NANENG 301 – Nano/Micro Fabrication techniques.

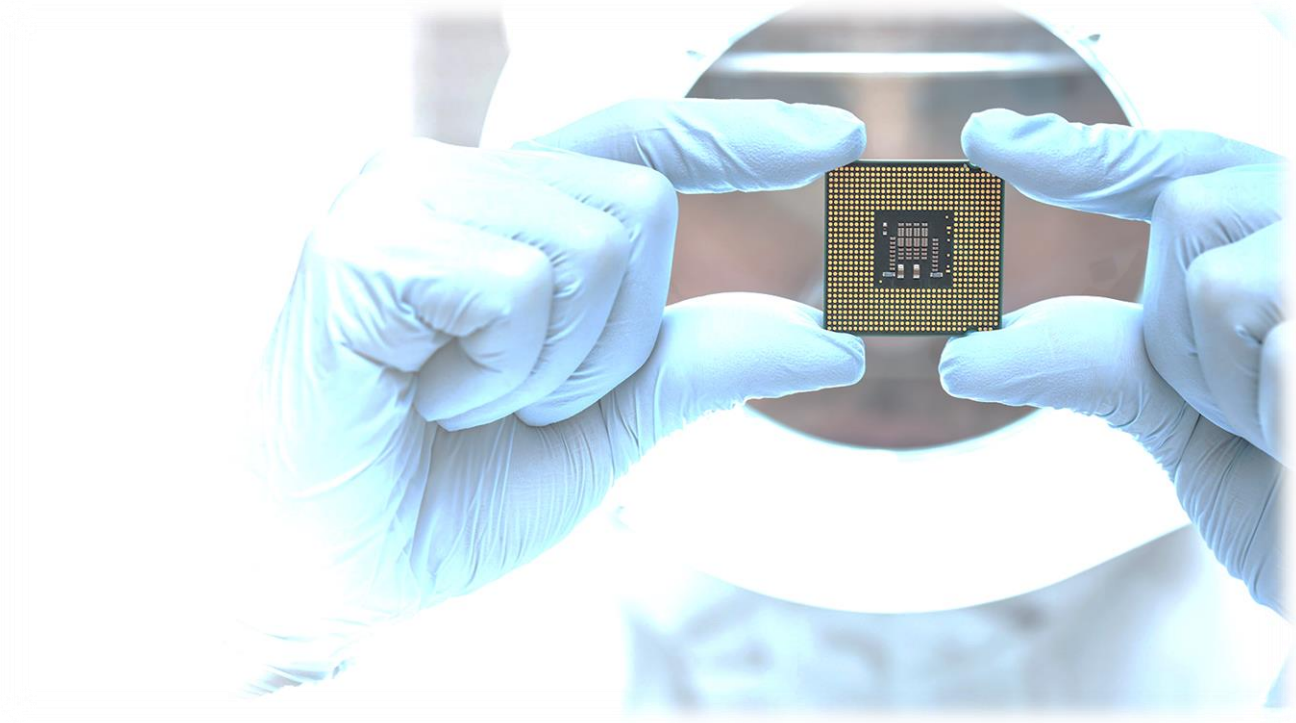
Final Project: Design and Fabrication of MOS-Structure

Supervisor: Dr. Yasmine Elogail, Eng. Kirollos Ernest & Eng. Alaa Mahmoud

University of science and technology at Zewail city.

Team members:

Eman Ashraf Teleb	201800854
Karim Hamdy Mostafa	201800519
Mohamed Abdelnaby	201800510



I- Overview:

In this paper, there is a full analysis of the fabrication flow of CMOS 250 nm technology; a short of complementary metal-oxide technology at which PMOS and NMOS are integrated on the same wafer. Among logic families, CMOS is widely used in electronics architectures for more than 90% of today's market. This paper is intended to design the process of an inverter, the basic CMOS structure. Due to the zero static power and robustness to noise advantages, CMOS is widely used regardless of its disadvantages such as the large area and complexity in fabrication that require extra sophisticated steps compared to NMOS technology. The proposed structure responds to high and low margins of voltage, within high margins the PMOS is off while the NMOS is on and connects the ground to the output resulting in a zero logic. On the other hand, with the low voltage margin the NMOS is disabled and the PMOS connects the VDD to the output charging the capacitance node. This is the use in digital architectures while in analog applications the inverter can be used as an amplifier with gain equals to -1. Nowadays, current research is working on getting a higher absolute gain by operating at the area between the high and low margins at which the two MOSFETs are in saturation [1]. There is full analysis and calculations for each fabrication step starting from choosing the wafer to get a full function CMOS inverter with metals connecting the isolated active regions to the inputs.

II- Wafer dimensions and preparations:

For the proposed structure (structure 1), the choice will be a lightly doped P wafer with $6x10^{16}cm^{-3}$ doping. As the (100) orientation has the best interface with oxide, we decided to choose it. For the cleaning process, the wafer is set for RCA cleaning to make sure the wafer becomes impurities free. The wafer thickness is chosen to be 2.5 micrometer and a width of 1.7 micrometer. Dimensions are shown in figure (1).

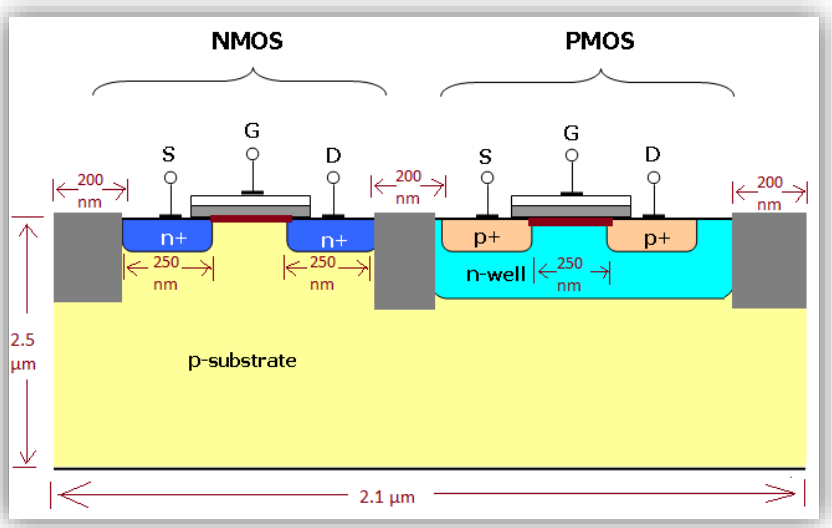


Figure (1) structure chosen dimensions.

III- Shallow trench isolation:

To ensure that there is no charge dissipation and each MOSFET works properly; with no effect on each other or with the near ICs if any, isolation is needed. This is done using Shallow trench isolation that is an oxide region between both MOSFETs and at the corners of the wafer. Firstly, an oxide layer is constructed using oxidation in a hot furnace with H_2O flow for high speed at 900 C to produce a thickness of 40 nm [2].

$$Si + 2H_2O \rightarrow SiO_2 + 2H_2$$

After that, the wafer is sent to a deposition furnace at 800 C to form a silicon nitride layer. Under low pressure, the LPCVD is done to ensure high uniformity of silicon nitride layer, a layer of 80 nm is grown.

$3SiH_4 + 4NH_3 \rightarrow Si_3N_4 + 12H_2$

While the silicon nitride layer performs a stress into the wafer, the oxide layer performs an opposite stress to relieve the silicon wafer.

Oxidation coefficients for silicon

	Dry		
Temperature (°C)	A (µm)	B (µm²/hr)	τ (hr)
800	0.370	0.0011	9
920	0.235	0.0049	1.4
1000	0.165	0.0117	0.37
1100	0.090	0.027	0.076
1200	0.040	0.045	0.027

Figure (2) Oxidation parameters [3].

Using figure (2), the calculations are as follows. $t = \frac{d_{ox}^2 + Ad_{ox}}{B} = \frac{0.04^2 + 0.165 * 0.04}{0.0117} = 42\ mins.$

After that, the wafer is sent to a deposition furnace at 800C to form a silicon nitride layer. Under low pressure, the LPCVD is done to ensure high uniformity of silicon nitride layer. A layer of 80 nm is grown in a 800C furnace. While the silicon nitride layer performs a stress into the wafer, the oxide layer performs stress in an opposite direction; so it is used to relieve the silicon wafer and ensure defects. Structure shown in figure (3).

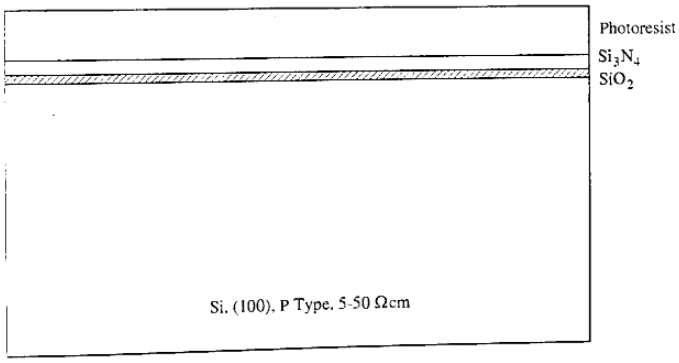


Figure (3) structure illustration [2].

After that the photoresist is placed getting ready for photolithography. The lithography liquid is spun for symmetry and to obtain a 1000 nm thickness of resist. Then baked in 100 C to drive off the solvents. Using A i-line of mercury arc lamp to be the stepper which will result in a contrast of 3.6 in the AZ-1350 and use projection printing and for k1=.7 and a numerical aperture of .8 in the medium the minimum feature is to be 319.4 nm.

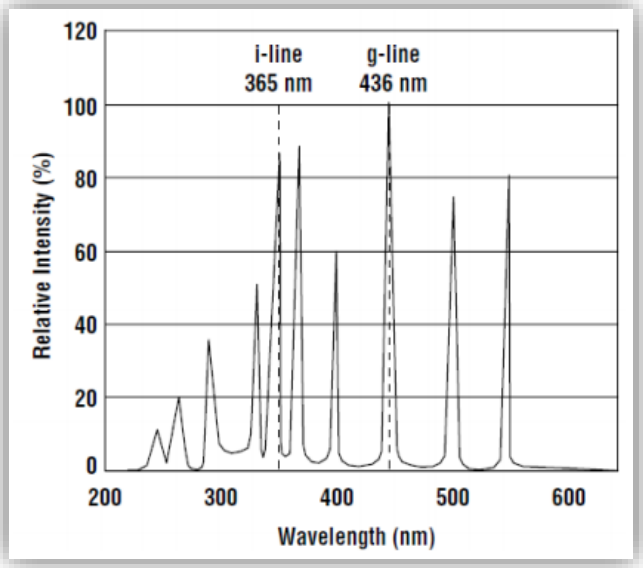


Figure (4) Relative intensity vs wavelength [3].

Table 8.1 Contrasts for several commercial resists for different wavelengths, λ

λ (nm)	AZ-1350	AZ-1450	Hunt 204
248	0.7	0.7	0.85
313	3.4	3.4	1.9
365	3.6	3.6	2
436	3.6	3.6	2.1

Data from Leers [13]. The AZ formulations are products of Shipley.

Figure (5). Different resist contract for different wavelengths [3].

Then the mask is used to pattern the LOCOS regions. Then the silicon nitride is etched. Using plasma dry etching due to its high directionality using fluorine plasma following the equation:

$$Si_3N_4+12F_3 \rightarrow SiF_4+2N_2$$

And etching the silicon oxide will flow the equation:

$$SiO_2+4F \rightarrow SiF_4+O_2$$

Assuming only a flux of ionic F equal to 10^{16} atmos/cm^{-3} and K=1, the etch rate for silicon nitride is $7.35x10^{-7}$ cm/sec and hence t1=10.88 sec and for silicon dioxide is $3.76x10^{-7}$ cm/sec so t2=15.92 sec.

After 26.8sec of plasma dry etching, the silicon nitride and silicon oxide are etched and silicon becomes an exposed region. Shown in figure (6).

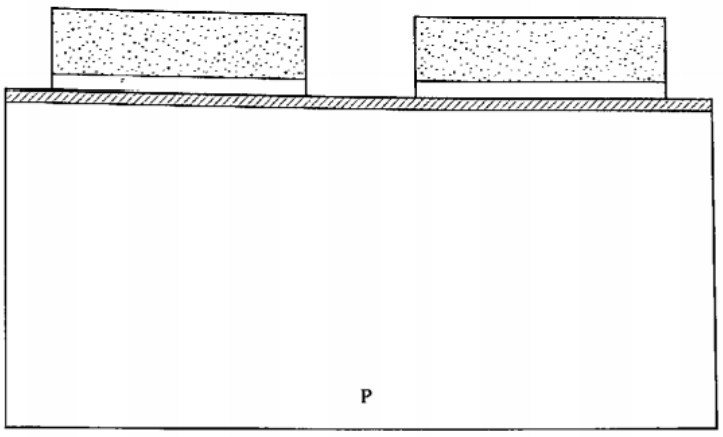


Figure (6). Silicon exposed after etching [2].

Then etch the silicon itself by ion-enhanced etching using bromine. Assuming Fc, flux chemical species, = $2x10^{18}$ atoms/cm^3 sec with sticking coefficient 0.01 And Fc, flux ionic species, = $2x10^{16}$ atoms/cm^3sec with sticking coefficient 1, The etch rate is $2x10^{-7}$ cm/sec.

For etching 500 nm it takes 250sec. While the photoresist is removed using AZ 351B as a developer.

For a better interface of silicon, an oxide layer of 20 nm is formed for 18.97 min. The regions of isolation are ready to be filled using high density plasma chemical vapor deposition to ensure filling the gaps with no knocks at a low temperature. Using CMD chemical mechanical polishing, a polish can be used to remove the excess of the oxide layer to get a smooth oxide layer. Structure shown in fig (7).

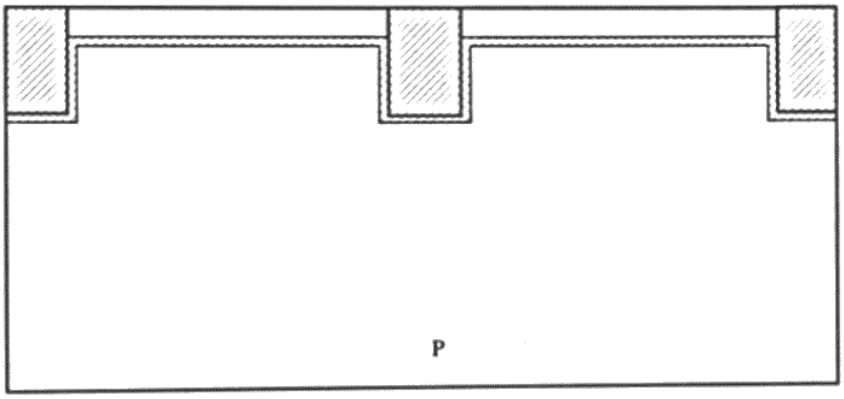


Figure (7). Structure after STI is done [2].

IV- N well formation:

The formation of the n well for the PMOS will undergo two steps; 1) ion implantation of the dopant, 2) drive-in diffusion to reach the needed junction depth. First, a mask is applied on the NMOS region to prevent any dopants from diffusing in this region. Then, phosphorus will be used as the implant in this process to be the dopant. Phosphorus is chosen as it has a high diffusion coefficient and therefore, it will need less energy for the next process. The implantation dose is typically in order of 10^{13} cm^{-2} to reach a concentration of 5 $x10^{16}$ cm^{-3}. The typical used energy for this process is 350 keV. The projection depth and the standard deviation in this process will not have a big effect on the result as a redistribution of dopants will occur in the drive-in diffusion. The next step is to make the dopants reach the required depth. This will occur through a drive in diffusion process at a temperature of 1000 degrees Celsius. In order for the dopants to get to the required depth of 0.5

microns. At this temperature, the diffusion coefficient of phosphorus will be 1.3627×10^{-14} cm^2/sec. From the equation of the junction depth, we numerically calculated the time it would need to diffuse for 0.5 microns which will be equal to 8.267 hours. The junction depth equation is as follows:

$$x_j = \sqrt{4Dt * \ln(\frac{Q_T}{C_B * \sqrt{\pi Dt}})}$$

V- **Gate formation:**

The first step is applying a photoresist on the NMOS region to implant a dose of boron. The energy used for implantation will not make the implants penetrate deeply into the substrate; it will only be on the surface of the silicon substrate. Process shown in figure (8).

Then, after removing this mask, another mask will be applied on the PMOS region to implant arsenic with a slightly higher energy because the arsenic is heavier than boron. Process shown in figure (9).

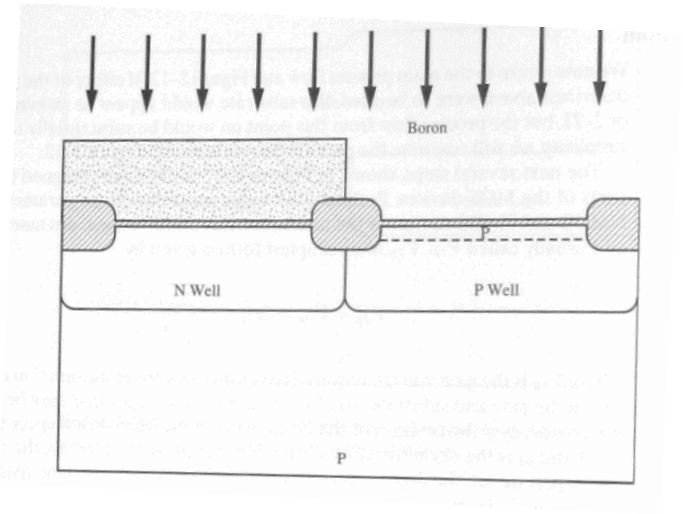


Figure (8). Masking NMOS to implant Boron [2].

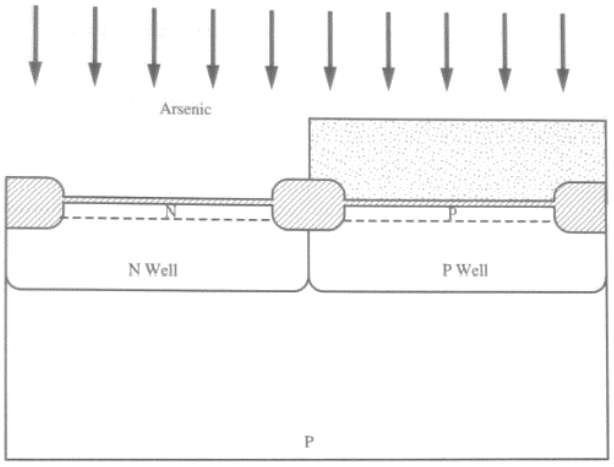


Figure (9). Masking PMOS to implant Arsenic [2].

After removing the second mask as well, an etching process will take place in order to reduce the thin layer of oxide on both the NMOS and PMOS active regions. HF is used due to its high selectivity as it will stop when the silicon is reached. Shown in figure (10), the wafer after applying HF.

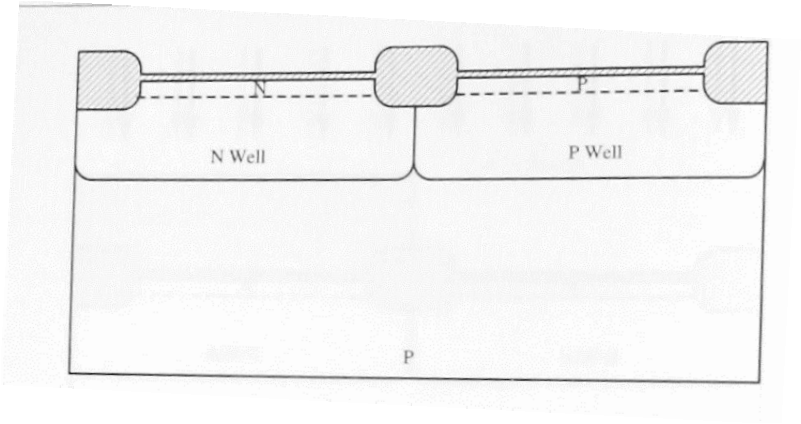


Figure (10). The wafer after HF application [2].

The dose, projection depth and straggle are calculated as follows. For ion implantation Junction depth x_j = 80 nm. As $N(x_j) = N_p * \exp\left(-\frac{x_j - R_p}{2\Delta R_p^2}\right) = N_B$ Rearranging we get, $x_j = R_p + \Delta R_p \sqrt{2\ln(\frac{N_p}{N_B})}$.

In addition, $R_p = \Delta R_p = a_0 + a_1 E + a_2 E^2$.

projected range

ion	a_0	a_1	a_2
B	27.5934	2.6199	-3.23x10^{-4}
P	3.0793	1.24786	6.0233x10^{-4}
As	6.4916	0.60656	-1.8822x10^{-4}

straggle

ion	a_0	a_1	a_2
B	13.46	0.8025	-2.162x10^{-4}
P	3.0032	0.5322	-5.5981x10^{-4}
As	2.3082	0.2245	-1.8085x10^{-4}

Figure (11). Projected range and straggle coefficients [3].

Substituting in the given equation and using the shown tables of figure (11), and using Bulk concentration N_B= 6x10^{16} cm^{-3} and Peak concentration N_p= 7x10^{17} cm^{-3} for NMOS and N_p = 8.4x10^{18} cm^{-3} for PMOS. We get the following values for Boron: R_p= 41.049 nm, ΔR_p=17.576 nm, E=5.1361 Kev and Φ=3.084x10^{12} cm^{-2}.

We get the following values for Arsenic: R_p= 37.55 nm, ΔR_p=13.502 nm, E=52.0435 Kev and Φ=2.843x10^{13} cm^{-2}. Then an oxide layer will be grown as the gate oxide; shown in figure (12). The grown oxide layer is 4.8 nm. Using figure (13), Operating at 1000 C, the rate is about 40 A/min, that is equal to 4 nm/min. The desired oxide thickness is 4.8 nm. So the time taken would be 1.2 mins.

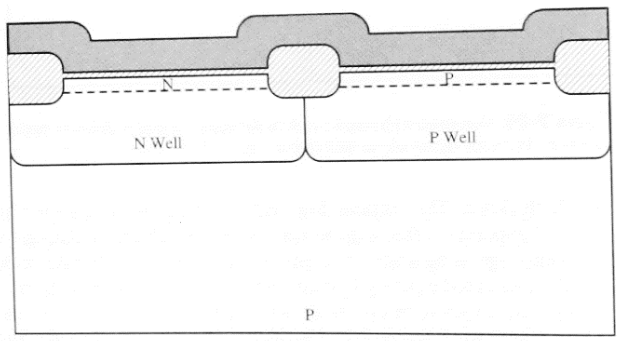


Figure (12). The grown oxide layer [2].

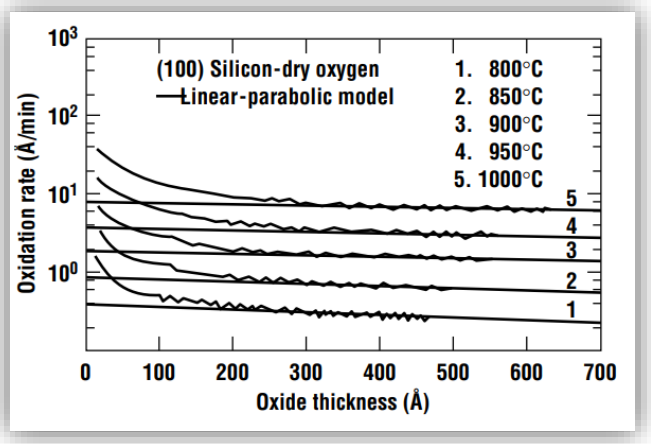


Figure (13). Oxidation rate vs oxide thickness [3].

Afterwards is the process of depositing and doping the polysilicon layer. The polysilicon has the thickness of 0.4 microns. Using LPCVD a layer of polysilicon will be deposited over the oxide layer. The source of the polysilicon will be silane. By thermal decomposition of silane, silicon is produced which will be amorphous or polycrystalline according to the decomposition temperature. The polysilicon is then highly doped with n-type doping using phosphorus or arsenic to increase its conductivity with a dose of $5*10^{15}$ cm^{-3}.

The LPCVD process: This process was chosen for its high uniformity and purity. The used temperature in this process is 580 Celsius. The deposition pressure used will be around 200 mtorr which will lead to a deposition rate of about 3.3 nm/min as shown in figure (14). Since the product will be polycrystalline, no further annealing is needed. In order to reach the required thickness of 0.4 microns, it will take 121.21 minutes.

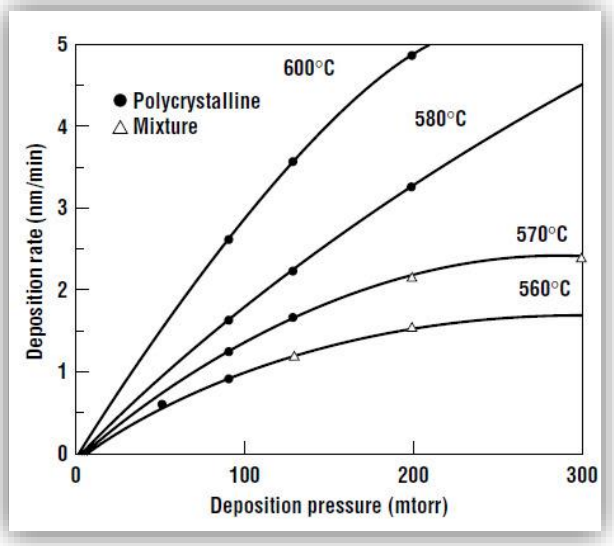


Figure (14). Deposition rate as a function of temperature vs deposition pressure [3].

VI- **Extension formation:** no extension would be needed as there will be no short channel effects with the given parameters.

VII- Source and drain formation:

Implanting source and drain areas needs to be done through a thin oxide layer to avoid channeling. A 10 nm oxide is grown before implanting source and drain to randomize the direction of implantation and thus reducing channeling. Using figure (13), Operating at 1000 C, the rate is about 60 A/min, that is equal to 6 nm/min. The desired oxide thickness is 8 nm. So the time taken would be 1.333 mins.

First, PMOS is protected using a mask and photoresist. And for NMOS, Arsenic is used for the junction to be shallow. Given desired junction depth of 80 nm, we can assume that R_p would be almost in the middle at 40 nm. Using figure (15), we get the energy for Arsenic to be 50 Kev. Using this energy and figure (16), we get ΔR_p of 7.5 nm. Implantation dose of $3x10^{15}$ cm^{-2} would be sufficient for Arsenic to penetrate the thin oxide and get implanted and still get easily masked later. Using $\varphi = \sqrt{2\pi}\Delta R_p N_p$ and the previously obtained parameters we get $N_p = 1.59x10^{21}$ cm^{-3}. Process shown in figure (17).

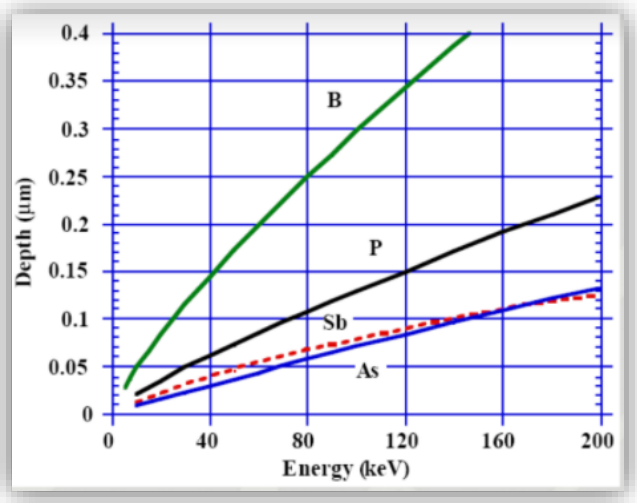


Figure (15). Energy vs. projection depth for different materials [3].

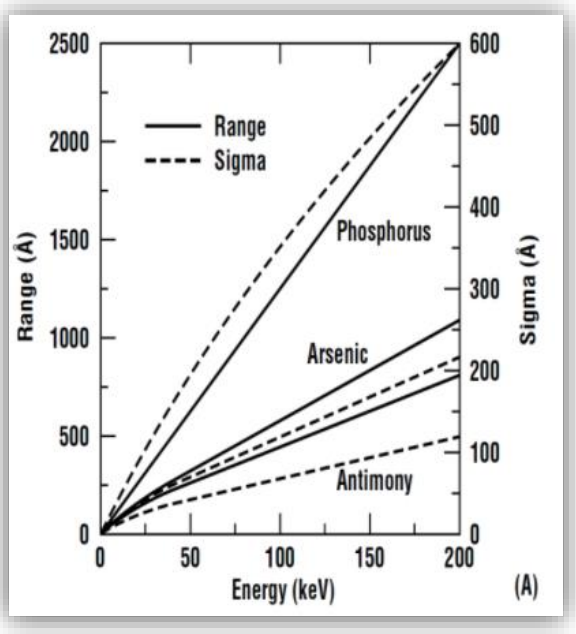


Figure (16). Range and sigma vs. Energy for different materials [3].

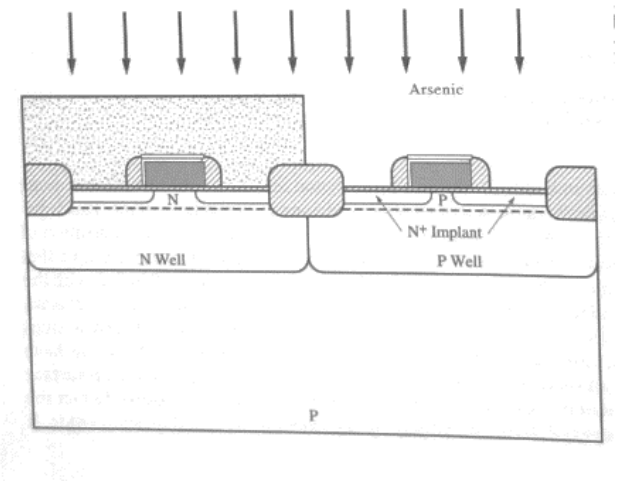


Figure (17). Protecting PMOS to implant Arsenic [2].

Now the NMOS is protected using a mask and photoresist. The PMOS source and drain are to be implanted with boron, as shown in figure (18). Using figure (15), the energy should be 7 Kev. Using this energy from figure (19), we get ΔR_p of 10 nm. Implantation dose $3x10^{15}$ cm^{-2} should be sufficient. Using the parameters we get, $N_p = 1.196x10^{21}$ cm^{-3}.

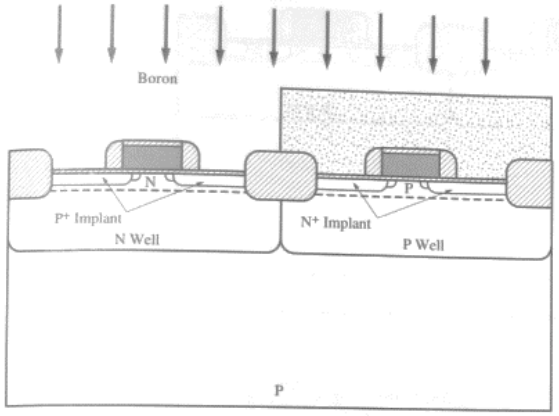


Figure (18). Protecting NMOS to implant Boron [2].

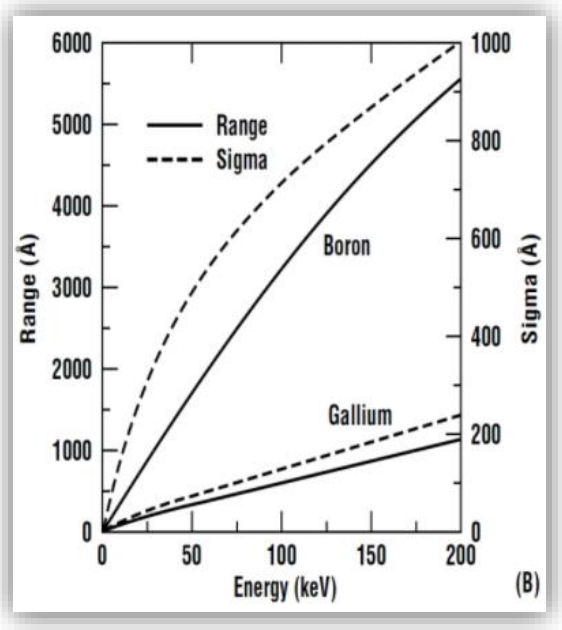


Figure (19). Range and sigma vs. Energy for different materials [3]

The energy used is lower because boron is much lighter than arsenic and thus requires less energy to obtain the same depth of 0.08 micrometer. Finally, the device would undergo rapid thermal annealing for 1 min at 1000 C to activate all implants, annealing damage, and diffusing the junctions to their depths.

References:

[1] W. Wilson, T. Chen and R. Selby, "A current-starved inverter-based differential amplifier design for ultra-low power applications," 2013 IEEE 4th Latin American Symposium on Circuits and Systems (LASCAS), 2013, pp. 1-4, doi: 10.1109/LASCAS.2013.6519040.

[2] J. D. Plummer, M. D. Deal, and P. B. Griffin, Silicon VLSI technology: fundamentals, practice, and modeling. Upper Saddle River, NJ: Prentice Hall, 2000.

[3] S. A. Campbell, Fabrication engineering at the micro and nanoscale. New York: Oxford University Press, 2013.