The Invariance of Characteristic Current Densities in Nanoscale MOSFETs and Its Impact on Algorithmic Design Methodologies and Design Porting of Si(Ge) (Bi)CMOS High-Speed Building Blocks

Timothy O. Dickson, *Student Member, IEEE*, Kenneth H. K. Yau, *Student Member, IEEE*, Theodoros Chalvatzis, Alain M. Mangan, Ekaterina Laskin, *Student Member, IEEE*, Rudy Beerkens, *Member, IEEE*, Paul Westergaard, Mihai Tazlauanu, *Member, IEEE*, Ming-Ta Yang, *Member, IEEE*, and Sorin P. Voinigescu, *Senior Member, IEEE*

***Abstract*—This paper provides evidence that, as a result of constant-field scaling, the peak f_T (approx. 0.3 mA/μm), peak f_{MAX} (approx. 0.2 mA/μm), and optimum noise figure $\mathrm{NF}_{\mathrm{MIN}}$ (approx. 0.15 mA/μm) current densities of Si and SOI n-channel MOSFETs are largely unchanged over technology nodes and foundries. It is demonstrated that the characteristic current densities also remain invariant for the most common circuit topologies such as MOSFET cascodes, MOS-SiGe HBT cascodes, current-mode logic (CML) gates, and nMOS transimpedance amplifiers (TIAs) with active pMOSFET loads. As a consequence, it is proposed that constant current-density biasing schemes be applied to MOSFET analog/mixed-signal/RF and high-speed digital circuit design. This will alleviate the problem of ever-diminishing effective gate voltages as CMOS is scaled below 90 nm, and will reduce the impact of statistical process variation, temperature and bias current variation on circuit performance. The second half of the paper illustrates that constant current-density biasing allows for the porting of SiGe BiCMOS cascode operational amplifiers, low-noise CMOS TIAs, and MOS-CML and BiCMOS-CML logic gates and output drivers between technology nodes and foundries, and even from bulk CMOS to SOI processes, with little or no redesign. Examples are provided of several record-setting circuits such as: 1) SiGe BiCMOS operational amplifiers with up to 37-GHz unity gain bandwidth; 2) a 2.5-V SiGe BiCMOS high-speed logic chip set consisting of 49-GHz retimer, 40-GHz TIAs, 80-GHz output driver with pre-emphasis and output swing control; and 3) 1-V 90-nm bulk and SOI CMOS TIAs with over 26-GHz bandwidth, less than 8-dB noise figure and operating at data rates up to 38.8 Gb/s. Such building blocks are required for the next generation of low-power 40–80 Gb/s wireline transceivers.**

***Index Terms*—Characteristic current densities, f_{MAX}, f_T, low-noise transimpedance amplifiers, MOS-CML, nanoscale CMOS, noise figure, operational amplifiers, output drivers, SiGe BiCMOS, SOI, wave-shape control.**

Manuscript received December 1, 2005; revised March 1, 2006. This work was supported by the Natural Sciences and Engineering Research Council of Canada, Micronet, STMicroelectronics, Taiwan Semiconductor Manufacturing Company, and Gennum, and by an equipment grant from Canada Foundation for Innovation and the Ontario Innovation Trust. CAD tools were provided by Canadian Microelectronics Corporation.

T. O. Dickson, K. H. K. Yau, T. Chalvatzis, A. M. Mangan, E. Laskin, P. Westergaard, and S. P. Voinigescu are with the Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, M5S 3G4 Canada (e-mail: sorinv@eecg.utoronto.ca).

R. Beerkens is with STMicroelectronics, Ottawa, ON, K2H 8R6 Canada.

M. Tazlauanu is with Quake Technologies, Ottawa, ON, K2K 2T8 Canada.

M.-T. Yang is with Taiwan Semiconductor Manufacturing Company, Hsin-Chu, Taiwan 300-77, R.O.C.

Digital Object Identifier 10.1109/JSSC.2006.875301

I. Introduction

THE measured maximum available power gain of single-transistor and cascode stages fabricated in state-of-the-art SiGe BiCMOS and 90-nm CMOS technologies exceeds 8 dB at 65 GHz [1]. Taking advantage of this outstanding transistor performance, we have recently demonstrated large levels of integration at 80 Gb/s in a PRBS generator with $2^{31}-1$ pattern length, implemented in 130-nm SiGe BiCMOS technology and operating from 3.3-V supply [2]. To date, only SiGe BiCMOS and III-V circuits have shown adequate performance for 40 Gb/s fiberoptic applications. However, they either operate from 3.3 V or higher supplies or/and dissipate significant power. CMOS circuits hold prospects for lower power but a 40-GHz latch, as needed in retimed OC768 systems, has so far proven beyond the reach of even the most advanced SOI and strained-silicon CMOS technologies [3].

In this work, we revisit CMOS and SiGe BiCMOS high-speed and low-noise topologies in the context of deep-submicron and nanoscale technologies and of operation from 2.5 V or lower supply voltages. The focus is on developing algorithmic design and scaling methodologies for the key building blocks that limit dynamic range—i.e., broadband input comparators and output drivers—and on minimizing the power dissipation of logic gates while achieving 40–80 Gb/s switching speeds.

A critical piece in realizing this goal is the evidence provided in the first part of the paper that the characteristic current densities of MOSFETs, corresponding to the peak f_T, peak f_{MAX}, and minimum noise figure $\mathrm{NF}_{\mathrm{MIN}}$, remain invariant over technology nodes and foundries. This is most likely caused by mobility degradation due to the vertical field [3], [4] and by the constant-field scaling rules [4] being applied consistently by most foundries following the International Technology Roadmap for Semiconductors (ITRS) [5].

While these and other authors [6]–[12] have explored the impact of scaling on the peak f_T, peak f_{MAX} values, and on the noise parameters of MOSFETs, it is for the first time that the scaling of the corresponding current densities is investigated. The implications on the design, scaling and porting of broadband and high-speed mm-wave circuits are discussed.

The paper is organized as follows. In Section II and the Appendix, simple analytical expressions are used to capture the

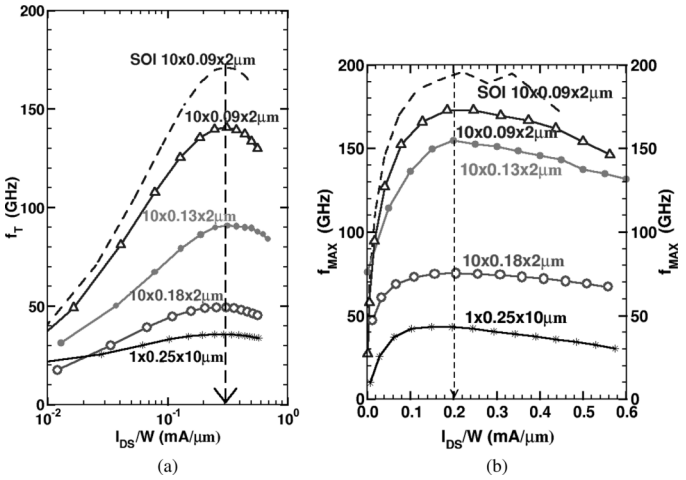


Fig. 1. Measured (a) f_T and (b) $f_{\rm MAX}$ as a function of drain current per micron of gate width for n-channel MOSFETs fabricated in different bulk and SOI technology nodes (0.25 μm from foundry A, 0.18 μm from foundry B, and 0.13-μm and 90-nm nodes from foundry C).

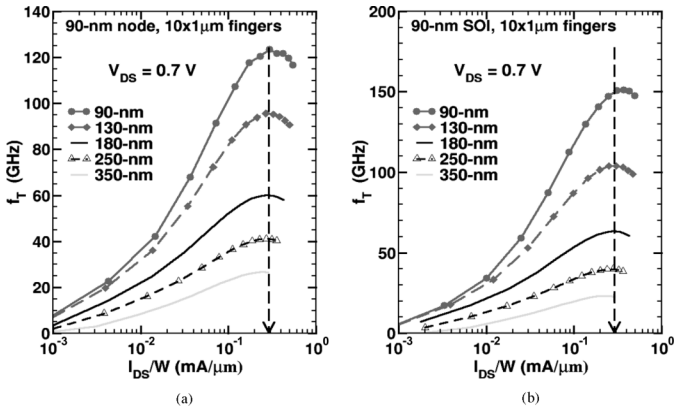


Fig. 2. Measured f_T as a function of drain current per micron of gate width for (a) bulk and (b) SOI n-channel MOSFETs fabricated with different drawn gate lengths in the 90-nm node.

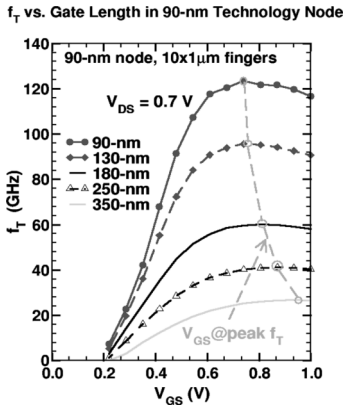


Fig. 3. Measured f_T as a function of gate-to-source voltage for bulk n-channel MOSFETs fabricated with different drawn gate lengths in the 90-nm node.

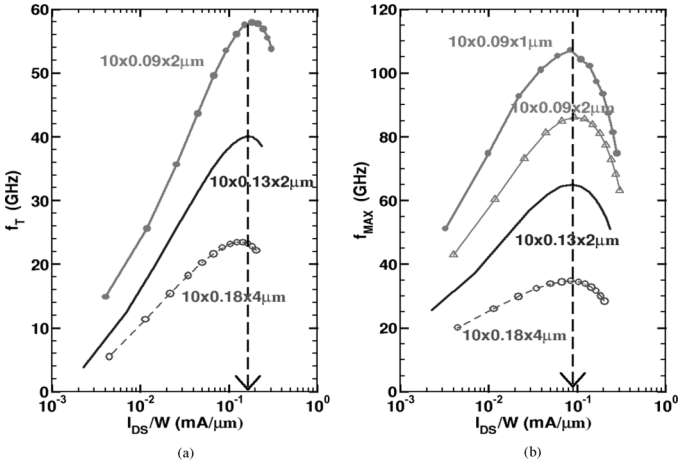


Fig. 4. Measured (a) f_T and (b) $f_{\rm MAX}$ as a function of drain current per micron of gate width for p-channel MOSFETs fabricated in different technology nodes (0.18 μm from foundry B, and 0.13-μm and 90-nm nodes from foundry C).

experimentally observed bias dependence of g_m, f_T, $f_{\rm MAX}$, and $\rm NF_{MIN}$ of Si bulk and SOI MOSFETs, and to prove that the characteristic current densities remain invariant over technology nodes, temperature, circuit topologies and, in a given technology node, over a wide range of transistor gate length and threshold voltage values. This is followed in Section III by a theoretical and experimental analysis of the scaling of high-speed current-mode logic (CML) and broadband low-noise amplifier topologies between technology nodes. Examples of high-speed CML and I/O circuits implemented in 130-nm SiGe BiCMOS, 130-nm, and 90-nm CMOS and SOI technologies are presented. These circuits demonstrate that excellent performance can be realized using the design and scaling techniques described in Sections II and III.

II. The Invariance of MOSFET f_T, $f_{\rm MAX}$, and $\rm NF_{MIN}$ Characteristic Current Densities

A. Transistors

We have recently observed that the measured peak $f_{\rm MAX}$ and peak f_T current densities of 130-nm Si nMOSFETs, 150-nm GaAs pHEMTs, and 100-nm InP HEMTs fall in the 0.2 mA/μm to 0.3 mA/μm range [9], [10]. Intrigued by these results, we have collected and scrutinized more closely the measured f_T and $f_{\rm MAX}$ characteristics for n-channel Si MOSFETs fabricated in the 250-nn, 180-nm, 130-nm, and 90-nm bulk and SOI technology nodes by different foundries. These data are summarized in Fig. 1 for nMOSFETs and confirm the earlier observation. We have found similar characteristic current densities going back to another paper of ours on the RF and noise figures of merit (FOMs) of 0.5-μm bulk CMOS technology [11], as well as in recently published data on 90-nm bulk CMOS from two other foundries [6], [12]. It is also interesting to notice that, unlike in heterojunction bipolar transistors (HBTs) where the peak $f_{\rm MAX}$ and peak f_T current densities coincide [9], [10], in Si, GaAs, and InP FETs, the peak $f_{\rm MAX}$ current density $(J_{\rm pfMAX})$ is systematically lower, around 0.2 mA/μm, than the peak f_T current density (J_{pfT}), typically 0.3 mA/μm. It became obvious that such behavior was not the result of mere serendipity and that it could be harnessed in more robust MOSFET circuit design for a wide range of high-speed digital, analog, mixed-signal, and RF applications.

More surprisingly, Fig. 2 illustrates that, in the 90-nm bulk and SOI CMOS node, these current densities remain constant for gate lengths varying between 90 nm and 350 nm, while, as

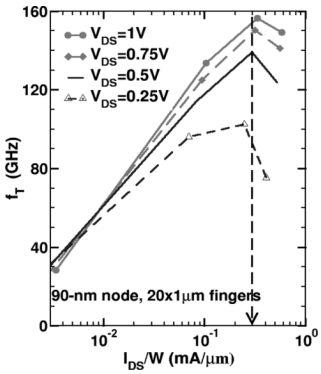


Fig. 5. Measured f_T as a function of drain current per micron of gate width for different drain–source voltages in a 90-nm nMOSFET.

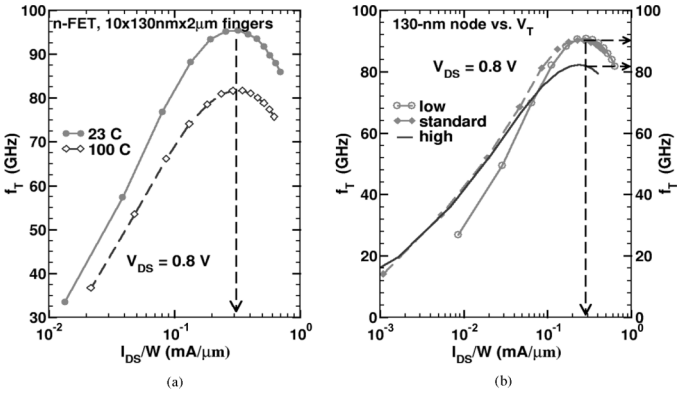


Fig. 7. Measured f_T as a function of drain current density for (a) bulk 130-nm nMOSFET at 23 °C and 100 °C, and (b) for standard-, low-, and high-V_T bulk 130-nm nMOSFETs.

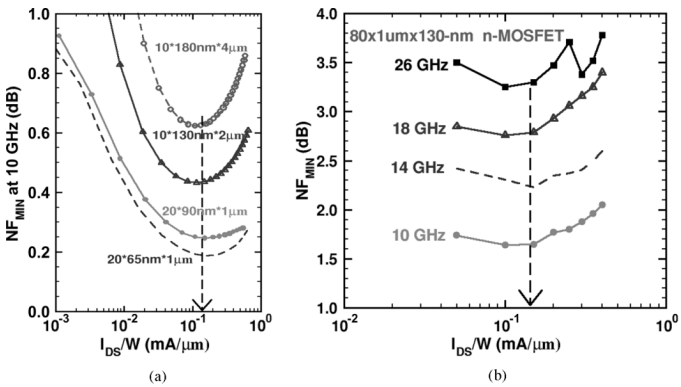


Fig. 6. (a) Simulated $\mathrm{NF_{MIN}}$ at 10 GHz as a function of drain current per micron of gate width in nMOSFETs fabricated in different technology nodes. (b) Measured $\mathrm{NF_{MIN}}$ as a function of drain current density at different frequencies for a $80 \times 1\ \mu\mathrm{m} \times 130$ nm MOSFET in the 130-nm node. The pad and interconnect have not been de-embedded from the measured noise figure.

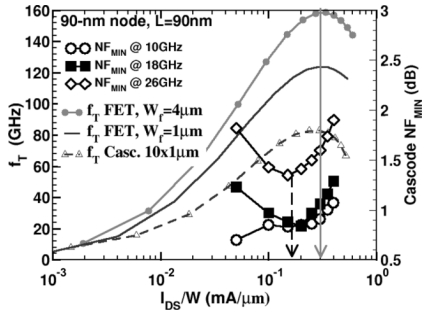


Fig. 8. Measured f_T and $\mathrm{NF_{MIN}}$ at different frequencies as a function of drain current per unit gate width for 90-nm n-channel MOSFETs and a 90-nm n-channel cascode.

shown in Fig. 3, the effective gate voltage at the peak varies widely. Similar behavior, measured over three different technology nodes from different foundries, is observed in p-channel devices for which peak f_T, shown in Fig. 4(a), and peak f_{MAX} values, reproduced in Fig. 4(b), and their corresponding current densities, are typically 40%–45% of those encountered in n-channel MOSFETs. This ratio closely follows that of the I_{ON} values of p- and n-channel devices [5]. However, J_{pfT} is less than half the I_{ON} for a particular device. We remind readers that I_{ON} is an important FOM for conventional CMOS logic [5].

Fig. 5 illustrates that, as long as devices are biased in the saturation (active) region, the characteristic current densities are only weakly dependent on the drain–source voltage through the channel-length modulation effect.

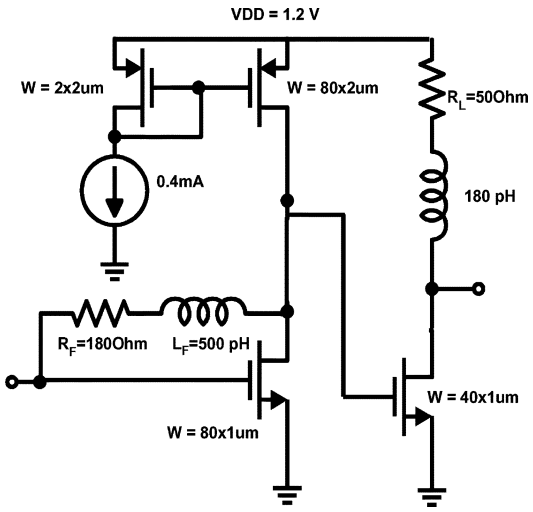
The simulated $\mathrm{NF_{MIN}}$ of n-channel MOSFETs in several technology nodes are plotted in Fig. 6(a), along with the measured $\mathrm{NF_{MIN}}$ of 130-nm nMOSFETs as a function of current density at different frequencies in Fig. 6(b). Unlike in HBTs, the optimum $\mathrm{NF_{MIN}}$ occurs at the same current density, approximately 0.15 mA/μm, irrespective of frequency [10] and of the technology node. The simulated and measured optimum noise figure current density (J_{OPT}) is in close agreement with measurement data reported by others in the 90-nm node [6], [7], and is relatively close to the peak-f_{MAX} bias of 0.2 mA/μm.

Additionally, the characteristic current densities remain largely unchanged over temperature and threshold voltage, as seen from the 130-nm node nMOS f_T plots in Fig. 7. For a detailed and intuitive analytical explanation of the underlying device physics responsible for the invariance of the characteristic current densities, the reader is referred to the Appendix.

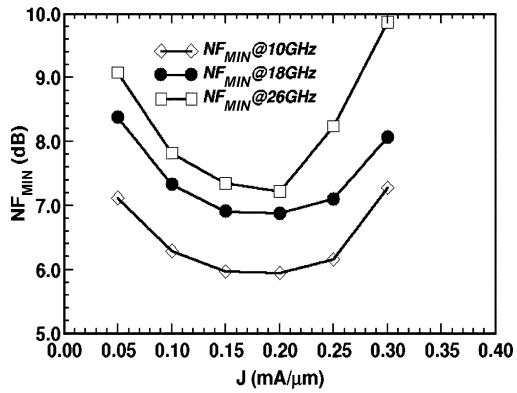
B. Topologies

Since both p-channel and n-channel MOSFETs exhibit these properties, one would expect that circuit topologies realized with combinations of MOSFETs will behave similarly. Indeed, this is confirmed by the measured f_T and $\mathrm{NF_{MIN}}$ characteristics of a 90-nm nMOS cascode, shown in Fig. 8 as functions of the drain current density. J_{pfT} and J_{OPT} remain unchanged from those of the single transistor. Furthermore, J_{OPT} is independent of frequency in the cascode stage as well as in a nMOS TIA stage with active pMOS load (Fig. 9). As indicated in the TIA schematics of Fig. 9(a), the size of the pMOSFET is set to ensure that it, too, is biased at its optimum noise current density.

With proper sizing and biasing of the SiGe HBT, the invariance of J_{pfT} and J_{pfMAX} can be extended even to MOS-HBT cascodes. For example, we recently reported a novel approach of biasing MOS-HBT cascodes at J_{pfMAX} to maximize the unity gain bandwidth (UGB) of operational amplifiers and to

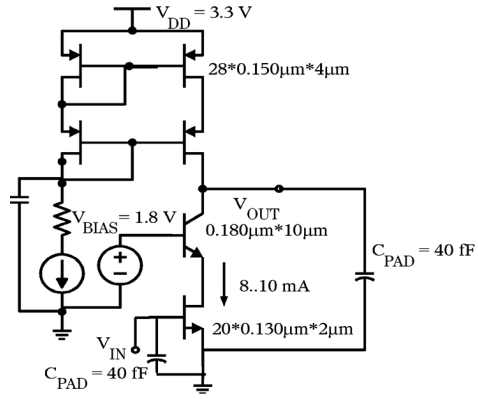


(a)

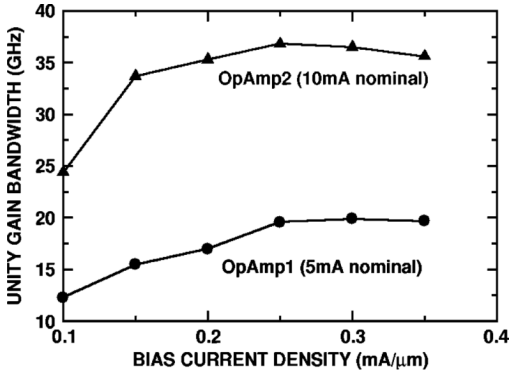


(b)

Fig. 9. (a) Schematic and (b) measured minimum noise figure versus drain current density for a single-ended 90-nm nMOS TIA at 10, 18, and 26 GHz.



(a)



(b)

Fig. 10. (a) BiCMOS operational-amplifier half-circuit schematic and (b) measured op-amp unity gain-bandwidth as a function of nMOSFET current density.

make it robust to process and bias current variation [13]. The schematic of the opamp half-circuit with cascode pMOS load is illustrated in Fig. 10(a). Both 5- and 10-mA versions of the half-circuit test structure [13] were fabricated in a 130-nm SiGe BiCMOS technology. The unity gain bandwidth—as high as 37 GHz—was measured on-wafer and plotted as a function of the nMOSFET current density in Fig. 10(b). The UGB of both versions varies by less than 15% when the bias current changes from 0.15 mA/μm to 0.35 mA/μm, making the design extremely robust to bias current and process variation which plagues analog design in nanoscale technologies. The same circuit was ported to a 180-nm SiGe BiCMOS technology with only minor changes [14]. The latter involve the resizing of the nMOSFET gate width from 20 μm to 25 μm in order to maintain the same gain and UGB. The current density of 0.2 mA/μm was preserved.

From these discussions and experimental evidence, it can be inferred that topologies such as the nMOS cascode, the nMOS-HBT cascode, the nMOS inverter with active pMOS load and the CMOS inverter can be treated in circuit design as composite transistors. When transistors are appropriately ratioed, each of these topologies can be described by its own f_T, $f_{\rm MAX}$, and $\rm NF_{MIN}$, which are usually different from those of the corresponding nMOSFET. For example, as seen in Fig. 8, the f_T of the nMOS cascode is about 40% lower than that of the transistor, while the f_T of the CMOS inverter is approximately 66% of the f_T of the nMOSFET [15]. However, Figs. 8, 9, and 10 indicate that the characteristic current densities of all these topologies are identical to those of the nMOSFET and remain invariant across technology nodes.

While conventional analog design in deep-submicron technologies is predicated on biasing at lower effective gate voltages or current densities to achieve higher DC gain [16], it becomes apparent that biasing in this region can yield circuits for which the noise figure and high frequency gain are rapidly changing with bias current, and which are more susceptible to temperature and statistical process variations affecting gate length and threshold voltage. In contrast, by employing constant-current-density biasing techniques in the peak-f_T or peak-$f_{\rm MAX}$ region, robust, bias-current-insensitive and process-insensitive analog, mm-wave, and high-speed designs can be achieved.

III. Scaling of Broadband Low-Noise and High-Speed Building Blocks

We have provided evidence in Section II for the invariance of the characteristic current densities of MOSFETs and of composite MOSFET topologies across technology nodes when MOSFETs are scaled according to the constant field scaling

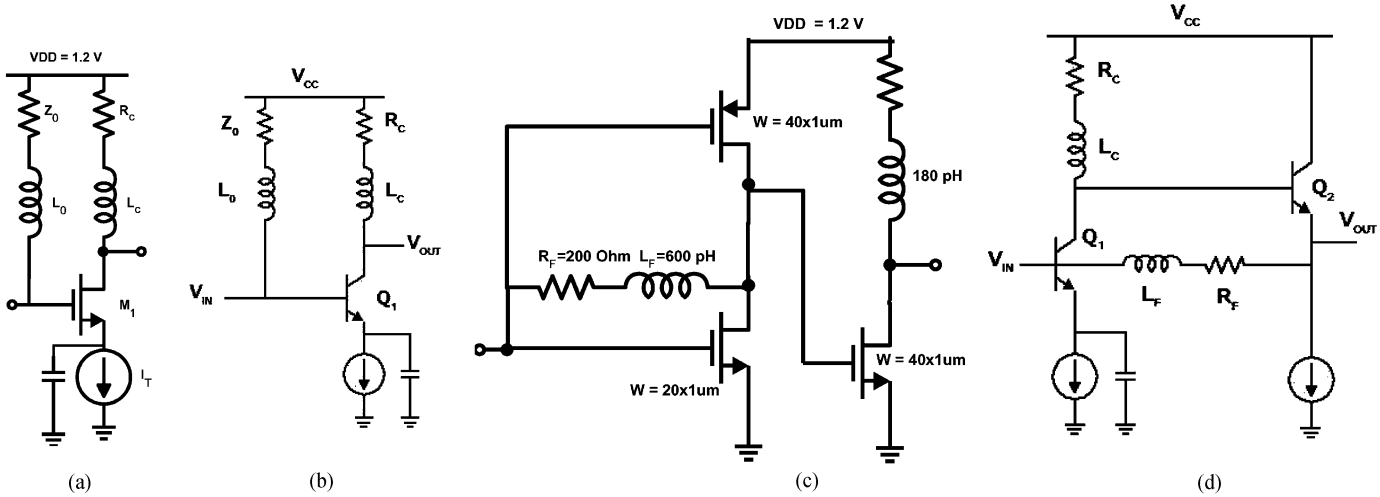


Fig. 11. (a) nMOS inverter. (b) SiGe HBT inverter. (c) CMOS-inveter TIA. (d) SiGe HBT TIA.

rules. Next, we demonstrate that the invariance of $J_{\rm OPT}$, J_{pfT}, and of the MOSFET capacitances per total gate width C_{gs}/W_G, C_{gd}/W_G, and C_{db}/W_G, allows for the scaling of optimally designed broadband low-noise amplifiers, CML gates and output drivers from one technology node to the next. This can be accomplished without adjusting the transistor size or bias current while improving performance. The application of the invariance of the $J_{\rm OPT}$ of MOS-MOS and MOS-HBT cascodes to tuned low-noise amplifiers and low-phase noise oscillators is discussed elsewhere [15], [17], [18].

A. Low-Noise Broadband Input Stages

As data rates increase and noise is integrated over larger bandwidths, it becomes critical to minimize internal circuit noise. Input amplifiers must have high bandwidth, low noise figure to improve sensitivity, and broadband impedance matching to avoid intersymbol interference stemming from signal reflections. Additionally, noise-impedance matching the input stage to the optimal source impedance $z_{\rm SOP}$ [19] will ensure the lowest possible noise figure. Meeting these goals simultaneously is not a trivial task, and requires attention to both device and topology optimization.

NMOSFET or SiGe HBT CML inverters (INV) with on-chip resistive matching, shown in Fig. 11(a) and (b), respectively, have been traditionally used as input comparators. Their noise factor can be expressed as a function of the signal source impedance Z_0 and of the transistor noise parameters:

$$F(Z_0) = 1 + \frac{1}{1+\left(\dfrac{\omega L_0}{Z_0}\right)^2} + R_N Z_0 \left| Y_{\rm COR} + \frac{2}{Z_0}\right|^2 + G_N Z_0. \tag{1}$$

The transistor noise resistance R_N, noise conductance G_N, and correlation admittance $Y_{\rm COR} = G_{\rm COR} + jB_{\rm COR}$ are functions of the device size and frequency [20]. These noise parameters can be recast as $R_N = R/W_G$, $G_N = G\omega^2 W_G$, $G_{\rm COR} = G_C \omega W_G$, and $B_{\rm COR} = B\omega W_G$, where G, R, G_C, and B are technology-dependent constants which characterize the geometry dependence at a given bias current density [20], [21], and W_G is the MOSFET gate width. Note that W_G is replaced with l_E in the case of an HBT circuit [1]. An optimal emitter length $l_{\rm EOPT}$ or gate width $W_{\rm GOPT}$ can be derived which minimizes the noise factor in (1) at frequencies up to ω:

$$l_{\rm EOPT}/W_{\rm GOPT} = \frac{1}{\omega}\frac{2}{Z_0}\sqrt{\frac{1}{\frac{G}{R}+G_C^2+B^2}}. \tag{2}$$

Instead of using an inverter topology as a low-noise amplifier, shunt-shunt (transimpedance) feedback can be employed when the optimal source impedance $z_{\rm SOP}$ of the transistor is higher than that of the generator. The noise factor of the TIAs in Fig. 9(a) and Fig. 11(c) and (d) is determined by considering the series combination of the feedback resistor R_F and feedback inductor L_F as a parallel feedback network across the transistor amplifier. The expression for the noise factor as a function of Z_0 and of the noise parameters of the composite transistor topology becomes [21]

$$F(Z_0) = 1 + R_N Z_0 \left| Y_{\rm COR} + \frac{1}{Z_0} + \frac{1}{R_F}\frac{1-j\omega_0}{1+\omega_0^2}\right|^2 + Z_0 G_N + \frac{Z_0}{R_F}\frac{1}{1+\omega_0^2} \tag{3}$$

with $\omega_0 = \omega L_F/R_F$. Again, an optimal emitter length or gate width can be derived at the angular frequency ω [21]:

$$l_E/W_{GOPT} = \frac{1}{\omega}\left[\frac{1}{Z_0}+\frac{1}{R_F}\frac{1}{1+\omega_0^2}\right]\sqrt{\frac{1}{\frac{G}{R}+G_C^2+B^2}}. \tag{4}$$

Comparing (2) and (4), and knowing that R_F is typically several times larger than Z_0, it follows that, by using shunt feedback for noise impedance matching, the size and bias current of the input transistor will be smaller than in the inverters of Fig. 11(a) and (b), leading to lower power dissipation and broader bandwidth [1].

Equations (2) and (4) point to a straightforward methodology for designing low-noise inverters or TIAs. First, $J_{\rm OPT}$ is determined. While $J_{\rm OPT}$ remains approximately 0.15 mA/μm for an

TABLE I
DESIGN PARAMETERS FOR NMOS AND CMOS TIAS IN 130-nm, 90-nm, AND 65-nm TECHNOLOGIES

	130nm CMOS	*90nm CMOS*	*90nm nMOS*	*65nm CMOS*	*65nm nMOS*
W (μm)	30 μm	30 μm	80 μm	30 μm	80 μm
R_F (Ohm)	200	200	200	200	200
I_{DS}(TIA)	6 mA	6 mA	16 mA	6 mA	16 mA
f_{3dB}(sim)	15.7 GHz	30.6 GHz	29.6 GHz	57.8 GHz	49. 4 GHz
Gain (sim)	9 dB @10GHz	8.4 dB	7.5 dB	8.9 dB	7.5 dB
V_{DD} (V)	1.4 V	1.2 V	1. 0V	1.2 V	1 V

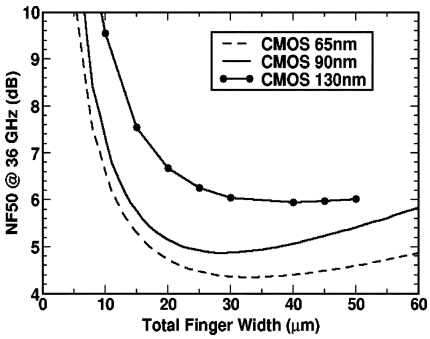
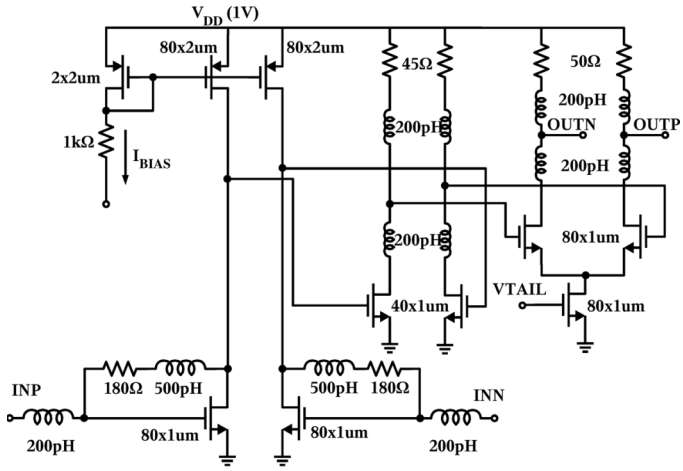
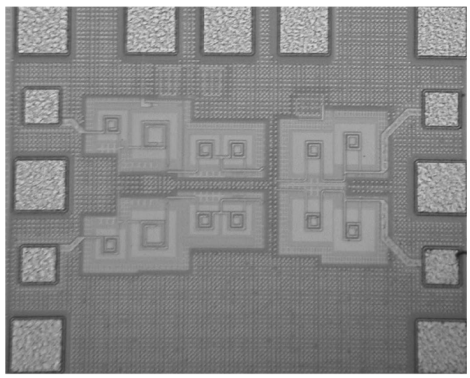


Fig. 12. Scaling of CMOS TIA from 130-nm to the 90-nm and 65-nm nodes.



(a)



(b)

Fig. 13. (a) Schematic and (b) layout of differential 90-nm nMOS TIA.

nMOSFET irrespective of frequency as discussed in Section II, it becomes a function of frequency in SiGe HBTs and should be selected at the desired 3-dB bandwidth of the amplifier ω [1], [21]. Technology constants R, G, G_C, and B can then be found for this bias point. Next, transistor Q1 or M1 is biased at $J_{\rm OPT}$ and sized using (2) or (4). Finally, concomitant noise and impedance matching can be achieved through loop gain optimization in the TIA, as described in [1], or by using a 50-Ω resistor at the input, in the case of the inverter.

Simulation results exploring the sizing and scaling of nMOS and CMOS TIA designs in three technology nodes are summarized in Table I and illustrated in Fig. 12. In all cases, a current density of 0.2 mA/μm was employed, corresponding to the peak $f_{\rm MAX}$ bias and close to the optimal noise bias. More interestingly, since $J_{\rm pfMAX}$, $J_{\rm OPT}$ and transistor capacitances per unit gate width are invariant across technology nodes, the square root term in (2) and (4), which depends solely on current density, also remains invariant across technology nodes. Therefore, the size and bias current of MOSFETs in the TIAs are practically constant from one technology node to the next while the noise figure and bandwidth are improved as designs are scaled.

From Table I, it becomes apparent that, if the f_T is adequate for the application, the CMOS inverter will require approximately 1/3 the size and bias current of an nMOSFET fabricated in the same technology node to achieve a certain noise resistance and optimal noise impedance while only a relatively small degradation in noise figure is incurred. This surprisingly little-known property of the CMOS inverter can significantly reduce the power dissipation of tuned [22] and broadband low-noise amplifiers.

To verify these findings, differential and single-ended versions of the low-noise broadband amplifier topologies presented in Fig. 9(a) and Fig. 11 were fabricated in 130-nm SiGe BiCMOS [29], 90-nm bulk and SOI CMOS technologies. Note that all circuits include a 50-Ω output driver. In each case, the TIA stage is biased close to $J_{\rm OPT}$. Fig. 13 shows the schematics and die photograph of the 90-nm bulk CMOS differential TIA. The single-ended S-parameters were measured on wafer and are reproduced in Fig. 14. The bandwidth is larger than 26 GHz and both S_{11} and S_{22} are better than -10 dB up to 30 GHz. The differential gain varies between 13 and 14 dB for $V_{\rm DD}$ values from 1 to 1.2 V. The measured single-ended eye diagrams at 25 Gb/s and 38.8 Gb/s are shown in Fig. 15 for 100 mV_{pp} single-ended inputs.

The 50-Ω noise figures for all nMOS, SOI and SiGe HBT amplifiers were measured up to 20 GHz, and are reported in Fig. 16.

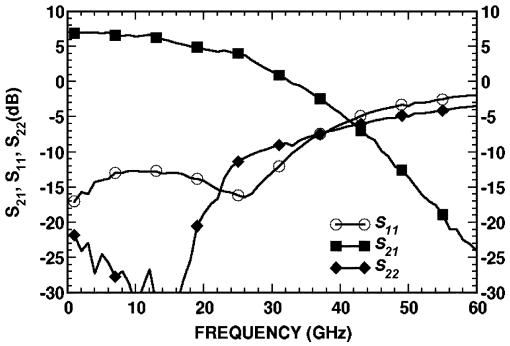
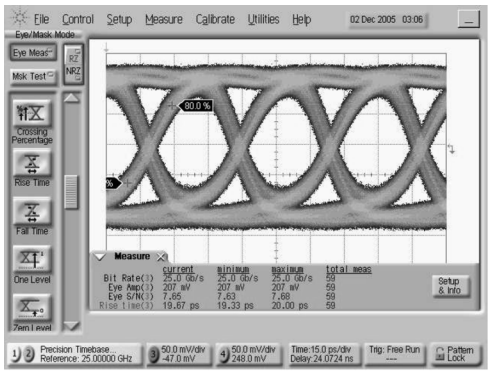
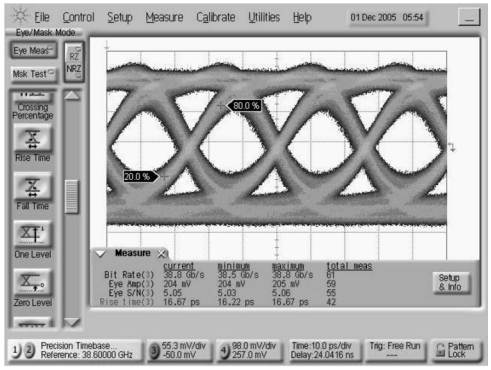


Fig. 14. Measured gain, input and output matching versus frequency for differential 90-nm nMOS TIA.



(a)



(b)

Fig. 15. 90-nm differential nMOS TIA 25 Gb/s and 38.8 Gb/s single-ended output eye diagrams with 100 mV_{pp} single-ended input.

All noise measurements are in single-ended mode with the unused input terminated in 50 Ω. The 90-nm nMOS and SiGe differential TIAs exhibit the lowest noise among the differential designs, both less than 10 dB at 10 GHz, with a slight advantage for the nMOS version. Fig. 16 also collects the noise figures of single-ended 90-nm bulk and SOI TIAs. As expected, their noise figures are 2–3 dB lower than those of the differential versions. Even though pMOSFETs are employed as active loads, these represent the highest bandwidth and lowest noise transimpedance amplifiers reported in CMOS [23]. When operated with only 6 mW power dissipation at 15 Gb/s, the 90-nm nMOS SOI TIA achieves a record 0.33 mW/Gbs [23].

When comparing the SiGe TIA and the SiGe inverter performance, simulations show and measurements confirm that the

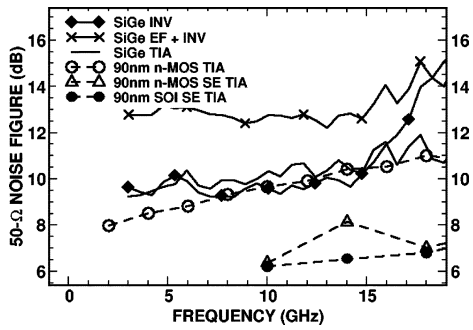
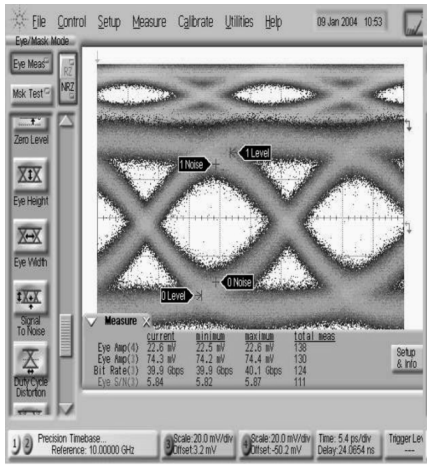
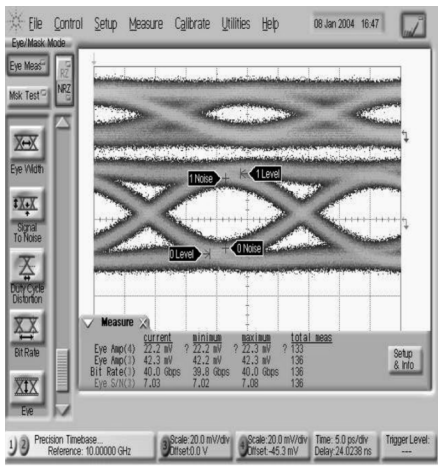


Fig. 16. Measured 50-Ω noise figures for various SiGe and nMOS differential and single-ended (SE) broadband input stages.



(a)



(b)

Fig. 17. 40-Gb/s 20-mV input (top) and corresponding output (bottom) eye diagrams for SiGe HBT (a) EF-INV and (b) TIA low noise input amplifiers. Output eye diagram Q factors are 5.8 and 7, respectively.

TIA has significantly better bandwidth (40 GHz compared to 11 GHz) and broadband input matching. Their noise figures are comparable below 15 GHz. To improve bandwidth, emitter-followers (EF) can be added to the input of a low-noise inverter, but at the expense of higher noise as seen in Fig. 16. The lower noise figure of the TIA results in higher sensitivity than that of the EF-INV even though the latter has larger gain. As demonstrated in Fig. 17(a), the EF-INV output eye diagram has a Q factor

of 5.8 for a 20-mVpp single-ended input (10-mVpp per side). The SiGe HBT TIA eye diagram of Fig. 17(b) has a Q factor of 7 for the same input amplitude while consuming 50 mW, 20 mW less than the EF-INV stage. These results prove the direct link between noise figure and sensitivity and the importance of low-noise design in wireline applications. We also note that, because of the larger bandwidth, the SiGe HBT TIA operates with higher sensitivity at higher data rates than the 90-nm nMOS and SOI TIAs.

B. Current-Mode Logic and Output Drivers

The open-circuit time constant (OCTC) of a chain of differential MOS-CML inverters and MOS-HBT (BiCMOS) cascode inverters [24] with a fanout of k can provide a useful metric for the ultimate digital speed of CMOS and SiGe BiCMOS technologies.

$$\tau_{MOSCML} = \frac{\Delta V}{I_T}\left[C_{gd}+C_{db}+\left(k+\frac{R_g}{R_L}\right)[C_{gs}+(1+g_mR_L)C_{gd}]\right] \quad (5)$$

$$\tau_{BiCMOSCML} \approx \frac{\Delta V}{I_T}\left[C_{\mu}+C_{cs}+\left(k+\frac{R_g}{R_L}\right)(C_{gs}+C_{gd})\right] \quad (6)$$

where I_T is the tail current, R_L is the load resistance, and ΔV is the logic swing. In (6) we have accounted for the fact that the g_m of the HBT is at least 10 times larger than that of the MOSFET at the same bias current. In HBT CML/ECL gates, the logic swing is typically set between 200 and 300 $\mathrm{mV_{pp}}$, independent of the technology node [25]. In MOS-CML or BiCMOS-CML, the minimum swing ΔV_{MIN} is determined from the condition that the MOS differential pair be fully switched:

$$\Delta V_{\mathrm{MIN}} = 2\left[V_{GS}(I_T) - V_{GS}\left(\frac{I_T}{2}\right)\right]. \quad (7)$$

Fig. 18(a) illustrates the measured fanout-of-1 delay as a function of the tail current density for MOS-CML inverters fabricated in the bulk 180-nm, 130-nm, and 90-nm nodes and for 90-nm SOI CMOS. For comparison, the delay of BiCMOS-CML cascode inverters is shown in Fig. 18(b). The delays were obtained by applying (5) and (6) to the measured small-signal parameters of nMOSFETs and SiGe HBTs biased at $I_T/2$. $V_{GS}(I_T)$ and $V_{GS}(I_T/2)$ used in (7) were determined from the measured DC characteristics. The SiGe HBTs used in the 180-nm, 130-nm and 90-nm nodes have similar performance, with an f_T of 160 GHz. While the 180-nm and 130-nm SiGe BiCMOS technologies are real, the 90-nm one is just a hypothetical process. All CMOS and BiCMOS inverters were designed to have a gain of 1.5 which ensures operation in a ring oscillator chain. Therefore, $\Delta V = 1.5^{*}\Delta V_{\mathrm{MIN}}$.

These results show that, in all CML gates, the tail current density corresponding to minimum gate delay is approximately 0.3 $\mathrm{mA}/\mu\mathrm{m}$, closely tracking the peak f_T current density. Furthermore, the delay changes by less than 10% when the tail current varies between 0.15 $\mathrm{mA}/\mu\mathrm{m}$ and 0.5 $\mathrm{mA}/\mu\mathrm{m}$. The results in Fig. 18 and the discussion in Section II provide a rigorous basis for the optimal sizing of MOSFETs and HBTs in (Bi)CMOS-CML gates [24], [25]

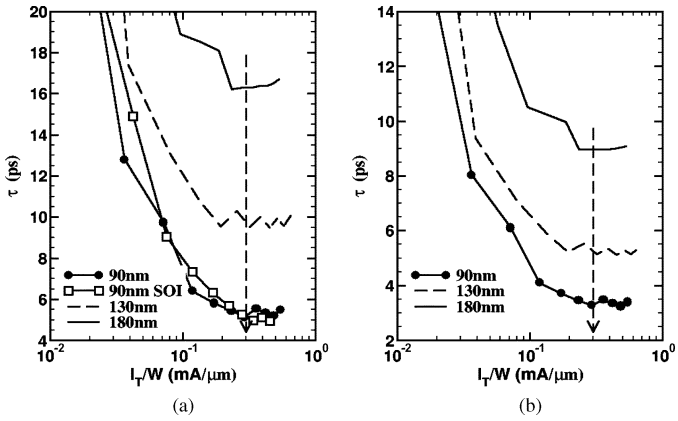


Fig. 18. (a) MOS-CML and (b) SiGe BiCMOS-CML fanout-of-1 delays as a function of tail current density for the 180-nm, 130-nm, and 90-nm nodes. The same $0.18\ \mu\mathrm{m} \times 2\ \mu\mathrm{m}$ SiGe HBT with 160 GHz f_T has been used in all nodes. The data were obtained by applying equations (5) and (6) to the DC and small-signal parameters of transistors extracted from S parameters measurements up to 65 GHz.

$$W_G = \frac{I_T}{J_{pfT}} = \frac{I_T}{0.3\ \mathrm{mA}/\mu\mathrm{m}},\ A_E = \frac{I_T}{1.5J_{pfT}} \quad (8)$$

where the W_G is the total gate width and A_E is the emitter area.

If the MOSFET is sized according to (8), this corresponds to $\sqrt{2}V_{\mathrm{EFF}} \leq \Delta V_{\mathrm{MIN}} \leq 2V_{\mathrm{EFF}}$ where V_{EFF} is the effective gate voltage of each transistor in the diff pair at $I_T/2 = 0.15\ \mathrm{mA}/\mu\mathrm{m}$. Both V_{EFF} at $0.15\ \mathrm{mA}/\mu\mathrm{m}$ and ΔV_{MIN} decrease by approximately $\sqrt{2}$ with each new technology node. The measured ΔV_{MIN} corresponding to the minimum delay bias is approximately 600 mV, 400 mV, and 320 mV in the 180-nm, 130-nm, and 90-nm nodes, respectively. Since the optimal-delay current density remains constant, it becomes possible to port designs from one technology node to the next, using the same tail current and device size while obtaining better performance due to the reduction in $\Delta\mathrm{V}$, as shown in Fig. 18. From another perspective, to achieve the same switching speed, the tail current can be reduced by $\sqrt{2}$ as technology scales, which, when combined with the lower supply voltage, allows for the power dissipation to be halved. We note from Fig. 18 that by adding the SiGe HBT option to a CMOS technology, the speed of CML gates is almost doubled, arguably a more economical solution than advancing one or two nodes along Moore's law.

Further power savings can be realized in both CMOS and BiCMOS circuits through the use of inductive peaking. Assuming that the output time constant in a CML gate dominates the overall bandwidth, the latter can be estimated as

$$BW_{3\mathrm{dB}} = \frac{1}{2\pi R_L C_L} = \frac{I_T}{2\pi C_L \Delta V}. \quad (9)$$

Shunt inductive peaking can improve bandwidth by up to 60% with constant group delay [26]:

$$L_{P,\mathrm{MAX}} = \frac{C_L R_L^2}{3.1} = \frac{C_L}{3.1}\left(\frac{\Delta V}{I_T}\right)^2 \tag{10}$$

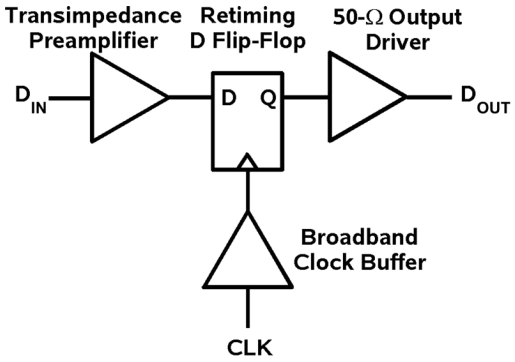
If $L_{P,\mathrm{MAX}}$ is the maximum inductance that can be realized in a particular backend with adequate self resonance frequency (SRF) for the intended data rate, the minimum tail current $I_{T,\mathrm{MIN}}$ to meet the bandwidth requirements is

$$I_{T,\mathrm{MIN}} = \Delta V \sqrt{\frac{C_L}{3.1\, L_{P,\mathrm{MAX}}}}. \tag{11}$$

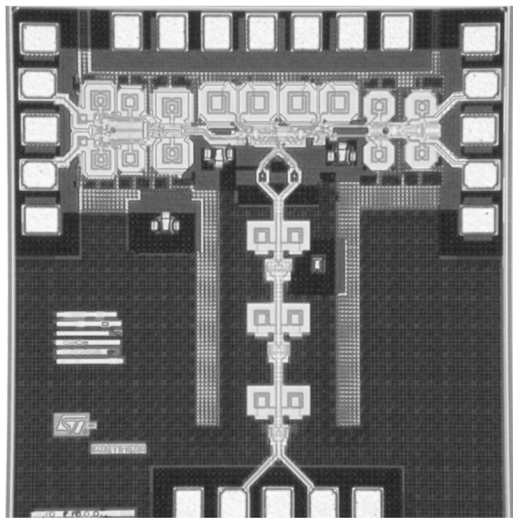
This reduction in power consumption can be realized without significant increase in die area by making use of stacked (or 3-D) inductors. As the number of metal layers continues to increase in modern backends, larger inductances can be implemented in a small area. Furthermore, as the inductor footprint over the substrate is reduced, the self-resonant frequency is improved making this technique effective even at very high data rates [2], [27]. Note that shunt-series peaking occurs almost by default due to the inductance of interconnect between stages. Hence, an even larger improvement in bandwidth is regularly obtained [2] without the need for area-intensive t-coils [28]. Equations (9)–(11) provide the underlying reasons why, for a given backend characterized by a fixed $L^*\mathrm{SRF}$ product, using bipolar devices with lower logic swing and output capacitance will result in smaller tail currents and lower power dissipation despite the higher supply voltage requirements.

An advantage of deep-submicron and nanoscale MOSFETs biased in the saturation region is that the device capacitances, g_m, f_T, and f_{MAX} vary little with respect to bias current or gate–source voltage once the effective gate voltage exceeds 250 mV. In addition, as we saw in Fig. 18, for tail current densities in the 0.15 mA/μm to 0.5 mA/μm range, the CML gate delay remains practically constant. This allows for the elegant implementation of output swing control in high-speed I/O drivers, without requiring source degeneration and without compromising bandwidth and the quality of the eye diagram.

To validate the scaling of CML gates and output stages, a 30-Gb/s backplane driver with controllable pre-emphasis [30] was designed and fabricated in a 130-nm SiGe BiCMOS technology [29] using only nMOSFETs and no SiGe HBT. MOSFETs in switching differential pairs were sized according to (8) and multi-metal 3-D inductors were employed for shunt-peaking bandwidth extension. Operation at 20 Gb/s with pre-emphasis, and up to 30 Gb/s with 2×300 $\mathrm{mV_{p-p}}$ swing was measured [30]. The circuit consumes 150 mW from a 1.5-V supply and marks the fastest driver with waveshape control reported to date in CMOS. For comparison, we note that 37.5 Gb/s operation was recently announced for a 130-nm CMOS transmitter by another group [28]. To reach this speed the authors employed series-shunt peaking, while waveshape and pre-emphasis control (which slows down the output node) was not included. The MOS-CML gates were biased at 0.15 mA/μm and the internal voltage swing was set to 500 $\mathrm{mV_{pp}}$ [28-slide set]. According to Fig. 18(a), this bias condition will result in at most 10% degradation in delay. The latter is compensated by series-shunt peaking (1.4 times more bandwidth than simple shunt peaking) and by the lower output swing of 2×260 mVpp.



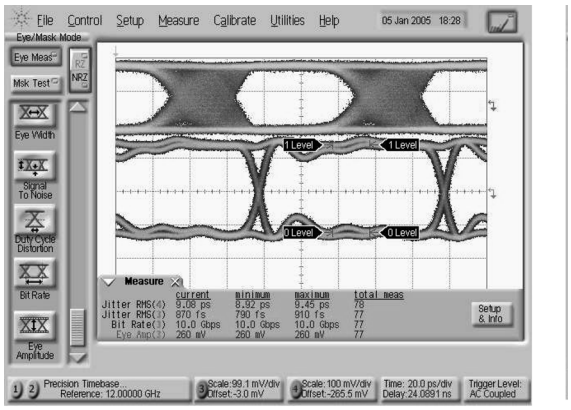
(a)



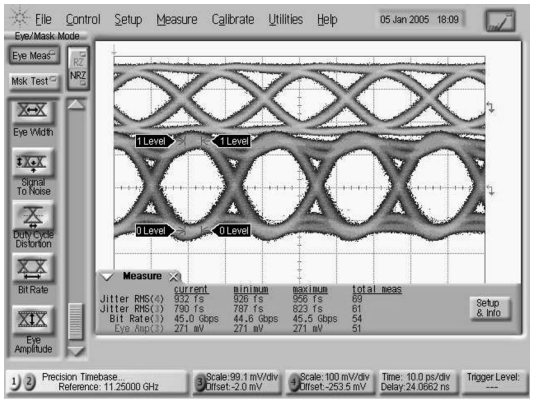
(b)

Fig. 19. (a) Schematic and (b) layout of broadband 49-Gb/s retimer.

Another experiment was conducted to verify the performance improvement that can be realized in a given technology node by the addition of a SiGe HBT option to a CMOS technology. A new 2.5-V, 49-Gb/s SiGe BiCMOS broadband retimer was designed and fabricated in the same 130-nm SiGe BiCMOS process as the 30-Gb/s CMOS driver. As shown in Fig. 19, the retimer employs the SiGe-HBT TIA discussed earlier, a SiGe BiCMOS flip-flop, a MOS-HBT output driver with series-shunt peaking and 5.5-ps rise time, and a 2.5-V broadband clock path consisting of three EF-INV stages that can be driven with a single-ended clock signal at frequencies up to 49 GHz. Eye diagrams at 12, 45, and 49 Gb/s with variable output swing between 2×300 $\mathrm{mV_{pp}}$ and 2×600 $\mathrm{mV_{pp}}$ are reproduced in Fig. 20 and Fig. 21. This is $2\times$ faster than the CMOS flip-flop in [28], and the fastest retimer and flip-flop operating in silicon at power supplies below 3.3 V. It proves that CML gate speed is doubled with respect to a CMOS-only circuit fabricated in the same node [28], [30], as predicted in Fig. 18.

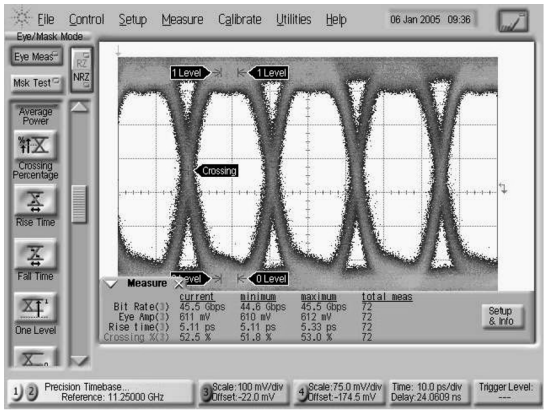


(a)

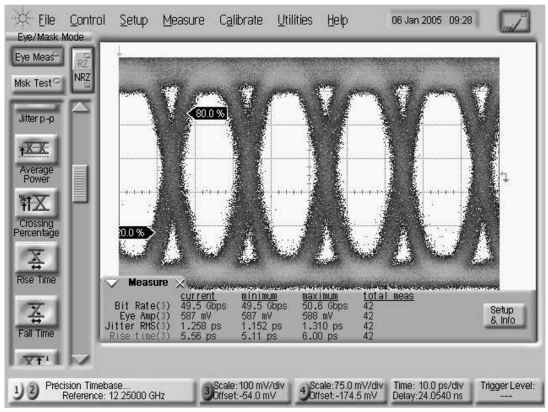


(b)

Fig. 20. (a) 10 Gb/s and (b) 45 Gb/s input (top) and 2 × 280 mVpp output (bottom) after retiming.



(a)



(b)

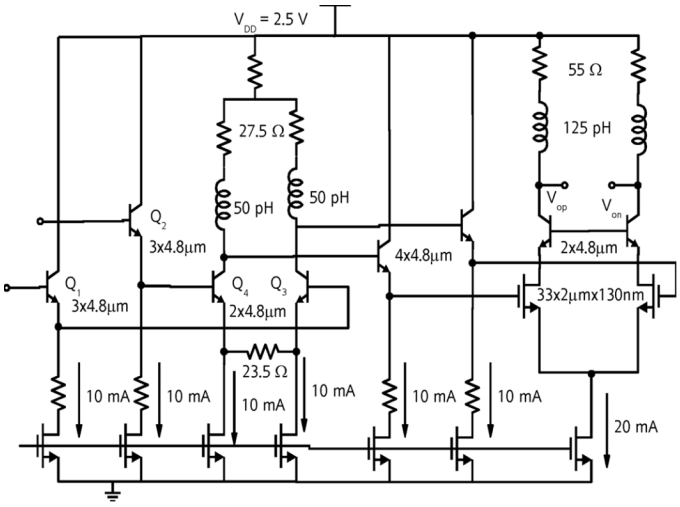
Fig. 21. (a) 45 Gb/s and (b) 49 Gb/s 2×600 mVpp retimed output.

Finally, an 80-GHz, 2.5-V BiCMOS output driver with preemphasis is presented in Fig. 22. The circuit was fabricated in the same 130-nm SiGe BiCMOS technology with 150-GHz f_T SiGe HBT [29], and consumes 200 mW. The output stage consists of a series-shunt inductively-peaked BiCMOS cascode inverter. The measured differential gain S_{21}, shown in Fig. 23, increases linearly by 7 dB from 10 GHz to 65 GHz, peaking above 10 dB in the 65- to 75-GHz range. More than 10 dB of gain control is realized as the tail current in the output stage is varied between 0.15 mA/μm and 0.5 mA/μm, with no noticeable change in bandwidth and degree of peaking. Moreover, because the output impedance of the MOS-HBT differential pair does not vary with bias current, the output return loss remains relatively constant and is lower than -10 dB up to 94 GHz as the tail current is adjusted, as also seen in Fig. 23. To the best of the authors' knowledge, this is the highest bandwidth broadband output driver and the fastest digital circuit using bulk MOSFETs on the high-speed path. Because it employs only lumped inductors smaller than 20 μm per side, the circuit occupies a die area smaller than 200 μm $\times$200 μm.

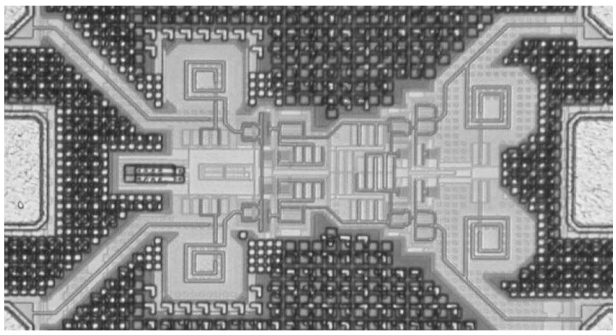
IV. Conclusion

Experimental data have proved that the peak f_T, peak f_{MAX}, and optimum $\mathrm{NF}_{\mathrm{MIN}}$ current densities are largely invariant over foundries, technology nodes, and basic circuit topologies, while the corresponding effective gate bias varies widely. Moreover, in the 90-nm bulk and SOI nodes, these current densities remain unchanged for gate lengths spanning the 90-nm to 350-nm range, covering virtually all analog and mixed signal applications. It was proposed, based on a simple analytical model, that this behavior is a direct result of carrier mobility degradation due to the large vertical electric field at the Si/SiO_2 interface in the MOSFET inversion layer. The electric field profile has so far been preserved across technology nodes due to the constant-field scaling rules being applied rather rigorously by all foundries following the ITRS. The exact values of the peak f_{MAX}, $\mathrm{NF}_{\mathrm{MIN}}$, and peak f_T current densities are determined by the interplay between the bias dependence of transconductance (dominated by mobility degradation due to the vertical field), output conductance, and gate–source and gate–drain capacitance per micron of gate width. Therefore, it is recommended that constant-current-density biasing be employed to design robust analog, mixed-signal, RF, and high-speed CMOS circuits that are less sensitive to process, temperature, and bias current variations, and which can be easily scaled from one technology node to another.

Algorithmic design and scaling methodologies, based on the invariance of the characteristic current densities in deep-submicron and nanoscale MOSFETs, were developed for the main circuit building blocks of a wireline transceiver. Remarkably, CMOS low-noise amplifiers, CML gates and output drivers have



(a)



(b)

Fig. 22. (a) Schematic and (b) layout of 80-GHz BiCMOS output driver with pre-emphasis and output swing control.

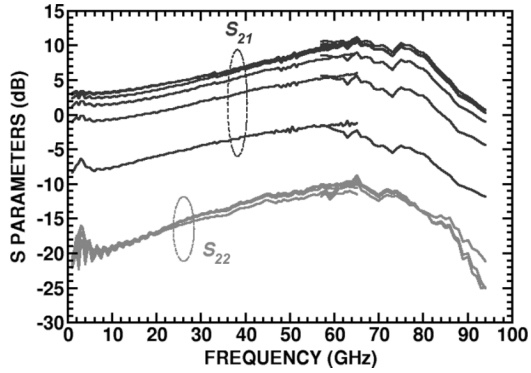


Fig. 23. Measured differential gain and output return loss of the 80-GHz BiCMOS output driver.

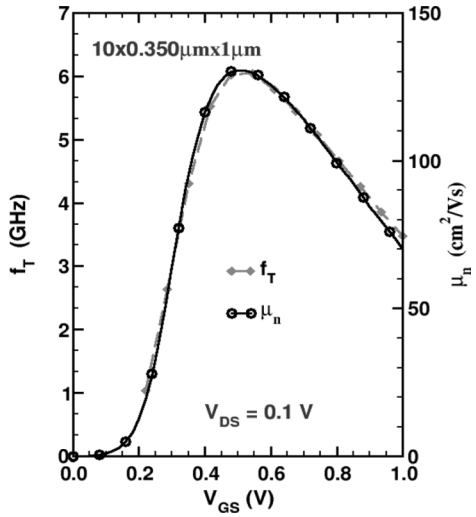


Fig. 24. Measured μ_n versus V_{GS} (extracted from the linear region of the $I_D - V_{GS}$ characteristics) and f_T versus V_{GS} characteristics at $V_{DS} = 0.1$ V for a 90-nm node nMOSFET with 350-nm gate length.

been shown to scale almost unchanged in terms of transistor size and bias current from 130-nm to 90-nm nodes while their noise, bandwidth and delay are improved. The first noise and input impedance-matched 90-nm CMOS single-ended and differential TIAs with over 26-GHz bandwidth and less than 8-dB noise figure were demonstrated. Eye diagrams were measured at data rates up to 38.8 Gb/s with 1-V supply. The algorithmic design methodology was also verified on MOS-HBT cascode operational amplifiers with up to 37-GHz unity gain bandwidth, 40-GHz SiGe HBT broadband low-noise amplifiers, a 49-Gb/s retimer and on an 80-GHz output driver with pre-emphasis fabricated in SiGe BiCMOS technology and operating from 2.5-V supply. As predicted by simple MOS-CML and BiCMOS-CML delay equations, the addition of a SiGe HBT option to a CMOS technology was experimentally verified to double the CML gate switching speed.

APPENDIX

Velocity Saturation or Mobility Degradation?

A well-known result of the constant-field scaling rules is that the channel inversion charge $Q_i/W = C_{\rm OX}\,(V_{GS} - V_T)$, gate capacitance $C_{\rm OX}\,L_{\rm min}$, and drain current per micron of total gate width, in either long-channel $I_{DS}/W = \mu_n E\,Q_i/W$ or in saturated velocity regime $I_{DS}/W = v_{\rm nsat}Q_i/W$, remain constant [4]. $L_{\rm min}$ represents the minimum gate length in a particular technology node. Furthermore, the shape of the mobility versus lateral and vertical electric field characteristics, and the saturation velocity, are universal material properties [3], [4]. It then follows that, if the two-dimensional electric field profile is preserved over technology nodes by foundries rigorously applying the ITRS [5] constant-field scaling guidelines, inversion charge, gate capacitance and drain current per micron of gate width will remain constant over technology nodes even in devices that exhibit significant mobility degradation and only a low degree of velocity saturation. The latter situation more accurately depicts deep-submicron MOSFETs.

The universal surface mobility curve for electrons in silicon can be expressed as [4]

$$\mu_{\rm n}(V_{GS}) = \frac{\mu_{\rm ac}\mu_{\rm sr}}{\mu_{\rm ac} + \mu_{\rm sr}} \tag{A1}$$

where $\mu_{\rm ac}[{\rm cm}^2/{\rm Vs}] \approx 330\ E_{\rm eff}^{-1/3}[{\rm MV/cm}]$ $\mu_{\rm sr}[{\rm cm}^2/{\rm Vs}] \approx 1450\ E_{\rm eff}^{-2.9}[{\rm MV/cm}]$ and $E_{\rm eff}$ is the effective vertical field. For devices fabricated in the 180-nm, 130-nm and 90-nm nodes, $E_{\rm eff}$ falls in the 1–2 MV/cm range where the surface scattering-

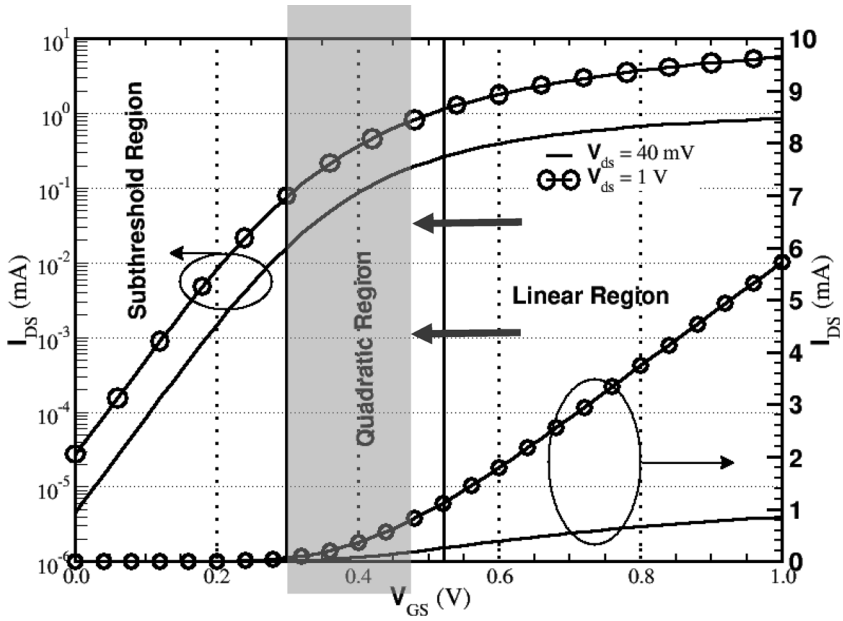


Fig. 25. Measured DC transfer characteristics of a 10×1 μm 90-nm nMOSFET.

limited mobility term μ_{sr} dominates over the acoustic-phonon-limited mobility $\mu_{\rm ac}$ [3] and leads to a rapid reduction in electron mobility as the gate voltage increases. Recent experimental data presented in the literature [3] indicate that the electron mobility in strained silicon channels, notwithstanding larger overall values, has the same dependence on the vertical field, with the onset of surface scattering occurring in the 1–2 MV/cm range.

A typical $\mu_n(V_{GS})$ curve, measured on n-FETs with 350-nm gate length using a technique based on the *I–V* characteristics at low V_{DS} as in [3], is reproduced in Fig. 24. This device is fabricated in the 90-nm node. Above $V_{GS} = 0.55$ V, electron mobility linearly decreases with increasing gate voltage. An identical shape is measured for the f_T versus V_{GS} characteristics at low V_{DS}, also shown in Fig. 24 on the left-side *Y* axis.

A closed-form expression for drain current that lumps together the vertical and lateral field dependence of mobility is present in all advanced MOSFET models [31], [32]. However, this has made it rather difficult in the past to actually pinpoint which effect is dominant. In deep-submicron MOSFETs mobility degradation due to the applied vertical field dominates [7]. The vertical field is typically one order of magnitude larger than the lateral electric field. In addition, the critical lateral field $E_C = v_{sat}/\mu_n$ for velocity saturation is not constant. It increases as the carrier mobility $\mu_{\rm n}$ degrades [4]. This phenomenon helps to delay the onset of velocity saturation in deep-submicron MOSFETs to ever larger lateral electric fields and larger V_{DS} as the gate voltage increases.

Analytical Derivation of Characteristic Current Density Scaling

It is helpful to capture the experimental trends described earlier in simple analytical equations that can provide insight for circuit design. The $I_{DS} - V_{GS}$ characteristics of very deep-submicron MOSFETs biased in saturation exhibit two distinct regions, as shown in Fig. 25. At low $V_{\rm eff}$ the well-known, long-channel square law applies with $\mu_n(V_{GS})$ remaining relatively constant with V_{GS}:

$$I_{\rm D} = \frac{\mu_{\rm n}(V_{GS})C_{\rm OX}}{2}\frac{W}{L}(V_{GS} - V_T)^2(1 + \lambda V_{DS}). \quad \text{(A2)}$$

At larger gate voltage, the characteristics become linear, as a result of mobility degradation. In the saturation region, the transconductance expression can be obtained by differentiating (A2) with respect to V_{GS}:

$$g_m \approx C_{\rm OX}\frac{W}{L}(V_{GS}-V_T)(1+\lambda V_{DS}) \times\left[\mu_{\rm n}(V_{GS})+\frac{\partial\mu_{\rm N}(V_{GS})}{2\partial V_{GS}}(V_{GS}-V_T)\right] \quad \text{(A3)}$$

$$\frac{\partial g_m}{\partial V_{GS}} \approx C_{\rm OX}\frac{W}{L}(1+\lambda V_{DS}) \times\left[\mu_{\rm n}(V_{GS})+2\frac{\partial\mu_{\rm N}(V_{GS})}{\partial V_{GS}}(V_{GS}-V_T) +1/2\frac{\partial^2\mu_{\rm N}(V_{GS})}{\partial V_{GS}^2}(V_{GS}-V_T)^2\right]. \quad \text{(A4)}$$

Equation (A4) indicates that, if the second derivative of mobility is neglected (a reasonable approximation), g_m reaches its peak when

$$V_{GS} - V_T \approx \frac{-\mu_{\rm n}(V_{GS})}{2\dfrac{\partial\mu_{\rm N}(V_{GS})}{\partial V_{GS}}}. \quad \text{(A5)}$$

Since the effective vertical field is a linear function of V_{GS} [4]

$$E_{\rm eff} = \frac{V_T + 0.2}{3t_{\rm OX}} + \frac{V_{GS} - V_T}{6t_{\rm OX}}. \quad \text{(A6)}$$

Equation (A5) can be recast as a function of $E_{\rm eff}$:

$$V_{GS} - V_T = V_{\rm effpeak}{}_{gm} \approx 3t_{\rm OX}\frac{-\mu_{\rm n}(E_{\rm eff})}{\dfrac{\partial\mu_{\rm N}(E_{\rm eff})}{\partial E_{\rm eff}}}. \quad \text{(A7)}$$

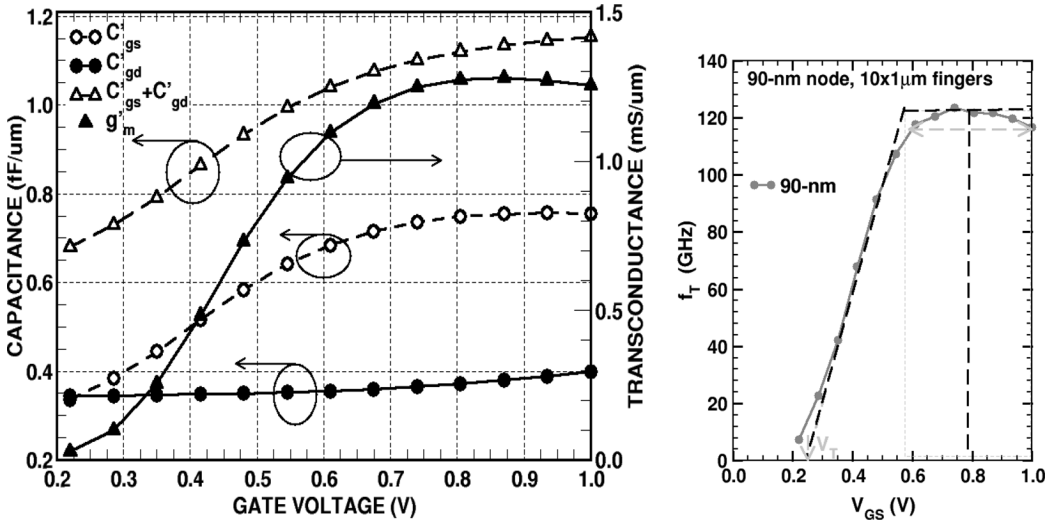


Fig. 26. Measured dependence of the small-signal transconductance, gate–drain and gate–source capacitances per unit gate width, and cutoff frequency as a function of the gate–source voltage in 90-nm nMOSFETs at $V_{DS} = 0.7$ V.

With the knowledge that the universal $\mu_n(E_{\rm eff})$ curve is a material property [3], [4], it follows that, at least in a first-order analysis, the effective gate voltage $V_{\rm effpeak}gm$ at which g_m reaches its peak value scales linearly with the gate oxide thickness, exactly as one would expect from constant field scaling rules. Substituting this effective gate voltage in the drain–current expression (A2) indicates that the drain current density I_{DS}/W at which peak g_m occurs is proportional to $t_{\rm ox}/L_{\rm min}$ and should remain constant over technology nodes if strict constant field scaling rules are applied. Beyond $V_{\rm effpeak}gm$, the transconductance is practically constant, although some interface scattering-induced degradation does occur. Therefore, as recently proposed in [16], the transconductance characteristics in the saturation regime can be divided into a region of linear dependence with V_{GS} (where intrinsic device voltage gain is relatively high as desired for opamp design) and one where g_m is flat. The simple $g_m - V_{GS}$ characteristics can then be integrated with respect to V_{GS} to obtain a very accurate estimate of the $I_{DS} - V_{GS}$ characteristics for deep-submicron MOSFETs [16].

The main FOMs for RF applications, f_T, $f_{\rm MAX}$, and minimum noise factor $F_{\rm MIN}$ can be expressed as functions of series resistances R_s, R_d, R_g, channel resistance R_i C_{gs}, g_{ds}, and g_m [33]. While exact analytical expressions of f_T, $f_{\rm MAX}$ and $F_{\rm MIN}$ as functions of all parameters of the small signal equivalent circuit are possible, they become too complex to interpret easily and are beyond our stated goal of providing an intuitive understanding of the observed experimental behavior. In order to make the analysis more tractable, the substrate network has been left out, resulting in the following expressions for f_T and $f_{\rm MAX}$ [34], [35]:

$$\frac{1}{2\pi f_T} = \frac{(C_{gs}+C_{gd})}{g_m} + (R_s+R_{\rm d})C_{gd} + (C_{gs}+C_{gd})(R_s+R_{\rm d})\frac{g_{ds}}{g_m} \quad \text{(A8)}$$

$$f_{\rm MAX} \approx \frac{f_T}{2\sqrt{(R_g+R_s+R_i)(g_{ds}+2\pi f_T C_{gd})}}. \quad \text{(A9)}$$

By differentiating (A8) with respect to the drain current per micron of gate width I'_{DS}, and imposing the condition for the derivative to be equal to zero, one can find an expression for the peak f_T current density and link it to the peak g_m current density:

$$\frac{\partial f_T}{\partial I'_{DS}} = \frac{1}{2\pi(C_{gs}+C_{gd})}\frac{\partial g_m}{\partial I'_{DS}} - \left(\frac{1}{g_m}+R_s+R_{\rm d}\right) \times f_T^2\frac{\partial C_{gd}}{\partial I'_{DS}} - \frac{f_T^2}{g_m}\frac{\partial C_{gs}}{\partial I'_{DS}} = 0 \quad \text{when } \frac{\partial g_m}{\partial I'_{DS}} > 0. \quad \text{(A10)}$$

Since, as illustrated in Fig. 26, C_{gd} and C_{gs} are monotonically increasing functions of I'_{DS}, f_T reaches its peak as a function of current density before g_m does.

The boundary between the two regions of the *I-V* or g_m–V characteristics corresponds roughly to the $V_{\rm eff}$ at which $f_{\rm MAX}$ reaches its peak value. Note that $f_{\rm MAX}$ depends on g_{ds} which increases with gate voltage and drain current. Once again, by taking the derivative of $f_{\rm MAX}$ with respect to the drain current density, one can prove that peak $f_{\rm MAX}$ occurs at a lower gate voltage and current density than peak f_T:

$$\frac{\partial f_{\rm MAX}}{\partial I'_{DS}} = \frac{1}{2\sqrt{(R_g+R_s+R_i)(g_{ds}+2\pi f_T C_{gd})}} \times \left[\frac{\partial f_T}{\partial I'_{DS}} - \frac{f_T}{\sqrt{g_{ds}+2\pi f_T C_{gd}}}\frac{\partial g_{ds}}{\partial I'_{DS}}\right] = 0 \quad \text{if } \frac{\partial f_T}{\partial I'_{DS}} > 0. \quad \text{(A11)}$$

The minimum noise factor $F_{\rm MIN}$ is given by [19]

$$F_{\rm MIN} \approx 1 + 2k_1\frac{f}{f_T}\sqrt{g_m(R_s+R_g)+k_2\left(1+\omega^2 R_i^2 C_{gs}^2\right)} \quad \text{(A12)}$$

where technology constants k_1 and k_2 depend on the parameters γ, β, and c that describe the drain and gate noise current sources of the FET [19], [31]. When differentiating (A12) with respect to the drain current density, and accounting for the fact that for

frequencies lower than f_T the frequency dependent term is negligible, one obtains

$$\frac{\partial F_{\mathrm{MIN}}}{\partial I'_{DS}} = \frac{k_1\sqrt{(R_g+R_s)g_m}f}{f_T}\left[\frac{1}{2g_m}\frac{\partial g_m}{\partial I'_{DS}} - \frac{1}{f_T}\frac{\partial f_T}{\partial I'_{DS}}\right] = 0 \quad \text{if } \frac{\partial f_T}{\partial I'_{DS}} > 0 \text{ and } \frac{\partial f_T}{\partial I'_{DS}} = \frac{f_T}{2g_m}\frac{\partial g_m}{\partial I'_{DS}} \quad \text{(A13)}$$

which indicates that J_{OPT} is lower than J_{pfT}, and that it is frequency independent, in good agreement with the experimental data presented in Figs. 6, 8, and 9.

Finally, it should be noted that the effective gate voltage and the characteristic current densities at which peak f_T, f_{MAX}, and the optimum F_{MIN} occur depend on the scaling of the equivalent oxide thickness (EOT) and on the constant-field scaling rules being applied. If a switch is made to constant-voltage scaling, the characteristic current densities are expected to increase in future nodes.

REFERENCES

[1] S. P. Voinigescu, T. O. Dickson, T. Chalvatzis, A. Hazneci, E. Laskin, R. Beerkens, and I. Khalid, "Algorithmic design methodologies and design porting of wireline transceiver IC building blocks between technology nodes," in *Proc. IEEE CICC*, San Diego, CA, Sep. 2005, pp. 111–118.

[2] T. O. Dickson, E. Laskin, I. Khalid, R. Beerkens, J. Xie, B. Karajica, and S. P. Voinigescu, "An 80-Gb/s $2^{31}-1$ pseudo-random binary sequence generator in SiGe BiCMOS technology," *IEEE J. Solid-State Circuits*, vol. 40, no. 12, pp. 2735–2745, Dec. 2005.

[3] S. E. Thompson, M. Armstrong, C. Auth, M. Alavi, M. Buehler, R. Chau, S. Cea, T. Ghani, G. Glass, T. Hoffman, C.-H. Jan, C. Kenyon, J. Klaus, K. Kuhn, Z. Ma, B. Mcintyre, K. Mistry, A. Murthy, B. Obradovic, R. Nagisetti, P. Nguyen, S. Sivakumar, R. Shaheed, L. Shifren, B. Tufts, S. Tyagi, M. Bohr, and Y. El-Mansy, "A 90-nm logic technology featuring strained-silicon," *IEEE Trans. Electron Devices*, vol. 51, no. 11, pp. 1790–1797, Nov. 2004.

[4] Y. Taur and T. H. Ning, *Fundamentals of Modern VLSI Devices*. Cambridge, U.K.: Cambridge Univ. Press, 1998, ch. 3, pp. 132–135.

[5] ITRS-2003. [Online]. Available: www.itrs.net/Common/2004Update/2004Update.htm

[6] B. Jangannathan, D. Greenberg, D. I. Sanderson, J.-S. Rieh, J. Pekarik, J. O. Plouchard, and G. Freeman, "Speed and power performance comparison of state-of-the-art CMOS and SiGe RF transistors," in *Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems, Dig. Tech. Papers*, Sep. 2004, pp. 115–118.

[7] K. Kuhn, R. Basco, D. Becher, M. Hattendorf, P. Packan, I. Post, P. Vandervoorn, and I. Young, "A comparison of state-of-the-art nMOS and SiGe HBT devices for analog/mixed-signal/RF circuit applications," in *Symp. VLSI Technology Dig. Tech. Papers*, Jun. 2004, pp. 224–225.

[8] N. Zamdmer, J. Kim, R. Trzcinski, J.-O. Plouchard, S. Narasimha, M. Khare, L. Wagner, and S. Chaloux, "A 243-GHz F_t and 208-GHz $F_{\max}$, 90-nm SOI CMOS SoC technology with low-power millimeter-wave digital and RF circuit capability," in *Symp. VLSI Technology Dig. Tech. Papers*, Jun. 2004, pp. 98–99.

[9] S. P. Voinigescu, D. S. McPherson, F. Pera, S. Szilagyi, M. Tazlauanu, and H. Tran, "A comparison of silicon and III-V technology performance and building block implementations for 10 and 40 Gb/s optical networking ICs," *Compound Semiconductor Integrated Circuits, Selected Topics in Electronics and Systems*, vol. 29, no. 1, pp. 27–58, 2003, World Scientific.

[10] S. P. Voinigescu, T. O. Dickson, R. Beerkens, I. Khalid, and P. Westergaard, "A comparison of Si CMOS, SiGe BiCMOS, and InP HBTs technologies for high-speed and millimeter-wave ICs," in *5th Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems, Dig. Tech. Papers*, Sep. 2004, pp. 111–114.

[11] S. P. Voinigescu, S. W. Tarasewicz, T. MacElwee, and J. Ilowski, "An assessment of the state-of-the-art 0.5 μm bulk CMOS technology for RF applications," in *IEDM Tech. Dig.*, 1995, pp. 721–724.

[12] W. Jeamsaksiri, A. Mercha, J. Ramos, D. Linten, S. Thijs, S. Jenei, C. Detcheverry, P. Wamback, R. Velghe, and S. Decoutere, "Integration of a 90 nm RF CMOS technology (200 GHz f_{MAX}—150 GHz f_T NMOS) demonstrated on a 5 GHz LNA," in *Symp. VLSI Technology Dig. Tech. Papers*, Jun. 2004, pp. 100–101.

[13] S. P. Voinigescu, R. Beerkens, T. O. Dickson, and T. Chalvatzis, "Design methodologies and applications of SiGe BiCMOS cascode opamps with up to 37-GHz unity gain bandwidth," in *IEEE CSIC Symp. Tech. Dig.*, Nov. 2005, pp. 283–286.

[14] G. Ng, B. Lai, P. Liu, and S. P. Voinigescu, "1 GHz OpAmp based bandpass filter," in *Si Monolithic Integrated Circuits in RF Systems, Tech. Dig.*, Jan. 2006, pp. 369–372.

[15] S. P. Voinigescu, T. O. Dickson, M. Gordon, C. Lee, T. Yao, A. Mangan, K. Tang, and K. H. K. Yau, "RF and millimeter-wave IC design in the nano-(Bi)CMOS era," in *Si-Based Semiconductor Components for Radio-Frequency Integrated Circuits (RFIC), 2006*, W. Z. Cai, Ed. New Delhi, India: Transworld Research Network, ISBN 81-7895-196-7.

[16] M. Vertregt and P. C. S. Scholtens, " Assessment of the merits of CMOS technology scaling for analog circuit design," in *Proc. ESSCIRC*, 2004, pp. 57–63.

[17] T. Yao, M. Gordon, K. Yau, M. T. Yang, and S. P. Voinigescu, "60-GHz PA and LNA in 90-nm RF-CMOS," presented at the IEEE RFIC Symp., San Francisco, CA, Jun. 2006, RMO2C-2.

[18] T. O. Dickson and S. P. Voinigescu, "SiGe BiCMOS topologies for low-voltage millimeter-wave frequency dividers and voltage-controlled oscillators," in *Si Monolithic Integrated Circuits in RF Systems, Tech. Dig.*, Jan. 2006, pp. 273–276.

[19] G. Vendelin, A. M. Pavio, and U. L. Rohde, *Microwave Circuit Design Using Linear and Nonlinear Techniques*, 2nd ed. New York: Wiley, 2005.

[20] S. P. Voinigescu, M. C. Maliepaard, J. L. Showell, G. E. Babcock, D. Marchesan, M. Schroter, P. Schvan, and D. L. Harame, "A scalable high-frequency noise model for bipolar transistors with application to optimal transistor sizing sizing for low-noise amplifier design," *IEEE J. Solid-State Circuits*, vol. 32, no. 9, pp. 1430–1439, Sep. 1997.

[21] H. Tran, F. Pera, D. S. McPherson, D. Viorel, and S. P. Voinigescu, "6-kOhm, 43-Gb/s differential transimpedance-limiting amplifier with auto-zero feedback and high dynamic range," *IEEE J. Solid-State Circuits*, vol. 39, no. 10, pp. 1680–1689, Oct. 2004.

[22] A. N. Karanicolas, "A 2.7-V 900-MHz CMOS LNA and mixer," *IEEE J. Solid-State Circuits*, vol. 31, no. 12, pp. 1939–1944, Dec. 1996.

[23] C. Kromer, G. Sialm, C. Berger, T. Morf, M. Schmatz, F. Ellinger, D. Erni, G.-L. Bona, and H. Jackel, "A 100 mW 4 × 10 Gb/s transceiver in 80 nm CMOS for high-density optical interconnects," in *IEEE ISSCC Dig. Tech. Papers*, Feb. 2005, pp. 334–335.

[24] T. O. Dickson, R. Beerkens, and S. P. Voinigescu, "A 2.5-V 45-Gb/s decision circuit using SiGe BiCMOS logic," *IEEE J. Solid-State Circuits*, vol. 40, no. 4, pp. 994–1003, Apr. 2005.

[25] H.-M. Rein and M. Moller, "Design considerations for very-high-speed Si bipolar IC's operating up to 50 Gb/s," *IEEE J. Solid-State Circuits*, vol. 31, no. 8, pp. 1076–1090, Aug. 1996.

[26] T. H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*. New York: Cambridge Univ. Press, 1998.

[27] T. O. Dickson, M.-A. LaCroix, S. Boret, D. Gloria, R. Beerkens, and S. P. Voinigescu, "30–100 GHz inductors and transformers for millimeter-wave (Bi)CMOS integrated circuits," *IEEE Trans. Microw. Theory Tech.*, vol. 53, no. 1, pp. 123–133, Jan. 2005.

[28] J. Kim, J.-K. Kim, B.-J. Lee, M.-S. Hwang, H.-R. Lee, S.-H. Lee, N. Kim, D.-K. Jeong, and W. Kim, "Circuit techniques for a 40-Gb/s transmitter in 0.13- μm CMOS," in *IEEE ISSCC Dig. Tech. Papers*, Feb. 2005, pp. 150–151.

[29] M. Laurens, B. Martinet, O. Kermarrec, Y. Campidelli, F. Deleglise, D. Dutartre, G. Troillard, D. Gloria, J. Bonnouvrier, R. Beerkens, V. Rousset, F. Leverd, A. Chantre, and A. Monroy, "A 150 GHz f_T/f_{MAX} 0.13-μm SiGe:C BiCMOS technology," in *Proc. BCTM*, 2003, pp. 199–202.

[30] P. Westergaard, T. O. Dickson, and S. P. Voinigescu, "A 1.5-V, 20/30-Gb/s CMOS backplane driver with digital pre-emphasis," in *Proc. IEEE CICC*, Orlando, FL, Oct. 2004, pp. 23–26.

[31] C. Enz, "An MOS transistor model for RF IC design valid in all regions of operation," *IEEE Trans. Microw. Theory Tech.*, vol. 50, no. 1, pp. 342–359, Jan. 2002.

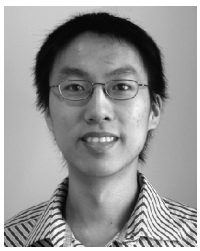
[32] L. Larson, "Silicon technology tradeoffs for radio-frequency/mixed-signal systems-on-a-chip," *IEEE Trans. Electron Devices*, vol. 50, no. 3, pp. 683–699, Mar. 2003.

[33] S. P. Voinigescu, M. Tazlauanu, P. C. Ho, and M. T. Yang, "Direct extraction methodology for geometry-scalable RF-CMOS models," in *Proc.IEEE ICMTS*, Mar. 2004, pp. 235–240.
[34] P. J. Tasker and B. Hughes, "Importance of source and drain resistance to the maximum f_T of millimeter-wave MODFETs," *IEEE Electron Device Lett.*, vol. 10, no. 7, pp. 291–293, Jul. 1989.
[35] H. M. J. Boots, G. Doornbos, and A. Heringa, "Scaling of characteristic frequencies in RF CMOS," *IEEE Trans. Electron Devices*, vol. 51, no. 12, pp. 2102–2108, Dec. 2004.

Timothy O. Dickson (S'01) received the B.S. (with highest honors) and M.Eng. degrees in electrical engineering from the University of Florida, Gainesville, in 1999 and 2002, respectively. He is currently working toward the Ph.D. degree at the University of Toronto, Toronto, ON, Canada.

From 2000 to 2002, he was a Graduate Research Assistant with the University of Florida, where he investigated isolation issues in RF and mixed-signal silicon-based technologies. He has held internships in the area of analog and RF circuit design at Dallas Semiconductor and Global Communication Devices. His research interests lie in the area of high-frequency integrated circuits for wireless and wireline communications.

Mr. Dickson was named an undergraduate University Scholar by the University of Florida in 1999. In 2004, he received the Best Student Paper Award at the VLSI Circuits Symposium in Honolulu. He holds a University of Toronto doctoral fellowship and is an Edward S. Rogers Sr. Scholar.



Kenneth H. K. Yau (S'03) received the B.A.Sc. degree, with distinction, in engineering physics from the University of British Columbia, Vancouver, BC, Canada, in 2003 and the M.A.Sc degree in electrical engineering from the University of Toronto, Toronto, ON, Canada, in 2006. He is currently pursuing the Ph.D. in electrical engineering at the University of Toronto.

He has held an internship with Broadcom Corporation, Vancouver, where he was involved with the development of embedded software for IP telephony. He was also involved with the development of a nano-scale vibration control system for the quadrupole magnets of the Next Linear Collider. His current research interests are in the area of device modeling for advanced MOSFETs and SiGe HBTs, and nano-scale device physics.

Mr. Yau has received numerous scholarships, including an undergraduate research award and a post-graduate scholarship from the Natural Sciences and Engineering Council of Canada (NSERC). He also received an achievement award from the Association of Professional Engineers and Geoscientists of British Columbia (APEGBC).

Theodoros Chalvatzis received the Diploma in electrical and computer engineering from the University of Patras, Greece, in 2001 and the M.A.Sc. degree in electrical engineering from Carleton University, Ottawa, ON, Canada, in 2003. He is currently working toward the Ph.D. degree at the University of Toronto, Toronto, ON, Canada.

From 2002 to 2003, he held an internship with Nortel Networks, where he worked on advanced wireless receivers. His research interests include high-speed analog-to-digital converters and digital circuits for wireless and wireline applications.

Alain M. Mangan received the B.A.Sc. degree in computer engineering from the University of Ottawa, Ottawa, ON, Canada, in 2003, and the M.A.Sc. degree in electrical and computer engineering from the University of Toronto, Toronto, ON, Canada, in 2005.

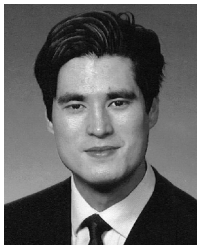
In 2001 and 2002, he performed device characterization work on 130-nm RF-CMOS, and 180-nm BiCMOS processes at Quake Technologies, Ottawa, for two work terms. His research interests include mm-wave device characterization for higher-frequency IC design.

Ekaterina Laskin (S'04) received the B.A.Sc. (Hons.) degree in computer engineering and the M.A.Sc degree in electrical engineering from the University of Toronto, Toronto, ON, Canada, in 2004 and 2006, respectively. She is currently working toward the Ph.D. degree at the Department of Electrical and Computer Engineering, University of Toronto.

Ms. Laskin was a University of Toronto Scholar from 2000 to 2004. She received the National Science and Engineering Research Counsel of Canada (NSERC) undergraduate student research award in industry and university in 2002 and 2003. She currently holds the NSERC Postgraduate Scholarship. Her research interests include the design of millimeter-wave and high-speed integrated circuits.

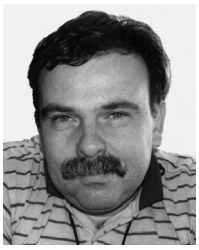
Rudy Beerkens (M'91) received the B.Sc. degree in electrical engineering from the University of Waterloo, Waterloo, ON, Canada, in 1986.

From 1986 to 2000, he was with Nortel Networks, Ottawa, ON, Canada, during which time he was involved in the areas of semiconductor manufacturing, BiCMOS process development, yield enhancement, device characterization and reliability. In 2000, he joined STMicroelectronics, Ottawa, ON, Canada, where he currently heads an integrated circuits design team. His scientific interests include high-speed broadband and millimeter-wave integrated circuits.



Paul Westergaard received the B.A.Sc. degree in electrical engineering from the University of Toronto, Toronto, ON, Canada, in 2000 and has completed requirements for the M.Eng. degree in electrical engineering from the University of Toronto in 2006.

In 1998, he spent 16 months at ATI Technologies Inc., Markham, Canada, where he was involved in the design and layout of high-speed interfaces and advanced logic cells. In 2001, he joined ViXS Systems Inc., Toronto, where he was responsible for designing digital signal processing algorithms and integrated circuits for consumer quality video filtering and compression. He is now with Torys LLP, Toronto, advising in the area of intellectual property and technology law. His graduate research interests included 30-Gbps CMOS backplane drivers, digital pre-emphasis equalization, and wave-shape control.



Mihai Tazlauanu (M'01) received the Ph.D. degree in electrical engineering from the Polytechnic University of Bucharest, Bucharest, Romania, in 1995 and the Ph.D. degree in physics from the University of Montreal, Montreal, QC, Canada, in 1997.

For ten years, he was with the Quality and Reliability Department, Research and Development Institute for Microelectronics, Bucharest. In 1996, he joined Semiconductor Insights, Kanata, Canada, where he was involved in process analysis and the development of leading-edge semiconductor technologies. Since 2001, he has been with Quake Technologies, Ottawa, ON, Canada, where he has been responsible for device evaluation and modeling as well as reliability qualification. His current research interests include RF model extraction for active and passive devices produced in advanced Si, Si-Ge, and III-V processes.

Ming-Ta Yang (M'96) was born in Tainan, Taiwan, R.O.C., in 1966. He received the B.S. and M.S. degrees in physics and applied physics from Chung-Yuan University, Chung-Li, Taiwan, in 1989 and 1992. During the 1989 academic year, he joined the Department of Physics, Chung-Yuan University, as a faculty member. In 1995, he received the Ph.D. degree in electrical engineering from the National Central University, Chung-Li, Taiwan. His Ph.D. dissertation was focused on the characterization and model of AlGaAs/InGaAs doped-channel heterostructure FETs and its application in monolithic microwave integrated circuits (MMICs).

In 1995, he joined Mosel Vitelic Inc., Hsin-Chu, Taiwan, where he worked on VLSI advanced technology development. He then joined the Taiwan Semiconductor Manufacture Company, Ltd. (TSMC) in 1998, where he set up a high-frequency characterization laboratory and organized a team to response for the development of characterization and Spice model both in mixed-signal and RF applications. His current research interests include the development of advanced HF characterization and model of devices both in the RF CMOS and high-speed SiGe HBT BiCMOS technologies. He has served on the advisory Workshop Committees of Compact Modeling for RF/Microwave Applications (CMRF) in 2004 and 2005.

Sorin P. Voinigescu (M'90–SM'02) received the M.Sc. degree in electronics from the Polytechnic Institute of Bucharest, Bucharest, Romania, in 1984, and the Ph.D. degree in electrical and computer engineering from the University of Toronto, Toronto, ON, Canada, in 1994.

From 1984 to 1991, he worked in R&D and academia in Bucharest, where he designed and lectured on microwave semiconductor devices and integrated circuits. From 1994 to 2002, he was with Nortel Networks and with Quake Technologies in Ottawa, ON, Canada, where he was responsible for projects in high-frequency characterization and statistical scalable compact model development for Si, SiGe, and III-V devices. He also led the design and product development of wireless and optical fiber building blocks and transceivers in these technologies. In 2000, he co-founded and was the CTO of Quake Technologies, the world's leading provider of 10 Gb Ethernet transceiver ICs. In September 2002, he joined the Department of Electrical and Computer Engineering, University of Toronto, as an Associate Professor. He has authored or co-authored over 60 refereed and invited technical papers spanning the simulation, modeling, design, and fabrication of high-frequency semiconductor devices and circuits. His research and teaching interests focus on nanoscale semiconductor devices and their application in integrated circuits at frequencies beyond 100 GHz.

Dr. Voinigescu received NORTEL's President Award for Innovation in 1996. He is a co-recipient of the Best Paper Award at the 2001 IEEE Custom Integrated Circuits Conference and at the 2005 Compound Semiconductor IC Symposium. His students have won the Best Student Paper Award at the 2004 IEEE VLSI Circuits Symposium and the 2006 SiRF Meeting.