F. H. Gaensslen

# Geometry Effects of Small MOSFET Devices

*The effects of diminishing MOS inversion channel length or width on device characteristics are discussed. As opposed to the geometric device size, an "electric device size" is established by normalizing all dimensions on an appropriately chosen depletion layer width. It is shown how this "electric size" governs the intensity of geometry effects. DC device modeling methods are reviewed with respect to their ease of application to electrically small devices. Finally, means for reduction of geometry effects are considered.*

## Introduction

The nonexistence of ideal, step-like changes at the edges of a MOSFET channel as discussed in [1] gives rise to geometry effects. In other words, transitional regions of finite width limit the spatial resolution. Because of their three-dimensional nature, geometry effects are mathematically hard to describe in the general case. The inversion channel geometry of a standard, planar MOSFET device is rectangular. In principle, this fact gives rise to two kinds of geometry effects. They can be studied independently in two specific cross sections of properly shaped devices, one in the direction of and the other perpendicular to the channel length direction. Consequently, these effects are referred to respectively as "short channel" and "narrow channel" effects.

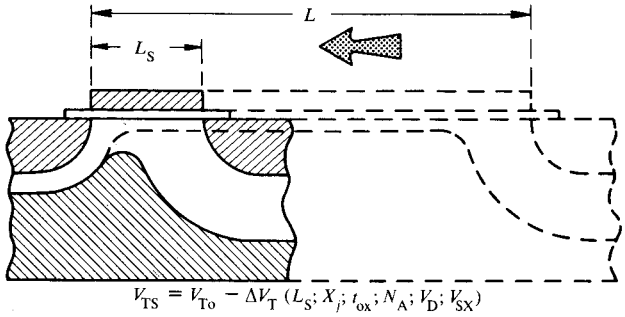


**Figure 1** Schematic device cross sections in channel length direction to illustrate the origin of short channel effects.

The following two sections describe short and narrow channel effects from a phenomenological point of view. In the fourth section the qualitative influence of geometry effects on various MOSFET channel characteristics is discussed. Then the relationship between physical device size and geometry effects is illustrated. The dc modeling of electrically small MOSFET devices is outlined in the sixth section, followed by a discussion of device designs with reduced geometry effects.

## Short channel effects [2-7]

The cross-sectional view of Fig. 1 illustrates what happens to a given n-channel device of channel length $L$, when its length is gradually reduced to $L_S$. The potentials assumed for source, gate, drain and substrate are 0, 0, 0.1 and $-V_{SX}$ volts, respectively. To isolate the effects of channel shortness, three-dimensional interaction has to be avoided; therefore the device is assumed to be very wide.

With the device turned off, a potential barrier keeps the two reservoirs of mobile carriers called source and drain separated from each other. As the channel length $L$ goes towards $L_S$, the space charge layers surrounding source and drain first touch each other and then start to merge. By this process the intervening potential barrier which determines the threshold of the device is lowered. Within the remaining channel portion the two transitional regions in front of source and drain have totally displaced the main portion of the channel in the center, the object of the

usual one-dimensional device description. Within the semiconductor the resolution limits, so to speak, set by the assumed parameter conditions have been exceeded. The new threshold voltage $V_{TS}$ for the short channel can be viewed as the large (*i.e.*, long and wide) channel threshold $V_{To}$ diminished by a correction term $\Delta V_T$ of the variables as shown in Fig. 1 at the bottom. With a decrease in channel length or substrate doping density the correction term increases. It is also enhanced by a greater diffusion depth, gate oxide thickness, or potential across the source and drain junctions.

In addition to the above decrease in threshold voltage, other second order effects also occur. With decreasing channel length but unchanged operating voltages the electric field intensities within the device increase. This can lead to changes in mobility or transconductance, to avalanching and to threshold voltage shifts caused by trapping of hot electrons in the gate dielectric.

### Narrow channel effects [8-10]

Another schematic cross-sectional view of a planar MOSFET device, this one in channel width direction, is shown in Fig. 2. The potentials assumed here for gate, substrate and channel are $V_{To}$, $-V_{SX}$ and 0 volts, respectively. The device is considered to be sufficiently long to eliminate three-dimensional interaction and to separate the narrow channel effects. As the channel width $W$ decreases towards $W_N$ the two depletion layer edge profiles start to overlap, causing a reduction of depletion layer depth underneath the remaining gate width. This reduction signals an increased gate voltage requirement for maintaining a given effective concentration of carriers per unit width within the channel. In other words, the long and wide or, briefly, large channel threshold voltage $V_{To}$ has to be increased by an amount $\Delta V_T$, as pointed out in Fig. 2, to arrive at the narrow channel threshold voltage $V_{TN}$. With a reduction in channel width and in substrate doping concentration the necessary correction term increases. Any increase in gate or field oxide thickness or in potential between channel and substrate causes a larger deviation from the wide or large channel threshold voltage value $V_{To}$.

### Qualitative short and narrow channel characteristics

Whether a device description is simple or complex, the threshold voltage always plays a dominating role. In general terms, the threshold voltage can be defined as the standardized onset of strong conduction. Several standards are in use, but fortunately they do not differ very widely. For a sufficiently long and wide n-channel device made within a substrate with uniform doping, the threshold voltage can be described by the summation of the following three terms:

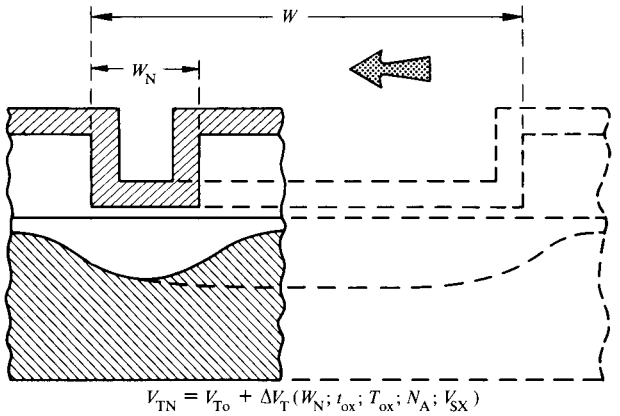


**Figure 2** Schematic device cross sections in channel width direction for discussing the cause of narrow channel effects.

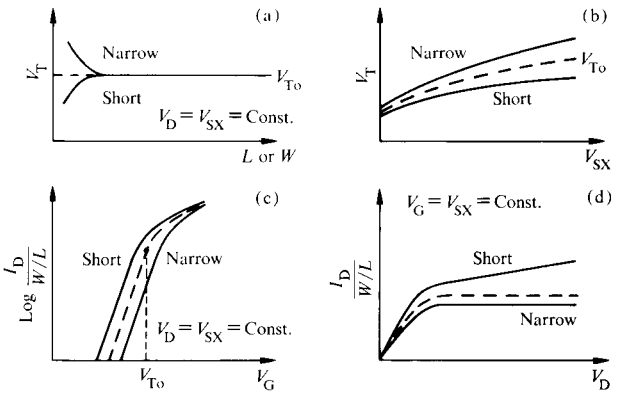


**Figure 3** Qualitative device characteristics showing separately the effects of an electrically short or narrow channel: (a) threshold voltage as function of channel length or width, (b) substrate sensitivity, (c) device transfer characteristic, and (d) drain characteristics.

$$V_{To} = V_{FB} + 2\phi_F + \frac{t_{ox}}{\epsilon_{ox}}\sqrt{2\epsilon_{Si}qN_A(|V_{SX}| + 2\phi_F)}.$$

The first two terms, the flatband voltage and the band bending in silicon, do not change with channel geometry. The last term is due to the fixed depletion charge in silicon. For a short and narrow MOS transistor this term becomes a function of channel length and width, drain potential, diffusion depth and field oxide thickness.

The qualitative influences of short and narrow channel effects on some enhancement mode device characteristics are separately shown in the four portions of Fig. 3. The same qualitative trends are also observed for depletion mode MOS channels. The respective large channel characteristics, those for a simultaneously long and wide

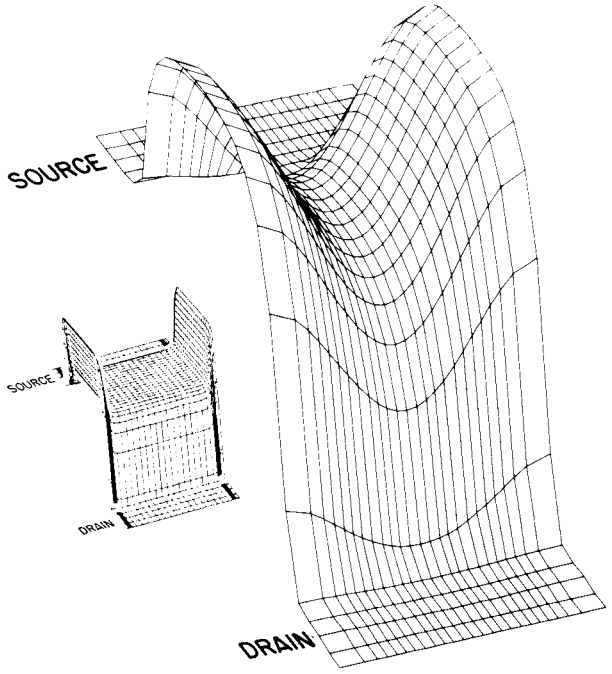


**Figure 4** Pseudo-three-dimensional view of the conduction band edge taken along the silicon/silicon dioxide interface of a simultaneously short and narrow MOSFET channel. Inset to the left shows similar view of a long and wide device at a different scale.

channel, are indicated in dashed lines. With decreasing length or width of the channel, as discussed above, the device threshold voltages are respectively lowered or raised. These relationships are plotted in Fig. 3(a).

A negative substrate potential of increasing magnitude, as indicated in Fig. 3(b), lowers the substrate sensitivity for a short channel device, while the one for a narrow channel is increased.

As shown in Fig. 3(c), the transfer characteristics are only shifted towards higher or lower currents for a short or narrow channel, respectively. The slope of the exponentially varying subthreshold current is unaffected by the altered barrier height due to the change in geometry.

The device output characteristics in Fig. 3(d) show for the short channel case increased drain currents at a given gate drive and some skew due to the dependency of the potential barrier on the drain voltage. The respective narrow channel output curve is only somewhat lower in current due to an increase in threshold voltage.

Normally only moderate short and narrow channel effects are tolerable. Therefore the superposition principle can be applied to solve the three-dimensional problem. A simultaneously short and narrow device can be approximated rather closely by applying together a short channel correction and a narrow channel correction to the long and wide channel case. A pseudo-three-dimensional view of the conduction band edge taken along the silicon/silicon dioxide interface for a simultaneously short and narrow MOSFET channel is shown in Fig. 4. The calculations for this drawing were done by computer using a numerical, two-dimensional device simulation program [11]. In front of the source a mathematical saddle point is apparent with curvatures of opposite sign in planes orthogonal to each other. In this view it can be readily seen how the narrowing of the channel raises the current-controlling potential barrier or the threshold. Conversely, shortening of the channel results in potential barrier lowering. Corner effects have very little influence on the threshold-determining potential barrier of the device which exists in the center of the channel. The drain corners, however, can be expected to enhance the second order high field effects mentioned earlier. Figure 4 also shows as an inset on the left a similar view of a long and wide device at a different scale.

### Geometric *versus* electric device size

In contrast to introductory statements made sometimes in papers on the subject of small MOS field-effect transistors, high lithographic resolution is absolutely no prerequisite for intense short or narrow channel effects. Based on the scaling principle [12] a simple argument can be made to prove this point. Given, an electrically short device with its threshold voltage reduced by a certain percentage relative to the large channel threshold $V_{T_0}$. If this device is scaled up properly (scaling factor $\alpha < 1$) by the following linear transformation of all distances $s$, voltages $V$, acceptor densities $N_A$ and currents $I$,

$$s' = \frac{s}{\alpha}; \qquad V' = \frac{V}{\alpha}; \qquad N_A' = \alpha N_A; \qquad I' = \frac{I}{\alpha},$$

the new, larger device suffers from essentially the same relative short channel effect. The primed quantities apply to the scaled-up device. The same holds true in the width direction for the narrow channel effect. Obviously it is not the mere geometric channel length or width that controls the relative intensities of these effects, but rather their "electric," "numeric" or "characteristic" counterparts. An electric length, in this context, has to be defined as the geometric length normalized on an appropriately chosen depletion layer width. In this way both the substrate doping concentration and the operating voltage range of the device can be brought into play.

To provide a meaningful definition of an "electric device size" in all three dimensions the same kind of normalization has to be applied also in the direction of the channel width.

Figure 5 shows a possible normalization approach. To simplify the discussion a uniformly doped substrate is chosen. Shown at the top are three substrate sensitivity curves for devices with channel geometries as indicated to the right. Points 1 and 2 define operating conditions referred to below.

For normalization in length direction (see condition 1 in Fig. 5) the depletion layer width $\bar{x}_D$ is taken with source and drain diffusions at the highest operating potential, called $V_D$. Since the built-in potential $V_{Bi}$ is a constant and does not scale, the exact substrate voltage for the scaled-up device $V'_{SX}$ has to be derived from the following relationship:

$$V'_{SX} + V_{Bi} = 1/\alpha(V_{SX} + V_{Bi}).$$

The electric channel lengths, $l$, for the original and the scaled-up device are numerically equal. As a consequence the two devices show the same short channel effect. The electric channel length has naturally no dimension and runs from zero to infinity. For the example drawn in Fig. 5 it equals one.

The same kind of reasoning applies in width direction (see condition 2 in Fig. 5). Because of the potential of source and drain the inversion channel also carries a potential of $V_D$. Assuming that the band bending in silicon ($2\phi_F$) about equals the built-in potential ($V_{Bi}$) in Condition 1, the same expression defines the depletion layer width needed here for normalization. The electric channel width, $w$, is also dimensionless and runs from zero to infinity, with the specific example shown at one.

In order to compare the electric size of devices, the same depletion layer width $\bar{x}_D$ has to be used for normalization in all three device dimensions. In summary, from a device design point of view, geometry effects do not get worse by downward scaling; they pose the same relative problem at any given dimension.

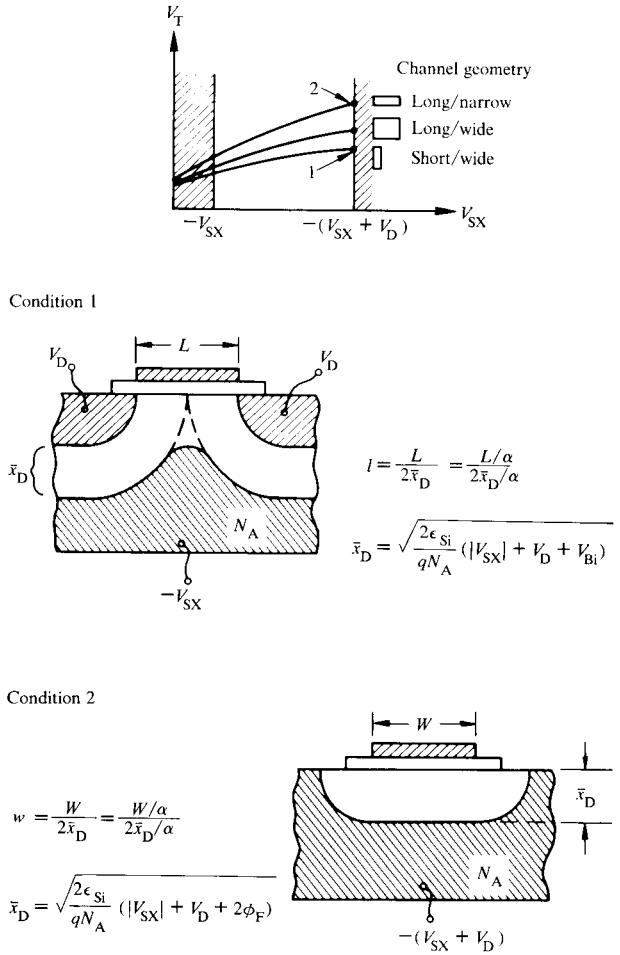


**Figure 5** Illustrations for the normalization of device dimensions in channel length and width directions, defining geometric *versus* "electric device size." Condition 1: Short/wide channel geometry. Condition 2: Long/narrow channel geometry.

## DC modeling of electrically small MOSFET devices

Modeling of MOSFETs and other semiconductor devices is commonly done by two fundamentally different methods: one uses a piecewise, analytical description of the device behavior and the other a continuous, numerical simulation.

The piecewise modeling employs separate equations within certain operating regimes such as subthreshold, linear, saturation and breakdown regions by accounting for the most dominant physical effect. Because of overlapping, this approach leads to continuity problems at the boundaries of these regimes. Once higher-order effects, such as geometry effects, are included, analytical models quickly get more complicated. They are not flexible at all with respect to conceptual device design changes. With their modest memory and computing requirements they are mainly geared towards application in circuit analysis programs. In designing devices they are most effective in exposing trends caused by wide-ranging physical parameter changes.

As things stand with the analytical approach the threshold voltage is of pivotal importance. The following two figures taken from [5] serve to highlight the complexities that arise from rigorously including only the short channel effect in the threshold voltage equation for a uniformly doped device. Figure 6(a) shows an actual, two-dimensional charge distribution in a short channel device. It is approximated by the simplified cross section shown in Fig. 6(b) with the assumption of overall charge neutrality.

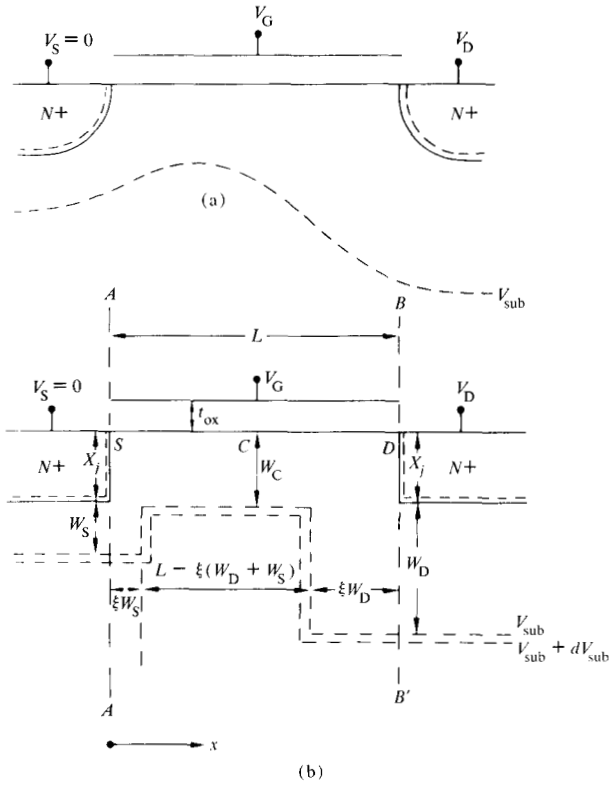


**Figure 6** (a) Schematic diagram of depletion charge profile of an *n*-channel IGFET. (b) Idealized equivalent charge profile for which the closed-form expression for the threshold voltage is obtained. Reprinted with permission from *Solid State Electron*. **16,** H. S. Lee, "An Analysis of the Threshold Voltage for Short-Channel IGFET's," Pergamon Press, Ltd. (1973).

$$V_{th} = V_{th\,long} - \frac{\xi}{C_{0x}L}\{A_0 + A_1X_j + A_2X_j^2\}$$

where

$$V_{th\,long} = \left(\Phi'_{MS} - \frac{Q_{ss\,eff}}{C_{0x}}\right) + B\phi_{FP} + \frac{1}{C_{0x}}\sqrt{[2K_{si}\epsilon_0 qN_A(|V_{sub}| + B\phi_{FP})]}$$

$$A_0 = K_{si}\epsilon_0\left\{\sqrt{(|V_{sub}| + B\phi_{FP})}\sqrt{(|V_{sub}| + KV_d + V_{Bi})} - (V_{sub} - B\phi_{FP}) + (KV_d + V_{Bi} - B\phi_{FP})\ln\left[\frac{\sqrt{(|V_{sub}| + B\phi_{FP})} + \sqrt{(|V_{sub}| + KV_d + V_{Bi})}}{\sqrt{(KV_d + V_{Bi} - B\phi_{FP})}}\right]\right\}$$

$$A_1 = \sqrt{(2K_{si}\epsilon_0 qN_A)}\{\sqrt{(|V_{sub}| + V_{Bi})} + \sqrt{(|V_{sub}| + KV_d + V_{Bi})} - \sqrt{(V_{Bi})} - \sqrt{(KV_d + V_{Bi})}\}$$

$$A_2 = -qN_A \ln\left\{\frac{\left[X_j + \sqrt{\left(\frac{2K_{si}\epsilon_0}{qN_A}(|V_{sub}| + KV_d + V_{Bi})\right)}\right]\left[X_j + \sqrt{\left(\frac{2K_{si}\epsilon_0}{qN_A}(|V_{sub}| + V_{Bi})\right)}\right]}{\left[X_j + \sqrt{\left(\frac{2K_{si}\epsilon_0}{qN_A}(KV_d + V_{Bi})\right)}\right]\left[X_j + \sqrt{\left(\frac{2K_{si}\epsilon_0}{qN_A}V_{Bi}\right)}\right]}\right\}$$

**Figure 7** Closed-form threshold equation including short channel effect. Reprinted with permission from *Solid State Electron*. **16,** H. S. Lee, "An Analysis of the Threshold Voltage for Short-Channel IGFET's," Pergamon Press, Ltd. (1973).

The analytical short channel threshold voltage result, as shown in Fig. 7, reduces the long channel threshold by a correction term which essentially amounts to a product of $1/L$ with a second order polynomial in $X_j$. The expressions for the coefficients $A_i$, shown at the bottom of Fig. 7, spoil this otherwise simple result. More simplifying assumptions usually yield simpler results, as shown by Yau [6] and Jeppson [9].

A more pragmatic but very effective approach with physical background has been advanced by Troutman and Fortino [7]. They invoke what is called a virtual cathode in an electron tube, do some potential barrier mechanics, and come up with the following analytical description of the short channel threshold voltage for a given substrate bias:

$$V_{TS} = V_{To} - (\alpha + \beta V_D).$$

The correction term takes here only a first order polynomial in $V_D$. The proximity coefficient $\alpha$ corrects only for the closeness of the two unbiased diffusion pockets, while the penetration coefficient $\beta$ gauges the additional drain bias dependency. For a given technology and substrate bias both coefficients can be determined from threshold *versus* drain voltage plots with channel length as parameter.

For numerical simulation of short or narrow channel effects with a two-dimensional program, the respective device cross section, as shown in Fig. 1 or Fig. 2, has to be mapped into an array of discrete elements. Over the range of this device cross section model the following set of partial differential equations has to be iteratively solved under the constraints of a set of boundary and input potential conditions:

$$\text{div grad } \Phi = -\frac{q}{\epsilon}(p - n + N_D^+ - N_A^-), \tag{1}$$

$$\text{div } \mathbf{j}_p = q(G_p - R_p), \tag{2}$$

$$\text{div } \mathbf{j}_n = q(R_n - G_n), \tag{3}$$

$$\mathbf{j} = \mathbf{j}_p + \mathbf{j}_n, \tag{4}$$

with

$$\mathbf{j}_p = q\mu_p p \text{ grad } \Phi + qD_p \text{ grad } p,$$

$$\mathbf{j}_n = q\mu_n n \text{ grad } \Phi - qD_n \text{ grad } n.$$

The first equation (Poisson) gives the relationship between charge and potential throughout the volume, while the next two force continuity of hole and electron current densities. The total transport of charge is summed up in the last equation. The program output consists for instance of discrete, two-dimensional distributions of mobile and immobile charges, of currents and potentials. Over the four operating regimes of the device, as mentioned above, the program responds to a set of input potentials with output current values as shown in Fig. 8 for a simulation of short channel effects [13]. Two measured

transfer curves are shown for each device with the drain voltage at 0.1 and 3.0 volts, respectively. Two-dimensional simulation results are indicated by crosses. The shortest of the three devices shown (channel length $L_1$) is easily identified by its drain voltage dependence (penetration effect) and by a general curve shift towards higher drain currents (proximity effect). Just by supplying the proper channel length input, a long channel simulation is easily adapted to a short channel case. Similar two-dimensional simulations of narrow channel effects have been published by Noble and Cottrell [10].

With diminishing electric channel length the drain voltage dependent penetration becomes more and more influential. This sensitivity can be very useful in establishing the effective length of a device channel very accurately, as demonstrated by Fig. 8. With all other parameters held constant, a fixed drain voltage step (0.1 to 3.0 V here) causes a separation in voltage between the exponential portions of the two transfer characteristics which is solely dependent upon the channel length. Matching this separation with two-dimensional simulations then means using the proper effective channel length. Hence numerical device simulation can be used effectively to refine the results of other physical channel length measurements. It cannot be overemphasized that channel length and width of physically small devices are very hard to determine, yet when it comes to describing the device behavior they are parameters of paramount importance. Geometry effects can be exploited to establish processing biases in length and width directions very accurately.

Numerical computer simulation is an indispensable tool for device design and understanding. It can be very flexible, at least in principle, with respect to conceptual and other changes. In disclosing trends of derived parameters, such as threshold voltages, two-dimensional simulations are less powerful.

### Device designs with reduced geometry effects

Short and narrow channel effects, even in view of their opposing signs, cannot be used jointly to compensate for geometric threshold voltage variations in integrated circuits. For self-aligned device designs the channel length and width tolerances are uncorrelated since they depend upon different masking steps. Separately they can be exploited to enhance circuit performance somewhat. But minimization of geometry effects by conceptually better device designs seems to be a most promising approach.

Narrow channel effects can be eliminated altogether by rather area-consuming circular or enclosed device structures. In this kind of device structure one of the diffusions

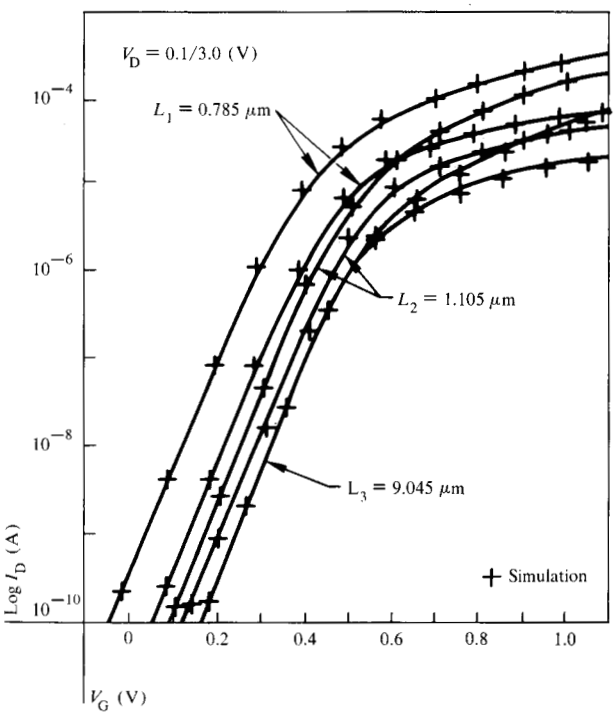


**Figure 8** Comparison of measured transfer characteristics (for drain voltages of 0.1 and 3.0 V) with calculated results of two-dimensional, numerical simulations for three MOSFETs of greatly different channel lengths [13].

and the channel surround the second diffusion. The original *V*-groove transistor holds the same advantage. In a normal, open device structure the cross section geometry and impurity profile have to be shaped in order to minimize the transitional region and thereby the narrow channel effect. Means towards this end are elements like recessed field oxide structures, field and channel implantations, or what is called a field shield. Numerical computer simulation is certainly the only tool that allows assessment of the overall behavior of device structures of this complexity.

To enhance the electric length of a given channel the separation between the depletion layer boundaries surrounding source and drain has to be increased. A similar array of approaches directed towards this goal is available here: ion implantation into the channel, the use of low doped source and drain diffusions, and recessing the channel into the substrate [14]. A combination of all three remedies should yield the best properties in regard to short channel effects and at the same time should also be rather immune to secondary high field effects such as avalanching and hot electron trapping, mentioned at the end of the section entitled "Short channel effects."

## Summary

The finite dimensions of practical MOS channels give rise to geometry effects of varying degree. Transitional regions along the edges of a channel limit the spatial resolution, causing what are labeled as short and narrow channel effects.

It has been shown that the relative intensities of these effects do not depend upon the geometric but rather on the "electric device size." Because of their three-dimensional nature geometry effects observed on electrically small MOSFET devices are mathematically hard to describe. The more elaborate, numerical simulation is clearly superior in exposing the impact of geometry effects on the characteristics of advanced, nonplanar device designs with multiple ion implantations.

Geometry effects can be exploited very effectively in conjunction with two-dimensional numerical simulations to establish the final channel dimensions and hence the overall processing biases. This is possible because of the rapidly increasing sensitivity of these effects to reductions in "electric channel size."

In digital circuit applications geometry effects are mostly detrimental since they merely widen the overall operating tolerances of devices. Even in view of their opposing signs, channel length and width variations cannot be used to compensate each other because they are statistically independent.

Looking towards the future, we certainly have to live with mild geometry effects, as we had to in the past. Where they cannot be tolerated they have to be controlled either by conceptually better device designs or at the expense of additional area.



## Acknowledgments

I would like to thank my colleagues, especially R. C. Jaeger and R. Merrill, for helpful discussions. The pseudo-three-dimensional simulations were done by R. Merrill using the two-dimensional simulation program of [11].


## References


1. F. H. Gaensslen, "Geometry Effects of Small MOSFET Devices" (invited talk, IEEE International Electron Device Meeting, Washington, DC), *IEDM Tech. Digest*, 512–515 (1977).
2. G. F. Neumark and E. S. Rittner, "Transition from Pentode- to Triode-like Characteristics in Field Effect Transistors," *Solid State Electron.* **10,** 299–304 (1967).
3. G. A. Armstrong, J. A. Magowan, and W. D. Ryan, "Two-Dimensional Solution of the D.C. Characteristics for the M.O.S.T.," *Electron. Lett.* **5,** 406–408 (1969).
4. M. B. Barron, "Computer Aided Analysis of Insulated Gate Field Effect Transistors," *Report No. 5501-1*, Stanford Electronics Laboratories, Stanford, CA (1969).
5. H. S. Lee, "An Analysis of the Threshold Voltage for Short-Channel IGFET's," *Solid State Electron.* **16,** 1407–1417 (1973).
6. L. D. Yau, "A Simple Theory to Predict the Threshold Voltage of Short-Channel IGFET's," *Solid State Electron.* **17,** 1059–1063 (1974).
7. R. R. Troutman and A. G. Fortino, "Simple Model for Threshold Voltage in a Short-Channel IGFET," *IEEE Trans. Electron Devices* **ED-24,** 1266–1268 (1977).
8. K. E. Kroell, "Geometry Effects on Field Effect Transistors," presented at European Solid-State Device Research Conference (ESSDERC), Munich, Germany, Sept. 1973.
9. K. O. Jeppson, "Influence of the Channel Width on the Threshold Voltage Modulation in MOSFET's," *Electron. Lett.* **11,** 297–299 (1975).
10. W. P. Noble and P. E. Cottrell, "Narrow Channel Effects in Insulated Gate Field Effect Transistors," *IEDM Tech. Digest*, 582–586 (1976).
11. P. E. Cottrell and E. M. Buturla, "Steady State Analysis of Field Effect Transistors Via the Finite Element Method," *IEDM Tech. Digest*, 51–54 (1975).
12. R. H. Dennard, F. H. Gaensslen, L. Kuhn, and H. N. Yu, "Design of Micron MOS Switching Devices," *IEDM Tech. Digest*, 168–170 (1972).
13. F. H. Gaensslen and A. LeBlanc, "Simulation of Micron MOS Device Characteristics," presented at European Solid-State Device Research Conference (ESSDERC), Munich, Germany, Sept. 1973.
14. K. Natori, I. Sasaki, and F. Masuoka, "An Analysis of the Concave MOSFET," *IEEE Trans. Electron Devices* **ED-25,** 448–456 (1978).



*Received November 7, 1978; revised May 2, 1979*



*The author is located at the IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598.*