COMSOL Multiphysics: simulation of a PN junction reverse biased photodiode

Giacomo Langfelder and Luca Pileri

29/11/2024

In this CAD class, you will learn how to design and simulate a silicon device. In particular, a photodiode will be used as a case study. The simplest approach is based on the design of a reverse-biased PN junction. As you will see during next classes, this is the basic structure to form the simplest photodiode.
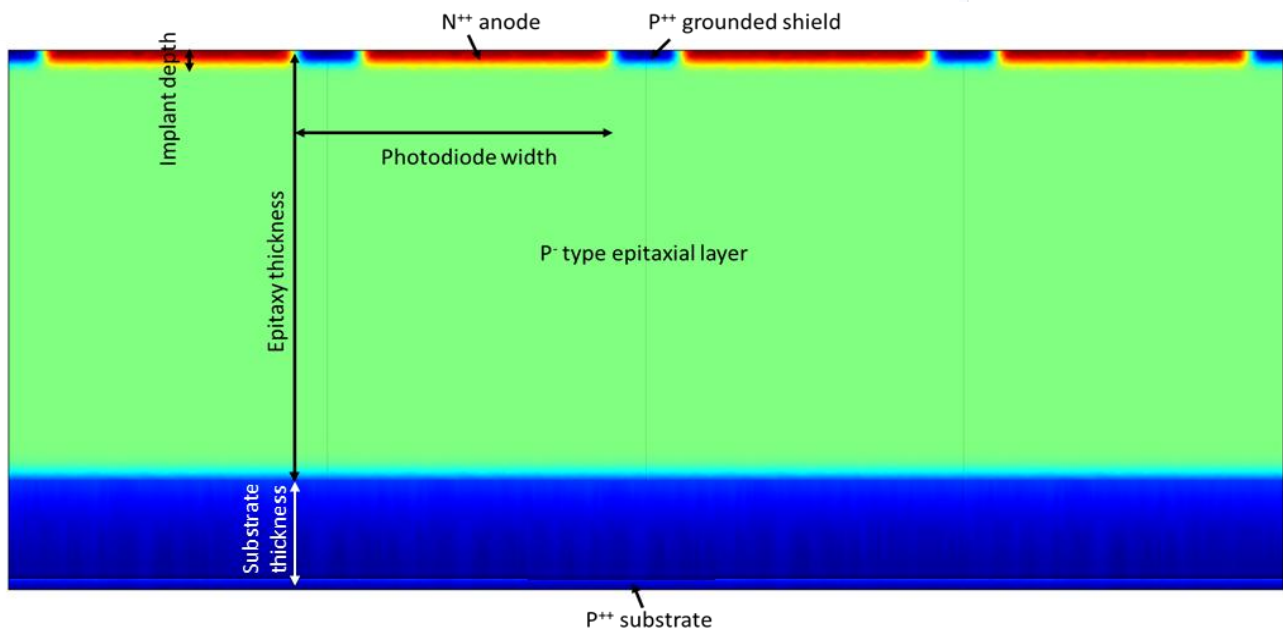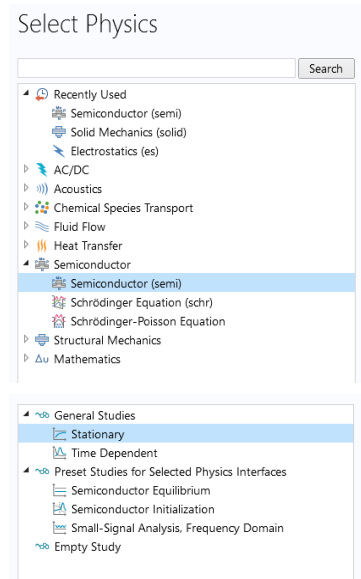
Begin by opening Comsol multiphysics.

Introduction

Choose a 2D model, then choose *Semiconductor* as the *Physics* to solve. Once this physics is added, choose the simulation study: select *Stationary*. We will indeed analyze the photodiode reverse biasing through quasi stationary steps.
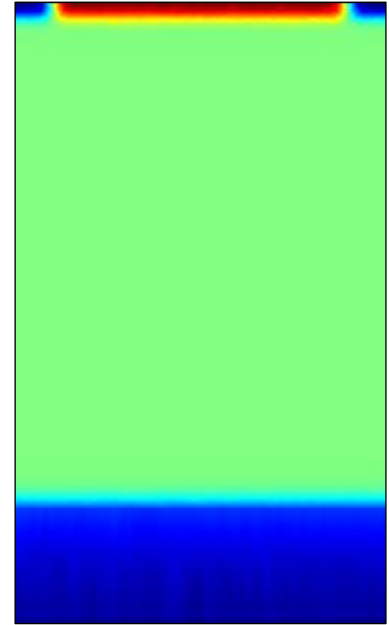
Click on *Done*, and the main Comsol interface will open.

As usual, we start by defining a set of parameters, to simulate a photodiode characterized by the structure shown below: the array is formed by heavily doped N-type anodes, separated by P-type implants used both to bias the epitaxial layer and to provide crosstalk shield between adjacent anodes. The epitaxy is grown on top of a heavily doped substrate which also prevents electrons diffusion towards the substrate itself.

In order to simulate this structure, we rely on a symmetric 2D simulation in such a way that the left and right boundary conditions can be conditions of continuity (i.e. Neumann conditions, where we set that the derivative of our variables has a fixed value, in this case null. Indeed, a null derivative means a continuity of the solution at the boundaries).

Name	Value	Description
lpix	3.5 [um]	Half pixel width
tepi	10 [um]	Epitaxial thickness
tsub	5 [um]	Substrate depth
timp	100 [nm]	Implant depth
lnplus	5 [um]	N+ width
lpplus	0.5 [um]	P+ width
impspace	0.5 [um]	Space between implates
wcont	0.2 [um]	Contact width
Vrev	3 [V]	Reverse voltage

The set of parameters that we define to assist our design is also shown in the Table above, with an obvious meaning for almost all of them. Just a note on *wcont*, which represents the width of the Ohmic contact: indeed, in the simulation we will not design the metal contact itself, but just set a Dirichlet boundary condition in those regions of the semiconductor where the electrostatic potential (one of the variables of our set of differential equations) is fixed. Note also that *lpix* here means half of the actual pixel width.

1.   Structure design

The simulation starts by designing a rectangle corresponding to the structure above. Doping profiles will be added later, so that we do not have to care about them now. The rectangle has a width corresponding to 2*lpix (7 μm) and a thickness (15 μm) corresponding to the sum of the epitaxy and the simulated portion of the substrate (no need to simulate the entire substrate, indeed, as it is a region with no active role in our device).

Instead, we have to define the contacts regions. These are *lines* corresponding to the silicon boundaries, where we expect that ohmic contacts will be placed. We thus add lines at the y-axis value corresponding to the silicon surface, and with a width corresponding to the typical minimum width of a metal layer in a CMOS process (in this case 200 nm: note that a larger width has only a negative effect on the so called Fill Factor, see next classes).

The same step needs to be repeated three times for the central anode contact on top of the N-type region (first figure below) and for the two lateral P+ contacts (for the second one, you can use the *Duplicate* command).

The structure design is already concluded. We can go to the *Materials* and *Boundary* Sections to complete our simulation settings. Note: for advanced simulations, where the behavior of the electric field at the boundaries is of high concern, you can try yourself designing the *real*istic metal contacts (e.g. Al or Au metal rectangles with a

thickness of about 500 nm on top of the silicon layer) and the oxide (SiO₂) overlapping the entire silicon structure (e.g. a 2-um oxide layer).

2.   Materials and Boundary Settings

As usual, right click on *Materials* and *Open Material Library*. Choose *Si – silicon* from the library (you find it in the *Semiconductors* part of the material tree) and add it to the model. The, close the material library.
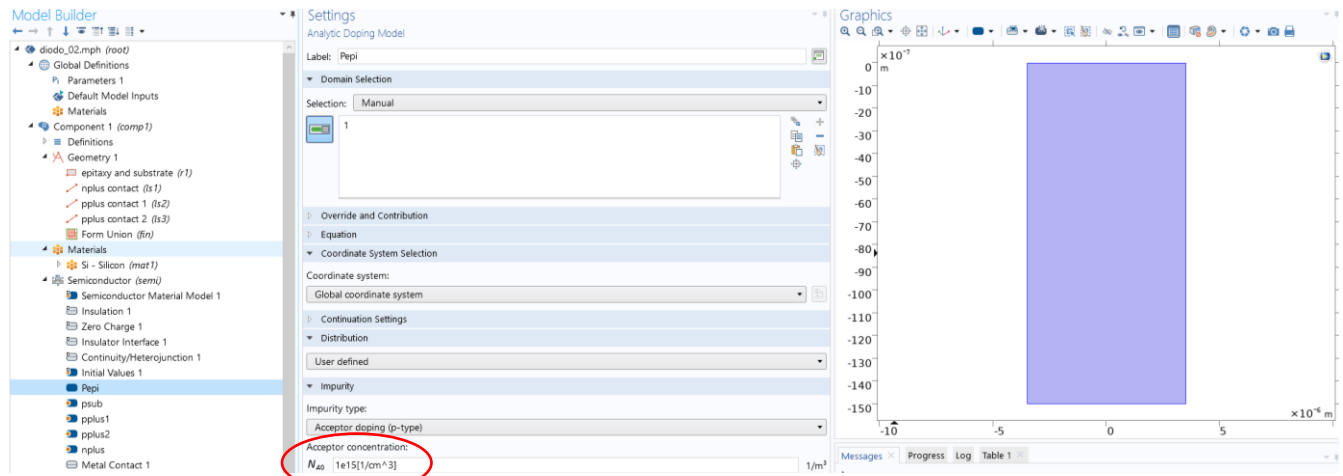
For what concerns boundary conditions, it is better first to have a look at the set of partial differential equations that the solver will compute in each single point of the mesh. As you know, solving a semiconductor problem includes effects of drift and diffusion, involving the solution of 3 equations (*Semiconductor material model* tab):

$$\rho + = q(p - n + N_d^+ - N_a^-)$$
$$\nabla \cdot \mathbf{J}_n = 0, \quad \nabla \cdot \mathbf{J}_p = 0$$
$$\nabla \cdot (-\epsilon_0 \epsilon_r \nabla V) = \rho \qquad \mathbf{J}_n = qn\mu_n \nabla E_c + \mu_n k_B T \nabla n + qnD_{n,th} \nabla \ln(T)$$
$$\mathbf{J}_p = qp\mu_p \nabla E_v - \mu_p k_B T \nabla p - qpD_{p,th} \nabla \ln(T)$$
$$E_c = -(V + \chi_0), \quad E_v = -(V + \chi_0 + E_{g,0})$$

the *Poisson equation*, the *Electrons continuity equation*, and the *Holes continuity equation*. This means that the three variables that are solved for in each mesh point are the electrostatic potential (*V* in the equation), the electron concentration *n*, and the hole concentration *p*. These are the same variables that should be set as boundary conditions (their value, Dirichlet, or their derivative, Neumann). For this reason, defining boundary conditions means defining doping and voltage values.

Before doing that, just set the device thickness (in the non-simulated dimension) to 7 μm (equal to lpix*2).
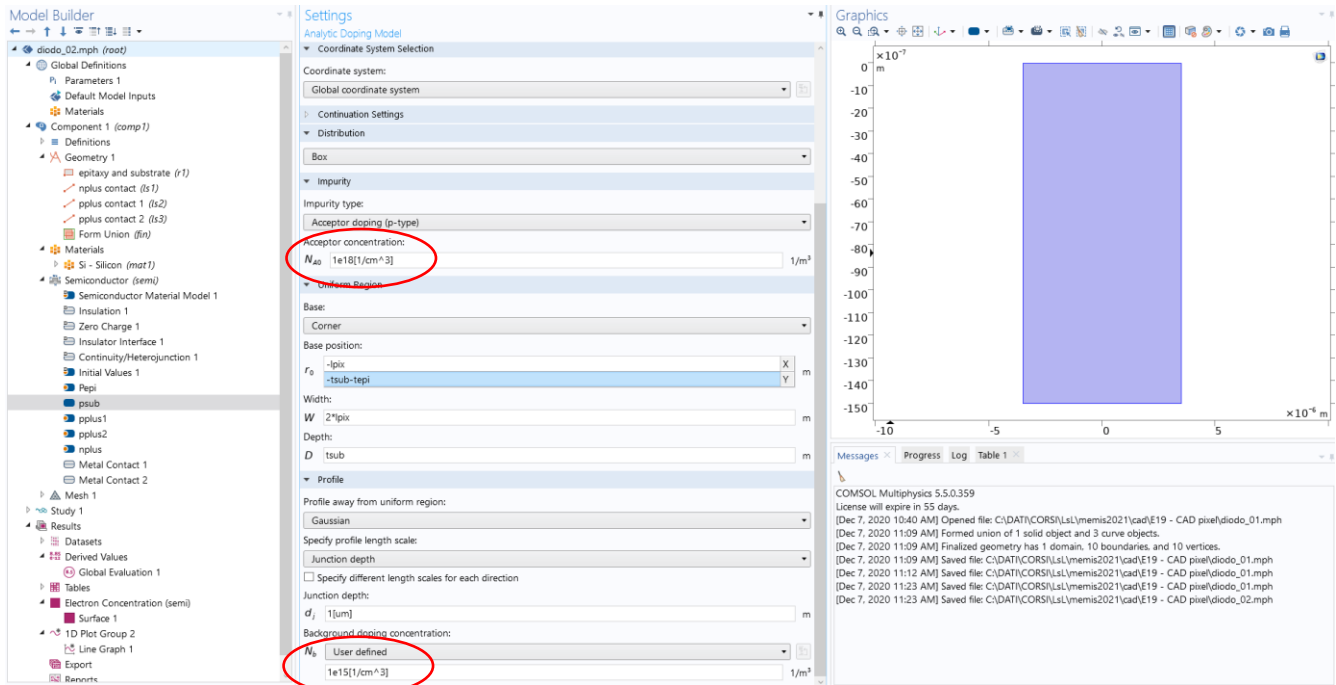
To define our doping regions, we start from the epitaxial layer doping: right click on *Semiconductor* → *Doping* → *Analytic doping profile*. As the epitaxial layer is the lowest doping, it will be overwritten by all other higher doping levels, so we can define it on our entire structure as below:
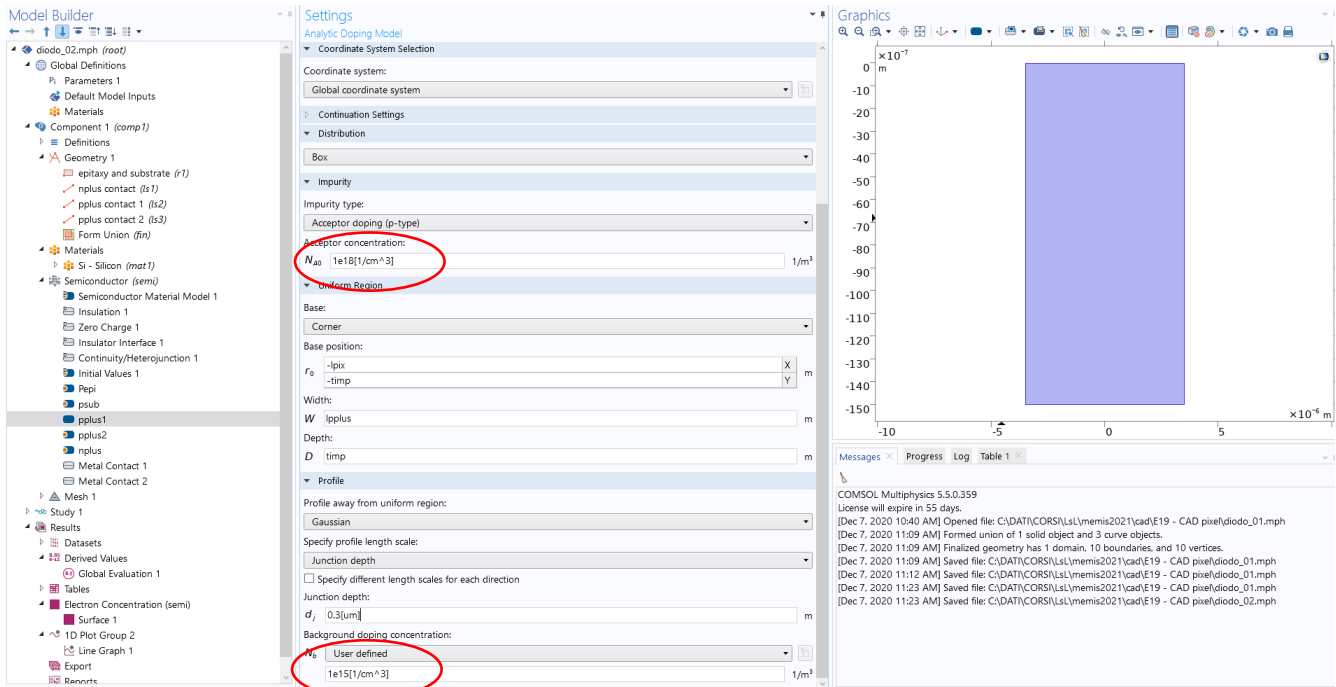

Take care in defining the doping value either in 1/m³, or to specify the 1/cm³ unit if you use it (as shown).

The next point is to add the substrate doping. We repeat the same procedure as above, but this time we select *Box* within the *Distribution* tab. This allows to define a sub-rectangle corresponding to the substrate doping. Note that, for the sake of being more realistic, the change of doping at the edge of the substrate facing the epitaxial layer is not abrupt but has a Gaussian decay profile with a junction distance from the nominal substrate edge of 1 um. This is defined in the Profile tab, which we leave as is for the sake of simplicity. The substrate doping is set to 10¹⁸ cm⁻³, as indicated, while the background oping corresponds to the epitaxial layer value.
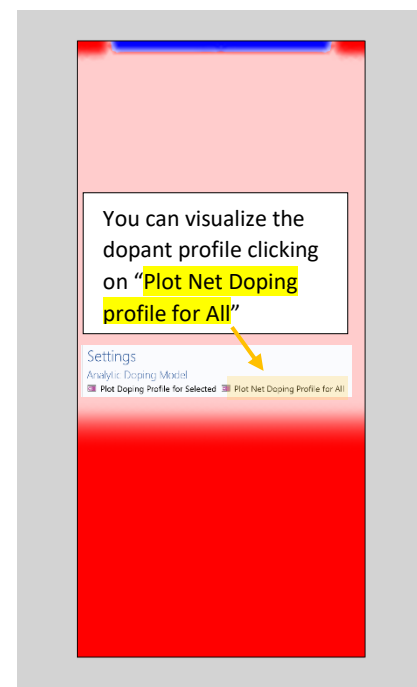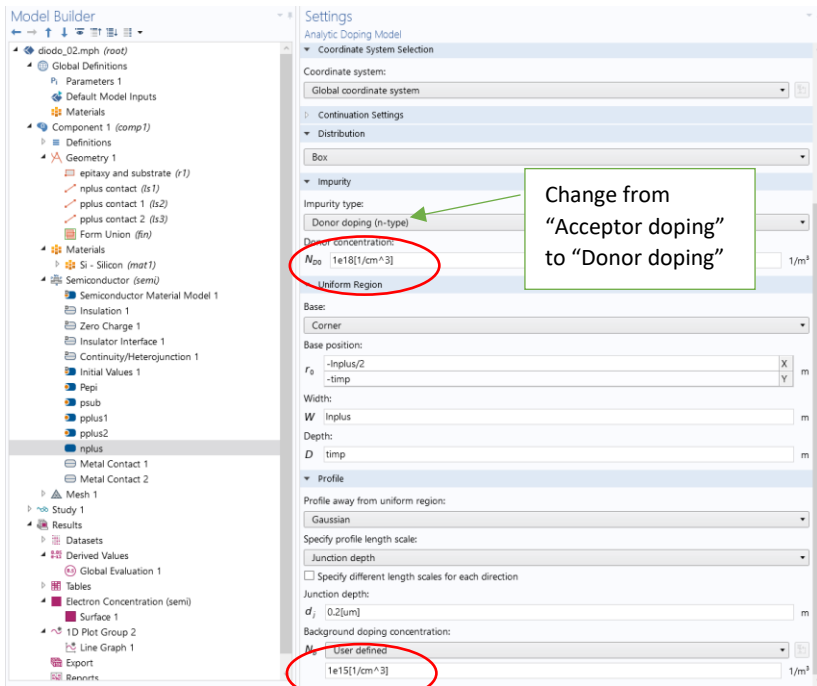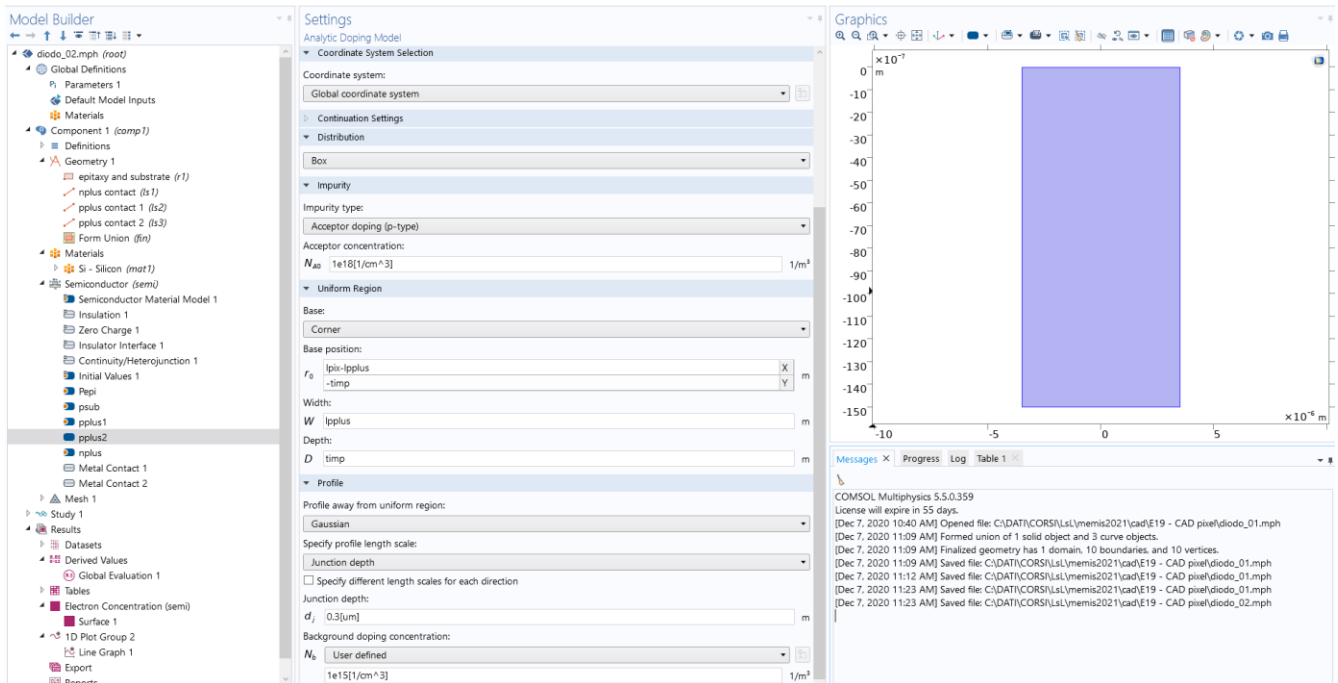
Note that doping values could have been more conveniently defined as parameters. I suggest you do this in your next designs, rather than using the approach that is shown here. It is definitely less risky…

In a similar way, we define the remaining doping regions, which correspond to the three implants (anode and two side P+ regions). With an obvious meaning of the definitions and the parameters, the procedure is shown in the pictures below.
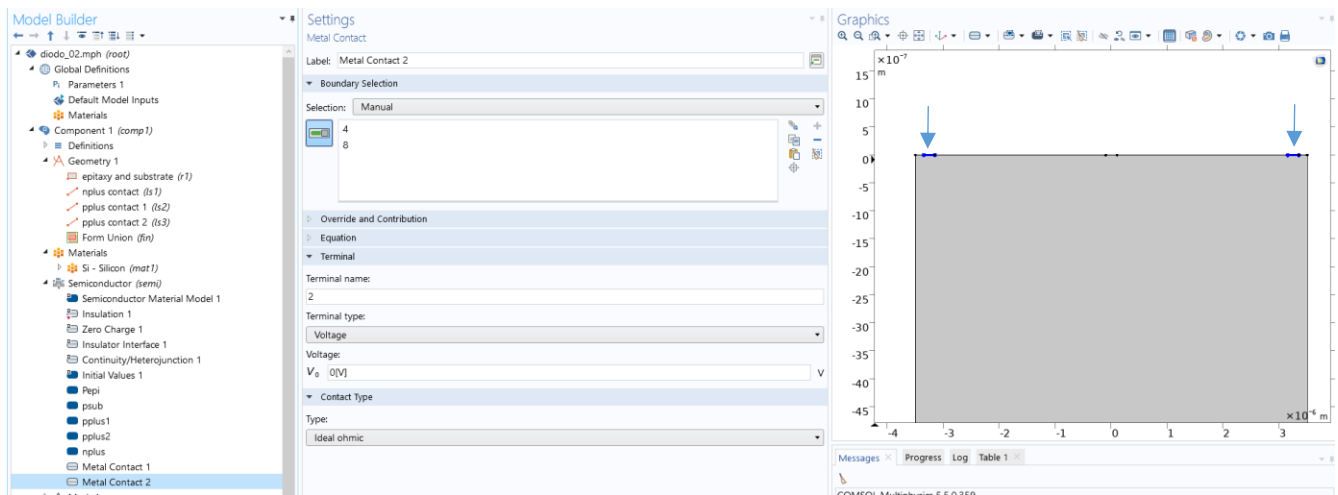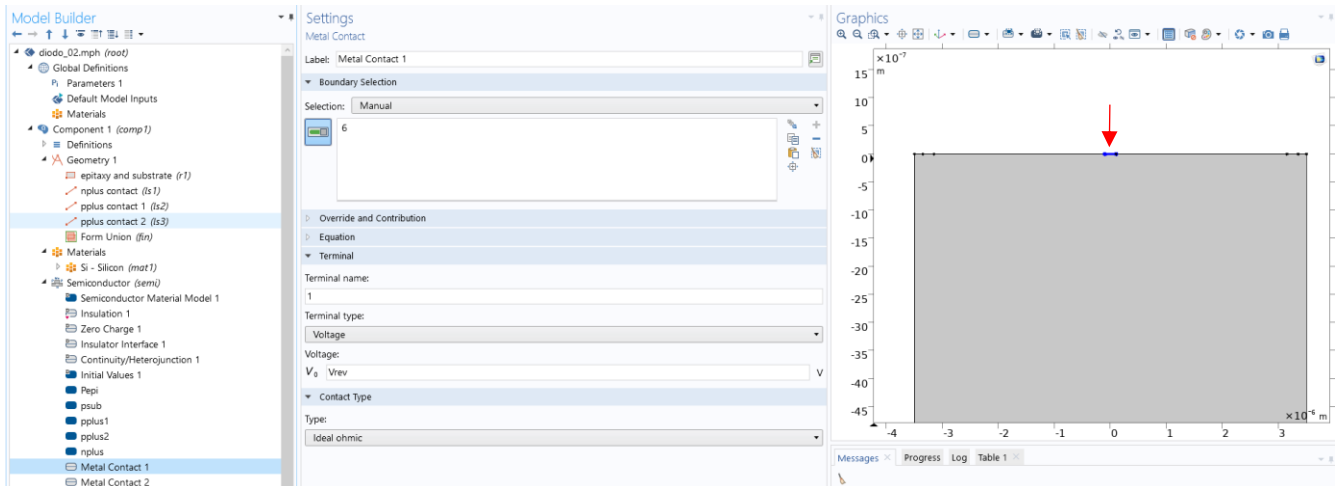

Note just that the implant depth of the Gaussian profile has been decrease to 300 nm (for the P-type) and 200 nm (for the N-type), in order to have a more realistic shallow implant for the N+ and P+ dopings.

Change from "Acceptor doping" to "Donor doping"

You can visualize the dopant profile clicking on "Plot Net Doping profile for All"

Finally, we have to set the biasing voltage: this is done by right clicking on *Semiconductor → Metal contact*. Here you can directly select the contact you want to bias, and the corresponding voltage. We assume a 3 V reverse Bias for the anode, and we ground the two P-type surface contacts.

The two P-type contact will be jointly put to ground with a single *Metal contact* selection (just add both of them in the selection box). Note that all contacts are set as *Ohmic*. In principle, you could also define *Shottky* contacts (but this is not the purpose, unless you want to implement a *Shottky* diode).
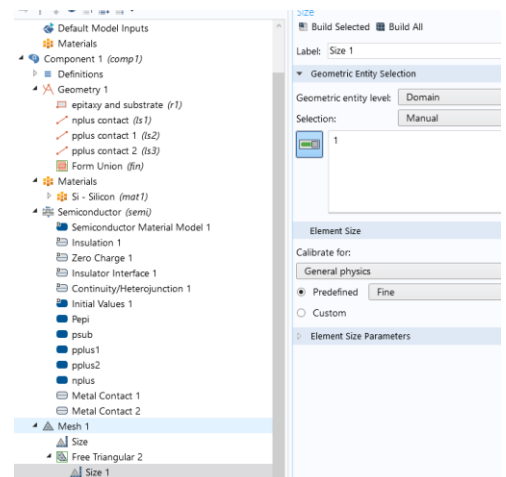
3.   Mesh

In this type of simulations, it is fundamental to have good refinement at the doping boundaries. Even a very simple structure like this photodiode can cause tremendous convergence issues to the solver.
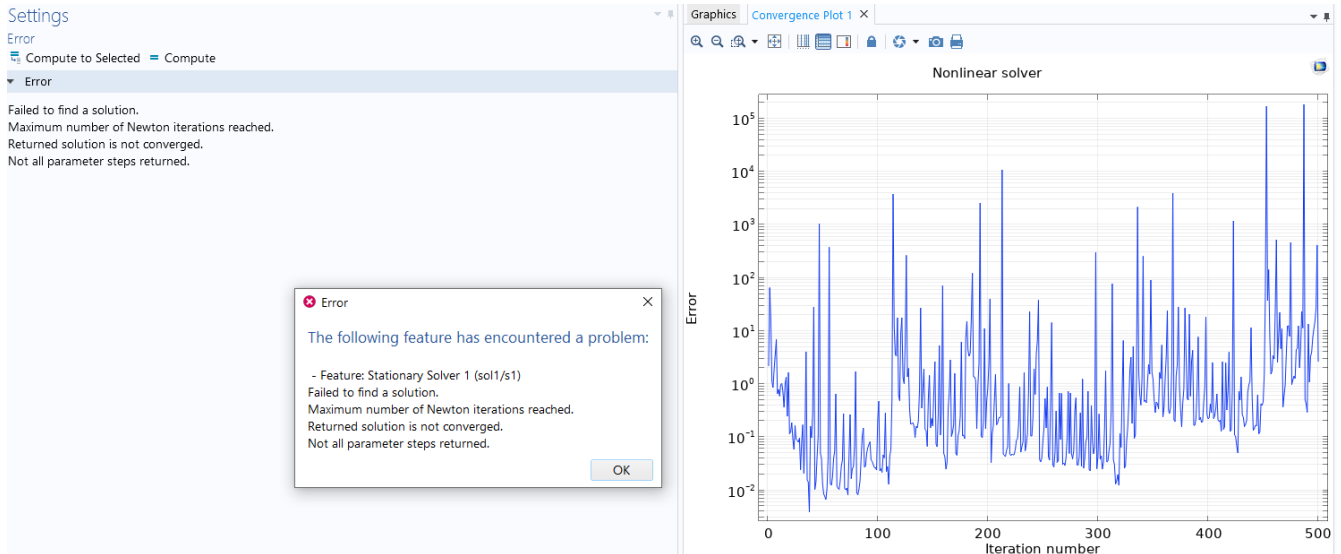
There are essentially two options:

(i)   refine the mesh everywhere (a bit silly, but quick to do…);

(ii)   refine the mesh only around the doping interfaces (a bit smarter, but it takes time to optimize it in details).

For the sake of simplicity, we initially start from the first options, also to let you see the problems that arise with this type of approach. Choose a *free triangular* distribution with a *fine* mesh size, as shown aside. We will later come back to the mesh settings to apply changes.
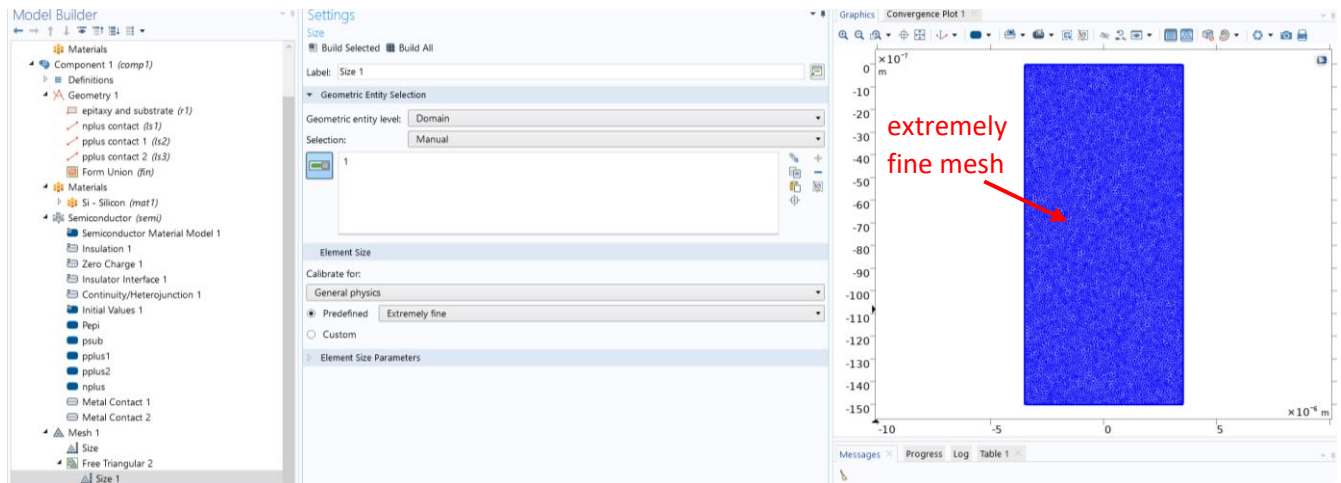
4.   Study

Right click on *Study* and click on *Compute*. While the solution is progressing (i.e. while the system tries to find a convergent solution for the drift-diffusion model on all mesh points), click on the *Convergence Plot* in the top part of the *Plot* window. You will see the solution accuracy plotted as a function of the iteration number (the solver tries to find a solution with an iterative procedure, known as Newton's method, that you may have studied, or will study, in other courses). As you see, the solution does not converge and at the end we get an error notice.
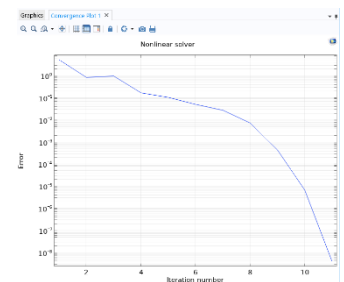

This is due to the fact that the mesh is not refined enough around those regions where doping concentration is changing abruptly (e.g. close to the implants). We need thus to refine our mesh. Go back to the mesh settings and select *Extremely Fine* as *Size*. This increases the number of points over your entire geometry, as shown.
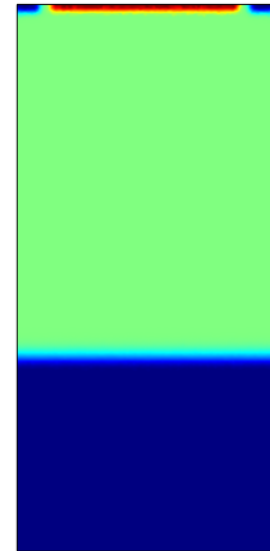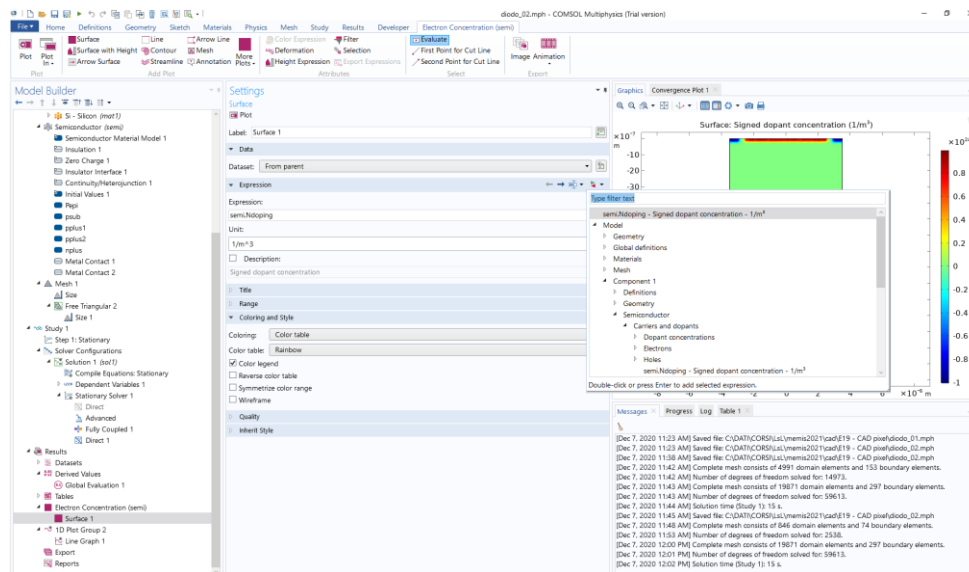

If we now repeat the study, the convergence plot should look much better and we should come to a solution, without error messages, in 11 iterations only.

Note: as already mentioned, increasing the mesh over the entire domain is not the most effective way from a computational-cost point of view. Defining different mesh size in different regions of the domain would be smarter. However, this goes beyond the purpose of this class, and we leave this for your personal curiosity.
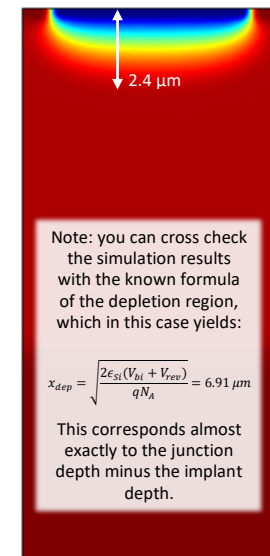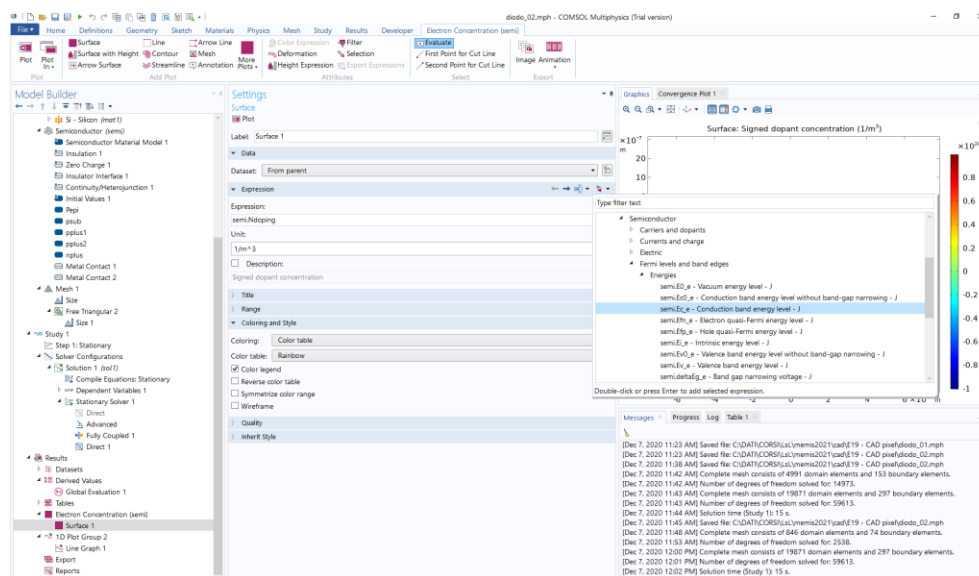
5.   Results analysis and parametric simulations

The result analysis starts by looking at our initial design. If you click on *Electron Concentration* → *Surface* and then in the *Expression* tab you select *Semiconductor* → *Carriers and Dopants* → *Signed dopant concentration* you will see the doping concentration plotted. If you prefer, you can also generate a new plot to keep the standard simulation results. You just need to Right click on *Results* → *2D Plot,* right click on the new plot section and then *Surface.* This effectively matches our initial purpose. Note, in particular, the slightly different implant depth of the N-type region with respect to the P-type region.
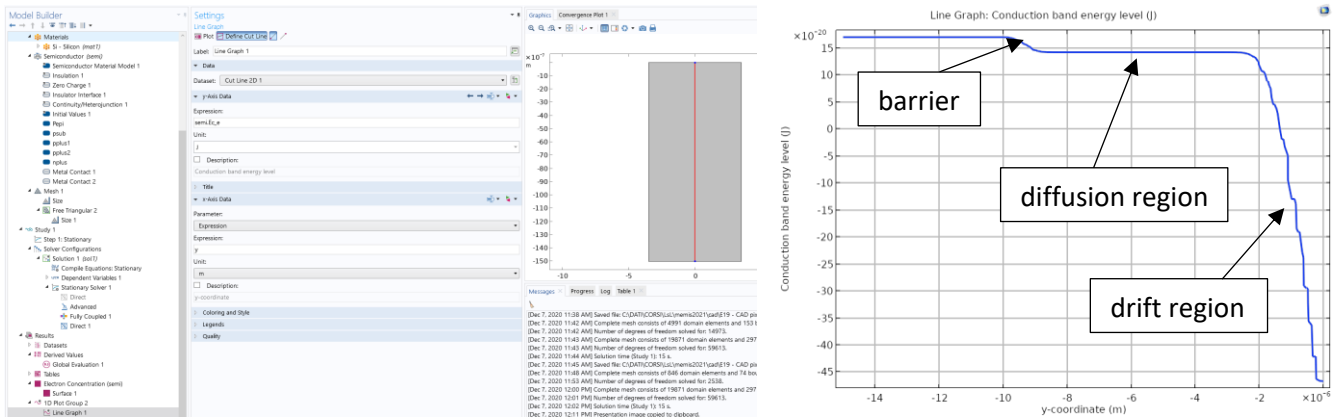

The next interesting thing to plot is the potential distribution, or the energy band graph. This can be obtained by selecting the *Conduction Band Energy* as shown in the figure below. As expected, the band has higher values close to the P-type regions, and "falls" towards the anode. Electrons sill indeed fall in this type of graph, being collected at the anode contact. You can use an identical procedure to plot the electric field, or the carrier density.

2.4 µm

Note: you can cross check the simulation results with the known formula of the depletion region, which in this case yields:

$$x_{dep} = \sqrt{\frac{2\epsilon_{Si}(V_{bi} + V_{rev})}{qN_A}} = 6.91 \, \mu m$$

This corresponds almost exactly to the junction depth minus the implant depth.

To evaluate this more quantitatively, we can draw a cut-line along a vertical cross section. To do this, right click on *Results* and click on *1D-plot group*. Right click than on *1D-plot group* and click on *Line Graph*. Then, define a cut line by clicking on *Define cut line* on the top part of the graph. Leaving the default vertical line through the device center is fine for our purpose. This will set the x-axis coordinate of our graph, which indeed corresponds to the y-coordinate of our geometry.
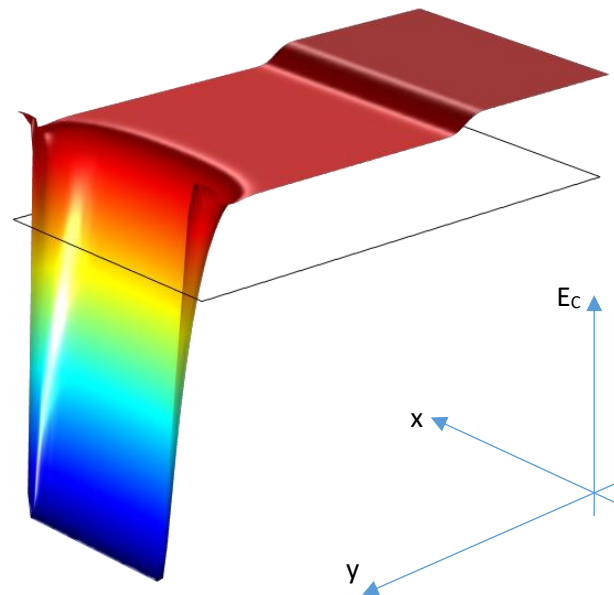
On the y-axis of our graph, we select again the *Conduction band energy*. We get the result shown below to the right, which effectively confirms the presence of an undepleted region (where diffusion will take place) and of a depleted region (where drift will take place). Additionally, we note a small energy barrier that prohibits electrons to flow towards the substrate (where they would recombine due to poor electrons lifetime there).

Note: the fact that the energy band appears a bit "fragmented" is just due to the limited number of mesh points.

Finally, if you want to have a better idea about the 2D value of the energy with a 3D-like plot, right click on *Surface 1* and click on *Height Expression*. This will render your energy band graph in a very intuitive way to understand how Electrons (and Holes) will move within this energy band.

Then, you can try lowering the epitaxial layer doping to check the changes in the extension of the depletion region (try e.g. $5 \cdot 10^{14}$ cm^{-3}).
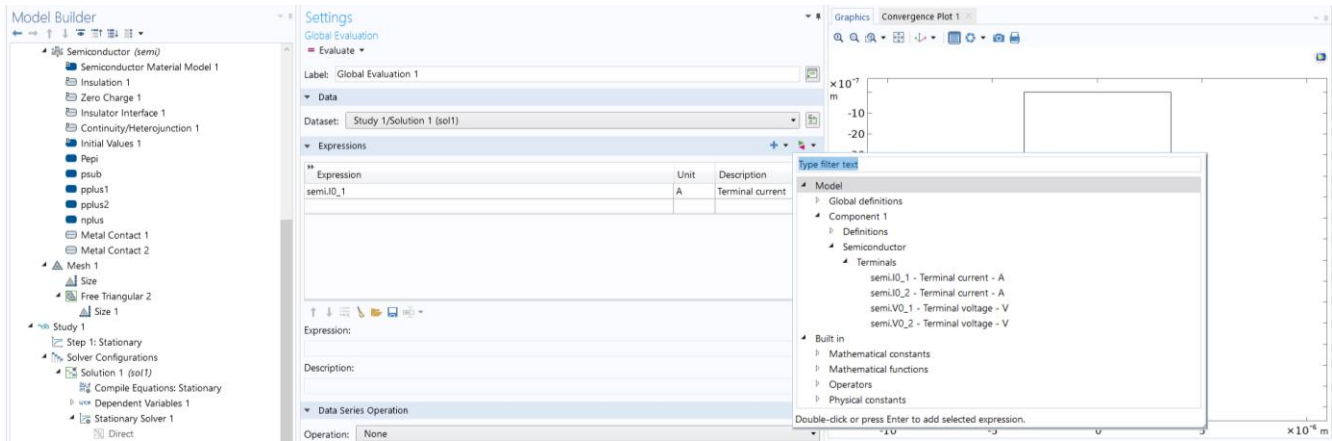

6. Dark current evaluation

The dark current is not a parameter that depends on the specific mesh point. As already learned (see past CAD classes), this is thus a *Global parameter*.

In order to evaluate it, we thus simply click on *Derived values* and click on *Global evaluation*. We then choose our variable as Semiconductor → Terminals → Terminal current 1 (GND electrodes). Clicking on "*evaluate*", we

find the dark current value of about 7 *fA* plotted in the Table. Note that this value corresponds to an out-of-plane thickness of 7 *μm*, as set at the beginning in the design phase.

Once again, we can cross-check this with theoretical predictions: if you go back to *Materials* → *Si – silicon* → *Shockley-Read-Hall recombination*, you find the value of the electrons lifetime (10 us). From the known equation, taking into account the depletion region only, this yields:

$$i_d = \frac{q\, n_i}{2\, \tau_e}\, x_{dep}\, A_{pix} = \frac{1.6\ 10^{-19}\ \ 1.08 \cdot 10^{10} cm^{-3}}{2 \cdot 10\ \mu s}\ 2.2\ \mu m \cdot (7\ \mu m \cdot 7\ \mu m) = 9.3\ fA$$
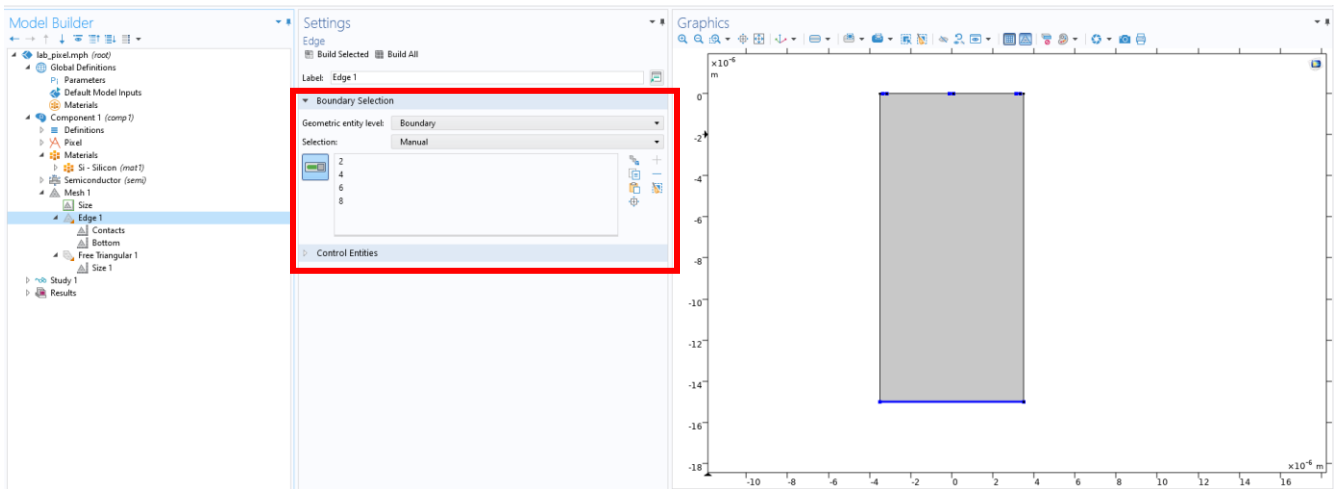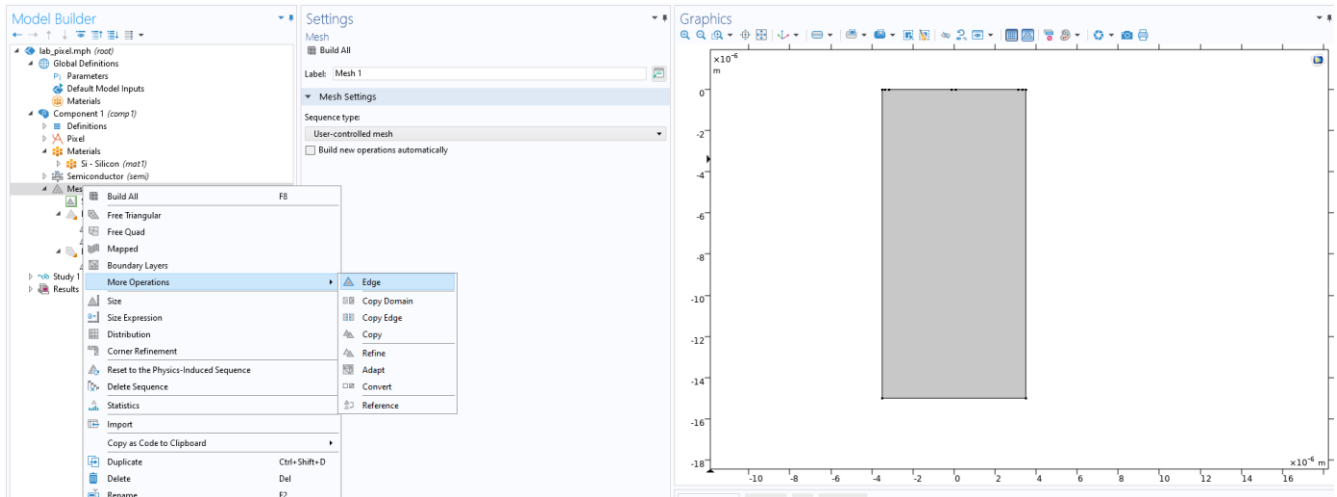
The results we're observing differ significantly from the simulated outcomes. This discrepancy might stem from improper boundary conditions at the top part of the device (excluding the metal region), which could be modelled more accurately. With the continuity condition in place, the software presumes that electrons and holes can traverse that boundary, thereby diminishing the current effectively collected at the contact. This example underscores the importance of never blindly accepting simulation results; always critique and validate them through theoretical analysis!

7.   Improving the mesh

As anticipated, it is not mandatory to define a fine mesh for the entire geometry. Indeed, high accuracy is only required near the doping implants which define the junction, whereas the epitaxial layer and the substrate can be meshed using larger element sizes. This can be done by defining a custom free triangular mesh whose "maximum element growth rate" is specified. This parameter defines the relative size of adjacent elements in the mesh: by defining a fine mesh for the contacts and then specifying a maximum growth rate the mesh will be built starting from the contacts with elements increasing in size across the substrate by a rate determined related to the maximum specified rate. This may help in case you are having convergence issues even with an extremely fine mesh.
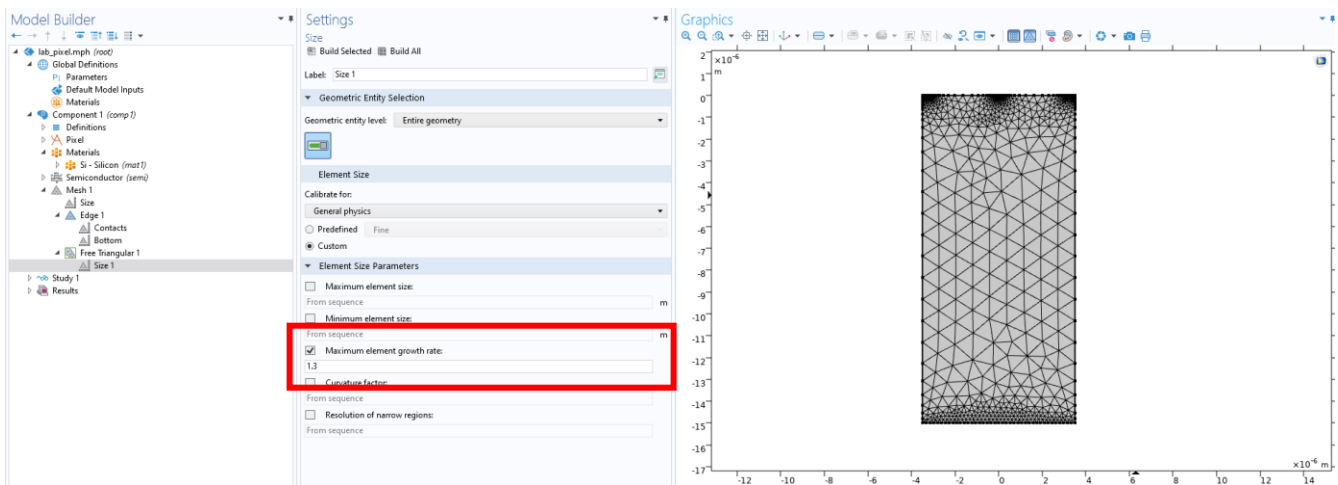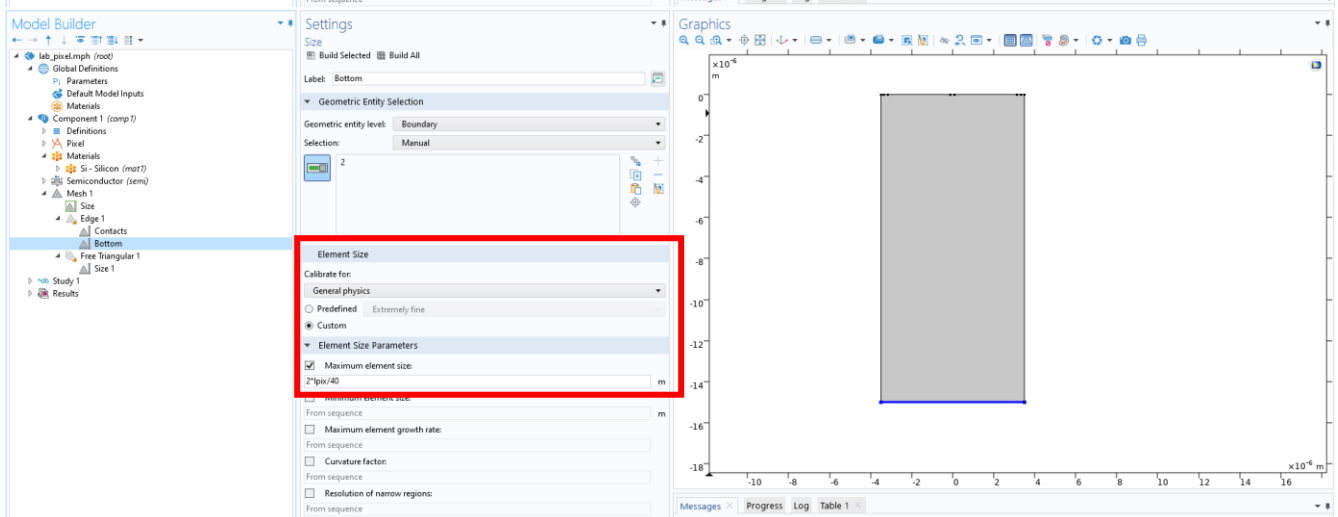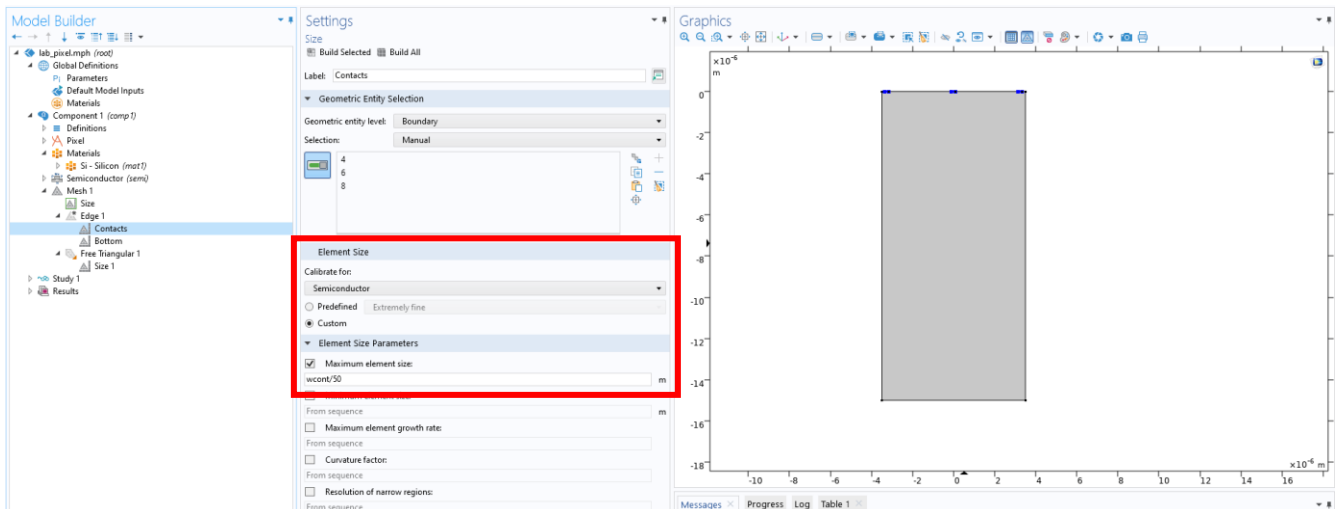
In order to achieve an acceptable accuracy also for the interface between substrate and epitaxial layer we will also define a custom mesh for the bottom side of the geometry, as well as for the contacts. The free triangular mesh will be then calculated by the software starting from both these boundaries and applying the specified maximum rate.

First, let's define the mesh of the edges. Right click on *Mesh 1* and add an Edge mesh by selecting *More Operations > Edge*. In *Edge 1* add all the boundaries we want to mesh to the selection.
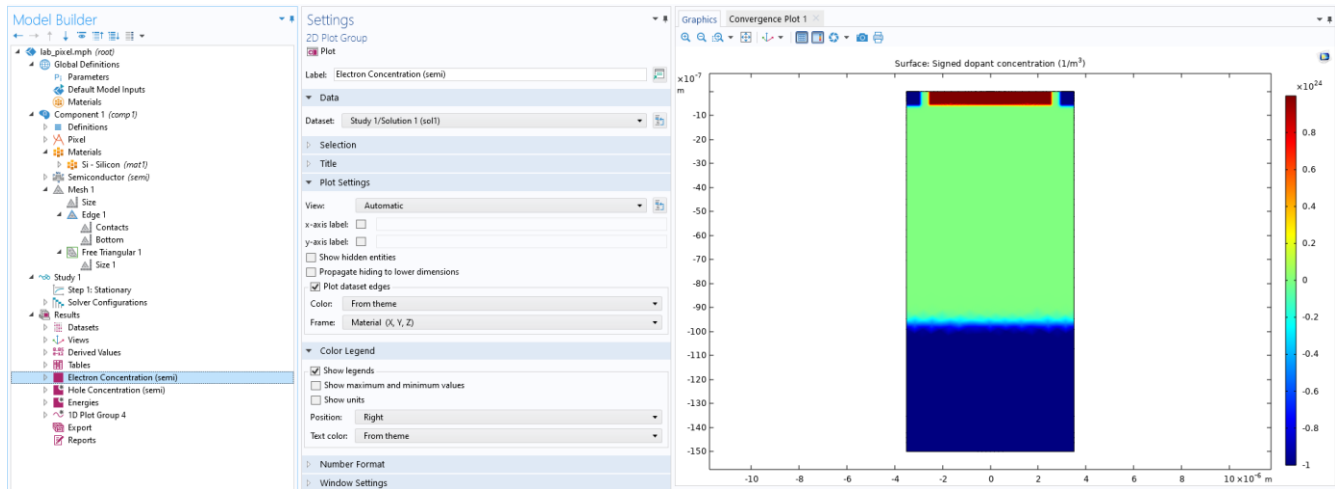
Then let's add two Sizes, one which we will call "Contacts" and one "Bottom" for the two boundaries we want to define. In each size, only add the boundaries you want to mesh (either the contacts or the bottom), under Element Size select Custom and specify the maximum element size to be a fraction of the contact and pixel width (set, for example, a few tenths of elements).

Next, add a *Free Triangular* mesh, within Domain Selection specify "Entire geometry", add a size, select a "Custom" element size, and check the "Maximum element growth rate". The default number is 1.3, which means that each element will be 30% larger at most with respect to adjacent elements (this is a *maximum* rate, so the software may decide to go smaller if needed). Once built, it should appear as in the figure: as you can see, the mesh is now much finer close to the contacts and to the bottom, but it gets coarser and coarser far from the boundaries.

You can now try to compute the stationary study, and if it does not converge try reducing the growth rate (remember that it should be always larger than 1) until it does. Keep in mind that, depending on how fine you defined the mesh on the boundaries, decreasing the growth rate may eventually lead to an excessive number of degrees of freedom, which may significantly slow down simulations. However, the smaller is the growth rate,

the more accurate will be the plotted results. The picture below is obtained with a growth rate of 1.05 and 95118 degrees of freedom.

This is only an example of how you can improve the mesh. You can try to come up with better ideas yourself!