

a single mask is used to pattern a nitride/oxide film formed on the bare silicon surface (*Mask #1*). The openings in the film become the *n*-well regions, and phosphorus is implanted into them (e.g., at 80 keV). Next, a thermal oxide is selectively grown on the *n*-well regions, to a thickness that is sufficient to block the boron implant used to form the *p*-wells. (The nitride mask layer is used to prevent the oxide from growing in the *p*-well regions.) The pad-oxide/nitride mask layer is then stripped, exposing the silicon in the *p*-well regions. The *n*-well regions are however, still covered with the oxide grown earlier. Thus, during the *p*-well implant (e.g., at 50 keV), this oxide masks the *n*-well regions, permitting the boron to enter the silicon only in the *p*-well areas.

Next the wells are driven in (e.g., at 1100°C for 500 min). At the conclusion of this drive-in step the concentrations in the wells for an 0.8- μm -CMOS process are $\sim 1 \times 10^{16}/\text{cm}^3$ for the *p*-well, and $\sim 3 \times 10^{16}/\text{cm}^3$ for the *n*-well. The *n*-well is selected to have a higher doping concentration (i.e., to improve the punchthrough performance of the PMOS devices and to eliminate the need for a separate channel-stop implant for the *n*-well). A higher concentration in both wells would still produce devices with relatively low capacitances at the bottoms of the source/drain-to-well junctions.

The channel-stop implant procedure is also usually included in the part of the processing sequence in which the wells are formed. In one reported twin-well process,¹³ only a *p*-well channel-stop implant (boron) is used, because the doping in the *n*-well is high enough ($3 \times 10^{16} \text{ cm}^{-3}$) that a separate channel stop implant is not required. In this case, an unmasked boron implant is performed following both well implants, but prior to field-oxide growth.

In a second approach,¹⁴ an additional mask can be used to provide both *n*-well and *p*-well channel stops (Fig. 8-8a). Note that in this procedure, a mask is saved by

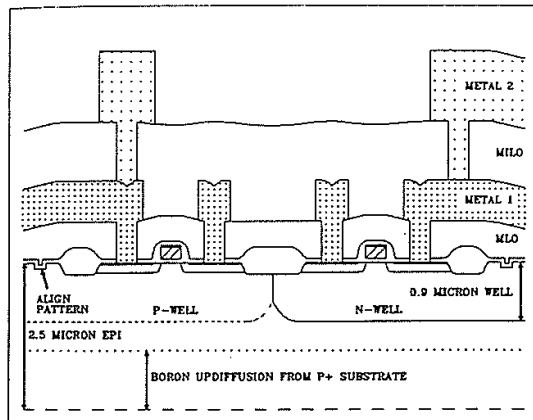


Fig. 8-6 Twin-well CMOS showing initial alignment mark.¹⁶ (© 1991 IEEE).

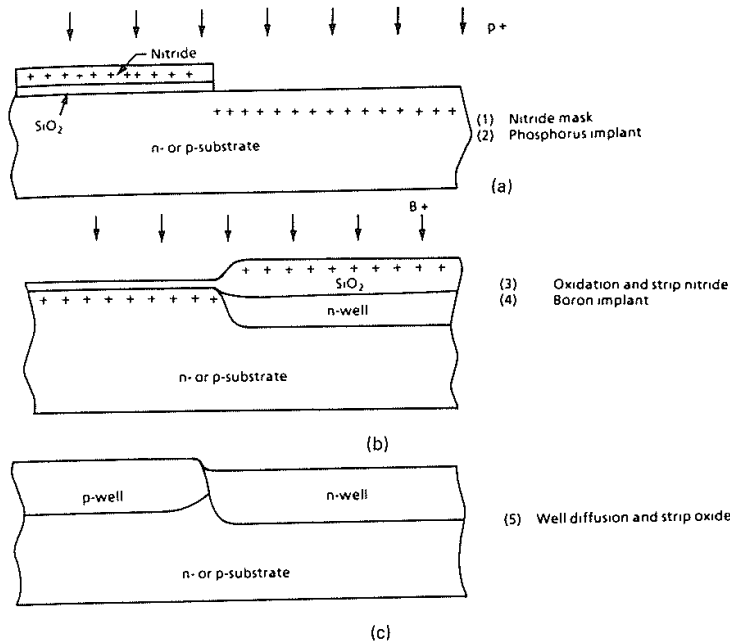


Fig. 8-7 Twin-well CMOS process using a single mask to form both wells. (a) P⁺ implant for the n-well. (b) B⁺ implant to form the p-well (blocked by the n-well oxide). (c) well drive-in for both wells.¹² (© 1982 IEEE).

performing a blanket boron implant. As a result, the boron channel stop dopant ends up in the n-well as well as in the p-well. Hence, the phosphorus channel stop dose must be increased to compensate for the boron's presence in the n-well field regions. The disadvantages of this method include the additional alignment step between the channel stops and the well masks, the interdiffusion that occurs during the oxidation step, and the asymmetry of the doping profiles. (The latter is due to the increased phosphorus concentration that must overcompensate the nonselective boron channel stop).

A variation of this method that requires no masks at all to create the channel stops has been developed to overcome the above drawbacks (Fig. 8-8b).^{14,15} In this sequence, both arsenic and phosphorus are co-implanted into the n-well regions. This places both the dopants that form the well and the channel stop into the n-well regions prior to implanting of the p-wells. An oxide is then selectively grown on the n-well regions and the boron dopant for the p-wells is implanted. After the wells have been driven in, the oxide over the n-wells is retained, and a second boron implant is carried out. This implant serves as both a channel stop in the p-well field regions and a punchthrough-prevention implant in the active regions of the p-well.

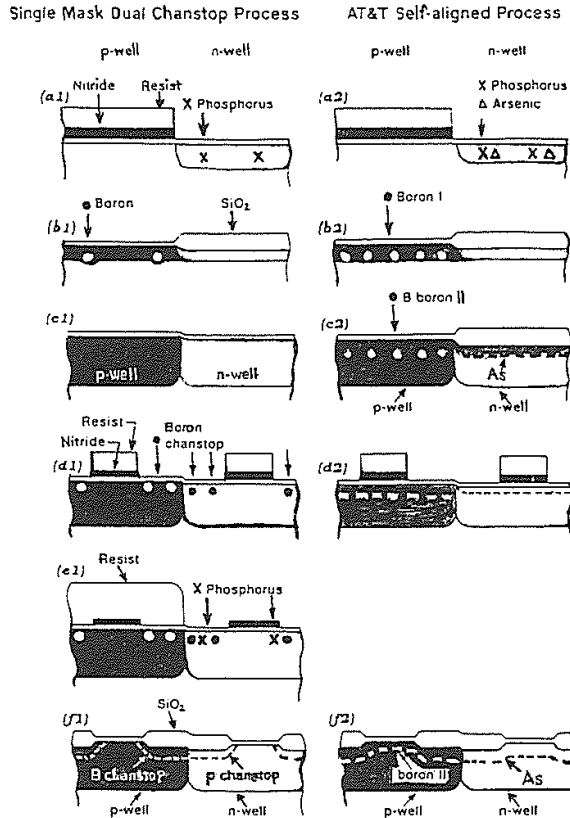


Fig. 8-8 (a) Twin-well CMOS with a single-mask dual channel stop process. (b) ATT twin-well CMOS process which requires no masks to form the dual channel stops (self-aligned process.)¹⁵ (© IEEE 1981).

It should also be noted that the minimum separation between two wells of the doping type at two different potentials is limited by electrical isolation. In single-well processes, to prevent punchthrough in the lightly-doped substrate between two such wells, adequate spacing between them must be allowed.

8.2.3 Split Well-Drive-In Process

In another report, the long, high-temperature drive-in step is performed in a novel manner, allowing the *p*-well transistors to be fabricated without use of a channel-stop

implant.¹⁷ After the p -well implant is carried out, the field oxide is first grown without driving the implanted well dopants. During this step, some of the implanted boron impurities segregate into the field oxide, leaving a surface concentration as low as would exist after any field oxide step. Next, the remainder of the well drive-in step is performed in a non-oxidizing ambient. The well dopant deep in the bulk now back-diffuses toward the surface, raising the Si-SiO₂ interface concentration by a factor of 2-3 (to $2\text{--}3 \times 10^{16} \text{ cm}^{-3}$ in this case, see Fig. 8-9). The well concentration at surface of the field regions is only about 1.5x lower than at the surface in the active regions. V_{TFISO} greater than 10 V (at I_D at $0.5 \text{ pA}/\mu\text{m}$ at $V_{\text{DS}} = 8 \text{ V}$) can be produced for isolation lengths of $1.2 \mu\text{m}$. It has been observed that for very narrow NMOSFETs (e.g., $0.35 \mu\text{m}$) fabricated with the split-well drive-in scheme, V_T is decreased by $\sim 100 \text{ mV}$ compared to wider channel devices (e.g., $5 \mu\text{m}$). This occurs because boron also diffuses from the channel regions to the field regions during the post-oxidation drive-in step. Such diffusion occurs because, as described earlier, the boron concentration at the silicon surface in the field regions has been reduced during the field oxide step by boron segregation effects.

8.3 RETROGRADE-WELL CMOS

As previously discussed, conventional wells are formed in single- and twin-well CMOS technology by implanting the well dopants and then diffusing them to the desired depth. However, lateral as well as vertical diffusion occurs, which has the effect of reducing packing density. For example, since the lateral diffusion is ~ 0.7 times the junction

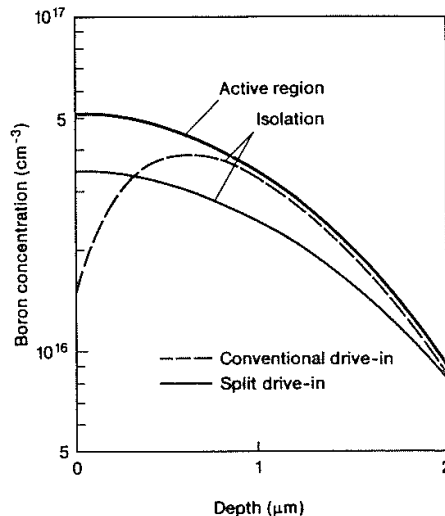


Fig. 8-9 SIMS profiles of boron well underneath field oxide and in the active area for split and non-split drive-in. Implanted boron dose: $8 \times 10^{12} \text{ cm}^{-2}$. The active area profile is the same for both drive-in cases.¹⁷ (© 1990 IEEE).

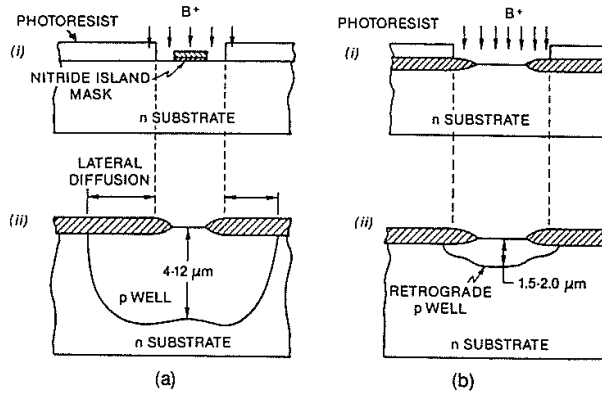


Fig. 8-10 Formation of (a) conventional and (b) retrograde p -well CMOS structures.¹⁹ (© 1981 IEEE).

depth, a typical diffused well of 3 to 4 μm results in a lateral spread of well dopant of ~ 2 to 3 μm . If a higher energy implant is used to place the dopants at the desired depth without significant additional diffusion (i.e., the high-energy implant is performed after the active area isolation structures [field oxide] is already in place, and only a short anneal at a relatively low temperature is used to activate the implant), much less lateral dopant spread occurs. This is because ion implantation is a more anisotropic dopant distribution process than diffusion. That is, the lateral implant straggle is $\sim 0.7 \mu\text{m}$ and barely exceeds one micron after all thermal processing.¹⁸ In addition, since the high-energy implanted wells are formed after the LOCOS isolation step, lateral spread of the well dopants is further reduced. Hence, the 7-9 μm n^+ -to- p^+ spacing required for a twin-well 1.2 μm CMOS process which uses conventional wells, can be reduced to $\sim 5 \mu\text{m}$ if retrograde wells are used instead.

Figure 8-10 depicts the cross-section of a high-energy implanted p -well in comparison with a conventional diffused well. In this example, the implanted p -well is formed by boron implantation at 400-600 keV, followed by a brief 30 min anneal at 1000°C (instead of the 6-20 hour drive-ins at 1100-1150°C used for conventional-well formation). Notice the lateral spread associated with the p -well is greatly reduced.

Such high-energy implants also cause the peak of the implant to be buried at a certain depth within the silicon substrate (depending on the implantation energy, but at roughly 1 μm for the energies most frequently used), and the impurity concentration decreases as the doping profile approaches the wafer surface (see Fig. 8-11). Since the well profile in this case is different from that of conventionally formed wells (in which the doping concentration after drive-in is highest at the surface, and decreases monotonically with depth), such deeply implanted wells are known as *retrograde wells*. It is estimated that, in order to avoid excessive spread of the active device electrical

parameters, the surface concentration after the retrograde implant should be at least five times smaller than its final value, which is set by shallower low-dose implants specifically designed for threshold and punchthrough control. Comparison of the doping profiles in conventional and implanted p -well profiles (obtained by simulation) is shown in Fig. 8-11.¹⁹ The retrograde-doping profile is retained by minimizing the temperature cycles of later process steps (e.g., by using RTP). Retrograde wells can be implemented on both bulk and epitaxial wafers.²⁰

Besides the potential benefit of increased packing density, retrograde wells offer the following advantages:¹

1. A retarded electric field is created in the parasitic vertical bipolar transistor, thereby reducing its current gain. This affords some protection against latchup.
2. Susceptibility to vertical punchthrough is reduced without having to significantly raise the surface doping concentration.
3. The conductivity in the bottom of the well is increased, which also provides additional latchup protection.
4. A higher threshold voltage can be achieved in the field regions of the p -wells if the boron implant is done following field oxidation (i.e., the boron does not segregate out into the oxide during the field oxide growth step).¹⁹
5. Lateral diffusion of the channel-stop boron is also eliminated, thereby reducing encroachment of the boron into the active regions. This lessens narrow width effects in the active devices.

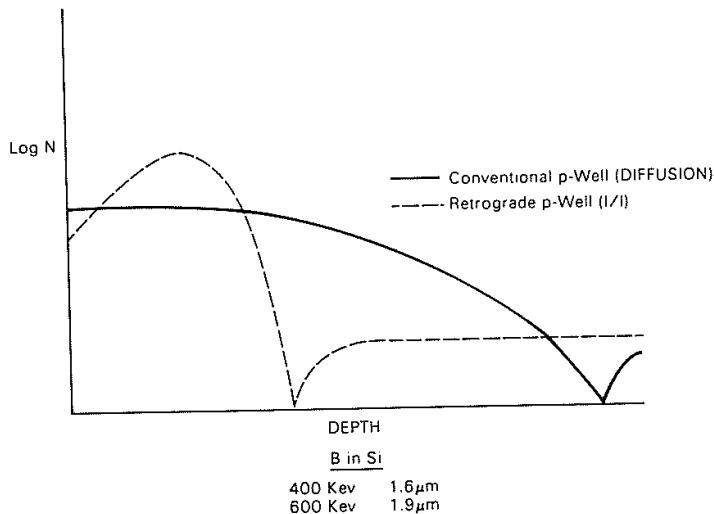


Fig. 8-11 Comparison of conventional and retrograde well p -well profiles.¹⁹ (© 1981 IEEE).

6. Greater flexibility exists in retrograde structures to optimize channel doping profiles for active device performance. That is, in conventional wells (especially shallow ones), the surface doping is primarily determined by the need to maintain sufficient vertical isolation. In retrograde structures, however, vertical isolation is secured by the high doping level towards the well bottom. Thus, it is possible to adjust the surface doping profile to a larger extent. This helps to achieve adequate short-channel device characteristics with retrograde wells.

7. The shallower well depth should also make the drain of the parasitic n^+ -to- n -well (or p^+ -to- p -well) parasitic FETs shallower. This should improve both the short-channel threshold roll-off and the punchthrough characteristics of these parasitic structures, as is discussed further in chap. 6.

However, several concerns associated with the retrograde process should also be mentioned. First, the process requires a thick photoresist as a masking material for the high-energy implant. Typical resist thicknesses are about $5\text{ }\mu\text{m}$, and special exposure and development techniques are required for this process. Second, there is concern about the annealability of the damage caused by the high-energy implant. Third, the retrograde implant may cause both the junction capacitance and the body factor to be significantly increased. The p^+ -to- n -well junction capacitances measured in retrograde-well structures formed with a 700 keV implant have been shown to be about twice the values obtained in conventional wells. While this can adversely impact the propagation delay, circuit performance is also impacted by a number of other parasitic capacitances, and the net result on typical CMOS circuit speed and power dissipation is often relatively small. For example, in a simulation of a 32-bit CMOS arithmetic logic unit (ALU), it was found that the circuit delay increases by about 7% as the p -channel junction capacitance increases by 30 percent.¹

When the retrograde well is formed by means of very high-energy implants, however, the doping concentration under the bottom of the source and drain regions is decreased, which reduces the junction capacitance. This implies that if such high-energy ion implanters ($\sim 1\text{ MeV}$) become available for production, some of the disadvantages of retrograde wells can be overcome.¹⁸ In addition, the higher doping that will be required in the wells for fabrication of submicron CMOS devices may mean a less severe junction capacitance penalty.

Although both p -type²¹ and n -type²² retrograde wells have been demonstrated (see Figs. 8-12a and 8-12b,²⁰ and Fig. 8-12c,²³ respectively), the p -type technology has been more widely implemented. This is because at least 600-keV beam energies are required for the formation of n -type (phosphorus) retrograde wells suitable for 0.5-0.8 μm CMOS. As of 1994 such machines were commercially available, but had only begun to be introduced into mainstream production environments. Retrograde p -type wells adequate for 0.5 μm devices require implants with energies of $\sim 300\text{ keV}$ at a dose of $\sim 1 \times 10^{13}\text{ cm}^{-2}$. Such p -well implants can be achieved by implanting singly ionized boron at 300 keV or doubly ionized boron at 150 keV. (The required p -well energy is smaller because boron has a larger projected range than either As or P; see Vol. 1, chap. 9). Although the doubly ionized boron approach is achievable with conventional

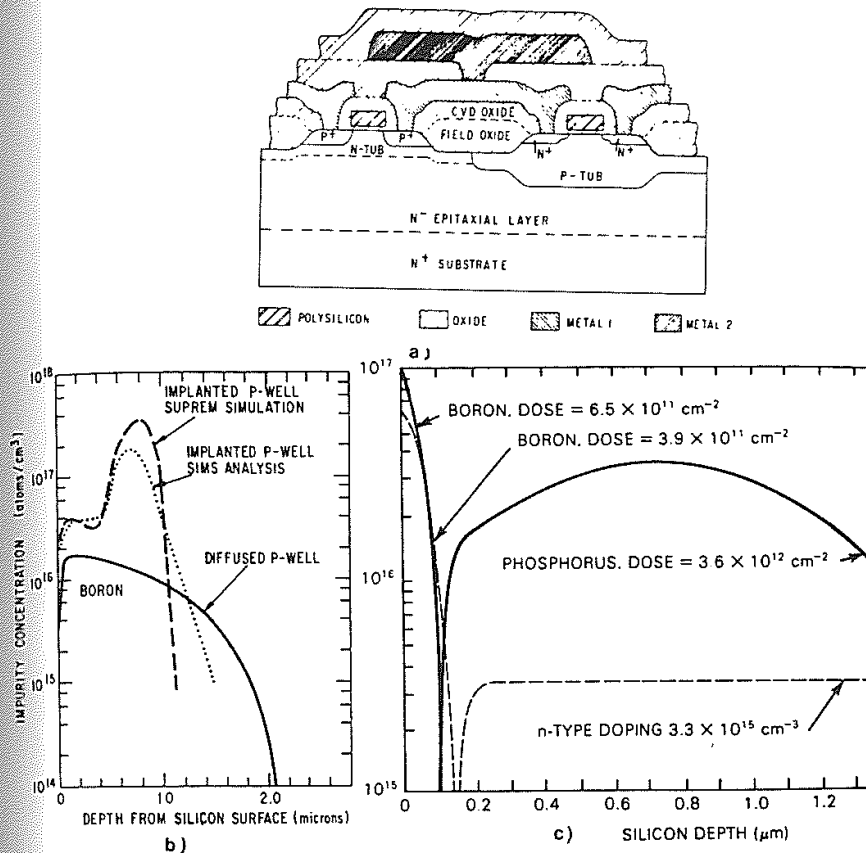


Fig. 8-12 (a) Twin-tub device cross section in which a retrograde *p*-well is used. General Electric AVLSI Process. (b) Retrograded *p*-well implanted impurity concentration profile. General Electric AVLSI process. Also shown is a conventionally thermally diffused well.²⁰ (© 1986 IEEE). (c) Simulated doping profiles in PMOS channel regions: dashed lines, lightly doped substrate; solid lines, retrograde *n*-well.²³ (© 1982 IEEE).

production implanters it is not a routine process to implement.* Subject to the precautions mentioned in the footnote, retrograde well CMOS processes have been

* One problem is that a large fraction of double-ionized ions, as high as 20%, can interact with gas molecules at the entrance of the acceleration column and become single-charged ions because of partial neutralization. These ions would only penetrate into silicon about half the desired range, distorting the profile and causing abnormalities of the threshold

the successfully implemented in a production environment with good results. However, these have primarily been limited to custom and advanced system applications, where higher costs of retrograde well processing are justified by the higher density and performance benefits offered by the retrograde-well device structures.

Two examples in which a 300 keV boron implant is used to form a retrograde *p*-well CMOS structure have recently been published. In the first, this *p*-well structure (together with a 1.2 μm epi layer) allowed an NMOS device-to-device spacing of less than 1 μm to be used when an optimized field implant was employed.²⁵ In the second, the *p*-well was fabricated with a 320 keV implant (using, in fact, doubly-charged boron accelerated at 160 keV), and a 5 min 1100°C RTP anneal. Together with the use of a 2-2.5 μm epi layer, a sub-half micron CMOS structure was fabricated (which permitted an n^+ to p^+ spacing of less than 2.5 μm).¹⁶

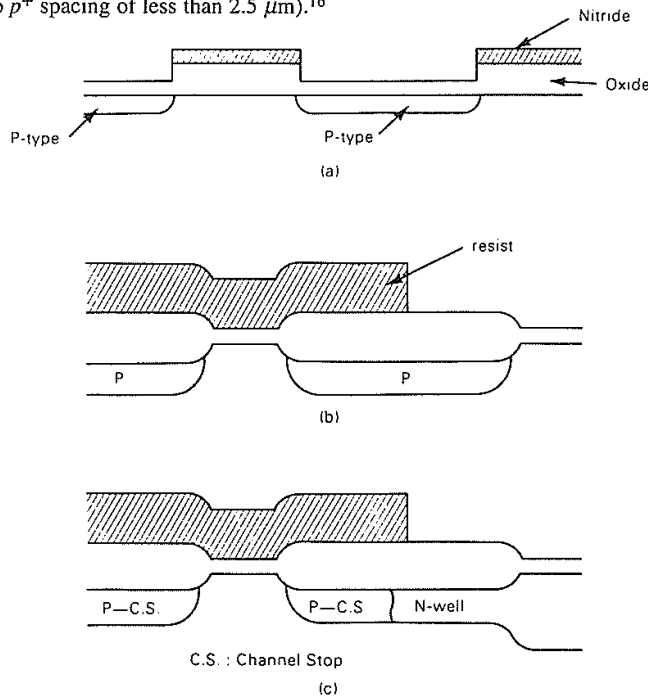


Fig. 8-13 A simplified retrograde *n*-well process. (a) Nitride mask for active area and B^+ implant for *p*-type channel stop.; (b) photoresist mask for *n*-well; and (c) high energy P^+ implant and any additional shallow implants in the *n*-well.¹⁸ (© 1986 IEEE).

* (cont.) voltage and other electrical parameters. Thus, if a high-current implanter is not available, a medium-current implanter of the post-analysis type should be used to eliminate single-charge ions before they reach the wafers.²⁴

However, an n -well retrograde CMOS process using a 700-keV phosphorus implant has also been demonstrated.¹⁸ The process sequence used to form this retrograde n -well CMOS structure (with the well implant performed after field oxidation) is shown in Fig. 8-13. The active areas are first defined using nitride masking. Next, a blanket low-energy boron channel-stop implant is performed. The field oxidation is then carried out and the retrograde n -well implant step is performed. A 30 minute, 1000°C anneal follows the implant. Note that the blanket boron channel-stop dopant is overcompensated in the n -well regions by the phosphorus implant. This retrograde-well process permits very small ($\sim 2\ \mu\text{m}$) n^+ -to- p^+ spacings.

Some advanced twin-well CMOS processes with both retrograde n - and p -wells, (using short [~ 5 min] RTP well-drive-in steps), have also been reported for submicron applications.²⁶ Scaling CMOS to $0.5\ \mu\text{m}$ and below reduces the well depths required to prevent vertical punchthrough and latchup. In addition, the well depth and well-dopant concentration can be reduced if additional, lower energy implants are precisely located with respect to the transistor structure to suppress punchthrough of the MOSFET channels and to enhance the field isolation characteristics.

One novel method of achieving such short well-drive-in times is to enhance the effective diffusion rate. It has been demonstrated that the diffusion of boron and phosphorus is substantially enhanced by oxynitridation in an NH_3 ambient with a thin oxide layer over the diffusion region.^{27*} These factors combine to produce a trend of lower implant energy in retrograde wells for scaled devices at and below $0.5\ \mu\text{m}$.

A twin-retrograde-well $0.7\text{-}\mu\text{m}$ -CMOS process for fabricating 1-Mbit SRAMs on p^+ -on- p^+ epi substrates has been reported.¹¹ The high energy implants also allow a restricted thermal budget to be used, thus reducing the up-diffusion from the p^+ substrate. This permits the use of a thinner epi layer (which also helps prevent latch-up), and also allows implantation of the channel-stop dopants after the field oxide has been grown and planarized (thus minimizing lateral diffusion of the boron channel-stop dopants).

The use of retrograde wells for both the n - and p -wells of a twin-tub CMOS process has also been reported.²⁹ The approach was called *quadruple-well CMOS* to emphasize the fact that for each type (n - or p -well), two wells were needed, one in the active area, and another, shallower one in the field region. The deep and shallow wells, however, are created simultaneously using the same higher energy implant steps. That is, since the implant is carried out after the field oxidation (and after the LOCOS nitride mask has been stripped), in regions of the wafer where no field oxide exists the implant penetrates deeply, but in regions where the implant must penetrate the field oxide, it enters the Si substrate less deeply (creating a shallower well there).

Another approach for fabricating retrograde wells without the need for very high ion implantation is to use epitaxy. Such a method for forming an n -well CMOS structure is shown in Fig. 8-14. Shallow, heavily doped wells are formed in the p -substrate by ion implantation of a slow diffusing dopant (typically arsenic) in order to prevent significant redistribution during subsequent processing. A p -type epitaxial layer is then grown to a thickness of a few microns, creating an n^+ buried layer. A low-energy phosphorus implant is then used to complete the n -wells. No drive-in is needed if it is

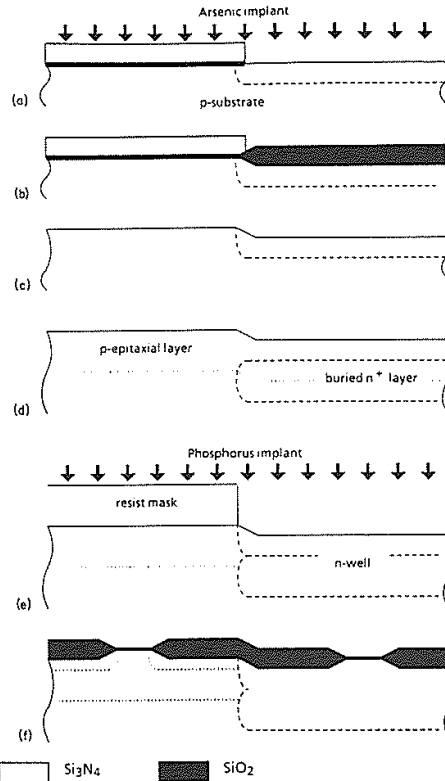


Fig. 8-14 Retrograde n -well architecture based on epitaxy. (a) Oxide/nitride well mask and buried layer implant. (b) Well oxidation. (c) Nitride/oxide strip. (d) Epitaxial growth. (e) Surface n -well implant. (f) Boron field implant and standard LOCOS field oxidation. From A.G. Lewis and J.Y. Chen, "Current Trends in MOS Process Integration," in N.G. Einspruch and G. Gildenblatt, Eds. *Advanced MOS Device Physics*, Academic Press, San Diego CA, 1989, Reprinted with Permission.

arranged that the surface phosphorus implant merges with the buried layer to complete the n -well. The formation of buried layers in this way is especially well-suited to BiCMOS technologies.

8.3.1 Retrograde Well Technologies to Improve Isolation

Since retrograde wells are formed by a high-energy implantation step following the LOCOS oxidation growth process, lateral diffusion of the well dopants is greatly reduced compared to a conventional well process. This results in significant saving of layout space. Isolation among opposite type devices is also expected to be improved.

Here we present data that supports this latter claim, as a supplement to the material on isolation presented in chap. 6.

The comparison of isolation performance between conventional-well and retrograde-well CMOS structures was performed using an *n*-well process and aluminum-gate MOSFETs.¹⁸ To establish a fair comparison, equal well depths were used in the active area of both the conventional-well and retrograde-well structures.

The process sequences used to form both the conventional-wells and retrograde-wells in an *n*-well CMOS process are shown in Figs. 8-15a and 8-15b, respectively.³⁰ The active areas are first defined in both structures using nitride masking. In the retrograde structures, a blanket low-energy boron channel-stop implant is performed next. In the case of the conventional well, the *n*-well resist mask is aligned to the patterned nitride, and the *n*-well implant (phosphorus) is carried out, followed by a blanket low-energy boron channel-stop implant. The field-oxide is now grown on both structures. In the retrograde-well structure, the *n*-well implant step is carried out after the field oxidation, followed by a short (~30 min) anneal at 1000°C. In both structures, the boron channel stop dopant is overcompensated in the *n*-well regions by the phosphorus implant.

Because the retrograde well is formed by high-energy ions penetrating the field oxide in the field regions the well depth there is only half of the well depth in the active regions. The shallower well depth in the field-regions of the retrograde-well structure (compared to the well field-region depth in the conventional-well structure) is advantageous. That is, it creates a shallower drain for the n^+ -to-*n*-well parasitic FET,

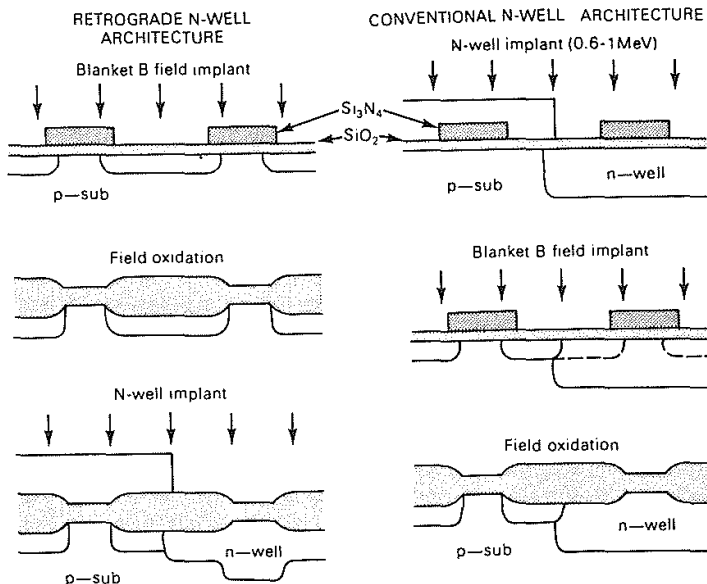


Fig. 8-15 Process architectures for retrograde and conventional *n*-well CMOS structures.³⁰

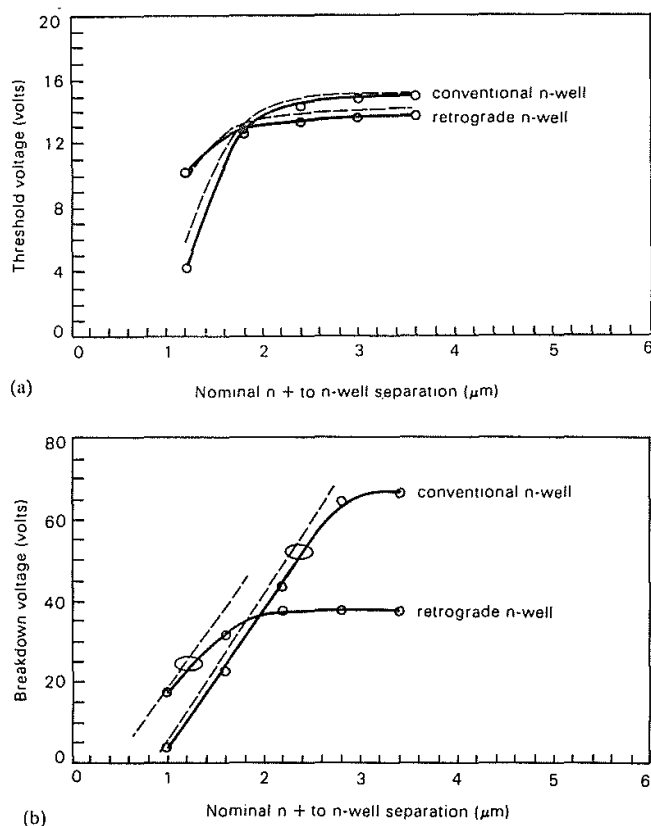


Fig. 8-16 Comparison of retrograde and conventional well structures for n^+ to n -well parasitic FETs: (a) field threshold voltage; and (b) n^+ -to- n -well breakdown voltage. Solid lines: experimental data, dashed lines: modeled results, breakdown voltage is modeled by punchthrough only.³⁷ (© 1987 IEEE).

thereby improving isolation. At the same time, the depth of the well in the active region is greater, so that the performance of the devices in the well are not significantly degraded because the n -well doping concentration near the surface remains sufficiently low. In addition, in the active regions the peak of implanted profile is near the bottom of the well, which is beneficial for vertical isolation. Meanwhile, the implant profile peaks near the Si surface in the field areas, which is again better for producing a channel stop and for preventing p^+ -to- p -substrate lateral punchthrough.

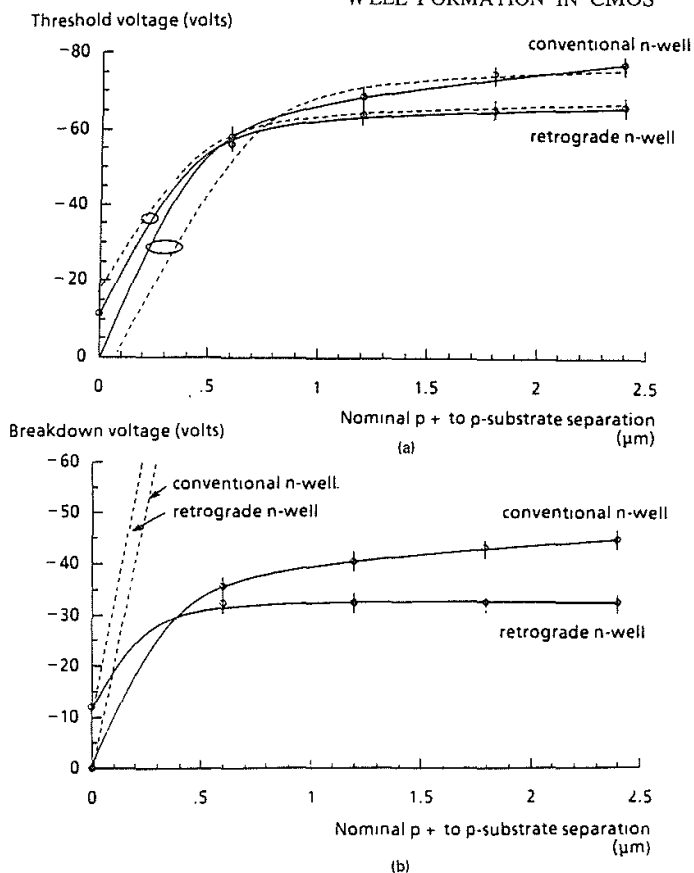


Fig. 8-17 Comparison of retrograde and conventional well structures for p^+ to p -well parasitic FETs: (a) field threshold voltage; and (b) p^+ -to- p substrate breakdown voltage. Solid lines: experimental data, dashed lines: modeled results, breakdown voltage is modeled by punchthrough only.³⁷ (© 1987 IEEE).

The measured and simulated $V_{T_{FISO}}$ s and breakdown voltages (all defined at a drain current in the parasitic FET test structure of 10 pA) for n^+ -to- n -well parasitic transistors in both the conventional and retrograde aluminum-gate CMOS test structures are shown in Figs. 8-16a and 8-16b. At large n^+ -to- n -well separations, avalanche breakdown at the well-substrate junction occurs before lateral punchthrough occurs. The lower value of avalanche breakdown in the retrograde structures is attributed to higher well doping. As described earlier, shallower retrograde wells should exhibit less propensity to degradation

caused by short-channel effects (e.g., to threshold lowering and DIBL). These predictions are confirmed from the experimental V_{TFISO} and punchthrough data.

Figure 8-17a gives the V_{TFISO} values for lateral p^+ -to- p -substrate parasitic devices. The absolute values are much larger than those for the n^+ -to- n -well devices. These higher values are due in part to the greater field doping required in the n -well to guaranty sufficient compensation of the boron. The larger values are also partly caused by the fixed charge at the Si-SiO₂ interface, which is generally positive and thus increases the magnitude of the threshold voltage of the p -channel transistors. It is also evident that short-channel effects are much less severe in these p -channel parasitic devices, especially for the retrograde structures; the thresholds do not decrease until the channel length approaches 0.5 μm . This effect is caused by the higher channel (n -well) doping in these parasitic transistors, and to the drain (substrate) doping being lower than the channel (n -well) doping.

Figure 8-17b depicts the breakdown voltages exhibited by the same p^+ -to- p -substrate devices, and comparison with Figs. 8-17a and b again shows the superior performance of these structures. In these types of parasitic FETs, the p -substrate, which is considered to be the drain, is not only deep but wraps around under the source (see Fig. 6-45). Thus, both lateral and vertical leakage paths may possibly exist between the p^+ source and the p -substrate. The limiting values of breakdown voltage for retrograde n -well structures is set by well to substrate avalanche breakdown, as is the case of the n^+ -to- n -well transistors. In conventional wells, however, the limit is set by p^+ -to- p -substrate vertical punchthrough breakdown. Thus, the maximum breakdown voltage shown on Fig. 8-17b for these devices is not the same as that shown in Fig. 8-16b, even though the drain-channel junction is in fact the same in both cases.

The application of retrograde wells for both the n - and p -wells to a twin-tub CMOS process has also been reported. The approach, as described in the previous section, was called *quadruple-well CMOS* to indicate that for each well type (i.e., n - or p -), two wells were needed, one in the active area, and another, shallower one in the field region (see Fig. 8-18). The two shallow wells actually function as self-aligned channel stops, in which the surface concentration of the well dopants is tightly restricted on each side of the well boundary. That is, as shown in Fig. 8-19 the reduction of the net impurity concentration is restricted to a small distance from the boundary of the well edge. Specifically, the surface doping concentration near the well boundary is decreased significantly only within a few tenths of a micron from the boundary, instead of a few microns, as occurs in a conventional twin-well structure (see Fig. 6-47). Experimental data shows that at a 4 μm n^+ -to- p^+ separation, subthreshold leakage in the isolation region is <1 pA for gate voltages between -15V and +20V (Fig. 8-20). Simulation data indicate that it is possible to obtain adequate isolation with a 2 μm n^+ -to- p^+ separation.

A CMOS process based on high energy implantations without epi-layers or trenches that allows n^+ -to- p^+ spacings as small as 1.5 μm has been reported. To realize such small isolation spacing, the n -well was formed with 700 keV, $1 \times 10^{13} \text{ cm}^{-2}$ phosphorus and 1.5 MeV, $3 \times 10^{13} \text{ cm}^{-2}$ boron implantations. A unique feature of the well structure in this report is that these same two implantations simultaneously form not only the two retrograde wells and the n -well and p -well channel stops, but the boron implant also

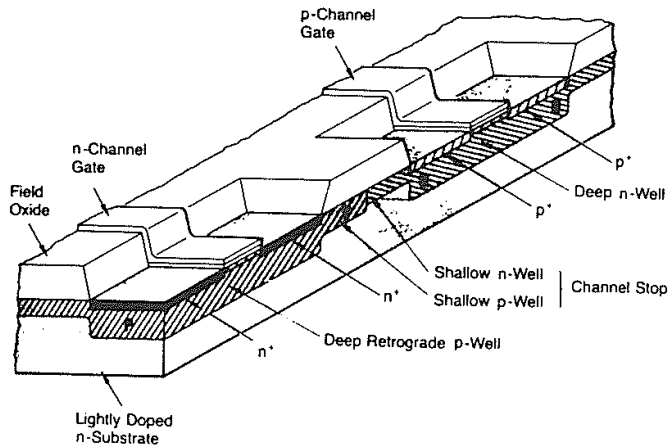


Fig. 8-18 Cross section of a quadruple-well CMOS device.³⁶ Reprinted with permission of VLSI Design.

formed a deep, buried p^+ layer under the phosphorus n -well profile. The p^+ layer eliminates the need for substrates with an epi layer, since the low resistivity substrate required to suppress latchup is created by the high energy implantation. It was also reported that the high energy implantation did not induce any excess leakage or measurably degrade MOS transistor characteristics.

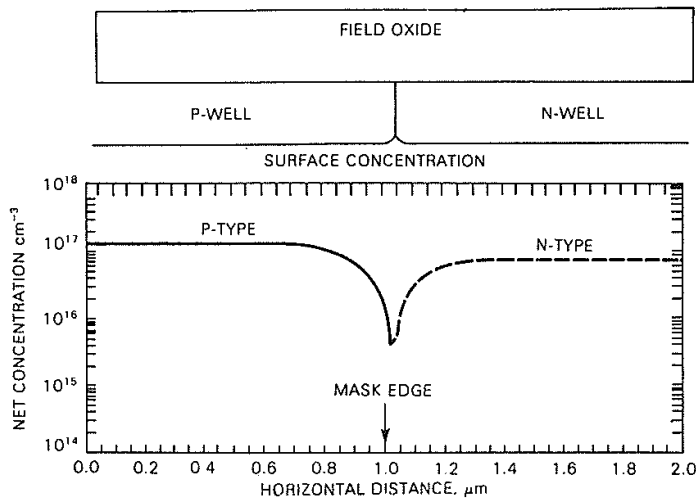


Fig. 8-19 Surface impurity concentration at the boundary of the p - and n -retrograde wells in quadruple-well CMOS.²⁹ (© 1986 IEEE).

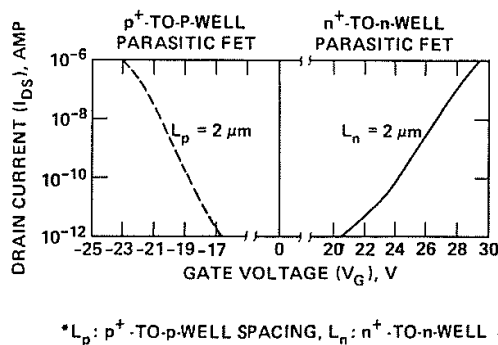


Fig. 8-20 Subthreshold characteristics of symmetric parasitic FETs in quadruple-well CMOS.²⁹ (© 1986 IEEE).

8.4 FUTURE TRENDS IN WELL FORMATION

It seems that high-energy ion implantation will find more application for generating profiled wells. Profiled wells are achieved by multiple implants at varying energies. They offer the ability to independently control several MOSFET characteristics, including threshold voltage, hot-carrier generation, punchthrough voltage, junction capacitance, latchup susceptibility, soft-error immunity, and the body-effect coefficient. The optimization of well profiles for improving device performance has been termed "well engineering."³¹

Unlike a conventional well, which has an almost flat impurity distribution, a profiled well can have an impurity profile tailored to independently optimize a variety of device characteristics. Figure 8-21 shows the impurity profile of a multiply-implanted well and the relation of the device characteristics to the profile.³²

In profile-implanted wells formed with high-energy implants, the wells are implanted after LOCOS isolation. In the n -well, multiple high-energy implants of phosphorus ions form a profiled well for active as well as isolation regions (with a thick photoresist serving as the p -well mask). Subsequently, low-energy implants are performed to adjust the PMOS V_T . A profiled p -well is next produced in the same way.

The deepest implant of the well in this example is fabricated with a 1 MeV implant. These implanted dopants form a heavily doped buried layer (10^{18} - 10^{19} cm^{-3}) at a depth of 2-4 μm that serves to reduce latchup susceptibility. Such a p^+ buried layer is more effective than an epitaxial layer in reducing substrate resistance and suppressing latchup, thus eliminating the need for an epi layer. The ability to use non-epi substrates when wafer sizes reach 150 mm and 200 mm represents a major cost savings, as the cost of 200 mm epitaxial wafers are ~\$100 US higher than bulk-CZ wafers. The replacement of epitaxial wafers with MeV implantation processing also alleviates problems with autodoping control and backside sealing.³⁵

The use of MeV implantation avoids the long (3-21 hour), high-temperature (1150-1200°C) well-formation drive-in step of the conventional process. This minimizes

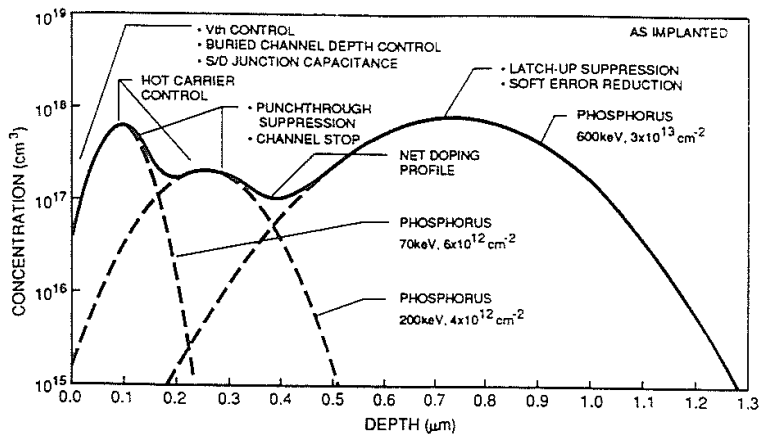


Fig. 8-21 The relation between device characteristics and the impurity concentration at various depths in multiple implanted well.³² Reprinted with permission of Solid State Technology.

thermally induced stress in large wafers that can cause wafer warpage, bow, and flatness degradation, as well as offering simplified processing with shorter fabrication times. Process simplification also results from the fact that the retrograde-well and field-isolation regions are formed by the same implant (after field oxidation). As noted in section 8.3.1, implantation through field oxide can also improve active device isolation characteristics. Employing MeV implantation to dope the bottom of the well more heavily leaves more room to tailor the doping profile at the top of the well in the channel area with lower energy implants.

Concern has existed that the high-energy implants may cause damage that will degrade device and circuit performance. Increased junction leakage is observed if the dose or energy are too low (i.e., $<3 \times 10^{14} \text{ cm}^{-2}$ and $\sim 1\text{--}1.5 \text{ MeV}$, respectively). Recent studies, however, indicate that if the dose and energy exceed $3 \times 10^{14} \text{ cm}^{-2}$ and 1.5 MeV , respectively, leakage due to implant damage is dramatically reduced with both boron and phosphorus implantation. In fact, when an optimum combination of dose, energy, and annealing process is used, less leakage is observed than in control devices having no buried layer. This behavior is believed to be caused by gettering of electrically active damage or impurity defects away from the depletion regions of active transistors by secondary defects near the concentration peak of the buried layer. The effect is called "self-gettering." For all of the above reasons it appears that high-energy implantation will become more widely used to form wells in device generations at and below $0.5 \text{ } \mu\text{m}$.

8.5 AFTER-GATE IMPLANTATION (AGI)

An alternative technique which exploits high-energy implantation to reduce the complexity of CMOS process called *after-gate implantation* (AGI) has been suggested. That is, by performing all implantations after the gate electrodes are patterned (wells, channel-stop doping, and source/drain), the number of mask steps can be reduced.

Figure 8-22 illustrates this approach, in which only seven masks are required to form a single p -well CMOS structure.³³ The LOCOS-isolation field-oxide regions are formed first, then the polysilicon gates are deposited and patterned. Using a mask that obscures the PMOS regions, the NMOS p -well implant, multiple-channel implants, and the n^+ source/drain implants are carried out. Then a mask is used to cover the NMOS areas, and the PMOS-channel and p^+ source/drain implants are performed. Since the well and channel implantations must penetrate the field oxide and the polysilicon gate, they must be carried out at higher energies. The p -well implant energy is 600 keV (at a dose of $1 \times 10^{13} \text{ cm}^{-2}$), and the NMOS devices receive low-dose channel implants of 135 keV, 175 keV, and 215 keV. The channel of the PMOS devices receives two low-dose, higher-energy phosphorus implants (at 220 keV and 440 keV). Such multiple-channel implants are used in both NMOS and PMOS devices to avoid unacceptable V_T variations caused by polysilicon gate thickness variations encountered in processing (and to improve the punchthrough resistance of the device structures). Studies also indicate

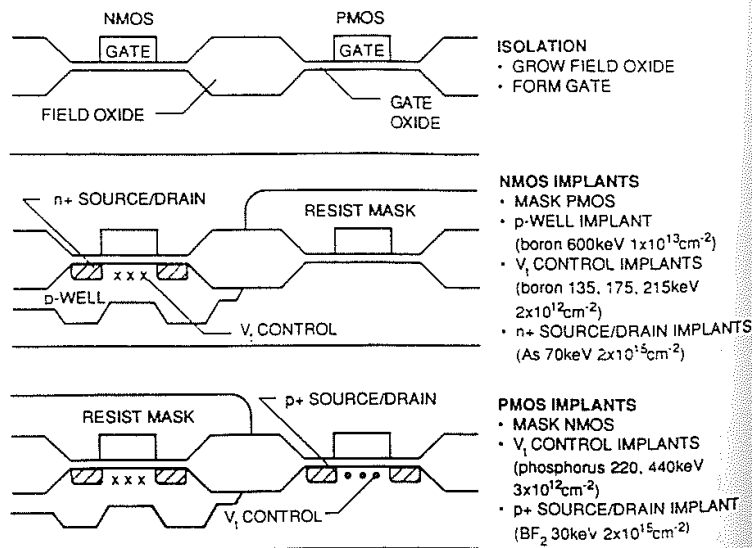


Fig. 8-22 After-Gate Implant (AGI) process sequence.³³ (© 1986 IEEE).

that the high energy AGI does not degrade the gate oxide integrity³² nor the MOS transistor characteristics.³⁴

The AGI process thus offers the benefits of implantation processing freedom and flexibility, as well as reduced CMOS process complexity. The low doses required for the well and channel implantations can be performed with a medium-current implanter which may also have a higher energy implantation capability.

REFERENCES

1. J.Y. Chen, "CMOS - The Emerging Technology," *IEEE Circuits and Devices Magazine*, March 1986, p. 16.
2. J.Y. Chen, *CMOS Devices and Technology for VLSI*, Prentice-Hall, Englewood Cliffs, N.J., 1990.
3. A.G. Lewis et al., *IEEE Trans. Electron Dev.*, **ED-34**, 2156, (1987).
4. T. Yamaguchi et al., *IEEE Trans. Electron Dev.*, **ED-31**, p. 205 (1984).
5. L.C. Parrillo "CMOS Active and Field Device Fabrication," *Semiconductor International*, April 1988, p. 64.
6. W. Wijaranakula et al., *J. Electrochem. Soc.*, December 1988, p. 3113.
7. H. Tsuya et al., *Japanese J. Appl. Phys.*, **22**, L16 (1983).
8. A. G. Lewis et al., "Vertical Isolation in Shallow *n*-Well CMOS Circuits," *IEEE Electron Dev. Lett.*, EDL-8, p. 107, 1987.
9. K. Tanno, F. Shimura, and T. Kawamura, *J. Electrochem. Soc.*, **128**, p. 395 (1981).
10. J. Manoliu and J.O. Borland, "A Submicron Buried Layer Twin-Well for High Density CMOS," *Tech. Dig. IEDM*, 1987, p. 20.
11. R. de Werd et al., *Tech. Dig. IEDM*, 1987, p. 532.
12. L.C. Parrillo et al., "Twin-Tub CMOS-II," *Tech. Dig. IEDM*, 1982, p. 706.
13. R.A. Chapman et al., "A 0.8 μm CMOS Technology," *Tech. Dig. IEDM*, 1987, p. 362.
14. M.-L. Chen et al. *Tech. Dig. IEDM* 1986, p. 256.
15. S.J. Hillenius and L.C. Parrillo, U.S. Patent No. 4,554,726, Nov. 25, 1985.
16. R.A. Chapman et al., "High Performance Sub-Half Micron CMOS Using Rapid Thermal Processing," *Tech. Dig. IEDM* 1991, p. 101.
17. Ch. Zeller, C. Mazure, & M. Kerber, "Field-Implant-Free Isolation by Double-Well Split Drive-In," *IEEE Electron. Dev. Letts.*, p. 215, May 1990.
18. R. Martin and J.Y. Chen, "Optimized Retrograde *n*-well for 1 μm CMOS, *Proc. Custom Integ. Circuits Conf.*, 1985, p. 199.
19. R.D. Rung, C.J. Dell'Oca, and L.G. Walker, "A Retrograde *p*-well for High Density CMOS," *IEEE Trans. Electron Dev.*, **ED-28**, p. 1115, 1981.
20. D.M. Brown et al., "Trends in Advanced CMOS Process Technology," *Proceedings of the IEEE*, December 1986, p. 1646.
21. S.R. Coombs, "Scalable Retrograde *p*-well CMOS Technology," *IEEE Trans. Electron Dev.*, **ED-28**, p. 346, 1981.
22. Y. Taur et al., *IEEE J. Solid-State Circuits*, **SC-20**, p. 123, 1985.

23. G.J. Hu et al., "Design and Fabrication of p -channel FET for 1- μm CMOS technology," *IEDM Tech. Dig.* 1982, p. 710.
24. A. Stolmeijer, "A twin-well CMOS process employing high-energy ion implantation," *IEEE Trans. Electron Dev.*, **ED-33**, April, 1986, p. 450.
25. A. Shida et al., 3.3V BiNMOS Technology Using NPN Transistors without Buried Layers," *Tech. Dig. IEDM*, 1991, p. 93.
26. Y. Mochizuki, "Special Report: High Energy Ion Implantation Technology," Nikkei Microelectronics, (Dec. 1991), p. 92.
27. M. Moslehi and K. Saraswat, "Thermal Nitridation of Si and SiO_2 for VLSI," *IEEE Trans Electron Dev.* **ED-32**, p. 116, January 1985.
28. P. Fahey, R.W. Dutton, and M. Moslehi, "Effects of Thermal Nitridation Processes on Boron and Phosphorus Diffusion in $\langle 100 \rangle$ Silicon," *Appl. Phys. Letts.*, **43**(7), p. 683 July, 1983.
29. J.Y. Chen, "Quadruple-well CMOS for VLSI Technology," *IEEE Trans. on Electron Dev.*, **ED-31**, p. 910, 1984.
30. A.G. Lewis, R.A. Martin, and J.Y. Chen, "Retrograde and conventional n -well CMOS technologies: A Comparison," Internal Report, EIL-87-17, Xerox Palo Alto Research Center, Oct. 1986.
31. K. Tsukamoto et al., *Nucl. Instrumentation & Methods*, B59/60, p. 584 (1991).
32. K. Tsukamoto et al., "High Energy Ion Implantation for ULSI: Well Engineering and Gettering," *Solid State Technology* P. 49, June 1992.
33. H. Mikoshiba et al., "A Novel CMOS Process Utilizing After-Gate Implantation Process," *Dig. of Tech. Papers, 1986 Symp. VLSI Tech.*, p. 41.
34. J. Pfister et al., "A Self-aligned Elevated Source/Drain MOSFET," *IEEE Trans. Electron Dev.*, September 1990, p. 365.
35. J.O. Borland and R. Koelsch, "MeV implantation technology: Next generation manufacturing with current generation equipment," *Solid-State Technology*, December 1993, p. 28.
36. J.Y. Chen, *VLSI Design* July 1984, p. 78.
37. A.G. Lewis et al., "Device Isolation in High Density LOCOS-Isolated CMOS," *IEEE Trans. Electron Devices*, **ED-34**, p. 1337, June 1987.

PROBLEMS

1. An n -well n^+ polysilicon-gate CMOS process starts with a substrate doping of $3 \times 10^{15}/\text{cm}^3$. The well doping near the surface of is approximately constant to a level of $3 \times 10^{16}/\text{cm}^3$. The gate oxide thicknesses are both 40 nm, and $Q_{\text{tot}} = 3 \times 10^{10}/\text{cm}^2$.
 - (a) Calculate the threshold voltages of the n - and p -channel transistors.
 - (b) Calculate the boron doses needed to shift the NMOS threshold to +1V and the PMOS threshold to -1V. Assume that the threshold shifts are achieved through shallow ion implantations.
2. In a p -well CMOS technology, the n -substrate has a doping concentration of $1 \times 10^{15}/\text{cm}^3$, and a p -well concentration near the surface of $5 \times 10^{16}/\text{cm}^3$. Calculate the source/drain junction capacitances (cm^{-2}) and body effect coefficients for both n - and p -channel MOSFETs assuming the gate oxide is 20 nm thick in both devices.

3. Describe the split well-drive process, and the advantages it offers compared to the conventional well-drive process.
4. Comparing retrograde well to conventional well, what is the most important difference in process technology and finished device structure? How does it affect device isolation?
5. In a (p^+ -diffusion)- n -well-(p -on- p^+ substrate)structure, the n -well is uniformly doped at $1 \times 10^{16} \text{ cm}^{-3}$ and is $1 \mu\text{m}$ deep, the p -epi thickness is $4 \mu\text{m}$ and is uniformly doped at $1 \times 10^{15} \text{ cm}^{-3}$, and the p^+ diffusion is $0.4 \mu\text{m}$ deep. If a substrate bias generator is needed and the p^+ diffusion and n -well are both biased at 5V , what is the maximum (in absolute value) substrate voltage that can be applied before punchthrough occurs?

