Exploring logic architectures suitable for TFETs devices

Juan Núñez and Maria J. Avedillo
Instituto de Microelectrónica de Sevilla, IMSE-CNM (CSIC/Universidad de Sevilla)
Av. Américo Vespucio s/n, 41092, Sevilla (Spain)
[jnunez,avedillo]@imse-cnm.csic.es

Abstract— Tunnel transistors are steep subthreshold slope devices suitable for low power operation so being potential candidates to overcome the power density and energy inefficiency limitations of CMOS technology, which are critical for IoT development. Although they show higher ON currents than CMOS at low supply voltages, currently TFETs do not reach those exhibited by CMOS at its nominal supply voltage and so they have being identified to be competitive for moderate operating frequencies. However, in many cases, architectural choices are not taken into account when benchmarking them against CMOS. In this paper we claim that the logic architecture should be selected in order to take full advantage of the specific characteristics of these devices. Widely used circuits are designed and evaluated showing how properly tuning the logic architecture results in raising the frequency up to which TFETs are competitive or in increasing power savings at lower frequencies.

Keywords— Tunnel transistors, Steep subthreshold slope, Low power, Low supply voltage, Fine-grained pipeline.

I. INTRODUCTION

Steep subthreshold slope (SS) devices are currently being investigated as potential candidates to solve the power and energy problem exhibited by CMOS devices. This limitation arises from their 60mV/decade minimum subthreshold slope (SS) which makes it impossible to lower the threshold voltage without producing unacceptable off-state leakage currents. As a consequence, circuit speed significantly degrades when supply voltage is reduced. This results in power density problems for high performance applications, requiring nominal supply voltages and energy inefficiency in low voltage applications. A smaller SS (SS<60mV/dec at room temperature) makes it possible to reduce threshold voltage while keeping leakage current under control, facilitating low voltage operation with acceptable speed and, thus, generating savings in power and energy.

Among steep SS devices, tunnel transistors (TFETs) [1]-[3] are one of the most attractive, and recent benchmarking of many beyond-CMOS reinorses that they are the leading low-power devices [4]. Intensive research is carried out in their development in different material systems, including Si/Ge TFETs, III-V TFETs and more recently, band-to-band TFETs based on two-dimensional transition metal dichalcogenides semiconductor [5], which paves the way for improved TFETs. Despite the achieved progress, their implementation still presents challenges. One of the major uncertainties of these devices is their ON current ($I_{ON}$) limitation. Although they show higher ON currents than CMOS at low supply voltages, currently TFETs do not reach those exhibited by CMOS at its nominal supply voltage. This is illustrated in Fig. 1 in which currents of TFETs and CMOS devices are compared. Two different projected tunnel transistor models have been used in this work. Both are available from the nanoHUB website [11]. One has been derived by Pennsylvania State University [12] and the other by Notre Dame University [13]. TFETs from Pennsylvania State University are look-up table based Verilog-A models for III-V interband TFETs based on calibrated Synopsys TCAD device simulations. Models with gate lengths of 20nm are available for a double gate GaSb-InAs Heterojunction TFET (TFETPSU). On the other hand, the TFETs models from Notre Dame University are based on the Kane-Sze formula for tunneling.-In this work, we use a model for a GaN/InN single gate TFET (TFETND).

A predictive CMOS technology has been also evaluated for comparison purposes, whose models have been obtained from the PTM web page [14]. The one selected has channel length similar to the available TFETs, namely: 22nm MOSFET devices for high performance (nominal $V_{DD}=0.8V$). Fig. 1 shows the $I_{ON}$ for the three selected devices, that is, their drain-to-source current when the gate-to-source and the drain-to-source voltages are the same (and equal to the supply voltage, $V_{DD}$). It can be observed that, although they show higher ON currents than CMOS at low supply voltages, current TFETs do not reach those exhibited by CMOS at its nominal supply voltage and, so, they have being identified to be competitive for moderate operating frequencies. Many works have shown power benefits for iso-performance or higher performance at iso-power within that limited frequency range [6]-[11]. However, in these works, architectural choices are not taken into account when benchmarking them against CMOS. A few works have pointed out that conclusions of the

Fig. 1 $I_{ON}$ current versus $V_{DD}$ for the evaluated set of transistors.
comparison can be distinct if those aspects are considered [3], [16]. The system/logic architecture should be selected in order to take full advantage of the specific characteristics of these devices. Fig. 1 shows $I_{ON}$ advantages for low supply voltages and thus they would be beneficial over CMOS when operated at such low voltages. In fact, [3] identifies the crossover supply voltage as an inexact but practical measure of the $V_{DD}$ value under which TFET power-speed is advantageous. At the system level, the number of cores can be increased in order to achieve required throughput at lower $V_{DD}$. At the logic level, pipelined registers can be added in order to achieve the required frequency at lower [16]. Through this architectural tuning, TFETs can provide power benefits also in higher performance applications or they can increase power savings. In this paper we explore the suitability of ultra-fined pipelined logic architectures (nanopipelined) based on clocked dynamic logic style for TFET circuit design and compare its potential as a power reduction technique in TFETs and CMOS technologies. The rest of the paper is structured as follows. Section II describes the motivation of this work and reports related previous work. Nanopipelined logic architecture and the evaluation experiment carried out are described in Section III. Section IV analyzes the obtained results. Finally some conclusions are given in Section V.

II. MOTIVATION AND PREVIOUS WORK

As a motivation for our work we have evaluated the power versus frequency behavior of FO4 inverters. For each target frequency and technology, minimum allowable $V_{DD}$ has been determined as the minimum supply voltage at which correct functionality, assuming a given logic depth $LD$, is observed with maximum logic swing degradation of 10%. Power has been measured at such $V_{DD}$ value. Fig. 2 depicts curves for two values of $LD$. Note logarithm scale has been applied to both axes and that power values have been normalized with respect to CMOS. It can be observed, as expected, that TFETs technologies are not able to achieve the largest frequency targets. It can be also observed that, in both cases, TFET devices exhibit advantages in terms of power for a given range of operating frequencies. There are not power savings under a given target frequency or over another given frequency. The exact position of this frequency region depends on the $LD$ value. This parameter has been halved in Fig. 2b with respect to Fig. 2a, leading to larger maximum frequencies up to which power performance is better than for CMOS. For instance, for TFETPSU, this frequency ($f_{EFF}$) is shifted up from 1.8GHz (marked with an arrow in Fig. 2a) to 3.7GHz (marked in Fig. 2b). Moreover, there are not only differences in terms of the position of the advantageous frequency range, but also in terms of the power reduction achieved by using the TFET technology. Fig. 3 depicts results of previous benchmarking experiment in a convenient way to illustrate this for CMOS and TFETPSU. Absolute power values, instead of normalized, are now shown. In addition, data for both $LD$ values are included on same picture. For that, let’s consider a frequency at which four circuits (CMOS-50, TFETPSU-50, CMOS-25 and TFETPSU-25) work (point $f_1$ in Fig. 3). At that frequency, power of TFET-50 is 62% of CMOS-50 (power saving of 38%), whereas power of TFET-25 is around 18% of CMOS-25 (power saving of 82%).-That is, power advantages are larger in the architecture with logic depth equal to 25. These results can be explained on the basis of the $I_{ON}$ versus $V_{DD}$ performance exhibited by each technology. For that architectural option, the TFET implementation can be operated around 0.25V. In the CMOS case, $V_{DD}$ does not reduce so much, explaining the larger differences. The results described herein illustrate our claims in previous section regarding the importance of suitably choosing a logic architecture to take full advantage of the TFETs devices characteristics. In particular, they show how, given a target frequency, logic depth should be selected such that required supply voltage is within the range in which TFETs are competitive with respect to CMOS in terms of current. Coming back to the example, note that assuming $LD=50$ is the original circuit, the $LD=25$ can be interpreted as pipelined versions with two stages. Of course there are power and delay overheads associated to the actual pipeline registers that should be taken into account. In [16] a simple model to
estimate the power reductions achieved when using pipeline to cut down logic depth, and taking into account the power overheads associated to the pipelined registers was developed. It was shown that in CMOS power savings cancel with the incorporation of a number of flip-flops equal to the 5% of the number of gates in the original circuit while this number rises to 90% for tunnel circuits. Simulation experiments of a simple adder tree validated the analysis. No power savings were obtained by the CMOS pipelined circuit while the TFET pipelined circuit saves 77% of power. These results indicate that TFETs transistors benefit more in terms of power than CMOS ones from logic depth reduction. This motivates the present work in which ultra-fine pipelined logic architectures without memory elements, that we denote as nanopipelining, are explored. In addition, speed limitations due to smaller $I_{ON}$ currents exhibited by these devices may make necessary this aggressive logic depth reduction, being also a motivation of the study.

III. EXPERIMENT DESCRIPTION

Dynamic logic is well suited to implement very fine grained pipelining for high performance applications, without applying conventional pipeline techniques which insert flip-flops to short down signal propagation paths in combinational logic. Instead, they rely on logic circuit styles, which naturally exhibit the capacity to block data propagation. Thus, they are well suited to implement pipeline architectures without memory elements [17]. Additionally, in [18], the suitability of the dynamic logic style for the implementation of TFET logic circuits is pointed out. In these fine grain pipeline logic styles, operating frequency (or throughput) depends both on the number of clock-phases and the number of gate levels per clock-phase. The clock period needs to accommodate all the phases and the duration of each phase is determined by the number of gate levels per clock-phase. Thus, from the point of view of ultra-high speed applications, a two-phase scheme with a single gate per clock-phase, as shown in Fig. 4a, is very attractive. As a case study, we have evaluated and compared the series interconnection of $LD$ logic gates implementing the functionality $A+B+C$. Both static and two-phase one-gate-per-phase dynamic realizations have been designed. The CMOS static realization of this circuit has been considered as a reference. Fig. 4b shows the schematic used for both CMOS and TFET dynamic gates, in which parasitic capacitances in the interconnections nodes have been included.

For each $LD$ value considered in this experiment (6, 8, 10 and 12) we have evaluated the maximum frequency achieved by the CMOS static implementation at two different supply voltages: its nominal supply voltage $V_{DD,CMOS}=0.8$V ($f_1$) and for $V_{DD,CMOS}=0.6$V ($f_2$). The second $V_{DD}$ value has been included in the experiment with the aim of not overestimating the power consumption of the conventional static CMOS reference implementation at the lower operating frequencies analyzed in this work. Since dynamic power (dominant at the frequencies of interest) is proportional to $V_{DD}^2$, power reduces almost 50% when supply voltage is reduced from its nominal value to 0.6V. Its value has been selected such that the frequency for the largest $LD$ value (12) is over 1 GHz. Then, we have evaluated the minimum $V_{DD}$ at which the dynamic circuits designed using CMOS devices and the TFETs counterparts can operate at those target frequencies $f_1$ and $f_2$.

Power consumption for all circuits has been evaluated at $f_1$ and $f_2$ by applying a long sequence of random input patterns considering the obtained minimum $V_{DD}$ values for each technology. Measured power includes clock power for the dynamic versions.

IV. RESULTS

Table I summarizes maximum frequencies for the static CMOS and minimum supply voltages for the other designs. Note that none of the TFETs statics designs can be operated at the $f_2$ frequencies. TFETND static design cannot be operated at the $f_1$ target either. Moreover, for the dynamic CMOS implementation, $V_{DD}$ reductions of 0.1V-0.15V and 0.25V-0.30V have been obtained for $f_1$ and $f_2$, respectively. Results for TFET dynamic circuits show that supply voltages can be reduced more than for CMOS implementations. Some issues should be discussed. First, $V_{DD}/2$ for TFETPSU is much lower than for TFETND, which can be explained from the differences of $I_{ON}$ current observed in Fig. 1 for both TFET devices. For the same reason TFETND cannot operate at very high frequencies at which TFETPSU can.

Fig. 5 shows power consumptions of the evaluated designs, which have been normalized with respect to those obtained for the static CMOS design. In Fig. 5a results are provided for $f_1$. It can be observed that power consumptions of the dynamic CMOS solutions are always larger than those of the static CMOS. Power savings of TFET circuits regarding their

<table>
<thead>
<tr>
<th>$LD$</th>
<th>$f_1$ (GHz)</th>
<th>$f_2$ (GHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>2.0</td>
<td>1.6</td>
</tr>
<tr>
<td>8</td>
<td>2.0</td>
<td>1.6</td>
</tr>
<tr>
<td>10</td>
<td>2.0</td>
<td>1.6</td>
</tr>
<tr>
<td>12</td>
<td>2.0</td>
<td>1.6</td>
</tr>
</tbody>
</table>

TABLE I. MAXIMUM OPERATING FREQUENCIES AND MINIMUM SUPPLY VOLTAGES FOR DIFFERENT $LD$ VALUES.
CMOS counterpart are evident. The static TFET PSU implementations exhibits power savings between 60% and 66% with respect to static CMOS. The dynamic TFET PSU designs, unlike CMOS, show power advantageous with respect to its static counterpart. They consume around half the power of the TFET PSU static solutions and around 20% of the power of the CMOS static implementations. Also note that TFET PSD and TFET PSU curves are always very close to each other but for $LD=6$, where $V_{DD}$ for TFET PSD is substantially larger than for the other values of $LD$.

Results at $f_2$ are depicted in Fig. 5b, where the most evident difference with respect to Fig. 5a is that power advantages of dynamic TFET PSD circuits have been significantly reduced, with power savings between 0% to 40%, since their supply voltages have been multiplied between 2 and 2.5 regarding the $f_1$ scenario. On the contrary, this does not happen for TFET PSU because its $V_{DD}$ has been scaled with a similar factor to the static CMOS design. Finally, although the dynamic CMOS design is closer to the static, it is not either competitive.

In all but one case, dynamic TFET implementations working at the maximum frequency exhibited by static CMOS at its nominal supply voltage ($f_1$) have been designed. These dynamic based solutions are competitive in terms of power with significant savings in the case of the PSU technology. Contrary, dynamic CMOS solutions do not reduce power since the power savings associated to $V_{DD}$ reductions do not compensate overheads due to clocks and to the precharge-evaluation operating principle of this logic style.

\section{V. Conclusions}

Simply comparing architecturally identical TFET and CMOS implementations can lead to misleading conclusions concerning the potential of the former to overcome CMOS limitations. It has been illustrated how properly tuning the logic architecture results in raising the frequency up to which TFETs are competitive or in increasing power savings at lower frequencies. In particular, preliminary results show that ultra fine-grained pipeline logic architectures, with their aggressive reduced logic depth, permits taking advantage of the good performance of TFETs at low supply voltage. In spite of power overheads associated to the dynamic logic style used, these TFETs implementations, unlike CMOS, are competitive in terms of power. Although evaluation has been carried at the schematic level, and further work to incorporate more realistic layout parasitic is required, the large advantages obtained allows to be optimistic concerning the potential of tunnel devices not only for low performance applications.

\section{Acknowledgment}

This work has been funded by Ministerio de Economía y Competitividad del Gobierno de España with support from FEDER under Project TEC2013-40670-P.

\section{References}