Scaling Limit of CMOS Supply Voltage from Noise Margin Considerations

Minjian Liu, Ming Cai and Yuan Taur
University of California, San Diego
9500 Gilman Dr. San Diego, CA,92093-0409
Email: miliu@ucsd.edu

***Abstract*— This paper investigates the scaling limit of CMOS supply voltage for maintaining the noise margin of NAND circuits subject to process tolerance induced threshold voltage variation. It is shown that for high performance ($V_t/V_{dd} < 1/3$) decananometer CMOS devices with ±20% gate length tolerance and corresponding short-channel threshold roll-off, the supply voltage cannot be lower than 0.5 V in order to keep logic state consistency in the worst-case switching scenario.**

I. Introduction

In the past three decades, MOSFETs have been scaled to ever smaller dimentions in order to achieve higher density and faster speed in VLSI technology. As CMOS devices are scaled to 50 nm gate length and below, power supply voltage (V_{dd}) needs to be scaled to below 1 V in order to manage the growth of active power. Extensive research has been done on power supply scaling limit set by the off-current (V_t) vs. overdrive (V_t/V_{dd}) conflict [1] [2]. In the digital circuit applications, noise margin requirement of static CMOS circuits sets another limit on the power supply voltage. This limit is due to the nonlinearity degradation which is shown in Fig. 1. The transfer curves of an inverter are plotted. When V_{dd} is reduced from 1 V ($\sim 40kT/q$) to 0.2 V ($\sim 8kT/q$), the V_{NM}/V_{dd} ratio shrinks 38%. CMOS's exhibit non-linearity only when the supply voltage is much higher than the electron thermal energy. A fundamental limit of the minimum power supply voltage is $3-4\ kT/q$.

But for logic circuits with multiple fan-in, multiple transfer curves corresponding to different input combinations must be considered [3]. This aggravates the loss of noise margin as shown in Fig. 2, where the noise margin square is obtained from the leftmost and the rightmost (flipped) transfer curves. Substantially higher V_{dd} than that for the inverter (Fig. 1) is needed to guarantee sufficient noise margin.

As CMOS's are approaching the scaling limit, process-induced V_t variation becomes increasingly prominent and can significantly increase the minimum supply voltage. To

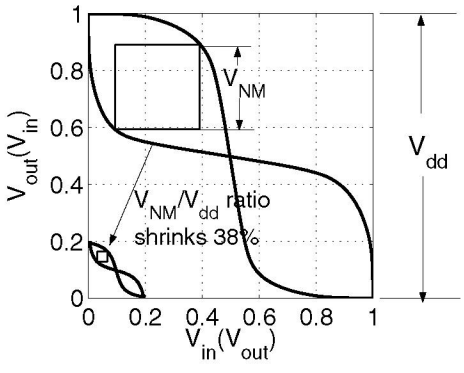
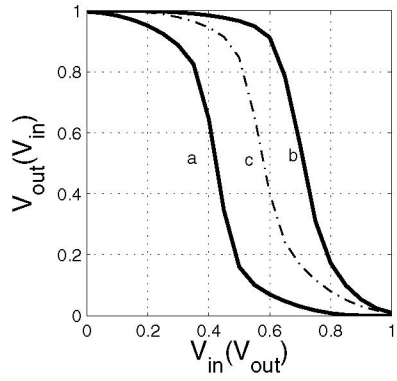
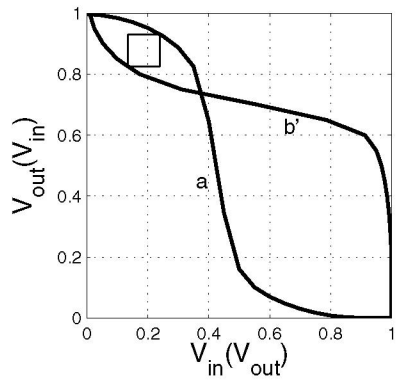


Fig. 1. The transfer curves of an inverter with V_{dd} of 1 V and 0.2 V respectively. For each V_{dd}, V_t is kept at 1/4 of V_{dd}. The slope of the linear transition region is degraded as V_{dd} decreases. Noise margin is the largest square that can fit inside the transfer curves.



(a)



(b)

Fig. 2. Noise margin for NAND logic circuits with multiple transfer curves depending on different combinations of inputs. In Fig.(a), one-gate-switching curve a is the leftmost curve with the lowest transition voltage and all-gates-switching curve b is the rightmost curve with the highest transition voltage. In Fig.(b), curve b' is the flipped curve of b in Fig.(a). Noise margin is the largest square that can fit into the "eye diagram".

TABLE I

V_t ROLL-OFF ASSUMPTION IN THE 2D DEVICE SIMULATION.

V_{ds} (V)	0.1	0.4	0.5	0.7	0.8
V_t (V) ($L = 8nm$)	0.151	0.089	0.070	0.032	0.013
V_t (V) ($L = 10nm$)	0.215	0.180	0.169	0.150	0.140
V_t (V) ($L = 12nm$)	0.259	0.236	0.230	0.218	0.212
V_t roll-off(V)	0.108	0.147	0.160	0.186	0.199

include V_t variation in the noise margin simulations, the worst case switching condition needs to be redefined. As shown in this paper, the new definition is based on both the input combination and the combination of CMOS's with different V_t. On the basis of this method, the minimum supply voltage condition was studied for high performance 10nm CMOS logic gate.

II. Device Modeling Assumptions

The simulation was carried out using the two-dimensional TCAD tool from SYNOPSYS. 10 nm is chosen as the channel length with the nominal V_t. $\pm 20\%$ channel length variation is assumed with a total V_t roll-off of 0.1 V ($V_{ds} = 0.1V$) as shown in Table I. The V_t roll-off is defined as the V_t difference between 12nm and 8nm CMOS's. V_t roll-off increases with drain voltage (supply voltage) due to the effect of drain induced barrier lowering (DIBL).

III. Noise Margin of CMOS NAND Circuits

In the binary digital logic, two distinct logic states are identified within a relatively small range of voltages. Logic state 1 lies in the voltage range between V_{Hmin} and V_{Hmax}. Logic state 0 lies in the voltage range between V_{Lmin} and V_{Lmax}. A combinatorial logic gate circuit is self-consistent when any possible combinations of inputs taken from the logic state ranges always produce an output state that lies in the correct logic state range.

In the logic circuit, inverters are added as buffers and they are adequate to push the logic state 1 to V_{dd} or the logic state 0 to ground. Therefore V_{Hmax} and V_{Lmin} are set to V_{dd} and 0 respectively. The definition of V_{Hmin} and V_{Lmax} is shown in Fig. 3 in terms of a 3-way NAND. Due to the V_t roll-off shown in Table I, a 12nm MOSFET has the lowest current drive and is hence the weak MOSFET. By the same token, an 8nm MOSFET is the strong MOSFET. The generation

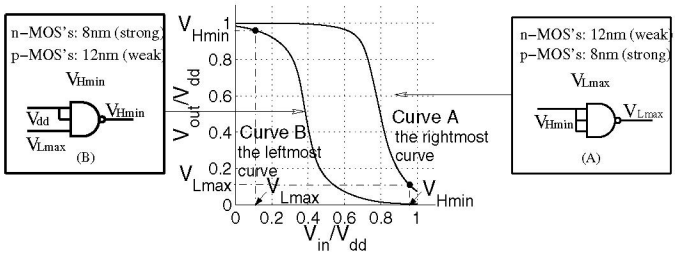
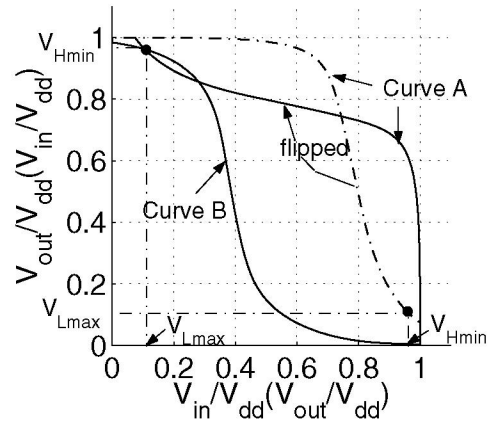
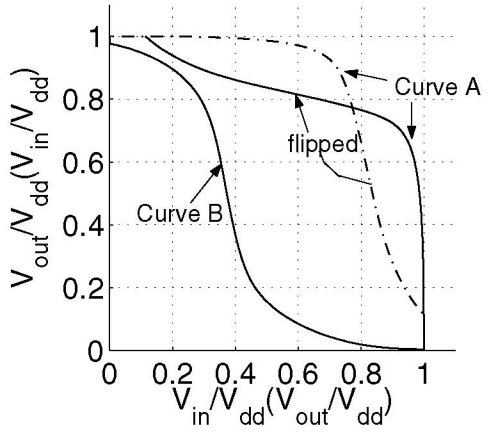


Fig. 3. The definition and the generation of the maximum low V_{Lmax} and the minimum high V_{Hmin} signal



(a) The transfer characteristics with the supply voltage above the minimum. Curve A is flipped to cross Curve B at (V_{Lmax},V_{Hmin}).



(b) The transfer characteristics with supply voltage below the minimum. The flipped Curve A no longer crosses Curve B. No results for V_{Lmax} and V_{Hmin} can be obtained.

Fig. 4. "eye diagram" for 3-way NAND gate logic circuit. V_{in} and V_{out} are normalized to the supply voltage. Curve A is generated by switching all inputs together. Curve B is generated by switching one input while tying the other two inputs at V_{dd}. Noise margin exists in (a), but not in (b).

of the maximum low signal V_{Lmax} requires strong pull-up MOSFETs (8nm-PMOS) and weak pull-down MOSFETs (12nm-NMOS) in the NAND circuit with all inputs tied at the minimum high signal V_{Hmin}. Based on the definition, the rightmost curve (Curve A) is generated by sweeping all inputs together from 0 to V_{dd}. Conversely, the generation of V_{Hmin} requires strong NMOS (8nm) and weak PMOS (12nm) with the logic inputs (1, 1, 0) as high as possible—(V_{dd}, V_{dd}, V_{Lmax}). The leftmost curve (Curve B) is generated by sweeping only one input while keeping the other inputs at V_{dd}. V_{Lmax} and V_{Hmin} are defined by the curves as shown.

Fig. 4 shows how the minimum supply voltage is determined from noise margin considerations. When the supply voltage is above the minimum limit in Fig. 4(a), curves A (flipped) and B have the intercepts and the noise margin is larger than zero. When the supply voltage is reduced too far, as shown in Fig. 4(b), the transfer curves do not intercept and

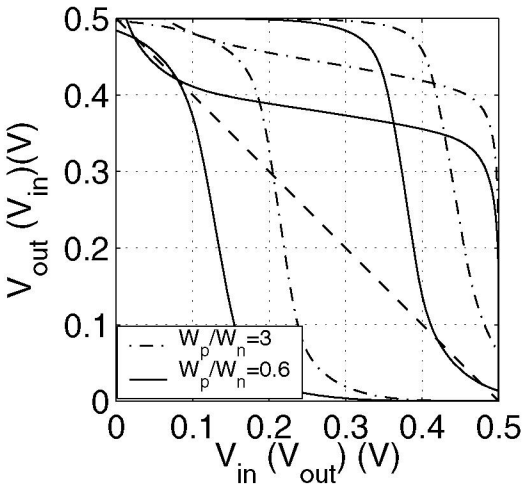


Fig. 5. Voltage transfer characteristics for different width ratios with fixed V_t. Noise margin is zero for width ratio of 3 (dash-dot lines) but larger than zero for width ratio of 0.6 (solid lines) where the leftmost curve and the flipped rightmost curve achieve better symmetry across the diagonal. Minimum V_{dd} is 0.5 V for $W_p/W_n = 3$ but smaller than 0.5 V for $W_p/W_n = 0.6$.

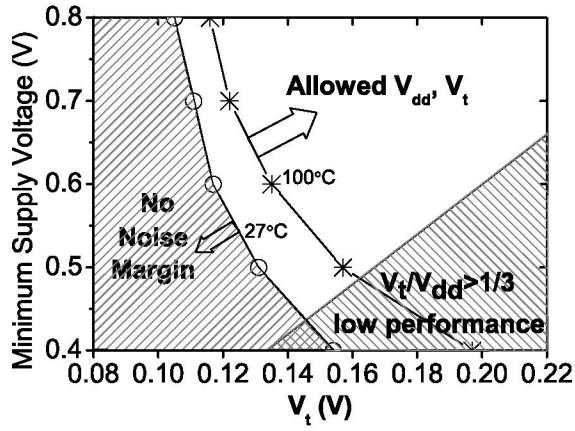


Fig. 6. Minimum supply voltage versus nominal V_t at 27°C and 100°C for 3-way NAND circuits. $V_t/V_{dd} < 1/3$ defines the high performance region. The minimum supply voltage is 0.43 V and 0.49 V at 27°C and 100°C respectively.

no consistent (V_{Lmax}, V_{Hmin}) can be defined. In that case, the output will eventually end up in the wrong state. When Curve B and the flipped Curve A have exactly one intersection, i.e., tangent to each other, the noise margin is zero, which corresponds to the minimum supply voltage condition.

Besides the dependence on V_t, minimum V_{dd} is also affected by the width ratio (W_p/W_n) of the NAND gate. As shown in Fig. 5 where voltage transfer characteristics are plotted with the same V_t but different width ratios. Minimum V_{dd} of 0.5 V is obtained at width ratio of 3 (dash-dot lines). As the width ratio decreases, the pull-down devices (NMOS) become stronger compared to the pull-up devices (PMOS), so the transfer curves shift left. Noise margin increases because the leftmost curve and the flipped rightmost curve achieve better symmetry across the diagonal as shown in Fig 5. When the width ratio reaches 0.6 (solid lines), the best symmetry is reached (width ratio smaller than 0.6 leads to the loss of symmetry) and the lowest minimum V_{dd} can be obtained. Therefore in the following noise margin study, W_p/W_n of 0.6 is used.

IV. Minimum Power Supply Voltage From Noise Margin Considerations

The simulation results of minimum V_{dd} as a function of the nominal V_t are plotted in Fig. 6 for a 3-way NAND gate at 27°C. V_t is varied by adjusting the gate work function so that the V_t roll-off is kept the same as that shown in Table I.

Higher V_t allows lower minimum V_{dd} because of the better non-linearity for higher V_t. The fundamental physics behind this is as follows: The slope of the voltage transfer curve is $V_{out}/V_{in} = g_m/g_{ds}$. Since $I_{ds} \propto (V_{gs} - V_t)^2$ where V_t is a function of V_{ds} due to DIBL, g_m is less sensitive to V_t variation than g_{ds}. The slope increases with higher V_t, leading to better non-linearity and allowing lower V_{dd}. High temperature degrades the noise margin condition through two aspects — the increase of kT/q factor and the decrease of V_t. By plotting the minimum V_{dd} vs. 100°C V_t, the latter effect is eliminated. The curve shows that the kT/q effect is increased for lower V_{dd}.

High V_t/V_{dd} ratio leads to low gate overdrive and poor performance. For high performance CMOS, $V_t/V_{dd} < 1/3$ is usually required. When both conditions are plotted in Fig. 6, they define a minimum supply voltage of 0.43 V for 27°C and 0.49 V for 100°C.

A more stringent assumption of V_t variation can decrease the minimum V_{dd}. Simulation of CMOS's with V_t roll-off of 100 mV at V_{ds} of 0.5 V (instead of 0.1 V) shows that minimum V_{dd} can be as low as 0.37 V for high performance 10nm CMOS.

V. Conclusion

Minimum V_{dd} of 0.5 V is obtained for high performance 10nm CMOS. The result can be applied to other channel length near scaling limit as long as the V_t roll-off is about the same level. For low power CMOS which allows $V_t/V_{dd} > 1/3$, a lower V_{dd} can be used with sufficient noise margin at the expense of switching performance.

VI. Acknowledgment

This work was supported in part by UMC under SRC Customization Project 2003-NJ-1145.001.

References

[1] Y. Taur and T.K.Ning,*Fundamentals of Modern VLSI Devices*, 1998.
[2] B. Davari, R. H. Denard and G. G. Shahidi, "CMOS Scaling for High Performance and Low Power—The Next Ten Years", *Proc. IEEE*, vol 83, pp. 595-606, Apr. 1995.
[3] D. J. Frank et al., "Device Scaling Limit of Si MOSFETs and Their Application Dependencies", *Proc. IEEE*, vol. 89, pp. 259-288, Mar. 2001.

