State of the Art 3D SiC Process and Device Simulation

Introduction

Silicon has long been the semiconductor of choice for high-voltage power electronics applications [1,2]. However, wide-bandgap semiconductors such as SiC have begun to attract attention because they are projected to have much better performance than silicon [3-7]. In comparison with silicon, wide-bandgap semiconductors offer a lower intrinsic carrier concentration (10 to 35 orders of magnitude), a higher electric breakdown field (4 to 20 times), a higher thermal conductivity (3 to 13 times), and a faster saturated electron drift velocity (2 to 2.5 times).

SiC exists in over 200 polytypes with different crystal structures with the same stoichiometry. However, only the 6H- and 4H-SiC polytypes are available commercially as both bulk wafers and custom epitaxial layers. Both of these polytypes have a hexagonal crystal structure and a bandgap in the neighborhood of 3eV. Of the two, 4H-SiC is now preferable due to the more isotropic nature of many of its electrical properties.

SiC 2D and 3D Simulations

Simulating SiC devices is more challenging compared to Silicon. Indeed very low intrinsic concentration combined with high doping values is usually detrimental to convergence. Silvaco’s 3D Process and Device solution can handle this situation thanks to state of the art mesh, discretization and solvers. In the following paragraphs we will explain what are the ingredients that we think are essential to accurately simulate these type of devices in 3D. Everything mentioned in these paragraph is implemented in Silvaco’s 3D Process and Device framework called Victory.

Mesh and Discretization

For faster and more accurate device simulation, it is best to start out with a mesh that has a perfect geometric dual and that dual is the Voronoi tessellation of the vertices. This means that the perpendicular bisectors of each tetrahedral mesh edge form a convex finite volume which does not overlap those of its neighbors. It is equivalent to saying the mesh is Delaunay (i.e. the interior of the circumsphere of each tetrahedron does not contain any vertices). The finite volume boxes should also not overlap any region or domain boundaries. When these two conditions are met, the finite volume discretization is relatively straightforward, and can be done without introducing extra cross coupling or discretization errors (which would degrade the convergence or accuracy). One way to achieve this is to simply create a finite difference style grid with Cartesian grid lines extending all the way through the domain. Unfortunately this tends to result in a relatively large number of grid points, an even larger Jacobi matrix, and a slow simulation (especially in 3D). Recent advances at Silvaco in mesh generation however have resulted in algorithms which can create and refine meshes whose finite volumes have these two important properties, without introducing excessive numbers of grid points. An example of Delaunay mesh is shown in Figure 1.

At the end of the process simulation, the device meshing algorithm creates a pure unstructured Delaunay remesh of the results of the process simulation. A Delaunay mesh of a particular set of points has certain special mathematical properties compared to any other mesh of the same points. One example of this in 3D is that the largest radius of the smallest sphere enclosing any tetrahedron is minimized for a Delaunay mesh; this leads to a high-quality mesh of compact elements.

Continued on page 2...
The Delaunay device mesh is conformal in the sense that a point sampling of the features of the process mesh is produced whose Delaunay mesh resolves those features. These features can be viewed as zero, one, two or three-dimensional and are defined recursively; these are the fundamental vertices, edges, polygons and polyhedra composing the structure. The conforming point sampling has the special property that the corresponding 3D Delaunay mesh has lower dimensional Delaunay meshes embedded in it; this property does not hold in general for arbitrary sets of 3D points.

The algorithm used to calculate the conforming point set has two main improvements over those commonly seen in the literature:

- The number of points required to generate the conforming point set is often considerably less than standard approaches. This allows fewer elements and finer control of those elements by breaking the usual relationship between local feature size within the structure and the element size in that region. This lower number of elements leads to decreased device simulation time.
- Features with acute angles are handled in a clean, general way compared to standard methods. This decreases the complexity of the implementation and represents the fundamentals of a fully general nD conformal Delaunay algorithm.

The conformal Delaunay algorithm also supports a general framework for producing arbitrarily locally refined meshes. Refinement schemes are implemented as plug-ins outside of the core meshing code itself and any number of these can be applied in an arbitrary order to a given mesh. A number of TCAD-specific approaches to refinement have been implemented; these include steadily decreasing element size within semiconductor regions according to distance to the insulator region or the junction, refinement on the variation of net or total doping. All these refinements are fully automatic.

**Solver**

Mesh and discretization are only part of the overall problem. Fast and accurate simulation will also depend on the solver used during the simulation. A new solver have been developed to best fit the work done on the mesh and discretization.

The PAM solver is a domain decomposition type linear system’s solver specifically designed for very large sparse linear systems. PAM solver is based on Yousef Saad’s Parallel Algebraic Recursive Multilevel Solver pARMS [8]. The parallelization is done with MPI (Message Passing Interface). Each MPI process handles the solution of part of the linear system and the MPI processes run in parallel. After each MPI process finishes with its part of the linear system, each solution is broadcast back to the main MPI process, and the solution to the global linear system is formed and returned.

The main advantage of the PAM solver is the fact that the domain decomposition approach leads to fast solution of smaller sized linear systems which is carried out in parallel, thus significantly reducing the total solution time for very large sparse linear systems. It is also an accurate and robust solver.

Another advantage is that it is a very diverse solver allowing the use of several combinations of global and local preconditioners and two iterative solvers. It can be used for a wide variety of linear systems with both non-symmetric and symmetric matrices arising from different types of applications. It also supports extended precision; all computations can be done in 64, 80, 128, 160, 256 bit precision.
Extended Precision

In a device simulation, the calculation of the electrostatic potential depends on the space charge density, which includes the difference \((n-p)\) of the electron and hole concentrations. Simultaneously, the calculation of current conservation implicitly depends on the sum \((n+p)\) of these concentrations. Consequently, the convergence and accuracy of a device simulation requires the numerical resolution of the quantities \((n-p)\) and \((n+p)\).

To resolve these quantities, the calculations must be performed with at least

\[
P \geq \frac{\ln n - \ln p}{\ln 2}
\]

significant bits of arithmetic precision. From the theory of carrier statistics, we can then estimate that for a convergent and accurate simulation, we need to maintain a precision of:

\[
P \geq \frac{1}{\ln 2} \left( \frac{E_g}{kT} \right).
\]

Therefore, the required precision \((P)\) depends on both the lattice temperature \((T)\) and on the bandgap \((E_g)\). In practice, the above inequality constitutes a loose upper bound, and we can frequently carry out a successful simulation with somewhat less precision than the relation indicates. This inequality also tells us that precision is more likely to be an issue with high-bandgap materials, and at low temperatures.

One way to skirt this issue is to consider only the majority carriers when solving the semiconductor equations. However, that option is unavailable if recombination is important, because recombination by definition involves both carrier types. In such a case, the most general solution is to carry out the simulation at a high level of precision.

When we discuss the precision of computer arithmetic, it helps to understand a little about the structure of floating-point numbers. Floating-point numbers are composed of three parts: the sign, the exponent, and the significand[9]. As its name suggests, the size of the significand corresponds to the number of significant bits in the number. In the above inequalities, \(P\) stands for the number of significant bits. The nominal precision of a floating point number is larger than the number of significant bits because it also counts the bits used for the sign and the exponent. In Victory Device and in Atlas, the various flags controlling the precision levels of the simulator and of the solver always refer to the nominal precision.

By default, Victory Device and Atlas use a nominal arithmetic precision of 64 bits. In order to support the simulation of high-bandgap materials in particular, these simulators can now run at several levels of extended precision as well. The level of precision is selected by means of a command-line flag (which in DeckBuild would be passed as one of the arguments to the SIMFLAGS param-

<table>
<thead>
<tr>
<th>Command-line flag</th>
<th>(None)</th>
<th>-80</th>
<th>-128</th>
<th>-160</th>
<th>-256</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal precision (bits)</td>
<td>64</td>
<td>80</td>
<td>128</td>
<td>160</td>
<td>256</td>
</tr>
<tr>
<td>Significant bits</td>
<td>53</td>
<td>64</td>
<td>104</td>
<td>128</td>
<td>209</td>
</tr>
</tbody>
</table>

Table 1. Arithmetic Precision Levels in Victory Device and Atlas.

For well-converged solutions, the run-time increases with the precision. This increase is especially significant at the highest precision levels. A 256-bit solution may take twenty times longer than a 128-bit one! For example, a simulation that consumed 10 hours using 256-bit arithmetic might take just a half-hour using 128-bit arithmetic. Memory requirements also increase with the precision. On the other hand, certain simulations that have difficulty converging at the lower precision levels are likely to run faster if the precision level is increased. Accordingly, the optimum precision level for a particular problem is likely to be the minimum level that yields good convergence during the solution.

4H-SiC has a bandgap of 3.26 eV. The second inequality above suggests that a simulation of a device composed of this material may require a precision as high as 185 bits. From the table, we can see that to get at least 185 bits of precision we must resort to 256-bit arithmetic. Now, 185 bits is a rough upper limit, so in some cases we may be able to simulate SiC devices using only 128-bit arithmetic without doing anything special, even though the precision available from 128-bit arithmetic is only 104 bits. In general, however, we can expect to either have to employ higher-precision arithmetic, or to take steps to reduce the precision requirement.

One way to reduce the precision requirement is to limit the intrinsic concentration affecting Shockley–Read–Hall recombination to at least a minimum value \((\text{NI, MIN})\). In heavy-doped regions, we can estimate the value of \(\text{NI, MIN}\) necessary to bring the spread in carrier concentrations within the scope of the available precision \(P\):

\[
\text{NI, MIN} = \frac{N_{\text{inf}}}{\sqrt{P}},
\]

where \(N_{\text{inf}}\) is the net doping. However, there is no reason to employ this limit if it is less than the nominal value of the intrinsic concentration,

\[
n_\text{i} = \sqrt{N_c N_v} \exp \left( \frac{E_g}{kT} \right),
\]

where \(N_c\) and \(N_v\) are the densities-of-states at the conduction and valence band edges.
For SiC with \( N_{net} = 1e19 \) and \( P = 104 \), the above estimate works out to \( \text{NI.MIN} = 2220 \). Practical experience has been that one can often get by with values of \( \text{NI.MIN} \) much less than this.

Raising \( \text{NI.MIN} \) of course will tend to raise the leakage current. For many typical high-bandgap structures, however, the calculated leakage current is so low that it is non-physical. Since electrons and holes are quantum particles, we can only have an integer number of them. Thus a calculated leakage current of \( 1e-27 \text{ A} \) corresponds to an average of one carrier crossing the device every five years. Most of the time, in this case, there will be zero current. If we increase \( \text{NI.MIN} \) such that the leakage current here is raised by a factor of a million, to \( 1e-21 \text{ A} \), then that corresponds to an average of one carrier crossing the device every three minutes. Again, we can say that most of the time there is zero current.

On the other hand, raising \( \text{NI.MIN} \) has little effect on the breakdown voltage calculated for a device. Consequently, if we are interested in determining the breakdown voltage, setting \( \text{NI.MIN} \) to a judiciously chosen value can speed up the calculation significantly.

Here is an example illustrating these points.

Figure 2 shows the top section of a SiC power device. This device has a long drift region about 160 µm in extent, with low doping, giving the device a high breakdown voltage. The width of the device is here assumed to be 1 µm. There are two contacts at the top of the device, a gate to the left, and a drain contact at the bottom.

To determine the breakdown voltage of this device, we perform a simulation in which the bias on the drain is gradually raised until breakdown occurs. In this simulation, we use the model of Hatakeyama et al [10] for anisotropic impact ionization, with the \(<0001>\) axis of the crystal oriented so that it is parallel to the long axis of the device. We also use the Shockley–Read–Hall [11, 12] model for recombination.

Figure 3 shows the results of this simulation, performed using various levels of arithmetic precision and various values for \( \text{NI.MIN} \). As predicted for this material, 256-bit arithmetic is required when \( \text{NI.MIN}=0 \). With higher values of \( \text{NI.MIN} \), we can reduce the precision of the arithmetic yet still obtain a converged solution. At \( \text{NI.MIN}=1e2 \) we can use 128-bit arithmetic, and the simulation takes only 5\% of the time needed for a 256-bit calculation.

From \( \text{NI.MIN}=0 \) to \( \text{NI.MIN}=1e2 \), the voltage predicted for the onset of breakdown rises from 14000 to 14080 V, a difference of just 0.6\%. The predictions for the leakage current...
vary greatly, but even the highest value of 1e−18 A is too small to be measured in a real-world device. In real-world terms, then, the predicted leakage current is effectively zero and the predicted breakdown voltage is 1.40e4 V.

Applications

In this section we will review and show 2D and 3D simulations of different types of structures using 3D Process and Device simulators. We will start to compare 2D and 3D breakdown voltage simulation of an identical structure in 2D and 3D for validation purposes. We will then show the simulation results of a Trench MOSFET. This particular structure will have a rounded bottom trench as well as a floating P region underneath the trench for improved blocking capability and thus increased breakdown voltage. We will then show also on a Trench MOSFET how we can optimize the breakdown voltage as a function of the trench shape. Finally, IGBT simulation exhibiting very high breakdown voltage will be demonstrated.

In order to create the 3D structure we used our 3D Process simulator, very suitable for 3D SiC power devices simulation since it is layout driven, accurate, fast and easy to use. After the process simulation is done a 3D structure is saved using a 3D tetrahedron mesh to ensure that any shape created during 3D process simulation is well conserved and transferred to the 3D Device simulator.

Comparison of 2D and 3D Simulation Using the Same Structure

The following example shows a comparison between 2D and 3D breakdown voltage simulation of a vertical SiC MOSFET.

Hatakeyama, et al. [10] have developed an anisotropic impact ionization model for materials with a hexagonal crystal structure, including 4H-SiC. The ionization coefficients in this model depend on the strength of the electric field and on the orientation of the field with respect to the optical axis of the crystal.

You can specify the orientation of the optical axis of the crystal in Victory Device by supplying values for ZETA and THETA parameters on a MATERIAL statement. ZETA defines the angle the optical axis makes with the y coordinate axis. THETA defines the angle that the projection of the optical axis onto the x-z plane makes with respect to the x-axis (measured clockwise around the y axis). The defaults are zeta=0 degrees and theta=0 degrees, making the <0001> direction parallel to the y-axis.

To make the optical axis parallel to the z-axis for 3D simulation so that the optical axis points toward the SiC surface for both 2D and 3D, we set ZETA=90 degrees and THETA=90 degrees on the MATERIAL statement.

Figure 4 shows the Net Doping in the 3D structure after process simulation. As shown in Figure 5, and as expected, the breakdown voltage is the same in 2D and 3D. Figure 6 shows the same impact ionization rate and location in the 2D structure and in a cutplane from the 3D structure, validating that breakdown voltage is the same for both 2D and 3D structures.

Comparison of 2D and 3D Simulation Using the Different Structure

The following example shows 3D simulation of a vertical DMOS where the bottom of the trench is rounded and with a floating P-type region under the trench to increase the breakdown voltage.
Because the structure in this case is 3D it is interesting to notice that doing 2D simulations will not allow accurate calculation of the breakdown voltage. Multiple 2D simulations were performed: one along the side of the device, another one along the diagonal, and finally one in between. All of them exhibit different breakdown voltage as shown in Figure 7, and also different from breakdown voltage from the 3D structure shown in Figure 8.

If we imagine the 2D structure as possessing a cylindrical symmetry about the $X=0$ axis, the area available for the lateral expansion of the depletion region is invariant under rotation about this axis. This will result in the same breakdown voltage in 2D and 3D. On the contrary, the 3D structure in this example does not possess such a cylindrical symmetry. In fact, the depletion region of the 3D structure can expand along the diagonal direction further than along the X and the Y direction. This leads to a higher breakdown voltage and has been verified by performing multiple 2D simulations.

In summary, the 3D structure provides a larger area for lateral depletion than the 2D one, and therefore exhibits a higher blocking capability. Therefore it has to be simulated in 3D.

Simulation results showing Electric Field, Impact Ionization Rate and Electron Current Density are shown respectively in Figures 9, 10, and 11.

It is interesting to notice that due to the presence of the P-type region underneath the trench, making a p-n junction, the maximum Electric Field and thus impact ionization rate is maximum at this junction location and not at the trench corner.
Breakdown Voltage Optimization Versus Trench Shape

The following example demonstrates the effect of the shape of a trench on IV and BV characteristics.

In this example we compare 3 different structures. One fully Manhattan (i.e. 90 degree layout and trench) versus 2 structures one having rounded trench edge and another one having rounded edge and angled trench.

In Figure 13, we clearly see the effect of the trench shape on BV characteristics. BV simulation reveals that the breakdown voltage can be increased by 30% using rounded trench edge and angled trench due to impact ionization occurring not anymore only at the corner of the trench as can be seen in Figures 16, 17 and 18.

In this example we did not focus only on breakdown voltage simulation but we also have simulated static performance of the structure. IdVg simulations shown in Figure 12 illustrate a 3D specific effect. Indeed a “hump” effect is observed on the IdVg characteristic as reported in [13]. This parasitic transistor results from current crowding at the edge of the trench as shown in Figures 14 and 15. In Figure 14 we can see current flowing along the edge of the trench when this trench is sharp, whereas we do not have any current when the trench rounded and angled. This “hump” effect is characterized in the IDVG curve as a parasitic transistor. This may be problematic in term of power consumption since it increases the off-current. To suppress this parasitic effect a non-Manhattan structure is used (including non 90 degree layout as well as angled trench).

In Figure 13, we clearly see the effect of the trench shape on BV characteristics. BV simulation reveals that the breakdown voltage can be increased by 30% using rounded trench edge and angled trench due to impact ionization occurring not anymore only at the corner of the trench as can be seen in Figures 16, 17 and 18.
Very High IGBT Breakdown Voltage Simulation

The following example demonstrates a 3D trench SiC IGBT simulation.

The particularity of this device is to have low doping long drift region of about 160um. This will lead to high breakdown voltage around 12KV. As expected and as mentioned earlier BV is affected by the shape of the trench. Maximum BV is obtained when the trench edge is rounded as shown in Figure 19 and 20.

Conclusion

Silicon based power devices are still dominant today in power electronics. However, wide bandgap semiconductors like SiC are now more and more used for high power, high-temperature applications because of superior thermal conductivity, lower intrinsic carrier concentra-

tion and better on-resistance compared to Silicon, which is a key figure of merit in power switching applications.

Silvaco anticipated the simulation needs years ago and is now able to provide a complete flow from process to device in 2D and 3D. The flow is now ready to adress the increasing needs of accurate SiC based devices simulation, to develop and optimize this promising technology.
Acknowledgement

The authors would like to express their gratitude to the members of the Silicon Carbide Group of AIST for supplying valuable advices validating our Process and Device flow.

References

cal Modeling and Scaling Properties of 4H-SiC Power De