Low Energy CMOS for Space Applications[1]

Ramesh Panwar and Leon Alkalaj
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California 91109

1 Introduction

The current focus of NASA's space flight programs reflects a new thrust towards smaller, less costly and more frequent space missions, when compared to missions such as Galileo, Magellan, or Cassini. Recently, the concept of a *microspacecraft* has been proposed, whereby a small, compact spacecraft that weighs tens of kilograms performs focused scientific objectives such as imaging. Similarly, a Mars Lander *micro-rover* project is under study that will allow miniature robots weighing less than seven kilograms to explore the Martian surface. To bring the microspacecraft and microrover ideas into fruition one will have to leverage compact 3D MCM technologies. Low energy CMOS will become increasingly important because of the thermodynamic considerations in cooling compact 3D MCM implementations and also from considerations of the power budget for space applications. In this paper, we show how the operating voltage is related to the threshold voltage of the CMOS transistors for accomplishing a task in VLSI with minimal energy. We also derive expressions for the noise margins at the optimal operating point. We then look at a low voltage CMOS (LVCMOS) technology developed at Stanford University which improves the power consumption over conventional CMOS by a couple of orders of magnitude and consider the suitability of the technology for space applications by characterizing its SEU immunity.

2 MOS Transistor Model

A MOS transistor has three regions of operation where the behavior is described by the following equations [1]:

$$I_{ds} = \begin{cases} \beta n V_T^2 e^{(V_{gs}-V_t)/nV_T} & \text{if } V_{gs} < V_t \\ \beta[(V_{gs}-V_t)V_{ds} - V_{ds}^2/2] & \text{if } 0 < V_{ds} < V_{gs} - V_t \\ \frac{\beta}{2}(V_{gs}-V_t)^2 & \text{if } 0 < V_{gs} - V_t < V_{ds} \end{cases}$$

where

I_{ds}	=	drain current
V_{gs}	=	gate-source voltage
V_{ds}	=	drain-source voltage
V_t	=	threshold voltage
V_T	=	thermal voltage
n	=	a process parameter between 1.2 and 1.4
β	=	transistor gain factor

[1]Partial support for this work was provided by the Data Systems Program NASA Code-R CSTI

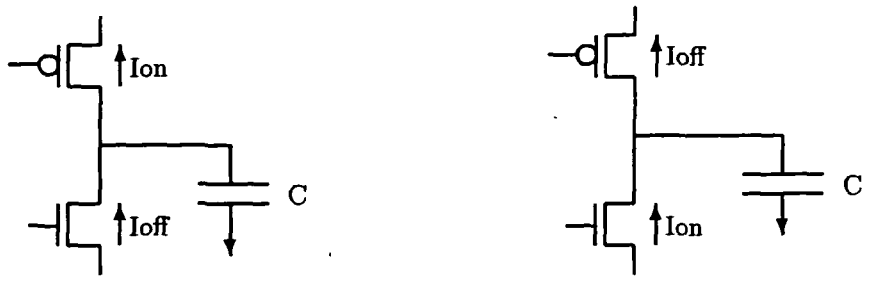

Figure 1: Gate capacitance charge/discharge

Note that the leakage current (I_{off}) through a transistor when the gate voltage is zero ($V_{gs} = 0$) is given by:

$$I_{off} = I_0 e^{-V_t/nV_T}$$

where $I_0 = \beta n V_T^2$. Thus the lower the threshold voltage (V_t), the higher the leakage current.

3 Gate Switching Time

Consider a complementary CMOS gate driving the gate of another transistor. The gate capacitance of the driven transistor will be charged or discharged by the driving gate. The charge needed to charge or discharge the gate capacitance is given by:

$$Q = CV$$

where C is the gate capacitance and V is the voltage swing at the gate (usually the operating voltage unless we are using differential mode logic). Let the time required to switch the gate capacitance be denoted by t_p. The charge Q is then given by:

$$Q = (I_{on} - I_{off})t_p$$

where I_{on} is the saturation current through the ON device and I_{off} is the leakage current through the OFF device. We thus have the following expression for the gate switching time:

$$t_p = \frac{CV}{\frac{\beta}{2}(V - V_t)^2 - I_0 e^{-V_t/nV_T}}$$

The charging/discharging of the gate capacitance by the ON/OFF currents is shown in Figure 1. We are now in a position to consider the energy required for VLSI computation.

4 VLSI Task Energy

The energy needed to accomplish a given task in VLSI is equal to the sum of the AC energy and the DC energy. The AC energy is consumed during computation due to transistors

switching state, while the DC energy is the energy that is dissipated due to leakage currents. Let us assume that we can attain a speedup $s(m)$ in a VLSI computation where m is the number of transistors, and let t_p denote the gate switching time of a transistor. The time required to complete a task is then expressed as $kt_p/s(m)$ where k is some constant for the task. The AC energy for the task is expressed as:

$$E_{ac} = \frac{1}{2} amCV^2 fkt_p/s(m)$$

where f is the frequency of operation and a is the average fraction of transistors that switch in a clock cycle. The DC energy for the task is given by:

$$E_{dc} = mI_{off}Vkt_p/s(m)$$

Thus the total energy for the task is given by:

$$E = E_{ac} + E_{dc}$$

By using the expressions for t_p and the leakage current I_{off}, it can be shown that:

$$E = \frac{km(afCV^2 + 2I_0Ve^{-V_t/nV_T})CV}{s(m)(\beta(V - V_t)^2 - 2I_0e^{-V_t/nV_T})}$$

This is minimized when

$$\frac{\partial E}{\partial V} = 0$$

which leads us to a cubic equation of the form:

$$c_3V^3 + c_2V^2 + c_1V + c_0 = 0$$

where the coefficients are given by:

$$\begin{aligned} c_3 &= af\beta C \\ c_2 &= -4af\beta V_tC \\ c_1 &= 3af\beta CV_t^2 - 4\beta I_0V_te^{-V_t/nV_T} - 6afI_0Ce^{-V_t/nV_T} \\ c_0 &= 4kI_0V_t^2e^{-V_t/nV_T} - 8I_0^2e^{-2V_t/nV_T} \end{aligned}$$

Note that this means that the optimal supply voltage is dependent not only on the threshold voltage but also on various other parameters like the frequency of operation, the activity of the logic, the capacitance of the nodes, and the transistor geometry. However, for sufficiently high threshold voltages where $V_t \gg nV_T$, the cubic equation reduces to the following quadratic:

$$V^2 - 4V_tV + 3V_t^2 = 0$$

whose solution is given by $V = 3V_t$. This result is independent of the activity and the frequency of operation and suggests that the threshold voltage should be set to a third of the supply voltage for optimal operation. We next show that the relative noise margins are unaffected by the operating voltage.

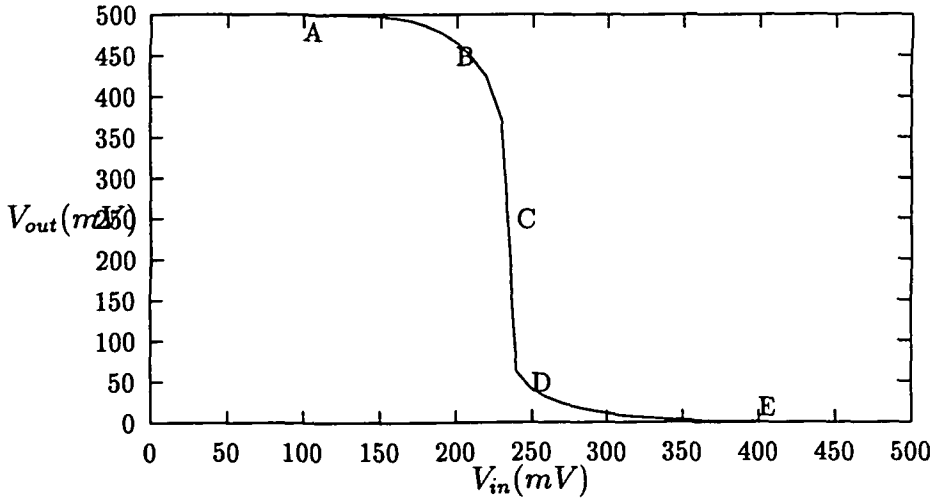

Figure 2: Regions of operation of an inverter

5 Noise Margins

The operation of a CMOS inverter can be divided into five regions [1]:

- Region A. In this region p-device is in the linear region while the the n-device is cut-off.
- Region B. In this region the p-device is in the linear region while the n-device is in saturation.
- Region C. In this region both the n- and p-devices are in saturation.
- Region D. In this region the p-device is in saturation while the n-device is in the linear region.
- Region E. In this region the p-device is cut-off and the n-device is in the linear mode.

The noise margins are defined to be:

$$NM_L = V_{IL} - V_{OL}$$

$$NM_H = V_{OH} - V_{IH}$$

where the voltages V_{IL} and V_{IH} are given by the unity gain points and the voltages V_{OL} and V_{OH} are 0 and V_{dd} for ideal inverters [1]. The relative noise margins are defined to be the ratio of the noise margins to the operating voltage and are given by:

$$RNM_L = \frac{NM_L}{V_{dd}}$$

$$RNM_H = \frac{NM_H}{V_{dd}}$$

The unity gain points occur in regions B and D and are given by setting the derivative dV_O/dV_I to -1. It can be shown that the output voltage in region B is given by [1]:

$$V_O = (V_I - V_{tp}) + [(V_I - V_{tp})^2 - 2(V_I - V_{dd}/2 - V_{tp})V_{dd} - \frac{\beta_n}{\beta_p}(V_I - V_{tn})^2]^{1/2}$$

Assuming $\beta_n = \beta_p$ and $V_{tn} = -V_{tp} = \alpha V_{dd}$, we have by taking the derivative and setting it to -1, the following expression for V_{IL}:

$$V_{IL} = \frac{1}{8}V_{dd}(3 + 2\alpha)$$

The noise margin NM_L is thus given by:

$$NM_L = \frac{1}{8}V_{dd}(3 + 2\alpha)$$

and the relative noise margin RNM_L is given by:

$$RNM_L = \frac{3}{8} + \frac{1}{4}\alpha$$

Similarly, in region D, the expression for the output voltage is given by [1]:

$$V_O = (V_I - V_{tn}) - [(V_I - V_{tn})^2 - \frac{\beta_p}{\beta_n}(V_I - V_{dd} - V_{tp})^2]^{1/2}$$

Again, by taking the derivative and setting it to -1, we have:

$$V_{IH} = \frac{1}{8}V_{dd}(5 - 2\alpha)$$

The noise margin NM_H is thus given by:

$$NM_H = \frac{1}{8}V_{dd}(3 + 2\alpha)$$

and the relative noise margin RNM_H is given by:

$$RNM_H = \frac{3}{8} + \frac{1}{4}\alpha$$

Thus at the optimal operating point for minimal VLSI energy, the relative noise margins would be given by:

$$RNM_L = RNM_H = 11/24$$

Note that this depends on several assumptions of which the most important ones are:

- The absolute values of the threshold voltages are equal.

- The β values of the n- and p-devices are equal.
- The supply and threshold voltages are such that all devices are capabale of operating in all 3 regions of operation (subthreshold, linear and saturation) and that the inverter characteristic shows 5 regions of operation as outlined above.

The reason for considering relative noise margins as opposed to noise margins is because the noise on the chip will scale down as the supply voltage is scaled down. Note that the relative noise margins for any computation at the minimal VLSI energy are independent of the supply voltage and this appears very encouraging since it allows us to scale the supply voltage as much as possible. We now show the operating characteristics of a low voltage CMOS technology that can be operated with a wide range of operating voltages.

6 Stanford LVCMOS

Low voltage CMOS devices built at Stanford University allow chips to be operated with supply voltages of several hundreds of millivolts to 5V. The devices are manufactured such that the zero-bias threshold voltage of the n- and p-devices is almost zero [2]. The bulk bias is then adjusted to set the threshold voltage to the appropriate value. The substrate of the p-devices is tied to a voltage $V_{pb} = V_{dd} + V_{bb}$ where V_{bb} is the extra bulk-bias voltage while the substrate of the n-devices is tied to a voltage $V_{nb} = -V_{bb}$. One disadvantage of this technology is that three voltages (V_{dd}, V_{pb}, V_{nb}) need to be routed on the chip besides the ground. The routing of four voltage levels may adversely affect the integration density of the devices. We are looking into the effect of the 4 rails on the integration density as opposed to the dual rail design in conventional CMOS. Figures 4-10 show some of the characterizations that have been done for this technology. One can see that the devices fail to saturate when operated at a supply voltage of 100mV with no extra bulk bias. However, the devices can be made to saturate at the same supply voltage by an extra bulk-bias of 0.5V. Increasing the extra bulk-bias decreases the saturation current through the devices since it increases the absolute value of the threshold voltage. This also makes the devices slower as shown by the inverter propagation delays in Tables 1 and 2. Tables 1 and 2 show the relative noise margins and inverter propagation delays for the LVCMOS technology at operating voltages of 100mV and 500mV respectively. The tables show that the inverter propagation delays for an operating voltage of 500mV degrade very gracefully with increasing bulk-bias but the propagation delays for an operating voltage of 100mV degrade very badly with increasing bulk-bias. Also, since the devices do not saturate at an operating voltage of 100mV without a significant amount of extra bulk-bias, it means that the devices cannot be operated at 100mV except for really slow applications. The inverter propagation delay degradation is less with increasing V_{bb} for higher operating voltages. For a 5V operating voltage, the inverter propagation delay remains virtually constant at 0.11ns. Figures 11-13 show the inverter characteristics for an inverter in this technology for three different supply voltages. For a low energy operation, one would choose an operating voltage of 500mV and set the threshold voltage to be a third of the operating voltage by adjusting the back-bias to achieve optimal operation. We next consider the SEU immunity of this technology.

$V_{bb}(V)$	RNM_L	RNM_H	$t_{inv}(ns)$
0	0.00	0.55	2
0.5	0.35	0.54	8
1.0	0.30	0.60	680
2.0	0.25	0.64	5840

Table 1: Relative noise margins and inverter propagation delays at 100mV

$V_{bb}(V)$	RNM_L	RNM_H	$t_{inv}(ns)$
0	0.20	0.60	0.34
0.5	0.26	0.58	0.41
1.0	0.38	0.48	0.53
2.0	0.44	0.48	1.00

Table 2: Relative noise margins and inverter propagation delays at 500mV

7 SEU Immunity of LVCMOS

We now characterize the SEU immunity of the LVCMOS technology by looking at the critical charge needed to cause an SRAM cell upset. Figure 3 shows the SPICE model used for the SEU immunity study. The particle strike is modeled by a pulsed current source with a triangular pulse having a 1:9 rise:fall shape. The pulse height is adjusted using a binary search to determine the critical charge at which the cell is just upset [3]. The critical charge for the upset (Q_c) is given by the area under the current pulse. Simulations were performed for an operating voltage of 500mV for different bias voltages. Simulations show that the critical charge is quite independent of the pulse width for pulse widths below 500ps. [The typical alpha particle pulse width is 200ps]. Table 3 shows the critical charge for various pulse widths and bias voltages for an operating voltage of 500mV.

The results show that this technology is much more prone to SEUs when operated at 500mV because the critical charge is much smaller than conventional CMOS. However, the advantge of this technology is that the devices can be operated at higher operating voltages

$V_{bb}(V)$	20ps	200ps	2ns
0	42fC	42fC	49fC
0.5	41fC	41fC	47fC
1.0	38fC	39fC	45fC
2.0	33fC	35fC	42fC

Table 3: Critical charge for various pulse widths at 500mV

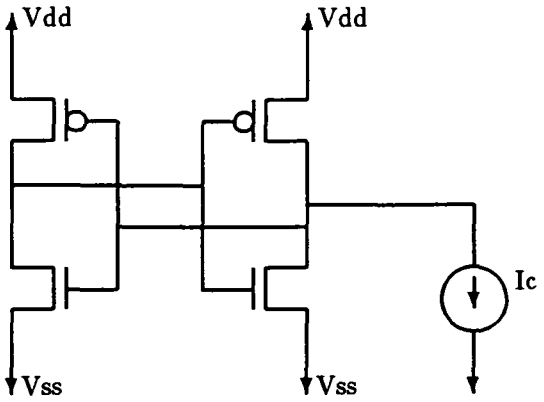

Figure 3: SRAM SEU model

$V_{dd}(V)$	Q_c
0.5	42fC
1.0	95fC
2.0	160fC
5.0	420fC

Table 4: Critical charge for various operating voltages

to ensure better SEU immunity. Table 4 shows the SEU immunity of the devices at $V_{bb} = 0V$ for various operating voltages when the SEU particle strike is modeled by a pulse of 200ps width. We see that the SEU immunity of the technology for an operating voltage of 5V is quite comparable to that of rad-hard CMOS. To achieve low energy operation, we can allow two different operating voltages for the chips. The chips will operate at a low voltage (say 500mV) in a clean environment and a higher voltage (say 5V) in a radiation intensive environment to ensure better SEU immunity. One can envision a system designed with a radiation detector that switches the operating voltage from 500mV to 5V when it encounters radiation and switches back to low voltage operation when the environment becomes clean again.

8 Conclusions

In this paper, we have presented the analysis for optimal supply voltage for achieving a task in VLSI with minimal energy. We also showed that the relative noise margins are independent of the operating voltage. We then showed some characteristics of the Stanford LVCMOS technology and showed how the logic speed depends on the operating voltage in this technology. The threshold voltage in this technology is adjusted by setting the back-bias appropriately. For optimal energy operation, the threshold voltage would have to be

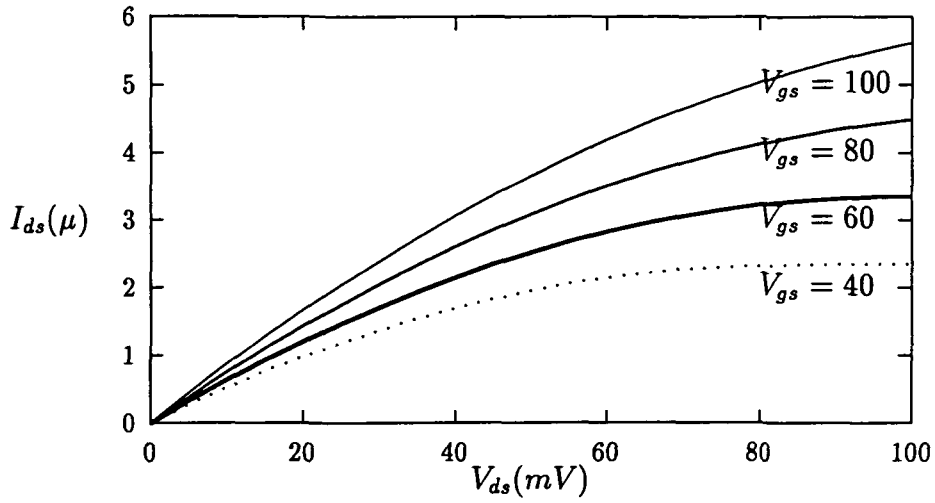

Figure 4: VI characteristic for $V_{bb} = 0V$ and 100mV operating voltage

set to a third of the operating voltage. We also characterized the critical charge for SEUs in this technology and showed it to be significantly less than conventional technology when the operating voltage was 500mV but comparable to rad-hard technology when the operating voltage is 5V. The benefit of the LVCMOS technology is that the supply voltage and the threshold voltage can both be changed and designed for more SEU tolerance. We propose the design of systems with two operating voltages - a lower operating voltage to achieve low energy in a clean environment and a higher operating voltage to achieve SEU tolerance in a radiation intensive environment. The results described in this paper represents work currently in progress. We need to explore further the tradeoffs between the energy efficiency and radiation tolerance by exploring not only the relationship between energy and SEU immunity but also the effect of fault tolerant design on the energy efficiency.

9 Acknowledgements

We acknowledge the discussions that we had with Martin Buehler, Bob Bunker and Bob Schober that helped us to understand the phenomena of SEUs in CMOS. We also acknowledge the discussions with Jim Burr who developed the LVCMOS technology and patiently answered the many questions that we had.

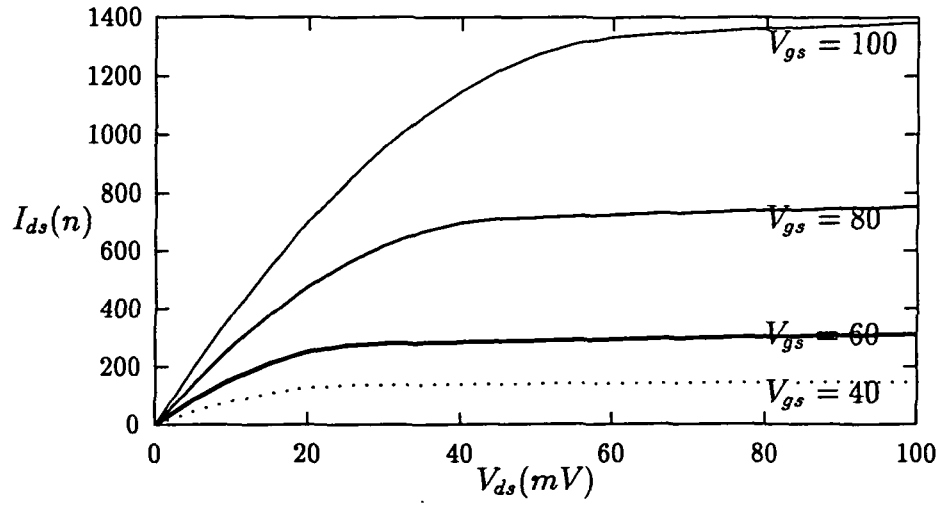

Figure 5: VI characteristic for $V_{bb} = 0.5V$ and 100mV operating voltage

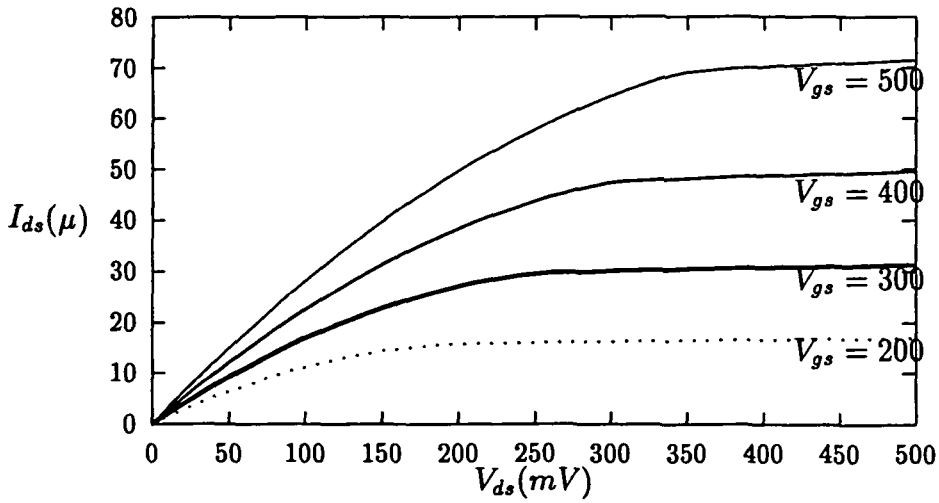

Figure 6: VI characteristic for $V_{bb} = 0V$ and 500mV operating voltage

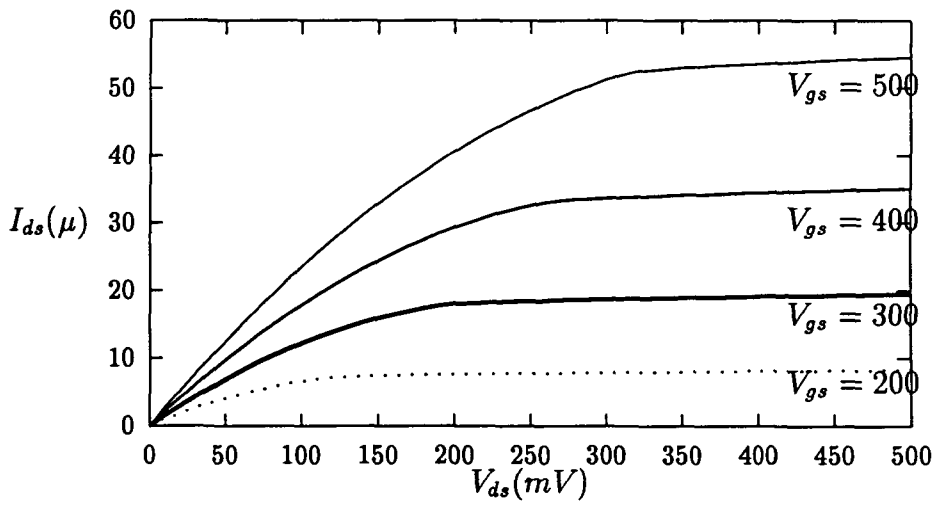

Figure 7: VI characteristic for $V_{bb} = 0.5V$ and 500mV operating voltage

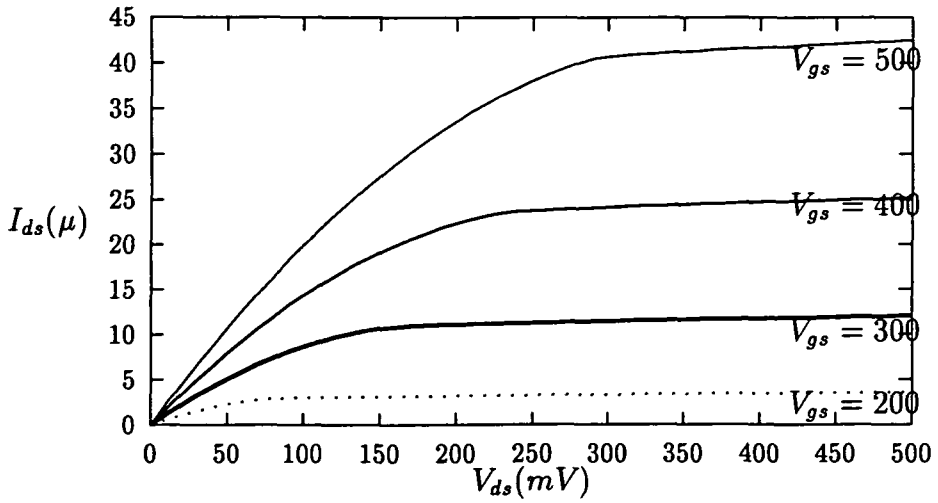

Figure 8: VI characteristic for $V_{bb} = 1V$ and 500mV operating voltage

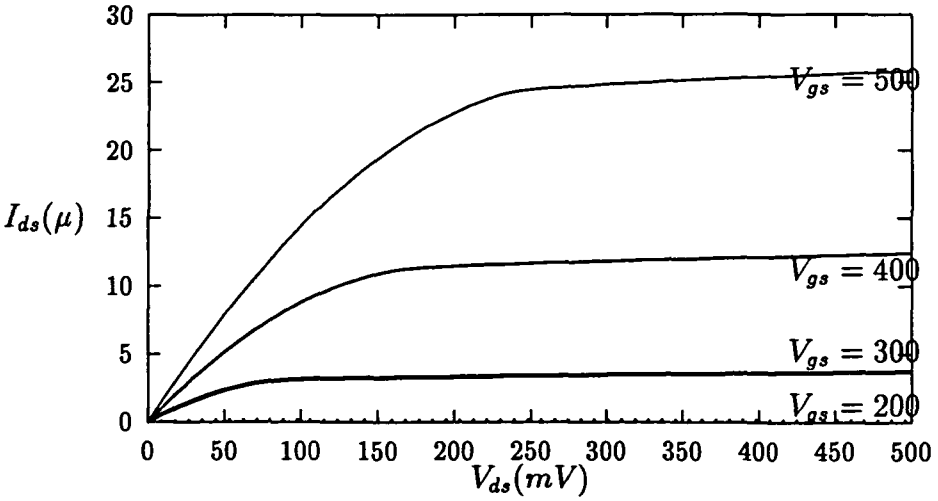

Figure 9: VI characteristic for $V_{bb} = 2V$ and 500mV operating voltage

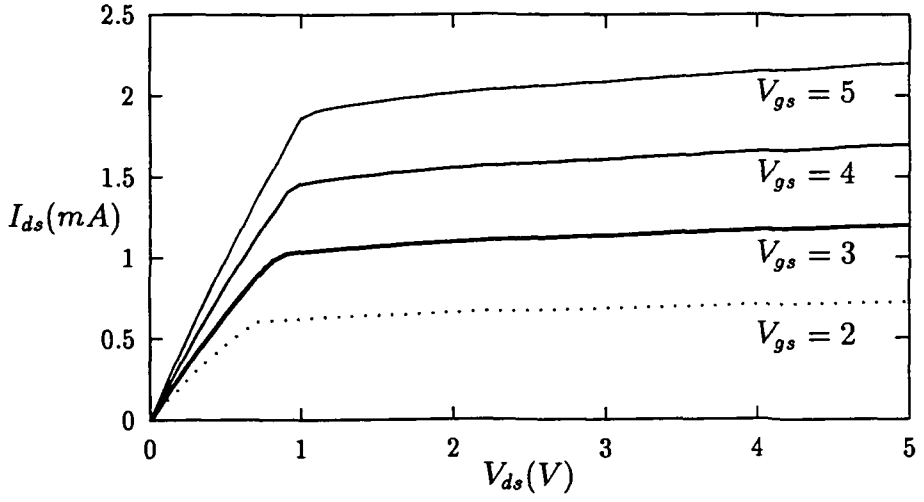

Figure 10: VI characteristic for $V_{bb} = 0V$ and 5V operating voltage

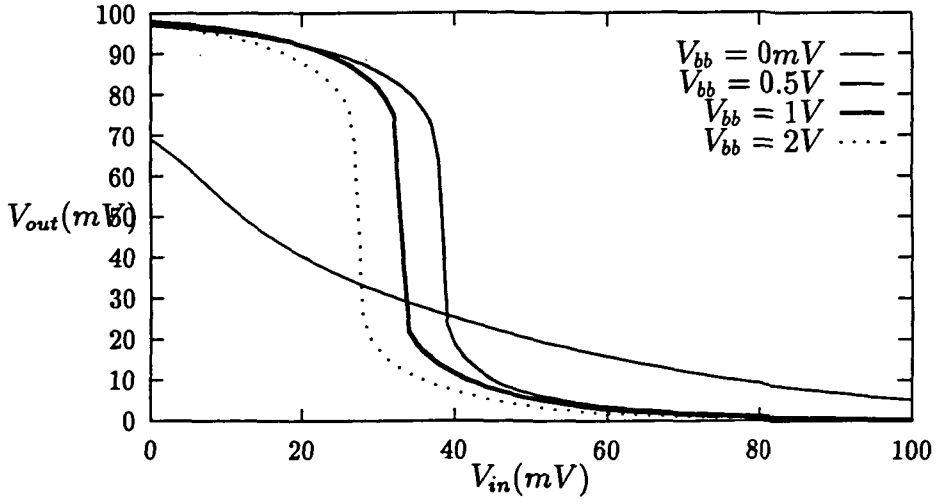

Figure 11: Inverter characteristic for 100mV operating voltage

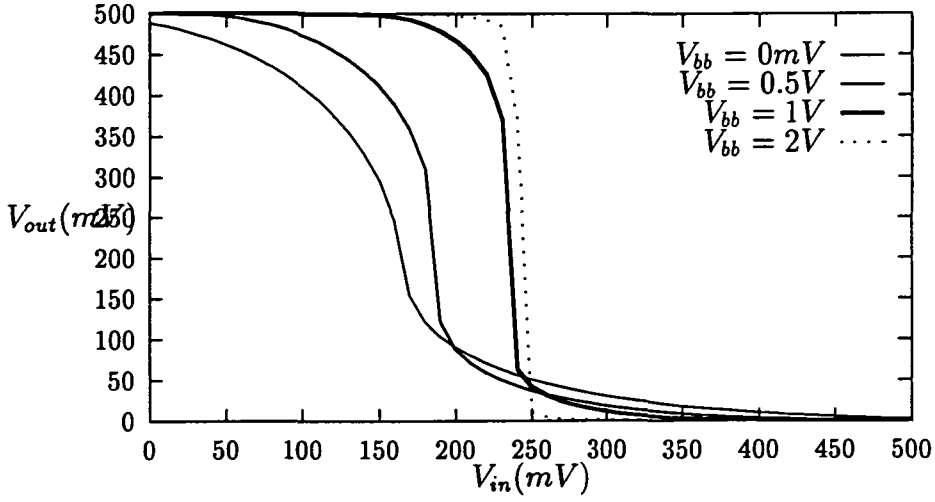

Figure 12: Inverter characteristic for 500mV operating voltage

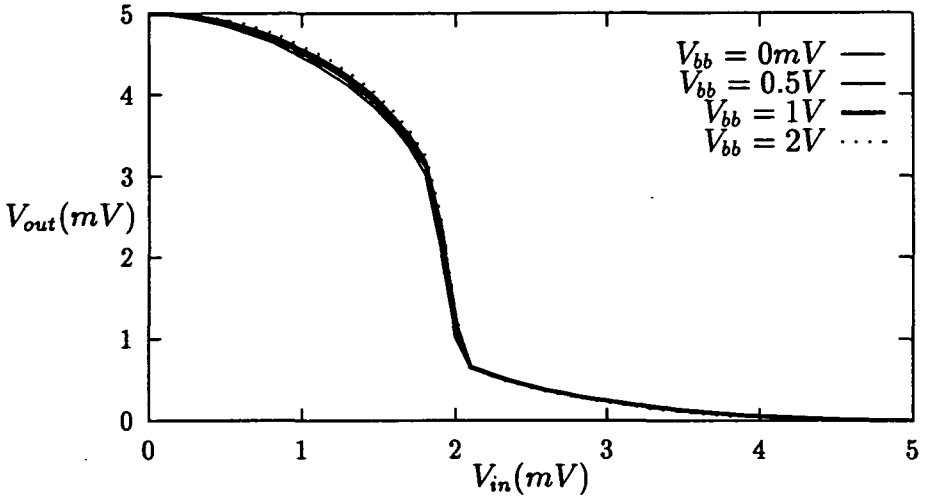

Figure 13: Inverter characteristic for 5V operating voltage

References

[1] N. Weste and K. Eshraghian, *Principles of CMOS VLSI Design*, Addison-Wesley, Reading, MA. 1988.

[2] J. Burr and A. Peterson, "Energy considerations in multichip-module based multiprocessors," *IEEE International Conference on Computer Design*, 1991.

[3] M. Buehler and B. Blaes, "Alpha-particle Sensitive Test SRAMs," *IEEE Transactions on Nuclear Science*, Vol. 37, No. 6, Dec 1990.