Low Threshold CMOS Circuits with Low Standby Current

Mircea R. Stan
University of Virginia, Electrical Engineering Department
Charlottesville, VA 22903
mircea@virginia.edu

Abstract

Multi-Voltage CMOS (MVCMSOS) is a design methodology for very low power supply voltages that uses low-threshold transistors in series with the supply rails. The control voltage on the gating transistors need to be outside of the low $V_{th}$ range (hence the name MVCMSOS) in order to reduce the standby current, but the resulting circuits operate at lower supply voltages and have a lower area overhead than the previously proposed Multi-Threshold CMOS (MTCMOS).

1 Introduction

Low-power design techniques try to reduce the power dissipation of high-performance systems (which impacts their packaging and heat removal costs) or the power consumption of portable equipment (which directly relates to size and battery life). Lowering the power supply voltage is one of the preferred methods as it has a quadratic effect on reducing power in CMOS [3] but this unfortunately increases the gate delay. Reducing $V_{th}$ of the transistors in order to preserve performance presents no problem in theory and has been reported [2, 6], but obtaining such small $V_{th}$ values with small variance remains a manufacturing challenge.

Even if the threshold can be accurately controlled, a big problem for low-power design is the exponential increase in the “off” current with reduced $V_{th}$. For circuits that operate continuously it has been shown that an optimal power-delay product can be obtained by reducing $V_{th}$ such that the DC power consumption becomes approximately equal to the dynamic power [2]. For event-driven applications though, which are idle for long periods of time, the large current in sleep mode becomes unacceptable and negatively impacts the overall power budget. Multi-threshold CMOS (MTCMOS) [7, 8] has been proposed as a general technique for reducing the standby current by using two types of transistors, low-threshold transistors for fast operation in normal mode and high-threshold transistors for reducing the sleep mode “off” current. The high-threshold transistors are placed as gating transistors in series with the supply, creating virtual power rails for the rest of the circuit being implemented with low-threshold transistors [7]. The high-threshold gating transistors are always “on” for normal operation and, if sized correctly, don’t have a significant impact on the circuit speed. With the circuit in sleep mode, the gating transistors are turned “off” and the overall leakage is limited by their small subthreshold current.

Figure 1: MVCMSOS inverters: A. with virtual power and ground, B. only with virtual ground, C. only with virtual power. Only low threshold transistors are used but the SLEEP signals need to be outside the $V_{ss} - V_{dd}$ range.

The rest of the paper shows how a straightforward circuit technique similar to MTCMOS but using only low-threshold transistors has better performance and less area by using control voltages larger than $V_{dd}$ for the PMOS and smaller than $V_{ss}$ for the NMOS. The advantages of MVCMSOS can be summarized as:

- For MVCMSOS the supply voltage can get much smaller than for MTCMOS due to the absence of high-threshold transistors, correct operation at 0.5V and lower being possible. The same effect could be obtained for MTCMOS by using a separate network driving the gating transistors with a separate (larger) $V_{dd}$.

- Because the gating transistors for MVCMSOS have low $V_{th}$ their size can be smaller than the high $V_{th}$ MTCMOS transistors for the same current drive. Furthermore, the MVCMSOS gating transistors can be distributed which results in a very efficient layout with minimal area overhead by using diffusion sharing. A similar diffusion sharing may be more difficult for MTCMOS depending on how well the implant registration mask can be positioned between two transistors.
The classical substrate pump used in memory design is the sleep voltage can be either very slow to avoid latch up. A possible on-chip circuit is the classical substrate pump used in memory design. The sleep voltage can be either very slow to avoid latch up. A possible on-chip circuit is generated on-chip or on-chip and the edges of the signal should be very slow to avoid latch up. A possible on-chip circuit is the classical substrate pump used in memory design.

The simulations in this paper use HSPICE with BSIM3v3 models for nine-stage ring oscillators in a 0.35μm CMOS technology. The models for high and low threshold transistors differ only in the Vth parameter in the model which is 0.332V and 0.232V for high and low threshold NMOS and −0.6734 and −0.2733 for high and low threshold PMOS, respectively. The PMOS is sized three times larger than the NMOS (1.8μ vs. 0.6μ) for symmetric noise margins, but probably not optimal for low power.

### 3 MTCMOS Sub-Volt Operation

The major advantage of MVC莫斯 is that it can operate at very low supply voltages. Figure 2 shows the output of regular high threshold, low threshold, MTCMOS and MVC莫斯 ring oscillators (using only NMOS gating transistors nine times minimum size) at three different Vdd values, 1V, 0.7V and 0.4V. The regular circuit with high threshold transistors ceases to function at 0.7V and MTCMOS ceases to function at 0.4V, but MVC莫斯 is just slightly slower than the regular low threshold circuit even at 0.4V. The delays of MTCMOS, MVC莫斯 and regular low threshold oscillators (divide by 18 for inverter delay) as a function of the power supply voltage are shown in figure 3.

MVC莫斯 in sleep mode requires control voltages lower than Vdd for the gating NMOS and higher than Vdd for the gating PMOS. Fortunately there is no hot-electrons problem since we only apply the larger voltages to the gates of the gating transistors. Latchup should not be a problem either if we can make sure that the rise and fall times for the control voltages are slow enough to reduce capacitive coupling. We have simulated the MTCMOS and MVCMOS ring oscillators with only NMOS gating transistors in sleep mode (see figure 4) and there is no reason to reduce the negative control voltage to less than −0.5V because at some point the subthreshold current becomes negligible and the leakage of the reverse biased drain junctions becomes dominant. This should be perfectly safe with no oxide breakdown for all current technologies. The sleep voltage can be either generated off-chip or on-chip and the edges of the signal should be very slow to avoid latchup. A possible on-chip circuit is the classical substrate pump used in memory design.

### 4 Sizing the Gating Transistors

The “virtual” power rails for MVCMOS (as for MTCMOS) have a much higher impedance than the true power rails and will unavoidably “bounce” which means that:

- The speed of the logic gates is reduced as the effective supply voltage is reduced by the bounce magnitude.
- The noise margins are reduced which becomes a problem especially when communicating with gates with different “virtual” rails.

Figure 5 shows the ring oscillator delay at Vdd = 0.5V for MVC莫斯 with both PMOS and NMOS, only NMOS or only PMOS gating transistors. Using both transistors is the slowest solution and, although the PMOS is sized three times larger than the NMOS, the circuit using only a gating NMOS is the fastest. Using just a gating NMOS has the lowest area overhead and is also the fastest but has a virtual ground which will bounce. Designers generally prefer having a noisy supply and maintain a “clean” ground which can be obtained using a gating PMOS even if it is a slower solution. A tool was proposed for sizing MTCMOS gating transistors for speed [3], but we believe that a more critical issue is sizing for noise margins. Figure 6 shows the variation of ground bounce with the size of the sleep transistors. The ground bounce is a noise problem because many logic gates share a centralized gating transistor, hence the same virtual ground, which is a characteristic carried over from the MTCMOS design style.

Instead of centralized gating transistors there is also the option of using distributed gating transistors, one for each logic gate. Such a distributed solution will completely eliminate the ground bounce noise problem but may result in a slower circuit as can be seen in figure 7. The layout with distributed sleep transistors can be done very efficiently using diffusion sharing with the logic transistors. In this case the layout for the pull-down network (PDN) for a gated inverter looks exactly like the PDN of a standard CMOS NAND gate and represents a small area increase. Distributing the gating transistors with every gate does not incur a large penalty which suggests the following design strategy:

- All the gates that are not on the critical path should have distributed gating transistors in order to eliminate any noise problems due to ground bounce.
Figure 4: Subthreshold current for MTCMOS with SLEEP at 0V and MVCMOS with SLEEP between 0V and −0.5V.

Figure 5: MVCMOS delay with one or two gating transistors at V_{dd} = 0.5V: A. both, B. PMOS only, C. NMOS only.

- Only gates that are critical should use common gating transistors for faster operation. A thorough analysis is required for noise and performance issues but since the number of transistors is small this should be easier.

- Gates that switch at the same time do not benefit (speed-wise) from sharing gating transistors, gates that switch at different moments do.

5 Conclusion

We have presented Multi-Voltage CMOS (MVCMOS), a design methodology for very low supply voltages with good performance and low leakage during sleep. The area overhead is smaller than for MTCMOS and the ground bounce problems can be eased by distributing the low threshold gating transistors. As suggested by a reviewer the gating transistors could be overdriven both "on" and "off" which requires even less width for the same current and reduces even more the area penalty for gating. More work is needed for analyzing dynamic and sequential MVCMOS circuits.

Acknowledgements

This work was supported in part by NSF CAREER grant MIP-9733140. I thank Sreekanth Nallagatla for helping with some of the initial simulations.

References