LARGE-SIGNAL RF SIMULATION AND CHARACTERIZATION OF ELECTRONIC DEVICES USING FERMI KINETICS TRANSPORT

By

Nicholas Charles Miller

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Electrical Engineering – Doctor of Philosophy

2017

ABSTRACT

LARGE-SIGNAL RF SIMULATION AND CHARACTERIZATION OF ELECTRONIC DEVICES USING FERMI KINETICS TRANSPORT

By

Nicholas Charles Miller

Design of radio frequency (RF) power amplifiers (PAs) for wireless communications requires small- and large-signal data collected from the underlying transistors, including scattering parameters (S-Parameters) and load-pull (LP), to determine optimal impedance targets. High speed devices operating with fundamental frequencies above 35 GHz present extreme challenges for measuring the harmonic signals resulting from nonlinear effects. Predictive physics based simulations in conjunction with compact modeling capabilities are promising alternatives to expensive and time-consuming measurements. To date, tools either exist in the electron transport domain or in the behavioral modeling domain and a key goal is to treat these problems simultaneously because they are strongly coupled at millimeter wave frequencies. Accurate physics based simulations of high speed and high power transistors require proper treatment of hot-electron, self-heating, and full-wave effects. The Boltzmann solver called Fermi kinetics transport (FKT) has been shown to capture all of these important physical effects. FKT can approach the accuracy of Monte Carlo methods while maintaining the computational efficiency of deterministic solvers. The latter trait allows simulation of large electronic devices such as the output stages of PAs. Previous work on FKT provided proof of concept results which demonstrated its versatility and accuracy as an electronic device simulation framework.

The purpose and contribution of this thesis is the use of FKT as a predictive TCAD tool to generate RF data required for PA design. This work begins with a thorough investigation of the underlying physical equations and their numerical solution for electronic device simulations. Included in this investigation is an analysis of the full-wave discretization technique called Delaunay-Voronoi surface integration (DVSI), a derivation of the FKT device equations and their discretization in energy- and real-space, and a detailed account on the numerical solution of the fully coupled non-

linear system of equations. The detail provided in this work is meant to provide future device engineers and researchers a thorough understanding of the numerical framework for their application and simulation needs. The FKT device simulator is then applied to real device geometries to generate useful data for RF circuit designers. Extensions of the FKT method required for large-signal LP simulations are presented with representative applications. Additionally, compact behavioral models are extracted directly from FKT device simulations, enabling a computationally efficient means for simulated LP data generation. The resulting TCAD tool is a promising simulation capability for high power RF transistor design and characterization. It is anticipated that PA design for 5G applications will use techniques like these in the near future.

To my wife, Samantha.

ACKNOWLEDGEMENTS

First and foremost, I would like to sincerely thank my advisor Professor John Albrecht for all the support and advice, both technical and personal, over my graduate school career. I would especially like to thank him for all his hard work these past few months on making my thesis readable. I wish to continue collaboration with him for many years.

Next, I extend a wealth of gratitude to my committee members Professor Timothy Grotjohn, Professor Donald Morelli, and Professor John Verboncoeur for their advice and input to my graduate career.

I would like to sincerely thank Dr. Matt Grupen at the Air Force Research Laboratory Sensors Directorate (AFRL) for his support and encouragement over the years and his time and effort editing my thesis. This thesis is based on his pioneering work. Matt amazes me with his deep technical discussions on solid state physics. However, they always seem to be laced with Seinfeld references. I was never sure if that was his or my fault.

I would also like to thank Dr. Bob Fitch and Dr. Gregg Jessen at AFRL for their collaboration and for furnishing measured data.

Many thanks to my colleagues at Michigan State University. In particular, I thank Marco Santia, Dr. Nandan Tandon, and Matt Hodek for their years of support and encouragement.

I also thank our colleagues at Tech-X corporation, Dr. David Smithe and Dr. Kris Beckwith, for their technical advice over the years.

Thank you to the SMART scholarship program for supporting my graduate education and research and enabling my collaborations with and future at AFRL.

I would be remiss to not thank my undergraduate research advisor Professor Shanker Balasubramaniam for his time and support in helping me learn computational electromagnetics/physics. Furthermore, I would like to thank Professor Ed Rothwell for connecting me with the CEM group and starting my research career. Finally, I would like to thank Dr. Andrew Baczewski, Dr. Naveen Nair, and Dr. Daniel Dault, each of whom devoted countless hours towards my fundamental un-

derstanding of computational electromagnetics/physics.

I would not be here without the support and encouragement of my parents, Michael and Debra Miller. Thank you for all you have done!

Finally and most importantly, I thank my wife Samantha for all her love, support, and encouragement. It is to her that I dedicate this thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Wireless communications is an ever-growing research area in the information age. Examples of wireless communications include commercial applications such as cellular telephones, wireless local area network (WLAN) and Wi-Fi, global positioning system (GPS), and military applications like radar sensors, electronic warfare (EW) systems, and satellite communications. A critical component common to all wireless systems is the radio frequency (RF) amplifier. As bandwidth and efficiency criteria evolve with more sophisticated technology, so too must the underlying transistor technologies which comprise the integrated RF amplifiers as well as the circuit topologies that exploit these capabilities.

At a system level, RF integrated circuits (RFICs) used for wireless communications are grouped into two modules: the transmitter (Tx) and the receiver (Rx). A basic illustration of both RFIC modules is presented in Figure 1.1. The transmission module is responsible for up-converting

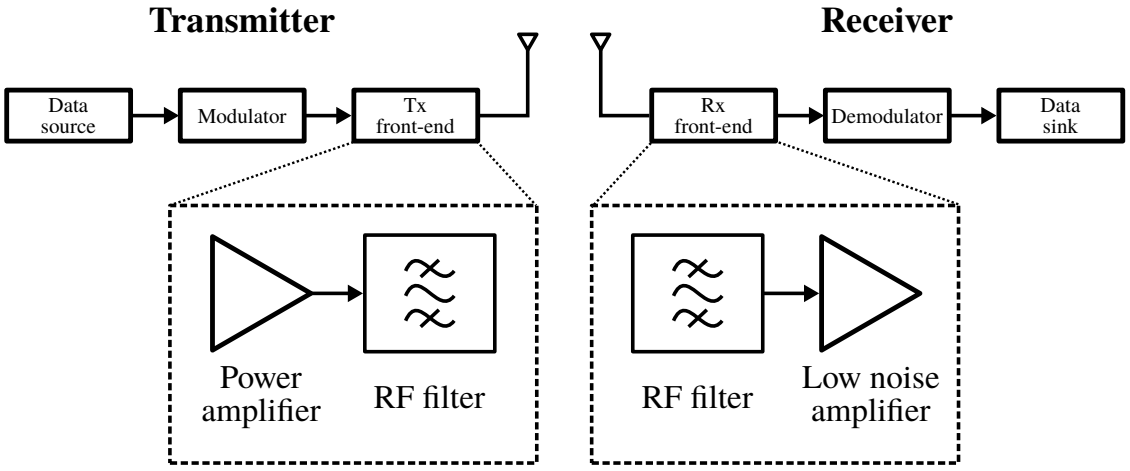


Figure 1.1: A system level diagram of an RFIC transmitter and receiver. The front-end components are further decomposed into the specific types of amplifiers required for the different modules.

(modulating) the signal to a specific frequency band, amplifying the signal with a desired amount of gain, and filtering the signal to the allocated frequency band of the system. Once the signal is received from the channel, the Rx module must filter out unwanted frequency content, amplify the

signal with the least amount of noise possible, then down-convert (demodulate) the signal to the base-band frequency range.

Power amplifiers (PAs) are used to amplify the signal before transmission with the antenna and are an important sub-component of the overall wireless communication process. Ever since the first operational example in the 1940s, transistors have been an important part of practical PAs and remain the only prospect for integrated amplification. The choice of semiconductors for the transistor technology has changed over time due to the changing demand of the PAs and the emergence of higher performance semiconductors that can be readily integrated. The biggest breakthrough which has allowed a wide expansion of semiconductor technology is the monolithic microwave integrated circuit (MMIC) process. The gallium arsenide (GaAs) MMIC amplifier was reported in 1976 and since then there has been tremendous progress in both LNAs and PAs [1]. The GaAs MMIC PA is still an integral part of every cell phone, which lends a calibration to the impact of the technology. Although GaAs was a major part of early RF semiconductor devices, there exist a wide range in transistor types and semiconductors. Included in the transistor types are the GaAs and silicon carbide (SiC) metal-semiconductor field effect transistors (MESFETs), silicon (Si) based metal-oxide-semiconductor FETs (MOSFETs), silicon germanium (SiGe) heterojunction bipolar transistors (HBTs) [2], and GaAs and gallium nitride (GaN) high electron mobility transistors (HEMTs) [3].

GaN HEMTs are a popular PA transistor technology choice due to their high electron saturation velocity, large semiconductor band-gap, high two-dimensional electron gas (2DEG) density in the channel, and high electron mobility [3]. A semiconductor with a large band-gap is superior to that with a lower band-gap as the former will suffer from band-to-band impact ionization at much higher electric fields, enabling higher operating voltages and greater output power in the same form factor. Furthermore, with a high electron saturation velocity, GaN HEMTs can operate at higher frequency compared to their GaAs counterparts which is critical for penetrating high power applications at millimeter-wave frequencies and higher. Finally, in terms of device fabrication, the ability of GaN technology to form heterojunctions allows for polarization induced high carrier

concentration and high electron mobility without modulation doping [4]. These two features allow for a high current density and low channel resistance, which also allows for high frequency operation [5]. These characteristics of GaN HEMTs have allowed implementation of high power and high frequency PAs [6, 7] and are important components for next generation radar and satellite communication RF front-ends.

Independent of the transistor technology and process, a wide variety of PA circuit topologies exist that have been developed over decades to exploit various physical benefits of each given technology. The most simple analog PAs are class A, class B, class AB, and class C amplifiers. The difference between the class A-C amplifiers is the operating point, with the class C amplifier boasting the highest theoretical efficiency at the expense of linearity and vice versa for class A. The next class types of PAs are the switching classes D and E. Class D amplifiers treat the transistor as an ideal switch, whereas the class E amplifier treats the transistor as a non-ideal switch and optimizes the load impedance for maximum efficiency [8]. More advanced types of amplifiers are the class F and inverse class F amplifiers. Class F amplifiers operate with class B biasing and are presented with a harmonic tuning matching network for high efficiency [9]. A transistor operating with a class B bias and with optimized harmonic termination impedance is designated as a class J amplifier. Advanced PA design includes broadband amplifier techniques and linearization techniques. Techniques for increasing bandwidth include reactive/resistive matching amplifiers and feedback amplifiers [1], as well as distributed amplifiers [10]. The techniques for amplifier linearization include feedback techniques [11, 12], feed-forward techniques [13], and efficiency enhancement techniques such as Chireix outphasing [14] and the Doherty amplifier [15].

Design of PAs is significantly more involved than the design of other types of amplifiers such as the low noise amplifier (LNA) and specific applications require drastically different designs. Basic requirements of PAs include high gain, high power added efficiency (PAE), and higher linearity [1]. Once the device technology, the type of PA, and the operating point are selected, the design of the PA often requires measured and simulated load-pull (LP) data to create models that will allow designed circuit performance to match the fabricated reality. LP refers to the evaluation

of figures of merit (FOMs) of a device under test (DUT) when the input and/or output loading conditions are changed [16]. The FOMs, whether measured or simulated, are provided for single transistors or DUTs rather than an entire PA module. Once the response data are available for multiple harmonics of the loading conditions, designers can then create the optimum loading conditions to maximize the desired FOMs. Although the LP concept is relatively straightforward, realizing the measurement system and producing the data can be incredibly time-consuming and sophisticated. The use of mechanical tuners to systematically change the loading conditions presented to the DUT has been in use since the mid-1970s [17]. When LP data for multiple harmonics are required for proper PA design, the use of mechanical tuners is a laborious measurement technique. Today, state-of-the-art measurement systems utilize active LP techniques for fast and efficient multi-harmonic data generation. The first instances of active LP used the open-loop active load, where a signal generator is used to inject incident waves to the DUT to mimic the scattered waves generated by impedance mismatches [18]. Closed-loop technology offers an improved version of the original active LP technique, where the reflected waves are generated directly from the incident waves through directional couplers [16]. The four-port differential time-domain (TD) LP technique is the state-of-the-art and offers powerful measurement capabilities [19].

Measured LP is the standard high frequency characterization used throughout the industry, but it does not come without drawbacks. The main difficulties of active LP are the high power and linearity required by the load amplifiers [16]. Furthermore, the data generated are only valid for a single DUT operating point and therefore bias and frequency sweeps are required for a broad range of device performance data. An alternative approach for PA design is utilization of computer aided design (CAD) tools with nonlinear device models. PA design with CAD tools offers a flexible and efficient approach to performance optimization and can be used to generate synthetic LP data or entire PA module characterization. In order to generate accurate and useful results, the nonlinear device model must be accurate over a wide range of operating points, frequency, loading conditions, and large-signal excitations. These nonlinear models can be grouped into two main categories: circuit based modeling and behavioral modeling.

Circuit based modeling of transistors has existed nearly as long as the devices have themselves. Many small-signal two-port models exist for a wide variety of transistors [1]. These models can be useful for circuit design, e.g., conjugate input matching. However, because PAs are driven by large-signal RF signal levels, these models do not accurately describe the nonlinear nature of the devices. Nonlinear circuit models are therefore critical for large-signal RF design of PAs. An example of a well-known nonlinear device model for GaN HEMT technology is the Angelov model [20, 21]. Keysight Technologies (a spin-off from Agilent Technologies), provides a circuit-based nonlinear GaN HEMT model called the DynaFET model [22, 23], which has garnered considerable attention for nonlinear compact modeling. Both examples of nonlinear GaN HEMT models require extensive measurements of devices to calculate parameters and train the models.

Opposite to circuit based modeling are behavioral nonlinear device models. Rather than develop equivalent circuits to represent the nonlinear transistor, behavioral modeling relies on approximations of the nonlinear analytic map between the input waves (voltages or currents) and the scattered waves. An example of a nonlinear behavioral model applicable to GaN HEMTs is the Polyharmonic Distortion Model (PHD) [24], which later was trademarked as the X-Parameter (XP) model [25, 26] and was used for high-efficiency PA design [27].

Regardless of the choice of the nonlinear device model (circuit based or behavioral), the vast majority of implementations utilize measurements to compute model parameters and/or train the model. An alternative to measurement based model extraction is physics based or technology computer aided design (TCAD) based model extraction. With TCAD based model extraction capabilities, the cost of sophisticated measurement systems and device fabrication can be overcome to allow efficient PA characterization and design. It is critical that the numerical TCAD models accurately describe the underlying physics of the semiconductor devices to produce quality and useful results. To the author's knowledge, there only exists one TCAD framework which can accurately produce nonlinear device models. This framework couples an accurate Monte Carlo (MC) method with harmonic balance (HB) [28] and has been used to extract XPs [29]. Although this TCAD framework boasts the accuracy of MC, it suffers from an immense computational burden

and cannot include full-wave electromagnetic (EM) effects in its present form, although it could be added. A greater limitation is that MC physical domains are limited to sizes of a few microns and PA output stages are often 10–100 times larger in reality.

The main contribution of this thesis is the utilization of a TCAD framework called Fermi Kinetics Transport (FKT), which captures hot-electron and full-wave effects, to simulate, characterize, and extract XP models of state-of-the-art GaN HEMT technology. FKT, which is a deterministic Boltzmann solver, offers a promising numerical framework for accurate, robust, and efficient electronic device simulation. The rest of this thesis is organized as follows. Background material required for this work is presented in Chapter 2. In Chapter 3, a complete discussion of the full-wave EM discretization method called Delaunay-Voronoi Surface Integration (DVSI), which is part of the FKT TCAD numerical framework, is presented. A thorough derivation of the governing equations of the FKT device simulator and their discretization in energy- and real-space is presented in Chapter 4. Numerical details required for solving the nonlinear device equations are presented in Chapter 5, including device simulation work-flow and boundary conditions. This chapter also includes an investigation into some numerical characteristics of the FKT equations and reports some advanced simulation techniques. Chapter 6 presents small- and large-signal RF simulations of GaN HEMT technology from DC to mm-wavelengths. In Chapter 7, a state-of-the-art GaN HEMT is characterized with the FKT device simulator. This chapter presents several small- and large-signal simulations of the HEMT technology. XP extraction of the GaN HEMT is also reported in this chapter. The compact model is imported into Advanced Design System (ADS) to generate data for RF circuit designers. A conclusion and discussion of future research enabled by this thesis are provided in Chapter 8.

CHAPTER 2

BACKGROUND

2.1 The Box Integration Method

Discretization of continuous systems of equations can be done in a multitude of ways. The standard approach for discretizing continuous equations which govern charge transport in semiconductor devices is the box integration method [30]. The idea of the box integration method is based upon Gauss' divergence theorem and can be dated as far back as the 1960s [31, 32, 33]. Given an arbitrary vector field $\vec{\mathbf{F}}$, its divergence integrated over a volume is equal to the flux of the vector field through the surface of the volume

$$\iiint_V \nabla \cdot \vec{\mathbf{F}} dV = \iint_S \hat{n} \cdot \vec{\mathbf{F}} dS. \tag{2.1}$$

Figure 2.1 illustrates the cross-section of a cylindrical volume. The 2D version of the surface integral for the cylindrical region is

$$\iint_S \hat{n} \cdot \vec{\mathbf{F}} dS = \int_{\phi=0}^{2\pi} d\phi \rho \hat{\rho} \cdot \vec{\mathbf{F}}. \tag{2.2}$$

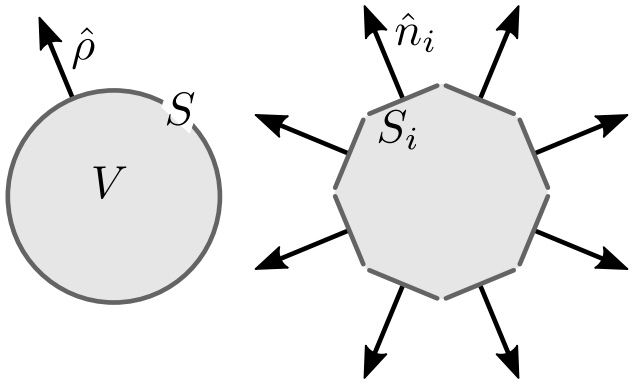


Figure 2.1: (left) A cross-section of a cylindrical volume. (right) A polygonal approximation of the cylindrical volume with eight finite surfaces.

For the discrete version of the cylindrical region shown in the right of Figure 2.1, the surface integral becomes

$$\iint_S \hat{n} \cdot \vec{\mathbf{F}} dS \approx \sum_i \int_{S_i} dS_i \hat{n}_i \cdot \vec{\mathbf{F}} \approx \sum_i F_i A_i, \tag{2.3}$$

if the vector projections of the field $\vec{\mathbf{F}}$ are approximated as constant over each surface of the polygon. This is the crux of the box integration method. Given a polyhedron which encloses a single point in space, the divergence integrated over the polyhedron is represented by a summation of fluxes across the polyhedron's faces. A special type of mesh is required to produce these types of polyhedra and is the topic of the following section.

2.2 The Delaunay-Voronoi Mesh

The basic principles of the box integration method were outlined in the previous section. To properly discretize the divergence of a vector field with the box integration method, a special type of mesh called a Delaunay mesh is required. A 2D Delaunay mesh is a set of polygons where the circumcircle of any polygon contains no points of the mesh. The points of each polygon lie on the perimeter of its respective circumcircle. The 3D analogue is a set of polyhedra with

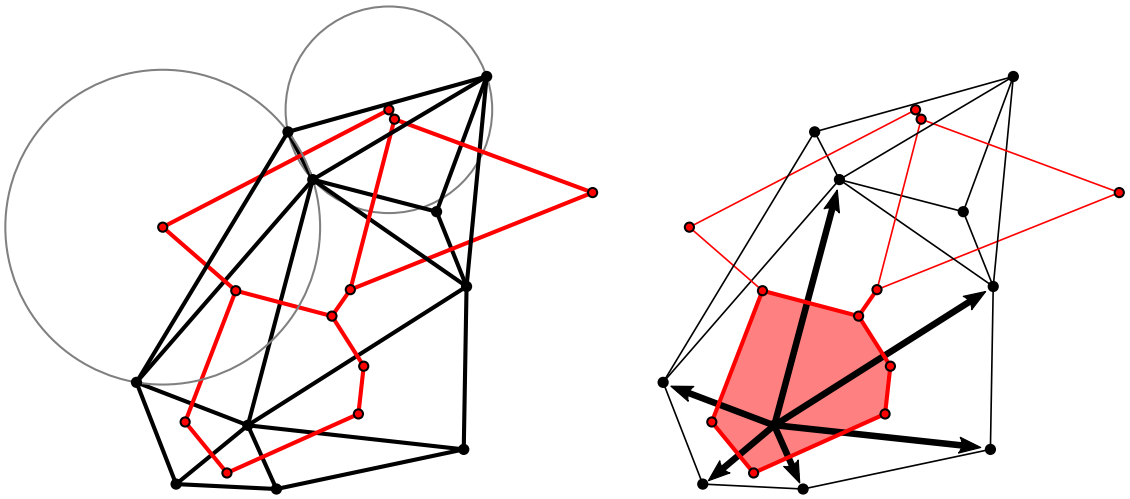


Figure 2.2: (left) The Delaunay triangulation of a random set of ten points. The Bowyer-Watson algorithm was used to create this triangulation. The Delaunay triangles are illustrated in black and the Voronoi cells in red. Only the Voronoi cells corresponding to interior Delaunay points are shown. Two circumcircles of the Delaunay triangles are also included. (right) One of the divergence stencils corresponding to the shaded Voronoi cell is outlined with arrows.

circumspheres. An example of a Delaunay triangulation of a random set of points is shown in Figure 2.2. The Bowyer-Watson algorithm [34, 35] was used to generate this triangulation. In the left of Figure 2.2, the Delaunay triangles are drawn in black and the Voronoi cells (most commonly referred to as the Voronoi diagram [36]) corresponding to the interior Delaunay nodes are drawn in red. Two circumcircles which inscribe triangles are also drawn in gray. This simple triangulation demonstrates the powerful relationship between the Delaunay triangles and the Voronoi polygons. Each Voronoi polygon corresponds to a single Delaunay node in the mesh. The faces of this Voronoi polygon (edges) correspond to edges which are connected to the Voronoi polygon's node. The normals of the Voronoi polygon's faces point in the *exact* same direction as their corresponding Delaunay edges, by construction. The importance of this relationship now becomes clear in terms of the box integration method. A divergence of a vector field integrated over each Voronoi polygon can be represented as a discrete summation of fluxes, defined on the Delaunay edges, across each face of the Voronoi polygon. One divergence stencil corresponding to the shaded Voronoi polygon is illustrated in Figure 2.2 (right).

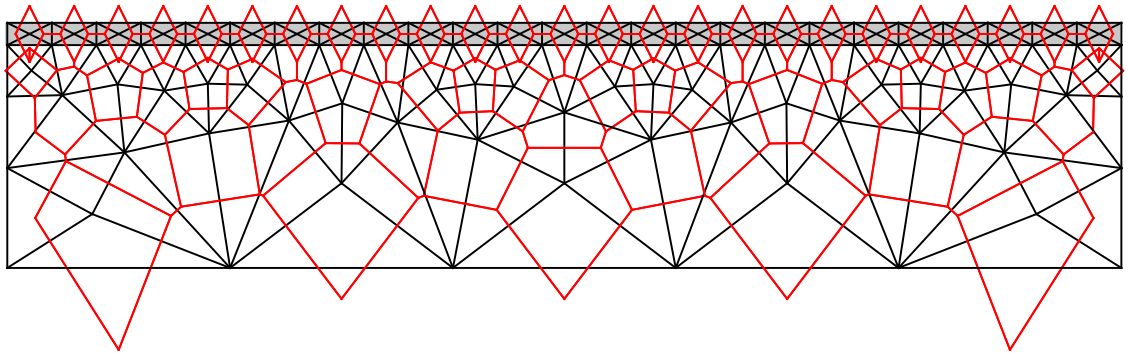
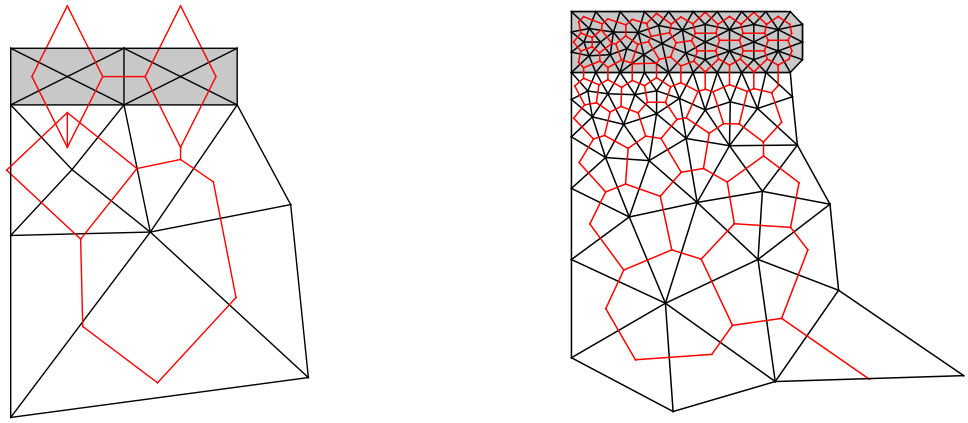


Figure 2.3: The Delaunay triangulation (black lines) and the corresponding Voronoi diagram (red lines) of the basic outline of a MESFET. This mesh was generated with the open source program Gmsh which uses the TetGen Delaunay triangulator.

There exists a wide range of algorithms to generate Delaunay triangulations. Included is the Bowyer-Watson algorithm, as well as other standard methods used in computational geometry [37, 38]. These methods are reliably used for the finite element method (FEM), among many other computational methods. An out-of-the-box Delaunay meshing algorithm, however, does not produce adequate meshes for simulating semiconductor devices with the box integration method.

To elaborate on the difficulties of generating a quality Delaunay-Voronoi (DV) mesh for discretization with the box integration method, a simple MESFET geometry is considered. The device consists of a thin semiconductor resting upon a larger substrate. These types of features are common in electronic devices. Figure 2.3 illustrates an attempt at meshing a basic MESFET with an open-source Delaunay meshing tool. The code is Gmsh [39], which uses the software TetGen [40] as its Delaunay mesher. It is evident that the Voronoi polygons span across the material interfaces as well as overall geometry boundaries. This is due to obtuse angles in triangles near the boundaries whose circumcenters fall outside of the domain boundary. A characteristic like this will not allow proper simulation of electronic devices as the charge densities must be uniquely defined in each material. Furthermore, numerical boundary conditions (BCs) rely on proper truncation of the Voronoi cells at mesh boundaries.



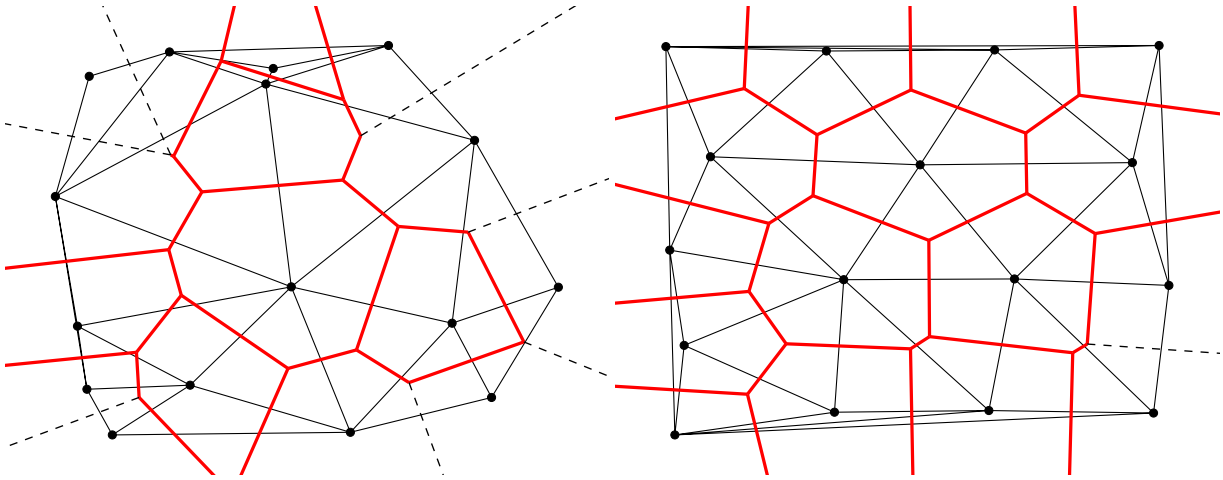
(a) (b)

Figure 2.4: A zoom of the Gmsh generated Delaunay mesh of the simple MESFET geometry. The gray shaded area is the semiconductor and the non-shaded region is the insulating substrate. (a) A coarse mesh which does not preserve material interfaces and geometry boundaries. (b) A refined mesh which over-meshes the charge transport direction.

As an attempt to fix the Voronoi polygons, the mesh is refined in the semiconductor region. Figure 2.4 illustrates (a) the original mesh and (b) the refined mesh in the top left corner of the device. The refined mesh produces conforming Delaunay triangles in both regions as well as Voronoi polygons which preserve geometry boundaries and can preserve material interfaces. This mesh, however, is not optimal for simulating electronic devices. The transport direction in field

effect transistors is parallel to semiconductor/substrate interfaces or heterojunctions. Therefore, the refined mesh in Figure 2.4b contains too many Delaunay edges which are perpendicular to the transport direction. Generating a Delaunay mesh with standard FEM software which preserves material interfaces and geometry boundaries will generate superfluous edges which leads to a very large number of degrees of freedom (DOF).

An alternative to simple mesh refinement could be sophisticated algorithms which ensure that circumcenters do not fall outside of their respective Delaunay triangles. One example is the centroidal Voronoi tessellation (CVT) algorithm [41]. The CVT algorithm attempts to move the points of the Delaunay triangulation to the centroids of their respective Voronoi polygons. Figure 2.5



(a) (b)

Figure 2.5: The (a) initial DV mesh, which is an input to the CVT algorithm and (b) the resulting DV mesh from the CVT algorithm.

presents an example of the CVT algorithm. Figure 2.5a is the initial Delaunay (and corresponding Voronoi diagram) triangulation of a random set of points. These mesh points are then moved according to the CVT algorithm to produce Figure 2.5b. Clearly, the CVT algorithm produces quality DV elements in the interior of the mesh. It does not prevent circumcenters from falling outside of the mesh boundary, however. Furthermore, because the algorithm is based upon moving mesh nodes, the CVT algorithm will not preserve the original device geometry. Several papers appear in the literature which address these problems. They report a stitching algorithm, which

aims to use the CVT algorithm in the interior of the geometry and stitch the resulting mesh onto a boundary conforming and preserving mesh [42, 43, 44].

A considerable amount of work exists in DV meshing algorithms specifically designed for semiconductor device geometries. A prominent algorithm which generates 3D Delaunay grids suitable for complex semiconductor device structures was pioneered by Conti, Hitschfeld, and Fichtner [45, 46, 47]. These algorithms decompose the global device geometry into sub-regions which can be refined in terms of the mesh. The sub-regions automatically preserve material interfaces and mesh boundaries and the interiors of the sub-regions are allowed to have circumcenters which fall outside of their respective element but still reside inside the sub-region. The commer-

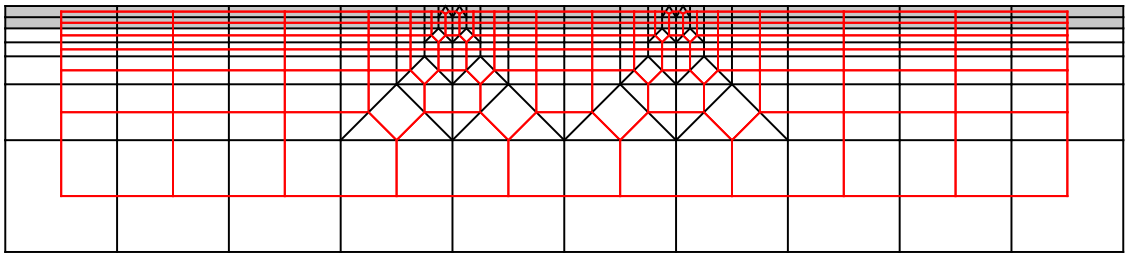
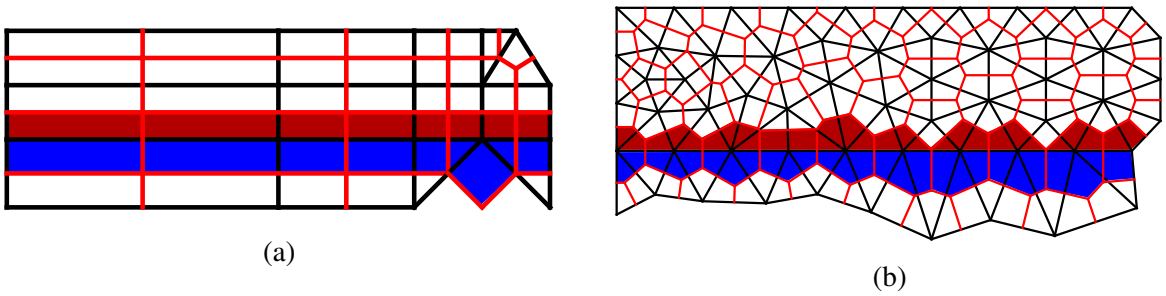


Figure 2.6: The Delaunay triangulation (black lines) and the corresponding polygons of the Voronoi diagram (red lines) of the basic outline of a MESFET. This mesh was generated with an in-house code based upon the work of Conti.

cial device simulator Sentaurus Device, a product of Synopsys, utilizes a separate meshing code which incorporates these algorithms in its meshing tools [48]. An example of a mesh generated by an in-house Air Force Research Laboratory, Sensors Directorate (AFRL) code is presented in Figure 2.6. Further progress on this meshing topic extends the work of Conti to produce boundary conforming meshes [49, 50].

Finally, a discussion on splitting the Voronoi mesh at material interfaces and truncating at mesh boundaries is required. Given that the triangles (2D) or tetrahedra (3D) on material interfaces or mesh boundaries do not have obtuse angles "pointing" towards the interface/boundary, the Voronoi mesh can be properly split or truncated. Figure 2.7 illustrates material interface truncation on (a) a mesh generated by the algorithm of [45] and (b) a mesh generated by Gmsh [39]. In both instances, the Voronoi cells corresponding to Delaunay nodes on the material interface have edges which are



(a) (b)

Figure 2.7: Two examples of mesh splitting at material interfaces. In both cases, the primary edges separating the red and blue shaded Voronoi polygons represent the material interface. (a) Material interface Voronoi splitting on a mesh generated by the algorithm of Conti. (b) The same material interface Voronoi splitting on a mesh generated by Gmsh.

exactly perpendicular to the material interface. The Voronoi edges intersect the Delaunay edges at exactly the circumcenter of the Delaunay edges — the midpoint of the edge. In 3D, the edges of a Voronoi polyhedron intersect at the circumcenters of the Delaunay triangles. The Voronoi diagram at the mesh boundaries uses the exact same recipe for truncation. The portion of the Voronoi cell which falls outside of the Delaunay mesh, however, is not included in the simulation domain. Figure 2.8 demonstrates the truncation of the Voronoi diagram at a boundary shaded in gray. The

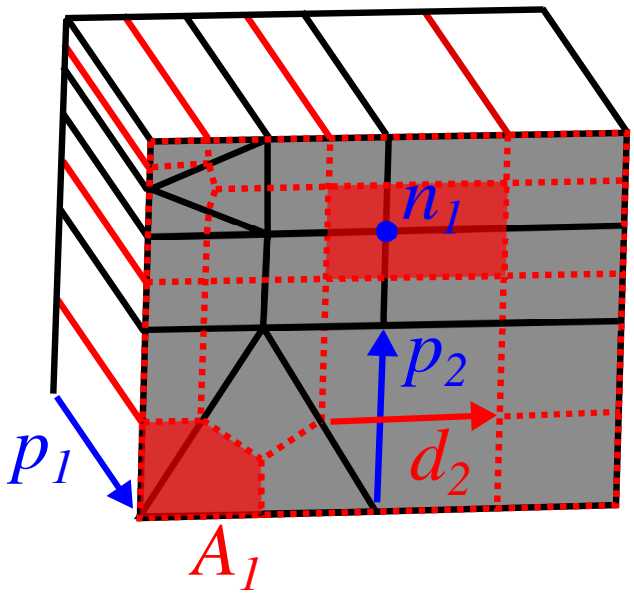


Figure 2.8: An example of Voronoi diagram splitting in 3D. The primary edges are drawn in black and the Voronoi edges are drawn in red. The interface of the Delaunay mesh is shaded in gray. The dashed red lines correspond to the Voronoi cells in the plane of the interface. Two Voronoi cells in the interface are shaded in red.

Voronoi edges in the plane of the boundary are drawn as dashed red lines. This boundary could be either a material interface or the end of the Delaunay mesh. If the boundary is a material interface, the Voronoi cells on the interface are used to calculate thermionic emission across the boundary. In the case of mesh boundaries, the Voronoi edges and cells on the plane are used for BCs. Both of these cases will be discussed in subsequent sections.

The following list summarizes the requirements of meshes suitable for semiconductor device simulation with the box integration method.

- The mesh must be Delaunay compliant, i.e., no nodes of any primary element fall within any of the circumcircles/circumspheres.
- Both the Delaunay triangulation and the Voronoi diagram must preserve mesh boundaries and material interfaces. In other words, the circumcenters of interior elements cannot cross mesh boundaries or material interfaces.

If these two criteria are met, the Voronoi cells corresponding to mesh nodes on boundaries/interfaces will have perpendicular edges to these boundaries/interfaces. These features are utilized to truncate the Voronoi diagram into material specific regions and the edges are used for device simulation BCs. An example of separating the Voronoi diagram into unique materials is illustrated in Figure 2.7.

2.3 Systems of Nonlinear Partial Differential Equations

Simulating charge transport by means of the FKT equations amounts to solving nonlinear partial differential equations (PDEs). Before complicated systems of nonlinear equations are derived and discretized, some important properties are discussed. Simple systems of nonlinear equations are presented in this section to discuss central concepts. First, the important topic of solving discrete nonlinear PDEs with iterative methods is presented. Then, characteristics of discrete nonlinear PDEs including convergence and stability are presented.

2.3.1 Numerical Solution of Discrete Nonlinear Equations

Subsequent chapters are devoted to the derivation and discretization of the nonlinear transport equations. Once a discrete set of equations is assembled, they must be solved for the independent variables. This section focuses on two prominent iterative methods, namely, the fixed point iteration method (FPM) and the Newton-Raphson or Newton's method. Given a set of N independent variables $\vec{x} = \{x_1, x_2, \cdots, x_N\}^T$ and N equations $\vec{F} = \{F_1(\vec{x}),\ F_2(\vec{x}),\ \cdots,\ F_N(\vec{x})\}^T$, the iterative methods attempt to calculate the independent variables which satisfy

$$\vec{F}(\vec{x}) = 0. \tag{2.4}$$

A very naive attempt at solving Eqn. (2.4) is a simple guess-and-check. The independent variables could be randomly sampled until they satisfy the residuals. However, there is no guarantee that this will ever produce the correct solution variables. Rather than randomly choosing the independent variables, the FPM and Newton's method provide "updates" to the solution variables in a systematic way. The following sections provide details on how the solutions are updated with the specific iterative methods.

2.3.1.1 Fixed Point Iteration Method

A set of independent variables $\vec{x}_P$ is a fixed point of the nonlinear equations $\vec{G}(\vec{x})$ if it satisfies $\vec{x}_P = \vec{G}(\vec{x}_P)$. The FPM attempts to use the fixed point to update the independent variables. To do so, the original nonlinear equations $\vec{F}(\vec{x}) = 0$ are algebraically converted to $\vec{x} = \vec{G}(\vec{x})$. If this manipulation is not possible, then the FPM cannot be used to iteratively solve the nonlinear equations. Given an initial guess of the independent variables $\vec{x}_0$, the FPM algorithm is

```
For m = 0 to M_max:
    Compute FP = G(x_m)
    Compute residual = x_m - FP
    If |residual| < tolerance: break
    Update solutions x_{m+1} = FP
```

Here, `M_max` represents a user-defined maximum iteration count and `tolerance` defines the convergence criterion of the nonlinear solver.

The benefit of using the FPM is that it only requires the algebraic manipulation of the nonlinear equations (if that manipulation is possible). It does not require the expensive calculation of a Jacobian matrix (see the following section on Newton's method). One drawback of the FPM is its slow convergence. The convergence of the FPM will be demonstrated in Section 2.3.1.3.

2.3.1.2 Newton's Method

Newton's method is a robust iterative solver for systems of nonlinear equations [51]. Given an initial guess of the independent variables $\vec{x}_0$ and the "change" in solution variables $\Delta\vec{x} = \vec{x}_{m+1} - \vec{x}_m$, the Newton algorithm is

```
For m = 0 to M_max:
    Compute residual vector "F"
    Compute Jacobian matrix "J"
    If |F| < tolerance: break
    Solve the linear system J*dx = -F
    Update solutions x^{m+1} += dx
```

Again, `M_max` represents a user specified maximum iteration count of the Newton algorithm and `tolerance` is the value that defines convergence of the nonlinear solver. The variable `dx` in the algorithm symbolizes the change in the independent variables $\Delta\vec{x}$.

This algorithm is a standard of solving nonlinear systems of equations because of its versatility and its quadratic convergence. However, Newton's method is known to fail if the initial guess is far from the solution. Source stepping is a common approach to alleviating convergence issues of Newton's method. For example, an applied bias of a device simulation is added incrementally and solutions to Newton's method are fed into the next solve. This method can also be applied to time stepping algorithms which require reduction of the time step. Newton's method requires the

calculation of the Jacobian matrix in the linear system

$$\begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \cdots & \frac{\partial F_1}{\partial x_N} \\ \frac{\partial F_2}{\partial x_1} & \frac{\partial F_2}{\partial x_2} & \cdots & \frac{\partial F_2}{\partial x_N} \\ \vdots & & \ddots & \vdots \\ \frac{\partial F_N}{x_1} & \frac{\partial F_N}{x_2} & \cdots & \frac{\partial F_N}{x_N} \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \\ \vdots \\ \Delta x_N \end{bmatrix} = - \begin{bmatrix} F_1(\vec{x}) \\ F_2(\vec{x}) \\ \vdots \\ F_N(\vec{x}) \end{bmatrix}. \tag{2.5}$$

Calculation of the Jacobian matrix and the solution of the linear system becomes expensive for large DOF.

2.3.1.3 Comparison of FPM and Newton's Method

A simple set of nonlinear equations is presented and solved with the FPM and Newton's method. This system of equations, with independent variables x and y, is

$$F_1(x,y) = 2x^2 + e^y - 7x = 0, \tag{2.6}$$

$$F_2(x,y) = 9y - e^x = 0. \tag{2.7}$$

The 2×2 Jacobian matrix required for Newton's method is

$$J_{ij}(x,y) = \begin{bmatrix} 4x-7 & e^y \\ -e^x & 9 \end{bmatrix}, \tag{2.8}$$

and the FPM functions are

$$G_1(x,y) = (2x^2 + e^y)/7, \tag{2.9}$$

$$G_2(x,y) = e^x/9. \tag{2.10}$$

The residuals $F_1(x,y)$ and $F_2(x,y)$ and their solution $F_1 = F_2 = 0$ are illustrated in Figure 2.9.

Figure 2.10 presents the norm of the residual $\sqrt{F_1^2 + F_2^2}$ at each iteration of the FPM and Newton's method. The red and blue lines illustrate the norm of the residuals calculated at each iteration of Newton's method and the FPM, respectively. Newton's method exhibits a quadratic rate of convergence and the FPM a linear rate of convergence. However, to achieve quadratic convergence,

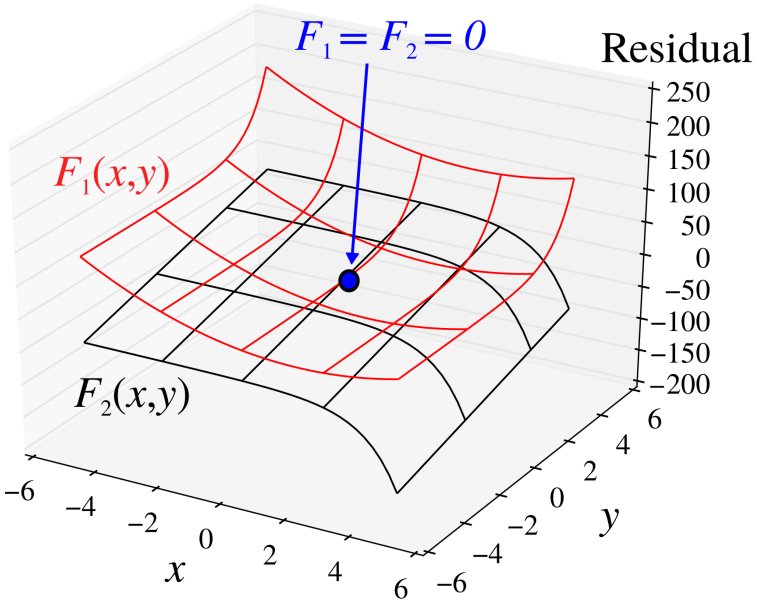


Figure 2.9: A set of nonlinear equations and its solution. The first and second residuals are plotted with red and black surfaces, respectively and the solution is illustrated with a blue circle.

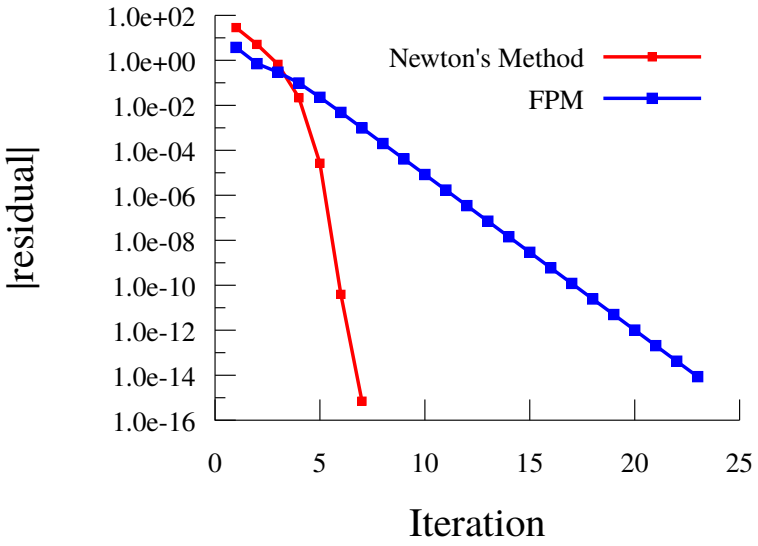


Figure 2.10: Residual vector norm calculated at each iteration of the FPM and Newton's method. Newton's method exhibits a quadratic rate of convergence. The FPM demonstrates a linear rate of convergence.

Newton's method must calculate a Jacobian matrix and solve the linear system of equations at each iteration. The FPM only requires the evaluation of the simple algebraic manipulation of the nonlinear equations. Newton's method is the preferred nonlinear solver for the discrete transport equations in subsequent chapters due to its robustness and strong convergence.

2.3.1.4 Other Iterative Methods

Solving the linear system of equations at each Newton iteration becomes burdensome with many DOF. Moreover, calculation of the Jacobian matrix can be difficult or even impossible in some cases. To overcome these obstacles, a class of nonlinear solvers called "quasi-Newton methods" was developed as an alternative to Newton's method. Among these is Broyden's method [52], which only requires the calculation of the Jacobian at the first iteration. Broyden's method still requires a linear solve at each iteration, however. Another important class of nonlinear solvers is the Newton-Krylov method [51]. Rather than directly solving the linear system of equations at each iteration with an LU factorization, these methods utilize Krylov subspace-based linear solvers such as GMRES. The calculation of the Jacobian can also be removed altogether with the Jacobian-free Newton-Krylov (JFNK) methods [53].

2.3.2 Mesh Convergence of Discretization Techniques

Discretizing systems of PDEs requires approximations of continuous derivatives in space and time. The box integration method is used for the FKT equations (see Section 2.1). An important characteristic of discretization techniques is the order of convergence of the methods. This is also known as mesh convergence. Quantifying mesh convergence requires evaluating the error in the numerical solutions on a series of meshes. The numerical solutions are compared to analytic solutions. If no analytic solutions are available, mesh convergence is quantified by using a highly discretized solution as the "exact" or reference solution.

Consider the 1D nonlinear first-order differential equation

$$\frac{dy}{dx} + 3y^2 = 0, \tag{2.11}$$

with the BC $y(x = 0) = 1$. The exact solution of the differential equation is

$$y_{\text{exact}}(x) = \frac{1}{1+3x}. \tag{2.12}$$

Because the solution function is singular at $x = -1/3$, the domain is bounded to $x > 0$. This differential equation is approximated with a backward difference

$$\frac{y_i - y_{i-1}}{\Delta x} + 3y_i^2 = 0 \tag{2.13}$$

and a central difference

$$\frac{y_{i+1} - y_{i-1}}{2\Delta x} + 3y_i^2 = 0. \tag{2.14}$$

Here, y_i is the i^{th} solution on a uniform 1D mesh with N nodes.

Eqns. (2.13) and (2.14) are solved with Newton's method. The Jacobian matrices of both discrete equations are illustrated in Figure 2.11. Both Jacobian matrices are calculated at the first Newton iteration with all initial solution variables set to zero. The $(0,0)$ element corresponds to the Dirichlet BC residual $y_0 - 1.0 = 0.0$. This BC is scaled by 5 in these plots. Figure 2.11a

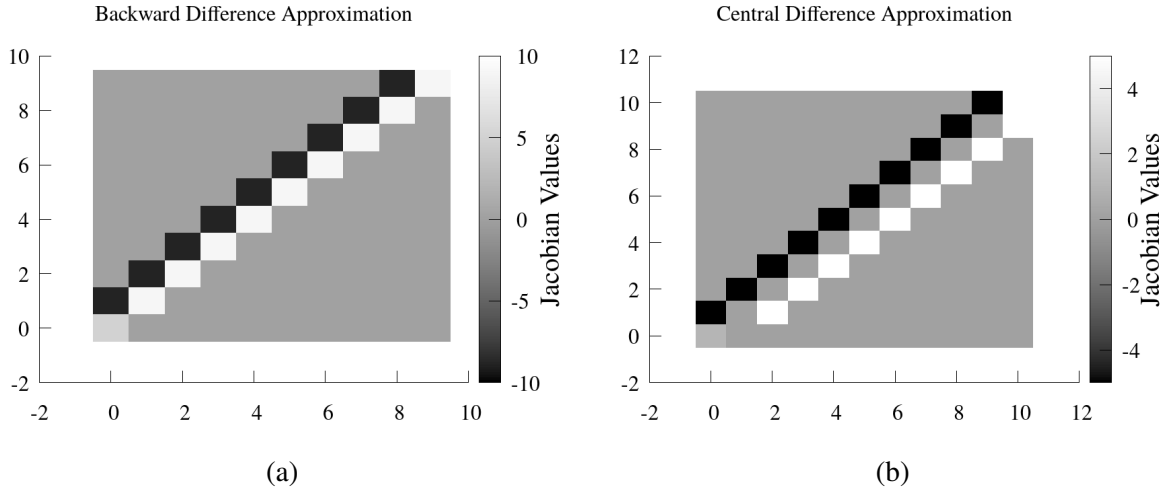


Figure 2.11: The Jacobian matrices of Eqns. (2.13) and (2.14) on a 10 and 11 node mesh, respectively. The last solution variable in (b) is governed by a backward difference approximation of Eqn. (2.11).

presents the backward difference Jacobian matrix on a 10 node mesh and Figure 2.11b presents the same for the central difference approximation on an 11 node mesh. The additional node and solution variable added to the central difference equations is solved with a backward difference approximation. The maximum and minimum values of the backward difference Jacobian correspond to the solution variables with the factor $1/\Delta x$. Figure 2.12 illustrates the discrete $y(x)$ solutions

calculated with the central difference approximation at each Newton iteration. The mesh size is $N = 11$. The analytic solution is drawn with a black dashed line.

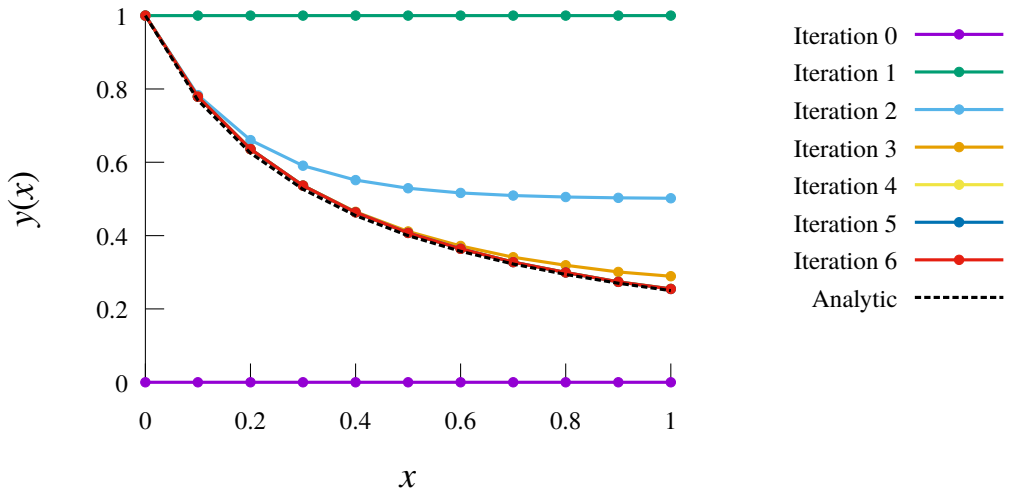


Figure 2.12: Discrete solutions of Eqn. (2.14) on an $N = 11$ node mesh. The solutions are plotted at each Newton iteration. The analytic solution is drawn as a dashed black line.

To quantify the mesh convergence of solving Eqns. (2.13) and (2.14), the L^2 error,

$$\varepsilon_k = \frac{\sqrt{\sum_i \left(y_i - y_{i,\text{exact}}\right)^2}}{\sqrt{\sum_i \left(y_{i,\text{exact}}\right)^2}}, \tag{2.15}$$

is calculated on a series of meshes. The order of convergence p is determined by the relation

$$\varepsilon_k < C N_k^{-p}, \tag{2.16}$$

where N_k is the number of mesh points (DOF) on the k^{th} mesh. Figure 2.13 presents the evaluation of Eqn. (2.15) on a series of meshes. The number of nodes in the series of meshes ranges from 10^2 to 10^5. Solutions are calculated from $x = 0$ to $x = 1$. The Dirichlet BC is $y_0 = 0$. Errors corresponding to the backward and central difference approximations are drawn in red and blue, respectively. It is clear that the backward difference and central difference approximations are first- and second-order discretization techniques. Second-order convergence of the central difference approximation is due to first-order error cancellation on the uniform mesh.

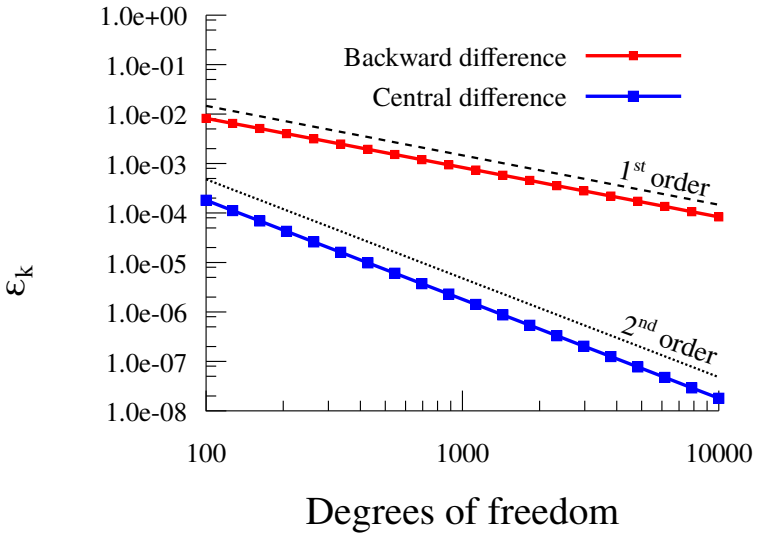


Figure 2.13: The error of the numerical solution of Eqn. (2.11) using a backward difference approximation (red line) and central difference approximation (blue line) of the spatial derivative. The backward difference approximation exhibits first-order mesh convergence. The central difference exhibits second-order mesh convergence.

2.3.3 Stability of Discrete Equations

Stability theory is a mathematical framework for analyzing the stability of nonlinear PDE solutions. The theory is prominent in systems theory and control engineering and is applied to the numerical solution of the nonlinear FKT equations. An autonomous system $\frac{dx}{dt} = f(x)$ has an equilibrium point x_e where $f(x_e) = 0$. The crux of stability theory [54] is as follows. The equilibrium point is

- stable if, for each perturbation of the equilibrium point, the time-integrated solutions $|x(t)|$ are bounded by a constant $\delta > 0$.
- unstable if it is not stable.
- asymptotically stable if it is stable and the perturbation returns to the equilibrium point as $t \to \infty$.

To elaborate on stability analysis of nonlinear equations, consider the autonomous system

$$\frac{dx}{dt} - 10x + 5xy = 0, \tag{2.17}$$

$$\frac{dy}{dt} - 3y - xy + 3y^2 = 0. \tag{2.18}$$

This nonlinear system has equilibrium points at $(x,y) = (0,0)$, $(0,1)$, and $(3,2)$. Figure 2.14

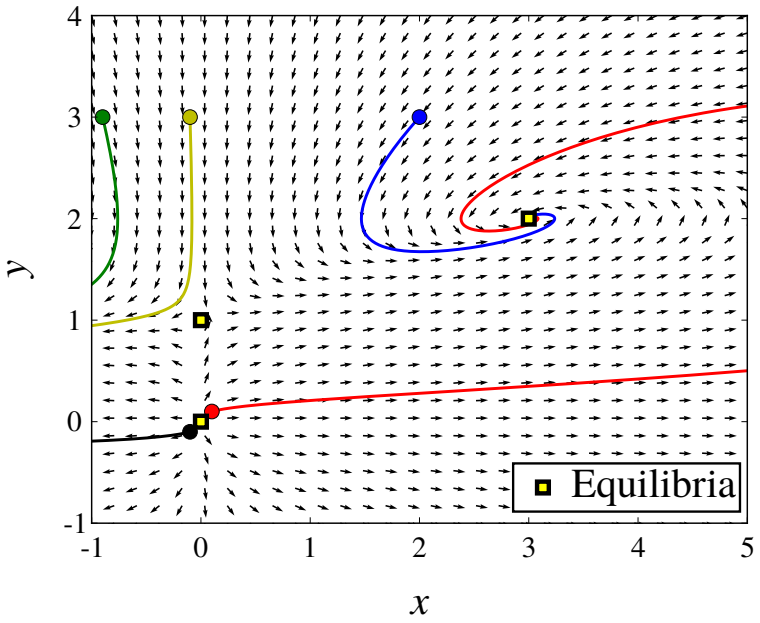


Figure 2.14: The direction field of Eqns. (2.17) and (2.18) and four trajectories with different initial conditions. The initial condition of each trajectory is illustrated with a solid circle.

presents the direction field of Eqns. (2.17) and (2.18). The equilibria are plotted at $(x,y) = (0,0)$, $(0,1)$, and $(3,2)$. The direction field is a vector field with components $(dx/dt, dy/dt)$. The solid lines in Figure 2.14 represent four trajectories of Eqns. (2.17) and (2.18) — the time evolution of the integrated solutions $(x(t), y(t))$. Each solid line starts at a different initial condition. It is clear from Figure 2.14 that the $(0,0)$ and $(0,1)$ equilibria are unstable. All trajectories diverge away from these points. The $(3,2)$ equilibrium point is asymptotically stable.

Direction fields and trajectory plots provide intuitive results for 2×2 systems of nonlinear equations. These tools do not apply to large systems of equations, however. Rather, the time evolution of a perturbation from an equilibrium point is analyzed by calculating

$$\Delta H(t) = \left| H(t) - H_{\text{equilib.}} \right|. \tag{2.19}$$

Here, $H_{\text{equilib.}}$ represents the equilibrium point of the N-dimensional system of nonlinear equations. $H(t)$ represents the time evolution of the integrated equations. Figure 2.15 presents the calculation of Eqn. (2.19) for each trajectory shown in Figure 2.14. All trajectories are compared

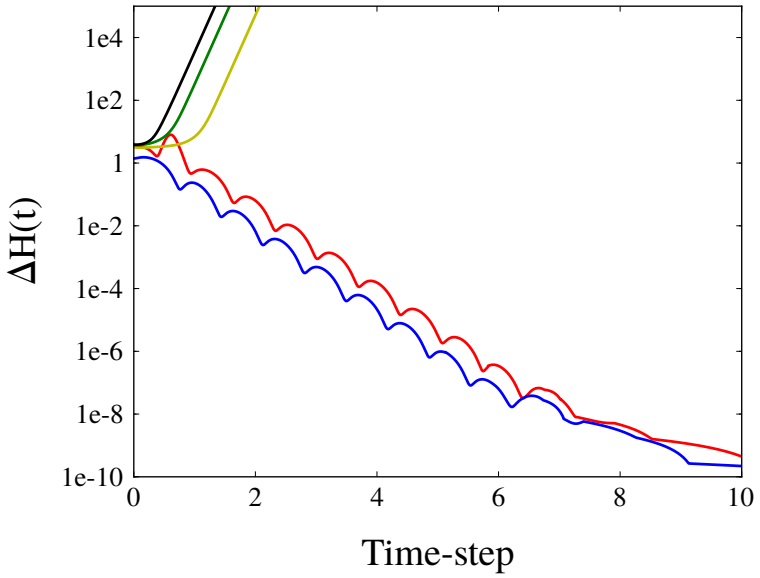


Figure 2.15: Calculation of Eqn. (2.19) for each perturbation presented in Figure 2.14. The five colors correspond to the five trajectories with different initial conditions. The black, green, and yellow lines indicate that the system of equations is not stable.

to the $H_{\text{equilib.}} = (3,2)$ equilibrium point. The perturbations drawn in red and blue converge to the $(3,2)$ equilibrium. The black, green, and yellow perturbations, however, diverge from the equilibrium points. This analysis provides quantitative stability results for large systems of nonlinear equations. The system of equations is stable if $\Delta H(t) \to 0$ as $t \to \infty$ for all perturbations.

2.4 RF Amplifiers

RF amplifier design relies heavily on accurate characterization and modeling of transistors. The transistor model must be valid over many operating conditions. For example, a circuit designer may require a valid model for a wide range of biases, a broad frequency range, several input powers, and various loading conditions. In what follows, the pertinent figures of merit (FOMs) for PA design are discussed. Linear two-port models and their uses in PA design are described in Section 2.4.2 and a discussion of large-signal PAs is provided in Section 2.4.3.

2.4.1 Amplifier Figures of Merit

The important FOMs for RF PA design include power gain, output power, the 1 dB gain compression point, *PAE*, and many others [1, 55]. The amplifiers are typically biased for peak transconductance, given by the maximum of

$$g_m = \frac{\partial I_D}{\partial V_{GS}}. \tag{2.20}$$

Here, V_{GS} is the gate-source bias and I_D is the steady-state drain current. The transconductance is calculated with a constant drain-source bias V_{DS}. Power gain of amplifiers is defined as the ratio of the output power to the input power. Gain is typically categorized as power gain G, available power gain G_A, and transducer power gain G_T. The definitions of the gain metrics are

$$G = \frac{P_L}{P_{in}}. \tag{2.21}$$

$$G_A = \frac{P_{avn}}{P_{avs}}, \tag{2.22}$$

$$G_T = \frac{P_L}{P_{avs}}, \tag{2.23}$$

Here, P_{in} is the power delivered to the network, P_{avs} is the power available from the source, P_{avn} is the power available from the network, and P_L is the power delivered to the load. Each gain is equal to the maximized gain, $G_T = G_A = G_P = G_{max}$, if the input and output are both conjugate matched to the device.

The output power is defined as the power delivered to the load P_L. For an ideal amplifier, the output power will increase linearly with respect to the input power. The output power of real transistors, however, will begin to saturate at high input power. This effect is quantified with the 1 dB gain compression point which is the point at which the gain decreases by 1 dB.

One of the most important FOMs in PA design is *PAE*. The PA stage consumes the most DC power in many wireless and Tx/Rx module applications and therefore it is critical to maximize the *PAE* for the optimal PA design. The *PAE* is defined as the ratio of the output to input power reference to the DC power level, or

$$PAE = \frac{P_L - P_{in}}{P_{DC}}. \tag{2.24}$$

This work will focus on the output power, power gain, and *PAE* for PA design purposes. There are many other important RF amplifier FOMs, however. Included is bandwidth, noise figure, input and output voltage standing wave ratio (VSWR), adjacent channel power ratio (ACPR), and stability [56, 1, 55].

2.4.2 Linear Two-Port Models

Linear two-port models are the standard representation of small-signal RF transistor responses. Figure 2.16 presents a general two-port model. Subscript 1 corresponds to the first port with voltage V_1 and current I_1. Here, the incident and scattered waves at both ports are

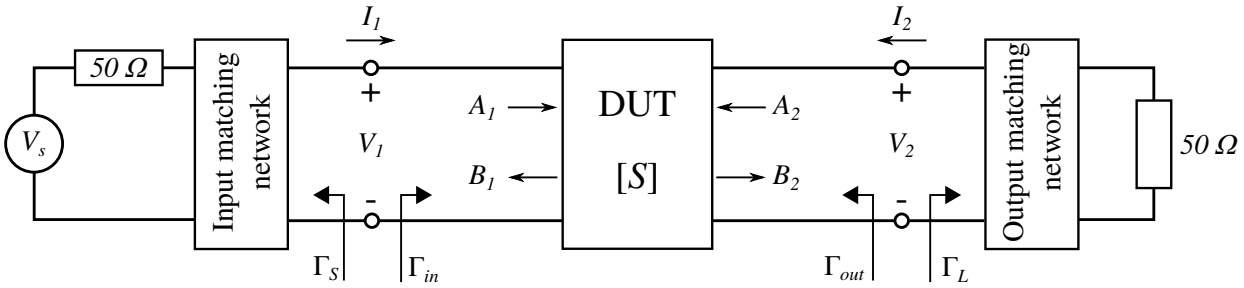


Figure 2.16: A linear two-port model of a DUT with input and output matching networks. This model is for a single fundamental frequency.

$$A_{1,2} = \frac{V_{1,2} + Z_0 I_{1,2}}{2}, \tag{2.25}$$

$$B_{1,2} = \frac{V_{1,2} - Z_0 I_{1,2}}{2}. \tag{2.26}$$

The reflection coefficient looking into the generator is Γ_S and Γ_{in} signifies the reflection coefficient looking into port 1. The same notation is used at port 2. The reflection coefficients looking into the generator and load are

$$\Gamma_S = \frac{Z_S - Z_0}{Z_S + Z_0}, \tag{2.27}$$

$$\Gamma_L = \frac{Z_L - Z_0}{Z_L + Z_0}. \tag{2.28}$$

The reference impedance is typically $Z_0 = 50\,\Omega$.

Scattering parameters (S-Parameters) describe the linear relation between the incoming waves A and outgoing waves B in the two-port model. The definition of S-Parameters is

$$\begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \tag{2.29}$$

where each coefficient is given by

$$S_{11} = \left.\frac{B_1}{A_1}\right|_{A_2=0}, \tag{2.30}$$

$$S_{21} = \left.\frac{B_2}{A_1}\right|_{A_2=0}, \tag{2.31}$$

$$S_{12} = \left.\frac{B_1}{A_2}\right|_{A_1=0}, \tag{2.32}$$

$$S_{22} = \left.\frac{B_2}{A_2}\right|_{A_1=0}. \tag{2.33}$$

The reflection coefficients looking into the input and output of the device are calculated from the S-Parameters

$$\Gamma_{in} = S_{11} + \frac{S_{12}S_{21}\Gamma_L}{1 - S_{22}\Gamma_L}, \tag{2.34}$$

$$\Gamma_{out} = S_{22} + \frac{S_{12}S_{21}\Gamma_S}{1 - S_{11}\Gamma_S}, \tag{2.35}$$

The input and output powers can now be defined in terms of the reflection coefficients and the S-Parameters. The power delivered to the network is

$$P_{in} = \frac{|V_S|^2}{8Z_0} \frac{|1-\Gamma_S|^2}{|1-\Gamma_S\Gamma_{in}|^2} \left(1 - |\Gamma_{in}|\right)^2, \tag{2.36}$$

the power available from the source is

$$P_{avs} = \frac{|V_S|^2}{8Z_0} \frac{|1-\Gamma_S|^2}{\left(1 - |\Gamma_S|^2\right)} \tag{2.37}$$

the power available from the network is

$$P_{avn} = \frac{|V_S|^2}{8Z_0} \frac{|S_{21}|^2 |1-\Gamma_S|^2}{|1-S_{11}\Gamma_S|^2 \left(1 - |\Gamma_{out}|^2\right)} \tag{2.38}$$

and the power delivered to the load is

$$P_L = \frac{|V_S|^2}{8Z_0} \frac{|S_{21}|^2 \left(1-|\Gamma_L|^2\right) |1-\Gamma_S|^2}{|1-S_{22}\Gamma_L|^2 |1-\Gamma_S\Gamma_{in}|^2}. \tag{2.39}$$

The small-signal gain can be calculated directly as

$$G = \frac{P_L}{P_{in}} = \frac{|S_{21}|^2 \left(1-|\Gamma_L|^2\right)}{\left(1-|\Gamma_{in}|^2\right) |1-S_{22}\Gamma_L|^2}, \tag{2.40}$$

as well as the available power gain

$$G_A = \frac{P_{avn}}{P_{avs}} = \frac{|S_{21}|^2 \left(1-|\Gamma_S|^2\right)}{|1-S_{11}\Gamma_S|^2 \left(1-|\Gamma_{out}|^2\right)}, \tag{2.41}$$

and the transducer gain

$$G_T = \frac{P_L}{P_{avs}} = \frac{|S_{21}|^2 \left(1-|\Gamma_S|^2\right) \left(1-|\Gamma_L|^2\right)}{|1-\Gamma_S\Gamma_{in}|^2 |1-S_{22}\Gamma_L|^2} \tag{2.42}$$

S-Parameters can also be converted to other representations of the linear two-port model, including the Z-, Y-, h-, and T- Parameters. For example, the small-signal current gain is one of the h-parameters

$$h_{21} = \frac{-2S_{21}}{(1-S_{11})(1+S_{22})+S_{12}S_{21}}. \tag{2.43}$$

The impedance and admittance parameters are also useful for de-embedding and circuit model extraction purposes.

S-Parameter simulation of a GaAs FET circuit model in ADS is presented to illustrate the use of small-signal circuit models. The layout of the S-Parameter simulation in ADS is shown in Figure 2.17. The gate to source bias is set by the parameter "Vlow" and the drain to source bias is set by "Vhigh". The simulation is over the 1–3 GHz frequency range and with input and output port impedances of 25 Ω. Figure 2.18 presents the S_{11} and S_{22} S-Parameters, the input and output reflection coefficients Γ_{in} and Γ_{out}, and the three power gains G, G_A, and G_T. The input and output reflection coefficients differ from S_{11} and S_{22} because Γ_S and Γ_L are nonzero. Furthermore, the three power gains produce different metrics over the frequency range.

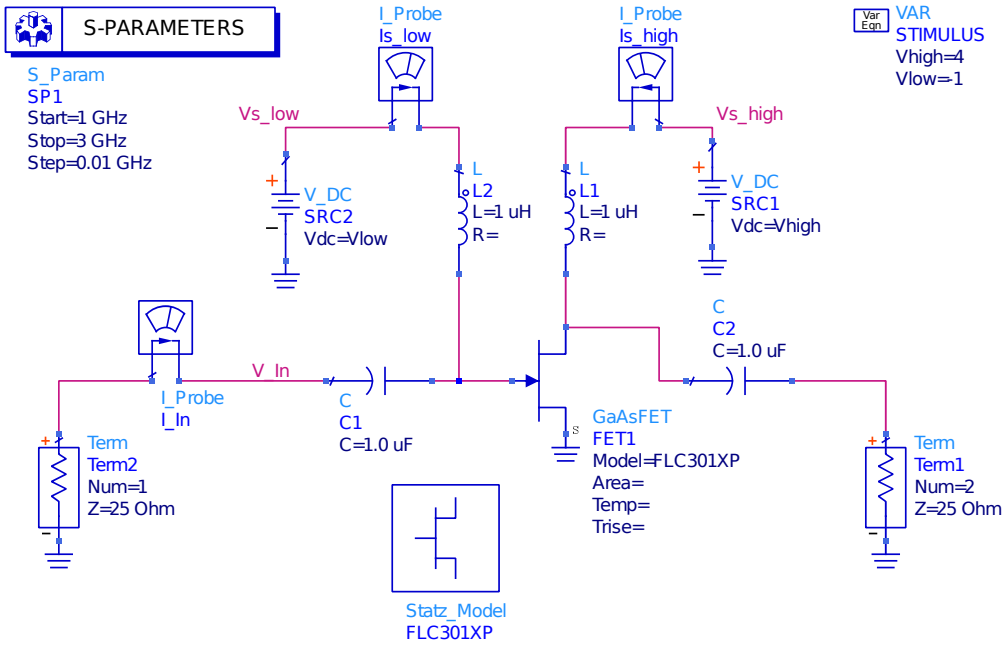


Figure 2.17: The ADS layout of S-Parameter simulation of a GaAs FET circuit model.

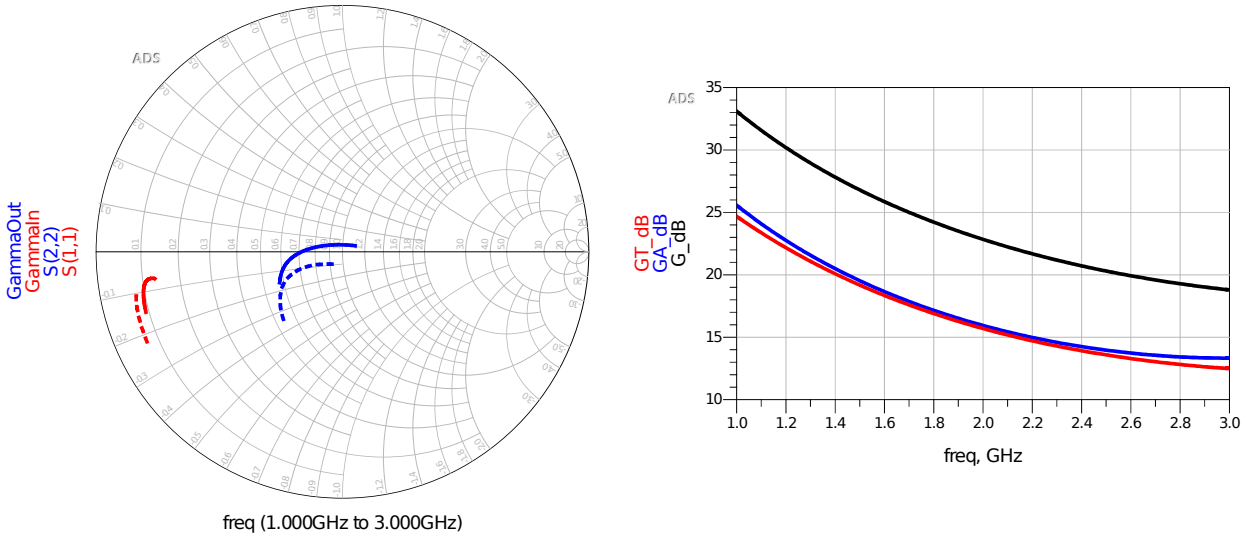


Figure 2.18: S-Parameters, reflection coefficients, and power gains of the GaAs FET circuit model simulated in ADS. The reflection coefficients on the Smith chart are the dotted lines. The ADS variables GammaIn and GammaOut and the power gain variables G_dB, GA_dB, and GT_dB correspond to Eqns. (2.34), (2.35), (2.40), (2.41), and (2.42), respectively.

2.4.3 Large Signal Analysis

The linear two-port models of the previous section do not accurately describe the nonlinear behavior of transistors. Because the models are linear, they do not depend on the input power level or the

loading conditions of the device. For power levels comparable to or above the 1 dB compression point, however, the measured S-Parameters will depend on input power level and loading conditions of the device. Small-signal S-Parameters can therefore not be used for large-signal simulation

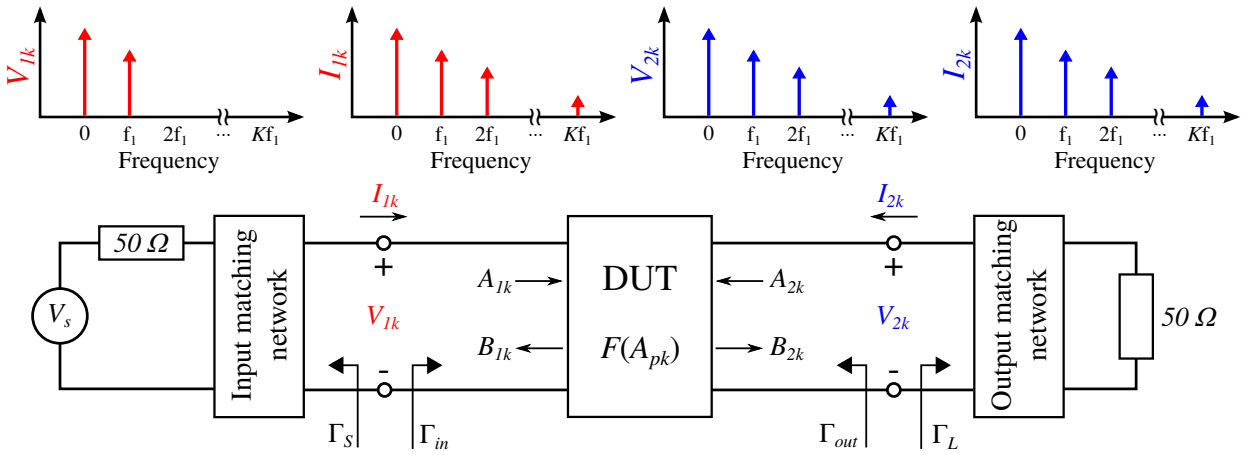


Figure 2.19: A nonlinear two-port model of a DUT with matching networks.

of devices.

Figure 2.19 illustrates a generalization of the two-port model to include the K harmonic response of a nonlinear device. In this figure, the input voltages and currents are designated as V_{1k} and I_{1k}, where the subscript 1 refers to port 1 and the subscript k is the index of the K harmonics. The fundamental frequency is $k = 1$ and the "DC" term is $k = 0$. Port 2 uses the same notation. Because transistors are inherently nonlinear, a large-signal applied to port 1 will generate multiple harmonics at port 2. This is the inter-modulation effect.

The power definitions for the nonlinear two-port model in Figure 2.19 are different than those for the linear analogue. Reflection coefficients Γ_S, Γ_{in}, Γ_{out}, and Γ_L are evaluated at the large-signal fundamental frequency ($k = 1$). The input power, power available from the source, and power delivered to the load are

$$P_{in} = \frac{1}{2}\mathrm{Re}\left\{V_{11}I_{11}^*\right\}, \tag{2.44}$$

$$P_{avs} = \frac{|V_S|^2}{8R_S}, \tag{2.45}$$

$$P_L = \frac{1}{2}\mathrm{Re}\left\{V_{21}I_{21}^*\right\}, \tag{2.46}$$

respectively. The factor R_S in the large-signal P_{avs} definition is the real component $R_S = \mathrm{Re}\{Z_S\}$ of the source impedance [56]. V_S is a large-signal tone at the fundamental frequency. The power gain and PAE are

$$G = \frac{P_L}{P_{in}}, \tag{2.47}$$

$$PAE = \frac{P_L - P_{in}}{P_{DC}}, \tag{2.48}$$

respectively and the DC power dissipation is

$$P_{DC} = V_{10} I_{10}. \tag{2.49}$$

An HB simulation layout of the GaAs FET model in ADS is illustrated in Figure 2.20. The fundamental frequency of the simulation is 1 GHz and the power available from the source P_{avs} is swept from 0 dBm to 25 dBm. Each FOM is plotted over the P_{avs} sweep in Figure 2.21. The x-axis of the plot is the input power P_{in}, in dBm, resulting from the applied P_{avs}. Here, the output power P_{out} is the solid black line, the power gain G is the solid red line, the transducer gain G_T is a dashed pink line, and the PAE is the solid blue line. The output power linearly responds to the input power at small-signal levels. However, the transistor begins to exhibit strong nonlinearity at the 1 dB compression point. This is approximately $P_{in} = 10$ dBm input power.

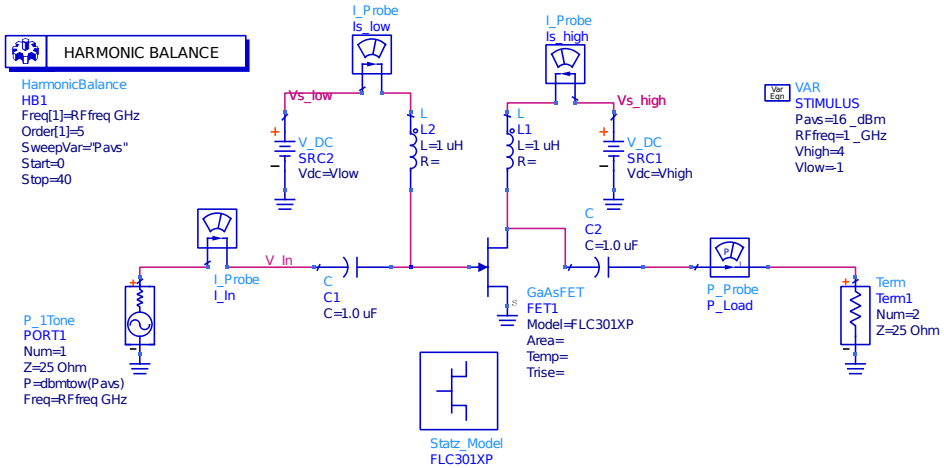


Figure 2.20: The layout of the HB simulation of a GaAs FET circuit model in ADS.

Characterizing large-signal responses of transistors is most commonly accomplished by plotting constant FOM contours on a Smith chart as a function of the load reflection coefficient, i.e.,

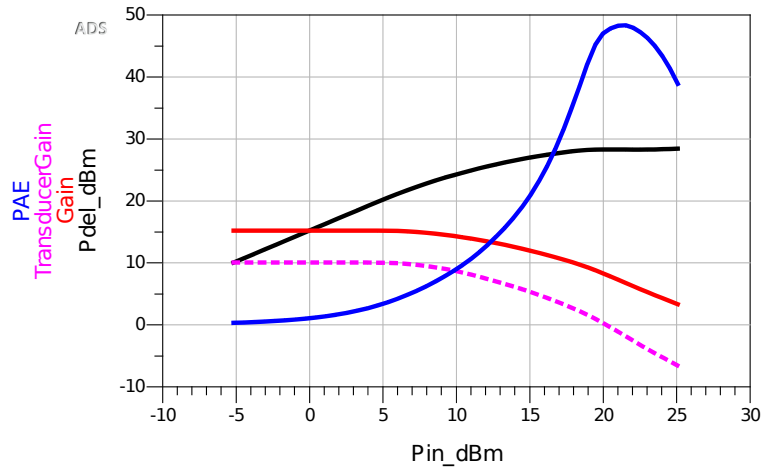


Figure 2.21: The output power "Pdel_dBm", power gain "Gain", transducer gain, and *PAE* of the GaAs FET circuit model calculator with HB in ADS. The 1 dB compression point is approximately 10 dBm input power.

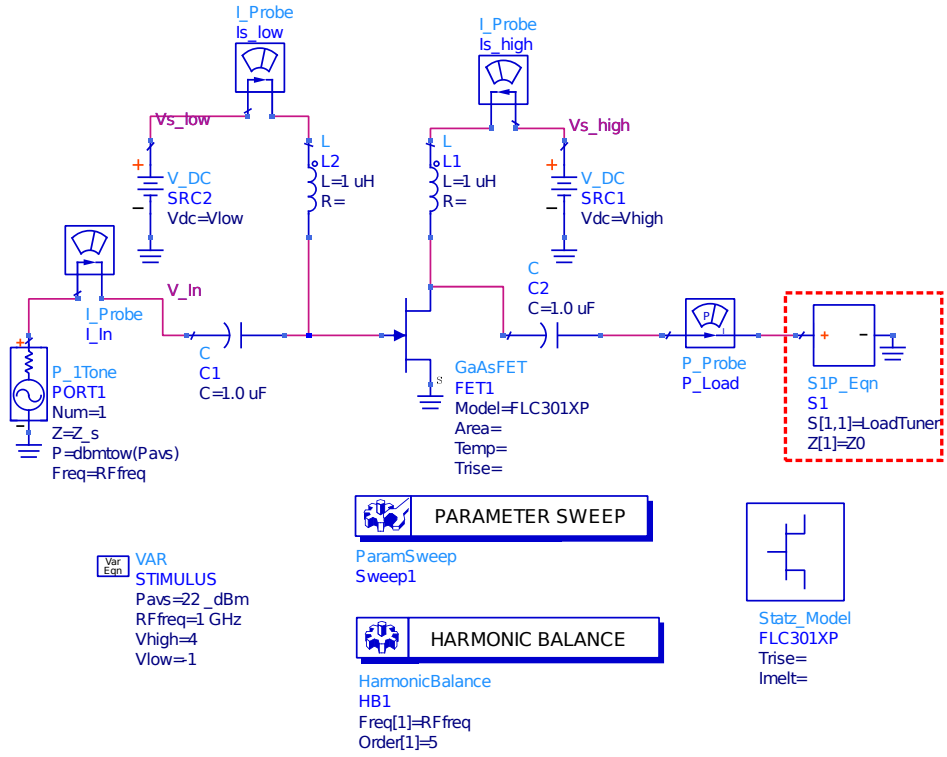


Figure 2.22: The layout of the LP simulation of a GaAs FET circuit model in ADS. HB is used to simulate the circuit and the load tuner (dashed red box) changes the reflection coefficient presented to the output of the transistor.

load-pull (LP). The FOM is plotted at the fundamental frequency. LP data is readily measured using systems ranging from computer-controlled electro-mechanical stub tuners to modern active

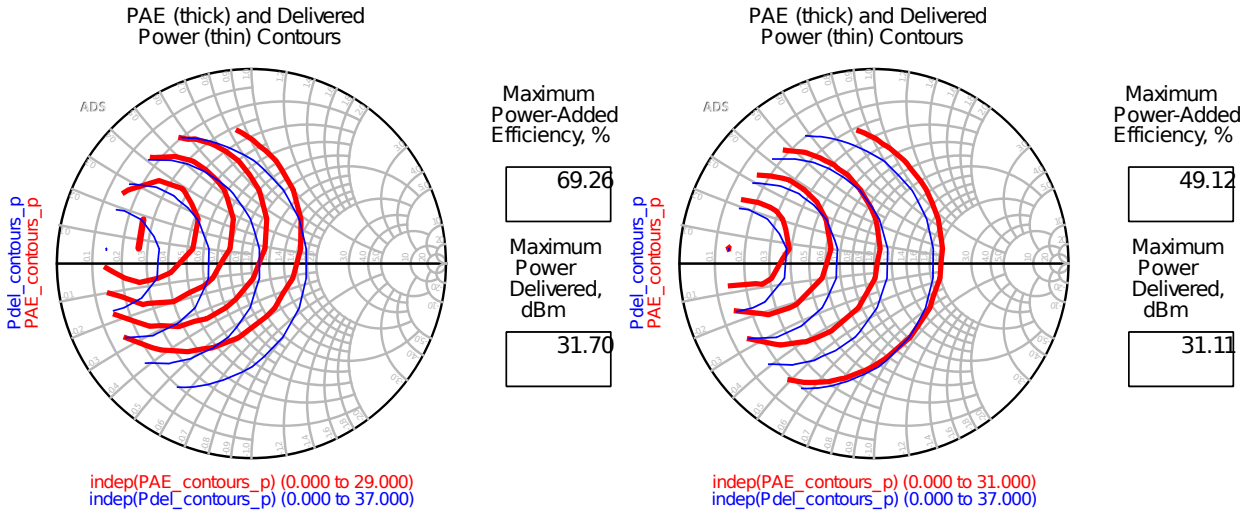


Figure 2.23: LP simulation of the GaAs FET circuit model in ADS at a center frequency of (left) 0.5 GHz and (right) 1 GHz. The red lines are the *PAE* contours and the blue lines are the output power contours. The *PAE* and output power FOMs are maximized with different load impedances.

LP test benches [16]. PA designers create suitable matching networks for the impedance design target synthesized from the LP data [57].

ADS offers a suite of simulation examples in its DesignGuide to help their users. Included in the kit is an example of LP simulation of the GaAs FET nonlinear circuit model. Figure 2.22 illustrates the GaAs FET LP simulation in ADS. This layout is similar to Figure 2.20. In the LP simulation, the "S1P_Eqn" component presents a series of reflection coefficients to the output port of the transistor. The HB simulation is repeated for each load impedance to calculate the output power and *PAE*. Finally, contours of each FOM are plotted on the Smith chart in Figure 2.23.

The LP contours in Figure 2.23 are calculated at a fundamental frequency of (left) 0.5 GHz and (right) 1 GHz. Both LP simulations were driven by 22 dBm power available from the source. Clearly, the load impedance which maximizes the power delivered to the load does not maximize the *PAE* at 0.5 GHz. The simulations at 1 GHz also exhibit the same effect. Furthermore, the optimal impedances are different at each fundamental frequency. This simple example demonstrates some of the difficulties of the PA design process.

2.5 X-Parameters

The PHD behavioral model is reviewed and discussed. This model was later trademarked under Agilent Technologies which is now Keysight Technologies [26]. XP modeling is a black-box frequency-domain technique and is the mathematically correct super set of S-Parameters [25]. The PHD model was first presented in [58] and later derived in [24]. The XP (PHD) model was later used extensively for black-box modeling of nonlinear devices [59, 60, 27, 61]. Using the notation in Figure 2.19, the XP model is now presented and discussed. A general nonlinear model describes the scattered waves at all ports in relation to the incident waves at all ports. Because the model is nonlinear, however, each scattered wave harmonic depends on all incident wave harmonics at each port. To simplify the model while retaining important features, the nonlinear function is linearized around the large-signal tone A_{11}. The general form of the XP model is

$$B_{pm} = \sum_{qn} S_{pq,mn} P^{m-n} A_{qn} + \sum_{qn} T_{pq,mn} P^{m+n} A^*_{qn} \tag{2.50}$$

with $T_{p1,m1} = 0 \ \forall (q,n)$. The incident and scattered waves are calculated from the port voltages and currents by

$$A_{qn} = \frac{V_{qn} + Z_0 I_{qn}}{2}, \tag{2.51}$$

$$B_{pm} = \frac{V_{pm} - Z_0 I_{pm}}{2}. \tag{2.52}$$

Here, the subscript q refers to the port number and n the harmonic index. For two-port nonlinear models, $q = 1,2$. The terms $S_{pq,mn}$ and $T_{pq,mn}$ are the coefficients of the XP model and $P = \text{phase}\{A_{11}\}$ is the phase of the large-signal tone.

Calculating the coefficients of the XP model is similar to calculating S-Parameters. For the S-Parameters, two simulations are required to calculate Eqns. (2.30) – (2.33). Because the scattered waves B_{pm} depend on both the incident waves A_{qn} and the complex conjugate of the incident waves A_{qn}, the XP model requires $(4K-1)$ simulations/measurements to calculate the $2K \times (4K-1)$ coefficients, $S_{pq,mn}$ and $T_{pq,mn}$, of a two-port model. This is demonstrated through the following examples.

2.5.1 Two ports, one harmonic

The XP model with $q = 1,2$ ports and $k = 1$ harmonics is

$$B_{11} = S_{11,11}A_{11} + S_{12,11}A_{21} + T_{12,11}P^2A_{21}^*, \tag{2.53}$$

$$B_{21} = S_{21,11}A_{11} + S_{22,11}A_{21} + T_{22,11}P^2A_{21}^*. \tag{2.54}$$

There are two incident waves and scattered waves at each port in this model. The incident wave A_{11} is the large-signal tone and A_{21} is a small-signal tone at the output port. This notation is used throughout the rest of this work.

If the large-signal tone A_{11} is the only non-zero incident wave, i.e., $A_{21} = 0$, then the $S_{11,11}$ and $S_{21,11}$ coefficients are trivially calculated. However, the remaining four XP coefficients are not calculated with a non-zero A_{21} and $A_{11} = 0$. The XP coefficients are only valid for one large-signal input A_{11}. Therefore, this term must remain constant when calculating the coefficients of the model.

The four remaining coefficients are calculated with two more simulations/measurements. Each simulation is driven by the large-signal tone A_{11} and a small-signal tone A_{21}. The A_{21} incident wave of the second simulation must be 90° out of phase with the first simulation in order to calculate the remaining XP coefficients. The two simulations yield four equations. These are Eqns. (2.53) and (2.54) with two orthogonal incident wave sets. The remaining XP coefficients are calculated by solving the 4×4 linear system outlined in the following.

Calculating the six XP coefficients of a two-port, one-harmonic model is outlined with the following. The notation a_i represents a set of incident waves corresponding to a single simulation/measurement of the nonlinear device. For the rest of this work, the large-signal tone is a cosine. This simplifies the model with $P = 1$. The notation A_{qn}^C and A_{qn}^S is used to represent two out of phase incident waves. The superscript C represents a cosine wave and the superscript S a sine wave. Finally, the scattered waves calculated with the simulation/measurement a_i are $B_{qn}(a_i)$. With these definitions, the XPs are calculated by:

1. Apply $a_0 = \{A_{11}, A_{21}^0 = 0\}$

- Simulate/measure the $\{B_{11}(a_0), B_{21}(a_0)\}$ scattered waves
- Compute $S_{11,11} = B_{11}(a_0)/A_{11}$ and $S_{21,11} = B_{21}(a_0)/A_{11}$.

2. Apply $a_1 = \{A_{11}, A_{21}^C\}$ and $a_2 = \{A_{11}, A_{21}^S\}$
   - Simulate/measure the $\{B_{11}(a_1), B_{21}(a_1)\}$ and $\{B_{11}(a_2), B_{21}(a_2)\}$ scattered waves
   - Compute $S_{12,11}$, $S_{22,11}$, $T_{12,11}$, $T_{22,11}$ by solving the following linear system

$$\begin{bmatrix} A_{21}^C & 0 & \left(A^C\right)_{21}^* & 0 \\ 0 & A_{21}^C & 0 & \left(A^C\right)_{21}^* \\ A_{21}^S & 0 & \left(A^S\right)_{21}^* & 0 \\ 0 & A_{21}^S & 0 & \left(A^S\right)_{21}^* \end{bmatrix} \begin{bmatrix} S_{12,11} \\ S_{22,11} \\ T_{12,11} \\ T_{22,11} \end{bmatrix} = \begin{bmatrix} B_{11}(a_1) - S_{11,11}A_{11} \\ B_{21}(a_1) - S_{21,11}A_{11} \\ B_{11}(a_2) - S_{11,11}A_{11} \\ B_{21}(a_2) - S_{21,11}A_{11} \end{bmatrix} \tag{2.55}$$

2.5.2 Two ports, two harmonics

The general algorithm for calculating the coefficients of the XP model becomes more apparent with a two-port, two-harmonic example. This system of equations is

$$\begin{aligned} B_{11} &= S_{11,11}A_{11} + S_{12,11}A_{21} + S_{11,12}A_{12} + S_{12,12}A_{22} + T_{12,11}A_{21}^* + T_{11,12}A_{12}^* + T_{12,12}A_{22}^* \\ B_{12} &= S_{11,21}A_{11} + S_{12,21}A_{21} + S_{11,22}A_{12} + S_{12,22}A_{22} + T_{12,21}A_{21}^* + T_{11,22}A_{12}^* + T_{12,22}A_{22}^* \\ B_{21} &= S_{21,11}A_{11} + S_{22,11}A_{21} + S_{21,12}A_{12} + S_{22,12}A_{22} + T_{22,11}A_{21}^* + T_{21,12}A_{12}^* + T_{22,12}A_{22}^* \\ B_{22} &= S_{21,21}A_{11} + S_{22,21}A_{21} + S_{21,22}A_{12} + S_{22,22}A_{22} + T_{22,21}A_{21}^* + T_{21,22}A_{12}^* + T_{22,22}A_{22}^* \end{aligned}$$

Computing the 28 XP coefficients of the two-port, two-harmonic model is described in the following. Again, the notation $a_i = \{A_{11}, A_{12}, A_{21}, A_{22}\}$ represents a single simulation/measurement of the nonlinear device. The coefficients are calculated by:

1. Apply $a_0 = \{A_{11}, 0, 0, 0\}$
   - Simulate/measure the $\{B_{11}(a_0), B_{12}(a_0), B_{21}(a_0), B_{22}(a_0)\}$ scattered waves

- Compute $S_{11,11} = B_{11}(a_0)/A_{11}$, $S_{11,21} = B_{12}(a_0)/A_{11}$, $S_{21,11} = B_{21}(a_0)/A_{11}$, and $S_{21,21} = B_{22}(a_0)/A_{11}$

2. Apply $a_1 = \{A_{11}, A_{12}^C, 0, 0\}$ and $a_2 = \{A_{11}, A_{12}^S, 0, 0\}$

   - Simulate/measure the $\{B_{11}(a_1), B_{12}(a_1), B_{21}(a_1), B_{22}(a_1)\}$ and $\{B_{11}(a_2), B_{12}(a_2), B_{21}(a_2), B_{22}(a_2)\}$ scattered waves
   - Compute $S_{11,12}$, $S_{11,22}$, $S_{21,12}$, $S_{21,22}$, $T_{11,12}$, $T_{11,22}$, $T_{21,12}$, $T_{21,22}$ by solving the following linear system

$$\begin{bmatrix} A_{12}^C & 0 & 0 & 0 & (A_{12}^C)^* & 0 & 0 & 0 \\ 0 & A_{12}^C & 0 & 0 & 0 & (A_{12}^C)^* & 0 & 0 \\ 0 & 0 & A_{12}^C & 0 & 0 & 0 & (A_{12}^C)^* & 0 \\ 0 & 0 & 0 & A_{12}^C & 0 & 0 & 0 & (A_{12}^C)^* \\ A_{12}^S & 0 & 0 & 0 & (A_{12}^S)^* & 0 & 0 & 0 \\ 0 & A_{12}^S & 0 & 0 & 0 & (A_{12}^S)^* & 0 & 0 \\ 0 & 0 & A_{12}^S & 0 & 0 & 0 & (A_{12}^S)^* & 0 \\ 0 & 0 & 0 & A_{12}^S & 0 & 0 & 0 & (A_{12}^S)^* \end{bmatrix} \begin{bmatrix} S_{11,12} \\ S_{11,22} \\ S_{21,12} \\ S_{21,22} \\ T_{11,12} \\ T_{11,22} \\ T_{21,12} \\ T_{21,22} \end{bmatrix}$$

$$= \begin{bmatrix} B_{11}(a_1) - B_{11}(a_0) \\ B_{12}(a_1) - B_{12}(a_0) \\ B_{21}(a_1) - B_{21}(a_0) \\ B_{22}(a_1) - B_{22}(a_0) \\ B_{11}(a_2) - B_{11}(a_0) \\ B_{12}(a_2) - B_{12}(a_0) \\ B_{21}(a_2) - B_{21}(a_0) \\ B_{22}(a_2) - B_{22}(a_0) \end{bmatrix} \tag{2.56}$$

3. Apply $a_3 = \{A_{11}, 0, A_{21}^C, 0\}$ and $a_4 = \{A_{11}, 0, A_{21}^S, 0\}$

- Simulate/measure the $\{B_{11}(a_3),\, B_{12}(a_3),\, B_{21}(a_3),\, B_{22}(a_3)\}$ and $\{B_{11}(a_4),\, B_{12}(a_4),\, B_{21}(a_4),\, B_{22}(a_4)\}$ scattered waves
- Compute $S_{12,11},\ S_{12,21},\ S_{22,11},\ S_{22,21},\ T_{12,11},\ T_{12,21},\ T_{22,11},\ T_{22,21}$ by solving the following linear system

$$\begin{bmatrix} A_{21}^C & 0 & 0 & 0 & (A_{21}^C)^* & 0 & 0 & 0 \\ 0 & A_{21}^C & 0 & 0 & 0 & (A_{21}^C)^* & 0 & 0 \\ 0 & 0 & A_{21}^C & 0 & 0 & 0 & (A_{21}^C)^* & 0 \\ 0 & 0 & 0 & A_{21}^C & 0 & 0 & 0 & (A_{21}^C)^* \\ A_{21}^S & 0 & 0 & 0 & (A_{21}^S)^* & 0 & 0 & 0 \\ 0 & A_{21}^S & 0 & 0 & 0 & (A_{21}^S)^* & 0 & 0 \\ 0 & 0 & A_{21}^S & 0 & 0 & 0 & (A_{21}^S)^* & 0 \\ 0 & 0 & 0 & A_{21}^S & 0 & 0 & 0 & (A_{21}^S)^* \end{bmatrix} \begin{bmatrix} S_{12,11} \\ S_{12,21} \\ S_{22,11} \\ S_{22,21} \\ T_{12,11} \\ T_{12,21} \\ T_{22,11} \\ T_{22,21} \end{bmatrix}$$

$$= \begin{bmatrix} B_{11}(a_3) - B_{11}(a_0) \\ B_{12}(a_3) - B_{12}(a_0) \\ B_{21}(a_3) - B_{21}(a_0) \\ B_{22}(a_3) - B_{22}(a_0) \\ B_{11}(a_4) - B_{11}(a_0) \\ B_{12}(a_4) - B_{12}(a_0) \\ B_{21}(a_4) - B_{21}(a_0) \\ B_{22}(a_4) - B_{22}(a_0) \end{bmatrix} \tag{2.57}$$

4. Apply $a_5 = \{A_{11},\, 0,\, 0,\, A_{22}^C\}$ and $a_6 = \{A_{11},\, 0,\, 0,\, A_{22}^S\}$

- Simulate/measure the $\{B_{11}(a_5),\, B_{12}(a_5),\, B_{21}(a_5),\, B_{22}(a_5)\}$ and $\{B_{11}(a_6),\, B_{12}(a_6),\, B_{21}(a_6),\, B_{22}(a_6)\}$ scattered waves
- Compute $S_{12,12},\ S_{12,22},\ S_{22,12},\ S_{22,22},\ T_{12,12},\ T_{12,22},\ T_{22,12},\ T_{22,22}$ by solving the following linear system

$$\begin{bmatrix} A_{22}^C & 0 & 0 & 0 & (A_{22}^C)^* & 0 & 0 & 0 \\ 0 & A_{22}^C & 0 & 0 & 0 & (A_{22}^C)^* & 0 & 0 \\ 0 & 0 & A_{22}^C & 0 & 0 & 0 & (A_{22}^C)^* & 0 \\ 0 & 0 & 0 & A_{22}^C & 0 & 0 & 0 & (A_{22}^C)^* \\ A_{22}^S & 0 & 0 & 0 & (A_{22}^S)^* & 0 & 0 & 0 \\ 0 & A_{22}^S & 0 & 0 & 0 & (A_{22}^S)^* & 0 & 0 \\ 0 & 0 & A_{22}^S & 0 & 0 & 0 & (A_{22}^S)^* & 0 \\ 0 & 0 & 0 & A_{22}^S & 0 & 0 & 0 & (A_{22}^S)^* \end{bmatrix} \begin{bmatrix} S_{12,12} \\ S_{12,22} \\ S_{22,12} \\ S_{22,22} \\ T_{12,12} \\ T_{12,22} \\ T_{22,12} \\ T_{22,22} \end{bmatrix}$$

$$= \begin{bmatrix} B_{11}(a_5) - B_{11}(a_0) \\ B_{12}(a_5) - B_{12}(a_0) \\ B_{21}(a_5) - B_{21}(a_0) \\ B_{22}(a_5) - B_{22}(a_0) \\ B_{11}(a_6) - B_{11}(a_0) \\ B_{12}(a_6) - B_{12}(a_0) \\ B_{21}(a_6) - B_{21}(a_0) \\ B_{22}(a_6) - B_{22}(a_0) \end{bmatrix} \tag{2.58}$$

2.5.3 Two ports, K harmonics

The system of equations of a two-port, K-harmonic XP model is

$$\begin{aligned} B_{11} &= S_{11,11}A_{11} + \cdots + S_{12,1K}A_{2K} + T_{12,11}A_{21}^* + \cdots + T_{12,1K}A_{2K}^* \\ &\vdots \\ B_{1K} &= S_{11,K1}A_{11} + \cdots + S_{12,KK}A_{2K} + T_{12,K1}A_{21}^* + \cdots + T_{12,KK}A_{2K}^* \\ B_{21} &= S_{21,11}A_{11} + \cdots + S_{22,1K}A_{2K} + T_{22,11}A_{21}^* + \cdots + T_{22,1K}A_{2K}^* \\ &\vdots \\ B_{2K} &= S_{21,K1}A_{11} + \cdots + S_{22,KK}A_{2K} + T_{22,K1}A_{21}^* + \cdots + T_{22,KK}A_{2K}^* \end{aligned}$$

The $2K \times (4K-1)$ coefficients of the XP model are calculated by

1. Apply $a_0 = \{A_{11}, 0, \cdots, 0\}$

   - Simulate/measure the $B_{pm}(a_0) = \{B_{11}(a_0), \cdots, B_{1K}(a_0), B_{21}(a_0), \cdots, B_{2K}(a_0)\}$ scattered waves
   - Compute $S_{11,11}, \cdots, S_{11,K1}, S_{21,11}, \cdots, S_{21,K1}$:

$$
\begin{aligned}
S_{11,11} &= B_{11}(a_0)/A_{11} \\
&\vdots \\
S_{11,K1} &= B_{1K}(a_0)/A_{11} \\
S_{21,11} &= B_{21}(a_0)/A_{11} \\
&\vdots \\
S_{21,K1} &= B_{2K}(a_0)/A_{11}
\end{aligned}
$$

2. Apply $a^C_{ij} = \{A_{11}, 0, \cdots, A^C_{ij}, \cdots, 0\}$ and $a^S_{ij} = \{A_{11}, 0, \cdots, A^S_{ij}, \cdots, 0\}$

   - Simulate/measure the $B_{pm}(a^C_{ij}) = \{B_{11}(a^C_{ij}), \cdots, B_{1K}(a^C_{ij}), B_{21}(a^C_{ij}), \cdots, B_{2K}(a^C_{ij})\}$ and $B_{pm}(a^S_{ij}) = \{B_{11}(a^S_{ij}), \cdots, B_{1K}(a^S_{ij}), B_{21}(a^S_{ij}), \cdots, B_{2K}(a^S_{ij})\}$ scattered waves
   - Compute the $S_{pi,mj} = \{S_{1i1j}, \cdots, S_{1i,Kj}, S_{2i,1j}, \cdots, S_{2i,Kj}\}$ and $T_{pi,mj} = \{_{1i,1j}, \cdots, T_{1i,Kj}, T_{2i,1j}, \cdots, T_{2i,Kj}\}$ coefficients by solving the following linear system

$$
\begin{bmatrix} D^C_{ij} & (D^C_{ij})^* \\ D^S_{ij} & (D^S_{ij})^* \end{bmatrix} \begin{bmatrix} S_{pi,mj} \\ T_{pi,mj} \end{bmatrix} = \begin{bmatrix} B_{pm}(a^C_{ij}) - B_{pm}(a_0) \\ B_{pm}(a^S_{ij}) - B_{pm}(a_0) \end{bmatrix} \tag{2.59}
$$

   where $D^C_{ij} = \left(A^C_{ij}\right) I$, $(D^C_{ij})^* = \left(A^C_{ij}\right)^* I$, and I is the $2 \times K$ identity matrix.

3. Repeat # 2 for $2K-1$ simulations to calculate all $2K \times (4K-1)$ XP coefficients.

2.6 Load-Dependent X-Parameters

The XP model of the previous section can be used to calculate the response of a nonlinear device to the inputs A_{pk}. The XP coefficients are valid for a single large-signal tone A_{11} and any small-signal perturbations A_{pk}, $pk \neq 11$. The XP model can be made load-dependent by the substitution

$$A_{2k} = \Gamma_{L,k} B_{2k} \tag{2.60}$$

Here, port 2 is the output port and $\Gamma_{L,k}$ is the reflection coefficient presented to the output port. The load reflection coefficient is specified for each harmonic. The reflection coefficient substitution and its consequences on calculating the output of the nonlinear model is discussed with the following examples. Again, the large-signal tone A_{11} is a cosine and the phase term is $P = 1$.

2.6.1 Two ports, one harmonic

The system of equations of a two-port, one-harmonic XP model with a load reflection coefficient is

$$\begin{aligned} B_{11} &= S_{11,11}A_{11} + S_{12,11}A_{21} + T_{12,11}A_{21}^* \\ B_{21} &= S_{21,11}A_{11} + S_{22,11}A_{21} + T_{22,11}A_{21}^* \\ A_{21} &= \Gamma_{L,1}B_{21} \end{aligned}$$

The XP coefficients are computed with the algorithm described in the previous section. The reflection coefficient substitution yields

$$\begin{aligned} B_{11} - S_{12,11}\Gamma_1 B_{21} &= S_{11,11}A_{11} + T_{12,11}A_{21}^*, \\ B_{21} - S_{22,11}\Gamma_1 B_{21} &= S_{21,11}A_{11} + T_{22,11}A_{21}^*. \end{aligned}$$

The scattered waves can no longer be calculated with a simple summation in Eqn. (2.50). The XP coefficients are calculated by solving the linear system

$$\begin{bmatrix} 1 & -S_{12,11}\Gamma_1 \\ 0 & 1 - S_{12,11}\Gamma_1 \end{bmatrix} \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix} = \begin{bmatrix} S_{11,11}A_{11} + T_{12,11}A_{21}^* \\ S_{21,11}A_{11} + T_{22,11}A_{21}^* \end{bmatrix}. \tag{2.61}$$

2.6.2 Two ports, K harmonics

The load-dependent XP model is generalized to two-ports and K harmonics. The system of equations for this XP model is

$$
\begin{aligned}
B_{11} &= S_{11,11}A_{11} + \cdots + S_{12,1K}A_{2K} + T_{12,11}A_{21}^{*} + \cdots + T_{12,1K}A_{2K}^{*} \\
&\vdots \\
B_{1K} &= S_{11,K1}A_{11} + \cdots + S_{12,KK}A_{2K} + T_{12,K1}A_{21}^{*} + \cdots + T_{12,KK}A_{2K}^{*} \\
B_{21} &= S_{21,11}A_{11} + \cdots + S_{22,1K}A_{2K} + T_{22,11}A_{21}^{*} + \cdots + T_{22,1K}A_{2K}^{*} \\
&\vdots \\
B_{2K} &= S_{21,K1}A_{11} + \cdots + S_{22,KK}A_{2K} + T_{22,K1}A_{21}^{*} + \cdots + T_{22,KK}A_{2K}^{*} \\
A_{21} &= \Gamma_1 B_{21} \\
&\vdots \\
A_{2K} &= \Gamma_K B_{2K}
\end{aligned}
$$

The reflection coefficient substitution yields

$$
\begin{aligned}
B_{11} - S_{12,11}\Gamma_1 B_{21} \cdots - S_{12,1K}\Gamma_K B_{2K} &= S_{11,11}A_{11} + \cdots + S_{11,1K}A_{1K} + \cdots + T_{12,1K}A_{2K}^{*} \\
&\vdots \\
B_{1K} - S_{12,K1}\Gamma_1 B_{21} \cdots - S_{12,KK}\Gamma_K B_{2K} &= S_{11,K1}A_{11} + \cdots + S_{11,KK}A_{1K} + \cdots + T_{12,KK}A_{2K}^{*} \\
B_{21} - S_{22,11}\Gamma_1 B_{21} \cdots - S_{22,1K}\Gamma_K B_{2K} &= S_{21,11}A_{11} + \cdots + S_{21,1K}A_{1K} + \cdots + T_{22,1K}A_{2K}^{*} \\
&\vdots \\
B_{2K} - S_{22,K1}\Gamma_1 B_{21} \cdots - S_{22,KK}\Gamma_K B_{2K} &= S_{21,K1}A_{11} + \cdots + S_{21,KK}A_{1K} + \cdots + T_{22,KK}A_{2K}^{*}
\end{aligned}
$$

The scattered wave outputs of the nonlinear device are calculated by solving the general linear system

$$
\begin{bmatrix} \mathbf{I} & -\mathbf{S}_1 \\ 0 & \mathbf{I} - \mathbf{S}_2 \end{bmatrix} \begin{bmatrix} B_{1i} \\ B_{2i} \end{bmatrix} = \begin{bmatrix} S_{11,ij}A_{1j} + T_{11,ij}A_{1j}^{*} + T_{12,ij}A_{2j}^{*} \\ S_{21,ij}A_{1j} + T_{21,ij}A_{2j}^{*} + T_{22,ij}A_{2j}^{*} \end{bmatrix} \tag{2.62}
$$

Here, the i^{th} rows of the $K \times K$ matrices $\mathbf{S}_1$ and $\mathbf{S}_2$ are $\left[S_{12,i1}\Gamma_1, \cdots, S_{12,iK}\Gamma_K\right]$ and $\left[S_{22,i1}\Gamma_1, \cdots, S_{22,iK}\Gamma_K\right]$, respectively. The matrix $\mathbf{I}$ is the $K \times K$ identity matrix.

CHAPTER 3

DELAUNAY-VORONOI SURFACE INTEGRATION

3.1 Introduction

DVSI is a computational electromagnetics (CEM) method that operates on unstructured meshes and is amenable to moving charges. The latter is paramount for any electronic device solver. This method has been successfully utilized to simulate semiconductor devices with full-wave electromagnetics coupled simultaneously to a novel Boltzmann equation solver for the charge transport including a full band-structure description for the electronic states [62]. Because the focus of that paper was the simulation of electronic devices and comparisons with other charge transport methods, complete detail of the EM discretization method was not included. Therefore, the purpose of this chapter is to present the numerical method of DVSI in sufficient detail to allow for independent implementations and to ensure the reproducibility of key results in the emerging literature.

In order to introduce DVSI in a CEM context, consider the two dominant numerical approaches: integral equation (IE) solvers and differential equation (DE) solvers. Both have a common goal of computing EM scalar and/or vector fields over a domain of interest. CEM methods can further be decomposed into two sub-categories: frequency-domain (FD) and TD methods. Because the ultimate goal is to develop a CEM method which couples to nonlinear electronic transport which is inherently in the time-domain, frequency domain methods are not considered in this discussion.

IE methods seek to solve an EM problem in which the unknown quantity of interest lies within an integral. These unknown quantities are typically surface currents on the boundaries of the computational domain. Electric and magnetic fields are computed through scalar and vector potentials. An IE is most often derived from Maxwell's equations using the Green's function approach [63]. The crux of the approach is deriving a Green's function which satisfies the linear Helmholtz equation subject to point sources. IEs are then built upon the principle of linear superposition. Although

this method is a staple of antenna design and other radiation (free-space) applications, the Green's function method cannot be applied when the point source is inherently nonlinear and is very limited when the dielectric medium is inhomogeneous. Other methods for deriving IEs from Maxwell's equations may produce viable options for coupling to electronic transport, but they are unknown to the author. For a more detailed discussion of IEs, please see [63] and references therein.

TD-DE methods are much better suited for representing and solving full-wave EM phenomena in electronic devices. Unlike IE methods, DE methods seek EM solutions directly from Maxwell's equations in differential form. The two most common TD-DE categories can be classified as unstructured and structured grid methods. Even though these two methods are both DE solvers, they differ greatly in their computational approach and sophistication. Unstructured grid (mesh) methods first decompose the computational domain into a collection of elements which are typically tetrahedra. The governing equations are then written in a form suitable for discretization. This amounts to writing Maxwell's equations in curl-curl form for the well known Finite Element Method Time Domain (FEMTD) [63] or adding flux terms to Maxwell's equations for the Discontinuous Galerkin Time Domain (DGTD) method [64]. Both of these methods seek a variational solution by approximating the quantities of interest with collections of basis functions which are carefully tailored to satisfy the governing equations and BCs, e.g., vector finite elements. When coupling these discretization methods to charge transport, it is not clear how to construct a basis set which approximates an equation with exponentially nonlinear dependence on the quantity of interest, such as the exponential relationship between carrier density in a semiconductor and the potential.

A separate discretization scheme, called the Finite Volume Method (FVM) [65], is a DGTD method with zeroth order basis functions. This method does not require a basis set to describe the solution variables and therefore does not have equivalent limitations with regard to nonlinear dependencies. However, special care must be taken when deriving the flux terms. An exhaustive search has yielded no viable options for representing full-wave EM in an electronic device simulator using DGTD, although attempts to construct the necessary flux terms are undoubtedly an active

area of research in the DG numerical methods community.

The other class of the DE methods is a structured grid method and is most commonly referred to as Finite Difference Time Domain (FDTD). It does not require a basis set since all partial derivatives are approximated with Taylor Expansions [63]. This makes FDTD a viable method to couple full-wave EM to nonlinear charge transport and implementations have been reported in the literature [66]. One drawback to using FDTD, however, is that the method represents the spatial computational domain with structured grids. Even if these structured grids are spaced non-uniformly, it still becomes burdensome when representing the complicated features in modern electronic devices, e.g., wrap-around gates or recessed Ohmic contacts.

The DVSI method in this work approximates Maxwell's equations in integral form with piece-wise constant vector projections and tessellates the computational domain with an unstructured, staggered mesh. The fact that DVSI tessellates the computational domain with an unstructured mesh is what makes DVSI well suited for these purposes. Parts of DVSI can be seen in the literature, including the well known Box Integration Method [30] in the computational electronics community and the Finite Integration Method (FIT) [67] and the Co-Volume Method [42] which are used for linear EM applications. It is worth mentioning that in certain limits of geometry, i.e., rectangular mesh elements, the DVSI method collapses exactly to a traditional Yee cell, which is the classical FDTD co-volume solution technique for this class of problems [68, 69]. DVSI is a fully general version for arbitrary geometries and completely unstructured discretization similar to an unstructured form of FDTD [70]. In regards to an unstructured FDTD formulation, DVSI is preferred for electronic device simulation due to its amenability to realistic device BCs. This will be discussed in subsequent sections.

In what follows, a brief discussion of the governing equations and the staggered meshes is presented in Section 3.2. Then, complete discussions of the discretization of Maxwell's rotational equations and Poisson's equation are presented in Sections 3.3 and 3.4, respectively. Each section contains benchmark tests that demonstrate and validate the numerical methods through physically relevant and quantitative test cases. Section 3.5 combines these methods and introduces the nec-

essary additional information that results in a numerical framework for representing full-wave EM for the target engineering application of electronic device simulation. This section also includes a benchmark test to validate the complete discretization of the relevant Maxwell's equations.

3.2 Physical Equations and Their Geometric Representation

The relevant physical quantities needed to simulate full-wave EM for the purposes of electronic device simulation are the electric field $\vec{\mathbf{E}}(\vec{\mathbf{r}},t)$, the magnetic field $\vec{\mathbf{H}}(\vec{\mathbf{r}},t)$, and the electric potential $\Phi(\vec{\mathbf{r}},t)$. Total electric fields are separated into rotational and irrotational components, i.e., their Helmholtz decomposition $\vec{\mathbf{E}}(\vec{\mathbf{r}},t) = \vec{\mathbf{E}}_{\text{rot}}(\vec{\mathbf{r}},t) + \vec{\mathbf{E}}_{\text{irr}}(\vec{\mathbf{r}},t)$. The Coulomb Gauge, $\nabla \cdot \vec{\mathbf{A}} = 0$, is used throughout and is mathematically convenient. The technique is started by expressing Maxwell's equations in their standard integral forms

$$\iint_S \hat{n} \cdot \left[\varepsilon \frac{\partial}{\partial t} \left(\vec{\mathbf{E}}_{\text{rot}}(\vec{\mathbf{r}},t) - \nabla\Phi(\vec{\mathbf{r}},t) \right) + \vec{\mathbf{J}}(\vec{\mathbf{r}},t) \right] dS = \oint_\Gamma \vec{\mathbf{H}}(\vec{\mathbf{r}},t) \cdot d\vec{\ell}, \tag{3.1}$$

$$\iint_S \hat{n} \cdot \mu \frac{\partial \vec{\mathbf{H}}(\vec{\mathbf{r}},t)}{\partial t} dS = -\oint_\Gamma \vec{\mathbf{E}}_{\text{rot}}(\vec{\mathbf{r}},t) \cdot d\vec{\ell}, \tag{3.2}$$

$$-\iint_S \hat{n} \cdot \varepsilon \nabla\Phi(\vec{\mathbf{r}},t)\, dS = \iiint_V \rho(\vec{\mathbf{r}},t)\, dV. \tag{3.3}$$

As alluded to in the introduction, the computational domain is represented with staggered unstructured meshes known as the Delaunay mesh and its dual Voronoi diagram. Figure 3.1 illustrates a two-dimensional example of the staggered meshes. The left highlights the Delaunay triangles in blue and the right highlights the Voronoi polygons in red. In three dimensions, the Delaunay mesh is comprised of finite edges, triangles, and tetrahedra which are called primary elements. The Voronoi diagram is made up of finite edges, polygons, and polyhedra which are then called dual elements. The relationship between the primary mesh and the dual mesh is exploited to discretize Maxwell's equations and is the focus of subsequent sections. The open source software TetGen was used to generate all Delaunay-compliant meshes for this work. For an in-depth discussion on mesh generation, there is extensive computer science literature that is actively seeking Delaunay-Voronoi mesh optimization. These are compliant not only with DVSI, but also with engineering

numerical methods across disciplines, e.g., solutions to the Navier-Stokes equations within the marker-and-cell algorithm.

3.3 Maxwell's Rotational Equations

To begin the discussion of solving Maxwell's equations in a bounded region with DVSI, the rotational equations (i.e., Ampere's and Faraday's laws) are discretized and the pertinent BCs for electronic device simulation are applied to obtain results. Two benchmark examples are presented which validate the numerical method quantitatively through comparisons with analytic results.

3.3.1 DVSI Discretization

Solving Maxwell's rotational equations, with proper BCs, is convenient with the relationship between the Delaunay and Voronoi meshes. When solving the rotational equations, it is assumed that the electric field $\vec{\mathbf{E}}$ is the *total* field comprised of both a rotational and irrotational component. Ultimately, in Section 3.5 the fields are decomposed into rotational and irrotational components, but initially only with total field versions of Eqns. (3.1) – (3.2).

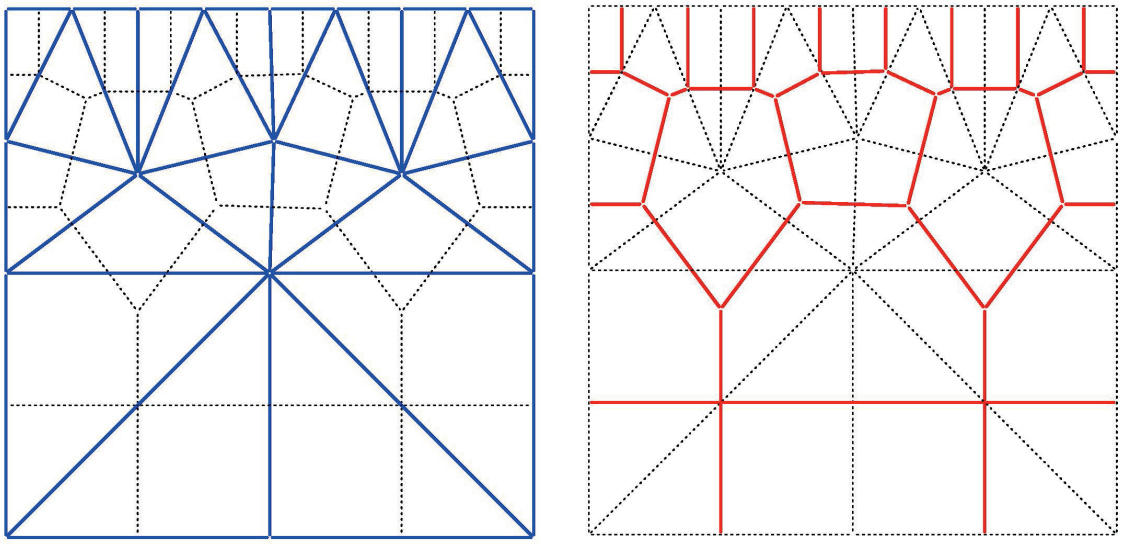


Figure 3.1: A two-dimensional example of the staggered DVSI mesh with the (left) Delaunay triangles drawn in blue and (right) the Voronoi polygons drawn in red.

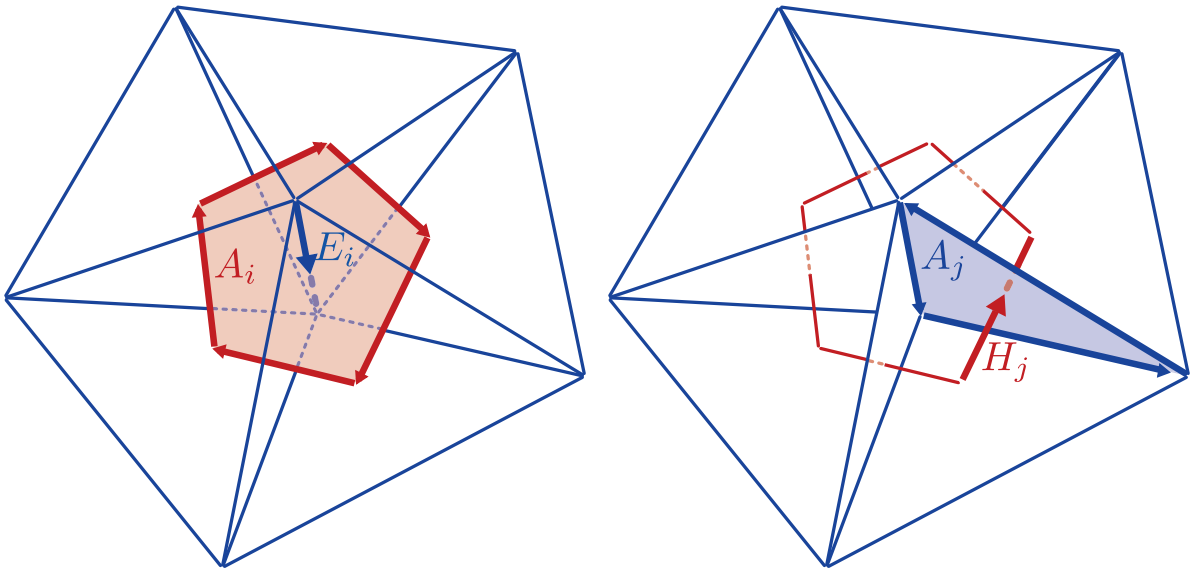


Figure 3.2: (left) Dual edges (shown as red arrows) associated with a primary edge (blue arrow). The dual polygon and dual edges are comprised of the circumcenters of each tetrahedron attached to the primary edge. (right) Primary polygon and edges associated with one of the dual edges.

$$\iint_S \hat{n} \cdot \left[\varepsilon \frac{\partial \vec{\mathbf{E}}(\vec{\mathbf{r}},t)}{\partial t} + \vec{\mathbf{J}}(\vec{\mathbf{r}},t) \right] dS = \oint_\Gamma \vec{\mathbf{H}}(\vec{\mathbf{r}},t) \cdot d\vec{\ell}, \tag{3.4}$$

$$\iint_S \hat{n} \cdot \mu \frac{\partial \vec{\mathbf{H}}(\vec{\mathbf{r}},t)}{\partial t} dS = -\oint_\Gamma \vec{\mathbf{E}}(\vec{\mathbf{r}},t) \cdot d\vec{\ell}. \tag{3.5}$$

Electric fields are represented as spatially constant vector projections on primary edges. Likewise, magnetic fields are represented as spatially constant vector projection quantities on dual edges. These relationships are depicted in Figure 3.2. The main points of DVSI discretization of the rotational equations are:

1. A mesh edge is always perpendicular to its associated polygon.

2. The solution variables representing electric and magnetic fields are spatially constant vector projections on mesh edges.

3. Materials reside within primary tetrahedra ensuring that primary edges do not cross dielectric boundaries, which is critical for applying BCs.

4. The entire volume retains the permeability of free space (we do not consider magnetic inhomogeneity as it is rare in semiconductor devices).

With this geometric description and these assumptions about electric and magnetic fields, the integral form of Maxwell's equations can now be written as the semi-discrete equations

$$\left(\varepsilon_i \frac{\partial E_i}{\partial t} + J_i\right) A_i = \sum_j H_j L_j, \tag{3.6}$$

$$\mu_0 \frac{\partial H_j}{\partial t} A_j = -\sum_i E_i L_i, \tag{3.7}$$

where E_i is short-hand notation representing the temporally dependent scalar field projection $E(\vec{\mathbf{r}}_i, t)$. The i^{th} primary edge has length L_i and the j^{th} dual edge has the length L_j. Areas A_i and A_j represent the polygons associated with mesh edges. The semi-discrete form of Maxwell's equations can be integrated in time with any method such as Forward Euler, Backward Euler, or the Crank-Nicolson method. The fully discrete Maxwell's equations are now written as

$$\begin{bmatrix} \mathscr{D}_{il}^{\mathscr{A}} & \alpha \mathscr{A}_{im} \\ \alpha \mathscr{F}_{jl} & \mathscr{D}_{jm}^{\mathscr{F}} \end{bmatrix} \begin{bmatrix} E_l^{n+1} \\ H_m^{n+1} \end{bmatrix} + \begin{bmatrix} -\mathscr{D}_{il}^{\mathscr{A}} & (1-\alpha)\mathscr{A}_{im} \\ (1-\alpha)\mathscr{F}_{jl} & -\mathscr{D}_{jm}^{\mathscr{F}} \end{bmatrix} \begin{bmatrix} E_l^{n} \\ H_m^{n} \end{bmatrix} + \begin{bmatrix} J_i^{N(\alpha)} A_i \\ 0 \end{bmatrix} = 0. \tag{3.8}$$

The superscript n represents the "known" field quantity at a previous time step and $n+1$ represents the next "unknown" field quantity in the time-marching method. The diagonal matrices are defined as $\mathscr{D}_{il}^{\mathscr{A}} = \varepsilon_i A_i \delta_{il}/\Delta t$ and $\mathscr{D}_{jm}^{\mathscr{F}} = \mu_0 A_j \delta_{jm}/\Delta t$, the Ampere matrix is defined as $\mathscr{A}_{im} = -c_{im} L_m$, and the Faraday matrix is defined as $\mathscr{F}_{jl} = c_{jl} L_l$. The tensors in the Ampere and Faraday matrices are collections of ± 1 to ensure a properly closed line integral around associated edges. For example, the Ampere constants are assigned to make primary edge vector projections follow the right hand rule and the Faraday constants are assigned to make dual edge vector projections follow the left hand rule. The constant α reflects the choice of the time integration method. Values $\alpha = 0, 1, 1/2$ will yield Forward Euler, Backward Euler, and Crank-Nicolson schemes, respectively. Current density time index $N(\alpha)$ is set to n for Forward Euler and $n+1$ for Backward Euler and Crank-Nicolson.

One important note about DVSI is that for simulations in which the geometry of interest is on the order of one or many wavelengths, this method would require a dense mesh to keep the spatially constant approximations valid. Because this paper is meant to highlight the numerical aspects of DVSI, dense discretizations are used to evaluate canonical CEM benchmarks which have analytic solutions. Simulating electronic devices does not retain this meshing requirement, since the majority of devices are inherently sub-wavelength in size. This is a tremendous meshing advantage of DVSI for semiconductor devices that are sub-wavelength in terms of obtaining full-wave solutions without over-meshing.

Also, since fluxes are defined as constant over polygons, an edge which does not physically touch its associated face is still valid as long as the area is computed correctly and the electric flux accounts for regions of different dielectric properties as described in the subsequent section. This is because the edge is closer to its associated dual face than any other edge due to the Delaunay nature of the mesh. This implies the need for strict adherence to Delaunay-compliant meshes. Validity of these mesh relationships also extends to node / volume relationships as well as boundary edges. Experience has shown that commercial tools claim to have $< 1\%$ non-Delaunay compliance in three dimensions. Given the previous statements on mesh requirements, even a small fraction of non-Delaunay compliant elements can lead to erroneous results with Box Integration or DVSI.

3.3.2 Dielectric Discontinuities

Dielectric discontinuities are easily addressed with DVSI. The area and dielectric constant of the i^{th} electric flux term is computed as a summation of the N_i local relative dielectric constants and areas of the electric flux, namely

$$\varepsilon_i A_i = \varepsilon_0 \sum_{k=0}^{N_i} \varepsilon_{r,k} A_k. \tag{3.9}$$

Sub-polygons are created within the main dual polygon where each sub-polygon resides completely within a single primary tetrahedron. An example of sub-polygon decomposition of the electric flux area is shown in Figure 3.3. The first three sub-polygons reside within one material, i.e., $\varepsilon_{r,1} = \varepsilon_{r,2} = \varepsilon_{r,3}$ and the other two are in another material with $\varepsilon_{r,4} = \varepsilon_{r,5}$. The total elec-

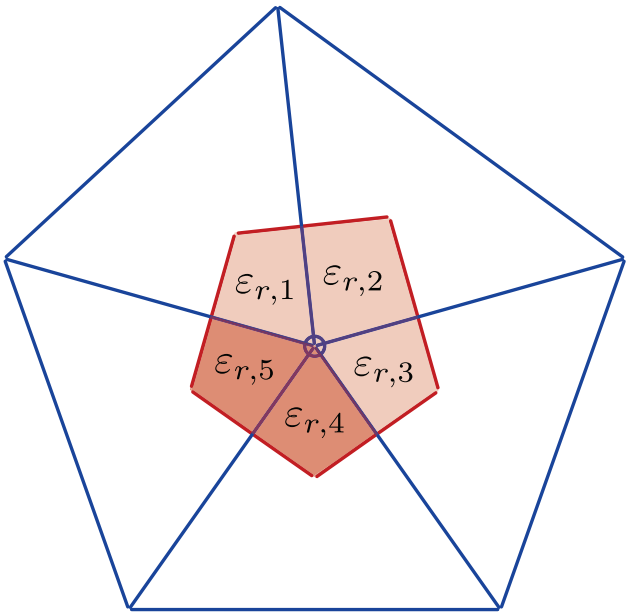


Figure 3.3: An illustration of the decomposition of a dual polygon which spans multiple dielectric regions. The dual polygon is split into multiple sub-polygons. Each is completely contained within a single primary tetrahedron and assigned the relative dielectric constant of that tetrahedron. The 3-D mesh was rotated such that the primary edge points into the page.

tric flux area is then calculated using Eqn. (3.9). Even though this example only illustrates five sub-polygons, the process is completely general to more complicated mesh elements.

3.3.3 Boundary Conditions

BCs are essential for any numerical simulation. In terms of DVSI, BCs are applied to primary and dual edges (as well as primary nodes in subsequent sections) which reside on the boundaries of the mesh. Each boundary dual edge connects the mid-point of its associated primary edge to the circumcenters of the boundary primary triangles associated with the same primary edge — see Figure 3.4 for an illustration. These relationships hold true for triangles which do not physically contain their circumcenters.

To properly analyze devices, BCs must be introduced to emulate sources and to truncate the computational domain. The former is accomplished by applying voltages across metal contacts in an electronic device. Before advancing to voltage BCs, examples are presented where EM fields will be excited by electric current densities to facilitate quantitative analysis of DVSI against

canonical examples.

3.3.3.1 Perfect Electrical Conductor

One important BC for CEM and electronic device simulations is the perfect electrical conductor (PEC). A DVSI implementation of PEC edges is quite straightforward. Rotational electric field edges which lie on PEC boundaries are always tangential to the surface. Because the physics states that $\hat{n} \times \vec{\mathbf{E}} = 0$ and $\hat{n} \cdot \vec{\mathbf{H}} = 0$, the electric field edges tangential to and magnetic field edges normal to PEC boundaries are simply removed from the DVSI solution set. Figure 3.4 illustrates the removal of DVSI edges at PEC interfaces.

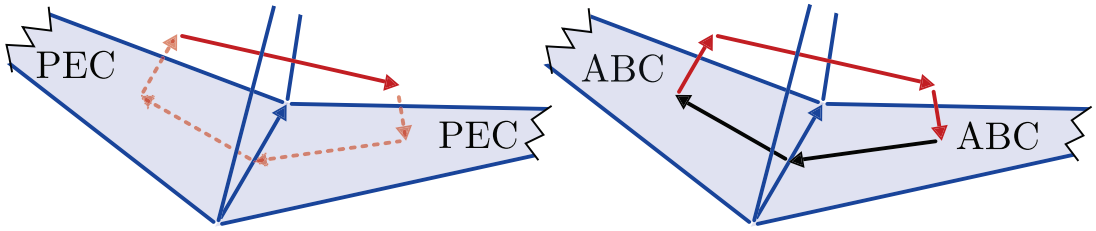


Figure 3.4: Illustrations of a primary edge on two different boundaries.

3.3.3.2 Impedance Boundary Condition

An impedance BC, otherwise known as a plane wave or first-order absorbing BC (ABC), exploits the relationship between TEM electric and magnetic fields or plane waves, given by

$$\vec{\mathbf{H}} = \frac{1}{Z}\hat{\mathbf{k}} \times \vec{\mathbf{E}}. \tag{3.10}$$

Here, $Z = \sqrt{\mu/\varepsilon}$ is the impedance of the matched exterior region. Utilizing this relationship between electric and magnetic fields requires knowledge of the wave vector. A simple resolution of this issue is assuming that the wave vector is in the direction of the boundary's normal $\hat{\mathbf{k}} \approx \hat{\mathbf{n}}$. This proves to be a convenient simplification as $\hat{\mathbf{n}} \times \hat{\mathbf{p}}_i$, where $\hat{\mathbf{p}}_i$ is the unit vector of the primary edge, is exactly the unit direction of in-plane dual magnetic field edge. The fully discrete DVSI

equations with the impedance BC are

$$\begin{bmatrix} \mathscr{D}^{\mathscr{A}}{}_{il} + \alpha \mathscr{I}_1 & \alpha \mathscr{A}_{im} \\ \alpha \mathscr{F}_{jl} & \mathscr{D}^{\mathscr{F}}{}_{jm} \end{bmatrix} \begin{bmatrix} E_l^{n+1} \\ H_m^{n+1} \end{bmatrix} + \begin{bmatrix} -\mathscr{D}^{\mathscr{A}}{}_{il} + \beta \mathscr{I}_1 & \beta \mathscr{A}_{im} \\ \beta \mathscr{F}_{jl} & -\mathscr{D}^{\mathscr{F}}{}_{jm} \end{bmatrix} \begin{bmatrix} E_l^n \\ H_m^n \end{bmatrix} + \begin{bmatrix} J_i^{N(\alpha)} A_i \\ 0 \end{bmatrix} = 0. \quad (3.11)$$

The impedance BC tensor takes the form $\mathscr{I}_1 = -\hat{\mathbf{d}}_f \cdot \left(\hat{\mathbf{n}}_f \times \hat{\mathbf{p}}_i\right) L_f \delta_{il} \delta_{lb} / Z$, where the in-plane dual edge with length L_f and unit vector $\hat{\mathbf{d}}_f$ is associated with the boundary primary edge with the unit vector $\hat{\mathbf{p}}_i$. There are always two boundary primary triangles (with different normals $\hat{\mathbf{n}}_f$) associated with a boundary primary edge. This implies that there are two in-plane dual edges associated with the i^{th} primary edge and is why a separate index f is applied to these edges. The first Kronecker delta ensures that the impedance boundary affects diagonal matrix elements and the second ensures that only primary electric field edges associated with the impedance boundary (with index b) are updated. The term $\beta = (1 - \alpha)$ is defined for brevity.

This BC serves two important functions in a device simulator. The first is a truncation to any open regions, i.e., the end of the physical device. The other is a model of an external load. This is a critical feature when dealing with devices that interface with other circuit components. In the device situation, impedance BCs are far more important than sophisticated methods of making perfectly matched layers or radiation BCs.

3.3.4 PEC Cavities with Ideal Dielectrics

To demonstrate the validity of this numerical framework, two PEC cavity results are presented and discussed. All numerical examples in this section solve the purely rotational fields of Maxwell's equations. To excite fields in the cavities, a current density is introduced on primary edges. These examples test all the aforementioned numerical aspects of the rotational solver except the impedance BC. This BC will be tested in Section 3.5.

The mathematical form of the current density used for these numerical examples is a modulated Gaussian, or

$$J_i(t) = \cos\left[\omega_0 (t - t_0)\right] \exp\left[-\frac{(t-t_0)^2}{2\sigma^2}\right]. \tag{3.12}$$

Here, $\omega_0 = 2\pi f_0$ is the center angular frequency of the Gaussian, $\sigma = 2.335/(2\pi BW)$ is Gaussian width associated with the bandwidth BW, and $t_0 = 6\sigma$ is the Gaussian delay. For each subsequent result, the time step was calculated with $\Delta t = 1/(S f_{\max})$, where $f_{\max}$ is the highest frequency in the simulation and S is a unit-less scale factor used to control the size of the time step. For each of the following benchmarks, the time step scale was chosen as $S = 400$.

The first result is a cubic PEC cavity filled 40% with air and 60% with bulk semiconductor ($\varepsilon_r = 5.9$, which represents bulk GaN). The dielectric discontinuity lies in the $z = 0.4$ m plane. Figure 3.5 (left) displays the computed resonant frequencies of the half filled cavity compared

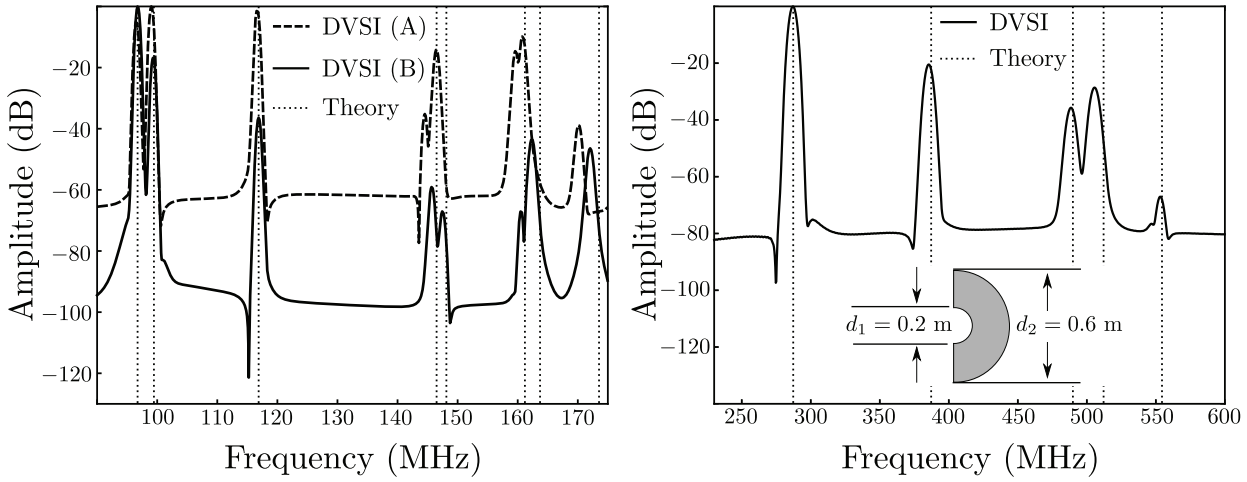


Figure 3.5: Results for two PEC cavities with ideal dielectrics.

with the analytic resonant frequencies [71] (dotted lines). Two different meshes were used to compute the DVSI results A and B. Mesh A is comprised of 6000 tetrahedra resulting in 7830 DOF and mesh B contains 20,250 tetrahedra resulting in 27,720 DOF. Local maxima of both mesh results were calculated to quantitatively assess the numerical method. Then, the error between the

computed resonant frequencies $f_{n,i}$ and the analytic frequencies $f_{a,i}$ was calculated with

$$\varepsilon = 100 \times \frac{\sqrt{\sum_i \left(f_{n,i} - f_{a,i}\right)^2}}{\sqrt{\sum_i f_{a,i}^2}}. \tag{3.13}$$

The global errors for the first eight resonant frequencies computed with meshes A and B were $\varepsilon = 1.36\%$ and $\varepsilon = 0.59\%$, respectively. This demonstrates the increasing accuracy of the higher resonant frequencies with mesh density, as expected.

To test the framework's ability to conform to curved features, the second result is a section of a cylindrical cavity illustrated in the inset of Figure 3.5 (right). The inner diameter is $d_1 = 0.2$ m and the outer diameter is $d_2 = 0.6$ m. The entire cavity is air filled. Figure 3.5 (right) shows the numerically computed resonant frequencies in comparison with analytic solutions [71]. The error between the analytic and the numerical resonant frequencies was again computed and global error for the first five frequencies of the cylindrical sector cavity was $\varepsilon = 0.68\%$. The fourth computed resonant frequency displayed the worst individual error of $\varepsilon = 1.25\%$. This is attributed to the mesh density in the cylindrical directions of the geometry, which was not uniform in the ρ, ϕ, and z directions.

3.4 Poisson's Equation

The next step in developing a DVSI full-wave EM framework is to discretize Poisson's equation. This section is solely devoted to solving irrotational problems where only the electric potential is the solution variable.

3.4.1 DVSI Discretization

The Delaunay-Voronoi mesh provides a convenient physical representation of Poisson's equation

$$-\iint_S \hat{n} \cdot \varepsilon \nabla \Phi\left(\vec{\mathbf{r}},t\right) dS = \iiint_V \rho\left(\vec{\mathbf{r}},t\right) dV. \tag{3.14}$$

As discussed in Section 3.2, a bounded domain is tessellated with two interweaving meshes. Each primary node, which is a point of the Delaunay mesh, is bounded by a dual polyhedron. An example is shown in Figure 3.6.

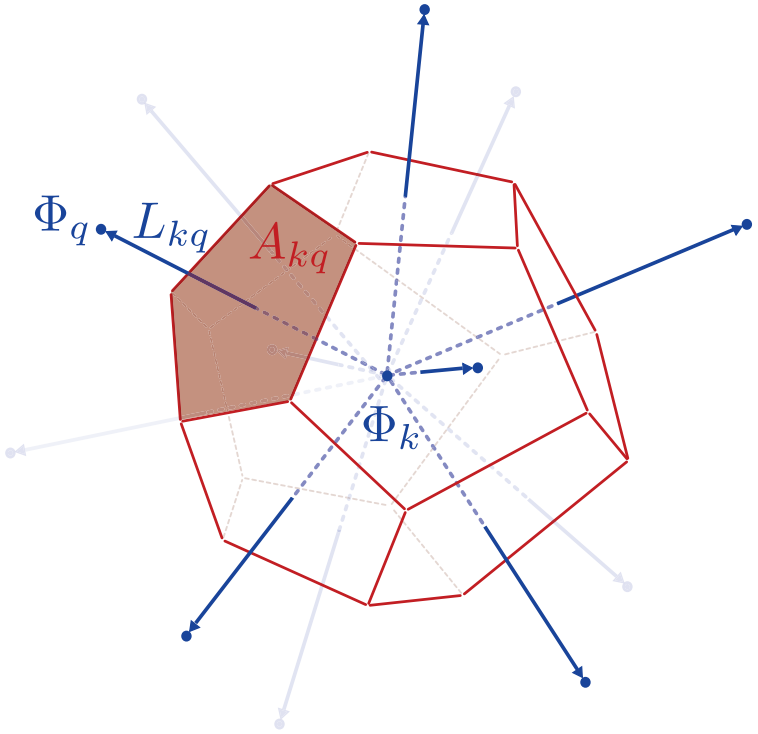


Figure 3.6: A visualization of the DVSI representation of Poisson's equation in three dimensions.

The key to discretizing Poisson's equation within DVSI is the fact that all primary edges connected to a primary node are always normal (or perpendicular) to the dual polygons of the bounding dual polyhedron. This is analogous to the divergence of the electric field out of the enclosing dual polyhedron. The gradient is discretized as a finite difference on the q^{th} primary edge associated with node k. Using these facts, Poisson's equation is written as

$$\mathscr{P}_{kp}\Phi_p^n \equiv \sum_q \varepsilon_{kq}\frac{A_{kq}}{L_{kq}}\left(\Phi_k^n - \Phi_q^n\right) = \rho_k^n V_k. \tag{3.15}$$

Here, the summation index q runs over all primary edges attached to primary node k. The same time notation is used as the previous section, i.e., $\Phi_k^n = \Phi\left(\vec{\mathbf{r}}_k, t_n\right)$ is the previous time step solution. Eqn. (3.15) states that electric scalar potentials are assigned to the nodes of a primary edge and the spatially constant electric field projection is defined as a finite difference of these potentials over

the primary edge with length L_{kq}. These potentials are driven by the charge density ρ_k defined inside the k^{th} dual polyhedron with volume V_k.

3.4.2 Boundary Conditions for Poisson's Equation

As with Maxwell's rotational equations, Poisson's equation cannot be solved without imposing BCs. Truncation of the computational domain is handled through Ampere's law and therefore only Dirichlet BCs will be discussed. This boundary will be used extensively to create a meaningful voltage source to drive electronic devices.

A Dirichlet BC specifies the values that the solution must take. In the case of Poisson's equation, these solutions are the electric scalar potentials. For all Dirichlet nodes listed with index b associated with the k^{th} primary node, Poisson's equation will take the form

$$\mathscr{P}_{kp}\Phi_p^n = \rho_k^n V_k + \sum_b \varepsilon_{kb}\frac{A_{kb}}{L_{kb}}\Phi_b^n. \tag{3.16}$$

3.4.3 Spherical Region with Dirichlet Surfaces

To demonstrate the validity of the DVSI Poisson solver while also displaying the usefulness of DVSI's unstructured grid, a spherical region with two Dirichlet surfaces at $r = 0.25$ m and $r = 2$ m is solved. A simple Gaussian charge model $\rho\left(\vec{\mathbf{r}}\right) = \exp\left(-\vec{\mathbf{r}}\cdot\vec{\mathbf{r}}\right)$ is applied in the spherical region. This allows the DVSI solver results to be compared to an analytic solution, derived by integrating the Laplacian in spherical coordinates with the Gaussian source term and a Dirichlet voltage of 0.25 V applied on both the inner and outer surfaces.

Figure 3.7 displays a slice of the DVSI mesh (left) and a cross section of the DVSI scalar potential solution (right). Mesh refinement near the center of the sphere was performed. The error was computed between the DVSI numerical solution and the analytic solution using Eqn. (3.13). Now, $f_{n,i}$ is replaced with $\Phi_{n,i}$, the numerical potential at node $\vec{\mathbf{r}}_i$, and $f_{a,i}$ is replaced with $\Phi_{a,i}$ which is the analytic solution at the same location. For the mesh shown in Figure 3.7 (left), the error

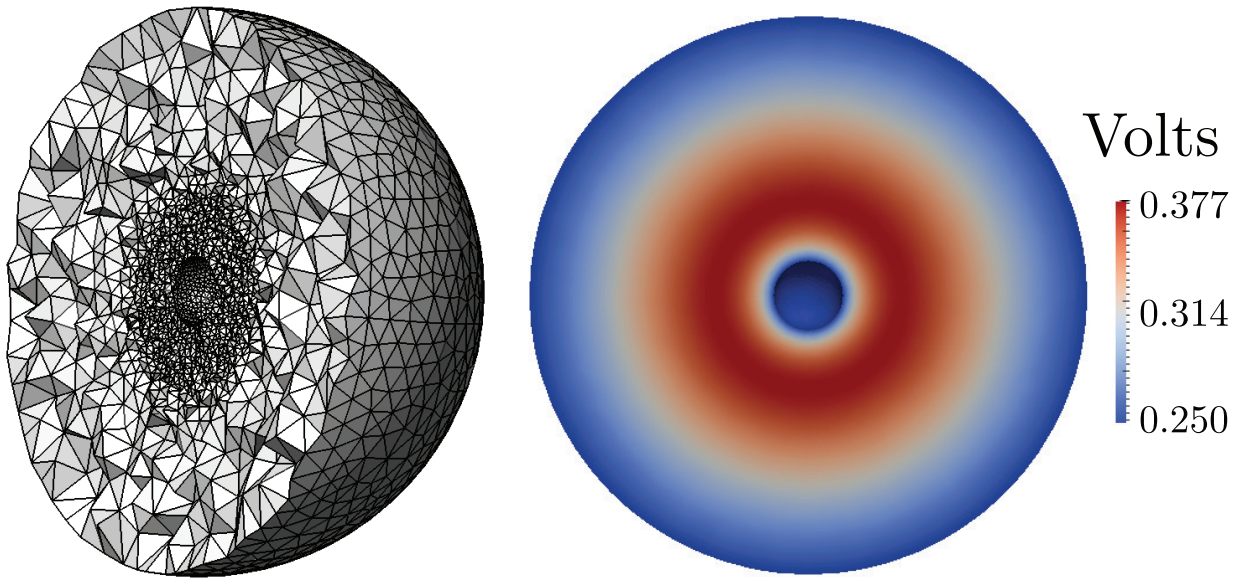


Figure 3.7: (left) A mesh of the spherical region with Dirichlet surfaces at $r = 0.25$ m and $r = 2$ m. (right) The numerical solution $\Phi(r)$ of Poisson's equation with a simple Gaussian charge density and a Dirichlet potential of 0.25 V applied at both surfaces.

was $\varepsilon = 2.5\%$, showing that the numerical solution, which resides on a completely unstructured mesh, agrees well with the analytic solution.

3.5 Complete Electromagnetics for Device Simulation

With the derivations and examples presented in Sections 3.3 and 3.4, a full-wave EM DVSI framework can now be presented. Here, the electric field, magnetic field, and the electric potential are solution variables that are solved simultaneously. Eqns. (3.1) – (3.3) completely describe EM phenomena in an electronic device.

3.5.1 DVSI Discretization

Proper simulation of electronic devices with full-wave effects requires solving the system of equations discussed in Sections 3.3 and 3.4 and shown in Eqns. (3.1) – (3.3). For the rest of this section, the rotational electric field will drop the "rot" subscripts and all electric field variables are assumed to be purely rotational. The fully discrete system of equations is

$$\begin{bmatrix} \mathscr{D}_{il}^{\mathscr{A}} & \alpha\mathscr{A}_{im} & \mathscr{P}_{ip}^{\mathscr{A}} \\ \alpha\mathscr{F}_{jl} & \mathscr{D}^{\mathscr{F}}{}_{jm} & 0 \\ 0 & 0 & \alpha\mathscr{P}_{kp} \end{bmatrix} \begin{bmatrix} E_l^{n+1} \\ H_m^{n+1} \\ \Phi_p^{n+1} \end{bmatrix}$$
$$+ \begin{bmatrix} -\mathscr{D}^{\mathscr{A}}{}_{il} & \beta\mathscr{A}_{im} & -\mathscr{P}^{\mathscr{A}}{}_{ip} \\ \beta\mathscr{F}_{jl} & -\mathscr{D}_{jm}^{\mathscr{F}} & 0 \\ 0 & 0 & \beta\mathscr{P}_{kp} \end{bmatrix} \begin{bmatrix} E_l^n \\ H_m^n \\ \Phi_p^n \end{bmatrix} + \begin{bmatrix} J_i^{N(\alpha)} A_i \\ 0 \\ -\rho_k^{N(\alpha)} V_k \end{bmatrix} = 0. \quad (3.17)$$

Eqn. (3.17) is a direct descendant of Eqns. (3.8) and (3.15). The additional tensor $\mathscr{P}_{ip}^{\mathscr{A}}$ couples the rotational electric and magnetic fields to the static electric potential. This tensor takes the form $\mathscr{P}_{ip}^{\mathscr{A}} \Phi_p^n \equiv \frac{\varepsilon_i A_i}{\Delta t L_i} \left(\Phi_{i,1}^n - \Phi_{i,2}^n \right)$, where the i^{th} primary edge points from node 1 with potential $\Phi_{i,1}$ to node 2 with potential $\Phi_{i,2}$.

3.5.2 Device Boundary Conditions

BCs for the complete EM equations will now be presented utilizing the information described in the previous two sections. Each BC previously described is now presented in the context of an electronic device simulation. Additional requirements needed for relevance to device simulation are also highlighted. A numerical example of a transmission line that requires these BCs for excitation and propagation is compared with analytic solutions.

3.5.2.1 Metal Contacts

Metal contacts are critically important BCs for electron devices and are both the regions where voltages (or currents) are applied to bias and drive the device as well as the regions that interact with external circuitry (e.g., driving loads). These contacts, in terms of a physical device, are certainly not perfect electrical conductors, but are sufficiently conducting to be treated as such in the majority of technologically relevant cases. Extensions to non-ideal contacts are an interesting research topic when the distributed effects of finite conductivity are important. For the purposes of

developing the numerical methods of DVSI and to allow comparison with analytic solutions, only the ideal cases when metals are PECs are considered.

The PEC BC was previously discussed in Section 3.3 and the only additional information needed is the definition of electric potentials. Because a PEC defines an equipotential surface, a Dirichlet BC is assigned to any potentials residing on this boundary. Two PEC surfaces can then be used to define a voltage, where one conductor is the reference potential which could be ground (or other reference potential offset).

3.5.2.2 Extrinsic Impedance Boundary Condition

The impedance BC is important for device simulation for several reasons. First, it can properly emulate resistive sheets which can be used to represent external loads. Second, the impedance BC can be used as a first-order domain truncation. An important point of the impedance BC is that it affects the total electric field. The discretized linear system including the impedance BC is

$$\begin{bmatrix} \mathscr{D}_{il}^{\mathscr{A}} + \alpha\mathscr{I}_1 & \alpha\mathscr{A}_{im} & \mathscr{P}_{ip} + \alpha\mathscr{I}_2 \\ \alpha\mathscr{F}_{jl} & \mathscr{D}_{jm}^{\mathscr{F}} & 0 \\ 0 & 0 & \alpha\mathscr{P}_{kp} \end{bmatrix} \begin{bmatrix} E_l^{n+1} \\ H_m^{n+1} \\ \Phi_p^{n+1} \end{bmatrix} + \begin{bmatrix} -\mathscr{D}_{il}^{\mathscr{A}} + \beta\mathscr{I}_1 & \beta\mathscr{A}_{im} & -\mathscr{P}_{ip} + \beta\mathscr{I}_2 \\ \beta\mathscr{F}_{jl} & -\mathscr{D}_{jm}^{\mathscr{F}} & 0 \\ 0 & 0 & \beta\mathscr{P}_{kp} \end{bmatrix} \begin{bmatrix} E_l^n \\ H_m^n \\ \Phi_p^n \end{bmatrix} + \begin{bmatrix} J_i^{N(\alpha)} A_i + \mathscr{I}_2 \left(\alpha\Phi_b^{n+1} - \beta\Phi_b^n \right) \\ 0 \\ -\rho_k^{N(\alpha)} V_k + \mathscr{P}_{kp}\delta_{pb} \left(\alpha\Phi_b^{n+1} - \beta\Phi_b^n \right) \end{bmatrix} = 0. \quad (3.18)$$

Again, the impedance BC tensor takes the form $\mathscr{I}_1 = -\hat{\mathbf{d}}_f \cdot \left(\hat{\mathbf{n}}_f \times \hat{\mathbf{p}}_i\right) L_f \delta_{il} \delta_{lb} / Z$ and the second impedance tensor is defined as $\mathscr{I}_2 = -\hat{\mathbf{d}}_f \cdot \left(\hat{\mathbf{n}}_f \times \hat{\mathbf{p}}_i\right) L_f \delta_{il} \delta_{lb} / \left(L_i Z\right)$, where L_i is the length of the i^{th} primary edge. The potentials in the source vector, with index b, account for any primary

nodes which lie on corners of the computational domain which connect Dirichlet boundaries and impedance boundaries directly.

3.5.2.3 Voltage Boundary Condition

One critical BC that may be difficult to implement without incorporating Poisson's equation is the voltage source, a staple of device engineering. This excitation boundary allows the application of voltages across different device contacts, such as the gate and drain of a field effect transistor (FET). It also allows a simulator to operate in conditions that replicate those in a physical device.

The voltage boundary can be any shape, but for simplicity it is assumed that the voltage boundary is a rectangular region connecting two PECs. The two PEC contacts are assigned constant potentials and the overlap between the contacts and the voltage excitation serves as Dirichlet line segments for the solution of Poisson's equation.

Coupling the voltage boundary into the rotational electric and magnetic fields is now considered. To do so, the voltage boundary is first declared as free of rotational electric fields. Dual magnetic fields in the voltage plane are then declared as unknowns and they are solved for using Ampere's law, Eqn. (3.1). Because dual magnetic field edges in the voltage plane correspond to exactly one rotational electric field edge, Ampere's law for the null rotational electric field edge is applied to solve for the in-plane dual magnetic field edge. This is mathematically written, with the Backward Euler time integration scheme, as

$$\mathscr{A}_{im} H_m^{n+1} + \mathscr{P}_{ip}^{\mathscr{A}} \left(\Phi_p^{n+1} - \Phi_p^n \right) + J_i^{n+1} A_i = 0. \tag{3.19}$$

In this version of Ampere's law for the voltage boundary, the line integrals of magnetic fields include the dual magnetic field edges in the voltage boundary plane. This is a subtle, but highly practical issue for device simulation.

With the use of the Poisson Dirichlet BC and the voltage boundary Ampere's law for the in-

plane magnetic fields, the discrete linear system with the voltage excitation is written as

$$\begin{bmatrix} \mathscr{D}_{il}^{\mathscr{A}} & \alpha\mathscr{A}_{im} & \mathscr{P}_{ip}^{\mathscr{A}} \\ \alpha\mathscr{F}_{jl} & \mathscr{D}_{jm}^{\mathscr{F}} & 0 \\ 0 & 0 & \alpha\mathscr{P}_{kp} \end{bmatrix} \begin{bmatrix} E_l^{n+1} \\ H_m^{n+1} \\ \Phi_p^{n+1} \end{bmatrix} + \begin{bmatrix} -\mathscr{D}_{il}^{\mathscr{A}} & \beta\mathscr{A}_{im} & -\mathscr{P}_{ip}^{\mathscr{A}} \\ \beta\mathscr{F}_{jl} & -\mathscr{D}_{jm}^{\mathscr{F}} & 0 \\ 0 & 0 & \beta\mathscr{P}_{kp} \end{bmatrix} \begin{bmatrix} E_l^n \\ H_m^n \\ \Phi_p^n \end{bmatrix} + \begin{bmatrix} J_i^{N(\alpha)} A_i + \mathscr{P}_{ib}^{\mathscr{A}} \left(\Phi_b^{n+1} - \Phi_b^n \right) \\ 0 \\ -\rho_k^{N(\alpha)} V_k + \mathscr{P}_{kb} \left(\alpha\Phi_b^{n+1} - \beta\Phi_b^n \right) \end{bmatrix} = 0. \tag{3.20}$$

The electric potentials added to the source vector reflect any Dirichlet voltages.

3.5.2.4 Two-Port Devices

In experiments and circuits, transistors are driven by time dependent voltages applied across contacts. For example, given a FET with source, gate, and drain metal contacts, the device operates with a voltage bias applied from the drain to the source and an oscillating voltage applied across the gate and the source. In terms of the complete DVSI equations, these form the voltage BCs.

The applied voltage $v_1(t)$ and the simulated current $i_1(t)$ are related to the port voltages in Figure 3.8 by $v_1(t) = v_1^+(t) + v_1^-(t)$ and $i_1(t) = \left[v_1^+(t) - v_1^-(t)\right]/Z_1$ [62]. It is assumed that the input and output ports are 50 Ω transmission lines. Given the two-port device network in Figure 3.8,

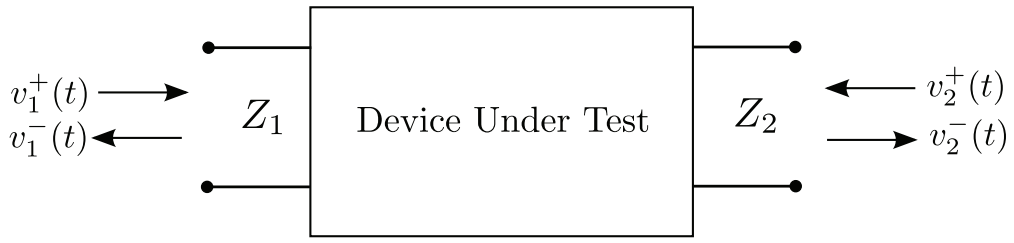


Figure 3.8: The two-port model of an electronic device. Input voltages are labeled with a subscript 1 and output voltages are labeled with a subscript 2. The input and output ports are ideal 50 Ω transmission lines.

inward and outward propagating waves are defined. These waves can then be used to characterize

the device with S-parameters, calculated with

$$v_n^{\pm}(t) = \frac{1}{2}\left[v_n(t) \pm Z_n i_n(t)\right], \tag{3.21}$$

$$S_{11}(\omega) = \frac{V_1^-(\omega)}{V_1^+(\omega)}, \tag{3.22}$$

$$S_{21}(\omega) = \frac{V_2^-(\omega)}{V_1^+(\omega)}. \tag{3.23}$$

Here, the spectral voltages $V_n^{\pm}(\omega)$ are the Fourier transforms of the transient voltages $v_n^{\pm}(t)$. The output current $i_2(t)$ is the transient current across the load Z_L. Customarily, the load Z_2 represents the characteristic impedance of a high frequency probe impedance matched to a network analyzer so that the voltage $v_2^+(t) = 0$. For demonstrative purposes, the load and output impedances are chosen to be purely resistive.

Before the transient voltages and currents are Fourier transformed, they must be computed in the simulation. Voltage drops are computed with a contiguous set of primary edges between metal contacts. An integral of the total electric field starting from the reference contact yields the proper voltage drop. Similarly, current is computed by integrating a contiguous, straight line of dual edges on which magnetic fields are defined. The line integral of the magnetic fields is parallel to the metal contacts and in the direction which corresponds to outward propagating power.

3.5.3 Lossless Transmission Line

To demonstrate the validity of the complete set of EM equations, a simple lossless transmission line is analyzed. This is a good candidate as it will require a transient voltage excitation, will test the impedance BC, and will be quantitatively validated via analytic S-parameters. The computational domain is bounded by two PEC plates located at the $z = 0$ and $z = 0.1$ mm planes, a voltage boundary at the $x = 0$ plane, a $Z_L = 50\ \Omega$ impedance boundary at the $x = 1$ mm plane, and two "open" boundaries at the $y = 0$ and $y = 0.1$ mm planes. These open boundaries declare in-plane magnetic fields to be null. The DVSI transmission line is excited with a Gaussian voltage drop between the PEC plates. The parameters of the Gaussian voltage $v(t) = V_0 \exp\left[-\frac{(t-t_0)^2}{2\sigma^2}\right]$ are

$V_0 = 0.5$ V, $\sigma = 2.5 \times 10^{-14}$ s, and $t_0 = 10\sigma$. The time step is chosen as $\Delta t = 0.1\sigma$ and the number of time steps is 100,000.

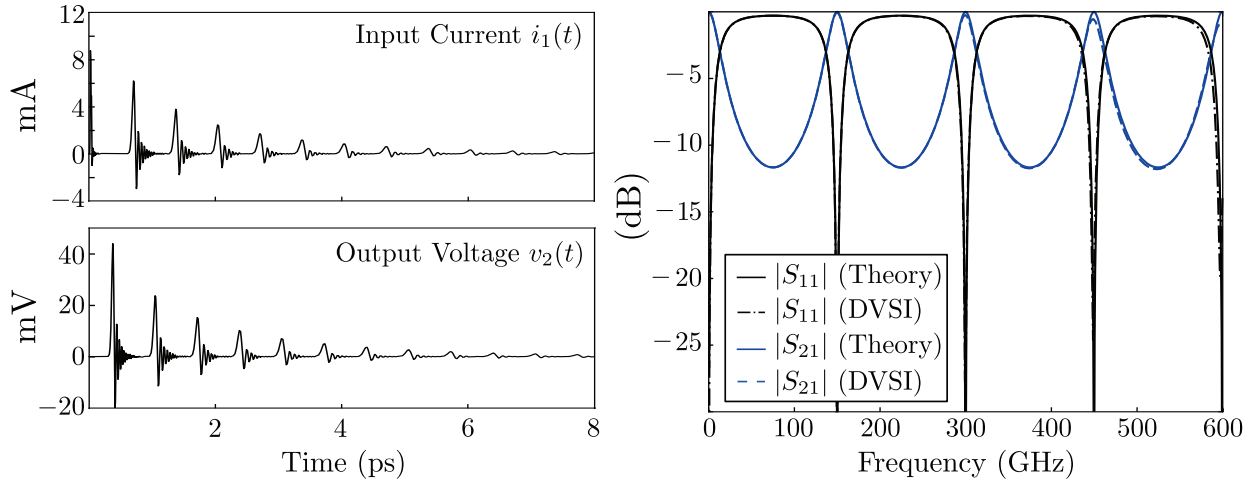


Figure 3.9: (left) The transient input current (top) and output voltage (bottom) calculated in the lossless transmission line. (right) The spectral reflection and transmission coefficients.

Figure 3.9 shows (left) the transient input current and output voltage and (right) the magnitude of the reflection and transmission coefficients $|S_{11}(\omega)|$ and $|S_{21}(\omega)|$, respectively. The solid lines are the analytic values [63]. Because this transmission line is lossless, the relation $|S_{11}(\omega)|^2 + |S_{21}(\omega)|^2 = 1$ is true for all frequencies. A good quantitative analysis of the full-wave transmission line simulation is the computed deviation of this relation from unity. Using the data presented in Figure 3.9 (right) and a modified version of Eqn. (3.13), the total deviation from unity was $\varepsilon = 3.7\%$. This result is a worst case scenario considering the fact that it was computed up to 600 GHz — twice the frequency corresponding to the electrical size of the transmission line. Electronic devices, which are the target application of this numerical framework, rarely operate at wavelengths below the electrical size of its gate width.

We now simulate the lossless transmission line on a series of meshes to estimate the DVSI method's order of convergence. Each mesh refinement decreases the edge length in the propagation direction. The number of tetrahedra range from 480 to 9600 and the total number of DOF ranges from 854 to 15,864. The DOF reflect the number of primary edges which define electric fields, the number of dual edges which define magnetic fields, and the number of primary nodes which define

electric potentials. The lossless transmission line operates with the same inputs as the previous example except with a modified Gaussian bandwidth of $\sigma = 10^{-12}$ s. The solution computed on each mesh was compared to the finest mesh solution to produce an error given by a modified version of Eqn. (3.13). These errors were then used to compute an error bound given by

$$|S_k - L| < CN_k^{-p}, \tag{3.24}$$

where S_k are the k solutions computed on each mesh, L is the solution computed on the finest mesh, N_k is the number of edges in the propagation direction for each mesh, and p is the order of convergence. The computed order of convergence was $p \approx 2.0$. As expected, this result is consistent with the order of convergence of the traditional Yee cell algorithm [70].

3.6 Conclusion

DVSI is an appropriate numerical framework for discretizing full-wave EM and coupling to charge transport. With the approximations outlined in this chapter, this numerical method is amenable to nonlinear charge transport on unstructured grids. The complete discretization of Maxwell's equations within DVSI was presented as well as the proper BCs for device simulation. Several benchmarks were presented to demonstrate the rotational and irrotational field descriptions independently and simultaneously. DVSI agreed well with analytic solutions for a number of numerical tests, including propagation in a transmission line using a voltage excitation. The importance of having a Delaunay-compliant mesh was emphasized as a requirement of the method in general. Although the fields are approximated to be constant vector projections on mesh edges, this is an appropriate level of discretization for electronic devices where the physical dimensions are substantially sub-wavelength. Finally, the framework allows for the solution of charge transport on the same mesh as the EM problem which is not typical of device solvers.

CHAPTER 4

FERMI KINETICS TRANSPORT

4.1 Introduction

Charge transport in semiconductors can be determined by the solution of the semi-classical Boltzmann transport equation (BTE) [72]

$$\frac{\partial f(\vec{\mathbf{k}},\vec{\mathbf{r}},t)}{\partial t}+\vec{\mathbf{v}}\cdot\nabla_r f(\vec{\mathbf{k}},\vec{\mathbf{r}},t)+\frac{\vec{\mathbf{F}}}{\hbar}\cdot\nabla_k f(\vec{\mathbf{k}},\vec{\mathbf{r}},t)=\left(\frac{\partial f(\vec{\mathbf{k}},\vec{\mathbf{r}},t)}{\partial t}\right)_{coll}. \tag{4.1}$$

The BTE is a seven-dimensional integro-differential equation and describes how a collection of particles responds to electromagnetic fields and scattering potentials. The distribution function $f(\vec{\mathbf{k}},\vec{\mathbf{r}},t)$ gives the occupation probability at the real-space location $\vec{\mathbf{r}}$ and reciprocal-space location $\vec{\mathbf{k}}$ at time t. It can be found by solving Eqn. (4.1). Clearly, approximations are required in order to yield tractable solutions.

A powerful statistical method for solving the BTE is the MC technique [73]. MC was first presented as a statistical approach to solving general integro-differential equations [74]. Since then, it has seen a wide variety of applications and is one of the preferred methods for accurate numerical simulation of semiconductor devices. Many varieties exist in the literature including the ensemble MC (EMC) method [75]. However, the EMC's stochastic algorithm is very computationally expensive. The cellular MC (CMC) implementation of the EMC method was developed to reduce the burdensome computational demands [76]. Despite scattering rate pre-computation and tabulation, CMC is still computationally intensive and can lead to long simulation times for large, complex devices. Combining MC charge transport with full-wave EM using, for example, the FDTD method [66], further increases the computational demands.

An alternative approach to solving the BTE is through deterministic methods. These rely on taking moments of the BTE. One of the first deterministic Boltzmann solvers for semiconductor device simulations was presented by Scharfetter and Gummel [77]. This work presented physi-

cal approximations to the phenomenological drift-diffusion (DD) model of charge transport. The physical size of the device technology was still relatively large and the numerical solution of the DD model produced satisfactory results. However, reduced device features produce larger electric fields and hot-electron effects, such as electron velocity saturation, rendering the DD approximation insufficiently accurate.

Two seminal works which extended the DD model to include carrier heating were presented by Stratton [78] and Blotekjaer [79]. These methods utilized different approximations of the BTE. Stratton proposed that the distribution function be split into even and odd components while Blotekjaer derived moments of the BTE without any assumptions on the form of the distribution function. Both methods required closure relations to generate a linearly independent system of equations [80]. A common closure relation is the prescription of electronic heat flow with Fourier's law [80]. Another common approximation used to specify momentum and energy in hydrodynamic models is parabolic band-structure [80]. However, this approximation will reduce the accuracy of the device solver when the electric field is large. To the best of the author's knowledge, no hydrodynamic or energy-transport model can include complete electronic band-structure. Some device solvers employ fitting parameters such as field- and temperature-dependent mobilities as an attempt to overcome the limitations of the parabolic band-structure [81].

FKT is a deterministic Boltzmann solver which has shown promising results in the literature. Historically, FKT was conceived by seeking an alternative method for calculating the electronic heat flow. Rather than use the combination of Fourier's law and an approximate thermal conductivity [80], it was proposed that a thermodynamic identity could be used as a closure relation for energy-transport models [82]. This closure relation provided a robust Boltzmann solver and it was later shown that FKT was completely amenable to electronic band-structure [83] and full-wave EM [62]. In particular, a version of FKT with GaAs band-structure, quantum scattering formalisms, and hot-electron effects was shown to reproduce the electron velocity saturation calculated by sophisticated MC methods [83] in a fraction of the computational time. FKT was also shown to accurately simulate a GaAs MESFET and GaN HEMT from DC up through mm-wave

frequencies without adjustable calibration parameters [83, 84].

In the rest of this chapter, a detailed discussion of the FKT device simulation framework is presented. In Section 4.1.1, a detailed derivation of the governing equations of charge transport is presented. Energy- and real-space discretization techniques are presented in Sections 4.2 and 4.3, respectively. Finally, Section 4.4 concludes this chapter.

4.1.1 Moments of the BTE

Following the method of Stratton [78], the collision term of the BTE is replaced with a momentum relaxation time approximation. The distribution function is split into even (equilibrium) $f_0(\vec{\mathbf{k}},\vec{\mathbf{r}},t)$ and odd $f_1(\vec{\mathbf{k}},\vec{\mathbf{r}},t)$ functions in $\vec{\mathbf{k}}$-space. This yields

$$\frac{\partial f(\vec{\mathbf{k}},\vec{\mathbf{r}},t)}{\partial t}+\vec{\mathbf{v}}\cdot\nabla_r f(\vec{\mathbf{k}},\vec{\mathbf{r}},t)-q\frac{\vec{\mathbf{E}}}{\hbar}\cdot\nabla_k f(\vec{\mathbf{k}},\vec{\mathbf{r}},t)=\left.\frac{\partial f_0(\vec{\mathbf{k}},\vec{\mathbf{r}},t)}{\partial t}\right|_{coll}-\frac{f_1(\vec{\mathbf{k}},\vec{\mathbf{r}},t)}{\tau_k}. \tag{4.2}$$

Here, the real-space external force $\vec{\mathbf{F}}$ is only due to an electric field, i.e., $\vec{\mathbf{F}}=-q\vec{\mathbf{E}}$. Eqn. (4.2) is solved with the method of moments. Electron density and electron kinetic energy density are defined as

$$n=\frac{1}{4\pi^3}\int d\vec{\mathbf{k}} f_0(\vec{\mathbf{k}}), \tag{4.3}$$

$$E_n=\frac{1}{4\pi^3}\int d\vec{\mathbf{k}} E(\vec{\mathbf{k}}) f_0(\vec{\mathbf{k}}), \tag{4.4}$$

and the electron flux density and electron kinetic energy flux density are defined as

$$\vec{\mathbf{J}}_n=\frac{1}{4\pi^3}\int d\vec{\mathbf{k}}\vec{\mathbf{v}} f_1(\vec{\mathbf{k}}), \tag{4.5}$$

$$\vec{\mathbf{K}}_n=\frac{1}{4\pi^3}\int d\vec{\mathbf{k}} E(\vec{\mathbf{k}})\vec{\mathbf{v}} f_1(\vec{\mathbf{k}}). \tag{4.6}$$

The symmetric (equilibrium) distribution function is the well-known Fermi-Dirac distribution

$$f_0(\vec{\mathbf{k}})=\frac{1}{\exp\left(\frac{E(\vec{\mathbf{k}})-F}{k_B T}\right)+1}, \tag{4.7}$$

which represents an electron gas with Fermi level F and temperature T. The term k_B is Boltzmann's constant. The non-equilibrium distribution function will be approximated with piece-wise

equilibrium distribution functions. Moments of the BTE will provide the non-equilibrium nature of the electron gases.

Solving the BTE with the method of moments amounts to integrating Eqn. (4.2) over all reciprocal-space. Using a general $\vec{\mathbf{k}}$-space dependent operator $O(\vec{\mathbf{k}})$, the BTE can be separated into symmetric (even) and antisymmetric (odd) equations, or

$$\frac{1}{4\pi^3}\int_{-\infty}^{\infty} d\vec{\mathbf{k}}O(\vec{\mathbf{k}})\left[\frac{\partial f_0}{\partial t}+\vec{\mathbf{v}}\cdot\nabla_r f_1 - q\frac{\vec{\mathbf{E}}}{\hbar}\cdot\nabla_k f_1 - \left.\frac{\partial f_0}{\partial t}\right|_{coll}\right]=0, \tag{4.8}$$

$$\frac{1}{4\pi^3}\int_{-\infty}^{\infty} d\vec{\mathbf{k}}O(\vec{\mathbf{k}})\left[\frac{f_1}{\tau_k}+\frac{\partial f_1}{\partial t}+\vec{\mathbf{v}}\cdot\nabla_r f_0 - q\frac{\vec{\mathbf{E}}}{\hbar}\cdot\nabla_k f_0\right]=0. \tag{4.9}$$

The multiplication of an even function with an odd function results in an odd function. Furthermore, the derivatives of even and odd functions (with respect to the dimension in which it is even or odd) result in odd and even functions, respectively. When an odd operator is used for a moment of the BTE, the LHS of Eqn. (4.8) is exactly zero. The LHS of Eqn. (4.9) is exactly zero when an even operator is used for a moment of the BTE. Taking moments of the BTE with the operators $\{1,\vec{\mathbf{v}},E(\vec{\mathbf{k}}),E(\vec{\mathbf{k}})\vec{\mathbf{v}}\}$ yields the electron continuity equation, particle flux density, energy conservation equation, and the kinetic energy flux density

$$\frac{\partial n}{\partial t}+\nabla\cdot\vec{\mathbf{J}}_n+C^n=0 \tag{4.10}$$

$$\vec{\mathbf{J}}_n=\frac{1}{4\pi^3}\int d\vec{\mathbf{k}}\vec{\mathbf{v}}\tau_k\left(\frac{q}{\hbar}\vec{\mathbf{E}}\cdot\nabla_k f_0-\vec{\mathbf{v}}\cdot\nabla_r f_0\right), \tag{4.11}$$

$$\frac{\partial E_n}{\partial t}+\nabla\cdot\vec{\mathbf{S}}_n+q\vec{\mathbf{E}}\cdot\vec{\mathbf{J}}_n+C^E=0, \tag{4.12}$$

$$\vec{\mathbf{K}}_n=\frac{1}{4\pi^3}\int d\vec{\mathbf{k}}E(\vec{\mathbf{k}})\vec{\mathbf{v}}\tau_k\left(\frac{q}{\hbar}\vec{\mathbf{E}}\cdot\nabla_k f_0-\vec{\mathbf{v}}\cdot\nabla_r f_0\right). \tag{4.13}$$

Appendix A.1 gives an in-depth derivation of Eqns. (4.10) – (4.13). The terms $\tau_k\frac{\partial\vec{\mathbf{J}}_n}{\partial t}$ and $\tau_k\frac{\partial\vec{\mathbf{K}}_n}{\partial t}$ can be ignored for devices operating up to the mm-wavelength regime as $\tau_k \ll \partial t$. For sub-mm-wavelength devices, i.e., terahertz devices, these terms must be considered.

4.1.2 An Alternative Treatment of Heat Flow

Eqns. (4.10) – (4.13), along with the definitions (4.3) – (4.7), are not a closed system of equations. The third moment of the BTE yields the kinetic energy flux density $\vec{\mathbf{K}}_n$. But, the total energy flux density $\vec{\mathbf{S}}_n = \vec{\mathbf{K}}_n + \vec{\mathbf{H}}_n$ is a combination of the kinetic energy flux density and the heat flux density or heat flow. Therefore, a prescription of the heat flow is required.

A common approximation of the electronic heat flow used in many other deterministic Boltzmann solvers is Fourier's law [80]

$$\vec{\mathbf{H}}_n = -\kappa(T)\nabla T, \tag{4.14}$$

where T is the temperature of the distribution of electrons and $\kappa(T)$ is a fictitious electronic thermal conductivity. The thermal conductivity is typically approximated by the Wiedemann-Franz law [80]

$$\kappa(T) = \left(\frac{5}{2} - p\right)\left(\frac{k_B}{q}\right)^2 q\mu nT. \tag{4.15}$$

Here, p is a tune-able correction factor.

In [80], it is reported that numerical issues in the device solver could be attributed to Fourier's law. It also reports that incorporating higher moments of the BTE may help alleviate some numerical issues. This thesis will focus on a different closure relation. In the seminal paper by Grupen [82], an alternative treatment of heat flow is presented. An analogy of the thermodynamic identity

$$k_B T\Delta\sigma = \Delta E_n - F\Delta n, \tag{4.16}$$

is used to calculate heat flow between Fermi distributions. This treatment of heat flow produces an extremely stable numerical framework. In subsequent sections which describe the spatial discretization of FKT, the second law of thermodynamics will be used to calculate the discrete heat flux density. It will be shown that the thermodynamic heat flow is easily obtained from the discrete particle and kinetic energy flux densities.

4.1.3 Self Heating Effects

Transistors used in PA circuits are heavily biased. Self heating effects are therefore critical for accurately simulating the devices. The final governing equation in the global FKT system is the lattice energy conservation equation [84]

$$\rho C_p \frac{\partial T_L}{\partial t} + \nabla \cdot \kappa \nabla T_L - C^E = 0, \tag{4.17}$$

where T_L is the temperature of the lattice and ρ, C_p, κ are the mass density, specific heat, and lattice thermal conductivity of the semiconductor crystal, respectively. The energy collision operator C^E is the source term of the lattice heating equation.

4.2 FKT Energy-Space Discretization

The first step in solving Eqns. (4.10) – (4.13) is the discretization of reciprocal-space. To avoid discretizing three-dimensions, reciprocal-space is converted to energy-space. Section 4.2.1 provides these details. The calculation of isosurface integrals is then presented in Section 4.2.2. Finally, the approximation and calculation of the piece-wise energy windows are presented in Sections 4.2.3 and 4.2.4. These steps are pre-computations in terms of the FKT device simulator. The resulting data corresponding to specific bulk materials are imported into the device simulator.

4.2.1 Converting Reciprocal-Space Integration to Energy-Space Integration

The densities, fluxes, and collision operators were presented in Section 4.1.1. Re-writing these equations with band-structure dependencies yields

$$n = \frac{1}{4\pi^3} \int d\vec{\mathbf{k}} f_0(E(\vec{\mathbf{k}})), \tag{4.18}$$

$$E_n = \frac{1}{4\pi^3} \int d\vec{\mathbf{k}} E(\vec{\mathbf{k}}) f_0(E(\vec{\mathbf{k}})), \tag{4.19}$$

$$\vec{\mathbf{J}}_n = \frac{1}{4\pi^3} \int d\vec{\mathbf{k}} \vec{\mathbf{v}} \tau_k \left(\frac{q}{\hbar} \vec{\mathbf{E}} \cdot \nabla_k f_0(E(\vec{\mathbf{k}})) - \vec{\mathbf{v}} \cdot \nabla_r f_0(E(\vec{\mathbf{k}})) \right), \tag{4.20}$$

$$\vec{\mathbf{K}}_n = \frac{1}{4\pi^3} \int d\vec{\mathbf{k}} E(\vec{\mathbf{k}}) \vec{\mathbf{v}} \tau_k \left(\frac{q}{\hbar} \vec{\mathbf{E}} \cdot \nabla_k f_0(E(\vec{\mathbf{k}})) - \vec{\mathbf{v}} \cdot \nabla_r f_0(E(\vec{\mathbf{k}})) \right). \tag{4.21}$$

$$C^n = \frac{1}{\left(4\pi^3\right)^2} \int_{\vec{\mathbf{k}}_i} \int_{\vec{\mathbf{k}}_f} d\vec{\mathbf{k}}' d\vec{\mathbf{k}} W_{\vec{\mathbf{k}},\vec{\mathbf{k}}'} \Big\{ \left(n_q+1\right) f_{0,i}(E(\vec{\mathbf{k}})) \left(1 - f_{f,0}(E(\vec{\mathbf{k}}'))\right) - n_q f_{f,0}(E(\vec{\mathbf{k}}')) \left(1 - f_{i,0}(E(\vec{\mathbf{k}}))\right) \Big\}, \tag{4.22}$$

$$C^E = \frac{1}{\left(4\pi^3\right)^2} \int_{\vec{\mathbf{k}}_i} \int_{\vec{\mathbf{k}}_f} d\vec{\mathbf{k}}' d\vec{\mathbf{k}} W_{\vec{\mathbf{k}},\vec{\mathbf{k}}'} E(\vec{\mathbf{k}}) \Big\{ \left(n_q+1\right) f_{0,i}(E(\vec{\mathbf{k}})) \left(1 - f_{f,0}(E(\vec{\mathbf{k}}'))\right) - n_q f_{f,0}(E(\vec{\mathbf{k}}')) \left(1 - f_{i,0}(E(\vec{\mathbf{k}}))\right) \Big\}. \tag{4.23}$$

Here, $W_{\vec{\mathbf{k}},\vec{\mathbf{k}}'}$ is a scattering probability between two $\vec{\mathbf{k}}$-space states computed from Fermi's golden rule and $n_q = 1/\left(\exp\left(\hbar\omega/k_B T_L\right) - 1\right)$ is the phonon occupation number with energy $\hbar\omega$ [84]. The variable T_L is the lattice temperature. The initial electron state has an occupation probability $f_{i,0}$ and the final state has an occupation probability $f_{f,0}$. Using the sifting property of the delta function, applying the relation $\nabla_k f_0(E) = \frac{df_0}{dE}\nabla_k E = \frac{df_0}{dE}\hbar\vec{\mathbf{v}}$, and using the delta function composition property, the densities, fluxes, and collision operators are

$$n = \int D(E) f_0(E) dE, \tag{4.24}$$

$$E_n = \int_E D(E) f_0(E) E dE, \tag{4.25}$$

$$\vec{\mathbf{J}}_n = \int_E \boldsymbol{\mu}(E) \cdot \left(q\vec{\mathbf{E}} \frac{df_0(E)}{dE} - \nabla_r f_0(E) \right) dE, \tag{4.26}$$

$$\vec{\mathbf{K}}_n = \int_E \boldsymbol{\mu}(E) \cdot \left(q\vec{\mathbf{E}} \frac{df_0(E)}{dE} - \nabla_r f_0(E) \right) E dE, \tag{4.27}$$

$$C^n = \int_E G(E) \left\{ \left(n_q+1\right) f_{0,i}(E) \left(1 - f_{f,0}(E-\hbar\omega)\right) - n_q f_{f,0}(E-\hbar\omega) \left(1 - f_{i,0}(E)\right) \right\} dE, \tag{4.28}$$

$$C^E = \int_E G(E) \left\{ \left(n_q+1\right) f_{0,i}(E) \left(1 - f_{f,0}(E-\hbar\omega)\right) - n_q f_{f,0}(E-\hbar\omega) \left(1 - f_{i,0}(E)\right) \right\} E dE, \tag{4.29}$$

where

$$D(E) = \frac{1}{4\pi^3} \sum_i \int d\vec{\mathbf{S}}_i \left. \frac{1}{|\nabla_k E(\vec{\mathbf{k}})|} \right|_{\vec{\mathbf{k}}=\vec{\mathbf{k}}_i}, \tag{4.30}$$

$$\boldsymbol{\mu}(E) = \frac{1}{4\pi^3} \sum_i \int d\vec{\mathbf{S}}_i \left. \frac{\vec{\mathbf{v}}(\vec{\mathbf{k}})\vec{\mathbf{v}}^T(\vec{\mathbf{k}})\tau_k(\vec{\mathbf{k}})}{|\nabla_k E(\vec{\mathbf{k}})|} \right|_{\vec{\mathbf{k}}=\vec{\mathbf{k}}_i}. \tag{4.31}$$

$$G(E) = \int_{\vec{\mathbf{k}}_i} \int_{\vec{\mathbf{k}}_f} d\vec{\mathbf{S}}' d\vec{\mathbf{S}} W_{\vec{\mathbf{k}},\vec{\mathbf{k}}'} \left. \frac{1}{|\nabla_k E(\vec{\mathbf{k}})|} \right|_{\vec{\mathbf{k}}=\vec{\mathbf{k}}_i} \left. \frac{1}{|\nabla_k E(\vec{\mathbf{k}})|} \right|_{\vec{\mathbf{k}}=\vec{\mathbf{k}}_f}, \tag{4.32}$$

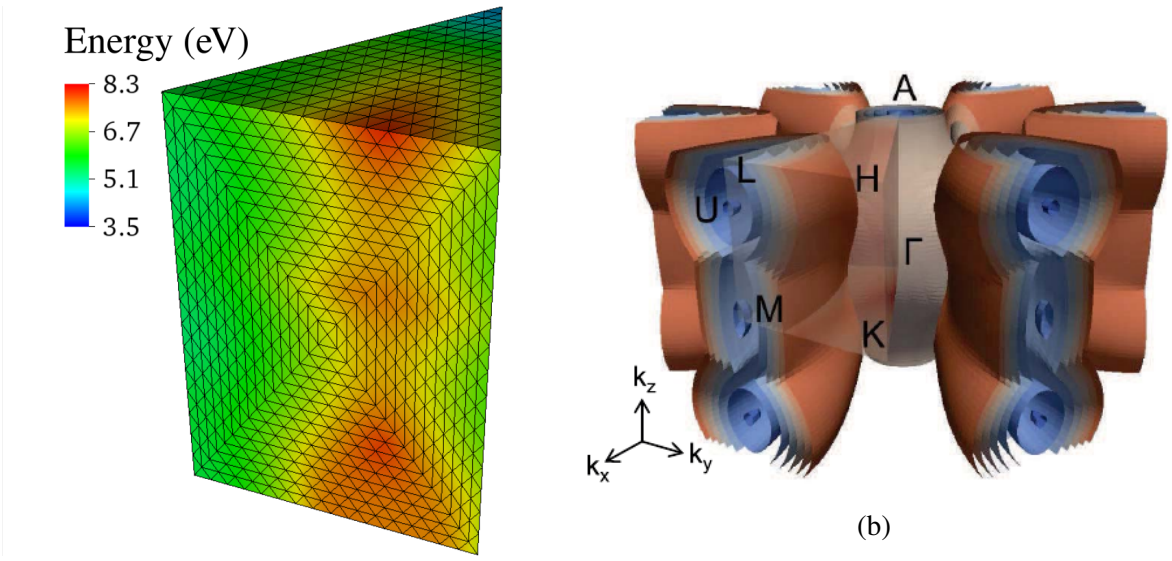
Here, $\boldsymbol{\mu}(E)$ is a tensor. Appendix A.2 provides an in-depth derivation of the reciprocal-space to energy-space conversion.

4.2.2 Isosurface Integrals

Eqns. (4.30) – (4.32) require integration over surfaces of constant electron eigen-energy, i.e., calculation of isosurface integrals. The calculation of bulk GaAs isosurface integrals for the FKT device simulator was first presented in [83]. It was later applied to bulk GaN in [84]. This work will present the process of determining surfaces of constant energy and the resulting isosurface integrations following the seminal work of Grupen.

Reciprocal space must be discretized to determine the surfaces of constant electron energy and calculate the isosurface integrals. To this end, a mesh of the irreducible edge of the Brillouin zone (IBZ) is generated and the band-structure is calculated on the nodes of the mesh. A k-space mesh of the wurtzite IBZ is presented in Figure 4.1a. Also illustrated is the electron energy calculated with the empirical pseudopotential method (EPM) [85]. Surfaces of constant energy are illustrated in Figure 4.1b [84]. An outline of the wurtzite IBZ is included with the energy isosurfaces. Details for calculating the isosurfaces of electronic band-structure are provided in [83].

The isosurfaces are divided into zones according to valleys of the GaN band-structure. An example of the first Γ-valley zone is presented in the panel diagram in Figure 4.2 (left). This zone is shaded in red. The DOS isosurface integral calculation, Eqn. (4.30), is also presented in Figure 4.2 (right) for the same zone. DOS data was obtained from Matt Grupen AFRL [84]. This



(a) (b)

Figure 4.1: (a) A mesh of the GaN IBZ and corresponding electron energy calculated with EPM and (b) the resulting surfaces of constant energy.

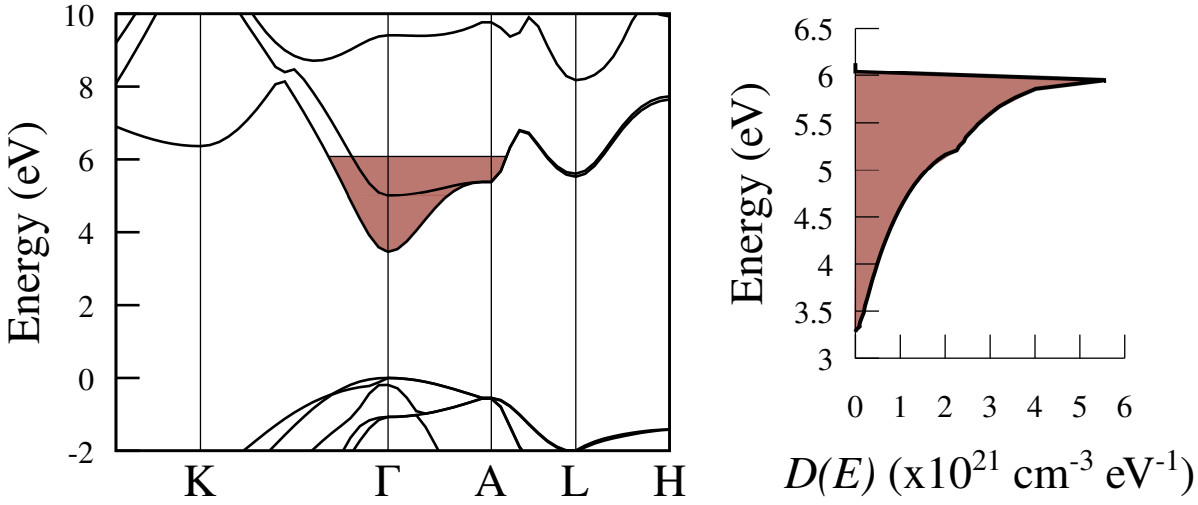


Figure 4.2: The (left) panel diagram of wurtzite GaN and (right) DOS calculated in the first Γ-valley. The EPM is used to calculate the electron energy. DOS data was obtained from Matt Grupen AFRL.

process is repeated for the other band-structure valleys. Figure 4.3 presents the DOS integrations for the first four valleys of wurtzite GaN. The data was obtained from Matt Grupen AFRL [84].

As described in [62] and [84], the isosurface integrals depend on FKT device simulator solution variables. Isosurface integration data must therefore be sampled over a range of FKT solution

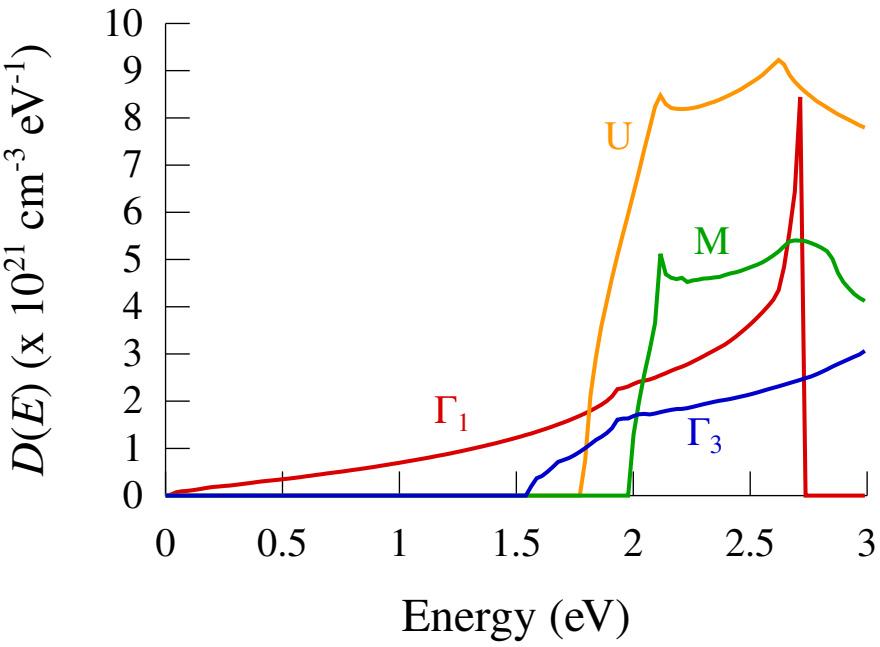


Figure 4.3: The DOS isosurface integrations calculated in the first four wurtzite GaN valleys. The DOS data was obtained from Matt Grupen AFRL.

variables. This data is interpolated in the device simulator to calculate the quantities required for residual and Jacobian construction. An example of the FKT solution variable dependence of the $\mu(E)$ isosurface integral is presented in Figure 4.4. Three integrations are presented with lattice temperatures of 300 K, 400 K, and 500 K. This data was obtained from Matt Grupen AFRL [84]. All isosurface integrals which depend on FKT solution variables are sampled in a similar manner.

4.2.3 Piece-Wise Energy Power Laws

The isosurface integration data of the previous section can be incorporated into the FKT device simulator in various ways. This work will focus on the method of Grupen [83]. Isosurface integration data is approximated with piece-wise energy power laws. Energy power laws allow the use of fast numerical Fermi-Dirac integral routines in the device simulator.

With the piece-wise energy power law approximation of the isosurface integrals, the densities,

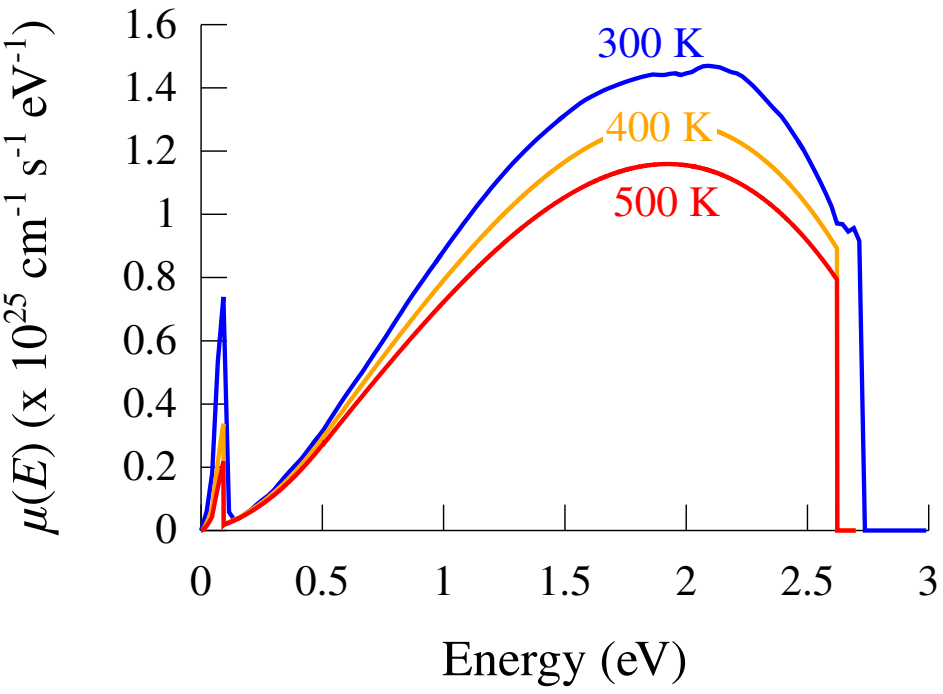


Figure 4.4: The $\mu(E)$ isosurface integrations calculated with different lattice temperature. The $\mu(E)$ data was obtained from Matt Grupen AFRL.

fluxes, and collision operators are

$$n \approx \sum_k \int_{E_{a,k}}^{E_{b,k}} A_k \left(E - E_{\rho,k}\right)^{\alpha_k} f_0(E) dE \equiv \sum_k n_k, \tag{4.33}$$

$$E_n \approx \sum_k \left\{ E_{\rho,k} n_k + \int_{E_{a,k}}^{E_{b,k}} A_k \left(E - E_{\rho,k}\right)^{\alpha_k+1} f_0(E) dE \right\}, \tag{4.34}$$

$$\vec{\mathbf{J}}_n \approx \sum_k \int_{E_{c,k}}^{E_{d,k}} D_k \left(E - E_{J,k}\right)^{\beta_k} \left(q\vec{\mathbf{E}} \frac{df_0(E)}{dE} - \nabla_r f_0(E) \right) dE \equiv \sum_k \vec{\mathbf{J}}_{n,k}, \tag{4.35}$$

$$\vec{\mathbf{K}}_n \approx \sum_k \left\{ E_{J,k} \vec{\mathbf{J}}_{n,k} + \int_{E_{c,k}}^{E_{d,k}} D_k \left(E - E_{J,k}\right)^{\beta_k+1} \left(q\vec{\mathbf{E}} \frac{df_0(E)}{dE} - \nabla_r f_0(E) \right) dE \right\}, \tag{4.36}$$

$$C^n \approx \sum_k \int_{E_{g,k}}^{E_{h,k}} G_k (E - E_{s,k})^{\gamma_k} \left\{ \left(n_q + 1 \right) f_{0,i}(E) \left(1 - f_{f,0}(E - \hbar\omega) \right) \right.$$

$$\left. - \, n_q f_{f,0}(E - \hbar\omega) \left(1 - f_{i,0}(E) \right) \right\} dE \equiv \sum_k C_k^n, \tag{4.37}$$

$$C^E \approx \sum_k \left\{ E_{s,k} C_k^n + \int_{E_{g,k}}^{E_{h,k}} G_k (E - E_{s,k})^{\gamma_k + 1} \left\{ \left(n_q + 1 \right) f_{0,i}(E) \left(1 - f_{f,0}(E - \hbar\omega) \right) \right. \right.$$

$$\left. \left. - \, n_q f_{f,0}(E - \hbar\omega) \left(1 - f_{i,0}(E) \right) \right\} dE \right\}. \tag{4.38}$$

Here, the flux isosurface integral matrix is $\boldsymbol{\mu}(E) \approx \mu(E)\mathbf{I}$ with $\vec{\mathbf{v}}(\vec{\mathbf{k}})\vec{\mathbf{v}}^T(\vec{\mathbf{k}}) \approx v(\vec{\mathbf{k}})^2\mathbf{I}$. The matrix $\mathbf{I}$ is the identity matrix. Figure 4.5 illustrates an approximation of the first Γ-valley DOS data. The

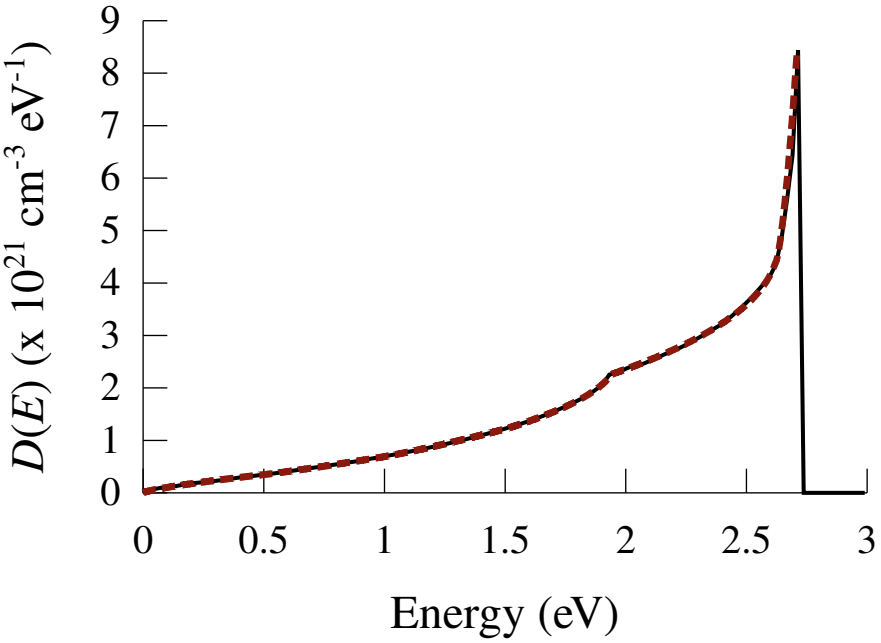


Figure 4.5: Approximation of the first Γ-valley DOS data with piece-wise energy power laws. The solid black line is the isosurface integration data and the dashed red line is the energy power law fit. Both data sets were obtained from Matt Grupen AFRL.

DOS integration data is plotted with a solid black line and the energy power law fits are plotted with dashed red lines. Figure 4.6 illustrates the approximation of the first Γ-valley $\mu(E)$ data.

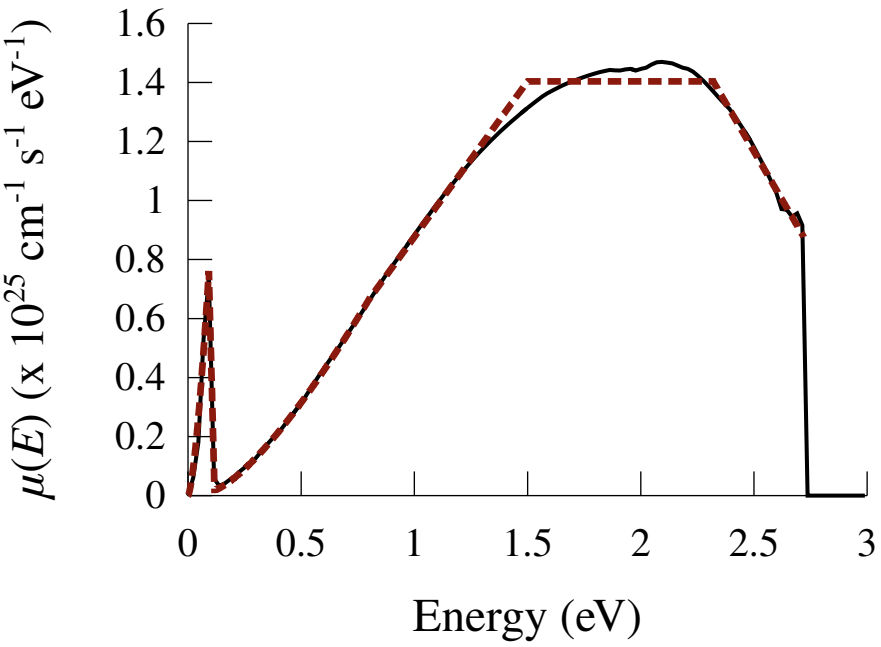


Figure 4.6: Approximation of the first Γ-valley $\mu(E)$ data with piece-wise energy power laws. The solid black line is the isosurface integration data and the dashed red lines are the energy power law fits. Both data sets were obtained from Matt Grupen AFRL.

Here, the integration data is plotted with a solid black line and the energy power law fits are plotted with dashed red lines. Because the $\mu(E)$ isosurface integral depends on solution variables of the FKT device simulator, energy power laws are required for each data set corresponding to samples of the solution variables.

4.2.4 Incomplete Fermi Integrals

The final component of the energy-space discretization is the evaluation of the energy integral over each piece-wise segment. The electron density and particle flux density in a single energy segment are

$$n_k = \int_{E_{a,k}}^{E_{b,k}} A_k \left(E - E_{\rho,k}\right)^{\alpha_k} f_0(E) dE, \tag{4.39}$$

$$\vec{\mathbf{J}}_{n,k} = \int_{E_{c,k}}^{E_{d,k}} D_k \left(E - E_{J,k}\right)^{\beta_k} \left(q\vec{\mathbf{E}}\frac{df_0(E)}{dE} - \nabla_r f_0(E)\right) dE, \tag{4.40}$$

Eqns. (4.39) and (4.40) are incomplete Fermi integrals of order α_k and β_k. With the substitutions

$$\frac{df_0(E)}{dE} = -\frac{df_0(E)}{dF}, \tag{4.41}$$

$$\varepsilon = \frac{E - E_c}{k_B T}, \tag{4.42}$$

$$\eta_{\rho,k} = \frac{F - E_c - (E_{\rho,k} - E_c)}{k_B T}, \tag{4.43}$$

$$\eta_{J,k} = \frac{F - E_c - (E_{J,k} - E_c)}{k_B T}, \tag{4.44}$$

$$a_k = \frac{E_{a,k} - E_{\rho,k}}{k_B T}, \tag{4.45}$$

$$b_k = \frac{E_{b,k} - E_{\rho,k}}{k_B T}, \tag{4.46}$$

$$c_k = \frac{E_{c,k} - E_{J,k}}{k_B T}, \tag{4.47}$$

$$d_k = \frac{E_{d,k} - E_{J,k}}{k_B T}, \tag{4.48}$$

the piece-wise densities and fluxes become

$$n_k = A_k (k_B T)^{\alpha_k + 1} F_{\alpha_k}\left(\eta_{\rho,k}, a_k, b_k\right), \tag{4.49}$$

$$E_{n,k} = E_{\rho,k} n_k + A_k (k_B T)^{\alpha_k + 2} F_{\alpha_k + 1}\left(\eta_{\rho,k}, a_k, b_k\right), \tag{4.50}$$

$$\vec{\mathbf{J}}_{n,k} = -qD_k \left(\vec{\mathbf{E}} (k_B T)^{\beta_k} F'_{\beta_k}\left(\eta_{J,k}, c_k, d_k\right) + \nabla_r \left[\frac{(k_B T)^{\beta_k + 1}}{q} F_{\beta_k}\left(\eta_{J,k}, c_k, d_k\right) \right] \right), \tag{4.51}$$

$$\begin{aligned} \vec{\mathbf{K}}_{n,k} = E_{J,k} \vec{\mathbf{J}}_{n,k} - qD_k \Bigg(& \vec{\mathbf{E}} (k_B T)^{\beta_k + 1} F'_{\beta_k + 1}\left(\eta_{J,k}, c_k, d_k\right) \\ & + \nabla_r \left[\frac{(k_B T)^{\beta_k + 2}}{q} F_{\beta_k + 1}\left(\eta_{J,k}, c_k, d_k\right) \right] \Bigg). \end{aligned} \tag{4.52}$$

Here, the incomplete Fermi integral of order α and its derivative are

$$F_\alpha(\eta, a, b) = \int_a^b \frac{\varepsilon^\alpha}{\exp(\varepsilon - \eta) + 1} d\varepsilon, \tag{4.53}$$

and

$$F'_{\alpha}(\eta,a,b)=\frac{dF_{\alpha}}{d\eta}=\alpha F_{\alpha-1}(\eta,a,b)+\frac{a^{\alpha}}{\exp(\varepsilon-\eta)+1}-\frac{b^{\alpha}}{\exp(\varepsilon-\eta)+1}. \tag{4.54}$$

The series expansion method [86] is preferred for the fast numerical evaluation of the incomplete Fermi integral of arbitrary order and parameter.

Because of the multiplication of the initial and final occupation probabilities, $f_{i,0}$ and $f_{f,0}$, the collision operators, Eqns. (4.37) and (4.38), cannot be represented as Fermi integrals. Therefore, the numerical routines for fast Fermi integral evaluation are not applicable for calculating the collision operators. Numerical quadrature of the piece-wise energy windows is a suitable method for evaluating the collision operator. The calculations of the collision operators are

$$\begin{aligned} C^{n} \approx \sum_{k}\sum_{i} G_{k}(E_i-E_{s,k})^{\gamma_k}\left\{\left(n_q+1\right)f_{0,i}(E_i)\left(1-f_{f,0}(E_i-\hbar\omega)\right)\right. \\ \left.-\,n_q f_{f,0}(E_i-\hbar\omega)\left(1-f_{i,0}(E_i)\right)\right\}w_i\Delta E_i, \end{aligned} \tag{4.55}$$

$$\begin{aligned} C^{E} \approx \sum_{k}\left\{E_{s,k}C^{n}_{k}+\sum_{i} G_{k}(E_i-E_{s,k})^{\gamma_k+1}\left\{\left(n_q+1\right)f_{0,i}(E_i)\left(1-f_{f,0}(E_i-\hbar\omega)\right)\right.\right. \\ \left.\left.-\;n_q f_{f,0}(E_i-\hbar\omega)\left(1-f_{i,0}(E_i)\right)\right\}w_i\Delta E_i\right\}, \end{aligned} \tag{4.56}$$

where w_i is the weight of the quadrature rule. Specific quadrature implementations include fast trapezoidal rules and Gaussian quadrature rules.

4.3 FKT Real-Space Discretization

After the discretization of energy-space, the next step in solving Eqns. (4.10) – (4.13) is discretizing real-space. The following provides the discretization details following the common practices of other deterministic Boltzmann solvers [87]. In particular, the box integration method [30] is used for the continuity equations and the Scharfetter-Gummel (SG) discretization technique [77] is used for the particle fluxes.

4.3.1 Continuity Equations

Discretizing the continuity equations requires integration over the semiconductor device domain. The integral form of the continuity equations on the DV mesh are

$$\iiint_V \left[\frac{\partial n}{\partial t} + C^n\right] dV + \iint_S \hat{n}\cdot\vec{\mathbf{J}}_n dS = 0, \tag{4.57}$$

$$\iiint_V \left[\frac{\partial E_n}{\partial t} + q\vec{\mathbf{E}}\cdot\vec{\mathbf{J}}_n + C^E\right] dV + \iint_S \hat{n}\cdot\vec{\mathbf{S}}_n dS = 0, \tag{4.58}$$

and

$$\iiint_V \left[\rho C_p \frac{\partial T_L}{\partial t} - C^E\right] dV + \iint_S \hat{n}\cdot\kappa\nabla T_L dS = 0. \tag{4.59}$$

Here, $\vec{\mathbf{E}} = \vec{\mathbf{E}}_{\text{rot}} - \nabla\Phi$ is the total electric field defined on a primary edge of the DV mesh. Using the box integration method [30] and backward Euler time discretization, the continuity equations become

$$\left[\frac{n_i^{t+1} - n_i^t}{\Delta t} + C_i^{n,t+1}\right] V_i + \sum_j J_{n,ij}^{t+1} A_{ij} = 0, \tag{4.60}$$

$$\left[\frac{E_{n,i}^{t+1} - E_{n,i}^t}{\Delta t} + C_i^{E,t+1}\right] V_i + q\sum_j \left(\mathrm{E}_j^{t+1} L_j\right) J_{n,ij}^{t+1} A_{ij} + \sum_j S_{n,ij}^{t+1} A_{ij} = 0. \tag{4.61}$$

$$\left[\rho_i C_{p,i} \frac{T_{L,i}^{t+1} - T_{L,i}^t}{\Delta t} - \sum_i C_i^{E,t+1}\right] V_i + \sum_j \left(T_{L,i}^{t+1} - T_{L,j}^{t+1}\right) \frac{\kappa_{ij} A_{ij}}{L_{ij}} = 0. \tag{4.62}$$

These transport equations are coupled in space and time with the discrete full-wave EM equations derived in Chapter 3.

4.3.2 Particle and Energy Fluxes

In Section 4.2.4, the particle flux density and kinetic energy flux density equations were discretized in terms of energy-space. Eqns. (4.51) and (4.52) are written without the arguments of the incomplete Fermi integrals for brevity. The fluxes in the k^{th} piece-wise energy window are

$$\vec{\mathbf{J}}_{n,k} = -qD_k \left(\vec{\mathbf{E}} (k_B T)^{\beta_k} F'_{\beta_k} + \nabla_r \left[\frac{(k_B T)^{\beta_k+1}}{q} F_{\beta_k}\right]\right), \tag{4.63}$$

and

$$\vec{\mathbf{K}}_{n,k} = E_{J,k}\vec{\mathbf{J}}_{n,k} - qD_k\left(\vec{\mathbf{E}}\,(k_BT)^{\beta_k+1}F'_{\beta_k+1} + \nabla_r\left[\frac{(k_BT)^{\beta_k+2}}{q}F_{\beta_k+1}\right]\right). \tag{4.64}$$

The kinetic energy flux density differs from the particle flux density by an order of energy. Therefore, the derivation of the real-space particle flux density discretization is directly amenable to the kinetic energy flux density.

4.3.2.1 The Original SG Discretization

The method of Scharfetter and Gummel [77] is utilized to discretize the fluxes. The original SG discretization technique was applied to the phenomenological DD equation. Given the projection of the DD equation onto a primary edge of the DV mesh

$$J_n(r) = -\mu_n\left(n(r)\mathrm{E}(r) + \frac{kT}{q}\frac{dn(r)}{dr}\right), \tag{4.65}$$

where r is the spatial dimension along the edge, the particle flux $J_n(r)$ and the electric field $\mathrm{E}(r)$ are approximated as spatially constant on the edge [77]. The mobility μ_n and electron temperature T are assumed to be spatially constant within a specific material region. With these approximations, Eqn. (4.65) is a linear first-order differential equation in terms of $n(r)$. Its solution on the i^{th} primary edge, with the BCs $n(r=r_0)=n_0$ and $n(r=r_1)=n_1$, is

$$J_{n,i} = \frac{\mu_n k_B T}{qL}\left[B(\xi)n_0 - B(-\xi)n_1\right], \tag{4.66}$$

where

$$\xi = \frac{q\mathrm{E}}{k_BT}L. \tag{4.67}$$

Appendix A.3 provides the complete derivation of the original SG discretization of the phenomenological DD model.

4.3.2.2 SG Discretization of the FKT Particle Flux

The SG discretization of the phenomenological DD model approximated the flux and electric field as spatially constant on a primary edge and solved a linear first-order differential equation with

density BCs at the nodes of the primary edge. The FKT particle flux requires approximations to its spatially dependent coefficients in order to use the same process. The projection of the k^{th} piece-wise FKT particle flux on the j^{th} primary edge is

$$J_{n,jk}(r) = -qD_{jk}\left(\mathrm{E}_j(r)\left(k_BT(r)\right)^{\beta_{jk}}F'_{\beta_{jk}}(r) + \frac{d}{dr}\left[\frac{\left(k_BT(r)\right)^{\beta_{jk}+1}}{q}F_{\beta_{jk}}(r)\right]\right), \tag{4.68}$$

where r represents the spatial coordinates along the edge from r_0 to r_1. The index k is associated to the primary edge with index j for all subsequent particle fluxes. With the assumption that the particle flux and electric field projections are spatially constant on the primary edge, i.e., $J_{n,jk}(r) \approx J_{n,jk}$ and $\mathrm{E}_j(r) \approx \mathrm{E}_j$, there are two possible discrete particle fluxes (out of many ways to discretize the equation) resulting from SG discretization. They are the generalized Einstein relation form

$$J^E_{n,jk} = \frac{qD_{jk}}{L_j}\left(\mathrm{Ein}_{n,jk}\right)_{\text{ave}}\left[B(\xi_{n,E,jk})N_{E,0,jk} - B(-\xi_{n,E,jk})N_{E,1,jk}\right], \tag{4.69}$$

with the density

$$N_{E,m,jk} = \left(k_BT_m\right)^{\beta_{jk}}F'_{\beta_{jk}}, \tag{4.70}$$

and Bernoulli function argument

$$\xi_{n,E,jk} = \frac{1}{\left(\mathrm{Ein}_{n,jk}\right)_{\text{ave}}}\left[\mathrm{E}_j + \Delta\mathrm{Ein}_{n,jk}\right]L_j, \tag{4.71}$$

and the inverse generalized Einstein relation form

$$J^I_{n,jk} = \frac{qD_{jk}}{L_j}\left(\frac{k_BT}{q}\right)_{\text{ave}}\left[B(\xi_{n,I,jk})N_{I,0,jk} - B(-\xi_{n,I,jk})N_{I,1,jk}\right], \tag{4.72}$$

with the density

$$N_{I,m,jk} = \left(k_BT_m\right)^{\beta_{jk}}F_{\beta_{jk}}, \tag{4.73}$$

and the Bernoulli function argument

$$\xi_{n,I,jk} = \frac{1}{\left(\frac{k_BT}{q}\right)_{\text{ave}}}\left[\left(\mathrm{Ein}^{-1}_{n,jk}\right)_{\text{ave}}\mathrm{E}_j + \Delta\left(\frac{k_BT}{q}\right)\right]L_j. \tag{4.74}$$

Appendix A.3.2 gives a complete derivation of both SG discretizations of the particle flux density and the definitions of the generalized Einstein relation form, Eqns. (34) – (36), and the inverse generalized Einstein relation form, Eqns. (43) – (46).

The SG discretization of the particle flux density is directly applicable to the kinetic energy flux density. The generalized Einstein relation form of the discrete kinetic energy flux density is

$$K^E_{n,jk} = \frac{qD_{jk}}{L_j}\left(\mathrm{Ein}_{E,jk}\right)_{\mathrm{ave}}\left[B(\xi_{E,E,jk})E_{E,0,jk} - B(-\xi_{E,E,jk})E_{E,1,jk}\right] + E_{J,jk}J^E_{n,jk}, \tag{4.75}$$

with the density

$$E_{E,m,jk} = (k_B T_m)^{\beta_{jk}+1} F'_{\beta_{jk}+1}, \tag{4.76}$$

and Bernoulli function argument

$$\xi_{E,E,jk} = \frac{1}{\left(\mathrm{Ein}_{E,jk}\right)_{\mathrm{ave}}}\left[\mathrm{E}_j + \Delta\mathrm{Ein}_{E,jk}\right]L_j, \tag{4.77}$$

and the inverse generalized Einstein relation form is

$$K^I_{n,jk} = \frac{qD_{jk}}{L_j}\left(\frac{k_B T}{q}\right)_{\mathrm{ave}}\left[B(\xi_{E,I,jk})E_{I,0,jk} - B(-\xi_{E,I,jk})E_{I,1,jk}\right] + E_{J,jk}J^I_{n,jk}, \tag{4.78}$$

with the density

$$E_{I,m,jk} = (k_B T_m)^{\beta_{jk}+1} F_{\beta_{jk}+1}(r), \tag{4.79}$$

and the Bernoulli function argument

$$\xi_{E,I,jk} = \frac{1}{\left(\frac{k_B T}{q}\right)_{\mathrm{ave}}}\left[\left(\mathrm{Ein}^{-1}_{E,jk}\right)_{\mathrm{ave}}\mathrm{E}_j + \Delta\left(\frac{k_B T}{q}\right)\right]L_j. \tag{4.80}$$

The generalized Einstein relation $\mathrm{Ein}_{E,jk}$ and the inverse generalized Einstein relation $\mathrm{Ein}^{-1}_{E,jk}$ of the kinetic energy flux density differ from their particle flux counterparts by an order β_{jk}. For example, the kinetic energy inverse is $\mathrm{Ein}^{-1}_{E,jk}(r) = \frac{F'_{\beta_{jk}+1}(r)}{F_{\beta_{jk}+1}(r)}$.

4.3.3 Heat Flow

The heat of a thermodynamic process is determined by the second law of thermodynamics. The second law states that the entropy of a thermodynamic system always increases, or

$$k_B T\Delta\sigma = \Delta E_n - F\Delta n = E_n(T_2) - E_n(T_1) - F\left[n(T_2) - n(T_1)\right]. \tag{4.81}$$

Here, ΔE_n is the internal energy of the system and $F\Delta n$ is the free energy [88]. The quantity $k_B T \Delta\sigma$ is the heat of the thermodynamic process. With $T_2 > T_1$, the heat of this thermodynamic process is positive.

Heat flow is required for closing the moments of the BTE. Therefore, an analogue of Eqn. (4.81) is needed. The internal and free energy rates are exactly the kinetic energy flux and the particle flux multiplied by the Fermi level, respectively. Physically, the discretized particle and kinetic energy flux densities represent the rates of particle and energy exchange between two non-equilibrium Fermi distributions connected by a DV primary edge. The fluxes are

$$J_{n,jk} = J_{0,jk}(T_{j,0}, T_{j,1}) - J_{1,jk}(T_{j,0}, T_{j,1}), \tag{4.82}$$

and

$$K_{n,jk} = K_{0,jk}(T_{j,0}, T_{j,1}) - K_{1,jk}(T_{j,0}, T_{j,1}), \tag{4.83}$$

where $J_{0,jk}(T_{j,0}, T_{j,1})$ represents particles flowing from distribution 0 with relative Fermi level $(F - E_c)_0$ and temperature $k_B T_0$ to distribution 1 with relative Fermi level $(F - E_c)_1$ and temperature $k_B T_1$. $J_{1,jk}(T_{j,0}, T_{j,1})$ represents the opposite flow of particles from distribution 1 to distribution 0. Figure 4.7 presents an illustration of the particle and kinetic energy exchange. Both terms of the fluxes depend on the Fermi distribution temperature at each node. This is due to the approximations made to the coefficients of the particle fluxes.

With the definitions (4.82) and (4.83), the heat flow is a direct analogue of Eqn. (4.81). The heat flowing from distribution 0 to 1 across the DV primary edge is

$$\begin{aligned} H_{n,jk} = {} & K_{0,jk}(T_{j,0}, T_{j,1}) - K_{0,jk}(T_{j,1}, T_{j,1}) \\ & - (F - E_c)_0 \left[J_{0,jk}(T_{j,0}, T_{j,1}) - J_{0,jk}(T_{j,1}, T_{j,1}) \right] \\ & - K_{1,jk}(T_{j,0}, T_{j,1}) + K_{1,jk}(T_{j,0}, T_{j,0}) \\ & + (F - E_c)_1 \left[J_{1,jk}(T_{j,0}, T_{j,1}) - J_{1,jk}(T_{j,0}, T_{j,0}) \right]. \end{aligned} \tag{4.84}$$

The first four terms of the heat flow represent the internal and free energy rates of distribution 0 and the last four represent the internal and free energy of distribution 1. The signs of each term are

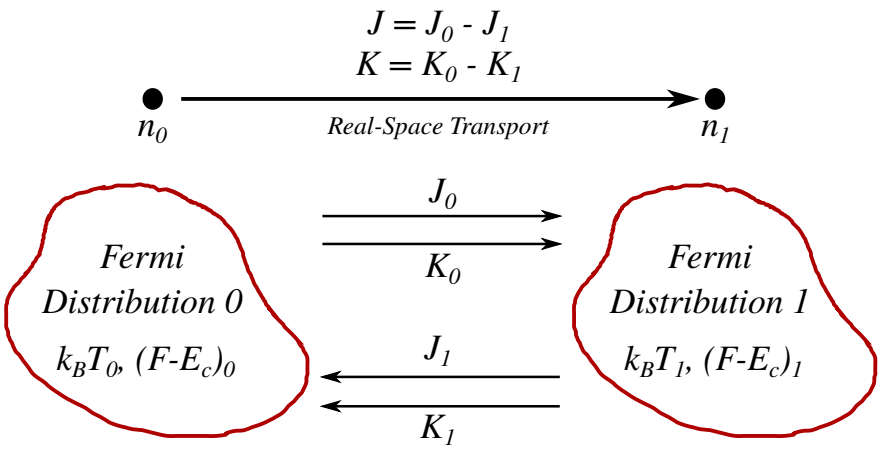


Figure 4.7: The particle and kinetic energy flux from node 0 with Fermi distribution 0 to node 1 with Fermi distribution 1 connected by a DV primary edge.

determined by the definition that heat flows from distribution 0 to distribution 1 (or from primary node 0 to primary node 1). If $T_0 > T_1$, then Eqn. (4.84) is a positive scalar quantity on the primary edge. If $T_1 > T_0$, then Eqn. (4.84) is a negative scalar quantity on the primary edge. Each case results in heat flow from higher temperature to lower temperature which is a consequence of the second law of thermodynamics.

4.4 Conclusion

FKT is a deterministic Boltzmann solver which enforces a thermodynamic heat flow as its closure relation. A thorough derivation of the FKT equations and their discretization in energy- and real-space was presented in this chapter. Section 4.2 provided the energy-space discretization techniques including the calculation of isosurface integrals, approximating the resulting integrations with piece-wise energy power laws, and converting the energy integrals into incomplete Fermi-Dirac integrals. Real space discretization techniques were presented in Section 4.3. The box integration method was used to approximate divergence operators on the DV mesh and the SG method was applied to discretize the FKT particle and energy flux densities. Finally, a detailed description of the heat flow algorithm was reported in Section 4.3.3.

CHAPTER 5

NUMERICAL SOLUTION OF THE FKT EQUATIONS

5.1 Introduction

A detailed derivation of the FKT device simulation equations and their discretization in energy- and real-space was presented in Chapter 4. This chapter focuses on solving the system of nonlinear FKT equations. In what follows, a detailed description of the nonlinear solver, the BCs used for device simulations, and important characteristics of the system of nonlinear equations are presented. The mesh convergence and stability of the FKT equations are investigated. Finally, advanced simulation techniques including higher-order discretization methods are presented.

5.2 The System of Nonlinear Equations

The governing equations of the FKT device simulator are Poisson's equation, Ampere's and Faraday's equations, electron continuity, energy conservation, and lattice heating. After real- and energy-space discretization, the residuals for n-type unipolar devices are

$$\mathrm{Pois}_i \equiv \sum_j \left(\Phi_i^{t+1} - \Phi_j^{t+1}\right) \frac{\varepsilon_{ij} A_{ij}}{L_{ij}} - q \sum_m \left(N_{D,i,m}^{+} - n_{i,m}^{t+1}\right) V_{i,m} = 0, \tag{5.1}$$

$$\mathrm{Amp}_i \equiv \left[\mathrm{E}_{\mathrm{rot},i}^{t+1} + \frac{\Phi_{i,0}^{t+1} - \Phi_{i,1}^{t+1}}{L_i} - \mathrm{E}_{\mathrm{rot},i}^{t} - \frac{\Phi_{i,0}^{t} - \Phi_{i,1}^{t}}{L_i}\right] \frac{\varepsilon_i A_i}{\Delta t} - q \sum_m J_{n,i,m}^{t+1} A_{i,m} - \sum_j H_j^{t+1} L_j = 0, \tag{5.2}$$

$$\mathrm{Far}_i \equiv \frac{H_i^{t+1} - H_i^{t}}{\Delta t} \mu_0 A_i + \sum_j \mathrm{E}_{\mathrm{rot},j} L_j = 0, \tag{5.3}$$

$$\mathrm{ElCont}_i \equiv \left[\frac{n_i^{t+1} - n_i^{t}}{\Delta t} + C_i^{n,t+1}\right] V_i + \sum_j J_{n,ij}^{t+1} A_{ij} = 0, \tag{5.4}$$

$$\mathrm{EnCons}_i \equiv \left[\frac{E_{n,i}^{t+1} - E_{n,i}^{t}}{\Delta t} + C_i^{E,t+1}\right] V_i + q \sum_j \left(\mathrm{E}_j^{t+1} L_j\right) J_{n,ij}^{t+1} A_{ij} + \sum_j S_{n,ij}^{t+1} A_{ij} = 0, \tag{5.5}$$

$$\mathrm{LattCons}_i \equiv \rho_i C_{p,i} \frac{T_{L,i}^{t+1} - T_{L,i}^{t}}{\Delta t} V_i - \sum_m C_{i,m}^{E,t+1} V_{i,m} + \sum_j \left(T_{L,i}^{t+1} - T_{L,j}^{t+1}\right) \frac{\kappa_{ij} A_{ij}}{L_{ij}} = 0. \tag{5.6}$$

Densities, fluxes, and collision operators with the subscript m are associated to unique Fermi distributions at the i^{th} semiconductor node of the DV mesh. The electron number and energy densities (with the time step superscript t removed for brevity) are

$$n_i = \sum_k A_{ik} \left(k_B T_i\right)^{\alpha_{ik}+1} F_{\alpha_{ik}}\left(\eta_{\rho,ik}, a_{ik}, b_{ik}\right) \equiv \sum_k n_{ik}, \tag{5.7}$$

$$E_{n,i} = \sum_k \left\{ E_{\rho,ik} n_{ik} + A_{ik} \left(k_B T_i\right)^{\alpha_{ik}+2} F_{\alpha_{ik}+1}\left(\eta_{\rho,ik}, a_{ik}, b_{ik}\right) \right\}, \tag{5.8}$$

and the particle and kinetic energy flux densities are

$$J_{n,ij} = \sum_k \frac{qD_{jk}}{L_j} \left(\text{Ein}_{n,jk}\right)_{\text{ave}} \left[B(\xi_{n,jk}) N_{0,jk} - B(-\xi_{n,jk}) N_{1,jk}\right] \equiv \sum_k J_{n,ijk}, \tag{5.9}$$

$$N_{m,jk} = \left(k_B T_m\right)^{\beta_{jk}} F'_{\beta_{jk}}\left(\eta_{J,m,jk}, c_{jk}, d_{jk}\right), \tag{5.10}$$

$$\xi_{n,jk} = \frac{1}{\left(\text{Ein}_{n,jk}\right)_{\text{ave}}} \left[\text{E}_j + \Delta\text{Ein}_{n,jk}\right] L_j, \tag{5.11}$$

$$K_{n,ij} = \sum_k \left\{ E_{J,jk} J_{n,ijk} + \frac{qD_{jk}}{L_j} \left(\text{Ein}_{E,jk}\right)_{\text{ave}} \left[B(\xi_{E,jk}) E_{0,jk} - B(-\xi_{E,jk}) E_{1,jk}\right] \right\}, \tag{5.12}$$

$$E_{m,jk} = \left(k_B T_m\right)^{\beta_{jk}+1} F'_{\beta_{jk}+1}\left(\eta_{J,m,jk}, c_{jk}, d_{jk}\right), \tag{5.13}$$

$$\xi_{E,jk} = \frac{1}{\left(\text{Ein}_{E,jk}\right)_{\text{ave}}} \left[\text{E}_j + \Delta\text{Ein}_{E,jk}\right] L_j. \tag{5.14}$$

These are the "generalized Einstein" particle fluxes derived in Appendix A.3.2. The script E is implied. The index k is associated to the primary edge with index j. This primary edge corresponds to the stencil associated to the i^{th} primary node. The functions $\left(\text{Ein}_{*,jk}\right)_{\text{ave}}$ and $\Delta\text{Ein}_{*,jk}$, which are discussed in Appendix A.3.2, are

$$\Delta\text{Ein}_{n,jk} = \frac{1}{L_j} \left(\left. \frac{k_B T}{q} \frac{F_{\beta_{jk}}}{F'_{\beta_{jk}}} \right|_{j,1} - \left. \frac{k_B T}{q} \frac{F_{\beta_{jk}}}{F'_{\beta_{jk}}} \right|_{j,0} \right), \tag{5.15}$$

and

$$\left(\text{Ein}_{n,jk}\right)_{\text{ave}} = \frac{1}{2} \left(\left. \frac{k_B T}{q} \frac{F_{\beta_{jk}}}{F'_{\beta_{jk}}} \right|_{j,1} + \left. \frac{k_B T}{q} \frac{F_{\beta_{jk}}}{F'_{\beta_{jk}}} \right|_{j,0} \right). \tag{5.16}$$

Here, the notation "$j,0$" represents node zero of the j^{th} primary edge. The Einstein functions for the kinetic energy flux density are

$$\Delta\text{Ein}_{E,jk} = \frac{1}{L_j}\left(\left.\frac{k_BT}{q}\frac{F_{\beta_{jk}+1}}{F'_{\beta_{jk}+1}}\right|_{j,1} - \left.\frac{k_BT}{q}\frac{F_{\beta_{jk}+1}}{F'_{\beta_{jk}+1}}\right|_{j,0}\right), \tag{5.17}$$

and

$$\left(\text{Ein}_{E,jk}\right)_{\text{ave}} = \frac{1}{2}\left(\left.\frac{k_BT}{q}\frac{F_{\beta_{jk}+1}}{F'_{\beta_{jk}+1}}\right|_{j,1} + \left.\frac{k_BT}{q}\frac{F_{\beta_{jk}+1}}{F'_{\beta_{jk}+1}}\right|_{j,0}\right). \tag{5.18}$$

See Section 4.2 for a detailed discussion on the energy-space discretization and Section 4.3 for details on the real-space discretization. With the definitions

$$J_{0,jk}\left(T_{p,j},T_{q,j}\right) = \frac{qD_{jk}}{L_j}\left(\text{Ein}_{n,jk}\right)_{\text{ave}}\left(T_{p,j},T_{q,j}\right)B(\xi_{n,jk}\left(T_{p,j}\right))N_{0,jk}\left(T_{p,j}\right), \tag{5.19}$$

$$J_{1,jk}\left(T_{p,j},T_{q,j}\right) = \frac{qD_{jk}}{L_j}\left(\text{Ein}_{n,jk}\right)_{\text{ave}}\left(T_{p,j},T_{q,j}\right)B(-\xi_{n,k}\left(T_{p,j}\right))N_{1,jk}\left(T_{p,j}\right), \tag{5.20}$$

$$K_{0,jk}\left(T_{p,j},T_{q,j}\right) = \frac{qD_{jk}}{L_j}\left(\text{Ein}_{E,jk}\right)_{\text{ave}}\left(T_{p,j},T_{q,j}\right)B(\xi_{E,jk}\left(T_{p,j}\right))E_{0,jk}\left(T_{p,j}\right), \tag{5.21}$$

$$K_{1,jk}\left(T_{p,j},T_{q,j}\right) = \frac{qD_{jk}}{L_j}\left(\text{Ein}_{E,jk}\right)_{\text{ave}}\left(T_{p,j},T_{q,j}\right)B(-\xi_{E,jk}\left(T_{p,j}\right))E_{1,jk}\left(T_{p,j}\right), \tag{5.22}$$

the heat flow and total energy flux densities are

$$\begin{aligned} H_{n,ij} = \sum_k \big\{ & K_{0,jk}(T_{j,0},T_{j,1}) - K_{0,jk}(T_{j,1},T_{j,1}) \\ & - (F-E_c)_0\left[J_{0,jk}(T_{j,0},T_{j,1}) - J_{0,jk}(T_{j,1},T_{j,1})\right] \\ & - K_{1,jk}(T_{j,0},T_{j,1}) + K_{1,jk}(T_{j,0},T_{j,0}) \\ & + (F-E_c)_1\left[J_{1,jk}(T_{j,0},T_{j,1}) - J_{1,jk}(T_{j,0},T_{j,0})\right]\big\}, \end{aligned} \tag{5.23}$$

$$S_{n,ij} = K_{n,ij} + H_{n,ij}. \tag{5.24}$$

5.2.1 Solution Variables

The solution variables of Eqns. (5.1) – (5.24) are the electric potential Φ_i^{t+1}, rotational electric field $\text{E}_{\text{rot},i}^{t+1}$, magnetic field H_i^{t+1}, the Fermi level relative to the conduction band minimum in each

material $(F - E_c)_i^{t+1}$, the electron temperature $(k_B T)_i^{t+1}$ of a discrete electron gas, and the lattice temperature $(k_B T_L)_i^{t+1}$. Assignment of the electron gas solution variables in real- and energy-space and electromagnetics solution variables are next presented.

5.2.1.1 Electron Gases in Real-Space

One or more electron gases are assigned to each Voronoi polygon (2D) or polyhedron (3D) in the DV mesh located in a semiconducting material. If a semiconductor mesh node lies on a material interface, then the Voronoi cell is split into sub-regions. Each sub-region corresponding to a unique material is assigned a separate electron gas. Primary nodes on the boundary of the mesh are truncated in a similar manner. Section 2.2 provides a discussion on material interface and mesh boundary splitting of the Voronoi diagram.

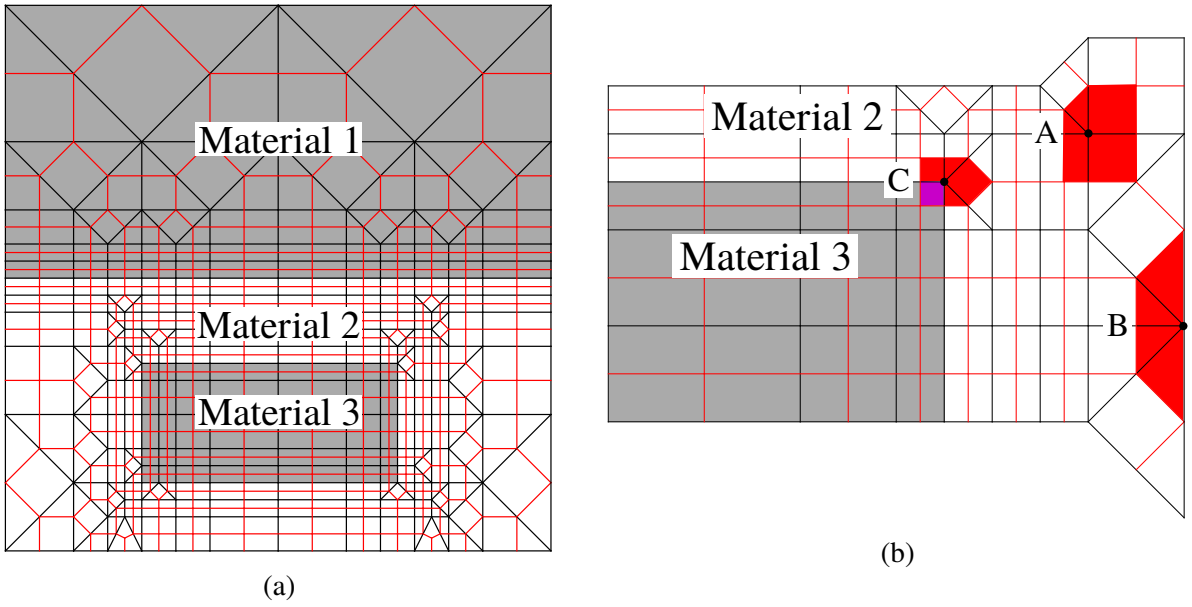


Figure 5.1: (a) An example of a device mesh generated by the Sentaurus meshing tool. (b) An illustration of the electron gases assigned to unique Voronoi polygons in the DV mesh.

Figure 5.1a presents an example of a DV mesh on which electron gases are assigned. Three Voronoi polygons are drawn in Figure 5.1b. The Voronoi polygon associated to gas A is completely enclosed in material 2 and therefore the entire polygon shaded in red represents the volume of the gas. Because gas B is associated to a node on the boundary of the mesh, the Voronoi polygon is

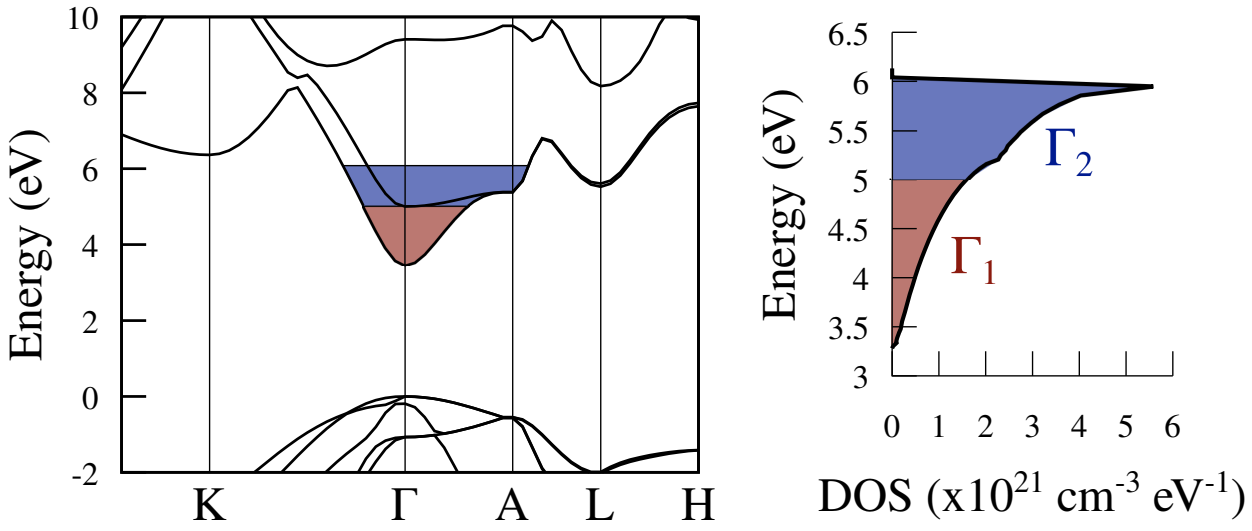


Figure 5.2: The (left) panel diagram of wurtzite GaN and (right) corresponding DOS isosurface integral of the first Γ-valley. The energy-space of the first Γ-valley is split into two distinct Fermi distributions doubling the number of electron gases in the semiconductor mesh.

truncated at the boundary. Node C is an important case in semiconductor device meshes. Because this mesh node resides on a material interface, it receives separate gases associated to the two unique materials. The Voronoi volume of gas C3 (the gas associated to material 3) is shaded in pink and the Voronoi volume of gas C2 (the gas associated to material 2) is shaded in red. This procedure generalizes to interfaces of an arbitrary number of materials.

5.2.1.2 Electron Gases in Energy-Space

Energy-space discretization was presented in Section 4.2. The band-structure of a bulk semiconductor is incorporated into the FKT device simulator through piece-wise energy power law fits of the isosurface integrals. For example, the power law fits are used to calculate the electron density and energy density, Eqns. (4.49) and (4.50). These densities are associated to a sub-region of the Voronoi cell corresponding to a unique material as described in the previous section.

The energy-space corresponding to an electron gas in a unique material can further be discretized into multiple Fermi distributions. Typically, valleys in the semiconductor band-structure are assigned separate distribution functions. Scattering between the distributions is calculated with

the collision operators C^n and C^E [83]. Figure 5.2 presents an illustration of a panel diagram and the DOS isosurface integral, Eqn. (4.30), of wurtzite GaN [84]. The band-structure was calculated with EPM [89]. The DOS data was obtained from Matt Grupen AFRL [84]. Only the first Γ-valley is included in the DOS calculation and the energy-space of the valley is split into two separate Fermi distributions labeled Γ_1 and Γ_2. The separation of the first Γ-valley into two distinct Fermi distributions yields two sets of solution variables $(F-E_c)^{t+1}$ and $(k_BT)^{t+1}$ at each semiconductor mesh node. Real-space transport only occurs between the same distributions in energy-space. Therefore, there is no particle flux between the Γ_1 and Γ_2 distributions. Rather, particles can drift and diffuse between Γ_1 at mesh node 1 and Γ_1 at mesh node 2 and then scatter from Γ_1 to Γ_2 at mesh node 2. The same is true for energy transport.

5.2.1.3 Electromagnetics

The solution variables corresponding to full-wave EM are the rotational electric field $\mathrm{E}_{\mathrm{rot},i}$ and the magnetic field H_i. The rotational electric field vector projections are defined on the primary edges of the Delaunay mesh and magnetic field vector projections are defined on the edges of the Voronoi diagram. A derived solution variable is the total electric field. On the j^{th} primary edge, the total electric field is

$$\mathrm{E}_j = \mathrm{E}_{\mathrm{rot},j} + \frac{\Phi_{j,0} - \Phi_{j,1}}{L_j}, \tag{5.25}$$

where the subscript $j,0$ corresponds to node zero of the primary edge. The primary edge points from node 0 to 1. A complete discussion of the full-wave EM discretization, DVSI, is presented in Chapter 3.

5.3 The Nonlinear Solver: Newton's Method

The coupled nonlinear system, Eqns. (5.1) – (5.24), is solved with Newton's method. See Section 2.3.1.2 for a detailed description on the general Newton's method solution of systems of nonlinear equations. The pseudo-code outlining the Newton algorithm in Section 2.3.1.2 illustrates the need to compute the residual vector and Jacobian matrix at each iteration. The residual vector

is simply the concatenation of the discrete nonlinear Eqns. (5.1) – (5.5). These equations are represented as Pois_i, Amp_i, Far_i, ElCont_i, EnCons_i, and LattCons_i, respectively. The Jacobian matrix J_{ij} is the partial derivative of the i^{th} residual with respect to the j^{th} solution variable.

The Newton linear system of an FKT device simulation is now discussed. The global nonlinear system consists of N_{phi} electric potential solution variables $\left\{\Phi_1, \Phi_2, ..., \Phi_{N_{\text{phi}}}\right\}$, N_F electron relative Fermi level solution variables $\left\{(F-E_c)_1, (F-E_c)_2, ..., (F-E_c)_{N_F}\right\}$, N_T electron temperature solution variables $\left\{k_BT_1, k_BT_2, ..., k_BT_{N_T}\right\}$, and N_{TL} lattice temperature solution variables $\left\{k_BT_{L,1}, k_BT_{L,2}, ..., k_BT_{L,N_{TL}}\right\}$. If full-wave EM is included in the device simulation, there are an additional N_{E} rotational electric field solution variables $\left\{\text{E}_{\text{rot},1}, \text{E}_{\text{rot},2}, ..., \text{E}_{\text{rot},N_{\text{E}}}\right\}$ and N_H magnetic field solution variables $\left\{H_1, H_2, ..., H_{N_H}\right\}$.

A linear system is solved at each Newton iteration. The linear system is comprised of the Poisson residuals Pois_i, then the Ampere residuals Amp_i, then Far_i, ElCont_i, EnCons_i, and LattCons_i. With this arrangement, the full-wave FKT Newton linear system is

$$\begin{bmatrix} \frac{\partial \text{Pois}_1}{\partial \Phi_1} & \frac{\partial \text{Pois}_1}{\partial \Phi_2} & \cdots & \frac{\partial \text{Pois}_1}{\partial k_BT_{L,N_{TL}}} \\ \frac{\partial \text{Pois}_2}{\partial \Phi_1} & \frac{\partial \text{Pois}_2}{\partial \Phi_2} & \cdots & \frac{\partial \text{Pois}_2}{\partial k_BT_{L,N_{TL}}} \\ \vdots & & \ddots & \vdots \\ \frac{\partial \text{LattCons}_{N_{TL}}}{k_BT_{L,1}} & \frac{\partial \text{LattCons}_{N_{TL}}}{k_BT_{L,2}} & \cdots & \frac{\partial \text{LattCons}_{N_{TL}}}{k_BT_{L,N_{TL}}} \end{bmatrix} \begin{bmatrix} \Delta\Phi_1 \\ \Delta\Phi_2 \\ \vdots \\ \Delta k_BT_{L,N_{TL}} \end{bmatrix} = - \begin{bmatrix} \text{Pois}_1 \\ \text{Pois}_2 \\ \vdots \\ \text{LattCons}_{N_{TL}} \end{bmatrix}. \tag{5.26}$$

The solution of Eqn. (5.26) determines the change in solution variables $\{\Delta\Phi_1, \cdots, \Delta\text{E}_{\text{rot},1}, \cdots, \Delta H_1, \cdots, \Delta(F-E_c)_1, \cdots, \Delta k_BT_1, \cdots, \Delta k_BT_{L,N_{TL}}\}$. The global solution variable set is updated according to

$$\begin{aligned} \Phi_1^{m+1} &= \Phi_1^m + \Delta\Phi_1 \\ &\vdots \\ \text{E}_{\text{rot},1}^{m+1} &= \text{E}_{\text{rot},1}^m + \Delta\text{E}_{\text{rot},1} \\ &\vdots \\ k_BT_{L,N_{TL}}^{m+1} &= k_BT_{L,N_{TL}}^m + \Delta k_BT_{L,N_{TL}}. \end{aligned} \tag{5.27}$$

5.4 Boundary Conditions

A list of the BCs required for electronic device simulations is presented in Table 5.1. Typical BCs include metal contacts and port BCs used to provide or absorb EM energy.

B.C.	Φ	$\mathrm{E}_{\mathrm{rot}}$	H	$F-E_c$	k_BT	k_BT_L
Ohmic	$\Phi_{\mathrm{CN}}+V_a$	0	0	$(F-E_c)_{\mathrm{CN}}$	k_BT_L	Soln.
Schottky	$\Phi_{\mathrm{SB}}+V_a$	0	0	Soln.	Soln.	Soln.
Lattice Heat Absorbing	Soln.	Soln.	Soln.	Soln.	Soln.	Soln.
AC Impedance (Ohmic)	$\Phi_{\mathrm{ACI}}^{t+1}$	0	0	$(F-E_c)_{\mathrm{CN}}$	k_BT_L	Soln.
AC Impedance (Schottky)	$\Phi_{\mathrm{ACI}}^{t+1}$	0	0	Soln.	Soln.	Soln.
Full-wave Voltage	Soln.	0	Soln.	Soln.	Soln.	Soln.
Full-wave Impedance	Soln.	Soln.	H_{imp}	Soln.	Soln.	Soln.

Table 5.1: A list of semiconductor device simulation BCs.

5.4.1 Ohmic Metal Contact

An essential metal BC found in almost every semiconductor device simulation is the Ohmic contact. Charge neutrality is enforced at a distribution on an interface between a semiconductor and an Ohmic contact. The solution variable $F-E_c$ of a charge distribution on an Ohmic contact interface is fixed to the charge neutral value $(F-E_c)_{\mathrm{CN}}$ and maintains the electric potential $\Phi_{\mathrm{CN}}+V_a$. Here, V_a is a voltage applied to the Ohmic contact. The charge neutral quantities are determined by the numerical solution of

$$\mathrm{Pois}_i|_{\mathrm{ohm.}} \equiv \sum_m \left(N_{D,i,m}^{+} - n_{i,m}^{t+1}\right) V_{i,m} = 0, \tag{5.28}$$

and the assignment

$$\Phi_{\mathrm{CN}} = -\chi_e + \Phi_M + (F-E_c)_{\mathrm{CN}}. \tag{5.29}$$

Here, the electron affinity χ_e and the work-function Φ_M are properties of the semiconductor and metal, respectively.

5.4.2 Lattice Heat Absorbing Boundary Condition

The heat absorbing (HA) BC is meant to emulate lattice heat flow out of the device domain. Without this BC, the simulated devices will heat up to nonphysical levels. Outgoing lattice heat flow is represented by the outward lattice temperature gradient $(\nabla T_L)_{\mathrm{HA}}$ on a boundary mesh node. The gradient flux is through the in-plane Voronoi polygons associated to primary nodes on the mesh boundary (see Figure 2.8). The total HA BC is

$$\begin{aligned}\mathrm{LattCons}_i|_{\mathrm{HA}} \equiv & \left[\rho_i C_{p,i}\frac{T_{L,i}^{t+1}-T_{L,i}^{t}}{\Delta t}-\sum_i C_i^{E,t+1}\right]V_i \\ & +\sum_j\left(T_{L,i}^{t+1}-T_{L,j}^{t+1}\right)\frac{\kappa_{ij}A_{ij}}{L_{ij}}+[\hat{n}\cdot(\nabla T_L)\,\kappa A]_{\mathrm{HA}}=0. \end{aligned} \tag{5.30}$$

Here, $\hat{n}$ represents the normal of the mesh boundary. The lattice temperature gradient at the mesh nodes of the HA boundary are reconstructed from the vector projections of the lattice temperature gradients onto the primary edges associated to the mesh nodes.

5.4.3 Quasi-Static AC Impedance Boundary Condition

The quasi-static solver is excellent for lower frequency simulations where full-wave effects are negligible. In order to calculate useful quantities including S-Parameters with the quasi-static solver, the AC impedance must be used as a termination opposite the excitation contact. The AC impedance BC is defined on metal contacts in the quasi-static solver. The integrated current flowing into the AC impedance contact is

$$i^{t+1}=\sum_k J_k^{t+1}A_k. \tag{5.31}$$

Here, J_k represents the particle flux density on the k^{th} primary edge pointing into the AC impedance metal contact. The area A_k is that of the in-plane Voronoi polygon associated to the primary edge (see Figure 2.8). This integrated current specifies a Dirichlet BC on the metal contact's potential via Ohm's law, i.e.,

$$\Phi_i^{t+1}\Big|_{\mathrm{ACI}}=Z_{\mathrm{ACI}}i_{\mathrm{ACI}}^{t+1}, \tag{5.32}$$

where $i_{\mathrm{ACI}}^{t+1}\equiv i^{t+1}-i^{t=0}$ represents only the AC component of the integrated current.

5.4.4 Full-Wave Voltage Port

The voltage BC is a specific case of the port BC in which EM energy enters the full-wave solver. The voltage port connects two metal contacts in the DUT. The electric field in the voltage port is strictly irrotational, i.e., $\mathrm{E}_{\mathrm{rot},i} = 0$. Ampere's law is solved on voltage port primary edges. The BC is

$$\mathrm{Amp}_i|_{\mathrm{volt.}} \equiv \left[\mathrm{E}_{i,\mathrm{irr}}^{t+1} - \mathrm{E}_{i,\mathrm{irr}}^{t}\right] \frac{\varepsilon_i A_i}{\Delta t} - q\sum_m J_{n,im}^{t+1} A_{im} - \sum_j H_j^{t+1} L_j = 0. \tag{5.33}$$

Here, the irrotational electric field is $\mathrm{E}_{i,\mathrm{irr}}^{t+1} = \frac{\Phi_{i,0} - \Phi_{i,1}}{L_i}$ where the subscript $i,0$ corresponds to node zero of the i^{th} primary edge. The primary edge points from node 0 to 1. This BC serves as the governing equation for the magnetic field defined on the i^{th} in-plane dual edge associated to a primary edge on the boundary (see Figure 2.8). This magnetic field is included in the line integral of Eqn. (5.33).

5.4.5 Full-Wave AC Impedance Port

The full-wave AC impedance (FWACI) BC is an energy sink port BC connecting two metal contacts. The magnetic field H_{FWACI} is defined on an in-plane dual edge associated to a primary edge on the boundary. The magnetic field is equal to the total electric field on the primary edge divided by the boundary's impedance Z_{FWACI}. The BC is

$$\mathrm{FWACI}_i \equiv \mathrm{E}_i^{t+1} - \mathrm{E}_i^{t=0} - Z_{\mathrm{FWACI}}\left(H_{i,\mathrm{FWACI}}^{t+1} - H_{i,\mathrm{FWACI}}^{t=0}\right) = 0, \tag{5.34}$$

In a manner similar to the quasi-static AC impedance BC, the subtraction of the solutions calculated at the first time step ensures that the BC only affects the AC components of the solution variables. This emulates a bias tee circuit.

5.4.6 Heterojunctions and Thermionic Emission Boundaries

As outlined in Section 5.2.1.1, one or more electron gases are assigned to semiconductor mesh nodes residing on heterojunctions. Figure 5.3 illustrates an example of a heterojunction between

two semiconductors. In this example, each electron gas is approximated with a single Fermi distribution in energy-space. Because the two materials have different electron affinities, there is a discontinuity in the conduction band across the heterojunction and thermionic emission is used to calculate real-space transport across the discontinuity. With both distributions' Fermi levels referenced to the higher conduction band, the particle flux and kinetic energy flux density thermionic emissions from semiconductor 1 to 2 are

$$J_{n,TE}^{1\to 2}(T_1) = \frac{1}{2\pi^2}\frac{m^*}{\hbar^3}(k_B T_1)^2 \int_0^\infty d\varepsilon_z \ln\left[1+\exp\left(\tilde{\eta}_1 - \varepsilon_z\right)\right], \tag{5.35}$$

and

$$K_{n,TE}^{1\to 2}(T_1) = \frac{1}{\pi^2}\frac{m^*}{\hbar^3}(k_B T_1)^3 \int_0^\infty d\varepsilon_z \varepsilon_z \ln\left[1+\exp\left(\tilde{\eta}_1 - \varepsilon_z\right)\right], \tag{5.36}$$

with the parameter

$$\tilde{\eta}_1 = \frac{(F_n - E_c)_1 + E_{c,1} - E_{c,\text{high}}}{k_B T_1}. \tag{5.37}$$

A numerical quadrature routine is required to evaluate Eqns. (5.35) and (5.36). The high side conduction band edge $E_{c,\text{high}}$ is associated to the charge distribution in the material with the smaller electron affinity at the heterojunction. The total thermionic emission particle, kinetic energy, heat, and total energy flux densities are

$$J_{n,TE} = J_{n,TE}^{1\to 2}(T_1) - J_{n,TE}^{2\to 1}(T_2), \tag{5.38}$$

$$K_{n,TE} = K_{n,TE}^{1\to 2}(T_1) - K_{n,TE}^{2\to 1}(T_2), \tag{5.39}$$

$$\begin{aligned} H_{n,TE} = {} & K_{n,TE}^{1\to 2}(T_1) - K_{n,TE}^{1\to 2}(T_2) \\ & - (F - E_{c,\text{high}})_1 \left[J_{n,TE}^{1\to 2}(T_1) - J_{n,TE}^{1\to 2}(T_2)\right] \\ & - K_{n,TE}^{2\to 1}(T_2) + K_{n,TE}^{2\to 1}(T_1) \\ & + (F - E_{c,\text{high}})_2 \left[J_{n,TE}^{2\to 1}(T_2) - J_{n,TE}^{2\to 1}(T_1)\right], \end{aligned} \tag{5.40}$$

and

$$S_{n,TE} = K_{n,TE} + H_{n,TE}. \tag{5.41}$$

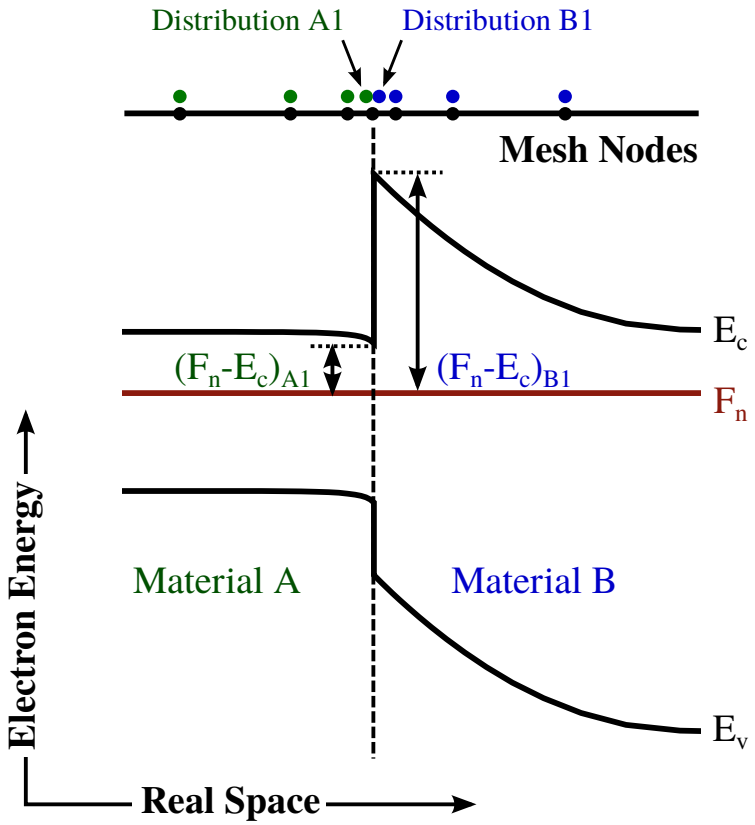


Figure 5.3: The band diagram across a two material heterojunction. The electron affinity of material A is greater than material B's electron affinity. The relative Fermi level solution variables are illustrated for each charge gas.

The thermionic particle flux and total energy flux densities are added to the electron continuity and energy conservation divergences, Eqns. (5.4) and (5.5). Figure 5.4 illustrates the particle divergence for the two unique charge distributions defined on a heterojunction mesh node. There is only a BTE particle flux ($J_{n,ij}$ and $K_{n,ij}$) between charges associated to the same material. Thermionic emission connects the two charge distributions at the heterojunction. With charge distributions $A1$ and $B1$ at the heterojunction mesh node, the electron continuity and energy conservation equations for both distributions are

$$\text{ElCont}_{A1}|_{\text{HJ}} \equiv \left(\frac{n_{A1}^{t+1} - n_{A1}^{t}}{\Delta t} + C_{A1}^{n,t+1} \right) V_{A1} + \sum_j J_{n,A1j}^{t+1} A_{A1j} + J_{n,TE}^{t+1} A_{TE} = 0, \tag{5.42}$$

$$\text{ElCont}_{B1}|_{\text{HJ}} \equiv \left(\frac{n_{B1}^{t+1} - n_{B1}^{t}}{\Delta t} + C_{B1}^{n,t+1} \right) V_{B1} + \sum_j J_{n,B1j}^{t+1} A_{B1j} - J_{n,TE}^{t+1} A_{TE} = 0, \tag{5.43}$$

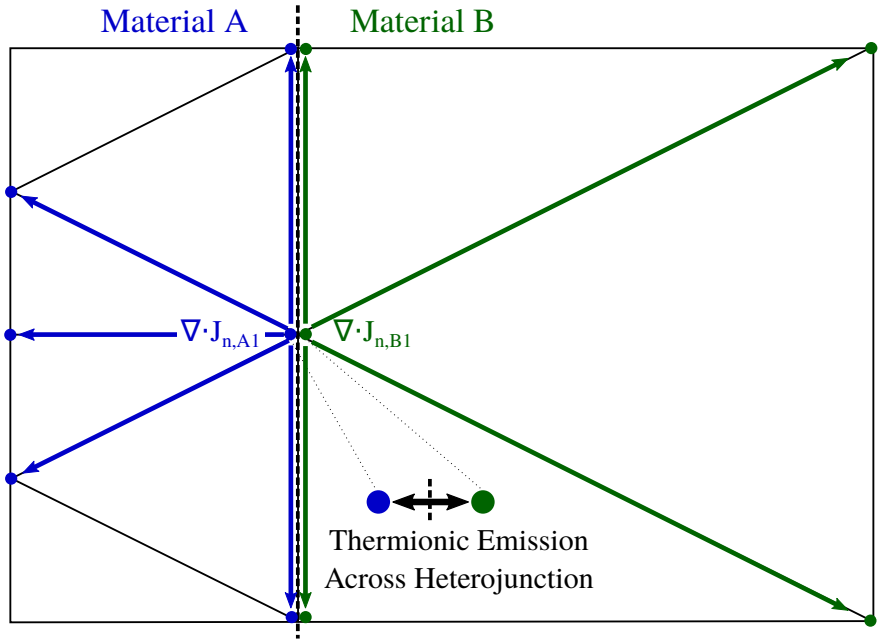


Figure 5.4: An example of the particle divergences associated to two distributions at a heterojunction. Particle fluxes are calculated between two electron gases in the same semiconductor.

$$\begin{aligned}\mathrm{EnCons}_{A1}|_{\mathrm{HJ}} \equiv & \left(\frac{E_{n,A1}^{t+1}-E_{n,A1}^{t}}{\Delta t}+C_{A1}^{E,t+1}\right)V_{A1} \\ & +q\sum_{j}\left(\mathrm{E}_{j}^{t+1}L_{j}\right)J_{n,ij}^{t+1}A_{A1j}+\sum_{j}S_{n,A1j}^{t+1}A_{ij} \\ & +S_{n,TE}^{t+1}A_{TE}+E_{\mathrm{pot}}J_{n,TE}^{t+1}A_{TE}=0. \end{aligned} \tag{5.44}$$

$$\begin{aligned}\mathrm{EnCons}_{B1}|_{\mathrm{HJ}} \equiv & \left(\frac{E_{n,B1}^{t+1}-E_{n,B1}^{t}}{\Delta t}+C_{B1}^{E,t+1}\right)V_{B1} \\ & +q\sum_{j}\left(\mathrm{E}_{j}^{t+1}L_{j}\right)J_{n,B1j}^{t+1}A_{B1j}+\sum_{j}S_{n,B1j}^{t+1}A_{ij}-S_{n,TE}^{t+1}A_{TE}=0. \end{aligned} \tag{5.45}$$

The last term in Eqn. (5.45) represents the potential energy required to overcome the barrier $E_{\mathrm{pot}} = \chi_A - \chi_B$ with $\chi_A > \chi_B$. The area A_{TE} corresponds to the Voronoi polygon in the heterojunction plane associated to the primary node on the interface (see Figure 2.8).

5.4.7 Schottky Metal Contact

Another important BC is the rectifying Schottky metal contact. This metal-semiconductor interface is a special type of heterojunction. The important characteristic of the BC is the Schottky barrier Φ_{SB}. With no applied bias, the solution variable $F - E_c$ of a charge distribution at this interface equals $-q\Phi_{\text{SB}}$. In the typical way, an applied bias which raises the Fermi level in the Schottky metal will increase the potential barrier of the metal-semiconductor interface. An applied bias which lowers the Fermi level will decrease the potential barrier. Transport across the discontinuity caused by the Schottky barrier is calculated with thermionic emission. The energy $-q\Phi_{\text{SB}}$ and the lattice temperature $k_B T_L$ are used to compute the thermionic emission from the metal into the semiconductor.

5.5 Semiconductor Device Simulation Work-Flow

In this section, the device simulation work-flow is discussed. Device solves include equilibrium, static, quasi-static, and full-wave simulations. Each device solve will be reviewed and discussed.

5.5.1 Work-Flow Overview

An overview of the semiconductor device simulation work-flow is provided in Figure 5.5. The input to any device simulation is the mesh of the DUT. Included in the DUT mesh are material specifications, dopant profiles and quantities, and BCs including Ohmic/Schottky contacts. The solve types are grouped into three distinct classes: Equilibrium, static, and transient solves.

The Newton linear system corresponding to a specific device solve will be presented in block matrix form. For example, the matrix corresponding to the partial derivative of the i^{th} Poisson residual with respect to the j^{th} electric potential solution variable is $\frac{\partial \text{Pois}_i}{\partial \Phi_j}$. The i^{th} diagonal of this matrix is

$$\frac{\partial \text{Pois}_i}{\partial \Phi_i} = \sum_j \frac{\varepsilon_{ij} A_{ij}}{L_{ij}}. \tag{5.46}$$

The derivative of the i^{th} Poisson residual with respect to the j^{th} Fermi level solution variable is

$$\frac{\partial \text{Pois}_i}{\partial (F - E_c)_j} = q \sum_m \frac{\partial n_{i,m}^{t+1}}{\partial (F - E_c)_j} V_{i,m}. \tag{5.47}$$

This Jacobian element requires partial derivatives of the densities with respect to the relative Fermi levels which amounts to derivatives of incomplete Fermi integrals in terms of the derived solution variables η. The numerical routine which calculations the incomplete Fermi integrals also returns the derivative of the Fermi integrals.

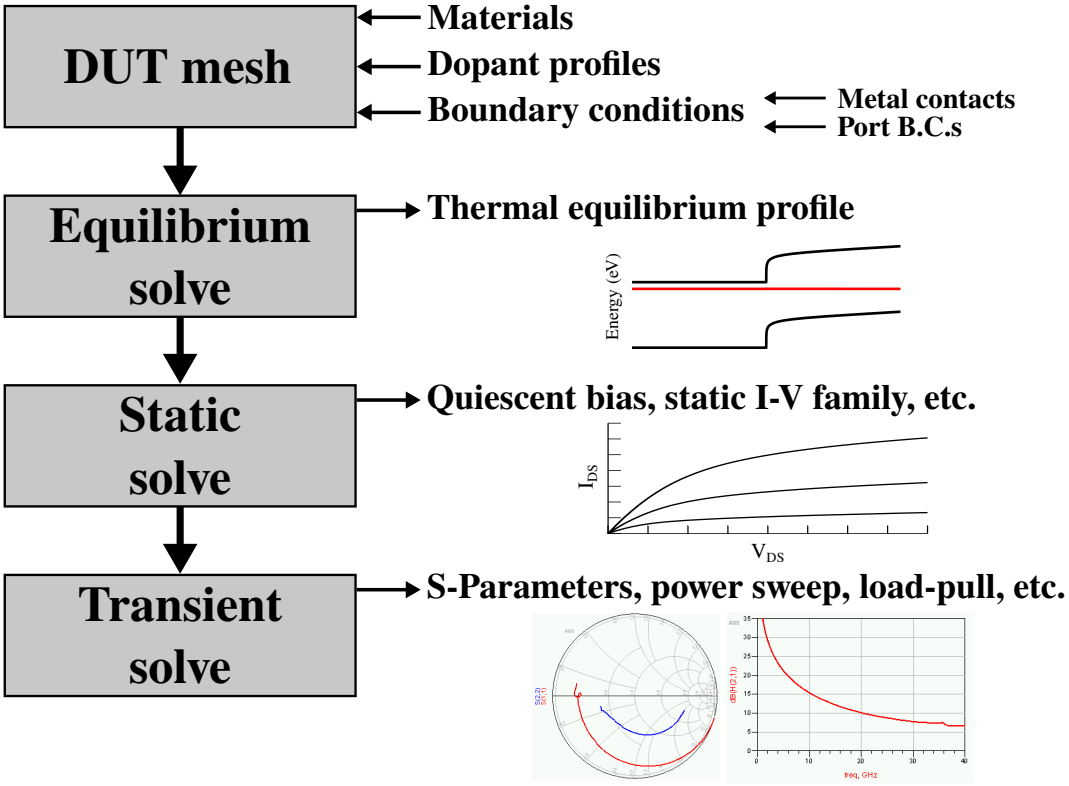


Figure 5.5: A general work-flow diagram of a semiconductor device simulation.

The following sections describe the solve types in more detail. To do so, a practical example is utilized — a GaN MESFET. Figure 5.6 presents an illustration of the 3D MESFET including the material definitions and the metal contacts. The Ohmic source and drain and the Schottky gate are the yellow regions at the top of the mesh. The Ohmic ground at the bottom of the mesh is not shown. The dark green region is the GaN channel and the gray region is the insulating

substrate. The transport direction (from source to drain) is along the x-axis in this example. The substrate/semiconductor interface is perpendicular to the z-axis.

A complete device simulation is conducted with the 3D GaN MESFET example. First, the thermal equilibrium profile is calculated with the equilibrium solver. Then, the quiescent bias is calculated and finally the S-Parameters are calculated with the quasi-static and full-wave solvers.

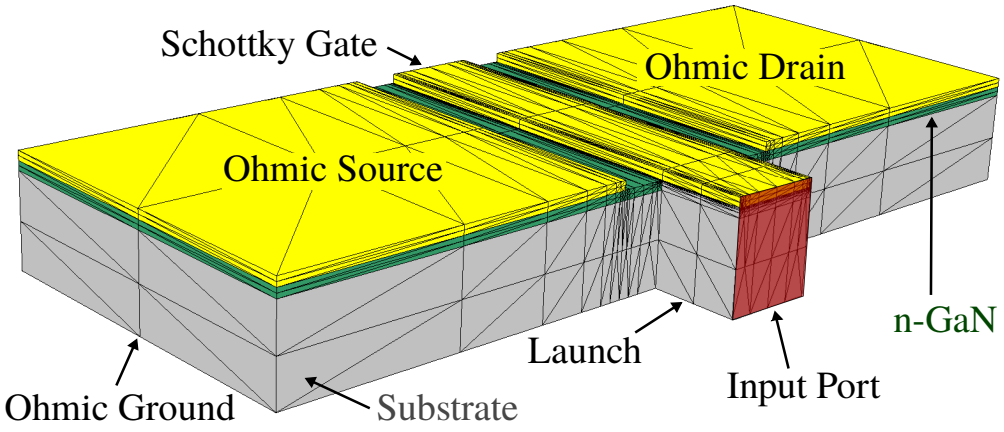


Figure 5.6: The 3D GaN MESFET example used to demonstrate the device simulation work-flow. The yellow regions at the top of the mesh are the Ohmic source and drain and the Schottky gate contacts. The bottom surface of the mesh is the Ohmic ground contact (not shown). The GaN channel is the dark green region and the insulating substrate is the gray region. The volume labeled "launch" represents the insulating region connected to the voltage port (the red surface). This is how EM energy is injected into the domain (see Section 5.4 for more information on the port BCs).

5.5.2 Equilibrium Solve

The equilibrium solve is the first component of the device simulation work-flow. This produces the electric potential and electron relative Fermi level profiles in thermal equilibrium. Only Poisson's equation, Eqn. (5.1), is solved to produce the thermal equilibrium solution profiles. The Newton linear system of this solve is

$$\left[\frac{\partial \text{Pois}_i}{\partial \Phi_j} \right] \left[\Delta \Phi_j \right] = - \left[\text{Pois}_i \right]. \tag{5.48}$$

Because the conduction band E_c will vary across the device in thermal equilibrium, the relative Fermi level solution variable $F - E_c$ must also be updated after each Newton iteration. This can be

accomplished by amending the diagonals of the Jacobian to be

$$\begin{aligned}\frac{\partial \text{Pois}_i}{\partial \Phi_i} &= \sum_j \frac{\varepsilon_{ij} A_{ij}}{L_{ij}} + q \sum_m \frac{\partial n_{i,m}^{t+1}}{\partial \Phi_i} V_{i,m} \\ &= \sum_j \frac{\varepsilon_{ij} A_{ij}}{L_{ij}} - q \sum_m \frac{\partial n_{i,m}^{t+1}}{\partial (E_c)_i} V_{i,m} \\ &= \sum_j \frac{\varepsilon_{ij} A_{ij}}{L_{ij}} + q \sum_m \frac{\partial n_{i,m}^{t+1}}{\partial (F - E_c)_i} V_{i,m}. \end{aligned} \tag{5.49}$$

The relation $E_c = -\chi - \Phi$ dictates that a positive change in the electric potential corresponds to a negative change in the conduction band. Furthermore, the electron Fermi level F is spatially constant at thermal equilibrium which allows for the final substitution in Eqn. (5.49). The thermal equilibrium solution variables are updated as

$$\Phi_i^{m+1} = \Phi_i^m + \Delta \Phi_i \tag{5.50}$$

$$(F - E_c)_i^{m+1} = (F - E_c)_i^m + \Delta \Phi_i. \tag{5.51}$$

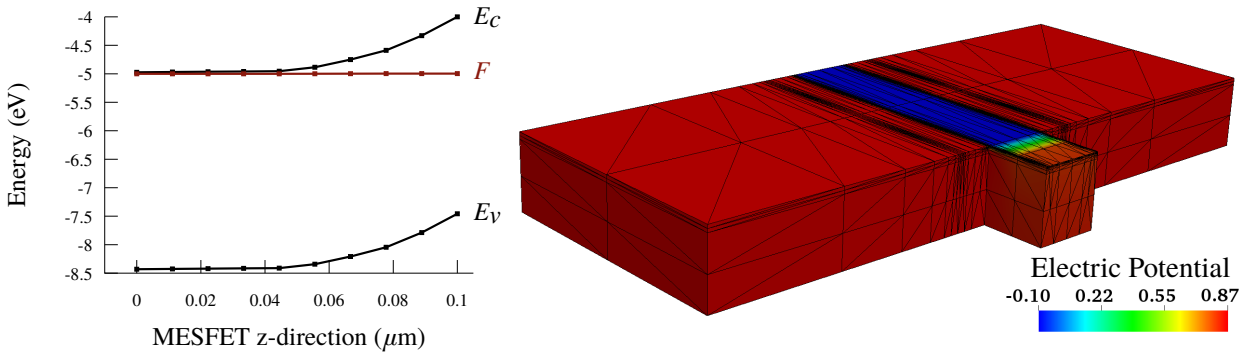
Here, the superscript m represents the Newton iterations.

After the Poisson stage of the thermal equation solve, the rest of the steady-state equations are added one at a time to allow all solution variables to relax to their numerical thermal equilibrium. The Newton linear system of the final equilibrium stage, comprised of the residuals Pois_i, ElCont_i, EnCons_i, and LattCons_i, is shown in Eqn. (5.52). The additional stages of the thermal equilibrium solves do not drastically change the electric potential or relative Fermi level solution variables.

The semiconductor band diagram and the electric potential profile of the 3D GaN MESFET example in thermal equilibrium are shown in Figures 5.7a and 5.7b, respectively. As shown in Figure 5.6, the metal contact located at the top middle of the device is a Schottky contact. In this example, the Schottky barrier is $\Phi_{\text{SB}} = 1.0\,V$.

5.5.3 Static Solve

The next solve in the semiconductor device simulation work-flow is the static or steady-state solve. Typically, an external voltage or current bias is applied to metal contacts to produce steady-state



(a) (b)

Figure 5.7: The thermal equilibrium solution variable profile of the 3D GaN MESFET example. (a) The semiconductor band diagram illustrating the relative Fermi level solution variable in the z-direction of the MESFET. (b) The spatial profile of the electric potential solution variable.

current flow. The static solver computes steady-state data including the current-voltage (I-V) family and transconductance. The quiescent bias of a transistor is also calculated with the static solver.

As described in previous sections, a linear system is solved at each Newton iteration in order to update the FKT device simulation solution variables. The linear system of the m^{th} static solve Newton iteration is

$$\begin{bmatrix} \frac{\partial \text{Pois}_i}{\partial \Phi_j} & \frac{\partial \text{Pois}_i}{\partial (F-E_c)_j} & \frac{\partial \text{Pois}_i}{\partial k_B T_j} & 0 \\ \frac{\partial \text{ElCont}_i}{\partial \Phi_j} & \frac{\partial \text{ElCont}_i}{\partial (F-E_c)_j} & \frac{\partial \text{ElCont}_i}{\partial k_B T_j} & \frac{\partial \text{ElCont}_i}{\partial k_B T_{L,j}} \\ \frac{\partial \text{EnCons}_i}{\partial \Phi_j} & \frac{\partial \text{EnCons}_i}{\partial (F-E_c)_j} & \frac{\partial \text{EnCons}_i}{\partial k_B T_j} & \frac{\partial \text{EnCons}_i}{\partial k_B T_{L,j}} \\ 0 & \frac{\partial \text{LattCons}_i}{\partial (F-E_c)_j} & \frac{\partial \text{LattCons}_i}{\partial k_B T_j} & \frac{\partial \text{LattCons}_i}{\partial k_B T_{L,j}} \end{bmatrix} \begin{bmatrix} \Delta\Phi_j \\ \Delta(F-E_c)_j \\ \Delta k_B T_j \\ \Delta k_B T_{L,j} \end{bmatrix} = - \begin{bmatrix} \text{Pois}_i \\ \text{ElCont}_i \\ \text{EnCons}_i \\ \text{LattCons}_i \end{bmatrix}. \tag{5.52}$$

The static FKT device simulation equations include Eqns. (5.1), (5.4), (5.5), and (5.6) with all time derivatives set to zero, i.e., $\frac{\partial}{\partial t} \to 0$. The Jacobian matrix, Eqn. (5.52), is more complex than the equilibrium Jacobian matrix. For example, the middle diagonal matrix element requires the partial derivative of the collision operator and the particle flux with respect to the relative Fermi level solution variable. This Jacobian matrix element is

$$\frac{\partial \text{ElCont}_i}{\partial (F-E_c)_j} = \frac{\partial C_i^{n,t+1}}{\partial (F-E_c)_j} V_i + \sum_j \frac{\partial J_{n,ij}^{t+1}}{\partial (F-E_c)_j} A_{ij}. \tag{5.53}$$

Derivatives of incomplete Fermi integrals are required for evaluating this Jacobian matrix element.

The particle flux is exponentially dependent on ratios of Fermi integrals making this matrix element highly nonlinear in terms of the solution variables.

Two static simulations of the GaN MESFET are now presented. Figure 5.8 illustrates the static

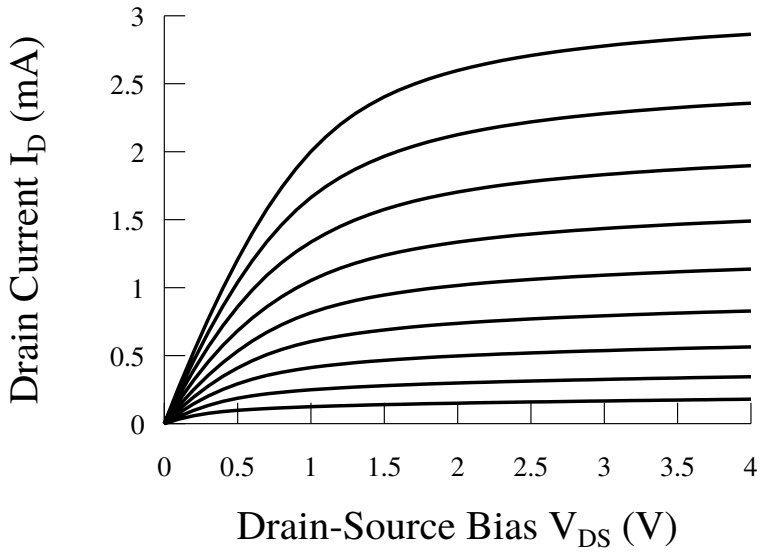


Figure 5.8: The I-V family calculated with the static solve. The gate-source biases start at $0.5\,V$ (top curve) and decrease by steps of $0.25\,V$.

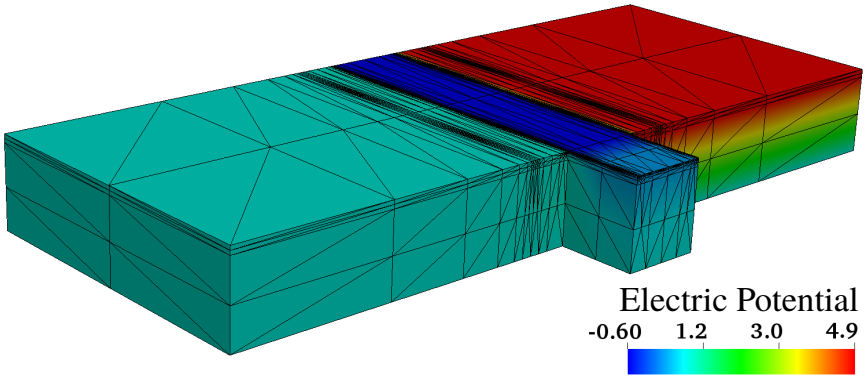


Figure 5.9: The spatial profile of the electric potential solution variable at a specific quiescent bias. The gate-source bias is $-0.5\,V$ and the drain-source bias is $4\,V$.

I-V family of the MESFET. The gate-source bias ranges from $V_{GS} = 0.5$ V to -1.5 V with a 0.25 V step. The drain-source bias is swept from $V_{DS} = 0$ to $4\,V$ for each gate bias. The electric potential solution variable profile of the MESFET with a quiescent bias of $V_{GS} = -0.5$ V and $V_{DS} = 4$ V is presented in Figure 5.9.

5.5.4 Quasi-Static Solve

After the quiescent bias of the device is calculated by the static solve, the quasi-static solve can be used to simulate the RF response without the inclusion of full-wave EM. The RF response metrics of transistors include small-signal S-Parameters. The simulated S-Parameters across a broad frequency range can be very useful to circuit designers for impedance matching at input and output ports in standard commercial software including ADS.

The linear system of the m^{th} Newton iteration of the quasi-static simulation retains the same form as Eqn. (5.52). However, because the time derivatives are now included in Eqns. (5.1), (5.4), (5.5), and (5.6), the Jacobian matrix is different. As an example, the middle diagonal element of the quasi-static Jacobian matrix is

$$\frac{\partial \text{ElCont}_i}{\partial (F-E_c)_j} = \left(\frac{\partial n_i^{t+1}}{\partial (F-E_c)_j} \frac{1}{\Delta t} + \frac{\partial C_i^{n,t+1}}{\partial (F-E_c)_j} \right) V_i + \sum_j \frac{\partial J_{n,ij}^{t+1}}{\partial (F-E_c)_j} A_{ij}. \tag{5.54}$$

The time step Δt is typically chosen according to the driving frequency of the electronic device simulation. A standard choice is $\Delta t = 1/(100 \times f_0)$, where f_0 is the fundamental frequency of the device simulation.

5.5.5 Full-Wave Solve

Full-wave and hot-electron effects are captured in an electronic device simulation by solving Eqns. (5.1) – (5.24). This solver can produce simulated S-Parameters which reflect the high frequency parasitics in an electronic device. These parasitics become important at high frequency. The

Newton linear system corresponding to a full-wave FKT device simulation is

$$\begin{bmatrix} \frac{\partial \mathrm{Pois}_i}{\partial \Phi_j} & 0 & 0 & \frac{\partial \mathrm{Pois}_i}{\partial (F-E_c)_j} & \frac{\partial \mathrm{Pois}_i}{\partial k_B T_j} & 0 \\ \frac{\partial \mathrm{Amp}_i}{\partial \Phi_j} & \frac{\partial \mathrm{Amp}_i}{\partial \mathrm{E}_{\mathrm{rot},j}} & \frac{\partial \mathrm{Amp}_i}{\partial H_j} & \frac{\partial \mathrm{Amp}_i}{\partial (F-E_c)_j} & \frac{\partial \mathrm{Amp}_i}{\partial k_B T_j} & 0 \\ 0 & \frac{\partial \mathrm{Far}_i}{\partial \mathrm{E}_{\mathrm{rot},j}} & \frac{\partial \mathrm{Far}_i}{\partial H_j} & 0 & 0 & 0 \\ \frac{\partial \mathrm{ElCont}_i}{\partial \Phi_j} & \frac{\partial \mathrm{ElCont}_i}{\partial \mathrm{E}_{\mathrm{rot},j}} & 0 & \frac{\partial \mathrm{ElCont}_i}{\partial (F-E_c)_j} & \frac{\partial \mathrm{ElCont}_i}{\partial k_B T_j} & \frac{\partial \mathrm{ElCont}_i}{\partial k_B T_{L,j}} \\ \frac{\partial \mathrm{EnCons}_i}{\partial \Phi_j} & \frac{\partial \mathrm{EnCons}_i}{\partial \mathrm{E}_{\mathrm{rot},j}} & 0 & \frac{\partial \mathrm{EnCons}_i}{\partial (F-E_c)_j} & \frac{\partial \mathrm{EnCons}_i}{\partial k_B T_j} & \frac{\partial \mathrm{EnCons}_i}{\partial k_B T_{L,j}} \\ 0 & 0 & 0 & \frac{\partial \mathrm{LattCons}_i}{\partial (F-E_c)_j} & \frac{\partial \mathrm{LattCons}_i}{\partial k_B T_j} & \frac{\partial \mathrm{LattCons}_i}{\partial k_B T_{L,j}} \end{bmatrix}$$

$$\times \begin{bmatrix} \Delta\Phi_j \\ \Delta \mathrm{E}_{\mathrm{rot},j} \\ \Delta H_j \\ \Delta (F-E_c)_j \\ \Delta k_B T_j \\ \Delta k_B T_{L,j} \end{bmatrix} = - \begin{bmatrix} \mathrm{Pois}_i \\ \mathrm{Amp}_i \\ \mathrm{Far}_i \\ \mathrm{ElCont}_i \\ \mathrm{EnCons}_i \\ \mathrm{LattCons}_i \end{bmatrix} \tag{5.55}$$

Full-wave simulations use the same choice of time step as quasi-static simulations. An example of the Jacobian matrix calculated in a full-wave simulation of the GaN MESFET is illustrated in Figure 5.14. The lattice heating equation is not included in this simulation.

Figure 5.10 presents the RF S-Parameters calculated with the quasi-static (dashed lines) and full-wave (solid lines) solvers. The frequency sweep is 1 to 200 GHz. The full-wave solver captures different small-signal parasitics in the MESFET. Proper treatment of the high frequency parasitics is paramount for accurate circuit model extraction.

5.6 Mesh Convergence of the Discrete FKT Equations

The mesh convergence of the discrete FKT device simulation equations is investigated. This numerical characteristic provides insight into how the errors of the device simulation solution variables converge with mesh refinement. Simulation of large electronic devices requires meshing strategies in order to yield accurate results without insurmountable computational demands. Fol-

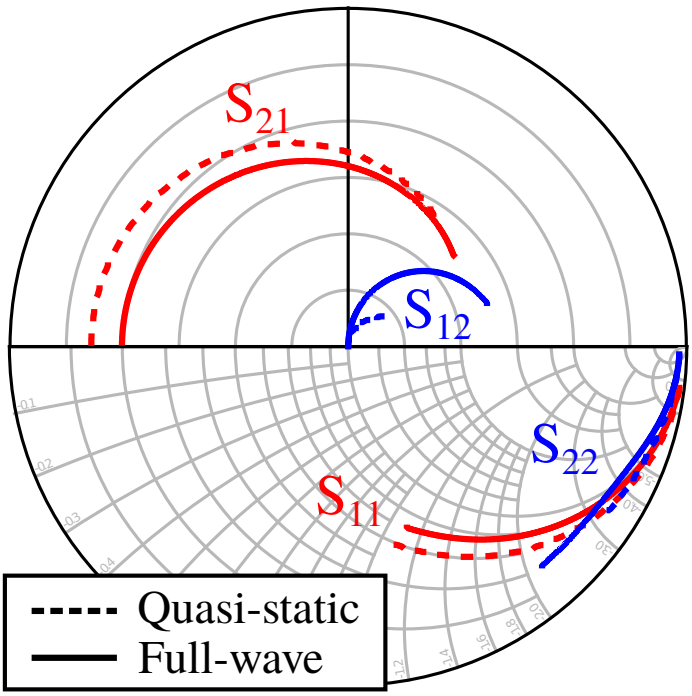


Figure 5.10: Simulated S-Parameters of the GaN MESFET. The operating point of the device is a gate-source bias of $-0.5\,V$ and a drain-source bias of $4\,V$. The top portion of the figure is a polar plot and the bottom half is a Smith chart plot. The dashed lines represent the S-Parameters calculated with the quasi-static solver and the solid lines are the S-Parameters calculated with the full-wave solver. The full-wave simulation produces significantly different S-Parameters than the quasi-static solver.

lowing Section 2.3.2, convergence of the discrete FKT equations is quantified by evaluating the i^{th} relative L^2 error

$$\varepsilon_{ik} = \sqrt{\frac{\sum_j w_j \left[\tilde{u}_{ij,k} - u_{ij,k}\right]^2}{\sum_j w_j u_{ij,k}^2}}, \tag{5.56}$$

on a series of meshes. Here, $\tilde{u}_{ij,k}$ and $u_{ij,k}$ are the i^{th} numerical and analytic solutions in the j^{th} element of the k^{th} mesh, respectively. The weights w_j are specified by the specific error calculation. The global error of all solution variables in the device simulation is calculated as

$$\varepsilon_k = \sqrt{\sum_i \varepsilon_{ik}^2}. \tag{5.57}$$

The order of convergence of the FKT equations is determined by the relation

$$|\varepsilon_k| < C N_k^{-p_k}, \tag{5.58}$$

N_k	ε_k	p_k
1041	0.12896775	3.00
1646	0.03903999	2.22
3034	0.01340268	1.38
6283	0.00572172	1.28

Table 5.2: Convergence of the static FKT solution variables in the GaN MESFET example with a gate-source bias of $-0.5\,V$ and drain-source bias of $4\,V$.

N_k	ε_k	$\Delta\varepsilon_k$
2277	0.05570511	
2893	0.02115591	-0.03454920
4513	0.00024652	-0.02090938
5535	0.00002885	-0.00021767

Table 5.3: Convergence of the I-V family of the GaN HEMT example calculated with the static FKT solver.

where N_k is the number of DOF in the mesh, p_k is the approximate order of convergence, and C is a constant.

Two numerical examples are presented to provide insight into the convergence of the discrete FKT equations. The first is a static simulation of the GaN MESFET. After a quiescent bias of $V_{GS} = -0.5\,V$ and $V_{DS} = 4\,V$ is calculated with the static solver, the global solution variables are saved on a series of five meshes. The first and last meshes are the most coarse and dense, respectively. Solutions on the dense mesh are used as the "analytic solutions" $u_{ij,k}$ in Eqn. (5.56).

Calculating the global solution variable error ε_k requires interpolation of the nodal solution variables on the dense mesh to the nodes of the series of meshes. To this end, linear interpolation is used for its simplicity and efficiency in the post-processing. After the dense mesh solution variables are interpolated to the series of meshes, the global solution variable error is calculated with numerical quadrature. Therefore, the weights w_j in Eqn. (5.56) represent the volumes of the quadrature simplices.

Table 5.2 reports the integrated errors for each mesh in the series. The first column lists the DOF corresponding to each mesh. The second and third columns list the solution variable errors

ε_k and the approximate order of convergence p_k for each mesh in the series. According to these results, the FKT device simulator exhibits approximately first-order convergence, i.e., $p_k = 1$. This is due to the SG discretization of the particle fluxes. The crux of the flux discretization is the approximation of the particle fluxes as spatially constant across the primary edges of the DV mesh. This approximation implies first-order mesh convergence.

A pragmatic example of FKT device simulation convergence is an investigation of the I-V family calculated by static simulation of a GaN HEMT on a series of meshes. The GaN HEMT is discussed in detail in Chapter 6. The device's currents are an important output of the device simulator. Therefore, it is a useful study to determine how the simulated currents are affected by mesh refinement. The error calculation of this example requires evaluation of Eqn. (5.56) at each drain-source bias. The I-V curves on the dense mesh are used as the "analytic" solutions. Table 5.3 lists the DOF, global errors in the I-V curves ε_k, and the corresponding change in the errors $\Delta\varepsilon_k$ compared to the previous mesh in the series. Figure 5.11 presents the calculated I-V curves on the series of four meshes. The change in the device's response becomes negligible after the third mesh

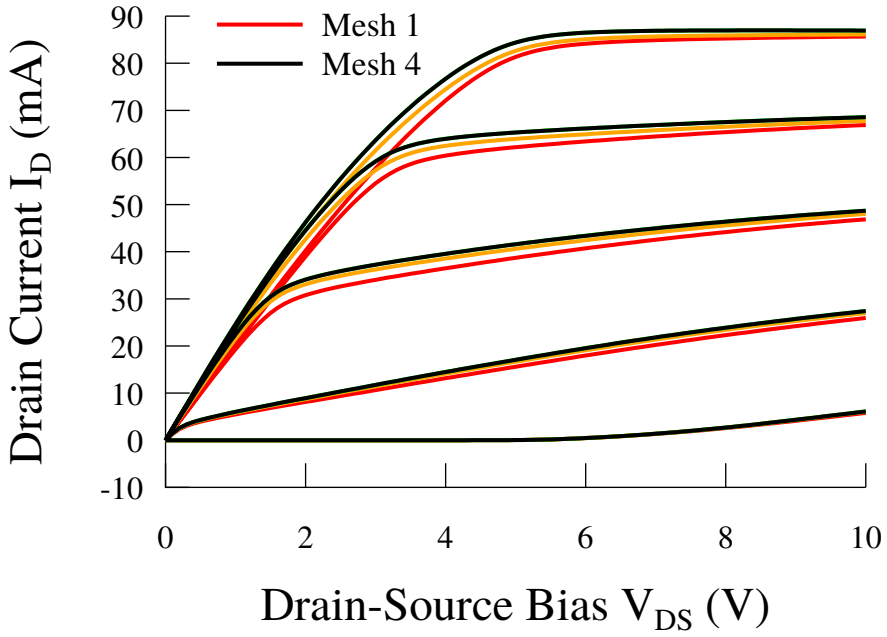


Figure 5.11: The I-V families of the GaN HEMT calculated with the static solver on a series of four meshes.

refinement. Furthermore, only the linear region of the device's I-V family is moderately affected by the mesh refinement. This type of mesh refinement may not be necessary when simulating the RF response at the peak transconductance of the saturated region of the I-V family.

5.7 Stability of the Discrete FKT Equations

Stability of nonlinear differential equations is a deep and rich subject in systems theory and engineering [54]. One part of stability theory is the analysis of equilibrium points. The analysis starts with a given equilibrium point of the autonomous system $\frac{dx_i(t)}{dt} = F_i(x_i(t))$, i.e., $F_i = 0$. The system of nonlinear differential equations are said to be stable if for a given perturbation of the equilibrium, the solutions return to the equilibrium point. A detailed discussion on stability of nonlinear equations was presented in Section 2.3.3.

The following examples analyze the stability of the discrete FKT equations. The equilibrium

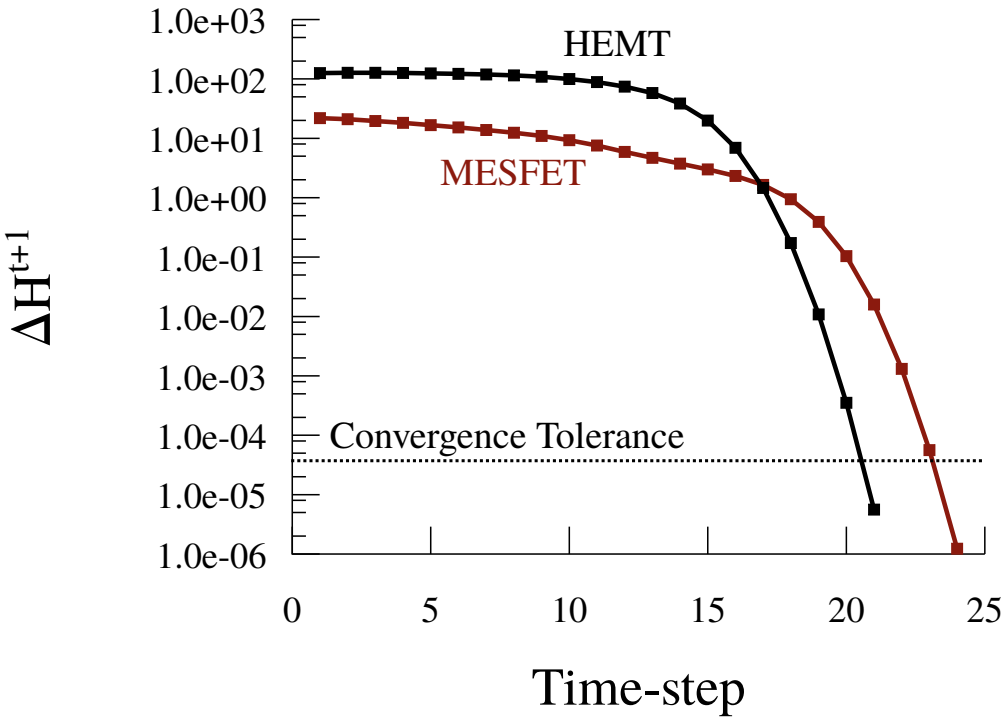


Figure 5.12: Deviation of the numerical enthalpy, Eqn. (5.60), for the GaN MESFET (red line) and HEMT (black line) examples. Both equilibria are perturbed by displacing two randomly chosen electron gases from their thermal equilibrium states.

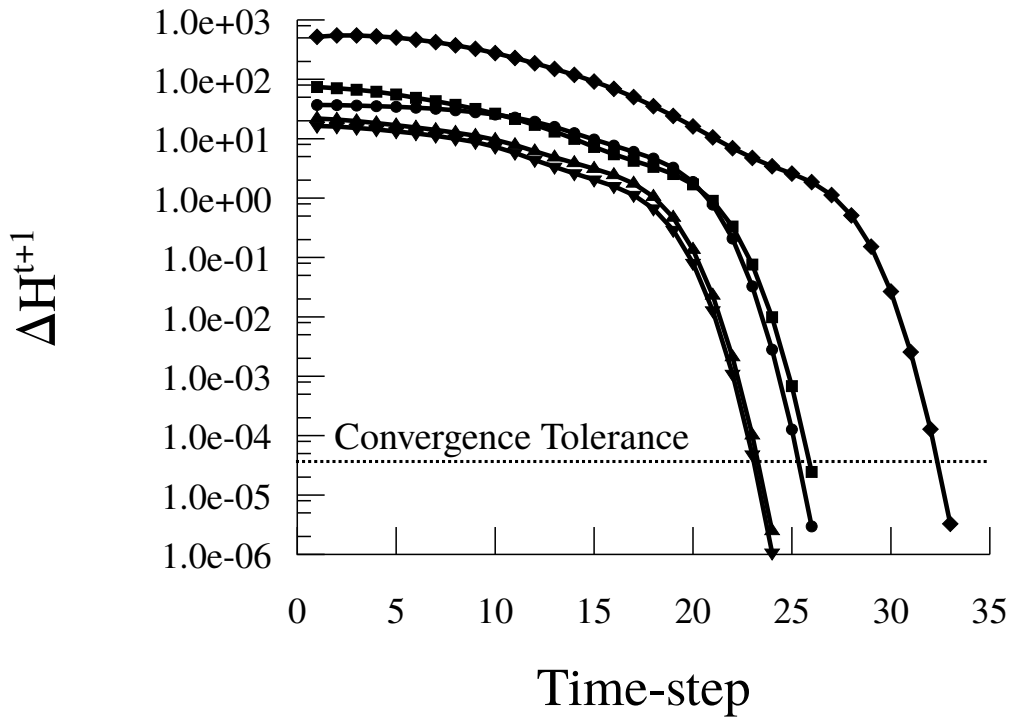


Figure 5.13: Five independent enthalpy analyses of the GaN MESFET example. Each perturbation moves three randomly chosen electron gases from their thermal equilibrium states. No perturbation was found to be unstable. The deviation from the thermal equilibrium enthalpy decayed exponentially near the convergence tolerance.

points are the solutions of the static solver at thermal equilibrium, i.e., no external biases. After the thermal equilibrium of the specific device example is computed, the total numerical enthalpy of the system is calculated as

$$H \equiv \sum_i \left[E_{n,i} + k_B T_i n_i\right] V_i. \tag{5.59}$$

The enthalpy is calculated from the electron kinetic energy density $E_{n,i}$, electron temperature $k_B T_i$ and electron density n_i defined in the DUT. The summation represents integration over the elements of the DV mesh. Perturbation amounts to displacing the FKT solutions from their thermal equilibrium states. For these purposes, two electron gases are randomly chosen and their relative Fermi levels and temperatures are perturbed. The perturbed solutions are used as initial conditions to the quasi-static solver. A transient simulation with no external biases is solved until the discrete residuals reach the global numerical tolerance of 10^{-9}. The deviation of the total numerical

enthalpy is quantified as

$$\Delta H^{t+1} = \left| H^{t+1} - H_{\text{equilib.}} \right|, \tag{5.60}$$

where H^{t+1} is calculated at each time step of the quasi-static solver and $H_{\text{equilib.}}$ is the numerical enthalpy of the thermal equilibrium solutions.

Two GaN device examples, a MESFET and a HEMT, are chosen to numerically demonstrate the stability of the discrete FKT equations. The GaN HEMT is discussed in detail in Chapter 6. Both device examples are simulated as closed systems, i.e., lattice heating is not included. Two electron gases are randomly chosen and perturbed from their thermal equilibrium states. Figure 5.12 presents the deviation of the enthalpy ΔH^{t+1} for both device examples. The red line represents the enthalpy deviation versus time for the GaN MESFET example and the black line presents the same for the GaN HEMT example. An approximate convergence line is illustrated in the figure. Figure 5.13 presents five more enthalpy deviation simulations of the MESFET example. Each perturbation moved an additional random electron gas from its equilibrium state. No perturbation simulation is found to be unstable. Furthermore, each enthalpy deviation approaches an exponential rate of decay near the convergence tolerance. This provides insight into the stability and robustness of the FKT device simulation equations.

5.8 Advanced Simulation Techniques

Simulating large electronic devices requires high computational efforts. For example, the DOF corresponding to a four-finger transistor could quadruple compared to a single gate transistor. Furthermore, devices with exotic materials like gallium oxide (β-Ga_2O_3) could require many Fermi distributions associated to band-structure valleys which would further increase the DOF. To accommodate the simulation of large and/or complex electronic devices, the TCAD tool must include advanced techniques to help alleviate some of the computational burden.

5.8.1 The Linear Solve

The most time consuming portion of a simulation of a large electronic device is the Newton linear solve. The Jacobian of an FKT device simulation with an arbitrary unstructured DV mesh is a sparse matrix. The sparsity comes from the fact that the discretization stencils only depend on nearest neighbors in the DV mesh. Figure 5.14 illustrates an example of an FKT device simula-

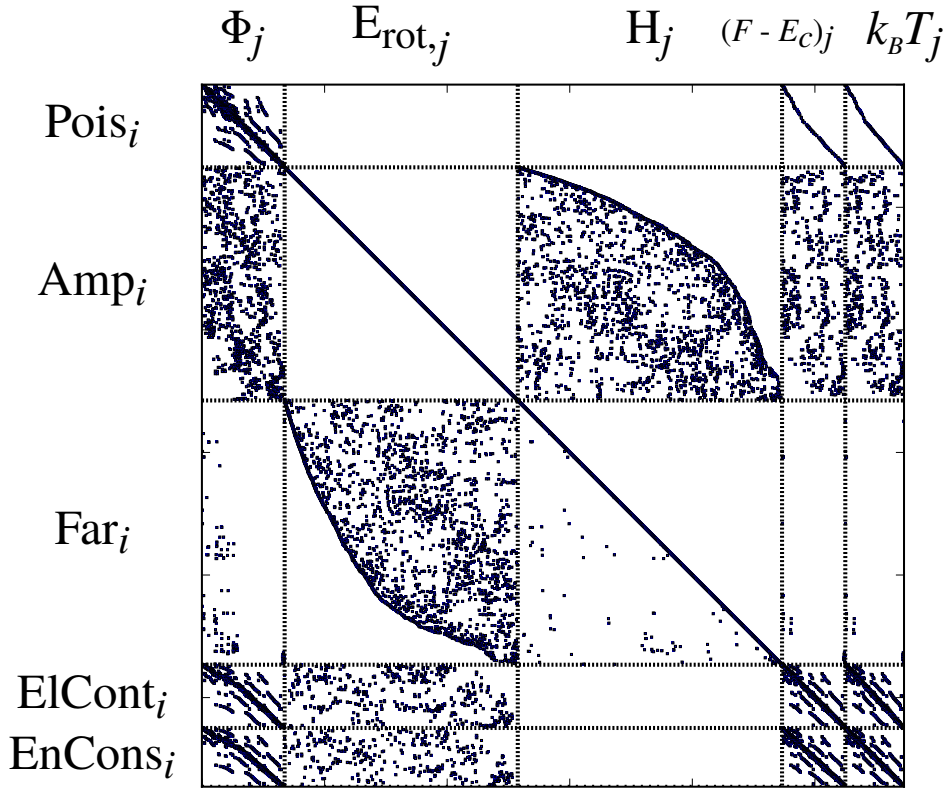


Figure 5.14: An example of the Jacobian matrix calculated in the FKT device simulator. The Jacobian matrix corresponds to a full-wave FKT solve. Dotted lines mark the boundaries of the residual rows and solution variable columns in the Jacobian matrix. The rows of the Jacobian correspond to Poisson's equation, Ampere's equation, Faraday's equation, electron continuity, and energy conservation residuals, respectively.

tion Jacobian. This Jacobian matrix corresponds to a full-wave simulation of the GaN MESFET first presented in Section 5.5.1 and illustrated in Figure 5.6. Lattice heating was not included in this FKT device simulation. The DV mesh of the MESFET is a standard example of a general device simulation mesh. Therefore, this Jacobian can be considered a standard example of an FKT device simulation Jacobian matrix. The sub-matrix blocks in the Jacobian matrix in Figure 5.14

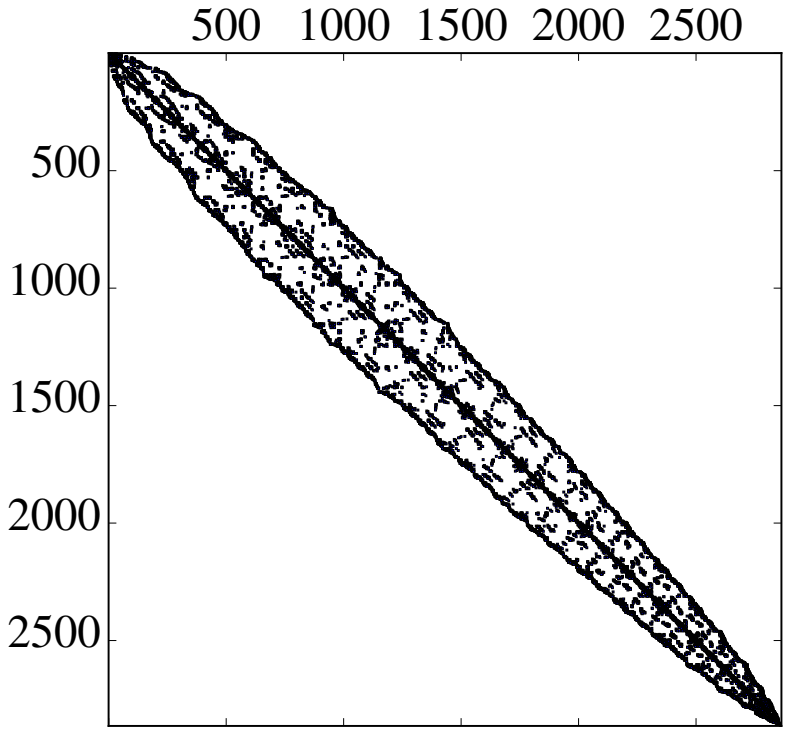


Figure 5.15: An example of the Jacobian matrix after RCM re-ordering.

correspond to the five equations of the FKT device simulation. The sub-matrix blocks are outlined with dashed lines. The governing equations of the full-wave simulation are Poisson's equation, Ampere's equation, Faraday's equation, electron continuity, and energy conservation. Located in the top left corner is the Poisson sub-matrix block corresponding to the elements $\frac{\partial \text{Pois}_i}{\partial \Phi_j}$.

One note about the Jacobian matrix in Figure 5.14 is about the Faraday sub-matrix block $\frac{\partial \text{Far}_i}{\partial \Phi_j}$. According to Eqn. (5.55), this sub-matrix should be zero. However, there are non-zero elements in this matrix block in Figure 5.14. These non-zero elements correspond to the full-wave voltage port BC described in Section 5.4.4. The voltage port BC requires magnetic field solution variables in the plane of the boundary. These in-plane magnetic field variables are governed by Ampere's law and their residuals are located at the magnetic field solution variable rows. This dependence on Ampere's law gives rise to the non-zero Jacobian elements corresponding to the electric potential.

The linear solve computation time can be reduced by re-ordering the solution variables to re-

duce the bandwidth of the Jacobian matrix. The bandwidth of a matrix is defined as the largest distance between the diagonal element and an off-diagonal element in a row. The overall bandwidth is the maximum distance of all rows. A banded matrix is one with a "reasonably small" bandwidth. A standard method to reduce the matrix bandwidth is the reverse Cuthill-McKee algorithm (RCM) [90]. RCM aims to permute the sparse matrix into a banded matrix form with as small of a bandwidth as possible. An implementation of the RCM algorithm is applied to the Jacobian matrix of the GaN MESFET example. The resulting banded matrix is illustrated in Figure 5.15. The bandwidth of the RCM re-ordered Jacobian matrix is 296.

5.8.2 Second-Order SG Discretization

Reduction of the Newton linear system size is another approach to alleviating some of the computational burden of simulating large devices. This could be accomplished by reducing the number of mesh elements in the simulation. Coarsening the mesh, however, could have adverse affects including inaccuracies in the simulated results and possible convergence issues.

A higher-order numerical framework would allow mesh coarsening without infringing on the integrity of the simulated results. In Section 5.6, the mesh convergence of the FKT device simulator was presented and discussed. The FKT equations exhibit first-order mesh convergence due to the SG discretization of the particle fluxes. Fluxes and electric field solution variables are approximated as spatially constant along the primary edges of the DV mesh which yields a first-order method. This section expands upon the SG discretization to develop a second-order method. In the following, the first-order particle flux device simulations are termed SG1 and the simulations with the second-order fluxes are termed SG2.

5.8.2.1 Derivation of the SG2 Particle Flux

The new discretization of the FKT particle flux is derived and discussed. This discretization is an extension of the original SG technique [77] which was presented in Section 4.3.2 and is derived in

Appendix A.3. The derivation starts with the phenomenological DD equation

$$J_n(r) = a_1 n(r) + a_2 \frac{dn(r)}{dr}, \tag{5.61}$$

where a_1 and a_2 are constant on a primary edge of the DV mesh. The flux $J_n(r)$ is approximated with a second-order polynomial which yields

$$J_2(r-r_0)^2 + J_1(r-r_0) + J_0 = a_1 n(r) + \frac{dn(r)}{dr}. \tag{5.62}$$

Here, J_0, J_1, and J_2 are coefficients of the polynomial. The original SG discretization technique solved an inhomogeneous linear differential equation for the electron density. Then, BCs were applied to yield the discrete form of the particle flux. This discretization follows the same approach. Eqn. (5.62) is a linear first-order differential equation with a spatially dependent inhomogeneity.

The solution of Eqn. (5.62) with the BCs $n(r=r_0) = n_0$, $n(r=r_1) = n_1$, $\left.\frac{dn}{dr}\right|_{r_0} = \frac{dn_0}{dx}$, and $\left.\frac{dn}{dr}\right|_{r_1} = \frac{dn_1}{dx}$ for $J_n \equiv J_2(r-r_0)^2 + J_1(r-r_0) + J_0$ is

$$J_n = \frac{a_2}{L}\left[B_1(\xi) n_0 + L B_3(\xi)\frac{dn_0}{dx} - B_2(\xi) n_1 - L B_4(\xi)\frac{dn_1}{dx}\right]. \tag{5.63}$$

Here, L is the length of the primary edge, the argument of the B functions is $\xi = \frac{a_1}{a_2}L$, and the B functions are

$$B_1(x) = \frac{x(4 + e^x(x-4) + 3x)}{4(2 + e^x(x-2) + x)}, \tag{5.64}$$

$$B_2(x) = \frac{x(4 + x + e^x(3x-4))}{4(2 + e^x(x-2) + x)}, \tag{5.65}$$

$$B_3(x) = -\frac{4 + e^x(x-4) + x(x+3)}{4(2 + e^x(x-2) + x)}, \tag{5.66}$$

$$B_4(x) = -\frac{4 + x - e^x(4 + x(x-3))}{4(2 + e^x(x-2) + x)}. \tag{5.67}$$

Eqn. (5.63) is the second-order discretization of the DD flux. In this derivation, the flux is approximated with a second-order spatial polynomial. Because a second-order spatial polynomial exhibits third-order mesh convergence, the overall discretization should be third-order. However, the FKT particle will be second-order due to approximations made to the coefficients a_1 and a_2. The density gradient terms $\frac{dn_m}{dx}$ are approximated with central differences across the primary edges.

The discrete flux, Eqn. (5.63), is similar to the original discrete particle flux, Eqn. (28), derived in the appendix. There are two sets of terms which represent particles flowing from node r_0 and node r_1 of the primary edge. This fact allows the heat flow algorithm to be applied to the second-order discrete flux. Finally, the discretization of the general DD flux is applied to the generalized Einstein FKT particle flux with the substitutions

$$a_1 \rightarrow -qD_{jk}\left[\mathrm{E}_j + \Delta\mathrm{Ein}_{n,jk}\right], \tag{5.68}$$

and

$$a_2 \rightarrow -qD_{jk}\left(\mathrm{Ein}_{n,jk}\right)_{\mathrm{ave}}. \tag{5.69}$$

Because the second-order discretization is arranged in a similar manner to the first-order discretization, the argument of the B functions is the same as the argument of the Bernoulli functions in the generalized Einstein flux. This argument is

$$\xi_{n,jk} = \frac{1}{\left(\mathrm{Ein}_{n,jk}\right)_{\mathrm{ave}}}\left[\mathrm{E}_j + \Delta\mathrm{Ein}_{n,jk}\right]L_j. \tag{5.70}$$

In the following, validation and simulation results are presented with the SG2 particle fluxes.

5.8.2.2 Validation of the Equations: Particle Flux Reconstruction

Validation of the second-order particle flux discretization in the FKT device simulator is performed by prescribing an analytic spatial profile to solution variables and calculating the discrete fluxes on primary edges of the DV mesh. The discrete particle flux is compared to the analytic particle flux to quantify a reconstruction error. The L^2 error described in Section 5.6 is used.

An example of a flux reconstruction with an analytic solution variable profile is illustrated in Figure 5.16. The left of the picture presents the structured cube mesh on which a flux reconstruction error is calculated. Also shown in the figure is the analytic electric potential solution variable profile. The reconstructed FKT particle fluxes are presented on the right of Figure 5.16. The vector plot is an interpolation of the nodal vector data loaded into a visualization software. Scalar vector projections onto primary edges of the DV mesh are compared to analytic vector projections

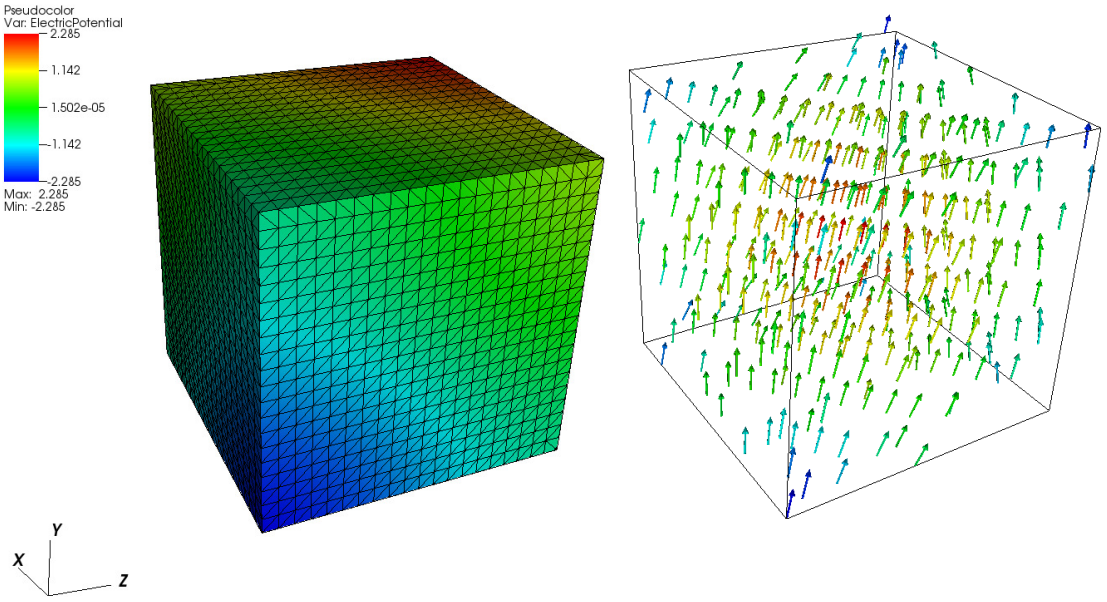


Figure 5.16: Flux reconstruction on a $10\times10\times10$ structured cube mesh. An analytic profile is prescribed for the electric potential, relative electron Fermi level, and electron temperature solution variables. The FKT particle fluxes are computed on all primary edges of the DV mesh and compared to analytic fluxes to quantify a reconstruction error.

to calculate the global L^2 error. This process is repeated on a series of meshes to quantify the mesh convergence of the FKT particle flux reconstruction.

Order analysis results are presented in Figure 5.17 for the SG1 and SG2 FKT particle flux reconstructions. Here, the L^2 errors of the reconstructed particle fluxes are presented versus the maximum primary edge length of the DV meshes. The errors corresponding to the SG1 particle fluxes are plotted in red and the errors corresponding to the SG2 particle fluxes are plotted in blue. As expected, the SG1 particle flux reconstructions exhibit first-order convergence. This result also demonstrates that the SG2 particle fluxes derived in this section exhibit second-order mesh convergence. As described in Section 5.8.2.1, the particle flux is approximated with a second-order polynomial and discretized by integrating the inhomogeneous first-order linear differential equation. Even though the particle flux is approximated with a third-order mesh convergent polynomial, the overall particle flux is second-order due to approximations of the flux coefficients. The flux coefficient approximations are described in Chapter 4 and Appendix A.3.

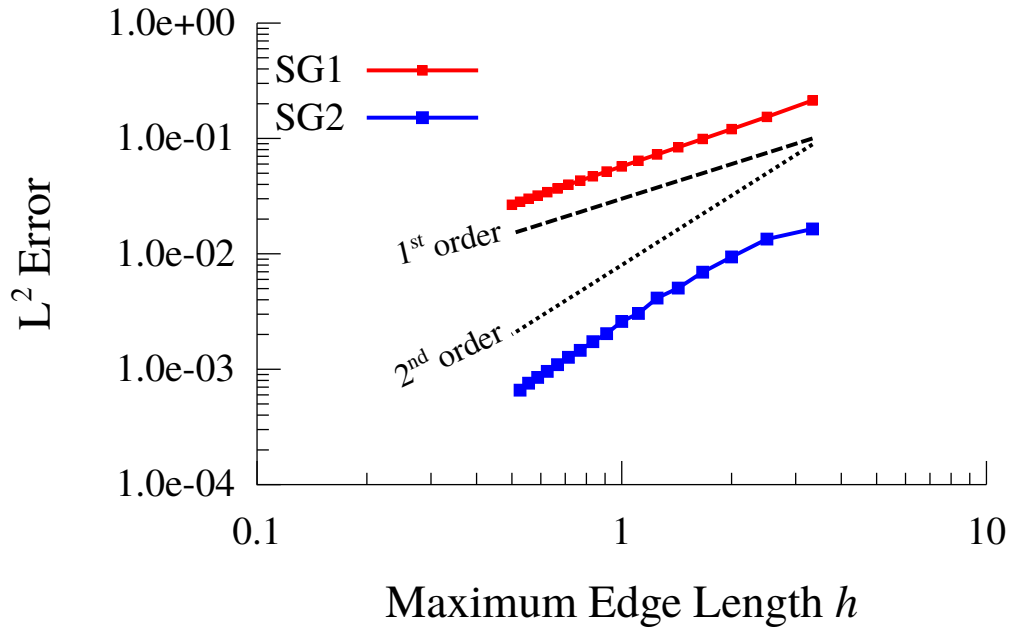


Figure 5.17: Order analysis of the FKT particle flux reconstruction. The errors corresponding to the SG1 particle fluxes are illustrated in red and the SG2 errors are illustrated in blue. As expected, the SG1 particle flux reconstructions exhibit first-order convergence and the SG2 particle flux reconstructions exhibit second-order convergence.

5.8.2.3 Four Particle System

Simulation of electronic devices with the SG2 FKT particle flux discretization begins with a simple 1D mesh with four nodes. The two boundary nodes are designated as Ohmic contacts. This simulation has two interior DOF. A dopant discontinuity exists at the middle of the device. The first two nodes are doped with $10^{12}\,\mathrm{cm}^{-3}$ donors and the second two are doped with $10^{14}\,\mathrm{cm}^{-3}$ donors. Figure 5.18 presents the I-V curves computed with the SG1 and SG2 FKT device simulators. The device currents calculated with the SG1 simulator are plotted in black and the device currents calculated with the SG2 simulator are plotted in red. The SG2 device simulator failed to converge at approximately $V_a = 0.2\,\mathrm{V}$. Furthermore, the slope of the anode current changes sign before the simulator crashes.

The previous results hint at a stability issue of the SG2 particle flux. To investigate this issue, the phase portrait of the two DOF system is calculated for stability analysis. Figure 5.19 presents the direction field drawn with black vectors and four trajectories drawn with blue lines. The x-axis

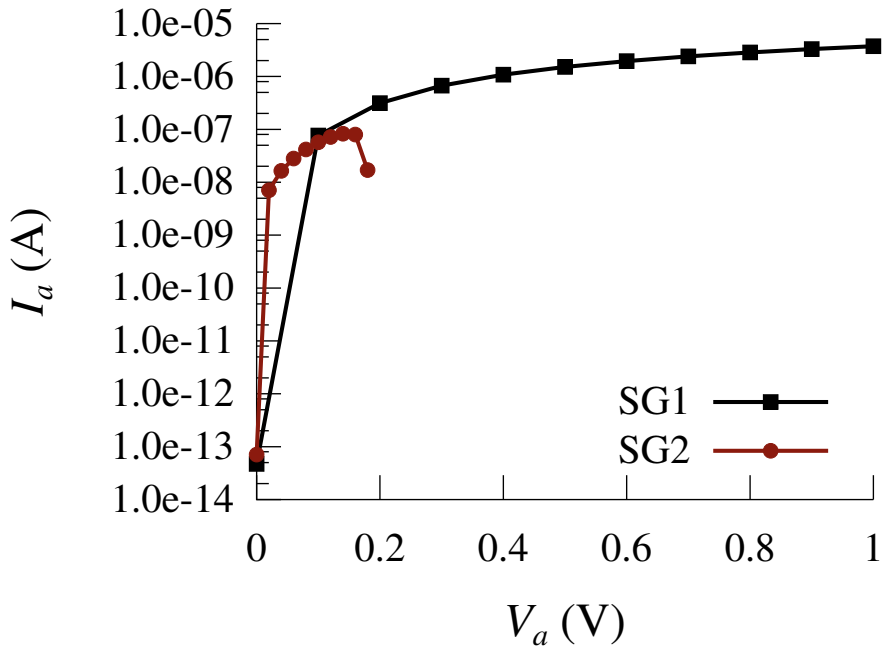


Figure 5.18: The simulated static I-V of the four particle system with two interior DOF. SG1 device simulation currents are plotted in black and SG2 currents are plotted in red.

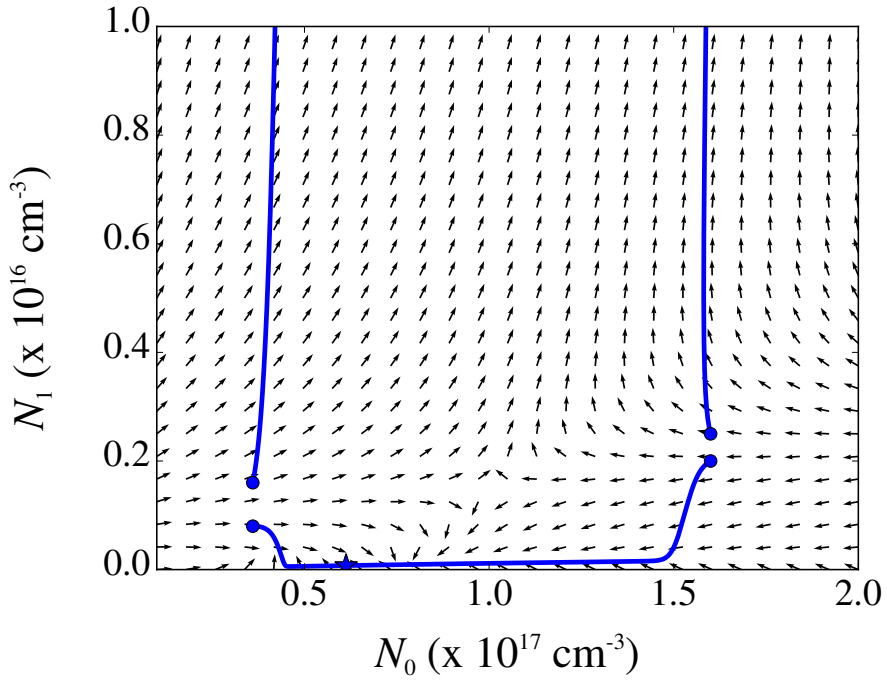


Figure 5.19: The phase portrait of the four particle, two DOF system. The direction field is plotted with black vectors and four trajectories are plotted with blue lines. Initial conditions of the trajectories are plotted with blue circles. There are three equilibria of the system of nonlinear equations, two of which are unstable.

represents the density N_0 of the first particle and the y-axis represents the second particle's density N_1. The trajectories represent the integrated SG2 nonlinear equations with four different initial conditions drawn with blue circles. There are several equilibria of the SG2 system of equations. The direction field indicates that two of the equilibria are unstable. These multiple and unstable equilibria are a known issue of higher-order methods in the computational fluid dynamics (CFD) community.

5.8.2.4 Flux Limiters

A well known mathematical result from the CFD community is Godunov's theorem. It states that linear numerical schemes for solving PDEs, having the property of not generating new extrema, can be at most first-order accurate [91]. The converse of this states that higher-order accurate numerical methods will generate new extrema. This fact is numerically demonstrated with the SG2 phase portrait in Figure 5.19. New extrema are generated by increasing the particle flux order from first to second. These equations are no longer monotone and therefore not unique and are unstable.

A substantial effort in the CFD community was devoted to the development of high-resolution schemes for fluid-flow simulations. Specifically, Harten developed a measure of discrete variation in the solution fields which is called total variational diminishing (TVD) [92]. A numerical technique called a "flux limiter" was developed in order to overcome the restrictions imposed by Godunov's theorem and enforce the monotonicity criterion set forth by Harten. Several versions of flux limiters exist in the literature with a concise summary provided in [93].

A flux limiter attempts to diminish spurious oscillations in solutions near discontinuities and/or sharp changes in the solution domain. The oscillations are a consequence of the new extrema generated by the higher-order method. In terms of the FKT device simulator, this includes sharp changes in electron density at dopant fronts and inflection points in the electric potential near metal contacts. The particle fluxes near these device features are restricted to enforce monotonicity. A smooth variation between low and high order is enforced based on numerical criteria. The

functional form of this variation is

$$J = J_{\text{low}} - \phi(r)\left(J_{\text{low}} - J_{\text{high}}\right). \tag{5.71}$$

Here, J_{low} represents the lower-order, monotonic flux and J_{high} represents the higher-order flux. The function $\phi(r)$ is the flux limiter. If $\phi(r) = 0$, then the overall flux is first-order. A second-order flux results from $\phi(r) = 1$. Standard practice enforces a specific functional form for the flux limiter $\phi(r)$ which smoothly varies from 0 to 1 [93]. An example of a flux limiter is the van Albada form

$$\phi(r) = \frac{r^2 + r}{r^2 + 1}. \tag{5.72}$$

Here, the term r represents successive gradients of electron density on a 1D mesh, or

$$r_i = \frac{N_i - N_{i-1}}{N_{i+1} - N_i}. \tag{5.73}$$

5.8.2.5 Higher-Order Simulation of a 1D Diode

The flux limiter described in the previous section is applied to the SG2 particle fluxes in the FKT device simulator. The J_{low} term of Eqn. (5.71) represents the SG1 particle flux and J_{high} represents the SG2 particle flux. The SG2 particle flux is given in Eqn. (5.63).

A 1D n^+-n diode is simulated with the SG2 FKT device simulator. The dopant density of the n^+ region is $3 \times 10^{17}\,\text{cm}^{-3}$ donors. The second region is lightly doped with $10^{12}\,\text{cm}^{-3}$ donors. Figure 5.20 presents an illustration of the thermal equilibrium band diagram calculated with the SG1 and SG2 device simulators. The conduction and valence bands are drawn in black and the electron Fermi level F is drawn in red. In this figure, the SG1 results are illustrated with solid lines and the SG2 results are illustrated with squares. This result demonstrates that the SG2 equations properly reproduce the thermal equilibrium band diagram.

The simulated I-V curves of the 1D diode are presented in Figure 5.21. Here, the SG1 results are plotted in black and the SG2 results are plotted in red. The SG2 current is similar to the current calculated with the SG1 device simulator. The difference resides in the fact that the SG2 device simulation is more accurate and therefore provides different results.

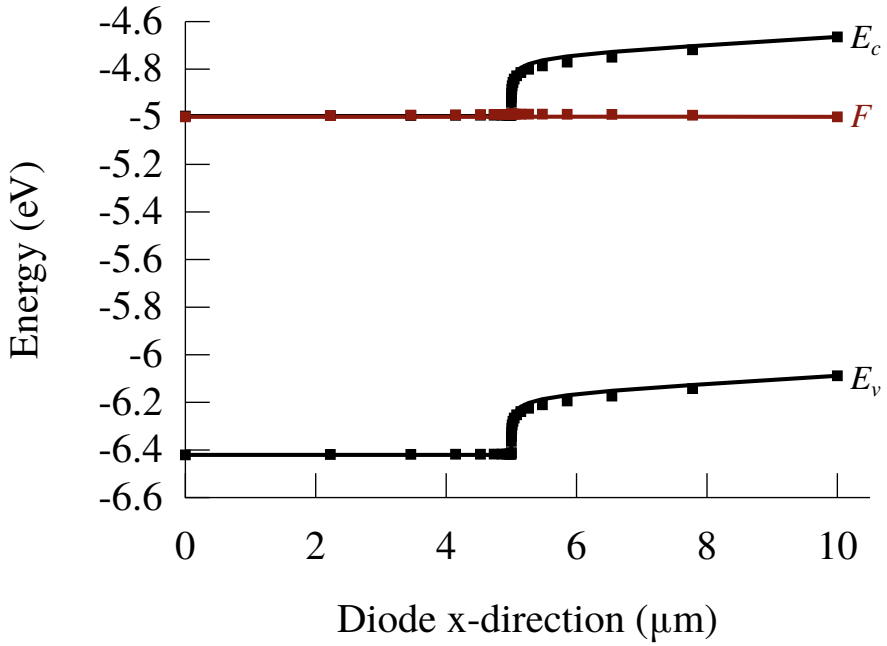


Figure 5.20: The simulated thermal equilibrium band diagram of the 1D diode. The SG1 results are plotted with solid lines and the results computed with the SG2 solver are plotted with squares.

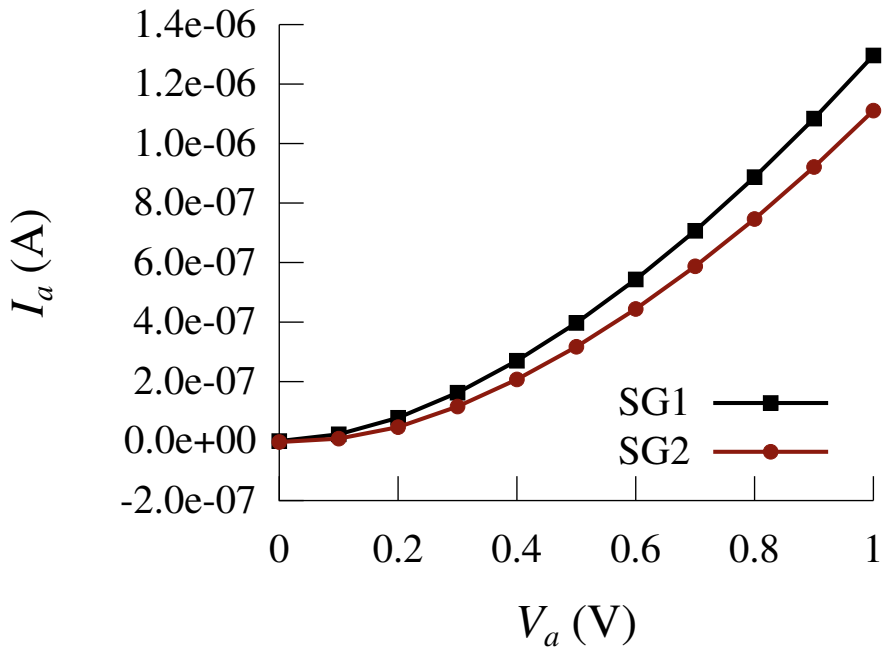


Figure 5.21: The simulated static I-V of the 1D diode. SG1 device simulation currents are plotted in black and SG2 currents are plotted in red.

Table 5.4 presents the error in the FKT solution variables versus the number of DOF. In a manner similar to the mesh convergence analysis in Section 5.6, the solutions on each mesh are

compared to the solutions on the most dense mesh. The first column of Table 5.4 lists the DOF for each mesh. The second through fifth columns present the relative error of the SG1 solutions, the approximate order of convergence of the SG1 errors, the relative error of the SG2 solutions, and the approximate order of convergence of the SG2 errors, respectively. As indicated by the

DOF	ε_k (SG1)	p_k (SG1)	ε_k (SG2)	p_k (SG2)
30	2.1e-1	1.2	3.7e-1	1.8
60	1.4e-1	1.2	2.7e-1	1.8
120	7.7e-2	1.2	1.1e-1	1.8
180	3.5e-2	1.2	2.4e-2	1.8

Table 5.4: Solution variable errors and approximate order of convergence of the SG1 and SG2 FKT device simulations of the 1D diode.

results in Table 5.4, the FKT device simulator with the SG2 particle fluxes exhibits higher-order convergence than the FKT device simulator with the SG1 particle fluxes. The order of convergence is not $p = 2.0$, however, because the flux limiter varies between 0 and 1 across the DV mesh. This yields particle fluxes between first- and second-order and the overall order of convergence for the 1D diode example is $p_k = 1.8$.

5.9 Conclusion

A detailed description of the numerical solution of the discrete FKT device simulation equations was presented in this chapter. The discrete system of nonlinear equations was summarized in Section 5.2, a detailed description of the FKT solution variables was presented in Section 5.2.1, and the Newton's method approach to solving the nonlinear equations was presented in Section 5.3. BCs used for simulating electronic devices were presented and discussed in Section 5.4. An overview of the semiconductor device simulation work-flow was presented in the next section. This included examples of the Jacobian matrices for static, quasi-static, and full-wave FKT device simulations. The mesh convergence and the stability of the FKT device simulation equations were then analyzed in Sections 5.6 and 5.7, respectively. The results presented in these sections provided quantitative evidence that the FKT device solver is first-order convergent and numerically robust

and stable for all device examples that were investigated. Finally, advanced simulation techniques were reported in the final section. Some aspects of the Newton linear solve were presented in Section 5.8.1. An extension of the SG discretization scheme was then presented which increased the mesh convergence to nearly second-order. Section 5.8.2 provided derivation and implementation details of the second-order SG scheme.

CHAPTER 6

RF SIMULATIONS OF GAN HEMT TECHNOLOGY

6.1 Introduction

Physics based modeling of GaN HEMT PAs offers a promising alternative to high frequency and large-signal measurements. However, the physics based TCAD tool must accurately capture the physical processes in the device to provide useful data to RF circuit designers. The salient features required for accurate large-signal RF simulations include electronic band-structure, hot-electron effects, self-heating, scattering, trapping, and full-wave EM.

PA devices with GaN technology are scaled to deep sub-micron gate lengths. This allows high frequency operation of the TR module. However, because the devices are scaled in physical size but the applied voltages cannot be scaled in a similar way, the GaN transistors are subject to large electric fields. This causes the electrons in the channel of the device to heat up substantially and begin to emit phonons. This is the hot-electron effect. Figure 6.1 illustrates the electron drift velocity as a function of the applied electric field calculated with MC [94, 95]. The electron velocity

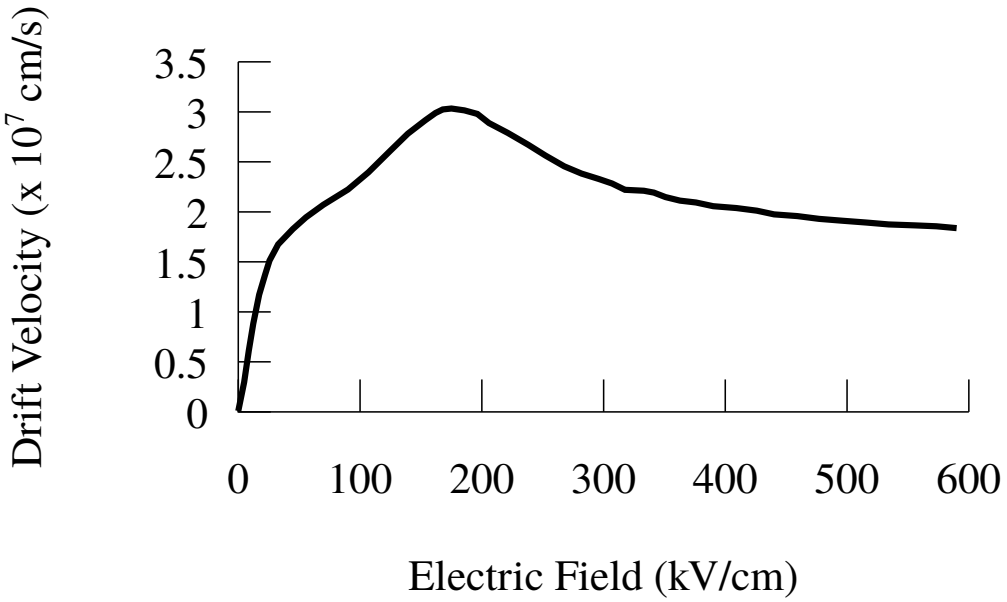


Figure 6.1: Electron drift velocity vs. electric field of bulk GaN calculated with MC.

peaks at approximately 200 kV/cm then begins to decrease in value. This is velocity overshoot and saturation — a result of the hot-electron effect. It is critical that the TCAD simulation tool captures hot-electron effects in order to produce accurate data.

The hot-electron effect can be mathematically explained with the second moment of the BTE. The Joule's heat source term in the energy conservation equation drives the electron temperature and is eventually balanced by the collision operator sink term. Because the large electric fields create hot electrons in the device which in turn interact with the lattice, it is necessary to include lattice heating in the device simulator. The lattice energy conservation equation provides a modest account for phonon temperature in the device.

Both physical effects, as well as full-wave EM, can be incorporated into deterministic and stochastic device simulators. As outlined in the introduction, two important categories of charge transport solvers are the hydrodynamic and MC methods. Both solvers have been applied to GaN HEMT simulations.

Hydrodynamic solvers are widely used in the device simulation community for analysis and characterization of power transistors. An AlGaN/GaN HEMT with surface traps was investigated with the commercially available DESSIS hydrodynamic software in [96]. This contribution did not include RF simulations, however. Self-heating and hot-electron effects were studied in AlGaN/GaN double-channel HEMTs in [97], but also did not include RF simulations. Finally, the two-dimensional device simulator Minimos-NT [98] was used to accurately simulate both the static and small-signal response of several GaN HEMTs. Hydrodynamic solvers attempt to incorporate electronic band-structure and hot-electron effects through fitting parameters. For example, a field-dependent mobility is tuned in [98] to model velocity saturation in GaN. Furthermore, hydrodynamic solvers rarely incorporate full-wave EM, with the exception of a 2D solver used to simulate THz plasma waves [99].

Many EMC and CMC methods have successfully simulated the responses of GaN HEMTs. Chronologically, this list (which is not claimed to be complete) begins with quantum corrected full-band CMC simulations [100]. Self-heating effects in sub-micron GaN HEMTs were calculated

with an electro-thermal MC method [101] and another MC method investigated the influence of source-gate distance on AlGaN/GaN HEMT performance [102]. Two separate electro-thermal MC analyses of GaN HEMTs were also presented in [103] and [104]. Finally, a full-band CMC solver was applied to large-signal RF simulations of GaN HEMTs in [28]. Although the MC simulations are accurate, the inclusion of full-wave EM is rare with the exceptions of [66] and [105]. A full-wave MC method, however, will suffer from long computational times.

FKT is a promising deterministic Boltzmann solver for large-signal RF GaN HEMT simulations. Features included and physical effects captured in the FKT device simulator are:

- Electronic band-structure calculated with EPM
- Hot-electron effects determined by higher moments of the BTE
- Self-heating effects by inclusion of the lattice heating equation
- Quantum mechanical scattering computed by Fermi's golden rule
- Full-wave electromagnetics calculated with Ampere's and Faraday's laws

The derivation, discretization, and analysis of the FKT equations was presented in Sections 3 – 5. In the following sections, a high speed GaN HEMT is simulated with the FKT device TCAD tool.

6.2 AlGaN/GaN HEMT Topology

Thermal equilibrium, static, and RF small- and large-signal FKT simulations of a single gate AlGaN/GaN HEMT are presented. The AlGaN/GaN HEMT is a high frequency device which demonstrated a record f_{max} of 230 GHz at a specific quiescent bias [106]. This device received considerable attention from the device simulation community. It was first used to study the effects of threading dislocations using the CMC method [107]. The GaN HEMT was also featured in an Institute of Electrical and Electronic Engineers (IEEE) magazine article describing modeling and simulation of terahertz devices with the CMC technique [108]. FKT device simulations of this device were also presented in [84] and [109].

The layout of the AlGaN/GaN HEMT is illustrated in Figure 6.2. This is the 3D mesh used for the FKT simulations. The stack of the device is 11 nm of GaN and 13 nm of $Al_{0.32}Ga_{0.68}N$. The interfaces of the stack are oriented in the z-direction of the device. Detailed illustrations of the HEMT stack are reported in [106, 107, 108]. The stack sits on a GaN buffer which is grown on a SiC substrate. The Ohmic source and drain contacts of the transistor are the yellow regions

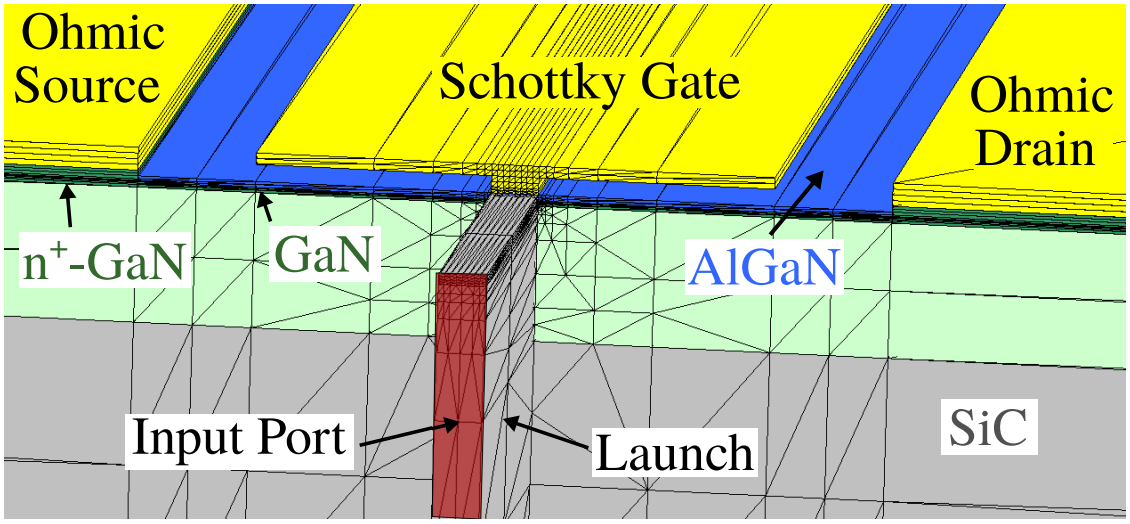


Figure 6.2: The 3D mesh used for FKT simulations of an AlGaN/GaN HEMT.

at the top left and right of the device. The middle contact is the Schottky gate. Not included in the illustration is the Ohmic ground contact at the bottom of the mesh. The SiC launch is located at the front of the device. The input port BC of the launch is shaded in red. Not illustrated is the output port BC connecting the Ohmic drain and the Ohmic ground metal contacts.

6.3 Thermal Equilibrium and Static Simulations

The thermal equilibrium of the HEMT is computed first with the FKT device simulator. As outlined in previous sections, the thermal equilibrium solution variables are the initial guess of the first static solve Newton iteration. Figure 6.3 presents the thermal equilibrium electric potential profile in the transistor. The Schottky T-gate of the device has an electric potential of approximately 0.65 V and is shaded in green. The barrier of the Schottky gate is $\Phi_{SB} = 1.2\,eV$. The blue region surrounding the T-gate is air. The Ohmic contacts both have an electric potential of

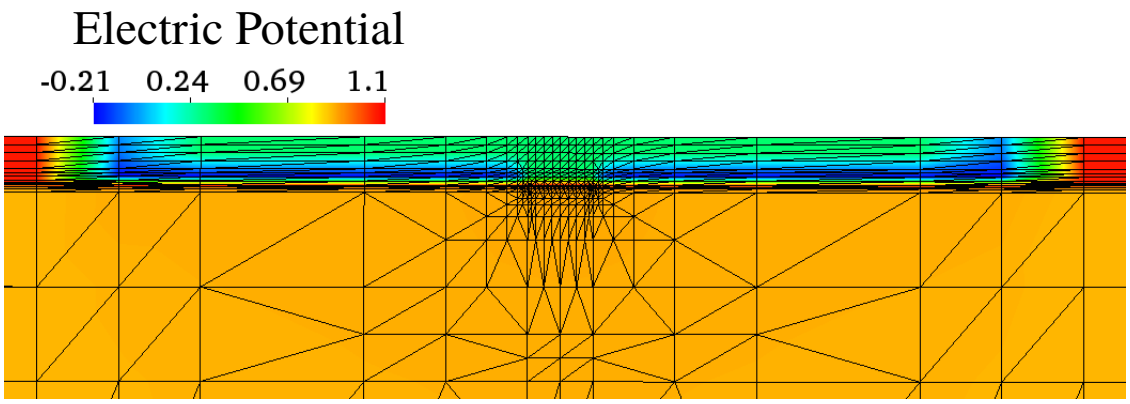


Figure 6.3: The thermal equilibrium electric potential profile of the AlGaN/GaN HEMT computed with the FKT device simulator.

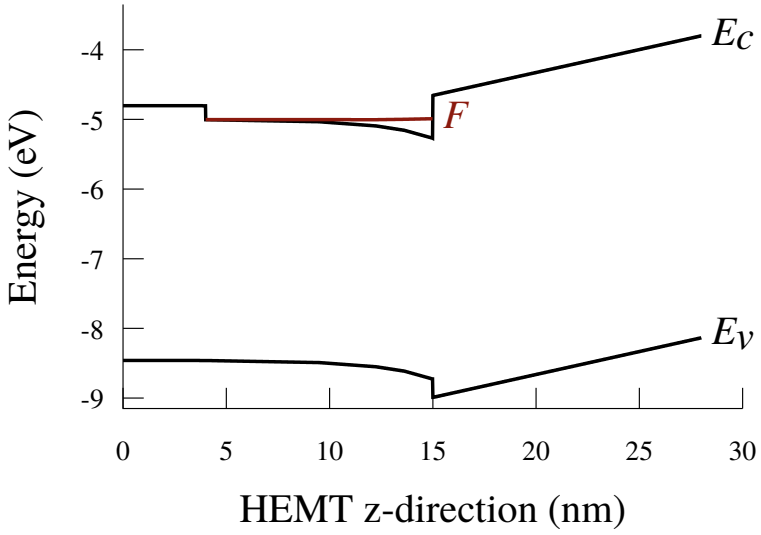


Figure 6.4: The thermal equilibrium band diagram (calculated across the stack under the Schottky gate) of the AlGaN/GaN HEMT computed with the FKT device simulator.

approximately 1.1 V — the red shaded regions.

A band diagram of the GaN HEMT computed with the FKT device simulator is presented in Figure 6.4. This result is calculated across the HEMT stack under the Schottky gate. The beginning of the GaN channel is located just before the 5 nm position in the band diagram. The AlGaN/GaN heterojunction is located at 15 nm in the plot. The polarization charge density resulting from this heterojunction yields the 2DEG of the HEMT. Highly degenerate electron gases reside near this junction in the GaN channel.

Static simulations of the AlGaN/GaN HEMT are presented based on the work in [84]. An

example of the quiescent bias of the HEMT calculated with FKT is illustrated in Figure 6.5. The static solver consists of Eqns. (5.1), (5.4), (5.5), and (5.6). Trapping effects are also included to properly determine the mobile electron densities in the GaN channel [62, 84]. Figure 6.6 presents the static I-V family of the AlGaN/GaN HEMT computed with the FKT device solver. The

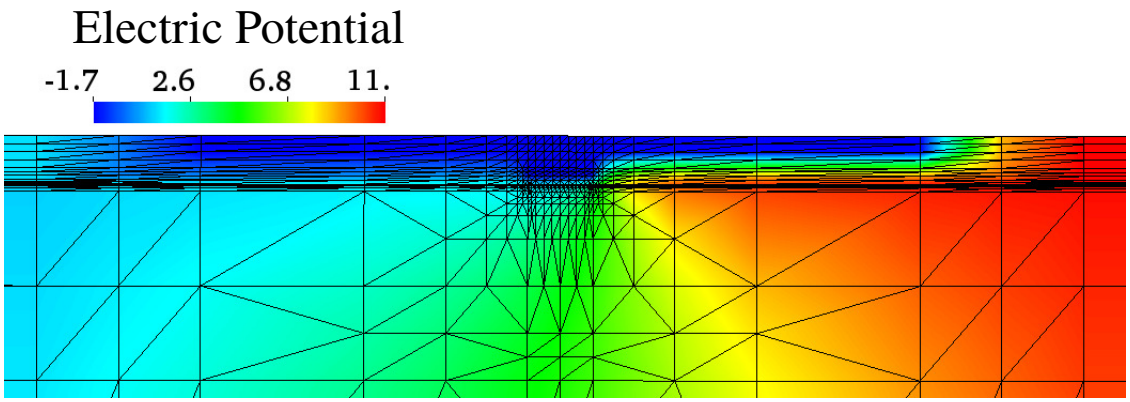


Figure 6.5: The electric potential profile of the AlGaN/GaN HEMT in quiescent bias. The gate-source bias is -1 V and the drain-source bias is 10 V.

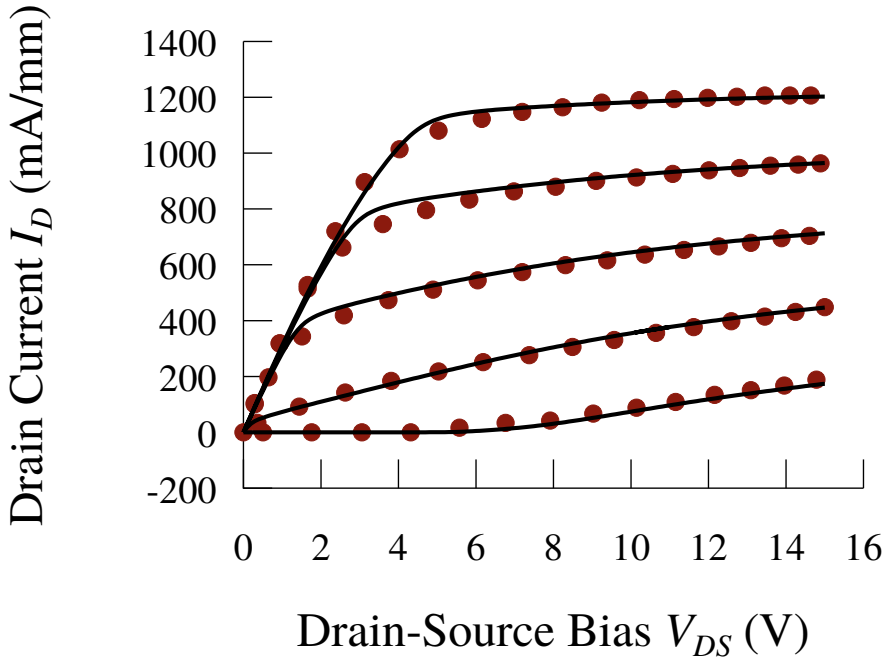


Figure 6.6: The static I-V family of the AlGaN/GaN HEMT. The solid black lines represent FKT simulations with different gate biases and the red dots are the measured equivalents. The gate biases range from 1 V to -3 V with steps of 1 V. This result was first presented by Matt Grupen AFRL.

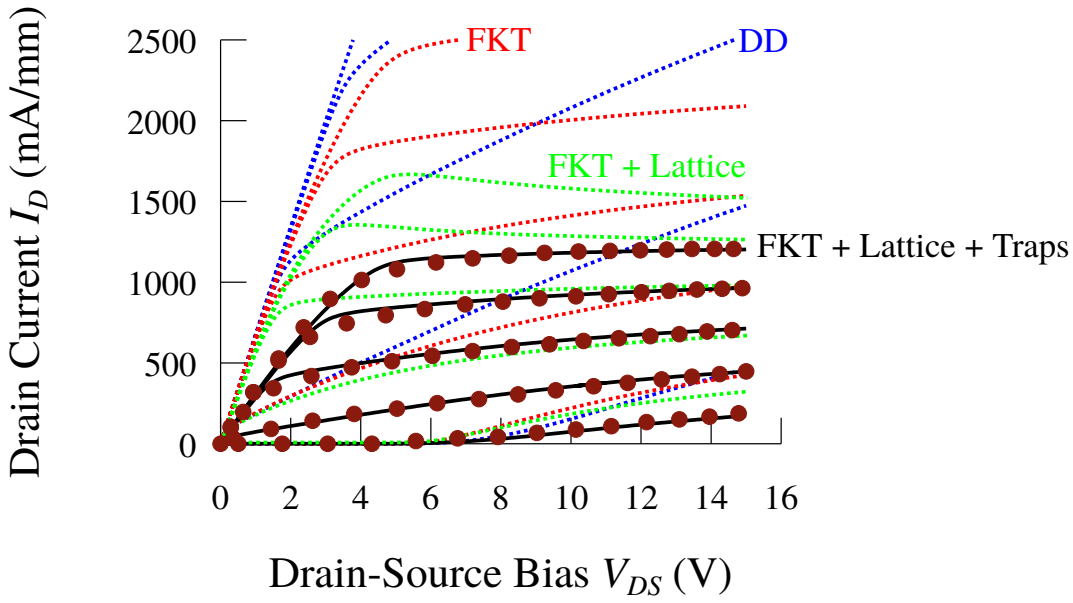


Figure 6.7: A comparison of the static I-V families of the AlGaN/GaN HEMT computed with the different FKT device solvers. The DD, FKT, FKT with lattice heating, and the FKT solver with lattice heating and trapping effects are illustrated with dashed blue, dashed red, dashed green, and solid black lines, respectively. The measurements are illustrated with red dots.

gate biases range from $V_{GS} = 1$ V to -3 V with steps of 1 V. The drain-source bias, V_{DS}, is swept from 0 to 15 V at each gate bias. FKT simulations are illustrated with solid black lines and the measurements [107] are illustrated with red dots.

Proper treatment of the electronic band-structure, the hot-electron and self-heating effects, the scattering, and the deep traps in the simulation framework is critical for accurately reproducing the measured data of the device. Figure 6.7 presents a comparison of several different device simulations. The DD (no energy transport, scattering, lattice heating or traps), FKT (no lattice heating or traps), FKT with lattice heating (no traps), and the complete device simulator are drawn with dashed blue lines, dashed red lines, dashed green lines, and solid black lines, respectively. Again, the measurements [107] are illustrated with red dots. The DD and FKT simulations severely overestimate the drain currents. The FKT simulation with lattice heating provides better results. Finally, the complete device simulation exhibits excellent agreement with the measurements. The deep traps are required to properly model the channel resistance.

Finally, the transconductance of the AlGaN/GaN HEMT is calculated. The gate-source bias,

V_{GS}, is swept from 1 V to device pinch-off at a specific drain-source bias, V_{DS}. The transconductance is calculated with Eqn. (2.20). Figure 6.8 presents the transconductance of the HEMT at a drain-source bias of $V_{DS} = 10\,\text{V}$. Here, the solid black line illustrates the drain current, I_D, and red line represents the transconductance, g_m. This simulation is particularly useful to RF circuit

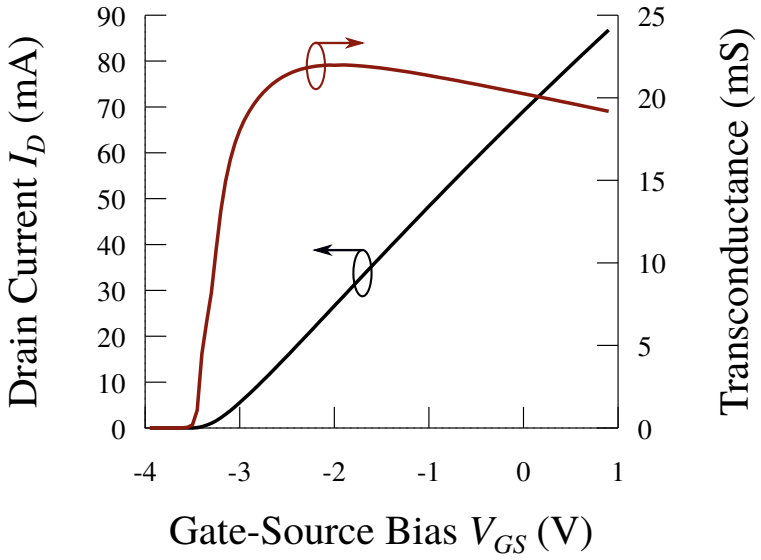


Figure 6.8: The transconductance of the AlGaN/GaN HEMT. The solid black line represents the drain current and the red line is the transconductance of the device. The drain-source bias is 10 V.

designers. First, it can be imported as the transconductance component of a small- or large-signal equivalent circuit. It is also used to determine the quiescent bias for RF simulations/measurements. The device is typically operated at peak transconductance to provide the largest power gain.

6.4 Small-Signal RF Simulations

Simulations of the AlGaN/GaN HEMT with low input power are next presented. The linear two-port simulation of the device follows the discussion in Section 2.4.2. To this end, the S-Parameters of the device are calculated with the quasi-static and full-wave solvers and used to quantify the small-signal power gain and small-signal current gain.

Calculating the S-Parameters, Eqns. (2.30) – (2.33), requires two FKT device simulations. For the first FKT simulation, a voltage port is assigned to the input of the device and an AC impedance port is assigned to the output of the device. The impedance presented to the output is chosen

such that $\Gamma_L = 0$. This ensures that $A_2 = 0$ and allows the calculation of S_{11} and S_{21}. The input port of the quasi-static simulation is the Schottky gate contact and the output port is the Ohmic drain contact. For full-wave simulations, the input port is the boundary connecting the Schottky gate and Ohmic ground contacts and the output port is the boundary connecting the Ohmic drain and the Ohmic ground contacts. The second simulation reverses the port BCs and presents the AC impedance to the input device such that $\Gamma_S = 0$. The final two S-Parameters S_{12} and S_{22} are calculated from this simulation. Because the reference impedance is $Z_0 = 50\,\Omega$, both input and output AC impedances are $50\,\Omega$.

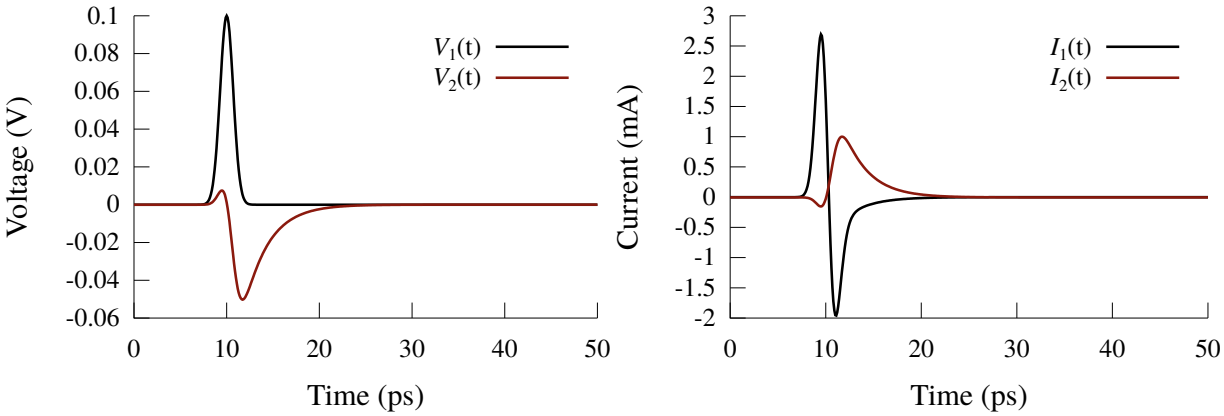


Figure 6.9: The (a) AC input and output voltages and (b) AC input and output currents of the AlGaN/GaN HEMT calculated with the quasi-static solver. The voltage BC is assigned to the input port and the AC impedance port BC is assigned to the output. For the quasi-static solver, the input port is the Schottky gate metal contact and the output port is the Ohmic drain metal contact.

Figures 6.9a and 6.9b present the transient input and output signals of the AlGaN/GaN HEMT computed with the quasi-static solver. Figures 6.10a and 6.10b present the same results computed with the full-wave solver. The input port is a voltage BC and the output port is terminated with a $50\,\Omega$ AC impedance BC. The Fourier transforms of these signals are used to calculate the incident waves, Eqn. (2.25), and scattered waves, Eqn. (2.26). The S_{11} and S_{21} S-Parameters are calculated with Eqns. (2.30) and (2.31), respectively. The relation between the output voltage $V_2(t)$ and output current $I_2(t)$ deserves special attention. The AC impedance BC requires a specific relationship between its integrated voltage and current. As discussed in Sections 5.4.3 and 5.4.5, the current

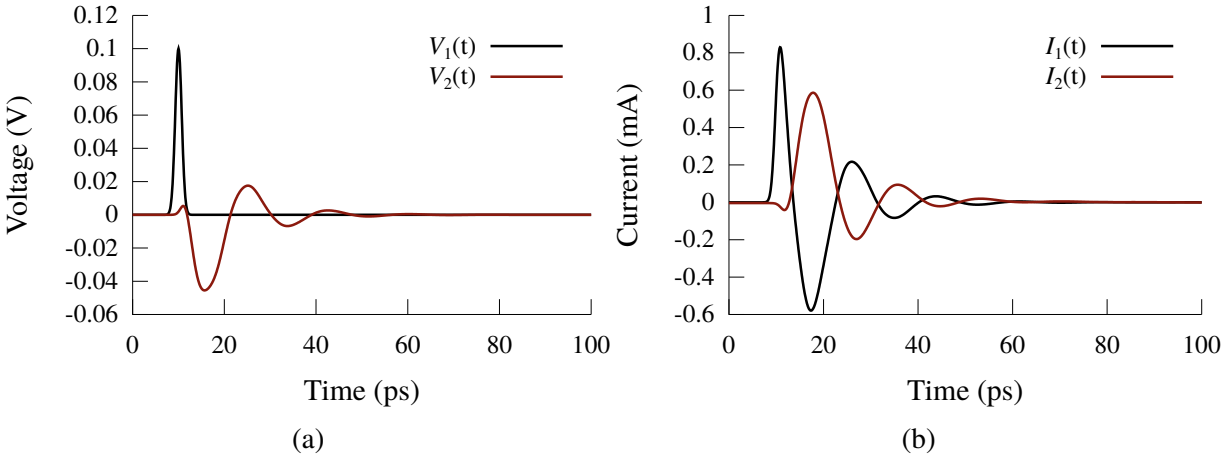


Figure 6.10: The (a) AC input and output voltages and (b) AC input and output currents of the AlGaN/GaN HEMT calculated with the full-wave solver. The voltage port BC is assigned to the input port and the AC impedance port BC is assigned to the output. For the full-wave solver, the input port connects the Schottky gate and the Ohmic ground and the output port connects the Ohmic drain and the Ohmic ground.

is prescribed to be exactly $-V_2(t)/50$. This relationship is qualitatively observed in Figures 6.9 and 6.10.

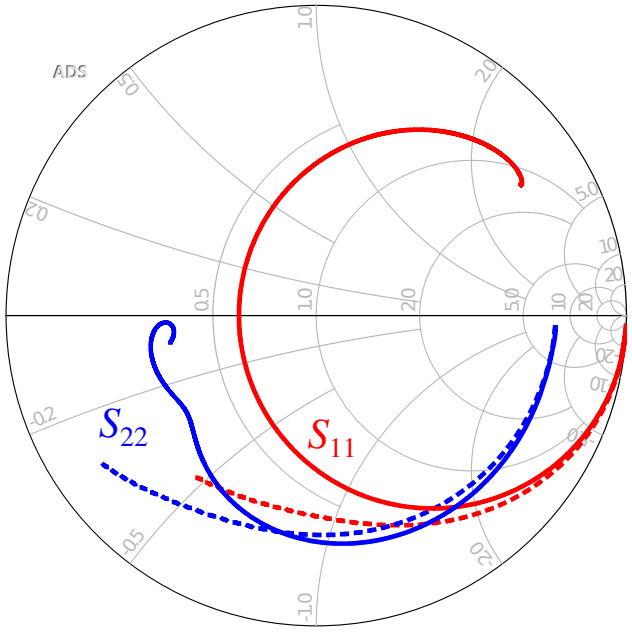


Figure 6.11: Two S-Parameters of the AlGaN/GaN HEMT computed with the FKT device simulator. The solid lines represent the full-wave results and the dashed lines represent the quasi-static results. The quiescent bias of the device is a drain-source bias of 6 V and no gate-source bias.

Two S-Parameters, S_{11} and S_{22}, are presented in Figure 6.11. Because the source and load

impedances are chosen to be $Z_S = Z_L = 50\,\Omega$, these S-Parameters represent the input and output reflection coefficients, i.e., $\Gamma_{in} = S_{11}$ and $\Gamma_{out} = S_{22}$. The S-Parameters are plotted versus frequency. Dashed lines represent the quasi-static results and solid lines represent the full-wave results. Substantial differences exist between the quasi-static and full-wave S-Parameters at high frequency. These differences are due to inductive coupling of the metal contacts.

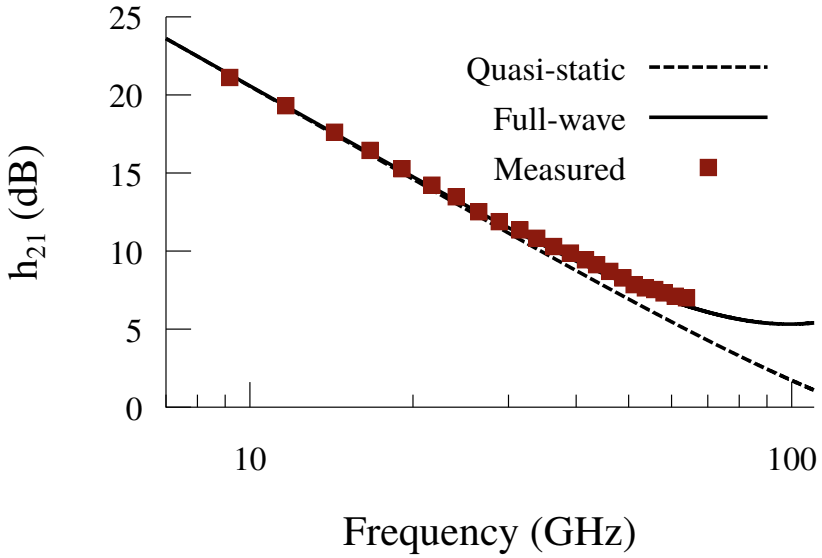


Figure 6.12: The simulated (black lines) and the measured (red squares) small-signal current gain of the AlGaN/GaN HEMT.

A comparison of the simulated and measured small-signal current gain is reported in Figure 6.12. The small-signal current gain is calculated with Eqn. (2.43). The dashed black line represents the quasi-static result and the solid black line is the full-wave result. The red squares represent the measurements [107]. It is evident that treatment of full-wave EM is required in the FKT simulator in order to properly capture the device parasitics at high frequency. This allows the device solver to compute accurate S-Parameters and provide useful data for RF circuit designers to synthesize high frequency matching networks for the GaN HEMT.

Finally, the small-signal power gain of the AlGaN/GaN HEMT is presented. Figure 6.13 reports the available power gain (black lines) calculated with Eqn. (2.41) and the transducer power gain (red lines) calculated with Eqn. (2.42). The dashed lines represent the quasi-static power gains and the solid lines are the power gains calculated with the full-wave solver. This result provides

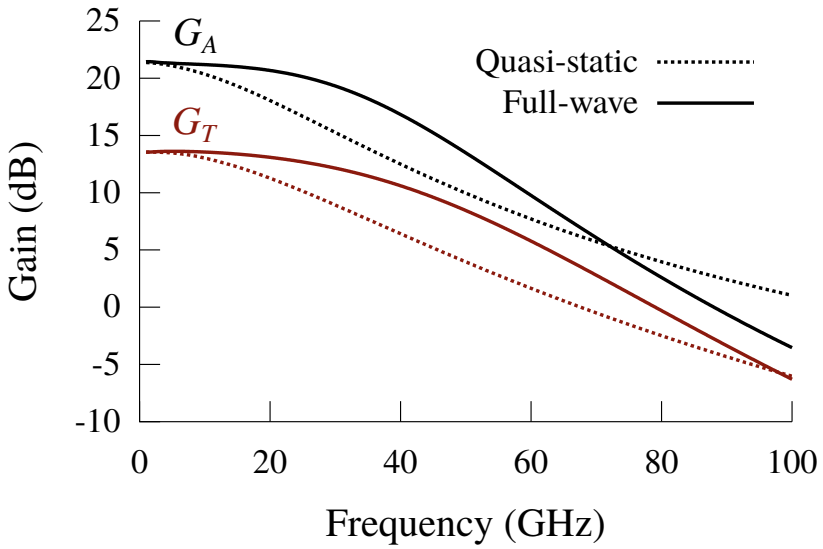


Figure 6.13: The simulated available power gain (black) and transducer power gain (red) of the AlGaN/GaN HEMT. Dashed lines represent the quasi-static power gains and the solid lines represent the full-wave power gains.

further evidence that the full-wave solver is crucial for simulating the small-signal response of GaN HEMT technology at high frequency.

6.5 Large-Signal RF Simulations

As outlined in Section 2.4.3, S-Parameters do not accurately model the nonlinear response of high power transistors. Large-signal simulations of the transistors, including LP and power sweeps, are therefore used for PA design. LP data determine the optimal loading conditions for maximizing specific FOMs and power sweep results provide further information for optimal operating conditions. This section focuses on utilizing the FKT device solver framework for large-signal RF simulations of GaN HEMT technology. The quasi-static AC impedance is tailored to simulate transistors with arbitrary reflection coefficients in the following subsection. Then, a discussion on calculating the large-signal FOMs with the FKT device solver is presented. Finally, this section concludes with LP and power sweep simulations of the GaN HEMT and comparisons of quasi-static and full-wave large-signal results.

6.5.1 Quasi-Static AC Load Impedance for LP Simulations

A critical component of large-signal RF simulations is the load impedance presented to the output port of the transistor. Because the load reflection coefficient is sampled over the Smith chart, the FKT device solver must incorporate impedances with reactive components. Section 5.4.3 discussed the quasi-static AC impedance BC. This BC uses the total integrated current flowing into the boundary, designated i^{t+1}, to specify a Dirichlet condition on the electric potential. The Dirichlet condition is

$$\Phi_i^{t+1}\Big|_{\text{ACI}} = Z_{\text{ACI}} i_{\text{ACI}}^{t+1}, \tag{6.1}$$

where $\Phi_i^{t+1}\Big|_{\text{ACI}}$ represents the electric potential on the quasi-static AC impedance BC. To sample the upper half of the Smith chart, the load is specified as a series RL impedance. In the TD, the BC is

$$\Phi_i^{t+1} = R i^{t+1} + L\frac{i^{t+1} - i^t}{\Delta t}, \tag{6.2}$$

where R and L are the real and imaginary components of the load impedance Z_L at the fundamental frequency. The subscript ACI is dropped for brevity. The inverse of Eqn. (2.28) gives the load impedance in terms of the reflection coefficient, or

$$Z_L = \frac{1+\Gamma_L}{1-\Gamma_L} Z_0. \tag{6.3}$$

Sampling the lower half of the Smith chart requires RC load impedances which are implemented in a similar manner.

6.5.2 Computing Large-Signal FOMs

The large-signal FOMs of the devices are computed at a single operating condition. Specifically, the GaN HEMT at a given quiescent bias is driven by a single sinusoid with a large-signal amplitude. Furthermore, a load-impedance is presented to the output port of the device with the AC quasi-static BC.

Figures 6.14 and 6.15 present the transient voltages and currents at both the input and output ports calculated with the quasi-static FKT device solver. The quiescent bias of the GaN HEMT is

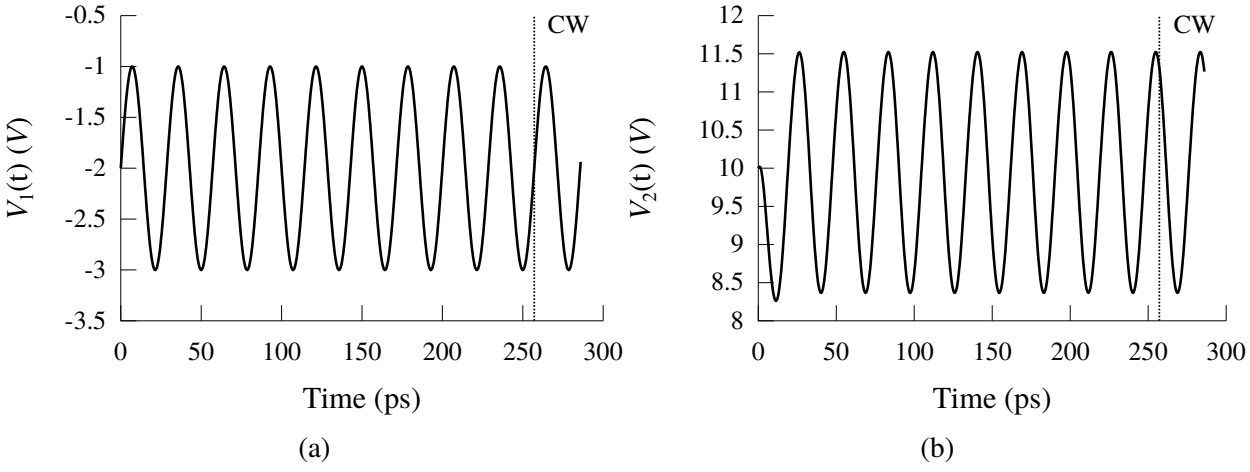


Figure 6.14: The (a) input and (b) output TD voltages of the AlGaN/GaN HEMT calculated with the quasi-static solver. The operating point of the HEMT is -2 V gate-source bias and 10 V drain-source bias. The large-signal tone (gate-source voltage) is 1 V in amplitude at 35 GHz.

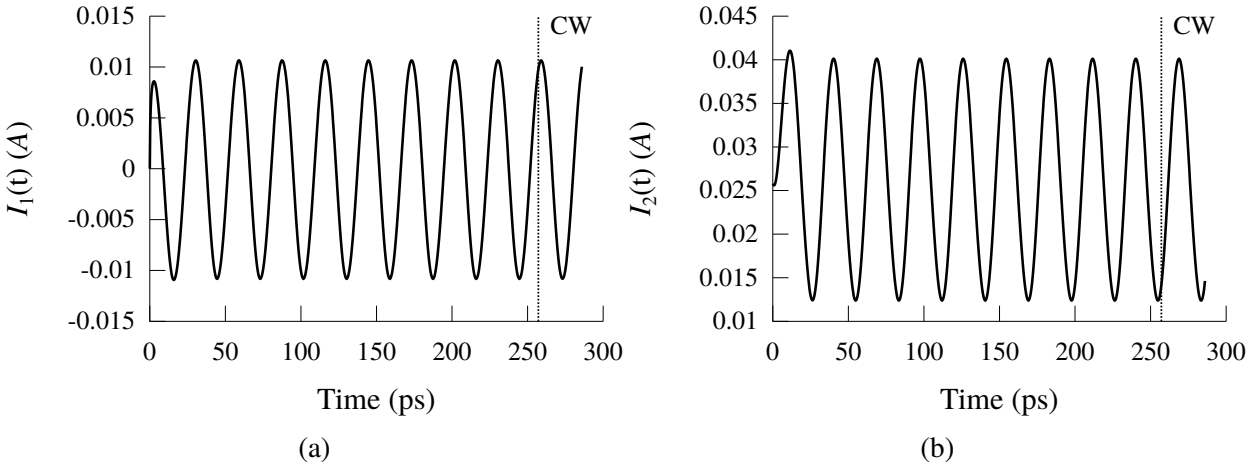


Figure 6.15: The (a) input and (b) output TD currents of the AlGaN/GaN HEMT calculated with the quasi-static solver. The operating point of the HEMT is -2 V gate-source bias and 10 V drain-source bias. The large-signal tone (gate-source voltage) is 1 V in amplitude at 35 GHz.

$V_{GS} = -2$ V and $V_{DS} = 10$ V. The HEMT is driven by a large-signal sinusoid with a 1 V amplitude at 35 GHz fundamental frequency. The load reflection coefficient presented to the output port is $\Gamma_L = 0.44 - 0.17j$ at 35 GHz. The device is simulated for 10 cycles of the driving tone to ensure that all transient signals are continuous wave (CW).

The CW portion of the transient signals are Fourier transformed to calculate their spectral content. These signals are interpolated such that the first discrete frequency component corresponds to the zeroth or DC frequency and the second discrete frequency component corresponds to the

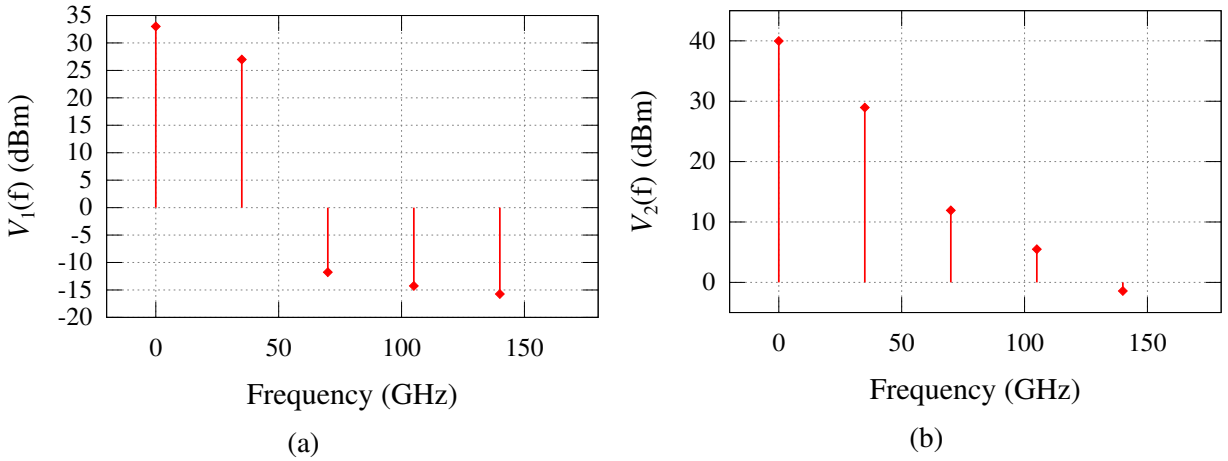


Figure 6.16: The magnitudes of the (a) input and (b) output voltage phasors of the AlGaN/GaN HEMT driven by a large-signal tone. These voltages are plotted in dBm. The five harmonics are computed by Fourier transforms of the transient voltages.

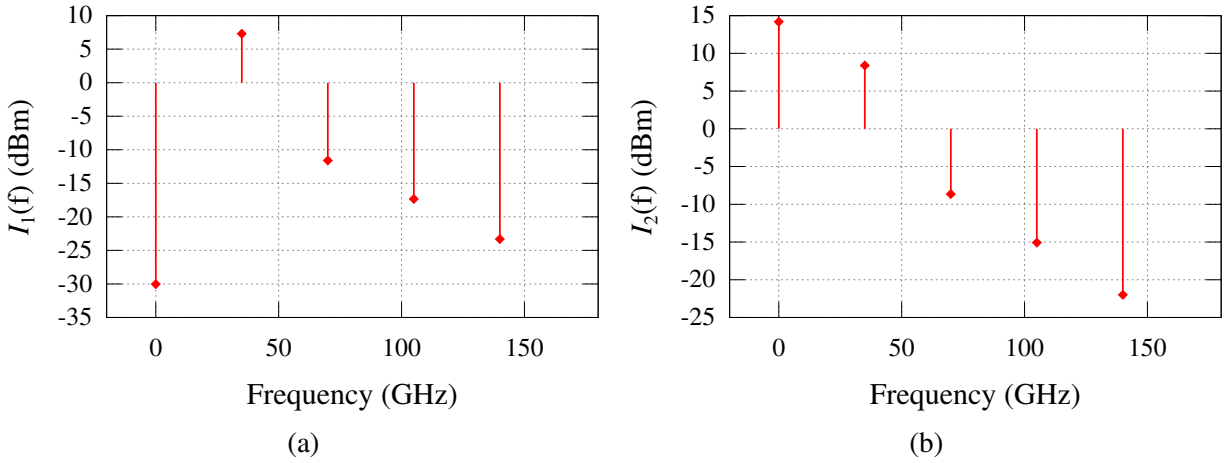


Figure 6.17: The magnitudes of the (a) input and (b) output current phasors plotted in dBm. These five harmonics are computed by Fourier transforms of the transient currents.

fundamental frequency of the large-signal tone. Figures 6.16 and 6.17 present the first five harmonics of the voltages and currents at the input and output ports of the GaN HEMT driven by the large-signal RF tone. The magnitudes of the spectral voltages and currents are plotted in dBm. The first harmonic of each plot is used to calculate the FOMs of the PA. The power available from the source and the input power of this particular operating condition are $P_{avs} = -2.04\,\text{dBm}$ and $P_{in} = -2.83\,\text{dBm}$, respectively. The transistor generates an output power of $P_{out} = 4.33\,\text{dBm}$ which corresponds to a power gain of $G = 7.17\,\text{dB}$ and a transducer power gain of $G_T = 6.39\,\text{dB}$.

6.5.3 LP Simulations

Optimal load impedances corresponding to specific operating conditions (bias, frequency, input power) are determined through LP simulations. To this end, the load reflection coefficient presented to the output port of the device is swept over a wide range of the Smith chart. Figure 6.18

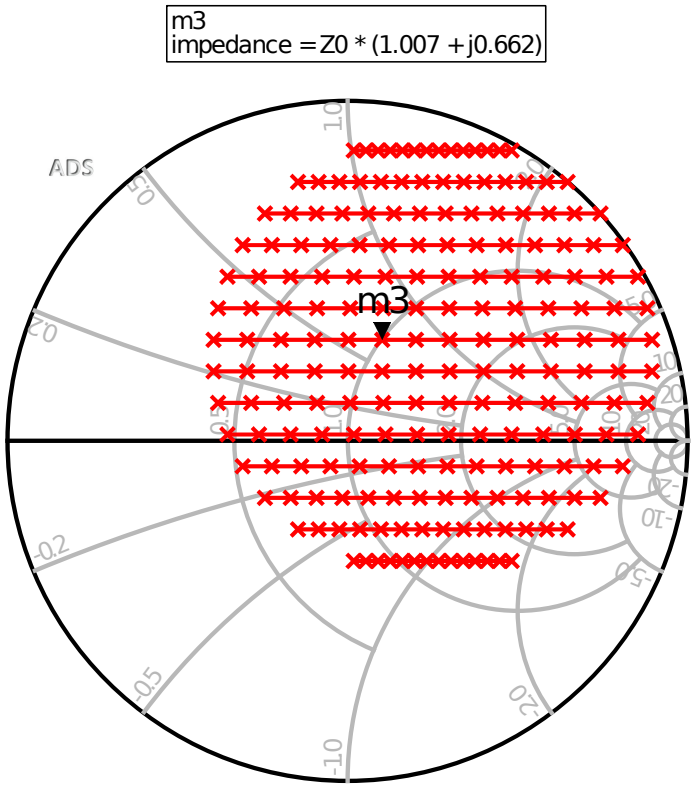
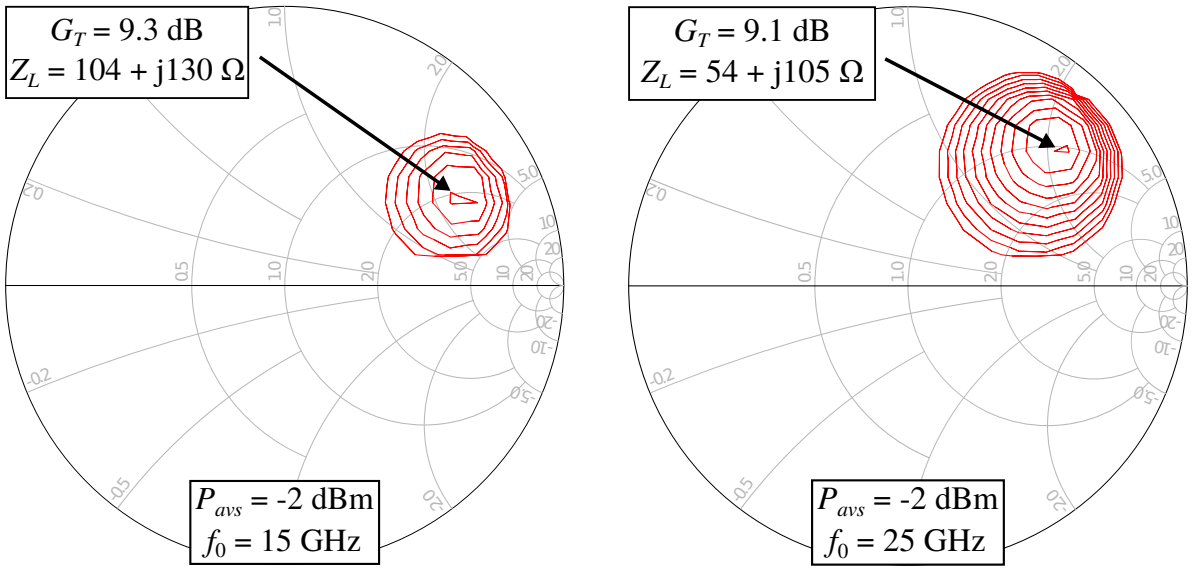


Figure 6.18: An example of reflection coefficient samples of the Smith chart that can be used for LP simulations.

presents a collection of reflection coefficients that can be used for LP simulations. The marker "m3" points to a reflection coefficient corresponding to an impedance of $Z_L = 50.4 + j33.1\,\Omega$.

The GaN HEMT is simulated with the FKT device solver for each reflection coefficient in the LP sample set. All desired FOMs are calculated from the CW voltages and currents. Finally, these FOMs are loaded into ADS and contours of each FOM are plotted on the Smith chart. The following sections will utilize this work flow for PA design.



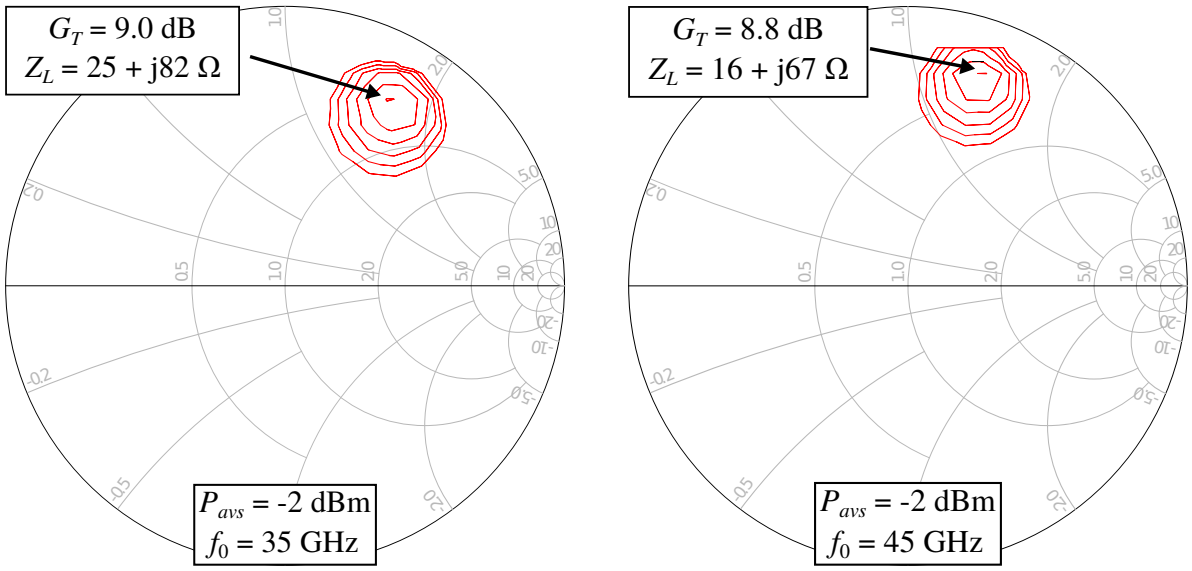
(a) (b)

Figure 6.19: LP contours calculated from FKT device simulations of the GaN HEMT biased as a class A amplifier (gate-source bias of 1 V and drain-source bias of 10 V) with -2 dBm power available from the source at a fundamental frequency of (a) 15 GHz and (b) 25 GHz. The optimal loads for maximum transducer gain are $104 + j130\,\Omega$ at 15 GHz and $54 + j105\,\Omega$ at 25 GHz. The peak gain is 9.3 dB at 15 GHz and 9.1 dB at 25 GHz. The step size of the contours is 0.2 dB.

6.5.4 Class A Amplifier

Large-signal simulation results of the GaN HEMT biased as a class A amplifier are presented in this section. The first step in the PA design process is LP simulation with moderate input power. Multiple LP data sets are collected from the GaN HEMT simulations at various fundamental frequencies. The quiescent bias of the amplifier is $V_{GS} = 1$ V and $V_{DS} = 10$ V. The amplifier is driven by a power available from the source of $P_{avs} = -2$ dBm.

Figure 6.19a presents the LP data calculated from FKT device simulations of the class A amplifier operating at 15 GHz. The FOM is the transducer gain of the device. Peak gain resides at a load impedance of $Z_L = 104 + j130\,\Omega$ at the fundamental frequency. The LP data set of the device operating at 25 GHz is next reported in Figure 6.19b. The load impedance corresponding to maximum transducer gain shifts to $Z_L = 54 + j105\,\Omega$ at 25 GHz. The last two sets of LP data are presented in Figures 6.20a and 6.20b. The former is calculated with a fundamental frequency of 35 GHz and the latter with 45 GHz.



(a) (b)

Figure 6.20: LP contours calculated from FKT device simulations of the GaN HEMT biased as a class A amplifier (gate-source bias of 1 V and drain-source bias of 10 V) with -2 dBm power available from the source at a fundamental frequency of (a) 35 GHz and (b) 45 GHz. The optimal loads for maximum transducer gain are $25 + j82\,\Omega$ at 35 GHz and $16 + j67\,\Omega$ at 45 GHz. The peak gain is 9.0 dB at 35 GHz and 8.8 dB at 45 GHz. The step size of the contours is 0.2 dB.

The next step in the PA design process is sweeping the input power with the optimal load impedance determined from the LP data. Because the input power depends on the loading condition as well as the intrinsic device parasitics, it is not trivial to sweep this quantity in the FKT device simulator. Therefore, the power available form the source, P_{avs}, is swept across a broad range and the resulting FOMs are plotted versus the calculated input power, P_{in}.

Figure 6.21a presents the input power sweep of the class A amplifier operating at a fundamental frequency of 15 GHz. In this figure, the output power, P_{out}, is plotted in black, the transducer gain, G_T, is plotted in red, and the *PAE* is plotted in blue. The power gain of the device is relatively constant at low input power. The 1 dB compression point is approximately -10 dBm. Figures 6.21b, 6.22a, and 6.22b report the input power sweeps of the HEMT at 25 GHz, 35 GHz, and 45 GHz, respectively. These simulations provide quantitative trends in the gain and efficiency of the GaN HEMT versus fundamental frequency. The HEMT is more efficient at lower frequency. Maximum *PAE* is achieved at an input power of approximately 10 dBm at 15 GHz. The gain of the

device is vastly reduced to approximately 2–3 dBm at 45 GHz. These high power results could be useful to circuit designers for determining trade-offs to meet specific criteria.

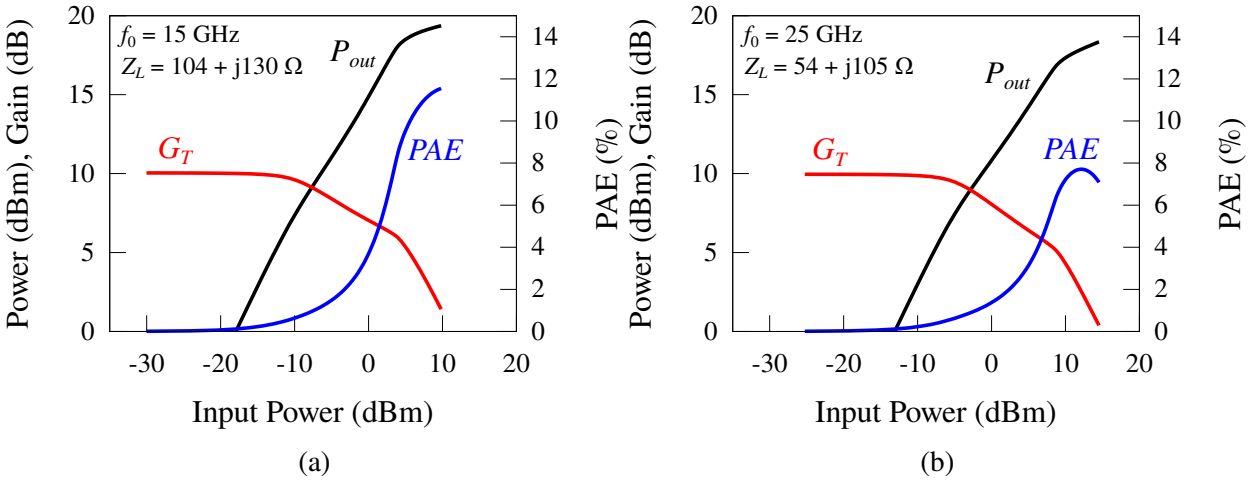


Figure 6.21: Simulated input power sweeps of the GaN HEMT biased as a class A amplifier (gate-source bias of 1 V and drain-source bias of 10 V) at a fundamental frequency of (a) 15 GHz and (b) 25 GHz. The optimal loads determined from the LP data are $104 + j130\,\Omega$ at 15 GHz and $54 + j105\,\Omega$ at 25 GHz. The peak *PAE* resides at an input power of approximately 10 dBm for both frequencies. The device operating at 15 GHz achieves the highest overall peak *PAE* of approximately 12%.

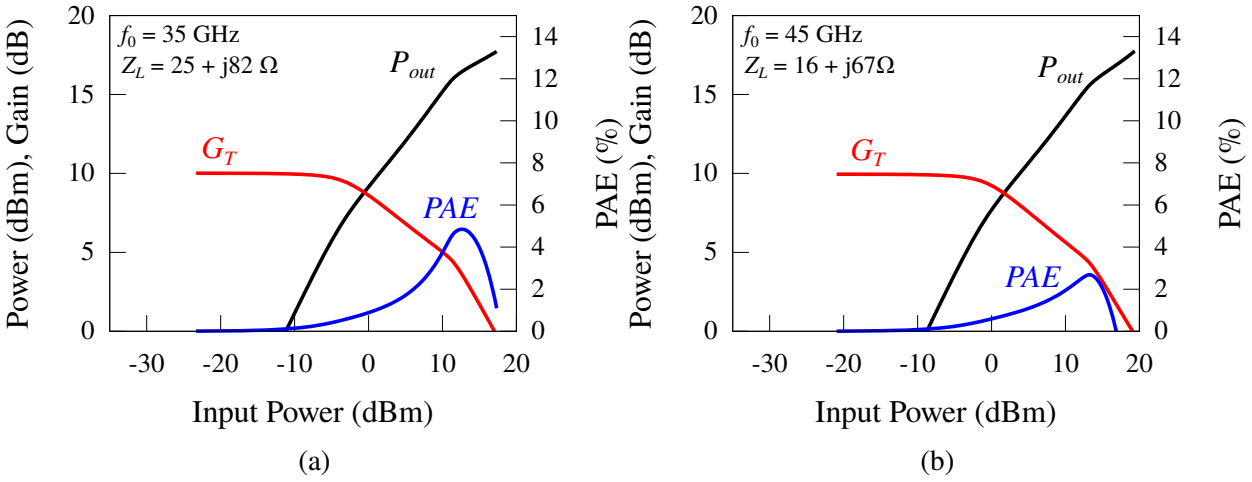


Figure 6.22: Simulated input power sweeps of the GaN HEMT biased as a class A amplifier (gate-source bias of 1 V and drain-source bias of 10 V) at a fundamental frequency of (a) 35 GHz and (b) 45 GHz. The optimal loads determined from the LP data are $25 + j82\,\Omega$ at 35 GHz and $16 + j67\,\Omega$ at 45 GHz. The *PAE* diminishes with increasing frequency. At 45 GHz, the peak *PAE* is between 2–3%.

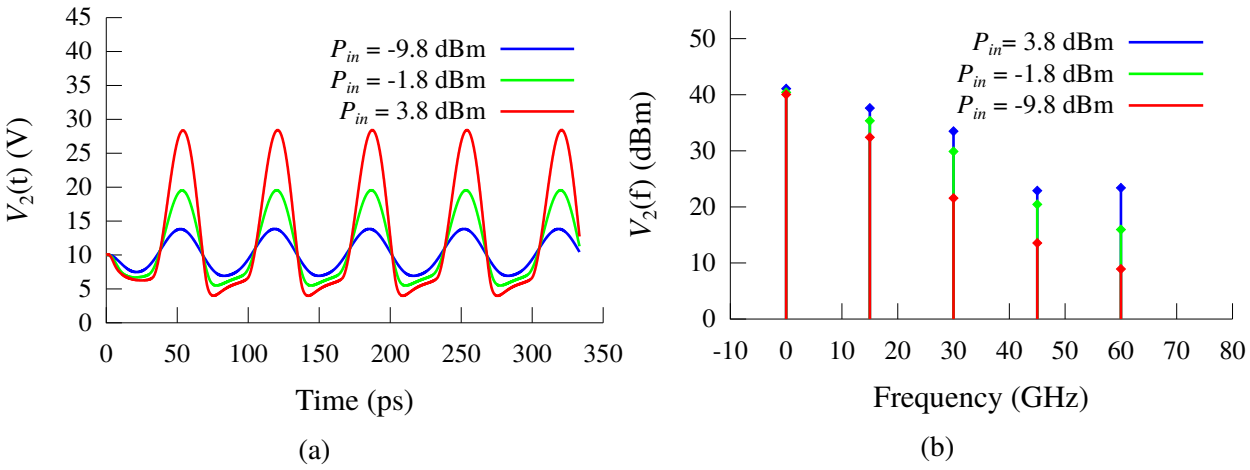


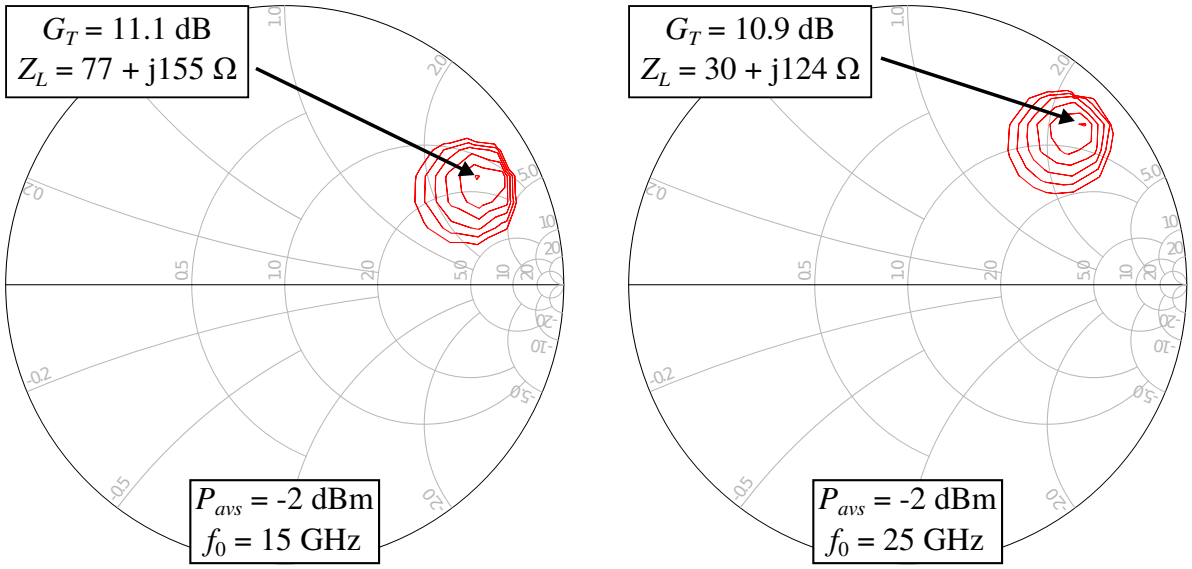
Figure 6.23: Simulated (a) output voltage transients and (b) output voltage spectral content of the GaN HEMT biased as a class A amplifier at 15 GHz. The input powers of the blue, green, and red lines are -9.8 dBm, -1.8 dBm, and 3.8 dBm, respectively. The transistor exhibits strong inter-modulation at each simulated input power.

The transient output voltages of the class A amplifier corresponding to various input powers are presented in Figure 6.23a. The fundamental frequency of these HEMT simulations is 15 GHz. Figure 6.23b illustrates the spectral content of the output voltages. The blue, green, and red lines represent the input powers $P_{in} = -9.8$ dBm, -1.8 dBm, and 3.8 dBm, respectively. Significant inter-modulation is observed in the response of the device at each input power.

6.5.5 Class AB Amplifier

The GaN HEMT is next biased as a class AB amplifier to improve efficiency. The gate-source bias is $V_{GS} = -2$ V. This operating point is between the linear class A bias and the nonlinear class B bias. The DC drain current is smaller at this bias resulting in a higher *PAE*.

The optimal load impedance for peak transducer gain of the class AB amplifier is determined with LP data calculated from FKT device simulations. Four LP data sets are presented in Figures 6.24a – 6.25b. The operating frequencies of the FKT device simulations are 15 GHz – 45 GHz with increments of 5 GHz. The optimal loading conditions for maximum transducer gain are highlighted in each figure. This impedance shifts across the Smith chart in a counter-clockwise fashion as the fundamental frequency increases. At 15 GHz, the optimal loading condi-



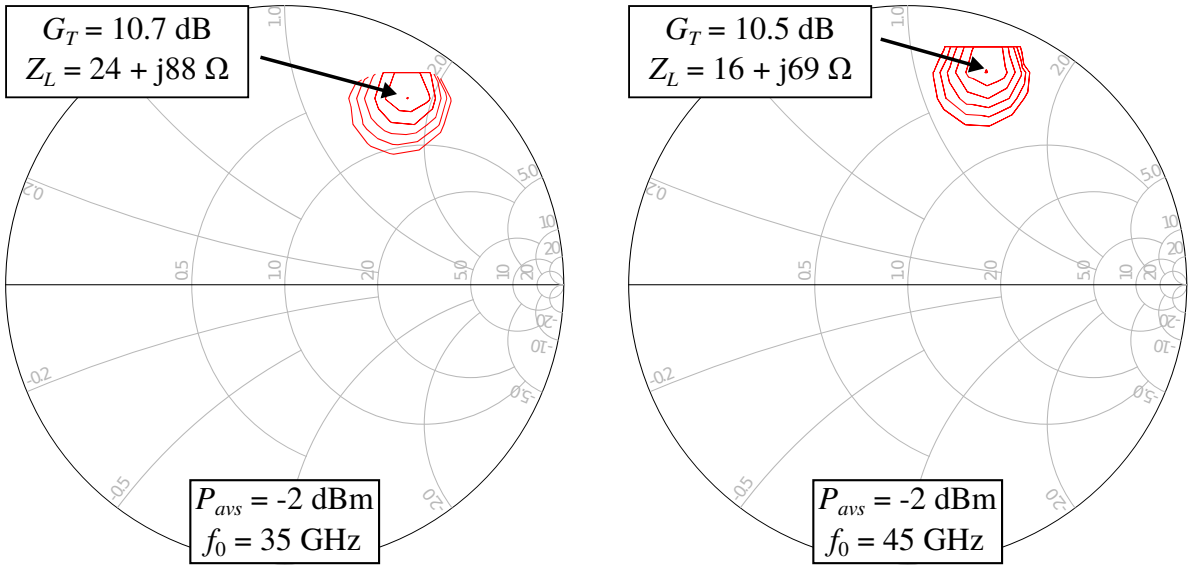
(a) (b)

Figure 6.24: LP data calculated from FKT device simulations of the GaN HEMT biased as a class AB amplifier (gate-source bias of -2 V and drain-source bias of 10 V) with -2 dBm power available from the source at a fundamental frequency of (a) 15 GHz and (b) 25 GHz. The optimal loads for maximum transducer gain are $77+j155\,\Omega$ at 15 GHz and $30+j124\,\Omega$ at 25 GHz. The peak gain is 11.1 dB at 15 GHz and 10.9 dB at 25 GHz. The step size of the transducer gain contours is 0.2 dB.

tion is $Z_L = 77 + j155\,\Omega$. As the frequency increases from 25 GHz and 35 GHz, the optimal load impedance decreases from $Z_L = 30 + j124\,\Omega$ to $Z_L = 24 + j88\,\Omega$. Finally, at 45 GHz, the optimal load impedance for peak transducer gain is $Z_L = 16 + j69\,\Omega$.

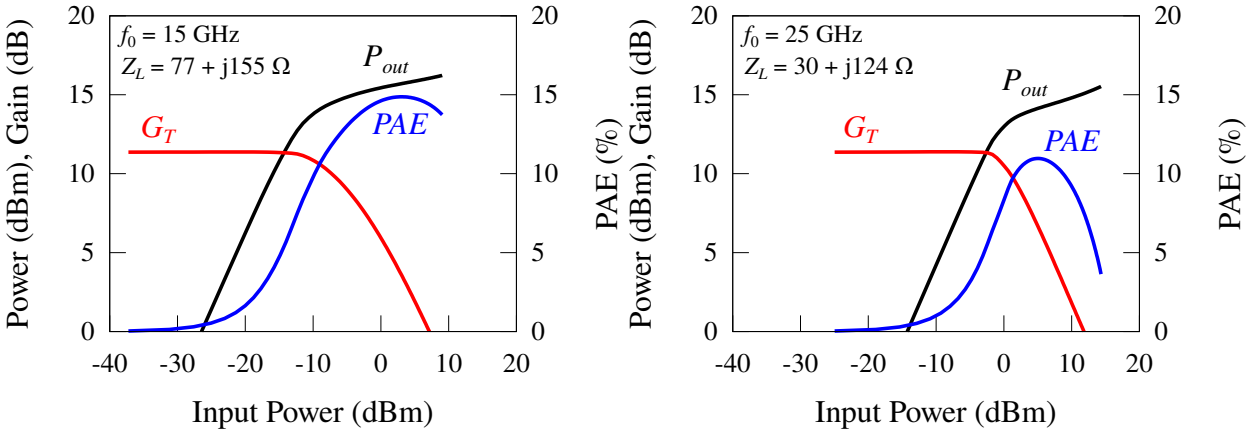
Power sweeps of the GaN HEMT biased as a class AB amplifier with the optimal loading conditions are next presented. The power available from the source, P_{avs}, is swept across a broad range and the resulting FOMs are plotted versus the calculated input power, P_{in}. Figure 6.26a reports the input power sweep of the HEMT at 15 GHz. The load impedance determined from the LP data is $Z_L = 77 + j155\,\Omega$. The output power, P_{out}, is plotted in black, the transducer gain, G_T, is plotted in red, and the *PAE* is plotted in blue. Figures 6.26b – 6.27b report the power sweep results for the amplifier at 25, 35, and 45 GHz, respectively. The load impedances used in these simulations are $Z_L = 30 + j124\,\Omega$, $24 + j88\,\Omega$, and $16 + j69\,\Omega$, respectively.

The power sweep results demonstrate that the class AB amplifier exhibits better efficiency than the class A amplifier, particularly at high frequency. At 45 GHz, the class AB amplifier



(a) (b)

Figure 6.25: LP data calculated from FKT device simulations of the GaN HEMT biased as a class AB amplifier (gate-source bias of -2 V and drain-source bias of 10 V) with -2 dBm power available from the source at a fundamental frequency of (a) 35 GHz and (b) 45 GHz. The optimal loads for maximum transducer gain are $24 + j88\,\Omega$ at 35 GHz and $16 + j69\,\Omega$ at 45 GHz. The peak gain is 10.7 dB at 35 GHz and 10.5 dB at 45 GHz. The step size of the transducer gain contours is 0.2 dB.



(a) (b)

Figure 6.26: Simulated input power sweeps of the GaN HEMT biased as a class AB amplifier (gate-source bias of -2 V and drain-source bias of 10 V) at a fundamental frequency of (a) 15 GHz and (b) 25 GHz. The optimal loads determined from the LP data are $77 + j155\,\Omega$ at 15 GHz and $30 + j124\,\Omega$ at 25 GHz. The peak *PAE* resides at an input power of approximately 0 dBm at 15 GHz and 5 dBm at 25 GHz. The device operating at 15 GHz achieves the highest overall peak *PAE* of approximately 15%.

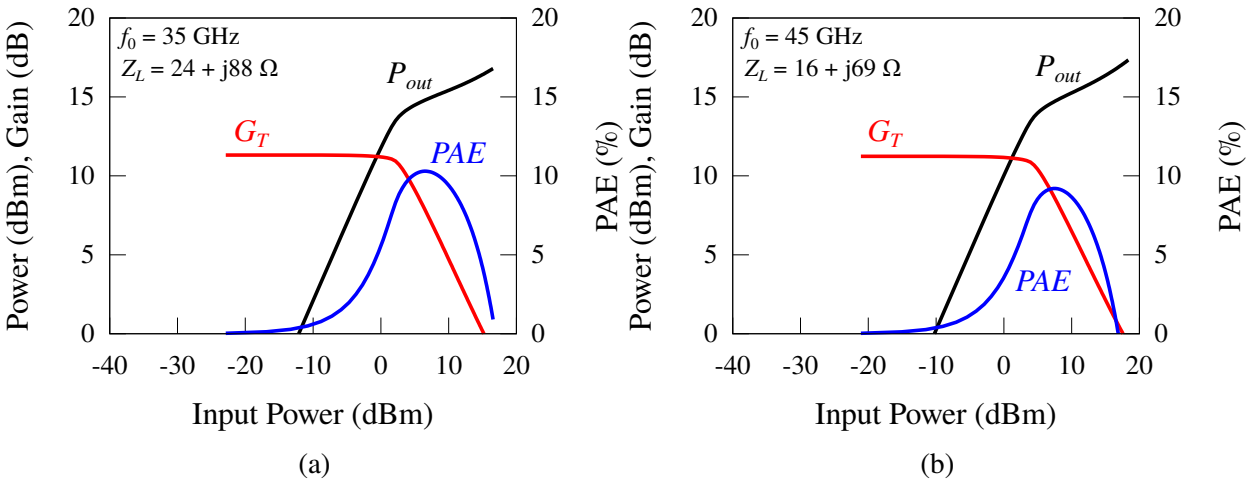


Figure 6.27: Simulated input power sweeps of the GaN HEMT biased as a class AB amplifier (gate-source bias of -2 V and drain-source bias of 10 V) at a fundamental frequency of (a) 35 GHz and (b) 45 GHz. The optimal loads determined from the LP data are $24 + j88\,\Omega$ at 35 GHz and $16 + j69\,\Omega$ at 45 GHz. The peak *PAE* resides at an input power of approximately 8 dBm at both 35 GHz and 45 GHz. The *PAE* of the class AB amplifier does not diminish like the class A amplifier at high frequency. At 45 GHz, the peak *PAE* of the device is approximately 8 – 9%.

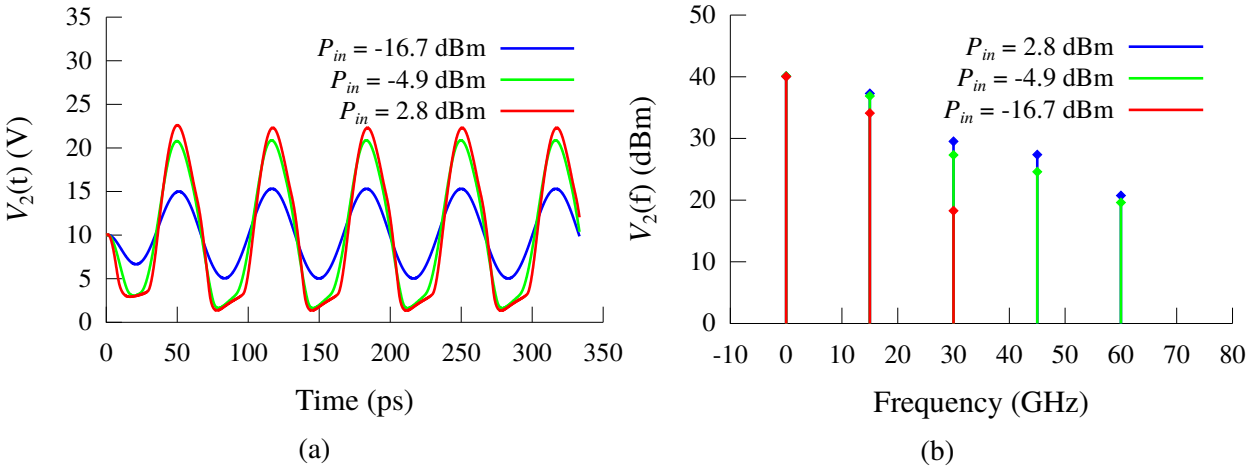


Figure 6.28: Simulated (a) output voltage transients and (b) output voltage spectral content of the GaN HEMT biased as a class AB amplifier at 15 GHz. The input powers of the blue, green, and red lines are -16.7 dBm, -4.9 dBm, and 2.8 dBm, respectively. Strong inter-modulation is observed at each input power.

demonstrates 10% *PAE* while the class A amplifier only demonstrates between 2 and 3%. It is concluded that the class AB amplifier is preferred for RF PA design due to its higher efficiency.

Several transient output voltages of the input power sweep at 15 GHz are presented in Figure 6.28a. Figure 6.28b illustrates the spectral content of the output voltages. The blue, green, and

red lines correspond to the input powers $P_{in} = -16.7$ dBm, -4.9 dBm, and 2.8 dBm, respectively. Significant inter-modulation is observed in the response of the class AB amplifier at each input power.

6.5.6 Quasi-Static vs. Full-Wave

Small-signal GaN HEMT results computed with the quasi-static and full-wave FKT device simulators were presented in Section 6.4. The full-wave simulations of the HEMT at high frequency exhibited vastly different parasitic effects than the quasi-static simulations. Proper treatment of high frequency parasitics is critical for accurate large-signal RF simulations. Comparisons of large-signal RF GaN HEMT simulations with the quasi-static and full-wave solvers are next presented.

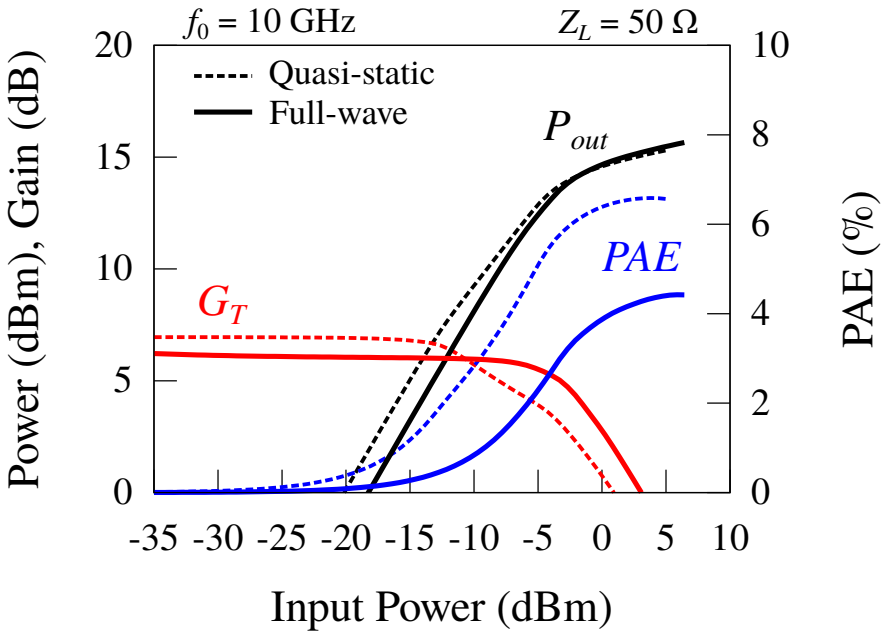


Figure 6.29: Input power sweeps of the GaN HEMT biased as a class AB amplifier calculated with the quasi-static (dashed lines) and the full-wave (solid lines) FKT device simulators. The gate-source bias is -2 V, the drain-source bias is 10 V, the fundamental frequency is 10 GHz, and the load impedance is 50 Ω.

An input power sweep of the AlGaN/GaN HEMT biased as a class AB amplifier at 10 GHz is reported in Figure 6.29. The gate-source bias is $V_{GS} = -2$ V and the drain-source bias is $V_{DS} = 10$ V. The load impedance presented to the output of the device is $Z_L = 50\,\Omega$. Quasi-static results

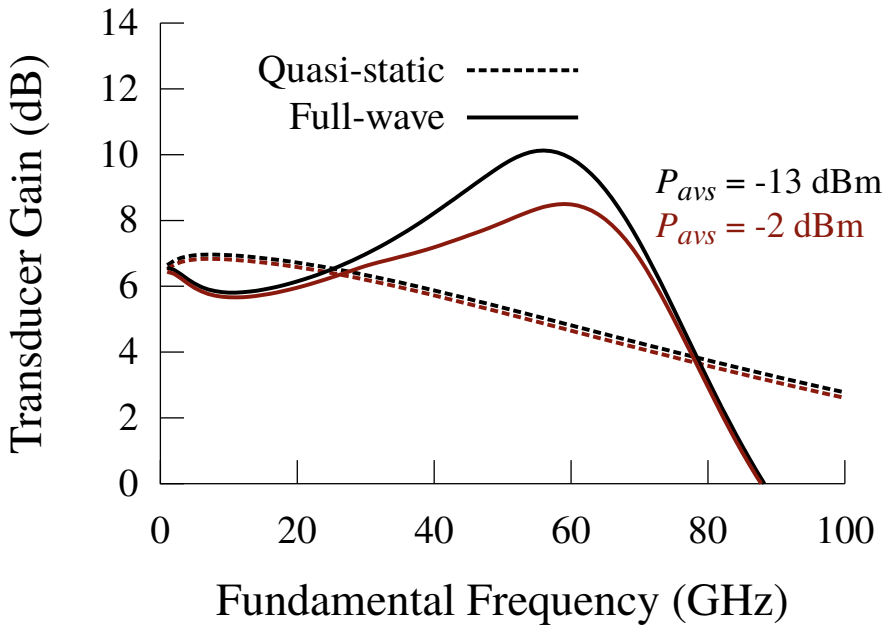


Figure 6.30: Quasi-static (dashed lines) and full-wave (solid lines) frequency sweeps of the GaN HEMT biased as a class AB amplifier. The gate-source bias is -2 V, the drain-source bias is 10 V, and the load impedance is 50 Ω. Black lines correspond to a power available from the source of -13 dBm and red lines correspond to a power available from the source of -2 dBm.

are illustrated with dashed lines and full-wave results are illustrated with solid lines. The full-wave result has a significantly different 1 dB compression point at $P_{in} = -4$ dBm compared to the quasi-static 1 dB compression point at $P_{in} = -10$ dBm. Furthermore, the *PAE* computed with the full-wave solver is vastly different than the *PAE* computed with the quasi-static solver. Clearly, proper treatment of full-wave EM is essential for accurately capturing the high frequency response of devices.

Figure 6.30 presents a set of simulated frequency sweeps of the same class AB amplifier. The fundamental frequency of the large-signal sinusoidal input tone is swept from 1–100 GHz. Quasi-static simulation results are illustrated with dashed lines and full-wave simulation results are illustrated with solid lines. Black lines correspond to a power available from the source of $P_{avs} = -13$ dBm and red lines correspond to $P_{avs} = -2$ dBm. As expected, the full-wave simulations exhibit a vastly different frequency response compared to the quasi-static simulations. The peak transducer gain of the HEMT is located at a fundamental frequency of approximately 55 GHz

according to the full-wave results. Quasi-static simulations do not provide the same qualitative features in the gain.

6.6 Discussion and Conclusions

This chapter applied the FKT TCAD framework to small- and large-signal RF simulations of an AlGaN/GaN HEMT. First, a detailed description of the HEMT was presented along with simulated thermal equilibrium results (i.e., no external bias). Next, static FKT simulations of the GaN HEMT were presented and compared to measurements [107]. Specifically, Figure 6.6 presented the static I-V family computed with the FKT device simulator. These results demonstrated excellent agreement with measurements. Figure 6.7 presented an interesting comparison of the I-V families calculated with different versions of the FKT device simulator. This comparison demonstrated that electronic band-structure, hot-electron effects, self-heating effects, and trapping effects all require proper treatment in the simulator in order to re-produce measured data.

The third section of this chapter focused on small-signal RF device simulations with the FKT TCAD tool. A detailed description was provided for simulating S-Parameters of electronic devices. Both quasi-static and full-wave simulation results were provided. The full-wave results differ from quasi-static at high frequency. It is concluded that the full-wave solver is essential for capturing the high frequency parasitics of devices. The main result of the small-signal RF section, which was first calculated and presented in [84], is the small-signal current gain in Figure 6.12. The simulations show excellent agreement with the measurements. The upward bending of the current gain is due to the inductive coupling of the metal contacts. This high frequency parasitic effect is captured with the full-wave device solver.

Large-signal RF simulations of the GaN HEMT were next presented in this chapter. Because large-signal RF simulations require reactive load impedances, the beginning of the section focused on the quasi-static AC impedance BC and its application as a series RL load. Next, a detailed description of the post-processing required to calculate PA FOMs was presented. Finally, several large-signal simulations of the GaN HEMT were reported for different amplifier classes and op-

erating frequencies. The LP data calculated from the device simulations were used to determine

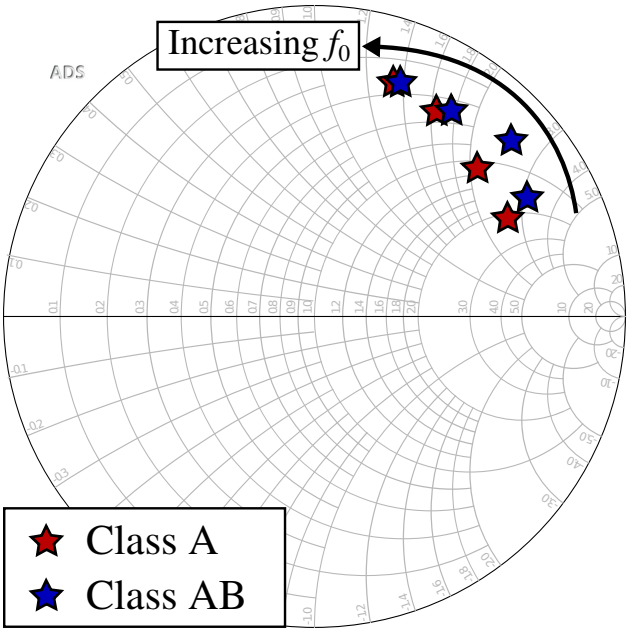


Figure 6.31: Summary of the optimal loads determined through FKT device simulations of the AlGaN/GaN HEMT biased as class A and class AB amplifiers. The fundamental frequency of both amplifiers was swept from 15–45 GHz with 5 GHz increments. The optimal load shifts in a counter-clockwise motion across the Smith chart with increasing frequency.

the optimal loading conditions. Figure 6.31 summarizes the optimal load reflection coefficients determined by the FKT device simulations. The fundamental frequencies were 15, 25, 35, and 45 GHz for both the class A and class AB amplifiers. The optimal reflection coefficients shift in a counter-clockwise motion across the Smith chart with increasing fundamental frequency. Reflection coefficients corresponding to the class A amplifier are illustrated with red stars and blue stars represent the class AB amplifier's optimal reflection coefficients. Input power sweep simulations were then presented using the load impedances corresponding to peak transducer gain.

Finally, the large-signal RF response of the GaN HEMT was simulated with the quasi-static and full-wave FKT device solvers in Section 6.5.6. The full-wave results exhibited different characteristics compared to the quasi-static results. This is due to the high frequency parasitics of the device which are not captured with the quasi-static solver. Full-wave simulations will be paramount for analysis of next generation devices operating at high frequency.

The FKT device simulator could be a valuable asset for high frequency design of GaN HEMT PAs. This chapter was devoted to demonstrating its use through a wide range of simulations that could be useful for RF circuit design. LP measurements at high frequency are expensive. The FKT device simulator is therefore an excellent candidate for calculating supplemental high frequency and high power LP data.

CHAPTER 7

CHARACTERIZATION AND COMPACT MODEL EXTRACTION OF GAN HEMTS

7.1 Introduction

The culmination of this thesis is the characterization and compact model extraction of a state-of-the-art AlGaN/GaN HEMT fabricated and measured at AFRL. This chapter is arranged as follows. In the first section, a description of the GaN HEMT is presented. Characterizations of the HEMT, including small- and large-signal simulations, are reported in the next several sections. Finally, an XP model is extracted from FKT device simulations and used for PA design in ADS.

7.2 AFRL AlGaN/GaN HEMT

The AlGaN/GaN HEMT is fabricated and measured at AFRL. All measurements of this device were provided by Bob Fitch AFRL and Gregg Jessen AFRL [6]. Figure 7.1 presents an illustration of the HEMT stack. The gate length of the device is 0.14 μm and the width is 75 μm. Total

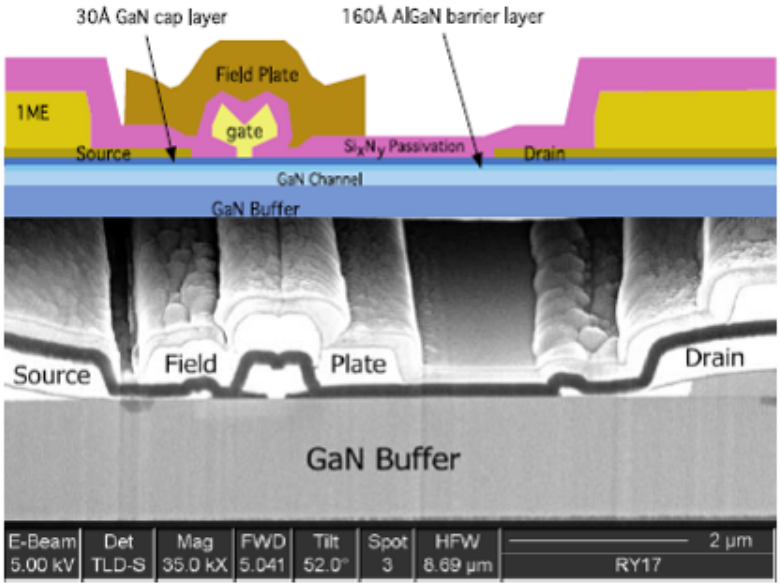


Figure 7.1: A cross-section of AFRL's standard GaN HEMT device.

periphery of the device is 300 μm — four gate fingers $\times$ 75 μm unit gate length. The stack of the HEMT consists of several layers. A 1.9 μm GaN buffer is first grown on the SiC substrate. The remaining stack consists of a 10 nm GaN channel, 1 nm of aluminum nitride (AlN), 15.4 nm of $Al_{0.278}Ga_{0.722}N$, and a 3 nm GaN cap.

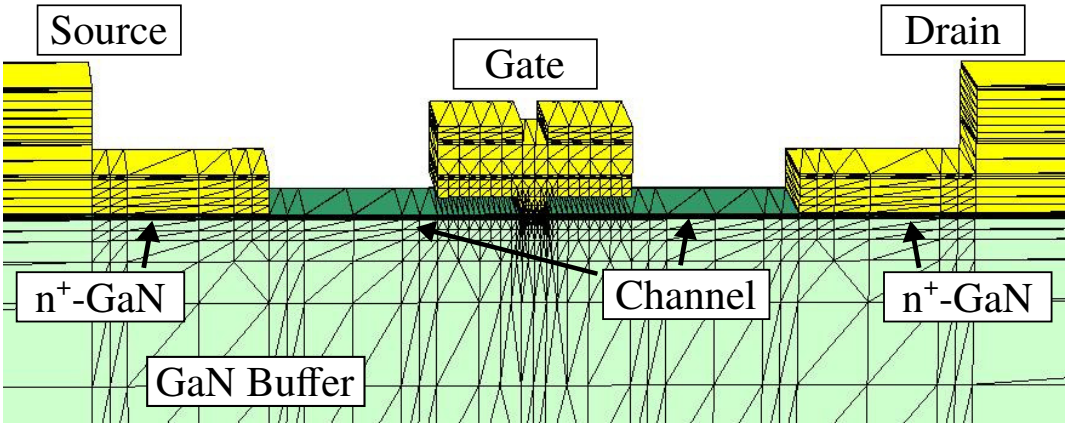


Figure 7.2: The DV mesh used to simulate the AFRL GaN HEMT in the FKT device simulator.

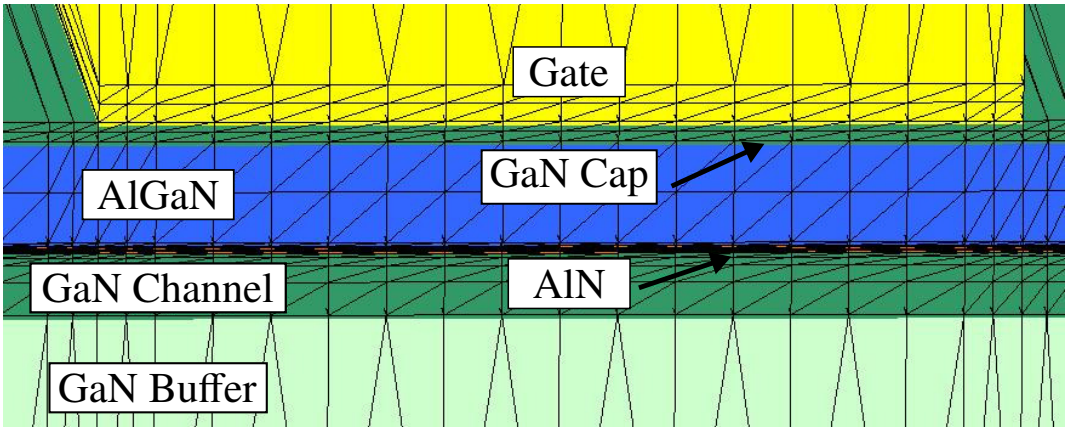


Figure 7.3: A zoom of the GaN HEMT stack under the Schottky gate. The stack consists of (from substrate to metal) a 1.9 μm GaN buffer, a 10 nm GaN channel, 1 nm of AlN, 15.4 nm of AlGaN, and a 3 nm GaN cap. The mole fraction of the AlGaN layer is $x = 0.278$.

The DV mesh on which the FKT device equations are solved is presented in Figures 7.2 and 7.3. This was generated with in-house meshing codes developed by Matt Grupen AFRL to simulate high speed electronic devices [62, 84]. These codes were developed following the work of Conti [46] described in Section 2.2. The silicon nitride (SiN) passivation on top of the GaN

cap and the air above the passivation layer are not illustrated in Figure 7.2. The n^+-GaN regions in Figure 7.2 mimic the contact resistances of the source and drain Ohmic metal-semiconductor interfaces.

As discussed previously, GaN HEMT technology is an excellent transistor choice for PAs due to its high power and high frequency capabilities. The HEMT does not require modulation doping because of spontaneous and piezoelectric polarization [3, 5]. The polarization charge densities, however, are not necessarily known after fabrication and must be determined for the FKT device simulations. The following sections on characterization of the HEMT discuss the determination of the unknown physical parameters. The physical parameters of the AFRL GaN HEMT are presented in Figure 7.4.

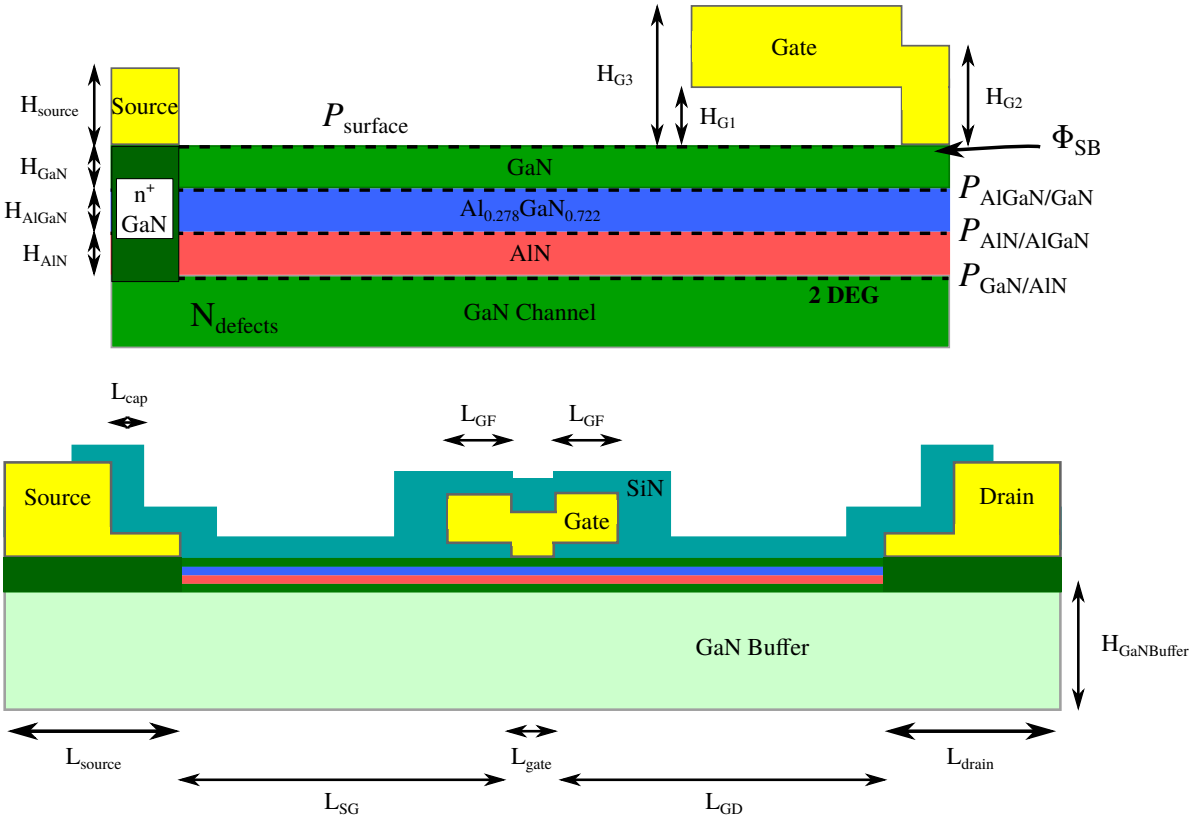


Figure 7.4: The physical parameters of the AFRL GaN HEMT. Most notable are the polarization charge densities P at the heterojunctions.

7.3 TLM Characterization

The unknown physical parameters of the AFRL GaN HEMT can be determined through what are known as transmission line measurements (TLM) [110]. For TLM analysis, the Schottky gate of the HEMT is removed to yield a resistor. A series of GaN resistors with variable spacing

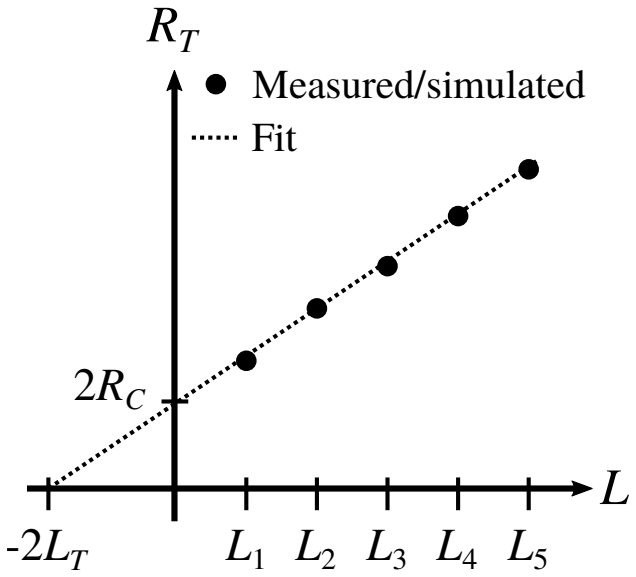


Figure 7.5: An example of determining the approximate access resistance with TLM. A series of resistors are measured/simulated and the total resistance of each resistor is calculated. The slope of the data determines the sheet resistance (access resistance) divided by the width of the resistors and the y-intercept determines twice the contact resistance.

between the Ohmic source and drain contacts are measured/simulated. The total resistance of each device is calculated with Ohm's law. An illustration of the set of total resistances is provided in Figure 7.5. The total resistance is given by

$$R_T = \frac{R_S}{W}\left(L + 2L_T\right), \tag{7.1}$$

where R_S is the sheet resistance of the 2DEG, L is the length between the Ohmic contacts, L_T is the transfer length, and W is the width of the resistors [110]. A linear fit of the data determines the slope and y-intercept. The sheet resistance (also called the access resistance) is calculated from the slope of the data by

$$R_S = \frac{\Delta R_T(L)}{\Delta L} \times W, \tag{7.2}$$

and the contact resistance (the resistance due to the Ohmic metal-semiconductor interface) is calculated from the y-intercept by

$$R_C = \frac{1}{2} R_T(L = 0). \tag{7.3}$$

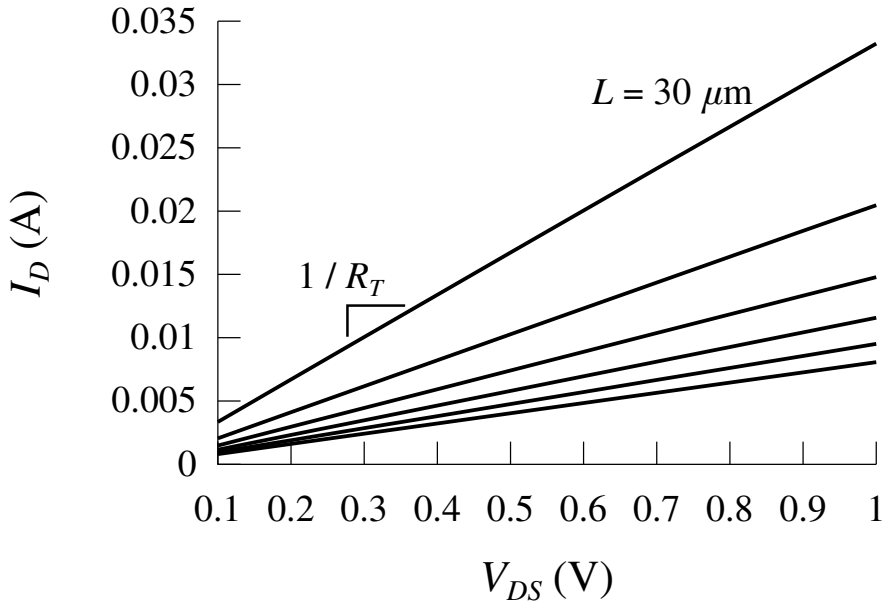


Figure 7.6: Simulated I-V curves of the TLM resistors computed with the FKT device solver. The lengths of the resistors vary from 5 μm to 30 μm with 5 μm increments. The I-V curve with the largest drain current at the drain-source bias of 1 V corresponds to the $L = 30\,\mu$m resistor.

The polarization sheet charge densities at the heterojunction interfaces are tuned to produce the best agreement between the simulated and measured total resistances of the GaN TLM resistors. To this end, several GaN resistors are simulated with the FKT device solver. The lengths of the resistors vary between 5 μm and 30 μm with 5 μm increments. The width of each resistor is 81 μm. Figure 7.6 presents the simulated I-V curves of the resistors. The drain-source bias, V_{DS}, is swept from 0 to 1 V, however, the resistance is calculated at low bias.

The total resistances of the devices are presented in Figure 7.7. The approximate access resistance and contact resistance calculated from the data are $R_S \approx 304\,\Omega$ and $R_C \approx 5.5\,\Omega$. These simulated results agree well with measurements and provide a good physical parameter set for RF simulations with the FKT TCAD tool.

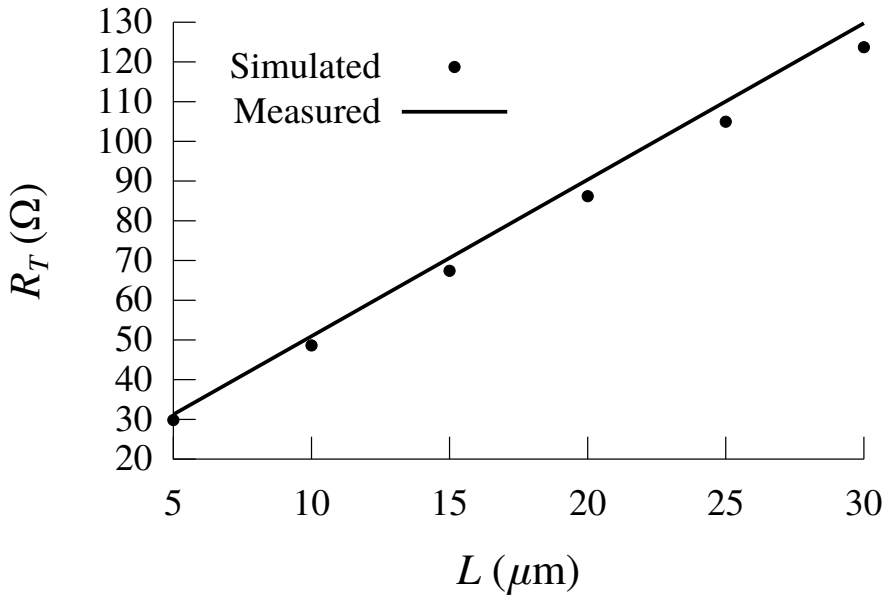


Figure 7.7: The simulated (black dots) and measured (solid line) total resistances of the GaN TLM resistors.

7.4 DC and Small-Signal Characterization

The initial physical parameter set determined by the TLM simulations is tuned for agreement with DC and small-signal measurements. DC I-V measurements provided by AFRL [6] are first used to tune the polarization charge densities at the heterojunctions. To improve the accuracy of the simulations, the polarization charge densities are allowed to vary along the channel. This provides additional DOF to fit the simulated I-V data onto the measurements. The variation of the 2DEG density across the GaN HEMT channel is presented in Figure 7.8. The source contact ends at 0.5 μm and the drain contact begins at 3.5 μm in the plot. The Schottky T-gate resides at the middle of the plot and the gate arms extend outward. The channel is approximated with five piece-wise polarization charge densities. The first polarization charge density covers the region between the source contact and the beginning of the T-gate arm. The second, third, and fourth regions follow the dimensions of the Schottky T-gate, i.e., the second starts at the beginning of the Schottky arm and ends at the beginning of the gate stem (the portion of the gate that is attached to the GaN cap). Finally, the fifth region extends from the end of the Schottky gate arm to the

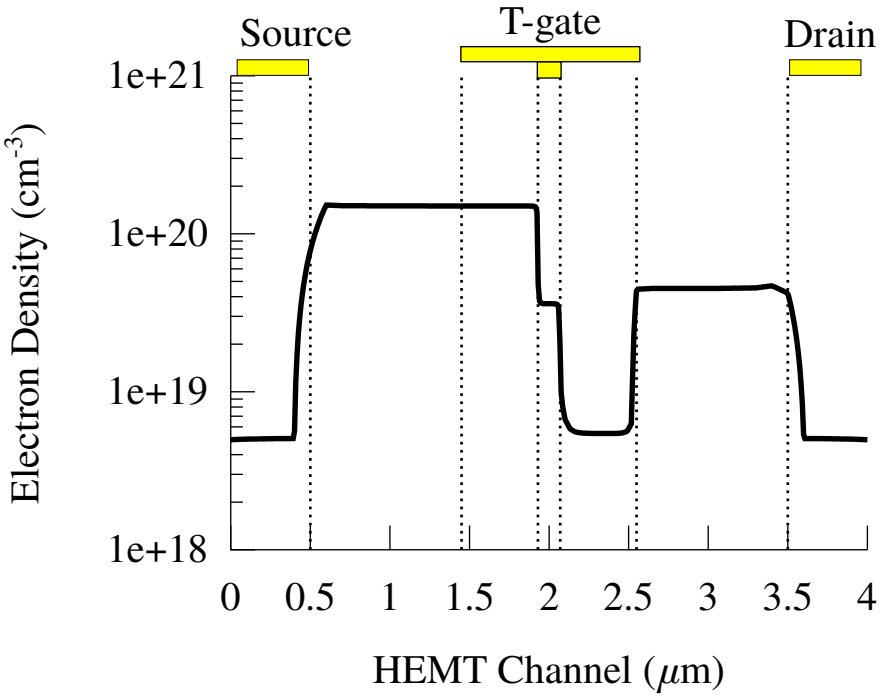


Figure 7.8: Electron density across the GaN HEMT channel. The channel begins at 0.5 μm and ends at 3.5 μm. Five different polarization charge densities are prescribed to the GaN/AlN heterojunction. The five regions are outlined with vertical dashed lines. The contacts of the device are illustrated at the top of the figure.

beginning of the drain contact. These regions are illustrated in Figure 7.8 with dotted lines.

The thermal equilibrium band diagram of the HEMT computed with the FKT device simulator is presented in Figure 7.9. The conduction band minimum is plotted with a black line and the electron Fermi level is plotted with a red line. This band diagram is calculated across the stack (from the GaN buffer to the Schottky gate). The 2DEG resides at 5 μm — the GaN/AlN interface.

The simulated I-V family of the AFRL GaN HEMT is presented in Figure 7.10. Here, the black lines represent the data computed with the FKT device simulator and the red dots represent the AFRL measurements. The drain-source bias, V_{DS}, is swept from 0 to 10 V and the gate-source bias, V_{GS}, is swept from +1 V to -3 V with 1 V steps. FKT simulations exhibit good agreement with the measurements. One physical effect that is not captured by the FKT device simulator is called the "kink effect". It is suggested that the kink effect could be induced by hot-electron trapping and field-assisted de-trapping via donor-like traps in the GaN buffer layer [111]. This trapping process

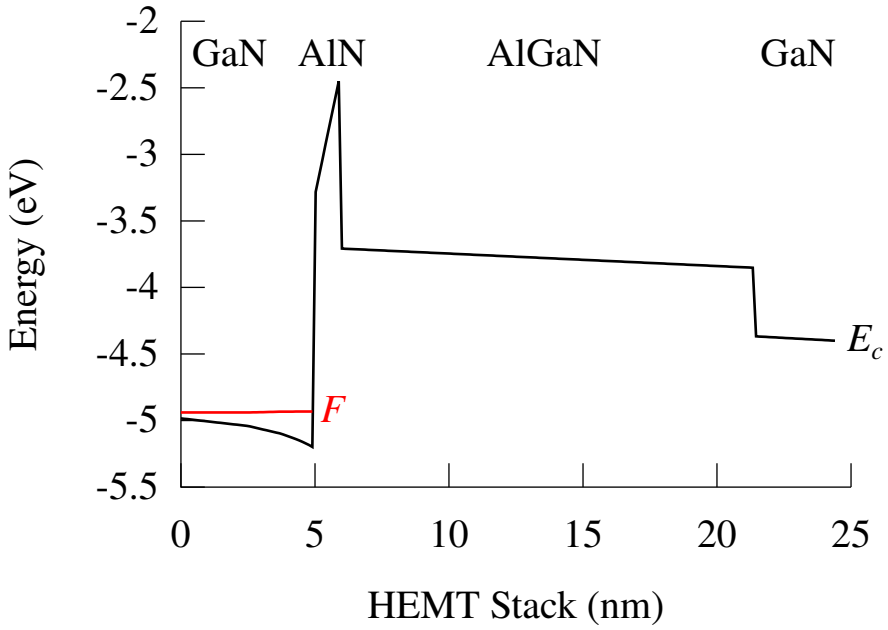


Figure 7.9: The thermal equilibrium band diagram of the AFRL GaN HEMT computed with the FKT device simulator.

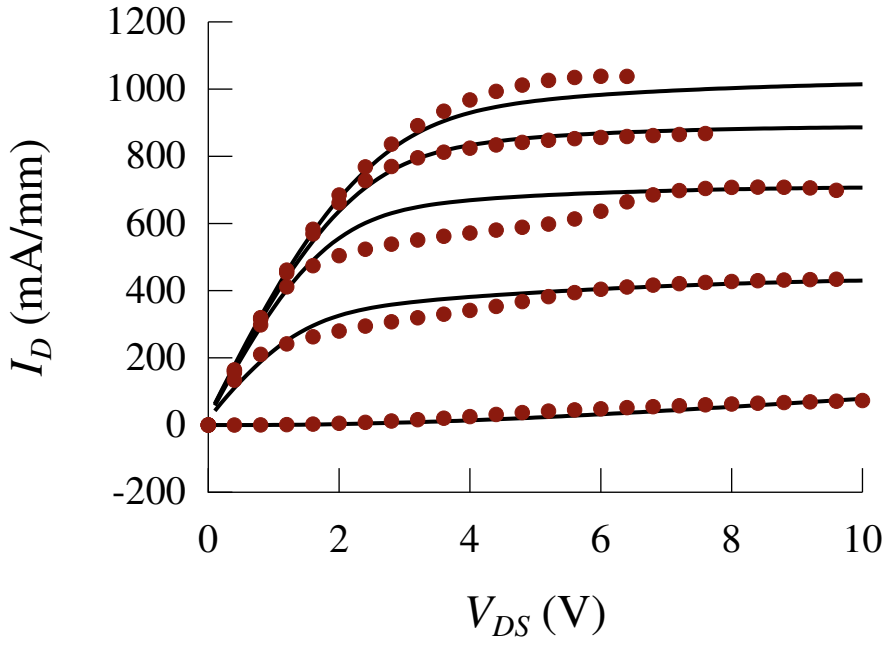


Figure 7.10: Simulated DC I-V curves (black lines) compared with AFRL measurements (red dots). The gate-source bias ranges from +1 V to -3 V with 1 V steps.

is much slower than the fundamental period of RF signals. Therefore, proper treatment of the kink effect is not required large-signal RF simulations of the GaN HEMT.

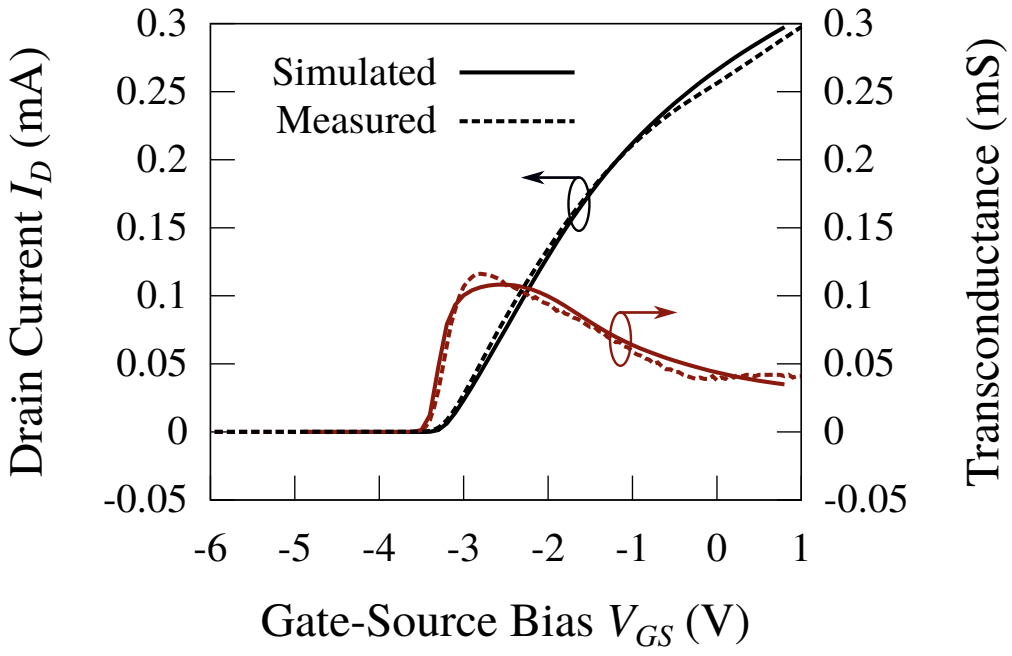


Figure 7.11: Simulated DC drain current (solid black line) and transconductance (solid red line) of the AFRL GaN HEMT. The measurements obtained from AFRL are plotted with dashed lines.

Figure 7.11 illustrates the drain current (black lines) and the transconductance (red lines) versus the gate-source bias, V_{GS}, of the AFRL GaN HEMT. FKT device simulations are drawn with solid lines and the AFRL measurements are drawn with dashed lines. FKT device simulations exhibit good agreement with the measurements.

The S-Parameters S_{11} and S_{22} are presented in Figure 7.12. The solid lines represent the results computed with the quasi-static FKT device simulator. Measured S-Parameters are drawn with dashed lines. Simulations agree well with measurements at lower frequency. At higher frequency, however, the simulations lose accuracy. Full-wave simulations are required for accuracy at high frequency.

The AFRL GaN HEMT is further characterized by simulating the small-signal current gain. Figure 7.13 presents the simulated h_{21} (black line) and the measured small-signal current gain (red boxes). FKT simulations exhibit more gain than the measured device at lower frequency. Furthermore, the slope of the measured current gain changes around 20–30 GHz. This is due to inductive coupling of the metal contacts. High frequency parasitic effects, including the inductive

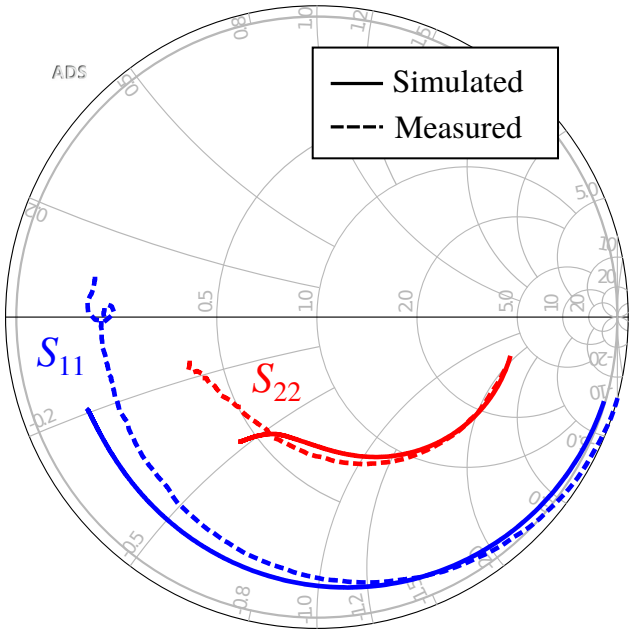


Figure 7.12: S-Parameters of the AFRL GaN HEMT. The solid lines represent the results computed with the quasi-static FKT device solver. The measured S-Parameters obtained from AFRL are drawn with dashed lines.

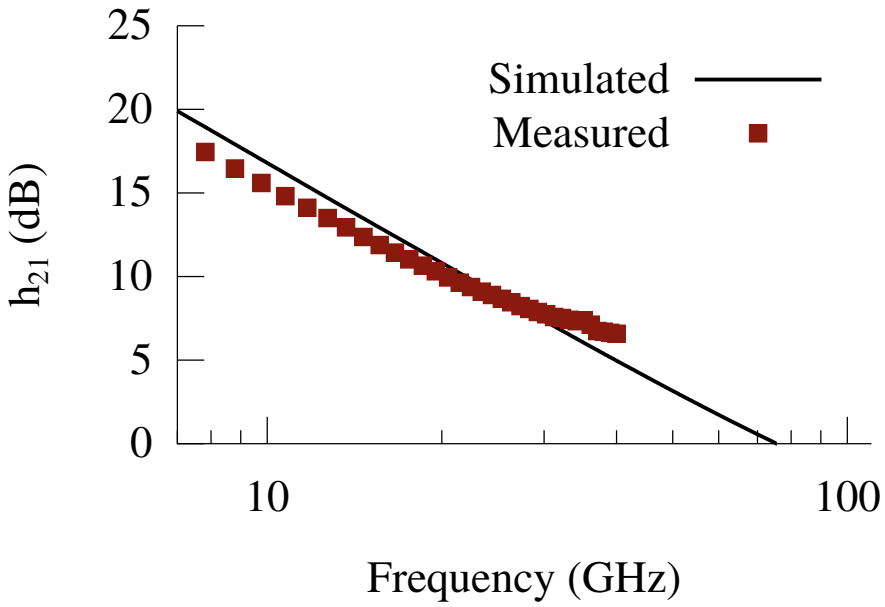


Figure 7.13: The small-signal current gain of the AFRL GaN HEMT. The FKT device simulation data (black line) compares reasonably well with the measurements (red boxes).

coupling of the metal contacts, can be captured with the full-wave FKT device simulator. This is the topic of future research.

7.5 Large-Signal Characterization

Large-signal simulations of the AFRL GaN HEMT with the FKT device simulator are presented in this section. The device was measured at AFRL. The HEMT operates as a class AB amplifier at 10 GHz.

An input power sweep of the AFRL GaN HEMT is illustrated in Figure 7.14. The simulated output power, P_{out}, is plotted with a black line, the simulated transducer gain, G_T, is plotted with a red line, and the simulated *PAE* is plotted with a blue line. Measurements of P_{out}, G_T, and *PAE* are presented with black, red, and blue squares, respectively. The gate-source bias is $V_{GS} = -2.8$ V and the drain-source bias is $V_{DS} = 21$ V. The load reflection coefficient is $\Gamma_L = 0.37 + j0.60\,\Omega$ at 10 GHz. The reflection coefficient looking into the source is $\Gamma_S = -0.27 + j0.61\,\Omega$ at 10 GHz. The FKT device simulations presented in Figure 7.14 agree well with measurements

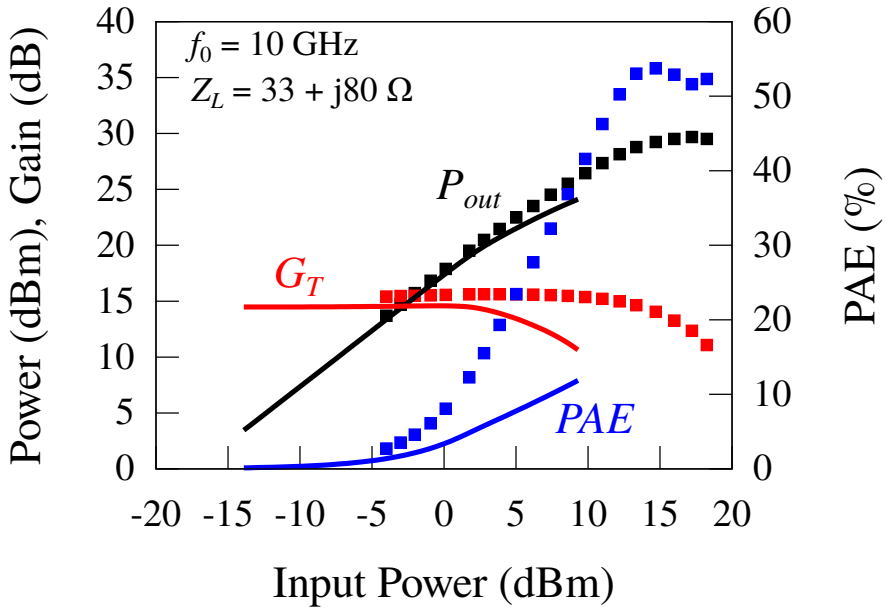


Figure 7.14: Input power sweep data of the AFRL GaN HEMT biased as a class AB amplifier at 10 GHz. The output power is plotted in black, the transducer gain is plotted in red, and the *PAE* is plotted in blue. FKT device simulation data are plotted with solid lines and measurements are plotted with squares.

at low input power. However, the simulations deviate from the measurements at higher input power. Most noticeably, the simulations reach compression at a significantly lower input power

than the measurements. This could be due to the load impedance presented to the output of the device. The simulations use a series RL load impedance, but this may not be the precise impedance that is experimentally presented to the HEMT. Full-wave FKT device simulations may also be required for better agreement with power sweep measurements. Simulated and measured *PAE* also deviate at higher input power. The FKT device simulator overestimates the static drain current at $V_{DS} = 21$ V leading to a lower simulated efficiency.

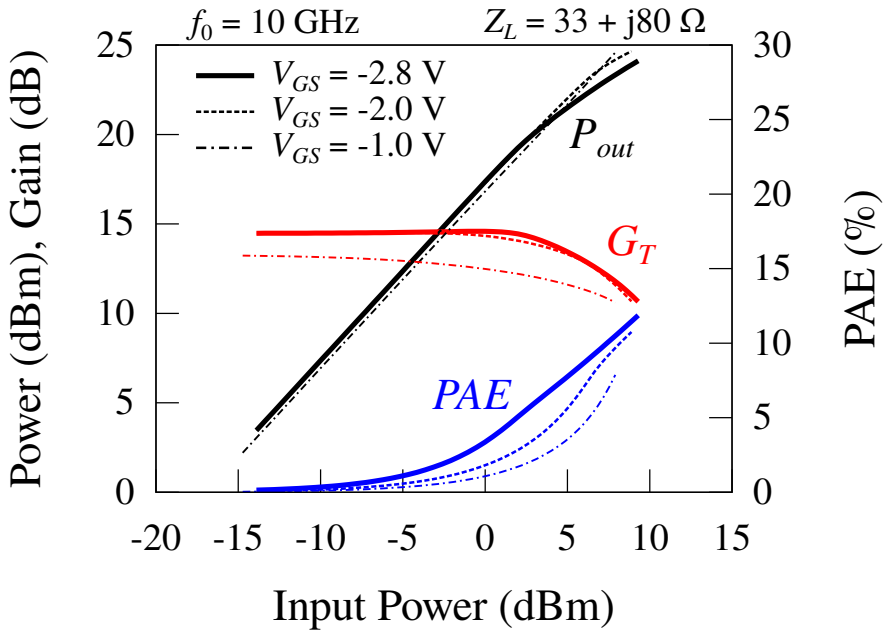


Figure 7.15: Simulated input power sweeps of the AFRL GaN HEMT at 10 GHz. The gate-source bias is changed from -2.8 V to -2 V (dashed lines) and -1 V (dot-dashed lines). The drain-source bias is held constant at 21 V. The gain and *PAE* of the device is reduced at -1 V gate-source voltage, as expected.

Large-signal characterization of the AFRL GaN HEMT continues by analyzing the power sweep results of the device with different operating points. Figure 7.15 illustrates simulated input power sweeps of the HEMT with different gate-source biases, V_{GS}. The output power, P_{out}, transducer gain, G_T, and the *PAE* are plotted with black, red, and blue lines, respectively. The fundamental frequency is 10 GHz and the load and source reflection coefficients are $\Gamma_L = 0.37 + j0.60\,\Omega$ at 10 GHz and $\Gamma_S = -0.27 + j0.61\,\Omega$ at 10 GHz, respectively. In Figure 7.15, the gate-source bias V_{GS} is changed from -2.8 V to -2 V (dashed lines) and -1 V (dot-dashed lines). As expected, the

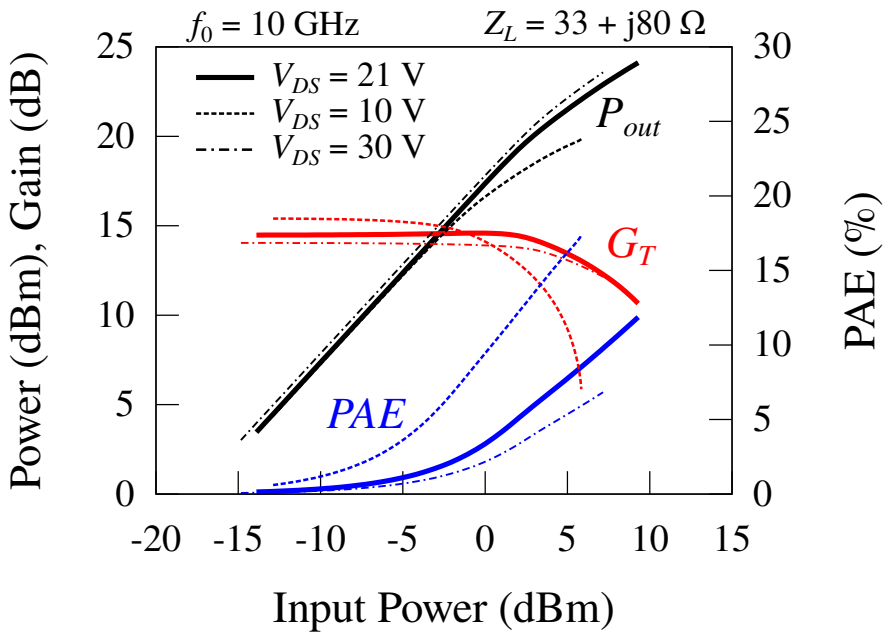


Figure 7.16: Simulated input power sweeps of the AFRL GaN HEMT at 10 GHz. The drain-source bias is changed from 21 V to 10 V (dashed lines) and 30 V (dot-dashed lines). The gate-source bias is held constant at -2.8 V. The small-signal gain of the device increases at 10 V drain-source bias. However, the device operating at this bias compresses at a lower input power.

PAE decreases as the gate-source bias increases. A larger drain current, I_D, results from a larger gate-source bias. An increase in the DC power consumed at the drain results in a decrease in the *PAE*. The quantitative decrease in the device efficiency determined by the FKT device simulations may be useful for RF circuit designers.

Figure 7.16 presents several input power sweep simulations with varying drain-source bias, V_{DS}. The output power, P_{out}, transducer gain, G_T, and the *PAE* are plotted with black, red, and blue lines, respectively. The fundamental frequency and loading conditions are the same as the previous results. Here, the drain-source bias, V_{DS}, is changed from 21 V to 10 V (dashed lines) and 30 V (dot-dashed lines). Decreasing V_{DS} has substantial effect on each device FOM. The low power transducer gain increases and the device compresses at a lower input power compared to $V_{DS} =$ 21 V. The *PAE* also increases because the DC drain current is reduced. FKT device simulations of the HEMT with a drain-source bias $V_{DS} = 30$ V exhibit similar effects. The transducer gain and *PAE* are lower than the $V_{DS} = 21$ V simulations.

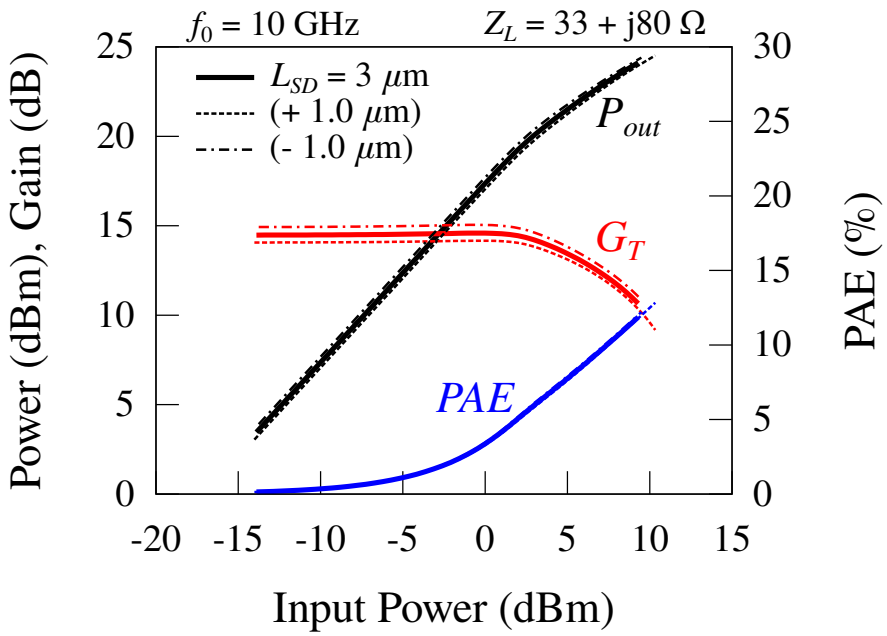


Figure 7.17: Simulated input power sweeps of the AFRL GaN HEMT at 10 GHz. The source-drain spacing is changed from 3 μm to 4 μm (dashed lines) and 2 μm (dot-dashed lines). These variations of the source-drain spacing have small impacts on the device power metrics.

The effect of the source-drain length, $L_{SD} = L_{SG} + L_{\text{gate}} + L_{GD}$ (see Figure 7.4), on the large-signal power metrics is investigated with the results in Figure 7.17. The output power, transducer gain, and *PAE* are calculated from three FKT device simulations with different spacing between the Ohmic contacts. The first simulation (solid lines) is the nominal source-drain length of $L_{SD} = 3.0\,\mu$m and the other two add/subtract 1.0 μm. The $L_{SD} = 4.0\,\mu$m input power sweep results are plotted with dashed lines and the $L_{SD} = 2.0\,\mu$m input power sweep results are plotted with dot-dashed lines. The transducer gain of the device is reduced by increasing the source-drain distance. In a similar manner, reducing the source-drain distance increases the transducer gain. However, a substantial decrease in the source-drain spacing (-33.3%) only moderately increases the low input power gain from 14.48 dB to 14.93 dB — a 3.1% increase. The +33.3% increase in the source-drain spacing yields a 2.9% decrease in gain from 14.48 dB to 14.06 dB. These simulation results could be valuable for RF circuit designers. Furthermore, measuring these results would require fabricating several more devices.

Finally, LP data are generated from FKT device simulations of the AFRL HEMT. Figure 7.18

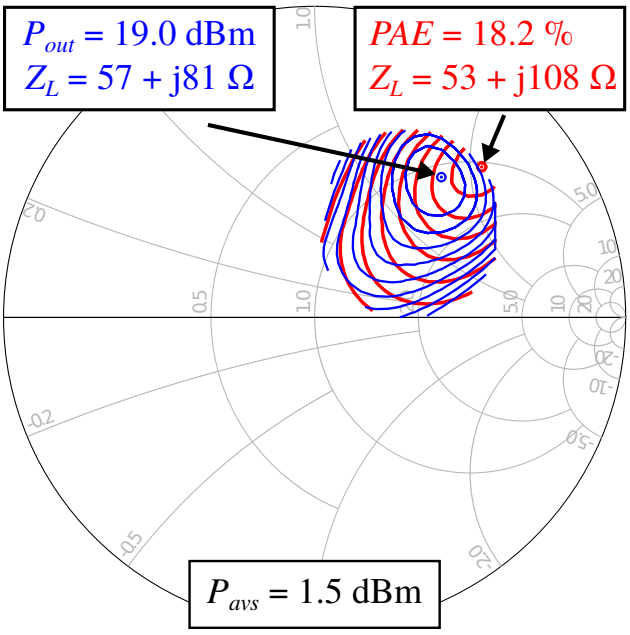


Figure 7.18: LP data generated from FKT device simulations of the AFRL GaN HEMT biased as a class AB amplifier at 10 GHz. The gate-source bias is -2.8 V, the drain-source bias is 10 V, and the power available from the source is 1.5 dBm. Output power contours are illustrated with blue lines and *PAE* contours are illustrated with red lines. The optimal load impedance for peak output power is $57 + j81\,\Omega$ at 10 GHz and the load impedance $53 + j108\,\Omega$ at 10 GHz corresponds to peak *PAE*. The peak output power is 19 dBm and the peak *PAE* is 18.2%.

reports the LP data of the HEMT biased as a class AB amplifier at 10 GHz. The gate-source bias is $V_{GS} = -2.8$ V and the drain-source bias is $V_{DS} = 10$ V. A power available from the source of $P_{avs} = 1.5$ dBm drives the FKT device simulations. The output power contours are illustrated with blue lines and the *PAE* contours are illustrated with red lines. Based on the LP data, the optimal load impedance for the peak output power of $P_{out} = 19$ dBm is $Z_L = 57 + j81\,\Omega$ at 10 GHz. A maximum *PAE* of 18.2% is achieved with a load impedance of $Z_L = 53 + j108\,\Omega$ at 10 GHz.

7.6 XP Extraction

Compact transistor models allow efficient PA design. Models can be imported into commercial CAD tools and used for simulated impedance matching, LP, power sweeps, etc. An example of a powerful CAD tool is ADS. In this section, the procedure for extracting an XP model with the FKT device simulator is presented and discussed.

The procedure for extracting a general XP model was presented in Section 2.5. A two-port model with K harmonics requires $4K-1$ simulations to calculate the $2K \times (4K-1)$ XP coefficients. This work extracts an XP model of the AFRL GaN HEMT biased as a class AB amplifier at 10 GHz. The gate-source bias is $V_{GS} = -2.8$ V and the drain-source bias is $V_{DS} = 10$ V. The large-signal input tone at the fundamental frequency has a power available from the source of $P_{avs} = 11.6$ dBm. Large-signal reflection coefficients looking into the source and into the load are $\Gamma_S = 50\,\Omega$ at 10 GHz and $\Gamma_L = 50\,\Omega$ at 10 GHz, respectively. These operating conditions result in an input power of $P_{in} = 4.4$ dBm.

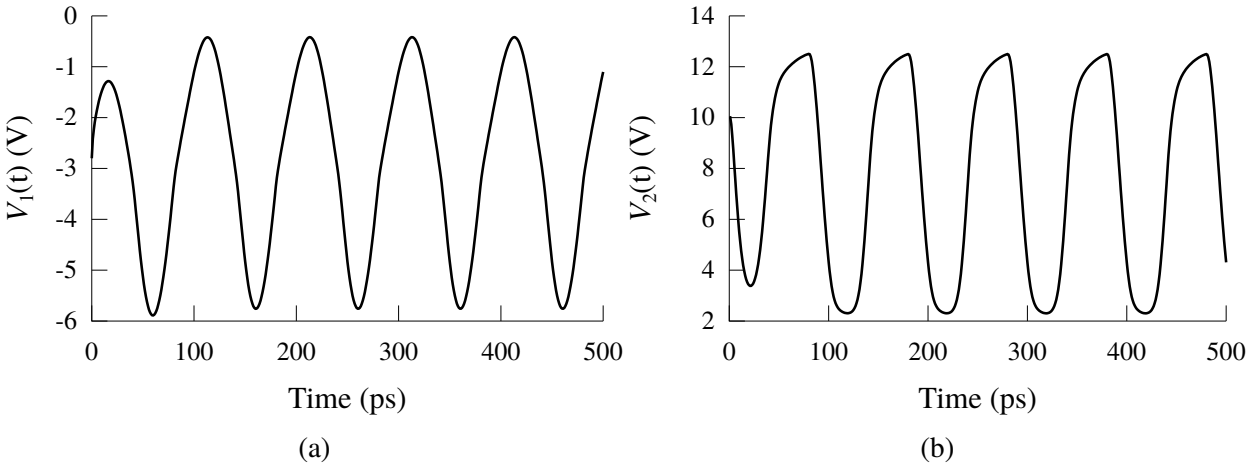


Figure 7.19: The (a) input and (b) output voltages of the AFRL GaN HEMT calculated with the quasi-static solver. The operating condition of the HEMT is a quiescent bias of a -2.8 V gate-source bias and 10 V drain-source bias. The large-signal incident wave at the fundamental frequency of the input port is 1 V in magnitude at 10 GHz.

The first XP extraction simulation applies only the large-signal tone A_{11} at the input of the device. The large-signal tone remains the same for all subsequent XP extraction simulations. The XP coefficients $\{S_{11,11},\ S_{11,21},\ \cdots,\ S_{11,K1}\}$ and $\{S_{21,11},\ S_{21,21},\ \cdots,\ S_{21,K1}\}$ are calculated with the transient voltages and currents resulting from the first FKT simulation. Figures 7.19 and 7.20 present the input and output voltages and currents of the AFRL GaN HEMT biased as a class AB amplifier with a power available from the source of $P_{avs} = 11.6$ dBm at 10 GHz. The incident harmonics, A_{pk}, and scattered harmonics, B_{pk}, calculated with Eqns. (2.51) and (2.52) are presented in Table 7.1. The transient gate voltage is chosen such that A_{11} is a pure cosine with a

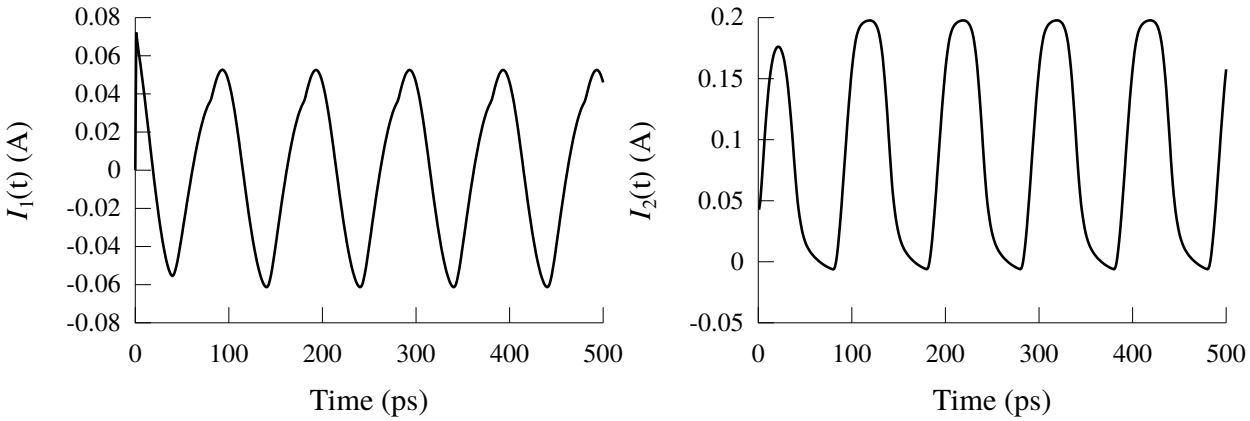


Figure 7.20: The (a) input and (b) output currents of the AFRL GaN HEMT calculated with the quasi-static solver. The operating condition of the HEMT is a quiescent bias of a -2.8 V gate-source bias and 10 V drain-source bias. The large-signal incident wave at the fundamental frequency of the input port is 1 V in magnitude at 10 GHz.

2 V amplitude. Inspection of Table 7.1 confirms that the transient gate voltage and current yield a

k	f_k (GHz)	A_{1k} (V)	A_{2k} (V)	B_{1k} (V)	B_{2k} (V)
1	10.0	$1.000\ \angle 0°$	-	$0.850\ \angle -93°$	$2.834\ \angle 117°$
2	20.0	-	-	$0.164\ \angle 116°$	$0.285\ \angle 77°$
3	30.0	-	-	$0.069\ \angle -124°$	$0.416\ \angle -173°$
4	40.0	-	-	$0.009\ \angle 153°$	$0.078\ \angle 129°$
5	50.0	-	-	$0.020\ \angle -37°$	$0.069\ \angle -77°$

Table 7.1: The incident and scattered harmonics of the AFRL GaN HEMT biased as a class AB amplifier with a power available from the source of 11.6 dBm at 10 GHz. The large-signal tone at the fundamental frequency of the input port is the only non-zero incident wave.

pure cosine large-signal incident wave A_{11}. Simulations with this form of the large-signal incident wave allow easy extraction of the XP coefficients because the phase of A_{11} is $P = 1$.

The remaining XP coefficients are calculated following the procedure described in Section 2.5. The incident harmonics are perturbed to calculate the higher-order XP coefficients. The large-signal tone A_{11} is not perturbed. Any change to A_{11} requires re-calculation of the XP model coefficients. The small-signal incident wave A_{21} is perturbed first. The GaN HEMT is simulated twice with two sets of non-zero incident waves $\{A_{11},\ A_{21}\}$ and $\{A_{11},\ A_{21}^*\}$. The tone A_{21}^* of the second perturbation simulation is 90° out of phase with the A_{21} tone of the first perturbation

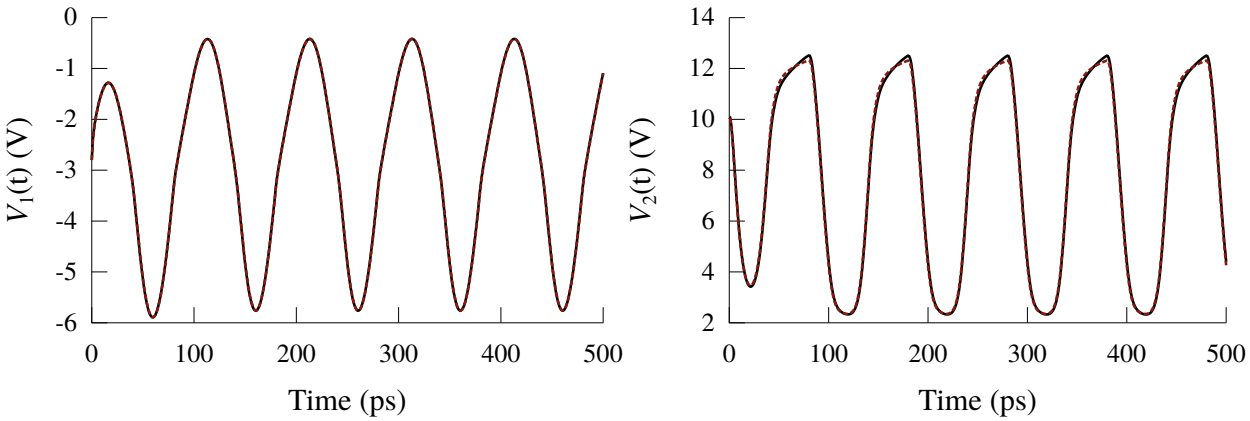


Figure 7.21: The (a) input and (b) output voltages of the AFRL GaN HEMT calculated with the quasi-static solver. The operating condition of the HEMT is a quiescent bias of a -2.8 V gate-source bias and 10 V drain-source bias. The solid black lines represent the large-signal simulation and the dashed red lines represent the first perturbation simulation of the first incident harmonic at the output port.

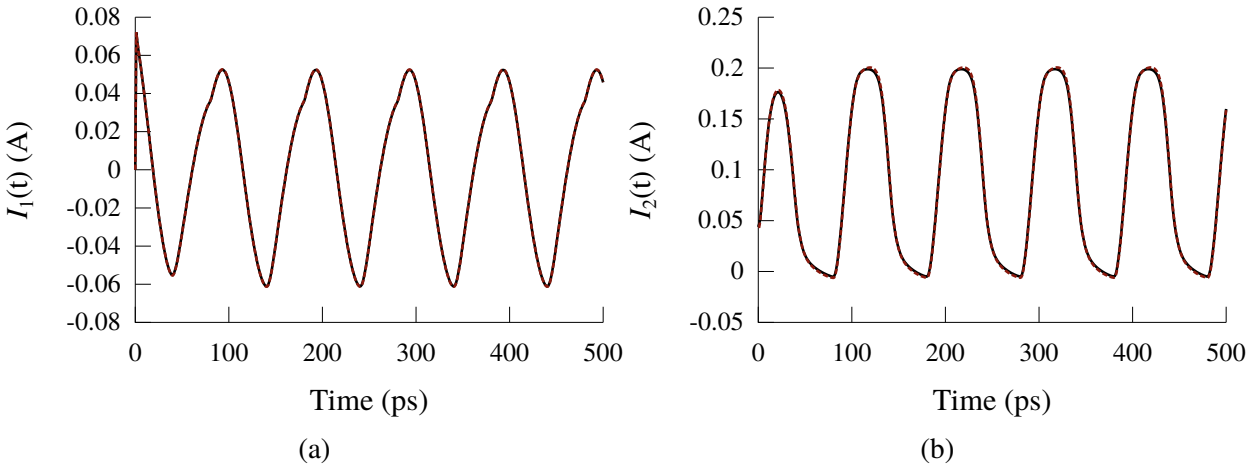


Figure 7.22: The (a) input and (b) output currents of the AFRL GaN HEMT calculated with the quasi-static solver. The operating condition of the HEMT is a quiescent bias of a -2.8 V gate-source bias and 10 V drain-source bias. The solid black lines represent the large-signal simulation and the dashed red lines represent the first perturbation simulation of the first incident harmonic at the output port.

simulation. Figure 7.21 presents the (a) input and (b) output voltages and Figure 7.22 illustrates the (a) input and (b) output currents of the two A_{21} perturbation simulations. The first perturbation applies a pure cosine as the small-signal A_{21} incident wave and the second applies a pure sine as the small-signal A_{21}^{*} incident wave. The spectral incident and scattered waves, A_{pk} and B_{pk}, of the

k	f_k (GHz)	A_{1k} (V)	A_{2k} (V)	B_{1k} (V)	B_{2k} (V)
1	10.0	$1.000 \angle 0^\circ$	$0.050 \angle 0^\circ$	$0.846 \angle -93^\circ$	$2.815 \angle 117^\circ$
2	20.0	-	-	$0.165 \angle 116^\circ$	$0.299 \angle 78^\circ$
3	30.0	-	-	$0.069 \angle -123^\circ$	$0.412 \angle -172^\circ$
4	40.0	-	-	$0.008 \angle 155^\circ$	$0.077 \angle 129^\circ$
5	50.0	-	-	$0.020 \angle -36^\circ$	$0.070 \angle -74^\circ$

Table 7.2: The incident and scattered harmonic phasors of the AFRL GaN HEMT biased as a class AB amplifier with a power available from the source of 11.6 dBm at 10 GHz. The large-signal simulation is perturbed with a small-signal tone at the second port and at the fundamental frequency.

k	f_k (GHz)	A_{1k} (V)	A_{2k} (V)	B_{1k} (V)	B_{2k} (V)
1	10.0	$1.000 \angle 0^\circ$	$0.050 \angle -90^\circ$	$0.856 \angle -93^\circ$	$2.833 \angle 118^\circ$
2	20.0	-	-	$0.165 \angle 115^\circ$	$0.290 \angle 74^\circ$
3	30.0	-	-	$0.070 \angle -124^\circ$	$0.417 \angle -172^\circ$
4	40.0	-	-	$0.009 \angle 153^\circ$	$0.078 \angle 130^\circ$
5	50.0	-	-	$0.020 \angle -38^\circ$	$0.070 \angle -78^\circ$

Table 7.3: The incident and scattered harmonic phasors of the AFRL GaN HEMT biased as a class AB amplifier with a power available from the source of 11.6 dBm at 10 GHz. The large-signal simulation is perturbed with a small-signal tone at the second port and at the fundamental frequency. This perturbation signal is out of phase with the first perturbation signal.

first and second A_{21} perturbation simulations are presented in Tables 7.2 and 7.3, respectively. For both simulations, the magnitude of the A_{21} phasor is 0.05 V. As confirmed by the data presented in Tables 7.2 and 7.3, the second perturbation is 90° out of phase with the first perturbation. The linear system solved to calculate the XP coefficients corresponding to the A_{21} terms is singular if the two perturbations are not orthogonal. Section 2.5.3 provides details for calculating the $S_{p2,m1} = \{S_{12,11}, \cdots, S_{12,51}, S_{22,11}, \cdots, S_{22,51}\}$ and $T_{p2,m1} = \{T_{12,11}, \cdots, T_{12,51}, T_{22,11}, \cdots, T_{22,51}\}$ XP coefficients given the two A_{21} perturbation simulations.

The perturbation procedure is repeated for all remaining higher-order incident harmonics, A_{pk}, to calculate their respective XP coefficients. The fifth incident harmonic at the input port, A_{15}, is included in the remaining perturbation simulations. Tables 7.4 and 7.5 present the incident and scattered wave phasors for the two orthogonal perturbation simulations. A pure cosine is applied for the first A_{15} perturbation and a pure sine is applied for the second simulation. Tables 7.4

k	f_k (GHz)	A_{1k} (V)	A_{2k} (V)	B_{1k} (V)	B_{2k} (V)
1	10.0	$1.000 \angle 0°$	-	$0.873 \angle -93°$	$2.854 \angle 117°$
2	20.0	-	-	$0.170 \angle 115°$	$0.283 \angle 73°$
3	30.0	-	-	$0.071 \angle -125°$	$0.424 \angle -175°$
4	40.0	-	-	$0.014 \angle 147°$	$0.084 \angle 120°$
5	50.0	$0.050 \angle 0°$	-	$0.041 \angle -123°$	$0.065 \angle -69°$

Table 7.4: The incident and scattered harmonic phasors of the AFRL GaN HEMT biased as a class AB amplifier with a power available from the source of 11.6 dBm at 10 GHz. The large-signal simulation is perturbed with a small-signal tone at the first port and at the fifth harmonic.

k	f_k (GHz)	A_{1k} (V)	A_{2k} (V)	B_{1k} (V)	B_{2k} (V)
1	10.0	$1.000 \angle 0°$	-	$0.873 \angle -93°$	$2.860 \angle 117°$
2	20.0	-	-	$0.170 \angle 114°$	$0.279 \angle 73°$
3	30.0	-	-	$0.075 \angle -125°$	$0.441 \angle -174°$
4	40.0	-	-	$0.012 \angle 118°$	$0.079 \angle 117°$
5	50.0	$0.050 \angle -90°$	-	$0.023 \angle 101°$	$0.095 \angle -78°$

Table 7.5: The incident and scattered harmonic phasors of the AFRL GaN HEMT biased as a class AB amplifier with a power available from the source of 11.6 dBm at 10 GHz. The large-signal simulation is perturbed with a small-signal tone at the first port and at the fifth harmonic. This perturbation is out of phase with the first perturbation at the input port and the fifth harmonic.

and 7.5 confirm that the two signals are orthogonal. The XP coefficients $S_{p1,m5} = \{S_{11,15}, \cdots, S_{11,55}, S_{21,15}, \cdots, S_{21,55}\}$ and $T_{p1,m5} = \{T_{11,15}, \cdots, T_{11,55}, T_{21,15}, \cdots, T_{21,55}\}$ are calculated with the two A_{15} perturbation simulations.

The two-port, five-harmonic XP model of the GaN HEMT biased as a class AB amplifier with a power available from the source of $P_{avs} = 11.6$ dBm at 10 GHz has 190 coefficients. The first 20 XP coefficients (sorted by the magnitude of the coefficients, $\sqrt{|S_{pq,mn}|^2 + |T_{pq,mn}|^2}$) are presented in Table 7.6. As expected, the largest coefficient, $S_{21,11}$, corresponds to the gain of the device. Because $T_{p1,m1} = 0 \; \forall (q,n)$, the fundamental scattered wave phasor at the output port is related to the incident wave phasors by

$$\begin{aligned} B_{21} = {} & S_{21,11}A_{11} + S_{21,12}A_{12} + \cdots + S_{22,1K}A_{25} \\ & + T_{21,12}A_{12}^* + \cdots + T_{22,1K}A_{25}^*. \end{aligned} \tag{7.4}$$

The coefficients in Table 7.6 with indices $(p,m) = (2,1)$ contribute to the fundamental scattered

p	q	m	n	$S_{pq,mn}$	$T_{pq,mn}$
2	1	1	1	2.834 $\angle 117°$	-
2	1	2	2	1.143 $\angle 82°$	0.328 $\angle 1°$
2	1	1	3	0.773 $\angle 57°$	0.777 $\angle 25°$
2	1	1	4	0.697 $\angle 94°$	0.627 $\angle 13°$
2	2	1	5	0.605 $\angle 102°$	0.657 $\angle 16°$
2	1	1	2	0.757 $\angle 119°$	0.472 $\angle -30°$
2	2	1	4	0.638 $\angle 100°$	0.616 $\angle 14°$
1	1	3	3	0.872 $\angle -137°$	0.077 $\angle 160°$
1	1	1	1	0.850 $\angle -93°$	-
1	1	5	5	0.845 $\angle -150°$	0.037 $\angle -150°$
1	1	4	4	0.834 $\angle -148°$	0.020 $\angle 25°$
2	2	1	3	0.557 $\angle 97°$	0.621 $\angle 12°$
2	2	3	3	0.711 $\angle -136°$	0.401 $\angle 88°$
2	1	3	3	0.695 $\angle 94°$	0.425 $\angle 96°$
1	1	2	2	0.806 $\angle -127°$	0.106 $\angle 60°$
2	2	5	5	0.793 $\angle -118°$	0.180 $\angle 171°$
2	2	3	4	0.677 $\angle 177°$	0.401 $\angle 92°$
2	1	1	5	0.496 $\angle 109°$	0.593 $\angle 17°$
2	2	1	2	0.530 $\angle 148°$	0.391 $\angle 8°$
2	1	4	4	0.650 $\angle 56°$	0.105 $\angle -10°$
2	1	3	4	0.489 $\angle 162°$	0.379 $\angle 90°$

Table 7.6: The 20 largest XP coefficients extracted from the AFRL GaN HEMT biased as a class AB amplifier with a power available from the source of 11.6 dBm at 10 GHz. The coefficients are sorted by their magnitude.

wave phasor at the output port, B_{21}. There is clearly inter-modulation between the input and output harmonics of this device. Eight of the top 20 largest XP coefficients correspond to the output term B_{21}. For a linear device, the only non-zero XP coefficient corresponding to B_{21} would be $S_{21,11}$.

7.7 PA Design with the XP Model in ADS

An XP model was extracted from FKT device simulations of the AFRL GaN HEMT in the previous section. In this section, the HEMT XP model is loaded into ADS for simple PA design. Figure 7.23 presents the XP simulation layout in ADS. An input file containing the XP information and coefficients is loaded into the component labeled "X2P". Because there are some small notational differences between the XP model in ADS and the XP model presented in this work, the

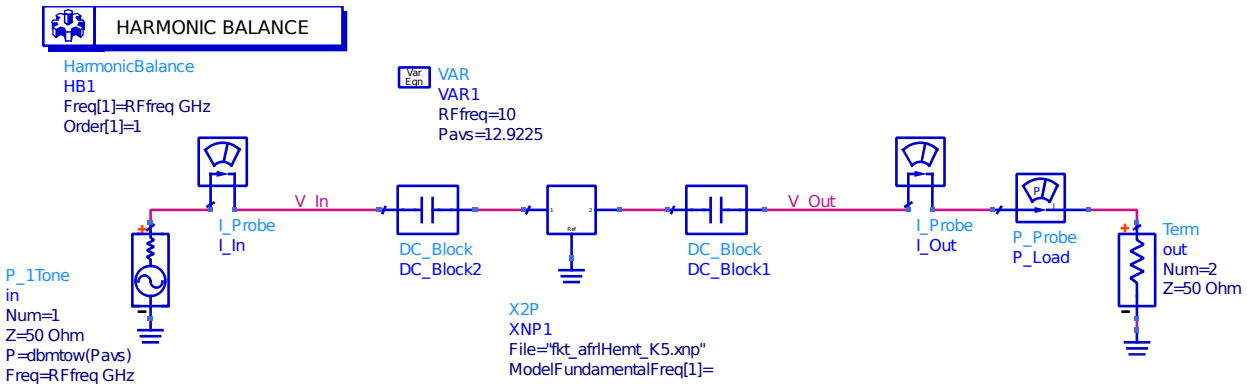


Figure 7.23: Layout of the XP model simulation in ADS. The components "P_1Tone" and "Term" represent the input and output ports of the circuit. Component "X2P" is the XP model which specifies and input file containing the model information and coefficients. Current and power probes are placed in the circuit to calculate the large-signal FOMs.

next section will focus on several practical considerations for the following ADS simulations.

7.7.1 Practical Considerations

Several practical considerations are presented in order to seamlessly apply the extracted XP model of this chapter to PA design in ADS. The XP model in Keysight's ADS is defined as

$$b_{ik} = X^B_{ik}\left(|a_{11}|\right)P^k + \sum_{(j,l)\neq(1,1)}\left(X^S_{ik,jl}\left(|a_{11}|\right)P^{k-l}a_{jl} + X^T_{ik,jl}\left(|a_{11}|\right)P^{k+l}a^*_{jl}\right), \tag{7.5}$$

where i, j are port indices, k, l are harmonic indices, a_{jl} are incident wave phasors, b_{ik} are scattered wave phasors, P is the phase of the large-signal phasor a_{11}, X^B_{ik} is the B-type XP coefficient, $X^S_{ik,jl}$ is the S-type XP coefficient, and $X^T_{ik,jl}$ is the T-type XP coefficient [26].

The first difference between the ADS XP model and the XP model in this work is the definition of the incident and scattered wave phasors. In this work, the phasors are defined in Eqns. (2.51) and (2.52). The ADS definitions of the incident and scattered waves are

$$a_{jl} = \frac{V_{jl} + Z_j I_{jl}}{\sqrt{8\mathrm{Re}\left\{Z_j\right\}}}, \tag{7.6}$$

$$b_{ik} = \frac{V_{ik} - Z^*_i I_{ik}}{\sqrt{8\mathrm{Re}\left\{Z_i\right\}}}. \tag{7.7}$$

Here, Z_i is the complex impedance of port i. The normalization of the ADS phasors follows the "power definition" [26].

Notationally, the ADS XP model, Eqn. (7.5), differs from the XP model presented in this work, Eqn. (2.50). First, the port and harmonic indices of the ADS XP coefficients are arranged differently than the ones presented in this work. However, this poses no issue when loading the coefficients into ADS. A more important distinction is the B-type XP coefficient, $X^B_{ik}(|a_{11}|)$. Using the ADS port and harmonic indices i, j, k, l, these coefficients are related to the $S_{p1,q1}$ terms by

$$X^B_{ik}(|a_{11}|)\,P^k = S_{i1,k1}A_{11}P^k. \tag{7.8}$$

The X^B coefficients define the product of the large-signal phasor, A_{11}, and the $S_{pm,qn}$ coefficients. This simple yet important distinction could yield significant differences in the gain computed from the XP model simulations.

7.7.2 Class AB Amplifier

An XP model is extracted from AFRL GaN HEMT simulations over a range of input power. ADS offers a wide range of independent variables of the XP model including fundamental frequency and quiescent bias. However, the operating point, $V_{GS} = -2.8\,\text{V}$ and $V_{DS} = 10\,\text{V}$, and the fundamental frequency of 10 GHz are constant in this XP model. The load impedance of the XP extraction simulations is $Z_L = 50\,\Omega$.

Validation of the XP model imported into ADS is first presented. The simulated FOMs of the ADS XP model are compared with the FOMs calculated directly from the FKT device simulations. This is a simple validation test as the XP model is extracted from the same FKT device simulations. Figure 7.24 presents the input power sweep results calculated in ADS and calculated directly from the FKT device simulations. The reported FOMs are output power, P_{out}, (black), power gain, G_P, (red), and PAE (blue). The FKT device simulation FOMs are plotted with solid lines and the ADS FOMs are plotted with squares. This result validates that the XP model is properly imported into ADS and that there is no mismatch in power definitions described in Section 7.7.1.

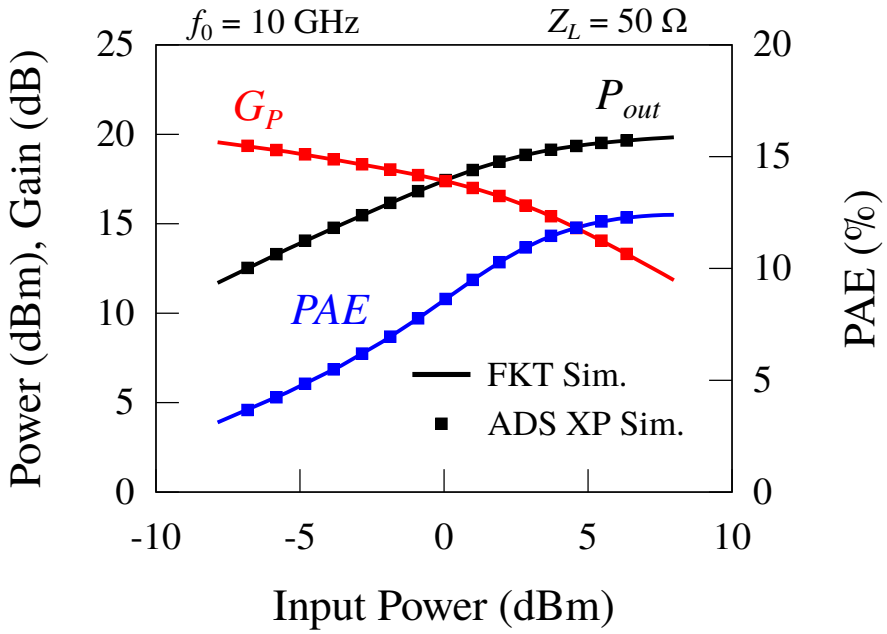
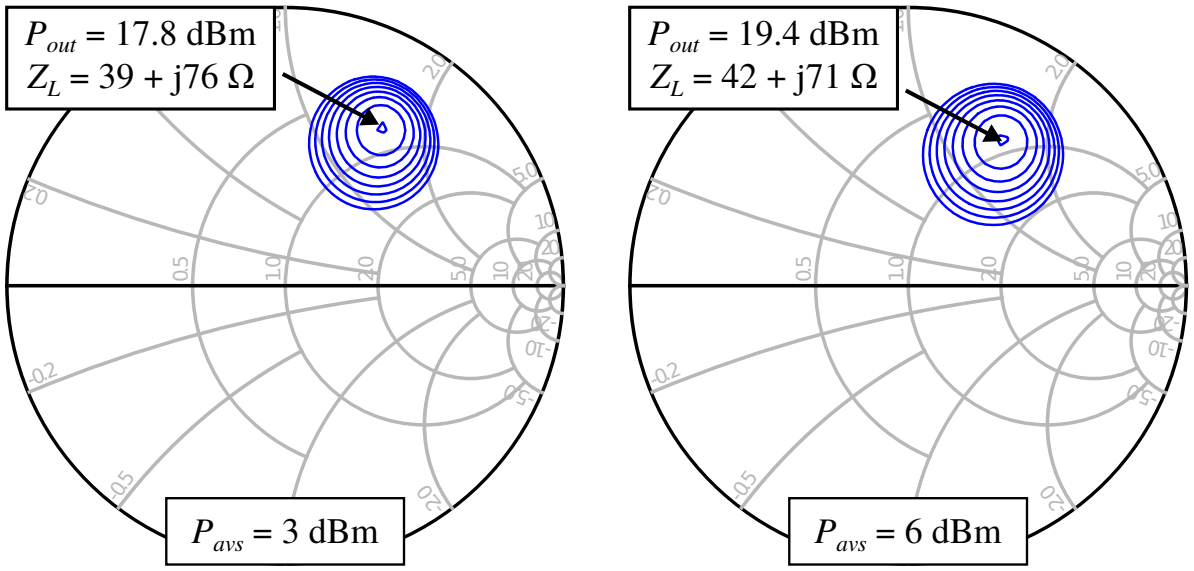


Figure 7.24: Validation of the input power dependent ADS XP model. This model is extracted from FKT simulations of the AFRL GaN HEMT and imported into ADS. Device parameters of the XP model include a gate-source bias of -2.8 V, a drain-source bias of 10 V, a fundamental frequency of 10 GHz, and a load impedance of 50 Ω.

To demonstrate the versatility of the XP model, several LP data sets are calculated in ADS with different input powers. These are proof-of-concept results generated from the XP model extracted at $\Gamma_L = 0$. Figures 7.25a – 7.26b present LP data generated from ADS simulations of the AFRL GaN HEMT XP model with $P_{avs} = 3$, 6, 9, and 12 dBm, respectively. These LP simulations determine the optimal load for maximum output power. At $P_{avs} = 3$ dBm, the optimal load is $Z_L = 39 + j76\,\Omega$ at 10 GHz. The optimal load for $P_{avs} = 6$ dBm is $Z_L = 42 + j71\,\Omega$ at 10 GHz. Finally, the optimal loads for $P_{avs} = 9$ dBm and $P_{avs} = 12$ dBm are $Z_L = 45 + j56\,\Omega$ and $Z_L = 42 + j46\,\Omega$, respectively, both at 10 GHz. The optimal load impedances corresponding to the powers available from the source of 3, 6, 9, and 12 dBm yield output powers of $P_{out} = 17.8$, 19.4, 20.0, and 20.4 dBm, respectively. A summary of the LP data versus input power is illustrated in Figure 7.27. Each reflection coefficient plotted on the Smith chart corresponds to the optimal load impedance determined from the ADS LP simulations. The optimal reflection coefficient shifts in a clockwise motion across the Smith chart as the input power increases.

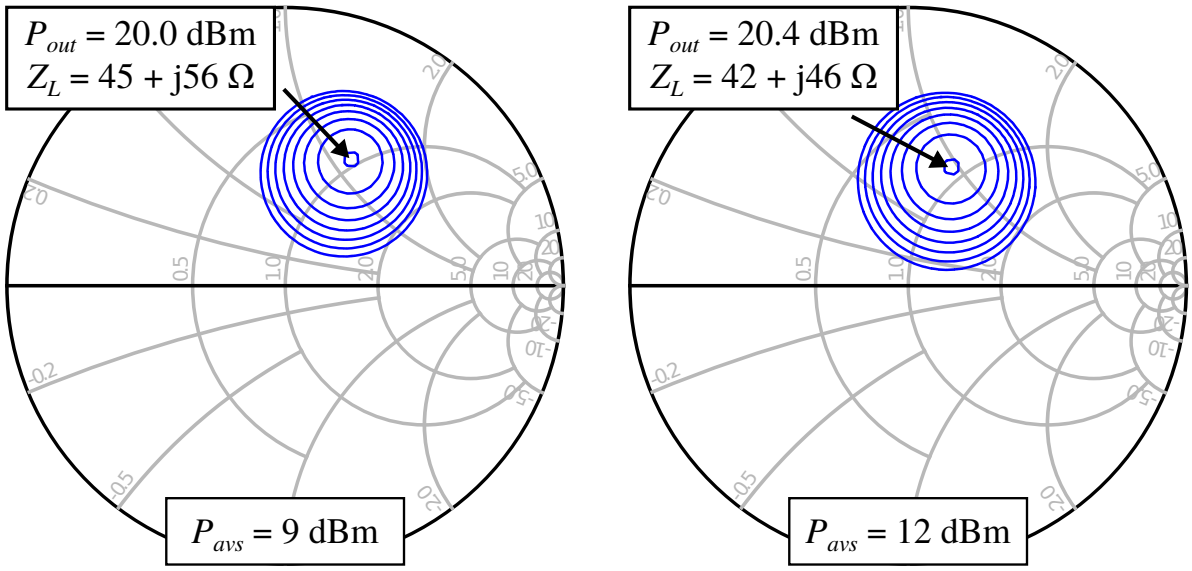


(a) (b)

Figure 7.25: Simulated LP data generated from ADS simulations of the AFRL HEMT XP model with a power available from the source of (a) 3 dBm and (b) 6 dBm. The power dependent XP model was extracted with a gate-source bias of -2.8 V, a drain-source bias of 10 V, a fundamental frequency of 10 GHz, and a load impedance of 50 Ω. The optimal loads are $39 + j76\,\Omega$ at 3 dBm and $42 + j71\,\Omega$ at 6 dBm, both at the fundamental frequency.

The transient input and output CW signals resulting from an ADS simulation of the XP model with the optimal load impedance at an input power of $P_{in} = 5.4$ dBm are presented in Figures 7.28 and 7.29. The input and output CW voltages are presented in Figures 7.28a and 7.28b and the input and output CW currents are illustrated in Figures 7.29a and 7.29b, respectively. The DC offset is not included in the CW signals. Clearly, the XP model reaches compression at the input power of $P_{in} = 5.4$ dBm as the output waveforms exhibit characteristics of a class B amplifier.

Finally, a comparison of the simulated LP data calculated directly from the FKT device simulations and the LP data computed with the XP model in ADS is reported in Figure 7.30. The magnitude of the large-signal incident phasor is held constant at $|A_{11}| = 0.75$ V for the LP simulations. Note that this definition of the incident wave follows Eqn. (2.51) and is different than Eqns. (7.6) and (7.7). Contours of the output power FOM computed from the FKT device simulations are plotted with red lines and output power contours computed from the ADS XP simulations are plotted with blue lines. Simulated LP data generated from the ADS XP simulations do not



(a) (b)

Figure 7.26: Simulated LP data generated from ADS simulations of the AFRL HEMT XP model with a power available from the source of (a) 9 dBm and (b) 12 dBm. The power dependent XP model was extracted with a gate-source bias of -2.8 V, a drain-source bias of 10 V, a fundamental frequency of 10 GHz, and a load impedance of 50 Ω. The optimal loads are $45 + j56\,\Omega$ at 9 dBm and $42 + j46\,\Omega$ at 12 dBm, both at the fundamental frequency.

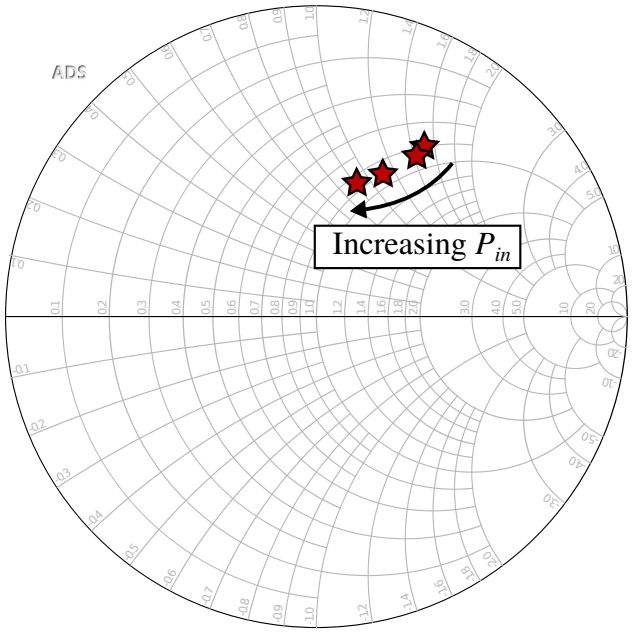


Figure 7.27: A summary of the optimal load impedances for maximum output power determined from the LP simulations of the AFRL GaN HEMT XP model in ADS.

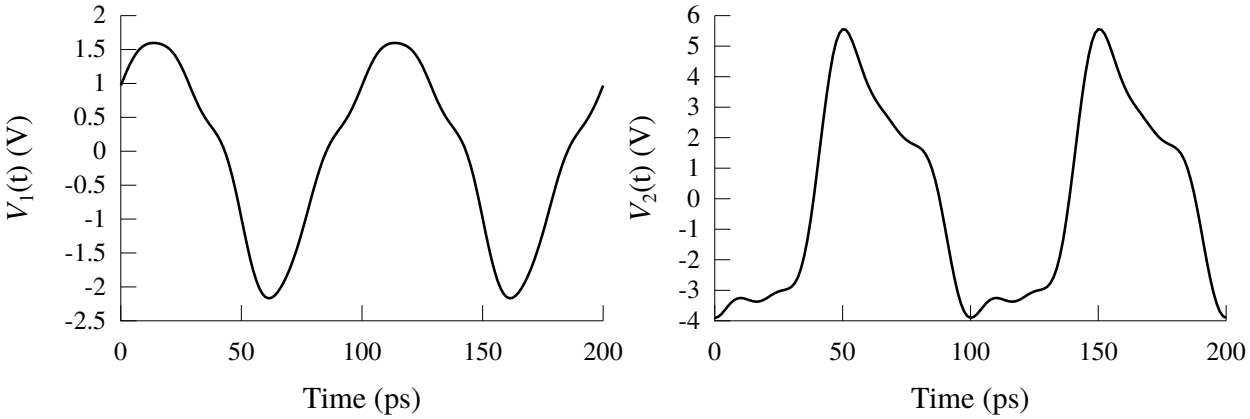


Figure 7.28: The (a) input and (b) output CW voltage signals calculated with the AFRL GaN HEMT XP model in ADS. Only the AC components of the waveforms are illustrated. The load impedance presented to the XP model is $42 + j46\,\Omega$ at 10 GHz. This is the optimal load for maximum output power determined from the LP simulations with a power available from the source of 12 dBm.

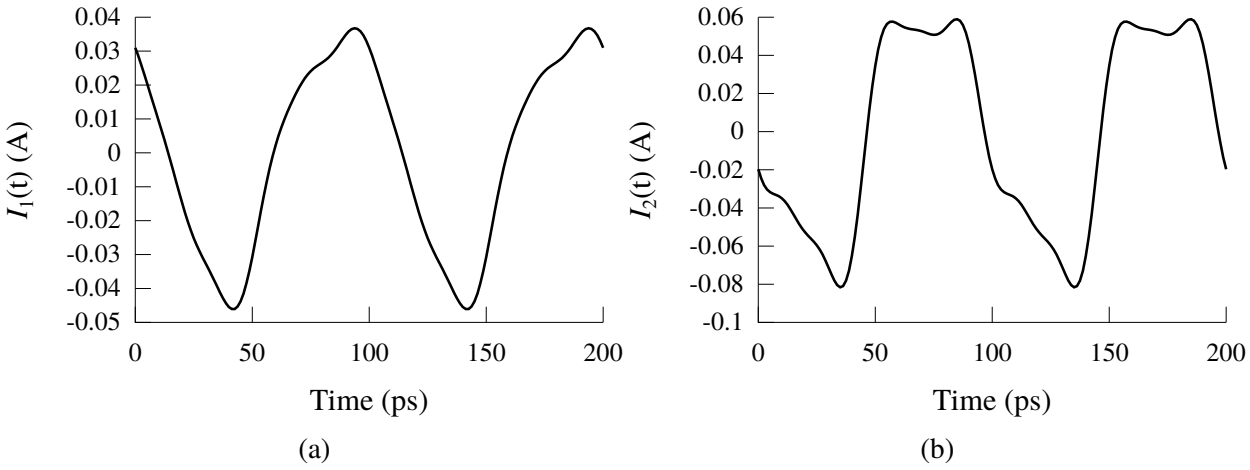


Figure 7.29: The (a) input and (b) output CW current signals calculated with the AFRL GaN HEMT XP model in ADS. Only the AC components of the waveforms are illustrated. The load impedance presented to the XP model is $42 + j46\,\Omega$ at 10 GHz — the optimal load for maximum output power determined from the LP simulations with a power available from the source of 12 dBm.

exactly replicate the LP data generated from the FKT device simulations. Clearly, the presence of a non-zero reflection coefficient alters the large-signal operating point of the HEMT. The XP model needs to be sampled over a range of load impedances to generate accurate LP data. A load-dependent XP model is available in ADS and the software interpolates the XP coefficients across

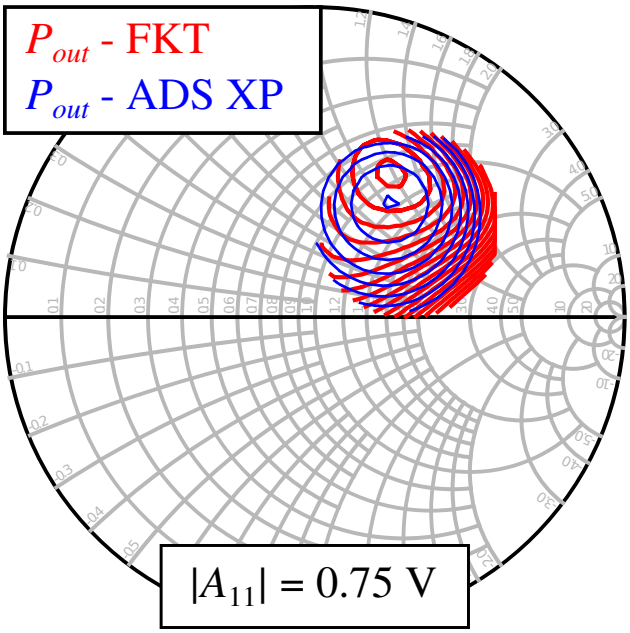


Figure 7.30: A comparison of the simulated LP data generated directly from the FKT device simulations (red lines) and from the ADS simulations of the XP model (blue lines). The reported FOM is the output power of the device. The magnitude of the large-signal incident phasor is held constant at 0.75 V for the LP simulations. Note that this definition of the incident wave is different than Eqns. (7.6) and (7.7).

the load impedance independent variable [112].

7.8 Discussion and Conclusions

This chapter presented a range of results that demonstrated the ability of the FKT device simulator to generate LP data for RF circuit design. The GaN HEMT used in the FKT simulations was presented in Section 7.2. A detailed description of the device layout was illustrated in Figures 7.1 – 7.4. Section 7.3 described the TLM characterization process and applied it to estimate unknown physical parameters in the GaN HEMT. The HEMT was further characterized in Section 7.4 by analyzing its DC and small-signal response. Comparison of the simulated and the measured I-V family was reported in Figure 7.10. The drain current, I_D, and the transconductance, g_m, of the HEMT versus the gate-source bias, V_{GS}, was presented in Figure 7.11. Simulations agreed well with the measurements. S-Parameters and the small-signal current gain were presented in Figures 7.12 and 7.13, respectively. The simulated DC and small-signal results provide confidence

that the FKT device simulator is adequately capturing the overall response of the AFRL GaN HEMT.

Large-signal simulations of the AFRL GaN HEMT with the FKT device simulator were presented in Section 7.5. A simulated input power sweep of the HEMT was compared to measurements in Figure 7.14. The impact of the operating point and other device features on the power response of the device were also investigated in Section 7.5. The simulated low power response of the device agreed well with measurements. However, the simulated high power response of the device did not agree with measurements. First, the simulated *PAE* was significantly lower than what was measured. This was expected as the simulated quiescent drain current was larger than measurements. Second, the simulated device compresses at a lower input power level compared to the measured device. These issues are the topic of future research which will aim to accurately reproduce the high power response of the AFRL GaN HEMT.

The remaining sections of this chapter focused on extraction and application of XP models. Section 7.6 provided a detailed description of the XP extraction process of the AFRL GaN HEMT simulations. Included in this section were data tables containing the incident and scattered wave phasors of the FKT device simulations. The XP extraction process injects incident wave signals at different harmonics of the device ports to calculate the coefficients of the model. The data tables provided quantitative examples of the signal injection process. Table 7.6 reported the 20 largest coefficients of the high power AFRL GaN HEMT XP model extracted at 10 GHz.

Section 7.7 presented proof-of-concept results demonstrating the use of the extracted XP model in ADS. The XP model was extracted over a range of input power levels at 10 GHz. The ADS XP model was used to generate the LP data reported in Figures 7.25a – 7.26b. ADS simulations of the XP model provided computationally efficient means to generate useful data for RF circuit design. Frequency and operating point dependencies could be incorporated into the XP model. This would be accomplished by extracting XP coefficients from FKT device simulations over a range of fundamental frequencies and DC biases. ADS interpolates the XP coefficients as functions of the input power, fundamental frequency, and DC bias independent variables. Loading conditions

must also be included in the XP model. Extraction of an XP model with operating condition dependencies from FKT device simulations is now possible with the work presented in this chapter.

The FKT device simulator is an excellent candidate for high frequency and high power simulations of GaN HEMT technology. Simulations of the HEMT technology can provide supplemental data for RF circuit design. Furthermore, the device simulator can generate measurement intensive data including LP at high frequency and frequency and/or power dependent LP. The simulation framework is amenable to XP model extraction which enables efficient power sweep and LP simulations in commercial CAD software like ADS.

CHAPTER 8

CONCLUSION AND FUTURE RESEARCH

This thesis presented high frequency and high power GaN HEMT simulations using the FKT device simulation TCAD tool. A thorough derivation and discretization of governing equations were presented along with a detailed discussion on solving the nonlinear device equations with the corresponding BCs. The FKT device simulator was used to generate useful data for RF circuit designers including S-Parameters, high frequency LP, and high frequency input power sweeps. Many of these simulated results would be difficult to measure at a fundamental frequency above 35 GHz. Finally, an XP model was extracted from FKT device simulations of a state-of-the-art GaN HEMT fabricated at AFRL. The XP model was imported into ADS for computationally efficient LP simulations.

8.1 Future Research

The main contribution of this work is utilizing the FKT device simulation TCAD tool developed by Grupen [84] to generate useful data for RF circuit designers. A list of research topics that are now possible based on the contributions of this thesis are

- Accurate simulation of the AFRL GaN HEMT at high input power
- Full-wave LP simulations
- Frequency, DC bias, and load dependent XP extraction
- Full-wave XP extraction
- Harmonic mixing simulations of state-of-the-art devices

APPENDIX

A.1 Moments of the BTE

In Section 4.1.1, moments of the BTE were separated into even and odd components. These components are

$$\frac{1}{4\pi}\int_{-\infty}^{\infty} d\vec{\mathbf{k}} O(\vec{\mathbf{k}}) \left[\frac{\partial f_0}{\partial t} + \vec{\mathbf{v}} \cdot \nabla_r f_1 - q\frac{\vec{\mathbf{E}}}{\hbar} \cdot \nabla_k f_1 - \left.\frac{\partial f_0}{\partial t}\right|_{coll} \right] = 0, \tag{1}$$

$$\frac{1}{4\pi}\int_{-\infty}^{\infty} d\vec{\mathbf{k}} O(\vec{\mathbf{k}}) \left[\frac{f_1}{\tau_k} + \frac{\partial f_1}{\partial t} + \vec{\mathbf{v}} \cdot \nabla_r f_0 - q\frac{\vec{\mathbf{E}}}{\hbar} \cdot \nabla_k f_0 \right] = 0. \tag{2}$$

Eqns. (1) and (2), in conjunction with the moment operators $\{1, \vec{\mathbf{v}}, E(\vec{\mathbf{k}}), E(\vec{\mathbf{k}})\vec{\mathbf{v}}\}$, are now used to derive the governing equations of FKT. It is worth stating that the left hand side of Eqns. (1) and (2) are exactly zero for integration over reciprocal-space with odd and even moment operators, respectively. For all of the following derivations, the $\vec{\mathbf{k}}$-space integrals are over the entire domain and therefore the limits of integration will be dropped.

Two vector calculus identities will be used in the following derivations. They are the scalar-scalar product gradient chain rule

$$\nabla(\psi\phi) = \phi\nabla\psi + \psi\nabla\phi, \tag{3}$$

and the vector-scalar product chain rule

$$\nabla \cdot \left(\psi\vec{\mathbf{A}}\right) = \vec{\mathbf{A}} \cdot \nabla\psi + \psi\nabla \cdot \vec{\mathbf{A}}. \tag{4}$$

A.1.1 Zeroth moment — Electron Continuity

The electron continuity equation is derived from Eqn. (1) with the unity moment operator. Using Eqn. (4.3), integration of the first term yields

$$\frac{1}{4\pi^3}\int d\vec{\mathbf{k}} \frac{\partial f_0}{\partial t} = \frac{\partial}{\partial t}\left(\frac{1}{4\pi^3}\int d\vec{\mathbf{k}} f_0\right) \equiv \frac{\partial n}{\partial t}. \tag{5}$$

Integration of the second term, using Eqns. (4) and (4.5), yields the divergence term of the electron continuity equation

$$\frac{1}{4\pi^3}\int d\vec{\mathbf{k}}\vec{\mathbf{v}}\cdot\nabla_r f_1 = \nabla_r\cdot\left(\frac{1}{4\pi^3}\int d\vec{\mathbf{k}}\vec{\mathbf{v}}f_1\right)\equiv\nabla\cdot\vec{\mathbf{J}}_n. \tag{6}$$

The term with the reciprocal-space gradient is more involved. Using, Eqn. (4), as well as the divergence theorem, this term integrates to

$$-\frac{q}{4\pi^3\hbar}\int d\vec{\mathbf{k}}\vec{\mathbf{E}}\cdot\nabla_k f_1 = -\frac{q}{4\pi^3\hbar}\int d\vec{\mathbf{A}}_k\hat{n}\cdot\vec{\mathbf{E}}f_1 = 0. \tag{7}$$

This term vanishes as the non-equilibrium distribution function goes to zero much faster than the surface area goes to infinity. Physically, this term must vanish as there are never an infinite number of particles in any system. Finally, the collision term is written as a collision operator

$$\frac{1}{4\pi^3}\int d\vec{\mathbf{k}}\left.\frac{\partial f_0}{\partial t}\right|_{coll}\equiv C_n, \tag{8}$$

and is treated using Fermi's golden rule. The zeroth moment of the BTE is

$$\frac{\partial n}{\partial t}+\nabla\cdot\vec{\mathbf{J}}_n+C^n=0. \tag{9}$$

A.1.2 Second Moment — Energy Conservation Equation

The electron energy conservation equation is derived from Eqn. (1) with the energy moment operator $E(\vec{\mathbf{k}})$. First, Eqn. (4.4) is used for integration of the first term. This yields

$$\frac{1}{4\pi^3}\int d\vec{\mathbf{k}}E(\vec{\mathbf{k}})\frac{\partial f_0}{\partial t} = \frac{\partial}{\partial t}\left(\frac{1}{4\pi^3}\int d\vec{\mathbf{k}}E(\vec{\mathbf{k}})f_0\right)\equiv\frac{\partial E_n}{\partial t}. \tag{10}$$

Next, using Eqns. (4) and (4.6), the divergence term of the energy conservation equation results from the integration

$$\frac{1}{4\pi^3}\int d\vec{\mathbf{k}}E(\vec{\mathbf{k}})\vec{\mathbf{v}}\cdot\nabla_r f_1 = \nabla_r\cdot\left(\frac{1}{4\pi^3}\int d\vec{\mathbf{k}}E(\vec{\mathbf{k}})\vec{\mathbf{v}}f_1\right)\equiv\nabla\cdot\vec{\mathbf{S}}_n. \tag{11}$$

Unlike the zeroth moment of the BTE, the second will yield a non-zero reciprocal-space gradient term. Using Eqns. (3) and (4) and the divergence theorem, the third term integrates to

$$-\frac{q}{4\pi^3\hbar}\int d\vec{\mathbf{k}}\vec{\mathbf{E}}\cdot E(\vec{\mathbf{k}})\nabla_k f_1 = -\frac{q}{4\pi^3\hbar}\left(\int d\vec{\mathbf{A}}_k\hat{n}\cdot\vec{\mathbf{E}}E(\vec{\mathbf{k}})f_1-\int d\vec{\mathbf{k}}\vec{\mathbf{E}}\cdot\left(\nabla_k E(\vec{\mathbf{k}})\right)f_1\right). \tag{12}$$

The first term of Eqn. (12) goes to zero for the same reasons as Eqn. (7). Then, with the relation $\hbar\vec{\mathbf{v}} = \nabla_k E(\vec{\mathbf{k}})$ and Eqn. (4.5), the fourth term integrates to

$$-\frac{q}{4\pi^3\hbar}\int d\vec{\mathbf{k}}\vec{\mathbf{E}}\cdot E(\vec{\mathbf{k}})\nabla_k f_1 = q\vec{\mathbf{E}}\cdot\vec{\mathbf{J}}_n. \tag{13}$$

This is the Joule's heat component of the energy conservation equation. Finally, the collision term is written as a collision operator

$$\frac{1}{4\pi^3}\int d\vec{\mathbf{k}}E(\vec{\mathbf{k}})\left.\frac{\partial f_0}{\partial t}\right|_{coll} \equiv C^E. \tag{14}$$

The energy conservation equation is

$$\frac{\partial E_n}{\partial t} + \nabla\cdot\vec{\mathbf{S}}_n + q\vec{\mathbf{E}}\cdot\vec{\mathbf{J}}_n + C^E = 0. \tag{15}$$

A.1.3 First and Third Moments — Particle and Kinetic Energy Flux Densities

The first and third moments do not require much attention before discretization. Using Eqns. (4.5) and (4.6) and the approximation that density of states, $d\vec{\mathbf{k}}$, is time-independent, the particle and kinetic energy flux densities are

$$\vec{\mathbf{J}}_n + \tau_k\frac{\partial\vec{\mathbf{J}}_n}{\partial t} = \frac{1}{4\pi^3}\int d\vec{\mathbf{k}}\vec{\mathbf{v}}\tau_k\left(\frac{q}{\hbar}\vec{\mathbf{E}}\cdot\nabla_k f_0 - \vec{\mathbf{v}}\cdot\nabla_r f_0\right), \tag{16}$$

$$\vec{\mathbf{K}}_n + \tau_k\frac{\partial\vec{\mathbf{K}}_n}{\partial t} = \frac{1}{4\pi^3}\int d\vec{\mathbf{k}}E(\vec{\mathbf{k}})\vec{\mathbf{v}}\tau_k\left(\frac{q}{\hbar}\vec{\mathbf{E}}\cdot\nabla_k f_0 - \vec{\mathbf{v}}\cdot\nabla_r f_0\right). \tag{17}$$

The terms $\tau_k\frac{\partial\vec{\mathbf{J}}_n}{\partial t}$ and $\tau_k\frac{\partial\vec{\mathbf{K}}_n}{\partial t}$ are approximately zero when $\tau_k \ll \partial t$. This proves to be an excellent approximation for up to sub-mm-wavelength devices.

A.2 Reciprocal-Space to Energy-Space Conversion

After taking moments of the BTE, the densities and fluxes require integration over the first Brillouin zone of reciprocal-space. Rather than engage in such a tedious calculation, FKT aims to make use of a couple well known delta function properties. Chapter 8 of [113] provides a detailed discussion of a similar derivation.

A.2.1 Two Delta Function Properties

There are two properties of the delta function which are needed to convert the reciprocal-space integration to energy-space integration. The first is the well-known sifting property of the delta function. This is

$$\int_{-\infty}^{\infty} n(t)\delta(t-T)dt = n(T) \tag{18}$$

The second identity is the composition of the delta function with another function

$$\int_{-\infty}^{\infty} \delta\left(f(\vec{\mathbf{r}})\right) g(\vec{\mathbf{r}}) d\vec{\mathbf{r}} = \sum_i \int_{-\infty}^{\infty} \frac{g(\vec{\mathbf{r}})}{|\nabla f(\vec{\mathbf{r}})|}\bigg|_{\vec{\mathbf{r}}=\vec{\mathbf{r}}_i} d\vec{\mathbf{S}}_i, \tag{19}$$

where $\vec{\mathbf{r}}_i$ and $d\vec{\mathbf{S}}_i$ represent the surface and surface differential element corresponding to a root of the scalar function $f(\vec{\mathbf{r}})$. This identity can be derived with the identity $\delta(ax) = \frac{1}{|a|}\delta(x)$ and a Taylor expansion at the function's roots.

A.2.2 Electron Density Conversion

To illustrate the conversion from reciprocal-space integration to energy-space integration, the electron density is derived. The flux equations follow the same derivation. Using the sifting property, Eqn. (18), the electron density becomes

$$n = \int_{-\infty}^{\infty} \left[\frac{1}{4\pi^3}\int d\vec{\mathbf{k}} f_0(E)\right] \delta(E - E(\vec{\mathbf{k}}))dE, \tag{20}$$

and Eqn. (19) is used to convert the delta function to

$$n = \int_{-\infty}^{\infty} \left[\frac{1}{4\pi^3}\sum_i \int d\vec{\mathbf{S}}_i \frac{1}{|\nabla_k E(\vec{\mathbf{k}})|}\bigg|_{\vec{\mathbf{k}}=\vec{\mathbf{k}}_i}\right] f_0(E)dE. \tag{21}$$

The surfaces with differential elements $d\vec{\mathbf{S}}_i$ are defined by the roots E_i of $E - E(\vec{\mathbf{k}}) = 0$, i.e., surfaces of constant electron eigen-energy or "isosurfaces."

A.3 SG Discretization

In Section 4.3.2, the discrete FKT particle flux was presented. This appendix includes the complete derivation of the SG discretization of both the phenomenological DD model and the FKT particle flux.

A.3.1 Discretization of the Classic DD Model

The vector projection of the phenomenological DD model for semiconductor device simulation defined on a primary edge of a DV mesh is

$$J_n(r-r_0) = a_1 n(r-r_0) + a_2 \frac{dn(r-r_0)}{dr}, \tag{22}$$

where r_0 and r_1 are the nodes of the primary edge and a_1 and a_2 are constant over the primary edge. Given the approximation that the particle flux and electric field are spatially constant on the primary edge, i.e., $J_n(r-r_0) \approx J_n$ and $\mathrm{E}(r-r_0) \approx \mathrm{E}$, the DD model becomes a first-order linear differential equation

$$J_0 = a_1 n(r-r_0) + \frac{dn(r-r_0)}{dr}. \tag{23}$$

Its solution, with the integrating factor $p(r-r_0) = \exp\left[\frac{a_1}{a_2}(r-r_0)\right]$, is

$$n(r-r_0) = \frac{J_0}{a_1} + C\exp\left[-\frac{a_1}{a_2}(r-r_0)\right]. \tag{24}$$

Here, C is an integration constant. Determining the constants J_0 and C require the two BCs $n(r-r_0)|_{r_0} = n_0$ and $n(r-r_0)|_{r_1} = n_1$. Applying these BCs yields the two equations

$$n_0 = \frac{J_0}{a_1} + C, \tag{25}$$

and

$$n_1 = \frac{J_0}{a_1} + C\exp\left[-\frac{a_1}{a_2}(r_1-r_0)\right]. \tag{26}$$

Solving Eqn. (25) for C, substituting into Eqn. (26), and re-arranging for J_0 yields

$$J_0 = a_1 \frac{n_1 - n_0\exp\left[-\frac{a_1}{a_2}(r_1-r_0)\right]}{1-\exp\left[-\frac{a_1}{a_2}(r_1-r_0)\right]}. \tag{27}$$

With the definition of the Bernoulli function $B(x) = \frac{x}{\exp(x)-1}$, the discrete particle flux is

$$J_n = J_0 = -\frac{a_2}{L}\left[B\left(\frac{a_1}{a_2}L\right)n_0 - B\left(-\frac{a_1}{a_2}L\right)n_1\right]. \tag{28}$$

Here, the length of the primary edge is $L = r_1 - r_0$. The Bernoulli function has the asymptotic forms

$$\lim_{x\to\infty} B(x) \to 0, \tag{29}$$

$$\lim_{x\to 0} B(x) \to 1. \tag{30}$$

Also, as the Bernoulli function argument approaches infinity, the Bernoulli function goes to negative infinity as $-x$.

A.3.2 Discretization of the FKT Particle Flux

With a spatially constant particle flux and electric field, the projection of the k^{th} piece-wise FKT particle flux on the j^{th} primary edge is

$$J_{n,jk} = -qD_{jk}\left(\mathrm{E}_j\left(k_B T(r)\right)^{\beta_{jk}} F'_{\beta_{jk}}(r) + \frac{d}{dr}\left[\frac{\left(k_B T(r)\right)^{\beta_{jk}+1}}{q} F_{\beta_{jk}}(r)\right]\right). \tag{31}$$

The spatial coordinate r is defined on a primary edge from r_0 to r_1. Solving Eqn. (31) requires choosing the dependent variable of the linear first-order differential equation. The following subsections explore two possibilities.

A.3.2.1 Generalized Einstein Relation Form

If the dependent variable of Eqn. (31) is

$$N_{E,jk}(r) = \left(k_B T(r)\right)^{\beta_{jk}} F'_{\beta_{jk}}(r), \tag{32}$$

then the FKT particle flux becomes

$$J_{n,jk} = -qD_{jk}\left(\mathrm{E}_j N_{E,jk}(r) + \frac{d}{dr}\left[\frac{k_B T(r)}{q}\frac{F_{\beta_{jk}}(r)}{F'_{\beta_{jk}}(r)}N_{E,jk}(r)\right]\right)$$
$$= -qD_{jk}\left(\left[\mathrm{E}_j + \frac{d}{dr}\left(\frac{k_B T(r)}{q}\frac{F_{\beta_{jk}}(r)}{F'_{\beta_{jk}}(r)}\right)\right]N_{E,jk}(r) + \left[\frac{k_B T(r)}{q}\frac{F_{\beta_{jk}}(r)}{F'_{\beta_{jk}}(r)}\right]\frac{dN_{E,jk}(r)}{dr}\right). \tag{33}$$

To make Eqn. (33) a linear first-order differential equation, the following approximations are made. With the definition of the generalized Einstein Relation

$$\mathrm{Ein}_{n,jk}(r) = \frac{k_B T(r)}{q}\frac{F_{\beta_{jk}}(r)}{F'_{\beta_{jk}}(r)}, \tag{34}$$

these approximations are

$$\frac{d}{dr}\left(\frac{k_B T(r)}{q}\frac{F_{\beta_{jk}}(r)}{F'_{\beta_{jk}}(r)}\right) \approx \frac{1}{L}\left(\frac{k_B T(r)}{q}\frac{F_{\beta_{jk}}(r)}{F'_{\beta_{jk}}(r)}\bigg|_{r_1} - \frac{k_B T(r)}{q}\frac{F_{\beta_{jk}}(r)}{F'_{\beta_{jk}}(r)}\bigg|_{r_0}\right) = \Delta\mathrm{Ein}_{n,jk}, \tag{35}$$

and

$$\frac{k_B T(r)}{q}\frac{F_{\beta_{jk}}(r)}{F'_{\beta_{jk}}(r)} \approx \frac{1}{2}\left(\frac{k_B T(r)}{q}\frac{F_{\beta_{jk}}(r)}{F'_{\beta_{jk}}(r)}\bigg|_{r_1} + \frac{k_B T(r)}{q}\frac{F_{\beta_{jk}}(r)}{F'_{\beta_{jk}}(r)}\bigg|_{r_0}\right) = \left(\mathrm{Ein}_{n,jk}\right)_{\mathrm{ave}}. \tag{36}$$

Eqn. (33) is solved in the same manner as Eqn. (23) with the substitutions

$$a_1 \to -qD_{jk}\left[\mathrm{E}_j + \Delta\mathrm{Ein}_{n,jk}\right], \tag{37}$$

and

$$a_2 \to -qD_{jk}\left(\mathrm{Ein}_{n,jk}\right)_{\mathrm{ave}}. \tag{38}$$

The discrete particle flux, with the BCs $N_{E,jk}(r)\big|_{r_0} = N_{E,0,jk}$ and $N_{E,jk}(r)\big|_{r_1} = N_{E,1,jk}$, is

$$J^E_{n,jk} = \frac{qD_{jk}}{L_j}\left(\mathrm{Ein}_{n,jk}\right)_{\mathrm{ave}}\left[B(\xi_{n,E,jk})N_{E,0,jk} - B(-\xi_{n,E,jk})N_{E,1,jk}\right], \tag{39}$$

with

$$\xi_{n,E,jk} = \frac{1}{\left(\mathrm{Ein}_{n,jk}\right)_{\mathrm{ave}}}\left[\mathrm{E}_j + \Delta\mathrm{Ein}_{n,jk}\right]L_j. \tag{40}$$

A.3.2.2 Inverse Generalized Einstein Relation Form

In the previous subsection, the discrete particle flux with the generalized Einstein relation was derived. If, however, the dependent variable of Eqn. (31) is

$$N_{I,jk}(r) = (k_B T(r))^{\beta_{jk}} F_{\beta_{jk}}(r), \tag{41}$$

then the FKT particle flux becomes

$$\begin{aligned} J_{n,jk} &= -qD_{jk}\left(\mathrm{E}_j \frac{F'_{\beta_{jk}}(r)}{F_{\beta_{jk}}(r)} N_{I,jk}(r) + \frac{d}{dr}\left[\frac{k_B T(r)}{q} N_{I,jk}(r)\right]\right) \\ &= -qD_{jk}\left(\left[\frac{F'_{\beta_{jk}}(r)}{F_{\beta_{jk}}(r)}\mathrm{E}_j + \frac{d}{dr}\left(\frac{k_B T(r)}{q}\right)\right] N_{I,jk}(r) + \left[\frac{k_B T(r)}{q}\right]\frac{dN_{I,jk}(r)}{dr}\right). \end{aligned} \tag{42}$$

Again, to make Eqn. (42) a linear first-order differential equation, the following approximations are made. With the definition of the inverse generalized Einstein Relation

$$\mathrm{Ein}^{-1}_{n,jk}(r) = \frac{F'_{\beta_{jk}}(r)}{F_{\beta_{jk}}(r)}, \tag{43}$$

these approximations are

$$\frac{d}{dr}\left(\frac{k_B T(r)}{q}\right) \approx \frac{1}{L}\left(\left.\frac{k_B T(r)}{q}\right|_{r_1} - \left.\frac{k_B T(r)}{q}\right|_{r_0}\right) = \Delta\left(\frac{k_B T}{q}\right), \tag{44}$$

$$\frac{k_B T(r)}{q} \approx \frac{1}{2}\left(\left.\frac{k_B T(r)}{q}\right|_{r_1} + \left.\frac{k_B T(r)}{q}\right|_{r_0}\right) = \left(\frac{k_B T}{q}\right)_{\mathrm{ave}}, \tag{45}$$

and

$$\frac{F_{\beta_{jk}}(r)}{F'_{\beta_{jk}}(r)} \approx \frac{1}{2}\left(\left.\frac{F_{\beta_{jk}}(r)}{F'_{\beta_{jk}}(r)}\right|_{r_1} + \left.\frac{F_{\beta_{jk}}(r)}{F'_{\beta_{jk}}(r)}\right|_{r_0}\right) = \left(\mathrm{Ein}^{-1}_{n,jk}\right)_{\mathrm{ave}}. \tag{46}$$

Again, Eqn. (33) is solved in the same manner as Eqn. (23), but with the substitutions

$$a_1 \to -qD_{jk}\left[\left(\mathrm{Ein}^{-1}_{n,jk}\right)_{\mathrm{ave}}\mathrm{E}_j + \Delta\left(\frac{k_B T}{q}\right)\right], \tag{47}$$

and

$$a_2 \to -qD_{jk}\left(\frac{k_B T}{q}\right)_{\mathrm{ave}}. \tag{48}$$

The discrete particle flux, with the BCs $N_{I,jk}(r)\big|_{r_0} = N_{I,0,jk}$ and $N_{I,jk}(r)\big|_{r_1} = N_{I,1,jk}$, is

$$J^I_{n,jk} = \frac{qD_{jk}}{L_j}\left(\frac{k_BT}{q}\right)_{\text{ave}}\left[B(\xi_{n,I,jk})N_{I,0,jk} - B(-\xi_{n,I,jk})N_{I,1,jk}\right], \tag{49}$$

with

$$\xi_{n,I,jk} = \frac{1}{\left(\frac{k_BT}{q}\right)_{\text{ave}}}\left[\left(\text{Ein}^{-1}_{n,jk}\right)_{\text{ave}}\text{E}_j + \Delta\left(\frac{k_BT}{q}\right)\right]L_j. \tag{50}$$

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] I. Bahl, *Fundamentals of RF and microwave transistor amplifiers*. Wiley, 2009.

[2] J. D. Cressler, "SiGe HBT technology: A new contender for Si-based RF and microwave circuit applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 46, no. 5, pp. 572–589, 1998.

[3] U. K. Mishra, P. Parikh, and Y.-f. Wu, "AlGaN / GaN HEMTs - An overview of device operation and applications," *Proceedings of the IEEE*, vol. 90, no. 6, pp. 1022–1031, 2002.

[4] O. Ambacher, "Role of spontaneous and piezoelectric polarization induced effects in group-III nitride based heterostructures and devices," *Phys. Stat. Sol. (B)*, vol. 381, pp. 381–390, 1999.

[5] U. K. Mishra, L. Shen, T. E. Kazior, and Y.-F. Wu, "GaN-based RF power devices and amplifiers," *Proceedings of the IEEE*, vol. 96, no. 2, pp. 287–305, 2008.

[6] R. C. Fitch, D. E. Walker, A. J. Green, S. E. Tetlak, J. K. Gillespie, R. D. Gilbert, K. A. Sutherlin, W. D. Gouty, J. P. Theimer, G. D. Via, K. D. Chabak, and G. H. Jessen, "Implementation of high-power-density X-band AlGaN/GaN high electron mobility transistors in a millimeter-wave monolithic microwave integrated circuit process," *IEEE Electron Device Letters*, vol. 36, no. 10, pp. 1004–1007, 2015.

[7] G. C. Barisich, A. C. Ulusoy, E. Gebara, and J. Papapolymerou, "A reactively matched 1.0-11.5 GHz hybrid packaged GaN high power amplifier," *IEEE Microwave and Wireless Components Letters*, vol. 25, no. 12, pp. 811–813, 2015.

[8] D. Sokal and A. Sokal, "Class E - A new class of high-efficiency tuned single-ended switching power amplifiers," *IEEE Journal of Solid-State Circuits*, vol. SC-10, no. 3, pp. 168–176, 1975.

[9] I. Song, A. C. Ulusoy, M. A. Oakley, I. Ju, M.-K. Cho, and J. D. Cressler, "Inverse class-F X-band SiGe HBT power amplifier with 44% PAE and 24.5 dBm peak output power," *Microwave and Optical Technology Letters*, vol. 58, no. 12, pp. 2868–2871, 2016.

[10] J. B. Beyer, S. N. Prasad, R. C. Becker, J. E. Nordman, and G. K. Hohenwarter, "MESFET distributed amplifier design guidelines," *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-32, no. 3, pp. 268–275, 1984.

[11] J. L. Dawson and T. H. Lee, "Automatic phase alignment for a fully integrated cartesian feedback power amplifier system," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 12, pp. 2269–2279, 2003.

[12] D. K. Su and W. J. McFarland, "An IC for linearizing RF power amplifiers using envelope elimination and restoration," *IEEE Journal of Solid-State Circuits*, vol. 33, no. 12, pp. 2252–2258, 1998.

[13] A. M. Smith and J. K. Cavers, "A wideband architecture for adaptive feedforward linearization," in *Vehicular Technology Conference*, 1998, pp. 2488–2492.

[14] M. P. van der Heijden, M. Acar, J. S. Vromans, and D. A. Calvillo-cortes, "A 19W high-efficiency wide-band CMOS-GaN class-E chireix RF outphasing power amplifier," in *Microwave Symposium Digest*, 2011.

[15] A. Agah, H.-t. Dabag, B. Hana, P. M. Asbeck, J. F. Buckwalter, and L. E. Larson, "Active millimeter-wave phase-shift Doherty power amplifier in 45-nm SOI CMOS," *IEEE Journal of Solid-State Circuits*, vol. 48, no. 10, pp. 2338–2350, 2013.

[16] A. Ferrero and M. Pirola, "Harmonic load-pull techniques," *IEEE Microwave Magazine*, no. June, pp. 116–123, 2013.

[17] J. M. Cusack and S. M. Perlman, "Automaic load contour mapping for microwave power transistors," *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-22, no. 12, pp. 1146–1152, 1974.

[18] Y. Takayama, "A new load-pull characterization method for microwave power transistor," in *IEEE MTT-S International Microwave Symposium Digest*, 1976, pp. 218–220.

[19] M. Marchetti, M. J. Pelk, K. Buisman, W. C. E. Neo, M. Spirito, and L. C. N. D. Vreede, "Active harmonic load-pull with realistic wideband communications signals," *IEEE Transactions on Microwave Theory and Techniques*, vol. 56, no. 12, pp. 2979–2988, 2008.

[20] I. Angelov, H. Zirath, and N. Rorsman, "A new empirical nonlinear model for HEMT and MESFET devices," *IEEE Transactions on Microwave Theory and Techniques*, vol. 40, no. 12, pp. 2258–2266, 1992.

[21] I. Angelov, V. Desmaris, K. Dynefors, P. Å. Nilsson, N. Rorsman, and H. Zirath, "On the large-signal modelling of AlGaN / GaN HEMTs and SiC MESFETs," in *Gallium Arsenide and Other Semiconductor Application Symposium*, 2005.

[22] J. Xu, J. Hom, M. Iwamoto, and D. E. Root, "Large-signal FET model with multiple time scale dynamics from nonlinear vector network analyzer data," in *Microwave Symposium Digest*, 2010, pp. 417–420.

[23] J. Xu, R. Jones, S. A. Harris, T. Nielsen, and D. E. Root, "Dynamic FET model - DynaFET - for GaN transistors from NVNA active source injection measurements," in *Microwave Symposium*, 2014, pp. 2–4.

[24] J. Verspecht and D. E. Root, "Polyharmonic distortion modeling," *IEEE Microwave Magazine*, no. June, pp. 44–57, jun 2006.

[25] D. Root, J. Verspect, J. Horn, and M. Marcu, "X-Parameters: Characterization, modeling, and design of nonlinear RF and microwave components," *Cambridge University Press*, 2013.

[26] Keysight, "www.keysight.com," 2017.

[27] Y. Wang, T. S. Nielsen, O. K. Jensen, and T. Larsen, "X-parameter based GaN device modeling and its application to a high-efficiency PA design," pp. 1–4, 2014.

[28] D. Guerra, "GaN HEMT modeling and design for millimeter and sub-millimeter wave power amplifiers through Monte Carlo particle-based device simulations," Ph.D. dissertation, 2011.

[29] D. Guerra, M. Saraniti, D. K. Ferry, and S. M. Goodnick, "X-parameters and efficient physical device simulation of mm-wave power transistors with harmonic loading," *International Microwave Symposium*, no. 2, pp. 1–4, 2014.

[30] A. Franz, G. Franz, S. Selberherr, C. Ringhofer, and P. Markowich, "Finite boxes - A generalization of the finite-difference method suitable for semiconductor device simulation," *IEEE Transactions on Electron Devices*, vol. 30, no. 9, pp. 1070–1082, sep 1983.

[31] G. E. Forsythe and W. R. Wasow, *Finite-difference methods for partial differential equations*. New York: Wiley, 1960.

[32] L. Fox, *Numerical solution of ordinary and partial differential equations*, 1st ed., Oxford, 1962.

[33] Y. Kurihara and J. J.L. Holloway, "Numerical integration of a nine-level global primitive equations model formulated by the box method," *Mon. Wea. Rev.*, vol. 95, no. 8, pp. 509–530, 1967.

[34] A. Bowyer, "Computing Dirichlet tessellations," *The Computer Journal*, vol. 24, no. 2, pp. 162–166, feb 1981.

[35] D. F. Watson, "Computing the n-dimensional Delaunay tessellation with application to Voronoi polytopes," *The Computer Journal*, vol. 24, no. 2, pp. 167–172, feb 1981.

[36] F. Aurenhammer, "Voronoi diagrams — A survey of a fundamental geometric data structure," *ACM Computing Surveys (CSUR)*, vol. 23, no. 3, pp. 345–405, sep 1991.

[37] D. T. Lee and B. J. Schachter, "Two algorithms for constructing a Delaunay triangulation," *International Journal of Computer & Information Sciences*, vol. 9, no. 3, pp. 219–242, jun 1980.

[38] L. P. Chew, "Constrained Delaunay triangulations," in *Proceedings of the third annual symposium on Computational geometry - SCG '87*. New York, New York, USA: ACM Press, 1987, pp. 215–222.

[39] C. Geuzaine and J.-F. Remacle, "Gmsh: A 3-D finite element mesh generator with built-in pre- and post-processing facilities," *International Journal for Numerical Methods in Engineering*, vol. 79, no. 11, pp. 1309–1331, sep 2009.

[40] H. Si, "TetGen, a Delaunay-based quality tetrahedral mesh generator," *ACM Transactions on Mathematical Software*, vol. 41, no. 2, pp. 1–36, feb 2015.

[41] Q. Du, V. Faber, and M. Gunzburger, "Centroidal Voronoi tessellations: Applications and algorithms," *SIAM Review*, vol. 41, no. 4, pp. 637–676, jan 1999.

[42] Z. Q. Xie, O. Hassan, and K. Morgan, "Tailoring unstructured meshes for use with a 3D time domain co-volume algorithm for computational electromagnetics," *Int. J. Numer. Meth. Engng*, vol. 87, no. July 2010, pp. 48–65, 2011.

[43] I. Sazonov, D. Wang, O. Hassan, K. Morgan, and N. Weatherill, "A stitching method for the generation of unstructured meshes for use with co-volume solution techniques," *Computer Methods in Applied Mechanics and Engineering*, vol. 195, no. 13-16, pp. 1826–1845, feb 2006.

[44] S. Walton, O. Hassan, and K. Morgan, "Advances in co-volume mesh generation and mesh optimisation techniques," *Computers & Structures*, vol. 181, pp. 70–88, mar 2017.

[45] N. Hitschfeld, P. Conti, and W. Fichtner, "Mixed element trees: A generalization of modified octrees for the generation of meshes for the simulation of complex 3-D semiconductor device structures," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 12, no. 11, pp. 1714–1725, 1993.

[46] P. Conti, N. Hitschfeld, and W. Fichtner, "An octree-based mixed element grid allocator for the simulation of complex 3-D device structures," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 10, no. 10, pp. 1231–1241, 1991.

[47] P. Conti, M. Tomizawa, and A. Yoshii, "Generation of oriented three-dimensional Delaunay grids suitable for the control volume integration method," *International Journal for Numerical Methods in Engineering*, vol. 37, no. 19, pp. 3211–3227, 1994.

[48] Synopsys, "Mesh Generation Tools User Guide," 2013.

[49] J. Krause, "On boundary conforming anisotropic Delaunay meshes," Ph.D. dissertation, 2001.

[50] N. Hitschfeld, L. Villablanca, J. Krause, and M. C. Rivara, "Improving the quality of meshes for the simulation of semiconductor devices using Lepp-based algorithms," *International Journal for Numerical Methods in Engineering*, vol. 58, no. 2, pp. 333–347, 2003.

[51] C. T. Kelley, *Solving nonlinear equations with Newton's method*. Philadelphia: SIAM, 2003.

[52] C. G. Broyden, "A class of methods for solving nonlinear simultaneous equations," *Mathematics of Computation*, vol. 19, no. 92, p. 577, oct 1965.

[53] D. Knoll and D. Keyes, "Jacobian-free Newton-Krylov methods: A survey of approaches and applications," *Journal of Computational Physics*, vol. 193, no. 2, pp. 357–397, 2004.

[54] H. K. Khalil, *Nonlinear systems*, 3rd ed. Upper Saddle River: Prentice Hall, 2002.

[55] D. Pozar, *Microwave engineering*, 4th ed. Wiley, 2005.

[56] G. Gonzalez, *Microwave transistor amplifiers: Analysis and design*, 2nd ed. Upper Saddle River: Prentice Hall, 1997.

[57] S. C. Cripps, *RF power amplifiers for wireless communications*, 2nd ed. Norwood: Artech House, 2006.

[58] D. E. Root, J. Verspecht, D. Sharrit, J. Wood, and A. Cognata, "Broad-band poly-harmonic distortion (PHD) behavioral models from fast automated simulations and large-signal vectorial network measurements," *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, no. 11, pp. 3656–3664, 2005.

[59] J. Horn, D. E. Root, and G. Simpson, "GaN device modeling with X-parameters," 2010.

[60] A. M. Peláez-pérez, S. Woodington, M. Fernández-barciela, P. J. Tasker, and J. I. Alonso, "Application of an NVNA-based system and load-independent X-parameters in analytical circuit design assisted by an experimental search algorithm," *IEEE Trans. Microwave Theory and Tech.*, vol. 61, no. 1, pp. 581–586, 2013.

[61] Y. Wang, T. S. Nielsen, O. K. Jensen, and T. Larsen, "GaN HEMT device modeling using X-parameters with emphasis on complexity reduction of harmonic load-pull measurement," 2015.

[62] M. Grupen, "Full wave electromagnetics and hot electron transport with electronic band structure for high speed semiconductor device simulation," *IEEE Trans. Microwave Theory and Tech.*, vol. 62, no. 12, pp. 2868 – 2877, 2014.

[63] J.-M. Jin, *Theory and computation of electromagnetic fields*. Wiley, 2010.

[64] J. Hesthaven and T. Warburton, "Nodal high-order methods on unstructured grids," *J. Computat. Phys.*, vol. 181, no. 1, pp. 186–221, sep 2002.

[65] R. J. LeVeque, *Finite volume methods for hyperbolic problems*. Cambridge University Press, 2002.

[66] K. J. Willis, S. C. Hagness, and I. Knezevic, "Multiphysics simulation of high-frequency carrier dynamics in conductive materials," *Journal of Applied Physics*, vol. 110, pp. 1–15, 2011.

[67] T. Weiland, "Time domain electromagnetic field computation with finite difference methods," *International Journal of Numerical Modeling: Electronic Networks, Devices, and Fields*, vol. 9, pp. 295–319, 1996.

[68] K. S. Yee, "Numerical solution of initial value problems of Maxwells equations," *IEEE Transactions on Antennas and Propagation*, vol. 14, pp. 302–307, 1966.

[69] N. K. Madsen, "Divergence preserving discrete surface integral methods for Maxwell's curl equations using non-orthogonal unstructured grids," *Journal of Computational Physics*, vol. 119, pp. 34–45, 1995.

[70] A. Taflove and S. C. Hagness, *Computational electrodynamics - The finite-difference time-domain method*, 3rd ed. Artech House, 2005.

[71] R. F. Harrington, *Time-harmonic electromagnetic fields*. Wiley-IEEE Press, 2001.

[72] K. Hess, *Advanced theory of semiconductor devices*, 2000, vol. 1.

[73] C. Jacoboni, *Theory of electron transport in semiconductors: A pathway from elementary physics to nonequilibrium green functions*, 2010.

[74] N. Metropolis and S. Ulam, "The Monte Carlo method," *Journal of the American Statistical Association*, vol. 44, no. 247, pp. 335–341, 1949.

[75] M. Fischetti and S. Laux, "Monte Carlo analysis of electron transport in small semiconductor devices including band-structure and space-charge effects," *Physical Review B*, vol. 38, no. 14, pp. 9721–9745, 1988.

[76] M. Saraniti and S. M. Goodnick, "Hybrid fullband cellular automaton/Monte Carlo approach for fast simulation of charge transport in semiconductors," *IEEE Transactions on Electron Devices*, vol. 47, no. 10, pp. 1909–1916, 2000.

[77] D. L. Scharfetter and H. K. Gummel, "Large-signal analysis of a silicon read diode oscillator," *IEEE Transactions on Electron Devices*, vol. ED-16, no. 1, pp. 64–77, 1969.

[78] R. Stratton, "Diffusion of hot and cold electrons in semiconductor barriers," *Physical Review*, vol. 126, no. 6, pp. 2002–2014, 1962.

[79] K. Blotekjaer, "Transport equations for electrons in two-valley semiconductors," *IEEE Transactions on Electron Devices*, vol. ED-17, no. 1, pp. 38–47, 1970.

[80] T. Grasser, T. W. Tang, H. Kosina, and S. Selberherr, "A review of hydrodynamic and energy-transport models for semiconductor device simulation," *Proceedings of the IEEE*, vol. 91, no. 2, pp. 251–274, feb 2003.

[81] Synopsys, "Sentaurus device user guide," 2013.

[82] M. Grupen, "An alternative treatment of heat flow for charge transport in semiconductor devices," *Journal of Applied Physics*, vol. 106, no. 12, p. 123702, 2009.

[83] ——, "Energy transport model with full band structure for GaAs electronic devices," *Journal of Computational Electronics*, vol. 10, no. 3, pp. 271–290, may 2011.

[84] ——, "GaN high electron mobility transistor simulations with full wave and hot electron effects," *IEEE Transactions on Electron Devices*, vol. 63, no. 8, pp. 3096–3102, 2016.

[85] T. K. Bergstresser and M. L. Cohen, "Electronic structure and optical properties of hexagonal CdSe, CdS, and ZnS," *Physical Review*, vol. 164, pp. 1069–1080, 1967.

[86] M. Goano, "Series expansion of the Fermi-Dirac integral over the entire domain of real j and x," *Solid-State Electronics*, vol. 36, no. 2, pp. 217–221, 1993.

[87] S. Selberherr, *Analysis and simulation of semiconductor devices*. Springer-Verlag Wien, 1984.

[88] K. Seeger, *Semiconductor physics: An introduction*, 9th ed. Springer-Verlag, 2004.

[89] S. Sridharan, A. Christensen, A. Venkatachalam, S. Graham, and P. D. Yoder, "Temperature- and doping-dependent anisotropic stationary electron velocity in wurtzite GaN," *IEEE Electron Device Letters*, vol. 32, no. 11, pp. 1522–1524, nov 2011.

[90] J. George and J.-H. Liu, *Computer solution of large sparse positive definite systems*. Prentice Hall, 1981.

[91] S. Godunov, "Different methods for shock waves," Ph.D. dissertation, Moscow State University, 1954.

[92] A. Harten, "High resolution schemes for hyperbolic conservation laws," *Journal of Computational Physics*, vol. 49, no. 3, pp. 357–393, mar 1983.

[93] M. S. Darwish and F. Moukalled, "TVD schemes for unstructured grids," *International Journal of Heat and Mass Transfer*, vol. 46, no. 4, pp. 599–611, 2003.

[94] B. Benbakhti, A. Soltani, K. Kalna, M. Rousseau, and J. C. De Jaeger, "Effects of self-heating on performance degradation in AlGaN/GaN-based devices," *IEEE Transactions on Electron Devices*, vol. 56, no. 10, pp. 2178–2185, 2009.

[95] S. Forster, B. Ubochi, and K. Kalna, "Monte Carlo simulations of electron transport in bulk GaN," in *2017 International Workshop on Computational Nanotechnology*, 2017.

[96] A. Brannick, N. A. Zakhleniuk, B. K. Ridley, L. F. Eastman, J. R. Shealy, and W. J. Schaff, "Hydrodynamic simulation of surface traps in the AlGaN/GaN HEMT," *Microelectronics Journal*, vol. 40, no. 3, pp. 410–412, 2009.

[97] X. D. Wang, W. D. Hu, X. S. Chen, and W. Lu, "The study of self-heating and hot-electron effects for AlGaN/GaN double-channel HEMTs," *IEEE Transactions on Electron Devices*, vol. 59, no. 5, pp. 1393–1401, 2012.

[98] S. Vitanov, V. Palankovski, S. Maroldt, R. Quay, S. Murad, T. Rödle, and S. Selberherr, "Physics-based modeling of GaN HEMTs," *IEEE Transactions on Electron Devices*, vol. 59, no. 3, pp. 685–693, 2012.

[99] S. Bhardwaj, B. Sensale-Rodriguez, H. G. Xing, and J. L. Volakis, "Full-wave hydrodynamic model for predicting THz emission from grating-gate RTD-gated plasma wave HEMTs," in *2015 73rd Annual Device Research Conference (DRC)*, 2015.

[100] S. Yamakawa, S. Goodnick, S. Aboud, and M. Saraniti, "Quantum corrected full-band cellular Monte Carlo simulation of AlGaN/GaN HEMTs," *Journal of Computational Electronics*, vol. 3, no. 3-4, pp. 299–303, 2004.

[101] T. Sadi, R. W. Kelsall, and N. J. Pilgrim, "Investigation of self-heating effects in submicrometer GaN/AlGaN HEMTs using an electrothermal Monte Carlo method," *IEEE Transactions on Electron Devices*, vol. 53, no. 12, pp. 2892–2900, 2006.

[102] S. Russo and A. D. Carlo, "Influence of the source-gate distance on the AlGaN/GaN HEMT performance," *IEEE Transactions on Electron Devices*, vol. 54, no. 5, pp. 1071–1075, 2007.

[103] S. Sridharan, A. Venkatachalam, and P. D. Yoder, "Electrothermal analysis of AlGaN/GaN high electron mobility transistors," *Journal of Computational Electronics*, vol. 7, no. 3, pp. 236–239, 2008.

[104] A. Ashok, D. Vasileska, O. L. Hartin, and S. M. Goodnick, "Electrothermal Monte Carlo simulation of GaN HEMTs including electron-electron interactions," *IEEE Transactions on Electron Devices*, vol. 57, no. 3, pp. 562–570, 2010.

[105] S. N. Tenneti, N. K. Nahar, and J. L. Volakis, "Full-wave modeling of THz RTD-gated GaN HEMTs," *Infrared Physics and Technology*, vol. 72, pp. 221–228, 2015.

[106] T. Palacios, A. Chakraborty, S. Heikman, S. Keller, S. P. DenBaars, and U. K. Mishra, "AlGaN / GaN high electron mobility transistors with InGaN back-barriers," *IEEE Elec*, vol. 27, no. 1, pp. 13–15, 2006.

[107] F. A. Marino, N. Faralli, T. Palacios, D. K. Ferry, S. M. Goodnick, and M. Saraniti, "Effects of threading dislocations on AlGaN / GaN high-electron mobility transistors," *IEEE Trans. Electron Devices*, vol. 57, no. 1, pp. 353–360, 2010.

[108] S. M. Goodnick and M. Saraniti, "Modeling and simulation of terahertz devices," *IEEE Microwave Magazine*, vol. 13, no. 7, pp. 36–44, 2012.

[109] N. C. Miller, J. D. Albrecht, and M. Grupen, "Large-signal RF GaN HEMT simulation using Fermi kinetics transport," in *2016 74th Annual Device Research Conference (DRC)*. IEEE, jun 2016, pp. 1–2.

[110] D. K. Schroder, *Semiconductor material and device characterization*, 3rd ed. Hoboken, NJ, USA: Wiley-Interscience, 2005.

[111] M. Wang and K. J. Chen, "Kink effect in AlGaN/GaN HEMTs induced by drain and gate pumping," *IEEE Electron Device Letters*, vol. 32, no. 4, pp. 482–484, 2011.

[112] G. Simpson, J. Horn, D. Gunyan, and D. E. Root, "Load-Pull + NVNA = Enhanced X-parameters for PA designs with high mismatch and technology-independent large-signal device models," in *72nd Microwave Measurement Symposium*, 2008, pp. 88–91.

[113] N. Ashcroft and N. Mermin, "Solid state physics," 1976.