# 1 Introduction

For more than 40 years, the evolution and growth of very-large-scale integration (VLSI) silicon-based integrated circuits (ICs) have followed from the continual shrinking, or scaling, of the CMOS technology, i.e., of the constituent transistors, or MOSFETs. Remarkably, this scaling of the CMOS technology (Taur and Ning, 2009), which has tracked "Moore's Law" (Moore (1965): the IC device packing density will double every 18 months) quite well since the 1960s, has not involved any major change in the basic, planar MOSFET structure on bulk silicon or, more recently, on partially depleted (PD) SOI, as shown in Fig. 1.1. This structure has simply been geometrically scaled to gate lengths of $L_g \cong 30$ nm, albeit with added complexity to the CMOS fabrication process as well as to the device doping profiles, etc., to control detrimental short-channel effects (SCEs), as described very well in Taur and Ning (2009). Most recently, however, such scaling has been slowed, and even stopped in terms of gate length, mainly because the complex doping profiles required cannot be achieved reliably, or with acceptable yield. The unavoidable randomness of the dopants in the silicon lattice causes, for nanoscale MOSFETs, prohibitive variations in device properties, e.g., the threshold voltage $V_t$. Indeed, this scaling slow-down is reflected well in Fig. 1.2 by comparison of an exemplary roadmap (SIA, 1994–2011, 2003 ITRS update) projection of high-performance (e.g., microprocessor or MPU) $L_g$ from the Semiconductor Industry Association (SIA) with the actual $L_g$ scaling that has been achieved, indicated by Intel's MPU scaling results (Natarajan *et al.*, 2008; Auth *et al.*, 2012) superimposed on the projection. Clearly, the technology lags the projection, which has been the general case. The most recent projection (SIA, 1994–2011, 2011 ITRS update) points to $L_g = 12.8$ nm in 2018 and $L_g = 5.9$ nm in 2026, whereas the latest reported achievement is minimum $L_g = 26$ nm in Intel's 22 nm-node FinFET technology (Auth *et al.*, 2012).

The IC industry is earnestly researching new device structures to enable continued CMOS scaling (SIA, 1994–2011, 2011 ITRS update), of which there are two main candidates: the planar fully depleted (FD) SOI MOSFET, shown in Fig. 1.3(a), likely with a thin underlying buried oxide (BOX) and heavily doped ground plane (GP) in the substrate (Liu *et al.*, 2010) as shown in Fig. 1.3(b); and the 3-D (also FD) FinFET (Hisamoto *et al.*, 1991), with simplified processing and a "quasi-planar" structure (Lindert *et al.*, 2001) as illustrated in Fig. 1.4. The former FD/SOI MOSFET operates with one gate (although the substrate can be considered as a second gate), but the FinFET uses two, or even three, gates. Both of these novel devices rely on an ultra-thin

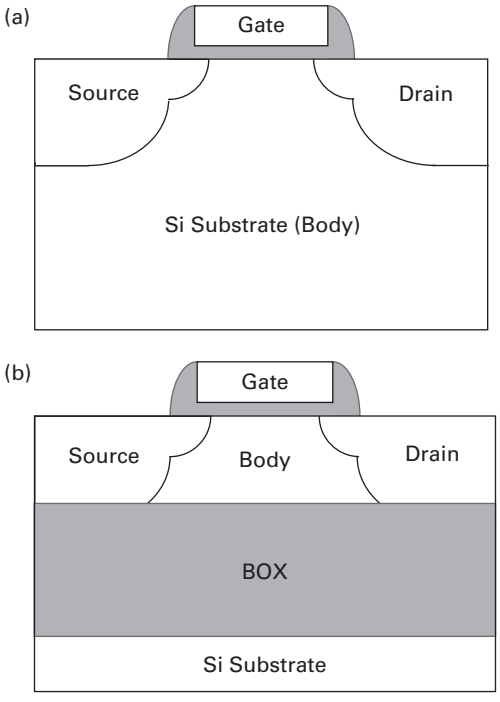


**Figure 1.1.** Cross-sectional views of classical MOSFET structures: (a) bulk-Si MOSFET; (b) PD/SOI MOSFET with thick BOX (buried oxide).

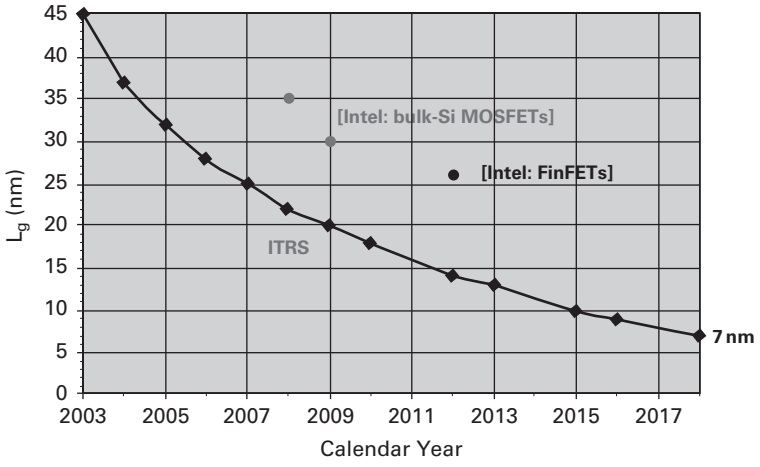


**Figure 1.2.** Exemplary SIA ITRS-projected scaling of high-performance CMOS gate length (SIA, 1994–2011, 2003 ITRS update). Generally, as indicated by the superposition of Intel's actual MPU scaling results (Natarajan *et al*., 2008; Auth *et al*., 2012), the projections lead the technology, and so they are modified periodically. For example, the 2011 roadmap (SIA, 1994–2011, ITRS 2011 update) projects $L_g$ = 12.8 nm, not 7 nm, for 2018, leading to 5.9 nm in 2026; $L_g$ = 7 nm is projected for 2024–25.

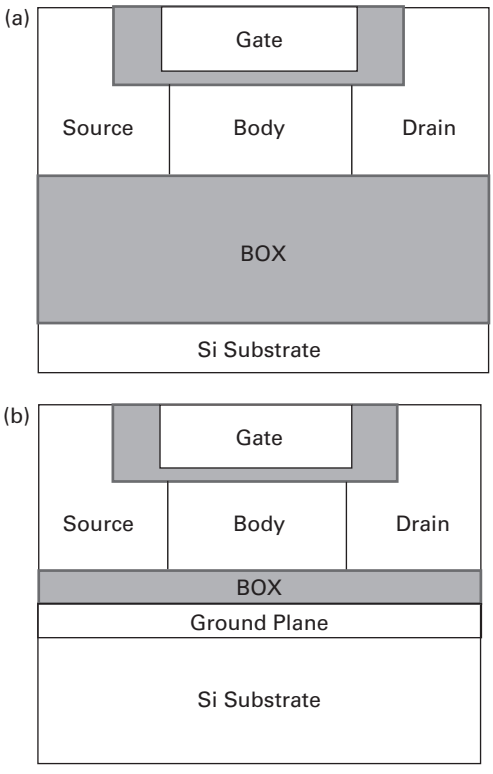


**Figure 1.3.** Cross-sectional views of basic, planar FD/SOI MOSFETs: (a) conventional device with thick BOX; (b) nanoscale device with thin BOX and GP, which can serve as a back gate.

body (UTB) to help control the SCEs, in contrast to the bulk-Si MOSFET which uses complex doping profiles to do so. This book describes the unique fundamentals of these two novel CMOS devices, with a bit more emphasis on the FinFET, which, in our opinion, has more potential. (In fact, Intel Corporation has already announced (Auth *et al.*, 2012) that their 22 nm-node CMOS technology introduced in 2012 (see Fig. 1.2), and future ones, will use FinFETs, or so-called trigate transistors (Kuhn, 2011).) Our book will refer to previous publications for the underlying basic MOSFET theory, especially Taur and Ning (2009), to which we consider our book a contemporary supplement.

## 1.1 Ultimate nanoscale CMOS

A scaling limit for classical CMOS, with bulk-Si or PD/SOI MOSFETs (Taur and Ning, 2009), is now palpable, and virtually defined (by Intel) to be $L_g \cong 30$ nm (see Fig. 1.2). No additional performance "boosters," such as strained-Si channels and metal/high-$k$ gate stacks, will enable a reliable classical technology at and beyond the 22 nm node with

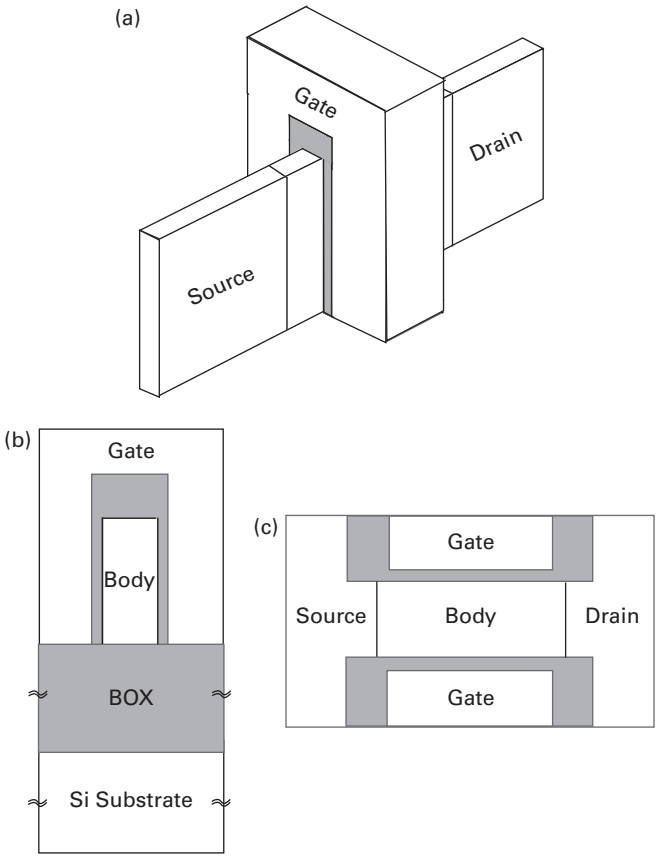


**Figure 1.4.** Basic quasi-planar FinFET structure: (a) 3-D view (the substrate is not shown); (b) 2-D cross-sectional S-D view (substrate is SOI); and (c) top cross-sectional view, including G-S/D spacers (not shown in (a)) that enable a G-S/D underlap (as indicated in (a)).

$L_g < 30$ nm. A new device structure, or structures, will be needed. This book is focused on the new devices that appear to have the most potential for nanoscale CMOS at 22 nm and beyond, i.e., the single-gate (SG) planar FD/SOI MOSFET probably with thin BOX, for which the substrate can be considered a second gate, and the quasi-planar (FD) FinFET, which can be double-gate (DG) or triple-gate (TG). Both of these new devices will require UTBs, which more than likely will be left undoped, thereby electrically coupling the gates. Our book describes the unique features and fundamentals of these UTB devices, for which the classical MOSFET physics (Taur and Ning, 2009) is inadequate, although still applicable in part. The discourse thereby gives useful insights on the optimal designs of UTB devices, as well as their ultimate utilities in nanoscale CMOS applications. Further, the book introduces the University of Florida's compact model for generic DG MOSFETs, UFDG (see the Appendix), which physically accounts for the unique UTB physics, as a useful aid in understanding the UTB-device fundamentals as well as in device and circuit design. Throughout the book, UFDG/Spice3 simulation

results are used to clarify the discussions of the fundamentals and to give added physical insights on their effects.

### 1.1.1 Planar FD/SOI MOSFET

The basic, planar FD/SOI MOSFET is shown in Fig. 1.3. This device has emerged from basic (PD and FD) SOI CMOS technology, which began its evolution in the 1980s. The process flow of the FD/SOI MOSFET is quite similar to that of the conventional (bulk-Si and PD/SOI) MOSFETs (Taur and Ning, 2009), except for the incorporation of the ~10 nm UTB and raised source/drain (S/D) regions (needed to keep series resistance as low as possible). The thin FD UTB enables electrical coupling between the (front) gate and the substrate (back gate). The thin-BOX option enhances this coupling, making $V_t$ significantly dependent on the substrate doping (i.e., that of the underlying GP) and its bias, as well as on the UTB ($t_{Si}$) and BOX ($t_{BOX}$) thicknesses. The SCEs and device scalability are governed by these thicknesses. Our discussions of the fundamentals of the FD/SOI MOSFET will describe dependences on $t_{BOX}$ (e.g., thin- versus thick-BOX features) as well as on $t_{Si}$.

Recent works on processing planar FD/SOI UTB MOSFETs, with thick and thin BOX, and on improving their electrical characteristics, reflect their potential for nanoscale CMOS. For example, Khakifirooz *et al*. (2010) demonstrated small, acceptable $V_t$ variation (see Fig. 1.5(a)), with minimal dependence on $t_{Si}$ (see Fig. 1.5(b)), enabled by undoped UTBs in thick-BOX FD/SOI MOSFETs with $t_{Si} < 10$ nm and $L_g < 35$ nm. Further, Khakifirooz *et al*. (2012) described thick-BOX devices fabricated with $t_{Si}$ as thin as 3.5 nm and the effective channel length ($L_{eff}$) varying down to 18 nm, and showed, as illustrated in Fig. 1.6, measured $I_{off}$ versus $I_{on}$ and SCEs for different $t_{Si}$ and substrate bias ($V_{GbS}$). These data show good control of SCEs and off-state current ($I_{off}$) with reasonable on-state current ($I_{on}$), and demonstrate the benefits of proper $V_{GbS}$ for SCE and $V_t$ control. The utility of substrate bias (and GP doping) for SCE and $V_t$ control, however, becomes more feasible (in terms of practical $V_{GbS}$) with thinner BOX, as demonstrated in Fig. 1.7 (Fenouillet-Beranger *et al*., 2009) via measured SCEs in nanoscale FD/SOI devices with $t_{BOX} = 20$ nm and grounded substrate with varying GP doping. Clearly, these samplings of nanoscale FD/SOI UTB MOSFETs imply feasibility of the planar device in future CMOS applications.

### 1.1.2 FinFET

The basic, quasi-planar (FD) FinFET is shown in Fig. 1.4. The intrinsic FinFET utilizes the third, vertical dimension, and is thus a bit more revolutionary than the planar FD/UTB MOSFET. The FinFET is, in essence, a planar MOSFET that is folded vertically, with the gate stack wrapped over the fin UTB and the device width being defined by the fin height. Aside from incorporating the vertical fin (the UTB width is now the minimum feature size, as opposed to $L_g$ in planar processing (Taur and Ning, 2009)) in the standard topology, the process flow of this 3-D device is not much

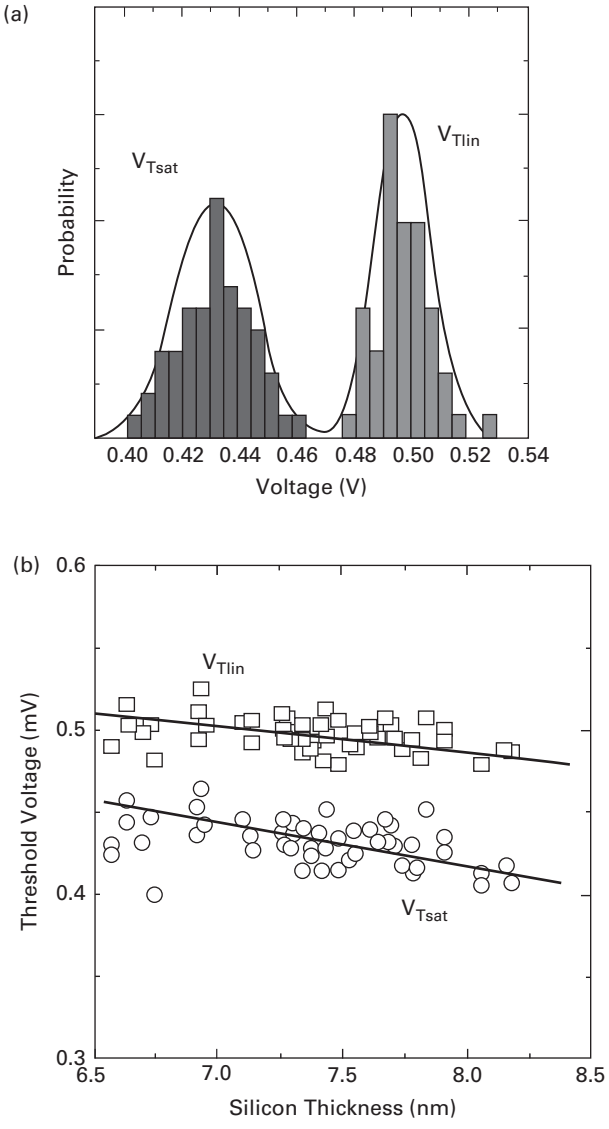


**Figure 1.5.** Measured variation in saturation threshold voltage ($V_{Tsat}$) and linear threshold voltage ($V_{Tlin}$) of planar FD/SOI nMOSFETs with $L_g < 35$ nm and $t_{Si} < 10$ nm: (a) within wafer variation, showing an acceptable 60–70 mV range; (b) variation due to change in $t_{Si}$, showing about a 25 mV/nm sensitivity. (After Khakifirooz *et al.*, 2010.)

different from that of the conventional MOSFETs. This novel device was conceived in 1991, but its development for CMOS did not begin in earnest until after 2000. The familiar FinFET mode is DG, with two active sidewall gates. A third gate can be activated on the top of the fin. The thin-fin UTB electrically couples the sidewall gates, like in the FD/SOI MOSFET, and its thickness ($t_{Si}$) governs the SCEs and the device

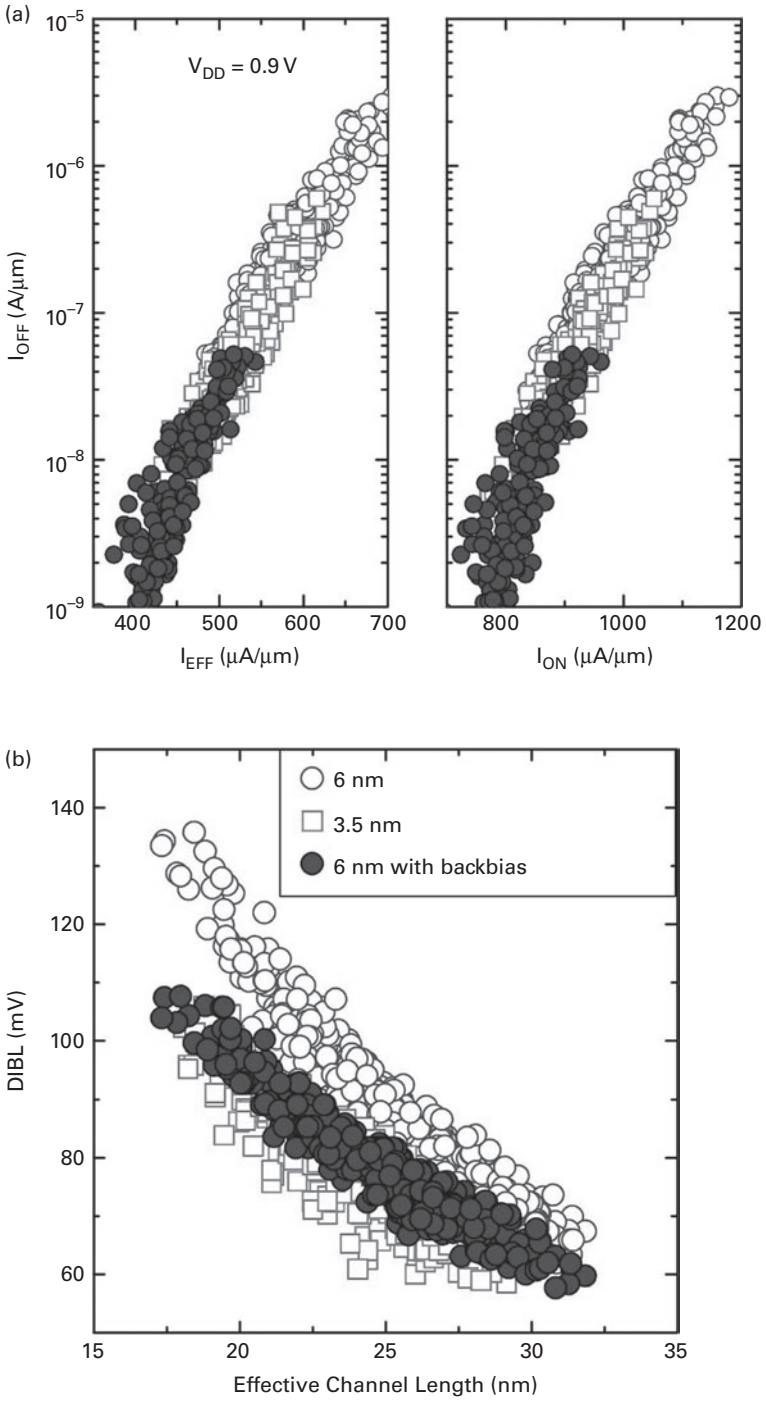


**Figure 1.6.** (a) Measured off-state current versus on-state and effective currents ($I_{EFF}$ is a CMOS inverter-based effective drive current (Na *et al*., 2002)) for planar FD/SOI MOSFETs with $t_{Si}$ = 6 nm, without (open circles) and with (filled circles) back-gate (substrate) bias ($V_{GbS}$, towards accumulation), and $t_{Si}$ = 3.5 nm without $V_{GbS}$ (squares), with the effective channel length ($L_{eff}$) varying down to 18 nm. (b) Measured DIBL versus $L_{eff}$ of the devices in (a), showing very good SCE control effected by either thinning $t_{Si}$ or by applying $V_{GbS}$. These results were obtained with devices having thick BOX (145 nm), and hence with higher $V_{GbS}$ (= −20 V) than would be needed for thin BOX. (After Khakifirooz *et al*., 2012.)

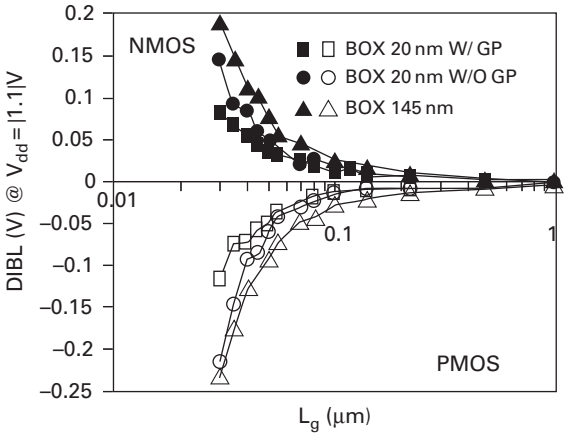


**Figure 1.7.** Measured DIBL versus gate length of planar FD/SOI nMOSFETs (filled symbols) and pMOSFETs (open symbols) with 145 nm-thick BOX (triangles), 20 nm thin BOX without ground plane (circles), and 20 nm thin BOX with ground plane (squares), showing improved SCE control with thin BOX and doping in the (grounded, i.e., $V_{GbS} = 0$ V) substrate; $t_{Si}$ is 8 nm and the high-*k*-based equivalent gate-oxide thickness is 1.7 nm. Note that the thin BOX enables SCE control without large $V_{GbS}$. (After Fernouillet-Beranger *et al*., 2009.)

scalability. The gate-source/drain (G-S/D) "underlap" indicated in the figure can be beneficial to the scalability of the FinFET, as well as to that of the FD/SOI MOSFET in Fig. 1.3.

Recent works on FinFET processing and design exemplify outstanding performance potential for nanoscale CMOS. Figure 1.8 shows measured current–voltage characteristics, reflecting excellent SCEs (DIBL ≅ 50 mV/V and subthreshold slope ≅ 70 mV/dec), and $I_{off}$ versus $I_{on}$ of Intel's 22 nm (bulk-Si) FinFET technology (Auth *et al*., 2012), which is currently in volume production with the first FinFET-based (multi-core) processor (Damaraju *et al*., 2012). The combination of the excellent SCE control and improved $I_{on}$ in Fig. 1.8 yields substantive speed–power improvement, e.g., 37% faster switching speed or ≅ 50% lower active power (via supply voltage scaling), as illustrated in Fig. 1.9 (Damaraju *et al*., 2012). While Intel's adoption of FinFETs for 22 nm-node technology utilizes bulk-Si substrates, scaling FinFETs to $L_g \cong 10$ nm will probably require (as we argue in Chapter 4) SOI (or SOI-equivalent) substrates. The scalability of SOI DG FinFETs, via $t_{Si}$ scaling, is indicated in Fig. 1.10 (Chang *et al*., 2003) by measured DIBL (drain-induced barrier lowering) of devices fabricated with $L_g$ varying from 20 nm to 150 nm (probably with $L_{eff} > L_g$) and $t_{Si}$ varying from 10 nm to 42 nm. For acceptable SCEs, e.g., DIBL ≅ 100 mV/V, with a lower limit of $t_{Si}$ = 4–5 nm (Trivedi and Fossum, 2005b), Fig. 1.10 implies SOI FinFET scalability down to $L_g < 10$ nm. Thus, perhaps even more so than for the planar FD/SOI MOSFET discussed in the previous section, these samplings of UTB FinFETs suggest viability for the device in nanoscale CMOS applications pushed to the end of the SIA roadmap (SIA, 1994–2011, 2011 ITRS update).

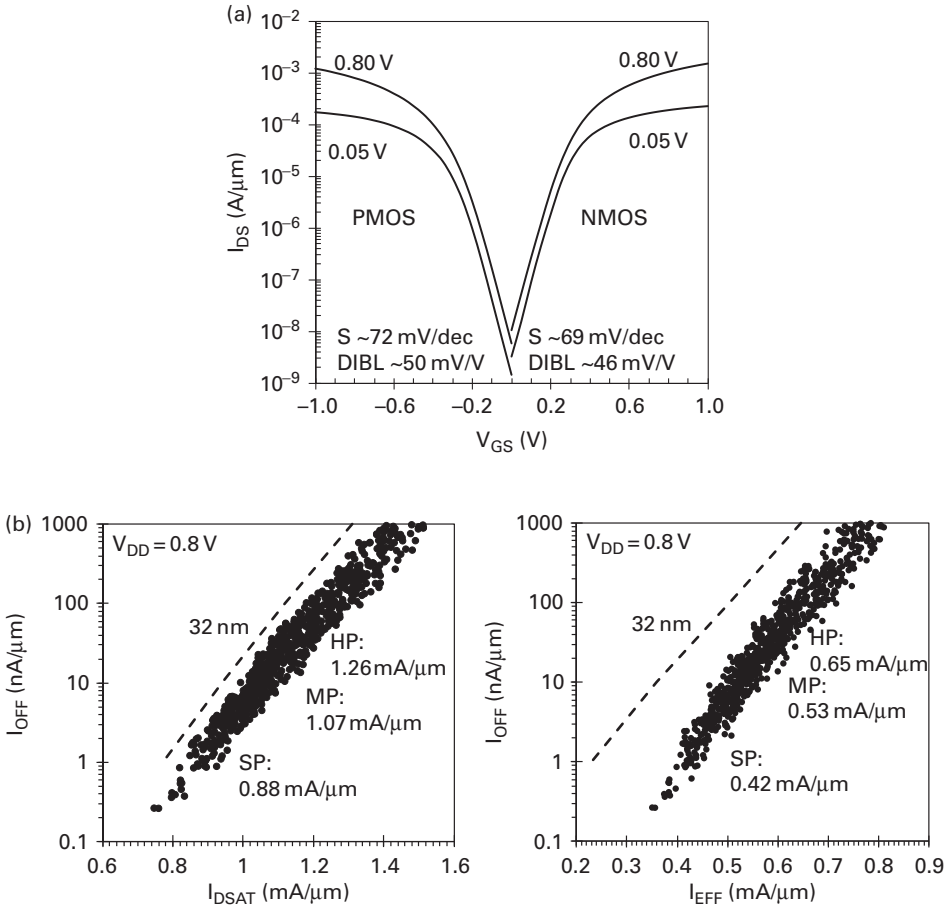


**Figure 1.8.** (a) Measured current–voltage characteristics of 22 nm-node CMOS FinFETs on bulk-Si substrate, with low supply voltage ($V_{DD}$) of 0.8V. Excellent SCE and $V_t$ control is achieved for $L_g$= 26 nm FinFETs with $t_{Si} \cong$ 8 nm, high-*k*-based effective gate-oxide thickness of 0.9 nm, and dual-work function metal gates. (b) Measured off-state current versus on-state ($I_{DSAT}$) and effective currents of nFinFETs in (a), exemplifying noticeable improvement in high-performance (HP), nominal-power (MP), and low-power (SP) devices compared to preceding 32 nm-node bulk-Si technology. (After Auth *et al*., 2012.)

## 1.2 Brief overview of the book

We wrote the book to be a supplement to existing books on classical CMOS devices, especially Taur and Ning (2009). Thus, we focus on the *unique fundamentals* of the two noted UTB devices, with references to Taur and Ning (2009) for the related classical MOSFET fundamentals. Brief overviews of the primary parts of the book are given below. Each chapter contains a set of exercises dealing with corresponding UTB-device fundamentals.

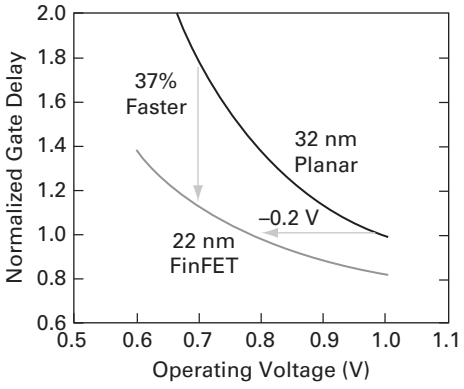


**Figure 1.9.** Measured normalized gate delay versus operating voltage (or active power) of the 22 nm-node FinFET CMOS of Fig. 1.8 compared to the preceding 32 nm-node planar bulk-Si technology, showing up to 37% faster switching speed at low $V_{DD}$ or 0.2V lower $V_{DD}$, and thereby ≅50% lower active power consumption, at similar switching speed. (After Damaraju *et al*., 2012.)

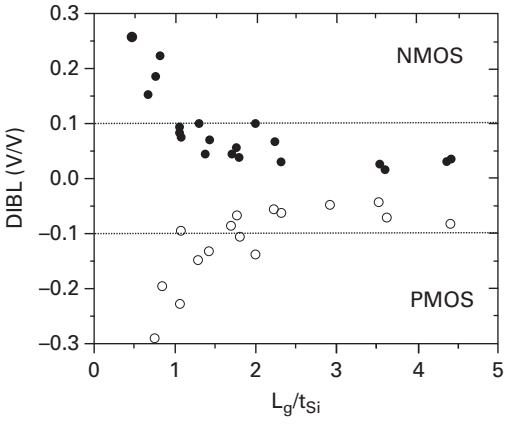


**Figure 1.10.** Measured DIBL versus the $L_g/t_{Si}$ ratio for DG FinFETs on SOI substrates. FinFETs with $L_g$ (< $L_{eff}$ probably) varying from 20 nm to 150 nm and $t_{Si}$ varying from 10 nm to 42 nm were measured. For acceptable DIBL ≅100 mV/V with a lower limit of $t_{Si}$ = 4–5 nm, scalability to $L_g$ < 10 nm is implied. (After Chang *et al*., 2003.)

## Chapter 2: Unique features of UTB MOSFETs

Chapter 2 describes the basic FD-UTB device theory. After a review of the classical SOI theory based on the depletion approximation and inversion/accumulation-charge sheets, a nonclassical, general analysis of the 1-D electrostatics based on Poisson's equation and the boundary conditions at the front and back surfaces of the UTB is described. The main result is a generic formalism for $V_t$ of UTB devices, accounting for the electrical coupling