Fabrication of Twin-well CMOS

Linh V. Phan
Microelectronic Engineering
Rochester Institute of Technology
Rochester, NY 16423

Abstract - A single-mask self aligned Twin-well process has been integrated into RIT's CMOS technology. These wells are self aligned to increase package density. The process has been simulated using TMA Suprem IV simulation tool. The simulated parameters were used in the actual fabrication. The wells are used to optimize both n- and p-channel active devices. The subthreshold leakage currents in isolated pmos and nmos devices are -1.28 pA/μm and 3.56 nA/μm of channel width, respectively when the devices were biased at < 5 volts. In addition, the twin-well process has produced active n- and p-channel FET's with excellent characteristics such as low threshold voltage, low subthreshold swing, and high transconductance.

I. INTRODUCTION

During the pass decade, CMOS has been a major technology because of its inherent low power characteristics. It has been noticed that the speed of CMOS, which has been traditionally been a slower technology, is now exceeded NMOS speed as a result of device size shrinkage. MOS scaling has shown that the saturation-current ratio between p- and n- channel FET's decreases as the channel length is reduced.[i] Furthermore, the ability to implement analog circuits and to diminished process complexity relative to NMOS has allowed CMOS to become the dominant technology in many VLSI applications such as large memories and microprocessors.

One major problem in bulk CMOS has been its layout density because both p- and n-channel devices must be fabricated and isolated properly. Conventionally, bulk CMOS has lower package density than does NMOS due to the large layout space required for the p- or n-well formed by diffusion. A twin-well technique has been developed to form two wells to optimize both n- and p-channel transistors.[ii] As previously mentioned, the wells are made by diffusion, and the interdiffusion of the two wells has significantly reduced the package density.

Another serious limitation in achieving high-density CMOS is the latchup phenomenon. Latchup is due to the regenerated switching in a p-n-p-n path caused by the parasitic p-n-p and n-p-n bipolar devices that are unavoidable in a bulk CMOS circuit. [iii] Various methods have been used to reduce latchup susceptibility. They include gold doping and neutron radiation to decrease minority-carrier lifetime, guardband around the devices to reduce lateral resistances and beta, and epitaxial layers on low-resistivity substrates to minimize the parasitic resistance. All of these techniques are only partially effective. Moreover, they can introduce leakage and add process complexity. For example, an epi. approach can only reduce the substrate resistance, not the parasitic resistance associated with the well. The additional cost and defect density are other concerns.

In addition to the parasitic field-oxide FET's and bipolar devices, which are the major problems in high density CMOS, designing less than two microns p-channel MOSFET's can be critical. In a twin-well approach, an n-well has been used to suppress punchthrough current in p-channel FET's at minimum expense of the n-channel performance. However, several other effects must be considered in a p-channel design. First, buried channel behavior is unavoidable because n^+-polysilicon is commonly used as the gate material. The corresponding negative work function difference (ϕ_{ms}) in conjunction with the relatively high channel doping concentration has caused a low p-channel threshold voltage. To achieve a higher threshold voltage, a boron implant is normally used to counterdoped the channel region. Consequently, a buried channel is formed at the channel surface. This buried layer, unless designed carefully, can result in poor turn-off characteristics and a strong short channel effect in the saturation region. Second, the p^+ source/drain junctions must be made shallow to minimize the short channel effect. Shallow p^+ junctions are difficult to form because boron atoms diffuse during heat treatment. Third, interconnect technology with good ohmic contacts to both p^+- and n^+-shallow junctions is also necessary for high density CMOS.

In this paper, the twin-well CMOS process and the resulting device structure are described. In addition, the electrical characteristics for the active devices fabricated using the twin-well CMOS process are also presented.

II. PROCESS SEQUENCE

The final twin-well CMOS is shown in figure 1. We first grow pad oxide on a lightly doped p-type substrate. The n-well areas for the pfet's are then defined (mask 1) by depositing and etching windows through the nitride, Si_3N_4. A thick photoresist is used to mask the n-channel devices when we implant phosphorous ions at an energy of 130 KeV and ϕ = 4e12 cm^{-2} to form n-well for the p-channel devices. We chose the phosphorous energy in order to place the implant peak deep into the substrate without implanting through the mask. After the n-well implant, the photoresist is then removed and field oxide is grown over the exposed substrate. The oxynitride is then removed, and the nitride is etched off. The field oxide is used as a mask over the p-channel devices during the p-well implant. For the p-well implant, we implant boron ions at an energy of 50 KeV and ϕ = 4e12 cm^{-2}. The p-well is implanted second to prevent boron from segregating into the oxide during the oxidation step of the well's mask. The wells are then driven-in for sixteen hours in inert ambient. These wells are used to optimize both n- and p-channel active FET's. The wells boundaries are self-aligned with the edge of the well oxide. After the drive in, the oxide is then completely removed.

The second pad oxide is then grown, and another nitride deposition is done. The nitride is used to define the active areas of the p- and n-channel devices through the second photo mask and active etch. A blanket, threshold adjust implant is implanted with boron ions at an energy of 35 KeV and ϕ = 2e12 cm^{-2}. In preparation for the formation of the gate, both the nitride and pad oxide are stripped. A 1000 angstroms of KOOI oxide is grown and stripped off to remove any residual oxynitride at the substrate surface. Five hundreds angstroms of gate oxide is then grown after fifteen minutes of TCA. TCA is done at 150 °C higher than oxidation temperature to remove contaminations in the oxidation furnace, hence reduces unwanted charges in the gate oxide. Polysilicon is then deposited and arsenic spin-on dopant is used to ensure that the gate is n^+-polysilicon. The n^+-polysilicon gates are defined using mask 3 photo and reactive ion etching. We then used two separate masks 4 and 5 to form optimized n^+ and p^+ source-drains (S/D). The n^+ S/D junctions are formed by an P^{31} implant, whereas the p^+ S/D junctions are implanted using B^{11} ions. After a CVD oxide deposition, The S/D implants are activated during the LTO densification step. The contact holes are then defined using mask level 6 and HF etch. Finally, a thin aluminum layer is sputter deposited for contacting shallow n^+- and p^+-junctions. A final 450 °C forming gas anneal is performed to achieve low contact resistance and a uniform metal-silicon interface.

A key issue in this process is the n-well oxidation required to reverse the well mask. This oxidation step not only saves one masking step, it also makes the two wells mutually self aligned. The completed Twin-well CMOS structure is shown in figure 1 below.

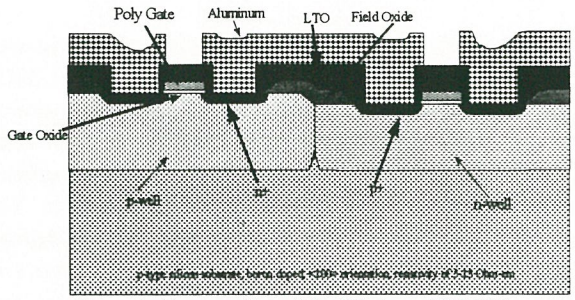

Figure 1. Cross section of final Twin-well CMOS

III. DEVICE STRUCTURE

The formation of a n-well is used to reduce $\beta_{p\text{-}n\text{-}p}$. The n-well is used in this process because it offers an additional advantage in reducing the buried channel behavior such as poor turn-off characteristics of the p-channel devices. The buried channel characteristics arise due to the boron doping needed to adjust the p-channel threshold voltage. Figure 2 shows the buried channel junction.

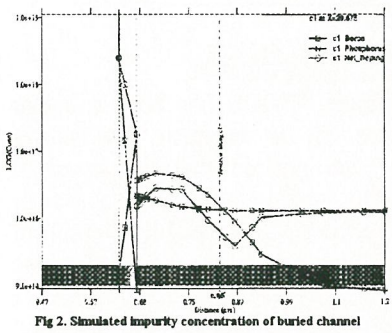

Fig 2. Simulated impurity concentration of buried channel

B_{11}^+ implant is also used to create the shallow S/D junctions in the p-channel MOSFET's. Figure 3 shows the doping profiles and the p^+ S/D junctions of the p-channel FET's. The simulated junction depth is ≤ 1.3 μm. The S/D junction for the n-channel FET's are implanted using P^{31} and the simulated dopant profiles are shown in figure 4.

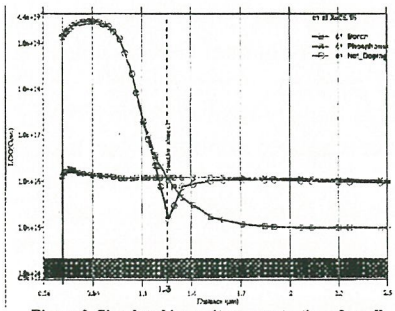

Figure 3. Simulated impurity concentration of nwell

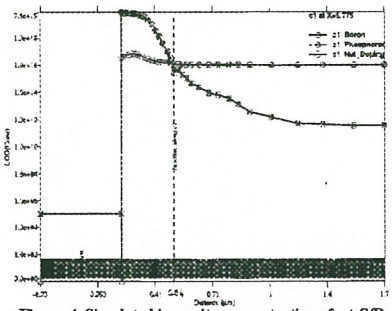

Figure 4. Simulated impurity concentration of n+ S/D

Device	Twin-well	Substrate
Vt	-1.0 V	-1.27 V
Id(leakage)	-1.28 pA/um	-1.88 pA/um
I_{DSAT}	-1.25 uA/um	-1.88 uA/um
SubVt - Swing	-147 mV/dec	- 240 mV/dec
Gm	0.43e-6 Mho/um	86.3e-9 Mho/um

Table T1. PMOS Test Device Characteristics.

Device	Twin-well	P-well
Vt	0.8 V	1.65 V
Id(leakage)	3.56 nA/um	15.6 nA/um
I_{DSAT}	0.34 uA/um	3.1 uA/um
SubVt - Swing	269 mV/dec	334 mV/dec
Gm	0.58e-6 Mho/um	0.89e-6 Mho/um

Table T2. NMOS Test Device Characteristics.

The S/D junctions are ≤ 0.54 μm.

IV. DEVICE CHARACTERISTICS

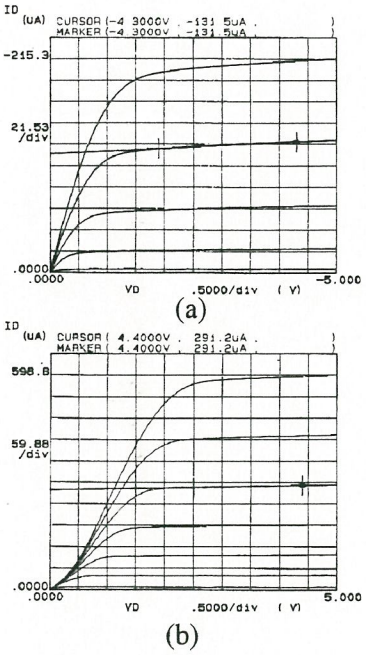

Fig. 5. I-V characteristics for 6 um CMOS FET's, (a) p-channel and (b) n-channel.

In this section, the electrical characteristics of the active devices in the twin-well CMOS are described. Figure 5 shows the typical I-V curves for both types of devices with 6 μm and 32 μm channel length and width, respectively. Notice that there is minimum punchthrough current in these FET's. The n- and p-channel FET's have a subthreshold swing of ~269 mV/dec and ~147 mV/dec, respectively. As seen in tables T1 and T2, the subthreshold swing of the devices in the twin-well are significantly improved relative to the devices in the substrate. As a result of the higher background concentration for the pmos, the buried layer thickness is significantly reduced, thereby improving the subthreshold characteristics. Additional device characteristics are summarized in tables T1 and T2. The current driving abilities were measured at $|V_{DS}| = 5$ volts, and the gate potential was at 2 volts. Although the junction capacitance and the body effect were not extrapolated, the expected body effect and the junction capacitance are expected to be high due to the relatively high doping concentration of the wells. In a VLSI circuit, the effect of junction capacitance is becoming less significant due to the dominant wiring and interwiring capacitances.[iv] The body effect in CMOS, although to a lesser degree when compared to that in NMOS, can degrade circuit performance in some applications. During the formation of the wells, higher energy implants may be used to push the highly doped region away from the active channel, thereby reducing the junction capacitance and the body effect. However, masking the implant may required a very thick resist.

As far as latchup is concerned, the n-well provides a heavily doped base region, which can reduce the current gain in the lateral p-n-p bipolar device. Notice from figure 1, the base region of the p-n-p device is the n-well and the emitter area is the small sidewall of the shallow p^+ junction. The heavily doped base region and the small emitter area have reduced the emitter efficiency, hence the current gain. Furthermore, the current gain of the vertical n-p-n devices is reduced as a result of deeper wells to widen the n-p-n base width.

V. SUMMARY and CONCLUSIONS

Twin-well process has been integrated into CMOS technology for high-density integrated circuits. These wells are self-aligned to increase package density. The p-channel device has subthreshold characteristics of -1.28 pA/μm of channel width and -147 mV/dec. It also has a threshold voltage of -1.0 volt and a transconductance of 0.43e-6 Mho/μm of channel width. The n-channel devices' threshold voltage and transconductance are 0.8 volts and 0.58e-6 Mho/μm of channel width, respectively. Furthermore, its subthreshold leakage current and swing are 3.56 nA/μm of channel width and 269 mV/dec, respectively. The only draw back of the twin-well approach is the high body effect.

REFERENCES

[i] R. Chwang and K. Yu, "CMOS - An n-well bulk CMOS technology for VLSI," VLSI Design, 4th quarter, p. 42, 1981.

[ii] L. C. Parrillo, R. S. Payne, R. E. Davis, G. W. Reutlinger, and R. L. Field, "Twin-tub CMOS - A technology for VLSI circuits," in IEDM Tech. Dig., p. 752, 1980.

[iii] D. B. Estreich and R. W. Dutton, "Modeling of latchup in CMOS integrated circuits," IEEE Trans. Computer-Aided Design, vol. CAD-1, p. 157, 1982.

[iv] J. Chen, "Quadrupole-Well CMOS for VLSI Technology," IEEE Trans. On Electron Devices, vol. ED-31, p. 918, 1984.

Linh V. Phan came from Vietnam in 1986. He received his BS in Microelectronic Engineering at Rochester Institute of Technology in May, 1997. He is currently employed with Motorola, Austin, Texas, as Device Integration Engineer.