NUMERICAL SIMULATION OF TUNNETT AND MITATT DEVICES IN THE MILLIMETER AND SUBMILLIMETER RANGE†

Chien-Chung Chen, Richard K. Mains, George I. Haddad and Heribert Eisele

Solid-State Electronics Laboratory
Department of Electrical Engineering and Computer Science
The University of Michigan, Ann Arbor, Michigan 48109

Abstract — Numerical simulation programs for two-terminal transit-time devices based on drift-diffusion and energy-momentum transport models, with valence band to conduction band tunneling incorporated, have been developed. These programs can deliver accurate TUNNETT and MITATT device simulation results in the millimeter and submillimeter range. As the simulation results show, while the energy-momentum program is more accurate in the higher frequency range, the drift-diffusion program, which demands less computer resources, is suitable for W band devices.

Simulation results for GaAs TUNNETT and MITATT devices by these two methods will be presented and compared. The results obtained so far indicate that these programs provide useful tools for high frequency device structure design and optimization. In particular the energy-momentum model is required for devices operating at THz frequencies.

1 Introduction

Tunnel injection transit-time (TUNNETT) and mixed tunneling and avalanche transit-time (MITATT) diodes are two potentially useful RF power sources at extremely high frequencies[1]. The valence band to conduction band tunneling by electrons in such devices has a profound effect on the device performance. In the past, many numerical simulation programs have been developed to simulate transit-time devices, mostly for simulation of impact ionization avalanche transit-time (IMPATT) diodes[2, 3]. The free charge carrier transport models used in such simulations are the drift-diffusion model and energy-momentum model[4]. Programs based on the drift-diffusion model are more popular because of its simplicity. The energy-momentum model, due to the large number of partial differential equations to solve, is less frequently adopted in simulation.

Simulation of the interband tunneling process is still rare today. Recently Dash and Pati developed a generalized method to include the tunneling mechanism in their simulation[5]. Their simulation was based on the simple drift model (no diffusion currents), and was limited to DC or small-signal analysis. And their simulation lacks verification with measurement data for real devices.

In this work both the large-signal drift-diffusion and energy-momentum programs are developed to simulate TUNNETT and MITATT diodes. Tunneling currents are modeled and incorporated

†This work was supported by the Army Research Office under contract No. DAAL03-92-G-0109.

in the transport models. A real W-band diode is simulated using both models, and simulation results are compared with the experimental measurement. A 300 GHz diode is also designed and simulated to show the capability and limitation of such devices at this frequency.

2 Transport Theory

The two most commonly used transport models for semiconductor device simulation are the drift-diffusion (DD) model and the energy-momentum (EM) model. In the DD model, charge carriers respond immediately to the external electric field and reach steady state without any delay in time, while the energy-momentum model requires relaxation times for carriers to adjust their energy and momentum before reaching steady state. Due to this transient nature in the EM model, free carriers exhibit velocity overshoot and undershoot effects in response to the change of electric field. This kind of transient response can also be observed in the Monte Carlo simulation. By using correct relaxation times, the EM model can depict carrier behavior in remarkable agreement with Monte Carlo observations[6].

Due to the simplicity in formulation, the DD transport model is still widely used for device simulation. As long as the time and spatial variations of the electric field are small, the device simulation results by the DD model agree with those by the energy-momentum model. Even at high frequencies where the EM model is more feasible, the DD simulation is still useful in providing preliminary insight of device operation.

Two-terminal transit-time device simulation programs have been developed in our laboratory. Both include the interband tunneling and impact ionization as the major carrier generation mechanisms. In the following we briefly describe the two transport models used in the two simulation programs developed in our laboratory.

2.1 Drift-Diffusion Model

This well-known model describes the free charge carrier transport by the following equations

$$J_p = qp\mu_p(E)E - qD_p(E)\frac{\partial p}{\partial x}$$
$$J_n = qn\mu_n(E)E + qD_n(E)\frac{\partial n}{\partial x}$$
$$\frac{\partial p}{\partial t} = -\frac{1}{q}\frac{\partial J_p}{\partial x} + G$$
$$\frac{\partial n}{\partial t} = \frac{1}{q}\frac{\partial J_n}{\partial x} + G$$

where the carrier mobilities, μ_p and μ_n, and the diffusion coefficients, D_p and D_n, are functions of the electric field E, and the free carrier generation rate G is given by

$$G = G_{II} + G_T + G_{th} - R_{th}$$

with G_{II}, G_T, G_{th} and R_{th} being the impact ionization generation, interband tunneling generation, thermal generation and recombination, respectively.

2.2 Energy-Momentum Model

In the EM simulation program developed by us, holes still follow the drift-diffusion equation, while electrons are described by a variation of the two-valley drift-diffusion equation with the mobilities and diffusion coefficients being functions of electron energies instead of the electric field. Two energy equations govern the electron energies. The following equations are solved by the EM simulation program:

$$J_p = qp\mu_p(E)E - qD_p(E)\frac{\partial p}{\partial x}$$

$$J_{n1} = qn\mu_{n_1}(w_1)E + qD_{n_1}(w_1)\frac{\partial n1}{\partial x}$$

$$J_{n2} = qn\mu_{n_2}(w_2)E + qD_{n_2}(w_2)\frac{\partial n2}{\partial x}$$

$$\frac{\partial w_1}{\partial t} = -\frac{J_{n1}E}{n_1} - \frac{w_1 - w_{th}}{\tau_{w1}(w1)}$$

$$\frac{\partial w_2}{\partial t} = -\frac{J_{n2}E}{n_2} - \frac{w_2 - w_{th}}{\tau_{w2}(w2)}$$

$$\frac{\partial p}{\partial t} = -\frac{1}{q}\frac{\partial J_p}{\partial x} + G_p$$

$$\frac{\partial n1}{\partial t} = \frac{1}{q}\frac{\partial J_{n1}}{\partial x} + G_{n1}$$

$$\frac{\partial n2}{\partial t} = \frac{1}{q}\frac{\partial J_{n2}}{\partial x} + G_{n2}$$

where $w_{th} = \frac{3}{2}k_B T$ is the thermal equilibrium energy for electrons, and G_p, G_{n1} and G_{n2} are given by

$$G_p = G_{II} + G_T + G_{th} - R_{th}$$

$$G_{n1} = G_T + G_{th} - R_{th} + \frac{n2}{\tau_{n_{2\to1}}(w2)} - \frac{n1}{\tau_{n_{1\to2}}(w1)}$$

$$G_{n2} = G_{II} - \frac{n2}{\tau_{n_{2\to1}}(w2)} + \frac{n1}{\tau_{n_{1\to2}}(w1)}.$$

Close examination of the above equations reveals that the momentum relaxation times for electrons are actually neglected in the program. This is justified by the fact that the momentum relaxation times are generally much smaller than the energy relaxation times.

3 Material Parameters

The material used for devices in this work is GaAs. Since these devices are intended for RF power generation, the material lattice temperature is usually very high. 500°K is assumed throughout this work. The majority of the material parameters used by the simulation programs are generated by a traditional Monte Carlo (MC) program[7]. This MC program provides the values of electron

velocities, diffusion coefficients, relaxation times and electron energies as functions of the electric field up to 1000 kV/cm.[1] Beyond 1000 kV/cm, the material parameters are extrapolated from the MC results. The methods used to evaluate the relaxation times are similar to those proposed by Stewart *et. al.*[6], while the electron (longitudinal) diffusion coefficients are calculated in the MC program using the method discribed by Fawcett[9]. Other electron parameters such as the material intrinsic concentration and electron-hole recombination lifetime and all the hole parameters are set according to empirical expressions or assumed values.

Two important material parameters used by the simulation programs are the impact ionization and tunneling rates. They have very profound effects on the device operation, but still lack adequate determination in the electric field range where the devices are operated (up to 3000 kV/cm). For the impact ionization rate, we use values that enhance the generation rate in the high field slightly to account for the bandgap narrowing effect in the heavily doped region. The impact ionization generation rate is expressed as

$$G_{II} = (\alpha_n J_n + \alpha_p J_p)/q$$

with α_n and α_p being given by

$$\alpha_n = A_n \exp\left[-(B_n/E)^2\right]$$

$$\alpha_p = A_p \exp\left[-(B_p/E)^2\right]$$

For GaAs at 500°K, we use

$$A_n = A_p = 2.7 \times 10^5 \quad \mathrm{cm}^{-1}$$

$$B_n = B_p = 7.1 \times 10^5 \quad \mathrm{V/cm}$$

As to the interband tunneling rate, we adopt the expression which complies with the highly idealized form proposed by Kane[10] as shown below

$$G_T = A_T E^2 \exp(-B_T/E)$$

For GaAs at 500°K, we use the following values for A_T and B_T

$$A_T = 1.5 \times 10^{20} \quad \mathrm{cm}^{-1}\mathrm{s}^{-1}\mathrm{V}^{-2}$$

$$B_T = 1.5785 \times 10^7 \quad \mathrm{V/cm}$$

In the EM program, all material parameters for electrons except the tunneling rate are tabulated according to the electron energy. In other words, they are treated as functions of the electron energy instead of the electric field. The tunneling rate is considered as a function of the electric field since it is related more to the electric field that electrons in the valence band experience than to the energy of electrons in the conduction band. Figure 1 shows the impact ionization and tunneling generation rates for 500°K GaAs that are used in the simulation programs.

[1]Actually our MC program's validity should not extend to this high field. For electric fields of hundreds of kV/cm, a more comprehensive MC technique that takes into account the actual band-structure should be used instead[8]. A similar program is currently under development in our laboratory.

4 Simulation Results and Discussion

Two transit-time device structures are simulated by the two programs discussed in the preceding sections. One is a 94 GHz diode which has already been fabricated and tested in our laboratory[11]. The other is a diode designed for 300 GHz operation. Simulation on the 94 GHz diode is to verify the validity of material parameters used in the programs, and by adopting the same material parameters for the 300 GHz simulation, the TUNNETT and MITATT diode's performance at higher frequencies can be investigated. Figure 2 shows the schematic diode structures.

Design of the 300 GHz diode comes from modification of the 94 GHz diode. For diodes operating at 300 GHz, the drift region becomes short, and the efficiency drops should the generation region remain the same. In order to reduce the generation region length, the n^+ doping concentration is set to 4×10^{18} cm^{-3}, which is close to the maximum n-type doping level achievable by MBE. The n-type drift region doping concentration is increased significantly to accommodate the high current density (250 kA/cm^2). Such high current density is required to obtain enough negative resistance from the diode to overcome the contact resistance. To understand the need of high bias current for high negative resistance, we write down the negative resistance of the diode as

$$-R = \frac{-G}{G^2 + B^2}$$

where B and $-G$ are diode's susceptance and negative conductance, respectively. At frequencies as high as 300 GHz, B is usually dominated by the diode's capacitance C and is much higher than $-G$. Therefore the diode's negative resistance is approximated by

$$-R \approx \frac{-G}{\omega^2 C^2}$$

In the above expression for negative resistance, the capacitance is essentially independent of the bias current density. To increase $-R$, we may try to increase $-G$. Increasing the bias current density has the effect of increasing the negative conductance since at higher current density the same amount of V_{rf} causes greater AC current. This explains the reason why we bias the 300 GHz diode at very high current density. Simulation shows the diode's bias voltage is about 6 Volts at 250 kA/cm^2. The DC power density is as high as 1.5 MW/cm^2. Dissipation of excessive heat is critical for such high power density.

The following presents the simulation results and discussion for the two diodes.

4.1 94 GHz Diode Simulation

The 94 GHz diode is a single-drift device with strong tunneling current. When biased under the DC condition, the bias voltage at $I_{dc} = 0.14$ A ranges from 9.5 to 11 Volts among different samples with 30 μm as the nominal diameter. Due to the small uncertainty of the diode's area from the undercut phenomenon during mesa etching, the actual current density used to test the diode is also uncertain. By assuming the diode area to be 5.6×10^{-6} cm^2, $J_{dc} = 25$ kA/cm^2 is used in the simulation. The maximum RF output power measured at 94 GHz is 33 mW with the efficiency being 2.65% and bias voltage being 9 Volts.

Both the DD and EM simulations include the heavily doped contact regions for more realistic results, although doing so makes the simulation process extremely slow. Simulating such contact regions is especially desirable for single-drift devices since the p^{++} region is part of the narrow yet important carrier generation zone. Specifying a lower doping for the p^{++} region or neglecting it in simulation leads to a very different result.

DC Simulation at 25 kA/cm^2

Figure 3 shows the electron concentration and electric field obtained from both the DD and EM DC simulation. The electron concentration in the drift region by the EM program is higher than by the DD program, which gives rise to the higher electric field and higher bias voltage by the EM program due to the higher space charge effect. Also the higher electron concentration in the drift region by the EM program needs lower electron drift velocity to achieve the same DC current. This explains why the optimum frequency by the EM program is slightly lower than by the DD program as shown in the AC simulation results later. It is also interesting to note that a large amount of electrons appear near the junction of the generation region and drift region by the EM program, which phenomenon can be explained by the significantly higher electron energy than in equilibrium in this region as shown in Figure 4. The higher electron energy also expands the effective avalanche region length. When the diode operating frequency gets higher, this avalanche region expanding effect may cause lower diode efficiency. In the drift region the slightly higher electron energy than in equilibrium leads to a slower electron velocity, which is also known as the velocity undershoot effect. On the other hand, the higher electron energy in the drift region means lower electron diffusion current, which is one of the reasons why the EM program usually predicts better diode efficiency than the DD program.

The DC bias voltage at 25 kA/cm^2 obtained by the DD and EM simulation is 10.5 and 10.7 Volts, respectively. The tunneling current is 9.1 kA/cm^2 by the DD program and 9.2 kA/cm^2 by the EM program.

AC Simulation at 25 kA/cm^2

Both DD and EM AC simulations at 94 GHz are carried out on the 94 GHz diode. Figure 5 shows the voltage and current waveforms for V_{rf} = 4.5 Volts at 94 GHz by both models. Comparison of the induced current waveforms by both models shows that electrons experience the velocity undershoot effect from ωt=90° to ωt=270° where the external voltage is decreasing, and the velocity overshoot effect for the rest of the cycle. The bias voltage at V_{rf} = 4.5 Volts is 8.9 Volts by the DD program and 9.2 Volts by the EM program.

AC simulation results at 94 GHz are summarized in Figure 6, which shows the RF output power and efficiency with and without considering the 1.5×10^{-6} Ω nominal contact resistance. The diode area is assumed to be 5.6×10^{-6} cm^2 for RF output power calculation. Even with the contact resistance being considered, both models predict better output power and efficiency than the measurement results. The bias voltage at which the maximum RF output power occurs is 8.9 Volts by both models. From the figure, it can be easily seen that the EM program gives higher efficiency, RF output power and larger RF voltage swing, although the differences are not very

significant. For frequencies up to W band, the DD program provides a very good tool for diode design and preliminary performance analysis, while the EM program is indispensable for more accurate analysis and better understanding of the device physics.

As mentioned in the discussion of DC simulation, the EM model is expected to give lower optimum frequency than the DD model. This is verified in Figure 7, in which the optimum frequency at $V_{rf} = 5$ Volts is 103 GHz by the DD model and 97 GHz by the EM model. At $V_{rf} = 5.5$ Volts, the optimum frequency by the EM model is 99 GHz, a slight increase from 97 GHz at $V_{rf} = 5$ Volts. With increased V_{rf} the diode's optimum frequency tends to increase. This is because the average electric field in the device tends to decrease as V_{rf} increases, and thus results in increased average electron velocity and optimum frequency.

4.2 Simulation of the Diode Designed for 300 GHz

Like the 94 GHz diode, the 300 GHz diode is a single-drift diode with heavily doped p^{++} contact region. For simulation accuracy, both the DD and EM programs include the heavily doped p^{++} and n^{+} contact regions, although doing so significantly slows down the simulation speed.

DC Simulation at 250 kA/cm^2

The DC simulation results by both programs are shown in Figure 8 and 9. The electron concentration predicted by the EM program significantly departs from the DD program's. Also the EM program predicts even higher electron energy above the equilibrium energy than in the 94 GHz, which indicates the electron velocity would be significantly slower than expected by the DD program due to the velocity undershoot effect. Therefore we anticipate the EM program will give a lower optimum frequency than the DD program. Also note the avalanche region expanding effect by the EM program. At 300 GHz the total diode length is short. The expansion of the avalanche region actually disperses the injected charge pulse spatially, and thus results in low efficiency.

The DC bias voltage at 250 kA/cm^2 is 6.33 Volts by the DD program and 6.36 Volts by the EM program with the respective tunneling currents being 60.1 and 56.7 kA/cm^2. The higher avalanche current by the EM program comes from the longer effective avalanche region.

AC Simulation at 250 kA/cm^2

Frequencies around 300 GHz have been swept by the two programs to determine the optimum frequency. Since the diode structure was originally designed with the help of the DD program, the maximum efficiency occurs at 300 GHz by the DD program. Due to the slower electron velocity in the drift region by the EM program, the optimum frequency by the EM program, 260 GHz, is significantly lower than by the DD program. We believe that in this extremely high frequency range electrons no longer follow the DD transport model. Therefore, in the following discussion only the EM program's results are presented.

A series of AC simulations at 260 GHz are carried out with small steps of V_{rf} to search for the

maximum output power. Figure 10 shows the current waveforms for V_{rf} = 2 Volts at 260 GHz. From it, one can see that at this high frequency impact ionization, instead of interband tunneling, still dominates the charge generation process. Two factors contribute to the dominance of impact ionization — one is the high bias current density, which enhances only the impact ionization; the other is the significant generation region expansion phenomenon for impact ionization at this high frequency.

Figure 11 shows the predicted diode output power, efficiency and negative resistance for a 10 μm diameter diode at 260 GHz with bias current density being 250 kA/cm^2 by the EM program. These P_{rf} and efficiency values are calculated without including any contact resistance. When the contact resistance is included in the calculation, P_{rf} and efficiency drop drastically since the diode's negative resistance is low. The actual maximum P_{rf} and efficiency after taking into account the nonzero contact resistance are plotted in Figure 12 with the load resistance also being shown. This figure reveals that the diode's performance is severely degraded by the contact resistance. From another point of view, as long as the contact resistance can be reduced to 5×10^{-7} $\Omega\cdot$cm^2, 18 mW of RF power can be provided by this diode at 260 GHz.

In addition to reduction of the contact resistance, many measures can be taken to improve the negative resistance at high frequencies. One way to increase the negative resistance is to increase the diode length without affecting the diode's operating frequency. For example, we may design the diode in such a way that the electric field in the drift region is lower than it is now to increase the electron drift velocity; or we may build the diode with InP, which has higher electron drift velocity at high electric field; or we may reduce the diode temperature by cooling it in liquid nitrogen to achieve high electron drift velocity. Another way is to build the diode using low bandgap materials such as InGaAs to increase the diode's negative conductance.

5 Conclusion

Interband tunneling has been included in both drift-diffusion and energy-momentum transport model programs to simulate operation of TUNNETT and MITATT diodes. Simulation using both programs is presented for an experimental diode. Both programs agree with the RF performance at 94 GHz that was measured in our laboratory. Simulation shows that the energy-momentum model gives slightly better diode performance at 94 GHz. A 300 GHz diode is also designed and simulated. The simulation results show that the drift-diffusion model deviates significantly from the energy-momentum model at such high frequencies. To accurately predict the diode's performance at terahertz frequencies, the energy-momentum program should be used. However, the fast drift-diffusion program is useful for preliminary diode structure design.

At terahertz frequencies, TUNNETT and MITATT diodes' performance is severely limited by the contact resistance. To extend the use of such devices into the terahertz range, the low-resistance ohmic contact technique is highly desired.

References

[1] G. I. Haddad *et. al.*, "Tunnel transit-time (TUNNETT) devices for terahertz sources," *Microw. Opt. Technol. Lett.*, Vol. 4, No. 1, pp. 23–29, 1991.

[2] P. E. Bauhahn, "Properties of semiconductor materials and microwave transit-time devices." Ph. D. Thesis, University of Michigan.

[3] R. K. Mains *et. al.*, "Simulation of GaAs IMPATT diodes including energy and velocity transport equations," *IEEE Trans. Electron Devices*, Vol. 30, No. 6, pp. 1327–1338, 1983.

[4] K. Bløtekjær "Transport Equations for Electrons in Two-Valley Semiconductors," *IEEE Trans. Electron Devices*, Vol. 17, pp. 38–47, 1970.

[5] G. N. Dash and S. P. Pati, "A generalized simulation method for MITATT-mode operation and studies on the influence of tunnel current on IMPATT properties," *Semicond. Sci. Technol.*, Vol. 7, pp. 222–230, 1992.

[6] R. A. Stewart *et. al.*, "Improved relaxation-time formulation of collision terms for two-band hydrodynamic models," *Solid-State Electronics*, Vol. 32, No. 6, pp. 497–502, 1989.

[7] C. Jacoboni and Paolo Lugli, "The Monte Carlo method for semiconductor device simulation," Springer-Verlag, Vienna, 1989.

[8] M. V. Fischetti and S. E. Laux, "Monte Carlo analysis of electron transport in small semiconductor devices including band-structure and space-charge effects," *Phys. Rev.*, Vol. 38, No. 14, pp. 9721–9745, 1988.

[9] W. Fawcett, "Non-ohmic transport in semiconductors," *Electrons in Crystalline Solids*, International Atomic Energy Agency, Vienna, pp. 531–618, 1973.

[10] E. O. Kane, "Zener tunneling in semiconductors," *J. Phys. Chem. Solid*, vol. 12, pp. 181–188, 1959.

[11] C. Kidner *et. al.*, "Tunnel injection transit-time diodes for W-band power generation," *Electron. Lett.*, Vol. 28, No. 5, pp. 511–513, 1992.

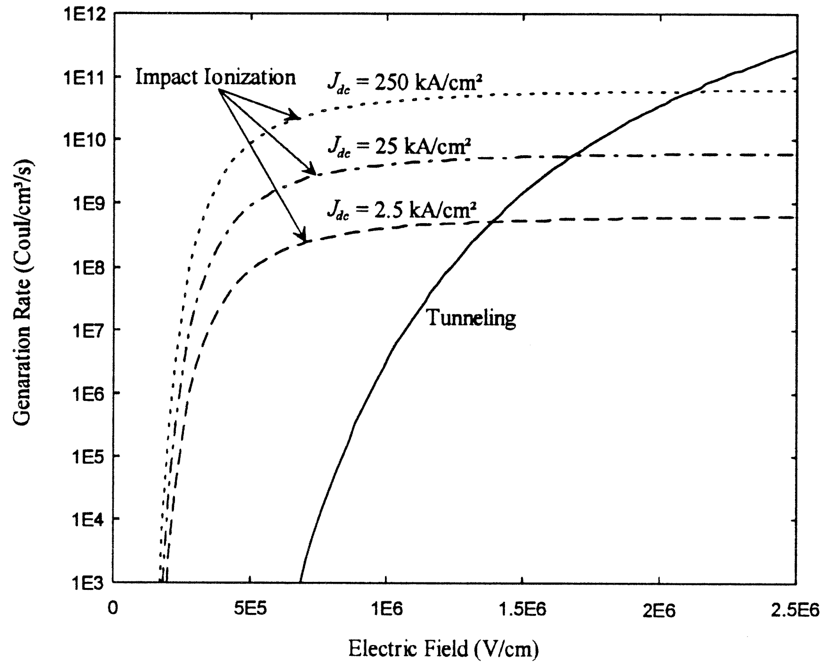

Figure 1: The impact ionization and tunneling generation rates for GaAs at 500°K.

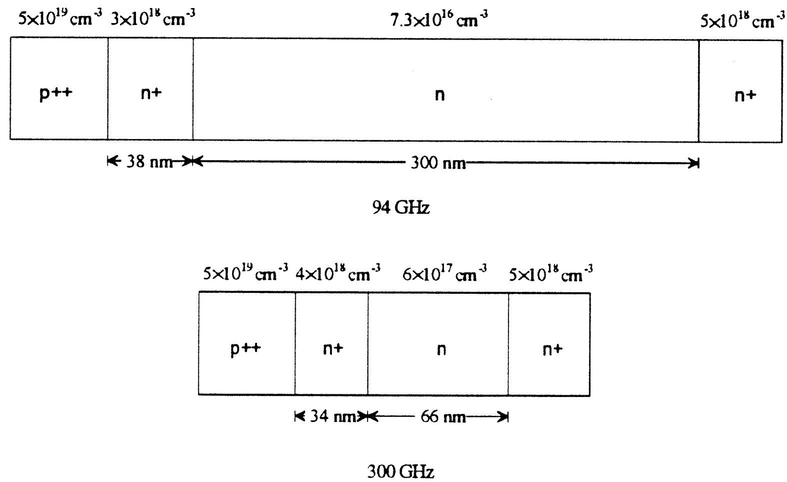

Figure 2: Doping structure of the two diodes simulated.

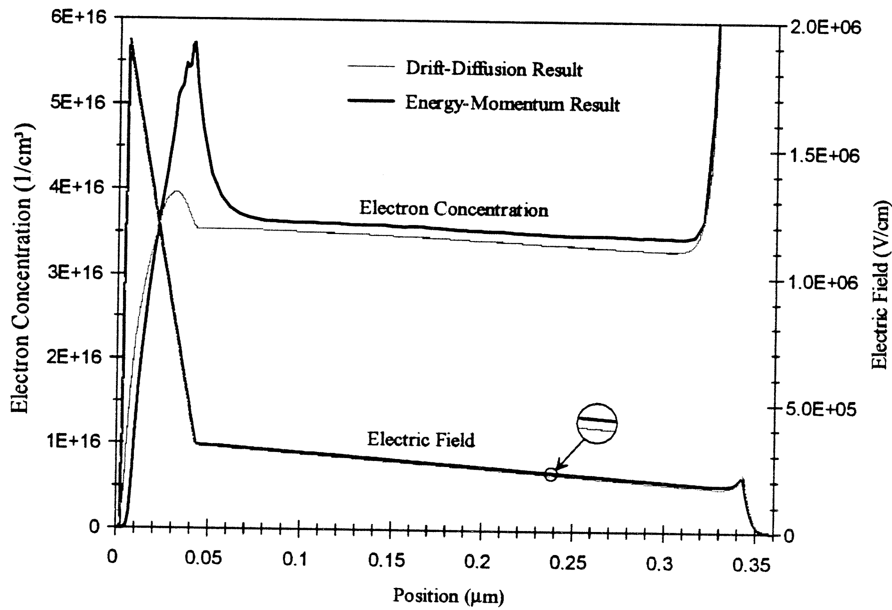

Figure 3: The DC solution for the 94 GHz diode's electron concentration and electric field.

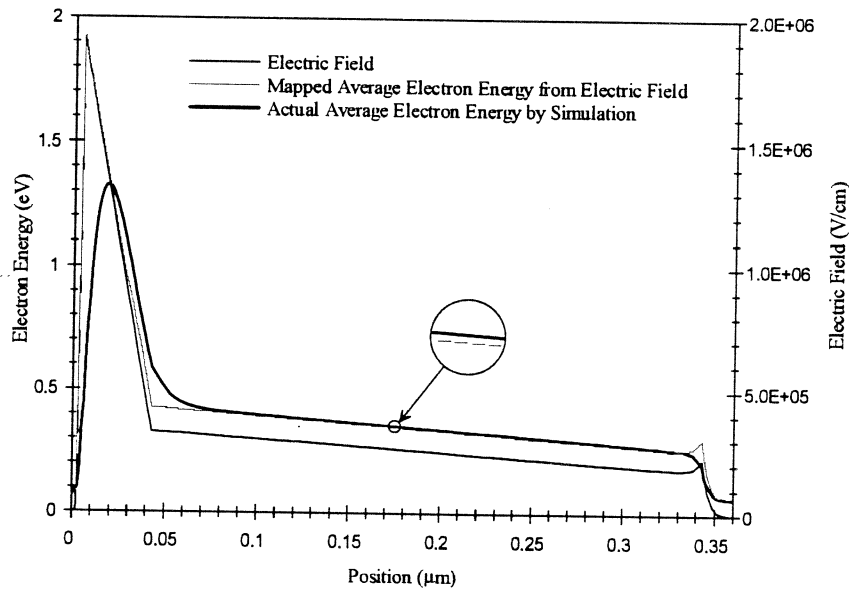

Figure 4: The electron energy in the upper valley for the 94 GHz diode by the EM program.

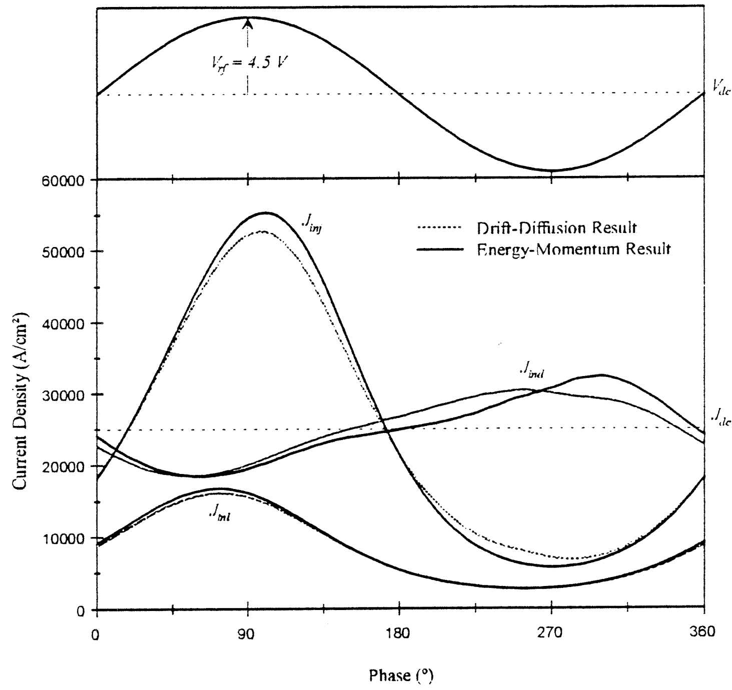

Figure 5: Voltage and current waveforms for V_{rf} = 4.5 Volts at 94 GHz.

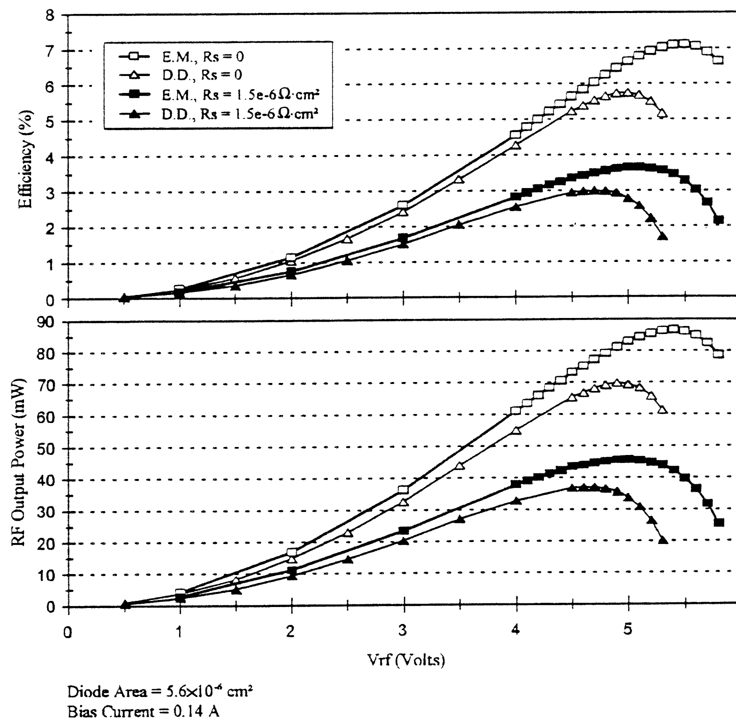

Figure 6: The 94 GHz large-signal simulation results.

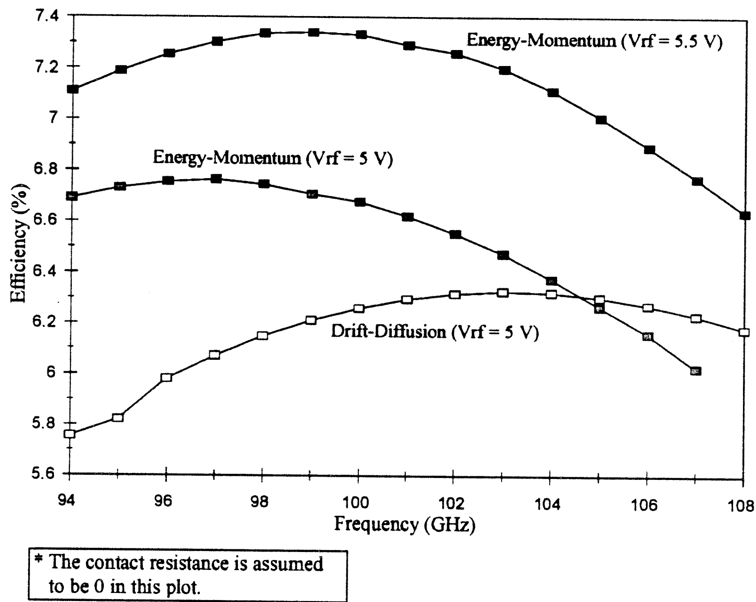

Figure 7: The 94 GHz diode efficiency versus frequency.

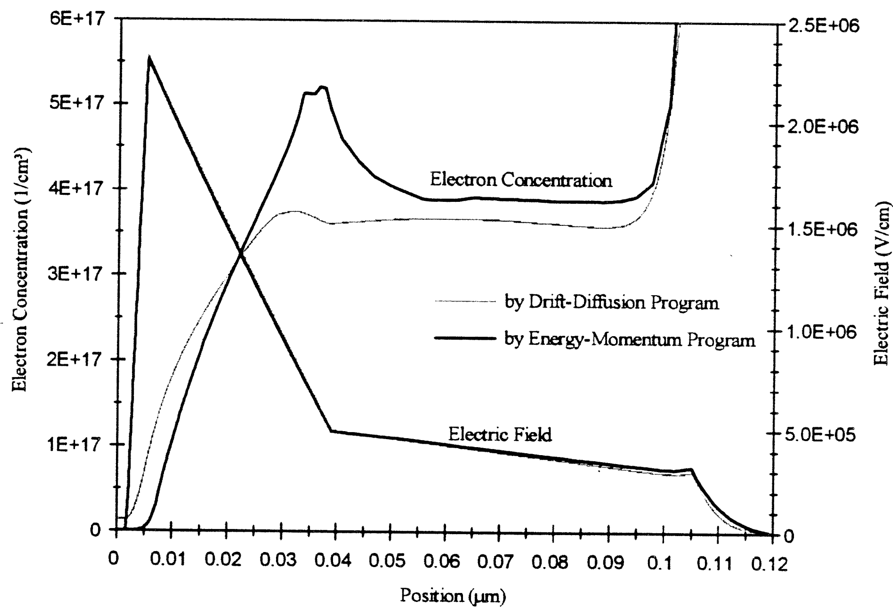

Figure 8: The DC solution for the 300 GHz diode's electron concentration and electric field.

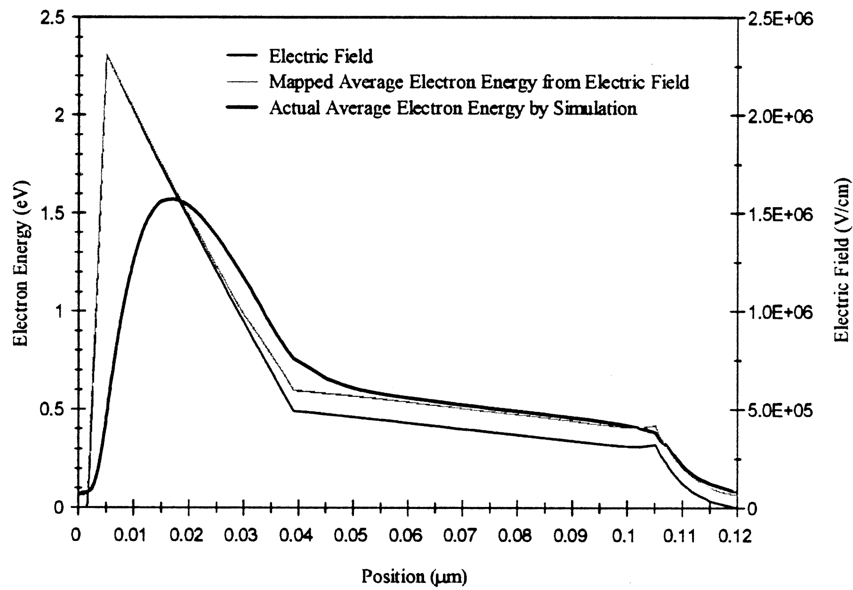

Figure 9: The electron energy in the upper valley for the 300 GHz diode by the EM program.

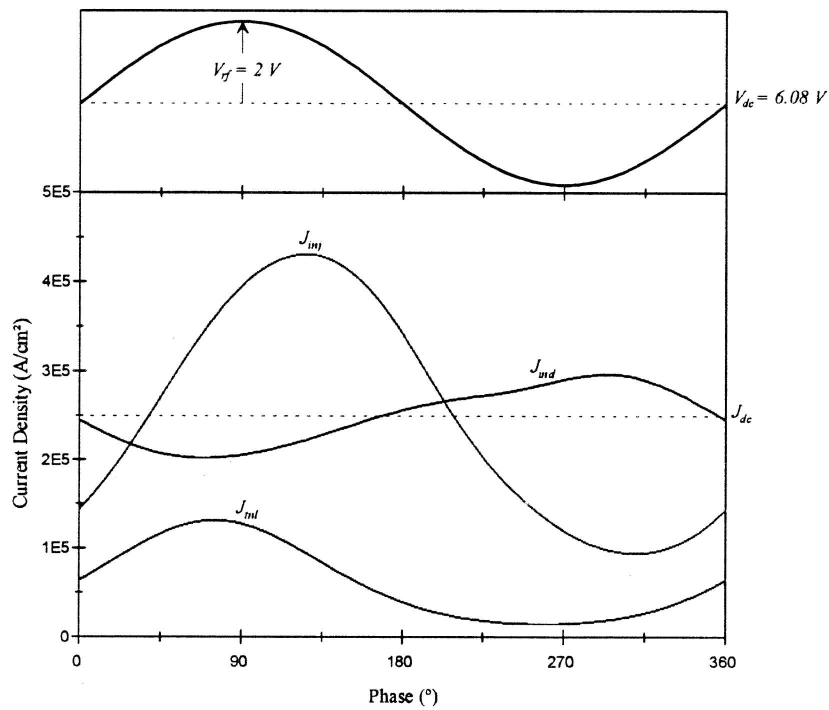

Figure 10: Voltage and current waveforms for V_{rf} = 2 Volts at 260 GHz.

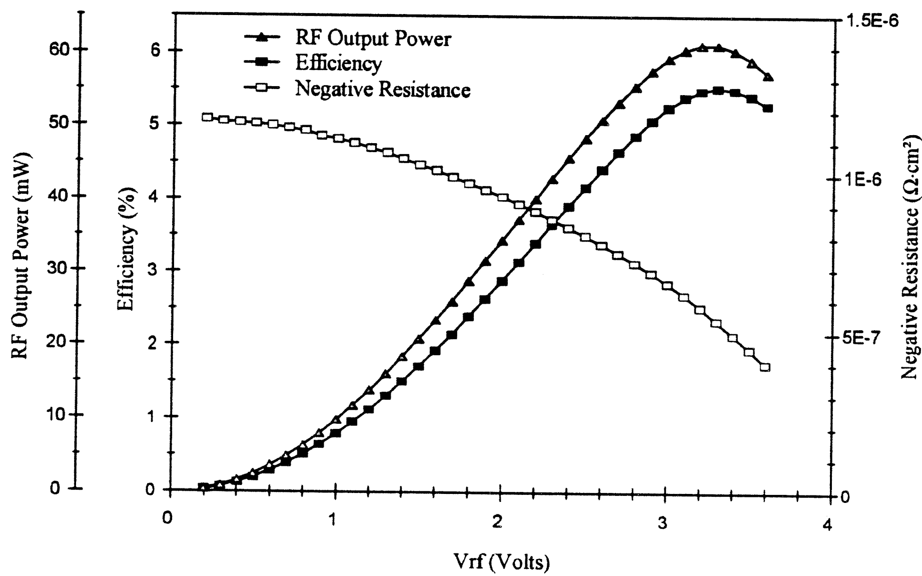

Figure 11: Predicted diode RF output power, efficiency and negative resistance at 260 GHz with current density 250 kA/cm^2 and diode diameter 10 μm.

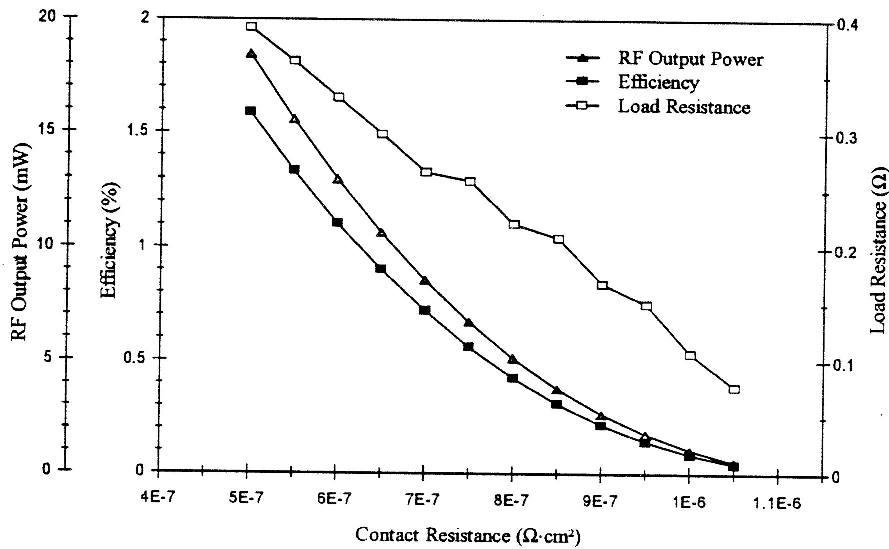

Figure 12: Predicted diode maximum RF output power and corresponding efficiency versus the contact resistance at 260 GHz with the diode diameter 10 μm.