Pergamon

0038-1101(95)00171-9

*Solid-State Electronics* Vol. 39, No. 4, pp. 425–430, 1996
Copyright © 1996 Elsevier Science Ltd
Printed in Great Britain. All rights reserved
0038-1101/96 $15.00 + 0.00

# ON THE LOWER BOUNDS OF CMOS SUPPLY VOLTAGE

GERHARD SCHROM, CHRISTOPH PICHLER, THOMAS SIMLINGER and SIEGFRIED SELBERHERR

Institute for Microelectronics, Technical University Vienna, Gusshausstrasse 27-29/E360, A-1040 Vienna, Austria

*(Received 24 November 1994; in revised form 2 February 1995)*

**Abstract**—The lower bounds of the supply voltage $V_{DD}$ of ultra-low-power CMOS technologies are investigated under the constraints of standard digital circuit design. After discussing the peculiarities of ultra-low-power CMOS processes, low-voltage device operation, and scaling benefits, the lower limits for $V_{DD}$ are narrowed down from two sides. First, a simple inverter based on idealised transistors is investigated analytically to determine a set of *absolute* lower bounds of $V_{DD}$ for a set of given design constraints, i.e. minimum gain and noise margins. Next, the feasibility and performance of ultra-low-power CMOS technologies are investigated using process and device simulation, followed by post-processing of the simulated $I$–$V$ and capacitance data, to determine a set of *achievable* lower bounds of $V_{DD}$. On the basis of state-of-the-art processes and special scaling, a set of possible ultra-low-power CMOS processes was developed and numerically analysed on the gate level. These numerical data are then related to the analytical results.

## NOTATION

| | |
|---|---|
| $U_T$ | thermal voltage $kT/q$ (V) |
| $V_{in}$ | inverter input voltage (V) |
| $V_{out}$ | inverter output voltage (V) |
| $V_{DD}$ | supply voltage (V) |
| $NM$ | noise margin (%) |
| $G$ | inverter gain |
| $F_U$ | unsymmetry factor $W_n/W_p$ |
| $W$ | channel width (m) |
| $V_T$ | threshold voltage (V) |
| $I_{on}$, $I_{off}$ | on/off drain current (A) |
| $t_d$ | inverter delay (s) |
| $t_l$ | leakage time (s) |
| $E_s$ | switching energy (J) |
| $P_{stat}$ | static power (W) |

## 1. INTRODUCTION

By drastically decreasing the supply and threshold voltages, a great reduction in power consumption can be achieved at the expense of an increase in gate delay. This can be compensated to a certain extent by employing parallelism in the system's design so that, for a given overall performance, the total power consumption is drastically reduced compared to conventional CMOS techniques[1,2]. Apart from possible new circuit design techniques, it is of great importance to ensure the portability of common design practice to ultra-low-power CMOS. Therefore, the analyses are carried out with respect to standard design constraints, such as minimum noise margins, gain, etc. This has to be done directly at the gate level because device-parameter based standard models are not accurate enough as the standard device parameters are themselves questionable when dealing with very small voltages[3]. At very low supply voltages, the transistors must be operated in the weak inversion region up to the onset of strong inversion, in order to achieve appropriate noise immunity and adequate speed at the same time. Therefore, the transition region between weak and strong inversion has a great impact on the performance of ultra-low-power CMOS, and traditional performance analysis methods based on compact modelling are inadequate. For this reason, we use a table-driven device model for circuit simulation, taking the $I$–$V$ data directly from the device simulation without any parameter fitting.

## 2. PROCESS TECHNOLOGY

The processes under consideration are recessed-well dual-gate processes with a very thin gate oxide (below 6 nm) to obtain controllably low threshold voltages. The source/drain (S/D) dopings are formed by single shallow implants and a conventional furnace anneal. The gate/source (G/S) and gate/drain (G/D) overlap capacitances can be controlled with a spacer formed prior to the S/D implants. As a consequence of the low voltages, ultra-low-power processes differ from conventional CMOS processes in several ways: because of the low $V_{DD}$, the hot carrier problem virtually does not exist and therefore an LDD process is not necessary; short-channel effects like DIBL are effectively reduced and no GIDL can occur; also, the effects of velocity saturation and mobility degradation become less severe. All these features extend the scalability of ultra-low-power CMOS processes. A major challenge is to achieve controllably low threshold voltages. Although the

adjustment of $V_{Tn,p}$ with a bulk bias seems very attractive, this method is unlikely to be accepted for digital circuit design because of the significant overhead (bulk-biasing would require two additional bus lines and prevent the use of local interconnects for the S/B contacts). Another problem can arise from the very thin gate insulator. If one uses very thin thermally grown gate oxides (below 6 nm), boron diffusion can considerably degrade the device behaviour. Also, boron segregation causes deleterious effects, especially in the sub-threshold region. On the other hand, there are several options for the gate insulator. Nitrides or oxynitrides may be good alternatives to conventional (pure $SiO_2$) gate oxides. Silicon nitride can be used as an effective diffusion barrier and ultra-thin $Si_3N_4$ layers with low defect densities are easier to fabricate[4]. The lower tunneling barrier compared to that of $SiO_2$ is still acceptable for ultra-low-power CMOS because of the low voltages and, also, because a controllably small gate current is allowed. As for very low $V_{DD}$, the devices must operate in the weak inversion region, the difference of the carrier mobilities $\mu_n$, $\mu_p$ can be roughly compensated by adjusting the threshold voltages to achieve symmetric inverter transfer characteristics. This compensation does not work, however, in the transient case because the speed is mainly determined by the strong inversion part of the input characteristics. The sub-threshold behavior is crucial because it determines the achievable ratio of $I_{on}/I_{off}$ which is limited by $e^{V_{DD}q/kT}$ and decreases as $V_{Tn,p}$ are made smaller. Therefore, "zero $V_T$" transistors are not desirable. On the other hand, if $V_{Tn,p}$ are too high, the speed becomes unacceptably low.

### 3. ANALYTICAL LOWER BOUNDS

To determine an absolute lower bound of the supply voltage we assume ideal MOSFETs with a gate swing of $S = \ln 10 \cdot U_T$, operating completely in the weak inversion mode. The drain currents of the $N$MOST and $P$MOST are then given by $I_{Dn} = I_{sn} e^{V_{GSn}/U_T} (1 - e^{-V_{DSn}/U_T})$ and $I_{Dp} = -I_{sp} e^{-V_{GSp}/U_T} (1 - e^{V_{DSp}/U_T})$, respectively[5], with $V_{BSn} = V_{BSp} = 0$ V. Note that in this case the threshold voltages $V_{Tn,p}$ can be modeled into $I_{sn,p}$. Setting $I_{Dn} + I_{Dp} = 0$ with $I_{sn} = I_{sp}$ yields an implicit equation for the transfer curve $V_{out}$ ($V_{in}$) of a symmetric inverter:

$$\frac{\sinh\left(\frac{V_{in} - V_{out}}{U_T}\right)}{\sinh\left(\frac{V_{in} - V_{DD}/2}{U_T}\right)} = e^{\frac{V_{DD}/2}{U_T}}, \qquad (1)$$

and the critical points where the gain $G = dV_{out}/dV_{in} = -1$ are determined by:

$$\frac{2\cosh\left(\frac{V_{in,c} - V_{out,c}}{U_T}\right)}{\cosh\left(\frac{V_{in,c} - V_{DD}/2}{U_T}\right)} = e^{\frac{V_{DD}/2}{U_T}}, \qquad (2)$$

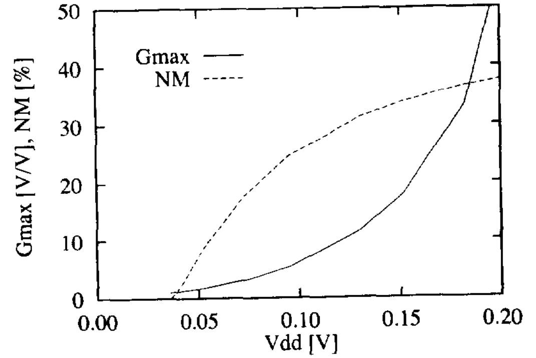


Fig. 1. Maximum gain and noise margins of a symmetrical inverter with ideal MOS transistors operating in the weak inversion mode.

and the noise margins $NM_H = NM_L = NM$ are computed as:

$$NM = \frac{V_{DD} - V_{out,c} - V_{in,c}}{V_{DD}} (\% V_{DD}). \qquad (3)$$

The maximum gain which occurs at $V_{in} = V_{out} = V_{DD}/2$ is given by:

$$-G_{max} = e^{\frac{V_{DD}/2}{U_T}} - 1. \qquad (4)$$

Solving eqns (1) and (2) numerically, together with eqns (3) and (4), yields noise margins and maximum gain as a function of the supply voltage. Figure 1 shows a plot of $NM$ and $G_{max}$ vs $V_{DD}$. For the design of digital circuits we have to impose certain constraints, i.e. to specify minimum values for $NM$ and $G_{max}$ at a nominal and maximum temperature, and to estimate the impact of an effective unsymmetry, $F_U = W_n/W_p$, as a consequence of minimum transistor size design. This is accounted for by a shift of the input voltage $\Delta V_{in} = -U_T \cdot \ln(F_U)/2$. These data are compiled in Table 1. Note that these numbers are absolute lower bounds which are not likely to be achieved with any CMOS process technology. Achievable values for $V_{DDmin}$ may be estimated by scaling the numbers from Table 1 by a factor of $S/(U_T \cdot \ln 10)$, where $S$ is an achievable average gate swing. Although this is not consistent with eqns (1) and (2), it can be used as a worst-case estimate for sub-threshold operation.

Table 1. Ideal-case minimum supply voltage $V_{DD}$ for given circuit design constraints

| Constraint | $V_{DDmin}$ (mV) ($T = 300$ K) | $V_{DDmin}$ ($U_T$) |
|---|---|---|
| $G_{max} > 1$ (ring oscillator) | 36 | 1.40 |
| $NM > 10\%$ (inverter) | 55 | 2.13 |
| $G_{max} > 4$ (standard design) | 83 | 3.22 |
| $F_U > 9$ (fan-in = 3) | 83 | 3.22 |
| $I_{on}/I_{off} > 10^4$ (dynamic logic) | 238 | 9.22 |

## 4. PROCESS AND DEVICE SIMULATION

To determine an achievable lower bound of the supply voltage, based on the guidelines in Section 2, a set of processes was designed and numerically evaluated. Both process and device simulation were done using VISTA with the SFC (simulation flow controller) to allow for quick process design and evaluation[6]. The simulated processes are a 0.35 $\mu$m process (A) for static logic and a 0.5 $\mu$m process (B) for dynamic logic. The processes were designed for proper d.c. characteristics but were not optimised for speed. Figures 2 and 3 show the doping profiles of static-logic CMOS devices. The dopant concentration under the gate is rather high and decays with increasing depth. For the electrical characterisation of the devices, MINIMOS[7] was used to calculate a matrix of drain currents $I_D(V_G, V_D)$ over a range of $V_G$ and $V_D$ for the *P*MOS and *N*MOS transistors. Based on these data, a fast and accurate table-driven d.c. analysis of simple gates and inverters is possible. The bulk effect could also be included but for the given devices and voltages it was found not to be significant. The dynamic behavior was estimated from capacitance data obtained by a.c. analysis with MINIMOS. The device characteristics for process A are shown in Figs 4 and 5. Figures 6 and 7 show the inverter transfer curves for processes A and B. Figures 8 and 9 show the noise margins and the inverter delay as a function of the supply voltage. From Fig. 8 it can be seen that a ring oscillator built with process A would work even at $V_{DD} = 80$ mV, and by using additional inverters at the gate inputs and outputs one could also design digital circuits for $V_{DD} < 100$ mV, but the overhead would be considerable. Figures 10 and 11 show the noise margins and the inverter delay as a function of the temperature. For process B, the ratios of $I_{on}/I_{off}$ in Table 2 are in the order of $10^4$, and the ratio of $\tau_l/\tau_d$ is about 2300, which is still acceptable for dynamic logic. For process A, it can be seen from Table 3 that for three-input NANDs with minimal transistors, the high-noise-margin, $NM_H$, is already very low. From these data we conclude that the achievable limits for the supply voltage will be at 200 mV for static logic and 500 mV for dynamic logic, with a fan-in of 3 at $T = 300$ K. The ultimate limit for the CMOS supply voltage is given by the thermal voltage as $V_{DD} > X \cdot kT/q$ where $X$ is a factor depending on the type of digital circuit technique and on the process technology.

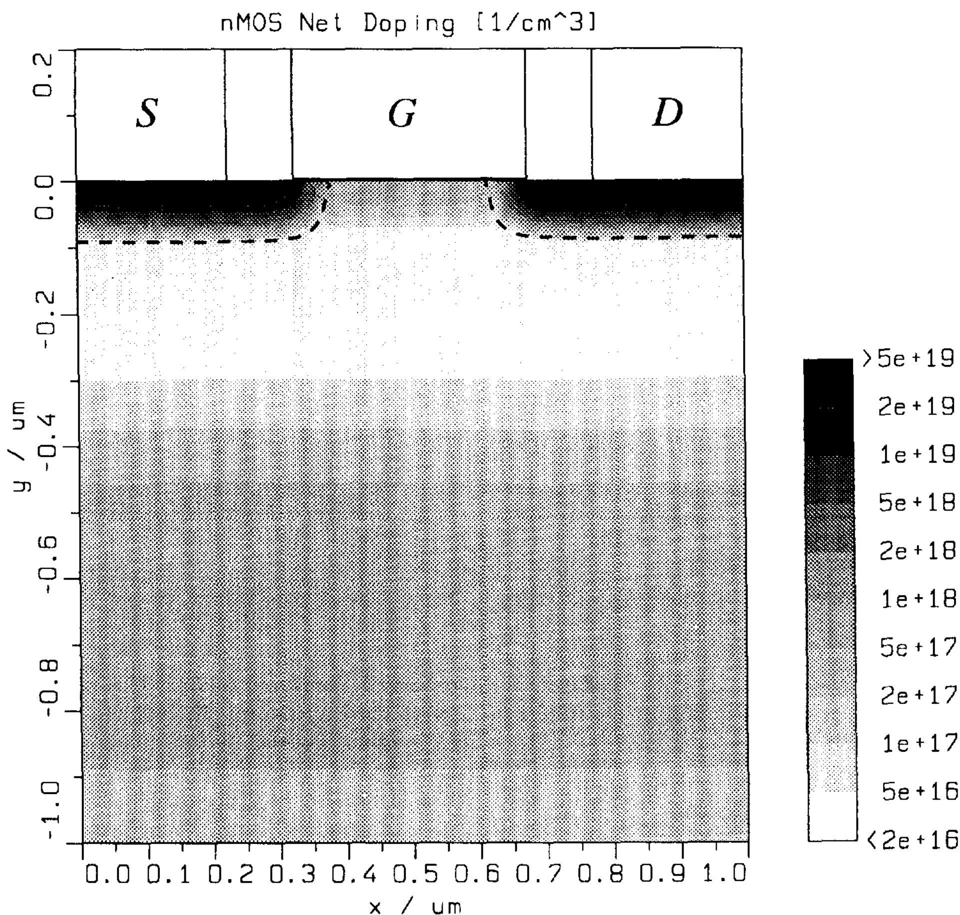


Fig. 2. NMOS doping profile, process A ($V_{DD} = 200$ mV).

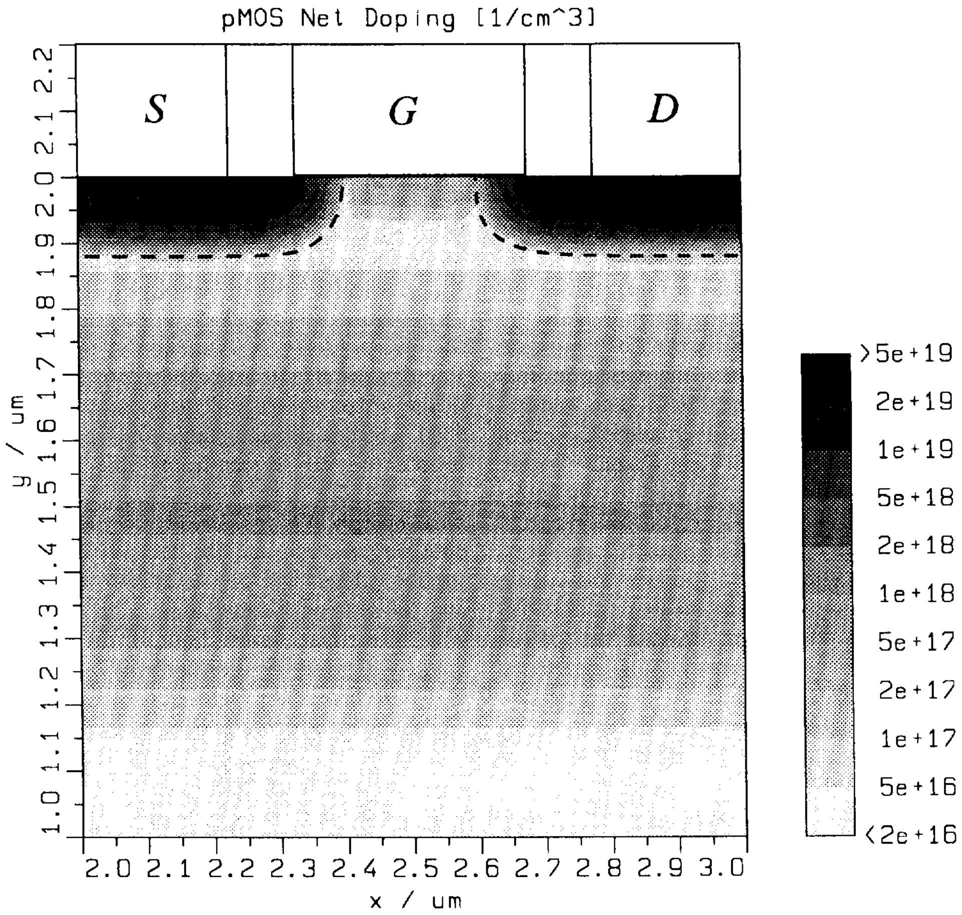


Fig. 3. PMOS doping profile, process A ($V_{DD} = 200$ mV).

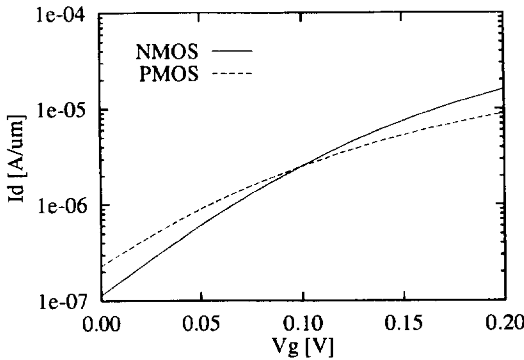


Fig. 4. Input characteristics, process A ($V_{DD} = 200$ mV).

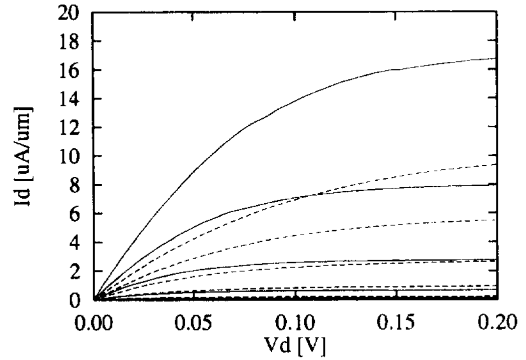


Fig. 5. Output characteristics, process A ($V_{DD} = 200$ mV).

We found that $X_{stat} < 8$ and $X_{dyn} < 20$ is sufficient for a fan-in of 3. To relate these data to the results in Table 1 we define a factor $Y$:

$$Y = \ln\sqrt{\frac{I_{on,n} I_{on,p}}{I_{off,n} I_{off,p}}}, \tag{5}$$

which should be slightly higher than the respective factor for $U_T$ in Table 1. In fact, for process A we obtain $Y = 4.17$, compared to 3.22, and $Y = 9.60$ for process B, compared to 9.22. This indicates that the on/off current ratio may be used as a quality criterion for a given process.

## 5. CONCLUSION

The lower bounds of ultra-low-power CMOS supply voltage were investigated both analytically and with process and device simulations, under the constraints of standard digital design. From these

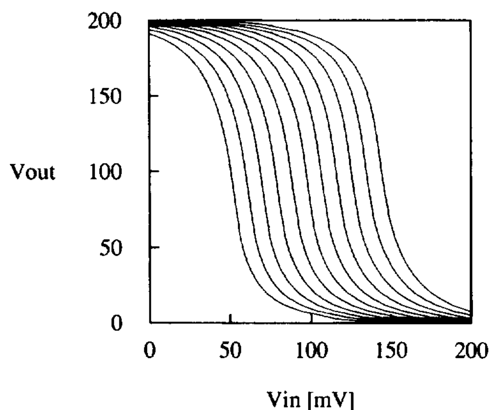


Fig. 6. Inverter transfer characteristics for $W_n/W_p = 0.1 \ldots 10$, process A at $V_{DD} = 200$ mV.

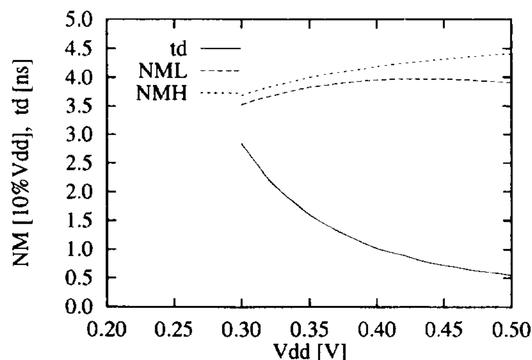


Fig. 9. Noise margins and delay time vs $V_{DD}$, process B.

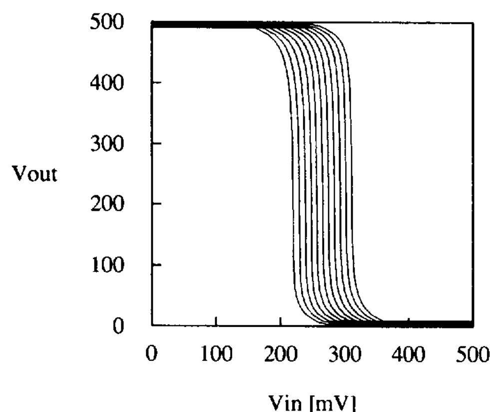


Fig. 7. Inverter transfer characteristics for $W_n/W_p = 0.1 \ldots 10$, process B at $V_{DD} = 500$ mV.

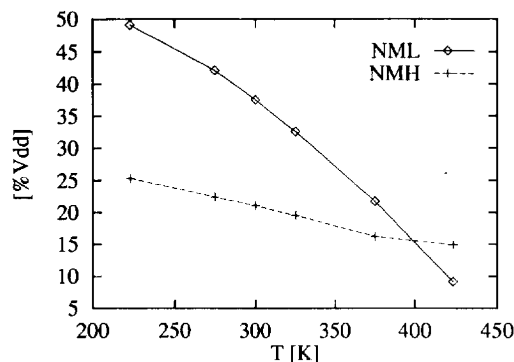


Fig. 10. Noise margins vs temperature, process A.

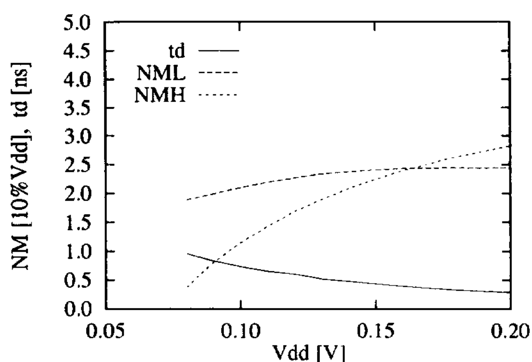


Fig. 8. Noise margins and delaytime vs $V_{DD}$, process A.

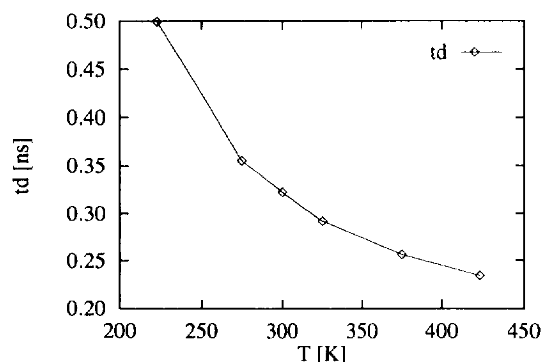


Fig. 11. Delay time vs temperature, process A.

Table 2. Simulated device characteristics. The threshold voltage was defined as $|I_D(V_T)| = 1\ \mu A\ \mu m^{-1}$

| Process | $V_{DD}$ (V) | $V_{T,n}$ (V) | $V_{T,p}$ (V) | $I_{off,n}$ (A $\mu m^{-1}$) | $I_{off,p}$ (A $\mu m^{-1}$) | $I_{on,n}$ (A $\mu m^{-1}$) | $I_{on,p}$ (A $\mu m^{-1}$) |
|---|---|---|---|---|---|---|---|
| A | 0.2 | 0.067 | −0.059 | $0.14 \times 10^{-6}$ | $0.27 \times 10^{-6}$ | $16.7 \times 10^{-6}$ | $9.4 \times 10^{-6}$ |
| B | 0.5 | 0.26 | −0.24 | $0.7 \times 10^{-9}$ | $2.8 \times 10^{-9}$ | $25.6 \times 10^{-6}$ | $16.7 \times 10^{-6}$ |

Table 3. Noise margins (in % $V_{DD}$) for a simple inverter and a 3-input NAND gate, and inverter delay, leakage time, switching energy, and static power consumption

| Process | $NM_{H,inv}$ | $NM_{L,inv}$ | $NM_{H,gate}$ | $NM_{L,gate}$ | $t_d$ | $t_l$ | $E_s$ | $P_{stat}$ |
|---|---|---|---|---|---|---|---|---|
| A | 28 | 23 | 13 | 39 | 0.29 ns | 7.2 ns | 0.65 fJ | 41 nW |
| B | 38 | 44 | 31 | 49 | 0.55 ns | 1.3 $\mu$s | 4.3 fJ | 0.88 nW |

analyses, absolute lower bounds and achievable bounds of $V_{DD}$ were determined as 83 and 200 mV for static logic, and 238 and 500 mV for dynamic logic, respectively. A simple quality criterion ($Y$ factor) was proposed to relate the numerical data to the analytical data.

## REFERENCES


1. D. Liu and C. Svensson, *IEEE J. Solid-St. Circuits* **28** 10 (1993).
2. S. Malhi and P. Chatterjee, *IEEE Circuits and Devices Mag.* p. 13 (March 1994).
3. G. Schrom, D. Liu, C. Pichler, C. Svensson and S. Selberherr, in *24th European Solid St. Device Res. Conf. ESSDERC'94* (Edited by C. Hill and P. Ashburn) p. 679. Editions Frontieres (1994).
4. T. Morimoto, H. S. Momose, S. Takagi, K. Yamabe and H. Iwai, in *Proc. 22nd Int. Conf. on Solid-St. Devices and Materials*, p. 361. Sendai, Japan (1990).
5. E. A. Vittoz, *23rd European Solid St. Device Res. Conf. EDDSERC'93* (Edited by J. Borel, P. Gentil, J. P. Noblanc, A. Nouailhat and M. Verdone) p. 927. Edition Frontieres, Gif-sur-Yvette Cedex, France (1993).
6. C. Pichler and S. Selberherr, in *Simulation of Semiconductor Devices and Processes* (Edited by S. Selberherr, H. Stippel and E. Strasser) Vol. 5, p. 25. Springer-Verlag, Wien/New York (1993).
7. C. Fischer, P. Habaš, O. Heinreichsberger, H. Kosina, P. Lindorfer, P. Pichler, H. Pötzl, C. Sala, A. Schütz, S. Selberherr, M. Stiftinger and M. Thruner, *MINIMOS 6.0 User's Guide*. Technical University, Vienna (1994).