LARGE SIGNAL DESIGN OF SILICON FIELD EFFECT TRANSISTORS FOR LINEAR RADIO FREQUENCY POWER AMPLIFIERS

Paolo Fioravanti

Emerging Technologies Research Centre

Dissertation submitted to De Montfort University for the degree of Doctor of Philosophy

August 1, 2006

Abstract

This thesis is devoted towards the design and analysis of Si RF power MOSFETs, providing novel insight into the device physics which form the basis for the estimation of the performance and the optimisation of these devices.

Extensive 2D numerical simulations are used in the assessment of the behaviour of these devices. It is demonstrated using both simulations and measurements that short channel effects play an important role in the flattening of the transconductance characteristic and in the reduction of the peak value of the transconductance with respect to the classic ideal limit due to velocity saturation.

Innovative methods for the estimation of the large signal performance of Si RF Power MOSFETs are developed. These include the identification of the amplifier load line and novel expressions for optimum input and output matching impedances as well as for power gain and linearity. These are shown to yield reasonably accurate predictions of performance when compared with the results of harmonic balance simulations and/or measurement.

A new figure of merit capable of addressing the RF trade-offs encountered during design is proposed. The approach is demonstrated via optimisation of the shield in an RF power VDMOSFET. A p+ implant is shown to be a viable alternative to the conventional method.

Finally, a procedure for the estimation of the electro-thermal performance of packaged devices is demonstrated. The possibility of heat extraction through the top side of a vertical DMOSFET is evaluated using this method. Simulation results show the potential of this approach for the reduction of the thermal resistance.

Acknowledgments

I would like to thank the following persons:

- ✓ My supervisors, Prof. M. M. De Souza and Dr. O. Spulber, for their close supervision and advice, their support and patience.
- ✓ Prof. S.N. Ekkanath Madathil and Dr. G. Cao for the inestimable and formative chats.
- ✓ All the past and present members of the centre for their help and friendship. In particular, I would like to thank Dr. K. Vershinin, Dr. N. Mogulnaia, Dr. M. Sweet, Dr. S. Hardikar, Dr. R. Cross, Mr. P. Taylor, Mr. T. Pease, Mr. M. Sperini, Mr. M. Cacheda Fondevila, Mr. R. Tadikonda, Mrs. D. Cross, Mr. D. Green, Mr. J. Nichols, Mr. D. Kumar, Mr. S. Atarah, Mr. S. Sicre, and Mr. D. Casterman.
- ✓ Dr. M. Cirstea, without whom I would not have come to the UK.
- ✓ All my dearest friends, which still might not exactly have understood what I have been doing all this time in the UK (!!): Mr. L. Renzi, Mr. M. Iacoboni, Mr. A. Prinzivalli, Mr. R. Di Bonaventura, Mr. G. Consorte, Dr. J. Ochando.
- ✓ All the members of my family, my grandmother Dina, my aunts Piera, Pina and Alessandrina, my uncles Angelo, Guglielmo and Tullio, my cousins Elena, Marco, Domenico and Enrico.
- ✓ My father, mother and brother in law Γιώργος, Βασιλική and Θάνος, and also Δέσποινα and Χρήστος.

Finally, special thanks to my father Giovanni, my mother Grazia, my brother Lorenzo and my sweet half, Penny, which have been always in my heart and in my thoughts, source of encouragement and strength, inestimable and constant supply of support, patience and love.

Table of Contents

1

Introduction

The demand for high frequency semiconductor devices has considerably increased in the last few years. Stringent specifications on linearity and efficiency need to be satisfied in the design of the next generation communication systems. Special attention is required in the design of the power amplifier stage to meet optimum trade-off between linearity and efficiency. While the implementation of complex circuital schemes certainly helps improve the performance of the system, the choice of an adequate device may drastically simplify the design of the amplifier. Strangely, the development of optimum power devices for communication systems does seem to be the only phase of the design where the performance trade-off is not fully taken into account. There is very little literature describing how these trade-offs could be resolved at a device level.
This is particularly true for Si RF power MOSFETs, where the lack of reliable physical device models complicates the design. Necessarily, the only viable approach for the optimisation of these devices is represented by the prediction of the device characteristics through a simulator. The typical device optimisation consists of the reduction of the capacitance content and the improvement of the transconductance. Although this is certainly a reasonable way to begin, the complexity of the relationship between these quantities and the actual RF performance of the device requires a more precise solution.

This thesis focuses on the development of Si MOSFETs for RF power applications. Beginning with an overview of the RF power amplifier and the requirements for RF power MOSFETs, this work provides a novel insight into the device physics and novel approaches for the estimation of the performance and the optimisation of these devices are proposed.

1.1 Objectives

This thesis focuses on the development of tools for the design of Si RF power MOSFETs. The goal is to concentrate on the resolution of the performance trade-

offs so commonly encountered during RF circuit design. In order to do so, the possibility of assessing the performance of a device before a prototype is realised is investigated. Analytical methods are developed in this work for the determination of the main RF performance figures of merit. When available, measurements have been used for the validation of these expressions.

In order to permit a more complete characterisation of the device behaviour before prototyping, this thesis also promotes the investigation of the effect of the package on the performance of the device. The analytical modelling of the thermal impedance of the package, supplemented by electro-thermal device simulations, is proposed for such a task.

This thesis also aims to provide a deeper understanding of the physical behaviour of Si RF power MOSFETs. Compact and equivalent models have been developed to facilitate the comprehension of main phenomena taking place in the devices considered.

1.2 Organisation

This thesis is organised in seven chapters.

Chapter 2 is an overview of the basic concepts for RF power amplifiers and Si devices.

In chapter 3, an analytical study of the transconductance characteristic of Si RF Power MOSFETs is presented. A compact model is developed to provide a deeper understanding of the physical phenomena involved in the rise, flattening and fall-off of the transconductance characteristic. A novel explanation of the transconductance flattening and the relation between drift region and peak value of the transconductance is proposed.

Chapter 4 describes innovative methods for the estimation of the large signal performance of Si RF Power MOSFETs. The amplifier load line is considered in order to take into account the non linear behaviour of the device. Novel expressions are derived for the optimum input and output matching impedances as

well as for power gain by including gate and drift region resistance in the equivalent model of the device. Procedures for the prediction of gain compression and linearity are discussed.

Chapter 5 proposes a new approach for the optimisation of devices for RF power applications. The approach relies on the definition of a novel figure of merit and is demonstrated on a conventional Si RF power VDMOSFET equipped with a shield plate. A novel device concept is also introduced and optimised by using the proposed figure of merit.

Chapter 6 describes the implementation of a novel procedure for the estimation of the electro-thermal performance of packaged devices. The possibility of heat extraction through the top side is discussed for a vertical DMOSFET.

Finally, the work is summarised in Chapter 7, along with possibilities for future work.

2

Basic RF Power Concepts – An overview

Abstract

This chapter constitutes an overview of the basic concepts which will be further utilised in the rest of this thesis.

Design issues, figures of merit and classification of RF power amplifiers are introduced. A brief description of the main Si RF power devices is provided. The general requirements of such devices for their use in RF power amplifiers are presented.

2.1 Introduction

High performance Radio Frequency (RF) power semiconductor devices are being achieved via the adoption of new technologies, mainly related to smaller features sizes and consequently device dimensions.

Increased frequency capabilities allow for communication systems with larger bandwidths. The quality of the signal is also an important issue. Signal distortion and noise reduction are of primary concern in communication systems in order to maximise transmission efficiency.

The explosive growth in wireless communication is driving towards the development of devices capable of high frequency. Power silicon MOSFETs are currently conquering the market in the sector of base station applications.

Among silicon devices the Lateral Double-diffused (LD)-MOSFET is commonly recognised as the best solution for RF applications. This is mainly due to its higher frequency capability with respect to the Vertical Double-diffused (VD)-MOSFET. On the other hand, the VDMOSFET has been shown to display better linearity [1][2] and to be better suited for broadband amplifier design [3].

In this chapter, the basic principles and concepts regarding the design and characterisation of RF power amplifiers are described in section 2.2. Specific requirements for the design of Si DMOSFETs for RF Power applications are discussed in section 2.3.

2.2 The power amplifier

RF Power Amplifiers (PAs) are a fundamental component of any communication system. The core of the power amplifier, the active device, is usually the most expensive component of the entire system. In order to reduce costs, the device is always operated at its performance limits.

Although the active device is typically non linear in nature, most of the RF PAs are derived as variations of linear designs. Nevertheless, a fundamental difference exists between these amplifier topologies. This difference lies in the design of the passive network between the active device and the load of the amplifier, the output matching network (OMN).

In RF PAs, optimum power transfer to the load is never achieved by conjugate matching. Since the device is operated at its maximum, all the assumptions of ideal behaviour fail to apply. The physical limits of the active device (maximum current and/or voltage) need to be taken into account. The load-line concept represents the solution to this problem. Its underlying principle is thoroughly described in Chapter 4.

A simplified RF PA configuration for RF applications is shown in Figure 2.1, where input matching network (IMN) and output matching network are in series with the characteristic impedance Z_C representing the transmission line. The role of the matching networks is to transform the transmission line impedance to present the active device with the optimum resistance for power transfer. Accessories passive networks (AN) might also be necessary in order to satisfy linearity and stability requirements for the amplifier.

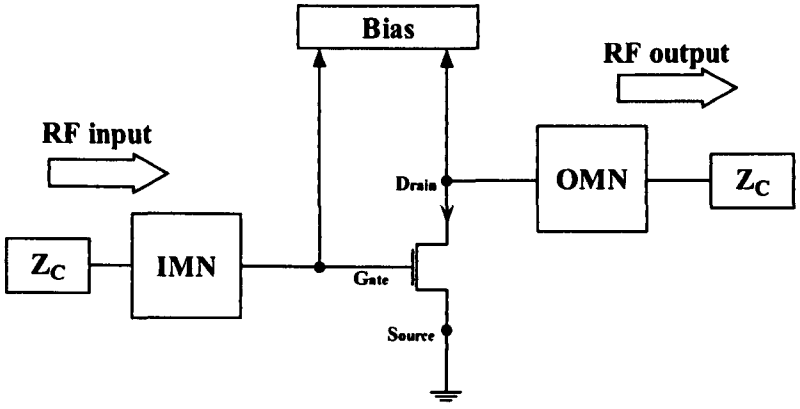


Figure 2.1: Simplified block diagram of a power amplifier.

A complete overview of linearity, efficiency and stabilisation schemes for AN of RF PAs can be found in [4][5]. Theory and principles of matching network design are well described in [6].

In the following, the main figures of merit used in the characterisation of power, linearity and frequency performance of RF PAs are described. Their determination provides the means for comparison of performance between different amplifier designs and/or active devices. A classic categorisation of the power amplifier is also described.

2.2.1 Power performance

Power gain

An expression for the power gain of a power MOSFET can be derived by using the simplified device model shown in Figure 2.2.

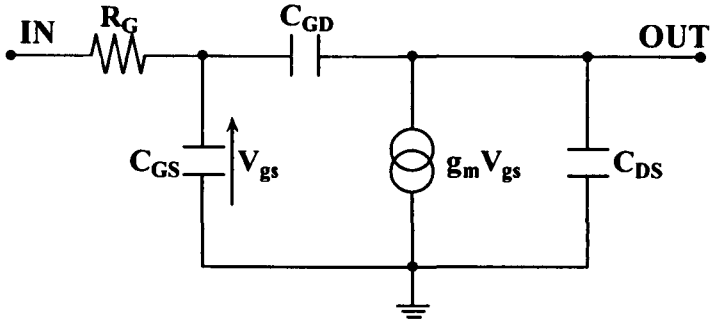


Figure 2.2: Simplified equivalent MOSFET model.

Considering the device output terminated by the conjugate match of the output impedance [7], the resolution of the equivalent circuit yields

$$G(\omega) = \frac{P_L}{P_{IN}} \approx \frac{1/2\left(\frac{\omega_T}{2\omega} i_{IN}\right)^2 g_{out}^{-1}}{\frac{i_{IN}^2 R_G}{2}} \approx \frac{\omega_T}{\omega^2 4 C_{GD} R_G}, \quad (2.1)$$

$$g_{out} \approx g_m \frac{C_{GD}}{C_{GS} + C_{GD}} = \omega_T C_{gd}, \quad (2.2)$$

where G is the power gain, P_L is the power dissipated on the load, P_{IN} is the power dissipated on the gate resistance, ω_T is the device cut-off frequency, i_{IN} is the input current and g_{out} is the resistive part of the output impedance.

The input power is calculated considering the series gate resistance as the only dissipative element of the input network in which, the Miller effect of the

feedback capacitance is also neglected. The output power expression is obtained by evaluating the output admittance of the device by considering that the output is conjugate matched. This implies that only half of the current from the current generator is available in the matching circuit for power transfer. The other half is dissipated in the device. The presence of the output matching circuit allows just the conductive, real part, component of the output admittance.

Gain compression point

The 1dB compression point P_{1dB} provides an indication of the amount of output power that can be delivered in linear conditions. It is a measure of the amount of power that can be handled by the power amplifier. P_{1dB} is defined as the RF output power corresponding to a reduction of 1dB in power gain with respect to the small signal power gain of the device.

Gain compression occurs when the output signal amplitude does not increase proportionally to the input signal. This is generally associated with the clipping of the output waveform and therefore with non linearity. For such a reason, the gain compression point can also be considered as an indication of device linearity.

Efficiency

Efficiency of an amplifier, also called drain efficiency, is defined as

$$\eta = \frac{P_{OUT}}{P_{DC}}, \qquad (2.3)$$

where P_{OUT} and P_{DC} represent the fundamental RF output and DC power respectively.

A more important figure of merit in RF Power applications is represented by the Power Added Efficiency (PAE), which also takes into account the power gain

$$PAE = \frac{P_{OUT} - P_{IN}}{P_{DC}} = P_{IN}\frac{G-1}{P_{DC}}. \qquad (2.4)$$

Both these efficiency measures are usually expressed in percentages. They relate the generated RF power to the DC power (dissipated as heat by the device.) As such, they are particularly important in applications for which power consumption and/or device heating are of concern.

2.2.2 Frequency performance

Cut off frequency

The cut-off frequency is extracted as the frequency at which the current gain of the amplifier reduces to unity. A classic expression for MOSFETs is

$$\omega_T \approx \frac{g_m}{C_{IN}} = \frac{g_m}{C_{GS} + C_{GD}}, \qquad (2.5)$$

where the cut off angular frequency ω_T is directly related to transconductance g_m and the input capacitance $C_{IN}=C_{GS}+C_{GD}$ of the device, with C_{GS} and C_{GD} representing the gate to source and gate to drain (feedback) capacitance.
Equation 2.5 is derived considering the simplified equivalent MOSFET model of Figure 2.2 by ignoring the effects of series gate resistance R_G.

Maximum operating frequency

The maximum frequency is defined as the frequency at which the power gain falls to unity. From Equation 2.1, the maximum angular frequency can be approximated as [7]

$$\omega_{max} \approx \frac{1}{2}\sqrt{\frac{\omega_T}{C_{GD}R_G}}. \qquad (2.6)$$

2.2.3 Linearity performance

Linearity performance is usually assessed via the determination of the intermodulation products. The intermodulation products are defined as "The additional frequencies at the output of a non-linear amplifier (or in general any non-linear network) when two or more sinusoidal signals are applied at the input" [8].
Imposing an input signal comprising several components at different frequencies, the non linear device behaviour yields the generation of undesired frequency components. These additional components are observed at frequencies which are the linear combination of the frequencies of the input signal. Of great importance

for the assessment of linearity performance are the frequency components which fall close or internally to the signal spectrum band.

An important measure of linearity is represented by the third order intermodulation distortion (IMD3). The IMD3 quantifies the amplitude of the output signal observed at the $2\omega_1$-ω_2 or $2\omega_2$-ω_1 frequency when a two-tone input signal of fundamental frequencies ω_1 and ω_2 is applied to the amplifier. IMD3 is typically measured in dBc, decibels to carrier level.

A simple formula for the assessment of IMD3 is provided in [9]

$$IMD3[dBc] = 20\log\left|\frac{\frac{3}{4}g_{m,3}V_{gs}{}^{3} + \frac{25}{8}g_{m,5}V_{gs}{}^{5}}{g_{m,1}V_{gs} + \frac{9}{4}g_{m,3}V_{gs}{}^{3} + \frac{25}{4}g_{m,5}V_{gs}{}^{5}}\right|, \qquad (2.7)$$

where the $g_{m,k}$ coefficients are related to the I_{DS}-V_{GS} output characteristic as

$$c_k = g_{m,k} = \frac{1}{k!}\frac{d^k I_{DS}}{dV_{GS}{}^{k}}\bigg|_{V_{GS}=V_{GS,Q}}. \qquad (2.8)$$

Equation 2.7 is derived from a 5th order expansion of the output current characteristic. The current generator is considered as the main cause of non linearity in the device; other sources of non linear behaviour are ignored.

2.2.4 Conventional amplifier classification

Amplification methods are categorised in literature depending on transmission technique (continuous or pulsed) and efficiency: classes A, B, AB and C represent the classes of amplifiers used for the amplification of time-continuous signals, while classes D, E, F and S (PWM) refer to signal transmissions where the transistor is used as a switch.

The following provides an overview of classes for time-continuous signal amplification, where the active device is assumed to display the idealised behaviour shown in Figure 2.3. The current is assumed to increase linearly from 0 to the saturated I_{MAX} value in correspondence of increases of gate voltage from the threshold (V_{th}) to the saturation voltage (V_{SAT}) values.

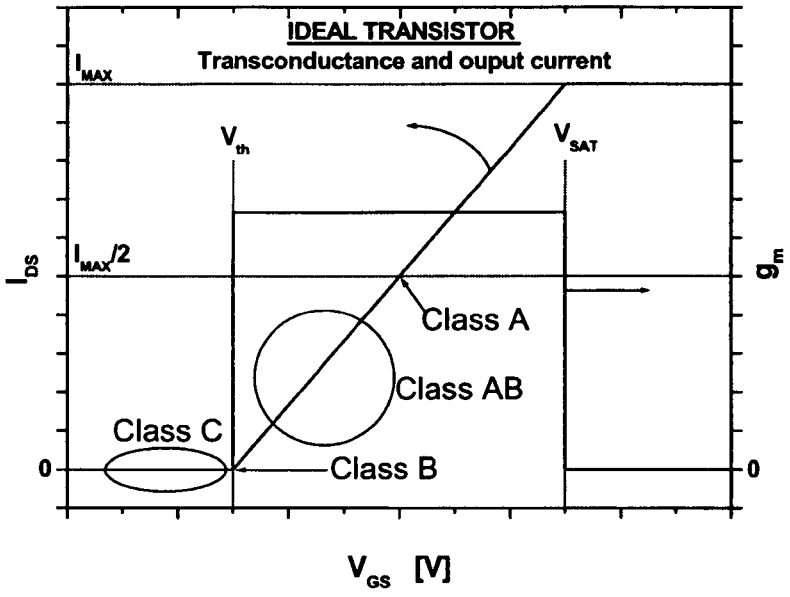


Figure 2.3: Transconductance and output characteristic of an ideal transistor highlighting the classic ABC classes of bias.

Class A amplifiers

The class A amplifier can be thought of as a small signal amplifier adapted for large signal transmission. As shown in Figure 2.3, the device is biased in the middle of the active region at $(V_{SAT}-V_{th})/2$ and $I_{MAX}/2$. This allows for unclipped current and voltage swings of maximum amplitudes of $(V_{SAT}-V_{th})/2$ and $I_{MAX}/2$ respectively. Waveform clipping occurs if the active device is driven beyond these limits, introducing distortion.

In class A the device is always biased on. It can be shown that this configuration allows for a maximum drain efficiency of 50%, which occurs when maximum voltage and current swings are achieved.

Class B amplifiers

In class B the amplifier is biased at its threshold voltage, as shown in Figure 2.3. The output signal is represented by the amplified, positive part of the AC input signal. By representing the output signal through its Fourier expansion and considering the so determined DC and fundamental frequency components, it can be shown that drain efficiency can rise up to a value of appreciatively 78%. The

maximum zero to peak value of the output current waveform is $I_{MAX}/2$ as in class A.

Class C amplifiers

In class C the amplifier is biased below the threshold voltage, as shown in Figure 2.3. The output signal is represented by a train of amplified AC pulses. The efficiency at the peak output power increases with respect to class B. In fact, it can be shown that the DC component of the signal is smaller than in class B. A typical compromise is the achievement of 85% efficiency. The reduction of conduction time also causes the output power to reduce with respect to class A and B.

Class AB amplifiers

In class AB the amplifier is biased over its threshold voltage, as shown in Figure 2.3. This situation represents a compromise between class A and class B amplifiers. Typical values of efficiency are around 60%. The output power is higher than in class B because of the increased conduction time of the active device.

2.3 Si RF Power MOSFETs

2.3.1 Device structures

Two main device structures can be identified in the Si RF Power market: the Vertical and the Lateral DMOSFETs. The simplified cross sections of these two devices are shown in Figure 2.4 and Figure 2.5. As in their high voltage counterparts, the breakdown voltage is sustained by a lightly doped n-drift region. On the other hand, the gate to n-drift region overlap is reduced in the RF devices in order to minimise capacitance content and improve power gain and frequency performance.

Contrary to LDMOSFETs, the reduction of the gate to drift region overlap via self-alignment of source and drift regions to the gate has just recently being reported in Vertical DMOSFETs [10].

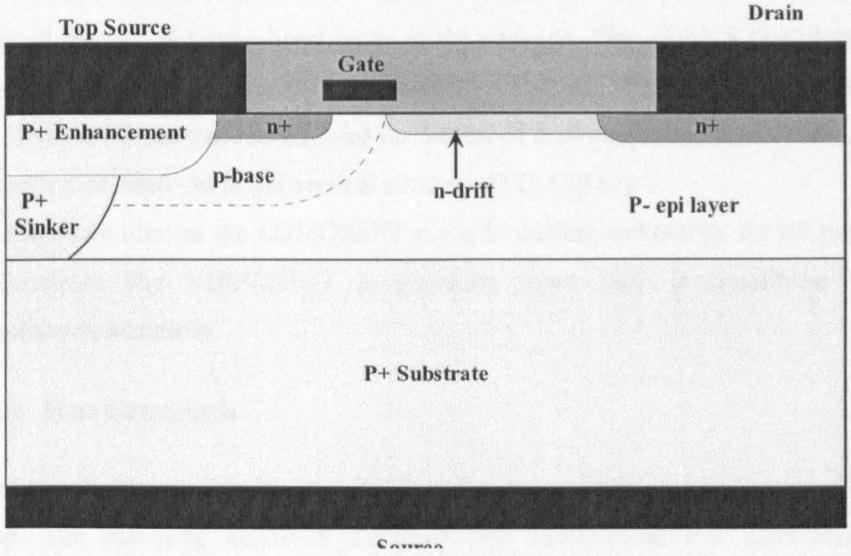


Figure 2.4: Cross section of an RF lateral power DMOSFET.

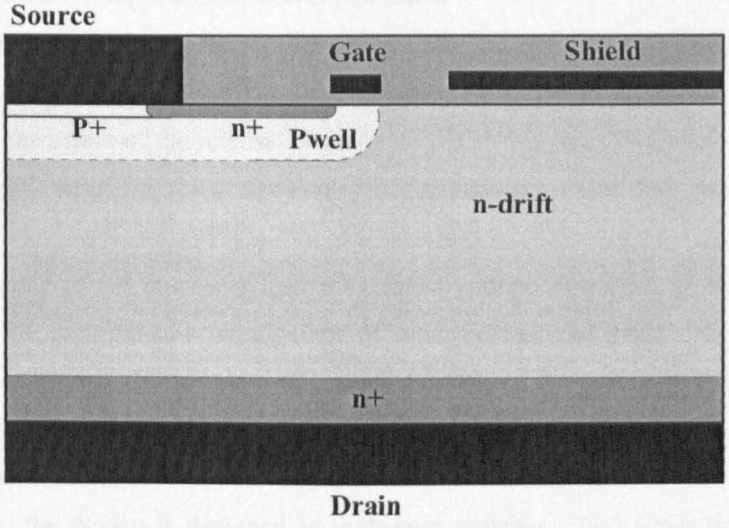


Figure 2.5: Cross section of an RF vertical power DMOSFET.

Two main differences can be highlighted between lateral and vertical DMOSFETs. These differences lie in the variations in used packages and in the drift region design.

The existence of a p+ sinker or a V-groove in LDMOSFET devices [11] permits to avoid the use of source bond wires to the package. This yields a considerable improvement in power gain at high frequency. Also, the drift region accessibility in LDMOSFET devices has allowed for advanced drift region designs which have not been paralleled yet in the vertical structure [12][13][14].

For these two reasons the LDMOSFET is the Si leading technology for RF power applications. The VDMOSFET is generally more used in broadband low frequency applications.

2.3.2 Requirements

As shown in Figure 2.1, the matching networks are always part of an amplifier setup. The matching networks transform the transmission line characteristic impedance into an equivalent pure resistance, allowing for maximum transfer of power and preventing the reflection of the signal.

Since the active device is mostly capacitive in nature, as acknowledged in Figure 2.2, matching networks generally display a resistive-inductive behaviour in order to cancel the effect of the internal capacitances of the device. This implies that in an RF Power amplifier the active device is generally presented with an inductive load.

The presence of the inductive load may cause voltage overshoot in the device, yielding the simultaneous application of high voltage and high current. This behaviour suggests the necessity of a careful design of the power MOSFET. The gate oxide thickness and the breakdown voltage must be adjusted in order to prevent device failure.

Typically, the device is designed to withstand voltages 2.5~3 times the supply voltage. The thickness of the gate is usually around 70nm for 28-48V applications. Such a thickness is required to sustain the electric field across the gate oxide, which for a 10V gate voltage and an oxide thickness of 70nm is in the

order of 10^5~10^6 volts per cm. In the sixth generation high voltage Motorola RF LDMOSFET [15], a gate oxide thickness of 30nm and a channel length of 0.5μm are used and still a drain to source breakdown voltage of 72V can be sustained.

High working frequency and high gain represent the most looked for characteristic in RF power devices. Achieving high working frequencies is directly linked to the decrease of the transit time of the carriers through the channel region. For this reason, short channel devices are largely used in RF applications. In short channel devices, the channel length reduction and the reach of velocity saturation in the channel can considerably reduce transit times, improving transconductance. Higher transconductance values have a positive effect on cut off frequency, maximum frequency and power gain, as acknowledged in Equations 2.1, 2.5 and 2.6. The improvement of transconductance also allows for the reduction of the device dimensions for a certain target output power.

Low capacitance content is highly desirable in an RF power MOSFET due to its positive effect on cut off frequency, maximum frequency and power gain. By improving the decoupling between input and output of the power amplifier, a low feedback capacitance also simplifies the design of matching and accessory networks and helps preventing oscillation by reducing undesired feedback.

Power gain and maximum operating frequency are also affected by the gate resistance value. Low gate resistances can be achieved by careful design of the device layout.

Power and frequency capabilities are also affected by the fabrication process and the device packaging. At these stages the parasitic capacitance content due to overlaps should be minimised and inductive/resistive wiring content reduced to a minimum.

Another important requirement in RF applications is represented by the linearity of the device, which has to be adequate to the target application. Modern communication systems require at least -30dBc IMD3 values for single channel communication systems [4][5][6]. At a device level, improvements of linearity are generally associated with constant values of the transconductance in

correspondence of large gate voltage variations [2]. Therefore, large transconductance bells are highly desired in RF power MOSFETs.

References

[1] De Souza M.M., Cao G., Sankara Narayanan E.M., Youming F., Manhas S.K., Luo J., Moguilnaia N., Progress in Silicon RF Power MOS Technologies – Current and Future Trends (INVITED), Proceedings of the International Caracas Conference on Device Circuits and Systems, 2002, pp. D-047.

[2] M. Trivedi, P. Khandelwal, and K. Shenai, Performance Modeling of RF Power MOSFET's, IEEE Transactions on Electron Devices, vol. 46, 1999, pp. 1794-1801.

[3] M. Trivedy, and K. Shenai, Comparison of RF Performance of Vertical and Lateral DMOSFET, IEEE ISPSD 1999, pp. 245-248

[4] S.C. Cripps, RF Power Amplifiers for Wireless Communications, Artech House, 1999.

[5] S.C. Cripps, Advanced Techniques in RF Power Amplifier Design, Artech House, 2002

[6] R. Gilmore, L. Besser, Practical RF circuit design for modern wireless systems, Artech House Publishers, 2003

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

[8] Beneking H., High Speed Semiconductor Devices, Chapman & Hall, 1994

[9] van der Heijden M.P., de Graaf H.C., de Vreede L.C.N., Gajadharsing J.R. and Burghartz J.N., Theory and Design of an Ultra-Linear Square-Law Approximated LDMOS Power Amplifier in Class-AB Operation, IEEE Trans. on microwave theory and techniques, 2002, pp.2176-2184

[10] N. Nenadovic, V. Cuoco, S.J.C.H. Theeuwen, H. Schellevis, G. Spierings, A. Griffo, M. Pelk, L.K. Nanver, R.F.F. Jos, J.W. Slotboom, RF Power

Silicon-On-Glass VDMOSFETs, IEEE Electron Device Letters, 2004, vol. 25, pp. 424–426

[11] O. Ishikawa-O, H. Yamada, and H. Esaki, A 2.45 GHz power LD-MOSFET with Reduced Source Inductance by V-Groove Connections, International Electron Devices Meeting, Technical Digest, p.166, 1985.

[12] J. Olsson, N. Rorsman, L. Vestling, C. Fager, J. Ankarcrona, H. Zirath, K.-H. Eklund, 1 W/mm RF Power Density at 3.2 GHz for a Dual-Layer RESURF LDMOS Transistor, IEEE Electron Device Letters, 2002, vol. 23, pp. 206-208

[13] J. Cai, C. Ren, N. Balasubramanian, J.K.O. Sin, High Performance Stacked LDD RF LDMOSFET, Proc. ISPSD, 2001, pp. 103-106

[14] M. Shindo, M. Morikawa, T. Fujioka, K. Nagura, K. Kurotani, K. Odaira, T. Uchiyama, I. Yoshida, High Power LDMOS for Cellular Base Station Applications, Proc. ISPSD, 2001, pp. 107-110

[15] H. Brech, W. Brakensiek, D. Burdeaux, W. Burger, C. Dragon, G. Formicone, B. Pryor, and D. Rice, Record Efficiency and Gain at 2.1GHz of High Power RF Transistors for Cellular and 3G Base Stations, IEEE IEDM 2003.

3

Transconductance behaviour in 28V Si RF Power MOSFETs

Abstract

An analytical study of the transconductance characteristic of Si RF Power MOSFETs is presented. A compact model highlights physical mechanisms previously neglected that explain the rise and flattening of the transconductance. Central to the model development is the derivation of a simple relation between the channel depletion charge and channel modulation. This strategy permits a simple yet comprehensive description of the intrinsic MOSFET, providing a clear insight into device behaviour. The flattening of the transconductance is demonstrated to occur due to an appreciably linear increase of the inversion charge with the gate to source voltage. A simplified formula is proposed for the peak transconductance value. The formula is shown to provide a clear explanation of the effect of the drift region resistance on the peak transconductance value.

3.1 Introduction

The power amplifier industry is increasingly looking for cost-effective, linear, Radio-Frequency (RF) Power devices for communications applications. Silicon MOSFETs lead the market for 28V mobile base station applications at frequencies up to 2.4GHz, displaying output power of over 100W at power gains of about 12dB [1]. Recent developments in ultra-thin silicon handling have led to the first RF power VDMOSFET capable of large signal operation at 2GHz [2], narrowing the gap historically separating the performance of vertical and lateral RF power MOSFETs [3]. In order to achieve an improvement in the power capability and linearity, an accurate understanding of the transconductance behaviour of these devices is of paramount importance. However, with the exception of [4], limited research has been conducted so far on the physical factors involved in the transconductance behaviour of RF Si Power MOSFETs. The reason for this can be identified in the lack of availability of suitable physical models.

The available models are empirical [5][6] or data based [7][8] and neither of these approaches provide a clear insight into the physical behaviour of the device. Analytical DMOS models developed in the seventies and eighties [9][10] are based on long channel approximations and consider a carrier velocity that saturates along the entire channel; these models have been shown to be not appropriate for submicron channel length DMOS transistors [11] and have been replaced by the more accurate enhancement/depletion (E/D) models and their derivatives [11][12][13][14]. In these models the channel region is modelled by using two serially connected MOSFET devices that share the same gate contact. More recently, improved accuracy has been achieved [15]-[29] by providing more sophisticated device descriptions and progressively including more physical phenomena.

However, the use of these models to obtain a singular analytical expression of the transconductance characteristic for the various regions of operation of the power MOSFET is complex and most importantly not able to provide a simple and clear tool for the explanation of the physical device behaviour.

A simplified model to represent Si RF Power MOSFETs has been proposed in [4], where the model used can be seen as a modification of the Level 1 MOSFET model to include the effects of the drift region and high frequency components. The work in [4] represents a singular effort to provide physical insight into the factors involved in the transconductance behaviour of RF Si Power MOSFETs. The transconductance flattening is proposed to be due mainly to the channel velocity saturation, as in [9] and [10].

In this work the explanation is extended to include channel modulation and short channel effects via a compact model. The idea is to obtain an expression for the transconductance which depends only on the intrinsic MOSFET design and on its bias. This permits a final expression for the transconductance in which the non-linear design-dependent drift resistance appears only through its effect on the bias and on the channel carrier velocity of the intrinsic MOSFET. This adds to the generality of the explanation of the transconductance behaviour and to its applicability to both vertical and lateral DMOSFETs, despite the possible complexity of the drift region design [30][31][32][33]. The same amount of current flows through the intrinsic MOSFET and the overall power MOSFET. Therefore, the idea behind this approach is to pay attention to the intrinsic MOSFET behaviour in a regime of non constant bias and in using the intrinsic MOSFET current equations to derive an expression for the overall power MOSFET transconductance. This approach, the transconductance expression and the explanation of the transconductance behaviour also appear in [34].

This chapter is organised as follows. A background review is given in section 3.2, highlighting some observed device behaviour. An intrinsic MOSFET compact model is presented in section 3.3 in order to include the main physical phenomena maintaining an acceptable level of simplicity. Model simplifications are then shown to yield a novel peak transconductance approximation. In section 3.4 the model is shown to permit a correct prediction of current and transconductance characteristics and is used to provide physical insights into the device behaviour. Simulation results [35] of a power VDMOSFET optimised for 28V RF applications are used to support the explanation. The schematic cross-section of

the device is reported in Figure 3.1, together with a simplified static equivalent lumped circuit model.

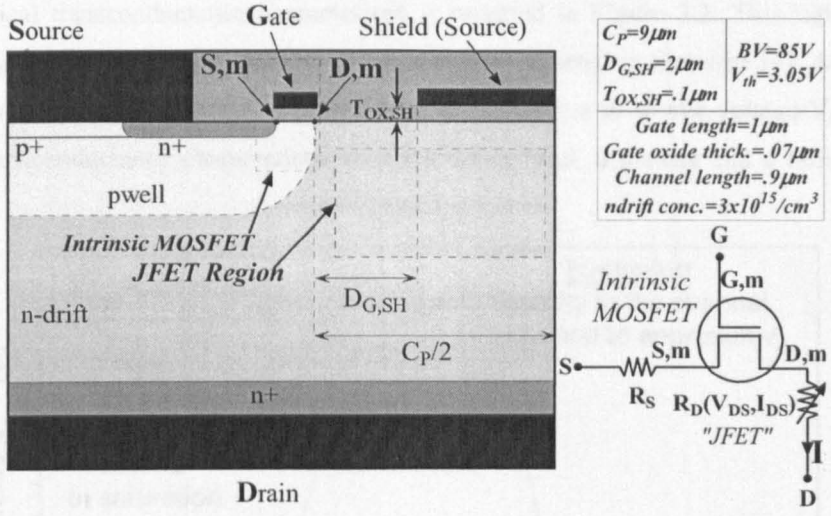


Figure 3.1: RF power VDMOSFET cross section and simplified static circuital equivalent model.

3.2 Background

In relation to the low power MOSFET, physical models for RF Si Power MOSFETs generally include a drift region resistance. Experimental determination of drift resistance is described in [36]-[39], which report the extension of low power parameter extraction techniques to the Power MOSFET device topology. The extracted data is typically used in the development of mathematical functions for modelling either the drift resistance or the potential $V_{D,m}$ at the drain end of the channel [14][16][17][29][40].

The modelling of an RF power silicon MOSFET consists of extracting a suitable model for the intrinsic MOSFET. The intrinsic MOSFET drain electrode is represented by the drain-end of the channel, D,m in Figure 3.1. The drift region can be modelled either as a JFET or as a variable resistor, whose value depends on the current level and on the drain to source voltage. The JFET source, drain and

gate electrodes are represented by the drain-end of the channel D,m, the drain and the source electrodes of the power MOSFET respectively.

A typical transconductance characteristic is reported in Figure 3.2. This figure highlights the region of operation of the device as reported in literature [4]. At a fixed drain to source voltage V_{DS} and for increasing gate to source voltage V_{GS}, the transconductance characteristic shows a rising front, a plateau and a falling front.

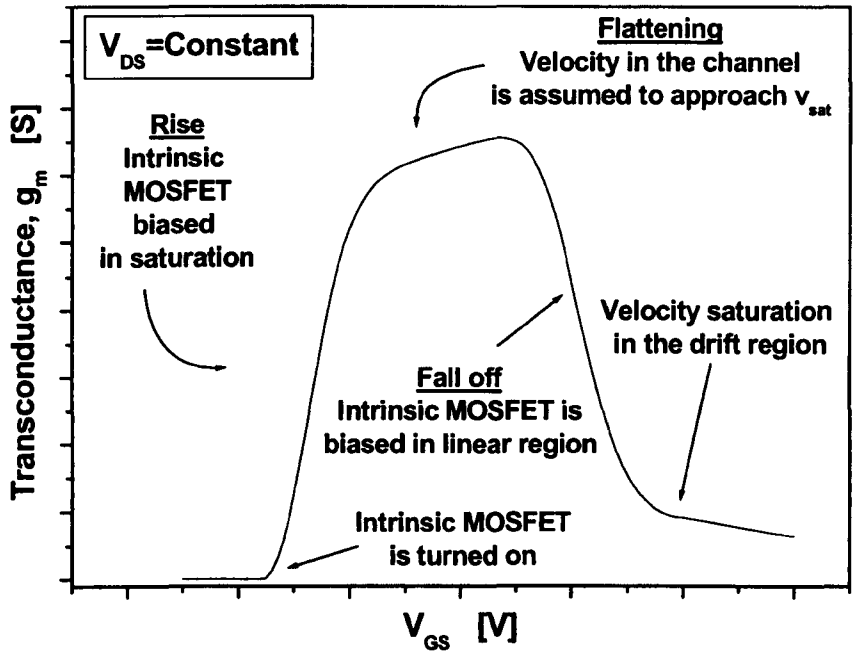


Figure 3.2: Typical RF Power MOSFET transconductance characteristic and its current understanding.

In [4] the rising and falling of the transconductance are respectively associated with the intrinsic MOSFET being biased in saturation and linear region respectively; the flattening is attributed to saturation velocity in the channel. In the fall-off regime, the transition from saturation to linear region of operation of the intrinsic MOSFET is attributed to the presence of the drift region resistance: the increase of V_{GS} raises the current level, causing a reduction of the voltage at the drain-end of the channel, i.e. the drain of the intrinsic MOSFET [4].

In the simple long channel approximation for the intrinsic MOSFET this leads to the ideal transconductance limit

$$g_{m,ideal} = WC_{OX}v_{sat}, \quad (3.1)$$

where C_{OX} is the gate oxide capacitance per unit area, W is the device total width and v_{sat} is about 10^7cm/s. Equation 3.1 implies that the velocity saturated transconductance value is independent on the drain to source voltage and on the drift region resistance. This is nonetheless not the case in Si RF power DMOSFETs where the peak g_m value is strongly affected by the drain-source voltage and by the drift region, as evident from Figure 3.3 and extensively demonstrated in Chapter 5.

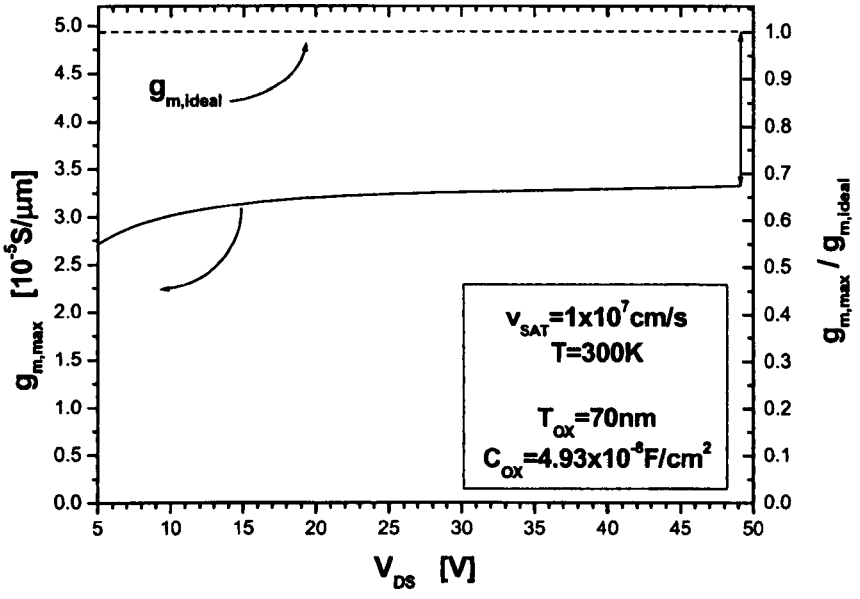


Figure 3.3: Simulated peak transconductance at several drain to source bias voltages. The data refers to the 28V Si RF Power MOSFET described in section 3.4.

Figure 3.3 also shows that the peak transconductance value $g_{m,max}$ remains always much below the ideal case of Equation 3.1. To account for this gap, as generally carried out in SPICE models, the saturation velocity v_{sat} is used as a fitting parameter: v_{sat} is in fact an equivalent maximum carrier velocity v_{max} whose value is extracted to fit the peak g_m value. This approach leads to a value of v_{max} which

is a function of the applied drain to source bias V_{DS} and therefore not physically equivalent to a constant saturation velocity. In fact, since $g_{m,max}$ is modelled as

$$g_{m,\max} = WC_{OX}v_{\max}, \qquad (3.2)$$

the results reported in Figure 3.3 imply that v_{max} increases with V_{DS}. The application of Equation 3.2 to extract v_{max} leads to a value that is about 66% of v_{sat} at V_{DS}=28V. This result is in contrast with the findings in section 3.4, where the maximum carrier velocity in the channel is shown to remain close to the saturation velocity when the intrinsic MOSFET is biased in saturation.

A more accurate inversion charge model and a physical velocity saturation value must be used in order to provide a deeper physical insight of the transconductance.

3.3 Model description

In this section a semi-analytical compact model is derived for the channel inversion charge of a power RF DMOSFET in the sub-threshold, linear and saturation regions of operation. The model is derived from the Level 3 model [41]-[45] by introducing a modified depletion charge model, ignoring narrow channel effects and implementing a simplified sub-threshold model. The choice of Level 3 as a starting point derives from the optimum compromise of this model between accuracy and simplicity.

Six parameters are added to take into account the effects of the non constant channel doping concentration. A smoothing function is used to avoid the first order current derivative discontinuity at the transition from linear to saturation region typical of Level 3 models. The proposed model does not make use of the velocity saturation as a fitting parameter: a physically significant value of 10^7cm/s is used for v_{sat}.

3.3.1 Channel modulation and Depletion charge

In this section an expression for the gate-controlled depletion charge of the intrinsic MOSFET of an RF power DMOSFET is obtained by extending a classical electrostatic model for a low power MOSFET.

The decrease of the channel depletion charge by the source and drain depletion regions is modelled by considering the simplified electrostatic model of the intrinsic MOSFET shown in Figure 3.4.

The depletion regions coming from the p-well/n+ junction and from the p-well/n-drift junction overlap with the gate induced depletion region, sharing part of the depletion charge. The unshared gate charge can be thought of as belonging to the trapezoidal area highlighted in Figure 3.4 [41], yielding:

$$Q_{depl} = Q_{depl,0} \cdot F_s , \qquad (3.3)$$

$$Q_{depl,0} = qN_{A,eff} W = \sqrt{2\varepsilon qN_{A,eff}(2\Psi_B)} , \qquad (3.4)$$

$$F_s = 1 - \frac{X_S + X_D}{2L_{CH}} , \qquad (3.5)$$

where $Q_{depl,0}$ is the depletion charge in the long channel case and F_s represents the fractional reduction of the depletion charge due to the drain and source depletion region overlap with the gate depletion region.

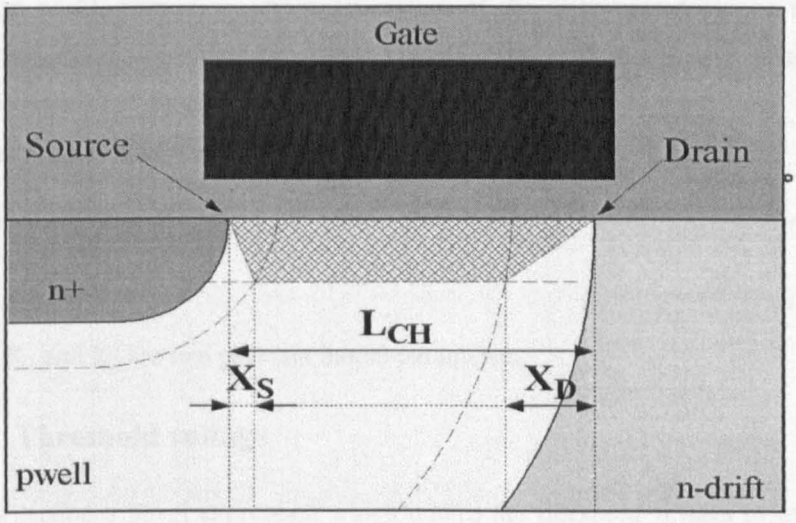


Figure 3.4: Electrostatic model of the intrinsic MOSFET.

In Equation 3.4, ε is the permittivity, q is the elementary charge, Ψ_B is the bulk potential and $N_{A,eff}$ is an effective p-region concentration which takes into account the effect of the varying doping profile. In Equation 3.5 L_{CH} is the channel length, X_S and X_D are the lateral extensions in the channel region of the depletion regions coming from source and drain respectively. Referring to Figure 3.4, F_S is calculated as the ratio between the total area depleted below the gate $A_{tot}=L_{CH}\cdot H$ and the trapezoidal area $A_{trapezoid}=[L_{CH}+ (L_{CH}-X_S-X_D)]\cdot H/2$, where H is the extension of the depletion region.

The following describes a novel strategy to directly relate the channel modulation ΔL with the gate depletion charge. An expression of the charge sharing factor F_S is obtained as a function of the channel modulation, relating depletion charge to ΔL through Equations 3.3.

Since the source is grounded, X_S is constant; X_D is instead a function of the voltage at the drain end of the channel and is approximated here as:

$$X_D = \frac{\Delta L}{K_1} + X_{D0}, \tag{3.6}$$

where X_{D0} is the minimum penetration of the drain depletion region in the channel, ΔL is the channel modulation and $K_1>1$ is a constant introduced to take into account the effect of the non constant doping profile. Expressing X_D as a function of ΔL permits a simple description of the drain side depletion region penetration in the channel. The factor $1/K_1$ is introduced to compensate for the over-estimation of X_D derived by using ΔL in Equation 3.6.

The factor F_S is then rewritten as:

$$F_s = 1 - \frac{\Delta L / K_1 + K_2}{2L_{CH}}, \tag{3.7}$$

$$K_2 = X_S + X_{D0}, \tag{3.8}$$

where K_1 and K_2 are two constant model parameters.

3.3.2 Threshold voltage

In this section a novel expression which relates the threshold voltage to channel modulation is described. The expression is derived by taking into account the

charge sharing factor described in section 3.3.1. Two model parameters are introduced to permit the fitting of the threshold voltage model to measured threshold voltage values.

From classic theory [46], the threshold voltage expression results as:

$$V_{TH} = V_{T0} + (F_S - 1)\gamma\sqrt{2\Psi_F}, \tag{3.9}$$

with

$$V_{T0} = V_{FB,eff} + 2\Psi_F + \gamma\sqrt{2\Psi_F}, \tag{3.10}$$

$$\Psi_F = \frac{K_B T}{q} \ln \frac{N_{A,eff}}{n_i}, \tag{3.11}$$

$$\gamma = \sqrt{2\varepsilon q N_{A,eff}} / C_{OX}, \tag{3.12}$$

$$C_{OX} = \varepsilon / t_{OX} \tag{3.13}$$

where Ψ_F is the Fermi potential, $V_{FB,eff}$ is an effective p-region flat-band voltage dependent on $N_{A,eff}$ and gate material, K_B is the Boltzmann constant, t_{OX} is the gate oxide thickness, T is the temperature in Kelvin and n_i is the intrinsic carrier concentration. V_{T0} represents the ideal threshold voltage in linear region of operation, corresponding to a condition in which source and drain depletion regions can be neglected. However, this ideal voltage is not directly measurable.

In order to facilitate the parameter extraction Equation 3.9 is rewritten by substituting F_S with Equation 3.7, resulting in the novel threshold voltage expression

$$V_{TH} = V_{TH\,max} - \frac{\Delta L / K_1}{2L_{CH}} \gamma\sqrt{2\Psi_F}, \tag{3.14}$$

with

$$V_{TH\,max} = V_{FB,eff} + 2\Psi_F + \gamma\sqrt{2\Psi_F}\left(1 - \frac{K_2}{2L_{CH}}\right), \tag{3.15}$$

where V_{THmax} is the measurable maximum threshold voltage displayed by the device. The value of V_{THmax} can be extracted as the threshold voltage of the intrinsic MOSFET in linear region. V_{TH} is the intrinsic MOSFET threshold voltage in saturation.

3.3.3 Inversion charge

This section describes simple models for the channel inversion charge in the sub-threshold and on-state of the intrinsic MOSFET. Three parameters are introduced to permit a straightforward fitting of the inversion charge. The continuity of the derivative of the charge is guaranteed by using two additional model parameters and an empirical smoothing function.

Inversion charge below threshold

The sub-threshold inversion charge model is derived by considering that the inversion charge concentration n for $\Psi_F<\Psi_S<2\Psi_F$ can be expressed as:

$$n = n_i e^{(\Psi_S-\Psi_F)/(K_B T)}, \tag{3.16}$$

Integrating the three-dimensional charge density n results in an effective inversion charge which can be expressed as [47]:

$$Q_{inv,av} = Q_0 e^{(V_{GS}-Vth)/(K_3 K_B T)}, \tag{3.17}$$

where $K_3>1$ is a fitting parameter and Q_0 is the inversion charge at the device threshold voltage V_{th}.

Inversion charge above threshold

In this section, the average inversion charge in the channel region is derived from the level 3 MOSFET drain current expression as:

$$Q_{inv,av} = \frac{I_{DS}}{v_{av}} + Q_0, \tag{3.18}$$

$$v_{av} = \mu_{eff} E_{long,av}, \tag{3.19}$$

$$E_{long,av} = \Delta V_{CH}^{*} / (L_{CH} - \Delta L), \tag{3.20}$$

$$\Delta V_{CH}^{*} = \min(V_{Dsat}, \Delta V_{CH}), \tag{3.21}$$

where v_{av} is the average carrier velocity in the channel, μ_{eff} is the effective channel mobility, $E_{long,av}$ is the average longitudinal electric field in the channel, L_{CH} is the channel length, V_{Dsat} is the channel pinch off voltage and ΔV_{CH} is the voltage drop across the inverted channel.

The average inversion charge in the channel is:

$$Q_{inv,av} = WC_{OX}[V_{GS} - V_{TH} - \frac{a\Delta V_{CH}^{*}}{2}] + Q_0, \tag{3.22}$$

$$a = (1 + \frac{F_S \gamma}{2K_4\sqrt{2\psi_F}}), \tag{3.23}$$

where $K_4>1$ is a correction factor introduced to avoid the overestimation of the inversion charge [46][48]. The expression of V_{Dsat} results in:

$$V_{Dsat} = V_{Dsat,1} + B - \sqrt{(V_{Dsat-1})^2 + B^2}, \tag{3.24}$$

$$V_{Dsat,1} = \frac{V_{GS} - V_{TH}}{a}, \tag{3.25}$$

$$B = E_{sat} / L_{CH}. \tag{3.26}$$

Equation 3.25 could also be used to calculate V_{Dsat}, but it would produce over-estimated values [46].

The expression used here is instead

$$V_{Dsat} = \frac{V_{GS} - V_{TH}}{aK_5}, \tag{3.27}$$

where $K_5>1$ is introduced to provide a smaller pinch-off voltage estimation with respect to Equation 3.25. To our knowledge, this is a novel approach. The use of K_5 and Equation 3.27 permits a simplified parameter extraction in comparison with the use of Equation 3.24: the average inversion charge in saturation becomes independent of a, allowing for an independent fit of the linear and saturation regions.

Smooth transition

In order to avoid derivative discontinuities in the modelled inversion charge the voltage drop across the inverted channel is empirically modelled as [49][50]:

$$\Delta V_{CH,smooth} = \Delta V_{CH}^{*}\left(1 - \frac{\ln[1 + e^{K_6(1 - V_{Dsat}/\Delta V_{CH})}]}{\ln[1 + e^{K_6}]}\right), \tag{3.28}$$

thus providing a smooth transition between saturation and linear regions in correspondence with V^{*}_{GS}, the gate to source voltage marking the transition between saturation and linear regions of operation. The transition velocity can be adjusted by varying the parameter K_6.

The continuity of the charge derivative at the transition from sub-threshold to saturation region can be guaranteed by imposing:

$$K_3 = \frac{Q_0}{K_B T W C_{OX}(1-\frac{1}{2K_5})}. \tag{3.29}$$

3.3.4 Transconductance expression

This section describes a novel formulation for the transconductance peak value. Expressing the current as:

$$I_{DS} = Q_{inv,av} v_{av}, \tag{3.30}$$

the transconductance results in

$$g_m = \left.\frac{\partial I_{DS}}{\partial V_{GS}}\right|_{V_{DS}=Const.} = g_{m1} + g_{m2}, \tag{3.31}$$

with

$$g_{m1} = \left.\frac{\partial Q_{inv,av}}{\partial V_{GS}}\right|_{V_{DS}=Const.} v_{av} \equiv Q'_{inv,av} v_{av}, \tag{3.32}$$

$$g_{m2} = Q_{inv,av} \left.\frac{\partial v_{av}}{\partial V_{GS}}\right|_{V_{DS}=Const.} = Q_{inv,av} v'_{av}. \tag{3.33}$$

Using Equations 3.22, 3.23 and 3.27, an approximated expression of the transconductance peak can be derived by rewriting the inversion charge in saturation:

$$Q_{inv,av-sat} = WC_{OX}(V_{GS} - V_{TH})(1-\frac{1}{2K_5}) + Q_0. \tag{3.34}$$

Equation 3.34 is then substituted into Equations 3.31-3.33, yielding

$$g_{m1} = WC_{OX} v_{av}(1 - V'_{TH})(1-\frac{1}{2K_5}), \tag{3.35}$$

$$g_{m2} = WC_{OX} v'_{av}(V_{GS} - V_{TH})(1-\frac{1}{2K_5}), \tag{3.36}$$

where V_{TH}' and v_{av}' are the derivatives of V_{TH} and v_{av} along V_{GS} respectively.

When the average carrier velocity is close to saturation, the variation of v_{av} and therefore its derivative will be small; in such conditions the total transconductance can be approximated as:

$$g_m \cong g_{m1} = g_{m.ideal} R_F , \tag{3.37}$$

$$R_F = (1 - \frac{1}{2K_5})(1 - V'_{TH}) , \tag{3.38}$$

where R_F represents a reduction factor with respect to the ideal transconductance case given in Equation 3.1.

A simplified version of Equation 3.38 can be obtained considering the threshold voltage expression in Equations 3.14 and the two piece linear approximation of the channel length modulation that will be described in section 3.4.1, Equations 3.45 and 3.46. Hence

$$V_{th} = V_{TH}\big|_{\Delta L=\Delta L_0} = V_{TH\max} - \frac{\Delta L_0 / K_1}{2L_{CH}} \gamma\sqrt{2\Psi_F} , \tag{3.39}$$

$$V_{TH} = V_{th} + \frac{\Delta L_0 / K_1}{2L_{CH}} \frac{(V_{GS} - V_{th})}{V^*_{GS} - V_{th}} \gamma\sqrt{2\Psi_F} , \tag{3.40}$$

$$V_{TH} = V_{th} - \frac{(V_{TH\max} - V_{th})}{V^*_{GS} - V_{th}} (V_{GS} - V_{th}) , \tag{3.41}$$

$$V'_{TH} = \frac{(V_{TH\max} - V_{th})}{V^*_{GS} - V_{th}} , \tag{3.42}$$

yielding the final peak transconductance expression

$$g_{m,\max} \cong g_{m.ideal} (1 - \frac{1}{2K_{5,eff}}) \frac{V^*_{GS} - V_{TH\max}}{V^*_{GS} - V_{th}} . \tag{3.43}$$

3.4 Results and discussion

In this section simulation and modelling results are present in three phases. First, MEDICI simulations are used to assess the amount of channel modulation, the voltage drop across the intrinsic MOSFET, the carrier velocity and the electric

field in the channel and the drift resistance. In the second phase, the extracted data is used to verify the applicability of the modelling equations to represent the behaviour of the intrinsic MOSFET of the considered Si RF power VDMOSFET. In the third phase the model is used to provide a physical insight into the transconductance behaviour in the various regions of conduction.

3.4.1 Device analysis

Channel length modulation

The amount of channel modulation ΔL is extracted by verifying the strong inversion condition along the channel length:

$$\left(F_{inv} = \Psi_S - V_D - V_{FB} - 2\Psi_B\right) > 0, \tag{3.44}$$

where V_D is the potential along the channel, Ψ_S is the surface potential and F_{inv} is here referred to as the inversion function, determined via MEDICI simulations along the channel length. The values of Ψ_S and V_D, have been extracted at several points along the channel length at several V_{GS} values at the operating V_{DS} voltage of 28V. The bias independent Ψ_B and V_{FB} have been extracted in the off-state imposing V_G=V_D=0V and for a small V_S>0V; such a bias permits the reduction of the drain and source side depletion region penetration in the channel and a more exact extraction.

The inversion function F_{inv} is positive in correspondence with parts of the channel region which are in strong inversion; solving F_{inv}=0 permits the determination of the inverted channel boundaries and subsequently the determination of the position of the pinch-off point and the amount of channel modulation. Figure 3.5 shows the F_{inv}=0 contours extracted from device simulations. Before reaching the threshold voltage, an inverted region at the centre of the gate is observed. There is no inversion at the drain and source ends of the channel region because of the effect of the drain potential and the high doping concentrations respectively. For increasing V_{GS}, the source and drain side inversion region boundaries extend towards the end of the channel region. At the threshold voltage all the source side channel region is strongly inverted, whereas channel modulation is observed at the

drain side. Channel modulation reduces for higher V_{GS}, becoming zero for $V_{GS}>V^*_{GS}$.

The extracted channel modulated length is shown in Figure 3.6 for $V_{GS}>V_{th}$, where ΔL reduces for increasing V_{GS} values.

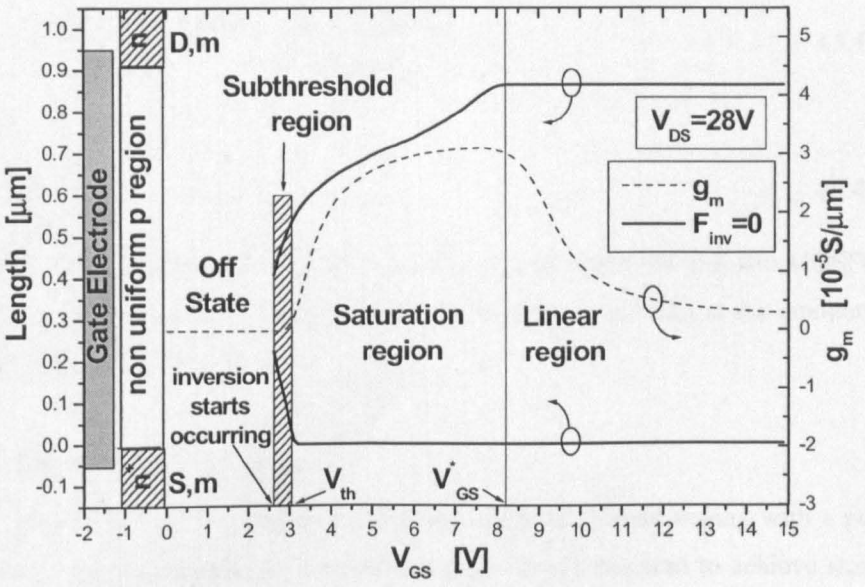


Figure 3.5: Inversion region boundaries as a function of V_{GS} for V_{DS}=28V.

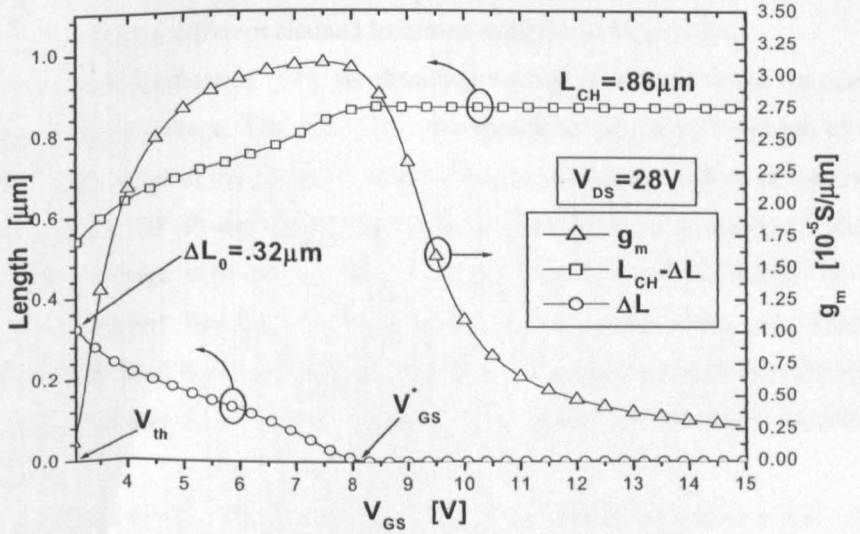


Figure 3.6: Length of the portion of the channel which remains in strong inversion as a function of V_{GS} for V_{DS}=28V.

Channel modulation reduces from a maximum ΔL_0 value at the threshold voltage to zero as the intrinsic MOSFET enters the linear region of operation at $V_{GS}=V^*_{GS}$. Figure 3.6 also shows that the amount of channel modulation ΔL can be approximated as a two piece linear function

$$\Delta L = \begin{cases} m(V_{GS} - V_{th}) + \Delta L_0, & V_{th} < V_{GS} \le V^*_{GS} \\ 0, & V_{GS} > V^*_{GS} \end{cases}, \tag{3.45}$$

with

$$m = \frac{-\Delta L_0}{V^*_{GS} - V_{th}}, \tag{3.46}$$

where V_{GS}^* represents the gate to source voltage at which the intrinsic MOSFET transits from saturation to linear region of conduction and ΔL_0 is the amount of channel modulation at V_{th}.

Threshold voltage extraction

RF Power DMOSFET devices display non uniform channel doping with a peak value at the source end. As a result, the gate voltage required to achieve strong inversion is not constant along the channel, but varies according to the doping concentration. This is acknowledged in Figure 3.5, where strong inversion is shown to occur in different channel locations at different V_{GS} values.

In a classic simplification [51], the threshold voltage is defined as the maximum local threshold voltage. Therefore, V_{th} corresponds to the strong inversion of the peak doping point in the channel. On the other hand, the complexity of threshold modeling for RF Power DMOSFET makes it difficult to precisely define a threshold voltage extraction method. Usually, RF Power DMOSFET device datasheets report threshold voltage values in saturation which are linearly extrapolated from the I_{DS}-V_{GS} characteristic. The values are typically extracted with the gate and drain terminal shorted to ensure that the device is operated in saturation.

In order to extract the threshold voltage at the considered application voltage V_{DS}, the Extrapolation method in the Saturation Region (ESR) described in [52] has been applied. The method consists of calculating the threshold voltage as the

offset of the line tangent to the $I_{DS}^{0.5}$-V_{GS} curve at the application voltage V_{DS} where the slope of the curve peaks. The accuracy has been verified with the $V_{th,max}$ determination in Figure 3.5, where $V_{th,max}$ is the gate to source voltage at which the inversion region boundary reaches the source.

Voltage drop across the intrinsic MOSFET and drift region resistance

The voltage drop across the channel as extracted from the two dimensional simulation results is reported in Figure 3.7, highlighting the potential at the drain ($V_{D,m}$) and at the source ($V_{S,m}$) end of the channel. ΔV_{CH} increases with V_{GS} up to small values of V_{GS} above V_{th}; it then decreases for higher values of V_{GS}.

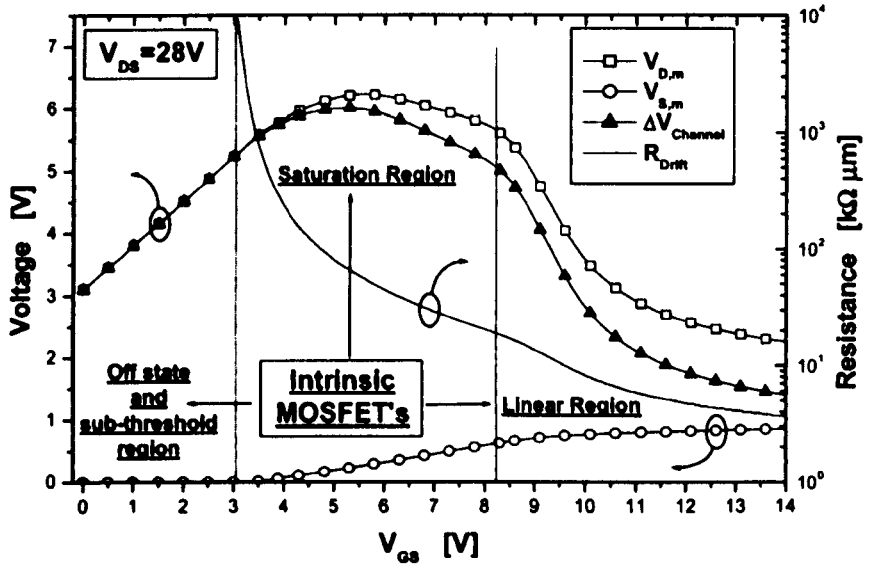


Figure 3.7: Voltage drop across the intrinsic MOSFET and calculated drift region resistance.

For small current levels and in the off state the drain of the intrinsic MOSFET can be considered floating: its potential is strongly affected by the gate potential and $V_{D,m}$ increases with V_{GS}. For increasing V_{GS} values above V_{th} a pinched-off channel is formed. The pinch-off point moves closer to the drain end of the channel and a larger current will flow through the intrinsic MOSFET drain. The influence of the gate potential reduces and $V_{D,m}$ decreases through the effect of the drift resistance. For higher V_{GS} values, the intrinsic MOSFET is biased in the linear region and all the current flows through the drain end of the channel. The

gate potential effect can be neglected and the intrinsic drain potential depends only on the drift resistance:

$$V_{D,m} = V_{DS} - I_{DS} R_{Drift}. \tag{3.47}$$

The non linear drift region resistance R_{Drift} is shown in Figure 3.7. R_{Drift} is calculated from Equation 3.47 from the simulated $V_{D,m}$ and I_{DS} values at the V_{DS}=28V. R_{Drift} progressively decreases for increasing V_{GS} respecting the inverse proportionality with I_{DS} acknowledged in Equation 3.47.

Carrier velocity in the channel

Carrier velocity is not constant along the channel length due to the varying channel doping. In order to include carrier velocity in a compact model formulation, the average or effective values need to be used. The average values are obtained from local, position dependent, simulation data as

$$v_{av} = \frac{1}{L_{CH} - \Delta L} \int_{0}^{L_{CH} - \Delta L} v(x)dx, \tag{3.48}$$

where v(x) is the position dependent carrier velocity extracted along the channel length. The application of Equation 3.48 yields the average carrier velocity reported in Figure 3.8.

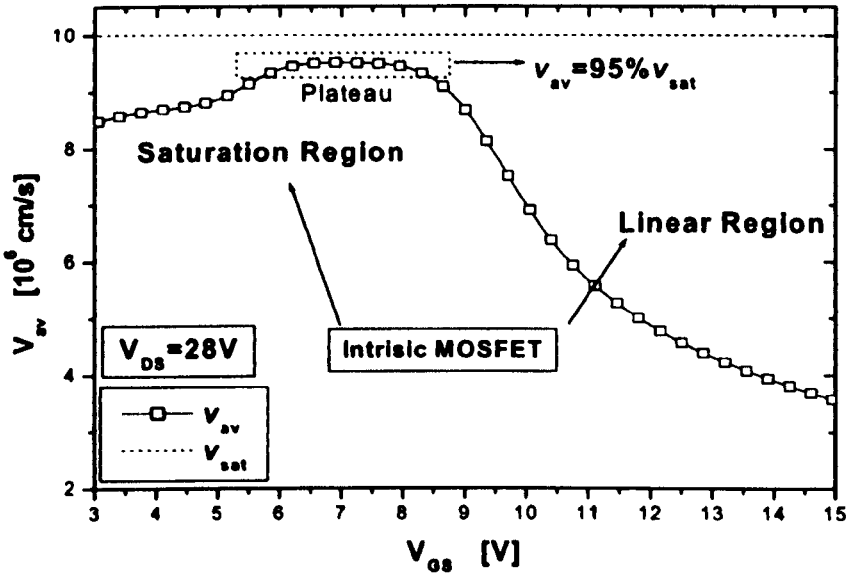


Figure 3.8: Average carrier velocity in the channel.

When the intrinsic MOSFET is biased in saturation the average carrier velocity remains above 85% of v_{sat}, indicating saturated velocity behaviour. In fact, it can be shown that the average longitudinal electric field in the channel $E_{av}=\Delta V/L_{CH}$ is at its saturation limit, which in silicon is around $5\cdot10^4$V/cm [53]. For increasing V_{GS} above threshold, v_{av} progressively increases until reaching a plateau close to the saturation velocity value. For higher gate to source voltages the intrinsic MOSFET is biased in the linear region with a progressive reduction of the voltage drop across the channel. This causes a reduction of the average longitudinal electric field in the channel, yielding a reduction of carrier velocity.

3.4.2 Model extraction and verification

The model parameters are extracted considering the I_{DS} versus V_{GS} characteristic at V_{DS}=28V and the observed threshold voltage value. The extraction is carried out in 7 steps:

1. V_{THmax} is extracted as the threshold voltage in linear region.
2. V_{th} is extracted as the threshold voltage at V_{DS}=28V.
3. $N_{A,eff}$ and K_2 are extracted by matching V_{THmax} with Equation 3.15.
4. K_1 is determined by fitting the modelled threshold voltage V_{TH} to V_{th} at $V_{GS}=V_{th}$.
5. K_4 and $K_{5,eff}$ are extracted by fitting the I_{DS}-V_{GS} trans-characteristic in the linear and saturation regions of conduction respectively.
6. K_6 is determined in order to achieve a smooth transition between linear and saturation regions of conduction. K_5 is adjusted to maintain a good saturation region fit.
7. K_3 is calculated according to Equation 3.29.

The application of steps 1 to 4 yields the definition of the threshold voltage V_{TH}. V_{TH} is not constant during device operation. In fact, during operation at V_{DS}=28V, the voltage drop across the intrinsic MOSFET channel ΔV_{CH} varies. This affects channel modulation and therefore the amount of the depletion charge controlled by the gate electrode, yielding variations of the threshold voltage. The minimum threshold voltage $V_{TH}=V_{th}$ is observed when the channel modulation is maximum

($\Delta L=\Delta L_0$) at $V_{GS}=V_{th}$. The maximum threshold $V_{TH}=V_{THmax}$ voltage is observed when the intrinsic MOSFET is biased in the linear region.

In step 5, the K_4 and K_{5eff} values are determined through an independent fit of the current characteristic in the linear and saturation regions of operation respectively. As shown Figure 3.9, the extracted K_4 and K_{5eff} values permit a good current match but determine an abrupt transition in the modelled current in correspondence with the transition of the intrinsic MOSFET from saturation to linear region of operation. The abrupt transition is a consequence of the abrupt variation of ΔV_{CH}^*.

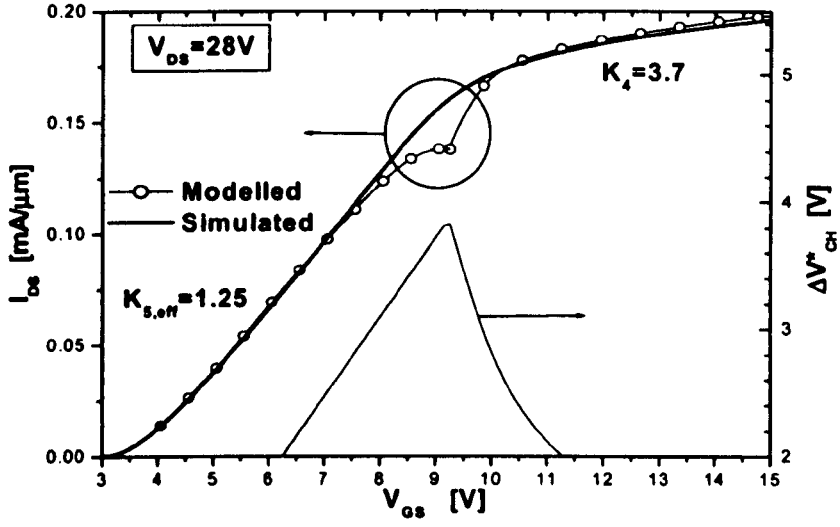


Figure 3.9: Extraction of K_4 and $K_{5,eff}$ via independent fitting of the output characteristic.

When K_6 is introduced in step 6, the abrupt transition in the current is removed. This result is achieved by providing a smooth approximation for the ΔV_{CH}^* determined during step 5. The introduction of the smoothing function leads to a reduction of K_5 with respect to the value of $K_{5,eff}$ determined in step 5. K_6 is finally extracted to fit the sub-threshold region current. The application of steps 5 to 7 provides a good prediction of the I_{DS} versus V_{GS} characteristic and the transconductance, as shown in Figure 3.10. The model proposed in section 3.3 represents an acceptable compromise between simplicity and accuracy, while providing physical insight into the intrinsic MOSFET behaviour in RF Power MOSFETs.

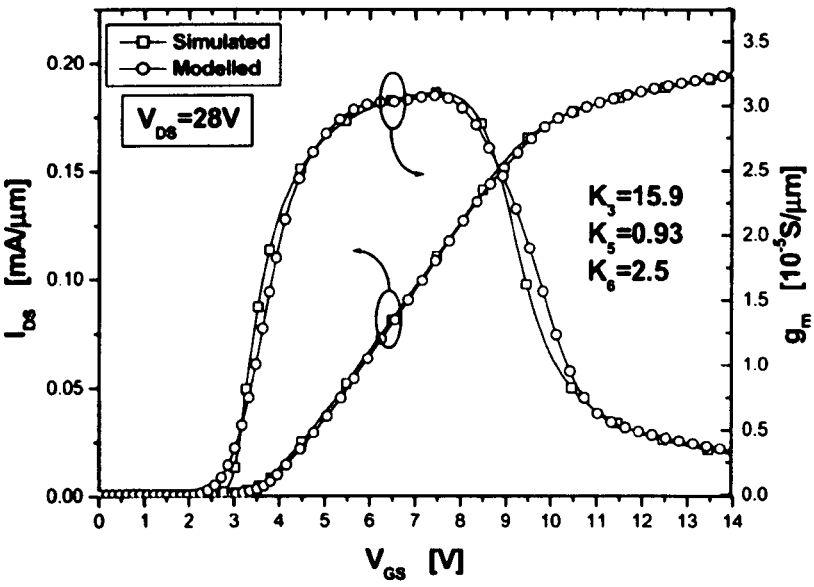


Figure 3.10: Modelled and MEDICI simulated transconductance and transfer characteristic.

The model has also been verified on measured data. Figure 3.11 shows modelled versus measured I_{DS}-V_{GS} and transconductance at V_{DS}=28V for a Polyfet SP204 RF power MOSFET.

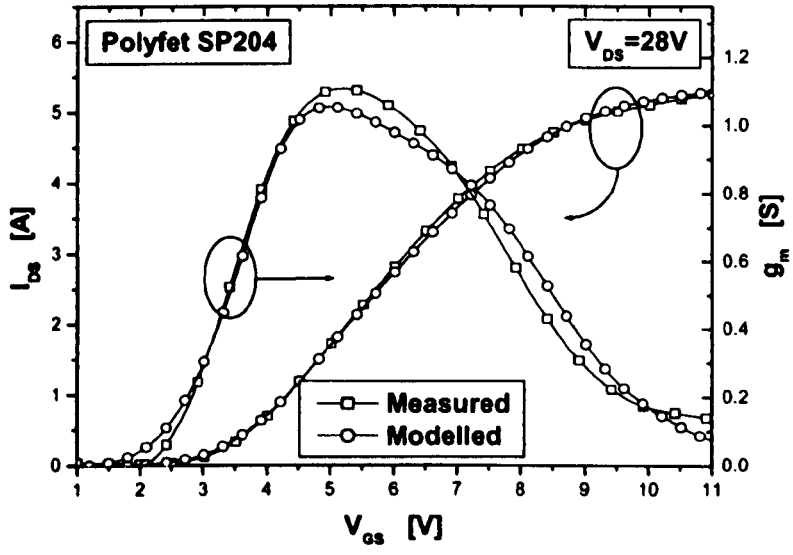


Figure 3.11: Modelled and measured transconductance and transfer-characteristic of a Polyfet SP204 device.

Measured IV characteristics have been used for the fitting of the model. In order to make possible the extraction of the model parameters it has been required to extrapolate information from measurements in several steps:

1- The I_{DS}-$V_{D,m}$ characteristics of the intrinsic MOSFET have been extracted. The JFET approach described in [4] has been used for the determination of $V_{D,m}$.

2- The I_{DS}-$V_{D,m}$ characteristics have been used to empirically estimate channel modulation via the determination of the Early voltage [43].

3- The carrier channel velocity has been estimated by using an empirical mobility model [54] and approximating the longitudinal electric field in the channel as $E_{long}=V_{D,m}/L_{CH}$.

3.4.3 Transconductance analysis

Using Equations 3.32 and 3.33 the two transconductance components have been calculated from the extracted v_{av} and $Q_{inv,av}$ values; the transconductance and its components g_{m1} and g_{m2} are reported in Figure 3.12. Equations 3.31 to 3.33 are rewritten here for convenience

$$g_m = \left.\frac{\partial I_{DS}}{\partial V_{GS}}\right|_{V_{DS}=Const.} = g_{m1} + g_{m2}, \tag{3.49}$$

$$g_{m1} = \left.\frac{\partial Q_{inv,av}}{\partial V_{GS}}\right|_{V_{DS}=Const.} v_{av} \equiv Q'_{inv,av} v_{av}, \tag{3.50}$$

$$g_{m2} = Q_{inv,av} \left.\frac{\partial v_{av}}{\partial V_{GS}}\right|_{V_{DS}=Const.} = Q_{inv,av} v'_{av}. \tag{3.51}$$

The transconductance can be ideally portioned into three main parts, corresponding to the three operating regions that the intrinsic MOSFET is biased in for increasing values of V_{GS}: sub-threshold, saturation and linear region.

In the sub-threshold region, for $V_{GS}<V_{th}$, the transconductance increases from zero to about 5-10% of the peak g_m value. In this region, the inverted charge in the channel and its derivative both increase exponentially, whereas the carrier velocity can be considered constant. This yields $g_{m2}=0$ and $g_m=g_{m1}$, as shown in Figure

3.12. Since the inversion charge dependence on V_{GS} is exponential, the transconductance will increase exponentially, explaining the sudden rise of the g_m characteristic.

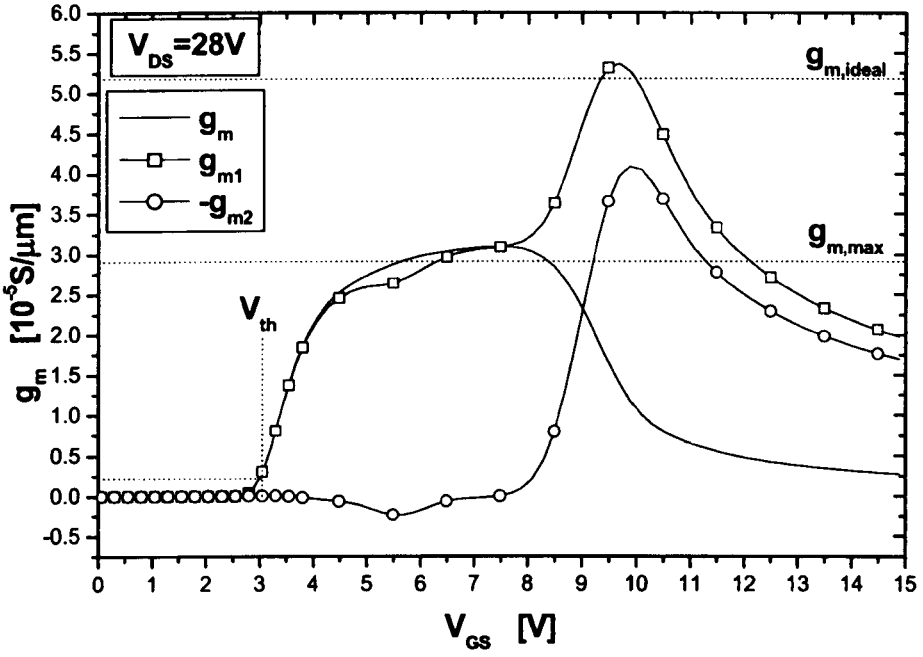


Figure 3.12: Total transconductance and its components, highlighting the ideal transconductance limit and the approximated $g_{m,max}$ value calculated using Equation 3.43.

In the saturation region, for $V_{th}<V_{GS}<V_{GS}^*$, the carrier velocity is saturated. As shown in Figure 3.12, small variations in velocity yield the small amplitude of the g_{m2} term which becomes zero in correspondence with the peak transconductance value, where $v_{av}'\approx 0$. The small amplitude of g_{m2} in saturation permits an approximation of the transconductance as $g_m \approx g_{m1}$, thus satisfying the basic assumption for the derivation of Equation 3.43.

Figure 3.12 also shows that the value of the transconductance peak is well predicted by Equation 3.43. The estimated error is around 6%, a considerable improvement with respect to the classic ideal g_m limit formula.

Although carrier velocity in the channel is saturated, the peak transconductance is considerably smaller than the ideal limit in case of velocity saturation. Since Equation 3.43 has been derived considering the effect of channel modulation, the

good prediction of the g_m peak implies that the reduction of the transconductance with respect to the ideal limit is due to the effect of channel modulation on the inversion charge in the channel. The same deduction applies to the entire flat region of the transconductance characteristic.

This implies that:

- The transconductance is limited by the rate at which inversion charge in the channel increases with the gate to source voltage ($Q'_{inv,av}$).
- Channel length modulation affects $Q_{inv,av}$ and $Q'_{inv,av}$, yielding $Q'_{inv,av}<WC_{OX}$.
- The ideal transconductance limit can be approached if channel length modulation involves a small part of the channel, that is when $Q'_{inv,av}\approx WC_{OX}$.

Therefore, more than the transconductance ideal limit, it is more appropriate to talk about the inversion charge limit.

In the linear region, for $V_{GS}^{*}<V_{GS}$, the reduction of ΔV_{CH} for increasing V_{GS} values results in a reduction of the longitudinal electric field and of the carrier velocity in the channel. This causes negative v_{av}' and g_{m2}. For small V_{GS} values above V_{GS}^{*}, the sudden reduction of the average carrier velocity leads $|g_{m2}|$ to increase faster than g_{m1}, causing the transconductance fall-off. For higher V_{GS} values, in correspondence with the compression of the output current, $|g_{m2}|$ and g_{m1} decrease appreciatively at the same rate.

Finally, Equation 3.43 is also useful in explaining the effect of the drift region resistance on the peak transconductance value reported in [55] and often observable during RF device optimisation. In fact, the peak g_m value is shown to be indirectly related to the drift resistance value, through the R_{drift} effect on V^{*}_{GS}. Since $g_{m,ideal}$, K_5 V_{THmax} and V_{th} depend only on the intrinsic MOSFET design, Equation 3.43 implies a relation of the peak transconductance value with the drift region resistance through V_{GS}^{*}. Therefore, since large drift resistances are associated with smaller transconductance fall-off points V^{*}_{GS}, Equation 3.43 clarifies the observed reduced g_m peaks associated to higher drift resistance values.

3.5 Summary

A compact model for a Si RF Power DMOS transistor that provides useful insight into the device physical behaviour is presented. In the model, channel modulation is directly linked to the variation of depletion charge in the channel and a novel expression for the charge sharing factor is presented. The model also includes a modified Level 3 inversion charge description, where smooth transitions between operating regions have been imposed, a simplified drain saturation voltage expression has been used and a physical, non fitted velocity saturation value considered. The proposed model has been shown to permit a good agreement with MEDICI simulations and measured data.

Based on the proposed model, a novel expression for the transconductance has also been proposed.

In saturation, although carrier velocity in the channel has been shown to be saturated, a large reduction of the peak value of the transconductance is observed with respect to the ideal transconductance limit. Consequently, it has been concluded that the reduction of the transconductance with respect to the ideal value and the transconductance flattening are mainly caused by the appreciatively linear increase of the channel inversion charge with the gate to source voltage.

Following a dual-piecewise linear approximation of channel modulation, a novel equation for the peak transconductance has been proposed. The formula permits to relate the effect of the drift region resistance on the peak transconductance value.

References

[1] A. Wood, W. Brakensiek, C. Dragon, and W. Burger, "120 Watt, 2 GHz, Si LDMOS RF Power Transistor for PCS Base Station Applications," IEEE MTT-S Digest, 1998, pp.707-710

[2] N. Nenadovic, V. Cuoco, S.J.C.H. Theeuwen, H. Schellevis, G. Spierings, A. Griffo, M. Pelk, L.K. Nanver, R.F.F. Jos, and J.W. Slotboom, "RF Power Silicon-On-Glass VDMOSFETs," IEEE Electron Device Letters, 2004, vol. 25, pp. 424–426.

[3] S.K.Leong, "VDMOS vs. LDMOS – How to choose," 2001 IEEE MTT-S International Microwave Symposium

[4] M. Trivedi, P. Khandelwal, and K. Shenai, "Performance Modeling of RF Power MOSFET's," IEEE Trans. Electron Devices, 1999, vol. 46, pp. 1794-1801.

[5] W. Curtice, J. Plá, D. Bridges, T. Liang, and E. Shumate, "A New Dynamic Electro-Thermal Nonlinear Model for Silicon RF LDMOS FETs," 1999 IEEE MTT-S International, pp. 419-422.

[6] J. Voulevi, J. Aikio, and T. Rahkonen, "Extraction of a Polynomial LDMOS Model for Distortion Simulations Using Small-Signal S-parameter Measurements," 2002 IEEE MTT-s Digest, pp. 2157-2160.

[7] V. Cuoco, M.P. v.d. Heijden, and L.C.N. de Vreede, "The "Smoothie" Data Base Model for the correct Modeling of Non-Linear Distortion in FET Devices," 2002 IEEE MTT-S Digest, pp.2149-2152.

[8] S. Akhtar, P. Roblin, S. Lee, X. Ding, S. Yu, J. Kasick, and J. Strahler, "RF Electro-Thermal Modeling of LDMOSFETs for Power-Amplifier Design," IEEE Trans. Microwave Theory and Techniques, 2002, pp. 1561-1570.

[9] H.J. Sigg, G.D. Vendelin, T.P. Cauge, and J. Kocsis, "D-MOS Transistor for Microwave applications," IEEE Trans. Electron Devices, 1972, vol. 19, pp. 45-53.

[10] E. Fong, D.C. Pitzer, and R.J. Zeman, "Power DMOS for High-Frequency and Switching Applications," IEEE Trans. Electron Devices, 1980, vol. 27, pp. 322-330.

[11] H. Masuda, T. Masuhara, M. Nagata, and N. Hashimote, "Device design of E/D gate MOSFET," J. Japan Soc. Appl. Phys., 1973, vol.42, pp.167-172.

[12] T.J. Rogers, S. Asai, M.D. Pocha, R.W. Dutton, and J.D. Meindl, "An experimental and theorethical analysis of double-diffused MOS transistors," IEEE J. Solid-State Circuits, 1975, vol. 10, pp.322-331.

[13] M.Y. Hong and D.A. Antoniadis, "Theoretical Analysis and Modelling of Submicron Channel Length DMOS Transistors," IEEE Trans. Electron Devices, 1995, vol. 42, pp. 1614-1622.

[14] J. Jang, O. Tornblad, T. Arnborg, Q. Chen, K. Banerjee, Z. Yu, and R.W. Dutton, "RF LDMOS Characterization and its Compact Modeling," Microwave Symposium Digest, 2001 IEEE MTT-S International, 2001, vol. 2, pp. 967-970.

[15] R. S. Scott, G. A. Franz, and J. L. Johnson, "An accurate model for power DMOSFETs including interelectrode capacitances," IEEE Trans. Power Electron., vol. 6, no. 2, pp. 192–198, Mar. 1991.

[16] J. Jang, T. Arnborg, Z. Yu, and R. W. Dutton, "Circuit model for power LDMOS including quasi-saturation," in Proc. SISPAD, 1999, pp. 15–18.

[17] N. Hefyene, E. Vestiel, B. Bakeroot, C. Anghel, S. Frere, A. M. Ionescu, and R. Gillon, "Bias-dependent drift resistance modeling for accurate DC and AC simulation of asymmetric HV-MOSFET," in Proc. SISPAD,2002, pp. 203–206.

[18] A. C. T. Aarts, M. J. Swanenberg, and W. J. Kloosterman, "Modeling of high-voltage SOI-LDMOS transistors including self-heating," in Proc. SISPAD, 2001, pp. 246–249.

[19] J. Victory, C. C. Mc Andrew, R. Thoma, K. Joardar, M. Kniffin, S.Merchant, and D. Moncoqut, "A physically-based compact model forLDMOS transistors," in Proc. SISPAD, 1998, pp. 271–274.

[20] Y. Kim, J. G. Fossum, and R. K. Williams, "New physical insights and models for high-voltage LDMOST IC CAD," IEEE Trans. Electron Devices, vol. 38, no. 7, pp. 1641–1649, Jul. 1991.

[21] Y. Chung and D. E. Burk, "A physically based DMOS transistor model implemented in SPICE for advanced power ID TCAD," in Proc. ISPSD, 1995, pp. 340–345.

[22] M. Y. Hong and D. A. Antoniadis, "Theoretical analysis and modelling of submicron channel length DMOS transistors," IEEE Trans. Electron Devices, vol. 42, no. 9, pp. 1614–1622, Sep. 1995.

[23] Y. Chung, "LADISPICE-1.2: A nonplanar-drift lateral DMOS transistor model and its application to power IC TCAD," Inst. Elect. Eng., Proc. Circuits Devices Syst., vol. 147, no. 4, pp. 219–227, 2000.

[24] C. M. Liu, F. C. Shone, and J. B. Kuo, "A closed-form physical backgate-bias dependent quasi-saturation model for SOI lateral DMOS devices with self-heating for circuit simulation," in Proc. ISPSD, 1995, pp. 321–324.

[25] Y. Subramanian, P. O. Lauritzen, and K. R. Green, "A compact model for an IC lateral diffused MOSFET using lumped-charge methodology," in Proc. Modeling Simulation Microsystems, 1999, pp. 284–288.

[26] N. D'Halleweyn, J. Benson, M. Swanenberg, and W. Redman-White, "A compact model for silicon-on-insulator LDMOST, including accumulation, lateral doping gradient and high-side behavior," in Proc. Silicon-on-Insulator Technology Devices, 2001, pp. 187–192.

[27] N. D'Halleweyn, L. Tiemeijer, J. Benson, M. Swanenberg, and W. Redman-White, "Charge model for SOI-LDMOST with lateral doping gradient," in Proc. ISPSD, 2001, pp. 291–294.

[28] A. Aarts, N. D'Halleweyn, and R. van Langevelde, "A Surface-Potential-Based High-Voltage Compact LDMOS Transistor Model," IEEE Trans. on Electron Devices, 2005, vol. 52, pp. 999-1007.

[29] M.B. Willemsen, and R. van Langevelde, "High-Voltage LDMOS compact model for RF applications," IEDM Technical Digest, 2005, pp.208-211.

[30] J. Olsson, N. Rorsman, L. Vestling, C. Fager, J. Ankarcrona, H. Zirath, and K.-H. Eklund, "1 W/mm RF Power Density at 3.2 GHz for a Dual-Layer RESURF LDMOS Transistor," IEEE Electron Device Letters, 2002, vol. 23, pp. 206-208.

[31] J. Cai, C. Ren, N. Balasubramanian, and J.K.O. Sin, "High Performance Stacked LDD RF LDMOSFET," in Proc. ISPSD, 2001, pp. 103-106.

[32] M. Shindo, M. Morikawa, T. Fujioka, K. Nagura, K. Kurotani, K. Odaira, T. Uchiyama, and I. Yoshida, "High Power LDMOS for Cellular Base Station Applications," in Proc. ISPSD, 2001, pp. 107-110.

[33] Y.C. Liang, S. Xu, C. Ren, and P.-D. Foo, "New Partial SOI LDMOS Device with High Power-Added Efficiency for 2 GHz RF Power Amplifier Applications," in Proc. IECON, 2000, vol. 2, pp.1001-1006.

[34] P. Fioravanti, O. Spulber, M.M. De Souza, "Transconductance behaviour in 28V Si RF Power MOSFETs", submitted for publication to IEEE Transactions on Electron Devices

[35] Technology Modeling Associated Inc., MEDICI User Manual, Sunnyvale, California, USA

[36] A. Ortiz-Conde, J.J. Liou, W. Wong, and F.J. Garcia Sanchez, "Simple Method for Extracting the difference between the drain and source series resistance in MOSFETs," Electronics Letters, Vol.30, pp. 1013–1015, 1994.

[37] A. Blaum, J. Victory, and C. C. McAndrew, "A Simple Physical Extraction Method for RD – RS of Asymmetric MOSFETS," International Conference on Microelectronic Test Structures (ICMTS), 1999, Vol. 12, pp. 141-146.

[38] J.A.M. Otten, and F.M. Klaassen, "A Novel technique to Determine the Gate and Drain Bias Dependent Series Resistance in Drain Engineering MOSFET's Using One Single Device," IEEE Transactions on Electron Devices, 1996, Vol. 43, pp. 1478-1488.

[39] N. Hefyene, C. Anghel, A.M. Ionescu, S. Frere, R. Gillon, M. Vermandel, B. Bakeroot and J. Doutreloigne, "An Experimental Approach for Bias-Dependent Drain Series Resistances Evaluation in Asymmetric HV MOSFETs," European Solid-State Device Research Conference (ESSDERC, pp.403-406), 2001.

[40] N. Hefyene, C. Anghel, A.M. Ionescu, S. Frere, R. Gillon, M. Vermandel, B. Bakeroot, and J.Doutreloigne, "Bias-dependence Drift resistance

Modeling for Accurate DC and AC Simulation of Asymmetric HV-MOSFET," in Proc. SISPAD, pp.203 - 206., 2002.

[41] L. Dang, "A simple Current Model for Short-Channel IGFET and Its Application to Circuit Simulation," IEEE J. Sol. St. Circ., 1979, Vol.14, pp. 358-367.

[42] D.M. Caughey, and R.E.Thomas, "Carrier Mobilities in Silicon empirically related to doping and field," in Proc. IEEE, pp.2192-2193, Dec.1967.

[43] A. Vladimirescu, and S. Liu, "The simulation of MOS integrated circuits," 1980, Memo UCB/ERL M80/7 Electronics Laboratory, University of California.

[44] G. Merckel, J. Borel, and N.Z. Cupcea, "An accurate large-signal MOS transistor model for use in computer-aided design," IEEE Trans. Electron Devices, 1972, vol. 19, p. 681-690.

[45] R. Swanson, and J. Meindl, "Ion-Implanted Complementary MOS Transistors in Low Voltage Circuits," IEEE J. So. St. Circ., 1972, vol. 7, pp.146-153.

[46] Y.Tsivids, "Operation and Modelling of t he MOS Transistor," Second Edition, Oxford University Press, 1999.

[47] N. Arora, "MOSFET Models for VLSI Circuit Simulation," Springer-Verlag, 1993.

[48] G. Merckel, J. Borel, and N.Z. Cupcea, "An accurate large-signal MOS transistor model for use in computer-aided design," IEEE Trans. on Electron Devices, vol. ED-19, pp. 681-690, 1972.

[49] J.A. Power, and W.A. Lane, "Enhanced spice MOSFET model for analog applications including parameter extraction schemes," in Proc. IEEE int. Conf. on Microelectronic Test Structures, 1990, vol.3, pp. 129-134.

[50] C.C. McAndrew, and B.K. Bhattacharyya, "O. Wing, A single-piece C_∞-continuous MOSFET model including subthreshold conduction," IEEE Electron Device Letters, 1991, vol.12, pp.565-567.

[51] M.D. Pocha, A.G. Gonzalez, and R.W. Dutton, "Threshold voltage Controllability in Double-Diffused-MOS Transistors, " IEEE Transactions on Electron Devices, vol. ED-21, pp. 778-784, 1974.

[52] A. Ortiz-Conde, F.J. García Sánchez, J.J. Liou, A. Cerdeira, M. Estrada, and Y.Yue, "A review of recent MOSFET threshold voltage extraction methods," Microelectronics Reliability, vol. 42, pp.583-596, 2002.

[53] D.A. Grant, J. Gowar, "Power MOSFETs: Theory and Applications," 1989, Wiley.

[54] T.C. Hong, P.K. Ko, C. Hu, 50-A gate MOSFET's at 77K, IEEE Trans. Electron Devices, ED-34, 1987, pp.2129

[55] S. Xu, C. Ren, L.Y.C., P.-D. Foo, and J.K.O. Sin, "Theoretical Analysis and Experimental Characterization of the Dummy-Gated VDMOSFET," IEEE Transactions on Electron Devices, 2001, pp. 2168-2176.

4

RF Performance Assessment of Si RF Power MOSFETs

Abstract

An innovative method for the estimation of large signal performance of Si RF Power MOSFETs is presented. The amplifier load line is used in the prediction of the large signal analysis. A procedure for the extraction of the optimum load line output characteristic is proposed. Power gain prediction is demonstrated by using a novel analytical power gain expression. Based on the analytical formulation, the load line characteristic is considered, permitting the assessment of gain compression. Finally, an extension of conventional prediction of linearity is proposed for the large signal case.

4.1 Introduction

Harmonic Balance (HB) simulations are the only viable approach to provide accurate RF performance estimation of devices in various applications. In the absence of HB, analytic expressions can provide a feasible alternative. A number of power gain expressions can be found in the literature for conventional MOSFETs [7][2][6]. Unfortunately, these expressions are not appropriate for Si RF power MOSFETs due to the substantial structural differences between these two applications. The most reliable analytical approaches available to date have been proposed in [4][5][7].

In this chapter an extension of [5] to deduce matching impedances and, for the first time, gain compression is proposed. Gain compression is assessed through the identification of the load line output characteristic of the amplifier. This output characteristic is also used in the assessment of device linearity. The work reported in this chapter also appears in [6].

The organisation of this chapter is as follows: the definition and extraction procedures for the load line characteristic are described in section 4.2. A new formulation for the power gain is proposed in section 4.3. The dependence on gate resistance and drain conductance is examined and the gain expression extended to the non linear current generator case to assess gain compression. A novel approach to estimate linearity is proposed in section 4.4, where the load line output characteristic is used in the inter modulation distortion estimation. A large signal approach and the extension of the analytical expression to the large signal case are also proposed.

4.2 Load line determination

Although the load line is extensively used in RF power amplifier design, real devices present non linear behaviour that make its exact determination non trivial.

Based on conventional theory, in this section a methodology is discussed for the determination of the load line. This procedure yields a direct relationship between input voltage and output current signal of an RF power amplifier which can be used for predicting the amplifier behaviour.

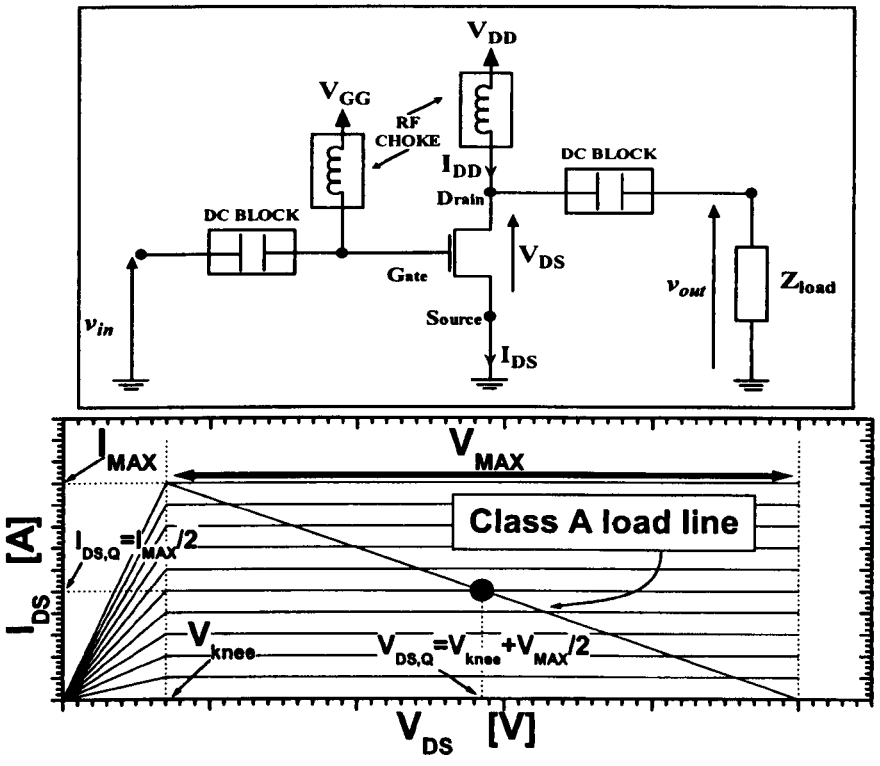


Figure 4.1: Power amplifier model and load line superimposed to the I_{DS}-V_{DS} characteristics in class A. The load line is the loci of output current and voltage that the amplifier configuration forces on the transistor. In class A the slope of the load line is $-1/R_{opt}$.

4.2.1 Background

A simple amplifier configuration is shown in Figure 4.1, where the transistor output bias is provided through an ideal RF choke, a load impedance Z_{load} is connected to the transistor through an ideal dc block capacitor and the transistor is considered as an ideal voltage controlled current source generator. The relationship between the output current and voltage of the power transistor can be expressed as

$$V_{DS} = V_{DD} + v_{out}, \quad (4.1)$$

$$I_{DS} = I_{DD} + \frac{v_{out}}{Z_L} = I_{DD} + \frac{V_{DS} - V_{DD}}{Z_L}. \quad (4.2)$$

Equation 4.2 represents a constraint on the transistor's output current and voltage that is imposed by the circuit. For a periodic input signal v_{in}, the steady state output current $i_{out}=I_{DS}-I_{DD}$ and voltage v_{out} will define a stationary trajectory around the device quiescent point. This trajectory is usually referred to as the load cycle.

A particular case of load cycle is obtained when a purely resistive load $Z_L=R_L$ is considered. This particular load cycle is the load line. When superimposed on the device I_{DS}-V_{DS} characteristics the load line is a line passing through the device quiescent point $(V_{DS,Q}, I_{DS,Q}) = (V_{DD}, I_{DD})$.

The concept of load line is particularly important in power amplifier design, where maximum linear power corresponds to a purely resistive load for the device current generator [4]. A resistive load for the device current generator is imposed by opportunely adjusting the device matching impedance in order to compensate for the internal device reactance as well as for the package parasitics. The value of this resistive load is chosen in order to yield maximum output current and voltage swings, therefore maximum linear power. Depending on the device bias, the optimum resistance value R_{opt} is determined considering the conduction angle α [4]:

$$\alpha = \arccos\left(-\frac{2I_{DS,Q}}{I_{MAX} - I_{DS,Q}}\right), \quad (4.3)$$

$$R_{opt} = \frac{V_{MAX}}{I_{MAX}} \frac{2\pi(1-\cos(\alpha/2))}{\alpha - \sin(\alpha)}, \quad (4.4)$$

where the conduction angle is a measure of the portion of the output signal in which conduction occurs (i.e. the input signal is above the device threshold voltage), and $V_{MAX}/2$ and $I_{MAX}/2$ are the amplitudes of the maximum voltage and current swings respectively. I_{MAX} corresponds to the current amplitude at the knee of the I_{DS}-V_{DS}. When the maximum voltage swing is limited by the device breakdown voltage V_{BR}, V_{MAX} is generally calculated as $V_{MAX}= V_{BR} - V_{knee}$.

However, in Si RF Power devices, the voltage swing is limited by the knee voltage and the application drain voltage, yielding

$$V_{MAX} = 2(V_{DS,Q} - V_{knee}). \quad (4.5)$$

4.2.2 Load line determination

Following the definition of an ideal load line, in a well designed linear power amplifier, the optimum resistance R_{opt} is imposed on the current generator causing the device load cycles to flatten onto the load line. This is an important result, since the load line can be easily extracted from static device measurements. In fact, the optimum load line is the line through the knee point of the I_{DS}-V_{DS} characteristic (V_{knee},I_{max}) and the bias point ($V_{DS,Q}$,$I_{DS,Q}$), as shown in Figure 4.2.

The superimposition of the load line on the I_{DS}-V_{DS} curves of the device permits the extrapolation of the output characteristic $I_{loadline}$ along the trajectory defined by the load line itself, as shown in Figure 4.2. The extraction is required because the load line, mathematically defined by Equation 4.2 for Z_L=R_{opt}, does not explicitly relate input (V_{in}) to output (I_{out}) of the power amplifier. The goal is to obtain an expression for the current along the load line as $I_{loadline}$=f(V_{GS}). As shown in Figure 4.2, this is achieved by extrapolating the intersection point between the I_{DS}-V_{DS} characteristics and load line for varying V_{GS} values.

A prerequisite for the correct determination of the load line characteristic is represented by the accurate extraction of the V_{knee} value to be used. However, when a non ideal RF power device is considered, even the identification of the knee voltage V_{knee} is problematic.

This task should not be carried out by visual examination of the output characteristics of the device. In the I_{DS}-V_{DS} characteristics reported in Figure 4.1, although a range of values can be graphically identified as suitable drain to source voltages, an accurate and easy V_{knee} determination is not possible.

The following describe a numerical approach for the accurate determination of V_{knee}. An accurate extraction of the knee voltage is essential to maximise the ideal maximum linearly delivered power P_{LIN}

$$P_{LIN} = 1/8 V_{MAX} I_{MAX} . \quad (4.6)$$

Since I_{MAX} has been defined as the current level at the knee of the I_{DS}-V_{DS} characterisitics, it can be expressed as

$$I_{MAX} = f(V_{GS,MAX}, V_{knee}), \quad (4.7)$$

where $V_{GS,MAX}$ is the maximum allowed gate voltage for the device, the determination of V_{knee} corresponds to finding the value that maximizes

$$P_{LIN} = 1/4[(V_{DS,Q} - V_{knee}) f(V_{GS,MAX}, V_{knee})] . \quad (4.8)$$

Equation 4.8 is a non linear but one dimensional problem that can be easily numerically solved.

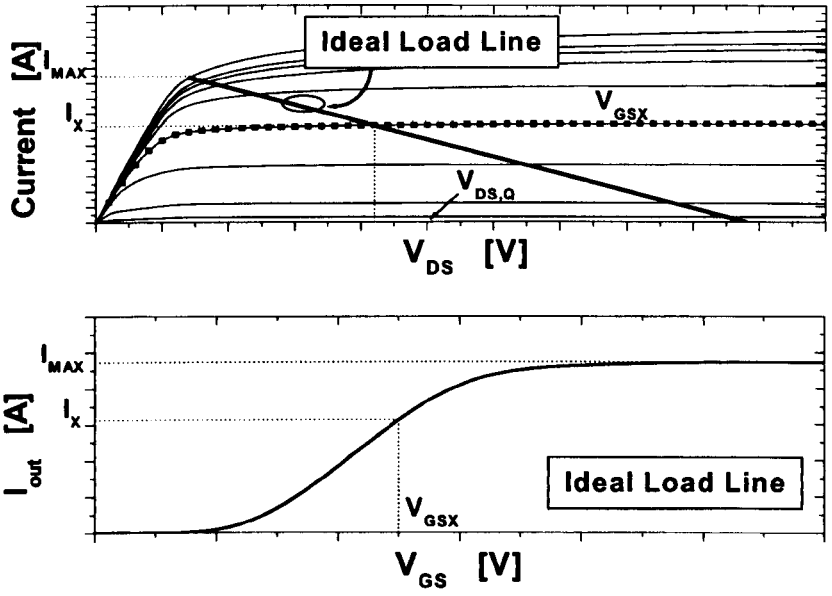


Figure 4.2: Superimposition of the load line to the I_{DS}-V_{DS} curves and extracted load line output trans-characteristic. The extraction is carried out considering a gate voltage at a time. For a certain gate voltage V_{GSX} and I_{DS}-V_{DS} curve is identified. The intersection of this curve with the load line permits an association of the current level I_X of the load line at the intersection point with the gate voltage V_{GSX} of the I_{DS}-V_{DS} curve. The locus of the (V_{GSX},I_X) points so determined is the load line output transfer-characteristic.

4.2.3 Results

Simulated I_{DS}-V_{DS} are used in this section for the determination of V_{knee} and of the load line output characteristic. The simulations are carried out in ADS [9] on the

publicly available model of the Polyfet SP204 [10]. This is a non linear model based on a lumped element equivalent circuit developed in standard Berkeley Spice format. The SP204 model is optimised for V_{DS}=28V and V_{GS}=3.16V, presenting a good match with the published s-parameters for this bias, as it will be shown in section 4.3, Figure 4.9. Thermal effects are not included.

Although not suitable to optimally represent the device behaviour for all possible biases, this model topology has been preferred to the state of the art electro-thermal model adopted by Motorola [11] due to the possibility of easily de-embedding the package effect and monitoring internal currents and voltages. The Motorola Electro-Thermal model (MET) is an empirical model which includes thermal effects, package and internal matching networks contributions. De-embedding the package and matching networks from the MET model is not possible due to the unavailability of detailed information on the device. Furthermore, the empirical nature of the MET model makes it impossible to monitor internal voltages and currents.

Extraction of the knee voltage

Equation 4.8 is used for the determination of the maximum linearly delivered power as a function of V_{knee}, where the drain to source bias is imposed at $V_{DS,Q}$=28V. The calculated values for a voltage sweep from 0V to V_{DS}=$V_{DS,Q}$ are shown in Figure 4.3. A clear maximum is identified in V_{knee}=7.17V.

Figure 4.3 helps to understand how to graphically identify the knee voltage when a non linear device is used. The knee voltage appreciatively corresponds to the intersection of the lines tangent to the current characteristic at V_{DS}=0V and V_{DS}=$V_{DS,Q}$.

Determination of the load line characteristic

The load line characteristic is compared in Figure 4.4 to the conventional device output characteristic at a constant V_{DS}=$V_{DS,Q}$; the latter is the current characteristic usually reported in device datasheets. The constant V_{DS} output characteristic yields a maximum current value which is considerably higher than the load line case. The smaller saturation value for the load line characteristic is

due to the knee of the I_{DS}-V_{DS} curves, which limits the current to the maximum value I_{MAX}. The two output characteristic yield different transconductance curves, with a narrower g_m bell observed for the load line case.

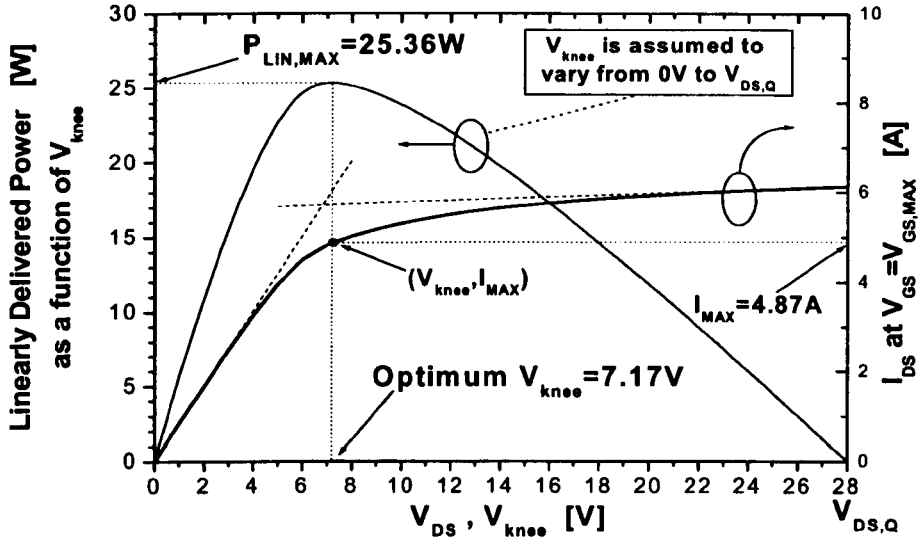


Figure 4.3: Determination of the knee voltage for a non ideal transistor. V_{knee} is determined in correspondence of the maximum of the calculated linearly delivered power.

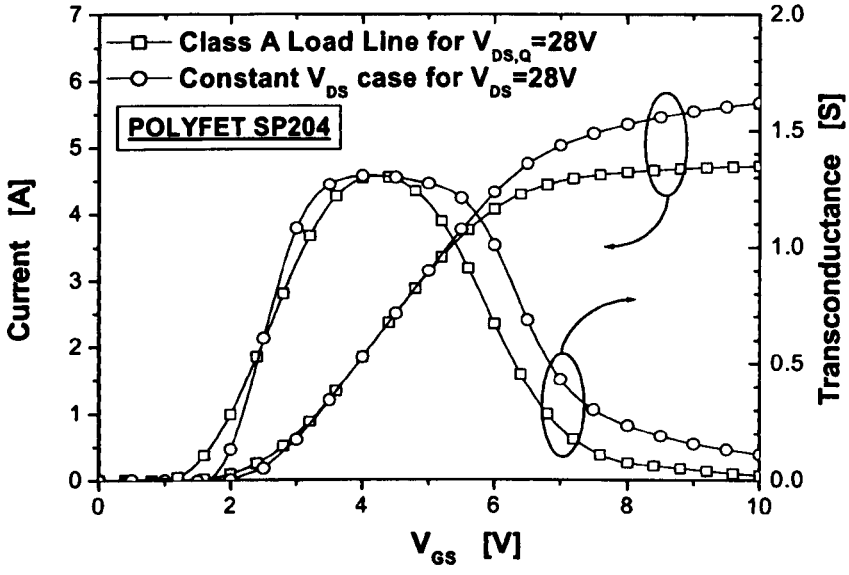


Figure 4.4: Drain current versus gate to source voltage and associated transconductance characteristics in the load line and in the constant drain to source voltage case.

Verification of the load line characteristic

The verification of the accuracy of the load line output trans-characteristic in representing the device load cycle is discussed in section 4.3, Figure 4.19. It is shown that the load line approximation permits a reasonably accurate prediction of the amplifier load cycle up to the 1dB gain compression point.

4.3 Power Gain Assessment

Generic analytic expressions and methodology for predicting the large signal gain of silicon RF power MOSFETs are presented. The expressions are derived from a model which includes input and output matching impedances, source inductance and gate resistance. Using the load line superimposed on a non linear current generator, this section demonstrates reasonably accurate predictions of gain and gain compression point.

4.3.1 Background

The effect of load impedance on load-pull contours was first analytically described by Cripps, who in 1983, demonstrated simplified equations that lead to good agreement with experiment [4]. Following Cripps, a power gain expression was derived in [5] based on the transistor model of Figure 4.5 under the assumption of a linear current generator, zero gate resistance and drain to source conductance.

The power gain and optimum source and load impedance were given as

$$G = \frac{g_m R_{opt}}{\omega^2 L_S [C_{GS} + C_{GD}(1 + R_{opt} g_m)]}, \tag{4.9}$$

$$Z_S = \frac{L_S g_m}{C_{GS} + C_{GD}(1 + R_{opt} g_m)} + j\frac{1 + \omega^2 L_S (C_{DS} R_{opt} g_m - C_{GS})}{\omega [C_{GS} + C_{GD}(1 + R_{opt} g_m)]}, \tag{4.10}$$

$$Z_L = \frac{(\omega^2 C_{DS} L_S R_{opt} - R_{opt} + j\omega L_S) g_m - \omega^2 C_{GS} L_S}{[j\omega(C_{DS} + C_{GD}) R_{opt} - 1] g_m + j\omega C_{GD}}, \tag{4.11}$$

where ω is the angular frequency, L_S is the source inductance, C_{GS} is the gate to source capacitance, C_{GD} is the gate to drain capacitance, C_{DS} is the drain to source capacitance, g_m is the transconductance value, R_{opt} is the load line optimum resistance, Z_L is the optimum load impedance and Z_S is the optimum source impedance. The value of the model parameters are extracted at the application frequency and bias. The transconductance value considered in Equations 4.9 to 4.11 varies with bias, reducing from the maximum g_m in class A, to $g_m/2$ in class B.

Another known expression that permits the calculation of the power gain is

$$G = \frac{g_m^2 R_L}{\omega^2 [C_{GS} + C_{GD}(1 + \frac{g_{m,MOS}}{g_{m,JFET}})]^2 R_G (1 + \omega^2 C_{DS}^2 R_L^2)}, \tag{4.12}$$

where R_L is the amplifier load resistance, $g_{m,MOS}$ and $g_{m,JFET}$ are the transconductance values associated with the intrinsic MOSFET and the JFET resistance of Si RF power MOSFETs [7].

Equations 4.9 to 4.11 may be considered as an improvement with respect to Equation 4.12, because Equation 4.12 does not include matching input and output impedances or the source inductance, which is particularly relevant for the VDMOS (Vertical Diffused MOS).

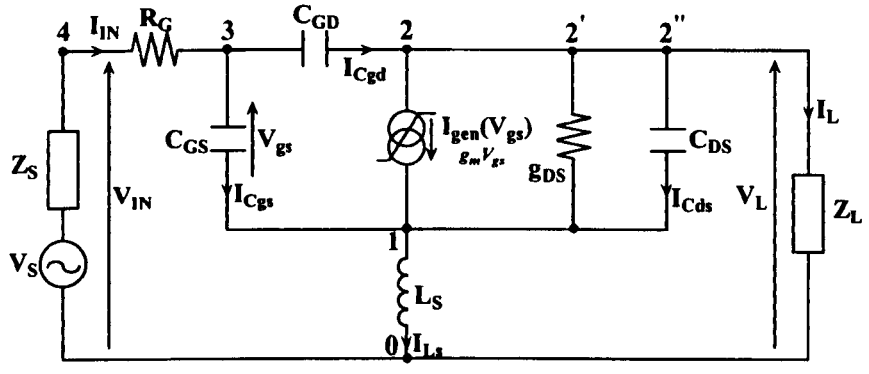


Figure 4.5: Transistor model including matching impedances, gate resistance, source inductance and a non linear voltage controlled current generator. In reference [5], R_G and g_{DS} were not considered and $I_{gen}(V_{gs})$ was assumed linear.

4.3.2 Theoretical analysis

Effect of Gate resistance on power gain

The expression in Equation 4.9 overestimates the power gain as can be concluded from the results in [5]. Hence to overcome this limitation, in this work the model was modified to include the impact of R_G. The drift conductance g_{DS} is also included to consider the power dissipated due to the drift resistance. Proceeding as in [5], the optimum load impedance Z_L is determined by forcing the current generator to see a real output impedance of value R_{opt}. R_{opt} is the load line optimum resistance associated with the maximum voltage and current swings. The optimum source impedance Z_S is determined as the conjugate match of the transistor's input impedance.

Referring to Figure 4.5, circuit analysis reveals the existence of five independent nodes, referred by numbers from 0 to 4. Node 2 is split as 2, 2' and 2'' to permit an easier identification of the currents. The imposition of R_{opt} on the current generator yields

$$V_{21} = -g_m R_{opt} V_{gs}, \tag{4.13}$$

Assuming that $V_{31}=V_{gs}$ is known,

$$V_{32} = V_{31} - V_{21} = V_{gs} + g_m R_{opt} V_{gs} = (1 + g_m R_{opt}) V_{gs}, \tag{4.14}$$

$$V_{RG} = V_{43} = R_G I_{IN}, \tag{4.15}$$

$$I_{CGD} = I_{32} = Y_{32} V_{32} = j\omega C_{GD}(1 + g_m R_{opt}) V_{gs}, \tag{4.16}$$

$$I_{CGS} = I_{31} = Y_{31} V_{31} = j\omega C_{GS} V_{gs}, \tag{4.17}$$

$$I_{CDS} = Y_{21} V_{21} = -g_{DS} g_m R_{opt} V_{gs}, \tag{4.18}$$

$$I_{gDS} = I_{2''1} = -j\omega C_{DS} g_m R_{opt} V_{gs}, \tag{4.19}$$

$$\begin{aligned} I_{L_S} &= I_{10} = I_{31} + I_{21} + I_{2'1} + I_{2''1} = \\ &= [j\omega C_{GS} + g_m(1 - g_{DS} R_{opt}) - j\omega C_{DS} g_m R_{opt}] V_{gs} \end{aligned}, \tag{4.20}$$

$$V_{LS} = V_{10} = Z_{10} I_{10} = j\omega L_S [j\omega C_{GS} + g_m(1 - g_{DS} R_{opt}) - j\omega C_{DS} g_m R_{opt}] V_{gs}. \tag{4.21}$$

Input and load voltages and currents can now be calculated as

$$V_{IN} = V_{40} = V_{43} + V_{31} + V_{10}, \tag{4.22}$$

$$I_{IN} = I_{32} + I_{31}, \tag{4.23}$$

$$V_L = V_{20} = V_{21} + V_{10}, \tag{4.24}$$

$$I_L = I_{32} - I_{21} - I_{2'1} - I_{2''1}, \tag{4.25}$$

resulting in

$$V_{IN} = R_G I_{IN} + V_{gs} + [\omega^2 C_{DS} L_S R_{opt} + j\omega L_S (1 - g_{DS} R_{opt})] g_m V_{gs} - \omega^2 C_{GS} L_S V_{gs}, \tag{4.26}$$

$$I_{IN} = j\omega (C_{GS} + C_{GD}) V_{gs} + j\omega C_{GD} R_{opt} g_m V_{gs}, \tag{4.27}$$

$$V_L = [\omega^2 C_{DS} L_S R_{opt} - R_{opt} + j\omega L_S (1 - g_{DS} R_{opt})] g_m V_{gs} - \omega^2 C_{GS} L_S V_{gs}, \tag{4.28}$$

$$I_L = [j\omega (C_{DS} + C_{GD}) R_{opt} - (1 - g_{DS} R_{opt})] g_m V_{gs} + j\omega C_{GD} V_{gs}. \tag{4.29}$$

Using Equations 4.26 to 4.29 the optimum source and load resistance are now given as

$$Z_S = R_G + \frac{L_S g_m (1 - g_{DS} R_{opt})}{C_{GS} + C_{GD} (1 + R_{opt} g_m)} + j \frac{1 + \omega^2 L_S (C_{DS} R_{opt} g_m - C_{GS})}{\omega [C_{GS} + C_{GD} (1 + R_{opt} g_m)]}, \tag{4.30}$$

$$Z_L = \frac{[\omega^2 C_{DS} L_S R_{opt} - R_{opt} + j\omega L_S (1 - g_{DS} R_{opt})] g_m - \omega^2 C_{GS} L_S}{[j\omega (C_{DS} + C_{GD}) R_{opt} - (1 - g_{DS} R_{opt})] g_m + j\omega C_{GD}}. \tag{4.31}$$

The input/output power and power gain expressions are derived by defining the input power as the power delivered to the transistor under conjugate match conditions and the output power as the power dissipated by the load-line resistance. These definitions yield

$$P_{OUT} = 1/2 R_{opt} (g_m V_{gs})^2, \tag{4.32}$$

$$P_{IN} = 1/2 \left[\begin{array}{c} (C_{GS} + C_{GD} (1 + R_{opt} g_m))^2 R_G + \\ + L_S g_m (1 - g_{DS} R_{opt}) (C_{GS} + C_{GD} (1 + g_m R_{opt}) \end{array} \right] \omega^2 V_{gs}^2, \tag{4.33}$$

$$G = \frac{R_{opt} g_m^2}{\left[\begin{array}{c} (C_{GS} + C_{GD} (1 + R_{opt} g_m))^2 R_G + \\ + L_S g_m (1 - g_{DS} R_{opt}) (C_{GS} + C_{GD} + C_{GD} R_{opt} g_m) \end{array} \right] \omega^2}. \tag{4.34}$$

It can be easily verified that expressions 4.30, 4.31 and 4.34 coincide with those in [5] for R_G=0Ω and g_{DS}=0S.

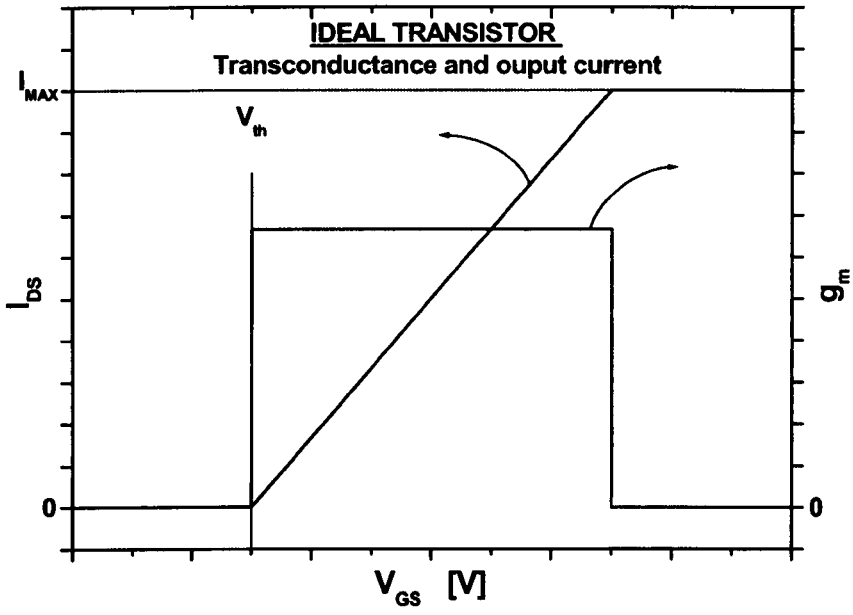


Figure 4.6: Transconductance and output characteristic of an ideal transistor.

Power Gain expressions in the nonlinear case

Equation 4.34 can be used only for the determination of power gain at small input signal levels, i.e. where the output is linearly proportional to the input. This occurs in the ideal transistor case. The derivation implicitly assumes a constant value of the transconductance and a linear drain current to input voltage relationship as shown in Figure 2.3. In this case only, the fundamental component of the device current I_{gen} remains proportional to the input signal V_{gs} through the transconductance value g_m:

$$I_{gen} = g_m V_{gs}. \tag{4.35}$$

However, real devices have transconductance values which are neither constant nor linearly dependent on the input voltage. In the nonlinear current generator case, Equation 4.35 cannot be considered valid. The fundamental component of the current (I_{gen}) has to be determined from the actual current waveform $I_{gen,wave}$

$$I_{gen,wave} = I_{out}\left(V_{GS} + V_{gs}(t)\right) \tag{4.36}$$

where I_{out} is the device output characteristic, V_{GS} is the gate bias voltage and $V_{gs}(t)$ is the input RF gate signal.

To obtain expressions for the input and load currents and voltages, $g_m V_{gs}$ has to be replaced with I_{gen} in Equations 4.26 to 4.29. The optimum source and load impedances, input and output power and power gain can then be calculated once the fundamental component of the current (I_{gen}) is known. However it is not possible to provide an explicit formula for these parameters. They are defined as:

$$Z_S = conj(V_{IN}/I_{IN}) \tag{4.37}$$

$$Z_L = V_L/I_L \tag{4.38}$$

$$P_{IN} = 1/2\,\mathrm{Re}[Z_S]\cdot I_{IN} conj(I_{IN}) \tag{4.39}$$

$$P_{OUT} = 1/2 R_{opt}\cdot I_{gen} conj(I_{gen}) \tag{4.40}$$

$$G = P_{OUT}/P_{IN} \tag{4.41}$$

The values of optimum load and source impedance are evaluated at the 1dB compression point. The input signal amplitude V_{gs}=V_{compr} at the 1dB compression point is determined from the 3rd order power series expansion of the output current characteristic, after [4]. In the one tone and two tone case it can be stated as

$$V_{compr,1\text{-}tone} = \sqrt{\frac{4c_1(1-10^{-0.05})}{3c_3}}, \tag{4.42}$$

$$V_{compr,2\text{-}tone} = \sqrt{\frac{4c_1(1-10^{-0.05})}{9c_3}}, \tag{4.43}$$

where c_1 and c_3 are the first and third order coefficients of the power series expansion of the current. The extraction of the optimum impedance will be clarified with a practical example in section 4.3.4.

4.3.3 Methodology for gain compression prediction

The gain compression cannot be estimated from the gain expressions in Equation 4.9, 4.12 and 4.34 since device non linearity is not taken into account. In this section an approach to predict gain compression based on the identification of the fundamental current component I_{gen} is described. I_{gen} is determined by calculating:

(i) the output characteristic I_{out} from the device I_{DS}-V_{DS} characteristics,

(ii) the current signal waveform $I_{gen,wave}$ from I_{out},

(iii) the fundamental component I_{gen} of the current signal from $I_{gen,wave}$,

Power gain is then determined by replacing $g_m V_{gs}$ with I_{gen} in Equations 4.26 to 4.29 and applying Equations 4.39 to 4.41.

Output current characteristic and signal waveform determination

Conventionally the I_{DS}-V_{GS} relationship used as the output current characteristic corresponds to a current measurement carried out at a constant V_{DS}. This output characteristic differs from the output characteristic extracted along the load line, as has been highlighted in section 4.3.4.

Since in a real amplifier the device operates along the load line characteristic, it follows that the correct estimation of power gain and gain compression can be achieved only if the output current characteristic I_{out} in Equation 4.36 is extracted along the load line. Using the conventional constant V_{DS} output characteristic instead of $I_{loadline}$ in the power calculation leads to a miscalculation of the fundamental component of the device generator current, as will be demonstrated later in section 4.3.4.

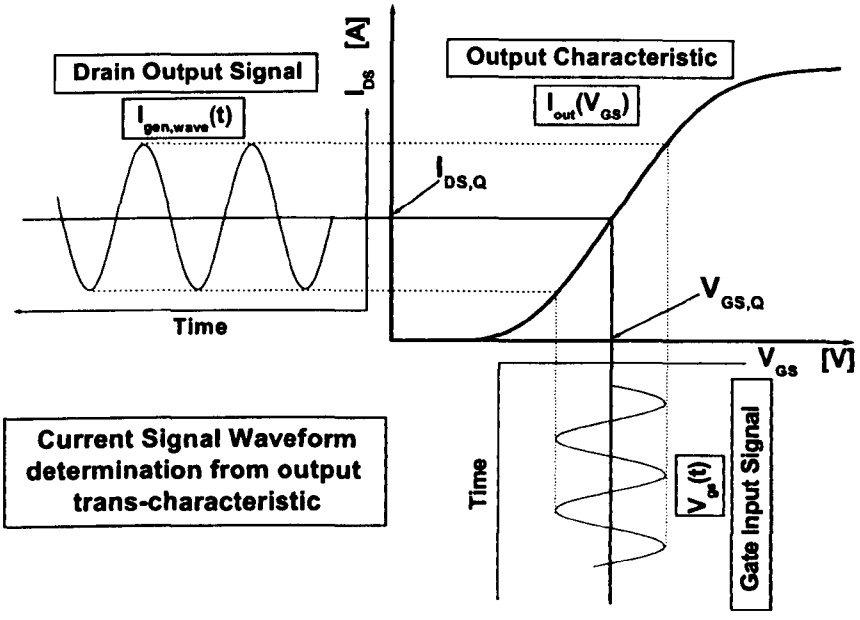


Figure 4.7: Amplification principle for a MOSFET in common source configuration: output drain current signal determination from the input gate voltage signal through the output trans-characteristic.

The load line characteristic $I_{loadline}$ is extracted from the device I_{DS}-V_{DS} curves as described in section 4.2. Once the output characteristic I_{out}=$I_{loadline}$ is determined, the output drain current waveform $I_{gen,wave}$ is determined from the device output transfer characteristic applying Equation 4.36, as demonstrated graphically in Figure 4.7. The current waveform is next used to determine the fundamental component of the generator current.

Fundamental generator current component

In this section two approaches based on the Fourier analysis of the current signal waveform are presented: the (FFA) Fourier Fundamental Approach, only suitable for single tone input signals, and the (FSA) Fourier Spectrum Approach, also suitable for multi-tone input-signals. For the sake of completeness, the method proposed in [7] that makes use of an rms value of transconductance is also described. This will be referred to as the Empirical Approach (EA).

EA – The Empirical Approach

In [7] the extension of gain to high input levels was carried out by substituting g_m with its rms value:

$$g_{m,rms} = \sqrt{1/T \int_0^T \left(\frac{\partial I_{gen,wave}}{\partial V_{GS}} \right)^2 dt} = \sqrt{1/T \int_0^T \left(\frac{\partial I_{out}(V_{GS,Q} + V_{gs}(t)}{\partial V_{GS}} \right)^2 dt} \qquad (4.44)$$

based on the empirical assumption that the generator current can be described as

$$I_{gen} = g_{m,rms} V_{gs} \qquad (4.45)$$

This expression is empirical in that it is not based on any justifiable physical or mathematical analysis; it can be thought of as derived from a linearization of the more general Equation 4.36. As such, its validity surely holds at small input signal levels, but not necessarily at large signal levels.

FFA - The Fourier Fundamental Approach

In the case of a periodic input waveform, it is possible to express the generator current waveform $I_{gen,wave}$ by its Fourier series expansion

$$I_{gen,wave} = I_{out}(V_{GS,Q} + V_{gs}(t) = I_0 + \sum_{k=1}^{\infty} I_K \sin(k\omega t + \phi_k) \quad (4.46)$$

with

$$I_0 = 1/T \int_0^T I_{gen,wave}(t)dt \quad (4.47)$$

$$I_k = \sqrt{A_k^2 + B_k^2} \quad (4.48)$$

$$\phi_k = a\tan(B_k / A_k) \quad (4.49)$$

$$A_k = 2/T \int_0^T I_{gen,wave}(t)\sin(k\omega t)dt \quad (4.50)$$

$$B_k = 2/T \int_0^T I_{gen,wave}(t)\cos(k\omega t)dt \quad (4.51)$$

where I_k represents the amplitude of the signal of the k-th multiple of the fundamental frequency, T is the period and ω is the angular frequency of the periodic input signal. For a single tone input of angular frequency ω, the determination of the current component at the fundamental frequency $I_{gen}=I_1$ is carried out using Equations 4.46 to 4.51.

FSA - The Fourier Spectrum approach

For a single-tone signal it is relatively easy to calculate I_{gen} from Equation 4.36 using the approach of Fourier series as given by Equations 4.46 to 4.51. However, in the case of two closely spaced signals the above approach is limited by excessive computational times. Hence a Fourier spectrum analysis is required. The decomposition of the current signal using the Fourier transform permits the identification of the signal spectrum and the identification of the frequency components of the current signal as

$$I(\omega) = \int_{-\infty}^{\infty} I_{gen,wave}(t)e^{-j\omega t}dt = \int_{-\infty}^{\infty} I_{out}(V_{GS,Q} + V_{gs}(t))e^{-j\omega t}dt\,. \quad (4.52)$$

The Fourier transform is carried out by applying the computationally efficient Fast Fourier Transform algorithm (FFT). The generator current component I_{gen} at the fundamental frequency can then be determined.

4.3.4 Results

In order to provide a benchmark for the assessment of the accuracy of the proposed expressions harmonic balance simulations have been used in this work. The simulations are carried out on the 28V Polyfet SP204 [10] in ADS [9] using the publicly available Polyfet model [10]. The model, which has been described in section 4.2.3, is shown in Figure 4.8 and Figure 4.9 to permit adequate reconstruction of measured IV-characteristics and s-parameters.

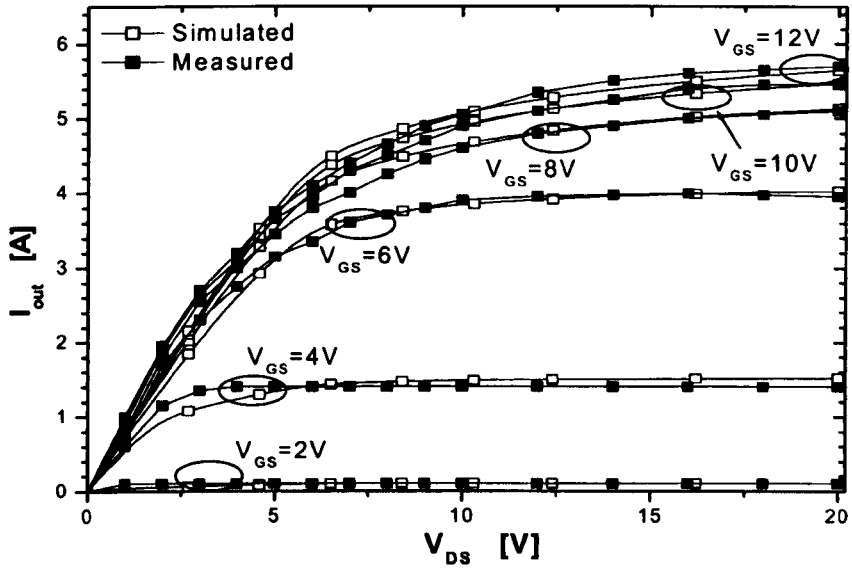


Figure 4.8: ADS simulated versus measured I_{DS}-V_{DS} characteristics for the Polyfet SP204.

In this work, model components related to the package have been removed to achieve direct correspondence between the calculated and simulated impedance values. Alternatively, impedance transformation can be carried out to include the effect of package [5]. Ideal components such as RF chokes, DC feeds and shorts instead of transmission lines have been used in the simulations. The basic schematic setup used in power gain simulations in the single and two tone input case is shown in Figure 4.10. The extraction of the device parameter values is next described, followed by the procedure for the determination of the optimum source and load impedance values in ADS. This is followed by the amplifier setup for the extraction of power gain in the single and two tone cases.

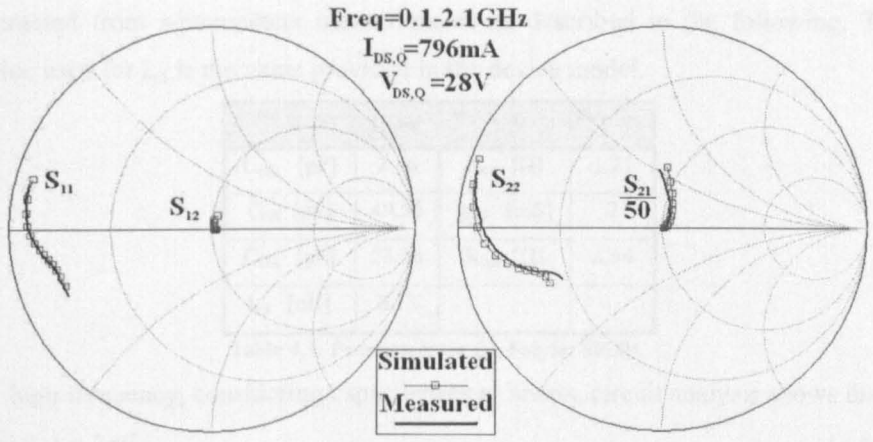


Figure 4.9: ADS simulated versus measured s-parameters for the Polyfet SP204.

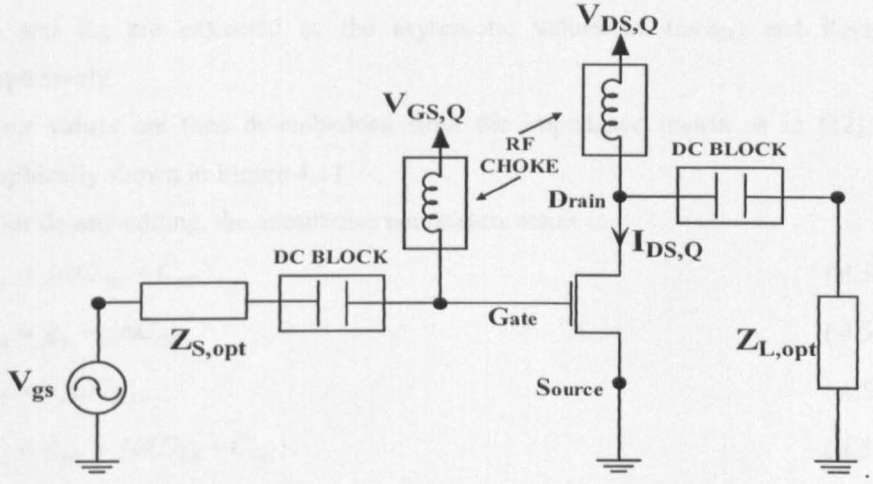


Figure 4.10: One-tone and two-tone ADS simulation setup.

Extraction of the Device Parameters Values

The data used for the prediction of power gain based on the proposed equations and approaches is represented from the set of I_{DS}-V_{DS} characteristics and the value of the model elements. The I_{DS}-V_{DS} data extracted from ADS DC simulations is used for the determination of the fundamental current component I_{gen}. The model elements values are extracted at the application voltage V_{DS}=28V from the results of a small signal simulation. Table 4.1 indicates the values used in the calculations throughout section 4.3.4. With the exception of L_S, these values have been

extracted from s-parameters measurements as described in the following. The value used for L_S is the value provided in the device model.

Parameter	Value	Parameter	Value
C_{GD} [pF]	2.56	R_G [Ω]	1.27
C_{GS} [pF]	49.56	g_{DS} [mS]	2
C_{DS} [pF]	23.36	R_{opt} [Ω]	8.54
L_S [nH]	0.13		

Table 4.1: Parameters for the Polyfet SP204.

At high frequency, considering capacitances as shorts, circuit analysis shows that

$$\mathrm{Im}(z_{21}) = 2\pi f L_S, \tag{4.53}$$

$$\mathrm{Re}(z_{11}) = R_G. \tag{4.54}$$

L_S and R_G are extracted as the asymptotic values of $\mathrm{Im}(z_{21})$ and $\mathrm{Re}(z_{11})$ respectively.

Their values are then de-embedded from the impedance matrix as in [12] as graphically shown in Figure 4.11.

After de-embedding, the admittance parameters result as

$$y_{11} = j\omega(C_{GS} + C_{GD}), \tag{4.55}$$

$$y_{21} = g_m - j\omega C_{GD}, \tag{4.56}$$

$$y_{12} = -j\omega C_{GD}, \tag{4.57}$$

$$y_{22} = g_{DS} + j\omega(C_{DS} + C_{GD}). \tag{4.58}$$

The model parameter values can now be extracted as

$$C_{GD} = -\frac{\mathrm{Im}(y_{12})}{\omega}, \tag{4.59}$$

$$C_{GS} = \frac{\mathrm{Im}(y_{11}) + \mathrm{Im}(y_{21})}{\omega}, \tag{4.60}$$

$$C_{DS} = \frac{\mathrm{Im}(y_{22}) + \mathrm{Im}(y_{12})}{\omega}. \tag{4.61}$$

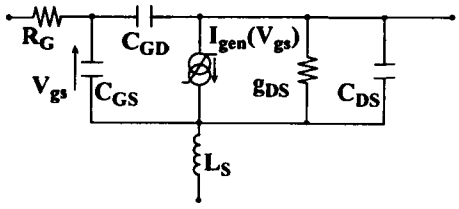
$$g_{DS} = \mathrm{Re}(y_{22}) \tag{4.62}$$

The class A R_{opt} value is calculated for a fixed supply voltage $V_{DS,Q}$=28V and a V_{knee}=7.17V as

$$R_{opt} = 2\frac{V_{DS,Q} - V_{knee}}{I_{MAX}}. \quad (4.63)$$

(a) $S' \rightarrow Z'$

(b) $Z' = \begin{Bmatrix} z'_{11} & z'_{12} \\ z'_{21} & z'_{22} \end{Bmatrix}$



(c) $Z = \begin{Bmatrix} z'_{11}-(R_G+j\omega L_S) & z'_{12}-j\omega L_S \\ z'_{21}-j\omega L_S & z'_{22}-j\omega L_S \end{Bmatrix}$

(d) $Z \rightarrow Y$

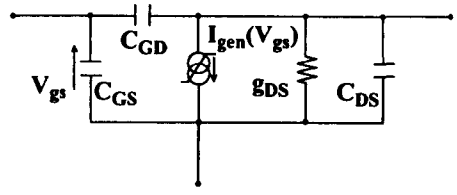


Figure 4.11: De-embedding source inductance and gate resistance from the small signal parameters circuit. The measured s parameters matrix S' is converted in the impedance matrix Z' (a,b). Z' refers to the equivalent circuit which includes R_G and L_S. R_G and L_S values are determined from the impedance matrix Z' using Equations 4.53 and 4.54. Their values are de-embedded from the impedance matrix in (c), after [12]. The impedance matrix Z corresponds to the equivalent circuit in which R_G and L_S have been removed. The admittance matrix Y (without the R_G and L_S contribution) is obtained from the Z parameter matrix in (d). The admittance matrix Y is then used for model parameter extraction by using Equations 4.59 to 4.61.

Optimum Source and Load impedance determination

The optimum source ($Z_{S,opt}$) and load ($Z_{L,opt}$) impedance values for class A bias are extracted from load and source pull simulations using ADS as the values leading to the maximum 1dB gain compression point. The schematics of the load and source pull simulation setups are shown in Figure 4.12 and Figure 4.14.

The optimum load impedance is determined first. As in real load pull measurements, $Z_{L,opt}$ is identified as the load impedance yielding the highest possible level of power delivered to the load for a constant input power level. As shown in Figure 4.12, an impedance tuner is used as the amplifier load. As the tuner impedance is varied, the corresponding power delivered to the load changes. Keeping the input power constant ensures that the variation of delivered power is associated only with the variation of the impedance. Measuring the delivered power for many tuner impedance values permits the identification of the loci of constant delivered power as a function of the load impedance. The optimum load

impedance is determined as the value corresponding to the maximum delivered power level. The load-pull contours are shown in Figure 4.13.

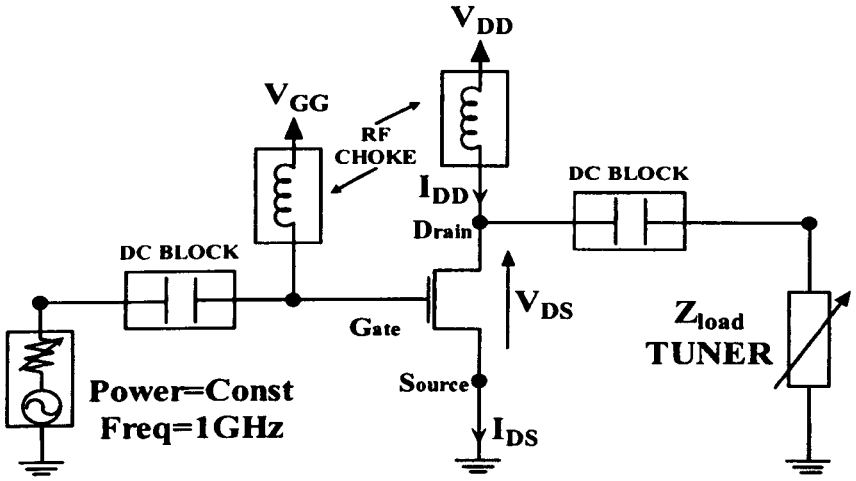


Figure 4.12: Load-pull simulation setup.

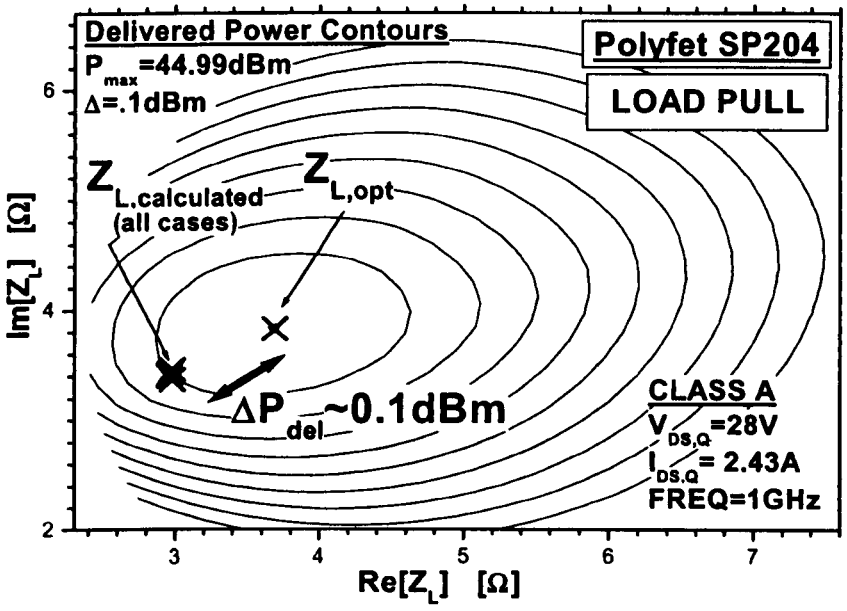


Figure 4.13: Load pull delivered power contours in the class A single tone simulation of the Polyfet SP204.

The optimum source impedance is determined analogously. As shown in Figure 4.14, for this case the impedance tuner is placed on the input side and the optimum load impedance at the output side of the amplifier. Keeping the input

power constant ensures that the variation of power delivered to the load is associated only with the variation of the tuner impedance. The loci of constant delivered power as a function of the source impedance are described in Figure 4.15. The optimum source impedance is determined as the value associated with the maximum delivered power level.

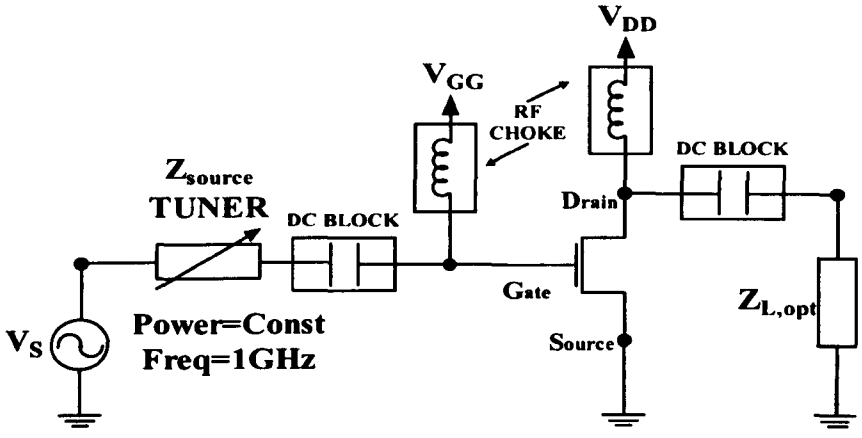


Figure 4.14: Source-pull simulation setup.

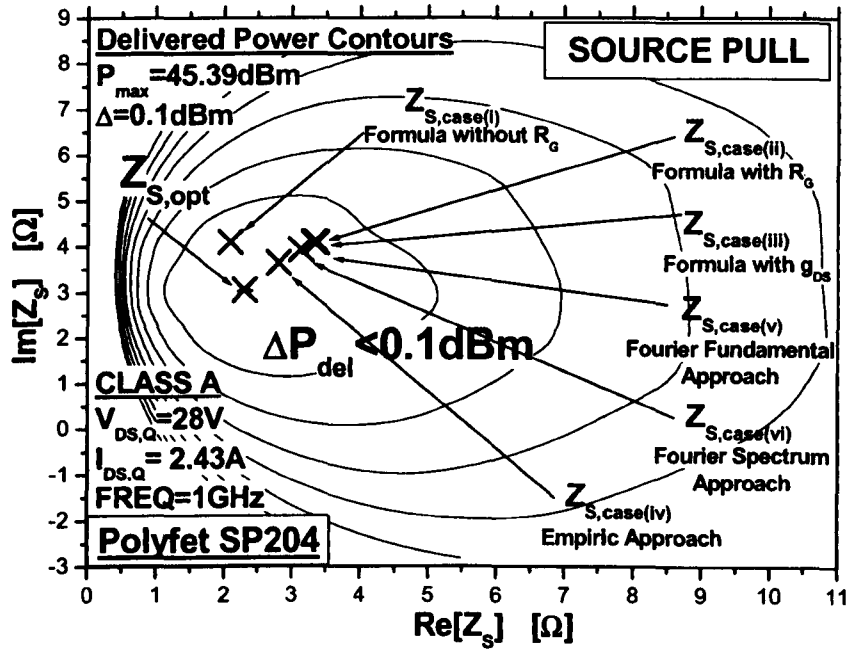


Figure 4.15: Source pull delivered power contours in the class A single tone simulation of the Polyfet SP204.

A dependence on the input power level is observed for load and source impedances determined from load and source pull simulations, as shown in Figure 4.16. The optimum load and source impedance values are $Z_{L,opt}$= 3.69+3.83j and $Z_{S,opt}$= 2.3+3.05j.

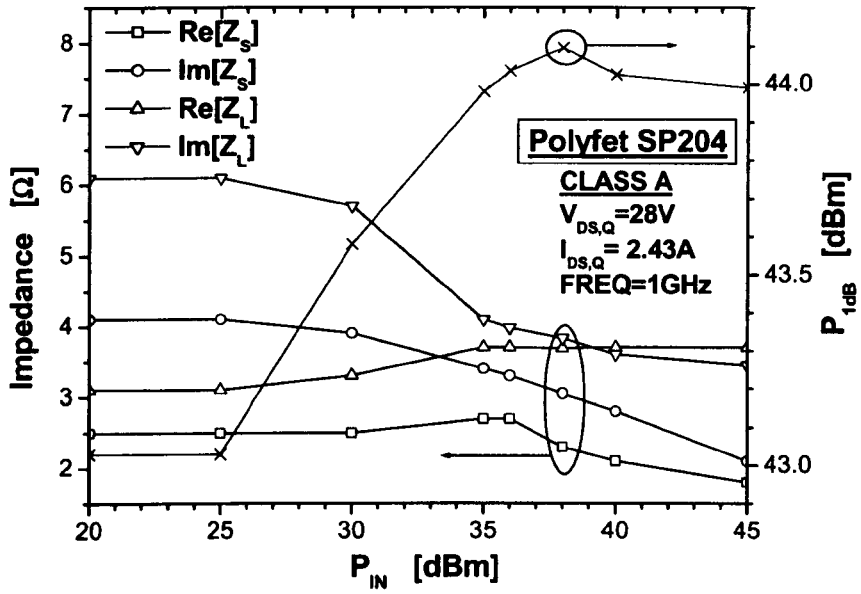


Figure 4.16: Dependence of the optimum load and source pull determined impedance on the input power level. The maximum P_{1dB} is found in correspondence of the optimum load and source impedance.

For a given P_{IN} power level, the data in Figure 4.16 has been extracted as follows:

1. Z_L has been determined from load pull simulations as the output impedance value that maximises the delivered power
2. Z_S has been determined from source pull simulations as the input impedance value that maximises the delivered power
3. P_{1dB} has been extracted from a single tone simulation of the Polyfet SP204.

The dependence of the optimum impedances on the input power level imposed on the load and source pull simulations is a consequence of the device non linearity. Low input power corresponds to small input signal amplitudes, therefore load and source impedance determined as such levels will be close to the small signal value. In Figure 4.16 the impedance values remain constant for power input

values up to 25dBm, corresponding to a 0.8V single tone signal amplitude. The constant impedance values reveal a linear response of the device for small input signal amplitudes. At higher input power the input signal amplitude increases, exposing the non linear device behaviour (waveform clipping and harmonic distortion generation). Therefore, the extracted impedance values implicitly consider the device non linearity, yielding larger 1dB gain compression point than in the small signal case.

The effect of optimum and small signal matching impedance effect on the device load cycles is shown in Figure 4.17. The impedances extracted at a small input signal cause the device to operate on a load line that is not the optimum class A. Also, a considerable deviation from the optimum load line is observed. This is a consequence of the non optimum values used, which yield the imposition of complex impedance instead of a pure resistor on the device current generator.

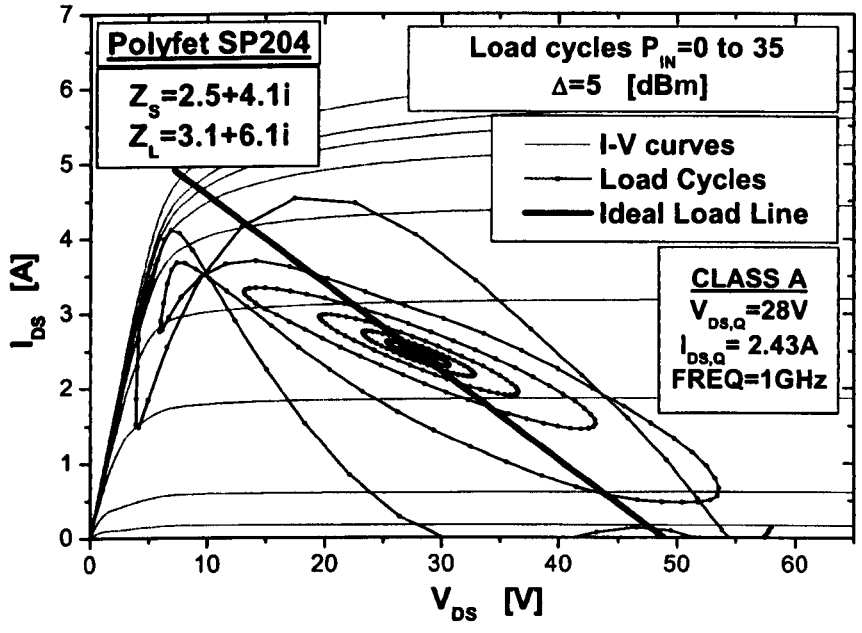


Figure 4.17: Ideal load line and load cycles for the load and source impedances determined at small input power. The non optimum impedances impose a non optimum load line to the device, resulting in lower maximum linear power.

Optimum Source and Load impedance in the single tone input signal case

Six different ways of calculating the optimum impedance values are now assessed. The first three assume a linear current generator in the equivalent model:

(i) without the inclusion of gate resistance [5];

(ii) including R_G (g_{DS}=0) as described in section 4.3.2.

(iii) including R_G and g_{DS}, as described in section 4.3.2.

Three methods consider the non linear current generator as described in section 4.3.3:

(iv) by using the Empiric Approach (EA);

(v) by using the Fourier Fundamental Approach (FFA);

(vi) by using the Fourier Spectrum Approach (FSA).

In (iv)-(vi) the calculations are carried out considering the output characteristic extracted along the load line rather than the fixed V_{DS} approach used to date.

An example of determination of the optimum load and source impedance values is demonstrated for the non linear current generator cases (iv)-(vi) with the device biased in class A at V_{DS}=28V. The Fourier Fundamental Approach is considered for this task. The calculated impedance values obtained from Equations 4.37 and 4.38 are shown in Figure 4.18. The source impedance displays a strong dependence on the input signal level, while the optimum load impedance is appreciatively constant. It can be shown that this behaviour follows from the existence of the feedback elements C_{GD} and L_S in the circuit. The optimum values for the source and load impedance are extracted in correspondence to the 1dB compression point. The V_{gs}=V_{compr} value at which gain compression occurs is calculated by using Equations 4.42 (single tone case), yielding a value of 2.43V.

The calculated optimum impedance values are compared with those from ADS simulations. The power delivered to the load, shown in Figure 4.13 and Figure 4.15, obtained by harmonic balance simulation in ADS is used to assess the accuracy of the calculated impedance values. The validation is based on the assumption that optimum impedance predictions correspond to a level of delivered power that is close to the maximum value.

The impedances calculated using all four approaches have been highlighted in Figure 4.13 and Figure 4.15, in the impedance plane of the delivered power contours. All the methods can be considered acceptable for the prediction the load and source impedance values. The errors in delivered power level are in fact always smaller than 0.1dBm, corresponding to a maximum error of 0.2% on the dBm value and 2.28% on the value in watts. The cause of the errors in source and load impedance prediction can be identified in the model simplification and to the load line approximation limitations, where the device output characteristic has been extracted along the ideal load line instead that along the actual load cycles, as shown in Figure 4.19.

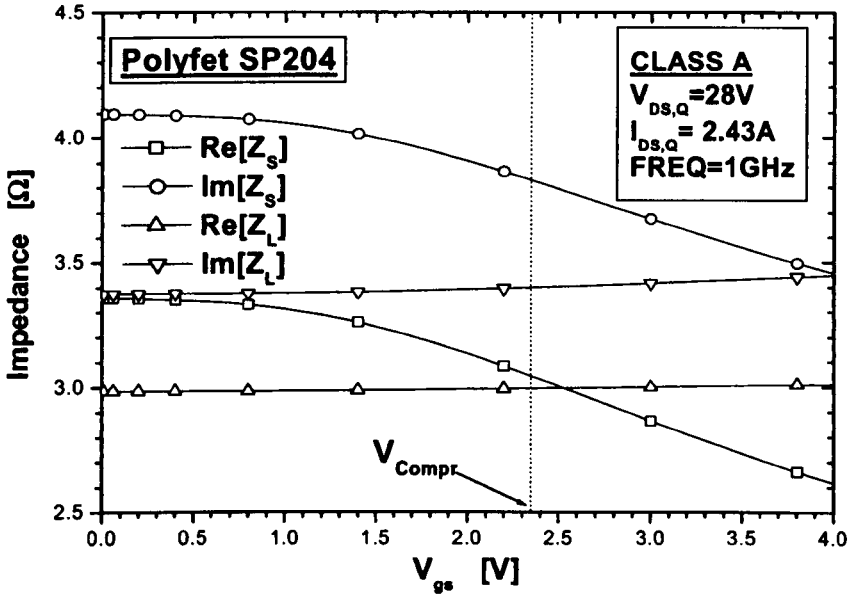


Figure 4.18: Calculated optimum source and load impedances in the single tone input signal case. The data refer to the application of the Fourier Spectrum Approach (FSA). V_{compr} is determined with Equation 4.42, after the determination of the power series coefficient c_1=1.2866[V^{-1}] and c_2=-0.0379[V^{-3}] from the fit of the load line output transc-characteristic.

In Table 4.2 it is observed that the calculated optimum load impedance remains practically constant in all cases. On the other hand, the introduction of the gate resistance and the drain conductance in the equivalent device model leads to differences in the calculated optimum source impedance values. The gate resistance causes an increase of the real part of the calculated source impedance. It

produces a substantial improvement in the prediction of the power gain value at small input signal levels. The drain conductance does not considerably affect power gain or impedance calculation, but adds to the generality of the model.

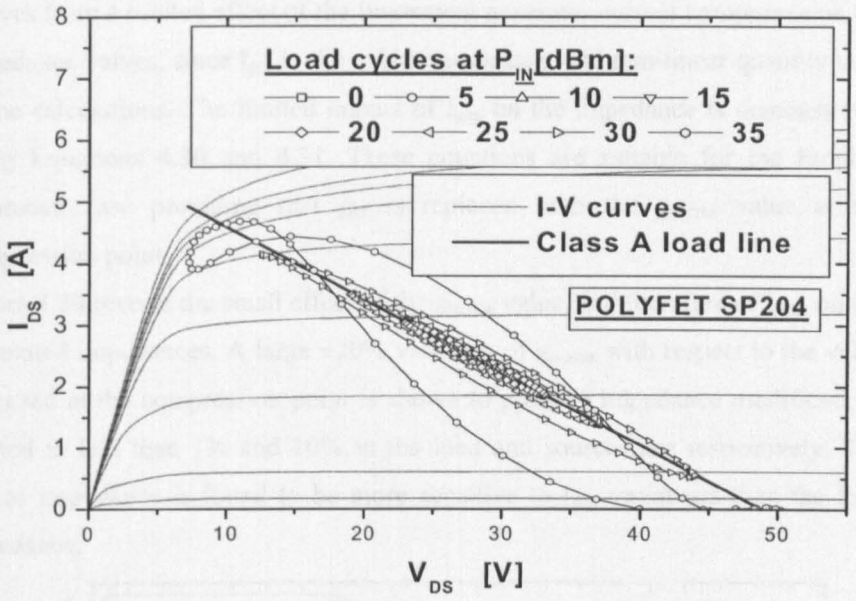


Figure 4.19: ADS HB simulated output signals superimposed the I_{DS}-V_{DS} curves and to the ideal load line in the single tone input signal case.

When the non linear current generator is considered in (iv)-(vi), the impedance values are extracted at the compression voltage. The real part of the source impedance prediction improves with the empirical method, but the device gain is overestimated.

Method	ZS [Ω]	ZL [Ω]	G [dB]
ADS HB simulation	2.3+3.05i	3.69+3.83i	12.38
(i) Work in [5] (No R_G and g_{DS})	2.09+4.1i	2.98+3.37i	14.28
(ii) This work, formula with R_G (g_{DS}=0S) (section4.3.2)	3.36+4. 1i	2.98+3.37i	12.90
(iii) This work, formula with R_G & g_{DS} (section4.3.2)	3.32+4.1i	2.96+3.42i	12.94
(iv) Empiric Approach (EA) with R_G & g_{DS}	2.80+3.641i	2.98+3.46i	13.13
(v) Fourier Fundamental Approach (FFA) with R_G & g_{DS}	3.32+4.09i	2.97+3.42i	12.25
(vi) Fourier Spectrum Approach (FSA) with R_G & g_{DS}	3.13+3.93i	2.97+3.43i	12.36

Table 4.2: Power gain at small input levels, optimum load and source impedance values.

Although the Empiric Approach (iv) will be later shown to produce miscalculated fundamental generator current component I_{gen}, source and load impedance predictions remain acceptable even in this case. This behaviour necessarily derives from a limited effect of the fundament generator current component on the impedance values, since I_{gen} is the only non-constant and non-linear quantity used in the calculations. The limited impact of I_{gen} on the impedance is demonstrated using Equations 4.30 and 4.31. These equations are suitable for the Empiric Approach case providing that g_m is replaced with the $g_{m,rms}$ value at the compression point.

Figure 4.20 reveals the small effect of the $g_{m,rms}$ value, and therefore of I_{gen}, on the calculated impedances. A large ±20% variation of $g_{m,rms}$ with respect to the value extracted at the compression point is shown to produce impedance modifications limited to less than 1% and 10% in the load and source case respectively. The source impedance is found to be more sensitive to I_{gen} variations than the load impedance.

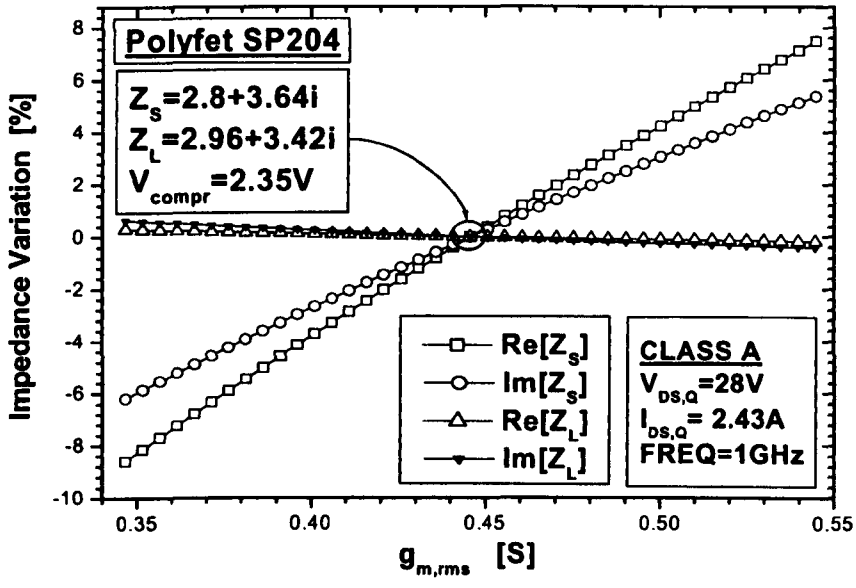


Figure 4.20: Effect of $g_{m,rms}$ variations on calculated source and load impedance in the single tone input signal case. The data is calculated applying Equations 4.30 and 4.31.

This last consideration provides a reading key for Figure 4.15 and Figure 4.16. In Figure 4.15 the calculated load impedances practically coincide. Since the

calculated load impedance displays small sensitivity to I_{gen} variations, the Z_L values mostly depend on the constant capacitance and resistance values of the equivalent lumped element circuit model used. The error in estimating $Z_{L,opt}$ can be attributed to the simple equivalent circuit topology used, that is a systematic error. In Figure 4.16, the increased sensitivity to I_{gen} causes the calculated Z_S values to spread out in the impedance plane, although remaining within the 45.29dBm delivered power contour. In this case, additionally to the systematic error due to modelling simplifications, errors in prediction also depend on the load line approximation, the accuracy of the load line extraction and the calculation of the fundamental component of the generator current, I_{gen}.

The fundamental generator current component I_{gen}

A comparison of calculated and ADS values of I_{gen} as a function of V_{gs} is shown in Figure 4.21. I_{gen} has been extracted from simulation by probing the device model internal current. This has been possible by the lumped element topology of the model used.

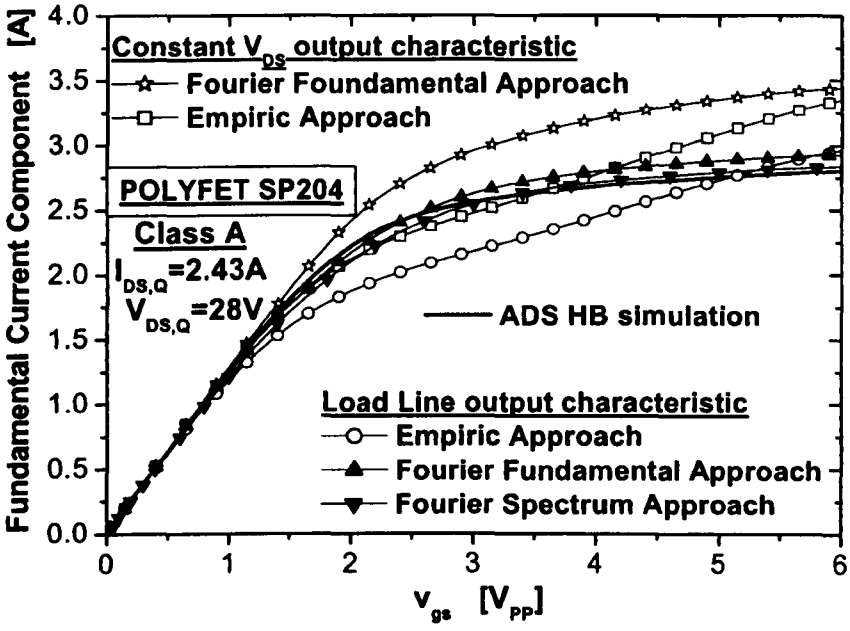


Figure 4.21: Calculation of the fundamental component of the generator current in the single tone input signal case.

When a constant V_{DS} output characteristic extracted at the quiescent drain voltage $V_{DS,Q}$ is used, the fundamental current component appears miscalculated regardless of the approach used: if a Fourier Fundamental Approach is used, the saturation value of the fundamental is considerably higher than in HB simulations. If the Empirical Approach of [7] is used, it leads to a calculated fundamental current component that does not saturate. This non physical behaviour is observed with the Empirical Approach even if the output characteristic extracted along the load line is used.

On the other hand, a good prediction of the fundamental component is achieved when $I_{loadline}$ is used in the Fourier Fundamental (FFA) or the Fourier Spectrum Approach (FSA). This implies that the load line extraction described in section 4.2 permits a good prediction of the actual load line cycles. This observation is acknowledged in Figure 4.19, where the device load cycles determined by ADS harmonic balance simulation remain close to the load line for input power levels $P_{IN} \leq 30$dBm. For input powers above 35dBm the load cycles deviate from the ideal load line. It is important to notice that at this power level the device is already beyond the 1dB gain compression point, which in this case corresponds to an input power level of 32dBm.

Power gain in the single tone input signal case

In this section the calculation of the power gain is carried out for a sweep of the input signal amplitude with the device biased in class A. The calculations are compared with the harmonic balance ADS simulation carried out as described at the beginning of this section. The output characteristic $I_{loadline}$ extracted along the load line is used. Power gain is compared for the three different methods described in section 4.3.3 for the determination of the fundamental component of the generator current:

(i) the Empiric Approach (EA);

(ii) the Fourier Fundamental Approach (FFA);

(iii) the Fourier Spectrum Approach (FSA).

The calculated power gain curves are compared with results of ADS simulation in Figure 4.22. The small signal power gain and 1dB gain compression point values are compared in Table 4.3.

A good prediction of power and 1dB compression point is achieved in the Fourier Fundamental and Spectrum Approach. Overestimation of the small signal power gain and underestimation of the 1dB compression point are instead found with the Empiric Approach. This behaviour is a consequence of the miscalculation of the fundamental generator current component shown previously in Figure 4.21.

Impact of class on one-tone power gain predictions

A verification of the proposed Fourier analysis based methods is carried out in Figure 4.23 and Figure 4.24 for different biases varying from class A ($I_{DS,Q}$=2.45A), mid class AB ($I_{DS,Q}$=1A), deep class AB ($I_{DS,Q}$=0.25A) and class B ($I_{DS,Q}$=0.02A). For each bias considered the load line characteristic and the optimum impedance values have been determined as described in the previous sections.

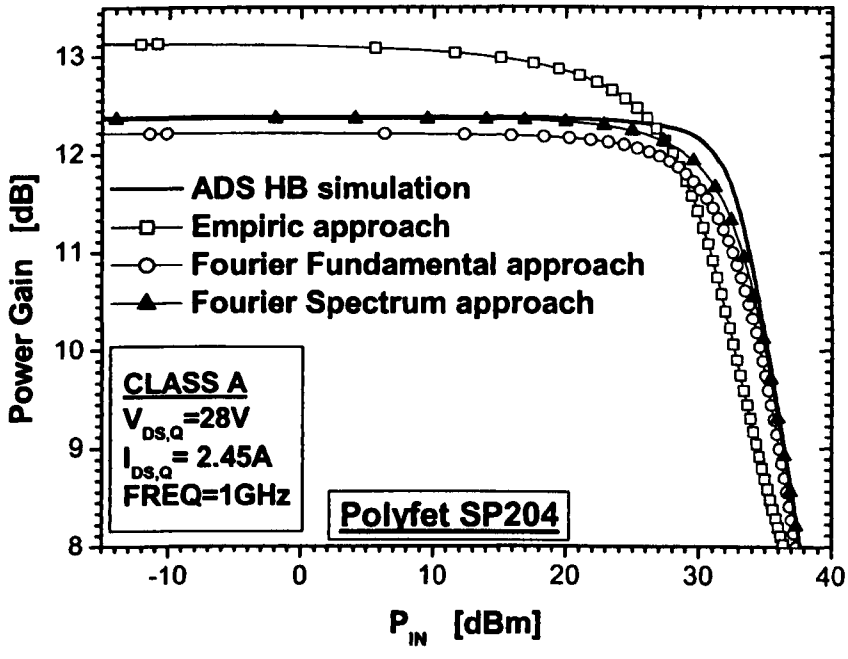


Figure 4.22: Comparison of the analytically calculated and ADS HB simulated power gains in the one tone input signal case.

Method	G [dB]	P1dB [dBm]
ADS HB simulation	12.38	44.1
Iloadline and Empiric Approach (EA)	13.13	39.76
Iloadline and Fourier Fundamental Approach (FFA)	12.25	43.09
Iloadline and Fourier Spectrum Approach (FSA)	12.36	43.706

Table 4.3: ADS HB simulated and calculated 1dB compression point in the single tone case.

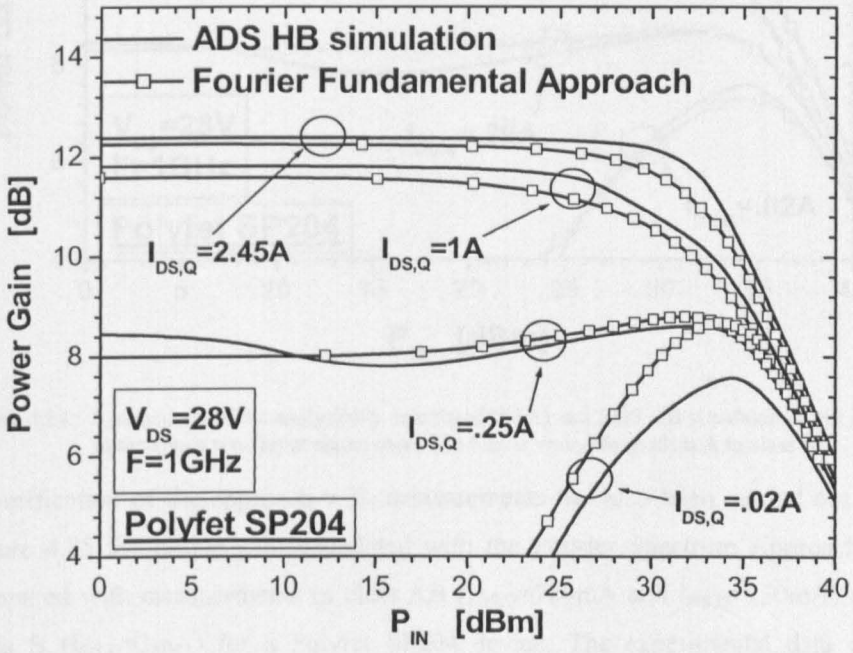


Figure 4.23: Comparison of the analytically calculated (FFA) and ADS HB simulated power gains in the one tone input signal case. The bias is varied from class A to class B.

The calculated results for the Fourier Fundamental (FFA) closely match the ADS values in class A and mid class AB. In deep class AB the precision of prediction reduces. This is a consequence of the load line approximation used in this work and of the load line current extraction method. At small gate voltages, due to the small currents involved, the extraction of $I_{Loadline}$ becomes challenging. This causes the power calculation to be less accurate in the small conduction angle deep AB and B classes. Similar considerations hold also for the Fourier Spectrum

Approach. However, the calculated results in this case show a considerable improvement of prediction in class B with respect to the FFA.

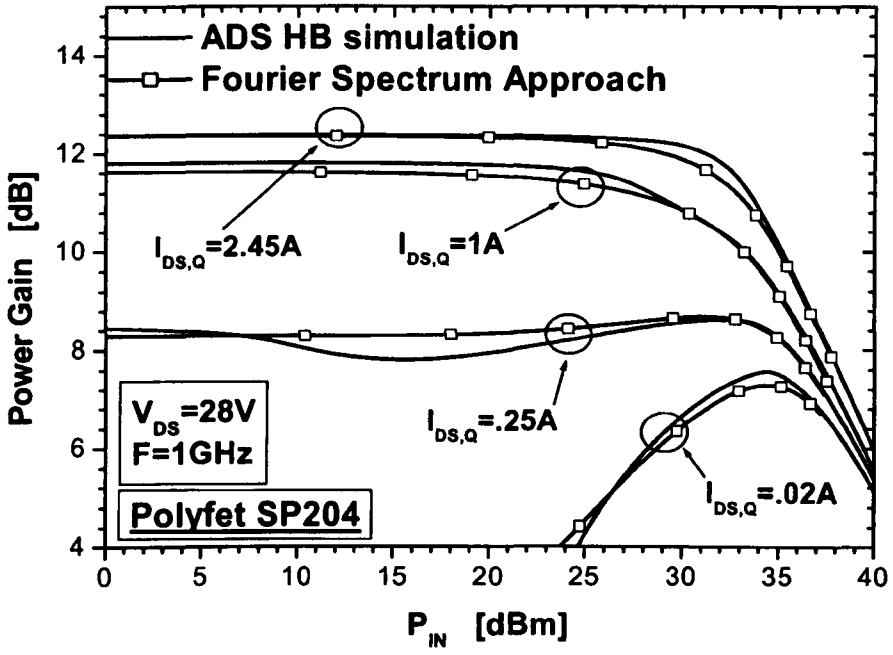


Figure 4.24: Comparison of the analytically calculated (FSA) and ADS HB simulated power gain in the single tone input signal case. The bias is varied from class A to class B.

A verification of the approach with measurements has also been carried out. In Figure 4.25 the power gain calculated with the Fourier Spectrum Approach is compared with measurements in class AB ($I_{DS,Q}$=780mA and $I_{DS,Q}$=250mA) and class B ($I_{DS,Q}$=3mA) for a Polyfet SP204 device. The experimental data was provided by Dr. J. Walker at SEMELAB PLC.

Measured I_{DS}-V_{DS} characteristics have been used in the extraction of the load line characteristic. The calculated power gain matches quite well with measurements in class AB, at least up to the 1dB gain compression point. On the other hand, the prediction in class B is not satisfactory. In this case, an accurate extraction of the load line is complicated by measurement errors. Furthermore, the quality of the impedance match in the measurement setup cannot cancel out the device capacitive content completely; hence the load to the current generator is not

purely resistive. This yields a deviation of the load cycles from the ideal load line, affecting the accuracy of the prediction.

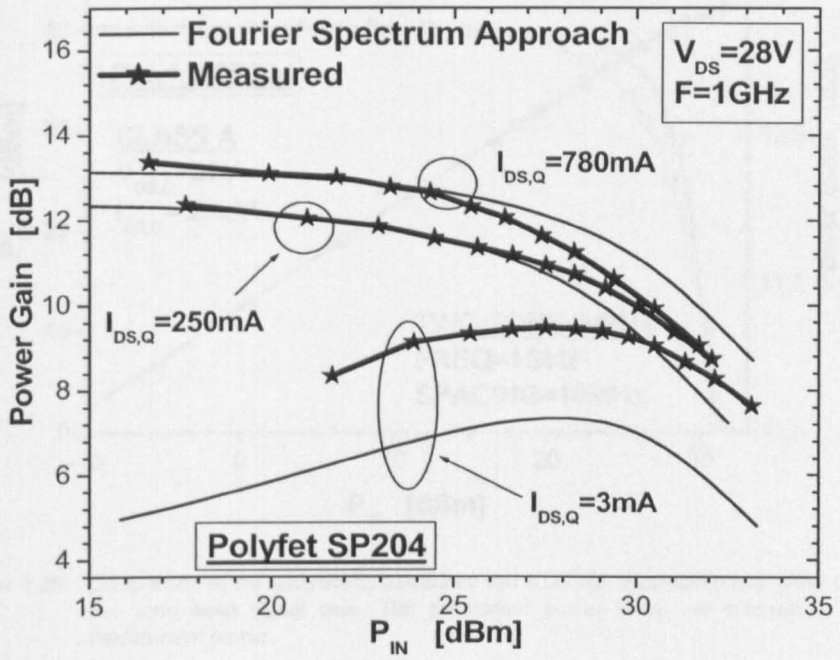


Figure 4.25: Comparison of the analytically calculated (FSA) and measured power gain in the single tone input signal case. The bias is varied from class AB to class B.

Class A Power gain in the two-tone input signal case

The validity of the prediction approach is demonstrated here in the class A two tone case only. A central frequency of 1GHz and a tone spacing of 10MHz have been considered for the input signal. The Fourier Spectrum Approach is suitable for this case. The calculation of power gain is shown in Figure 4.26. Good agreement with both power gain and gain compression point is obtained (Table 4.4).

Method	G [dB]	P_{1db} [dBm]
ADS HB simulation	12.367	41.65
Iloadline and Fourier Spectrum Approach (FSA)	12.37	41.56

Table 4.4: ADS HB simulated and calculated maximum power gain and 1dB compression point in the two tones input signal case.

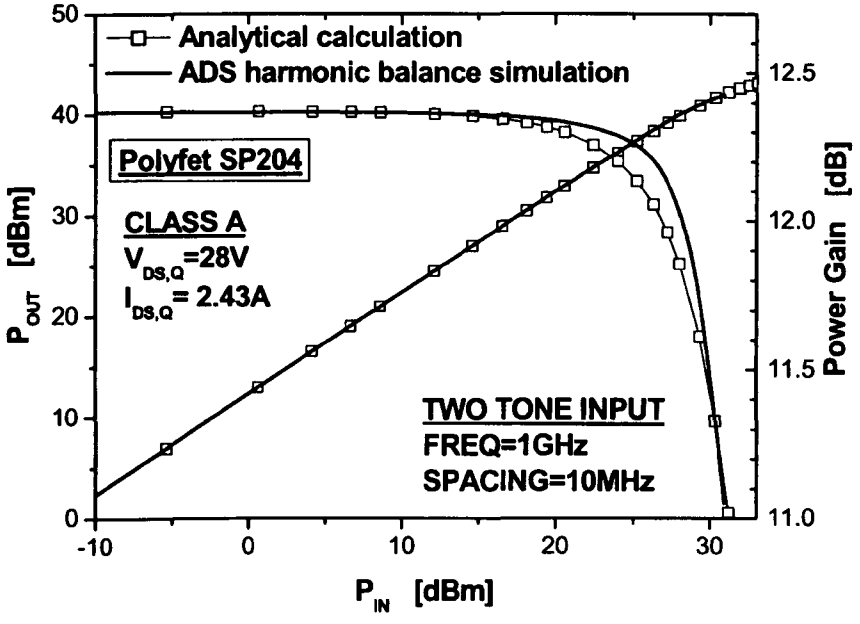


Figure 4.26: Comparison of the analytically calculated and ADS HB simulated power gains in the two tone input signal case. The considered power levels are referred to total fundamental power.

4.3.5 Applicability and limitations of the analysis

The equations presented in section 4.3.2 and 4.3.3 hold in all cases of continuous signal transmission for which a load line can be defined.

Load line approximation

With the load line approximation it has been assumed that the drain current and voltage at the RF frequency follow a trajectory close to the load line. This reasonable assumption does not fully correspond to the real case. As it has been shown in section 4.3.4, there exists a deviation of the trajectory from the load line. This deviation produces small errors at medium input signal power levels permitting a good estimation of the 1dB gain compression point. However, a large deviation from the ideal load line is observed at large power levels as a consequence of severe waveform clipping which reduces the accuracy of prediction.

Model simplification and harmonic balance approach

The methods described and the equations presented all assume the application of a known signal V_{gs} directly to the current generator of Figure 4.5. Constant values for the device capacitances are also considered in the analysis, making the current generator the only source of non linearity. In reality, the capacitors cannot be considered constant and V_{gs} is unknown but needs to be determined from the known amplifier input signal V_S. The calculation is simple in the ideal linear case. However, when the device non linearity is considered, a harmonic balance approach is required. This corresponds to the resolution in both the time and frequency domain of the equation

$$\frac{V_S}{2\,\mathrm{Re}[Z_S]} = I_{IN}, \tag{4.64}$$

which is derived considering the impedance Z_S as the conjugate match of the device input impedance. Substituting Equation 4.27 into 4.64 and I_{gen} to $g_m V_{gs}$, Equation 4.64can be rewritten highlighting linear and non linear parts according to the harmonic balance methodology [6]:

$$\underbrace{\frac{V_S}{2\,\mathrm{Re}[Z_S]}}_{\text{Linear side}} = \underbrace{\begin{bmatrix} C_{GD}(V_{GS}, V_{DS}, f))V_{gs} + \\ +j\omega C_{GD}(V_{GS}, V_{DS}, f)R_{opt} I_{gen} + j\omega(C_{GS}(V_{GS}, V_{DS}, f) \end{bmatrix}}_{\text{Non Linear side}}. \tag{4.65}$$

The solution of Equation 4.64 requires a harmonic balance non linear optimisation approach, which is beyond the scope of analytic modelling.

4.4 Assessment of linearity

This section describes an extension of known methods for the estimation of linearity. An analytical formulation is derived from the classical small signal analysis, where the load line output characteristic is again used to describe the large signal case. Novel generalised analytical expressions are proposed for the determination of inter modulation distortion. The expressions permit the inclusion of higher degree intermodulation products, going beyond the conventional 9^{th} order limit. A method for estimating linearity based on the implementation of the Fourier Spectrum Approach (FSA) is also presented. The proposed methods are shown to improve prediction with respect to conventional small signal formulation.

4.4.1 Background

When an input RF signal is applied to a non linear device, additional frequency components can be observed at the device output. Observing the spectrum of the output signal, two different non linear components are identifiable: harmonic distortion and inter modulation distortion. Harmonic distortion is represented by the existence of frequency components which are multiples of the original input signal frequency in the output signal spectrum. Inter Modulation Distortion (IMD) is observed for multi tone input signals and causes additional frequency components to be observed in the same range of frequency as the input signal.
Whilst harmonic distortion can be easily eliminated from the output spectrum through a simple low band pass filtering, this is not possible for IMD since it takes place in the application frequency range. This makes Inter Modulation Distortion the most important measure of linearity in RF communication applications.
IMD is calculated after the determination of the Inter Modulation distortion Products (IMP), which are defined as 'the additional frequencies at the output of a non-linear amplifier (or in general of any non-linear network) when two or more sinusoidal signals are applied at the input [13].

For a two tone input signal of fundamental frequencies $f_2>f_1$, corresponding to a carrier frequency $f_0=(f_1+ f_2)/2$ (central frequency) and modulating frequency $f_m=(f_2-f_1)/2$ (half the frequency spacing $\Delta=f_2-f_1$):

$$v_{gs}(t) = A[\cos(2\pi f_1 t) + \cos(2\pi f_2 t)] = 2A\cos(2\pi f_0 t)\cos(2\pi f_m t), \quad (4.66)$$

considering the n-th order power series approximation of the output signal waveform $i_{gen,\,wave}$

$$i_{gen,wave}(t) = c_0 + \sum_{k=1}^{n} c_k v_{gs}^k(t), \quad (4.67)$$

the IMD product expressions are determined by highlighting the frequency components of the output signal waveform. The complexity of the expressions increases with the order of the power series approximation. However, it has been shown in [14] that 9th order $i_{gen,\,wave}$ power series approximation yield satisfactory estimations of linearity in RF Power double diffused Si MOSFETs. Such a high order is necessary to properly take into account the abrupt power gain compression displayed by these devices. It has to be noted that the accurate linearity in [14] is achieved after a fitting procedure based on the analysis of intermodulation distortion measurements.

The IMP expressions relate the intermodulation products to the input signal amplitude A and the c_k coefficients in Equations 4.66 and 4.67. Substituting Equation 4.66 in 4.67 and applying Werner's trigonometric rules results in a relation between $i_{gen,\,wave}$ and frequency, input signal amplitude and power series coefficients. The terms of this relation can be arranged in different groups: the harmonics, presenting frequencies that are multiples of the original input signal frequencies, and the inter modulation distortion products, presenting frequencies that are a linear combination of the original input signal frequencies.

The so obtained intermodulation product expressions for a certain order i will result in

$$IMPi = \sum_{k=1}^{n} p_{ik} c_k' A^k, \quad (4.68)$$

$$c_k' = \begin{cases} c_k & \text{for } k \text{ odd} \\ 0 & \text{for } k \text{ even} \end{cases}, \quad (4.69)$$

where p_{ik} are the inter modulation product coefficients. The p_{ik} values determined up to the 9th order can be found in [14] and are reported in Table 4.5.

n	1	3	5	7	9
IMP1	1	9/4	25/4	1225/64	3969/64
IMP3	-	3/4	25/8	735/64	1323/32
IMP5	-	-	5/8	245/64	567/32
IMP7	-	-	-	35/64	567/128
IMP9	-	-	-	-	63/256

Table 4.5: Intermodulation Products Coefficients.

As shown in [15], considering that the coefficients c_k can be expressed as

$$c_k = g_{m,k} = \frac{1}{k!}\frac{d^k I_{DS}}{dV_{GS}^{\ k}}\bigg|_{V_{GS}=V_{GS,Q}}, \tag{4.70}$$

an expression for the third order intermodulation distortion is obtained as the ratio of the third order intermodulation product term and the term at the fundamental frequency:

$$IMD3[dBc] = 20\log\left|\frac{\frac{3}{4}g_{m,3}V_{gs}^{\ 3} + \frac{25}{8}g_{m,5}V_{gs}^{\ 5} + \frac{735}{64}g_{m,7}V_{gs}^{\ 7} + \frac{1323}{32}g_{m,9}V_{gs}^{\ 9}}{g_{m,1}V_{gs} + \frac{9}{4}g_{m,3}V_{gs}^{\ 3} + \frac{25}{4}g_{m,5}V_{gs}^{\ 5} + \frac{1225}{64}g_{m,7}V_{gs}^{\ 7} + \frac{3969}{64}g_{m,9}V_{gs}^{\ 9}}\right|. \tag{4.71}$$

In a similar way, the fifth order intermodulation distortion results in

$$IMD5[dBc] = 20\log\left|\frac{\frac{5}{8}g_{m,5}V_{gs}^{\ 5} + \frac{245}{64}g_{m,7}V_{gs}^{\ 7} + \frac{567}{32}g_{m,9}V_{gs}^{\ 9}}{g_{m,1}V_{gs} + \frac{9}{4}g_{m,3}V_{gs}^{\ 3} + \frac{25}{4}g_{m,5}V_{gs}^{\ 5} + \frac{1225}{64}g_{m,7}V_{gs}^{\ 7} + \frac{3969}{64}g_{m,9}V_{gs}^{\ 9}}\right|. \tag{4.72}$$

A visualisation of the intermodulation products for the determination of IMD3 and IMD5 is shown in Figure 4.27.

The limits of Equations 4.71 and 4.72 derive from the hypotheses on which they are based. As pointed out in [14], it is supposed that:

(i) the system response is independent of the previous state of the system (quasi-static assumption),

(ii) the bandwidth is limited to the vicinity of the input RF signal envelope itself (weak distortion),

(iii) the input RF signal envelope signal variations are slow when compared to the RF time domain (narrow fractional bandwidth).

The quasi-static hypothesis (i) prevents Equations 4.71 and 4.72 to predict the non symmetrical low and high side intermodulation products observable during measurements (low and high side intermodulation products correspond to values of the spectrum extracted at frequencies respectively lower or higher than the central signal f_0, as shown in Figure 4.27).

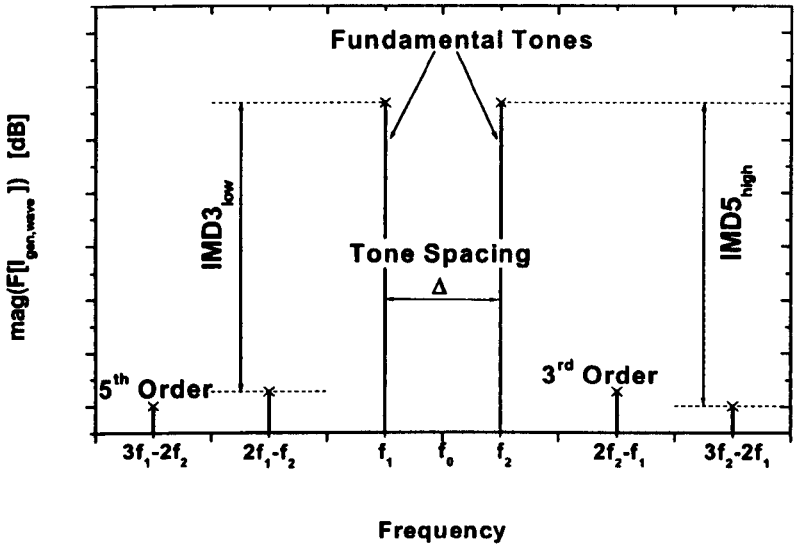


Figure 4.27: Typical two tone output spectrum detail in the proximity of the central frequency $f_0=(f_1+f_2)/2$, for a signal spacing $\Delta=f_2-f_1>0$. Highlighted are the low side third order ($IMD3_{low}$) and high side fifth order ($IMD5_{high}$) intermodulation distortion. Low and high sides corresponds to frequencies respectively lower or higher than f_0. On each side, Intermodulation distortion is determined as the difference in decibel between the intermodulation product amplitude and the amplitude of the fundamental frequency.

The narrow fractional bandwidth assumption and the weak distortion approximation yield the coefficients c'_k to be null in correspondence with even k values, as acknowledged in Equation 4.69. Hypotheses (ii) and (iii) imply that the even order distortion products remain out of band. This leads to a inaccurate linearity prediction in a class B or deep class AB power amplifier, where second degree effects play an important role [14].

Contrary to the work in [14], the c_k values in Equations 4.68-4.72 are not determined by fitting the measured intermodulation distortion. They are instead

calculated from the output characteristic, after [15], allowing for the intermodulation distortion prediction prior to any RF measurement. However, the so calculated c_k coefficients do not take into account information on the phase of the intermodulation products, affecting the prediction accuracy. Also, the c_k values of Equations 4.68-4.72 are extracted at the bias voltage and determined as derivatives of the output current. Therefore, they are small signal parameters. This implies that their validity holds just for small amplitudes of the input envelope.
Finally, it is important to point out that the determination of linearity through the application of Equations 4.71 and 4.72, is carried out considering the device transfer-characteristic I_{DS}-V_{GS} at a constant V_{DS}=$V_{DS,Q}$, as acknowledged by Equation 4.70.

4.4.2 Methodology for linearity estimation

The limitations of the conventional analytical approach for inter modulation distortion prediction have been identified in section 4.4.1. In particular, the conventional approach has been shown to be based on an implicit small signal assumption and on the use of the constant V_{DS}=$V_{DS,BIAS}$ output characteristic.
In order to overcome these limitations, two approaches to predict inter modulation distortion are proposed in this section: the Large Signal Fitting Approach (LSFA) and the Fourier Spectrum Approach (FSA).

Large Signal load line Fitting

Linearity is determined after calculating:

(i) the load line output characteristic I_{out}=$I_{loadline}$ from the device I_{DS}-V_{DS} characteristics, as described in section 4.3.3,

(ii) large signal coefficients from the power series approximation of the load line output characteristic $I_{loadline}$,

(iii) the intermodulation product amplitudes IMPi.

Step (ii) consists of the n-th degree power series fitting centred at the quiescent bias voltage $V_{GS,Q}$ of the output characteristic $I_{loadline}$. This corresponds to the determination of the large signal c_{LS-k} coefficients of

$$I_{out} = c_{LS-0} + \sum_{k=0}^{n} c_{LS-k} (V_{GS} - V_{GS,Q})^k \,. \tag{4.73}$$

Step (iii) consists in the application of

$$IMPi = \sum_{k=1}^{n} p_{ik} c'_{LS-k} A^k \,, \tag{4.74}$$

$$c'_{LS-k} = \begin{cases} c_{LS-k} & for\ k\ odd \\ 0 & for\ k\ even \end{cases}, \tag{4.75}$$

where the p_{ik} are the inter modulation product coefficients.

Intermodulation distortion is then determined as

$$IMDn = 20\log\left(\frac{IMPn}{IMP1}\right), \tag{4.76}$$

where IMP1 is the single tone (low or high side tone amplitude) fundamental current component and IMPn is the single tone n-th order intermodulation product amplitude.

A novel expression for p_{ik} is presented here in order to extend the application of Equation 4.74 to orders higher than the 9th, for which the p_{ik} coefficients are available:

$$p_{ik} = 2^{1-i} \frac{i!}{\left(\frac{i-1}{2}\right)!\left(\frac{i+1}{2}\right)!} \frac{i!}{\left(\frac{i-k}{2}\right)!\left(\frac{i+k}{2}\right)!} \,. \tag{4.77}$$

The derivation of Equation 4.77 is described in the following.

Considering a two tone signal

$$v_{gs}(t) = 2A\cos(2\pi f_0 t)\cos(2\pi f_m t) \,, \tag{4.78}$$

the corresponding output current signal can be expressed from Equation 4.73 as

$$i(t) = I_{out}(v_{gs}(t) + V_{DS,Q}) - I_{out}(V_{DS,Q}) \,, \tag{4.79}$$

$$i(t) = \sum_{k=1}^{n} c_{LS-k} v_{gs}(t)^k \,, \tag{4.80}$$

where I_{out} is the device output characteristic.

After substitution of Equation 4.78 into Equation 4.80, and expansion of the n-th power of the cosine function using [14]

$$(\cos\theta)^n = \frac{1}{2^{n-1}}\sum_{j=0}^{n-1}\binom{n}{j}\cos[(n-2j)\theta], \tag{4.81}$$

the i-th degree component of the current i_i results as

$$i_i = c_{LS-i}v_{gs}(t)^i, \tag{4.82}$$

$$i_i = 2^{2-i}A^i c_{LS-i}\left[\sum_{k_m=0}^{n-1}\binom{i}{k_m}\cos[(i-2k_m)2\pi f_m t]\right]\left[\sum_{k_0=0}^{n-1}\binom{i}{k_0}\cos[(i-2k_0)2\pi f_0 t]\right]. \tag{4.83}$$

Assuming a narrow band and weak distortion approximation, the k-th order intermodulation component of the current i_i is identified as the component corresponding to the $\cos(k\cdot 2\pi f_m)\cos(2\pi f_0)$ term, which result as

$$i_{ik} = \left[2^{2-i}\binom{i}{(k-i)/2}\binom{i}{(k+1)/2}\right]c_{LS-i}A^i\cos(k\cdot 2\pi f_m t)\cos(2\pi f_0 t). \tag{4.84}$$

The p_{ik} coefficient can now be determined as half the square brackets term in Equation 4.84. The half factor derives from considering symmetrical the contribution from low and high side current spectrum.

Fourier Spectrum Analysis

Linearity is determined from the current output signal frequency spectrum $F[I_{gen,wave}]$. $F[I_{gen,wave}]$ is determined in three steps by calculating:

(i) the load line output characteristic $I_{loadline}$ from the device I_{DS}-V_{DS} characteristics, as described in section 4.3.3

(ii) the current signal waveform $I_{gen,wave}$ from $I_{loadline}$, as described in section 4.3.3

(iii) the Fourier Transform of $I_{gen,wave}$, implementing Equation 4.52 through a FFT algorithm

Third order low and high side intermodulation distortion are then calculated from the signal spectrum $F[I_{gen,wave}]$ as:

$$IMD3_{low}[dBc] = 20\log\left|\frac{F[I_{gen,wave}(t)]\big|_{f=(2f_1-f_2)}}{F[I_{gen,wave}(t)]\big|_{f=f_1}}\right|. \tag{4.85}$$

$$IMD3_{high}[dBc] = 20\log\left|\frac{F[I_{gen,wave}(t)]\big|_{f=(2f_2-f_1)}}{F[I_{gen,wave}(t)]\big|_{f=f_2}}\right|. \quad (4.86)$$

Analogously, the low and high side fifth order Intermodulation distortion result as:

$$IMD5_{low}[dBc] = 20\log\left|\frac{F[I_{gen,wave}(t)]\big|_{f=(3f_1-2f_2)}}{F[I_{gen,wave}(t)]\big|_{f=f_1}}\right|. \quad (4.87)$$

$$IMD5_{high}[dBc] = 20\log\left|\frac{F[I_{gen,wave}(t)]\big|_{f=(3f_2-2f_1)}}{F[I_{gen,wave}(t)]\big|_{f=f_2}}\right|. \quad (4.88)$$

4.4.3 Results

In order to provide a benchmark for the assessment of the accuracy of the proposed methodology, harmonic balance simulations have been used. The simulations are again carried out in ADS [9] on the Polyfet SP204 model [10]. The basic schematic setup used in power gain simulations is shown in Figure 4.28. Ideal components such as RF chokes, DC feeds and shorts instead of transmission lines have been used in the simulations.

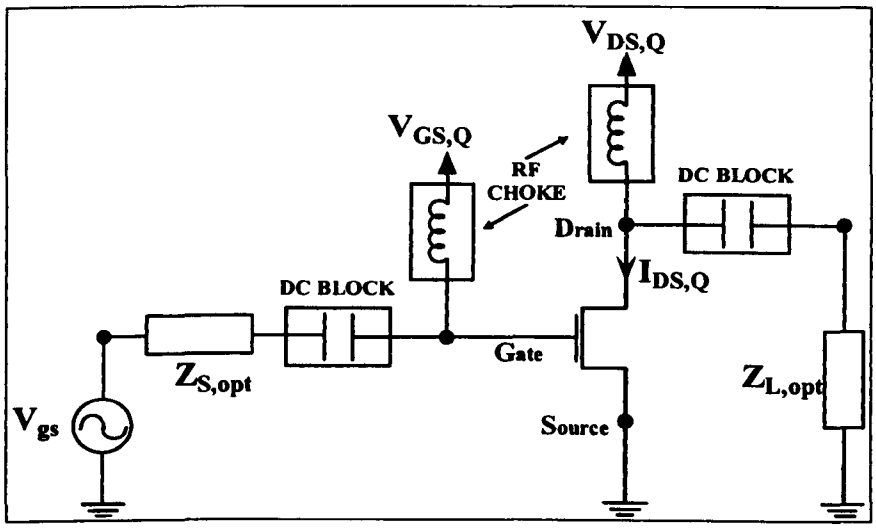


Figure 4.28: One-tone and two-tone ADS simulation setup.

Monitoring internal voltages and currents has permitted the determination of linearity considering only the generator current. Also, it has been possible to express linearity as a function of the peak to peak envelope amplitude. This allows for a direct comparison between harmonic balance simulations and the analytic calculations, permitting correct assessment of the accuracy of the prediction of Equations 4.71-4.72 and 4.85-4.88.

Analytical calculations vs HB simulations

Intermodulation distortion is compared for:

(i) the 'Small Signal Formula' of Equations 4.71 and 4.72, where the parameters are extracted from the constant $V_{DS}=V_{DS,Q}$ output characteristic,

(ii) the 'Small Signal Formula' as in (i), but considering the load line current characteristic $I_{loadline}$,

(iii) the Large Signal load line Fitting Approach (LSFA), as described in section 4.4.2, but considering the constant $V_{DS}=V_{DS,Q}$ output characteristic,

(iv) the Large Signal load line Fitting Approach (LSFA), as described in section 4.4.2,

(v) the Fourier Spectrum Analysis method, as described in section 4.4.2,

(vi) HB simulation.

Cases (i) and (ii) as well as cases (iii) and (iv) differ only in the output characteristic considered. The two output characteristics have been previously shown in section 4.2, Figure 4.4. The validity of the load line characteristic in representing the device load cycle has been demonstrated in section 4.3. The comparison with the constant $V_{DS}=V_{DS,Q}$ output characteristic is carried out in order to assess which kind of miscalculation derives from its employment in a prediction of linearity.

When compared to the constant V_{DS} output characteristic, the load line characteristic presents a considerably reduced maximum current and a narrower transconductance. The peak transconductance value is not considerably different

in the two cases. The reduction of the g_m bell width in the load line case implies a possible degradation of linearity with respect to the constant V_{DS} case.

The calculated and ADS simulated third order intermodulation distortion are plotted in Figure 4.29 and Figure 4.30.

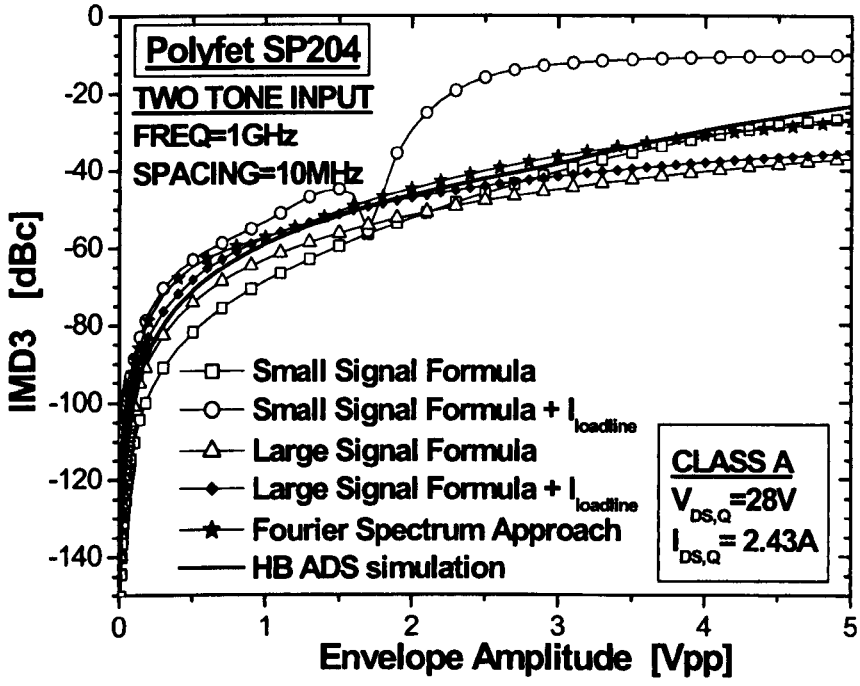


Figure 4.29: Comparison of the ADS HB simulated and calculated third order intermodulation distortion.

The degradation of linearity associated with considering the load line instead of the constant V_{DS} output characteristic is acknowledged by the small signal conventional algorithm as well as by the large signal algorithm. Both small and large signal approaches are capable of capturing trend behaviours, associating larger third order intermodulation distortion to the load line case.

However, when the load line is used, a comparison with HB results reveals that the small signal approach displays considerably larger errors in estimating linearity than the large signal approach. In particular, the small signal approach reveals a sweet spot that is not observed in ADS HB simulation. This behaviour derives from the incapacity of the small signal approach of correctly reproducing the device output characteristic, as shown in Figure 4.31.

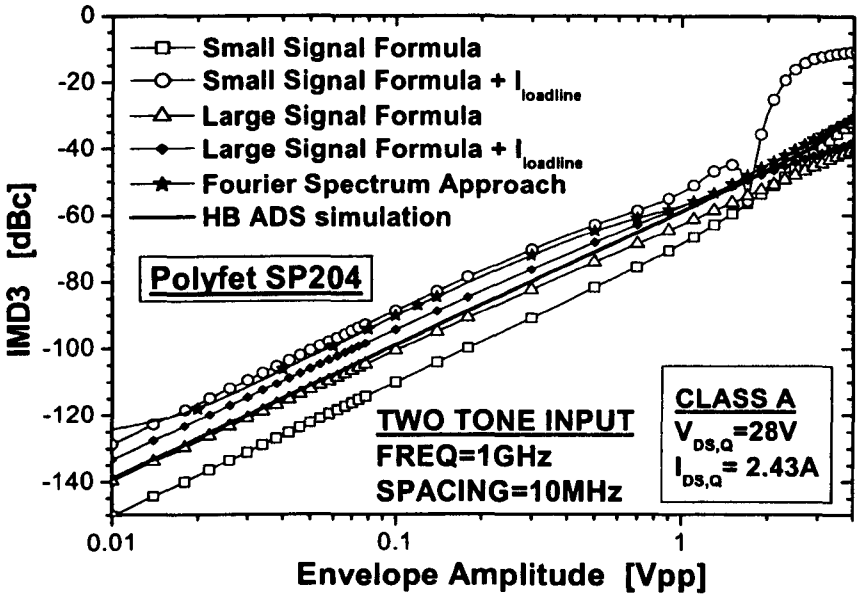


Figure 4.30: Comparison of the ADS HB simulated and calculated third order intermodulation distortion. A logarithmic scale is used for the two tone envelope peak to peak voltage.

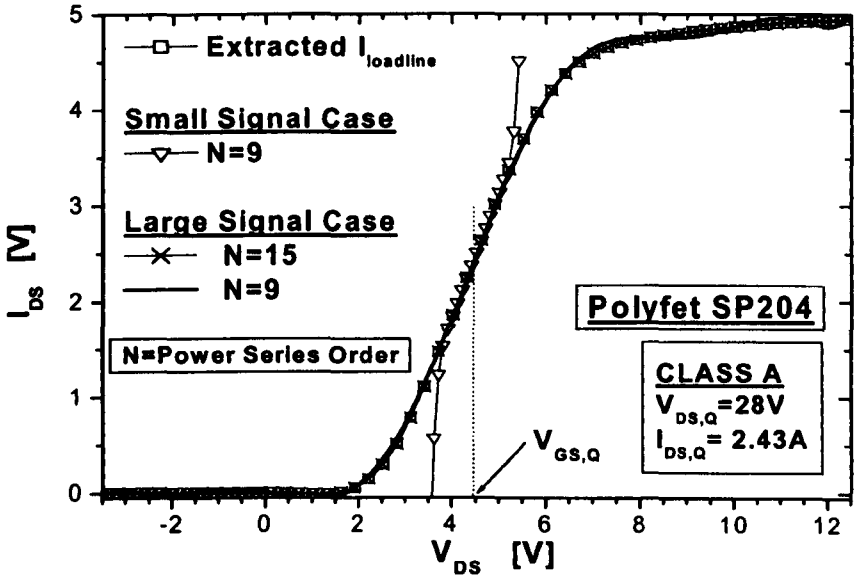


Figure 4.31: Load line current characteristic reconstruction in the small signal (ii) and in the large signal case (iv). The small signal case approximation holds only for small voltages amplitudes. In the large signal case an acceptable fit can be achieved already for 9th order power series approximations. Oscillations in the current prediction can be avoided by increasing the order of the power series.

On the other hand, the large signal fitting approach permits a correct reconstruction of the output characteristic which leads to the improved estimation of IMD3. The analytical expression of the p_{ik} coefficients presented in section 4.4.2 permits the simple extension of the IMD3 calculation to power series of order beyond the 9^{th}, yielding improved reconstruction of the output characteristic. Consequently, the 3^{rd} order intermodulation distortion estimation is more accurate than in the small signal case and valid for a larger range of two tone envelope amplitudes.

However, for two tone envelope amplitudes above 2Vpp, Figure 4.29 and Figure 4.30 show that the large signal approach prediction worsens. This is a consequence of the modelling simplification used in the derivation of the approach and in particular of the narrow bandwidth as well as weak distortion assumptions and of the impossibility of including phase information in the intermodulation distortion calculation.

Good linearity predictions are instead achieved at such signal amplitudes in the Fourier Spectrum Analysis case (v), where no particular assumption apart from the steady state approximation is considered. The correct prediction in this case is a consequence of the proper consideration of the intermodulation current phase. This information is in fact available after the application of the Fourier transform to the calculated current signal waveform.

Unfortunately, the Fourier Spectrum Analysis is unsuitable for resolving small intermodulation products amplitudes. As pointed out in [16] this is a typical problem in computer aided analysis and is a consequence of numerical errors taking place when small intermodulation products are considered.

4.4.4 Applicability and limitations of the analysis

In the proposed methods, the amplifier linearity is calculated as the current generator linearity, neglecting the effects of the non linear capacitances. As for the power gain calculation of section 4.3, it is not possible to correctly identify the actual input waveform to the device current generator.

The load line fitting large signal approach removes the small signal limitation of the conventional approach, but maintains the quasi-static, narrow fractional bandwidth and weak distortion assumptions of the conventional method.

Although no assumptions of the output signal spectrum are considered, the quasi-static hypothesis is maintained in the Fourier Spectrum Analysis case. As pointed out in [16], the limit of any Fourier based approach for the determination of linearity is constituted by the large computational times involved. A very large number of data points might be necessary in order to correctly determine the signal spectrum at two closely spaced frequencies. Other calculation issues derive from the introduction of aliasing and interpolation errors [16].

4.5 Summary

This chapter demonstrates a way to use the load line concept towards the prediction of the amplifier performance. An extraction procedure has been proposed and described for the load line output characteristic. It has been shown that only the load line characteristic permits correct prediction of the device performance.

An analytic formulation of the large signal input and output power of Si RF power MOSFETs has been presented. Improved power gain and optimum matching impedance expressions are provided. The inclusion of the effect of gate and drift region resistance has been shown to improve the accuracy of the power gain prediction. The effects of a realistic non linear current generator have been taken into account in a novel methodology for the prediction of gain compression. The methodology has been demonstrated for A, B and AB classes of operation for both the single and two tone input case. The use of the proposed analytic formulation and basic Fourier analysis has been shown to permit reasonably accurate prediction of gain and gain compression.

Following the load line output characteristic extraction, an analytical formulation has been proposed to predict intermodulation distortion. The approach is based on

the large signal fitting of the output characteristic where an arbitrary high degree of approximation can be used for the fitting polynomial. The approach has been shown to correctly estimate linearity. Improved prediction of intermodulation distortion has then been demonstrated by using a Fourier analysis based approach.

References

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

[2] S.A. Maas, Nonlinear Microwave and RF circuits, Artech House microwave library, 2003

[3] R. Gilmore, L. Besser, Practical RF circuit design for modern wireless systems, Artech House Publishers, 2003

[4] S. Cripps, "A Method for the Prediction of Load-Pull Contours in GaAs MESFETs", 1983 IEEE MTT-S Int. Microwave Symp. Dig., pp. 221-223, June 1983.RF Power Amplifiers for Wireless Communications. Norwood, MA: Artech House, 1999.

[5] J. Walker, Analytic expressions for the optimum source & load impedance and associated large-signal gain of an RF power transistor, IEEE MTT-S, 2003, vol. 3 , pp. 1725 – 1728.

[6] P. Fioravanti, O. Spulber, M.M. De Souza, "Analytical Large Signal Modeling of Silicon RF Power MOSFETs", submitted to IEEE Transactions on MTT.

[7] Trivedi, P. Khandelwal, K. Shenai, Performance Modeling of RF Power MOSFET's, IEEE Transactions on Electron Devices, 1999, vol. 46, pp. 1794-1801

[8] S.C. Cripps, RF Power Amplifiers for Wireless Communications, Artech House, 1999.

[9] Agilent EEsoft ADS version 2003c. Help guide.

[10] Polyfet website, www.polyfet.com.

[11] W. Curtice, J. Plá, D. Bridges, T. Liang, E. Shumate, A New Dynamic Electro-Thermal Nonlinear Model for Silicon RF LDMOS FETs, 1999 IEEE MTT-S International, pp. 419-422

[12] G. Dambrine, A. Cappy, F. Heliodore, E. Playez, A new method for determining the FET small-signal equivalent circuit, IEEE Trans. Microwave Theory and Techniques, 1988, vol. 36, pp. 1151-1159

[13] Beneking H., High Speed Semiconductor Devices, Chapman & Hall, 1994

[14] S.C. Cripps, Advanced Techniques in RF Power Amplifier Design, Artech House, 2002

[15] van der Heijden M.P., de Graaf H.C., de Vreede L.C.N., Gajadharsing J.R. and Burghartz J.N., Theory and Design of an Ultra-Linear Square-Law Approximated LDMOS Power Amplifier in Class-AB Operation, IEEE Trans. on microwave theory and techniques, 2002, pp.2176-2184

[16] K.Mayaram, D.C.Lee, S.Moinian, D.Rich, J.Roychowdhury, Computer-Aided Circuit Analysis Tools for RFIC Simulation: Algorithms, Features, and Limitations, IEEE Trans. on ciruicts and systems-II, Analog and digital signal processing, vol.47,no. 4, April 2000

5

RF based optimisations

Abstract

The effect of a shield plate in a vertical RF Power DMOSFET is analysed. Different shield topologies are studied by providing physical insight into the effect of the design parameters on the device behaviour. The dependence of the peak value of the transconductance on the drift region resistance is identified and explained.

An innovative approach for device optimisation is proposed and implemented in the identification of the best possible design. The approach is based on the definition of a novel figure of merit for the resolution of the device design trade offs via determination of power gain and linearity. The optimum shield design is identified and its RF performance assessed.

A novel device concept is presented. The device is equipped with a grounded p-region which replaces the shield plate. The novel concept is optimised and compared to the shielded RF VDMOSFET.

5.1 Introduction

Due to their small area and simple fabrication technology, field plates have been extensively used in power device design as planar junction termination in order to increase the junction breakdown voltages [1]. In [2] and [3] an analytic description has been carried out to model the effect of the plate on the electric field distribution. The field plate in vertical RF power devices has been used for several years [4][5][6][7]. The plate is shorted to the source electrode and is placed over the drift region between the gates of adjacent cells. By being shorted to the source, the field plate is grounded in RF power applications. Hence the MOSFET is biased in common source configuration. By that, the source is connected to the circuit ground while gate and drain are connected to the input and the output of the amplifier respectively. The grounded field plate, often called shield plate, is not a Faraday shield but provides capacitive decoupling between input and output ports of the common source power amplifier by significantly reducing the feedback capacitance. In recent designs, the shield plate has been used to increase the radio frequency performance of vertical MOSFET [8][9][10], where the gate electrode is cut to cover the channel region only. In these devices the shield plate permits a large reduction of the feedback capacitance content and an increase of the drift doping concentrations. It has been shown that, with respect to an unshielded VDMOSFET, a vertical DMOS device equipped with a shield plate presents a more uniform potential distribution which determines a significant reduction of the electric field and an improvement in breakdown voltage capability [9][10]. However, up do date there has been no discussion on the effect of position or dimension of the shield on RF performance. In [8] the possibility of optimising the shield plate in vertical structures is proposed, but not further investigated.

Typical power device optimisations rely on the resolution of the breakdown voltage (BV) versus on resistance ($R_{DS,on}$) trade off. On the other hand, RF Power devices for communications applications require aggressive design rules to achieve best possible gain and linearity. These are two additional figures of merit

in comparison to power devices. Unlike BV-$R_{DS,on}$ however, gain and linearity are inaccessible in device simulators [11][12]. Consequently, addressing the trade off between power gain and linearity results in the major challenge for the RF engineer. This trade off is evident at the device level in terms of capacitance content and transconductance behaviour. A decrease in gate to drain capacitance is a clear sign of increased power gain. As pointed out in [7], the flatness of the transconductance characteristic is an indicator of good linearity performance. However, although relevant, these remarks are not suited to be used as optimisation criteria at the device design level.

This chapter examines the effects on the electrical performance of a standard RF Power VDMOSFET that occur with a field plate connected to the source. A novel figure of merit is introduced in order to address the linearity versus power gain trade offs and to permit the identification of an optimum RF design. A novel device concept is also proposed. The new design can be though as an evolution of the shield plate design to further reduce the feedback capacitance content.

The novel figure of merit is presented in section 5.2 and then extensively applied throughout the chapter. The examined shield plate configurations are described in section 5.3. The shield effect analysis begins with the description of the effect on the breakdown voltage in section 5.3.2, and is followed by the determination of the optimum shield plate dimensions in section 5.3.3. In this section only the shield position is taken into account and the other design parameters (drift doping, cell pitch, etc.) remain unchanged. The variation of these design parameters for RF performance optimisation is then described in section 5.3.4. The optimum shielded VDMOSFET is then compared with the unshielded one in section 5.3.5.

The novel GP-VDMOSFET design concept is proposed in section 5.4 and optimised in section 5.4.3. The effects of the grounded p-region and of the shield plate on the VDMOSFET performance are compared in section 5.4.4.

5.2 RF optimisation: the Unified Figure of Merit

The value of a figure of merit could be strictly related to its predictive capability without compromising usability. This section proposes a new figure of merit to simplify the optimisation of RF power devices. The goal is to define a physically derived and reasonably accurate tool for addressing trends and RF performance tradeoffs.

A convenient definition for the novel figure of merit is

$$UFM = -\left(G_{[dB]}\right)^{\alpha}\left(IMD3_{[dBc]}\right)^{\beta}. \qquad (5.1)$$

where $IMD3_{[dBc]}$ is the third order intermodulation distortion expressed in dBc, $G_{[dB]}$ is the power gain expressed in dB, α and β are positive constants that represent the optimisation weights. The acronym UFM stands for Unified Figure of Merit, which acknowledges the composite nature of the figure of merit.

The adjustment of the optimisation weights in Equation 5.1 permits to give more relevance to either power gain or linearity during the device optimisation process. If power gain is the main optimisation target then $\alpha>\beta$; if linearity is the main optimisation criterion then $\alpha<\beta$. As in communication, in the case of power gain and linearity having same importance, Equation 5.1 can be rewritten as

$$UFM = -G_{[dB]}IMD3_{[dBc]}. \qquad (5.2)$$

Equation 5.2 is extensively used in this chapter for the optimisation of the shield VDMOSFET device. Since the UFM intends to represent an ideal performance limit of the device, the best power gain and the best linearity performance are considered in its calculation. Suitable procedures for these tasks are described in this section.

5.2.1 Desired linear output power and device width

When working with simulators [11][12], the device is generally defined through the description of its two dimensional cross-section. The device extension in the third dimension is typically assumed to be 1μm. The simulation results are therefore referred to this device width.

In order to estimate RF performance, the total device width has to be considered. A simple estimation of its value can be obtained as a function of the target output linear power for the device that is being designed. From the maximum linear power expression [4][5], given the desired maximum output linear power $P_{OUT,linear}$, the total device width results in

$$W = \frac{8P_{OUT,linear}}{V_{MAX} I_{MAX-1\mu m}}, \quad (5.3)$$

$$V_{MAX} = 2(V_{DS,Q} - V_{knee}), \quad (5.4)$$

where $I_{MAX-1\mu m}$ is the maximum linear current extracted at the knee voltage V_{knee}, with the suffix -1μm indicating that the values refer to a device gate width of 1μm, and V_{MAX} is the maximum voltage swing. The total device width is an extension of the device in the third dimension required for the generation of the target output power (current in power applications).

Once the device width has been determined, the device parameters can be dimensioned considering the parallel contribution of the number of 1μm devices required to obtain such a total width. The higher the device width, the larger current, transconductance and capacitance values result, whereas the load line resistance decreases with W.

5.2.2 Maximum power gain Determination

The power gain expression proposed in Chapter 4

$$G = \frac{R_{opt} g_m^2}{\left[\begin{array}{c}(C_{GS} + C_{GD}(1 + R_{opt} g_m))^2 R_G + \\ + L_S g_m (1 - g_{DS} R_{opt})(C_{GS} + C_{GD} + C_{GD} R_{opt} g_m)\end{array}\right]\omega^2}, \quad (5.5)$$

is used by considering an equivalent circuital model and imposing the optimum load line resistance R_{opt} on the device current generator. An expression for the maximum device power gain can be obtain from Equation 5.5 by substituting the optimum load line resistance R_{opt} with the device output resistance R_{out} and imposing g_{DS}=0S. As pointed out in [16], a similar procedure is operated in order to achieve the maximum gain expression in the small signal case. Following the

resolution of the equivalent device model described in Chapter 4, as also resulting from the analysis of the device equivalent model in [16], R_{out} can be expressed as

$$R_{out} = \frac{C_{GS} + C_{GD}}{g_m C_{GD}}. \qquad (5.6)$$

When a value different from R_{opt} is imposed on the current generator, a reduction of the maximum linear power is observed. On the other hand, the imposition of R_{out} yields the maximum power gain at small signal. The differences between the two cases are qualitatively highlighted in Figure 5.1. Using the maximum g_m value yields larger gain and therefore a better prediction of the maximum device gain.

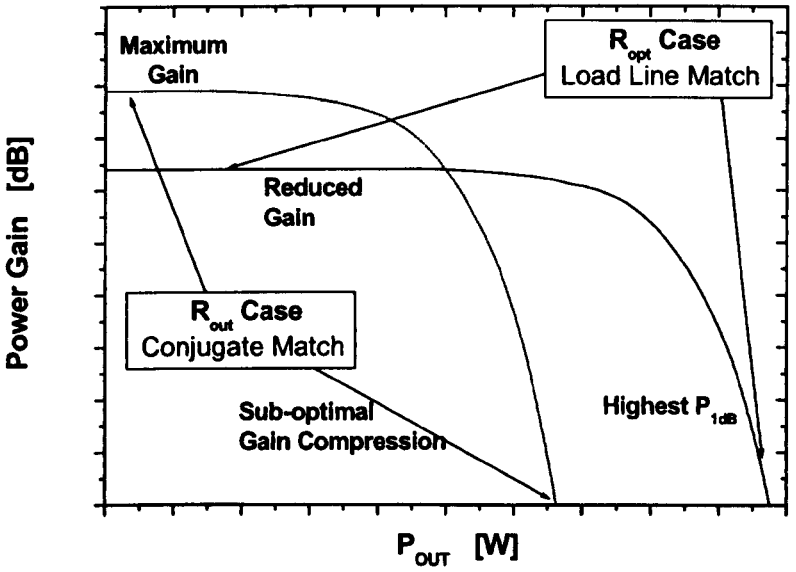


Figure 5.1: Qualitative comparison of power gain for maximum linear power and maximum gain.

Substituting R_{opt}, the maximum power gain expression for the UFM determination gives

$$G = \frac{g_m}{2C_{GD}[2(C_{GS} + C_{GD})R_G + L_S g_m]\omega^2}, \qquad (5.7)$$

where ω is the angular frequency, L_S is the source inductance, R_G is the gate resistance, g_m, C_{GS} and C_{GD} are the transconductance, the gate to source capacitance and the gate to drain capacitance of the device respectively. The gain

calculated applying Equation 3.3 represents the maximum gain that the considered device is capable of. The values to be used in Equation 3.3 are those extracted at the application voltage, where g_m is the maximum transconductance value. For the UFM calculation, Equation 3.3 has to be calculated in class A, where the maximum power gain is achieved.

5.2.3 Linearity determination

Linearity is calculated in class A by applying the large signal load line fitting approach presented in Chapter 4 and described by

$$IMDn = 20\log\left(\frac{IMPn}{IMP1}\right), \tag{5.8}$$

$$IMPi = \sum_{k=1}^{n} p_{ik} c'_{LS-k} A^k, \tag{5.9}$$

$$c'_{LS-k} = \begin{cases} c_{LS-k} & \text{for } k \text{ odd} \\ 0 & \text{for } k \text{ even} \end{cases}, \tag{5.10}$$

$$p_{ik} = 2^{1-i} \frac{i!}{\left(\frac{i-1}{2}\right)!\left(\frac{i+1}{2}\right)!} \frac{i!}{\left(\frac{i-k}{2}\right)!\left(\frac{i+k}{2}\right)!}, \tag{5.11}$$

with

$$V_{signal}(t) = A[\cos(2\pi f_1 t) + \cos(2\pi f_2 t)] = 2A\cos(2\pi f_0 t)\cos(2\pi f_m t), \tag{5.12}$$

where A is the amplitude of the two tone input signal V_{signal} and the c_{LS-k} terms are the coefficients of the large signal power series representation of the output characteristic.

The choice of class A for the calculation of linearity is motivated by the fact that this class presents the least third order intermodulation distortion at medium-low input signal amplitudes. In this chapter, the linearity values used in the determination of the UFM are obtained by considering an input signal amplitude of 1Vpp. This value is found to be appropriate for Si Power RF MOSFETS biased in class A, permitting the two tone signal to remain (unclipped) in the flat transconductance region.

5.3 The Shielded VDMOSFET

This section describes the effect of the shield plate on the behaviour of the device. The Unified Figure of Merit presented in section 5.2 is used for the optimisation of the shielded device. Finally, RF performance assessment for the optimised shielded VDMOSFET design is carried out.

5.3.1 Device Description

An unshielded 28V VDMOS silicon device for RF Power applications has been considered as a starting point for this analysis. The gate oxide thickness, gate length and channel length values are 70nm, 1μm and 0.9μm respectively. The cell pitch CP, defined here as the minimum distance between adjacent p-well regions, is 9μm. The n-drift doping concentration is $1.25 \cdot 10^{15}$atoms/cm^3. The device displays a breakdown voltage of 85V and a threshold voltage V_{th} of 3.4V. The threshold voltage value has been extrapolated at the constant application voltage V_{DS}=28V as the offset voltage of the line tangent to the I_{DS}-V_{GS} curve where the slope of the curve peaks.

Two main shield plate topologies, shown in Figure 5.2, are described:

(i) The "MOS" shield, with the plate positioned over the drift, in its two versions:
   a. The "Drift" shield, laid continuously over the drift region
   b. The "Split" shield, cut in correspondence of the drift region centre

(ii) The "Overlap" shield, with the plate positioned over the gate electrode.

The shield plate is always shorted to the source electrode.

In the MOS shield cases, the shield plate position in the device is described by the parameters T_{OX}, $D_{Gate,Shield}$ and L_{Shield} in Figure 5.2, where T_{OX} represents the oxide thickness of the MOS structure constituted by the shield plate over the drift region and $D_{Gate,Shield}$ is the distance between gate and shield electrodes. In the Overlap shield, the shield plate position is described by the parameter T_{Over} which represent the oxide thickness between gate and Overlap shield.

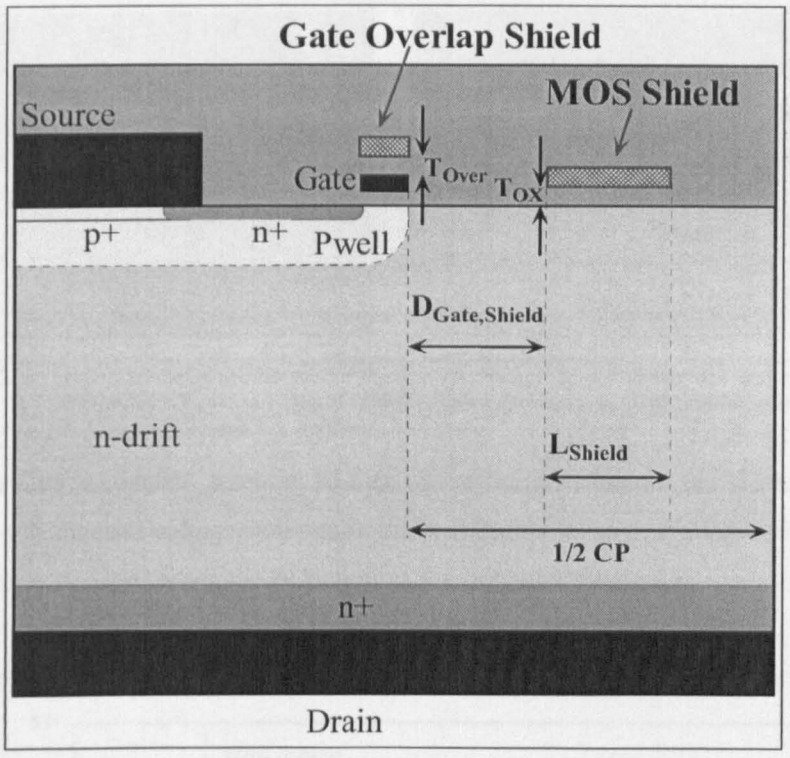


Figure 5.2: Considered shield plate topologies.

5.3.2 Breakdown voltage, potential and electric field distributions

This section reports a brief overview of the effect of the shield plate on the modification of breakdown voltage, potential and electric field distribution in RF power VDMOSFETs. Field plates are extensively used in power device design; their effect on the electric field distribution and on the breakdown voltage is well known and understood [1][2][3]. In a VDMOSFET equipped with a MOS shield, the shield plate leads to a more uniform distribution of potential and electric field, determining a significant reduction of the electric field and an improvement in breakdown voltage capability [9][10] with respect to an unshielded VDMOSFET. The effect of the shield plate on the potential distribution is shown in Figure 5.3.

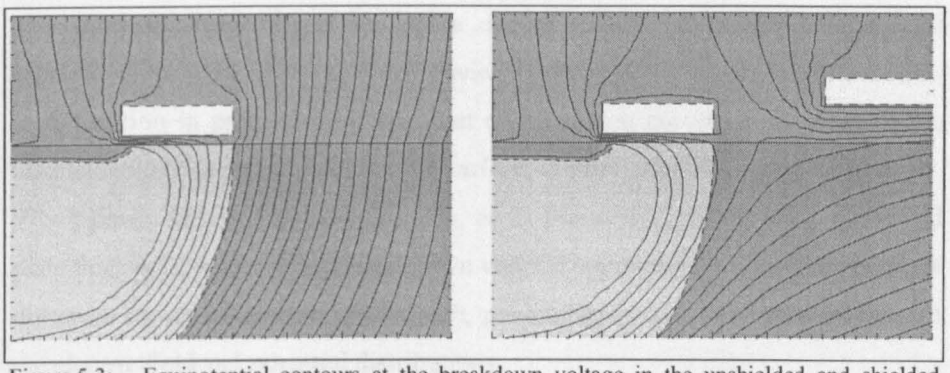


Figure 5.3: Equipotential contours at the breakdown voltage in the unshielded and shielded VDMOSFET cases. The shielded device presents a drift shield plate with $D_{Gate,Shield}$=1μm and T_{OX} =100nm.

By creating a constant potential surface above the drift region, the shield plate forces the constant voltage contours in the drift region to bend, closing within the device itself. This causes the reduction of the potential gradient in the proximity of the channel, yielding the peak electric field reduction with respect to the unshielded structure acknowledged in Figure 5.4.

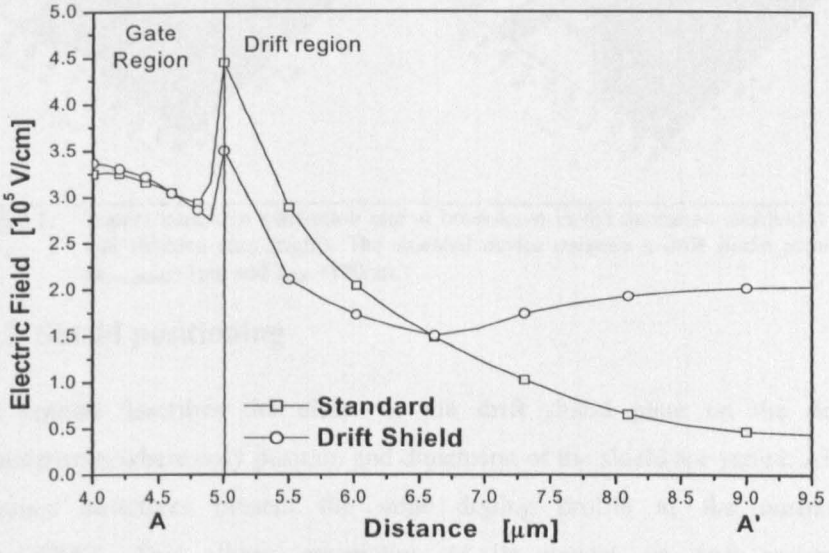


Figure 5.4 Maximum electric field at breakdown voltage in the unshielded and shielded VDMOSFET cases. The shielded device presents a drift shield plate with $D_{Gate,Shield}$=1μm and T_{OX} =100nm. The electric field amplitude is extracted at the silicon-oxide interface of the shielded VDMOSFET, from the beginning of the channel to the centre of the drift region.

The analysis of the impact ionization rate of Figure 5.5 reveals a change of breakdown location: in the shielded device the breakdown occurs at the p-well/n-drift junction in proximity of the point of maximum curvature, whereas in the unshielded device it takes place at the surface, in proximity of the gate end.

When positioned overlapping the gate, as in the overlap shield case, the shield plate does not contribute to a breakdown voltage improvement. The gate electrode obstructs the shield to drift region path, preventing a substantial modification of the electric field and potential distributions.

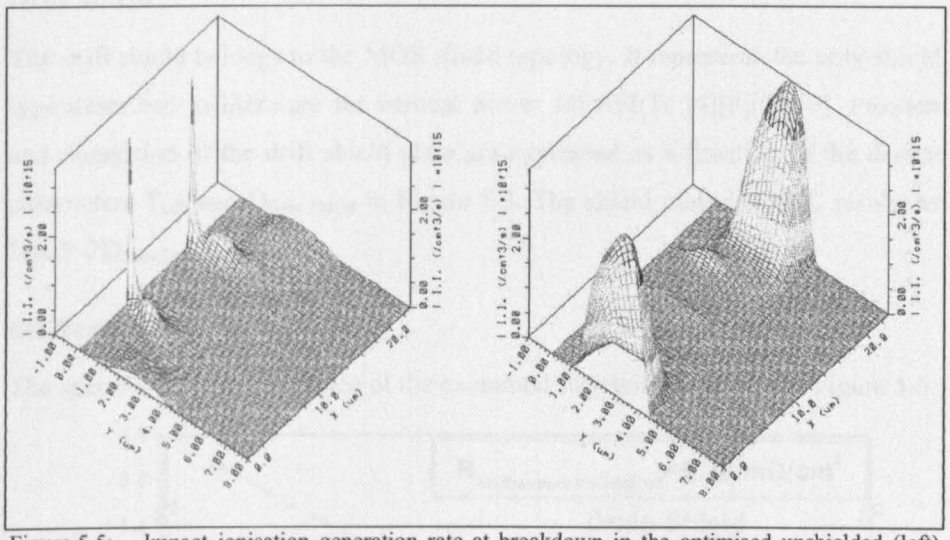


Figure 5.5: Impact ionisation generation rate at breakdown in the optimised unshielded (left) and shielded case (right). The shielded device presents a drift shield plate with $D_{Gate,Shield}$=1μm and T_{OX} =100nm.

5.3.3 Shield positioning

This section describes the effect of the drift shield plate on the device characteristics where only position and dimension of the shield are varied. All the examined structures present the same doping profile as the unshielded VDMOSFET. This allows association of its impact on drift resistance, transconductance and capacitance, power gain and linearity.

In this section it is shown that the shield plate insertion always causes a degradation of transconductance and an improvement of gate to drain capacitance. These are opposing effects in terms of RF performance. The degradation of the transconductance is expected to reduce the linearity of the device and also to

reduce power gain, since a smaller g_{max} implies reduced current capabilities. The reduction of the gate to drain capacitance is expected to enhance gain and maximum frequency. The UFM values are estimated to resolve this trade off. In the RF based calculations the device gate width is set at 29mm. This length allows for 20W of maximum linear RF output power in the case an unshielded VDMOSFET. For a fair comparison, all the device configurations examined in this section present the same breakdown voltage of 85V.

Drift Shield

The drift shield belongs to the MOS shield topology. It represents the only shield type described in literature for vertical power MOSFETs [4][8][9][10]. Position and dimension of the drift shield plate are expressed as a function of the design parameters T_{OX} and $D_{Gate,Shield}$ in Figure 5.2. The shield plate length L results as $L=CP-2D_{Gate,Shield}$.

On Resistance

The specific on-state resistance of the examined structures is plotted in Figure 5.6.

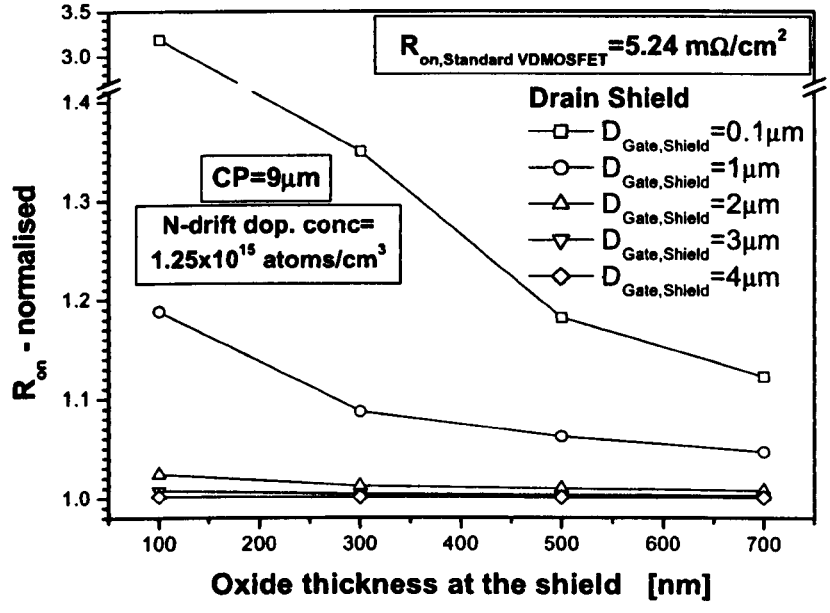


Figure 5.6: Specific on resistance for the drift shielded devices with respect to the optimised standard VDMOSFET.

Increased specific on-state resistance is observed with the shield plate with respect to the unshielded VDMOSFET. The increase is due to an augmented JFET resistance associated with the cell spacing. With respect to the unshielded device, the channel of the JFET constituted by the p-wells of adjacent cells and the n-drift is reduced by the depletion region formed under the grounded shield plate. Figure 5.6 shows that the reduction of T_{OX} and $D_{Gate,Shield}$ determines the increase of the on resistance. This follows from the proximity and the extension of the drift shield over the drift region, which causes the expansion of the depletion region of the shield MOS structure.

Transconductance

Figure 5.7 shows the peak transconductance values of the examined devices, where degradation of the transconductance is observed when the drift shield plate is used; the values are extracted in the saturation region at the operating voltage of V_{DS}=28V.

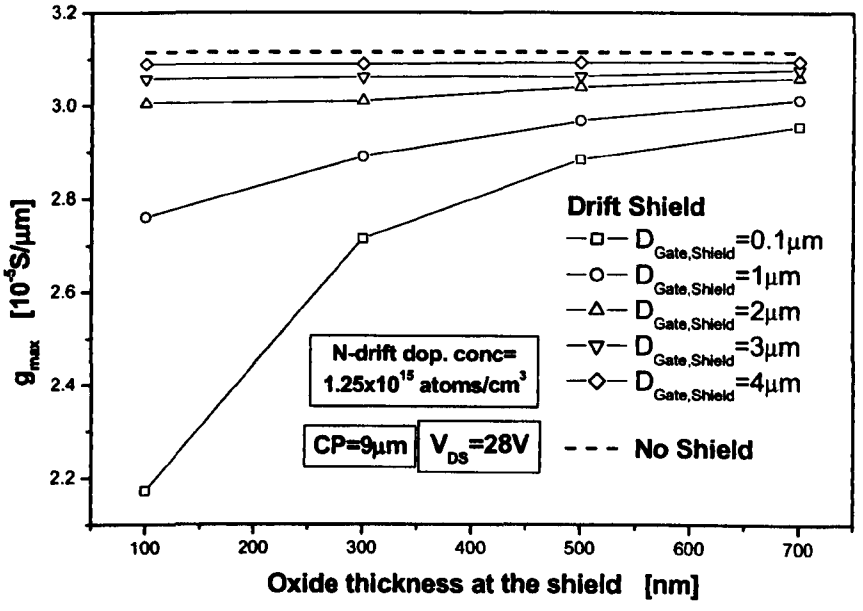


Figure 5.7: Effect of the drift shield plate on the peak transconductance at V_{DS}=28V.

The degradation of the transconductance increases with the reduction of T_{OX} and $D_{Gate,Shield}$. In both cases the shield causes an increase of the drift region resistance.

For a certain V_{DS} this yields a reduction of the potential drop across the channel with respect to the optimised standard VDMOSFET, as shown in Figure 5.8.

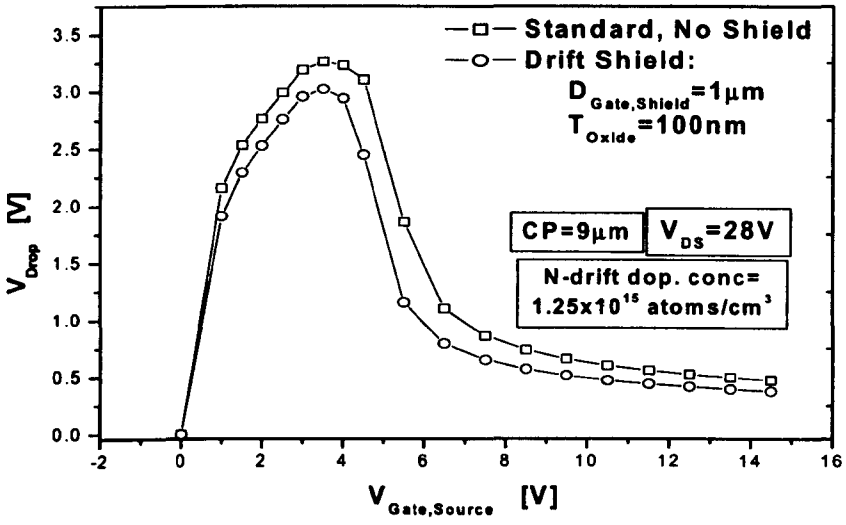


Figure 5.8: Effect of the drift shield plate on the voltage drop across the MOSFET channel.

Furthermore when the shield plate is used, the reduced voltage drop across the channel with respect to the optimised standard VDMOSFET yields an earlier fall-off of the transconductance, as shown in Figure 5.9. As described in [7], the intrinsic MOSFET moves back into the linear region for lower gate to source voltages.

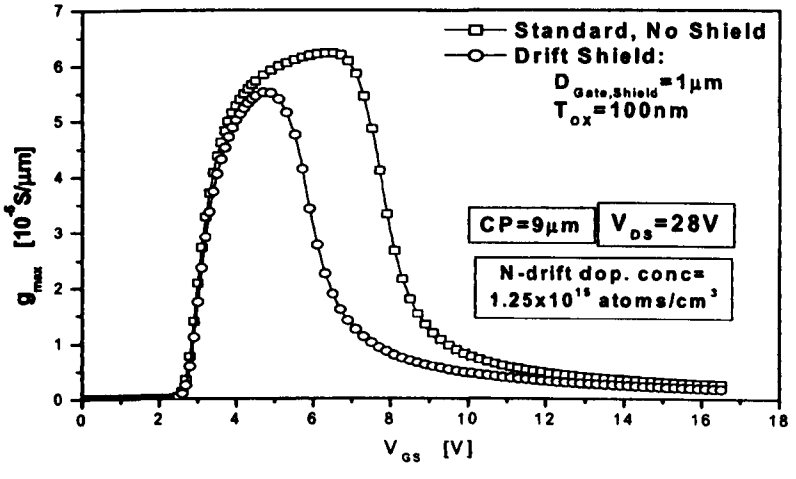


Figure 5.9: Transconductance versus gate to source voltage for the unshielded and shielded VDMOSFET.

With the shield plate, the peak transconductance is lowered with respect to the unshielded VDMOSFET. As explained in Chapter 3, this behaviour is related to the earlier fall-off point V^*_{GS} and higher drift resistance caused by the insertion of the shield.

This results in

$$g_m \cong g_{m.ideal} R_F, \tag{5.13}$$

$$g_{m,ideal} = WC_{OX} v_{sat}, \tag{5.14}$$

$$R_F = (1 - \frac{1}{2K}) \frac{V^*_{GS} - \max(V_{TH})}{V^*_{GS} - V_{th}}, \tag{5.15}$$

where W is the device width, C_{OX} is the gate oxide capacitance per unit area, v_{sat} is the carrier saturation velocity, K is a constant larger than unity, $\max(V_{TH})$ and V_{th} represent the maximum threshold voltage displayed by the device and the threshold voltage at the application voltage respectively, $g_{m,ideal}$ is the ideal transconductance limit and R_F is a reduction factor. Equation 5.15 relates V^*_{GS} with the peak transconductance value: larger drift region resistances give smaller values of V_{GS}^* which, according to Equation 5.15, are associated with smaller values of the peak transconductance.

Capacitance

The capacitors defined in the devices examined are represented in Figure 5.10. By considering that the shield plate is shorted to the source electrode and by solving the equivalent capacitance circuit associated with the capacitors reported in Figure 5.11, expressions for the gate to drain, gate to source and drain to source capacitance are derived. The expressions are used to clarify the effect of the design parameters on the capacitance content.

The gate to drain capacitance results in:

$$C_{GD} \approx (\frac{1}{C_{Gmos}} + \frac{1}{C_{depl}})^{-1} + (\frac{1}{C_{SHmos}} + \frac{1}{C_{GSHox}})^{-1} + C_{GDov}, \tag{5.16}$$

$$C_{depl} = C_{depl-1} + C_{depl-2}, \tag{5.17}$$

where C_{Gmos} is the MOS capacitor defined by the gate electrode over the channel region, C_{depl} is the capacitance associated with the depletion region of the p-

well/n-drift junction and C_{GDov} is the parasitic MOS capacitor defined by the gate electrode over the n-drift region. C_{depl-1} in C_{depl} is distinguished from C_{depl-2}. C_{depl-1} is associated with the curvature of the p-well/n-drift junction and is strongly affected by the cell spacing and the potential redistribution caused by the shield insertion, unlike C_{depl-2}.

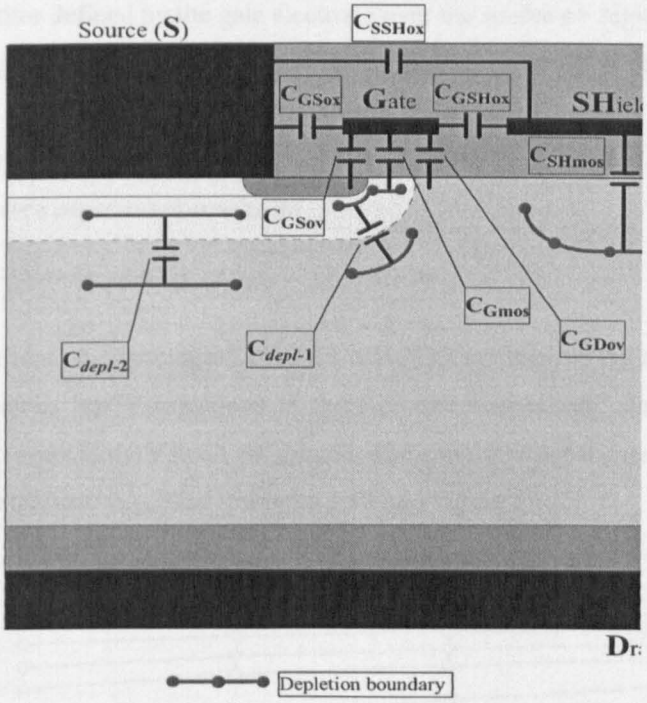


Figure 5.10: Capacitors intrinsically defined in the examined shielded structures.

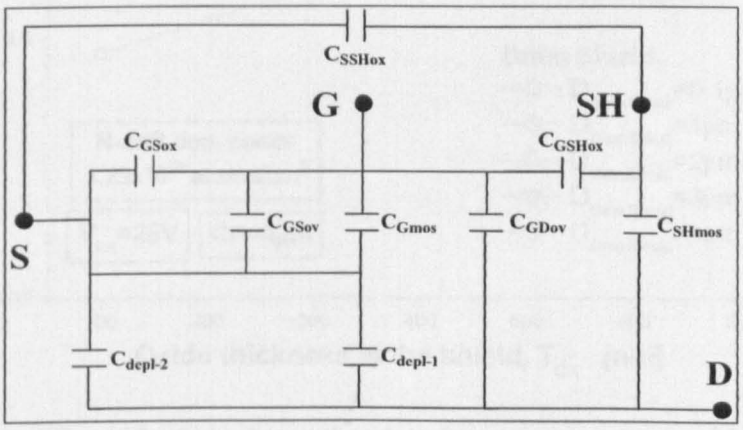


Figure 5.11: Equivalent capacitance circuit for the shielded devices.

The gate to source capacitance results in:

$$C_{GS} \approx C_{GSox} + C_{GSov} + C_{Gmos} + C_{GSHox} + C_1, \tag{5.18}$$

$$C_1 = \left(\frac{1}{C_{SHmos}} + \frac{1}{C_{GDov}}\right)^{-1}, \tag{5.19}$$

where C_{GSox} is the oxide capacitance between the gate and the source and C_{GSov} is the MOS capacitor defined by the gate electrode over the source n+ region. C_{GSHox} is the oxide capacitance between gate and shield, C_{SHmos} is the MOS capacitance defined by the shield electrode over the n-drift region and C_{SSHox} is the oxide capacitance between source and shield.

The drain to source capacitance results in:

$$C_{DS} \approx C_{depl} + \left(\frac{1}{C_{S\,SHox}} + \frac{1}{C_{SHmos}}\right)^{-1}, \tag{5.20}$$

The approximations in Equations 5.16, 5.18 and 5.20 are derived by neglecting the effects of series type connections of three or more capacitors, which cause negligible effects due to their small magnitude. The simulated capacitance content of the devices examined is plotted in Figure 5.12 and Figure 5.13.

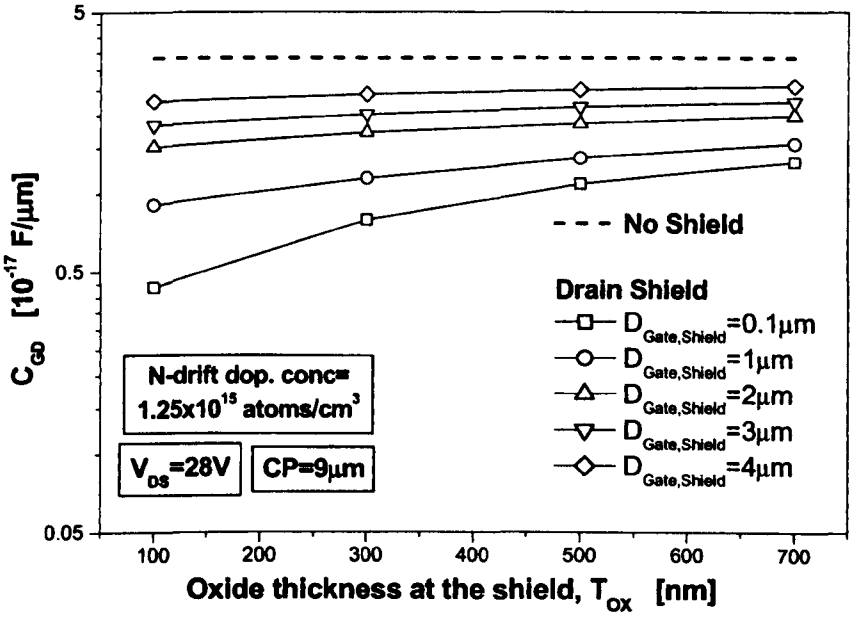


Figure 5.12: Effect of the shield plate on the gate to drain capacitance for unmodified doping profiles.

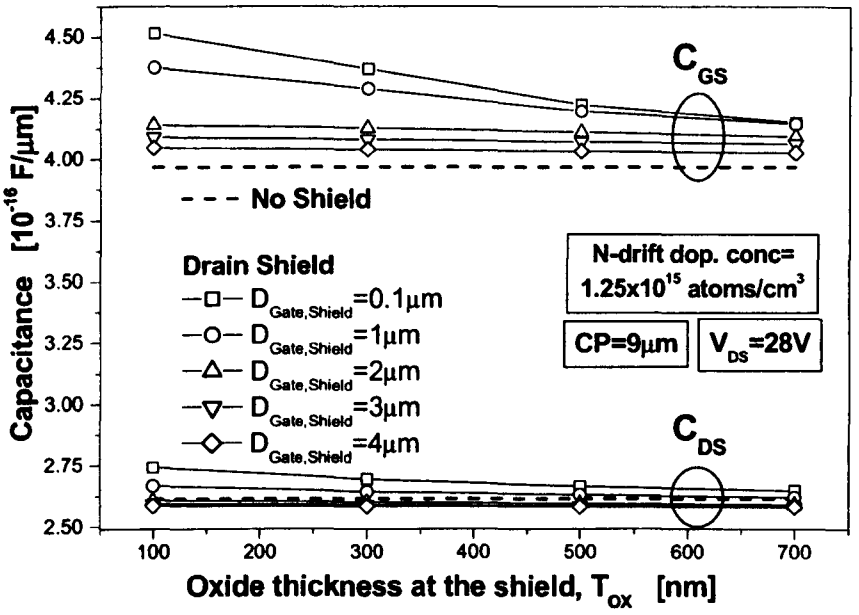


Figure 5.13: Effect of the shield plate on the gate to source and drain to source capacitance for unmodified doping profiles.

The reduction of the electric field caused by the shield plate allows for an increased depletion of the reverse biased junctions, resulting in the decrease of $C_{depl\text{-}1}$. Consequently, C_{GD} is reduced in accordance with Equations 5.16 and 5.17, as shown in Figure 5.13. The effect of the shield increases for small $D_{Gate,Shield}$ and T_{Oxide} and determines the associated reduction of gate to drain capacitance reported in Figure 5.12. The second term in Equation 5.16 has a lower order effect on C_{GD} which derives from the small values of C_{GSHox}. The shield plate, being shorted to the source electrode, increases the gate to source capacitance due to the contribution of C_{GSHox} and C_{SHmos}, in accordance with Equations 5.18 and 5.19. C_{GSHox} and C_{SHmos} increase with the reduction of $D_{Gate,Shield}$ and T_{Oxide} respectively, leading to the associated C_{GS} increase of Figure 5.13. According to Equations 5.18 and 5.19, C_{GSHox} produces a direct increase of C_{GS}, while C_{SHmos} increases C_1 and has a secondary effect on C_{GS}. As acknowledged in Equations 5.20, the drain to source capacitance is strongly affected by the value of C_{SSHox}. For small C_{SSHox} corresponding to large values of $D_{Gate,Shield}$, C_{DS} is approximately C_{depl}; since the shield insertion reduces the p-well/n-drift depletion capacitance C_{depl}, C_{DS} can

become smaller than in the unshielded case. For large C_{SSHox} the drain to source capacitance increases in spite of the reduction of C_{depl}.

It has to be noted that the shield plate also causes a reduction of charge on the gate electrode with respect to the unshielded VDMOSFET case. In the MOS shield case, the reduction of gate charge is proportional to the effect of the shield on the drift region. Large influence of the shield on the drift region is associated with large values of the C_{SHmos} capacitance and therefore to large amount of charge on the shield. The shield affects the gate charge through C_{GSHox}. An accurate description of this mechanism will be described later for the overlap shield topology, where the gate charge reduction is the only cause of the reduction of C_{GD}. Such a description has not been provided in the MOS shield case, where both charge variations in the electrodes and internally to the device take place.

In order to quantify the influence of charge variations on capacitance, a more complex device model than the one reported in Figure 5.11 should be used. The model should take into account the resistive paths in the device and non-linear capacitance effects.

RF based analysis

The calculated power gain and third order intermodulation distortion, determined as described in section 5.2, are reported in Figure 5.14 and Figure 5.15. The Unified Figure of Merit values are reported in Figure 5.16.

Figure 5.14 shows a power gain improvement with the shield plate which is associated with the reduction of C_{GD} with respect to the optimised standard VDMOSFET. Figure 5.15 shows the degradation of linearity caused by the negative effect of the shield on the transconductance characteristic.

Examining the calculated UFM values of Figure 5.16, an optimum shield can be identified in the device with a gate to shield distance of 2μm and an oxide thickness of 100nm. For such a choice of the design parameters the degradation of C_{GS} and g_m is within 5%, while C_{GD} is about 45% of the unshielded case. At the optimum, the UFM reaches a value of 951, much higher of the 816 in the unshielded VDMOSFET case.

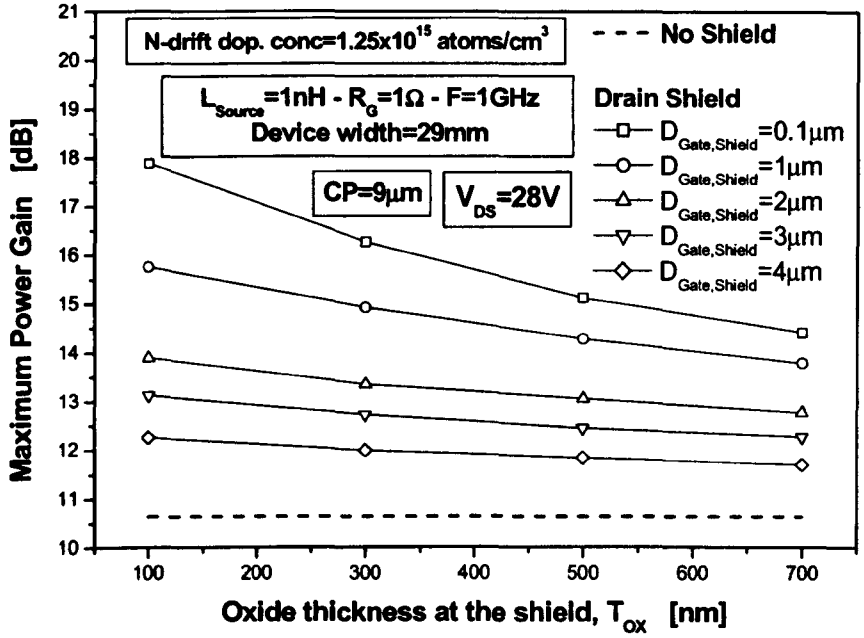


Figure 5.14: Effect of the drift shield plate insertion on the class A power gain Gmax for unmodified doping profiles.

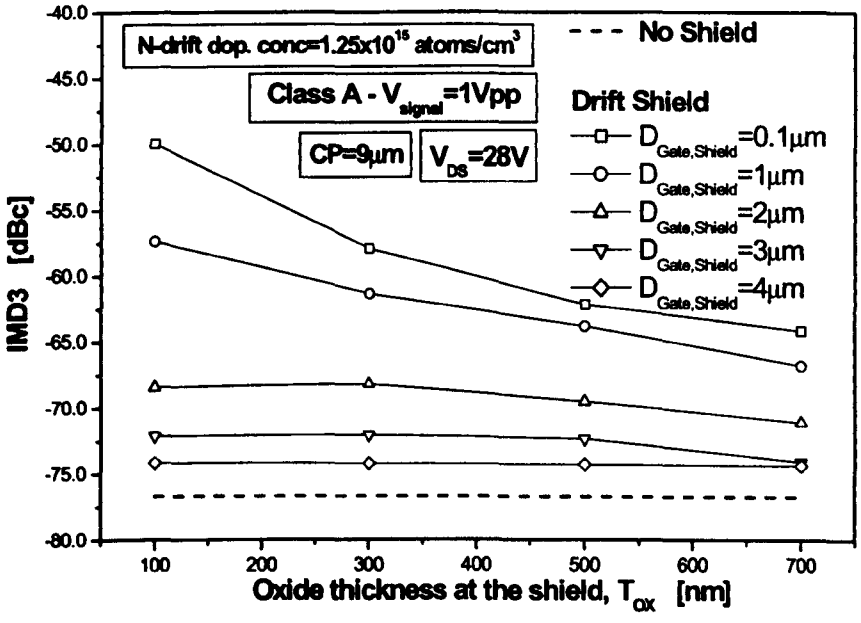


Figure 5.15: Effect of the drift shield plate insertion on the optimum class AB IMD3[dBc] for unmodified doping profiles.

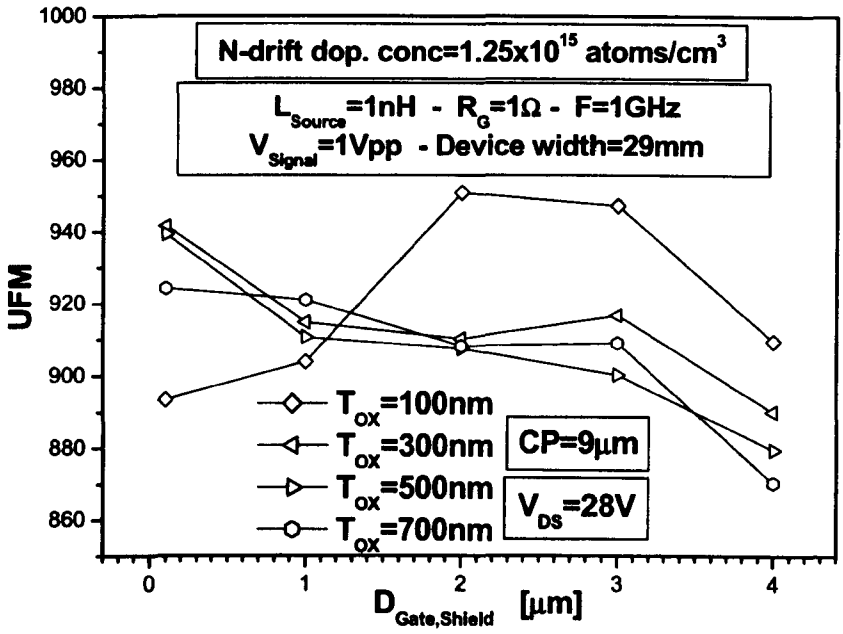


Figure 5.16: Effect of the drift shield plate insertion on the Unified Figure of Merit for unmodified doping profiles.

Split Shield

The split shield belongs to the MOS shield topology. The split shield effect is examined to establish if any RF improvement can be achieved with respect to the drift shield configuration by shortening the shield-plate/drift-region overlap. Its position in the device can be described by the parameters T_{OX}, $D_{Gate,Shield}$ and L_{Shield} of Figure 5.2, where T_{OX} represents the oxide thickness of the MOS structure constituted by the shield, $D_{Gate,Shield}$ is the distance between gate and shield electrodes and L_{Shield} is the length of the shield plate.

Being a MOS shield, it presents the same overall trend as the drift shield. The only difference with the drift shield is due to the reduced length of the shield plate. For constant $D_{Gate,Shield}$ and T_{OX} values, smaller L_{Shield} causes a reduced effect on the drift region leading to a less degraded g_m but also reducing the improvements on C_{GD}.

Figure 5.17 to Figure 5.24 show the effect on the main device characteristics extracted from simulation results and on the calculated maximum power gain,

linearity and UFM values of the split shield design parameters T_{OX}, $D_{Gate,Shield}$ and L_{Shield}.

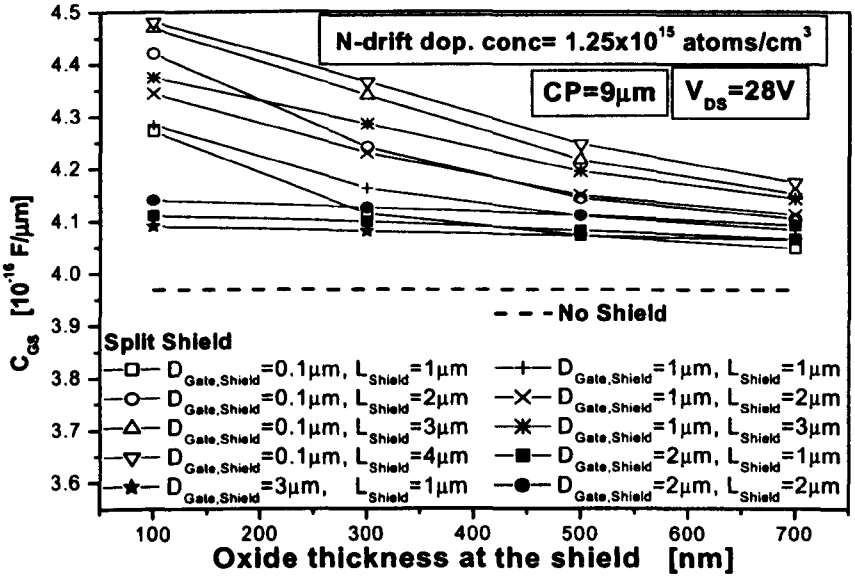


Figure 5.17: Effect of the split shield plate on the gate to source capacitance for unmodified doping profiles.

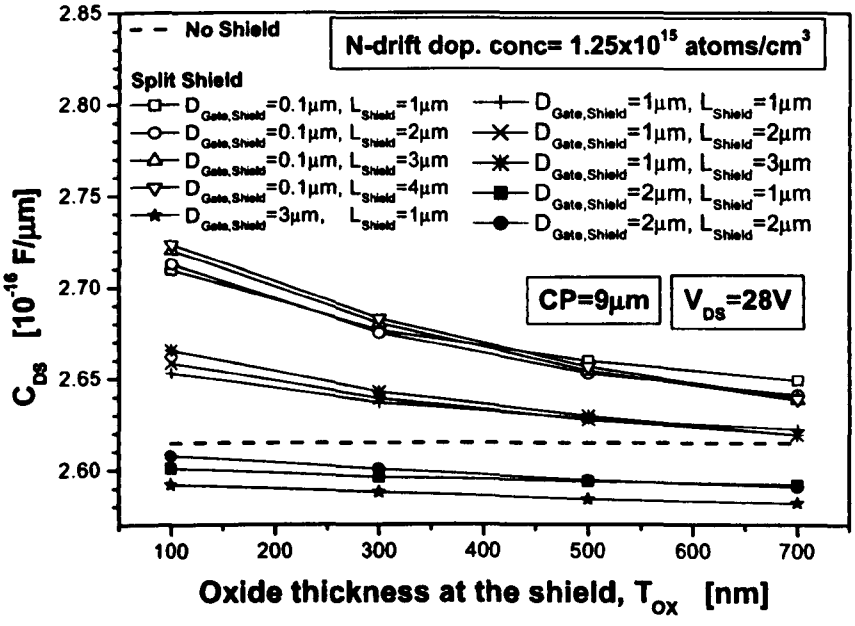


Figure 5.18: Effect of the shield plate on drain to source capacitance for unmodified doping profiles.

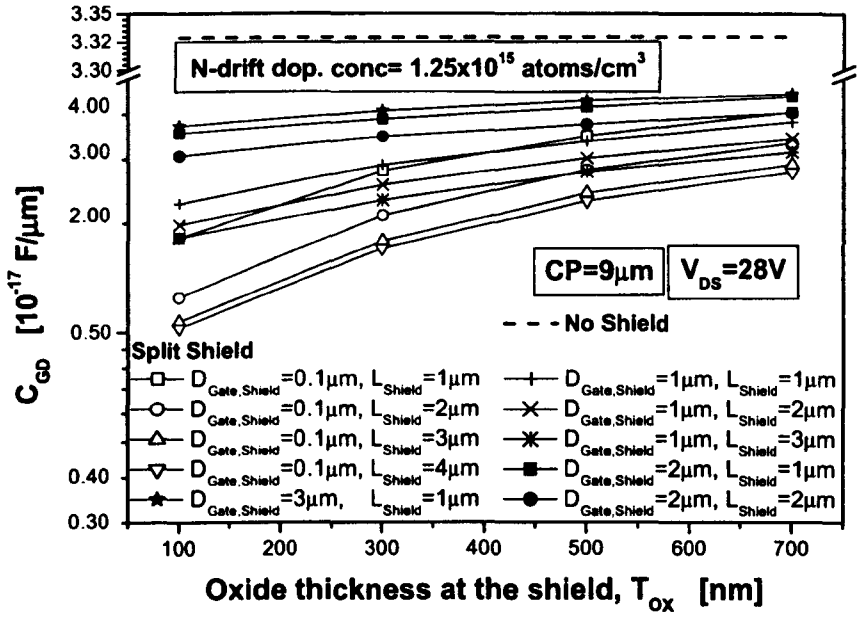


Figure 5.19: Effect of the shield plate on the gate to drain capacitance for unmodified doping profiles.

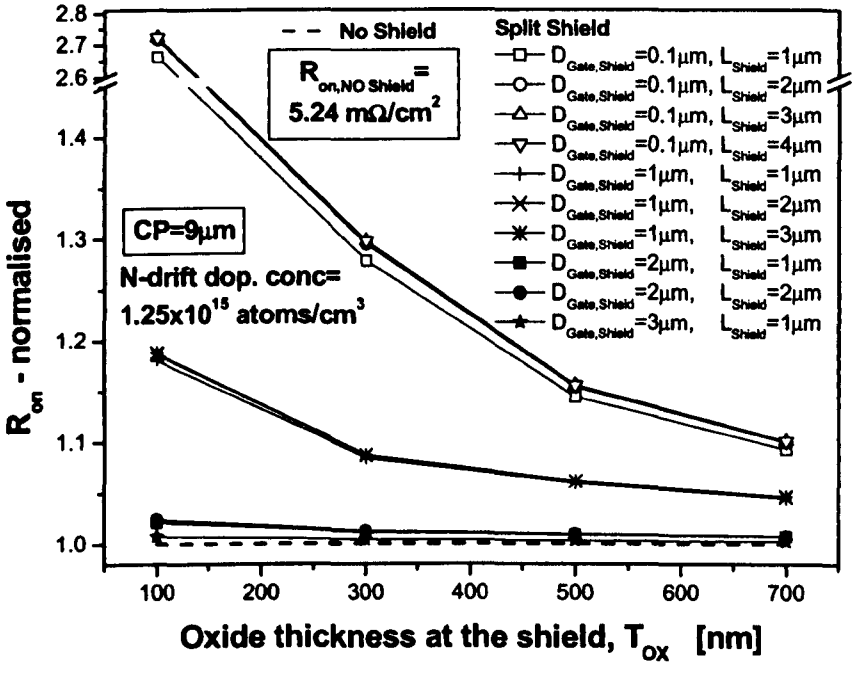


Figure 5.20: Specific on resistance for the split shielded devices with respect to the optimised standard VDMOSFET.

With respect to the drift shield, for constant $D_{Gate,Shield}$ and T_{OX} values, smaller L_{Shield} causes a reduced effect on the drift region leading to a less degraded g_m but also reducing the improvements in C_{GD}. The g_m improvement is due to a smaller JFET resistance associated with the reduced drift region depletion width caused by the split shield. The reduction of the depletion width causes the increase of C_{depl-1}, yielding higher C_{GD} values. The L_{Shield} reduction also produces a reduced C_{SHmos}, thus improving C_{GS} and C_{DS} according to Equation 5.18 - 5.19 and Equation 5.20 respectively.

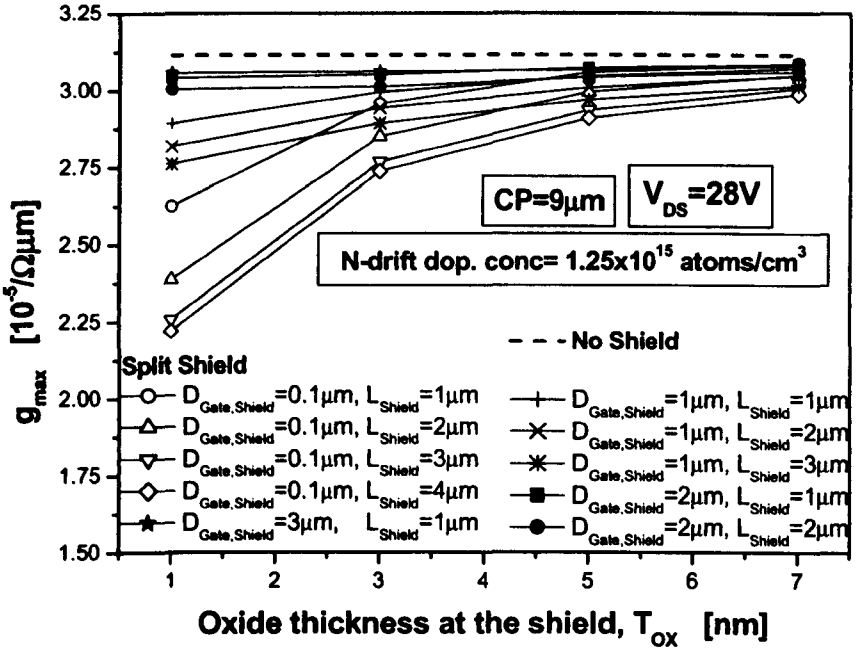


Figure 5.21: Effect of the split shield plate on the peak transconductance at VDS=28V.

The calculated third order inter-modulation distortion and power gain values for the considered split shield configurations are reported in Figure 5.23 and Figure 5.22.

As in the drift shield case, the degradation of g_m caused by the split shield insertion worsens IMD3. The improvement of C_{GD} yields an increased Power Gain.

The evaluation of the UFM values reported in Figure 5.24 shows that the optimum split shield geometry corresponds to $D_{Gate,Source}$=0.1μm, T_{OX}=300nm and L_{Shield}=2μm.

With respect to the drift shield case with $D_{Gate,Source}$=2μm and T_{OX}=100nm, the reduced split shield length permits a less degraded transconductance together with a low value of the gate to drain capacitance. The resulting UFM is about 949, slightly smaller than the value of 951 in the optimum drift shield case. This result is achieved with an improved power gain but with degradation of IMD3 with respect to the drift shield case, suggesting that the split shield is more appropriate than the drift one for applications at which high frequency is more a requirement than linearity.

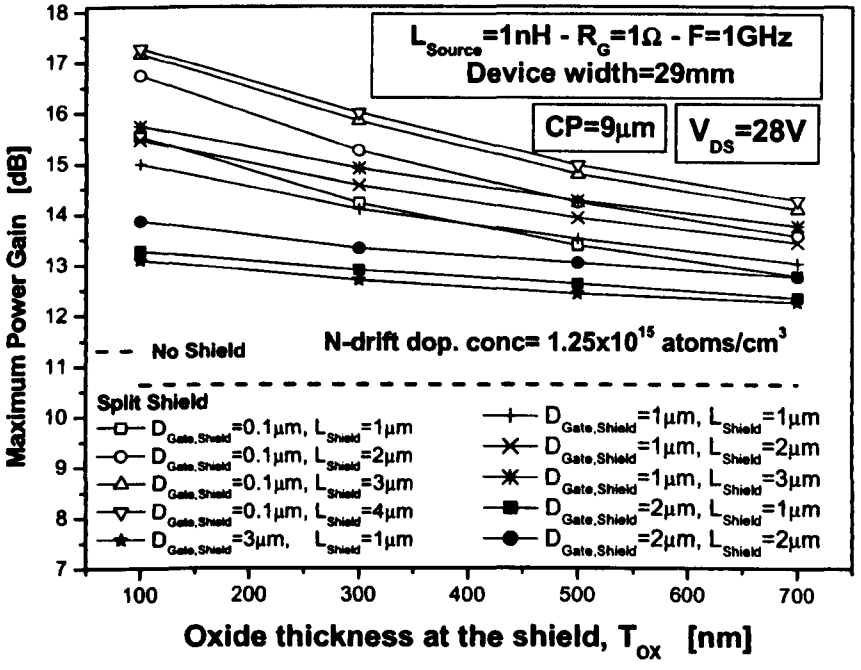


Figure 5.22: Effect of the split shield plate insertion on the class A maximum power gain Gmax for unmodified doping profiles.

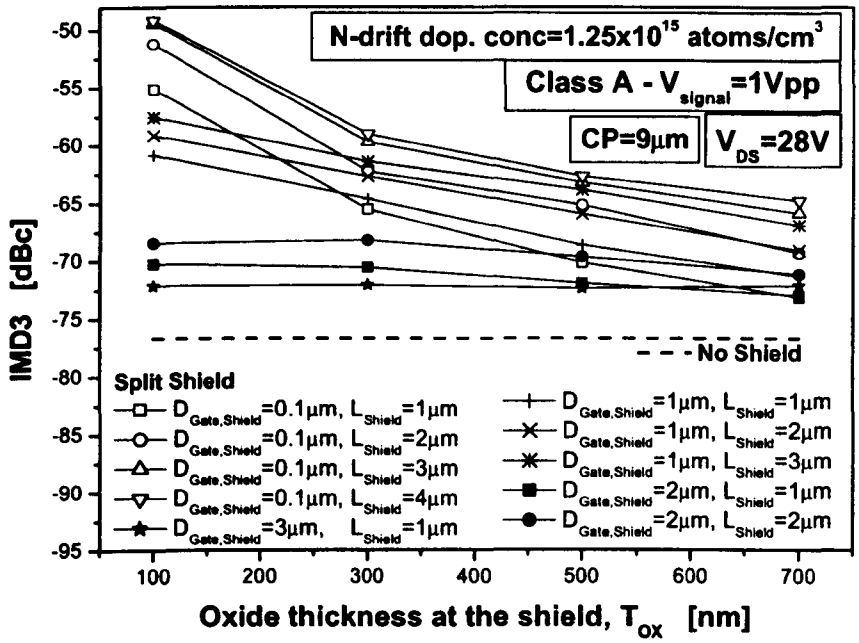


Figure 5.23: Effect of the drift shield plate insertion on the optimum class AB IMD3[dBc] for unmodified doping profiles.

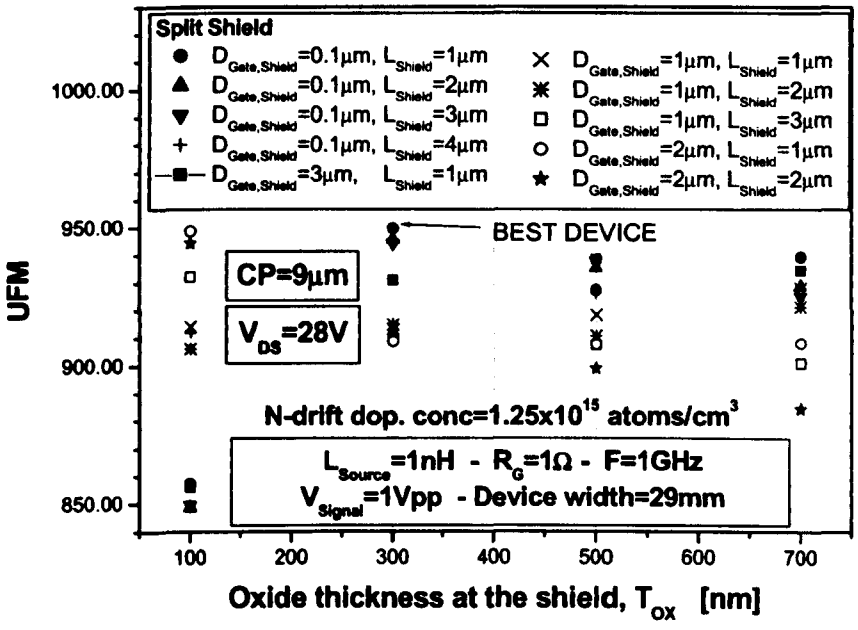


Figure 5.24: Effect of the split shield plate insertion on the Unified Figure of Merit for unmodified doping profiles. The calculated values are normalised to the optimised standard VDMOSFET.

Gate Overlap shield

The position of the overlap shield can be described by the parameter T_{Over} of Figure 5.2, where T_{Over} is the oxide thickness separating gate and shield electrodes. The overlap shield is placed over the gate electrode and presents the same length as the gate electrode.

This shield plate configuration yields very small effects on the drift region because of the interposition of the gate electrode between shield and drift region. Nevertheless, a gate to drain capacitance reduction is observed. This section examines the mechanism causing the capacitance reduction and addresses the effects of the gate overlap shield on the transconductance.

Capacitance

In the gate overlap shield case, the modification of the electric field distribution in the drift region is quite small. Figure 5.25 reports drain and gate to source capacitances slightly increased with respect to the unshielded VDMOSFET.

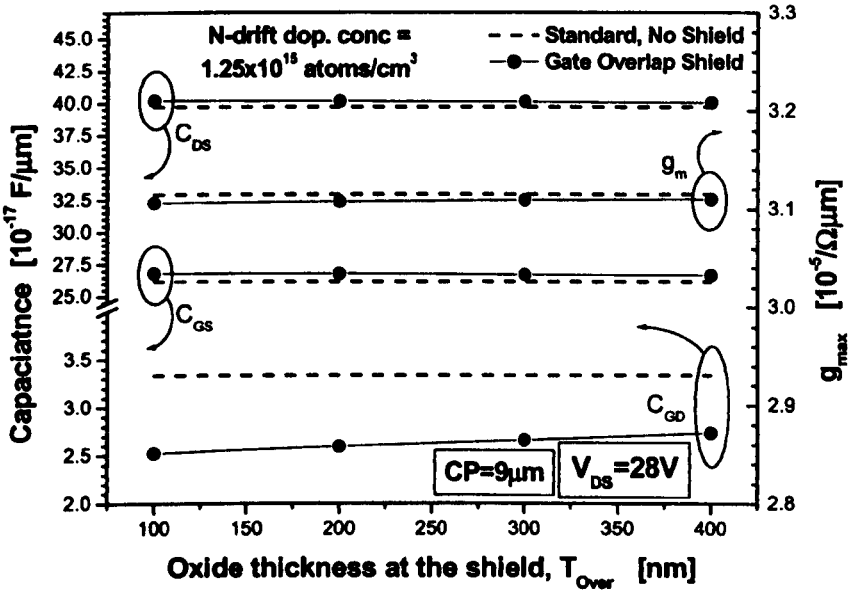


Figure 5.25: Effect of the overlap shield plate insertion on capacitance and g_m for unmodified doping profiles.

Analogous to the MOS shield case, the equivalent capacitance model in Figure 5.26 can be used to clarify the effect of the design parameter T_{over} on the device performance. With respect to unshielded case, the increases of C_{GS} and C_{DS} in Figure 5.25 are due to the added capacitance contributions C_{GSH} and C_{SHmos} respectively. In Figure 5.25 a large reduction of C_{GD} is revealed. This reduction does not derive from a shield effect on the depletion of the p-well/n-drift junction. Instead, C_{GD} is reduced because of the decrease of the total charge on the gate electrode, as shown in Figure 5.27. This is a consequence of a strong capacitive coupling between gate and shield electrodes. The increase of C_{GD} with the increase of gate to shield oxide thickness T_{Over} reported in Figure 5.25 derives from a reduction of such capacitive coupling.

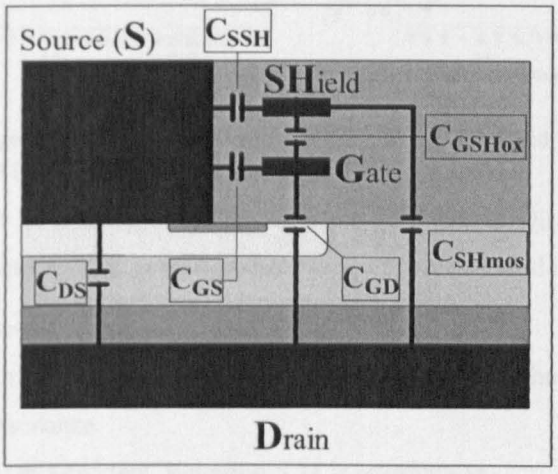


Figure 5.26: Simplified capacitance model in the shield overlap case.

The model in Figure 5.26 has been used to explain the gate charge reduction. The charge on the shield electrode (Q_{SH}) increases for increasing drain to source voltage, as shown in Figure 5.27. This increase depends on the existence of a small (MOS) capacitance between shield and drain electrode, C_{SHmos} of Figure 5.26. The gate to shield oxide capacitance C_{GSHox} couples part of the Q_{SH} charge on the gate electrode, reducing the charge on the gate $Q_{G,Overlap}$. The amount of charge coupled is Q_{GSHox}, the charge on C_{GSHox}.

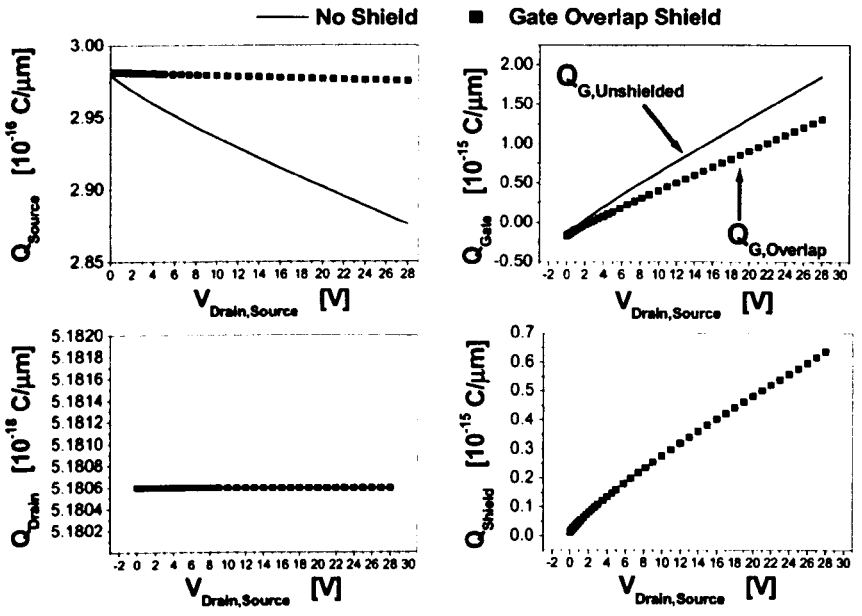


Figure 5.27: Charge on the electrodes of the unshielded VDMOSFET and with a gate overlap shield presenting oxide thickness T_{Over}=100nm.

Therefore, when an overlap shield is used, the gate charge $Q_{G,Overlap}$ is reduced with respect to the $Q_{G,Unshielded}$ charge observed in the unshielded device case:

$$Q_{G,Overlap} = Q_{G,Unshielded} - Q_{GSHox} \cdot \tag{5.21}$$

The gate charge reduction described by Equation 5.21 causes the reduction of the gate to drain capacitance.

In order to verify this concept, Equation 5.21 is rewritten as

$$\Delta Q = Q_{G,Unshielded} - Q_{G,Overlap} = Q_{GSHox} \cdot \tag{5.22}$$

According to Equation 5.22, the validity of the charge coupling mechanism is proved if ΔQ can be shown to match Q_{GSHox}, where ΔQ is the variation of charge on the gate electrode caused by the overlap shield plate. This is demonstrated in Figure 5.28, where the error in estimating the gate charge variation ΔQ shows a mean value of just 3%. In Figure 5.28 ΔQ is calculated comparing the simulation results in the shielded and unshielded case. Q_{GSHox} is instead calculated using just the simulation results of the shielded device, through the resolution of the capacitance model of Figure 5.26. The capacitance C_{SHmos} is calculated first, yielding the determination of Q_{GSHox}, as described in Equations 5.24 to 5.28.

To further verify the charge coupling mechanism, Figure 5.28 also shows that the simulated C_{GD} is well matched by the gate to drain capacitance calculated from charge reduction:

$$C_{GD,calculated} = \frac{\partial(Q_{G,Unshielded} - Q_{GSHox})}{\partial V_{DS}}. \tag{5.23}$$

The shield to drain MOS capacitance C_{SHmos} is calculated from the total drain to shield capacitance C_{DSH} as

$$C_{SHmos} = C_{DSH} - (1/C_{GD} + 1/C_{GSHox})^{-1}, \tag{5.24}$$

$$C_{GSHox} = \frac{\varepsilon WL}{T_{Over}}, \tag{5.25}$$

where the gate to shield oxide capacitance C_{GSHox} is calculated with the classic plane capacitor formula.

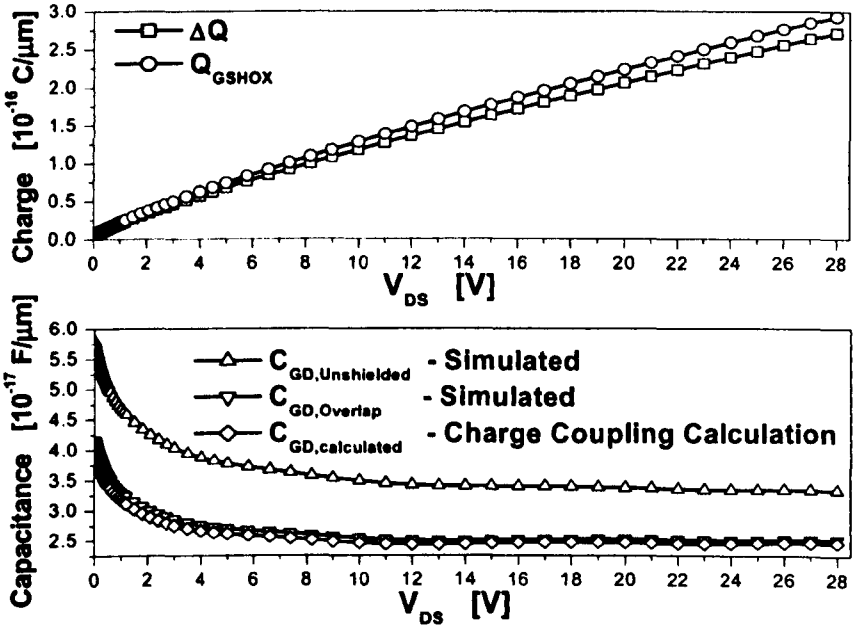


Figure 5.28: Validation of the model assumption by calculation of charge and capacitance.

Q_{GSHox} is determined from Q_{SH} considering all the parallel capacitance contributions between shield and gate and applying a charge partitioning calculation here:

$$Q_{GSHox} = \frac{Q_{SH}}{1+K}, \tag{5.26}$$

$$K = \frac{C_2}{C_{GSHox}}, \tag{5.27}$$

$$C_2 \approx \left(\frac{1}{C_{GD}} + \frac{1}{C_{SHmos}} \right)^{-1}. \tag{5.28}$$

The approximation in Equation 5.28 derives by neglecting the effects of series type connections of three or more capacitors, which cause negligible effects due to their small magnitude.

Transconductance

Figure 5.25 reports the peak g_m value to reduce with the increase of gate to shield oxide thickness T_{Over}. The gate overlap shield plate insertion determines a small reduction of the peak transconductance value and a slightly earlier fall-off with respect to the standard VDMOSFET. This behaviour is a main consequence of the decrease of carrier concentration in the drift region, in correspondence with the end of the channel, as shown in Figure 5.29. Since the reduction takes place along the current path, the drift-resistance is increased, causing the earlier transconductance fall-off. The associated reduction V_{Drop} affects the g_m peak value through the reduction of the transconductance fall-off point V^*_{GS}, analogously to the drift shield.

RF based analysis

As illustrated in Figure 5.25, with respect to the standard-unshielded VDMOSFET, the use of the gate overlap shield plate permits a C_{GD} reduction of about 25% at a limited cost for g_m. The maximum effect of the overlap shield is achieved for observed for the minimum considered oxide thickness T_{Over}=100nm, as reported in Figure 5.30. The maximum UFM value of 895 is observed for the minimum considered oxide thickness T_{Over}=100nm.

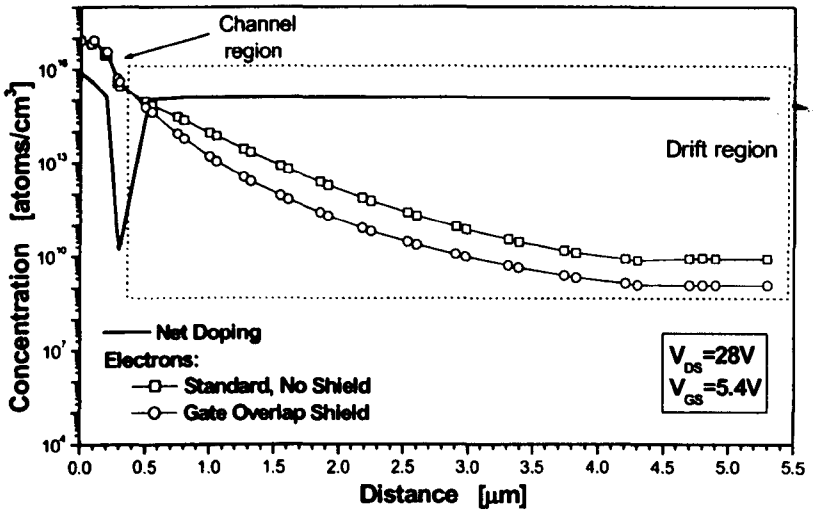


Figure 5.29: Overlap shield effect on carrier concentration at $V_{GS}=V_{th}+2V$ and $V_{DS}=28V$ in the top of the device. The shielded devices present $D_{Gate,Shield}=2\mu m$, $T_{Over}=100nm$. Doping profiles are unchanged for the two devices.

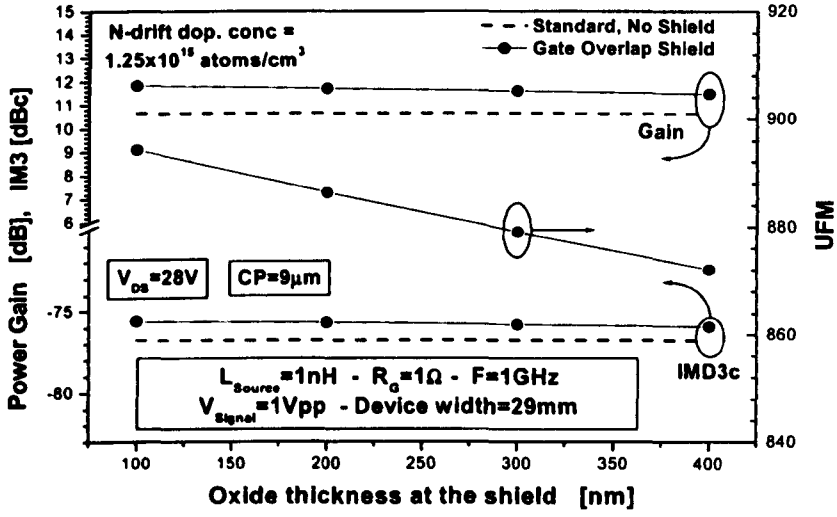


Figure 5.30: Effect of the overlap shield plate insertion on linearity, power gain and UFM for unmodified doping profiles.

5.3.4 Optimum shield plate design

In this section the effects of the variations of the design parameters on the performance of the shielded RF power VDMOSFET are analysed, leading to the identification of an optimum shield plate design. The optimum drift shield plate

configuration is used as a starting point for this analysis, since it displays the maximum observed UFM value.

For fair RF performance comparisons, unless otherwise stated, all the device configurations examined in this section present the same breakdown voltage of 85V.

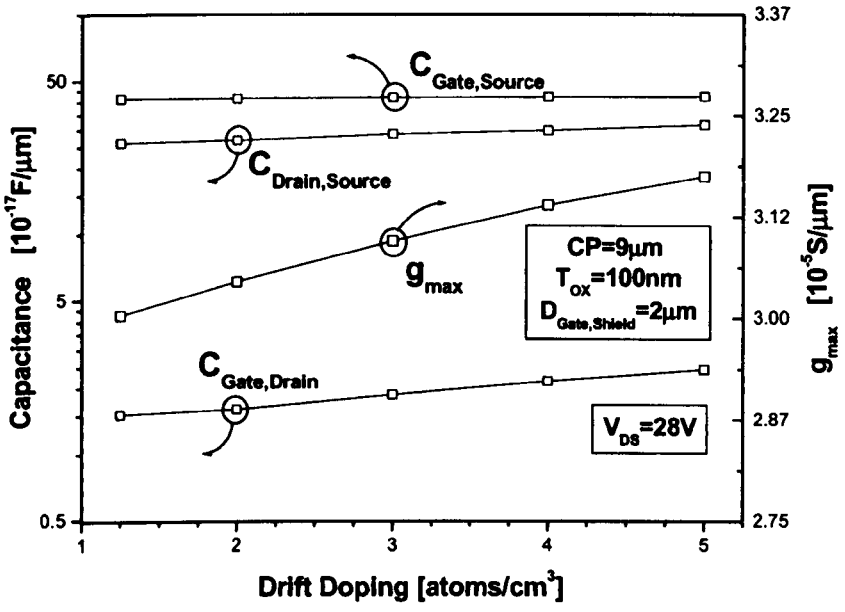


Figure 5.31: Effect of the drift doping on capacitance content and transconductance peak value of the shielded VDMOSFET.

Effect of the drift doping concentration

The effect of the drift-doping on capacitance and transconductance for the optimum shielded device is reported in Figure 5.31. Mostly, the doping increase affects gate to drain capacitance and transconductance. For high drift doping concentrations, C_{GD} increases because of shallower p-well/n-drift depletion regions, whereas g_m improves because of the reduction of the drift region resistance. The reduction of the drift region resistance determines an increase of the voltage drop across the channel. This produces an increase of the peak transconductance value and delays the transconductance fall-off.

Figure 5.32 and Figure 5.33 show the calculated RF-related values for devices' widths dimensioned to generate 20W of RF output power.

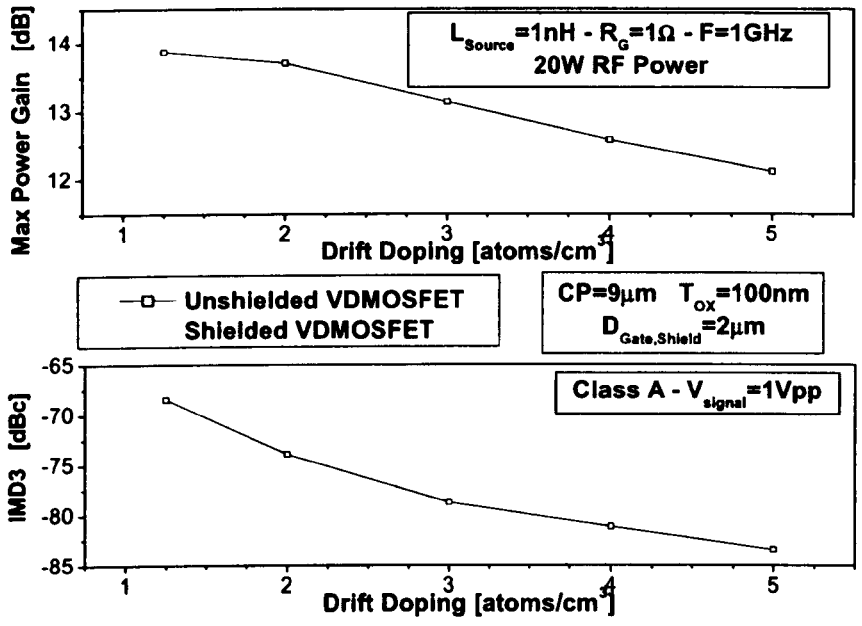


Figure 5.32: Effect of the drift doping on power gain and intermodulation distortion of the shielded VDMOSFET.

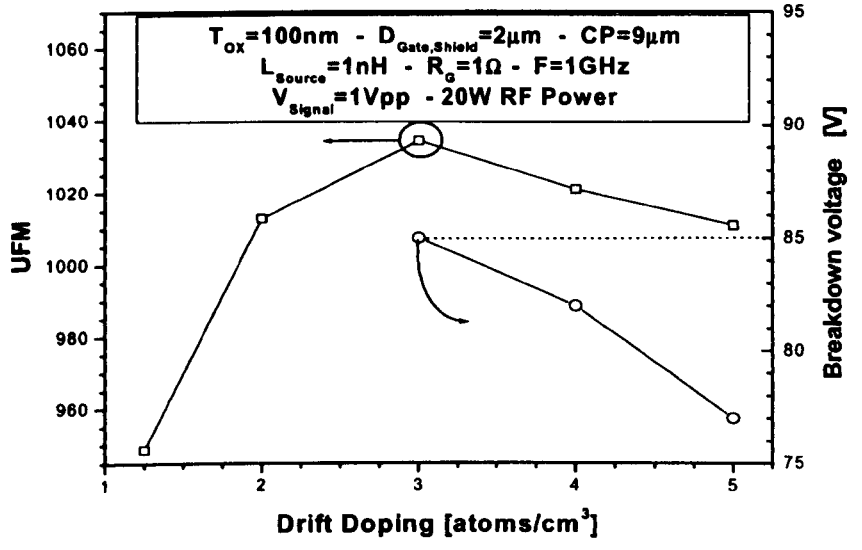


Figure 5.33: Effect of the drift doping on the UFM of the shielded VDMOSFET.

Figure 5.32 shows the effects of capacitance and transconductance variations on power gain and linearity. It can be seen that the power gain degrades and linearity improves for increased n-drift doping concentrations. The determination of the

UFM permits to resolve this RF performance trade-off. Figure 5.33 reveals an optimum n-drift doping concentration of $3 \cdot 10^{15}$atoms/cm^3. For higher n-drift doping concentrations the breakdown voltage cannot be kept at the target value of 85V, yielding devices unsuitable for 28V RF applications. Therefore, the results in Figure 5.33 show that the best RF performance is associated with the maximum drift doping concentration for which the breakdown voltage specifications remain satisfied.

Minimising the oxide thickness

It has already been shown that the optimum shield plate positioning is associated with the smallest considered oxide thickness T_{OX}. The minimum shield oxide thickness that can be used depends on the quality of the oxide and on the maximum voltage drop across the shield MOS structure. Depending on these factors, the minimum T_{OX} has to be dimensioned in order to avoid the shield oxide breakdown. The simulation results in Figure 5.34 indicate a maximum voltage drop across the shield of about 30V. The minimum oxide thickness at the shield is accordingly set to 30nm, considering a critical electric field of 10^7V/cm for the silicon dioxide.

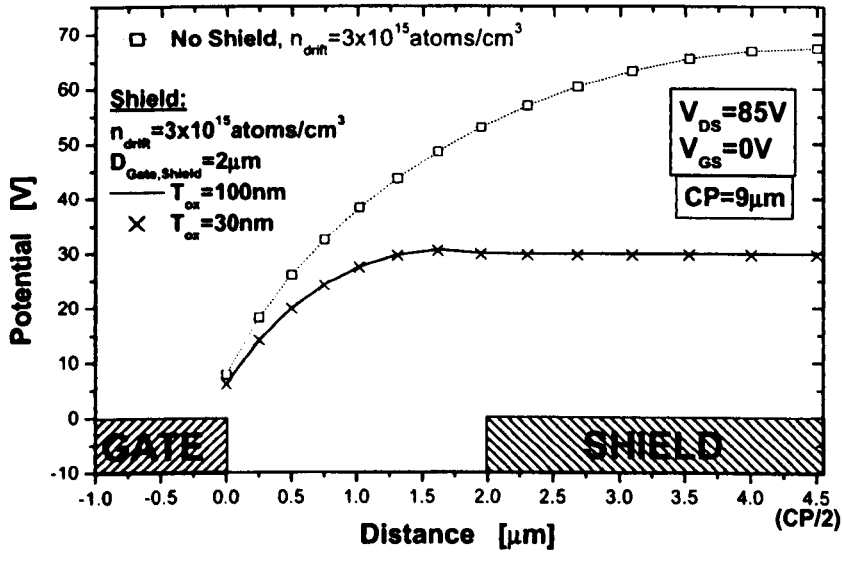


Figure 5.34: Potential at the silicon-oxide interface.

The reduction of the oxide thickness at the shield plate causes the degradation of linearity and the improvement of power gain shown in Figure 5.35. The maximum UFM is achieved at the minimum possible T_{OX} value of 30nm.

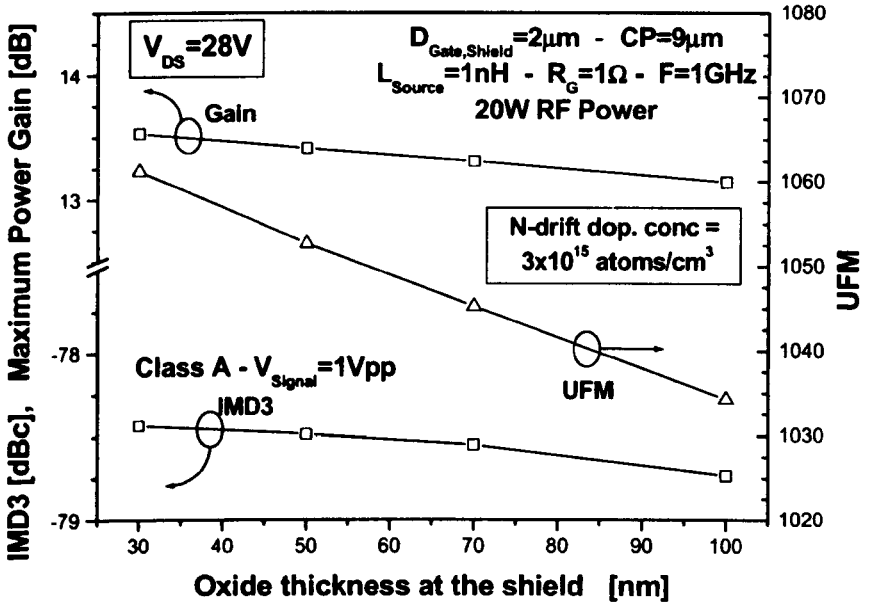


Figure 5.35: Effect of drift doping on linearity, power gain and UFM.

Adjusting the device Cell Pitch

The effect of the cell pitch on capacitance and transconductance for the optimum shielded device is reported in Figure 5.36. The cell pitch variation causes small modification on the device capacitance, but a noticeable change in peak transconductance value. This behaviour follows from the reduction of the drift resistance: increased cell spacing yields a smaller JFET resistance and a reduction of the p-well/n-drift depletion region width. The transconductance is observed to reach a saturated value at large cell pitches. This corresponds to adjacent cells that do not considerably affect each other: the JFET resistance becomes independent of the cell pitch.

The p-well/n-drift depletion region width reduction associated to increasing cell spacing causes a slight increase of C_{depl} and C_{GD}. The gate to drain capacitance

increase and the improvement in transconductance are responsible for the power gain reduction and linearity improvements shown in Figure 5.37.

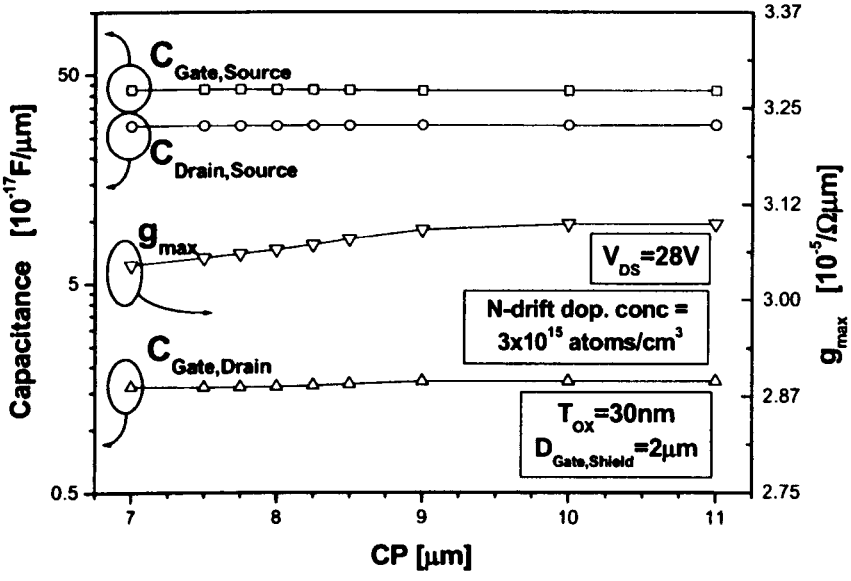


Figure 5.36: Effect of the cell spacing on capacitance content and transconductance peak value of the shielded VDMOSFET.

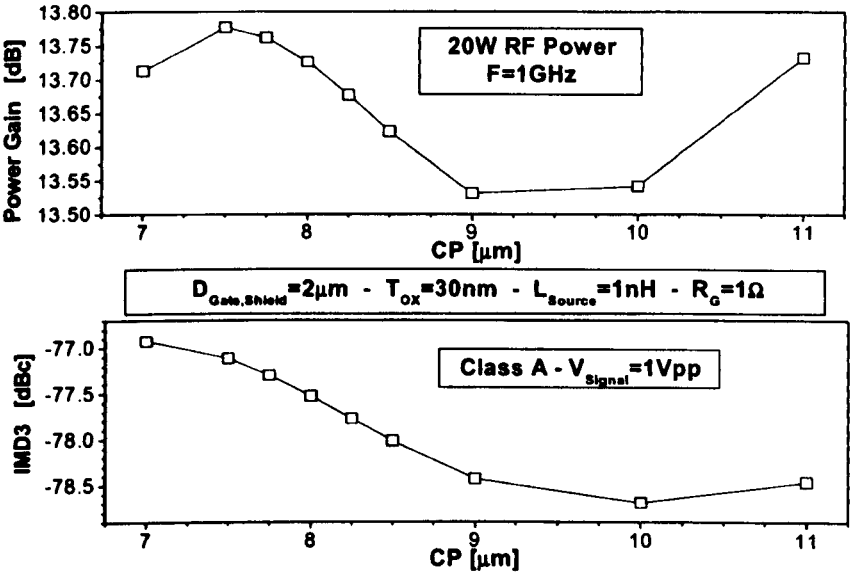


Figure 5.37: Effect of the cell spacing on power gain and intermodulation distortion of the shielded VDMOSFET.

The RF performance trade-off is resolved considering the UFM values in Figure 5.38. The UFM reveals an optimum cell pitch of 8μm. Although the UFM improves along with cell pitches above 9μm, CP cannot be arbitrarily increased since the device breakdown voltage falls below the target 85V value.

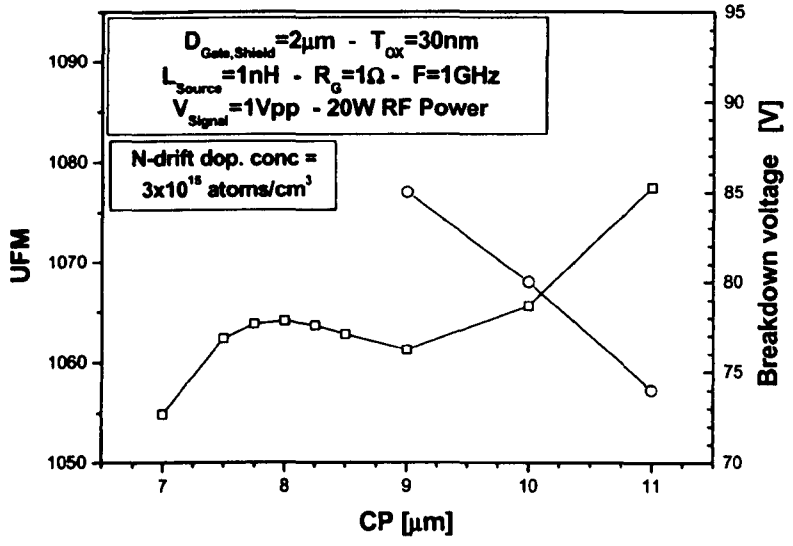


Figure 5.38: Effect of cell spacing on UFM of the shielded VDMOSFET.

Optimised device

The optimised shielded VDMOSFET presents a drift doping concentration of $3 \cdot 10^{15}$atoms/cm^3, a shield oxide thickness T_{OX} of 30nm, a cell spacing CP of 8μm and $D_{Gate,Shield}$=2μm. The gate oxide thickness, gate length and channel length values are 70nm, 1μm and 0.9μm respectively.

The device displays a breakdown voltage of 85V and a threshold voltage V_{th} of 3.05V. The threshold voltage value has been extrapolated using the at constant V_{DS}=28V as the offset voltage of the line tangent to the $I^{0.5}{}_{DS}$-V_{GS} curve where the slope of the curve peaks (ESR method [19]).

The device IV output characteristics are reported in Figure 5.39 for a swing of V_{DS} and V_{GS} values. The device displays good current saturation properties. The transconductance characteristic at V_{DS}=28V and the capacitance content of the device are reported in Figure 5.40.

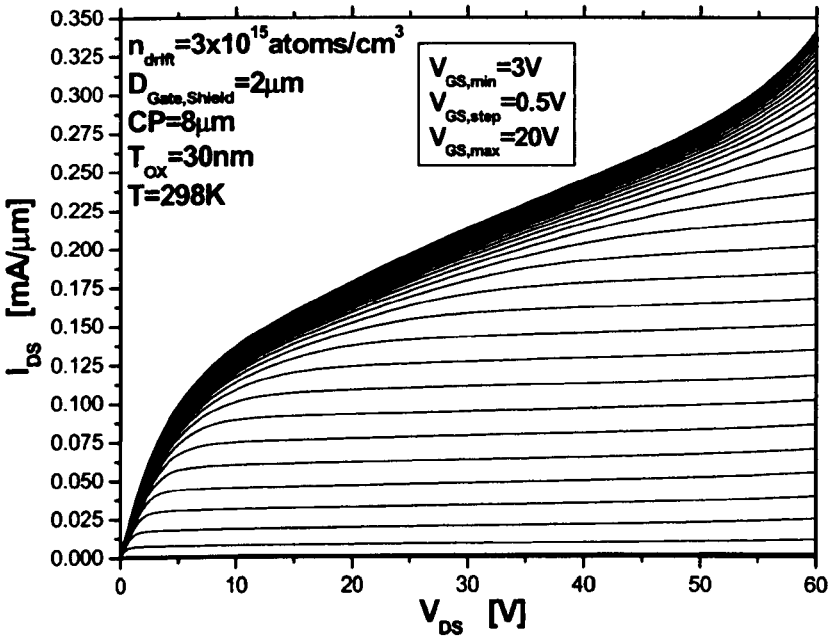


Figure 5.39: Output current vs. voltage characteristics of the shielded VDMOSFET optimised for 28V RF applications.

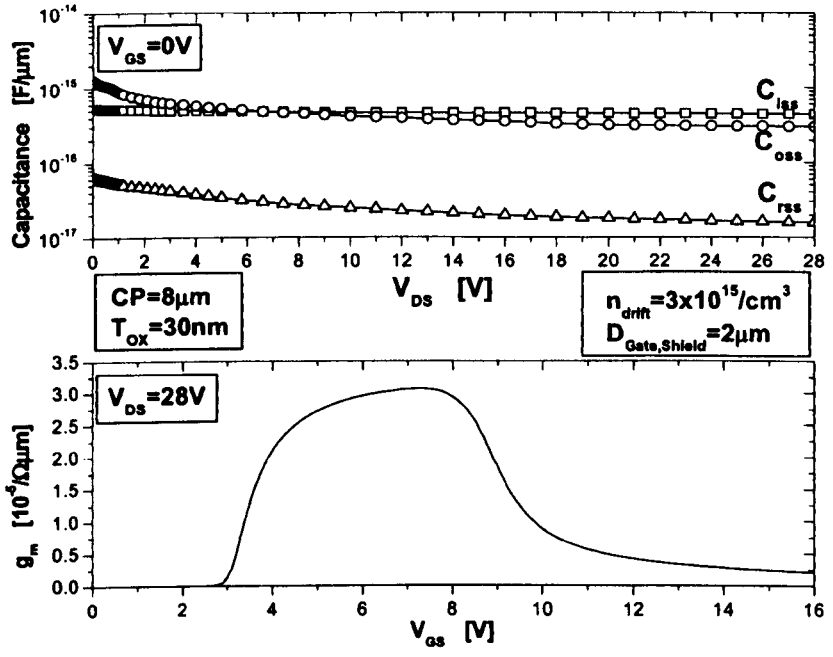


Figure 5.40: Capacitance vs. drain to source voltage and transconductance versus gate to source voltage of the shielded VDMOSFET optimised for 28V RF applications.

5.3.5 Shielded vs. unshielded VDMOSFET

This section presents a comparison of RF performance of the optimum shielded and the unshielded VDMOSFET. The devices present the same breakdown voltage of 85V; the total gate width of the devices is dimensioned for a 20W RF output power at an application voltage V_{DS}=28V. The main characteristics of the considered devices are reported in Table 5.1.

	UNSHIELDED	SHIELDED
C_{iss} [F/μm] (V_{DS} =28V V_{GS}=0V F=1MHz)	$4.30 \cdot 10^{-16}$	$4.40 \cdot 10^{-16}$
C_{oss} [F/μm] (V_{DS} =28V V_{GS}=0V F=1MHz)	$2.95 \cdot 10^{-16}$	$2.96 \cdot 10^{-16}$
C_{rss} [F/μm] (V_{DS} =28V V_{GS}=0V F=1MHz)	$3.33 \cdot 10^{-17}$	$1.61 \cdot 10^{-17}$
$g_{m,max}$ [S/μm] (V_{DS} =28V)	$3.16\ 10^{-5}$	$3.07 \cdot 10^{-5}$
W [mm]	29	23
L_S [nH]	1	
R_G [Ω]	1	

Table 5.1: Device characteristics.

A comparison of the frequency capability is carried out through the calculation of power gain and linearity by the application of the Fourier Spectrum Analysis described in Chapter 4.

Figure 5.41 shows that the shielded device has an improvement in maximum operating frequency F_{max} as well as in cut off frequency F_{cut}. While F_{max} is not considerably enhanced, a large improvement of F_{cut} is observed. This is a consequence of the small input capacitance C_{iss} associated with the shielded device, according to the well know cut off frequency approximation

$$F_{cut} \approx \frac{g_m}{2\pi C_{iss}}, \tag{5.29}$$

$$C_{iss} = C_{GS} + C_{GD}. \tag{5.30}$$

Considering the value per unit length reported in Table 5.1, the input capacitance C_{iss} is in fact reduced from about 12.5pF in the unshielded VDMOSFET to about 9.3pF in the shielded case.

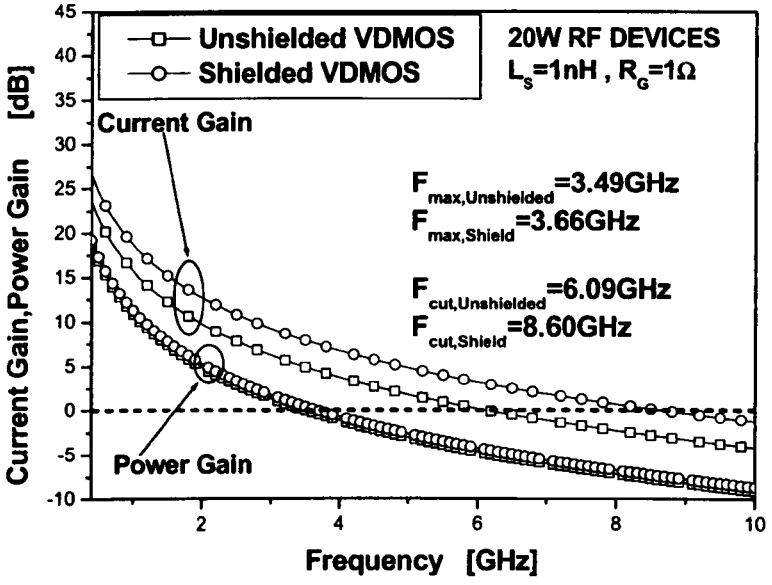


Figure 5.41: Current gain versus frequency for the optimised drift shielded and standard VDMOSFETs. The calculation includes the values of input and output matching impedances as a function of frequency and refers to a class A amplifier.

Figure 5.42 and Figure 5.43 compare the calculated power gain and linearity for shielded and unshielded VDMOSFETs, with the devices biased in class A.

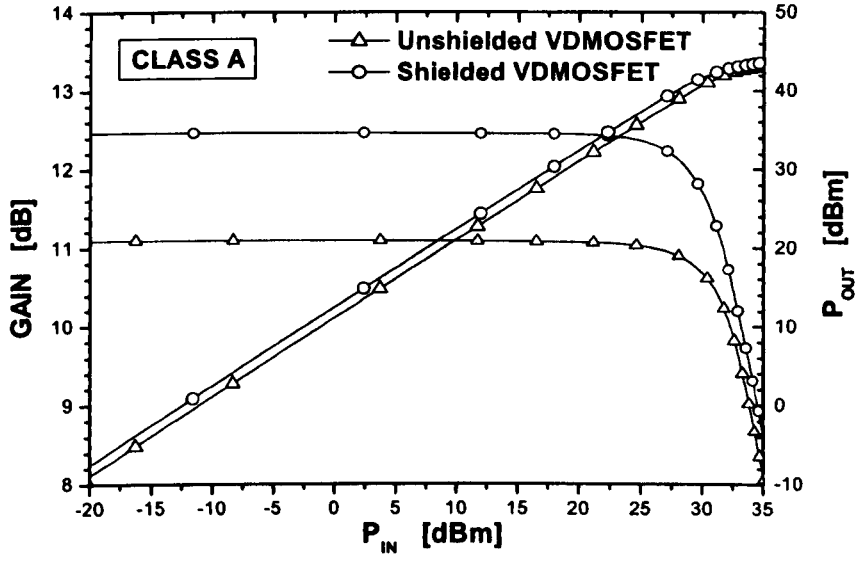


Figure 5.42: Class A power gain and output power versus input power for the optimised shielded and unshielded VDMOSFET.

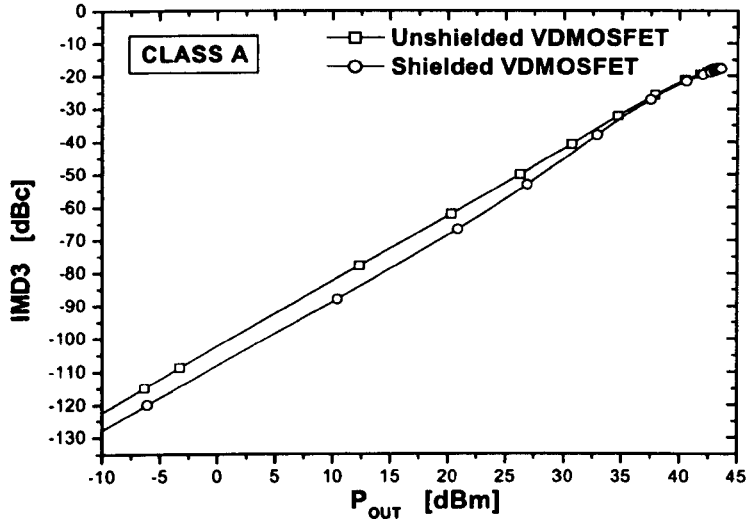


Figure 5.43: Class A third order intermodulation distortion versus output power for the optimised shielded and unshielded VDMOSFET.

At a frequency of 1GHz, the shielded device displays an enhancement of the power gain of about 1.5dB, corresponding to a 45% increase with respect to the unshielded VDMOSFET.

Linearity is also improved with the shielded VDMOSFET. As acknowledged in Figure 5.43, an average 5dBc reduction of IMD3 is observed with respect to the unshielded device over a large output power span.

The main RF power figures of merit for the shielded and unshielded VDMOSFET are summarised in Table 5.2. An overall performance improvement is observed with the shield plate.

	UNSHIELDED	SHIELDED
G [dB]	11.09	12.46
$I_{DS,Q}$ [A]	1.37	1.46
$V_{GS,Q}$ [V]	5	6
F_{cut} [GHz]	6.09	8.6
Fmax [GHz]	3.45	3.66
P_{1dB} [W]*	16.54	16.45
IMD3* [dBc]	-18.19	-19.45
PAE* [%]*	38.92	37.33

Table 5.2: Class A performance of the 20W shielded and unshielded VDMOSFETs. '*' indicates that the extraction is carried out at the 1dB compression point.

5.4 The GP-VDMOSFET

Large improvement in RF performance can be achieved in a VDMOSFET when a shield plate is used. The main effect of a shield plate design is the modification of the potential and electric field distributions where, by creating a depleted region between the gates, the shield plate considerably reduces the electric field at the top of the device, improving the breakdown voltage and decreasing the feedback capacitance.

In this section, the novel Grounded P-region VDMOSFET (GP-VDMOSFET) is presented. Instead of using a shield plate, a depleted region in the drift is created with a grounded p-region at the top of the device, between adjacent p-wells. The grounded p-region allows for a deeper penetration of the depletion region in the drift than in the shield plate case, causing further reduction of the feedback capacitance. On the other hand, this enlarged depletion region causes degradation of the transconductance, yielding an RF performance trade-off.

5.4.1 Device description

The unshielded VDMOS presented in section 5.3 has been considered as a starting point for this analysis. As shown in Figure 5.44, the p-region position in the grounded p-region VDMOSFET can be described by the parameters D_P and W_P, where D_P represents the depth and W_P is the implant window opening of the p region; the distance between the p-region and the p-well follows $D_{GP}=CP/2-L_P/2$, where L_P is the lateral extension of the p-region implant: $L_P \approx W_P+1.5*D_P$.

In this device, attention has to be paid in the design of the p-region in order to avoid a possible punch through at high drain to source voltages. Punch through is likely to occur for low p-region doping concentrations and small p-region contact to n-drift distances, as will be described in section5.4.3.

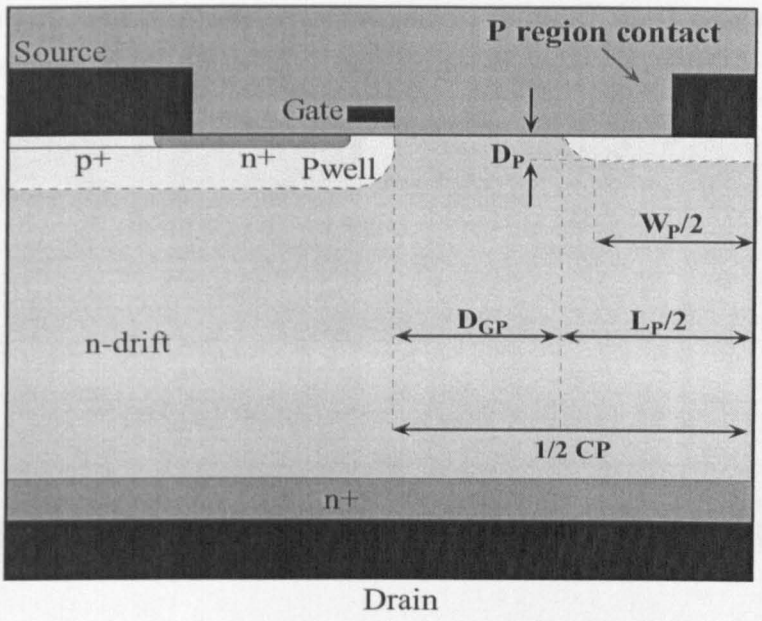


Figure 5.44: Considered device topology for the GP_VDMOSFET.

5.4.2 Similarities with the drift shield design

Analogous to the MOS shield case described in section 5.3, the p-region design determines a depletion region in the drift region of the VDMOSFET. The depletion region forces a more uniform potential distribution in the device, as shown in Figure 5.45.

The more uniform potential distribution yields a reduction of the peak electric field at the drain-end of the channel region reported in Figure 5.46. This causes the grounded p-region VDMOSFET breakdown to occur in correspondence to the point of maximum curvature of the p-region, as in the shield plate case.

The p-region effect on resistance and capacitance is analogous to the effect determined by the shield plate. Figure 5.46 also shows the increased effect of the grounded p-region with respect to the shield design in reducing the electric field at the channel drain end. This results from a deeper penetration of the depletion region in the drift when a grounded p-region is used.

The p-region/n-drift depletion narrows the JFET channel associated with the cell spacing, causing a resistance increase. Such an increase is responsible for the

reduction of the voltage drop across the intrinsic MOSFET and therefore for the degradation of the transconductance characteristic shown in Figure 5.47.

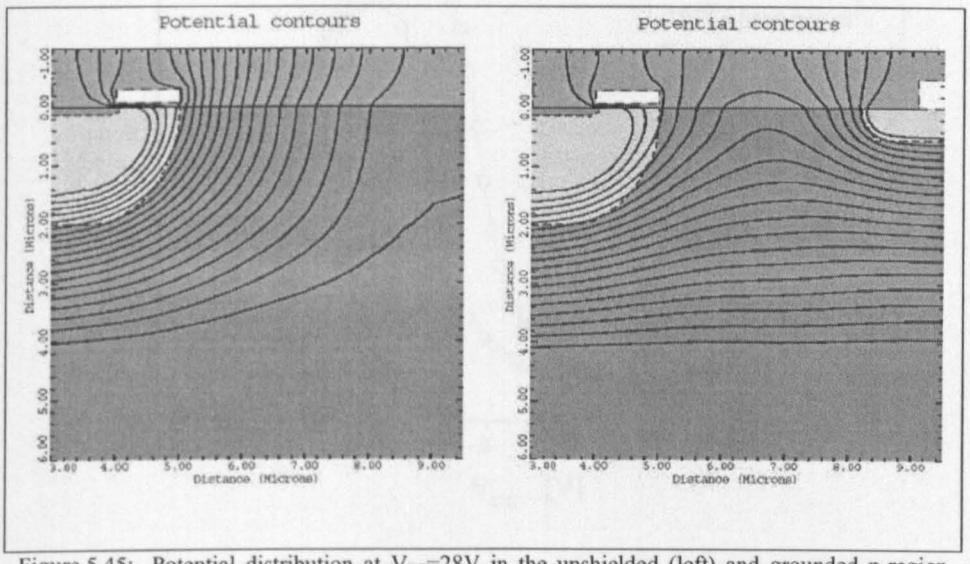


Figure 5.45: Potential distribution at V_{DS}=28V in the unshielded (left) and grounded p-region VDMOSFET (right). For the grounded p-region device the p-region doping is 10^{19} atoms/cm^3, W_P=2μm and T_P=0.5μm.

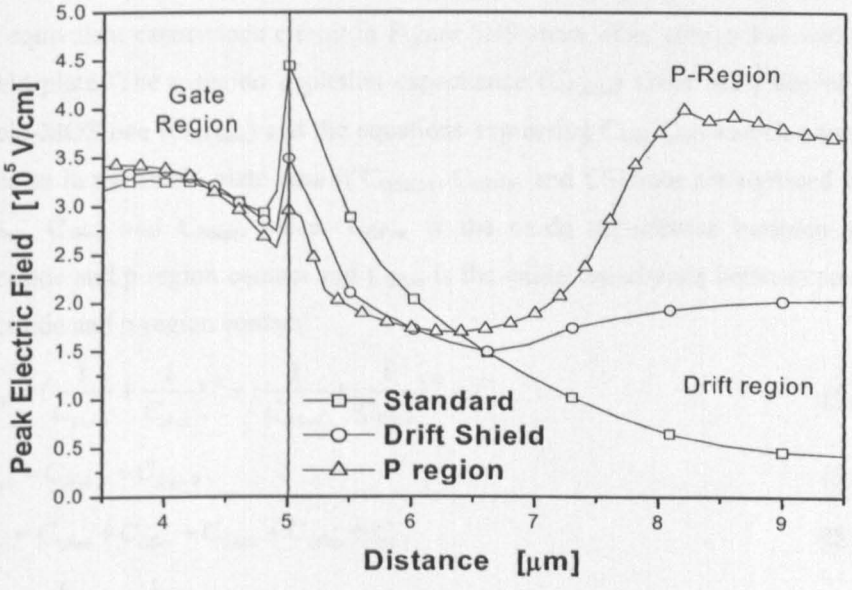


Figure 5.46: Peak electric field at breakdown for the unshielded, shielded and grounded p-region VDMOSFET. The grounded p-region device has a p-region doping of 10^{19} atoms/cm^3, W_P=2μm and T_P=0.5μm.

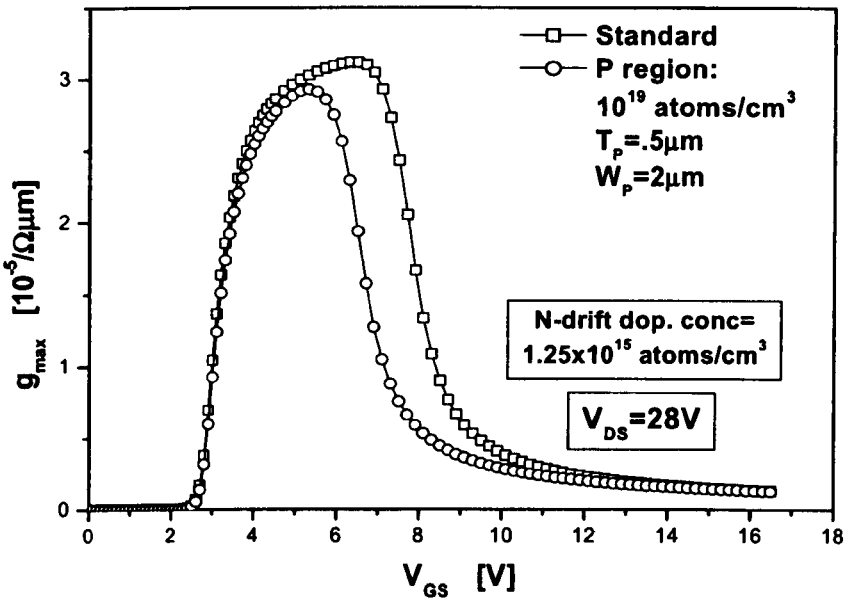


Figure 5.47: Transconductance versus gate to source voltage for the standard and grounded p-region VDMOSFET for unmodified drift doping profiles.

The analysis of the capacitance content in Figure 5.48 and the determination of the equivalent capacitance circuit in Figure 5.49 show other similarities with the shield plate. The p-region depletion capacitance ($C_{P,depl}$) takes the place of the shield-MOS one ($C_{SH,mos}$) and the equations expressing C_{GS}, C_{GD} and C_{DS} are the same as in the shield plate case if C_{GSHox}, C_{SSHox} and CSHmos are replaced with C_{GPox}, C_{SPox} and C_{Pdepl}, where C_{GPox} is the oxide capacitance between gate electrode and p-region contact and C_{SPox} is the oxide capacitance between source electrode and p-region contact:

$$C_{GD} \approx \left(\frac{1}{C_{Gmos}} + \frac{1}{C_{depl}}\right)^{-1} + \left(\frac{1}{C_{Pdepl}} + \frac{1}{C_{GPox}}\right)^{-1} + C_{GDov}, \tag{5.31}$$

$$C_{depl} = C_{depl-1} + C_{depl-2}, \tag{5.32}$$

$$C_{GS} \approx C_{GSox} + C_{GSov} + C_{Gmos} + C_{GPox} + C_1, \tag{5.33}$$

$$C_1 = \left(\frac{1}{C_{Pdepl}} + \frac{1}{C_{GDov}}\right)^{-1}, \tag{5.34}$$

$$C_{DS} \approx C_{depl} + \left(\frac{1}{C_{SPox}} + \frac{1}{C_{Pdepl}}\right)^{-1}. \tag{5.35}$$

As in the shield plate case, the p-region effect on the potential and electric field distribution causes a reduced C_{depl} with respect to the standard VDMOSFET, which yields a considerable reduction of C_{GD}, as it will be shown in section 5.4.3.

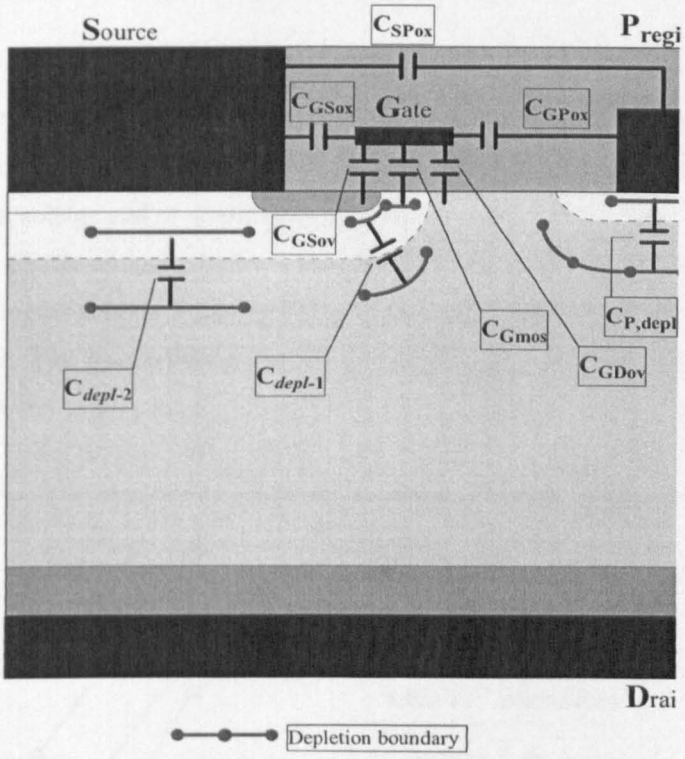


Figure 5.48: Capacitors intrinsically defined in the grounded p-region VDMOSFET.

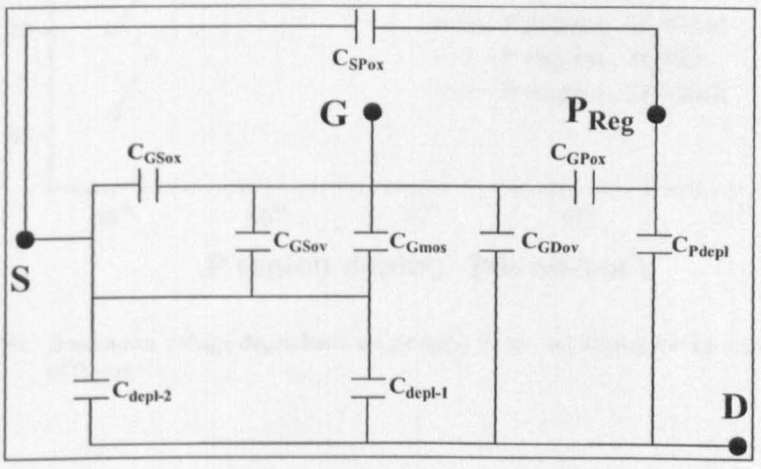


Figure 5.49: Equivalent capacitance circuit for the grounded p-region design.

5.4.3 Device optimisation

In this section the optimisation of the grounded p-region VDMOSFET is carried out by using the Unified Figure of Merit.

Avoiding the p-region punch through

An appropriate choice of doping concentration and dimension of the p-region permits to avoid the occurrence of punch-through, which would otherwise limit the breakdown voltage and/or the forward bias safe operating area.

The p-region punch through occurs when the p-region/n-drift depletion region reaches the p-region contact, allowing for an electron flow from the drain to the p-region contact. The punch-through is likely to occur for low p-region doping concentrations and small p-region contact to n-drift distances, as shown in Figure 5.50.

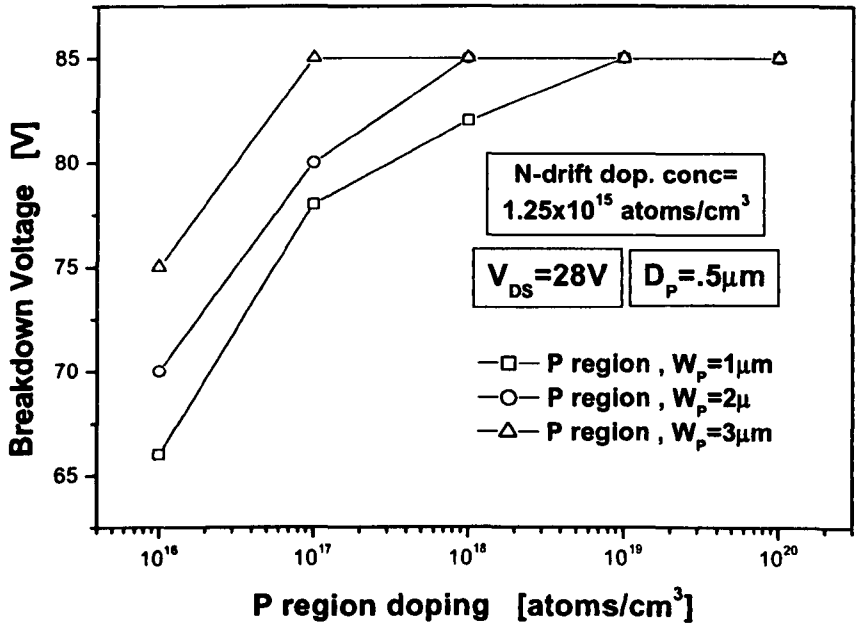


Figure 5.50: Breakdown voltage dependence on p-region width and doping for a p-region depth of 0.5μm.

Adjusting the p-region depth can prevent the p-region punch-through, as shown in Figure 5.51. Lower p-region depths are required for higher p-region concentrations.

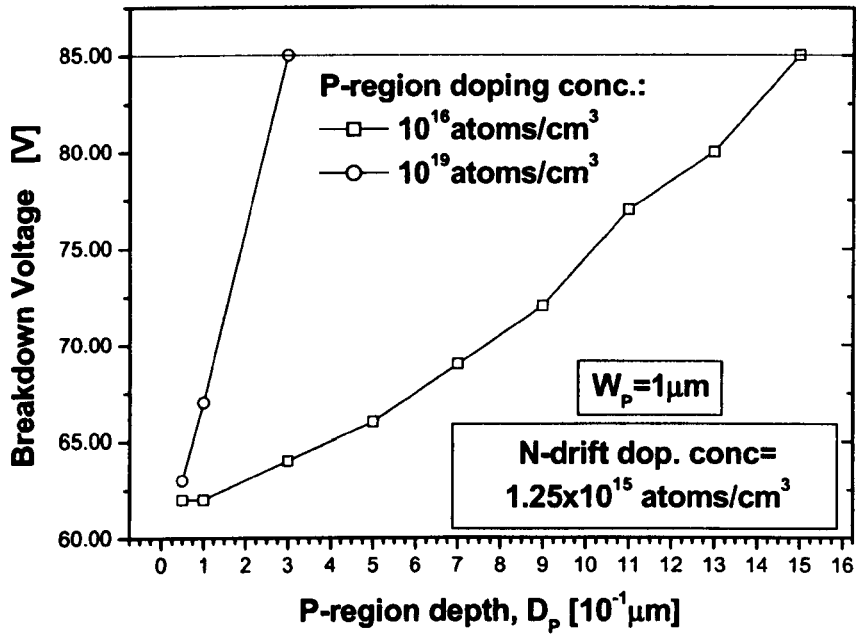


Figure 5.51: Effect of p-region depth and doping on the breakdown voltage for a p-region width of 2μm.

P-region doping and width effect

The doping of the p-region has a direct effect on the extension of the depletion region and therefore on both capacitance and resistance content of the device. The increase of the p-region doping concentration and/or of the p-region width causes the increase of the p-region/n-drift depletion region, yielding an increased drift resistance and a decreased C_{Pdepl}. Consequently, transconductance and gate to drain capacitance are reduced with respect to the standard VDMOSFET, as shown in Figure 5.52. Their reduction yields the increase of power gain and the degradation of linearity reported in Figure 5.53 and Figure 5.54.

The resulting RF trade-off is analysed by calculating the UFM. Figure 5.55 shows that improved performance is achieved for a p-region of small width and doping concentrations of about 10^{19} atoms/cm^3.

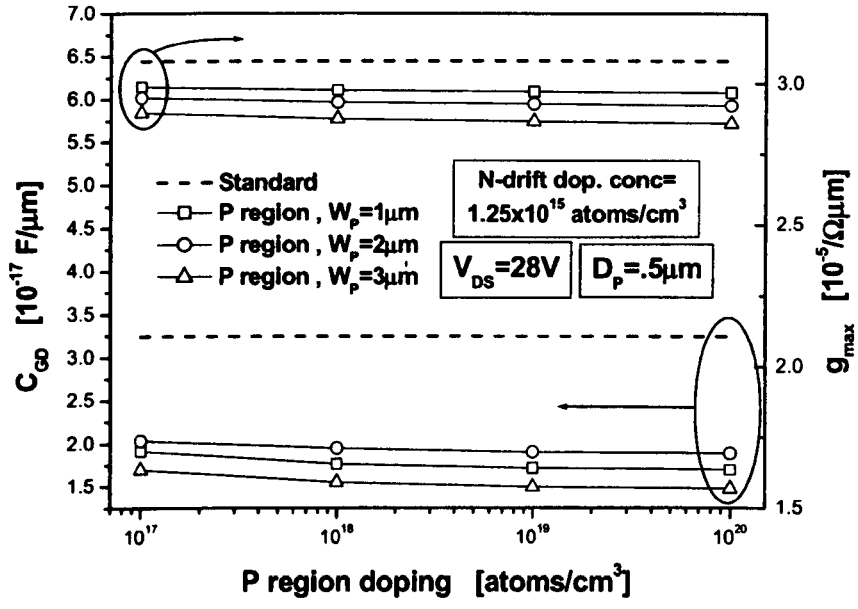


Figure 5.52: Effect of the p-region doping concentration on the gate to drain capacitance and peak transconductance values at V_{DS}=28V.

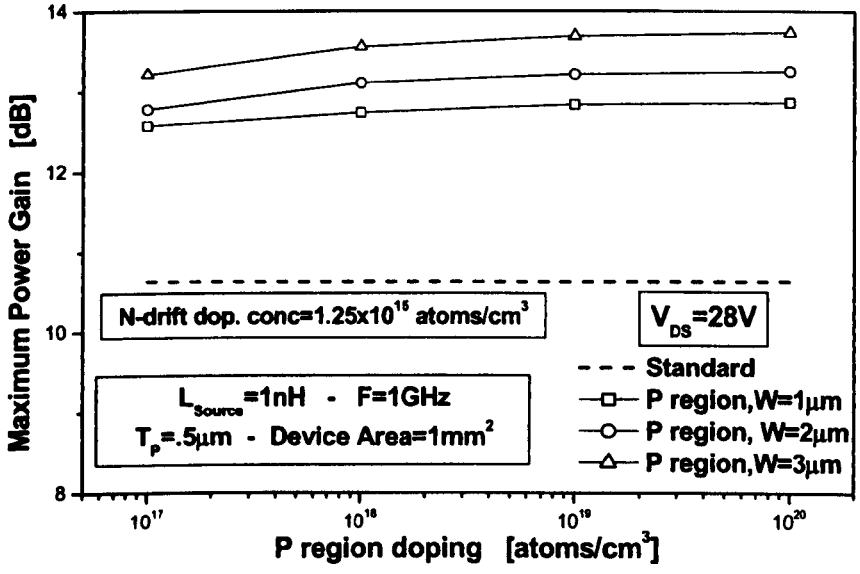


Figure 5.53: Effect of the p-region width and doping concentration on power gain.

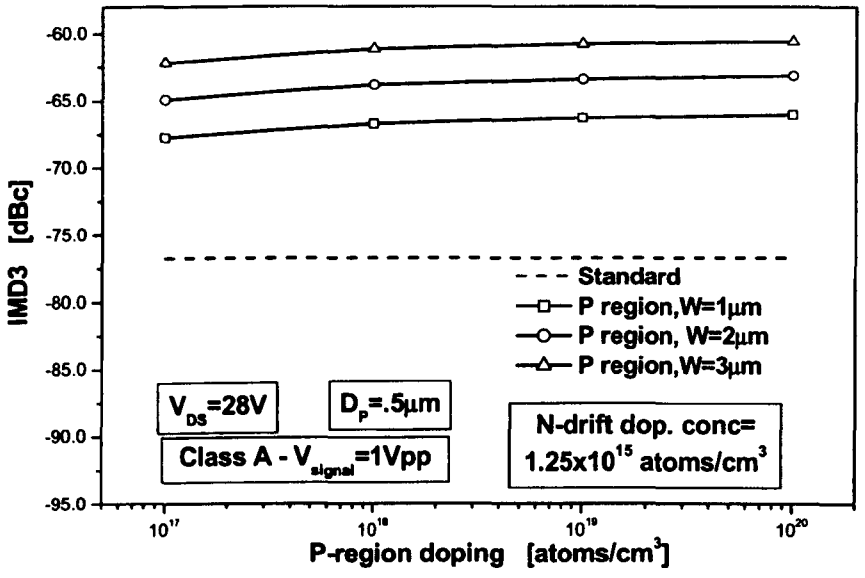


Figure 5.54: Effect of the p-region width and doping concentration on linearity.

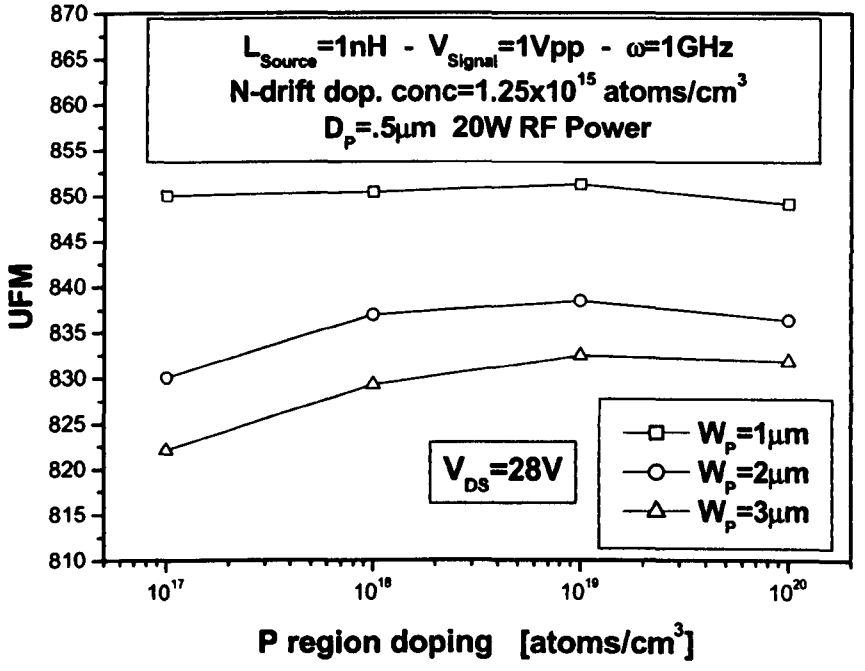


Figure 5.55: Effect of the grounded p-region doping concentration on the Unified Figure of Merit.

Effect of the p-region depth

A p-concentration of 10^{19} atoms/cm^3 and a window opening 2μm are considered in this section in order to limit the occurrence of punch through of the p-region in correspondence of small p-region depths D_P. The effect of the p-region implant depth D_P on C_{GD} and g_{max} is shown in Figure 5.56.

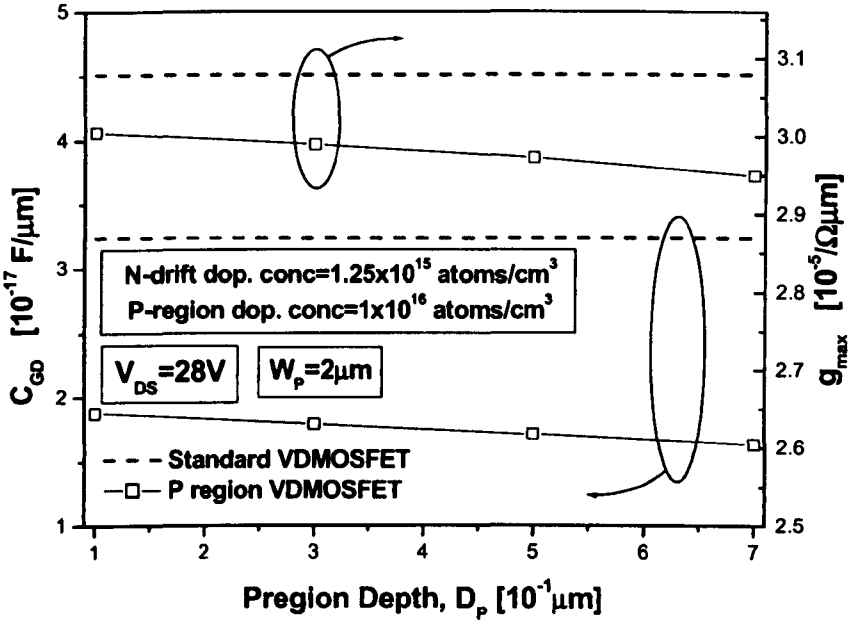


Figure 5.56: Effect of the p-region depth on the gate to drain capacitance and peak transconductance values at V_{DS}=28V.

An increased implant depth causes the p-region/n-drift depletion region to penetrate further in the drift region; the drift resistance increases and reduces the peak transconductance. The effect of the p-region on the potential and electric field distribution also increases, yielding the reduction of C_{depl} and C_{GD}.

In Figure 5.57, the calculated power gain, linearity and UFM values show that shallow p-region implants yield improved RF performance.

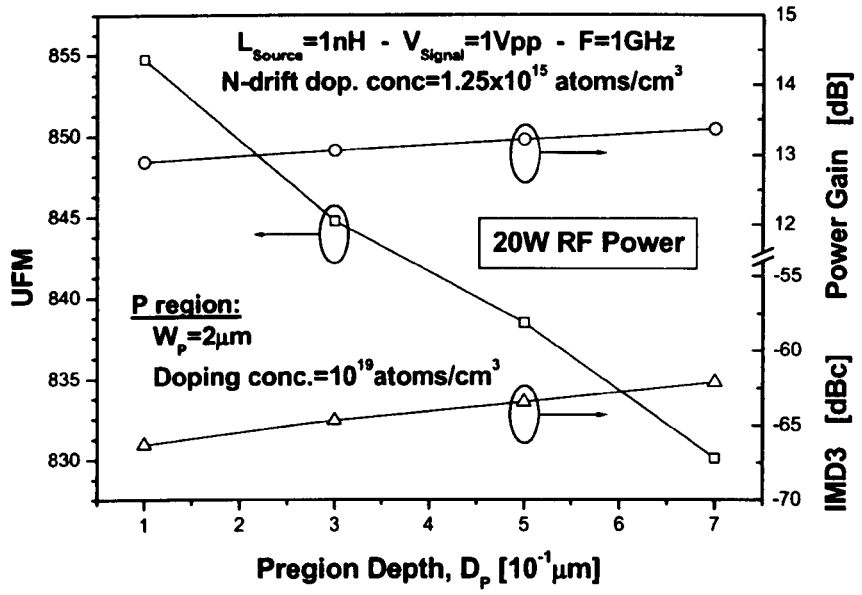


Figure 5.57: Effect of the p-region depth on power gain, linearity and Unified Figure of Merit for a p-region doping concentration of 10^{19}atoms/cm^3.

Effect of the n-drift doping concentration

Raising the drift-doping concentration causes the p-well/n-drift as well as the p-region/n-drift depletion regions to reduce, decreasing the drift resistance. For a certain bias, higher doping concentrations lead to higher electric fields across the depletion regions, causing a reduction of the breakdown voltage. Also, higher drift doping concentrations cause a deeper penetration of the p-region/n-drift depletion regions in p-side of the junction, which can lead to the punch through of the p-region.

The width and doping concentrations of the p-region are again kept constant at W_P=2μm and 10^{19}atoms/cm^3 respectively to prevent the p-region punch through. In order to achieve a breakdown voltage of 85V, the p-region depth is adjusted to compensate for raising drift doping concentrations. For fixed n-drift and p-region doping concentrations, the best RF performance will be associated to the smaller p-region depth for which p-region punch-through does not occur.

Figure 5.58 reports the effect of the p-region depth on the breakdown voltage as a function of the n-drift doping concentration. If the p-region depth is increased, a

breakdown voltage of 85V can be achieved for drift doping concentration as high as $4 \cdot 10^{15}$atoms/cm^3. For higher concentrations the p-region punches through even if the depth of the p-region is increased.

The increase of drift doping causes an increase of C_{GD} because of shallower p-well/n-drift depletion regions and the increase of g_m because of the reduction of the drift region resistance.

Figure 5.59 shows the gate to drain capacitance and the peak transconductance of the p-region VDMOSFET for increasing drift doping concentration where the p-region depth is the minimum depth which allows for avoidance of the occurrence of punch through.

The increase of D_P causes the gate to drain capacitance to remain almost constant despite the increasing doping concentrations. The negative effect of D_P on the transconductance is compensated by the reduction of drift resistance associated with the increase of drift doping.

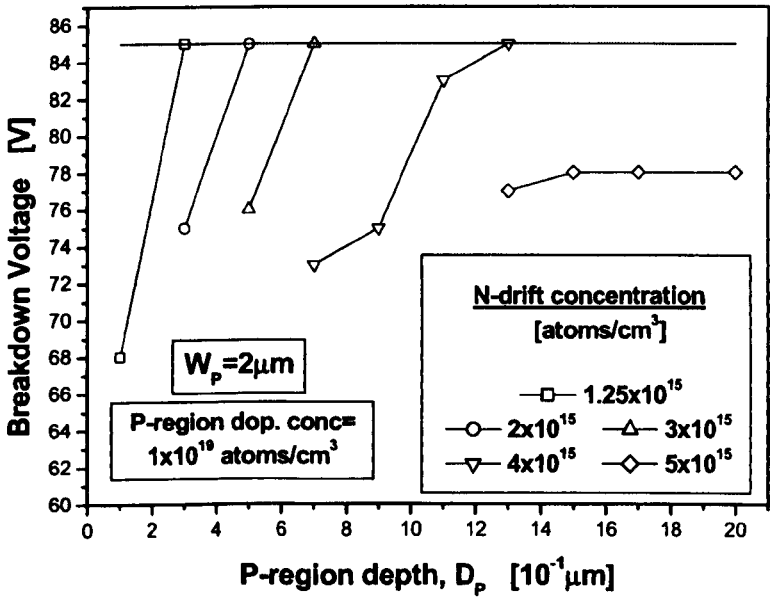


Figure 5.58: Effect of the p-region depth on the breakdown voltage as a function of the n-drift doping concentration. The p-region width and doping concentrations are kept constant at W_P=2µm and 10^{19}atoms/cm^3 respectively.

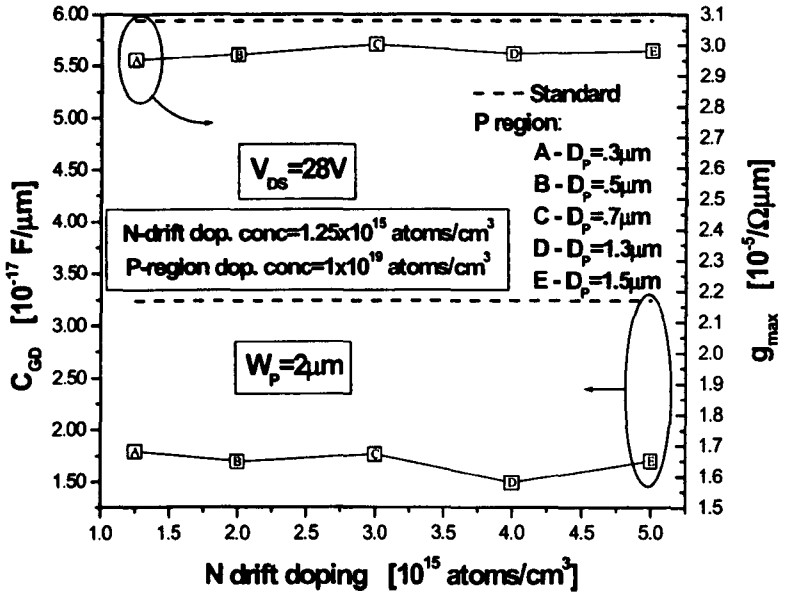


Figure 5.59: Effect of the p-region depth on the gate to drain capacitance and peak transconductance values at V_{DS}=28V.

Figure 5.60, Figure 5.61 and Figure 5.62 demonstrate the calculated power gain, linearity and UFM respectively. Again, higher power gain is achieved in correspondence with lower C_{GD} values, whereas linearity improves for increasing drift doping concentrations.

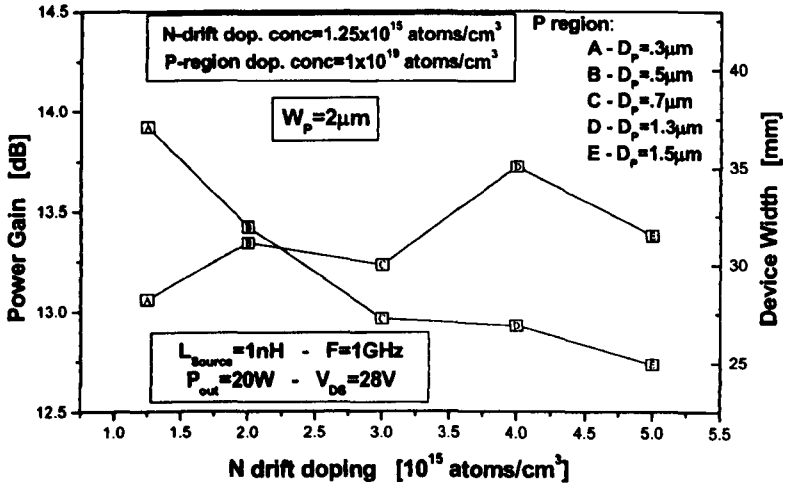


Figure 5.60: Effect of the n-drift doping on G_{max}[dB] for a p-region doping concentration of 10^{19}atoms/cm^3.

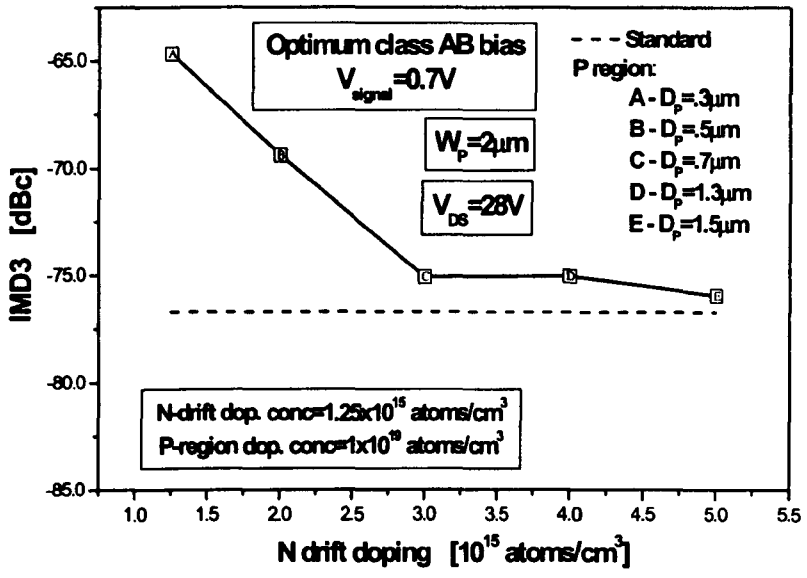


Figure 5.61: Effect of the n-drift doping on IMD3.

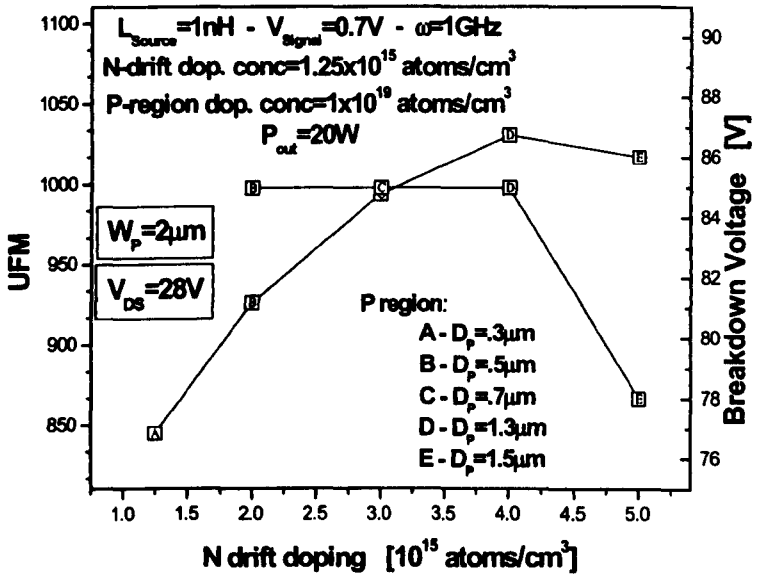


Figure 5.62: Effect of the n-drift doping on the Unified Figure of Merit in the grounded p-region VDMOSFET.

In Figure 5.62, the grounded p-region VDMOSFET shows a maximum UFM value at a drift doping concentration of $4 \cdot 10^{15}$atoms/cm^3.

Cell pitch adjustment

The effect of the cell pitch on capacitance and transconductance of the grounded p-region device is reported in Figure 5.63. The peak value of the transconductance improves with larger values of the cell pitch, due to the reduction of the JFET resistance. The feedback capacitance is instead found to degrade because of the reduced effect of the grounded p-region on the potential distribution in the drift.

The reduction of the p-well/n-drift depletion region width associated with increasing cell spacing causes a slight increase of C_{depl} and C_{GD}. The gate to drain capacitance increase and the improvement in transconductance yield the power gain reduction and linearity improvements shown in Figure 5.64.

The RF performance trade-off between linearity and power gain is resolved considering the UFM values in Figure 5.65. The UFM reveals an optimum cell pitch of 8.5μm.

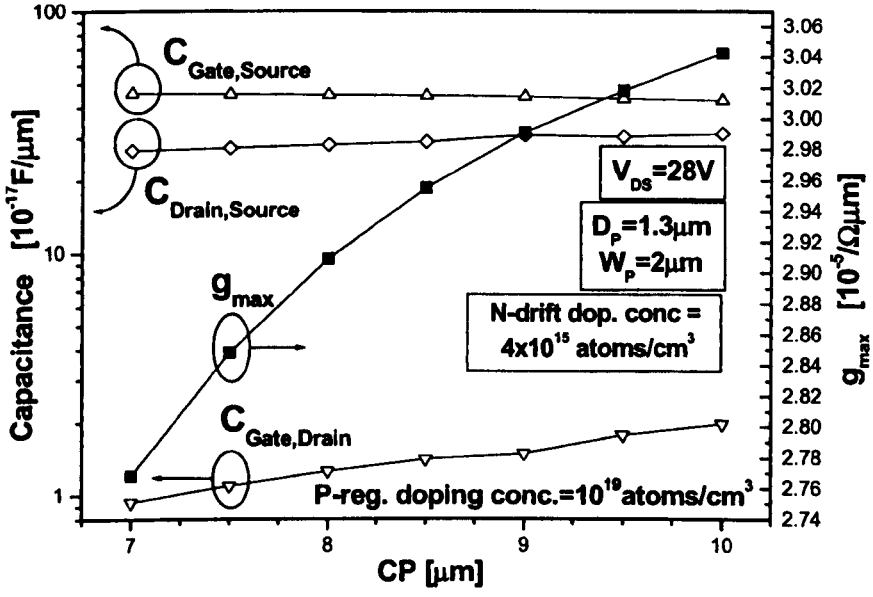


Figure 5.63: Effect of the cell spacing on capacitance content and peak value of the transconductanc of the grounded p-region VDMOSFET.

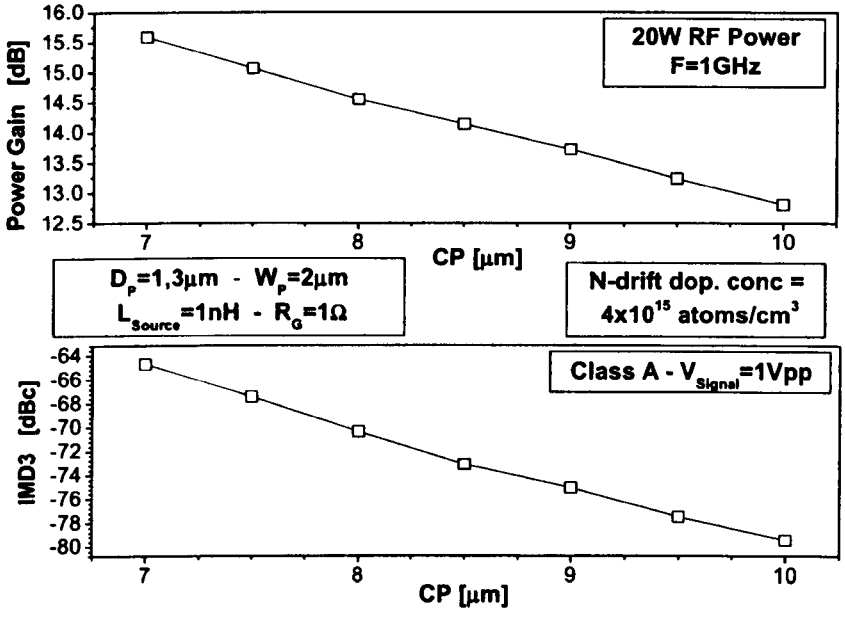


Figure 5.64: Effect of the cell spacing on power gain and intermodulation distortion of the grounded p-region VDMOSFET.

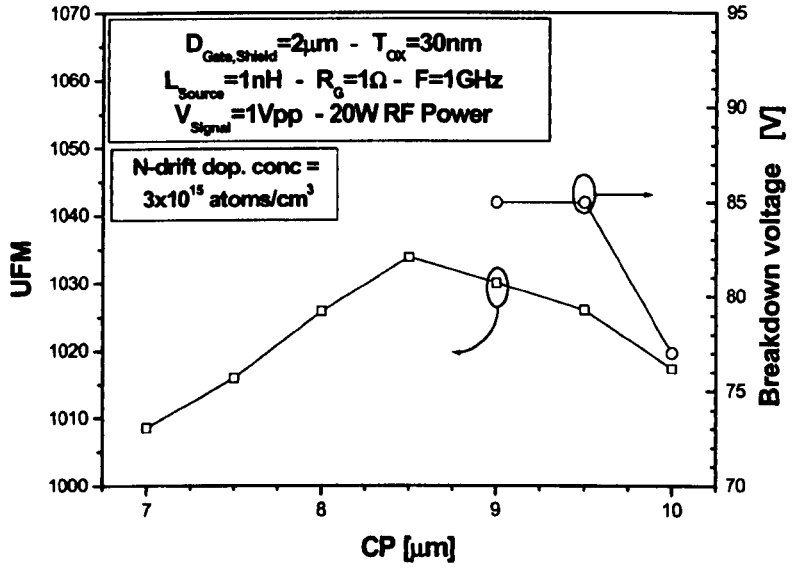


Figure 5.65: Effect of cell spacing on UFM of the grounded p-region VDMOSFET.

Optimised device

The optimised GP-VDMOSFET has a drift doping concentration of $4{\cdot}10^{15}$atoms/cm^3, W_P=2μm, D_P=1.3μm, CP=8μm and a p-region doping concentration of 10^{19} atoms/cm^3. The gate oxide thickness, gate length and channel length values are 70nm, 1μm and 0.9μm respectively.

The device displays a breakdown voltage of 85V and a threshold voltage V_{th} of 3.1V. The threshold voltage value has been extrapolated using the at constant V_{DS}=28V as the offset voltage of the line tangent to the $I^{0.5}{}_{DS}$-V_{GS} curve where the slope of the curve peaks (ESR method [19]).

The device IV output characteristics are reported in Figure 5.66 for a swing of V_{DS} and V_{GS} values. The device displays good current saturation properties. The transconductance characteristic at V_{DS}=28V and the capacitance content of the device are reported in Figure 5.67.

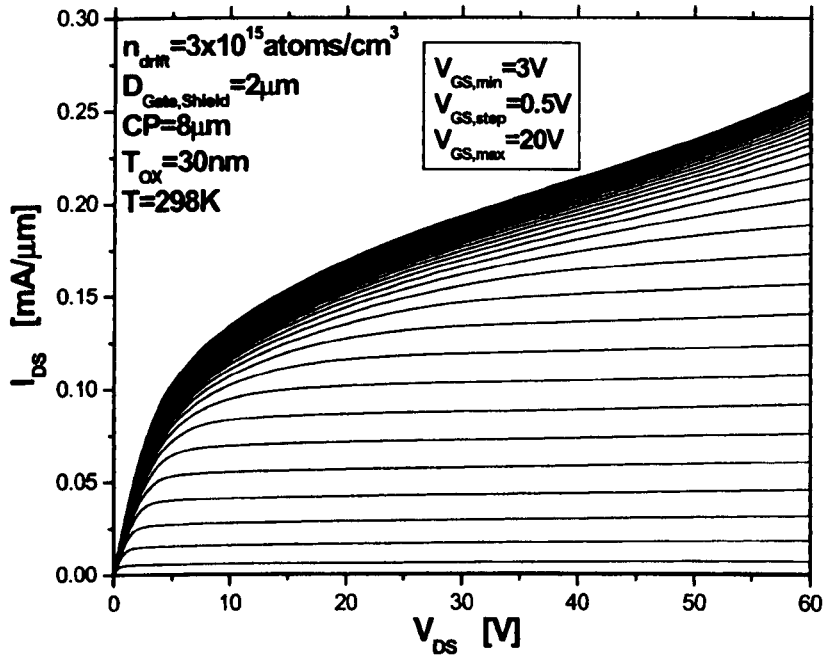


Figure 5.66: Output current vs. voltage characteristics of the grounded p-region VDMOSFET.

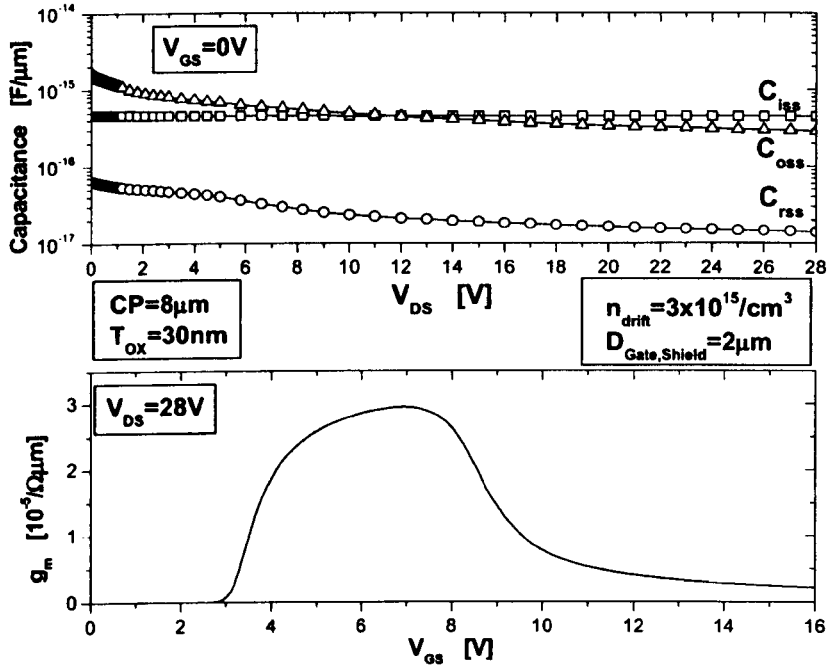


Figure 5.67: Capacitance vs. drain to source voltage and transconductance versus gate to source voltage of the grounded p-region VDMOSFET.

5.4.4 Shield plate versus grounded p-region

This section presents a comparison of RF performance of the optimum shielded and the unshielded VDMOSFET. The devices present the same breakdown voltage of 85V; the total gate width of the devices is dimensioned for 20W RF output power at an application voltage of V_{DS}=28V. The main characteristics of the considered devices are reported in Table 5.3.

	SHIELDED	Pregion
C_{iss} [F/µm] (V_{DS} =28V V_{GS}=0V F=1MHz)	$4.40 \cdot 10^{-16}$	$4.45 \cdot 10^{-16}$
C_{oss} [F/µm] (V_{DS} =28V V_{GS}=0V F=1MHz)	$2.96 \cdot 10^{-16}$	$2.87 \cdot 10^{-16}$
C_{rss} [F/µm] (V_{DS} =28V V_{GS}=0V F=1MHz)	$1.61 \cdot 10^{-17}$	$1.40 \cdot 10^{-17}$
$g_{m,max}$ [S/µm] (V_{DS} =28V)	$3.07 \cdot 10^{-5}$	$2.95 \cdot 10^{-5}$
W [mm]	23	23
L_S [nH]	1	
R_G [Ω]	1	

Table 5.3: Device characteristics.

A comparison of the frequency capability is carried out through the calculation of power gain and linearity by the application of the Fourier Spectrum Analysis described in Chapter 4.

With respect to the optimum shielded device, Figure 5.68 shows that the GP-VDMOSFET displays higher maximum operating frequency F_{max} but a reduced cut off frequency F_{cut}. The enhancement of F_{max} mainly derives from the feedback capacitance reduction reported in Table 5.3. The F_{cut} degradation is due to the increase of the input capacitance C_{iss} and the degradation of the transconductance.

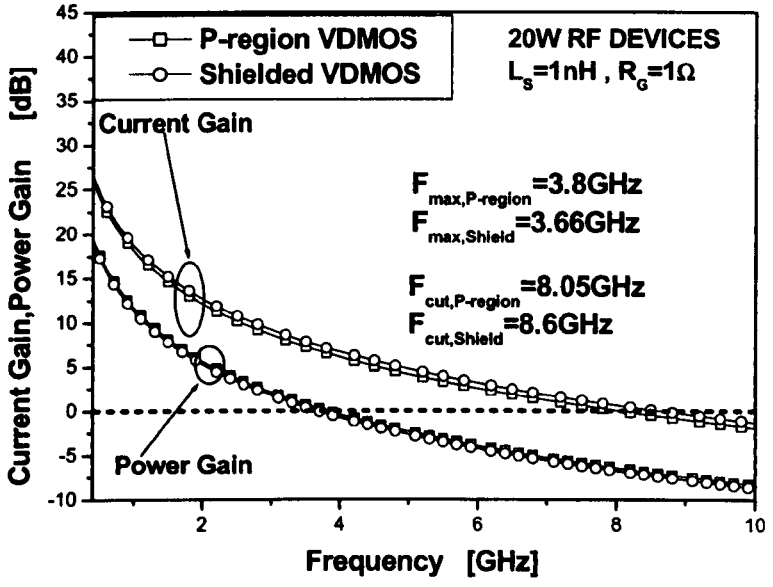


Figure 5.68: Current gain versus frequency for the optimised drift shielded and grounded p-region VDMOSFETs. The calculation includes the values of input and output matching impedances as a function of frequency and refers to a class A amplifier.

Figure 5.69 and Figure 5.70 compare the calculated power gain and linearity for the shielded and grounded p-region VDMOSFETs, with the devices biased in class A.

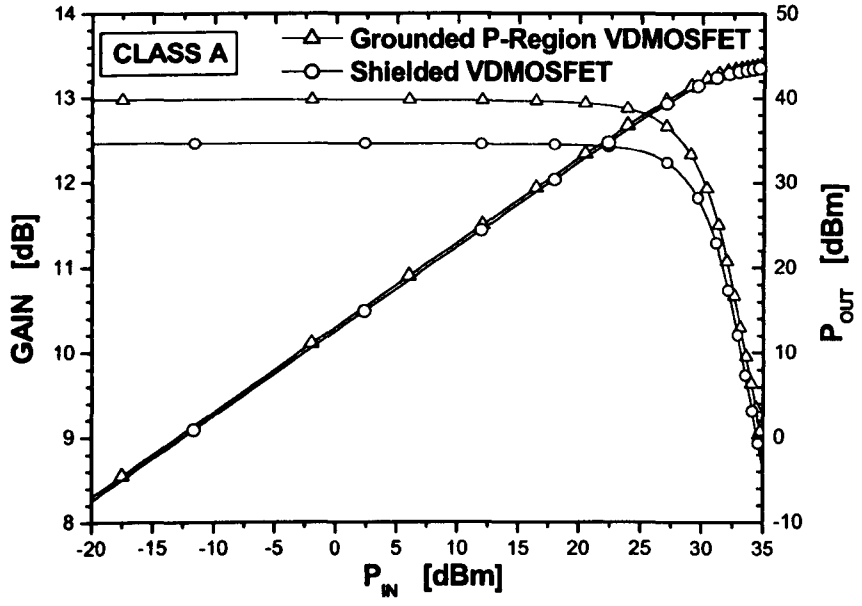


Figure 5.69: Class A power gain and output power versus input power for the optimised shielded and grounded p-region VDMOSFET.

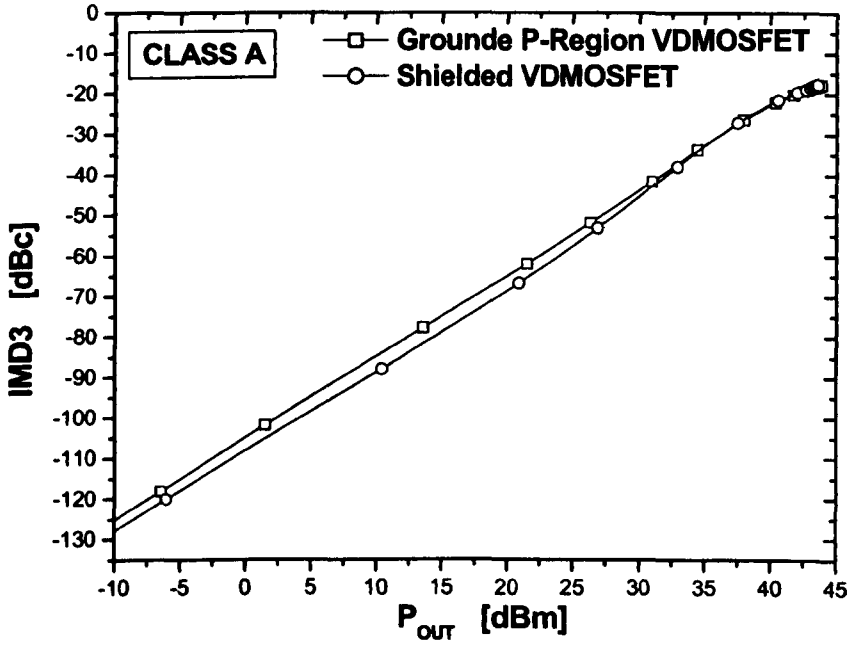


Figure 5.70: Class A third order intermodulation distortion versus output power for the optimised shielded and grounded p-region VDMOSFET.

At a frequency of 1GHz, the grounded p-region device displays an enhancement of the power gain in excess of 0.5dB, corresponding to a 4.3% increase with respect to the unshielded VDMOSFET.

Linearity is however slightly degraded with respect to the shielded VDMOSFET. As acknowledged in Figure 5.70, an average 2dBc reduction of IMD3 is observed with the GP_VDMOSFET with respect to the ahielded VDMOSFET.

The main RF power figures of merit for the shielded and grounded p-region VDMOSFET are summarised in Table 5.4.

	SHIELDED	P-REGION
G [dB]	12.46	13
$I_{DS,Q}$ [A]	1.46	1.42
$V_{GS,Q}$ [V]	6	6.08
F_{cut} [GHz]	8.6	8.05
Fmax [GHz]	3.66	3.8
P_{1dB} [W]*	16.45	16.61
IMD3* [dBc]	-19.45	-19.7
PAE* [%]*	37.33	40.2
* extracted at the 1dB compression point.		

Table 5.4: Class A performance of the 20W shielded and grounded p-region VDMOSFETs.

5.5 Summary

A systematic analysis of the effect of shield geometry and topology on a conventional RF power VDMOSFET has been carried out. Three different shield plate configurations were considered. The effects of the design parameters on breakdown voltage, capacitance and transconductance have been studied. It has been highlighted that the shield plate generally causes a reduction of the of the transconductance peak value. This reduction has been directly linked with the effect of the drift resistance on the transconductance fall-off voltage. The variation of capacitance caused by the shield plate has been clarified via an equivalent capacitance model.

A novel shield plate configuration where the shield is placed over the gate electrode has been presented and analysed. The observed feedback capacitance reduction has been explained through a previously unconsidered charge coupling mechanism.

An innovative RF based optimisation of the shield plate has been carried out. It has been shown how the shield plate determines trade-off of the RF performance and how these can be resolved by using the novel Unified Figure of Merit, yielding the identification of an optimum RF power device. Maximum power gain and linearity expression have been proposed for the determination of this figure of merit. The RF performance of the optimum shielded device has been assessed. Large improvements with respect to the unshielded VDMOSFET have been observed.

A novel device concept has been presented, the GP-VDMOSFET. The effect of p-region and shield plate designs on the device performance has been compared. It has been shown that the p-region permits larger reductions of the feedback capacitance than the shield plate. An RF based optimisation of the GP_VDMOSFET has been carried out. The possible occurrence of p-region punch through has been taken into account. P-region punch through has been prevented by opportunely adjusting the p-region doping, thickness and length. The GP_VDMOSFET has been shown to represent a viable alternative to the shielded VDMOSFET.

References

[1] Basavana Goud C., Bhat K.N., Two-dimensional Analysis and Design Considerations of High-Voltage Planar Junctions Equipped with Field Plate and Guard Ring, IEEE Transactions on Electron Devices, 1991, pp. 1497-1504

[2] Brieger K.-P., Falck E., Gerlach W., The Contour of an Optimal Field Plate – An Analytical Approach, IEEE Transactions on Electron Devices, 1988, pp. 684-688

[3] Chung S.K., Yoo D.C., Choi Y.I., An Analytical Method for Two-Dimensional Field Distribution of a MOS Structure with a Finite Field Plate, IEEE Transactions on Electron Devices, 1995, pp. 192-195

[4] Ishikawa O., Esaky H., A High-Power High-Gain VD-MOSFET Operating at 900MHz, IEEE Transactions on Electron Devices, 1987, pp. 1157-1162

[5] A. Wood, C. Dragon and Wayne Burger, High performance silicon LDMOS technology for 2GHz RF power amplifier application, IEEE International Electron Devices Meeting, pp.87-90, 1996.

[6] H. F. F. Jos, Novel LDMOS Structure for 2 GHz High Power base station applications", Proceedings of the European Microwave Conference, 1998.

[7] H.Brech, W.Brakensiek, D. Burdeaux, W. Burger, C. Dragon, G. Formicone, B. Pryor, D. Rice, Record Efficiency and Gain at 2.1GHz of High Power RF Transistors for Cellular and 3G Base Stations, IEEE IEDM 2003

[8] Nenadovic N., Cuoco V., Theeuwen S.J.C.H., Schellevis H., Spierings G., Griffo A., Pelk M., Nanver L.K., Jos R.F.F., Slotboom J.W., RF Power Silicon-On-Glass VDMOSFETs, Electron Device Letters, 2004, vol. 25, pp. 424–426

[9] Shuming Xu, Changhong Ren, Pang-Dow Foo, Yong Liu, Yi Su, Dummy Gated Radio Frequency VDMOSFET with High Breakdown Voltage and Low Feedback Capacitance, ISPSD, 2001, pp. 385-388

[10] Shuming Xu, Changhong Ren, Liang Y.C., Pan-Dow Foo, Sin J.K.O., Theoretical Analysis and Experimental Characterization of the Dummy-Gated VDMOSFET, IEEE Transactions on Electron Devices, 2001, pp. 2168-2176

[11] Medici User Manual, Technology Modelling Associates, Sunnyvale, California, USA.

[12] Atlas User Manual, Silvaco International, Santa Clara, California, USA.

[13] Trivedi, P. Khandelwal, K. Shenai, Performance Modeling of RF Power MOSFET's, IEEE Transactions on Electron Devices, 1999, vol. 46, pp. 1794-1801

[14] S. Cripps, "A Method for the Prediction of Load-Pull Contours in GaAs MESFETs", 1983 IEEE MTT-S International Microwave Symposium. Digest, June 1983, pp. 221-223.

[15] S.C. Cripps, RF Power Amplifiers for Wireless Communications, Artech House, 1999.

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

[17] H. Masuda, T. Masuhara, M. Nagata, N. Hashimote, Device design of E/D gate MOSFET, J. Japan Soc. Appl. Phys., 1973, vol.42, pp.167-172

[18] T.J. Rogers, S. Asai, M.D. Pocha, R.W. Dutton, J.D. Meindl, An experimental and theorethical analysis of double-diffused MOS transistors, IEEE J. Solid-State Circuits, 1975, vol. 10, pp.322-331

[19] A. Ortiz-Conde, F.J. García Sánchez, J.J. Liou, A. Cerdeira, M. Estrada, and Y.Yue, "A review of recent MOSFET threshold voltage extraction methods," Microelectronics Reliability, vol. 42, pp.583-596, 2002.

6

Package Thermal Effect

Abstract

A rigorous modelling procedure to estimate the heat extraction of a packaged VDMOSFET, taking into account the thermal resistance at both device and package level has been developed.

Simulation results reveal that a significant reduction in the thermal resistance can be realised in a conventional power VDMOSFET by heat extraction from the top side of the device (through source and gate) rather than through the conventional bottom drain.

A novel package topology for heat extraction from the top side is described. Device simulations suggest that the new package makes wafer thinning less significant for improvement of thermal resistance.

6.1 Introduction

Si RF Power MOSFETs are generally available as discrete semiconductors in three-terminal packages. In RF Power applications, due to the high frequency involved, the effect of the package can substantially degrade the electrical performance of the device.

For such a reason, Power RF packages are more complex than other discrete packages. The package header typically requires a ceramic layer to provide isolation between the package mounting base and the leads. The package and leads are made of copper, normally covered in gold metallization. The leads have to be considered as a transmission line at high frequency. The die is attached to the header with gold eutectic. Wire bonds, also in gold, are used to connect the top-side die contacts to the package leads. Shape and length of the wire bond wires, insertion angles and relative positions of the strip lines are chosen to minimise package parasitics, which can be considerably large due to the high frequency of application.

The parasitics and the heat extraction properties of a particular package noticeably affect the RF performance. Insufficient heat extraction can significantly reduce carrier mobility and even compromise package integrity. Parasitic source inductance can considerably reduce power gain at high frequency.

It is the packaged device performance that attracts customers; therefore it is important to predict the effect of the package at the early stages of a design. In Chapter 4 it has been shown that reasonably accurate predictions of RF performance can be achieved by inclusion of the package dependent source impedance in the power gain calculation. On the other hand, prediction of the effect of the package on the device internal temperature distribution and of the effect of temperature on the device performance is a complex issue. For such a task, 3-dimensional finite element analysis of packaged devices is desirable. However, this approach is impractical due to high computational resource requirements. Consequently, thermal evaluation is usually carried out either at the device level, ignoring the package contribution, or at the package level,

considering a uniform temperature distribution within the device. A combination of simultaneous device and package simulations remains inaccessible to date [1]. In this chapter, a unified thermal modelling procedure is proposed for estimating the total thermal resistance of a packaged device. The procedure is described in section 1.2. In section 6.3 the approach is used to compare the effectiveness of heat extraction from the top side of a recessed gate VDMOSFET, through source and gate, with the conventional approach of heat extraction from the bottom-drain side. The description of the method and the concept of source side heat extraction also appear in [2]. Finally, in section 6.4 some considerations of the package effect on the performance of RF power Si MOSFETs are discussed.

6.2 Package thermal effect

This section describes the integration of an approximated analytical method into electro-thermal simulations of semiconductor devices. The method permits the determination of thermal capacitance and resistance, allowing for packaged device characterisation.

6.2.1 Background

Thermal extraction in power devices takes place mainly through conduction. Heat is generated by self heating inside the die and is transferred out of the die through the device metal routing layers, the soldering-bonding layers and the package base. In package design, simplified approaches are often used to provide estimates of the package thermal resistance and capacitance [3]. Complicated package heat flow patterns can always be converted in equivalent thermal circuit diagrams. The analogy between the diffusion of heat and the transport of electrical charge permits the resolution of these circuits with classic circuit analysis methodologies. This yields high level insight, providing a link between design parameters and

thermal performance. Also in [3], these approaches are shown to produce adequate prediction of thermal performance in simple packaging configurations.

Thermal resistance and capacitance are measures of the capability of a given structure for transferring and accumulating heat respectively. The general definition of thermal resistance is the ratio between a variation of temperature ΔT and the associated heat flow P_{th}

$$R_{th} = \frac{\Delta T}{P_{th}}, \tag{6.1}$$

where the heat flow is the variation of heat stored with time $P_{th}=Q/\Delta t$.

The thermal capacitance is the ratio between the accumulated heat Q and the associated temperature variation

$$C_{th} = \frac{Q}{\Delta T}. \tag{6.2}$$

When the heat flow can be approximated as one dimensional, as depicted in Figure 6.1, simple geometry dependent definitions of thermal resistance and capacitance result in

$$R_{th} = \frac{h}{k \cdot A}, \tag{6.3}$$

$$C_{th} = c \cdot d \cdot V = c \cdot d \cdot h \cdot A, \tag{6.4}$$

where d, c, k respectively are the density, the specific heat and the thermal conductivity of the heat flow medium and h, A, V correspond to the height (in the direction of the heat flow), the area, and the volume (V=Ah) of the element involved in the heat conduction.

A more complex situation encountered in semiconductor packages is represented by the heat flowing from a small source of area A_{heat} to a large sink. In this case the heat flow cannot be approximated as one dimensional and Equations 6.3 and 6.4 fail. The heat flow takes place in a three dimensional fashion. In order to determine equivalent thermal resistance and capacitance, a spreading angle approximation is usually carried out in this case. The spreading angle α, as shown in Figure 6.2, is a measure of the deviation of the heat flow from a one-dimensional behaviour.

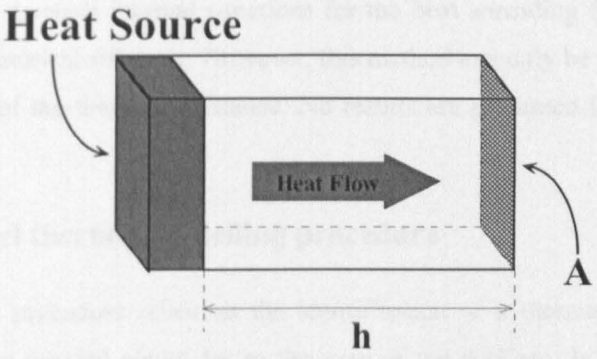


Figure 6.1: One dimensional approximation of the flow channel involved in the heat conduction.

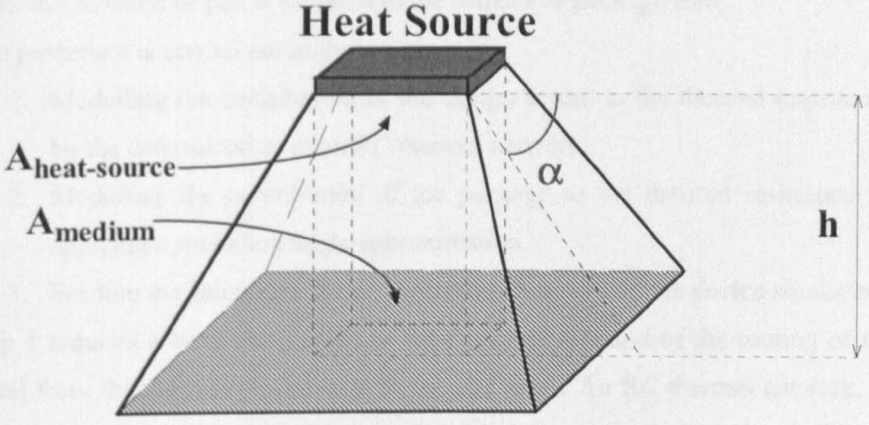


Figure 6.2: Truncated-pyramid approximation of the flow channel involved in the heat conduction.

The analytical approach derived from the spreading angle approximation is based on the identification of the final area A_{medium} of the medium involved in the heat conduction on the side opposite to the heat source. This area can be easily determined geometrically. An equivalent area is then calculated permitting the application of Equations 6.3 and 6.4. The equivalent area A_{eq} is determined as the average between A_{heat} and A_{medium} or is calculated from the average dimensions of the heat sources. In [4] it has been shown through Finite Elements Analysis that the spreading angle approximation for a spreading angle of 45° yields an average 11% error on the estimation of the thermal resistance.

If required, improved accuracy can be achieved with the graphical method described in [5] and summarized in [6]. The method is developed from the

solution of the Laplace thermal equations for the heat spreading from a circular source to a cylindrical substrate. However, this method can only be applied for the determination of the thermal resistance. No results are presented for the thermal capacitance.

6.2.2 A novel thermal modelling procedure

The modelling procedure relies on the identification of a thermal RC network from the device thermal electrodes to the case of the package. In this study, an electrode is addressed as thermal whenever it belongs to the thermal extraction path; that is when its pad is soldered to the respective package lead.

The procedure is carried out in three main steps:

1. Modelling the contribution of the design layout to the thermal impedance by the determination of an RC thermal network.
2. Modelling the contribution of the package to the thermal resistance by applying a spreading angle approximation.
3. Feeding the calculated thermal resistance values into the device simulator.

Step 1 requires a basic description of the device layout and of the routing of the metal from the thermal electrode(s) to the pad areas. An RC thermal network, as shown in Figure 6.3, can be determined following an approach similar to [7]. In [7] an equivalent thermal network has been used to model a staggered via network configuration of a thin film multi-chip module, which closely resembles a multi-level metal routing in a device layout. Oxides and metal connections are considered as parallel thermal resistances which serially connect different metal routing levels. The thermal resistance and capacitance of each parallel element can be calculated as in Equations 6.3 and 6.4.

In step 2 the thermal model is extended to include the package. The repetitive application of the spreading angle rule yields the thermal resistance and capacitance values from the device pad(s) to the package case as a serial contribution of the package layers.

Thermal resistance and capacitance values are calculated from of the different layers using Equations 6.3 and 6.4, where the element section A is replaced by the equivalent element area A_{eq} determined as

$$L_{medium} = L_{heat} + 2\tan(\alpha)h, \tag{6.5}$$

$$W_{medium} = W_{heat} + 2\tan(\alpha)h, \tag{6.6}$$

$$A_{eq} = \sqrt{L_{heat}W_{heat}} \cdot \sqrt{L_{source}W_{source}}, \tag{6.7}$$

where h is thickness of the medium, L and W represent the rectangular dimensions of the surfaces involved in the heat conduction. The suffixes heat and source indicate for each layer the initial and final areas along the thermal path.

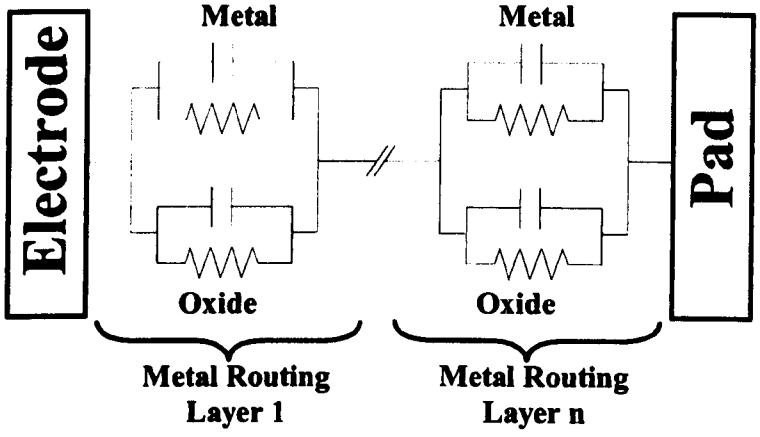


Figure 6.3: Step 1 thermal network used in the thermal modelling.

In step 3, the total thermal impedance can be calculated by solving the equivalent thermal network determined in steps 1 and 2. For inclusion in the device simulator [8] in the form of a lumped element contacting the device thermal electrode(s), the calculated thermal resistance is multiplied by the equivalent device length; this is because device simulators usually assume a reference width of 1μm for any considered device. A fully coupled electro-thermal simulation can now be carried out taking into account the thermal effect of the package on the device performance.

6.2.3 Results

Device and package description

The schematic cross-section of the device considered in this section is reported in Figure 6.4. The device is a recessed gate VDMOSFET displaying a maximum current of 75A and a 100V breakdown voltage. A die size of 7.5mm^2 has been estimated as appropriate for the given power rating. The die size corresponds to the active area of the device.

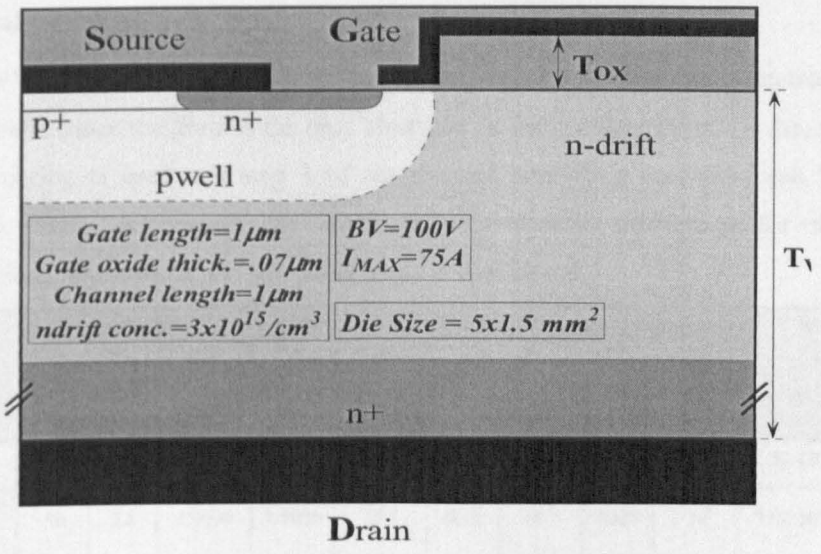


Figure 6.4: Recessed gate power MOSFET cross section.

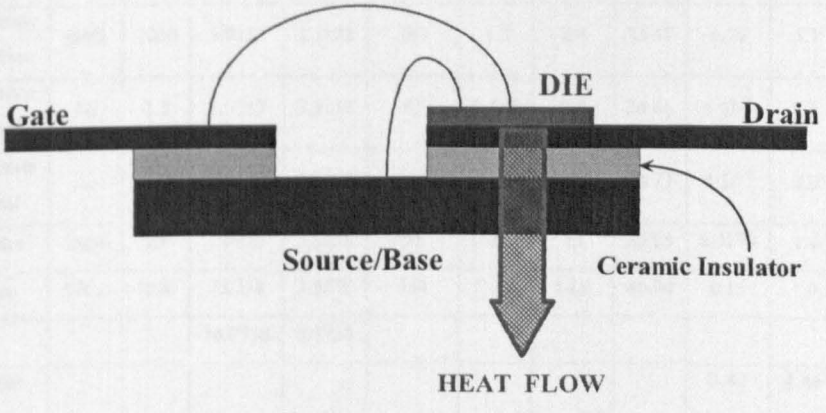


Figure 6.5: Conventional internally isolated VDMOS package.

Figure 6.5 represents a conventional internally isolated package layout of a VDMOSFET. The drain pad is directly soldered to the package drain lead, while gate and source are wire-bonded. An isolation layer exists in the thermal extraction path between the drain lead and the base of the package. The properties of the drain to source isolation in the package are crucial to guarantee optimal performance: the insulator layer has to electrically isolate the drain from the source while thermally conducting the heat from the drain to the source/base of the package.

Thermal modelling

When a VDMOSFET is placed in the conventional package, the thermal electrode is the drain. Since the drain is the only electrode on the bottom side of the die, no metal routing is used and step 1 of the thermal modelling procedure can be skipped. Table 6.1 shows the results of spreadsheet calculation for the conventional package. A 45° spreading angle is considered.

Material		h	L	W	k	h	d	A_{eq}	R_{th}	Cth
		µm	mm	mm	W/mK	J/gK	g/cm^3	mm^2	K/W	J/K
Pad	Au	0.3	5	1.5	317	0.13	18.9	7.501	$1\cdot10^{-3}$	$5.52\cdot10^{-6}$
Substrate metal	Au	3.8	5.0006	1.5006	317	0.13	18.9	7.528	$1\cdot10^{-3}$	$7.02\cdot10^{-5}$
Substrate metal	Ni	2.5	5.0082	1.5082	91	0.445	8.9	7.569	$3\cdot10^{-3}$	$7.49\cdot10^{-5}$
Substrate isolation	BeO	1000	5.0132	1.5132	281	1.3	2.9	13.67	0.26	$5.1\cdot10^{-2}$
Substrate metal	Ni	2.5	7.0132	3.5132	91	0.445	8.9	24.66	$1\cdot10^{-3}$	$2.5\cdot10^{-4}$
Substrate metal	Au	3.8	7.0182	3.5182	317	0.13	18.9	24.73	$4\cdot10^{-4}$	$2.3\cdot10^{-4}$
Solder	AuSn	25	7.0258	3.5258	57	0.015	15	25.03	0.0179	$1.4\cdot10^{-4}$
Base	WCu	1500	7.0758	3.5758	234	0.21	14.9	40.94	0.15	0.19
			10.0758	6.5758						
Total									**0.44**	**$4.46\cdot10^{-6}$**

Table 6.1: Calculation of the package thermal resistance and capacitance.

The total thermal resistance and capacitance are calculated as the series of the thermal resistances and capacitances of the layers respectively, resulting in $R_{th,pack}$=0.44K/W and $C_{th,pack}$=4.46·10^{-6}J/K. For inclusion in the device simulator, these values are scaled considering the total device width W=1m, resulting in $R_{th,scaled}$=44·10^{4}K·μm/W and $C_{th,scaled}$=4.46·10^{-12}J/(μm·W). The calculated thermal capacitance and resistance are then included in the device simulator as lumped elements connected to the thermal electrode of the device.

Electro-thermal packaged device simulations

Electro-thermal simulations including the package effect have been carried out to extract the device thermal resistance, to identify the location of the device hot spot and to assess the influence of wafer thickness (T_{WAF}) and recessed gate oxide thickness (T_{OX}) on the thermal resistance. The extraction of the total packaged device thermal resistance $R_{th,total}$ is demonstrated in Figure 6.6. The total thermal resistance value is extracted as the slope of the maximum die temperature versus dissipated power curve, resulting in $R_{th,total}$=0.8K/W for a wafer thickness of 150μm. By repeating the $R_{th,total}$ for devices with several T_{OX} values, it has been verified that the recessed gate oxide thickness does not have any effect on the total device thermal resistance. This result is however expected, since the gate oxide does not belong to the heat extraction path.

Figure 6.7 shows the existence of a direct proportionality between wafer thickness and thermal resistance. Thermal resistance decreases with the wafer thickness. The minimum $R_{th,total}$=0.6K/W corresponds to a considered wafer thickness of 50μm. The linear relation between T_{WAF} and $R_{th,total}$ in Figure 6.7 permits to extrapolate an ideal lower bound for the thermal resistance reduction that can be achieved by wafer thinning, $R_{th,limit}$=0.48K/W. This value corresponds to a wafer in which the substrate has been completely removed, leaving only the epi-region. The ideal limit depends on both package and device and corresponds to the serial contribution of package and epi-region.

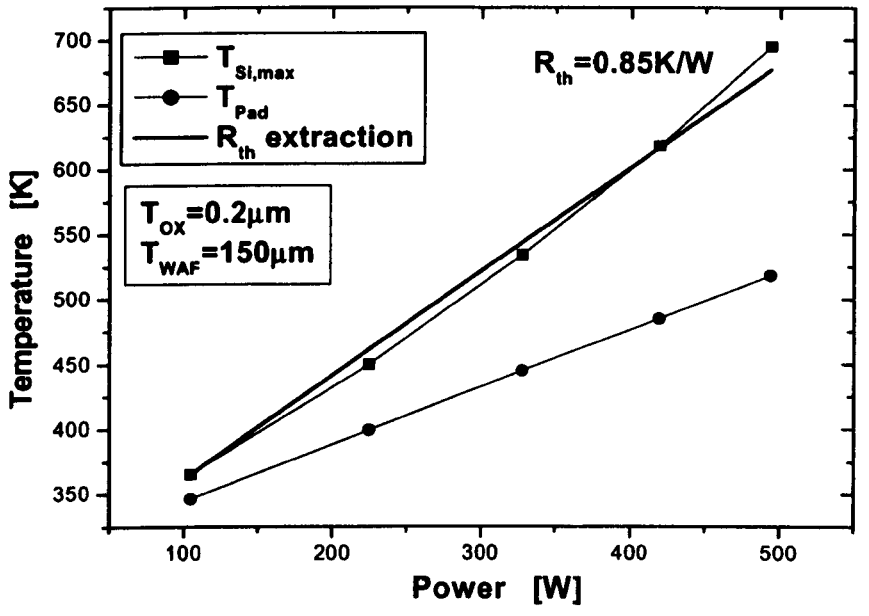


Figure 6.6: Simulated maximum device temperature and package pad temperature versus DC output power. The total device thermal resistance is extracted as the linear fit slope of the maximum device temperature curve.

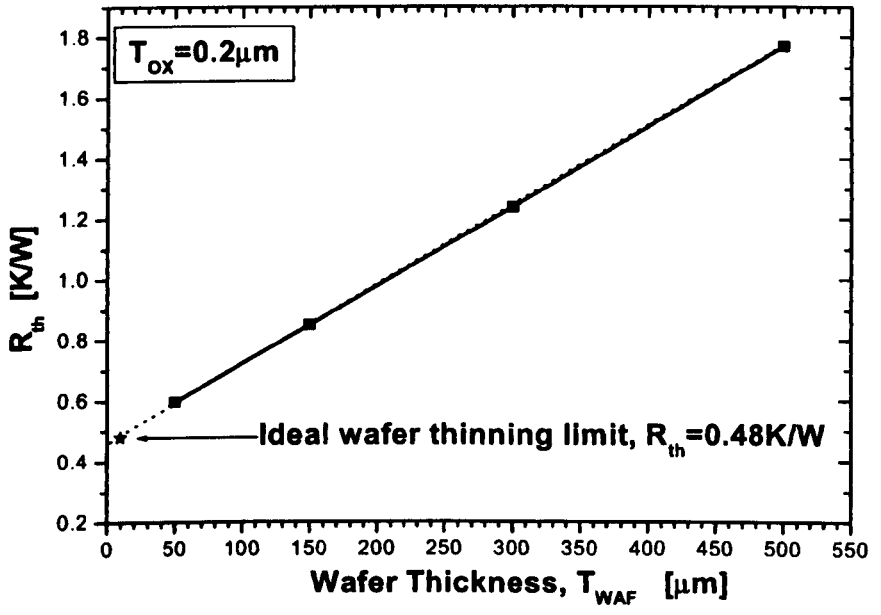


Figure 6.7: Thermal resistance dependence on wafer thickness as extracted from electro-thermal simulation results.

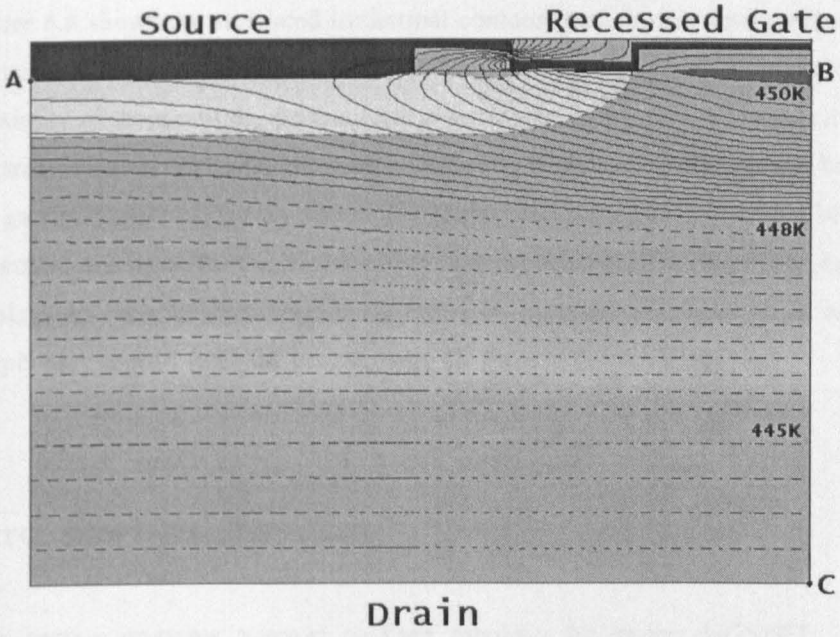


Figure 6.8: Bottom side heat extraction: 2-D Medici cross-section detail at V_{DS}=10V and V_{GS}=5V of a 50μm thick VDMOSFET with isothermal contour lines.

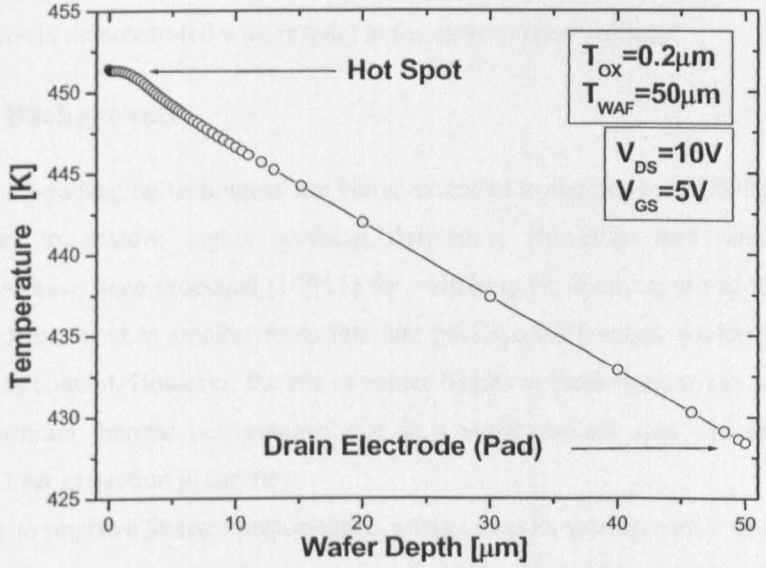


Figure 6.9: Simulated temperature distribution in the device as at V_{DS}=10V and V_{GS}=5V along the line B-C of Figure 6.8.

Figure 6.8 shows the simulated isothermal contour lines for the device biased in the on state when the package thermal effect is included. The hot spot is placed in proximity of the top of the device. Heat is generated in this region because of the electric resistance caused by the current crowding in the accumulation layer below the gate and by the JFET resistance due to the cell spacing [9]. Since the heat is generated mainly at the top of the device and extracted at the bottom, the device displays an internal temperature gradient, as reported in Figure 6.9, where temperature reduces with the proximity of the thermal electrode, the drain.

6.3 Source side heat extraction

This section proposes a novel package topology for power MOSFETs. The package design concept relies on a heat extraction that takes place from the gate and source side, close to the device hot spot. Wafer thinning approaches become unnecessary for the improvement of thermal performance. Improved thermal resistance is demonstrated with respect to the conventional package.

6.3.1 Background

Lately, IC-packaging techniques are being extended to the power MOSFET case. In order to achieve higher working frequency, chip-scale and wafer-scale packages have been proposed [10][11] for switching PC-board applications. This approach has lead to smaller footprints and package thicknesses, yielding lower parasitics content. However, the use of solder bumps in these designs has lead to a non optimum thermal performance due to a small contact area and therefore limited heat extraction properties.

In order to improve thermal performance, a large contact area approach and a two-sided cooling have also been presented [12]. This approach has lead a considerable reduction of the thermal resistance but at the cost of a die size increased well above the actual active device area. Enlarged pad sizes have been

required to guarantee reduced thermal resistance and to facilitate surface mounting of these devices directly to the PC-board.

On the other hand, similar progress in packaging of larger power MOSFETs has not taken place yet. Although the device hot spot is situated in proximity of the top of the die, thermal extraction typically occurs through the bottom. The entire device die belongs to the thermal path. Consequently, improved thermal performance is normally achieved through a reduction of the thermal path inside the die achieved by wafer thinning approaches. Final device thicknesses as low as 40μm have been reported [13]. This kind of thermal performance improvements is however associated to increased difficulties in wafer handling and device packaging stages due to increased wafer fragility.

6.3.2 Novel Package and Device Layouts

A package layout suitable for heat extraction from the source side is represented in Figure 6.10. The device top side is bonded to the package header.

In the novel package, large bonding areas permit good thermal extraction through both gate and source pads. The heat can flow out of the source pad directly into the base of the package through the soldering layers, whereas coming out of the gate pad it encounters a layer of ceramic insulator. This insulating layer has to electrically isolate gate and source metals in the package while conducting heat out of the device.

A suitable device layout for the package described in Figure 6.10 is reported in Figure 6.11. The device considered is the recessed gate power MOSFET described in section 1.2.

The device layout has a source pad which is bigger than the gate. This is due to the package presenting an insulation layer in the heat extraction path of the gate, which makes the source extraction path preferable to the gate extraction path. More explanations of this concept are provided in section 6.3.3.

The metal routing is articulated in three layers and their two interconnections. The three layers are represented by the device electrodes (layer 0), the connections between correspondent electrodes (layer 1) and the pad layer (layer 2).

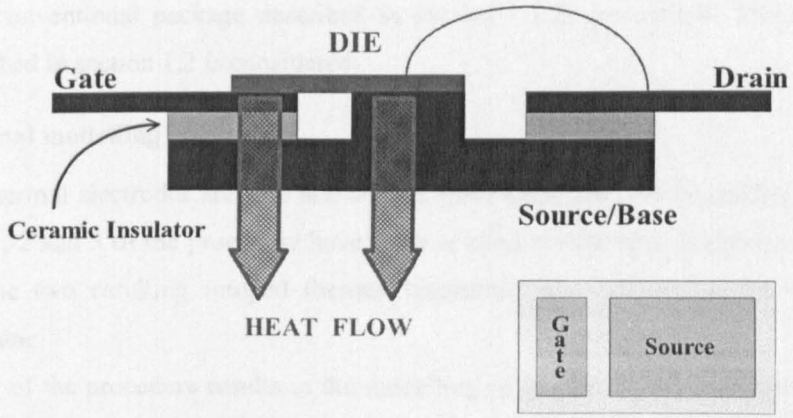


Figure 6.10: VDMOS die and package for Top Side Heat Extraction.

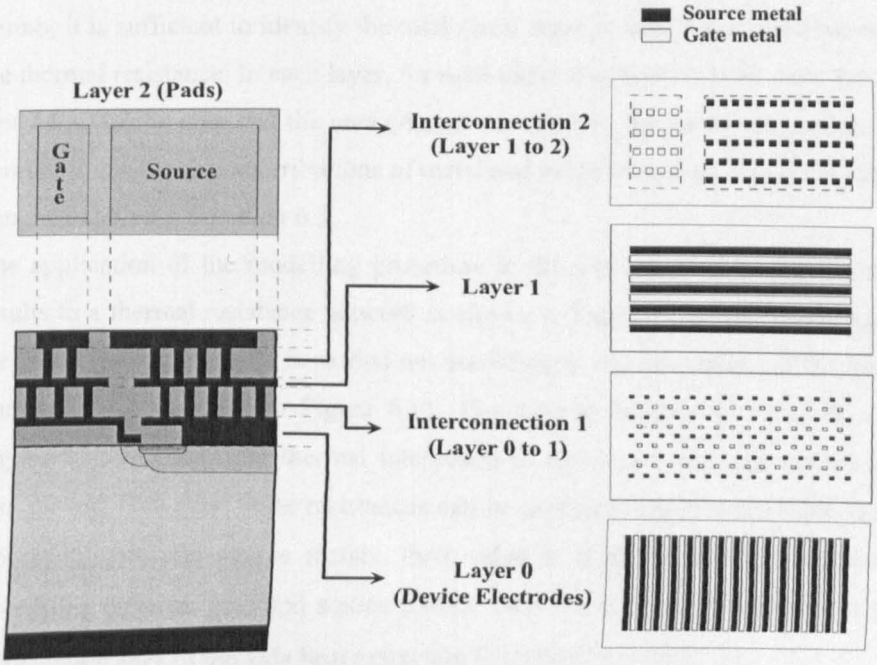


Figure 6.11: Metal layers layout in the Top Side Heat Extraction case. Source and gate electrodes (Layer 0) are connected (Layer 1) before being routed to the respective pads (Layer2).

6.3.3 Results

In this section the effect of the proposed package on the total thermal impedance is assessed. Also, a comparison between the performance of the proposed package

with conventional package described in section 1.2 is carried out. The device described in section 1.2 is considered.

Thermal modelling

The thermal electrodes are gate and source; there exist two heat extraction paths. Step 1, 2 and 3 of the procedure have to be applied to both heat extraction paths, and the two resulting lumped thermal resistances are included in the device simulator.

Step 1 of the procedure results in the modelling of the five layers comprising the metal routing described in Figure 6.11. It is clear that the more detailed the information on the layout, the more accurate the final result. However, as a rule of thumb, it is sufficient to identify the total metal areas in each layer and determine the thermal resistance. In each layer, for each thermal extraction path, once the die area (A_{die}) of the chip and the area (A_{metal}) occupied by the metal are known, the parallel of the thermal contributions of metal and oxide (area A_{die}-A_{metal}) are easily obtained applying Equation 6.3.

The application of the modelling procedure to the top side heat extraction case results in a thermal resistance network as shown in Figure 6.12; the modelling of the metal routing in step 1 is carried out according to the description of the metal routing layout provided in Figure 6.11. The thermal resistances R_1 to R_{n+1} of Figure 6.12 represent the thermal interaction in each layer between source and gate metals. However, these resistances can be neglected. Due to the thick oxide separating gate and source metals, their value is high and implies a thermal decoupling between gate and source metals. Step 2 can proceed similarly to the conventional case of top side heat extraction described in section 1.2.

The thermal capacitance and resistance calculations for the source and gate thermal extraction paths are reported in Table 6.2 to Table 6.9. Table 6.2 and Table 6.3 describe the calculations for the parallel contribution to the thermal resistance from the source and gate electrodes through two metal levels.

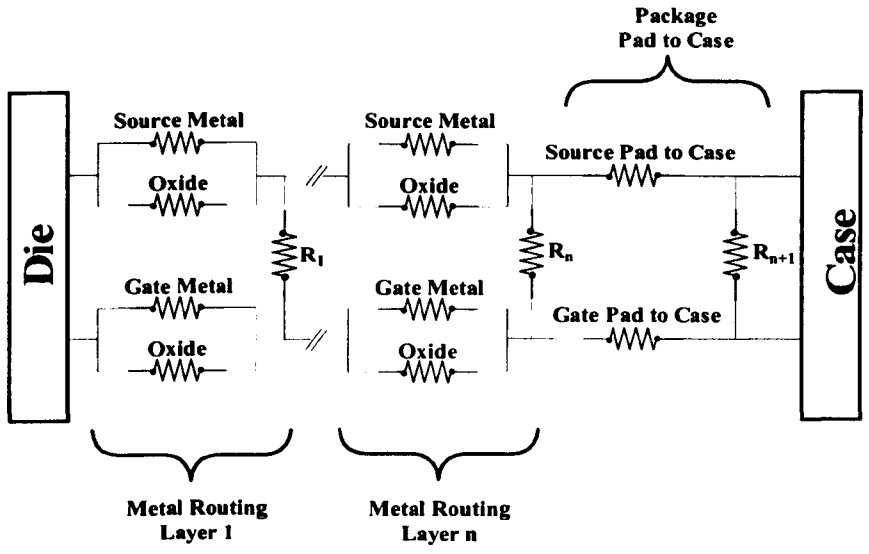


Figure 6.12: Thermal resistance equivalent model in the top-side heat extraction case.

Calculations are carried out on the basis that the Aluminium occupies just 20% of the die area in layer 0/1 and 40% on layer 1. Due to the symmetrical structure of the routing layers, the oxide contribution to the thermal resistance is symmetrically divided between source and gate thermal extraction path. This yields to the consideration of an oxide area which is 30% of the total die size in layer 0/1 and 10% in layer 2, for both the thermal paths. In Table 6.4-Table 6.5 and Table 6.7-Table 6.8, layer 1/2 and layer 2 calculations are reported for the source and gate extraction paths respectively. The reference area is the associated pad area. Finally, Table 6.6 and Table 6.9 describe the step 2 calculation associated with the package layer contributions. A 45° spreading angle is considered.

The resolution of the thermal resistance equivalent model yields a value of the thermal resistance from $R_{th,source}$=0.51K/W for the source and $R_{th,gate}$=4.86K/W for the gate thermal extraction paths. These values are mainly dependent on the heat path through the package; the electrodes to pad thermal resistances are only about 0.011K/W and 0.082K/W respectively for the source and gate case. Approximating the total package thermal resistance as the parallel of the contributions of the two heat extraction paths, a final value of $R_{th,pack}$=0.46K/W is

obtained. This value corresponds to a gate pad that presents the minimum allowed size.

Material		h	L	W	k	h	d	A	R_{th}	Cth
		µm	mm	mm	W/mK	J/gK	g/cm^3	%	K/W	J/K
Metal	Al	1	5	1.5	170	0.896	2.7	20	0.003	$1.8 \cdot 10^{-6}$
Oxide	SiO_2	1	5	1.5	1.4	1	2.27	30	0.31	$2.5 \cdot 10^{-6}$
Total									0.003	$4.3 \cdot 10^{-6}$

Table 6.2: Source and Gate extraction path, step 1: layer 0/1 thermal impedance calculation.

Material		h	L	W	k	h	d	A	R_{th}	Cth
		µm	mm	mm	W/mK	J/gK	g/cm^3	%	K/W	J/K
Metal	Al	1	5	1.5	170	0.896	2.7	40	0.001	$3.6 \cdot 10^{-6}$
Oxide	SiO_2	1	5	1.5	1.4	1	2.27	10	0.952	$7.6 \cdot 10^{-7}$
Total									$1.9 \cdot 10^{-4}$	$4.3 \cdot 10^{-6}$

Table 6.3: Source extraction path, step 1: layer 1 thermal impedance calculation.

Material		h	L	W	k	h	d	A	R_{th}	Cth
		µm	mm	mm	W/mK	J/gK	g/cm^3	%	K/W	J/K
Metal	Al	1	4.2	1.5	170	0.896	2.7	20	0.004	$1.5 \cdot 10^{-6}$
Oxide	SiO_2	1	4.2	1.5	1.4	1	2.27	80	0.141	$5.7 \cdot 10^{-6}$
Total									$4.52 \cdot 10^{-4}$	$7.2 \cdot 10^{-6}$

Table 6.4: Source extraction path, step 1: layer 1/2 thermal impedance calculation.

Material		h	L	W	k	h	d	A	R_{th}	Cth
		µm	mm	mm	W/mK	J/gK	g/cm^3	mm^2	K/W	J/K
Metal	Al	1	4.2	1.5	170	0.896	2.7	6.8	$9.4 \cdot 10^{-4}$	$7.6 \cdot 10^{-6}$
Total									$9.4 \cdot 10^{-4}$	$7.6 \cdot 10^{-6}$

Table 6.5: Source extraction path, step 1: layer 2 thermal impedance calculation.

Material		h	L	W	k	h	d	A_{eq}	R_{th}	Cth
		µm	mm	mm	W/mK	J/gK	g/cm^3	mm^2	K/W	J/K
Step1			4.2	1.5					0.011	$1.3\cdot10^{-6}$
Metal	Au	3.8	4.2	1.5	317	0.13	18.9	0.65	0.001	$2.3\cdot10^{-5}$
Solder	AuSn	25	4.2076	1.5076	57	0.015	15	0.70	0.067	$2.2\cdot10^{-6}$
Base	WCu	1500	4.2576	1.5576	234	0.21	14.9	3.425	0.43	$7.2\cdot10^{-5}$
			7.2576	4.5576						
Total									0.511	$8.0\cdot10^{-7}$

Table 6.6: Source extraction path, step 2: layer 2 thermal impedance calculation.

Material		h	L	W	k	h	d	A	R_{th}	Cth
		µm	mm	mm	W/mK	J/gK	g/cm^3	%	K/W	J/K
Metal	Al	1	0.3	1.5	170	0.896	2.7	5	0.065	$1.0\cdot10^{-7}$
Oxide	SiO_2	1	0.3	1.5	1.4	1	2.27	40	1.984	$4.0\cdot10^{-7}$
Total									0.063	$5.1\cdot10^{-7}$

Table 6.7: Gate extraction path, step 1: layer 1/2 thermal impedance calculation.

Material		h	L	W	k	h	d	A	R_{th}	Cth
		µm	mm	mm	W/mK	J/gK	g/cm^3	mm^2	K/W	J/K
Metal	Al	1	0.3	1.5	170	0.896	2.7	1.5	0.013	$5.1\cdot10^{-7}$
Total									0.013	$5.1\cdot10^{-7}$

Table 6.8: Gate extraction path, step 1: layer 2 thermal impedance calculation.

Optimum sizing of the gate pad is chosen from the gate size effect on the total package thermal resistance. Figure 6.13 shows that increased gate pad sizes cause a decrease of $R_{th,gate}$ but an increase of $R_{th,source}$. The minimum $R_{th,pack}$ is obtained for the smallest considered pad size length of 0.3mm. The gate heat extraction path can be thought of as a secondary heat path, since it is associated with a higher thermal resistance.

Material		h	L	W	k	h	d	A	R_{th}	Cth
		µm	mm	mm	W/mK	J/gK	g/cm^3	mm^2	K/W	J/K
Step1			0.3	1.5					0.082	$2.3 \cdot 10^{-7}$
Metal	Au	3.8	0.3	1.5	317	0.13	18.9	0.758	0.0261	$1.6 \cdot 10^{-6}$
Solder	AuSn	25	0.307	1.507	57	0.015	15	9.05	0.862	$1.7 \cdot 10^{-7}$
Lead	WCu	1400	0.357	1.557	234	0.21	14.9		0.940	$3.6 \cdot 10^{-6}$
Isolation	Al_2O_3	100	0.557	1.757	57	0.21	14.9	18.37	2.203	$7.2 \cdot 10^{-6}$
Base	WCu	1400	0.757	1.957	234	0.21	14.9		0.586	$2.9 \cdot 10^{-5}$
			3.357	4.557						
Total									4.701	$9.2 \cdot 10^{-8}$

Table 6.9: Gate extraction path, step 2: layer 2 thermal impedance calculation.

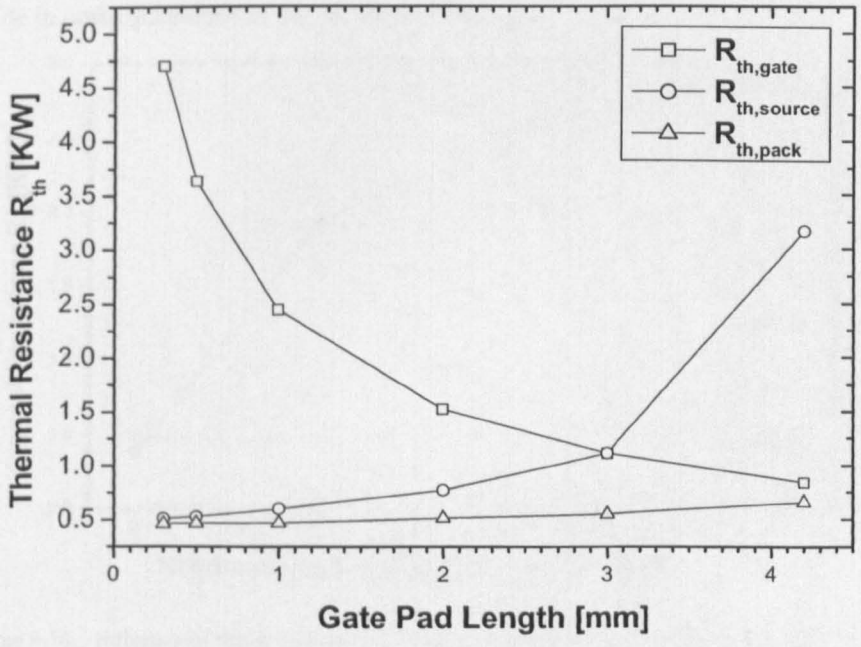


Figure 6.13: Influence of gate pad length on the thermal resistance of the gate ($R_{th,gate}$) and source ($R_{th,source}$) heat extraction paths. The total package thermal resistance ($R_{th,pack}$) is approximated as the parallel of $R_{th,source}$ and $R_{th,gate}$.

Electro-thermal packaged device simulations

Electro-thermal simulation results show that in the novel package the wafer thickness T_{WAF} does not produce any substantial effect on the total thermal resistance.

On the other hand, an increase of the thermal resistance can be shown to be associated for larger recessed gate oxide thicknesses. The gate oxide is in fact part of the gate heat extraction path, therefore its increase has a direct effect on $R_{th,gate}$ and to the total thermal resistance. The choice of a gate pad smaller than the source one, and therefore of the gate heat extraction path as secondary with respect to the source one, limits the effect of the thermal resistance variation associated with the increase of the gate oxide. As a result, Figure 6.14 shows that the thermal resistance becomes appreciatively independent of the thickness of the oxide in correspondence to the recession of the gate.

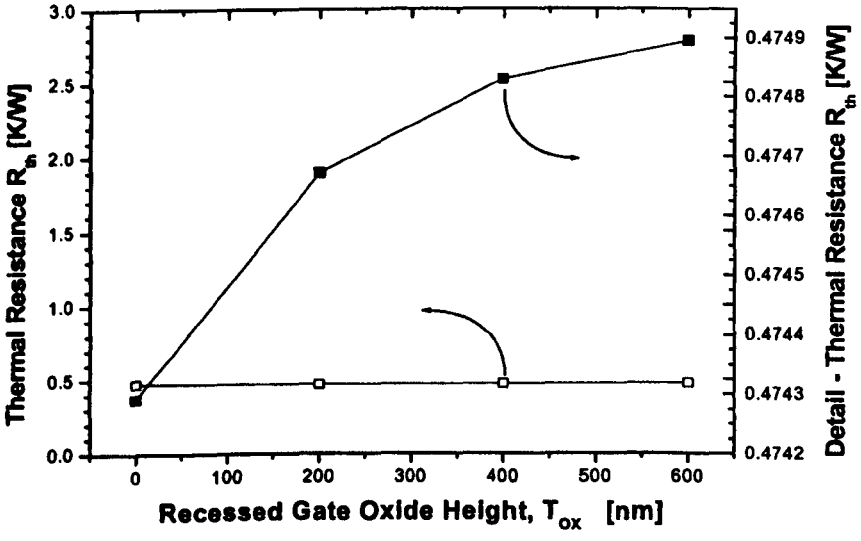


Figure 6.14: Influence of the recessed gate oxide thickness on the total thermal resistance.

The total thermal resistance value obtained in the novel package case, $R_{th,total}$=0.475K/W, is lower than in the conventional package case. In the top side case, the thermal resistance does not depend upon the device wafer thickness. The heat extraction from the top-source results in more efficiency than from the bottom-drain side even if the device wafer thickness is reduced. With respect to

the conventional package case and T_{WAF}=50μm, a 20% reduction of the thermal resistance is observed when the novel package is used. In the considered case, the total thermal resistance for the novel package is comparable to the conventional package ideal limit achievable by wafer thinning.

Figure 6.15shows the simulated isothermal contour lines for the device biased in the on state when the novel package effect is considered. The hot spot location is unchanged with respect to the conventional package case (Figure 6.8). The heat flows from the hot spot to gate and source electrodes without involving the entire epi layer. Since the metal routing provides a low thermal resistance path, the temperature on the device pads is maintained close to the maximum temperature in the silicon. The proximity of the thermal electrodes to the main heat generation point determines the device to operate in almost isothermal conditions, as also shown in Figure 6.16.

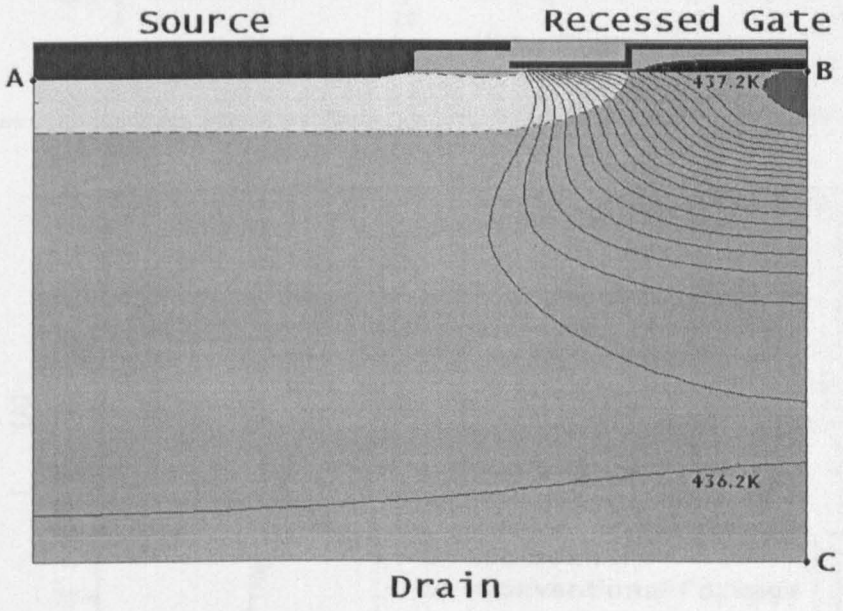


Figure 6.15: Top side heat extraction: 2-D Medici cross-section detail at V_{DS}=10V and V_{GS}=5V of a 50μm thick VDMOSFET with isothermal contour lines.

Figure 6.16 reports the temperature variation along the device thickness in both the considered package cases. A considerable temperature reduction is achieved at the top of the device when the novel package is used. At a device level, lower

temperature implies a less degraded mobility and an improvement of current capabilities. Accordingly, Figure 6.17 shows an improved transfer-characteristic for the novel package.

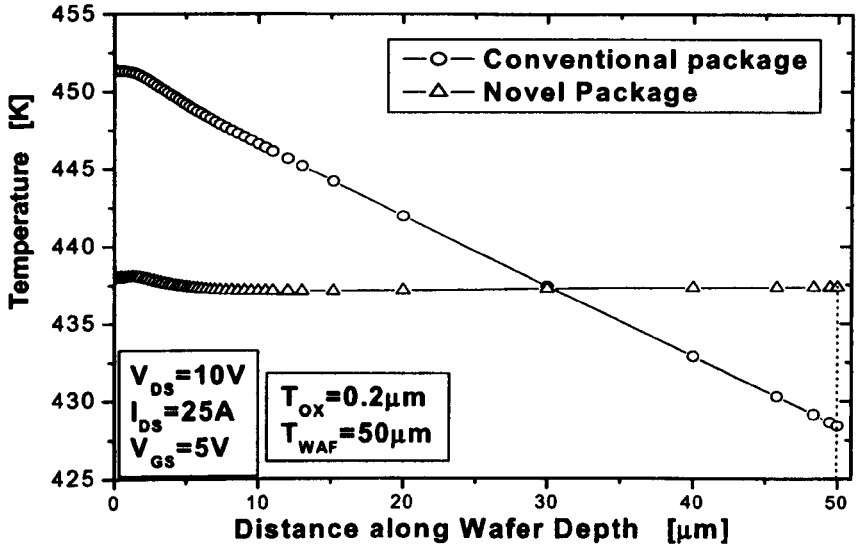


Figure 6.16: Simulated temperature distribution in the device as at V_{DS}=10V and V_{GS}=5V along the line B-C of Figure 6.8. and Figure 6.15.

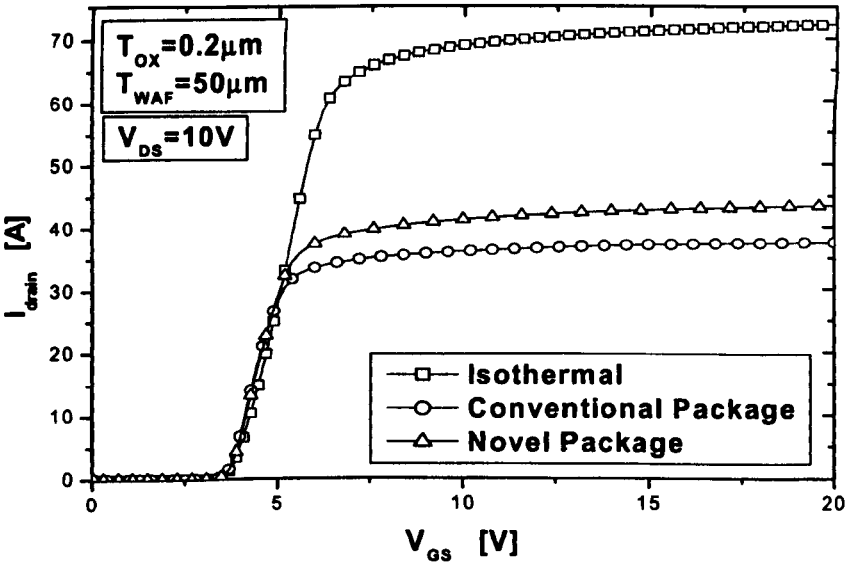


Figure 6.17: Simulated transfer-characteristics at V_{DS}=10V comparing the packaged devices presenting the lowest thermal resistances for bottom and top side heat extraction with the isothermal case.

6.4 Effect of top-side heat extraction in RF applications

This section presents some considerations on the package effect on the electro-thermal performance of RF power Si VDMOSFETs. The considered device is the optimum shield device identified in Chapter 5 and reported in Figure 6.18. The device width W is dimensioned to permit the generation of 20W of RF linear power, yielding W=21mm and a total active area of 0.75mm^2. The novel package topology proposed in section 6.3 is considered. The device layout for this package is shown schematically in Figure 6.19.

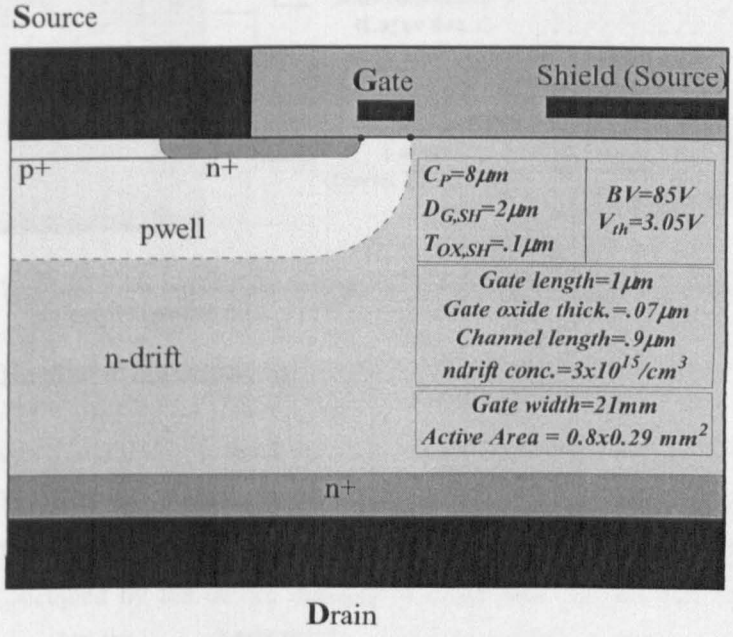


Figure 6.18: Shielded RF power VDMOSFET cross section.

The novel package is preferred to the conventional one mostly for the absence of a source bond wire and for the associated source inductance reduction which makes it more suitable for RF power applications. Another preference factor is that the use of the toxic and expensive Beryllium Oxide (BeO) is avoided. Adequate heat extraction can be achieved even if the thermally less conductive Alumina (Al_2O_3) is used.

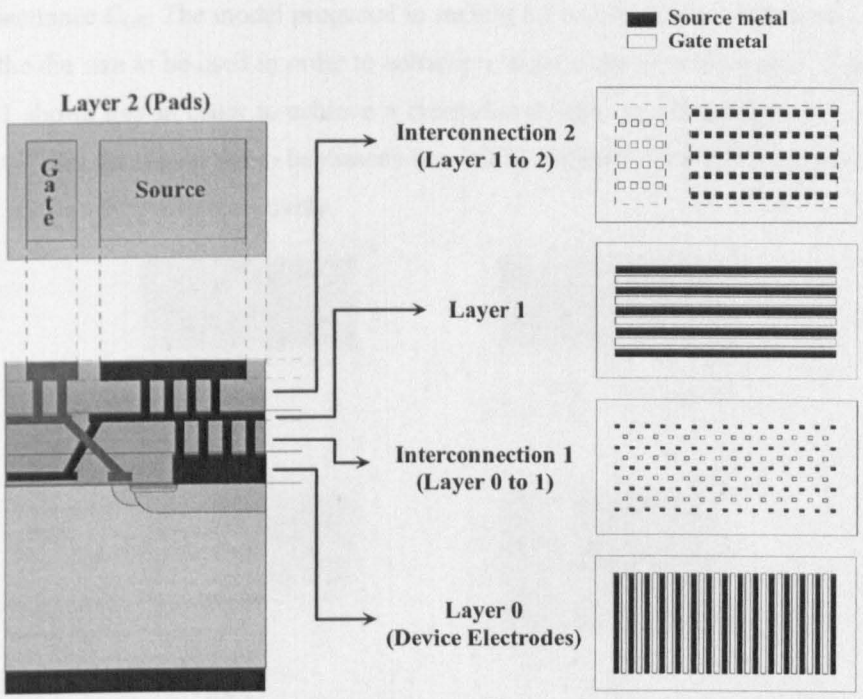


Figure 6.19: Metal layers layout of the Si RF power VDMOSFET of Figure 6.18 for the package proposed in section 6.3.

6.4.1 Die size dimensioning

Si RF power MOSFETs present active areas much smaller than normal power MOSFETs. Although dimensioned for the generation of 20W of RF power, the active area of the shielded VDMOSFET of Figure 6.18 is about one thirtieth of the area occupied by the device considered in sections 1.2 and 6.3. The small active areas of Si RF power MOSFETs makes heat extraction a serious problem.

In order to reduce the thermal impedance, the die size can be over-dimensioned with respect to the active device area, as shown in Figure 6.20. A larger contact area is used to facilitate heat extraction.

Figure 6.21 shows the effect of the over-dimensioning of the die size on the package thermal resistance and capacitance content. The die size increase corresponds to the increase of the source pad area only; the gate pad dimension is maintained at its minimum in order to improve heat extraction, as explained in section 6.3, but also to avoid excessive degradation of the gate to source

capacitance C_{GS}. The model proposed in section 1.2 can be used to obtain an idea of the die size to be used in order to achieve a target thermal performance. Figure 6.21 shows that in order to achieve a thermal resistance of value between 2 and 3K/W, the die size needs to be dimensioned to be between three and two times as big as the active area resectively.

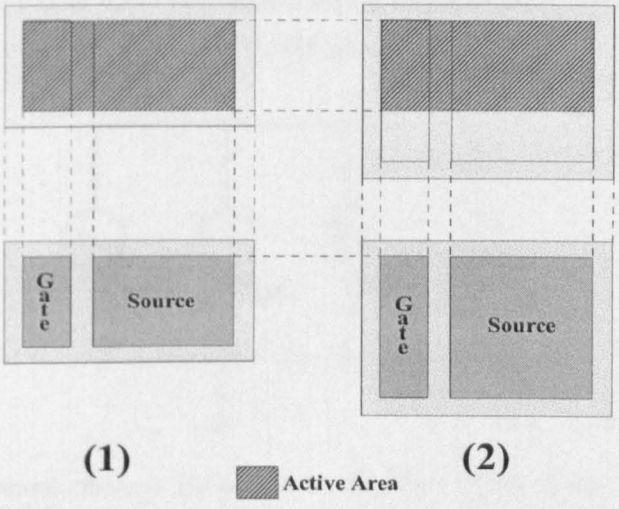


Figure 6.20: Top device layout highlighting the portion of the die occupied by the active area. (1) represents the case of an active area which is appreciatively the die size, whereas in (2) the die size is increased to increase the contact area and improve heat extraction.

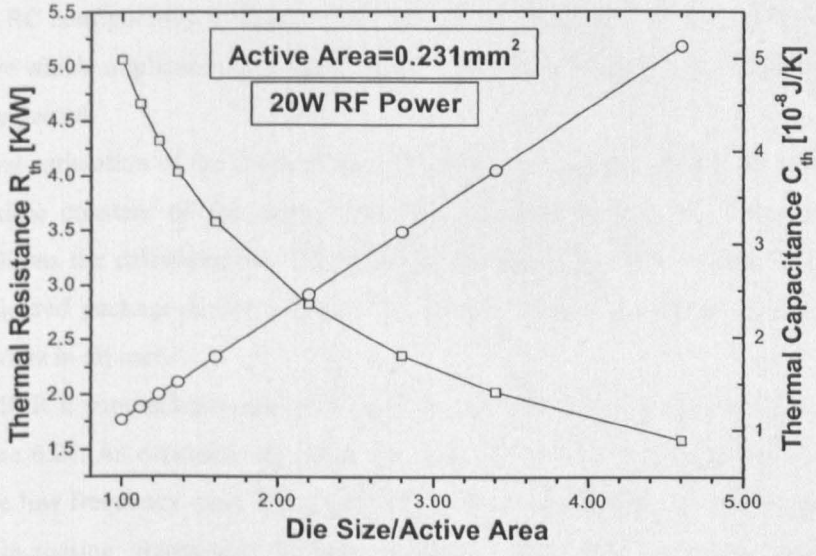


Figure 6.21: Influence of the die size on the total thermal resistance and capacitance.

6.4.2 RF thermal response

Thermal resistance and capacitance permits an estimation of the device temperature once the power being dissipated in the device is known. In SPICE models this is typically done by including a thermal sub-circuit of the kind represented in Figure 6.22. The internal device temperature results from separate contributions of the DC power and the RF power in the device.

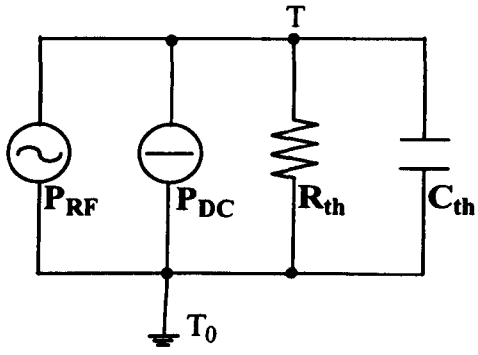


Figure 6.22: Thermal sub-circuit for the determination of the internal device temperature raise $\Delta T=T-T_0$.

In RF power applications the DC power depends on the device bias point, whereas the RF power on the amplitude of the output signal and on frequency. The low-pass RC configuration in Figure 6.22 suggests the existence of a cut-off frequency above which amplitudes of the temperature variations caused by the RF signal can be neglected.

A first estimation of the thermal cut-off frequency $F_{cut,th}$ can be carried out from the time constant of the circuit $\tau=R_{th}\cdot C_{th}$, resulting in $F_{cut,th}\approx 5/\tau$. Figure 6.23 highlights the calculated die size effect on the thermal cut-off frequency for the considered package-device design. The cut-off frequency remains smaller than 100MHz in all cases.

The SPICE simulation results of the thermal circuit of Figure 6.22 are reported in Figure 6.24. As expected, maximum thermal effect of the RF signal corresponds to the low frequency case. The amplitude of the RF temperature variation reduces for increasing frequency. In the considered case, the maximum observed temperature amplitude is below 5K at 100MHz and below 0.5K at 1GHz.

At high frequency the thermal heating is therefore only due to the DC power and the internal temperature becomes independent of the thermal capacitance. The device temperature remains constant at the DC value even in the case of large signal RF sweeps of input and output power.

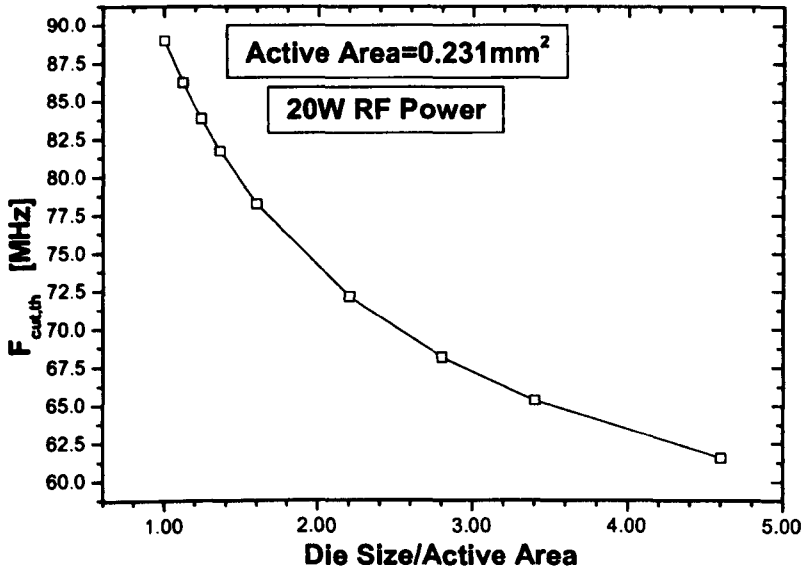


Figure 6.23: Influence of the die size on the thermal cut off-frequency.

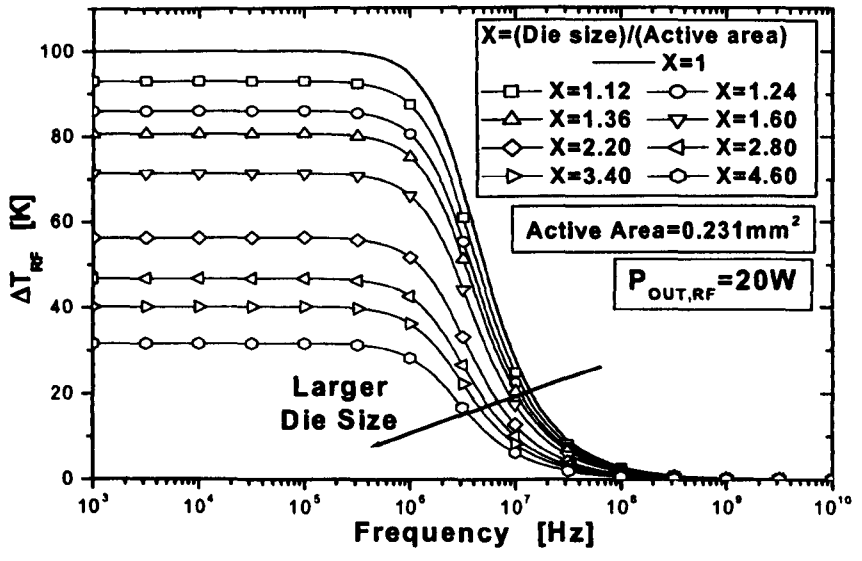


Figure 6.24: Temperature variation as a function of frequency and die size for a dissipated RF power of fixed amplitude $P_{OUT,RF}$=20W.

6.5 Summary

A modelling procedure to determine the thermal resistance of packaged devices has been described. Adequate prediction of the thermal resistance has been verified.

A novel package topology has been presented. The package design concept relies on a heat extraction that takes place from the gate and source side of the die, close to the device hot spot. Wafer thinning approaches become unnecessary for the thermal performance improvement. By using the proposed modelling procedure, improved thermal resistance is demonstrated with respect to the conventional packaging case. Improved thermal performance is achieved without the toxic and costly beryllium oxide. The absence of the source wire makes the novel package particularly attractive for packaging of RF devices.

References

[1] M. Rencz, V. Szekely, A. Poppe, K Torki, B. Courtois: Electro-thermal simulation for the prediction of chip operation within the package, SEMITHERM Symposium, 2003, pp.168-175

[2] P. Fioravanti, O. Spulber, P.P. Ray, G. Cao, D. Hinchley, M.M. De Souza, Thermal assessment of paskaged VDMOSFET: Source versus Drain side heat extraction, Proceeding ISPS, 2004, pp. 115-120

[3] B. Guenin, The use of Spreadsheets in Packaging Thermal Calculations, SEMITHERM Symposium, Tutorial, 1998

[4] B.Guenin: The 45° Heat Spreading Angle – An Urban Legend?, Electronics Cooling Magazine, Vol.9, No 4, November 2003

[5] D.P. Kennedy. "Heat Conduction in a Homogeneous Solid Circular Cylinder of Isotropic Media." IBM Report TR 00.15072.699, 1959

[6] R. Tummala, E. Rymaszewski, and A. Klopfenstein, Eds., Microelectronics Packaging Handbook, 2nd Ed., Kluwer Academic Publishers, 1998

[7] F.Christiaens, E.Beyne, J.Berghmans: Modelling of Thermal Vias in Thin Film Multichip Modules, I-THERM IV,1994, pp.101-107

[8] MEDICI User Manual, Technology Modelling Associated, Inc

[9] C.Hu, M.-H. Chi, V.M.Patel: Optimum Design of Power MOSFET's, TED, Vol. ED-31, No.12, December 1984, pp. 1693-1700

[10] T. Sammon, H. Shoefield, A. Arzumanyan, D. Kinzer, A New Generation of Wafer Level Packaged HEXFET Devices, Proceedings of PCIM Europe, 2000

[11] Siliconix Introduces MICRO FOOT, News release, 28, November 2000, Vishay-Siliconix web site

[12] A. Sawle, M. Standing, T.Sammon, A. Woodworth, DirectFET – A Proprietary New Source Mounted Power Package for Board Mounted Power, Proceedings of PCIM Europe, 2001

[13] J.A. Herbsommer, H. Safar, W. Brown, E.W. Lau, D.P. Farrel, P. Gammel, O. Lopez, G. Terefenko, Improved Electrical and Thermal Performance of Ultra-thin RF LDMOS Power Transistors, IEEE MTT-S, 2003, pp. 213-216

7

Conclusions

The work presented in this thesis describes the development and the implementation of an innovative set of tools that facilitates the development of a Si RF power MOSFET. It has been shown that a complete characterisation of the electro-thermal behaviour of the device can be carried out before the fabrication of a prototype. The possibility of predicting RF performance with reasonable accuracy, non accessible via device simulation alone, is expected to reduce time and therefore costs associated to the development of novel device concepts for RF power applications.

In this thesis, novel insight into the physical phenomena taking place in Si RF power devices is provided. Qualitative equivalent models to support the explanation have been used throughout this thesis in order to provide a deeper understanding of the link between design parameters and device performance. A simplified, yet powerful compact model has been proposed to highlight the cause of the peculiar behaviour exhibited by the transconductance of Si RF power MOSFETs.

7.1 Main achievements

The major results that have been presented in this thesis are summarised here.

- ***Intrinsic MOSFET.*** A compact model for the intrinsic MOSFET of an RF power MOSFET has been proposed. It has been demonstrated that the proposed model permits correct reconstruction of the output transfer- as well as transconductance characteristic.
- ***Transconductance flattening.*** An explanation of the transconductance flattening mechanism has been proposed. The occurrence of channel modulation has been shown to reduce the rate at which the inversion charge in the channel increases with the gate to source voltage, reducing the peak value and causing the flattening of the transconductance characteristic.

- ***Transconductance peak value.*** An expression has been proposed for the peak value of the transconductance. It has been demonstrated that this expression permits accurate prediction of the peak value of the transconductance. The formula has been shown to provide for the first time a clear link between the resistance of the drift region and the peak value of the transconductance.
- ***Load line characteristic.*** A procedure for the extraction of the optimum amplifier load line is proposed in the case of non linear behaviour of the active device used in the amplifier.
- ***Optimum matching impedance and power gain.*** Expressions for the optimum source and load impedance as well as power gain of an RF power amplifier have been proposed. The expressions explicitly include the effect of the internal capacitance, the source inductance and the gate and drift region resistance of the device.
- ***Gain compression.*** Reasonably accurate gain and gain compression predictions are demonstrated by using an analytically derived approach. The non linear behaviour of the current generator is taken into account.
- ***Linearity.*** Expressions for the estimation of intermodulation distortion have been proposed. The classical small signal analysis has been extended to properly account for large signal inputs. The proposed expressions include the effects of higher order intermodulation products than previously reported. For the first time the load line output transfer characteristic is considered in the prediction of the intermodulation distortion. A novel approach for the prediction of intermodulation distortion based on the Fourier analysis and on the load line transfer characteristic is also proposed.
- ***Device optimisations.*** A Unified Figure of Merit is proposed for the resolution of the trade-offs encountered during device design via the determination of power gain and linearity.

- ***Shield plate.*** A previously not considered charge coupling mechanism is shown to contribute to the reduction of the feedback capacitance when a shield plate is used.
- ***GP-VDMOSFET.*** A grounded p-region between the p-wells of adjacent cells is proposed in order to achieve a low feedback capacitance. The novel concept is shown to provide a valid alternative to the shielded VDMOSFET.
- ***Package thermal modelling.*** A procedure for the determination of the thermal impedance of packaged devices is demonstrated. The proposed thermal modelling procedure is used to assess the possibility of extracting heat through the source side of a power VDMOSFET. Simulation results show that significant reduction of the thermal resistance can be achieved.

7.2 Future work

- ***Intrinsic MOSFET.*** The proposed intrinsic MOSFET model has been fitted to the simulated and measured data only at the application voltage. Similar to the development of low power MOSFET models of second and third generation, the model can be extended to include the dependence of the fitting parameters on voltage and mathematical conditioning for improved convergence properties. Dependence on temperature can also be included. Finally, the model can be implemented in commercial RF SPICE simulators capable of harmonic balance analysis and fully validated versus measured performance of several devices.
- ***Package thermal modelling.*** The thermal modelling procedure can be fully validated versus measured thermal impedance data. A measurement setup can be arranged including infrared thermal sensors to monitor transient variations of temperature at the thermal junction. The measurement setup should also include a temperature controlled chuck

equipped with a thermocouple to monitor the temperature at the base of the package.

- ***Matching impedances, power gain and linearity.*** Validation of the matching impedance, power gain and linearity prediction has been strongly limited by the difficulties in obtaining accurate characterisation of the RF performance of the device. It has been observed that published datasheet data and published model parameters for the devices considered in this work do not correspond well to measured performance. In order to fully validate the analytical approach, accurate information on device and package are required, as for example total device width and gate oxide thickness. Package layout can be included in a 3D electro-magnetic simulator in order to determine and then de-embed its effect. Isothermal IVs and s-parameter measurement can be carried out takning into account the thermal resistance of the device, according to the DC power dissipated for the particular bias condition. Measured load and source pull contours can be measured to precisely establish the value of the optimum matching impedances. Ad hoc text fixtures could then be designed and used in the measurement of power gain and intermodulation distortion.

To conclude, the analysis of the limitation in gain compression and linearity prediction has shown that the effect of the non-linear device capacitance has not been considered and that the input signal is supposed to be applied directly to the current generator of the device. In order to overcome these limitations a harmonic balance optimised approach should is the best solution.

List of Figures

List of Tables