A critical review of recent progress on negative capacitance field-effect transistors

Cite as: Appl. Phys. Lett. 114, 090401 (2019); https://doi.org/10.1063/1.5092684
Submitted: 13 February 2019 . Accepted: 14 February 2019 . Published Online: 08 March 2019

Muhammad A. Alam ©, Mengwei Si ©, and Peide D. Ye ©

ARTICLES YOU MAY BE INTERESTED IN

On the validity and applicability of models of negative capacitance and implications for MOS applications
Applied Physics Letters 113, 042904 (2018); https://doi.org/10.1063/1.5036984

"Negative capacitance" in resistor-ferroelectric and ferroelectric-dielectric networks: Apparent or intrinsic?
Journal of Applied Physics 123, 105102 (2018); https://doi.org/10.1063/1.5016152

Ferroelectric or non-ferroelectric: Why so many materials exhibit “ferroelectricity” on the nanoscale
Applied Physics Reviews 4, 021302 (2017); https://doi.org/10.1063/1.4979015
A critical review of recent progress on negative capacitance field-effect transistors

The remarkable progress of electronics in the 20th century sometimes obscures the dramatic story of repeated reinvention of the underlying device technology. The reinventions were catalyzed by the limits of power dissipation and self-heating for the corresponding device technologies. In the 1950s, when the vacuum tubes reached the power dissipation limits, the more power efficient bipolar transistors took over. In the 1980s, bipolar transistors were replaced by an even more power efficient technology based on complementary metal-oxide-semiconductor (CMOS) field-effect transistors (FETs). As power consumption, self-heating and scaling considerations threaten the scaling of CMOS at the twilight of Moore’s law, it is not surprising that researchers are once again looking for a more scalable and energy-efficient replacement, such as tunnel FETs, nano-electromechanical-FETs, spin-FETs, phase-FETs. The negative capacitance field-effect transistor (NC-FET) proposed by Salahuddin and Datta5 is a recent entry to the list.

The so-called Boltzmann tyranny defines the fundamental thermionic limit of the subthreshold slope (SS) of a MOSFET at 60 mV/dec at room temperature, and therefore precludes lowering of the supply voltage and the overall power consumption. As shown in Fig. 1(a), a negative capacitance field-effect transistor (NC-FET) adds a thin-layer of ferroelectric (FE) material to the existing gate oxide of a MOSFET.

The theory suggests that the consequence of this “trivial” change can be dramatic with complete disappearance of ferroelectric hysteresis (AV), Fig. 1(b). The internal voltage at the FE-oxide interface would be larger than the gate voltage, so that the SS will reduce below the Boltzmann limit of 60 mV/dec at room temperature, as shown in Fig. 1(c). As a result, the on-current (Ion) would be reached at a lower supply voltage (VDD) and the power consumption would be reduced significantly. Moreover, Fig. 1(d) shows that unlike a traditional MOSFET, the threshold voltage (Vth) would actually increase as VDD increases, making transistor scaling easier. The elegant simplicity of the device concept and the urgent need for a new “transistor” at the twilight of the Moore’s law have inspired many researchers in industry and academia to explore the physics and technology of the NC-FET, and since 2008, hundreds of papers have been published.

Despite the simplicity of the original NC-FET theory, the experimental data accumulated over the years [Fig. 1(e)] show a relatively broad scatter. This level of scatter is not unexpected for a fundamentally new class of transistor. However, in the context of the frantic pace of activities, scatter in the published data, challenges of characterization, and emergence of a diversity of models used to interpret the results, some researchers have asked thoughtful and interesting questions regarding the physics and viability of the device technology which are summarized in the following discussion.

The questions related to the basic issue of polarization of the thin ferroelectric layer are shown in Fig. 1(a). In a NC-FET, the series addition of a sufficiently large positive (i.e., gate or depletion) capacitor is expected to stabilize the FE in the zero polarization state. In the original NC-FET theory, the NC effect (also called “quasi-static NC”) is realized without polarization switching. In contrast, the “transient NC” effect requires and is associated with real polarization switching. Theoretically, the zero polarization state can be interpreted either by a single-domain or a multi-domain model. In the single domain approximation, each unit cell of the ferroelectric (FE) material to the existing gate oxide of a MOSFET.

The remarkable progress of electronics in the 20th century sometimes obscures the dramatic story of repeated reinvention of the underlying device technology. The reinventions were catalyzed by the limits of power dissipation and self-heating for the corresponding device technologies. In the 1950s, when the vacuum tubes reached the power dissipation limits, the more power efficient bipolar transistors took over. In the 1980s, bipolar transistors were replaced by an even more power efficient technology based on complementary metal-oxide-semiconductor (CMOS) field-effect transistors (FETs). As power consumption, self-heating and scaling considerations threaten the scaling of CMOS at the twilight of Moore’s law, it is not surprising that researchers are once again looking for a more scalable and energy-efficient replacement, such as tunnel FETs, nano-electromechanical-FETs, spin-FETs, phase-FETs. The negative capacitance field-effect transistor (NC-FET) proposed by Salahuddin and Datta5 is a recent entry to the list.

The so-called Boltzmann tyranny defines the fundamental thermionic limit of the subthreshold slope (SS) of a MOSFET at 60 mV/dec at room temperature, and therefore precludes lowering of the supply voltage and the overall power consumption. As shown in Fig. 1(a), a negative capacitance field-effect transistor (NC-FET) adds a thin-layer of ferroelectric (FE) material to the existing gate oxide of a MOSFET.
consistent, i.e., a single model must interpret all the relevant experiments. The goal is to organize and compare the results of various experiments and modeling efforts published to date, (b) to use the information gathered to answer a set of important questions in the field, and (c) to suggest a protocol for reporting NC-FET experiments. In this rapidly evolving field, we cannot offer conclusions, but simply provide some starting points for a coherent discussion.

In essence, there are three types of questions regarding a NC-FET.

1. Can a capacitor be negative? Is there any device that unambiguously demonstrated negative capacitance?

   The answer is yes. The negative capacitance associated with a micro-electro-mechanical (MEM) switch can be unambiguously stabilized at any position within the unstable region, demonstrating the existence and the utility of the negative capacitance. A number of experimental results and theoretical calculations support this concept.

2. Given the domain dynamics, can a FE-based capacitor or FET show negative capacitance? Are the reported transient and steady-state experiments conclusive?

   Over the years, four types of experiments shown in Fig. 2 have addressed this question.

   In small-signal measurements [Fig. 2(a)], the total capacitance of the stack (at \( V_G = 0 \)) is reported to be larger than the capacitance of the dielectric layer, suggesting the validity of the single-domain approximation. Such DC enhancement was not observed in a multi-domain ferroelectric hafnium zirconium oxide (HZO) externally connected to a commercial DE capacitor, or in a FE/DE stack.8,14 The small-signal measurement is considered as a quasi-static measurement without triggering the polarization switching. To explain the apparent discrepancy, some researchers have suggested that the metal interlayer (needed to measure the internal node voltage) fundamentally alters the FE-polarization.8,9 A precise mathematical formulation of the essential difference between the two structures is still being formulated.

   In transient RC measurements [Fig. 2(b)], a voltage drop across the ferroelectric capacitor is observed when applying a voltage pulse.15–18 Initially, the phenomenon was interpreted by a single-domain Landau-Khalatnikov (L-K) model with renormalized parameters, and was taken as an unambiguous proof of the existence of a negative capacitance effect in the ferroelectric insulator.15 Recently, other groups have argued that multi-domain variants, such as the Kolmogorov-Avrami-Ishibashi (KAI) model19 or Preisach-Miller models20–22 can also explain the experimental observations.

   FIG. 1. (a) Schematic image of a NC-FET with ferroelectric and conventional dielectric as the gate stack. (b) The fundamental difference in transfer characteristics of a Fe-FET versus a NC-FET which has an anti-clockwise hysteresis or zero-hysteresis, respectively. (c) Expected steep-slope less than 60 mV/dec at room temperature for a NC-FET. (d) Expected negative DIBL and negative drain resistance for a NC-FET. (e) Summary of the reported representative data in the literature in terms of SS versus hysteresis in transfer characteristics: [Si,26,28,33–36,41–45 Ge/GeSn,9–13,2D,9–41 InGaAs9–13]. SS is plotted as the larger SS in forward and reverse gate sweeps and only when both are available. Data without explicitly reported hysteresis are plotted with 1 mV hysteresis. (f) Summary of reported switch times of representative ferroelectric films versus the electric field by different characterization methods in the literature: [BTO (R-C),90 PZT/PNZT (R-C),89,91 P(VDF-TrFE) (R-C), FE:HfO2 (I-V),78,92–94 FE:HfO2 (ring oscillator),43,95–97 P(VDF-TrFE) (R-C), FE:HfO2 (I-V),78,92–94 FE:HfO2 (ring oscillator),43,95–97 P(VDF-TrFE) (R-C), FE:HfO2 (optical)43].

   FIG. 2. Four types of experiments have been used to characterize the negative capacitance effect. (a) Two configurations for the small signal measurement: (i) The internal metallic node separating \( C_{DE} \) and \( C_{GE} \) capacitances is used to measure the voltage/capacitances and (ii) the total capacitance of the FE/DE stack is measured and compared to \( C_{DE} \). (b) Transient RC measurement: Unlike typical RC decay, the voltage across the ferroelectric capacitor may actually increase, (c) Ramp pulse measurement: Voltage change across the \( C_{DE} (\Delta V_{DE}) \) could be larger than the voltage change in \( V_{pulse} (\Delta V_{pulse}) \), as a consequence of voltage amplification. (d) Transistor measurement: The sub-60 mV/dec subthreshold slope at room temperature and the hysteresis-free and negative differential resistance are signatures of the NC-FET.
attributed the transient voltage-drop to the delay in domain flipping associated with polarization switching and discharging of the dielectric components.\textsuperscript{39,51,23,24} This alternate explanation suggests that the transient experiments may not be able to conclusively distinguish between single and multi-domain dynamics.

A third type of measurement involves applying a ramp voltage pulse to the series combination of FE and DE capacitors, as shown in Fig. 2(c). A differential voltage amplification on the DE capacitor by the FE capacitor was observed.\textsuperscript{31} Once again, the result can be explained by either a single domain or a multi-domain switching model.\textsuperscript{22}

Although the small signal, pulse, and ramp voltage experiments have not produced a definitive conclusion, they highlighted the need to distinguish between samples with and without internal nodes, thin vs. thick ferroelectrics, one- vs. two-dimensional analysis, and the importance of leakage current in interpreting the diversity of the results reported to date.\textsuperscript{36–39}

The fourth and final type of experiments ([Fig. 2(d)]) involve fabricating an NC-FET and directly measuring its subthreshold slope, on-current, drain-induced barrier lowering (DBL), and output conductance, as shown in Figs. 1(b)–1(d). The scatter in the NC-FET data shown in Fig. 1(e) could be understood to result from considerations discussed next.

A common pitfall is to confuse NC-FETs with ferroelectric FETs (Fe-FETs). One must not confuse NC-FETs with Fe-FETs; they are structurally identical, but functionally distinct. A Fe-FET has hysteretic I-V characteristics, but a NC-FET does not, see Fig. 1(b). In a NC-FET, the total gate capacitance is positive, which means that the negative capacitance state of the ferroelectric insulator is stabilized in a single state according to the quasi-static NC model.\textsuperscript{7} In a Fe-FET, the total gate capacitance is negative (if we use the NC concept to understand the ferroelectric switch), so that the transistor switches between two states with the corresponding hysteresis in the transfer characteristics as highlighted in Fig. 1(b). In addition, if the ferroelectric polarization switching in a Fe-FET happens in the “subthreshold” region, a SS with deep sub-60 mV/dec at room temperature may be observed. However, the deep sub-60 mV/dec in a Fe-FET comes from the ferroelectric polarization switching instead of the negative capacitance effect. Although a NC-FET is fundamentally different from a Fe-FET, a Fe-FET-based logic switch with the hysteresis window less than half of the operating voltage may still offer higher on-current and lower off-current.\textsuperscript{7} Ultimately, its adoption as a logic switch will depend on the variability of the hysteresis window and the fundamental speed of (single or multiple) domain switching.

In the literature, both the Fe-FET (sometimes interpreted as an unstabilized NC-FET) and the quasi-static NC-FET have been studied and reported. The first experimental reports explored the question of “negative capacitance” and steep-switching associated with Si Fe-FETs that used a thick PVDF-TrFe) polymer as the ferroelectric insulator.\textsuperscript{16–18} The discovery of ferroelectric HfO\textsubscript{2} was an exciting advance, because it enabled CMOS compatible processing of a ferroelectric-gated MOSFET.\textsuperscript{25,30} Quasi-static or stabilized hysteresis-free Si NC-FETs with sub-60 mV/dec SS at room temperature with ferroelectric hafnium zirconium oxide (HZO) as the ferroelectric gate insulator have been reported since 2014.\textsuperscript{31–34} After these works, Si NC-FETs were studied with various gate stacks and structures, which fall into the category of either a Fe-FET (unstabilized NC-FET)\textsuperscript{35–40} or a steep-slope hysteresis-free NC-FET\textsuperscript{41–47} with minimum SS down to ~40 mV/dec at room temperature.

NC-FETs with alternate channel materials have also been reported. For example, NC-FETs with a Ge channel were demonstrated.\textsuperscript{48–54} A steep subthreshold-slope and a nearly hysteresis-free performance have been observed.\textsuperscript{54–59} The first reported 2D NC-FET applied P(VDF-TrFE) polymer as a ferroelectric insulator and MoS\textsubscript{2} as the channel material, but the fabricated device was unstable and double-sweep transfer characteristics were not measured.\textsuperscript{51} In 2017, 2D MoS\textsubscript{2} steep-slope and hysteresis-free NC-FETs were demonstrated by careful capacitance matching design.\textsuperscript{50–52} Unstabilized NC-FETs with a 2D MoS\textsubscript{2} channel were also reported. Although they achieved a sub-60 mV/dec SS at room temperature, they featured a counterclockwise hysteresis.\textsuperscript{46–48} NC-FETs using other low dimensional materials such as carbon nanotubes\textsuperscript{60} and WSe\textsubscript{2}\textsuperscript{61} have also been reported. Finally, a NC-FET with a III-V semiconductor as the channel was also demonstrated, but hysteresis-free and sub-60 mV/dec SS have not been achieved simultaneously.\textsuperscript{62–64}

Experiments must be interpreted self-consistently. A single-domain L-K theory anticipates the simultaneous occurrence of reduced SS, negative DBIL, negative drain resistance (NDR)\textsuperscript{56–60} and noise-suppression of the drain-current as shown in Figs. 1(c) and 1(d) and demonstrated experimentally in Refs.\textsuperscript{57} and \textsuperscript{68}, for example. Thus, if multiple groups were to report these features associated with a single device, it would support the existence of NC-FET operation.

It is important to understand that “multi-domain models” derived from Ginzburg-Landau theory have so far not been able to explain the observed NDR, negative DBIL, and hysteresis-free sub-60 mV/dec slope directly and self-consistently. Rather, the specialized multi-domain models (e.g., KAI,\textsuperscript{19} Miller,\textsuperscript{20–22} and/or Modified Miller\textsuperscript{16}) interpret the steady state response by suggesting that all steady-state measurements are in fact time-dependent, defined by the sweep rate of measurement. Then, they interpret the “DC” subthreshold slope as a consequence of time-dependent phase-lag, associated with ferroelectric polarization switching.\textsuperscript{41–47} Also, some models attempt to explain the hysteresis-free operation and NDR by invoking non-ideal charge trapping to compensate the counterclockwise hysteresis.\textsuperscript{23–24} Unfortunately, the charge trapping leads to substantially different forward and reverse subthreshold sweeps\textsuperscript{7} and cannot explain (essentially) hysteresis-free operation seen in many experiments, as in the bottom left corner of Fig. 1(c). A self-consistent explanation for the observed features remains an important goal for “multi-domain” theory of NC-FET operation.

3. Even if a FE-DE can be stabilized in the NC state, are the dimensions suitable for ultra-scale transistors beyond the 5 nm technology node? Would it switch fast enough? Given the unique physics of the gate stack, would the technology be reliable and immune from gate dielectric breakdown, negative bias temperature instability, hot carrier degradation, and other perspectives?

The scaling questions are device specific. For example, several groups have reported that the parasitic gate-drain capacitance of a FinFET actually improves the capacitance matching and reduces the
subthreshold slope.66,67 On the other hand, the quantum capacitance of ultra-thin body transistors may negate some of the improvement. Recently, researchers from Global Foundries have reported integrating doped hafnium oxide ferroelectric layers into state-of-the-art 14 nm Si FinFET technology and demonstrated that 101 stage ring oscillators show improved SS and actually reduce the active power consumption of the circuits.19

The frequency dependence is another question of interest. Quasi-static NC-FET models argue that transistor operation does not require domain switching,20 so that FE switching speed may not be relevant for NC-FET operation. An in-depth recent analysis7 however, concludes that a NC-FET will always switch slower than the corresponding Fe-FET. This limit reflects the fact that while the amount of polarization switching necessary for a NC-FET is substantially smaller than that of a Fe-FET, the internal field in the NC-FET is also substantially smaller than that of the Fe-FET. Therefore, one can view the Fe-FET switching speed [see Fig. 1(t)] as the upper limit of NC-FET switching. A recent experiment has reported a 3.6 ns single pulse response and a 100 ps multi-pulse response of a HfO2-based ferroelectric switch,46 suggesting the possibility of achieving near GHz operation. Single ultrashort pulse measurements could be just limited by obtaining sufficient inversion charges to support FE switching. Finally, a report based on the optical characterization of polarization switching6 and recent experiments by Global Foundries5,59 indicate that tens-of-GHz switching may be possible. Figure 1(f) summarizes the representative reports in terms of the ferroelectric switching speed and compares NC-FETs to the current Si transistor technology.84 The unification of various characterization methodologies and quantification of the damping coefficient in the L-K equation are essential for future progress regarding this topic. Unless new data show otherwise, one may be cautiously optimistic regarding the switching speed of these transistors. Finally, reliability issues place several important constraints on device operation.81–84 The voltage application at the dielectric node of the gate-stack suggests that dielectric breakdown considerations would restrict the NC-FET operation at the same on-current with the same interface field, but at a reduced operating voltage. It has been suggested that a V-shaped field profile in the gate-stack would lead to bias-temperature instability (BTI) issues related to collection of the tunneling, soft-breakdown, and hot carrier injection (HCI)-induced current at the dielectric/ferroelectric interface.85,86 Fortunately, the interface of defect generation, parasitic gate-drain capacitance, and negative capacitance is likely to suppress the Negative-Bias Temperature Instability (NBTI) degradation—the most important reliability concern for modern MOSFETs.8 Also, since the HZO transition temperature is sufficiently high, self-heating induced changes in the Landau-coefficients may not be an important issue. Transistor reliability is fundamentally important and establishing NC-FET reliability would be an important goal for future research.

To summarize, the discussion above related to these three questions highlights the fact that a fragmented approach (that only emphasizes the reduction of the subthreshold slope) has created a confusing mix of results in the field. The NC-FET concept must be self-consistently validated by (a) a combination of dielectric and ferroelectric thicknesses, (b) a broad set of sweep ranges and rates, and (c) comprehensive reports of transfer, output, and noise characteristics, demonstrating hysteresis-free operation, negative DBIL, negative drain resistance, and suppression of 1/f noise. It is important to report frequency response of an isolated NC-FET to establish a lower limit of operation of these transistors. Reliability studies of NC-FETs, i.e., stability of charge accumulation and threshold voltage, voltage acceleration and ferroelectric dielectric stack breakdown, are urgently needed. Obviously, as the transistor technology scales below 5 nm node, the critical device dimensions are extremely small, and the questions of integrating sufficiently thick FE-layers into a gate stack as well as ensuring high speed and reliability would become increasing important research topics.

The NC-FET concept provides a unifying perspective to a broad range of device phenomena collectively known as Landau switches, and it allows arbitrary tailoring of the energy landscape.74,100 Although the validity of quasi-static NC is still being debated, the concept of NC—if conclusively demonstrated—will have broad implications for device physics. Indeed, its conceptual demonstration would open up a broad class of applications including electro-chemical sensing and MEMS-based actuation.7 In this regard, the experience of thick ferroelectric films should inform, but not constraint future research in the field. The HfO2-based ferroelectric films are relatively new to the material/device communities, therefore their properties may be substantially different from traditional ferroelectric materials. The domain dynamics of such a constrained thin film is indeed not known. Given the urgency of finding a new low-power switch, NC-FET research justifiably merits the broad attention and the in-depth analysis it has received from the device physics community over the last decade.

See supplementary material for a summary of all the representative models.

The authors acknowledge valuable discussions with Supriyo Datta, M. S. Lundstrom, K. Karda, S. Salahuddin, Suman Datta, A. I. Khan, K. K. Ng, S. K. Gupta, J. van Houdt, and many other experts in the field. The work was supported in part by ASCENT, one of the six centers in JUMP, a semiconductor Research Corporation (SRC) program sponsored by DARPA.

REFERENCES

5S. Salahuddin and S. Datta, Nano Lett. 8, 405 (2008).

Applied Physics Letters


