

A TWO-DIMENSIONAL FINITE ELEMENT ANALYSIS
OF THE STATIONARY SEMICONDUCTOR DEVICE EQUATIONS

by

PATRICK PABLO CHAVEZ

B. A. Sc., The University of British Columbia, 1995

A THESIS SUBMITTED IN PARTIAL FULFILMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

Department of Electrical and Computer Engineering

We accept this thesis as conforming to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

July 1997

© Patrick Pablo Chavez, 1997

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Electrical and Computer Engineering

The University of British Columbia
Vancouver, Canada

Date August 15, 1997

Abstract

The ability to model the steady-state field inside active structures, such as a transistor, is an important aspect of monolithic microwave integrated circuit (MMIC) design. This paper focuses on such an active zone of semiconductor material, and presents a finite element analysis of the classical semiconductor equations. The semiconductor equations are very nonlinear and govern the potential and carrier density distributions in semiconductor materials. A previously developed finite element method (FEM) formulation of these equations, referred to as the current conservation model, is re-introduced and re-derived with compact matrix notation. It is shown how this formulation can be solved with the Newton-Raphson iterative scheme. Then, a newly developed FEM formulation, referred to as the advection-diffusion model, of the continuity equations is described in detail. It is shown by example how this formulation solved with Gummel's iterative technique is very numerically robust. These two different solution methods of the steady-state system of coupled Poisson and continuity equations are combined into a final solution algorithm that exploits their strengths. As a specific example, GaAs MESFETs are the focus of implementation, and the resulting potential field and carrier density distributions are used to calculate various MESFET parameters such as electrode currents, voltage gain, capacitances, and conductances. Finally, various extensions to the FEM approach involving the application of the method of moments (MoM) are briefly discussed and partially demonstrated. These extensions are intended to compensate for the assumptions and simplifications, mainly with respect to the artificial boundary conditions, used in the original stand-alone FEM formulations.

Table of Contents

Abstract	ii
List of Figures	v
1 Introduction	1
1.1 The Big Picture.	1
1.2 Objectives and Motivation	5
1.3 Thesis Structure.	6
2 Problem Specifications	8
3 FEM Formulation	12
3.1 Elements	12
3.2 Discretization	15
3.2.1 Poisson Equation	16
3.2.2 Continuity Equations	17
3.2.2.1 Current Conservation Model	17
3.2.2.2 Advection-Diffusion Model	19
4 Solution Methodologies	22
4.1 Introduction to Iterative Techniques	22
4.2 Gummel's Method	25
4.3 Newton-Raphson Method	31
5 Application	37
5.1 Mesh Spacing	37
5.2 Simulations	38
5.2.1 Example 1	41
5.2.2 Example 2	47

6	Extensions to FEM Model	57
6.1	Method of Moments and Integral Equations	57
6.2	Parasitic Capacitances and Green's Integral Equation	58
6.2.1	Complex Images	59
6.2.2	Multipipe Model	60
6.2.3	Results.	60
6.3	Artificial Boundary Condition Adjustments	62
6.3.1	Green's Integral Equation	62
6.3.2	Boundary Integral Equation	66
7	Conclusion	68
7.1	Accomplishments	68
7.2	Future Work	69
	Bibliography	71
	Appendix A Derivation of Shape Functions	77
	Appendix B Capacitance Calculations for a Multi-Port System	80

List of Figures

Figure 1	Layout of a monolithic microwave integrated circuit [1].	1
Figure 2	2D cross section of a MESFET device.	10
Figure 3	Six adjacent quadratic triangular elements.	12
Figure 4	Local (left) and global (right) shape functions for Node i.	13
Figure 5	Direct iterations; Newton-Raphson iterations.. . . .	23
Figure 6	Direct iterations; Newton-Raphson iterations.. . . .	24
Figure 7	Direct iterations; Newton-Raphson iterations.. . . .	25
Figure 8	(a) Schottky barrier strip and (b) results for $N_d = 5 \times 10^{15} \text{ cm}^{-3}$	28
Figure 9	Electron drift velocity versus electric field intensity.	41
Figure 10	MESFET device for Example 1.	42
Figure 11	Mesh grid for Example 1 - 959 nodes, 448 elements.. . . .	42
Figure 12	Potential field.	44
Figure 13	Electron carrier distribution.	44
Figure 14	Current density field.	45
Figure 15	I_d vs. V_d for Example 1.	46
Figure 16	I_d vs. V_g ($V_d = 2.5 \text{ V}$) for Example 1.	46
Figure 17	Mesh grid for Example 2 - 1453 nodes, 688 elements.	48
Figure 18	Electron carrier concentration and potential distributions.	49-50
Figure 19	Current density field.	51
Figure 20	I_d vs. V_d for Example 1.	52
Figure 21	(a) Drain and (b) gate current versus gate voltage.	53
Figure 22	Electrode charges for varying potentials on (a) the drain, (b) the gate, and (c) the source.	54
Figure 23	Small signal model of MESFET in Example 2.	56

Figure 24	MESFET with strip conductor feeds.	57
Figure 25	Modified small signal model of MESFET in Example 2.	61
Figure 26	Gain versus frequency plots for circuit models.	62
Figure 27	Modified mesh grid for Example 2 - 1677 nodes, 800 elements.	63
Figure 28	Drain current versus drain-source voltage.	63
Figure 29	Artificial boundary adjustment results.	65
Figure B1	Circuit model for two strip conductors.	81

1 Introduction

1.1 The Big Picture

As with other electrical technology, microwave technology has been following a path of smaller size and increased complexity. Monolithic microwave integrated circuits consist of a planar substrate on which passive and/or active circuit elements are constructed. Passive elements such as microstrip transmission lines, inductors, capacitors, and resistors take the form of various deposited layers of metal, dielectric, and resistive films on the surface of the substrate. In addition to this, a substrate made of semiconductor material may have active devices grown directly into it through ion implantation. For example, gallium arsenide substrates, which are the most common of the semiconductor substrates, allow for the construction of GaAs FETs which have many applications such as low-noise amplifiers, high-gain amplifiers, broadband amplifiers, mixers, oscillators, phase shifters, and switches [1]. Figure 1 shows a simple layout of a monolithic microwave integrated circuit (MMIC).

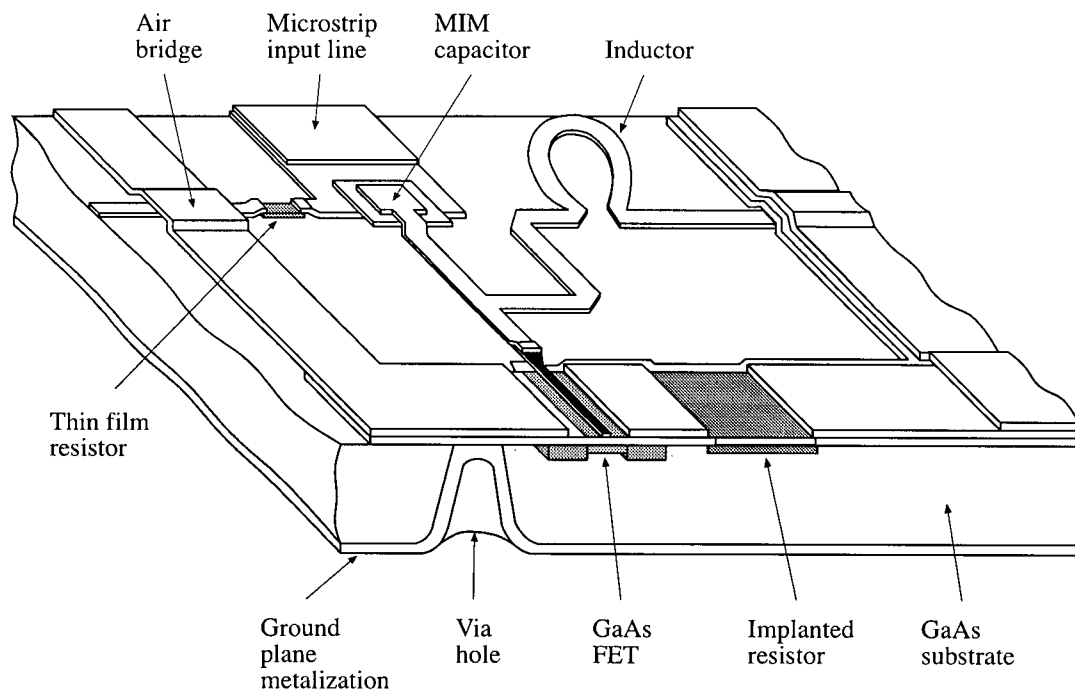


Figure 1. Layout of a monolithic microwave integrated circuit [1].

The situation of many circuit elements operating at microwave frequencies and occupying such a compact environment that exists in MMICs leaves little room for error in the design and calls for considerable precision in the fabrication of MMICs. Unlike conventional circuit analysis in which a circuit is simply treated as the sum of its components whose operation is independent of their environments, a more thorough analysis is required. Such an analysis must take into account the electromagnetic effects of mutual field coupling among the components, discontinuities and sharp edges in the structures of the components, biasing circuitry, and packaging. Extensive use of specialized computer-aided design (CAD) software that treats the analysis as an electromagnetics problem is vital.

In order to predict the voltage-current characteristics of a microwave circuit, the electromagnetic fields must be modeled. Dynamic and static fields induced by passive elements obey Helmholtz' equation and Laplace's equation, respectively, both of which are linear differential equations [1]. Dynamic and static fields induced by active elements also obey Helmholtz' equation and Poisson's equation (Laplace's equation with a source term), respectively, but are coupled to the variable charge distributions (electrons and holes) which are governed by the continuity equations [2]. The continuity equations are nonlinear.

Theoretically, it is possible to deal with the problem of calculating the fields occurring in an MMIC such as the one shown in Figure 1 by simply solving the appropriate equation with specified boundary conditions. Practically, due to the complexity of such a problem and limited computing resources, this is simply not feasible. The complexity resides in the diversity of boundary shapes, types of boundary conditions, uncertainty in material parameters, and nonlinearities in active regions that exist in a typical MMIC. As a result, it becomes necessary to break down the problem into smaller ones. A smaller problem may consist of focusing on a single transmission line, capacitor, inductor, or transistor which occupies just a portion of the circuit. In order not to neglect coupling to the rest of the circuit, "appropriate" artificial boundary conditions surrounding the particular problem domain must be employed.

Once a sub-problem is defined, a solution technique must be applied to the governing equation. In some cases analytic solutions are available [3], but in the vast majority of cases,

analytic solutions are simply not attainable. In such cases, numerical techniques have offered the only possible route to a solution. The finite element method, finite difference method, and method of moments are the three main classes of such techniques. Each has its own strengths and weaknesses.

The method of moments has proven to be ideal for linear problems [4]-[8]. It requires an analytic solution of the linear differential equation at hand for a point source. This solution is known as a Green's function. Based on the equivalence theorem which states that any static or dynamic electromagnetic field in a source-free domain can be fully represented by an equivalent distribution of charge or current sources, respectively, and taking full advantage of the principle of superposition of point source solutions, a Green's integral equation that represents the solution of the linear differential equation in terms of source distributions is set up. Intimately related to the Green's integral equation is the boundary integral equation, which also requires knowledge of the Green's function and gives an expression of the field satisfying the governing equation over the entire domain in terms of the field over a closed contour in or around the domain. The method of moments involves discretizing the integral equation by way of an approximate basis function expansion of the true unknown solution and a minimization of the weighted inner product residuals reducing the problem to the determination of the unknown expansion coefficients via a matrix inversion.

Finite element and finite difference methods appear similar to each other on the surface but, in reality, have completely different mathematical bases. They both solve differential equations, and they both require that the solution domain be finite. The fundamental difference between the two methods lies in the fact that finite difference methods approximate the differential equation operator and differential equation domain at hand with a difference operator and a finite number of discrete points spread over the original domain, while finite element methods approximate the solution with an expansion of known basis functions and unknown coefficients over the entire domain [9]-[11]. Finite element methods are formulated either directly from the original physical arguments used to derive the differential equations, known as a variational formulation, or based on minimization of weighted inner product residuals, known as a weak formulation. In this sense,

the method of moments and the finite element method are closely related. Generally speaking, finite element methods are more flexible than finite difference methods when dealing with inhomogeneous solution domains, irregularly shaped boundaries, and derivative boundary conditions. On the other hand, finite difference methods are more easily implemented in terms of formulation and programming since they are grid dependent, while finite element methods are coordinate-independent with mapping built into their formulation.

The limitation of the method of moments is its dependency on the availability of an analytic expression of the Green's function. The Green's function satisfies boundary conditions at infinity and at other boundaries such as multi-layered substrate interfaces and the ground plane. Unfortunately, with added complexity, in the form of irregularly shaped boundaries and lossy materials, an integral equation representation of the solution quickly becomes increasingly difficult to derive [12]. Finite difference and finite element schemes do not suffer from such a limitation. On the other hand, moment method solutions of integral equations are typically more efficient than finite difference or finite element solutions to the differential equation of the same problem, and they naturally handle infinite domains. So, whenever dealing with relatively simple passive structures located in infinite domains, the method of moments has the upper hand. For more complex structures, the finite difference or finite element method must be employed [13], [14]. Also, since the principle of superposition does not apply to nonlinear problems, integral equation solutions do not exist, and the method of moments is not applicable to active structures such as transistors. In contrast, since they do not depend on integral equations, finite difference and finite element methods in conjunction with iterative solution techniques offer possibilities to reach the solution in a nonlinear domain.

In referring back to the original problem of modeling an entire MMIC one subregion at a time, it becomes apparent that a choice of numerical solution is beneficial. For subregions containing simple, passive structures such as planar strip-like conductors, the method of moments along with the appropriate integral equation are appropriate. For subregions of higher complexity and/or containing active structures, the finite element method would seem to be the better choice over the finite difference method. The reason for this is that it is possible to exploit the

mathematical similarities of both the method of moments and the finite element method in order to directly couple the finite element solution in a finite subregion to moment method solutions in any adjacent infinite or finite subregions. Taking into account subregion interactions overcomes the problem of neglecting the mutual coupling of MMIC components.

1.2 Objectives and Motivation

For amplifier and digital logic applications, MMIC layouts include transistors which take the form of doped subregions of the semiconductor substrate. This paper focuses on one such active subregion of the MMIC problem and presents a finite element analysis of the governing classical semiconductor device equations. Two different finite element method (FEM) formulations of the steady-state system of coupled Poisson and continuity equations are described, and two different iterative techniques are used to solve for the potential and carrier density distributions inside the active subregion. As a specific example, GaAs MESFETs are the focus of implementation, and the resulting field distributions are used to calculate various MESFET parameters such as electrode currents, voltage gain, capacitances, and conductances. Additionally, techniques based on the MoM that link the finite element domain to any exterior domains are briefly discussed.

To this day, finding a general solution to the tightly coupled set of nonlinear elliptic semiconductor equations has presented quite a challenge to the scientific community. There are many aspects to this challenge. The first and most important one is to be able to ensure robustness of the numerical solution. Since the equations are nonlinear and tightly coupled and since the parameters of semiconductor devices, such as doping density distributions, bias conditions, and geometries, vary widely depending on the device, it is difficult to develop a general solution algorithm that is guaranteed to reach a solution in all cases. Another important aspect is the rate of convergence to the solution. Almost all solution methods implement iterative schemes, and large systems of equations must be solved simultaneously at each iteration. This task demands considerable computing power and speed, and therefore efficiency of the solution algorithm is

important. Equally important is the accuracy of the final solution, and this requirement often forces a trade-off with the speed of convergence. Generally speaking, almost all solutions, the vast majority of which implement either finite element or finite difference schemes, involve the discretization of the semiconductor equations over the domain of the device, which results in a grid of points or a mesh spanning and representing the domain. The trade-off occurs in the coarseness of the mesh. That is, a finer mesh implies more equations, which require longer solution times per iteration, but results in more accurate solutions, while a cruder mesh implies fewer equations, which require less time to solve per iteration, but results in less accurate solutions.

Applications of finite difference techniques have been studied and fine tuned over at least the past forty years in finding solutions to nonlinear elliptic differential equations and about thirty years in finding solutions to the semiconductor equations in particular [15]-[20]. Barnes and Lomax first introduced an FEM solution in [21], although they could not claim superiority of this solution over the highly refined finite difference schemes. Since their work, relatively little has been published regarding the application of the FEM specifically to the semiconductor equations, [22]-[25]. The purpose of this thesis is to expand on the work of Barnes and Lomax in [22] and [23] by first generalizing their FEM approach with compact FEM notation carried over from Zienkiewicz [9] and Buchanan [10]. Following this, a new FEM steady-state formulation referred to as the advection-diffusion model is described. Advantages and disadvantages of both formulations are pointed out, and a final solution algorithm is presented that adopts the best features of both of them. The final algorithm is used to simulate and analyze GaAs MESFET devices similar to those in [22].

1.3 Thesis Structure

Sections 2-4 give a detailed description of the problem and a derivation of two different routes to a solution. The final solution approach offering the best performance involves a combination of two FEM models, the advection-diffusion model and the current conservation model, that are numerically solved by two different techniques, Gummel's method and the

Newton-Raphson method, respectively. Following this in Section 5, implementation considerations, such as stability and accuracy restrictions, of the FEM formulations are addressed. Two simple GaAs MESFET problems are introduced and fully analyzed in order to demonstrate the proposed techniques. Finally, various extensions to the FEM approach involving the application of the method of moments (MoM) are briefly discussed and partially demonstrated in Section 6. These "add-ons" attempt to compensate for the assumptions and simplifications, mainly with respect to the artificial boundary conditions, used in the original stand-alone FEM formulations. Ultimately, the goal of such extensions is to account for the effects of any exterior structures on the interior domain.

2 Problem Specifications

The equations governing potential and carrier density distributions in the active semiconductor regions of an MMIC are the Poisson and continuity equations written as follows [23]:

$$\nabla \cdot \epsilon \nabla \phi = -\rho \quad (1)$$

$$q \partial p / \partial t = -\nabla \cdot \mathbf{J}_p + qG$$

$$q \partial n / \partial t = \nabla \cdot \mathbf{J}_n + qG \quad (2)$$

where

$$\rho = q(N_d + p - n) \quad (3)$$

$$\mathbf{J}_p = q(p\mu_h \mathbf{E} - D_p \nabla p)$$

$$\mathbf{J}_n = q(n\mu_e \mathbf{E} + D_n \nabla n) \quad (4)$$

$$\mathbf{E} = -\nabla \phi \quad (5)$$

and

G = net generation rate

N_d = donor minus acceptor ion densities

ϕ = potential

n, p = electron and hole carrier densities

$\mathbf{J}_n, \mathbf{J}_p$ = electron and hole current densities

D_n, D_p = electron and hole diffusion coefficients

q = electron charge, 1.602×10^{-19} C

ϵ = semiconductor permittivity

μ_e, μ_h = electron and hole mobilities

t = time.

The modeling problem is reduced to a two-dimensional cross section of the region, and steady state conditions are enforced, i.e. it is assumed that no variations of the potential and carrier density in the z coordinate and time dimension exist [19]. Also, the problem is further reduced to that of a majority-carrier device, for which the contribution of minority-carrier transport to the total current is negligible. In the case of GaAs MESFETs, electron transport dominates, and neglecting the hole continuity equation is therefore a valid approximation. Since generation and recombination phenomena depend on the existence of both electrons and holes, as is discussed in most texts on semiconductor theory such as Pulfrey and Tarr [26] and Markowich [2], the generation-recombination term is also omitted in the following analysis of a majority-carrier device. The generation-recombination term introduces a nonlinear coupling between the electron and hole continuity equations, so its omission simplifies the finite element formulation somewhat. Whether or not inclusion of this term would diminish the effectiveness of the FEM approaches described here for simulation of semiconductor devices in general has not been investigated. Regardless, how to incorporate this term into the two formulations described here is mentioned where appropriate. With these simplifications, equations (2) and (3) reduce to

$$0 = \nabla \cdot \mathbf{J}_n \quad (6)$$

$$\rho = q(N_d - n) \quad (7)$$

for electron majority carrier devices. For hole majority carrier devices, a similar expression exists.

For GaAs MESFETs with submicron gate lengths, which are of interest in MMIC technology and which will be the focus of later examples, even small applied bias voltages ($\sim 1-2$ V) generate very large electric fields ($\sim 10^4$ V/cm) thus causing electrons to reach their saturation velocities throughout much of the device. Since the saturation velocity for electrons in GaAs is approximately 10^7 cm/s [26], transit times for electrons from one end of a one-micron-long MESFET to the other is on the order of 10^{-11} s. Such carrier transit times give an indication that operating frequencies of a submicron gate MESFET can reach well within tens of gigahertz, the range of millimeter-wave fields. In fact, it is this high frequency responsiveness along with

their compact size that makes MESFETs so attractive for microwave amplifier applications in MMIC designs [1].

Also, in place of a full transient analysis of the equations, it is reasonable to use a steady-state analysis to approximate the behavior of the dynamic fields at operating frequencies which are typically well below the inverse of the transit times (in this case ~ 100 GHz). Unfortunately, since this is a nonlinear problem, use of modal and frequency analysis in place of the transient analysis is not viable when dealing with higher operating frequencies.

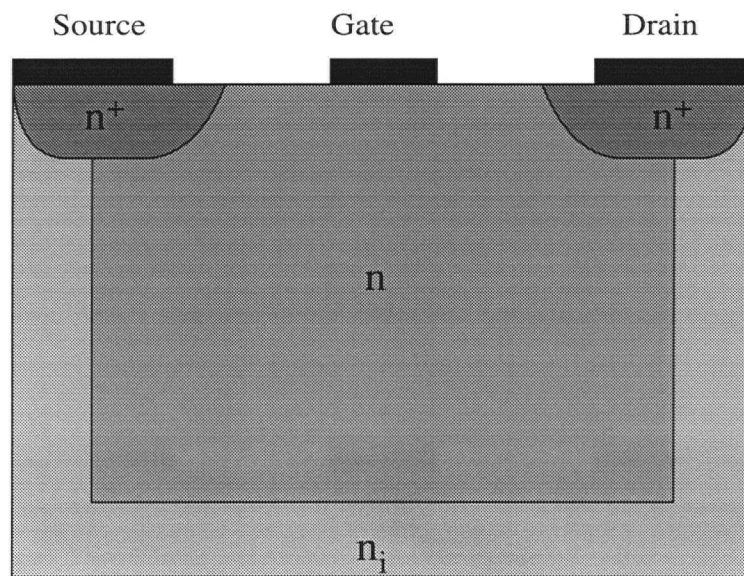


Figure 2. 2D cross section of a MESFET device.

A MESFET device consists of a gate, drain, and source as shown in Figure 2. Under the drain and source metal electrodes, heavily doped ($> 10^{17} \text{ cm}^{-3}$) n^+ regions ensure very thin Schottky barrier-induced depletion regions under the electrodes which are typically dealt with in problems by substituting them with ohmic contacts [2], [23] (See Section 5.3.1). The region under the gate is n-type with doping levels around 10^{15} - 10^{17} cm^{-3} . Surrounding these active regions is a semi-insulating GaAs substrate which ideally has an intrinsic carrier concentration of $2.25 \times 10^6 \text{ cm}^{-3}$. During operation, a current is directed from the drain to the source along a channel around the gate depletion region. By varying the gate voltage, the size of the Schottky barrier depletion region changes determining the amount of current and the operating state of the

device. FETs typically have two operating states, the active regime and the saturation regime. In normally-on, or depletion, devices the depletion region does not reach the buffer zone under equilibrium conditions, while in normally-off, or enhancement, devices this depletion region extends into the substrate buffer at equilibrium. For amplifier applications, a MESFET must be forced into the saturation regime, in which the electron channel is commonly referred to as being "pinched off" between the semi-insulating substrate and the gate depletion region, by applying a negative bias voltage on the gate. This phenomenon will be demonstrated in a depletion GaAs MESFET through simulation using the FEM.

3 FEM Formulation

The finite element method consists of approximating an unknown function, in this case potential and carrier densities, by a low order polynomial with unknown coefficients over a subregion, or element, in the solution domain. All the elements together span the entire domain, and qualitatively speaking, each element interacts with its neighboring elements through their common boundaries. The behavior of the function at the edges of those elements bordering the solution domain is pre-defined by the solution domain's boundary condition specifications. In solving for the polynomial coefficients by way of a matrix inversion, the boundary conditions are, in effect, propagated throughout the solution domain via element boundary interactions as dictated by the governing system equations. A brief discussion of this process is explained in the following sections, and details can be found in various books on finite element analysis such as Zienkiewicz [9], Buchanan [10], and Burnett [11].

3.1 Elements

For two-dimensional problems, an element has a polygonal shape. In practice, triangular and quadrilateral elements are the most popular because they are best suited for approximating boundaries of irregularly shaped domains. In addition to this, formulation of the polynomial approximations for these elements is straightforward as will be shown next.

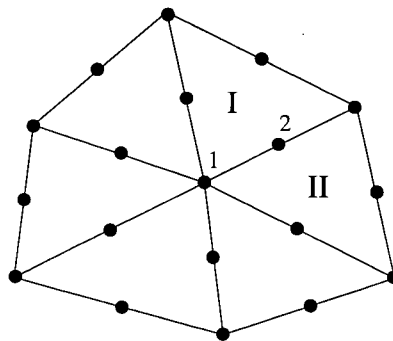


Figure 3. Six adjacent quadratic triangular elements.

Quadratic triangular elements are used in the FEM simulations that will be presented here. Figure 3 shows six such interconnected elements. Each is made up of six points, three at the vertices and the other three about half-way between vertex pairs. Over each element, a function g can be approximated with a function g_a defined as

$$g \approx g_a = \sum_n g_n N_n = [N][g] . \quad (8a)$$

where g_n is the value of the function at node n , N_n is a function of cartesian coordinates x and y for node n , and $[N][g]$ is the matrix representation of the summation. N_n is a basis function commonly referred to as a local shape function in FEM theory. The term “local” implies “per element”. A “global” shape function for a particular node is simply the summation of all the local shape functions for that node. Figure 4 shows a linear (three-node triangular elements) shape function for a node both locally and globally.

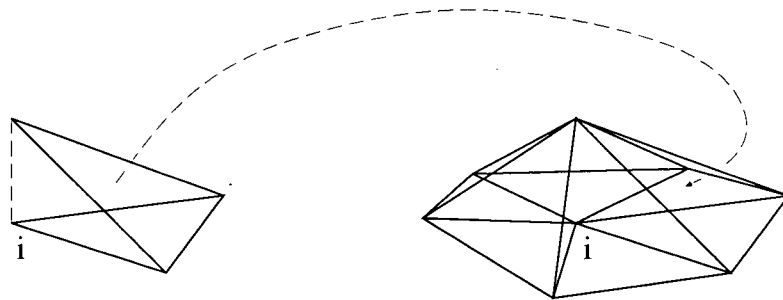


Figure 4. Local (left) and global (right) shape functions for Node i .

From (8a), the gradient of g can be written as

$$\nabla g \approx \nabla g_a = \nabla[N][g] \quad (8b)$$

where ∇ operates on each component of $[N]$. Throughout this paper, $\nabla[N]$ is a two-row matrix with the top row defined as $\partial[N] / \partial x$ and the bottom row defined as $\partial[N] / \partial y$, and $\nabla[N]^T$ is a

two-column matrix with the left column defined as $\partial[N]^T / \partial x$ and the right column defined as $\partial[N]^T / \partial y$. It follows that $\nabla[N]^T = (\nabla[N])^T$.

Derivation of N_n depends on the type of element, and for triangular elements, the shape functions are often defined in terms of area coordinates rather than cartesian coordinates. A basic property of both local and global shape functions is that each N_n has value 1 at node n and value 0 at node m for $m \neq n$. Also, local shape functions for nodes of each element or global shape functions for nodes of all elements must satisfy the following property:

$$\sum_n N_n = 1 \quad , \quad (9)$$

where N_n is local, and $n = 1, 2, \dots$ number nodes per element; or N_n is global, and $n = 1, 2, \dots$ total number of nodes.

For a detailed discussion of area coordinates and the derivation of shape functions for various different elements, the reader is directed to Zienkiewicz [9] and Burnett [11]. As an extension to an exercise in Buchanan [10], a simple, yet general, method for determining shape functions for virtually any type of element is proposed in Appendix A. Although this method was not actually used in the FEM code here, it serves well as a complete and simple explanation of shape functions as an alternative to that which is given in most literature.

A three node linear element with a three term expansion is the simplest triangular element. The higher the order of the element, i.e. the more terms used in (8), the better the unknown function can be represented over that element. Fewer higher order elements than lower order elements are needed to obtain the same desired accuracy in the final approximate solution. On the other hand, lower order elements require fewer nodes, and in either case, the exact solution will be approached as the elements become infinitesimally small. The choice of using quadratic triangular elements over linear triangular elements is based on the fact that they allow for the crudest possible expansion of the mobility gradient term in the advection-diffusion model of the continuity equations, while linear elements lead to a zero-valued mobility gradient term and are therefore unsuitable. A detailed explanation regarding this topic will be given in Section 3.2.2.2.

3.2 Discretization

At this point, either a variational or a weighted residual Galerkin approach is used to complete the FEM formulation. In many cases such as this one, a variational form of the system equations is unavailable, so Galerkin's method is employed here.

Most linear time-independent systems can be represented by an equation $f(g,x,y) = f_1(g,x,y) + f_2(x,y) = 0$ and appropriate boundary conditions, where f_1 is a linear function of g , and f_1 , f_2 , and g are functions of spacial coordinates, x and y . It is desired to determine the unknown function g , so it is replaced by an approximate expansion g_a leaving the expansion coefficients g_n as the only unknowns yet to be determined. Since f_1 is linear with respect to g , $f(g_a,x,y)$ can be written as $\sum_n f_1(N_n,x,y)g_n + f_2(x,y) = [f][g] + f_2(x,y)$, where $[f]$ is a row vector of functions $f_1(N_n,x,y)$, and $[g]$ is a column vector of the unknown coefficients. The residual is defined as $R = f(g_a,x,y) - f(g,x,y) = [f][g] + f_2(x,y)$, and the set of coefficients $[g]$ is sought so as to minimize the residual over the solution domain. This is done by setting weighted inner products of the residual to zero on an element-by-element basis which results in a matrix equation representation of the system equations for each element spanning a subdomain Ω_e . Galerkin's method consists of using the shape functions as the weighting functions, and the result looks like this:

$$\int_{\Omega_e} [N]^T([f][g] + f_2(x,y)) dA = [A_1][g] + [F_1] ,$$

where Ω_e is the region spanned by the element, $[N]^T$ is a column vector of shape functions, matrix $[A_1]$ is called the local stiffness matrix, and $[F_1]$ is the local forcing vector of the element. The j th component of $[g]$ corresponds to the approximate value of g at node j , the i th component of $[F_1]$ corresponds to a force at node i , and the ij th component of $[A_1]$ can be interpreted as a coupling factor, or stiffness, between nodes i and j . A global stiffness matrix and forcing vector representing the whole system are formed by the superposition of the local stiffness matrix and

forcing vector component contributions of all the elements spanning the entire solution domain. For instance, the total, or global, stiffness between nodes 1 and 2 in Figure 3 is equal to the sum of the two local stiffnesses between these nodes which are defined in the local stiffness matrices of elements I and II. The resulting matrix equation is

$$[A][g] = [F],$$

where $[A]$ and $[F]$ are the global stiffness matrix and forcing vector, respectively, and $[g]$ is now a column vector of all the unknown coefficients. After forcing essential boundary condition values on nodes located at domain boundaries in this equation, the solution for the coefficients is given by

$$[g] = [A]^{-1}[F]$$

For construction of $[A]$, numerical quadrature is used to evaluate the inner product integrals. Gauss-Legendre quadrature is best suited for rectangular elements, while numerical formulae based on area coordinates are ideal for triangular elements (See [9], [10], and [11]). For large domains with many elements, $[A]$ becomes very sparse, and a specialized matrix inversion routine that takes advantage of this property should be employed in finding $[g]$. Discretization of the semiconductor equations is described next.

3.2.1 Poisson Equation

Using the notation in (8) to approximate the potential as $[N][\phi]$ and the electron carrier density as $[N][n]$ for each element, and inserting these expressions and (7) into (1) gives

$$\nabla \cdot \epsilon \nabla [N][\phi] = -q(N_d - [N][n]) .$$

Taking the inner products of this expression yields the matrix expression

$$\int_{\Omega_e} [N]^T \nabla \cdot \epsilon \nabla [N] dA [\phi] - q \int_{\Omega_e} [N]^T [N] dA [n] = - q \int_{\Omega_e} [N]^T N_d dA$$

Using the Green-Gauss theorem on the first term, which is essentially integration by parts in more than one dimension, leads to the FEM expression of the Poisson equation:

$$\int_{\Omega_e} \nabla [N]^T \epsilon \nabla [N] dA [\phi] + q \int_{\Omega_e} [N]^T [N] dA [n] = q \int_{\Omega_e} [N]^T N_d dA + \int_{\Gamma_e} [N]^T \epsilon \nabla [N] \cdot \mathbf{n} dS [\phi] \quad (10)$$

where Γ_e is the contour tracing Ω_e , and \mathbf{n} is the unit vector normal to Γ_e . The last term is an integration along the outer contour of the element and represents a Neumann boundary condition on the potential, i.e. a Dirichlet boundary condition on the normal electric field component. This term is zero for insulating boundaries since, by definition, no drift current, which depends on the electric field, can exist across these boundaries. (10) is coupled to the continuity equations, and for each quadratic triangular element, it is a set of six equations in twelve unknowns, $[\phi]$ and $[n]$.

3.2.2 Continuity Equations

Discretization of the electron continuity equation can be done in two ways yielding two different numerical models which are referred to here as the current conservation model and the advection-diffusion model. Each has its advantages and disadvantages in terms of implementation considerations, and the final solution algorithm makes use of both models. Following is a detailed derivation of each model.

3.2.2.1 Current Conservation Model

The current conservation model was first introduced and demonstrated by Barnes and Lomax in [23] and [22], respectively, and is re-derived and rewritten here with compact notation adopted from Zienkiewicz [9] and Buchanan [10]. Approximating the electron carrier density by

$[N][n]$, and inserting this into (4) gives an approximate expression, \mathbf{J}_a , for the net current density:

$$\mathbf{J}_a = q([N][n]\mu_e \mathbf{E} + D_n \nabla[N][n]).$$

Inserting this into (6) gives

$$0 = \nabla \cdot \mathbf{J}_a = q \nabla \cdot ([N][n]\mu_e \mathbf{E} + D_n \nabla[N][n]).$$

Taking the inner product of this expression yields a matrix expression, and applying the Green-Gauss theorem results in the final form:

$$\int_{\Omega_e} [N]^T \nabla \cdot \mathbf{J}_a \, dA = 0$$

$$\int_{\Omega_e} \nabla[N]^T \cdot \mathbf{J}_a \, dA = \int_{\Gamma_e} [N]^T \mathbf{J}_a \cdot \mathbf{n} \, dS \quad (11a)$$

$$q \int_{\Omega_e} \nabla[N]^T \cdot [N] \mu_e \mathbf{E} \, dA [n] + q \int_{\Omega_e} \nabla[N]^T D_n \nabla[N] \, dA [n] = \int_{\Gamma_e} [N]^T \mathbf{J}_a \cdot \mathbf{n} \, dS \quad (11b)$$

This model has the desirable property, from a circuit analysis point of view, that current is conserved regardless of mesh coarseness. This property is made possible by defining the surface integral term on the right as the net current directed out of an element through its nodes (See Section 5.2). For an interior node of a system of elements such as Node 1 in Figure 3, this term is set to zero indicating a zero net global current through the node. The same is done for exterior nodes at insulating boundaries. Thus, interior and insulator nodes obey Kirchhoff's current law, or $I_{in} = 0$ where I_{in} is the net current through each of these nodes. For an exterior node, c , at a non-insulating boundary of the domain, i.e. a contact boundary, there exists a non-zero net current, $I_c = \int N_c \mathbf{J}_a \cdot \mathbf{n} \, dS$, which is the sum of the local surface integral contributions and where N_c is the global shape function for node c . Letting the total current through the contacts be written as $\sum_i I_i$ and using the global form of (11a), an expression for the total contact current can be

derived in the following manner:

$$I_{\text{tot}} = \sum_i I_i = \sum_c I_c + \sum_{\text{in}} I_{\text{in}} = \int \sum_c N_c \mathbf{J}_a \cdot \mathbf{n} \, dS + 0 = \int \nabla \sum_i N_i \cdot \mathbf{J}_a \, dA.$$

where I_{tot} is the total current through all the nodes, and I_i denotes net current through any node.

Substituting the identity of (9) into this expression gives

$$\sum_c I_c = \int \sum_c N_c \mathbf{J}_a \cdot \mathbf{n} \, dS = \int \nabla \sum_i N_i \cdot \mathbf{J}_a \, dA = \int \nabla(1) \cdot \mathbf{J}_a \, dA = 0.$$

So the net current throughout the system is zero, and current is conserved globally.

The surface integral in (11) represents a boundary condition on \mathbf{J}_a , which is the sum of both the drift and diffusion current densities. Again, for insulating boundaries, it has value zero since no current may exist across them.

3.2.2.2 Advection-Diffusion Model

Another formulation of the continuity equations begins with inserting (4) into (6) and expanding the result with the chain rule:

$$\begin{aligned} 0 &= \nabla \cdot \mathbf{J}_n = \nabla \cdot (qn\mu_e \mathbf{E} + D_n \nabla n) \\ &= qn\mu_e \nabla \cdot \mathbf{E} + qn \nabla \mu_e \cdot \mathbf{E} + q\mu_e \nabla n \cdot \mathbf{E} + qD_n \nabla^2 n \end{aligned} \quad (12)$$

This equation is referred to as an advection-diffusion equation in Brooks and Hughes [27] and as the mass transport equation in Buchanan [10], which describes the fate of a concentration of some substance as it flows in some medium and, at the same time, diffuses into the medium. In this case the substance is the electron, and the medium is the conduction band in the semiconductor. The last term represents the diffusion, the third term represents the flow with $\mu_e \mathbf{E}$ as the electron velocity, and the first two terms represent an interaction between the electrons and their

surroundings. Using (1), (3), and (7) and assuming that the permittivity ϵ remains constant over the span of an element allows substitution of $\nabla \cdot \mathbf{E}$ in the first term:

$$0 = qn\mu_e(q(N_d - n)/\epsilon) + qn\nabla\mu_e \cdot \mathbf{E} + q\mu_e\nabla n \cdot \mathbf{E} + qD_n\nabla^2 n \quad (13)$$

This substitution removes some coupling between the continuity equation and the Poisson equation and is the key step in the derivation of the advection-diffusion model of the continuity equations. Now, the first term is nonlinear due to the n^2 term and the mobility μ_e , which is a nonlinear function of the electric field intensity when taking velocity saturation into account. The second and third terms represent a heavy nonlinear coupling to Poisson's equation through both μ_e and \mathbf{E} , while the first term represents a nonlinear coupling to Poisson's equation only through μ_e .

Substituting the carrier density n with $[N][n]$ in (13) and taking the inner products gives

$$\begin{aligned} 0 = & \int_{\Omega_e} [N]^T [N][n] \mu_e (q(N_d - [N][n])/\epsilon) dA + \int_{\Omega_e} [N]^T [N][n] \nabla\mu_e \cdot \mathbf{E} dA \\ & + \int_{\Omega_e} [N]^T \mu_e \nabla [N] \cdot \mathbf{E} dA [n] + qD_n \int_{\Omega_e} [N]^T \nabla^2 [N] dA [n] ; \end{aligned}$$

and applying the Green-Gauss theorem to the last term leaves

$$\begin{aligned} & \int_{\Omega_e} [N]^T [N][n] \mu_e q(N_d - [N][n])/\epsilon dA + \int_{\Omega_e} [N]^T [N] \nabla\mu_e \cdot \mathbf{E} dA [n] \\ & + \int_{\Omega_e} [N]^T \mu_e \nabla [N] \cdot \mathbf{E} dA [n] - qD_n \int_{\Omega_e} \nabla [N]^T \nabla [N] dA [n] \\ & = - qD_n \int_{\Gamma_e} [N]^T \nabla [N] \cdot \mathbf{n} dS [n] \end{aligned} \quad (14)$$

The first two terms of this highly nonlinear equation offer several possibilities in the derivation of an iterative solution method as will be discussed for the application of Gummel's method (See Section 4.2). Some of these variations result in numerical convergence, and others may not.

It is also important to address the mobility gradient $\nabla\mu_e$ in the second term of (14). In

building the stiffness matrix for a given element, $\nabla\mu_e$ is calculated in the following manner: First, with the electric field intensity $|\mathbf{E}| = (\nabla\phi \cdot \nabla\phi)^{1/2}$ where $\nabla\phi = \nabla[\mathbf{N}][\phi]$, the values of $\mu_e(|\mathbf{E}|)$ at each node of the element are approximated and are referred to as $[\mu_e]$; second, with these nodal values and the shape functions of the element, interpolation (See (8).) is used to write an approximate representation of μ_e as $[\mathbf{N}][\mu_e]$; and finally, the expression for $\nabla\mu_e$ is given by $\nabla[\mathbf{N}][\mu_e]$. For a linear triangular element, the electric field, or $-\nabla\phi$, is a constant. It follows that $\nabla\mu_e$ must be zero rendering linear triangular elements inadequate for approximating (15) on an element by element basis. Therefore, six node, quadratic triangular elements are applied to this problem instead, yielding a constant $\nabla\mu_e$ for each element. Of course, the higher the order of the element, the better $\nabla\mu_e$ can be approximated over the span of the element.

The right hand term is a Neumann boundary condition on the carrier density n . Again, for insulating boundaries this term must be zero since any gradient in n would imply a diffusion current, and by definition no current can exist across such boundaries. For some boundaries it may be desired to force Dirichlet conditions on the potential (which implies a non-zero potential flux across the boundary) but also to force a zero net current condition across the boundary. In the current conservation model, this is achieved simply by setting the boundary condition on the current density to zero. To impose such a condition in the advection-diffusion model, the Neumann boundary condition on the carrier density, i.e. the diffusion current condition, must be set equal to the negative of the drift current. In other words, setting (4) to zero gives

$$\begin{aligned} \mathbf{J}_n &= q(n\mu_e\mathbf{E} + D_n\nabla n) = 0 \\ qD_n\nabla n &= -qn\mu_e\mathbf{E} \end{aligned}$$

So

$$- qD_n \int_{\Gamma_e} [\mathbf{N}]^T \nabla[\mathbf{N}] \cdot \mathbf{n} \, dS [n] = - q \int_{\Gamma_e} [\mathbf{N}]^T [\mathbf{N}] \mu_e \nabla[\mathbf{N}][\phi] \cdot \mathbf{n} \, dS [n] \quad (15)$$

for a zero net current across the boundary in the advection-diffusion model.

4 Solution Methodologies

The Poisson equation and electron continuity equation together form a set of coupled, nonlinear equations to which iterative techniques must be applied in order to obtain solutions for $[n]$ and $[\phi]$. First, a qualitative explanation of both direct iteration, of which Gummel's method [15] is a subset, and Newton-Raphson iteration in one variable is given. Following this are descriptions of Gummel's method and the Newton Raphson method in their application to the two semiconductor models derived previously: the current conservation model, (10) and (11), and the advection-diffusion model, (10) and (15). These iterative techniques may or may not successfully converge to solutions depending on the FEM model to which they are applied. Taking advantage of the best features of two combinations of model and iterative scheme, a solution algorithm is proposed that offers good robustness and performance.

4.1 Introduction to Iterative Techniques

The following introduction to iterative techniques is based on a similar discussion in Zienkiewicz [9]. A single variable function, Kx , for which K is a function of x is examined. Finding the values of x that satisfy $Kx + f = 0$, where f is assumed to be constant, can be done using direct iteration or the Newton-Raphson method. Direct iteration begins with an initial guess to the solution x_0 . Next, the first iteration consists of finding a new solution x_1 from

$$x_1 = -f / K(x_0),$$

which hopefully is closer to the actual solution than is x_0 . Repeating this procedure, n iterations are carried out until x_n converges to a particular solution. Gummel's method is a type of direct iteration scheme with the added simplification of decoupling the Poisson and continuity equations as will be seen in Section 4.2. Newton Raphson iteration begins with linearizing the equation

around some initial guess x_0 by way of a truncated Taylor series expansion which gives the expression

$$K(x_0)x_0 + f + \Delta(K(x_0)x_0)\Delta x_0 = 0 \quad ,$$

where $\Delta x_0 = x_1 - x_0$ and $\Delta(K(x_0)x_0)$ is known as the Jacobian which is, in effect, the function that is tangent to Kx at x_0 . The first iteration involves first solving for Δx_0 and then for x_1 which is a new solution and guess for the next iteration. Again, repeating this procedure, n iterations are carried out until x_n converges to a particular solution. Next, performance characteristics of these two schemes on three different types of functions Kx are shown graphically.

Figure 5 shows how successive iterations of each method approach the solution of x for $K_1x = -f_1$. The initial guess is denoted as $x = a_0$.

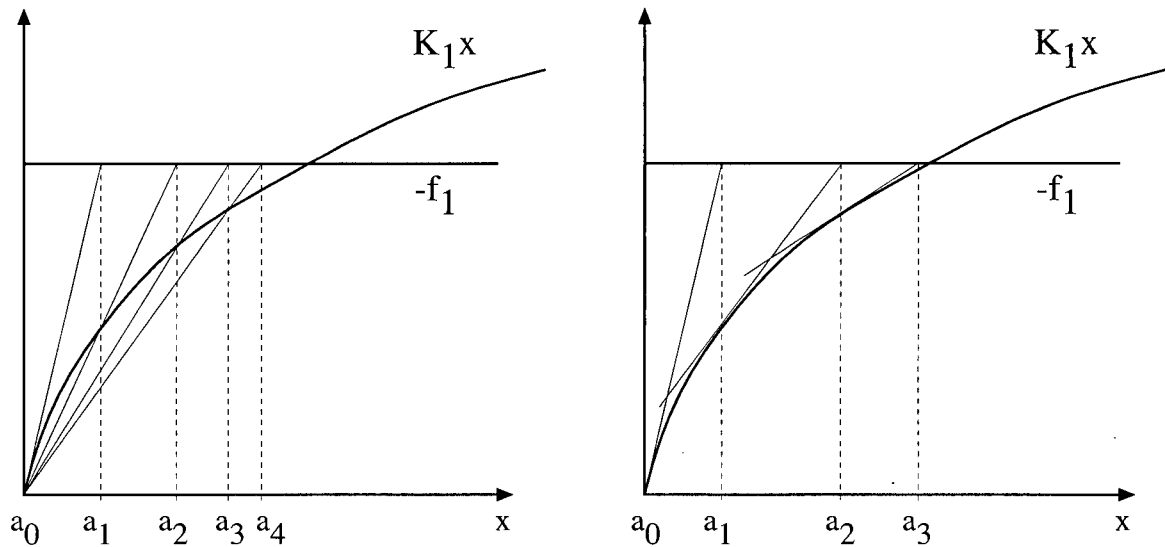


Figure 5. Direct iteration ; Newton-Raphson iteration.

It is clear that the Newton-Raphson method has a much faster convergence rate than direct iteration for this case. Figure 6 shows the progress of a few iterations of both methods in the solution of x for $K_2x = -f_2$.

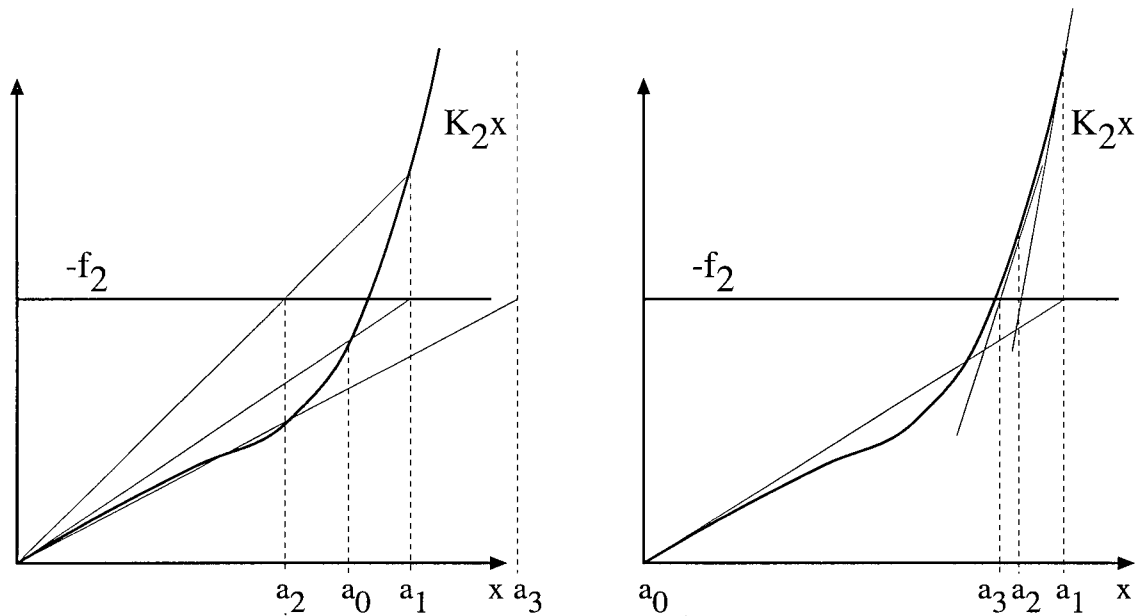


Figure 6. Direct iteration ; Newton-Raphson iteration.

In this case, the Newton Raphson method still converges, but direct iteration does not converge to the solution regardless of the proximity of the initial guess to the exact solution. Finally, Figure 7 shows the behavior of each method for the $K_3x = -f_3$ case. Here, direct iteration converges to the solution, while the Newton-Raphson iterations possibly diverge unless the initial guess is close enough to the exact solution. Ideally in this situation, direct iteration could be used to reach a suitable initial guess after which Newton-Raphson iteration could be applied to rapidly reach the solution. The cases shown here will serve as analogies to the performances of Gummel's method and the Newton-Raphson method applied to the two different FEM models in the following two sections.

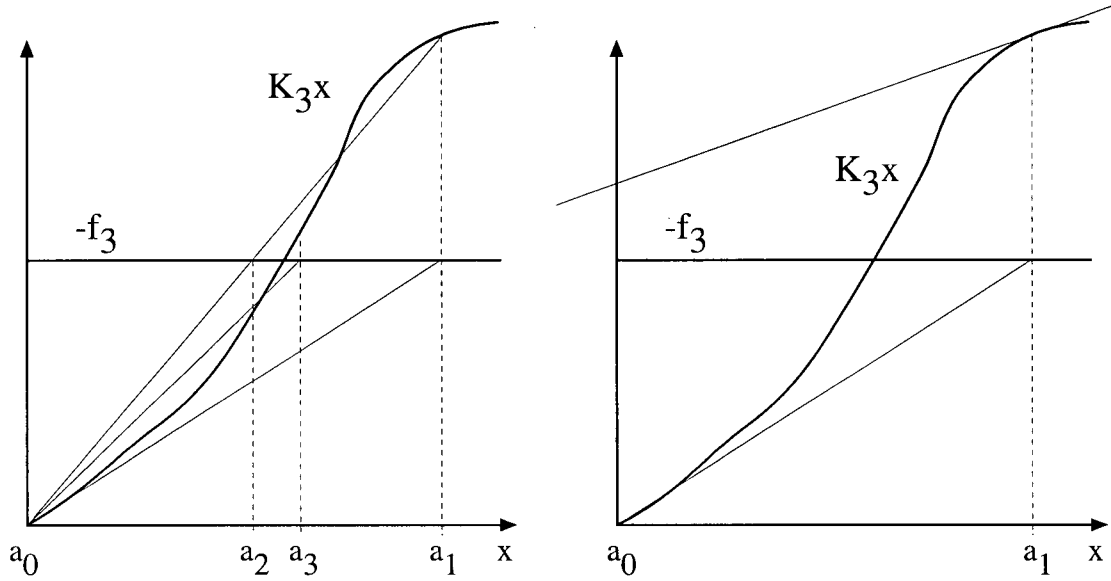


Figure 7. Direct iteration ; Newton-Raphson iteration.

4.2 Gummel's Method

Gummel first introduced his scheme in [15] as a general solution approach to semiconductor equations. It involves decoupling the equations and iteratively solving for $[n]$ and $[\phi]$ independently. Gummel's method is actually a Gauss-Seidel method, which is described in most numerical analysis books such as Froberg [28], with specific application to the semiconductor equations. One iteration consists of two steps: 1) using an initial guess for the carrier density $[n]^i$ in Poisson's equation, find a new potential $[\phi]^{i+1}$ and 2) using $[\phi]^{i+1}$ in the continuity equation, find a new carrier density $[n]^{i+1}$. The index i denotes the iteration number. These two steps are repeated until two consecutive iterations yield a pair of $[n]$'s and $[\phi]$'s with no significant changes between them. Such a change is defined by some norm, and a tolerance level is defined to be the maximum value of change for which the iterations may end (See Section 5.2).

In terms of the iterative notation and leaving out the surface integral term, (10) becomes

$$\int_{\Omega_e} \nabla[N]^T \epsilon \nabla[N] dA [\phi]^{i+1} + q \int_{\Omega_e} [N]^T [N] dA [n]^i = q \int_{\Omega_e} [N]^T N_d dA.$$

Assuming ϵ is constant over each element and letting the doping density N_d be represented by its nodal values as $[N][N_d]$ allows this expression to be written in this compact form

$$[A_l][\phi]^{i+1} = [B_l]_p ([N_d] - [n]^i)$$

where

$$[A_l] = \int_{\Omega_e} \nabla[N]^T \nabla[N] dA ; \quad (16)$$

$$[B_l]_p = (q/\epsilon) [B_l] ; \quad (17)$$

$$[B_l] = \int_{\Omega_e} [N]^T [N] dA . \quad (18)$$

It should be noted that if a continuous mathematical expression for N_d is available, it should not be approximated by $[N][N_d]$ but instead should be employed directly in the integration over the element's domain. Building the global matrices $[A]$ and $[B]_p$ from the local matrix contributions and incorporating any boundary conditions lead to a solution of $[\phi]^{i+1}$:

$$[\phi]^{i+1} = [A]^{-1} [B]_p ([N_d] - [n]^i) . \quad (19)$$

Formulating the continuity equation current conservation model, (11), in the iterative notation allows for three possible configurations (Again, the boundary term is not shown here.):

$$([C_l]^i + [D_l]) [n]^{i+1} = 0 ; \quad (20)$$

$$[C_l]^i [n]^{i+1} = - [D_l] [n]^i ; \quad (21)$$

$$[D_l] [n]^{i+1} = - [C_l]^i [n]^i ; \quad (22)$$

where

$$[D_I] = qD_n[A_I] ; \quad (23)$$

$$[C_I]^i = q \int_{\Omega_e} \nabla[N]^T [N] \mu_e^i \mathbf{E}^i dA = -q \int_{\Omega_e} \nabla[N]^T [N] \mu_e^i \nabla[N] [\phi]^i dA . \quad (24)$$

The number of possible configurations depends on the number of different stiffness matrices available. In this case, the two stiffness matrices $[C_I]$ and $[D_I]$ can be arranged in three different combinations on the left side of the equation as shown in (20), (21), and (22) since ${}_2C_1 + {}_2C_2 = 3$, where ${}_nC_m$ is the number of possible combinations of n elements in groups of m . It should be noted that D_n is considered to be constant over the span of the element, and $[A_I]$ is defined in (16). Putting (20) through (22) in global terms with $[C]$ and $[D]$, incorporating boundary conditions, and solving yields

$$[n]^{i+1} = ([C]^i + [D])^{-1} [F] ; \quad (25)$$

$$[n]^{i+1} = -([C]^i)^{-1}([D] [n]^i + [F]); \quad (26)$$

$$[n]^{i+1} = -[D]^{-1}([C]^i [n]^i + [F]). \quad (27)$$

In the above equations, the forcing vector $[F]$ accounts for the boundary condition contributions to the matrix formulation.

As a quick evaluation, a simple problem consisting of a thin strip of GaAs with a metal contact having zero voltage at one end and an artificial boundary also set to zero voltage at the other end was set up to test the convergence of combinations of any one of (25)-(27) with (19). The doping density is uniform and the strip is partitioned into 100 quadratic triangular elements as shown in Figure 8(a).

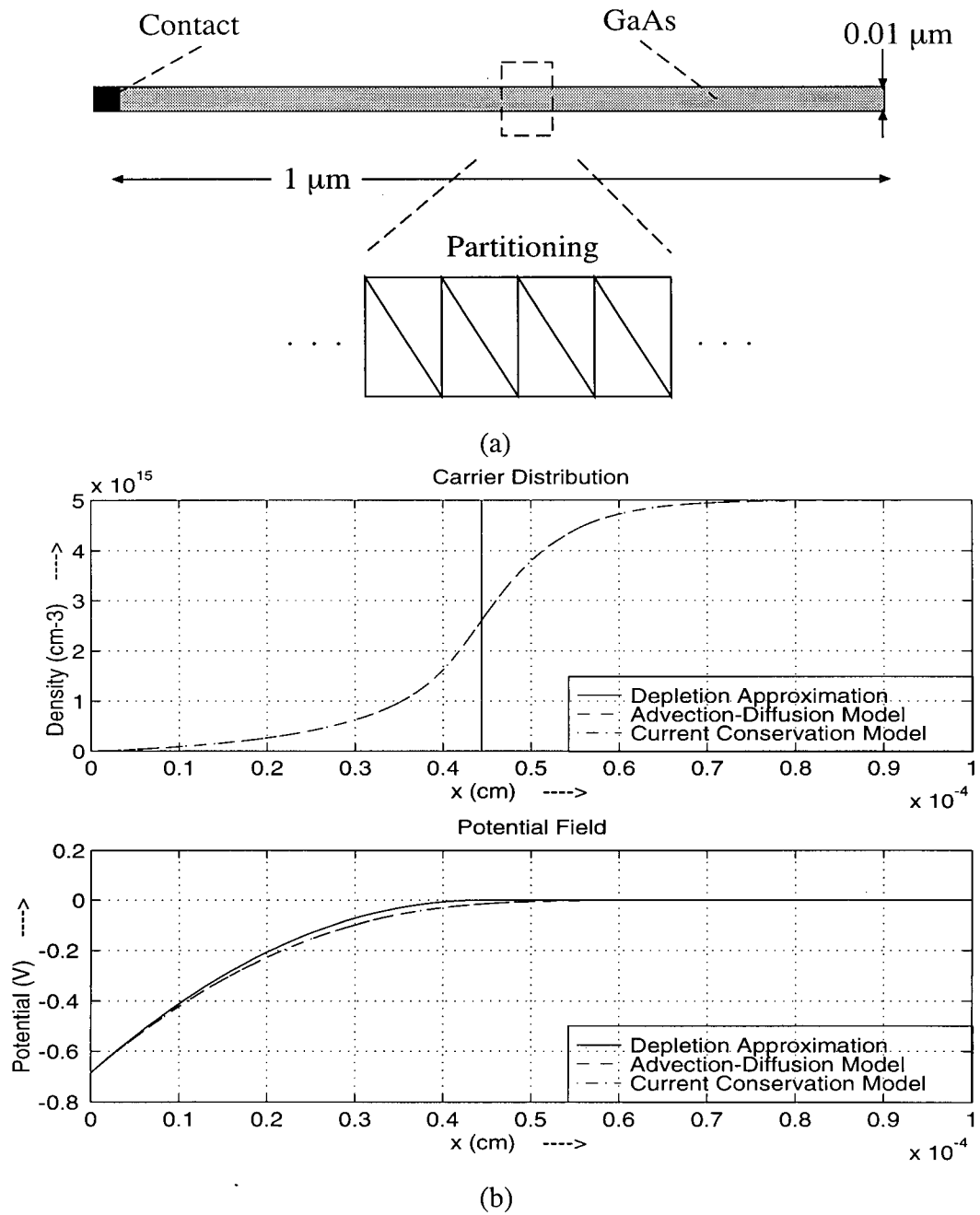


Figure 8. (a) Schottky barrier strip and (b) results for $N_d = 5 \times 10^{15}\ \text{cm}^{-3}$.

The metal contact introduces a negative voltage boundary condition

$$-\psi_s = -(\phi_b - \phi_0) + V, \quad (28a)$$

where ϕ_b is the barrier potential of the semiconductor, $\phi_0 = kT/q \ln(N_c/N_d)$, and V is the voltage

applied to the contact conductor. The barrier potential for GaAs is taken to be 0.8 V [22], and the contact conductor is at zero volts. The voltage boundary condition, in turn, also implies a carrier boundary condition n_s just under the contact in the semiconductor. The relationship between ψ_s and n_s is

$$n_s = N_d e^{-q\psi_s/kT}. \quad (28b)$$

Explanation of (28) is given in Pulfrey and Tarr [26]. Through these boundary conditions, a Schottky barrier will be induced at the contact end of the strip. At the opposite end of the strip, the carrier density is set equal to the doping level N_d and the potential to zero. These are necessary artificial boundaries that simulate boundary conditions at infinity.

In the absence of any educated guess of the carrier density, charge neutrality ($n = N_d$), was used as an initial guess in the first iteration. The (19), (25) combination converged to a solution only when $N_d < 10^{15} \text{ cm}^{-3}$, otherwise no combinations of any one of (25)-(27) and (19) reached convergent solutions with higher doping levels, $N_d > 10^{15} \text{ cm}^{-3}$. In fact, even when the initial guess of the carrier density was very close to the actual solution, all these combinations diverged from the solution. Thus, it is concluded that the current conservation model solved using Gummel's method is useless for any practical problems!

Putting the advection-diffusion model, (14), which has four different stiffness matrices, into Gummel's iterative form results in fifteen ($4C_1 + 4C_2 + 4C_3 + 4C_4 = 15$) different possible configurations, as opposed to the three possible configurations for the current conservation model shown in (20)-(22) and discussed earlier. After applying each of these configurations with (19) to the test problem defined earlier and shown in Figure 8(a), it was found that only two forms lead to converging iterations regardless of N_d . The first is given as

$$\begin{aligned} ([E]_i + [F]_i + [G]_i + [D]_i) [n]^{i+1} &= [G]_i [N_d] ; \\ [n]^{i+1} &= ([E]_i + [F]_i + [G]_i + [D]_i)^{-1} [G]_i [N_d] ; \end{aligned} \quad (29a)$$

and the second is slightly different:

$$\begin{aligned} ([E_1]^i + [G_1]^i + [D_1]) [n]^{i+1} &= [G_1]^i [N_d] - [F_1]^i [n]^i ; \\ [n]^{i+1} &= ([E]^i + [G]^i + [D])^{-1} ([G]^i [N_d] - [F]^i [n]^i) ; \end{aligned} \quad (29b)$$

where the corresponding local stiffness matrices are

$$[E_1]^i = q \int_{\Omega_e} [N]^T \mu_e^i \nabla [N] \cdot \nabla [N] [\phi]^i dA ;$$

$$[F_1]^i = q \int_{\Omega_e} [N]^T [N] \nabla \mu_e^i \cdot \nabla [N] [\phi]^i dA ;$$

$$[G_1]^i = (q^2/\epsilon) \int_{\Omega_e} [N]^T [N] [n]^i \mu_e^i [N] dA ;$$

and $[D_1]$ is defined in (23). Note that ϵ is assumed to be constant over the span of the element, \mathbf{E} has been replaced with $-\nabla[N][\phi]^i$, and $[A_1]$ is defined in (16). It is also important to note that for the case of a zero net current boundary condition in (29), (15) contributes to the stiffness matrix, not the forcing vector!

Both combinations of (19) and one of (29a) and (29b) are very robust, and in all cases (varying N_d and contact voltages) tested on the strip, charge neutrality gave a sufficiently good initial guess of the carrier density to ensure convergence in the Gummel iterations. (29b) tended to reach a solution in slightly fewer iterations as compared to (29a). On the other hand, this scheme is more computationally intensive since $[E_1]^i$, $[F_1]^i$, and $[G_1]^i$ must be recalculated for each iteration. Although it has not been done here, inclusion of a generation-recombination term would create an additional nonlinear term as a function of carrier densities that would also have to be recalculated for each iteration.

Tan et al. [29] introduce an application of Scharfetter-Gummel (S-G) interpolation in conjunction with a seven-point Gaussian quadrature for the construction of a stiffness matrix and forcing vector for triangular elements that lead to a higher accuracy and convergence rate of the

solution when compared to the classical Gummel finite element formulation. The S-G interpolation is only used when applying the quadrature on the potential and carrier values of the previous iteration and therefore requires relatively minor modification to the original classical formulation in a computer program. Application of this scheme was briefly tested on the strip problem as well as the MESFET examples in Section 5.2, but there was no noticeable improvement in the accuracy or convergence rate of the solution. Nevertheless, [29] shows that the Scharfetter-Gummel scheme really shines for higher doping conditions, and due to its easy implementation in already existing FEM formulations that apply Gummel's method, it is well worth mentioning here.

For later reference, the current conservation Gummel approach and the advection-diffusion Gummel approach (using (29b)) are abbreviated as CCG and ADG, respectively.

4.3 Newton-Raphson Method

The Newton-Raphson method applied to the semiconductor equations consists of iteratively solving the Poisson and continuity equations simultaneously and implicitly. Its application to solving the current conservation model is derived here. The general form of a discretized nonlinear system can be written as

$$[K([a])] = [P([a])][a] + [F([a])] = 0 \quad (30)$$

where $([a])$ indicates that the matrix or vector is a function of the coefficients $[a]$. Linearizing $[K([a])]$ by way of a truncated Taylor series expansion around some point $[a]^i$ gives

$$[K([a])] = 0 \approx [K([a]^i)] + [K([a]^i)]_T \Delta[a]^i .$$

Defining $[a]^i$ as the result of iteration i allows an iterative process to be described by

$$\Delta[a]^i = - ([K([a]^i)]_T)^{-1} [K([a]^i)] , \quad (31)$$

$$[a]^{i+1} = [a]^i + \Delta[a]^i \quad (32)$$

where

$$[K([a])]_T = d[K([a])] / d[a] . \quad (33)$$

(31) and (32) describe one iteration of the Newton-Raphson procedure. Also, it is important to note that Dirichlet boundary conditions are applied to (31) by enforcing a no-change condition for a boundary node value, i.e. $\Delta a_j = 0$ where j refers to a Dirichlet boundary node. The core of this method is the Jacobian or tangential matrix $[K([a]^i)]_T$. Using (30), (33), and the chain rule leads to

$$[K([a])]_T d[a] = d[K([a])] = [P([a])] d[a] + d[P([a])] [a] + d[F([a])] \quad (34)$$

Rewriting the last two terms as

$$d[P([a])] [a] = [X([a])] d[a] \quad (35)$$

$$d[F([a])] = [Y([a])] d[a] \quad (36)$$

and substituting into (34) yields an expression for $[K([a])]_T$:

$$[K([a])]_T = [P([a])] + [X([a])] + [Y([a])] . \quad (37)$$

Combining the Poisson and current conservation continuity models, (10) and (11), into one local matrix representation gives

$$[K([a])] = [P([a])][a] + [F([a])] = \begin{bmatrix} [A_1] & [B_1]_p \\ 0 & [C_1] + [D_1] \end{bmatrix} \begin{bmatrix} [\phi] \\ [n] \end{bmatrix} + \begin{bmatrix} [B_1]_p [N_d] \\ 0 \end{bmatrix} = [0] \quad (38)$$

where $[A_1]$, $[B_1]_p$, $[C_1]$, and $[D_1]$ are defined in (16), (17), (24), and (23), respectively. In this case the $[F([a])]$ term is not actually a function of $[a]$, and it follows from (36) that $[Y([a])] = 0$. $[P([a])]$ on the other hand has the drift term $[C_1]$ which is a function of $[\phi]$, thus

$$d[P([a])][a] = \begin{bmatrix} 0 & 0 \\ 0 & d[C_1] \end{bmatrix} \begin{bmatrix} [\phi] \\ [n] \end{bmatrix} \quad (39)$$

So $d[C_1][n]$ must be evaluated in order to find $[X([a])]$ as defined by (35). Following are the details of a full derivation of $d[C_1][n]$ into the form, $[\Omega_1] d[\phi]$:

Through (24) and the chain rule

$$d[C_1][n] = -q \int_{\Omega_e} \nabla[N]^T [N][n] (d\mu_e) \nabla[N][\phi] dA - q \int_{\Omega_e} \nabla[N]^T [N][n] \mu_e \nabla[N] d[\phi] dA \quad (40)$$

Now $d\mu_e$ is evaluated:

$$\begin{aligned} d\mu_e &= (d\mu_e / d[\phi]) d[\phi] = (d\mu_e / d|E|) (d|E| / d[\phi]) d[\phi] \\ &= (d\mu_e / d|E|) (d(\nabla\phi^T \nabla\phi)^{1/2} / d[\phi]) d[\phi] \\ &= (d\mu_e / d|E|) (d(\nabla\phi^T \nabla\phi)^{1/2} / d(\nabla\phi)) (d(\nabla\phi) / d[\phi]) d[\phi] \\ &= (d\mu_e / d|E|) ((\nabla\phi^T \nabla\phi)^{-1/2} \nabla\phi^T) (d(\nabla[N][\phi]) / d[\phi]) d[\phi] \\ &= (d\mu_e / d|E|) (\nabla\phi^T / |E|) (\nabla[N] d[\phi] / d[\phi]) d[\phi] \\ &= (d\mu_e / d|E|) |E|^{-1} (\nabla[N][\phi])^T \nabla[N] d[\phi] \end{aligned}$$

Substituting this result into (40) gives

$$d [C_1][n] = -q \int_{\Omega_e} \nabla[N]^T \cdot [N][n] \left((d\mu_e / d|E|) |E|^{-1} (\nabla[N][\phi])^T \nabla[N] d[\phi] \right) \nabla[N][\phi] dA \\ - q \int_{\Omega_e} \nabla[N]^T \cdot [N][n] \mu_e \nabla[N] d[\phi] dA .$$

Rearranging terms leaves

$$d [C_1][n] = [\Omega_1] d[\phi] \\ = -q \int_{\Omega_e} \nabla[N]^T \cdot \nabla[N][\phi] [N][n] (d\mu_e / d|E|) |E|^{-1} (\nabla[N][\phi])^T \nabla[N] dA d[\phi] \\ - q \int_{\Omega_e} \nabla[N]^T \cdot [N][n] \mu_e \nabla[N] dA d[\phi] .$$

Now, (39) can be rewritten as

$$d[P([a])] [a] = [X([a])] d[a] = \begin{bmatrix} 0 & 0 \\ 0 & d[C_1] \end{bmatrix} \begin{bmatrix} [\phi] \\ [n] \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ [\Omega_1] & 0 \end{bmatrix} \begin{bmatrix} d[\phi] \\ d[n] \end{bmatrix} .$$

It follows from (37) that

$$[K([a])]_T = [P([a])] + [X([a])] + [Y([a])] = \begin{bmatrix} [A_1] & [B_1]_p \\ [\Omega_1] & [C_1] + [D_1] \end{bmatrix} . \quad (41)$$

(38) and (41) are used to build the global Newton-Raphson formulation as defined by (31) and (32):

$$\begin{bmatrix} \Delta[\phi]^{i+1} \\ \Delta[n]^{i+1} \end{bmatrix} = \begin{bmatrix} [A] & [B]_p \\ [\Omega]^i & [C]^i + [D] \end{bmatrix}^{-1} \left(\begin{bmatrix} [A] & [B]_p \\ 0 & [C]^i + [D] \end{bmatrix} \begin{bmatrix} [\phi]^i \\ [n]^i \end{bmatrix} + \begin{bmatrix} [B]_p [N_d] \\ 0 \end{bmatrix} \right) \\ \begin{bmatrix} [\phi]^{i+1} \\ [n]^{i+1} \end{bmatrix} = \begin{bmatrix} [\phi]^i \\ [n]^i \end{bmatrix} + \begin{bmatrix} \Delta[\phi]^{i+1} \\ \Delta[n]^{i+1} \end{bmatrix} \quad (42)$$

Note that only $[C]$ and $[\Omega]$ must be computed for each iteration. On the other hand for an N -node problem, $2N$ equations must be solved simultaneously at each Newton-Raphson iteration as

opposed to having to solve two sets of N equations at each Gummel iteration. This places more demand on computing power both in terms of memory and speed. It should also be noted that any nonlinear term, possibly to account for generation-recombination in the continuity equations, would follow the same basic procedure as that of the nonlinear drift term for determination of $[K([a])]_T$.

This Newton-Raphson current conservation technique, (42), does converge to a solution given a "suitable" initial guess, as opposed to the implementation of Gummel's method to the same model which, for doping levels beyond a relatively low point, diverges from the solution regardless of the initial guess (See Section 4.2). Unfortunately, using charge neutrality and zero potential everywhere as an initial guess most often leads to diverging results and is therefore not "suitable". On the other hand, for a good initial guess (42) converges very rapidly to the solution. In fact, Newton-Raphson methods converge quadratically near the solution, while direct iteration methods converge only linearly as the solution is approached [28]. Combining the robustness of the Gummel approach to the advection-diffusion model with the efficiency of the Newton-Raphson approach to the current conservation model yields a good system for solving the coupled Poisson-continuity equations. As it turns out, relatively few iterations of Gummel's method are required to achieve an appropriate initial guess for which the Newton-Raphson method can then be implemented to quickly converge to the solution (See Section 5.2). Applying this procedure to the strip test of the previous section requires a total of about seven iterations (four Gummel iterations and three Newton-Raphson iterations) as opposed to nine iterations required when applying only ADG. Figure 8(b) shows the carrier density distribution and potential field solutions to the strip problem for both models as well as analytic solutions based on the depletion approximation which is discussed in most semiconductor theory texts [26], [2]. Throughout the remainder of this paper, the Newton-Raphson current conservation approach and the combined Newton-Raphson current conservation and Gummel advection-diffusion approach are referred to as CCNR and ADCC, respectively.

Applying the Newton-Raphson method to the advection-diffusion model was not considered here because there is no apparent advantage it could offer. On observation of (14), it is

clear that a Newton Raphson approach would be quite computationally intensive, much more so than either the Gummel approach to the advection-diffusion equations (ADG) or even CCNR, due to the many nonlinear terms that would have to be calculated at each iteration. In other words, even for a best case scenario in which it offered better convergence properties than CCNR, its computational inefficiency would be more than enough to diminish its usefulness. Also, the current conservation model has the nice property of current conservation which the advection-diffusion model does not.

Qualitatively speaking, the current conservation model's stiffness matrix is analogous to K_2 in Section 4.1, and Figure 6 shows iteration behavior analogous to that of the CCG and CCNR approaches when applied to the Schottky barrier strip test. Also, the advection-diffusion model's stiffness matrix is analogous to either K_1 or K_3 for which Gummel's method proves to be very robust. The ADCC scheme ensures convergence by first applying direct iteration (Gummel's method) to the K_1 or K_3 case (advection-diffusion model) until a reasonable initial guess is achieved before implementing Newton-Raphson iterations on the K_2 case (current conservation model) for faster convergence.

5 Application

Examples of application of the previously described ADG and ADCC methods are demonstrated here. In each example, the methods are used to compute the potential field and carrier distribution throughout a particular MESFET device. From these results various electronic properties of the device can be derived such as currents, conductances, and capacitances. Example 1 demonstrates current and small signal conductance calculations based on the analyses in [22]. This example is also used to compare the performances and results of ADG and ADCC. Example 2 is a full analysis of a MESFET which describes, in detail, the derivation of a small signal circuit model from the potential field and carrier distribution computed with ADCC. Preceding the examples, in the following section, certain restrictions regarding the mesh coarseness of the FEM problem are specified that ensure, to a certain degree, convergence and accuracy of the computed solution.

5.1 Mesh Spacing

For numerical stability and physically meaningful results, general statements can be made concerning restrictions on mesh coarseness. The Debye length L_D is the characteristic decay length of charge in a semiconductor (See [26] and [2]) and is given by

$$L_D = (kT\epsilon / q^2 N_d)^{1/2} .$$

Qualitatively speaking, if the lower order polynomial expansion of an element that has dimensions of the order h is to reasonably approximate a decaying charge distribution, it is safe to say that h should not be larger than the Debye length. This restriction is necessary to achieve physically meaningful results. The phrase “dimensions of the order h ” refers to the approximate distance between any two nodes of an element.

For numerical stability, another limitation on the size factor h of an element is Reynold's number:

$$R_c = vh/D \leq 2,$$

where v is the velocity and D is the diffusivity of the carriers. In mathematics, this is also referred to as Peclet's number and is of importance in flow problems in general. Attempts have been made [27] to remove this restriction by modifying the weighting functions used in the standard FEM formulation; these are known as Petrov-Galerkin techniques. Put simply, the standard Galerkin FEM discretization (as demonstrated in this paper) can be shown to, by its very nature, add artificial negative diffusion to flow equations which, when combined with large drift constants, i.e. a large Peclet number, may lead to spurious node-to-node oscillations or "wiggles" in the computed solution. These wiggles are strictly numerical artifacts and, in the case of nonlinear flow problems, may even degrade the convergence and stability of an iterative solution scheme. The Petrov-Galerkin FEM discretization "solves" this problem essentially by adding positive artificial diffusion to cancel the negative artificial diffusion in the standard formulation. Unfortunately, such diffusive schemes lower the order of accuracy of the numerical solution from second order to first order. In fact, it is questionable to use such wiggle-suppression schemes at all (See Gresho and Lee [29].), and if wiggles do appear, simple local mesh refinement is probably the safest remedy.

Fortunately, in the case of GaAs the drift velocity of the electron carriers cannot exceed approximately 2×10^7 cm/s (barring velocity overshoot or any other quantum phenomena [26]), so Reynold's number does not dictate the mesh fineness for the submicron gate MESFET problem as much so as does the Debye length.

5.2 Simulations

Following are two examples of FEM simulations for submicron gate MESFETs. Similar to the demonstrations in [22], the first example briefly analyzes a MESFET for the main purpose of

comparing the performances of the ADG and ADCC schemes. The second describes a complete analysis of another MESFET with a thinner active region of higher doping concentration in which conductances and capacitances of the device at a certain bias are predicted. In both examples, doping levels are relatively low in comparison to practical scenarios. This is done in order to keep the number of nodes at "reasonable" levels for implementation on a conventional workstation. The FEM algorithm is programmed entirely in Professional MATLAB v4.2, and all FEM simulations are carried out on a Sun SPARC 10 workstation. Before proceeding to the examples, some details concerning current calculations, norms, and the electron mobility expression are specified here.

As in Barnes and Lomax [23] following the ideology of the current conservation model, currents $[I]$ and charges $[Q]$ travelling through and existing on all nodes are expressed conveniently as

$$[I] = \int_{\Gamma} [N]^T \mathbf{J}_a \cdot \mathbf{n} \, dS = ([C] + [D]) [n] \quad (43)$$

and

$$[Q] = \int_{\Gamma} [N]^T \epsilon \nabla [N] \cdot \mathbf{n} \, dS [\phi] = \epsilon [A] [\phi] + \epsilon [B]_p [n] - q \int_{\Omega} [N]^T N_d \, dA \quad (44)$$

where local stiffness matrices $[C]$, $[D]$, $[A]$, and $[B]_p$ are defined in (23), (24), (16), and (17), respectively. These representations simply state that a node's current and charge are the accumulations of weighted normal component contributions of electric current density and electric flux density, respectively, over the portion of the boundary, Γ , spanned by that node's global shape function. It should be clear that these definitions satisfy the integral form of Maxwell's equations [30] for steady-state current and charge as node spacing becomes infinitesimal, i.e.

$$\int_{\Gamma} \mathbf{J} \cdot \mathbf{n} \, dS = I \quad \text{and}$$

$$\int_{\Gamma} \mathbf{D} \cdot \mathbf{n} \, dS = Q .$$

For comparing node potentials resulting from the last two iterations, a scaled Euclidean norm is used:

$$\Delta\phi = (1/\max([\phi]^i)) \cdot (\sum_n (\phi_n^i - \phi_n^{i-1})^2)^{1/2}$$

For comparing node carrier densities from the last two iterations, the following norm is used:

$$\Delta n = (\sum_n (1 - n_n^{i-1}/n_n^i)^2)^{1/2}$$

In contrast to the scaled Euclidean norm, this one places more importance on local nodal changes regardless of the values of the carrier densities at those nodes. Since carrier densities vary on exponential scales, this is a more suitable norm. In the above definitions, i is the iteration number, and n is the node number

As in [22], the following expression is used for the nonlinear mobility term:

$$\mu_e = (\mu_0 + v_s |E|^3 / |E_0|^4) / (1 + |E|^4 / |E_0|^4) , \quad (45)$$

where μ_0 is the low field intensity electron mobility, v_s is the electron saturation velocity, and $|E|$ is the electric field intensity. The constants are chosen as $\mu_0 = 5000 \text{ cm}^2 \text{ s}^{-1} \text{ V}^{-1}$, $v_s = 0.85 \times 10^7 \text{ cm s}^{-1}$, and $|E_0| = 4000 \text{ V cm}^{-1}$. Figure 9 shows a plot of the electron drift velocity with respect to electric field intensity.

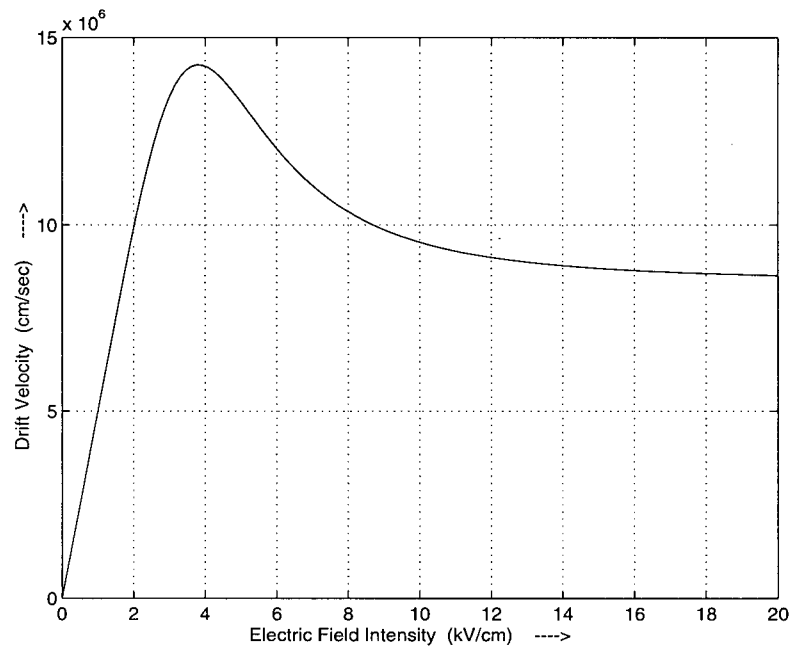


Figure 9. Electron drift velocity versus electric field intensity.

5.2.1 Example 1

Figure 10 shows the configuration of the cross section of the GaAs MESFET to be analyzed. The cross section is $1 \times 2 \mu\text{m}$, and the gate length is $0.4 \mu\text{m}$. There is an active region with a uniform doping concentration of $N_d = 5 \times 10^{15} \text{ cm}^{-3}$, and a buffer zone with a uniform doping concentration of $N_d = 1 \times 10^{13} \text{ cm}^{-3}$ and thickness $0.2 \mu\text{m}$. In order to reduce the number of necessary nodes, the Schottky barriers under the source and drain are neglected, and ohmic contacts are assumed in place of heavily doped wells that have the same effect.

The mesh used for this problem was generated using a semi-automatic mesh generation algorithm based on Zienkiewicz and Phillips [31] and is shown in Figure 11. As can be seen, the mesh is finer near the contact edge singularities and underneath the gate since these are the regions of maximal gradients in carrier density and potential field.

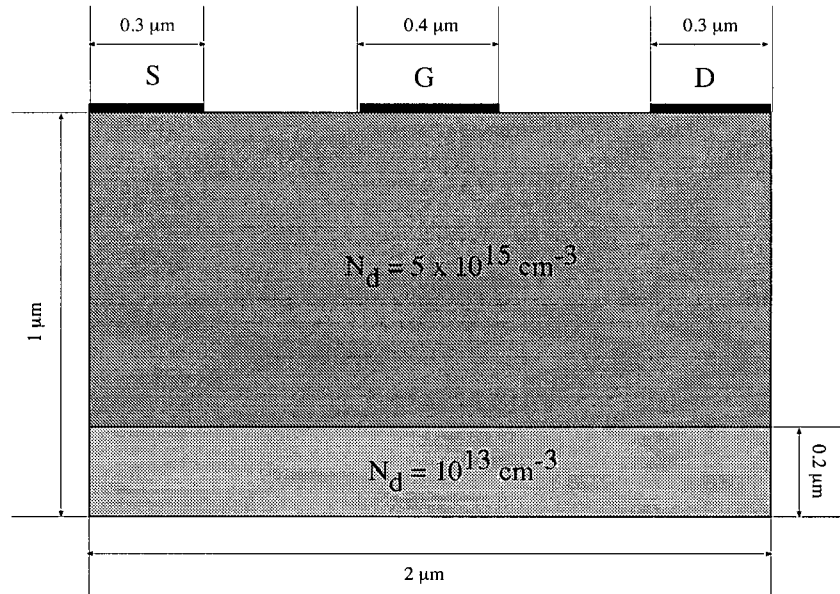


Figure 10. MESFET device for Example 1.

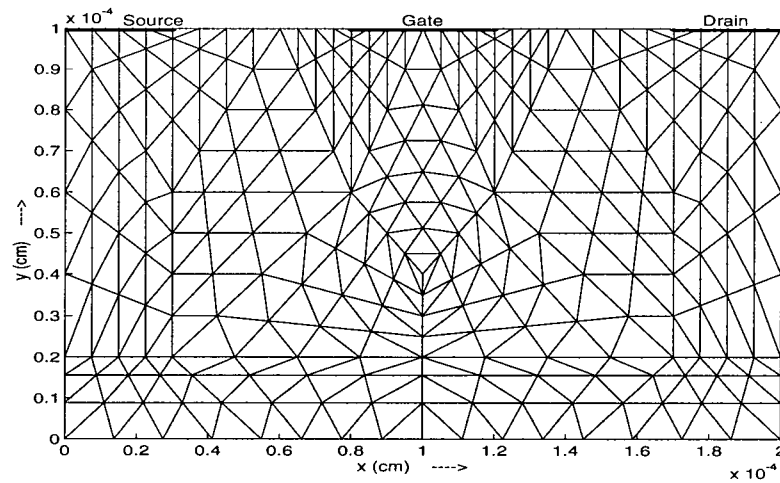


Figure 11. Mesh Grid for Example 1 - 959 nodes, 448 elements.

Dirichlet boundary conditions exist along the contacts and just below the buffer zone. The boundary potentials at the source and drain (ohmic contacts) are set equal to the voltage on the source and drain contacts, V_d and V_s , while the boundary potential along the gate boundary is set to -0.6852 V using (28a). $V_s = 0$ V is used as the reference voltage throughout this example. The boundary carrier densities at the source and drain contacts are set equal to the doping level forcing charge neutrality as required for ohmic contacts, while the carrier density n_s just under the gate

contact is $1.62 \times 10^4 \text{ cm}^{-3}$ satisfying (28b). The boundary under the buffer zone is given an artificial condition of zero potential and charge neutrality to approximate the effects of a deep substrate. Of course, the artificial charge neutrality condition results in “leaking” current through this artificial boundary, and current conservation is not preserved. This is done intentionally in order to keep the number of ADG iterations low. It turns out that inclusion of the term (15), which is required to specify zero net current across a boundary in the advection-diffusion model, increases the number of iterations necessary to reach a solution with ADG, in this case by a factor of four.

Neumann boundary conditions are enforced on the potential and carrier density at the semiconductor-air borders between the contacts, and along the left and right edges of the device. These boundary values are zero suggesting no current, diffusion and drift, across the boundaries, i.e. perfect insulation. Applying these conditions is accomplished simply by neglecting the surface integral terms in the FEM formulations for both ADG and ADCC.

Figures 11 and 12 show the potential field and electron carrier distribution calculated with the ADCC method for $V_s = V_g = 0 \text{ V}$ and $V_d = 3 \text{ V}$. The Schottky barrier spans most of the region under the gate leaving only a thin channel of a high concentration of electrons just underneath. The channel is “pinched off” at the drain side by a depleted region of carriers.

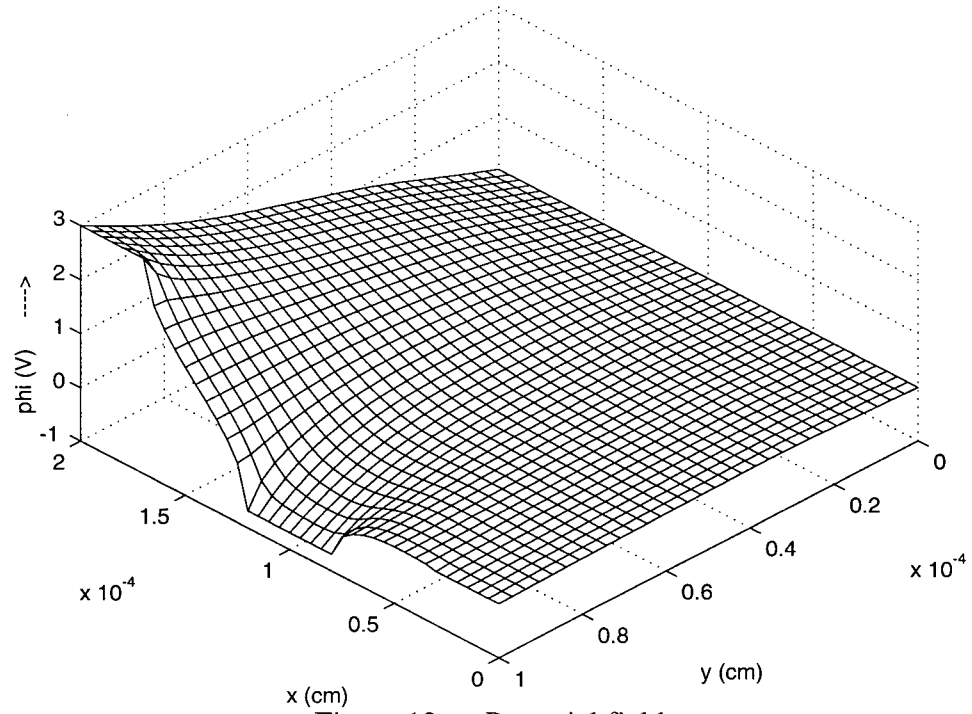


Figure 12. Potential field.

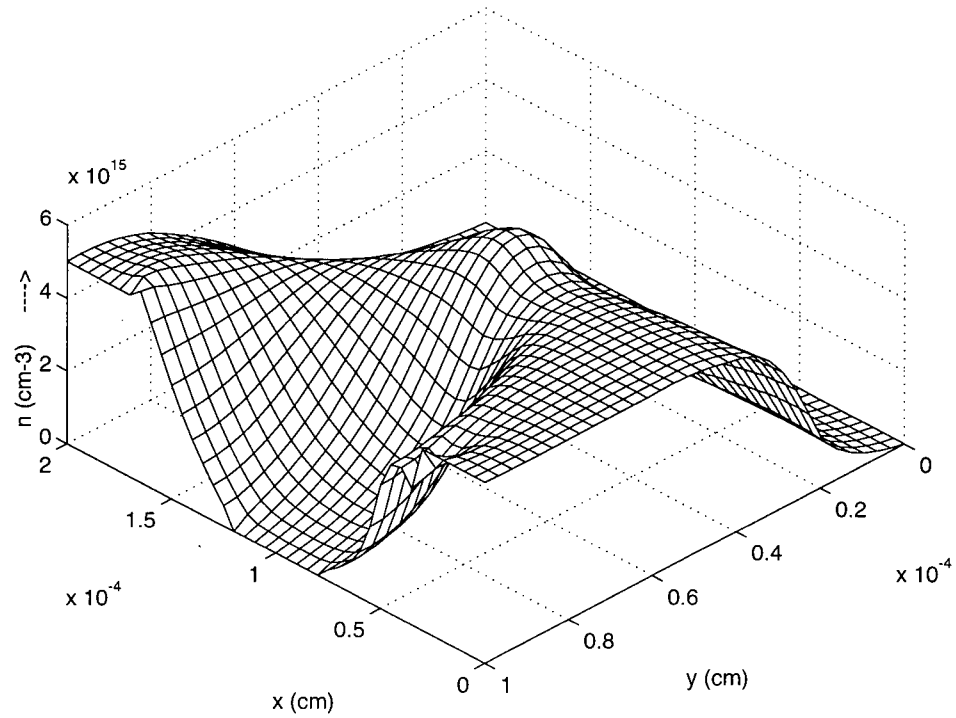


Figure 13. Electron carrier distribution.

Figure 14 shows a plot of the current density field, and it is clear that the concentration of current is greatest along the narrow channel underneath the barrier. Also, after closer inspection of Figure 13, a “bump” indicating an accumulation of carriers above the $5 \times 10^{15} \text{ cm}^{-3}$ level exists along this channel. This bump and the adjacent valley of carriers constitute what is known as a stationary or Gunn domain, and is a result of velocity saturation effects as predicted by one-dimensional FET analysis [26]. As the drain voltage is increased further in the saturation regime, the stationary domain absorbs the bulk of the extra potential difference along the channel.

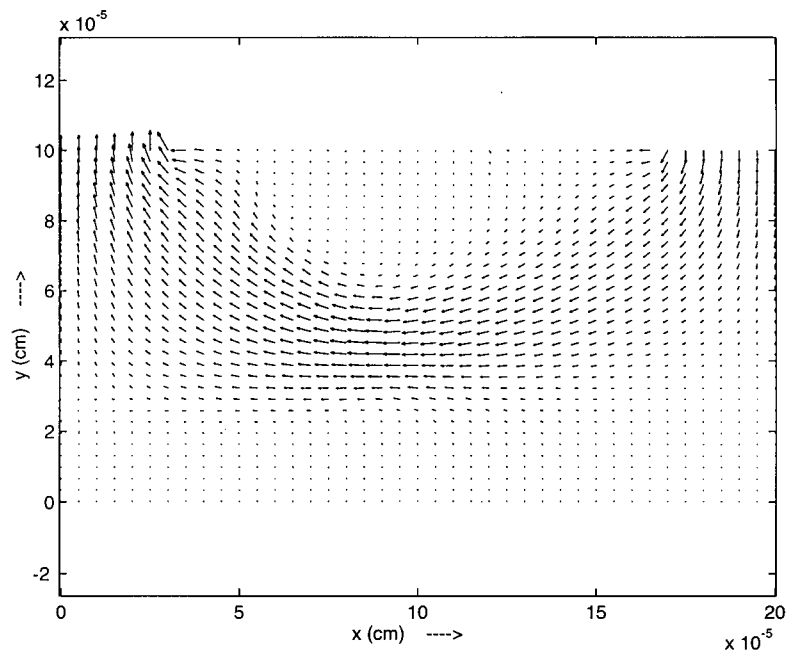


Figure 14. Current density field.

Simulations were carried out for varying gate and drain voltages. A plot of I_D vs. V_D , for $V_g = 0, 0.5 \text{ V}$, is given in Figure 15. As expected, there is an active regime and a saturation regime of operation. There is also a region between these that shows negative resistance. Such regions are typically avoided in the design of amplifiers since they may cause the device to oscillate. Figure 16 shows a plot of I_D versus V_g for $V_D = 2.5 \text{ V}$.

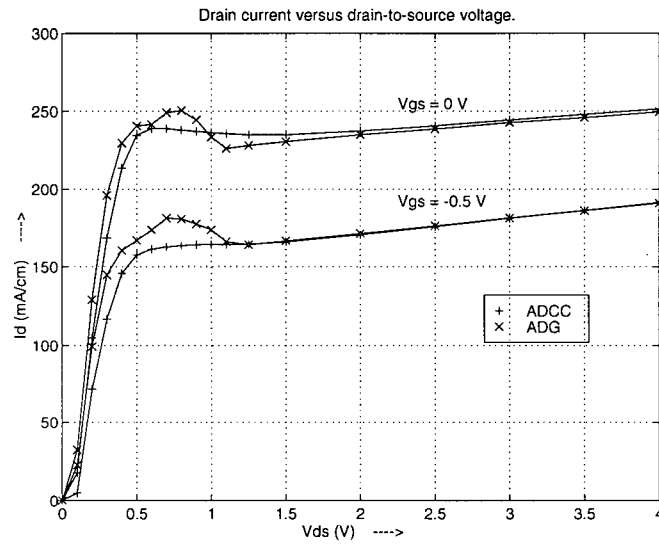


Figure 15. I_D vs. V_D for Example 1.

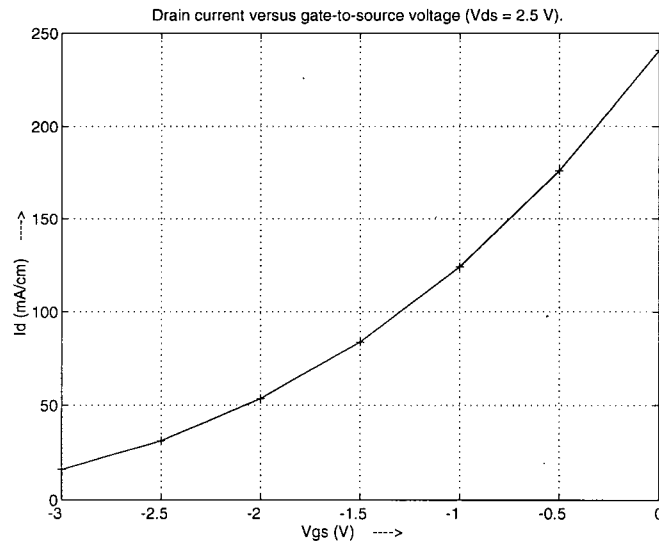


Figure 16. I_D vs. V_g ($V_D = 2.5$ V) for Example 1.

From these two plots, a quick calculation of the voltage gain G of this device operating in saturation mode, or $V_D > 2$ V, and at $V_g = 0$ can be made:

$$G = (\partial I_D / \partial V_g) / (\partial I_D / \partial V_D) \approx 116.3 / 7.2 \approx 16.$$

Figure 15 also shows drain current versus drain-to-source voltage results for both the ADG and the ADCC methods for $V_D = 0, 0.5$ V. The ADG and ADCC curves closely match each other

in the saturation regime but differ considerably in the triode regime and between the triode and saturation regimes. It is assumed that the crude finite element approximation of the term containing the gradient of the mobility, $\nabla\mu_e$, in the advection-diffusion equations may introduce significant error in representing the drift current term, particularly when the electric field intensity along the channel is in the range of the “hill” in Figure 9, i.e. where μ_e is highly nonlinear. Since the current conservation model does not suffer from such approximations and has fewer terms to approximate than the advection-diffusion formulation, the ADCC plot is assumed to be a better approximation of the true current. For this problem, the ADG method required between thirty and forty iterations, while the ADCC method required only about eight iterations for change tolerances $< 10^{-4}$. Of the ADCC iterations, the first two or three were ADG iterations and the remaining were CCNR iterations.

It should be emphasized that these results as well as those for the strip test in Section 4.2 and shown in Figure 7(b) come from two completely different FEM formulations, the current conservation model and the advection-diffusion model, solved by two completely different techniques, the Newton-Raphson method and Gummel’s method, respectively. So, the fact that the ADG and ADCC results agree quite nicely, particularly in the saturation regime, confirms with some certainty that the results are indeed valid as approximate solutions to the MESFET problem.

5.2.2 Example 2

Next, a device similar to the one in the previous example is fully analyzed. A “full analysis” constitutes predicting the small signal characteristics of the MESFET for the purpose of constructing a small signal circuit model useful for implementation in circuit analysis applications. The circuit model used here is adopted from Pulfrey and Tarr [26], and a similar model is presented by Curtice [32].

The device used in this example is the same as the one described in Example 1 with a few modifications: The active layer is $0.4\text{ }\mu\text{m}$ thick with doping density $N_d = 2 \times 10^{16}\text{ cm}^{-3}$, and the buffer zone is $0.6\text{ }\mu\text{m}$ thick. Also, rather than forcing a neutral charge condition, a zero net current condition is forced on the lower artificial boundary. This ensures current conservation in the

system assuming the current definition of (43).

The mesh used is shown in Figure 17. It is made up of 1453 nodes and 688 elements. Potential and carrier density field solutions were computed for a series of bias conditions using ADCC. The number of iterations required to reach the solutions starting from an initial guess of zero potential and neutral charge is about forty, of which only the last three or four are CCNR iterations. For a sufficiently good initial guess, CCNR iterations reached a solution within three or four iterations.

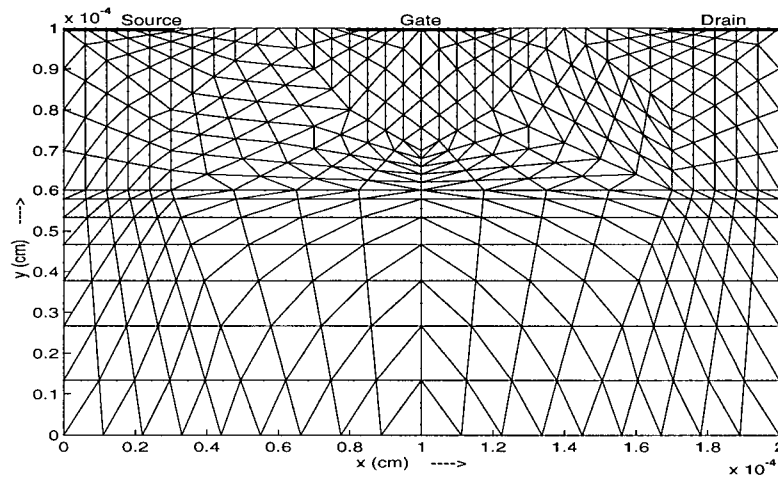


Figure 17. Mesh grid for Example 2 - 1453 nodes, 688 elements.

Figure 18 shows how the potential and carrier fields change as the drain voltage is varied. As the drain voltage is increased from zero voltage (Figure 18(a)-(b)), the depletion region expands on the drain side and begins to pinch-off the channel of carriers. In Figure 18(c)-(d), the channel is completely pinched off and there is relatively little change in the carrier density distribution as the drain voltage is further increased. As a result, the drain current increases very slowly for higher drain voltages, and the device is described as being in the saturation regime. It should be noted that, unlike in Example 1, there is no Gunn domain present in this device. Typically, very small scale MESFET devices do not have this trait, the reason for this being that carrier transit times are so small such that charge accumulations do not have enough time to form before the carriers leave the device. A small carrier wall exists along part of the artificial boundary in all cases. It is a result of satisfying the zero net current boundary condition across this boundary, the potential for which

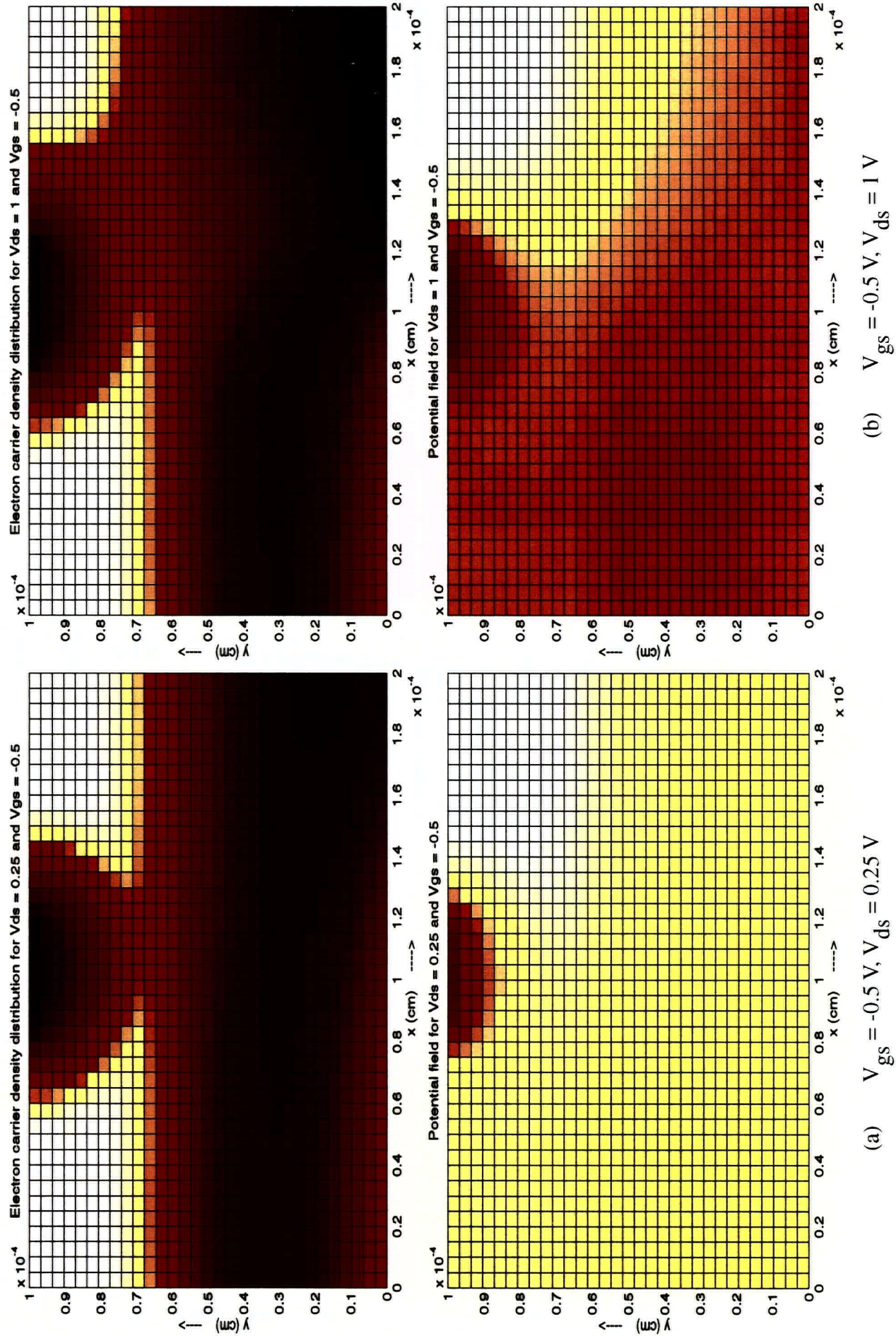
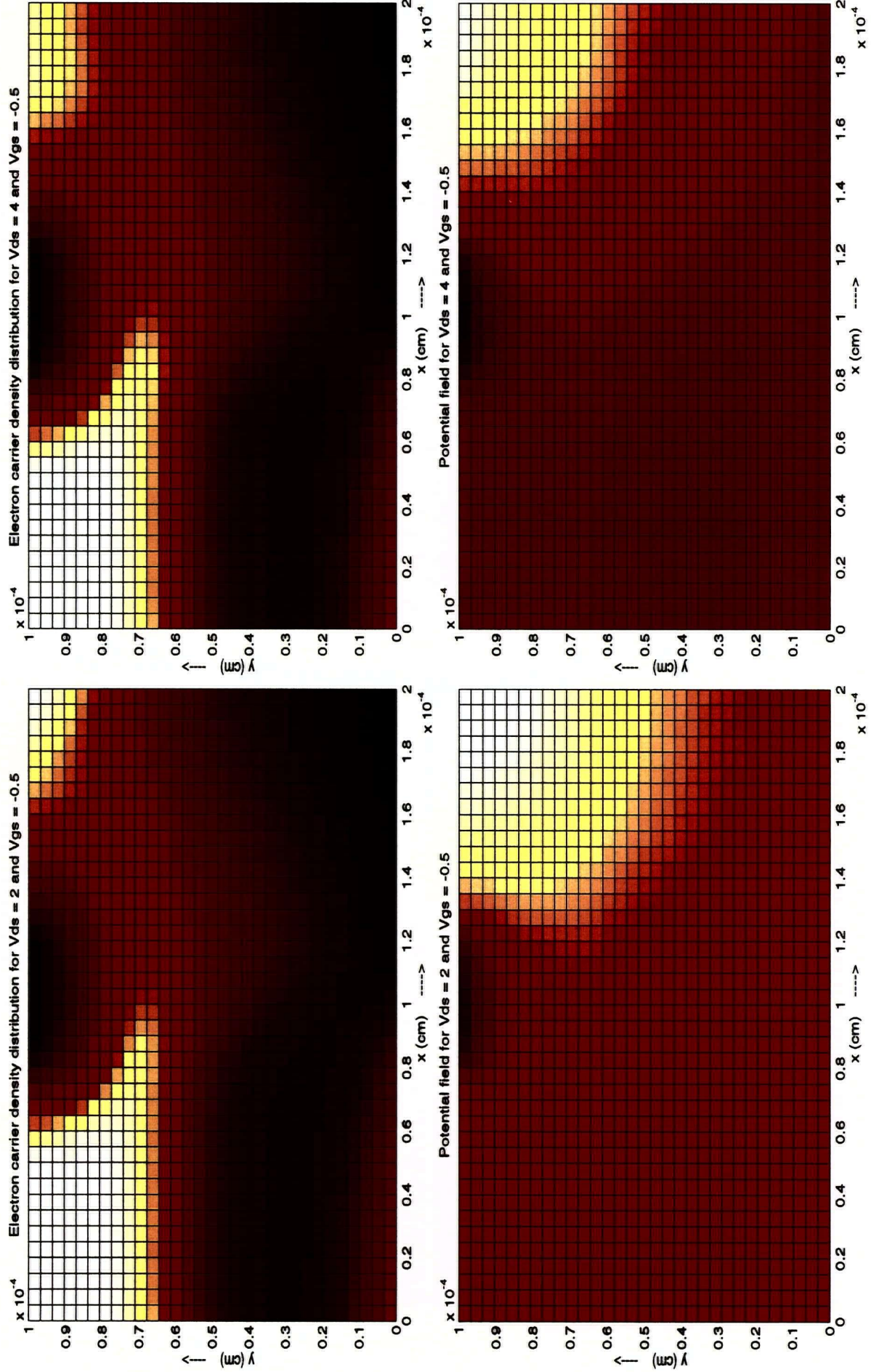


Figure 18. Electron carrier concentration (top) and potential (bottom) distributions.



(c) $V_{gs} = -0.5$ V, $V_{ds} = 2$ V

(d) $V_{gs} = -0.5$ V, $V_{ds} = 4$ V

Figure 18. cont.

$= V_{ds}$ or N_d

$= V_{gs} + V_{bi}$ or 16210 cm^{-3}

has a Dirichlet condition. The carrier wall supplies the necessary diffusion current that cancels out the drift current caused by the potential gradient on the boundary and is therefore solely a by-product of the artificial boundary condition with no physical meaning.

A plot of the current density field is given in Figure 19, and as expected, the majority of the current exists mainly along a narrow channel from the drain to the source and between the Schottky barrier-induced depletion region and the buffer zone.

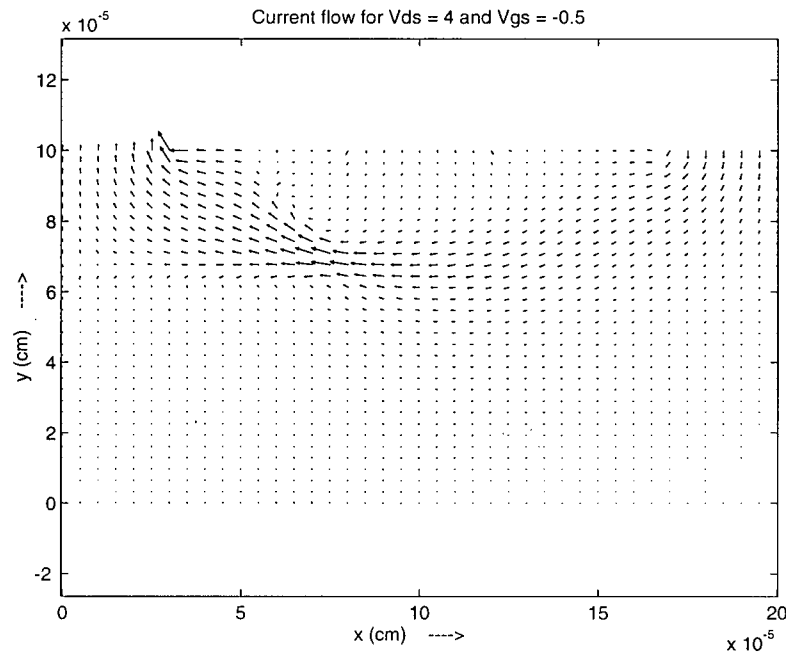


Figure 19. Current density field.

Drain current versus drain voltage plots for different gate voltages are shown in Figure 20. As expected, both a triode and saturation regime are apparent from these plots. Unlike the device in Example 1, there is no region of negative differential resistance in any of the current plots. It is believed that there is a correlation between the existence of a Gunn domain and the appearance of a negative differential resistance in the plots. Under certain bias conditions and depending on its dimensions (particularly the gate length) and doping levels, a MESFET device may behave similarly to a transferred electron device (TED) operating in the limited space-charge accumulation (LSA) mode [33]. TEDs, also known as Gunn diodes, are typically also made of GaAs and are

used for their negative resistance characteristics in amplifier and oscillator applications.

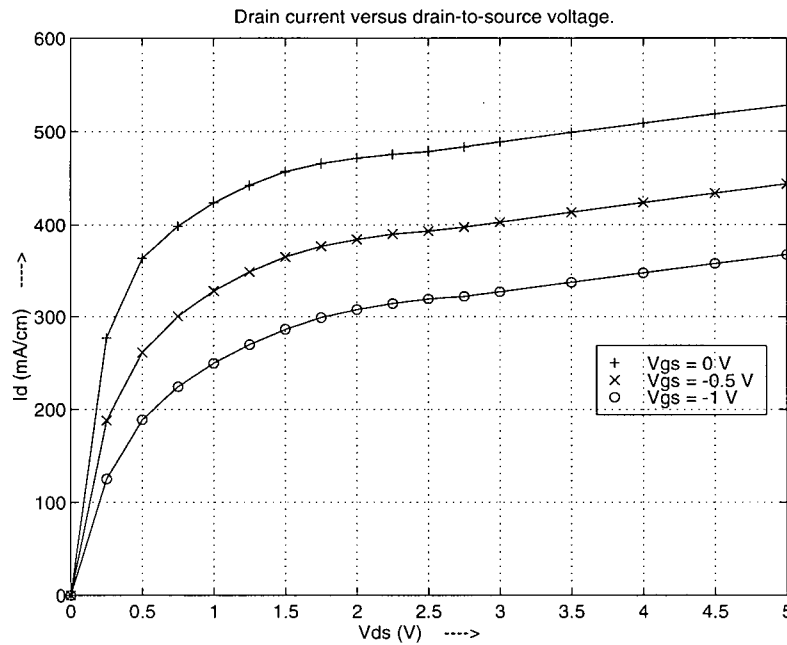


Figure 20. I_d vs. V_d for Example 1.

For a small-signal characterization, the bias conditions of interest are chosen to be $V_{gs} = -0.5$ V and $V_{ds} = 4$ V. All small-signal conductances and capacitances are sought in order to build a small-signal model of the MESFET. From the data used to create the plots in Figure 20, the small signal conductance g_{ds} is roughly approximated as

$$g_{ds} = \partial I_d / \partial V_{ds} \approx \Delta I_d / \Delta V_{ds} = 2.06/1 \text{ mho/m} = 2.06 \text{ mho/m} \quad (46)$$

where $\Delta I_d = I_d(V_{ds}=4.5) - I_d(V_{ds}=3.5)$ and $\Delta V_{ds} = 4.5 - 3.5 = 1$. Next, the drain current must be calculated for different gate-source potentials. Figure 21(a) shows a plot of the results. From this data the transconductance g_m can be computed:

$$g_m = \partial I_d / \partial V_{gs} \approx \Delta I_d / \Delta V_{gs} = 16.17/1 \text{ mho/m} = 16.17 \text{ mho/m} \quad (47)$$

Similarly, the conductance g_{gs} is given as (See Figure 21(b).)

$$g_{gs} = \partial I_g / \partial V_{gs} \approx \Delta I_g / \Delta V_{gs} = 15.97/1 \text{ mho/m} = 15.97 \text{ mho/m} \quad (48)$$

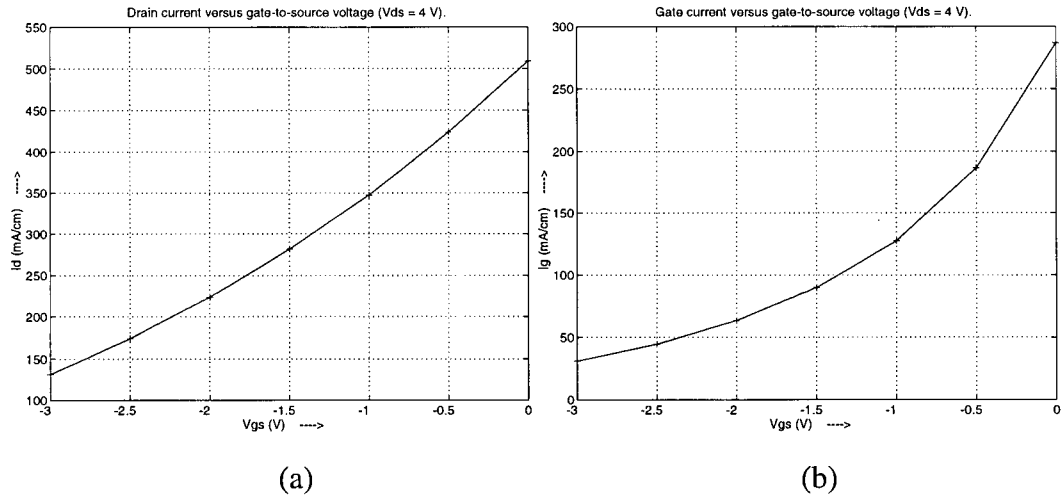


Figure 21. (a) Drain and (b) gate current versus gate voltage.

By linearizing the currents and voltages around the DC bias points, the method outlined in Appendix B can be applied for finding the small signal capacitances of the device. This is done simply by replacing $[V]$ and $[Q]$ with perturbations $\Delta[V]$ and $\Delta[Q]$, respectively, so that the resulting capacitance matrix satisfies

$$\Delta[Q] = [C] \Delta[V],$$

where $\Delta[Q] = [\delta Q_d \ \delta Q_g \ \delta Q_s]^T$ and $\Delta[V] = [\delta V_d \ \delta V_g \ \delta V_s]^T$.

Using (44), electrode charges are computed from the carrier and potential solutions for several electrode potentials in the vicinity of the bias voltages on the drain, gate, and source. Figure 22 shows plots of the charges on each electrode for varying electrode potentials.

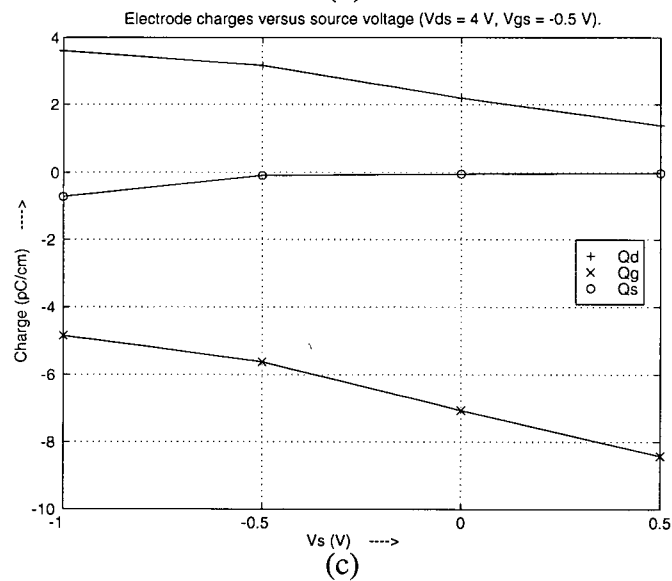
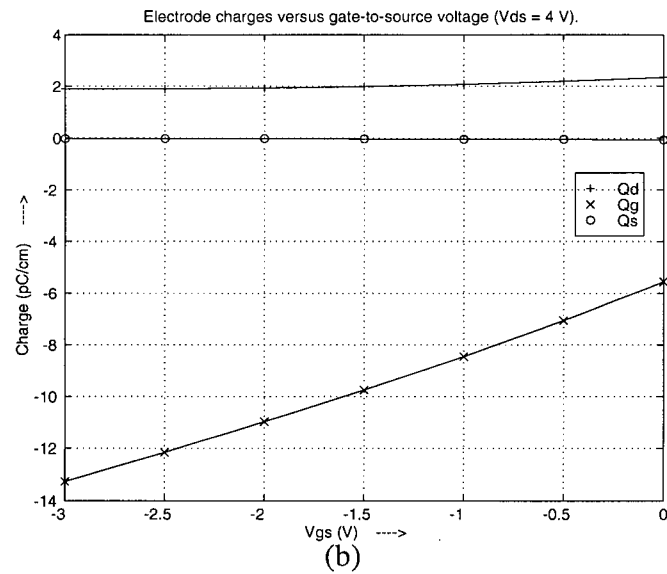
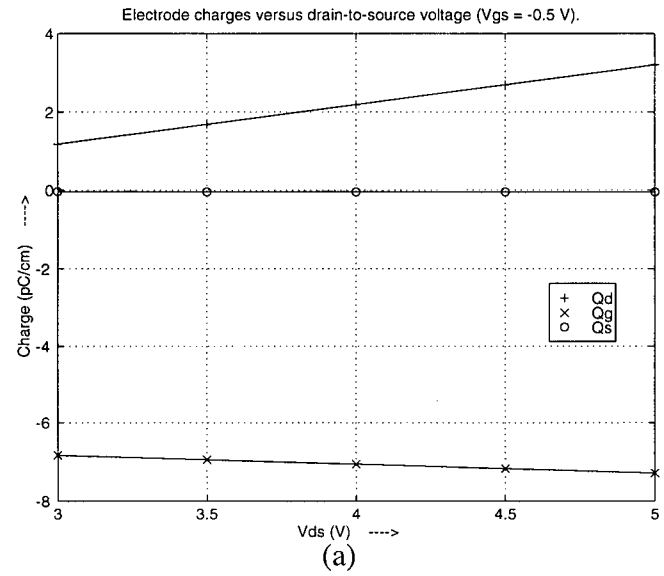


Figure 22. Electrode charges for varying potentials on (a) the drain, (b) the gate, and (c) the source.

From the data in the plots, the capacitance matrix is constructed:

$$[C] = \begin{bmatrix} 100.04 & 26.66 & -179.22 \\ -22.7 & 288.9 & -280.52 \\ -0.26 & -3.28 & 6.6 \end{bmatrix} \text{ pF/m}$$

As discussed in Appendix B, the diagonal terms give the capacitances to ground. Since this is an active device, this capacitance matrix is not symmetric. So, the next step is to decide which off-diagonal terms are to be used to represent the mutual capacitances between the drain, gate, and source terminals. Assuming that the device is in common source mode, the gate acts as the input, and it is desired to represent the behavior of the charges on the gate as they redistribute themselves in response to a change in the applied input voltage. This behavior can be expressed by the capacitance elements C_{21} and C_{23} [34]. Also, since the manner in which the charge redistributes itself at the output (the drain terminal) in response to a change in the drain voltage is sought, the capacitance element C_{13} is used as the mutual capacitance between the drain and the source, which in this case is an additional capacitance to ground. The resulting mutual capacitances and self-capacitances for this MESFET device in common source configuration are given as follows:

$$\begin{aligned} C_{GS} &= 280.5 \text{ pF/m} \\ C_{GD} &= 22.7 \text{ pF/m} \\ C_{DS} &= 179.2 \text{ pF/m} \\ C_{GG} &= 288.9 \text{ pF/m} \\ C_{DD} &= 100.0 \text{ pF/m} \end{aligned} \tag{49}$$

Compiling the results of (46)-(49) yields the small signal circuit model shown in Figure 23. An exponential term is introduced in the current source expression to take into account the time it takes for an electron to travel from the source to the drain. This delay time is taken to be approximately 10 ps, which is the time required for an electron to travel 1 μm at a velocity of $1 \times 10^7 \text{ cm/s}$.

τ = transit time (~ 10 ps)

R_G and R_D = parasitic resistances

$g_m = 13$ mho/m

$g_{gs} = 16.2$ mho/m

$g_{ds} = 2.1$ mho/m

$C_{GG} = 288.9$ pF/m

$C_{DD} = 100.0$ pF/m

$C_{DS} = 179.2$ pF/m

$C_{GS} = 280.5$ pF/m

$C_{GD} = 22.7$ pF/m

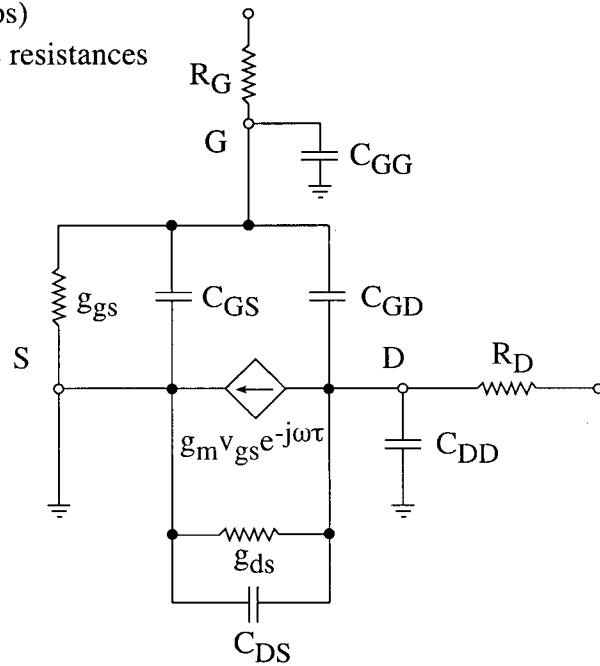


Figure 23. Small signal model of MESFET in Example 2.

It should be noted that the rough approximations of the partial derivatives in the above analysis were used in the context of a simple example with the intention of demonstrating direct application of the sampled data points in the plots. For applications in circuit analysis software, such as SPICE, each set of sampled data could be replaced by a polynomial curve fit. It follows that all conductances and capacitances could then be represented by continuous polynomial expressions as functions of the terminal voltages.

6 Extensions to FEM Model

In monolithic microwave integrated circuits (MMICs), MESFETs are implanted into an intrinsic GaAs substrate, and strip-like conductors feed the source, gate, and drain of the devices. Typically, source and drain strips are on the order of tens of microns in width, extending far past the feeds of the MESFET's active region (See Figure 24). These strip conductors introduce considerable capacitances of their own that are not completely accounted for in the previously described FEM analysis. These capacitances are parasitic degrading the performance of the device in terms of speed and power consumption [32], [35]-[37]. Following are descriptions and recommendations of how Green's integral equations and boundary integral equations (BIE) solved with the method of moments (MoM) can be used in conjunction with the FEM model of the MESFET to give a more accurate model of the device. The substrate and ground plane shown in Figure 24 extend infinitely along the x-axis.

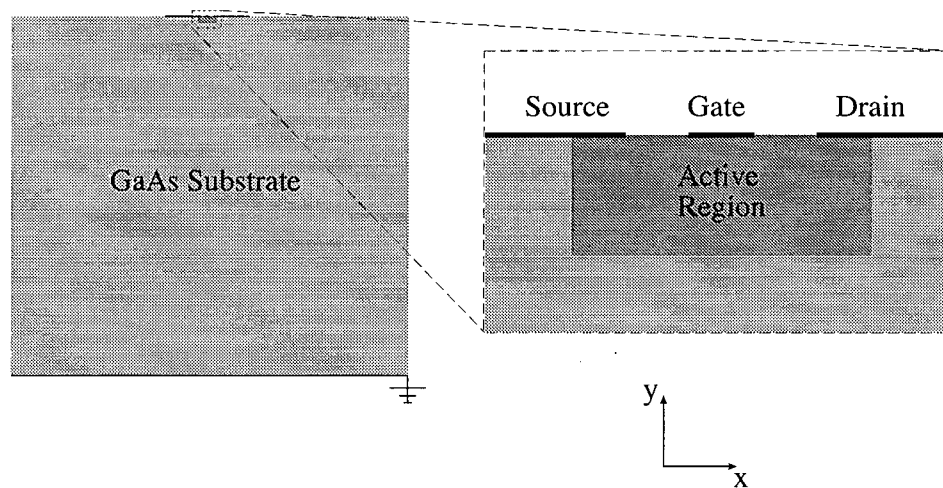


Figure 24. MESFET with strip conductor feeds.

6.1 Method of Moments and Integral Equations

The method of moments (MoM) is a very efficient numerical technique originally

introduced by Harrington [5] for finding approximate solutions to integral equations of the form:

$$a(r) = \int_{\Omega} b(r') c(r, r') dr'$$

where r is a position in space, $a(r)$ is known in some domain Ω , $b(r')$ is unknown and to be determined, and $c(r, r')$ is a known function, typically a Green's function for a particular linear differential equation satisfying a set of boundary conditions. The MoM consists of expanding the unknown $b(r')$ into a sum of N known basis functions with N unknown coefficients, setting up a set of N equations for N known values of $a(r)$, and solving for the unknown coefficients by a matrix inversion.

6.2 Parasitic Capacitances and Green's Integral Equation

The strip electrodes of a MESFET can be thought of as having free charges sitting on their surfaces when bias potentials are applied to them. As in [6], the goal is to find these charges in order to determine the parasitic capacitances of the strip electrodes which are then simply added together with the MESFET capacitances resulting from the FEM simulation in the circuit model. The big approximation and assumption that should be noted here are that the GaAs substrate is essentially charge-free (a perfect insulator) and that the strips and that the charge-free air-substrate-ground configuration can be decoupled from the active region of the MESFET. This reduces the problem to a linear Poisson equation which governs the behavior of the static field resulting from a set of free charges restricted to some finite domain (the strip surfaces).

A general solution of the potential field for the linear Poisson equation is the Green's integral equation .

[5]:

$$\phi(r) = \int_{\Omega} \rho(r') G(r, r') dr' \quad (50)$$

where $\phi(r)$ is the potential field, $\rho(r')$ is the charge distribution on some domain Ω , and $G(r, r')$ is the Green's function of the system. In this case, the strip surfaces constitute the domain of the problem, and the Green's function [38] satisfies the boundary conditions that exist at the substrate interfaces and the condition of zero potential at infinity. Since the potentials on the electrodes and the Green's function are known, the charge distribution can be found using the MoM.

As in , New York: Macmillan, 1968.

[6], the strip conductors are approximated to be infinitesimally thin. In contrast to , New York: Macmillan, 1968.

[6], the problem is solved in two dimensions rather than three, and complex images and a multipipe model of the strips are used to give a much more compact and efficient MoM solution to the problem rather than using conventional image theory and pulse basis functions. Brief overviews of these techniques follow.

6.2.1 Complex Images

Complex images [39] are used for the purpose of determining an approximate closed-form Green's function for a multilayered substrate. For the case of a point charge above a ground plane, conventional image theory states that the potential field above the ground plane is identical to the field for the case in which the ground plane is replaced by a mirror image of the original charge with a sign opposite that of the original charge. This conveniently allows the Green's function of the original system to be represented as the sum of two free-space Green's functions which have simple, known analytic expressions.

When dealing with a point charge embedded in a multilayered (more than two layers) structure, conventional image theory (, New York: Macmillan, 1968.

[6]) leads to representation of the Green's function by a singly (three layers) or multiply (more than three layers) infinite series, which often converges very slowly. The use of complex images allows for a much more compact (typically ~ 5 terms) approximate Green's function. A complex image has both a complex position and a complex amplitude. Since static fields are real, the complex images of a point source always occur in complex conjugate pairs. Complex images are computed numerically by approximating the Green's function in the spectral domain with complex exponentials using Prony's method. The spectral domain Green's function is computed using the transmission line method outlined in [12].

6.2.2 Multipipe Model

The multipipe model [7] is used for the determination of an appropriate expansion of $\rho(r')$. Each strip is represented by a series of pipes each with constant unknown charge. Gauss-Chebyshev quadrature is used to approximate the Green's integral equation, (50), by a finite summation which essentially results in the replacement of the unknown continuous charge distribution by a finite series of point charges with unknown values. Then, applying the results of Wheeler's work in which he derives a single pipe model of a strip in free space assigns a radius to each of the individual point charge locations, and hence pipes are formed. The result is a series of pipes with different radii and unknown charges that takes the place of the original strip. The charges on each pipe are related to the charge distribution on the strip by a simple relationship. With forced potentials on the strips and the appropriate Green's function, the charges are found using the MoM. The moment matrix resulting from such an approach is very compact for high computational efficiency.

6.2.3 Results

Extending the problem of Example 2, the source, gate, and drain of the MESFET are fed by strips numbered 1, 2, and 3, respectively. Strip 1 and Strip 3 are each $10\text{ }\mu\text{m}$ wide, and Strip 2 is $0.4\text{ }\mu\text{m}$ wide. In applying the MoM, thirteen pipes are used to model each of Strip 1 and Strip

3, while six are used for Strip 2 yielding a 32x32 moment matrix. Five complex image terms are sufficient to accurately approximate the Green's function of the air-substrate-ground plane system. Using the procedure outlined in Appendix B, the capacitances of the strips are calculated to be

$$C_{11} = C_{33} = 230.0 \text{ pF/m}$$

$$C_{22} = 148.7 \text{ pF/m}$$

$$C_{12} = C_{23} = 72.9 \text{ pF/m}$$

$$C_{13} = 98.2 \text{ pF/m}$$

Superposition of these parasitic capacitances with those computed in Example 2 gives the modified circuit model shown in Figure 25. The extra capacitance has the overall effect of slowing

τ = transit time (~ 10 ps)

R_G and R_D = parasitic resistances

$g_m = 13 \text{ mho/m}$

$g_{gs} = 16.2 \text{ mho/m}$

$g_{ds} = 2.1 \text{ mho/m}$

$C'_{GG} = 437.6 \text{ pF/m}$

$C'_{DD} = 330.0 \text{ pF/m}$

$C'_{DS} = 277.4 \text{ pF/m}$

$C'_{GS} = 353.4 \text{ pF/m}$

$C'_{GD} = 95.6 \text{ pF/m}$

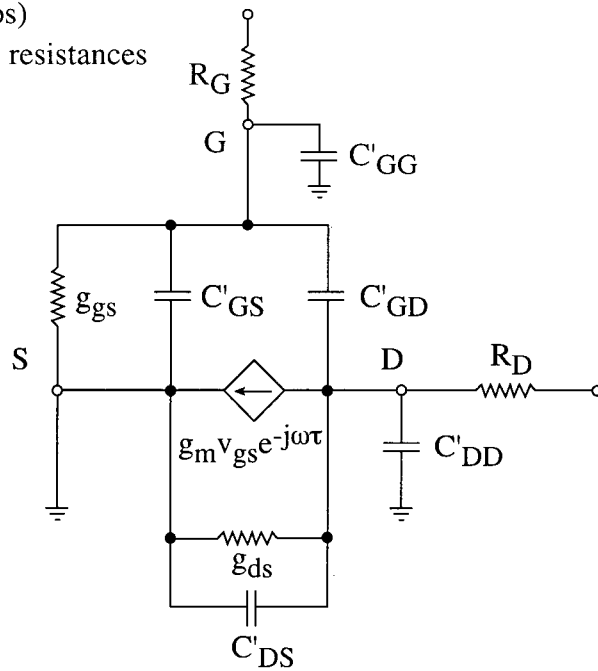


Figure 25. Modified small signal model of MESFET in Example 2.

the device down. Figure 26 compares magnitude versus frequency plots of the transfer functions for both models. The analytic expression for the transfer function of the circuit model is

$$\frac{V_d}{V_g} = \frac{g_m e^{-j\omega\tau}}{g_{ds} + j\omega(C_{DD} + C_{DS}) - \frac{\omega^2 C_{GD}(C_{GG} + C_{GS}) - j\omega C_{GD}g_{gs}}{g_{gs} + j\omega(C_{GD} + C_{GG} + C_{GS})}}$$

As can be seen in Figure 26, inclusion of the parasitic capacitances in the modified circuit model causes the 3 dB cutoff frequency to decrease by more than one half, from 7 GHz to 3 GHz. This illustrates the level of importance of parasitic capacitances in the performance of a MESFET.

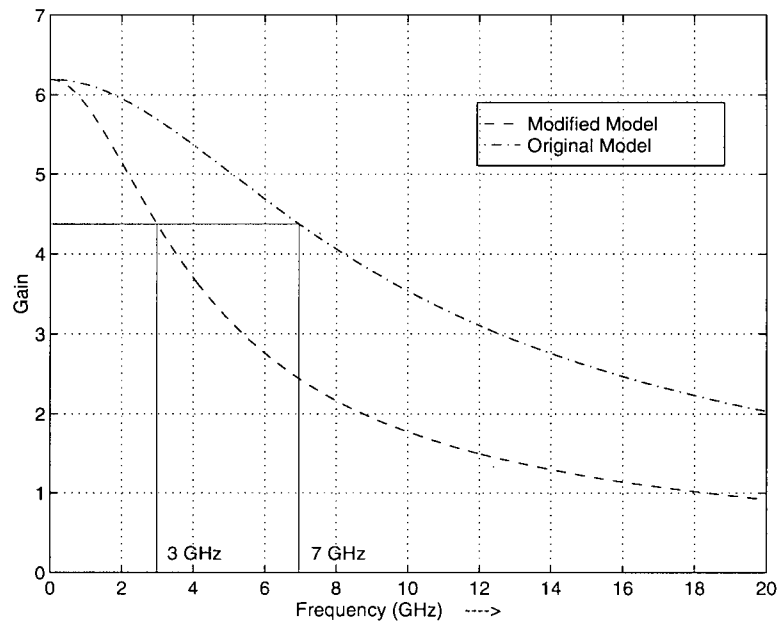


Figure 26. Gain versus frequency plots for circuit models.

6.3 Artificial Boundary Condition Adjustments

In both Example 1 and Example 2, the potential is set to zero on the artificial boundary in order to emulate the effects of the ground plane of the substrate which is actually much farther away. The thickness of the substrate is typically on the order of 100 μm making it impractical to include it in the solution domain of the FEM problem due to the many extra nodes that would be required in such a situation. In addition to this, the zero diffusion and drift current conditions on the left and right boundaries are not entirely realistic since some current does, in reality, travel outside of the active region of the device and into and out of the portions of the electrode strips that

extend beyond the active region as depicted in Figure 24. Following, possible solutions to these deficiencies using integral equation techniques are proposed.

6.3.1 Green's Integral Equation

A simple approach is to modify the FEM solution domain by extending the buffer zone around the entire active region. Zero net current boundary conditions are imposed on the boundary around the buffer zone. Using the Green's integral equation and MoM as explained earlier, the charges on the strip conductors are calculated. From these charges the field along the artificial boundary is computed and is used as a potential boundary condition in the FEM model.

Applying this procedure to the MESFET in Example 2, the buffer zone is made to surround the active region and is $1\text{ }\mu\text{m}$ thick. The mesh is shown in Figure 27. The potential and carrier density fields computed with FEM and the new artificial boundary conditions for $V_{ds} = 4\text{ V}$ and $V_{gs} = 0\text{ V}$ are shown in Figure 29. Drain current versus drain-source voltage for $V_{gs} = 0$ is plotted in Figure 28. There is considerable difference between this current plot and the one in Figure 20 even though the device being modelled is the same.

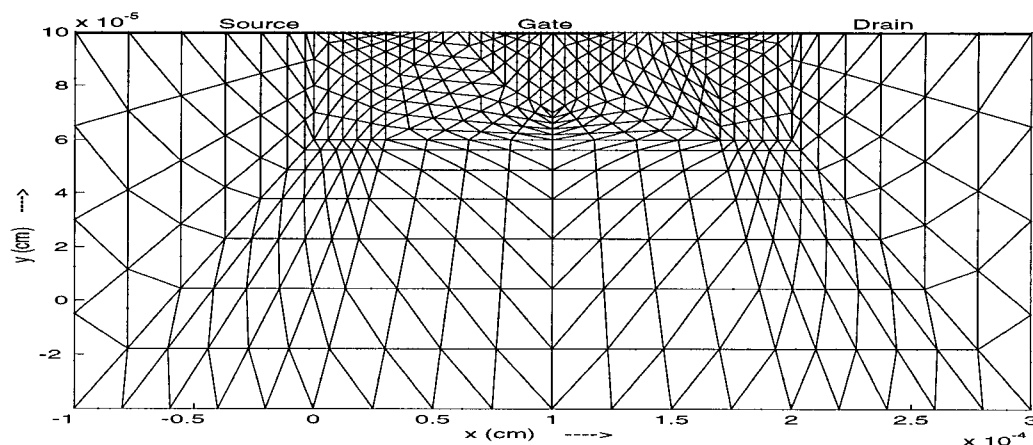


Figure 27. Modified mesh grid for Example 2 - 1677 nodes, 800 elements.

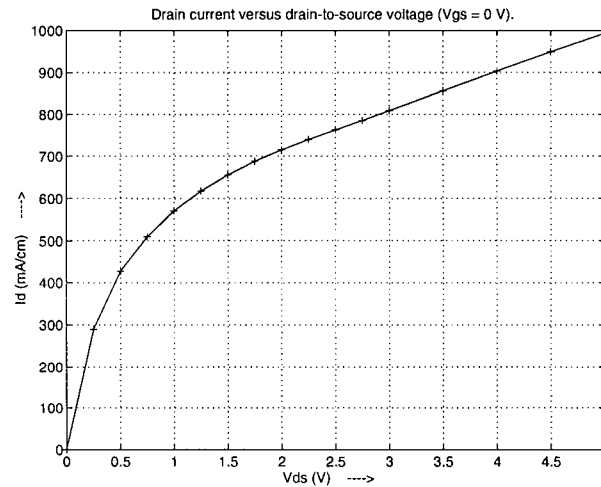


Figure 28. Drain current versus drain-source voltage.

These results are questionable and still under investigation due to the following reasons. Recall (See Section 2.) that ohmic contact conditions were applied at the boundaries between the source and drain conductors, and the active region of the MESFET in place of the heavily doped wells that exist in reality. This simplification greatly reduced the number of necessary nodes in the model (See Section 5.1). Unfortunately in this case, this simplification introduces a large step discontinuity in the carrier concentration condition along the source and drain portions of the boundary. Specifically, the portion of the boundary between the active region and the source and drain conductors has a carrier density condition of $2 \times 10^{16} \text{ cm}^{-3}$, while the portion of the boundary between the buffer region and the source and drain conductors has a carrier density condition of $1.621 \times 10^4 \text{ cm}^{-3}$ due to the Schottky barrier. This discontinuity results in unrealistic, nonphysical current behavior in the vicinity of the step. Of course, if the ohmic contact approximations would be replaced by the actual heavily doped contact wells, this abrupt step discontinuity does not occur, but the number of nodes required as a result of this would be costly. This scenario has not been investigated here. Another potential problem is that the artificial carrier walls that appear along most of the buffer zone boundary are quite high (See Figure 29.), and it is questionable that they have no effect on the true carrier and potential distributions in and around the active region.

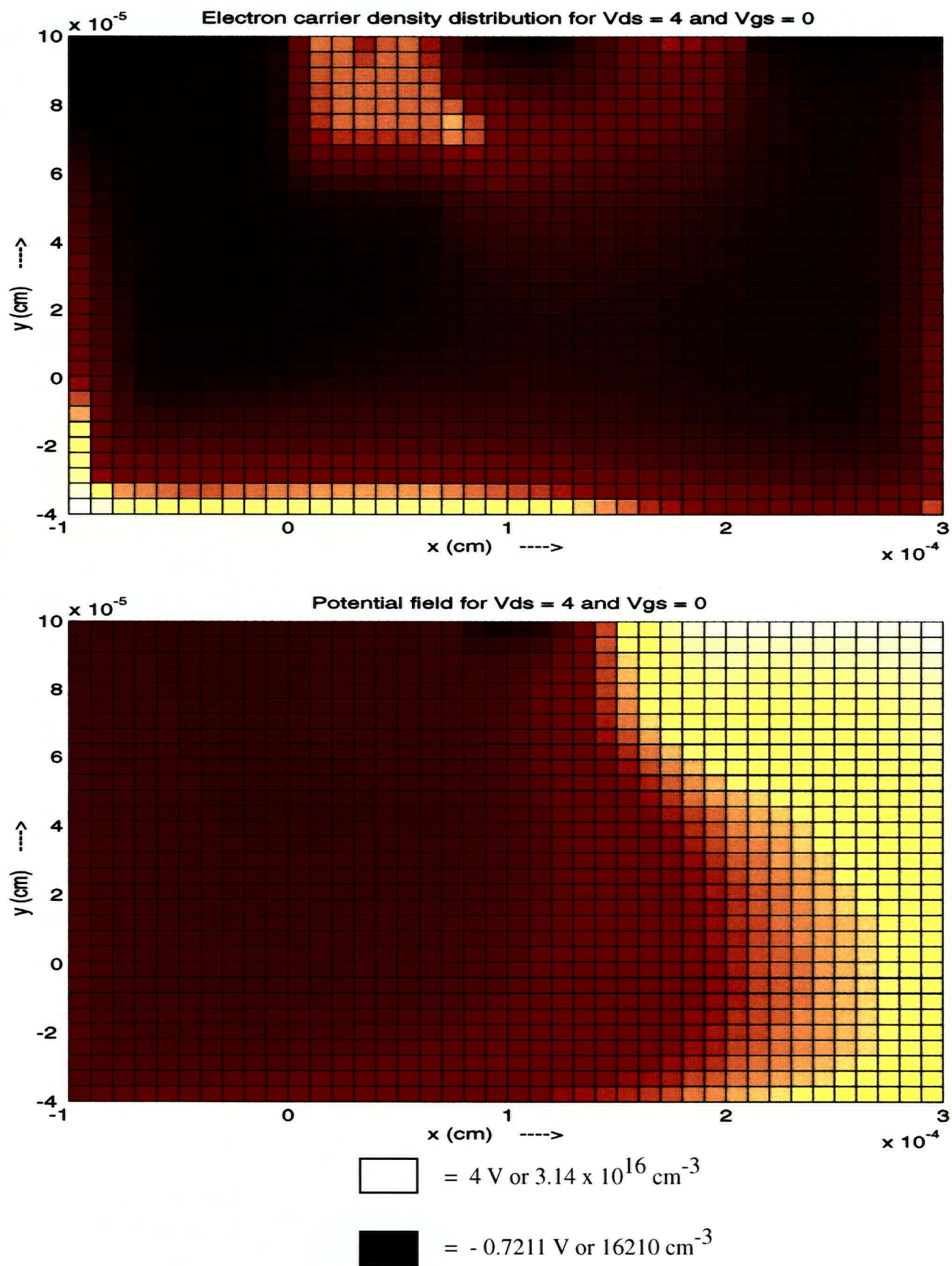


Figure 29. Artificial boundary adjustment results.

6.3.2 Boundary Integral Equation

Instead of using the Green's integral equation to find the potential along the artificial boundary, the above approach could be further improved by incorporating an MoM matrix of a boundary integral equation (BIE) directly into the original FEM formulation thereby coupling the interior of the FEM domain to the exterior.

The BIE expresses the relationship between the potential at a point on the boundary and the potential on the entire boundary. As outlined in [9], the BIE that satisfies the Laplace equation is

$$-\phi(r)/\epsilon = \int_{\Gamma} \phi(r') (\nabla G(r,r') \cdot \mathbf{n}) dr' - \int_{\Gamma} G(r,r') \alpha(r') dr' \quad (51)$$

where $G(r,r')$ is the Green's function satisfying the exterior boundary conditions, $\phi(r)$ is the potential along the boundary Γ , and $\alpha(r') = (\nabla \phi(r') \cdot \mathbf{n})$, the potential flux normal to Γ . Using the MoM to discretize this equation, a matrix expression of α at a finite number of points in terms of the potential at those same points is derived:

$$\begin{aligned} [A] [\phi] &= [B] [\alpha] \\ [\alpha] &= [B]^{-1} [A] [\phi] \end{aligned} \quad (52)$$

where $[\phi]$ and $[\alpha]$ are nodal values of ϕ and α , respectively, along Γ . An approximate expression for α is obtained by adding some shape functions $[M]$ to (52), so $\alpha \approx [M] [\alpha]$. Next, although [9] uses a variational formulation to complete the derivation, it is shown here how the familiar Galerkin's method can also be used to arrive at a similar final result. Substituting the approximation for α and (52) into the surface term in (10) (global version) results in a new expression for this term:

$$\int_{\Gamma} [N]^T \epsilon \nabla [N] [\phi] \cdot \mathbf{n} dS = - \int_{\Gamma} [N]^T \epsilon \alpha dS = - \left(\int_{\Gamma} [N]^T \epsilon [M] dS [B]^{-1} [A] \right) [\phi] \quad (53)$$

The minus sign is introduced since the normal vector of the interior FEM domain points in the opposite direction to the normal vector of the exterior BIE domain. The term enclosed in the brackets is a matrix, and it can be thought of as the stiffness matrix for a "boundary element" comprised of and coupling all the nodes on the outer boundary of the FEM problem. This stiffness matrix differs from the one resulting from the variational approach in that it is not symmetric. The boundary element is treated as any other finite element, and the components of its stiffness matrix are added to the global stiffness matrix in the same manner. This hybrid method has proven effective in open electromagnetic scattering problems [40], and its application to the MESFET problem is currently being studied.

7 Conclusion

This paper presented a two-dimensional finite element analysis of the stationary semiconductor device equations. Exploiting the best features of two different FEM formulations, the current conservation model and the advection-diffusion model, along with two different numerical solutions, Gummel's method and the Newton Raphson method, a robust and efficient solution approach was developed. This technique was put to the test in two sample problems in which potential and carrier density distributions were computed inside GaAs MESFETs. In order to demonstrate the usefulness of these field distributions for the purpose of circuit analysis, various MESFET parameters such as electrode currents, voltage gain, capacitances, and conductances were calculated and incorporated into a small-signal circuit model of one of these devices. Implementation considerations, such as stability and accuracy restrictions, of the FEM solution were also addressed. Finally, various extensions to the FEM approach involving the application of the method of moments (MoM) were briefly discussed and partially demonstrated. These extensions consist of Green's integral and boundary integral equation solutions that are implemented in such a way as to compensate for the limitations of a finite domain and to improve physical accuracy of the FEM solution. The ultimate goal of such extensions is to couple the finite domain to the rest of the circuit of which it is a sub-region.

7.1 Accomplishments

The ability to model the steady-state fields inside semiconductor materials is an important step in properly accounting for active regions in MMICs. Through the development of a new formulation, it has been shown that the finite element method can be used reliably in the solution of the classical semiconductor equations at least in the case of majority carrier devices, which are most prominent in MMIC designs. Additionally, it has also been suggested that the choice of a finite element approach over a finite difference scheme has its advantages in light of the "big picture" of

MMIC modeling. The similarity in the mathematical structures of the finite element method and the method of moments together with the inherent flexibility of boundary handling in the finite element method offers a direct means for the coupling of the two methods. This is necessary if the mutual field coupling among all the components of an MMIC is to be considered. Failure to account for these component interactions may lead to inaccurate performance predictions in the overall design of an MMIC.

7.2 Future Work

Application of the FEM formulations presented in this paper is not restricted to majority-carrier devices and can easily be applied in the simulation of other semiconductor devices. Also, these FEM formulations, as they stand, are not restricted to just two dimensions and, theoretically, can be applied to three dimensional problems as well, without any modification of the expressions or notation shown here. This would be possible by simply using three-dimensional finite elements in place of the two-dimensional elements implemented here. These two points open the door to a substantial amount of further investigation of and testing for the effectiveness of this finite element analysis in dealing with a myriad of semiconductor problems.

Also, the application of boundary integral equations to account for exterior effects on the finite element domain has yet to be accomplished. Hybrid boundary integral finite element methods have successfully been applied to linear open electromagnetic scattering problems and open static field problems. Use of a boundary integral equation discretized with the method of moments leads to a convenient boundary element which is incorporated into the finite element formulation as any other element. For linear problems, this leads to one matrix equation that can be solved simultaneously; iterative techniques are not required. Herein lies the advantage of this hybrid method, as opposed to other existing iterative hybrid schemes, for dealing with nonlinear problems (nonlinear interior, linear exterior), since any other iterations besides the already necessary iterations due to the nonlinear aspect of the problem, e.g. Gummel or Newton-Raphson iterations, would prove to be quite costly in terms of speed. The difficulties with the boundary

integral approach are rooted in the challenge of accurately evaluating integrals over singularities (self-terms) and choosing appropriate basis functions both of which are necessary for an accurate moment method formulation. There is much room for experimentation in this area.

Bibliography

- [1] D. M. Pozar, *Microwave Engineering*, Addison-Wesley Publishing Company Inc., 1990.
- [2] P. A. Markowich, *The Stationary Semiconductor Device Equations*, New York: Springer-Verlag, 1986.
- [3] J. Svacina, *Analysis of Multilayer Microstrip Lines by a Conformal Mapping Method*, IEEE Transactions on Microwave Theory and Techniques, Vol. 40, No. 4, pp. 769-772, April 1992.
- [4] C. Wei, R. F. Harrington, J. R. Mautz, and T. K. Sarkar, *Multiconductor Transmission Lines in Multilayered Dielectric Media*, IEEE Transactions on Microwave Theory and Techniques, Vol. 32, No. 4, pp. 439-449, April 1984.
- [5] R. F. Harrington, *Field Computation by Moment Methods*, New York: Macmillan, 1968.
- [6] N. G. Alexopolous, J. A. Maupin, and P. T. Greiling, *Determination of the Electrode Capacitance Matrix for GaAs Fet's*, IEEE Transactions on Microwave Theory and Techniques, Vol. MTT-28, pp. 459-466, May 1980.
- [7] G. E. Howard, J. J. Yang, and Y. L. Chow, *A Multipipe Model of General Strip Transmission Lines for Rapid Convergence of Integral Equation Singularities*, IEEE Transactions on Microwave Theory and Techniques, Vol. 40, No. 4, pp. 628-636, April 1992.

- [8] K. S. Oh, D. Kuznetsov, and J. E. Schutt-Aine, *Capacitance Computations in a Multilayered Dielectric Medium Using Closed-Form Spatial Green's Functions*, IEEE Transactions on Microwave Theory and Techniques, Vol. 42, No. 8, pp. 1443-1453, August 1994.
- [9] O. C. Zienkiewicz, *The Finite Element Method*, London: McGraw-Hill, 1977.
- [10] G. R. Buchanan, *Schaum's Outline of Theory and Problems of Finite Element Analysis*, New York: McGraw-Hill, 1995.
- [11] D. S. Burnett, *Finite Element Analysis*, Reading, Massachusetts: Addison-Wesley, 1987.
- [12] R. Crampagne, M. Ahmadpanah, and J. Guiraud, *A Simple Method for Determining the Green's Function for a Large Class of MIC Lines Having Multilayered Dielectric Structures*, IEEE Transactions on Microwave Theory and Techniques, Vol. MTT-26, No. 2, pp. 82-87, February 1978.
- [13] Z. Pantic and R. Mittra, *Quasi-TEM Analysis of Microwave Transmission Lines by the Finite-Element Method*, IEEE Transactions on Microwave Theory and Techniques, Vol. MTT-34, No. 11, pp. 1096-1103, November 1986.
- [14] A. Khebir, A. B. Kouki, and R. Mittra, *Higher Order Asymptotic Boundary Condition for the Finite Element Modeling of Two-Dimensional Transmission Line Structures*, IEEE Transactions on Microwave Theory and Techniques, Vol. 38, No. 10, pp. 1433-1437, October 1990.

- [15] H. K. Gummel, *A Self-consistent Iterative Scheme for One-dimensional Steady State Transistor Calculations*, IEEE Transactions on Electron Devices, Vol. Ed-11, pp. 455-465, 1964.
- [16] A. de Mari, *An Accurate Numerical One-Dimensional Solution of the PN-Junction Under Arbitrary Transient Conditions*, Solid State Electronics, Vol. 11, pp. 1021-1053, 1968.
- [17] M. Adler, *A Method for Achieving and Choosing Variable Density Grids in Finite Difference Formulations and the Importance of Degeneracy and Band Gap Narrowing in Device Modeling*, Numerical Analysis of Semiconductor Devices, Dublin: Boole Press, 1979.
- [18] A. F. Franz, G. A. Franz, S. Selberherr, C. Ringhofer, and P. Markowich, *Finite Boxes - A Generalization of the Finite Difference Method Suitable for Semiconductor Device Simulation*, IEEE Transactions, Vol. ED-30, pp. 1070-1082, 1983.
- [19] M. S. Mock, *Basic Theory of Stationary Numerical Models*, Process and Device Modeling, North-Holland: Elsevier Science Publishers B. V., pp. 159-195, 1986.
- [20] R. E. Bank., W. M. Coughran, Jr., W. Fichtner, D. J. Rose, and R. K. Smith, *Computational Aspects of Semiconductor Device Simulation*, Process and Device Modeling, North-Holland: Elsevier Science Publishers B. V., pp. 229-265, 1986.
- [21] J. J. Barnes and R. J. Lomax, *Two-Dimensional Finite Element Simulation of Semiconductor Devices*, Electron. Letters, Vol. 10, pp. 341-343, August 8, 1974.

- [22] J. J. Barnes, R. J. Lomax, and G. I. Haddad, *Finite-Element Simulation of GaAs MESFET's with Lateral Doping Profiles and Submicron Gates*, IEEE Transactions on Electron Devices, Vol. Ed-23, pp. 1042-1048, September 1976.
- [23] J. J. Barnes and R. J. Lomax, *Finite-Element Methods in Semiconductor Device Simulation*, IEEE Transactions on Electron Devices, Vol. Ed-24, pp. 1082-1089, August 1977.
- [24] P. E. Cottrell and E. M. Buturla, *Steady State Analysis of Field Effect Transistors via the Finite Element Method*, International Electron Devices Meeting Technical Digest, Washington, DC, pp. 51-54, December 1975.
- [25] G. L. Tan, X. L. Yuan, Q. M. Zhang, W. H. Ku, and A. J. Shey, *Two-Dimensional Semiconductor Device Analysis Based on New Finite-Element Discretization Employing the S-G Scheme*, IEEE Transactions on Computer-Aided Design, Vol. 8. No. 5, pp. 468-478, May 1989.
- [26] D. L. Pulfrey and N. G. Tarr, *Introduction to Microelectronic Devices*, Englewood Cliffs, New Jersey: Prentice Hall, 1989.
- [27] A. N. Brooks and T. J. R. Hughes, *Streamline Upwind/Petrov-Galerkin Formulations for Convection Dominated Flows with Particular Emphasis on the Incompressible Navier-Stokes Equations*, Computer Methods in Applied Mechanics and Engineering, Vol. 32, pp. 199-259, 1982.
- [28] C Froberg, *Introduction to Numerical Analysis*, Reading, Massachusetts: Addison-Wesley, 1965.

- [29] P. M. Gresho and R. L. Lee, *Don't Suppress the Wiggles - They're Telling You Something!*, Computers and Fluids, Vol. 9, pp. 223-253, Great Britain: Pergamon Press Ltd., 1981.
- [30] S. V. Marshall and G. G. Skitek, *Electromagnetic Concepts and Applications*, Englewood Cliffs, New Jersey: Prentice Hall, 1990.
- [31] O. C. Zienkiewicz and D. V. Phillips, *An Automatic Mesh Generation Scheme for Plane and Curved Surfaces by 'Isoparametric' Co-ordinates*, International Journal for Numerical Methods in Engineering, Vol. 3, pp. 519-528, 1971.
- [32] W. R. Curtice, *A MESFET Model for Use in the Design of GaAs Integrated Circuits*, IEEE Transactions on Microwave Theory and Techniques, Vol. MTT-28, pp. 448-456, May 1980.
- [33] S. Y. Liao, *Microwave Circuits and Devices*, Prentice Hall Inc., 1980.
- [34] T. Takada, K. Yokoyama, M. Ida, and T. Sudo, *A MESFET Variable-Capacitance Model for GaAs Integrated Circuit Simulation*, IEEE Transactions on Microwave Theory and Techniques, Vol. MTT-30, pp. 719-723, 1982.
- [35] R. L. Van Tuyl, C. A. Liechti, and C. A. Stolte, *Gallium Arsenide Digital Integrated Circuits*, Air Force Avionics Lab., AFWAL, WPAFB, Technical Report AFAL-TR-26-264, April 1977.
- [36] J. Maupin, P. Greiling, and N. Alexopoulos, *Speed Power Tradeoff in GaAs FET Integrated Circuits*, paper presented at 1st Specialty Conference Gigabit Logic for Microwave Systems, Orlando, Florida, May 1979.

- [37] R. L. Van Tuyl, C. A. Liechti, R. E. Lee, and E. Gowen, *GaAs MESFET Logic with 4-GHz Clock Rate*, IEEE Journal on Solid State Circuits, Vol. SC-12, pp. 485-496, October 1977.
- [38] W. E. Boyce and R. C. DiPrima, *Elementary Differential Equations and Boundary Value Problems*, John Wiley & Sons Inc., 1986.
- [39] Y. L. Chow, J. J. Yang, and G. E. Howard, *Complex Images for Electrostatic Field Computation in Multilayered Media*, IEEE Transactions on Microwave Theory and Techniques, Vol. 39. No. 7, pp. 1120-1125, July 1991.
- [40] P. P. Silvester and G. Pelosi, *Finite Elements for Wave Electromagnetics Methods and Techniques*, New York: IEEE Press, 1994.
- [41] W. R. Smythe, *Static and Dynamic Electricity*, McGraw-Hill, 1968.

Appendix A Derivation of Shape Functions

The following derivation of shape functions is an extension of Problem 3.2 in [10]. A function f in a two-dimensional space can be approximated by a polynomial of the form

$$f(x,y) \approx C_1 g_1(x,y) + C_2 g_2(x,y) + \dots + C_M g_M(x,y) \quad (A.1)$$

where $g_i(x,y)$ is some arbitrary known function of space, and C_i is some arbitrary constant. Given values of this function at M specific points in space gives a set of M equations and M unknowns:

$$\begin{aligned} f(x_1, y_1) &= C_1 g_1(x_1, y_1) + C_2 g_2(x_1, y_1) + \dots + C_M g_M(x_1, y_1) \\ f(x_2, y_2) &= C_1 g_1(x_2, y_2) + C_2 g_2(x_2, y_2) + \dots + C_M g_M(x_2, y_2) \\ &\vdots \\ f(x_M, y_M) &= C_1 g_1(x_M, y_M) + C_2 g_2(x_M, y_M) + \dots + C_M g_M(x_M, y_M) \end{aligned}$$

or

$$[f] = [X][C]$$

where $f_i = f(x_i, y_i)$ and $X_{ij} = g_j(x_i, y_i)$. So, the solution for the coefficients is given by $[C] = [X]^{-1}[f]$. Now, (A.1) can be rewritten as

$$f(x,y) \approx [g][C]$$

where $[g]$ is a row vector with $g_i = g_i(x,y)$. Substituting for $[C]$ in this expression gives

$$f(x,y) \approx [g][X]^{-1}[f] = [N][f] = N_1(x,y)f(x_1,y_1) + N_1(x,y)f(x_2,y_2) + \dots + N_M(x,y)f(x_M,y_M) \quad (A.2)$$

where $N_i(x,y)$ is the shape function for point i . So, (A.2) expresses the original function in terms of it's values at a finite number of points, M , and the shape functions at these points. Note that this procedure is not limited to two dimensions and is just as easily applicable to higher dimensions.

As a simple example, an expression for the shape functions of a linear triangular element is derived. Here, the three points, or nodes, at which the function to be approximated is known are located at the three vertices of the triangle, (x_1,y_1) , (x_2,y_2) , and (x_3,y_3) . Next, $[g]$ is chosen to be $[1 \ x \ y]^T$. This is the same as saying that $g_1(x,y) = 1$, $g_2(x,y) = x$, and $g_3(x,y) = y$ in (A.1).

The matrix $[X]$ is constructed:

$$[X] = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}$$

It follows that

$$[N] = [1 \ x \ y] \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}^{-1} \quad (A.3)$$

Solving for an analytic expression for $N_i(x,y)$, requires some considerable amount of algebra. Analytic solutions become much more difficult to obtain for higher order elements. Nevertheless, in practice there is no need to have analytic expressions for the shape functions, and (A.3) is ideally suited for numerical applications. Once $[X]^{-1}$ has been computed, it is a simple matter to calculate the value of the shape functions at any point in space simply by plugging the coordinates

into $[g]$ and solving (A.3). This method for computing shape functions can be applied to a space of any number of dimensions. On the other hand, choosing an appropriate set of functions $[g]$ and sampling points is quite arbitrary. For triangular elements of any order, Pascal's triangle can be used to find the proper $[g]$.

Appendix B Capacitance Calculations for a Multi-Port System

A circuit model composed of N nodes interconnected by capacitors representing the capacitive coupling for a system of N conductors is sought. It is assumed that a means is available for calculating charges, Q_n , on the conductors given any arbitrary set of potentials, V_n , forced on the conductors. The charge on conductor 1 is defined as

$$Q_1 = C_{11}V_1 + C_{12}V_2 + \dots + C_{1N}V_N \quad ,$$

Similarly, for the charge on conductor 2

$$Q_2 = C_{21}V_1 + C_{22}V_2 + \dots + C_{2N}V_N \quad ,$$

and so forth for the charges on the remaining $N-2$ conductors. Putting these relationships between the conductor voltages and charges into matrix form gives

$$[Q] = [C][V]$$

where $[Q]$ is an $N \times 1$ column vector with components Q_n , $[V]$ is an $N \times 1$ column vector with components V_n , and $[C]$ is an $N \times N$ matrix with components C_{nm} . $[C]$ is referred to as the capacitance matrix [41]. By definition, the N components of the first column of $[C]$ are simply the values of the charges on the N conductors for the case of a potential of value one on the first conductor and zero on the rest. The N components of the second column of $[C]$ are the values of the charges on the N conductors for the case of a potential of value one on the second conductor and zero on the rest, and so on. The magnitude of each individual component of $[C]$, $|C_{nm}|$, is the value of the circuit capacitor between node n and m for $n \neq m$, mutual capacitance, or the value of the capacitor from n to ground for $n = m$, self-capacitance, in the equivalent circuit model.

For example, two microstrip conductors are represented by the equivalent circuit model as

shown in Figure B1.

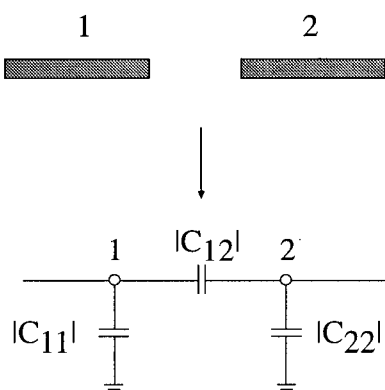


Figure B1. Circuit model for two strip conductors.

It should be noted that for a passive system, $C_{nm} = C_{mn}$, due to reciprocity.