# In Quest of the Next Switch

Suman Datta
Department of Electrical Engineering
University of Notre Dame
Notre Dame, IN, USA
sdatta@nd.edu

Nikhil Shukla
Department of Electrical Engineering
University of Notre Dame
Notre Dame, IN, USA
nshukla@nd.edu

***Abstract*—Conventional CMOS scaling and the Moore's law have been the cornerstone of progress in computing hardware technology. However, with dimensional scaling expected to end soon, there is a pressing need to find the next transistor solution that can continue to support the technology revolution. Will this hardware solution be an enhanced or an augmented version of MOSFET or a switch based on a radically new switching mechanism. Ultimately, do we require a complete deviation from the Boolean paradigm itself? In this invited paper, we will review some of the actively pursued state-of-the-art transistor and related concepts. While it remains unclear which of these options will eventually make it into commercial products, we will argue- based on lessons learnt from the past two decades of transistor development- that sustained and systematic research with careful benchmarking remains the key to success in the quest for the new switch.**

***Keywords—FinFET; Quantum Well MOSET; Tunnel FET; Phase-FET; NC FET; Oscillators; $VO_2$***

## I. Introduction

Technology scaling has led to an unprecedented level of integration with billions of high-speed nano-transistors on a single chip reducing the cost per function. On the device technology front, device engineers have achieved through continued scaling, new transistor breakthroughs, and introduced innovations at a rapid pace followed by successful launch of commercially successful products such as high performance microprocessors that presently power over a billion mobile smartphones, personal computers, and massive data centers.

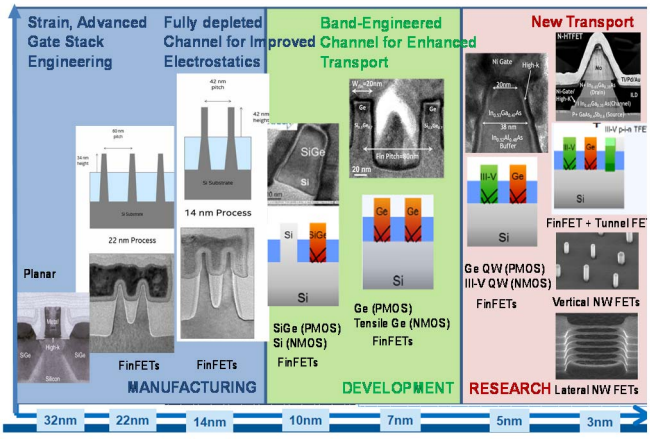


**Fig. 1:** Logic transistor innovation and its evolutionary roadmap.

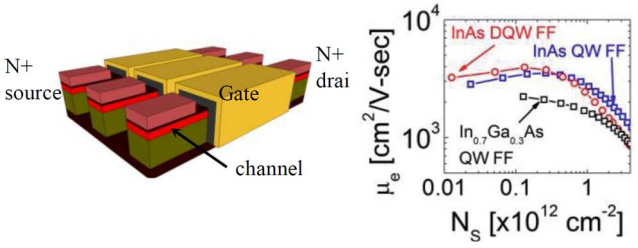


**Fig. 2:** Quantum-well FinFET and its experimental room

We have achieved this by overcoming seemingly fundamental barriers to transistor scaling imposed by natural laws of physics allowing us to sustain an aggressive scaling roadmap in a persistent effort towards enhancing the energy efficiency of the transistors. We will outline this incredible journey that spans over last sixteen years of logic transistor research and development. We will argue that serendipity has always played a key role in shaping this remarkable path of transistor innovation. Recent innovations in silicon complementary metal oxide semiconductor (CMOS) transistors that have achieved successful commercial implementations include strained channel transistors, high-k metal gate transistors and non-planar Tri-gate transistors as illustrated in Fig. 1 [1-3].

The road ahead in our quest for the next transistor / switch is fraught with complex challenges. New transistor architectures such as gate-all-around vertical and horizontal nanowires, new band engineered channels such as quantum-well FinFETs are transistors. Some of these designs and architectures are even venturing to utilize the concept of stabilizing the "negative capacitance" in a ferroelectric dielectric integrated within the gate stack to 'boost' the performance of a conventional transistor. We will summarize the current state-of-the-art of these new transistors in the next sections. At this stage, it's not clear which of these options will find their way into commercial products. As history has proven before, the research will find its way into technologies in unexpected ways, provided we remain vigilant and committed.

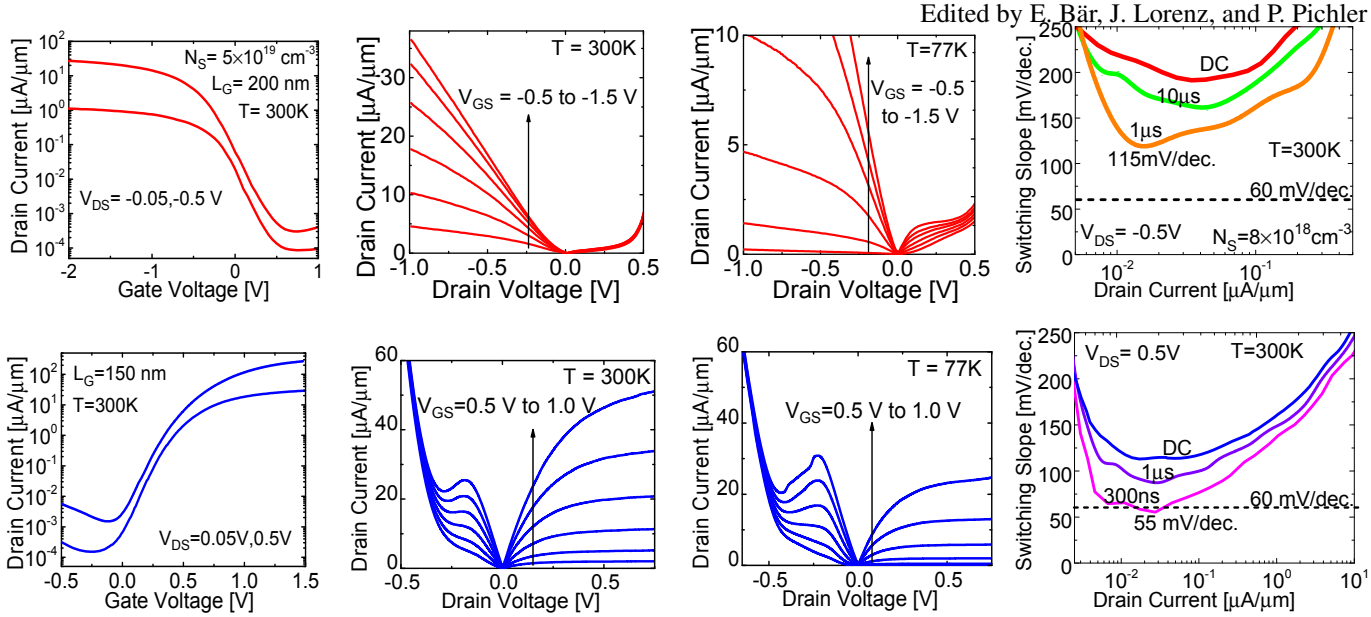


**Fig. 3:** Transfer, output and switching characteristics of (a-d) PTFET and (e-h) NTFET. All measurements are at T=300K, except the additional T=77K data in (c) and (g). NDR is visible in PTFET output characteristics at T=77K, due to the suppression of trap response. Improvement in switching performance from pulsed mode measurements is evident from (d, h).

## II. BOLTZMANN TRANSISTORS

### *A. Quantum-Well (QW) FinFETs*

Power constrained scaling of CMOS technology places several stringent requirements on the performance metrics of the underlying device, such as drive-current, leakage current, on-to-off state current ratio, subthreshold slope, Drain Induced Barrier Lowering (DIBL), all which have to be satisfied simultaneously. The key requirement however, is to reduce dynamic power dissipation which can be achieved by scaling down the supply voltage ($V_{DD}$) because power dissipation is proportional to the square of $V_{DD}$. At reduced $V_{DD}$, however, traditional silicon CMOS is limited by reduced drive currents. In this context, III−V compound semiconductor material systems are being actively researched because they afford competitive drive currents at lower supply voltages. The advantage primarily stems from the reduced effective mass in III−V materials which results in higher mobility. Recently, InGaAs and InAs heterostructure FinFETs have been demonstrated as an n-channel alternative to Si FinFET [4−6]. InAs single and dual Quantum Well (QW) FinFETs exhibit peak mobility >3,000 $cm^2$/V-sec. Projected short-channel InAs QW FinFETs are expected to show 15% higher drive current with comparable short channel effects as Silicon FinFETs, making them a potential candidate for future transistors.

### *B. Inter-Band Tunnel FETs (TFETs)*

In the future, even high mobility transistors will face supply voltage $V_{DD}$ scaling challenges as we near 500mV supply due to the fundamental 60mV/decade sub-threshold swing limitation in MOSFETs arising from the fundamental Boltzmann limit. Consequently, any further threshold voltage scaling leads to severe increase in static power dissipation. TFETs with gated band-to-band tunneling at the source-channel junction [7] potentially leads to filtering of the high energy tail of the Femi-Dirac distribution of electrons in the source of TFET. This band-pass filtering action at the source-channel junction leads to an effective "cooling" of the carrier population and results in sub-*kT/q* or sub-60 mV/dec of sub-threshold swing at room temperature. Thus, TFETs enable aggressive $V_{DD}$ scaling without degrading the ON-state performance of the transistors. Figure 3 shows the experimental transfer ($I_{DS}$-$V_{GS}$) and output ($I_{DS}$-$V_{DS}$) characteristics for the fabricated PTFET and NTFET. $GaAs_{0.35}Sb_{0.65}$ channel PTFETs exhibit $I_{ON}$ =30μA/μm at $I_{ON}/I_{OFF}$ = $10^5$. The PTFET output characteristics exhibit negative differential resistance (NDR) and saturation at low temperature as contribution from mid-gap $D_{it}$ is suppressed. $In_{0.65}Ga_{0.35}As$ channel NTFET shows $I_{ON}$ =275μA/μm at $I_{ON}/I_{OFF}$=3×$10^5$. The switching slope (SS) in the fabricated TFETs surpasses the Boltzmann limit of 60 mV/decade at room temperature. This dilution of SS is a consequence of mid-gap $D_{it}$ with slow response time. To suppress the response of mid-gap $D_{it}$, we perform pulsed $I_{DS}$-$V_{GS}$ measurement on TFETs with the input gate-voltage $V_{GS}$ pulse having rise time varying from 10μs down to 300 ns to assess actual switching performance. Fig. 3(d,h) plots the SS as a function of drain current for various gate pulse rise times. We achieve SS=55mV/decade for NTFET and SS=115mV/decade for PTFET at room temperature in pulsed mode measurements. Consequently, engineering high-quality scaled gate dielectrics and tunnel barriers in the As-Sb system is fundamental to the realization of high $I_{ON}$ with steep switching slope demonstration in TFETs.

## III. LANDAU TRANSISTORS

While TFETs harness an alternate transport mechanism within conventional semiconductors, a new paradigm for enabling *sub*-kT/q switching devices exploits of an entirely new class of materials with unique properties, enabled by novel materials physics, and unavailable in conventional semiconductors. Such phenomena include insulator-metal transitions, ferroelectricity, nano-electro-mechanical barriers; and, in this context, their corresponding transistor implementations can be broadly classified as Landau transistors [10] because of the similar double well energy landscape associated with phase transitions which is described by the phenomenological Landau theory [11]. Here, we will describe two Landau transistor concepts: The Phase-FET based

on the phenomenon of abrupt insulator-metal transition (IMT), and the Negative Capacitance (NC) FET that harnesses negative capacitance in a composite ferroelectric dielectric gate stack, both of which are being actively pursued as steep-slope transistor options.

### *A. Phase FETs*

Phase-FETs utilize the phenomenon of an abrupt insulator-metal transition (IMT) that arises from collective interactions in certain materials, and manifests itself as orders of magnitude

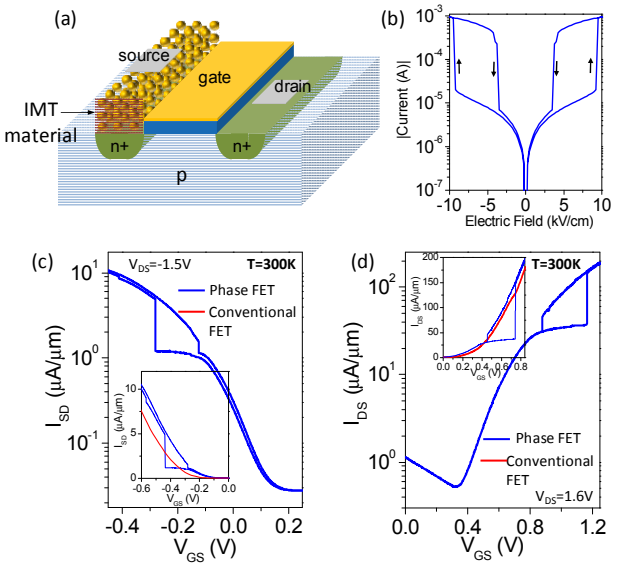


**Fig. 4:** (a) Schematic of the Phase-FET consisting of the IMT material integrated with the source of a conventional transistor. **(b)** Typical I-V characteristics of a two-terminal $VO_2$ device exhibiting abrupt switching associated with the IMT. **(c)(d)** Transfer characteristics $I_{DS}$-$V_{GS}$ of p-type and n-type Phase FET. The conventional transistors for the p-type and n-type devices are p-Ge QW FinFET and n-$In_{0.7}Ga_{0.3}As$ FinFET, respectively. Insets show the comparison of the Phase-FET with their respective conventional transistors at matched OFF current.

change in resistivity. The change in resistivity is induced due to a fundamental change in the band-structure, and subsequent modulation of the bandgap as well as amplification of the carrier concentration. For example, in the prototypical IMT material vanadium dioxide ($VO_2$), an electrical stimulus applied to a two-terminal $VO_2$ device collapses the 0.6eV bandgap in the insulating state, transforming $VO_2$ into a metal (no bandgap) accompanied by four orders of magnitude change in resistivity [12].

The concept of the phase-FET entails the integration of such IMT materials into the source of a conventional transistor (Fig. 4(a)) [13]. During operation, the gate-terminal voltage $V_{GS}$ modifies the current flowing through the series combination of the transistor channel and the IMT material, triggering the abrupt and reversible phase transition, and enabling the steep *sub*-kT/q switching characteristics. Further, the high-resistivity insulating state of the IMT material exponentially reduces the OFF-state leakage of the conventional FET while the orders of magnitude reduction in resistivity following transformation to the metallic state ensures negligible reduction in the ON state drive current of the FET enabling the Phase-FET to exhibit a higher ON-OFF ratio in comparison to the conventional FET; the higher ON-OFF ratio can be translated into higher ON current at matched OFF current of the MOSFET. Thus, such a transistor design seamlessly incorporates the abrupt switching phenomenon in IMT materials with the superior field-effect dynamics of conventional semiconductor based transistors to deliver improved performance.

Figure 4(c) shows the transfer ($I_{DS}$-$V_{GS}$) and output ($I_{DS}$-$V_{DS}$) characteristics of a Phase-FET consisting of a n-type multi-channel $In_{0.7}Ga_{0.3}As$ QW FinFET in series with $VO_2$. It can be observed from the inset of Fig. 4(c) that the Phase-FET delivers a 20% higher ON current in comparison to the conventional transistor at matched OFF state current. Further, generality of the Phase-FET design makes it applicable to a p-type solution as well. Fig. 4(d) shows the p-type Phase FET constructed using a Ge QW FinFET and $VO_2$, exhibiting steep-switching characteristics and ~60% improvement in ON current in comparison to the conventional transistor at matched OFF-state current [13]. The steep-slope and higher ON-OFF current ratio enables $V_{DD}$ scaling and subsequent reduction in dynamic power.

### *B. Negative Capacitance (NC) FETs*

While the Phase FET induces internal voltage amplification at the source side of a conventional MOSFET through IMT, the NC FET aims to exploit the concept of negative capacitance in a ferroelectric-dielectric gate stack to create an internal step-up transformer between the applied gate-voltage $V_{GS}$ and the surface potential $\Psi_s$. Negative capacitance, as predicted by the Landau mean-field theory, is an unstable region in the polarization (P) versus electric-field (E) characteristics of a ferroelectric where dP/dE<0, and thus can be considered as a negative capacitance (NC) region. Salahuddin *et al.* [14] proposed that such an unstable NC region could be stabilized using a dielectric capacitance in series with the ferroelectric. The series combination of a positive capacitance (dielectric and semiconductor band bending) and an effective negative capacitance (ferroelectric) in the gate stack would amplify the internal node voltage $\Psi_s$ allowing the NC FET to exhibit steep slope and increase the net gate capacitance ($C_{ox}$) enabling the NC FET to deliver higher ON current in comparison to the conventional FET.

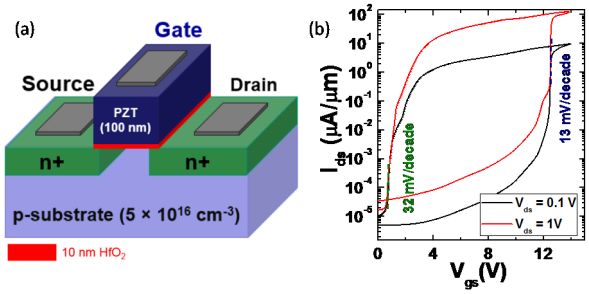


**Fig. 5**: (a) Schematic of the NC FET showing the ferroelectric (PZT) and dielectric ($HfO_2$) incorporated into the transistor gate stack. (b) Experimental $I_{DS}$-$V_{GS}$ characteristics of a PZT gated FET with channel length $L_{ch}$ =10 μm for $V_{DS}$= 0.1 V and $V_{DS}$= 1 V.

Figure 5(a)(b) shows the schematic and transfer characteristics of an NCFET demonstrated by monolithically integrating a polycrystalline $PbZr_{0.52}Ti_{0.48}O_3$ (PZT) ferroelectric with a conventional Silicon planar MOSFET (10nm $HfO_2$ gate dielectric). The transfer characteristics exhibit *sub*-kT/q switching (turn ON: 13mV/decade; turn OFF: 32mV/decade) owing to the switching of the ferroelectric [15].

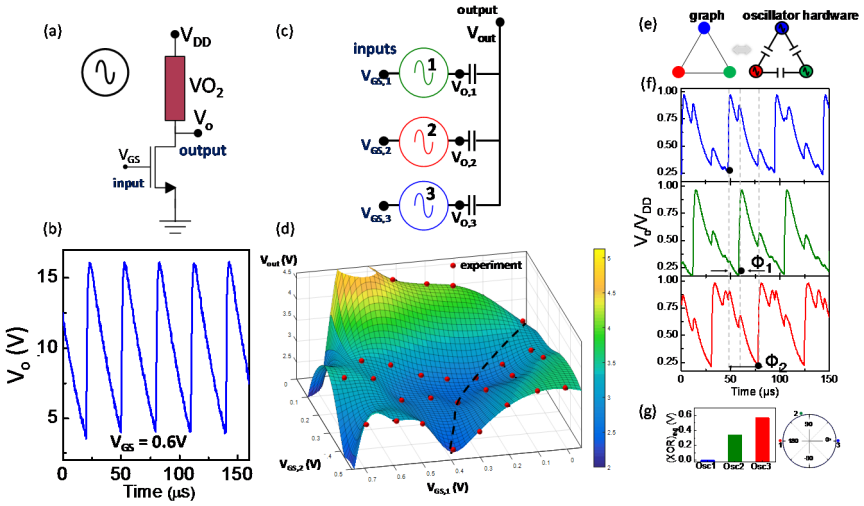


**Fig. 6: (a)** Circuit schematic of a $VO_2$ based relaxation oscillator consisting of a two-terminal $VO_2$ device in series with the channel of a FET **(b)** Time domain waveform of the $VO_2$ oscillator **(c)** Capacitively coupled $VO_2$ oscillators (star configuration). **(d)** Output voltage as a function of inputs for star configuration; the attractor like characteristics enable associative computing **(e)** $VO_2$ coupled oscillator hardware mapped to a representative graph (node ≡ oscillator; edge ≡ coupling capacitor) **(f)** time-domain waveform of the coupled oscillators. **(g)** The distinct phase dynamics among the oscillators enables coloring of the graph. Phase is measured using time averaged XOR metric.

The relatively large hysteresis which is a concern for scaling and low voltage operation arises due to the capacitance mismatch between the ferroelectric and dielectric, and can be alleviated through scaling of the ferroelectric. Additionally, scaled ferroelectrics are also necessary to maintain compatibility with the tight fin-pitch at scaled technology nodes. Scaling conventional PZT to the *sub*-10nm thickness is a challenge. Consequently, this has spurred the earnest development of $HfO_2$-$ZrO_2$ alloys [16] which are able to exhibit ferroelectricity at scaled dimensions with appropriate materials engineering.

## IV. NON-BOOLEAN COMPUTING ARCHITECTURES

Boolean logic has been the back-bone of information processing spanning over last six decades. However, there are certain computationally hard problems like combinatorial optimization and associative processing (e.g. computer vision) wherein this conventional paradigm is fundamentally inadequate. It has been proposed that dynamical systems like coupled oscillator networks which incorporate inherent parallelism in their synchronization dynamics can enable an efficient computational paradigm for solving such problems. Such systems are envisioned as analog co-processors that would augment the conventional CMOS microprocessor. The electrically induced phase transition in $VO_2$ incorporates a fundamental instability between the insulating and the metallic states. This instability can be exploited through a negative feedback engineered using a series resistance (MOSFET channel) to realize low power relaxation oscillator (Fig.6 (a)(b)). Further, since the computational fabric lies in the synchronization dynamics of the oscillators, a capacitive coupling scheme is used to enable the exchange of reactive power among the oscillators while preventing them from disturbing each other's quiescent point [17]. The phase synchronization dynamics of such capacitively coupled $VO_2$ oscillators provides an experimental test-bed for solving computationally hard problems. In particular, the solution to two specific problems: a) Higher order distance norm computation for applications like saliency detection (Fig. 6(c)(d)) [17][18] b) Vertex coloring of graphs (Fig. 6(e)(f)(g)) is experimentally demonstrated.

## ACKNOWLEDGEMENTS

This work was supported by National Science Foundation (NSF) ASSIST Nanosystems ERC under Award Number EEC-1160483, SRC DARPA sponsored StarNet LEAST centre, NSF Visual Cortex on Silicon EXPEDITION under Award Number -1317560.

## REFERENCES

[1] Thompson, Scott E., et al. "A logic nanotechnology featuring strained-silicon." *IEEE Electron Device Letters* 25.4 191-193 (2004)

[2] Chau, R, *et al.* "Benchmarking nanotechnology for high-performance and low-power logic transistor applications." *IEEE Transactions on Nanotechnology* 4.2 153-158 (2005).

[3] Kavalieros, J., *et al.* "Tri-gate transistor architecture with high-k gate dielectrics, metal gates and strain engineering." *2006 Symposium on VLSI Technology, 2006. Digest of Technical Papers.*

[4] A. V. Thathachary *et al*, "Indium Arsenide (InAs) Single and Dual Quantum-Well Heterostructure FinFETs", IEEE Symposia on VLSI Technology and Circuits, Kyoto, 2015

[5] A. V. Thathachary *et al*., "Impact of Sidewall Passivation and Channel Composition on InxGa1-xAs FinFET Performance", IEEE Electron Device lett., vol. 36, no. 2, pp 117-119, Feb. 2015

[6] A. V. Thathachary, *et al.,* "Electron Transport in Multigate InxGa1-x As Nanowire FETs: From Diffusive to Ballistic Regimes at Room Temperature", Nano Letters 2014 Feb 12;14(2):626-33

[7] A. C. Seabaugh and Q. Zhang, "Low-voltage tunnel transistors for beyond CMOS logic," Proc. IEEE, 98(12), pp. 2095-110, 2010

[8] R. Bijesh, *et. al*., "Demonstration of InGaAs/GaAsSb near broken-gap tunnel FET with $I_{ON}$ =740 µA/µm, $g_m$=700 µS/µm and gigahertz switching performance at $V_{DS}$=0.5V," 2013 Int. Elect. Dev. Meet. (IEDM), 28.1-.4,

[9] R. Pandey *et. al*., "Demonstration of p-type $In_{0.7}Ga_{0.3}As/GaAs_{0.35}Sb_{0.65}$ and n-type $GaAs_{0.4}Sb_{0.6}/In_{0.65}Ga_{0.35}As$ complimentary Heterojunction Vertical Tunnel FETs for Ultra-Low Power Logic," VLSI technology symposium (VLSIT), pp. T206-T207, 2015

[10] Jain, Ankit, and Muhammad Ashraful Alam. "Prospects of hysteresis-free abrupt switching (0 mV/decade) in Landau switches." *IEEE Transactions on Electron Devices* 60.12 (2013): 4269-4276.

[11] J.-C. Tolédano and P. Tolédano, *The Landau theory of phase transitions*. Singapore: World Scientific, 1987.

[12] Madan, H., *et al.*, "26.5 Terahertz electrically triggered RF switch on epitaxial $VO_2$-on-Sapphire (VOS) wafer." *IEEE International Electron Devices Meeting (IEDM)*, pp. 9-3 (2015).

[13] N. Shukla, et al., "Synchronized charge oscillations in correlated electron systems", *Sci. Reports* 4, Article number: 4964 (2014)

[14] S. Salahuddin and S. Datta. "Use of negative capacitance to provide voltage amplification for low power nanoscale devices." *Nano letters* 8.2 (2008): 405-410.

[15] S. Dasgupta *et al.,* "Sub-kT/q Switching in Strong Inversion in $PbZr_{0.52}Ti_{0.48}O_3$ Gated Negative Capacitance FETs", IEEE J. Exploratory Solid-State Comp. Dev. and Cir., vol. 1, pp 43-48 (2015).

[16] Müller, Johannes, et al. "Ferroelectricity in simple binary $ZrO_2$ and $HfO_2$." *Nano letters* 12.8, 4318-4323 (2012).

[17] N. Shukla, *et al.*, "Pairwise Coupled Hybrid Vanadium Dioxide-MOSFET (HVFET) Oscillators for Non-Boolean Associative Computing", IEEE International Electron Device Meeting (IEDM) Technical Digest, pp. 673-676, (2014)

[18] N. Shukla et al., "Ultra-low powered coupled oscillators for computer vision applications" VLSI Symposium, Hawaii, (2016).