frac{W_n}{L_n}\left(U_{gs} - U_{th}\right)^2, \tag{2.17}$$

indicating that the drain current is independent on U_{ds} in the ideal case. Since a MOSFET operating in saturation produces a current in response to its gate-source overdrive voltage, a figure of merit that indicates how well a device converts a voltage to a current can be defined. This figure of merit, which is called the "transconductance" and denoted by g_m, can be obtained by using the derivation of the drain current by the change in the gate-source voltage as shown below

$$g_m = \left.\frac{\partial I_d}{\partial U_{gs}}\right|_{U_{ds}=const.} = \mu_n C_{ox} \frac{W}{L}(U_{gs} - U_{th}). \tag{2.18}$$

In a sense, g_m represents the sensitivity of the device: for a high g_m, a small change in U_{gs} results in a large change in I_d. The direct relationship between g_m and I_d can be described with the following equation:

$$g_m = \sqrt{2\mu_n C_{ox} \frac{W}{L} I_d} \,. \tag{2.19}$$

Though I_d is theoretically independent on U_{ds} in the saturation region, the maximum allowable U_{ds} is limited by the avalanche-effect. With an oversized U_{ds}, the MOSFET can be destroyed rapidly. The breakdown appears first on the side of drain, since the largest difference between the gate-potential and the channel-potential exists at this point. Today, the breakdown voltage of a MOSFET, which is denoted by U_{br}, drops dramatically with the decreased dimensions of transistors. The typical value of a 120 nm CMOS technology is, for instance, only about 2.5 V.

In reality, for a MOS transistor, several second-order effects are not negligible. The "body effect" must first be considered, which means that the threshold voltage U_{th} of a MOS device can be varied by a bulk-voltage U_B. As U_B drops, U_{th} increases. Furthermore, the effective channel length gradually decreases as the potential difference between the gate and the drain increases, resulting in an enhanced I_d with an increased U_{ds}. Therefore, the expression of the I_d in the saturation region should be corrected and given as

$$I_d = \frac{1}{2}\mu_n C_{ox} \frac{W_n}{L_n}\left(U_{gs} - U_{th}\right)^2 \cdot (1 + \lambda U_{ds}) \,, \tag{2.20}$$

where λ is the channel-length modulation coefficient. This effect is called the "channel-length modulation".

The small-signal model of a MOSFET is given in Fig. 2.19. It can be seen from the output characteristic that a MOS device operates as a voltage-controlled current source in the saturation region. Therefore, a voltage-dependent current source equal to $g_m \cdot U_{gs}$ is incorporated at the output. The channel-length effect is modelled by the drain-source resistance R_{ds}, whose value is approximately $1/(\lambda \cdot I_d)$. The bulk potential, which influences the threshold voltage and hence the gate-source overdrive, is modelled by a second current source $g_{mb} \cdot U_{bs}$, where g_{mb} is the secondary transconductance caused by bulk potential. That is, the bulk is regarded here as the second gate. Furthermore, to predict the RF behaviour, the device capacitances between every two of the four terminals are also included in the complete small-signal model of a transistor.

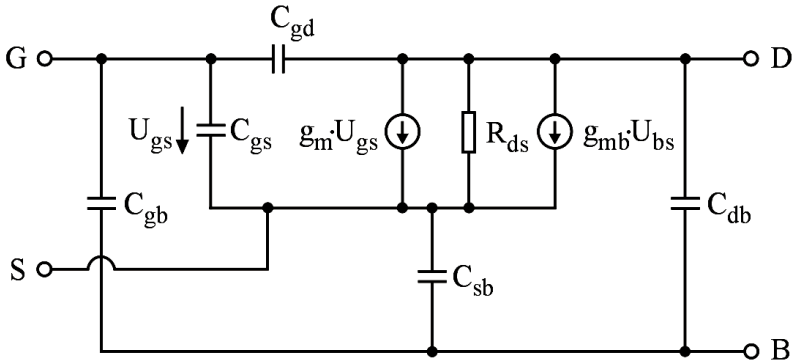


Fig. 2.19. The complete MOS small-signal model.

2.3.2 HiVP Configuration

As mentioned above, CMOS technology is widely used due to lots of benefits. However, the bottleneck of this technology is the breakdown voltage. The gate-drain voltage of a MOS transistor is especially critical due to the field distribution along the channel and the extreme thin gate oxide. Therefore, a single MOS device, especially one of the deep-submicron CMOS technologies, is not suitable for power amplifier design. Recently, the HiVP configuration [23] becomes attractive to solve this problem. The HiVP is a high voltage and high power device configuration in which several devices are connected DC and RF in series, so that the large output voltage can be divided by all the cascaded devices. With some methods it is even possible to share the large voltage equally [24].

Figure 2.20 illustrates the schematic of a CMOS HiVP configuration. Several devices (e.g. three devices as shown here in Fig. 2.20) are connected in series. Therefore, the currents flowing through the transistors are all the same. Ideally, the transistors have the same operating points. The resistors R_1-R_3 are used as the voltage divider to ensure that U_{ds} of all the transistors are exactly the same. Clearly, the DC voltage U_{ds} of T1-T3 are equal to the DC voltage drop of R_1-R_3, respectively, However, the DC voltage U_{ds} of T3 is equal to the DC voltage drop of R_3 plus the gate-source voltage of T3. Therefore, R_1 and R_2 should have the same value while R_3 must be much smaller than the others. The DC voltages between gate and source of all the devices are always the same and equal to U_{gg}, because the transistors are absolutely identical and the same current flows through them. The resistor R_3 serves simultaneously as a feedback to allow the gate voltages of T1 – T3 to swing with the RF output signal. Therefore, the voltages between the gate and the drain of each device can remain smaller than the maximum allowable voltage for the transistors. The two inductors L_1 and L_2 serve as RF chokes which feed DC power to the gate and drain, respectively.

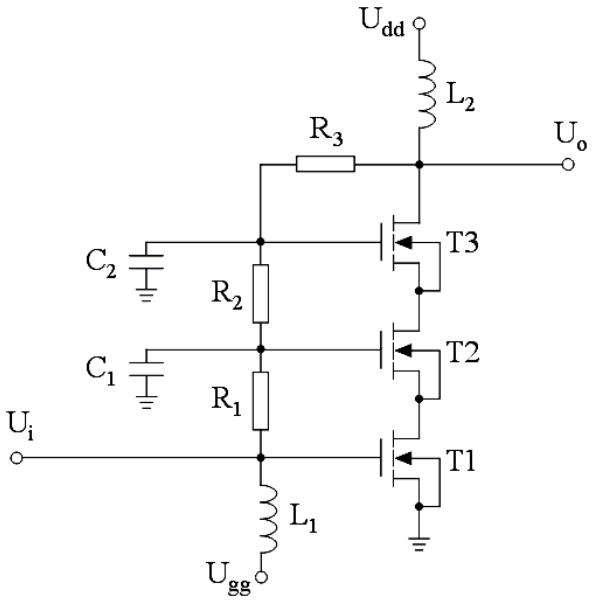
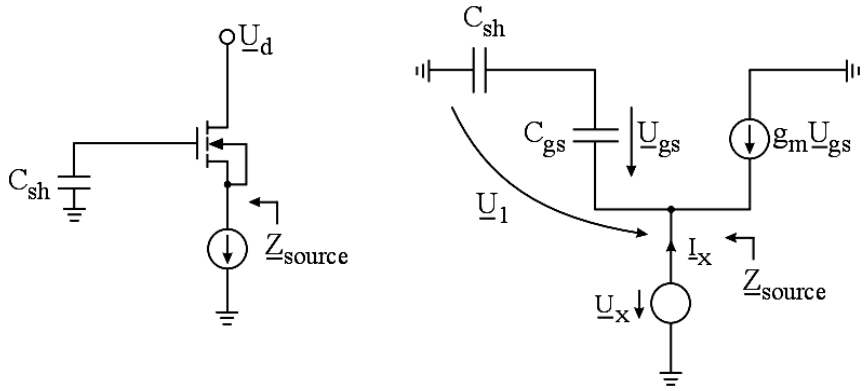


Fig. 2.20. A HiVP configuration with three transistor devices.

As mentioned above, the main function of a HiVP structure is to realize a voltage division. That means the drain voltages of the transistors are different. Since the same current flows through all the transistors in a row, the voltage difference can only be fulfilled by the increasing drain impedance from the lower transistor to the upper one. In this sense, the HiVP configuration operates also as a power combiner. The output power of this configuration is the

combined sum of all the power achieved from each individual transistor. In this manner, a high output power is achievable.

The impedance level seen at the drains of T2 and T1 can be adjusted by the capacitors C_2 and C_1, which are connected between the gates of the floating MOS transistors and the ground. Noting that the drain impedance of these two transistors is equal to the impedance at the source input of the upper device, the calculation of this impedance can be done by using a current source at the source of the device as shown in Fig. 2.21 (a). The small-signal equivalent circuit is shown in Figure 2.21 (b). C_{sh} symbolizes C_1 and C_2 in Fig. 2.20 and Z_{source} indicates the impedance seen at the source.



(a) Calculation of the source impedance using a current source
(b) The small-signal equivalent circuit of a)

Fig. 2.21. The calculation of the impedance seen at the drain.
(C_{sh}: The shunt capacitance between gate and ground)

Noting that

$$\underline{U}_1 = -\underline{U}_X , \tag{2.21}$$

following equations can be obtained

$$\underline{U}_{gs} = \frac{1/\, j\omega C_{gs}}{1/\, j\omega C_{gs} + 1/\, j\omega C_{sh}} \cdot \underline{U}_1$$

$$= -\frac{C_{sh}}{C_{sh} + C_{gs}} \cdot \underline{U}_X . \tag{2.22}$$

On the other hand

$$\underline{I}_x = -\underline{U}_{gs} \cdot \left(g_m + j\omega C_{gs} \right). \tag{2.23}$$

Substituting equation (2.22) into equation (2.23) yields

$$\underline{I}_X = \frac{C_{sh}}{C_{sh} + C_{gs}} \cdot \left(g_m + j\omega C_{gs} \right) \cdot \underline{U}_X . \tag{2.24}$$

This leads to the impedance at the source input of the device

$$\underline{Z}_{source} = \frac{\underline{U}_X}{\underline{I}_X} = \frac{1}{g_m + j\omega C_{gs}} \cdot \frac{C_{sh} + C_{gs}}{C_{sh}}. \tag{2.25}$$

If the transistor has a gate width of several millimetres, the gate-source capacitance normally has a value in the range of several pikofarad. Therefore, $j\omega C_{gs}$ at several Gigahertz is much smaller than g_m (normally > 1). The equation above can approximately be rewritten in

$$\underline{Z}_{source} \approx \frac{1}{g_m} \cdot \frac{C_{sh} + C_{gs}}{C_{sh}} = \frac{1}{g_m} \cdot \left(1 + \frac{C_{gs}}{C_{sh}}\right). \tag{2.26}$$

Observe from equation (2.26) that C_{sh} is the unique but useful variable to adjust the impedance level seen at the drains of each individual transistor. The smaller C_{sh}, the larger is Z_{source}. For a high output power, the voltage swing of the drain of the top device must be large enough (e.g. peak-peak voltage > 10 V). On the other hand, the voltage swing of the drain of the bottom device must remain small (peak value smaller than 1.5 V) because the source of the bottom device is directly connected to the ground. For an equal distribution of the large drain voltage of the top device, the higher devices must have a larger voltage swing. As mentioned above, the larger voltage swing can only be obtained with a higher impedance level seen at the drain. According to equation (2.26), the gate of a higher device must be connected with a smaller capacitor, which means that the following condition must be fulfilled [24].

$$C_2 < C_1 \tag{2.27}$$

The HiVP configuration can be regarded as a single transistor [23]. The gate of the first transistor can be considered as the input and the drain of the top transistor can be considered as the output. The difference between such a configuration and a conventional single device is the ability to carry a larger voltage. Therefore, a large supply voltage can be applied to generate a larger output power.

2.3.3 LDMOS Devices

The lateral double diffused MOS (LDMOS) power transistor is a development of the MOS-technology, which has an increased breakdown voltage. Today, silicon LDMOS technology has a strong position in base station applications due to its benefits, which are improved efficiency; higher peak-power capability and lower cost-per-watt performance. Furthermore, LDMOS power amplifiers are widely used in modern wireless communication systems also due to their better intermodulation distortion (IMD) performance compared to other competing technologies [25].

The operating theories of an LDMOS transistor are based on those of a conventional MOS device. However, compared to a conventional MOS transistor, the breakdown voltage of an LDMOS transistor is significantly increased by using a long n-drift region at the drain terminal. Fig. 2.22 shows the device structure of a LDMOS transistor [26]. The drift region is low doped and hence has relatively high resistivity. Therefore, the high electrical field strength at the drain can be degraded. The usage of this drift region can boost the breakdown voltage of the transistor, even up to 70 V, so that a higher supply voltage can be used for a power amplifier design to obtain a large output power. Today, the commercial LDMOS

transistors typically have a supply voltage of 26 V and are available with power outputs of more than 100 W at 2 GHz.

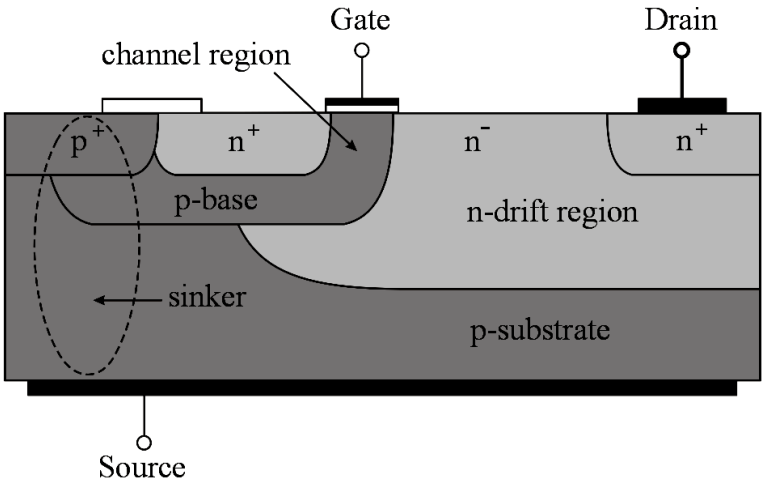


Fig. 2.22. The cross section of an LDMOS transistor [26].

On the other hand, the additional n-drift region involves a relatively large output capacitance which leads to an optimum output impedance of the device well below 50 Ohm. Therefore, a broadband output matching is quite difficult using an LDMOS device. The short channel region, called the p-base, is formed by lateral diffusion of a p-type implantation that will enhance RF performance of the transistor [27]. The source terminal is directly structured on the back side of the wafer, so that bonding wire of the source terminal can be eliminated, thus the source inductance can be greatly reduced. A thicker gate oxide is implemented, which reduces the feedback capacitor between the drain and the gate, so that the transit frequency is enhanced.

2.4 Introduction and Analyses of Conventional Circuit Design Concepts

In this section, the design concept as well as the functional principle of the common-source (CS) single-ended stage is first introduced. According to open-circuit time constants method, the high-frequency response of this circuit is discussed. It can be proven, that the problem often occurring in analog circuits, the Miller effect, causes the limitation for the bandwidth. Therefore, this effect should be diminished in the radio frequency circuits. The cascode circuit is introduced on this account, which effectively eliminates the Miller effect. The functional principle of the cascode circuit is provided in subsection 2.4.4. Today, the differential circuit is widely used in practice, its benefits and drawbacks are described in subsection 2.4.5.

2.4.1 Common-Source Single-Ended Stage

In case of a common-source single-ended stage, the gate and the drain of the MOS transistor are separately used as the input and the output of the circuit, while the source and the bulk of the transistor are directly connected to the ground. The schematic of such a configuration is depicted in Fig. 2.23, where U_i and U_o denote the input and the output voltage, respectively. The supply voltage is indicated by U_{dd}.

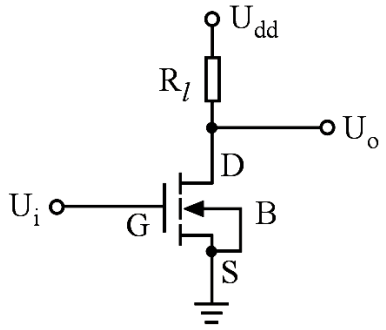


Fig. 2.23. The schematic of a common-source single-ended stage.

The MOSFET converts variations in its gate-source voltage to a variation of the drain current, which can pass through the load R_l to generate an amplified output voltage. Both the large-signal and the small-signal analysis show, that the gain (G) of this common-source stage is proportional to the transconductance of the transistor g_m and the load impedance as expressed below [22]

$$G = -g_m R_l , \tag{2.28}$$

when the body effect is neglected. The minus sign in equation (2.28) indicates that the input and the output voltage are just 180° out-of-phase.

In practice, the load R_l shown in Fig. 2.23 is generally a resistive load, which can directly be realized with a resistor. It can also be realized by using a MOSFET that operates in the deep triode region. Moreover, the load can also be a diode-connected MOSFET to obtain a relatively linear gain, or be a current-source load to maximise the output voltage swing etc. Sometimes, the load is connected at the source of the transistor device to form a source degeneration circuit. In this manner, a linear behaviour of a common-source stage can be obtained.

2.4.2 Estimation of the High-Frequency Bandwidth

In microwave region, it is sometimes desired to approximately predict the bandwidth of the proposed radio frequency circuits. Leastwise, it is demanded to have the insight to find the components which are responsible for the limitation of the circuit bandwidth. Two such approximate methods are open-circuit time constants (OCTC) and short-circuit time constants (SCTC), which are used to estimate the high-frequency 3-dB point and the low-frequency 3-dB point, respectively [28].

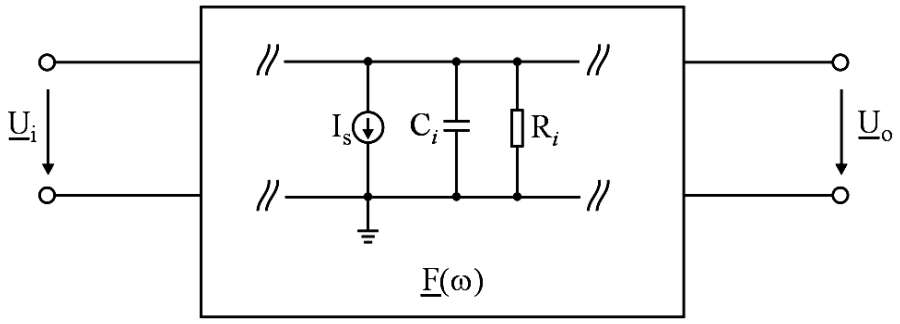


Fig. 2.24. A two-port network.

Using OCTC, let's consider the high-frequency bandwidth of an arbitrary two-port network shown in Fig. 2.24. This network consists only of sources, n resistors R_i and n capacitors C_i, where i is an arbitrary integer between 1 and n. The resistor R_i in Fig. 2.24 denotes the effective resistance facing ith capacitor C_i with all the other capacitors removed (open-circuited). The all-pole transfer function $\underline{F}(\omega)$ of this two-port network can approximately be described as

$$\underline{F}(\omega) = \frac{\underline{U}_o}{\underline{U}_i} = \frac{a_0}{(\tau_1 j\omega + 1)(\tau_2 j\omega + 1)\cdots(\tau_n j\omega + 1)}, \tag{2.29}$$

whereas τ_i are the various time constants, determined by the product of the resistors and capacitors $R_i \cdot C_i$. Multiplying out the terms in the denominator leads to the following polynomial

$$b_n \cdot (j\omega)^n + b_{n-1} \cdot (j\omega)^{n-1} + \cdots b_1 \cdot (j\omega)^1 + 1, \tag{2.30}$$

where the coefficient b_n is the product of all the time constants and b_1 is the sum of them. Generally, near the 3-dB limit-frequency, the first-order term in equation (2.30) dominates over the higher-order terms, therefore, the reasonable approximation of the transfer function can be simplified as

$$\frac{\underline{U}_o}{\underline{U}_i} \approx \frac{a_0}{\left(\sum_{i=1}^{n} \tau_i\right) \cdot j\omega + 1}. \tag{2.31}$$

As shown later, (2.31) could further be simplified if only the dominant RC-constant arising from the input is considered. The estimated high-frequency bandwidth of the system $\omega_{3\text{-dB}}$ is then simply the reciprocal of the effective time constant, given as

$$\omega_{3-dB} \approx \frac{1}{\left(\sum_{i=1}^{n} \tau_i\right)}. \tag{2.32}$$

Thus, each time constant represents a local bandwidth degradation term. Naturally, not all capacitors in a network belong to the OCTC calculation. For instance, the large coupling capacitors used to connect the output of one stage to the input of the next stage is not counted to such calculation. As a consequence, the removal of a capacitor that belongs to the OCTC calculation should lead to an increase of the high-frequency power gain.

2.4.3 High Frequency Response and Miller Effect

The gain of a single-ended common-source stage expressed in equation (2.28) is only valid for low-frequency application. To show the frequency response of the circuit in the high-frequency region, the capacitance of the MOSFETs given in Fig. 2.19 must also be taken into account. The complete small-signal equivalent circuit of a single-ended common source stage is depicted in Fig. 2.25, where R_g and R_l are the gate and load resistance of the transistor, respectively.

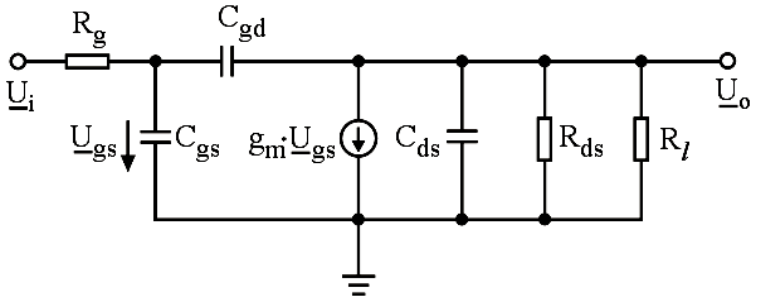


Fig. 2.25. The small-signal equivalent circuit of a single-ended common source stage.

Known from the OCTC method, the effective resistance R_i facing each capacitor should first be calculated, under the condition that all the other capacitors are open-circuited. Obviously, only R_g is connected to C_{gs}, as well as the parallel circuit of R_{ds} and R_l to C_{ds}. An important phenomenon that occurs in many analog circuits, namely the "Miller effect" [22], arises from C_{gd}. Usually, the Miller effect is caused by connecting a capacitance across two nodes that have an inverting gain between them. Using a test current source, the resistance facing C_{gd} can be calculated as follows [28]

$$R_{C_{gd}} = R_g + \left(R_l \,||\, R_{ds}\right) + g_m \cdot R_g \cdot \left(R_l \,||\, R_{ds}\right). \tag{2.33}$$

According to equation (2.31), the gain of the common-source stage in dependence of the frequency can approximately be described as follows.

$$\frac{\underline{U}_o}{\underline{U}_i} = \frac{-g_m \cdot R_l}{\left[C_{gs} \cdot R_g + C_{ds} \cdot \left(R_l \,||\, R_{ds}\right) + C_{gd} \cdot R_{C_{gd}}\right] \cdot j\omega + 1} \tag{2.34}$$

The largest limit-factor in equation (2.34) for the bandwidth is associated with C_{gd}. Even though the value of C_{gd} is conventionally small, however, its effect is Miller-multiplied by the gain g_m and the resistances.

Actually, the gate-drain capacitor C_{gd} in Fig. 2.25 provides a feedforward path that conducts the input signal to the output at very high frequencies, resulting in a slope in the frequency response. Therefore, C_{gd} is a significant capacitor which belongs to the OCTC calculation. It is worth to invest some effort to mitigate its effect by somewhat improved circuit topologies.

2.4.4 Cascode Circuits

As mentioned in section 2.4.3, the Miller effect is the most significant factor for the degradation of the bandwidth. Therefore, it is important for RF circuits design to figure out how to mitigate the Miller effect. One possibility is to prevent the coupling-capacitance between the input and the output across the gain stage. The cascode circuit is such a candidate, which effectively eliminates the Miller effect and therefore is widely used for microwave circuit design. The schematic of a cascode circuit is depicted in Fig. 2.26.

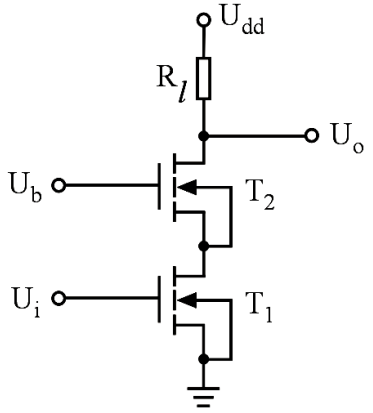


Fig. 2.26. The schematic of a cascode circuit.

Obviously, the cascode topology is a cascade of a common-source stage (T_1) and a common-gate stage (T_2). The input device T_1 generates a small-signal drain current proportional to U_i and the cascode device T_2 simply routes the current to R_l. Therefore, the voltage gain of a cascode stage is just equal to that of a common-source single-ended stage.

Intuitively, the output is now at the drain of the T_2 while the input is at the gate of T_1. There is no capacitance directly across these two nodes; hence the Miller multiplication is dramatically reduced. Using the high-frequency mode of the cascode circuit, which is depicted in Fig. 2.27, the suppression of the Miller effect is analytically presented. In this figure, the main parasitic capacitances of both transistors are shown. Though C_{gd1} is still connected between the gate and the drain of the input device, the resistance facing it has been partially changed. In this case, the gate resistance R_g facing it remains, however, the other resistance facing it is now the resistance of the parallel circuit comprising R_{ds1} and the resistance seen looking toward the source of T_2. Similar to the calculation of the resistance seen at the source of the upper device in the HiVP configuration, the resistance seen looking toward the source of T_2 can be calculated with the result of $(1/g_{m2}) \| R_{ds2}$, neglecting the body effect of the devices. The total resistance facing C_{gd1} can now be expressed as

$$R_{C_{gd1}} = R_g + R_{ds1} \| \left(\frac{1}{g_{m2}} \| R_{ds2} \right) + g_{m1} \cdot R_g \cdot \left(R_{ds1} \| \left(\frac{1}{g_{m2}} \| R_{ds2} \right) \right). \tag{2.35}$$

Since $1/g_{m2}$ has a much smaller value than R_{ds1} and R_{ds2}, the equation above can approximately be rewritten in

$$R_{C_{gd1}} = R_g + \frac{1}{g_{m2}} + g_{m1} \cdot R_g \cdot \frac{1}{g_{m2}}. \tag{2.36}$$

Compared to the parallel resistance $R_l \,||\, R_{ds}$ in equation (2.33) for the common source single-ended stage, $1/g_{m2}$ also has a much smaller value, leading to a significant degradation of the Miller effect caused by C_{gd1}.

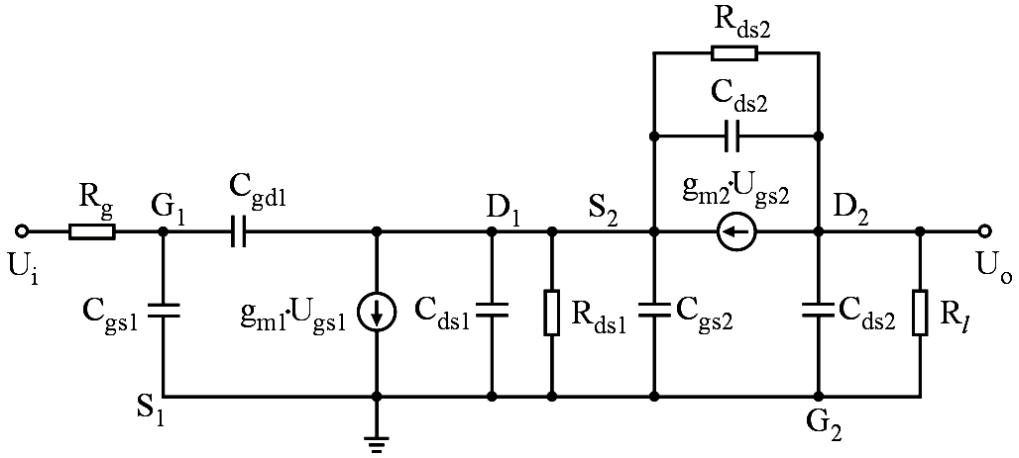


Fig. 2.27. The small-signal circuit of a cascode circuit.

Clearly, since one more transistor is applied, the cascode circuit has more parasitic capacitances than the single-ended stage. However, since the Miller multiplication is definitively removed from the entire configuration, every capacitance in the cascode circuit makes only a small contribution to the effective time constant, resulting in a relatively small sum. Therefore, the bandwidth of the circuit is increased.

Another important benefit of the cascode circuit is its higher output impedance, compared to a single-ended common-source stage. The output impedance of a cascode circuit can be described as follows [29]

$$\begin{aligned} R_{out} &= \left(1 + g_{m2} \cdot R_{ds2}\right) \cdot R_{ds1} + R_{ds2} \\ &\approx g_{m2} \cdot R_{ds2} \cdot R_{ds1}, \end{aligned} \tag{2.37}$$

assuming $g_{m2} R_{ds2} >> 1$. This result can be regarded that the output impedance of T_1 is greatly increased by a factor of $g_{m2} \cdot R_{ds2}$. Therefore, the attenuation of the RF signal at the output of the circuit is reduced.

Despite the benefits of a cascode circuit mentioned above, the drawbacks of such a configuration are not negligible. Assuming both transistors T_1 and T_2 operate in saturation region, the output voltage swing of the cascode circuit is decreased by at least the overdrive voltage of T_2. Therefore, the maximum achievable output power is also reduced. Just on this account, cascode topologies extended to three or even more stacked devices are usually not used in practice, though they have even larger output impedance than the conventional cascode circuit.

2.4.5 Differential Amplifier

The differential operation has become an important choice in today's high-performance analog and mixed-signal circuits. The basic schematic of a differential pair is presented in Fig. 2.28. Two identical single-ended devices T_1 and T_2 are connected in parallel. They also have the same supply voltage U_{dd} and identical loads, i.e. $R_{l1} = R_{l2} = R_l$. U_{i1}, U_{i2}, U_{o1} and U_{o2} separately denote the input as well as the output voltages of the T_1 and T_2. The differential pair employs a current source I_S to minimise the dependence of the sum of the bias currents I_{d1} and I_{d2} flowing through the devices on the input common-mode level.

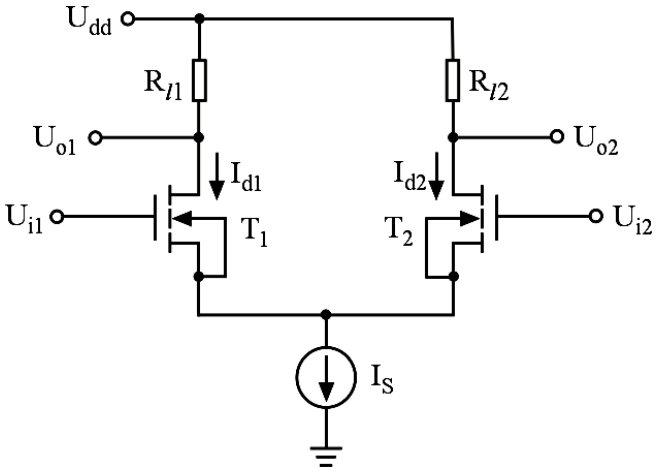
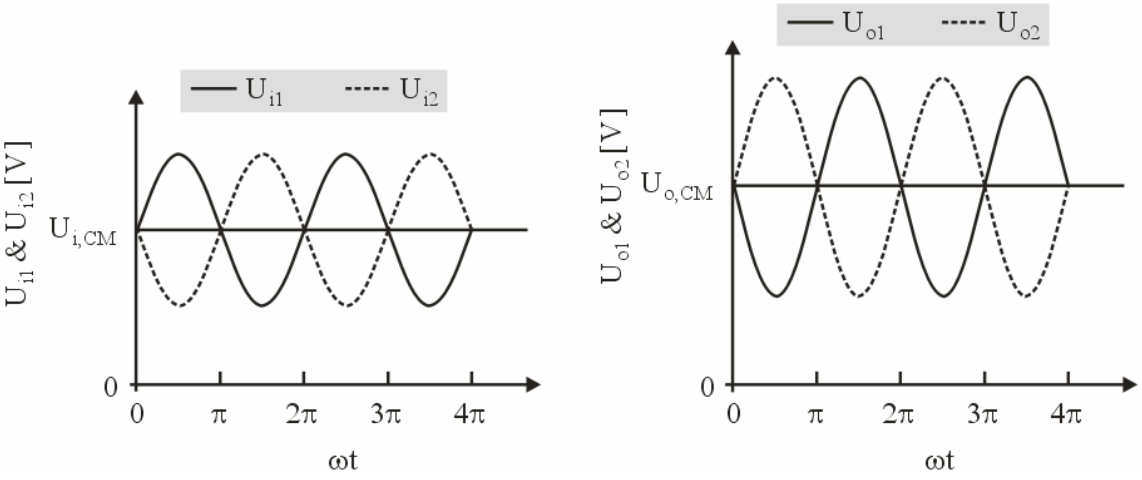


Fig. 2.28. The basic schematic of a differential amplifier.

The waveforms in the time domain of the input and the output voltages are given in Fig. 2.29. The two input signals have the same center potential named common-mode (CM) level, but are 180° out of phase as shown in Fig. 2.29 (a). The input common-mode voltage level is denoted as $U_{i,CM}$ in this figure. The output signals of the two signal-ended stages are also 180° out of phase with each other and separately with their input signals as shown in Fig. 2.29 (b).



(a) The input voltages (b) The output voltages

Fig. 2.29. The waveforms of the input and the output signals of a differential pair.

Compared to a single-ended stage, an important advantage of differential operation is the higher immunity to the environmental noise, since the environmental noise disturbs each individual CM signal, but not the differential output signal. The other useful property of the differential signalling is the doubled achievable voltage swings, as the peak to peak swing for $U_{o1}-U_{o2}$ is equal to $2\cdot\left[U_{dd}-\left(U_{gs}-U_{th}\right)\right]$, while the maximum output swing of U_{o1} or U_{o2} is only equal to $U_{d}-\left(U_{gs}-U_{th}\right)$ [22].

Using small-signal analysis, the voltage gain of the differential pair can be given with the following equation, where $g_{m,d}$ is the transconductance of the differential pair,

$$\frac{U_{o2}-U_{o1}}{U_{i1}-U_{i2}}=-g_{m,d}R_{l}. \tag{2.38}$$

According to equation (2.19), $g_{m,d}$ in equation (2.38) is $1/\sqrt{2}$ times that of a single-ended stage, if the single-ended transistor has the same dimension and is also biased at I_S, since the bias current in the differential pair must be equally divided by the two parallel transistors. In this respect, the gain of a differential pair is smaller than that of a single-ended stage. On the other hand, a differential pair, which has the same device-dimensions and the same load impedance as a single-ended stage, can achieve the same gain as the single-ended stage only at the cost of twice the bias current.

Similar to the single-ended common-source stage, the load of a differential pair can also be selected as resistive loads, diode-connected loads, or current-source loads, according to different applications.

2.5 Impedance Matching

As mentioned in chapter 2.1, in order to obtain the maximum output power, the reference impedance (usually 50 Ohm) must be transformed to the optimum input and output impedance Z_{in_opt} and Z_{out_opt} of the transistor. Matching networks should be used at the input and the output of the amplifier circuit. Considering the specific problems of RF power amplifiers, the matching circuits must meet the following requirements as much as possible:

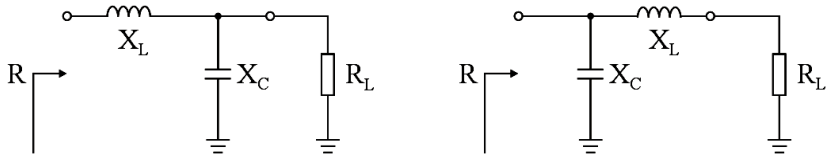
1. Transform the source and load impedance in the optimal operating impedance.
2. Maintain the specified amplitude- and phase-frequency response over a certain frequency range.
3. Attenuate the higher harmonics. Thus, only low- or band-pass circuits are employed.
4. Have insignificant power loss, i.e. maintain high efficiency. This can be done by using only lossless circuit elements.
5. Other requirements related to cost, size, weight, reliability and practicability.

A very useful graphical aid to the analysis of the impedance transformations is provided by the Smith chart, which is basically a plot of all passive impedances in a reflection coefficient chart of unit radius [30]. Furthermore, using the Smith chart, the frequency dependence of the S-parameters and other amplifier characteristics can be presented.

In this section, discrete matching networks are first introduced, which are above all indispensable for the integrated circuits (IC) design in the frequency region of Gigahertz. On the other hand, microstrip lines already find extensive use as passive circuit elements. Therefore, characteristics of microstrip transmission lines are also presented here. Based on the knowledge of microstrip matching networks, the multi-section transmission line impedance transformer is discussed.

2.5.1 Discrete Matching Networks

There are several different methods for the design of matching networks with lumped elements. The simplest lumped matching networks are the two-reactance matching networks, typically the L-circuits shown in Fig. 2.30.



(a) L-matching network for $R < R_L$ (b) L-matching network for $R > R_L$

Fig. 2.30. The discrete L-matching networks.

Assuming the impedances to be matched R_L and the target impedance R are purely resistive, and X_L as well as X_C are the reactance of the passive elements at the operating frequency *f*, the design equations for the matching network in Fig. 2.30 (a) can be given as [31]

$$X_L = R\sqrt{\frac{R_L}{R} - 1}\,; \qquad X_C = R_L / \sqrt{\frac{R_L}{R} - 1}\,. \tag{2.39}$$

This circuit is only used for $R < R_L$. If $R > R_L$, the input and output of the network of Fig. 2.30 (a) must be interchanged as shown in Fig. 2.30 (b). The design equations are then

$$X_L = R_L\sqrt{\frac{R}{R_L}-1}\,; \qquad X_C = R/\sqrt{\frac{R}{R_L}-1}\,. \tag{2.40}$$

The quality factor, which denotes the ratio of the stored average energy and the energy loss, is also an important characteristic of the matching networks. Assuming the circuit has a bandwidth of Δf around the center frequency f_0, there is

$$Q = \frac{f_0}{\Delta f}\,. \tag{2.41}$$

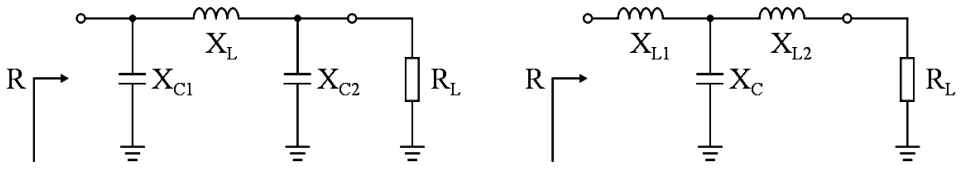
The equation above means, higher Q implies smaller bandwidth. For the L-matching networks, the quality factor is determined by the source resistance R_L and the first passive component connected with it. For instance, the quality factor for the matching network shown in Fig. 2.30 (a) is determined by the R_L and parallel connected passive component X_C. It can be calculated with

$$Q = \frac{R_L}{X_C}\,. \tag{2.42}$$

Similarly, the quality factor of the L-matching networks shown in Fig. 2.30 (b) is determined by the source resistance R_L and series connected passive component X_L, which can be calculated as follows

$$Q = \frac{X_L}{R_L}\,. \tag{2.43}$$

In practice, the three-reactance discrete matching networks are mostly used, since the values obtained in the two-reactance matching networks may be impractical and there is no design flexibility. For instance, there is no freedom to vary the quality factor Q, hence to change the bandwidth. The typical three-reactance matching networks are π- and T-matching networks as shown in Fig. 2.31.



(a) π-matching networks (b) T-matching networks

Fig. 2.31. The three-reactance matching networks.

In case of $R_L\left(1+Q^2\right) > R$, the design equations for the π-matching network in Fig. 2.31 (a) can be given as follows [31]

$$X_{C1} = \frac{R}{Q}\,; \quad X_{C2} = \frac{R_L}{\sqrt{\frac{R_L}{R}\left(1+Q^2\right)-1}}\,; \quad X_L = \frac{QR}{1+Q^2}\left[1+\frac{1}{Q}\sqrt{\frac{R_L}{R}\left(1+Q^2\right)-1}\right]. \tag{2.44}$$

The quality factor Q can be chosen according to the circuit bandwidth, harmonic attenuation and efficiency. Increasing Q can enhance the filtering effect of the matching network, but reduces the bandwidth and efficiency. Under the condition $R\left(1+Q^2\right)>R_L$, the design equations for the T-matching network shown in Fig. 2.31 (b) can be given as [31]

$$X_{L1}=QR;\quad X_{L2}=R_L\sqrt{\frac{R}{R_L}\left(1+Q^2\right)-1};\quad X_C=\frac{R\left(1+Q^2\right)}{Q+\sqrt{\frac{R}{R_L}\left(1+Q^2\right)-1}}. \tag{2.45}$$

Also here, the quality factor Q can in principal be chosen arbitrarily, but the recommended values of Q usually range from 1 to 10.

2.5.2 Impedance Transforming Property of a Transmission Line

Microstrip lines are used extensively in building microwave transistor amplifiers because they are easily fabricated using printed-circuit techniques. Network interconnections and the placement of passive and transistor devices are easily realized on its metal surface. The superior performance characteristics of the microstrip line make it one of the most important mediums of transmission in microwave transistor amplifiers. Since the microstrip line belongs to the most relevant members of transmission lines, the impedance transforming property of a transmission line is first discussed in this subsection.

Fig. 2.32 illustrates a lossless transmission line terminated in an arbitrary load impedance $\underline{Z}_L$, where the characteristic resistance Z_0 indicates the ratio of voltage to current for such a travelling wave and β is the imaginary part of the propagation constant. The transmission line has a length of *l*. The relationship between the to be matched impedance $\underline{Z}_L$, the goal impedance $\underline{Z}_{in}$ and all these parameters is to be found. In order to simplify the following calculation, the position of $\underline{Z}_L$ is defined as zero, so that the transmission line is located in the negative axis as shown in Fig. 2.32.

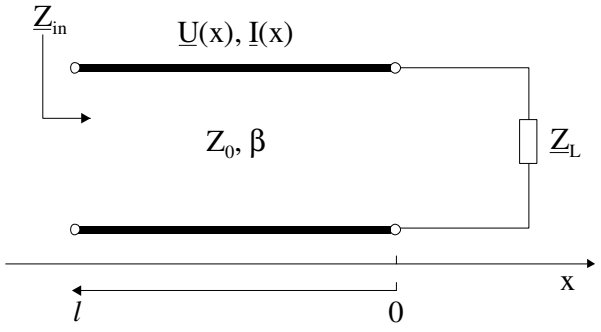


Fig. 2.32. The lossless transmission line impedance transformer.

Assume that an incident wave of the form $U_0^+e^{-\beta x}$ is generated from a source at $x<0$. In case that $\underline{Z}_L\neq\underline{Z}_0$, a reflected wave is excited. The total voltage on the line is then the sum of the incident and the reflected waves and is described by the following equation [32]

$$\underline{U}(x)=\underline{U}_0^+e^{-j\beta x}+\underline{U}_0^-e^{j\beta x}, \tag{2.46}$$

where U_0^+ and U_0^- indicate the magnitudes of the incident and the reflected voltages, respectively. The total current on the line is

$$\underline{I}(x) = \frac{\underline{U}_0^+}{Z_0} e^{-j\beta x} - \frac{\underline{U}_0^-}{Z_0} e^{j\beta x}. \tag{2.47}$$

The total voltage and current at the load are related by the load impedance, so, at the point of $x = 0$, we must have the following relationship

$$\underline{Z}_L = \frac{\underline{U}(0)}{\underline{I}(0)} = \frac{\underline{U}_0^+ + \underline{U}_0^-}{\underline{U}_0^+ - \underline{U}_0^-} Z_0. \tag{2.48}$$

Solving for U_0^- from the above equation, leads to

$$\underline{U}_0^- = \frac{\underline{Z}_L - Z_0}{\underline{Z}_L + Z_0} \underline{U}^+. \tag{2.49}$$

We define the voltage reflection coefficient $\underline{\Gamma}$ as the amplitude of the reflected voltage wave normalized to the amplitude of the incident voltage wave. From equation (2.49), we obtain

$$\underline{\Gamma} = \frac{\underline{U}_0^-}{\underline{U}_0^+} = \frac{\underline{Z}_L - Z_0}{\underline{Z}_L + Z_0}. \tag{2.50}$$

According to equations (2.46) and (2.47), the input impedance seen looking toward the load is

$$\begin{aligned} \underline{Z}_{in} &= \frac{\underline{U}(-l)}{\underline{I}(-l)} = \frac{\underline{U}_0^+ e^{-j\beta(-l)} + \underline{U}_0^- e^{j\beta(-l)}}{(1/Z_0)\cdot\left(\underline{U}_0^+ e^{-j\beta(-l)} - \underline{U}_0^- e^{j\beta(-l)}\right)} \\ &= \frac{\underline{U}_0^+ \left(e^{j\beta l} + \underline{\Gamma} e^{-j\beta l}\right)\cdot Z_0}{\underline{U}_0^+ \left(e^{j\beta l} - \underline{\Gamma} e^{-j\beta l}\right)} = \frac{1 + \underline{\Gamma} e^{-2j\beta l}}{1 - \underline{\Gamma} e^{-2j\beta l}} Z_0. \end{aligned} \tag{2.51}$$

Using equation (2.50), the above equation can be rewritten as

$$\underline{Z}_{in} = Z_0 \frac{\underline{Z}_L + jZ_0 \tan \beta l}{Z_0 + j\underline{Z}_L \tan \beta l}. \tag{2.52}$$

This equation denotes the most important impedance transforming property of a transmission line, which is frequently used for many circuit developments. In practice, there are several different types of transmission line, e.g. strip line, waveguide, coax, etc. In this work, microstrip line is applied.

2.5.3 Microstrip Geometry and Characteristic Parameters

A microstrip line is a transmission line consisting of a strip conductor and a ground plane separated by a dielectric medium. Fig. 2.33 (a) illustrates the microstrip geometry. The

dielectric material serves as a substrate and is sandwiched between the strip conductor and the ground plane. The substrate has a thickness of h while the conductor has a width of w and a thickness of t. The dielectric constant ε and the relative dielectric constant of the substrate ε_r are related by $\varepsilon = \varepsilon_r \cdot \varepsilon_0$, where $\varepsilon_0 = 8.855 \times 10^{12} F/m$ denotes the vacuum dielectric constant. The approximated electromagnetic field lines distributed in the microstrip lines are shown in Fig. 2.33 (b).

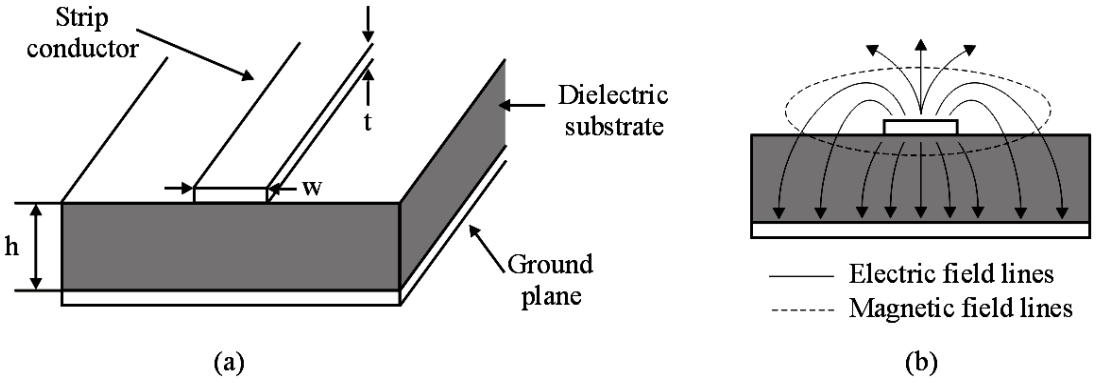


Fig. 2.33. The microstrip geometry (a) and the field configuration (b) [32].

Clearly, the electromagnetic field lines in the microstrip are not contained entirely in the substrate. Therefore, the propagating mode in the microstrip is not a pure transverse electromagnetic mode (TEM) but a quasi-TEM. Assuming a quasi-TEM mode of propagation in the microstrip line, the phase velocity is given by

$$v_p = \frac{c}{\sqrt{\varepsilon_{eff}}}. \tag{2.53}$$

The parameter c is the velocity of the electromagnetic wave in the free space (299792458 m/s) and ε_{eff} is the effective relative dielectric constant of the microstrip, which is dependent on the substrate thickness and the conductor width. Since some of the field lines are in the dielectric region and some are in air, the relation $1 < \varepsilon_{eff} < \varepsilon_r$ always exist. The design formulas for ε_{eff} is given by [32]

$$\varepsilon_{eff} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2} \cdot \frac{1}{\sqrt{1 + 12h/W}}, \tag{2.54}$$

and the characteristic impedance Z_0 of microstrip line is

$$Z_0 = \frac{60}{\sqrt{\varepsilon_{eff}}} \ln\left(\frac{8h}{W} + \frac{W}{4h}\right) \quad \text{for } W/h \le 1, \tag{2.55}$$

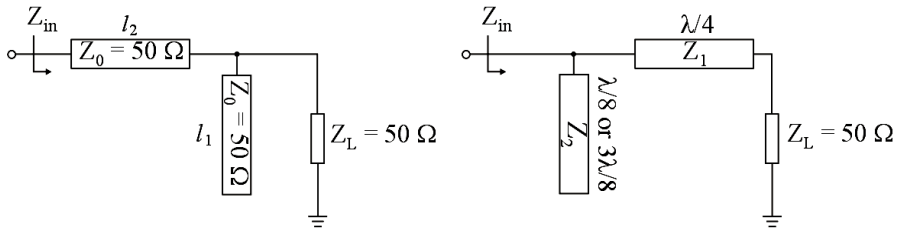
whereas

$$Z_0 = \frac{120\pi}{\sqrt{\varepsilon_{eff}} \cdot \left[W/h + 1.393 + 0.667 \ln\left(W/h + 1.444\right)\right]} \quad \text{for } W/h \ge 1. \tag{2.56}$$

These results are curve-fit approximations to rigorous quasi-static solutions [33].

2.5.4 Single-Stub Tuning

The impedance transforming properties of microstrip lines shown in equation (2.52) can be used in the design of matching networks. A microstrip line can be applied as a series transmission line, as an open-circuited stub, or as a short-circuited stub. There are mainly two variables to define a microstrip line, namely the length l and the characteristic impedance Z_0, if the substrate is determined. Therefore, there are typically two types of combinations of a series transmission line and a shunt stub as shown in Fig. 2.34. Both of these two combinations can transform a 50 Ω resistor into any value of the input impedance.



(a) with fixed characteristic impedance (b) with fixed length

Fig. 2.34. Two typical microstrip matching networks using a series transmission line and a shunt stub.

In Fig. 2.34 (a), a short- or an open-circuited stub with the length of l_1 is connected in parallel with the 50 Ohm load Z_L. It is followed by a series microstrip transmission line, which has a length of l_2. The characteristic impedance of these two sections of lines is fixed on 50 Ohm, only the lengths of them are varied.

The other possibility of the microstrip matching network is shown in Fig. 2.34 (b). This network uses a series quarter-wave line with characteristic impedance Z_1 followed by a short- or open-circuited shunt stub having a length of λ/8 or 3λ/8 and the characteristic impedance Z_2, where λ is the wave-length at the given frequency and substrate. Clearly, in this configuration, the lengths of the lines are determined, only their characteristic impedances are variable.

Design processes for the microstrip matching networks mentioned above are introduced in [34]. There are also matching networks with combinations of a series transmission and a series stub [32], however, the shunt stub is especially easy to fabricate in microstrip form.

2.5.5 Quarter-Wave and Multi-Section Matching Networks

The quarter-wave transmission line, indicated by l in Fig. 2.32 equalling to $\lambda / 4$, is a very useful impedance transformer. From (2.52), the characteristic impedance Z_0 of a quarter-wave transmission line can be calculated with

$$Z_0 = \sqrt{Z_{in} \cdot Z_L} \,. \tag{2.57}$$

Obviously, such a transformer can easily be realized. However, it can match a real load only to real input impedance. If the center frequency f_0 is given, the bandwidth of a quarter-wave transmission line can be calculated as follows [32]

$$\frac{\Delta f}{f_0} = 2 - \frac{4}{\pi}\cos^{-1}\left[\frac{\Gamma_t}{\sqrt{1-\Gamma_t^2}} \cdot \frac{2\sqrt{Z_0 Z_L}}{|Z_L - Z_0|}\right], \tag{2.58}$$

where Γ_t is the maximum value of voltage reflection coefficient that can be tolerated as shown in Fig. 2.35. The parameter θ in the figure is equal to $\pi f / 2 f_0$, while θ_t and π-θ_t are the lower and the upper edges of the passband, respectively; θ_t can be calculated using the following equation.

$$\cos\theta_t = \frac{\Gamma_t}{\sqrt{1-\Gamma_t^2}} \cdot \frac{2\sqrt{Z_0 Z_L}}{|Z_L - Z_0|} \tag{2.59}$$

The maximum voltage reflection coefficient Γ_m depends on the ratio of Z_L/Z_{in}.

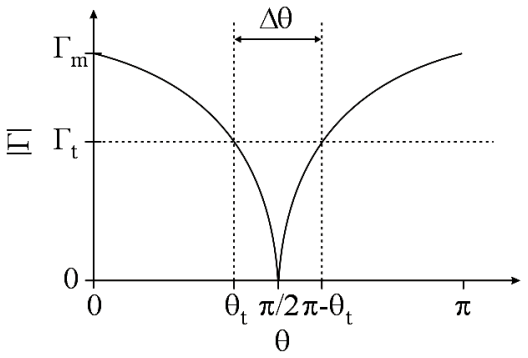


Fig. 2.35. The definition of the bandwidth of a microstrip line.

In practice, multi-section transformers are widely used. As shown in Fig. 2.36, N sections of quarter-wave transmission lines are connected in a row. They have different characteristic impedances $Z_1 \ldots Z_N$. Typically, all values of Z_n [$n \in (1 : N)$] increase or decrease monotonically across the transformer. Γ_n are the partial reflection coefficients. Just like multi-reactance discrete matching networks shown above, the main conception of such matching is retaining a lower quality factor Q, which implies a broader bandwidth for a given resonance frequency. Therefore, the bandwidth of the matching networks can theoretically be increased if more sections of transmission lines are used.

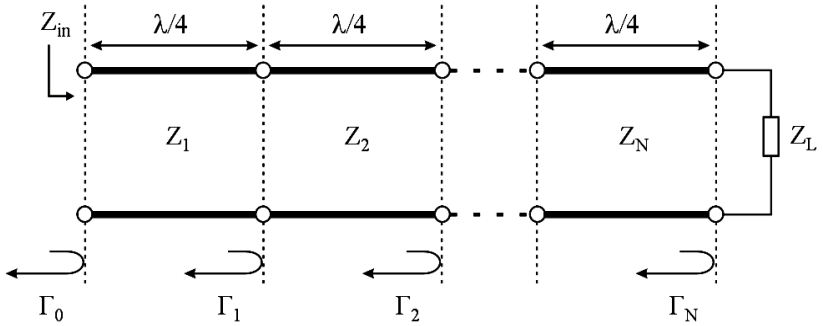


Fig. 2.36. Multi-section microstrip matching networks.

As the length of each section of transmission lines is fixed on λ/4 at a given frequency, the different characteristic impedances of all the sections must be determined. There are mainly

two types of multi-section transformer, namely, the binomial multi-section transformer and the Chebyshev multi-section transformer. For the binomial multi-section transformer, the characteristic impedance can be defined with the following equations [32]

$$\ln\frac{Z_1}{Z_{in}} \approx 2^{-N} C_1^N \ln\frac{Z_L}{Z_{in}} \quad \text{and} \quad \ln\frac{Z_{n+1}}{Z_n} \approx 2^{-N} C_n^N \ln\frac{Z_L}{Z_{in}} \text{ for } n \geq 1. \tag{2.60}$$

The bandwidth of the binomial N-section transformer can be given by

$$\frac{\Delta f}{f_0} = 2 - \frac{4}{\pi}\cos^{-1}\left[\frac{1}{2}\left(\frac{\Gamma_m}{|A|}\right)^{1/N}\right], \tag{2.61}$$

with

$$A = 2^{-N}\frac{Z_L - Z_{in}}{Z_L + Z_{in}}. \tag{2.62}$$

To define the characteristic impedance of a Chebyshev N-section transformer, θ_m should first be calculated with the following equation [32]

$$\sec\theta_m \approx \cosh\left[\frac{1}{N}\cosh^{-1}\left(\left|\frac{\ln(Z_L/Z_{in})}{2\Gamma_m}\right|\right)\right], \tag{2.63}$$

if Γ_m is fixed. Let the overall reflection coefficient Γ be equal to the Chebyshev polynomials $T_n(x)$, resulting in

$$\Gamma(\theta) = 2e^{-jN\theta}\left[\Gamma_0\cos N\theta + \Gamma_1\cos(N-2)\theta + \cdots + \Gamma_n\cos(N-2n)\theta + \cdots\right] = \Gamma_m e^{-jN\theta} T_N(\sec\theta_m\cos\theta). \tag{2.64}$$

Equating similar terms in $cosn\theta$, Γ_n can be obtained. The following work is to determine the characteristic impedance of the N section transmission lines with

$$\Gamma_0 \approx \frac{1}{2}\ln\frac{Z_1}{Z_{in}} \quad \text{and} \quad \Gamma_n \approx \frac{1}{2}\ln\frac{Z_{n+1}}{Z_n} \text{ for } n \geq 1. \tag{2.65}$$

The bandwidth of the Chebyshev N-section transformer can be given by

$$\frac{\Delta f}{f_0} = 2 - \frac{4\theta_m}{\pi}. \tag{2.66}$$

The binomial multi-section transformer provides the best flatness of the response near the design frequency, while the Chebyshev multi-section transformer has a poorer flatness but a larger bandwidth.

There are also other methods of determining the characteristic impedance of cascaded quarter-wavelength impedance transformer, typically by means of checking the available tables [35]. Another possibility is to use different software, one of which is LLsmith used in this work.

2.6 Biasing Network

The purpose of a DC bias is to select the proper quiescent point and hold the quiescent point constant over variation in transistor parameters and temperature. It determines the output power, amount of distortion, voltage headroom, efficiency, gain of the stage, noise of the stage, and class of operation (class A, AB, etc.). For some classes of power amplifiers, it is also expected that the biasing network is helpful to suppress the higher harmonics. The proper bias point is therefore a trade-off between all of these factors. Selecting the optimum bias point can sometimes be difficult, and it will vary depending on the requirements of the amplification stage.

The DC bias point can be set in a number of ways. The fundamental biasing network for power amplifier design is already shown in Fig. 2.1. Generally, a sufficiently high output power is required by a power amplifier design, the supply voltage U_{dd} can therefore not be fed to the drain of the MOSFET through a resistive load or any active load as mentioned in section 2.4.1, since such loads consume relatively large voltage headroom. Hence an inductor (L_2 in Fig. 2.1) is mostly used between U_{dd} and the drain of the MOSFET, which serves as RF choke and simultaneously feeds DC power to the power device. The inductance must be large enough so that the current flowing through the power device remains substantially constant. Such a biasing concept is primarily suitable for the power amplifier circuits used in cellular phones, where usually only a low power supply is available.

As introduced before, the different classes of the power amplifiers, especially of the current source mode power amplifiers are mainly determined by the gate-source-voltage of the MOSFET. The gate voltage U_{gg} is likewise fed to the gate through another large inductor L_1, as shown in Fig. 2.1. Another possibility to determine the gate voltage is using a resistive voltage divider, which allocates a part of U_{dd} to be used as U_{gg}, where the resistors must also be large enough to avoid the RF power loss.

Due to lots of merits, class AB power amplifiers are widely used. As shown in Fig. 2.8, the most critical harmonic to deal with in a class AB power amplifier is the second one. So the removal of the second harmonic voltage at the transistor output is of significant importance. This can be realized by using a low- or band-pass output matching network, or by using a suitable biasing network, e.g. the quarter-wave short-circuited stub. The quarter-wave short-circuited, high impedance line provides the DC path for the supply voltage and presents an open circuit to the RF signals at the fundamental frequency. Simultaneously, the second harmonic is shorted by it. The narrowest practical line, which means the line having the largest characteristic impedance, should be used for the $\lambda/4$ short-circuited stub to avoid unwanted RF coupling. On the other hand, the maximum allowable current must also be considered, since it is desired, that a large DC current flows through this line to the drain of the transistor. The bypass capacitors C_b shown in Fig. 2.1 are used as RF short circuits. Several parallel bypass capacitors between 1 pF and 100 nF can be applied to couple RF signals of different frequencies to the ground.

(a) series resonator (b) parallel resonator

Fig. 2.37. The resonators used in the bias network.

Another method to realize the harmonic removal by the biasing is the use of the series or parallel LC-resonant circuits [32] between U_{dd} and the drain of the MOSFET as shown in Fig. 2.37. Theoretically, such biasing networks also consume no voltage headroom. The resonant frequency f_{res} for both the series and the parallel resonators can be calculated with the following equation

$$f_{res} = \frac{1}{2\pi\sqrt{LC}}. \tag{2.67}$$

The quality factors Q of them are separately given with

$$Q = \frac{1}{R}\cdot\sqrt{\frac{L}{C}} \qquad \text{for series resonance,} \tag{2.68}$$

and

$$Q = R\cdot\sqrt{\frac{C}{L}} \qquad \text{for parallel resonance.} \tag{2.69}$$

In practice, f_{res} is just defined at the operating frequency of the amplifier circuit. Therefore, the resonant circuit provides an open circuit to the RF signal at the fundamental frequency and simultaneously shorts all the other harmonics of the output signal. Additional LC-resonant circuits tuned to the higher harmonics can also be used in the biasing network to form a determined waveform of the output voltage. Such biasing concepts are often used, e.g. in the class F power amplifiers as shown below.

To meet the high requirement of linearity, power amplifiers of CDMA transmitters normally operate 6 dB backed off from the saturation. It is therefore an important issue to increase the PAE in lower output power levels. The Doherty amplifier was attempted to improve the efficiency using the technique of load-line modulation of a carrier amplifier through a peaking amplifier attached to a quarter-wave transmission line [36] [37]. Another possibility developed in recent years is to use a so-called adaptive biasing network [38] [39], which supplies a low quiescent current at the low output power level for high efficiency, and a higher quiescent current at high output power level for high linearity.

2.7 Design Parameters of the Power Amplifier

This chapter introduces some basic concepts and parameters in the analysis and design of the microwave transistor power amplifiers. The most important considerations in a microwave power amplifier design are power gain, stability, bandwidth, 1-dB compression point, intermodulation distortion (IMD), adjacent channel power ratio (ACPR) and power added efficiency (PAE).

2.7.1 Power Gain

There are several different power gain equations appearing in the literature and are used in the design of microwave amplifiers [34], namely the transducer power gain G_T, the operating power gain G_P and the available power gain G_A. They are defined as follows

$$G_T = \frac{P_L}{P_{AVS}} = \frac{power\ deliverd\ to\ the\ load}{power\ available\ \ from\ the\ source} \tag{2.70}$$

$$G_P = \frac{P_L}{P_{IN}} = \frac{power\ deliverd\ to\ the\ load}{power\ input\ to\ the\ network} \tag{2.71}$$

$$G_A = \frac{P_{AVN}}{P_{AVS}} = \frac{power\ available\ \ from\ the\ network}{power\ available\ \ from\ the\ source}. \tag{2.72}$$

These definitions are all the same, if the input and output of the power amplifier circuit are both conjugately matched to the source impedance and to the load impedance respectively. In this case, the maximum power gain can be obtained.

2.7.2 Stability

Stability is an extremely important issue for the power amplifier design. Oscillation can occur if either the input or the output impedance has a negative real part. This is a serious problem, because the power device used in the power amplifier circuit can be destroyed under such a condition.

Defining Γ as the reflection coefficient, the case of negative real part of the input and output impedance can be denoted by $|\Gamma_{in}| > 1$ and $|\Gamma_{out}| > 1$. According to Fig. 2.38, the input reflection coefficient Γ_{in} and the output reflection coefficient Γ_{out} of a two port network can be described with the following equations [32].

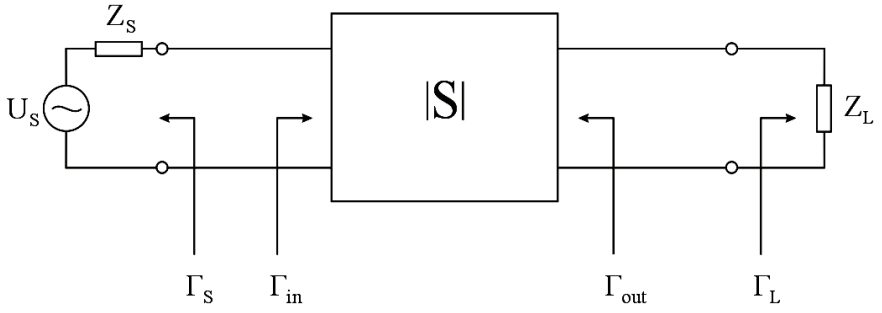


Fig. 2.38. A two-port network with reflection coefficient.

$$\Gamma_L = \frac{Z_L - Z_0}{Z_L + Z_0} \tag{2.73}$$

$$\Gamma_S = \frac{Z_S - Z_0}{Z_S + Z_0} \tag{2.74}$$

$$\Gamma_{in} = \frac{Z_{in} - Z_0}{Z_{in} + Z_0} = S_{11} + \frac{S_{12}S_{21}\Gamma_L}{1 - S_{22}\Gamma_L} \tag{2.75}$$

$$\Gamma_{out} = \frac{Z_{out} - Z_0}{Z_{out} + Z_0} = S_{22} + \frac{S_{12}S_{21}\Gamma_S}{1 - S_{11}\Gamma_S}, \tag{2.76}$$

where the S-parameters are those of the power amplifier circuit. Because Γ_{in} and Γ_{out} depend on the source and load matching networks, the stability of the amplifier depends on Γ_S and Γ_L. Two types of stability, namely unconditional and conditional stability can be defined as follows [32]:

1. *Unconditional stability*: The network is unconditionally stable if $|\Gamma_{in}| < 1$ and $|\Gamma_{out}| < 1$ are fulfilled for all passive source and load impedances.
2. *Conditional stability*: The network is conditionally stable if $|\Gamma_{in}| < 1$ and $|\Gamma_{out}| < 1$ are fulfilled only for a certain range of passive source and load impedances.

Since the S-parameters of the circuits are dependent on the frequency, it is possible for an amplifier to be stable at some frequencies but unstable at other frequencies.

Applying equation (2.75), the condition of $|\Gamma_{in}| = 1$ can be rewritten as follows

$$\left| \Gamma_L - \frac{(S_{22} - \Delta \cdot S_{11}^*)^*}{|S_{22}|^2 - |\Delta|^2} \right| = \left| \frac{S_{12} \cdot S_{21}}{|S_{22}|^2 - |\Delta|^2} \right|, \tag{2.77}$$

with $\Delta = S_{11}S_{22} - S_{12}S_{21}$. The signs of * indicate the conjugate complex number. In the complex Γ plane, equation (2.77) can be regarded as a circle with the center of C_o and a radius of R_o, which are described by

$$C_o = \frac{(S_{22} - \Delta \cdot S_{11}^*)^*}{|S_{22}|^2 - |\Delta|^2} \tag{2.78}$$

$$R_o = \left| \frac{S_{12}S_{21}}{|S_{22}|^2 - |\Delta|^2} \right|. \tag{2.79}$$

This circle is named output stability circle. Similarly, the input stability circle can be given with its center of C_i and the radius of R_i described by

$$C_i = \frac{(S_{11} - \Delta \cdot S_{22}^*)^*}{|S_{11}|^2 - |\Delta|^2} \tag{2.80}$$

$$R_i = \left| \frac{S_{12}S_{21}}{|S_{11}|^2 - |\Delta|^2} \right|. \tag{2.81}$$

Given the S-parameters of the device, we can plot the input and output stability circles in the Smith chart to denote where $|\Gamma_{in}| = 1$ and $|\Gamma_{out}| = 1$.

Mostly, the stability circle divides the Smith chart into two parts denoting that the circuit is only conditionally stable. One of the two parts is inside the circle and the other is outside the circle. The next question is then, which part belongs to the stable region, inside or outside. The following two rules are available:

1. If $|S_{11}| < 1$, the center of the Smith chart for load is in the stable region. Then, if the stability circle includes the center point, the stable region is inside this circle. Conversely, if the stability circle doesn't include the center, the stable region is outside this circle.
2. If $|S_{11}| > 1$, the center of the Smith chart for load is in the instable region. Then, if the stability circle includes the center point, the stable region is outside this circle. Conversely, if the stability circle doesn't include the center point, the stable region is inside this circle.

The same relationship is also valid for $|S_{22}|$ and the Smith chart for source. To research whether the circuit is conditional stable in the entire frequency band of interest, stability circles should be plotted for all the frequencies, the rules described above are also repeated for each frequency to see, where the circuit is stable and where not.

Similarly, the prerequisites for unconditional stability are: 1) all the stability circles for all the frequencies are outside the Smith charts, not only for source but also for load, and 2) $|S_{11}| < 1$, as well as $|S_{22}| < 1$ are fulfilled for all the frequencies.

Conveniently, there are two mathematical methods to test whether a circuit is unconditionally stable. The first one is the so-called $K - \Delta$ test, where it can be shown that a circuit is unconditionally stable if *Rollet's condition*, defined as

$$K = \frac{1 - |S_{11}|^2 - |S_{22}|^2 + |\Delta|^2}{2|S_{12}S_{21}|} > 1, \tag{2.82}$$

along with the auxiliary condition that

$$|\Delta| = |S_{11}S_{22} - S_{12}S_{21}| < 1 \tag{2.83}$$

are simultaneously satisfied. The other mathematical method applies the unconditional stability factor μ. The circuit is unconditionally stable, if

$$\mu = \frac{1 - |S_{11}|^2}{|S_{22} - S_{11}^* \cdot \Delta| + |S_{21}S_{12}|} > 1 \tag{2.84}$$

is obtained for the whole frequency band [40], where Δ is given in equation (2.83). Here, only a single parameter is used compared to the $K - \Delta$ test. Therefore, the unconditional stability factor μ is commonly used.

2.7.3 Gain Compression and 1-dB Compression Point

Transistors and diodes are nonlinear devices, and the nonlinearity can be used in practice for functions like amplification, detection and frequency conversion [41] etc. However, nonlinear device characteristics can also lead to undesired responses such as gain compression

and the generation of spurious frequency components. These effects may produce increased losses, signal distortion and possible interference with other channels.

Assuming u_i and u_o are the input and the output voltage respectively, the output response for a general nonlinear two-port network can be modeled as a Taylor series [42] shown below

$$u_o = a_0 + a_1 u_i + a_2 u_i^2 + a_3 u_i^3 + \cdots , \tag{2.85}$$

where a_0, a_1, a_2 ··· are the Taylor coefficients. If a sinusoidal signal $u_i = U_0 \cdot \cos\omega_0 t$ is applied to the input of the nonlinear network, the output voltage is then

$$\begin{aligned} u_0 &= a_0 + a_1 U_0 \cos\omega_0 t + a_2 U_0^2 \cos^2 \omega_0 t + a_3 U_0^3 \cos^3 \omega_0 t \\ &= (a_0 + \frac{1}{2} a_2 U_0^2) + (a_1 U_0 + \frac{3}{4} a_3 U_0^3)\cos\omega_0 t + \frac{1}{2} a_2 U_0^2 \cos 2\omega_0 t + \frac{1}{4} a_3 U_0^3 \cos 3\omega_0 t + \cdots . \end{aligned} \tag{2.86}$$

This result leads to the voltage gain at the fundamental frequency ω_0 of $a_1 + 3/4 \cdot a_3 \cdot U_0^2$. In most practical amplifiers a_3 is typically negative, so that the gain of the amplifier tends to decrease for large values of U_0 or of P_{in}. This effect is called gain compression, or saturation, which can be observed in the typical power transfer characteristic of a power amplifier as shown in Fig. 2.39. Physically, the gain compression occurs due to the fact that the instantaneous output voltage of an amplifier is limited by the power supply voltage used to bias the power device. To quantify the linear operating range of the amplifier, 1 dB compression point is defined as the power level for which the output power has decreased by 1 dB from the ideal characteristic. This point is normally denoted by P_{1dB} and stated in terms of either input power or output power.

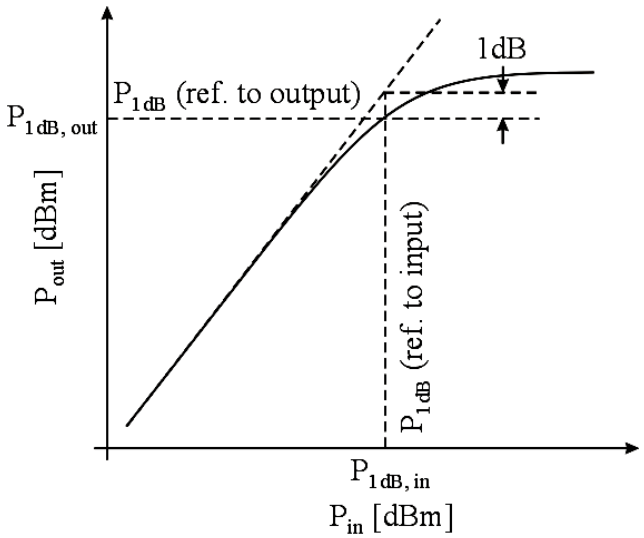


Fig. 2.39. The conventional power transfer function and the 1 dB compression point.

2.7.4 Intermodulation Distortion

If a single-tone signal is applied to the input of the power amplifier, harmonics can occur leading to a signal distortion if those components are in the passband of the amplifier. But usually, these harmonics lie outside the passband of the amplifier, and so do not interfere with the desired signal at the frequency of interest.

It is much more critical to consider a two-tone input voltage. Even in early radio systems, the interfering effects of an active neighboring channel in mildly non-linear communications systems were well known. In those simpler modulation schemes, the use of two sinusoidal signals to represent two active channels was considered adequate. For instance, if the input voltage is $u_i = U_0(\cos\omega_1 t + \cos\omega_2 t)$, consisting of two closely spaced frequencies ω_1 and ω_2, from equation (2.85), the output voltage u_0 can be obtained as follows:

$$\begin{aligned} u_0 &= a_0 + a_1 U_0^2 \\ &+ a_1 U_0(\cos\omega_1 t + \cos\omega_2 t) + \frac{9}{4} a_3 U_0^3(\cos\omega_1 t + \cos\omega_2 t) \\ &+ \frac{1}{2} a_2 U_0^2(\cos 2\omega_1 t + \cos 2\omega_2 t) + a_2 U_0^2\left[\cos(\omega_1 - \omega_2)t + \cos(\omega_1 + \omega_2)t\right] \\ &+ \frac{1}{4} a_3 U_0^3(\cos 3\omega_1 t + \cos 3\omega_2 t) + \frac{3}{4} a_3 U_0^3\left[\cos(2\omega_1 - \omega_2) + \cos(2\omega_2 - \omega_1) + \cos(2\omega_1 + \omega_2) + \cos(2\omega_2 + \omega_1)\right] \\ &+ \cdots \end{aligned} \tag{2.87}$$

It can be seen that the output spectrum consists of harmonics in form of $m\omega_1 \pm n\omega_2$, where m and n are integers. These combinations of the two input frequencies are called intermodulation products, and the order of a given product is defined as |m| + |n|. The output spectrum is shown in Fig. 2.40, assuming $\omega_1 < \omega_2$. Obviously, if ω_1, ω_2 are close, all the second-order products will be far from ω_1 or ω_2, and can easily be filtered from the fundamental signal. The critical terms come from the third-order products of equation (2.87). They are the cubic terms of $3\omega_1$, $3\omega_2$, $2\omega_1 + \omega_2$, $2\omega_2 + \omega_1$, $2\omega_1 - \omega_2$ and $2\omega_2 - \omega_1$, where the last two difference terms are located close to the two fundamental signals of ω_1 and ω_2, so that they can not easily be filtered by the passband filter at the output of the power amplifier. Such intermodulation products will generate distortion of the output signal. Therefore, the third-order intermodulation distortion (IMD3) is a very important specification of a power amplifier design, which should be as low as possible. Usually, the power difference between the signals at the frequencies $2\omega_1 - \omega_2$ and ω_1 is called IMD3 lower, while that between the signals at the frequencies $2\omega_2 - \omega_1$ and ω_2 is called IMD3 upper.

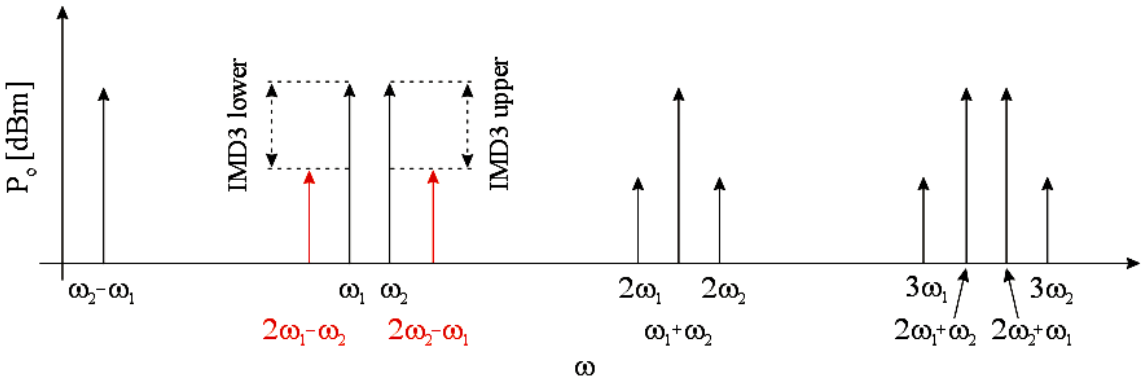


Fig. 2.40. The output spectrum of second and third-order two-tone intermodulation products.

Naturally, the fifth-order products of $3\omega_1 - 2\omega_2$ and $3\omega_2 - 2\omega_1$, which also exist in the output signal but are not shown in equation (2.87) and Fig. 2.41, are also located near to the fundamental signals. However, they have generally much lower magnitude.

2.7.5 Intercept Point

According to equation (2.87), the voltage associated with the third-order products U_0^3 is increased as the magnitude of the input voltage U_0 rises. Since the power is proportional to the square of the voltage, the output power of third-order products must increase as the cube of the input power. So for small input powers the third-order intermodulation products must be very small, but will increase quickly as input power increases.

Fig. 2.41 presents this effect graphically by plotting the output power for the first- and third-order products versus input power on logarithmic scales. The output power of the first order product has a slope of unity, whereas the line describing the response of the third-order products has a slope of 3. Both the first- and third-order responses will exhibit compression at high input powers. The extension of their idealized responses which are shown with dotted line in Fig. 2.41 will intersect, typically at a point above the 1 dB compression point. This point is named the third-order intercept point, denoted by IP3, especially by iIP3 referred to the input power and by oIP3 referred to the output power. In practice, many circuit components follow the approximate rule that oIP3 is 12 to 15 dB higher than $P_{1dB,out}$ [42].

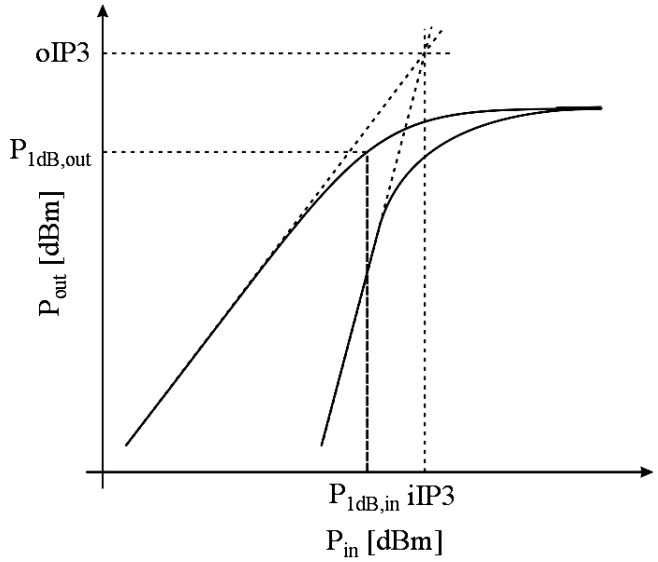


Fig. 2.41. The power transfer function and the third-order intercept point.

2.7.6 ACPR

As introduced in section 2.7.4, intermodulation distortion describes the interfering effects of an active neighboring channel for the simpler modulation schemes, where two sinusoidal signals represent two active channels. As the modulation becomes more complex, it becomes less obvious that the sinusoidal representation will adequately simulate the problem of interference from the neighbor channel. For many of the current and future transmission standards, e.g. IS-95 code division multiple access (CDMA), wideband code division multiple access (WCDMA) etc., ACPR becomes the most important test parameter for characterizing the distortion of subsystems and the possibility that a given system may cause interference with a neighboring radio. ACPR is the logical extension of the distortion measurement where the two sinusoidal signals used in the IMD measurement are replaced by a given modulated signal, which has a given bandwidth of Δf. The bandwidth and the location are functions of the standards being employed.

Usually, ACPR is defined as the ratio of the average power in the adjacent frequency channel to the average power in the transmitted frequency channel. Fig. 2.42 provides the graphic presentation of ACPR, where the output power is demonstrated in dependence of the frequency *f*. In this figure, f_c indicates the center frequency. Similar to IMD, we define the ratio of the power in the left bandwidth to the power in the bandwidth of the main signal as ACPR lower and the ratio of the right side as ACPR upper. ACPR measurements on the same transmitter can provide different results depending on the statistics of the transmitted signal. Different peak-to-average ratio (PAR) values of the input signal have a different impact on the non-linear components of the transmitter, e.g. on the power amplifier, and therefore a different impact on the ACPR as well. Higher PAR cause more interference in the adjacent channel, and hence lower ACPR. Sometimes, ACPR is also named the adjacent channel leakage ratio (ACLR).

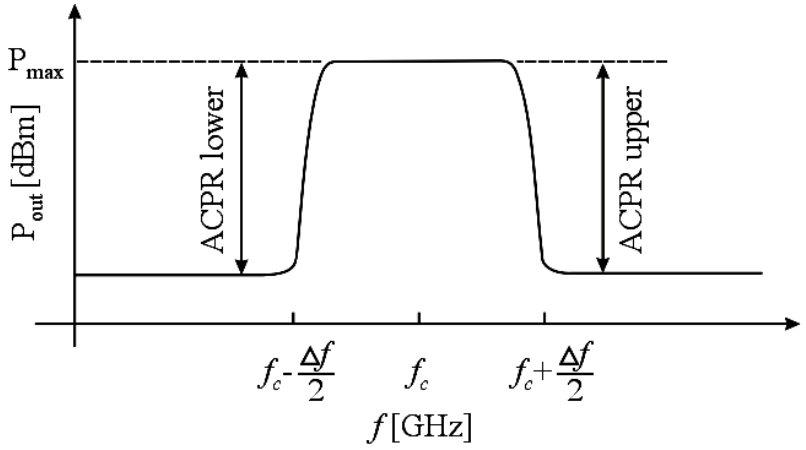


Fig. 2.42. The output spectrum and ACPR

2.7.7 Power Added Efficiency

High efficiency is an important power amplifier characteristic which allows for smaller power sources and reduced cooling requirements. The original efficiency is defined as output efficiency which can be calculated as

$$\eta = \frac{RF\ \ output\ \ power}{DC\ \ power}. \tag{2.88}$$

This definition is also called drain efficiency, if FET transistors are used as power devices. The other definition of the efficiency is the power added efficiency (PAE), which is in practice mostly used and described as follows

$$PAE = \frac{RF\ \ output\ \ power - RF\ \ input\ \ power}{DC\ \ power} = \frac{P_{out} - P_{in}}{P_{dc}}, \tag{2.89}$$

where P_{out} is the RF output power and P_{in} is the RF input power. This definition includes the consideration of the RF driving power, which can be quite substantial in an RF power amplifier. In general, if the RF power gain is less than 10 dB, the drive power requirements will start to take a serious bite of the PAE of the power amplifier.

3. Design of the CMOS Driver and Power Amplifiers

Using standard deep-submicron CMOS technology, the driving amplifiers and the power amplifiers adopted in cellular phones are designed. The design processes are described in this chapter. The simplified transceiver system applied in a cellular phone is shown in Fig. 3.1. To ensure that the transmitter has sufficiently high output power, a power amplifier (PA) is applied in the end-stage of the transmitter. Furthermore, in order to provide sufficiently high power gain, several driving stages are necessary. On the other hand, a mobile transmitter requires a large dynamic gain control range; hence programmable-gain amplifiers (PGA) or variable-gain amplifiers (VGA) must be adopted. If a PGA or VGA is also designed at the same operating frequency as that of the power amplifier, it can simultaneously be used as driver or predriver of the power amplifier.

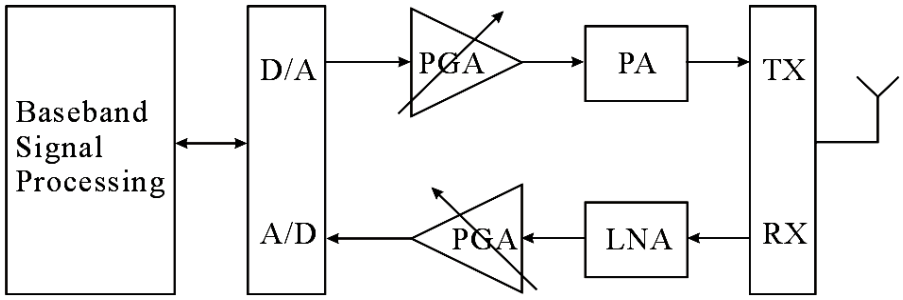


Fig. 3.1. Simplified transceiver system in a cellular phone.

Today, CMOS is the technology of choice for a higher integration level and lower cost because it is capable of implementing a significant amount of digital signal processing and because the vast majority of today's integrated circuits are implemented in this technology. However, it also shows drawbacks for the design of a logarithmic PGA due to its square-law transfer characteristic in the saturation region and the design of a power amplifier with large output power due to the low breakdown voltage. Therefore, novel circuit concepts not only for PGAs but also for the PAs are introduced in this chapter to overcome these problems.

The design of a radio frequency logarithmic PGA using a 0.12-µm CMOS technology is first described in section 3.1. In section 3.2, a class A HiVP power amplifier for GSM mobile communication system is presented, which is also developed in a 0.12-µm CMOS process.

3.1 A Logarithmic Programmable-Gain Amplifier

Using a 0.12-µm CMOS technology, two radio frequency PGAs are designed in this section, one is designed at 1 GHz and the other is designed at 2 GHz. They should control the overall gain of the transmitter and simultaneously be adopted as a driving stage of the power amplifier. Following specifications should be fulfilled by these two PGAs:

- operating frequency: 1 GHz or 2 GHz
- dynamic gain control range: > 50 dB
- maximum gain: at least 8 dB
- logarithmic gain variation
- gain control step: 3 dB
- maximum output power: 10 dBm
- adaptive DC power consumption

The two logarithmic PGAs should be realized in a standard 0.12-µm CMOS technology. However, in case that a MOS transistor works in the saturation region, a logarithmic gain variation implemented in a CMOS technology can only be fulfilled, if an approximate rational function instead of an exponential function is used [43]. However, the maximum gain control range obtained in this manner is limited within 30 dB. For an even larger gain control range, parasitic bipolar transistors can be used for accurate exponential transformation [44]. It has also been attempted to develop a logarithmic PGA for a larger dynamic gain variation simply in a CMOS technology, several cascaded PGA stages have been used [45]. However, such a design concept has an essential drawback. Usually, the gain varies not only in the positive range but also in the negative range. In case that a larger negative gain should be realised, more series amplifier stages should be applied, resulting in even larger power consumption. A more reasonable behaviour should be an adaptive power consumption, which means the DC power consumption should be reduced when the gain becomes smaller, or in other word, when the attenuation becomes larger. In this work, a novel parallel structure to design a logarithmic PGA is presented.

In subsection 3.1.1, an overview of the conventional PGA and VGA design is introduced. The new circuit configuration to design a logarithmic PGA is then described in subsection 3.1.2. Finally, the simulation results, the layout of the circuits as well as the experimental results of the proposed PGA circuit are presented. A conclusion is given in subsection 3.1.4.

3.1.1 Introduction of the PGA and VGA

A gain variation block has become an indispensable function block for many mobile communication systems in order to maximize the overall system dynamic range. In the CDMA systems, for instance, the mobile transmitter is required to provide at least 80 dB dynamic gain control range [46] - [48]. Up to now, a large gain control range is realized in two stages, namely, in an intermediate frequency (IF) stage which is typically at 70 MHz, and in a radio frequency (RF) stage designed at the operating frequency of the transceiver.

Mostly, a PGA has discrete gain steps as shown in Fig. 3.2 (a), since the gain of a PGA is controlled by a discrete digital signal. Oppositely, the gain of a VGA changes continuously with the variation of the biasing voltage or current as shown in fig. 3.2 (b). A discrete gain variation causing indeed phase discontinuity; however, the amount of it is well below what is currently specified in the 3GPP UMTS specifications. Therefore, it is not critical to use a PGA designed at 2 GHz in the UMTS transmitter applications.

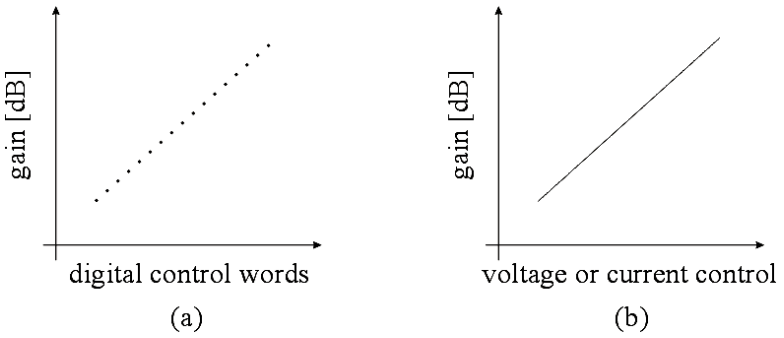


Fig. 3.2. Different operational functions of the PGA and VGA.
(a) discrete gain variation; (b) continuous gain variation

There are four different methods reported so far to realize the gain variation. They are shortly introduced below.

(1) Current steering technique

The gain variation can be realized by using the current steering technique [49], [50]. Fig. 3.3 shows the schematic of a conventional VGA using such technique.

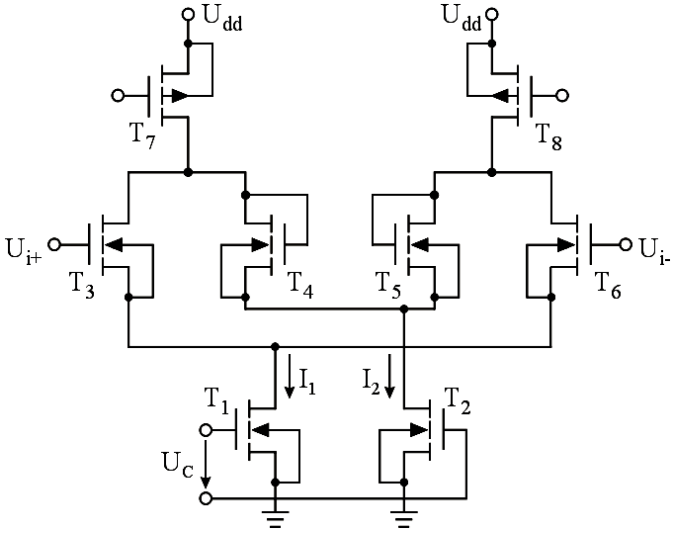


Fig. 3.3. The schematic of the VGA using current steering technique.

The amplifier is composed of an input source-coupled pair (T_3, T_6) and the diode connected loads (T_4, T_5). I_1 is the current flowing through the input pair, and I_2 is the current flowing through diode-connected loads. In case that T_3 – T_6 have the same dimension, the differential gain can be approximately expressed as [50]

$$G \approx \sqrt{\frac{I_1}{I_2}} \cdot \frac{1}{\frac{1}{r_{ds-p}\sqrt{\mu_n C_{ox}\left(W_n / L_n\right) \cdot I_2}} + 1}, \tag{3.1}$$

where $r_{ds\text{-}p}$ is the drain-source resistance of PMOS transistors (T_7, T_8). In equation (3.1), μ_n is the carrier mobility; C_{ox} is the gate oxide capacitance per unit area; W_n and L_n are the channel width and the channel length of T_3 – T_6, respectively. Obviously, the gain can be controlled by the current ratio I_1 / I_2.

(2) Variable transconductance topology

As shown in equation (2.28), the gain of a single-ended FET amplifier is proportional to the load *R* and the transconductance of the MOS transistor g_m. Therefore, it is an effective method to vary the gain by changing the transconductance of the transistors [51], [52]. There are several different methods to control g_m. An example is shown in Fig. 3.4. Two transistors T_1 and T_2 are connected in series. They share the same bias current, the total power consumption of the current-reused amplifier is minimised. To achieve a higher gain, both T_1 and T_2 are in common-source configuration. The gate width and the bias current of the transistors are chosen to maximise g_m while maintaining low power consumption.

The gain-controlled mechanism is achieved by adjusting the bias voltage U_C at the gate of T_2. As the gate voltage of T_2 decreases, the drain voltage of T_1 is suppressed, moving the transistor bias out of its high gain region.

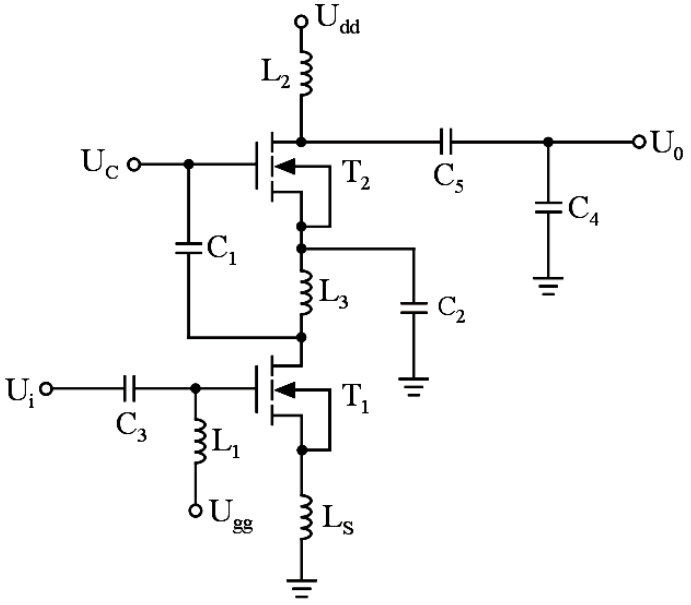


Fig. 3.4. The schematic of the VGA using variable transconductance topology.

(3) Attenuator topology

A gain variation can be obtained using additional attenuator at the input of the amplifier stages [53], [54]. The condition is that the amplifier stages should first be developed to possess the maximum gain. The Π or T shape attenuation circuits [53] can be applied, which are depicted in Fig. 3.5 (a) and (b). In these circuits, low field resistance of the zero-biased FETs are used as variable resistors, whose values are controlled by the gate voltages. The equivalent circuits of such attenuators are presented in Fig. 3.5(c) and (d).

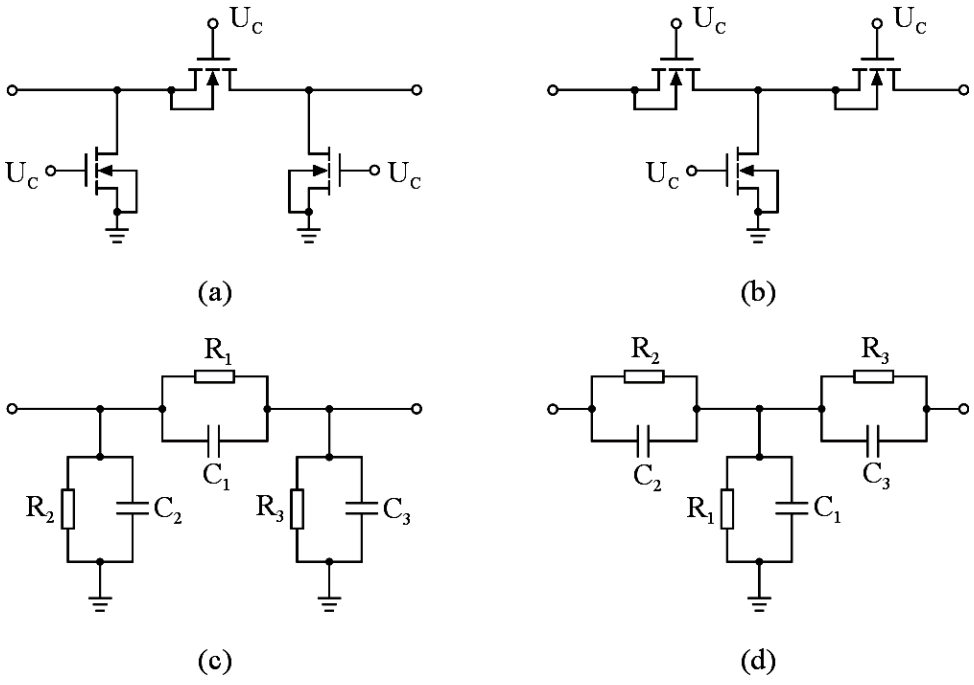


Fig. 3.5. Basic circuit topologies of variable attenuators (a), (b) and their equivalent circuits (c), (d).

(4) Feed back topology

A gain variation can also be achieved by a feedback topology [55], [56]. For instance, a VGA with resistor-network feedback is shown in Fig. 3.6.

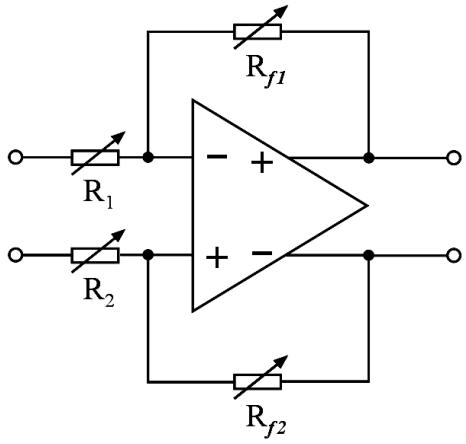


Fig. 3.6. A VGA with resistor-network feedback.

Obviously, the values of the resistors R_1, R_2, R_{f1} and R_{f2} are adjustable in this amplifier circuit. The differential inverting configuration indicates that its voltage gain can be varied by changing the ratios of R_{f1} / R_1 and R_{f2} / R_2. Thus signal amplification and attenuation can easily be realized. High linearity can be achieved if the loop gain is very large and the resistor network is linear [55].

Another possible feedback configuration is shown in Fig. 3.7. Clearly, R_d is the load of the small signal amplifier circuit. A variable resistor R_s is used at the source of the MOSFET, which indicates a feedback configuration. The gain variation can be achieved by varying this source resistance.

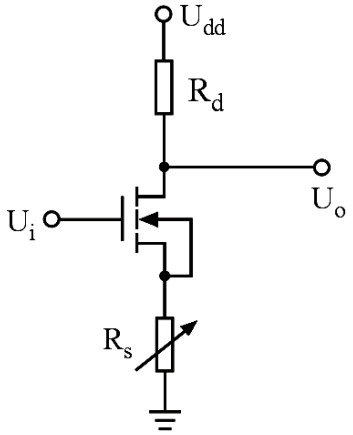


Fig. 3.7. The source feedback transistor stage.

The small signal gain of the circuit shown above can be given as follows [56]

$$G = -\frac{g_m R_d}{1 + g_{ds}(R_s + R_d) + g_m R_s}, \tag{3.2}$$

where g_{ds} is the drain-source-admittance and g_m is the transconductance of the transistor. In case that g_{ds} is very small, equation (3.2) can be approximated as:

$$G = -\frac{g_m R_d}{1 + g_m R_s} \approx -\frac{g_m R_d}{g_m R_s} = -\frac{R_d}{R_s}. \tag{3.3}$$

Therefore, by changing the source resistance, the gain or attenuation of this transistor stage can be controlled.

3.1.2 Design Concept of the Radio Frequency PGA

In order to achieve a large logarithmic gain control range realized in a CMOS technology, a new circuit configuration for the PGA design [57] is introduced in this work. In principal it belongs to the variable transconductance topology. Several amplifier cells containing transistor devices with different gate widths are connected in parallel. The variation of the transconductance is therefore realized by the turning on or turning off the different amplifier cells.

As an example, the block diagram of a PGA circuit with four amplifier cells is illustrated in Fig. 3.8. These parallel connected amplifier cells have the common load and they are digitally controlled by a 2 to 4 demultiplexer. B_1 and B_2 are the two inputs of the demultiplexer. At any given time, only one output of the demultiplexer has a high voltage level; hence only one amplifier cell of the four is turned on. In Fig. 3.8, the voltages U_{DD} = 2.5 V and U_{dd} = 1.5 V are the power supplies used for the amplifier cells and for the demultiplexer, respectively. In this work, 18 amplifier cells must be connected in parallel, in order to realize the specifications of a gain control range larger than 50 dB and a gain control step of 3 dB. These amplifier cells must be controlled by a 5 to 18 demultiplexer. The output digital control words (DCW) of the demultiplexer are defined as 1 to 18 implying the increasing gain.

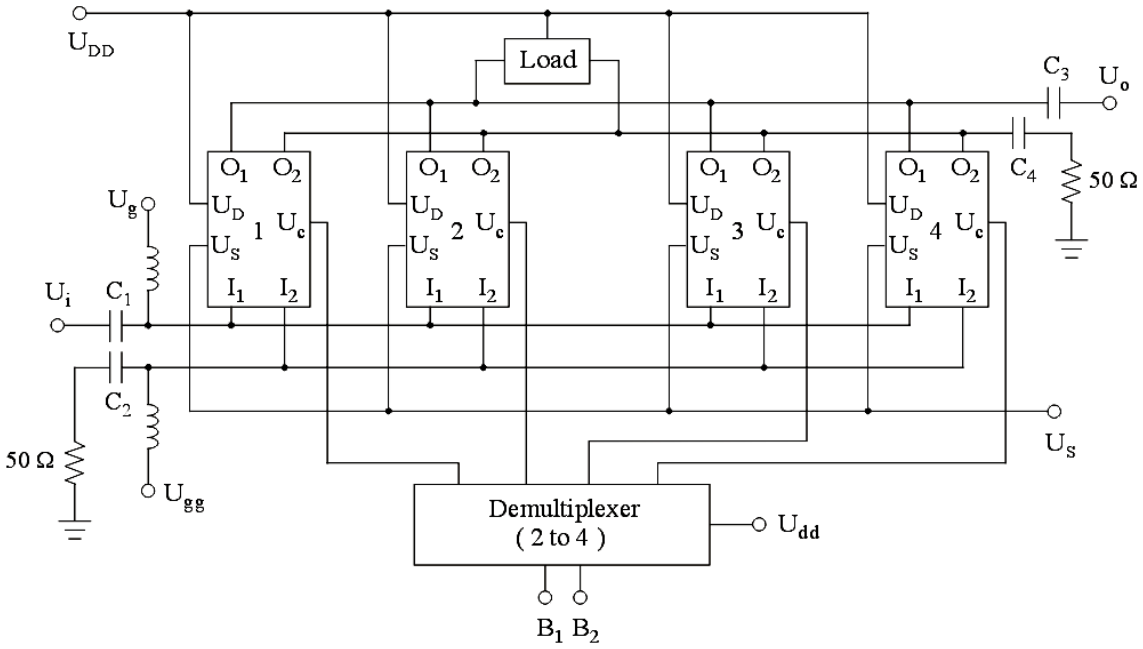


Fig. 3.8. The block diagram of the proposed PGA.

Fig. 3.9 presents the schematic of a single amplifier cell. To obtain the higher immunity to noise and crosstalk coming from substrate and power supply, differential configuration is proposed for all the amplifier cells. Cascode structure is used in the differential pairs to minimise the Miller effect. Guard rings are employed around the devices to reduce the substrate noise. The transistors T_1 – T_4 have all the same dimensions and the gate width of them is determined according to the gain expected from each cell. The transistor T_C is used as the control-element. It works as the current source of the differential pair and the gate of it is connected to one of the eighteen outputs of the 5 to 18 demultiplexer. Therefore, this cell is activated as soon as the connected demultiplexer output has a high voltage level. Moreover, single-ended input and output are expected from this PGA, therefore, the input I_2 of each amplifier cell is connected with a 50 Ohm (the reference impedance) resistor through the input RF coupling C_2, whereas the output O_2 of each amplifier cell is also connected with a 50 Ohm resistor through the output RF coupling C_4 as shown in Fig. 3.8.

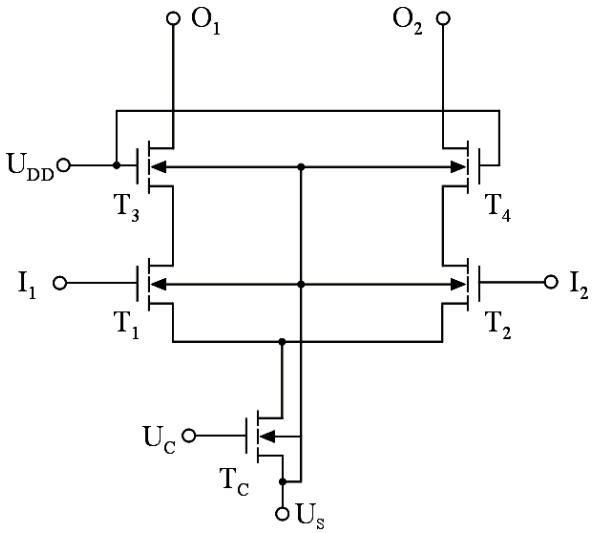


Fig. 3.9. The schematic of an amplifier cell.

The common load of the amplifier cells are attempted in two ways, they are the active load comprising a current mirror and the passive load comprising an LC tank as introduced in Fig. 2.37. The active load is in many cases preferred, since it requires less chip area and it needs no inductor which causes coupling with the other inductors used in the other circuits of the system. However, the current mirror consumes quite large voltage headroom. For instance, a pair of p-channel transistors having a gate width of 200 µm is first used as the load of the PGA circuit; simulation shows that a DC voltage of 0.6 V is consumed on it. Therefore, in case that a higher output power is demanded, passive load with LC tank is mostly preferred. In the given components library, there is a fixed inductance of 4.7 nH. Using equation (2.67), two different capacitances can be determined for the two different resonance frequencies. They are 5.4 pF for the frequency of 1 GHz and 1.35 pF for the frequency of 2 GHz, respectively. Simulations show that such a LC tank consumes voltage headroom of only 0.02 V.

The transistors of amplifier cell 1 are first dimensioned for the largest gain of 9 dB. Gate widths of 250 µm and of 180 µm are chosen for the transistors T_1 – T_4 and for T_C in cell 1, respectively. Since variable transconductance topology is proposed in this PGA and g_m is nearly proportional to the gate width of the transistor known from equation (2.18), transistors with gradually reduced gate width are selected for the other amplifier cells, according to the 3 dB gain steps.

Since the bias current flowing through the transistor devices is also proportional to the gate width, the adaptive DC power consumption can automatically be fulfilled with this configuration of parallel connected amplifier cells. On the other hand, more parallel amplifier cells require indeed more chip area; however, since the most amplifier cells are designed for the attenuation, devices with only very small dimensions are used in these cells. Especially in case that the LC tank is used and also integrated on the chip, compared to the chip area occupied by the inductor and the capacitor applied in the LC tank, the chip area invested for the most attenuation cells is almost negligible.

Usually, a PGA is applied in a feedback loop to form an automatic-gain control (AGC). The PGA is implemented in an analog radio frequency transceiver IC, while the rest of the AGC is realized in a digital base-band modem (modulator/demodulator) IC. In this design, the input bits of the demultiplexer are used as the digital interface to the outside. The design logic and the schematic of the 5 to 18 demultiplexer are presented in the appendix.

3.1.3 Simulation and Measurement Results

Simulation for the PGA operating at 1 GHz is first implemented with the simulation tool Cadence. The largest gain should be obtained at the maximum digital control word, i.e. at 18 in decimal form and 10010 in binary form. The gate width of the active devices used in the different amplifier cells are determined step by step, according to the maximum gain and the 3 dB gain steps. The forward transmission coefficient S_{21} is first simulated, which is used here to indicate the gain. The simulation results are presented in Fig. 3.10.

Obviously, a maximum gain of 9 dB is obtained at the first gain step at about 1 GHz. The following gain steps are also achievable with the reduced digital control words. The dB-linear gain variation is realized in this manner. The 3-dB bandwidth of this circuit is about 240 MHz. Since the bandwidth is mainly determined by the common LC-tank, all the amplifier cells have a similar bandwidth.

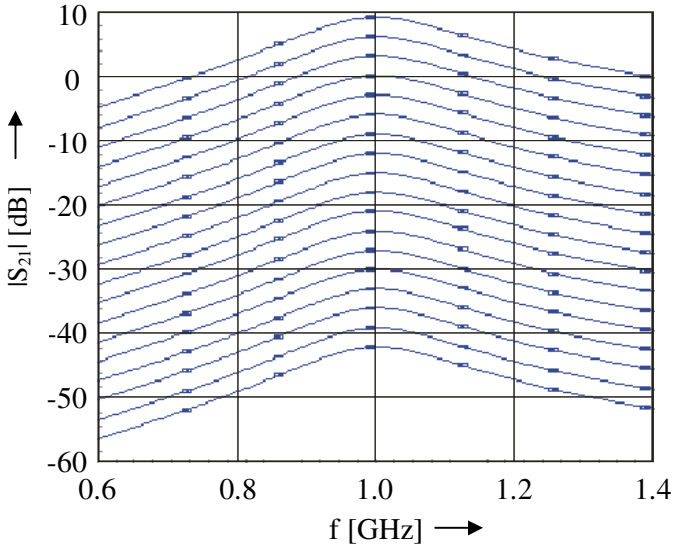


Fig. 3.10. Simulation results of the magnitude of S_{21}.

The power transfer function simulated at the maximum gain (9 dB) is shown in Fig. 3.11. This simulation is implemented at 1 GHz. It can be seen that the maximum output power reaches approximately 10 dBm, while the 1-dB compression point is located at about 7 dBm.

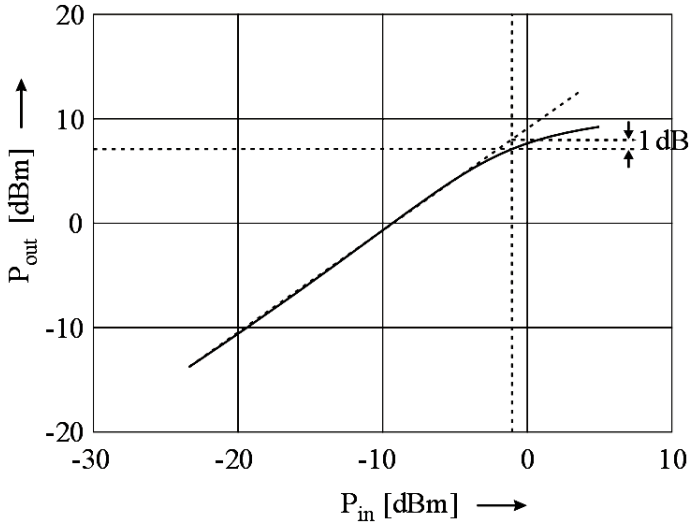


Fig. 3.11. The simulated output power versus the input power.

The two PGAs separately designed at 1 GHz and 2 GHz actually have the same circuit concept. They differ from each other only by the different capacitance used in the load. Therefore, similar results have also been achieved in the simulations for the PGA operating at 2 GHz, except a much larger bandwidth obtained for it due to the lower quality factor, since the capacitance in its LC tank is four times smaller than that used in the PGA for 1 GHz.

The layout of the PGA circuit operating at 1 GHz is shown in Fig. 3.12. The active area occupies 0.5 x 0.4 mm^2. The pads for the supply voltages U_{DD} and U_{dd} are also indicated in the figure. The 5 to 18 demultiplexer is also integrated on the chip. Clearly, it only uses a very small chip area. The pads B_1 – B_5 indicate the five inputs of this demultiplexer. Moreover, the PGA designed at 2 GHz uses a smaller capacitor in its load, while all its other parts are identical to those in the PGA designed at 1 GHz.

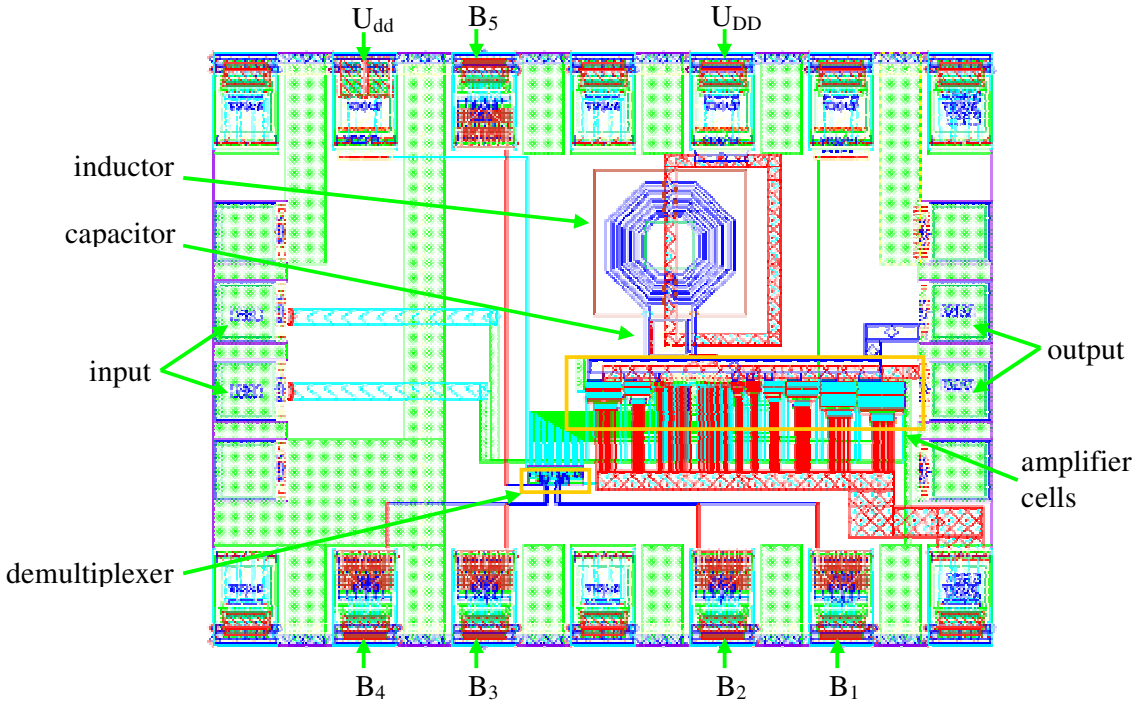


Fig. 3.12. The layout of the PGA operating at 1 GHz.

The PGA test chips are fabricated in the Infineon standard 0.12-µm CMOS technology. A micrograph of the chip for the PGA circuit operating at 1 GHz is shown in Fig. 3.13. Due to the passivation layer on the chip surface, only the inductor and the capacitor are observable. The different pads are also indicated in the micrograph.

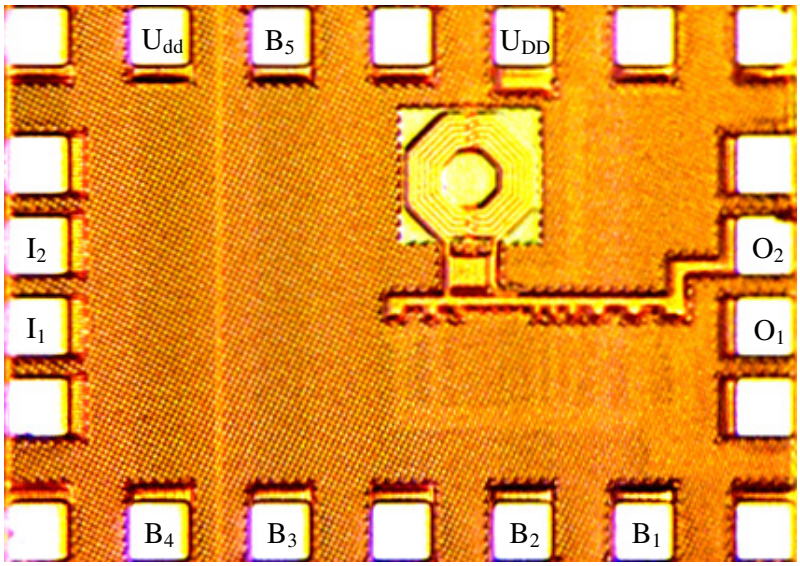


Fig. 3.13. The micrograph of the PGA chip operating at 1 GHz.

On-wafer measurements are implemented. The pads of B_1 – B_5, as well as the pads for the supply voltages U_{DD} and U_{dd} are contacted using two GPPGPPG microprobes of the company CASCADE MICROTECH, where G symbolizes the ground and P symbolizes the power. The DC power and signals are fed through the P pads into the circuit. The other two microprobes of GSSG are used to contact the input and the output pads, where S symbolizes the signals. Not only RF signals but also the DC biasing voltages are provided through the signal pads.

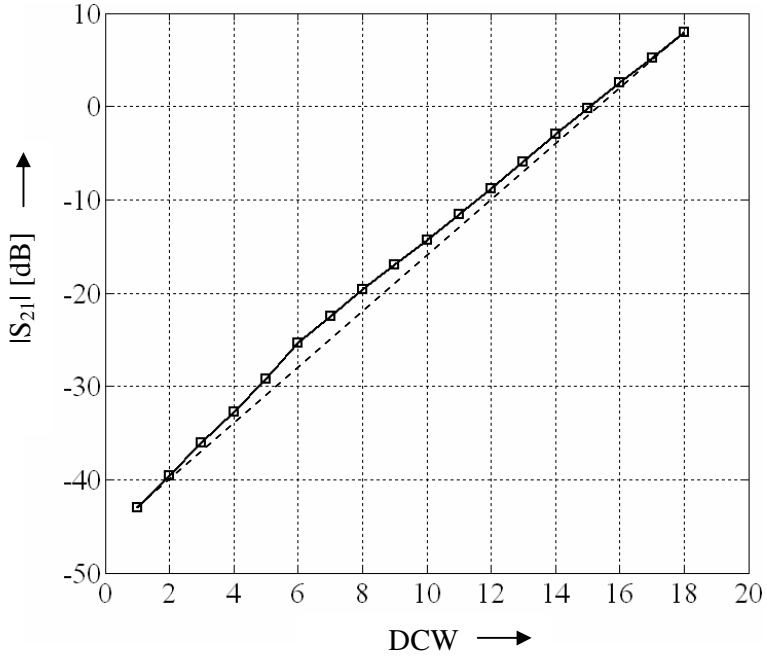


Fig. 3.14. The measured gain versus digital control words at the frequency of 1 GHz.

Using HP 8510B Network Analyzer, the forward transmission coefficient S_{21} indicating the voltage gain is measured according to different digital control words (DCW). Fig. 3.14 shows the measurement results obtained at 1 GHz. The gain can approximately be varied from - 43 dB up to 8 dB with a control step of about 3 dB. A gain control range of 51 dB is achieved. This result implies that the technique of parallel amplifier cells is effective for a dB-linear PGA design.

The measured gain offset ΔG versus digital control words is shown in Fig. 3.15. Theoretically, an even larger dynamic range can be realized. However, the attenuation range can sometimes be limited at a high frequency due to the direct coupling between the input und the output of the circuit through the silicon substrate.

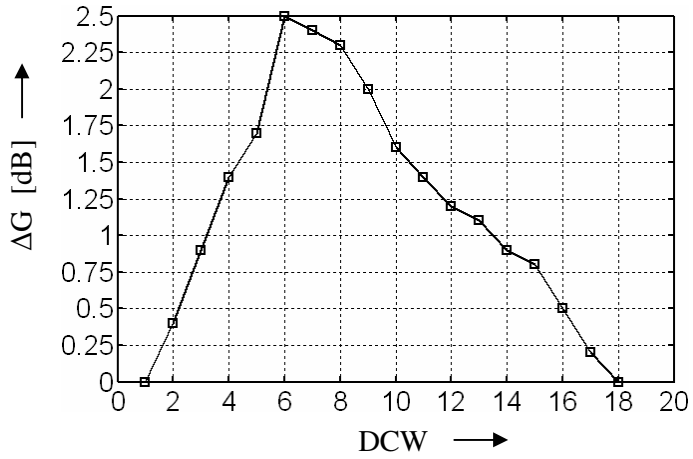


Fig. 3.15. The measured gain offset versus DCW at the frequency of 1 GHz.

Since all the amplifier cells have the same load, the PGA has a nearly constant bandwidth while the gain changes. The frequency response of the PGA is shown in Fig. 3.16, which is measured with the gain of - 4 dB and -16 dB at 1 GHz. A 3-dB bandwidth of about 200 MHz is obtained. It is a little bit smaller than the simulated value mentioned above.

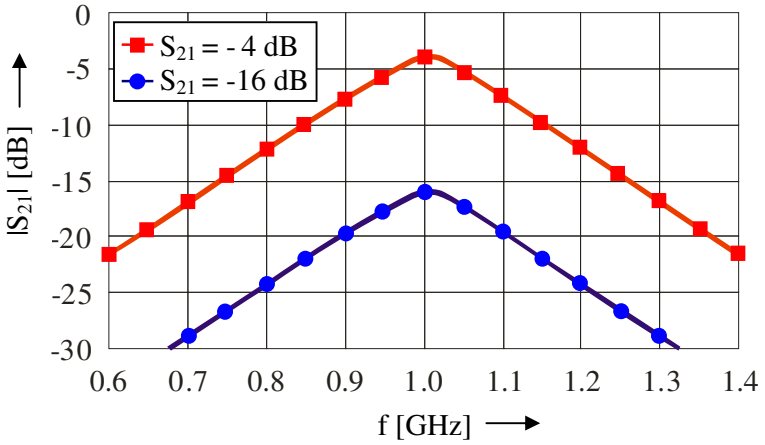


Fig. 3.16. Measured frequency response of the PGA
with the gain of - 4 dB and -16 dB at the frequency of 1 GHz.

During the S-parameter measurements, HP 4155B Semiconductor Parameter Analyzer is used for the DC current measurement, which simultaneously provides the DC supply voltage.

The measured DC current versus the DCW is presented in Fig. 3.17. Obviously, it can be reduced with decreased voltage gain. Adaptive power consumption is obtained.

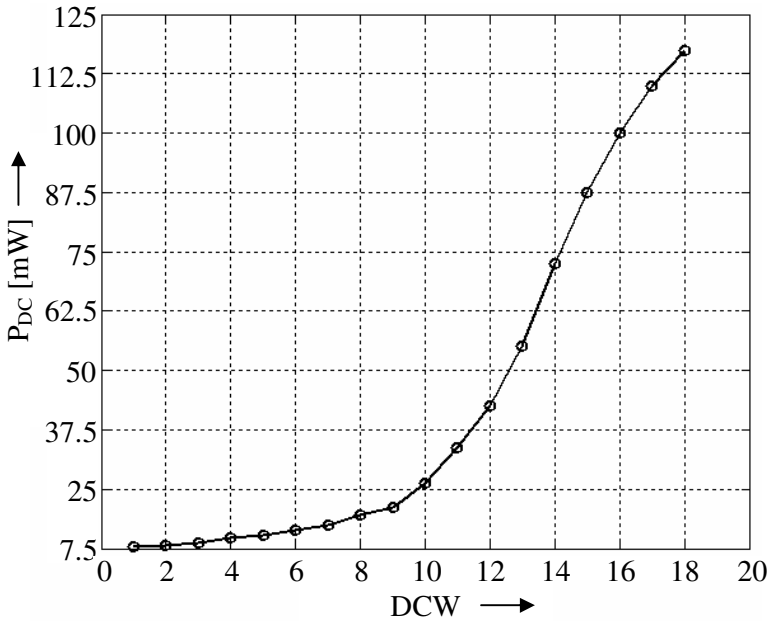


Fig. 3.17. The DC current consumption versus the DCW at the frequency of 1 GHz.

Fig. 3.18 shows the input-referred output power and the oIP3 measured at 1 GHz and with the gain of 8 dB. The maximum output power reaches the level higher than 9 dBm, while the 1-dB compression point is located at 8 dBm. The oIP3 is as high as 22 dBm.

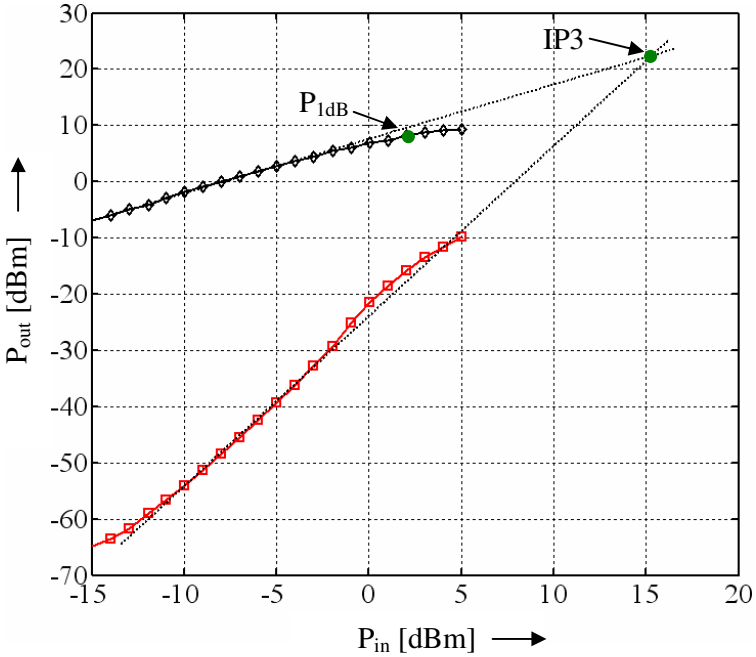


Fig. 3.18. Measured output power and IP3
versus the input power at the frequency of 1 GHz.

The measured results of the third-order and fifth-order intermodulation distortion (IMD3 and IMD5) are shown in Fig. 3.19, as a function of the PGA output power. IMD3 better than - 40 dBc is accomplished when the PGA output power is lower than about 3 dBm. However, it can be degraded if the output power increases. Compared with an IF PGA, this is a typical critical feature of an RF PGA. Therefore, an RF PGA or VGA should usually be linearized using a predistorter [58].

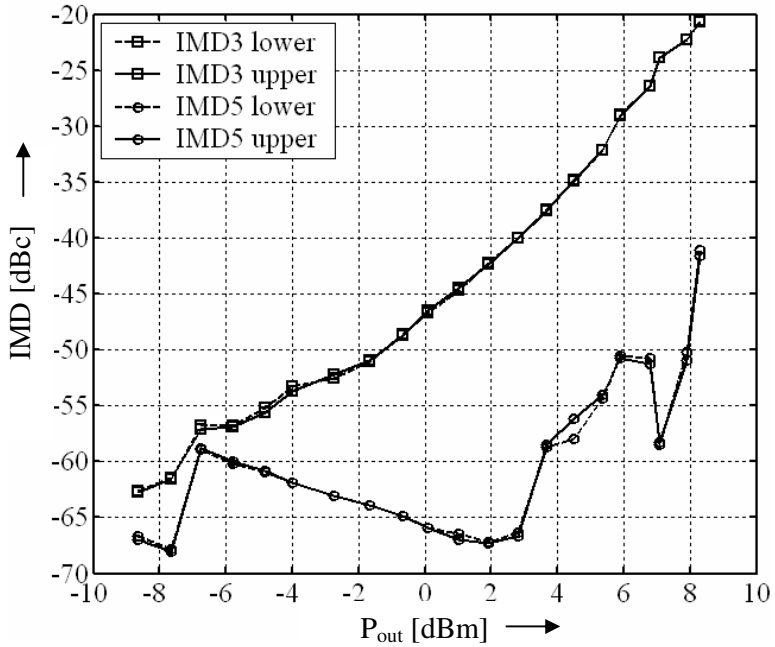


Fig. 3.19. Measured IMD at the frequency of 1 GHz.

It can be seen; almost all the requirements of the PGA circuit are fulfilled. Similar measurement results are also obtained for the other PGA operating at 2 GHz. The single significant difference is that the PGA working at higher frequency has a much larger bandwidth of about 700 MHz.

3.1.4 Conclusion

In this section, an overview of the PGA and VGA design concepts are first given. A novel configuration using parallel amplifier cells is further introduced for the design of a PGA with large logarithmic gain variation. Using this new circuit concept, radio frequency PGAs are designed and fabricated in this work. A standard 0.12-µm CMOS technology is used. The realized PGA has a gain control range of 51 dB and a gain control step of 3 dB. This is the largest linear gain control range reported so far for a radio frequency PGA in a CMOS technology. The maximum output power reaches 9 dBm, while the 1-dB compression point is located at 8 dBm. A high linearity denoted by the oIP3 of 22 dBm at the maximum gain is also achieved in the measurement. Adaptive DC power consumption is also achieved. The configuration of parallel amplifier cells is proven to be a feasible design concept for logarithmic programmable gain amplifiers. Certainly, it can also be applied in many other device technologies.

3.2 A High Voltage/High Power Class A Power Amplifier

The power amplifier is an indispensable component in the cellular phone, which is simultaneously a design bottleneck in the deep-submicron CMOS process. Usually, large output voltage is required to achieve large output power. However, on the other hand, the maximum allowable voltage of a MOSFET in a deep-submicron CMOS process is well limited. For instance, the supply voltage of a conventional 0.12-µm CMOS transistor is only 1.5 V and the breakdown voltage is close to 2.5 V. The gate-drain voltage of a MOS transistor is especially critical due to the field distribution along the channel and the extreme thin gate oxide. This restriction limits the maximum allowable output power of a single conventional transistor. Recent years have seen worldwide efforts to develop the CMOS power amplifiers. Almost all of them have been implemented in 0.35-µm [59] [60] or 0.18-µm [61] CMOS processes. In this work, however, a class A power amplifier using 0.12-µm CMOS technology is attempted.

To solve the problem of low breakdown voltage, the HiVP structure is applied in this design. Using a 0.12-µm standard CMOS technology, a microwave class A power amplifier is developed in this work. The original expected specification of it is the RF output power of 2 W (33 dBm), obtained with a DC power supply of 3.6 V.

In subsection 3.2.1, the design concept of the CMOS power amplifier is first demonstrated. The simulation and experimental results of the proposed power amplifier circuit are presented in subsection 3.2.2, where, according to the measurement results and the results of re-simulation, the difficulty to design a CMOS power amplifier, especially with a low power supply, is analysed and explicated. Suggestions to achieve even larger RF output power are described. Moreover, to enhance the efficiency of the power amplifier circuit on lower power levels, the technique of digitally controlled gate width is introduced in subsection 3.2.3, which is followed by a conclusion given in subsection 3.2.4.

3.2.1 Design and Simulation of a CMOS HiVP Power Amplifier

As mentioned above, an integrated class A power amplifier using 0.12-µm CMOS technology is proposed in this work. The simplified schematic of the power amplifier circuit is shown in Fig. 3.20, which is mainly composed of the HiVP structure introduced in subsection 2.3.2 and the input and the output matching networks M_I and M_O. The supply voltage U_{dd} shown in the schematic is 3.6 V.

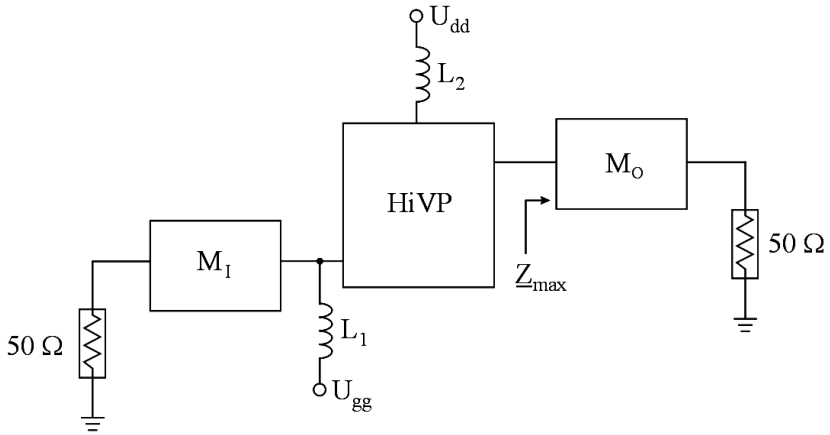


Figure 3.20. Simplified schematic of the CMOS HiVP power amplifier.

As mentioned above, the transistors should ideally have identical operating points since the same current flows through them. Therefore, the numbers of the transistors used in the HiVP structure must be limited since the supply voltage is limited. In this work, three NMOS transistors are connected in series. The DC-simulation result of the HiVP structure is depicted in Fig. 3.21 compared with that of a conventional single transistor, which has the same gate width of 4.5 mm as those transistors used in the HiVP structure.

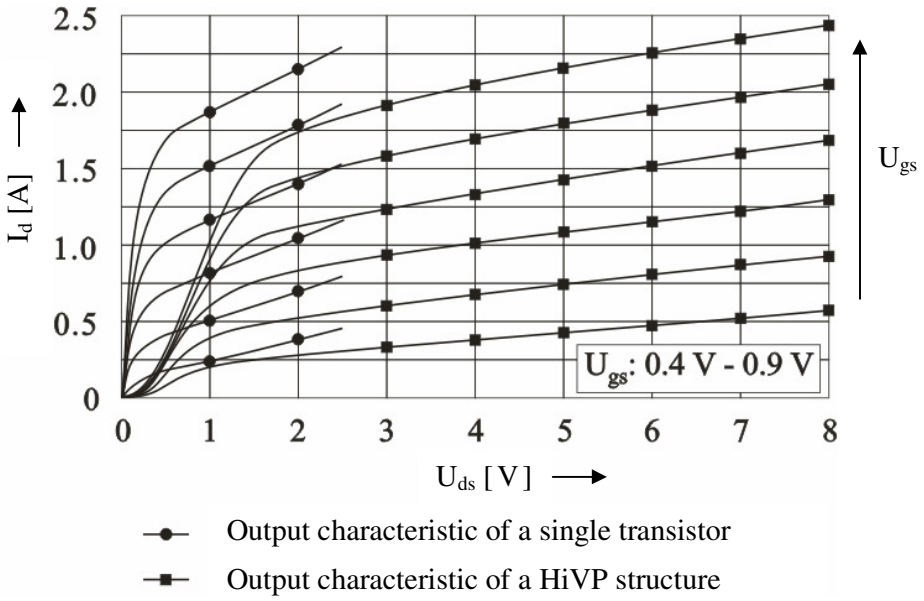


Fig. 3.21. DC simulation results of a single MOSFET and of a HiVP structure.

Obviously, they have the similar output characteristics, which indicate that a HiVP structure can be considered and used as a single transistor device. The gate of the bottom transistor can be regarded as the input and the drain of the top transistor as the output of the HiVP structure. The difference between a HiVP configuration and a single conventional device is that the former can carry a much larger DC voltage.

A shortcoming of the HiVP configuration can be observed in Fig. 3.21. Compared to a single transistor, the lines of the output characteristic of the HiVP structure are shifted to the right. Since the supply voltage U_{dd} is divided by all the transistors in the HiVP structure, at a small value of U_{dd} the transistors are not turned on. According to the load-line matching theory introduced in section 2.1, the maximum range for the voltage swing is required for the maximum achievable output power. In case that the lines of output characteristic are shifted to the right, this condition is definitely damaged. The more transistors are used in the HiVP structure, the more critical is this problem. This is another important reason to limit the number of the active devices used in the HiVP structure as far as possible.

Furthermore, the input and the output matching networks, M_I and M_O, are used in order to obtain a complex conjugate matching at the input and a power matching at the output of the circuit. To increase the matching flexibility and avoid excessive power loss of on-chip inductors, the input and output matching networks are designed to be off chip. The output matching network is especially important. Due to the low DC supply voltage, an impedance transformation at the output, namely from the 50 Ohm load to a much smaller resistance, is necessary. In case that the required output power is 2 W, the maximum value of the transformed impedance $\underline{Z}_{max}$ can be calculated as follows.

$$\underline{Z}_{\max} = \frac{U_{dd}^2}{2P_{\max}} = \frac{\left(3.6\ V\right)^2}{2 \cdot 2\ W} = 3.24\ \Omega \tag{3.4}$$

With the 3.24 Ohm load, the peak RF current will not exceed $U_{dd}/R_{max} \approx 1$ A, and the DC drain current bias must be set approximately to this value. Since the peak drain current is the sum of the bias and peak RF current, the transistor must be designed to supply about 2 A with minimum voltage drop. To fulfil this condition, a gate width of 3.4 mm is first selected in the simulation. In this case, in order to obtain the output power of 2 W, the peak value of the drain voltage of the top device must be higher than 11 V [24], which requires at least four transistors to be used in the HiVP structure. In this work, however, it is attempted to use fewer transistor to minimize the problems with more transistors as mentioned above. In this case, however, the ability of the HiVP structure to sustain high voltage is degraded; hence the drain voltage of the top device (e.g. T3 in Fig. 2.20) must be reduced. On the other hand, to guarantee the expected output power, the current flowing through the HiVP structure must further be increased. This means that the large output power should be obtained more by an even larger current rather than by a higher voltage. Two modifications should be implemented here. The first one is going on to reduce the optimum output impedance (even less than 3.24 Ohm). The ratio of the impedance transformation from the load resistance of 50 Ohm to the optimum output impedance is enhanced. Therefore, the ratio of the voltage transformation through the output matching network is also increased. In this case, the voltage amplitude of 14 V required on the load resistance can be accomplished by a lower drain voltage of T3. The other modification is going on to increase the gate width of the transistors, in order to allow a larger current to flow through the HiVP structure. Finally, transistors having a gate width of 4.5 mm are selected in this work.

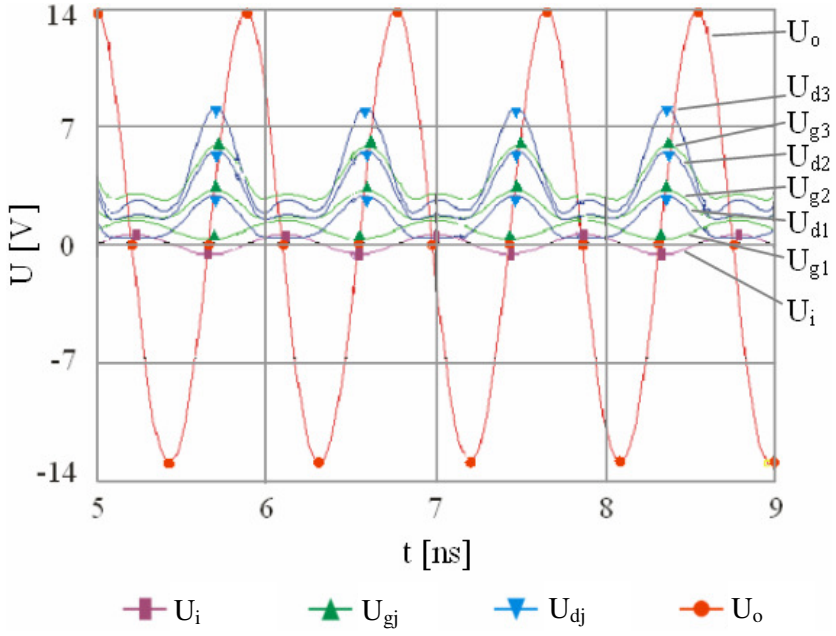


Fig. 3.22. Simulated voltage waveforms of the HiVP structure at the frequency of 900 MHz.

A class A power amplifier is designed. The optimized gate voltage of U_{gg} = 0.8 V is selected by the simulation. A large signal simulation is implemented at the frequency of 900 MHz, which is approximately the center frequency of the GSM uplink band in Europe. A

sinusoidal signal having amplitude of 0.5 V is fed into the input of the HiVP structure, namely into the gate of the bottom device T1. The simulated waveforms of the output voltage U_o as well as the gate voltages U_{gj} and the drain voltages U_{dj} of all the three transistors are presented in the time domain as illustrated in Fig. 3.22, where j is the integer between 1 and 3.

The RF output voltage on the load resistance of 50 Ω has positive amplitude of 14 V and negative amplitude of -13 V, a slight distortion is generated. The accomplished output power is approximately 32 dBm, which is also shown by the power-transfer (P_{out}-P_{in}) simulation. The voltage at the drain of the top device U_{d3} is equal to the supply voltage U_{dd} added by an RF voltage swinging around it. The maximum value of the voltage obtained of at the drain of the top device is 7.8 V. On the other hand, the lowest value of the U_{g1} is 0.3 V. Obviously, U_{d3} can be equally divided by all the three transistors used in the HiVP structure. The gate-drain voltages of all the transistors remain smaller than 2.5 V. All the transistors operate in safe operating area. The equal voltage division can be achieved by the optimization of the resistances used in the voltage divider as described in [24].

As mentioned above, transistors with gate width of 4.5 mm are employed in the proposed HiVP power amplifier. Therefore, compact layout for such large transistors is required. On the other hand, the transistors must be able to sustain a large current. The 0.12-µm CMOS technology used in this work has six metal layers. The materials and the thickness of these metal layers are provided in Fig. 3.23. Obviously, the top metal layer, namely the metal 6 has the largest thickness, compared to the other five metal layers. Therefore, it has the capability to sustain a much larger current. To obtain a compact layout for the large transistors and hence for the entire HiVP structure, all these metal layers must be used.

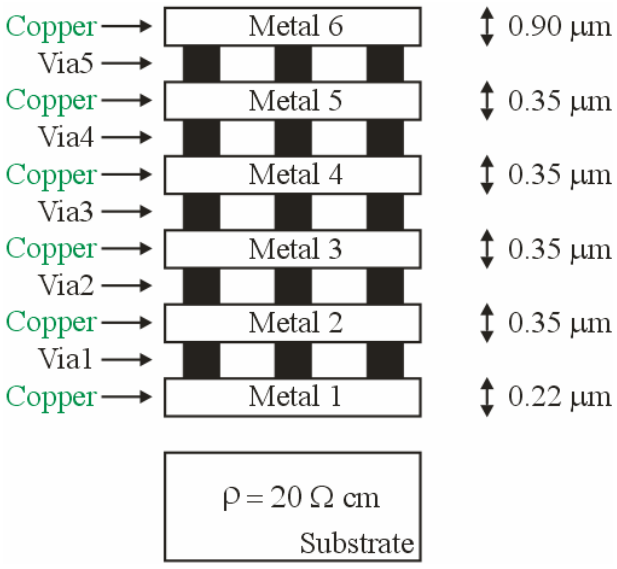


Fig. 3.23. The description of the metal layers of the Infineon 0.12-µm CMOS technology.

In Fig. 3.24, the layout concept for a transistor having a gate width of 4.5 mm is illustrated. To realize such a large transistor, 496 small transistor cells are connected in parallel, each of which has a gate width of 9.07 µm. They are equally distributed in a 16 x 31 matrix. In this manner, significant voltage attenuation on a long poly-Si line is avoided. Both the drain and the source are constructed using metal 3 – metal 6. In order to mitigate the parasitic capacitance rising from the overlap of the different metal layers used for the drain and the source, lots of contacts are applied to connect the metal layers.

$W_{D,S}$ in Fig. 3.24 indicates the maximum width of the metal lines, which connect the drain and source of every transistor. They are determined according to the magnitude of the current

flowing through the transistor and the current-carrying-ability of metal 3 – metal 6. Observing the layout given in Fig. 3.24, the drain is arranged at the top of each transistor and the source is located at the bottom. The current flows downward through the transistor and is allocated by all the transistor cells. The largest drain current exists only on the top of each transistor and can be decreased on the way. Oppositely, the source current is still relatively small on the top of each transistor but it gathers itself top down. Therefore, the largest source current exists only at the bottom of each transistor. As shown in Fig. 3.24, the maximum line width $W_{D,S}$ is only used at the top of the transistor-layout, which is reduced top down. On the other side, the line width for the source starts above with a very small value, increases however top down up to $W_{D,S}$ at the bottom of the transistor-layout. Obviously, due to the arrangement of this staggered form, the layout area can significantly be reduced compared to the conventional design concept of the MOS transistors, in which the width of the metal lines used for the drain and the source remains the same top and down.

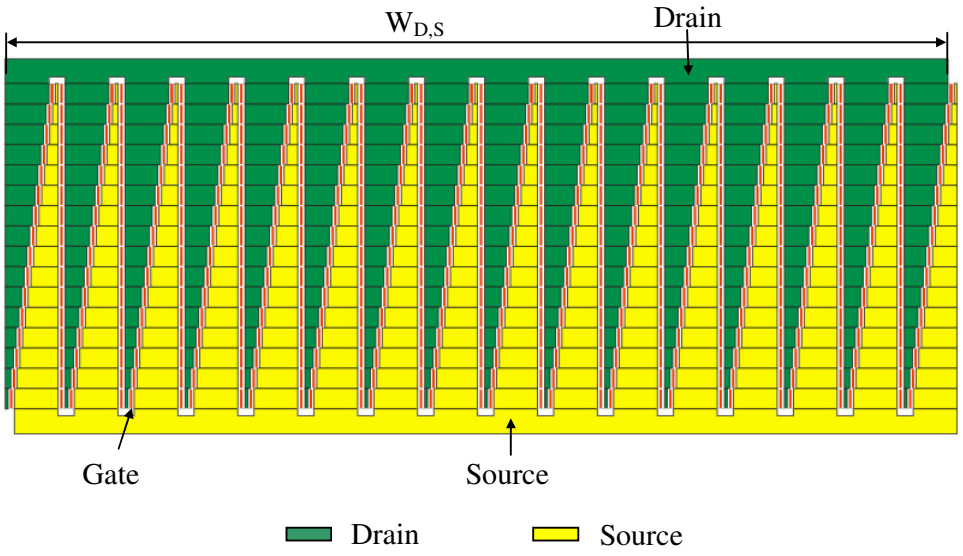


Fig. 3.24. The layout concept for NMOS transistors.

Additionally, a guard ring is used around each transistor to reduce the substrate noise. Therefore, the lowest metal layer, namely metal 1 is reserved to connect the guard ring to the ground. Furthermore, metal 2 is used to connect all the gates of the basic transistor cells.

Another emphasis of the layout design concerns the gate distribution. Assuming the input signal comes from the left of the layout and runs to the right, there is a long distance between the leftmost transistor cells and the rightmost transistor cells due to the large dimension of the transistor layout. This results in a significant phase difference between the input signals of the transistor cells on the left side and the input signals of the transistor cells on the right side. Therefore, the amplified output signals of these transistor cells partly compensate each other; a degradation of the output power is generated. In order to minimise the phase difference, the so called H-structure is used, which is illustrated in Fig. 3.25. In this figure, a chip area is assumed to be divided into 16 parts, which are arranged in a 4 x 4 square matrix. The small circles in the figure indicate the inputs of every part. Apparently, the center input signal is first fed to the middle of the chip area. It is then routed equally into all the directions. The time delay from the center input to all the part-inputs is identical, so that no phase difference exits between the input signals of all the 16 parts. The decisive factor of the using of H-structure is that the entire chip area can be divided into an n x n square matrix. Since the transistor cells used in Fig. 3.24 are distributed in a 16 x 31 matrix (not square matrix), the phase difference between the input signals of the transistor cells can not totally be deleted.

However, using the H-structure, the maximum route-difference of the input signals among all the transistor cells is significantly reduced. The degradation of the output power due to the phase difference is negligible.

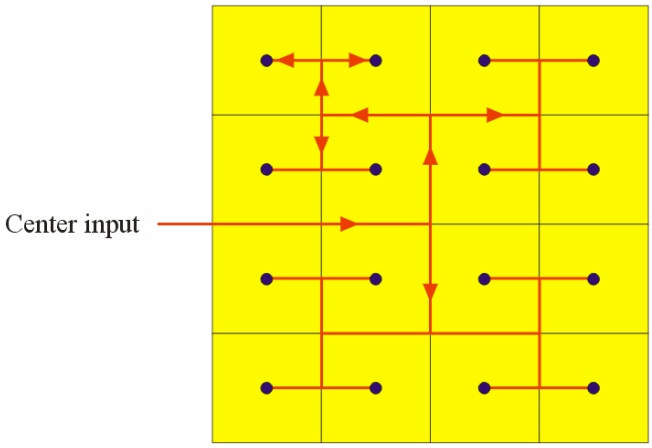


Fig. 3.25. The H-structure.

In the HiVP structure, it is required to isolate the NMOS transistors from the substrate, so that the bulk of every transistor can directly be connected to the source of the transistor as shown in Fig. 2.20. In this manner, the bulk voltage can then swing with the source voltage and the voltage difference between every two terminals of the transistors can be maintained smaller than the breakdown voltage. The 0.12-µm CMOS technology used in this work is a triple-well technology, where the NMOS transistors can easily be isolated from the substrate.

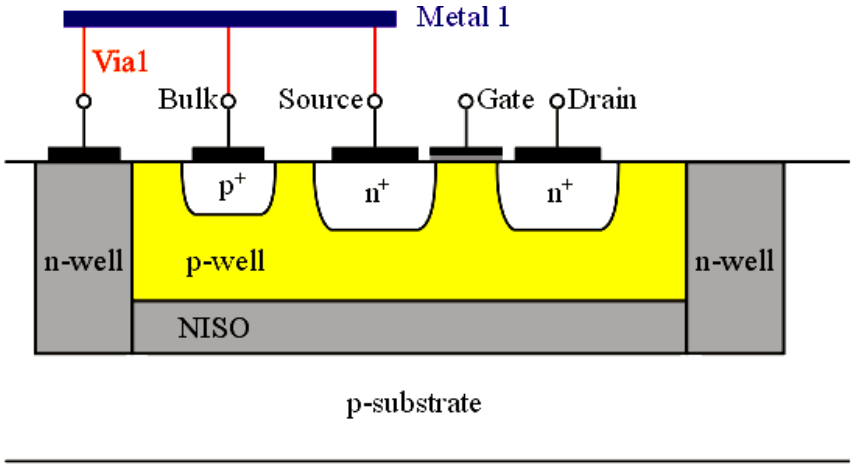


Fig. 3.26. The cross-section of the isolated NMOS transistors.

The cross-section of the isolated NMOS transistors is illustrated in Fig. 3.26. The NMOS transistor is located in the p-well, which is separated by the NISO layer from the p-substrate in the vertical direction and is embraced by the n-well layer in the horizontal sides. The p-well has a determined width of 2 µm according to the design rules. In this manner, a NMOS transistor is totally isolated from the substrate; hence the bulk of it can be connected with the source terminal using metal layers. However, a p-n-junction is constructed due to the direct contact of the n-well and the p-well, which causes additional power loss. To avoid this secondary effect, the n-well layer must be connected to a higher positive voltage level than

the p-well. Another possibility to avoid the power loss is to short the n-well and the p-well directly, also using the metal layers as shown in Fig. 3.26. The NMOS transistors used in the schematic must therefore be symbolized as presented in Fig. 3.27, where the transistor symbol is complemented by the n-well. Additionally, a diode must be arranged at the border of the n-well. In this figure, the p-n-junction of the diode is shorted.

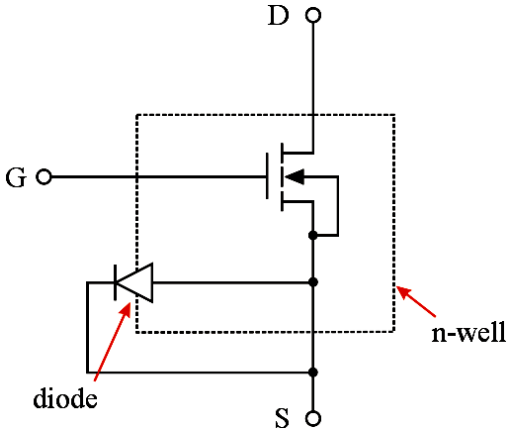


Fig. 3.27. The complete symbol of the isolated NMOS transistors.

The conclusive layout of the transistor with a gate width of 4.5 mm is demonstrated in Fig. 3.28. This layout occupies a chip area of 250 µm x 645 µm.

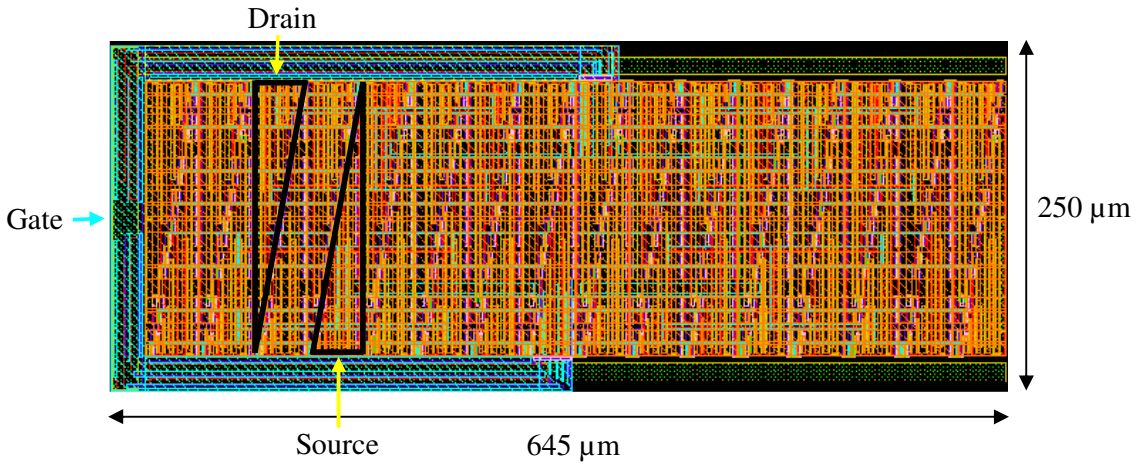


Fig. 3.28. The conclusive layout of the NMOS transistor with a gate width of 4.5 mm.

Finally, the layout of the HiVP structure using three active devices is presented in Fig. 3.29. The different terminals and the active as well as the passive components are indicated. The drain, source, as well as the input and the output of this configuration are also indicated in the figure. Obviously, three large transistors are applied here, which are connected in series and occupy most of the chip area. The voltage divider constructed by the resistors and the capacitors used to determine the drain impedance of the different active devices are arranged on the left of the chip. Metal-insulator-metal (MIM) capacitors are used in this design, which perform high capacitance density, low current leakage and high breakdown. It can also be seen, only the HiVP structure is integrated on this chip. As mentioned above, the input and output matching networks will be realized on the printed circuit board (PCB) later. Bonding

pads are therefore necessary at the terminals of the chip, which are used as connecting points for the terminals of the HiVP and the conductor lines on the PCB.

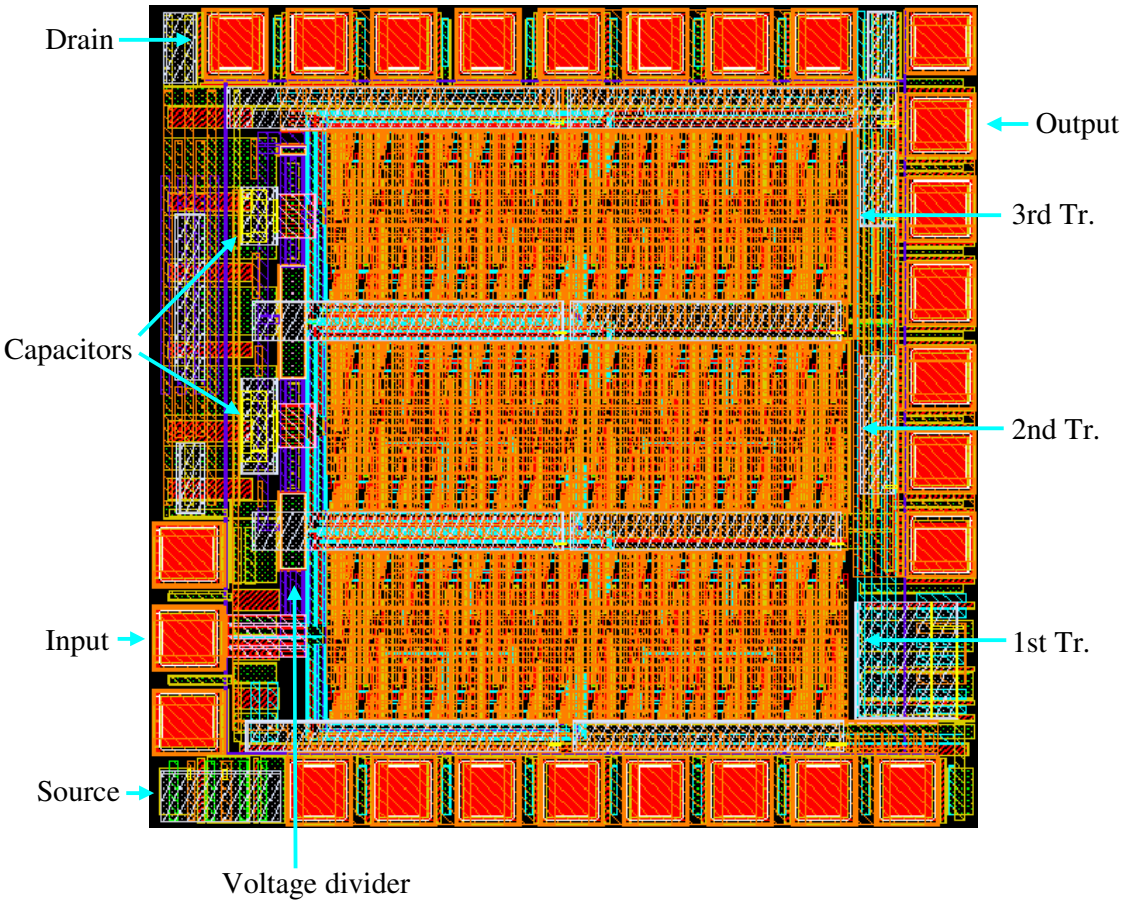


Fig. 3.29. The layout of the HiVP structure using the 0.12-µm CMOS technology.

Also due to the large current flowing through the active devices, numerous bonding wires are required at the drain and source terminals of the HiVP structure, which require numerous pads at these terminals. The total chip area is 975 µm x 970 µm.

3.2.2 Measurement Results

The micrograph of the chip with its bonding wires is shown in Fig. 3.30. In spite of the passivation layer on the chip surface, the three transistors, the voltage divider and the capacitors of the HiVP configuration can well be recognised, compared with the chip layout.

As mentioned above, the HiVP chip should be mounted on a PCB, where the input and the output matching networks of the HiVP power amplifier are built. The layout the PCB is given in Fig. 3.31. The passive components to be mounted are also indicated in the layout. The inductor L_1 and the capacitor C_1 construct the input matching network, while L_2, C_2 construct the output matching network. Several parallel decoupling capacitors C_{b1} – C_{bn} distributed between 1 pF and 100 nF are applied to couple RF signal of different frequencies to the ground. A large inductor L_3 is used as the RF choke at the drain of the HiVP device, and a large capacitor C_3 operates as the coupling capacitor at the output. All these passive components are in the form of surface mount device (SMD) and will be soldered on the metal

lines. The RF choke for the gate of the HiVP device and the coupling capacitor at the input of the amplifier circuit are included by the Bias-Tee. The data of these passive components are listed in Table 3.1.

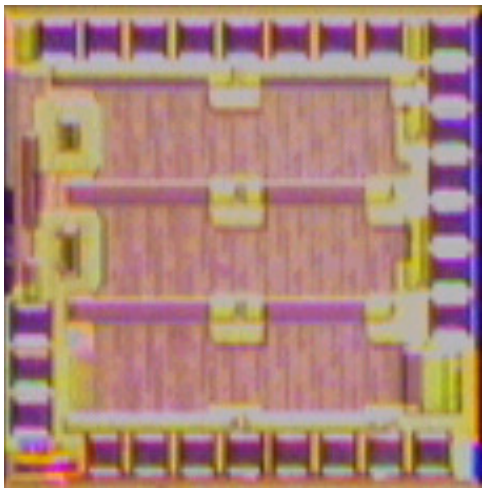


Fig. 3.30. The micrograph of the HiVP chip.

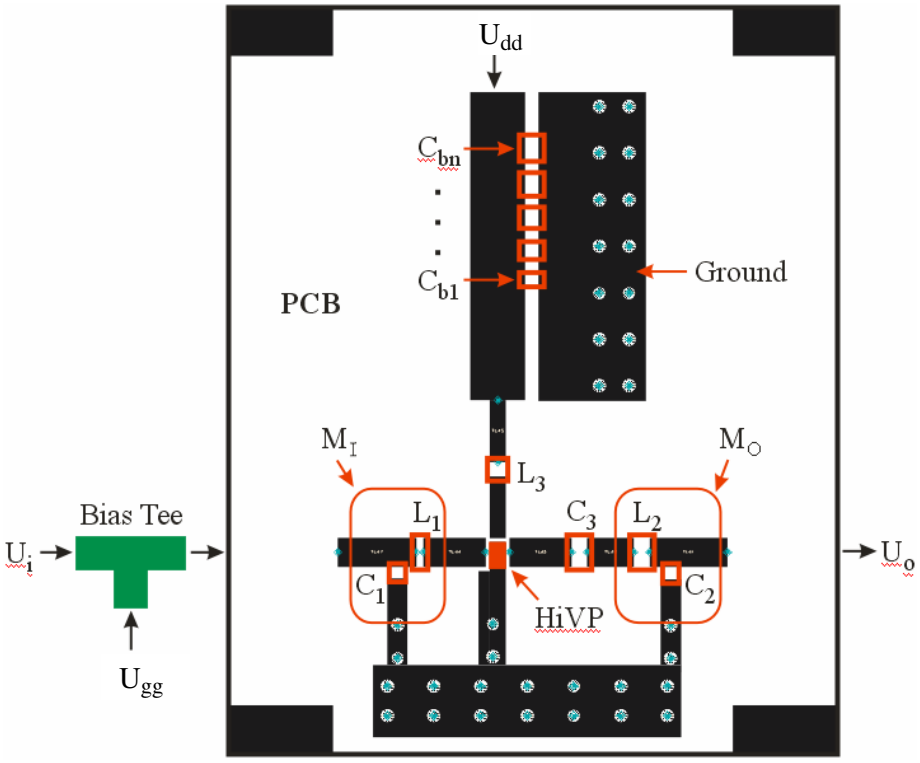


Fig. 3.31. The layout of the printed circuit board.

Table 3.1
The passive components mounted on the PCB

L_1	C_1	L_2	C_2	L_3	C_3
3 nH	4 pF	2.2 nH	13 pF	100 nH	47 pF

Using the bias voltages U_{gg} = 0.8 V and U_{dd} = 3.6 V, the output power, the power gain and the power added efficiency are measured at 900 MHz. The measurement results are illustrated in Fig. 3.32. It can be seen that the maximum output power reaches 29.5 dBm (0.9 W) at the frequency of 900 MHz, while the 1-dB compression point is located at about 27 dBm. The small-signal gain is about 11.5 dB, which is degraded with the increased power level. A maximum power added efficiency of about 34.5 % is achieved in the high output power range.

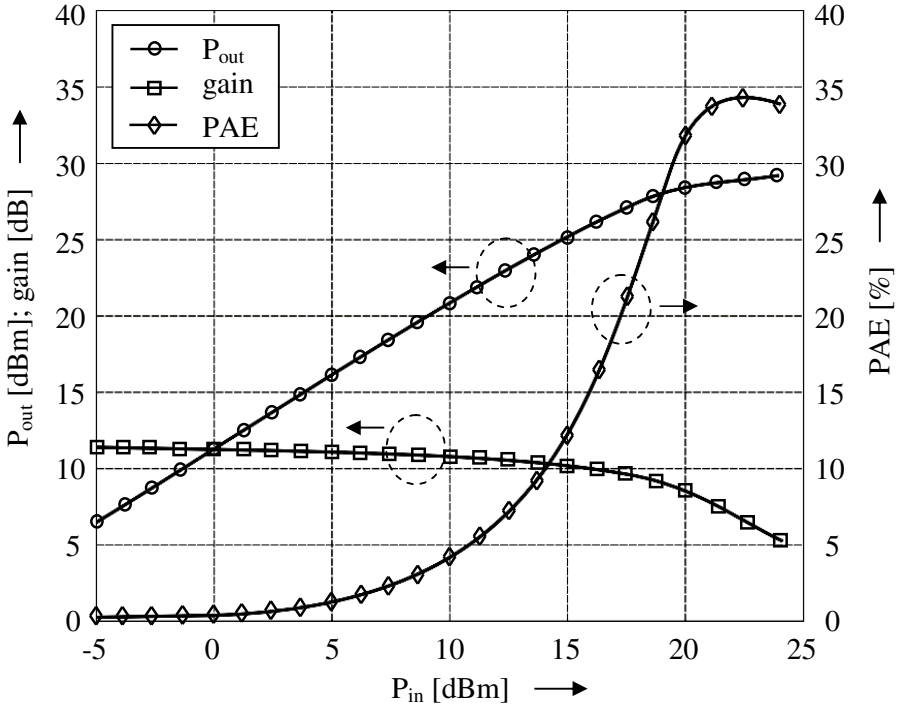


Fig. 3.32. The measured output power, gain and PAE at the frequency of 900 MHz.

Comparisons are done between the power amplifier realized in this work and the other power amplifiers achieved in recent years. Due to the most critical problem of low breakdown voltage, the CMOS power amplifiers realized up to now are mostly implemented in 0.35-µm or 0.6-µm technologies. Only very few works have been attempted with the technologies, in which the gate width of the MOS transistors is smaller than 0.18-µm. Results of the comparisons are summarized in Table 3.2. Obviously, though the expected output power of 2 W is not fulfilled, the HiVP power amplifier realized in this work performs comparable characteristics in comparison to the other works.

Table 3.2
Comparisons of CMOS power amplifiers

Author	**Technology [CMOS]**	**Frequency [GHz]**	**P_{out} [dBm]**	**PAE [%]**	**U_{dd} [V]**
Sowlati et al. [62]	0.18-µm	2.4	23	42	2.4
Reyneart et al. [63]	0.18-µm	1.75	27	34	3.3
Knopik et al. [64]	0.13-µm	2	21	18	2.8
This work	0.12-µm	0.9	29.5	34.5	3.6

In the future, following efforts can be devoted in order to achieve even larger RF output power. The RF choke L_3 mounted on the PCB is selected very large in this work so that it acts

like a current source. As the current flowing through the HiVP structure is very large, the power loss in L_3 can no longer be neglected. The lumped electrical model of an inductor is shown in Fig. 3.33. It consists of an inductance L_s with a series resistance R_s and a parasitic capacitance C_p.

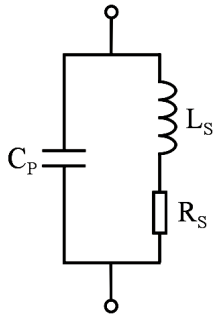


Fig. 3.33. Lumped electric model of an inductor.

The series resistance R_s causes the loss of supply voltage leading to reduced output power. It increases with the increased inductance L_s and is approximately proportional to the square root of the frequency [65].

On the other hand, the self-resonance occurs due to the parasitic capacitance C_p. For an inductor, only the energy stored in the magnetic field is of interest. Therefore, the electric energy present in the parasitic capacitance has to be subtracted from the magnetic energy. At the self-resonant frequency, the energy in the magnetic field is equal to the energy in the electric field and the quality factor of the inductor becomes zero. For the frequencies above the self-resonance frequency, no net magnetic energy is available from an inductor to any external circuit. The quality factor of an inductor can be presented as follows.

$$Q = \frac{\omega L_s}{R_s}\left[1-\left(\frac{\omega}{\omega_0}\right)^2\right], \tag{3.5}$$

where ω_0 is the self-resonance frequency of the inductor given in equation (2.67). Seen from the equation above, Q increases with the frequency at lower frequencies. However, at higher frequencies, the self-resonance effect will be responsible for a decrease of the quality factor and at the resonance frequency, the quality factor becomes zero.

In order to reduce the parasitic resistance of the RF chock and increase its self-resonance frequency, relatively small inductance values can be attempted to be used in the future. However, the drawback of lowering the inductance of L_3 is the current swing in it and this AC current has to be delivered by the power supply. Of course, the voltage source itself is not able to deliver the high frequency AC current. One solution is to use the some decoupling capacitors C_{b1} – C_{bn} as shown in Fig. 3.31, and the larger the current swing is; the larger should be the capacitance. But this method requires very high quality capacitors, which are still capacitive up to the operating frequency. The most off-chip capacitors have parasitic inductance of their package, in series with the capacitor itself. For instance, if a capacitor of 40 pF has a parasitic inductance of about 0.6 nH, then it behaves as an inductance rather than a capacitance at the frequencies above 1 GHz. The real solution for the problem of the AC power supply current is the use of differential structures, which is widely used in many commercial power amplifiers. As introduced in subsection 2.4.5, two amplifiers are placed in parallel and are driven by two out-of-phase input signals. Therefore, the current through the DC-feed is only DC. In case that only single-ended HiVP structure is used in this work, the

RF chock L_3 can only be reduced reasonably. A compromise between the parasitic resistance, self-resonance frequency and the AC current must be found.

Additionally, gate width of 3.4 mm was selected in the previous work achieving the RF output power of 26 dBm at 900 MHz. In this work, output power of 29.5 dBm is obtained by using the transistors having gate width of 4.5 mm. Therefore, even larger gate width can be attempted in the future in order to reach the expected output power of 33 dBm.

Finally, it can be attempted to measure the source resistance of the HiVP structure, namely the resistance occurring between the source of T1 and the ground. Known from the re-simulation, a reduction of the bias voltage and hence of the operating current can be caused due to this source resistance, therefore the output power is reduced. This resistance comes from the bonding wires which connect the source pads to the conductor lines of the PCB and also from the conductor lines themselves. Though several bonding wires are placed in parallel, the parasitic resistance and inductance are not significantly reduced due to the proximity effect [65]. In case that several conductors are placed closely, the presence of nearby currents will change the flux in and around the conductor and it will cause mutual coupling between the conductors. For instance, two metal traces M_1 and M_2, which respectively have resistance of R_0 and inductance of L_0, are in parallel. The currents flowing through them have the same direction. Due to the parallel nature, the total resistance and inductance should normally be $R_0/2$ and $L_0/2$, respectively. However, these values can only be obtained if the distance of these two traces is very large. If they are close to each other, the magnetic field of the two traces will add to each other. Therefore, though the total inductance is smaller than L_0, it is much larger than $L_0/2$. On the other hand, the currents flowing through the two traces are distracted by each other. This means that the currents are pushed outwards by each other and the cross-section for the current are reduced. Therefore, though the total resistance is smaller than R_0, it is much larger than $R_0/2$.

3.2.3 HiVP Design Concept with Adjustable DC Power Consumption

Usually, high power added efficiency is only obtained at high RF power levels. It is also proven by the HiVP power amplifier realized above. As shown in equation (2.16), the DC drain current is proportional to the gate width of the MOS transistors. Therefore, the DC power consumption remains high also if the RF output power is very low, resulting in very low power added efficiency at these power levels.

In order to improve the power added efficiency achieved at low power levels, a novel transistor design concept with adjustable gate width is introduced in this work. A MOS transistor having large gate width can be divided into several parallel branches, each of which can, for example digitally, be turned on or turned off. Therefore, the DC current flowing through the transistor is adjustable. Using the same concept, the HiVP structure is modified and shown in Fig. 3.34.

Obviously, the HiVP structure shown in Fig. 2.20 is here equally divided into n branches. The transistors Ti (i is an arbitrary integer between 1 and 3) having gate width of w are respectively substituted by n transistors Ti' having gate width of w', where

$$w = n \cdot w' . \tag{3.6}$$

It can be seen that each branch of the HiVP structure includes its own voltage divider comprising the resistors R_i' and the shunt capacitors C_i', which are used to determine the drain impedance of the transistors. Because the voltage dividers included by all the branches are connected in parallel, the condition of

$$R_i' = n \cdot R_i \tag{3.7}$$

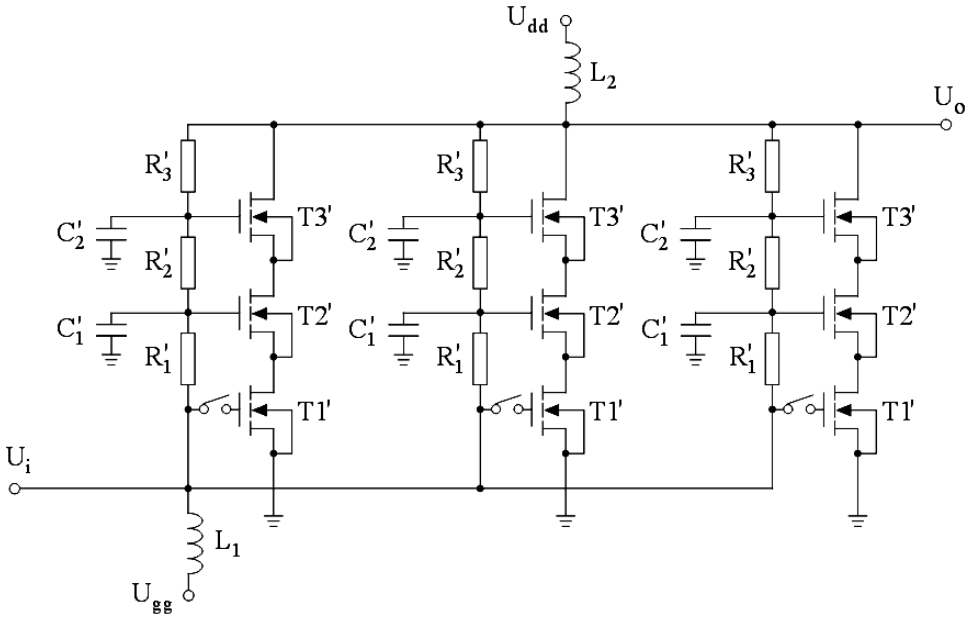


Fig. 3.34. The HiVP structure with n branches.

must be fulfilled, assuming R_i and C_i are the resistors and the shunt capacitors used in the original HiVP structure depicted in Fig. 2.20. Furthermore, the relationship between C_i' and C_i must be found out. From equation (2.26), the shunt capacitors can be described as follows.

$$C_{shunt} = \frac{C_{gs}}{\underline{Z}_{source} \cdot g_m - 1} \tag{3.8}$$

Since the transistors used in each branch of the modified HiVP structure are n times smaller than those used in the original one, C_{gs} and g_m are reduced to C_{gs}/n and g_m/n, respectively. Similar to R_i', the impedance seen at the source of the upper transistors $\underline{Z}_{source}$ must also be increased to $n \cdot \underline{Z}_{source}$, since all the HiVP branches are connected in parallel. Inserting these analyses into equation (3.8), following relationship can approximately be found out:

$$C_i' = \frac{1}{n} \cdot C_i \tag{3.9}$$

As shown in Fig. 3.34, the input ports, namely the gates of all the parallel HiVP branches are connected by the switches. These switches can be digitally controlled. In case of low RF output power, several switches can be turned off; hence the corresponding HiVP branches are disconnected. In this manner, the DC current flowing through the HiVP structure is decreased; the power added efficiency of the power amplifier circuit is therefore increased at the low power levels. There are also other methods to disconnect the HiVP branches, for instance, the disconnection at the drain of each branch. However, since the metal lines used to connect the gates of the HiVP branches are relatively thin, it is more feasible to disconnect the gate line than to disconnect the drain line.

In this work, the integer n equal to 10 is selected. Therefore, the gate width of the transistors T1'-T3' used in each HiVP branch is 450 µm. Simulation is also done at the frequency of 900 MHz. In case that all the branches are turned on, the maximum RF output power of 2 W can also be obtained during the simulation. With the degraded output power,

the HiVP branches are gradually turned off. Finally, only one branch is applied for the lowest output power.

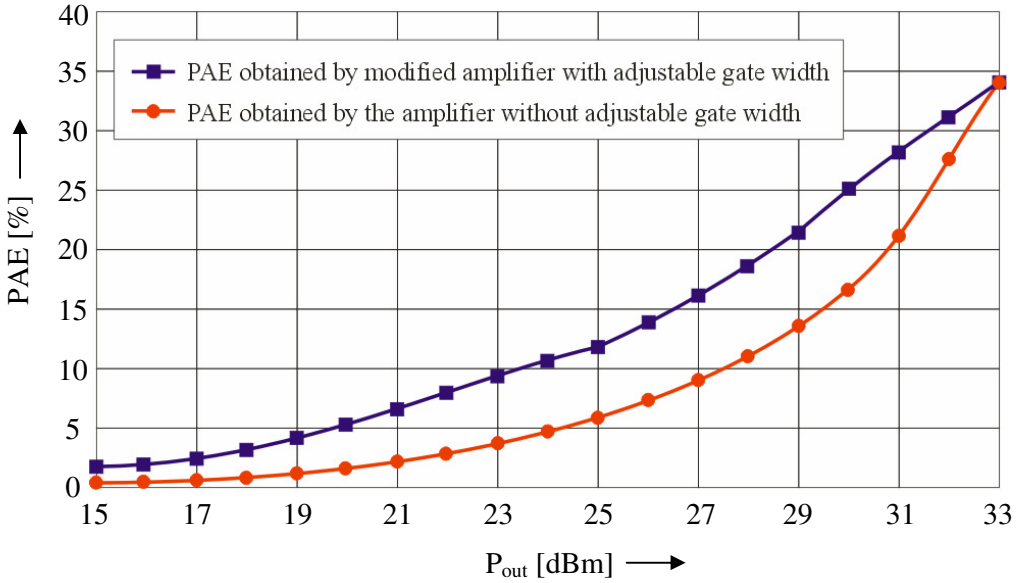


Fig. 3.35. The simulated power added efficiency obtained for the modified HiVP power amplifier with adjustable gate width.

The simulated power added efficiency versus the output power is presented in Fig. 3.35. Clearly, using the technique of adjustable gate width, the power added efficiency of the HiVP power amplifier can significantly be improved in the low power range.

3.2.4 Conclusion

The most important advantage of the HiVP configuration is the equal division of a large voltage, which solves the problem of low breakdown voltage of the transistors in deep-submicron CMOS technology. In this work, the methods to improve the performance of the HiVP structure are described. Moreover, using the adjustable gate width technique, the power added efficiency of the HiVP power amplifier can clearly be improved in the low power range. Therefore, using HiVP configuration, deep-submicron CMOS technology is theoretically a feasible process to realize power amplifiers applied in cellular phones of mobile communications systems. A monolithic integration of such transceiver system in a deep-submicron CMOS technology becomes possible. In this case, the fabrication of such transceiver systems will be simplified and the process cost will be reduced. Certainly, the HiVP configuration can also be adopted in other ranges, e.g. satellite communication and split-ring resonator microplasma [66], as well as in the other technologies, e.g. GaAs-FET technology.

4. Design of the Broadband LDMOS Power Amplifiers

At the beginning of the 1990s, digital mobile radio frequency communication began to make first steps into the market. Since that time, several different generations of mobile communications systems exist simultaneously on the market. Therefore, there is a strong demand to design multi-frequency multi-standard transceivers, which can operate in different systems. For a long time, such transmitters could only be implemented by replicating the radio frequency transmitters and receivers for each operating band. Today, efforts have been devoted to directly develop the broadband circuits and passive components used in the multi-frequency multi-standard transceivers [67]-[69]. In this work, the design of the broadband power amplifiers for several downlink applications is introduced.

Using Motorola LDMOS device MRF21030SR3, a single-ended broadband radio frequency power amplifier is first designed and fabricated in this work. Section 4.1 describes the design process for this amplifier in detail. Due to the compromise of the linearity and the power added efficiency, a class AB power amplifier is proposed. Not only the simulation but also the measurement results are presented. Furthermore, using the same power devices, the same biasing networks, the same operating points and the same impedance matching networks, a broadband balanced power amplifier is designed. In section 4.2, the functional principles of the balanced structure, as well as the benefits of it are introduced. Excellent performances are achieved during the measurements.

4.1 Design of a Broadband LDMOS Single-Ended Class AB Power Amplifier

Using Motorola LDMOS device MRF21030SR3, a single-ended broadband radio frequency power amplifier for the base station applications is first proposed, which can simultaneously be used in GSM1800 and UMTS systems in Europe. Known from Fig. 1.1, this power amplifier should have a bandwidth of at least 400 MHz. Furthermore, following specifications should also be fulfilled:

- maximum output power larger than 43 dBm
- small signal gain larger than 10 dB
- maximum power added efficiency higher than 30 %
- high linearity indicated by ACPR better than 40 dBc

In this section, the design process to develop such a broadband power amplifier is introduced in detail. It includes the choice of simulation model of the power devices, determination of the operating point, improvement of the stability and design of the impedance matching networks, etc. Simulations are first implemented, in order to predict and optimize the circuit performance. The proposed power amplifier circuit is then fabricated in the laboratory and measured. Good agreements are obtained between the simulation results and the measurement results.

4.1.1 Selection of the simulation model of LDMOS transistors

The device manufacturer Motorola provides two different models for the simulation, namely the Root-model and the MET-model. The former is a table-form model, which is previously derived as “HP Root FET Model” by Agilent. This model can predict the behavior of the PA circuits in dependence on the operating point, the frequency and the power. Naturally, it is only adapted for the simulation if the measured data are available in some fixed ranges.

The MET-model is an electric-thermal model, in which the dynamic process of the heating is considered. The MET-model of the Motorola LDMOS transistors is an empirical nonlinear large-signal model. The transfer characteristic $I_d - U_{gs}$ as well as the transconductance g_m are modelled in terms of each operating point and each temperature. With this model, small-signal-, large-signal-, harmonic balance-, noise- and transient-simulations can be implemented. Therefore, such model of power devices is suitable for the design of power amplifiers used in base stations. In this work, the MET-model of the LDMOS transistors is selected for all the simulations. The simulation of temperature behaviour can be turned on or turned off, in dependence on the necessity.

4.1.2 DC Simulation and Selection of the Operating Points

To obtain the high output power and simultaneously achieve the high power added efficiency mentioned above, a class AB power amplifier is developed in this work. It is therefore necessary to select the suitable operating points for the LDMOS transistor. Using the transistor model, DC simulations are first implemented. The recommended drain-source voltage U_{ds} for MRF21030SR3 is 26 V. The simulated transfer characteristic of the transistor, which denotes the relationship between the drain current I_d and the gate-source voltage U_{gs} for the given drain-source voltage $U_{ds} = 26\ V$, is shown in Fig. 4.1.

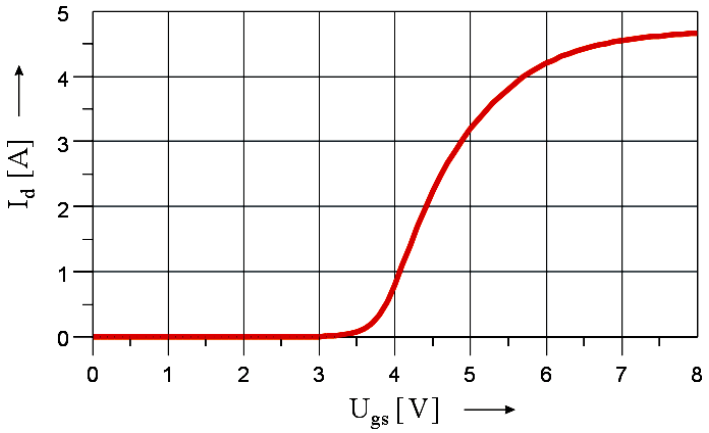


Fig. 4.1. The simulated transfer characteristic of MRF21030SR3 at U_{ds} = 26 V.

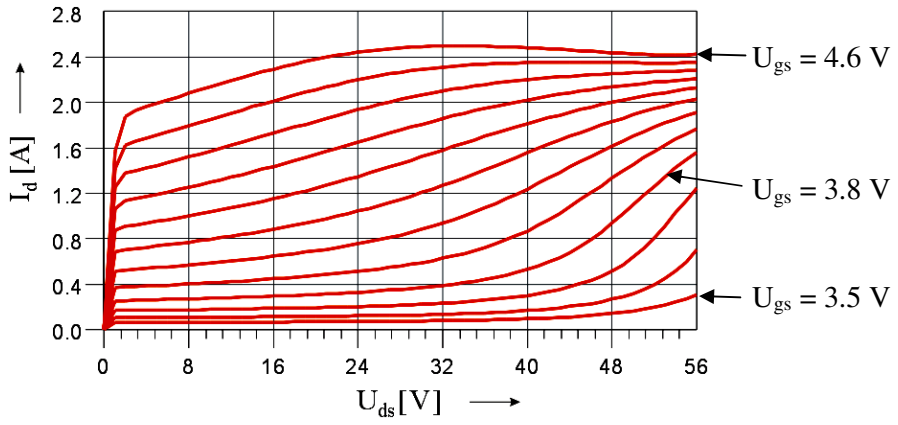


Fig. 4.2. The simulated output characteristic of MRF21030SR3 for U_{gs} between 3.5 V and 4.6 V.

Obviously, for a class AB power amplifier the gate voltage should be tuned between 3.5 V and 4.6 V, approximately. The simulated output characteristic, which denotes the relationship between the drain current I_d and the drain-source voltage U_{ds} for U_{gs} between 3.5 V and 4.6 V, is shown in Fig. 4.2. In this work, the optimum gate voltage is fixed by the compromise between the linearity and the efficiency after lots of simulations, and under the condition that the expected maximum output power must be accomplished. Finally, the optimal operating points of $U_{ds} = 26\ V$ and $U_{gs} = 3.8\ V$ are selected.

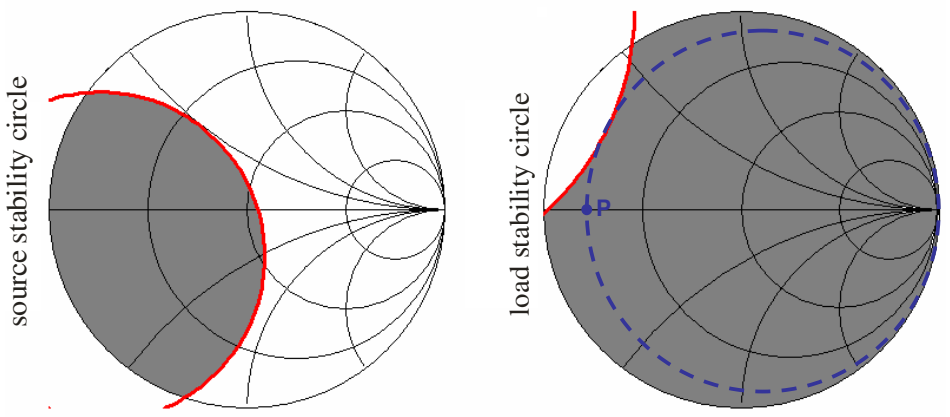
4.1.3 Advanced Stability Improvement

As mentioned above, almost all the transistors are potentially unstable within a large bandwidth. Therefore, it is important to research the stability characteristic of the used active device, before the matching network is developed. In case that the unconditional stability is not available by the transistor itself, actions must be taken in order to improve the conditional stability as far as possible.

The stability characteristic of MRF21030SR3 can be simulated with the S-parameter and be demonstrated in the Smith charts. In Fig. 4.3 (a), the simulated input and the output stability circle, which here are also named as source and load stability circle respectively, are presented. The simulated magnitudes of the S-parameters are also listed in it. The frequency, at which the simulation is done, is 2140 MHz, just the center frequency of the UMTS band. The grey areas in the Smith charts illustrate the stable regions. Obviously, this transistor is not unconditional stable at this frequency. Therefore, effort must be devoted to improve the conditional stability.

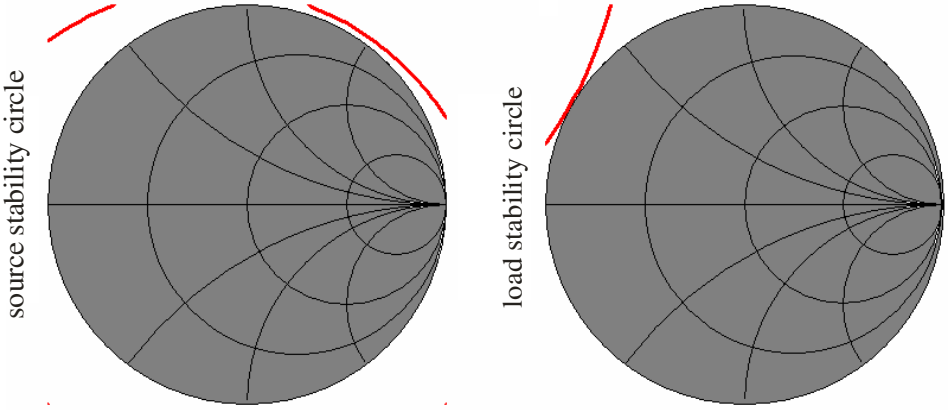
In Fig. 4.3 (a), a dotted circle is drawn to be tangent to the load stability circle. It crosses the real axis at the point P, which indicates the resistance of $R = 6.5\ \Omega$ in the Smith chart. Obviously, by adding a series resistance of 6.5 Ohm, all the impedance points located in the instable region in this Smith Chart will be shifted into the dotted circle, which is just entirely included in the stable region. Therefore, a series resistor of 6.5 Ohm directly connected at the output port of the transistor can improve the stability of the power amplifier circuit dramatically. Fig. 4.3 (b) shows the resimulated results for the stability with such a resistor at the drain of the transistor. Apparently, it now becomes unconditionally stable at the frequency of 2140 MHz.

There are altogether four methods of resistive loading to improve the circuit stability [34]. They are series or parallel resistors connected directly at the input or at the output of the transistors as shown in Fig. 4.4. Just like the example above, the functional principle of all the four methods is to shift the transformed input and output impedance into the stable region by using the additional resistors. However, the mostly used method in practice is a parallel resistor connected at the output of the power device. A series resistor connected at the input or at the output of the transistor decreases the power gain dramatically, and a parallel resistor connected at the input of the transistor can produce a significant deterioration in the noise performance of the power amplifier. Therefore, the preferred method is just to connect a resistor at the device output in parallel as shown in Fig 4.4 (d). However, this method has also its disadvantages. Firstly, the parallel resistor consumes additional DC power; therefore, the power added efficiency is greatly decreased. Secondly, to ensure the high frequency performance of the power amplifier, resistors only with very small dimensions can be applied in the circuit, for instance, resistors as surface mount device (SMD) should be used. However, such devices can easily be destroyed by a large current resulting in the instability of the power amplifier again.



frequency	mag.(S_{11})	mag.(S_{12})	mag.(S_{21})	mag.(S_{22})
2140 MHz	0.427	0.06	6.872	0.979

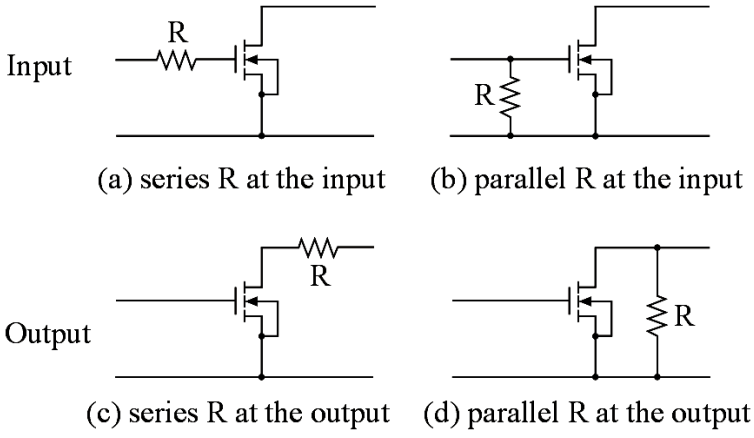
(a)



frequency	mag.(S_{11})	mag.(S_{12})	mag.(S_{21})	mag.(S_{22})
2140 MHz	0.45	0.053	6.037	0.762

(b)

Fig. 4.3. The stability improvement with series resistor simulated at 2.14 GHz.
(a) without the series resistor at the output; (b) with the series resistor at the output.



(a) series R at the input (b) parallel R at the input

(c) series R at the output (d) parallel R at the output

Fig. 4.4. Four methods of resistive loading for the stability improvement.

A modified configuration [70] for the stability improvement is introduced in this work by adding a series SMD capacitor C to the shunt resistor as shown in Fig 4.5.

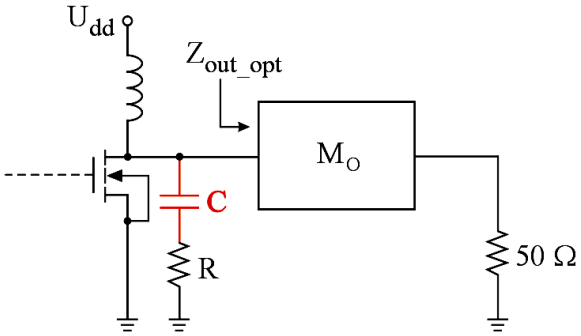
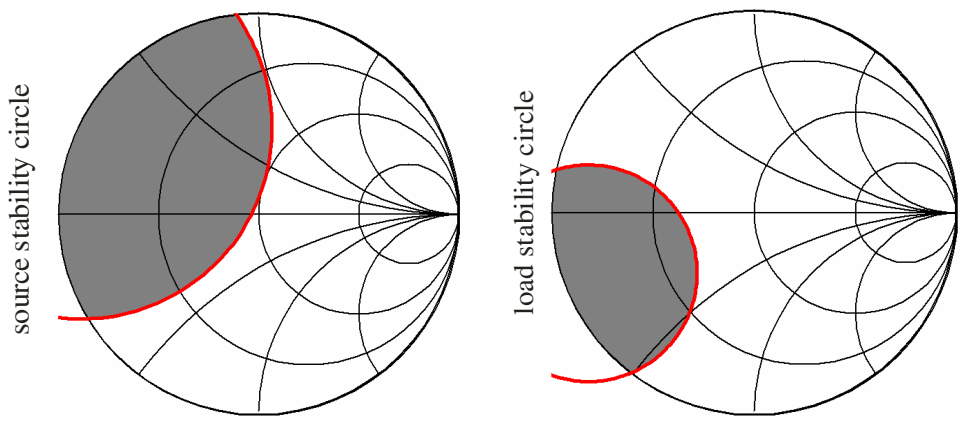
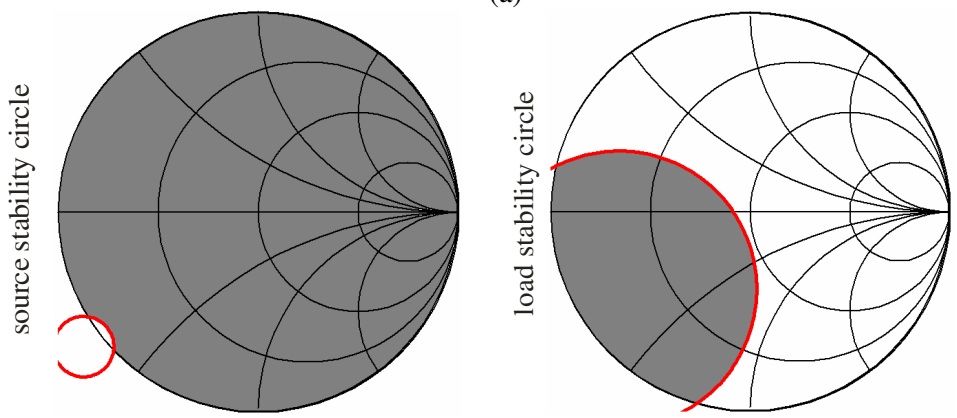


Fig. 4.5. The advanced method for the stability improvement.



frequency	mag.(S_{11})	mag.(S_{12})	mag.(S_{21})	mag.(S_{22})
1850 MHz	1.048	0.033	3.337	1.004

(a)



frequency	mag.(S_{11})	mag.(S_{12})	mag.(S_{21})	mag.(S_{22})
2140 MHz	1.005	0.026	2.589	0.686

(b)

Fig. 4.6. The stability simulated at 1.85 GHz.
(a) original stable region; (b) improved stability with shunt resistor and series capacitor.

The series capacitor prevents a DC current flowing through the shunt resistor R, and reduces the total DC power consumption; so that a higher power added efficiency can be guaranteed. Moreover, the series capacitor protects the shunt resistor against a large current ensuring the stability improvement obtained by it.

Since the active device MRF21030SR3 is recommended by the manufacturer only for the power amplifier design used in the UMTS band, it is more important to observe, how stable this device is in the downlink frequency band of the GSM1800 system. Using ADS2003C, the source stability circle and the load stability circle of the LDMOS transistor are simulated at 1850 MHz, the center frequency of the GSM1800 system. The simulation results are shown in Fig. 4.6 (a). According to the rules introduced in section 2.7.2 and the magnitude of the S-parameters listed below, the stable regions are determined to be inside the stability circles. They are depicted with grey in the Smith charts. Obviously, the stable regions only occupy small areas in both Smith charts. Therefore, a 50 Ohm shunt resistor with an additional series capacitor of 47 pF is directly connected at the drain of the transistor, in order to improve the stability. Fig. 4.6 (b) illustrates the new simulation results, where it is clear to see, that the stable regions both for the source and for the load are greatly enhanced. The other benefits of this structure, such like the enhancement of power added efficiency, will be shown later in the large signal simulation.

4.1.4 Design of the Matching Networks and the S-parameter Simulation

The optimum input and output impedance, namely Z_{in_opt} and Z_{out_opt}, of the power device MRF21030SR3 measured in the UMTS band, are given in Table 4.1. These values are determined using the load-pull measurement introduced in section 2.1 and are provided by the manufacturer Motorola. The 50 Ohm reference impedance from the source and the load should be transformed to these values, respectively. Obviously, the optimum output impedance is in the entire frequency band well below 50 Ohm, the reference impedance. The typical characteristics of the LDMOS transistors is shown here again.

Table 4.1
Optimum input and output impedance of MRF21030SR3

f (MHz)	**Z_{in_opt}** (Ohm)	**Z_{out_opt}** (Ohm)
2110	15.3 - j9.4	3.7 - j0.78
2140	14.6 - j9.4	3.4 - j0.37
2170	14.3 - j8.8	3.0 + j0.13

In order to achieve a broadband power amplifier, several different impedance matching networks are proposed in this work. The first one is carried out with the lumped elements, i.e. with the discrete inductors and capacitors as introduced in subsection 2.5.1. Since the real part of Z_{in_opt} and Z_{out_opt} are smaller than 50 Ohm, the L-matching network shown in Fig. 2.30 (a) is used both for the input and for the output matching networks. In the second attempt, matching networks consisting of an open-circuited stub followed by a series microstrip transmission line shown in Fig. 2.34 (a) are selected. The third possibility to realize the input and output matching networks applied in this work is using a single-section transmission line as described in subsection 2.5.5. The S-parameters of the power amplifier circuits with these matching networks are simulated, whereas the 3-dB bandwidth can be found out in the S_{21} curve. The simulation results show that the bandwidth obtained by using these three matching networks is not large enough to simultaneously cover the in UMTS and GSM1800 systems in Europe.

Finally, the multi-section transmission lines are used as microstrip matching networks both for the input matching and for the output matching. There are lots of different possibilities for the multi-section impedance transformer. For instance, both the input and the output matching networks use a two-section impedance transformer, respectively, or both of them use a three-section impedance transformer, etc. Also the combinations of two- and three- section impedance transformer can be adopted. Moreover, for example, only for a simple two-section impedance transformer design, there are also many different alternatives; since arbitrary inter-impedance point can be selected.

To optimize the performance of the power amplifier circuit, lots of matching networks are attempted in this work, until the satisfactory results are achieved. All the different transformers can first be designed in LLsmith, in order to obtain the characteristic impedance of each section transmission line and the length of them in form of percentage of the wavelength. Furthermore, the length of the transmission lines can be described in form of millimetre by using equations (2.53) and (2.54). Observed from these equations, the phase velocity v_p is slightly dependent on the width of the transmission lines. Different v_p, and hence different wavelength for each line, can separately be determined using these two equations. These different matching networks are one by one adopted in the power amplifier circuit for the S-parameter simulations in ADS2003C. Usually, the calculated design parameters can not directly provide the best simulation results. To fulfil all the circuit specifications, optimisations are always necessary. Finally, a combination of matching networks, which contains a three-section impedance transformer at the input and a two-section impedance transformer at the output, is regarded as the best selection for this work. It must be emphasised that this final result shown here is obtained considering not only the 3-dB bandwidth, but also the other specifications, such like the stability, the maximum output power, the linearity and the power added efficiency. The schematic of this final power amplifier circuit is demonstrated in Fig. 4.7.

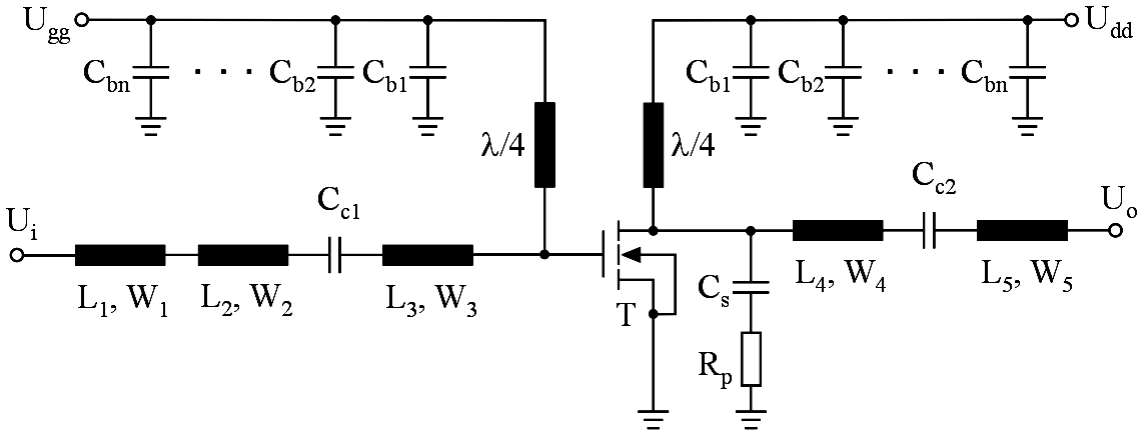


Fig. 4.7. The schematic of the broadband power amplifier using multi-section impedance transformer for the input and output matching networks.

In Fig. 4.7, T indicates the single transistor MRF21030SR3 used as the power device. The bias voltages for this single-ended power amplifier are U_{gg} = 3.8 V and U_{dd} = 26 V. As introduced in section 2.6, two sections of quarter-wave transmission line have separately been applied at the gate and at the drain of the transistor as biasing networks, which present open circuits to the RF signal at the fundamental frequency and simultaneously short the second harmonic. Both of them have a line-width of only 0.8 mm indicating a high characteristic impedance of about 80 Ohm on the RO4003 substrate. According to the data sheet of this substrate, a current of 3.2 A can be sustained by such a line with the width of 0.8 mm, if it has

a conductor-thickness of 35 µm. As shown in Fig. 4.2, the drain current is only about 0.35 A at the selected operating point. However, it can dramatically be enhanced with an increased RF input power. Therefore, the quarter-wave transmission line, especially used at the drain should be sufficiently wide. Moreover, these lines can also not be too narrow due to some mechanical structures, e.g. the bypass capacitors C_{b1} – C_{bn} must be soldered between these lines and the ground area. Therefore, these lines must be wide enough for the soldering. In this work, the bypass capacitors are selected from 1 pF to 100 nF. The parameters L_x and W_x in Fig. 4.7 denote the length and width of each section of transmission line, where x is an integer between 1 and 5. The design parameters are listed in Table 4.2. A shunt resistor of 50 Ohm connected with a series capacitor of 47 pF is used at the drain of the transistor, in order to improve the stability of the circuit.

Table 4.2
Design parameters of the broadband power amplifier

L_1	W_1	L_2	W_2	L_3	W_3	L_4
18 mm	3 mm	14 mm	4.5 mm	10 mm	13.6 mm	15 mm
W_4	**L_5**	**W_5**	**C_{c1}**	**C_{c2}**	**C_S**	**R_p**
10.2 mm	14 mm	3 mm	47 pF	47 pF	47 pF	50 Ohm

The simulation results of the S-parameters for this power amplifier in the frequency band of 1 GHz – 3 GHz are shown in Fig. 4.8. A maximum small-signal gain of about 12.5 dB is accomplished, while the 3-dB bandwidth reaches 1 GHz. Obviously, a broadband power amplifier is realized, whose bandwidth covers simultaneously the UMTS band and the GSM1800 band in Europe. The input matching and the output matching of this amplifier circuit indicated by the magnitude of S_{11} and S_{22} are mostly located between 0 dB and -10 dB.

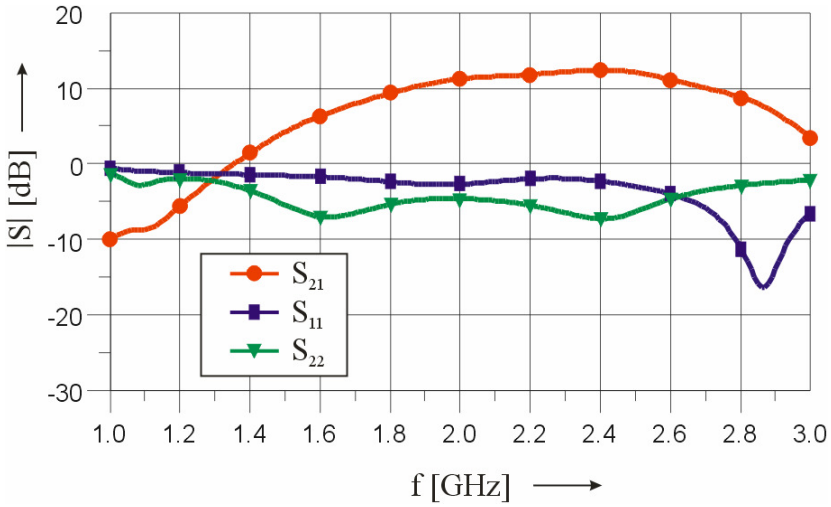


Fig. 4.8. Simulated S-parameters of the power amplifier with matching networks realized with multi-section impedance transformer.

The unconditional stability factor µ is first simulated in the frequency range and shown in Fig. 4.9. The stability curves for the following three cases are presented here:

1. circuit without any resistive loading
2. circuit with a 50 Ω shunt resistor connected at the drain of the device.
3. circuit with a 50 Ω shunt resistor and a 47 pF series capacitor at the drain of the device.

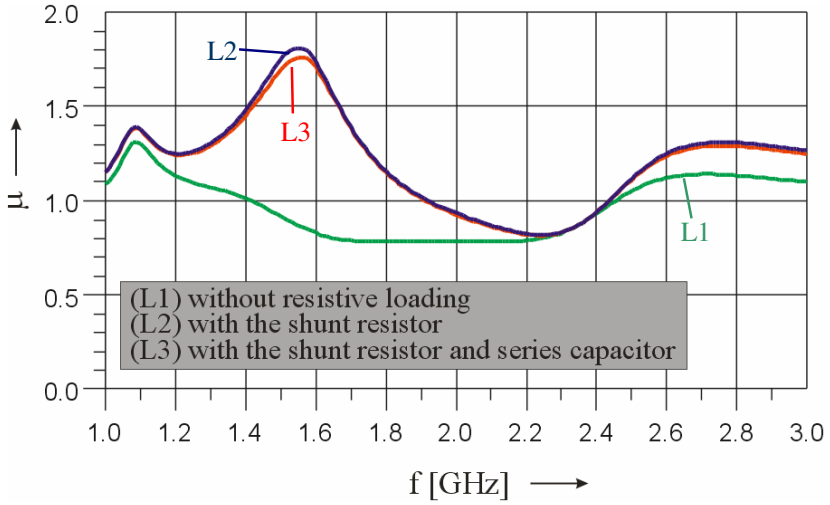


Fig. 4.9. Simulated unconditional stability factor μ versus the frequency.

Obviously, with a shunt resistor, the stability in the GSM1800 band is greatly improved. The stability improvement is hardly interrupted in case that a sufficiently large series capacitance is used. However, the condition that $\mu > 1$ is not achieved in the whole frequency band, especially in the UMTS band. This circuit is not unconditionally stable. This assessment can also be confirmed in Fig. 4.10, where the stability circles in the frequency range between 1 GHz and 3 GHz are demonstrated. Simulation shows that the magnitudes of S_{11} and S_{22} remain smaller than 1 in the entire frequency range.

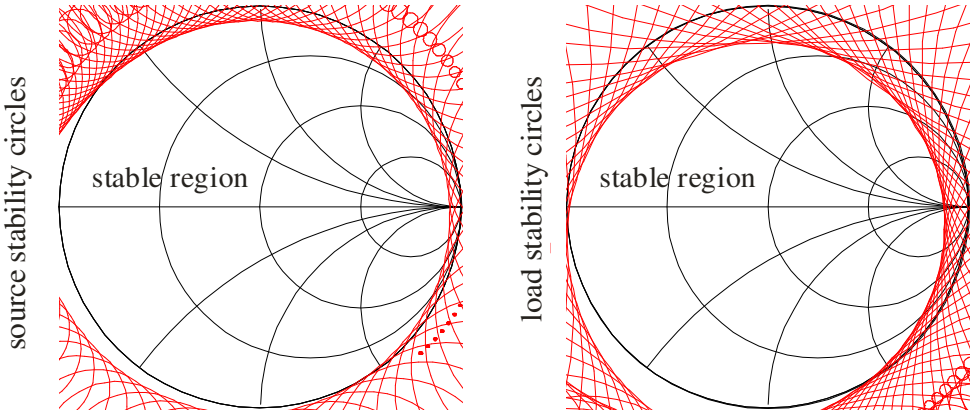


Fig. 4.10. Simulated stability circles versus frequency.

Altogether, the multi-section impedance transformer is the single effective tool to realize the matching networks for a broadband power amplifier. Using an adequate resistive loading, especially a shunt resistor at the output of the active device, it is feasible to design a LDMOS power amplifier being conditional stable, even in a very large frequency band.

4.1.5 Simulation for Large Signal Response of the Broadband Power Amplifier

In this work, the large signal behaviours, such as the maximum output power, the large signal power gain and the power added efficiency, are simulated. For a broadband power amplifier, it is not only important to see whether the proposed power amplifier circuit

provides a relatively constant small signal gain in a wide frequency range. It is for instance also important to observe whether large output power can be accomplished in different frequency bands. In this subsection, several large signal simulations for the broadband power amplifier shown in Fig. 4.7 are presented.

Using harmonic balance simulation of ADS2003C, the power transfer functions, which indicate the relationship between P_{in} and P_{out}, are simulated both in the UMTS band and in the GSM1800 band. The simulation results are shown in Fig. 4.11. A maximum output power of about 45.5 dBm is obtained at 2140 MHz, with a 1-dB compression point of about 45 dBm. The 6 dB back-off point is then fixed at the point of P_{out} = 39 dBm, correlating with the input power of 27 dBm. The maximum output power achieved at 1850 MHz is about 44.5 dBm, with a 1-dB compression point of about 43.7 dBm. Obviously, though the maximum output power and 1-dB compression point at 1850 MHz are a little bit lower than those at 2140 MHz, reasonable large output power are obtained in both frequency bands. Using these simulation results, the large signal power gain versus P_{in} obtained at these two different frequencies is calculated and presented in Fig. 4.11. They are saturated in the higher input power levels. Similar to the simulation results shown in Fig. 4.8, the power gain in the GSM1800 band is slightly lower than that in the UMTS band.

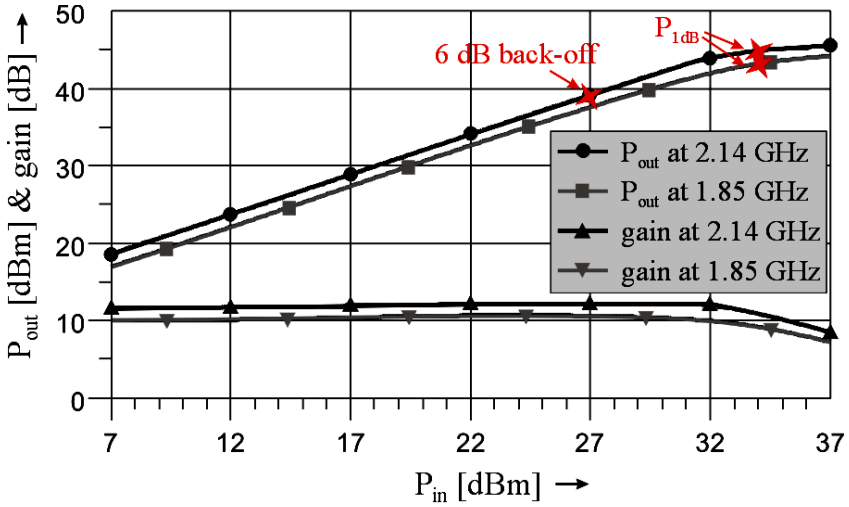


Fig. 4.11. Simulated power transfer function and large signal power gain.

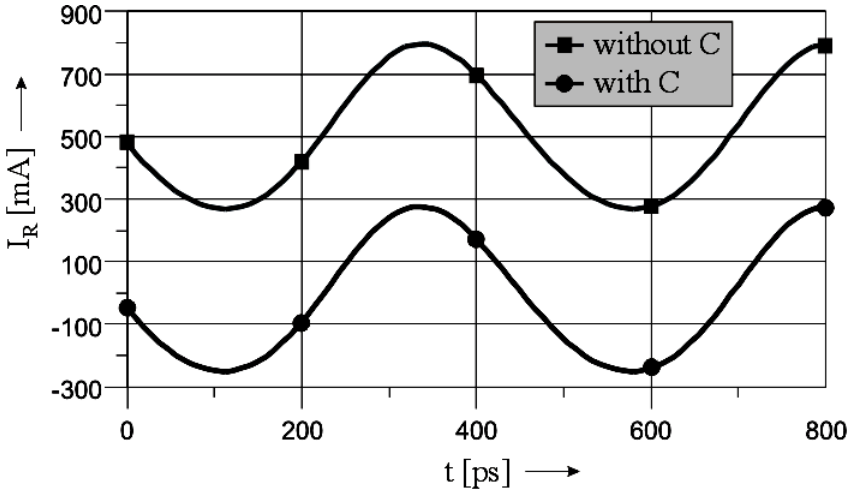


Fig. 4.12. Simulated waveform of the current flowing through the shunt resistor.

As mentioned in subsection 4.1.3, a series capacitor and a shunt resistor are directly connected at the drain of the LDMOS transistor in order to improve the stability. Fig. 4.12 depicts the simulated waveform of the current I_R, which flows through the shunt resistor R_p with or without the capacitor C_s. It is simulated in case that the power amplifier circuit has an output power of 40 dBm. Obviously, using the series capacitor C_s the DC current having a magnitude of about 0.52 A is deleted.

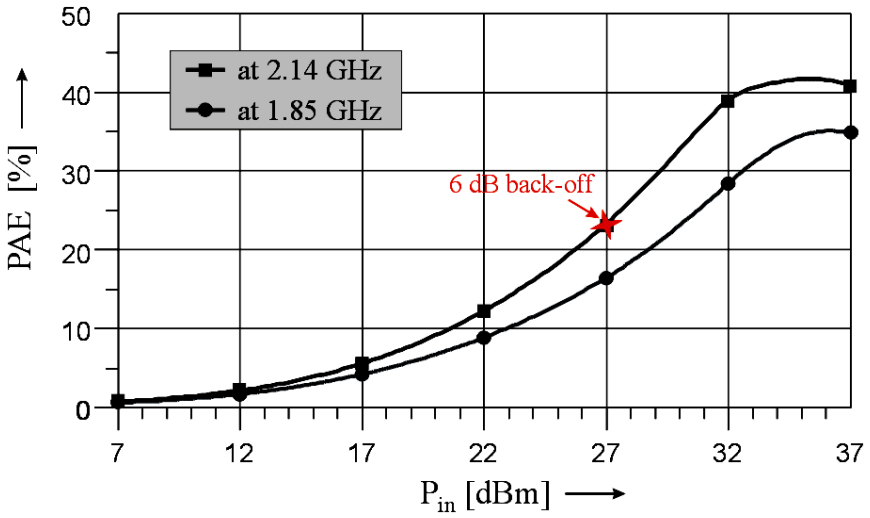


Fig. 4.13. Simulated power added efficiency of the single-ended power amplifier.

The simulated power added efficiency depending on the input power is shown in Fig. 4.13. Simulation is also done at two different frequencies. A maximum PAE of 41 % is achieved at 2140 MHz, and 35 % at 1850 MHz. To meet the high requirement of linearity, power amplifiers of CDMA transmitters usually operate 6 dB backed off from the saturation. Therefore, it is sometimes more interesting to observe the power added efficiency obtained at the 6 dB back-off point. As analysed above, the input power related 6 dB back-off point is located at 27 dBm. A PAE of 22 % is obtained at this point.

4.1.6 Experiments and Measurement Results

According to the circuit schematic introduced in Fig. 4.8 and the design parameters given in Table 4.2, a broadband single-ended class AB LDMOS power amplifier is fabricated using the Motorola LDMOS transistor MRF21030SR3. As mentioned in the design process, this power amplifier is built on a 0.81 mm-thick RO4003 substrate, which has a relative permittivity of 3.38 and a conductor thickness of 35 µm. The photograph of this circuit is presented in Fig. 4.14.

In this circuit, straight medium adaptors (SMA) are applied as the input port and the output port, which have a coaxial structure and perform perfect RF closeness. Just the same as the simulation, the bias voltages used for the measurements are U_{gg} = 3.8 V and U_{dd} = 26 V. Obviously, multi-section transmission-lines are used both at the input and at the output of this circuit, in order to yield a broadband matching. By reason of symmetry, the shunt resistor of 50 Ohm connected with a series capacitor of 47 pF, which is used in the simulation for the stability-improvement, is replaced by two parallel shunt resistors, each of which is 100 Ohm and connected with a series capacitor of 47 pF. They are separately soldered on both sides of the conductor line, which is directly connected to the drain of the LDMOS transistor.

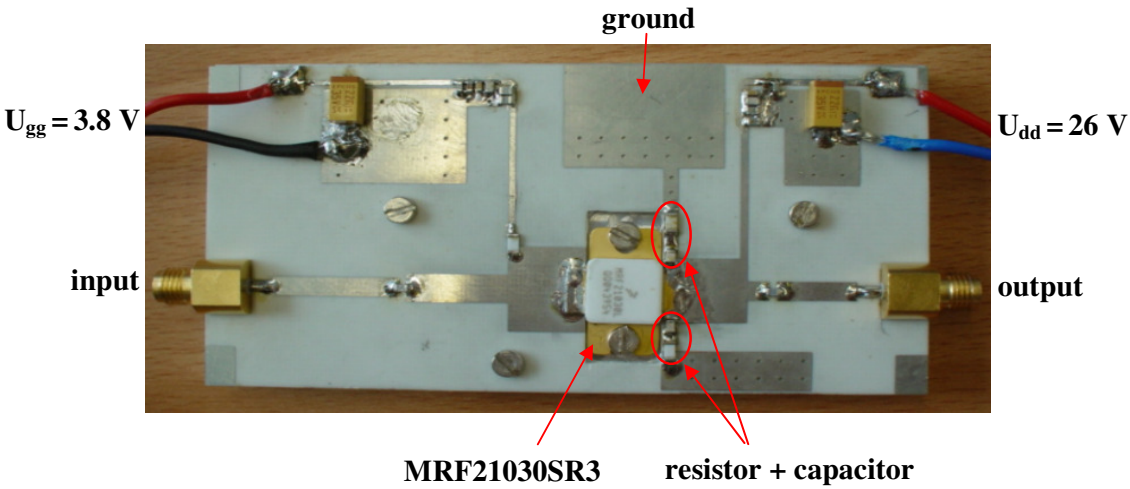


Fig. 4.14. The photograph of the fabricated single-ended power amplifier.

Using *HP 8510B Network Analyzer*, the S-parameters of the amplifier circuit are measured. The measurement results are shown in Fig. 4.15. Obviously, a gain of about 13 dB is achieved in a large frequency range. This power amplifier circuit has a 3-dB bandwidth of 1 GHz with the centre frequency of 2.1 GHz. In terms of the bandwidth, there is a good agreement between the simulated and the measured results. The difference is that the flatness of the measured S_{21} curve is not as good as that of the simulated result and all the curves are shifted to the lower frequency range about 200 MHz. However, just due to this shift, both the UMTS band and the GSM1800 band are perfectly covered by the bandwidth of this power amplifier. Other than the simulated results, the measured gain in the GSM1800 band is a little bit larger than that in the UMTS band. Moreover, it can be seen that the input and the output matching of this circuit are not as desired lower than -10 dB in the entire frequency range. This is a typical feature obtained for a single-ended power amplifier, since the optimum input and output impedance provided by the manufacturer concentrate mainly on the maximum output power, not on the VSWR.

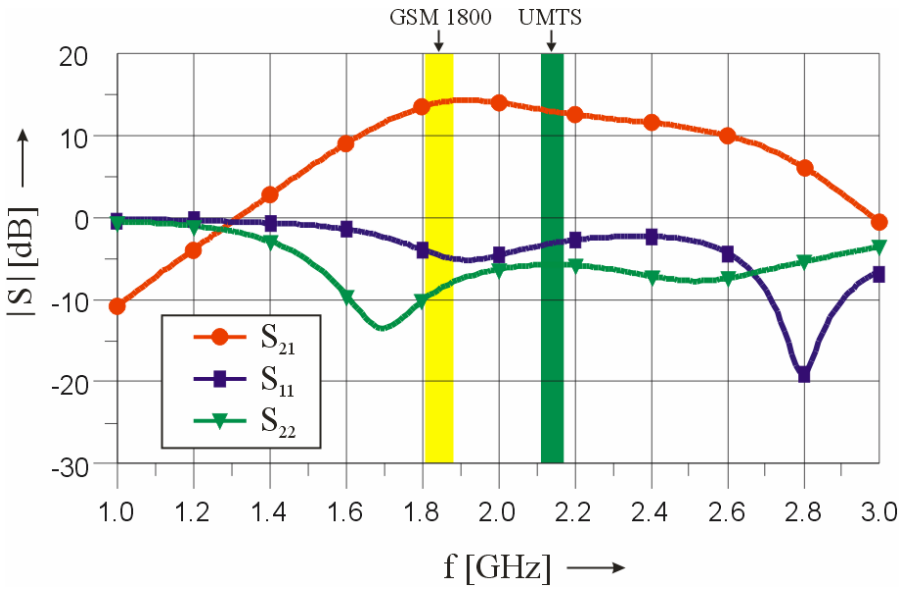


Fig. 4.15. Measured S-parameters of the single-ended power amplifier.

Using *HP 8664A Synthesized Signal Generator* and *Agilent E4419B Power Meter*, the output power P_{out} of the power amplifier versus the input power P_{in} is measured, not only at 1.85 GHz but also at 2.14 GHz. The measurement results are presented in Fig. 4.16.

Obviously, almost the same maximum output power of about 43.5 dBm (22.4W) is achieved at both frequencies, while the 1-dB compression points are located at 42.7 dBm. The power gain versus P_{in} can be calculated and shown in the same figure. A linear gain of about 13 dB is also obtained during this measurement. Also in the large signal measurement, the power gain obtained in the GSM1800 band is slightly higher than that in the UMTS band.

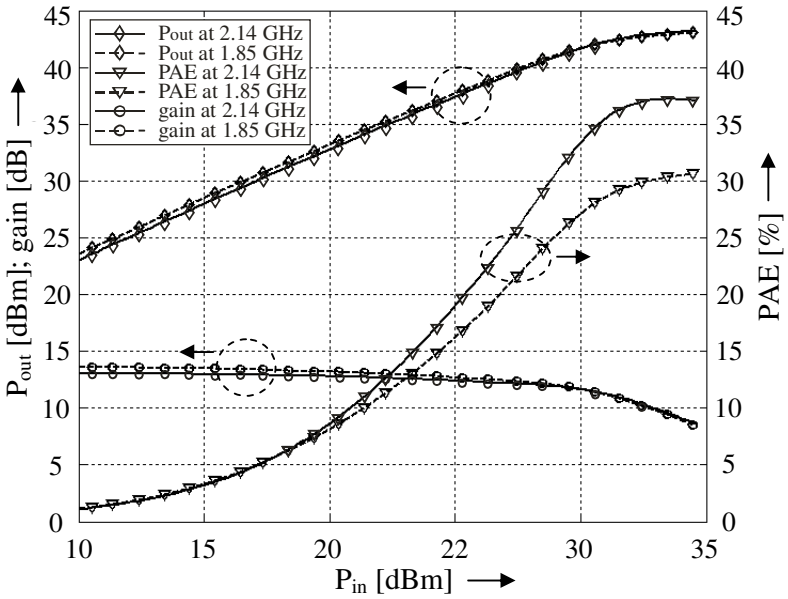


Fig. 4.16. Measured output power, power added efficiency and power gain versus the input power.

The power added efficiency is also measured and shown in Fig. 4.16. The maximum power added efficiency achieved at the frequency of 1.85 GHz is 30.5 % for the maximum output power. At the frequency of 2.14 GHz, the power added efficiency is as high as 37.2 % for the maximum output power. Though the same maximum P_{out} are obtained at the two frequencies, the maximum PAE are quite different, since the DC power consumed by the transistor are not the same at these two different frequencies. In case that a RF signal is fed at the input, the DC current flowing through the transistor is not equal to that shown in Fig. 4.2, but much larger; and it doesn't remain the same if the frequency changes. Furthermore, at the frequency of 2.14 GHz PAE of 18 % is obtained at the 6 dB back-off point from the 1-dB compression point, where P_{out} is equal to 36.7 dBm. The measured values of the output power and PAE are all somewhat lower than those of the simulations, but in a reasonable range.

Using *MS 2668C Spectrum Analyser*, the third order intermodulation distortion of this single-ended power amplifier is measured with 1 MHz offset frequency. The two input signals are generated by using *HP 83650B* and *HP 8664A Synthesized Signal Generator*. Due to the limited output power of these signal generators, the lower and upper IMD3 versus the output power is only measured up to P_{out} = 40 dBm as shown in Fig. 4.17. At the frequency of 2140 MHz, IMD3 better than -35 dBc is obtained at the 6 dB back-off point.

Using *Anritsu MG3700A vector signal generator* and *Anritsu MS2781A signal analyzer*, the spectrum of the output signal is measured, whose results are shown in Fig. 4.18. The measurement was first implemented at 2.14 GHz for a 5 MHz single-carrier W-CDMA-signal with a peak to average ratio (PAR) larger than 10 dB, while the peak output power is about

39 dBm. The average power of - 4.25 dBm shown in the figure below is measured with 33 dB attenuation at the power amplifier output. Obviously, ACPR of about 45 dBc is obtained at this power level, with which the 3GPP ACPR requirement is met. There is a rule of thumb obtained from the practice, namely, at the power level, where the IMD3 is about 30 dBc, ACPR of 45 dBc can be approximately fulfilled. This rule is here again proven.

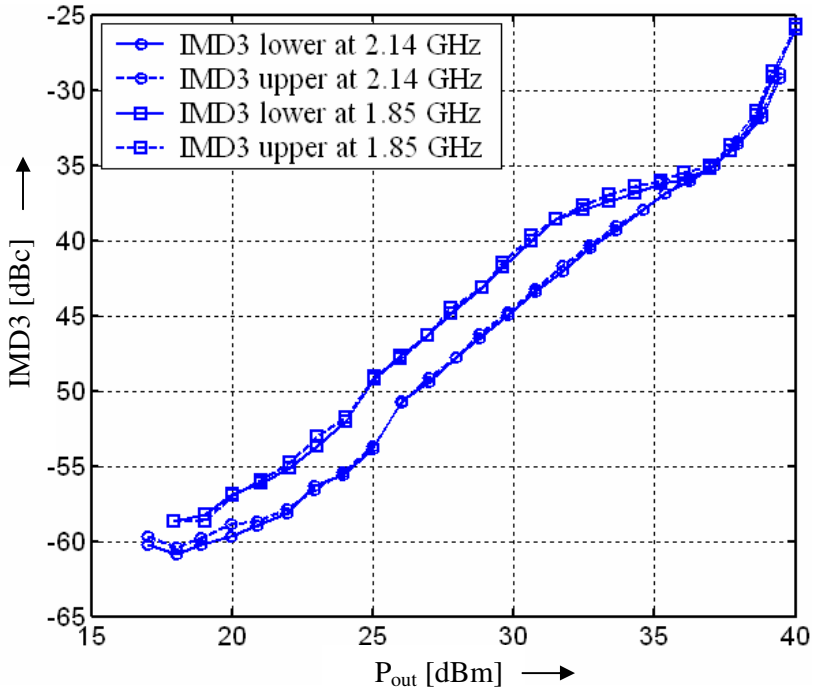


Fig. 4.17. Measured third order intermodulation distortion versus the output power.

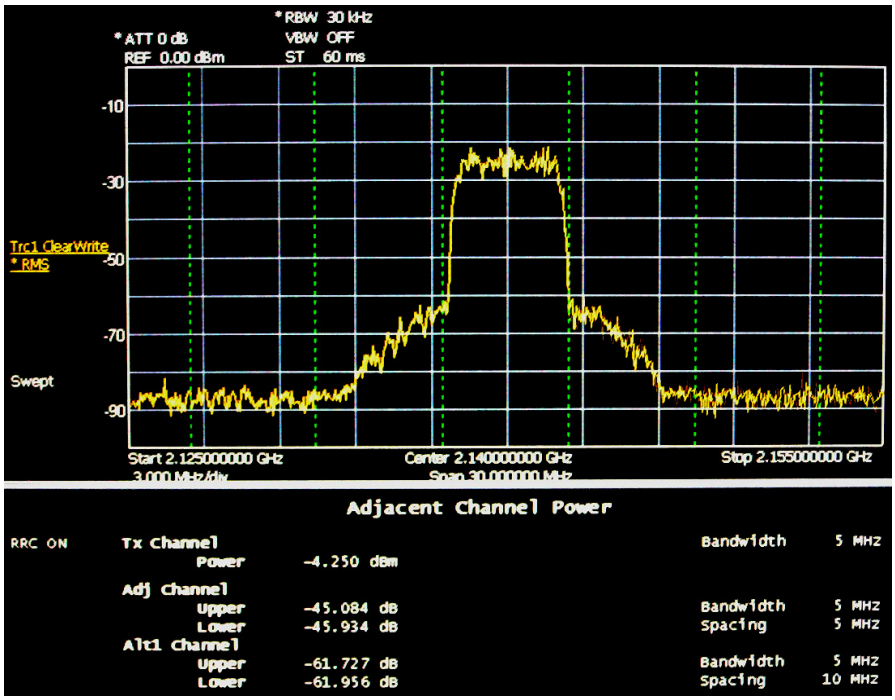


Fig. 4.18. Measured ACPR at 2.14 GHz and with 5 MHz offset as well as 39 dBm peak output power.

Furthermore, the measured ACPR versus the peak output power is given in Fig. 4.19. Clearly, ACPR performance is degraded with the increased power level. It is also evident that there is a slight difference between the ACPR lower and ACPR upper.

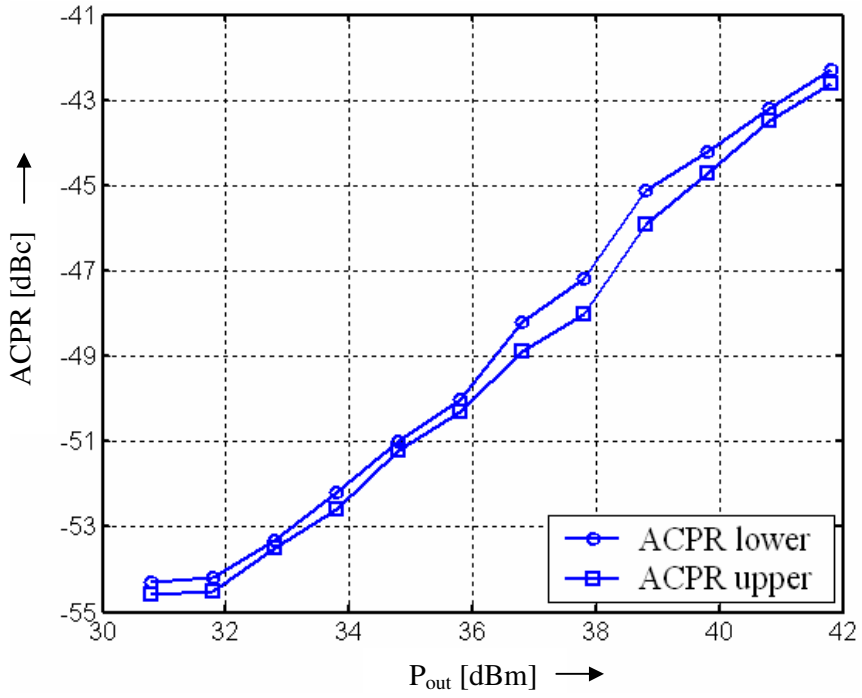


Fig. 4.19. Measured ACPR versus the output power of the single-ended power amplifier.

Comparisons are done between the power amplifier realized in this work and the other LDMOS power amplifiers achieved in recent years. Results of the comparisons are summarized in Table 4.3. Obviously, the power amplifier realized here has the largest bandwidth.

Table 4.3
Comparisons of LDMOS power amplifiers

Author	**Center Frequency [GHz]**	**Bandwidth [MHz]**	**P_{out} [dBm]**	**PAE [%]**	**U_{dd} [V]**
Cassan et al. [71]	3.5	550	46	36.7	28
Bagger et al. [72]	2.1	400	43	45	28
This work	2.1	1000	43.5	37.2	26

4.2 Design of a Broadband LDMOS Balanced Class AB Power Amplifier

Since the designed power amplifier will be connected with the other components as well as the antenna in the transmitter, it is demanded that the proposed power amplifier should have very low voltage standing wave ratio (VSWR) at its input and output, which indicate excellent cascading ability. Moreover, an even larger output power up to 50 W (47 dBm) is expected. Therefore, a balanced power amplifier is introduced in this work, which contains two power devices connected in parallel, providing the possibility for even higher maximum achievable output power.

In this section, the functional principle of a balanced structure is introduced. Using the same LDMOS transistors and the same matching networks obtained in the design of the broadband single-ended power amplifier, a broadband LDMOS balanced class AB power amplifier is developed. The simulation processes as well as the measurement results are demonstrated below.

4.2.1 Balanced Structure

The balanced structure, whose configuration is shown in Fig. 4.20, is attractive due to the very low VSWR and the excellent cascading ability. Two identical single-ended amplifiers T_1 and T_2 are connected in parallel through two quadrature hybrid couplers, which operate as a power divider at the input and a power combiner at the output, respectively.

Quadrature hybrid couplers are 3-dB directional couplers with a 90° phase difference in the outputs of the through and coupled arms [32]. Therefore, at the input of the balanced amplifier, the two reflected signals from the two individual amplifiers are 180° out of phase. They compensate each other, so that the balanced amplifier has ideally a VSWR of 1 at its input. The same operational mode is also valid for the two reflected signals at the output. Therefore, VSWR of a balanced amplifier depends on the coupler, not on each individual amplifier. The forward transmitted signals of the two individual amplifiers are again in phase at the output, due to the same length of their routes.

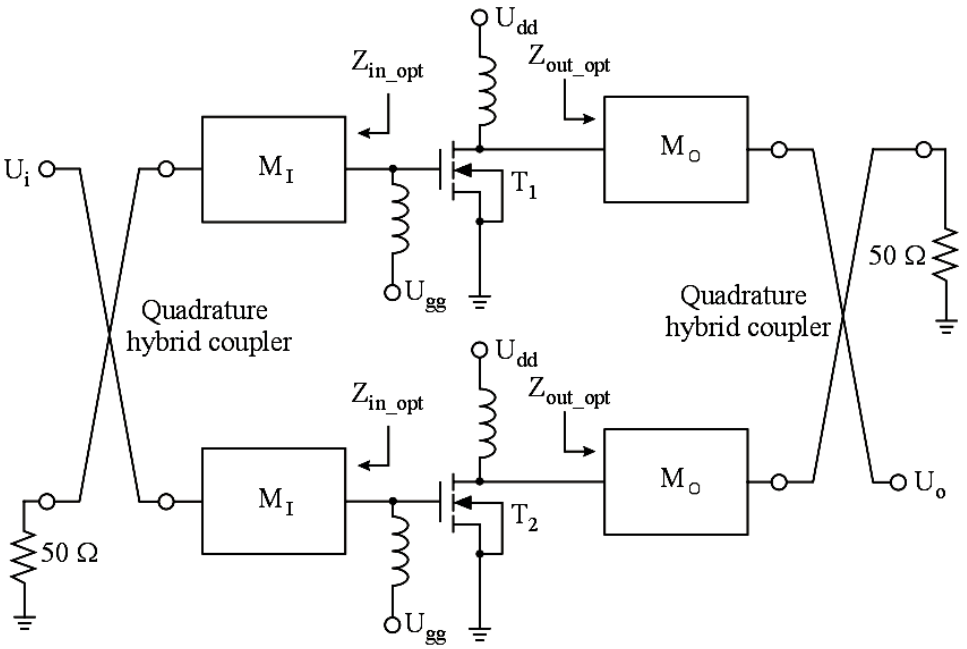


Fig. 4.20. The balanced structure.

The magnitudes of the S-parameters for a balanced amplifier can be given as follows [34]

$$|S_{11}| = 0.5 \cdot |S_{11T1} - S_{11T2}| \tag{4.1}$$

$$|S_{21}| = 0.5 \cdot |S_{21T1} + S_{21T2}| \tag{4.2}$$

$$|S_{12}| = 0.5 \cdot |S_{12T1} + S_{12T2}| \tag{4.3}$$

$$|S_{22}| = 0.5 \cdot |S_{22T1} - S_{22T2}|, \tag{4.4}$$

where S_{11T1}, S_{11T2} ··· S_{22T1} and S_{22T2} indicate the S-parameters of the individual amplifiers T_1 and T_2. It can be seen that the magnitude of S_{11} and S_{22} should be zero, if the amplifiers T_1 and T_2 are absolutely identical. Excellent stability can therefore also be expected. Moreover, in case of entire identity of two individual amplifiers, the gain of a balanced power amplifier remains same to that of a single-ended one, according to equation (4.2). The same is with the reverse transmission coefficient S_{12} as shown in equation (4.3). Though the gain of a balanced PA will not be doubled in spite of the use of two transistors, the maximally accomplishable output power of a balanced power amplifier is twice that of a single-ended amplifier.

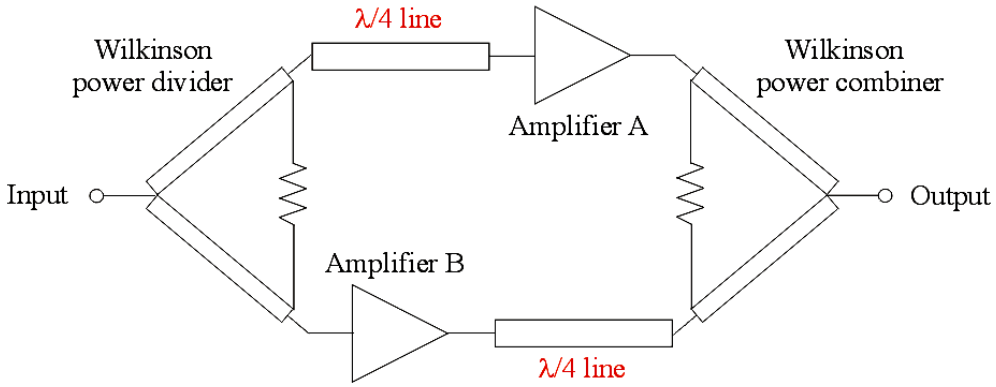


Fig. 4.21. The balanced structure using Wilkinson couplers.

A typical quadrature hybrid coupler can be realized with a branch-line coupler [32]. Because the input signal of a branch-line coupler is equally divided and further transmitted to the two output ports. There is a 90° phase difference between two output signals. In this work, however, Wilkinson couplers [32] are used as the power divider and as the power combiner. The schematic of a balanced structure using Wilkinson couplers is shown in Fig. 4.21. Since the two output signals of a Wilkinson divider are in phase, it is necessary to shift the phase of the input signal before one single-ended amplifier (e.g. T_1) and the phase of the output signal after the other single-ended amplifier (e.g. T_2) for 90°, respectively. Quarter-wave transmission-lines can be used for such phase-shift. The 180° out-of-phase of the two reflected signals at the input and at the output of the balanced amplifier is generated again. The in-phase signals at the output of the balanced amplifier, which are forward transmitted from the two individual amplifiers, can also be obtained in this manner.

4.2.2 Simulation of the LDMOS Balanced Power Amplifier

Similarly, using the MET-model of Motorola LDMOS device MRF21030SR3, a broadband balanced radio frequency power amplifier is simulated in ADS2003C. As mentioned above,

the same biasing networks and the same impedance matching networks as those of the single-ended power amplifier presented in Fig. 4.7 are also used in this design. The same operating points of U_{gg} = 3.8 V and U_{dd} = 26 V are also selected here for the two individual power amplifiers. Simulations for the S-parameters and for the large signal behaviours are implemented. The simulation results are shown below.

The simulated S-parameters are demonstrated in Fig. 4.22. A maximum small-signal gain of about 12.5 dB is accomplished. Using the multi-section impedance transformers for the input and the output matching networks, the proposed balanced power amplifier remains to be broadband. Both the UMTS and the GSM1800 bands in Europe are covered by its bandwidth. Compared to the single-ended power amplifier, the bandwidth of the balanced power amplifier is slightly diminished, because the gain in some frequency ranges is a little reduced due to the additional attenuation on the Wilkinson power couplers. Obviously, the magnitudes of S_{11} and S_{22} are in a large frequency range lower than -10 dB. The largest benefit of the balanced structure, namely very low VSWR, is clearly proven again in this simulation. Therefore, this balanced power amplifier has perfect cascading ability and the excellent stability.

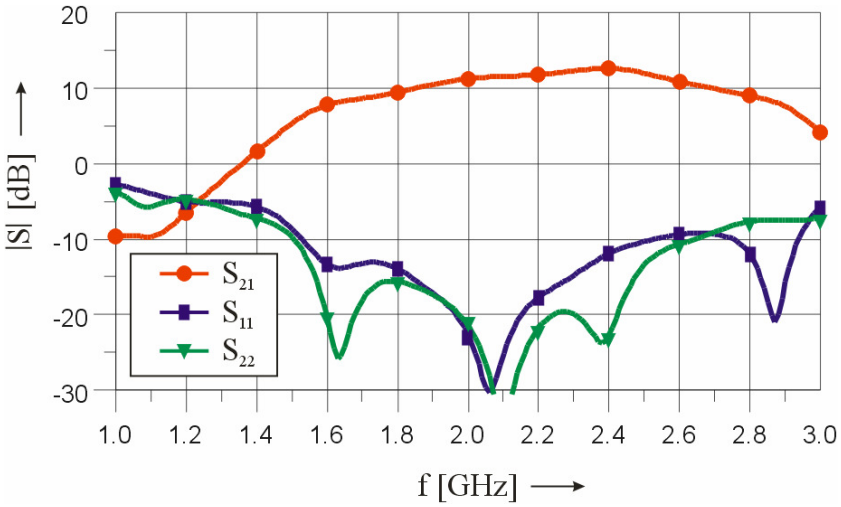


Fig. 4.22. Simulated S-parameters of the broadband balanced power amplifier.

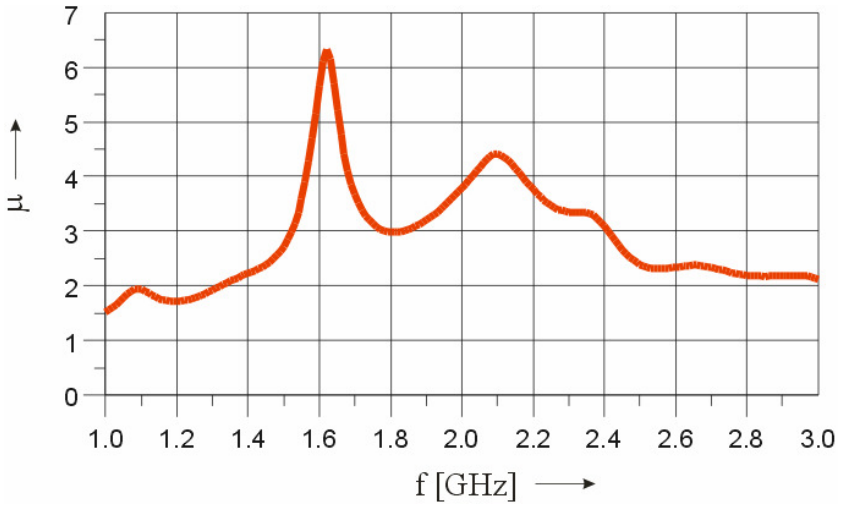


Fig. 4.23. Simulated unconditional stability factor μ of the balanced power amplifier.

The unconditional stability factor μ simulated in the frequency range between 1 GHz and 3 GHz is shown in Fig. 4.23. Obviously, it is in the whole frequency range even larger than 1.5. This balanced structure is therefore unconditional stable. However, an important point must be emphasised here. Though the balanced structure is unconditional stable, each individual single-ended power amplifier used in the balanced structure must also be stable. The methods for stability improvement introduced in section 4.1.3 are absolutely necessary.

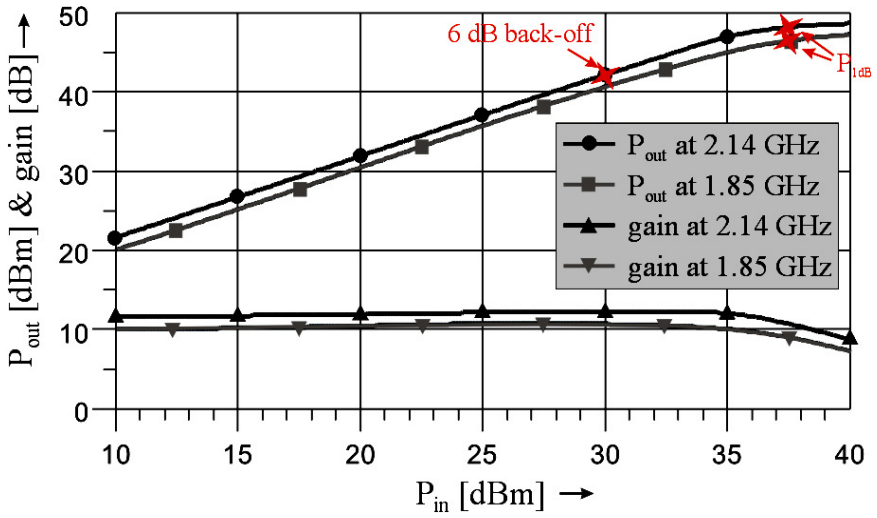


Fig. 4.24. Simulated power transfer function and large signal power gain of the balanced power amplifier.

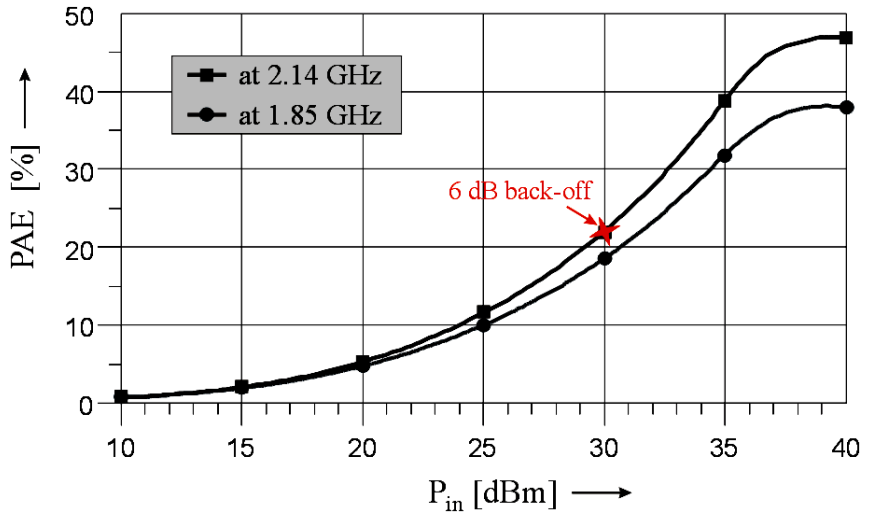


Fig. 4.25. Simulated power added efficiency of the balanced power amplifier.

The power transfer function of this proposed balanced power amplifier, which indicates the output power versus the input power, is simulated at two different frequencies. The simulation results are shown in Fig. 4.24. The maximum output power of about 48.6 dBm is obtained at the frequency of 2140 MHz, with a 1-dB compression point of about 48 dBm. The 6 dB back-off point is therefore as high as 42 dBm, as denoted in this figure. The maximum output power achieved at 1850 MHz is about 47.2 dBm, with a 1-dB compression point of about 46.6 dBm. Obviously, though the maximum output power and 1-dB compression point at 1850 MHz are slightly lower than those at 2140 MHz, reasonable large output power is

obtained in both frequency bands. This balanced power amplifier is proven to be useful both for the UMTS and for the GSM1800 systems in Europe indeed. Moreover, it is clear to see that the maximum output power of the balanced power amplifier is just about 3 dB higher than the simulated values of the single-ended power amplifier at both frequencies, since the two power devices are connected in parallel.

Applying the simulation results given above, the large signal power gain of this balanced amplifier can be calculated at the two different frequencies mentioned above. They are also presented in Fig. 4.24 related to the input power. Note that the gain of a balanced power amplifier is approximately the same as that of the single-ended one presented in subsection 4.1.6. Similar to the simulation results shown in Fig. 4.22, the large signal power gain in the GSM1800 band is slightly lower than that in the UMTS band. Both of them are saturated in the higher power levels.

The simulated power added efficiency versus the input power is shown in Fig. 4.25. Obviously, the maximum PAE reaches 47 % at 2140 MHz and 38 % at 1850 MHz. A PAE of 21.5 % is obtained at 2140 MHz and at the point of P_{in} = 30 dBm. This input power correlates with P_{out} = 42 dBm, the 6 dB back-off point as presented in Fig. 4.24.

4.2.3 Experiments and Measurement Results

A balanced power amplifier is fabricated using two Motorola LDMOS transistor MRF21030SR3. Similar to the single-ended amplifier, it is also fabricated on the RO4003 substrate. The photograph of the balance amplifier is shown in Fig. 4.26. The bias voltages are U_{gg} = 3.8 V and U_{dd} = 26 V. Obviously, Wilkinson couplers are used as the power divider at the input and as the power combiner at the output. Two λ/4-lines are used before the upper amplifier and after the lower amplifier respectively, in order to tune the signal phase as introduced above.

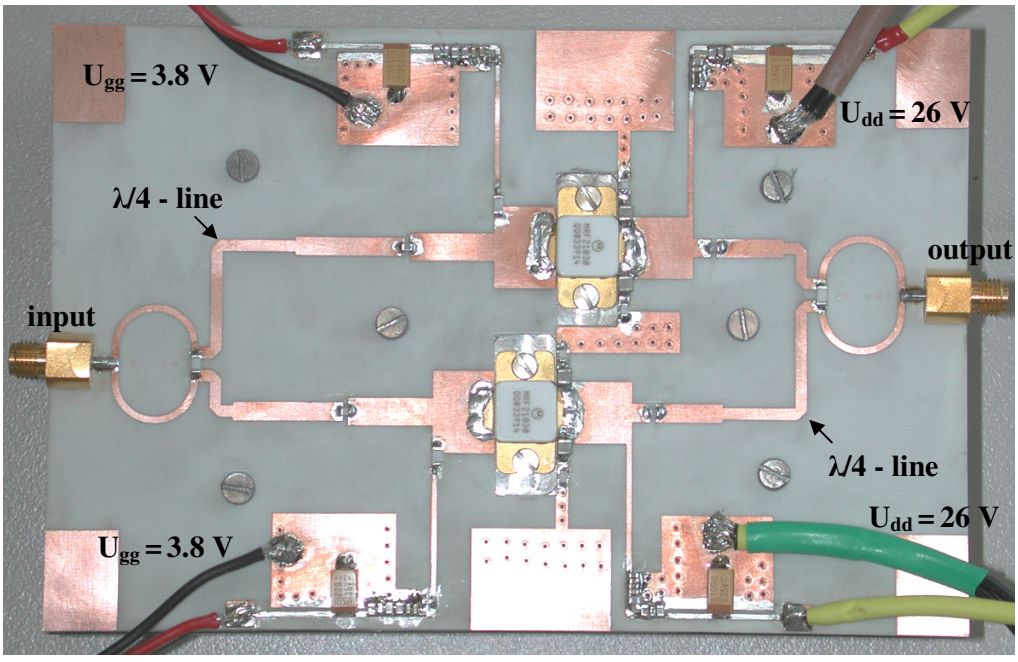


Fig. 4.26. The photograph of the fabricated balanced power amplifier.

Measured S-parameters are shown in Fig. 4.27. A linear gain of about 12 dB is obtained. This power amplifier has a 3-dB bandwidth of more than 830 MHz with the center frequency

of 2 GHz. Compared to the simulation results, the center frequency is shifted to the lower frequency range about 200 MHz and the gain is degraded in high frequency range due to the additional attenuation on the couplers. However, the obtained bandwidth is sufficiently large to simultaneously cover the UMTS and the GSM1800 down-link bands in Europe. Magnitudes of S_{11} and S_{22} of this power amplifier are in the frequency range of interest between -10 dB and -30 dB.

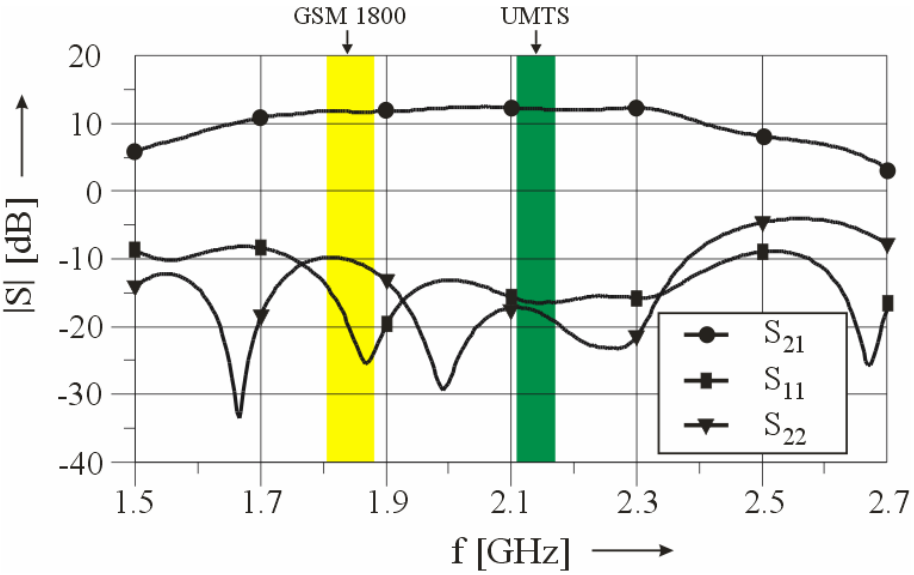


Fig. 4.27. Measured S-parameters of the LDMOS balanced power amplifier.

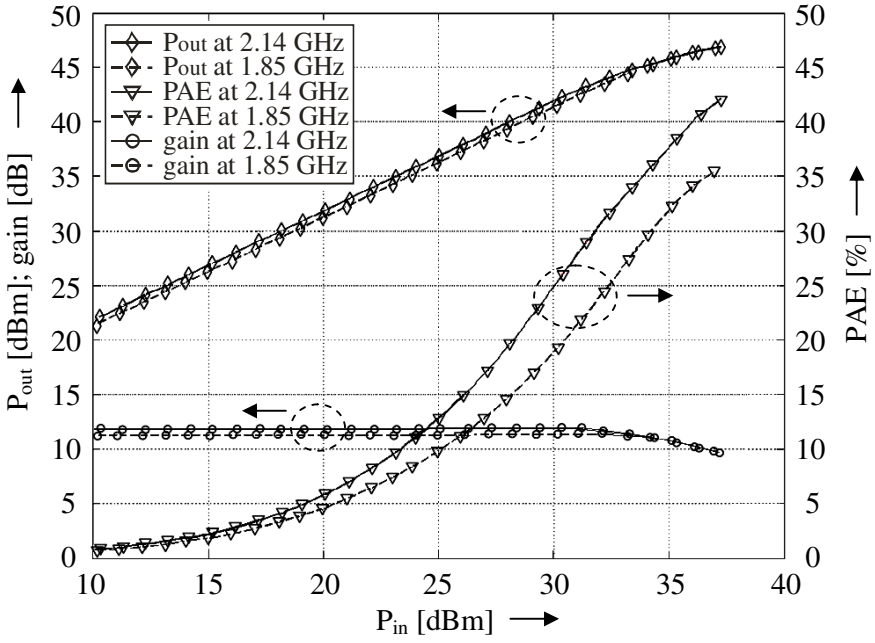


Fig. 4.28. Measured output power, PAE and power gain of the balanced power amplifier.

Fig. 4.28 presents the input-referred output power P_{out}, power added efficiency and the power gain measured at 1.85 GHz and 2.14 GHz, respectively. The maximum output power of 50 W (47 dBm) is accomplished at both frequencies, while the 1-dB compression points

are located at about 46 dBm. The power added efficiency of 35.5 % is achieved at 1.85 GHz for the maximum output power. At 2.14 GHz, the power added efficiency is as high as 42.5 % for the maximum output power. Even with the consideration of 6 dB back off from the 1-dB compression point, namely at the point of P_{out} = 40 dBm, a PAE of 20 % is obtained at this frequency. The measured values of the output power and the PAE are all lower than those of the simulations. However, all the requirements in terms of output power, the PAE and the power gain are fully fulfilled.

The third order intermodulation distortion versus the output power is measured with 1 MHz offset frequency. It is also measured in the UMTS and the GSM1800 bands, respectively. Due to the limited output power of the two signal generators, the measurements are only implemented up to P_{out} = 40 dBm, but this is just the 6 dB back off point in the UMTS band. The measurement results are shown in Fig. 4.29. Obviously, the third order intermodulation is degraded with the increased power level. However, IMD3 better than - 42.5 dBc at 1.85 GHz and than - 39 dBc at 2.14 GHz are obtained when the output power is lower than 40 dBm. High linearity is yielded for this balanced LDMOS power amplifier.

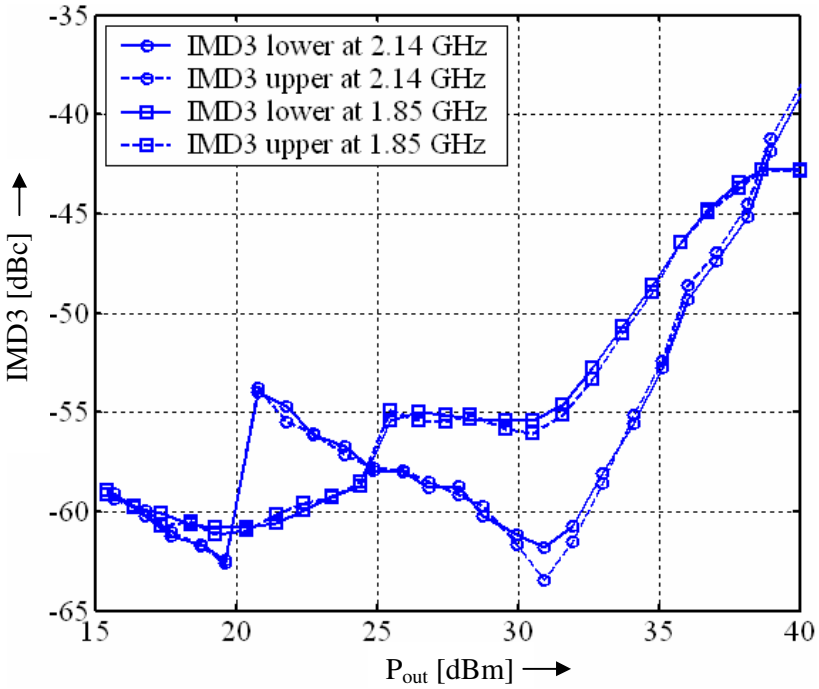


Fig. 4.29. Measured third order intermodulation distortion versus the output power.

The spectrum of the output signal is measured using a 5 MHz single-carrier W-CDMA-signal, which has a center frequency of 2.14 GHz and a peak to average ratio (PAR) larger than 10 dB. The measurement results are shown in Fig. 4.30. The measurement is first done in case that the peak output power is about 45.8 dBm. The average power of 2.87 dBm shown in the figure above is measured with 33 dB attenuation at the power amplifier output. Obviously, ACPR of about 45 dBc is obtained at this output power level, with which the 3GPP ACPR requirement is met.

Furthermore, the measured ACPR in dependence on the peak output power is presented in Fig. 4.31. It can be seen that this power amplifier performs high linearity in the whole power range. The difference between the ACPR lower and ACPR upper are as small as negligible. The symmetry of this balanced amplifier is better than that of the single-ended one presented in subsection 4.1.6.

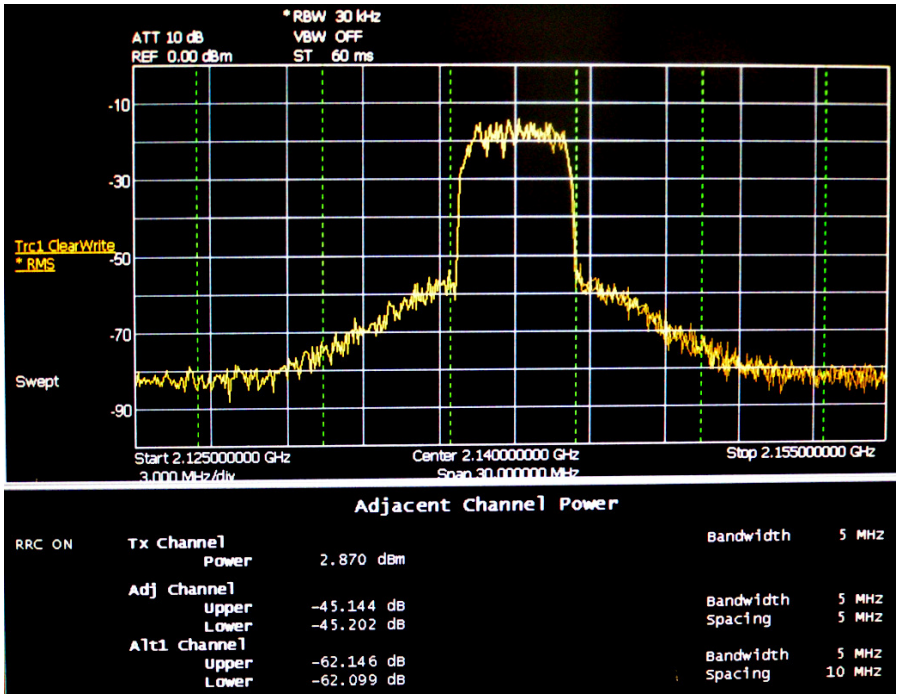


Fig. 4.30. Measured ACPR at 2.14 GHz and with 5 MHz offset as well as 45.8 dBm peak output power.

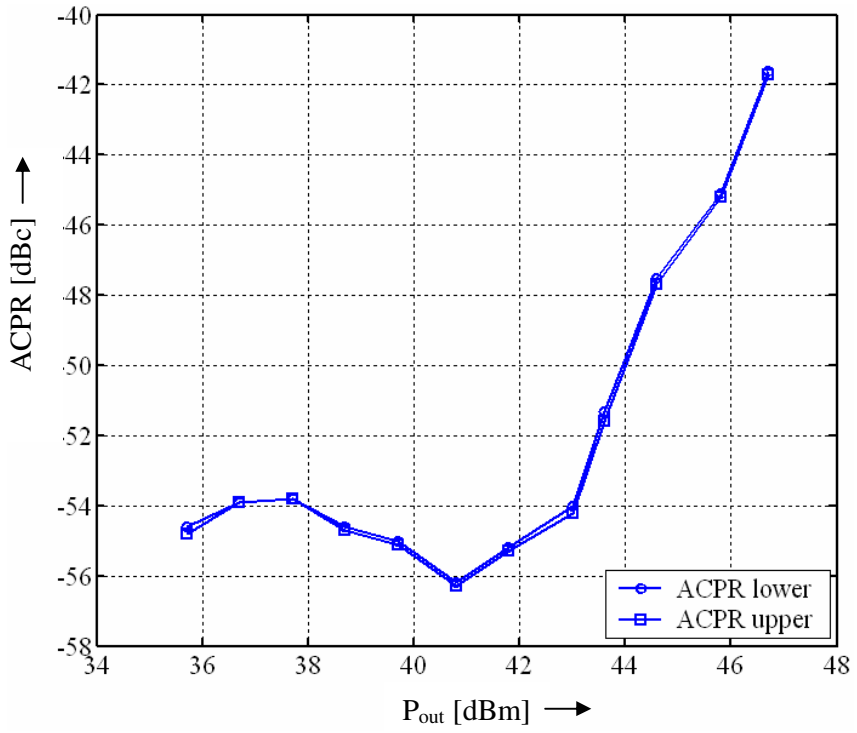


Fig. 4.31. Measured ACPR versus the output power of the balanced power amplifier.

Finally, a 50 W broadband balanced power amplifier using LDMOS transistors MRF21030SR3 is developed, which can simultaneously be adopted in the UMTS and in the GSM1800 systems in Europe. The bandwidth realized by this balanced power amplifier is

also larger than the bandwidth obtained by the other works listed in table 4.3. High efficiency, high linearity and outstanding cascading ability prove this design concept to be an ideal candidate for multi-standard multi-frequency base-station transmitters.

5. Summary

The provision of information and services at any time and any place is supported by the development of modern mobile communications. Every wireless communications system includes a transmitter, which mostly requires a power amplifier circuit in its radio frequency stage. In this work, radio frequency power amplifiers and their driving stages not only for the base stations but also for the cellular phones are developed. Both the hybrid and the integrated power amplifier circuits are designed with MOS transistors, which are regarded as the most cost-efficient devices due to the materials with lower cost and relatively simple process for the fabrication. Thanks to the development of modern semiconductor technologies, these transistors are now fast enough for designing the radio frequency circuits operating at several GHz. However, MOS transistors also perform drawbacks in designing some radio frequency circuits. For instants, CMOS devices perform exponential or logarithmic transfer functions only in the subthreshold region. In this region, however, the current is very small and hence the MOS devices are only suitable for very low power applications. Therefore, it is generally difficult to realize a logarithmic PGA with high output power in the CMOS technologies. Moreover, with the decreased size scaling to deep-submicron, the breakdown voltage of the CMOS transistors becomes lower and lower. Therefore, the ability of these transistors to sustain a large output voltage decreases and the difficulty of designing power amplifiers using CMOS technologies increases. In order to enhance the breakdown voltage of the MOS transistors, additional low doped n-drift region is used at the drain-terminal of the LDMOS transistors. The breakdown voltage of such transistors is boosted even up to 70 V, so that the design of power amplifiers for base station applications by using MOS transistors becomes feasible. But on the other hand, the optimum output impedance of the LDMOS transistors is usually very low, so that these transistors are principally not suitable for designing the broadband power amplifiers. To solve the problems mentioned above, modified circuit topologies are introduced in this work for designing different power amplifier circuits and their driving stages. The functional principle of these circuit topologies are described in detail. Simulation and experimental results of these circuits are also presented.

Today, CMOS is the technology of choice for a higher integration level and lower cost because it is capable of implementing a significant amount of digital signal processing and because the vast majority of today's integrated circuits are implemented in this technology. In this work, efforts are devoted to design the power amplifiers and their driving stages for cellular phones using standard deep-submicron CMOS technology.

Firstly, an RF CMOS logarithmic programmable gain amplifier is designed using a 0.12-µm CMOS technology. This work presents a novel circuit concept for the PGA design, where several amplifier cells are connected in parallel which are digitally controlled by a demutiplexer. These amplifier cells separately employ the transistors with different gate width; and at any given time, only one amplifier cell is turned on. In this manner, different gain and adaptive power consumption can be achieved. In order to obtain the high immunity to the environmental noise, differential operation is adopted in the amplifier cells. The cascode configuration is also applied in the circuit, which effectively suppresses the Miller effect and increases the output impedance of the active devices. Instead of active load, an LC-tank is applied as the common load for all the amplifier cells. It is proven to consume much less voltage headroom than the active load; hence a higher output power can be accomplished. The PGA circuit is fabricated and is measured at RF. The gain of the PGA can be varied between -43 dB and 8 dB with a gain control step of about 3 dB. A large gain control range of 51 dB is realized. This is the largest logarithmic gain control range reported so far for the CMOS radio frequency PGA. Adaptive power consumption is also achieved by using the parallel configuration of the amplifier cells. With the largest gain of 8 dB, the maximum

output power reaches 9 dBm and the 1-dB compression point is located at 8 dBm, while the oIP3 is as high as 22 dBm.

Furthermore, a CMOS power amplifier for the application in a cellular phone is proposed in this work. To solve the problem of low breakdown voltage, the HiVP structure is employed in this design. The HiVP structure is a High Voltage and High Power device configuration in which several transistor devices are connected DC and RF in series, so that the large output voltage can be divided by all the cascaded devices. However, the number of the transistors used in the HiVP structure must be restricted under the consideration of equal division of the supply voltage and the shift of the output characteristic. Since the same current flows through all the transistors, the different drain voltage can only be obtained with different drain impedance, which can be adjusted by the shunt capacitance connected at the gate of the upper transistor. In case that high output power is required, the DC current flowing through the HiVP circuit can even be larger than 1 A; hence transistors with gate width of several mm are required. The design approach to build compact layout for the large transistors is described in this work. Using the staggered layout concept described in Fig. 3.24, the chip area can effectively be decreased. Additionally, the phase difference along the gate lines of the numerous amplifier cells can be diminished by using the H-structure. In this work, microwave HiVP class AB power amplifiers are attempted in the deep-submicron CMOS technologies. To increase the matching flexibility and avoid excessive power loss of the on-chip inductors, the matching networks are implemented on the PCB. The fabricated power amplifier circuit is measured and the final measurement results are summarized and analysed as follows. With a supply voltage of 3.6 V, the maximum output power of 29.5 dBm is obtained at 900 MHz, while the maximum power added efficiency reaches 34.5 %. This amplifier circuit belongs to the very few efforts over the world to attempt microwave power amplifiers in deep-submicron CMOS technologies. The realized characteristics of this power amplifier are comparable in comparison to the other works presented in recent years, or even better. Moreover, the design concept for transistors with adjustable gate width is introduced in this work, where the power added efficiency achieved on the low power levels is greatly improved.

Using Motorola LDMOS transistor MRF21030SR3, broadband power amplifiers used in the base stations of the mobile communications systems are also developed in this work. A single-ended class AB power amplifier is first designed, which can simultaneously be used in the GSM1800 and in the UMTS base stations of Europe. The stability is one of the most important issues for such broadband power amplifier design, since almost all the transistors are potentially unstable within a large bandwidth. A method for the stability improvement is demonstrated in this work, where a shunt resistor with an additional series capacitor is directly connected at the drain of the LDMOS transistor. This series capacitor protects the shunt resistor against a large current; hence the stability improvement obtained by the shunt resistor is guaranteed. Simulations also show that the series capacitor reduces the total DC power consumption; hence the power added efficiency of the amplifier circuit is enhanced. In order to fulfil the specification for the bandwidth of this multi-frequency single-ended power amplifier, not only discrete but also distributed impedance matching networks are attempted. The S-parameters of these power amplifier circuits with different matching networks are simulated. Comparing the simulation results, the matching network consisting of multi-section transmission lines is determined to be the best candidate to realize a broadband power amplifier. Moreover, quarter-wave transmission lines determined at the frequency of 2 GHz is applied as the biasing networks, which provide the DC path for the biasing voltage and simultaneously an open circuit for the RF signal at the fundamental frequency. The even harmonics, which cause significant signal distortions, can be shorted by such biasing networks. The proposed power amplifier circuit is fabricated on the RO4003 substrate and measured. With a supply voltage of 26 V, the maximum output power reaches 43.5 dBm both in the GSM1800 band and in the UMTS downlink band. A linear power gain of about 13 dB

is achieved in a large frequency band. The highlight of this work is the 3-dB bandwidth of 1 GHz with the center frequency of 2.1 GHz. This is the largest bandwidth reported so far in the LDMOS technology. The maximum power added efficiency larger than 37 % is accomplished in the UMTS band and than 30 % in the GSM1800 band. High linearity is also obtained, which is denoted by the adjacent channel power ratio (ACPR) lower than - 40 dBc within the whole power level.

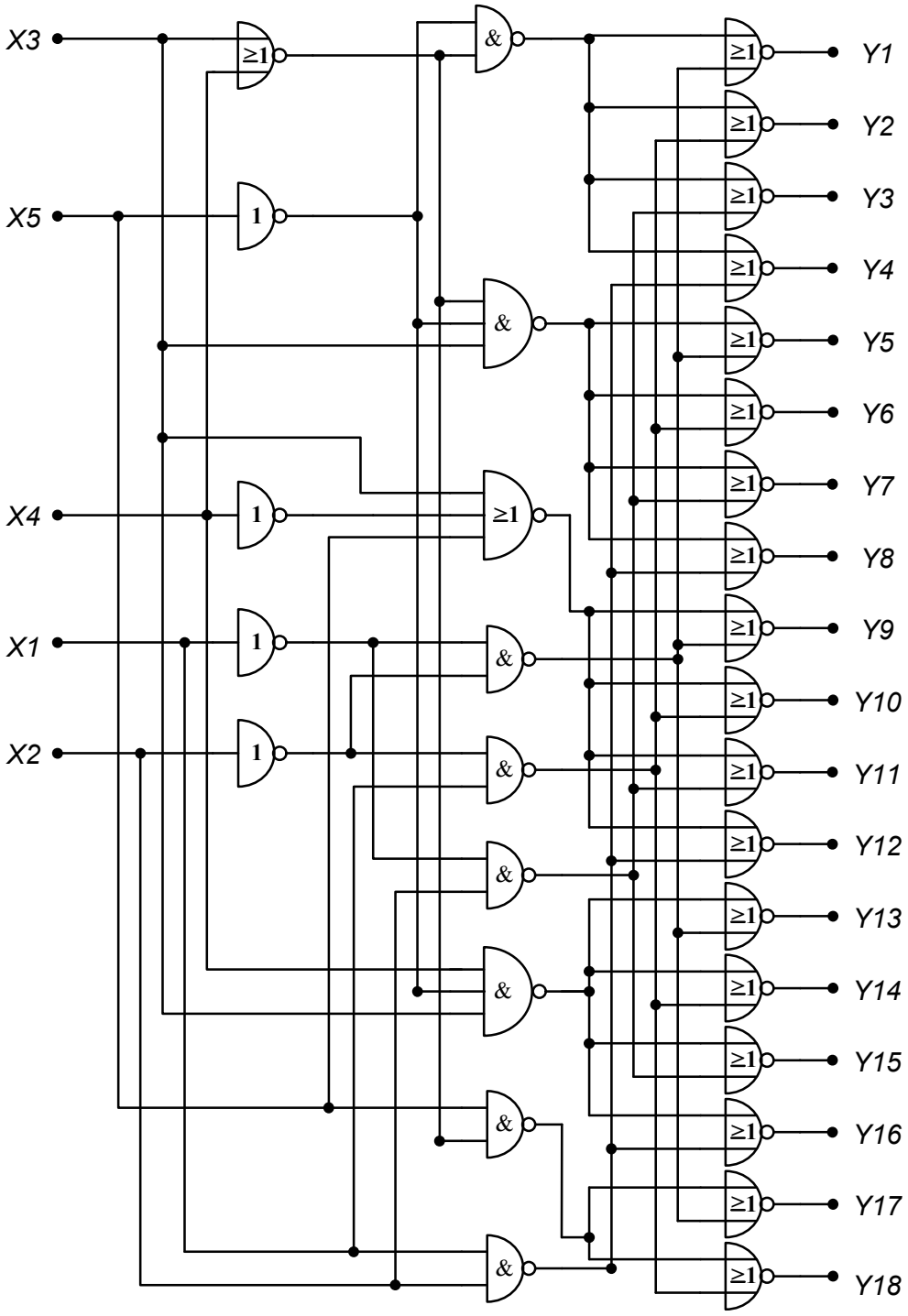
The voltage standing wave ratio (VSWR) is an essential problem in the design of a single-ended power amplifier, since the output matching network is generated according to the power matching, not according to the conjugate complex matching. In order to solve this problem, a balanced LDMOS broadband class AB power amplifier is designed. Two identical single-ended amplifiers are connected in parallel through two quadrature hybrid couplers, which operate as a power divider at the input and a power combiner at the output, respectively. The two reflected signals from the two individual amplifiers are 180° out of phase both at the input port and at the output port of the balanced amplifier. They compensate each other, so that a VSWR of 1 can theoretically be accomplished at both ports. In this work, the quadrature hybrid couplers are realized with Wilkinson power dividers and additional quarter-wave transmission lines. Using the same active device, the same matching networks and the same biasing networks obtained for the single-ended one, the balanced power amplifier is fabricated. Good performances are obtained in the measurements. With a supply voltage of 26 V, the maximum output power reaches 47 dBm (50 W) both in the GSM1800 and in the UMTS downlink bands. A linear power gain of about 12 dB is achieved in a large frequency band. It has a 3-dB bandwidth of 830 MHz with the center frequency of 2 GHz. The magnitudes of S_{11} and S_{22} are lower than -10 dB in the entire frequency band. Therefore, the realized balanced power amplifier has superior cascading ability. The maximum power added efficiency higher than 42 % is accomplished in the UMTS band and than 35 % in the GSM band. High linearity is also obtained in the measurement, which is denoted by the ACPR lower than - 40 dBc within the whole power level. It is proven that this realized LDMOS balanced power amplifier is suitable for the multi-standard multi-frequency mobile communication systems.

Appendix

A1. Logic of the 5-to-18 demultiplexer

$Y1 = \overline{X5} \cdot \overline{X4} \cdot \overline{X3} \cdot \overline{X2} \cdot \overline{X1} = \overline{\overline{X5} \cdot \overline{X3} + X4 + \overline{X2} \cdot \overline{X1}}$
$Y2 = \overline{X5} \cdot \overline{X4} \cdot \overline{X3} \cdot \overline{X2} \cdot X1 = \overline{\overline{X5} \cdot \overline{X3} + X4 + \overline{X2} \cdot X1}$
$Y3 = \overline{X5} \cdot \overline{X4} \cdot \overline{X3} \cdot X2 \cdot \overline{X1} = \overline{\overline{X5} \cdot \overline{X3} + X4 + X2 \cdot \overline{X1}}$
$Y4 = \overline{X5} \cdot \overline{X4} \cdot \overline{X3} \cdot X2 \cdot X1 = \overline{\overline{X5} \cdot \overline{X3} + X4 + \overline{X2} \cdot X1}$
$Y5 = \overline{X5} \cdot \overline{X4} \cdot X3 \cdot \overline{X2} \cdot \overline{X1} = \overline{\overline{X5} \cdot \overline{X4} \cdot X3 + \overline{X2} \cdot \overline{X1}}$
$Y6 = \overline{X5} \cdot \overline{X4} \cdot X3 \cdot \overline{X2} \cdot X1 = \overline{\overline{X5} \cdot \overline{X4} \cdot X3 + \overline{X2} \cdot X1}$
$Y7 = \overline{X5} \cdot \overline{X4} \cdot X3 \cdot X2 \cdot \overline{X1} = \overline{\overline{X5} \cdot \overline{X4} \cdot X3 + X2 \cdot \overline{X1}}$
$Y8 = \overline{X5} \cdot \overline{X4} \cdot X3 \cdot X2 \cdot X1 = \overline{\overline{X5} \cdot \overline{X4} \cdot X3 + \overline{X2} \cdot X1}$
$Y9 = \overline{X5} \cdot X4 \cdot \overline{X3} \cdot \overline{X2} \cdot \overline{X1} = \overline{X5 + \overline{X4} + X3 + \overline{X2} \cdot \overline{X1}}$
$Y10 = \overline{X5} \cdot X4 \cdot \overline{X3} \cdot \overline{X2} \cdot X1 = \overline{X5 + \overline{X4} + X3 + \overline{X2} \cdot X1}$
$Y11 = \overline{X5} \cdot X4 \cdot \overline{X3} \cdot X2 \cdot \overline{X1} = \overline{X5 + \overline{X4} + X3 + X2 \cdot \overline{X1}}$
$Y12 = \overline{X5} \cdot X4 \cdot \overline{X3} \cdot X2 \cdot X1 = \overline{X5 + \overline{X4} + X3 + \overline{X2} \cdot X1}$
$Y13 = \overline{X5} \cdot X4 \cdot X3 \cdot \overline{X2} \cdot \overline{X1} = \overline{\overline{X5} \cdot X4 \cdot X3 + \overline{X2} \cdot \overline{X1}}$
$Y14 = \overline{X5} \cdot X4 \cdot X3 \cdot \overline{X2} \cdot X1 = \overline{\overline{X5} \cdot X4 \cdot X3 + \overline{X2} \cdot X1}$
$Y15 = \overline{X5} \cdot X4 \cdot X3 \cdot X2 \cdot \overline{X1} = \overline{\overline{X5} \cdot X4 \cdot X3 + X2 \cdot \overline{X1}}$
$Y16 = \overline{X5} \cdot X4 \cdot X3 \cdot X2 \cdot X1 = \overline{\overline{X5} \cdot X4 \cdot X3 + \overline{X2} \cdot X1}$
$Y17 = X5 \cdot \overline{X4} \cdot \overline{X3} \cdot \overline{X2} \cdot \overline{X1} = \overline{X5 \cdot \overline{X3} + X4 + \overline{X2} \cdot \overline{X1}}$
$Y18 = X5 \cdot \overline{X4} \cdot \overline{X3} \cdot \overline{X2} \cdot X1 = \overline{X5 \cdot \overline{X3} + X4 + \overline{X2} \cdot X1}$

A2. Schematic of the 5-to-18 demultiplexer



References

[1] W. C. Jakes, "Microwave mobile communications," AT&T, 1974, reprinted by IEEE Press, Piscataway, NJ.

[2] J. G. Proakis, "Digital communications," McGraw-Hill, New York, 1989.

[3] S. C. Cripps, "RF power amplifier for wireless communications," 1999 Artech House, Inc.

[4] D. R. Lohrmann, "Amplifier has 85% efficiency while providing up to 10 Watts power over a wide frequency band," *Electronic Design*, vol. 14, pp. 38-43, March 1966.

[5] A. D. Artym, "Switching mode of high frequency power amplifiers," *Radiotekhnika*, vol. 24, pp. 58-64, 1969.

[6] V. V. Gruzdev, "Calculation of circuit parameters of single-ended switching-mode tuned power amplifiers," *Trudy MEI*, vol. 2, pp. 124-128, 1969.

[7] I. A. Popov, "Switching mode of single-ended transistor power amplifier," *Poluprovodnikovye pribory v tekhnike svyazi*, vol. 5, pp. 15-35, 1970.

[8] V. B. Kozyrev, "Single-ended switching-mode tuned power amplifier with filtering circuit," *Poluprovodnikovye pribory v tekhnike svyazi*, vol. 6, pp. 152-166, 1971.

[9] H. Kobayashi, J. M. Hinrichs, and P. M. Asbeck, "Current-mode class D power amplifiers for high-efficiency RF applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 49, no. 12, Dec. 2001, pp. 2480-2485.

[10] N. O. Socal, and A. D. Socal, "Class E – a new class of high-efficiency tuned single-ended switching power amplifiers," *IEEE J. Solid-State Circuits*, vol. SC-10, pp. 168-176, June 1975.

[11] A. V. Grebennikow, and H. Jaeger, "Class E with parallel circuit – a new challenge for high-efficiency RF and microwave power amplifiers," *IEEE MTT-Digest*, pp. 1627-1630, 2002.

[12] A. Grebennikov, "Switched-mode tuned high-efficiency power amplifiers: historical aspect and future prospect," *IEEE Radio Frequency Integrated Circuits Symposium*, pp. 49-52, June, 2002.

[13] D. Milosevic, J. Tang, and A. Roermund, "Explicit design equations for Class-E power amplifiers with small DC-feed inductance," *European Conference on Circuit Theory and Design*, vol. 3, pp. 111-113, August, 2005

[14] V. J. Tyler, "A new high-efficiency high power amplifier," *Marconi Review*, vol. 21, 1958, pp. 96-109.

[15] S. Tu, and C. Toumazou, "Highly efficient CMOS class E power amplifier for wireless communications," *Proceedings of the 1998 IEEE International Symposium on Circuits and Systems*, vol. 3, pp. 530-533, June, 1998.

[16] F. Fortes and M. Rosario, "A second harmonic class F power amplifier in standard CMOS technology," *IEEE Transactions on Microwave Theory and Techniques*, vol. 49, no. 6, June. 2001, pp. 1216-1220.

[17] J. E. Lilienfeld, US Patent 1,745,175, 1930.

[18] O. Heil, British Patent 439,457, 1935.

[19] C. T. Sah, "A new semiconductor tetrode – the surface-potential controlled transistor," *Proc. IRE*, vol. 49, 1961, pp. 1623-1634.

[20] A. S. Grove, and D. J. Fitzgerald, "Surface effects on p-n junctions: characteristics of surface space-charge regions under nonequilibrium conditions," *Solid-State Electron.*, vol. 9, pp. 783-806, 1966.

[21] H. J. M. Veendrick, "MOS ICs," VCH Verlagsgesellschaft mbH, 1992.

[22] B. Razavi, "Design of analog CMOS integrated circuits," McGraw-Hill. Inc., 2001.

[23] A. K. Ezzeddine, and H. Huang, "The high voltage/high power FET (HiVP)," *IEEE Radio Frequency Integrated Circuits Symposium*, 2003, pp. 215-218.

[24] **L. Wu**, R. Tao, U. Basaran, J. Luger, I. Dettmann, and M. Berroth, "The integrated 2W high voltage/high power 0.12-µm RF CMOS power amplifier," *34th European Microwave Conference (EuMC 2004)*, Amsterdam, Netherlands, pp. 451-454, October 12-15, 2004.

[25] S. R. Novis, and L. Pelletier, "IMD parameters describe LDMOS device performance", *Microw. RF*, vol. 37, pp. 69-74, July 1998.

[26] F. Lepine, A. Adahl, and A. Zirath, "L-Band LDMOS power amplifiers based on an inverse class-F architecture," *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, no. 6, June. 2005, pp. 2007-2012.

[27] F. Jian, L. Zhaoji, and Z. Bo, "Study for safe operating area of high voltage LDMOS," *in Int. Business of Elec. Product Reliability and Liability Conf.*, Apr. 2004, pp. 15-18.

[28] T. Lee, "The design of CMOS radio-frequency integrated circuits," Cambridge University Press, 1998.

[29] P. R. Gray, and R. G. Meyer, "Analysis and design of analog integrated circuits," John Wiley &Sons, Inc, pp. 226, 3rd. Edition, 1992.

[30] P. H. Smith, "Transmission line calculator," Electronics, vol. 12, no. 1, pp. 29-1, January 1939.

[31] M. Albulet, "RF power amplifiers," Noble Publishing Corporation, 2001.

[32] D. M. Pozar, "Microwave engineering," John Wiley & Sons, Inc, 2nd ed., 1998.

[33] I. J. Bahl, and D. K. Trivedi, "A designer's guide to microstrip line," Microwaves, May 1977, pp. 90-96.

[34] G. Gonzalez, "Microwave transistor amplifiers," Second edition, Prentice Hall Inc., New Jersey, 1997.

[35] L. Young, "Table of cascaded homogeneous quarter-wave transformers," *IEEE Trans. on Microwave Theory and Techniques*, vol. 7, no. 2, April. 1959, pp. 233-237.

[36] W. H. Doherty, "A new high efficiency power amplifier for modulated waves," *Proceedings of the IRE*, vol. 24, no. 9, pp. 1163-1182, 1936.

[37] I. Dettmann, **L. Wu**, and M. Berroth, "Comparison of A single-ended class-AB, a balanced and a Doherty power amplifier", *17th Asia-Pacific Microwave Conference (APMC 2005)*, Suzhou, China, vol. 2, pp. 1167-1170, 2005.

[38] Y. S. Noh, "An intelligent power amplifier MMIC using a new adaptive bias control circuit for W-CDMA Applications", *IEEE Journal of solid-state circuits*, vol. 39, no.6, pp. 967-970, June, 2004.

[39] S. Reed, Y. Wang, F. Huin, and S. Toutain, "HBT power amplifier with dynamic base biasing for 3G handset applications," *IEEE Microwave and Wireless Components Letters*, vol. 14, no. 8, pp. 380-382, August, 2004.

[40] M. L. Edwards, and J. H. Sinsky, "A new criteria for linear 2-Port stability using a single geometrically derived parameter," *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-40, December 1992, pp. 2803-2811.

[41] M. E. Hines, "The virtures of nonlinearity-detection, frequency conversion, parametric amplification, and harmonic generation," *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-32, September 1984, pp. 1097-1104.

[42] D. M. Pozar, "Microwave and RF Design of Wireless Systems," 2000, John Wiley & Sons, Inc.

[43] R. Harjani, "A variable gain CMOS amplifier with exponential gain control," *in Symp. VLSI Circuits Dig. Tech. Papers,* June 2000, pp. 146-149.

[44] W. C. Song, C. J. Oh, G. H. Cho, and H. B. Jung, "High frequency/high dynamic range CMOS VGA," *Electron. Lett.,* vol. 36*,* pp. 1096-1098, Jun 2000.

[45] P. Orsatti, F. Piazza, and Q. Huang, "A 71-MHz CMOS IF-baseband strip for GSM," *IEEE Journal of Solid-State Circuits,* vol. 35, no. 1*,* pp. 104-108, Jan. 2000.

[46] Recommended Minimum Performance Standards for CDMA2000 Spread Spectrum Mobile Station, 3GPP2 Doc. C.S0011-A

[47] Recommended Minimum Performance Standards for Dual-Mode Wide-band Spread Spectrum Cellular Mobile Station, IS-98A.

[48] Recommended Minimum Performance Standards for 1.8 to 2.0 GHz Code Division Multiple Access (CDMA) Personal Stations, ANSI J-Standard 018.

[49] S. Aggarwal, A. Khosrowbeygi, and A. Daanen, "A single-stage variable-gain amplifier with 70-dB dynamic range for CDMA 2000 transmit application," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 6, pp. 911-917, 2003.

[50] R. Saito, K. Hosoda, A. Hyogo, and T. Maruyama, H. Komuraki, H. Sato, K. Sekine, "A 1.8-V, 73-dB dynamic-range CMOS variable gain amplifier," *ESSCIRC 2004-29th European Solid-State Circuits Conference*, pp. 301-304, 2003.

[51] Y. S. Wang, and L. H. Lu, "5.7 GHz low-power variable-gain LNA in 0.18-μm CMOS," *Electronic Letters*, vol. 41, no. 2, pp. 66-68, 2005.

[52] Q. Chaudhry, R. Alidio, G. Sakamoto, and T. Cisco, "A SiGe MMIC variable gain cascade amplifier," *IEEE Microwave and Wireless Components Letters*, vol. 12, no. 11, pp. 424-425, 2002.

[53] K. R. Nary, and R. L. Van-Tuyl, "An MMIC amplifier for automatic level control applications," *IEEE Microwave and Millimetre-Wave Monolithic Circuits Symposium, Digest of Papers*, pp. 73-76, 1990.

[54] M. Detratti, J. P. Pascual, M. L. De-La-Fuente, J. Cabo, and J. L. Garcia, "A GaAs monolithic linear-in-dB wide-dynamic-range variable-gain amplifier with matching compensation for 1.95-GHz applications," *Microwave and Optical Technology Letters*, vol. 44, no. 3, pp. 251-257, 2005

[55] K. Nishikawa, and T. Tokumitsu, "An MMIC low-distortion variable-gain amplifier using active feedback," *IEEE Transactions on Microwave Theory and Techniques*, vol. 43, no. 12, pt. 2. pp. 2812-2816, 1995.

[56] M. Anowar Masud, and H. Zirath, "A 45 dB variable gain low noise MMIC amplifier," *35th European Microwave Conference*, pp. 669-672, 2005.

[57] **L. Wu**, U. Basaran, R. Tao, M. Berroth, and Z. Boos, "A 2 GHz CMOS dB-Linear programmable-gain amplifier with 51 dB dynamic range," *35th European Microwave Conference (EuMC 2005)*, Paris, France, October 3-7, 2005, pp. 617-620.

[58] G. Hau, T. Nishimura, and N. Iwata, "High efficiency, wide dynamic range variable gain and power amplifier MMICs for wide-band CDMA handsets," *IEEE Microwave and wireless components letters,* vol. 11, no. 1, pp. 13-15, Jan. 2001.

[59] C. Hsiao, C. Kuo, and Y. Chan, "Integrated CMOS power amplifier and down-converter for 2.4 GHz Bluetooth applications," *IEEE Radio and Wireless Conference*, pp. 29-32, August, 2001.

[60] C. Fallesen, and P. Asbeck, "A 1 W CMOS power amplifier for GSM-1800 with 55 % PAE," *IEEE MTT-S International Microwave Symposium*, vol. 2, pp. 911-914, May, 2001.

[61] Y. Ding, and R. Harjani, "A CMOS high efficiency 22 dBm linear power amplifier," *IEEE Custom Integrated Circuits Conference*, pp. 557-560, Oct. 2004.

[62] T. Sowlati, and D. Leenaerts, "A 2.4-GHz 0.18-µm CMOS self-biased cascade power amplifier," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 8, pp. 1318-1324, Aug. 2003.

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

[64] V. Knopik, B. Martineau, and D. Belot, "A 20 dBm CMOS class AB power amplifier design for low cost 2GHz-2.45GHz consumer applications in a 0.13-µm technology," *IEEE International Symp. on Circuits and Systems*, vol. 3, pp. 2675-2678, May. 2005.

[65] P. Reynaert, and M. Steyaert, "RF Power Amplifiers for Mobile Communications," 2006, Springer.

[66] J. A. Hopwood, "Microplasma: physics and applications," *Plasma Science Committee of the National Academies*, Boston, USA, Sep. 27, 2003.

[67] X. Jing, Z. Du, and K. Gong, "A compact multiband planar antenna for mobile handsets," *Antennas and wireless propagation letters*, vol. 5, Issue 1, Dec. 2006, pp. 343-345.

[68] J. Anguera, C. Puente, C. Borja, N. Delbene, and J. Soler, "Dual-frequency broad-band stacked microstrip patch antenna," *Antennas and wireless propagation letters*, vol. 2, Issue 1, 2003, pp. 36-39.

[69] C. Monzon, "A small dual-frequency transformer in two section," *IEEE Trans. On microwave theory and techniques*, vol. 51, no. 4, pp. 1157-1161, Apr. 2003.

[70] **L. Wu**. U. Basaran, I. Dettmann, M. Berroth, T. Bitzer, and A. Pascht, "A broadband high efficiency class-AB LDMOS balanced power amplifier," *35th European Microwave Conference (EuMC 2005)*, Paris, France, October 3-7, 2005, pp. 1079-1082.

[71] C. Cassan, and P. Gola, "A 3.5 GHz 25 W silicon LDMOS RFIC power amplifier for WiMAX applications," *RFIC Symposium*, Issue 3-5, pp. 87-90, June 2007.

[72] R. Bagger, P. Andersson, and C. D. Shih, "20 W LDMOS power amplifier IC for linear driver application," *IEEE MTT-S International Microwave Symposium*, Issue 3-8, pp. 1075-1078, June. 2007.