Review

RF MOSFET: recent advances, current status and future trends

Juin J. Liou a,*, Frank Schwierz b

a School of EE and CS, University of Central Florida, Orlando, FL 32816-2450, USA
b Fachgebiet Festkörperelektronik, Technische Universität Ilmenau, Germany

Received 15 April 2003; accepted 6 May 2003

Abstract

Recent advances in complementary metal oxide semiconductor (CMOS) processing, continuous scaling of gate length, and progress in silicon on insulator have stirred serious discussions on the suitability of metal-oxide semiconductor field-effect transistors (MOSFETs) for RF/microwave applications. This paper covers the recent advances and current status of mainstream CMOS as the dominating technology in very large scale integration, future trends of RF MOSFETs, and applications of MOSFETs in RF electronics. Aspects of RF MOSFET modeling are also addressed. Despite some lingering debates, the prospects for RF MOS with operating frequencies in the lower GHz range are very promising.

© 2003 Elsevier Ltd. All rights reserved.

Contents

1. Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1882

2. Evolution, current status, and future trends . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1883
2.1. Historical background of the MOSFET . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1883
2.2. Current status of MOSFET development . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1884
2.2.1. Mainstream CMOS . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1884
2.2.2. RF MOSFET . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1885
2.3. Future trends . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1886
2.3.1. Mainstream CMOS . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1886
2.3.2. RF MOSFET . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1887

3. Applications of MOS in RF electronics . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1888

4. Issues of RF MOSFET modeling . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1889
4.1. Modeling of three-terminal RF MOSFET . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1890
4.1.1. BSIM model . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1890
4.1.2. Philips MOS model . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1891
4.1.3. EKV model . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1891
4.1.4. Model comparison . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1892
4.2. Modeling of four-terminal RF MOSFET . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1892

*Corresponding author. Tel.: +1-407-823-2786; fax: +1-407-823-5835.
E-mail address: liou@pegasus.cc.ucf.edu (J.J. Liou).

0038-1101/S - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/S0038-1101(03)00225-9
1. Introduction

It is beyond anyone’s imagination as to how the metal-oxide semiconductor field-effect transistor (MOSFET) could have evolved in the past 40 years. Moreover, it is very likely that a decade from now you won’t recognize a MOSFET as it is today. Although some pundits have predicted that the evolution of semiconductor technology to smaller dimensions will slow down as the feature size hit the 0.1-µm mark, the developments in fact have been speeding up recently. The ITRS (International Technology Roadmap for Semiconductors) [1] recently revised its projection for the 2003 technology node from 120 nm (ITRS 1999 issue) to 100 nm (ITRS 2001, 2002 update), and the targeted gate length for MOSFETs in high-performance logic circuits for 2003 from 85 nm (ITRS 1999) to 45 nm (ITRS 2001). Fig. 1 shows the evolution of the DRAM capacity and the minimum MOSFET feature size (i.e., the gate length), together with the ITRS targets up to the year 2016. Continuous MOSFET scaling and simultaneous increase of chip size led to more and more complex integrated logic circuits with enhanced speed and performance. For example, microprocessors manufactured by Intel Corp. have seen the number of transistors per chip increases from 2300 in 1971 (Intel 4004) to more than 10^8 in 2003, as illustrated in Fig. 2. The clock frequency of the processors, on the other hand, has been increased from 1 MHz in 1971 (Intel 4004) to about 3 GHz in 2003 (Pentium 4).

This leads us to look at the issue of transistor speed. If a processor shows a high clock frequency, it must contain fast transistors. Indeed, continuous scaling not only made the MOSFETs smaller, but also much faster. Roughly speaking, transistor speed is inversely proportional to the gate length. In the past, MOSFETs have been considered as “slow” devices. Several reasons contributed to this conviction. First, the electron mobility, a measure of how fast the free electrons can move in a semiconductor, in Si is by nature lower than in GaAs and other compound semiconductors. Second, the inversion channel of a MOSFET is located very close to the Si/SiO_2 interface, thereby subjected to the effects of interface roughness, crystal imperfections, and interface traps. As a result, the mobility of free electrons traveling in the inversion channel is further degraded. Fig. 3
shows the effective electron mobility in the MOS channel versus the vertical electric field (i.e., the field caused by the gate voltage). It can be seen that the mobility decreases with increasing field, due to the fact that a larger field, i.e. a larger positive gate voltage, will attract the electrons in the inversion layer more strongly to the Si/SiO₂ interface and thus induce more interface scattering on the electrons. Table 1 compares the low-field electron mobilities in a Si MOS inversion channel, bulk Si, and bulk GaAs (the doping concentration is $10^{17}$ cm$^{-3}$ in all cases).

Besides the low-field mobility in MOSFET channels, a second reason for the inferior MOSFET RF performance in the past was the relatively large gate length. In the early 1980s, the gate length of production-stage MOSFETs was in the range of 1.5–2.5 μm, while RF GaAs metal-semiconductor field-effect transistors (MESFETs) with gate lengths between 0.25 and 0.5 μm were quite common and commercially available.

Thanks to the aggressive feature size reduction in the past few years, however, MOSFETs are now qualified as RF devices. Today, the MOSFET gate length is comparable or even smaller than that of III–V RF FETs. Currently, deep-submicrometer complementary metal oxide semiconductor (CMOS) processes typically reach several tens of GHz cutoff frequency $f_T$ and show relatively low noise figures, making them a serious alternative to the traditional III–V compound devices [2–4]. In addition, MOSFETs offer very large scale integration (i.e., the possibility to realize systems-on-chip) and high reliability. It should be noted that the progress toward MOSFET’s capability for RF operation is a spin-off from the developments of very large scale integration (VLSI) and is based on the tremendous research and development efforts of the Si VLSI chipmakers. Much of the advance achieved in increasing the switching speed of MOSFETs for fast logic circuits has been adopted for RF MOSFETs. Fig. 4 shows the evolution of the market share of the different semiconductor technologies. Note that in the early 2000s about 90% of the market belongs to CMOS and MOSFETs. Since the R&D budgets for enhancing the performance of a certain semiconductor technology are normally related to its market share, one can get an idea on the tremendous amounts of money spent on the development of modern and future MOSFETs.

In this paper, we will start with a brief history of the MOSFET. Because RF MOSFET development benefited from the advances in Si VLSI mentioned above, we will address both mainstream CMOS and RF MOSFETs. Issues discussed include the prospects and future trends of RF MOSFETs, combination of MOSFETs with SiGe heterojunction bipolar transistors (HBTs, SiGe BiCMOS), and applications of MOSFETs in RF electronics. Finally, important issues pertinent to RF MOSFET modeling will be addressed.

2. Evolution, current status, and future trends

2.1. Historical background of the MOSFET

The first transistor, nowadays known as a MESFET, was proposed in a patent filed by Lilienfeld in 1926 [5] who, two years later, presented the idea of the depletion-mode MOSFET [6]. These concepts were made possible by Lilienfeld’s understanding of conductivity modulation by a transverse electric field, which he repeatedly used to describe signal amplification of the transistors in the patents. However, it is still not known if Lilienfeld ever made a working, amplifying transistor.

The notion of the inversion-mode MOSFET was proposed by Heil in 1935 [7]. Nevertheless, the first MOSFET was not fabricated until 1960 by Kahng and Atalla [8,9]. The immense delay between the idea and the fabrication was due to the technical difficulties to obtain a good oxide and the lack of fundamental and basic semiconductor notions. After World War II, researchers at Bell Labs were trying unsuccessfully to make field effect devices and apparently they were not aware of the previous work by Lilienfeld and Heil. Ironically, this unsuccessful research on field-effect devices led to the
birth of the bipolar transistor [10–14]. Using the ideas of the "point contact transistor", Shockley completed the description of the bipolar junction transistor on 23 January 1948 and filed it for a patent on 26 June 1948 [15]. Shockley and his two Bell Labs colleagues, Brattain and Bardeen, received the Nobel Prize in Physics in 1956 for the "invention of the transistor." Soon the bipolar transistor became the dominating device in semiconductor electronics and maintained this position for more than two decades.

Interests in the MOSFET were resuscitated in 1960 by Kahng and Atalla [8,9], who presented the first successful realization of a silicon inversion-channel MOSFET using thermally grown oxide for the gate insulator. The MOSFET dramatically increased its importance three years later when Wanlass and Sah invented the CMOS circuit [16–18]. Due to their compactness and low power dissipation, MOSFETs have been the most widely used semiconductor device since the 1980s (see Fig. 4).

It is understandable that the evolution of integrated circuits has always followed but somewhat behind that of semiconductor devices. In 1959, Kilby fabricated a flip-flop from a single chip of germanium using gold wires for interconnections. This is called today a hybrid integrated circuit and is the predecessor of the monolithic integrated circuit. Kilby, who received the Nobel Prize in Physics 2000 for the "invention of the integrated circuit", gave in 1976 [19] his personal account about his previous invention of the integrated circuit at Texas Instruments. The development of the planar technology at Fairchild in the late 1950s allowed Noyce to invent in 1959 [20] the monolithic silicon integrated circuit. Moore, a cofounder of both Fairchild and Intel, predicted in 1965 [21] that the number of devices per integrated circuit would increase exponentially with increasing time. This astonishing prediction, which has been named the Moore's Law [22] and is still valid today, was based on only five data points available from Fairchild.

While most integrated circuits fabricated during the 1960s and 1970s were based on bipolar transistors, MOSFET-based circuits gained importance and took over the leading role in the early 1980s.

2.2. Current status of MOSFET development

2.2.1. Mainstream CMOS

The basic structure of MOSFETs, shown in Fig. 5 consisting of a single gate, a semiconducting substrate (frequently called body or bulk), and heavily doped source and drain regions, has not changed much in the past 20 years [23]. Only the dimensions and other features have been scaled down continuously to meet the demands of higher speed and increased compactness. A somewhat significant modification to the MOSFET during the last 10 years was the introduction of the silicon on insulator (SOI) technology, where the transistor body is separated from the semiconducting wafer by an insulating layer. The gate length, which is directly related to the effective channel length, is a main feature controlling the MOSFET performance. Reducing the gate length, however, requires the scaling of other features such as the oxide thickness, drain/source junction depth, and substrate doping density [24–26]. For example, when the gate length is reduced, the oxide thickness needs to be reduced so that the gate can have a better control of the channel inversion and can give rise to a larger oxide capacitance. Typically, the gate is made of a heavily doped polysilicon, and silicides are frequently deposited underneath the drain/source contacts to reduce the contact resistance.

Unfortunately, undesirable effects called short-channel effects occur in modern deep submicrometer MOSFETs. To alleviate such effects, several additional features have been added to the basic MOSFET structure. For example, to minimize hot-carrier effects due to the high electric field near the drain junction, lightly doped drain/source (LDD) regions are implemented in MOSFETs [27]. Another important phenomenon associated with the short-channel effects is the threshold voltage roll-off (i.e., threshold voltage starts to decrease when the channel length is reduced below a critical value). This can be effectively eliminated by the use of halo or pocket-implant MOSFET [28].

Many foundries, including IBM, Intel, TSMC, Philips, Motorola, and LSI Logic, are gearing up to start volume production of 100- to 90-nm technology node processes as early as 2003. In these processes, gate lengths can be selectively etched down to about half this technology node. In fact, Intel's prototype 90-nm process has already produced functional circuits with MOSFET gate length of 50 nm.

The primary aim of the entire research and development efforts in VLSI electronics has been to decrease the cost per logic function and per stored bit, and simultaneously to increase the switching speed of logic.
gates. So far this aim has in fact been achieved and resulted in new generations of microchips containing more transistors and having a higher clock frequency than their predecessor.

2.2.2. RF MOSFET

An important measure of RF transistor is the cutoff frequency $f_T$. This is the frequency at which the small-signal current gain $h_{21}$ of the transistor rolls off to unity (i.e., 0 dB). Fig. 6 shows the cutoff frequency versus the gate length of n-channel MOSFETs, with the upper limit as of May 2002 indicated [23]. For today's 50- to 100-nm gate-length MOSFETs, the cutoff frequency can reach an astonishing 200 GHz [29,30]. Applying a frequently used rule of thumb that the cutoff frequency should be around 10 times the transistor's operating frequency, one could use these devices to design integrated circuits operating up to 20 GHz, an operating frequency higher than that for the great majority of modern RF electronics. This, however, is only half of the picture because a high cutoff frequency is not the only requirement for a good RF transistor. Other figures of merit have to be considered as well. For example, a high maximum frequency of oscillation $f_{\text{max}}$ which is the frequency at which the transistor's unilateral power gain $U$ rolls off to unity (i.e., 0 dB), is often desirable. Small noise figure $NF_{\text{min}}$ and high output power $P_{\text{out}}$ are also critical for RF noise and power applications, respectively.

Properly designed III–V RF field-effect transistors (MESFET and HEMT) show comparable $f_{\text{max}}$ and $f_T$, but typically with $f_{\text{max}} > f_T$. The situation is different for Si MOSFETs. For these devices, one could either realize short-channel transistors for high $f_T$ but substantially lower $f_{\text{max}}$ or long-channel transistors for rather low $f_T$ but higher $f_{\text{max}}$. Table 2 illustrates such a trade-off using two Si MOSFETs reported in 1998 and 2000. The trade-off resulted from the fact that a very high $f_{\text{max}}$ can only be achieved with transistors having a high $f_T$ and a low gate resistance $R_G$. For III–V FETs, metal gates with multi-finger mushroom structures are frequently used to minimize $R_G$. The gates of Si MOSFETs are made of polysilicon, which has a much higher resistivity than a metal. Furthermore, for MOSFETs in digital applications, multi-finger and mushroom gates are not common. Reducing $R_G$ is imperative for RF MOSFETs because $R_G$ not only limits the power gain attainable at a certain frequency (and thus $f_{\text{max}}$), but also sets a lower limit to the minimum noise figure.

There are several means to minimize the gate resistance of Si MOSFETs:

- deposition of silicides on top of the polysilicon gate,
- metal overgates on top of the polysilicon gates,
- multi-finger gates with small finger width.

Recently, considerable progress on increasing the maximum frequency of oscillation of Si MOSFETs has been made. Fig. 7 shows the maximum frequency of oscillation as a function of gate length for Si MOSFETs reported up to 2001 [33]. Highlights are the 50-nm SOI-MOSFET [34] with an $f_{\text{max}}$ of 193 GHz ($f_T$ 178 GHz) and the 80-nm SOI-MOSFET [35] with $f_{\text{max}}$ of 185 GHz and $f_T$ of 120 GHz. In Fig. 8, the reported maximum frequency of oscillation of Si MOSFETs is plotted versus the cutoff frequency. Considerable improvement of $f_{\text{max}}$ since January 2001 can be clearly seen.

![Fig. 6. Reported cutoff frequency versus gate length of MOSFETs [23].](image1)

![Fig. 7. Reported maximum frequency of oscillation versus gate length of MOSFETs [33].](image2)
minimum noise figures $N_{\text{Fmin}}$ of experimental Si MOS-FETs are given in Fig. 9 [23]. Only noise figures up to about 15 GHz had been reported, and MOSFETs become too noisy at even higher frequency for practical microwave applications.

Due to the low breakdown voltage, the standard MOSFET structure, shown in Fig. 5, can only be used in relatively low power applications. High-power microwave amplifiers used in the base stations for mobile communication systems, however, are designed for maximum operating (and thus breakdown) voltage to deliver maximum output power. For such applications, a different MOSFET structure called the LDMOSFET (laterally-diffused MOSFET) is frequently used [23]. The cross-section of a typical high-power LDMOSFET is shown in Fig. 10. It consists of a $p^+$-Si substrate on top of a lightly doped $p^-$-layer grown epitaxially. While the $n^+$-source region extends to the gate, the $n^+$-drain region is spatially separated from the gate. The conductive connection between the $n^+$-drain and the channel region underneath the gate is realized by an $n^+$-LDD (lightly doped drain) region which is frequently called the drift region. This is the main region contributing to a high breakdown voltage. Typically, the gate length of LDMOSFETs is in the range of 0.3–1.0 μm and the gate oxide is several ten nm thick. As such, the technology, especially the lithography and the gate oxide deposition, of the LDMOSFET is much more relaxed than that of the small signal microwave MOSFET. Depending on the specific designs, typical $f_T$ of LDMOSFETs is around 5–15 GHz, and the drain-to-source breakdown voltage $B_{\text{VDSS}}$ is around 20–40 V.

2.3. Future trends

2.3.1. Mainstream CMOS

The road to advances beyond a decade into the future has always been obscure and has stimulated much speculation as to where miniaturization must end. Thus far, new ideas have regularly met the challenges posed by new problems. One thing for sure is that the downscaling will continue at least for a couple of years. The ITRS targets for the year 2016 for high-performance logic circuits (microprocessors in desktop computers) are a physical MOSFET gate length (i.e., the transistor’s gate length after all fabrication steps are finished) of 9 nm and microprocessors with a chips size of 140–280 mm² containing more than 3 billion transistors. To reach this target and ensure that the transistor still operates satisfactorily, however, requires many changes and innovations. The most likely changes are [39]:

**Fig. 8.** Reported maximum frequency of oscillation of MOSFETs versus the cutoff frequency [33].

**Fig. 9.** Measured data of minimum noise figure versus frequency, with four lowest figures indicated with (a), (b), and (c) reported in [36–38].

**Fig. 10.** Schematic cross-section of a typical LDMOSFET structure [23].
(a) To increase the mobility and improve performance, silicon will be mixed with a semiconductor like germanium to produce a strained monocrystalline silicon layer at the wafer surface that lets electrons move faster.

(b) To reduce the gate leakage current, gate oxide will be made of materials with higher dielectric constant than today's silicon dioxide.

(c) To reduce the gate resistance and improve channel control, gates will be made of metals, instead of polysilicon.

(d) For better control of MOSFET's on and off states and reduce power consumption, more than one gate will be used.

The technology of putting a Si strained layer on a silicon wafer seems to be closest to commercialization. The process is as follows. Germanium atoms replace some of the silicon atoms near the wafer's surface, then a thin layer of silicon is grown on top of this SiGe layer. Because Ge atoms are larger than Si atoms, the distance between the atoms in the SiGe layer is larger than that in pure Si. When the top Si layer is grown, its atoms line up with the SiGe atoms below, and it becomes strained in the two directions parallel to the plane of the wafer. Finally, MOSFETs are fabricated in this top strained silicon layer. The tensile strain forces in the strained silicon layer change the energy band structures. As a result, the effective mass of electrons and holes is reduced and the phonon scattering is decreased. Because of these, much higher free-carrier mobilities in Si strained layer are possible. Several research groups reported a 60% increase in the mobilities [40].

As mentioned earlier, the oxide thickness decreases with decreasing gate length. For a 100-nm process, the oxide thickness should be around or even less than 1.5 nm. Such a thin oxide allows for a considerable amount of current to flow from the gate to channel, thereby increasing the power consumption and degrade the functionality of MOSFETs. The solution to this problem seems to be straightforward—replace the silicon dioxide with a material having a higher dielectric constant $k$. A gate over a thick high-$k$ material can control the channel just as effective as one over a thinner low-$k$ material. Consequently, using a high-$k$ dielectric allows one to increase the oxide thickness and reduce the gate leakage. Promising candidates for high-$k$ dielectrics are oxinitrides and hafnium dioxide.

Today's MOSFETs have heavily doped polysilicon gates. When the gates are biased, a depletion region forms near the interface of polysilicon gate and oxide layer. Since the oxide is very thin, this depletion region in effect increases considerably the thickness of the dielectric and therefore reduces the effectiveness of the controlling mechanism of the gate. The second drawback of polysilicon gates is, as mentioned above, their relatively high series resistance. Metal gates can minimize these two problems. But before choosing a suitable metal, factors such as thermal stability and work function need to be considered. Recently, ruthenium–tantalum alloy has been suggested as a possible gate metal, which has the advantage of adjustable work function to meet the required threshold voltage of MOSFETs [41].

In the past few years, more and more researches have been devoted to investigating innovative MOSFETs having more than one gate to enhance the control of ultra-small MOSFETs. Several vertical and planar structures have been suggested and tested, and the frontrunner is an approach called the double-gate MOSFET under development by many semiconductor companies [42–44]. In Fig. 11, two double-gate MOSFET versions are shown. On the left is a planar double-gate transistor investigated by IBM researches, and the other is a vertical double gate FET commonly called FinFET favored by AMD, TSMC, and others. The FinFET is built by etching the silicon layer of an SOI wafer to form narrow vertical fins on the wafer surface. The narrow fin forms the channel, the source and drain are formed at end of the fin, and the gate drapes over both sides of the fin (i.e., two gates). Since the fin is made extremely thin, no region of the fin escapes the influence of the gate, and no leakage path for the carriers to flow between the source and drain when the device is off. Currently, Intel is pushing for a modified FinFET version called the Trigate FET. Here, an almost quadratic cross-section of the fins is striven for and the gate surrounds the fin from the vertical surfaces and from the top. Thus, there are three gates controlling the current path in the fin. Such a device is schematically shown in Fig. 12.

2.3.2. RF MOSFET

The advanced MOSFET concepts described above are focused on developing transistors for high-level integrated circuits, such as microprocessors, ASICs, etc. How they will be used for RF MOSFETs is still unclear.

![Fig. 11. Schematic of a planar double-gate MOSFET (left) and a vertical double-gate MOSFET (FinFET, right).](image-url)
Currently, the preferred RF MOSFET structures are conventional single-gate bulk or single-gate SOI MOSFETs. There is no doubt, however, that RF MOSFETs will benefit from the progress made in Si VLSI. Besides the MOSFET, another important issue for RF electronics is good passive on-chip components, i.e. capacitors, inductors, and resistors [45], which can be made using the CMOS process.

Recent progress on the SiGe HBT, although in principle being a competing device, will help to expand the market share of MOSFETs in RF electronics. SiGe HBTs offer excellent RF properties (high speed, i.e. both high $f_T$ and $f_{max}$ and low noise figure), as evidenced by the maximum frequency of oscillation versus cutoff frequency shown in Fig. 13 [33] and the noise performance given in Table 3.

During the last few years, considerable efforts have been devoted to realize SiGe HBTs from standard CMOS processes [48–51]. This makes pure Si-based RF ICs possible, where SiGe HBTs are used in the critical parts which demand RF performance (such as frequency and noise performance) while Si MOSFETs are employed in other components. Fig. 14 shows the $f_T$ and $f_{max}$ values obtained from the CMOS-compatible SiGe HBTs as a function of emitter width [33]. An interesting feature of SiGe HBTs is the fact that both $f_T$ and $f_{max}$ between 50 and 100 GHz can be realized with an emitter width larger than 0.4 µm. This is different from the RF Si MOSFET, which requires a much smaller gate length to reach such high frequency limits.

### 3. Applications of MOS in RF electronics

In the following, civil RF applications that currently have large market volumes or are expected to create mass markets in the near future are described. Furthermore, RF MOSFET applications in wireless communications are discussed.

The operating frequencies of civil RF applications can range from a few hundreds MHz to 100 GHz, but most systems having mass markets operated at frequencies below 6 GHz. The number of units sold in these markets is in the order of millions per year. Table 4 summarizes various applications and their frequency spectrums.

Examples for civil communications systems in the higher frequency range are the Hiperlink system (17 GHz), direct-to-satellite communication (20 GHz down
FETs are not the best devices for this application due to the large on-current and low on-resistance. A supply voltage of 3 V, RF power transistors possess high efficiency at low output power combined with a high efficiency at a limited supply voltage of 3 V has been established as a standard. To deliver a high efficiency, additional low-noise transistors used in wireless communications are required. The noise requirements for any low-noise amplifier, the use of low minimum noise figure transistors is desired. The noise requirements for transistors such as to whether such devices will find widespread applications in industry.

Another requirement for the handset is the reduction of power consumption. At present, a supply voltage of 3 V has been established as a standard. To deliver a high output power combined with a high efficiency at a limited supply voltage of 3 V, RF power transistors possessing a large on-current and low on-resistance are required. MOSFETs are not the best devices for this application due to the relatively low output power density. GaAs transistors (especially GaAs HBT) dominate this application. To date, the dominant power RF transistor used in base stations of wireless communications systems with operating frequencies up to 2.5 GHz is the Si LDMOSFET, which in the last several years has replaced all other competing Si and GaAs transistors. Si LD-MOSFETs only become the advantages of moderate cost, high reliability and extremely high output power. Aside from cellular phones, several other civil RF systems with operating frequencies below 20 GHz are potential fields for the application of RF MOSFETs as well. For example, Bluetooth, in which the requirements on transistor's RF performance are quite moderate, is predestined for RF MOSFETs. In fact, all-MOS Bluetooth products are already commercially available. For other applications below 20 GHz but with more stringent requirements concerning RF performance, the combination of SiGe HBTs and CMOS (SiGe BiCMOS) is currently a heavily discussed possibility in industry.

Although RF MOSFETs seem to fulfill most of the performance requirements for civil RF systems with operating frequencies up to 6 GHz, debates still existed as to whether such devices will find widespread applications in the market. Nevertheless, there is a growing belief that the importance of MOSFETs in RF applications will increase in the near future and for many years to come.

### 4. Issues of RF MOSFET modeling

MOSFET has four terminals: drain, source, gate, and body. The number of free carriers in the channel is mostly controlled by the field induced from the gate voltage, but a change in the body potential can also affect the number of channel carriers. At low frequencies, the impedance of the junction capacitance is so large that the substrate impedance may not be seen from the drain terminal, and a MOSFET can be modeled as a three-terminal device. In RF MOSFETs, however, the distributed $R-C$ network composed of the depletion capacitances and the substrate resistances, and the ac current passing through this $R-C$ network, make the concept of three-terminal device invalid.

It is much more difficult to explain and predict the behavior of a four-terminal device than that of a three-terminal one. A three-terminal device can be treated as a two-port network, where four complex numbers are sufficient to characterize the device. On the other hand, nine complex numbers are required to characterize a three-port network, such as a four-terminal device. Unfortunately, there is no established measurement technique to perform on-wafer three-port S-parameter measurements. Even if three-port measurements are possible, the thre
possible, many bias combinations for the ac characteristics will need to be considered, and the measurements and parameter extractions become impractical and too time consuming.

In spite of these problems, intensive efforts have been devoted to RF MOS modeling and parameter extraction. In the following, we will first briefly introduce the strategies for RF modeling and parameter extraction of two-port networks. Then the modeling of four-terminal MOSFETs based on the macro-modeling approach will be discussed. Only the approaches and concepts will be given in this section, as the detailed information of device physics and RF modeling can be found in the literature.

4.1. Modeling of three-terminal RF MOSFET

We will first discuss RF III–V FET models, as these models have been developed in the past to address high-frequency issues and can be readily extended for RF MOSFETs.

Only when both a proper model and a good set of model parameters are used together can one obtain accurate and meaningful circuit simulation results. Thus it is imperative to develop a well-constructed model and its parameter extraction method in order to effectively describe and predict the device characteristics. A model is normally represented by an equivalent circuit consisting of elements such as resistances, capacitances, inductances, and current sources. These elements are described by mathematical equations. All of the constants and coefficients associated with the equations, called the model parameters, ideally should have physical origins. However, some of them may result from empirical experiences to enhance the model accuracy. The model parameters need to be determined from extraction techniques based on experimental data from MOSFETs measured at different biased condition and frequencies. At low frequencies, these measurements are usually done utilizing Z, Y, or H matrices. This is because it is easier to measure the voltage or current with open or short circuits. However, there are many problems with this at high frequencies. By terminating the port with a cable of characteristic impedance, the S-parameter technique measures power waves propagating into and being reflected from the device and thus is the easiest and the most reliable way to characterize high-frequency networks.

A conventional small-signal equivalent circuit of III–V FETs is shown in Fig. 15 [59]. This equivalent circuit, which consists of capacitances $C$, inductances $L$, resistances $R$, and a current source involving transconductance $g_m$ and delay time $\tau$, can be divided into two parts:

1. Intrinsic elements: $g_m$, $g_{ds}$, $C_{gs}$, $C_{gd}$, $C_{ds}$, $R_s$, and $\tau$, which are functions of biases.

2. Extrinsic elements: $L_g$, $R_g$, $C_{pg}$, $L_s$, $R_s$, $C_{pd}$, and $L_d$, which are independent of biases.

The extrinsic elements are first determined from data measured at cold-FET conditions. Then, de-embedding the extrinsic elements yields the values of intrinsic elements.

Small-signal simulation based on the model in Fig. 15, unfortunately, cannot be used to reliably extrapolate the true large signal performance of RF MOSFETs. This is due to the fact that the small signal model, while simple and straightforward, does not guarantee charge conservation in a semiconductor device [60]. Charge conservation is absolutely required for accuracy and convergence of circuit simulation. This means that after controlling voltages are allowed to vary and then returned to their original values, the values of the current and charge should remain unchanged regardless of the variation history. As a result, accurate RF MOSFET modeling should be based on the large signal model. As the equivalent circuit in Fig. 15 does not contain a complete set of intrinsic capacitances, we cannot build a large-signal model assuring charge conservation from this small-signal model. The equivalent circuit of a typical large signal model for III–V FETs is illustrated in Fig. 16. It composes of nonlinear intrinsic components such as the current and charge sources, rather than linearized intrinsic components such as resistances and capacitances. Based on this, a more accurate charge-conserving small-signal model can then be developed for RF simulation [61].

The Shichman–Hodges model, normally refer to as level 1 MOSFET model, has simple equations, yet it is only valid for long-channel devices and does not have charge conservation. In the past 20 years, such a model has evolved into the following three mainstream, public-domain models: BSIM from UC Berkeley [62], MOS model from Philips Laboratories [63], and EKV model.
These models are basically large signal ones, so the issue of charge conservation does not present a problem. The BSIM, MOS, and EKV models have been implemented in several commercially available SPICE simulators, as illustrated in Table 5.

An overview and brief comparison of the three models are presented below.

4.1.1. BSIM model

As illustrated in Table 6, the BSIM model has gone through several improvements and versions in the past, starting from the Berkley Level-I model for 5.0-µm MOSFETs to the latest BSIM4 for 0.1-µm MOSFETs. Many physical effects have been added during the BSIM evolution. For example, the BSIM3 was developed from a coherent quasi-two-dimensional analysis of the MOSFET, and it takes into account the effects of many device and process variables for good model scalability and predictability.

4.1.2. Philips MOS model

MOS11 model is the latest version in Philips MOS model family, but MOS9 model is still being used in the industry. Such models were developed intended for the simulation of circuit behavior with emphasis on analog applications. The models give a complete description of all transistor-action related quantities, such as nodal currents and charges, noise-power spectral densities, and weak avalanche currents. The equations describing these quantities were based on the gradual channel approximation with a number of first-order corrections for small-geometry effects. The consistency is maintained by using the same carrier-density and electric-field expressions in the calculations of all model quantities.

4.1.3. EKV model

EKV model is a relatively new model, which is just beginning to be implemented in some commercial circuit simulators. Two EKV versions existed: the old EKV v2.6 and more recent EKV v3.0. The model is somewhat unique in that it uses the substrate contact, rather than the source contact, as the voltage reference. This greatly simplifies the description of the MOS device behavior and has given the EKV model a greater hope of fundamentally eliminating the asymmetry problems in the source referencing models. As a result, the complexity of the model structure is greatly reduced, with a small number of physically relevant parameters. Despite the simplicity, the EKV model is quite accurate and generates continuous second derivatives. Furthermore, the

Table 5

| Availability of BSIM, MOS, and EKV models in commercially available software tools |
|----------------------------------------|----------------|--------------|--------------|--------------|--------------|
| BSIM3 | BSIM4 | MOS9 | MOS11 | EKV |
| HSspice | Available | Available | Available | Available | Available |
| Spectre | Available | Available | Available | Available | Available |
| SmartSpice | Available | Available | Available | Available | Available |
| Pspice | Available | | | | |

Table 6

| Comparison of the various BSIM versions |
|----------------------------------------|----------------|--------------|--------------|--------------|--------------|
| Model | Geometry (µm) | Ids | Subthreshold | Rout | Scalability |
| Berkeley Level-1 | 5.0 | Poor | NA | Fair | Poor |
| Berkeley Level-2 | 2.0 | Poor | Poor | Fair | Fair |
| Berkeley Level-3 | 1.0 | Fair | Poor | Poor | Poor |
| BSIM-1 | 0.8 | Good | Fair | Poor | Fair |
| BSIM-2 | 0.25 | Good | Good | Fair | Fair |
| BSIM-3 | 0.15 | Good | Good | Good | Good |
| BSIM-4 | 0.1 | Good | Good | Good | Good |
The EKV model is built around a hierarchal structure which allows for a clean link between hand calculations (using a subset of the model equations) and the full fledged circuit simulation. This model is also the first to introduce parameters to specifically address the issue of MOS mismatch.

### 4.1.4. Model comparison

Table 7 summarizes the major differences in modeling approach of the BSIM, Philips MOS, and EKV models. For example, only the MOS11 model includes diffusion tendency in describing the dc current, and the EKV model is the only one using the bulk terminal as the reference. These differences give rise to the different levels of complexity, robustness, and accuracy among these models.

It is fairly safe to say that the number of model parameters increases with increasing model complexity. As illustrated in Fig. 17, the number of intrinsic model parameters is indeed increased exponentially over the past 40 years. It is also indicated that the EKV and Philips MOS models possess less model parameters than the BSIM model.

### 4.2. Modeling of four-terminal RF MOSFET

Theoretically, it is possible to develop a four-terminal MOSFET model for RF applications. However, the measurements of four-terminal devices require much effort, and the parameter extraction procedures are too difficult due to the complicated theoretical $Y$-parameter expressions of the four-terminal equivalent circuit. The macro-modeling approach can alleviate this problem. This approach makes use of a commercially available MOSFET model core, such as the BSIM4, MOS11, or EKV model, and adds lumped-element equivalent circuit extensions. In other words, the model core and lumped components compose an equivalent circuit representing an RF four-terminal MOSFET. One of these lumped components is the gate resistance, which consists of the distributed gate electrode resistance as well as the non-quasi-static element. The effect of substrate using lumped components should also be incorporated into the macro model. Fig. 18 shows a macro RF MOS model consisting of the gate resistance and a sub-circuit accounting for the effect of substrate including two capacitances and five resistances.

Usually, parameters of the model core are extracted with dc and low-frequency $C-V$ measurements using the same procedures available in the commercial model. The extraction of the lumped elements is then carried out using measured $S$-parameters at a given dc bias condition. The gate resistance has a great influence on the real part of $Y_{11}$, and the substrate network affects the real part of $Y_{22}$ most significantly. Thus, the gate resistance and substrate-related elements are extracted from...
Re\((Y_{11})\) and Re\((Y_{22})\) directly or using a fitting technique. Results of frequency-dependent \(Y\)-parameters modeled and measured from a 0.18-\(\mu\)m n-MOSFET are given in Fig. 19 [67].

The approach of macro-modeling discussed above provides a useful compromise between accuracy and efficient. These models, accompanied by appropriate parameter extraction processes, show fairly good agreements with measured RF data. However, such an approach cannot overcome some of the limitations inherited in the MOSFET model core. For example, \(g_m\) and \(g_{ds}\) predicted by commercial models is not yet very accurate, this is especially troublesome for RF circuit simulation where the higher order derivatives of the current should be smooth and correct. There are also errors in predicting the intrinsic \(C-V\) characteristics of short-channel MOSFETs. Moreover, scalable models for the added lumped elements are needed for the whole model to be scalable and predictive.

5. Conclusions

The evolution, state of the art, as well as future trends of RF MOSFETs have been presented and discussed. Aspects of RF MOSFET modeling and three widely used MOS models were also reviewed. There is no doubt that MOSFETs are progressing quickly and becoming a strong contender in the RF applications traditionally dominated by III-V devices. Today’s RF MOSFETs already fulfill most of the performance requirements for civil RF systems operating in the frequency range up to 6 GHz. For example, a low-end wireless system such as Bluetooth is well suited for MOS technology. In high-end systems, the combination of MOSFETs and CMOS-compatible SiGe HBTs (i.e., SiGe BiCMOS) is a promising option.

References


