

PERFORMANCE ADVANTAGES OF SUBMICRON SILICON-ON-INSULATOR DEVICES FOR ULSI

JAMES C. STURM

Department of Electrical Engineering, Princeton University, Princeton, NJ 08544

ABSTRACT

In this paper the various performance advantages of SOI and SOS structures for submicron ULSI circuits will be described. In addition to the traditional speed and radiation-hardness advantages, there are several significant advantages of thin-film SOI compared to bulk structures for the submicron scaling of MOS transistors. These advantages include short-channel threshold voltage stability, improved sub-threshold slope, increased saturation current, and reduced hot electron effects. Both theory and data from several groups will be presented to illustrate these effects. Since these advantages all fall in areas that are critical limits to device engineering and scaling in the submicron regime, the motivation for using SOI should be even stronger as devices are scaled to smaller dimensions in the future. Consideration of SOI or SOS on the ULSI scale will require a technology capable of low defect films with a film thickness of 1000 Å or less, however.

The prospects for minority carrier devices such as bipolar transistors for BI-MOS will also be discussed. Both device structure (lateral vs. vertical) and material quality are issues that must be addressed.

INTRODUCTION

The traditional advantages of Silicon-on-Insulator (SOI) or Silicon-on-Sapphire (SOS) for MOS transistors are well known. These advantages of radiation hardness, device isolation, and low capacitances have driven the application of SOI/SOS processes where these virtues are required.

In this paper we will not focus on these traditional niche advantages, but rather discuss the advantages of SOI/SOS for submicron Ultra-Large-Scale Integrated circuits as a whole. It will be shown that through the use of ultra-thin films, SOI MOS transistors potentially offer several substantial advantages with respect to their bulk counterparts. These areas of performance favoring SOI include short channel threshold voltages, subthreshold slopes, current drive, and hot electron effects. In addition to these performance advantages, various challenges and problems must also be addressed. These include both material challenges for less than 1000Å films and device challenges for bipolar structures.

SHORT CHANNEL THRESHOLD VOLTAGES

The increase in integration levels of MOS technology has been marked by a steady decline of both lateral and vertical dimensions. For MOS transistors, this corresponds primarily to a thinner gate oxide (t_{ox}) and a shorter gate length (L). Assuming a scaling factor s , we can write that t_{ox} and L scale as $1/s$.

$$t_{ox} \rightarrow t_{ox} \quad (1)$$

$$L \rightarrow \frac{L}{s} \quad (2)$$

To shrink depletion regions along with device geometries (to prevent punchthrough, etc.), the decline in dimensions has been marked by an increase in channel doping (N_A).

$$N_A \rightarrow sN_A \quad (3)$$

Using a simple formula for depletion region width (X_d)

$$X_d = \left(\frac{2\epsilon_{si}(V + \phi_{bi})}{N_A q} \right)^{1/2} \quad (4)$$

it can easily be seen that the scaling of X_d depends on power supply voltage as well as N_A . However, a combination of technological (5 V standard) and fundamental (built-in voltages, subthreshold slopes) factors have caused power supply voltages to remain relatively constant despite shrinking dimensions. Besides increased electric fields, a constant power supply implies that X_d will scale as $N_A^{-1/2}$, or

$$X_d \rightarrow \frac{X_d}{\sqrt{s}} \quad (5)$$

This is an important result of constant voltage (CV) scaling: channel length decreases faster than depletion regions. This relative increase in the size of depletion regions has a marked effect on threshold voltages, and can be understood according to a model first proposed by Yau [1].

In a MOS transistor, the voltage on the gate serves to support both the charge in the channel as well as that in the substrate depletion region under the channel region (fig. 1a). At threshold, the inversion charge is virtually zero and the gate supports only the space charge in the substrate. By calculating the field required to support this charge one can easily determine the threshold voltage. In a short channel device depletion regions will occupy a relatively larger portion of the device as just discussed. Hence a considerable part of this space-charge can be partitioned to be supported by the source and drain regions rather than by the gate (fig. 1b). This reduction in charge supported by the gate leads to a reduced threshold voltage. The threshold voltage then becomes dependent on gate length, a circuit designer's nightmare. Variable threshold voltages will affect noise margin, power, speed, and are not tolerable for circuit design.

Shown in Figure 2 is a short channel transistor fabricated in a thin SOI film. Note that if the film is thinner than the depletion region thickness, a fractionally smaller part of the depletion region charge is supported by the source or drain than in a conventional bulk FET. This should lead to more stable threshold voltages as channel length is decreased. In Figure 3 are data of Colinge [2] for a n-channel MOS transistors fabricated in SOI films of only $0.1\mu\text{m}$ thickness. The film thickness is sufficiently small to insure that the film is depleted completely through at threshold. Shown for comparison are similar transistors fabricated in bulk substrates. Note that the threshold reduction is indeed much smaller in the SOI. The results can be accurately modeled using either Yau's model or a 2-D device simulator [3]. It should be

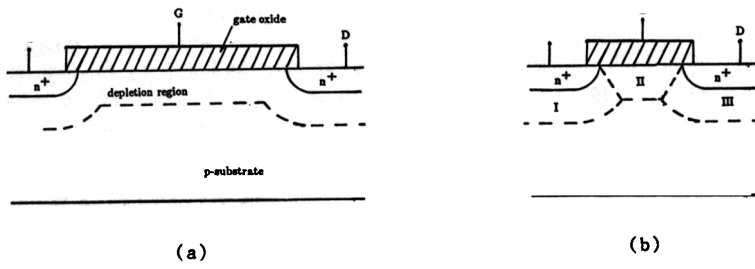


Fig. 1. Cross section of depletion regions in long-channel (a) and short-channel (b) bulk MOS devices

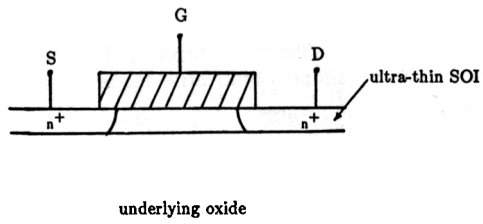


Fig. 2. Cross section of short-channel SOI device

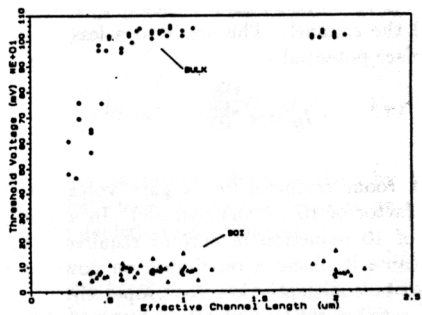


Fig. 3. Threshold voltage vs channel length for SOI transistors in 1000Å films and for simultaneously processed bulk transistors [2]

noted that when an SOI film is "depleted through", the threshold voltage becomes a function of the substrate or "back-gate" potential, and is usually lower than that of a corresponding bulk transistor [4]. This effect explains the lower V_T for long channels in the data of Colinge.

SUBTHRESHOLD SLOPES

When the gate voltage of a MOS transistor is decreased below the threshold voltage, the transistor current does not drop abruptly to zero as predicted by simple theory, but rather decreases with an exponential dependence on gate voltage. It can be shown that in the subthreshold region the physics of an MOS transistor look like that of a bipolar transistor, with source, channel surface potential (ϕ_{surf}), and drain substituting for emitter, base, and collector, respectively. Relying on bipolar device physics, the drain current should depend on ϕ_{surf} as

$$I_D \propto e^{\frac{q\phi_{surf}}{kT}} \quad (6)$$

In a MOS transistor, one cannot directly control ϕ_s , but only capacitatively couple to it through the gate. Using a simple capacitive model (Fig. 4a), it is straightforward to show a coupling efficiency $\frac{1}{n}$, where

$$\frac{d\phi_s}{dV_G} = \frac{1}{n} \quad (7)$$

$$n = \frac{C_{ox} + C_{dep}}{C_{ox}} \quad [5] \quad (8)$$

C_{ox} and C_{dep} represent the capacitances of the gate oxide and substrate depletion region, respectively. (The effective capacitance due to interface states at the silicon-silicon dioxide interface has been neglected for simplicity.) Any capacitances beside C_{ox} which are connected to the surface channel (such as C_{dep}) are parasitic capacitances which make it difficult for V_G to control ϕ_{surf} . Hence $n \geq 1$ implies a non-ideal coupling between the gate and the channel. This leads to a less steep dependence of I_D on gate voltage than on surface potential.

$$I_D \propto e^{\frac{qV_G}{nkT}} \quad (9)$$

In the best case ($n=1$) at room temperature, a gate voltage swing of 60 mV is required to reduce I_D by a factor of 10 (60 mV/decade). In a more typical situation, $n \sim 2$ and hence a factor of 10 reduction in current requires a 120 mV gate swing. This slope is important because it places a restriction on power supply scaling. For transistor "OFF" current to be sufficiently low for proper circuit operation, the gate voltage swing required below threshold is determined by the subthreshold slope. For example, a subthreshold slope of 120 mV/decade and the circuit requirement of an OFF current $< 10^{-5}$ of the current at threshold would require a threshold voltage of at least 0.8 V.

In an SOI transistor where the film depletes through well below threshold, any field lines from the silicon channel surface must terminate at the conductor under the underlying oxide. Hence, the depletion capacitance of the bulk model is replaced by the series combination of capacitors represented by the SOI film (C_{body}) and the underlying oxide. (Fig. 4b) Because the thickness of the underlying oxide is typically very thick (1000's of Å) compared to the gate oxide thickness (100's of Å), this parasitic capacitance is usually negligible, and hence the coupling of gate voltage to the channel should be very efficient ($n \sim 1$), leading to a subthreshold slope at room temperature of 60 mV/decade. Such an effect has indeed been observed by Colinge for MOSFET's fabricated in ultra-thin SOI films. (Fig. 5) [6]

It should be noted, however, that interface states can add an effective capacitance which will degrade subthreshold slopes. This effect is well documented in bulk MOSFET's for interface states at the gate oxide interface [5]. An effective capacitance C_{it} ($= q^2 N_{it}$ where N_{it} is the interface state density) is caused by interface states. This reasoning can be extended to predict effect of interface states at the back of a thin SOI film on the subthreshold characteristics of the top surface FET. Although the density of interface states at the top silicon-silicon dioxide interface are typically fairly low (owing to the high quality of thermal oxides), states at the lower Si - SiO₂ interface (not a thermally grown interface) can be many times higher in density [7]. The effect can be modelled by the addition of an interface state capacitance C_{it} representing these states at the bottom of the SOI film. (fig. 4).

Because of the potentially high value of C_{body} (representing the capacitance across the thin SOI film), the effect of these interface states can be considerable. While much effort has been focused on the growth of high quality gate oxides on top of SOI material, it can thus be seen that for high performance SOI MOSFET's in ultra-thin films, a high quality lower oxide interface will also be necessary.

SATURATION CURRENT

To a first approximation, the current in a MOS transistor can be expressed as

$$I_D = \frac{W}{L} C_{ox} \mu (V_G - V_T - \frac{V_D}{2}) V_D \quad (10)$$

before saturation, and

$$I_{D,SAT} = \frac{W}{2L} C_{ox} \mu (V_G - V_T)^2 \quad (11)$$

in saturation. (The drain saturation voltage is given by $V_{D,SAT} = V_G - V_T$). It is well known, however, that these expressions are only an approximation because they assume a fixed substrate depletion region charge.

The gate voltage of a MOS transistor supports both the inversion charge in the channel (which carries the drain-source current) as well as the space charge in the substrate depletion region under the channel. Because this depletion region grows as one goes from source to drain, a much smaller fraction of the gate voltage can support electrons in the inversion layer (channel) at the drain than at the source. This effect is also known as the "body effect" which states that threshold voltage rises as the

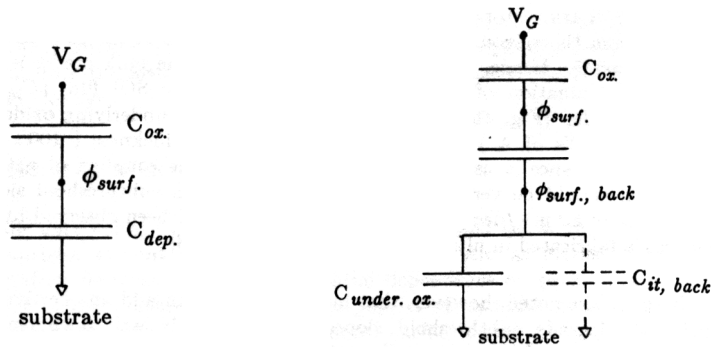


Fig. 4. Capacitance divider model of bulk (a) and SOI (b) FET's in the subthreshold region.

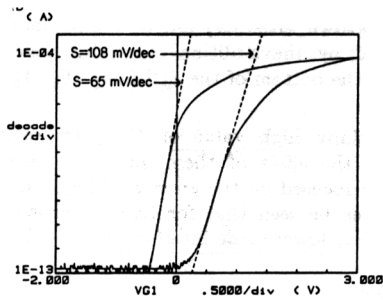


Fig. 5. Subthreshold curves for SOI FET's made in thin (900\AA , 65 mV/dec.) and thick (4500\AA , 108 mV/dec.) SOI films [6]

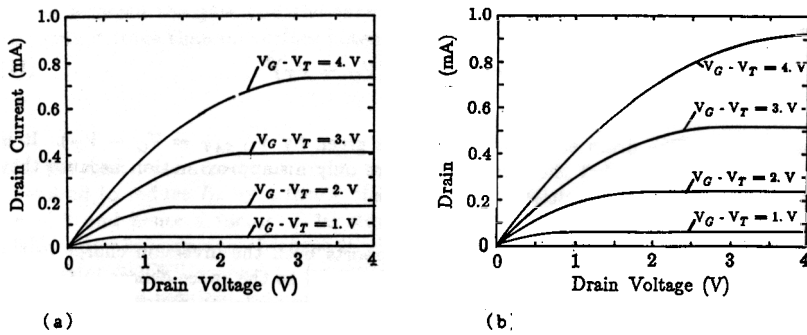


Fig. 6. Modelled performance of bulk (a) and ultra-thin film SOI (b) FET characteristics assuming constant channel mobility.

channel-substrate bias is increased. With these considerations in mind, the usual textbook analysis develops a standard expression which gives a drain current smaller than that predicted by (10) and (11) for all values of gate and drain voltage [8]. Sometimes equation (10) is still used, but a factor $K' (<1)$ is added to account for this discrepancy.

Consider, however, an SOI transistor which depletes through its film before reaching inversion. The "substrate" charge in the depleted film is constant, independent of the channel bias. If the underlying oxide is thick, it can be shown that (10) and (11) are indeed the proper expressions for the transistor drain current -- predicting increased current over an otherwise identical bulk transistor [9]. Also predicted are increased drain saturation voltages. Shown in figure 6 are simulations to demonstrate the magnitude of this effect. We have chosen device structure parameters values typical of a 1- μm technology (200 Å gate oxide, $4 \times 10^{16} \text{cm}^{-3}$ channel doping), but have simulated long channel devices so that short channel effects could be neglected. We have chosen an 800 Å SOI thickness to insure that the film is depleted through well below threshold, and that (10) and (11) describe the SOI drain current. Although the bulk and SOI transistors will have different threshold voltages (for reasons discussed earlier), the curves have been plotted as a function of $V_G - V_T$ (as opposed to simply V_G) to provide a fair comparison. Finally, the same fixed, electron, mobility was used in both devices, and zero substrate bias was assumed in the bulk transistor.

As can be seen in figure 6, a substantial increase in the drain current of the SOI device can be expected. Although these results were derived from a simple 1-D analytical model, two dimensional simulation using the PISCES simulator (10) showed similar results. The expected increase in saturation current of the ultra-thin film SOI device due to this fixed substrate charge effect is 29% over the range $V_G - V_T = 0$ to 4 V (fig. 8a).

There is a second effect which may be expected to further increase the current in ultra-thin film SOI FET devices -- the effect of vertical field on surface mobility. In MOS devices, increasing the vertical field in an MOS devices confines the carriers to a region where more scattering sites are present -- leading to reduced mobilities (fig. 7)[11]. Since the vertical field required for inversion increases in a bulk device as doping is increased, surface mobility decreases for higher doped channels, a well known effect. Since FET's with gate lengths of 0.5 μm will require channel doping on the order of 10^{17}cm^{-3} , it can be seen that a maximum surface mobility of only 500 $\text{cm}^2/\text{V}\cdot\text{s}$ is possible. This reduction becomes even more severe near the drain of a device since the larger channel-substrate depletion region causes larger surface electric field.

In a SOI device in an ultra-thin film with a thick underlying oxide, the surface electric field at inversion can be less than that in a corresponding bulk device. This is because there is less space-charge charge under the channel to support the electric field at the surface. Further, as the channel voltage increases near the drain, this substrate charge (and hence the electric field) remain fixed. Hence, this should first lead to higher effective mobilities in thin SOI devices, and, second, to no mobility degradation as one moves toward the drain. This effect was incorporated into a home-made device simulator which used the gradual channel approximation (good for long-channel devices) and calculated an effective surface mobility for each point along the channel depending on the local vertical electric field. The dashed line in fig. 7 was used for

mobility which corresponds to

$$\mu = 1035 - 365 \log \left(\frac{E}{10^4} \right) \quad (12)$$

where E is in units of V/cm and μ in $cm^2/V-s$.

This effect has little impact near the source and of the device for large gate voltages (where the channel charge is much larger than the substrate charge), but it is always significant near the drain where the channel charge is small and the substrate charge in a bulk device is largest. The predicted bulk and SOI saturation currents are shown in fig. 8b. Over the range $V_G - V_T = 0$ to $4V$, the SOI saturation current is an average of 43% larger than the bulk current.

To date, we have modelled the effects of fixed substrate charge and reduced vertical field on long channel devices. However, the concepts of more carriers in the channel and higher mobility should translate into higher performing short channel devices as well. (However, due to 2-D effects, the gradual channel approximation is not adequate for modelling such devices).

During switching transients, MOS transistors spend a large part of their time in saturation. Hence our predicted increase in saturation currents for fully depleted SOI films should translate directly into faster circuits. Although it has long been argued that SOI/SOS loses much of low substrate capacitance advantage at submicron geometries (when interline capacitances begin to dominate), this increased saturation current should keep SOI/SOS prominent as a high speed technology. It is here timely to note that the world's fastest room temperature CMOS circuits to date known to this author have been reported in SIMOX SOI at this conference [12].

MATERIAL CHALLENGES FOR ULSI FET TECHNOLOGY

Other advantages of silicon-on-insulator for submicron VLSI technology include reduced hot-electron effects [13] and the potential for low leakage currents [14]. These will not be reviewed here on account of space. Virtually all of these advantages (short-channel threshold voltages, subthreshold slopes, saturation current, hot-electrons...) depend on total depletion of the SOI film body. Since submicron FET's will need channel dopings of $\sim 10^{17} cm^{-3}$, total film depletion will require films of under 1000\AA in thickness. For ULSI application, gate breakdown voltages need to be uniformly high. Many people have observed poor gate oxide yields in SOI films, and have attributed this to dislocations or other defects in the film. The responsible defects need to be identified and eliminated. Finally, good subthreshold slopes will require a bottom silicon-insulator interface with few interface states. Altogether, the requirement of a high quality film of $< 1000\text{\AA}$ thickness, with electronic quality interfaces, should provide challenging direction for the materials community. Fortunately, requirement of very thin films is consistent with emerging SIMOX technology.

MINORITY CARRIER DEVICES AND STRUCTURES

Up to here this paper has focused on MOS devices. However, in many large scale circuits the dominant delays come not from individual gate delays but from delays in

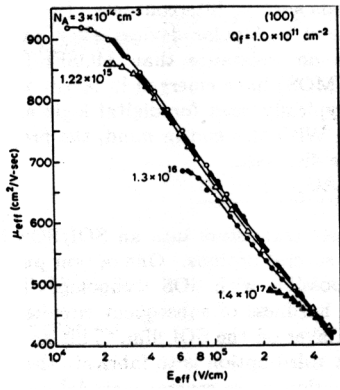


Fig. 7. Effective surface mobility as a function of average electric field in the inversion layer for various values of substrate doping ($Q_f = 1 \times 10^{11} \text{ cm}^{-2}$) [11]. The dotted line represents the relationship used for subsequent transistor modelling.

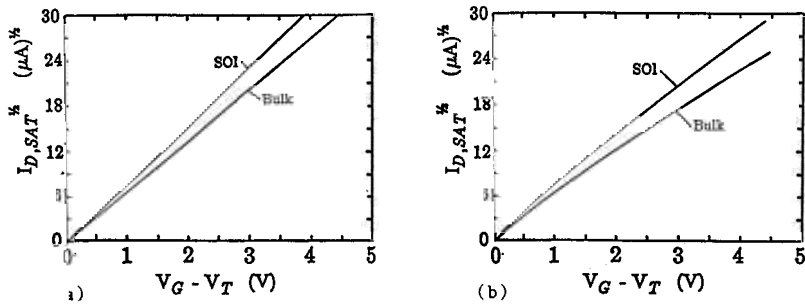


Fig. 8. Modelled saturation current vs gate voltage for bulk and ultra-thin film SOI FET's using constant mobility (a) and mobility dependent on vertical field (b).

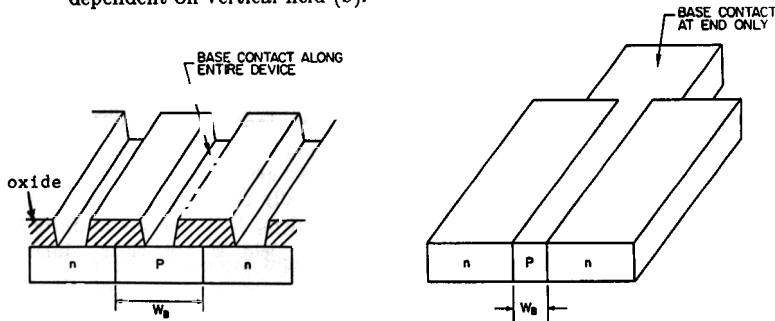


Fig. 9. Diagrams of various strategies for lateral bipolar transistors.

driving large capacitive loads such as interconnects. Because of the minority carrier nature of the current transport in bipolar devices, they have an intrinsically higher transconductance and lower on resistance than MOSFET's. Thus, many merged bipolar-MOS processes (BI-CMOS) have emerged in recent years [15,16]. In these circuits, MOS transistors are typically used for digital logic and bipolar transistors are used for driving large loads. With this end in mind, the prospects of bipolar devices in SOI/SOS films will now be discussed.

The integration of bipolar transistors into an SOI/SOS MOS process leaves the process/device designer with several options. One option put the bipolar structure in the substrate. This is not possible with SOS technology, however, and in any case would degrade the radiation hardness of subsequent circuits. A second option would involve growing an epitaxial layer on the SOI film. This greatly adds to process complexity and cost, however. A third option is to fabricate the bipolar structure directly in the thin SOI film. This option poses several material and device design challenges which we will now discuss.

Bipolar devices depend on the transport and lifetime of minority carriers. Minority carriers are inherently a more sensitive probe of semiconductor materials than majority carriers, however. For example, majority carrier properties such as mobilities are dependent on the scattering site density, and at room temperature are sensitive down to approximately the 10^{15} – $10^{16}cm^{-3}$ range. On the other hand, minority carrier lifetimes in silicon can be affected by defect concentrations as small as $10^{11}cm^{-3}$ or less. Thus while many SOI techniques yield material with excellent majority carrier properties such as mobility, it should not be surprising that most SOI techniques to date yield material with fairly poor minority carrier properties. Recently Kroll [17] prepared diodes, (a simple minority carrier device), in SOI material from several different sources. He then measured the diode quality factor in low level forward bias, a rough measure of minority carrier properties (1.00 → good, 2.00 → bad). As can be seen from the data in Table 1, all techniques but wafer bonding yielded relatively poor results. In the case of SIMOX, this low lifetime is readily understandable. The implantation process can introduce damage as well as unwanted impurities into the SOI film. Zone melting and recrystallization are known to dissolve part of the cap layer which can degrade the minority carrier properties of the SOI films and interfaces [7,18]. Since the SOI film in the wafer bonding method comes directly from a virgin substrate, it is not surprising that these films yielded the best results. Clearly, further work is needed to identify and remove the lifetime-killing centers in SIMOX or zone-melting SOI if these materials are to be used for BI-CMOS circuits. These sites could be dislocations, isolated contaminants, metals gettering by the dislocations, etc.

With respect to process yield considerations for minority carrier SOI devices, one other concern can be raised. Conventional bulk processes rely heavily on gettering processes, either deep within the wafer bulk (intrinsic gettering) or at the back wafer surface (extrinsic gettering). Since SOI films are usually separated from the substrate by oxides which are in general good diffusion barriers, how one might incorporate gettering into an SOI process to improve ULSI yield is not clear. Innovative approaches will be required.

Assuming materials challenges can be met, the fabrication of high performance bipolar devices directly in thin film is still not straightforward. Vertical bipolar structures typically require at least a micron of silicon. Vertical structures can be made in

Table 1. Diode quality factor for junctions in various SOI films. Data is from Kroll [17] except for the ZMR films (J.C. Sturm).

<u>MATERIAL</u>	<u>DIODE QUALITY FACTOR (n)</u>
Simox	1.45
Wafer Bonding	1.15
Bulk Epi (control)	1.07
ZMR	1.6 — 2.0

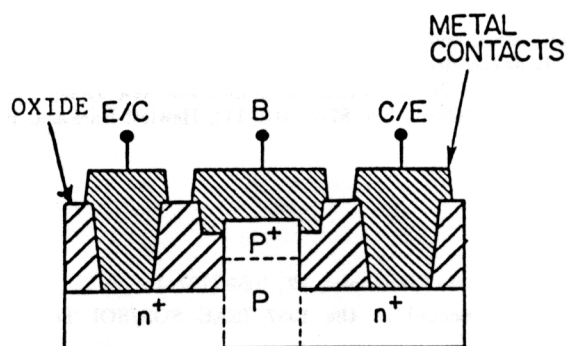


Fig. 10. Lateral bipolar transistor with a self-aligned base contact [20]

thinner films, these devices will have a fairly high collector resistance, decreasing their utility as load drivers. A device which is better suited to thin films is a lateral bipolar transistor. (Fig. 9) Such a structure should merge well with existing MOS fabrication processes. A serious problem in these devices, however, is lateral base resistance [19]. Assuming a $0.3 \mu\text{m}$ film and a $0.3 \mu\text{m}$ wide base with a doping of 10^{18}cm^{-3} , the base would have an approximate resistance of $3\text{K}\Omega/\mu\text{m}$ of base length. Assuming the base is contacted from one end, current crowding would render all but the first micron or so of the device useless. One alternative may be to use a self aligned process to contact the base along its entire length while retaining a minimum geometry base width (fig. 10) [20]. Clever merged bipolar MOS-devices using both the base and a gate to control the current may also be helpful [19].

SUMMARY

Ultra-thin SOI films of less than 1000 \AA thickness have many potential advantages for submicron MOS transistors. Several problems normally encountered in scaled MOS devices are alleviated in these structures. Those include short channel threshold voltages, subthreshold slopes, saturation currents, surface mobilities, and hot electrons. However, the use of SOI films for ULSI devices will only become widespread if the exacting yield and uniformity requirements of ULSI can be met. Dislocations, gettering approaches, and film lifetimes are all significant issues. Nevertheless, the significance of the advantages described above should provide a powerful incentive for SOI to expand beyond its traditional application areas.

ACKNOWLEDGMENTS

The collaboration and assistance of L. Pfeiffer, K. West (recrystallization) and J. Jo are gratefully acknowledged. This work also would not have been possible without the generous support of N.S.F. (P.Y.I. program), the New Jersey Commission on Science and Technology (Grant No. 87-240040-11), Hewlett-Packard, Emerson Electric, and Rucker & Kolls.

REFERENCES

- [1] L.D. Yau, *Solid State Electronics*, **17**, 1059 (1974).
- [2] J.P. Colinge, presented at the 1987 IEEE SOS/SOI Technology Workshop, Durango, CO, 1987.
- [3] J.P. Colinge, private communication.
- [4] H.-K. Lim and J.G. Fossum, *IEEE Trans. Electron Devices* **ED-30**, 1244 (1983).
- [5] R.J. van Overstraeten, G. Declerck, and G.L. Brouk, *IEEE Trans. Electron Devices* **ED-20**, 1150 (1973).
- [6] J.P. Colinge, *IEEE Electron Device Lett.* **EDL-7**, 244 (1986).
- [7] J.C. Sturm, J.D. Plummer, and J.F. Gibbons, *Appl. Phys. Lett.* **46**, 1171 (1985).
- [8] See for example, R.S. Muller and T.I. Kamins, *Device Electronics for Integrated Circuits*, (John Wiley & Sons, New York, 1977), p. 356.
- [9] H.-K. Lim and J.G. Fossum, *IEEE Trans. Electron Devices* **ED-32**, 446 (1985).

- [10] M.R. Pinto, C.S. Rafferty, and R.W. Dutton, *Pisces II: Poisson and Continuity Equation Solver*, Stanford Electronics Laboratories, Stanford University, 1984.
- [11] S.C. Sun and J.D. Plummer, *IEEE Trans. Electron Devices* **ED-27**, 1497 (1980).
- [12] P.K. Vasudev, this conference.
- [13] J.P. Colinge, *IEEE Trans. Electron Devices* **ED-34**, 2173 (1987).
- [14] J.C. Sturm, M.D. Giles, J.F. Gibbons, *IEEE Electron Device Lett.* **EDL-5**, 151 (1984).
- [15] O. Minato, T. Masuhara, T. Sasaki, Y. Sakai, and T. Hayashida, *ISSCC Tech. Dig. Pap.*, 256 (1982).
- [16] J. Miyamoto, S. Saitoh, H. Momose, H. Shibata, K. Kanzaki, and S. Kuhyama, *IEDM Tech. Dig.*, 63 (1983).
- [17] W.A. Kroll, J.F. Buller, G.V. Rouse, and R.D. Cherney, *IEEE Circuits & Devices Magazine*, **2**, 20 (July 1987).
- [18] C.I. Drowley and T.I. Kamins, in *Laser-Solid Interactions and Transient Thermal Processing of Materials*, edited by J. Narayan, W.L. Brown, R.A. Lemons, (Elsevier Science, New York, 1983), pp. 511-516.
- [19] J.P. Colinge, *IEEE Trans. Electron Devices* **ED-34**, 845 (1987).
- [20] J.C. Sturm, J.P. McVittie, J.F. Gibbons, and L. Pfeiffer, *IEEE Electron Device Lett.* **EDL-8**, 104 (1987).