Evaluation of MOBILE-based gate-level pipelining augmenting CMOS with RTDs

Juan Núñez, María J. Avedillo, José M. Quintana
IMSE-CNMC-SIC and University of Seville. Av. Américo Vespucio s/n, 41092, Seville (Spain)

ABSTRACT
The incorporation of Resonant Tunnel Diodes (RTDs) into III/V transistor technologies has shown an improved circuit performance: higher circuit speed, reduced component count, and/or lowered power consumption. Currently, the incorporation of these devices into CMOS technologies (RTD-CMOS) is an area of active research. Although some works have focused the evaluation of the advantages of this incorporation, additional work in this direction is required. We compare RTD-CMOS and pure CMOS realizations of a network of logic gates which can be operated in a gate-level pipeline. Significant lower average power is obtained for RTD-CMOS implementations.

Keywords: Resonant tunneling diode, Nano-pipeline, Emerging technologies, Logic circuits, Power efficiency.

1. INTRODUCTION
Resonant tunneling devices (RTDs) are nowadays considered the most mature type of quantum-effect devices. They are already operating at room temperature and they exhibit very attractive characteristics as high-speed operation and low power consumption. RTDs are very fast non linear circuit elements which have been integrated with transistors to create novel quantum devices and circuits. This incorporation of tunnel diodes into III/V transistor technologies has shown an improved circuit performance: higher circuit speed, reduced component count, and/or lowered power consumption [1],[2],[3].

The degree of development of resonant tunneling devices is very different. RTDs fabricated in III-V are, undoubtedly, the most mature and most reported circuits based on resonant tunneling use them combined with different types of transistors. Since the currently dominant technologies use silicon, plenty of efforts are being devoted to develop devices with negative resistance in this material. The resulting diodes have provided worse performance than those achieved in III-V technologies. Currently, the realization of tunnel diodes in silicon is a very active research area where progresses are expected. In fact, it has been suggested that the addition of RTDs to CMOS technology could extend its life and even make investments rentable. It has been shown the integration of Resonant Interband Tunneling Diodes (RITDs) with standard CMOS[4] and SiGe HBT[5]. It has been also reported a RITD with a cutoff frequency of 20GHz, allowing for the first time, applications of mixed signal, RF and high speed logic circuits[6]. Simpler and compatible with CMOS process to fabricate Tunneling Diodes have been recently reported: structures that do not need Ge[7], and a fabrication process based on CVD (Chemical Vapor Deposition) instead of MBE[8] (Molecular Beam Epitaxy) is presented. Another explored option is the development of procedures for III-V RTDs compatible with silicon substrates. Recently importance advances have been achieved in this area, such us the Tunneling Diodes in III-V materials and Ge using ART (Aspect Ratio Trapping)[9].

Some works have focused the evaluation of the advantages of incorporating RTDs to CMOS technologies. The keeper transistor of the domino logic gates is replaced by an RTD[10], which significantly increases the noise immunity without affecting the area, delay and power consumption. A static memory cell[11] consisting of the incorporation of a well-known DRAM cell topology with a pair of RTDs has been reported. This structure improves the performance of a typical 6-transistors SRAM cell in terms of the static power consumption in three orders of magnitude.

However, in our opinion, additional work in this direction is required. In particular, in the field of logic circuits, estimations of performance/count devices improvements obtained through the addition of RTDs have been evaluated for a set of logic functionalities (combinational gates and flip-flops)[12], but without taking into account their usage in gate networks. This is a key point, because as it will be explained in next section, RTD logic gates allow the implementation of a pipeline at the gate level. That is, each gate is a pipeline stage and thus it should be compare to CMOS logic styles
operating in a similar way. Moreover, up to our knowledge there is a lack of recent studies in this area and as a consequence no data involving current technologies are available. We claim that on the basis of the mentioned edge-triggered behaviour of MOBILE gates a simpler single phase scheme without latches is possible. This paper addresses this issue and contributes to provide results on how RTD-CMOS realizations compare to pure CMOS gate-level pipelining.

This paper is organized as follows: in Section 2, RTD logic networks based on the MOnostable to BIstable operation principle are described. In Section 3, we present the RTD-CMOS network which has been evaluated. A brief description of the experiment is also given. In Section 4, a comparison in terms of the average power consumption with pure CMOS realizations of these structures is described. Finally, some key conclusions are given in Section 5.

2. RTD-BASED MOBILE LOGIC GATES

Logic circuit applications of RTDs are mainly based on the MOnostable-BIstable Logic Element (MOBILE) which exploits the negative differential resistance of their $I-V$ characteristic (Figure 1a). The MOBILE (Figure 1b) is a rising edge-triggered current controlled gate which consists of two RTDs connected in series and driven by a switching bias voltage, $V_{CK}$. When $V_{CK}$ is low, both RTDs are in the on-state (or low resistance state) and the circuit is monostable. Increasing to an appropriate maximum value ensures that only the device with the lowest peak current switches (quenches) from the on-state to the off-state (the high resistance state). Output is high if the driver RTD is the one which switches and it is low if the load does. Logic functionality can be achieved if the peak current of one of the RTDs is controlled by an input.

In the configuration of the rising (falling) edge-triggered inverter MOBILE shown in Figure 1c (Figure 1d), the peak current of the driver (load) RTD can be modulated using the external input signal. RTD peak currents are selected such that the value of the output depends on whether the external input signal is “1” or “0”. Assuming the same peak current density, $j_p$, for all the RTDs, the peak current is proportional to their area. Figures depict required area relationships.

For $V_{CK}$ high (low), the output node of rising (falling) edge-triggered MOBILE maintains its value even if the input changes. That is, this circuit structure is self-latching allowing to implement pipeline at the gate level. In other words, network operation speed is independent of logic depth but is determined by the clock frequency at which single gates can be operated. Cascaded rising edge-triggered MOBILE gates operated in a pipelined fashion use a four-phase clocking scheme. It has been demonstrated that a network of MOBILE-based gates can be operated with a single clocked bias signal [13]. To achieve this operation, rising edge-triggered gates and falling edge-triggered gates are alternated and latches are added at each stage to remove the return-to-reset behavior.

However, we have realized that it is not necessary to remove the return-to-reset behavior to ensure the correct operation of a MOBILE network of gates which alternates rising and falling edge-triggered stages. It is enough to keep the output

Figure 1. RTD MOBILE circuits. (a) RTD $I$-$V$ characteristic and symbol. (b) Basic MOBILE. (c) Rising edge-triggered MOBILE inverter. (d) Falling edge-triggered MOBILE inverter.
of each MOBILE stage until it has been evaluated by the next one. Thus, each latch is replaced by a static inverter. A detailed analysis of the operation of this architecture shows that the ‘non-return-to-reset’ is not necessary to ensure a correct operation. Moreover, it is enough to maintain the output of each MOBILE stage until the next one has evaluated. Thus, the latches reported in [13], which exhibit a large static consumption that limits their practical usage, can be substituted by a simpler circuit. Moreover, simulations of alternating rising and falling edge-triggered MOBILE gates without inter-stage elements show correct operation. This is explained because the decision on which output the MOBILE will give is taken when \( V_{CK} \) is approximately equal to \( 2V_P \). For this value of the clock voltage, the output of the previous MOBILE stage has not reached the reset value yet and, thus, it can be properly evaluated. In spite of this, since the MOBILE operating principle is very sensitive to load, an inter-stage element is advantageous to increase robustness and to ease design. This element takes care of fan-out and isolates MOBILE gates. Figure 2a depicts the proposed architecture.

Figure 2b shows HSPICE simulated waveforms corresponding to an interconnection of rising and falling edge-triggered MOBILE inverters with static inverters as inter-stage elements. Evaluation problems could occur for those cases in which the output of the previous MOBILE stage differs from its corresponding reset value. In these situations, the active edge of the clock signal forces a change in the output of previous stage which is, at the same time, being evaluated. In the shown waveforms, this happens for the falling edge-triggered MOBILE (the one with \( V_{OUT,2} \) as input and \( V_{OUT,3} \) as output) in the marked clock transition. Note that the right input value (zero) is taken in spite of \( V_{OUT,1} \) being reset to zero (and so \( V_{OUT,2} \) reset to one) by that transition. Note that in addition to the already mentioned advantages of including an inter-stage element, the small delay introduced by the inverter favors correct operation.

The operation of the proposed interconnection scheme of MOBILE gates has been experimentally validated. Up to our knowledge it is the first time a working single phase MOBILE network is reported. A three-stage chain of MOBILE inverters have been fabricated. The first and the third follower are falling edge-triggered, whereas the second one is rising edge-triggered. They have been implemented with MOS-NDR devices (circuit made up of transistors that emulate the RTD \( I-V \) characteristic) and the MOBILE gate topology from [14] in a 1.2V-130nm CMOS commercial technology.

Figure 3 depicts experimental waveforms in which the clock (a sinusoidal clock of 100MHz) and the input have been captured using an oscilloscope. They have been represented along with the digital version of the output in the logic.
analyzer Agilent 16902B. Note that the same sequence of the input is observed on the output as expected (there are six inverting stages in total since we are using inverters after each MOBILE inverter), with a delay between them, corresponding to three semi periods of the clock signal. This delay is associated to the consecutive evaluation of the three MOBILE stages. Moreover, the return to the reset value is observed. In order to validate circuit operation for faster clock transitions, and since increasing the frequency of the sinusoidal clock was not possible due to dynamic limitations of the pads, we have used a pulse generator.

3. **STUDY DESCRIPTION**

The power performance of a network of inverters implemented with RTDs and commercial CMOS transistors has been evaluated. We have compared it with True Single Phase Clock (TSPC) CMOS realizations, since they also implement gate-level pipeline. Variations of TSPC have been proposed and widely applied in the design of high speed applications.

The study uses transistors from a standard 130nm CMOS process. For the RTDs, we have started the experiments using a model from LOCOM\(^{[15]}\). This model corresponds to a III-V RTD and has been experimentally validated (\(j_P=21\text{KA/cm}^2\), \(V_P=0.21\text{V}, C=4\text{fF/\mu m}^2\)). Sizing of networks to be compared has been carried out through HSPICE\(^{\circledR}\) parametric simulations. 4-stage chains of inverters with fan-out 1 (LOAD1), 2 (LOAD2) and 3 (LOAD3) at each stage have been simulated at three normalized frequencies, \(f_{\text{NORM}} \{0.16, 0.20, 0.24\}\). This frequency range has been selected on the basis of reported pipelined circuits in similar CMOS technologies. For each architecture, we have varied parameters taking a discrete number of samples of each one in a given range. Among the simulated circuits, the one such that its Monte-Carlo simulation (modeling both transistor and RTD intra-chip variations) shows correct operation and minimizes average power has been selected. Parameters included in the design space exploration for each logic style (RTD-CMOS and TSPC) as well as simulation conditions are described below.

3.1 **RTD-based circuits**

Supply voltage has been explored in the range from 0.6V to 0.8V. Transistor lengths and widths have been set to the minimum values associated to the technology (\(L_{\text{MIN}}=0.12\mu\text{m}\) and \(W_{\text{MIN}}=0.16\mu\text{m}\)). PMOS transistor width is \(K\) times the NMOS one (\(K=3.5\) for this technology). We have varied the RTD areas \(f_{X,R}\) and \(f_{X,F}\) assuming that: \(f_{D,R}=f_{X,R}\) and \(f_{L,R}=1.5f_{X,R}\) for the rising edge-triggered inverter and \(f_{L,F}=f_{X,F}\) and \(f_{D,F}=1.5f_{X,F}\) for the falling edge-triggered inverter. \(f_{X,R}\) and \(f_{X,F}\) have been varied from 0.04\(\mu\text{m}^2\) to 0.4\(\mu\text{m}^2\) (This lower limit has been fixed assuming that the minimum RTD that could be fabricated would be \(0.2\mu\text{m} \times 0.2\mu\text{m} (0.04\mu\text{m}^2)\) to be in agreement to the technology node we are using for transistors.

![Figure 3. Experimental results of the fabricated MOS-NDR MOBILE chain.](http://proceedings.spiedigitallibrary.org/)

\(^{*}\) \(f_{\text{NORM}} = f_{\text{FO4}}\), where \(FO4\) is the FO4-inverter delay of the technology.
3.2 TSPC circuits

For the TSPC network, $V_{DD}$ has been varied by taking seven equispaced points from 0.6V to 1.4V. Transistor lengths have been fixed to the minimum. A typical CMOS sizing scheme has been assumed. A design parameter $W$ has been defined which corresponds to the width of a basic NMOS transistor. When $m$ transistors are connected in series their widths are multiply by $m$. $W$ has been varied from 0.16µm to 1.6µm.

3.3 Simulation Setup

Ideal clock waveforms for each structure have been applied. For the RTD-based circuits, we have considered a clock in which the rising, falling, hold and reset times are the same. In TSPC structures a pulse train clock with a duty cycle of 50% has been used. Standard mismatch models from the technology have been used with the MOS transistors. Since there are no mismatch models available for the RTDs, Gaussian distributions (relative error of $\pm 10\%$) have been associated to the peak voltage, intrinsic capacitance and the peak current density of each device. Variations of the supply voltage ($\pm 10\%$) around its nominal value have been considered.

4. EVALUATION

Circuits derived by the above described design exploration experiment have been evaluated and compared. We have measured the average power consumption at different frequencies ($f_{\text{NORM}} 0.16, 0.20$ and $0.24$) and with distinct load conditions (fan out 1, 2 and 3 for each stage). Figure 4a depicts simulation results for RTD-CMOS networks using two models for the RTDs: the already mentioned LOCOM RTD ($j_p=21\text{KA/cm}^2$, $C=4\text{fF/}\mu\text{m}^2$) and a model exhibiting the characteristics of a silicon RITD reported in [16] ($j_p=218\text{KA/cm}^2$, $C=6\text{fF/}\mu\text{m}^2$). Significant improvement of the performance of the RTD-CMOS networks is obtained while increasing the operation frequency. At $f_{\text{NORM}}=0.24$, the average power for the RTD-CMOS networks is 26% (RTD LOCOM) and 52% (RITD on Si) of the measured for the

Figure 4. Comparison between average power consumptions of the RTD-CMOS and the TSPC designs.
TSPC architecture. Note the small difference between the average powers of both RTD-CMOS chains in spite of the fact that their respective $j_p$ values differ in one order of magnitude. This is explained since the network with the silicon RITD (derived by the sizing experiment) has RTD areas smaller than the LOCOM network to operate at the same speed. This is in agreement with previous works on MOBILE operation speed\cite{17}.

For load 2 (Figure 4b), at $f_{\text{NORM}}=0.24$, power consumption of the RTD-CMOS networks is reduced up to 13% (RTD LOCOM) of the TSPC. For load 3 (Figure 4c), at $f_{\text{NORM}}=0.24$, measured average power is 10% of the TSPC. The RTD-CMOS designs have been also validated using a buffer instead of a CMOS inverter as inter-stage element which allows greater logic flexibility.

5. CONCLUSIONS

Realizations of RTD-CMOS logic networks working on the basis of the MOBILE operating principle have been introduced, which simplifies the inter-stage element reported in previous single-phase architectures and translates in power, area and clock load advantages. A comparison to only transistor implementations using TSPC logic style, well suitable for gate-level pipeline, like the proposed RTD-CMOS, has been carried out. Operation of RTD-CMOS realizations is static, overcoming one of the main drawbacks of TSPC CMOS. Very significant power savings have been obtained for the RTD-CMOS at the target frequencies, which compares favorably with other TSPC-based logic styles.

6. ACKNOWLEDGEMENT

This work has been funded in part by the Spanish Ministry of Science and Innovation with support from ERDF under Projects TEC2007-67245 and TEC2010-18937 and in part by the Consejería de Innovación, Ciencia y Empresa, Junta de Andalucía under Project TIC-2961.

REFERENCES


