In the name of God

Silvaco Project - Diffusion

Fabrication Course

Amirhossein Zahedi 404211388

February 2026

Introduction:

In this project, we will simulate diffusion processes using Silvaco ATHENA. It has several stages, including wafer initialization, pre-deposition, drive-in, and ion implantation. The objective is to analyze how thermal diffusion and ion implantation affect junction depth (Xj), sheet resistance (Rs), and surface concentration (Ns). Practical constraints such as solid solubility limits, mobility degradation, and thermal budget limitations are considered. The results are analyzed to determine which process targets are realistic, difficult, or physically impossible.

The project evaluates:

- The formation of diffusion profiles under constant-source (pre-dep) and limited-source (drive-in) conditions.
- The dependence of junction depth and sheet resistance on thermal budget
- Trade-offs between Xj and Rs for different processing targets
- Multi-step diffusion strategies
- Replacement of pre-deposition with ion implantation

Theory:

Diffusion is the redistribution of atoms from regions of high concentration of mobile species to regions of low concentration. In silicon processing, diffusion is governed by Fick's First and Second Laws**.**

Fick's first law: $F(x,t) = -D\frac{\partial C(x,t)}{\partial x}$

Where J is the dopant flux, D is the diffusion coefficient, C is the concentration, and x is the depth. D is itself $D_i = D_0\, exp(-\frac{E_a}{kT})$ and it is strongly temperature dependent. Ea is the activation energy, and k is the Boltzmann constant.

Fick's second law: $\frac{\partial C(x,t)}{\partial t} = D\frac{\partial^2 C(x,t)}{\partial x^2}$

This equation describes how concentration evolves with time.

Two major regimes are used in diffusion:

1. Pre-deposition or constant-source diffusion:

During pre-deposition, the wafer surface is exposed to a constant dopant source. The surface concentration (Ns) reaches the solid solubility limit and remains fixed; however, the total dose (Q) increases with time**.** $C(x,t) = N_s\left[erfc\left(\frac{x}{2\sqrt{Dt}}\right)\right]$, $Q = 2N_s\sqrt{\frac{Dt}{\pi}}$

2. Drive-in or limited source diffusion:

After removing the dopant source, redistribution occurs without additional dopant supply, and the solution becomes Gaussian: $C(x,t) = \frac{Q}{\sqrt{\pi Dt}}exp\left(-\frac{x^2}{4Dt}\right)$

In this process, profile spreads and surface concentration (Ns) decrease; meanwhile, junction depth will increase approximately with $\sqrt{t}$.

Some parameters are important in diffusion. Surface concentration and total dose were mentioned. Junction depth and sheet resistance have critical roles, too. Junction depth is the depth at which the concentration is equal to the background doping. Sheet resistance is also the resistance of the wafer that is diffused: $R_s = \frac{1}{q\int_0^\infty \mu(N)C(x)dx}$

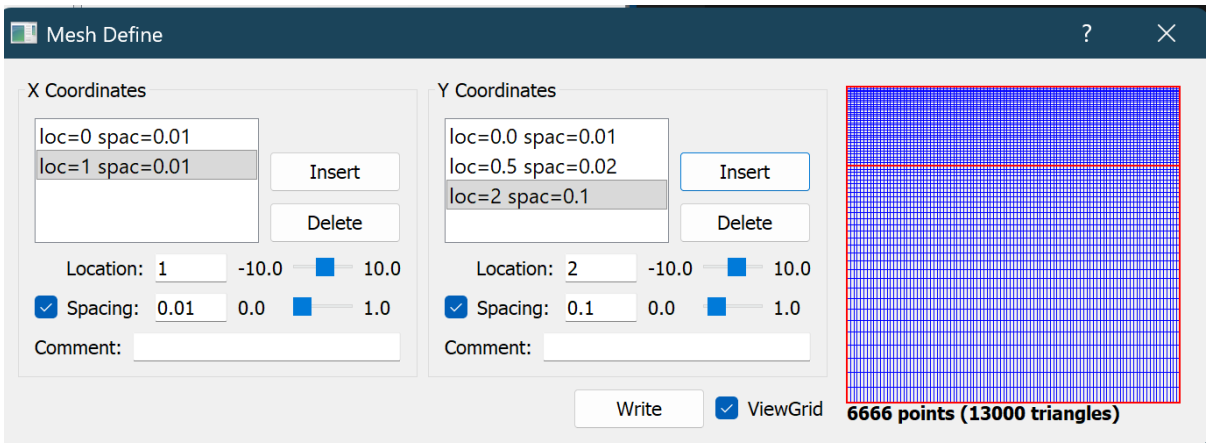
$\mu\big(n(x)\big)$ is the mobility of carriers, and it is concentration-dependent. At high doping, mobility decreases due to impurity scattering, but Rs does not decrease linearly with increasing dose.

Solid solubility limit: At high concentrations, dopant atoms cannot remain electrically active beyond a certain limit.

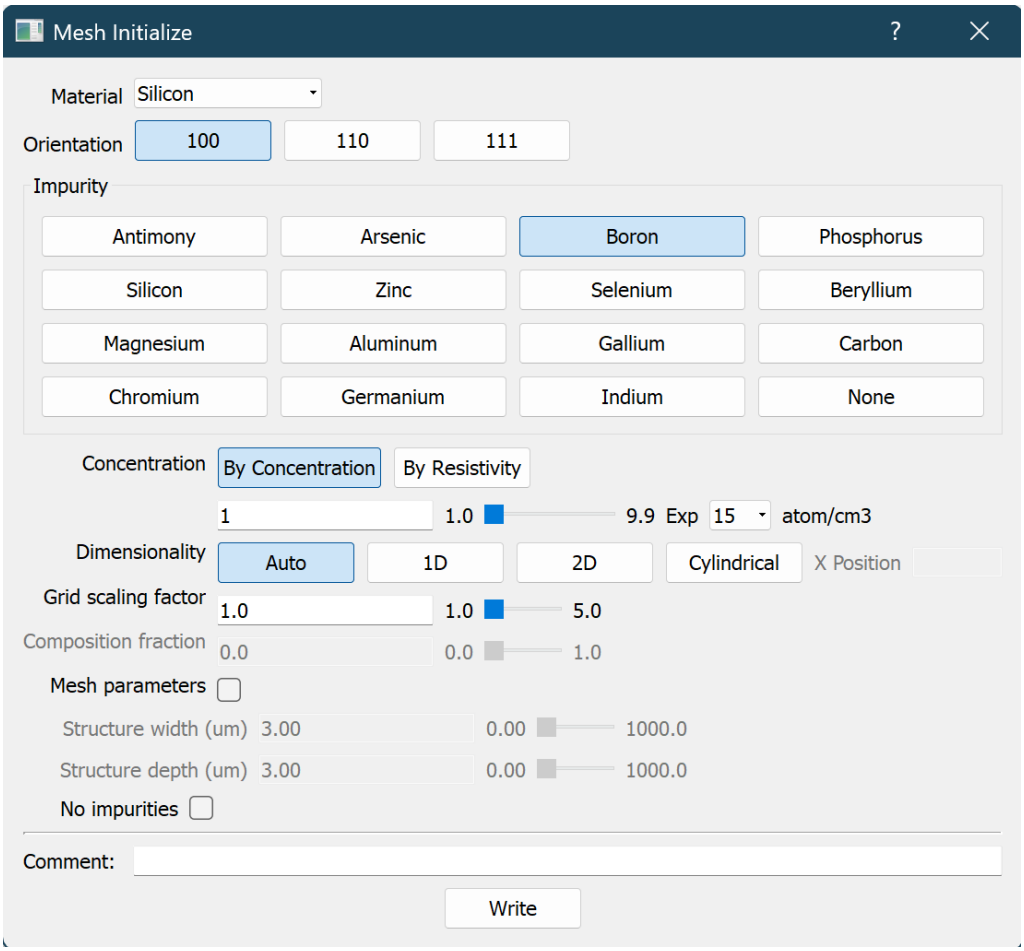
- ➔ Increasing drive-in time → increases Xj → decreases Ns → increases Rs
- ➔ Increasing pre-dep time → increases Q → decreases Rs → increases Xj after drive-in

Main Project:

To begin the process, the mesh should first be defined and initialized.



A <100> silicon wafer doped with Boron at 1e15 at the temperature of 900 C. To define the mesh, since the near-surface regions are more important, we will initially make them denser.



Mesh

line x loc = 0.0 spacing=0.01

line x loc = 1.0 spacing=0.01

line y loc = 0.0 spacing = 0.01

line y loc = 0.5 spacing = 0.02

line y loc = 2.0 spacing = 0.10

Initializing the Wafer

init silicon c.boron=1e15 orientation=100

Part a:

Based on Plumer, the solid solubility of Phosphorous at 900 C is about 4*10^20.

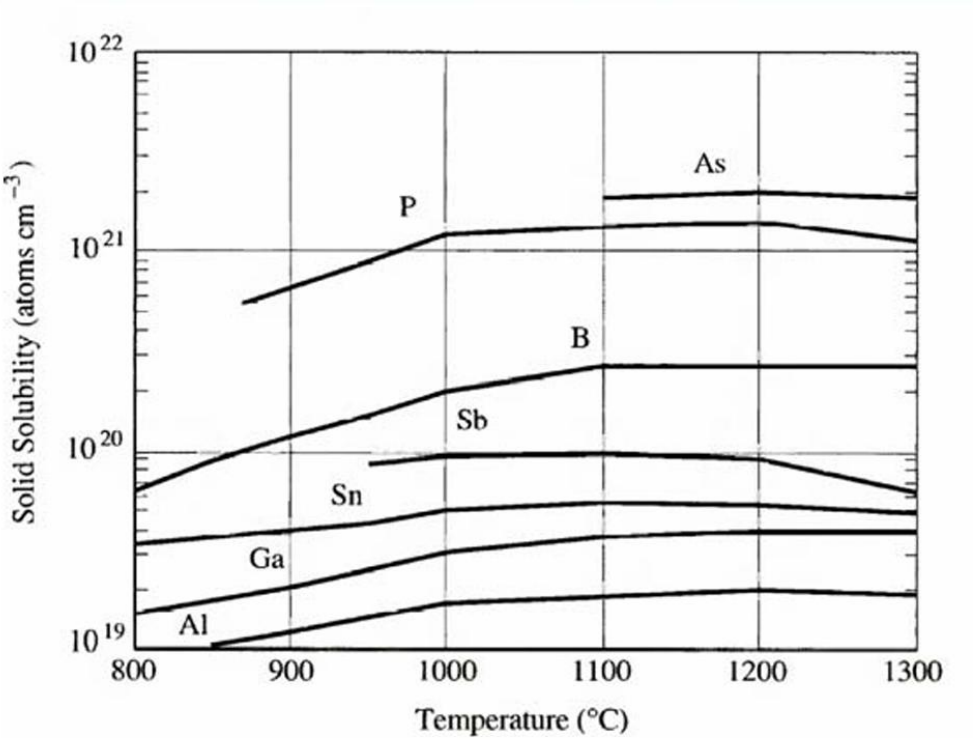
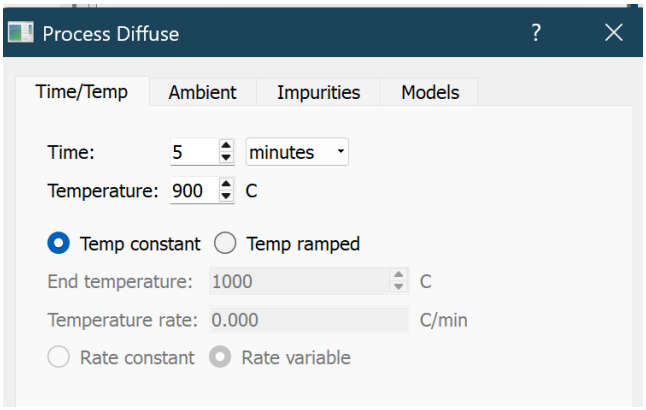
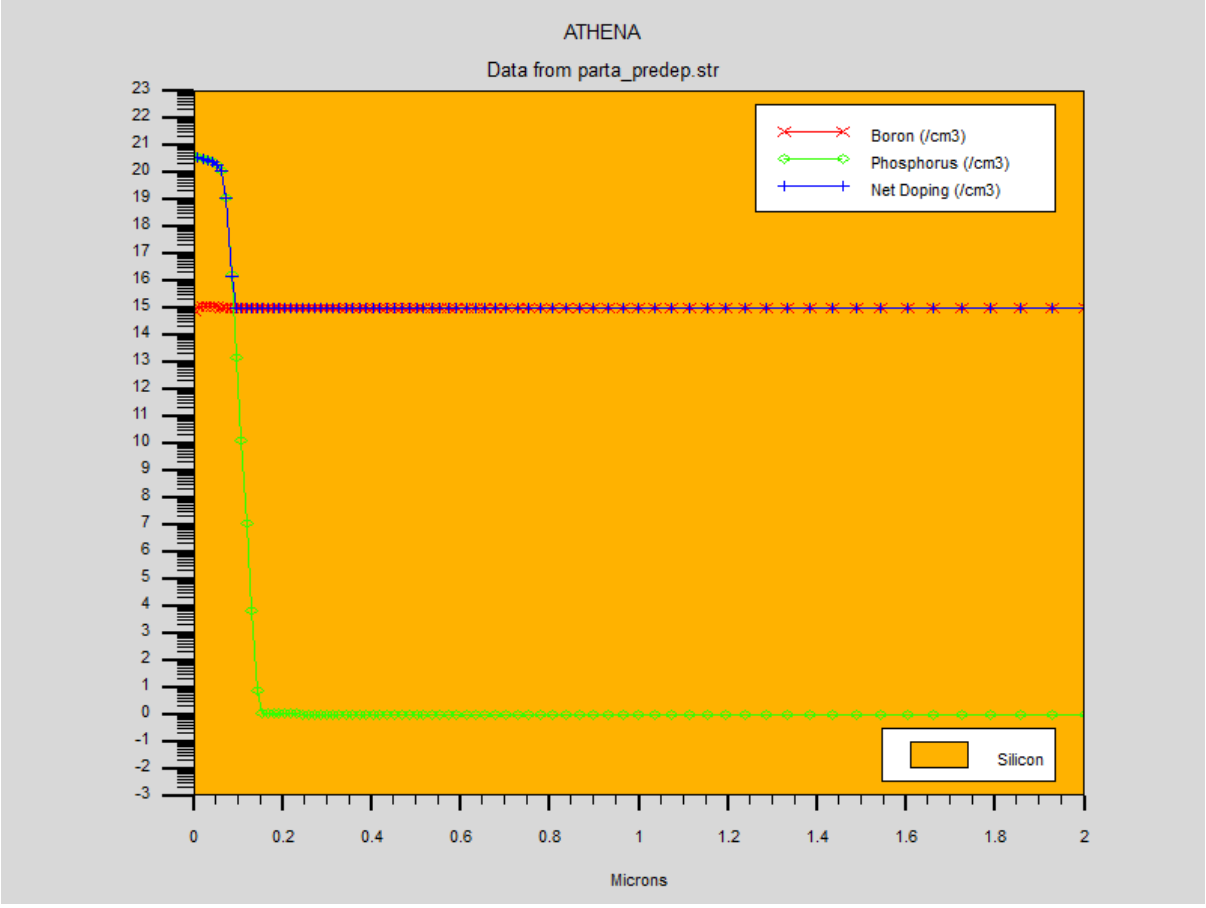


Figure 7–4 Solid solubility curves for various dopants in silicon. These values are the equilibrium solubilities at each temperature and may not be achieved in device doped regions. (After [7.3].)

We perform the pre-dep step at 900 °C with a phosphorus concentration of 4*10^20 for 5 minutes.



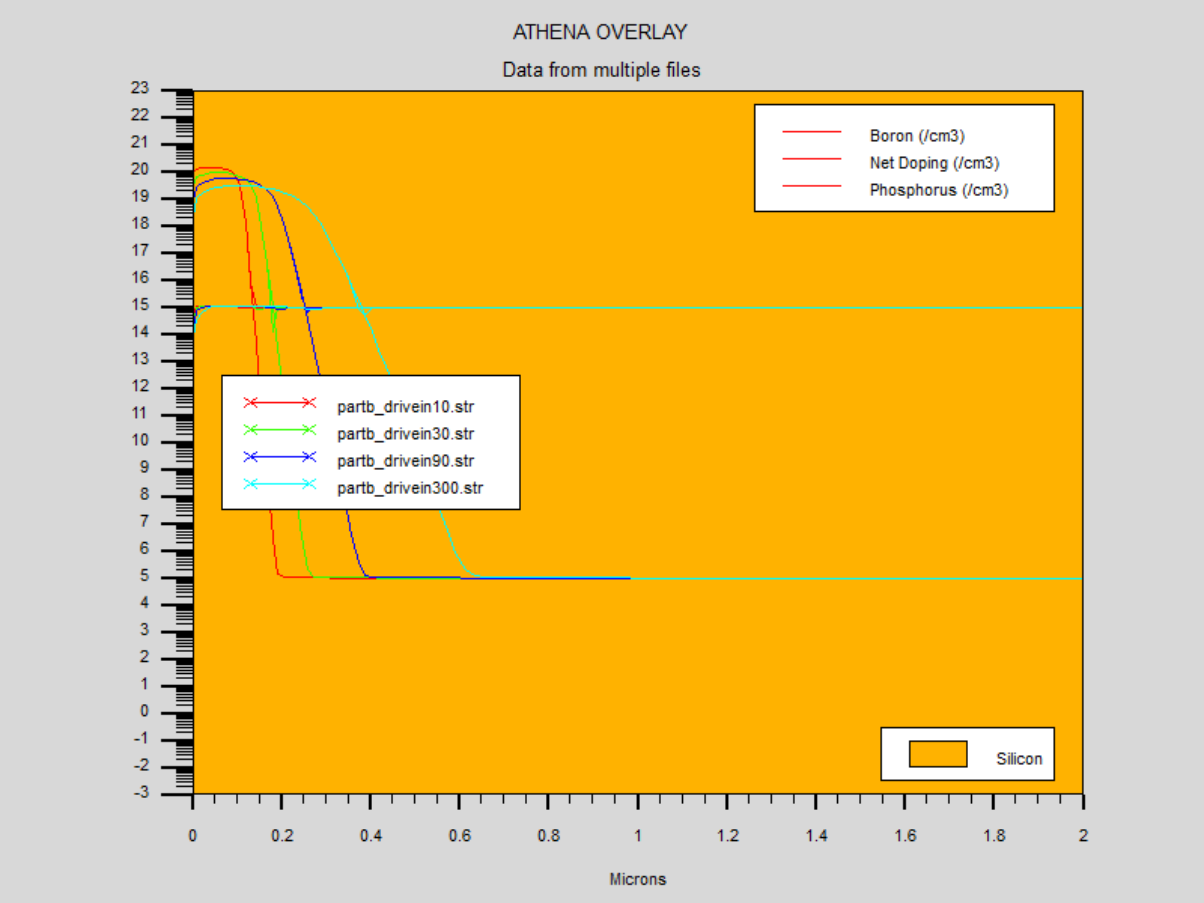
The final doping concentration vs depth is:



We see that the surface concentration remains fixed at the solid solubility, while the dopant distribution widens and the dopant dose increases over time. Junction depth (Xj) is about 0.1 um where the dopant concentration reaches the background concentration.

Part b:

In this part, we perform the drive-in step for 10.30, 90, and 300 minutes. It is important to note that diffusion is cumulative; each time you diffuse, the amount increases.

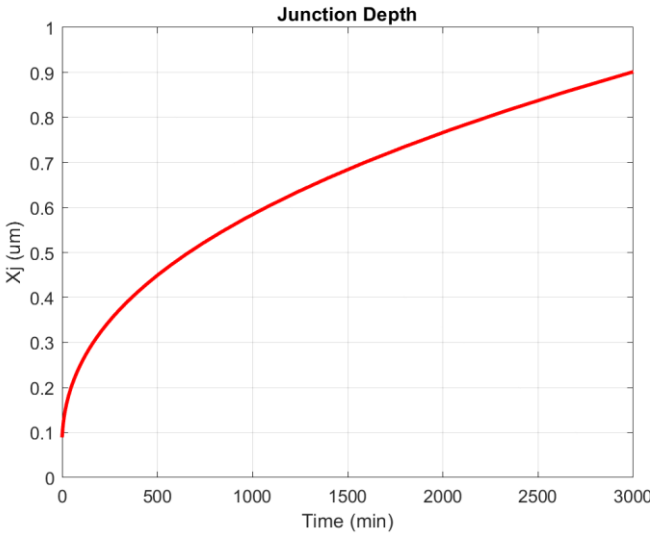


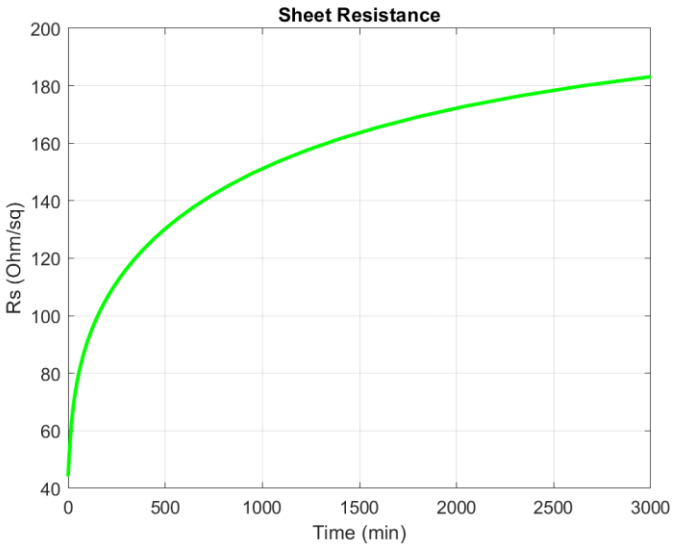
It can be seen that the likely Gaussian distribution widens over time, and its peak also decreases. It means that the junction depth and sheet resistance increase with time, and the surface concentration decreases with time spent in the drive-in process. However, the dose won't change.

Part c:

In this part, we will run a sweep of the drive-in step at a constant temperature to plot junction depth (Xj), sheet resistance (Rs), and surface concentration (Ns) over time. As you can see on the main deck, I used a time loop with 100 steps, ranging from 0.01 to 3000 minutes on a logarithmic scale because the change in speed is faster at the beginning and decreases over time. I saved the 4 columns of data of time, Ns, Rs, and Xj from Silvaco and plotted them in MATLAB for those 100 steps.

For sheet resistance and junction depth, time is plotted linearly, but for surface concentration, it is plotted logarithmically to better show the change over time.

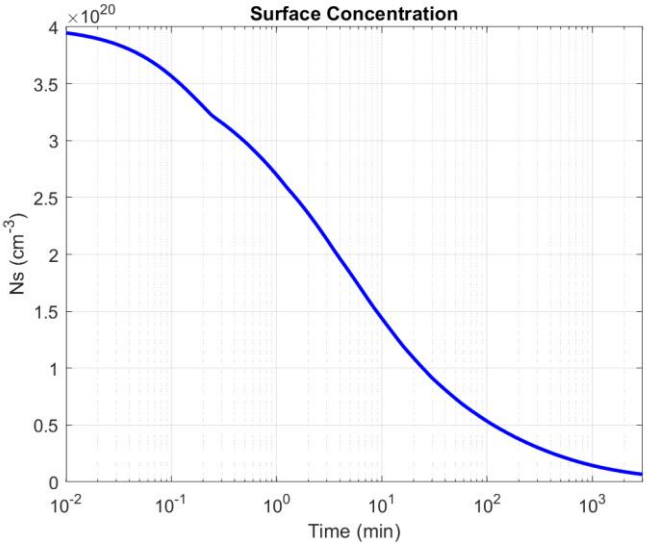




It can be seen that Xj, which is the junction depth, will increase because the dopants are diffusing in silicon, but the growth speed is likely proportional to $\sqrt{t}$. So the speed will decrease. From the chart, we can see that Xj starts at 0.1 μm due to the pre-dep step.

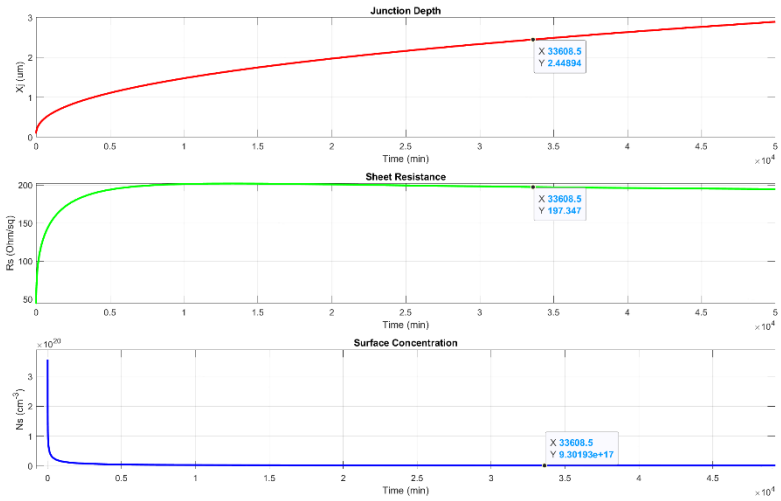
Sheet resistance is also similar: it increases, but the rate of increase decreases over time.

In surface concentration, this decrease occurs faster at the beginning, but the rate of decrease also decreases exponentially.

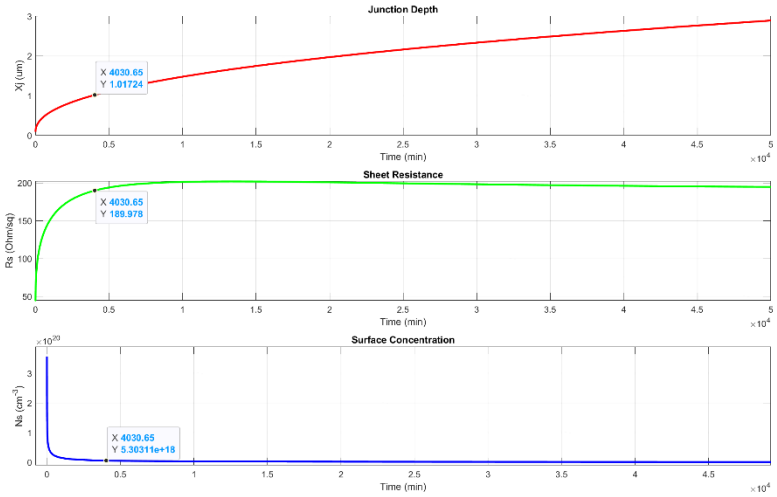


For reaching Xj = 2.5 microns, more time is needed.

It can be seen that, for Xj = 2.5 microns, we need 34000 minutes. At this time, Rs is 200 Ohm/sqr, and Ns is 9.3e17.



Since for 2.5 micron, 34000 minutes is 23 days, and it's impossible, for the next steps (As TA gave permission), I will set my target Xj to 1 micron and continue. For 1 micron, time is 4000 minutes, Rs is 190 ohm/sqr, and Ns is 5.3e18. The process is still impossible, but I have to continue.



If we want to decrease Rs while keeping Xj constant, we can increase the dopant dose in the pre-dep step because Rs is proportional to 1/Q.

If we want to keep Rs constant and increase Xj, as we know, Xj is proportional to $\sqrt{Dt}$. So, to multiply Xj by n, we should take more time: O (n^2). Also, to keep Rs constant, the dopant doses should be multiplied by n, since Rs increases with time.

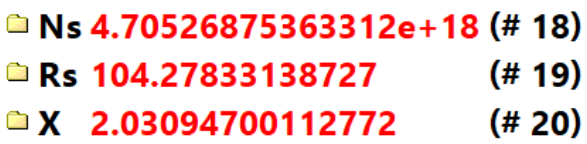
Now we consider n and m for 2, 3, 4, and 10, based on Xj = 2.5 μm and Rs = 200 Ω/square.

Target Xj	Target Rs	Related Ns	Drive-In Time	Pre-Dep Time
n = 2	m = 2	4.95e18	17000	23
n = 2	m = 3	8.57e18	15000	45
n = 2	m = 4	1.42e19	11000	70
n = 2	m = 10	3.34e19	8000	220
n = 3	m = 2	2.76e18	40000	26
n = 3	m = 3	5.05e18	35000	52
n = 3	m = 4	8.30e18	30000	90
n = 3	m = 10	2.22e19	22000	300
n = 4	m = 2	1.56e18	80000	26
n = 4	m = 3	3.00e18	72000	55
n = 4	m = 4	5.22e18	62000	100
n = 4	m = 10	1.66e19	43000	370
n = 10	m = 2	2.31e17	650000	26
n = 10	m = 3	5.09e17	560000	55
n = 10	m = 4	1.12e18	500000	120
n = 10	m = 10	5.54e18	350000	650

To fill the chart, I will use different dopant doses and manually and mathematically vary the drive-in times to determine whether the condition is possible. I won't do multiple pre-deps and increased temp as they are for part e.

For example, if n and m are 2, by multiplying the time of both pre-dep and drive-in, mathematically we can achieve Rs/2 and 2*Xj. Let's see: drive-in: 136000 mins; pre-dep: 20 mins.

It seems that mathematics can be pretty accurate. I will use it for the other rows of the table as well. But I will fill the table instead of explaining.



Mathematics won't work because it is a simplification and does not account for secondary effects.

By analyzing the chart, we can see that most, if not all, are impossible because they require a thermal budget for days to be met. So it can be understood that only considering pre-dep time to increase the dose and drive-in time can't make all the parameters that we want. As Xj increases, the drive-in time increases significantly. And to decrease Rs, the dose should be increased significantly. Surface concentration is mostly correlated with dose, but with a small slope.

Part d:

Considering all 16 conditions, none are easy or normal to achieve. Most of them are impossible, and maybe one or two are possible but hard and expensive.

Instead of copying the table again, I colored the impossible rows red and the expensive rows yellow. There is no possible and ok row. All yellow rows have a drive-in time of more than 8000 minutes (5.5 days), and they are also very expensive given the thermal budget. So I will choose n=2 and m=10, and n=2 and m=4 because they have the fewest drive-in times, although their pre-dep time is also quite high.

Part e:

As noted in the previous section, the condition in which Xj was the least was selected based on the drive-in time. However, based on the calculations, we understood that to reduce Rs, we have to increase the dopant dose, and to increase Xj, we have to extend the drive-in step. A very long-time drive-in is not possible due to wafer damage, high cost, thermal budget constraints, and oxide degradation. And a long-time pre-dep step doesn't work well either because it leads to rapid solid solubility, dopant clustering, and mobility degradation. So by changing times as we did in part d, nothing could have happened. Now, increasing temperature, multiple pre-dep drive-in steps, and ways to reduce negative diffusion, such as a thin layer of silicon nitride, can help us make some of those impossible tasks possible. Using multiple pre-deps increases the rate of dose increase; higher doses are more achievable; peak surface stress is lower; controlling Xj is easier; and it avoids clustering better than a one-time long process. Increasing the temperature will also increase Xj because the diffusion constant D increases exponentially with temperature. 1 hour at 1010 C is equal to 10 hours at 900 C, according to

the formulas. I will use the combination for two of the conditions mentioned in the table of part c.

I'm going to compare methods of increasing Xj separately in this part.

1- No method. 10 mins pre-dep & 200 mins drive-in:

Ns 5.07119094523643e+19 (# 21)
Rs 68.3845909979674 (# 22)
X 0.36815346245926 (# 23)

2- Increasing drive-in time. 10 mins pre-dep & 2000 mins drive-in:

Ns 1.35731848491977e+19 (# 21)
Rs 113.365394155761 (# 22)
X 0.837394539618015 (# 23)

3- Increasing pre-dep time. 100 mins pre-dep & 200 mins drive-in:

Ns 1.18719620301589e+20 (# 21)
Rs 15.577516423291 (# 22)
X 0.587179011916037 (# 23)

4- Using silicon nitride. 10 mins pre-dep & 200 mins drive-in:

Ns 1.1109373536509e+20 (# 24)
Rs 28.6032403931107 (# 25)
X 0.39440628913676 (# 26)

5- Doing multiple times (three times). 10 mins pre-dep & 200 mins drive-in:

Ns 7.51157542765544e+19 (# 30)
Rs 31.768815040035 (# 31)
X 0.543786750996531 (# 32)

6- Increasing temperature to 1000. 10 mins pre-dep & 200 mins drive-in:

Ns 4.58508187577418e+19 (# 22)
Rs 22.3074579758457 (# 23)
X 1.35214247111331 (# 24)

Based on the extracted data using the methods described, it seems that, for increasing Xj, the most effective method by a significant margin is temperature, and after that, it's drive-in time; however, it will also increase Rs. To decrease Rs, the most effective method is to increase pre-dep time, then increase temperature and silicon nitride, and then repeat this multiple times.

So now, for two of the impossible conditions, I will use increasing temperature, silicon nitride, and multiple pre-dep drive-in diffusions to reach what I wanted.

n = 4 and m = 4: Using the setting below, I extracted these parameters, and yes, it is possible.

diffus time=5 minutes temp=900 c.phosphor=4e20

deposit nitride thick=0.2

diffus time=220 temp=1000 nitro

etch nitride all

diffus time=10 minutes temp=950 c.phosphor=4e20

diffus time=322 temp=1100 nitro

Ns 6.4548078968593e+18 (# 27)
Rs 45.4641939117162 (# 28)
X 3.99352304928651 (# 29)

n = 10 and m = 10: Using the setting below, I extracted these parameters, and yes, it is possible.

diffus time=10 minutes temp=950 c.phosphor=4e20

deposit nitride thick=0.2

diffus time=600 temp=1100 nitro

etch nitride all

diffus time=10 minutes temp=950 c.phosphor=4e20

diffus time=600 temp=1150 nitro

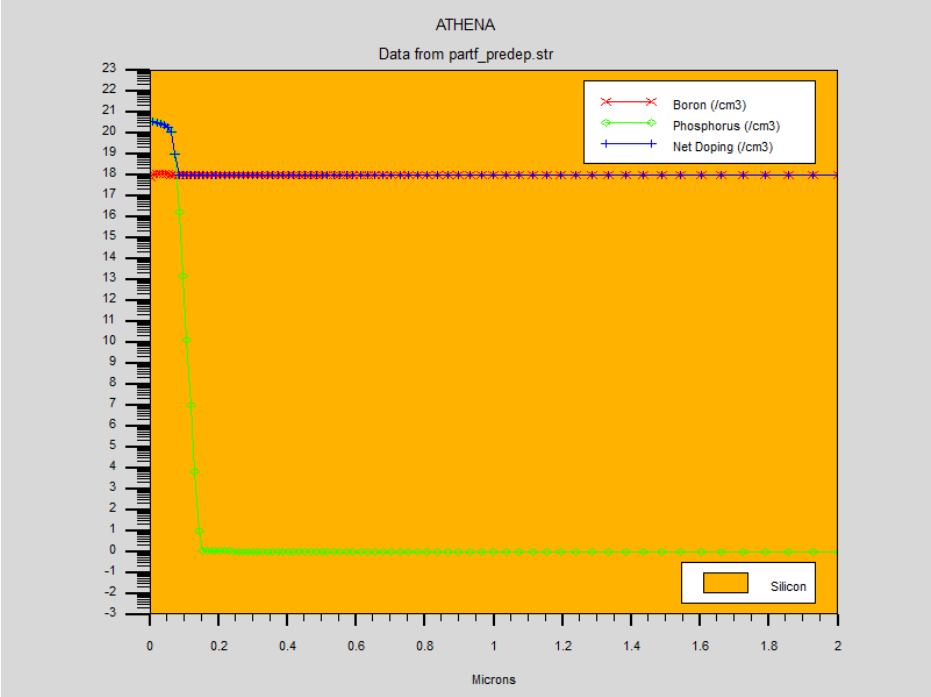
Ns 5.79482624165215e+18 (# 28)
Rs 19.8260716891983 (# 29)
X 10.0789296325284 (# 30)

It seems that other conditions may also apply, using different temperature methods, multiple diffusions, nitride, time, etc.

Using multiple pre-deps can be better than a long-time pre-dep because of the thermal budget and the time spent considering effectiveness. When the surface density reaches the solid solubility, the rate of increase slows, making multiple pre-deps a more logical method.

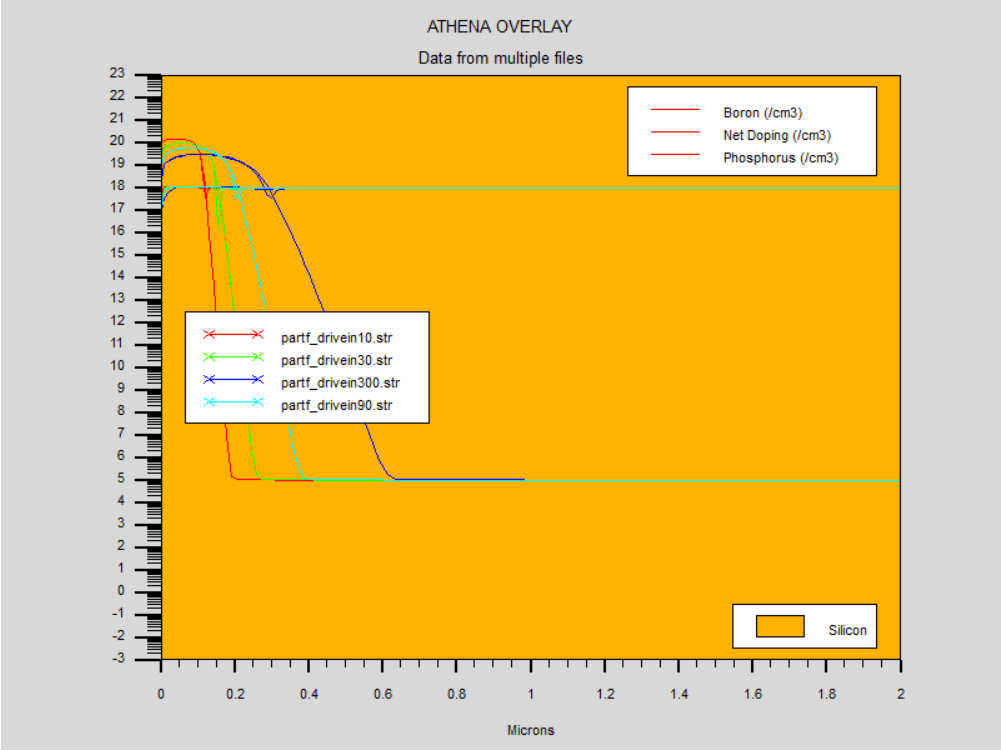
Part f:

In this part, set the doping background to 1e18 and repeat all previous steps.



From the pre-dep step, we see that, due to the increased doping background, Xj is smaller than in the previous case, since Xj is where the background doping equals the diffused dopant concentration. However, there isn't a significant gap.

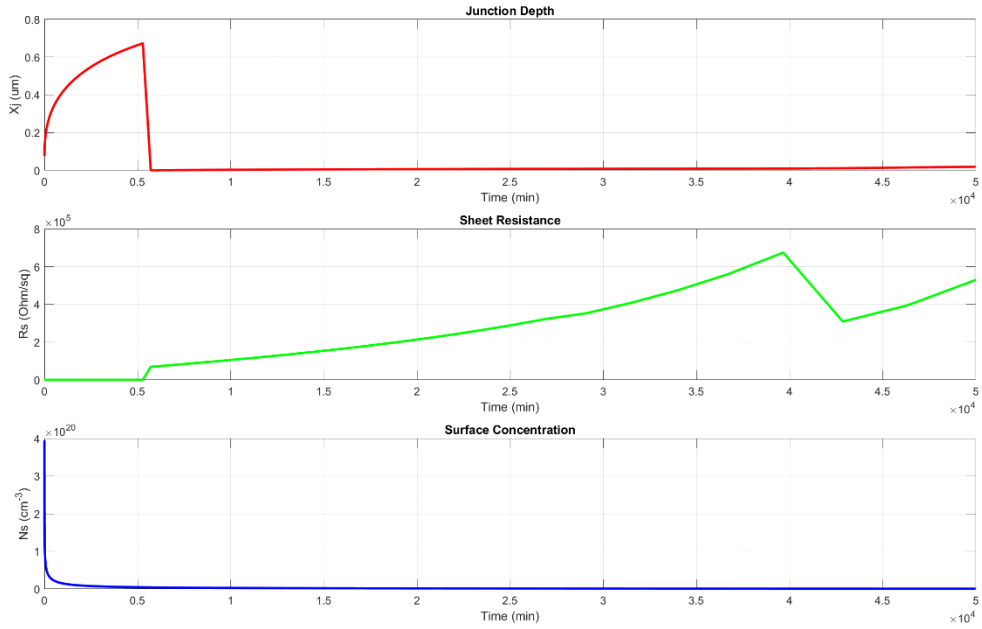
In the next part, I will discuss the drive-in for 10-, 30-, 90-, and 300-minute durations.



As I discussed in part b, this likely Gaussian distribution widens by time as Xj increases. Also, the peak and Ns will decrease as the time spent in the drive-in increases.

The important matter is the comparison of this figure with the figure of part b. Since the doping background is a lot higher and it's like a border for dopant distribution, it is like a cut line in which this line has cut the Gaussian from above, and it means that compared to part b figure, Xjs of the above figure are less at the same time. For example, Xj was 4 um for 300 minutes, but in the new situation it's near 3 um.

Let's see Ns, Rs, and Xj in the new time sweep:



First, let's explain the charts before 5000 minutes, in which the charts are ok and the same as part c. However, at 5000 mins, Xj was 1.1 um; now it's 0.65 um due to the doping background. And also Rs is higher by about 320 instead of 200. The main question is what happens after that time, when Xj approaches zero, and Rs increases rapidly to the order of 10^5. We know that Xj is where the doping concentration is equal to background doping, so when time is spent, and the Gaussian distribution widens, its concentration at the peak or even the parts before, becomes equal to the background, and Xj becomes nonsense. The diffusion becomes over-diffused. For Rs, because the net doping gets negative, it grows exponentially. So we have a drive-in time limit for high-doping backgrounds.

Now with this new doping background, it will be more difficult to reach nXj and Rs/m that part c and d want. This time, the only variable parameters are not pre-dep and drive-in times. Changes in temperature and multiple steps can also be used. I will make that chart again, including new columns of temperature and the second step. First, to reach 1 um, I will increase the pre-dep time to increase the dopant dose and prevent the problem in the previous part. With 15 minutes of pre-dep, 10000 is needed for the drive-in to reach 1 um of Xj. Rs becomes 177, and Ns becomes 5.67e18. Let's use our options to complete the table.

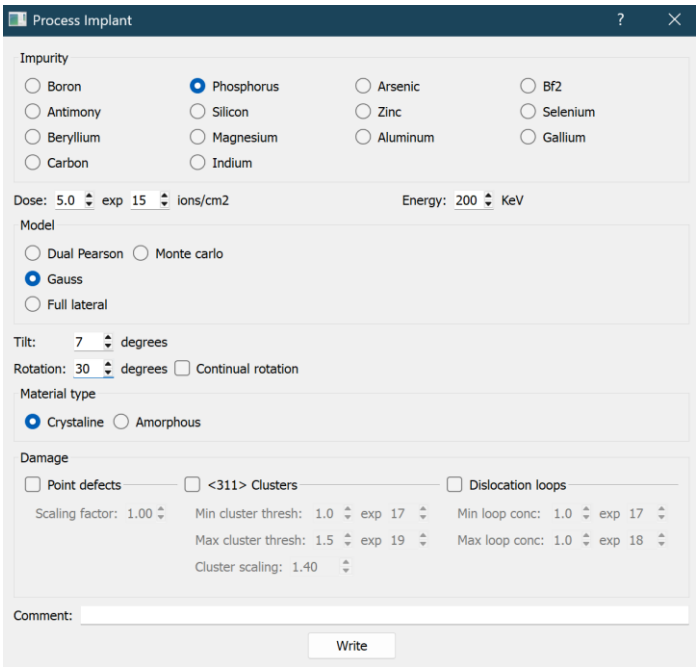
Target Xj	Target Rs	Related Ns	Drive-In Time1	Pre-Dep Time1	Temp1 (P-D)	Drive-In Time2	Pre-Dep Time2	Temp2 (P-D)
n = 2	m = 2	5.86e18	250	13	950-1100	-	-	-
n = 2	m = 3	9.25e18	200	19	950-1100	-	-	-
n = 2	m = 4	1.16e19	190	27	950-1100	-	-	-
n = 2	m = 10	2.74e19	110	19	1000-1100	-	-	-
n = 3	m = 2	3.71e18	750	8	1000-1100	-	-	-
n = 3	m = 3	5.57e18	580	10.5	1000-1100	-	-	-
n = 3	m = 4	7.26e18	500	13	1000-1100	-	-	-
n = 3	m = 10	2.09e19	280	35	1000-1100	-	-	-
n = 4	m = 2	2.49e18	600	11	1000-1150	-	-	-
n = 4	m = 3	4.26e18	420	15	1000-1150	-	-	-
n = 4	m = 4	5.07e18	380	17	1000-1150	-	-	-
n = 4	m = 10	1.50e19	220	45	1000-1150			
n = 10	m = 2	1.03e18	1100	2	1100-1240	1000	1.5	1105-1210
n = 10	m = 3	1.43218	1100	2	1100-1240	1000	2.5	1100-1200
n = 10	m = 4	2.04e18	900	3	1100-1200	920	3	1100-1200
n = 10	m = 10	6.25e18	530	7	1100-1200	530	7	1100-1200

For increasing Xj, my utility was temperature and drive-in time, although pre-dep time was also important. For decreasing Rs, increasing pre-dep time, and decreasing drive-in time would help. It seems that using temperature and multiple steps, while handling the timing and Xj and Rs we want, is not that hard, and those previously impossible processes are now possible and acceptable.

Conditions in the green rows are possible with the parameters that I identified. The yellow one is possible, but it's hard and requires a lot of time and thermal budget. Red ones are really hard to achieve and almost impossible because achieving High Xj and High Rs simultaneously is difficult. For more Xj, temperature may not be sufficient, and we are limited by wafer damage. The related Ns in all conditions are not that much different. For impossible ones, maybe adding more steps can be helpful for reducing time and thermal budget.

Part g:

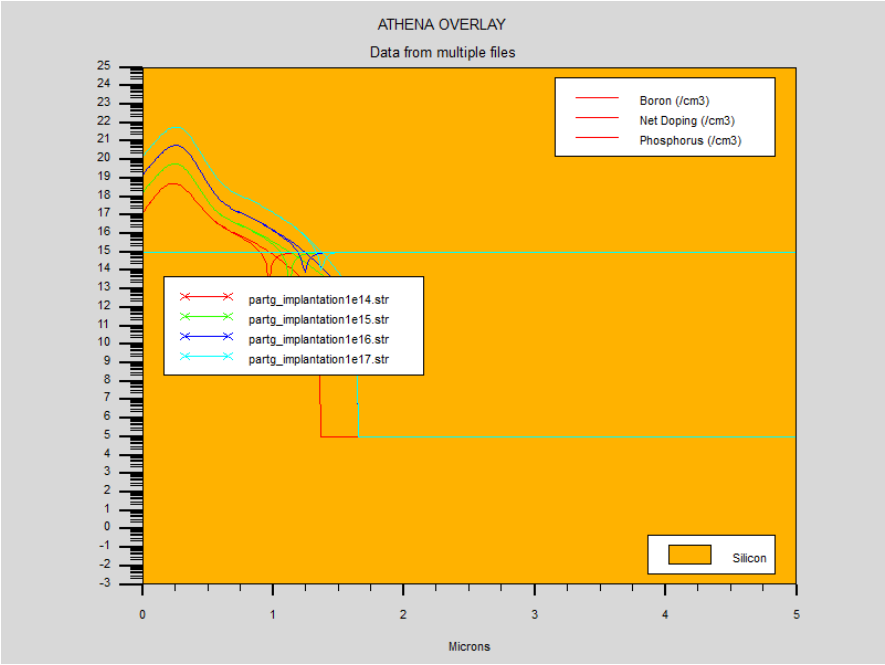
For this part, I will use 1e15 Boron as the doping background. During implantation, the dose, energy, tilt, and rotation should be identified. Energy defines the depth of implanted dopants. Dose is accurate and counted specially using nitride as the shield. This is the main difference between diffusion and implantation: implantation is closed-loop, whereas diffusion is open-loop in the case of dopant doses. Tilt is the angle between the ion beam and the wafer surface normal (vertical axis). It reduces channeling. In crystalline silicon, ions can travel deeper along crystal directions (like ⟨100⟩). This causes deeper-than-expected junction depth (Xj) and broader dopant profile. A small tilt (typically 7°) breaks the alignment with the crystal channels and yields a more Gaussian, predictable profile. Rotation is the azimuthal angle of the wafer around its surface normal during implantation. It rotates the wafer in-plane and is usually combined with tilt. It is used to further randomize ion entry direction and suppress channeling along specific crystal axes.



Balance between Xj and Rs:

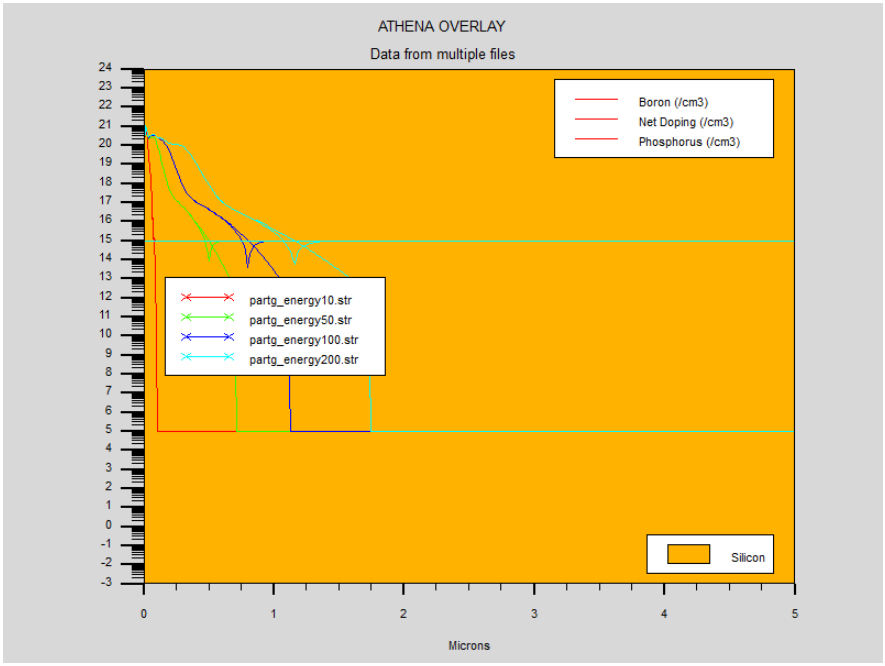
Depositing nitride will keep the dopants in the wafer, and after annealing, the nitride can be etched away. In implantation energy controls the initial depth before annealing. More energy means implanting deeper, so energy controls Xj. Dose controls Rs because it is proportional to 1/Q. Higher dose means lower Rs, higher surface concentration, and a greater risk of clustering, which can be reduced by a thin layer of nitride. Another effective parameter on Xj is annealing time. More time spent means deeper junction, more spread dopants, and also an increase in Rs. Usually, increasing Xj over time increases Rs too. For example, if we want deeper Xj and lower Rs, using the highest allowed energy (200 keV), a moderately high dose (5e15), and a short, high-temperature anneal (1000–1050°C) can achieve this. It seems that in diffusion and annealing, dose and depth are coupled, but in implantation, they are not.

Dose comparison:



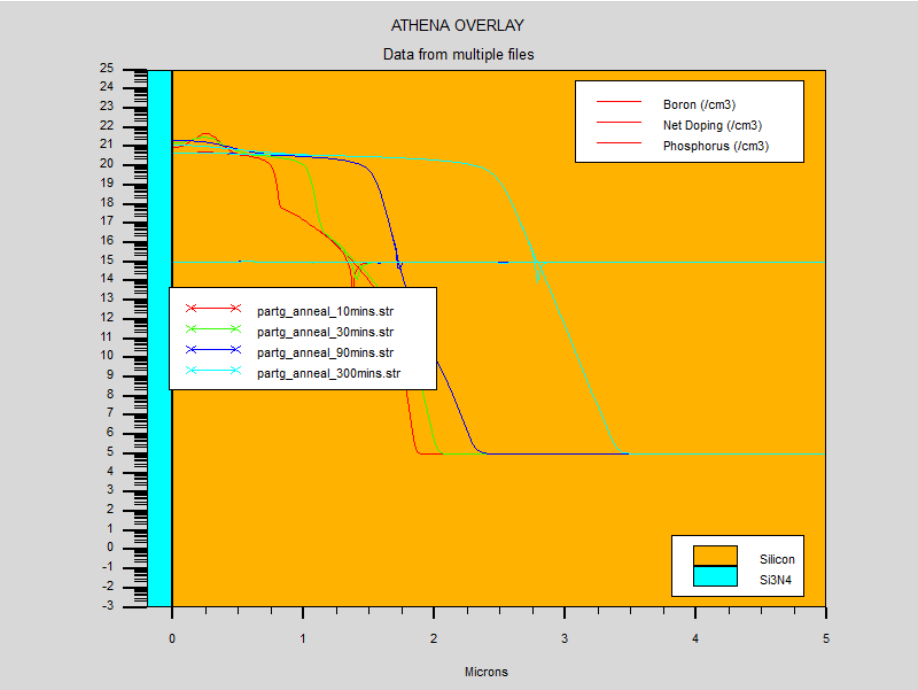
It can be seen that increasing the dose increases the peak of the implantation distribution and Xj. It is logically coherent with the increase of Ns and other concentrations.

Energy comparison:



As we can see, energy increases and identifies the depth. The distribution widens based on the amount of dopant energy.

Anneal time comparison:



After annealing, the distribution becomes approximately a step function, and as the annealing time increases, the step widens and Xj increases. So Xj is based on all of them, dose, energy, and annealing time.

Rs comparison:

Rs_1e14	376.063572954876	(# 16)
Rs_1e15	59.6512213833352	(# 20)
Rs_1e16	6.77562634654984	(# 24)
Rs_1e17	0.819622427417344	(# 28)
Rs_t10	2.39152275051389	(# 38)
Rs_t30	1.86800784822787	(# 42)
Rs_t90	1.36288984361718	(# 46)
Rs_t300	0.914866443933047	(# 50)
Rs_E10	37.9099406751268	(# 68)
Rs_E50	18.8230202770995	(# 72)
Rs_E100	12.4299824773816	(# 76)
Rs_E200	9.19132992157738	(# 80)

Rs decreases significantly with increasing dose. It also reduces with annealing time, but not to a considerable extent. Increasing energy reduces Rs, too. At last, the most effective parameter is the dopant dose.

Secondary effects:

In Silvaco simulations of ion implantation and subsequent annealing, several secondary effects can be observed beyond the ideal Gaussian dopant profile predicted by simple analytical models. One common effect is channeling, which appears as a deeper "tail" in the dopant distribution when tilt and rotation angles are not properly selected. Implantation damage is also visible, particularly at higher energies and doses, leading to enhanced diffusion during annealing due to transient enhanced diffusion (TED). Dopant clustering and partial activation may occur at high concentrations, reducing the electrically active dopant fraction and increasing the extracted sheet resistance (Rs) compared to the implanted dose. During drive-in or anneal steps, segregation of dopants at material interfaces can modify the near-surface concentration profile. Additionally, concentration-dependent diffusivity causes non-ideal broadening of the profile, deviating from simple $\sqrt{Dt}$ behavior. These effects demonstrate that TCAD results reflect complex physical mechanisms rather than idealized analytical solutions, and careful model selection is necessary to obtain realistic junction depth (Xj) and Rs values.

From my simulations, these effects can be seen:

1- In the energy sweep (10–200 keV), higher energies show a deeper and longer dopant tail instead of a perfectly symmetric Gaussian. This is caused by channeling.

2- In the annealing comparison (10–300 minutes), diffusion is not strictly proportional to $\sqrt{t}$. The initial anneal steps show stronger-than-expected spreading. This is due to implantation damage, which creates excess silicon interstitials. During early annealing, these defects enhance dopant diffusion (TED), causing additional junction movement beyond classical diffusion predictions.

3- In the dose variation (1e14–1e17 cm^{-2}), higher doses do not scale linearly in electrically active concentration. At very high doses: Dopants cluster, some dopants become electrically inactive, and solid solubility limits are approached. This causes the net active doping to saturate and affects sheet resistance (Rs). The peak concentration increases, but not in proportion to the dose.

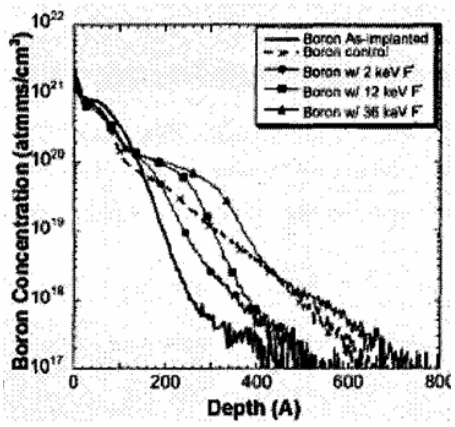
4- In your anneal curves, the high-concentration region spreads differently from the low-concentration tail. This indicates concentration-dependent diffusion, where diffusivity increases at high dopant concentrations due to pair diffusion mechanisms. This deviates from constant-D analytical models.

Part h (Bonus):

-Verification of Results:

B. Vanderpool - *Reduction of shallow boron and phosphorus diffusion by co-implantation*

This paper shows how diffusion profiles vary with dose and anneal conditions.



For Example, this figure of the names paper is similar to the energy vs. depth figure for the implantation in part f. It should be noted that the variation may be due to the above figure being for Boron and to the simplification in Silvaco.

H. Mehran et al. - *Phosphorus Diffusion in Silicon: Impact of Doping and Oxidation*

In this paper, there are some figures showing depth vs time and depth vs dose.

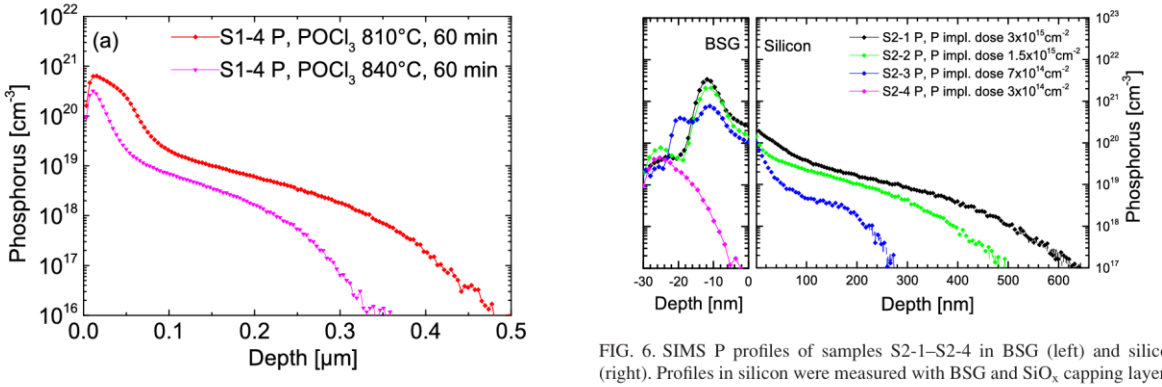


FIG. 6. SIMS P profiles of samples S2-1–S2-4 in BSG (left) and silicon (right). Profiles in silicon were measured with BSG and SiO_x capping layers.

These figures are very similar to the time- and dose-diffusion distributions I simulated in the previous parts.

H. Wagner et al. - *IMPROVING THE PREDICTIVE POWER OF MODELING THE EMITTER DIFFUSION BY FULLY INCLUDING THE PHOSPHSILICATE GLASS (PSG) LAYER*

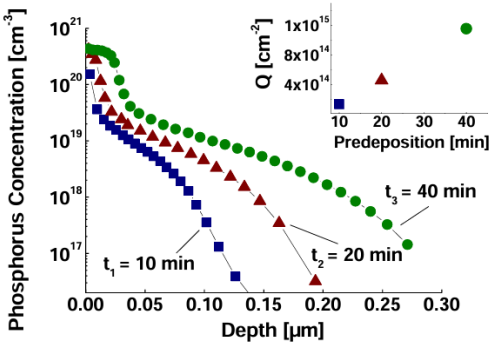


Figure 1 | Phosphorus diffusion profiles in silicon, measured with ECV, obtained with various predeposition times t, while the temperature T = 840 °C and the flows of $POCl_3$ and O_2 were kept constant. The inset shows the dose of phosphorus (Q) that entered the silicon, calculated by integration over the ECV profile.

This figure from the named paper shows diffusion distribution vs time, which is similar to our results.

The deviation arises because TCAD tools solve the semiconductor equations numerically using finite-element or finite-volume discretization on a mesh. Diffusion is modeled using concentration- and temperature-dependent diffusivity, clustering effects, segregation at interfaces, and sometimes transient enhanced diffusion (TED). Analytical solutions, in contrast, assume simplified boundary conditions (e.g., constant surface concentration or limited-source Gaussian diffusion), constant diffusivity, and one-dimensional ideal behavior. Real experimental data also include process variations such as furnace non-uniformity, oxide growth interactions, implantation damage, and measurement uncertainty. Therefore, differences between TCAD, analytical solutions, and fabricated devices are expected.

In addition, numerical factors influence the results. Mesh resolution affects gradient accuracy near the junction. Time-step control during annealing impacts dopant redistribution. Mobility and bandgap narrowing models alter extracted Rs. Solver tolerances and convergence criteria also contribute to small numerical discrepancies. Thus, TCAD results represent a physics-based numerical approximation rather than an exact analytical solution.

I have already completed the sensitivity analysis and discussed the temperature- and time-dependence of Xj and Rs, both in pre-dep and drive-in, in part f. I have shown in the table how effective their changes could be and how sensitive they are to temperature and time. Because diffusion is exponentially dependent on temperature, A small increase in temperature significantly increases junction depth (Xj) and Rs decreases as higher diffusion reduces surface concentration gradients and increases electrically active dopant depth. Temperature is the most sensitive parameter in diffusion control. Sheet resistance depends strongly on both surface concentration and dopant activation. Careful adjustment of thermal budget allows meeting the assigned junction depth while maintaining acceptable Rs, demonstrating consistency between TCAD modeling and semiconductor diffusion theory.

-Importance of Meshing:

Meshing is a critical aspect of TCAD simulations because all physical equations (diffusion, Poisson, and continuity equations) are solved numerically at discrete mesh nodes. The accuracy of extracted parameters such as junction depth (Xj) and sheet resistance (Rs) strongly depends on proper mesh resolution, particularly in regions with steep doping gradients. Fine mesh spacing is essential near the silicon surface, around the projected range of implanted dopants, and at the metallurgical junction where concentration changes rapidly. Insufficient resolution in these regions can lead to incorrect Xj extraction, inaccurate Rs values, and underestimation of peak electric fields. Conversely, using an excessively fine mesh throughout the entire structure significantly increases computation time and may cause convergence issues without improving accuracy. An effective strategy is to apply non-uniform meshing: dense near the surface and junction, and progressively coarser in the bulk where gradients are small. Adaptive refinement based on impurity gradients further improves numerical reliability. A mesh convergence study, by refining the mesh and verifying that key results change by less than a few percent, ensures that the simulation outcomes are physically meaningful and not artifacts of discretization. For the simulations in this project, I refined the mesh, and the results were almost identical.

-Solver Models and Methods:

Solver models and numerical methods in Silvaco ATLAS significantly influence both the physical accuracy and numerical stability of device simulations. Different physical models, such as concentration-dependent mobility (e.g., CVT or Lombardi), Shockley-Read-Hall (SRH) recombination, Auger recombination, bandgap narrowing (BGN), and Fermi–Dirac statistics, affect the predicted carrier distribution, sheet resistance (Rs), and junction behavior. For heavily doped regions, enabling bandgap narrowing and Fermi statistics is essential; otherwise, carrier concentration and built-in potential may be underestimated. Similarly, advanced mobility models reduce mobility at high doping levels, thereby increasing the extracted Rs compared with constant-mobility assumptions. Recombination models primarily affect leakage currents and transient behavior, but can also influence convergence under high-injection conditions.

Regarding numerical methods, the choice between Gummel iteration and Newton's method affects convergence and simulation time. Gummel iteration is more stable for initial biasing and strongly nonlinear problems, but converges more slowly. Newton's method is faster and more accurate near the solution but may fail without a good initial guess. Using block or fully coupled solvers improves robustness when solving strongly coupled electrostatic and transport equations. Mesh density and scaling also interact with solver performance; overly coarse meshes reduce accuracy, while excessively fine meshes can lead to convergence difficulties.

Overall, advanced physical models increase realism but also computational cost and sensitivity to mesh quality. A balanced approach, using appropriate high-doping models and starting with Gummel before switching to Newton, provides accurate and stable results. Comparing extracted parameters such as Xj, Rs, and electric field profiles under different model selections helps ensure that conclusions are physically consistent rather than solver-dependent.

By the command setting and methods and models part, there are many options to change and see how the process changes.

-Customized Solver Scenarios:

Standard default settings in Silvaco ATHENA and ATLAS are suitable for basic simulations, but several practical device scenarios require customized physical models and numerical methods to obtain reliable results.

First, in heavily doped regions (e.g., $\geq 1\times10^{18}$ cm^{-3}), standard Boltzmann statistics and constant mobility models become inaccurate. Under these conditions, Fermi-Dirac statistics, bandgap narrowing (BGN), and concentration-dependent mobility models must be enabled. Without these corrections, carrier concentration, built-in potential, and sheet resistance (Rs) are underestimated. Numerically, such high doping gradients can cause convergence instability, so using Gummel iteration initially and then switching to Newton's method improves robustness.

Second, in shallow junction or ion-implanted structures with steep concentration gradients, default meshing and diffusion settings are often insufficient. Implant damage, transient

enhanced diffusion (TED), and clustering effects require advanced diffusion models (e.g., pair diffusion or fully coupled models in ATHENA). A coarse mesh near the surface leads to incorrect junction depth (Xj) extraction. Adaptive mesh refinement and smaller time steps during annealing are necessary to properly capture rapid dopant redistribution.

Third, in high-field or reverse-biased device simulations (e.g., breakdown or strong depletion conditions), standard drift-diffusion with basic mobility models may fail to predict peak electric fields accurately. High-field mobility saturation, impact ionization, and field-dependent mobility models must be activated. Additionally, fully coupled Newton solvers with tighter convergence tolerances are often required because the equations become strongly nonlinear. Without these adjustments, the simulation may either fail to converge or significantly underestimate breakdown voltage and electric field peaks.

Overall, customized solver configurations are essential when dealing with heavy doping, steep gradients, or strong electric fields. Proper selection of physical models, mesh refinement, time stepping, and numerical methods ensures that simulation results reflect physical behavior rather than numerical artifacts.

Sources:

1. Diffusion PowerPoint, 2. ChatGPT, 3. Gemini, 4. Plumer Book, 5. ATHENA User Manual, 6. ATHENA Tutorial PowerPoint, 7. TA Session, Silvaco.com