Group IV Nanotube Transistors for Next Generation Ubiquitous Computing

Hossain M. Fahad[*], Aftab M. Hussain[*], Galo A. Sevilla Torres[*], Sanjay K. Banerjee[‡], and Muhammad M. Hussain[*]
[*]Integrated Nanotechnology Lab, Electrical Engineering, CEMSE Division,
King Abdullah University of Science and Technology, Thuwal, 23955-6900, Saudi Arabia.
[‡]Microelectronics Research Center, University of Texas at Austin, J. J. Pickle Research Campus,
10100, Burnet Rd, Austin, Texas, 78758, USA.

ABSTRACT

Evolution in transistor technology from increasingly large power consuming single gate planar devices to energy efficient multiple gate non-planar ultra-narrow (< 20 nm) fins has enhanced the scaling trend to facilitate doubling performance. However, this performance gain happens at the expense of arraying multiple devices (fins) per operation bit, due to their ultra-narrow dimensions (width) originated limited number of charges to induce appreciable amount of drive current. Additionally arraying degrades device off-state leakage and increases short channel characteristics, resulting in reduced chip level energy-efficiency. In this paper, a novel nanotube device (NTFET) topology based on conventional group IV (Si, SiGe) channel materials is discussed. This device utilizes a core/shell dual gate strategy to capitalize on the volume-inversion properties of an ultra-thin (< 10 nm) group IV nanotube channel to minimize leakage and short channel effects while maximizing performance in an area-efficient manner. It is also shown that the NTFET is capable of providing a higher output drive performance per unit chip area than an array of gate-all-around nanowires, while maintaining the leakage and short channel characteristics similar to that of a single gate-all-around nanowire, the latter being the most superior in terms of electrostatic gate control. In the age of big data and the multitude of devices contributing to the *internet of things*, the NTFET offers a new transistor topology alternative with maximum benefits from performance-energy efficiency-functionality perspective.

Keywords: silicon, germanium, nanotube, tunnel, epitaxy

1. INTRODUCTION

The demand for ultra-fast and multi-functional computing technology in an ultra-mobile form factor is continually driving the silicon transistor industry to smaller and smaller technology nodes. In the past, physics-induced limitations (off-state leakage heat-dissipation, degraded channel control – DIBL, SS) on single gate planar architecture transistor made it difficult to continue scaling beyond the 32 nm technology node, where the critical dimension was the transistor gate length (L_g). With the introduction of non-planar FinFET architecture in 2011 [1], L_g – scaling resumed the aggressive scaling of doubling performance every eighteen months while maintaining low power consumption and improved transistor characteristics at ultra-short channels (14 nm and below), thereby pro-longing the life of silicon and other group IV transistor materials. FinFET scaling up-to the 5 nm technology node may be viable [2], however, economics will be the deciding factor in determining the direction of the transistor industry and at this point there are two, assuming conventional and inexpensive channel materials are used in the subsequent technology nodes:

1) Continue transistor scaling on a new type of non-planar transistor architecture (gate-all-around nanowires or 'hanging fins').
2) Stop transistor scaling at or near the physical limit of the conventional material system and build upwards i.e. 3D integrated circuits (IC).

In current generation ICs, the overall computation latency is being dominated by wire propagation delay rather than transistor delay, as was the case in the past. This may be addressed by either boosting the performance at the individual transistor level (option 1) or by reducing the interconnect wire length by integration of all necessary compute blocks on a single 3D package (option 2). 3D ICs will require major overhaul in fabrication technologies such as die-die alignment,

Micro- and Nanotechnology Sensors, Systems, and Applications VI, edited by Thomas George, M. Saif Islam, Achyut K. Dutta, Proc. of SPIE Vol. 9083, 90830K · © 2014 SPIE
CCC code: 0277-786X/14/$18 · doi: 10.1117/12.2050096

new forms of in-situ chip heat management, etc. Since the industry has already transitioned to non-planar transistor architecture, exploring new topologies to boost on-state performance may be a viable option. With the above in mind, this paper describes the concept of a unique nanotube based transistor topology using conventional channel materials. Through extensive device simulations and analytical calculations, we show that such a transistor architecture merges the unique low-leakage, strong quantum-confinement effects with on-state drive performance capability per unit of device area. Towards the end of the paper, a highly integrated and industry-compatible process for fabricating 1D nanotubes is also discussed.

2. SILICON NANOTUBE FIELD EFFECT TRANSISTOR

2.1 Concept

The silicon nanotube transistor is a vertically oriented 3D device topology consisting of an ultra-narrow dimension hollow silicon body that is gated by inner/outer core/shell gate stacks as depicted in Figure 1. This type of gate strategy induces strong electrostatic charge control in the ultra-thin (< 10 nm) silicon body leading to volume-inverted mode of charge transport [3, 4]. Here, the channel charge centroid becomes surface delocalized and shifts into the volume of the silicon nanotube reducing charge scattering due to relative distance between the edge (interface) of the channel surface. Volume inversion enables strong charge-carrier energy quantization for high transistor on-state drive current coupled with a strong electrostatics keeps off-state leakage and performance degrading short channel effects such as the DIBL and SS at a minimum. Such volume-localized charge transport is inherent in state-of-the-art gate-all-around nanowires with diameters well below 10 nm. In order to leverage high drivability from gate-all-around nanowire technology, multiple device arraying is necessary to function as a single transistor. This arraying process is already in practice with current generation FinFETs. The problem here is that, although a single nanowire/fin may exhibit ultra-low off-state leakage and excellent short channel behavior, arraying 'adds' up both the off-state and on-state currents from the individual nanowires/fins. So although a higher transistor on-state current is achieved, the off-state leakage also goes up in proportion. In addition to this, arraying also tends to increase the silicon chip-area as a trade-off for on-state drive performance gain. This may be controlled up-to a certain extent by controlling the nanowire/fin pitch. However, this imposes both fabrication as well as physics-induced constraints that are responsible for higher parasitic capacitances and resistances that are detrimental to overall chip latency.

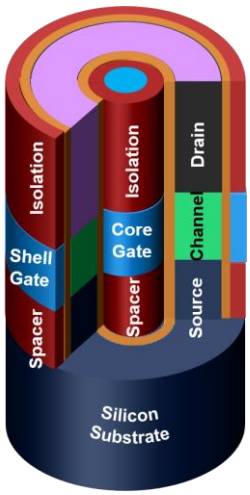


Figure 1. 3D representation of the silicon nanotube transistor concept and the volume inversion capability in ultra-thin nanotubes.

2.2 Silicon Nanotube Transistors for Conventional CMOS

Using 3D device modeling and semi-classical charge transport models with quantization effects, device performance of a single nanotubes (NTFET) are compared to gate-all-around nanowire (GAA NWFET) transistors. Figure 2 compares the I_d-V_g transfer characteristics for n and p channel devices. In this comparison with silicon as the channel material, the nanotube has a thickness of 10 nm and the nanowire diameter is fixed at 20 nm. The reason for the wide nanowire choice

is to make the device performance comparison at matched off-state leakage current (I_{off}) ~ 10 nA. The nanotube has an inner core gate diameter of 100 nm. A key observation to make from Figure 2 is that unlike GAA NWFETs, the performance of silicon NTFETs can be scaled up by varying the inner core-gate diameter (Figure 1) without increasing off-state leakage or short channel effects. Controlling the NTFET inner core-gate dimensions comes down to the contact technology scalability. Nonetheless, the silicon nanotube transistor offers a competitive edge over gate-all-around nanowires as well as a new insight in 1D transistor.

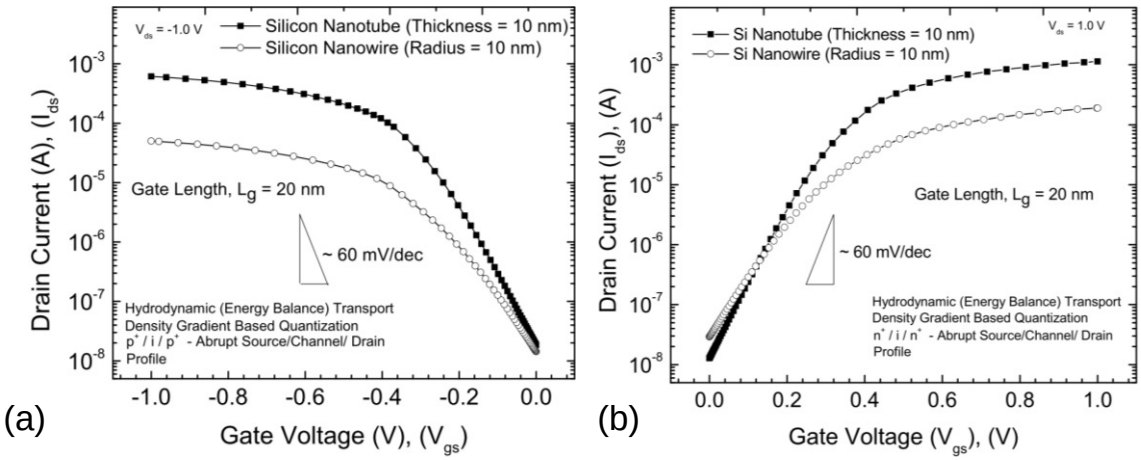


Figure 2. Performance comparison between a single nanotube and a single gate all-around nanowire for (a) p-channel and (b) n-channel transistor. The dimensions of the nanowire are chosen to make a fair comparison at matched off-state leakage current (I_{off}).

2.3 Silicon Nanotube Tunnel Transistors

Ultra-low power mobile consumer applications emphasizes on long standby battery-lifetime and form-factor. However, given the increasing trend of integrating a multitude of functionalities in portable consumer electronics, the next step is to have high performance computing capability. The introduction of FinFETs since 2011 has seen a dramatic improvement in battery lifetime, where portable electronics can go a lot farther on a single charge. Although, computing performance has been increasing steadily it has not seen a dramatic change as much as the off-state leakage control. However, the emphasis on long battery-lifetime is of paramount importance and various materials and device technologies are being investigated to provide even longer standby-time in next-generation portable electronics.

Current generation transistors utilize classical over-the-barrier charge transport physics. Due to a fundamental limit, the transistor power supply (V_{dd}) needed to drive the transistors cannot be scaled down in proportion to the transistor gate length. This is because in such 'classical' devices, charges are injected from the source to the drain terminal by lowering a potential barrier which is classically limited to kT/q or a 60 mV/ dec transistor sub-threshold swing (Figure 3(a,c)). Figure 3(d) shows a roadmap for low power systems requirements for the next decade with a 5% forward projected increase in power supply capacity from batteries for portable electronics. As it can be seen, classical over-the-barrier transistors using silicon channel even with multiple-gate architectures will not work due to their inherent fundamental limitation.

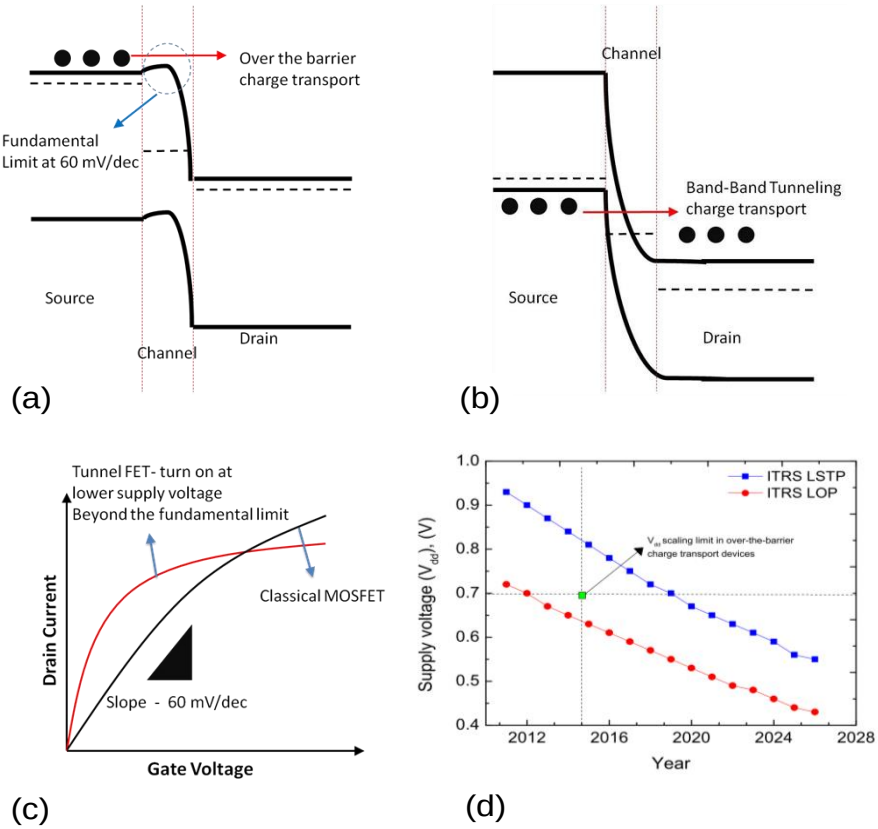


Figure 3. (a) Classical over-the-barrier charge transport devices. (b) Band-band charge tunneling transistors. (c) Advantage of tunnel FETs over classical FETs – steeper sub-threshold slope means transistors can turn on with smaller voltages. (d) ITRS projected power supply requirements for future ultra-low power mobile computing platforms.

In order to meet the power supply targets of Figure 3(d), tunnel transistors are likely possible transistor candidates in the ultra-low power computing regime. A tunnel FET is a gated p-i-n diode that utilizes the large band-edge overlap of a highly doped source and an intrinsic channel enabling charge carriers to tunnel through the forbidden gap from the source to the channel which are then collected by the drain (Figure 3(b)). This is a quantum-mechanical phenomenon that is a function of charge effective mass (m^*), material band-gap (E_g) and band-edge overlap ($\Delta\Phi$). Although tunnel transistors are capable of ultra-low power operation, they are incapable of high on-state drive performance. Most of the reported demonstrations report sub-60 mV/dec turn-on characteristics albeit many of them are point-slope values. The best reported drive currents from tunnel FETs are still in µA/µm range as opposed to the current generation transistors with drive performance in the mA/µm range. Several innovative solution have been proposed to increase the output performance such as the use of strain and hetero-junctions [5-7]. Some of the more promising demonstrations include a gate-all-around topology around a III-V nanowire and the use of a high-K/metal gate stack [8, 9]. In a recent publication, we have shown that a single silicon nanotube can out-perform arrays of nanowires for tunnel transistor application [10]. The key point here was that at a matched off-state leakage current, the performance of a single nanotube is scaled up by a factor of 10x due to the inner core gate compared to a nanowire while showing nearly similar turn-on behavior. What the nanotube architecture is effectively doing is physically enhancing the drivable on-state tunnel current in an area efficient manner. This becomes visibly impactful when considering ultra large scale device integration and density on a chip-level for high performance computing in mobile applications.

2.4 Junction-less silicon nanotube transistors

In recent years, the concept of a uniformly doped ultra-thin-body multiple gate transistor gained popularity [11, 12]. The reason being that a junction-less device configuration with similarly doped source, channel and drain regions becomes immune to performance variability that can occur due to random unwanted dopants in the channel especially at ultra-

short transistor channel lengths. These devices operate like a gated resistor with the accumulation of majority carriers in the channel unlike traditional inversion mode transistors. The main drawback of these devices is the use of expensive high work-function or incompatible gate materials, as such accumulation mode devices are normally-ON. Nonetheless the unique architecture of the nanotube may prove advantageous for such an application compared to GAA nanowires.

2.5 $Si_{1-x}Ge_x$ alternate channel nanotube transistors

Alternate channel materials using group IV materials such as $Si_{1-x}Ge_x$, $Si_{1-x}Sn_x$, $Ge_{1-x}Sn_x$, Ge offer the unique advantage of high charge carrier mobility and injection velocity on account of their low effective mass (m^*) compared to plain silicon [13-15]. Traditionally, p-channel transistors have shown typically low drive currents compared to their n-channel counterparts due to the fact that holes are effectively heavier than electrons. However, alloying silicon with germanium reduces the in-direct silicon band-gap by the straining the hole valence band, leading to lowering of the hole effective carrier mass (m^*_h). When these material benefits are combined with the nanotube architecture, ultra-high performance PMOS transistors can be created that can out-perform GAA NWFET arrays using the same channel material.

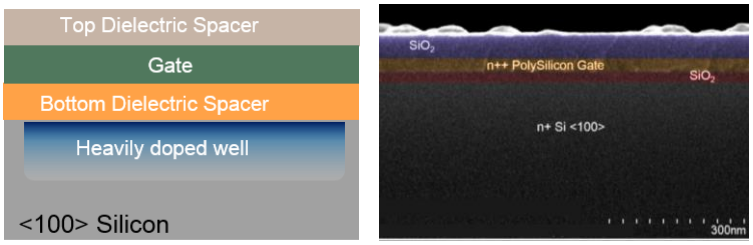
3. INTEGRATION PROCESS FLOW

Figure 4 details key process steps of a unique and integrated flow to fabricate vertical highly aligned 1D silicon nanotube transistors. This is a highly compatible CMOS process that offers several key potential advantages [16]:

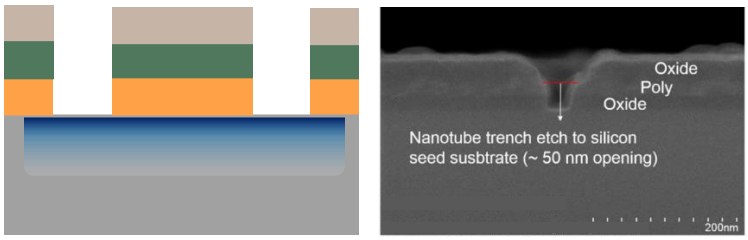
(i) The device channel length (L_g) is defined by the thickness of the deposited material enabling the formation of ultra-short channel devices. Such a process is immune to bottlenecks due to lithographic constraints and the shortest gate length is defined by the minimum thickness of the deposited material. Atomic layer deposited gate materials can be theoretically used to achieve sub-nm gate length devices.

(ii) The nanotube thickness can be controlled by the deposited gate dielectric thickness allowing ultra-narrow width nanotube FETs. Initial nanotube trenches can be defined using conventional lithography. Using high-K gate dielectric enables the use of thick materials while maintaining a low effective oxide thickness.

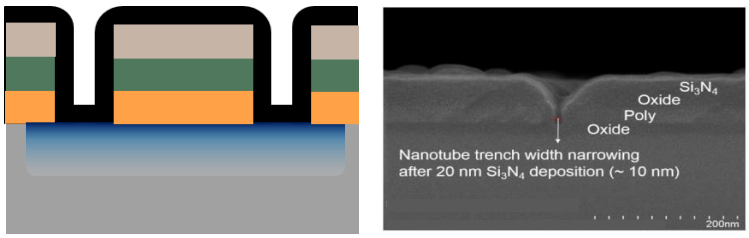
(iii) Epitaxial growth allows the formation of highly abrupt source/channel and drain/channel junctions. This requires a low temperature epitaxial growth process to reduce unwanted dopant diffusion across the junctions. In addition, a good interface with the gate dielectric and the epitaxial material is important in ensuring low interface defect density (D_{it}).



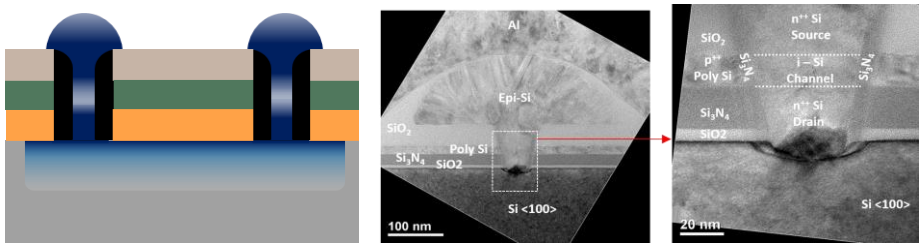
(a) Deposition of gate stack.



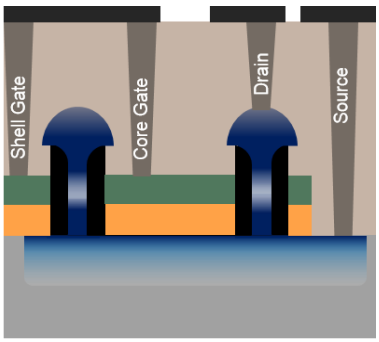
(b) Nanotube trench etch



(c) Gate dielectric deposition



(d) Selective epitaxial silicon nanotube growth with in-situ source/drain doping



(e) Contact Formation

Figure 4. Integrated process flow to fabricate vertical highly aligned 1D silicon nanotube transistors.

4. CONCLUSION

In this paper, we have demonstrated the benefits of a having a nanotube architecture for field effect transistor applications on a variety of platforms such as classical CMOS, tunnel FETs, junction-less and alternate channel transistors. Compared to popular state-of-the-art gate-all-around nanowires, the nanotubes provide a promising alternative in bringing high performance computing capability to next generation ultra-mobile consumer applications using conventional group IV channel materials.

REFERENCES

[1] Bohr, M., "The evolution of scaling from homogeneous era to the heterogeneous era," Int. Elect. Dev. Meet. 1.1, 1-6 (2011).

[2] Lee, H., Yu, L., Ryu, S.-W., Han, J.-W., Jeon, K., Jang, D.-Y., Kim, K.-H., Lee, J., Kim, J.-H., Jeon, S.-C., Lee, G.-S., Oh, J.-S., Park, Y.-C., Bae, W.-H., Lee, H.-M., Yang, J.-M., Yoo, J.-J., Kim, S.-I , and Choi, Y.-K., "Sub-5nm all-around gate FinFET for ultimate scaling," Proc. VLSI Symp. Tech. 58-59 (2006).

[3] Fahad, H. M., Smith, C. E., Rojas, J. P., and Hussain, M. M., "Silicon nanotube field effect transistor with core-shell gate stacks for enhanced high-performance operation and area scaling benefits," Nano Lett. 11(10), 4393-4399 (2011).

[4] Fahad, H. M., and Hussain, M. M., "Are nanotube architectures more advantageous than nanowire architectures for field effect transistors?" Sci. Rep. 2, 475 (2012).

[5] Boucart, K., Ionescu, A. M., and Riess, W., "Asymmetrically strained all-silicon tunnel FETs featuring 1V operation, " Proc. Eur. Sol. Stat. Dev. Res. Conf. 452-456 (2009).

[6] Ionescu, A. M., and Riel, H., "Tunnel field-effect transistors as energy-efficient electronic switches," Nature 479 (7373), 329-337 (2011).

[7] Tomioka, K., Yoshimura, M., and Fukui, T., "Steep-Slope tunnel field-effect transistors using III-V nanowire/Si heterojunction," Proc. VLSI Symp. Tech. 47-48 (2012).

[8] Tomioka, K., Yoshimura, M., and Fukui, T., "Sub 60 mV/decade switch using an InAs nanowire-Si heterojunction and turn-on voltage shift with a pulsed doping technique," Nano Lett. 13(12), 5822-5826 (2013).

[9] Borg, B. M., Dick, K. A., Ganjipour, B., Pistol, M.-E., Wernersson, L.-E., and Thelander, C., "InAs/GaSb heterostructure nanowires for tunnel field-effect transistors," Nano Lett. 10(10), 4080-4085 (2010).

[10] Fahad, H. M., and Hussain, M. M., "High-performance silicon nanotube tunneling FET for ultralow-power logic applications," IEEE Trans. Elect. Dev. 60(3), 1034-1039 (2013).

[11] Colinge, J.-P., Lee, C.-W., Afzalian, A., Akhavan, N. D., Yan, R., Ferain, I., Razavi, P., O'Neill, B., Blake, A., White, M., Kelleher, A.-M., McCarthy, B., and Murphy, R., "Nanowire transistors without junctions," Nature Nanotech. 5, 225-229 (2010).

[12] Colinge, J.-P., Lee, C.-W., Ferain, I., Akhavan, N. D., Yan, R., Razavi, P., Yu, R., Nazarov, A. N., and Doria, R. T., "Reduced electric field in junctionless transistors," Appl. Phys. Lett. 96, 073510-1-3 (2010).

[13] Fang, W. W., Singh, N., Bera, L. K., Nguyen, H. S., Rustagi, S. C., Lo, G. Q., Balasubramanium, N., and Kwong, D.-L., "Vertically stacked SiGe nanowire array channel CMOS transistors, " IEEE Elect. Dev. Lett. 28(3), 211-213 (2007).

[14] Hussain, A. M., Fahad, H. M., Singh, N., Rader, K. R., Schwingenschlogl, U., and Hussain, M. M., "Exploring SiSn as channel material for LSTP device applications," Dev. Res. Conf. 93-94 (2013).

[15] Gupta, S., Chen, R., Magyari-Kope, B., Lin, H., Yang, B., Nainani, A., Nishi, Y., Harris, J. S., and Saraswat, K. C., "GeSn technology: Extending the Ge electronics roadmap, " Int. Elect. Dev. Meet. 16.6.1-16.6.4 (2011).

[16] Rojas, J. P., and Hussain, M. M., "Sub-15 nm nano-pattern generation by spacer width control for high density precisely positioned self-assembled device nanomanufacturing," IEEE Conf. Nanotech. 20-23 (2012).