Received 20 January 2017; revised 23 March 2017; accepted 28 March 2017. Date of publication 30 March 2017;
date of current version 24 April 2017. The review of this paper was arranged by Editor C. C. McAndrew.

Digital Object Identifier 10.1109/JEDS.2017.2689738

GAAFET Versus Pragmatic FinFET at the 5nm Si-Based CMOS Technology Node

YA-CHI HUANG[1], MENG-HSUEH CHIANG[1] (Senior Member, IEEE), SHUI-JINN WANG[1], AND JERRY G. FOSSUM[2] (Life Fellow, IEEE)

1 Institute of Microelectronics, Department of Electrical Engineering, National Cheng Kung University, Tainan 701, Taiwan
2 Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611, USA

CORRESPONDING AUTHORS: M.-H. CHIANG (e-mail: mhchiang@mail.ncku.edu.tw) and J. G. FOSSUM (e-mail: fossum@tec.ufl.edu)

This work was supported in part by the Ministry of Science and Technology, Taiwan, in part by the National Center for High-Performance Computing, in part by the National Nano Device Laboratories, and in part by the National Chip Implementation Center.

ABSTRACT Speed and power performances of Si-based stacked-nanowire gate-all-around (GAA) FETs and pragmatic ultra-thin-fin FETs at the 5nm CMOS technology node are projected, compared, and physically explained based on 3-D numerical simulations. The respective device domains are also used to compare integration densities based on 6T-SRAM layouts. Predicted comparable performances and densities, with considerations of the complexity/cost of GAAFET processing versus that of the FinFET with pragmatic simplifications, suggest that the FinFET is the better choice for the future.

INDEX TERMS FinFET, GAAFET, G-S/D underlap, ultra-thin body.

I. INTRODUCTION

The pervasive presumption regarding the continued scaling of CMOS technology is that gate-all-around (GAA) FETs, with vertical or lateral nanowire channels/ultra-thin bodies (UTBs), will replace FinFETs as the primary device at or around the 7nm node [1]–[4]. In fact, the latest (2015) ITRS calls the GAA device the "ultimate structure" [1]. The more complex and costly processing of GAAFETs, clearly implied in [5] and [6], could be prohibitive but has not been defined nor seriously assessed at this point in time. The purported superiority of the GAAFET is based solely on its better control of short-channel effects (SCEs) at gate lengths of 12-14nm and below, which suggests better performance. Indeed, for very low-voltage operation ($V_{DD} \sim 0.6V$), the achievement of a steep subthreshold slope (SS) with low DIBL can yield, with low off-state current (I_{off}), relatively high on-state current (I_{on}) and perhaps faster CMOS speed [7].

The FinFETs generally assumed for comparisons with GAAFETs in support the noted presumption are tri-gate devices on bulk-Si substrates. The possible lack of I_{off} and V_t control in such FinFETs due to S-D punch-through and its (random) high–doping remedy [8] is typically ignored. In this paper, we use physical insights and 3-D numerical device simulations to argue that a pragmatically designed FinFET [7], with relatively simple processing, is a viable alternative device at 5nm (as well as 7nm), which, in the processing-performance tradeoff, can obviate the stacked-nanowire GAAFET (as well as other GAA structures). Comparisons of speed, power, and integration density, based on Sentaurus [9] I-V and C-V simulations of the two Si-based devices, and considerations of the severe complexity and high cost of the GAAFET technology relative to that of the pragmatic FinFET underlie our argument. Indeed, our argument is commensurate with recent GAA works that suggest serious issues of high parasitic capacitance, low I_{on}, excessive variability, and high series resistance [10], [11] that would undermine performance and/or necessitate more complex processes such as co-optimization of junction formation and nanowire release [5] and an inner spacer technology to reduce parasitic capacitance [6].

We note that this work presents a basic succinct benchmark, with simplifications in the simulations as well as in the extent of the benchmark, to provide clear insights into the results. Our FinFET-GAAFET comparison implies general conclusions that are not inconsistent with a more thorough comparison [12], which may involve more uncertainty in the device structures and physical modeling.

II. DEVICE DESIGNS/DOMAINS AT 5NM

We describe the design of a pragmatic FinFET (PFFET) at the 5nm node, giving physical insights to guide the design, and then define a counterpart GAAFET based on it and previous GAA-technology studies [2], [3]. The PFFET [7] is designed as follows:

*SOI structure (possibly on a bulk-Si wafer), which eliminates the noted S-D punch-through issue;

*undoped fin UTB/channel, which eliminates random-doping effects;

*metal gate on classical SiON; no high-k dielectric, suggested by predominant bulk inversion in the on- as well as the off-states which undermines the high-k benefit [7];

*G-S/D underlap, which can augment SCE control without much increase in series resistance [13];

*no lattice strain, the viability of which is dubious at the 5nm node [14].

The PFFET effects a good tradeoff between CMOS performance and process complexity for nanoscale gate lengths [7]. We assume a stacked-nanowire GAAFET with the same features noted above for the PFFET.

The device structures are illustrated in Figs. 1 and 2; their pertinent dimensions are given in Table 1. We assume a gate length $L_g = 12$nm for both devices at the 5nm node, consistent with the typical L_g-vs.-gate pitch relation in advanced nanoscale CMOS technologies [1]–[3]. Without high k, we let the gate-oxide thickness be 1.0nm for both devices. We assume uniform doping (N_{SD}) in the S/D regions, and a lateral Gaussian doping profile [$N_{SD}(y)$] in the spacer regions. The spacer length $L_{sp} = 4$nm and the lateral straggle assumed for $N_{SD}(y)$ define the effective G-S/D underlap [13], which we estimate yields a nominal effective channel length $L_{eff} \cong 16$nm for weak-inversion conditions; it shrinks to about L_g for strong-inversion conditions as the Debye length in the spacer diminishes, thereby enabling a beneficial tradeoff for series resistance (or I_{on}) versus SCE control (or I_{off}). (We note that L_{eff} is roughly defined by the locations where N_{SD} drops to 10^{19}cm^{-3}, but this location varies with L_{eff} [13].) The bodies are not intentionally doped; a low doping density ($N_B \sim 10^{15}$cm^{-3}) exists due to the natural dopants in silicon.

The channel/UTB thickness (t_{Si}) is crucial. It must be thin enough to adequately suppress SCEs, but not so thin to cause prohibitive V_t variations due to quantum carrier-confinement (QM) effects. For the PFFET, these criteria imply $t_{Si} \sim L_{eff}/2 > \sim 4$nm [7]. For the GAAFET, which we model as quadruple-gate (with a square UTB cross-section of side t_{Si}), first-order analysis of the 3-D Poisson equation in the UTB shows that t_{Si} can be ~1.4x thicker for the same SCE control [15]; and an approximate 2-D extension of the 1-D QM analysis in [16] shows that t_{Si} must be at least ~1.3x thicker to avoid excessive V_t variations. Note then that the latter QM-based requirement does not undermine the good SCE control in the GAAFET. Accordingly, we choose $t_{Si} = 6$nm for both devices, meaning that the GAAFET will give better suppression of SCEs, and hence

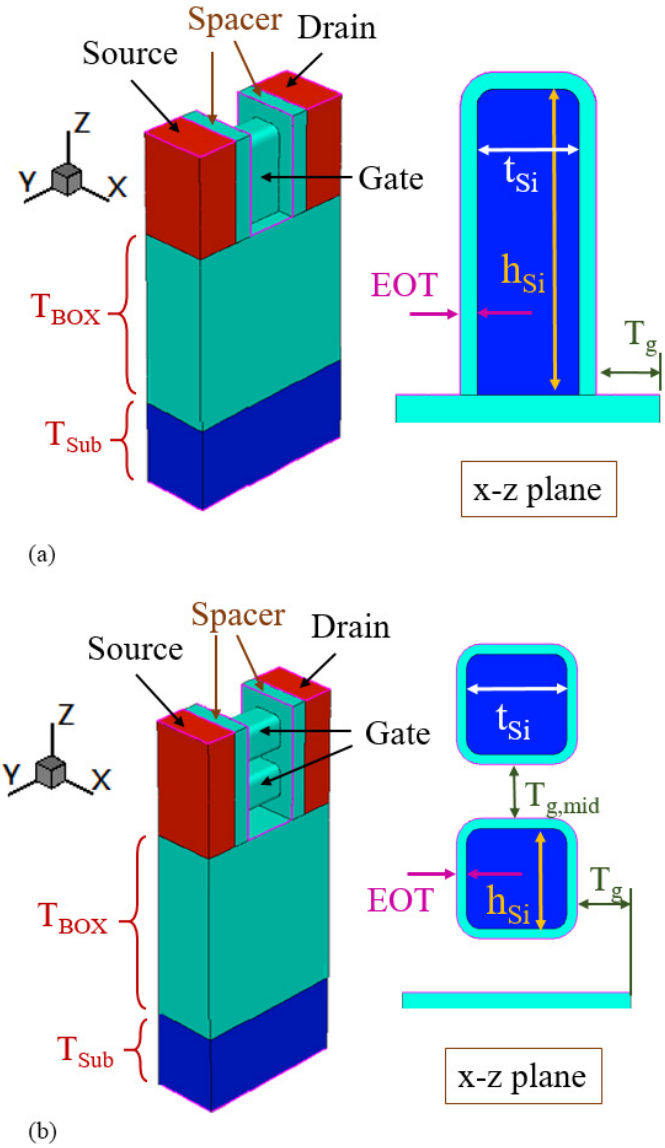

FIGURE 1. 3-D views and channel cross-sections of (a) PFFET and (b) GAAFET (not to scale).

tend to higher I_{on} at low V_{DD}. Assuming a reasonable fin aspect ratio, we let the PFFET height be $h_{Si} = 24$nm, and we presume the same height for the composite GAAFET stack, allowing two 6nm nanotubes as shown in Fig. 2.

We use Sentaurus [9] for 3-D numerical simulations of the n-channel devices operating at $V_{DD} = 0.6$V. We neglect the QM effects, which are small for the assumed t_{Si} as noted above. (In fact, supplemental Sentaurus simulations of the two devices with QM effects validate this simplification.) Classical drift-diffusion transport using the Philips unified mobility model, surface scattering, and high-field velocity saturation is selected for the simulations, but velocity overshoot is accounted for by assuming an effective saturated drift velocity $v_{sat(eff)} = 3 \times 10^7$cm/s [17] greater than v_{sat}. Other device-related parameters are listed in Table 1.

TABLE 1. Parameters of the simulated PFFET and GAAFET.

Design rule		Symbol	Value (nm)
Perpendicular (x-axis)	Contact width	C_W	10
	Equivalent oxide thickness	EOT	1
	Contact edge to diffusion	C_D	10
	Poly to poly	P	7.5
	Poly to dif. ext.	T_g	5
	Channel thickness	t_{Si}	6
	Fin pitch	FP	14
Longitudinal (y-axis)	Gate length	L_g	12
	Contact length	C_L	10
	Gate to contact	L_{sp}	4
	Gate pitch	GP	33
Vertical (z-axis)	Channel height	h_{Si}	24
	Gate to gate thickness	$T_{g,mid}$	4.6
	BOX thickness	T_{BOX}	100
	Substrate thickness	T_{Sub}	20

III. INTEGRATION DENSITY COMPARISON

Using the device domains described in the previous section, we first compare integration densities of the PFFET and GAAFET technologies based on the 6T-SRAM cell. The fin pitch (FP) and contacted gate pitch (GP) are typically scaled by 0.70x and 0.78x for each technology node [18], respectively, which gives 14nm for FP and 33nm for GP at the 5nm node based on FP (42nm) and GP (70nm) of the 14nm node. We use different numbers of fins per device for both technologies, denoting the pull-up, pass-gate, and pull-down transistor-width ratio as PU:PG:PD. The three typical designs of SRAM bitcells are high-density (HD), high-performance (HP), and low-voltage (LV), with assumed ratios of 1:1:1, 1:2:2, and 1:1:2, respectively [19]. A schematic of a 6T-SRAM bitcell is shown in Fig. 3, along with the layout of the HD design.

We find comparable PFFET and GAAFET cell areas for all designs. In fact, with the same h_{Si} for the PFFETs and the GAAFETs in the HD SRAM cell, we find exactly equal cell areas ($0.008\mu m^2$).

IV. SIMULATION RESULTS - PERFORMANCE COMPARISON

We next compare the predicted drive currents (I_{on}) and intrinsic delays (CV/I) of the PFFET and GAAFET technologies, targeted at high-performance (HP) and low-operation-power (LOP) applications, and give physical insights on the results. The leakage currents (I_{off}) were set to 10nA and 0.1nA for HP and LOP, respectively, by tuning the metal-gate work functions. (We define I_{off} and I_{on} as the actual currents in the PFFET and 2-tube GAAFET; normalizing by an effective device width is meaningless and misleading because of predominant bulk inversion in the undoped channels.)

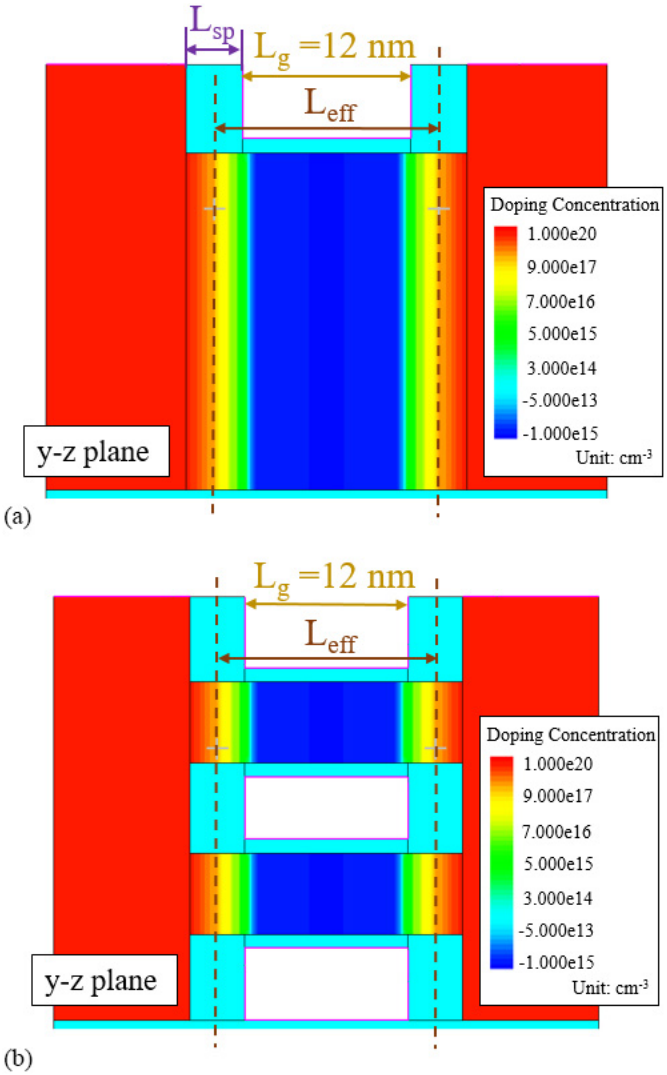

FIGURE 2. 2-D cross-sections along y-z plane of (a) PFFET and (b) GAAFET (not to scale).

We list the simulated characteristics for the nominal L_{eff} = 16nm HP nMOS devices in Table 2 (pMOS trends are similar). As expected, better suppression of SCEs in the GAAFET yields better DIBL and SS, yielding lower V_t and higher I_{on} for the same I_{off}. However, due in part to higher series resistance in the GAA device structure [11], the PFFET I_{on} is nonetheless comparable to that of the GAAFET, and the G-S/D capacitance (C) is also comparable, even with a larger gate area, due to more GAA-structure parasitics; so CV/I of the PFFET and GAAFET are about the same. (We note further that the mentioned supplemental Sentaurus simulations predicted slightly larger QM effects for the GAAFET with more spatial confinement, thus rendering the PFFET performance better comparatively.) We also note that a rigorous analysis in [20] of respective SOI-device parasitic capacitances showed, at the 7nm node, that the GAAFET has slightly higher capacitance than the FinFET, thus solidifying our conclusion regarding comparable speeds.

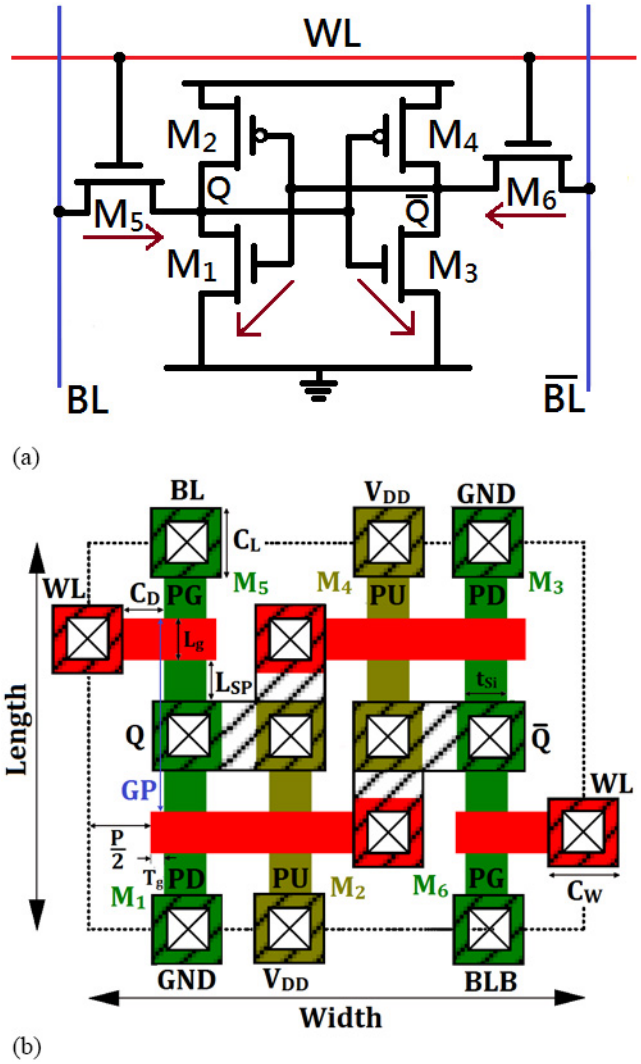

FIGURE 3. **(a) Circuit schematic of the 6T-SRAM bitcell. (b) Layout of the high-density cell (not to scale).**

TABLE 2. **Simulated characteristics of the nominal HP (I_{off} = 10nA) nMOS devices (L_{eff} = 16nm).**

	I_{on} (μA)	SS (mV/dec)	DIBL (mV/V)	V_t (mV)	C (10^{-17}F)	CV/I (ps)
PFFET	41.60	76.45	65.46	132	2.865	0.413
GAAFET	42.10	67.58	21.82	110	2.834	0.404

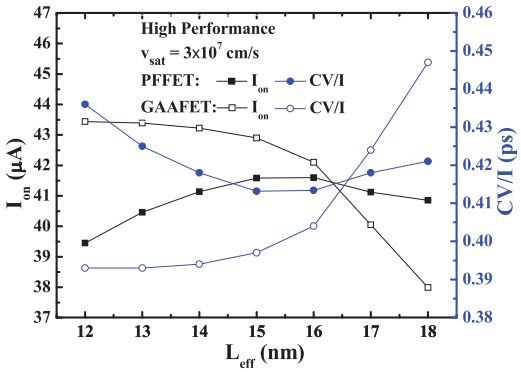

FIGURE 4. **Simulated drive currents and intrinsic delays of nMOS PFFETs and GAAFETs versus L_{eff} for the HP application (I_{off} = 10nA).**

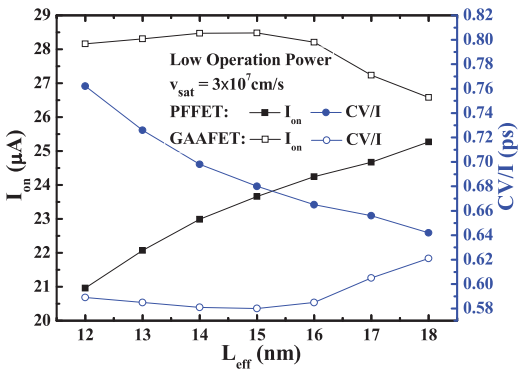

FIGURE 5. **Simulated drive currents and intrinsic delays of nMOS PFFETs and GAAFETs versus L_{eff} for the LOP application (I_{off} = 0.1nA).**

Finally, we note that additional simulations assuming high-k dielectrics (with EOT = 0.7nm) confirm our earlier comment about no high-k benefit at the ultimate nodes. For the HP GAAFET, I_{on} only increases from 42.1μA to 46.6μA with high k, and CV/I also increases from 0.40ps to 0.43ps, still longer than that of the PFFET. A recent study noted an issue of current reduction in GAAFETs due to surface-roughness and remote-Coulomb scattering associated with high-k gate dielectric [10]. This issue further stresses the no high-k benefit. The same study also noted another issue of additional parasitic G-S/D capacitance due to the protruded gate between two vertically stacked tubes of the GAAFET [10], which is not present in a comparable FinFET. This additional capacitance is not included in our GAAFET simulations (with uniform spacer thickness from top to bottom), suggesting that the PFFET would compare even more favorably than we have shown.

Since the lateral $N_{SD}(y)$ straggle, linked to technology and yield, defines the effective G-S/D underlap and L_{eff}, we now check the impact of the straggle on performance. Figures 4 and 5 show the drive currents and intrinsic delays of PFFETs and GAAFETs versus L_{eff} (via different straggles) for HP and LOP applications, respectively. For HP, comparable speeds are predicted as in Table 2, and in fact the PFFET can be even faster for L_{eff} > 16nm. Increasing the G-S/D underlap increases I_{on} via improved SCEs until increased S-D resistance becomes predominant and lowers it; the optimal PFFET L_{eff} is 15-16nm. Decreasing the GAAFET L_{eff} improves its performance, but note that it cannot be much less than 15-16nm due to processing limitations needed to avoid random S/D dopants in the undoped UTB/channel.

For LOP, the nominal (L_{eff} = 16nm) PFFET I_{on} is 14% lower than that of the GAAFET, and its CV/I is 14% longer. The superior SCEs of the GAAFET do yield a significant performance gain here, but note in Fig. 5 how a longer L_{eff} and the improved SCEs it affords benefit both I_{on} and CV/I of the PFFET, rendering it more competitive with the

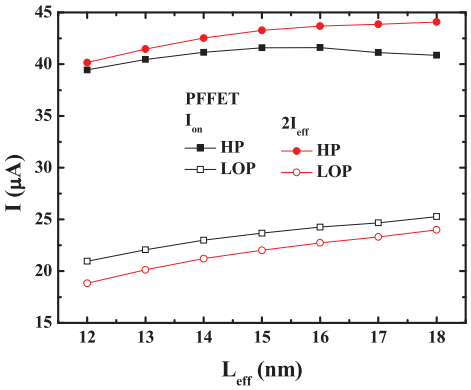

FIGURE 6. Simulated I_{on} and I_{eff} of nMOS PFFETs versus L_{eff} for the HP and LOP applications. (Note that we plot here and in Fig. 7 2xIeff, instead of I_{eff}, to better illustrate the comparisons with I_{on}).

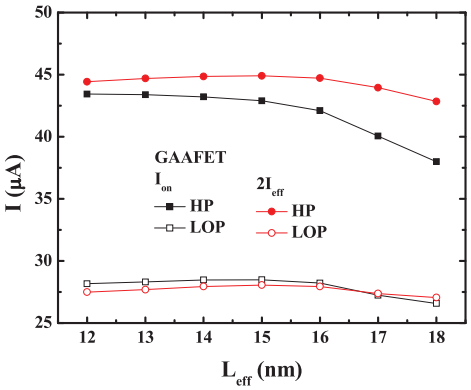

FIGURE 7. Simulated I_{on} and I_{eff} of nMOS GAAFETs versus L_{eff} for the HP and LOP applications.

GAAFET. This insight regarding the PFFET with G-S/D underlap further confronts well the fact that the GAAFET, with better SCE control, tends to perform relatively better for shorter L_g and lower V_{DD} [11].

Perhaps a better factor for the intrinsic delay (speed) of these two advanced CMOS devices operating at low V_{DD} is the "effective drive current" I_{eff} [21], in lieu of I_{on}. We show in Figs. 6 and 7 predicted comparisons of I_{eff} and I_{on} versus L_{eff} for the HP and LOP PFFETs and GAAFETs in Figs. 4 and 5. Both currents show similar trends for all cases, still suggesting longer L_{eff} for the PFFET for more performance competitiveness.

V. CONCLUSION

This work has demonstrated comparable integration densities and speed-power performances for the pragmatic Si-based FinFET and GAAFET technologies at the 5nm CMOS node. Optimizing the lateral S/D-spacer doping straggle and the G-S/D underlap it defines was shown to benefit the PFFET in this comparison. The severe processing complexity and cost that can be inferred for the GAAFET relative to the PFFET, coupled with our comparable performance and density results, suggest that the PFFET is the better choice for future nanoscale CMOS, especially if monolithic 3-D integration is exploited, as expected [1], to extend Moore's law after 2021. Furthermore, we note that the DG PFFET can, still with relatively easy processing, be designed with independent gates and bias offset to allow V_t adjust [7] for SOC applications, as is being done with the FD-SOI MOSFET [22].

REFERENCES

[1] *ITRS 2.0 Publication: 5_2015 ITRS 2.0_More Moore.pdf*. p. 14. Accessed on Aug. 3, 2016 [Online]. Available: http://www.itrs2.net/

[2] C. Pan *et al.*, "Technology/system codesign and benchmarking for lateral and vertical GAA nanowire FETs at 5-nm technology node," *IEEE Trans. Electron Devices*, vol. 62, no. 10, pp. 3125–3132, Oct. 2015, doi: 10.1109/TED.2015.2461457.

[3] R. Kim, U. E. Avci, and I. A. Young, "CMOS performance benchmarking of Si, InAs, GaAs, and Ge nanowire n- and pMOSFETs with L_G=13 nm based on atomistic quantum transport simulation including strain effects," in *IEDM Tech. Dig.*, Washington, DC, USA, Dec. 2015, pp. 34.1.1–34.1.4, doi: 10.1109/IEDM.2015.7409824.

[4] R. Stevenson, "Rise of the nanowire transistor," *IEEE Spectr.*, vol. 53, no. 2, pp. 9–11, Feb. 2016, doi: 10.1109/MSPEC.2016.7419783.

[5] H. Mertens *et al.*, "Vertically stacked gate-all-around Si nanowire CMOS transistors with dual work function metal gates," in *IEDM Tech. Dig.*, San Francisco, CA, USA, Dec. 2016, pp. 19.7.1–19.7.4, doi: 10.1109/IEDM.2016.7838456.

[6] S. Barraud *et al.*, "Vertically stacked-nanowires MOSFETs in a replacement metal gate process with inner spacer and SiGe source/drain," in *IEDM Tech. Dig.*, San Francisco, CA, USA, Dec. 2016, pp. 17.6.1–17.6.4, doi: 10.1109/IEDM.2016.7838441.

[7] J. G. Fossum and V. P. Trivedi, *Fundamentals of Ultra-Thin-Body MOSFETs and FinFETs*. New York, NY, USA: Cambridge Univ. Press, 2013.

[8] J. G. Fossum, Z. Zhou, L. Mathew, and B.-Y. Nguyen, "SOI versus bulk-silicon nanoscale FinFETs," *Solid-State Electron.*, vol. 54, no. 2, pp. 86–89, Feb. 2010, doi: 10.1016/j.sse.2009.12.002.

[9] *Sentaurus Device User Manual Ver. I-2013.12*, Synopsys, Mountain View, CA, USA, Dec. 2013.

[10] M. Karner *et al.*, "Vertically stacked nanowire MOSFETS for sub-10nm nodes: Advanced topography, device, variability, and reliability simulations," in *IEDM Tech. Dig.*, San Francisco, CA, USA, Dec. 2016, pp. 30.7.1–30.7.4, doi: 10.1109/IEDM.2016.7838516.

[11] M. Rau *et al.*, "Performance projection of III-V ultra-thin-body, FinFET, and nanowire MOSFETs for two next-generation technology nodes," in *IEDM Tech. Dig.*, San Francisco, CA, USA, Dec. 2016, pp. 30.6.1–30.6.4, doi: 10.1109/IEDM.2016.7838515.

[12] A. Veloso *et al.*, "Gate-all-around nanowire FETs vs. triple-gate FinFETs: on gate integrity and device characteristics," *ECS Trans.*, vol. 72, no. 2, pp. 85–95, 2016, doi: 10.1149/07202.0085ecst.

[13] V. Trivedi, J. G. Fossum, and M. M. Chowdhury, "Nanoscale FinFETs with gate-source/drain underlap," *IEEE Trans. Electron Devices*, vol. 52, no. 1, pp. 56–62, Jan. 2005, doi: 10.1109/TED.2004.841333.

[14] L.-T. Pang, K. Qian, C. J. Spanos, and B. Nikolic, "Measurement and analysis of variability in 45 nm strained-Si CMOS technology," *IEEE J. Solid-State Circuits*, vol. 44, no. 8, pp. 2233–2243, Aug. 2009, doi: 10.1109/JSSC.2009.2022217.

[15] I. Ferain, C. A. Colinge, and J.-P. Colinge, "Multigate transistors as the future of classical metal–oxide–semiconductor field-effect transistors," *Nature*, vol. 479, pp. 310–313, Nov. 2011, doi: 10.1038/nature10676.

[16] V. P. Trivedi and J. G. Fossum, "Quantum-mechanical effects on the threshold voltage of undoped double-gate MOSFETs," *IEEE Electron Device Lett.*, vol. 26, no. 8, pp. 579–582, Aug. 2005, doi: 10.1109/LED.2005.852741.

[17] D. J. Frank, S. E. Laux, and M. V. Fischetti, "Monte Carlo simulation of a 30 nm dual-gate MOSFET: How short can Si go?" in *IEDM Tech. Dig.*, Dec. 1992, pp. 553–556, doi: 10.1109/IEDM.1992.307422.

[18] S. Natarajan *et al.*, "A 14nm logic technology featuring 2nd-generation FinFET, air-gapped interconnects, self-aligned double patterning and a 0.0588 μm^2 SRAM cell size," in *IEDM Tech. Dig.*, San Francisco, CA, USA, Dec. 2014, pp. 3.7.1–3.7.3, doi: 10.1109/IEDM.2014.7046976.

[19] T. Huynh-Bao *et al.*, "A comprehensive benchmark and optimization of 5-nm lateral and vertical GAA 6T-SRAMs," *IEEE Trans. Electron Devices*, vol. 63, no. 2, pp. 643–651, Feb. 2016, doi: 10.1109/TED.2015.2504729.
[20] J. Lacord *et al.*, "Parasitic capacitance analytical model for sub-7-nm multigate devices," *IEEE Trans. Electron Devices*, vol. 63, no. 2, pp. 781–786, Feb. 2016, doi: 10.1109/TED.2015.2506781.
[21] M. H. Na, E. J. Nowak, W. Haensch, and J. Cai, "The effective drive current in CMOS inverters," in *IEDM Tech. Dig.*, San Francisco, CA, USA, Dec. 2002, pp. 121–124, doi: 10.1109/IEDM.2002.1175793.
[22] Q. Liu *et al.*, "FDSOI CMOS devices featuring dual strained channel and thin BOX extendable to the 10nm node," in *IEDM Tech. Dig.*, San Francisco, CA, USA, Dec. 2014, pp. 9.1.1–9.1.4, doi: 10.1109/IEDM.2014.7047014.

YA-CHI HUANG received the B.S. and M.S. degrees in electronics engineering from National Ilan University, I-Lan, Taiwan, in 2014 and 2015, respectively. She is currently pursuing the Ph.D. degree with the Department of Electrical Engineering, Institute of Microelectronics, National Cheng Kung University, Tainan, Taiwan.

MENG-HSUEH CHIANG (S'97–M'01–SM'07) received the B.S. degree in electrical engineering from National Cheng Kung University, Tainan, Taiwan, in 1992, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Florida, Gainesville, FL, USA, in 1995 and 2001, respectively.

He is currently a Professor with the Department of Electrical Engineering, National Cheng Kung University.

SHUI-JINN WANG received the B.S., M.S., and Ph.D. degrees in electrical engineering from National Cheng Kung University, Tainan, Taiwan in 1978, 1980, and 1985, respectively.

He is currently a Distinguished Professor with the Department of Electrical Engineering, National Cheng Kung University.

JERRY G. FOSSUM (S'69–M'71–SM'79–F'83–LF'09) was born in Phoenix, AZ, USA. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Arizona, Tucson. During his graduate program he was a NASA Predoctoral Trainee.

In 1971, he joined the Sandia Laboratories, Albuquerque, NM, USA, as a Technical Staff, where he was engaged in various semiconductor device design and modeling activities, including the development of silicon solar cells. In 1978, he moved to the University of Florida, Gainesville, where he is currently a Distinguished Professor Emeritus of Electrical and Computer Engineering. He has authored or co-authored approximately 300 papers published in technical journals and conference proceedings, and a book entitled *Fundamentals of Ultra-Thin-Body MOSFETs and FinFETs* published in 2013. He has directed the research of 35 Ph.D. students. His general area of expertise continues to be semiconductor device theory, modeling, and design, with applications to SOI/multigate MOSFETs and solar cells.

Dr. Fossum was a recipient of the Best Paper Award at the IEEE International SOI Conference in 1992 and the IEEE/EDS J. J. Ebers Award for "Outstanding contributions to the advancement of SOI CMOS devices and circuits through modeling" in 2004.