Silicon power MOSFET at low temperatures: A two-dimensional computer simulation study

Hua Ye a,*, Changwoo Lee a, James Raynolds a, Pradeep Haldar a, Michael J. Hennessy b, Eduard K. Mueller b

a College of Nanoscale Science and Engineering, State University of New York at Albany, Albany, NY 12203, USA
b MTECH Laboratories, LLC, Ballston Spa, NY 12020, USA

Received 6 July 2006; received in revised form 25 January 2007; accepted 30 January 2007

Abstract

Understanding how the structure of the unit-cell affects the cryogenic performance of a Si power Metal Oxide Semiconductor Field Effect Transistor (MOSFET) is an important step toward optimizing of the device for cryogenic operations. In this paper, numerical simulations of the Si power Double Diffused MOSFET (DMOS) are performed at room temperature and cryogenic temperatures. Physically based models for temperature dependent silicon properties are employed in the simulations. The performances of power DMOS with various unit-cell structures are compared at both room temperature and low temperatures. The effect of the cell structure on the on-resistance and breakdown voltage of the device are analyzed. The simulation results suggest that the device optimized for room temperature operation can be further optimized at cryogenic temperatures.

Keywords: Cryogenic power electronics; Power MOSFET; Numerical simulation

1. Introduction

Extensive research has been focused on development of cryogenic power electronics system recently [1–11]. The interest of investigating cryogenic power electronics comes from the idea of building power conditioning systems with much higher efficiency by cooling power devices down to cryogenic temperatures and by combining them with High Temperature Superconductor (HTS) components in order to gain synergistic benefits [12]. The development of superconducting hybrid power electronics system is becoming an enabler in future military system where substantial electric power will be needed to operate combatant ships, electric air aircraft, and directed energy weapons [13]. Cryogenic power electronics is of interest to deep space exploration where low temperature is the norm than the exception.

Due to the better electronic, electrical, and thermal properties of certain semiconductor materials at low temperatures [14–16], cryogenic power electronics is expected to have better efficiency, higher speed, reduced leakage current, and reduced latch-up susceptibility [15–17]. For instance, as Jackson et al. argued, it was not likely that existing devices were truly optimized for cryogenic environment and developments of optimized cryogenic devices were required [2].

Numerical simulation study of the Si power MOSFET’s at cryogenic temperatures can be used as an important step toward optimizing power electronics for cryogenic operations. Very few numerical simulations of Si power MOSFET under cryogenic temperatures have been reported in the published literature [21,22]. In this paper, we compare...
the performances of power DMOS' with various unit-cell structures at both room temperature and low temperatures using two-dimensional numerical simulations. The dependence of the cell structure on the threshold voltage, the breakdown voltage, and the on-resistance of the device are analyzed. The simulation results suggest that the device optimized for room temperature operation can be further optimized at cryogenic temperatures.

2. Device simulation description

Power DMOS' are vertical conduction devices which comprise a repetitive array of unit-cells arranged in a topological layout (such as linear, square array, circular array, atomic lattice layout, and hexagonal array). Therefore, it is necessary to simulate the performance of the elementary unit-cell. Fig. 1a and b show the baseline unit-cell of a medium-voltage DMOS and the half unit-cell used in the numerical simulation due to symmetry, respectively. This information was obtained from first-order analytical calculations. In the baseline unit-cell, the p-base depth was 4 \textmu m and n+ source depth was 1 \textmu m, yielding a channel length of 2.4 \textmu m (with lateral diffusion as 80\% of the vertical depth). Gaussian profiles have been assumed for n+ source and p-base with the peak concentrations of $3 \times 10^{17}$ and $1 \times 10^{19}$ cm$^{-3}$, respectively. The starting epi-layer thickness was 52 \textmu m and the doping concentration was $3 \times 10^{14}$ cm$^{-3}$. The n+ substrate had a concentration of $1 \times 10^{19}$ cm$^{-3}$ with the actual thickness of n+ substrate was 300 \textmu m. However, in the simulation, a thickness of 20 \textmu m was used in order to save the computational time. Due to very high doping concentration in the n+ substrate, which behaves like a conductor, this approximation had minimum effect on the simulation results. The gate oxide thickness was assumed to be 100 nm with a gate-length of 12 \textmu m in the baseline unit-cell. In this study, the topological design was assumed to be circular cell in square array. Cylindrical symmetry was assumed in simulating the half unit-cell, which gave the characteristic current values for a circular unit-cell. In order to give the current density (A/cm$^2$), the unit-cell current values will be multiplied by the number of unit-cells that can be accommodated in 1 cm$^2$ area (e.g., for the circular baseline unit-cell with a diameter of 30 \textmu m, a multiplier of $1.11 \times 10^5$ will be used).

When a positive voltage is applied at the gate electrode, a conductive path (inversion channel) is created across the p-base region underneath the gate. The total on-resistance ($R_{on}$) of a power DMOS is determined by all the resistive components as given by [23]:

$$R_{on} = R_{Source} + R_{CH} + R_A + R_{JFET} + R_D + R_{Sub}$$

where $R_{Source}$ is the n+ source resistance, $R_{CH}$ is the channel resistance, $R_A$ is the accumulation layer resistance, $R_{JFET}$ is the contribution from the drift region between the p-base regions, $R_D$ is the drift region resistance, and $R_{Sub}$ is the substrate resistance. However, due to the high doping concentrations in the source and substrate regions, $R_{Source}$ and $R_{Sub}$ are negligible for high voltage power devices [24].

In this simulation work, the physically-based two and three-dimensional device simulation software ATLAS [25] has been used. The operation of the semiconductor device is modeled by a set of Partial Differential Equations (PDEs) with appropriate models for various temperature dependent semiconductor properties. ATLAS uses Box Integration Method to solve these simultaneous PDEs numerically on a discrete grid of points which is referred to as mesh points. All the simulations in this study have been performed in two dimensions. A variable mesh spacing has been used by refining the mesh in the critical regions (such as in the channel and junctions regions) in order to increase the accuracy without overly lengthening the computational time. Cylindrical symmetry was also employed for this purpose.

The drift-diffusion models for the steady-state are based on the three PDEs:

\begin{align*}
    &R_{Source} + R_{CH} + R_A + R_{JFET} + R_D + R_{Sub} \\
    &\text{(a) Cross-sectional view of the DMOS unit-cell showing the dimensions and impurity concentrations and (b) half unit-cell used in the simulation.}
\end{align*}
Poisson’s equation: \[ \nabla \cdot (\varepsilon \nabla \Psi) = q(n - p - N_h^0 + N_A^0) \]  
Electron continuity equation: \[ \nabla \cdot (\vec{J}_n) = qR \]  
Hole continuity equation: \[ \nabla \cdot (\vec{J}_p) = -qR \]  
where \( \Psi \) is the electrical potential, \( n \) and \( p \) are the concentrations of electrons and holes, respectively, \( D_n \), \( D_p \), \( \mu_n \), \( \mu_p \) are diffusivity and mobility of electrons and holes, respectively, \( \varepsilon \) is the electric permittivity of the material, \( q \) is electron charge, \( R \) is the net electron–hole recombination rate, \( J_n \) and \( J_p \) are electron and hole current density, respectively, \( N_h^0 \) and \( N_A^0 \) are the concentration of the ionized donor and ionized acceptor.

In order to model the device characteristics at low temperatures, incomplete ionization and Fermi–Dirac statistics are employed in the simulations. At room temperature, impurities are usually assumed to be fully ionized. However, impurity freeze-out effects kick in at low temperatures. Moderately doped semiconductors are modeled using Fermi–Dirac statistics with appropriate degeneracy factors for conduction and valence bands [26]. For degenerately \( (>3 \times 10^{18} \text{ cm}^{-3}) \) doped semiconductors, complete ionization is assumed [25]. The use of Boltzmann statistics is normally justified in semiconductor device theory and simple to be implement in the analysis; however, in order to more accurately simulate the low temperature performance of semiconductors, Fermi–Dirac statistics is used in this study. The effective density of states in silicon are considered temperature dependent [25]. In the simulations, the temperature dependence of the bandgap energy for silicon is considered and taken from Sze [27]. Bandgap narrowing was not considered in this study since it has little effect on the performance of power DMOS at low temperatures.

The temperature dependence of carrier mobility is the most important factor that affects the performance of power DMOSs at low temperatures. At lower temperatures, the vibration of the lattice is reduced resulting in less frequent scattering from the lattice. Therefore, carrier mobility increases as temperature decreases. In order to accurately model carrier mobility, various scattering mechanisms need to be considered. For low field bulk mobility, Klaassen’s [28,29] unified mobility model is used in this study. Klaassen’s model considers the effects of lattice scattering, impurity scattering, carrier–carrier scattering, and impurity clustering effects at high concentration. The model agrees well with the empirical data for a wide range of donor and acceptor concentrations with temperature dependence over the range of 70–500 K. For the mobility degradation within the inversion layers, Lombardi et al.’s [30] model is used to include the effects of surface acoustic phonon scattering and surface roughness scattering. In our simulation, the Klaassen model is used to calculate bulk mobility instead of Lombardi et al.’s original model. For high field velocity saturation, Caughey and Thomas’s model [31] is used to calculate the carrier mobility. The saturation velocities are considered as temperature-dependent [32], where the saturation velocity of silicon increases about 20% when cooled from 300 K to 77 K.

Shockley–Read–Hall (SRH) and Auger recombination were considered in this simulation work. SRH recombination accounts for phonon transitions in the presence of a trap (or defect) within the bandgap of the semiconductor. The SRH carrier lifetime is considered to be concentration and temperature-dependent as proposed by Klaassen [28]. Auger recombination describes a three-particle transition whereby a mobile carrier is either captured or emitted [33]. When the reverse bias is sufficiently large, the electric field in the space charge region will be high enough to accelerate free carriers up to sufficiently high kinetic energy to cause avalanche breakdown. At lower temperatures the mean free path of carriers increases, giving them more energy for a given electric field prior to collisions, thus resulting in a reduced avalanche breakdown voltage. The phenomenological local electric field model proposed by Selberherr [33] was employed in this study. Numerical simulation of breakdown at low temperatures can be difficult because the depletion regions have too few carriers to start an avalanche due to the extremely low intrinsic carrier concentration at 77 K. In order to overcome this numerical difficulty, a small beam of light was artificially shined on the top surface of the DMOS to provide carriers which can start the avalanche when conditions are right [25]. This technique is routinely practiced even at room temperature where extremely low intrinsic carrier concentration causes numerical convergence problems such as simulation studies in certain SiC based devices. The wavelength of this artificial light was 800 nm and the beam intensity was 0.001 Watt/cm² in the simulations. We checked the effects of different beam intensities on simulation results in order to make sure that this artificial light did not affect the accuracy of the breakdown voltage.

3. Simulation results and discussions

3.1. Baseline unit-cell

Steady-state \( I–C \) characteristics, threshold characteristics, and breakdown characteristics of baseline power DMOS unit-cell have been simulated as shown in Fig. 2. It is shown in Fig. 2a and b that the on-resistance of the power DMOS unit-cell decreases with temperature. This is due to the higher electron mobility at lower temperatures. In Fig. 2b, the specific current \( (A/cm²) \) is calculated by multiplying the drain current of the unit-cell by the number of unit-cells that can be accommodated in a 1 cm² area. The specific on-resistance \( (\Omega \text{ cm}²) \) is calculated by dividing the drain–source voltage with the corresponding specific current. Fig. 2c shows that, at 77 K, the DMOS has higher threshold voltage (from 4.75 V at 300 K to 6.0 V at 77 K), transconductance, and current saturation level.
The intrinsic carrier concentration decreases 30 orders of magnitude at 77 K resulting in increased band-bending in the p-base region, which needs to be compensated by a higher gate voltage in order to inverse the channel [34].

The higher transconductance at 77 K is due to the increased channel mobility whereas the higher current saturation level is due to the higher electron mobility in all the device regions. Fig. 2d shows that the blocking voltage of the baseline unit-cell decreases significantly from 660 V at 300 K to 510 V at 77 K, a 23% reduction. At lower temperatures, the mean free path of the carrier increases giving free carrier more energy for a given electric field prior to collision, thus resulting in a reduced avalanche breakdown voltage.

Although the on-resistance of the baseline unit-cell shows moderate improvement at 77 K, it is far less than the improvement we expected. As shown in Fig. 2a, the drain–source current of the baseline unit-cell decreases significantly from 660 V at 300 K to 510 V at 77 K, a 23% reduction. At lower temperatures, the mean free path of the carrier increases giving free carrier more energy for a given electric field prior to collision, thus resulting in a reduced avalanche breakdown voltage.

The on-resistance improvement of power DMOS at lower temperature is a direct result of the increased carrier mobility. Fig. 4a compares the electron mobility distribution within the baseline half unit-cell at 300 K and 77 K under a drain–source voltage of 3 V. Due to the impurity scattering, the electron mobility in the high doping region does not increase much at 77 K. The electron mobility in the accumulation and the channel regions has moderate increase at 77 K due to surface scattering. While electron mobility increases by a factor of over 10 in the lightly doped epi-layer at 77 K compared to that at 300 K, it only increases by a small amount in the JFET region for the baseline unit-cell. Fig. 4c shows that, in the JFET region, while the horizontal electrical fields are low, the vertical electrical fields (parallel to the current flow) are high at both temperatures. This high parallel electrical field in the JFET region corresponds to the limited improvement of electron mobility at 77 K in this region as shown in Fig. 4b. We believe the reason for the limited electron mobility improvement in JFET region is due to electron velocity saturation. Fig. 4d shows the electron velocities in the JFET region with various drain–source voltages at 77 K. It is clearly shown that the peak electron velocities in the JFET region saturate at about $1.3 \times 10^7$ cm/s. Therefore, in order to achieve the highest
cryogenic improvement, the structure of the baseline unit-cell needs to be modified in such a way that the electrical field in the JFET region can be reduced. Fortunately, the electrical field in the JFET region in a unit-cell optimized for room temperature operation is normally much smaller than that in the baseline cell.

Fig. 3. (a) Electrical field and (b) electrostatic potential of the baseline unit-cell with $V_{ds} = 3$ V and $V_{gate} = 10$ V at 300 K.

Fig. 4. (a) Electron mobility distribution of the baseline unit-cell with $V_g = 10$ V and $V_{ds} = 3$ V at 300 K and 77 K. (b) Electron mobility along the vertical line through the center of the baseline unit-cell with $V_g = 10$ V and $V_{ds} = 3$ V at 300 K and 77 K. (c) Electrical field along the vertical line through the center of the baseline unit-cell with $V_g = 10$ V and $V_{ds} = 3$ V at 300 K and 77 K. (d) Electron velocity along the vertical line through the center of the baseline unit-cell with $V_g = 10$ V, $V_{ds} = 3$ V, 10 V, and 30 V at 77 K.
3.2. Gate-length optimization

The effect of the gate-length on the cryogenic improvement of unit-cell was studied by conducting numerical simulations of unit-cells with various gate-lengths while other design parameters remained constant. It is well known that a larger gate-length leads to smaller JFET region resistance but larger accumulation and channel resistance because of the longer path along the oxide-Si surface as well as smaller equivalent channel density [23]. Optimization is the best approach to find the best trade-off between these major on-resistance components. As the gate-length increases, the electrical field in the JFET region decreases; therefore, we can also expect on-resistance improvement at cryogenic temperatures.

Fig. 5a shows, that as the gate-length increases to 72 μm, not only the DMOS on-resistance decreases substantially (compared to Fig. 2b from 12 μm) at room temperature, but also the much higher (10x) cryogenic improvement is achieved. Fig. 5b and c show the electrical field and electron mobility along the center vertical line of the 72 μm-gate-length unit-cell at 300 K and 77 K. Compared with the baseline unit-cell, under a drain–source voltage of 3 V, the vertical electrical field is reduced by a factor of 10. Without the existence of the high parallel electrical field in the JFET region, the electron mobility in the region increases more than 10 times when cooled to 77 K without suffering from the velocity saturation effect. Since the drift region and the JFET region resistance contribute to the majority of the on-resistance of the unit-cell, the on-resistance of the DMOS decreases by a factor of 10 at 77 K.

Fig. 6 shows the plot of the specific on-resistance vs. the gate-length, where the value of specific on-resistance for a particular gate-length is taken as the value at $V_{ds} = 0.5$ V. At 300 K, the optimized value of the gate-length is found to be 80 μm while the optimized value of the gate-length is 72 μm at 77 K. However, the difference is less than 0.5% between the 80 μm and 72 μm gate-length cells at 77 K. This means gate-length optimized for 300 K is effectively optimized for 77 K operation. The gate-length optimization of DMOS at cryogenic temperatures has been analyzed before by Singh using an analytical approach [34]. He concluded that the power DMOS with optimized gate-length design for room temperature operation would have more than 15% higher on-resistance than that optimized for 77 K operation. We believe the possible discrepancy might come from his calculation of JFET region resistance, where mobility in the region was assumed to follow the same temperature dependence as that in the epi-layer. It has been shown in our simulations that it does not always hold true for smaller gate-lengths where parallel electrical field is high.

3.3. Comparison with experimental results

Experimental characterizations were performed on commercially available power MOSFETs and compared to the
simulation results. The tested device is a 1000 V vendor rated power MOSFET with a measured breakdown voltage of 1100 V. The tested device has a die area of 1.9 cm². In Fig. 7, both the measured current and on-resistance have been normalized to 1 cm² for comparison with simulation results. The power MOSFETs were mounted on the Direct Bonded Copper (DBC) and wire bonded for electrical connections. A Tektronix 371A curve tracer was used to measure the $I-V$ characteristics of the integrated module at both room temperature and liquid nitrogen temperature. Since the curve tracer provides pulsed current with a pulse width of 250 µs and repetition rate of 15 Hz, little temperature rise in the MOSFET was observed due to minimum heat dissipation during testing. The power DMOS cell structure used in the simulation has an epi-layer doping concentration of $1.25 \times 10^{14}$ cm$^{-3}$, epi-layer thickness of 84 µm, and gate-length of 70 µm and the simulation found its breakdown voltage to be 1120 V. As shown in Fig. 7, the simulation results matched with experimental measurement very well at both room temperature and 77 K even though the exact cell structure of the tested devices is not known. The simulation correctly predicted more than 14x reduction of the on-resistance for this particular device. At 77 K, the breakdown voltage of the tested devices was measured to be 860 V, which is also very close to the breakdown voltage of the simulated cell, 884 V.

### 3.4. Epi-layer thickness optimization for cryogenic operation

One of the purposes of this paper is to answer the question: “Does the commercially available power MOSFETs also give optimized on-resistance although its breakdown voltage is degraded?” Our analysis in the gate-length optimization showed that no additional optimization of gate-length is needed for cryogenic operation. However, we found that epi-layer thickness could be reoptimized to gain the most performance improvement at 77 K.

As mentioned earlier, the resistance from the epi-layer contributed to the majority of the on-resistance for medium to high voltage power MOSFETs. Higher voltage blocking capability requires thicker epi-layer and lower doping concentration. Both factors increase device on-resistance. Device design therefore requires optimization of a combination of the epi-layer thickness and the doping concentration so that lowest specific on-resistance can be achieved. For certain epi-layer doping levels, there is a corresponding threshold thickness, above which a maximum blocking voltage will be reached no matter how thick the epi-layer is; on the other hand, if the thickness is smaller than the threshold value, the blocking voltage will decrease. The cell with this threshold thickness will have the lowest on-resistance with the highest reachable blocking voltage for a specific epi-layer doping level. The commercially available devices have an optimized combination of epi-layer doping level and thickness for room temperature operation. However, we found that the optimized design for room temperature operation is not optimal at 77 K. In our analysis, we conducted simulations for cells with epi-layer doping concentrations from $1.25 \times 10^{14}$ cm$^{-3}$ to $3.5 \times 10^{14}$ cm$^{-3}$. This covers power MOSFETs with breakdown voltages between 500 V and 1200 V at room temperature. The epi-layer thickness under consideration ranged from 32 µm to 112 µm.

The simulation results of blocking voltage dependence on epi-layer thickness for various epi-layer doping concentrations at both 300 K and 77 K are shown in Fig. 8. They clearly demonstrate that there exists a threshold epi-layer thickness for any specific doping level at certain temperature. Increasing the thickness above this threshold value will increase the specific on-resistance without increasing the blocking voltage. It also clearly shows that the threshold epi-thickness for certain doping levels at 77 K is significantly smaller than that at 300 K. It indicates that commercially available devices, which are optimized for room temperature operation, are not optimal for cryogenic operation. The epi-layer thickness of device optimized for room temperature operation can be reduced significantly without reducing its blocking voltage at 77 K. This is because, at 77 K, the mean free path of carriers increases giving free carrier more energy for a given electric field prior to collision, which effectively reduces the critical electrical field for avalanche breakdown. As the critical field decreases, the threshold epi-layer thickness required to reach critical field before punch-through also decreases.
due to the triangular shape of the electrical field near the p-base and epi-layer junction. The summary of epi-layer thickness optimization for various doping levels at both 300 K and 77 K are listed in Table 1. It indicates while more than 8–13 times on-resistance improvement (depending on the voltage rating) can be expected by cooling the commercially available power MOSFETs to 77 K, the device optimized for 77 K by reducing the epi-layer thickness can provide additional 5–10% on-resistance reduction depending on its voltage rating.

4. Conclusion

Two-dimensional numerical simulation study of the Si power DMOS’ at cryogenic temperatures is presented in this paper. In the simulation, Fermi–Dirac statistics and incomplete ionization are considered. Physical based models for temperature dependent silicon properties are employed. The simulation results are compared with experimental measurements with good agreement. It is suggested that any medium to high voltage power MOSFETs will get huge on-resistance improvement (8–13 times, depending on voltage rating) by cooling down to 77 K. The simulation results also suggest that the gate-length of a power DMOS cell optimized for room temperature operation is sufficiently optimized for cryogenic operations. However, the epi-layer thickness optimal for room temperature operation can be reduced without significantly reducing its breakdown voltage at 77 K. This leads to another 10% in on-resistance reduction.

References


