A gate sizing and transistor fingering strategy for subthreshold CMOS circuits

Morteza Nabavi^{a)} and Maitham Shams^{b)}

Department of Electronics, Carleton University,

1125 Colonel By Drive, Ottawa, Ontario K1S 5B6, Canada

a) nabavi@doe.carleton.ca

b) shams@doe.carleton.ca

Abstract: Parallel Transistor Stacks (PTS) has been shown to be an effective technique for improving the speed of digital circuits operating in the subthreshold region which comes at the cost of power consumption and area. However, our experience shows that using PTS is not beneficial in all cases. In this paper, we present a methodology to identify whether using PTS is beneficial (or not) in a particular CMOS technology and what transistor sizing can be employed to maximize the circuit speed. Our technique is based on analyzing the Current-Over-Capacitance (COC) ratio of PMOS and NMOS transistors. The results of incorporating the proposed methodology in a 4-bit comparator and a 19-stage inverter ring oscillator, using 90 nm CMOS technology, illustrate 26% and 40% extra improvement compared to the blind use of PTS, respectively.

Keywords: VLSI, CMOS, logic design, subthreshold circuits

Classification: Integrated circuits

References

- [1] B. H. Calhoun, A. Wang, and A. P. Chandrakasan, "Modeling and Sizing for Minimum Energy Operation in Sub-threshold Circuits," *IEEE J. Solid-State Circuits*, vol. 40, no. 9, pp. 1778–1786, Sept. 2005.
- [2] M. Muker and M. Shams, "Designing digital subthreshold CMOS circuits using parallel transistor stacks," *IET Electronics Letter*, vol. 47, no. 6, pp. 372–374, March 2011. Also see the featured interviews on page 354 in the same issue.
- [3] J. M. Rabaey, A. Chandrakasn, and B. Nikolic, *Digital Integrated Circuits: A Design Perspective*, 2nd ed., Pearson Education, Toronto, 2003.

1 Introduction

Digital circuits operating in the subthreshold region benefit from very low power consumption at the cost of speed [1]. In order to improve the speed of subthreshold circuits, Muker and Shams [2] introduced the Parallel Transistor Stacks (PTS) structure that, for example, enhances the speed of a 9-stage

ring oscillator by 2.63 times compared to the conventional transistor sizing optimization method at a supply voltage of 0.3 V in a 65 nm LP CMOS technology. However, there are CMOS technologies in which using PTS blindly may not improve the speed of a subthreshold circuit. In this paper, we describe a methodology to identify the case when using PTS is beneficial for improving the speed of a digital circuit operating in the subthreshold region. We also elaborate on obtaining the appropriate transistor sizing where applying PTS is advised.

2 Propagation delay

The propagation delay (D) for charging or discharging a node is given by [3]:

$$D = \frac{CV_{DD}}{2I}, \quad (1)$$

where C is the load capacitance, I denotes the average current, and V_{DD} is the supply voltage. As (1) indicates, higher speeds (i.e., lower propagation delays) can be achieved by increasing the current and lowering the capacitance. Therefore, the highest operating frequency is attained when the Current-Over-Capacitance (COC) ratio is maximized. This is often exactly or close to the transistor's minimum width in modern nanometer CMOS technologies, as a result of the so-called Inverse-Narrow-Width-Effect (INWE) [2].

Fig. 1 shows the COC ratio versus the width of both PMOS and NMOS transistors in a 130 nm CMOS technology. The transistors are biased for maximum current drive in the subthreshold region (i.e., $V_{DS} = |V_{GS}|$). The capacitance includes the total gate and junction capacitance of the transistor. Based on this figure, in order to achieve a higher COC ratio we either use the minimum transistor width (point A) or use PTS locked to the minimum width. Nevertheless, the maximum COC ratio is not always associated with the minimum transistor width, as shown in Fig. 1 (b). The COC ratio for the PMOS transistor shown in this figure consists of two regions. In Region 1

Fig. 1. Current over capacitance ratio as a function of width for transistors in a 130 nm CMOS technology: a) NMOS, b) PMOS.

the COC ratio corresponding to the widths larger than the minimum width have a lower value compared to the COC ratio at the minimum width. On the other hand, the widths located in Region 2 have a COC ratio larger than the minimum width. Based on this figure, we use PTS for transistor widths in Region 1 and avoid PTS for transistor widths in Region 2. Applying PTS on the transistors in Region 2 will result in a lower circuit speed.

3 Methodology

The following steps explain our methodology on how to identify the cases when to use and when not to use PTS in designing a logic circuit.

- 1- Plot the COC ratio versus width for each transistor type (PMOS and NMOS) for the given technology kit (e.g., Fig. 1).
- 2- Based on Step 1, identify the transistor width (W_{INWE}) associated with the local maximum COC ratio, which typically appears near the minimum width due to INWE [2]. As shown in Fig. 1, point A denotes W_{INWE} for NMOS transistor and point C denotes W_{INWE} for PMOS transistor.
- 3- Apply a transistor sizing optimization method such as Logical Effort [3] to optimize the circuit performance.
- 4- Find the optimum width (W_{opt}) for each transistor in the circuit based on Step 3.
- 5- Compare COC_{opt} , the COC ratio corresponding to W_{opt} (Step 4), with COC_{INWE} , the COC ratio corresponding to W_{INWE} (Step 2).
 - a. If $COC_{opt} < COC_{INWE}$, use PTS, and split the transistor of width W_{opt} into multiple (K) transistors of width W_{INWE} , where $K = W_{opt}/W_{INWE}$ rounded to the nearest digit. For example, if Logical Effort finds that an NMOS transistor should have a width of $1.6\mu\text{m}$ (point B in Fig. 1(a)) we split it into $K = W_{opt}/W_{INWE} = 10$ transistors at the width of 160 nm (point A in Fig. 1(a)).
 - b. Else If $COC_{opt} \geq COC_{INWE}$, using PTS results in a slower circuit. In this case we keep the transistor width at W_{opt} . For example, if Logical Effort finds a transistor of width $8\mu\text{m}$ (point E in Fig. 1(b)), we don't use PTS, as the COC ratio at point E has already a higher value compared to the COC ratio at 160 nm (point C in Fig. 1(b)).

Note that Logical Effort (LE) is a much simpler optimization method compared to exhaustive blind simulations. Although LE may not result in the exact optimum solution, especially after extracting layout parasitic, it gives a very good initial point to start with. Iterative applications of LE may be

needed for fine tuning. However, LE assumes that the current of a transistor is linearly proportional to its width. This assumption is not valid for MOSFETs operating in the subthreshold region in most modern CMOS technologies due to the INWE [2]. By using PTS, a linear-relationship between the current of the transistor and its width is established, enabling one to use LE for transistor sizing of subthreshold circuits [2].

4 Results

A 19-stage inverter ring oscillator running at 0.2 V has been examined to verify the correctness of the proposed methodology in 130 nm and 90 nm CMOS technologies. We explored the effect of incorporating PTS on each inverter as shown in Fig. 2 (b-d). The results are shown in Table I.

Considering a standard inverter shown in Fig. 2 (a), the simulation results for the 130 nm technology shown in Table I reveals that applying PTS on only PMOS transistors (column *P-PTS*) of each inverter in the ring oscillator reduces the frequency of oscillation by 3%. On the other hand, applying PTS on only NMOS transistors (column *N-PTS*) improves the frequency by 120%, while applying PTS on both NMOS and PMOS transistors increases the frequency only by 100%. This result is expected as the COC ratio corresponding to the PMOS transistor of 8 μm width (point E in Fig. 1 (b)) is more than the COC ratio compared to 160 nm (point C in Fig. 1 (b)), and the NMOS transistor at the width of 4 μm (point D in Fig. 1 (a)) has much lower COC ratio compared to point A. Therefore, in this case, the optimum sizing

Fig. 2. Inverter configurations for a 19-inverter ring oscillator: (a) Standard, (b) PMOS PTS, (c) NMOS PTS, (d) PMOS and NMOS PTS.

Table I. Frequency, Delay, and Power of 19-stage ring oscillator and a 4-bit comparator in 130 and 90 nm technology with the standard, only PMOS PTS (P-PTS), only NMOS PTS (N-PTS), both NMOS and PMOS PTS (PN-PTS), and PTS based on the proposed methodology (OPT-PTS).

Technology	Metrics	Standard	P-PTS	N-PTS	PN-PTS	OPT-PTS
Ring Oscillator						
130 nm	Frequency (KHz)	364.8	357.2	806.1	753.9	806.1
	Power (nW)	10.3	9.821	23.45	21.73	23.45
90 nm	Frequency (MHz)	6.4	5.3	11.6	9.15	11.6
	Power (nW)	165.9	138.4	327.6	269.5	327.6
4-bit Comparator						
130 nm	Delay(ns)	908.3	862.0	573.5	548.3	531.0
	Power (pW)	772	768.6	1401	1403	1390
90 nm	Delay(ns)	56.0	74.7	36.2	48.6	34.2
	Power (nW)	9.76	8.32	17.12	15.2	16.79

(column *OPT-PTS*) is achieved by applying PTS on NMOS transistors, and avoiding applying PTS on PMOS transistors. Note that by applying PTS the power consumed by the ring oscillator increases by 127%.

In the 90 nm technology, the COC ratio for both transistor types follow the same behavior as that of the 130 nm technology. Simulation illustrates that the corresponding COC ratio of the NMOS transistor versus width has a sharp decay, whereas the COC ratio corresponding to the PMOS transistor increases after a certain width (1 μm). The simulation results on a 19-stage ring oscillator shown in Table I for the 90 nm confirm that applying PTS on all PMOS transistors of 8 μm width results in 16% lower frequency compared to that of the standard sizing. On the other hand, applying PTS on only NMOS transistors improves the speed almost by 81%. Applying PTS on both PMOS and NMOS transistors increases the speed only by 41% compared to the standard case. Again here, for achieving the maximum speed we advise applying PTS just on NMOS transistors and not PMOS transistors. As shown in Table I, the power consumed by the ring oscillator increases compared to the standard case from 165.9 nW to 327.6 nW.

We also applied our methodology on implementing a 4-bit comparator. The simulation results after Logical Effort optimization are shown in Table I (column *Standard*).

By running Logical Effort on this complex circuit, it suggests PMOS transistors with widths in both regions (Fig. 1 (b)). According to our methodology, for NMOS transistors, as all suggested widths have lower COC ratio than the minimum width, we apply PTS on all NMOS transistors in both CMOS technologies. Applying PTS on PMOS transistors depends on whether the Logical Effort locates the width sizes in Region 1 or Region 2. Regardless the fact that PTS should be applied on certain transistors, in Table I we list the simulation results of applying PTS only on all PMOS and only on all NMOS transistors in the *P-PTS*, and *N-PTS* columns, respectively. Column *PN-PTS* lists the simulation results for applying PTS on both NMOS and PMOS transistors. The last column, *OPT-PTS*, illustrates the results of optimum sizing (maximum speed) based on our methodology. This column shows 41% increase of speed compared to the standard sizing, which is 3% more than applying PTS on all transistors blindly. The reason that our methodology in this case doesn't show considerable improvement compared to the blind PTS is that most of the PMOS transistors are in Region 1.

The result of applying the proposed methodology, i.e. selective application of PTS, on the 4-bit comparator in the 90 nm technology is also shown in *OPT-PTS* column of Table I. The speed is increased by around 39% compared to standard sizing, while applying blind PTS only improves the speed by 13%. Note that applying PTS only on all PMOS transistors (*P-PTS*) degrades the speed by 33%, while applying PTS only on all NMOS transistors (*N-PTS*) improves the speed by 35% compared to the standard sizing.

The power consumed by the 4-bit comparator in both 130 nm and 90 nm technologies shows 80% and 72% increase when PTS is applied compared to the standard case.

5 Conclusion

Although using parallel transistor stacks (PTS) gives rise to considerable speed benefit in the subthreshold circuits, there are cases where PTS should be avoided. In this paper we proposed a methodology to advise selective application of PTS on some transistors in a circuit. Applying the proposed methodology on different circuits illustrates up to 40% increase in speed compared to when PTS is applied on all transistors blindly. Incorporating our methodology shows 38% to 121% speed improvement compared to the conventional transistor sizing optimization method. Besides, it is important to note that using PTS increases the circuit area up to 50% and according to our simulations the power consumption is linearly proportional to the speed.