JUDY HOYT: Get started. Just a couple of reminders. Homework number 4 is due today up front. You can bring that up at the end of the hour. Homework number 3 is going back. I've got it in the back there in that orange folder. So you can pick up your homework before you leave today. And there are a couple people who haven't picked up prior homework, so they're in the back of that orange folder.

We've got one handout for today. That's handout 27. Today we're going to be talking about the Suprem4 four process simulator. This is where we are, November 2 here on election day. And you're handing in homework number 4.

The solutions to homework number 3, you're getting your graded homework number 3s back. The solutions will be on the web later today, and I'll bring hard copies next time. They didn't get printed or xeroxed in time.

OK, so let's go on to handout number 27, which is today's handout for today's lecture. This handout or this lecture is on so-called quote unquote "introduction to the "Suprem4 four process simulator." Now, you've already really been introduced to it. Fortunately, you've been using it for the last couple of homeworks in your homework sets. And your TA introduced it in a practical sense a number of weeks ago when she gave a lecture when you had your first Supreme homework.

But I want to talk about it in a little more detail and maybe give you some examples of what it can do, to give you an idea. So in the very first lecture, we discussed an example of a CMOS process flow. And we drew a series of cartoons, PowerPoint artist's conception of what things would look like.

At this point in the class, at this point in the class, we've covered some of the fundamentals and the modeling of a lot of the key processes in that CMOS flow. We've talked about thermal oxidation, diffusion of dopants. We just had four lectures on ion implantation and transit-enhanced diffusion. And we've talked a little bit about how these processes fit together.

So today I want to review some of these concepts and some CMOS process modules and CMOS flows now in the context of what we know about models and in the context of using the Suprem4 four simulator. This is just an example of an n MOSFET and a p MOSFET fabricated on a wafer. This was the cartoon that we showed earlier in this course, in the very first lecture. And what we can do today is we don't need to use cartoons anymore. As you know, you can use the Suprem4 simulator to generate this type of information reasonably accurately.

So let's go on to slide number 2. And I have a series of examples I wanted to go through. And some of these things, since you've already run the simulator, you may be familiar with. But we'll go into some things you may not have seen.

The first three are 1D. Supreme 4, it's a two-dimensional simulator. But it can be run in a one-dimensional mode where you're primarily interested in what happens in depth. But in general, you run it in 2D mode. But to run things quickly, to understand basic physics concepts, it's often easier to run the simulator in 1D mode and then later on run a two-dimensional example. Obviously, when you do it in 2D mode you have a lot more calculations to do and it's a lot slower.

So first, I want to talk a little bit about boron segregation at a gate oxide silicon interface and compare a couple issues. And I'm using this as an example for you to understand a little bit about gridding issues that can take place when you use the simulator. So a very practical issue of keeping an eye on how fine your grid is.

The second example, we'll talk about comparing some arsenic and phosphorus implant profiles, the different types of profiles, to actual SIMS data. We'll see which ones do a better job at actually simulating or reproducing what the data looks like. And then we'll talk about diffusion models, introduce the different types of diffusion models that are available. And again, these three are 1D. The last example-- I hope we have time-- is 2D example of an older technology. It's a 200-nanometer gate length n MOSFET just to go through some of the processing of that.

So let's go on to slide number 3. This is an example, and if you've already done the homeworks, you've probably seen this. This is an example of what's called the mesh or the grid. And Suprem4 uses a triangular grid, triangular mesh. So it's a series of nonoverlapping triangular elements. You see each one of these. If you look at it carefully, it can be drawn as a triangle.

The mesh is very important because what we do in the mesh or the grid is that we-- to frame four tracks, the values of all parameters numerically along a moving adjustable grid. So at these nodes, at these grid points is tracking the values of all-important parameters, the interface-- what the material is, the interface between two materials. It tracks the dopant concentration value. A lot of different things are tracked at each grid point.

If you look at this grid, so this is a grid that belongs to a device structure. Let me see if I have in the next slide. Oh, that's a blow up. Yeah, let's just go a little bit of a blow up on slide number 4. You can see the grid a little bit better. This belongs to a device structure where over here on the left-hand side there's actually a gate. There's a polysilicon gate. There's a space there. I don't know if you can see. Here's a gate oxide. And there's some metal up here. And this is the silicon.

This is an [INAUDIBLE] [? plane and ?] diffused junction. And you say, why does the grid look like it does? Why would it look like this? You see? In this region here, near the channel and near the surface, you see a very fine grid. The mesh is very fine. The spacing between the grid points is very small. Whereas, you get down here in the silicon substrate, and the mesh is very coarse. The spacing between points on which you do the solution is very large.

So why would that be? Does anybody have any idea why up here I would have a fine grid, and down here I have a coarse grid?

AUDIENCE: Because the term-- [? does ?] this change anything?

JUDY HOYT: Right, because up here at the junction you probably have some doping profiles that are changing very rapidly in depth. So you need a lot of grid points to maintain good accuracy and good fidelity in reproducing those profiles. Down here, there's really not much action. Probably just constant doping of boron. So there's no need to have very fine mesh.

Even today's computers are only so powerful. So Supreme tries to do intelligent gridding and not waste and put in a lot of grid points where you don't need it. If you use a very fine mesh everywhere, you'd say, well, that's the best, right? Make it fine everywhere. The problem with that is it would take too long to the solution, and it's really a waste of time.

So just going back to slide 3, this is exactly what we just said. The grid needs to be fine near the active device region, near the interface, or where any quantity of interest is changing very rapidly, depending on what you're interested in knowing about, doping profiles, interstitials, vacancies. Wherever they change rapidly, you need a fine grid. For diffusion, very often Supreme does a finite difference solution for oxidation. It uses a finite element type of technique, again, both on this grid.

OK, let's go on to slide number 5. Actually, I probably should have shown this first. Of that mesh that we just showed, this is the actual device structure. It's half of an n MOSFET. And now it becomes pretty obvious what you were looking at in terms of the grid. The yellow here is the silicon. This is the polysilicon gate.

The blue is the gate oxide. This is an oxide a spacer, a sidewall spacer right next to it. There's a metal contact here to the silicon. And this is the sort of LOCOS region, or the isolation region, in between devices. This is either the source or the drain, and you can see contours here corresponding to arsenic doping. So this is the actual structure.

Very often, in Supreme, things are symmetrical. So another way, besides doing intelligent gridding, if you have a symmetrical device, you can often only simulate half of it and then just project that about the center line here because the MOSFETs are very often symmetrical in the way they're fabricated. So sometimes that helps you save time in the simulation.

So again, once you start-- we haven't had do very many two-dimensional simulations. But once you start doing them, you'll see how much time it can take to simulate all these diffusion, especially if you're doing the most sophisticated models. So you have to take advantage of intelligent gridding.

I want to start on slide number 6 the first example now, which is a simple one-dimensional example that looks at the behavior of boron and how it behaves and segregates at an oxide silicon interface. And this is a Suprem4 input file. As you know, all of you have done your homework, Suprem4 takes text files as input. You can write the text in any text editor you want as long as it doesn't leave a lot of spurious characters that confuse the program. And the program has a parser that reads in the text and interprets the text in terms of the commands that Suprem4 knows about.

Whenever you see a dollar sign, of course, that's just a comment. So Supreme ignores that. So this is only for your own notation. So the first command here is called the mesh command, and it tells you that a parameter called grid dot fac equals 0.04. And then the grid factor has to do with the fineness of the grid. The smaller the number, the finer the mesh.

And so we're defining in this one statement what we want to make the default mesh size in a uniform sense.

AUDIENCE: [INAUDIBLE].

JUDY HOYT: 0.04, it's not in units. I think it's a multiplier. So I think it can be anything from 1 to 10 to maybe 0.01. If you look up in the Supreme manual, it will tell you. I don't remember the exact definition of grid dot fact, but when grid dot fac equals 1, it gives you a certain grid spacing. My experience has been that, for most modern devices, a grid dot fac of 1, it's a pretty large spacing, larger than we would probably want to have. I think I have some examples where we varied this in here, so we can look for the different grid dot facs. And you'll see how it comes out.

The next thing is pretty simple. It says an initialize statement, this says just create the silicon substrate. That's all it does. It tells you the orientation in the z-axis is 100. So that's the orientation of the wafer. And it's boron doped, and it's 1E17 per cubic centimeter. And again, if you need to know things like boron equals 1E17, you wonder what the units are, if you go into the Suprem4 manual, it'll tell you what the default units are for any variable.

This is an implant statement. You just did a homework on ion implantation, where we're implanting boron. So this is the species. This is the dose, the number of atoms per square centimeter, the energy. And you're specifying a tilt angle of seven degrees.

Depending on the dopants, there will be different default models. A lot of the dopants default to the Pearson-IV model. In other words, if you don't say what model to use, it automatically uses default Pearson-IV. I think arsenic defaults to dual Pearson. You have to look up in the Supreme manual and figure out what the default is if you don't specify. So remember, just because you didn't say a model, Supreme has to assume one of the models. And so that's the default model. You need to become familiar with what that is.

This statement here is deposing. It's a deposition statement. So it's not a growth. It's not thermal oxidation. It's literally depositing oxide just by plunking it, (WHIT), on the surface. You're not consuming any silicon. And we haven't talked about this type of process before, but it's called chemical vapor deposition. The next several lectures we'll talk about that.

So this is not thermal oxidation. We're just plunking an oxide down at a certain temperature, 600. And we're telling it the thickness in microns, 0.005 microns. So it's very thin. It's 50 angstroms. DY, if you look in the Supreme manual under the deposition statement for oxide, DY tells the grid spacing in the y direction that you want to use in the oxide.

So DY is the grid spacing specific to the oxide. And look how small it is. It's 0.0001. Well, the reason it's so small is because the thickness is so small. You don't want one grid point in the oxide. You'd like to have several. So we're telling it what the gridding to use just in the oxide.

And then there's a simple statement in which you select what variable you want to plot. In this case, you want to plot the log10 of boron. You can do some fancy-- put in titles and all that stuff. The Save File command saves the file in a format that Supreme can then later use, read in and use later on. It's not necessarily a format that you would find very useful. But Supreme uses it. So Save File into an out dot file is-- defines which file you want it to write to. And this is the particular name we chose, B1 dot inp. You can put it wherever you want.

Here, you're simulating the diffusion. So this is simulating diffusion for a time of 17 minutes at a temperature of 850. And if I specify an inner ambient, it's just like a diffusion, like an argon or nitrogen. If I specify dry O2, then you're actually doing an oxidizing. So oxidation is done, as you know now, by using the diffusion statement and specifying what type of [INAUDIBLE]-- wet O2, dry O2, whichever.

And then, after this amount of diffusion, 850 for 17 minutes, you're going to plot what the profile looks like of the boron after that. So we have an as-implanted plot. And we have a certain line type and color, here, color number one. And we have after diffusion at 17 minutes at 850.

This is a simple file. It's a type of file that you've been using that your TAs been creating for you. And you've been running them. Now we want you to start to understand what those files do so you can create them and modify them yourself. And remember, you can always go to the Suprem4 manual to look up any of these commands, deposition, select, Save file. All those are explained in the manual.

So this is the output of that file on slide number 7. This is what it looks like. Of course, I've doctored it up. I've added a little color and things to make it look a little better, but it's the basic output. What that file puts out is a plot. On the y-axis is the boron concentration. The x-axis is distance. And Supreme assumes that 0-- it puts 0 at the original silicon interface, or surface, I should say. So if you deposit layers on top of that, they will be sort of on the negative x-axis.

Remember, we deposited 50 angstroms of oxide. That's what I'm showing in the yellow here. If you go back one minute and you look at deposition oxide, thickness is 0.005 microns. That's 50 angstroms. So that's what this is, from here to here in this yellow region.

And there are two profiles that are shown here. The black one, the dark one is Supreme plotted the boron profile immediately after ion implantation. So that's the as-implanted. And the red one is the boron profile after that anneal that we did, 17 minutes inert at 850. Of course, it's diffused. The peak has gone down a little. It's broadened out. And there has been a little bit of segregation at the oxide silicon interface.

The diffusion of boron and oxide is relatively slow. It's about a factor of 2,000, roughly, a factor of 1,000 slower than diffusion in silicon. So you see a fair amount of motion of the boron and silicon. But when it hits that wall, hits that oxide, it doesn't diffuse very fast. So you're not going to get-- you see in the silicon there's very little boron after 17 minutes has actually diffused in.

If you want to know the segregation coefficient, you can look it up in the appendix. They have an appendix in Suprem4 for each of the dopants. For boron segregation, which is the ratio of the concentration of boron in the silicon to that in the oxide at 850. Supreme calculates it from this formula. You can modify either one of these constants, by the way, if you want to. And I calculate it at about a factor of 10 to 1, or 0.1. So that would be this peak value to the value in the oxide.

So that's what that simple file puts out. Now, let's look at a couple of other special cases. So we're going to now do something similar. But instead of depositing the oxide straight down by deposition, from chemical vapor deposition, we're going to form the oxide by thermal oxidation. This is the type of model that we've talked about before. And we're going to display the output.

Here on slide number 8, I'm showing a new Supreme input file, or command file. And this time, instead of saying deposition, we're going to say to do a diffusion. Time equals 5-- that's five minutes-- at temperature of 850 in dry O2. So it's going to grow five minutes worth of oxide at 850 in dry O2. And then we're going to plot it using color number 2. And then we're going to do another five minutes and plot it again.

So we can do sequential. So that will be after 10 minutes. So you can look at it and see how the oxide grows and see what happens to the boron profile as the oxide grows. So we're going to display it at 5, 10, and 17 minutes. We're going to display the boron profile in the oxide.

So on slide number 9, this is after five minutes of now thermal oxidation. We have an initial oxide that's formed. And this is what the boron looks like. And you notice the boron near the surface has been depleted, again, because of that segregation. And now you're oxidizing, so you're consuming silicon. So you're moving into the silicon. This interface is moving towards the right.

And same segregation as the previous case. And now we have a little bit of boron that's been incorporated in the silicon, not so much by diffusion because, again, but by consumption because, remember, we initially implant it with boron. And then you start oxidizing. If you go back to the command file, read in the as-implanted structure.

So remember, we did a save file of B1.imp. How did I get that? Well, that save file-- so this is a way to just save time. After we did the as-implanted, we saved the file in something called B1 dot inp. And then we're going to call it back later, that same file, in a subsequent Supreme run. So here, I'm loading in a file called B1.imp.

If you didn't want to do that, you could repeat the simulation of the implant. It's not going to take very many minutes. But if you have a complex two-dimensional simulation, you might as well save it. It takes an hour to run. Save it. And you can always use it again as the input to another Supreme simulation, another diffusion or oxidation.

OK, so we are going to-- in this here back on slide number 8, we're going to load that in, take whatever oxide was on there off just to make sure it's not there, and diffuse it at 850 for five minutes in dry O2. We're going to grow a couple of different thicknesses of oxide.

So I've consumed this, and that's how a lot of the boron has gotten in there. And you see the segregation from this point, from here to here, that ratio is about a factor of 10. That's what we expect because we said k0 for segregation of boron in Suprem4 is about a factor of 10.

Go to the next slide, slide number 10. After another five minutes, so I have a total of 10 minutes of oxidation. Again, you see, now you've grown a little more oxide. The yellow is thicker. And you've consumed more of the boron, and the segregation is causing this effect. So this is not really diffusion because the diffusion, look how rapidly the boron drops. If it were diffusing, this would go straight through the oxide. But it's not because there is slow boron diffusion in this assumption in the oxide. You see the boron profile in the oxide.

And if we finally go to the last, which is after the full 17 minutes, we have grown now 50 angstroms of gate oxide at 850 in 17 minutes. And the red is what it looks like, the boron looks like due to thermal oxidation. You see a lot of the boron has been incorporated into the oxide because of segregation effects and consumption.

The black is as-implanted, so that was how it was as-implanted. If I thermally oxidize it, I get the red curve. You can see it's depleted at the surface, and it's piled up in the oxide. Instead of thermally oxidizing, let's say I didn't want to grow my gate oxide, I just deposit it, 50 angstroms at 850. The boron profile is quite different. It's the blue.

So why is that? So there's two different cases here. In the case of if I do an implant and I thermally oxidize it, I consume silicon. And I suck some of that boron into the oxide, and then it segregates by segregation. So you see? The boron concentration is relatively low at the surface.

If instead of thermally oxidizing I just implant the boron, place an oxide down there by a chemical vapor deposition process, then I don't consume any of the silicon. So the boron stays a lot higher. So I have about a factor of two more boron at the surface. So you can imagine you would get different device characteristics due to this slightly different ways of forming this oxide because it has an impact on the boron profile by different diffusion. Well, the diffusion coefficient's the same, but it's really different segregation and the fact that in one case we're consuming silicon. In the other case, we're not.

OK, that's just a simple example. I had to use-- you see all these points. I had to use a really fine grid in order for you to see this effect. [INAUDIBLE]. So let me talk a little bit about that grid dot fac and the gridding.

When you have very shallow profiles, which is mostly what you're doing these days for CMOS, and thin oxides-- oh, OK, here we go. This gives us the answer. If you use the default grid, then the grid spacing is about 0.1 micron. So there's about 0.1 micron. That's actually bigger than a lot of devices. So it's much too crude. So we can multiply that by a small number, like the grid dot fac and make that grid spacing much, much smaller.

So in this particular example, I'm showing you where we did that same implant of boron and then 850, 17-minute oxidation, with a grid dot fac of 0.4. It's still very coarse. You can see. This doesn't really have much of a shape to it. There's a point here where we've solved. There's a point here. There's a point here. But it's just very jagged looking because the grid is too coarse.

And in fact, in the oxide you're getting almost no detail or information at all because you really don't have any grid points in the oxide. You've got one at the interface and one at the surface. So for this particular example, a grid fac of 0.4 is way too crude. So this is something you always want to check by doing a quick test. What kind of a grid do I want?

If you go on to the next slide, slide number 13, here's a factor of 10, finer grid, grid dot fac of 0.04. Still, it looks pretty good. Now, it's starting to look like a smooth curve. It's not all jagged looking. And it has a shape that you would imagine could be associated with diffusion and-- ion implantation and diffusion.

But the oxide grid, in the oxide, it's still too coarse because the oxide is thin. It's only on the order of 10 to 50 angstroms during this oxidation. So I'm only getting what? One point in the oxide. I have one at the interface, one in the oxide, and one here. That's still not enough.

So if we go to the next slide, slide number 14, and as I mentioned before, we use this particular statement to make a very fine grid, and just inside the oxide. We don't want to make it that fine everywhere. Then we use this-- you can insert a statement called methoddy dot oxide equals 0.0005. That puts in a 5 angstrom grid in the oxide, which kind of makes sense. You have a 50 angstrom oxide you're growing. You'd like to have something like a 5 angstrom grid in that oxide.

And now you can start to see what it really looks like. Remember, before I had one point here, one point here, and one here. It's a solution looked completely different. This is what it would have looked like if you just hadn't bothered to put in a special grid in the oxide. You'd think it was a triangular sort of profile. It's not triangular at all. The boron actually looks sort of flat topped in the oxide and then goes down very rapidly.

So you really need to keep an eye on your grid. And you typically want to run problems or solutions for a number of different grid spacings to get a feel for what grid is appropriate and what grid gives you physically realistic results.

OK, so that's a little bit of a trivial example on oxidation, but I just wanted to make the point of how important knowing your grid is because you can get a solution. You look at it, and you think it shows you something. And in fact, it could be physically not meaningful if you haven't used the right grid.

So let's go on to slide number 15 and example number two. I want to talk about ion implant modeling. Remember, there's a number of different analytic models that we talked about. We said, you can use a Gaussian. You can use a Pearson-IV. You can use dual Pearson. And that there are tables in the literature that give the first three moments, the RP, the delta RP, and the skewness. Those tables have been tabulated since the 1970s or so. And they've been updated over time-- of the implant distributions. And they've been produced by theory and also by fitting to experiment.

This is the most common analytic formulation that you will see in simulators. It's called the Pearson-IV. We talked about that distribution. And we said it does a pretty good job of replicating profiles into amorphous case. It does not model channeling all that well.

Channeling requires the use of more parameters, generally by curve fitting. So a Pearson-IV is a single Pearson distribution. If you want to do channeling, you'll often see dual Pearson, meaning two different Pearson-IV profiles are added together. And lookup tables are used.

One of the things you need to be aware of are inaccuracies in stopping powers, particularly the electronic stopping at low energies and even nuclear stopping powers. At very low energies, a lot of these stopping powers are not known that well. So if you're trying to simulate energies below, say, 10 keV, 20 keV, depending on the dopant, you may not get very good results.

Dopants that are widely used in Silicon technology, they're constantly updating these tables to try to make them a little better, more accurately represent reality in experiments. But some dopants that are not used that often at low energies, you'll simulate, and you won't get a very good simulation compared to the actual data.

So let's go on to slide number 16, and this is an example of a Supreme input file where we're using tabulated moments. So Supreme has in its database tables of all these moments. And it's using tabulated moments to evaluate and to plot for you an implant distribution.

So here's an example of ion-implanted phosphorus. So we define our mesh. Here's a grid fac equal 0.1. So it's not all that fine of a mesh. The substrate is boron. We're going to implant phosphorus at a particular dose, [? 8014. ?] Energy, not all that low, about 30 keV, but somewhat low, a 7-degree tilt. And we're saying to use a Gaussian model, very, very simple model. Save it in a file called Gaussian.saved, and then plot it.

And then we're going to plot it. In addition to plotting the phosphorus that it calculated, Supreme has a nice way of-- using a command called profile, you can input into Supreme xy data, a matrix of concentration versus depth. And that's called the profile statement. It reads in phosphorus. And the name of the file-- this is the particular name. It's kind of ugly, but you can call it whatever you want-- and use that to plot the phosphorus.

So we can not only plot the Supreme simulation, we can also plot the actual SIMS data. So this happens to be an implant for which I have actual SIMS data that I have entered into this. And we're going to compare that data to the Gaussian solution.

So if you go on to slide number 17, you'll see. This is for that relatively low-energy phosphorus implant. The SIMS data here is shown in red. This is actual data obtained from a company called Charles Evans. I think we've talked about Evans Associates in this class. They used a cesium beam with a primary beam energy of 2 keV to profile. And this is what they got.

And the black line is what Supreme says it should look like using a Gaussian simulation. And these are default values that are tabulated inside the Suprem4 simulator using the standard LSS range parameters, RP and delta RP. Remember, you only get two for a Gaussian.

So compared to the SIMS, it's not all that great. The simulation does not do a very good job. It's too simplified of a model. The stopping power is apparently too small, and the range is overestimated by the theoretical calculation compared to what was experimentally observed.

Well, we can try another model. Same SIMS data. This time, we try the black model, which is the dual Pearson, OK? Dual Pearson uses more parameters. It simulates this sort of channel tail a little bit better. But again, the range and the skewness are both overestimated, way too much skewness.

So what does this tell us? Well, this basically tells us that these tabulated moments are not perfect for all dopants. So take them with a grain of salt, especially phosphorus. Phosphorus [INAUDIBLE] used that much in CMOS technology, right? Shallow junctions for n-type are typically made by arsenic because phosphorus diffuses too fast. So not a whole lot of energy and manpower has gone into improving the moment tables for shallow phosphorus.

So if you happen to be using phosphorus in your experiments, take it with a grain of salt if you're using these tabulated values. According to my data, it doesn't fit that well. Of course, there might be something wrong with my data. That's another possibility. So exactly what's reality, it's not always clear.

But interestingly, on the next slide, exact same data this time compared to Monte Carlo simulation. Now, these Monte Carlo simulations take a lot longer. It's not an analytic. It's a numerical solution. You have to follow each ion into the silicon and see where it lands up and then statistically create a profile. And that's what the black is, OK?

And you can see the black looks kind of jaggy because, again, you're statistically creating this profile. But interestingly, the black does a reasonably good job of agreeing with the SIMS. Doesn't mean they're right, but it kind of gives me some confidence that, when I see a simulation agrees pretty well with the SIMS data, at least in terms of its range and its delta RP, its width here, that's probably a pretty good sign. So there's enough physics in this Monte Carlo simulation to reproduce the data pretty well.

But notice, this is how you tell Supreme to do a Monte Carlo implant. You give it the implant statement. You tell it the species, the dose, the energy. Now, here you tell it how you want to do it. You see? In prior, I had said Gaussian or dual Pearson, or if you say nothing, it defaults probably to Pearson-IV. Here, you said Monte Carlo.

Now, an important thing when you do Monte Carlo is to specify the number of ions you want it to shoot into the sample and to follow. So this is 25,000 ions, not very many, and tilt and rotation. You need at least 10,000 to get good accuracy. Probably, 100,000 is better because you can see here, after I'm two decades beyond the peak, I start to get a lot of noise. So the noise comes in. So I'm really only getting two decades of smooth data. That's with 25,000 ions.

In a one-dimensional simulation, this isn't too bad. It might only take 5, 10 minutes. No big deal. But now this is just 1D. If I'm doing this in 2D across the channel of a MOSFET, that 5, 10 minutes goes to be three or four hours or more. So you have to really compromise to a certain extent.

So what people will do sometimes is fit the 1D profile using Monte Carlo and then fit to it an analytic solution by changing some of the default parameters. And then use that analytic in the two-dimensional simulation as a way of getting around spending so much time. So that's just an example of how actual Supreme outputs compare to real data.

Let's take another case. So that was phosphorus. As I said, phosphorus is not all that widely used. Here's the case for 30 kilovolts, same energy, but this time arsenic, using the dual Pearson analytic. Again, the SIMS data is the red line. And the calculated is the black. And this time, with the dual Pearson, it does a pretty good job. You got the range almost just right and even the broadening, not quite, but pretty close.

The shape are much more accurate than for phosphorus. I don't know why, but I'm guessing it's because shallow arsenic junctions are the way people make junctions these days. And so people have updated in Supreme these tables, these tables of moments to fit better the experimental data. Arsenic is also heavier, dominated by nuclear stopping, which is a little better understood than electronic stopping.

Remember, we said electronic stopping powers always have a little bit of fuzziness about them. Phosphorus is going to be lighter. It has more electronic stopping. Maybe that's why. But again, take the profiles you get with a grain of salt unless you've checked them out experimentally.

Slide 21, that same data, this time with Monte Carlo. Excellent job. This looks even better. It's got the broadening a little better. So it looks quite good using Monte Carlo. Of course, Monte Carlo took longer to generate, longer to simulate.

Another advantage of the Monte Carlo besides the fact that it seems to be pretty accurate, it has pretty good comparison to data, you can generate profiles of [INAUDIBLE] interstitials and vacancies. And we need those profiles. If you want to use those profiles, you can use them in subsequent diffusion simulations in order to simulate TED.

If you do an implant and you don't tell the simulator to use a damage model, it's not going to be able to simulate TED because it needs to have some damage model. So Monte Carlo is also very useful for that purpose.

OK, so let's go on to slide 22. So those were a couple of examples of simple oxidation, simple ion implantation. Now we get to the more-- a little more exciting models, a little more complicated. There are three major diffusion models in Suprem4. And these are what they're called.

The way you invoke a different model in Suprem4 is you use the method statement. And there are three methods for solving the differential equations associated with diffusion in Suprem4. There's PD. That's Partial Differential equation. PD dot Fermi is a method that takes into account the impact of the Fermi level, just like the name would suggest on the dopant diffusion coefficient.

For example, this is an equation that Supreme uses for n-type dopants. You're very familiar with this. It has an n over ni, and it may be an n over ni squared depending on the dopant. And it has stored in it values for D0, D minus, D double minus for all the dopants. It models concentration-dependent diffusion, but it does not model TED or OED.

So it's quite simple. It's very fast. And the nice thing is there's relatively few parameters, just a few D parameters. And you can get your answer. So that's the nice thing about it. And it does model concentration-dependent. So this is the first thing you would use because it's relatively fast.

The next method, the next level of complication is PD.trans. And it takes into account the Fermi level. So it already has this model built in, the Fermi model, as well as the impact of nonequilibrium interstitial and vacancy profiles on the diffusion, but not the other way around. It will not show you the impact of dopant diffusion on the motion of interstitials and vacancies. So it's not fully coupled, but it does take into account the impact of the ions on the dopant diffusion.

It's very useful for OED. So people tend to use it for oxidation-enhanced diffusion. Or you can use it for transient-enhanced diffusion with relatively low doping concentrations. If you have higher doping concentrations, then the coupled diffusion of the pair, of the dopant plus the point defect actually affects the point defect profile. And so you need something more sophisticated than PD.trans.

And this is the model it uses, the basic concept that you should be familiar with now, if you've done your homework, is that the diffusivity is the unperturbed diffusivity times something in parentheses, where it depends on F sub i. So it takes into account the enhancement in CI over CI star that can happen when you do an oxidation or the enhancement in CV or CV star if you do a nitridation or if you do an implant. And these things get enlarged or suppressed. So it'll take that into account.

The last method you can use takes the longest in terms of it can be very long simulation times, PD.full. The name comes from the fact that it is fully-coupled diffusion. And what does that mean? It means that the interstitials and vacancies impact the flux of the dopants, and vice versa. The flux of the dopants impact the interstitials and vacancy diffusion.

This should be used most of the time if you are interested in transient-enhanced diffusion. And in high concentrations, in general, for instance, the emitter push effect, we talked about how phosphorus can pair with interstitials and, by diffusion, drag the interstitials into the substrate.

When it goes substitutional, it releases those interstitials. Those interstitials then cause the boron base to broaden. That was the emitter push effect. You'll never be able to get that out of the Fermi model. There's no way. And even the trans model won't have that. So to simulate the things like emitter push effect, those types of fully coupled cases, you need this PD.full.

So those are three different statements, and you can compare the results from all three statements for a given situation. Now, one thing I should say. As you go from here to here to here, you're invoking more physics, more chemistry, and also more parameters. That's one problem with this. Sure, you can model the emitter push effect. But there's a lot of parameters you need to know, the diffusion rates of the vacancies and the interstitials, their recombination of interfaces. If those parameters are in Supreme, they always have default parameters. But they may not be accurate.

So again, take everything you simulate with a grain of salt. If somebody has input something in there to the best of their knowledge, that might be the interstitial diffusion diffusivity. But maybe it wasn't well characterized at the temperature under the conditions that you're using. So you can get lots of different profiles, but how accurate are they represent experiment is something that you need to determine for yourself.

So here's an example. Let's look at a couple of examples of different diffusion profiles. On page 23, this is a simple 1D arsenic implant, and then we're going to do a long-time diffusion. So the simplest thing [INAUDIBLE] is an implant arsenic at 30 keV, the certain dose, intermediate dose, 2.6e14, and diffuse it in a furnace at 1,000 degrees for 30 minutes. Before we did the diffusion, we put down an oxide cap, not by oxidation, but by deposition. And so that oxide cap helps prevent the arsenic from evaporating out of the wafer.

So this is the actual command that was used that specifies the point defect model, methodpd dot Fermi. So we're using the Fermi diffusion model. And this is the diffusion statement, 1,000 degrees for 30 minutes inert.

I have some SIMS data here, which is shown in the red. And then the black dots or the black circles are the simulation using a Gaussian model. Here's a Gaussian and a Fermi. And you can see it actually reproduces the data amazingly well. It means that somebody calibrated Supreme pretty accurately to this particular furnace. The Fermi models good enough.

And you say, well, how about all that-- why didn't you use PD.trans or fully coupled? How did you get away with such a simple 30-second simulation? Well, the anneal is long. So there's not going to be any big damage or TED effects after-- normal diffusion is going to dominate this. You need concentration dependency how box-like it is.

So you need to use the Fermi model for Fermi-level effects. But you don't need to go and invoking necessarily TED or anything. So if you know you have a case where TED effects are not that prevalent, you might as well use the Fermi model. It'll go a lot faster, and in a two-dimensional simulation, that could save you a lot of time.

Here's that same simulation here on slide 24. But this time, we used the fully coupled model. So when we did the simulation, instead of methodpd dot Fermi, we said methodpd dot full. And the simulation is the black line. It agrees still very well with the data, maybe not quite as well. But to within all the experimental uncertainties, it really does a good job.

So the-fully coupled model doesn't make much difference-- that should say difference, not different, difference-- because 30 minutes, again, is much longer than the time scale for TED. At 1,000 degrees with this dose, I'm guessing TED time scale is probably less than a second or so. And in fact, if you go on to slide 25, we can even figure out roughly what it is. At 1,000 degrees, this was for 1e14 phosphorus. At 40 keV, TED is less than a second.

So for 2.6e14, again, it's 2.6 times that, almost the same range, not much difference. So we're talking on the order of a few seconds. So a 30-minute anneal, clearly, TED is not going to be very important. That's why you get the same results. But this is a sanity check you can do with Supreme. Make sure you get the same result for that 30-minute anneal with PD.Fermi and PD.full. If you don't, then there's something missing in your understanding of what Supreme is doing. So there's a simple example.

Now let's do something different. On slide 26, instead of 1,000 degrees for 30 minutes, I'm going to do ten seconds, 1,000 degrees for ten seconds. OK, now you're in a range where you can imagine there might be some TED effects. And this is the simulation we're getting.

The as-implanted here is shown in the black. And I believe this was implanted with a Monte Carlo. In fact, you can tell it's Monte Carlo because, see all this jaggedness in the as-implanted? You'd never get that out of an analytic solution, right? The jaggedness comes from the statistical nature of Monte Carlo. So that's the Monte Carlo as-implanted in black.

PD.Fermi, which does no TED, shows not very much motion of the arsenic for a 10-second RTA. PD.full you get a fair amount of broadening. So that's got the TED built into it. Now, exactly which one is more accurate, I really can't say. In fact, my experience with arsenic has been on our rapid thermal anneals, that Supreme tends to overestimate a little bit the amount of motion compared to what we see by SIMS, if you actually do a SIMS profile of 1,000 degree 10 second.

Of course, there's always the question of how well is your RTA calibrated. Rapid thermal annealing machines are extremely difficult to calibrate. They are nonequilibrium environments, right? It's not like a hot furnace where everything's the same temperature. The wafer is the only thing that gets hot. And measuring its temperature is kind of an art. So it's hard to say.

So you have to be careful, and typically when you're doing rapid thermal annealing and simulating TED, you want to compare it to a couple of experimental results just to make sure that it makes sense. Now, I can make this PD.full look like whatever I want by changing some of the internal parameters in the Suprem4 simulator.

If I change things like the diffusivity of the interstitials and vacancies, I can change what this profile looks like. So you have some latitude there. This is using default parameters that are built into Suprem4.

Let me go on slide 27 to another related example, where we're going to talk about how-- we talked last time about some clever experiments that have been done at the IEDM, where the order of the anneals-- we have the exact same anneals, but we're going to reverse the order which we do them makes a difference because that's going to dissolve a lot of 311 defects.

We're going to show how Supreme can actually simulate that. And first, I'm going to just remind you what the usual order of making a MOSFET is just from last time. Usually, you start with a wafer. You do your isolation. This could be your shallow trench isolation we talk about as oxide. Oop, that didn't come out very well.

I'm going to implant the boron profile. This is an n MOSFET. I'm going to implant that early on the super steep retrograde. Grow the gate oxide, so we're growing the gate oxide. Form the polysilicon gate by deposition and etching. The source drain extensions are now implanted.

Now, the source drain extensions bring with them-- remember, we talked about the reverse short channel effect. They're going to introduce a certain number of 311 defects that are going to cause TED of the boron that's underneath here.

Put in the spacers. Now, the spacers, if they're nitride-- a lot of people today are using nitride and not oxide-- if they're nitride and they're LPCVD nitride, that's a pretty high-temperature process, about 800 degrees, one of the worst temperatures you can possibly use for TED.

Why is that? Well, we saw that the time that TED lasts at 800, it can be quite long. And CI over CI star can be quite large. So an hour at 800 can cause a lot of transient-enhanced diffusion. And then after that, typically after you make the spacers, you do this deep source drain implant here for the contact regions. And then you do a final rapid thermal anneal, usually 1,000 to 1,050, maybe 10 to 30 seconds, something in that range.

So the important thermal steps are, this nitride at 900, this nitride thermal budget at 800, and then a rapid thermal anneal at around 1,000. And now we're going to compare with Supreme, how Supreme thinks in 1D sense this would go. So we're going to just look at the effect of TED just on the arsenic diffusion itself. We're not going to look at the effect on the boron diffusivity. That would require a two-dimensional simulation.

So this is a simulation of an arsenic source drain extension implant and diffusing it using the usual order, which is 800 degrees C step for the nitride deposition, followed by a deep source drain, and then followed by 1,000 degrees C 5-second RTA to activate everything.

So here's the arsenic Monte Carlo implant. And you can see on this scale-- this is the distance in microns and the-- they didn't use a very fine grid. Good enough to get this profile, but for the as-implanted, you can see it's kind of ugly looking. And that's part of the reason the grid was a little bit too coarse. But again, with Monte Carlo, we were trying to speed up the processing a little.

So here's 2 keV, very shallow peaks, at 1e15. And then the blue line is 800 degrees, 30 minutes anneal. And then the red line is 800 degrees, 30 minutes, followed by a rapid thermal anneal at 1,000 for five seconds, so 1D simulation.

So almost all the diffusion really takes place at the 800 degree 30 minutes. And that's because of TED. That's not ordinary 800 degrees C. So in order to model this, this must be a PD.full. So it must be taking into account the transit-enhanced diffusion. Well, this is trans. I'm sorry. This is PD.trans. You can either use trans or full. This particular one is PD.trans.

OK, so that's what it looks like, very little contribution of the RTA. And your junction depth here is about 0.1 micron. That's what this predicts. So why is that? If you look on slide 29, this is a plot-- Supreme, in addition to the dopants, you can output the interstitials and vacancies. So here's a normalized concentration versus depth. And the blue line here is for CI over CI star at 800 degrees.

And so you can see, CI over CI star after this anneal, at the 800-degree anneal, that is, it's pretty large. It's somewhere between 20 and 40 and near the surface. It's pretty big. That's the blue line. The red line, which I apologize you can't see. You have to write it on your-- it's the same as your x-axis, basically. The red line is just about one. You can't even see it on the scale.

So CI over CI star at 1,000 degrees, after five seconds is one, which there isn't much enhancement left. And we saw that. Remember, the enhancement time, the amount of time that it lasts is relatively short. So this is after five seconds. So TED is already-- the interstitial concentration-- the 311s have all dissolved. And the interstitial concentration has gone back to CI star. So after five seconds, we still have a lot of enhancement at 800 degrees, and that's why we get all that TED in the 800-degree C profile.

So now, on slide 30, we're going to do something a little different. We're going to activate the source drain extension implant prior to the nitride spacer at 800. So what we're going to do instead of doing the usual put the nitride down right after implanting, we're going to do a 5-second rapid thermal anneal after implanting.

So we send it out for implant. We bring it back. We do a 5-second RTA first, and then we do another five seconds-- and then we do 800 degrees plus 30 minutes. We do that second. So we're changing the order. And you can see the red line now is after the implant plus 5 seconds at 1,000, you diffuse this far.

And then if you add the 830 minutes, it only goes that far. It doesn't go very far because five seconds is 1,000 is really enough to dissolve all the 311s, right? The enhancement-- the TED time at 1,000 degrees is only a couple of seconds.

So I dissolve all the 311s. Then you put it back in the furnace at 800, and you don't get-- this is just normal 800 degrees C diffusion. There's not many CI over CI*s now been reduced because I got rid of all those 311s. So the junction depth now, here, instead of being out here at 0.1, is only 0.07.

So the exact same amount of time that the wafer spent in the RTA and in the furnace is just that the order of the operations was changed. And this is reduced, the junction depth is reduced because CI over CI star is reduced by about 3x during the 1,000-degree C step.

So this is something-- these are all simulations. Supreme can simulate the fact that 311s are generated by the damage and that they dissolve at different temperatures, at different rates. And then so it can take into account these types of effects. Again, it's using the trans model.

The accuracy, of course, you always take with a grain of salt, whether the junction depth is really 0.06, 0.04, I wouldn't put my life on it, to be honest, because there's a lot of parameters in Supreme. But the point is, you can fit your data-- assuming you get some data, you can fit data that shows the effect of the order of the different implants. There's enough physics built into it. You just have to get the right parameters.

So the final process, just to show that what we did is the extension implant followed by 5-second RTA and 1,000 high-temperature nitride spacer at 800 and then another 5-second RTA. The reason I needed to add the last RTA is to activate the deep source drain. Remember, the deep source drain goes in last after the nitride spacer. So this is five seconds at 1,800, five seconds at 1,000. This is what it would look like.

The final five seconds at 1,000 really doesn't contribute much to further motion. It's already done most of its motion because TED, at that point, is completely over. And if you've done it the usual order, 800 degrees, 30-minutes spacer, plus 5 seconds to anneal everything, you get this junction depth. So this the Supreme predicts that reversing the order of the anneals makes a difference. Again, it's something you'd have to check experimentally just to be sure.

OK, so that was a TED example. And then for the fourth example, this will be the only example I'll do where we do two-dimensional, a two-dimensional simulation. And this I took this right from the Supreme manual. So you can run this yourself. This particular example is not one that I made up. I took it out of the example file, and this is on the computer. This example is one of the canned examples that comes with Suprem4.

It's a 200-nanometer gate length, so that's 0.2 micron gate length and MOSFET. It's kind of old fashioned now. But it uses something called self-aligned silicide, and we're going to talk about silicides in the next few lectures. So you'll get an idea.

And I've put the [? input ?] file and columns. So this is the first part of the file, and this is the second part. Again, if you want to see the [? actual ?] file, it's in the Suprem4 directory. So this is TSUPREM-4. So example, here's the mesh using a fairly coarse grid, grid dot fac of 0.9.

You can also define the mesh in different directions, in the y direction, in the x direction, so a little more sophisticated definition of the grid. We start out by growing the gate oxide. So here's 850, 25 minutes. And it's in dry O2 plus a little bit of HCl. So there's an HCl oxidation. And then you can plot the structure to see what it looks like after gate oxidation.

And this says, whenever you see the word source, what that says is-- and then you have the file name that says, take all the commands from that file name and run them now. So t Suprem4, in order to-- let's say you do the same operation in Supreme over and over again. You always do a plot. You want to do several different plots. Rather than putting all those plot commands in the main file over and over again, you create sub files.

And this file called S4EX10P.input is just a series of commands that defines the colors and things like that. So it enables you to call this file whenever you want it to create a plot. So it's a way of cleaning up your Supreme input file. So if you want to know what this particular file does, you have to go look at that command file.

Then you deposit some materials or depositing polysilicon. You tell it the thickness. You're etching the polysilicon to the right and to the left. This is to make a gate. And here you're calling that file, again, that plotting file, OK? Now, one of your homework problems that you just handed in was on-- or we just handed back to you, homework 3, was different methods. This is using the compressed method as far as solving for the oxidation rate goes. Here's 850 oxidation in dry O2 and HCl again.

So this is what we call re-ox. So you formed the gate now. So the initial gate oxidation is just to grow a very thin gate oxide. You formed the gate, and you etched it. And then, in the course of etching, sometimes you introduce damage right near the corners of where the channel is going to be. Remember? This is going to be my channel. So you often introduce some plasma damage down here.

You [INAUDIBLE] up the oxide there. You do things that are not necessarily good for the gate oxide. So people often, at this point, will, in this technology, say, 0.18 micron technology, would take the wafer at this point, put it back in the oxidation furnace, and do what's called a gate re-ox. And this looks like we're doing a re-ox here at 850 for 25 [? minutes. ?]

So that re-ox is going to help-- if you thinned the oxide at all in the course of your etching, it's going to help boost that back up. And it's going to grow a little oxide along the edge here of the polysilicon. So it's a way of dealing-- it was introduced for reliability considerations. Re-ex, as it says, take the gate, put it back in, and do another oxidation.

Then you form the sidewall spacer. Here, we're depositing oxide. It's using an oxidation-- a deposition process. And then you do the deep source drain implant. This is 1e15 energy of 60 keVs, fairly deep. Then you're going to deposit some titanium. I won't go through the details of this.

But in the next few lectures, we'll talk about the fact that-- when we get to the silicide lecture-- you can take metals, put them on silicon, and react them at a certain temperature and form a metallic phase. It's called a silicide. This is titanium disilicide, which forms a good contact. And then we look at the final structure.

So that's the file, and you can go ahead and run that. And let me show you some examples of what comes out of that. Oh, actually, this slide, number 33, remember, I was saying you wanted to repeat this over and over again, these commands? This was that file, S4EX10P.input.

This is just a plotting sequence file that you use over and over again that tells it what colors to use and how to label things according to what material is. So rather than type this in every time in your Suprem4, you can put it in a separate file. That's just for your reference.

So let's go to slide 34. This is from running that example. After gate oxidation, what does it look like? Well, you have an xy structure, so y is in depth. And there's a certain gate oxide grown everywhere across on the structure. That's relatively simple.

After gate patterning, you have deposited this green layer everywhere and then etched it off everywhere to the right of this one line. So now, that's what the polysilicon gate looks like with the gate oxide underneath it after gate patterning, OK?

And notice, it doesn't model etching in any sophisticated way. You just tell it where to cut it off. So it's not modeling any shape effects of the etch or anything like that. That's not that type of a program. We'll talk more about that when we talk about how to model etching.

After gate re-ox, remember, I said we're going to put it back in the furnace and subject it to a step of 850 for 25 minutes to oxidize all around and to beef up this oxide. But interestingly, look what it's done to the gate oxide. You see what it's done. There's like a bird's beaking effect. Some of the oxidant has diffused underneath the polysilicon and diffused it. And you've got a little thicker oxide now here on this part underneath the gate compared to the center of the gate.

That's not necessarily a good thing if it's too thick here because what do we know about the gate oxide thickness determines the threshold voltage in the device. So if I have a thicker gate oxide here than here, I'm going to turn on my channel at different places differently. So that's not necessarily a good thing. We want a uniform gate oxide.

Here, of that same structure now, after I've formed a spacer, here's the oxide spacer. And we did a source drain extension. I'm sorry. I first did the 5e13 source drain extension. Then you formed the spacer, and then you did the deep source drain. These contours, by the way, correspond to arsenic in the substrate. Each one is a different arsenic concentration. You can see how it's shaped. This region here corresponds to the extension, This. Region to the deep source drain. This particular simulation uses PD.Fermi, so the Fermi model, very simple.

And then after silicidation, Supreme doesn't necessarily go to a very good model of siliciding, but you see it's reacted. We'll talk in subsequent lectures about what silicidation is all about. I mostly want you to get a feel for how a two-dimensional simulation of such a structure actually looks.

There's the final structure here on slide 39, showing the arsenic concentration contours. So here, this region in between is your channel. Here's the polysilicon gate in green, and this is your gate oxide. And because of the reoxidation, we have a little bit of the smile effect. And see the way the device seems to be smiling, has a little smile here? That's actually a very happy device.

But it's not actually all that good. Smiling is not very good. In fact, you want a gate oxide that's perfectly flat, has the same concentration-- same thickness as much as possible all the way across it. It got unflat because we did the re-ox, and a little bit of oxidation took place from the corners.

So these days, for the very shallow, the very shortest channel devices, reoxidation is not such a popular thing to do because it does create this nonuniformity in the oxide thickness. In fact, you can go back. What I did was I went back here on page 40 and said, oh, OK, is there some way we can deal with this and maybe make the re-ox have a little bit less effect on the-- a little less of this extra oxidation right at the corner?

So this is really what we're talking about just to give you a zoom in on slide 40. Here, actually, here's a zoom in of what it is. This nonuniform oxide thickness took place under the gate. And the real origin is the lateral oxidation under the poly during the gate re-ox step. This is when the gate re-ox was done at 850 for 25 minutes.

So in fact, from the very first lecture of class, this is a real TDM of a real device to show you this effect is real. See the way this device is smiling also? So Supreme didn't just make that up. The oxidant did get under here and oxidize in this region. So the vt is going to be nonuniform across this device. So you have to actually watch out for this. This does really happen, just to show you that Supreme is based in reality with respect to some of these things.

On slide 43, actually did a little example of how we can change the amount of this effect that happens, the amount of this nonuniformity. It may not be very clear if I'm looking at this. But on the left-hand side, what was done was, when we did the etch, we etched all the way down, OK, and then did a re-ox.

So on the left-hand side, you're etching like this. And here's your gate oxide like this. And we're basically etching in the example where we told Supreme we act as if we etched all the way down. So the oxide-- the prior to re-ox.

So when we put it in the furnace on the left-hand side-- oh, sorry-- when we put it in the furnace for oxidation, this oxide was gone at that corner. And that's what you end up with on the left-hand side. So you can imagine, when that oxide is gone, then you can get a fair amount of attack in here laterally of the oxidant getting underneath there.

So what was done here on the right-hand side instead, 50 of the 70 angstroms were etched. So we left a lot of the oxide on there and then just oxidized it-- re-ox was done at a lower temperature, 825 for a shorter time, 12 minutes. So you get some of the benefits of the re-ox without some of that extra.

So in this example, when we did the etching, it didn't remove it all. It looked like this when it went into the oxidation furnace. It looked like that and then got oxidized. So, by refining the etch process, we didn't etch all the way down. And if you look carefully, I think on the next slide, it'll be maybe a little more obvious.

The lateral nonuniformity is quite a bit reduced. This oxide thickness is much more uniform going across here. There's still a little bit of it, but 825 for 12 minutes, there's a lot less of that lateral nonuniformity. So it's an example of a process-- you can use Suprem4 to optimize reasonably efficiently and reasonably accurately in this two-dimensional model.

OK, so let me go on to summarize. The simulators, like Suprem4, which, by the way, I didn't tell you, but Suprem4 was originally written at Stanford. And sometimes that's called S Supreme. That's where it came out of originally in the 1980s. It was then commercialized in a small company called TMA. That company was bought out by another company called Avant. And Avant later sold the technology to a company called Synopsis. Synopsis is a big design house. They make a lot of CAD software for designing chips. But they also support TSUPREM-4.

So if you want to have questions about the simulator, you need to contact people at this company called Synopsis. There are other simulators out there. Suprem4 is one of the most popular, but there's another company called Silvaco, which makes a competing product, which is called Athena.

In any case, these simulators like this, they've been developed over the years to enable what we call physically accurate or robust or correct simulations of complex processes. However, I've tried to tell you to take everything with a grain of salt. Don't believe it just because you simulate it in Supreme and it looks like that.

One thing, you might have done it wrong. You might have used the wrong grid factor. Supreme itself has parameters in it that are unknown, that somebody just stuck in there. Some graduate student writing the program said, oh, I don't know exactly what this parameter is. I'll put this number in. It's rough.

Well, you need to find out what parameter is in there. And if you go to the appendix in the manual, it'll tell you in general what all the numbers are that it's using. And you can decide whether those numbers-- whether you like them or not.

The nice thing about these simulations from Suprem4 or Athena, you can feed them into a device simulator, such as Medici or whatever, and then predict IV and CV characteristics, the actual electrical characteristics of the device. So they are designed to be coupled, the process simulator and the device simulator. And that's a very nice feature.

There's a lot of new understanding that's being developed and issues related to TED and OED and other anomalous effects, low energy ion implants. I showed you some low-energy phosphorus where Supreme doesn't do a very good job of modeling the data. Highly tilted implants with shadowing, oxidation of trenches.

Supreme can handle most of these situations because it has the physics built in. Exactly the accuracy, that's always the question mark. Whenever you're running these simulators, keep in mind the basic physical models. There are a lot of parameters that must be known accurately, and we don't know them all accurately. So keep your eyes open.

A big caveat is keep the mesh or the grid in mind. When your grid is fine enough, you should be able to get the exact same solution with minor changes to the grid. So I run it for a certain grid. I cut the grid spacing in half. The solution should look identical. If all of a sudden the solution looks much smoother when I cut the grid in half, the spacing or much finer, then obviously the original solution didn't have a fine enough grid.

So you've got to do these sanity checks. Run it once with a certain grid. Then use twice the grid points. Does it look a lot smoother? Well, that means you probably didn't do a good enough job the first time. And then do it again. And eventually, it'll converge. It has about the same smoothness regardless of your grid. In which case, your grid is probably fine enough. Your solution should be independent of the grid. Otherwise, that's not physically realistic.

There's a lot of different process integration schemes. I just showed you one, just changing the order of the device, of the anneals. The simulators help us to understand these interaction between the various steps and how to optimize the overall technology in ways that you could never do if you just did this by hand.

So you've used Suprem4 now. You're going to use it again in homework number 5. So I think gives you some idea of how powerful this tool is. But if you are going to really use it in your research, make sure you read the manual. It is very important because there's a lot of little things in there you need to know about.

OK, that's it for today. Homework number 4, please bring up front. Homework number 3 is in the orange folder in the back by the TA. And I guess that's it. We'll meet on-- oh, somebody have the sign up sheet for your final project? Oh, great. Thanks. Make sure you sign up.