Journal of
Low Power Electronics
Vol. 13, 1–18, 2017

Back to the Future: Digital Circuit Design in the FinFET Era

Xinfei Guo*, Vaibhav Verma, Patricia Gonzalez-Guerrero, Sergiu Mosanu, and Mircea R. Stan

Department of Electrical and Computer Engineering, University of Virginia, Charlottesville, VA, 22904, USA

(Received: 3 June 2017; Accepted: 13 June 2017)

It has been almost a decade since FinFET devices were introduced to full production; they allowed scaling below 20 nm, thus helping to extend Moore's law by a precious decade with another decade likely in the future when scaling to 5 nm and below. Due to superior electrical parameters and unique structure, these 3-D transistors offer significant performance improvements and power reduction compared to planar CMOS devices. As we are entering into the sub-10 nm era, FinFETs have become dominant in most of the high-end products; as the transition from planar to FinFET technologies is still ongoing, it is important for digital circuit designers to understand the challenges and opportunities brought in by the new technology characteristics. In this paper, we study these aspects from the device to the circuit level, and we make detailed comparisons across multiple technology nodes ranging from conventional bulk to advanced planar technology nodes such as Fully Depleted Silicon-on-Insulator (FDSOI), to FinFETs. In the simulations we used both state-of-art industry-standard models for current nodes, and also predictive models for future nodes. Our study shows that besides the performance and power benefits, FinFET devices show significant reduction of short-channel effects and extremely low leakage, and many of the electrical characteristics are close to ideal as in old long-channel technology nodes; FinFETs seem to have put scaling back on track! However, the combination of the new device structures, double/multi-patterning, many more complex rules, and unique thermal/reliability behaviors are creating new technical challenges. Moving forward, FinFETs still offer a bright future and are an indispensable technology for a wide range of applications from high-end performance-critical computing to energy-constraint mobile applications and smart Internet-of-Things (IoT) devices.

Keywords: FinFET, FDSOI, Planar, VLSI, Scaling, Sizing, Digital Design.

1. INTRODUCTION

The continuous scaling of planar CMOS devices has delivered increasing performance and transistor densities. However, it also reached a point where increased leakage current, fluctuation of device characteristics and short channel effects became serious obstacles to further scaling. This was mainly because deeply-scaled planar devices became increasingly influenced by the drain potential as the gate lost the ability to fully control the channel; this led to transistors that were never fully off and leaked continuously. To solve this problem, gate oxides were aggressively thinned and high-k dielectric gate materials were adopted to increase the gate-channel capacitance, but the gate-related issues, such as gate leakage and gate-induced drain leakage (GIDL) increased.[1,2] FinFET devices became attractive for sub-30 nm nodes[3,4] because of their unique channel structure with good gate control that enables a much improved short channel control, thus requiring little or no doping in the channel. The threshold voltage V_t can be scaled down in FinFETs for both improved device performance and a much lower operation voltage. Lower channel doping also reduces dopant ion scattering, thus leading to better drive currents and decreases random dopant fluctuations (RDF).[5–7] FinFETs back-end-of-line (BEOL) fabrication is fully compatible with planar devices in both bulk and SOI varieties, which reduces the need for new, FinFET-specific developments in that area. However, the introduction of FinFETs has brought a few changes and challenges in digital circuit design due to their unique gate structure and electrical properties. This has also impacted the circuit design decisions and some of the available design tradeoffs. For example, FinFET devices have a significant amount of parasitics

*Author to whom correspondence should be addressed.
Email: xg2dt@virginia.edu

that need to be modeled precisely and be carefully considered in the layout of all circuits, especially in SRAM and analog circuits. From a circuit design aspect, in addition to the extra effort needed to address the impact of parasitics at the layout level, new circuit techniques are needed in the area of body-biasing and memory read/write assist in SRAMs to replace techniques that worked well in planar but are inefficient for FinFET. The double/multipatterning also requires tool vendors and designers to work together to make sure the layout coloring is correct (colors refer to different exposures of the same layer while performing multipatterning). New constraints have been added to FinFET design, such as width quantization and self-heating effects, for which designers need to make early decisions in the design cycle. In this paper, we analyze these aspects at both the device and circuit levels. To study these challenges, we simulate across multiple technology nodes which cover a wide range of gate lengths and also substrates including both SOI and Bulk. For FinFET, we simulate with both $1 \times$ nm[a] industry-standard node and a 7 nm predictive node. This paper aims to provide a detailed analysis and global view of how FinFETs differ from previous technology nodes and what are the implications on circuit design. We restrict our focus to digital circuits, but several of the findings can be applied to analog design as well.

The paper is organized as follows. Section 2 discusses the FinFET basics and how FinFETs are different from planar technologies at the device level. We address the changes and challenges FinFETs have introduced for circuit design in Section 3. In Section 4, we summarize all the challenges from the designers' perspective. Section 5 concludes the paper.

2. FINFET DEVICE

2.1. Scaling and Sizing

2.1.1. FinFET Structure

Compared to conventional planar devices (bulk or SOI), FinFET devices have unique 3-D gate structures that enable some special properties for FinFET circuit design which will be detailed in the following sections. Illustrated in Figure 1 is a planar device and a FinFET device (the substrate is not included in the figure). While the channel of the planar device is horizontal, the FinFET channel is a thin vertical fin with the gate fully "wrapped" around the channel formed between the source and the drain. The current flows parallel to the die plane whereas the conducting channel is formed around the fin edges.

[a]In advanced technology nodes the "numbering" scheme is somewhat arbitrary, while in older technologies the node "number" used to denote the smallest feature size, usually the transistor gate length, in modern technologies the node number does not refer to any one feature in the process, and foundries use slightly different conventions; we use $1\times$ to denote the 14 nm–16 nm FinFET nodes offered by several foundries.

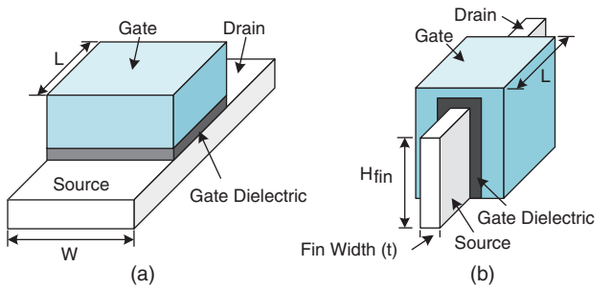

Fig. 1. Illustration of structural differences (no substrate): (a) planar device; (b) FinFET device.

With this structure, the gate is able to fully deplete the channel thus having much better electrostatic control over the channel.

FinFETs can be classified by gate structure or type of substrate. Different gate structures lead to two versions of FinFET—Shorted-gate (SG) FinFETs and Independent-gate (IG) FinFETs. In SG devices, the left and right sides are connected together in a wrap-around structure as in Figure 1; this can serve as a direct replacement for the planar devices which also have one gate, a source and a drain (three terminal-devices). In IG FinFETs, the top part of the wraparound gate structure is etched out and this results two separate left and right sides that can act as independent gates and can be controlled separately.[8,9] Although IG FinFETs offer more design options, the fabrication costs are also higher in general. Depending on the substrate, the FinFETs can be either SOI or bulk FinFETs as illustrated in Figure 2. SOI FinFETs are built on SOI wafers and have a lower parasitic capacitance and slightly less leakage. Bulk FinFETs are more familiar to designers, the fabrication costs are relatively lower, and they also have better heat transfer rate to the substrate compared to SOI FinFETs,[8] thus bulk FinFETs are usually preferred for most digital applications. The fabrication of both types of FinFET devices is compatible with those of the conventional planar devices fabricated on either bulk or SOI wafers.

2.1.2. Device Geometry and Sizing

Unlike planar technologies for which the transistor width is a continuous value fully under the control of the circuit designer, in FinFET technologies device widths are quantized into units of whole fins. The effective gate width

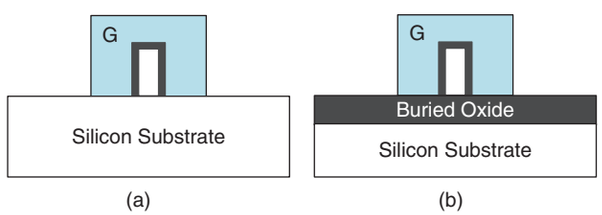

Fig. 2. Cross section view of structural differences between (a) Bulk FinFET and (b) SOI FinFET.

of a FinFET device is roughly $n(2H_{fin} + t)$, where n is the number of fins, t is the fin width and H_{fin} is the fin height as illustrated in Figure 1(b). Since the gate of a FinFET device is designed to achieve good electrostatic control over the channel, and because of the etching uniformity requirements, the fin dimensions (e.g., height H_{fin}) are not under designer control, and thus the device width cannot have an arbitrary value as in planar technologies. Wider transistors with higher on-currents are obtained by using multiple fins, but the range of choices is limited to integer values. This is known as the *width quantization* issue.[10–12] This quantization issue doesn't allow flexibility in terms of device sizing which becomes problematic especially in analog design and SRAMs. The designers need to adapt to this new constraint during the design phase.[13] An alternative solution would be for the foundry to provide the designers with multiple versions of FinFET with different fin heights.[14] For example,[15] did an early attempt by exploring the design space of FinFETs with double fin heights and showed that the lack of continuous sizing can be somewhat compensated; this method though has many uncertainties from both fabrication costs and manufacturing difficulties, so it is unlikely to become widely available. In summary, for digital circuits, width quantization might not be a big issue since most of the cell designs can be adapted to use the limited choice of device widths available.

2.1.3. FinFET Device Scaling—Fin Height

As discussed in the last section, fin height determines the overall width of a device. This is a very important parameter for circuit designers but they don't actually have control over it. Smaller fin heights offer more flexibility in terms of sizing, but this would lead to more fins, which means more silicon area. In contrast, FinFET devices with taller fins offer less flexibility with sizing but have a smaller silicon footprint and the increasing fin heights for successive FinFET nodes combines with the lateral scaling to actually accelerate "Moore's Law"—style scaling; but this might also result in larger short-channel effects and some structural instabilities.[1,8] In addition, taller devices could also lead to an increase in unwanted capacitance. This indicates that there are some opportunities for device-circuit codesign that are unlikely to become available for fabless companies but could become important for vertically-integrated companies that have their own fabs. An example of such involvement can be to analyze the design space of current versus capacitance for different fin heights. As the technology node approaches the sub-10 nm scale, this type of analysis is more and more important since the fabrication difficulties are increasing, and the design tradeoffs might drastically change.[16]

2.2. Bulk versus FDSOI versus FinFET Devices

In this section, we present some study results across multiple technology nodes (from 130 nm to 7 nm) which include both real technology nodes that are used in industry, and also predictive nodes which are widely used in academia but not tied to any specific foundry. As for the 7 nm node, we use a recently released predictive 7 nm PDK[17] which is based on current realistic assumptions for the 7 nm technology node but is not tied to or verified by a specific foundry. We believe that this analysis will provide us with a good insight on how FinFET devices are right now (with industry PDKs) and how good these devices are likely to be in the future (with predictive PDKs) as we move forward compared to the planar devices.

From a digital circuit designer's perspective, whether the technology is planar or FinFET, whether it is bulk or SOI, the parameters of interest are the same—how much current can one transistor drive, leakage, DIBL, GIDL and so on. Summarized in Table I are device parameters we extracted based on extensive simulation results across multiple technology nodes.

2.2.1. Device Models

Device models are critical for circuit designers to run simulations and make design decisions. They need to be accurate and efficient in terms of simulation time and complexity. The fact that fins are 3D structures that rise above the substrate means that they are more strongly affected by their immediate environment than planar devices. This results in a number of challenges during the modeling process. For example, the interaction between the device and its surroundings needs to be accurately modeled. Besides, the unique gate structure leads to increased gate capacitance and also to more components when modeling the parasitic capacitance and resistance compared to the planar devices.[19,20] These capacitance and resistance values are crucial since the inaccuracy caused during extracting R and C parasitic will lead to

Table I. Summary of device parameters across multiple technology nodes (extracted from I–V curves).

Technology	Physical length L_g (nm)	Nominal V_{dd} (V)	I_n/I_p (Saturation)	Subthreshold slope (mV/dec)	DIBL parameter	GIDL slope (mV/dec)	Channel length modulation λ (/V)
130 nm bulk	120	1.2	4.24	92.07	0.53	3346	0.246
Bulk[18a]	45	1.0	1.45	98.3	1.61	286	0.387
28 nm FDSOI	30	1.0	3.21	84.2	0.993	198.42	0.260
1 × nm bulk FinFET	14	0.8	0.99	71.1	0.485	429.79	0.256
7 nm bulk FinFET[17a]	20	0.7	0.90	67.6	0.745	2220.6	0.203

Note: [a]Predictive nodes.

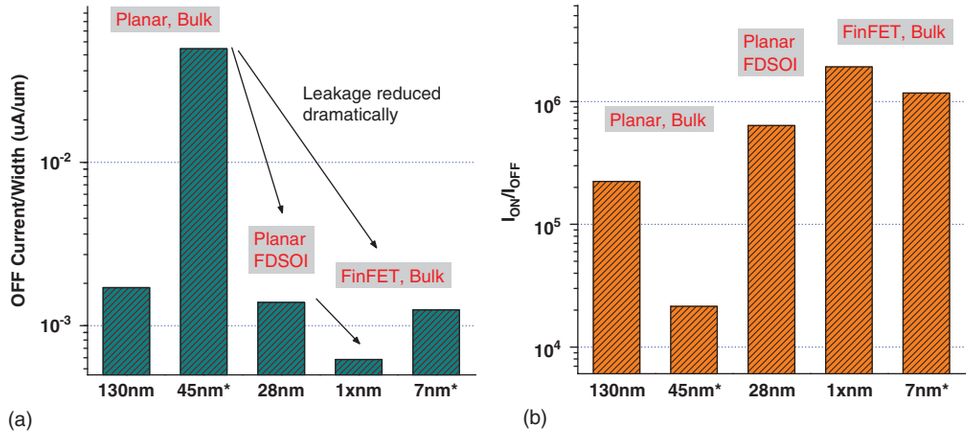

Fig. 3. (a) Leakage current with technology scaling; (b) I_{on}/I_{off} ratio with technology scaling.
Note: *Predictive technology nodes (45 nm and 7 nm).

mis-characterization and under/over-estimated design margins. Figure 4 shows an example of how FinFET parasitic capacitance is accounted for a 2-finger device. It is clear that more components contribute to both intrinsic capacitance (in the SPICE models) and extraction capacitance (accounted during extraction). For example, the gate capacitance includes gate to top of fin diffusion, gate to substrate between fins, gate to diffusion inside channel, gate to diffusion between fins, gate to contact, and so on. Similarly, the Fin-to-Fin capacitance is also newly introduced for FinFET devices. The complexity of modeling has been increasing as the device dimensions shrink. Coupling and Miller effects are more pronounced in these devices as well.

The FinFET structure brings new modeling challenges. In a planar device, the source and drain are self-aligned with the gate and often intrude slightly under it. In FinFET devices there is a spacer between the gate and the source and drain, which are usually raised and have a strain caused by a SiGe layer that creates a lattice mismatch. This means there are much more complex parasitic capacitance and resistance structures and more model calibrations are required to achieve good accuracy. As for the designer, the simulation efficiency also matters and it depends on the levels of model complexity, but thanks to the fast solvers and accurate extraction tools recently developed, the simulation time has remained tractable.

2.2.2. Leakage

One of the driving forces that leads the industry to move from bulk planar to FDSOI or FinFET technologies is the difference in leakage. With every new process generation the doubling of gate density is also associated with a doubling of the amount of leakage current.[21] This is also clear from the simulation results in Figure 3(a) where the subthreshold current (OFF current) per unit width is plotted for different technology nodes. It can be seen from the plot that, when scaling from 130 nm to 45 nm, the leakage current increases significantly, due to the fact that the channel depth underneath the gate becomes larger and a significant volume of the channel is too far away from the gate and there is a subsequent loss of electrostatic control. FDSOI and FinFET on the other hand achieve much better leakage results because the gate has much better control over the channel in these technologies. Our simulations show that 28 nm FDSOI and 7 nm FinFETs have comparable

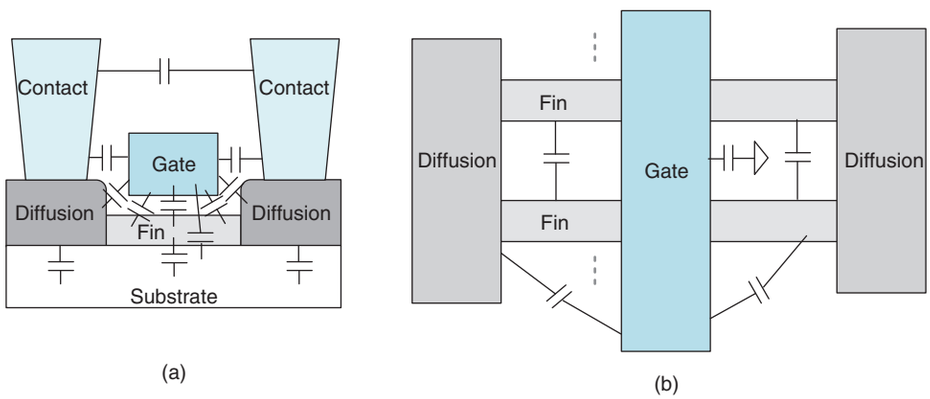

Fig. 4. Capacitance components for a FinFET device: (a) Cross-section view and (b) Top view.

leakage numbers. However, $1 \times$ nm bulk FinFET shows a reduction of leakage of at least 50%. This can be due to the fact that FDSOI and FinFET use different mechanisms to reduce leakage. In FDSOI, leakage reduction is achieved by making the channel thinner, by limiting its depth with the help of an insulating layer, while in FinFET it is achieved by making the gate wrap around the channel.

Another way of explaining the leakage reduction in FinFET devices is to look into the subthreshold slope. The sub-threshold slope also measures how fast the device can switch from OFF to ON, and the lower bound is 60 mV/dec at room temperature. Table I shows that, together with the move to FDSOI and FinFET, the subthreshold slope value has actually improved with scaling and this has resulted in a significant benefit for continuously improving frequency, active power, leakage power or a combination of the three over the past few years.[22]

2.2.3. I_{on}/I_{off} Ratio

The I_{on}/I_{off} ratio is an important figure of merit for having high performance (higher I_{on}) and low leakage power (lower I_{off}) for the devices. Since the leakage current (I_{off}) has been significantly reduced in FinFET devices, their I_{on}/I_{off} ratio is superior to bulk, as shown in Figure 3(b). This has also enabled a continuous performance improvement.

2.2.4. DIBL

Drain-Induced-Barrier Lowering (DIBL) is a short-channel effect that appears as the distance between the source and drain decreases to the extent that they become electrostatically coupled. The drain bias affects the potential barrier to carrier flow at the source junction, resulting in subthreshold current increase. To characterize it, we use the DIBL parameter, which is defined in Eq. (1) and corresponds to the change of leakage current due to V_{ds}. The smaller this parameter, the better the DIBL behavior is. It is shown in Table I that FinFETs achieve very good DIBL behaviors compared to bulk devices. In particular, the 1xnm FinFET device has the lowest DIBL effect among all five technology nodes considered.

$$\Delta \log(I_{off}) = (\text{DIBL Parameter}) \times V_{ds} \tag{1}$$

2.2.5. Channel Length Modulation (CLM)

Channel length modulation (CLM) is another short-channel effect that is caused by large drain biases. It is characterized by the CLM parameter λ which is generally proportional to the inverse of the channel length. Smaller λ means less CLM effect. Table I shows that CLM has been getting worse as the channel length shrinks in planar devices even by increasing the doping density. When technology switched from planar to FDSOI and FinFET, CLM has been improved due to the better control over the channel. Especially, in 7 nm technology node, the CLM effect is the smallest and is as good as a relatively old long-channel technology (130 nm).

2.2.6. GIDL

The introduction of high-k/metal-gate stacks in planar devices has led to substantial reduction in the gate leakage and has exposed other leakage mechanisms such as gate-induced drain leakage (GIDL) as primary gate-related leakage mechanisms.[23] GIDL occurs due to the high reverse bias between the silicon body and the drain junction (a PN-junction) near the gate edge at a nearzero or a negative gate bias.[24] GIDL usually increases as the gate length (L_g) decreases due to the floating body effect and is usually pronounced in short-channel devices. In this paper, we pick the GIDL slope to quantify this effect; the larger this slope the lesser GIDL effect the device has. Interestingly, the results in Table I indicate that as the technology switched to FinFET, GIDL has actually improved. The suppression of GIDL can be explained by the light doping of the channel and better junction placement gradient as suggested in Ref. [23]. In conclusion, FinFETs are superior to planar devices in terms of I_{on}/I_{off}, DIBL, CLM, GIDL, and thus appear to be a true "back to the future" reset of most of the metrics that were getting worse with every new technology node for bulk planar technologies!

2.2.7. W_p/W_n Ratio

Another interesting aspect for FinFET technologies is that the pull up network (PUN) and the pull down network (PDN) can become very symmetric. PMOS and NMOS devices with the same number of fins have very comparable driving strength, and the conventional 2:1 or 3:1 sizing strategy is not be applicable (or necessary) in the FinFET case. This can be seen from the I_n/I_p ratio in Table I, which is very close to 1 for the FinFET nodes. Figure 5 further demonstrates this. It plots the voltage transfer curve (VTC) under different supply voltages for a FinFET inverter with $W_p/W_n = 1$. It shows that the small-signal gain (which is the slope of the transfer curve when the input is equal to the mid-point voltage) is close to ideal (very high gain), and the curves are very balanced in all cases which further demonstrates that the ratio of 1:1 is optimal for FinFET logic.

The reason behind this fact is due to the unique fabrication process for FinFET. As opposed to planar structures which can only be fabricated in a single plane due to process variation and interfaces traps, FinFETs can be fabricated with their channel along different directions in a single die. This results in enhanced hole mobility. The N type FinFETs implemented along plane $\langle 100 \rangle$ and the P type FinFETs fabricated along plane $\langle 110 \rangle$ lead to faster logic gates since it combats the inherent mobility difference between electrons and holes.[1,25,26] Moreover, since the gate has very good control over the channel, doping concentrations can be much lower than in planar devices, thus allowing to reduce the random dopant

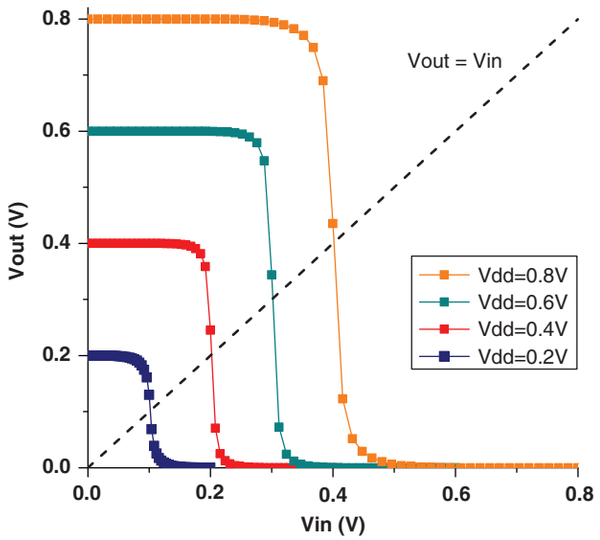

Fig. 5. VTC curves under different supply voltages for a 1xnm FinFET inverter (PMOS and NMOS are sized equally).

fluctuations (RDF),[7] mitigating the impact of mobility on current.

The symmetric PUN and PDN introduce ease in terms of physical design and sizing but it also brings slight changes in design decisions and standard cell design.

2.2.8. Alpha-Power Law

The long-channel MOSFET model (Shockley model), assumes that carrier mobility is independent of the applied fields, since the lateral or vertical electric fields were low.[27] However, for short-channel MOSFETs, the velocity of carriers reaches a maximum saturation speed due to carriers scattering off the silicon lattice. This also leads to a degradation in mobility that depends on the gate to source voltage V_{gs}.

The drain current I_d is quadratically dependent on the drain to source voltage (V_{ds}^2) in the long-channel regime and linearly dependent on V_{ds} when fully velocity saturated due to an electric field higher than a critical electric field $E_c = V_c/L_g$,[28] where V_c is the corresponding critical voltage and L_g is the gate length. A moderate supply voltage is when the transistor operates between the long-channel regime and velocity saturation. The complete model, called the α-power law model, is presented in Eq. (2):

$$I_{ds} = \begin{cases} 0, & V_{gs} < V_t \quad \text{(Cutoff)} \\ I_{dsat}\dfrac{V_{ds}}{V_{dsat}}, & V_{ds} < V_{dsat} \quad \text{(Linear)} \\ I_{dsat}, & V_{ds} > V_{dsat} \quad \text{(Saturation)} \end{cases} \tag{2}$$

where $I_{dsat} = P_c(\beta/2)(V_{gs} - V_t)^{\alpha}$ and $V_{dsat} = P_v(V_{gs} - V_t)^{\alpha/2}$. The exponent α is called the velocity saturation index, and ranges from 1 for fully velocity saturated transistors to 2 for transistors with long channel or low supply voltage.

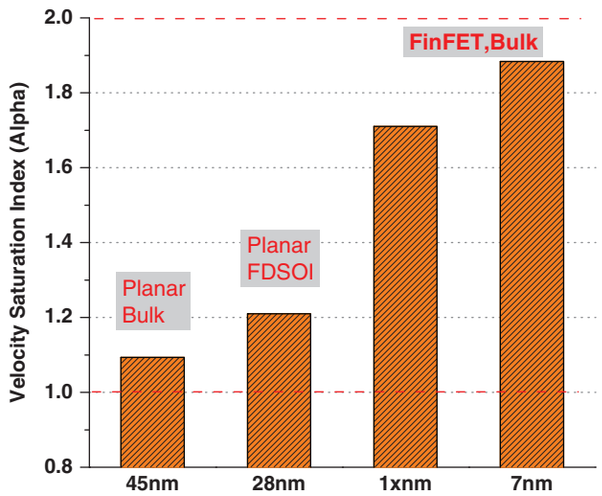

Fig. 6. Velocity saturation index (α) for different technologies.

We performed $I_{ds} - V_{gs}$ simulations for the base NMOS transistors of four different technologies and determined their respective velocity saturation index α. The results obtained, summarized in Figure 6, suggest that, as we switch to FinFETs, devices behave increasingly more according to the long-channel model, again, in a "back to the future" way.

2.3. FinFET Fabrication

In the previous section we studied the device parameters of FinFET versus planar technology nodes and found out that FinFET devices stand out in almost all the metrics. Besides, the process technology of FinFET is relatively straightforward and compatible with conventional planar device fabrication process.[29] But there are still challenges, for example, fin shape control and recess of shallow trench isolation (STI) oxide are still critical in the integration of FinFETs. Due to the space limit and the focus of this paper, we list only a few fabrication advances and challenges in the FinFET era in this section.

2.3.1. Double/Multi-Patterning

Although technologies keep scaling to the order of a few nanometers, lithography still uses 193 nm wavelength light, which makes printability and manufacturability more challenging due to increased distortion. Beyond 20 nm the use of multi-patterning is required for device fabrication. Using multi-patterning technology, a single layout is decomposed into two or more masks and manufactured through two or more exposure steps. These masks are then combined to get the original intended layout. By decomposing the layout into two or more masks as shown in Figure 7, the pitch size is effectively doubled thereby enhancing the resolution.[30] To achieve this, on the design side, color (mask) assignments are used. Several techniques of multi-patterning include Litho-Etch-Litho-Etch Double Patterning (LELE DP), Spacer-is-Metal Self-Aligned Double Patterning (SIM

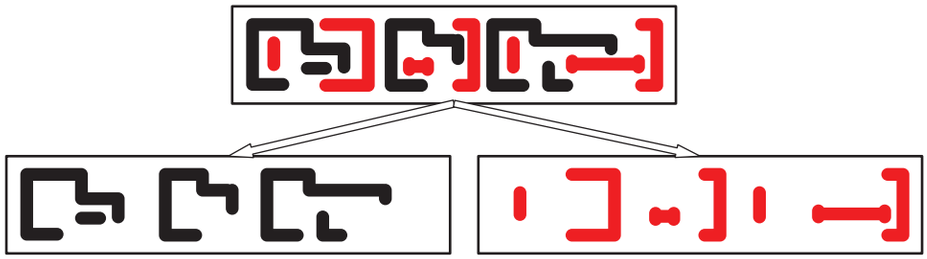

Fig. 7. Layout decomposition: A single layer is decomposed in two or more masks to enhance the resolution.

SADP), Litho-Etch-Litho-Etch-Litho-Etch Triple Patterning (LELELE TP) and Spacer-is-Dielectric Double Patterning (SID SADP). To use these techniques the designer can include the colored masks per layer that must be multipatterned or use a colourless flow where the foundry performs the decomposition.[31]

2.3.2. Fin Formation

Although multipatterning brings new fabrication challenges, some of the known fabrication steps from the planar technology can be repurposed to achieve new required shapes like the 3D fins. Sidewall spacer deposition steps from planar processes are utilized to perform self-aligned double patterning (SADP). Similarly, the steps used to form Shallow trench isolation (STI) can be extended to fabricate fins by additional etching of STI areas and thereby exposing Si fins. Fins are fabricated in a regular fashion over a large area. Thereafter unwanted fins are excised and the remaining fins become a part of active areas of the devices. Hence FinFET fabrication becomes compatible with old planar CMOS processes using repurposing of existing steps, plus a few extra steps.

2.3.3. Shape of the Fins

Several studies have shown that FinFET performance is affected by the cross-sectional area of the fin, therefore the fin shape. Intel's 22 nm node microprocessor was built with FinFET sidewalls sloping at about 8 degrees from vertical which makes more sturdy devices among other advantages.[26] Figure 8 shows the main types of fins analyzed in the literature. Experimental data shows that a FinFET with a rectangular cross-sectional area has better short channel effect metrics, in particular sub-threshold slope, GIDL and DIBL if compared with a triangular or trapezoidal cross-sectional area.[32] On the other hand a triangular fin can reduce leakage current by 70% if compared with a rectangular fin.[33]

2.3.4. Middle-End-of-Line (MEOL)

Middle-end-of-line (MEOL) is a new term introduced in the FinFET era. It refers to the intermediate process steps that complete the transistor formation (Front-end-of-line: FEOL) before contacts and interconnect formation (Back-end-of-line: BEOL).[34] MEOL is necessary to provide better cell level connections with restricted patterning capabilities and multipatterning.[35] The introduction of MEOL increases the complexity of fabrication and modeling as well. For circuit designers, the added new parasitic effects from MEOL need to be considered during the design process since these parasitics have been demonstrated to be one of the dominant sources.[36] MEOL parasitics have been usually accounted at the logic gate-level parasitic extraction step using the standard EDA tools. For physical design engineers, the added MEOL means more complex design rules and longer debugging process, also, the layout tools must automate conformance to rules as much as possible.

2.4. Summary—What Have We Learned So Far?

The studies discussed in the previous sections show that FinFET devices outperform planar devices (bulk and SOI) in almost all aspects. In particular, much less leakage current enable a wide range of applications from high-end to energy-constrained applications. Better I_{on}/I_{off} ratio have led to continuous performance improvement compared to planar at the same node. FinFET devices also provide improved sub-threshold and short-channel behavior. An added advantage of the FinFET is that it can be easily fabricated along different channel planes in a single die, and this makes sizing strategy simpler. The added MEOL made the transition from planar devices to FinFETs slightly more complex in terms of the fabrication and parasitics but the

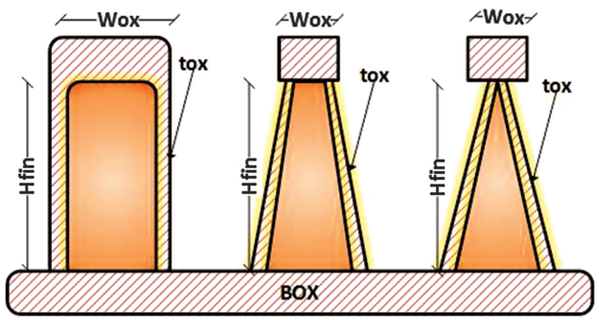

Fig. 8. Left side: A fin with a vertical slope which presents better short channel metrics.[32] Middle: A standard fin with some degree of inclination as the one used in the 22 nm Intel's node.[26] Right side: A fin with a triangular cross-sectional area that can help to reduce the leakage.[33]

back-end of the process is essentially the same, and therefore the part of the design flow associated with the physical implementation remains similar.[7]

3. FINFET CIRCUITS

Since FinFET devices have much better electrostatic properties and other metrics than planar devices, new logic and wider design space exploration opportunities become available. In this section, we discuss these new changes that FinFETs have introduced at the circuit level.

3.1. Logic Styles

As discussed in Section 2.1.1, FinFETs come in two flavors—short-gated (SG) and independent-gated (IG). For IG FinFETs, the top part of the gate is etched out, resulting in two independent gates. Because the two independent gates can be controlled separately, IG-mode FinFETs offer more design styles.[8,9] Although the gates are electrically isolated, their electrostatics are highly coupled. The threshold voltage of either of the gates can be easily influenced by applying an appropriate voltage to the other gate. Shown in Figure 9 is one example of different flavors of 2-input NAND gate implemented using SG/IG gate or a hybrid of both (modified from Ref. [8]). In SG mode, FinFET gates are tied together, so they work the same as the planar devices; In IG mode, one device (with two gates) is driven by two independent signals, and some logic functions can be realized by one device; in IG-Low Power mode, one gate is disabled and acts as the reverse-biased back-gate. The designers can even mix the two types of devices and balance the tradeoff if it is allowed by the foundry. But IG gate requires one more step of etching in the fabrication step.

3.2. Body Effect

Adaptive Body Biasing (ABB) has been used by circuit designers as an effective design technique to reduce the impact of die-to-die and within-die variations by changing the NMOS and PMOS threshold voltages independently in order to maximize performance.[37] FinFETs fabricated in bulk or SOI processes receive little benefit from controlled body effect because the channel in the FinFET is mostly in the top of the fin, away from the body. Thus the body bias techniques is not applicable to FinFET circuit design anymore.[38] To validate the above argument, we apply both reverse and forward body bias to a 2-finger transistor and simulate the ON current for both 1 × nm and 7 nm nodes, with the results are shown in Figure 10. The ON current doesn't change with the body voltage, as expected, and it indicates that FinFET devices are largely insensitive to the body effect. On one hand, this reduces the available design knobs, on another hand, this can mitigate the stack effect. In the following sections, we present two solutions to address these two separate aspects.

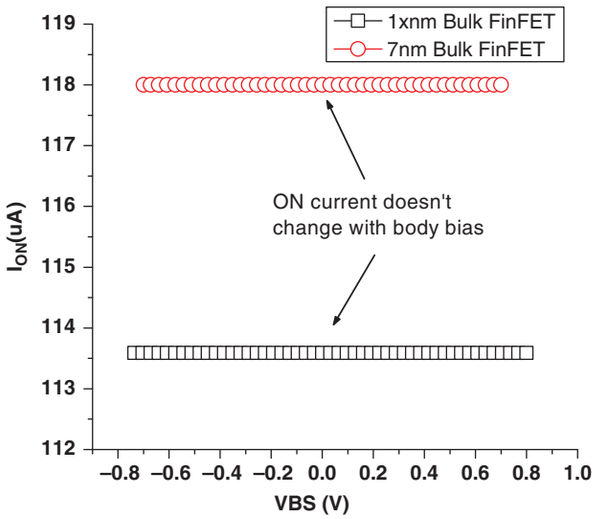

Fig. 10. ON current versus Body bias for a 2-finger 1 × nm and 7 nm NMOS transistor.

3.2.1. Gate Overdrive with Split-Circuit Biasing to Substitute for Body Biasing in FinFET

In this section, a circuit topology is presented which substitutes body biasing, but doesn't rely on the body (or back plane) voltage to do so. Additionally, the topology presented here does not require varying the voltage swing to

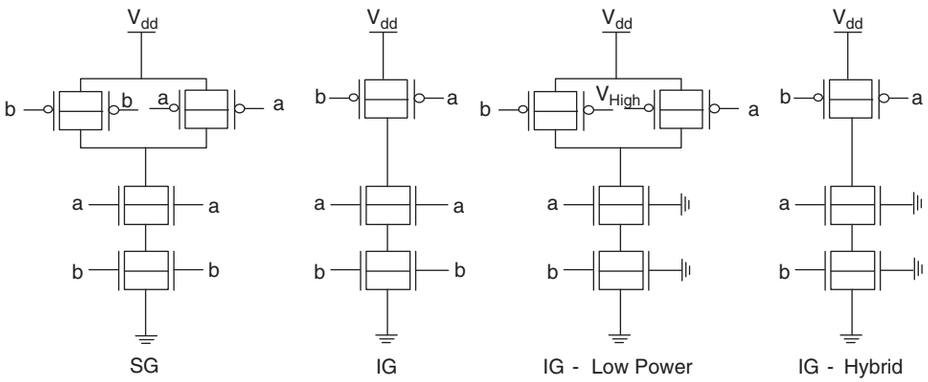

Fig. 9. Different FinFET logic styles: 2-input NAND gate designs with SG and IG devices.

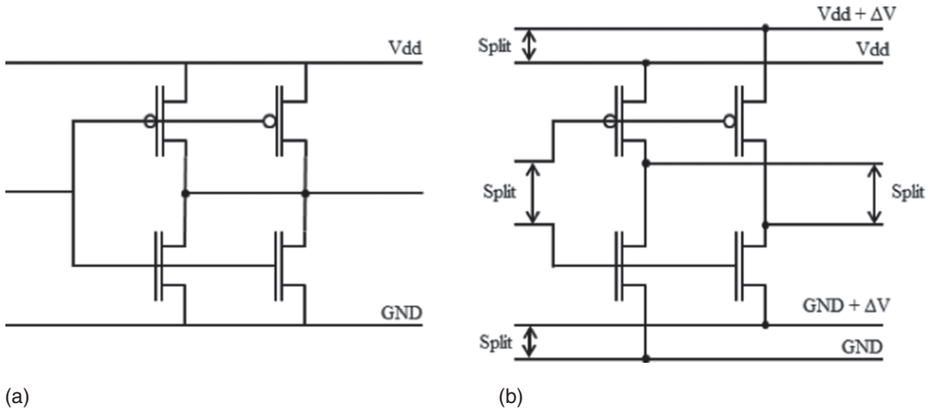

Fig. 11. (a) Static CMOS inverter with 2 fins per transistor and (b) Split-circuit inverter with split inputs, outputs, and supply rails.[39]

modulate performance and power as in DVS. The effect is achieved by splitting the inputs, outputs, and supply rails of a gate and applying a small difference between the two sets of supply rails, which will either overdrive some device gates (in forward bias) or decrease the leakage current for 'off' transistors (in reverse bias).[39]

Figure 11 shows an inverter implemented with the regular topology and the proposed split-circuit biasing topology. The idea is to regulate V_{gs} to mimic the threshold shift achieved by body biasing. In order to do this, two supply voltage domains are needed. One domain will be the nominal domain, with voltage swings from 0 to V_{dd}. The second domain will have the same differential, but both the ground and supply rail will be shifted up by some bias voltage, ΔV, such that the voltage swings of gates receiving this supply domain will swing from $0 + \Delta V$ to $V_{dd} + \Delta V$. The inputs, outputs, and supply rails of a traditional CMOS topology are split such that the number of each are doubled. Any two corresponding inputs will carry the same logic, but one is shifted up by some bias voltage, ΔV. Under forward bias, the higher inputs will drive the gates of the NMOS while the lower inputs will drive the NMOS; this will result in a higher V_{gs} for half of the NMOS and a higher absolute value of V_{gs} for half of the PMOS, and therefore a higher I_{on} for half of the devices. Under reverse bias, ΔV will be negative. The power rail voltages will be obtained by two off-chip supplies which have the same differential (V_{dd}), and one on-chip charge pump to maintain the voltage separation between the two domains (ΔV). Figure 12 shows the performance and static power response of a butterfly module of FFT for the 7 nm and 20 nm FinFET nodes to the split-circuit biasing, which enables a wide range of performance (For example, at a forward bias of 0.2 V the delay of the butterfly module reduced to 58% of the nominal delay with 7 nm). This confirms that split-circuit biasing gives effective control over device current post-fabrication in FinFET technology which cannot benefit significantly from controlled body effect.

3.2.2. *Stack Height as a Potential Design Knob in FinFET Circuit Design*

In some logic cells, NAND gate for example, several transistors are connected in series and stacked. In planar CMOS circuit, stack height is limited by the body effect; due to the body effect, the voltage between source and body of the top stacked transistor will increase the threshold voltage and will lead to performance degradation; if the stack height keeps increasing, the pull down current will become smaller and the circuit will become slower or might not

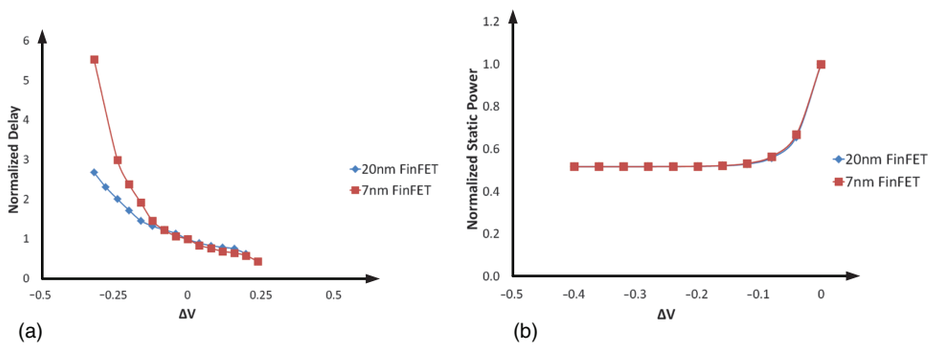

Fig. 12. Simulation results (using the predictive FinFET nodes[40]) with the Split-biasing circuit for a FFT butterfly module (a) Normalized delay from a change in inputs to a change in outputs; (b) Normalized static power in standby mode.[39]

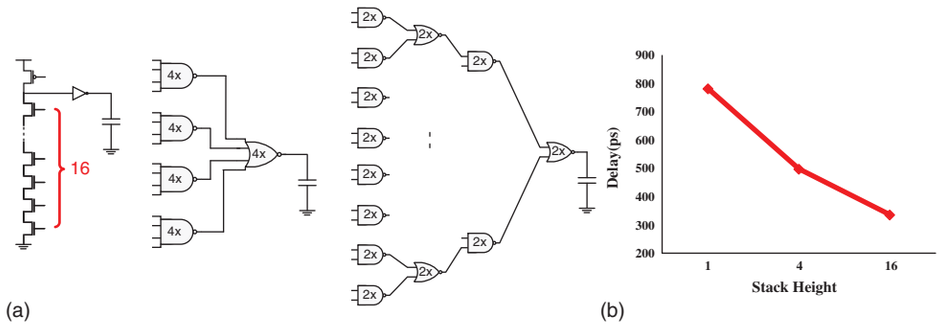

Fig. 13. (a) 16-input AND gate implemented with different stack height (1, 4 and 16); (b) 16-input AND delay simulations with different stack height (interconnect capacitance is considered).

even function correctly. For FinFET logic due to the insensitivity to the body effect as discussed above, the stack effect will be minimal and this can lead to higher stack logic cells with potential of increasing the fan-in and reducing the logic depth, thus further reducing delay and leakage paths. Our first attempt of simulating a 16-input AND gate confirms the above assumption. Shown in Figure 13(a) is a 16-input AND gate implemented with different stack heights and logic depths. Figure 13(b) shows the simulated delay in $1 \times$ nm FinFET technology corresponding to different stack height. The results suggest that a stack height of 16 and a corresponding logic depth of 2 stages achieves the best performance. Another benefit of increasing the stack height is the reduction of leakage. If we assume that the leakage with stack height of 16 design is $16I$, where I is the leakage of the unit-sized transistor, then the leakage for a stack height of 2 is $(16+8+4+2)I$, which is much larger. In summary, due to the fact that the stack effect is weak in FinFET logic, designers can increase the stack height with a relative relaxed margin to balance the tradeoffs of area, delay and leakage.

3.3. Standard Cell Libraries

There are many tradeoffs that need to be considered when developing standard cell libraries. For example, logic offerings such as the max number of logical inputs on complex gates, flip-flop and latch offerings, clk buffers, drive strength for each cell and so on. As discussed in the previous sections, FinFET devices have several unique intrinsic device characteristics, and these bring several changes to the standard cell library designers. First, with planar transistors, designers can arbitrarily change transistor width in order to manage drive current. With FinFETs, due to the width quantization fact as discussed in Section 2.1.2, they can only add or subtract fins to size it and change the current. Second, since body biasing is generally ineffective, as discussed in last section, this might lead to more logical inputs on complex gates in FinFET libraries. Coming to the physical design, the FinFET devices have periodic structures, and the optimal W_p/W_n ratio is almost 1:1, thus the FinFETs layout looks more regular, and the PMOS and NMOS regions are symmetric. The standard cell template height (in the number of M1 wiring tracks) usually comes in several flavours. For example, a high density library might be 9 tracks tall, a high performance library might be 13 tracks tall, and a power optimized library might be 10.5 tracks tall. But in FinFET, the additional constraint of fitting a fixed number of fins within a cell complicates this Ref. [4]. Especially in most FinFET technologies, fin and metal pitches are different and have not tended to line up. Power rail connections at the top and bottom of the cell typically force the removal of 1 fin each, and typically 2 additional fin tracks must be removed in the center of the cell to accommodate gate input connections, all of these make compact FinFET cell design very complex. In addition,[4] also pointed out that to meet the multiple patterning requirement, the coloring process need to be conducted during the design of the standard cells, coloring also needs to meet density solutions (each color mask must have reasonably consistent density across the chip).

3.4. Logical Effort

The logical effort method is an approximate, simplified model to analyze the delay of a gate. The normalized delay is expressed as:

$$d = f + p = g \cdot h + p \tag{3}$$

where p is the parasitic delay, i.e., the delay of the gate driving no external load, and f is the effort delay, expressed as the product of logical effort g and fanout h. The logical effort g is proportional to the complexity of a gate as a more complex gate leads to higher gate delay. The fanout h is the ratio of the output load capacitance to the input capacitance of a gate.

We estimated the g and p for an inverter, a 2-input NAND and a 2-input NOR for different technologies using simulation. For this, we use a simple simulation setup consisting of *fanout of* 1 and *fanout of* 4 gate delay chains. The results obtained are summarized in Table II. The values of g and p have been normalized to the respective inverter values for each technology. The table shows that the g and p values vary slightly across technologies

Table II. Normalized logical effort g and parasitic delay p values.

	7 nm FinFET			1 × nm FinFET			28 nm FDSOI			130 nm Bulk			Textbook		
	INV	NAND	NOR	INV	NAND	NOR	INV	NAND	NOR	INV	NAND	NOR	INV	NAND	NOR
g	1.00	1.35	1.59	1.00	1.06	1.34	1.00	1.11	1.52	1.00	1.14	1.54	1	1.33	1.67
p	1.68	2.59	3.38	0.62	1.30	0.95	2.90	4.21	4.52	0.49	0.96	0.80	1	2	2

depending on transistor sizing for different technologies. Measured normalized delays for different gates are presented in Figure 14 which shows that gates maintain a similar trend for increase in complexity across different technologies. NOR gates with stacked PMOS are slower than NANDs (stacked NMOS) even in FinFETs where the ratio of ON current in NMOS to PMOS is close to 1 as shown in Table I.

3.5. Thermal Effect Inversion (TEI)

Thermal behavior is one of the important device characteristics that affect the design decisions like margins, floorplan and cooling costs. It has been shown in the literature that temperature characteristics of FinFET-based circuits are fundamentally different from those of conventional bulk CMOS circuits.[41] In a bulk technology, if the transistor operates in the super-threshold region, the delay increases with the temperature, and in the near/sub-threshold region, the delay decreases with the increasing temperature. While in FinFET, it has been reported that the circuits run faster at higher temperatures in all supply voltage regimes (including the super-threshold one), and this is called the Temperature Effect Inversion (TEI) phenomenon.[41] In both planar devices and FinFET devices, the threshold voltage decreases at the higher temperature, and the mobility of charge carriers in the channel decreases due to the ionized impurity and phonon scattering.[42] TEI

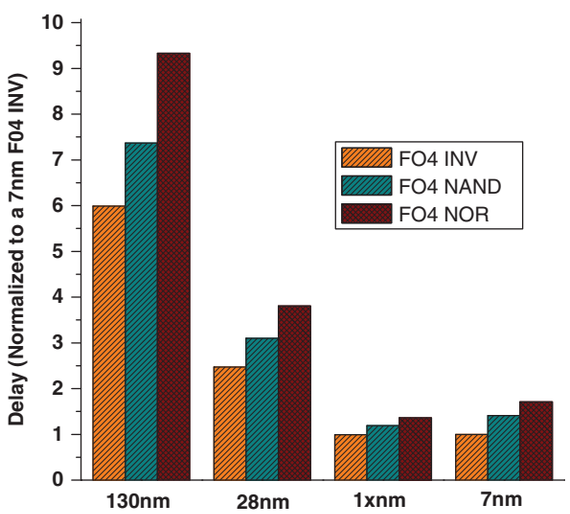

Fig. 14. Simulated FO4 delays for Inverter, 2-input NAND and 2-input NOR gates in different technology nodes (all values are normalized to the 7nm FO4 INV delay).

happens due to the fact that FinFET channels are usually undoped or lightly doped, so they exhibit only a small change in mobility with temperature. It has been shown in Ref. [43] that TEI's inflection voltage approaches nominal supply and the impact of this effect can no longer be safely discounted when scaling into future FinFET and FDSOI devices with smaller feature sizes. To validate this, we simulate the delay vs. temperature for a 9-stage ring oscillator in multiple technology nodes. The simulation results are shown in Figure 15; the results show that for all technologies, the increased temperature slow down the devices if they work under near and sub-threshold region. Interestingly, for the 28 nm FDSOI node, TEI appears across all voltages, and for 1xnm bulk FinFET node, the TEI effect has already approached 0.7 V, which is only 0.1 V below the nominal voltage (0.8 V). Similarly, for 7 nm bulk FinFET, the inversion starts from around 0.6 V (0.1 V below the nominal voltage of 0.7 V). We can conclude that the TEI effect is indeec becoming increasingly important in current and future technologies as it will cover all of the operating voltage ranges.

The TEI effect introduces new tradeoffs and also challenges in circuit design. On one hand, a higher temperature increases the leakage and cooling budget, but, on the another hand, it helps with the performance. The benefits of TEI can be maximized with the assist of novel power management techniques that can dynamically tune the voltage or frequency based on the real-time temperature[43] or novel algorithms that can determine the maximum performance under power constraints.[44] Since thermal issues also emerge as important reliability concerns throughout the system lifetime, the TEI effect can compensate some of the performance degradation introduced by reliability threats such as BTI and EM.[42,45] The optimal operating temperature can be exploited to reduce design cost and runtime operating power for overall cooling with the proper utilization of the TEI effect.

3.6. SRAM Design

SRAMs are one of the most area and power hungry components on a chip. The never-ending demand for packing more functionality per area and the requirement of higher performance from processing units leads to continuous scaling of devices.[46] This scaling trickles down to smaller bitcells and enables an increase in memory array density in terms of number of bits stored per area. Hence from the density point of view, minimum sized transistors are desired in bitcells. This translates to a 1:1:1

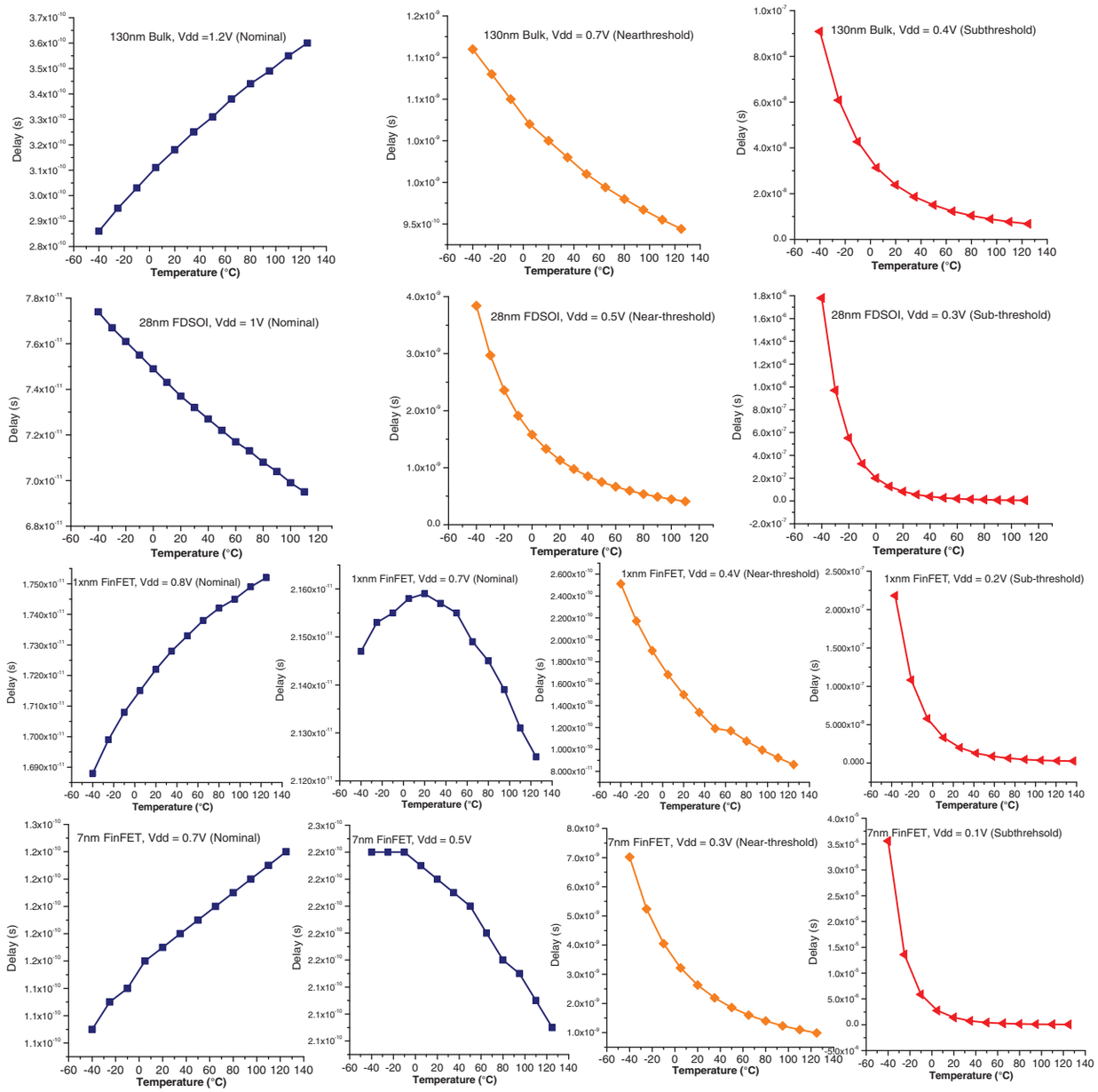

Fig. 15. Simulated temperature characteristics (delay vs. temperature) in multiple technology nodes for a 9-stage ring oscillator.

(PU:PG:PD) fin bitcell for FinFETs (where PU is the size of the Pull-up PMOS, PD is the size of the Pull-down NMOS, and PG is the size of the pass-gate NMOS in a 6T SRAM cell). The 1:1:1 bitcell provides highest array density but it suffers from flaws in terms of lower read stability and writability.[46,47] The constant need for voltage scaling to lower power further exacerbates SRAM readability and writability issues. This calls for alternate bitcells like the Low Voltage (LV) 1:1:2 cell and High Performance (HP) 1:2:2 cell[46] along with read and write assist techniques to improve SRAM metrics. Several assist techniques[48,49] have been proposed and studied to improve SRAM performance and lower operational *V*min. These techniques focus on improving PD:PG strength ratio for read assists and PG:PU strength ratio for write assists. These techniques become increasingly necessary in the era of FinFET SRAM design because transistor width quantization in terms of number of fins decreases device level sizing options to improve SRAM bitcell functionality.

3.7. Variability and Reliability

A reduced feature size causes statistical fluctuations in nanoscale device parameters which are known as process variations. They lead to mismatched device behaviors and

degrade the yield of the entire die. In planar devices, a number of dopants must be inserted in the channel which lead to Random Doping Fluctuations (RDF) causing significant variations in threshold voltage. In FinFETs, since the channel is undoped or lightly doped, this reduces the statistical impact of RDF on V_t. The variability associated with line-edge roughness (LER), the random deviation of gate line edges from the intended ideal shape, which results in non-uniform channel lengths, is also lower in FinFETs. But other process variations do appear in FinFETs. Since they have small dimensions and lithographic limitations, these devices suffer physical fluctuations on gate length, fin thickness or oxide thickness.[1,50,51] Overall, FinFETs emerge superior to planar devices by overcoming RDF and LER, which are two major sources of process variation.

Besides process variations, which represent the time-zero process variability, time-dependent variations (aging) such as Bias Temperature Instability (BTI), Hot Carrier Injection (HCI) and Electromigration (EM) also appear to be critical for reliability considerations. These aging issues conspire to worsen metrics like performance, power and lifetime. As the technology scaling is reaching the nanoscale FinFET regime, the transistors become more susceptible to voltage stress due to the increased effective field associated with the scaling of the thin oxide. Similarly, the shrinking geometries of metal layers render higher current densities, and the tremendous number of transistors within a compact area results in higher power densities. Together, these lead to increased on-chip temperatures which potentially accelerate the wearout effects.[52,53] Besides, the thermal resistance (R_{th}) of the multi-gate topology and the reduced gate pitch in FinFET devices exacerbate self-heating which will accelerate aging.[54] Figure 16 shows our simulation results with the industrial aging models; the results show a very significant performance degradation under accelerated stress condition, and if we scale this to the normal operating condition (nominal V_{dd} and normal on-chip temperature), the degradation is still much larger than that in the planar devices. For interconnect reliability, EM no longer can be signed off using aggressive margins, a comprehensive thermal-aware EM signoff methodology needs to be adopted for FinFET designs. New types of EM rules that are dependent on the direction of current flow, metal topology, via types, co-vertical metal overlaps etc. are required to address the potential reliability issues.[55]

3.8. Interconnect

As the devices become smaller and smaller, the interconnect becomes more and more dominant in determining circuit performance. This is because of the yield and EM requirements, the interconnect can't scale at the same rate as the transistors. As interconnect is becoming more compact at each node below 20 nm,[56,57] the interconnect RC

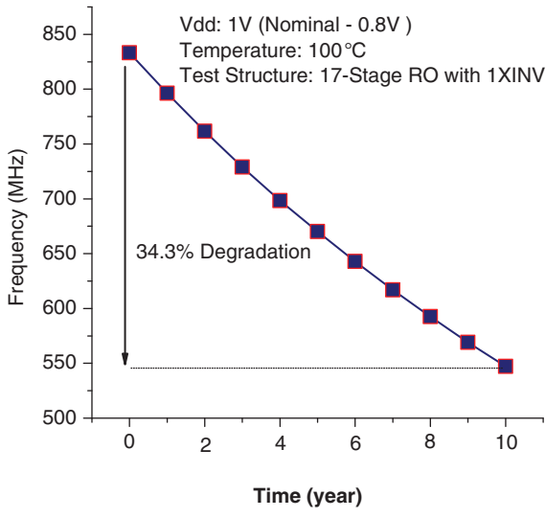

Fig. 16. Aging simulation with 1 × nm bulk FinFET with foundry-provided aging models (BTI + HCI).

parasitic delay will affect the performance in a more significant way and become one of the bottlenecks on the scaling roadmap. To address this, interconnect materials such as Aluminum, cobalt (Co) or ruthenium (Ru) could be better alternatives due to the better sheet resistance, but there are also cost and reliability considerations in the interconnect scheme design.[58] The pitch size of the metal lines also doesn't scale down that much as the technology moves into the sub-20 nm regime due to the RC parasitic and coupling consideration as well. For designers, since they don't have control over the materials and design rules, the only knob they have is the dimension of the wire. This requires to consider interconnect capacitance in the early design phase even before the physical design. The FinFET PDKs usually provide relatively accurate wire models to account this.

3.9. Power and Energy

FinFETs provide improvements in power and energy consumption since they overcome the leakage problems of planar devices and deliver better performance. To further investigate this aspect, we simulate a NAND-based ring oscillator[59] across multiple technologies. The duty cycle of the ring oscillator can be tuned and in our case, it is set as 10%. Shown in Figure 17(a) is the simulated delay versus V_{dd}, in which the values of each node are normalized to the delay at their own nominal voltages. It shows that FinFETs provide a significant performance advantage at any operating voltages, and the reduced performance due to lowering the voltage is much lower in FinFETs compared to other technology nodes as well. Figure 17(b) presents the energy versus V_{dd} plot, similar normalization is applied. As it shows, although the minimum energy optimal points are similar for all the technologies (around 0.2–0.3 V range), the energy of

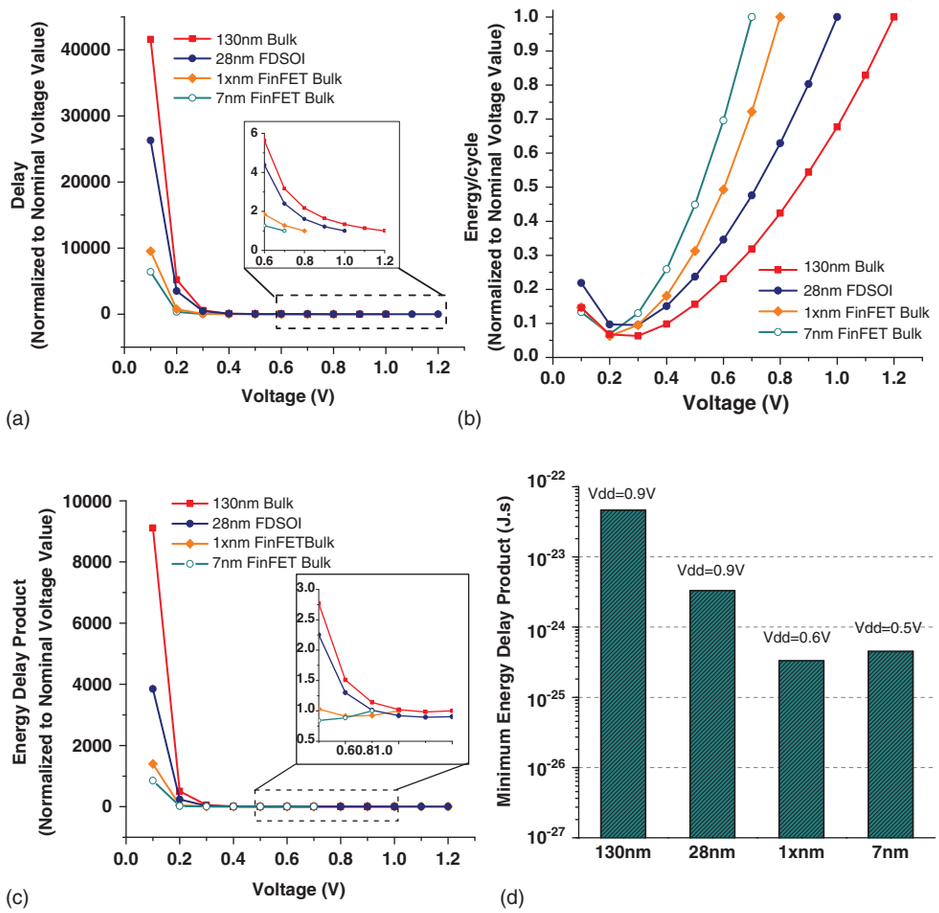

Fig. 17. (a) Delay versus V_{dd}; (b) Energy/cycle versus V_{dd}; (c) Energy Delay Product (EDP) versus V_{dd} and (d) Minimum EDP values across multiple technology nodes (simulated with the same NAND-based ring oscillator structure).

FinFET scales the best with voltage; in other words, as the voltage is scaled down, FinFETs offer more energy savings than planar devices. In Figure 17(c), the energy delay product versus V_{dd} is plotted. FinFETs offers the best energy efficiency for circuit operating under a wide range of voltages since, as the voltage scales down, the energy delay product doesn't change significantly for FinFETs compared to planar devices. Figure 17(d) presents the minimum energy delay product across the four technology nodes. As technology scales, the EDP improves as expected.

The above study shows that FinFETs provide more options for performance versus other metrics tradeoffs. For example, since FinFETs offer very good energy efficiency over a wide range of voltages, voltage scaling techniques can be very effective as designers strive to maximize performance per mW without hurting energy. FinFET-based design will be able to support wider use of dynamic voltage frequency scaling (DVFS) and enable a wider range of applications from high-end performance critical systems to energy-constraint devices.

4. SUMMARY—DIGITAL CIRCUIT DESIGN WITH FINFETS

We have shown in previous sections that FinFET devices offer significant performance improvements and power reduction compared to planar devices. Digital circuit design with FinFET broadens the design window once again. Operating voltage continues to scale down, short channel effects are reduced significantly, the process variation have been improved, the FinFET devices have lower leakage power in standby mode, etc.

Although FinFET devices offer advantages in many dimensions, they also bring challenges in the design process. FinFET devices have non-standard shapes and require complex modeling of the parasitics in the TCAD tools. Moreover, the physical layout-dependent effects have a significant impact on the metrics. Therefore, the design tools and design flows need to be able to assist the designers to build circuits that accurately correlate to the models. During the design process, extraction plays a big role to obtain accurate timing analysis and power estimation for FinFETs, so enhancement to the foundational

EDA tools, in particular SPICE simulations, extraction and physical verification that operate on part of the design below the first metal layer are required.[7] Interconnect resistance is becoming more important, so IR drop and power-grid design becomes more critical. Besides, to meet the double/multipatterning requirements, the standard cell, floorplanning, placement and route (P and R) need to be colored correctly. For example, during power planning, all power rails need to be free of double patterning violations. Similarly, all the placement of standard cells and hard macros need to be double patterning-compliant. Physical verification (e.g., DRC) engines need to be able to check and guide the designers to meet the double patterning rules. More verifications are required, and more checkpoints need to be inserted during the design phase to make sure the design specification is met.

For custom designers and standard cell designers, all of the blocks require a redesign due to the following reasons. First, the options of sizing are less granular due to the width quantization fact in FinFET, getting more drive strength will require more fins in parallel. Second, the thermal behavior and options available to circuit designers are different than what they may be used to with planar devices. For example, body biasing will be impractical, thermal effect inversion (TEI) fact introduces new tradeoffs, higher fan-in and complex logic are possible due to the insensitivity to the stack effect. As dozens of new and complicated design rules arise for FinFET devices, physical design efforts are increasing, but the bright side for FinFET devices is the more regular layout and equal P and N regions, and because of this, the foundry usually provides a template layout on which fingers and gates are already placed, physical designers don't need to start from scratch, but the layout tools still need to automate conformance to rules as much as possible.

FinFETs also offer more design options for trading performance with other metrics. As discussed in Section 3.9, one major design optimization benefit of FinFETs is much higher performance with the same energy budget. Similarly, they consume much lower power and energy to achieve equal performance to planar devices. This essentially gives designers the ability to extract the highest performance for the lowest power, which is a critical optimization for battery-powered devices. Since FinFETs have lower leakage and can operate faster, the circuit can afford to have more and fine-grained power gating structures to further save power in standby mode. Runtime techniques like DVFS can be used with a lower cost to maximize energy efficiency. On top of all these benefits, the circuit can operate in near-threshold to save energy with lower performance penalties.[60]

As more transistors fit on one chip in the FinFET era, the design flow needs to be able to handle big designs which have billions of transistor at a fullchip level, thus optimizing the runtime and reducing peak memory are necessary, and more parallelism is required. Because of the increased complexity and number of instances on chip, an increasing number of signoff corners are required to cover process and environmental variations. Addressing these new challenges together with the new, more complex design-for-manufacturing rules, including double/multipatterning, along with the increasing design scale, require close collaboration between the foundry, tool vendors and designers to fully take advantages of what FinFETs have to offer.

5. CONCLUSION

FinFETs present a new frontier for the electronics industry and have enabled high performance and power sensitive applications ranging from small portable devices to supercomputers. In this paper, we studied the changes since the advent of the FinFET devices and addressed the challenges we face with these devices. FinFETs offer benefits in many dimensions such as the significantly improved power and performance metrics and lesser short-channel effect. FinFETs endeavour to offer advantages of future scaled devices while offsetting the problems introduced by many generations of planar CMOS scaling. But new challenges also appear due to many unique properties which FinFETs have shown. Adapting to the new challenges and fully benefiting from FinFETs will require the growing knowledge and design experiences and this paper attempts to add to that knowledge base.

Acknowledgments: This work was supported by NSF grants CCF 1619127 and CCF 1543837, by DARPA under the UPSIDE and PERFECT programs and by the Center for Future Architecture Research (C-FAR), one of six SRC STARnet Centers, sponsored by MARCO and DARPA. The authors would like to thank the anonymous reviewers.

References

1. D. Bhattacharya and N. K. Jha, FinFETs: From devices to architectures. *Advances in Electronics* 2014 **(2014)**.
2. B. H. Calhoun, Y. Cao, X. Li, K. Mai, L. T. Pileggi, R. A. Rutenbar, and K. L. Shepard, Digital circuit design challenges and opportunities in the era of nanoscale CMOS, *Proceedings of the IEEE* **(2008)**, Vol. 96, pp. 343–365.
3. J. Warnock, Circuit design challenges at the 14 nm technology node, *Proceedings of the 48th Design Automation Conference*, ACM **(2011)**, pp. 464–467.
4. R. Aitken, G. Yeric, B. Cline, S. Sinha, L. Shifren, I. Iqbal, and V. Chandra, Physical design and FinFETs, *Proceedings of the 2014 on International Symposium on Physical Design*, ACM **(2014)**, pp. 65–68.
5. J.-H. Lee, Bulk FinFETs: Design at 14 nm node and key characteristics, Nano Devices and Circuit Techniques for Low-Energy Applications and Energy Harvesting, Springer **(2016)**, pp. 33–64.
6. B. Yu, L. Chang, S. Ahmed, H. Wang, S. Bell, C.-Y. Yang, C. Tabery, C. Ho, Q. Xiang, and T.-J. King, et al., FinFET scaling to 10 nm gate length. *IEEE International Electron Devices Meeting, 2002, IEDM'02* 251 **(2002)**.

7. J. Kawa, Designing with FinFETs: The opportunities and the challenges, *Synopsys White Paper*, Synopsys **(2012)**, pp. 1–8.

8. P. Mishra, A. Muttreja, and N. K. Jha, FinFET circuit design, Nanoelectronic Circuit Design, Springer **(2011)**, pp. 23–54.

9. A. Muttreja, N. Agarwal, and N. K. Jha, CMOS logic design with independent-gate FinFETs, *25th International Conference on Computer Design, 2007, ICCD 2007, IEEE* **(2007)**, pp. 560–567.

10. F. Sheikh and V. Varadarajan, The impact of device-width quantization on digital circuit design using FinFET structures. *Proc. EE241 Spring* 1 **(2004)**.

11. J. Gu, J. Keane, S. Sapatnekar, and C. Kim, Width quantization aware FinFET circuit design, *Custom Integrated Circuits Conference, 2006. CICC'06. IEEE* **(2006)**, pp. 337–340.

12. W.-K. Yeh, W. Zhang, Y.-L. Yang, A.-N. Dai, K. Wu, T.-H. Chou, C.-L. Lin, K.-J. Gan, C.-H. Shih, and P.-Y. Chen, The observation of width quantization impact on device performance and reliability for high-*k*/metal tri-gate FinFET. *IEEE Transactions on Device and Materials Reliability* 16, 610 **(2016)**.

13. B. Swahn and S. Hassoun, Gate Sizing: FinFETs versus 32 nm Bulk MOSFETs, *Design Automation Conference, 2006 43rd ACM/IEEE, IEEE* **(2006)**, pp. 528–531.

14. T.-L. Lee, C. C. Yeh, C.-Y. Chang, and F. Yuan, FinFETs with different Fin heights, US Patent 9,425, 102 **(2016)**.

15. C.-H. Lin, C.-S. Chen, Y.-H. Chang, Y.-T. Zhang, S.-R. Fang, S. Santra, and R.-B. Lin, Design space exploration of FinFETs with double Fin heights for standard cell library, *2016 IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*, IEEE **(2016)**, pp. 673–678.

16. Re-Engineering The FinFET: http://semiengineering.com/re-engineering-the-finfet/.

17. L. T. Clark, V. Vashishtha, L. Shifren, A. Gujja, S. Sinha, B. Cline, C. Ramamurthy, and G. Yeric, ASAP7: A 7-nm FinFET predictive process design kit. *Microelectronics Journal* 53, 105 **(2016)**.

18. FreePDK45 from NCSU: https://www.eda.ncsu.edu/wiki/freepdk45:contents.

19. S. S. Rodriguez, J. C. Tinoco, A. G. Martinez-Lopez, J. Alvarado, and J.-P. Raskin, Parasitic gate capacitance model for triple-gate FinFETs. *IEEE Transactions on Electron Devices* 60, 3710 **(2013)**.

20. N. Lu, T. B. Hook, J. B. Johnson, C. Wermer, C. Putnam, and R. A. Wachnik, Efficient and accurate schematic transistor model of FinFET parasitic elements. *IEEE Electron Device Letters* 34, 1100 **(2013)**.

21. C. C. Hu, Modern Semiconductor Devices for Integrated Circuits, Part 7: MOSFETs in ICs–Scaling, Leakage, and Other Topics **(2011)**.

22. P. Royer, P. Zuber, B. Cheng, A. Asenov, and M. Lopez-Vallejo, Circuit-level modeling of FinFET sub-threshold slope and DIBL mismatch beyond 22 nm, *2013 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD), IEEE* **(2013)**, pp. 204–207.

23. P. Kerber, Q. Zhang, S. Koswatta, and A. Bryant, GIDL in doped and undoped FinFET devices for low-leakage applications. *IEEE Electron Device Letters* 34, 6 **(2013)**.

24. S. Cho, J. H. Lee, S. Ouchi, K. Endo, M. Masahara, and B.-G. Park, Design of SOI FinFET on 32 nm technology node for low standby power (LSTP) operation considering gate-induced drain leakage (GIDL). *Solid-State Electronics* 54, 1060 **(2010)**.

25. T. Chiarella, L. Witters, A. Mercha, C. Kerner, M. Rakowski, C. Ortolland, L.-Å. Ragnarsson, B. Parvais, A. De Keersgieter, S. Kubicek, et al., Benchmarking SOI and bulk FinFET alternatives for planar CMOS scaling succession. *Solid-State Electronics* 54, 855 **(2010)**.

26. W. P. Maszara and M. R. Lin, FinFETs-technology and circuit design challenges, *2013 Proceedings of the ESSCIRC (ESSCIRC)*, September **(2013)**, pp. 3–8.

27. N. H. Weste and D. M. Harris, CMOS VLSI Design: A Circuits and Systems Perspective, Pearson Addison-Wesley **(2005)**.

28. T. Sakurai and A. R. Newton, Alpha-power law MOSFET model and its applications to CMOS inverter delay and other formulas. *IEEE Journal of Solid-State Circuits* 25, 584 **(1990)**.

29. H. Zhou, Y. Song, Q. Xu, Y. Li, and H. Yin, Fabrication of bulk-Si FinFET using CMOS compatible process. *Microelectronic Engineering* 94, 26 **(2012)**.

30. FinFET and Multi-Patterning Aware Place and Route Implementation: http://go.mentor.com/4h_c2.

31. Mastering the Magic of Multi-Patterning: http://go.mentor.com/4gue4.

32. Y. Liu, K. Ishii, M. Masahara, T. Tsutsumi, H. Takashima, H. Yamauchi, and E. Suzuki, Cross-sectional channel shape dependence of short-channel effects in Fin-type double-gate metal oxide semiconductor field-effect transistors. *Japanese Journal of Applied Physics* 43, 2151 **(2004)**.

33. B. D. Gaynor and S. Hassoun, Fin shape impact on FinFET leakage with application to multithreshold and ultralow-leakage FinFET design. *IEEE Transactions on Electron Devices* 61, 2738 **(2014)**.

34. A. Biddle and J. S. Chen, FinFET technology-understanding and productizing a new transistor. *A Joint Whitepaper from TSMC and Synopsys* **(2013)**.

35. M. Rashed, N. Jain, J. Kim, M. Tarabbia, I. Rahim, S. Ahmed, J. Kim, I. Lin, S. Chan, H. Yoshida, et al., Innovations in special constructs for standard cell libraries in Sub 28 nm technologies, *2013 IEEE International Electron Devices Meeting (IEDM), IEEE* **(2013)**, pp. 9–7.

36. C.-S. Lee, B. Cline, S. Sinha, G. Yeric, and H.-S. P. Wong, 32-bit processor core at 5-nm technology: Analysis of transistor and interconnect impact on VLSI system performance, *IEEE International Electron Devices Meeting (IEDM), 2016*, IEEE **(2016)**, pp. 28–3.

37. J. W. Tschanz, J. T. Kao, S. G. Narendra, R. Nair, D. A. Antoniadis, A. P. Chandrakasan, and V. De, Adaptive body bias for reducing impacts of die-to-die and within-die parameter variations on microprocessor frequency and leakage. *IEEE Journal of Solid-State Circuits* 37, 1396 **(2002)**.

38. W.-T. Chang, S.-W. Lin, C.-T. Shih, and W.-K. Yeh, Back bias modulation of UTBB FDSOI, bulk FinFET, and SOI FinFET, *2016 IEEE International Nanoelectronics Conference (INEC)*, IEEE **(2016)**, pp. 1–2.

39. A. Whetzel and M. R. Stan, Gate overdrive with split-circuit biasing to substitute for body biasing in FinFET and UTB FDSOI circuits, *2016 IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*, IEEE **(2016)**, pp. 467–472.

40. PTM Model: http://ptm.asu.edu/.

41. W. Lee, Y. Wang, T. Cui, S. Nazarian, and M. Pedram, Dynamic thermal management for FinFET-based circuits exploiting the temperature effect inversion phenomenon, *Proceedings of the 2014 International Symposium on Low Power Electronics and Design*, ACM **(2014)**, pp. 105–110.

42. W. Sootkaneung, S. Chookaew, and S. Howimanporn, Combined impact of BTI and temperature effect inversion on circuit performance, *2016 IEEE 25th Asian Test Symposium (ATS)*, IEEE **(2016)**, pp. 310–315.

43. Y. Zu, W. Huang, I. Paul, and V. J. Reddi, T_i-states: Processor power management in the temperature inversion region, *49th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO), 2016*, IEEE **(2016)**, pp. 1–13.

44. E. Cai and D. Marculescu, TEI-turbo: Temperature effect inversion-aware turbo boost for FinFET-based multi-core systems, *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design*, IEEE Press **(2015)**, pp. 500–507.

45. E. Cai, D. Stamoulis, and D. Marculescu, Exploring aging deceleration in FinFET-based multi-core systems, *2016 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, IEEE **(2016)**, pp. 1–8.

46. D. Burnett, S. Parihar, H. Ramamurthy, and S. Balasubramanian, Finfet sram design challenges, *2014 IEEE International Conference on IC Design Technology* May **(2014)**, pp. 1–4.

47. W. P. Maszara, Finfets: Designing for new logic technology, Micro- and Nanoelectronics: Emerging Device Challenges and Solutions, CRC Press **(2014)**, pp. 113–136.

48. B. Zimmer, S. O. Toh, H. Vo, Y. Lee, O. Thomas, K. Asanovic, and B. Nikolic, Sram assist techniques for operation in a wide voltage range in 28-nm cmos. *IEEE Transactions on Circuits and Systems II: Express Briefs* 59, 853 **(2012)**.

49. T. Song, W. Rim, S. Park, Y. Kim, G. Yang, H. Kim, S. Baek, J. Jung, B. Kwon, S. Cho, H. Jung, Y. Choo, and J. Choi, A 10 nm finfet 128 mb sram with assist adjustment system for power, performance, and area optimization. *IEEE Journal of Solid-State Circuits* 52, 240 **(2017)**.

50. E. Baravelli, M. Jurczak, N. Speciale, K. De Meyer, and A. Dixit, Impact of LER and random dopant fluctuations on FinFET matching performance. *IEEE Transactions on Nanotechnology* 7, 291 **(2008)**.

51. X. Wang, A. R. Brown, B. Cheng, and A. Asenov, Statistical variability and reliability in nanoscale FinFETs. *2011 IEEE International Electron Devices Meeting (IEDM), IEEE* 5 **(2011)**.

52. S. Taghipour and R. N. Asli, Aging comparative analysis of high-performance FinFET and CMOS flip-flops. *Microelectronics Reliability* 69, 52 **(2017)**.

53. X. Guo and M. R. Stan, Implications of accelerated self-healing as a key design knob for cross-layer resilience. *Integration, the VLSI Journal* 56, 167 **(2017)**.

54. H. Jiang, S. Shin, X. Liu, X. Zhang, and M. A. Alam, The impact of self-heating on HCI reliability in high-performance digital circuits. *IEEE Electron Device Letters* 38, 430 **(2017)**.

55. Reliability Challenges in 16 nm FinFET Design: http://semiengineering.com/reliability-challenges-16nm-finfet-design/.

56. Interconnect Challenges Rising: http://semiengineering.com/interconnect-challenges-grow-3/.

57. N. Lu and R. A. Wachnik, Modeling of resistance in FinFET local interconnect. *IEEE Transactions on Circuits and Systems I: Regular Papers* 62, 1899 **(2015)**.

58. S. Tu, Putting the pieces together in the materials space: Advanced materials solutions for 10 nm and beyond. *SEMICON Taiwan* **(2015)**.

59. A. Wang, A. P. Chandrakasan, and S. V. Kosonocky, Optimal supply and threshold scaling for subthreshold CMOS circuits, *2002 Proceedings IEEE Computer Society Annual Symposium on VLSI*, IEEE **(2002)**, pp. 7–11.

60. N. Pinckney, L. Shifren, B. Cline, S. Sinha, S. Jeloka, R. G. Dreslinski, T. Mudge, D. Sylvester, and D. Blaauw, Near-threshold computing in FinFET technologies: Opportunities for improved voltage scalability, *Proceedings of the 53rd Annual Design Automation Conference*, ACM **(2016)**, p. 76.

Xinfei Guo

Xinfei Guo *received the B.S. degree in Microelectronics from Xidian University, Xi'an, China, in 2010 and the M.S. degree in Electrical and Computer Engineering from the University of Florida, Gainesville, FL, in 2012. Currently, he is a Ph.D. candidate in Computer Engineering at the University of Virginia, Charlottesville, VA. He has a broad interest in digital circuit and microarchitectures. His current research focus on Reliability (Wearout and Accelerated Recovery Techniques), Cross-layer power and reliability co-design methodology, Low-power and Energy-efficient design. He is a student member of the IEEE and ACM and a recipient of the 2017 IEEE Circuits and Systems (CAS) Pre-Doctoral Scholarship. He also received the Best Paper Award at SELSE 2017 and A. Richard Newton Young Student Fellowship in 2013.*

Vaibhav Verma

Vaibhav Verma *received his B.E. in Electrical and Electronics Engineering from Birla Institute of Technology and Science-Pilani, Hyderabad, India in 2013. He worked as an R&D Engineer at Synopsys, India from 2013 to 2016. He is currently working towards his Ph.D. in Electrical Engineering at the University of Virginia. His research interest lies in energy-efficient circuit design and high-performance, low-power memory design.*

Patricia Gonzalez-Guerrero

Patricia Gonzalez-Guerrero *received her M.S. degree in Electrical Engineering from the University of Virginia in 2014. She worked as an ASIC design and verification engineer in the ESSN R&D in Hewlett Packard, Costa Rica. She received, her B.S. degree in electronics engineering from Pontifical Xavierian University, Bogota, Colombia. She is currently working towards her Ph.D. at the University of Virginia. Her main research interests are system level integration, applications influence in system level design, subthreshold digital design for ultralow power-low frequency systems and mixed signal design.*

Sergiu Mosanu

Sergiu Mosanu *received the B.Sc. degree in Electrical and Computer Engineering from Jacobs University, Bremen, Germany, in 2011 and the M.Sc. degree in Information and Media Technologies from Hamburg University of Technology, Hamburg, Germany, in 2014. He is currently a Ph.D. candidate in Electrical Engineering at the University of Virginia, in Charlottesville, VA as part of the High-Performance Low-Power Laboratory. He is passionate about reconfigurable hardware, heterogeneous computing and emerging computer architectures. His current research activities focus on security for small low-power IoT devices, processing in memory techniques as well as testing and verification of RISC-V systems. During his studies he also did several research oriented internships at the German Aerospace Center, German Center for Artificial Intelligence, Micron Technology, and collaborated with Vector Foiltec and Airbus.*

Mircea R. Stan

Mircea R. Stan *received the Ph.D. (1996) and the M.S. (1994) degrees in Electrical and Computer Engineering from the University of Massachusetts at Amherst and the Diploma (1984) in Electronics and Communications from Politehnica University in Bucharest, Romania. Since 1996 he has been with the Charles L. Brown Department of Electrical and Computer Engineering at the University of Virginia, where he is now a professor. Professor Stan is teaching and doing research in the areas of high-performance low-power VLSI, temperature-aware circuits and architecture, embedded systems, spintronics, and nanoelectronics. He leads the High-Performance Low-Power (HPLP) lab and is a co-director of the Center for Automata Processing (CAP). He has more than eight years of industrial experience, has been a visiting faculty at UC Berkeley in 2004–2005, at IBM in 2000, and at Intel in 2002 and 1999. He has received the NSF CAREER award in 1997 and was a co-author on best paper awards at SELSE 2017, ISQED 2008, GLSVLSI 2006, ISCA 2003 and SHAMAN 2002. He was the chair of the VLSI Systems and Applications Technical Committee (VSA-TC) of IEEE CAS in 2005–2007, general chair for ISLPED 2006 and for GLSVLSI 2004, technical program chair for NanoNets 2007 and ISLPED 2005, and on technical committees for numerous conferences. He is a Senior Editor for the IEEE Transactions on Nanotechnology since 2014, and was an AE for the IEEE Transactions on Nanotechnology in 2012-2014, IEEE Transactions on Circuits and Systems I in 2004–2008 and for the IEEE Transactions on VLSI Systems in 2001–2003. He was Guest Editor for the IEEE Computer special issue on Power-Aware Computing in December 2003 and a Distinguished Lecturer for the IEEE Circuits and Systems (CAS) Society in 2012–2013 and 2004–2005, and for the Solid-State Circuits Society (SSCS) in 2007–2008. Professor Stan is a Fellow of the IEEE, a member of ACM, and of Eta Kappa Nu, Phi Kappa Phi and Sigma Xi.*