NCFET to Rescue Technology Scaling: Opportunities and Challenges

Hussam Amrouch*, Victor M. van Santen*, Girish Pahwa†, Yogesh Chauhan†, Jörg Henkel*

*Department of Computer Science, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

†Electrical Engineering Department, Indian Institute of Technology Kanpur, Kanpur, India

{amrouch, victor.santen, henkel}@kit.edu, {girish, chauhan}@iitk.ac.in

(Invited Paper)

Abstract—Negative Capacitance Field Effect Transistor (NCFET) is one of the promising emerging technologies that may overcome the fundamental limits of conventional CMOS technology. NCFET features a ferroelectric (FE) layer within the transistor's gate, which internally amplifies the voltage, allowing NCFET to operate at a lower voltage while sustaining performance at considerable energy savings. In this work, we raise awareness that n- and p-NCFET transistors are *asymmetrically* affected by the FE layer and show, for the first time, how this asymmetry results in unbalanced circuit performance (e.g., longer fall than rise propagation delay, reduced noise margins). As NCFET are meant to maintain performance while reducing power, we present a solution by scaling the number of fins in n-NCFET to regain symmetry. We optimize iteratively in conjunction with supply voltage scaling to find the minimal energy consumption while maintaining performance. In our first case study, we achieve at least 34% lower power consumption and thus 34% higher energy efficiency as the circuit exhibits identical propagation delay. However, our second case study reveals that NCFETs can consume 3x more power and energy than the FinFET design. In summary, not considering the asymmetry and replacing FinFET with current-matched NCFET results in unreliable circuits (timing violations). This work exemplifies how the power and energy consumption of a NCFET circuit might surpass that of a FinFET, if circuits are designed considering asymmetry and circuit metric matching.

Index Terms—Emerging technology, Negative Capacitance, NCFET, FinFET, Beyond-CMOS

I. INTRODUCTION

Negative Capacitance Field Effect Transistor (NCFET) is an emerging technology, which is an evolution of a regular FinFET that has recently become compatible with the existing CMOS fabrication [1]. NCFET includes a ferroelectric (FE) layer within the gate stack of a transistor, which acts under certain conditions as a negative capacitance. The latter results in an internal voltage amplification, which allows the transistor to have a larger gain without the need to increase the operating voltage [2]. This, in turn, enables NCFET to feature a sub-threshold swing (SS) that goes beyond the Boltzmann limit of 60mV/dec at room temperature, which is the fundamental limit of MOSFET [3] [4]. NCFET opens the door for developing ultra-low power circuits due to the internal voltage amplification provided by the FE layer. For example, while the transistor operates at a supply voltage of $V_{DD} = 0.47V$, the FE layer manifests itself as negative capacitance and thus amplifies the voltage, so that charge density in the channel is equivalent to a gate potential of $V_G = 0.7V$. In other words, the transistor is able to operate at a much lower voltage without a loss in the performance, which provides circuits with considerable power savings.

This internal voltage amplification is shown in Fig. 1b. As can be noticed, the ferroelectric layer amplifies the gate voltage V_G applied to a transistor, which increases the charge density in the transistor's channel leading to a higher ON-current (I_{ON}). The integration of the differential gain (A_V) curve shown in Fig. 1b over V_G provide the average gain V_{avg}). In Fig. 1a, we show the average again (A_{avg}) at different voltages starting from the nominal voltage, which is 0.7V in the studied 7nm FinFET technology, all the way down to 0.2V. Note that, A_V is subject to the internal transistor capacitance C_{int} (i.e. the baseline MOS capacitance of the underlying transistor) and the capacitance of added ferroelectric layer C_{fe}. The matching between those two capacitances determines the obtained differential gain, which is expressed as [5]:

$$A_V = \frac{|C_{fe}|}{|C_{fe}| - C_{int}} \qquad (1)$$

It is noteworthy that both A_V as well as A_{avg} depend on the voltage applied to the NCFET, which in the context of digital circuits is typically the supply voltage V_{DD} itself.

The potential benefits of the NCFETs have been corroborated by several different experimental works both at the device level as well as at the circuit level [3], [6], [7] – especially after the discovery of ferroelectricity in HfO$_2$-based materials [8]. This made NCFET compatible with the existing CMOS fabrication process because HfO$_2$ is the used material in current technologies to fabricate high-κ gates for MOSFET transistors. Importantly, the recent demonstration of ferroelectric integrated 14 nm node NC-FinFETs by GlobalFoundries has opened doors to the production of NCFETs at advanced technology nodes [1]. Further, NCFETs with gate all around nanowire and nanosheet geometries have been also recently demonstrated [4], [9].

When NCFETs are employed, the main goal is to decrease the power consumption of the circuit through lowering the supply voltage V_{DD} while maintaining circuit's performance.
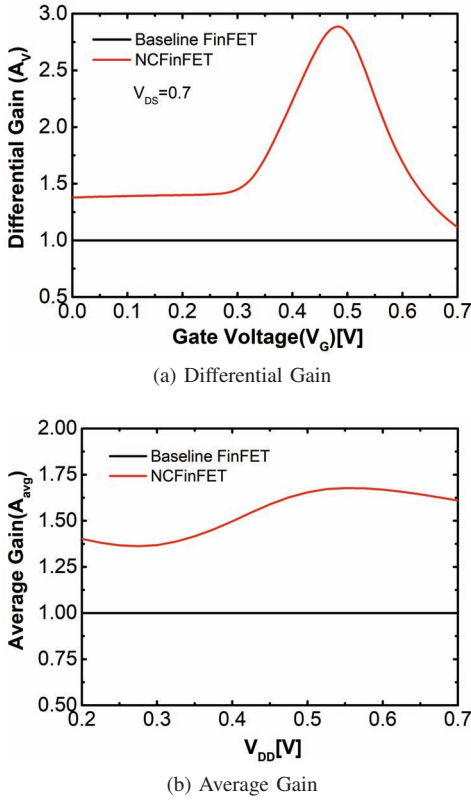
(a) Differential Gain

(b) Average Gain

Fig. 1. (a) Voltage amplification (i.e. gain ratio) at the gate of an NCFET transistor provided by the ferroelectric layer. (b) Differential amplification (gain) ratio of different operating voltages (V_{DD}).

Fig. 2. (a) V_{dd} reduction possible in NCFET transistors, with a ferroelectric layer thickness of $t_{fe} = 4nm$, while I_{ON} of NCFET still matches original current. (b) NCFET transistors are calibrated to match I_D at same V_{th} to ensure that the circuits switch at the same time.

This increases energy efficiency, as propagation delays remain constant (same performance), yet less dynamic and leakage power are consumed due to the lower V_{DD}. Note that lower voltage mitigates self-heating problems in advanced technology structures like nanowires [10] as well as mitigates transistor aging [11], [12]. The methodology for iso-performance low power NCFET circuits is to match the ON-current I_{ON} of baseline transistor and NCFET transistor [2] [3]. Fig. 2(a) shows how I_{ON} is matched at $V_{match.n} = 0.52V$ for nMOS and $V_{match.p} = 0.47V$ for pMOS (absolute value is used). When the transistor is operated at its corresponding V_{match}, the drain currents $I_D = I_{ON}$ match. Therefore, for the same load capacitance C_{load} the same charging time (i.e. propagation delay) is required. Thus, the common assumption [2] [5] is that the circuit features the same performance.

However, in this work, we show how side-effects of NCFETs break that assumption. NCFETs feature an increase in the gate capacitance due to the added ferroelectric layer, which acts as a plate capacitor and nMOS and pMOS transistors benefit *asymmetrically* from the introduction of the ferroelectric layer. In this work, we reveal, for the first time, the hidden impact of asymmetric ON-currents in p-NCFET and n-NCFET on circuit performance (propagation delay).

II. RELATED WORK

In the following, we summarize the state of the art into two main categories.
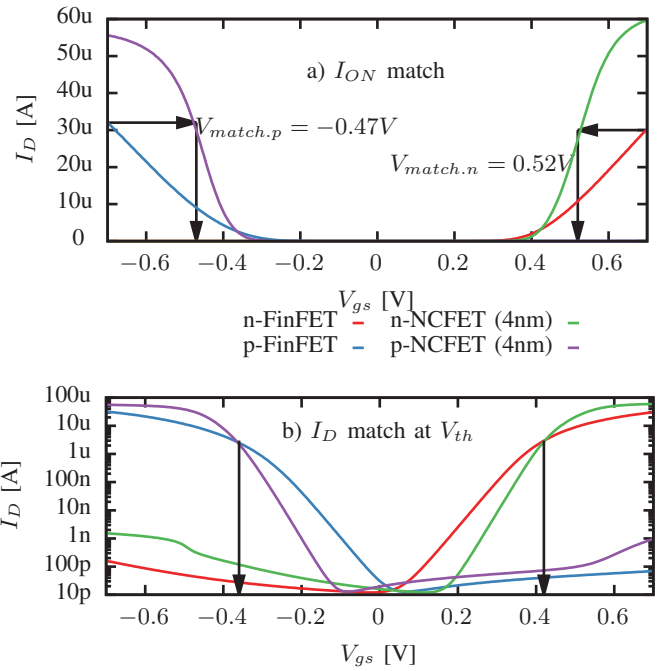
(1) Device-Level: The initial concept of NCFET was first proposed in 2008 [2]. After which, several works [3], [6], [7] demonstrated different manufactured prototypes with clear steep-slope electrical properties (i.e. beyond Boltzmann limit of $< 60mV$ sub-threshold swing). The majority of device-level work focused on the advancement of the manufacturing or semiconductor technology, report a single transistor type (either n-type or p-type) as an example [3] [6] [9] and thus state of the art is unaware of the potential asymmetry between p-NCFET and n-NCFET device.

The work in [7] shows p-NCFET and n-NCFET current plots with evidence of asymmetry, but the initial FinFET transistors were already asymmetric. Therefore, the authors do not mention or consider the additional asymmetry introduced by the ferroelectric layer itself. As their test structures are already designed to mitigate asymmetry, they report balanced performance in their ring oscillators and SRAM cells comparable to their nominal FinFET metrics. Other works also report asymmetry values [1], but do not mention the asymmetry. Employing NCFET in large-scale circuit design is proposed in our work [5], where standard cells are transitioned to NCFETs. However, we did not consider asymmetry, leading to sub-optimal results as the synthesis tool cannot fully compensate for the NCFET asymmetry.

(2) System-Level: When it comes to the system level, there are two key improvements with respect to high performance and low-power applications [13]: (1) NCFET circuits can operate at a higher frequency at the same operating voltage compared to conventional (counterpart) FinFETs. (2) NCFET
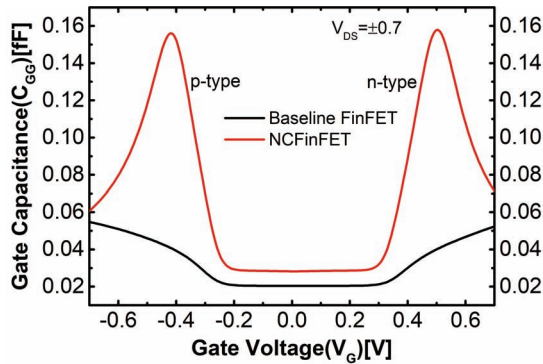
Fig. 3. Comparison of gate capacitance C_{gg} in NCFET and regular Fin-FET transistors. As shown, the NCFET transistor has a much higher gate capacitance due to the negative capacitance effect caused by the additional ferroelectric layer inserted inside the transistor gate stack.

Fig. 4. Delay of a NAND2 cell for FinFET and NCFET. It highlights how the actual circuit performance (with taking C_{gg} and I_{ON} asymmetry increase into account) is significantly worse than the nominal circuit performance. This might result in timing violations.

circuits can operate at a lower voltage, yet they can be still clocked with the same frequency as in the conventional circuits. Dynamic Voltage Frequency Scaling (DVFS) technique, as a power management technique, is an effective way to reduce the energy consumption of an application at runtime. DVFS typically aims to select the minimum operating voltage $V_{DD.min}$ that sustains the requested frequency f_{clk}. Reducing V_{DD}, in FinFET, results in reducing the total power consumption from reductions in both dynamic and leakage power. However, such a well-known voltage dependency becomes inverse for leakage power in NCFET technology due to the negative DIBL effect. With such opposed dependencies (dynamic and leakage) with respect to the voltage, total power follows the dominant power component while scaling the voltage, which leads to a novel trade-off [14]. Recently, energy-optimization in NCFET-based processors were proposed to explore the new trade-offs that NCFET technology brings to the system level [15]

III. NEW CHALLENGES IN NCFET CIRCUIT DESIGN

In practice, employing NCFETs in circuits is not just simply replacing FinFET transistors with NCFET transistors. NCFET transistors do not just have the voltage amplification but also other differences to FinFET transistors (as discussed in the previous sections). These side-effects should be considered, when NCFETs are introduced in circuits.

A. Impact of Capacitance Increase in NCFET circuits

The promise of employing NCFET circuits is to reduce the power consumption of the circuit, while maintaining the performance the circuit [2] [5]. Reducing V_{DD} lowers both the dynamic as well as the leakage power of circuit [5]. This is due to the fact, that charging a capacitance (gate capacitance, wire capacitance, etc.) requires less carriers to charge that capacitance to a lower potential (i.e. voltage).

The intuitive approach when employing NCFET in circuits is a simple two-step process:

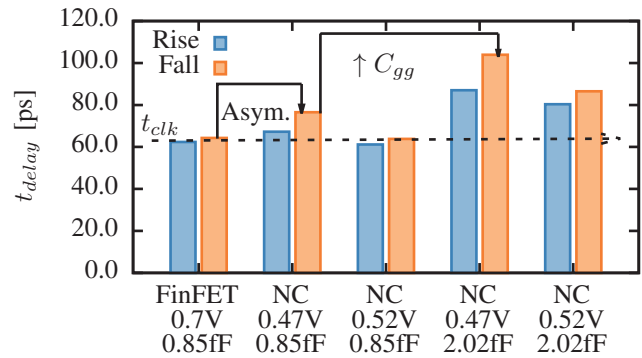1) Deposit an additional ferroelectric layer onto all transistors to convert the FinFET into NCFET transistors

2) Operate circuit at iso-performance V_{DD}, i.e. V_{match}.

Intuitively, this approach does result in lower power consumption of the circuit due to the reduced V_{DD} at the same performance. However, the actual circuit metrics (e.g., delay) are worse than nominal, i.e. those of the original FinFET design. This is due to two side-effects. The known but not considered asymmetry, discussed in Section III-B as well as the well-known gate capacitance (C_{gg}) increase, which is discussed now. The NCFET transistors do provide comparable ON-currents (I_D) to the FinFET at nominal V_{DD}, but they have to drive a higher capacitive load (C_{load}).

In a digital circuit, each standard cell charges the gates of the transistors in the following cell (in the logic path). However, with the introduction of the ferroelectric layer in the transistor, it is well-known that the capacitance of these gates C_{gg} is significantly higher [7] [1]. As seen in Fig. 2(a), we reach the same current at $V_{DD} = 0.47V$. At the same time, according to Fig. 3, at 0.47V C_{gg} is approximately 3x higher. The same current (as the intuitive approach matched I_D) now has to charge a larger load capacitance, which does take longer. Fig. 5 shows how a NAND2 cell has prolonged delays due to the asymmetry and the increase in C_{gg}. Due to the asymmetry fall times are prolonged, while rise times are barely affected, which creates an *unbalanced* cell (longer fall than rise times). C_{gg} increases both rise and fall times. Such an unexpected increase in the propagation delay of circuit leads to timing violations in the circuit [5]. As the designer is unaware of sub-nominal performance, the clock frequency is too high and the NCFET-based circuits do not deliver stable result before the clock cycle ends (timing violations).

B. Asymmetry in NCFET

Next to the well-known increase in C_{gg} NCFETs have another side-effect, which must be considered during circuit design. NCFETs benefit asymmetrically from the ferroelectric layer in nMOS and pMOS. While the p-NCFET transistors gain significantly from the ferroelectric layer, the n-NCFETs do not see a voltage increase as significant. Fig. 2(a) shows how for I_{ON} matching a p-NCFET needs just 0.47V while
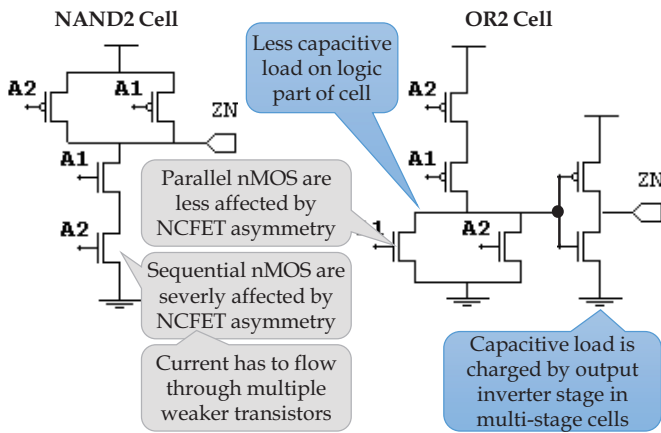
Fig. 5. Impact of NCFET changes with the topology of the cells. Cells which have nMOS transistors in sequentially arranged, have the current flow through multiple weakened nMOS transistors. Thus, these cells exhibit a stronger shift in delay and power due to the asymmetry. If the cells are multi-stage cells (e.g., building a OR cell from NOR + INV), the increased C_{load} (due to higher C_{gg}) is shouldered by the output stage and thus the topology within the logic part of the cell matters less.

a n-NCFET requires 0.52V. When the circuit is operated at $V_{match.p} = 0.47V$, the n-NCFET cannot reach the desired I_D and thus underperforms. Fig. 2 shows how the single fin n-NCFET transistor provides $11.6\mu A$ at $V_{match.p} = 0.47V$, while the single fin p-NCFET can provide $31.6\mu A$ close to the desired $30\mu A$. If instead $V_{match.p} = 0.52V$ is selected, the p-NCFET provides $43.7\mu A$, which is 45% higher current than the n-NCFET. Therefore, at any voltage there is an asymmetry between n-NCFET and p-NCFET[1]. *This is the first-time that such an asymmetry is explicitly mentioned for NCFET transistors.* Previous work like [1] might have been aware, but since I_D curves are frequently plotted in log-scale, it is not visible in these works. Other works like [7] have a visible asymmetry, however no previous work has mentioned or considered this phenomenon.

1) Asymmetry in Standard Cells: Standard cells are the basic blocks from which every large digital circuit (e.g., microprocessors) is build. Therefore, studying the impact of the asymmetry in NCFET within standard cells is important for digital circuit design. The key circuit metrics of a standard cell are its propagation delay (for each path from each input pin to each output pin) and its power consumption (again for each path). Therefore, during the characterization of a cell library, this delay and power information is stored for each cell per path and under different operating conditions (load capacitance C_{load} and signal slew t_{slew}) [5].

The asymmetry of NCFET affects both delay and power of standard cells. We start by discussing delay in the following, which is then followed by the impact on dynamic and leakage power.

Fig. 5 shows how the topology of the standard cells matters

[1]Note, that at lower V_{DD} n-NCFET is weaker, while at higher V_{DD} it is actually stronger.

for propagation delay. If the NAND cell on the left discharges C_{load} a conductive path from ZN to V_{DD} through the two nMOS transistors is formed. This is undesirable for NCFET designs, as n-NCFET are up to three times weaker than n-FinFET (comparing NCFET at 0.47V with FinFET at 0.7V (see Fig. 2(a))) and thus this discharge takes up three times longer. Comparing this to charging C_{load} via one of the pMOS transistors, it does not matter if p-NCFET at $V_{match.p}$ or p-FinFET at nominal V_{DD} are used, as their ON-currents are identical. Therefore, rise delay is unaffected, while fall delay is prolonged.

2) Solutions of NCFET Asymmetry: In this section, we discuss the multiple solutions to NCFET asymmetry (i.e. to strengthen the n-NCFET), which are as follows:

(a) Change circuit topology, e.g. use multi-stage cells.
(b) Decrease the dopant concentration of n-NCFET.
(c) Decrease the work function of n-NCFET.
(d) Increase t_{fe} to increase the voltage amplification.
(e) Increase the number of fins for n-NCFET.

The topology of a cell could be changed to a multi-stage cell (solution (a)). This ensures that the impact of the asymmetry is minimized, as now the C_{load} is shouldered by the output stage inverter. This does not just minimize the load on the first section of the cell, but also changes the topology of the first logic part of the cell into a more favorable one. For example, building a NAND2 from an AND2 + INV removes the sequential n-NCFET transistors and thus further reduces the impact of asymmetry on the delay of the cell. However, (a) is not always practical due to constraints in area and power (multi-stage cells consume more of both due to the higher number of transistors compared to single-stage cells).

The solutions (b)-(d) change the n-NCFET itself, without altering the p-NCFET type to remove the asymmetry by strengthening the transistor and thus increasing I_Dn-NCFET). Decreasing dopant concentration (b) or decreasing the work function (c) is unpractical, as this affects the V_{th} of the transistor. To gain a 3x increase in I_D such a large decrease in V_{th}(n-NCFET) is necessary, that leakage of the n-NCFET transistor becomes too high. Furthermore, the large difference in V_{th}(n-NCFET) and V_{th}(p-NCFET) would result in unequal switching of the digital circuit (n-NCFET switches earlier or later than p-NCFET), which frequently cannot be tolerated (e.g., in clock trees, where the duty cycle must be at 50%).

Altering the thickness of the ferroelectric layer t_{fe} solely for n-NCFET results in hysteresis. Increasing t_{fe} beyond 4 nm results in hysteresis effects [16], in which the gate voltage of the transistor does not fall in the same manner as it has risen. This results in timing issues (e.g., increased delays) and thus cannot be tolerated. Additionally, designers already use the thickest hysteresis-free t_{fe} in p-NCFET to gain as much voltage amplification as possible to reduce V_{DD} as much as possible. Therefore, there is no additional headroom solely for n-NCFET to exploit and any additional t_{fe} used in n-NCFET does lead to hysteresis issues. Lastly, increases in t_{fe} decreases V_{th}, which again results in intolerably high leakage.
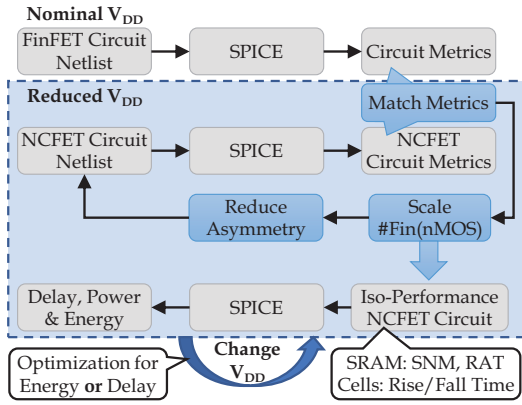
Fig. 6. An overview of our NCFET circuit design approach. We optimize for circuit metrics instead of per transistor I_D matching, as circuit metrics are ultimately what matter for a circuit designer. Additionally, our approach allows us to use less than triple the amount of fins for circuit metric matching, thus retaining as much as power savings as possible when transitioning from FinFET to NCFET.
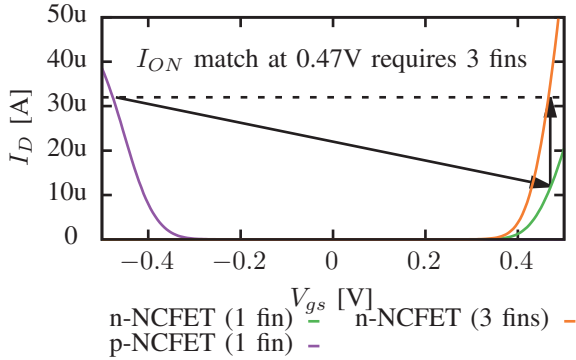

Fig. 7. For a single fin, I_{ON}(p-NCFET) = $31.6\mu A$ at $-0.47V$, while I_{ON}(n-NCFET) = $11.6\mu A$ at $0.47V$. Thus, to match I_{ON} of the p-NCFET at 1 fin at $0.47V$ with the n-NCFET transistor, the n-NCFET requires 3 fins.

Our proposed solution is therefore, to increase the number of fins of the weaker NCFET transistors (n-type in our case). Note that if the p-NCFET should be weaker than n-NCFET in future or other NCFET-technologies our work analogously applies with reversed types. We increase fin numbers for the weaker type. While this does induce an area and power overhead, it is a practical solution. The area overhead of adding fins to a transistor is smaller than adding additional transistors (like in (a)) as no additional wiring or entire transistors are necessary. The power overhead mitigates the power benefit of NCFET slightly (later in the evaluation, we evaluate how much precisely), but it does not overcome it, i.e. a power benefit still remains. To have the same performance as p-NCFET the number of fins has to be tripled (see Fig. 5). However, we show within this work, that such a high increase in the number of fins is not necessary.

IV. NCFET CIRCUIT DESIGN

This section discusses our approach of finding the optimal circuit design when employing NCFET transistors. When transitioning from FinFET to NCFET transistors, NCFET asymmetry and a C_{gg} increase complicate circuit design. Our process flow (shown in Fig. 6) describes our approach. In

our work, we base ourselves on the modeling provided in [16], which integrated the modeling of the ferroelectric layer into the industry-standard FinFET compact transistor model BSIM-CMG [17]. This work therefore allows us to model NCFETs with BSIM-CMG and therefore in circuit simulators (e.g., SPICE).

A. Match Circuit Metrics

First we simulate the FinFET circuit in the SPICE circuit simulator. By simulating at nominal voltage ($V_{DD} = 0.7V$), we obtain the nominal circuit metrics. These metrics describe the performance of the circuit and differ from circuit to circuit. For example, for standard cells, the circuit metrics are propagation delay and power consumption, while for SRAM cells, the metrics are static noise margin, read access time, write margin and critical charges.

Previous approaches in NCFET circuit design focus on matching I_{ON}, as intuitively, this should result in similar circuit metrics. However, as shown in Section III (especially Fig. 4) NCFET side-effects result in different circuit metrics, even if the transistors match with respect to I_{ON}. Instead of I_{ON} matching, the approach matches circuit metrics, as these define circuit performance and reliability. First the nominal circuit metrics are obtained. Then all FinFET transistors are replaced with NCFET transistors and V_{DD} is reduced to $V_{match.p}$. Then the NCFET circuit is simulated in SPICE to find the current circuit metrics. We consider a metric to be comparable if within 10%. 10% deviation from nominal is chosen to allow for the wide difference in FinFET and NCFET transistors. Almost all transistor properties (V_{th}, I_{ON}, I_{OFF}, SS, C_{gg}, C_{gd}, etc.) change so it is unrealistic to demand a perfect match in circuit metrics across an entire range of conditions (temperature, t_{slew}, C_{load}). So instead of trying the impossible perfect overlap between two curves (e.g. delay over C_{load} for NCFET and FinFET), we define a curve in the 10% range around nominal FinFET curve to be comparable and thus satisfactory.

B. Find #fins and V_{DD} for optimal Energy Efficiency

1) Find #fins for n-NCFET: Our next step is to change the number of fins for the n-NCFET transistors to increase their I_D and thus reducing the imbalance in our circuit. Additionally, increasing I_D counteracts the impact of the increase in C_{load} due to higher C_{gg}. Our goal is not to find a unique number of fins for each n-NCFET in the circuit, as this would result in an unfeasible large design space to explore (#nMOS would be the #dimensions to explore). Instead, the current number of fins for each n-NCFET is individually scaled with scaling factor s. As fin numbers are discrete, they are round to the nearest number of fins. According to Fig. 7: $s = 3$, which is therefore the starting point of our simulation. The resulting updated NCFET circuit netlist has the original number of fins (#fins(p)) for all p-NCFET, while the $\#fins(NCFET.n) = \lfloor s \cdot \#fins(FinFET.n) \rfloor$. Then the circuit metrics for various s (e.g. $s \in [0, 5]$ in steps of 0.5) are obtained to find the s value which has the closest match to the nominal circuit metrics.

Our approach does not change $\#fins(NCFET.p)$. In our experience, cell libraries try to put the majority of C_{load} on the nMOS transistors. For instance, the NAND2 cell in Fig. 5 is a regular cell with parallel pMOS and sequential nMOS, i.e. more load on the nMOS. Yet the OR2 cell, uses an inverter output stage to offload C_{load} to the inverter, thus removing load from the earlier logic with parallel pMOS to the nMOS in the inverter. Therefore, when switching to $V_{DD} = V_{match.p}$, pMOS transistors are I_{ON} matched and can handle the C_{load} and there is no need to scale $\#fins(NCFET.p)$.

2) Scale V_{DD} to match circuit metrics: If even the best matching s still results in insufficient circuit metrics (compared to nominal), then we do not change the circuit further. Instead, we increase $V_{DD} > V_{match.p}$. This increases the strength of both n-NCFET and p-NCFET and allows both to deal better with the increased C_{load}. We perform voltage scaling instead of further transistor scaling as otherwise the area overhead of our additional fins would become too high. Scaling voltage reduces the power benefits from transitioning to NCFET, but introduces no area overhead.

After increasing $V_{DD} > V_{match.p}$ in 50mV steps[2], we perform the same scaling $\#fins(NCFET.n)$ loop to try to find optimal s and reach comparable circuit metrics. Once comparable circuit metrics are reached (following potentially multiple voltage scaling steps), we have the first iso-performance circuit. This circuit features the same number of $\#fins(NCFET.p)$, a higher $\#fins(NCFET.n)$ and is within 10% of the original nominal circuit metrics. After comparing metrics, we report power in SPICE to check if due to the employment of NCFET transistors still power benefits exist despite power overheads due to higher $\#fins(NCFET.n)$ and $V_{DD} > V_{match.p}$. As we guarantee iso-performance, any power benefits immediately translate to better energy efficiency as well. This circuit could be the final result as it probably (see next Section) saves power and energy while maintaining performance and reliability to nominal levels.

However, our approach can be further optimized by continuing to increase V_{DD} in 50mV steps. Each step the approach checks if an even more efficient design in terms of power can be found. For example, higher V_{DD} might reduce the necessary $\#fins(NCFET.n)$ reducing the power overhead induced by the additional fins. Yet, increasing V_{DD} introduces its own power overhead. Therefore, a trade-off between V_{DD} overhead and $\#fins(NCFET.n)$ overhead must be found. We thus continue to increase V_{DD} as long as the total power of the circuit decreases, while maintaining iso-performance. As soon as total power of the circuit increases while increasing V_{DD}, we stop as we found the most efficient design $V_{DD} = V_{opt}$ at given s.

Note, that each circuit might have a unique V_{opt} due to its unique circuit topology. Therefore, when optimizing entire cell libraries, an (weighted according to cell occurrence) average

of V_{opt} should be used, as all cells have to ultimately operate at the same V_{DD}.

3) Cell characterization versus Analytic Solution: It could be argued, that because n-NCFET require $s = 3$ according to Fig. 7, the solution is simple and can be found analytically. However, due to the voltage dependency of the voltage amplification A_{avg} and gate capacitance C_{gg} (Fig. 1), this is near impossible. Fig. 1b and 1a show how both A_{avg} and V_{gg} fluctuate significantly over V_{DD}. Analytically obtaining rise and fall propagation delay of the cell would be challenging. The rise and fall time of the input signal (signal slew t_{slew}) determines how long we spend at a given voltage, i.e. how long we spend at a given A_{avg} and C_{gg}. The transistors which are directly connected to the inputs depend on the signal slew of the input, but transistors later in the logic path (e.g., the transistor of the output inverter in the OR cell in Fig. 5) depend on the output slew of input transistors. Additionally, these output slews depend on the topology of the cells and the C_{gg} of the output transistors (as this is what the input transistors charge). All these complications and interactions cannot be solved analytically (e.g., with integer linear programming or machine learning). Therefore, in this work, we opted to use cell characterizations tools, which internally use SPICE to model each transistor at each time step in full detail (i.e. considering all these shifting properties A_{avg}, C_{gg}, t_{slew} via NCFET BSIM-CMG). While cell characterization definitely is a brute-force solution, it is the only viable solution, which considers voltage dependencies, topology dependencies and transistor interactions and thus results in accurate estimations for the right #fins and V_{opt}.

C. Evaluation of NAND and OR Cells

In this work we evaluate two case studies with the standard cells NAND2 and OR2. NAND2 and OR2 are the most frequently used cells and are polar opposites of each other in terms of circuit topology (see Fig. 5), with sequential and parallel transistors swapped between nMOS and pMOS transistors. Additionally, the studied OR2 cell is multi-stage, while the studied NAND2 is not. Our approach supports larger circuits, but these would be synthesized based on such standard cells instead of being built from the transistor-level up and then simulated in SPICE.

We study the cells from the ASAP7 [18] cell library with the modified model from [16] [5] to support the ferroelectric layer, which is modeled via Landua-Khalatnikov theory [19]. We then use SiliconSmart, a commercial cell characterization tool, which acts as a front-end for HSPICE v2017-3 to characterize the standard cells with respect to delay, power and energy.

1) NAND2 Cell: We chose the `NAND2xp5` from the ASAP7 [18] library, as it is the most frequently occurring cell in microprocessors synthesized with ASAP7. For high $t_{slew} = 0.867ps$ and low $C_{load} = 0.02fF$ the cell exhibits exactly the behavior we would expect (see Table I). Pin B has the largest shift in the fall time (discharge via nMOS path), which is exactly where the sequential nMOS transistors play a role. In fact, the NCFET NAND2 at $V_{match.p} = 0.47V$ is

---

[2]50mV is chosen to limit computational effort as an entire cell library has to be characterized again for each voltage step.

TABLE I

NAND2 AND OR2 CELLS AT WORST-CASE CONDITION (HIGH t_{slew} AND LOW C_{load}) RESULTS IN UNBALANCED CELLS. FOR NAND2, FIN-SCALED RESULTS ARE SUFFICIENT TO REACH COMPARABLE (WITHIN 10%) PERFORMANCE. I_{ON} MATCHING WITH V_{match} AND SOLE V_{DD} SCALING ARE INSUFFICIENT FOR ISO-PERFORMANCE. IN OR, NEITHER I_{ON} MATCHING NOR V_{DD} SCALING REACH NOMINAL FINFET DELAY WITHIN 10%.

NAND2xp5	0.7V FinFET	0.47V NCFET	Difference	0.52V NCFET	0.57V NCFET	s=3 0.57V	s=3 0.55V
Pin A Fall Time [ps]	113.1	180.6	-67.5	131.3	104.2	65.35	71.75
Pin A Rise Time [ps]	41.04	114	-72.96	81.55	64.39	106.1	10.85
Pin B Fall Time [ps]	**118.9**	**233**	-114.1	**194.1**	**169.7**	**117.2**	**122.2**
Pin B Rise Time [ps]	28.4	90.97	-62.57	66.45	51.0	84.41	85.51
OR2x2	0.7V FinFET	0.47V NCFET	Difference	0.52V NCFET	0.56V NCFET	s=3 0.56V	s=5 0.54V
Pin A Fall Time [ps]	19.81	74.68	-54.87	35.72	17.14	50.68	75.49
Pin A Rise Time [ps]	**165.3**	**300.3**	-135	**242.9**	**210.4**	**176.7**	**180.9**
Pin B Fall Time [ps]	25.67	113.8	-88.13	80.59	66.5	111	140.8
Pin B Rise Time [ps]	152.7	285.1	-132.4	231.4	200.8	170.4	178.5

TABLE II

POWER OF NAND2 AND OR2 CELL AT WORST-CASE CONDITION. OPTIMIZED NAND2 CELLS CONSUME MORE (0.57V) AND LESS (0.55V) POWER THAN NOMINAL. OR2 CELL AT WORST-CASE CONDITION. BOTH OPTIMIZED ISO-PERFORMANCE NCFET OR2 CONSUME MORE POWER.

NAND2xp5	Power FinFET 0.7V	Power NCFET 0.47V	Power NCFET 0.52V	Power NCFET 0.57V	Power s=3 0.57V	Power s=3 0.55V
Pin A fall [nW]	0.4154	0.00767	0.0497	0.209	1.289 (+210%)	0.082 (-80%)
Pin A rise [nW]	0.7136	0.0896	0.1492	0.2789	0.565 (- 20 %)	0.468 (-34%)
Pin B fall [nW]	0.4911	0.0217	0.0595	0.1467	1.066 (+117%)	0.153 (-69%)
Pin B rise [nW]	0.5403	0.0334	0.0766	0.1706	0.444 (- 18 %)	0.219 (-59%)
OR2x2	Power FinFET 0.7V	Power NCFET 0.47V	Power NCFET 0.52V	Power NCFET 0.56V	Power s=3 0.56V	Power s=5 0.54V
Pin A fall [nW]	0.5428	0.2326	0.349	0.5297	1.157 (+113%)	1.406 (+159%)
Pin A rise [nW]	0.4754	0.1716	0.2489	0.3604	0.509 (- 7 %)	0.508 (+ 7 %)
Pin B fall [nW]	0.5689	0.024	0.3348	0.4621	0.976 (+ 71%)	1.238 (+117%)
Pin B rise [nW]	0.4187	0.1166	0.1782	0.2678	0.401 (- 4 %)	0.416 (- 0.5%)

about twice (187%) the delay compared to FinFET NAND2. Even at $V_{match.n} = 0.52V$, the original NCFET design is too slow, even though both NCFET transistor types now operate at or above matching I_{ON}. Therefore, regardless at which V_{match} the FinFET are replaced with NCFET, the circuit metrics do not match and propagation delay underestimations of 64% (see Pin B Fall at 0.52V) and up to 87% (Pin B Fall 0.47V) occur.

TABLE III

NAND2 CELL AT BEST-CASE CONDITION (LOW t_{slew} AND LOW C_{load}) RESULTS IN BALANCED DELAY AND ISO-PERFORMANCE AT 0.47V

NAND2xp5	0.7V FinFET	0.47V NCFET	Difference
Pin A Fall Time	10.71	10.25	0.46
Pin A Rise Time	9.736	8.106	1.63
Pin B Fall Time	10.21	9.49	0.72
Pin B Fall Time	8.991	7.157	1.834

This trend becomes slightly better towards high C_{load}, but as long as t_{slew} is high, the original NCFET is too slow at 0.47V.

Worst-case timing: A different behavior emerges for low $t_{slew} = 0.02ns$ and low $C_{load} = 0.02fF$ (see Table III). At these conditions, the NCFET cell is fast enough at $V_{match.p}$ to be comparable to the nominal FinFET cell despite the asymmetry, which shows the importance of taking the t_{slew} and C_{load} into account. We evaluate 7 x 7 = 49 combinations of (t_{slew}, C_{load}) per pin (A and B) and rise/fall, which results in $49 \cdot 2 \cdot 2 = 196$ combinations. In 113 (57.6%) of these combinations the $V_{match.p} = 0.47V$ is sufficient for correct timing. For $V_{match.n} = 0.52V$, 153 cases (78%) result in correct timing.

Cell for worst-case timing: In order to obtain a cell, which does meet timing of 0.7V FinFET, we need to understand synthesis/timing. Synthesis cannot know if a output signal falls or rises, i.e. it will take at given conditions (e.g., high t_{slew} and low C_{load}) the worst time (i.e. the maximum of the four combinations: pin A/B rise/fall). Our approach finds the NCFET cell in which its maximum is smaller than the FinFET maximum. This is shown in the far right of Table I, in which a $V_{DD} = 0.57V$ cell with n-NCFET fin scaling factor of $s = 3$ meets the timing of FinFET 0.7V cell. Importantly, simply scaling V_{dd} to 0.57V is insufficient as the delay for pin B fall time is still too high due to the asymmetry in NCFET. In the $s = 3$ case, the NCFETs delays are much more symmetrical due to the scaling of the n-NCFET fins. To meet timing without fin scaling V_{DD} has to be increased to 0.64V, which consumes significantly more power than the fin-scaled NAND2 cell at 0.57V.

Power Consumption: The power consumption for the worst-case condition (i.e. condition of Table I) is shown in Table II. This shows that NCFET consume considerably less power if timing meets at 0.47V (57% of all cases), 0.52V (78%) or even 0.57V (96%). However, in the worst-case, which our algorithm had to select, the two fall cases actually consume more power than regular FinFET. Keep in mind, that his occurs in 7 out of 196 cases (3.5%), which are all very unlikely (super long t_{slew} with very low C_{load}). Operating with comparable (up to 10% deviation, as mentioned in Section IV-A) delay for this absolute worst-case condition results in $s = 3$ at 0.55V. The 0.02V difference in V_{DD} results in significantly lower power

consumptions, which provide power benefits over the FinFET design. The large difference in power consumption is because are close to the C_{gg} peak shown in 3) and the high voltage amplification at 0.57V. Note the difference to from 0.55V to 0.52V is still quite significant, which highlights how sensitive NCFET are to voltage.

In summary our approach finds the appropriate NAND2 cells in NCFET, which can meet timing (i.e. results in reliable operation) and still saves power and energy over the original FinFET design.

2) OR2 Cell: The OR2x2 cell is a cell with complementary topology to the NAND2 cell. Therefore, we expect different propagation delay results. For 34 (17% of the total 196) of the evaluated $V_{match.p}$ is sufficient with respect to timing. For $V_{match.n}$, 86 cases (44%) result in correct timing. These values are significantly lower than the NAND2 cell, which is unintuitive. However, the NAND2 cell fails timing almost exclusive in pin B fall delay (frequently by a large margin), while the other delays remain positive. The OR2 cell instead fails very homogenously across both pins and for both rise and fall times, which is expected due to the more balanced topology.

Power Consumption: The two candidates for optimized cells are (1) s=3 0.56V and (2) s=5 0.54V. While (1) operates at higher V_{DD} it uses less fins with (2) scaling the number of fins with 5x to operate slightly lower V_{DD}. Both (1) and (2) are within 10% of the maximum delay for pin A rise and have no higher value at other cases (even though pin B rise comes close). Unfortunately, both cases have significantly higher power consumptions for pin A and pin B fall delays. *This highlights for the first-time, that NCFET might not always be beneficial when comparable circuit metrics are desired.* The fin-scaling is also not the issue, as the NCFET cell at 0.67V would match FinFET delays, but consumes up to 7x more power than the FinFET designs.

In summary, the OR2 cell cannot be implemented more efficiently in NCFET than in FinFET. The topology does not lend itself to NCFETs and must be redesigned from the ground up to be better suitable for NCFETs with their higher C_{gg} (more load for the initial stage in multi-stage cell designs) and weaker n-NCFET.

V. Conclusion

This work revealed the hidden impact on circuits by the asymmetry between nMOS and pMOS when ferroelectric layers are introduced in their gates. We proposed a circuit design technique which mitigated this asymmetry (if necessary and possible), while ensuring the circuit metrics remained at nominal level (i.e. comparable to the FinFET implementation). Our circuit design technique scaled the number of fins for nMOS transistors and carefully selected V_{DD} to find a circuit design and operating voltage, which maintained the circuit's performance of NAND cells but reduced the energy consumption by at least 34%. These energy saving are less optimistic, than previously reported as we considered the negative side-effects of NCFETs with their newly revealed

asymmetry and well-known gate capacitance increase. For OR2 an increase of power and energy consumption of about 3x is reported, highlighting how NCFET are not universally better than traditional FinFET designs.

References

[1] Z. Krivokapic, U. Rana1, R. Galatage, A. Razavieh *et al.*, "14nm ferroelectric FinFET technology with steep subthreshold slope for ultra low power applications," in *IEEE International Electron Devices Meeting (IEDM)*, Dec. 2017, pp. 15.1.1–15.1.4.

[2] S. Salahuddin and S. Datta, "Use of negative capacitance to provide voltage amplification for low power nanoscale devices," *Nano letters*, vol. 8, no. 2, pp. 405–410, 2008.

[3] K.-S. Li, P.-G. Chen, T.-Y. Lai, C.-H. Lin *et al.*, "Sub-60mv-swing negative-capacitance finfet without hysteresis," in *2015 IEEE International Electron Devices Meeting (IEDM)*. IEEE, 2015, pp. 22–6.

[4] C.-J. Su, T.-C. Hong, Y.-C. Tsou, F.-J. H. P.-J. Sung1and *et al.*, "Ge nanowire FETs with HfZrOx ferroelectric gate stack exhibiting SS of sub-60 mv/dec and biasing effects on ferroelectric reliability," in *IEEE International Electron Devices Meeting (IEDM)*, Dec. 2017, pp. 15.4.1–15.4.4.

[5] H. Amrouch, G. Pahwa, A. D. Gaidhane, J. Henkel *et al.*, "Negative capacitance transistor to address the fundamental limitations in technology scaling: Processor performance," *IEEE Access*, vol. 6, pp. 52 754–52 765, 2018.

[6] M. H. Lee, S.-T. Fan, C.-H. Tang, P.-G. Chen *et al.*, "Physical thickness 1.x nm ferroelectric HfZrOx negative capacitance FETs," in *IEEE International Electron Devices Meeting (IEDM)*, Dec. 2016, pp. 12.1.1–12.1.4.

[7] K. Li, Y. Wei, Y. Chen, W. Chiu *et al.*, "Negative-capacitance finfet inverter, ring oscillator, sram cell, and ft," in *2018 IEEE International Electron Devices Meeting (IEDM)*, Dec 2018, pp. 31.7.1–31.7.4.

[8] J. Muller, T. S. Boscke, U. Schroder, S. Mueller *et al.*, "Ferroelectricity in simple binary zro2 and hfo2," *Nano letters*, vol. 12, no. 8, pp. 4318–4323, 2012.

[9] M. H. Lee, K. . Chen, C. . Liao, S. . Gu *et al.*, "Extremely steep switch of negative-capacitance nanosheet gaa-fets and finfets," in *2018 IEEE International Electron Devices Meeting (IEDM)*, Dec 2018, pp. 31.8.1–31.8.4.

[10] O. Prakash, S. Manhas, J. Henkel, and H. Amrouch, "Impact of NBTI Aging on Self-Heating in Nanowire FET," in *DATE*, 2020.

[11] S. Mishra, H. Amrouch, J. Joe, C. K. Dabhi *et al.*, "A simulation study of nbti impact on 14-nm node finfet technology for logic applications: Device degradation to circuit-level interaction," *IEEE Transactions on Electron Devices*, vol. 66, no. 1, pp. 271–278, Jan 2019.

[12] H. Amrouch, S. B. Ehsani, A. Gerstlauer, and J. Henkel, "On the efficiency of voltage overscaling under temperature and aging effects," *IEEE Transactions on Computers*, 2019.

[13] M. Rapp, S. Salamin, H. Amrouch, G. Pahwa *et al.*, "Performance, power and cooling trade-offs with ncfet-based many-cores," in *Proceedings of the 56th Annual Design Automation Conference 2019*. ACM, 2019, p. 41.

[14] S. Salamin, M. Rapp, H. Amrouch, G. Pahwa *et al.*, "Ncfet-aware voltage scaling," in *2019 IEEE/ACM International Symposium on Low Power Electronics and Design (ISLPED)*. IEEE, 2019, pp. 1–6.

[15] S. Salamin, M. Rapp, H. Amrouch, A. Gerstlauer *et al.*, "Energy optimization in ncfet-based processors," in *2020 Design, Automation & Test in Europe (DATE)*. IEEE, 2020, pp. 1–6.

[16] G. Pahwa, T. Dutta, A. Agarwal, S. Khandelwal *et al.*, "Analysis and compact modeling of negative capacitance transistor with high on-current and negative output differential resistancepart ii: Model validation," *IEEE Transactions on Electron Devices*, vol. 63, no. 12, pp. 4986–4992, 2016.

[17] M. V. Dunga, C.-H. Lin, A. M. Niknejad, and C. Hu, "Bsim-cmg: A compact model for multi-gate transistors," in *FinFETs and Other Multi-Gate Transistors*. Springer, 2008, pp. 113–153.

[18] L. T. Clark, V. Vashishtha, L. Shifren, A. Gujja *et al.*, "Asap7: A 7-nm finfet predictive process design kit," *Microelectronics Journal*, vol. 53, pp. 105 – 115, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S002626921630026X

[19] K. M. Rabe, M. Dawber, C. Lichtensteiger, C. H. Ahn *et al.*, "Modern physics of ferroelectrics: Essential background," in *Physics of Ferroelectrics*. Springer, 2007, pp. 1–30.